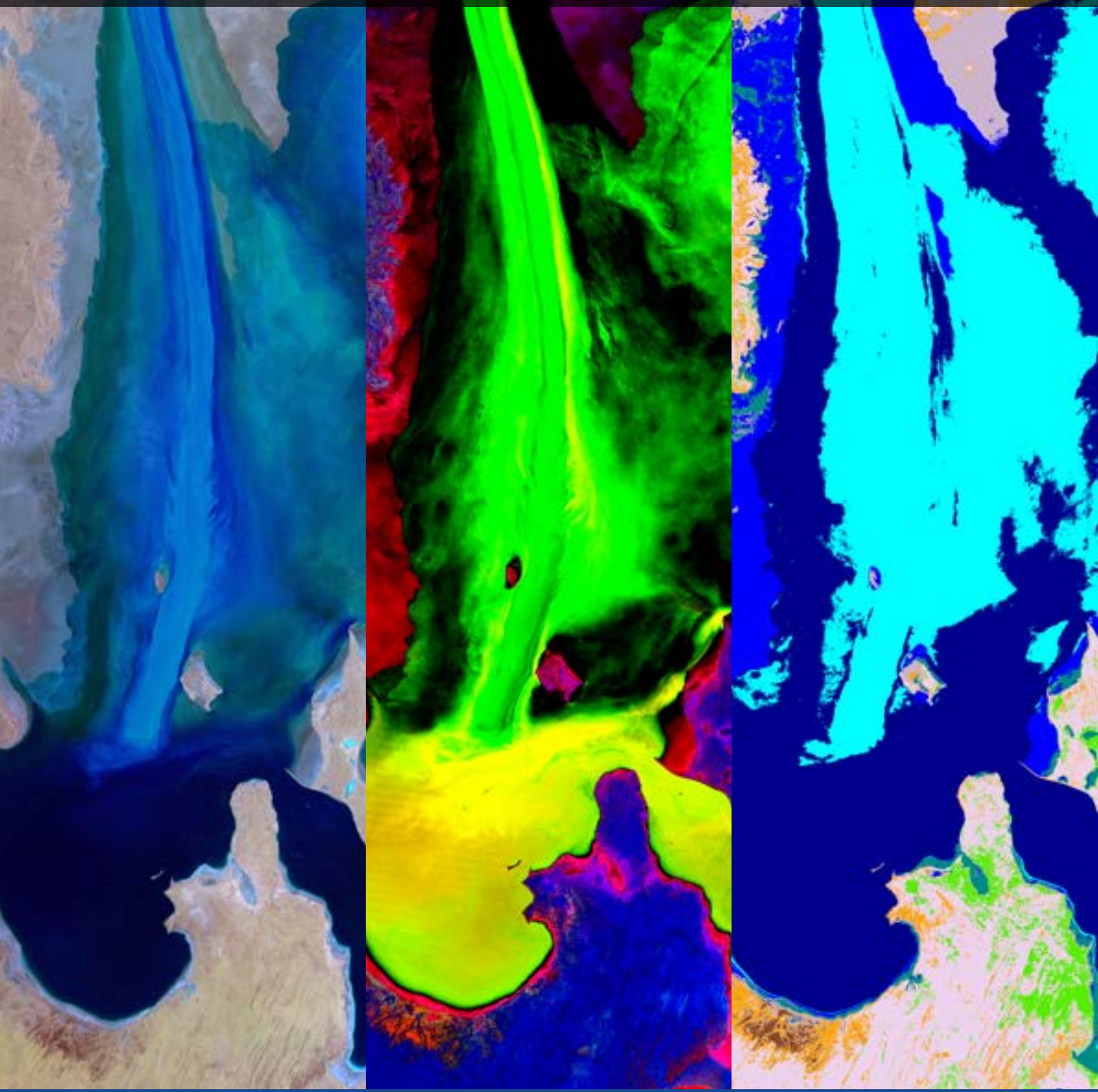
Earth Observation:
Data, Processing and Applications

Volume 2: Processing



This publication is available in PDF format at *eopublications.au*.
We welcome your comments regarding the readability and usefulness of this report.
To provide feedback, please email *feedback@eopublications.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
Volume 2A 978-0-9943019-5-6
Volume 2B 978-0-6482278-0-9
Volume 2C 978-0-6482278-1-6
Volume 2D 978-0-6482278-2-3
Volume 2X 978-0-6482278-4-7

Copyright:
All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, save for content supplied by third parties, and logos. Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. The full licence terms are available from *https://creativecommons.org/licenses/by/4.0/legalcode*. A summary of the licence terms is available from *https://creativecommons.org/licenses/by/4.0/*

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Recommended Citation for Volume 2:
CRCSI (2020) *Earth Observation: Data, Processing and Applications. Volume 2: Processing* (Eds. Harrison, B.A., Jupp, D.L.B., Lewis, M.M., Sparks, T., Mueller, N., Byrne, G., Phinn, S.F., Hudson, D.A.) CRCSI, Melbourne.

Latest Revision:
October 2025

Background image on previous page: Landsat-5 image of Lake Eyre, South Australia, acquired on 9 May, 2009, while the lake was filling. Three images are juxtaposed: a colour composite of bands 5, 4, 2 as RGB on the left, a principal component transformation (with PC3, PC4 and PC1 displayed as RGB) in the centre and a 10 class ISODATA classification on the right.

Source: Norman Mueller, Geoscience Australia

About this Series

This text, *Earth Observation: Data, Processing and Applications*, is an Australian Earth Observation (EO) community undertaking, which has been jointly sponsored by the Cooperative Research Centre for Spatial Information (CRCSI), the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Geoscience Australia (GA) and the Bushfire and Natural Hazards Cooperative Research Centre (BNHCRC). This publication describes EO data, processing and applications in an Australian context, and includes a wide range of local case studies. It aims to:

- Consolidate EO teaching resources in Australia;
- Encourage a greater number of tertiary courses involving EO in Australia; and
- Promulgate Australia's growing dependence on EO data.

This new text substantially revises and updates two previous books published by CSIRO, namely *Introduction to Remotely Sensed Data* and *Introduction to Image Processing*, which were written to support users of the microBRIAN Image Processing system and used as tertiary texts throughout Australasia for over a decade. The new series comprises multiple volumes, with the option to expand into the future. Unlike the previous publications, the new text is envisaged as an ongoing, community resource, involving a broader range of professionals from academia, government and industry, with relevance to a wider audience. Each volume has been prepared by an editorial panel of local experts, supported by contributing authors who have written and/or reviewed application chapters and case studies. Planning and production of the series has been overseen by advisors from the sponsor organisations.

This series currently comprises three volumes:

Volume 1: Data

This volume outlines the basics of EO in terms of energy sources, data acquisition, sampling characteristics, and image availability. Discussion focuses on the scientific foundations of EO. This volume also considers the interactions of electromagnetic radiation with various targets and media, as relevant to understanding optical, thermal, passive microwave and radar imagery. Volume 1 comprises three sub-volumes:

- 1A—Basics and Acquisition
- 1B—Image Interpretation
- 1X—Appendices

Volume 2: Processing

Volume 2 describes the various options involved with image representation, analysis, transformation, integration and modelling, including details of relevant algorithms, with emphasis on their underlying mathematical and statistical principles. It comprises six sub-volumes:

- 2A—Basic Image Operations
- 2B—Image Rectification
- 2C—Image Transformations
- 2D—Image Integration
- 2E—Image Classification
- 2X—Appendices

Volume 3: Applications

This final volume introduces the Australian environment in terms of topography, climate, ecoregions, land use and vegetation dynamics, then covers a broad range of application areas reliant on EO data. Specific case studies are included to demonstrate individual applications as appropriate. Volume 3 currently comprises two sub-volumes:

- 3A—Terrestrial Vegetation
- 3B—Surface Waters

This structure has been designed to easily enable expansion of volumes in the future.

The resulting publication aims to be both comprehensive and current. It will also enable consistent definitions to be established within the Australian EO sector. Ongoing inclusion of methods, applications and case studies, which utilise a wide range of data sources, will enhance its value to our whole community.

About this Volume

Image processing allows digital image data to be displayed, enhanced, analysed or output in colour or grey-scale format. Most Earth Observation (EO) image data is generated directly by sensing the object of interest, as occurs with remote sensing scanners or digital cameras carried on satellite and aircraft. Various aspects of EO data acquistion are detailed in Volume 1.

Digital imagery is manipulated using image processing operations to enhance the information it contains. Processing may involve enhancing or detecting features of interest, changing the image geometry, extracting patterns, removing noise, or reducing data volume. Image processing uses mathematical operations with digital image data to achieve these results flexibly and quickly. Images are typically large datasets, so while the actual processing operations are often quite simple, specialised hardware and software are generally required to perform the functions efficiently.

Volume 2 in this series describes the various options involved with image representation, analysis, transformation, integration and modelling, including details of relevant algorithms, with emphasis on their underlying mathematical and statistical principles. It comprises six sub-volumes:

Volume 2A — Basic Image Operations

Volume 2A introduces image processing in the context of EO imagery, and is presented in three parts:

- Introduction: describes digital image formats, processing stages and calibration requirements;
- Interpretation: introduces basic processes for improving image contrast, colour, presentation and geometry; and
- Analysis: summarises methods for deriving statistics, classifications and segmentations of EO imagery.

Volume 2B — Image Rectification and Registration

Volume 2B describes concepts and processes for correcting EO image geometry, including relevant cartographical background, factors impacting the geometry of EO imagery, models for image rectification, and methods for resampling imagery.

Volume 2C — Image Transformations

Volume 2C details a range of image transformations commonly used with EO imagery, in terms of:

- Mathematical Foundations—filtering operations, linear operations and non-linear operations;
- Single Channel Rescaling—linear rescaling, non-linear rescaling and destriping;
- Single Channel Filtering—smoothing, interpolation, edge enhancement, surface highlights, and texture; and
- Multiple Channel Operations—Principal Components Analysis, channel ratios and vegetation indices.

Volume 2D — Image Integration

Volume 2D considers methods that integrate multiple EO images with other spatial datasets, including multiple image operations, change detection methods and time series analysis.

Volume 2E — Image Classification and Analysis

Volume 2E provides extensive background information for classifying EO images within the structure:

- Background—classification approaches, relevant statistical methods, preliminary considerations;
- Class Definition—enumerating classes, generating class statistics, and measuring separation between classes;
- Allocation—clustering approaches, classification methods and knowledge-based classifiers; and
- Interpretation—labelling and aggregating classes, and assessing classifications.

Volume 2X — Appendices

Volume 2X supplements the information in Volume 2 with detailed discussion of brightness correction models, BRDF models, colour systems and sequences, filters, and matrix algebra.

Preface

It has been suggested that business schools do three things for students[1]:

- provide them with a vocabulary that enables them to talk with authority about subjects they do not understand;
- transmit to them a set of principles that have demonstrated an ability to withstand any amount of disconfirming evidence; and
- provide a ticket of admission to a job where they can learn something about these subjects.

Some would say that these characteristics apply to current approaches to education in many other disciplines, including Earth Observation.

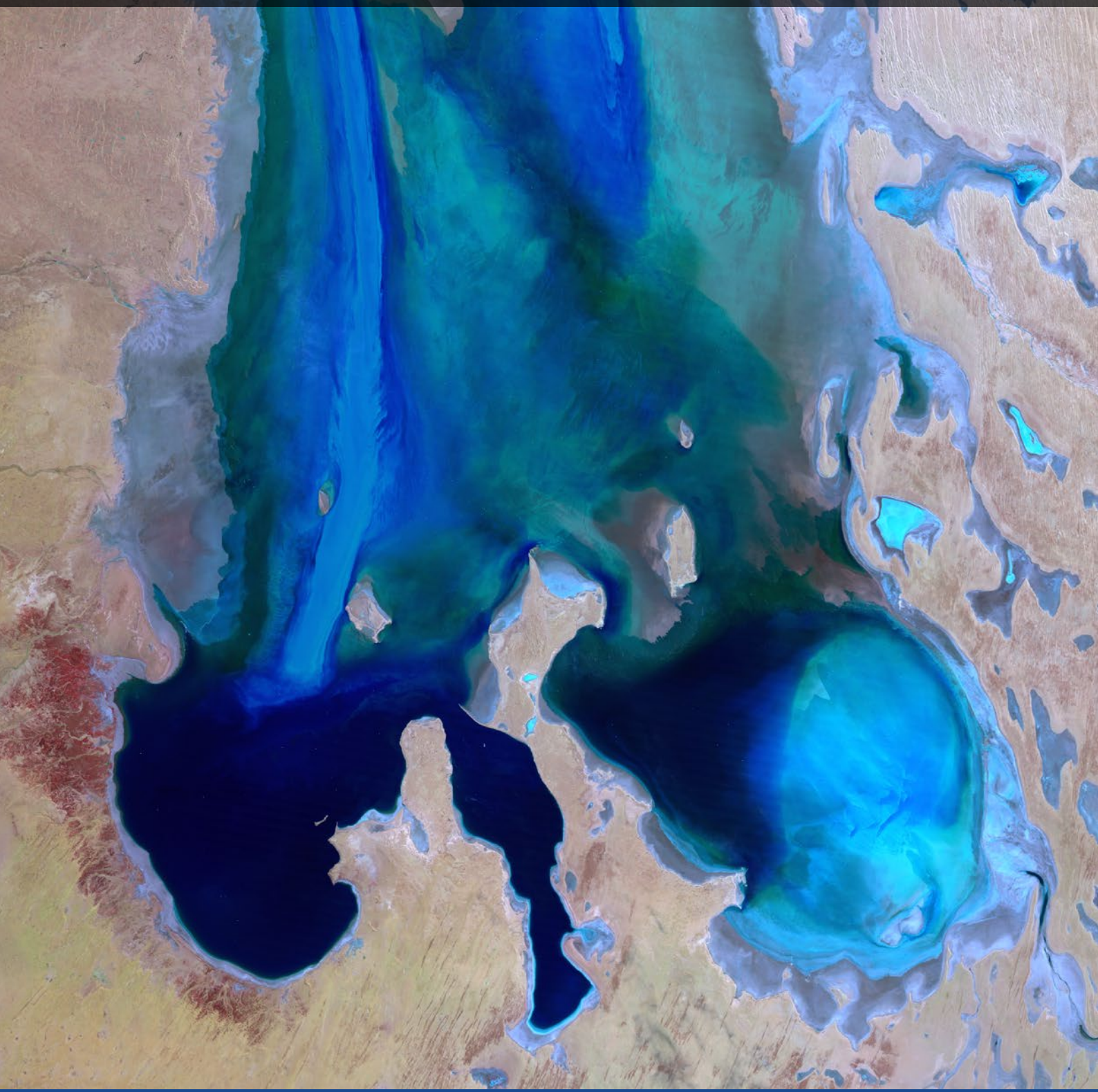
It is the aspiration of the authors of this text that students will learn more than the vocabulary of Earth Observation from these pages; hopefully they will acquire the knowledge to develop robust analytical and synthetic skills that will enable this field to advance and expand with insight and integrity.

A range of additional resources to assist readers is compiled in various Appendices. In particular, numerous online tutorials are now available from a range of sources to complement the material presented in this series.

1 Comments by Russell L. Ackoff in an interview with Robert J. Allio as reported in: Allio, R.J., and Ackoff, R.L. (2003). Iconoclastic management authority, advocates a 'systematic' approach to innovation. Strategy and Leadership 31 (3), 19–26.

Earth Observation:
Data, Processing and Applications

Volume 2A: Processing—Basic Image Operations



This publication is available in PDF format at eopublications.au.
We welcome your comments regarding the readability and usefulness of this report.
To provide feedback, please email *feedback@eopublications.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
978-0-9943019-5-6

Copyright:
All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, save for content supplied by third parties, and logos. Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. The full licence terms are available from *https://creativecommons.org/licenses/by/4.0/legalcode*. A summary of the licence terms is available from *https://creativecommons.org/licenses/by/4.0/*.

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Recommended Citation for Volume 2A:
CRCSI (2017) *Earth Observation: Data, Processing and Applications. Volume 2A: Processing—Basic Image Operations*. (Eds. Harrison, B.A., Jupp, D.L.B., Lewis, M.M, Sparks, T., Phinn, S.F., Mueller, N., Byrne, G.) CRCSI, Melbourne.

Latest Revision:
October 2025

Background image on previous page: Landsat-5 image of Lake Eyre, South Australia, acquired on 9 May, 2009, while the lake was filling. Three images are juxtaposed: a colour composite of bands 5, 4, 2 as RGB on the left, a principal component transformation (with PC3, PC4 and PC1 displayed as RGB) in the centre and a 10 class ISODATA classification on the right.
Source: Norman Mueller, Geoscience Australia

Acknowledgements

Production of this series of texts would not have been possible without the financial support of CSIRO, CRCSI, GA and BNHCRC, input from members of the editorial panels and direction from members of the various advisory panels.

Volumes 1 and 2 of this series are based on text originally published in Harrison and Jupp (1989, 1990, 1992 and 1993)[1]. Many illustrations and some text from these publications have been reproduced with permission from CSIRO. Much of the text in Section 3 has been derived from Jupp and Strahler (1996)[2].

Other contributors are gratefully acknowledged:

- reviewers: Alfredo Huete and Laurie Chisholm (various sections), David Hudson and Joshua Sixsmith (parts of Sections 1 and 2), and Tim Malthus (Section 3);
- illustrations: Norman Mueller kindly supplied most of the lovely images; other contributors of graphical material include: Tony Sparks, Megan Lewis, Fuqin Li, David Jupp, Stephen Sagar, Medhavy Thankappan, Edward King, Karen Joyce, Guy Byrne, Chris MacLellan, Tim Malthus, Neil Flood, Thomas Cudahy, Michael Caccetta, Ken Clarke and Dale Roberts;
- Jason Mazur (MIrage Digital) for permission to reproduce Figure 5.15 in this document;
- Paul Cockram and Matej Hrcek, who 'translated' the Harrison and Jupp (1989, 1990) files from Quark Express 2.0;
- Joost Kuckartz for formatting of references and equations;
- Daniel Rawson (Accessible Publication & Template Design) for layout and formatting; and
- Carl Davies (CMDphotographics) for selected graphical illustrations.

We thank those owners of copyrighted illustrative material for permission to reproduce their work. Credits for individual illustrations are provided below the relevant graphic.

All volumes in this series are covered by the copyright provisions of CC BY 4.0 AU.

1 Harrison, B.A., and Jupp, D.L.B. (1989) *Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual.* CSIRO, Melbourne. 156pp.
Harrison, B.A., and Jupp, D.L.B. (1990) *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.
Harrison, B.A., and Jupp, D.L.B. (1992) *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual.* MPA, Melbourne.
Harrison, B.A., and Jupp, D.L.B. (1993) *Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual.* MPA, Melbourne.

2 Jupp, D.L.B., and Strahler, A. (1996) *Image Brightness and BRDF Workshop Issues.* CSIRO Earth Observation Centre, Canberra.

Table of Contents

Acknowledgements i

Introduction 1

1 Digital Images **3**
- **1.1** Image Structure 3
- **1.2** Image Data Formats 4
  - **1.2.1** Band sequential (BSQ) 5
  - **1.2.2** Band interleaved by pixel (BIP) 5
  - **1.2.3** Band interleaved by line (BIL) 5
- **1.3** Data Volume 6
- **1.4** Metadata 6
- **1.5** Pixel Quality Flag 9
- **1.6** Hierarchical Storage Structures 10
- **1.7** Further Information 11
- **1.8** References 11

2 Processing Stages **13**
- **2.1** Data Provenance 13
  - **2.1.1** Data products 14
  - **2.1.2** Information products 14
- **2.2** Processing Methodologies 17
- **2.3** Modelling 17
  - **2.3.1** Modelling Overview 17
  - **2.3.2** Forward and inverse models 18
  - **2.3.3** Empirical or analytical models 18
  - **2.3.4** Semi-empirical or semi-analytical models 18
  - **2.3.5** Physics-based models 18
- **2.4** Further Information 19
- **2.5** References 19

3 Calibration **21**
- **3.1** About Calibration 21
- **3.2** Optical Image Distortions 23
  - **3.2.1** Spectral 23
  - **3.2.2** Spatial 25
  - **3.2.3** Radiometric 30
  - **3.2.4** Temporal 33
- **3.3** Optical Image Corrections 36
  - **3.3.1** Geometric 36
  - **3.3.2** Radiometric 37
- **3.4** Analysis Ready Data 43
- **3.5** Further Information 43
- **3.6** References 43

Interpretation 47

4 Contrast **49**
4.1 Channel statistics 49
4.1.1 Histograms 49
4.1.2 Variance 51
4.2 Contrast enhancement 53
4.2.1 Linear stretching 53
4.2.2 Non-linear stretching 60
4.2.3 Statistical stretching 64
4.2.4 Lookup tables 65
4.3 Further Information 66
4.4 References 66

5 Colour **67**
5.1 Colour basics 67
5.1.1 Colour description 67
5.1.2 Models for colour ordering 70
5.1.3 Colour reproduction 74
5.2 Colour enhancement 74
5.2.1 Pseudo-colouring 74
5.2.2 Colour composite imagery 75
5.3 Further Information 87
5.4 References 87

6 Presentation **89**
6.1 Image Display 89
6.1.1 Display Hardware 89
6.1.2 Grey-scale Imagery 90
6.1.3 Colour Composites 91
6.2 Hardcopy 91
6.2.1 Printer Output 91
6.2.2 Photographic Output 93
6.3 Further Information 93
6.4 References 93

7 Geometry **95**
7.1 Image Coordinate System 95
7.2 Shape and Orientation 96
7.2.1 Modifying image size 96
7.2.2 Changing image orientation 99
7.3 Rectification and Resampling 100
7.4 Image grids 103
7.4.1 Sensor-specific custom grids 103
7.4.2 Sensor-agnostic nested grids 103
7.5 Further Information 104
7.6 References 104

Analysis 105

8 Statistics **107**

8.1 Spectral 107

8.1.1 Histograms 107

8.1.2 Variance 109

8.1.3 Crossplots 109

8.1.4 Covariance 112

8.2 Spatial 114

8.2.1 Local variance 115

8.2.2 Variograms 116

8.2.3 Autocorrelation and variograms 116

8.3 Radiometric 120

8.3.1 EMR interactions with target and sensor 120

8.3.2 EMR interactions with the atmosphere 121

8.3.3 Sensor operation 121

8.4 Temporal 124

8.5 Further Information 126

8.6 References 126

9 Classification **129**

9.1 Characterising Classes 134

9.1.1 Defining image classes 135

9.1.2 Comparing classes and classifications 141

9.1.3 Selecting image channels 147

9.2 Allocating Pixels 147

9.2.1 Density slicing 148

9.2.2 Parallelepiped classification 148

9.2.3 Minimum distance classifier 149

9.2.4 Maximum Likelihood classifier 151

9.3 Labelling Classes 153

9.3.1 Clustering tools 153

9.3.2 Aggregating Classes 154

9.3.3 Post-classification processing 157

9.4 Verifying Results 157

9.5 Further Information 161

9.6 References 161

10 Segmentation **163**

10.1 Image Spatial Patterns 163

10.2 Spectral Themes 164

10.3 Ancillary Data Boundaries 165

10.4 Recombining Segments 165

10.5 Cleaning Boundaries 166

10.6 Further Information 171

10.7 References 171

List of Figures

Figure 1.1 Components of a digital image 3

Figure 1.2 Image data formats 4

Figure 1.3 Earth Observation equipment metadata 7

Figure 1.4 Earth Observation result metadata 8

Figure 2.1 The concept of tidal attribution and Lowest (LOT) and Highest (HOT) Observed Tide. 16

Figure 2.2 Roebuck Bay Example 16

Figure 3.1 Precision versus trueness 21

Figure 3.2 Geometry of illumination and viewing positions 23

Figure 3.3 AVHRR pixel geometry and correction of spatial distortions 25

Figure 3.4 Example of Point Spread Function 28

Figure 3.5 Effect of Landsat-7 ETM+ SLC malfunction 29

Figure 3.6 Systematic radiometric miscalibration 31

Figure 3.7 Estimated annual % gain change per year for Landsat-7 ETM+ calibrators 32

Figure 3.8 Validation sites for vicarious sensor calibration 33

Figure 3.9 Seasonal solar illumination variations 34

Figure 3.10 Temporal variations in EO imagery 35

Figure 3.11 Geometric and radiometric corrections 36

Figure 3.12 Geometric correction of CASI imagery 37

Figure 3.13 Surface reflectance correction 38

Figure 3.14 ASTER image mosaic 39

Figure 3.15 Configuration of trial corner reflectors 40

Figure 3.16 Corner reflectors 40

Figure 3.17 CR response in SAR image 41

Figure 3.18 CR network 41

Figure 3.19 Target visibility measure 42

Figure 4.1 Grey-scale image display 49

Figure 4.2 Frequency histogram 50

Figure 4.3 Cumulative histogram 50

Figure 4.4 A bi-modal histogram 50

Figure 4.5 Histogram stretching for image display 51

Figure 4.6 Histograms with equal ranges and differing variances 51

Figure 4.7 Example image 52

Figure 4.8 Contrast enhancement 54
Figure 4.9 Linear stretching 56
Figure 4.10 Bi-modal histogram stretching 58
Figure 4.11 Logarithmic and exponential functions 60
Figure 4.12 Enhancing high image values 62
Figure 4.13 Logarithmic Stretching Example 63
Figure 4.14 Normalised stretch 64
Figure 4.15 Histogram equalisation 64
Figure 4.16 LUT operations for rescaling image data values 65
Figure 4.17 User-defined stretch 65
Figure 5.1 Relative light absorption of human cones 68
Figure 5.2 Opponent process theory of colour vision 68
Figure 5.3 Perception of intensity changes 69
Figure 5.4 Effect of adjacent intensities on perceived brightness 69
Figure 5.5 Perception of chromaticity changes 69
Figure 5.6 Effect of colour intensity change on perceived hue 70
Figure 5.7 Additive and subtractive colour mixing 70
Figure 5.8 Relationship between additive and subtractive primary colours 71
Figure 5.9 RGB colour cube 71
Figure 5.10 CIE imaginary primaries 72
Figure 5.11 CIE chromaticity diagram 73
Figure 5.12 Pseudo-colouring of image data 74
Figure 5.13 Formation of a colour composite image 76
Figure 5.14 Shark Bay colour composite example 79
Figure 5.15 Aerial photograph of Shark Bay 80
Figure 5.16 Example image channels for colour composites 82
Figure 5.17 Standard colour composite images 83
Figure 5.18 Thermal colour composite using Sentinel-2 imagery 84
Figure 5.19 Non-standard colour composite images 85
Figure 5.20 Spectral bands in Landsat sensors 86
Figure 6.1 Additive colour mixing 90
Figure 6.2 Generating a colour LUT from RGB intensities 90
Figure 6.3 Mirroring the LUT 91
Figure 6.4 Dithering 92
Figure 7.1 Standard image coordinate system 96

Figure 7.2 Subsetting an image 96
Figure 7.3 Joining multiple images 97
Figure 7.4 Blocking an image 97
Figure 7.5 Sub-sampling an image 98
Figure 7.6 Enlarging an image 98
Figure 7.7 Reversing pixels and/or lines in an image 99
Figure 7.8 Defining image origin 99
Figure 7.9 Transposing an image 99
Figure 7.10 Rotating an image 100
Figure 7.11 Earth Observation scene geometry 101
Figure 7.12 Three-stage rectification process 102
Figure 7.13 Overlaying map data 103
Figure 7.14 GA data cube for ARG25 product 104
Figure 8.1 Effect of rescaling on image histogram 108
Figure 8.2 Example image channels 110
Figure 8.3 Example crossplots 111
Figure 8.4 Format of variance/covariance matrix 113
Figure 8.5 Spectrally equivalent images with differing spatial patterns 114
Figure 8.6 Impact of pixel size on feature discernment 115
Figure 8.7 One-dimensional variogram 117
Figure 8.8 Two-dimensional variogram 118
Figure 8.9 Artificial example images 119
Figure 8.10 Local standard deviation results derived using average function 120
Figure 8.11 Variogram results for example images 120
Figure 8.12 Continental mosaic of ASTER imagery 123
Figure 8.13 Trends in chlorophyll-a concentration 124
Figure 8.14 Time series example 125
Figure 9.1 Supervised classification 131
Figure 9.2 Unsupervised classification 131
Figure 9.3 Hybrid or mixed mode classification 132
Figure 9.4 Study site 133
Figure 9.5 Vegetation Classification 134
Figure 9.6 Theme and seed values 135
Figure 9.7 Locating a training area 136
Figure 9.8 Five patch types 136

Figure 9.9 Generating themes using image statistics 138
Figure 9.10 Comparison of class seeds 139
Figure 9.11 Canonical Variates Analysis (CVA) 141
Figure 9.12 Assessing iterative classifications 142
Figure 9.13 Dendrogram example 143
Figure 9.14 Spectral MST 145
Figure 9.15 Density slicing 148
Figure 9.16 Parallelepiped classification 148
Figure 9.17 Distance thresholds for minimum distance classifiers 149
Figure 9.18 Channel gates 150
Figure 9.19 Migrating means process 150
Figure 9.20 Maximum Likelihood classification 152
Figure 9.21 Allocation using Maximum Likelihood classifier 153
Figure 9.22 Highlighting class deficiencies 154
Figure 9.23 MST colour sequences. 156
Figure 9.24 Heron Island Reef example 159
Figure 9.25 Image classification sequence 159
Figure 9.26 Classified image 160
Figure 10.1 Complementing an image mask 164
Figure 10.2 Including and excluding regions in spectral segmentation 164
Figure 10.3 Image segmentation with ancillary data boundaries 165
Figure 10.4 Conceptual model of domains relevant to EO image interpretation 167
Figure 10.5 Image sampling resolution relative to imaged objects 167
Figure 10.6 Typical GEOBIA classification workflow 168
Figure 10.7 Example image 169
Figure 10.8 Primary segmentation 169
Figure 10.9 Classification of land and water 170
Figure 10.10 Secondary segmentation 170

List of Tables

Table 1.1 Example of data compression rates from run length encoding 6

Table 1.2 Pixel quality flag 9

Table 1.3 Examples of pixel quality flag values 10

Table 2.1 Image processing stages 17

Table 4.1 Calculation of variance statistics 52

Table 5.1 Equivalent Landsat band numbers 86

Table 6.1 Subtractive colour mixing 92

Table 8.1 Covariance matrix calculation 111

Table 9.1 The classification process 130

Table 9.2 Comparison of image and cover classes 134

Table 9.3 Minimum spanning tree linkages 144

Table 9.4 Minkowski metrics 146

List of Excurses

Excursus 1.1—Metadata Standards 8
Excursus 2.1—Intertidal Extents Model 15
Excursus 3.1—Australian Geophysical Observing Network 39
Excursus 4.1—Channel Statistics 52
Excursus 5.1—Human Colour Perception 68
Excursus 5.2—Colour Composition 75
Excursus 5.3—Landsat band numbers 86
Excursus 8.1—Band Correlation Statistics 110
Excursus 8.2—Spatial Statistics 119
Excursus 8.3—Continental Scale Calibration of ASTER imagery 122
Excursus 9.1—Example Classification 133
Excursus 9.2—Colour Selection 156
Excursus 9.3—Mosaic Model Classification 157
Excursus 10.1—Geographic Object-Based Image Analysis 166

Introduction

Image processing allows digital image data to be displayed, enhanced, analysed or integrated with other datasets. A wide variety of image processing operations can be used to enhance the information context of imagery, such as highlighting or detecting features of interest, changing the image geometry, extracting patterns, removing noise, analysing time series trends or reducing data volume. Image processing methods rely upon repetitive mathematical operations to achieve these results.

While most image datasets can be processed on individual computers using appropriate software, extremely large datasets may require specialised hardware and software to perform image processing operations with flexibility and efficiency. Increasingly the storage and processing of very large image datasets is undertaken in internet-based (or 'cloud-based') data stores, using open access software libraries.

The following three sections describe:

- digital image structure and formats (see Section 1);
- processing stages and methodologies (see Section 2); and
- calibration approaches (see Section 3).

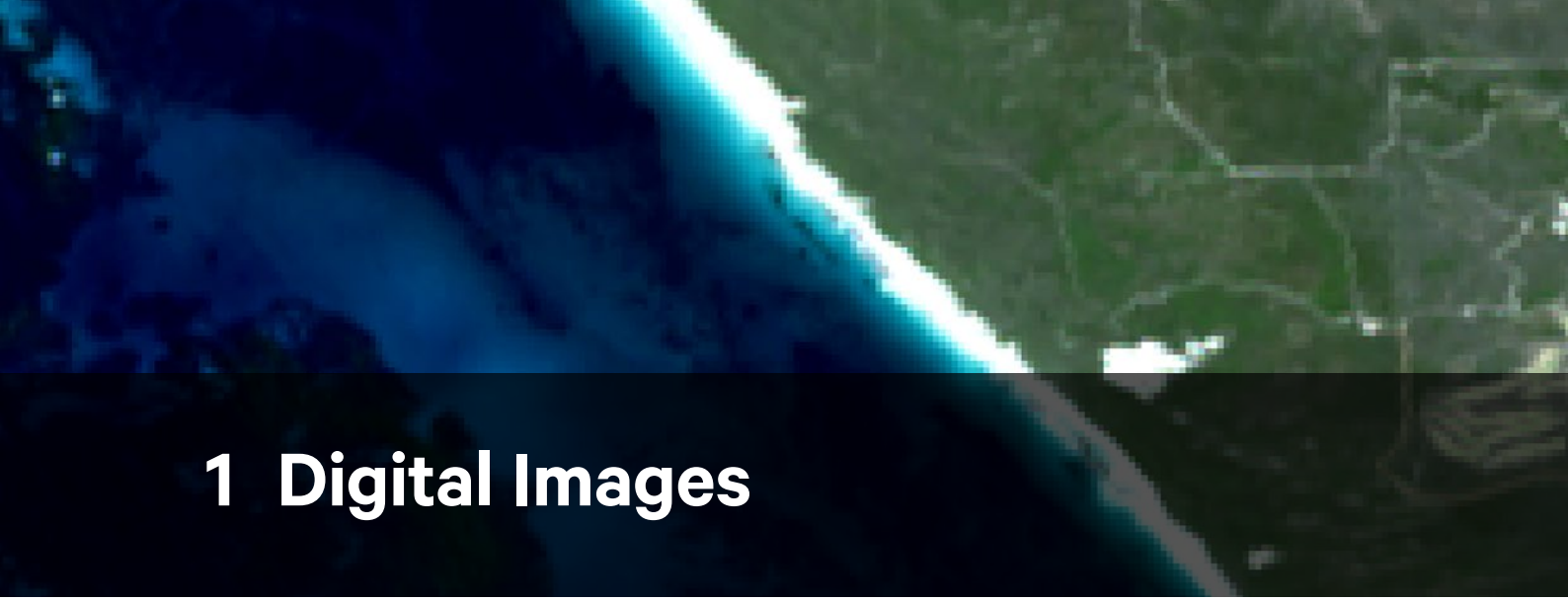
Contents

1 Digital Images 3

2 Processing Stages 13

3 Calibration 21

Background image on previous page: Landsat-5 image of Wilson's Promontory, Victoria, acquired on 30 January 2011, shown using bands 3, 2, 1 as RGB.
Source: Norman Mueller, Geoscience Australia



1 Digital Images

Any pictorial or spatial data may be represented as a digital, or numeric, image. An image needs to be in digital form to be processed by computer software. Spatial data, such as a digital photograph, is represented as an image using a pattern of non-overlapping cells (or a 'tessellation') with each cell being associated with the observed 'colour' of some portion of the original object or picture. In remotely sensed data, colour usually refers to radiance in discrete wavelength bands or channels. The differences between digital and photographic images, as well as various aspects of image resolution (spectral, spatial, temporal and radiometric), are described in Volume 1B—Section 1 of this series.

1.1 Image Structure

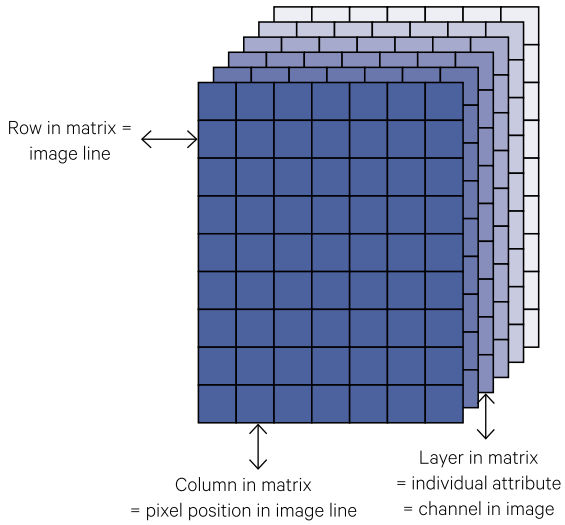
Image processing systems assume that the tessellation used to record the values in an image consists of a regular cell size, and usually represent the tessellation as a grid. In image processing each cell in the image grid is referred to as a 'pixel' (picture element). A single channel in a digital image can be considered as a two-dimensional table, or matrix, with the pixels across a line representing columns, and the lines down the image representing rows. Similarly, in a multi-channel image, each channel can be viewed as an attribute or dimension in a three-dimensional table.

The entries in the table represent an observation of some attribute or variable (see Volume 1A—Section 1). In the case of remotely sensed data, the attribute (or image channel) is a scaled record, in a given wavelength band, of the energy reflected or emitted by the Earth's surface. This concept is illustrated in Figure 1.1. Image pixel locations are generally referenced by their pixel (column) and line (row) positions relative to an origin in the top-left corner of the image (see Section 7 below).

Digital images allow the use of a large range of mathematical and statistical operations to enhance, reduce or identify particular features in spatial data. Such operations analyse the colour, or spectral, characteristics and/or the pattern, or spatial attributes of an image as described in Sections 8 to 10 below and Volumes 2C and 2E. The geometric properties of image data also allow operations to be applied to modify the spatial relationship between features by changing image shape, scale and/or orientation (see Section 7 and Volume 2B). Consideration of image geometry is important when comparing multiple images from the same or different sensors, including time series datasets (see Volume 2D). Ongoing improvements in the capabilities of image processing software enable some geometric operations to be applied to imagery concurrently with other processes.

Figure 1.1 Components of a digital image

A multi-channel image comprises three data dimensions—pixels, lines and channels. Pixels (or 'elements' or 'samples') along the line can be considered as columns in a matrix. Lines in the image are like rows of a matrix, with each channel being an attribute layer in the matrix.



Background image: Landsat-8 image of Deepene Beach, WA, acquired on 17 December 2016. This natural colour composite is based on bands 4, 3 and 2 as RGB.
Source: Norman Mueller, Geoscience Australia

Earth Observation (EO) image pixels contain estimates of surface radiance in particular wavelengths. However, image channels could also contain values that indicate other spatial attributes. Such attributes can be represented as either a scaled measure of some parameter, such as temperature or height at each pixel, or a nominal value that indicates a category, such as soil type, land use or cadastral region. The process of integrating other data sources with remotely sensed imagery is detailed in Volume 2D.

1.2 Image Data Formats

A wide range of file formats have been devised to store image data. This section briefly includes the most commonly encountered generic formats. Various forms of metadata are often added before or after these generic formats in specific proprietary and industry standard file types (see Section 1.3).

As introduced above, image data can be viewed in terms of three dimensions of data:

- pixels across each line;
- lines down the image; and
- channels (or bands) of pixels and lines (see Figure 1.1).

These image data dimensions are generally recorded using one of three standard ordering formats:

- Band Sequential (BSQ);
- Band Interleaved by Pixel (BIP); or
- Band Interleaved by Line (BIL).

The implementation of each format is illustrated in Figure 1.2d, Figure 1.2e and Figure 1.2f respectively. In any of these file formats, all pixels in an image channel have the same fixed range of possible values, such as:

- one bit (range 0–1);
- one byte (usually 8 bits: 0–255);
- one integer (0–32767);
- a signed integer (-32768–32767); or
- floating point (positive or negative numbers including decimal places).

Each file format generally reserves one value in the potential data range to indicate missing values, which is called the 'null' (or no data) value.

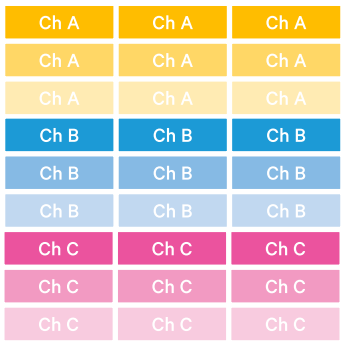
Figure 1.2 Image data formats

a. Example image data for channel A

b. Example image data for channel B

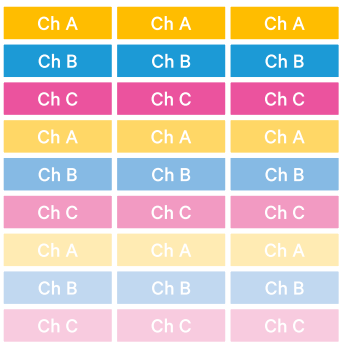
c. Example image data for channel C

d. Band sequential file format (BSQ)



e. Band interleaved by pixel format (BIP)

f. Band interleaved by line format (BIL)



Adapted from: Chuvieco and Huete (2010)

1.2.1 Band sequential (BSQ)

This format records all the data for each band or channel in the image as a continuous block of information (see Figure 1.1d). Within each channel block, the ordering is usually by line, that is:

channel 1:
line 1: pixel 1, pixel 2, ... pixel m
line 2: pixel 1, pixel 2, ... pixel m
...
line n: pixel 1, pixel 2, ... pixel m

channel 2:
line 1: pixel 1, pixel 2, ... pixel m
line 2: pixel 1, pixel 2, ... pixel m
...
line n: pixel 1, pixel 2, ... pixel m
...
...

channel z:
line 1: pixel 1, pixel 2, ... pixel m
line 2: pixel 1, pixel 2, ... pixel m
...
line n: pixel 1, pixel 2, ... pixel m

This format allows data for an individual channel to be retrieved quickly but requires several read operations to get all the data for all channels relating to a particular pixel or line.

1.2.2 Band interleaved by pixel (BIP)

This format records all channels for each image pixel column as sequential file data (see Figure 1.1e):

line 1:
pixel 1: channel 1, channel 2, ... channel n
pixel 2: channel 1, channel 2, ... channel n
...
pixel z: channel 1, channel 2, ... channel n

line 2:
pixel 1: channel 1, channel 2, ... channel n
pixel 2: channel 1, channel 2, ... channel n
...
pixel z: channel 1, channel 2, ... channel n
...
...

line m:
pixel 1: channel 1, channel 2, ... channel n
pixel 2: channel 1, channel 2, ... channel n
...
pixel z: channel 1, channel 2, ... channel n

I'm not dumb. I just have a command of thoroughly useless information (Calvin, Calvin and Hobbes)

This format allows rapid access to columns of image data for all channels so is useful for data requiring pixel-based processing that uses multiple channels, such as many hyperspectral analyses. Access to complete channel information, however, is slower than with BSQ format data.

1.2.3 Band interleaved by line (BIL)

This format records all channels for each image line as consecutive data in the file (see Figure 1.1f):

line 1:
channel 1: pixel 1, pixel 2, ... pixel m
channel 2: pixel 1, pixel 2, ... pixel m
...
channel z: pixel 1, pixel 2, ... pixel m

line 2:
channel 1: pixel 1, pixel 2, ... pixel m
channel 2: pixel 1, pixel 2, ... pixel m
...
channel z: pixel 1, pixel 2, ... pixel m
...
...

line n:
channel 1: pixel 1, pixel 2, ... pixel m
channel 2: pixel 1, pixel 2, ... pixel m
...
channel z: pixel 1, pixel 2, ... pixel m

The BIL format allows all data for a particular image line to be accessed quickly but separates the data for each individual channel so that complete channel information is retrieved more slowly than for BSQ. This format is most appropriate for EO data since many processing operations require the analysis of both multiple channel values and groups of pixels, that is, it is a good compromise between BSQ and BIP. Since various correction techniques process the data for individual image lines as discrete entities, this format enables efficient data access.

1.3 Data Volume

The size of an uncompressed image file, in bytes, can be computed as:

number of pixels x *number of lines* x *number of channels* x *number of bytes per pixel* + *metadata*

To reduce the amount of disk space required to store an image data file, data compression methods can be used in conjunction with various data formats. These methods can be divided into two categories, which describe the impact of compression on the original data:

- loseless—maintains the original data, which can be recovered by decompression; and
- lossy—reduces the original data by permanently removing information that the compression algorithm considers to be redundant, such that the removed information cannot be recovered by decompression. Lossy compression algorithms are suited to visual representation of imagery but should not be used when further processing of image data is anticipated, or for any form of data where specific values represent particular categories or null pixels.

One of the most commonly used compression methods for thematic data is 'run-length encoding' which records the number of sequential pixels along a line which have the same value, then the value itself, rather than one value for each pixel. This lossless method reduces the number of values to be stored in the image file when adjacent pixels in the line frequently have the same value but increases the time required to read or write the file. Such methods are most appropriate for categorical data sets, however their effectiveness on unclassified EO imagery depends on the sensor and ground features being imaged. In general terms, in imagery where adjacent pixels rarely have the same values, there is less reduction in data volume when using this method to encode data values (see Table 1.1).

Table 1.1 Example of data compression rates from run length encoding

Different images were compressed using run length encoding to demonstrate the resulting reduction in data volume. All images were stored in 16 bit integer format, with scaled reflectance values.

Image	Description	Data volume reduced by
Landsat-8 of rural NSW	Full scene, including null pixels on outside edges	25%
	Subset without null pixels	27%
Landsat-5 of NSW south coast	Full scene, including null pixels on outside edges	48%
	Subset with 1/3 water pixels	47%
	Subset of agricultural land	58%
	Subset of classified image containing mixed rural/woodland with minor urban component	78%

Source: Tony Sparks, Icon Water

1.4 Metadata

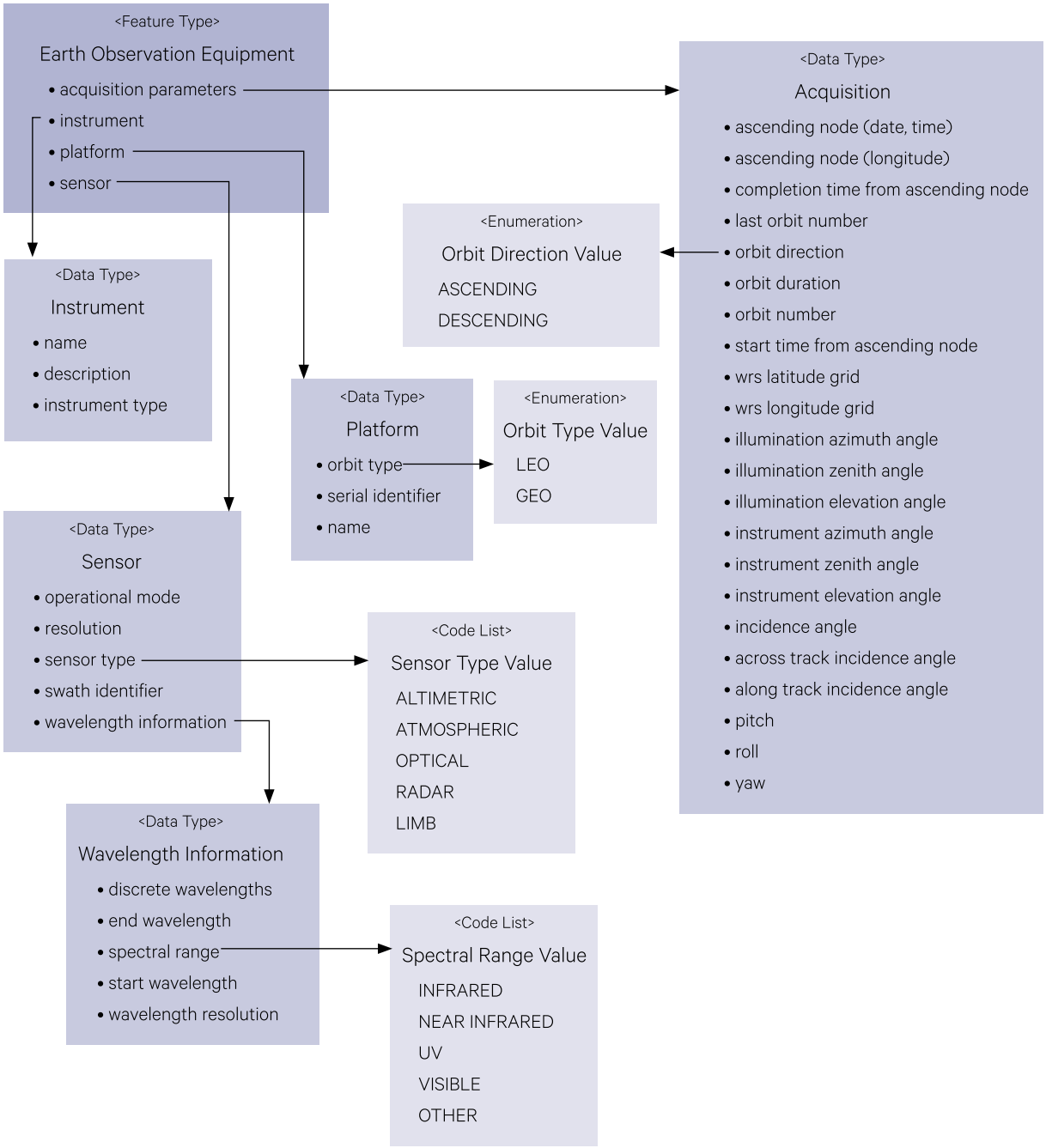
EO image files generally include additional information relating to the source and size of the image and its processing history. This information is called metadata (data about data) and is typically written in an image file as one or more header records before the image data. The goal of metadata for geophysical datasets is that the data becomes more portable, more accessible and 'self-describing' (Retscher *et al.*, 2011). For example, metadata elements for EO datasets include information relating to the acquisition platform, acquisition sensor, data processing history and spatial representation.

Standards for metadata of EO datasets have been developed to maximise consistency in data transfer and distinguish between:

- structural metadata—describes how data components are recorded as data structures; and
- descriptive metadata—describes the data components.

For example, the Open Geospatial Consortium (OGC) metadata describing the equipment used during EO data acquisition is shown in Figure 1.3 and the description of the EO result is given in Figure 1.4 (OGC, 2012). The Generic Earth Observation Metadata Standard (GEOMS) is commonly used to exchange validation data used in conjunction with EO datasets (Retscher *et al.*, 2011).

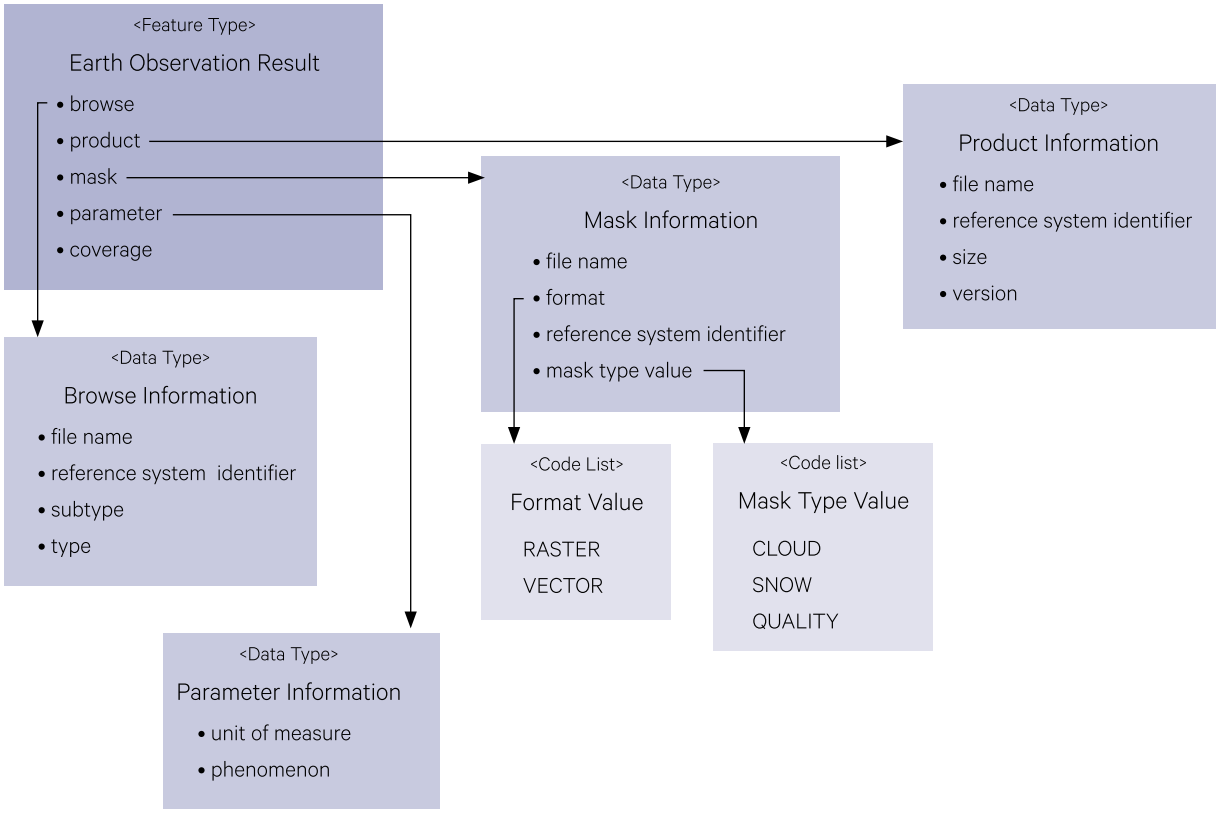
Figure 1.3 Earth Observation equipment metadata



Adapted from: OGC (2012) 8.1.2

The Sensor Web, envisaged as encircling the globe with electronic sensing devices during the next century, will add a new level of sophistication for observing and measuring the Earth, and calibrating and validating EO imagery (van Zyl *et al.*, 2009; Bröring *et al.*, 2011). Accessing and interrogating a system of this scale and complexity will take metadata to a new level of sophistication and utility (Hu *et al.*, 2014).

Figure 1.4 Earth Observation result metadata



Adapted from: OGC (2012) 8.1.4

Excursus 1.1—Metadata Standards

Source: David Hudson, Geoscience Australia

The term 'metadata' is used to describe data about data. There are many different types of metadata and many different metadata standards.

The most common standards used for EO data are:

- data labelling standards:
  - all spatial data—ISO 19115-1 (*https://www.iso.org/standard/53798.html*)
  - raster data—ISO 19115-2 (*https://www.iso.org/standard/39229.html*)
- data transport standards:
  - picture data—Web Map Service (*http://www.opengeospatial.org/standards/wms*)
  - vector data—Web Feature Service (*http://www.opengeospatial.org/standards/wfs*)
  - raster data—Web Coverage Service (*http://www.opengeospatial.org/standards/wcs*)
- EO product standard:
  - CEOS Analysis Ready Data for Land (CARD4L)—(CEOS, 2017)

In the broader context of spatial data, ANZLIC (Australia and New Zealand Land Information Council) is sponsoring the Foundation Spatial Data Framework (FSDF). This ongoing, collaborative initiative provides a common reference to assemble and maintain foundation level spatial data in Australian and New Zealand. The FSDF aims to deliver national coverage of the most current and authoritative foundation spatial data to serve the widest possible variety of users in both government and commercial activities. Ten themes for this standardised and quality-controlled data have been defined:

- administrative boundaries;
- elevation and depth;
- geocoded addressing;
- imagery;
- land cover and land use;
- land parcel and property;
- place names;
- positioning;
- transport; and
- water (ANZLIC, 2017).

1.5 Pixel Quality Flag

Some EO image suppliers include an additional channel with the image data to indicate the data quality of the original image pixel or derived products. For example, many recent Landsat products from Geoscience Australia (GA) include a Pixel Quality Flag value, which is the binary representation of the cumulative sum of 16 test conditions (Sixsmith, 2012). The test conditions used to determine each bit value are summarised in Table 1.2. An example of interpreting this channel is provided in Table 1.3. Such pixel quality flags assist users in checking that the imagery or derived products are reliable and suitable for analysis and interpretation (Sixsmith *et al.*, 2013).

Table 1.2 Pixel quality flag

Example of pixel quality flag values used for Landsat images by Geoscience Australia.

Test Condition	Bit Position (right to left)	Pixel Quality Flag equals 0 if this pixel is	Pixel Quality Flag equals 1 if this pixel is
Band 1 is saturated	0	Saturated in Band 1	Not saturated in Band 1
Band 2 is saturated	1	Saturated in Band 2	Not saturated in Band 2
Band 3 is saturated	2	Saturated in Band 3	Not saturated in Band 3
Band 4 is saturated	3	Saturated in Band 4	Not saturated in Band 4
Band 5 is saturated	4	Saturated in Band 5	Not saturated in Band 5
Band 6-1 is saturated (This is designed to match Landsat-7 ETM+. The thermal band for Landsat-5 TM corresponds to Band 6-1 and this value is duplicated in Band 6-2.)	5	Saturated in Band 6-1	Not saturated in Band 6-1
Band 6-2 is saturated (This is designed to match Landsat-7 ETM+. The thermal band for Landsat-5 TM corresponds to Band 6-1 and this value is duplicated in Band 6-2.)	6	Saturated in Band 6-2	Not saturated in Band 6-2
Band 7 is saturated	7	Saturated in Band 7	Not saturated in Band 7
Contiguity	8	Null value in at least one band	All bands for this pixel contain non-null values
Land or Sea (based on vector data from TM World Borders (*http://thematicmapping.org/downloads/world_borders.php*) spliced with Australian Coastal Boundary using a 100 m buffer into the sea.)	9	Sea	Land
Cloud (based on Automatic Cloud Cover Assessment (ACCA) algorithm; Irish, 2000; Irish ***et al.***, 2006)	10	Cloud	No Cloud
Cloud (Fmask; based on TOAR rather than NBAR as input reflectance; (Zhu and Woodcock, 2012)	11	Cloud	No Cloud
Cloud Shadow (only computed for cloud detected by ACCA algorithm using in-house algorithm)	12	Cloud Shadow	No Cloud Shadow
Cloud Shadow (Fmask; based on TOAR rather than NBAR as input reflectance; (Zhu and Woodcock, 2012))	13	Cloud Shadow	No Cloud Shadow
Topographic Shadow (This feature has not been implemented yet so bit 14 is set to value 0)	*14*	*Topographic Shadow*	*No Topographic Shadow*
To be determined (Default untested value currently set to 0)	*15*	*N/A*	*N/A*

Table 1.3 Examples of pixel quality flag values

Four examples of pixel quality flag values defined in Table 1.2.

Example	Binary Code (for bits 0 to 15)																Equivalent Decimal value	Interpretation of Binary Code
	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0		
1	0	0	1	1	0	0	1	1	1	1	1	0	1	1	1	0	13294	Bands 1 and 5 are saturated (bits 0 and 4 = 0) No null values (bit 8 = 1) Land pixel (bit 9 = 1) Clouds detected by ACCA and Fmask (bits 10 and 11 = 0) No cloud shadow (bits 12 and 13 = 1)
2	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	4095	No saturated bands (bits 0–7=1) No null values (bit 8 = 1) Land pixel (bit 9 = 1) No cloud detected (bits 10 and 11 = 1) Cloud shadow detected by ACCA and Fmask (bits 12 and 13 = 0)
3	0	0	1	1	0	0	1	1	0	1	1	0	1	1	1	0	13166	Bands 1, 5 and 7 are saturated (bits 0, 4 and 7 = 0) No null values (bit 8 = 1) Land pixel (bit 9 = 1) Clouds detected by ACCA and Fmask (bits 10 and 11 = 0) No cloud shadow (bits 12 and 13 = 1)
4	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16383	Clear Land Pixel—no saturation, null values, cloud, or cloud shadow

1.6 Hierarchical Storage Structures

A range of digital storage structures has been developed to access large images more efficiently. Several hierarchical data structures have been developed to optimise storage and retrieval of large image-based datasets. Three implementations of such structures include:

- pyramids of downscaled image tiers which effectively store image at a range of scales (see Volume 1B—Section 1.2). These tiers enable imagery to be rapidly displayed at a range of scales;
- mosaics (also called overviews or tilesets) of smaller images which together cover a larger area, possibly with overlapping extent and different geometric properties (see Section 1.7); and
- datacubes of calibrated, conformal image layers through time (see Section 3.4).

Each of these structures offer different advantages for displaying, interrogating and processing large image-based datasets. As the volume of image data increases and its spatial resolution decreases, efficient access to data is becoming critical for many applications. Some interchange formats used with EO imagery include:

- GeoTIFF (Georeferenced Tagged Image File format)—an open file format based on TIFF for georeferenced raster imagery and embraced by NASA Earth Science data systems;
- HDF (Hierarchical Data Format)—multi-object file format used for all NASA EO data products;
- NetCDF (Network Common Data Form)—set of self-describing, generic data formats (and software libraries) for multidimensional, array-oriented scientific data; and
- JPEG—(Joint Photographic Experts Group)—most common standard for digital photography. Many JPEG implementations use lossy compression, which render smooth variations in tone and colour but reduces image fidelity (see Section 1.7).

1.7 Further Information

Metadata

Open Geospatial Consortium (OGC): (OGC, 2012)

The Generic Earth Observation Metadata Standard (GEOMS): Retscher *et al.* (2011)

Foundation Spatial Data Framework (FSDF): *http://www.anzlic.gov.au*

Sensor Web

Hu *et al.* (2014)

Virtual Constellations: *http://ceos.org/ourwork/virtual-constellations/*

Data structures

GEOTIFF: *http://trac.osgeo.org/geotiff/* (*https://earthdata.nasa.gov/standards/geotiff*)

HDF: *http://nsidc.org/data/hdfeos/*

NetCDF: *https://www.unidata.ucar.edu/software/netcdf/*

JPEG: *https://jpeg.org/jpeg2000/index.html*

Tileset Concepts and Terminology: *http://www.microimages.com/documentation/FeatureSummaries/2010/Tileset%20Terminology.pdf*

1.8 References

ANZLIC (2017). *Australian and New Zealand Land Information Council (ANZLIC) website*. Retrieved from *http://www.anzlic.gov.au*.

Bröring, A., Echterhoff, J., Jirka, S., Simonis, I., Everding, T., Stasch, C., Liang, S., and Lemmens, R. (2011). New Generation Sensor Web Enablement. Sensors, 11(3), pp. 2652. doi:*http://dx.doi.org/10.3390/s110302652*.

CEOS (2017). *CEOS Analysis Ready Data for Land (CARD4L) Description Document*. Retrieved from: *https://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiR-_OCosLXAhWLnJQKHRpUCOsQFggoMAA&url=http%3A%2F%2Fceos.org%2Fdocument_management%2FMeetings%2FPlenary%2F30%2FDocuments%2F5.5_CEOS-CARD4L-Description_v.22.docx&usg=AOvVaw2r8F89fp--1nxlALHTQsAi*.

Chuvieco, E., and Huete, A. (2010). *Fundamentals of Satellite Remote Sensing*. CRC Press, Boca Raton.

Hu, C., Guan, Q., Chen, N., Li, J., Zhong, X., and Han, Y. (2014). An Observation Capability Metadata Model for EO Sensor Discovery in Sensor Web Enablement Environments. *Remote Sensing*, 6(11), pp. 10546.

Irish, R. R. (2000). *Landsat 7 automatic cloud cover assessment*. Paper presented at the SPIE 4049, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Orlando, United States.

Irish, R. R., Barker, J. L., Goward, S. N., and Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm. *Photogrammetric Engineering & Remote Sensing*, 72(10), pp. 1179-1188. doi:*http://dx.doi.org/10.14358/PERS.72.10.1179*.

OGC (2012). *Earth Observation Metadata Profile of Observations and Measurements* (OGC 10-157r3) (J. Gasperi, F. Houbie, A. Woolf, and S. Smolders Eds.). Open Geospatial Consortium.

Retscher, C., de Maziere, M., Meijer, Y., Vik, A. F., Boyd, I., Niemeijer, S., Koopman, R. M., and Bojkov, B. (2011). *The Generic Earth Observation Metadata Standard (GEOMS), Version 1.0*. Retrieved from *http://avdc.gsfc.nasa.gov/PDF/GEOMS/geoms-1.0.pdf*.

Sixsmith, J. (2012). *Unlocking the Landsat Archive Work Package 4.1. Pixel Quality Documentation* (GA TRIM Number D2012-212824). Geoscience Australia.

Sixsmith, J., Oliver, S., and Lymburner, L. (2013). *A hybrid approach to automated landsat pixel quality*. Paper presented at the IEEE Geoscience and Remote Sensing Symposium (IGARSS), Melbourne, Australia.

van Zyl, T. L., Simonis, I., and McFerren, G. (2009). The Sensor Web: systems of sensor systems. *International Journal of Digital Earth*, 2(1), pp. 16-30. doi:*http://dx.doi.org/10.1080/17538940802439549*.

Zhu, Z., and Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, pp. 83-94. doi:*http://dx.doi.org/10.1016/j.rse.2011.10.028*.

2 Processing Stages

Image processing techniques can be applied to a wide variety of data in a diverse range of disciplines. These techniques rely on 'spectral' and/or 'spatial' analysis. In spectral analysis, the values of individual pixels in an image are used to identify particular features. In spectroscopy, for example, identification of a sample is based on its spectral response in a large number of channels. Alternatively, in spatial analysis, the patterns formed by groups of pixels in an image may be used to identify or enhance features on the basis of their shape, size or proximity. For example, image spatial patterns are particularly important in visual interpretation of images. Image processing of EO data combines both spectral and spatial analyses to interpret, analyse and model image data.

2.1 Data Provenance

At all stages of processing EO imagery, analysts need to be mindful of the image data provenance, that is:

- its intrinsic resolutions and extents in terms of spectral, spatial, radiometric and temporal properties (see Volume 1); and
- any pre-processing that has been applied to the data (see Section 3 below).

Selection of appropriate processing sequences for calibration, analysis, modelling and integration of EO datasets requires both an understanding of data provenance and insight into the assumptions and constraints of the component algorithms. All EO products should be produced using mature, operational processing pathways, which are documented in terms of:

- data sources (primary EO inputs, metadata parameters and any ancillary data used);
- major algorithms, described in a relevant Algorithm Theoretical Basis Document (ATBD);
- processing sequence, as a series of steps from input to output;
- validation of product or its underlying algorithm; and
- statement of accuracy and limitations relating to product and/or input data.

It should be noted that image data can represent two fundamental types of information. The value of each pixel in a given image channel is derived from either:

- a continuous range of values for some characteristic, such as the acquired radiance in optical or thermal imagery; or
- a set of categories, such that different categories (or classes) are represented by different channel values. The results from classifying EO imagery into surface cover classes, for example, are generally encoded into a single classification channel (see Section 9 and Volume 2E).

The types of image processing tools and methods that are appropriate for continuous data may not be appropriate for categorical data and, conversely, processes that can be applied to categorical data may not be suitable for continuous data. In this context, knowledge of data provenance is essential for meaningful and reliable results.

Image products derived from EO data sources have traditionally been grouped in terms of their processing history. Here, image products will be considered in terms of two broad types:

- data products—corrected image observations from EO sensors (see Section 2.1.1); and
- information products—derived from EO data products for specific applications (see Section 2.1.2).

Background image: Astronaut photo of Roebuck Bay, just south of Broome, WA, taken on 11 June 2015 from the International Space Station (ISS) at an altitude of 402 km. **Source:** NASA. Retrieved from: *https://eol.jsc.nasa.gov/SearchPhotos/photo.pl?mission=ISS044&roll=E&frame=885*

2.1.1 Data products

In the context of EO imagery, data products comprise corrected image observation values acquired from EO sensors. The observation values are presented in grid cells and represented as radiance and/or digital numbers. The corrections applied to these values include geometric, radiometric, and atmospheric adjustments to compensate for known distortions in the original EO image data.

Preparation of data products does not involve interpretation models or classification procedures. These products provide EO data for users to process for their own purposes and correspond to the traditional EO image scenes, such as terrain-corrected (also known as ortho-corrected) and map-oriented (with no terrain correction) imagery. Examples of GA EO data products include the Australian Reflectance Grid (ARG) suite of products. The ARG25 product, for example, delivers an estimate of surface reflectance from Landsat TM/ETM+ data using physical rather than empirical models. This product aims to deliver imagery in which differences in the reflective value(s) between images that were acquired at different times by different sensors[2] will be primarily due to on-ground changes in biophysical parameters rather than artefacts of the imaging environment. These products are processed to deliver standardised surface reflectance measures and, in future, are intended to be conformal to the National Nested Grid (NNG; see Section 7).

2.1.2 Information products

Information products derived from EO data are also available for specific applications. These products have been developed to highlight specific features of interest for particular users and comprise interpreted values derived from classification and/or modelling techniques applied to the corrected image values. These are represented as grid cell values and can depict a broad range of information, such as:

- vegetation indices (NDVI, EVI—see Volume 3A);
- temperature (K);
- position (x, y coordinates);
- fractional cover; or
- land cover (e.g. water, bare soil, vegetation type).

Data channels in these products can be either continuous (such as range of NDVI values) or categorical (such as land cover categories). Examples of information products produced by GA include the Australian Geographic Reference Grid (AGRI), Sentinel Hotspots, the Dynamic Land Cover Dataset (DLCD) and the Intertidal Extent Model (see Excursus 2.1).

You can have data without information,
but you cannot have information without data.
(Daniel Keys Moran)

2 Notwithstanding differences due to variations in spectral bandwidth between sensors.

Excursus 2.1—Intertidal Extents Model

Source: Stephen Sagar, Geoscience Australia
Further Information: (Sagar *et al.*, 2017)

An example of continental scale modelling utilising the Landsat archive

The InterTidal Extents Model (ITEM) (Sagar *et al.*, 2017) provides valuable information regarding the exposed intertidal zone—the area between the land and sea that can be observed between the highest and lowest tide. Traditionally, surveying this dynamic landscape has been difficult, especially when mapping large expanses of shoreline—some 50,000 km for the Australian coast.

The process used to create the ITEM product leverages the full 28 year time series archive of Landsat data, managed in a countrywide, analysis ready data platform, known as the Australian Geoscience Data Cube (AGDC; see Section 3.4). To fully exploit the time series of data, ITEM moves away from a single scene analysis method to a full ensemble modelling framework by implementing two key concepts:

- through the use of an ancillary tidal model, the ensemble of Landsat observations is reframed and reordered from the traditional time series domain to one based on tidal height; and
- the full reordered ensemble of data is then examined using a median estimator to account effectively for data noise and incompleteness.

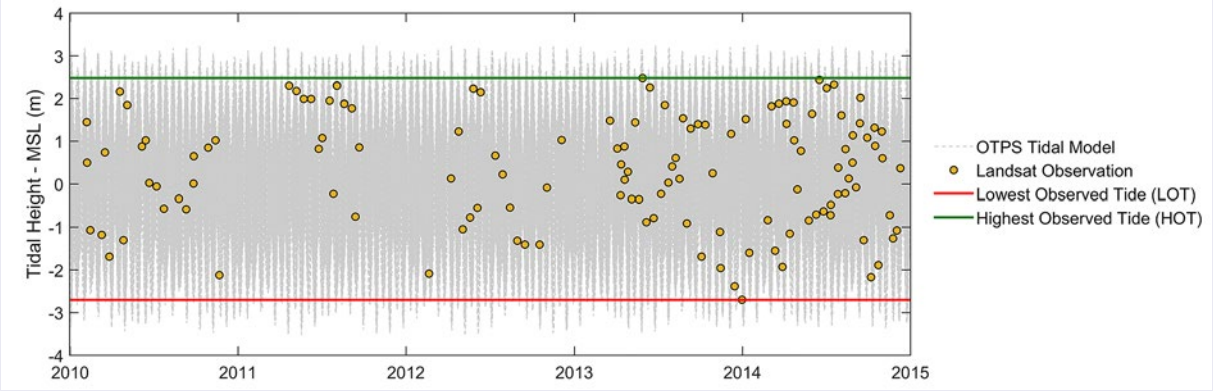
This concept illustrates how EO data analysis can be driven by other environmental attributes, enabling a new range of applications to grow from the traditional time series approach.

Processing steps

1. The modelling process to derive the product utilises a continental scale tidal model developed by Oregon State University (Egbert and Erofeeva, 2002, 2010) applied to the 221 AGDC cells that cover the Australian coastline.
2. Each Landsat tile in these coastal cells is attributed with a tidal height, relative to Mean Sea Level (MSL), corresponding to the time of observation acquisition. Tidal heights are modelled at a single constant location for each 1° by 1° cell, determined manually and based on the complexity of the coastline contained within the cell.
3. The time series of observations for each AGDC cell is then sorted based on the modelled tidal height, and split into subsets representing 10% percentile intervals of the Observed Tidal Range (OTR) for that cell. Due to the Sun synchronous nature of the various Landsat sensor observations, it is unlikely that the full physical extent of the tidal range in any cell will be observed. Hence, terminology has been adopted for the product to reflect the highest modelled tide observed in a given cell (HOT) and the lowest modelled tide observed (LOT; see Figure 2.1). These measures are relative to MSL and have no consistent relationship to Lowest Astronomical Tide (LAT) and Highest Astronomical Tide (HAT).
4. Each tile observation is masked for pixel quality based on the AGDC PQA layer to exclude pixels flagged for cloud, band saturation and contiguity (see Section 1.5). The Normalised Difference Water Index (NDWI) (McFeeters, 1996) is calculated for each tile, and a pixel-based median NDWI composite is derived for each 10% percentile interval of the cell's OTR.
5. NDWI composites derived for each cell are used to estimate the extent of exposed substratum at each 10% interval of the OTR through a classification of water/non-water pixels, and combined to create the ITEM Relative Extents Model (see Figure 2.2).
6. Variability of the NDWI values contributing to the tidal interval median composites is analysed for each pixel. This produces a confidence layer output that indicates regions in which the observed water/land interactions are not driven by tidal factors. This layer helps to identify modelling errors and issues such as coastal instability over time.

Figure 2.1 The concept of tidal attribution and Lowest (LOT) and Highest (HOT) Observed Tide.

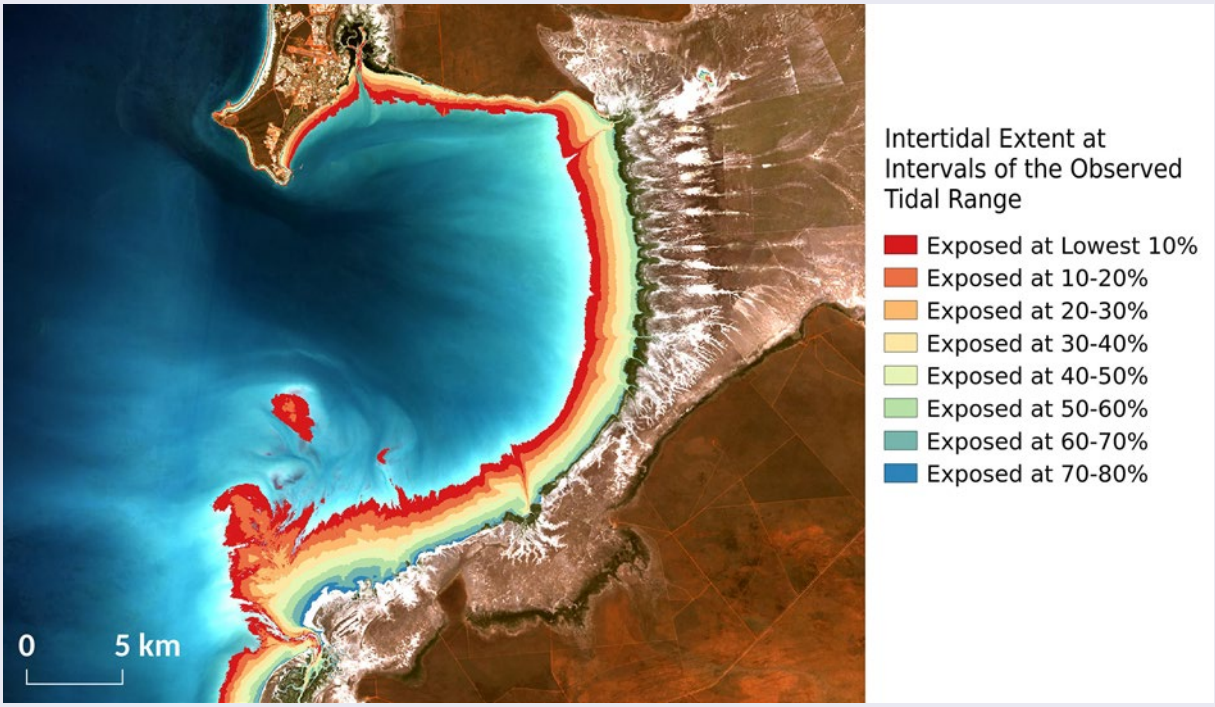
Modelled tidal heights for Landsat observations shown in relation to the full tidal model. Note how the observed tidal range (HOT-LOT) differs from the full tidal range.



Source: Stephen Sagar, Geoscience Australia.

Figure 2.2 Roebuck Bay Example

Example of the ITEM Relative Extents Model showing the intertidal flat extents and topography at Roebuck Bay, WA, a site of ecological significance for migratory shorebirds.



Source: Stephen Sagar, Geoscience Australia.

The modelling approach enables the creation of a continuous intertidal extent map for the full Australian coastline. Since tidal extent is a function of the tidal height and the shoreline elevation, the data set also provides relative topography, and thus a framework to derive coastal elevation models. This is an example of how EO products can act as a complement to existing data, such as ship-based bathymetry, enabling a more realistic representation and a deeper understanding of Australia's vast coastline.

The knowledge provided by ITEM can contribute to a broad range of applications, including environmental monitoring applications for migratory bird species, habitat mapping in coastal regions, hydrodynamic modelling, and examining the geomorphology of the intertidal zone.

2.2 Processing Methodologies

Although some image processing exercises will not include all stages, processing of EO imagery can be considered within the context of three broad functional stages:

- descriptive interpretation—this is generally an initial process of visually delineating and identifying features or patterns in an image. Digital image processing can be useful during this stage to improve feature contrast and colour (see Sections 4 and 5) or modify image geometry (see Section 7 and Volume 2B);
- exploratory data analysis—involves using statistical tools to identify features and patterns in the image. This stage produces results such as land cover maps from EO imagery; and
- modelling—relates image values to data from other sources and accesses a wide variety of transformations to determine the relationships between different data types.

Table 2.1 uses some example applications to illustrate the relationship between different processing stages and spectral/spatial data attributes. Image processing is primarily used for exploratory data analysis but draws on both colour and shape (that is spectral and spatial, data and tools) and can be used to prepare data for both descriptive activities and modelling.

This Volume introduces image processing methods in terms of these three stages. Sections 4 to 6 cover the basic techniques that are relevant to visual image interpretation, including image enhancement, geometry, display and hardcopy. The image analysis topics of statistics, classification, and segmentation are described in Sections 8 to 10. Various image transformations, which may be used for interpretation, analysis or modelling, are detailed in Volume 2C. The detailed processing involved with more advanced functions of image rectification and classification are covered in Volumes 2B and 2E respectively. Processing of multiple images and their integration with other datasets is discussed in Volume 2D.

Table 2.1 Image processing stages

Stages of processing spectral and spatial data can be considered as description, analysis and modelling. Image processing is primarily used for exploratory data analysis and involves consideration of both the spectral and spatial attributes of image data.

Data Attribute		Stages of Processing		
Attribute	Example	Description	Analysis	Modelling
Colour	Spectra	Compiling spectral libraries, eg. AU-SPECCHIO	Compiling image time series, eg. GA data cube (see Section 7.4.2)	Quantifying identified image features, eg. oil spill analysis (see Volume 3B)
Imagery	Spatial Patches	Identifying major spatial features, eg. labelling land cover categories	Exploratory data analysis methods, eg. analysing area and distribution of land covers over time	Developing mathematical model to predict image values, eg. bio-optical modelling of water constituents (see Volume 3B)
Shape	Spatial Patterns	Interpreting spatial patterns, eg. aerial photography interpretation	Spatial analysis methods, eg. mathematical morphology	Inferring three-dimensional structure, eg. stereology/tomography or lidar modelling of forests (see Volume 3A)

2.3 Modelling

In many application areas, physics-based models are being developed and standardised to relate image values to measurable surface properties. To ensure that the value of EO imagery is optimised, such methods need to be fostered and endorsed by the wider scientific community. Processing methodologies for specific application areas are detailed in Volume 3.

This section overviews aspects of modelling that are relevant EO analyses. Subsequent volumes contain examples of specific models derived using EO datasets in various applications. The following text and sub-sections provides relevant background information on modelling terminology, prerequisites, and constraints.

2.3.1 Modelling Overview

Models describe various aspects of the physical world, including interactions and dynamics, using mathematics. They are typically applied to current datasets, which quantify current conditons, to predict future datasets, which should infer future conditions. Mathematical models can be developed which relate known real world properties to EO measurements. The modelled relationship can then to used to infer one entity from the other, either as a:

- forward model—to compute likely EO measurements from known real world properties; or as an
- inverse model—to compute likely real world properties from EO measurements (see Section 2.3.2).

A range of mathematical and algorithmic tools are used to develop appropriate models. Typically simpler, more robust models are less flexible and transferable than complex and computationally intensive models. Mathematical models are generally grouped into three categories:

- empirical or analytical models—use simple stochastic relationships, such as regression curves, between input and output data (see Section 2.3.3);
- semi-empirical or semi-analytical models—use *a priori* knowledge to determine a relationship between input and output data (see Section 2.3.4); and
- physics-based models—use knowledge about physical relationships between input and output data to define a relationship (see Section 2.3.5).

The unique perspective of EO imagery offers an attractive opportunity to model different aspects of planet Earth using appropriately-scaled data. A wide variety of models have been developed to interact with EO datasets for specific calibration, such as corrections for atmospheric and BRDF effects (see Section 3.3.2 and Appendices 1 and 2) and myriad applications, including phenology, water optics, carbon cycling, and climate studies. Some of these models are discussed in the context of specific application areas in Volume 3.

2.3.2 Forward and inverse models

Forwards models use data from the 'real world' to estimate a set of 'modelled' data, which may relate to a future time or state, or an alternate measurement mode. For example, in the context of hydrological EO, forward models start with known water properties and determine the upwelling radiation that would be measured by EO sensors.

Inverse models essential invert the forward modelling process, that is, they estimate 'real world' data from a set of 'modelled' data. In the context of EO imagery, inversion algorithms operate on a per pixel basis to determine the best combination of input parameters that would explain the spectral measurements in an EO image (Dekker *et al.*, 2001).

2.3.3 Empirical or analytical models

By definition, empirical models are based on data, rather than theory. Relevant measurements (or observations) of the 'input' variable and 'output' variable are used to develop, calibrate and test the relationship between them. The resulting model correlation enables an input value to be converted to an output value. Such models need to accurately relate the input and output variables and define the reliable constraint range(s) for input values. Empirical models for the remote sensing of surface waters, for example, relate the optical properties of water to their observed optical measurements in EO imagery (see Volume 3B).

2.3.4 Semi-empirical or semi-analytical models

Like empirical models, semi-empirical or semi-analytical models are based on data, but involve assumptions, approximations or generalisation to simplify the modelling process. *A priori* knowledge, for example, may be used to define the relationship between input and output data. In the context of EO, this may involve selecting the most appropriate spectral bands to develop a model. However, while such tailoring of data may improve model accuracy, the resulting modelling relationships may be less transferable to other locations and other sensors.

2.3.5 Physics-based models

Physics-based models use known physical laws to define the relationship between input and output variables. Physics-based models of the water column, for example, attempt to account for the energy transfer process through both the atmosphere and the waterbody (see Volume 3B). These models are more sophisticated, flexible and transferrable than empirical or semi-empirical models, but are signficantly more computationally intensive and require more detailed parameterisation data.

Essentially, all models are wrong, but some are useful
(George E. P. Box)

2.4 Further Information

Image Processing

The Pixel Page: *http://home.earthlink.net/~castleman/*

Image Processing Systems

ERDAS Imagine: *http://www.hexagongeospatial.com*

ERMapper: *http://www.hexagongeospatial.com*

ENVI: *http://www.harrisgeospatial.com*

TNTmips: *http://www.microimages.com*

2.5 References

Dekker, A. G., Brando, V., Anstee, J. M., Pinnel, N., Kutser, T., Hodgenboom, E., Peters, S., Pasterkamp, R., Vos, R., Olber, C., and Malthus, T. J. (2001). Imaging Spectrometry of Waters. Chapter 11 in '*Imaging Spectrometry: Basic Principles and Prospective Applications*' (Eds: F. van der Meer, and S. M. de Jong). Springer, Netherlands.

Egbert, G.D., Erofeeva, S.Y., 2010. *The OSU TOPEX/Poseiden Global Inverse Solution TPXO* [WWW Document]. TPXO8-Atlas Version 10. URL *http://volkov.oce.orst.edu/tides/global.html*.

Egbert, G.D., Erofeeva, S.Y., 2002. Efficient Inverse Modeling of Barotropic Ocean Tides. *J. Atmospheric Ocean. Technol.* 19, 183–204. doi: *https://doi.org/10.1175/1520-0426(2002)019%3C0183:EIMOBO%3E2.0.CO;2*

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432. doi:*10.1080/01431169608948714*

Sagar, S., Roberts, D., Bala, B., Lymburner, L., 2017. Extracting the intertidal extent and topography of the Australian coastline from a 28 year time series of Landsat observations. *Remote Sens. Environ.* 195, 153–169. doi:*10.1016/j.rse.2017.04.009*

3 Calibration

Various characteristics of EO imagery are introduced in Volume 1. While an observation recognises both the existence of the observed object and the action of recording something about it, a measurement provides a scaled record of such interactions and one that can be directly compared with similarly scaled measurements (see Volume 1A—Section 1). In order to relate EO observations to physical measurements of a target, images need to be calibrated to a known standard. Alternatively, to enable inter-comparisons, multiple images can be calibrated to a relative standard.

Section 3.1 provides an overview of calibration in the context of EO imagery while Section 3.2 reviews the need for calibration of remotely sensed imagery in terms of spectral, spatial, radiometric and temporal factors. Specific processing procedures that can be used to calibrate EO imagery are introduced in Section 3.3.

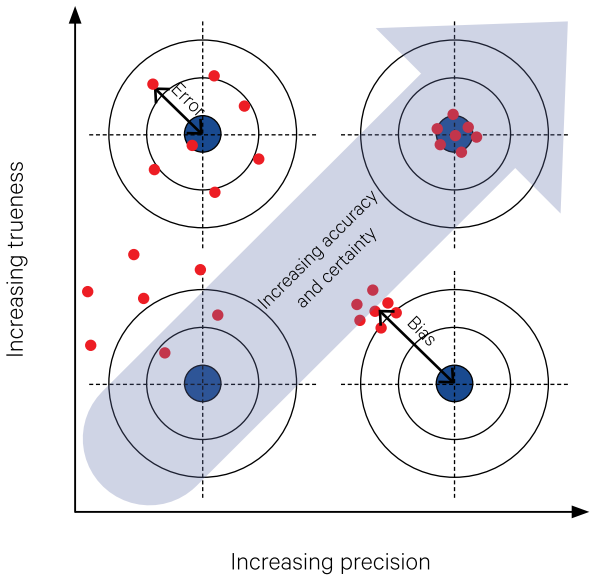
3.1 About Calibration

At this point, it is worth pondering the fact that all measurements of physical quantities are subject to uncertainties. Accordingly, it is inevitable that repeated measurements of the same quantity would yield variable results. Thus, every measurement implicitly comprises both a numeric value to estimate the quantity and a degree of uncertainty to indicate the inherent variability of the estimate. An accurate estimate of 'truth' generally involves repeated measurements of some traceable reference standard (see Figure 3.1; also ISO 57251).

When you can measure what you are speaking about and express it in numbers you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.
(Lord Kelvin, 1883)

Figure 3.1 Precision versus trueness

Trueness—closeness of agreement between the average value from a large series of test results and an accepted true value. Precision—closeness of agreement between independent measurements of a quantity under the same conditions.



Background image: Landsat-5 TM image of Lake Eyre acquired on 28 March 2011, displayed using bands 5,4,2 as RGB. A slanting striping pattern, resulting from imperfect calibration of the 16 TM detectors, is visible in the central water features. **Source:** Norman Mueller, Geoscience Australia

As illustrated in Figure 3.1, to appreciate the uncertainty inherent in all measurements, it is necessary to distinguish between:

- Trueness—closeness of agreement between the average value from a large series of test results and an accepted true value; and
- Precision—closeness of agreement between independent measurements of a quantity under the same conditions.

For EO data, the Committee of Earth Observation Satellites (CEOS) and Malthus and Li (2015) define relevant terminology as:

Calibration—process of quantitatively defining the responses of a system to known, controlled signal inputs;

Cross-Calibration—process of transferring the well-known radiometric calibration of one satellite sensor to another poorly-calibrated sensor via near-simultaneous imaging of a common ground target;

Traceability—property of a measurement result, which relates that result to a stated metrological reference through an unbroken chain of calibrations of a measuring system or comparisons, each contributing to the stated measurement uncertainty;

Uncertainty—parameter that characterizes the dispersion of the quantity values attributed to a measured mean, based on the information used;

Vicarious Calibration—techniques that make use of natural or artificial sites on the surface of the Earth for post (launch or flight) calibration of airborne or space-borne sensors; and

Validation—process of assessing, by independent means, the quality of the data products derived from the system outputs.

Being acquired by remote sensors, EO data necessarily involve one or more levels of inference or indirection (see Volume 1A—Section 1). As detailed in Volume 1, mechanisms for image acquisition and sampling inevitably produce distortions within the resulting data. These distortions can occur in terms of four intrinsic attributes:

- spectral—specific wavelengths recorded;
- spatial—location, area and scale of record;
- radiometric—energy levels recorded; and
- temporal—time of record.

Examples of image distortions include wide scan angles (resulting in larger ground pixels at the edges of a scan than in the centre) and varying Sun and observer positions (resulting in brightness variations which are unrelated to target brightness). Such distortions can impact interpretation, analysis and modelling of individual images and sets of imagery.

Typical processing scenarios include:

- single images, where classification and other image processing is predicated on consistent values from similar target surface covers;
- mosaics of images, which are intended to produce a consistent composite image for spectral image processing; and
- multiple images, such as time series datasets or imagery acquired by different sensors, where changes in the target surface are being detected between images and should not be confused with brightness variations resulting from imaging artefacts.

In each case, the underlying processing algorithms assume consistency between image values, and variation due to imaging artefacts will violate this assumption and undermine the credibility of analysis results. Thus, before EO data can be analysed for particular applications, the recorded observations need to be calibrated to some acceptable standard.

Calibration can be either:

- absolute—relates the image data range to a physically-measured parameter; or
- relative—rescales the radiance measurements of one or more images to match those of a reference image.

The primary objective in EO image calibration is to ensure a spatial and temporal consistency, or 'stationarity', in which a specific target with a particular reflectance will be the same at all locations and in all images within a given dataset. Image calibration is essential when multiple images are being compared digitally (see Volume 2D). Obviously, the degree to which this needs to be accomplished depends on the specific application. Ultimately, the value of spending time and effort on image correction depends on the value of the final product, as judged by the client, since the total cost of any image product needs to reflect its value to the end user.

As discussed in Volume 1B—Section 9, from an economic perspective the analysis of EO imagery involves a range of costs, including project planning and management, image and ancillary data acquisition, data processing and integration, and product presentation. When a project aims to deliver a low-cost product, the benefits of correcting for effects such as image brightness variations or atmospheric correction may not be justified (Song *et al*., 2001). Conversely, where product accuracy is crucial, the effort required to correct for image distortions would be appropriate. Given the extensive range of approaches and methods that can be used to model or correct for image distortions, the 'best' method will vary for each specific application, its intended use and the 'value' of the processing results to the user.

The importance of EO image calibration is increasingly recognised and significant effort has been expended in recent years to standardise procedures and establish reference sites for this purpose. In Australia, the TERN AusCover Remote Sensing Data Facility has compiled *Good Practice Guidelines* for calibration and validation activities of EO data products (Held *et al.*, 2015). Such guidelines are essential for the generation of biophysical and geophysical products from EO datasets, especially when the underlying data for a given product is derived from a variety of sensors.

3.2 Optical Image Distortions

The following sub-sections consider image distortions in terms of:

- spectral (see Section 3.2.1);
- spatial (see Section 3.2.2);
- radiometric (see Section 3.2.3); and
- temporal (see Section 3.2.4) factors.

While calibration processing is traditionally discussed in terms of radiometric and geometric corrections, the following structure has been adopted to emphasis both the sources of image distortions that are relevant to EO and the need to select appropriate calibration algorithms. In practice, many processing systems and data suppliers provide packaged solutions for image calibration; however, an awareness of the inherent assumptions underlying the convenience of such options is essential for maximising the value and accuracy of EO products.

3.2.1 Spectral

Spectral resolution, density and extent of EO datasets are introduced in Volume 1B—Section 1.1. In the context of spectral characteristics, calibration relates to standardising the reflectance observed by a remote sensor to account for radiance variations resulting from changes in atmospheric conditions and/or illumination and viewing interactions. Note that the term spectral calibration is used to ensure that an optical sensor measures reflectance within a defined range of wavelengths, a process that needs to be applied to the sensor before data acquisition.

The basic geometry used to describe illumination and viewing positions in EO is illustrated in Figure 3.2 (see also Volume 1B—Section 3.5). Ideally, the Earth's surface would be scanned with the illumination and viewing positions directly overhead each image pixel, with no interference from the atmosphere or surface topography. The calibration process attempts to convert EO data to this ideal, or at least reduce the pixel-to-pixel variation that results from atmospheric and illumination/viewing geometry factors. Such calibration is particularly important when multiple image acquisitions are being analysed.

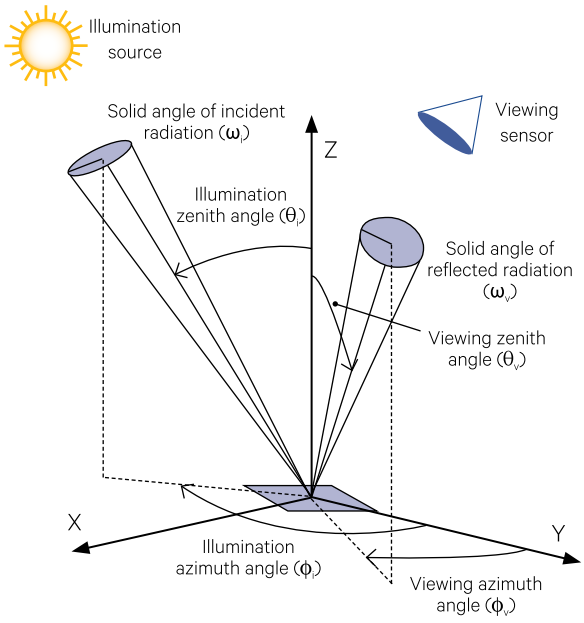
Optical, shortwave and thermal image data from airborne and satellite platforms have varying degrees of angle-dependent brightness variation which change with Sun position, atmospheric conditions, land surface type and sensor characteristics (see Volume 1B). If left uncorrected, these broad level variations can make it difficult to use these data with standard image processing and interpretation methods.

The impact of these brightness variations is most obvious in scanners with wide scan angles (for example, airborne scanners can have scan angles (field of view: FOV) up to ±90°). In imagery acquired from a nadir-viewing position, such brightness variations derive from the:

- atmospheric component phase functions (see Appendix 1); and
- angle-dependent reflectance from the surface of Earth (or the Bi-Directional Reflectance Distribution Function: BRDF; see Appendices 1 and 2 and Volume 1B—Section 3).

Figure 3.2 Geometry of illumination and viewing positions

Basic geometry of illumination source (i) and viewing sensor (v) relative to a target surface



Brightness variations can also occur in imagery acquired by off-nadir viewing angles. In this case, the sensor may view the sides of ground features, such as the woody structures of trees or shadows beneath the canopy. This effect is emphasised for foliage with vertically-oriented leaves (such as *Eucalyptus* trees).

The atmospheric contribution to scan line brightness variations results from two factors:

- varying atmospheric path lengths at different parts of the scan; and
- variations in the atmospheric scattering phase variation (the composite phase function of the atmospheric constituents), which describes the way radiation is scattered out of a volume relative to the direction it enters.

In optical remote sensing, air molecules and aerosols account for much of the scattering, and the effects are greatest at shorter wavelengths. For example, if the aerosol atmospheric turbidity is high, the atmospheric brightening in the blue-green wavelengths will be visible and significant in most remotely sensed data. In thermal image bands, the angular variation is predominantly due to the changing path length; thus the major atmospheric effect is due to water vapour along the path.

These atmospheric and surface effects are present in all satellite data, as well as in aerial photography and scanner data, depending on the variation in the phase angle between the Sun, target and sensor among the pixels in the resulting image. The impact of topography on EO values depends on scale (or pixel size), leading to quite complex image brightness variations (see Appendix 1).

The land surface itself contributes to the broad level brightness variations common to all EO images. The base 'colour' and brightness is a function of the spectral properties of the material types making up the scene (e.g. reflectance and transmittance from leaves, or soil grain mineralogy) plus structural effects which give rise to angular variations (see Volume 1B). The angular effects can be ascribed to three main factors:

- volume effect—as the relative illumination and view angles vary within an image, changes in radiation path lengths and extinction on complex surfaces result in a Volume BRDF that has some similarity to the variations induced by the atmospheric phase function. This effect depends on surface structure and in vegetation it varies with leaf density, leaf angular distributions and total leaf area;
- occlusion (or hotspot) effect—is a more specific effect induced by the fact that the shadows cast by the Sun represent parts of the surface that are occluded (not visible to the Sun). Some of these occluded areas are not visible to the sensor either and these can be considered as those shadows that would be cast by the Sun if it could be located in both the viewing and illumination positions. As the Sun and sensor positions move, different surface areas become occluded to one or both positions, resulting in image brightness variations. For example, when the viewing and illumination positions are aligned, no shadows are visible and the imaged scene appears bright; conversely, when the viewing and illumination angles are 'opposed' many shadows are observed, resulting in lower image brightness; and
- specular (or glint) effect—is most pronounced on water surfaces. This refers to the surface 'mirror' (or Fresnel) reflectance in which the radiation is usually unaltered by the surface material from which it is reflected. On land surfaces it is a composite of reflections diffused from facets of varying angular positions and specularity. In Australia, *Eucalyptus* leaves are especially specularly reflective due to their waxy coating. On water, glint is probably the major component of the BRDF and the hotspot effect does not occur.

Atmospheric and surface effects interact to produce a composite effect that varies with Sun, target and sensor geometry. The effects are more obvious in aerial photography and video images where the central perspective means that at high Sun angles the hotspot and specular points are within the angular radius of the frame. For airborne scanners and satellite borne scanners it is possible to avoid the brightest hotspot and glint points by careful consideration of the scan geometry. However, since the variations are significant for all angles, the effects cannot be ignored in any image.

The scene and atmospheric brightness variations are all minimised if the sensor scans at right angles to the principal plane (the plane formed by the vectors joining the Earth's centre to the sensor and to the Sun). For an airborne scanner, this means flying in a level attitude (neutral pitch), with the Sun directly in front of, or behind, the flight line. If the hotspot or specular points need to be avoided at high Sun times, the scanning sensor may be tilted forward or backward. However, these 'avoiding' strategies are not necessarily easy to implement with acquisition of airborne imagery.

EO imagery is commonly pre-processed to correct for atmospheric effects, and may also be corrected for BRDF (Schaaf *et al.*, 2002) and topographic effects (Richter *et al.*, 2009). The impact of these corrections is illustrated in Section 3.3.2 below. Appropriate processing techniques can also correct for missing and misaligned data in an image (see Volume 2C). When significant portions of an image are reconstructed by image processing, the accuracy of the 'infilled' data should be verified independently. Specific procedures used in image calibration to account for brightness variations are considered in Section 3.3.2 and Appendix 1.

3.2.2 Spatial

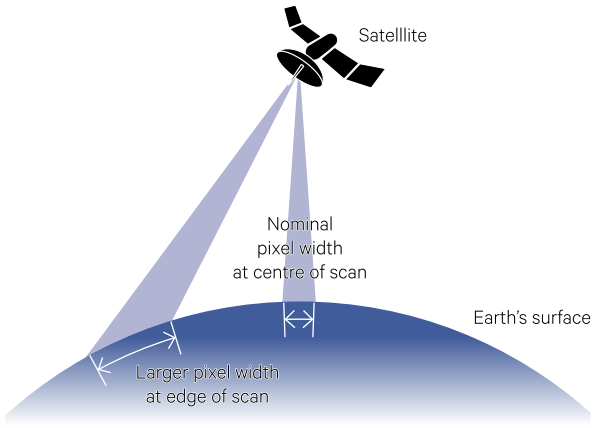
Factors determining the spatial resolution, extent and density of EO imagery are detailed in Volume 1B—Section 1.2. Geometric distortions are inherent to EO imagery due to the image acquisition process (see Volume 2B for details). These distortions broadly result from the composite impact of:

- variations in platform height, direction, speed and stability (most relevant to airborne imagery);
- variations in the viewing position of the sensor when acquiring each pixel in the image;
- scanning mechanism of the sensor (including errors in operation); and
- Earth curvature and rotation.

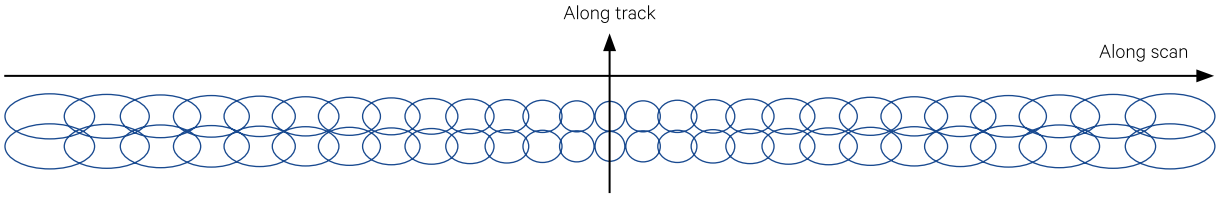
The net effect of these factors is image data with pixels of varying size so that each pixel does not represent the same ground area. This is illustrated in Figure 3.3 for the pixel geometry of AVHRR imagery. In addition, as described in Volume 1B—Section 2, characteristics of the imaging optics can introduce artefacts. Such artefacts mean that not all parts of a ground pixel area are imaged uniformly and a given image pixel may include contributions from neighbouring pixel areas.

Figure 3.3 AVHRR pixel geometry and correction of spatial distortions

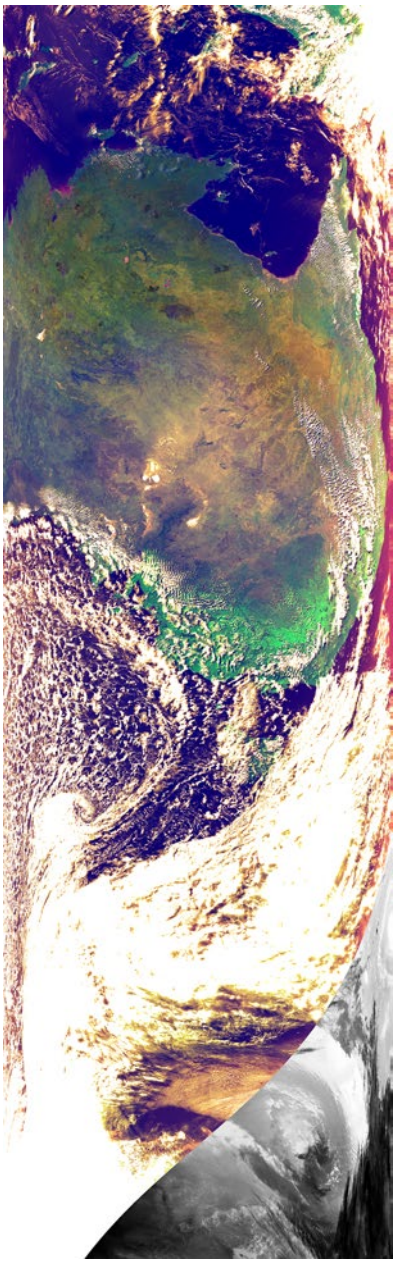
a. For wide scan angles, as occurs with AVHRR, pixel size increases away from the centre of an image swath. This effect is further exaggerated by Earth curvature, especially for sensors imaging wide swaths such as AVHRR.



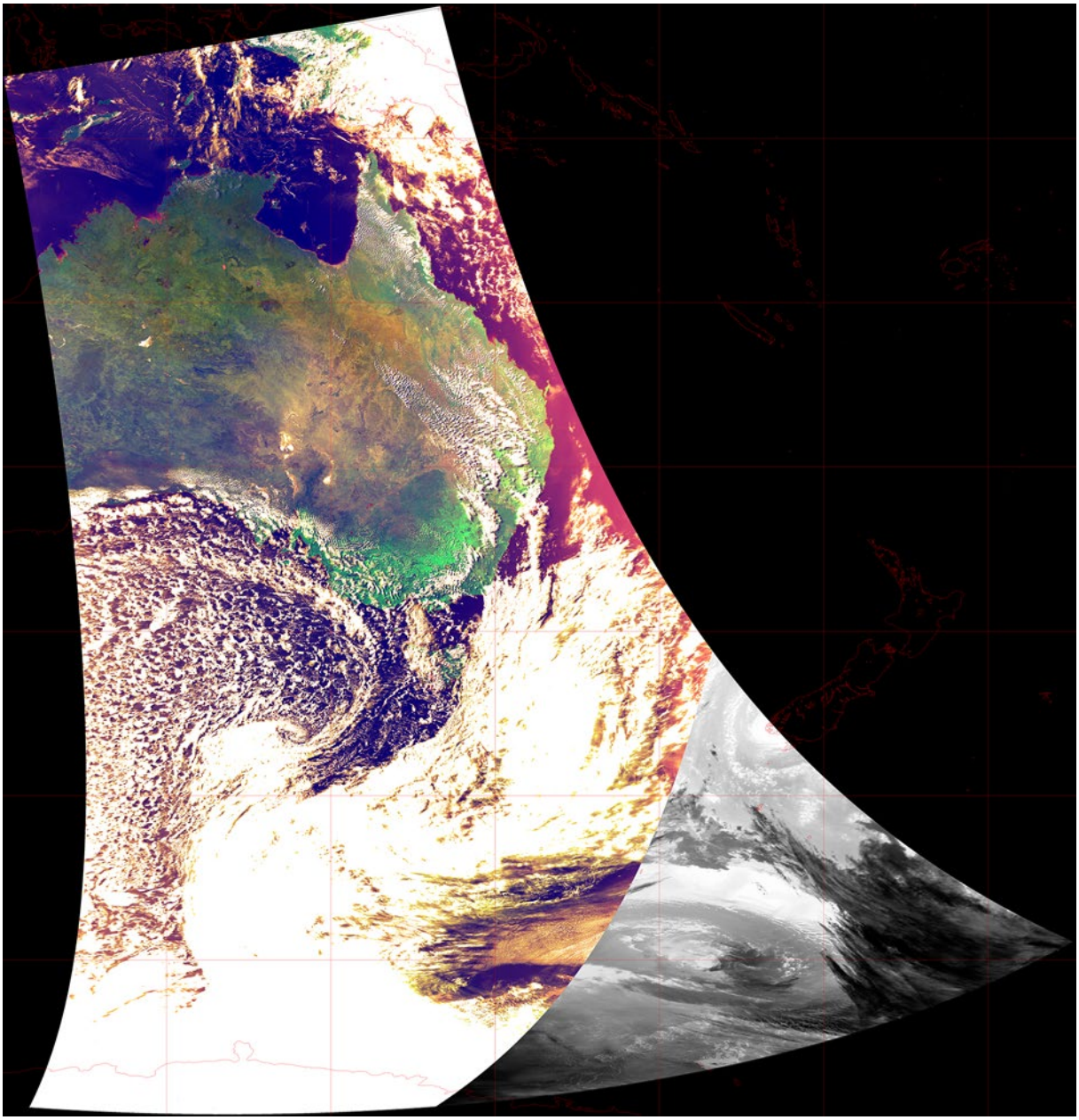
b. When looking vertically down to Earth at the centre of each scan line (or nadir), the sensor detects radiance from a circular ground area. As the sensor looks away from the image centre (off-nadir viewing), a larger ground area is imaged in both the along-scan and along-track directions. A greater increase in imaged area occurs in the along-scan direction, resulting in the optical pixel shape changing from a circle to an ellipse. Near the edges of the image swath, the ground area being recorded as a single pixel is significantly larger than at the centre, resulting in distinct geometric distortions in such imagery (see Volume 2B). Note that there are over 2,000 pixels per line in AVHRR imagery; here the number is significantly reduced for simplicity.



c. This example shows the original image swath acquired by NOAA-19 on 3 September 2017 before and after geometric rectification. For daytime acquisitions, two reflectance channels are used to create a false colour composite. Once the Sun has set, a thermal channel is used, resulting in the greyscale section in the southeast of this image.



d. Rectified Image swath for image shown in Figure 3.3c resampled to 0.01° grid (~1 km pixel size) and aligned with geographic projection.

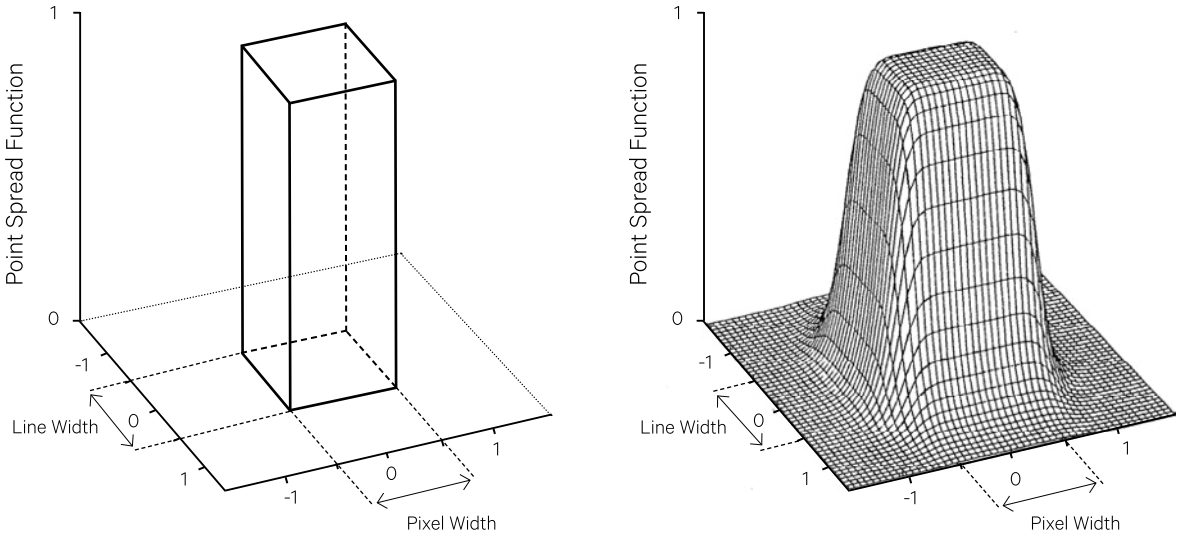


Adapted from: a and b. Breaker (1990), c and d. Sourced from Edward King, CSIRO

Figure 3.4 Example of Point Spread Function

a. Idealised PSF—the radiance measured for a single pixel should form a stepped function (Cracknell, 1998)

b. Modelled PSF for AVHRR pixel—X and Y axes in this figure indicate the central pixel covering the area demarked by elements -0.5 to +0.5 and lines -0.5 to +0.5. Quite clearly, a significant proportion of the radiance recorded for this single AVHRR image pixel originates from the adjacent pixels.



Adapted from: Mannstein and Gesell (1991),

One measure of imaging fidelity is the Point Spread Function (PSF; see Volume 1B—Section 2.1.1.1). The PSF of an 'ideal' pixel is compared with the modelled PSF for an AVHRR pixel in Figure 3.4, showing the extent to which neighbouring pixels can contribute to the recorded radiance values of the central pixel.

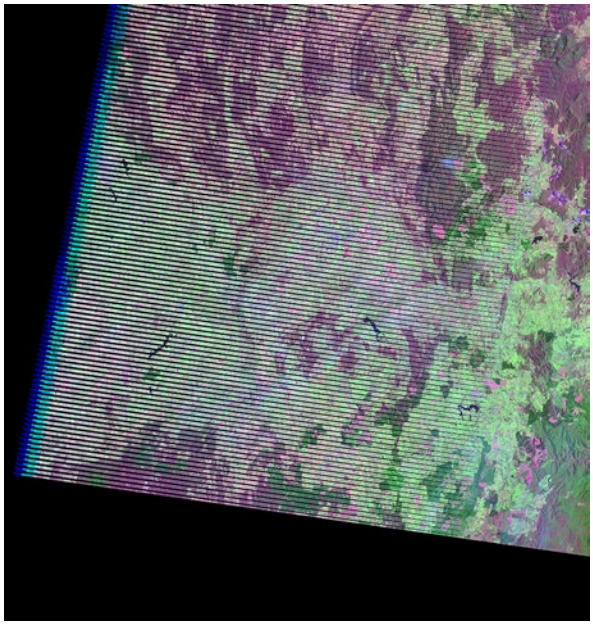
Spatial distortions in images can change their shape, orientation, scale and internal proportions, and do not allow them to be accurately registered with other spatial datasets, including other EO images from the same or different sensors. These distortions are most obvious in images acquired by scanners with wide-scan angles (see Figure 3.3).

Some distortions also result from scanner malfunction, as occurred with the failure of Landsat-7's ETM+ Scan Line Corrector (SLC) in 2003. The SLC corrects the action of the scanner for the motion of the satellite. With the SLC inoperable (termed SLC-Off) the scan lines are not parallel, but form a zig-zag pattern, leaving regular null-data gaps that widen from the centre of the image to the eastern and western edges of the swath (Barsi *et al.*, 2007; USGS, 2017). The resulting missing data is estimated to comprise 22% of the scene area. Two examples of images impacted by this anomaly are shown in Figure 3.5.

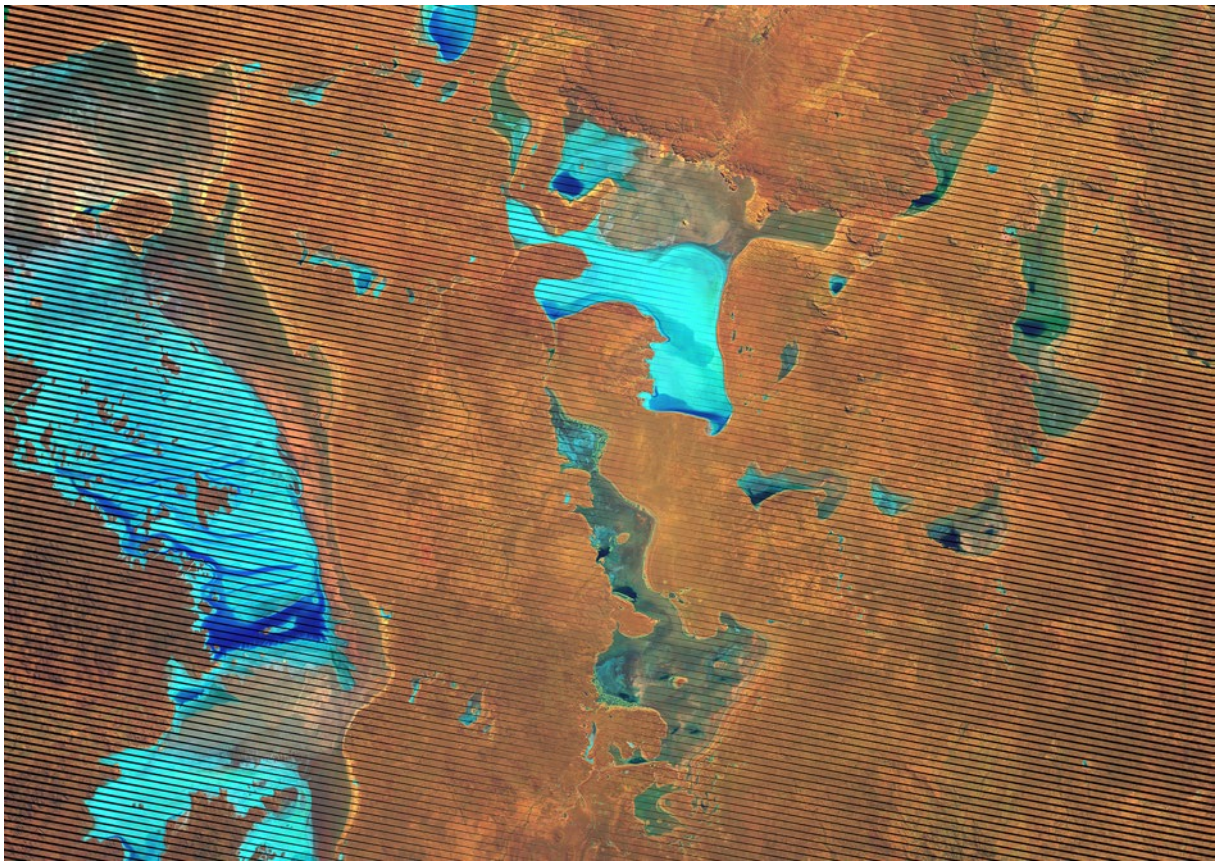
Figure 3.5 Effect of Landsat-7 ETM+ SLC malfunction

These images shows the pattern of missing data that characterised all Landsat-7 ETM+ imagery following the failure of the scan line corrector (SLC) on 31 May 2003.

a. Western portion of Landsat-7 image over Bathurst, NSW, (path 90 row 83) acquired on 2 January 2012.. Bands 5, 4, 2 and RGB.



b. Landsat-7 image over Lake Gairdner and Woomera, SA, acquired on 10 July 2017. (5,4,2 as RGB).



Source: a. Tony Sparks, Icon Water, b. Norman Mueller, Geoscience Australia

The geometry of an EO image can be modified to remove known distortions and register the image to a different (map or image) projection (see Figure 3.3). This is particularly important when the imagery is to be used with GPS or GIS data, or when two or more images are to be compared digitally. Image distortions in airborne data that result from aircraft instability can be avoided using stabilised platforms and inertial navigation systems. Geometric distortions in satellite borne sensors are well-defined and routinely corrected by data suppliers before distribution of data products.

The accuracy of any geometric rectification exercise is critical to the locational accuracy of image features. Whenever multiple images are to be processed as a single dataset, particularly for image time series, image registration needs to be accurate to ensure that detected differences are due to ground changes rather than misregistration (see Volume 2D). Ideally, a spatially calibrated image would represent a regular grid placed on the Earth's surface. While in reality this ideal is impossible to achieve, geometric correction should produce an image in which specific locations can be confidently situated, and pixels in different image acquisitions can be reliably compared. Specific processing procedures for geometric rectification of EO imagery are discussed in Sections 3.3.1 and 7, and Volume 2B.

3.2.3 Radiometric

The concepts of radiometric resolution, density and extent are introduced in Volume 1B—Section 1.3. Radiometric calibration accounts for variations in the measured reflectance due to imperfect operation of sensor components (see Volume 1A—Section 13). These variations may be:

- random—data is not acquired correctly, or at all, for some pixels, creating missing or spurious values for individual pixels, sequences of pixels, or whole lines in the image. Missing values can be interpolated from neighbouring pixels for image continuity, but the interpolated values are merely cosmetic. Spurious values may be more difficult to detect, but various algorithms have been proposed to correct specific datasets (see Sections 3.3.2 and 8.3 below and Volume 2C).
- systematic—miscalibration between multiple detectors in a sensor can result in visible striping artefacts in the acquired image. Examples of this radiometric distortion include Landsat TM 16-line horizontal striping and ALI vertical striping (see Figure 3.6).
- gradual—aging may result in deterioration of individual detectors over time, resulting in radiometric variations in multi-date imagery from the same sensor.

Pixel digital numbers, which are assumed to be proportional to millivolts of radiation, are related to measured radiance by an appropriate gain factor for each imaged channel:

$$L_j = G_j dn_j$$

where

L_j is the radiance data in band j ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$);
G_j is the gain factor for band j; and
dn_j is the digital number.

Thus, to both maintain precision and avoid overflow of saturated pixels, image data is supplied in (SI) units of $W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$ multiplied by the gain factor (see Volume 1A—Section 13.3).

EO sensors are routinely calibrated prior to launch (for satellites), flights (for airborne) or field use (for ground-based sensors). Sensing instruments can, however, change during operation, due to a range of factors, including aging or failure of components. Since the operation of satellite sensors can vary after launch and/or degrade in orbit, their radiometric performance is routinely checked post-launch, via well-established on-board calibration processes, with the aid of both on-board calibrators and ground reference points (Malthus and Li, 2015).

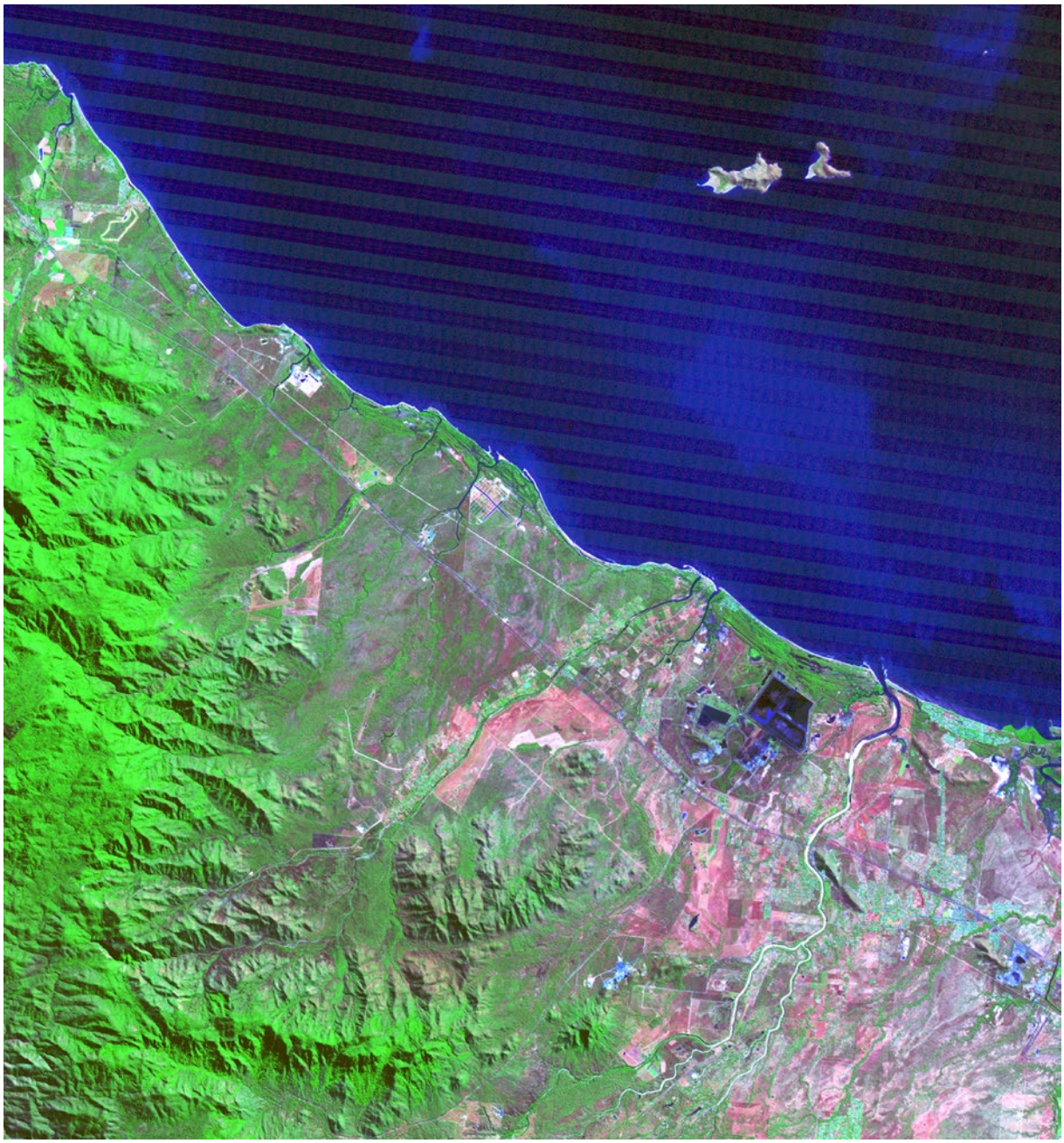
For example, the on-board calibration system for Landsat-7 ETM+ relies on three components to calibrate optical bands to within 5% absolute uncertainty (Markham *et al.*, 2004):

- Internal Calibrator (IC)—two lamps used to check detector efficiency against a dark, reference background after each scan line;
- Partial Aperture Solar Calibrator (PASC)—auxiliary optics that became saturated in most bands, so is no longer used for calibration; and
- Full Aperture Solar Calibrator (FASC)—a deployable diffuser panel that provides absolute calibration against the rising Sun.

This calibration system yields reliable estimates of the uncertainty of the absolute radiometric calibration associated with each image channel (see Figure 3.7). Once in orbit, Landsat-7 ETM+ was briefly cross-calibrated against Landsat-5 TM, since initially both satellites were in 'tandem' orbits and thus acquired imagery over the same targets only 10–30 minutes apart (Teillet *et al.*, 2001). Various research teams also regularly perform vicarious calibration of ETM+ and TM reflective bands (Thome *et al.*, 2004). A detailed description of the more recent Landsat-8 calibration is referenced in Section 3.5.

Figure 3.6 Systematic radiometric miscalibration

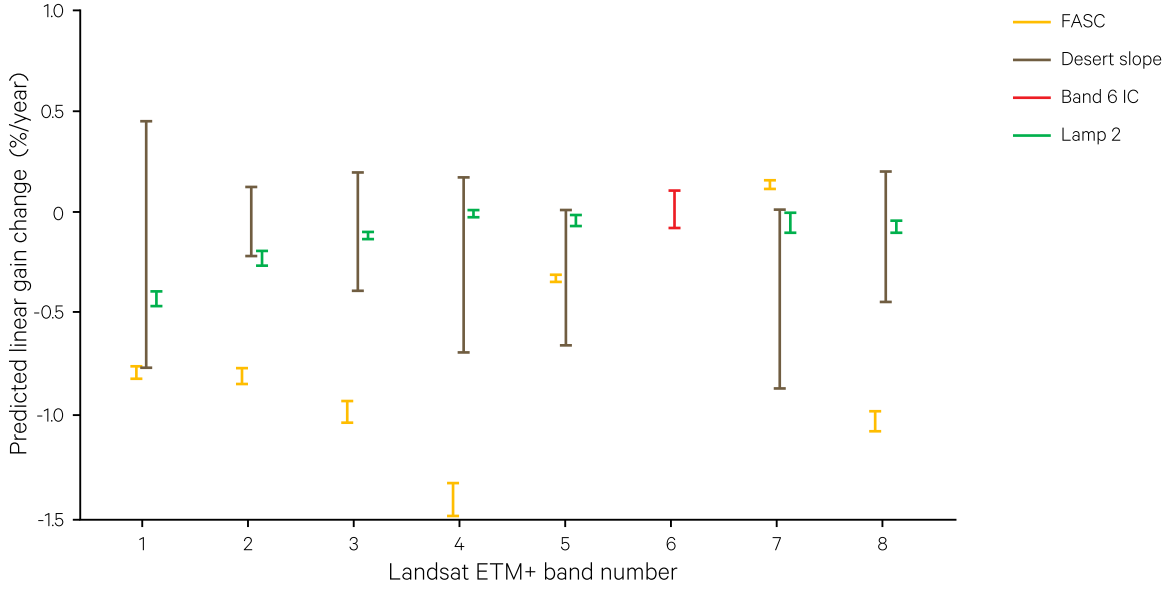
Landsat-5 TM image acquired on 31 July 2003 over northern fringe of Townsville and Paluma Range National Park, Queensland, displayed using bands 5, 4, 2 as RGB. This image shows a striping pattern that was typical of the 16 detector swath in this sensor. In this case, the striping is most visible in the ocean.



Source: Tony Sparks, Icon Water

Figure 3.7 Estimated annual % gain change per year for Landsat-7 ETM+ calibrators

Error bars are the 95% confidence limits in the slope estimate. In the reflective bands, the Full Aperture Solar Calibrator (FASC) gain change is clearly different than the lamp 2 and the desert sites. The lamp 2 change is significant, but could be due to changes in the Internal Calibrator system. The desert sites slopes (straight average) are not significantly different than zero, so indicate there is no change in gain. However, they do not preclude the small changes indicated by lamp 2.



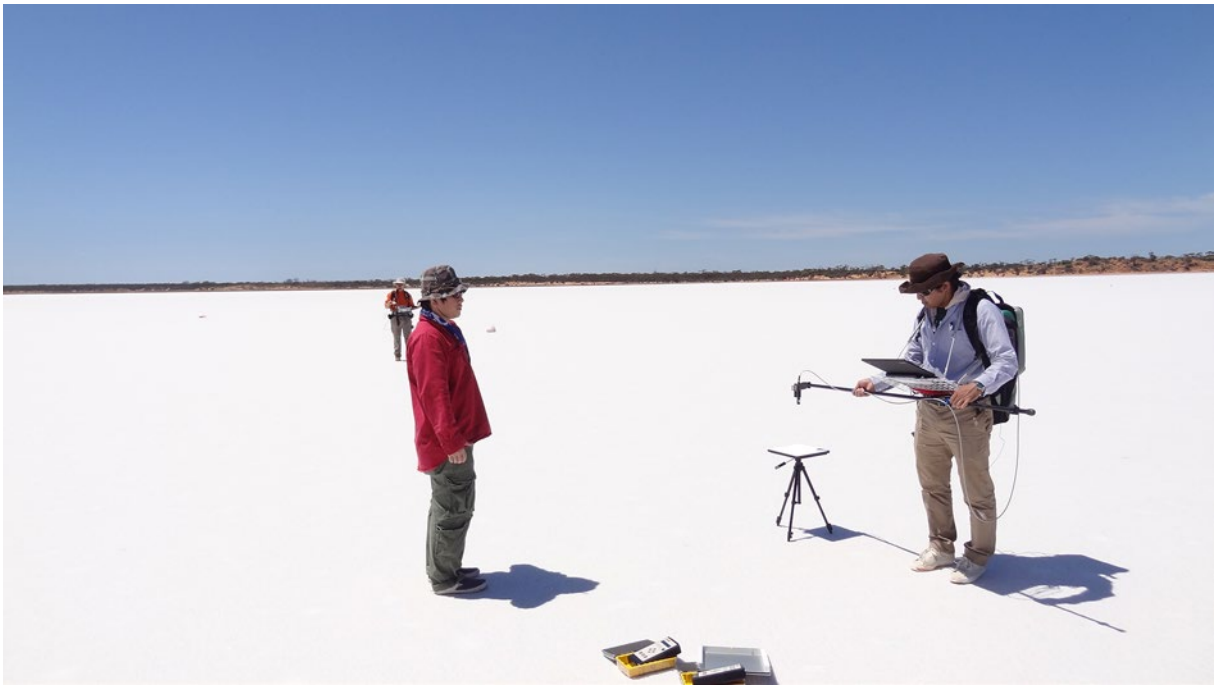
Source: Barsi *et al.* (2007) Figure 11

Most satellite sensor calibration systems have proven to be consistent within individual images and can maintain consistency for several months, but degradation can occur over longer time periods (Thome *et al.*, 1997). Additionally ground-based references (Slater *et al.*, 1987; Slater *et al.*, 1996) and celestial bodies (Kieffer and Wildey, 1985; Kieffer *et al.*, 2003) are used to calibrate satellite sensors over time. Ideal reference points are spatially uniform, spectrally invariant, bright, flat, diffuse surfaces, preferably situated in elevated, inland, arid locations (Teillet *et al.*, 2007). As illustrated in Figure 3.8, dry salt lakes offer suitable reference sites in Australia (Malthus and Li, 2015), and deserts and glaciers provide similar opportunities elsewhere. Some of the approaches used for correcting radiometric distortions are outlined in Section 3.3.2 and Section 8.3.

Figure 3.8 Validation sites for vicarious sensor calibration

a. Lake Frome, SA

b. Lake LeFroy, WA



Source: a. Sue Campbell and Guy Byrne, CSIRO, b. Chris MacLellan, University of Edinburgh

3.2.4 Temporal

To reliably analyse the trends in EO time series imagery, all pixel values must be directly comparable. In addition to correcting the spectral and radiometric distortions mentioned in Sections 3.2.1 and 3.2.3 above, all multi-temporal imagery requires pre-processing to correct for:

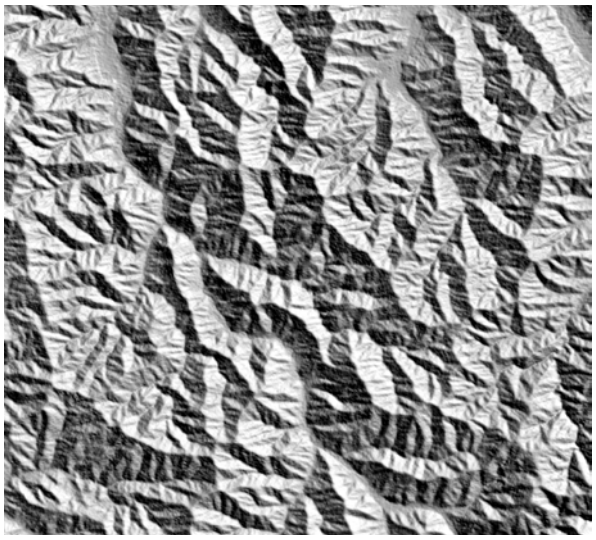
- diurnal and seasonal variations in solar zenith angles, which impact both the overall image brightness and topographic shading (see Figure 3.9); and the
- variations in atmospheric conditions, such as water vapour (as cloud, cloud shadow or snow) or aerosols (see Appendix 1).

Figure 3.9 Seasonal solar illumination variations

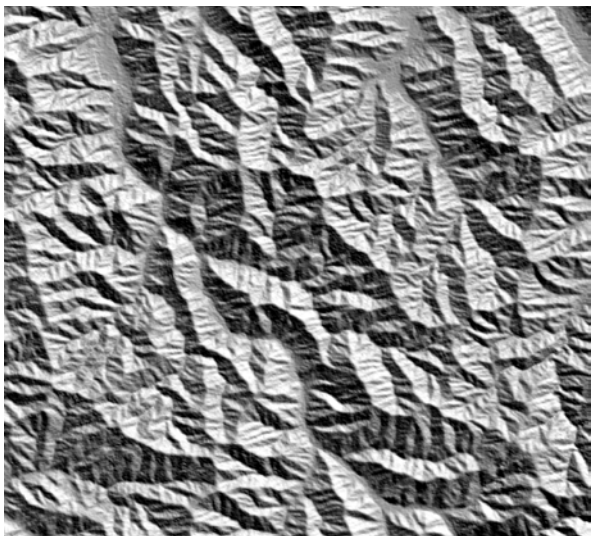
An elevation image was used to simulate the surface shading effects for different seasonal sun positions.

a. Summer (22 December)

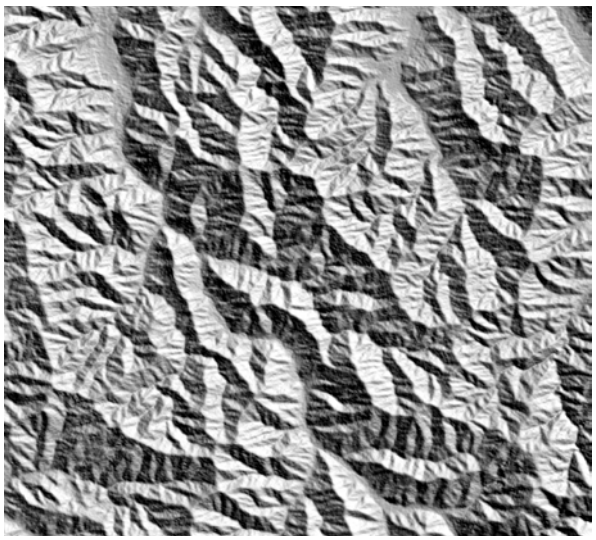
b. Autumn (22 March)



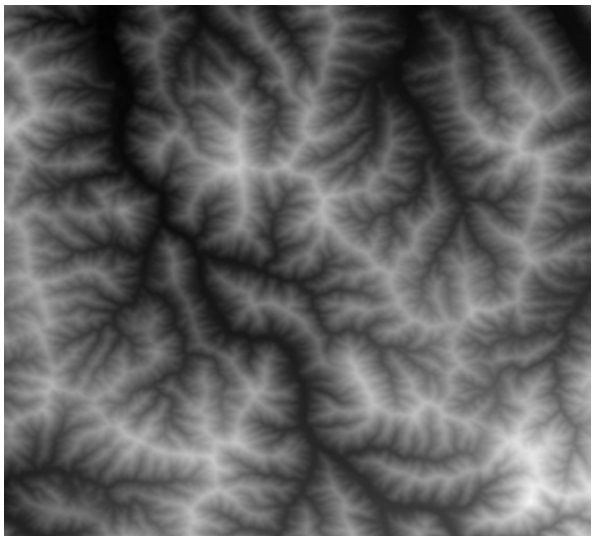
c. Winter (22 June)



d. Spring (22 September)



e. Elevation image displayed with light shades corresponding to high elevation and dark shadows indicating low elevation



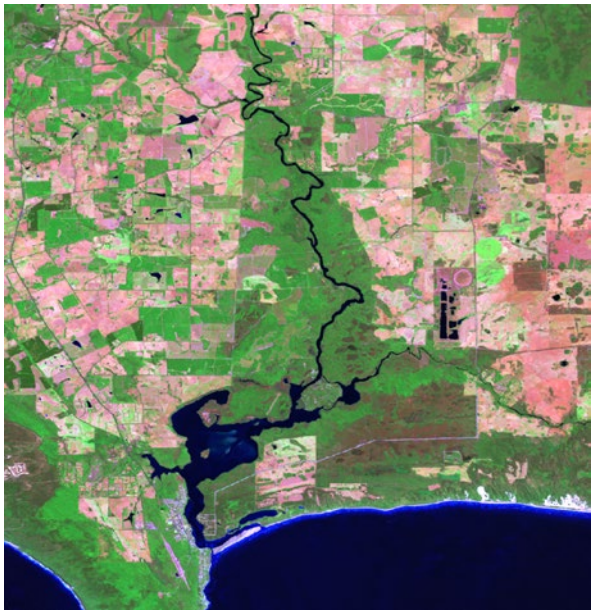
Source: Fuqin Li, Geoscience Australia

Obviously, multi-temporal imagery is likely to show seasonal changes in vegetation cover and inundation, which may or may not be the focus of a particular change detection analysis. Imagery from different sensors is often acquired at different times of the day, which may also reflect diurnal cycles in vegetation, soil moisture and/or water cover. An example of the variations that can occur in imagery acquired in different seasons is shown in Figure 3.10. Specific recommendations for selection of multi-temporal imagery to detect surface cover changes are addressed in Volume 2D.

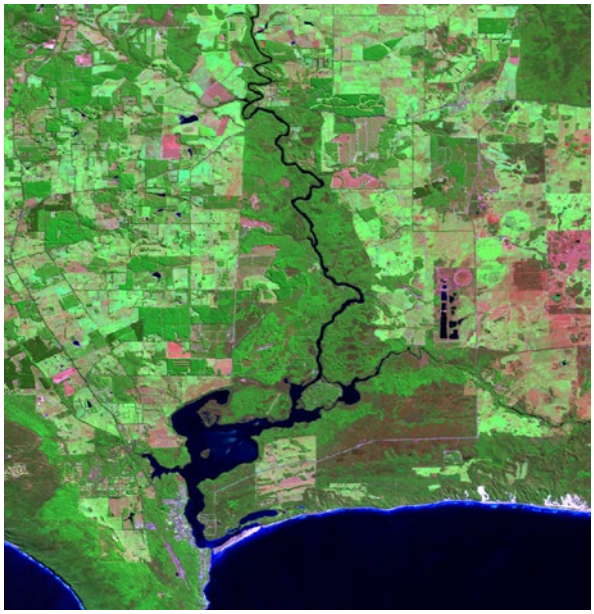
Figure 3.10 Temporal variations in EO imagery

Landsat-8 imagery acquired over Augusta, WA, in four different seasons. Images are displayed as false colour composites, with band 6 as red, band 5 as green, and band 3 as blue.

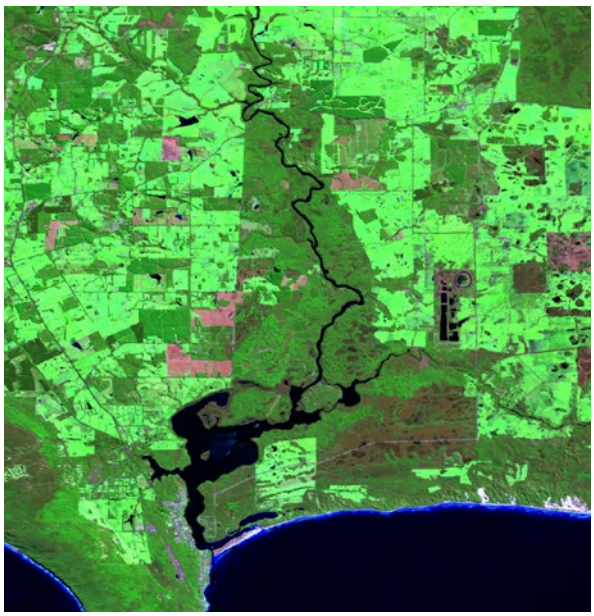
a. Summer—13 January 2015



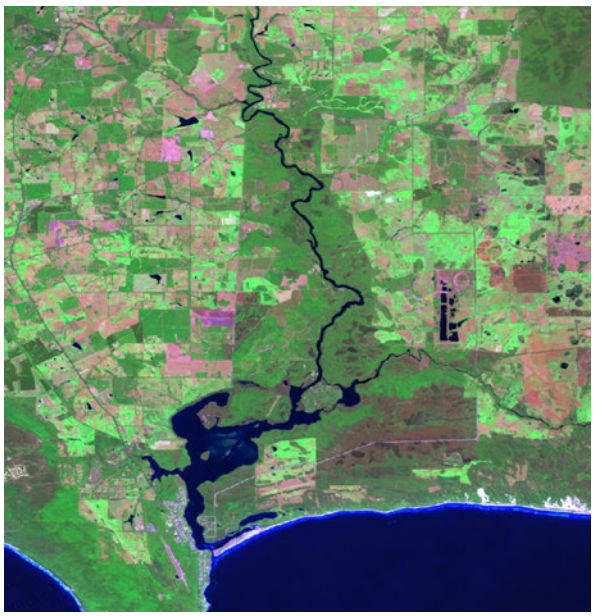
b. Autumn—19 April 2015



c. Winter—25 August 2015



d. Spring—13 November 2015



Source: Norman Mueller, Geoscience Australia

3.3 Optical Image Corrections

Prior to interpretation and analysis, EO imagery is routinely corrected for the various distortions discussed in Section 3.2. Correction methods are grouped into two broad categories:

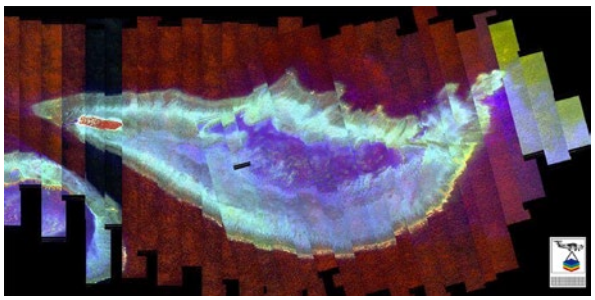
- geometric correction—removing or minimising spatial distortions in the image (see Section 3.3.1); and
- radiometric correction—removing or minimising spectral, radiometric and temporal distortions (see Section 3.3.2).

An example of CASI imagery, acquired from multiple flight lines, that has been corrected for both geometric and radiometric distortions is shown in Figure 3.11. Such corrections are essential in order to digitally compare multi-date images (see Volume 2D).

Figure 3.11 Geometric and radiometric corrections

CASI imagery acquired on 1 June 2002 (western portion) and 3 June 2002 (eastern portion) over Heron Island, Queensland, displayed as natural colour composite (bands 643.7 nm, 522.0 nm, 459.6 nm as RGB).

a. Raw image data—multiple flight lines visible with inconsistent spatial and radiometric features



b. Corrected image—all original images calibrated to surface reflectance , with cross track and deglint corrections applied. (Note: deglinting results in exposed features, such as the island, being obscured.) Image geometry has been rectified to a consistent grid comprising 1 m ground resolution elements.

Source: Karen Joyce, James Cook University

3.3.1 Geometric

Geometric correction accounts for spatial distortions in EO imagery that result from platform and sensor movement and operation, and/or Earth curvature and rotation (see Section 3.2.2 above and Volumes 1A and 2B). Such corrections generally involve three logical processing stages:

- rectification—the process of modelling geometric characteristics of EO imagery so that the image geometry accurately represents the geometric features of the Earth's surface;
- registration—the process of geometrically matching different spatial datasets, such as imagery and/or maps, so that positions in one dataset may be accurately located in others; and
- resampling—modifying the geometry of an image (which may be from either an EO or map data source). This process usually involves rectification and/or registration.

The rectification process involves using well-established models for the sensor and platform, often in conjunction with ground control points (GCP—any features in the study area that can be detected and precisely located on both a map and image), to regularise the image grid. GCP can also be used to obtain external reference information when platform tracking and attitude data are not known precisely.

The process of rectification maps the relationship between locations in the acquired image and a regular, defined (but potentially arbitrary) grid, whereas the process of registration relocates the image data into another defined grid, such as a map projection or another image, possibly with reference to cartographical models. Parameters from both rectification and registration models are used to resample imagery, enabling it to overlay a convenient pre-defined grid.

Image providers generally supply imagery that has been corrected for known geometric distortions and resampled to a defined grid. Such ortho-corrected datasets should state the calibration accuracy and uncertainty, and need to be used with awareness of their pre-processing assumptions and constraints. Multi-temporal studies require all imagery to be resampled to a common grid (see Volume 2D).

An example of the impact that the rectification and resampling processes can have on image geometry is shown in Figure 3.12. The original CASI airborne scanner image over Heron Island Reef displays significant geometric distortion due to platform instability during image acquisition (see Figure

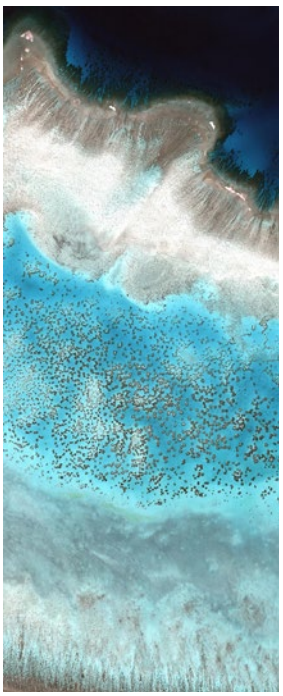
3.12a). Before the image can be usefully compared with other spatial data, these geometric distortions need to be corrected (see Figure 3.12b). Processing methods involved in image rectification, registration and resampling are introduced in Section 7 below and detailed in Volume 2B.

Figure 3.12 Geometric correction of CASI imagery

Imagery acquired by the airborne scanners can exhibit significant geometric distortion due to aircraft movement during image acquisition (see Figure 3.11 for image details). In this case the image also needed to be flipped top to bottom (see Section 7.2.2).

a. Original image

b. Corrected image



Source: Karen Joyce, James Cook University

3.3.2 Radiometric

As detailed in Section 3.2above and Volume 1, the radiance measured by EO optical sensors includes the target reflectance plus a range of effects related to image acquisition, including atmospheric scattering and absorption, surface BRDF, terrain illumination and scanner operation. These acquisition-related effects may be corrected either by simple transformations to reduce their impact or by applying physics-based models with appropriate data (see Section 8.3). For example, the impact of correcting Landsat TM/ETM+ imagery for atmospheric and BRDF effects is illustrated in Figure 3.13.

The logical components of radiometric correction include:

- atmospheric correction—uses vicarious calibration data (see Section 3.1) to correct for atmospheric effects, either by:
  - 'top-down' modelling (atmospheric correction models convert 'top-of-atmosphere' sensor data to reflectance from the ground); or
  - 'bottom up' modelling (radiative transfer models and atmospheric properties convert target reflectances measured on the ground to top-of-atmosphere radiance).
- surface BRDF minimisation—reduces differences due to changes in viewing and illumination positions within and between images, using BRDF parameters or co-incident image acquisitions (Collings *et al.*, 2010; Wu *et al.*, 2010; Collings *et al.*, 2011; Flood *et al.*, 2013). Such methods are detailed in Appendices 1 and 2.
- terrain illumination correction—modifies terrain shading effects using geometrical parameters derived from accurately registered Digital Surface Models (DSM) (as in Li *et al.*, 2012). Topographic correction procedures are further discussed in Appendix 1.
- instrument calibration—uses on-board and/or vicarious reference data to correct for detector differences. Correction of instrument errors may involve destriping algorithms to remove systematic patterns (such as histogram equalisation—see Volume 2C), or specialised algorithms for random distortions (see Section 8.3).

Absolute radiometric correction methods convert pixel digital values to radiance at the sensor, before using atmospheric/BRDF models to estimate surface reflectance(s) (Paolini *et al.*, 2006). An example of multi-scene cross-calibration of 140 ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) images is shown in Figure 3.14. This processing was undertaken as part of a project developing a suite of mineral maps for Australia based on hyperspectral and multispectral EO imagery (Cudahy *et al.*, 2008; see also Excursus 8.3). The seamless mosaic was generated using:

- a physical model that normalises (most of) the intrinsic EO image complications, such as topographic shading, surface directional scattering, solar illumination and residual atmospheric and instrument effects;
- transformation of the ASTER mosaic to (apparent) surface reflectance using least-squares regression between the ASTER data and atmospherically-corrected HyMap airborne imagery; and
- validation of the accuracy of the ASTER cross-calibration based on 60 coincident regions of interest.

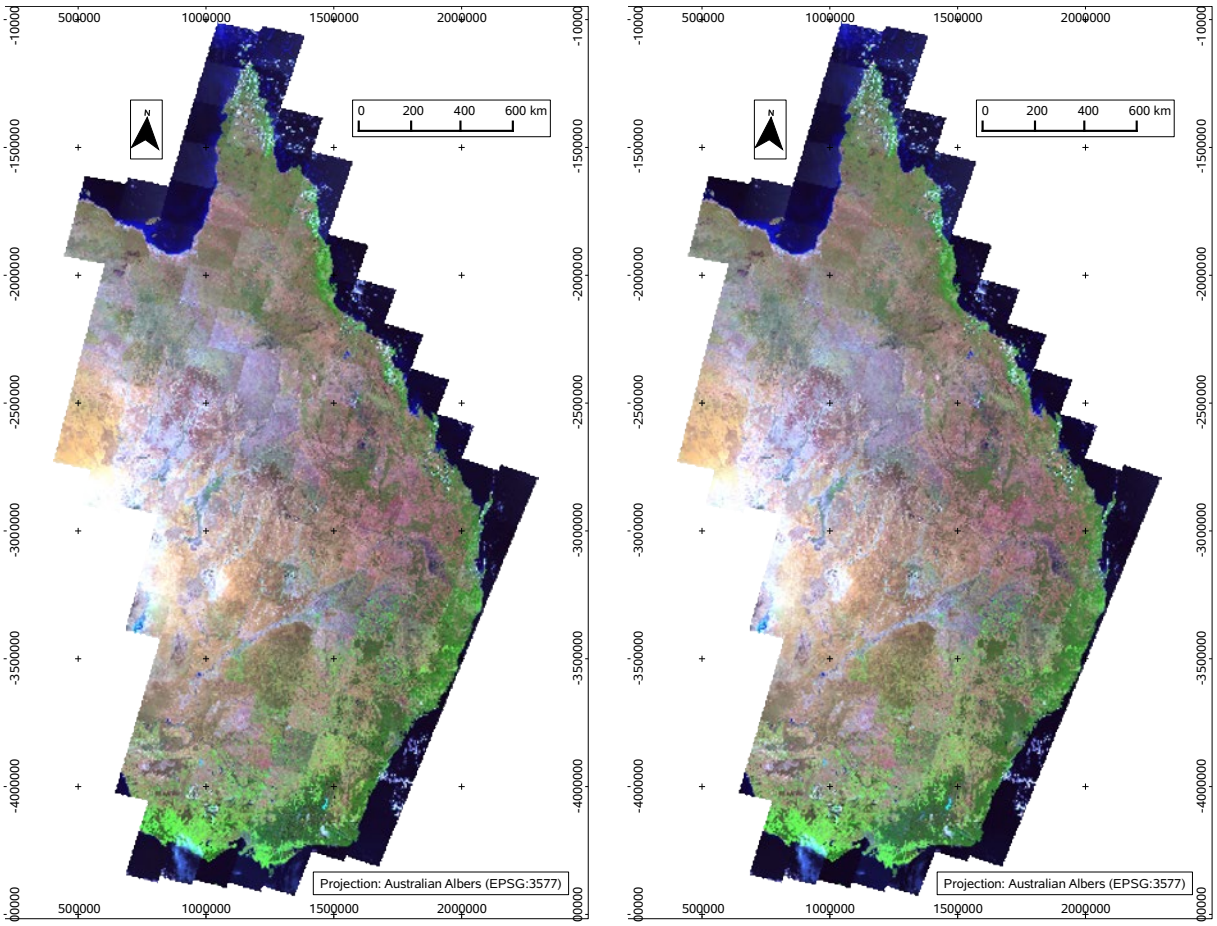
Figure 3.13 Surface reflectance correction

This example of surface reflectance correction applied to Landsat TM imagery for eastern Australia uses a method described by Flood *et al.* (2013). This method uses the 6S Radiative Transfer Model to correct for atmospheric effects and bi-directional reflectance modelling to remove the effects of topography and angular variation in reflectance. Note that some atmospheric effects remain in the corrected imagery due to limitations in input parameters. For example, aerosol modelling assumes a fixed Aerosol Optical Depth (AOD) of 0.05 which would not be valid in smoky or hazy conditions.

This Landsat TM mosaic of eastern Australia was compiled from imagery acquired between August and October 2006. These months experienced prolonged dry weather so vegetation changes were minimal.

a. Standard top-of-atmosphere reflectance

b. Corrected surface reflectance image



Source: Flood *et al.* (2013) using imagery from USGS

Integrated modelling approaches are now being used to generate a range of standard, calibrated image products, such as the ARG25 version of Landsat TM imagery supplied by GA, which is corrected for multiple radiometric distortions (see Section 3.4). Analysis of statistics related to radiometric distortions are further discussed in Section 8.3.

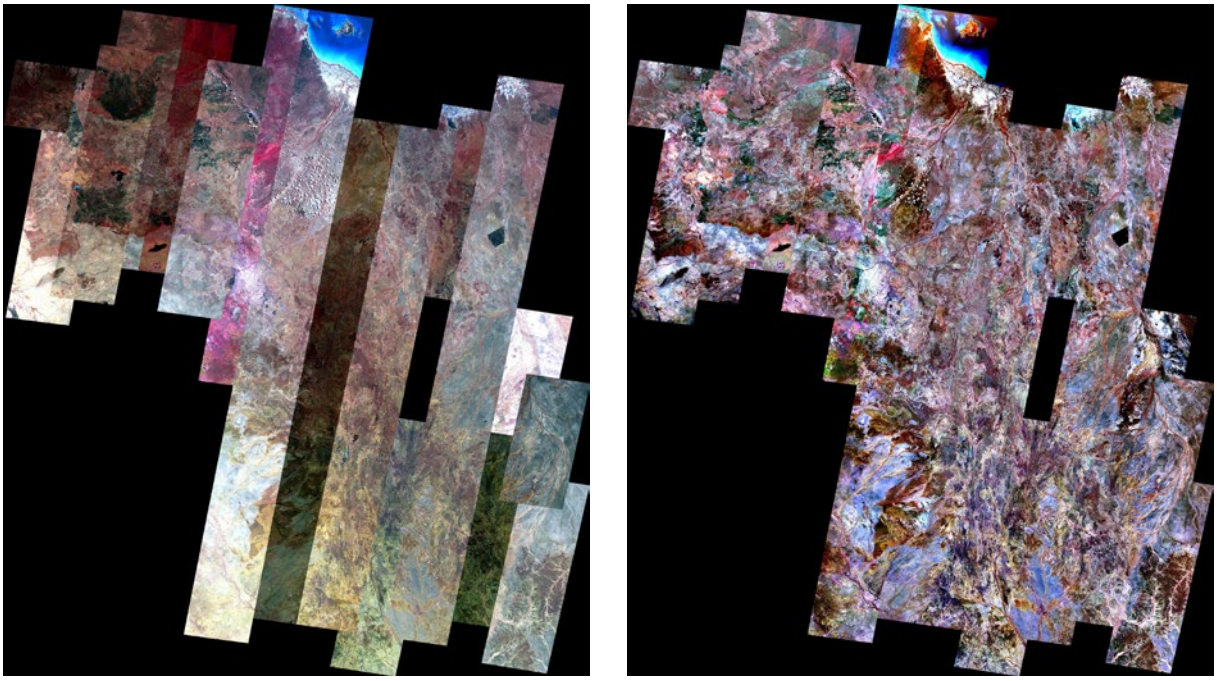
One network of corner reflectors that has been established in Australia to calibrate SAR instruments is described in Excursus 3.1.

Figure 3.14 ASTER image mosaic

This mosaic over the Mount Isa region, Queensland comprises 140 ASTER image scenes acquired over a five year period.

a. Level 1B (radiance-at-sensor) mosaic with SWIR cross-talk correction

b. CSIRO cross-calibrated mosaic after reduction to (apparent) surface reflectance.



Source: Thomas Cudahy, C3DMM

Excursus 3.1—Australian Geophysical Observing Network

Source: Medhavy Thankappan, Geoscience Australia

Calibration of Synthetic Aperture Radar instruments using corner reflectors

The performance of Synthetic Aperture Radar (SAR) instruments is verified by internal and/or external calibration. In this context, calibration is the process of characterising the end-to-end performance of the SAR system to measure the backscattered signal. Internal calibration involves characterisation of the radar system performance using signals from devices built into the sensor system, while external calibration involves the use of ground-based points or distributed targets.

In external calibration, the performance of the SAR instrument is related to a known measurement standard. Point targets, such as radar reflectors or active transponders, can be used for this purpose. A radar reflector is a passive device that reflects incoming radar energy back to the source and can take different forms. The trihedral radar reflector, commonly referred to as a 'corner reflector', which strongly reflects the incident radar energy by virtue of its shape, is structurally robust, more tolerant of alignment errors during deployment compared to other forms, and a popular choice for ground-based calibration of SAR satellites.

The Radar Cross Section (RCS) of an imaged target is a measure of the size of that target as seen by the imaging radar. Distributed targets of known RCS such as agricultural fields, tropical rain forests, or boreal forests, could also be used for calibration. Corner reflectors are routinely used, as they have low maintenance and are available at low cost compared to active devices such as transponders, which also need power for operation. Corner reflectors also exhibit a large RCS relative to their small size, which is maintained over a wide range of incidence angles, ensuring their proper identification in the SAR image.

The Australian Government invested AUD$23 million in 2010, towards building the Australian Geophysical Observing System (AGOS), to enable highly accurate spatial and temporal estimation of large-scale

ground deformation. GA implemented the geospatial component of AGOS, which includes Global Navigation Satellite System (GNSS) instrumentation, a network of high precision GNSS monuments, a robotic GNSS antenna calibration facility, a Synthetic Aperture Radar (SAR) data repository and a network of passive radar corner reflectors in areas specific to AGOS research interests, and designed such that they can be used to monitor crustal deformation and perform routine calibration of SAR instruments on spaceborne platforms.

In order to identify the optimal design of corner reflectors to be deployed as part of the AGOS infrastructure and ensure suitability for calibration applications, GA manufactured 18 corner reflector prototypes with different sizes and material finishes for evaluation. A triangular trihedral design was chosen for the corner reflectors because of the simplicity of manufacture, long term structural rigidity and relative stability for large RCS. Garthwaite *et al.* (2015a) provide a comprehensive report on the design and size choices for the corner reflectors.

The corner reflector prototypes were characterised at the Defence Science and Technology Organisation (DSTO) ground reflection range in Adelaide, by comparing actual RCS measurements with the expected theoretical values and quantifying the change in RCS at different azimuth and elevation angles (Drake and Hatty, 2013). The setup of the corner reflectors for characterisation at the ground reflection range is illustrated in Figure 3.15 and Figure 3.16. Results from the characterisation of the corner reflectors showed that the RCS performance of the prototypes was comparable to theoretical values (Thankappan *et al.*, 2013).

In order to evaluate actual field performance, the corner reflector prototypes were then temporarily deployed at a test site north of Canberra for a five month period. Performance testing involved data acquisitions using SAR satellites at X- and C-band to verify that the observed RCS of the corner reflectors are comparable to theoretical values for calibration of SAR instruments. SAR image acquisitions over the test site showing the response from the corner reflectors can be seen in Figure 3.17.

Figure 3.15 Configuration of trial corner reflectors

Schematic of the experimental setup for corner reflector RCS measurements at the DSTO ground reflection range. The vertical axis that bisects the reflector corner apex also bisects the centre of the turntable used to rotate the reflector in azimuth.

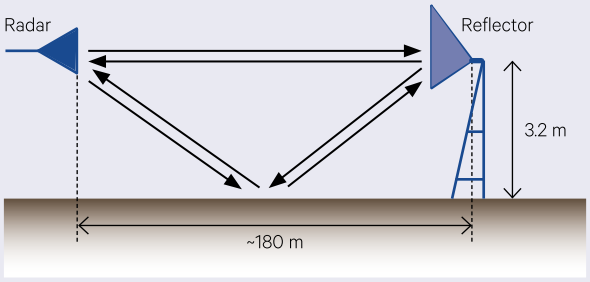


Figure 3.16 Corner reflectors

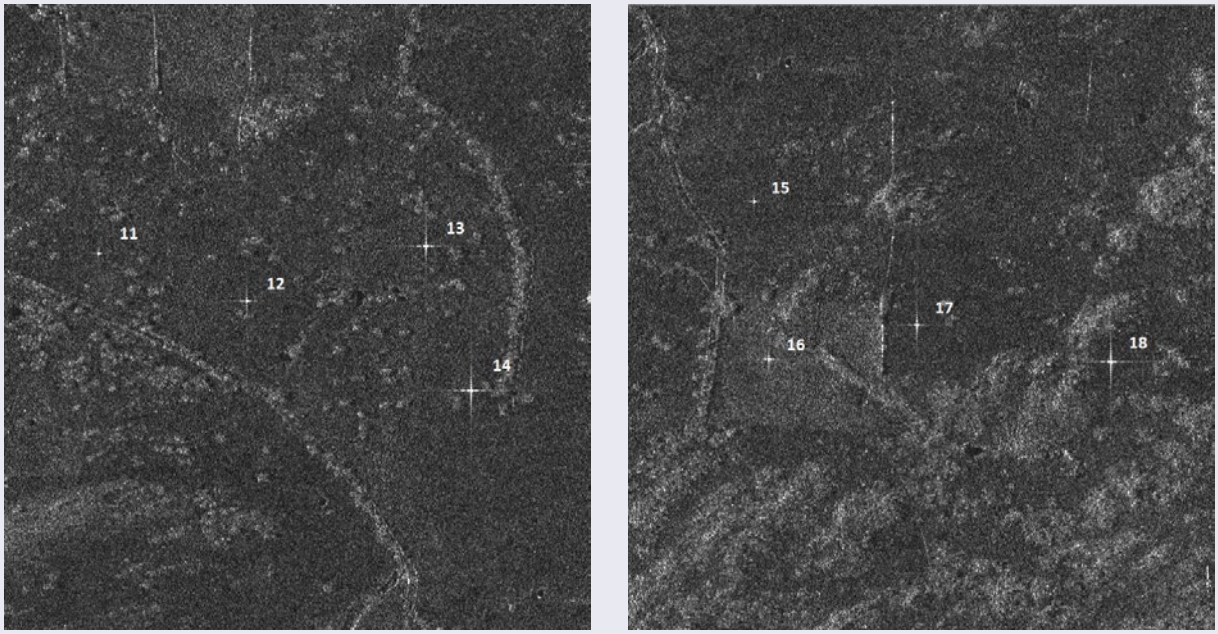
Aperture views of a 1.5 m corner reflector at the DSTO ground reflection range with different alignments:

a. Azimuth cut measurement

b. Elevation cut measurement

Figure 3.17 CR response in SAR image

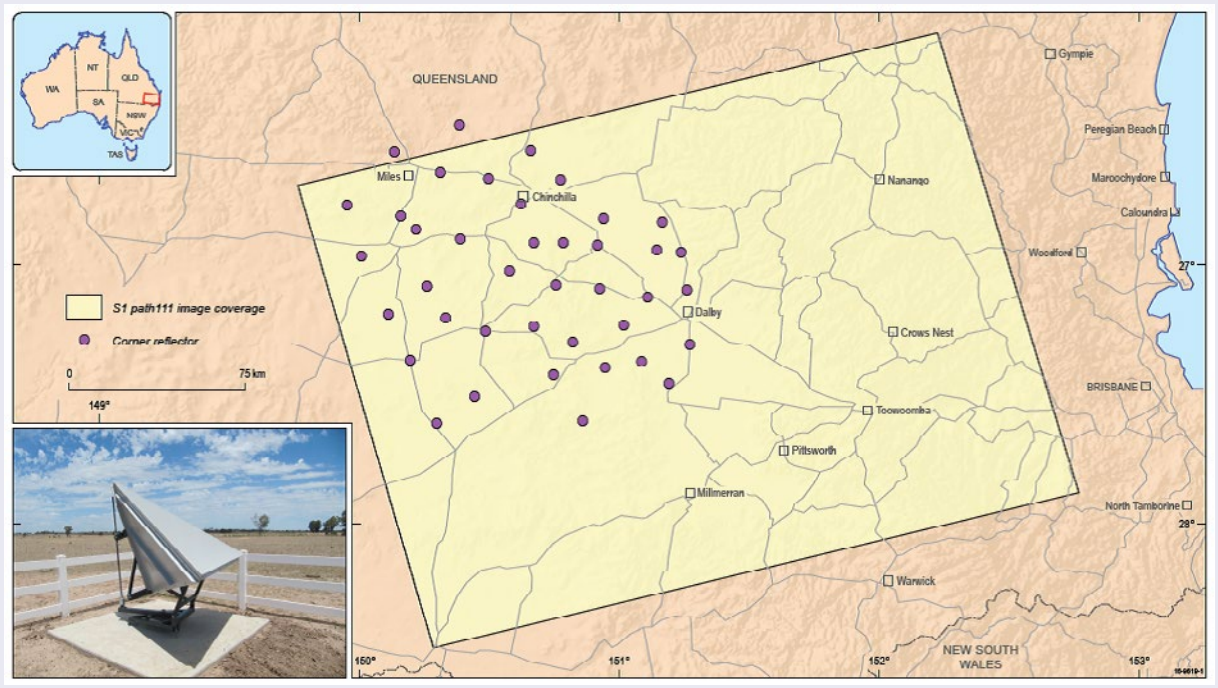
Impulse response of some of the corner reflector prototypes in TerraSAR-X images over Gunning test site on 29 December 2013, labelled by corner reflector number.



Following the performance evaluation, a total of 40 corner reflectors (CR) were deployed permanently in Queensland, Australia, as shown in Figure 3.18. The array includes 34 CR of 1.5 m, 3 CR of 2.0 m and 3 CR of 2.5 m dimension. Through the design process and the precision manufacturing techniques employed, the CR of different sizes are highly suitable for calibration and validation of spaceborne SAR instruments operating in X-, C- and L-band frequencies.

Figure 3.18 CR network

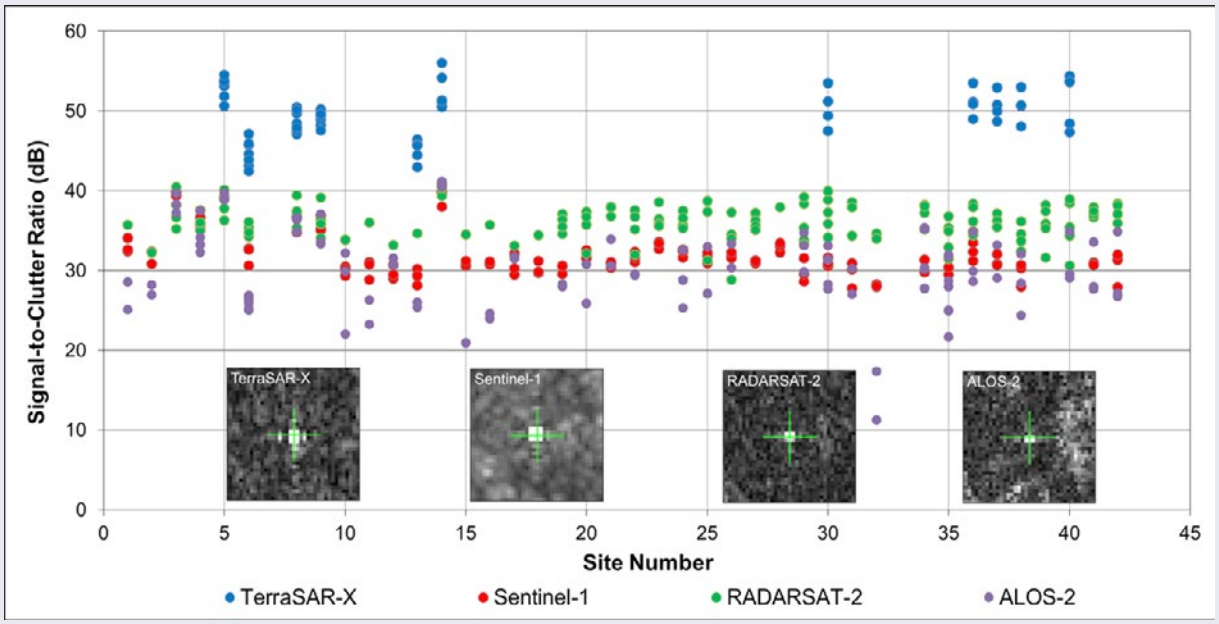
Map of the Surat Basin in Queensland, Australia showing the distribution of 40 corner reflector sites. Inset photo shows a typical 1.5 m corner reflector deployed in the network, which is designed for adjustment in azimuth and elevation.



A commonly used measure of target visibility in a SAR image is the Signal-to-Clutter Ratio (SCR). The point target response from the Queensland corner reflector array has been assessed using SCR for SAR datasets from TerraSAR-X, RADARSAT-2, Sentinel-1A and ALOS-2 satellites (see Figure 3.19). Generally, the SCR for calibration targets should be at least 20 dB (preferably ~30 dB) in order to minimise errors in computation of the calibration factor but also the signal should not be saturated. SCR for each corner reflector was calculated using the oversampled impulse response image and was found to exceed 20 dB irrespective of the frequency bands studied.

Figure 3.19 Target visibility measure

Signal-to-Clutter Ratio (SCR) for SAR datasets from TerraSAR-X, Sentinel-1A, RADARSAT-2 and ALOS-2 satellites.



The AGOS corner reflector array is the largest array of corner reflectors (40) spread over a 100 km by 100 km area that is capable of supporting calibration of SAR satellites operating in multiple frequencies including X-, C- and L-bands. It is currently supporting the European Space Agency's Fiducial Reference Measurement for SAR (FRM4SAR) initiative and being used routinely for calibration of the Sentinel-1A and Sentinel-1B SAR satellite constellation. Other international satellite operators have expressed interest in using this infrastructure for ongoing calibration of their SAR missions.

The permanently deployed corner reflector infrastructure presents an opportunity for independent calibration and comparison of SAR instruments on current and future satellite missions, and is considered an important Australian contribution to global satellite calibration/validation efforts (Garthwaite *et al.*, 2015b).

3.4 Analysis Ready Data

Traditionally, geometric and atmospheric corrections of EO imagery have been variously shared by data suppliers and users. Images were corrected individually using the best available methods, with the same image being corrected multiple times by multiple users. While this approach has enabled decades of advances in EO applications, it is inadequate for ensuring access to the expanding archive of EO data.

The volume of global EO data is clearly growing rapidly in terms of data sources, frequency and archives. Increased awareness of this trend has necessitated significant changes in the established data supply chains. Several suppliers have embraced the concept of 'Analysis Ready Data', which are calibrated and organised in a readily accessible format for immediate analysis. Such data are intended to be:

- geo-located—all pixels aligned geometrically;
- cross-calibrated—all pixels comparable through time and across sensors; and
- ordered—all pixels sequenced in time for easy retrieval.

The availability of 'Analysis Ready Data' reduces inefficiencies in data handling, calibration and processing and ensures a consistent base between users. It is anticipated that integrated data archives will extend analysis and integration opportunities for EO imagery and allow its full value to be realised. For example, CEOS Analysis Ready Data for Land (CARD4L) aims to deliver 'satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets' (CEOS, 2017).

One implementation of analysis ready data is the Digital Earth Australia (DEA; formerly Australian Geoscience Data Cube, AGDC). This initiative organises data into stacks of consistent, time-stamped geographic 'tiles', which can be rapidly manipulated in a high performance computing environment. Pixel flags record data quality decisions and a relational database tracks all tiles, pixels and metadata for the data cube. Data cube technology offers a streamlined approach to the varied and voluminous flow of data from the global EO satellite fleet and is expected to become the standard delivery vehicle for EO data and information products. Further information on the DEA is provided in Section 7.4.2.

3.5 Further Information

Digital Earth Australia (DEA; formerly Australian Geoscience Data Cube, AGDC): *http://www.ga.gov.au/about/projects/geographic/digital-earth-australia*

TERN AusCover Good Practice Guidelines: Held *et al.* (2015)

Jensen (2016) Chapter 6

USGS Landsat Missions: *https://landsat.usgs.gov*

Landsat-8 Calibration: *https://landsat.usgs.gov/landsat-8-l8-data-users-handbook-section-3*

CEOS Analysis Ready Data for Land (CARD4L): (CEOS, 2017)

A number of operational, commercial and research systems are already being used extensively, such as the Google Earth Engine (Gorelick *et al.*, 2017): *https://earthengine.google.com/*

3.6 References

Barsi, J. A., Markham, B. L., Helder, D. L., and Chander, G. (2007). *Radiometric calibration status of Landsat-7 and Landsat-5*. Paper presented at the SPIE 6744, Sensors, Systems, and Next-Generation Satellites XI, Florence, Italy. *http://dx.doi.org/10.1117/12.738221*.

Breaker, L. C. (1990). Estimating and removing sensor-induced correlation from advanced very high resolution radiometer satellite data. *Journal of Geophysical Research: Oceans*, 95(C6), pp. 9701-9711. doi:*http://dx.doi.org/10.1029/JC095iC06p09701*.

CEOS (2017). *CEOS Analysis Ready Data for Land (CARD4L) Description Document*. Retrieved from: *https://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiR-_OCosLXAhWLnJQKHRpUCOsQFggoMAA&url=http%3A%2F%2Fceos.org%2Fdocument_management%2FMeetings%2FPlenary%2F30%2FDocuments%2F5.5_CEOS-CARD4L-Description_v.22.docx&usg=AOvVaw2r8F89fp--1nxlALHTQsAi*.

Collings, S., Caccetta, P., Campbell, N., and Wu, X. (2010). Techniques for BRDF Correction of Hyperspectral Mosaics. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10), pp. 3733-3746. doi:*http://dx.doi.org/10.1109/TGRS.2010.2048574*.

Collings, S., Caccetta, P., Campbell, N., and Wu, X. (2011). Empirical Models for Radiometric Calibration of Digital Aerial Frame Mosaics. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7), pp. 2573-2588. doi:*http://dx.doi.org/10.1109/TGRS.2011.2108301*.

Cracknell, A. P. (1998). Synergy in remote sensing - what's in a pixel? *International Journal of Remote Sensing*, 19(11), pp. 2025-2047. doi:*http://dx.doi.org/10.1080/014311698214848*.

Cudahy, T., Jones, M., Thomas, M., Laukamp, C., Caccetta, M., Hewson, R., Rodger, A., and Verrali, M. (2008). *Next Generation Mineral Mapping: Queensland Airborne HyMap and Satellite ASTER Surveys 2006–2008*. CSIRO Exploration and Mining Report P2007/34.

Drake, P., and Hatty, S. (2013). *Radar Signature Characterisation Measurements of Triangular Trihedral Calibrators Rep*. DSTO-CR-2013-0212. Defence Science and Technology Organisation, Salisbury, Australia. 132 pp.

Flood, N., Danaher, T., Gill, T., and Gillingham, S. (2013). An Operational Scheme for Deriving Standardised Surface Reflectance from Landsat TM/ETM+ and SPOT HRG Imagery for Eastern Australia. *Remote Sensing*, 5(1), pp. 83. doi:*http://dx.doi.org/10.3390/rs5010083*.

Garthwaite, M. C., Nancarrow, S., Hislop, A., Thankappan, M., Dawson, J. H., and Lawrie, S. (2015a). *Design of radar corner reflectors for the Australian geophysical observing system: a single design suitable for InSAR deformation monitoring and SAR calibration at multiple microwave frequency bands. Record 2015/03*. Geoscience Australia, Canberra. *http://dx.doi.org/10.11636/Record.2015.003*

Garthwaite, M. C., Hazelwood, M., Nancarrow, S., Hislop, A., and Dawson, J. H. (2015b) A regional geodetic network to monitor ground surface response to resource extraction in the northern Surat Basin, Queensland. *Aust. J. Earth Sci*. 62, 4.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. doi:*https://doi.org/10.1016/j.rse.2017.06.031*.

Held, A., Phinn, S. R., Soto-Berelov, M., and Jones, S. (2015). *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products*. TERN AusCover. Retrieved from *http://data.auscover.org.au/xwiki/bin/view/Good+Practice+Handbook/WebHome*.

Jensen, J.R. (2016) *Introductory Digital Image Processing. A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc.

Kieffer, H. H., Stone, T. C., Barnes, R. A., Bender, S., Eplee, R. E. J., Mendenhall, J., and Ong, L. (2003). *On-orbit radiometric calibration over time and between spacecraft using the Moon*. Proceedings of: Sensors, Systems and Next-Generation Satellites VI. SPIE 4.

Kieffer, H. H., and Wildey, R. L. (1985). Absolution calibration of Landsat instruments using the moon. *Photogrammetric Engineering and Remote Sensing*, 51(9), pp. 1391–1393.

Li, F., Jupp, D. L. B., Thankappan, M., Lymburner, L., Mueller, N., Lewis, A., and Held, A. (2012). A physics-based atmospheric and BRDF correction for Landsat data over mountainous terrain. *Remote Sensing of Environment*, 124, pp. 756-770. doi:*http://dx.doi.org/10.1016/j.rse.2012.06.018*.

Malthus, T. J., and Li, F. (2015). Calibration of Optical Satellite and Airborne Sensors. Chapter 4 in *'AusCover Good Practice Guidelines (A technical handbook supporting calibration and validation activities of remotely sensed data products)'* (Eds: A. Held, S. R. Phinn, M. Soto-Berelov, and S. Jones). TERN AusCover, St Lucia, Australia.

Mannstein, H., and Gesell, G. (1991). *Deconvolution of AVHRR data*. Paper presented at the 5th AVHRR Data Users' Meeting, Tromsø, Norway.

Markham, B. L., Thome, K. J., Barsi, J. A., Kaita, E., Helder, D. L., Barker, J. L., and Scaramuzza, P. L. (2004). Landsat-7 ETM+ on-orbit reflective-band radiometric stability and absolute calibration. *IEEE Transactions on Geoscience and Remote Sensing*, 42(12), pp. 2810-2820. doi:*http://dx.doi.org/10.1109/TGRS.2004.836389*.

Paolini, L., Grings, F., Sobrino, J. A., Jiménez Muñoz, J. C., and Karszenbaum, H. (2006). Radiometric correction effects in Landsat multi-date/multi-sensor change detection studies. *International Journal of Remote Sensing*, 27(4), pp. 685-704. doi:*http://dx.doi.org/10.1080/01431160500183057*.

Richter, R., Kellenberger, T., and Kaufmann, H. (2009). Comparison of Topographic Correction Methods. *Remote Sensing*, 1(3), pp. 184. doi:*http://dx.doi.org/10.3390/rs1030184*.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, 83(1), pp. 135-148. doi:*http://dx.doi.org/10.1016/S0034-4257(02)00091-3*.

Slater, P. N., Biggar, S. F., Holm, R. G., Jackson, R. D., Mao, Y., Moran, M. S., Palmer, J. M., and Yuan, B. (1987). Reflectance- and radiance-based methods for the in-flight absolute calibration of multispectral sensors. *Remote Sensing of Environment*, 22(1), pp. 11-37. doi:*http://dx.doi.org/10.1016/0034-4257(87)90026-5*.

Slater, P. N., Biggar, S. F., Thome, K. J., Gellman, D. I., and Spyak, P. R. (1996). Vicarious Radiometric Calibrations of EOS Sensors. *Journal of Atmospheric and Oceanic Technology*, 13(2), pp. 349-359. doi:*http://dx.doi.org/10.1175/1520-0426(1996)013%3C0349:vrcoes%3E2.0.co;2*.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., and Macomber, S. A. (2001). Classification and Change Detection Using Landsat TM Data. Remote Sensing of Environment, 75(2), pp. 230-244. doi:*http://dx.doi.org/10.1016/S0034-4257(00)00169-3*.

Teillet, P. M., Barker, J. L., Markham, B. L., Irish, R. R., Fedosejevs, G., and Storey, J. C. (2001). Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sensing of Environment*, 78(1), pp. 39-54. doi:*http://dx.doi.org/10.1016/S0034-4257(01)00248-6*.

Teillet, P. M., Barsi, J. A., Chander, G., and Thome, K.J. (2007). *Prime candidate Earth targets for the post-launch radiometric calibration of space-based optical imaging instruments*. Paper presented at the SPIE 6677, Earth Observing Systems XII, San Diego, United States. *http://dx.doi.org/10.1117/12.733156*.

Thankappan, M., Garthwaite, M. C., and Williams, M. L., Hislop, A., and Dawson, J.H. (2013). *Characterisation of Corner Reflectors for the Australian Geophysical Observing System to Support SAR Calibration*. Proc. ESA: Living Planet Symposium, Edinburgh, 9–13 September 2013.

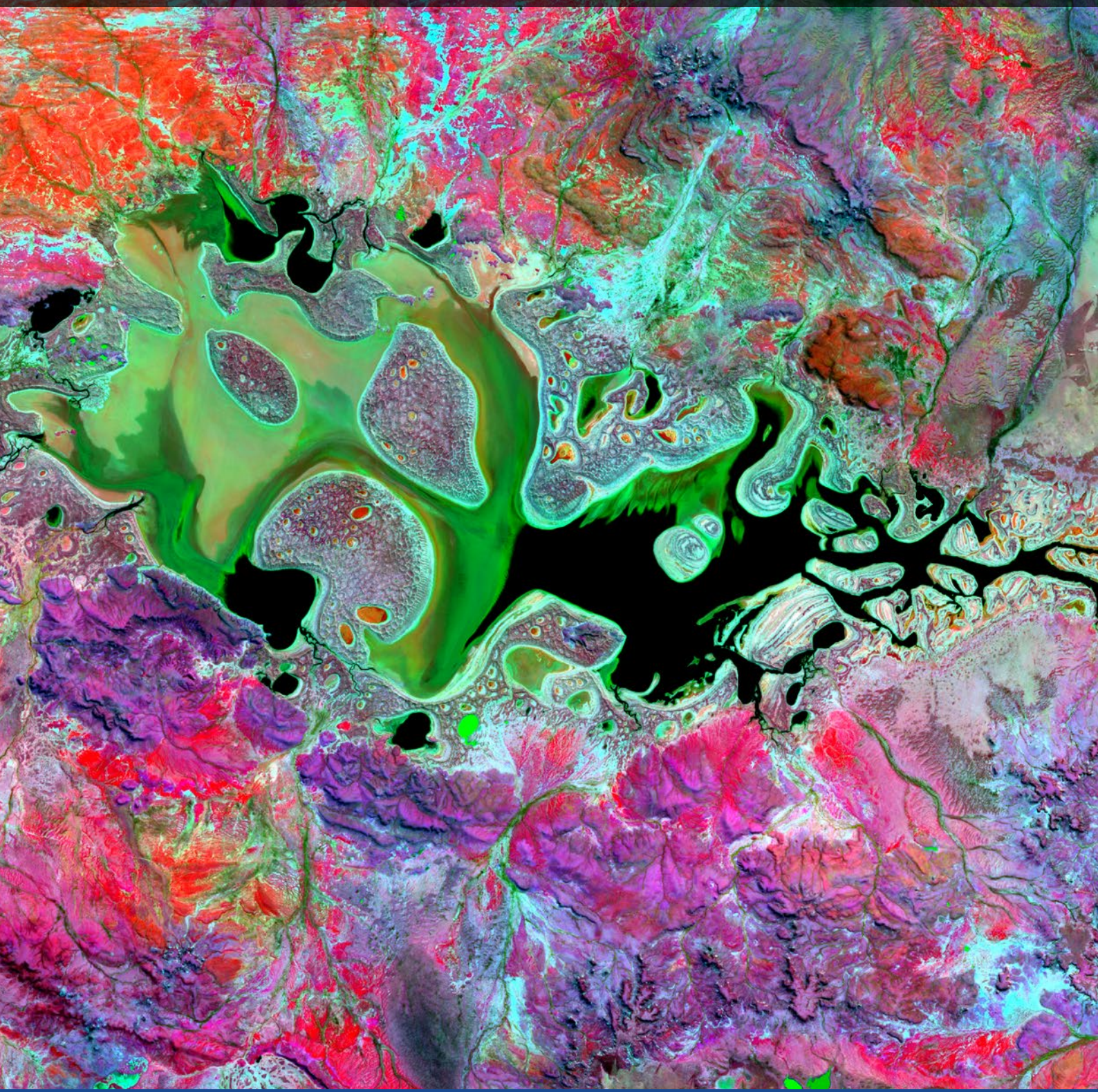
Thome, K., Markham, B., Barker, J., Slater, P. N., and Biggar, S. (1997). Radiometric calibration of Landsat. *Photogrammetric Engineering & Remote Sensing*, 63(7), pp. 853-858.

Thome, K. J., Helder, D. L., Aaron, D., and Dewald, J. D. (2004). Landsat-5 TM and Landsat-7 ETM+ absolute radiometric calibration using the reflectance-based method. *IEEE Transactions on Geoscience and Remote Sensing*, 42(12), pp. 2777-2785. doi:*http://dx.doi.org/10.1109/TGRS.2004.839085*.

USGS (2017). *SLC-off Products: Background*. Retrieved from *https://landsat.usgs.gov/slc-products-background*.

Wu, X., Collings, S., and Caccetta, P. (2010). *BRDF and illumination calibration for very high resolution imaging sensors*. Paper presented at the IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

Interpretation



While image processing is largely concerned with computer manipulation of data values, a fundamental part of analysing imagery involves visual interpretation. The human eye can only differentiate a limited number of discrete colours and an even more restricted set of grey shades. Display representation of the image data therefore needs to maximise the contrast between available display colours to provide optimal differentiation of image features. Various techniques for image enhancement have been developed to highlight features of interest in displayed data.

Interpretation is generally the initial process of visually delineating and identifying features or patterns in an image. The following four sections cover the basic techniques that are relevant to visual image interpretation, namely:

- contrast—enhance contrast between image features and background (see Section 4);
- colour—using colour to highlight specific features or compare multiple channels (see Section 5);
- presentation—methods for displaying image data on monitor devices and producing hardcopy imagery, such as inkjet or laser plots or photographic products (see Section 6).
- geometry—changing image shape, scale, and orientation (see Section 7); and

Contents

4 Contrast 49

5 Colour 67

6 Presentation 89

7 Geometry 95

Background image on previous page: Infrared colour composite over Lake Carnegie, a large ephermal lake in central WA, acquired by Landsat-8 on 8 July 2017. The vibrant colours of this non-standard colour composite derive from bands 10 (TIR 1), 5 (NIR) and 6 (SWIR 1) as RGB respectively (see also Figures 5.16 to 5.19). **Source:** Norman Mueller, Geoscience Australia.

4 Contrast

Contrast relates to the ability to differentiate a feature from its background. In image processing, contrast enhancement can be conducted in a number of ways depending on the type of features being identified in an image. These techniques can improve either the specific contrast of a selected feature or the overall image contrast.

Numeric values recorded in image or raster format may be represented as a picture by associating each image value with an intensity and/or colour on a display or hardcopy device. A simple grey-scale representation of an image channel is illustrated in Figure 4.1.

Figure 4.1 Grey-scale image display

Any spatial data can be represented pictorially by associating each data value with a grey shade or colour.

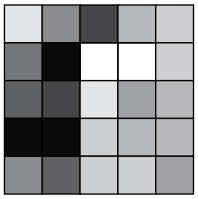
a. Image pixel values

1	5	8	3	2
6	9	0	0	2
7	8	1	4	3
9	9	2	3	3
5	7	2	2	4

b. Grey-scale lookup table

0
1
2
3
4
5
6
7
8
9

c. Grey-scale image



Source: Harrison and Jupp (1990) Figure 3

A finite number of grey shades or colours can be represented using display screens or hardcopy devices. Remote sensors are designed to detect reflectance/emission levels over a wider range than would occur within a single image scene, so the data range for a single remotely sensed image rarely covers the full potential range of image values (such as 0–255 for 8-bit (or 'byte') data sources—see Section 1). To maximise the contrast in a pictorial representation of image data, we need to match the range and distribution of values in the image (or selected features of interest within the image) to the available range of grey shades or colours in the display device.

4.1 Channel statistics

The range of data values in an image channel is usefully summarised by its frequency histogram (see Section 4.1.1) while the distribution of values within the data range is indicated by the variance statistic (see Section 4.1.2). These simple statistics are commonly used by image processing systems to improve image contrast.

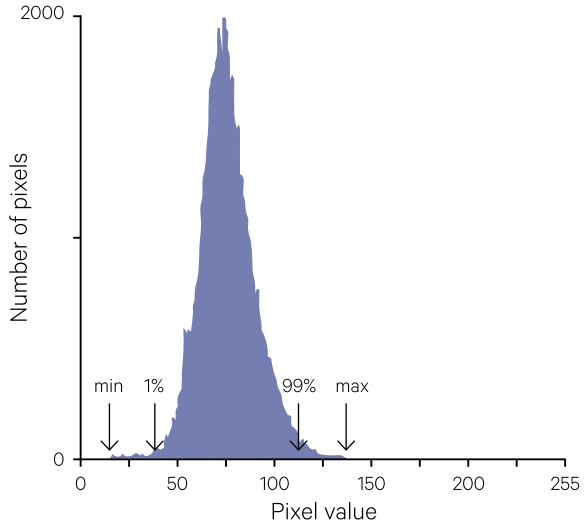
4.1.1 Histograms

The range and distribution of data values in any image channel may be represented as a histogram. The usual form for histograms in image processing is the 'frequency histogram' shown in Figure 4.2, with the horizontal axis representing the image data range and the vertical axis indicating the number of pixels. A frequency histogram is formed by plotting the number of pixels that contain each digital value in a channel against the full range of digital values. In the example in Figure 4.2, the most common image channel value (the 'mode'—maximum count) is 74 for 2,131 pixels. Ancillary statistics such as the absolute minimum, maximum and mode values in each channel are usually reported separately. The total number of pixels used to generate the histogram is also usually reported (61,440 in this example). The frequency histogram and its related statistics are used extensively in image processing to define the effective data range for image display and to transform or identify particular features.

Background image: Panchromatic image of Canberra, ACT, acquired by Landsat-8 on 10 August, 2017. **Source:** Norman Mueller, Geoscience Australia

Figure 4.2 Frequency histogram

A frequency histogram is formed by plotting the number of pixels at each image value in a channel. In this example, for a total of 61,440 pixels, the minimum channel value is 15 and the maximum value is 137. The largest number of pixels, or mode, occurs at image value 74 (2,131 pixels).

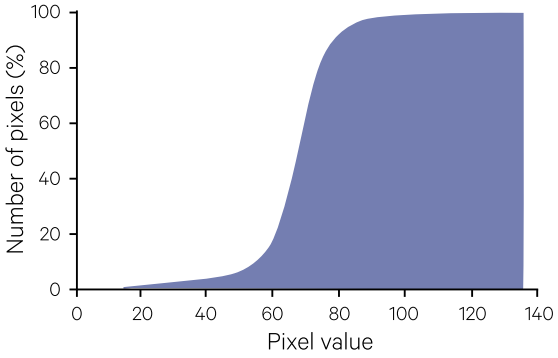


Source: Harrison and Jupp (1990) Figure 4

Percentage points (or percentiles) for the histogram summarise the histogram distribution in terms of cumulative statistics. A cumulative histogram, as illustrated in Figure 4.3, could be derived manually from the full range of percentiles if required.

Figure 4.3 Cumulative histogram

This demonstrates a (stylised) cumulative histogram for the image data used in Figure 4.2. 1% of pixels have values below or equal to 44, 20% have values below or equal to 66, while 95% of pixels have values below or equal to 89.



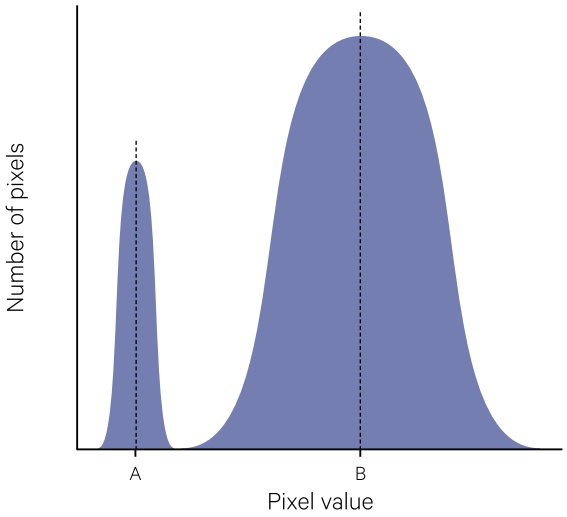
Source: Harrison and Jupp (1990) Figure 5

The shape of the frequency histogram is also significant in interpreting image data. The example histogram (see Figure 4.2) has the mode closer to the minimum than to the maximum value producing a skewed distribution. This is caused by a larger spread of values close to the maximum value, which have relatively small numbers of pixels. This feature is also reflected in the relatively large difference between the values of the 95% and 99% points (89 and 113 respectively—see Figure 4.3). The degree of 'skewness' can be readily assessed in most cases by comparing the median or modal pixel values for a channel with its mean value. Skewed data distributions commonly occur in EO imagery with the mode usually being closer to the minimum value.

Channel histograms can also be discontinuous or multi-modal in EO imagery. Discontinuities may be due to poor calibration techniques or inadequate rescaling operations. Multiple modes in a histogram can occur when an image contains several different features that have discrete reflectance levels in the wavelength range of the channel. For example, water and vegetation features have very different reflectance levels in near infrared wavelengths (approximately 700 nm–1100 nm) so an image containing water and land features could produce a near infrared channel histogram similar to the one illustrated in Figure 4.4 (see also Figure 4.10). When highly contrasting features constitute an image, it is possible for the distribution to become disjoint (that is separated) along the pixel value (horizontal) axis.

Figure 4.4 A bi-modal histogram

A bi-modal histogram occurs when an image channel contains two features that have distinctly different (non-overlapping) radiance values. For example, the near infrared reflectance for water is typically low (such as histogram peak labelled A) while vegetation reflectance is higher (such as histogram peak labelled B; see also Figure 4.10).



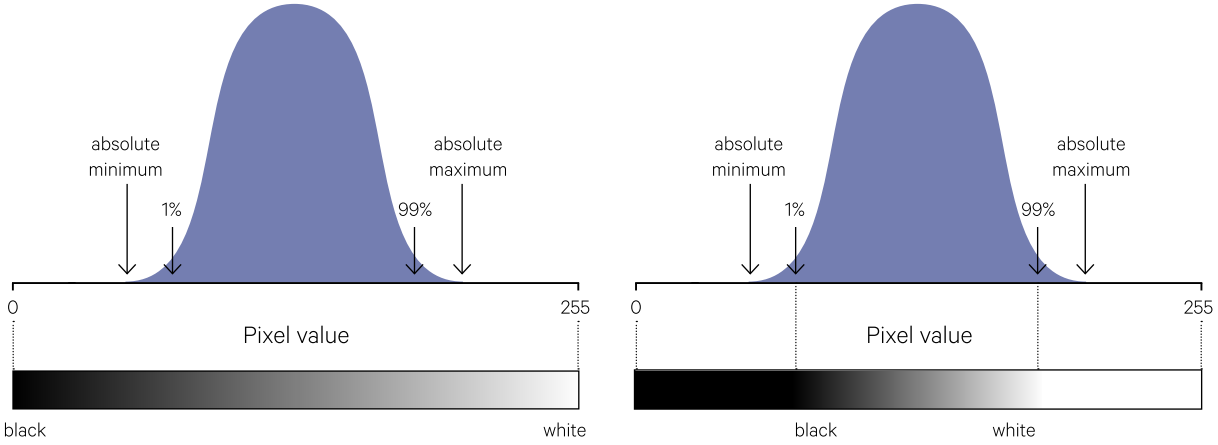
Source: Harrison and Jupp (1990) Figure 6

Various image processing operations modify image representation by altering the image histogram range or shape. When displaying an image, for example, it is desirable to achieve as much contrast as possible. Any display device has a finite number of grey levels between the darkest and brightest colours (see Section 6.1).

Figure 4.5 Histogram stretching for image display

a. 'Unstretched' histogram with the theoretical minimum being mapped to black and the theoretical maximum being mapped to white. This assignment does not consider the actual range of values that occurs in a particular image channel.

b. Stretched histogram display where minimum and maximum values of the image channel are used to define contrast limits (see also Figure 4.8).



Source: Harrison and Jupp (1990) Figure 7

In the case of the 8-bit (byte) format channel data shown in Figure 4.2 and Figure 4.3, if pixels with value 0 are displayed as black and pixels with value 255 are displayed as white, the displayed image would only contain mid-grey shades. To utilise the available contrast range of the display device, usually the minimum or 1% value is displayed as black and the 95% or 99% value as white (see Figure 4.5). This process is called image stretching and is further discussed in Section 4.2.1.

The histogram shape may also be altered to improve image contrast (see Section 4.2). This is useful in some applications to more clearly delineate particular features. Other aspects of generating and interpreting image histograms are considered in Section 8.1.1.

4.1.2 Variance

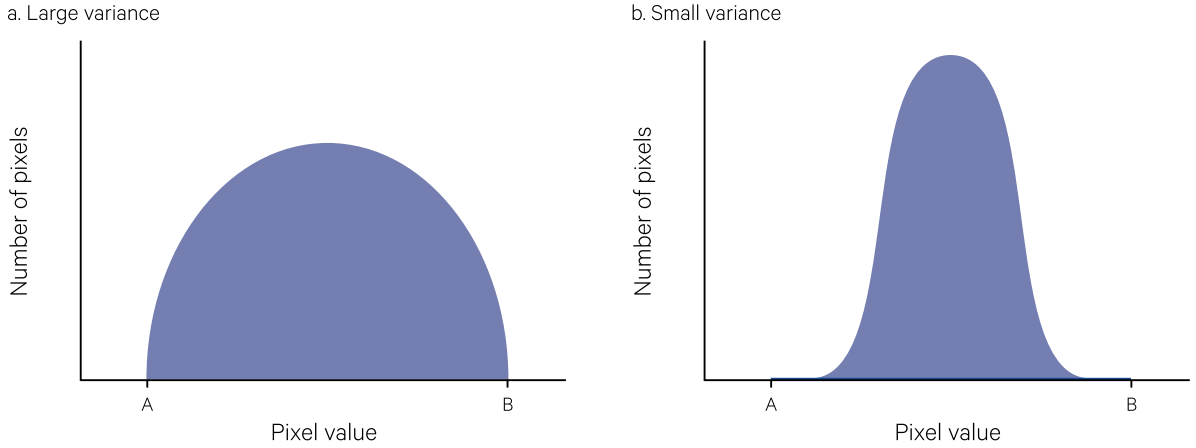
While the range of data values in a channel is an indication of the channel variability, it does not indicate the way the values are distributed between the minimum and maximum values. Figure 4.6 shows two histograms with the same range but different variance values.

Figure 4.6 Histograms with equal ranges and differing variances

Pixel value A indicates the channel minimum value, while value B is the maximum value in the channel.

a. Large variance

b. Small variance



Source: Harrison and Jupp (1990) Figure 8

A statistical measure called variance (σ^2) summarises the differences between all pixel values and the mean value for the channel. Mathematically, channel variance is the average of the squares of the deviations of the pixel values from their mean value and is calculated as:

$$\text{channel variance}\left(\sigma^2\right) = \frac{\sum\left(x-\mu\right)^2}{n}$$

where

x is pixel value in channel;
n is total number of pixels in channel (excluding null pixels); and
μ is true mean of the channel data.

Calculation of these statistics is illustrated in Excursus 4.1 for a simple example image channel. Another statistic used to indicate channel variability is equal to the square root of the variance value, and is called the 'standard deviation' (σ). Use of variance and standard deviation statistics in image analysis are further discussed in Section 8.1.2.

Excursus 4.1—Channel Statistics

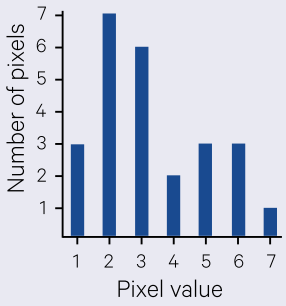
Using the variance equation given above, channel statistics are computed for an example image channel comprising 25 pixels (5 pixels in 5 rows). The example image values and the image frequency histogram are shown in Figure 4.7.

Figure 4.7 Example image

a. Example image values

1	2	3	3	4
3	3	5	6	3
2	2	2	6	5
5	4	2	2	1
7	6	3	2	1

b. Frequency histogram



Source: Harrison and Jupp (1990) Figure 9

Calculation of the variance statistic for this channel is detailed in Table 4.1. As is frequently the case in EO images, the distribution of pixel values about the mode is not symmetric.

Statistical measures of central tendency:

mode (most common value) = 2
median (midpoint of range) = 4
mean (average; μ) = 83/25 = 3.32

Statistical measure of data extent:

range (maximum - minimum) = 7-1 = 6

Statistical measures of dispersion from mean:

variance (σ^2) = 73.33/25 = 2.933

standard deviation $\left(\sigma\right) = \sqrt{2.933} = 1.713$

Table 4.1 Calculation of variance statistics

Pixel value	Frequency count	Pixel value x frequency count	(Mean value – pixel value)^2	(Mean value – pixel value)^2 x frequency count
1	3	3	5.29	15.75
2	7	14	1.69	11.83
3	6	18	0.09	0.54
4	2	8	0.49	0.98
5	3	15	2.89	8.67
6	3	18	7.29	21.87
7	1	7	13.69	13.69
Total	25	83		73.33

4.2 Contrast enhancement

Since any display (or hardcopy) device can only show a finite number of grey shades or colours, in order to optimise contrast in the viewed image we need to match the range of image values to the available range of shades/colours in the device. Various techniques can be used to match these ranges including:

- linear stretching (see Section 4.2.1);
- non-linear stretching (see Section 4.2.2);
- statistical stretching (see Section 4.2.3); and
- lookup tables (LUT; see Section 4.2.4).

4.2.1 Linear stretching

For a single image channel, if pixels equal to the potential minimum image range value are displayed as black and those near the potential maximum value in the image are displayed as white, with the values between these minimum and maximum values being linearly mapped to discrete shades of grey between black and white, the resulting image may not provide adequate contrast between features of interest. If the dynamic range of the data does not cover the potential range, the variance in the displayed image—and hence its overall contrast—may be increased by linear stretching.

A display device using an 8-bit colour space can represent 256 shades of grey in a displayed image, while a 10-bit colour space represents 1024 shades (see Section 6.1). If an image contains no pixels with values equal to the display-minimum value there will be no pixels showing as true black in the displayed image. Similarly, if the display-maximum value is set at 254 for an 8-bit colour space but there are no values above 200 in the image, the 55 'brightest' shades of grey on the display screen will not be used in displaying the image—only mid-grey and dark grey shades would be displayed.

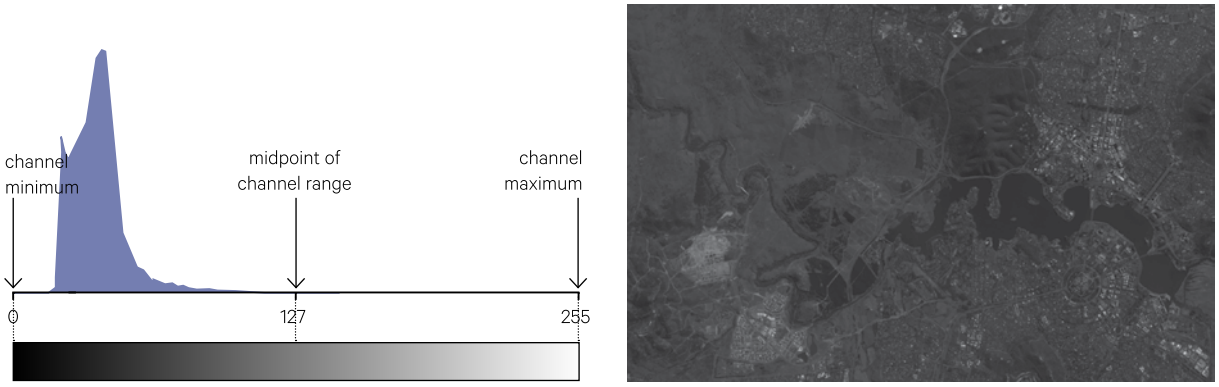
If the actual minimum and maximum values are displayed as black and white respectively, with the intermediate values again being linearly assigned to the grey shades in between, we would be utilising the full dynamic range of the display device but, due to the histogram distribution shape, most pixels would still be displayed as mid-grey shades. Alternatively, display limit values close to the 1% minimum and 99% maximum in the channel (that is, the values which are respectively above and below only 1% of all pixel values—or the central 98% of values) would reduce the range of image values that are displayed by matching this range with the major data variations in the image. This process is illustrated in Figure 4.5. In some cases the 2% and 98% values are used as the display minimum and maximum values (see Figure 4.8c).

This simple method of contrast enhancement is referred to as a linear stretch since the image values are being proportionally stretched to match the available display range. The minimum and maximum values selected for a stretch depend on the nature of the image data. In most cases the 1% and 99% (or 1% and 95%) values render displayed images with good contrast.

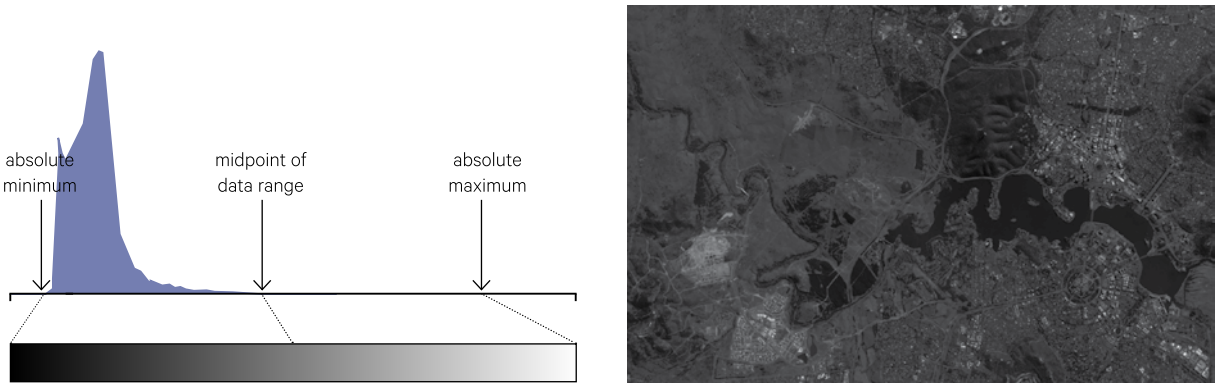
Figure 4.8 Contrast enhancement

This panchromatic image of central Canberra was acquired by Landsat-8 on 10 August 2017.

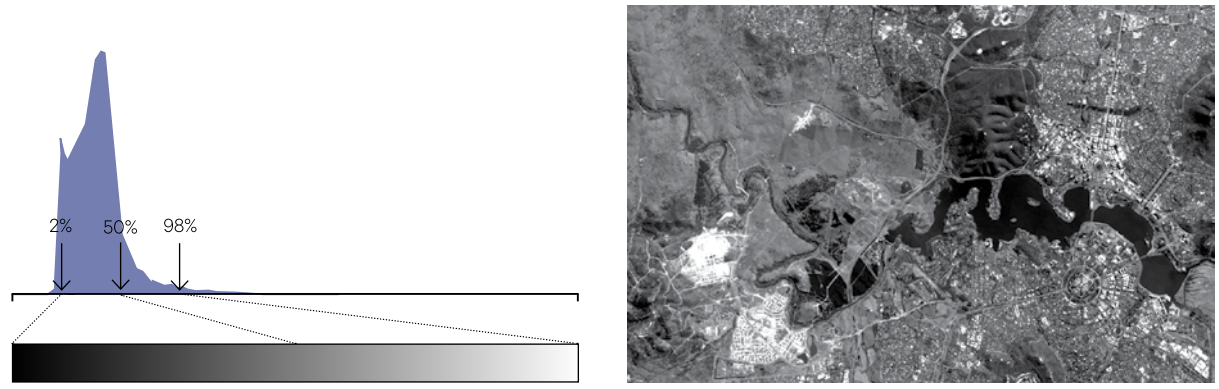
a. Unstretched image contains no true white or black pixels with most of the image being dark grey.



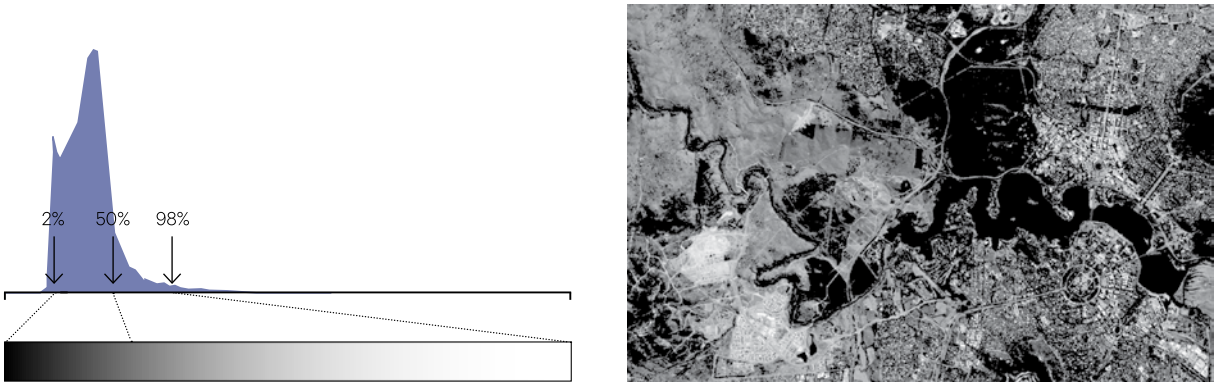
b. An image stretched between its minimum and maximum values contains some black and white pixels but most are still mid to dark grey.



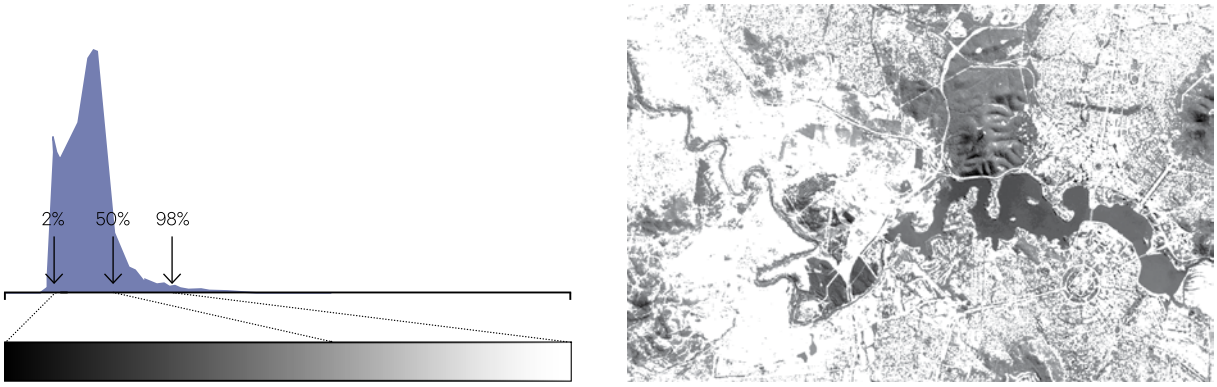
c. An image stretched between the 2% and 98% values uses the available range of grey shades more effectively.



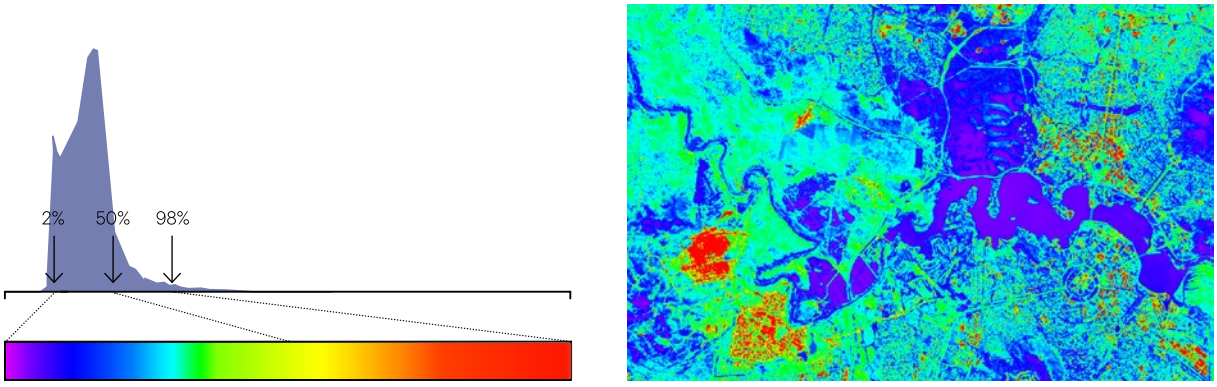
d. A non-linear (high Gaussian) stretch maps the high image values to a greater proportion of grey shades than the low image values. This improves the contrast in brighter areas at the expense of darker areas.



e. Non-linear (low Gaussian) stretch maps the low image values to a greater proportion of grey shades—this improves the contrast in darker areas but reduces contrast in lighter areas.



f. Pseudo-colour enhancement assigns different colours to different image values (see Section 5.2.1).

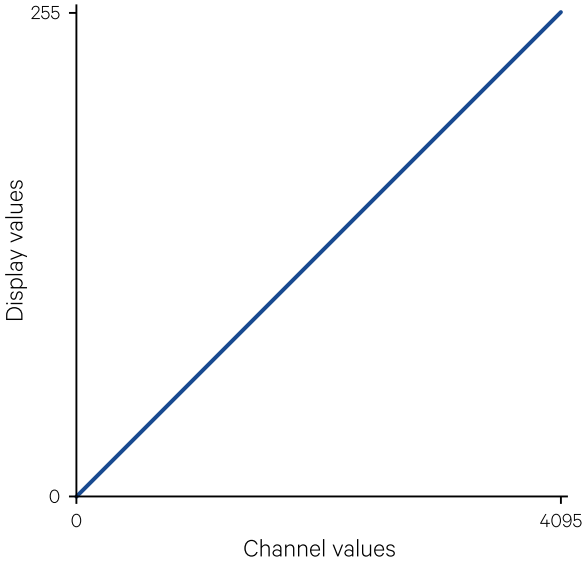


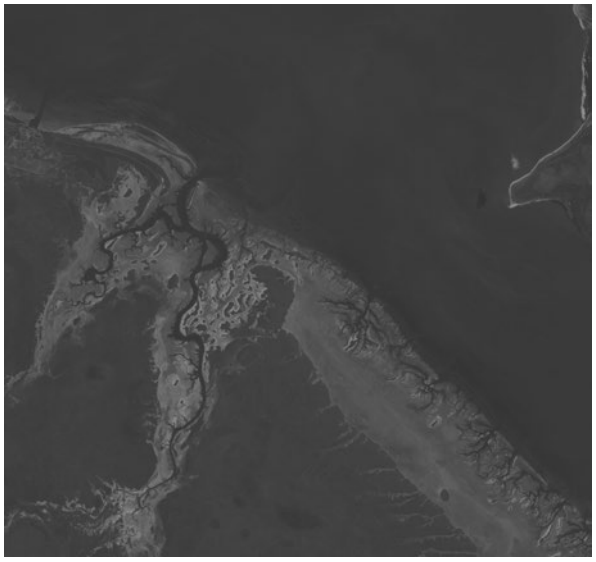
Source: Norman Mueller, Geoscience Australia

Figure 4.9 Linear stretching

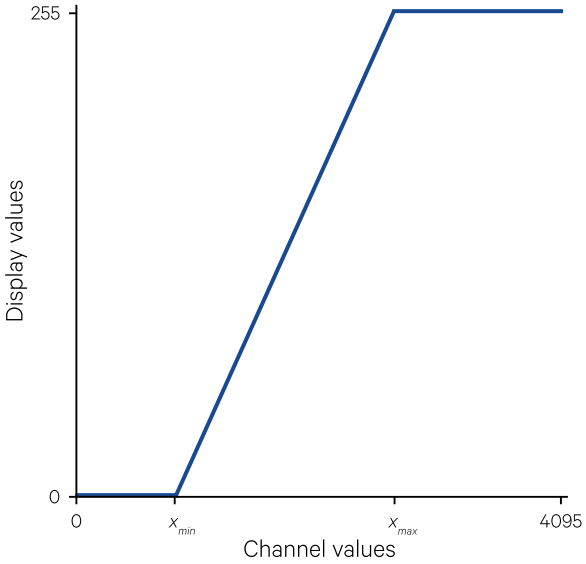
This panchromatic band was acquired by Landsat-8 on 7 July 2017 over Gulf of Carpentaria, northwest of Borroloola, NT.

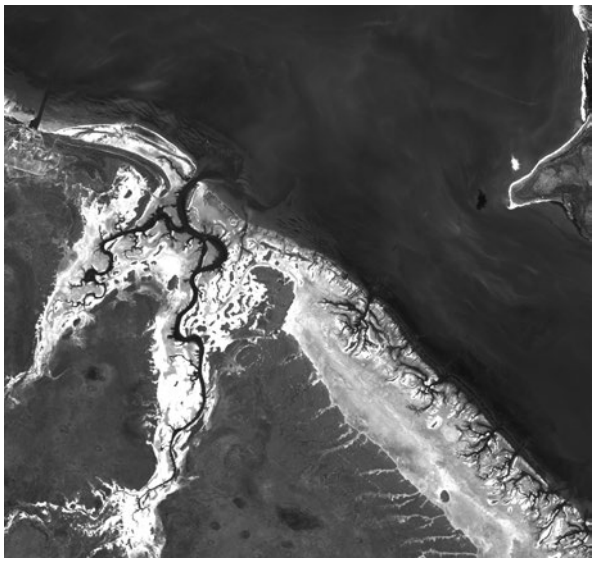
a. 12-bit Image channel displayed with value 0 as display minimum and value 4,095 as display maximum.





b. The range of image values between the nominated image minimum (x_{min}) and maximum (x_{max}) are linearly spread between the display minimum and maximum values. Image pixels with values below x_{min} are displayed as the display minimum and those with values above x_{max} are displayed as the maximum display value.





Source: Norman Mueller, Geoscience Australia

Linear stretching increases the overall contrast in a displayed image by increasing its variance. As illustrated in Figure 4.9, for an image channel being mapped to the display values 0–255, the range of displayed values, *y*, are related to the image value, *x*, as:

$$y = \begin{cases} 0 & \text{if } x < x_{min} \\ 255 \times \dfrac{x - x_{min}}{x_{max} - x_{min}} & \text{if } x_{min} \le x < x_{max} \\ 255 & \text{if } x \ge x_{max} \end{cases}$$

The variance of the displayed image values (y^2) will be approximately equal to:

$$\sigma_y^{\,2} = \left(\frac{255}{x_{max} - x_{min}} \right)^2 \times \sigma_x^{\,2}$$

so that, as the range in the image values is reduced, the variance of the displayed image increases, with values below x_{min} and above x_{max} being 'clipped' (that is, given the minimum and maximum values respectively). The contrast of a specific feature, such as a boundary, can be similarly increased by selecting minimum and maximum values that are appropriate to that feature.

The linear stretching technique assumes that the image histogram is uni-modal (see Section 4.1). Images with multi-modal histograms will require a different method for selecting these minimum and maximum values to optimise contrast in one of the value ranges. For example, Figure 4.10 shows a histogram of a thermal infrared channel for a daytime image which contains both land (warm) and water (cool) pixels. The unstretched image is shown in Figure 4.10a, and a linear stretch across both histogram peaks is shown in Figure 4.10b. The latter does not highlight features in either land or water. A stretch of values within the histogram peak with the lower values would highlight differences in the water but lose contrast in the land pixels (see Figure 4.10c), while a stretch of values spanning the histogram peak with higher values would produce good contrast in the land areas but cause water pixels to be displayed as black (see Figure 4.10d). (In this case, image segmentation methods could be used to separate the image into land and water regions, then rescale each segment separately before recombining into one image—see Sections 8, 9 and 10).

The contrast within a specific feature or region in an image may also be improved by using the minimum and maximum (or better, the 1% and 99%) values of that feature. In some cases these may be easily derived from the channel histogram as in the example in Figure 4.10. However, more commonly the appropriate minimum and maximum values would be identified using interactive training on the feature (see Section 9.1.2).

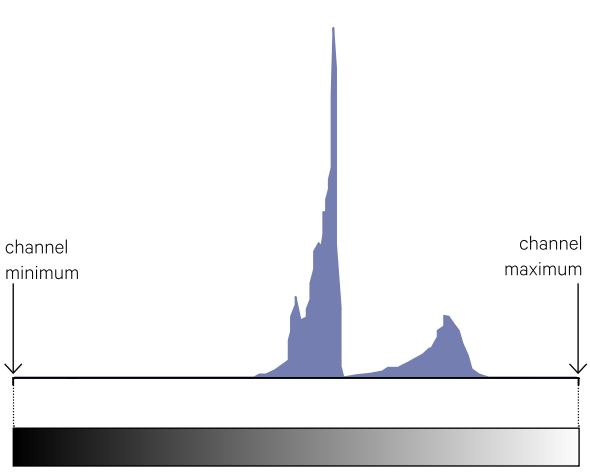
The channel data range may also be displayed as an inverse grey-scale by associating the contrast maximum value with black and the minimum with white (see Figure 4.10e). This display mode is usually referred to as an inverse or negative channel and is often represented as a negative channel number. Inverse channel display is useful for some temperature channels that are coded as increasing image numbers for decreasing temperature.

Finally, either pseudo-colouring (see Section 5.2.1 and Figure 4.10f) or density slicing (see Section 9.2.1) can also be used to visually enhance single image channels by adding colour.

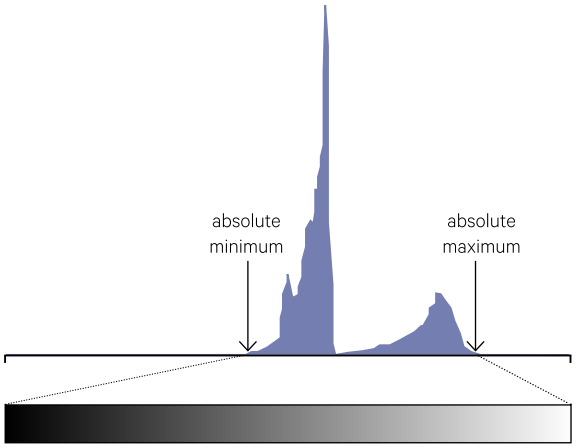
Figure 4.10 Bi-modal histogram stretching

This Landsat-8 thermal channel over Shark Bay, WA was acquired on 3 August 2017.

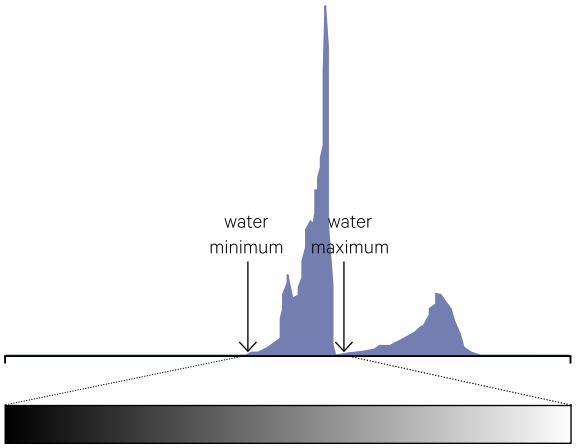
a. Original unstretched image

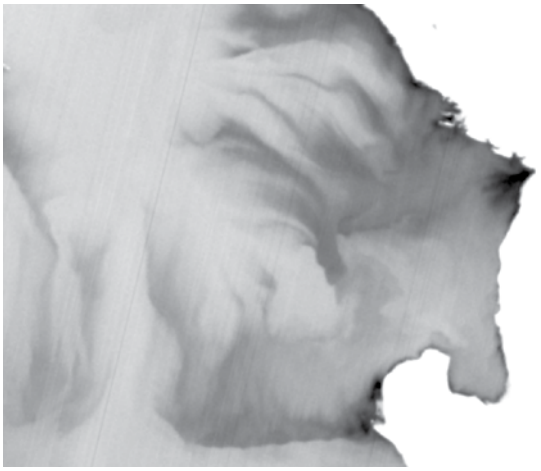


b. Linear stretch over full range of image data channel

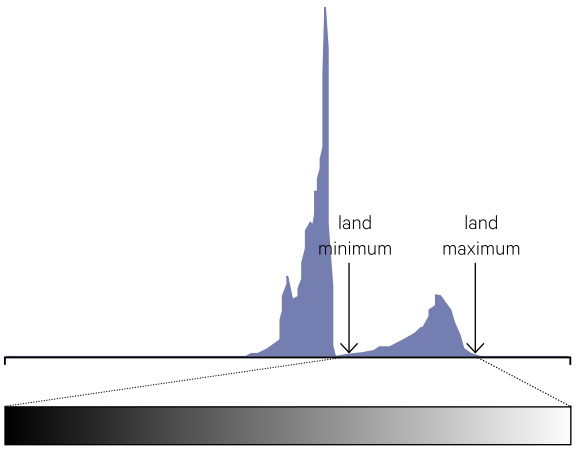


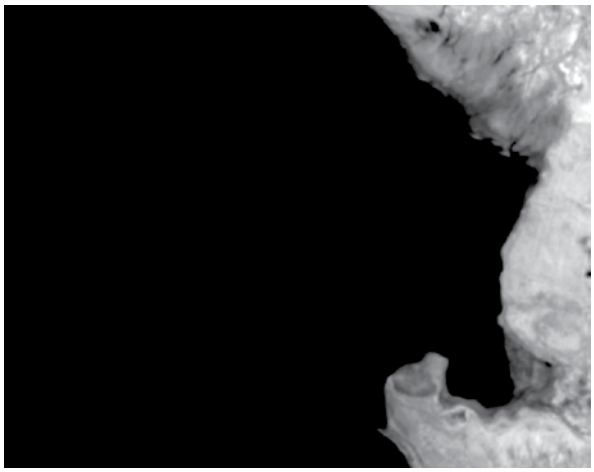
c. Stretching a range of low channel values defined by the first histogram peak improves contrast in the water but loses all contrast in the land areas.



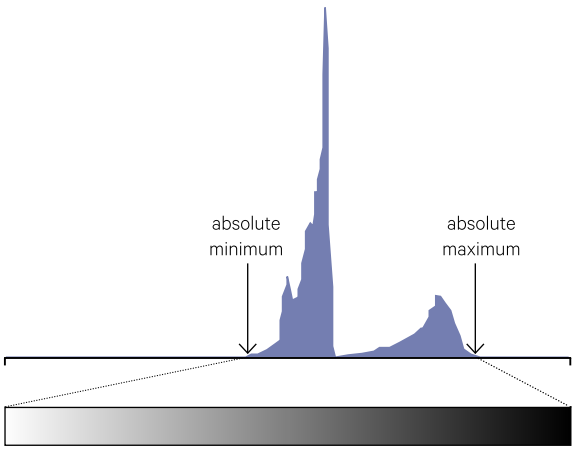


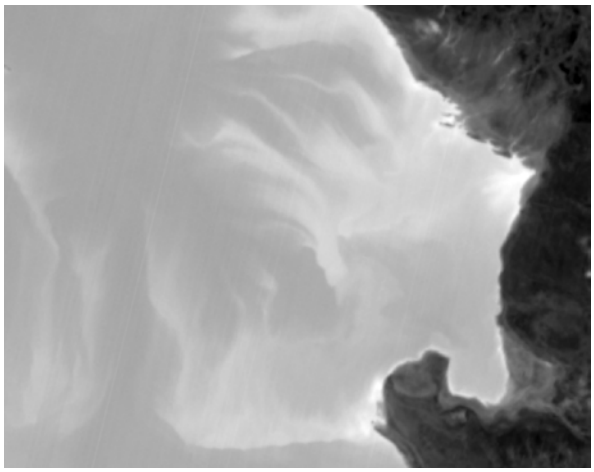
d. Linear stretching of the second histogram peak loses all water contrast to maximise contrast in land features.



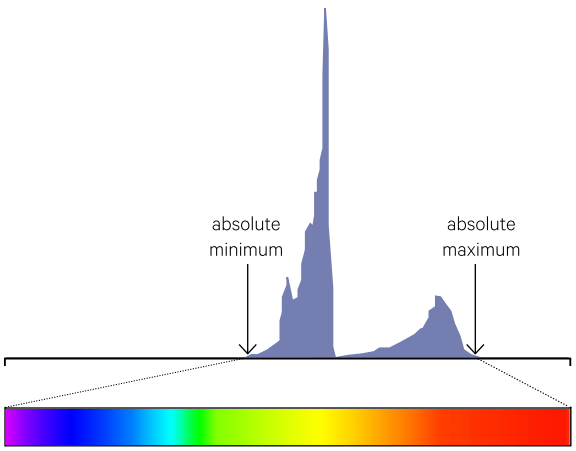


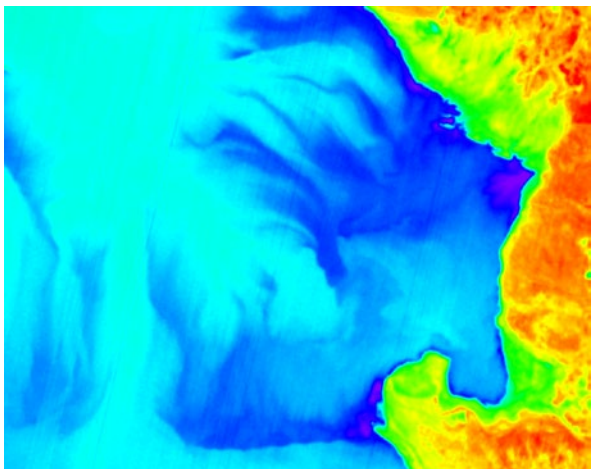
e. The image data range can also be inverse scaled so that high image values indicate low temperatures (darker) and low image values indicate high temperatures (lighter).





f. Pseudo-colouring of the full histogram range (with a different colour for each image value) improves visual discrimination over most of the image.





Source: Norman Mueller, Geoscience Australia

4.2.2 Non-linear stretching

The mapping of image values (within the defined contrast range) to the available display grey shades (or colours) may also be done in a non-linear fashion as illustrated in Figure 4.8d and Figure 4.8e. Image data are typically rescaled in a non-linear way using the logarithmic or exponential functions.

A logarithmic function is illustrated in Figure 4.11a. This graph represents the equation of:

$$y = \log_{10}(x)$$

where

$$x = 10^{y}$$

The exponential function is expressed as:

$$y = \exp_{10}(x)$$

where

$$\exp_{10}(x) = 10^{y}$$

(see Figure 4.11b).

These functions are often expressed more generally as $\log_a(x)$ or a^x where the choice of a affects the scaling and properties of the functions. When a is the natural constant 'e', the logarithmic function, $\log_e(x)$ or *ln*, is referred to as the Napierian logarithm and has special properties:

$$a^{x} = e^{\log(a^{x})} = e^{x \log(a)}$$

The base 10 number system computes these functions with the simply remembered values of:

$$\log_{10}(1) = 0 \qquad \exp_{10}(0) = 1$$
$$\log_{10}(10) = 1 \qquad \exp_{10}(1) = 10$$
$$\log_{10}(100) = 2 \qquad \exp_{10}(2) = 100$$

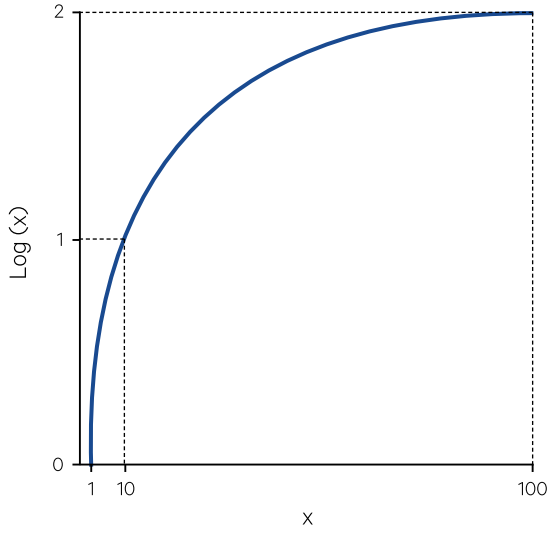
as illustrated in Figure 4.11.

As these figures and sample values show, the exponential is the inverse function of the logarithm so can be used to convert logarithmically-scaled data to a linear scale. Similarly, the log of an exponent produces linear scaling, that is:

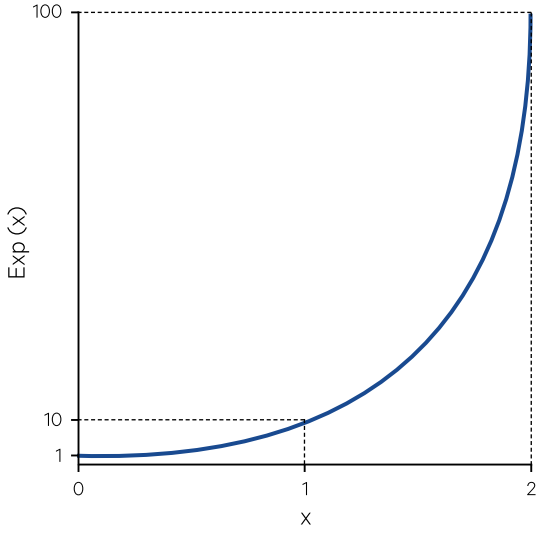
$$x = \log(\exp(x)) = \exp(\log(x))$$

Figure 4.11 Logarithmic and exponential functions

a. Logarithmic function: y = $\log_{10}$ (x)



b. Exponential function: y = $\exp_{10}$ (x)



Source: Harrison and Jupp (1990) Figure 65

Non-linear stretches are frequently implemented using an exponential function in the form:

$$y = \begin{cases} A & \text{if } x \le x_{min} \\ A + B \times \left(x - x_{min} \right)^{\gamma} & \text{if } x_{min} < x \le x_{max} \\ A + B \times \left(x_{max} - x_{min} \right)^{\gamma} & \text{if } x > x_{max} \end{cases}$$

where

x is an image pixel value;
x_{min} is the selected minimum image value for stretching;
x_{max} is the selected maximum image value for stretching;
y is the output (displayed) grey-scale value;
A is the minimum output level;
B is the scaling factor between the input and output ranges; and
γ is the gamma stretch parameter.

A linear stretch is equivalent to a gamma value of 1.0. A gamma stretch value (γ) less than one produces the result shown in Figure 4.8e, where contrast is increased in the low values (dark areas in the displayed image) but reduced in the high values (bright areas). Conversely, a gamma value greater than one results in contrast being improved in the high values (bright areas) and reduced in the low values (dark areas; see Figure 4.8d and Figure 4.12b). Gamma stretch parameters less than 1.0 are usually required during image printing to improve matching of colour intensities between the display screen and the printer (see Section 6).

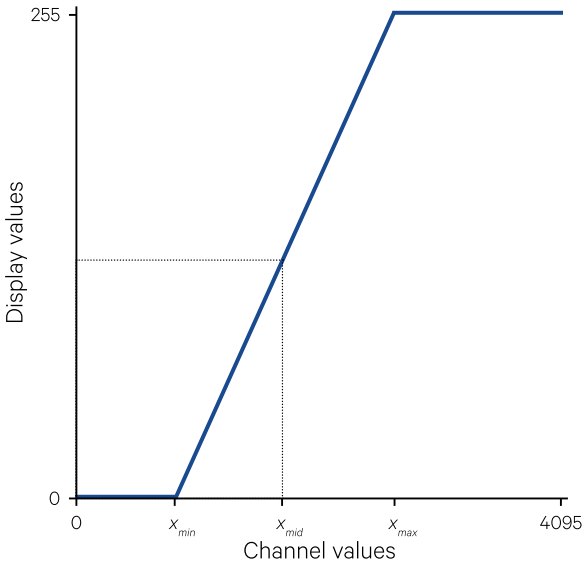
In terms of image contrast, a log transformation produces a similar result to a gamma function when the gamma value is less than one (see Figure 4.13), while the exponential transformation results are similar to a gamma function with value greater than one. Logarithmic and exponential stretches can also be implemented as image transformations as detailed in Volume 2C. Other rescaling methods, such as destriping, are also discussed in Volume 2C.

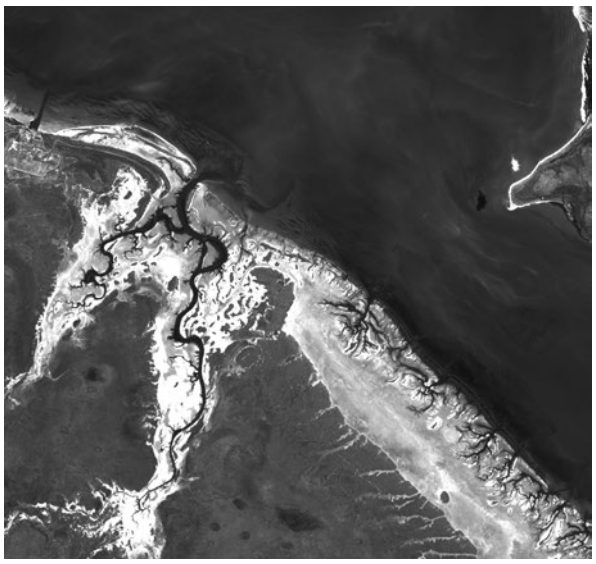
There are dark shadows on the earth,
but its lights are stronger in the contrast.
(Charles Dickens)

Figure 4.12 Enhancing high image values

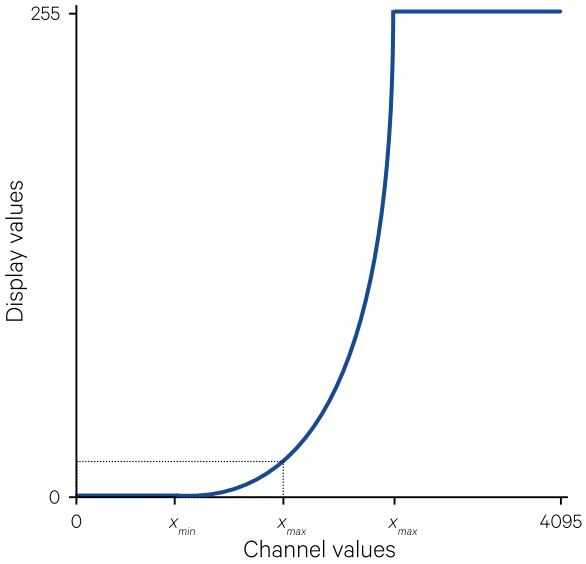
This panchromatic band was acquired by Landsat-8 on 7 July 2017 over Gulf of Carpentaria, northwest of Borroloola, NT.

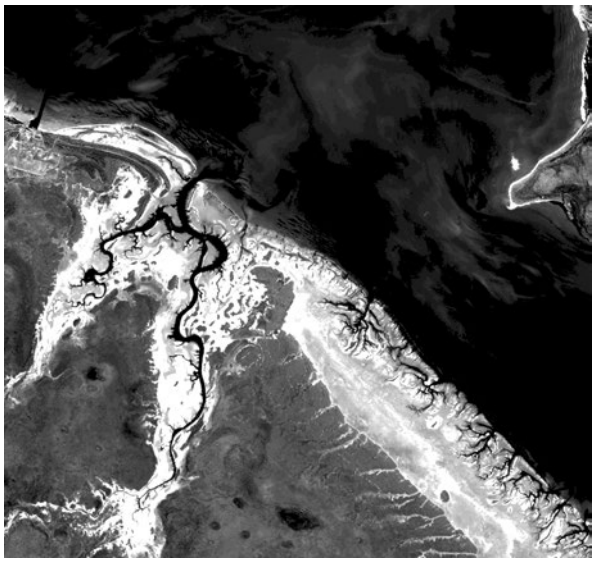
a. Linear stretch of panchromatic channel (see Figure 4.9)





b. Same image with high Gaussian stretch, which increases contrast in features with high image values.



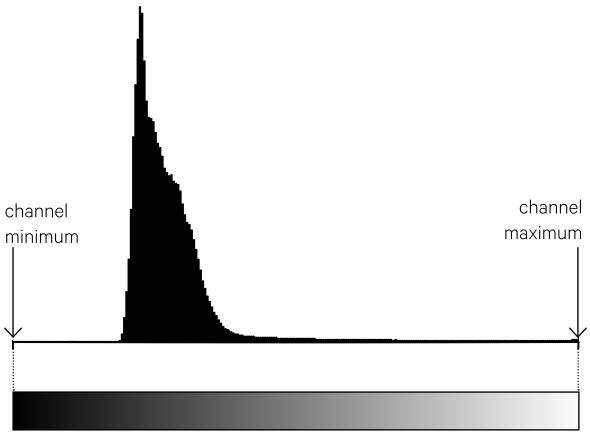


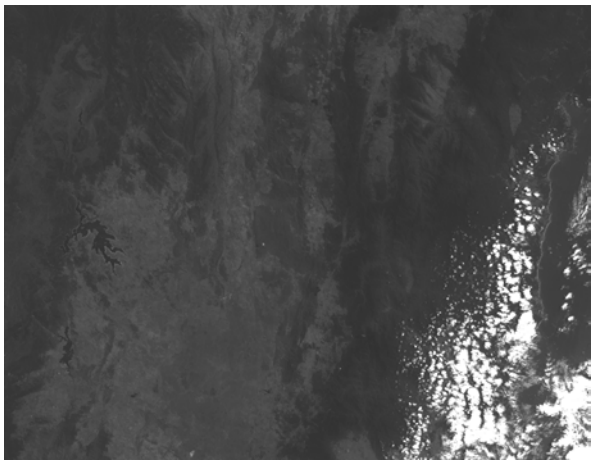
Source: Norman Mueller, Geoscience Australia

Figure 4.13 Logarithmic Stretching Example

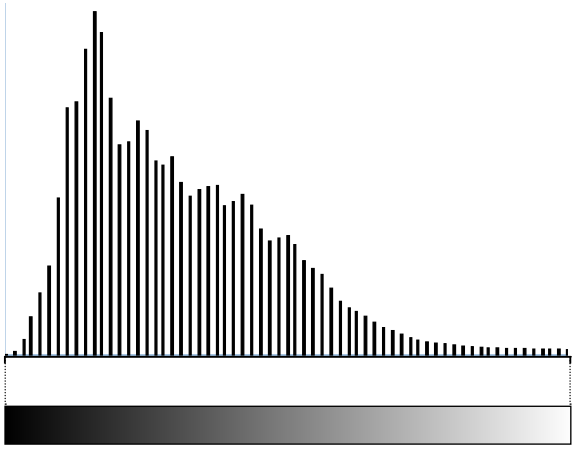
This example uses an 8-bit (byte) format image channel displayed using an 8-bit colour space. This image was acquired by Landsat-5 on 20 January 2010 over southeastern NSW. The town of Cooma is near the centre of this image, with Canberra in the northwest and the coastline visible in the east. Band 1 (visible blue) is used to differentiate between spectral features of mountainous terrain with forest cover, cleared pasture, water bodies, and cloud.

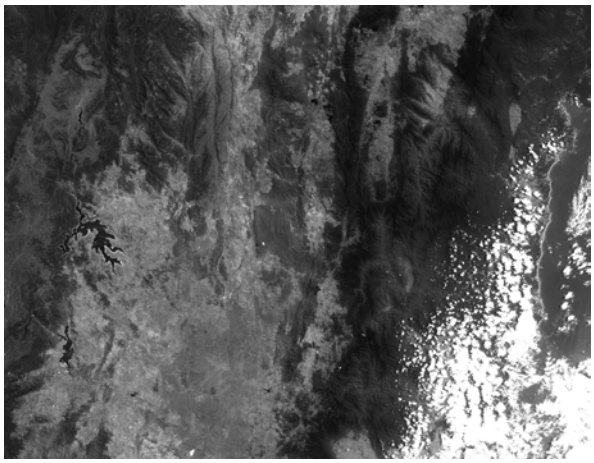
a. Unstretched image



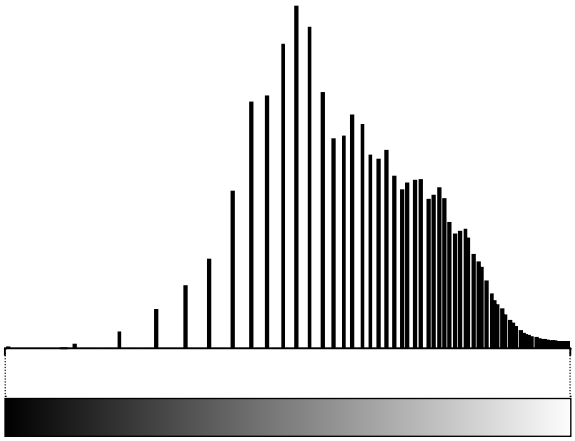


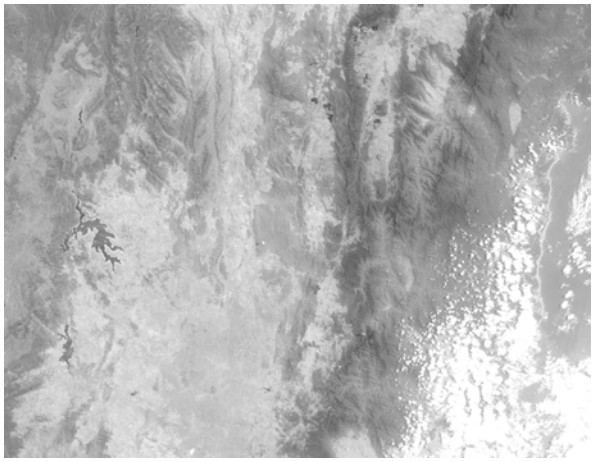
b. Linear stretch





c. Logarithmic stretch, which increases contrast in low image values.





Source: Tony Sparks, Icon Water

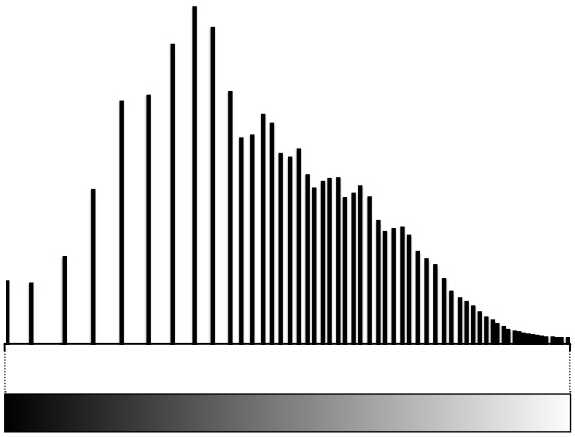
4.2.3 Statistical stretching

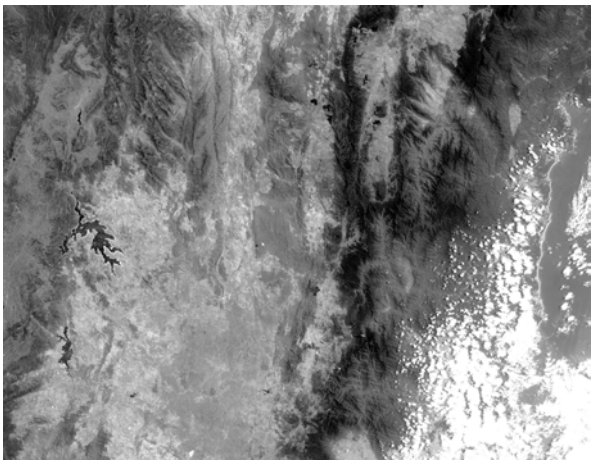
4.2.3.1 Normalisation

Some image processing systems offer an option during image display to fit each image channel to a normalised distribution based on sample channel statistics (see Figure 4.14). For normally distributed data, 95% of data lies within two standard deviations of the mean. The (sample) channel statistics can be used to determine a minimum value equivalent to the mean value minus one standard deviation and a maximum value equivalent to the mean value plus one standard deviation. These statistically-determined minimum and maximum values thus span 95% of the data range. This channel distribution can also be used to spread image values appropriately between the selected minimum and maximum values (see Figure 4.14).

Figure 4.14 Normalised stretch

Normalised stretch applied to image shown in Figure 4.13.





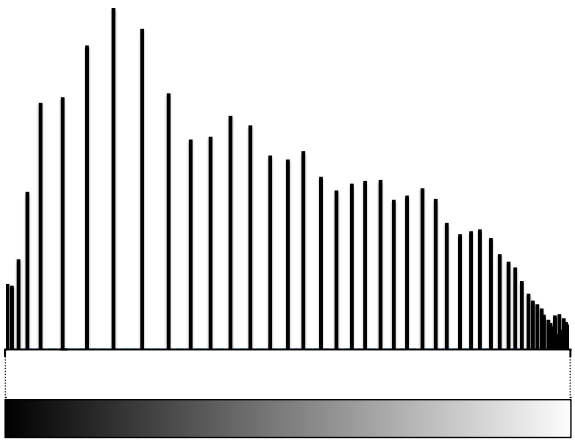
Source: Tony Sparks, Icon Water

4.2.3.2 Histogram equalisation

As the name implies, this technique attempts to equalise the number of pixels assigned to each value in the image. If values were distributed equally in the rescaled image channel it would have a flat histogram. In EO imagery, however, the initial distribution of image values over a restricted data range generally disallows this result. Histogram equalisation is a useful method for both increasing overall variance and improving contrast between those values occurring most frequently in an image (see Figure 4.15).

Figure 4.15 Histogram equalisation

Histogram equalisation applied to image shown in Figure 4.13



Source: Tony Sparks, Icon Water

This operation is based on the grey level (or bin) histogram counts, n_1, n_2, ... n_N, for each image value x_j, where j = 1 to N. For a byte image, the rescaled values x_j' are computed as:

$$x_j' = \frac{255 \times \sum_{i=1}^{j} n_i}{\sum_{i=1}^{N} n_i}$$

A linear stretch can be included in the rescaling by defining:

n_j = image minimum value for $j < x_{min}$ and
n_j = image maximum value for $j > x_{max}$.

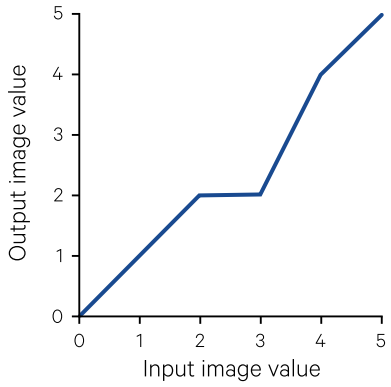
In many image processing systems, histogram equalisation can be applied as a statistical stretch during the image display process.

4.2.4 Lookup tables

A lookup table (LUT) is an expedient method for implementing a range of rescaling operations by allowing a defined output value for each possible input value in the image. A LUT is a single column table in which rows represent input data values and table entries indicate the output values to assign to each input value. The definitions can represent a wide variety of contrast enhancements such as piecewise linear stretch or sawtooth stretching (also called 'overlapping contrast enhancement') as illustrated in Figure 4.16. The impact of a user-defined LUT on a grey scale image is shown in Figure 4.17. A random LUT could also be generated to highlight small contrast differences between features by chance.

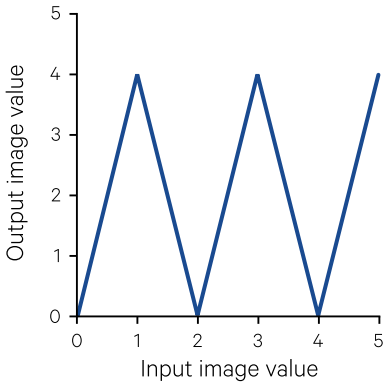
Figure 4.16 LUT operations for rescaling image data values

a. Piecewise linear rescaling



LUT entry	Output Value
0	0
1	1
2	2
3	2
4	4
5	5

b. Sawtooth stretch

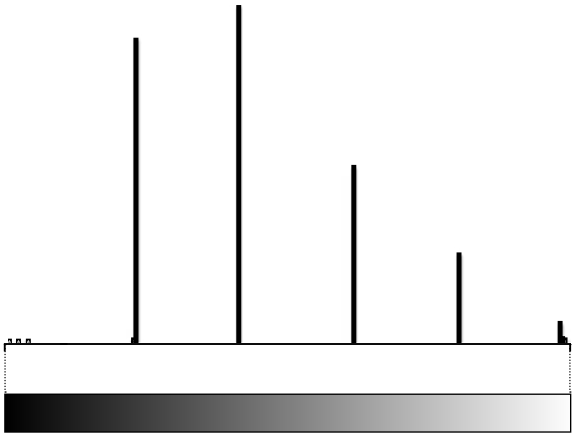


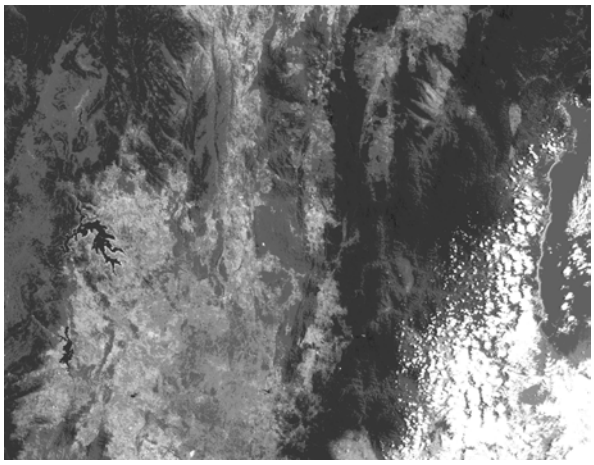
LUT entry	Output Value
0	0
1	4
2	0
3	4
4	0
5	4

Source: Harrison and Jupp (1990) Figure 10

Figure 4.17 User-defined stretch

User-defined stretch applied to image in Figure 4.13





Source: Tony Sparks, Icon Water

LUT can be used to implement statistical rescaling operations such as histogram equalisation (see Section 4.2.3.2). The use of LUT is further discussed in the context of pseudo-colouring (Section 5.2.1), image display (Section 6.1) and image transformations (Volume 2C).

4.3 Further Information

Brightness and Contrast: *https://www.tutorialspoint.com/dip/brightness_and_contrast.htm*
http://www.dspguide.com/ch23/5.htm

Castleman (1998) Chapter 5.

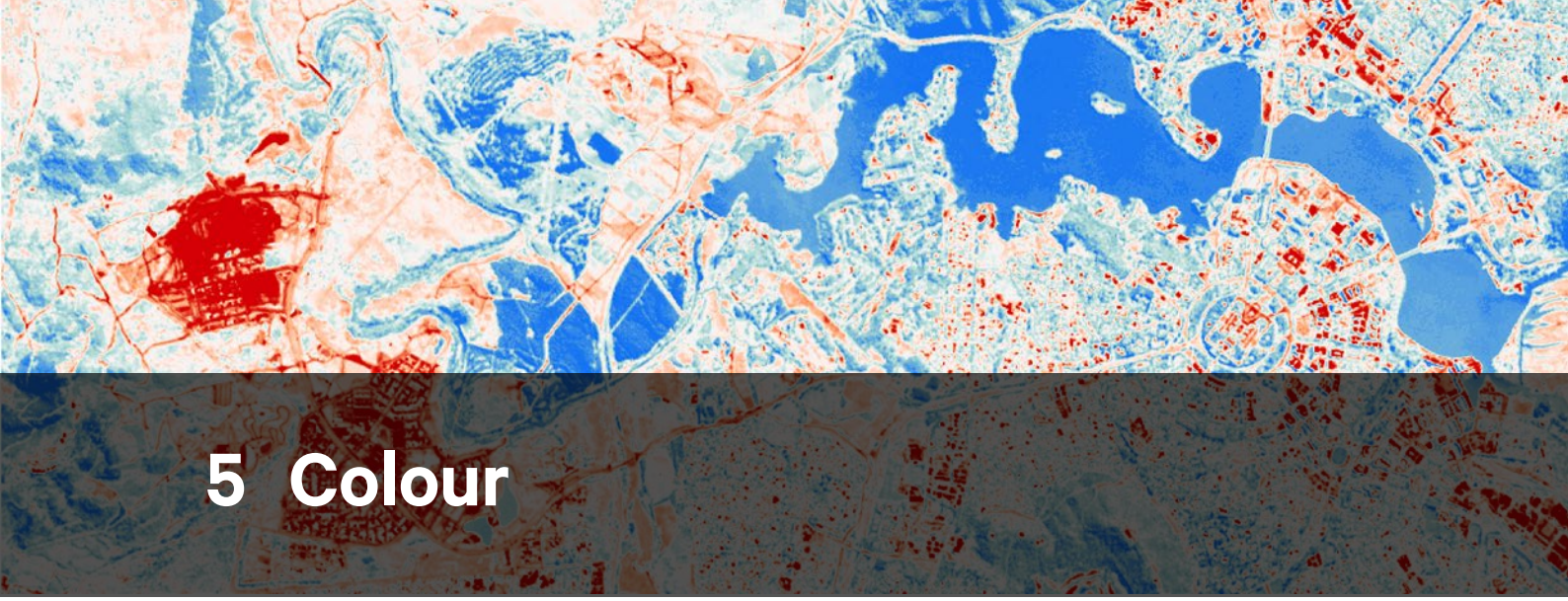
Jensen (2016) Chapter 4

4.4 References

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice Hall. 667 pp.

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Jensen, J.R. (2016) *Introductory Digital Image Processing. A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc.



5 Colour

Colour is an important aspect of image presentation and interpretation. Discussion so far has referred to representing image data as shades of grey between black and white. Since the human eye is not particularly good at detecting grey shades, image data can be interpreted more easily if displayed as a range of different colours. The type and range of colours used to represent data values in a colour image will directly affect our visual interpretation of the data by enhancing or reducing our perception of its patterns and features. Various contrast enhancement transformations discussed above attempt to improve this displayed colour range.

Colour look-up tables can be used to represent single channel image data using pseudo-colour techniques as described in Section 5.2.1 or by density slicing (Section 9.2.1), while multi-channel imagery can be represented as colour composites (Section 5.2.2). The broader issues of colour description, ordering and reproduction which are relevant to image representation are discussed initially in Section 5.1.

5.1 Colour basics

5.1.1 Colour description

Our common language descriptions for different colours are often vague and inconsistent making it difficult to use such descriptions to precisely define a particular colour. This situation is further exacerbated by the differences between the measured and perceived brightness of colours and the variations in visual perception within the human population. The latter factors are further discussed in Excursus 5.1.

Before considering some of the systems that attempt to quantitatively define colours, we will look at aspects of individual colours that may be used to describe our perception of them. Three attributes that are frequently used to describe colour are hue, value and chroma.

The term hue describes the attribute that is commonly referred to as colour, that is the combination of different wavelengths used to produce a colour. A 'natural' ordering of hues follows the colour spectrum (with increasing wavelength) from violet, through indigo, blue, green, yellow, and orange to red. The colours in this spectrum represent the spectral hues, that is, those produced by mixing adjacent wavelengths in the spectrum. Non-spectral hues are formed by mixing two or more spectral hues, such as purple (blue plus red) or brown.

The perceived brightness (lightness or darkness) of a colour is referred to as its value or intensity. This is different to the measured lightness (or luminance) of a colour. Intensity depends on the amount of energy present in all wavelengths of a colour.

Chroma describes the saturation or purity of a colour. This indicates the range of wavelengths around the average colour wavelength in which most energy is carried. A single wavelength would produce a spectrally pure colour and hence high saturation value, whereas a mixture of wavelengths, such as in a pastel colour, would have low value. The full spectrum of wavelengths, when combined, appears as white. This attribute is often linked with the colour value so can be difficult to describe and identify. It can also be described as the amount of grey in a colour since a high chroma (or vivid) colour can be made duller by mixing with grey.

The perceived hue, value and chroma of any colour are all affected by the patch size and shape of the colour, and the colours of surrounding patches. These factors are discussed in Excursus 5.1 and present obvious problems for colour description in systems that rely only on these perceptual attributes. Excursus 5.2 further considers the problems of colour interaction and how these may be resolved to achieve a balance of contrast and harmony in digital imagery.

Background image: Pseudo-colouring of Landsat-8 panchromatic image of Canberra, ACT, acquired on 10 August, 2017.
Source: Norman Mueller, Geoscience Australia

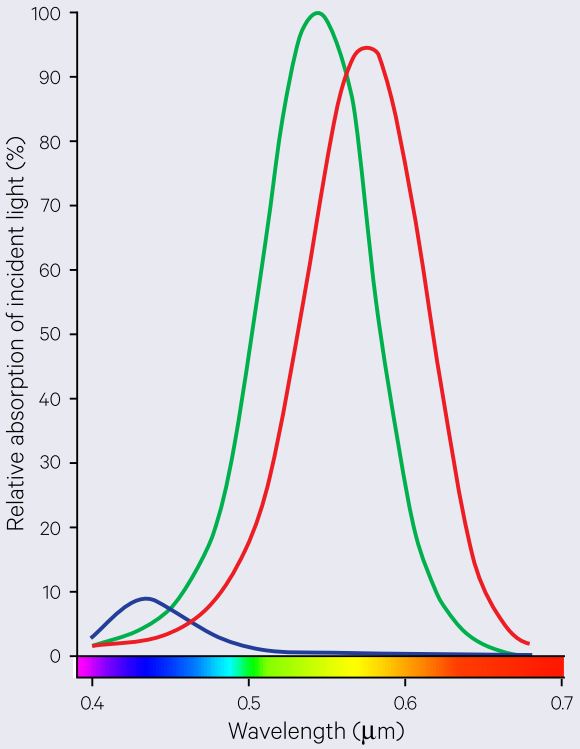
Excursus 5.1—Human Colour Perception

The precise mechanism of human colour vision is not well established. Various theories have been developed to try to explain this fundamental phenomenon. The currently accepted theory marries two previously competitive views as a two-stage process. These involve the physical, photochemical detection of a colour signal and its subsequent psychological processing.

The physical components of vision rely on two types of photochemical cells in the retina of the eye. The rods detect intensity changes independently of colour. The physical sensation of colour is believed to be due to the spectral sensitivities of three types of cones. These are commonly referred to as the blue, green and red receptors, which respond to short, medium and long wavelengths in the visible spectrum respectively (see Figure 5.1). These photochemical observations have been confirmed by various psychophysical colour mixing and matching experiments. This interpretation is often referred to as the tristimulus theory of colour vision and supports the observation that most colours can be produced from these three primary colours (see Appendix 3).

Figure 5.1 Relative light absorption of human cones

In the human eye, the relative absorption of light differs for different cones, with blue cones being significantly less responsive than green or red cones.

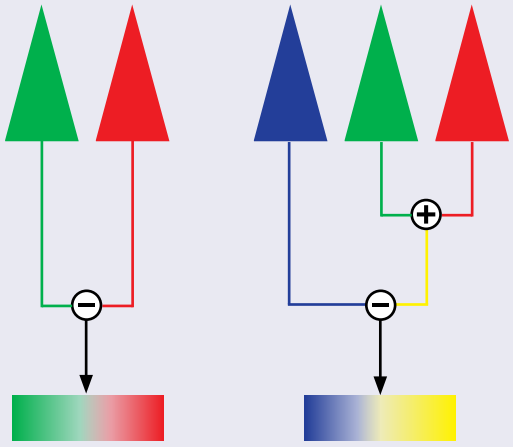


Source: Harrison and Jupp (1989) Figure 9 [Adapted from Cornsweet (1970)]

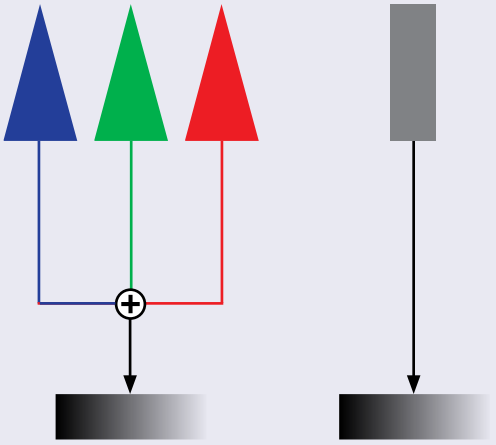
Figure 5.2 Opponent process theory of colour vision

This theory explains colour vision in terms of three colour planes. In the chromatic system, one plane (derived from the comparing the signals from red and green cones) transitions from red to green while the other creates a yellow signal from the red and green cones then compares this signal with the blue cone signal to produce a blue/yellow transition. In the achromatic system, signals from all cones are combined to produce a colour luminosity scale, which is used with the luminosity signal from the rods to perceive changes in brightness.

a. Chromatic system



b. Achromatic system



After the signals from these receptors have been coded, the brain analyses their differences to produce the psychological impression of colour. This analysis is explained by the opponent process theory (Hering, 1964), which proposes that colours are perceived in terms of three sets of 'opposing' colour pairs:

- red-green—red (R) and green (G) signals are compared to give a R versus G output signal (see Figure 5.2a);

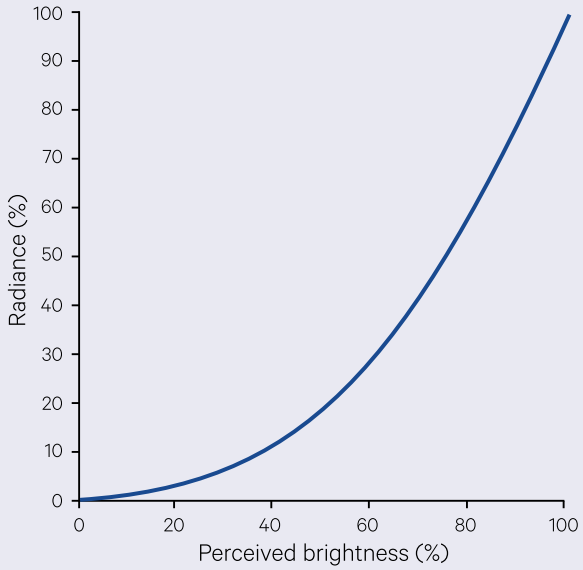
- yellow-blue—red and green are summed to produce a yellow signal (Y), which is then compared with the blue receptor signal (B) to produce a Y versus B response (see Figure 5.2a); and
- black-white—all receptor signals (R, G and B) are summed to produced a colour luminosity, which is processed in conjunction with an achromatic luminosity signal from the rod cells (see Figure 5.2b).

This recoding of the original signals reduces the transmitted colour information down to two colour difference signals and a luminance signal. A similar coding system of luminance and 'chrominance' was used in colour televisions to improve transmission efficiency (Padgham and Saunders, 1975).

Each individual may also see a particular colour differently depending on the sensitivities of his or her visual sensors. Figure 5.3 indicates the non-linear relationship between perceived and measured brightness. The context of a given colour intensity will also impact its perceived brightness (see Figure 5.4).

Figure 5.3 Perception of intensity changes

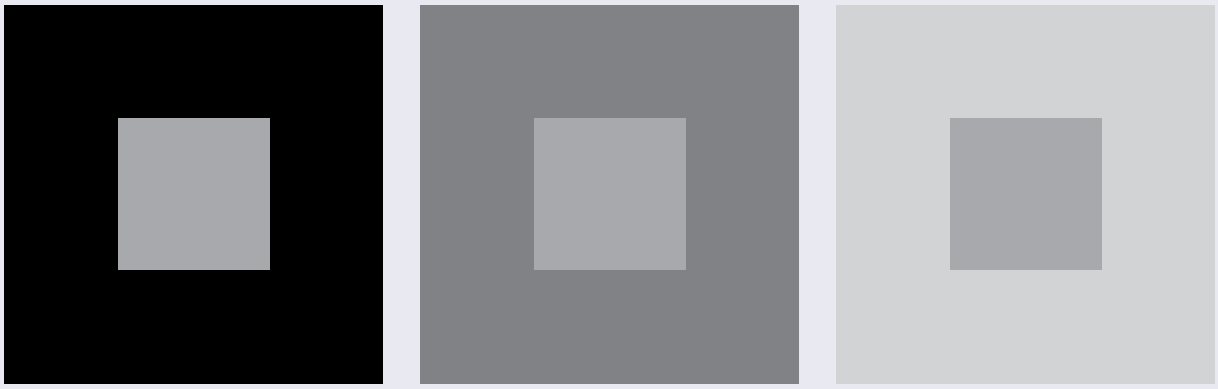
Differences in high intensities (brighter colours) are not well perceived, whereas smaller differences in low intensities (darker colours) are more easily perceived



Source: Harrison and Jupp (1990) Figure 39 [Adapted from Land (1977)]

Figure 5.4 Effect of adjacent intensities on perceived brightness

The central squares below all have the same brightness (40%) but appear as lighter and darker shades depending on the adjacent intensities (outer squares left to right: 100%, 60% and 20%).

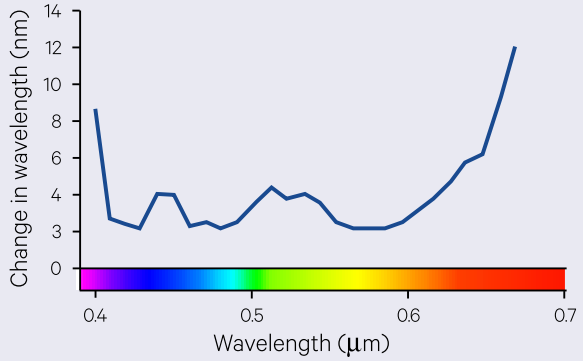


The spectral sensitivity of the human eye also varies with wavelength, being most sensitive to yellow-green colours. This means that a light radiating yellow wavelengths at a particular energy level will appear brighter than a light radiating the same amount of energy in red wavelengths. This phenomenon has obvious implications for generating imagery with apparently uniform brightness over a range of colours.

Another interesting variable in human vision is the threshold levels for perceiving chromaticity changes at different wavelengths. As shown in Figure 5.5, changes in chromaticity at wavelengths corresponding to blue/green or orange are more readily perceived than at wavelengths corresponding to blue, yellow/green or red.

Figure 5.5 Perception of chromaticity changes

The extent to which changes in chromaticity are perceived varies with wavelength.

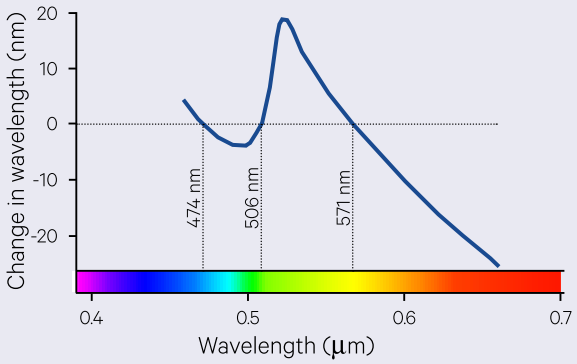


Adapted from: Cornsweet (1970)

It has been demonstrated that the human perception of spectral hue, like brightness and saturation, depends on both the wavelength and intensity of the stimulus ('Bezold-Brucke Phenomenon'). As illustrated in Figure 5.6 for a constant wavelength, hue varies with intensity except at three invariant points: 474 nm, 506 nm and 571 nm (Padgham and Saunders, 1975). Perceived hue has been observed to change with exposure time (Cornsweet, 1970) and colour saturation can also affect perceived brightness ('Helmholtz-Kohlrausch Effect'; Padgham and Saunders, 1975).

Figure 5.6 Effect of colour intensity change on perceived hue

The apparent change in wavelength at lower luminance levels is plotted against that wavelength at high luminance. Perceived hue varies with wavelength except at 474 nm, 506 nm and 571 nm.



Adapted from: Padgham and Saunders (1975)

5.1.2 Models for colour ordering

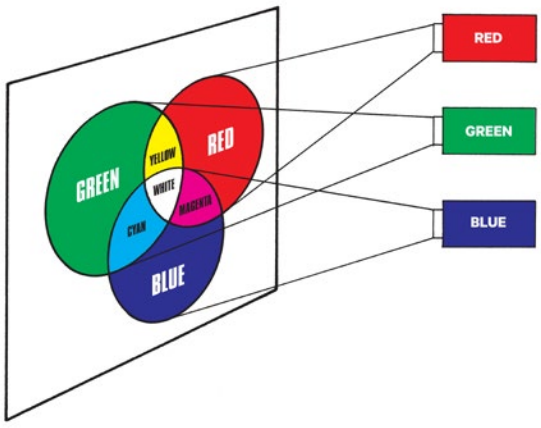
Many and various systems have been proposed to quantify or identify colours and colour differences (Niblack, 1986). These include generic models such as the simple RGB (red, green, blue) colour cube, the HSI (hue, saturation, intensity) colour space and various chromaticity diagrams as well as specialised systems such as the Munsell colour system, the CIE (Commission Internationale de l'Eclairage) models and industrial chromaticity standards.

The systems that are used to describe or identify colours rely on either technical characteristics, such as how a colour is produced or physically measured in terms of wavelengths, or perceptive qualities, that is, how we actually see it. This perception depends on both the physical colour and our visual processing of it, the latter being affected by both surrounding colours and lighting conditions. Some colour description systems attempt to use both physical and perceptual colour attributes. Colour spaces in which the perceptual difference between two colours is linearly related to the Euclidean, or linear, distance between them are termed uniform colour spaces. When image colours are selected from a uniform colour space, the visual discrimination of processing results is optimised.

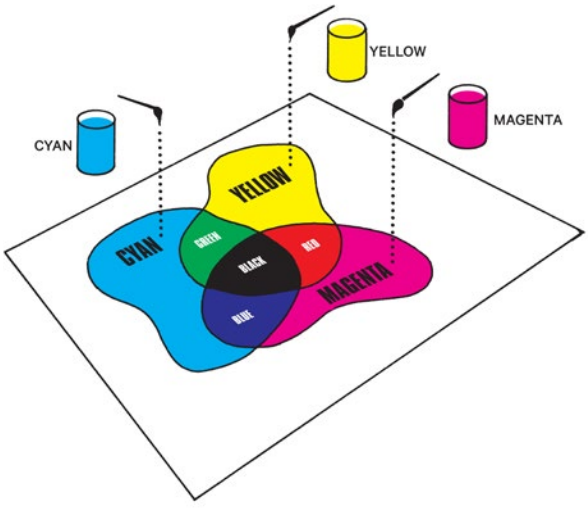
Why do two colours, put one next to the other, sing?
Can one really explain this? No.
Just as one can never learn how to paint.
(Pablo Picasso)

Figure 5.7 Additive and subtractive colour mixing

a. The additive primaries (red, green and blue) are added to a black background to give other colours.



b. The subtractive primaries (yellow, magenta and cyan) subtract from a white background to produce other colours.

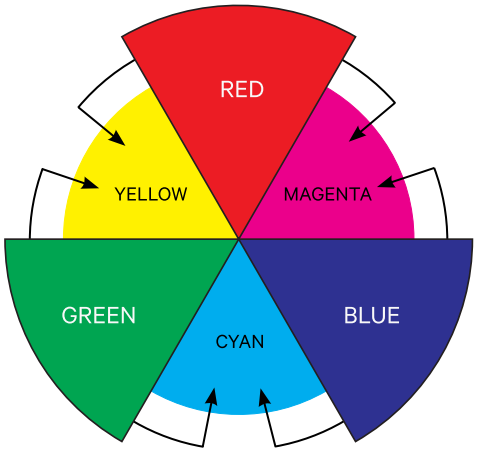


Source: Harrison and Jupp (1990) Figures 10 and 11

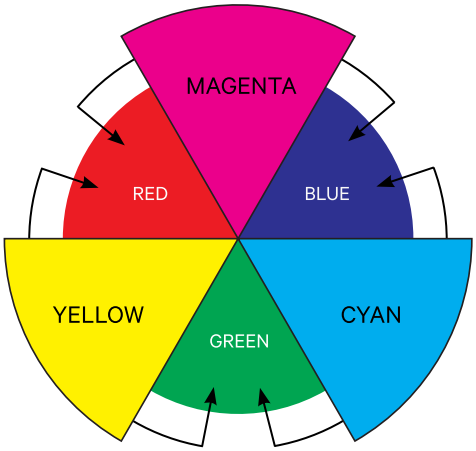
One of the simplest systems to represent colours relies on the additive primaries: red, green and blue (see Appendix 3 and Volume 1). The additive primary colours represent the basic hues that can be added together to produce white and all other colours (see Figure 5.7a and Figure 5.8a). Colour representation devices that operate by adding coloured light to a black base, such as projectors and display monitors, use additive primaries. Pairs of these primary colours can be added together to form the subtractive primaries, yellow (red plus green), magenta (red and blue) and cyan (green and blue). Film dyes and print media use the subtractive primaries to produce other colours by 'removing' their complementary colours from white (see Figure 5.7b and Figure 5.8b).

Figure 5.8 Relationship between additive and subtractive primary colours

a. Additive primaries—can be combined to form brighter, secondary colours; added all together they give white light



b. Subtractive primaries—produce secondary colours by removing their complementary colours from white; added all together they produce black



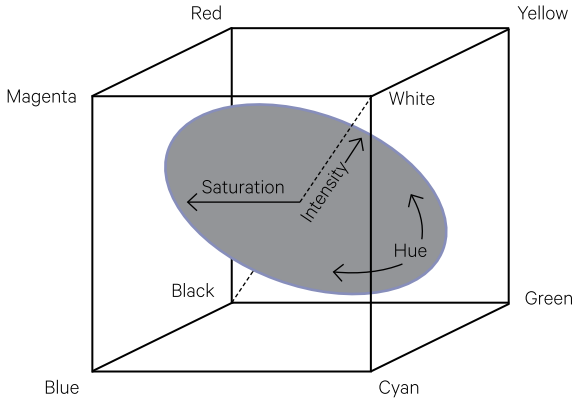
Source: Harrison and Jupp (1990) Figure 12

These colour components can be viewed as three dimensions forming a colour cube as illustrated in Figure 5.9. Due to the complementary nature of the additive and subtractive primaries, the two-dimensional 'corners' of this space represent the subtractive primaries yellow, magenta and cyan, with black and white at opposite corners of the space. Thus, the axis of grey shades is the three-dimensional line from black to white where the proportions of each primary colour are equal. A limitation of this system for defining colours is that the visual perception of colour differences is not necessarily related linearly to differences measured within the colour space.

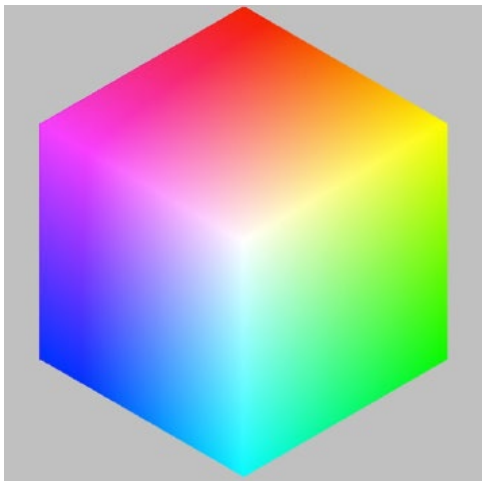
Figure 5.9 RGB colour cube

The colours formed by differing proportions of the additive primary colours, red green and blue, can be represented as a three-dimensional data space. The diagonal axis in this space, referred to as the intensity axis, represents the shades of grey from black to white that have equal proportions of each primary colour. The subtractive primaries occur at the diagonal 'corners' on each two-dimensional plane through two axes. An infinite number of planes that are orthogonal to the intensity axis can be defined. Colours on each plane have equal intensity, with saturation increasing away from the axis and hue varying around it.

a. Orientation of RGB colour cube



b. Colour planes on surface of RGB colour cube



Source: a. Harrison and Jupp (1990) Figure 11; b. Frank Clark. Retrieved from: *http://www.frank-t-clark.com/Professional/Papers/ColorHCW/ColorHCW.html)*

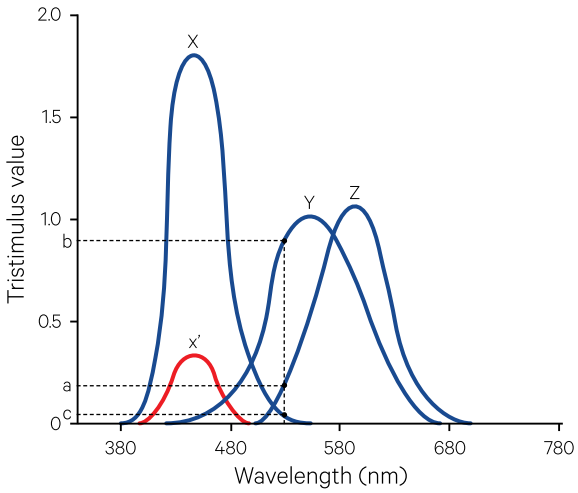
The colour coordinates defined by the RGB colour space can be expressed in terms of the perceptual attributes of hue, value (or intensity) and chroma (or saturation). The RGB axes which represent colours as Cartesian coordinates (see Volume 2B) can be mathematically rotated to cylindrical or triangular coordinates for planes of equal intensity in the cube. Since colour intensity changes along the intensity axis of the cube, as indicated in Figure 5.9, planes which are orthogonal to this axis may be defined on which the colours have theoretically equivalent intensity, with hues changing around the axis and saturation increasing radially away from it. These projections are further derived in Appendix 3.

Colour composite images are sometimes converted to the HSI colour space to manipulate the colour attributes individually. This conversion involves coding the hue, saturation and intensity coordinates as three image channels. The transformations for converting between RGB and HSI coordinates are detailed in Appendix 3. The HSI colour space is used as a basis for a number of specialised colour systems such as the Munsell System (Munsell, 1929).

A more quantitative approach to colour description was developed by the Commission Internationale de l'Eclairage (CIE). This system algebraically converts the proportions of the three primary colours required to match a given colour to the proportions of three artificial or imaginary primaries. These values are referred to as the CIE tristimulus values and are labelled X, Y and Z. The artificial primaries can be expressed as three curves as shown in Figure 5.10. For example, a colour of a single wavelength of 530 nm would be represented by the values a in X, b in Y and c in Z. Colours can be matched directly with a mixture of 1, 2 or 3 primaries, or by first adding with one or more primaries and then matching with the remaining primaries. The x' curve in the graph represents that red light needs to be added to the single wavelength colours between 400 nm and 500 nm to match the colours produced by mixing the corresponding levels of X and Y.

Figure 5.10 CIE imaginary primaries

A single wavelength colour (such as 530 nm) can be expressed in terms of its X, Y and Z component values (for example, values a, b and c for the wavelength 530 nm). The x' curve indicates the amount of red light which needs to be added to single wavelength colours between 400 nm and 500 nm to match the colours which are produced by mixing the corresponding levels of X and Y.



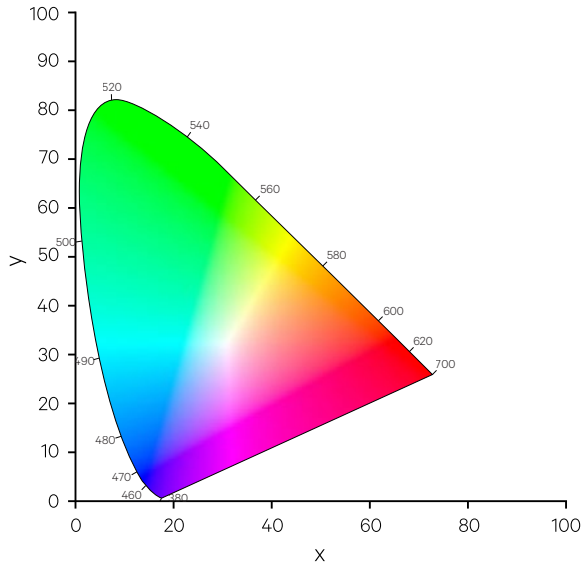
Source: Harrison and Jupp (1990) Figure 12 [Derived from CIE 1931 Standard Colorimetric Observer: CIE 1986]

This system can be used to produce a colour space that contains all perceptible colours. These are usually plotted in a normalised three-dimensional space in which two dimensions (x and y) define colour (or chromaticity) and the third defines luminance (see Appendix 3 for details). A two-dimensional plot of the colour dimensions is referred to as a chromaticity diagram or triangle, which represents changes in hue and saturation for a given intensity level. This plot forms the characteristic shape shown in Figure 5.11.

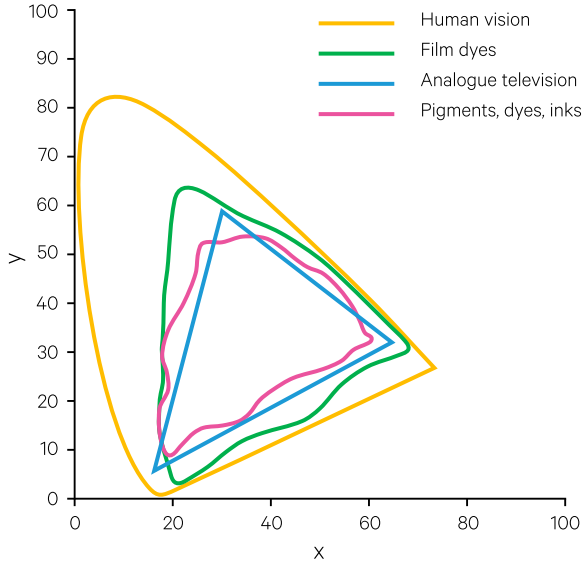
Colour is my day-long obsession, joy and torment (Claude Monet)

Figure 5.11 CIE chromaticity diagram

a. This two-dimensional representation of an intensity plane shows the ranges of colour hues that are visible to the human eye. The central white point represents the achromatic point at a given intensity level (see Appendix 3).



b. The ranges of visible colour hues are compared with those that can be reproduced by different media.



Source: a. I, Sakurambo. Retrieved from: *https://commons.wikimedia.org/wiki/File:CIExy1931.svg*. b. Harrison and Jupp (1990) Figure 13 [Adapted from Baldwin (1984) and Foley and van Dam (1984)]

This colour space is used to quantitatively measure colour differences but does not represent perceived colour differences linearly as they are not related linearly to the physical measurement of brightness. The interior and boundary of the curved region represents the visible range of colours. The pure colours form the boundary of this shape. The colour of the standard light source or illuminant, which is intended to represent sunlight, is located at the achromatic point, or white, where x, y and z have equal intensities of 1/3. (This point is defined in colorimetry as the colour of a blackbody radiator at 6504K; see Volume 1A—Section 2.11.) Again, saturation can be represented as the distance away from the achromatic point and hue is the angle of this direction.

The CIE colour space allows us to determine the purity of colours, that is, the distance of a colour from the standard white light source or illuminant (which varies along a locus with temperature) relative to its distance from the dominant wavelength at the edge of the visible region. Complementary colours can also be determined as those that lie on opposite sides of a straight line through the achromatic point. The diagonal is also used to define colour gamuts. A mixture of any two colours can produce the colours represented by a straight line between them. A mixture of any three colours produces the gamut, or range, of colours contained in the triangle bounded by the connecting lines between the colours.

The CIE chromaticity coordinates can also be converted to an approximately uniform colour space using dimensions of lightness (L*), redness-greenness (u*) and blueness-yellowness (v*) as defined in Appendix 3. The perceptual difference between two colour stimuli can then be approximated as the Euclidean distance between the points representing them in this uniform space. Other specific colour systems derived from these general models are described in Appendix 3.

5.1.3 Colour reproduction

Section 5.1.2 discussed a number of systems that have been devised to describe the range of humanly perceptible colours. However, the various display or hardcopy devices that are available for image processing can only represent a subset of these colours. The RGB technologies used to achieve a mixture of colours in different display screens are chosen to produce the widest possible range, or gamut, of colours (see Section 6).

Digital display devices typically span a greater range of colours than those achievable using colour films or inks (Baldwin, 1984). However, the use of the x' primary in colour matching (see Figure 5.10) indicates that there are some colours which cannot be readily created by adding lights. Colours outside the gamut triangle could only be produced by adding a 'negative' component of one of the primaries (see Appendix 3) so are represented by a colour on the edge of the triangle. For analogue colour television, for example, this means that the high purity blue and green colours are most restricted.

In display devices, brightness and contrast controls, as well as basic chromaticity differences between the RGB primaries used in display technologies, affect both the range of colours that can be represented and the degree of visual discrimination between them. The gamut of colours produced by hardcopy devices is affected by properties of the inks, tones and hardcopy media (paper or film) that are used.

The small size of the print or film gamut relative to display monitors means that the colours of displayed imagery cannot be fully reproduced in hardcopy. If an exact reproduction is required, a reduced gamut of colours must be used on the display device.

5.2 Colour enhancement

5.2.1 Pseudo-colouring

For a single channel of data, pseudo-colouring techniques can be used to display the image values as different colours rather than different grey shades (see Figure 4.10f). These techniques use lookup tables (LUT) to assign a colour to each image value. This operation is illustrated in Figure 5.12 for the image shown in Figure 4.1.

Figure 5.12 Pseudo-colouring of image data

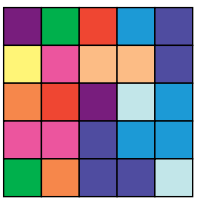
Spatial data can be represented pictorially by associating each data value with a particular colour. The selected colours may not be related to the data values but allow better visual discrimination of spatial patterns in the data than grey shade display techniques (see Figure 4.1).

a. Image pixel values

1	5	8	3	2
6	9	0	0	2
7	8	1	4	3
9	9	2	3	3
5	7	2	2	4

b. Grey scale lookup table

c. Grey scale image



Source: Harrison and Jupp (1990) Figure 14

The pseudo-colour table may be defined so that each image value has a discrete colour (see Figure 4.10f) or blocks of image values can be assigned the same colour using density slicing (see Section 9.2.1). This enhancement technique allows small changes in image values to be clearly highlighted as different colours. The pseudo-colour enhancement is most effective if applied to image data that has been contrast-stretched so that the colour differences apply over the bulk of the image data. Pseudo-colouring is also useful for interpretation and presentation of transformation results where these are recorded in a single image channel (see Volume 2C). Suggestions for selecting colours are provided in Excursus 5.2 and Excursus 9.2.

Excursus 5.2—Colour Composition

The selection of distinctive and effective compositions of colours involves an understanding of colour interactions in terms of contrast, harmony and rhythm. Both the quantity and surroundings of a colour will change its perceived hue, value and chroma.

Colour contrast can be defined in terms of perceived temperature with red, orange and yellow being considered warm hues and blues and greens as cool. The warm colours are generally more apparent in an image, especially against a neutral (grey) background, whereas cool colours are less 'active'. When colour is being used to represent a physical scale, such as temperature or yield, it should be remembered that dark colours are instinctively associated with low values and bright colours are often assumed to have high values. Strong, controlled variations in value and chroma, such as topographic shading, are usually interpreted as spatial differences in an image.

Harmony describes the sense of 'relatedness' among colours. A sequence of colours that has been selected from a natural progression in a colour model, such as a constant hue with changing value, renders an immediately perceived relationship between the image features portrayed in these colours. Traditionally, multiple hue harmonies have been established by selecting adjacent or close colours on a hue circle, or by pairing opposite or complementary colours (Greenberg *et al.*, 1982).

The final visual rhythm of an image will depend not only on the colours that have been selected but also on the relative amounts of each colour and their spatial arrangement. In image processing, this final balancing is only achievable by visual assessment of a 'painted' image.

Selection of colours should also consider our perceptive sensitivities as briefly mentioned in Appendix 3.1. Blue hues are generally not well perceived especially for thin features. The perceived colour hue of a 'patch' changes with its background and extent, with warm colours generally being too dominant for large patches. Blue on a black background, or its complementary mixture of yellow on a white background, are difficult to distinguish. Background colours that are mixtures of several spectral hues (such as brown) tend to mask fine detail. Simultaneous contrast, that is, the occurrence of one colour on two different and 'opposite' colour backgrounds, typically occurs in thematic maps and image classifications. However, such contrasting combinations reduce the discrimination between different hues so that nearly identical colours can be misread due to the immediate colour contexts in which they appear.

Colour value is a very useful attribute for distinguishing between colours but is easily forgotten so is not well suited for colours keys and is best limited to about five steps. The difference between measured and perceived intensity or value is most relevant when varying the value of a colour hue, with changes in low intensity values being less obvious than those at high intensities (see Excursus 5.1). A sub-conscious association is usually made between the perceived brightness difference of colours and the relationship between the values they represent. The value of background colours should not be too high as these will tend to draw attention away from the feature colours. Different hues also appear to have different intensities even in pure colours; for example, yellow always appears less saturated than other hues (see Figure 5.1).

The perceived chroma of a colour tends to increase with patch size so that selecting a colour from a small sample area may be misleading, especially for background colours. This can also lead to confusion between two similar hues or two values of the same hue that have important differences of chroma, since large areas of a lower chroma colour may appear similar or identical to small areas of the higher chroma colour.

5.2.2 Colour composite imagery

Multi-channel image data provide another opportunity for colour image display. In the same way that the human eye perceives colours as the proportions of the blue, green and red wavelengths it detects (see Excursus 5.1), colour monitors use blue, green and red colour primaries to create a full gamut, or range, of colours on a display screen. A colour composite image is formed by associating three different image channels with these three colour primaries: one channel is displayed as a blue scale, another as a green scale and a third as a red scale (see Figure 5.13). The display device then effectively registers these three colour scales as though they were three transparencies being viewed in front of a uniform light source. The resulting composite colours of a pixel depend on its shades of red, green and blue in the individual channels. Usually each channel of a colour composite image is individually contrast-enhanced to improve contrast in the displayed image.

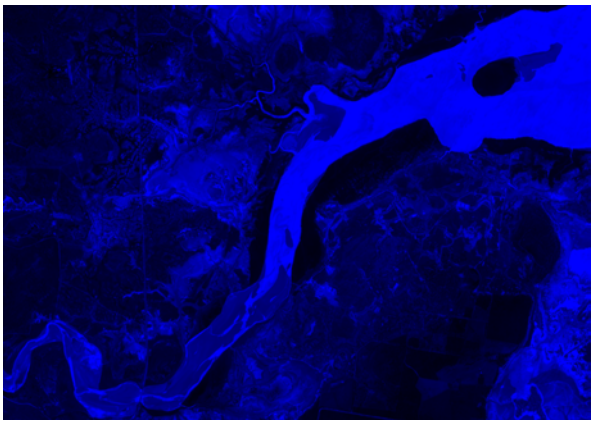
Figure 5.13 Formation of a colour composite image

Three image channels can be viewed simultaneously by associating each channel with a different primary colour in a colour composite image. Additive colour devices, such as colour monitors, use the primary colours blue, green and red while subtractive colour devices, such as plotters and photographic media, combine their inverse primaries yellow, magenta and cyan. This example uses channels 3 (visible green), 8 (NIR) and 11 (SWIR) from a Sentinel-2 image over the coastal town of St Lawrence, in central Queensland, that was acquired on 28 August 2015.

a. In the additive colour system (used by devices such as colour monitors), one channel is displayed as a blue scale, a second channel is shown as a green scale, while the third is displayed as a red scale (see second column below). These primary colours are effectively treated as registered transparencies over a light source so add together to form the lighter composite colours in Figure 5.13c below.

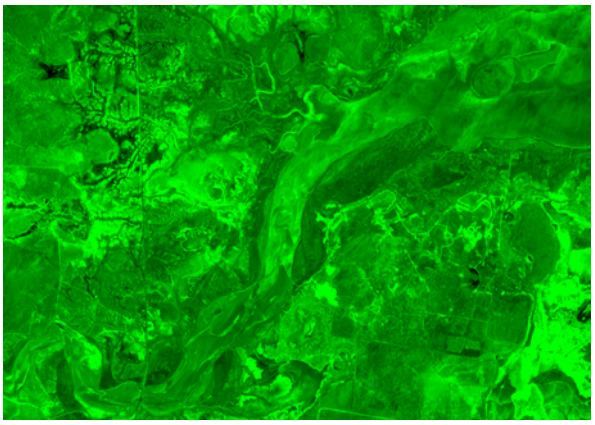
Band 3 grey scale

Represented as blue

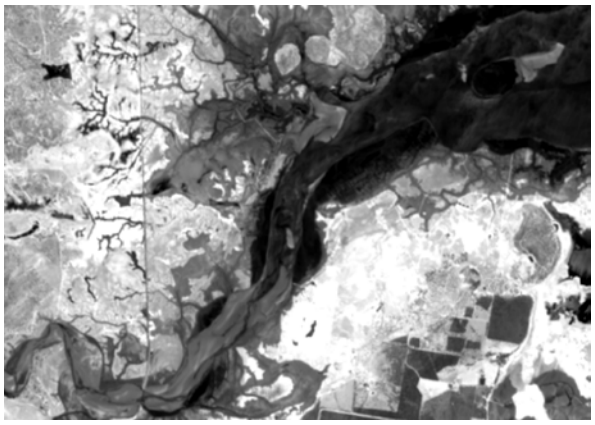


Band 8 grey scale

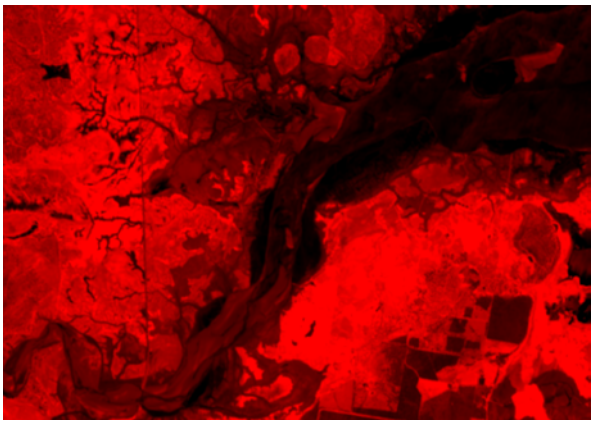
Represented as green



Band 11 grey scale



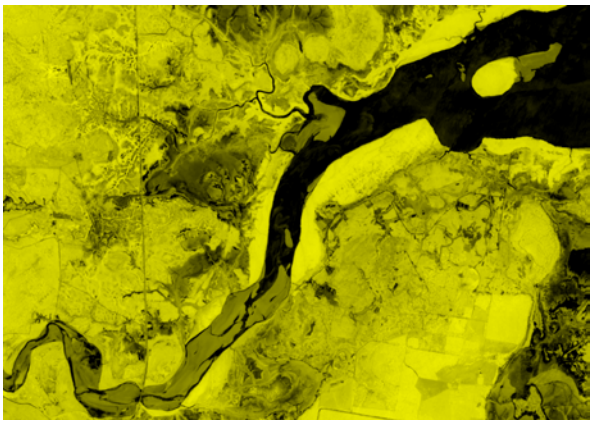
Represented as red



b. In the subtractive colour system (used by plotters and photographic media), the equivalent colour composite is formed from a yellow scale (inverse of the 'blue' channel), a magenta scale (inverse of the 'green' channel) and cyan scale (inverse of the 'red' channel) as shown in second column below. The inverse of a channel is commonly indicated as the negative channel number.

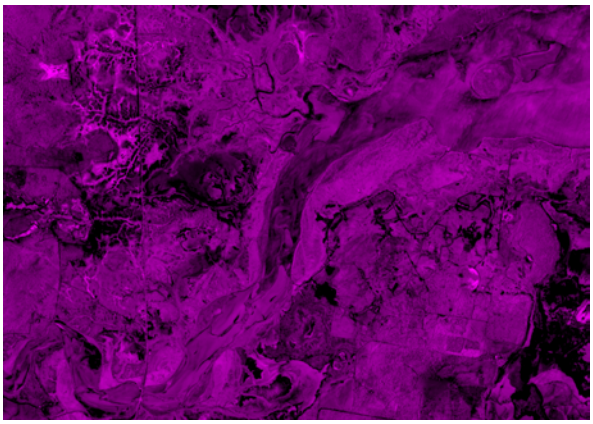
Inverse of band 3 as greyscale

Represented as yellow

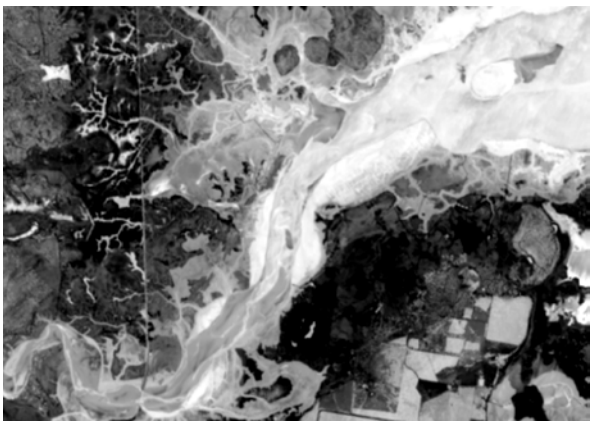


Inverse of band 8 as greyscale

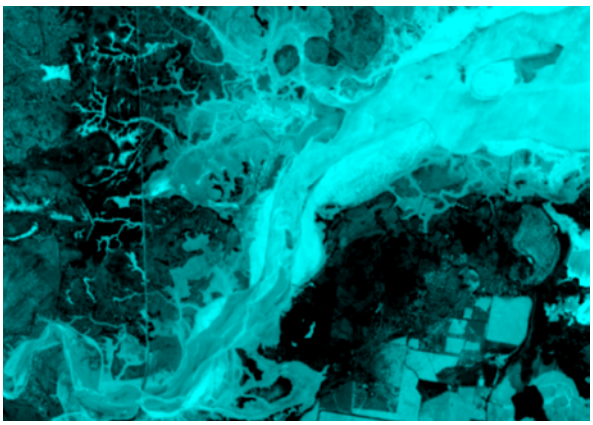
Represented as magenta



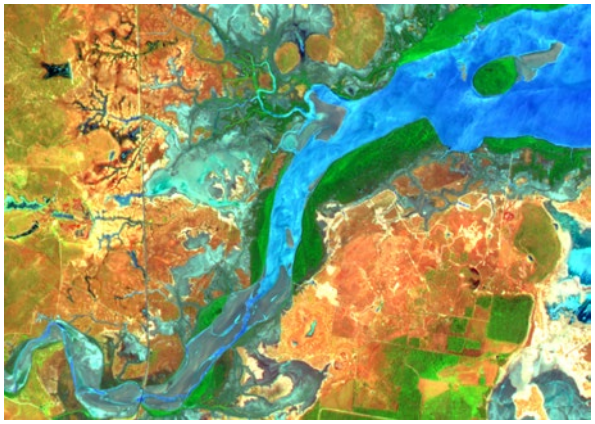
Inverse of band 11 as greyscale



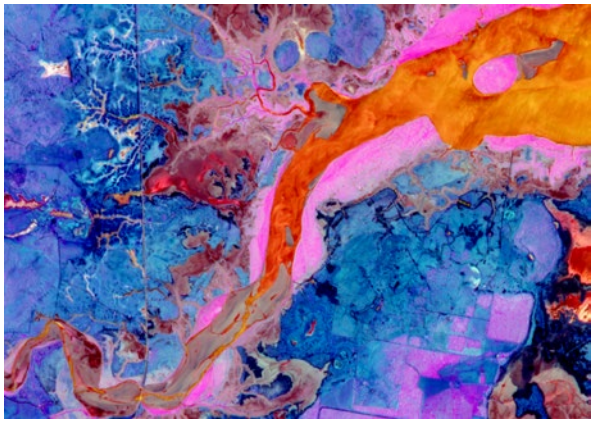
Represented as cyan



c. Colour composite formed as described in Figure 5.13a and Figure 5.13b.



d. An inverse colour composite is formed In the additive system by displaying the inverse of each channel (shown in first column of Figure 5.13b above) to the blue, green and red primaries. In the subtractive system, the same result is achieved by mapping the intensities of each channel (shown in the first column of Figure 5.13a) directly to the subtractive primaries: yellow, magenta and cyan.



Source: Norman Mueller, Geoscience Australia

A 'true', 'real' or 'natural' colour composite is formed by simultaneously displaying:

- a blue data channel (that is, recorded reflectances were in the wavelengths of blue light) as a 'blue scale' (with shades of blue, instead of grey, representing the channel values);
- a green data channel as a green scale; and
- a red channel as a red scale (see Figure 5.14d).

This should resemble the actual colours of the scanned object. With EO data, the colours of a 'true' colour composite image would be similar to those of the Earth's surface viewed from a high altitude aircraft (see Figure 5.15).

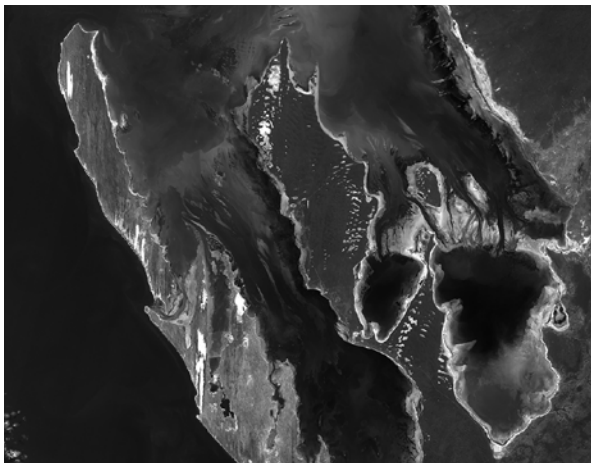
In nature, light creates the colour.
In the picture, colour creates the light.
(Hans Hofmann)

Figure 5.14 Shark Bay colour composite example

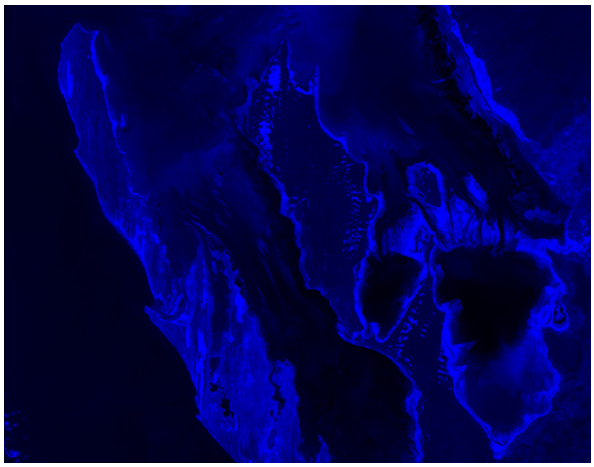
Three image channels can be viewed simultaneously on an additive colour device by associating each channel with a different additive primary colour (red, green and blue) in a colour composite image. These primary colours (a, b and c) are effectively treated as registered transparencies over a light source, so add together to form lighter composite colours (see Figure 5.14d overpage).

This example shows selected visible channels from a Landsat-8 image over Shark Bay, WA, acquired on 3 August 2017. Since blue wavelengths reflect from greater water depths than green or red wavelengths, most water features are coloured blue. Green wavelengths are also reflected from shallow water resulting in aqua colours for water adjacent to land. The red soils of this region have the highest reflectance in the red band. Cloud in the lower left corner of the image is bright in all visible channels so combines to appear white. These bands combine in a natural colour composite to produce feature colours that are comparable to an aerial photograph (see Figure 5.15 overpage).

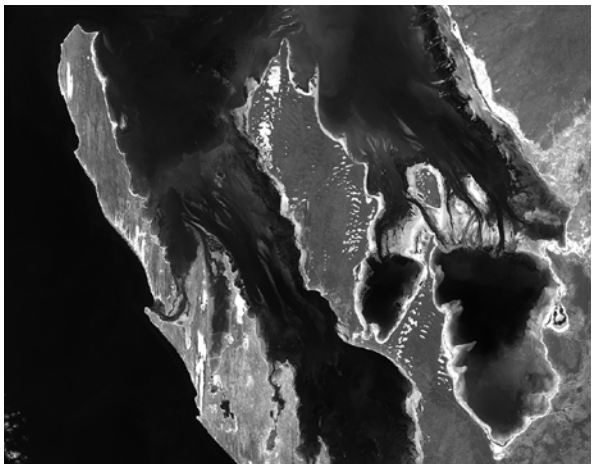
a. Greyscale of band 2 (visible blue)



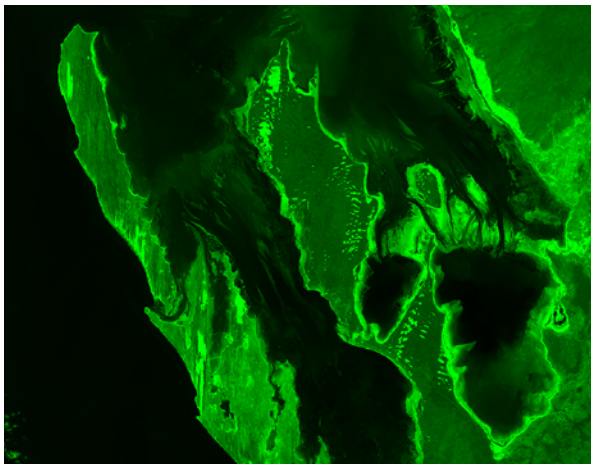
Represented as blue



b. Greyscale of band 3 (visible green)

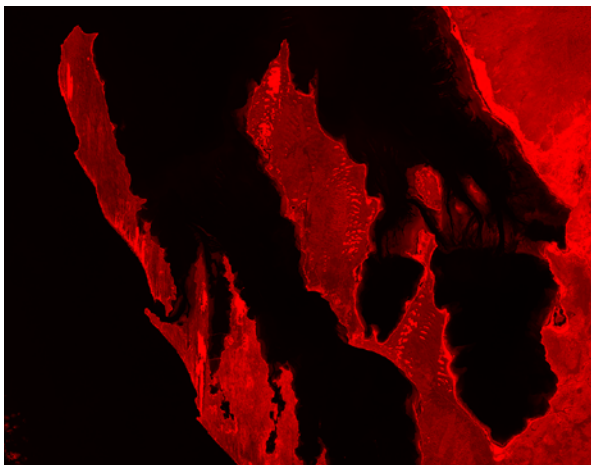


Represented as green

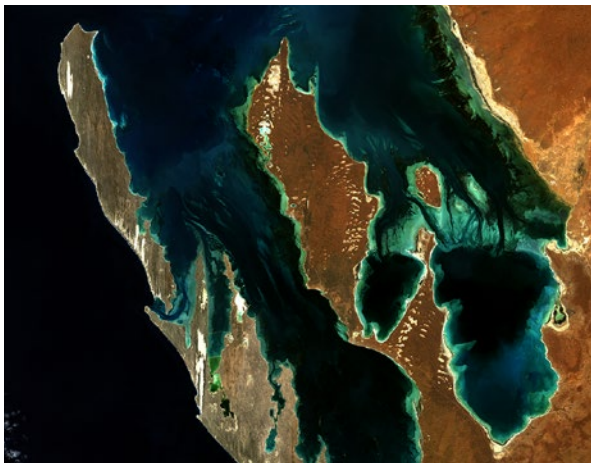


c. Greyscale of band 4 (visible red)

Represented as red



d. Colour composite of bands 2, 3 and 4 as RGB respectively



Source: Norman Mueller, Geoscience Australia

Figure 5.15 Aerial photograph of Shark Bay

This oblique aerial photograph over Shark Bay shows the true colours of ground features.

Source: Jason Mazur ©Mirage Digital. Used with permission.

Any other mapping of channels to the blue, green and red primaries is called a false colour composite since features in the resultant colour image do not have their true colours. Since visible bands tend to be highly correlated in EO imagery, including one or more non-visible bands in a colour composite increases the information content of the displayed image. This increased information content improves contrast between image features and thus enables better visual discrimination of those features.

Many EO sensors record multiple bands in visible and infrared wavelengths. For example, eight (of the 11) bands acquired by Landsat-8 are shown in Figure 5.16 over a large ephemeral lake, Lake Carnegie, in the arid centre of Western Australia. While some bands are more similar than others, each provides different information about the ground features. Different combinations of these bands can be used to create a range of colour composites (see Figure 5.17). When selected with knowledge of the reflectance characteristics of particular ground features, specific band combinations highlight similarities or differences between features of interest.

A standard false colour composite image for EO data displays a green data channel as blue, a red data channel as green, and a near infrared data channel as red as illustrated in Figure 5.17b. This combination of data channels produces a colour composite with similar colour representation to a colour infrared photograph with a characteristic set of colours for different land cover features:

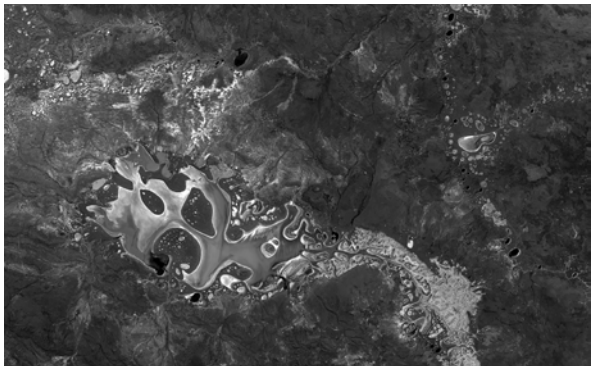
- vegetation appears red due to high reflectance in a near infrared reflectance channel, which is displayed as a red-scale;
- exposed red-brown soils appear green since they have highest reflectance in the red channel, which is displayed as a green-scale;
- water appears blue since its highest reflectance is in the green channel, which is displayed as a blue-scale; and
- bare areas appear white, since they have high reflectance in all of these channels.

Another informative colour composite displays SWIR as red, near infrared as green and visible green or visible blue as blue; see Figure 5.17c). For comparison, a colour composite of Lake Carnegie using Sentinel-2 image bands 11 (SWIR), 5 (red edge) and 3 (green) as RGB is shown in Figure 5.18. With imagery containing thermal data, an attractive and useful colour composite is formed by assigning TIR to red, NIR to green, and a visible channel to blue (see Figure 5.17d).

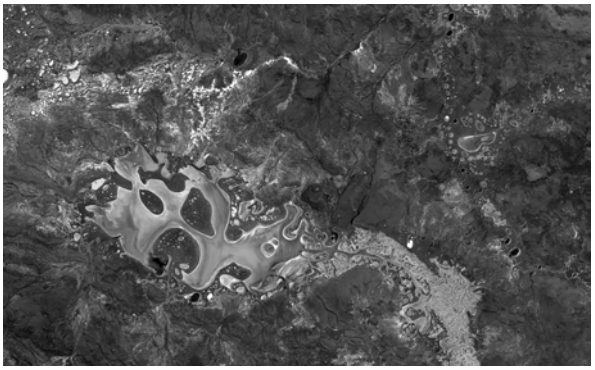
Figure 5.16 Example image channels for colour composites

Individual channels from Landsat-8 image over Lake Carnegie, WA, acquired on 8 July 2017.

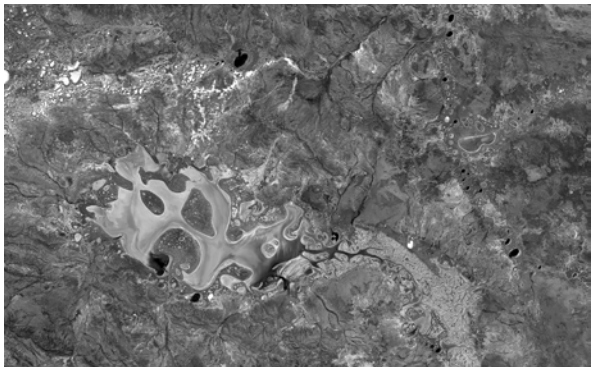
a. Blue channel (band 2: 0.452–0.512 μm)



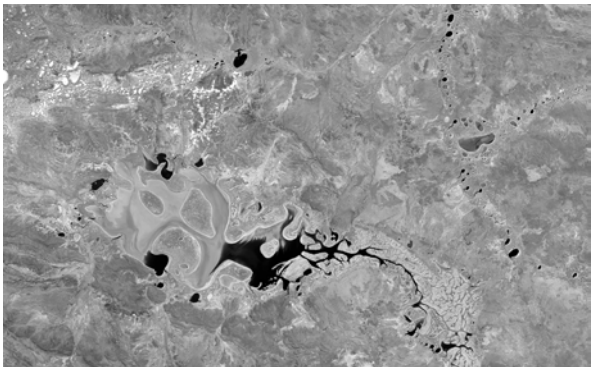
b. Green channel (band 3: 0.533–0.590 μm)



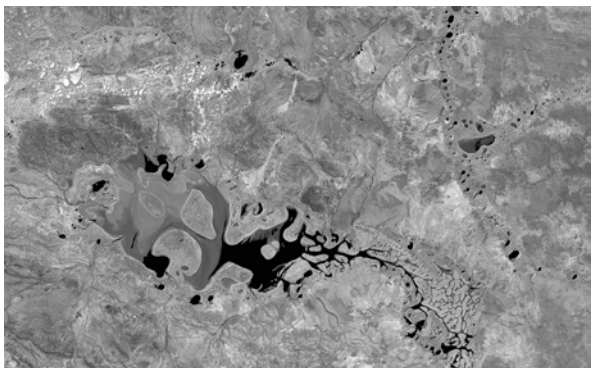
c. Red channel (band 4: 0.636–0.673 μm)



d. NIR channel (band 5: 0.851–0.879 μm)

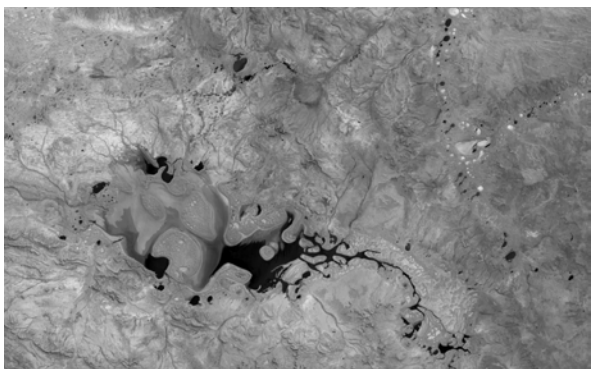


e. SWIR1 channel (band 6: 1.566–1.651 μm)

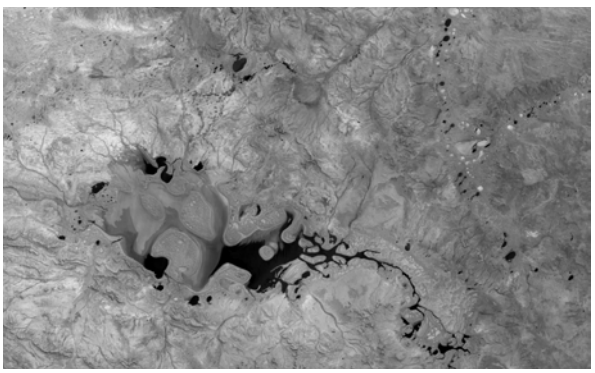


f. SWIR2 channel (band 7: 2.107–2.294 μm)

g. TIR1 channel (band 10: 10.60–11.19 μm)



h. TIR2 channel (band 11: 11.5–12.51 μm)



Source: Norman Mueller, Geoscience Australia

Figure 5.17 Standard colour composite images

These colour composite images were formed from selected triplets of the Landsat-8 channels shown in Figure 5.16.

a. A natural colour composite has colours that are similar to the colours of the original image scene since:

- a channel of red radiance values (OLI band 4) displays as red;
- a channel of green radiance values (OLI band 3) displays as a green scale; and
- a channel of blue radiance values (OLI band 2) displays as a blue scale.

As the name implies, natural colour composites represent the true colours of ground features. In this image, vegetation is sparse so land areas are dominated by the red soil background. Water in the salt lake varies in colour depending on sediment levels, with muddy water appearing green or brown.

b. A standard false colour composite for EO data associates:

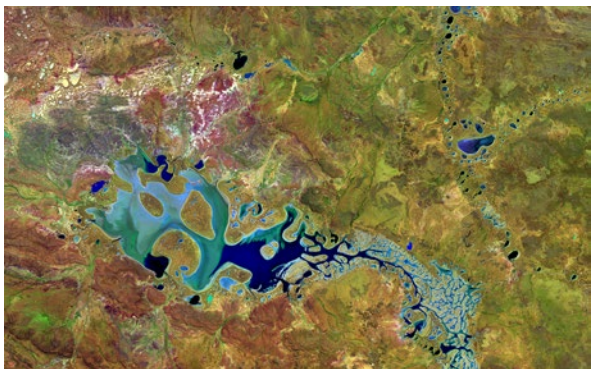
- a NIR channel (OLI band 5) with a red scale;
- a red channel (OLI band 4) with a green scale; and
- a green channel (OLI band 3) with a blue scale.

The colours of features in the resulting colour composite are similar to those of a colour infrared photograph. For example, vegetation appears as shades of red, deep clear water appears dark blue, shallow or muddy water is light blue and bare red soil appears green.

c. A SWIR false colour composites can be formed from a wide range of channel/colour combinations. In this example:

- SWIR 1 channel (OLI band 6) is displayed as a red scale;
- NIR channel (OLI band 5) is displayed as a green scale; and
- visible green channel (OLI band 3) is displayed as a blue scale.

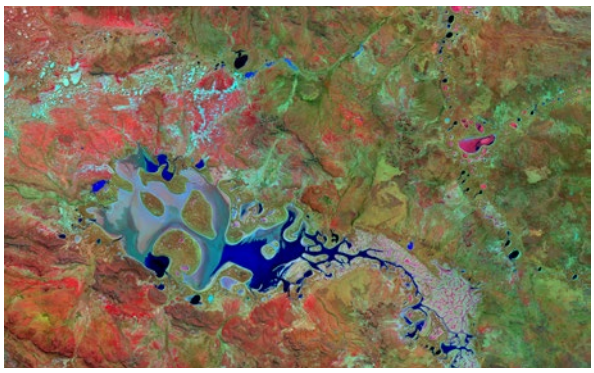
This colour composite renders 'intuitive' colours for ground features, with barer soils showing as red (due to higher SWIR reflectance), vegetation as green (due to high NIR reflectance) and muddy water as blue (due to high green reflectance). Compared with the natural colour composite (see Figure 5.17a above), this image highlights the differences in vegetation to the north and south of the lake.



d. In this thermal false colour composite:

- TIR 1 channel (OLI band 10) is displayed as a red scale;
- NIR channel (OLI band 5) is displayed as a green scale; and
- visible green channel (OLI band 3) is displayed as a blue scale.

Since this colour composite displays a thermal channel as red, hot bare, land areas are clearly visible as red tones and 'hot' lakes appear pink. Vegetation shows as green (due to high NIR reflectance) and most lakes appear as shades of blue (due to high green reflectance).



Source: Norman Mueller, Geoscience Australia

Figure 5.18 Thermal colour composite using Sentinel-2 imagery

This Sentinel-2 image over Lake Carnegie, WA, was acquired on 12 March 2017 and is displayed using bands 11, 5, 3 as RGB. These bands render a thermal colour composite that is comparable to Figure 5.17d.

Source: Norman Mueller, Geoscience Australia

Different combinations of bands can be used to enhance certain features by forcing them to appear as a more prominent colour. Non-standard colour composites may be formed by displaying different combinations of image channels as different primary colours, as illustrated in Figure 5.19. These composites can be valuable to highlight particular features or emphasise the differences between a pair of similar image bands. Since the various Landsat sensors have differing channel numbers, a summary of equivalent bands is included for reference in Excursus 5.3.

A number of statistical methods for selecting the most appropriate image bands for colour composites on the basis of their correlation have been proposed. For example, Sheffield (1985) suggested an algorithm based on image covariance statistics for selecting band combinations for colour composites from imagery with more than three channels. Another approach, the 'Optimum Index Factor' (OIF; Chavez *et al.*, 1982), selects the triplet of bands with the maximum sum of standard deviations and the minimum correlation between pairs of bands (see Section 8.1). Other methods for selecting sets of bands with minimum correlation are discussed in Volume 2C (in the context of image transformations) and Volume 2E (in the context of image classification).

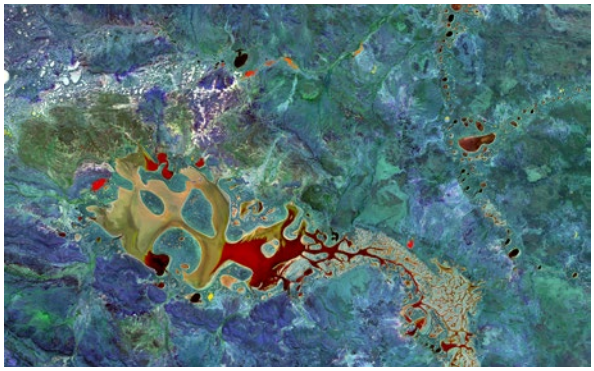
Figure 5.19 Non-standard colour composite images

These unusual colour composite images were formed from selected triplets of the Landsat-8 channels shown in Figure 5.16.

a. In this non-standard colour composite:

- green channel (OLI band 3) is displayed as a red scale;
- NIR channel (OLI band 5) is displayed as a green scale; and
- SWIR 1 channel (OLI band 6) is displayed as a blue scale.

This composite image uses the same triplet of bands as the SWIR false colour composite shown in Figure 5.17c, but displays the bands as non-intuitive colours so that vegetation appears as blue shades and water appears as red shades. Contrasting colours can enhance discrimination of some features.



b. In this non-standard colour composite:

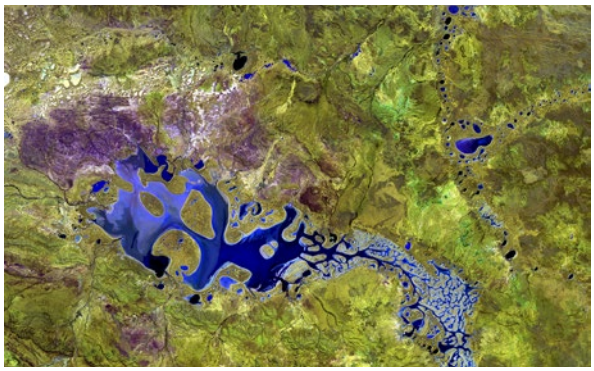
- NIR channel (OLI band 5) is displayed as a red scale;
- SWIR 1 channel (OLI band 6) is displayed as a green scale; and
- a blue channel (OLI band 2) is displayed as a blue scale.

This combination of bands can be used to highlight land versus water features.

c. In this non-standard colour composite:

- SWIR 2 channel (OLI band 7) is displayed as a red scale;
- SWIR 1 channel (OLI band 6) is displayed as a green scale; and
- blue (OLI band 2) is displayed as a blue scale.

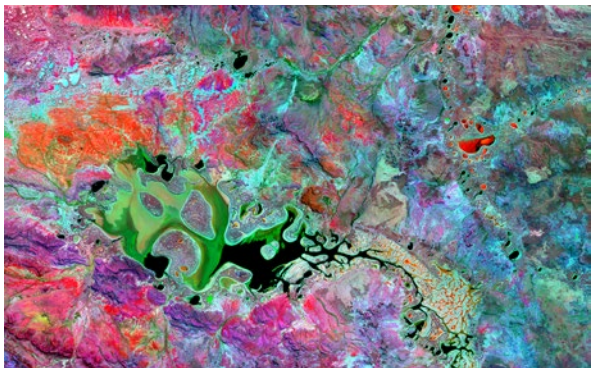
By using both SWIR bands, this composite highlights the differences in their information content. Since the SWIR bands are displayed as red and green, a range of yellow hues show changes in SWIR reflectances, possibly due to changes in mineralogy. Water clearly contrasts with land by appearing as blue tones.



d. In this non-standard colour composite:

- TIR 1 channel (OLI band 10) is displayed as a red scale;
- NIR channel (OLI band 5) is displayed as a green scale; and
- SWIR 1 (OLI band 6) is displayed as a blue scale.

This vibrant combination of bands differentiates between ground features on the basis of their reflectance properties in different infrared wavelengths.



Source: Norman Mueller, Geoscience Australia

Excursus 5.3—Landsat band numbers

Source: Norman Mueller, Geoscience Australia

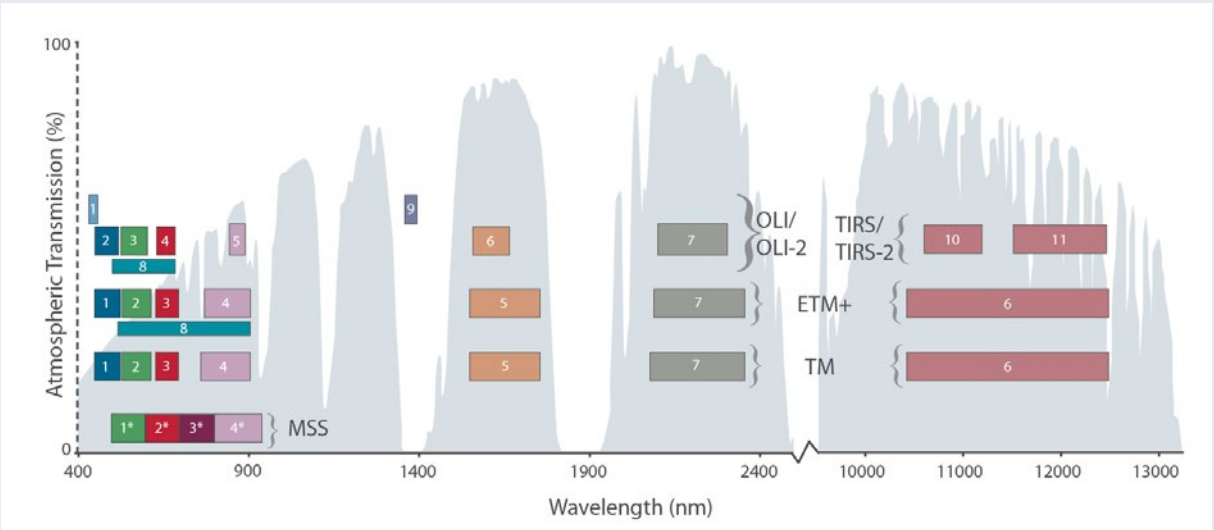
The Landsat-8 OLI sensor has additional spectral bands compared to Landsat-4/5 TM and Landsat-7 ETM+. A coastal blue band was added as Band 1 for Landsat-8 OLI, resulting in a change to the band numbering order. The spectral bands used in various Landsat sensors are shown in Figure 5.20.

Table 5.1 summarises the equivalence of spectral band numbers in these sensors. This summary should be useful for producing comparable colour composites from imagery acquired by different Landsat sensors.

Table 5.1 Equivalent Landsat band numbers

Spectral band	Sensor		
	Landsat-4/5 TM	Landsat-7 ETM+	Landsat-8 OLI
Coastal blue	-	-	1
Visible blue	1	1	2
Viaible green	2	2	3
Visible red	3	3	4
Near infrared	4	4	5
Shortwave infrared 1	5	5	6
Shortwave infrared 2	7	7	7
Thermal infrared 1	6	6	10
Thermal infrared 2	-	-	11
Panchromatic	-	8	8
Cirrus (1.363–1.384 mm)	-	-	9

Figure 5.20 Spectral bands in Landsat sensors



Source: NASA. Retrieved from: *https://landsat.gsfc.nasa.gov/wp-content/uploads/2016/10/all_Landsat_bands.png*

5.3 Further Information

Colour Society of Australia: *http://www.coloursociety.org.au*

Optimum Index Factor: *http://spatial-analyst.net/ILWIS/htm/ilwisapp/optimum_index_factor_functionality_algorithm.htm*

Gonzalez and Woods (2017) Chapter 7

Hunt (1987, 2004)

Levine (1985)

5.4 References

Baldwin, L. (1984). Color Considerations. *BYTE Magazine*, 227-249.

Chavez, P. S., Berlin, G. L., and Sowers, L. B. (1982). Statistical method for selecting Landsat MSS ratios. *Journal of Applied Photographic Engineering*, 8, pp. 23–30.

Cornsweet, T. N. (1970). *Visual Perception*. Academic Press, New York.

Foley, J. D., and van Dam, A. (1984). *Fundamentals of Interactive Computer Graphics*. Addison-Wesley.

Gondalez and Woods (2017). *Digital Image Processing*. Pearson Education, Inc.

Harrison, B. A., and Jupp, D. L. B. (1989). *Introduction to Remotely Sensed Data. Part ONE of the microBRIAN Resource Manual* (156 pages). CSIRO Australia, Melbourne.

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Hunt, R.W.G. (1987). *Measuring Colour*. Ellis Horwood Ltd. .

Hunt R.W.G (2004). *The Reproduction of Colour*. 6th edn. John Wiley and Sons, West Sussex.

Land, E. (1977). The retinex theory of colour vision. *Scientific American*, 237(6), pp. 108-128.

Levine, M.D. (1985) *Vision in Man and Machine*. McGraw-Hill, Inc. 574 pp.

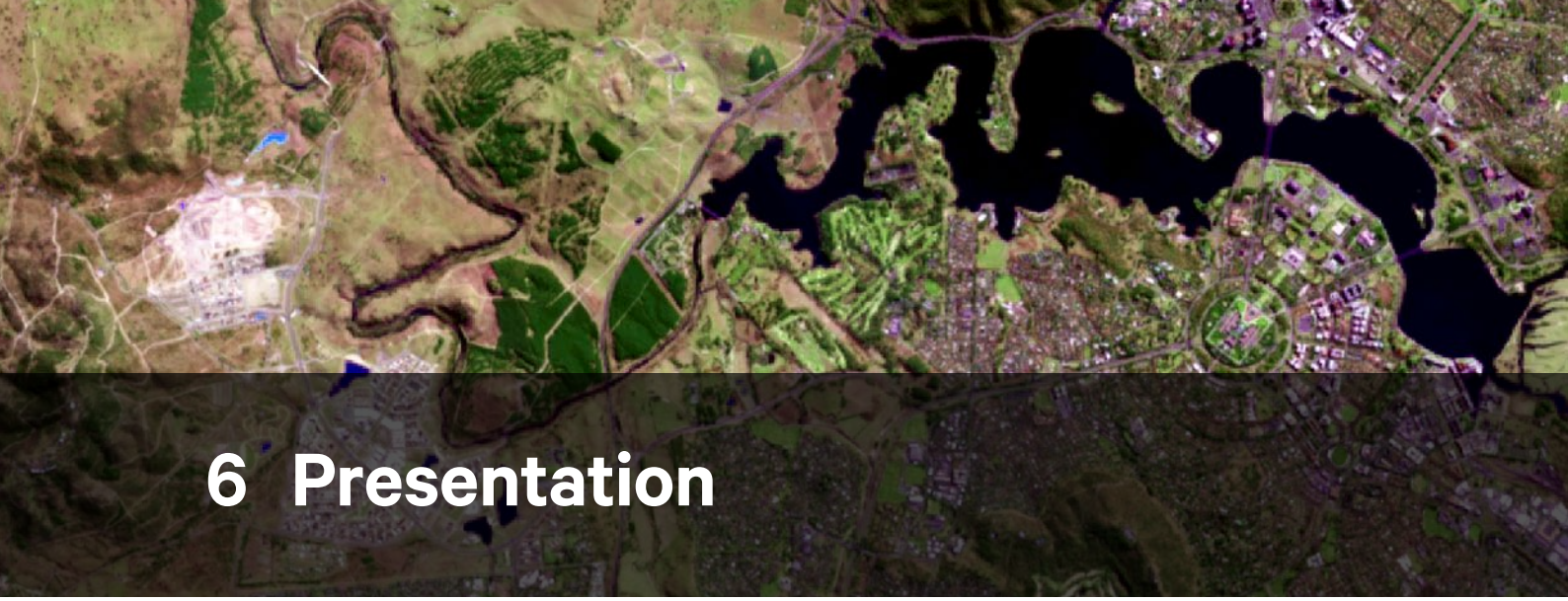
Munsell, A. H. (1929). *Munsell Book of Color*. Munsell Colour Co., Inc.

Niblack, W. (1986). *An Introduction to Digital Image Processing*. Prentice-Hall International, New Jersey.

Padgham, C. A., and Saunders, J. E. (1975). *The Perception of Light and Colour*. Academic Press, New York.

Sheffield, C. (1985). Selecting band combinations from multispectral d. *Photogrammetric Engineering and Remote Sensing*, 51(6), pp. 681–687.

Stone, M. C., Cowan, W. B., and Beatty, J. C. (1988). Color gamut mapping and the printing of digital color images. *ACM Transactions on Graphics*, 7(4), pp. 249-292. doi:*http://dx.doi.org/10.1145/46165.48045*.



6 Presentation

The techniques commonly used to represent spatial data as a picture are discussed in Sections 4 and 5, while mechanisms to alter image geometry are presented in Section 7. This Section considers options for viewing imagery, namely:

- image display (see Section 6.1); and
- image hardcopy (see Section 6.2).

6.1 Image Display

EO images are typically displayed on computer screens, either for visual interpretation, interactive enhancements and interrogations, or assessment of other processing procedures. This section introduces common features of image processing systems that are relevant to image display:

- using display hardware (see Section 6.1.1);
- displaying grey scale imagery (see Section 6.1.2); and
- displaying colour composite imagery (see Section 6.1.3).

Other commonly available display features include zooming and annotation of the displayed image.

6.1.1 Display Hardware

The technology behind computer display screens has changed significantly in recent decades from monitors using cathode ray tubes (CRT) to liquid crystal display (LCD) to light emitting diodes (LED). We will not venture into engineering aspects of these different technologies, but simply observe that this development path has resulted in display hardware with significantly higher resolution and flexibility for much lower cost.

For image display, any display screen is treated as a grid, which can be activated at discrete points to represent individual image pixels. All colour screens effectively involve activation of blue, green and red 'sub-pixels' to create colour images. These three primary colours are integrated by the eye and interpreted as a single colour (see Section 5). Specific hardware features of image display screens determine the number of colours that may be displayed on the screen at any one time.

EO imagery is generally represented on display screens in terms of four display modes:

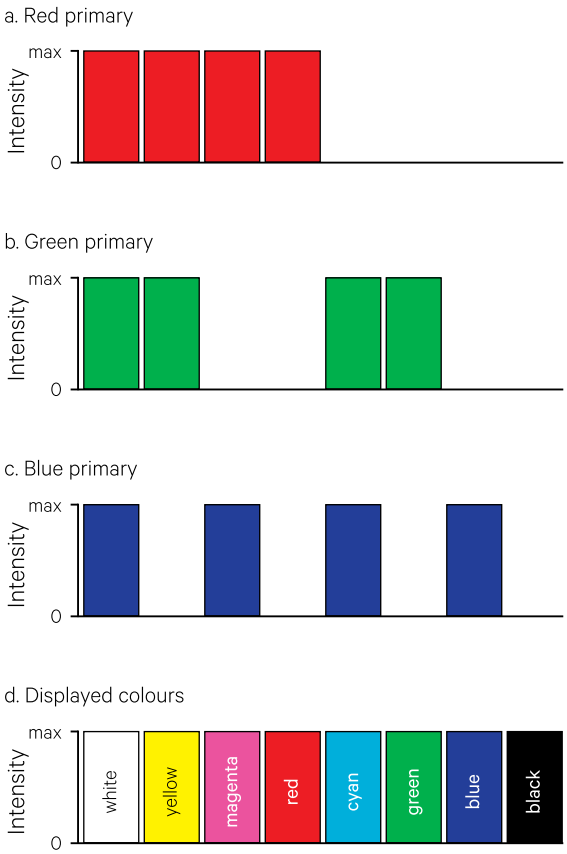
- bit-map—all pixels are presented as either black or white; this mode is used to define image masks (see Section 10);
- grey scale—represents a mapping of values in one image channel to shades of grey on the display screen (see Section 4);
- pseudo-colour—shows a grey scale image (one image channel) as range of colours rather than grey tones (see Section 5.2.1); and
- colour composite—operates by associating three image channels with the three additive primary colours (see Section 5).

Where the three channels being displayed in colour mode represent the same data values (that is, they are the same data channel), the proportions of blue, green and red at each screen pixel are identical, so add together to give a shade of grey. Some simple examples of colour mixing are illustrated in Figure 6.1.

Background image: Colour composite of Landsat-8 image of Canberra, ACT, acquired on 10 August, 2017, with bands 6, 5, 3 displayed as RGB, and sharpened using band 8 (panchromatic). **Source:** Norman Mueller, Geoscience Australia

Figure 6.1 Additive colour mixing

A range of colours can be produced on the display monitor using the additive primary colours, red green and blue. This example only shows mixing of the additive primaries to create all additive and subtractive primary colours, plus black and white.



Source: Harrison and Jupp (1990) Figure 27 [Adapted from Richards (1986)]

An intermediate display operation can occur in most image processing systems whereby each image value is assigned a different output value or colour using a colour lookup table (LUT; see Section 4.2.4 and Volume 2C). Thus, by re-defining the values in the LUT, we can interactively change the colours of the displayed image without altering the image values. Various display programs in image processing software make extensive use of this colour LUT feature for interactive image enhancements, such as pseudo-colouring (see Sections 5 and 9).

To maintain a displayed image involves continually refreshing the display screen for every pixel of every line in the display. Early CRT monitors used screen phosphors with relatively low persistence (that is, they faded easily) so needed to be refreshed (or rewritten) frequently to avoid flickering. While the standard refresh rate for CRT screens was 30 frames per second, newer screen technologies have superior persistence so require less frequent refresh cycles.

6.1.2 Grey-scale Imagery

6.1.2.1 Contrast stretching

Minimum and maximum values for contrast stretching may be determined using:

- image histogram statistics:
  - over the whole image (see Section 4); or
  - interactive training on a particular feature (see Section 9.1.1.1); or
- interactive screen enhancements (see Section 6.1.2.2).

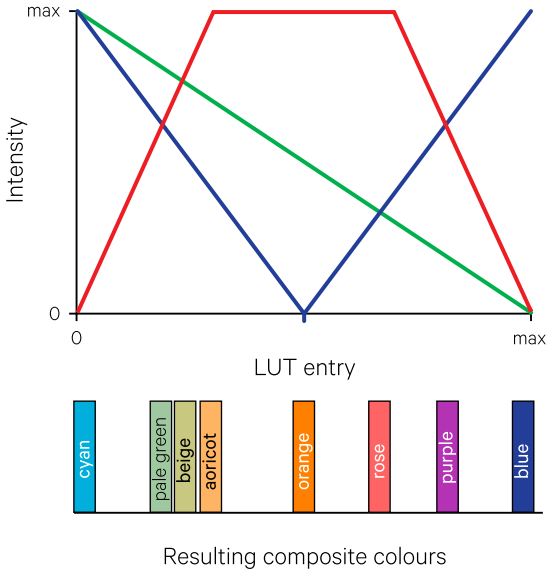
Interactive enhancements use the hardware LUT feature to modify the displayed grey shades of the image values in display memory. In some image processing systems, interactive contrast stretching also allows a non-linear stretch to be applied to the displayed image, for example via gamma function values. The interactive stretching operations are generally monitored using a grey scale bar showing current minimum and maximum values and/or a two-dimensional graph of image versus display values.

6.1.2.2 LUT manipulation

Lookup tables (LUT) allow pseudo-colouring of a grey scale image via a LUT editing facility. For example, a LUT editor could permit a 256 entry LUT to be defined graphically by specifying the intensities of the blue, green and red primaries for each entry value. An example LUT generated this way is shown in Figure 6.2.

Figure 6.2 Generating a colour LUT from RGB intensities

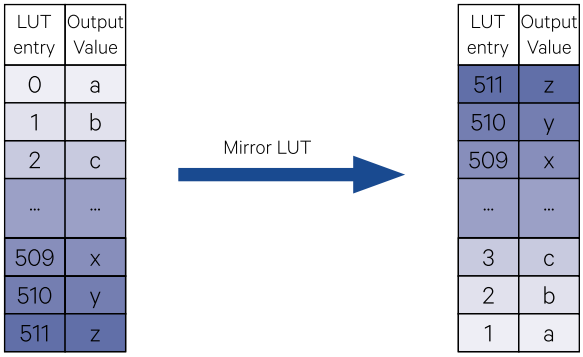
In the additive colour system, varying intensities of red, green and blue can be graphically defined for each LUT entry. These intensities are summed for each entry to produce a set of graduating colours.



Source: Harrison and Jupp (1990) Figure 30

Figure 6.3 Mirroring the LUT

The order of entries can be reversed to alter colour assignment in the displayed image. This type of operation can be applied to all or part of the LUT. Other operations such as rotation of entries can also be used to systematically alter the LUT.



Source: Harrison and Jupp (1990) Figure 31

Other manipulations which may be applied to LUTs include mirroring the full set of entries, that is, reversing the order of colour values between entries 0–511 to apply to entries in the order 511–0 as illustrated in Figure 6.3. The LUT entries may also be rotated in variable step sizes. These operations are most easily understood when applied to a standard colour table (see Section 9.2.1).

6.1.3 Colour Composites

Colour composite imagery may be formed by assigning any three image channels to the blue, green and red colour primaries (see Section 5.2.2). Most image processing systems allow the inverse of a channel (that is, showing low values as bright, and high values as dark) to be displayed (for both colour and grey scale images). The inverse of a channel is commonly indicated by specifying the negative of that channel (see Figure 5.13b in Section 5). One or two colour primaries may also be de-activated in a colour composite display (commonly by entering '0' rather than a channel number). Generally, menu-defined minimum and maximum values for appropriate channels of the image entry being displayed are automatically used unless local stretch values are active. In most image processing systems, a non-linear stretch can also be defined before displaying the image, such as by specifying an appropriate gamma function values for each channel.

Some image processing systems allow spatial data to be displayed as 'RGBI' composites, associating three image bands with the red, green and blue colour primaries and a fourth data source to be shown as intensity. This function is convenient for embedding other spatial information (such as topography) into a colour composite (see Volume 2C).

6.2 Hardcopy

Hardcopy of image data during or after processing is valuable for various stages of interpretation, field checking and final presentation. This usually involves writing the image as a colour or grey scale picture to a hardcopy device. To overlay a hardcopy image onto a map requires that the image data be appropriatedly scaled and often rectified (that is, geometrically modified) in some way before presentation (see Section 7 and Volume 2B).

The basis for colour representation in hardcopy devices is different from that used in display devices (see Sections 5 and 6.1, plus Appendix 3). Whereas a display monitor derives its full gamut, or range, of colours from the additive primaries (blue, green, and red), hardcopy devices use the subtractive primaries (yellow, magenta and cyan) to create a large range of colours as illustrated in Figure 5.13. Section 6.2.1 discusses some features of printing technology that can be used to modify the range of output colours, while Section 6.2.2 briefly introduces photographic output options.

6.2.1 Printer Output

Printers use a fine grid of discrete, coloured dots in the subtractive primaries to form imagery in much the same way display devices activate the additive primaries associated with screen pixels (see Sections 5 and 6.1. Various technologies have been developed to apply colour to a plotting medium (paper or transparent film). For example, to represent individual image pixels, ink jet printers 'squirt' fine drops of ink, whereas laser printers use laser technology to apply toner to the output media.

Most printers create their colour gamuts using consumables of the three subtractive primaries: yellow, magenta and cyan. As illustrated in Figure 5.7b and Figure 5.8b, by mixing combinations of these three colours in any printed dot we can create the eight different colours listed in Table 6.1. Due to the variable qualities of consumables and the printing mechanism used, mixing yellow, magenta and cyan often creates a dark brown colour, rather than true black, so most printers provide a black option in addition to the three primary colours. Occasionally other consumable colours are also used to enable a greater variety of output colours.

Table 6.1 Subtractive colour mixing

Ink Used in Dot			Resulting Dot Colour
Yellow	Magenta	Cyan	
off	off	off	White
on	off	off	Yellow
off	on	off	Magenta
off	off	on	Cyan
on	on	off	Red
on	off	on	Green
off	on	on	Blue
on	on	on	Black

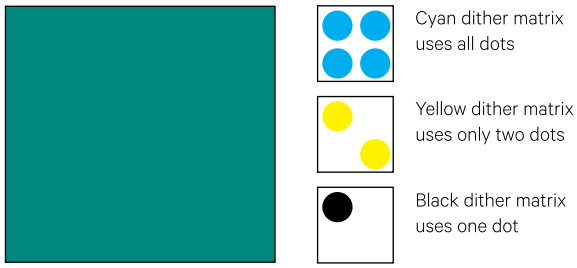
Printers are sometimes classified in terms of whether the amount of colour applied to each dot can be varied:

- binary or bilevel printers (such as inkjet printers and early laser printers) only have two levels of intensity for each primary colour—on or off—resulting in the eight colours listed in Table 6.1. To create a larger apparent range of colours, these printers use techniques such as dithering (see Figure 6.4). Dithering, however, necessarily reduces the effective resolution of the printer (see Appendix 4).
- multi-level printers (such as modern laser printers) have multiple levels of intensity for each primary colour so can directly print a greater range of colours—from 8 to 64 tones for each primary colour, depending on printer technology. These printers require less dithering, so can retain higher output resolution.
- continous tone printers (such as dye sublimation printers or film recorders) offer the largest range of intensities.

Continuous tone images (contone) comprise a seemingly infinite range of grey tones or colours and are only truly rendered by continuous tone printers. However, many laser printers now blend the applied primary colour within printer dots to create exact shades, rather than rely on the eye to integrate them. This blending of primaries within dots results in smoother colour gradations and reduced graininess. Other differences between printers relate to the variation that can be achieved in dot size, shape or placement, which primarily impact image sharpness.

Figure 6.4 Dithering

The dithering technique is implemented by grouping together blocks of dots (or nibs), then using the blended colour of each group of dots to represent a single pixel in the image. The size of a block can vary and affects both the number of colours that can be represented and the effective resolution of the printed output (see Appendix 4). In this example, a teal sample colour would be rendered using 100% cyan, 50% yellow, 25% black and 0% magenta.



Colour prints may be formed from EO imagery using pseudo-colouring or colour composites. As discussed in Section 5.2.2, to display a colour composite image using the additive primaries of blue, green and red, one image channel is represented as a blue scale, another as a green scale and a third as a red scale. To print the same image we need to map our image data to the subtractive primaries yellow, magenta and cyan. That is:

- the channel displayed as blue is printed as an inverse yellow scale (with low image values being plotted as high intensity and vice versa);
- the channel displayed as green is printed as an inverse magenta scale; and
- the channel displayed as red is printed as an inverse cyan scale.

The combinations of intensities in our additive primary colours can thus be approximated in subtractive colour media by representing their inverse intensities in the complementary subtractive primary. This inverse mapping is handled automatically in most image processing systems so colour composites only need to be specified in terms of the blue, green and red primaries for printing. To ensure differentiation of colours in a printed output, some image processing systems also allow multiple user-defined colour maps to be defined for particular printer configurations. A given image, for example, could have one LUT for printing to a laser printer and another LUT for printing to an inkjet printer.

A final consideration in printing imagery is the differences between the additive and subtractive colour spaces. The way colours are formed and perceived in these two systems is physically different so it is often difficult to print an image using exactly the same colours as appear on the display screen (see Section 5.1.3). Our perception of the brightness of an object is non-linearly related to the measured amount of light, or luminance, from it, with discrimination being greater between darker light levels (see Figure 5.3). Similarly, the lightness which is perceived as the mid-tone in a linear sequence of varying grey levels between black and white is darker than the real mid-tone (Drury, 1987). Other aspects of colour perception are introduced in Appendix 3.

The additive display system starts with a black screen and creates colours by adding coloured light to this black background (see Figure 5.13). Small changes in intensity in the dark shades are perceived more readily than in the light shades.

In subtractive colour systems, inks or dyes are added to a white background to produce various colours, that is they are 'subtracted' from white. On the white background, small additions of colour have more visual impact in the bright colours than in darker colours. Thus, while the two renditions of colour may be theoretically comparable, our perception makes them appear different. To compensate for this, a non-linear contrast stretch is usually employed when printing imagery (see Section 4.2.2).

The entertainment is in the presentation.
(John McTiernan)

6.2.2 Photographic Output

Photographic output of image data can be produced using colourwriting (or photowriting) techniques. This output can be in the form of positive or negative transparencies or prints. Film writers use a point of variable intensity to expose a small area of colour (or black and white) film. The size of the area exposed and the number of pixels and lines that can be written to a negative both depend on the selected aperture size. Image pixel intensity information is used with a film-response calibration function to determine the intensity of exposure at each point.

6.3 Further Information

Dithering: *http://webstyleguide.com/wsg2/graphics/dither.html*
https://www.tutorialspoint.com/dip/concept_of_dithering.htm

Hunt (2004)

6.4 References

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Hunt R.W.G (2004). *The Reproduction of Colour*. 6th edn. John Wiley and Sons, West Sussex.

Richards, J. A. (1986). *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag, Berlin, Germany.

7 Geometry

Image data have properties relating to both colour and geometry. EO imagery can display a variety of geometric distortions due to characteristics of the scanning mechanism used to acquire the data (see Volume 1A). These distortions are summarised in Section 3.2.2 above.

The basic concepts and simple manipulations relating to image geometry, which are relevant to a range of image processing operations, are introduced in Sections 7.1 and 7.2. Image rectification and registration may be required in remote sensing exercises to locate image features on maps, locate field sites in images, or create output imagery that are scaled to overlay a map or some other image (see Sections 7.3 and 7.4). In particular, image registration is essential for multi-temporal studies, where multiple images need to be geometrically registered for change detection analyses, or for integrating map data (such as cadastral boundaries) with an image (see Volume 2D). Mosaicking of multiple scenes for large area studies also requires accurate image rectification and registration.

The detailed methodologies required to rectify the geometry of EO images or register such imagery with other spatial data are described in Volume 2B. Registered or integrated image datasets are required for multi-date analyses (see Volume 2D.

7.1 Image Coordinate System

Figure 1.1 (see Section 1 above) represents an image as a n-dimensional table or matrix of numbers, with image lines forming rows in the table, pixels along a line forming columns in the table, and channels in the image corresponding to different data attributes or variables. Entries in the table can be identified by specifying their corresponding column and row numbers, counting from column 1 at the leftmost column and row 1 as the top row. Similarly, the location of a particular pixel in an image is traditionally recorded as the number of pixel columns or samples from the leftmost pixel and the number of lines from the top line as illustrated in Figure 7.1a.

This coordinate system differs from the traditional planar system used for two-dimensional graphs where all positive locations are above and to the right of a central origin (see Figure 7.1b). The image coordinate system effectively places the origin at the top-left corner of the image with the centre of the top-left pixel being location (1,1). Other locations are then counted to the right of, and down from, this origin.

In the context of image locations, a 'pixel' position corresponds to a column in the image matrix and a 'line' position corresponds to a row. Typically, both the size of an image, and locations within it, are given as a coordinate pair in the order '(*pixels across, lines down*)' in the same way as the planar system specifies '(*x, y*)'. The definition of the image coordinate system is further discussed in Volume 2B.

Image geometry may be altered by the simple operations of modifying image size and orientation (see Section 7.2) or the more complex procedures of rectification and registration (see Section 7.3, and Volume 2B).

Geometry is not true, it is advantageous.
(Henri Poincare)

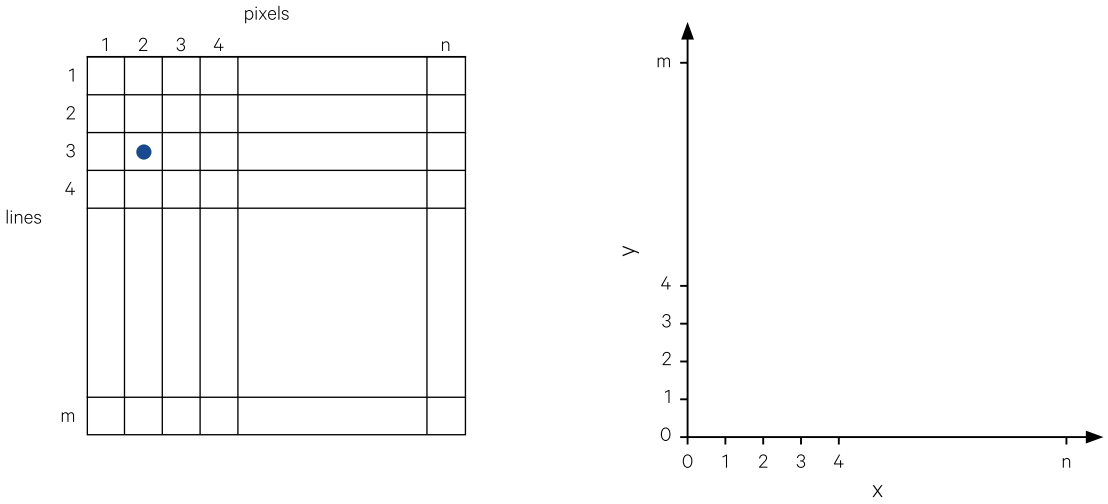
Background image: Colour composite of Landsat-8 image of Canberra, ACT, acquired on 10 August, 2017, displayed using bands 6, 5, 3 displayed as RGB, and sharpened using band 8 (panchromatic). This image has been overlaid with map information showing roads and water features.
Source: Norman Mueller, Geoscience Australia

Figure 7.1 Standard image coordinate system

A standard image coordinate system is compared with the traditional planar coordinate system.

a. Image coordinates are relative to an origin in the top left corner of the image. Location ● is in pixel 2 and line 3 of the image or '(2,3)'. The same coordinate referencing system is used for all channels in the image.

b. Most map coordinates are based on a planar system with the origin in the lower left corner of the map.



Source: Harrison and Jupp (1990) Figure 15

7.2 Shape and Orientation

7.2.1 Modifying image size

The size of image files may be altered in a number of ways:

- subsetting a large image into one or more smaller image(s) (see Section 7.2.1.1);
- joining several images into a larger image (see Section 7.2.1.2);
- blocking an image into a smaller image by:
  - averaging the values of several pixels in the input image to a single value in the output image or
  - sub-sampling the input image to represent a larger pixel (see Section 7.2.1.3); and
- enlarging the scale of an image by duplicating single pixel values in the input image to become multiple pixels in the output image (see Section 7.2.1.4).

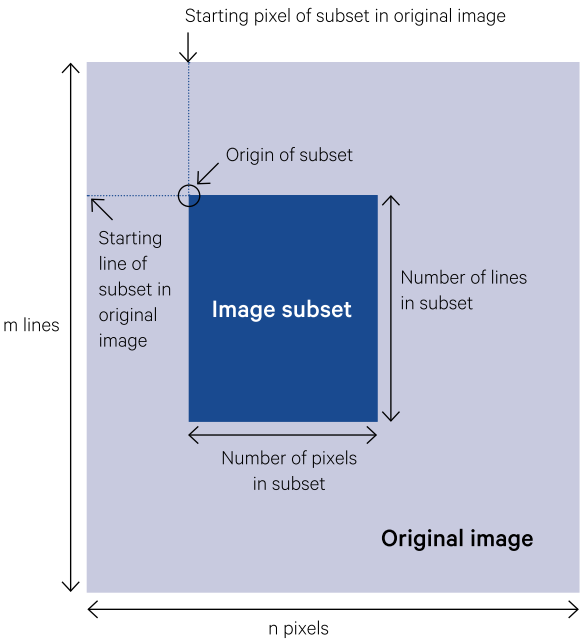
Modifying image scale potentially introduces artefacts into the output image. The impact of upscaling (creating larger pixels by blocking) and downscaling (creating smaller pixels by enlarging) EO imagery is considered in Volume 1B—Section 2.4.

7.2.1.1 Subsetting

Image subsetting simply involves extracting one or more portions of an image file to become separate image files or 'subsets'. This portion is usually specified as a starting location, that is the origin in terms of the original image coordinates (see Figure 7.1), and the size of the new image subset (the number of pixels and number of lines). This operation generally assumes that the number of channels is not modified between images, although the output image could contain fewer channels if required. Figure 7.2 illustrates the parameters required for image subsetting.

Figure 7.2 Subsetting an image

A sub-area within an original image can be defined as a given number of pixels and lines from an origin pixel location.



Source: Harrison and Jupp (1990) Figure 16

7.2.1.2 Sewing and mosaicking

A reciprocal action to image subsetting is image sewing or mosaicking where several smaller images are joined together to form a larger image. A mosaic can be viewed as a picture or design formed by cementing coloured objects onto a surface (Ahuja and Schachter, 1983). Multiple input images, possibly with different geometries, can be mosaicked onto a new grid, or base map, using image resampling (see Section 7.3). This process requires that the relationship between each input image and the base map be defined precisely, generally by applying a multi-stage modelling approach as detailed in Volume 2B.

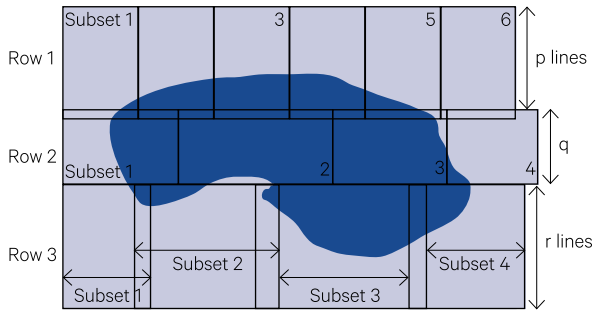
Image sewing usually refers to the operation of patching subsets together assuming that all subsets are compatible in terms of scale and orientation. This operation does not alter the shape or scale of the (sub)image in any way (see Figure 7.3). By contrast, image mosaicking allows the subsets that are being joined to be altered geometrically, in terms of shape, scale and orientation, before joining (see Section 7.3).

Image sewing simply requires that the origins of subsets being joined be specified relative to a consistent image coordinate system. The operation of image mosaicking can additionally use a range of sophisticated models to account for geometric differences between the image subsets being joined and/or modify the scale of the output image relative to the input images. This process is especially relevant to EO data where the along-line and along-track distortions can vary significantly over a complete image scene and hence differ between subsets of any given scene (see Section 7.3 and Volume 2B).

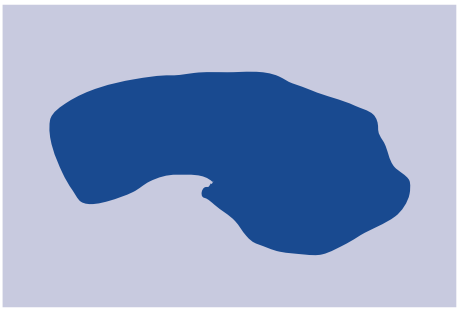
Figure 7.3 Joining multiple images

Within a row of subsets, the number of pixels in, and the number of pixels being copied from, each image can be different, but the same number of lines is copied from each input image. This ensures that areas of overlap in the input images are represented only once in the output image.

a. Input images



b. Output image



Source: Harrison and Jupp (1990) Figure 17

7.2.1.3 Reducing

Image size and scale may be reduced (without subsetting) by the operations of blocking or sub-sampling. Reducing image size is sometimes called image 'minification'.

Image blocking reduces image size by averaging the values of multiple pixels in the input image into a single pixel in the output image. The selected block size determines the number of pixels that are averaged. The effect of a 2 pixel by 3 line block on an input image of 10 pixels and 10 lines is shown in Figure 7.4 In this case, since the blocking factor is not evenly divisible into the number of lines in the input image, any partial blocks are not included in the output image. By default, blocking generally starts from the first pixel in an image.

Figure 7.4 Blocking an image

This example uses a single channel image of 10 pixels and 10 lines with a 2x3 blocking factor to produce an output image with 4 pixels and 3 lines.

a	b	5	6	7	6	6	5	4	3
c	d	4	5	6	6	5	5	5	4
e	f	3	7	6	7	6	5	6	7
3	4	4	5	5	7	7	6	6	7
4	4	4	4	5	6	6	7	7	7
5	5	4	5	6	6	5	6	6	5
5	4	4	4	5	6	7	8	4	3
4	3	3	4	5	7	9	6	2	1
4	2	3	8	7	9	9	7	3	1
2	3	3	9	9	8	7	6	2	2

where (a+b+c+d+e+f)/6 = A

A	5	6	5	5
4	4	6	6	6
4	4	7	8	2

Source: Harrison and Jupp (1990) Figure 18

Figure 7.5 Sub-sampling an image

a. An input image with 20 m pixel size is sub-sampled to simulate an output image with 40 m pixels by dropping alternate pixels and lines in the image

Input image pixel size = 20 m

a	b	c	d	e	f	g	h
i	j	k	l	m	n	o	p
q	r	s	t	u	v	w	x
y	z	A	B	C	D	E	F
G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V
W	X	Y	Z	1	2	3	4
5	6	7	8	9	10	11	12

Output image pixel size = 40 m

a	c	e	g
q	s	u	w
G	I	K	M
W	Y	1	3

b. An input image with 20 m pixel size is sub-sampled to simulate an output image with 30 m pixels.

Input image pixel size = 20 m

a	b	c	d	e	f	g	h
i	j	k	l	m	n	o	p
q	r	s	t	u	v	w	x
y	z	A	B	C	D	E	F
G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V
W	X	Y	Z	1	2	3	4
5	6	7	8	9	10	11	12

Output image pixel size = 30 m

a	c	d	f	g
q	s	t	v	w
y	A	B	D	E
O	Q	R	T	U
W	Y	Z	2	3

Source: Harrison and Jupp (1990) Figure 19

In image analysis, blocking can be used to develop spatial statistics by simulating the effect of a larger pixel size (see Section 8.2). The process may also be applied after resampling to a very small pixel size to produce a more continuous image (see Volume 2B).

The size and 'density' of an image may also be reduced by sub-sampling. This process effectively represents a larger pixel size by selecting the value of a single pixel in the input image rather than using the average of a block of pixels as is the case for image blocking. Figure 7.5a and Figure 7.5b demonstrate this process for converting from a 20 m pixel size to output pixel sizes of 40 m and 30 m respectively.

7.2.1.4 Enlarging

The reverse process to reducing image size is image enlargement, where a single pixel in the input image is duplicated to represent multiple pixels in the output image. This process is sometimes called image 'magnification' and has the effect of changing image scale by spatially sub-dividing the input pixels as shown in Figure 7.6a. This enlargement process may be performed in conjunction with precise scaling as described in Volume 2B, or more simply by using a consistent multiplicative factor. Figure 7.6b shows the result of enlarging an image by converting from a 50 m pixel to a 20 m pixel using along-line resampling (see Volume 2B). Similarly, this process can be used to convert a single line and/or pixel image into a larger image to match the extent of another image as shown in Figure 7.6c.

Figure 7.6 Enlarging an image

A single input pixel could be expanded to a larger number of output pixels by specifying an output pixel width that is smaller than the specified input width in the same proportion as the required expansion.

a. Duplicate each pixel to represent a smaller pixel size

Input image pixel size = 40 m

a	b	c	d
e	f	g	h
i	j	k	l
m	n	o	p

Output image pixel size = 20 m

a	a	b	b	c	c	d	d
a	a	b	b	c	c	d	d
e	e	f	f	g	g	h	h
e	e	f	f	g	g	h	h
i	i	j	j	k	k	l	l
i	i	j	j	k	k	l	l
m	m	n	n	o	o	p	p
m	m	n	n	o	o	p	p

b. Enlarge image by changing relative pixel size from 50 m to 20 m

Input image pixel size = 50 m

a	b	c	d
e	f	g	h
i	j	k	l
m	n	o	p

Output image pixel size = 20 m

a	a	b	b	b	c	c	d	d	d
a	a	b	b	b	c	c	d	d	d
e	e	f	f	f	g	g	h	h	h
e	e	f	f	f	g	g	h	h	h
e	e	f	f	f	g	g	h	h	h
i	i	j	j	j	k	k	l	l	l
i	i	j	j	j	k	k	l	l	l
m	m	n	n	n	o	o	p	p	p
m	m	n	n	n	o	o	p	p	p
m	m	n	n	n	o	o	p	p	p

c. Matching single pixel, or line of pixels, to a larger image

Input pixel depth = 60 m (1 line)

a	b	c	d	e

Output pixel size = 120 m (6 lines)

a	b	c	d	e
a	b	c	d	e
a	b	c	d	e
a	b	c	d	e
a	b	c	d	e
a	b	c	d	e

Source: Harrison and Jupp (1990) Figure 20

7.2.2 Changing image orientation

Image orientation may be modified in a number of ways:

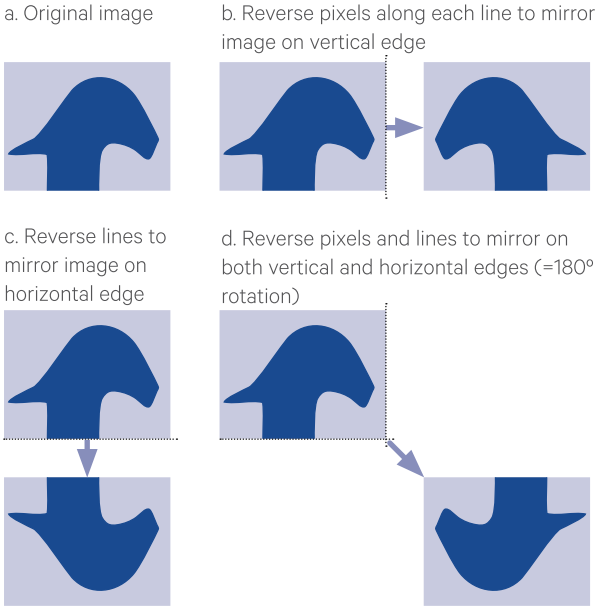
- reversing pixels and/or lines (see Section 7.2.2.1);
- transposing image along a diagonal (see Section 7.2.2.2); and
- rotating the image origin (see Section 7.2.2.3).

As noted in the sub-sections below, different combinations of these operations can be used interchangeably to achieve a particular result. The development of polynomial models that involve image rotation, and the process of resampling imagery, are introduced in Section 7.3, and detailed in Volume 2B.

7.2.2.1 Reversing pixels and/or lines

The reversal of pixel order along an image line produces a mirror image of the original image representation at a vertical edge. Similarly, the order of lines down an image can be reversed by simply mirroring the original image along the first or last line. Such processes are illustrated in Figure 7.7.

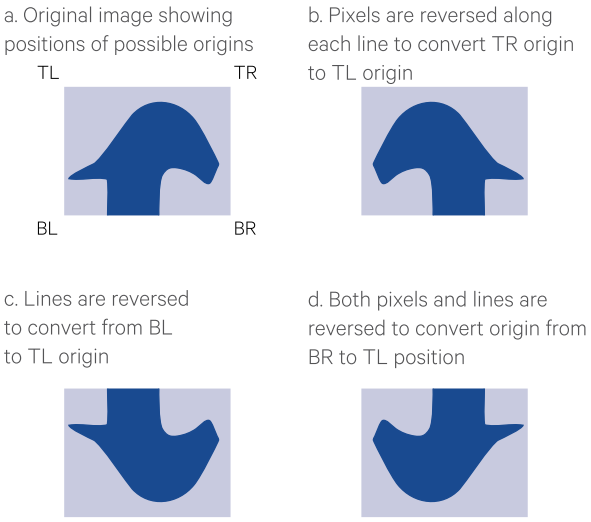
Figure 7.7 Reversing pixels and/or lines in an image



Source: Harrison and Jupp (1990) Figure 21

Since these operations do not involve image rotation, the required reversals can generally be specified by indicating which corner of the original image is to be the top-left corner in the processed image. The corner positions are typically labelled relative to the original image orientation, that is, top-left (TL), top-right (TR), bottom-left (BL) and bottom-right (BR) (see Figure 7.8). Lines and/or pixels are then reversed and/or swapped as required to produce an output image with the defined origin in the top-left position.

Figure 7.8 Defining image origin

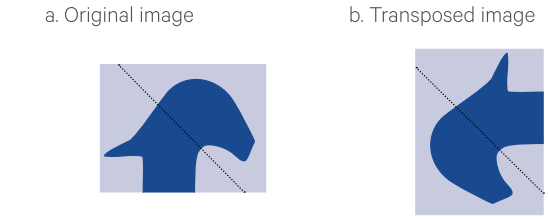


Source: Harrison and Jupp (1990) Figure 22

7.2.2.2 Transposing an image

This process involves rewriting the image with the rows of the input image being columns in the output image and columns of the input image being output rows (see Figure 7.9). The same result could be achieved by rotating the image 90° (see Figure 7.10b) then reversing pixels along each line (see Figure 7.7). For square images, the effect of this operation is to flip the image about the diagonal axis between the top-left corner and the bottom-right corner.

Figure 7.9 Transposing an image

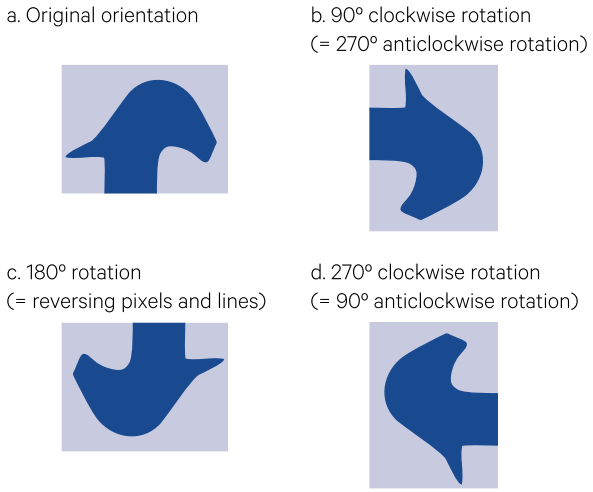


Source: Harrison and Jupp (1990) Figure 23

7.2.2.3 Rotating image origin

The image origin can also be altered by rotating an image by 90°, 180° or 270° in a clockwise or anti-clockwise direction (see Figure 7.10). These simple rotation angles can be applied to an image using combinations of pixel and/or line reversals (see Section 7.2.2.1) and image transpositions (see Section 7.2.2.2) as illustrated in Figure 7.8, Figure 7.9 and Figure 7.10. A rotation of 90° in a clockwise direction is equivalent to transposing the image, then reversing pixels (see Figure 7.10b). Rotating by 180° (see Figure 7.10c) is the same as reversing both lines and pixels (see Figure 7.7d and Figure 7.8d above), while a 270° rotation in a clockwise direction (see Figure 7.10d) gives the same results as reversing the lines after transposing an image.

Figure 7.10 Rotating an image



Source: Harrison and Jupp (1990) Figure 24

7.3 Rectification and Resampling

Previous sub-sections introduce the concepts of geometric manipulation of imagery and describe operations which may be required to modify the scale or orientation of imagery before processing. Those operations described thus far, however, do not actually change the 'shape' of the image nor correct for geometric distortions.

EO imagery contains a variety of geometric distortions which need to be accounted for before registering with different image sources or map data (see Volume 2B—Section 2). Image processing software can implement a range of models to account for known geometric distortions in satellite and aircraft scanner imagery as well as map projection models which compensate for the spatial inconsistencies between different map coordinate systems (see Volume 2B—Section 3). These models can be used in conjunction with control point modelling techniques to achieve accurate registration between different coordinate systems (see Volume 2B—Section 4). The fundamentals of map projections are introduced in Volume 2B—Section 1.

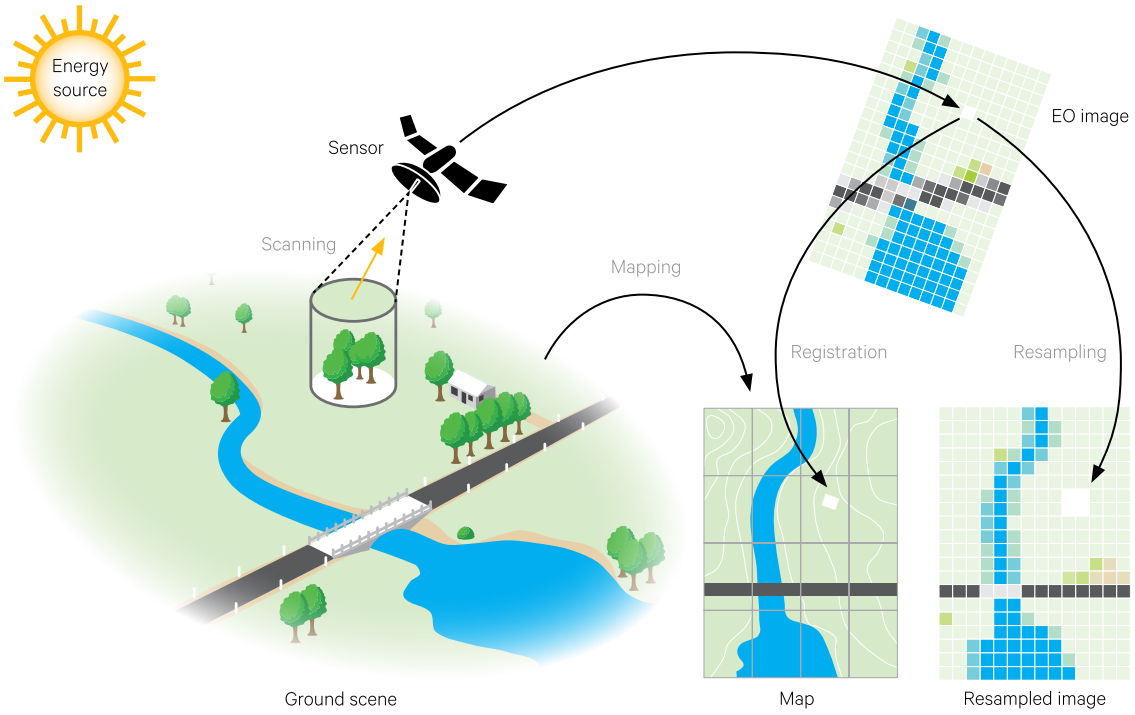
Image rectification involves modelling the geometry and topology of an image. Image registration models the relationship between the geometry of an image and the geometry of some other spatial representation of the imaged scene (such as a map or another image). Image resampling uses the modelled relationship between two 'geometries' to rewrite the geometry of one dataset to match the other, for example, to overlay an image onto another image or a map (see Volume 2B—Section 5). This process changes the spatial relationship between objects in the image, and may be compared to selectively stretching sections of the image as if it were a rubber sheet. Resampling can also be applied to multiple input images, to create a mosaic (see Section 7.2.1.2).

The rationale for image rectification and registration in the context of EO images is illustrated in Figure 7.11. The left side of this figure shows a ground scene being imaged by an EO sensor. Each optical pixel—a remote measurement that is integrated over a 'volume' on the surface being imaged—relates to one geometric pixel in the EO image (see Volume 1B—Section 1.2). In Figure 7.11, one optical pixel is shown as a cylinder. The relationship between the geometry of the resulting image and the geometry of the scene being imaged may be described by modelling the optical and physical relationship between the scene and the sensor. As illustrated on the right side of Figure 7.11, this relationship is commonly defined relative to a map of the scene.

I think the universe is pure geometry—
basically a beautiful shape twisting around
and dancing over space-time.
(Antony Garrett Lisi)

Figure 7.11 Earth Observation scene geometry

This illustration shows a ground scene being imaged by a remote sensor. A 'volume' on the surface that is observed as one optical pixel is highlighted by the cylinder. The corresponding geometric pixel for this location is shown as a white square in the EO image. While the EO image contains recognisable features from the ground scene, its geometry rarely matches a standard map projection. To rectify the image to map coordinates, the locations of specific features needs to be identified on a map. For example, the location of the highlighted optical pixel is shown here as a white square on the map. To allow EO imagery to be digitally or physically registered to a map, it then needs to be resampled to match the map geometry. As shown in this example, the original optical pixel may be represented by multiple geometic pixels in the resampled image.



Adapted from Harrison and Jupp (1992) Figure 1

As detailed in Volume 2B, transformations between two different geometries, such as an EO image and a map, usually involve three stages:

- a nominal coordinate transformation which corrects for known image distortions;
- a nominal map transformation, which accounts for inconsistencies in map coordinates; and
- a polynomial transformation to relate the two nominally transformed geometries.

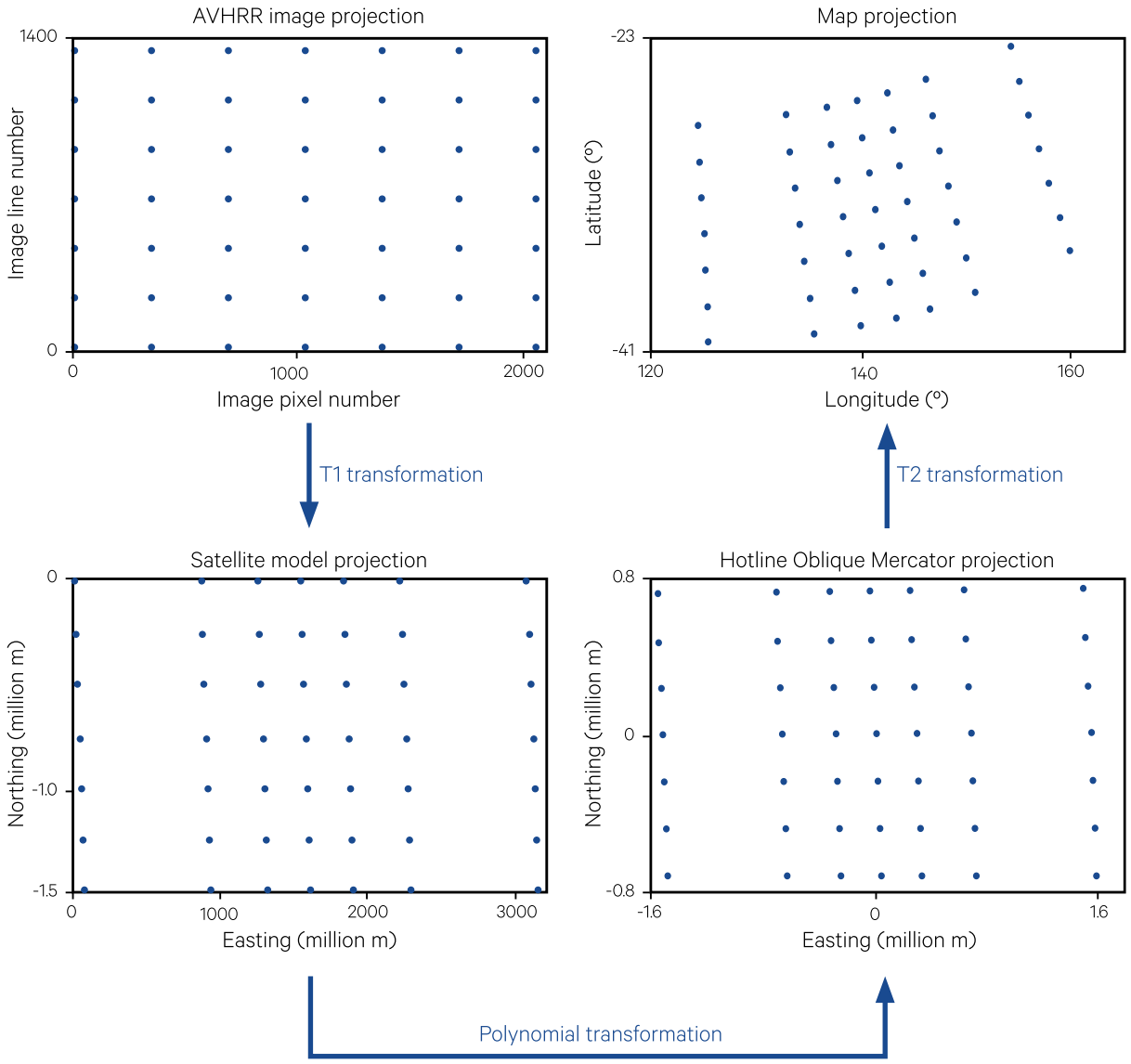
To convert from coordinates in an image to a map, for example, the modelling sequence would be:

- convert from EO image coordinates to nominal image coordinates using a pre-defined image model;
- convert from nominal image coordinates to nominal map coordinates using a polynomial model derived from ground control points; and
- convert from nominal map coordinates to coordinates of selected map projection using a pre-defined map model.

This three-stage approach allows the relationship between the two nominally transformed geometries to be modelled by a lower order (and thus more stable) polynomial model. Figure 7.12 illustrates this process for an AVHRR image being registered to geographical coordinates. Development of appropriate rectification and registration models for EO imagery is described in Volume 2B.

Figure 7.12 Three-stage rectification process

This schematic shows the three stages of modelling from evenly-spaced pixel locations in an AVHRR image to geographical coordinates. The T1 transformation uses a pre-defined satellite model to convert from image coordinates to a grid based on metres Similarly, the T2 transformation uses pre-defined cartographic models to convert between map coordinates and a metre grid. The relationship between the two modelled grids can then be computed as a polynomial model of lower order than would have been required to map directly from image to map coordinates. A corresponding set of inverse models are used to convert from map to image coordinates.



Source: Harrison and Jupp (1992) Figure 31

Generally, the scene geometry is represented by an independent mapping exercise so that the image is viewed as being registered with the scene when it is registered to a map of the scene. To sensibly overlay the image onto the map, or map data onto an image, both the map and image data need to be viewed using the same projection and datum (see Figure 7.13). This can be done either by resampling the image data to the same projection as the map or, for visual presentation only, converting the map and/or image data to a matching projection during the image display process. The latter option assumes that the map and image data are represented in projections that can be 'interconverted' during image display. The resampling process for EO imagery is detailed in Volume 2B.

Figure 7.13 Overlaying map data

Aerial imagery acquired in 2016 and orthorectified to the ACT (Stromlo) Grid, which is based on the Australian Geodetic Datum 1966 (AGD66)[3]. In this example, a portion of this imagery over central Canberra is overlain with a GIS layer that shows the actual extent of lakes. The GIS layer, however, is stored in MGA 55 coordinates relative to the Geodetic Datum of Australia 1994 (GDA94). MGA coordinates based on GDA94 are shifted approximately 200 m northeast of MGA coordinates based on AGD66.

a. Overlay of Lakes GIS layer without matching the image coordinate system (in this case due to a different datum) results in an overlay shift to the northeast of approximately 200 m relative to image features.

b. Overlay of Lakes GIS layer after transformation to the same coordinate reference system used in the image.

Source: Tony Sparks, Icon Water, based on ACT Government orthorectified imagery supplied by AAM (Ref. No.: 26121A) CC BY 4.0.

7.4 Image grids

Resampled imagery implicitly conforms either to a:

- custom grid system—whose scale and geometry is customised for a specific EO sensor (see Section 7.4.1); or a
- nested grid configuration—which is generic for a range of sensors and products (see Section 7.4.2).

7.4.1 Sensor-specific custom grids

To allow EO data to be overlaid with map data or other imagery, the original geometry of the data is modified to conform to a selected grid. The image grid is based on an appropriate map coordinate system and a grid cell size that suits the spatial resolution of the original data.

Custom grids used by many EO data suppliers differ for each data source and do not necessarily conform to a common underlying grid cell size or map coordinate system. Consequently, when applications require that imagery from different EO sources be compared, one or both images need to be further resampled to a common map projection and/or custom grid. If not conducted with care and consistency, this process may introduce additional artefacts, especially when multiple images are being analysed. Accordingly, a new range of EO image products is being offered, which attempt to standardise image grids between different data sources (see Section 7.4.2).

7.4.2 Sensor-agnostic nested grids

Traditionally EO imagery has been reprojected onto a grid whose scale and geometry is customised for a specific EO sensor and conforms to a relevant map projection (see Section 7.4.1). Customised grids are not necessarily compatible between different sensors, however, thus complicating comparison of imagery from different sources. To address this problem, the Australian and New Zealand Land Information Council (ANZLIC) have developed standards for a National Nested Grid (NNG), with scalable grid cell sizes (ANZLIC, 2012). The grid cell sizes of products comprising this suite are designed to be scalable (such as 25 m, 250 m and 1000 m) and thus facilitate easy data comparison between different products. This format will greatly simplify comparison of image datasets acquired by different sensors at varying spatial scales.

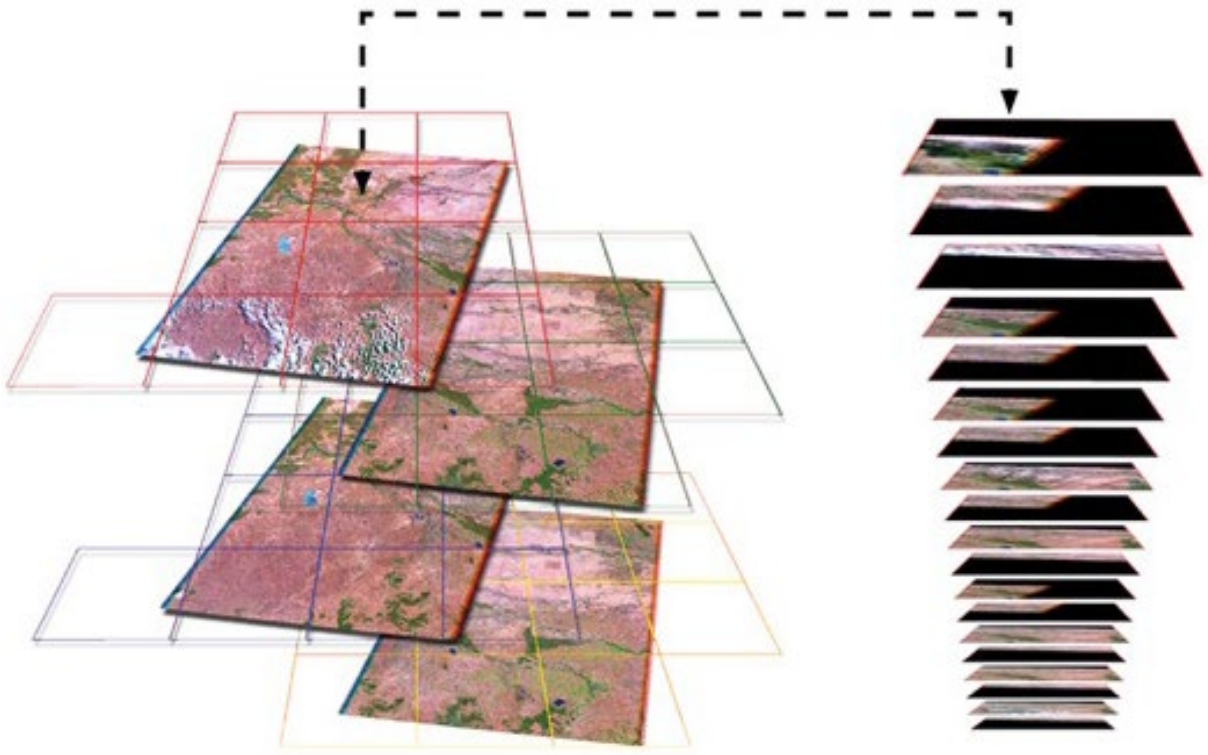
Accordingly, major EO product suppliers are planning to offer EO products that share a common underlying nested grid base with scalable cell sizes. Nested grids will allow EO data to become 'sensor agnostic' and readily comparable with spatial data from other sources. This feature is particularly relevant given the growing time series of EO imagery and will significantly simplify the process of monitoring long term changes in land and water resources.

3 For further information on the ACT projection, please visit: *https://www.planning.act.gov.au/tools_resources/survey-data-maps/surveying-data/surveyors_information/coordinate_system/act_standard_grid_parameters*

For example, the ARG25 is the first NNG-compatible product to be released by GA in the DEA (see Section 3.4). This product is structured as a data cube, with time as the third dimension (see Figure 7.14). The surface reflectance data for each imaged date and time are stacked in chronological order, with Landsat TM/ETM+ sources being interleaved as appropriate to their acquisition time. Interrogation of this data cube allows the time series for individual grid cells, as well as their heritage and data quality, to be extracted and compared. The resulting stack of surface reflectance grids is effectively consistent over space and time. This product greatly reduces the processing requirements for numerous studies, and enables the development of a new set of tools for identifying and quantifying environmental change in land and water resources across Australia (see Volume 3).

Figure 7.14 GA data cube for ARG25 product

The data cube organises images from multiple sources and times, such as the images on the left, into stacks of tiles on the right, with the top of the stack being the most recent observations.



Source: Alex Ip, Geoscience Australia.

7.5 Further Information

Intergovernmental Committee on Surveying and Mapping: *http://www.icsm.gov.au/index.html*

Jensen (2016) Chapter 7

7.6 References

Ahuja, N., and Schachter, B. J. (1983). *Pattern Models*. John Wiley and Sons, New York.

ANZLIC (2012). *National Nested Grid (NNG) Specification Guideline – 2012*. ANZLIC NNG Workgroup.

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Harrison, B. A., and Jupp, D. L. B. (1992). *Image Rectification and Registration. Part FOUR of the microBRIAN Resource Manual (draft only)*. MPA, Melbourne.

Jensen, J.R. (2016) *Introductory Digital Image Processing. A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc.

Analysis

Image analysis involves using tools based on statistics and/or numerical analysis to identify features and patterns in the image. The following sections introduce the image analysis topics of:

- statistics (see Section 8);
- classification (see Section 9);
- and segmentation (see Section 10).

These topics are considered in greater detail in Volumes 2C, 2D and 2E.

Contents

8 Statistics 107

9 Classification 129

10 Segmentation 163

Background image on previous page: Landsat-8 image over Weipa, Queensland, acquired on 8 August 2013. This composite image displays bands 6, 5, 3 as RGB and has been sharpened using the panchromatic band. Weipa is situated on the western coast of Cape York Peninsula, facing the Gulf of Carpentaria, and is known for numerous bauxite mines. **Source:** Norman Mueller, Geoscience Australia

8 Statistics

Sections 4 to 7 above introduce the concept of a digital image as a spatial dataset having properties of both colour and geometry. Being in numeric form, digital image data may also be processed using standard statistical techniques to effectively analyse its 'colour' properties. These techniques may be categorised into four groups:

- spectral—techniques that effectively disregard the spatial properties of image data so that an image channel is treated as a one-dimensional array of numbers. These techniques derive statistics that describe the image in terms of the range and variation of data values they contain (see Section 8.1).;
- spatial—describe the data values of an image as they relate to its spatial patterns. These statistics can be used to examine the spatial patterns within the image data and quantify the effects of changes in image spatial resolution (see Section 8.2);
- radiometric—methods to analyse components of image (or ground calibration) values that are not due to surface reflectance, such as atmospheric components, BRDF, terrain-related variations in illumination and scanner inconsistencies (see Section 8.3); and
- temporal—ways to standardise multi-date image data, potentially comprising two to tens of thousands of images, to enable meaningful quantitative comparisons and analysie (see Section 8.4).

8.1 Spectral

A digital image represents spatial attributes as arrays of numbers. Statistics relating to the distribution of values in each channel, or the similarity of values (or correlation) between pairs of channels, can be derived from a digital image. These statistics can then be used to analyse image quality, extract particular spectral features, reduce data volume or transform the data in a variety of ways.

While image data may relate to a wide range of attributes, in EO imagery data channels (or bands) typically represent a measure of spectral radiance or reflectance. Although the following discussion refers to 'spectral' image values, these statistics could be derived for any image attribute data.

8.1.1 Histograms

The distribution of data values in any image channel may be represented as a histogram as introduced in Section 4.1.1. The usual form for histograms in image processing is the 'frequency histogram' shown in Figure 4.2, with the horizontal axis representing the image data range and the vertical axis indicating the number of pixels in the image with each data value.

As the scale of a printed or displayed histogram may not be sufficiently large to enable values to be read exactly, ancillary statistics relating to the absolute minimum, maximum and mode are usually reported separately. In the example in Figure 4.2, no pixels had values below 15 or above 137. The total number of pixels used to generate the histogram is also usually reported: 61440 in this case. If the whole image was used to generate the histogram, the total number of pixels in a channel will equal the number of pixels per line multiplied by the number of lines minus the number of null pixels in that channel. In many image processing systems, histograms may also be generated for sub-regions of an image, on the basis of a sub-range of values in one or more channels (called a spectral theme) or a sub-area (called a training patch or training set). The use of spectral themes and training patches are further discussed in Section 9.1 in the context of image classification.

Percentage points (or percentiles) for the histogram are also generally provided. These values summarise

Background image: Landsat-8 image over Weipa, Queensland, acquired on 8 August 2013. This composite image displays bands 6, 5, 3 as RGB and has been sharpened using the panchromatic band. Weipa is situated on the western coast of Cape York Peninsula, facing the Gulf of Carpentaria, and is known for numerous bauxite mines. Source: Norman Mueller, Geoscience Australia

the histogram distribution in terms of cumulative statistics (see Figure 4.2 and Figure 4.3). The percentage points provide a 'normalised' view of histogram values that is useful for operations that require histograms to be compared between channels (see Volume 2D—Section 1.4).

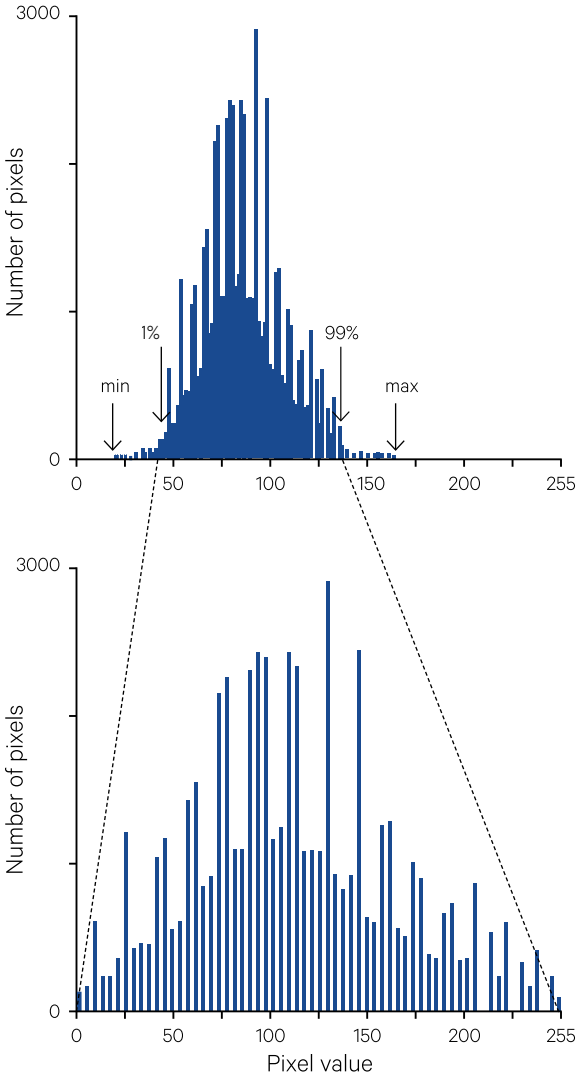
The histogram shape is also useful in interpreting the image data. Skewed data distributions commonly occur in remotely sensed imagery. Image acquisition and calibration operations can sometimes truncate the upper or lower limits of the available data range. For example, during image acquisition the sensor gains may not be set correctly to capture the full range of radiance values from a particular ground scene. When the scanner has been set to measure a much larger range of intensities than occurs in the scene, the range of intensity values in the channel will be relatively narrow. In EO imagery, a narrow range can also be due to poor illumination (such as a low Sun angle). Alternatively, the minimum radiance level detected by a sensor may be higher than the minimum radiance from an image scene and/or the maximum level being detected may be lower than the maximum radiance causing the discrimination between very low and/or very high values to be lost. The image histograms can be used to identify such problems. Histograms should also be checked after all scaling operations, such as transformations, to ensure that the scaled value range of the output image has not been compressed or truncated unnecessarily (see Volume 2C).

Discontinuous histograms are sometimes encountered in EO imagery. Where image data have been recalibrated from 8 bits (values in the range 0–255) to a 9 or 10 bit data range, discontinuities may occur in the image data range so that certain values within the range are not represented in the image. This effect is usually due to pre-processing with an inadequate recalibration algorithm that does not completely account for sensor differences. For similar reasons, image values may be unevenly distributed as shown in Figure 8.1a.

It is the mark of a truly intelligent person to be moved by statistics.
(George Bernard Shaw)

Figure 8.1 Effect of rescaling on image histogram

The original image histogram (top) has a mostly continuous histogram between values 15 (1% minimum) and 137 (99% maximum). When the image values are rescaled to fill the full 8-bit (byte) range (0 to 255), the histogram of the rescaled image (below) is discontinuous.



Source: Harrison and Jupp (1990) Figure 40

Discontinuities may be visible in imagery that has been stretched, especially when the image data span a small range (see Volume 2C). The cause of this effect is shown in Figure 8.1 where a channel, which contains only 60 different pixel values between the 1% and 99% points, is re-written to fill the whole 8-bit (byte) image data range (256 values). Each of the original values maps to a unique value in the expanded range to produce only 60 different output values. Discontinuities in image histograms may also be due to spectral features of the image itself as illustrated in Figure 4.4, or radiometric irregularities, such as sensor striping (see Section 8.3).

8.1.2 Variance

The calculation and interpretation of variance (σ^2) and standard deviation (σ) statistics for an image channel are introduced in Section 4.1.2. The standard equation for variance is given as:

$$\textit{channel variance}\left(\sigma^2\right) = \frac{\sum\left(x-\mu\right)^2}{n}$$

where

x is pixel value in channel;
n is total number of pixels in channel (excluding null pixels); and
μ is true mean of the channel data.

This equation can only apply when the real mean of the data values is known exactly, rather than just being estimated from a set of sample values. In image processing, statistics are frequently generated by sampling image values and are usually interpreted to be representative of larger image areas. Thus, the mean value used to calculate variance is only an estimate of the real data mean and the equation is usually modified to use one less degree of freedom as:

$$\frac{\sum\left(x-\overline{x}\right)^2}{n-1}$$

where

$$\textit{channel mean}\left(\overline{x}\right) = \frac{\sum\left(x\right)}{n}$$

An alternative formula for variance is:

$$\frac{\sum\left(x\right)^2 - n\left(\overline{x}\right)^2}{n-1}$$

This form can produce very unstable results (such as negative variance values) so should not be used. The following equations are used to update mean and variance statistics in some image processing systems:

$$\overline{x}_n = \overline{x}_{n-1} + \mathrm{e}_{n-1}$$

$$\sigma_n^{\,2} = \frac{n-1}{n}\sigma_{n-1}^{\,2} + \left(n-1\right)\mathrm{e}_{n-1}^{\,2}$$

where

$$\mathrm{e}_{n-1} = \frac{x_n - \overline{x}_{n-1}}{n}$$

and

n is the number of samples.

For large values of n, as typically occur in image analysis, the difference between dividing by n or n–1 is not significant.

Null pixels should be excluded from all statistical calculations in image processing. In most image processing systems, when the variance of a single channel is being computed, a pixel is only treated as masked or excluded from the image if its value equals the null value in that channel. In some multi-channel operations, however, a pixel will be excluded if it has a null value in any channel.

If a channel has a low variance (and hence low standard deviation) we would expect that most pixel values would be close to the mean value (see Section 4.1.2). The standard deviation is related to the probability of a pixel having a particular value in the channel range. A theorem developed by Chebyshev allows us to define, for any data distribution, a lower bound estimate of the probability that a random variable falls within a particular range of the mean; namely, there is greater than or equal to 75% probability that a random pixel value in the range will be within 2σ units of the mean, or at least 89% probability that it will be within 3σ units (Walpole, 1974). Such probabilities can be determined more precisely if the probability distribution of the data is known.

Channel standard deviation is a particularly important statistic for checking training patches when developing an image classification (see Section 9 and Volume 2E). Typically, variance of individual image classes as well as the overall variance for a classification can be computed during the classification process.

8.1.3 Crossplots

Sections 8.1.1 and 8.1.2 above discussed image statistics relating to single channels. These statistics were derived by treating the image channel as a one-dimensional array of numbers, with no account being taken of the location of the pixel values. Most imagery, however, contains multiple channels. Various statistics, which indicate the correlation between channels, may be calculated in the same way, by considering the multi-channel image as a multi-dimensional matrix (see Figure 1.1). For example, the six optical data channels in a Landsat TM image can be considered as a six-dimensional matrix.

The simplest way to view the correlation between channels is to construct crossplots for pairs of channels (see Excursus 8.1). A crossplot is a two-dimensional graph with the horizontal axis showing the values of one channel, the vertical axis showing the values of the second channel and the points on the graph indicating the 'paired values' of pixels in the image. These points can also be colour-coded to indicate the relative number of pixels with each paired value. The crossplots in Figure 8.2 illustrate

the correlation between channels A, B and C in the example image, with the number of pixels having each paired value indicated by the number of dots.

To define the relationship between the two channels, a regression function can be fitted to the crossplotted data. The significance of the function can also be computed and, if relevant, the regression line is displayed on the crossplot as illustrated in Figure 8.2b. It may also be useful to compute the number of pixels at each paired set of values represented on a crossplot by cross-tabulation of the channels.

Correlation between channels indicates a degree of redundancy in image data. If channels are highly correlated the image data can be reduced to a smaller number of dimensions (or channels). This may be advantageous for efficiency in processing, transmission or storage, or to display the three most significant dimensions. Principal Component Analysis (PCA) allows image data to be reduced to fewer dimensions (see Volume 2C). While correlation between channels may simplify image handling, it may also indicate limitations for feature recognition. In the case of Landsat TM for example, the visible channels TM1, TM2 and TM3 are usually highly correlated. The same applies to the near infrared and shortwave infrared channels, TM5 and TM7. This means that in a broad sense there may be only two or three different data dimensions for identifying image features. This aspect of channel correlation is most significant for the design of sensors and choice of airborne scanner channels for particular remote sensing projects.

Excursus 8.1—Band Correlation Statistics

For simplicity, this example is based on a three-channel image containing the pixel values shown in Figure 8.2. Mean and variance statistics (see Section 8.1.2) for these channels are:

$n = 25$

$x_A = 3.3$

$x_B = 4.2$

$x_C = 3.12$

$\sigma^2_A = 2.938$

$\sigma^2_B = 118/25 = 4.72$

$\sigma^2_C = 56.64/25 = 2.2656$

The crossplot graphs in Figure 8.3 were constructed by locating the channel values of each pixel in the image along the relevant axes. For example, in Figure 8.3a, for each pixel in the image, the channel A value was located on the horizontal axis, the value of channel B on the vertical axis, then a dot is marked on the graph at the intersection of these two values. On each crossplot, this process is shown for the values of pixel 1, line 5 in the image.

Figure 8.2 Example image channels

Channel A

	1	2	3	3	4
	3	3	5	6	3
	2	3	3	6	5
	5	4	2	2	1
pixel 1, line 5 →	7	6	3	2	1

Channel B

	3	6	7	8	7
	2	4	7	5	6
	1	1	3	4	5
	1	1	2	5	5
pixel 1, line 5 →	2	3	4	6	7

Channel C

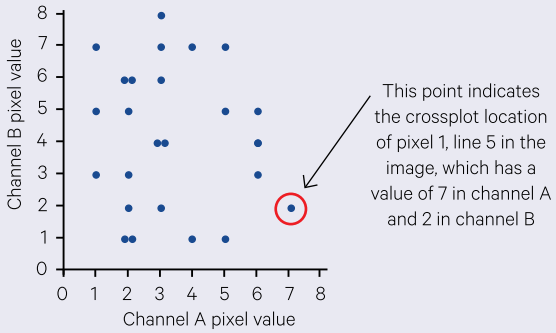
	1	3	4	3	3
	2	3	4	5	4
	2	2	3	6	5
	4	4	2	1	2
pixel 1, line 5 →	6	5	2	1	1

Source: Harrison and Jupp (1990) Figures 41

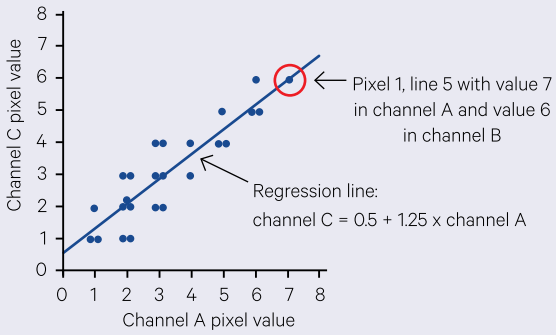
Figure 8.3 Example crossplots

The image channels in Figure 8.2 are crossplotted to show the extent of correlation between channels.

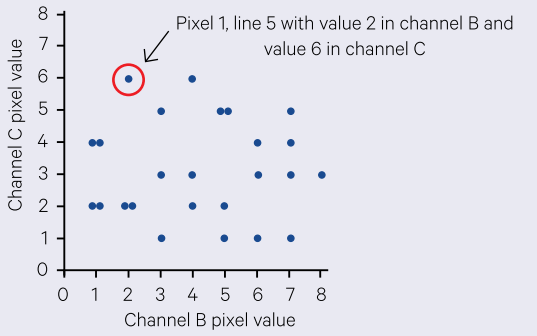
a. Channel B versus Channel A



b. Channel C versus Channel A



c. Channel B versus Channel C



Source: Harrison and Jupp (1990) Figures 42

The distribution of the points on the crossplots immediately indicates the similarity, or correlation, between the two plotted channels. Figure 8.3a shows little correlation between channels A and B, with the points being widely distributed around the graph. The value of a pixel in channel A is unlikely to indicate its value in channel B. However, Figure 8.3b shows a strong correlation between channels A and C. With this type of relationship, it would be possible to calculate the position of the regression line between the two channels and use the values of channel A to predict approximate values in channel C. Figure 8.3c again shows little correlation between channels B and C. This is not surprising considering that channels A and C are very similar and channels A and B are largely uncorrelated. In the case of channels A and C, the correlation is described as 'positive', since an increase in a value in one channel relates to an increase in the other channel. Channels can also be negatively correlated if there is a trend of increasing values in one channel being related to decreasing values in the other.

Table 8.1 details the calculations required to compute the covariance between channels A and B (see Section 8.1.4). When applied to all channel pairs, the covariance statistics for the example image channels are:

$$v_{AB} = -9.6 / 25 = -0.384$$

$$v_{AC} = 57.544 / 25 = 2.3$$

$$v_{BC} = -3.6 / 25 = -0.00144$$

Table 8.1 Covariance matrix calculation

Statistics for channels A and B are based on the example images shown in Figure 8.2.

x_A	$x_A - \overline{x}_A$	x_B	$x_B - \overline{x}_B$	$\left(x_A - \overline{x}_A\right) \times \left(x_B - \overline{x}_B\right)$
1	-2.3	3	-1.2	2.76
2	-1.3	6	1.8	-2.34
3	-0.3	7	2.8	-0.84
3	-0.3	8	3.8	-1.14
4	0.7	7	2.8	1.96
3	-0.3	2	-2.2	0.66
3	-0.3	4	-0.2	0.06
5	1.7	7	2.8	4.76
6	2.7	5	0.8	2.16
3	-0.3	6	1.8	-0.54
2	-1.3	1	-3.2	4.16
2	-1.3	1	-3.2	4.16
2	-1.3	3	-1.2	1.56
6	2.7	4	-0.2	-0.54
5	1.7	5	0.8	1.36
5	1.7	1	-3.2	-5.44
4	0.7	1	-3.2	-2.24
2	-1.3	2	-2.2	2.86
2	-1.3	5	0.8	-1.04
1	-2.3	5	0.8	-1.84
7	3.7	2	-2.2	-8.14
6	2.7	3	-1.2	-3.24
3	-0.3	4	-0.2	0.06
2	-1.3	6	1.8	-2.34
1	-2.3	7	2.8	-6.44
Total				-9.60

In this example, channels A and C have a high covariance value, which is consistent with the strong correlation visible in the crossplot of these two channels. Similarly, both A versus B and B versus C, which had no visible correlation in the crossplots, have low covariance values.

Accordingly, for our example image, the covariance matrix is:

Channel	A	B	C
A	2.938	-0.384	2.300
B	-0.384	4.720	-0.014
C	2.300	-0.014	2.266

and the correlation matrix is:

Channel	A	B	C
A	1.0	-0.10312	0.89148
B	-0.10312	1.0	-0.00404
C	0.89148	-0.00404	1.0

8.1.4 Covariance

The distribution of data values in a single channel could be represented pictorially by a channel histogram or mathematically by the variance statistic. Similarly, the correlation between two channels can be shown pictorially by a crossplot or represented mathematically by a statistic called covariance.

The covariance between two channels i and j (v_{ij}) is calculated as:

$$v_{ij} = \frac{\sum_{k=1}^{n}\left(x_{ik} - \overline{x}_i\right)\times\left(x_{jk} - \overline{x}_j\right)}{n}$$

where

n is the total number of pixels in the image;
x_{ik} is value for pixel k in channel i;
$\overline{x}_i$ is the mean of pixel values in channel i;
x_{jk} is value for pixel k in channel j; and
$\overline{x}_j$ is the mean of pixel values in channel j.

In most statistical analyses in image processing, the mean values are estimated from a sample of the image values so that a degree of freedom is lost in calculating the covariance. This equation is then usually implemented as:

$$v_{ij} = \frac{\sum_{k=1}^{n}\left(x_{ik} - \overline{x}_i\right)\times\left(x_{jk} - \overline{x}_j\right)}{n-1}$$

As mentioned before, any null pixels should be excluded from computation of image statistics. For covariance statistics, a pixel will be excluded if it has the null value in either channel i or channel j. A covariance matrix $\mathbf{C}_{n-1}$ based on n–1 samples, that is:

$$\frac{1}{n-1}\times\sum\left(\boldsymbol{x}_n - \overline{\boldsymbol{x}}_{n-1}\right)\times\left(\boldsymbol{x}_n - \overline{\boldsymbol{x}}_{n-1}\right)^T$$

can be updated using the equations:

$$\boldsymbol{C}_n = \frac{n-1}{n}\boldsymbol{S}_{n-1} = \left(n-1\right)\boldsymbol{e}_{n-1}\boldsymbol{e}_{n-1}{}^T$$

where

$$\boldsymbol{e}_{n-1} = \frac{\boldsymbol{x}_n - \overline{\boldsymbol{x}}_{n-1}}{n}$$

$$\overline{\boldsymbol{x}}_n = \overline{\boldsymbol{x}}_{n-1} + \boldsymbol{e}_{n-1}$$

and

$\overline{\boldsymbol{x}}_{n-1}$ is the vector of channel mean values for n–1 samples.

Negative covariance values indicate negative correlation between the channels; that is, high values in one channel are associated with low values in the other. Variance values, however, can never be negative.

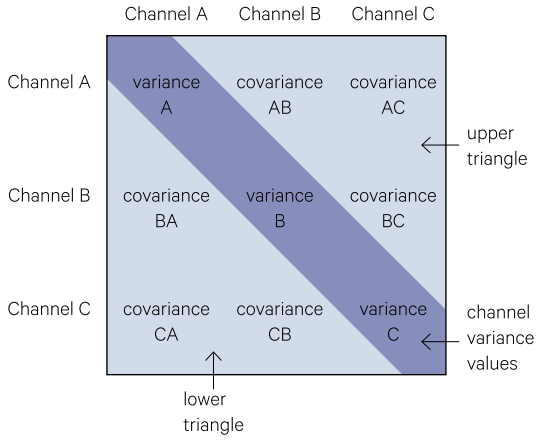
Large positive covariance values indicate a strong positive correlation between the two channels, while a large negative covariance would indicate that a strong negative correlation exists between the channels. The covariance between two identical channels (that is each pixel has the same value in both channels) would be equal to the variance of the channel. In a uniformly random image, that is, when every image value has equal probability of occurring at any pixel and with independent channels, the variance in each channel would approximate:

$$\frac{1}{12}\left(\text{maximum image value-minimum image value}\right)^2$$

and the covariance values should be close to zero. Thus, for a random image with values in the range 0–255, the variance would be approximately 5419.

To evaluate the variance within and between all channels in an image, these variance and covariance values are usually presented in a table of the form shown in Figure 8.4. This table is called a covariance matrix. Since the covariance AB is the same as BA, the matrix is always symmetric about the diagonal of variance values. As such it may also be written as a triangular matrix with either the upper or lower triangle omitted.

Figure 8.4 Format of variance/covariance matrix



Source: Harrison and Jupp (1990) Figure 43

The calculation of all values in the variance/covariance matrix can be summarised in the equation:

$$v_{ij} = \frac{\sum_{k=1}^{n}\left(x_{ik} - \overline{x}_i\right)\times\left(x_{jk} - \overline{x}_j\right)}{n}$$

for

$i = 1$ to m
$j = 1$ to m
$k = 1$ to n

where

m is the number of channels in image; and
n is the number of pixels in the image

or

$$v_{ij} = \frac{\sum_{k=1}^{n}\left(x_{ik} - \overline{x}_i\right)\times\left(x_{jk} - \overline{x}_j\right)}{n-1}$$

where

$\overline{x}_i$ and $\overline{x}_j$ are sample-based estimates of the real mean values.

Another representation of variance between channels is the correlation matrix. This is a normalised form of the covariance matrix that has values of 1 along the diagonal with all off-diagonal values being in the range -1 to 1. Elements of the correlation matrix are calculated from the covariance matrix using the formula:

$$c_{ij} = \frac{v_{ij}}{\sqrt{v_{ii}\times v_{jj}}}$$

where

c_{ij} is the correlation between channels i and j;
v_{ij} is the covariance between channels i and j;
v_{ii} is the variance of channel i; and
v_{jj} is the variance of channel j.

As with the covariance matrix, values close to 0 indicate little correlation between channels. A value of 0 would mean no correlation at all. The advantage of using a correlation matrix is to have the values normalised within a fixed range. In the covariance matrix 'large' positive or negative values indicate strong correlation between channels. In the correlation matrix we have a scale to clearly define 'large', with values close to 1 or -1 showing strong positive or negative correlation respectively. Diagonal values are always equal to 1 since each channel is completely correlated with itself. This would only occur for off-diagonal values if each pixel has the same values in both channels. This normalising allows correlation matrices to be directly compared between imagery, an operation that is not valid using covariance matrices.

Principal Component transformations completely remove the correlation between image channels, that is, create an image that has all zero values for the off-diagonal entries in its covariance matrix. Such computations may be based on either a covariance or correlation matrix (see Volume 2C).

8.2 Spatial

Section 8.1 discusses various statistics that may be derived from the spectral values of digital imagery to assist in its analysis and interpretation. (Here 'spectral' refers to both radiance and other measurable attributes of a pixel.) These spectral statistics are computed without any regard for the spatial arrangement of those values within the image. For example, Figure 8.5 illustrates two images with the same distribution of spectral values, and hence exactly the same spectral statistics, but displaying totally different spatial patterns.

Figure 8.5 Spectrally equivalent images with differing spatial patterns

Both images contain 12 pixels with value 1 and 13 pixels with value 5.

5	5	5	5	5
5	1	1	1	5
5	1	1	1	5
5	1	1	1	5
5	1	1	1	5

5	1	5	1	5
1	5	1	5	1
5	1	5	1	5
1	5	1	5	1
5	1	5	1	5

Source: Harrison and Jupp (1990) Figure 44

Visual interpretation of imagery is largely dependent on the delineation of spatial patterns. Such patterns may not be easily extracted from digital image processing despite being clearly visible to a human interpreter. However, various filtering and texture operators can be used to manipulate the spatial information in imagery (see Volume 2C).

Spatial statistics may also be derived from image values to indicate the image spatial structure. Various models and measures have been proposed to describe the spatial variation within an image (Ahuja and Schachter, 1983; Levine, 1985). For example, spatial statistics have been to forecast large-scale erosion and deposition patterns in arid landscapes (Pickup and Chewings, 1988).

Spatial structure and statistics of digital imagery are significantly influenced by many factors including the size, shape and density of objects in the original scene and the relationship between the size of the objects and the size of the image pixels (Woodcock and Strahler, 1987). Various aspects of image spatial resolution are discussed in Volume 1. Distinct objects tend to be spectrally homogeneous and separable from other objects and background. In this case, the extent to which adjacent pixels in the image are more spectrally similar than expected, that is, the degree to which they are 'spatially autocorrelated', is determined by the size of an image pixel relative to the objects being imaged. The mean distance at which the spatial autocorrelation disappears is related to the size, spacing and shape of the objects being imaged.

The nature and causes of spatial variation in remotely sensed imagery are investigated by Woodcock *et al.* (1988a, 1988b) using the variogram measure with simulated and real data. Jupp (1988b, 1988a) provide a mathematical framework for modelling the effect of image pixel size on spatial image covariance then apply these results to show that the spatial structure of an image depends on the underlying spatial structure of the scene that it represents.

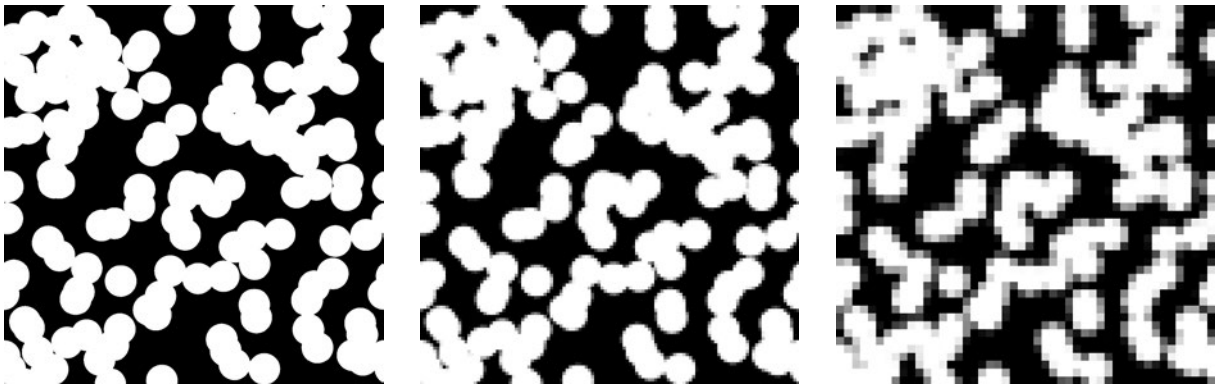
These authors have defined two classes of image spatial resolution that relate image pixel size to object size in the original scene. The H-resolution (high) case occurs when the image pixel size is smaller than the object size so that many pixels will represent the object. For example, a large agricultural field imaged by Landsat OLI would contain many pixels. Similarly, a sugar cane plantation, with fields spanning 200 m by 100 m, imaged by an airborne camera system would contain many pixels per field. When such fields contain a homogeneous land cover, adjacent pixels within the field would be spectrally similar, that is, spatially autocorrelated. The L-resolution (low) case occurs when the image pixel size is larger than the object size so each pixel includes several objects. This case exists in Landsat OLI scale imagery (with 30 m pixels) of urban areas.

Figure 8.6 shows six artificial images of a woodland with decreasing spatial resolution. If a tree crown is the object of interest, Figure 8.6a to Figure 8.6e are examples of the H-resolution case since the pixel size is smaller than the tree crown size, while Figure 8.6f is L-resolution. Despite its 'blockiness', the distribution of individual crowns is still discernible in Figure 8.6e.

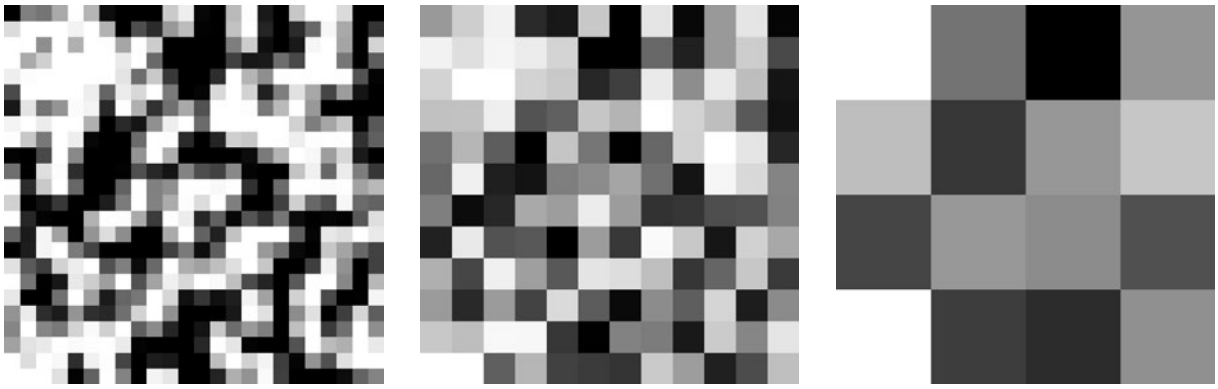
Figure 8.6 Impact of pixel size on feature discernment

In this synthetic image, circles represent tree crowns of 10 m diameter on a contrasting background. The ground area of each image is approximately equal to 120 m x 120 m, that is, the area of 4 x 4 Landsat TM/ETM+/OLI multispectral pixels.

a. 10 cm pixel b. 1 m pixel c. 2.5 m pixel



d. 5 m pixel e. 10 m pixel f. 30 m pixel



Source: David Jupp, CSIRO

The differences between these resolution classes and the properties within them can be determined using spatial statistics. The commonly accepted statistics for analysing image spatial structure and variation are:

- local variance (see Section 8.2.1);
- spatial covariance (see Section 8.2.2); and
- autocorrelation and variograms (see Section 8.2.3).

8.2.1 Local variance

Local variance may be computed for each pixel and averaged over the whole image. If the local neighbourhood of a pixel is defined as those pixels that are adjacent to it, then the values of the pixel and its eight neighbours can be used to calculate a local variance statistic for that pixel. The average of the local variance values over the whole image indicate the extent to which adjacent pixels have similar spectral values in a particular channel.

The effect of changes in pixel resolution on image spatial structure can be determined by computing the local variance values for a set of images with increasing pixel size. A larger pixel size may be simulated for an image using the image blocking technique described in Section 7.2.1.3. The effect of this process is illustrated in Figure 8.6 above.

The average local variance statistics (namely, the mean, standard deviation, minimum, maximum and pixel count) can be computed using a 3×3 pixel neighbourhood centred on every pixel in the image (see Volume 2C for more details of this type of 'filter-based' processing). The local standard deviation (square root of the local variance) is also referred to as the mean 'texture' value and is calculated as:

$$\text{Texture} = \sqrt{\frac{1}{2} \times \frac{\sum \left(x_{ij} - x_{kl} \right)^2}{n}}$$

where

x_{ij} is the central image value at column i and row j;
x_{kl} is the pixel value at column k and row l, where $k = I \pm 1$ and $l = j \pm 1$; and
n is the number of pixels involved in the calculation—usually eight but sometimes less if there are null pixels within the 3×3 window.

Where *n* equals 8, this equation becomes:

$$\sqrt{\frac{\sum\left(x_{ij}-x_{kl}\right)^2}{16}}$$

These spatial statistics may be computed iteratively with image blocking being used in simulate larger pixels between iterations. It is useful to graphically summarise the texture statistics from this iterative sequence by plotting pixel resolution versus local variance as illustrated in Excursus 8.2. The major repeating patterns in the image have been found to occur at a resolution of between two-thirds and three-quarters of the pixel size in the image with maximum variance. This technique has application for determination of the optimal pixel size for studying specific features (Woodcock, 1985).

These statistics may also be used to estimate the values for local pixel-to-pixel uncorrelated noise for each channel. Such noise estimates may then be used as weights in a Principal Components Analysis (see Volume 2C). The local spatial variance can also be used to determine threshold values for image despiking (see Volume 2C) by defining the 'normal' level of local variation in image values. For this purpose, the local standard deviation would be normalised (by dividing by the channel mean) then multiplied by 100 to represent the percentage variation equivalent to one standard deviation. The threshold level would then be some multiple of this to represent an appropriate number of standard deviation units (see Volume 2C).

8.2.2 Variograms

The variogram of an image is the expected squared grey-level difference between two pixels which are a given distance and direction apart (Ahuja and Schachter, 1983). This measure may be estimated using a large number of randomly-selected pixel pairs in the same row or pixel column in the image. The variogram statistic is usually computed for a range of distances, or lags, between the pixel pairs. The variogram is calculated as:

$$\frac{\sum\left(x_i-x_{i+j}\right)^2}{n}$$

where

x_i is the pixel value at column *i*;
x_{i+j} is the pixel value at column *i+j*;
j is the lag distance in pixel columns along an image line; and
n is the number of pixel pairs involved in the calculation.

A similar statistic for quantifying spatial variation within an image is *spatial covariance*. This is calculated as:

$$\frac{\sum\left(\left(x_i-x\right)\times\left(x_{i+j}-\overline{x}\right)\right)}{n}$$

where

$\overline{x}$ is the mean value in the image channel;
x_i is the pixel value at column *i;*
x_{i+j} is the pixel value at column *i+j;*
j is the lag distance in pixel columns along an image line; and
n is the number of pixel pairs involved in the calculation. (Again, n–1 is often used instead of n when the mean value used for channel *x* is only an estimate rather than the real mean value.)

This formula assumes that the mean of the channel is spatially invariant (or 'stationary'). In practical image terms, this would imply that the local mean and variance do not vary systematically over an image. For remotely sensed imagery, variable illumination and reflectance geometry generally invalidate this assumption. The variogram provides a more relaxed measure of spatial variance and has been commonly used in geostatistics for this purpose. When the data are stationary the variogram is related to spatial covariance as:

$$V_j = 2\left(Cov_0 - Cov_j\right)$$

where

j is the distance lag between pixels;
V_j is the variogram for distance lag *j* over the image;
Cov_0 is the spatial covariance for no lag, that is, the variance of the channel; and
Cov_j is the spatial covariance for a distance lag *j* over the image.

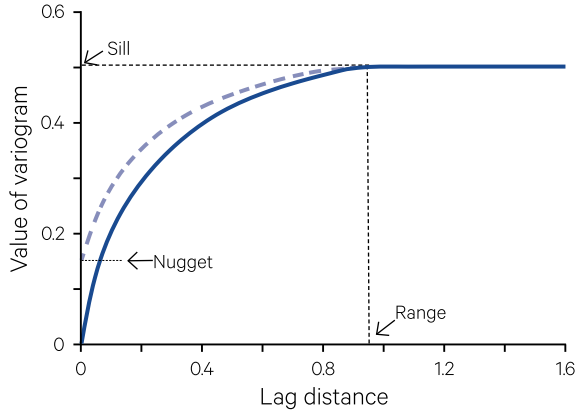
As the name implies, the semi-variogram equals half the variogram:

$$SV = Cov_0 - Cov_j$$

8.2.3 Autocorrelation and variograms

Spatial autocorrelation is another measure of spatial variation. It is computed as the covariance between pixels in a channel divided by the channel variance. These results can be displayed graphically as a 'correlogram' by plotting autocorrelation value versus distance (Cliff and Ord, 1981). This statistic however also assumes a stationary mean value over the channel so it is an inappropriate measure for remotely sensed imagery.

Figure 8.7 One-dimensional variogram



Source: Harrison and Jupp (1990) Figure 46 [Adapted from: Jupp *et al.* (1988a)]

The squared grey-level differences can be computed for pairs of pixels which are a range of distances apart along the same scan line for each line in the image (or set of images). In image processing systems, the distance is often specified as a lag factor, with multiple lags being computed as required. As the length of the lag determines the number of pixel pairs which will be included in the calculation for a given image width, it is advised that the lag distance should be no greater than one tenth of the image line length to retain sufficient confidence in the resulting statistics. These results are best presented graphically by plotting variogram value versus lag distance as shown in Figure 8.7.

Jupp *et al.* (1988a) report there are five especially significant parameters that describe the one-dimensional variogram:

- The 'sill' at which the variogram flattens off to a level equal to twice the regional data variance. (Variogram data are sometimes presented as the 'semi-variogram'—variogram/2—so that the sill height equals the regional variance). The existence and size of the sill are important parameters. The lack of a sill could indicate a trend in the data (such as a scan-angle brightness effect across the image) and/or that the pixel spatial resolution is too small for the scales of pattern in the image (that is, the L-resolution case described above).
- The 'range' value of lag *j* at which the variogram reaches the sill. This parameter is often related to the size or scale of the largest elements (objects) in the scene that produce the correlation structure.
- The 'nugget effect', which, if present, is expressed as a finite limit for the variogram as the lag *j* tends to zero. The variogram value is numerically zero at lag 0, however uncorrelated noise in the data always produces variance. For an image of an uncorrelated noise series, the variogram will have no range and the nugget effect will equal the sill.
- The 'derivative at the origin' is diagnostic of the essential variability of the data at the finest scale. If the limit of the derivative at the origin as lag *j* tends to zero is near zero, the underlying process will be 'smooth'. If the derivative limit is high, the process will be 'rough'. Jupp *et al.* (1988a) have interpreted this parameter as the amount of edge and boundary at the finest scale.
- The 'anisotropy' of the variogram appears if the image is defined in more than one dimension. It reveals a covariance structure such that values tend to be more similar in one or more preferred directions than in others. This will commonly arise in a remotely sensed image when the Sun is not overhead and three-dimensional objects in the ground scene cast shadows (dark shapes) that are elongated in one direction.

Jupp *et al.* (1988b) used an artificial image of a woodland scene to demonstrate how the covariance and variogram statistics are modified. Woodcock *et al.* (1988b) summarise their interpretations of variograms for remotely sensed imagery as:

- The sill is related to the proportion of the area covered by objects, which is a function of their number or density;
- The range is related to the size of objects in the scene. The shape and range of the variogram are more closely related to the area of objects than to their shape;
- The shape of a variogram is related to the variance of the size of objects in the scene. A more rounded or gradually sloping shape is characteristic of higher variance in the size of objects; and
- Increasing the image pixel size has the following effects on variograms:
  - the height of the sill is reduced;
  - the range of influence is increased; and
  - the height of the variogram at the distance equal to one pixel unit increases relative to the sill.

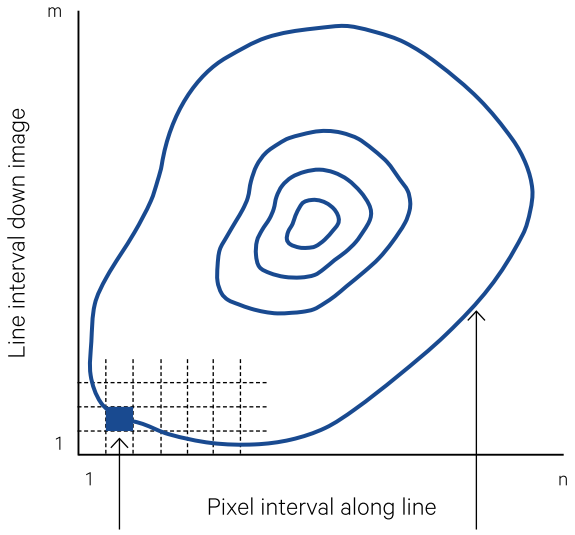
Curran (1988) discusses the use of the semi-variogram (*variogram/2*) in remote sensing analyses and gives examples of computing the spatial variance statistic along transects within specific features. Curran also suggests how spatial statistics can be used to determine an appropriate spatial resolution (for remotely sensed pixels and ground sample plots) for different image features and the number of sample points required to achieve a given confidence level. Spatial statistics of individual image features (such as an agricultural field) can also be determined after isolating the feature using subsetting (see Section 7.2.1.1) and/or segmentation techniques (see Section 10) and possibly re-orienting the resulting image to align image lines with the feature (see Section 7.3 and Volume 2C).

Woodcock *et al.* (1988a) report that one-dimensional variograms are easier to display and interpret, although two-and higher-dimensional displays may also be produced. Variogram statistics computed along an image line are one-dimensional. However, image data could be rotated or transposed to determine these statistics in a different direction (see Section 7.2.2). A two-dimensional variogram is usually displayed as a contour plot as shown in Figure 8.8 and can be useful for revealing anisotropy in data. This plot is constructed by entering the mean squared difference of all pixels pairs that were a particular distance apart in pixels and lines. Contour lines are then drawn between cells with similar values. The interval at which the contour structure breaks down indicates the size of the major repeating patterns in the image (Woodcock, 1985).

Spatial statistics offer an important tool for analysing the scale of an image relative to the objects it represents. The 'optimum' pixel size for image data obviously depends on the cover types being investigated. A finer pixel size is not always preferable since the additional information may simply be increasing the variation within a feature rather than improving its discrimination from other objects or the background. For a woodland area for example, a 2 m pixel would sub-sample individual tree crowns and report changes in density within each crown. If the object of the study is to map woodland as one homogeneous category, this detail will simply complicate the task. A larger pixel size however will integrate multiple crowns into a single pixel size and thereby highlight changes in density of tree cover rather than leaf cover. This integration is most effective for woodland areas when the pixel size is at least 50 m. Thus coarser scale imagery can often resolve broad area composite features such as woodland more effectively than finer scale data.

Figure 8.8 Two-dimensional variogram

The size of the major repeating pattern in the image corresponds to the point at which the contour structure fails.



Source: Harrison and Jupp (1990) Figure 48

The choice of 'optimal' scale for remotely sensed imagery (disregarding the logistical questions of cost and availability) can be summarised by two questions relating to the object of interest:

- Is the feature of interest causing variance between adjacent pixels? If not, is the feature too large or too small relative to the pixel size?
- At what scale does spectral variance in an image relate principally to the object of interest?

The spatial analyses described in this Section attempt to provide answers to these questions. The ability to 'see' features in imagery by visual interpretation often confuses the issue of an appropriate scale for digital data. The type of analyses undertaken in digital image processing requires different criteria for the selection of optimal image resolution for a particular application.

Excursus 8.2—Spatial Statistics

Source: David Jupp, CSIRO

Three example images are illustrated in Figure 8.9. These images were artificially created to represent an area of 80 m square with 50% cover (that is, half the area is covered by a single 'type' of object and the other half is background). The object values are distributed to represent three scales of pattern:

- disks of 8 m diameter;
- disks of 1 m diameter; and
- random distribution.

The centres of the disks are randomly (or Poisson) distributed. The random distribution of disks is essentially objects with diameter a little smaller than a pixel.

In these images, every pixel either has value 0 (background) or 1 (disk) and the proportion of disk is simple the image mean. In general, if the cover of disks is q, then the image mean is q for all cases (random, 1 m or 8 m disks), and the image variance is:

$$\begin{aligned} \text{var} &= \frac{1}{N}\sum_{j=1}^{N}(x_j - q)^2 \\ &= q - q^2 \\ &= q(1-q) \end{aligned}$$

Since the image standard deviation is the square root of the image variance, for 50% cover, q = 0.5, var = 0.25 and standard deviation = 0.5. As occurs with the mean, the standard deviation is the same for all three cases. So the overall mean and standard deviation do not respond at all to the different disk sizes.

In an EO image of a ground scene, the disks could be considered to be representative of tree crowns so that the large and small disks would be like simplified images of woodland vegetation with big and little trees respectively. Each image is 512 pixels by 512 lines, and as it covers an area of 80 m by 80 m, the pixel size is equivalent to 15.625 cm.

The images were blocked four ways to produce images at four new resolutions:

- 2×2: 31.25 cm pixel;
- 4×4: 62.5 cm pixel;
- 8×8: 125 cm pixel; and
- 16×16: 250 cm pixel.

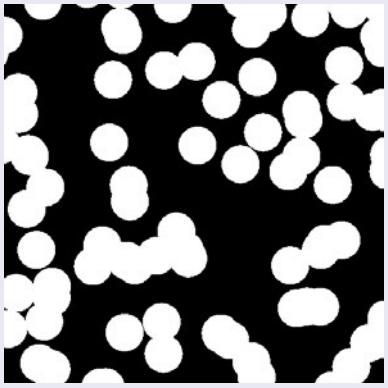
Then, local variance statistics were computed for the five scales (original plus four blockings listed above).

The mean local standard deviation, or texture, can be computed as the average of the square root of local variance (as detailed in Section 8.2.1 above). Figure 8.10 represents the results of the average function graphically. For the large tree image, the maximum pixel size (250 cm) is much smaller than the object size (8 m) so the variance is still increasing at a blocking factor of 16. However, in the small tree image the variance peaks at a pixel size around 62.5 cm, which is between one half and three-quarters of the object size (1 m). The object size in the random image is effectively one pixel, so the variance decreases rapidly with blocking. Unlike the overall mean and variance, the three disk models therefore have very different behaviour with block variance.

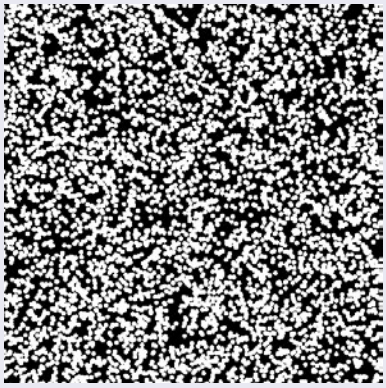
Figure 8.9 Artificial example images

These three artificial images comprise 512 x 512 square pixels and cover an area of 80 m square. Percent cover within each image is 50%.

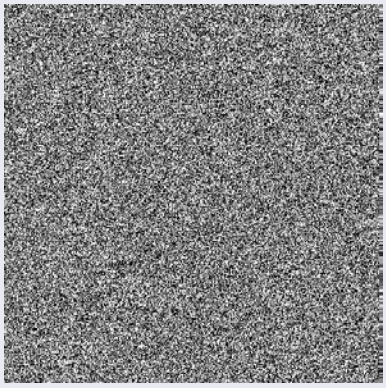
a. Large disks (8 m diameter)



b. Small disks (1 m diameter)



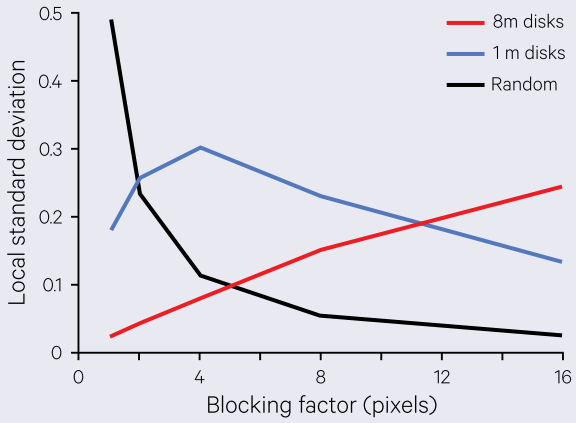
c. Random image



Source: David Jupp, CSIRO

Figure 8.10 Local standard deviation results derived using average function

Object size for large disks (see Figure 8.9) is much larger than the pixel size so variance continues to increase over the distance range used in this example. With the small disk image, variance peaks at a distance approximately equivalent to two-thirds of the object size. The pixel size effectively equals object size in a random image so variance only decreases.



Source: Harrison and Jupp (1990) Figure 45

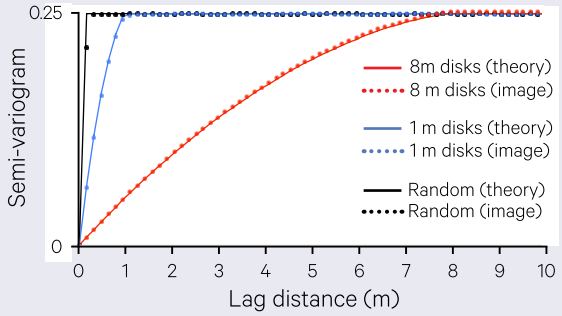
The semi-variogram statistics also show different behaviours for the three disk models. The three artificial images in Figure 8.9 were processed with lags from 1 pixel up to 52 pixels (or 8.125 m). Theoretical model results (Jupp *et al.*, 1989) for the equivalent continuous functions were also computed.

These results are plotted in Figure 8.11 and show good agreement. In this example, given that the cover fraction (q) is 50%, the sill of the variogram occurs at value 0.25 or, as shown above, $q(1-q)$.

Accordingly to theory (Jupp *et al.*, 1989), the range values correspond to disk size, with the curve for the 1 m disk image reaching the sill at a lag distance of 1 m and the curve for the 8 m disk image reaching the sill at a lag distance of 8 m. The random image has a much narrower range as it is effectively a disk of approximately diameter 15 cm (one image pixel).

Figure 8.11 Variogram results for example images

Semi-variogram values are plotted against lag distances to indicate spatial variance of three images with different object sizes and distributions (see Figure 8.9). The curves derived from images with 8 m and 1 m disks match theoretical results.



Source: David Jupp, CSIRO

8.3 Radiometric

In the context of EO imagery, the term 'radiometric statistics' is sometimes used to refer to variations in radiance values resulting from detector inconsistency in both image data values and calibration measurements (see Section 3.2.3). In this Section, we use the term to embrace all variations in EO imagery that distort true surface reflectance measurements (see Section 3.3.2).

The process of EO image acquisition is detailed in Volume 1. This process involves an EO sensor, carried on a travelling platform, observing the undulating surface of the Earth through a changing atmosphere, with a moving illumination source. While the goal of EO is to quantify surface reflectance, the actual process of image acquisition invariably results in measurements that include factors related to atmospheric composition, surface illumination and instrument operation. The major factors that interfere with measurement of surface reflectance during EO image acquisition can be summarised as:

- EMR interactions with target and sensor (see Section 8.3.1);
- EMR interactions with the atmosphere (see Section 8.3.2); and
- sensor operation (see Section 8.3.3).

A wide variety of approaches are used to correct EO imagery for these factors. While most implementations of these methods attempt to correct for all of the major factors in a single, integrated model, we discuss the three major factors individually below. An example of radiometric calibration applied to a continental scale dataset is given in Excursus 8.3.

8.3.1 EMR interactions with target and sensor

As discussed in Volume 1B—Section 3, the radiance measured by EO sensors is impacted by several interactions that can occur between the incident radiance and the imaging target. These interactions also vary with wavelengths of the EMR energy, the texture and type of surface materials, and the three-dimensional relationship between illumination source, target and sensor. Cumulatively, these interactions result in quantifiable variations in 'scene brightness' in EO imagery (see Appendix 1).

In this context, 'correction' of scene brightness variations changes the spectrum of values for each image pixel to what would have been measured by the EO sensor if the geometry of the illumination source and sensor were 'fixed', rather than varying over the image extent. Two corrections that are often applied to EO imagery attempt to compensate for:

- surface orientation (see Volume 1B—Section 3.5); and
- anisotropy (the differences in geometrical reflective properties of targets when they are measured in different directions; see Volume 1B—Section 3.6).

Surface orientation is generally related to surface topography, but can also vary with sensor view angle. Terrain illumination correction uses geometrical parameters derived from accurately registered Digital Surface Models (DSM) to modify topographic shading in imagery (Li *et al.*, 2012). These algorithms are limited both by the resolving power of the EO data with respect to hill slopes and the accuracy of the elevation data (in all dimensions).

Anisotropy can be quantified by the Bidirectional Reflectance Distribution Function (BRDF; Nicodemus *et al.*, 1977). BRDF correction reduces the differences due to varying Sun and sensor positions in EO images, and can be applied to the variation within a single image and across a set of images. Some of the models that have been used for BRDF correction are detailed in Appendix 2. Accuracy of these models depends on the spatial resolution of the available calibration data.

The three main approaches to modelling scene brightness, namely empirical models, measurement models and semi-empirical models, are detailed in Appendix 1.

8.3.2 EMR interactions with the atmosphere

Radiance detected by EO sensors includes effects due to EMR scattering and absorption by atmospheric gases, aerosols and/or water vapour (see Section 3.2.3 above and Volume 1B—Section 4). A range of methods are used to correct for atmospheric path radiance in EO imagery, with variable success. Commonly used methods include:

- subtract 'dark values'—determining the image value corresponding to a very dark object with virtually no discernible radiance, such as a deeply shaded hillside or valley. Any image values below the dark value are assumed to represent atmospheric radiance alone. The image data can then be rescaled to only contain values related to surface reflectance (see Volumes 2C and 2D for details).
- atmospheric normalisation—between two or more images can be based on a relative or standardised reference (Coppin *et al.*, 2004). Some methods that have been proposed for atmospheric normalisation of EO image sequences include:
  - simple regression (SR)—matching reflectance values for the full image scene to a known reference using spectral channel crossplots, then use the regression equation to rescale image values (Jensen, 1983; see Volume 2C);
  - pseudo-invariant feature (PIF)—selects landscape features with near-constant reflectance through time and rescales image ranges so that all features have comparable values in all images (Hall *et al.*, 1991a, Furby and Campbell, 2001). An automated version of this approach called Automatic Scattergram-Controlled Regression (ASCR) uses scattergrams (or channel crossplots; see Section 8.1) to select pixels with near-constant reflectance (Elvidge *et al.*, 1995); and
  - radiative transfer models (RT)—simulate the impact of the atmosphere on reflected or emitted radiance using parameters derived from imagery and field data (Ahern *et al.*, 1988; see Volume 1B—Section 5 and Appendix 1).

Radiative transfer models such as MODTRAN (Berk *et al.*, 2003) account for Mie, Rayleigh and aerosol scattering effects and atmospheric gas absorption using estimates of atmospheric optical depth (visibility), water vapour, and concentrations of major gases (such as ozone and carbon dioxide). Such models estimate a range of optical parameters, such as transmittance (for Sun/sensor directions), path radiance, atmospheric albedo, and the proportion of diffuse to total irradiance (see Appendix 1). Their accuracy is strongly dependent on the quality of available aerosol data to determine the atmospheric profile at the time of image acquisition. Such methods are commonly used to correct hyperspectral imagery where estimates of atmospheric water vapour can be derived from image spectra.

8.3.3 Sensor operation

Imaging sensors can introduce a range of radiometric distortions as reviewed in Section 3.2.3. These sensor-related distortions can be:

- random—such as missing or spurious pixels or lines;
- systematic—such as multiple detectors recording inconsistent measurements; or
- gradual—such as detector sensitivity deteriorating with time.

The impact of such distortions is removed or reduced by the process of radiometric calibration (see Section 3.3.2). EO imagery can be radiometrically calibrated to either an absolute or relative standard.

Absolute calibration relates image values directly to a known reference and enables pixel observations to be converted into radiance measurements. Such methods require precisely calibrated ground reference points which, ideally, are simultaneously measured during image acquisition, and allow removal of calibration differences between detectors and channels. A range of methods have been proposed for absolute calibration of various sensors, such as Teillet *et al.* (1990) for NOAA AVHRR, Slater *et al.* (1996) for EOS sensors, and Teillet *et al.* (2001) using hyperspectral data for multiple sensors.

Relative calibration simply attempts to equalise the variation between multiple detectors. For example, whiskbroom sensors operate by recording multiple lines with different detectors in each imaging scan (see Volume 1A—Section 14). While efforts are made to calibrate these detectors, sensitivity differences still exist and result in a characteristic multi-line, horizontal striping pattern in the imagery (see Section 3.2.3 for details). Similarly, pushbroom scanners can introduce vertical striping patterns into imaged scenes when individual detectors along a scan line deliver different radiance values for the same target reflectance (see Figure 3.6).

One relative method for 'destriping' images with horizontal or vertical striping involves equalising the mean and variance values for the set of detectors. If the mean and variance values of each detectors were accumulated separately, then the image could be rescaled so each detector has a mean and variance equal to either the average values for the whole image or to a reference sensor. A similar method, known as 'histogram equalisation', matches the histograms of each detector to a selected reference histogram (such as the mean). These processes are further discussed in Volume 2C.

Excursus 8.3—Continental Scale Calibration of ASTER imagery

Source: Michael Caccetta, CSIRO
Further Information: Caccetta *et al.* (2013)

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), launched in 1999, acquired global imagery with particular value for geological applications. Over Australia, this imagery enabled continental scale mapping of surface mineral abundance and composition (see Volume 1B—Excursus 6.1).

3,600 scenes of ASTER L1B radiance at sensor data, acquired between May 2000 and March 2008, were converted into a single, calibrated mosaic of apparent surface reflectance by correcting each scene for effects due to both instrument characteristics and atmospheric/illumination characteristics. This extensive dataset included approximately 400 different dates with significant radiometric variation occurring between adjacent scenes (see Figure 8.12a, Figure 8.12c and FFigure 8.12e). To generate a calibrated reflectance dataset, processing required use of:

- new statistical cross-calibration procedures, which compensated for between-scene variations, to create a single mosaic; and
- an extensive set of independent, calibrated reflectance imagery (Hyperion satellite imagery) to convert the mosaic to apparent reflectance (Caccetta *et al.*, 2013).

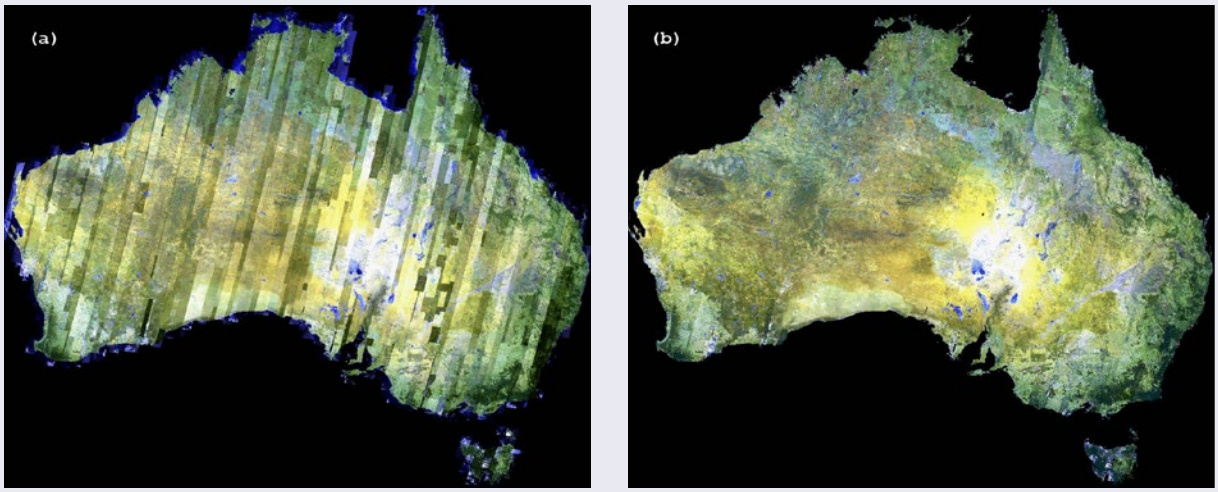
For ASTER optical bands, the seamless, continental mosaic resulting from the cross-calibration process is illustrated in Figure 8.12b, Figure 8.12d and Figure 8.12f. This cross-calibrated mosaic was then converted to apparent reflectance by determining a linear transformation based on robust least squares regression between the ASTER mosaic and a set of 18 co-located Hyperion scenes. A similar approach was used to generate a continental emissivity mosaic from ASTER L2 surface emissivity data.

From these two continental mosaics (reflectance and emissivity), a suite of geoscience information products have been derived, primarily indicating mineral group content and composition (for example, see Volume 1B—Excursus 6.1). These geoscience maps are available from: *http://doi.org/10.4225/08/51400D6F7B335*.

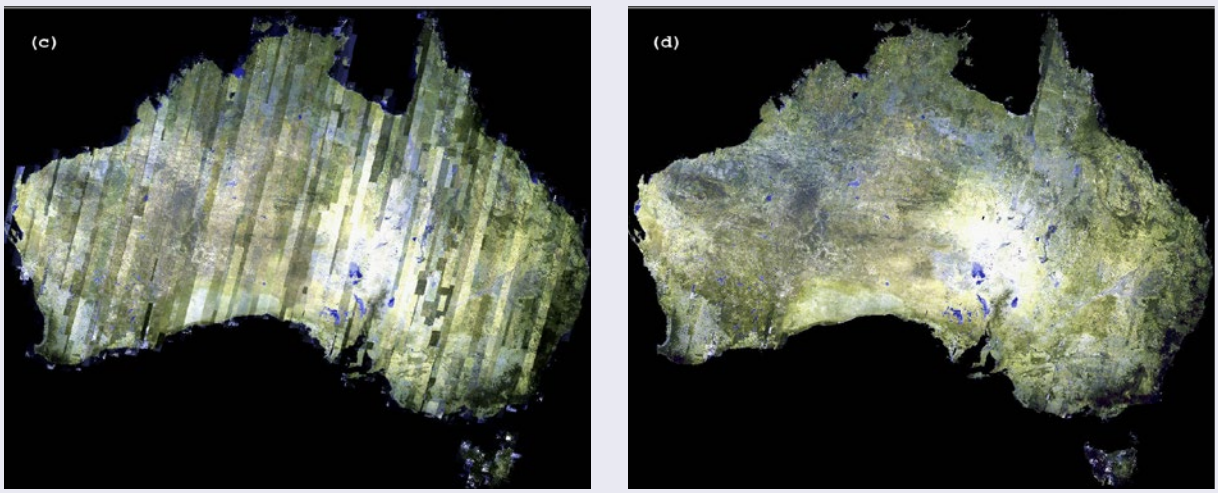
Figure 8.12 Continental mosaic of ASTER imagery

3,600 ASTER L1B scenes (a, c and e) were cross-calibrated to create a seamless continental mosaic of apparent reflectance (b, d and f).

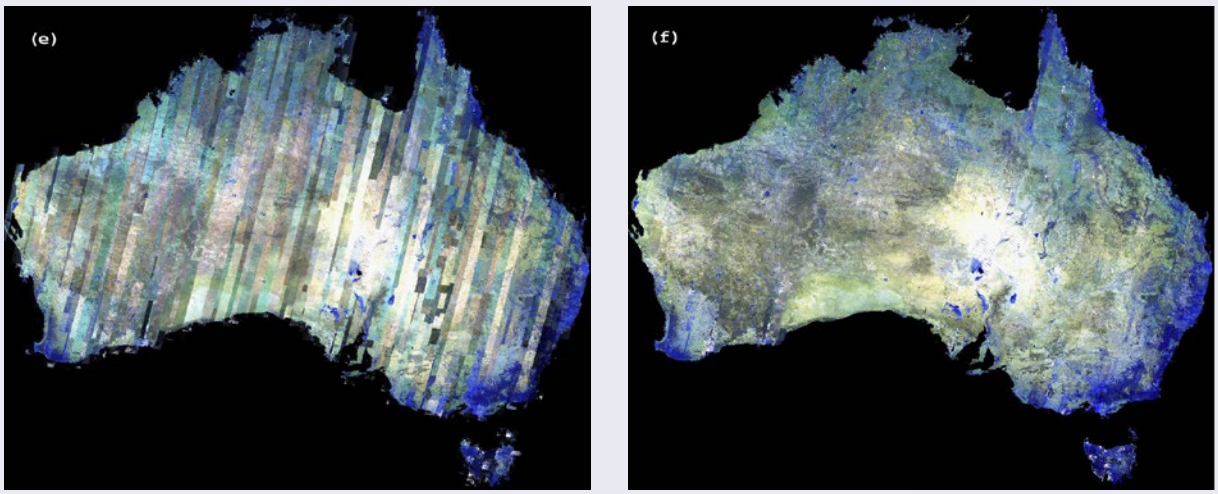
a. and b. ASTER bands 7 (SWIR), 4 (SWIR), 1 (green) shown as RGB



c. and d.ASTER bands 8 (SWIR), 5 (SWIR), 2 (red) shown as RGB



e. and f. ASTER bands 9 (SWIR), 6 (SWIR), 3 (NIR) shown as RGB



Source: Michael Caccetta, CSIRO

8.4 Temporal

Time series datasets differ from many other sets of measurements in that they are assumed to represent a non-random sequence. By definition, a time series dataset comprises consistent measurements taken at equally spaced time intervals. Image time series, then, represent a sequence of two (or more)-dimensional measurement layers, or multivariate time series, whose values are consistent in both space and time. For example, Figure 8.13 shows a temporal profile plot of mean chlorophyll concentration near the mouth of the River Murray, as estimated from MODIS imagery.

Most time series analyses aim to identify those factor(s) that drive changes in the data series measurements and, based on this understanding, predict future values in the series. Each of these goals requires that the underlying pattern(s) in the data be described precisely. The identified patterns can then be related to their drivers and extrapolated to future scenarios.

It is assumed that the measurements in time series datasets include both systematic pattern (often resulting from identifiable components) and random noise (or error). The latter generally obscures the systematic pattern to some extent and a range of filtering or screening algorithms are commonly used to clarify the systematic data content. The underlying systematic pattern results from both seasonal (repeating) trends and longer term trends. Thus, where the time series is the sum of these three components, it can be represented by the equation:

$$X_t = m_t + s_t + Y_t$$

where

X_t is the time series;
m_t is the long term trend;
s_t is the seasonal trend; and
Y_t is random noise.

In some time series, a multiplicative model may be more appropriate to represent the interactions between components:

$$X_t = m_t \times s_t \times Y_t$$

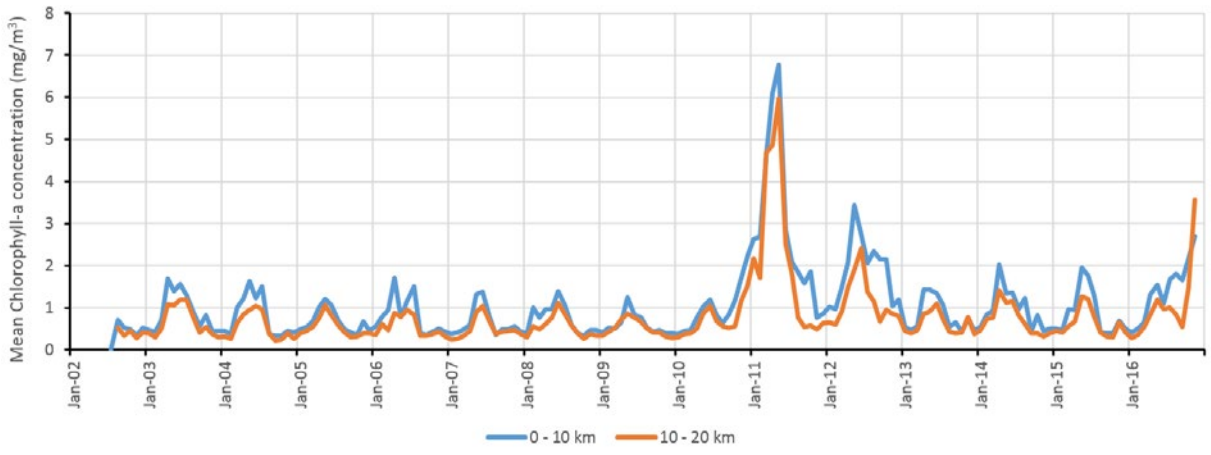
In such cases, logarithmic transformations are used to convert the data into a linear space so that an additive model can be fitted (see Volume 2C).

Where the number of equally spaced observations in a time series is infinite, it can be viewed as a stochastic process for statistical purposes. A time series is considered stationary when its statistical properties are time-invariant, that is the mean, variance and covariance are constant. The goal of time series analyses is to extract the long term (m_t) and seasonal trends (s_t) so that the remaining noise component (Y_t) is stationary.

Many time series datasets demonstrate a seasonal trend based on annual cycles (whether driven by climatic or economic events) and a longer term trend that carries the mean of the seasonal trend up or down over time. There are many examples of time series datasets in which the general shape and duration of the seasonal shape is discernible from year to year, but the amplitude increases (or decreases) with the longer-term trend (see Figure 8.14).

Figure 8.13 Trends in chlorophyll-a concentration

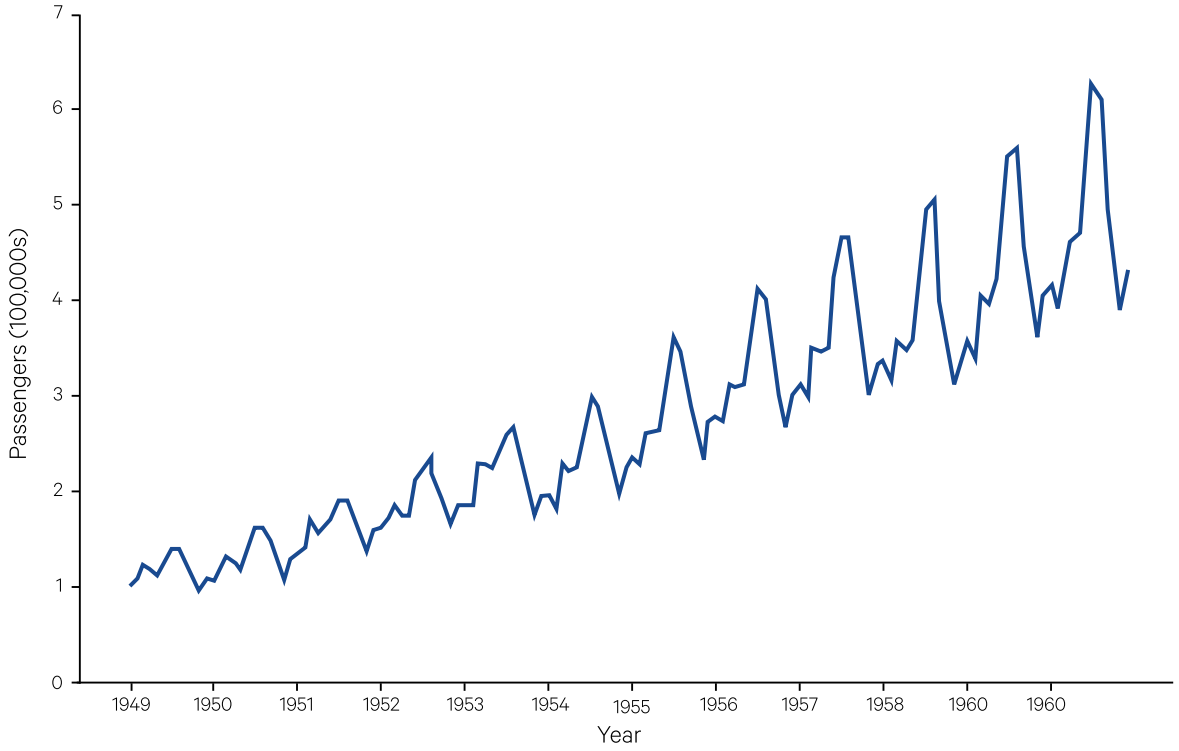
Temporal profile of ocean chlorophyll-a concentration near the mouth of the River Murray, in Australia, as measured from MODIS Aqua Level 3 chlorophyll a data product. The graph shows that high (2011) and normal (2012) river flows result in substantial increases in ocean chlorophyll a concentrations at least 20 km from shore, as compared to the concentrations observed in very low flow years (2002–2010).



Source: Ken Clarke, University of Adelaide (adapted from data underlying Figure 8 in Auricht *et al.*, 2017)

Figure 8.14 Time series example

This example, based on international airline passenger numbers from 1949 to 1960 inclusive, shows both seasonal trends with each year and a longer term over the 12 year period.



Adapted from: *http://www.statsoft.com/Textbook/Time-Series-Analysis#identifying* [Derived from Box and Jenkins (1976)]

Trend analysis requires some form of local averaging of the dataset to eliminate non-systematic components, such as median smoothing (see Volume 2C), before representing the trend as a mathematical function. Standard time series methods determine the trend first, then use that relationship to focus on seasonal perturbations. Time series models can be assessed by the randomness of their error distribution and/or predictive ability.

Seasonality analysis identifies the seasonal dependency, or correlation, of a measurement on preceding measurements. This characteristic is called serial dependence or serial correlation and can be examined using autocorrelation techniques. Once adjusted for the long term trend, standardised seasonal averages should sum to zero. Seasonality can be removed from a time series by block averaging or seasonal differencing (Reinert, 2010).

The traditional approach to statistical analysis of time series datasets can be summarised in terms of three stages:

- identification—pre-processing of data to achieve stationarity. This requires an initial assessment of the data to understand the overall trend, the periodicity and the likely relationship (additive or multiplicative) between components. This assessment should determine whether the data need to be differenced or transformed.
- estimation—of seasonal and longer term trends; and
- verification—by overfitting or residual analysis (Box and Jenkins, 1970).

Autoregressive (AR) processes assume that any measurement in a time series can be estimated from a linear combination of specific, preceding measurements plus an error factor. This requires that the series be stationary, and a stationary series is not serially dependent. Stationarity is generally achieved by differencing each measurement in the series with that prior measurement that is offset in time by a selected lag interval. This process can be considered as a type of linear filter. Log transformation may also be required to stabilise variance.

Moving average (MA) processes assume that each observation involves random error plus the cumulative impact of prior errors. With certain conditions, the moving average equation can be inverted into an autoregressive form. Moving averages can be computed from the linear combination of measurements within a specified time interval, that is:

Moving average for measurement

$$\text{at time } \frac{k}{2} = \sum_{i=0}^{k-1} \frac{\text{value}_{k-i}}{k}$$

Alternatively, an exponentially weighted moving average can be used which weights 'current' time measurements more heavily. An AutoRegressive Moving Average process (ARMA) effectively combines both processes.

Where observations in the time series are incomplete or noisy and result from an unobservable process, Markov models may be more appropriate for analysing trends. Discrete state variables, such as EO time series, can be modelled using hidden Markov Models (Rabiner, 1989; Viovy and Saint, 1994). Examples of studies focussing on EO time series include Scarth *et al.* (2015), Schmidt *et al.* (2016) and Frantz *et al.* (2016). Analysis of EO time series datasets is considered in detail in Volume 2D—Section 3.

8.5 Further Information

Time Series Analysis Tools for Satellite Imagery:
TimeSat—*http://web.nateko.lu.se/timesat/timesat.asp*
Google Earth Engine—*https://earthengine.google.com/*

SEOS: eLearning Tutorials: *http://www.seos-project.eu/modules/*

LEOWorks—Image Processing Tutorials: *http://leoworks.terrasigna.com/files/LEOWorks4_beta_manual.pdf*

8.6 References

Ahern, F. J., Brown, R. J., Chilar, J., Gauther, R., Murphy, J., Neville, R. A., and Teillet, P. M. (1988). Radiometric correction of visible and infrared remote sensing at the Canada Centre for Remote Sensing. '*Remote Sensing Yearbook 1988/89*' (Eds: A. Cracknell, and L. Hayes), pp. 101-127. Taylor and Francis, London, United Kingdom.

Ahuja, N., and Schachter, B. J. (1983). *Pattern Models*. John Wiley and Sons, New York.

Auricht, H. C. C., Clarke, K. D., Lewis, M. M., and Mosley, L.M. (*In press*). Have droughts and increased water extraction from the River Murray (Australia) reduced coastal ocean productivity? *Marine and Freshwater Research,* 68, pp. 1–14.

Berk, A., Anderson, G. P., Acharya, P. K., Hoke, M. L., Chetwynd, J. H., Bernstein, L. S., Shettle, E. P., Matthew, M. W., and Adler-Golden, S. M. (2003). *Modtran 4 Version 3 Revision 1 User's manual*. Airforce Research Laboratory, Hanscom, United States.

Box, G., and Jenkins, G. (1970). *Time series analysis: Forecasting and control*. Holden-Day, San Francisco, United States.

Caccetta, M., Collings, S., and Cudahy, T. (2013). A calibration methodology for continental scale mapping using ASTER imagery. *Remote Sensing of Environment*, 139, pp. 306–317.

Cliff, A. D., and Ord, J. K. (1981). *Spatial Processes: Models and Applications*. Pion, London.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9), pp. 1565-1596. doi:*http://dx.doi.org/10.1080/0143116031000101675*.

Curran, P. J. (1988). The semivariogram in remote sensing: An introduction. *Remote Sensing of Environment*, 24(3), pp. 493-507. doi:*http://dx.doi.org/10.1016/0034-4257(88)90021-1*.

Elvidge, C. D., Yuan, D., Weerackoon, R. D., and Lunetta, R. S. (1995). Relative radiometric normalization of Landsat Multispectral Scanner (MSS) data using a automatic scattergram-controlled regression. *Photogrammetric Engineering and Remote Sensing*, 61(10), pp. 1255-1260.

Frantz, D., Röder, A., Udelhoven, T., and Schmidt, M. (2016). Forest Disturbance Mapping Using Dense Synthetic Landsat/MODIS Time-Series and Permutation-Based Disturbance Index Detection. *Remote Sensing*, 8(4), pp. 277. doi:*http://dx.doi.org/10.3390/rs8040277*.

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Jensen, J. R. (1983). Urban/suburban land use analysis. In '*Manual of Remote Sensing*' (Ed: R. N. Colwell). Second Edn. American Society for Photogrammetry and Remote Sensing (ASPRS); John Wiley & Sons, Inc., Falls Church, Virginia.

Jupp, D. L. B. (1988a). *Fastrack Classification*. Paper presented at the First microBRIAN User Group meeting, Brisbane, Australia.

Jupp, D. L. B. (1988b). *The mosaic model for microBRIAN Classification*. Paper presented at the First microBRIAN User Group meeting, Brisbane, Australia.

Jupp, D. L. B., Mayo, K. K., Harrison, B. A., Hutton, P. G., and Ahmad, W. (1988a). Remote sensing and image processing on personal computers : the CSIRO (Australia) microBRIAN system. In '*Desktop Planning : Advanced Microcomputer Applications for Physical and Social Infrastructure Planning*'. Hargreen Pub. Co., Melbourne, Australia.

Jupp, D. L. B., Strahler, A. H., and Woodcock, C. E. (1989). Autocorrelation and regularisation in digital images. II. Simple image models. *IEEE Trans. Geosci. And Remote Sensing*, 27, pp. 247–258.

Jupp, D. L. B., Walker, J., Harrison, B. A., and Reece, P. (1988b). *The Effects of Vegetation Structure in Remotely Sensed Data and its Implications for Operational Monitoring of Large-scale Environmental Changes in Kakadu National Park*. CSIRO Division of Water Resources Research, Canberra, Australia.

Levine, M. D. (1985). *Vision in Man and Machine*. McGraw-Hill Inc., USA.

Li, F., Jupp, D. L. B., Thankappan, M., Lymburner, L., Mueller, N., Lewis, A., and Held, A. (2012). A physics-based atmospheric and BRDF correction for Landsat data over mountainous terrain. *Remote Sensing of Environment*, 124, pp. 756-770. doi:*http://dx.doi.org/10.1016/j.rse.2012.06.018*.

Nicodemus, F. E., Richmond, J. C., and Hsia, J. J. (1977). *Geometrical considerations and nomenclature for reflectance* (NBS Monograph 160). US Department of Commerce, National Bureau of Standards.

Pickup, G., and Chewings, V. H. (1988). Forecasting patterns of soil erosion in arid lands from Landsat MSS data. *International Journal of Remote Sensing*, 9(1), pp. 69-84. doi:*http://dx.doi.org/10.1080/01431168808954837*.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), pp. 257-286. doi:*http://dx.doi.org/10.1109/5.18626*.

Reinert, G. (2010). *Time Series*. Department of Statistics, University of Oxford. Retrieved from *http://www.stats.ox.ac.uk/~reinert/time/notesht10short.pdf*.

Scarth, P., Armston, J., Flood, N., Denham, R., Collet, L., Watson, F., Treyithick, B., Muir, J., Goodwin, N., Tindall, D., and Phinn, S. (2015). Operational Application of the Landsat Timeseries to Address Large Area Landcover Understanding. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci, XL-3/W3,* pp. 571–575. doi:*http://dx.doi.org/10.5194/isprsarchives-XL-3-W3-571-2015*.

Schmidt, M., Pringle, M., Devadas, R., Denham, R., and Tindall, D. (2016). A Framework for Large-Area Mapping of Past and Present Cropping Activity Using Seasonal Landsat Images and Time Series Metrics. *Remote Sensing*, 8, pp. 312. doi:*http://dx.doi.org/10.3390/rs8040312*.

Slater, P. N., Biggar, S. F., Thome, K. J., Gellman, D. I., and Spyak, P. R. (1996). Vicarious Radiometric Calibrations of EOS Sensors. *Journal of Atmospheric and Oceanic Technology*, 13(2), pp. 349-359. doi:*http://dx.doi.org/10.1175/1520-0426(1996)013%3C0349:vrcoes%3E2.0.co;2*.

Teillet, P. M., Fedosejevs, G., Gauthier, R. P., O'Neill, N. T., Thome, K. J., Biggar, S. F., Ripley, H., and Meygret, A. (2001). A generalized approach to the vicarious calibration of multiple Earth observation sensors using hyperspectral data. *Remote Sensing of Environment*, 77(3), pp. 304-327. doi:*http://dx.doi.org/10.1016/S0034-4257(01)00211-5*.

Teillet, P. M., Slater, P. N., Ding, Y., Santer, R. P., Jackson, R. D., and Moran, M. S. (1990). Three methods for the absolute calibration of the NOAA AVHRR sensors in-flight. *Remote Sensing of Environment*, 31(2), pp. 105-120. doi:*http://dx.doi.org/10.1016/0034-4257(90)90060-Y*.

Viovy, N., and Saint, G. (1994). Hidden Markov models applied to vegetation dynamics analysis using satellite remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 32(4), pp. 906-917. doi:*http://dx.doi.org/10.1109/36.298019*.

Walpole, R. E. (1974). *Introduction to Statistics*, Second Edn. Macmillan, New York.

Woodcock, C. E. (1985). *Understanding Spatial Variation in Remotely Sensed Imagery*. (PhD thesis), University of California.

Woodcock, C. E., and Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, 21(3), pp. 311-332. doi:*http://dx.doi.org/10.1016/0034-4257(87)90015-0*.

Woodcock, C. E., Strahler, A. H., and Jupp, D. L. B. (1988a). The use of variograms in remote sensing: I. Scene models and simulated images. *Remote Sensing of Environment*, 25(3), pp. 323-348. doi:*http://dx.doi.org/10.1016/0034-4257(88)90108-3*.

Woodcock, C. E., Strahler, A. H., and Jupp, D. L. B. (1988b). The use of variograms in remote sensing: II. Real digital images. *Remote Sensing of Environment*, 25(3), pp. 349-379. doi:*http://dx.doi.org/10.1016/0034-4257(88)90109-5*.

9 Classification

Image classification is commonly used for feature or pattern recognition and usually involves converting a continuous tone image into a thematic map. As with most image processing tools, the suitability of image classification for a particular interpretation task will depend on the nature of the application and the resolution of features to be identified in the image. Land cover mapping exercises are frequently based on image classification since all the available channels in an image can help to discriminate between different cover types. Geological analyses, however, often require delineation of linear features so may be based on enhancement techniques rather than classification.

As discussed in Volume 1, different land cover features reflect and emit electromagnetic energy differently and are thus recorded as different spectral values in an EO image. Spectral pattern recognition involves identifying typical spectral values for the various image features and grouping together pixels that have values similar to the feature values. The process of image classification implicitly involves two sets of classes:

- image classes, defined by spectral and/or spatial attributes of one or more image channels; and
- interpreter categories, which are related to ground features in the imaged scene.

Image classification attempts to define a set of image classes that represent interpreter categories as closely as possible. In broad terms, this exercise involves a number of logical stages:

- characterising image classes (see Section 9.1);
- allocating each pixel to an image class (see Section 9.2);
- grouping and labelling image classes to represent interpreter categories (see Section 9.3); and
- assessing the accuracy of labelled interpreter categories (see Section 9.4).

While these logical stages are listed as sequential steps, in practice two or more stages may be integrated into a single image processing algorithm, and such algorithms can differ between image processing systems. Also, 'recipes' for classification can imply that it is a sequential process, however in practice several of the stages involved are iterative—classes are defined, pixels allocated to classes, allocations are assessed then more classes may be defined or some may be merged (see Table 9.1). This iterative nature may be time-consuming, but is essential for a reliable and informative result from classification of EO imagery. The interpreter is encouraged to view the whole process as one of becoming acquainted with the image data and understanding what it has to offer to a particular application. The term 'exploratory data analysis' embraces this spirit of enquiry.

It is once again the vexing problem of identity within variety; without a solution to this disturbing problem there can be no system, no classification.
(Roman Jakobson)

Background image: ISODATA classification of Landsat-8 image over Weipa, Queensland, into 10 land cover classes using all optical bands. The original image is shown in the banner of Section 8. **Source:** Norman Mueller, Geoscience Australia

Table 9.1 The classification process

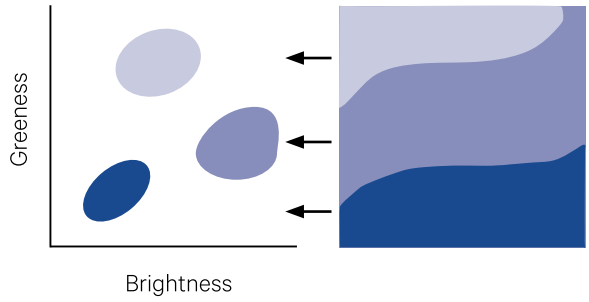
Logical Stage	Potential sub-stages	Example Process	If results of sub-stage not satisfactory
Characterise image classes	Pre-process to highlight feature(s) being classified	Despiking; greenness index; Principal Components Analysis (PCA); image segmentation	Trial other processing options
	Select most informative channels	Feature selection	Remove redundant channels or trial different channel set
	Define initial image classes	Interactive training patches; auto-seeding	Select more classes, possibly using different method(s)
Allocate pixels to image classes	Select and apply allocation algorithm	Minimum Distance classifier; Maximum Likelihood classifer	Trial different allocation algorithm(s)
	Check that final set of classes represents image variation	Create residual image	Generate more classes, possibly using additional (or different) channels, then re-allocate pixels to classes
	Check that class seeds represent all pixels allocated to them	Mean migration; spectral transfer	Recompute class seeds or generate additional classes
Group and label image classes to represent interpreter classes	Analyse class separation using spectral and spatial tools	Canonical Variates Analysis (CVA); co-occurrence matrix analysis	Remove redundant classes and split diverse classes then re-allocate pixels to classes
	Group similar image classes to represent interpreter categories using statistical tools	Spatial and spectral MST; dendograms	Review image classes and interpreter categories
	Select appropriate colour sequences to show grouped interpreter classes	Minimum Spanning Tree; interactive painting	Review available colour systems
Verify interpreter categories with independent information	Select sample pixels in image and reference data set and cross-tabulate	Interactive training; automatic algorithms to select appropriate samples; locate reference sample points in image	Select more samples for poorly represented classes
	If interpreter categories are available as a reference image channel, cross-compare classification channel with reference channel	Channel cross tabulation algorithms	Refine classification channel or reference channel
	Analyse cross tabulation(s) for level of agreement	Contingeny table manipulation; Kappa statistics	Select more samples and/or refine classification

The class development phase is an important preliminary step but is often misunderstood. Its aim is to represent the spectral variation in the image—not necessarily to build ground feature categories as such. Although training patches may be taken with ground features in mind, the classification process simply uses them to represent the structure of the image rather than an interpreter's view of the structure of the imaged scene. Thus, it is important to differentiate between image classes, which can be characterised by their spectral and/or spatial attributes, and interpreter categories, which are quite independently based on Earth surface features.

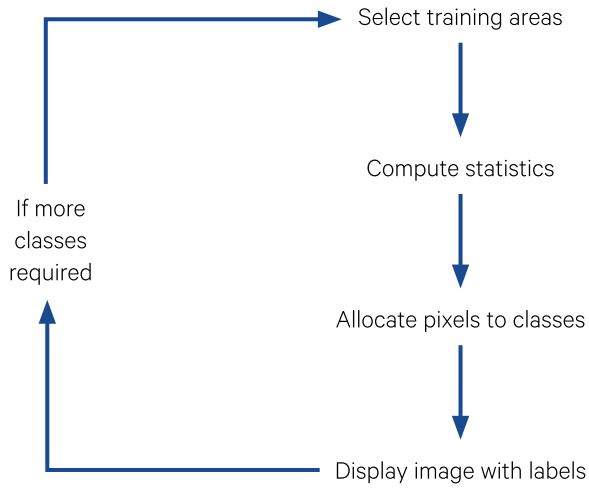
'Supervised' classification procedures have traditionally required sets of sample values, which represent the various ground features in the image (such as a particular vegetation or soil type). The sample values are then used to identify pixels with similar values on the assumption that these pixels are likely to be the same ground feature categories (see Figure 9.1). This approach is valid when the interpreter is sure of the number and range of image classes that exist, but can result in misclassifications when more classes occur than expected, or the attributes of some classes are not consistent with their assumed character. A prudent interpreter should be mindful that EO imagery offers a unique perspective of our planet, and thus can present new information about the ground features being mapped.

Figure 9.1 Supervised classification

a. The spectral values of features that are visible in the image are located on an image crossplot



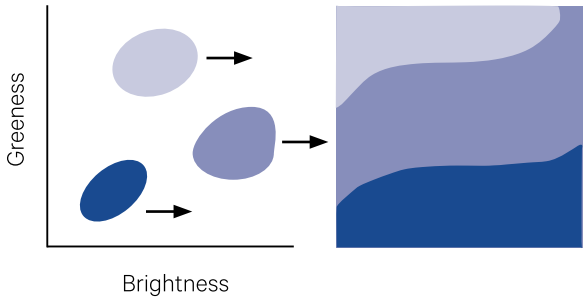
b. Summary of processes



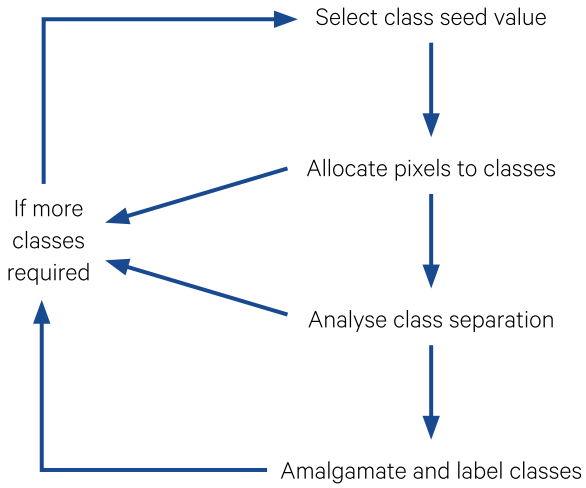
Source: Harrison and Jupp (1995) Figures 2a and 68

Figure 9.2 Unsupervised classification

a. Pixel clusters are identified on the image crossplot then related to features visible in the image



b. Summary of processes

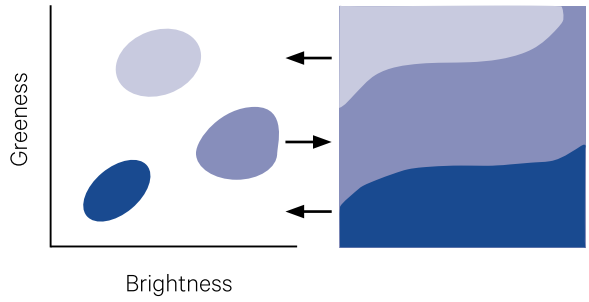


Source: Harrison and Jupp (1995) Figures 2b and 69

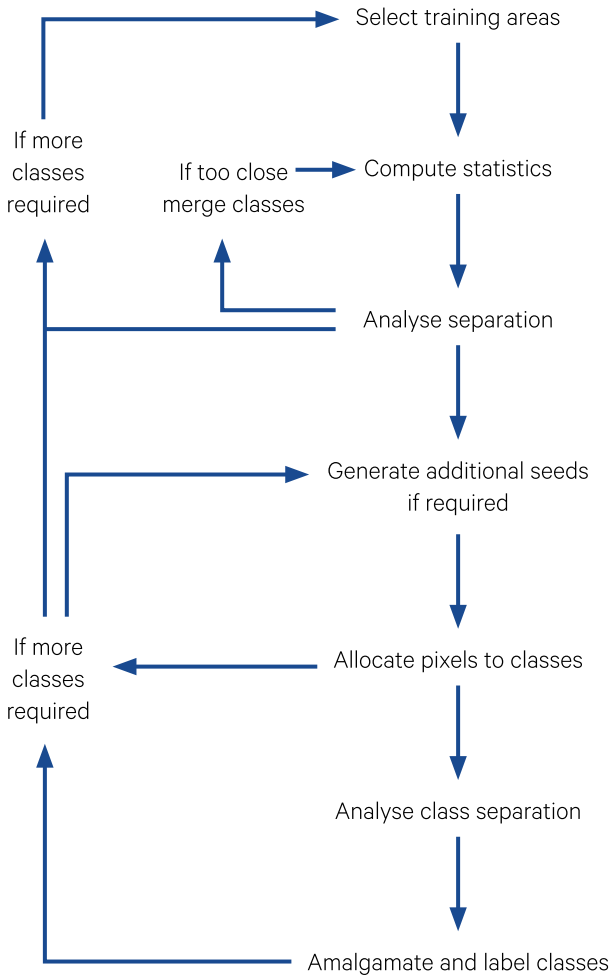
'Unsupervised' methods simply define groups of pixel values based on spectral statistics (using various 'auto-seeding' methods) then subsequently associate category labels with each group (see Figure 9.2). While some supervision is generally required in image classification to ensure that the final classes present a valid mapping of ground features, it is often beneficial to explore the subtle variability of EO imagery using various statistical tools along the way. A number of hybrid classification methods, which combine aspects of both supervised and unsupervised classification, have also been proposed (see Figure 9.3; also Excursus 9.3 and Volume 2E).

Figure 9.3 Hybrid or mixed mode classification

a. A combination of supervised and unsupervised methods are used to create classes



b. Summary of processes



Source: Harrison and Jupp (1995) Figures 2c and 70

The following sub-sections provide an overview of standard image classification methods in terms of:

- characterising classes (see Section 9.1):
  - defining classes—describes how image classes can be defined using interactive training or feature statistics, and also introduces some of the class statistics that are useful to differentiate between classes (see Section 9.1.1);
  - selecting channels—introduces popular distance measures that can be used to assess the most appropriate channels to use for image classification (see Section 9.1.2); and
  - pre-processing—reviews some of the pre-processing options that may be appropriate prior to classifying EO imagery (see Section 9.1.3);
- allocating pixels to classes (see Section 9.2)—introduces four popular approaches:
  - density slicing and painting of single channels (Section 9.2.1);
  - Parallelepiped classification for multi-channel data (Section 9.2.2);
  - Nearest Neighbour classifier (Section 9.2.3); and
  - Maximum Likelihood classifier (Section 9.2.4).
- labelling image classes (see Section 9.3)
  - clustering tools—to assess initial classifications (see Section 9.3.1) and relate image classes to interpreter categories (see Section 9.3.2);
  - aggregating classes—to represent identified clusters (see Section 9.3.2); and
  - post-classification processing—options to improve cohesion of classified image (see Section 9.3.3); and
- verifying the final classification (see Section 9.4).

One implementation of a hybrid classification methodology, the Mosaic Model approach, is introduced in Excursus 9.3. For a more considered overview of the statistics, algorithms, and philosophical issues relevant to Image Classification and Analysis, please refer to Volume 2E. An example of image classification is provided in Excursus 9.1.

Excursus 9.1—Example Classification

Source: Megan Lewis, University of Adelaide
Further information: Lewis (1998)

A vegetation mapping methodology was used to relate Landsat-5 multi-date imagery to ecological classifications. This study was conducted at Fowlers Gap Arid Zone Research Station, a 39,000 ha pastoral property in far western NSW, Australia. The climate in this region is arid. The vegetation is typical of much of the southern Australian arid zone, comprising low open chenopodiaceous shrublands, and some low open *Acacia* and *Casuarina* woodlands and grasslands on the plains (Mabbutt *et al.*, 1973, and Burrell, 1973).

Australian arid landscapes, particularly in their dry state, typically lack spectral variability, with most differentiation being related to soil brightness and vegetation cover. While the use of single-date imagery for classification depends on adequate spectral contrast between different vegetation types at that time, the use of seasonally-contrasting images takes advantage of any differences between vegetation types to improve discrimination. Because this arid vegetation varies considerably between summer and winter and wet and dry conditions, two combined Landsat images were used as a basis for classification of land cover types. The images show contrasting conditions: widespread actively growing grass and ephemeral vegetation after summer rains (see Figure 9.4a) compared with more restricted perennial vegetation growth in mid-winter of the same year (see Figure 9.4b). In total, 12 optical channels—six from each image (3 visible, 1 NIR, 2 SWIR)—were used to classify vegetation.

One hundred 6.25 ha sites were sampled, representative of the diversity of vegetation types and land systems present at Fowlers Gap. The samples were sited within relatively homogeneous areas. At each site a wheelpoint apparatus (Mentis, 1981) was used to estimate percentage cover of all physical (bare ground, stone) and biotic (ephemerals, grass and perennial shrub and tree species) ground cover components. Fifty sites were used as 'training areas' to calibrate relationships between spectral and field data, and the remaining fifty were used as validation sites to test the spectral classification.

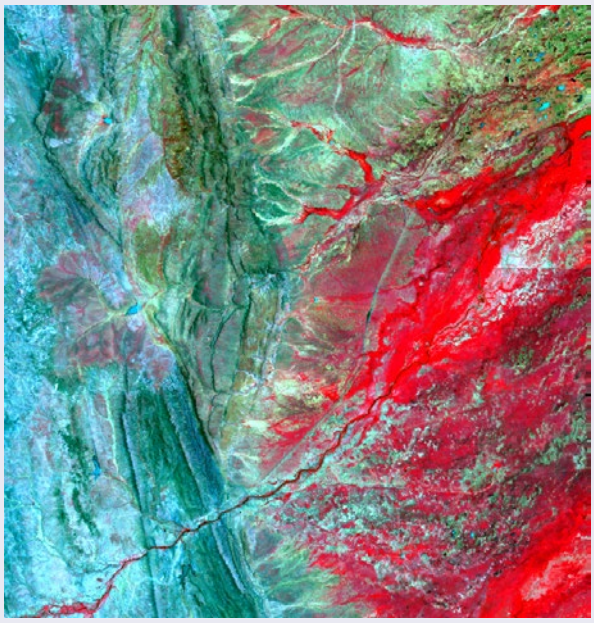
The field site data and multi-date imagery were used to:

- create nine groups of land cover based on agglomerative hierarchical clustering of field data;
- relate cover and spectral classes via a discriminant function; and
- map vegetation using a maximum likelihood classification of multi-date Landsat TM imagery (see Figure 9.5).

Figure 9.4 Study site

False colour composites of Landsat-5 TM images used to classify vegetation of Fowlers Gap Research Station, NSW.

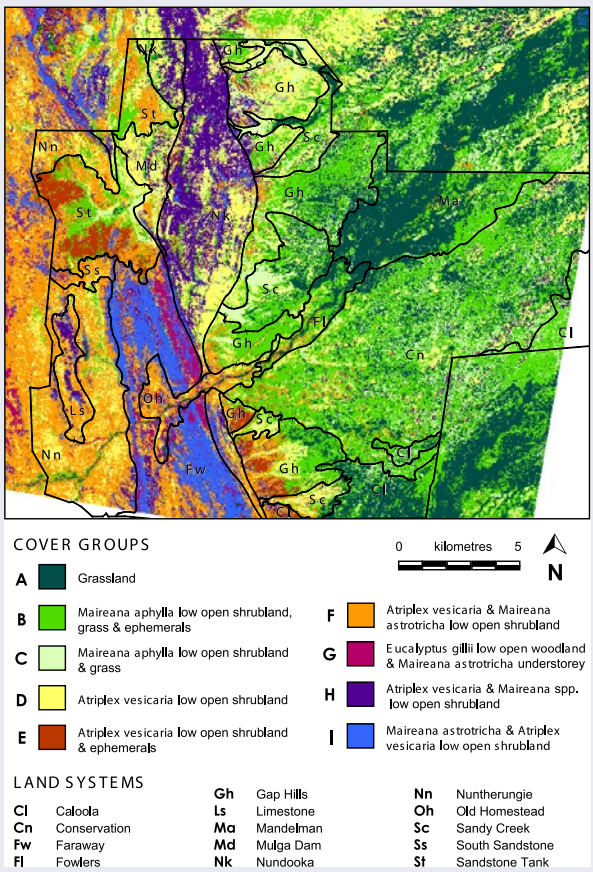
a. Summer image (30 January 1993)



b. Winter (23 June 1993)

Figure 9.5 Vegetation Classification

Maximum likelihood classification of two dates of Landsat-5 TM imagery, overlain with land systems previously mapped from black and white photography by Mabbutt (1973).



A comparison of mapped image class with ground class for the 50 verification sites (see Table 9.2) yielded a Kappa of 0.31 (s.e. 0.08, p<0.001). Although these figures suggest overall agreement between predicted and actual vegetation class, the accuracy of mapping for individual vegetation types varied considerably from 100% for the *Maireana astrotricha–Atriplex vesicaria* shrublands and *Eucalyptus gilii* open woodland, to 15% for *Acacia vesicaria* shrublands. These figures need to be treated with caution however, and have not been used to assess the accuracy of individual classes since the number of samples in each class is low.

Table 9.2 Comparison of image and cover classes

This matrix compares image class versus cover class for 50 independent validation sites.

Cover Class	Image Class									Total
	A	B	C	D	E	F	G	H	I	
A	3	3								6
B		5	1	2		2				10
C		4		1						5
D		1		1		4		1		7
E	1		1			1				3
F				2		3				5
G							2			2
H					1	3		4	2	10
I									2	2
Total	4	13	2	6	1	13	2	5	4	50

9.1 Characterising Classes

Remote sensing image interpretation based on spectral classification relies on the observation that different land covers reflect or emit electromagnetic energy in different and often unique ways, while similar land covers have similar spectral signatures. As detailed in Volume 1, the spectral discrimination between different land covers in an EO image will be determined by:

- the spectral and spatial resolutions of the image;
- the radiometric characteristics of the sensing device; and
- the date and time of image acquisition.

To spectrally identify a particular feature in an image requires that its spectral values be different from those of other features in that image. While spatial attributes of image pixels can also be considered during image classification (see Excursus 10.1), for simplicity the following sections primarily focus on spectral attributes.

When contemplating classification of an EO image, an interpreter is encouraged to consider factors relating both to image provenance and project management:

- Does the EO image have appropriate spectral and spatial resolutions and extents to identify the desired output categories? If not, it would be wise to either source imagery with more appropriate pixel size/channels/dates or radiometric detail, or adapt interpreter expectations to what the selected image can deliver (see Volume 2E—Section 1);
- What is the most appropriate method for defining classes? This question is best answered by exploring the image patterns and statistics, but options may be constrained by project budget (see Section 9.1.1 and Volume 2E—Section 4);
- Which image channels will best differentiate between ground features of interest to the project? The most relevant channels should maximise contrast and separation between classes (see Section 9.1.2 and Volume 2E—Section 5); and

- Should the image be pre-processed before classification to remove areas that are not of interest to the project or highlight specific features? Fewer channels and pixels would reduce the 'cost' of processing and, given the potential for misclassification, project focus should be restricted to regions 'known' to the interpreter (or covered by the available reference resources). Options for pre-processing are discussed in Section 9.1.3 and Volume 2E—Section 3.4.

Ultimately, an ill-considered approach to image classification will deliver an inadequate result. Selection of suitable imagery combined with careful management of the classification process, however, can yield valuable information. It is the responsibility of the interpreter to ensure that, subject to project constraints, the final classification reflects the information content of the original EO data set. When constraints are imposed, they should be appropriately acknowledged as the source of shortcomings in the final classification, rather than leaving users to conclude that these are the natural and inevitable consequence of using EO imagery.

9.1.1 Defining image classes

As detailed in Volume 2E, any class can be characterised by either:

- its current membership; or
- a rule that determines its membership.

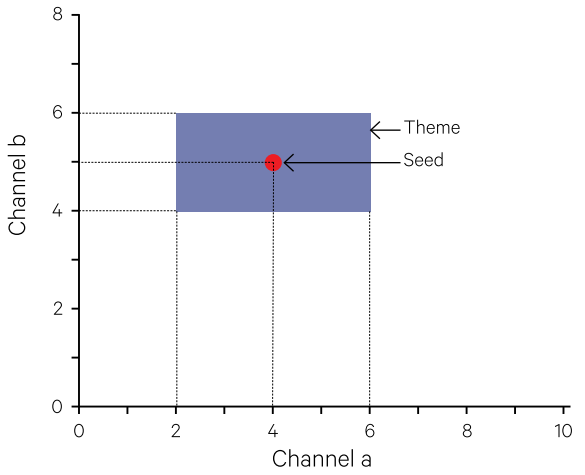
Both of these approaches can be used to define image classes. For example, once a classification has been completed, an image channel can be created in which each pixel has the value of its class 'number'. This 'classification channel' then effectively summarises the membership of each class. Alternatively, when membership is defined by an underlying rule, the relevant spectral attributes of the class need to be enunciated. This requires either:

- a valid range of values to be defined in one or more channels, such that pixels within the range(s) can belong to that class. In this text, such ranges are called 'themes'; or
- a central value is defined in one or more channels, which can be linked to a statistical criterion to determine a spectral boundary for the class. In this text, these central values are called 'seeds'.

Figure 9.6 shows the difference between seed and theme values for a two-channel image using a spectral crossplot (see Section 8.1.3).

Figure 9.6 Theme and seed values

A theme is a range of values in each channel. A seed is the central value in each image channel. In this example, the theme values are 2–6 in Channel a, and 4–6 in Channel b, while the seed values are 4 in Channel a, and 5 in Channel b.



Source: Harrison and Jupp (1990) Figure 49

The range of values for a theme can be defined by two values, the upper and lower bounds, in each image channel, with these bounds being referred to as the minimum and maximum values of the theme in a given channel. A theme can also be used to generate seed values using feature statistics as described in Section 9.1.1.2. Seeds are principally used for image classification and are also required for statistical analyses during the classification procedure (see Section 9.3 below and Volume 2E).

Sample values that represent image classes may be generated either by using:

- training patches (see Section 9.1.1.1); or
- feature statistics (see Sections 8.1 and 9.1.1.2).

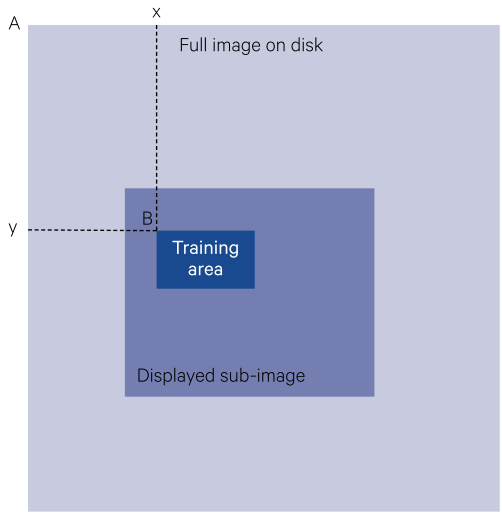
9.1.1.1 Training patches

The values of features that are visible in a displayed image may be determined interactively. This process is often referred to as training, with the resulting sample values sometimes being called a training patch (or more commonly a 'training set'). This terminology relates to the simplest form of image classification where each visually-defined sample is assumed to represent a particular ground feature. Once each pixel in the image is associated with one of these sample values, that is, the computer is 'trained' to identify the ground features based on their spectral values, the image is considered to be classified in terms of these features.

In most image processing systems the screen cursor may be used to identify an individual pixel or an enclosed area. An area is generally represented either as a rectangle, defined by two diagonally opposite corners in the image, or a polygon. As illustrated in Figure 9.7, cursor location is reported in terms of image pixels relative to the whole image, not just the displayed portion (see Sections 6.1 and 7). In many image processing systems, training areas may also be defined by specifying the locations of points along their boundaries. These locations can often be expressed either in terms of image co-ordinates or as registered map co-ordinates (see Section 7 and Volume 2B).

Figure 9.7 Locating a training area

A training area is selected using the screen cursor on the displayed image window with the location of the area being reported relative to the whole image.



Source: Harrison and Jupp (1990) Figure 50

Some image processing systems also allow training patches to be defined as:

- a line—to include those pixels intercepted by a straight line between two defined points;
- a string—ragged lines, which may be useful to sample long, thin features such as roads and rivers; and/or
- a polygon—closed area with irregular outline (see Figure 9.8 and Volume 2E—Section 4.1.1).

Figure 9.8 Five patch types

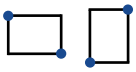
a. Point

single pixel recorded by column and row (referred to as pixel and line) position in the image

b. Line

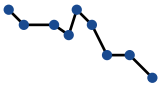
straight line interpolated between any two points

c. Box



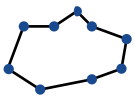
rectangle interpolated between two diagonally opposite corner points

d. String



series of straight lines interpolated between pairs of adjacent, irregularly spaced points to form a ragged line

e. Polygon



closed string which outlines a single region

Source: Harrison and Jupp (1995) Figure 45

In most image processing systems, when the training area comprises a single pixel, the digital values of all specified image channels are determined. For an area or a line, these values are extracted for all pixels in the defined region and results are often summarised in terms of the mean, standard deviation, minimum value, maximum value and/or number of pixels for each channel. If some pixels in a training area contain the null value in selected channels, the reported statistic may vary between channels, since null pixels are not used to generate image statistics.

The reported values can generally be saved as seeds or themes for subsequent use with classification or segmentation processes. Themes are commonly used in a number of image processing operations:

- to generate various statistics, including seeds (see Section 9.1.1.2);
- to implement a simple parallelepiped classification (see Section 9.2.2); or
- to define regions for image segmentation (see Section 10).

Interactive query of image values is a basic process in image analysis. As well as being used in image enhancement, classification and segmentation, sample values of selected features are required for a number of image transformations (see Volume 2C). In many image processing systems, several patches can be grouped to represent a single feature, then cumulative statistics can be generated for all patches associated with each feature type. Interactive training may also be used for identifying appropriate display parameters for particular features (see Section 9.1.1.2). Other options for using training patches with image classification are detailed in Volume 2E—Section 4.

9.1.1.2 Feature statistics

The feature identification process described in Section 9.1.1.1 can provide a basic statistical summary for pixels located within a defined training area. Other statistics that may be generated for training areas include channel histograms, and covariance and correlation matrices.

The histogram percentage points of an image feature may be used to determine display parameters to optimise the contrast of that feature in the image (typically the 1% and 99% values of the feature are selected as the minimum and maximum values for image display; see Section 4 above). Similarly, the covariance matrix of an image feature may be used to tailor a principal components transformation to that feature (see Volume 2C).

In most image processing systems, channel crossplots can be generated for:

- a complete image;
- a sub-image area (specified by screen cursor location or user-defined coordinates); and/or
- a limited spectral range as defined by a theme.

As detailed in Section 8.1.3, crossplots graphically summarise the correlation between pairs of image channels. When based on a training area, the crossplot also clearly indicates the degree of dispersion of image values in a pair of channels for the selected image region. Use of a spectral theme when generating a crossplot will obviously clip or truncate the plotted spectral space to the upper and lower bounds of the theme in the crossplotted channels, but indicates the 'density' of paired values within a particular spectral region.

Crossplots may also be used in conjunction with channel histograms to 'automatically' generate image theme values. This non-interactive mode may be used to create sample values of 'unidentified' image features as an initial step in image classification (see Excursus 9.3 below or Volume 2E). For example, the two most significant channels in an image can be crossplotted (as in Figure 9.9b), then histogram intervals corresponding to 10% or 20% of pixels (that is divisions of the 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% or 20%, 40%, 60%, and 80% points from the channel histograms) are overlaid onto the crossplot to form a grid (see Figure 9.9d). Any cells in the grid containing crossplotted values can be considered as active. The image values corresponding to the percentage points can then be used to form a theme for each active cell as illustrated in Figure 9.9e. Class seeds derived from these themes are shown relative to the image crossplot in Figure 9.10c below in term of original image channels and the first two principal components (see Volume 2C).

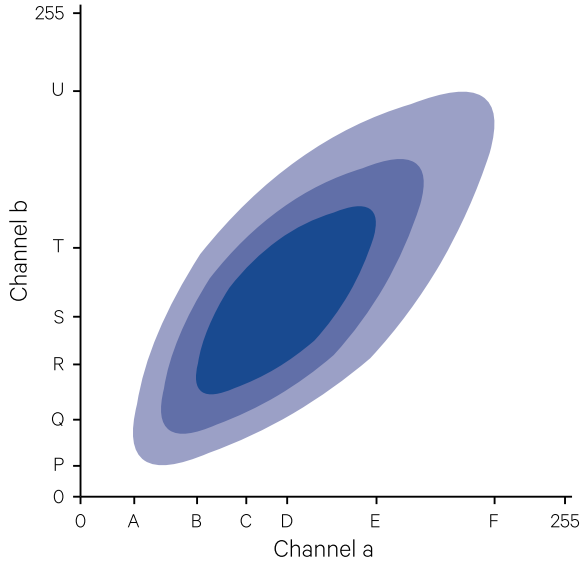
Figure 9.9 Generating themes using image statistics

Themes for auto-seeding can be generated using an image crossplot and histogram percentage points for the two most significant image channels.

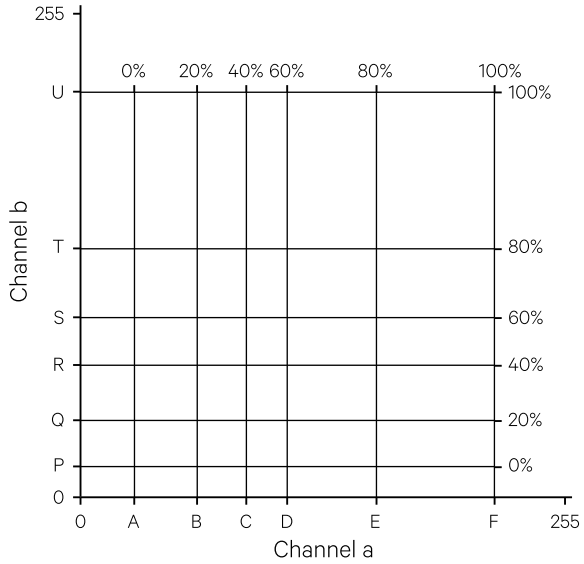
a. Histogram percentage points can be used to stratify each image channel into equal-size categories. For example, in channel a, 20% of the pixels have values in the range A–B.

Histogram % point	0	20	40	60	80	100
Channel a value	A	B	C	D	E	F
Channel b value	P	Q	R	S	T	U

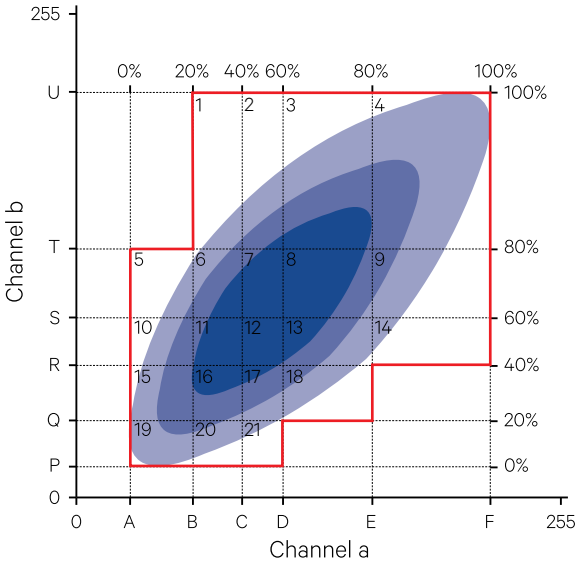
b. Crossplot of image values in channels a and b



c. Histogram percentage points from two channels are paired to form themes.



d. The percentage points can be superimposed on a crossplot to indicate which cells formed by theme ranges are 'active' in the image, that is, they do contain paired values which occur in the image.



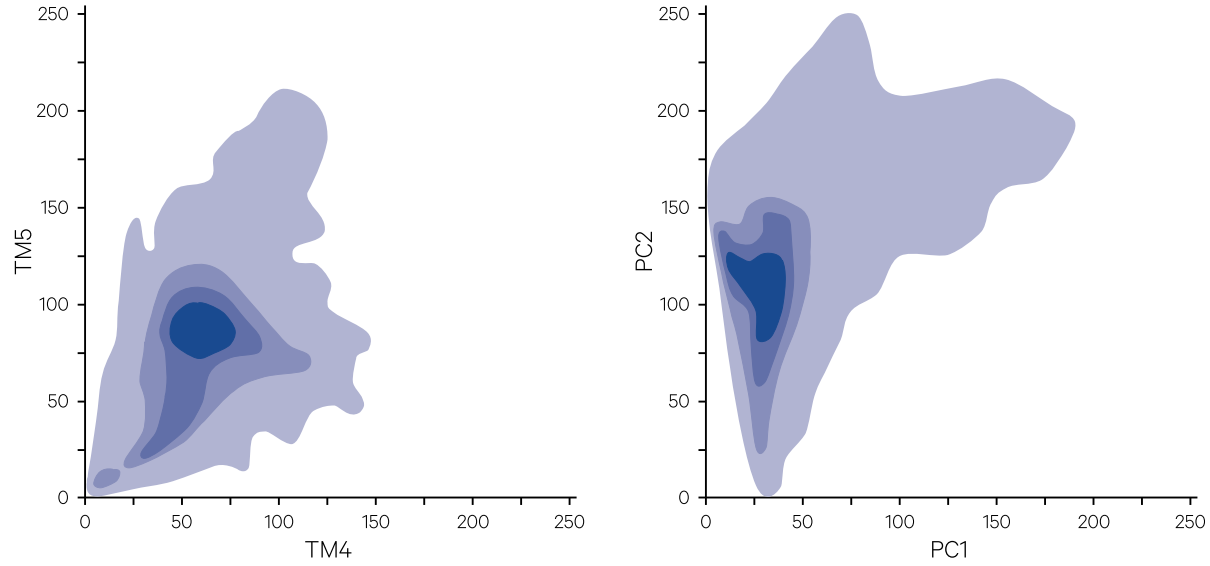
e. The ranges for each cell can then be used to define themes in those channels (see also Figure 9.10).

Theme number	Channel a	Channel b		
	Min	Max	Min	Max
1	B	C-1	T	U
2	C	D-1	T	U
3	D	E-1	T	U
4	E	F	T	U
5	A	B-1	S	T-1
6	B	C-1	S	T-1
7	C	D-1	S	T-1
8	D	E-1	S	T-1
9	E	F	S	T-1
10	A	B-1	R	S-1
11	B	C-1	R	S-1
12	C	D-1	R	S-1
13	D	E-1	R	S-1
14	E	F	R	S-1
15	A	B-1	Q	R-1
16	B	C-1	Q	R-1
17	C	D-1	Q	R-1
18	D	E	Q	R-1
19	A	B-1	P	Q-1
20	B	C-1	P	Q-1
21	C	D	P	Q-1

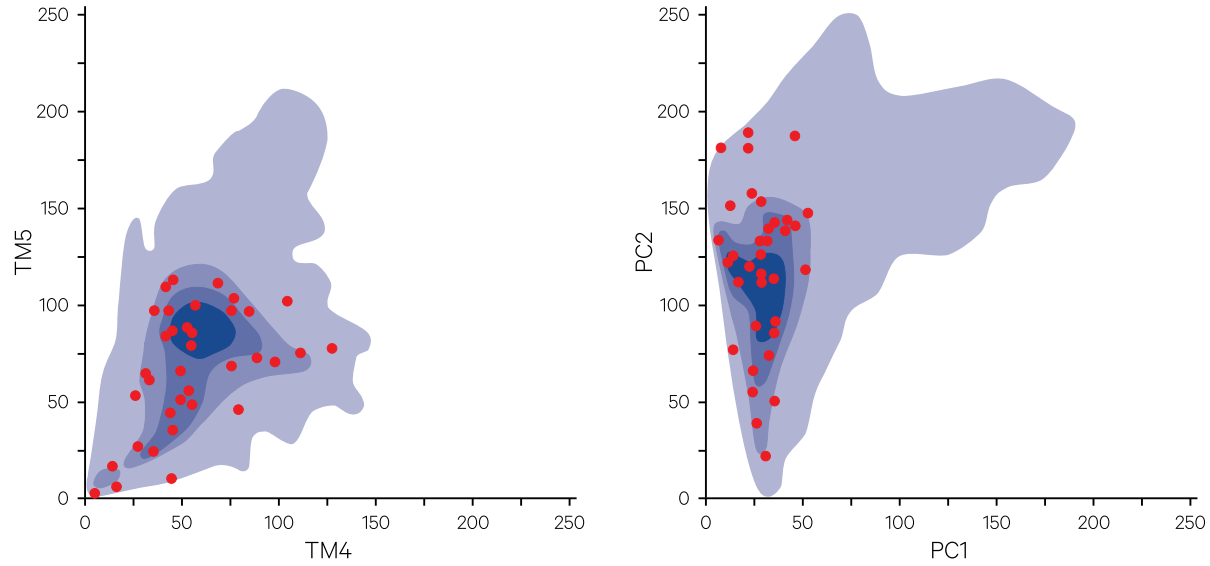
Source: Harrison and Jupp (1990) Figure 51

Figure 9.10 Comparison of class seeds

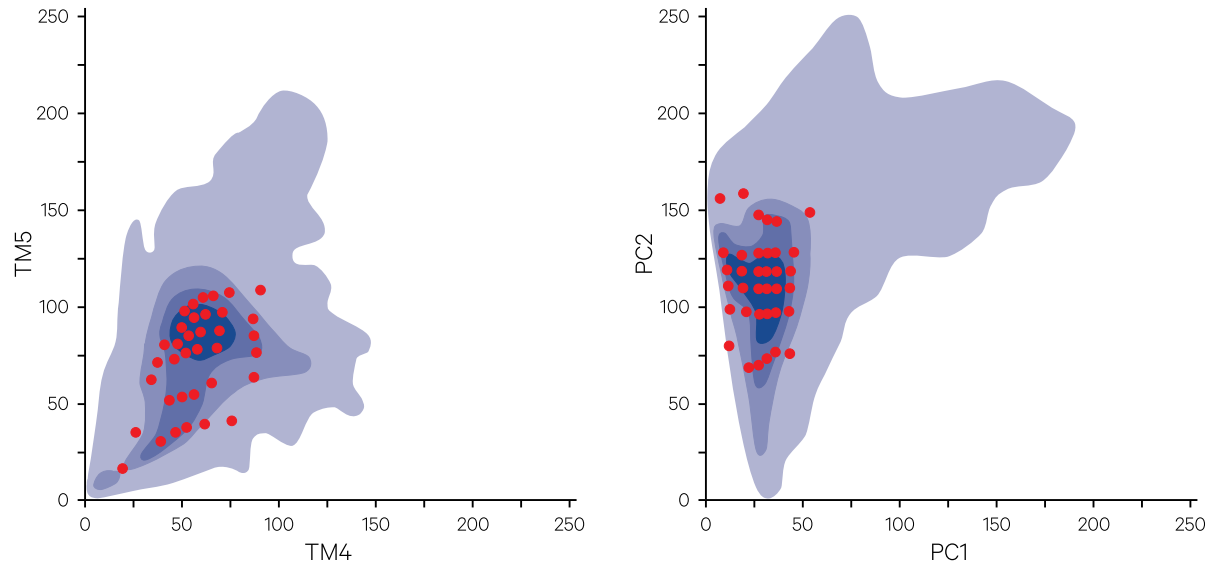
a. Example crossplots for a Landsat TM image in terms of original channels (left) and transformed to principal components (right)



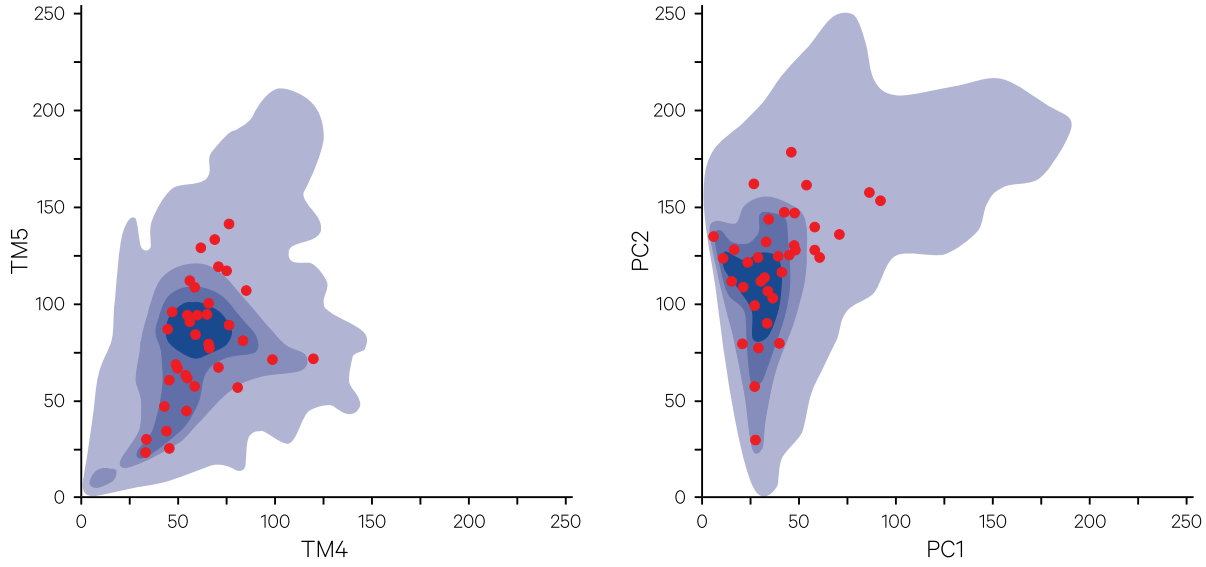
b. Class seeds derived from training patches



c. Class seeds derived from autoseeding themes (see Figure 9.9)



d. Class seeds derived from unsupervised classification



Source: Harrison and Jupp (1995) Figure 55

Various statistics can be computed using image themes. All pixels that satisfy a theme in one or more images would be used to compute the theme statistics of mean, standard deviation, minimum and maximum values, as well as the associated number of pixels. The mean values of a set of themes can often be used as seeds in image classification (see Section 9.2). Class seeds generated using this method effectively represent unidentified image features. The philosophy of the Mosaic Model approach utilises such seeds as sample class values (see Excursus 9.3). After image pixels are grouped or allocated to the classes, the features they represent can be identified interactively and an appropriate label assigned to each class. Pixel counts from themes, that is, the number of pixels in each theme, can also be converted to estimates of ground area using the results of image rectification (see Section 7 and Volume 2B).

A range of other approaches have been proposed to automatically generate themes or seeds for image classes (see Volume 2E—Section 4.2), including:

- crossplotting first principal component (PC1) and second principal component (PC2) (or the first two canonical variates: CV1 and CV2—see Section 9.1.2.1) to locate local maxima points then using 'spectral transfer' to associate new channels values with each class (see Section 9.2.3);
- AMOEBA clustering classifier (Bryant, 1979, 1989, 1990), which uses spatial and spectral characteristics to derive class statistics from homogeneous patches in the image;
- non-parametric distance measures—can use predefined ranges to cluster pixels values around arbitratry seeds, such that each pixel which cannot belong to an existing class becomes a new class seed (eg. Skidmore, 1989);
- semi-automatic approaches based on region-growing away from defined pixels (eg. Bolstad and Lillesand, 1992);
- use clustering algorithms to group pixels into classes defined by gradients shown on channel crossplots (eg. Baker *et al.*, 1991);
- locate values of concentrated pixels as shown on three-dimensional contour plots, that is, a two channel crossplot on which pixel density is highlighted (Goldberg and Shlien, 1976; Schachter *et al.*, 1976; Narendra and Goldberg, 1977; Goldberg and Shlien, 1978). For example, Watershed classification (Watson, 1987; Watson *et al.*, 1992) views the contour plot as a landscape, and uses connectivity principles to define spectral regions as separate 'catchments' (Hashim, 1996); and
- in some implementations of the minimum distance classifier (see Section 9.2.3), unclassified pixels can become new seeds. Subsequently, the mean migration process (see Figure 9.19 below) can be used to redefine class seed values as the mean values of their member pixels.

Three different sets of class seeds that have been generated for a Landsat TM image are shown in Figure 9.10 on crossplots of TM4 vs TM5 and PC1 vs PC2. Seeds derived from training patches tend to be concentrated where pixel values are concentrated whereas the unsupervised class seeds sample a broader cross-section of the image crossplot values.

9.1.2 Comparing classes and classifications

Once a class is defined by its membership, a range of statistics can be computed to further characterise the class, quantify its separation from other classes, and determine its inherent stability and cohesion. Relevant measures of class separation are detailed in Volume 2E—Section 5.2. These metrics can be computed for classifications including and excluding selected channels to determine those channels which are most effective in separating a set of pre-defined classes. Similarly, the 'value' of individual classes can be tested by computing these statistics for a classification that includes them compared with a classification that does not.

9.1.2.1 Classification statistics

A range of statistics related to image classification are detailed in Volume 2E—Sections 2 and 5. Here we provide an overview of two basic methods that can be used to analyse and compare classifications—Canonical Variates Analysis (CVA; see Volume 2E—Sections 2 and 5) and Numerical Taxonomy (see Volume 2E—Section 6). Interested readers are encouraged to refer to Volume 2E for a full explanation of the underlying assumptions and limitations of these tools.

CVA allows the image data space to be rotated and rescaled to maximise the differences between sets of classes relative to the variation within them by defining linear combinations of selected image attributes (see Figure 9.11). The total covariance matrix (***T***) is the sum of the between-class covariance matrix (***B***) and the within-class covariance matrix (***W***), that is:

$$T = B + W$$

where

B is computed from the class mean (or seed) values; and

W is the pooled (or average) within-class variability.

In image terms:

image being classified = mean image + residual image

(minus the offset used to create the residual image; this is 127 for 8-bit (byte) images) and:

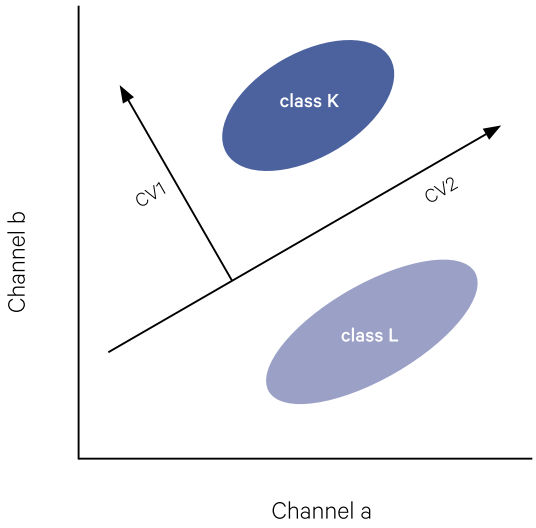
T is the covariance matrix of the image being classified;

B is the covariance matrix the 'mean' image, in which pixel values are replaced by the seed (or mean) values of the class that pixel has been allocated into (see Section 9.2.3); and

W is the covariance matrix the 'residual' image, which is the difference between the original image being classified and the 'mean' image (see Section 9.2.3).

Figure 9.11 Canonical Variates Analysis (CVA)

To maximise the separation between class K and class L, we can rotate and rescale channel axes so that the first canonical variate (CV1) gives greatest separation between class distributions. CV2 is defined as orthogonal (perpendicular) to CV1 (see also Volume 2E—Sections 2 and 5)



Source: Harrison and Jupp (1995) Figure 23

In a 'good' classification, the ratio of the within-class variation to the between-class variation (see Volume 2E—Section 5.1.3) will be small, and the residual image will have the structure of a random sample from a normal multivariate population with a zero mean vector and variability summarised by the within-class covariance matrix. The residual image is never spatially random but it should not contain significantly coherent patterns (see Section 9.2.3).

Various statistical tests based on CVA can be used to compare alternative classifications. These tests and their underlying statistics are presented in greater detail in Volume 2E—Sections 2 and 5. When derived for different classifications of an image (whether differences in classifications stem from using different classes, channels and/or allocation algorithms), such statistical tests can indicate which classification provide the best overall separation between classes. Some commonly used tests for comparing classifications include:

- minimum Trace (sum of diagonal elements) of ***W***:

  Tr ***W***

  This indicates that the within-class sum of squares is small (and is the same as selecting the classification with max Tr ***B***). This statistic implies that the individual classes are spherical and homogeneous, so outliers or small clusters near large ones could render misleading results. Also, it is not invariant to scaling so gives different results for standardised and non-standardised data.

- Wilk's λ or Likelihood ratio statistic:

  $$\text{minimum } \frac{|\boldsymbol{W}|}{|\boldsymbol{T}|}$$

  where

  $|\boldsymbol{W}|$ is the determinant of $\boldsymbol{W}$

  which is the same as testing min $|\boldsymbol{W}|$ or max $|\boldsymbol{T}|/|\boldsymbol{W}|$. This statistic assumes that all classes have the same basic shape but do not have to be spherical. This test involves more computations but is generally preferred over min Tr $|\boldsymbol{W}|$.

- largest root criterion:

  $$\lambda_{max}\left(\boldsymbol{W}^{-1}\boldsymbol{B}\right)$$

  The largest overall ratio of between-to-within variation is used to indicate that one classification is better than another.

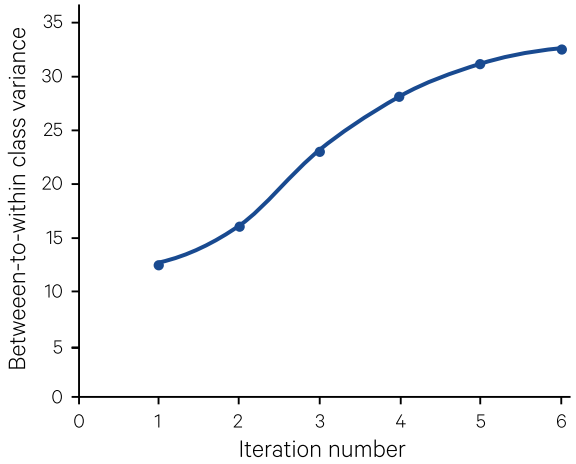
- Hotelling's trace criterion:

  $$\max\left(\text{Tr}\left(\boldsymbol{W}^{-1}\boldsymbol{B}\right) = \sum_{i=1}^{n}\lambda_i\right)$$

  The total between-to-within variance ratio for the individual variables is used to assess class separation. This criterion is more sensitive to image dimensionality.

- As new spectral classes are added to a classification, one indication of the value of additional classes is the trend in between-class to within-class variation, which tends to stabilise for the 'right' number of classes. This can be plotted for each iteration as shown in Figure 9.12.

Figure 9.12 Assessing iterative classifications



Source: Harrison and Jupp (1995) Figure 27 [Adapted from McVicar *et al.* (1989)]

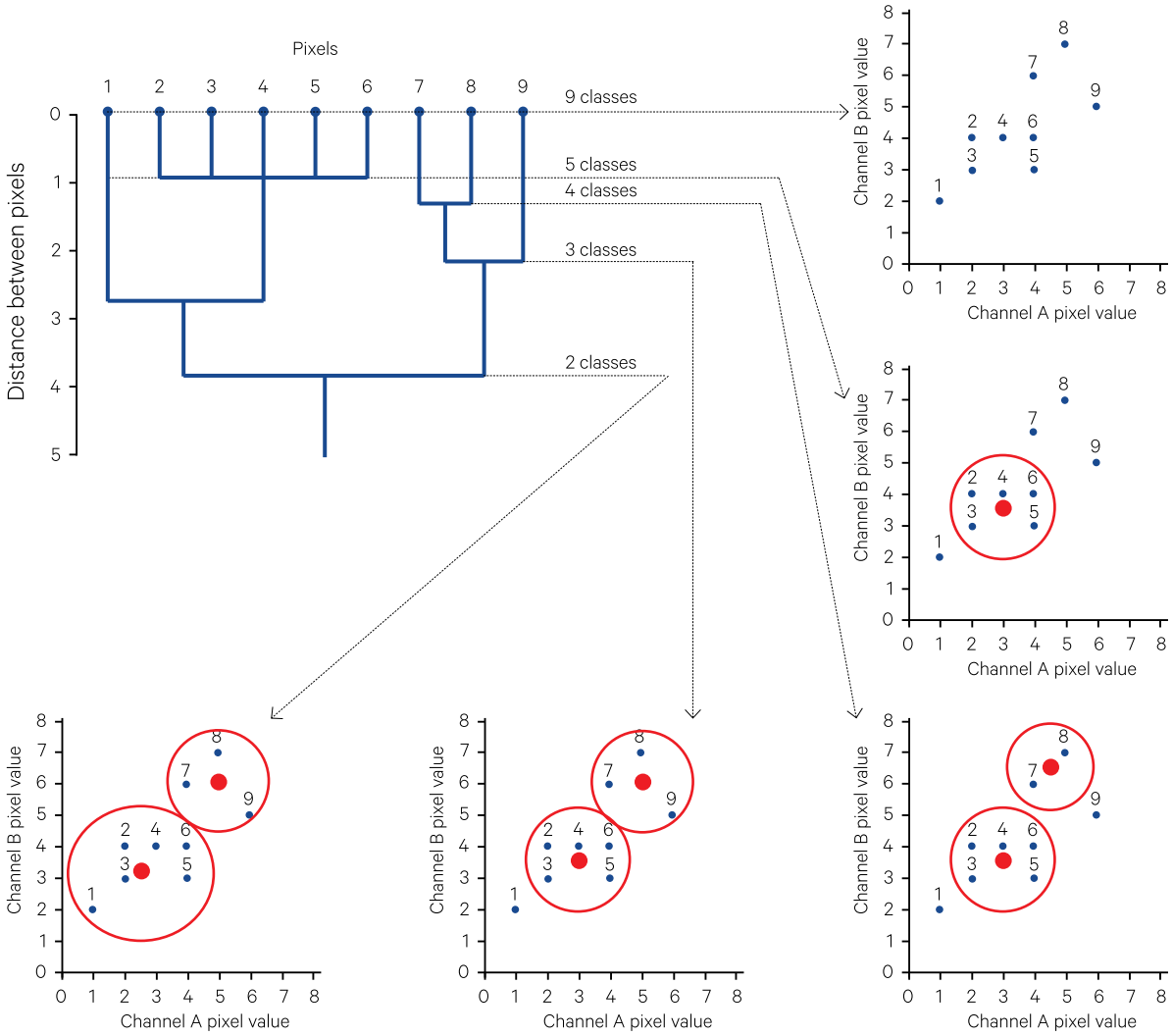
Another approach to analysing class separation uses tools developed for numerical taxonomy. In this context, numerical taxonomy embraces methods that systematically group classes using their numerical characteristics. For image classification, most taxonomic methods use some measure of the distance between classes to determine how to combine them into groups. Methods either start with one group (all classes) and systematically divide this group into multiple sub-groups in a hierarchical fashion, or alternatively, agglomerative methods start with multiple groups (such as a set of image classes) and, on the basis of some 'similarity' measure, cluster the individual classes into related groups. Such methods are detailed in Volume 2E—Section 6.1.

The hierarchical interconnections between classes can be viewed using either:

- dendrograms—a tree diagram, which successively links classes to form larger groups, with linkages being ordered by the distance between groups (see Figure 9.13). Different algorithms use different methods for determining intra-class distance and different strategies for forming groups so can and do create different dendrograms (see Volume 2E—Section 6.1.1). However, the first and last linkages are often similar for dendrograms generated by different algorithms; or
- minimum spanning trees (MST)—a connected, non-cycling graph that links all members of a set. In the context of image classification, MST can be based on both the spectral and spatial similarity of classes. A variety of methods can be used to generate MST, using different axes and different similarity measures (see Volume 2E—Section 6.3).

Figure 9.13 Dendrogram example

This tree diagram progressively groups nine pixels (shown as blue dots) into groups based on some measure of the distance between them. The mean values of the merged groups are shown as red dots.



Source: Adapted from Richards (1986)

For example, the Mahalanobis distances in a selected set of channels between pairs of classes can be computed and ordered in terms of increasing value (see Table 9.3). These ordered distances can be used to construct both a spectral minimum spanning tree (MST) and a dendrogram of taxonomic linkages. To create the MST, the pairs of classes are plotted and linked on a crossplot of selected image channels (or transformed channels such as CV1 versus CV2). The closest pair of classes (eg. classes 9 and 17 in Table 9.3) are joined first, then the second pair (classes 3 and 6 in Table 9.3) and so on until all linked pairs are connected (see Figure 9.14). It is also informative to use different coloured links for different ranges of distances between class pairs when creating the MST.

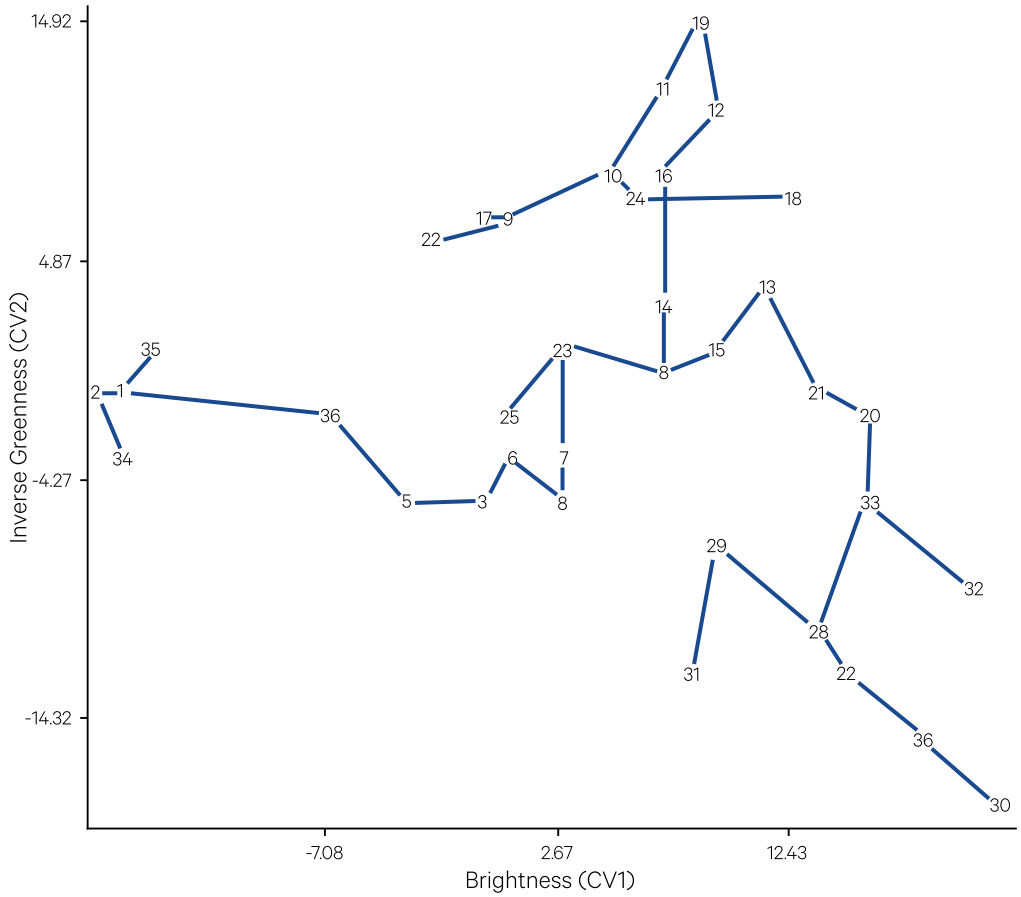
The MST and dendrogram of clusters are useful tools to assess the distance between classes or to evaluate training sets, that is to 'learn' how classes are separated in different attributes, during the initial exploratory phase of image classification. If two class seeds are separated by less than two units in Mahalanobis distance, they could be considered to represent a single statistical class. The MST is valuable during the classification labelling stage to identify classes which form gradients rather than discrete, spectrally distinct clusters (see Section 9.3).

Table 9.3 Minimum spanning tree linkages

Linkage	Class a	Class b	Mahalanobis distance
1	9	17	2.64
2	3	6	2.852
3	21	20	4.078
4	8	35	5.689
5	28	27	5.952
6	4	7	6.225
7	15	13	6765
8	20	33	7.130
9	6	4	7.188
10	5	3	7.206
11	9	22	8.637
12	27	26	10.668
13	16	12	10.680
14	8	14	11.118
15	23	25	11.887
16	11	10	13.276
17	10	9	15.841
18	26	30	16.332
19	7	23	16.476
20	36	5	16.876
21	28	29	18.238
22	33	32	20.799
23	13	21	22.548
24	14	16	22.585
25	33	28	23.997
26	23	8	24.301
27	10	24	29.686
28	24	18	29.725
29	29	31	31.463
30	12	19	38.313
31	19	11	42.595
32	1	36	46.874
33	1	35	77.683
34	1	2	88.883
35	2	34	202.427

Figure 9.14 Spectral MST

In this example, class numbers are graphed by their brightness (CV1) and inverse greenness (CV2) values. Classes are then linked in order of increasing Mahanobis distance between them (see Table 9.3).



Source: Harrison and Jupp (1995) Figure 59

Spatial statistics can also be used to compare classifications. For example, the class co-occurrence matrix, which quantifies the frequency with which pairs of classes occur in adjacent pixels in the image, can be used to construct a spatial MST. This effectively links classes on the basis of their spatial proximity (see Volume 2E—Section 6.4).

9.1.2.2 Measuring class separation

Before mentioning some statistics that are commonly used to measure class separation, we should pause and reflect on what they mean. Given that a class represents a group of pixels, how do we measure the distance between two groups? There are many options for this, including:

- distances between class seeds;
- some 'sensible' average of the set of distances between the values of each pixel in one class and the values of each pixel in another class;
- weighting factors that account for different data ranges in different channels; and/or
- measures that consider the distribution of pixels in each class.

Some of the more commonly encountered distance metrics are based on the Minkowski series of distance functions summarised in Table 9.4. These metrics simply measure the separation between class seeds (which are effectively assumed to be class centres) in the n-dimensional space defined by *n* image channels. Obviously, it is important when using these metrics that the class seeds do represent class centres.
A standardised version of Euclidean distance, the Karl-Pearson distance (K^2), uses weights derived from channel standard deviation (s_k) to account for variability between channels:

$$\sum_{k=1}^{n_c} \frac{\left(x_{ik} - x_{jk}\right)^2}{s_k^2}$$

Table 9.4 Minkowski metrics

Metric	Equation	Description
Generic series of distance functions ($\|I.I\|_p$)	$d_p\left(X_i,X_j\right)=\left(\sum_{k=1}^{n_c}\left\|x_{ik}-x_{jk}\right\|^p\right)^{1/p}$	p is greater than or equal to 1 (p =1 for city block; p =2 for Euclidean distance; p = ∞ for Chebychev)
City block (taxicab, Manhattan or L_1)	$d_1\left(X_i,X_j\right)=\sum_{k=1}^{n_c}\left\|x_{ik}-x_{jk}\right\|$	X_i is a member of class i, X_j is a member of class j, n_c is the number of attribute (or channel) values at each pixel, x_{ik} is the k'th component of X_i, x_{jk} is the k'th component of X_j, and | | denotes the absolute value (or unsigned difference).
Euclidean distance (L_2)	$d_2\left(X_i,X_j\right)=\left(\sum_{k=1}^{n_c}\left\|x_{ik}-x_{jk}\right\|^2\right)^{1/2}$	$\sum\left\|x_{ik}-x_{jk}\right\|^2$ is commonly written as $\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|^2$
Chebychev (L_∞)	$d_\infty\left(X_i,X_j\right)=\mathrm{Max}_{k=1,n_c}\left\|x_{ik}-x_{jk}\right\|$	
With channel weighting function	$d_p\left(X_i,X_j\right)=\left(\sum_{k=1}^{n_c}w_k\left\|x_{ik}-x_{jk}\right\|^p\right)^{1/p}$	W_k are pre-determined weighting values for the channels

Other class separation statistics consider the distribution of values around class centres. For example, the Mahalanobis distance (D^2; Mahalanobis, 1936; Rao, 1948) can be computed as:

$$d_p\left(X_i,X_j\right)=\sqrt{\left(\boldsymbol{x}_i-\boldsymbol{x}_j\right)^T\boldsymbol{W}^{-1}\left(\boldsymbol{x}_i-\boldsymbol{x}_j\right)}$$

where $\boldsymbol{W}$ is the selected class covariance matrix, $\boldsymbol{W}^{-1}$ is its inverse, and each set of channel values for individual pixels or classes in the image can be considered as a one-dimensional array, or vector, $\boldsymbol{x}_k$. This distance effectively measures the separation between class seeds and assumes all classes can be characterised by equal, normal distributions. Mahalanobis distance is particularly relevant to measuring inter-class distances when the Minimum Distance classifier has been used to allocate pixels to classes and can be shown to be equivalent to the Euclidean distance in the Canonical Variate transformed space (see Section 9.1.2.1 and Volume 2E).

Separation between two overlapping probability distributions can be measured by a statistics called Divergence (see Volume 2E—Section 5.2). This is most relevant to classifications defined by Maximum Likelihood allocation (see Section 9.2.4). For EO analyses, the Transformed Divergence is more commonly used. This is more sensitive than average divergence to classes which are not well-separated (Swain and Davis, 1978; Richards, 1986; Thomas *et al.*, 1987):

$$D'_{ij}=k\left[1-\mathrm{e}^{-D_{ij}/8}\right]$$

where k scales the D'_{ij} values between 0 for identical classes and k for maximum separation.

Another particularly robust class separation statistic for classifications based on Maximum Likelihood allocation is the Jeffries-Matusita distance (Swain and Davis, 1978):

$$J_{ij}=\sqrt{\int\sqrt{P\left(\boldsymbol{x}\,|\,class\ i\right)}-\sqrt{P\left(\boldsymbol{x}\,|\,class\ j\right)}\,dx}$$

For normally-distributed classes this becomes:

$$J_{ij}=\sqrt{k\left(1-\mathrm{e}^{-B_{ij}}\right)}$$

where k scales the J_{ij} values between 0 for identical classes and k for maximum separation (that is, no overlap). When the underlying populations are normally distributed ($N\left(\mu,\boldsymbol{W}\right)$):

$$B_{ij}=\frac{1}{2}\log\left(\frac{\left|\frac{1}{2}\left(\boldsymbol{W}_i+\boldsymbol{W}_j\right)\right|}{\sqrt{\left|\boldsymbol{W}_i\right|}\sqrt{\left|\boldsymbol{W}_j\right|}}\right)+\frac{1}{8}\left(\mu_i-\mu_j\right)^T\left(\frac{\boldsymbol{W}_i+\boldsymbol{W}_j}{2}\right)^{-1}\left(\mu_i-\mu_j\right)$$

B_{ij}, the Bhattacharyya distance, is a natural extension of the Mahalanobis distance to cases where classes have different covariance matrices. The general expression for the Bhattacharyya distance for a non-normal distribution is:

$$e^{-B_{ij}}=\int_X\sqrt{P\left(\boldsymbol{x}\,|\,Class\ i\right)P\left(\boldsymbol{x}\,|\,Class\ j\right)}\,dx$$

9.1.3 Selecting image channels

Ideally, to maximise the efficiency of the image classification process, the image being classified would not contain a significant degree of redundant information. Accordingly, it is useful to assess the intrinsic dimensionality of an image prior to classification using various 'feature selection' methods, which highlight the most informative image channels (see Section 9.1.3.1 and Volume 2E—Section 5.3.1). In some cases, pre-processing of imagery is appropriate prior to classification (see Section 9.1.3.2).

9.1.3.1 Feature selection

While modern computing power happily copes with large datasets, a compromise has traditionally been required between classification accuracy and computational efficiency. Greater classification efficiency is achieved by using as few image channels as possible, that is, identifying the subset of channels that provide maximum discrimination between classes. Two criteria that have been used for this purpose are:

- minimum error rate; and
- maximum separation between classes (McKay and Campbell, 1982b; Young and Fu, 1986).

Relevant measures of class separation are introduced in Section 9.1.2 and detailed in Volume 2E—Section 5.2. The best subset of image channels is commonly assumed to be the one that optimises some selected class separability measure. For example, Mahalanobis distances between classes can be computed with and without selected channels to assess the contribution being made by one or more channels to class separation. Similarly, from a set of potential channel combinations, for example, the combination that maximises the Jeffries-Matista distance (J) could be regarded as the 'best' (see Section 9.1.2.2 and Volume 2E—Section 5.3.1).

McKay and Campbell (1982a) discussed three broad types of method for variable selection, namely:

- methods that use stepwise F-tests;
- methods associated with CVA; and
- methods which evaluate all possible subsets (see Volume 2E—Section 2).

They concluded that a combination of CVA and 'all-subsets' methods provides the most effective basis for channel selection, although for an image with a large number of channels, the 'all-subsets' approach may not be feasible. Some of the tests that are used to compare CVA results are reviewed in Section 9.1.2.1 and Volume 2E—Section 5.1.3.3. When using such statistics, however, it is worth remembering that pooled statistics may overlook the significance of small classes that represent specific ground features.

9.1.3.2 Pre-processing

It may also be appropriate to apply one or more image transformations to imagery prior to classification (see Volume 2E—Section 3.4). Commonly used pre-processing includes:

- removal of 'noise' pixels using despiking routines (see Volume 2C);
- segmentation to remove pixels unrelated to a particular study (see Section 10);
- rescaling selected image channel(s) to emphasise selected portions of the data range (see Volume 2C);
- channel ratios to highlight specific ground feature characteristics, such as vegetation greenness (see Volume 2C and 3A); and
- Principal Component Analysis (PCA) to reduce the dimensionality of the image (see Volume 2C).

9.2 Allocating Pixels

Image classification involves reducing the variation contained within an EO image to a set of classes that can be labelled as scene features, such as land cover. The degree to which different features can be discriminated in an image largely depends on the resolution of the image data itself, but is also affected by the flexibility of the classification method being applied.

Various methods can be used to allocate pixels to classes during image classification. Four of the most commonly encountered approaches include:

- Density slicing and painting of single channels (Section 9.2.1);
- Parallelepiped classification for multi-channel data (Section 9.2.2);
- Nearest Neighbour classifier (Section 9.2.3); and
- Maximum Likelihood classifier (Section 9.2.4).

The first two allocation methods listed above rely on themes to define valid data ranges for classes, while the last two methods determine class membership on the basis of spectral separation from class seeds.

All allocation approaches may be applied in a 'supervised' or 'unsupervised' (or mixed) mode, where the terminology refers to whether the spectral statistics used to define the classes were derived from examples of the ground feature categories to be mapped (the supervised mode) or defined numerically to describe the variation in the data without reference to a pre-determined set of ground feature categories (the unsupervised mode). It is very difficult to 'train' a classifier to recognise categories whose definition is not related to the physics being observed in the EO image. It is also useful to investigate aspects of the data that might not be included in an *a priori* categorisation of the image.

A range of other model-based and knowledge-based classification approaches are described in Volume 2E—Section 8:

- contextual classifiers (see Volume 2E—Section 8.1);
- mixture classifiers (see Volume 2E—Section 8.2);
- decision trees (see Volume 2E—Section 8.3);
- expert systems (see Volume 2E—Section 8.4); and
- neural networks (see Volume 2E—Section 8.5).

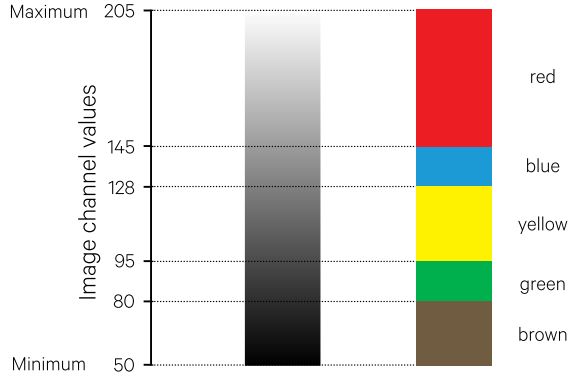
9.2.1 Density slicing

The simplest form of image classification involves grouping ranges of consecutive values in a single image channel to represent different categories. As with most classification results, these categories are typically displayed as different colours so this process also performs an image enhancement function in a a similar fashion to pseudo-colouring (see Figure 9.15 and Figure 5.12).

Most image processing systems utilise the colour LUT of the colour display system to instantaneously redefine colours of the image data being stored in display memory (see Section 6). For example, after a grey-scale image has been displayed, defined ranges of image values could be associated with a selected display colour then reset to their original grey tones.

Figure 9.15 Density slicing

A single image channel can be pseudo-coloured by assigning ranges of image values to specific colours.



Source: Harrison and Jupp (1990) Figure 52

In most image processing systems, the painted colours can be considered as image categories or classes and summarised in a single channel in which each painted pixel in the image is given a class number associated with its painted colour and unpainted pixels are set to the null value. This 'classification channel' can be displayed as a grey-scale image for interactive colouring or 'cleaned' using a modal filter (see Volumes 2C and 2E).

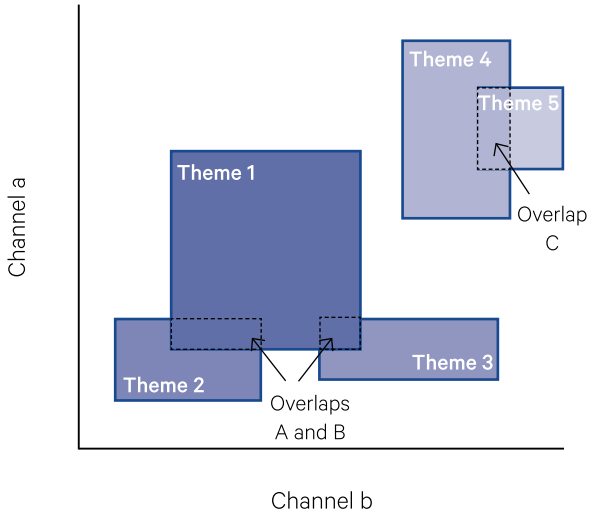
Interactive density slicing is used for labelling and aggregating classes when the classification is represented as a single channel (see Section 9.3 and Volume 2E). The technique is also useful to enhance single channel imagery or transformed data channels (see Volume 2C).

9.2.2 Parallelepiped classification

A simple method for multi-channel image classification involves the use of themes defined by minimum and maximum values (see Section 9.1.1.2). While different implementations of this technique vary, pixels are generally mapped into the first theme that they satisfy. As shown in Figure 9.16, when the categories formed using this process are shown on a crossplot, they are represented as rectangular boxes. This characteristic has resulted in this form of classification being referred to as 'parallelepiped' classification. The process can be viewed as a multi-channel form of density slicing with the requirement that a pixel fall into the specified data range in each image channel.

Figure 9.16 Parallelepiped classification

This simple classification technique uses multi-channel themes to define image classes. If theme values overlap, pixels are allocated to the first theme they satisfy. In this case overlap areas A and B would be assigned to Theme 1, while overlap area C would be assigned to Theme 4.



Source: Harrison and Jupp (1990) Figure 53

Themes may be defined using interactive interrogation (as described in Section 9.1.1.1) or feature statistics (see Section 9.1.1.2). Each image pixel is checked against these theme bounds and associated with a particular theme (generally the first one it satisfies). The operation is performed using the image file values and considers all image channels. Being a disk-based process, this classification method is slower than interactive density slicing (see Section 9.2.1).

The allocation process, that is associating a pixel with one of the themes, may be applied by colour 'overlaying' on a displayed (colour composite) image or by writing an output image to disk using the colour codes associated with each theme. This simple classification technique is also useful for defining appropriate bounds of particular features for image segmentation (see Section 10.2). Since certain transformations are more effective when applied to a pre-segmented image, some implementations of image operations, such as ratioing or destriping, may also allow themes to be specified to implement image segmentation during processing (see Volume 2C).

9.2.3 Minimum distance classifier

A Minimum Distance (or Nearest Neighbour) classifier allocates each pixel in an image to one of the available classes using a minimum distance rule, that is the class whose seed values are deemed to be the 'closest'. As introduced in Section 9.1.2 and detailed in Volume 2E—Sections 5 and 7, various metrics can be used to determine the distance between two classes. Some of the metrics commonly encountered in minimum distance classifiers are illustrated in Figure 9.17.

One of the principal advantages of the minimum distance classifier is its computational efficiency, which generally is linearly proportional to the number of classes. Many implementations of this classifier allow new seeds to be created automatically when a pixel cannot be allocated to an existing class seed.

A common algorithm for this purpose is the 'City Block' metric (see Table 9.1 and Figure 9.17b), which measures the distance between a pixel and a class seed weighted by channel gate values. In mathematical terms, this means that a pixel with attribute values $\boldsymbol{x}$ is allocated to the decision region D_i from which it has the minimum measurement $dist(\boldsymbol{x}, D_i)$:

$$dist\left(\boldsymbol{x}, D_i\right) = \sum_{j=1}^{n_c} \frac{\left|x_j - m_{ij}\right|}{gate_j}$$

where

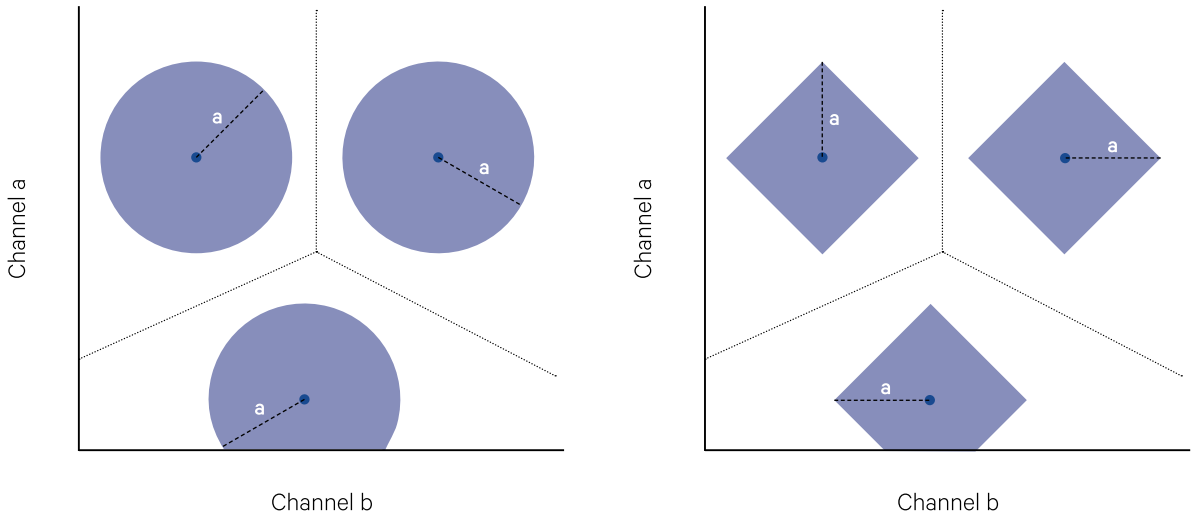
- $\boldsymbol{x}$ is the vector of attributes (or channel values) for the pixel to be allocated,
- D_i is the image attribute decision region for class I,
- x_j is the value of attribute j for the pixel to be allocated
- m_{ij} is the value of attribute j representing class i,
- $gate_j$ is a weight supplied for each attribute or channel, and
- n_c is the number of attributes (or channels) used for the allocation.

Figure 9.17 Distance thresholds for minimum distance classifiers

Different methods can be used to measure distance in image spectral space. Two of the thresholds that are commonly used to determine distance in the minimum distance classifier are:

a. Euclidean distance, where threshold distance = a from the class seed in all directions

b. City block metric, where threshold distance = a from the class seed only in the directions of the channel axes



Source: Harrison and Jupp (1995) Figure 31

Another condition that may be applied during allocation is that the pixel value be within a specified distance of the class seed in each image channel involved in this process (see Figure 9.18a). That is, the attributes must satisfy the condition:

$$L_{ij} < x_j < U_{ij}$$

where

$$L_{ij} = m_{ij} - gate_j \text{ and}$$

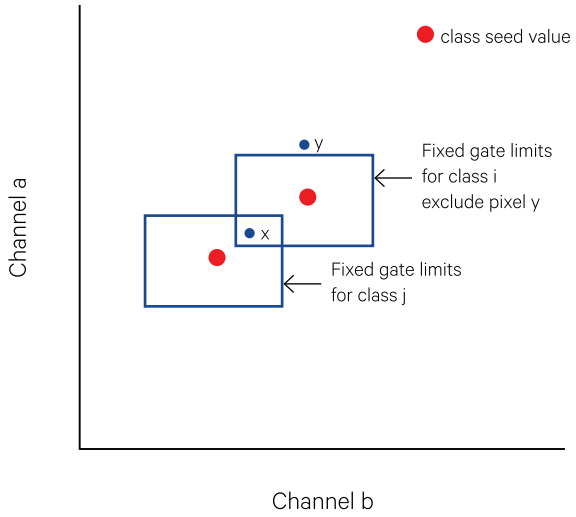
$$U_{ij} = m_{ij} + gate_j$$

For example, the class 'gate' value for a given channel *j* ($gate_j$) may be proportional to the within-class standard deviation in that channel (see Figure 9.18b).

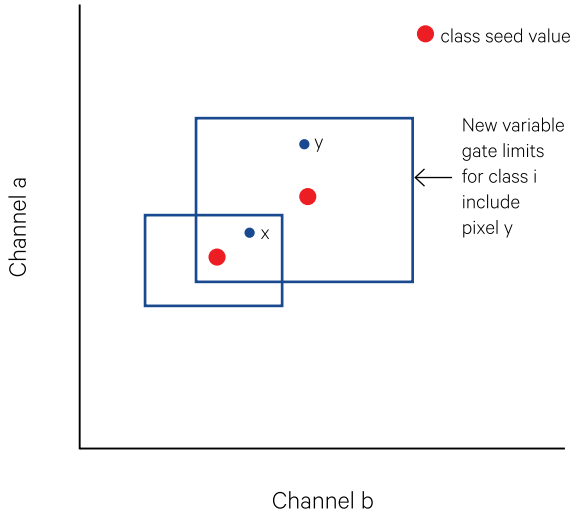
Figure 9.18 Channel gates

Channel gates are shown for two classes, i and j, in relation to two pixels, x and y.

a. Fixed gates use the same gate value for all classes.



b. Variable gates allow different classes to attract outliers at varying distances.



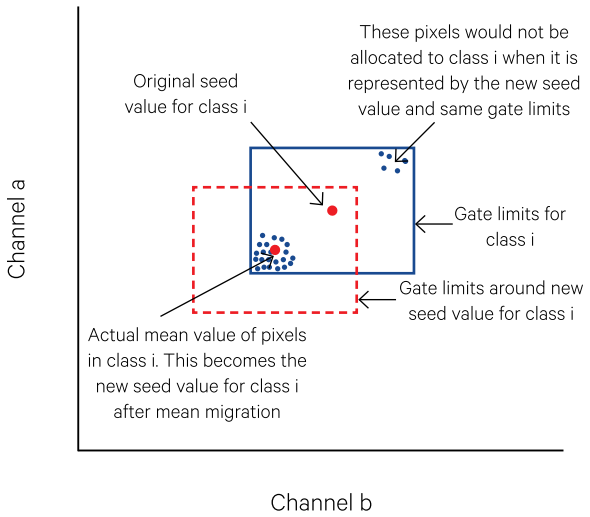
Source: Harrison and Jupp (1995) Figure 29

The results of this classification can be represented as a 'mean' image, in which each classified pixel is given the seed values of its class in each image channel involved in the allocation. For the classification to adequately represent the original image data, the mean image should be similar to the raw data image. Similarity can be measured by a 'residual' image in which the value of each classified pixel is the (shifted) difference between the raw data and mean image values (Jupp and Mayo, 1982). When the mean image has captured the variation of the raw data image, the residual image should appear as random noise. Accordingly, locations where pattern is visible in the residual image may indicate where additional classes are required.

Another image processing operation that is often used in conjunction with the minimum distance classifier is called 'mean migration'. This process is particularly valuable when new seeds have been created automatically, based on a seed value that is not central to its membership. Migrating means, as the name implies, simply involves computing the actual mean for the pixels allocated to a given class, then using this mean as the new seed value. The function of this step is to centre the seed within the current class membership as shown in Figure 9.19. The mean migration process should be followed by another allocation stage so that all pixels are given the 'opportunity' to be associated with the closest class seed, especially when there are several class seeds in close spectral proximity (see Volume 2E—Section 6.2.2).

Figure 9.19 Migrating means process

During mean migration, the actual mean of all pixels allocated to a class is computed. In this example, most pixels allocated to class i are not close to the class seed, so the mean of the pixels in that class 'migrates' the class seed closer to them. In many cases, some pixels originally allocated to that class would no longer be included in it.



Source: Harrison and Jupp (1995) Figure 38

The same process can be used to compute class seed values in terms of a different set of image channels. This 'spectral transfer' operation is particularly useful when a classification developed using a transformed set of channels (such as PCs) is to be transferred back to the original image channels (see Volume 2E—Sections 4.2.4 and 5.3.2).

9.2.4 Maximum Likelihood classifier

Although widely used for classifying EO imagery, the Maximum Likelihood classifier is conceptually more complex and computationally more expensive than most other classification approaches. The foundational statistics for Maximum Likelihood are only briefly introduced below, but the interested reader can discover considerable detail on this topic in Volume 2E—Sections 2 and 7.

While the classifiers described in Sections 9.2.1 to 9.2.3 effectively define decision regions in the image spectral space that correspond to interpreter categories, with multiple decision regions often becoming associated with a single interpreter category during the labelling phase (see Section 9.3), Maximum Likelihood classification relies on probability methods to directly associate interpreter categories with image classes via the following logical steps:

- assume a parametric form for the class distributions;
- estimate the parameters of the distributions from training sets;
- compute the posterior probability that a pixel belongs to each class; then
- allocate that pixel to the class for which it has highest probability of membership.

This method rests on the assumption that pixels known to be in an interpreter category have attributes that are (independent) samples from an underlying probability distribution which is characteristic of that category. Using conditional probabilities (see Volume 2E—Section 2.3.2.1), this distribution may be denoted:

$$P\left(\boldsymbol{x} \mid C_j\right)$$

where $\boldsymbol{x}$ is the set of attributes observed for the pixel. The notation indicates the probability that a pixel from scene class C_j will have the attributes $\boldsymbol{x}$, that is, the probability of a pixel having the values $\boldsymbol{x}$ given the condition that it is a sample from scene class C_j. The overall probability of pixels in the image taking a specific value is assumed to follow the mixture distribution:

$$P\left(\boldsymbol{x}\right) = \sum_{j=1}^{N} q_j P\left(\boldsymbol{x} \mid C_j\right)$$

where the q_j are the relative class densities. The 'reciprocal' probability, which allocates pixels with known attribute values but unknown class membership to one from a set of possible classes, is referred to as the posterior probability or likelihood depending on whether the Bayesian or Likelihood approach is taken (see Volume 2E—Sections 2 and 7). Thus:

$$P\left(C_j \mid \boldsymbol{x}\right)$$

is the probability that a pixel with attribute values $\boldsymbol{x}$ is a member of class j. Bayes' Theorem (see Volume 2E—Section 2.3.2.1) derives this conditional probability as:

$$P\left(C_j \mid \boldsymbol{x}\right) = \frac{q_j P\left(\boldsymbol{x} \mid C_j\right)}{P\left(\boldsymbol{x}\right)}$$

where

q_j is the probability of class j occurring; this can be derived from some additional information about the class, such as the expected frequency of occurrence, and is often referred to as the 'prior' probability;
$P(\boldsymbol{x})$ is the probability of attribute values $\boldsymbol{x}$ occurring in the image and can be estimated from the distribution of all pixels in the image; and
C_j denotes scene class j.

Prior probabilities can be used to introduce non-numerical data into the Maximum Likelihood classifier as a 'cascade' of probabilities. Such implementations of the Maximum Likelihood classifier allow the prior probabilities to be derived from non-image sources, for example a GIS data plane (see Volume 2E—Section 9.5.3).

Since this classifier tends to associate one image class with one interpreter category, the extent to which it is appropriate for a particular project depends heavily on whether the assumptions about class distributions are valid for the image being classified. In general, only the Gaussian distribution is used and the method is often called the 'Gaussian Maximum Likelihood' method. That is (see Volume 2E—Section 2.4):

$$P\left(\boldsymbol{x} \mid C_j\right) = \left(2\pi\right)^{-\frac{n_c}{2}} \left|\boldsymbol{S}_j\right|^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}\left(\boldsymbol{x}-\mu_j\right)^T \boldsymbol{S}_j^{-1}\left(\boldsymbol{x}-\mu_j\right)}$$

where

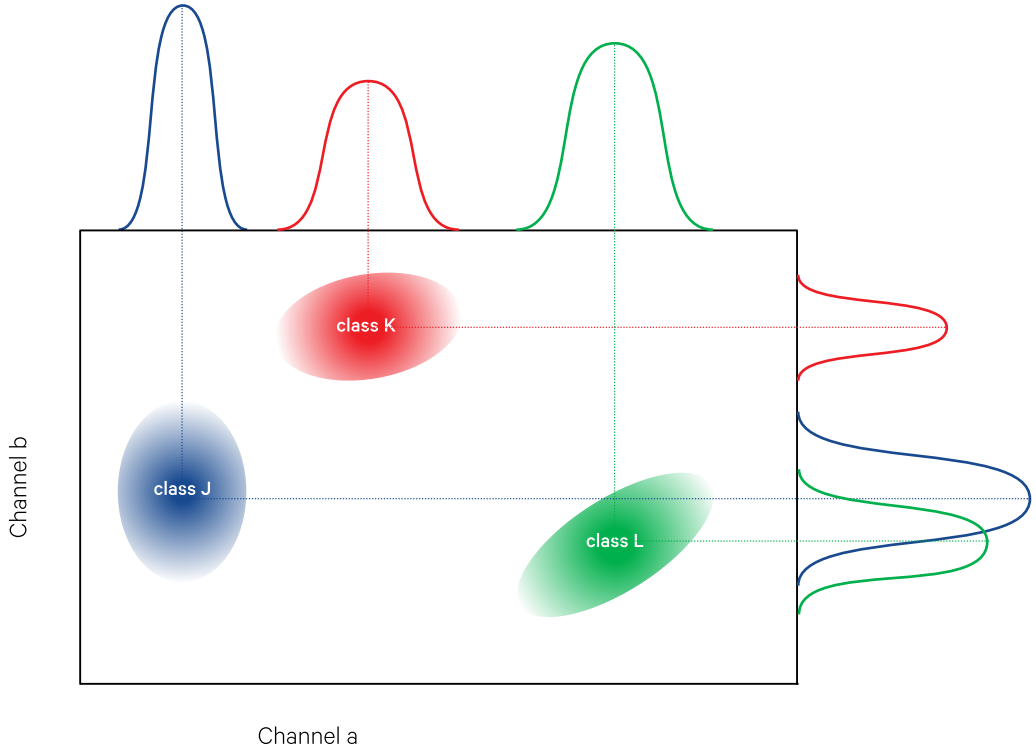
$\boldsymbol{x}$ is the vector of attribute values for the pixel;
C_j denotes scene class j;
n_c is the number of attributes;
$\boldsymbol{S}_j$ is the variance/covariance matrix for class j; and
μ_j is the mean vector for class j.

Other distributions that can be used for Maximum Likelihood classification are discussed in Volume 2E—Section 7.2.3.

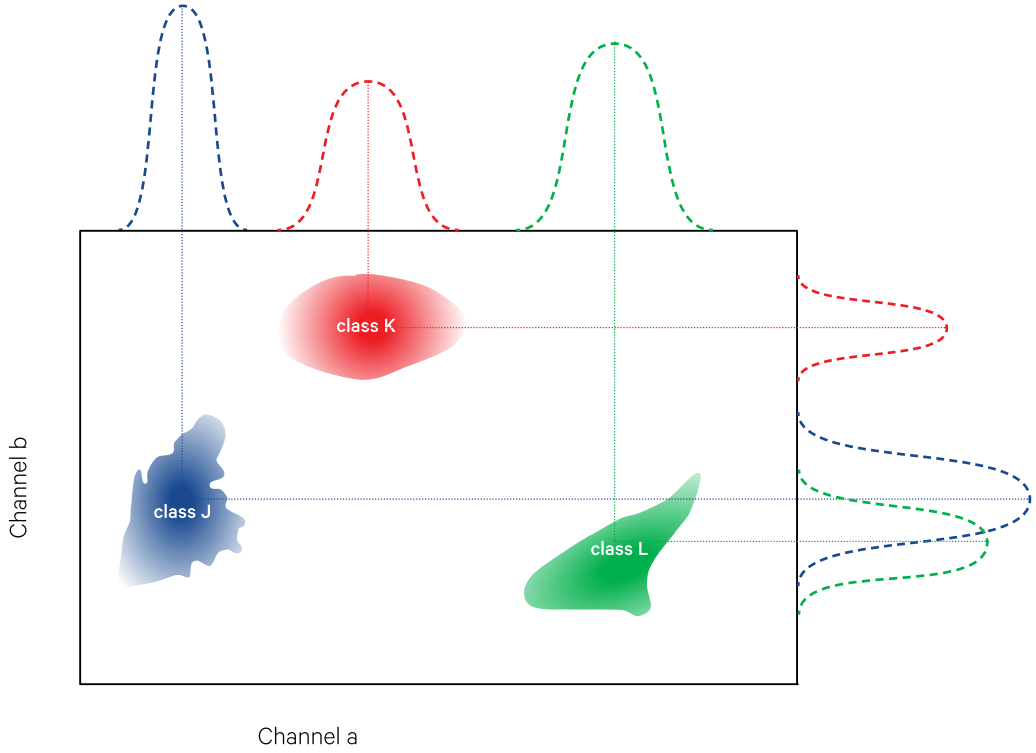
The parameters defining this distribution are generally estimated from training. If some estimates for the class prior probabilities are given, pixels can be allocated to the available classes. With no estimates available for these parameters, they are often taken as equal for each class and normalised to sum to one. The underlying philosophy of this approach assumes that the interpreter categories (such as land cover types) are well represented in the image data as normal distributions around class mean values (see Figure 9.20a). Unfortunately, in most EO images, the actual distributions of image values associated with particular ground features are more complex (see Figure 9.20b) .

Figure 9.20 Maximum Likelihood classification

a. Maximum likelihood classification is based on the premise that image classes can be modelled as statistical distributions.



b. Within a given range of band values, the distribution of class values in EO imagery can be quite irregular.



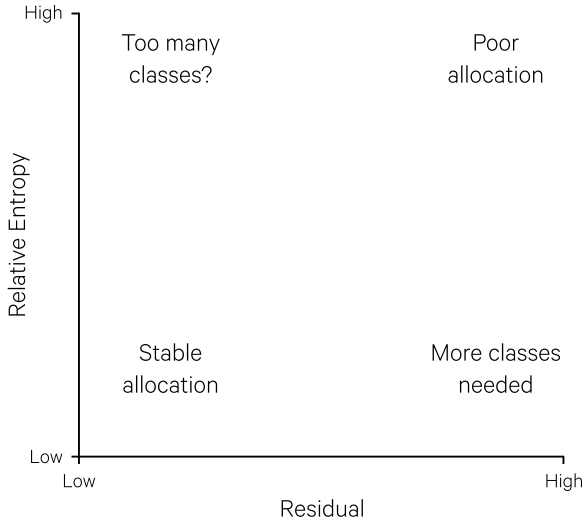
Since Maximum Likelihood classifiers are computationally expensive, especially for large numbers of classes and/or channels, other approaches are often preferred, particularly when combined with post-allocation labelling (see Section 9.3 and Volume 2E—Section 7). Since it is possible that some image pixels do not actually belong in any of the 'available' classes, some versions of the Maximum Likelihood classifier define a rejection probability, below which a pixel will not be allocated to the closest class. Additional statistics may be provided by some implementations of the Maximum Likelihood classifier such as:

- residual—indicates whether a pixel is typical of its allocated class. The residual value is high when a pixel is considered to be a significant 'probability-based distance' from the class mean value, relative to the class distribution;
- relative entropy—indicates whether a pixel has high probability of belonging to several classes; and
- minimum support ratio—equals the difference in probabilities between the allocated class and the next likely class. This highlights those classes whose allocation may change with different prior probabilities.

For example, as illustrated in Figure 9.21, while low 'entropy' and low 'residual' values would indicate stably allocated pixels, those pixels with high 'residual' and high 'entropy' values might be poorly allocated among the available classes. Conversely, pixels with low 'residual' and high 'entropy' values would suggest that there is considerable overlap between classes, and high 'residual' and low 'entropy' values might indicate the need for more classes.

Figure 9.21 Allocation using Maximum Likelihood classifier

Pixels with stable allocations have low residual values and low relative entropy values.



Source: Harrison and Jupp (1995) Figure 34

For more information related to the Maximum Likelihood classifier, please refer to Volume 2E.

9.3 Labelling Classes

The labelling stage involves assigning labels to the allocated spectral classes. The allocation results can commonly be represented as a single classification channel, in which each class is represented by a different numeric value. A classification channel displayed as a grey-scale image can then be interpreted using interactive painting (see Section 9.1). Class labelling is often an interactive process, allowing an interpreter to highlight the spatial extent and continuity of individual spectral classes. It is generally prudent during labelling to associate multiple spectral classes with a single ground-based category.

While the labelling stage implies that the allocation stage has been completed, the iterative nature of the whole classification process necessitates an assessment of each iteration to determine when it has reached completion. Just as statistical tools are available to select the most appropriate image (or transformed) channels (see Section 9.1.3), a range of methods have been developed to highlight where there are too few or too many classes to efficiently capture the intrinsic image variation. These methods are introduced in Section 9.1.2 and discussed in greater detail in Volume 2E. For example, to assess the contribution of a particular class or group of classes to the overall classification, some of the statistics summarised in Section 9.1.2.1 could be computed for classifications with and without those classes (see also Figure 9.12).

9.3.1 Clustering tools

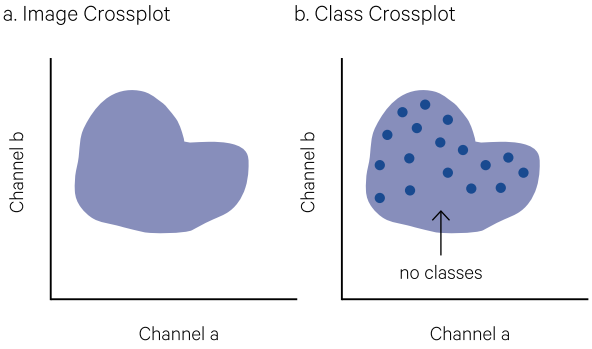
Various spectral and spatial analyses of the classes can also be used at this stage to combine spectral classes into a single interpreter category. It may be appropriate to apply relevant transformations to seed values in order to sort classes by a specific criterion (such as vegetation greenness). Graphically plotting class seed values on crossplots of selected pairs of image (or transformed) channels is a valuable way of viewing the spread of classes in an image classification.

Some of the tools that are generally available for classification analysis include:

- sorting classes by a relevant attribute, such as brightness or greenness. Since most classifiers create class seeds in an arbitrary order, this may require that the classification be spectrally transferred onto channels that are ordered in some recognisable gradient, such as PC1 or a vegetation index.
- crossplotting the values of class seeds using pairs of significant channels (see Section 8.1.3)—comparing crossplots of image channels with crossplots of class seeds can be a useful way of highlighting spectral regions in the image that are not represented by the current set of classes (see Figure 9.22).
- canonical variates analysis (CVA) highlights class separation by maximising the differences between sets of classes relative to the variation within them. As with PCA, the first two canonical variates (CV1 and CV2) are the most informative with respect to class separation. Accordingly, to visualise class distribution and separation it is most useful to plot class seeds on a crossplot of CV1 versus CV2 (see Section 9.1.2.1 and Volume 2E—Section 5.1.3 and 5.3).
- taxonomic clustering of the basic image classes to group them into a given number of interpreter classes (see Section 9.1.2.1). Given the variations that can occur between distance metrics and clustering algorithms (see Volume 2E—Section 6), the first and last (or earliest and latest) aggregations are likely to be the most consistent and reliable.
- minimum spanning tree (MST) orders class pairs by their spectral or spatial proximity (see Section 9.1.2.1):
  - the Mahalanobis distance between class pairs can be used to create a spectral MST (see Section 9.1.2.2); and
  - the co-occurrence matrix, which reports the frequency with which each pair of classes occur as adjacent pixels, shows the spatial similarity of classes. Co-occurrence results can be used to generate a spatial MST in which the linkages indicate spatial association between pairs of classes (see Section 9.1.2.2 and Volume 2E—Section 6.4).

Figure 9.22 Highlighting class deficiencies

Comparing crossplots of image channels with crossplots of class seeds can be a useful way of highlighting spectral regions in the image not represented by the current set of classes.



Source: Harrison and Jupp (1995) Figure 57

9.3.2 Aggregating Classes

The process of aggregating classes, and associating each group of clases with an interpreter label, usually involves interactive image painting of a displayed classification channel to highlight selected classes and potential clusters (see Sections 5.2.1 and 6.1.2.2). Given that most image classifications involve large numbers of classes, it is useful to approach the interactive labelling stage in a structured way using the various clustering tools mentioned above and in Section 9.3.1. At this stage any ancillary data related to the imaged scene and/or the desired classes (such as relevant maps) would be consulted.

Some suggestions to prepare for labelling include:

- order classes so that visual interpretation of the classification channel is intuitive, for example, lighter grey shades would relate to higher reflectance or vegetation greenness;
- plot classes on crossplots of CV1 vs CV2. The relationship between the original channels and each canonical variate is summarised in the CVA transformation matrix (see Volume 2E—Sections 2 and 5). For most EO image classification involving land cover, CV1 is associated with brightness and CV2 is associated with vegetation greenness. Awareness of the gradients along the axes of the CV crossplot will help to label the image classes;
- compute spectral and spatial MST and plot each onto crossplots of CV1 vs CV2 using different colours to represent different spectral/spatial distance ranges. When distance information is included in this way it is obvious to the interpreter which classes are closest;
- compute dendrogram of classes and note the largest and smallest clusters; and/or
- highlight on the MST and/or dendrogram those classes that contain the largest numbers of pixels.

Interactive painting involves display of a classification channel as a grey scale image, then changing selected classes to specific colours. By highlighting a class this way, both its spatial extent and cohesion can be visualised. At this stage, any ancillary data related to the imaged scene and/or the desired interpreter classes (such as relevant maps) would also be consulted. Some guidelines for interactive painting include:

- large classes are best coloured first and labelled using ancillary information so that their region of the MST can be identified;
- using a different colour, the 'closest' class can then be highlighted. If these two classes form coherent patterns, they are likely to belong to the same group. Reference to ancillary information can help to decide when classes should be combined. This process is repeated for all 'adjacent' classes;
- given that classes tend to be linked along environmental gradients in the MST, it is valuable to identify the extremes of the MST early in this process, unless those classes are too small or indistinct; and
- the danger in moving along a particular gradient is that there is a tendency to keep adding classes to one group. It is beneficial to move to a different part of the MST once several classes have been combined.

The result of this process should be that all image classes become associated with several class groups, which are assumed to represent the interpreter classes. The interpreter class groups are often presented as a thematic image and methods for selecting appropriate colours for this are discussed in Excursus 9.2. Those groups can also be recorded as a classification channel, so that the spectral transfer process could be used to compute a mean and residual image for the groups. These images could then be used in the same way as described in Section 9.3.1. to determine whether the aggregated groups are spectrally homogeneous (see Volume 2E—Section 9).

Excursus 9.2—Colour Selection

One of the principal uses of image processing is to identify image features and patterns and highlight their differences and similarities. Colour is frequently used to differentiate classes (or identified numeric relationships) in visual assessment and interpretation. An inappropriate selection of colours for this task can mean that the perceived colour differences do not consistently reflect differences in the numeric values.

The selection of effective colour sequences requires consideration of:

- our perceptual discrimination of colours (see Excursus 5.1);
- their representation on the display/hardcopy device;
- the size and shape of the features being represented;
- our perception of appropriate colours for the ground features being represented;
- the relationship between the different features that we are trying to portray; and
- the interactions between different colours in adjacent features.

In image painting or presentation of classification results, colour selection can be associated with the progression of image values to form natural sequences. The selection of colours using uniform colour spaces can attempt to compensate for the differences between measured and perceived properties of colours. Sequences are best selected using some logic such as following a continuous path in a colour space model and/or restricting colours to planes of constant value or intensity. For example, all colours may have the same hue, or be selected from the sequence of mixtures between two complementary hues. Colours are best spaced at equal perceptual distances in any colour model (as discussed in Appendix 3, linear interpolation between colours in different models produces different results in colour spacing due to the non-linear scaling of the models).

Progressions in the RGB colour space can be easily defined in most image processing systems. The expression of composite colours in terms of their component proportions of the red, green and blue primaries allows a sequence to be developed by varying only one component primary. A number of sequences can also be linked to cover a larger range of colours. For example, starting with a reddish brown, we can increment the amount of red up to the maximum level then add increasing amounts of green to stop at a partially saturated yellow. The colours in this sequence will display a natural progression due to their relationship in the colour space.

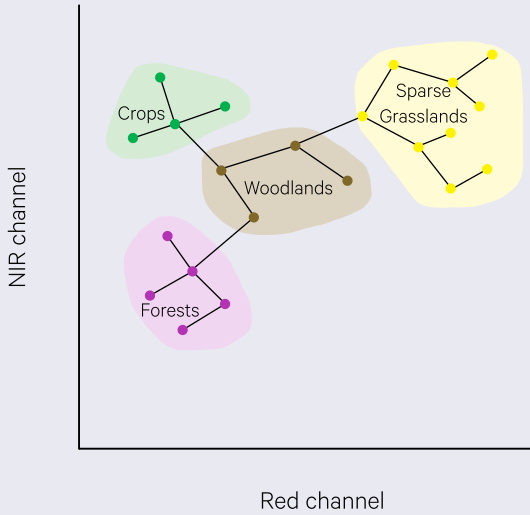
The gradual changes between interpreter categories, such as forest to woodland to open woodland to grassland could be effectively represented as a sequence of related colours such as bright green through darker greens to browns. These sequences may associate specific colour groups with different land covers, for example:

- purples for forest classes;
- browns for woodland classes;
- greens for crop classes; and
- yellows for sparse grasslands and bare ground classes.

The key 'node' points in the MST can be used as the point of changing colour sequences as shown in Figure 9.23. Thus, the MST not only provides a useful base for defining the progression but it can also be used as an effective colour key and implicitly incorporate some spectral (or spatial) information in the descriptive labels. Some methods for using different colour spaces to determine equal perceptual intervals between colours are given in Appendix 3.

Figure 9.23 MST colour sequences.

The minimum spanning tree, which is used for classification labelling in the Mosaic Model approach (see Excursus 9.3 and Volume 2E), can be used to coordinate colour selection for the final classes. In this example, a sequence of yellows are used to map sparse grassland classes, a sequence of browns map woodland classes, a sequence of purples map forest classes and a sequence of greens map crop classes.



Source: Harrison and Jupp (1990) Figure 56

9.3.3 Post-classification processing

To improve the cohesion of a labelled classification, it may be appropriate to apply a 3x3 modal filter to the interpreter group classification channel before final presentation. This process replaces the central value of the filter with the mode (most popular value) of its neighbouring pixels (see Volume 2C), often with constraints such as a minimum number of pixels having the same interpreter class. Modal filtering as a post-classification process is a very effective method of removing small, spurious patches, which enables greater compression of classification results and reduces computational needs for any subsequent GIS processing. Given that for most EO imagery the optical pixel size is larger than the geometric pixel size (see Volume 1), this process should not be viewed as undermining data fidelity, but simply removing imaging artefacts in the classification context. Other options for post-classification processing are discussed in Volume 2E—Section 9.

9.4 Verifying Results

Class labelling based purely on statistical data is rarely satisfactory for a detailed interpretation of EO imagery, so verification with ground data (or 'ground-truthing') within a multi-stage statistical sampling regime is usually required. Most image processing systems contain statistics routines for the final or verification stage of the classification process.

Image processing systems typically provide a number of options for generating sample pixels and comparing the purity of spectral class groupings against the labels supplied by an interpreter. Where an ancillary data channel represents an alternative 'classification', possibly based on GIS data, the classification and an ancillary data channels may also be cross tabulated. This process allows every pixel in the image to be included in the tabulation.

The verification stage effectively checks the degree to which the labelled classification represents ground features. Many analyses cross tabulate the spectral classes versus the labelled classes to determine the degree of agreement. This cross tabulation is commonly called a contingency table. A range of statistics can be derived from the contingency table to summarise the extent to which the spectral classes correspond to the labels groups (see Excursus 9.1). This topic is detailed in Volume 2E—Section 10.

Excursus 9.3—Mosaic Model Classification

Source: Harrison and Jupp (1990)
Further information: Volume 2E; Jupp *et al.* (1985a, 1985b)

Section 9.1 above introduced the concepts of supervised, unsupervised and hybrid or 'mixed mode' classification. The processes involved with each of these approaches are summarised in Figure 9.1, Figure 9.2 and Figure 9.3 respectively. In one version of hybrid classification, the underlying model used for both an EO image and the land cover it represents is the 'mosaic' model (Jupp, 1988). This can be considered as a mixed mode approach to classification (see Figure 9.3). The mosaic model approach provides a sophisticated methodology for flexible class generation, labelling and verification. This model views the image or imaged scene as a composite of coloured patches or 'components' that do not overlap. In this context, for EO imagery the term 'colour' can be considered as a generalised one for spectral signature, but in other data sets it may simply be a set of features or attributes attached to each patch.

Mosaic modelling

Both the scene and an image of the scene can be modelled as a mosaic, although the nature and purpose of the two models may be different. The extent to which the two models can be reconciled defines the value of EO and image processing as a means for inferring the scene 'mosaic' from the image mosaic model.

A patch model of the underlying scene is usually based on a 'logical' assignment of labels or descriptions to the patches. These may be geographical labels or descriptions of contents in terms of land cover categories. The scene model describes the land surface from a particular viewpoint and often for a particular purpose or application. Patches composing the models for real scenes obviously need to be hierarchical in that a 'patch' at one scale, such as an area of open forest, is composed of smaller components at a larger scale, such as tree crowns and areas of bare soil. Similarly, patches at one scale may be part of larger units at a smaller scale.

The patch model of an image considers the pattern and form of the image. Its purpose is to accurately represent the image structure and form rather than incorporate a viewpoint or particular application. The pixel size of an image effectively defines the finest scale for the patch size in the image model. Patches larger than individual pixels are usually required in EO analyses. Geographic mapping typically imposes a logical or descriptive patch structure on the landscape. This logical description of the image may be based on different considerations from the spectral patch structure. One objective of exploratory image analysis is to establish the degree to which they might be brought into coincidence. If possible, this is accomplished by the way logical class labels are assigned to the image-based patches.

Methodology

The mosaic model can be realised by first classifying the image into classes, which effectively creates patches of connected groups of pixels with similar image properties. Given the variability of the real world, these patches are usually rather ragged and spotty, however (if necessary) such problems can be overcome with subsequent processing.

Class seed values may be generated in three ways:

- interactive training;
- theme statistics; or
- unsupervised classification.

The result of the allocation stage is not assumed to be a one-to-one classification of ground features, but should be a mosaic model that represents the spectral (colour) variation in an image precisely. In this case, the 'residual' image (the difference between the raw data image and the 'mean' image—Section 9.2.3) has no significant unmapped features, that is, there would be no coherent features visible in the residual image. Volume 2E describes various statistical analyses that may be used during the generation of the image mosaic model to achieve these requirements.

The Mosaic Model approach effectively combines the processes known as sub-class modelling and post-classification labelling (Haralick, 1976; Strahler *et al.*, 1981). Sub-class modelling, as the name implies, models the image as a set of spectral classes. These classes are defined statistically with the goal of representing the spectral variation of the image. Once pixels have been allocated to these defined sub-classes, the interpreter assigns an appropriate ground feature label to each class in the post-classification labelling stage (see Section 9.3). A major advantage of the mosaic model approach is the flexible way in which the image mosaic can be labelled for different interpretation purposes. Once our image model represents the essential spectral variation of the image, the classes may be aggregated in different ways to indicate different scene features or different levels of detail within the features.

Spectral ambiguities, which cause different ground features to be represented by the same image values, become a problem in the labelling stage. This is effectively the point at which the scene model cannot be reconciled with the image mosaic model by simply using spectral data. In such cases, additional image channels, patch shape or ancillary data may need to be used to resolve ambiguities (see Volume 2E).

For example, two image patches that have the same colour can occur in different contexts and represent different land covers. This commonly occurs in EO imagery for the tropical land covers of mangroves and rainforest. In images that contain both land covers, spectral classification results are typically confused since both cover types represent healthy vegetation that absorbs blue and red wavelengths and strongly reflects near infrared radiation. In this case however, ancillary data such as a 5 m elevation contour could be used to separate the two land covers, which typically occur at different elevations.

The Mosaic Model classification approach can be viewed as a hierarchical model in terms of the level of detail contained within both the image mosaic and the labelled mosaic (or scene interpretation). This method can also be applied iteratively, using progressive aggregation of classes. For example, for an initial classification comprising 100 spectral classes, the first aggregation may reduce the number of classes to 60, then the second to 40, followed by the third and final aggregation resulting in 25 labelled groups.

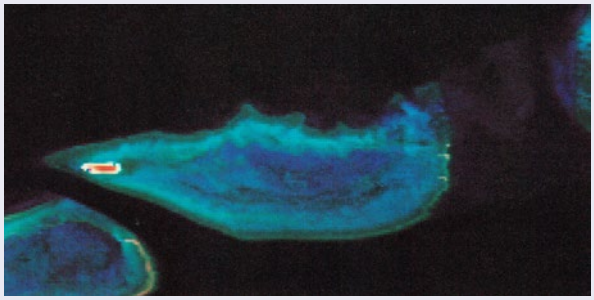
Obviously, a greater level of detail in the image mosaic (indicated by class separability not simply the number of classes) will require more effort during both the classification and labelling stages, while a larger number of reliably labelled groups in the labelled mosaic will require more effort during the labelling stage. The value of this detail however depends on the purpose(s) of the analyses. Volume 2E describes the sequence of operations required to produce and verify a detailed labelled mosaic.

Example classification

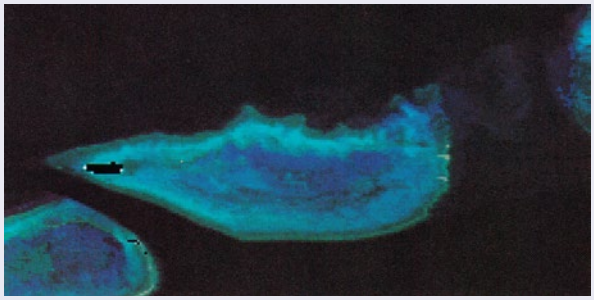
A Landsat MSS image of Heron Island Reef, Queensland, was classified in terms of reef cover classes (Jupp *et al.*, 1985a; Jupp *et al.*, 1985b). The original, mean and residual images for this classification are shown in Figure 9.24.

Figure 9.24 Heron Island Reef example

a. Original image—Landsat MSS image of Heron Island Reef.



b. Mean image—classification into 85 classes with each pixel being represented by the mean class value.



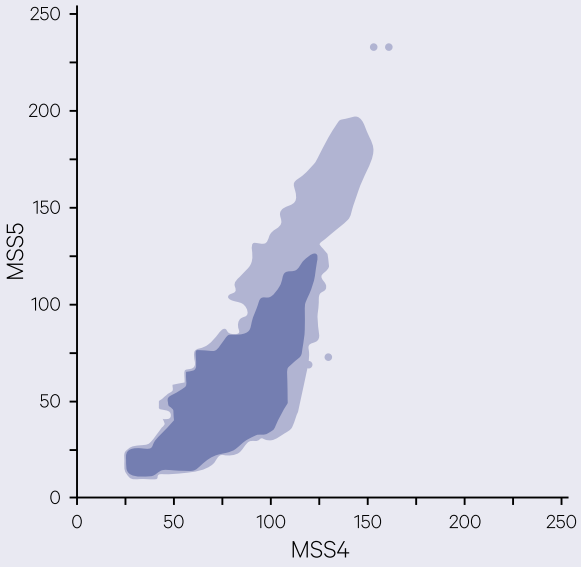
c. Residual image represents the difference between the mean and original images.

Figure 9.25 illustrates the results of image classification using image crossplots which correspond to the original, mean and classified images. (Pixels are actually allocated to classes on the basis of their values in all four image channels but the data variation is shown in only two dimensions here for simplicity of presentation.) The crossplot of the mean image represents the same data range as the original image but with only 85 different image values. The effectiveness of this data reduction stage can be easily seen in the colour similarities between the original and mean images in Figure 9.24a and Figure 9.24b respectively. This data simplification has not significantly reduced the variation in the image data; it simply represents it in a form that can be manipulated more easily.

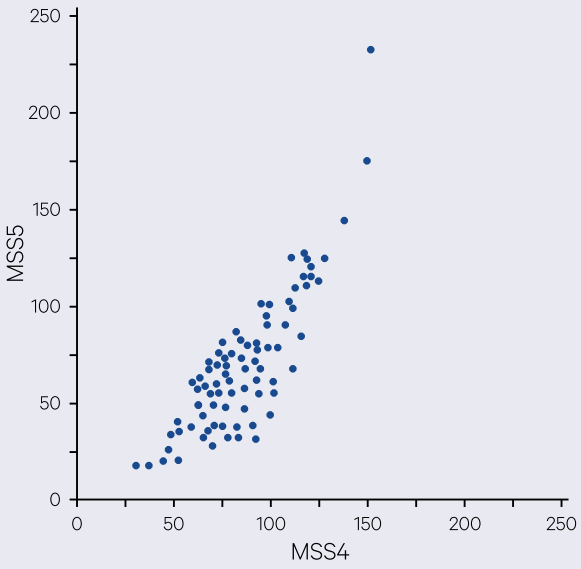
Figure 9.25 Image classification sequence

An example of Mosaic model classification based on a Landsat MSS image of Heron Island Reef.

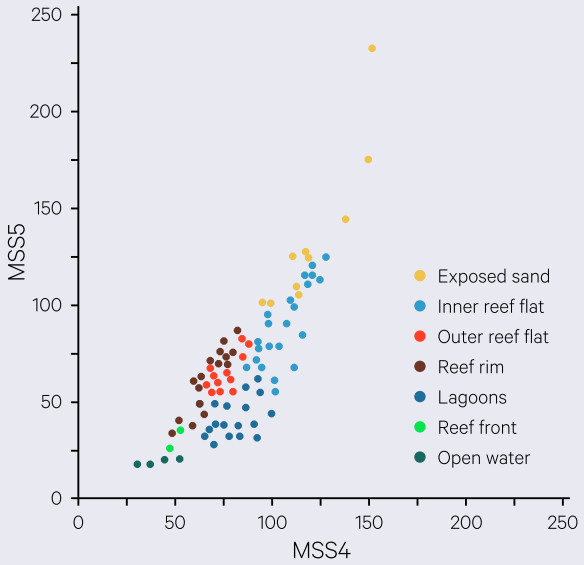
a. A crossplot of bands 4 and 5 in the original image shows the distribution and variation of data values.



b. A crossplot of the mean image (in which each pixel assumes the seed values of the class to which it has been allocated) shows that the classification summarises the distribution of the original image data range but contains less variation. In this case only 85 classes are used.



c. The mosaic model classification methodology allows the 85 classes to be flexibly grouped and labelled. This crossplot shows a labelled mosaic of seven broad categories.

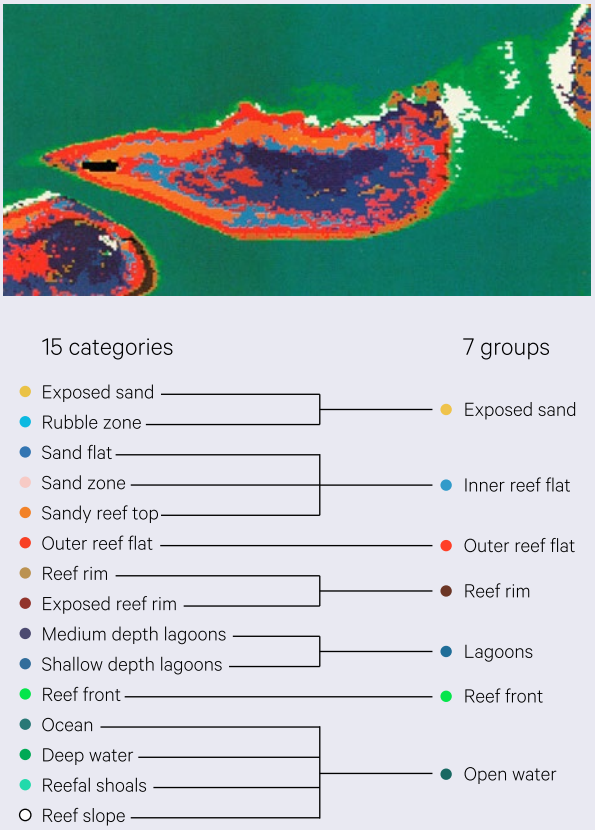


Source: Harrison and Jupp (1990) Figure 54

The labelling stage has reduced the previous 85 spectral classes into 7 major (and 15 minor) categories that represent differences in reef cover type (see Figure 9.26). As introduced in Section 9.3.1 and detailed in Volume 2E, the procedure for aggregating classes can involve statistical analyses of both spectral and spatial similarity between classes in conjunction with interactive labelling. One advantage of the Mosaic Model approach is that these categories could be redefined at a later stage to provide a more, or less, detailed analysis of reef cover, or to represent a different ground feature such as water depth.

Figure 9.26 Classified image

The classification has been labelled into 15 categories which can be further reduced to 7 broad groups.



Source: Harrison and Jupp (1990) Plate 12

9.5 Further Information

Image Classification

SEOS: eLearning Tutorials: *http://www.seos-project.eu/modules/classification/classification-c00-p01.html*

GIS Geography: *http://gisgeography.com/image-classification-techniques-remote-sensing/*

9.6 References

Baker, J. R., Briggs, S. A., Gordon, V., Jones, A. R., Settle, J. J., Townshend, J. R. G., and Wyatt, B. K. (1991). Advances in classification for land cover mapping using SPOT HRV imagery. *International Journal of Remote Sensing*, 12(5), pp. 1071-1085. doi:*http://dx.doi.org/10.1080/01431169108929711*.

Bolstad, P. V., and Lillesand, T. M. (1992). Semi-automated training approaches for spectral class definition. *International Journal of Remote Sensing*, 13(16), pp. 3157-3166. doi:*http://dx.doi.org/10.1080/01431169208904108*.

Bryant, J. (1979). On the clustering of multidimensional pictorial data. *Pattern Recognition*, 11(2), pp. 115-125. doi:*http://dx.doi.org/10.1016/0031-3203(79)90057-8*.

Bryant, J. (1989). A fast classifier for image data. *Pattern Recognition*, 22(1), pp. 45-48. doi:*http://dx.doi.org/10.1016/0031-3203(89)90037-X*.

Bryant, J. (1990). Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm. *AMOEBA clustering revisited*, 56, pp. 41-47.

Burrell, J. P. (1973). *Vegetation of Fowlers Gap Station. Lands of Fowlers Gap Station*, New South Wales. Fowlers Gap Arid Zone Research Station Series No. 3 (Ed: J. A. Mabbutt), University of New South Wales, Sydney.

Goldberg, M., and Shlien, S. (1976). A Four-Dimensional Histogram Approach to the Clustering of Landsat Data. *Canadian Journal of Remote Sensing*, 2(1), pp. 1-11. doi:*http://dx.doi.org/10.1080/07038992.1976.10854944*.

Goldberg, M., and Shlien, S. (1978). A Clustering Scheme for Multispectral Images. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(2), pp. 86-92. doi:*http://dx.doi.org/10.1109/TSMC.1978.4309905*.

Haralick, R. M. (1976). Automatic remote sensor image processing. 'Digital picture analysis', Vol. 11, pp. 5-63. Springer-Verlag, Berlin and New York.

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Hashim, M. (1996). Image Classification with Higher-Dimension Watershed Classifier: A Test to Malaysian National Land Use Classification Scheme. *Bulletin Ukur, Penerbitan Akademik Fakulti Kejuruteraan dan Sains Geoinformasi*, 7(2), pp. 141-148.

Jupp, D. L. B. (1988). *The mosaic model for microBRIAN Classification*. Paper presented at the First microBRIAN User Group meeting, Brisbane, Australia.

Jupp, D. L. B., Heggen, S. J., Mayo, K. K., Kendall, S. W., Bolton, J. R., and Harrison, B. A. (1985a). *The BRIAN Handbook. An Introduction to Landsat and the BRIAN (Barrier Reef Image Analysis) System for Users, CSIRO Division of Water and Land Resources, Natural Resources Series No.3*. CSIRO, Canberra, Australia.

Jupp, D. L. B., and Mayo, K. K. (1982). The use of residual images in Landsat image analysis. *Photogrammetric Engineering and Remote Sensing*, 48(4), pp. 595-604.

Jupp, D. L. B., Mayo, K. K., Kuchler, D. A., Heggen, S. J., Kendall, S. W., Haywood, M. J., and Ayling, T. (1985b). *A Landsat-based interpretation of the Cairns Section of the Great Barrier Reef Marine Park. Interpretation of Landsat data by Computer-based Classification and Labelling, CSIRO Division of Water and Land Resources, Natural Resources Series No.4*. CSIRO, Melbourne, Australia.

Lewis, M. M. (1998). Numeric classification as an aid to spectral mapping of vegetation communities. *Plant Ecology*, 136 (2), pp. 133–149.

Mabbutt, J. A., Burrell, J. P., Corbett, J. R., and Sullivan, M. E. (1973). *Land Systems of Fowlers Gap Station*. Lands of Fowlers Gap Station, New South Wales. Fowlers Gap Arid Zone Research Station Series No. 3 (Ed: J. A. Mabbutt), Sydney.

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings National Institute of Science, India*, 2(1), pp. 49-55.

McKay, R. J., and Campbell, N. A. (1982a). Variable selection techniques in discriminant analysis: I. Description. *British Journal of Mathematical and Statistical Psychology*, 35(1), pp. 1-29. doi:*http://dx.doi.org/10.1111/j.2044-8317.1982.tb00638.x*.

McKay, R. J., and Campbell, N. A. (1982b). Variable selection techniques in discriminant analysis: II. Allocation. *British Journal of Mathematical and Statistical Psychology*, 35(1), pp. 30-41. doi:*http://dx.doi.org/10.1111/j.2044-8317.1982.tb00639.x*.

McVicar, T. R., Ahmad, W., and Hill, G. (1989). *Landsat Thematic Mapper data for management of Conondale State Forest, Queensland*. Paper presented at the QRSA 1989 Remote Sensing Seminar, Brisbane, Australia.

Mentis, M. T. (1981). Evaluation of the wheel-point and step-point methods of veldt condition assessment. *Proc. Grassland Soc. Southern Africa*, 16, pp. 89–94.

Narendra, P. M., and Goldberg, M. (1977). A non-parametric clustering scheme for landsat. *Pattern Recognition*, 9(4), pp. 207-215. doi:*http://dx.doi.org/10.1016/0031-3203(77)90005-X*.

Rao, C. R. (1948). Tests of Significance in Multivariate Analysis. *Biometrika*, 35(1/2), pp. 58-79. doi:*http://dx.doi.org/10.2307/2332629*.

Richards, J. A. (1986). *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag, Berlin, Germany.

Schachter, B. J., Davis, L. S., and Rosenfeld, A. (1976). Scene segmentation by cluster detection in color spaces. *SIGART Bulletin*(58), pp. 16-17. doi:*http://dx.doi.org/10.1145/1045264.1045267*.

Skidmore, A. K. (1989). Unsupervised training area selection in forests using a nonparametric distance measure and spatial information. *International Journal of Remote Sensing*, 10(1), pp. 133-146. doi:*http://dx.doi.org/10.1080/01431168908903852*.

Strahler, A. H., Franklin, J., Woodcock, C. E., and Logan, T. (1981). *FOCIS: A forest classification and inventory system using Landsat and digital terrain data* (NFAP 255). Nationwide Forestry Applications Program, USDA, Houston, United States.

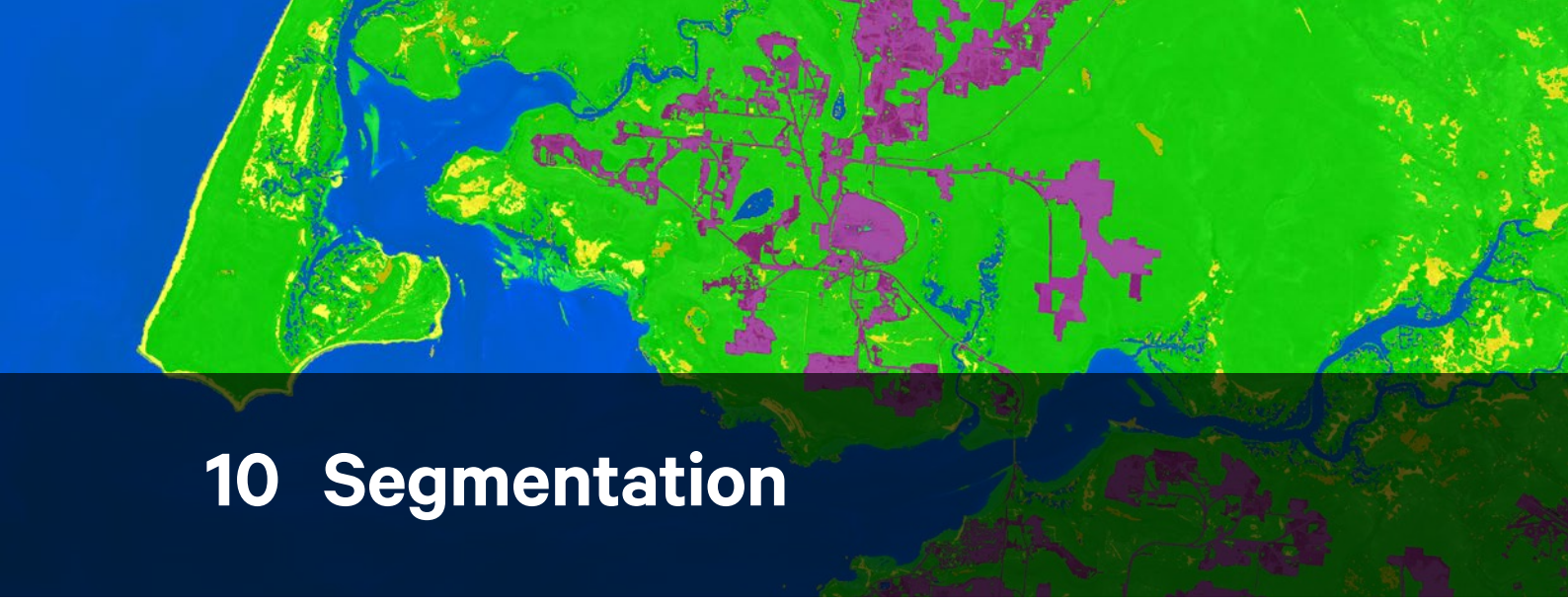
Swain, P. H., and Davis, S. M. (1978). *Remote Sensing: The Quantitative Approach*. McGraw Hill, New York, United States.

Thomas, I. L., Ching, N. P., Benning, V. M., and D'Aguanno, J. A. (1987). A review of multi-channel indices of class separability. *International Journal of Remote Sensing*, 8(3), pp. 331-350. doi:*http://dx.doi.org/10.1080/01431168708948645*.

Watson, A. I. (1987). A new method of classification for Landsat data using the 'watershed' algorithm. *Pattern Recognition Letters*, 6(1), pp. 15-19. doi:*http://dx.doi.org/10.1016/0167-8655(87)90044-4*.

Watson, A. I., Vaughan, R. A., and Powell, M. (1992). Classification using the watershed method. *International Journal of Remote Sensing*, 13(10), pp. 1881-1890. doi:*http://dx.doi.org/10.1080/01431169208904237*.

Young, M. T., and Fu, K.-S. (1986). *Handbook of Pattern Recognition and Image Processing*. Academic Press, Berkeley, United States.



10 Segmentation

Segmenting describes the process of dividing an image into two or more regions on the basis of:

- its spatial patterns (see Section 10.1);
- its pixel values (see Section 10.2); or
- ancillary data (see Section 10.3).

The segmentation may be required before various forms of image analysis, for example to remove cloud or other features that are not of interest prior to classifying an image. Alternatively, stratifying an image into regions, possibly using ancillary data such as cadastral boundaries, allows statistics to be derived for each sub-region. Recent advances in image processing systems allow sophisticated forms of image segmentation to be used in conjunction with standard classification approaches to analyse EO imagery (see Excursus 10.1).

If the regions can be clearly defined on the basis of pixel values, such as cloud or water features (see Figure 4.10), they may be separated using spectral themes. For regions that are visible from spatial patterns in the image (for example, irrigation channels, roads or rivers), the boundaries may also be defined by manually tracing with the screen cursor on a displayed image. When the regions are independent of, or indistinct within, the image, these boundaries may be defined from some ancillary data source, such as a map, using a line-digitising process. To incorporate digitised map boundaries into the image requires that a transformation exists between the coordinate systems of the relevant map and image data. The digitised strings would be converted into image coordinates and then be used to segment the imagery.

Segmentation algorithms analyse the spatial pattern relationships in an image to define boundaries for regions then allocate pixels to those regions based on their position in the image. This process attempts to mimic the role of an aerial photography interpreter and aims for spatially-coherent segments, starting with either:

- a single pixel, then 'growing' a region until boundaries are detected; or
- defined boundary edges, then grouping those pixels they enclose.

In most image processing systems, segmented regions may be identified in two ways in an output image. Pixels in the segments may be set to:

- a 'mask' or null value in all channels (in which case they can be 'ignored' in subsequent processing); or
- a user-specified value in a selected image channel.

Segmenting images using the null value allows regions of the image to be removed from the image and excluded from further processing, as may be required when only a portion of an image is being studied. Alternatively, segmenting with user-defined values allows the creation of ancillary data channels.

10.1 Image Spatial Patterns

Where a clearly visible spatial pattern in the image forms a boundary for image segmentation, cursor tracing of the boundary on the displayed image can be used to define the regions for segmentation. Linear features such as roads, rivers or irrigation channels, which frequently indicate cadastral or management zones, are usually selected for this type of segmentation. Due to 'jitter' problems with human operators, this method is generally less satisfactory than spectral segmentation for features such as coastlines where spectral differences (such as a near infrared channel) could be used to clearly differentiate between land and water.

Background image: Segmentation of Landsat-8 image over Weipa, Queensland, into four major land cover groups: water, vegetation, bare and infrastructure. The original image is shown in the banner of Section 8. **Source:** Norman Mueller, Geoscience Australia

Implementations of spatial segmentation can vary between image processing systems. Some of these options may include:

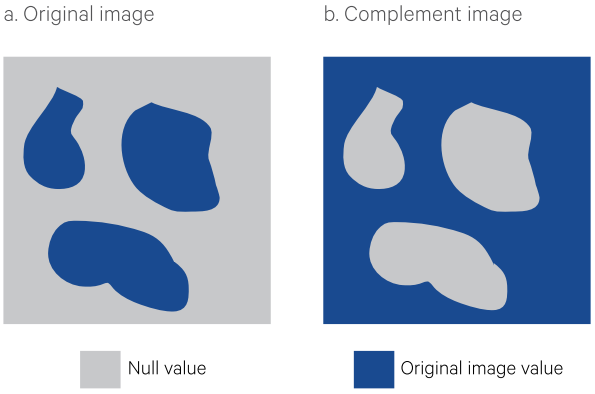
- excluding pixels inside a defined region by setting them to the null value;
- excluding pixels outside a defined region by setting them to the null value; or
- identifying pixels inside a defined region by setting them to a defined value in one or more channels.

Pixels that fall onto the region boundary are typically processed in the same way as those inside the region.

The segmented image can be considered as a mask, in which one or more regions are defined by those pixels set to either a null value or a user-specified value. A segmentation mask can often be complemented as illustrated in Figure 10.1. This process transfers null 'status' from the excluded pixels to the included pixels and sets the excluded pixels to the values in a specified image channel.

Figure 10.1 Complementing an image mask

Pixels that have been excluded from an image (shown in Figure 10.1a) can be complemented to their original value (Figure 10.1b).



Source: Harrison and Jupp (1990) Figure 57

10.2 Spectral Themes

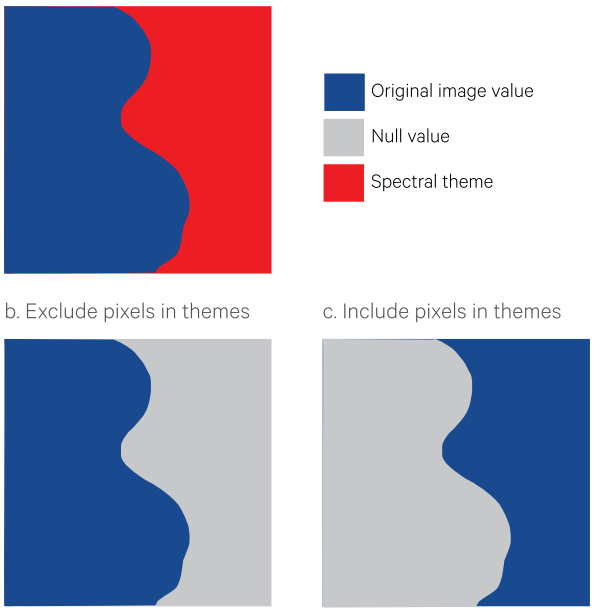
The development of spectral themes to represent particular image features is discussed in Sections 9.1 and 9.2. Spectral segmentation represents a simple form of feature classification where one or more themes can be used to define a feature for the purposes of including it in or excluding it from an image. The themes may define those features that are to be included or excluded from the image, depending on which are easier to develop.

Image segmentation is often an intermediate step in image processing to reduce data volume so it is usually preferable to leave some unwanted pixels in the image rather than remove some required ones. Spectral segmentation may be best effected using particular image channels to define a feature. A mask defined using one set of channels may be transferred to another (registered) set of channels using image pasting as defined in Volume 2D—Section 1.2.

Spatial and spectral segmentation themes can also be combined. For example, spectral classes defined for a deep water feature may also detect land pixels of shadows in steep terrain. In this case a spatial mask can be used to remove the land pixels, often without needing to be accurately located.

Most image processing systems allow images to be segmented spectrally by setting selected pixel values to the null value in all channels of the output image (see Figure 10.2). Typically, pixels that are either included in, or excluded from, the defined themes can be replaced with the values of unmasked pixels. In addition, the complement of the digitised image can often be computed.

Figure 10.2 Including and excluding regions in spectral segmentation



Source: Harrison and Jupp (1990) Figure 58

10.3 Ancillary Data Boundaries

When the boundaries required for image segmentation are defined by some ancillary data source, such as a map, the polygonal region boundaries may be digitised and registered with the image for segmentation purposes. The process required for registration of map and image data is introduced in Section 7 and detailed in Volume 2B. Once a registration model exists, the digitised boundary can be converted to image coordinates and incorporated as an ancillary data channel in the image.

Typically, either one selected channel, or all channels, of an image are processed. Many image processing systems offer the choice of:

- including, excluding or separating from each polygonal region those pixels that constitute its boundary. If the boundary string pixels are to be separated, a unique value can often be specified for them in the output image;
- treating areas of overlap between regions in different ways, such as retaining either the first or last region value, or setting them to ('flagging as') a particular value; and/or
- 'cleaning' boundaries prior to segmentation by removing any closed regions smaller than a specified size.

Pixels in each region may be assigned values in a number of ways as illustrated in Figure 10.3. This assumes that each region has a value associated with it in the vector file, which can simply be assigned to the region pixels to create an ancillary data channel (see Figure 10.3a). If the pixels in the regions were copied from the input images into a new null image, the resulting channels would only include non-null values inside the polygonal regions (see Figure 10.3b). Alternatively, to exclude the polygonal regions from the output image, only those pixels would be set to the null value (see Figure 10.3c).

Figure 10.3 Image segmentation with ancillary data boundaries

a.Set region pixels to another value

b. Copy region pixels from image

c. Set region pixels to null value

Source: Harrison and Jupp (1990) Figure 59

Regions can only be defined by closed boundary vector data. Where ancillary data are in the form of (non-closed) lines or (isolated) points, image pixels on the lines or at the points can often be assigned to a particular value. Where point or line data are sufficiently dense, filter interpolation may be used to fill in any missing pixel values and create a continuous image channel (see Section 10.5 or Volume 2C). Typically, pixels that coincide with multiple points and/or line locations in a vector file can only be set to the first or last point/line value or given a separate value. If this overlapping occurs frequently it may be worth initially registering the ancillary data to a finer image grid to separate the locations as different pixels. In the case of non-nominal data, image blocking could be used to average the values into a larger pixel size (see Section 7.2.1.4).

10.4 Recombining Segments

The complementary operation to image segmentation allows 'null' regions to be transferred between images for the purposes of segmentation or recomposition ('cutting and pasting'). This process naturally assumes that both input images are registered and have the same size. The null mask may be defined by one or more channels in the input image(s) and will only be transferred to the same channel(s) in the output image. Such processes are further discussed in Volume 2D—Section 1.2.

10.5 Cleaning Boundaries

The boundaries of segmented regions may be 'cleaned' using various filtering operations (see Volume 2C). In particular, Section 9 in Volume 2C describes the process of interpolating selected 'missing' values. In conjunction with thickly spaced, but discontinuous, ancillary data points or lines, this operation can create a continuous data channel.

Alternatively, isolated pixels of a selected value (usually equal to the null value by default) may be iteratively removed using a modal filter option. As mentioned in Section 10.3, image processing systems often allow ancillary data boundaries to be 'cleaned' before merging with image data.

Excursus 10.1—Geographic Object-Based Image Analysis

Source: Guy Byrne, Geoscience Australia

Geographic Object-Based Image Analysis (GEOBIA) represents a synthesis of the approaches to land cover mapping used by traditional visual interpretation of imagery and statistically-driven, pixel-based image analysis. A class or object can be described based on a hierarchy of relationships to other classes within the scene, and this hierarchy can be multidimensional. Hay and Castilla (2008) and Blaschke *et al.* (2014) suggest that GEOBIA's ability to define an image model based on the explicit specification of its member elements and their interrelationships (or 'ontology') constitutes a paradigm shift in digital image analysis.

In addition to standard spectral classification descriptors, some discriminates that can be used to describe the relationship between objects and classes within the Definiens GEOBIA Software include:

- geometry (extent, shape);
- position (location, within specified region or not);
- texture; and
- relations to neighbour objects, sub-objects and super-objects (existence of, number of, border to, relative area of, distance from, overlap, aggregration indices).

Using visual interpretation, a landscape is mapped into meaningful units based on the analyst's *a priori* understanding of the scene. That understanding is fundamentally based on the topological relationships between the visible elements within the scene. More recently, pixel-based image analysis uses a variety of statistical and machine learning engines to partition an image into a number of spectrally distinct classes, which are then labelled according to their ontological characteristics.

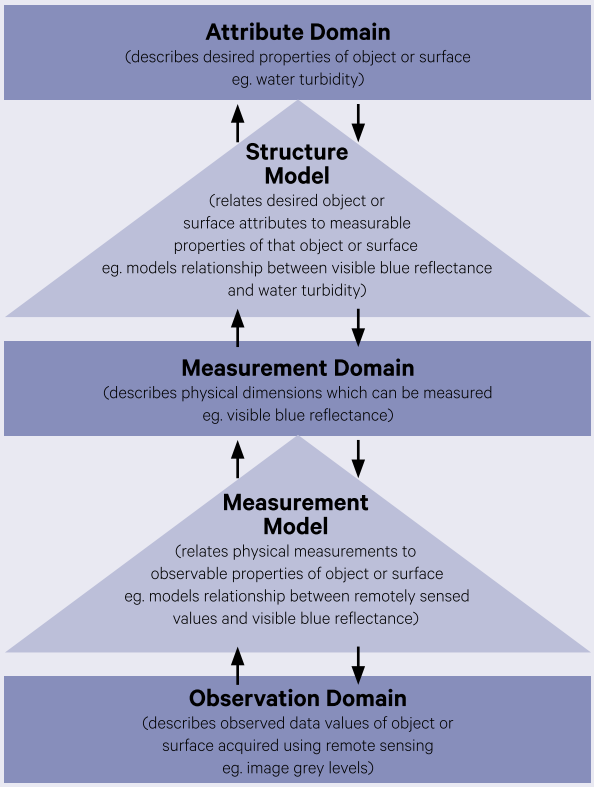
Figure 10.4 shows a conceptual model of the relationship between the observation and attribute domains relevant to EO image interpretation (for a full discussion of this model, please refer to Volume 1A—Section 1). The interplay between the scales of objects, or features, in the landscape and the resolution of the EO image fundamentally defines the sorts of question an interpreter can ask of the data. GEOBIA is most effective when the image pixel size is smaller than the image features to be identified. This scaling relationship between pixels and features has been described as the H-resolution sensor scene model (Strahler *et al.*, 1986) and is depicted in Figure 2.12c (see also Volume 1B—Section 2.2).

To be useful, segments must be measurable, substantial, accessible, differentiable and accountable.
(Philip Kotler)

Using GEOBIA, a classification model is iteratively built through a hierarchy of binary classifications. A feature of interest is 'assigned to a class' by identifying single, or combinations of, attributes based on its spectral, spatial and geometrical properties and ontological relationships. A simplified conceptual model of a typical GEOBIA workflow is shown in Figure 10.6.

Figure 10.4 Conceptual model of domains relevant to EO image interpretation

Interpreting imagery from EO sensors is an indirect process, whereby a measurement model is used to transform EO observations to measurements of some measurable property, and a structure model is used to relate those measurements to application-specific attributes.



Adapted from: Harrison and Jupp (1989) Figure 33

Geographic Object-Based Image Analysis (GEOBIA) is a sub-discipline of Geographic Information Science (GIScience) devoted to developing automated methods to partition remote sensing imagery into meaningful image-objects, and assessing their characteristics through spatial, spectral and temporal scales, so as to generate new geographic information in GIS-ready format. (Hay and Castilla, 2008)

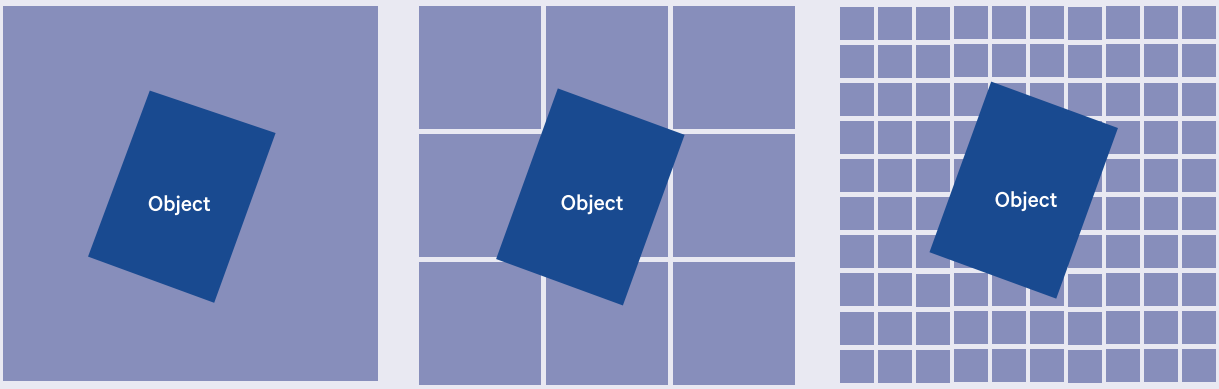
Figure 10.5 Image sampling resolution relative to imaged objects

Depending on image spatial resolution, a single object can be imaged by a single pixel, several pixels or many pixels:

a. When the object being imaged is smaller than an image pixel, then sub-pixel techniques are most appropriate;

b. When the size of the object being imaged is similar to the pixel size, then pixel-by-pixel methods are relevant; or

c. When the object being imaged is much larger than image pixels, object-based techniques are most appropriate.



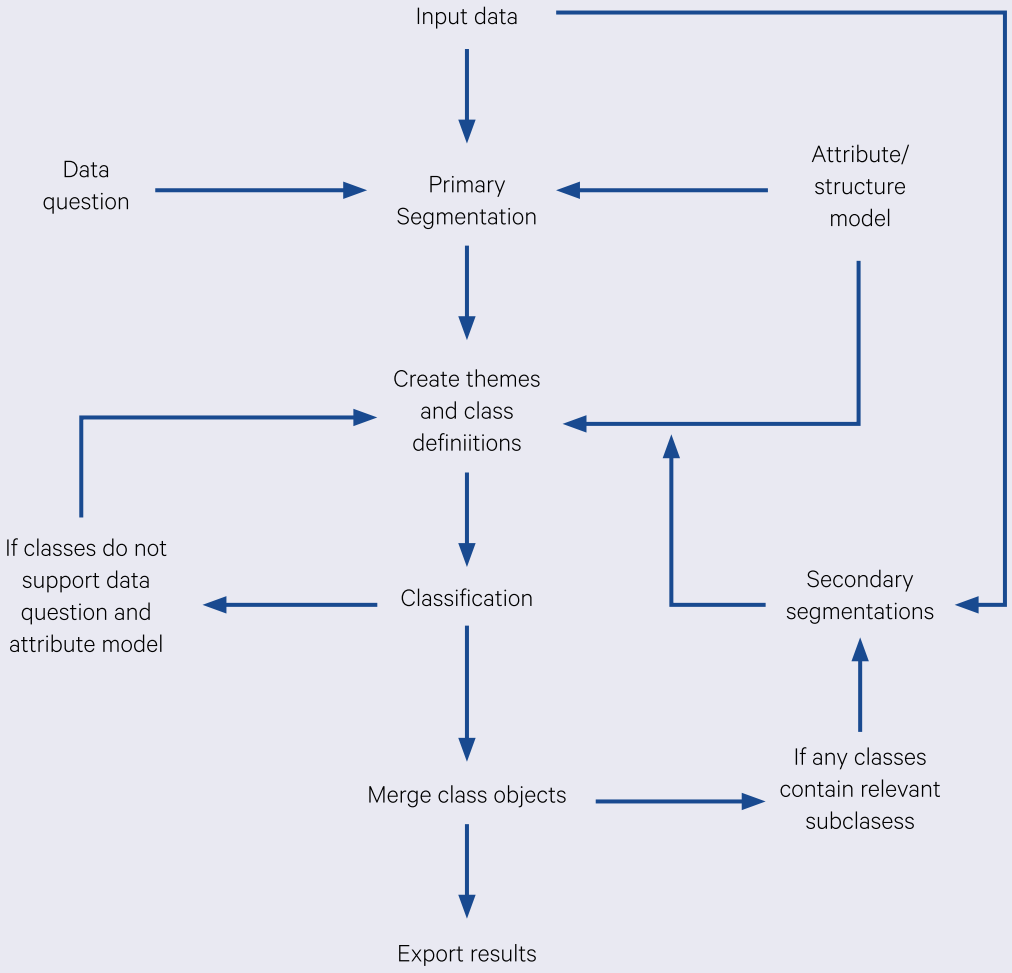
Single 30 m pixel

Nine 10 m pixel

One hundred 1 m pixels

Figure 10.6 Typical GEOBIA classification workflow

This classfiication workflow is similar to the hybrid classification model introduced in Section 9 (see Figure 9.3). In this case, the input data can include EO image bands, GIS shape files and raster surfaces.



The GEOBIA processing model is particularly useful for mapping features that are spectrally similar. For example, vegetation growing at the margins of forests is often spectrally confused with riparian vegetation. However, by including consideration of the shape, size and location of these two types of vegetation, they can be more easily distinguished.

Using a hypothetical example image of Jervis Bay, NSW (see Figure 10.7), the GEOBIA processing sequence to separate spectrally similar features would be:

- create an initial set of objects using a primary segmentation (see Figure 10.8);
- use a NIR spectral theme to partition the objects into groups that represent land and water then merge contiguous objects within these groups (see Figure 10.9);
- to differentiate specific vegetation types, define a secondary segmentation using criteria most appropriate to the land group (see Figure 10.10); then
- progressively classify the land group into a series of sub-classes until all the spectrally-distinct and relevant scene model covers are assigned.

Figure 10.7 Example image

180-day geomedian product over Jervis Bay, derived from Landsat-8 imagery acquired between 1 January 2016 and 28 June 2016, displayed using bands 5,4,3 as RGB.

Figure 10.8 Primary segmentation

This initial segmentation created 110,850 objects for land and 17,469 objects for water within the image shown in Figure 10.7. A subset of this image is shown here for clarity.

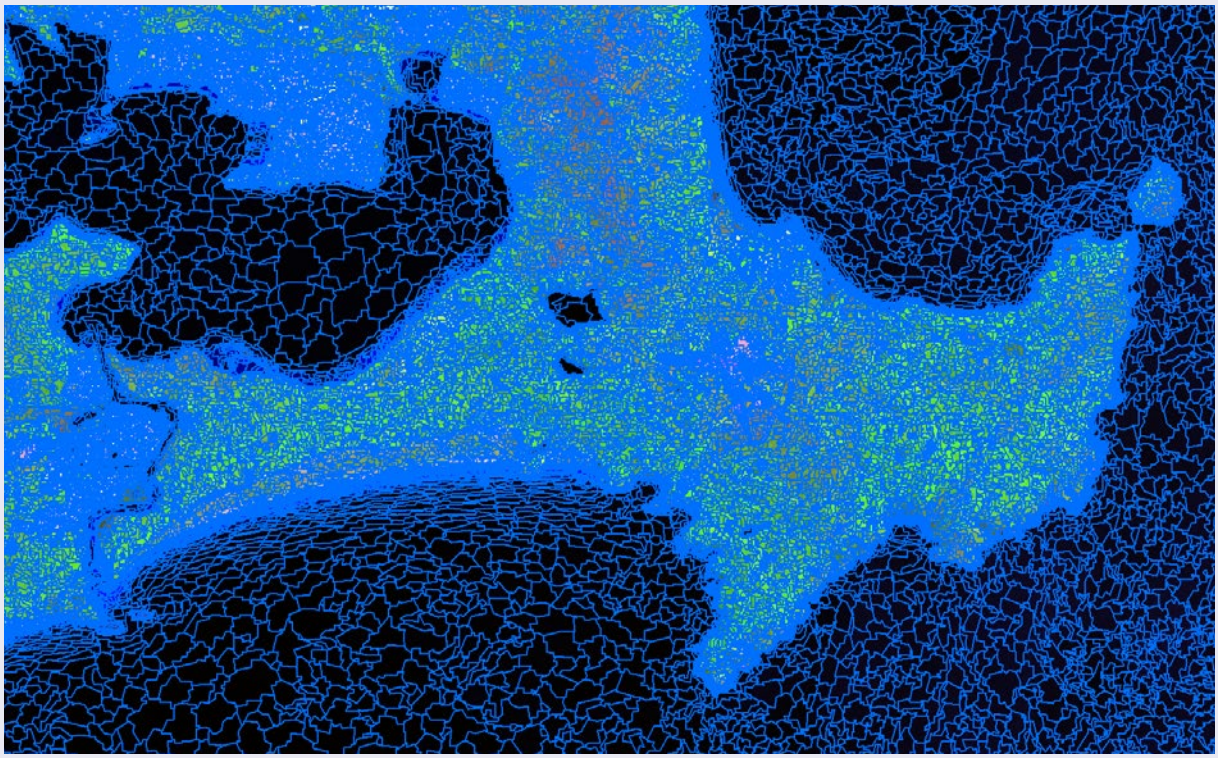


Figure 10.9 Classification of land and water

The segments defined by the primary segmentation stage are merged into two distinct groups that separate land from water. At this stage, the land group (grey) comprises 27 contiguous objects and the water group (aqua) comprises 260 contiguous objects.

Figure 10.10 Secondary segmentation

A secondary segmentation process uses a set of attributes (such as appropriate spectral bands and scaling parameters) that are relevant to the land group to define specific classes. Similarly, to classify the water group, the most appropriate attributes would be used. In this case, 49,563 objects were created in the land group and 1,127 objects were created in the water group. Individual classes can then be refined using new criteria until their mappings are deemed satisfactory. Again a subset is shown for clarity.

In the case of the spectrally-similar forest margin and riparian vegetation classes, the topological landscape relationships unique to each vegetation group need to be understood and then described using sets of rules that 'pull' the mixed spectral class apart. For example, since forests often grade into pasture, an adjacency metric may be a useful class descriptor. In other locations proximity to water or streamlines could be used. As an intermediate stage, a number of smaller sub-classes may be required to fully describe the geospatial relationships of the riparian and forest boundary classes. Such classes could then be merged to define the final set of classes prior to validation.

GEOBIA has been used successfully for many applications including mapping mangroves (Kamal *et al.*, 2015), seagrass (Roelfsema *et al.*, 2014), and coral reefs (Phinn *et al.*, 2012), and monitoring mine site rehabilitation (Bao *et al.*, 2014).

10.6 Further Information

Geographic Object-Based Image Analysis

Blaschke *et al.* (2014)

Johansen *et al.* (2010)

10.7 References

Bao, N., Lechner, A. M., Johansen, K., and Ye, B. (2014). Object-based classification of semi-arid vegetation to support mine rehabilitation and monitoring. *Journal of Applied Remote Sensing*, 8(1), pp. 083564. doi:*http://dx.doi.org/10.1117/1.JRS.8.083564*.

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R. Q., van der Meer, F., van der Werff, H. M. A., van Coillie, F., and Tiede, D. (2014). Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS *Journal of Photogrammetry and Remote Sensing*, 87, pp. 180–191. doi:*http://dx.doi.org/10.1016/j.isprsjprs.2013.09.014*.

Harrison, B. A., and Jupp, D. L. B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual* (256 pages). CSIRO Australia, Melbourne.

Hay, G. J., and Castilla, G. (2008). Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. Chapter 1.4 in '*Object-Based Image Analysis*. Lecture Notes in Geoinformation and Cartography.' (Eds: Blaschke T., Lang S., and H. G.J.). Springer, Berlin, Heidelberg.

Johansen, K., Bartolo, R., and Phinn, S. (2010). Special Feature—Geographic Object-Based Image Analysis. J. *Spatial Science*, 55(1), pp. 3–7. doi:*http://dx.doi.org/10.1080/14498596.2010.494653*.

Kamal, M., Phinn, S., and Johansen, K. (2015). Object-based approach for multi-scale mangrove composition mapping using multi-resolution image datasets. *Remote Sensing*, 7(4), pp. 4753–4783.

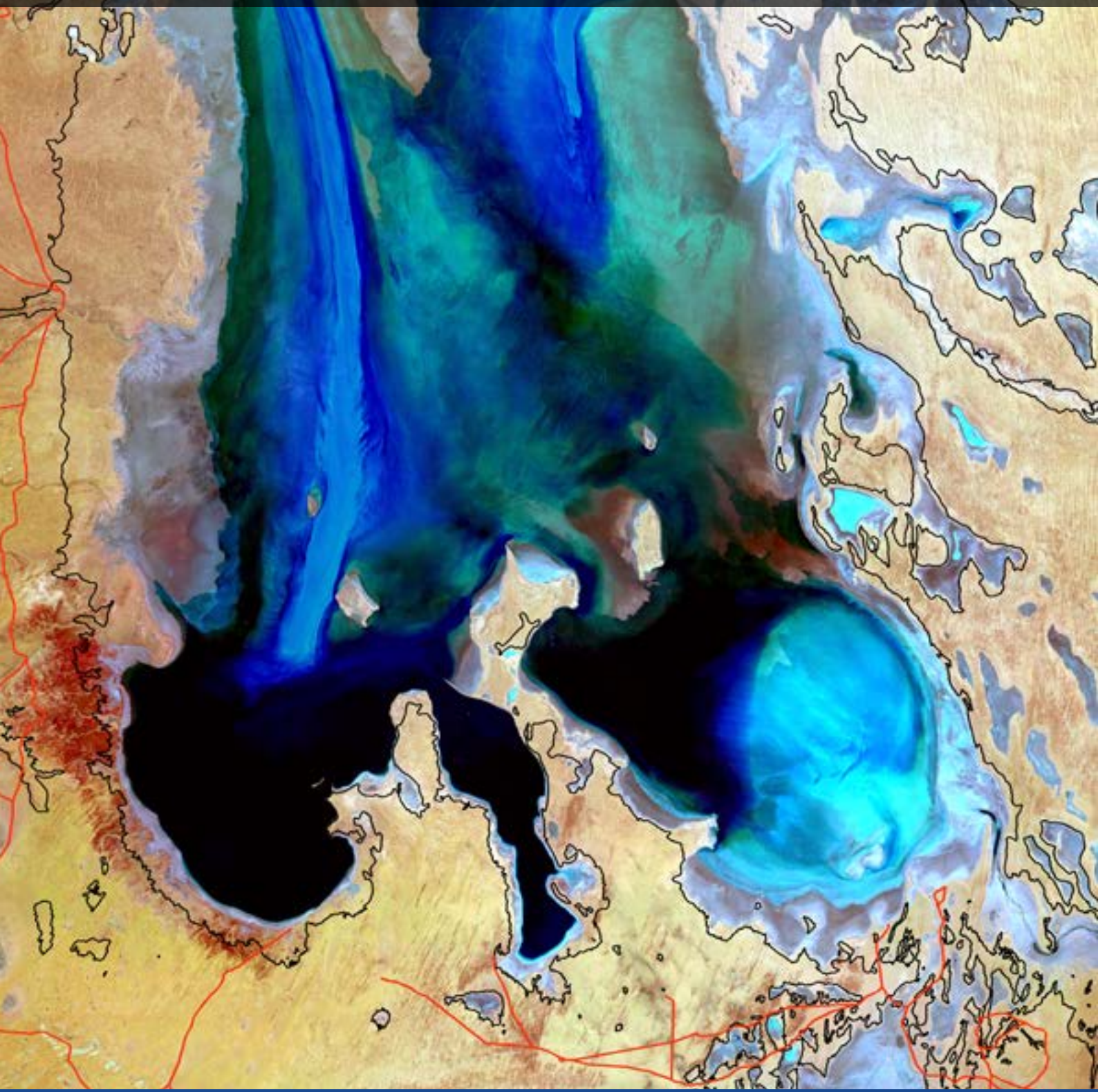
Phinn, S., Roelfsema, C., and Mumby, P. J. (2012). Multi-scale Object Based Image Analysis for Mapping Geomorphic and Ecological Zones on Coral Reefs. J. *Remote Sensing*, 33(12), pp. 3768–3797. doi:*http://dx.doi.org/10.1080/01431161.2011.633122*.

Roelfsema, C., Lyons, M., Kovacs, E. M., Maxwell, P., Saunders, M. I., Samper-Villarreal, J., and Phinn, S. (2014). Multi-Temporal Mapping of Seagrass Cover, Species and Biomass: A Semi-Automated Object Based Image Analysis Approach. *Remote Sensing of Environment*, 150, pp. 172–187. doi:*http://dx.doi.org/10.1016/j.rse.2014.05.001*.

Strahler, A. H., Woodcock, C. E., and Smith, J. A. (1986). On the nature of models in remote sensing. *Remote Sensing of Environment*, 20(2), pp. 121–139. doi:*http://dx.doi.org/10.1016/0034-4257(86)90018-0*.

Earth Observation:
Data, Processing and Applications

Volume 2B: Processing—Image Rectification



This publication is available in PDF format at *eopublications.au*.
We welcome your comments regarding the readability and usefulness of this report.
To provide feedback, please email *feedback@eopublications.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
978-0-6482278-0-9

Copyright:
All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, save for content supplied by third parties, and logos. Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. The full licence terms are available from *https://creativecommons.org/licenses/by/4.0/legalcode*. A summary of the licence terms is available from *https://creativecommons.org/licenses/by/4.0/*.

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Recommended Citation for Volume 2B:
CRCSI (2018) *Earth Observation: Data, Processing and Applications. Volume 2B: Processing—Image Rectification*. (Eds. Harrison, B.A., Jupp, D.L.B., Lewis, M.M., Sparks, T., Mueller, N., Byrne, G.) CRCSI, Melbourne.

Latest Revision:
October 2025

Background image on previous page: Landsat-5 image of Lake Eyre, South Australia, filling on 9 May, 2009, displayed using bands 5, 4, 2 as RGB and overlaid with contour lines (black) and roads (red).
Source: Norman Mueller and Erin Teifer, Geoscience Australia

Acknowledgements

Production of this series of texts would not have been possible without the financial support of CSIRO, CRCSI, GA and BNHCRC, input from members of the editorial panels and direction from members of the various advisory panels.

Volumes 1 and 2 of this series are based on text originally published in Harrison and Jupp (1989, 1990, 1992 and 1993)[1]. Many illustrations and some text from these publications have been reproduced with permission from CSIRO.

Other contributors are gratefully acknowledged:

- reviewers: Daniel Jaksa (Section 1), Craig Smith, Wang Lan-Wei (Section 4);
- illustrations: Norman Mueller kindly supplied most of the lovely images; other contributors of graphical material include: Tony Sparks, Megan Lewis, Guy Byrne, James Storey, Chris Evenden, Raj Menon, Geoscience Australia, ICSM, Spookfish, Nearmap, Dunn's Aerial Photography, Inc (Wellington, Florida, USA), Lacy Larson (Palm Beach International Airport);
- excursus: Karen Joyce, Peter Scarth, Rohan Fernando, Wang Lan-Wei, Tony Sparks, Norman Mueller, Guy Byrne;
- Daniel Rawson (Accessible Publication & Template Design) for layout and formatting; and
- Carl Davies (CMDphotographics) for selected graphical illustrations.

We thank those owners of copyrighted illustrative material for permission to reproduce their work. Credits for individual illustrations are provided below the relevant graphic.

All volumes in this series are covered by the copyright provisions of CC BY 4.0 AU.

[1] Harrison, B.A., and Jupp, D.L.B. (1989) *Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual.* CSIRO, Melbourne. 156pp.
Harrison, B.A., and Jupp, D.L.B. (1990) *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.
Harrison, B.A., and Jupp, D.L.B. (1992) *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual.* MPA, Melbourne.
Harrison, B.A., and Jupp, D.L.B. (1993) *Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual.* MPA, Melbourne.

Table of Contents

Volume 2B: Processing—Image Rectification a

Acknowledgements i

Background 1

1 Cartography **3**

1.1 Coordinate Systems 4

1.1.1 Cartesian 4

1.1.2 Polar 5

1.1.3 Origins 6

1.2 Fundamentals of mapping 7

1.2.1 Geodesy 7

1.2.2 Projections 9

1.2.3 Scale 20

1.3 Map Coordinate Systems 21

1.3.1 Geographical coordinates 21

1.3.2 Universal Transverse Mercator 22

1.4 Mapping standards 25

1.4.1 Australia 25

1.4.2 Global 25

1.5 Further Information 26

1.6 References 26

2 Image Geometry **29**

2.1 Pixel Size 31

2.2 Sampling and Recording Rate 33

2.3 Viewing Perspective 34

2.4 Platform Stability 38

2.5 Image Grid Origin 43

2.6 Image Registration 43

2.7 Further Information 49

2.8 References 49

Rectification 51

3 Transformation Models **53**

3.1 Nominal Models 53

3.1.1 Map 53

3.1.2 Image 55

3.2 Arbitrary Mathematical Models 57

3.2.1 Affine 59

3.2.2 Bilinear 64

3.2.3 Quadratic 64
3.2.4 Cubic 65
3.3 Multiple Models 66
3.4 Further Information 70
3.5 References 70

4 Control Point Modelling **71**
4.1 Selecting Control Points 73
4.1.1 Selection criteria 73
4.1.2 Selection process 76
4.2 Fitting Arbitrary Models 80
4.2.1 Method of least squares and polynomial models 86
4.2.2 Predictive error 87
4.2.3 Detecting outlier points 88
4.2.4 Selecting the optimum model 89
4.2.5 Checking the extent of control 90
4.3 Further Information 92
4.4 References 92

Resampling and Registration 93

5 Resampling Methods **95**
5.1 Along-line Resampling 102
5.2 Nearest Neighbour Resampling 103
5.3 Bilinear Interpolation 108
5.4 Cubic Convolution 109
5.5 Further Information 110
5.6 References 110

6 Registration **111**
6.1 Registering EO Images with Other Spatial Datasets 118
6.2 Mosaicking EO Images 122
6.3 Locating Sample Sites in EO imagery 126
6.4 Further Information 132
6.5 References 132

List of Figures

Figure 1.1 Earth Observation scene geometry 3

Figure 1.2 Two-dimensional Cartesian coordinate system 4

Figure 1.3 ECEF rectangular geocentric coordinate system 5

Figure 1.4 Polar coordinate system 5

Figure 1.5 Four possible origins for Cartesian coordinate systems 6

Figure 1.6 AUSGeoid2020 7

Figure 1.7 Modelling the shape of the Earth 8

Figure 1.8 Definition of Earth shape parameters 8

Figure 1.9 Conceptual basis of a map projection 9

Figure 1.10 Major map projection surfaces 10

Figure 1.11 Secant contact between projection surface and sphere 11

Figure 1.12 Aspects for cylindrical projection 12

Figure 1.13 Graticules 13

Figure 1.14 Australia Report Map 13

Figure 1.15 Perspectives for azimuthal projections 14

Figure 1.16 Azimuthal maps of South Pole 14

Figure 1.17 Equirectangular or Plate Carrée 'projection' 16

Figure 1.18 Mercator projection 17

Figure 1.19 Transverse Mercator projection 18

Figure 1.20 Robinson projection 19

Figure 1.21 Albers Equal Area projection 19

Figure 1.22 Lambert Conformal Conic projection 20

Figure 1.23 Measurement of geodetic and geocentric latitude 22

Figure 1.24 Measurement of Longitude 22

Figure 1.25 Universal Transverse Mercator grid 23

Figure 1.26 Map Grid of Australia 23

Figure 1.27 Cross section of MGA zones 24

Figure 2.1 Image rectification 30

Figure 2.2 Optical versus geometric pixel sizes 31

Figure 2.3 Pixel size calculations 32

Figure 2.4 Effect of wide scan angle 32

Figure 2.5 Sensor delay offset for Landsat MSS 33

Figure 2.6 Panoramic view distortion and Earth curvature 34
Figure 2.7 Spatial distortions in continental scale imagery 34
Figure 2.8 Earth rotation skew 36
Figure 2.9 Daily acquisitions for MODIS/Aqua and MODIS/Terra 37
Figure 2.10 Effect of camera orientation in aerial photography 38
Figure 2.11 Relief effects on aerial imagery 38
Figure 2.12 Platform attitude variations 39
Figure 2.13 Platform altitude variation 40
Figure 2.14 Platform velocity variation 41
Figure 2.15 Original image data showing individual flight lines 42
Figure 2.16 Geometric correction of CASI imagery 42
Figure 2.17 Corrected image mosaic 42
Figure 2.18 Image coordinate system 43
Figure 2.19 Actionable data 45
Figure 2.20 Multi-angle viewing 46
Figure 2.21 Camera system on aerial platform 47
Figure 2.22 Spookfish high resolution image of Federation Square, Melbourne 48
Figure 3.1 Local Transverse Mercator projection 54
Figure 3.2 Hotine Oblique Mercator projection 55
Figure 3.3 Determining pixel size adjustment factor 56
Figure 3.4 Geometric effects of polynomial models 58
Figure 3.5 Shift operation 60
Figure 3.6 Skewing operation in X direction 61
Figure 3.7 Rotation operation 61
Figure 3.8 Origin adjustment for image rotation 62
Figure 3.9 Rescaling operation 63
Figure 3.10 Reflection operation 63
Figure 3.11 Bi-directional skew 64
Figure 3.12 Quadratic transformation 64
Figure 3.13 Cubic transformation 65
Figure 3.14 Generalisation of three-stage registration model 66
Figure 3.15 Three-stage registration model 67
Figure 3.16 Locations of control points 68
Figure 3.17 Three-stage conversion from AVHRR image to map coordinates 69
Figure 4.1 Control point modelling 71

Figure 4.2 Control point modelling sequence 72
Figure 4.3 Examples of suitable ground features for control point survey 74
Figure 4.4 Precise placement of control point at road intersection 74
Figure 4.5 Control point distribution 75
Figure 4.6 Cotter Dam wall on 24 December 2012 77
Figure 4.7 GCP used for accuracy assessment. 79
Figure 4.8 Orthorectified image 79
Figure 4.9 Model stability and complexity 80
Figure 4.10 Landsat MSS imagery 81
Figure 4.11 GCP locations shown on resampled, greyscale images 82
Figure 4.12 Resampled images 85
Figure 4.13 Predictive error 87
Figure 4.14 Q-Q plot 89
Figure 4.15 Predictive variance image 91
Figure 5.1 The resampling problem 95
Figure 5.2 Interpolation example 96
Figure 5.3 Interpolation functions for resampling methods 97
Figure 5.4 Resampling methods 98
Figure 5.5 Effects of resampling methods 100
Figure 5.6 Along-line resampling 102
Figure 5.7 Nearest neighbour resampling 103
Figure 5.8 Nearest neighour interpolation 104
Figure 5.9 Potential artefacts from nearest neighbour resampling 104
Figure 5.10 Picking and binning approach 105
Figure 5.11 Output image or base map 106
Figure 5.12 Example grid 106
Figure 5.13 Sub-pixel behaviour 108
Figure 5.14 Bilinear interpolation 108
Figure 5.15 Cubic convolution interpolation 109
Figure 6.1 GCP work packages 113
Figure 6.2 Example survey site 114
Figure 6.3 Surveyed GCP for AGRI 114
Figure 6.4 Accuracy assessment 115
Figure 6.5 AGRI national mosaic 116
Figure 6.6 Landsat-8 estimate of current Global Land Survey (GLS) horizontal error 117

Figure 6.7 Georegistering two sources of EO data 118
Figure 6.8 Overlaying map information on EO imagery 119
Figure 6.9 Datum differences 120
Figure 6.10 Projection differences 121
Figure 6.11 Uncorrected image scenes 123
Figure 6.12 Scene mosaic 124
Figure 6.13 Surface reflectance mosaic 125
Figure 6.14 Pixel mosaic 125
Figure 6.15 Lake George validation site 128
Figure 6.16 Spectral sampling protocol 128
Figure 6.17 Lake George field sampling on 3 May 2018 129
Figure 6.18 Location of Lake George field site 130
Figure 6.19 Landsat-8 OLI field site pixels 130
Figure 6.20 Lake George spectra transect lines overlaid on Landsat-8 OLI image grid 131
Figure 6.21 Field spectra versus Landat-8 OLI spectra on 3 May 2018 131
Figure 6.22 Comparison of field and satellite reflectances for each band 131

List of Tables

Table 1.1 Comparison of common map projections 15

Table 1.2 Map scales 21

Table 2.1 Pixel sizes for satellite data sources 33

Table 4.1 Verification errors in selected GCP 79

Table 4.2 15 GCP used to rectify path 89 scene 82

Table 4.3 25 GCP used to rectify path 90 scene 83

Table 4.4 Model parameters 84

Table 4.5 Residuals after removing five GCP for path 90 scene 84

Table 4.6 Error associated with polynomial models fitted to 32 GCP 91

Table 4.7 Error associated with polynomial models fitted to 150 GCP 92

Table 5.1 Resampling methods 99

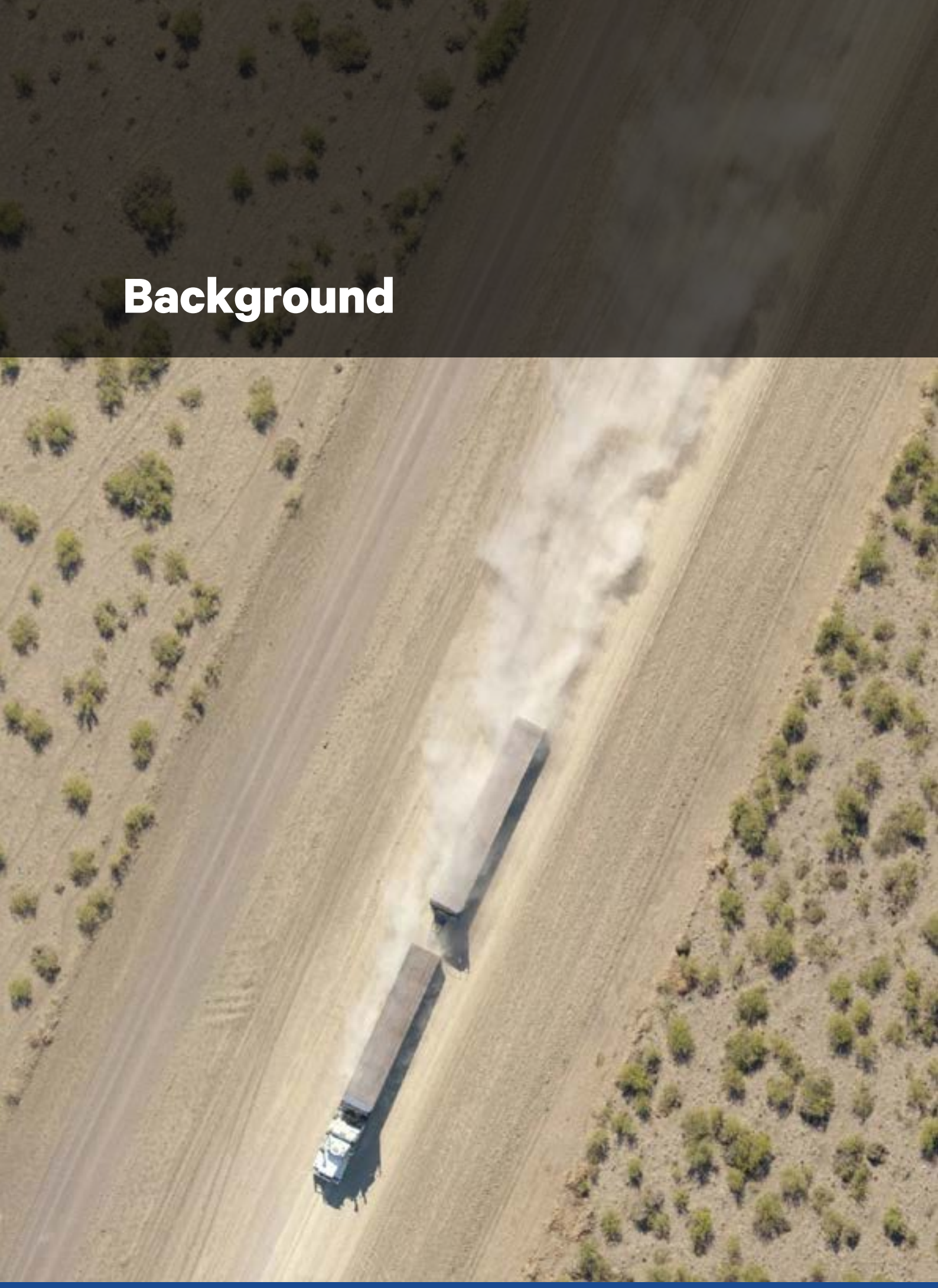
Table 6.1 ALOS/PRISM characteristics 112

Table 6.2 Map scale and image resolution 118

List of Excurses

Excursus 1.1—Common Map Projections 15
Excursus 2.1—Generation of CASI-2 Image Mosaic 42
Excursus 2.2—Geospatial Imagery Redefined 44
Excursus 3.1—Estimating Image Pixel Width and Depth 56
Excursus 3.2—Using Multiple Models for Image Rectification 68
Excursus 4.1—Map Accuracy Standards 73
Excursus 4.2—Control Point Selection in AGRI project 74
Excursus 4.3—Registration of Pleiades-1 Satellite Imagery 77
Excursus 4.4—Image Rectification Example 81
Excursus 5.1—'Picking and Binning' Approach 105
Excursus 5.2—Basis of the 'Picking and Binning' Approach 107
Excursus 6.1—The Australian Geographic Reference Image 111
Excursus 6.2—Overlaying Map and Image Data 121
Excursus 6.3—Multi-scene Image Mosaic 123
Excursus 6.4—Validation of Surface Reflectance Data 127

Background



A problem common to all remote sensing is the need to accurately locate observations to the ground, a process called 'geo-referencing' or, because satellite observations often form images, 'image rectification' and in some cases 'orthorectification'. To ensure that observations taken at different times and from different satellites and instruments can be compared, accurate and consistent geo-referencing is essential. Monitoring is therefore only possible with consistent and accurate geo-referencing.
(Lewis et al., 2011)

Image Rectification of Earth Observation (EO) data relies on an understanding of the shape of the Earth's surface (and its geometric representation in various forms of maps), and the diverse factors that result in the geometry of EO imagery. Before introducing specific methods for image rectification, the following sections review the underlying principles of:

- cartography—the art and science of making maps (see Section 1); and
- image geometry—factors that determine and distort the shape and spatial interrelationships of EO imagery (see Section 2).

Contents

1 Cartography 3

2 Image Geometry 29

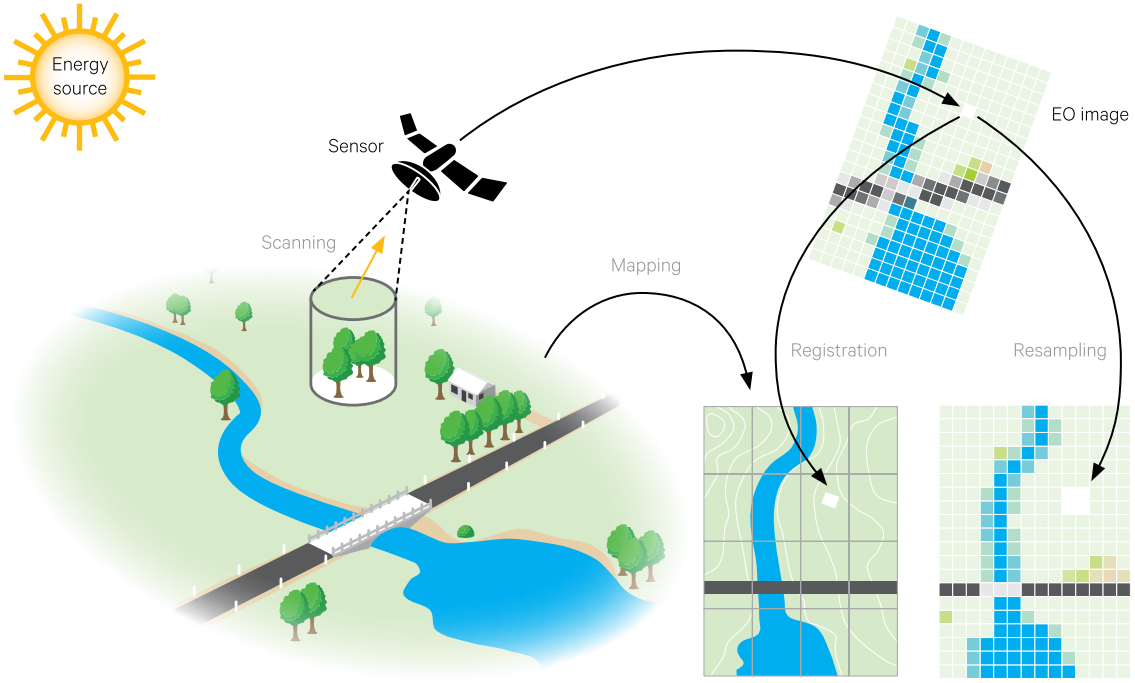
Background image on previous page: High resolution aerial image of a road train near Gurley, NSW, acquired on 25 June 2017. This image has a pixel size of 1 cm. **Source:** © Spookfish

1 Cartography

Cartography is concerned with representing the spatial relationships between features on the Earth's surface. This activity can be considered as involving any form of spatial data acquisition (traditionally limited to surveying and geodesy) as well as the graphical representation of the data in map form. For practical reasons maps are usually depicted on a flat surface.

Figure 1.1 Earth Observation scene geometry

This illustration shows a ground scene being imaged by a remote sensor. A 'volume' on the surface that is observed as one optical pixel is highlighted by the cylinder. The corresponding geometric pixel for this location is shown as a white square in the EO image. While the EO image contains recognisable features from the ground scene, its geometry rarely matches a standard map projection. To rectify the image to map coordinates, the locations of specific features need to be identified on a map. For example, the location of the highlighted optical pixel is shown here as a white square on the map. To allow EO imagery to be digitally or physically registered to a map, it then needs to be resampled to match the map geometry. As shown in this example, the original optical pixel may be represented by multiple geometric pixels in the resampled image.



Source: Harrison and Jupp (1992) Figure 1

Background image: Australian continent and adjacent islands shown on Lambert Conformal Conic projection centred on 140°E and 25°S, with two Standard Parallels at 18°S and 26°S (see Section 1.2.2). **Source:** Chris Evenden, Geoscience Australia

Various projection methods are available to represent three-dimensional information derived from the irregular surface of the Earth as a two-dimensional (or planar) surface—either on hardcopy (such as paper) or in digital form (such as an EO image). However, projection transformations unavoidably misrepresent the directions, distances, areas and/or shapes of features on the original, curved surface. While current cartographic techniques allow maps to be produced to a high level of accuracy, the constraints implicit in any projection mean that a two-dimensional map (in digital or hardcopy format) cannot faithfully represent the actual three-dimensional surface of Earth, but merely offers a 'view' of it.

Earth Observation (EO) imaging devices represent the Earth's surface by creating pixel values that integrate the radiances of various components within a volume on the surface as illustrated in Figure 1.1 (see Volume 1A—Sections 13 to 15 for details of the imaging processes used by EO sensors). Each imaged volume can contain one or more surface features such as trees, bare soil or houses. In this way, an EO image is a two-dimensional view, or map, of the Earth's three-dimensional surface from the perspective of the scanning platform. To locate an image pixel relative to other information about the Earth's surface requires that its 'volume' can be identified as a single point on the surface. For practical reasons, ground identification points are generally located using maps of the Earth's surface. Ideally, each pixel in an image would be associated with the location of its volume on the map. In practice, however, a generalised model between the image and a map is developed which can be applied to each pixel in the image.

Various coordinate systems and map projections can be used to represent information about the Earth's surface. The sub-sections below describe background information relevant to these two aspects of mapping.

1.1 Coordinate Systems

Coordinate systems allow location, topology and metrics to be defined for spatial data. Such systems precisely describe the relative positions of features on the Earth's surface by defining distance and direction algebraically. Complex geometries and geometric relationships can be represented using coordinate systems. Once defined in a mathematical system, any geometry can be modified in a precise and consistent manner.

Two major categories of coordinate systems are used to describe the three-dimensional Earth:

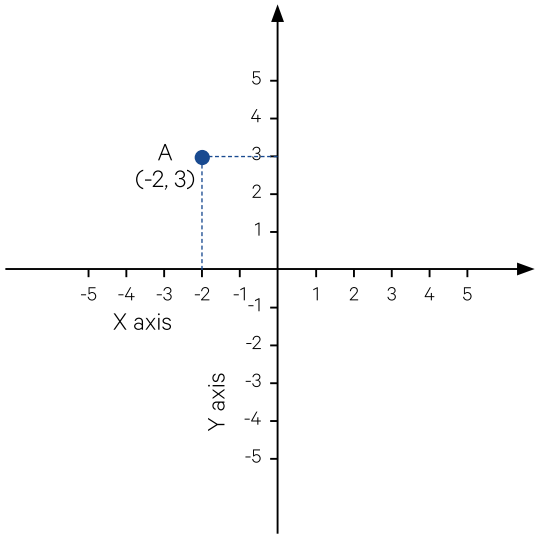
- Cartesian (or rectangular) coordinates (see Section 1.1.1); and
- polar coordinates (see Section 1.1.2).

1.1.1 Cartesian

The standard two-dimensional, Cartesian coordinate system defines two perpendicular axes, usually labelled X and Y, in a planar surface (see Figure 1.2). The crossover point of these axes defines the origin when both X and Y values equal zero. As illustrated in Figure 1.2, values on the X-axis are referenced as being negative to the left of the origin and positive to the right; values on the Y-axis are positive above the origin and negative below. Individual coordinate systems may differ in the scaling or orientation of axes, or position of origin.

Any point in this linear, two-dimensional space, or grid, can be referenced by a pair of numbers or 'coordinates', which indicate its position relative to these two axes. For example, in Figure 1.2, point A is directly above value -2 on the X-axis and directly to the left of value 3 on the Y-axis. This position would then be referenced as '-2, 3'. A three-dimensional coordinate system allows an additional value Z to be associated with each X, Y location. In terms of representing the three-dimensional geometry of a horizontal surface, the third dimension would be a value above or below the surface for each location.

Figure 1.2 Two-dimensional Cartesian coordinate system



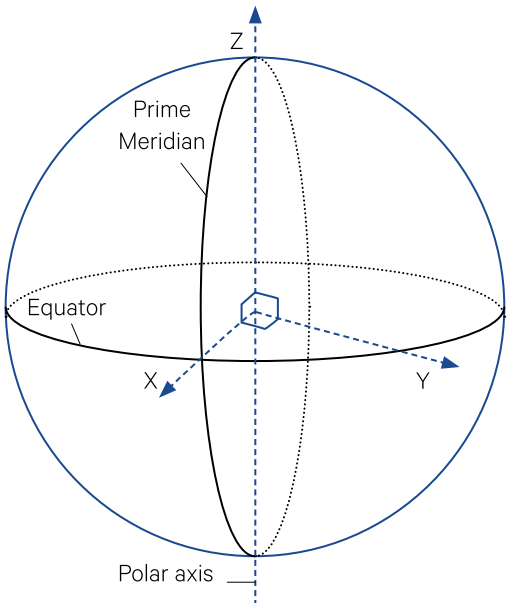
Source: Harrison and Jupp (1992) Figure 2

A three-dimensional coordinate system can also be used to define locations in space. If the origin of the coordinate system is at the centre of a solid, such as the Earth, the three-dimensional coordinates can be used to define locations on the surface. This is the system used by all Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS). Earth-centred, Earth-Fixed (ECEF) Rectangular Coordinates (now more commonly known as geocentric) use a Cartesian coordinate system with:

- the origin at the centre of the ellipsoid;
- the Z axis as the mean axis of Earth rotation (with positive values to North);
- the X axis at the intersection of the equatorial and prime meridian planes; and
- the Y axis is mutually perpendicular to the X and Z axes and defines the equatorial plane (see Figure 1.3).

The Cartesian coordinate system forms the basis for many spatial referencing systems. However, both the way the system is implemented, and the relationship between the implementation and the ground distances and directions, varies with, and between, different map and image data sets.

Figure 1.3 ECEF rectangular geocentric coordinate system



Source: Harrison and Jupp (1992) Figure 3

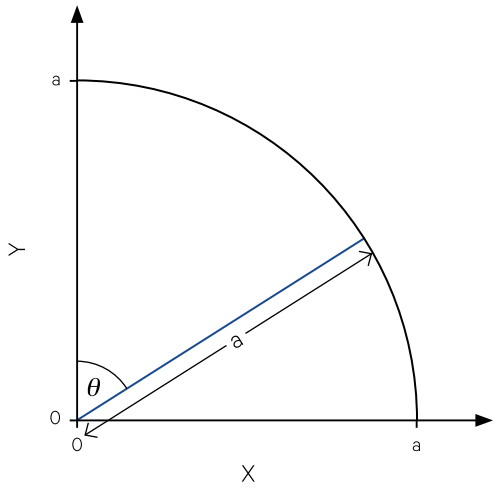
1.1.2 Polar

The location of a point on a plane can also be represented by the straight-line distance from the origin to the point, together with the angle between one axis and the distance line (see Figure 1.4). The angle is typically measured from North in a two-dimensional polar coordinate system. If the distance is constant, as occurs when all points being located lie on the circumference of a circle, only an angular measurement is required to identify a particular position.

A similar approach can be used to describe locations on a three-dimensional surface by defining two distances from an origin and two orthogonal angles. In this case, the first angle indicates a direction to the location when it is projected onto a selected plane, and the second angle is a direction from this reference plane to the location. When the surface is a sphere and the origin is the centre of the sphere, all distances are equal to the radius of the sphere so only the angular measurements are required. These three-dimensional coordinates are known as spherical coordinates.

Figure 1.4 Polar coordinate system

Polar coordinates are specified by their linear distance on a plane from a defined origin and their angular distance on that plane from a defined reference axis.



Source: Harrison and Jupp (1992) Figure 4

This type of three-dimensional angular coordinate system is used to define latitude and longitude locations on the Earth's surface (assuming a constant Earth radius). Latitude measurements effectively project point locations on the spherical surface onto a two-dimensional plane or circle that intersects the poles and Greenwich, England. The angle between the projected point and the equator is then used as latitude with longitude being the angle between this reference plane and the original position. The effect of the non-spherical shape of the Earth on these measurements is discussed in Section 1.3.1.

Maps codify the miracle of existence.
(Nicholas Crane)

1.1.3 Origins

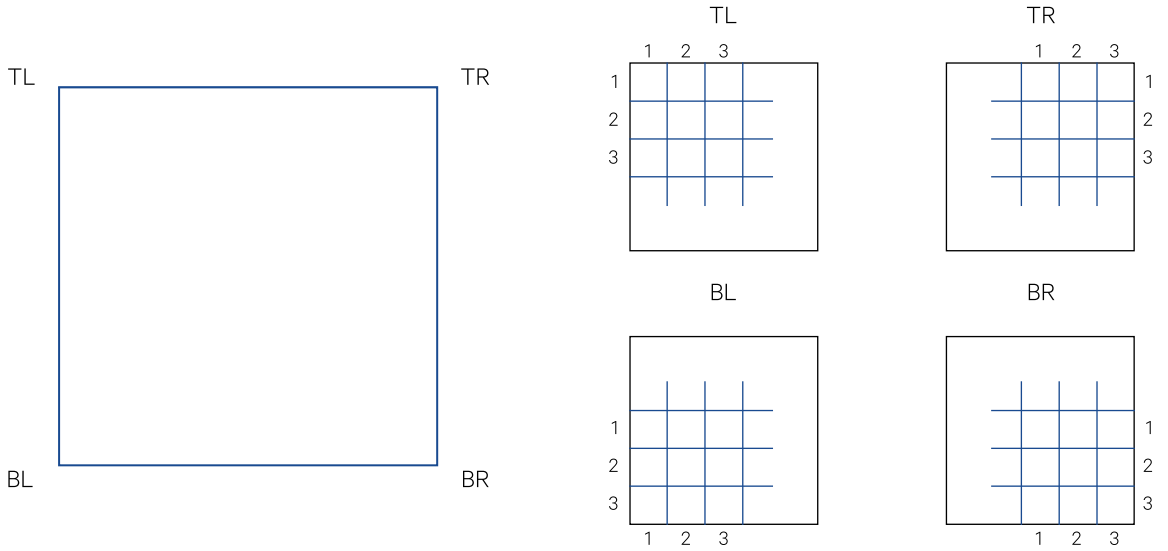
A common difference between coordinate systems is the origin position used for coordinate referencing. Four fundamental grid-addressing schemes can be used as the basis of a planar coordinate system. In image processing these are commonly referred to by the origin positions of:

- top-left (TL);
- top-right (TR);
- bottom-left (BL); and
- bottom-right (BR; see Figure 1.5).

An origin in position BL represents the traditional grid coordinate system and most map systems. For most image data a TL grid applies but TR, BL or BR may also be the 'natural' grid for an image depending on the direction of scanning relative to the target surface. Accepted conventions for spatial referencing in maps and imagery are discussed below.

Figure 1.5 Four possible origins for Cartesian coordinate systems

TL: top left; TR: top right; BL: bottom left; BR bottom right



Source: Harrison and Jupp (1992) Figure 5

1.2 Fundamentals of mapping

Cartography visually represents selected features on the surface of the Earth as a map. As such it is both an art and a science. Key concepts underlying the science of cartography include:

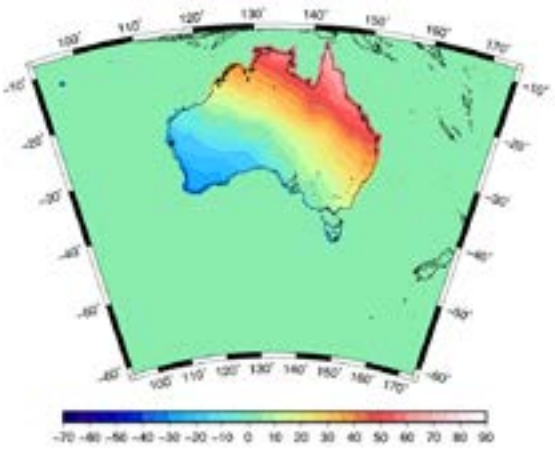
- geodesy—mathematical models of the Earth's shape and size (see Section 1.2.1);
- projections—methods to project from three-dimensions to two-dimensions (see Section 1.2.2); and
- scale—ratio of distances on the map to those it represents on the ground (see Section 1.2.3).

1.2.1 Geodesy

Geodesy is defined as 'the science which deals mathematically with the size and shape of the Earth and the Earth's external gravity field' (ASPRS, 1984). An understanding of the geometric characteristics of the globe is necessary for any accurate Earth surface mapping exercise, especially for large area studies. In EO, the large area coverage of many satellite data sources requires that Earth's shape be considered in the process of image geometric rectification (see Section 3.1).

Figure 1.6 AUSGeoid2020

This gravimetric geoid model for the Australian region was fitted *a posteriori* to the Australian Height Datum for determining heights from Global Navigation Satellite Systems (GNSS). Units are in metres.



Source: Geoscience Australia (2018)

The irregular form of the actual surface of Earth makes it difficult to model the shape mathematically (see Volume 1A—Section 3.3). Position referencing systems require that the location of a point can be defined independently of variations in surface elevation. Mean sea level is generally used as the datum for gravity-related elevation, so a surface equivalent to undisturbed mean sea level over the whole globe would form an ideal base as a level surface for surveying and mapping exercises.

> *For the execution of the voyage to the Indies,*
> *I did not make use of intelligence,*
> *mathematics or maps.*
> *(Christopher Columbus)*

The Earth's topography-free shape is called the 'geoid' and is defined as an equipotential surface due to gravity and rotation, that is, every point on the surface is perpendicular to constant gravitational forces. This surface is ideal for surveying purposes but is too irregular for map projections. Variations in the type and density of crustal and sub-surface materials at different locations on the Earth's surface cause the strength and direction of its gravitational forces to vary (see Volume 1A—Section 8). Accordingly, the geoid is not a regular, mathematical surface, but a smoothly undulating one (see Volume 1A—Excursus 3.1). For example, AUSGeoid2020, a gravimetric geoid that models the variations in gravitational forces across the Australian continent is illustrated in Figure 1.6.

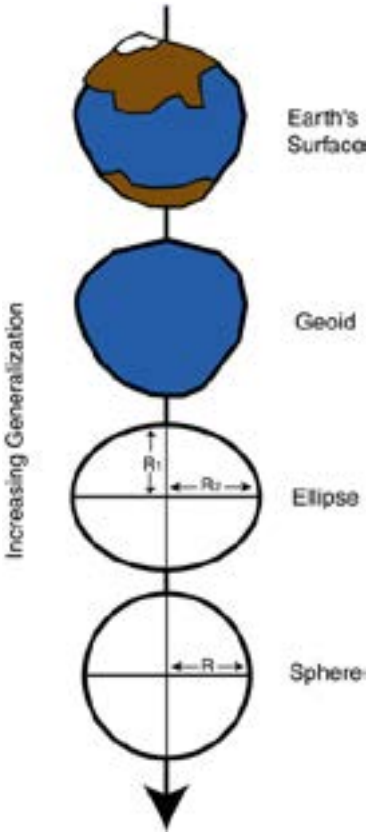
For simplicity in mapping, a mathematical surface called the ellipsoid (or, in older texts, a spheroid) is defined that closely matches the shape of the geoid for a selected portion of the surface (see Figure 1.7). The shape of the ellipsoid is defined by its major and minor axes, which respectively lie on the equatorial and polar planes. By convention, these shape parameters are given as radius values and are referred to as the semi-major axis (*a*) and the semi-minor axis (*b*) as illustrated in Figure 1.8a. The degree of flattening at the poles is indicated by the ratio:

$$f = \frac{a - b}{b}$$

and is usually expressed as 1 / *f*. The constants *a* and *f* are most commonly used to define an ellipsoid.

Figure 1.7 Modelling the shape of the Earth

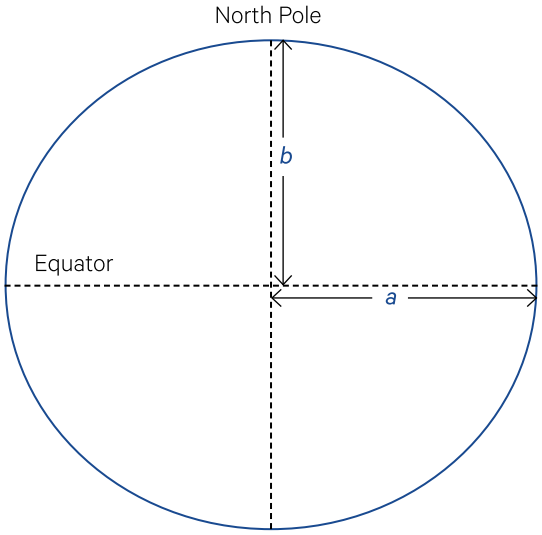
The shape of Earth is irregular. With increasing accuracy, it can be modelled by a sphere, an ellipse or a geoid.



Source: Pennsylvania State University (CC BY-NC-SA 3.0). Retrieved from *https://www.e-education.psu.edu/geog486/node/1883*

Figure 1.8 Definition of Earth shape parameters

The degree of flattening of the globe (*f*) is determined from the major (*a*) and minor (*b*) axes such that *f* = (*a*-*b*)/*b*. The geoid varies from the ellipsoid due to variations in the 'pull' of gravity on the surface. The actual shape of Earth's surface results from topography above or below the geoid.



Source: Harrison and Jupp (1992) Figure 6.

A range of ellipsoids has been proposed at different times and for different regions. The advent of space geodesy has helped refine ellipsoid estimates. Ellipsoidal parameters are used to model Earth's shape and size for both geodetic coordinate conversions and satellite rectification models.

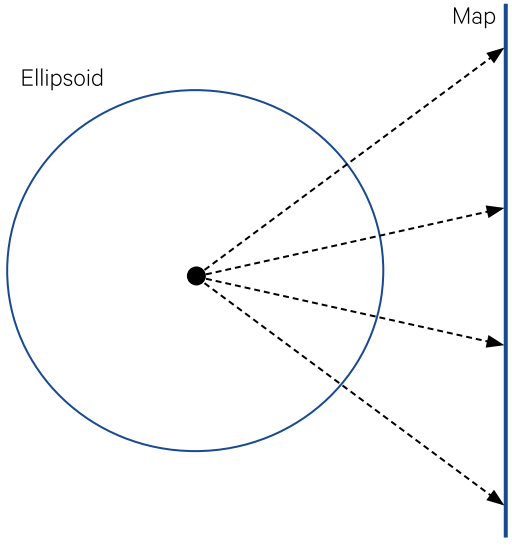
1.2.2 Projections

Cartography attempts to represent the curved surface of the Earth on a flat sheet or map. Conceptually, the Earth's surface is 'projected' onto a surface that could be flattened (this is referred to as a 'developable surface'). This process can be viewed as light rays from a defined source (or 'projection centre') transferring spatial information from a transparent globe onto a developable surface using the laws of perspective (see Figure 1.9). Differences in such 'geometric' map projections occur when both the location of the projection centre and the shape and position of the projection surface vary with respect to the Earth's globe. Some map projections introduce algorithmic components that 'tinker' with the geometrical relationships between projection centre, globe surface and developable surface to generate maps with particular properties.

Since the projection process involves transferring three-dimensional information onto two-dimensions, distortions are unavoidable. To fit onto the map, different regions of the Earth's surface may need to be compressed or stretched, and such distortions may vary in different parts of a map (Bolstad, 2008). Thus, the art of map-making involves minimising the distortions on a given map to best suit its intended purpose.

Figure 1.9 Conceptual basis of a map projection

A map is effectively the result of systematically projecting information from the Earth's surface onto a defined surface from the perspective of a defined source. This process is most easily visualised as light rays from a source being projected through a transparent globe onto a flat surface, or onto a surface that can be flattened, such as a cone or a cylinder.



Adapted from: Bolstad (2008) Figure 3-26

Three fundamental and mutually exclusive properties have been defined for map projections:

- equidistance—accurate representation of distances. A map is equidistant if scale is preserved along a set of standard lines; that is, distances along those lines are proportional to the distance between corresponding locations on the Earth's surface.
- conformality—accurate representation of shapes and angles. A conformal (or orthomorphic) map locally preserves angles so that any two lines on the Earth's surface intersect at the same angle as the projections of those lines on the map. This property also preserves shape at a point, so that small objects are only slightly distorted but larger features have increasing distortion. and
- equivalence—accurate representation of areas. An equivalent (or equal-area) projection maintains areal relationships of regions so may be used to compare the areas of different features (Richardus and Adler, 1972; Castleman, 1979).

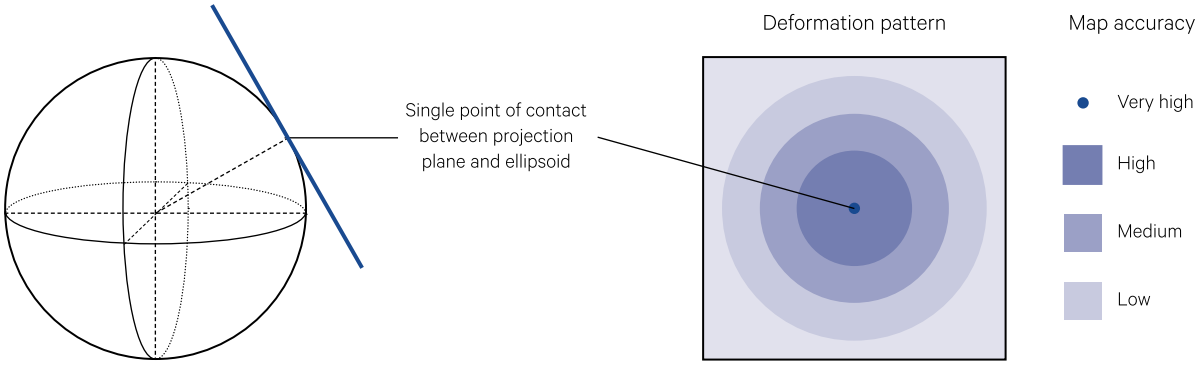
Various map projections have been developed for different purposes. For example, some applications, such as navigation, require that angular information be accurately represented on a map, while others need distances or areas to be truthfully represented.

To produce a map, the curved surface of the Earth is projected onto a flat surface (or a surface that can be flattened), typically a plane, a cylinder or a cone (the latter two being 'cut and unrolled' to form a flat surface; see Figure 1.10). Depending on the type of projection used, only small areas around a point or along a line on the map, where the projection surface 'touches' the globe, accurately represent the Earth's surface. When using a map, it is essential to be aware of its projection characteristics and the effect they may have on the accuracy of any measurements taken from the map (Robinson *et al.*, 1978).

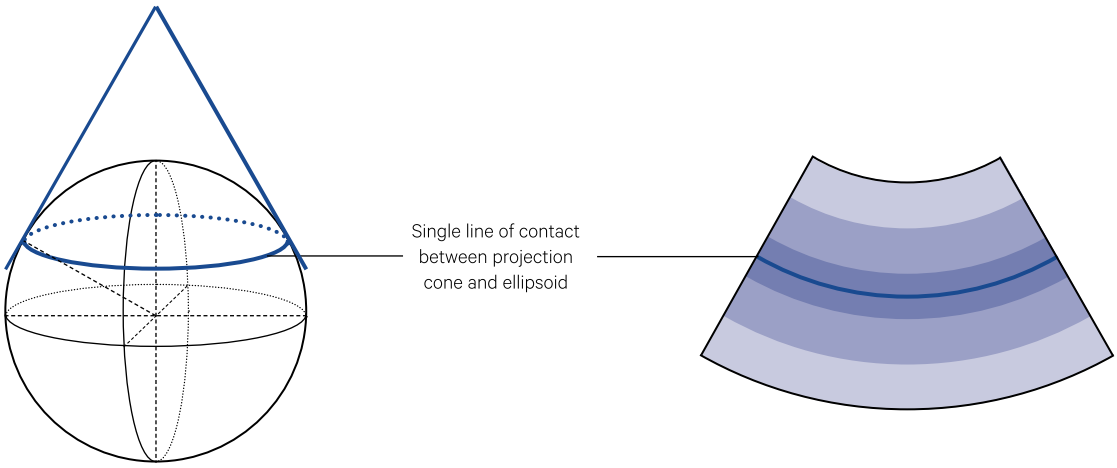
Figure 1.10 Major map projection surfaces

The three commonly used projection surfaces are shown with the deformation patterns that occur in the resultant maps. Spatial accuracy on maps Is highest where the projection surface most closely contacts the globe.

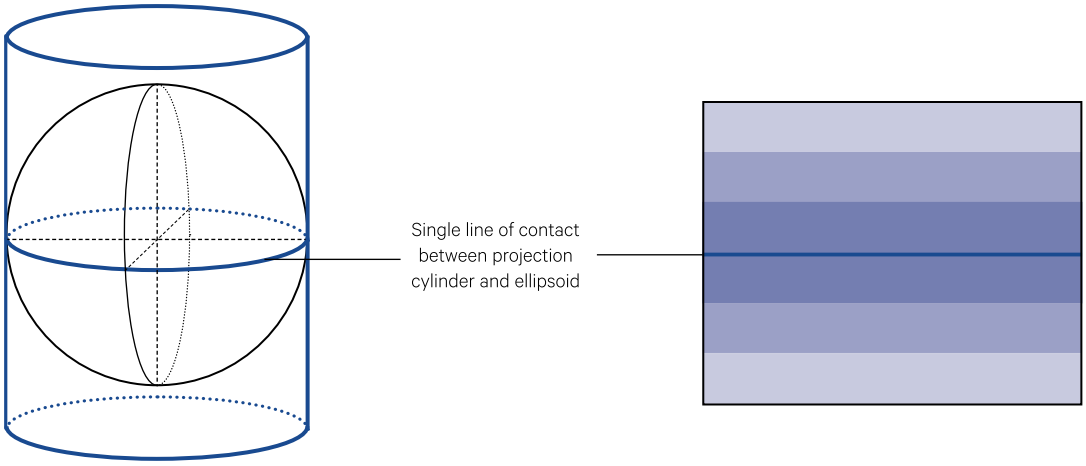
a. Planar projection



b. Conical projection



c. Cylindrical projection



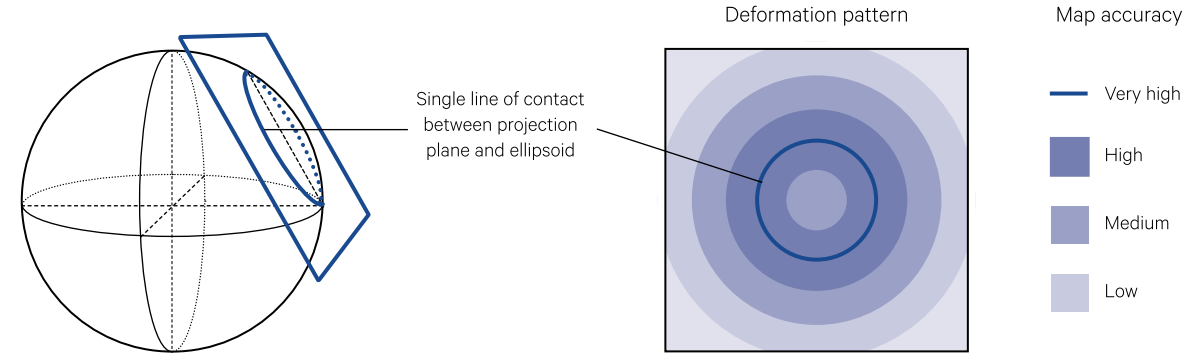
Adapted from: Richardus and Adler (1972) and Robinson *et al.* (1978)

These three projection surface types can be further differentiated by allowing the map surface to contact the Earth at one, two or more points or lines. Figure 1.10 shows the case of the single point/line of contact or 'tangential coincidence'. To increase the extent of contact between surfaces (and hence the area of minimum distortion) 'secant coincidence' results in a line of contact instead of a point for a plane projection (see Figure 1.11a), and two lines of contact when the projection surface is a cone or cylinder (see Figure 1.11b/c). A series of successive projection surfaces, or 'polysuperficial coincidence', would further increase the extent of contact between the map and Earth surfaces and consequently result in larger areas of minimum distortion in the map (Richardus and Adler, 1972).

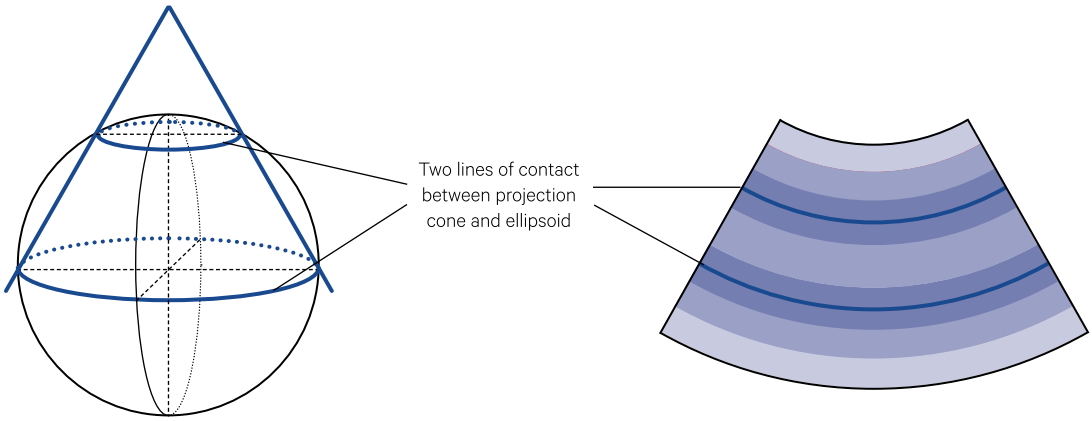
Figure 1.11 Secant contact between projection surface and sphere

Secant contact between developable surfaces and the globe is shown for the three commonly used projection surfaces. The deformation patterns indicate that greater proportions of the resulting maps have higher spatial accuracy than can be achieved using tangential contact (see Figure 1.10).

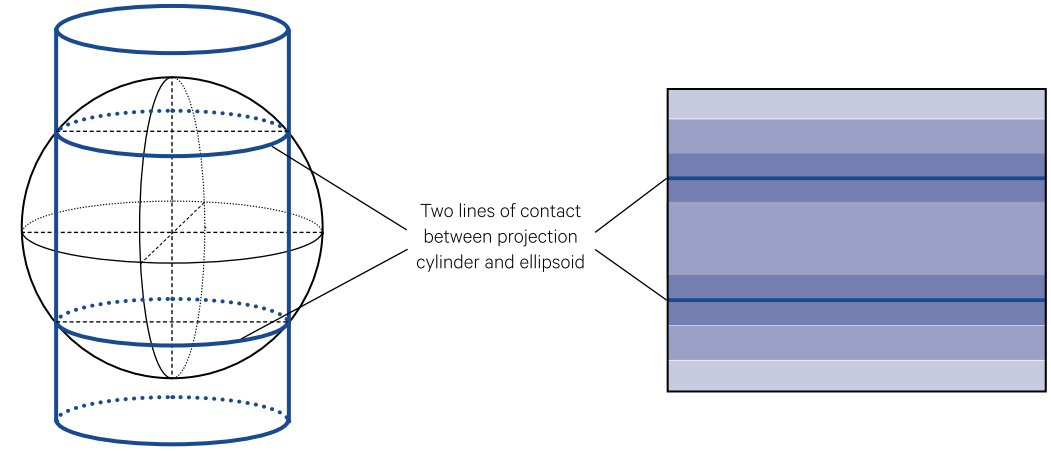
a. Secant projection plane with one circle of contact between projection plane and sphere



b. Conical projection with two circles of contact between projection cone and sphere



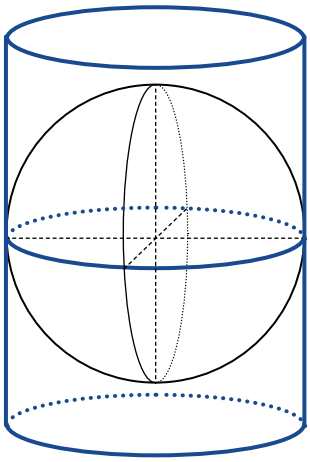
c. Cylindrical projection with two circles of contact between projection cylinder and sphere



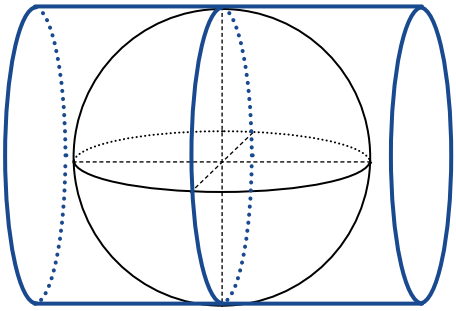
Adapted from: Richardus and Adler (1972) and Robinson *et al.* (1978)

Figure 1.12 Aspects for cylindrical projection

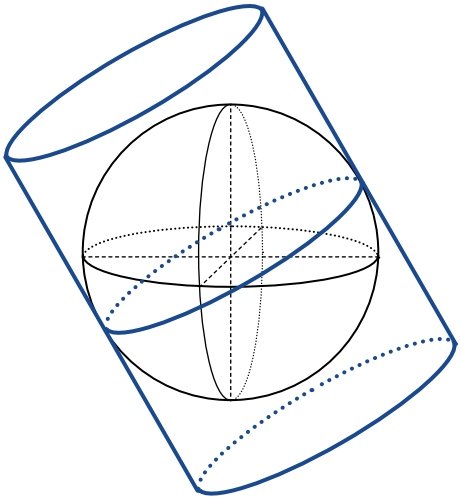
a. Normal cylindrical projection—axis of symmetry for cylinder is Earth's axis of rotation



b. Transverse cylindrical projection—axis of symmetry for cylinder is orthogonal to Earth's axis of rotation, so aligns with the equatorial plane.



c. Oblique cylindrical projection—axis of symmetry for cylinder at an oblique angle to Earth's axis of rotation. In this example, the oblique angle is 30° counter clockwise.



Source: Harrison and Jupp (1992) Figure 9

When using maps to rectify small scale EO imagery, it is important to consider the perspective, or aspect, of the projection surface used to construct the map. There are three possible aspects:

- normal—perpendicular to the plane of the equator;
- transverse—parallel to the plane of the equator; and
- oblique—neither perpendicular nor parallet to the plane of the equator.

Figure 1.12 illustrates these aspects for a cylindrical projection.

Cylindrical projections with normal or transverse aspect present lines of latitude and longitude as a grid (see Figure 1.13a). This means that angles can be measured on the map accurately, but distance is only true along the standard parallel(s). Distance and areal measurements based on cylindrical projections become less accurate as they move away from the standard parallel(s) and become severely distorted in polar regions. The Mercator projection is the most commonly used cylindrical projection, especially for navigation. Pseudo-cylindrical projections, such as the Robinson projection, present meridians of longitude as curved lines (see Excursus 1.1).

On conic projections with normal aspect, meridians of longitude appear as equally-spaced straight lines radiating from the apex of the cone, while parallels of latitude are equally-spaced arcs of circles (see Figure 1.13b). Representation of area and shape is increasingly less accurate away from the standard parallel(s), and distance is only true along the standard parallel(s). An example of a normal conic projection is the Australian Report Map produced by Geoscience Australia (GA), which is based on two Standard Parallels at 18°S and 36°S (see Figure 1.14).

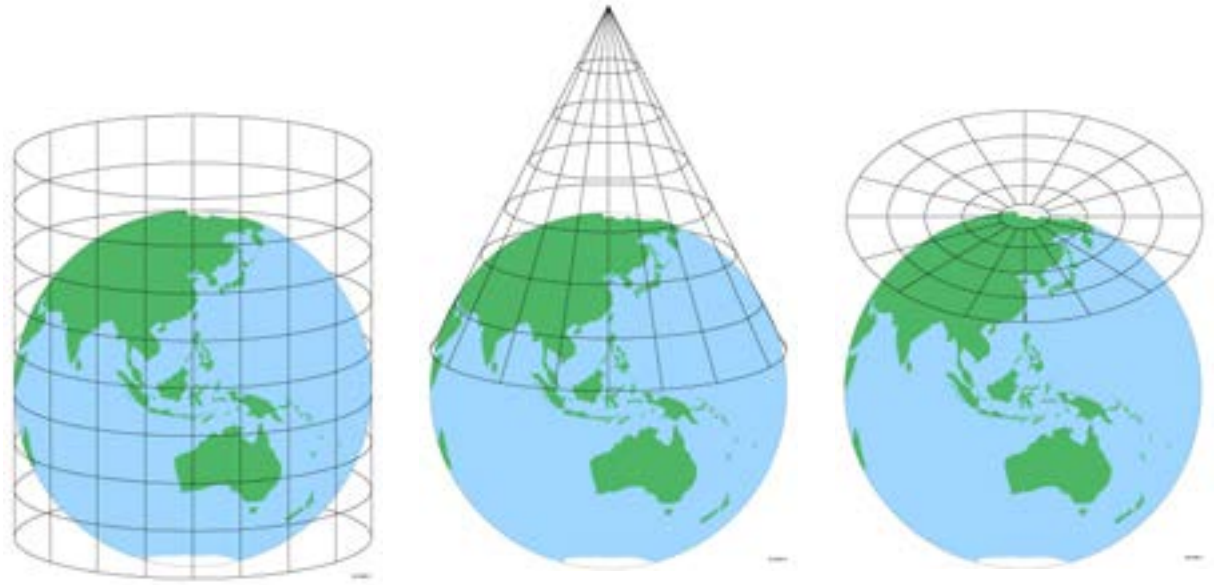
Figure 1.13 Graticules

Different map projections result in differing geometric relationships between lines representing parallels of latitude and meridians of longitude.

a. Cylindrical projection shows graticule as a grid.

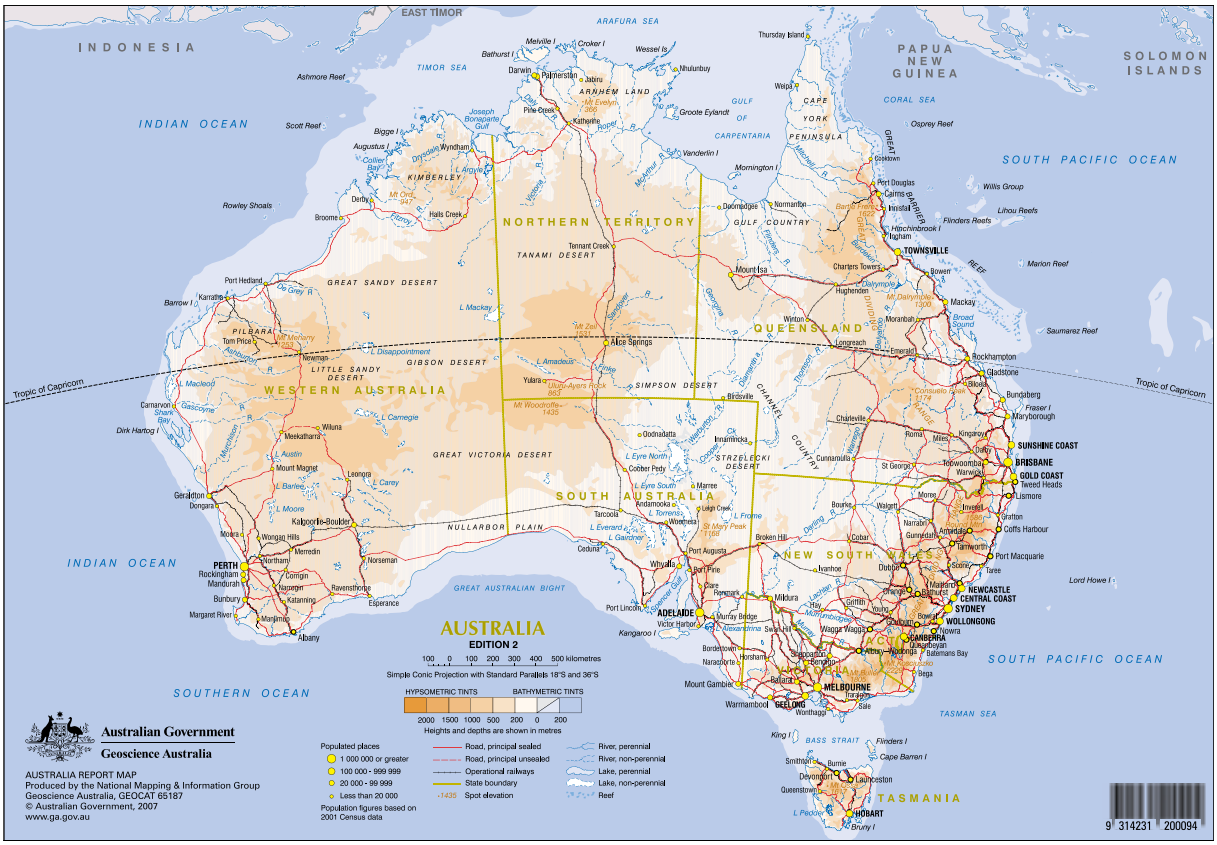
b. Conic projection with normal aspect shows meridians of longitude as evenly-spaced straight lines radiating from the apex of the cone, with parallels of latitude as equally-spaced arcs.

c. Planar projection with the contact point at the North pole, shows meridians of longitude as straight lines radiating from the central point, with parallels of latitude as concentric circles around that point.



Source: Geoscience Australia

Figure 1.14 Australia Report Map



Source: Geoscience Australia.

When the projection surface is a plane contacting the globe at a single point, the resulting map is described as azimuthal. On azimuthal projections with the contact point at the North or South Pole, meridians of longitude appear as straight lines radiating from the central point, and parallels of latitude are represented as concentric circles around that point (see Figure 1.13c). The perspective of an azimuthal projection can vary with the location of the projection centre, including:

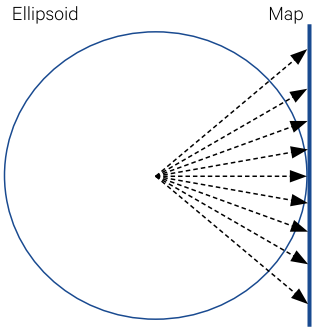
- gnomonic—at the centre of the ellipsoid (see Figure 1.15a);
- stereographic—diametrically opposed to the tangent point (see Figure 1.15b); or
- orthographic—at infinity (see Figure 1.15c).

Distortions in an azimuthal projection are minimal near the centre point but increase away from it. Similarly, compass directions are only true from this point on the map (Bolstad, 2008).

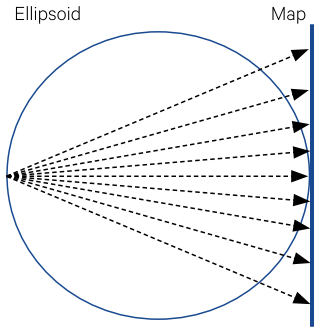
Characteristics of some commonly encountered map projections are introduced in Excursus 1.1

Figure 1.15 Perspectives for azimuthal projections

a. Gnomonic perspective



b. Stereographic perspective



c. Orthographic perspective

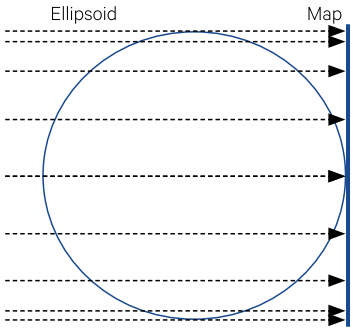
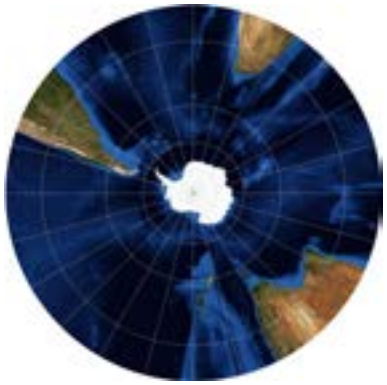
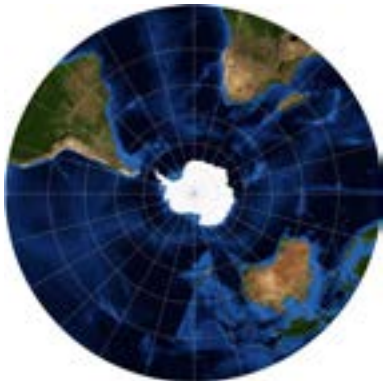


Figure 1.16 Azimuthal maps of South Pole

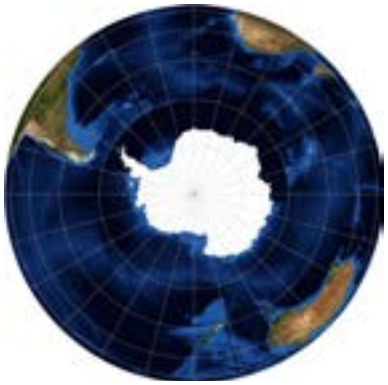
a. Gnomonic perspective with 65° edge radius



b. Stereographic perspective with 90° edge radius



c. Orthographic perspective with 90° edge radius



Created using G.Projector 2.3.0 (Schmunk, 2018)

Excursus 1.1—Common Map Projections

Source: ICSM (2018)
Further Information: *http://www.icsm.gov.au/*

Features of some of the more frequently used map projections are summarised in Table 1.1. The following sub-sections briefly describe these projections and provide example maps, which highlight their differenes.

Table 1.1 Comparison of common map projections

Projection	Type	Properties	Representation of				Dates from	Commonly used for
			Areas	Directions	Shapes	Distance		
Equirectangular	Pseudo-projection	None	Not true	Not true	Not true	Consistent along meridians	BC	GIS
Mercator	Cylindrical	Conformal and true direction	Increasingly less accurate away from equator	True within small areas	True within small areas but large areas are distorted	Increasingly less accurate away from equator	1569	Navigation, Equatorial region maps; World maps
Robinson	Pseudo-cylindrical	None	Not true	True along all parallels and standard meridian	Not true	Constant along equator and other parallels but scales vary	1963	Global maps
Albers Equal Area Conic	Conic	Equivalent	True	True in limited areas	True near standard parallels	Only true along standard parallels	1805	Continental maps
Lambert Conformal Conic	Conic	Conformal	Distortion increases away from standard parallels	Reasonably accurate	True near standard parallels	Only true along standard parallels	1772	Mid-latitude regional maps; Aeronautical charts
Stereographic	Azimuthal	Conformal	Distortion of large shapes increases away from centre point	Only true from centre point	Distortion increases away from centre point	Not true	before AD 150	Polar regions

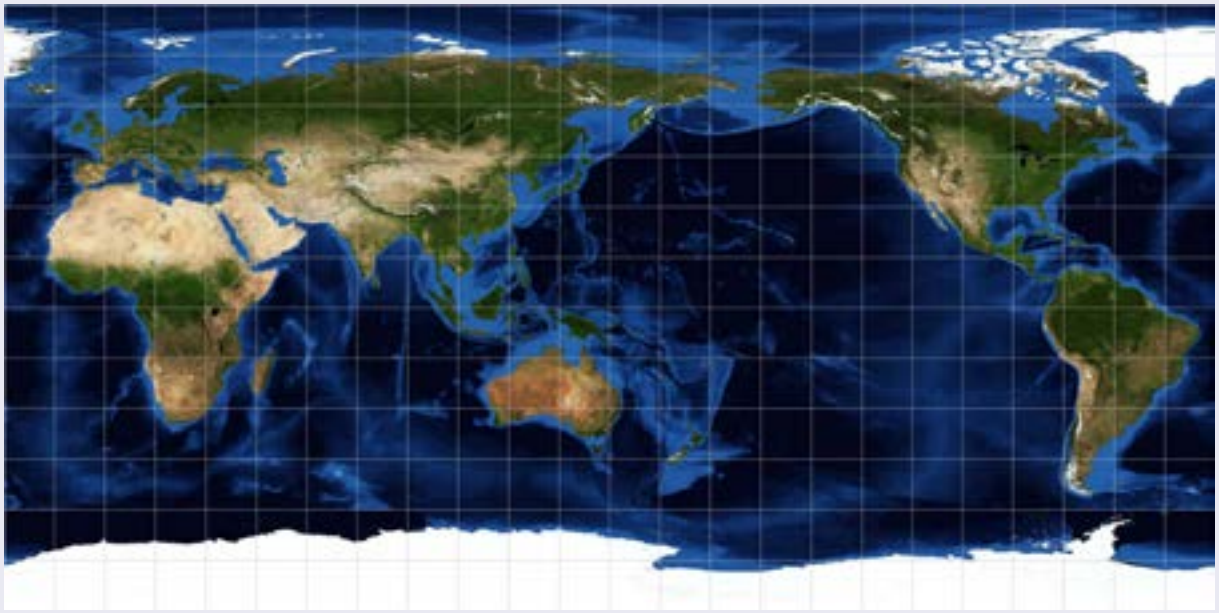
Source: Adapted from ICSM (2018)

Equirectangular, Geographical or Plate Carrée

This simplistic, cylindrical pseudo-projection has been used for nearly 2,000 years (see Figure 1.17). With the equator as the standard parallel, the intersections of latitude and longitude lines form equally-sized squares. Areal and shape representations are distorted but distances along meridians are consistent. Although this representation is not considered a true projection, it is widely used in computing applications including GIS.

Figure 1.17 Equirectangular or Plate Carrée 'projection'

This simplistic, cylindrical 'projection' is centred on 150°E with the equator as the standard parallel.



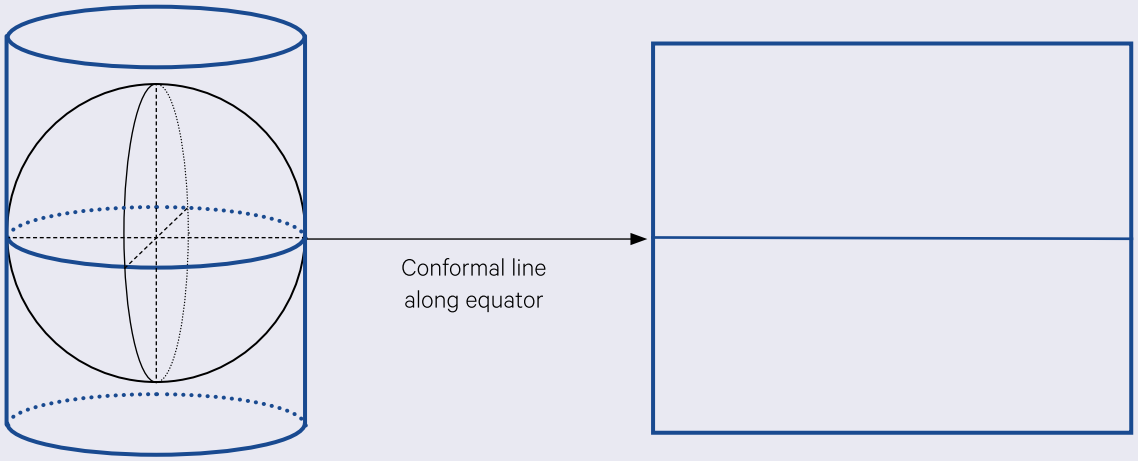
Created using G.Projector 2.3.0 (Schmunk, 2018)

Mercator

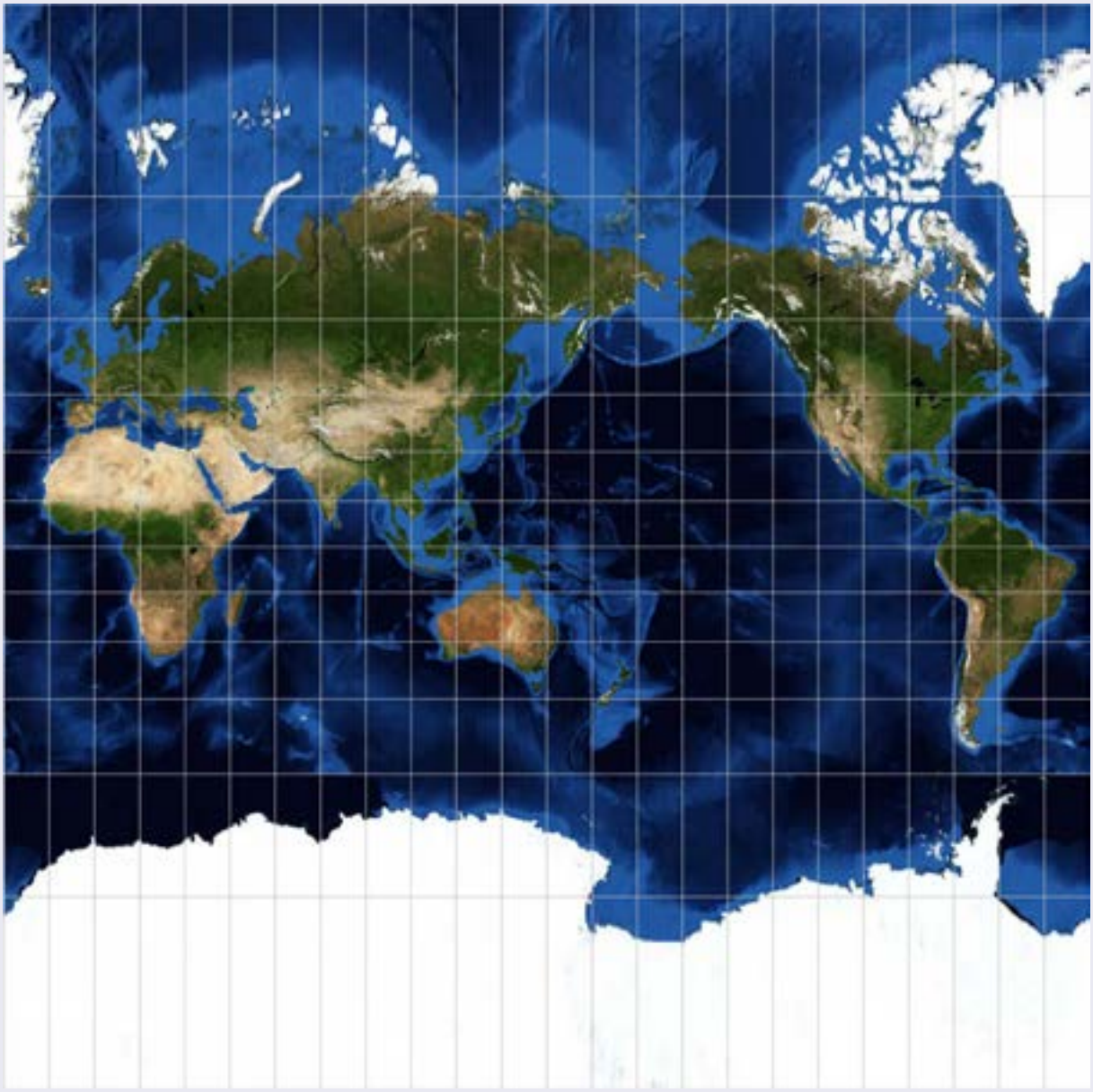
The Mercator projection is a cylindrical projection with a normal aspect and the equator as the standard parallel (see Figure 1.18a). This projection accurately represents the globe's features at the equator, that is, it is said to be 'conformal' at the equator only. Scale, and hence distance measurements, are only correct at the equator. While lines of longitude are equally spaced, the separation between lines of latitude increases towards the poles (see Figure 1.18b). The Mercator projection was developed for navigation so is orthomorphic, or angle-preserving.

Figure 1.18 Mercator projection

a. A normal cylindrical projection unfurls to a rectangular planar map with its conformal line along the equator (proportions shown here for one hemisphere only).



b. Mercator projection centred on 150°E with the equator as the standard parallel



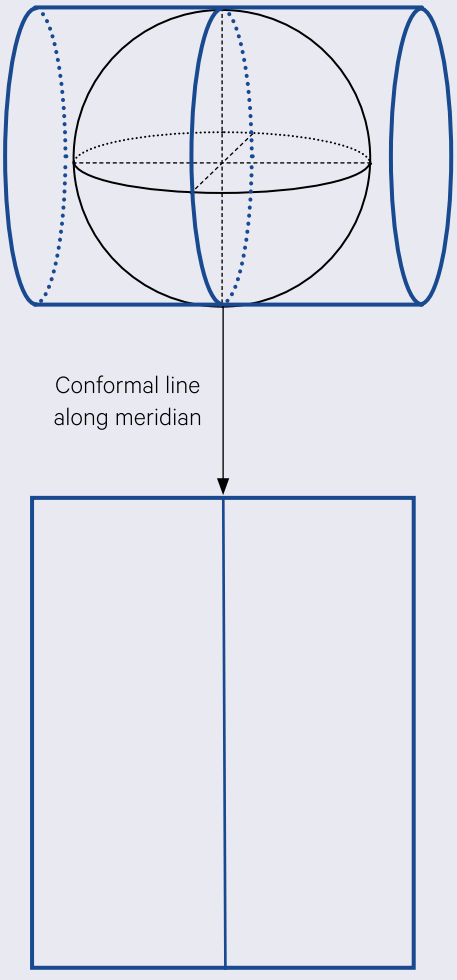
Created using G.Projector 2.3.0 (Schmunk, 2018)

Transverse Mercator

The Mercator projection can also be produced with a transverse aspect as illustrated in Figure 1.19. The Transverse Mercator projection is conformal along the meridian to which it is tangential. Since scale exaggeration increases away from the standard meridian, it is generally used for mapping a small area around that meridian. A global series of Transverse Mercator maps are defined by the Universal Transverse Mercator (UTM) system as described in Section 1.3.

Figure 1.19 Transverse Mercator projection

A Transverse Mercator projection unfurls to a rectangular planar map with its conformal line along the selected meridian of longitude. This projection is generally used for one hemisphere of the globe (as shown) or less.



Robinson

The Robinson projection is a pseudo-cylindrical projection based on the equator as the standard parallel. As with the Mercator projection, areas and shapes are accurately represented close to the equator (up to 15° away). However the pseudo-cylindrical basis for this projection reduces distortions to an acceptable level in mid-latitude regions between 15° and 45° north and south of equator, and reduced distortion exists in polar regions. Unlike the Mercator projection, both latitude and longitude lines are evenly spaced and only the central meridian appears as a straight line. This projection compromises on distortions to produce an attractive map, but it is not conformal, equal area, equidistant or perspective (ICSM, 2018).

Figure 1.20 Robinson projection

This pseudo-cylindrical projection is centred on 150°E with the equator as the standard parallel.

Created using G.Projector 2.3.0 (Schmunk, 2018)

Albers Equal Area Conic

As the name implies, this projection represents areas on the map in proportion to their areas on the ground. Introduced in 1805, Albers Equal Area Conic projection is used for mapping large regions with accurate areal representation (see Figure 1.21). Distances and scale are only true on the standard parallels, while directions are considered to be acceptably accurate.

Figure 1.21 Albers Equal Area projection

Conic projection centred on 135°E and 25°S with two standard parallels 18°S and 26°S.

Created using G.Projector 2.3.0 (Schmunk, 2018)

Lambert Conformal Conic

Lambert Conformal Conic, a conic projection based on two standard parallels of latitude, represents shapes accurately in areas near the standard parallels (see Figure 1.22). Distances, however, are only true along these conformal lines. This projection is used for regional maps for locations between 20° and 60° from the equator and also for aeronautical charts. Parallels of latitude appear as curved lines on all conic projections.

Figure 1.22 Lambert Conformal Conic projection

Conic projection centred on 135°E and 25°S with two standard parallels 18°S and 26°S.

Created using G.Projector 2.3.0 (Schmunk, 2018)

Stereographic

This planar projection preserves shapes in the central areas of the map, but directions are only true from the centre, and areas are not equal (see Figure 1.16). Stereographic maps are most commonly used for polar regions.

1.2.3 Scale

Maps vary not only in terms of their underlying geodetic base and projection, but also in scale. Map scale is represented as a ratio of distance on the map to its equivalent distance on the ground. For example a scale of 1:10,000 means that 1 cm on the map represents 10,000 cm (or 100 m) on the ground (see Table 1.2). In general terms, scale can be described as:

- local—less than 1:50,000;
- regional—1:100,000 to 1:250,000;
- continental—1:1,000,000 to 1:5,000,000; and
- global—greater than 1:5,000,000.

While only one map at 1:5,000,000 covers Australia, the same coverage requires 516 maps at 1:250,000. Thus a map with a larger scale (such as 1:10,000) covers a smaller area, so shows more detail for ground features, while smaller scale maps (such as 1:10,000,000) covers a larger area with less detailed information. When map data is being used to rectify EO imagery, it is important that an appropriate scale be selected to ensure that ground control points are measured with sufficient accuracy (see Table 1.2).

Table 1.2 Map scales

Scale	Ground equivalence of 1 cm on map	Extent (longitude/ latitude)	Common usage	Appropriate to rectify
1:10,000	100 m	0.05 x 0.05	Street directories	Airborne imagery
1:25,000	250 m	0.125 x 0.125	Agriculture	SPOT/Sentinel-2
1:50,000	500 m	0.25 x 0.25	Walking	Landsat TM/OLI Sentinal-2
1:100,000	1 km	0.5 x 0.5	Topographic maps	Landsat MSS
1:250,000	2.5 km	1.5 x 1.0	Long distance navigation; Visual Terminal Charts (VTC)	Landsat MSS
1:500,000	5 km	3.0 x 2.0	Terminal Area Charts (TAC)	MODIS
1:1,000,000	10 km	6.0 x 4.0	International Map of the World series (IMW); World Aeronautical Charts (WAC)	MODIS/AVHRR
1:5,000,000	50 km	continental	General reference and tourism	n/a
1:10,000,000	100 km	global	Global summary	n/a

Sources:
Extent: XNatMap. Retrieved from: *https://www.xnatmap.org/a dnm/docs/1genmap/mapid/mapid.htm*
Usage: Geoscience Australia. Retrieved from: *http://www.ga.gov.au/scientific-topics/national-location-information/topographic-maps-data/basics/what-is-map-scale*

1.3 Map Coordinate Systems

Various coordinate systems may be used for position referencing on the Earth's surface. The fundamental system, to which all other systems are related, references locations to the geographical coordinates of latitude and longitude (see Section 1.3.1). Another map coordinate system commonly used for smaller area mapping is Universal Transverse Mercator (UTM; see Section 1.3.2).

Analytical geometry has never existed.
There are only people who do linear geometry badly,
by taking coordinates,
and they call this analytical geometry.
(Jean Dieudonne)

1.3.1 Geographical coordinates

The equator is defined as the only 'great circle' of the Earth that is perpendicular to the axis of rotation, and hence equidistant from the poles. A great circle is any circle drawn on the Earth's surface, the plane of which intersects the centre of the Earth. A 'small circle' differs from a great circle in that its plane does not intersect the centre of the Earth.

Parallels of latitude are 'small circles' parallel to the equator that measure distance north and south to the poles. Latitude ranges from 90° at the North Pole through 0° at the equator, to -90° at the South Pole. On a spherically-defined Earth, meridians of longitude form 'great circles' around the globe, which pass through the poles. Where an ellipsoid is used to define the Earth, meridians of longitude are ellipses that have the same semi-major and semi-minor axes that are used to define the ellipsoid. Meridians are numbered from 0° at the prime meridian, which passes near Greenwich, England, and span 360° around the globe. Measurements east from Greenwich are given as positive degrees or degrees east, and measurements west are given as negative degrees or degrees west. The system of meridians of longitude and parallels of latitude is referred to as a graticule on a map. In image processing systems, geographical coordinates are generally given as decimal degrees (to at least four decimal places; Trevithick, 2015), rather than in degrees, minutes and seconds.

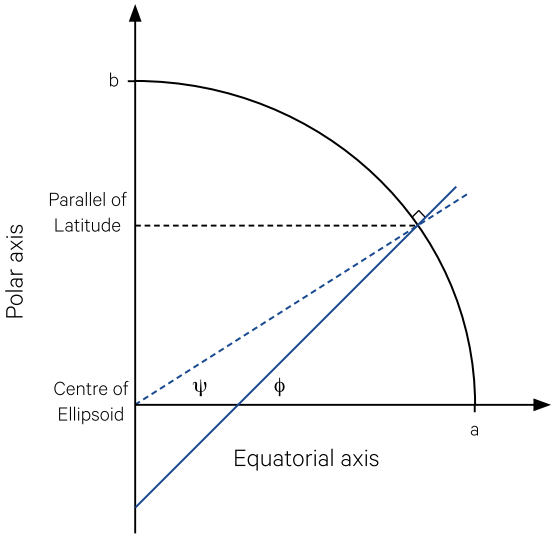
The non-spherical shape of the Earth allows latitude to be measured in a number of ways. The most commonly used method for mapping is geodetic latitude, which measures the angle between the equatorial plane and the normal to the spheroid at a selected point (see Figure 1.23). If the Earth were a perfect sphere this angle would occur at the centre of the Earth. Since geodetic latitude is relative to the surface of the spheroid, different datums and/or ellipsoid models can result in different measured geodetic latitude values for a given point on the Earth's surface.

Geocentric latitude measures an angle at the centre of the Earth's ellipsoid between the equator and the geocentric radius at a point (see Figure 1.23). This angle is less than the geodetic latitude (except at the poles and at the equator where the two are equal) and is used for satellite tracking and other astronomical measurements. The actual difference in terms of ground distances between geodetic and geocentric latitude is significant at mid-latitudes.

Geodetic longitude is the same as geocentric longitude (see Figure 1.24). Geocentric latitude and longitude can be considered as the polar coordinate equivalents of a three-dimensional Cartesian coordinate system that has the origin at the centre of the spheroid (see Section 1.1.1).

Figure 1.23 Measurement of geodetic and geocentric latitude

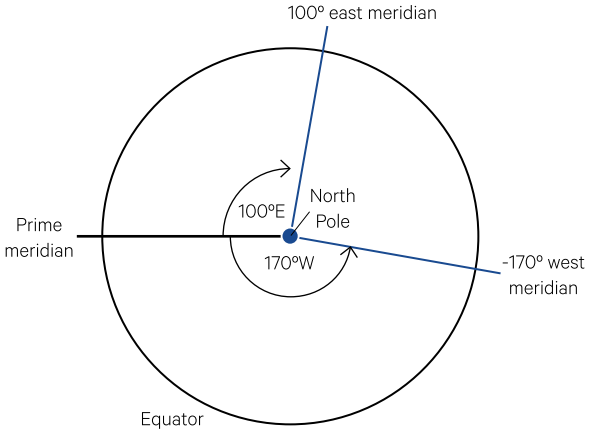
Reference framework for geodetic latitude (ϕ) compared with geocentric latitude (ψ).



Source: Harrison and Jupp (1992) Figure 11

Figure 1.24 Measurement of Longitude

Looking at the Earth's globe from above the North Pole, longitude is measured in degrees east or west from the Prime Meridian, which passes through Greenwich, England.



Source: Harrison and Jupp (1992) Figure 11

1.3.2 Universal Transverse Mercator

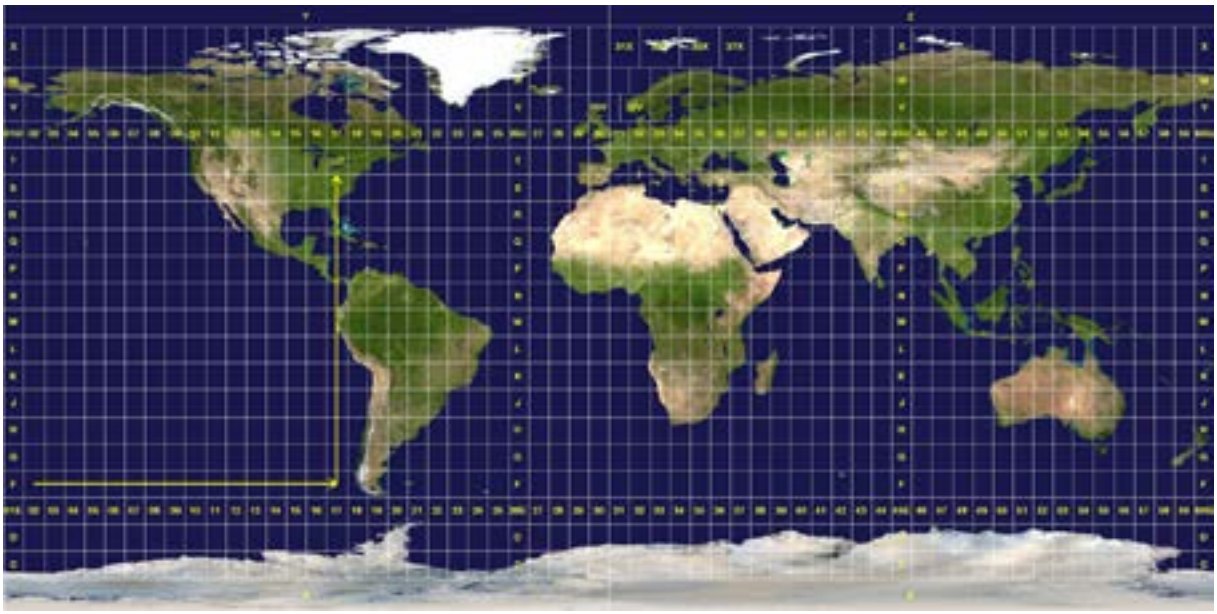
The Universal Transverse Mercator (UTM) system uses a series of 60 Transverse Mercator projections around the globe as the basis of a rectangular grid coordinate system with each UTM zone spanning 6° of longitude (see Figure 1.25). To minimize the scale variations that occur away from the central meridian, the projection for each zone is made secant at two small circles that are parallel to, and 180 km east and west from, the central meridian. Scale is then constant along the meridians that intersect these circles and distortion within the entire zone is minimal. Distance measurements from UTM-based maps are very accurate, since the distortion inherent to the map projection is generally less than human measurement error.

UTM zones extend from 80° North of the equator to 80° South and are numbered eastward around the equator with zone 1 being centred on the 177° meridian. Although the UTM system has global applications, different countries use different spheroids, or projecting globes, so local mapping grids based on UTM have regional limits.

Coordinates on the Map Grid of Australia 2020 (MGA2020) are derived from a UTM system based on the Geodetic Datum of Australia (GDA; see Volume 1A—Excursus 3.1). This system uses precise reference points to register the projection with surveyors' plans. This grid covers Australia (and the territories administered by Australia) but does not include Heard Island, McDonald Islands or the Australian Antarctic Territory (see Figure 1.26). MGA zones are numbered from zone 46 with the central meridian as 93°E to zone 59 with the central meridian as 171°E.

Figure 1.25 Universal Transverse Mercator grid

The global grid for Universal Transverse Mercator (UTM) defines 60 zones that each span 6° of longitude.



Source: Jan Krymmel (Wikimedia Commons). Retrieved from: https://commons.wikimedia.org/wiki/File:Utm-zones.jpg#file

In the UTM grid, zones are 6° wide with a half-degree overlap at the equator (totalling 725 km). The true origin of each zone is the intersection of its central meridian with the equator. The rectangular coordinates, called Eastings and Northings, measure distance within each zone as metres east and north from a false zone origin. In the northern hemisphere the false origin is a point 500,000 metres west of the true origin; in the southern hemisphere, to maintain positive coordinate values, the false origin is 500,000 metres west and 1,000,000 metres south of the true origin.

The cross-section of MGA is shown in Figure 1.27 for two adjacent zones. Map scale is true only along the standard meridians for each zone, where the scale factor (k) equals one. The scale factor is slightly reduced in the centre of the zone, between the standard meridians, which means that, proportionally, map distances are slightly less than the ground distances they represent. Between the standard meridians and the zone boundary, the scale factor is slightly enlarged.

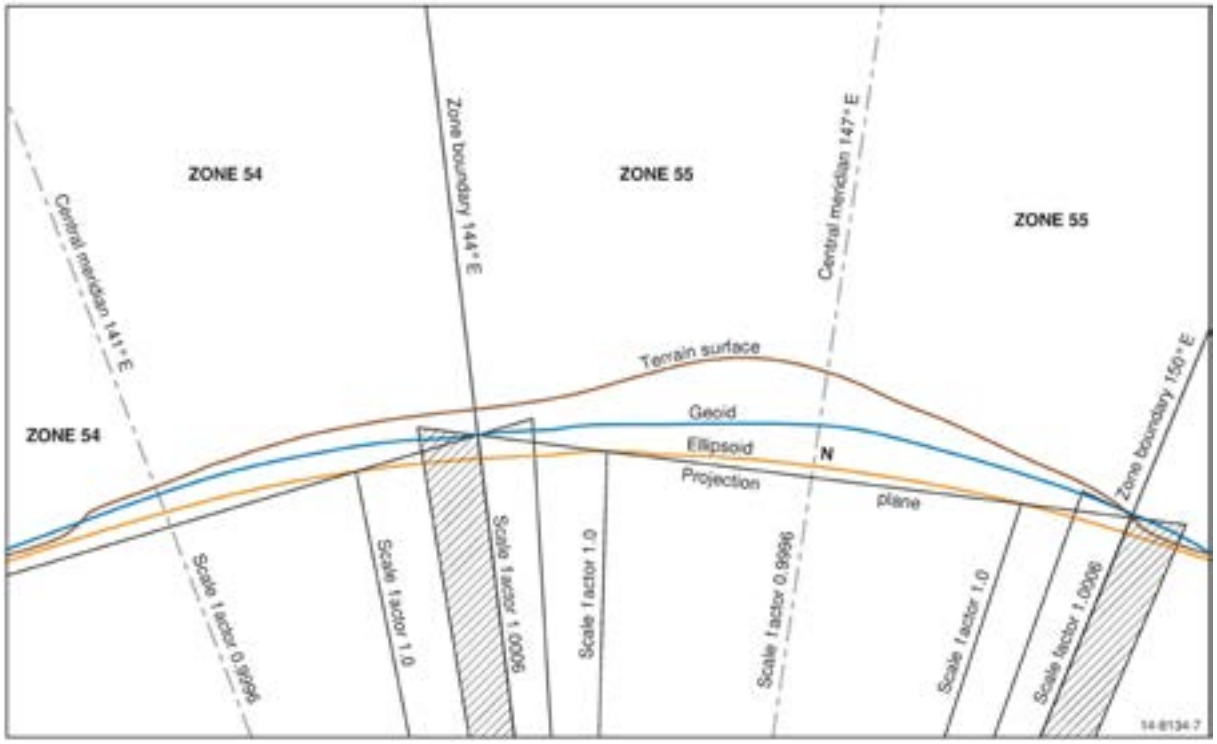
Figure 1.26 Map Grid of Australia

Based on the International Map of the World (IMW), the 1:1,000,000 scale map series indicative index shows the 50 map sheets covering Australia, with the alphanumeric identification grid in blue characters. In this series, the two letter combination identifies the hemisphere and a 4° band of latitude, while the two digit number identifies a 6° band of longitude.

Source: xnatmap.org

Figure 1.27 Cross section of MGA zones

This cross-sectional view of two adjacent MGA zones shows the variation in scale factor (k) away from the two standard meridians. In the centre of the zone, between these meridians, map scale is slightly reduced, whereas between the standard meridians and the zone boundaries, map scale is slightly enlarged.



Source: ICSM (2014) page 50

Pre-metric UTM maps measured distances in yards rather than metres and defined a system of 72 zones around the globe. These coordinates can be accurately converted to and from UTM metre coordinates and latitude and longitude if required. Whenever UTM coordinates are given on a map, these should be used to define locations rather than latitude and longitude since they can be read more accurately. On maps where the meridians or parallels are represented by curved lines, great care must be taken reading any latitude and longitude values.

As detailed in Sections 2 and 3, the perspective of the EO image needs to be considered when rectifying global and continental scale imagery. The satellite orbital path can be viewed as an Oblique Mercator Projection, with the orbit itself being considered as an expanded great circle encompassing the Earth. Full scenes of these large area images cover several UTM zones, so UTM coordinates cannot be used to rectify them. The distortions inherent to a map projection are particularly important for large areas. An image that is effectively produced as an Oblique Mercator Projection aligns most closely with an Oblique Mercator map. This is also relevant when using such images in conjunction with map overlays after rectification.

Some continental imagery is distributed in tailored projections, which often need to be reprojected for compatibility with other data sets. For example, higher resolution MODIS land products are typically supplied in the MODIS Sinusoidal Projection to minimise distortion (NASA, 2018). Image processing methods to rectify and resample EO imagery are detailed in subsequent sections.

1.4 Mapping standards

Mapping standards are defined and updated by relevant authorities to ensure that positioning systems are accurate and consistent. Ongoing changes to mapping standards are necessary for several reasons:

- movement of tectonic plates—for example, Australia's tectonic plate moves to the north-north-east by approximately 7 cm per year (see Volume 1B—Excursus 10.1);
- improvements in the measurement of reference points on the Earth's surface—for example, differences between the International Terrestrial Reference Frame 1992 (ITRF1992) and ITRF2014 result in variations of approximately 9 cm in ellipsoidal heights in Australia (see Volume 1A—Excursus 3.2); and
- deformation of the Earth's crust due to subsidence and tectonic activity (see Volume 1A—Section 3.2).

In addition, with the increasing spatial accuracy of the ubiquitous positioning devices based on the Global Navigation Satellite System (GNSS; see Volume 1B—Section 10), it is imperative that local mapping standards be regularly updated for compatibilty with global systems (see Section 1.4.1).

1.4.1 Australia

Detailed parameters for Australian mapping standards are defined by ICSM (2018). The datum being used in Australia in 2018 is designated the Geocentric Datum of Australia 1994 (GDA94; see Volume 1A—Excursus 3.1 and Excursus 3.2). This datum was adopted by the ICSM in 1994 (ICSM, 2014) to replace the previously used Australian Geodetic Datums, AGD66 and AGD84 (NMCA, 1972; Lauf, 1983).

GDA94 is a geocentric datum that is part of the Global Geodetic Reference Frame (GGRF) and compatible with the Global Navigation Satellite System (GNSS). Its definition comprises:

- reference frame: ITRF92 (International Terrestrial Reference Frame 1992) at epoch of 1994.0
- ellipsoid: Geodetic Reference System 1980 (GRS80):
  - semi-major axis (*a*): 6,378,137 m; and
  - inverse flattening (1/*f*): 298.257222101.

This datum "*includes all the areas contained within Australia's marine jurisdiction within 200 nautical miles of Australia and its external territories, and the areas of Australia's continental shelf beyond 200 nautical miles as confirmed by the United Nations Commission on the Limits of the Continental Shelf. The areas include Cocos (Keeling) Island, Christmas Island, Norfolk Island and Macquarie Island but excludes Heard-McDonald Island and the Australian Antarctic Territory (AAT)*" (ICSM, 2014).

However, given the dynamic nature of our planet, in particular the movement of tectonic plates, by 2020 the difference between GDA94 coordinates and ITRF92 coordinates will be approximately 1.8 m. Accordingly, the GDA94 is being replaced by GDA2020, which will align more closely with global satellite positioning systems. GDA2020 is also based on the GRS80 ellipsoid but will be aligned with a realisation of ITRF2014 at epoch 2020.0 (GA, 2018b). The transition process to GDA2020 is described in Volume 1B—Excursus 10.1 (see also GA, 2018c). The extents of GDA2020 are the same as those for GDA94.

The standard map coordinate system used in Australia is the Map Grid of Australia (MGA94), based on the Universal Transverse Mercator (UTM) projection and GDA94. This replaces the previous Australia Map Grid (AMG) systems of AMG66 (based on UTM projection and AGD66) and AMG84 (based on UTM and AGD84). With GDA2020, MGA94 will be replaced by MGA2020.

1.4.2 Global

The World Geodetic System 1984 (WGS84) is a geocentric, Earth-fixed terrestrial reference system, ellipsoid and geodetic datum that was developed by the US Department of Defence and is now maintained by US National Geospatial Intelligence Agency (NGA). It has been updated several times with the latest version (G1762) being aligned to the GNSS realisation of ITRF2008 (IGb08). Globally distributed monitoring stations are used to update the reference frame for tectonic movement (see Figure 2.2 in NGA, 2014).

WGS84 integrates various models and parameters, fundamentally including the shape and size of Earth (semi-major axis (*a*) = 6,378,137 m and inverse flattening (1/*f*) = 298.257223563), as well its gravity and magnetic fields (NGA, 2014). WSG84 is commonly used for global imagery so offers a relevant cartographic base for continental-scale EO products, but presents limitations to regional-scale applications in Australia (GA, 2018d).

The European Survey Petroleum Group (EPSG—a consortium of positioning professionals from European oil companies) defined a Geodetic Parameter Data Set, which links map projections to coordinate reference systems (IAOGP, 2018). The EPSG::4326 defines geographic coordinates based on the first version of WGS84. EPSG::7665 defines geographic coordinates based on the latest version, WGS84 (G1762).

1.5 Further Information

Metadata

Open Geospatial Consortium (OGC): (OGC, 2012)

The Generic Earth Observation Metadata Standard (GEOMS): Retscher *et al.* (2011)

Foundation Spatial Data Framework (FSDF): *http://www.anzlic.gov.au*

Datums and Projections

Geoscience Australia: *http://www.ga.gov.au/earth-monitoring/geodesy/geodetic-datums.html*

ICSM: *https://www.icsm.gov.au/education/fundamentals-mapping/overview-fundamentals-mapping*

Map Projections: *https://geographx.co.nz/map-projections/*

Distortions on Map Projections Using Gedymin Profiles: *http://geoawesomeness.com/wp-content/uploads/2017/04/Gedymin-faces-Map-Projections.png*

Global Map Projector software (Schmunk, 2018), NASA: *https://www.giss.nasa.gov/tools/gprojector/*

Datum Standards

Geocentric Data of Australia (GDA2020): *http://www.ga.gov.au/scientific-topics/positioning-navigation/datum-modernisation*

ICSM (2014)

NGA (2014)

WGS84: *http://www.ga.gov.au/scientific-topics/positioning-navigation/wgs84*

Coordinate Systems and Projections

PenState College of Earth and Mineral Sciences: *https://www.e-education.psu.edu/geog486/l7.html*

Robinson *et al.* (1978)

Snyder (1987)

1.6 References

ASPRS (1984). *Multi-lingual Dictionary of Remote Sensing and Photogrammetry*. Falls Church, Virginia.

Bolstad, P. (2008) *GIS Fundamentals. A First Text on Geographic Information Systems*. 3rd Edn. AtlasBooks, White Bear Lake, Minnesota.

Castleman, K.R. (1979). *Digital Image Processing*. Prentice-Hall Inc. New Jersey.

Harrison, B.A., and Jupp, D.L.B. (1992). *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual*. MPA, Melbourne.

GA (2018a). Geoscience Australia webpage: *Australia's Maritime Jurisdiction: http://www.ga.gov.au/metadata-gateway/metadata/record/70362/*

GA (2018b). Geoscience Australia webpage: *Geocentric Datum of Australia 2020 (GDA2020)*: *http://www.ga.gov.au/scientific-topics/positioning-navigation/geodesy/datums-projections/gda2020*

GA (2018c). Geoscience Australia webpage: *Datum Modernisation in Australia*: *http://www.ga.gov.au/scientific-topics/positioning-navigation/datum-modernisation*

GA (2018d). Geoscience Australia webpage: *What are the limitations of using World Geodetic System 1984 in Australia? http://www.ga.gov.au/scientific-topics/positioning-navigation/wgs84*

IAOGP (2018). *International Association of Oil and Gas Producers EPSG Geodetic Parameter Dataset*: *http://www.epsg-registry.org*

ICSM (2014). *Geocentric Datum of Australia Technical Manual*. Version 2.4, Intergovernmental Committee on Surveying and Mapping (ICSM) Permanent Committee on Geodesy (PCG), ISBN 0-9579951-0-5.

ICSM (2018). *Intergovernmental Committee on Surveying and Mapping* website: *http://www.icsm.gov.au/*

Lauf, G.B. (1983). *Geodesy and Map Projections*. TAFE Publications Unit, Melbourne.

Lewis, A., Wang, L.-W., and Coghlan, R. (2011). *AGRI: The Australian Geographic Reference Image. A Technical Report*. GeoCat#72657. Geoscience Australia, Canberra. ISBN 978-1-921954-47-4.

NASA (2018). MODIS Land webpage: *https://modis-land.gsfc.nasa.gov/MODLAND_grid.html*

NGA (2014) *National Geospatial-Intelligence Agency Standardization Document*. Department of Defense World Geodetic System 1984. Version 1.0.0. NGA. STND.0036_1.0.0_WGS84. *http://earth-info.nga.mil/GandG/publications/NGA_STND_0036_1_0_0_WGS84/NGA.STND.0036_1.0.0_WGS84.pdf*

NMCA (National Mapping Council of Australia). (1972). *The Australian Map Grid Technical Manual. Special Publication # 7.* AGPS. Canberra.

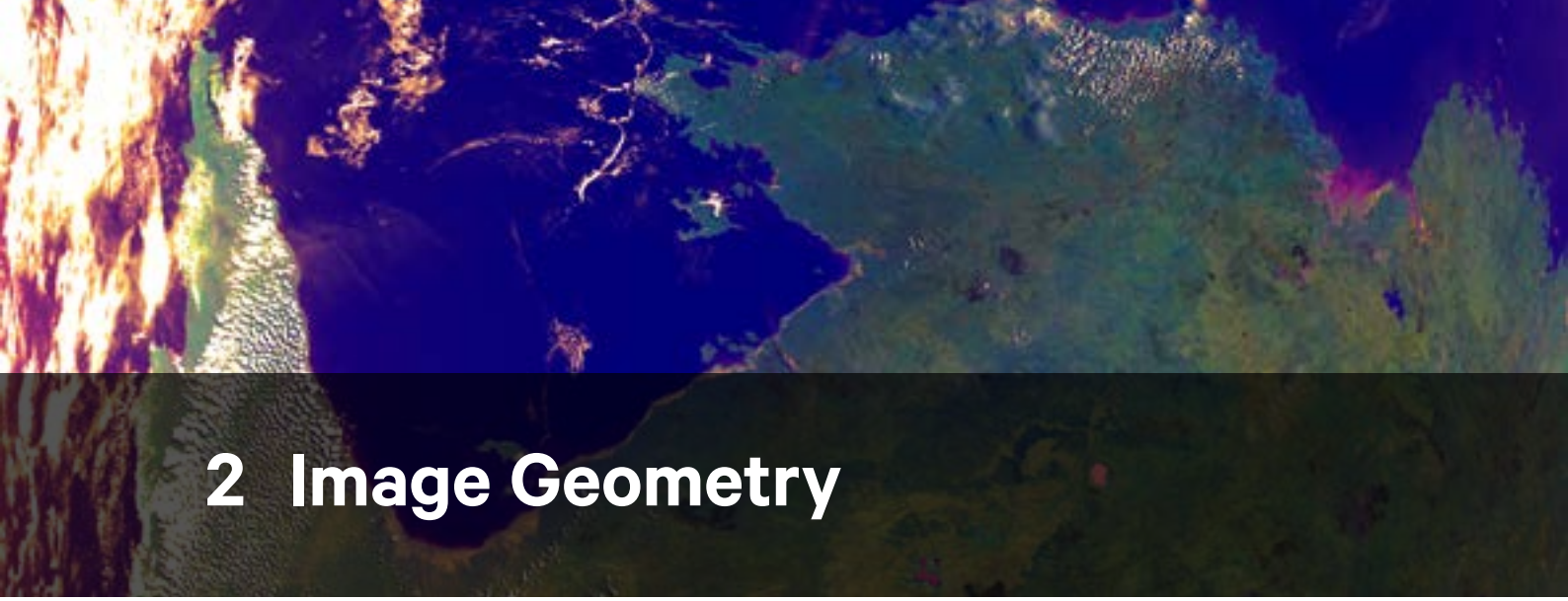
Richardus, P. and Adler, R.K. (1972). *Map Projections for Geodesists, Cartographers and Geographers.* North-Holland Pub. Co. Amsterdam.

Robinson, A., Sale, R., and Morrison, J. (1978). *Elements of Cartography.* 4th Edn. Wiley. NY.

Schmunk, R.B. (2018). G.Projector (Global Map Projector) Software Version 2.3.0, NASA. Retrieved from: *https://www.giss.nasa.gov/tools/gprojector/*

Snyder, J.P. (1987). *Map Projections—A Working Manual.* USGS Prof. Paper 1395. USGS, Washington. *https://pubs.usgs.gov/pp/1395/report.pdf*

Trevithick, R. (2015). Field data collection and management for Earth Observation image validation. Ch 3 in *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products.* Version 1.1. (Eds: Held, A., Phinn, S., Soto-Berelov, M., and Jones, S.). TERN AusCover, ISBN 978-0-646-94137-0.



2 Image Geometry

Various geometric distortion effects can be identified in EO imagery. As introduced in Volume 2A—Section 7, these may be related to characteristics of the sensor and/or the platform used to acquire the imagery, or to the shape or rotation of the Earth. Certain types of distortion are systematic and thus may be modelled with appropriate parameters as discussed in Section 3.

Other sources of distortion that relate to the stability of the imaging platform are more difficult to quantify. An understanding of these sources of distortion and the methods used to correct them in EO imagery is essential for geometrically registering imagery to another image or map with defined accuracy and thus:

- obtaining EO data that is 'fit-for-use';
- understanding the limitations of EO imagery—what can and cannot be done with it;
- selecting appropriate image correction algorithms; and
- critically assessing datasets derived from EO imagery.

Image grid cells are implicitly located relative to a defined reference position, or datum (see Sections 1.1.3 and 2.5). EO data is initially acquired using scanning mechanisms that cannot maintain a precise grid size so most image products are rectified (corrected for known geometric distortions) before distribution. To ensure that the geometry of the resulting image is consistent, the EO data is generally resampled (written to another image grid) to align with a selected map coordinate system (see Figure 2.1).

The perspective of the EO sensor determines the most appropriate map coordinate system for rectifying and resampling EO imagery (see Section 1). Regional scale imagery, such as Landsat scenes, can be easily contained within one UTM (Universal Transverse Mercator) zone, so this projection is generally used to rectify regional scale imagery (see Section 1.3.2). Continental scale imagery, such as MODIS or AVHRR, spans several UTM zones, so UTM maps are not appropriate to rectify those images. The satellite orbital path can be viewed as an Oblique Mercator Projection, with the orbit itself being considered as an expanded great circle encompassing the Earth, so this projection is more appropriate for continental scale imagery (see Section 3.1).

Let no man ignorant of geometry enter here.
(Plato)

Background image: Raw data image over northern Australia as acquired by NOAA-19 on 3 September 2017, displayed using bands 2, 2, 1 as RGB (see also Figure 2.7). **Source:** Edward King, CSIRO

Figure 2.1 Image rectification

Landsat TM images over Adelaide, SA, displayed using bands 3, 2 and 1 as RGB.

a. Raw format image showing ~10° skew due to Earth rotation during image acquistion

b. Image after georegistration to AMG coordinates

c. Map based on AMG coordinates

Source: Megan Lewis, University of Adelaide

Geometry is the science of correct reasoning on incorrect figures
(George Polya)

2.1 Pixel Size

The instantaneous field of view (IFOV) of a detector is the view angle that is used for a single radiance reading (see Volume 1B—Section 1.2). The IFOV is largely determined by the dimensions of the detector's aperture (which is rectangular for some imaging systems, thus giving different angles for along-line and along-track dimensions). In conjunction with the focal length and altitude of the detector, a linear IFOV can be computed for the ground distance spanned by this view angle. However, this distance assumes that the detector has perfectly uniform response over the view angle and no displacement occurs in its position relative to the Earth's surface being imaged. The Modulation Transfer Function (MTF) of an imaging system describes its spatial frequency response so is used to account for characteristics of the aperture, optics and the imaging detectors, and blurring effects due to platform movement (see Volume 1B—Section 2.1.1.3).

A better estimate of the ground area whose radiance is measured by a detector is the Effective IFOV (EIFOV), or optical pixel size, which is determined by adjusting the IFOV by the MTF. The extent of the EIFOV is defined by the points at which the distribution of detected radiance is reduced by 50% due to the system MTF, that is, only half the radiance of a point on the perimeter of this area is being measured by the detector. The EIFOV generally covers a larger area than the IFOV (Slater, 1980).

The geometric pixel size describes the ground area represented by a pixel in the context of a continuous image. This measure is usually smaller than the optical pixel (see Figure 2.2). The average geometric pixel width along an image line can be computed from the total scan angle (or Field of View: FOV), the platform height, and the number of pixels per scan line (see Figure 2.3a) as:

$$\frac{2h\left(\tan\frac{\theta}{2}\right)}{N}$$

or, at the centre of a line, the width of the pixel is:

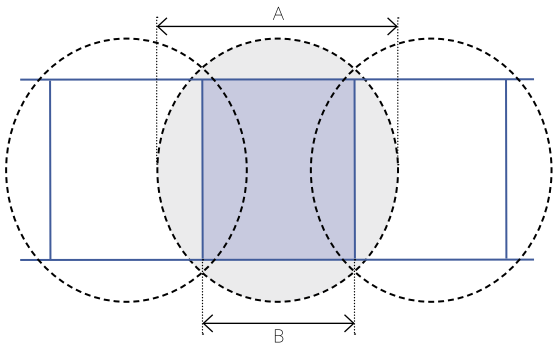
$$2h \times \tan\left(\frac{\theta}{2N}\right)$$

where

h is the platform height;
θ is the total scan angle (FOV); and
N is the number of pixels per scan line.

Figure 2.2 Optical versus geometric pixel sizes

The optical pixel width (that is, the along line width of each pixel as imaged by the sensor) is labelled A. To account for overlapping pixels in the image scaling, the geometric pixel size is given as B.

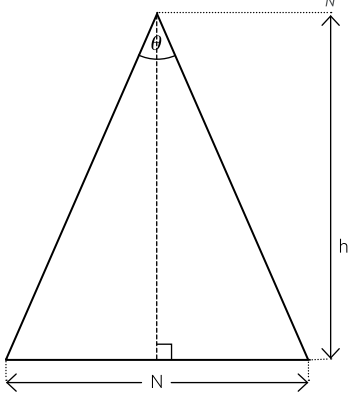


Source: Harrison and Jupp (1989) Figure 35

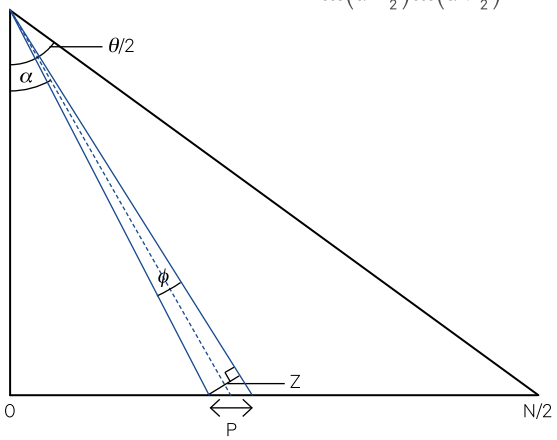
Geometry is the most complete science.
(David Hilbert)

Figure 2.3 Pixel size calculations

a. Average geometric width of a single pixel = $\frac{2h\left(\tan\frac{\theta}{2}\right)}{N}$

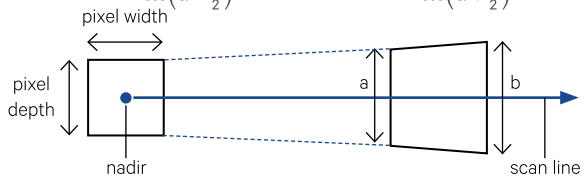


b. Ground width of a single pixel, $P = \frac{h\sin\phi}{\cos\left(\alpha-\frac{\phi}{2}\right)\cos\left(\alpha+\frac{\phi}{2}\right)}$



c. Pixel size increases along scan line with distance from nadir.

Pixel depth $a = \frac{2h\tan\frac{\phi}{2}}{\cos\left(\alpha-\frac{\phi}{2}\right)}$ and pixel depth $b = \frac{2h\tan\frac{\phi}{2}}{\cos\left(\alpha+\frac{\phi}{2}\right)}$

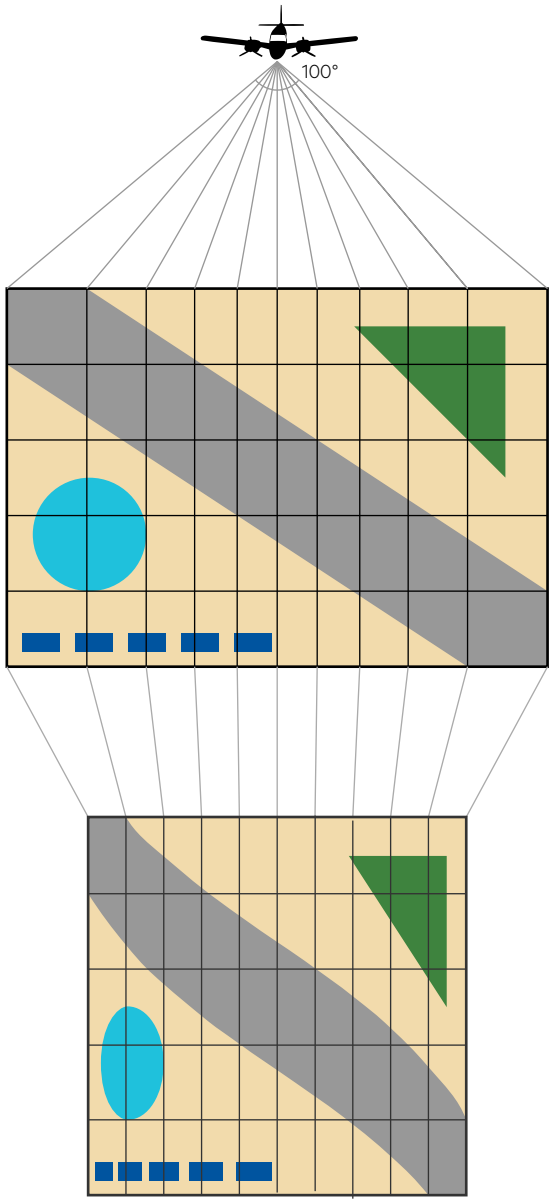


Source: Harrison and Jupp (1992) Figure 15

For an image formed by an electromechanical scanner, the total scan angle and the number of pixels per scan line are usually constant so pixel width can vary with platform altitude. As discussed in Volume 1B—Section 1.2, the actual ground width of pixels along a line can vary significantly for scanners with wide FOV (see Figure 2.4).

Figure 2.4 Effect of wide scan angle

Panoramic distortion means that ground pixel width increases significantly away from a vertical view. This results in image features showing lateral distortion when displayed with a constant pixel width.



Adapted from: Harrison and Jupp (1992) Figure 18

A better estimate of the ground width (P) of a pixel at a particular position along a line can be computed using the dimensions illustrated in Figure 2.3b, where:

$$Z = \frac{h\sin\phi}{\cos\left(\alpha-\frac{\phi}{2}\right)}$$

then

$$P = \frac{Z}{\cos\left(\alpha+\frac{\phi}{2}\right)}$$

$$= \frac{h\sin\phi}{\cos\left(\alpha-\frac{\phi}{2}\right)\cos\left(\alpha+\frac{\phi}{2}\right)}$$

where

ϕ is the IFOV (or EIFOV angle);
N is the number of pixels per scan line;
h is the platform height; and
α is the angular displacement from nadir to the centre of ϕ.

At nadir, $\alpha = 0$ so:

$$P = \frac{h\sin\phi}{\cos^2\left(\frac{\phi}{2}\right)} = 2h\tan\left(\frac{\phi}{2}\right)$$

The average geometric pixel width given above effectively assumes that ϕ equals θ / N for each pixel (where θ is the FOV).

The variation in depth due to look angle within a non-nadir pixel can also be computed as illustrated in Figure 2.3c, with the side closest to nadir, *a*, being

$$a = \frac{2h\tan\frac{\phi}{2}}{\cos\left(\alpha - \frac{\phi}{2}\right)}$$

and the side away from nadir, *b*, being

$$b = \frac{2h\tan\frac{\phi}{2}}{\cos\left(\alpha + \frac{\phi}{2}\right)}$$

These calculations have been applied to two satellite data sources to produce Table 2.1.

Table 2.1 Pixel sizes for satellite data sources

Attribute	Landsat-1/2/3 MSS	AVHRR-9/10
IFOV	0.086 mrad	1.3 mrad
FOV	11.56°	55.3999°
Altitude	913 km	850 km
Pixels per scan line	3230	2048
Average geometric pixel width	57.224 m	1203.261 m
Ground pixel width at nadir	78.518 m	1105.0001 m
Ground pixel width at edge of scan	79.385 m	3426.9041 m

The geometric pixel depth is dependent on the speed of travel of the scanner's platform and can be simply computed from:

platform velocity x *mirror period*

where

platform velocity is defined in distance per unit time; and
mirror period is the time taken to image one scan line.

Since the mirror period is constant for most scanning devices, platform velocity is the major factor varying pixel depth.

2.2 Sampling and Recording Rate

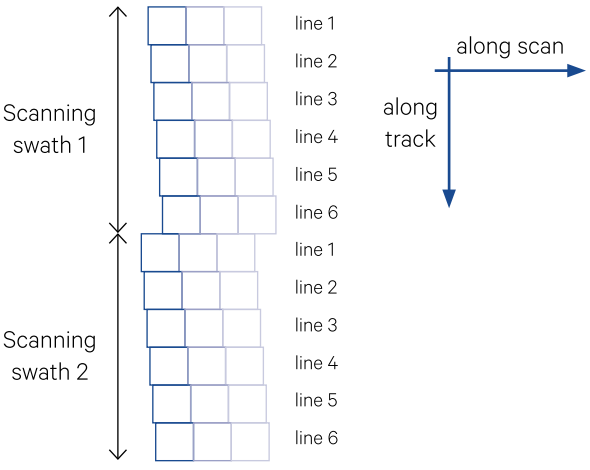
Inconsistencies in the rate of pixel sampling along an image line can cause geometric distortions of sub-areas in the acquired image. If the scanning mechanism used by the sensor does not operate at a constant speed, the ground areas being imaged as individual pixels along a line will not be of equal size and thus do not represent a regular grid pattern. This distortion occurs in electromechanical scanners that use oscillating mirrors, such as Landsat MSS.

The sequential nature of recording sensor values in a multi-sensor scanner also affects image geometry. When a scanner uses a bank of sensors to scan multiple lines in a single swath (that is, one side-to-side scan across the full image scene) a delay occurs between recording the reflectance values of each line's sensor. This distortion is evident in both Landsat MSS and TM imagery and is illustrated in Figure 2.5 for the simpler case of an MSS image.

For similar reasons, a small delay occurs between recording different channels in an image so that the effective ground coverage of each channel is shifted slightly. The total shift is generally less than half a pixel between the first and last channels so is often disregarded.

Figure 2.5 Sensor delay offset for Landsat MSS

Six sensors for each band allow six image lines to be recorded during each west to east scan. Due to sensor delay, the ground locations for the start of each line are slightly offset in the along scan direction as the satellite travels south. In addition, after every scanning swath of six image lines, Earth's rotation introduces a counter offset to the west.



Adapted from: Colwell (1983) Figure 21-25

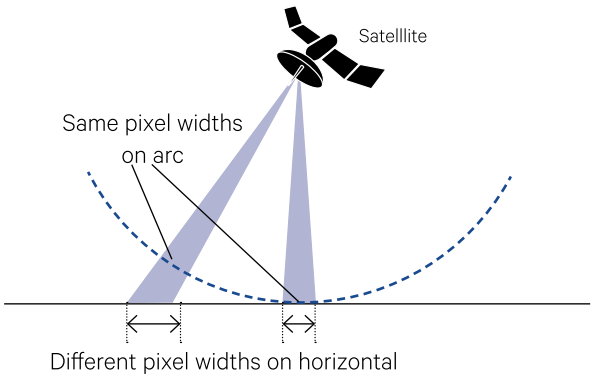
2.3 Viewing Perspective

As discussed in Volume 1A—Section 14 and Volume 1B—Section 1.2, image scanners usually determine the area of a pixel by a constant angle of view (instantaneous field of view: IFOV). When the total scan angle (or field of view: FOV) is large, as occurs for continental scale images (such as MODIS and AVHRR) or airborne scanner images, the ground area corresponding to a constant angle of view increases considerably away from the vertical. This panoramic view effect is illustrated in Figure 2.6a. In aircraft scanner imagery, which is typically acquired with a very wide scan angle, this effect produces a characteristic S-bend distortion as illustrated in Figure 2.4.

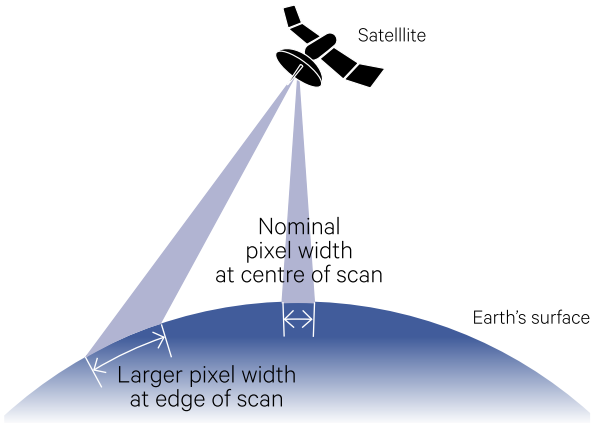
The three-dimensional nature of the Earth's surface becomes relevant to image geometry in the same way that it affects map 'accuracy'. This effect is most significant in continental scale imagery when it combines with panoramic view distortion as shown in Figure 2.6b and Figure 2.7.

Figure 2.6 Panoramic view distortion and Earth curvature

a. Panoramic View Effect—while all pixels have the same width when measured on an arc centred on scanner position, edge pixels are wider when measured along the horizontal plane



b. Earth Curvature—for a constant view angle (IFOV), pixels at the edge of the scan cover a larger ground area.

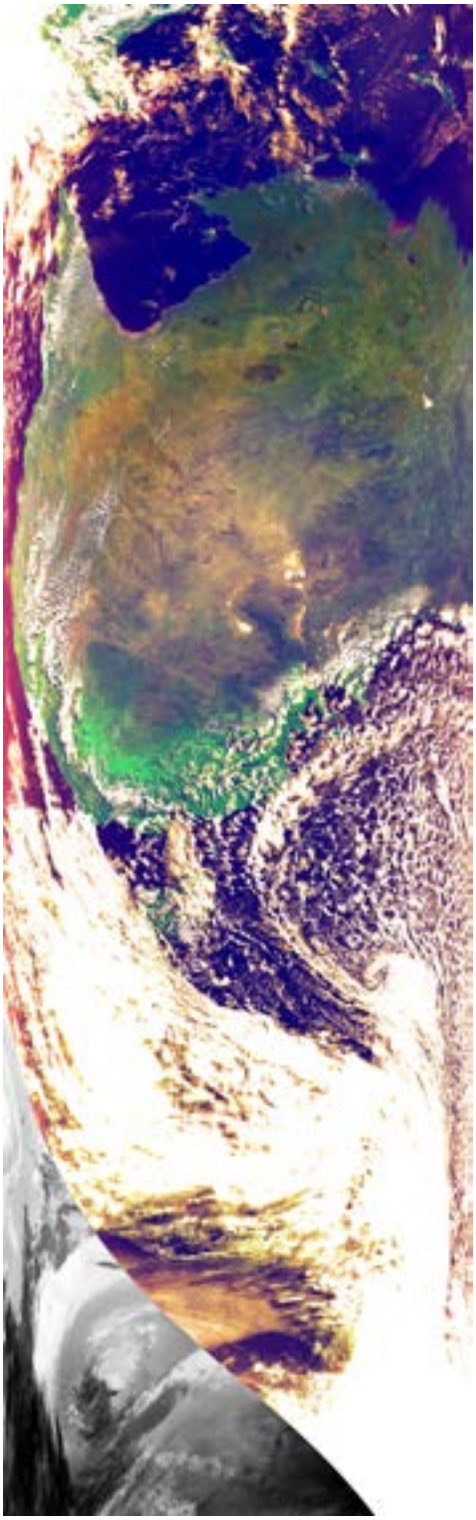


Source: Harrison and Jupp (1992) Figure 17

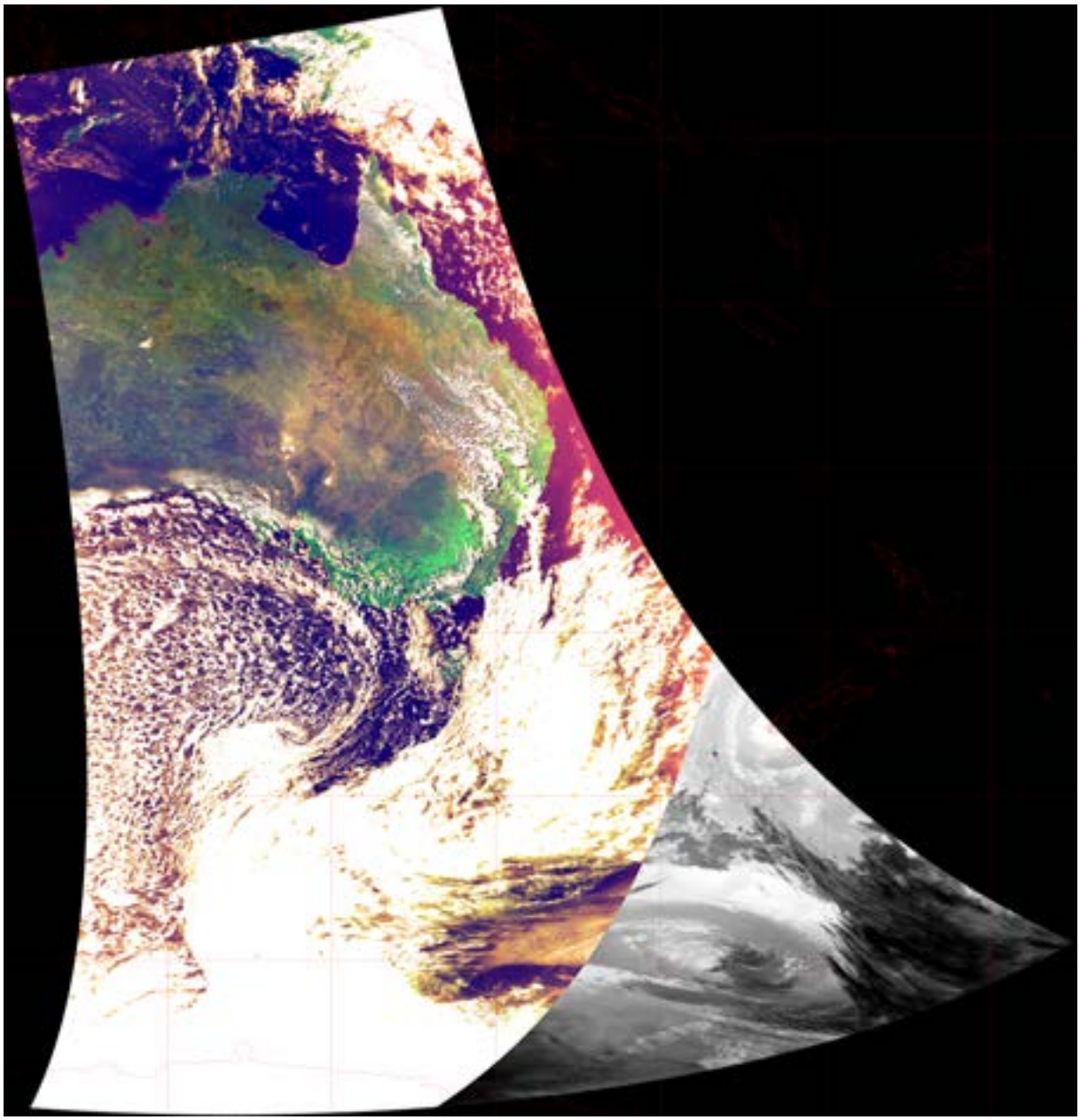
Figure 2.7 Spatial distortions in continental scale imagery

This example shows an image acquired by NOAA-19 on 3 September 2017 before and after geometric rectification. For daytime acquisitions, two reflectance channels are used to create a false colour composite (bands 2, 2, 1 as RGB). Once the Sun has set, a thermal channel is used, resulting in the grey scale section in the southeast of the image.

a. Original image swath (before flipping left to right)



b. Rectified Image swath resampled to 0.01° grid (~1 km pixel size) and aligned with geographic projection.



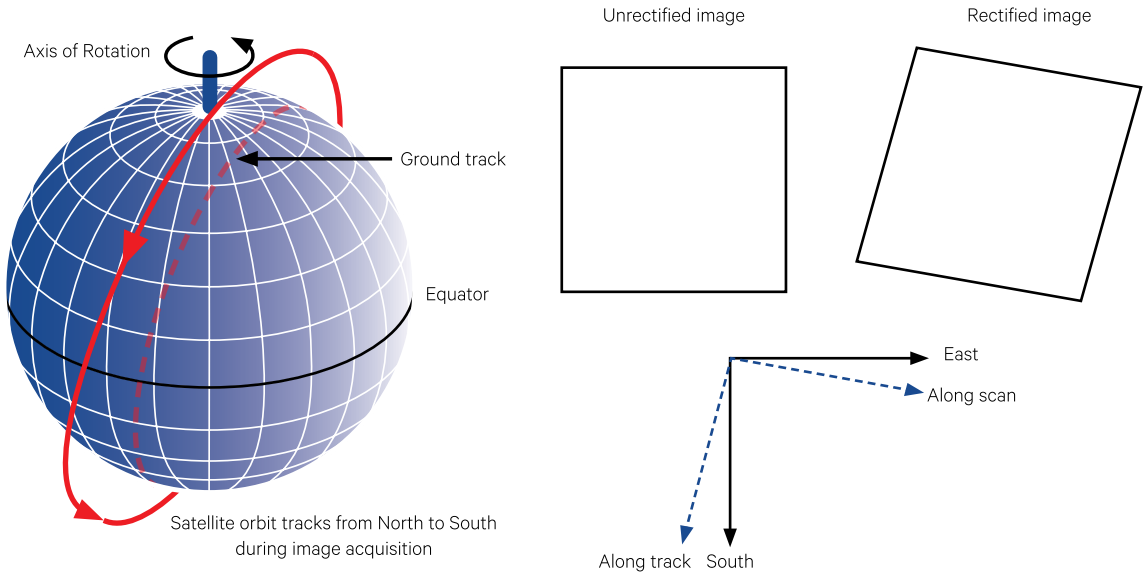
Source: Edward King, CSIRO

The rotation of the Earth as the satellite is orbiting results in a characteristic skew distortion. The direction of skew depends on the satellite direction: sensor platforms imaging from north to south (descending mode), such as Landsat, SPOT or MODIS (Terra), skew towards the west from top to bottom in the image (that is, the skew is in northwest/southeast direction—see Figure 2.8a); those imaging from south to north (ascending mode), such as MODIS (Aqua), skew towards the east from top to bottom in the image (that is, the skew is in northeast/southwest direction—see Figure 2.8b). This effect varies with latitude, being greatest at the equator and minimum near the poles (see Figure 2.9).

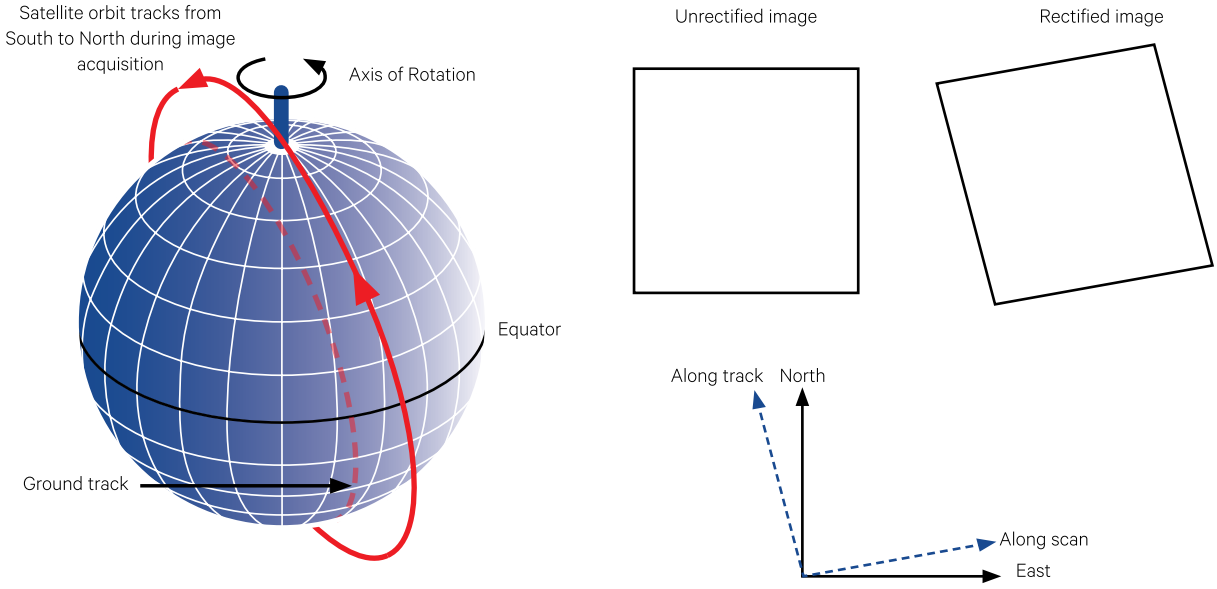
The only thing you sometimes have control over is perspective. You don't have control over your situation. But you have a choice about how you view it.
(Chris Pine)

Figure 2.8 Earth rotation skew

a. Descending Mode



b. Ascending Mode



Source: Harrison and Jupp (1992) Figure 19

While the Earth is frequently referred to as a sphere, it is actually flatter at the poles, so its shape is approximated more closely by a spheroid (see Section 1.2.1). If a satellite is orbiting at a constant orbital radius, this non-spherical shape affects the altitude of the satellite above the Earth's surface. This relative altitude change can also affect image geometry.

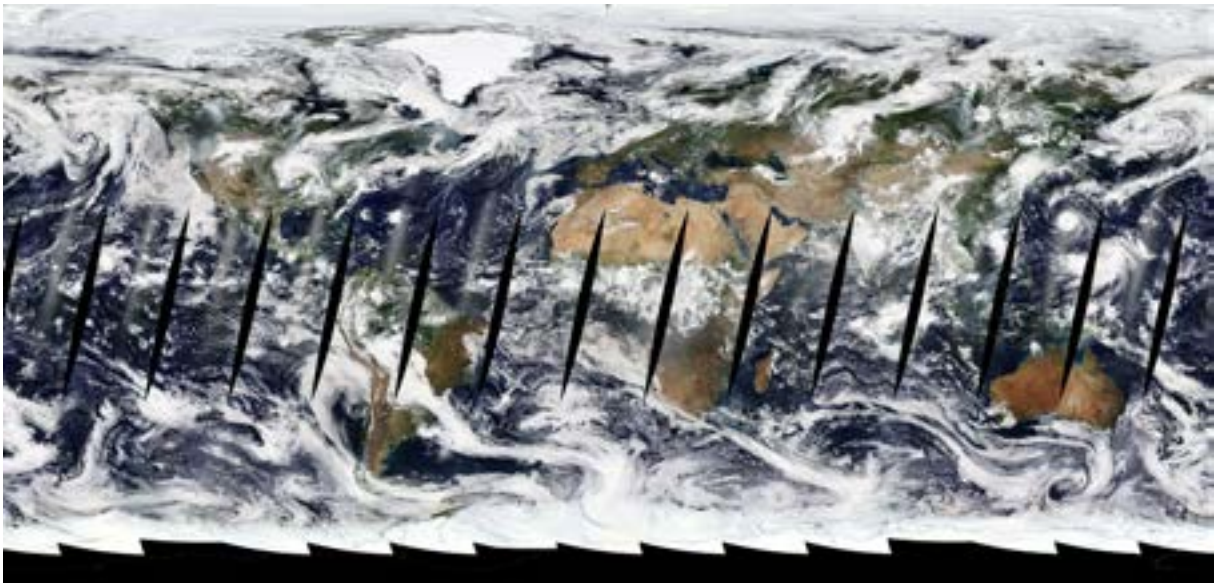
Scanning devices use the continuous forward movement of the platform during image acquisition to determine the separation between lines in an image. However, this movement also results in another skew distortion. As each line is being imaged, the platform moves forward causing the end of the line to be shifted, relative to the start of line, in the direction of travel. In general terms, the direction of this skew is perpendicular to the direction of travel of the platform. In terms of satellite data, image lines from descending mode platforms acquire a skew towards the south (Figure 2.8a), while images obtained from platforms travelling in ascending mode display a northerly skew along each line (Figure 2.8b).

The tilt of the EO device can also impact the size and scale of imaged features as illustrated in Figure 2.10. An off-nadir viewing angle is also selectable on some satellite-borne scanners (such as SPOT) to enable more frequent acquisition of ground areas from adjacent satellite overpass tracks, resulting in similar image geometric distortions.

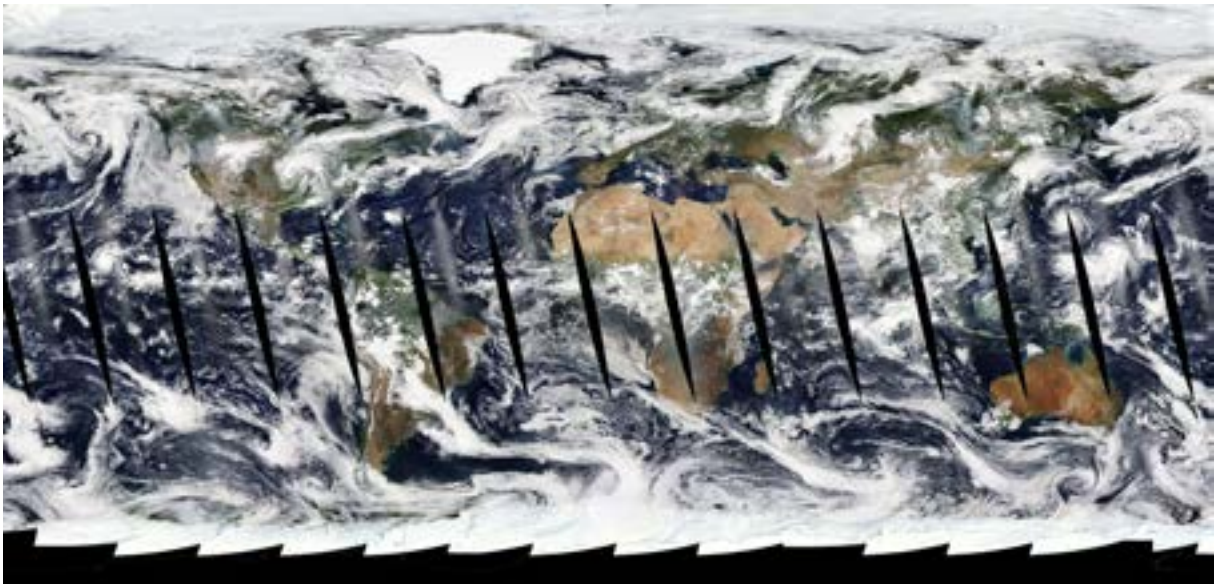
Figure 2.9 Daily acquisitions for MODIS/Aqua and MODIS/Terra

These illustrations generated by NASA Worldview show the image paths acquired on 20 August 2018 by the MODIS sensor, which is carried by both the Aqua and Terra satellites (shown at 5 km resolution).

a. MODIS/Terra collects imagery during its descending mode so image strips are skewed in northeast/southwest direction.



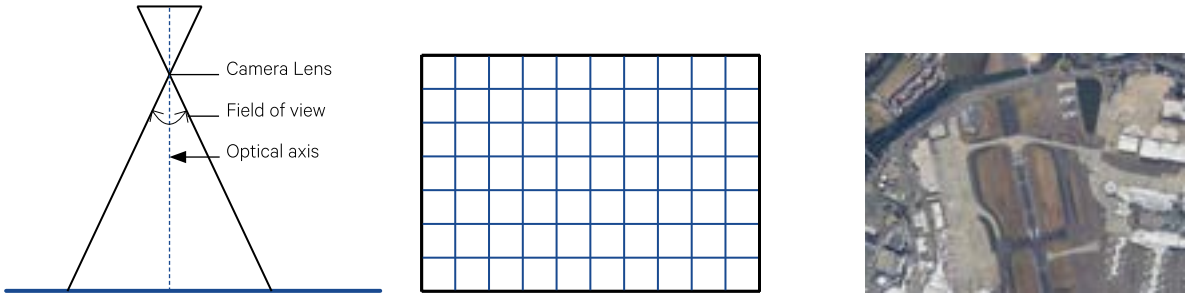
b. MODIS/Aqua collects imagery during its ascending mode so image strips are skewed in northwest/southeast direction.



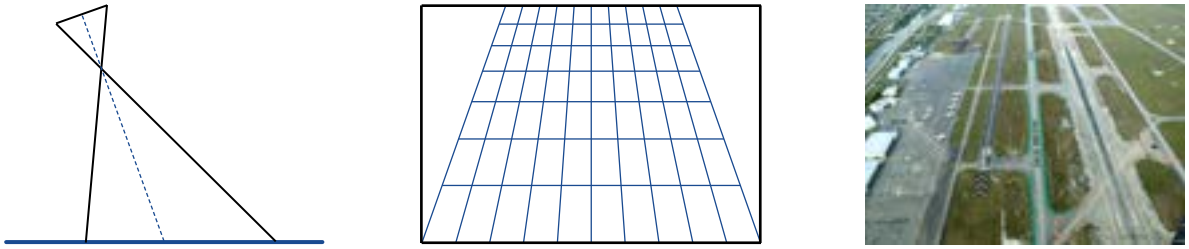
Source: NASA Worldview. Retrieved from: *https://worldview.earthdata.nasa.gov*

Figure 2.10 Effect of camera orientation in aerial photography

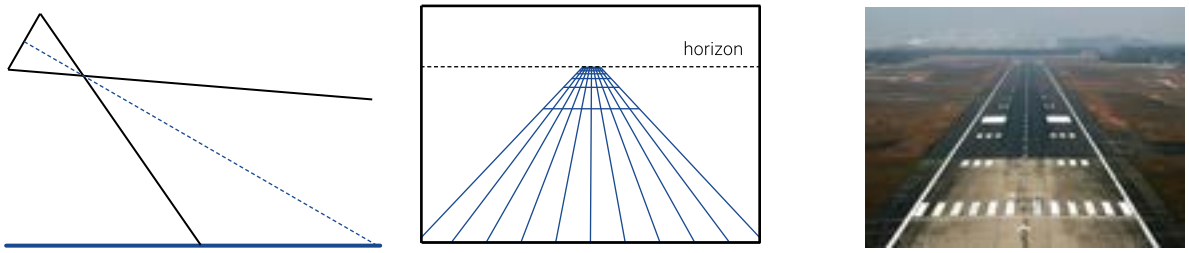
a. Vertical camera orientation yields relatively regular image grid.



b. Low oblique camera orientation produces skewed image grid.



c. High oblique camera orientation results in severely skewed image grid with horizon visible in image.



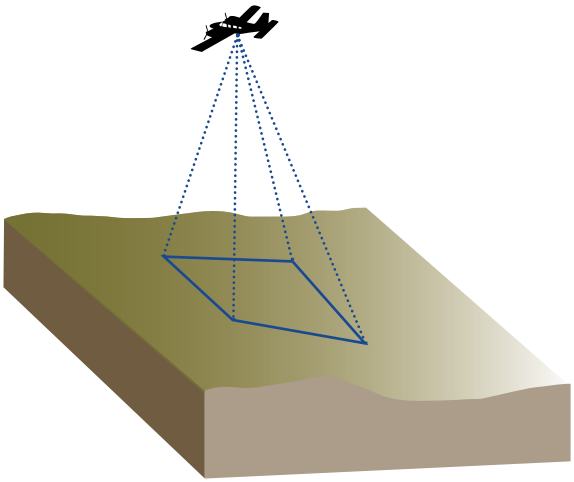
Adapted from: Wolf and Dewitt (2000)
Image sources: a. Perth Airport Runway 16R acquired on 3 May 2017, © Spookfish.
b. Palm Beach International Airport Runway 28L, © Dunn's Aerial Photography (Wellington, Florida, USA).
c. Trivandrum International Airport runway 32R photographed on 31 January 2007, © Raj Menon.

2.4 Platform Stability

Sources of distortion that are related to the stability of an imaging platform are more difficult to quantify. These may be due to variations in platform attitude (pitch, roll and yaw), altitude, velocity and heading. Considering the acquisition time for an individual Landsat image is less than half a minute, the variation of any of these parameters within an image is only likely to be significant for the most precise mapping requirements. Similarly, the altitude of satellite platforms and the narrow limits to attitude variation which are permitted before re-alignment mean that the impacts of pitch, roll or yaw of the platform are unlikely to seriously distort such images. However, the converse reasons, combined with surface relief variations, make these factors very significant for imagery acquired using aircraft platforms (see Figure 2.11).

Figure 2.11 Relief effects on aerial imagery

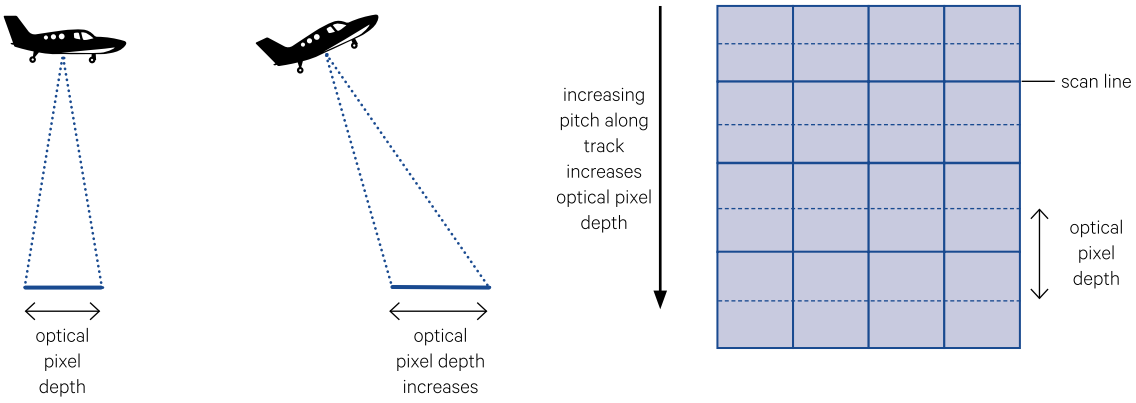
Both platform stabiity and surface relief impact the ground area imaged in each image pixel.



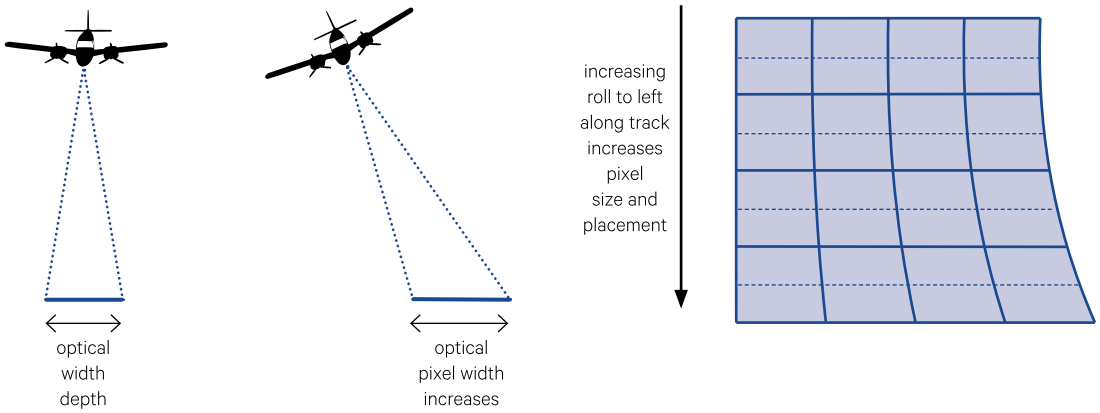
The pitch of a platform is defined as an angular movement about the scan line direction (that is, perpendicular to the direction of platform movement) as shown in Figure 2.12a. A constant pitch has the effect of displacing the complete image scan line. If the pitch is forward (that is, lowering the front of the platform), this displacement is away from the direction of travel of the platform; if the pitch is raising the front of the platform, the displacement is towards the direction of travel. A variation in pitch expands or contracts the ground distance between lines of the image depending on the direction and rate of change of the pitching movement. These effects are much more significant for imagery acquired from aircraft platforms than for spacecraft platforms.

Figure 2.12 Platform attitude variations

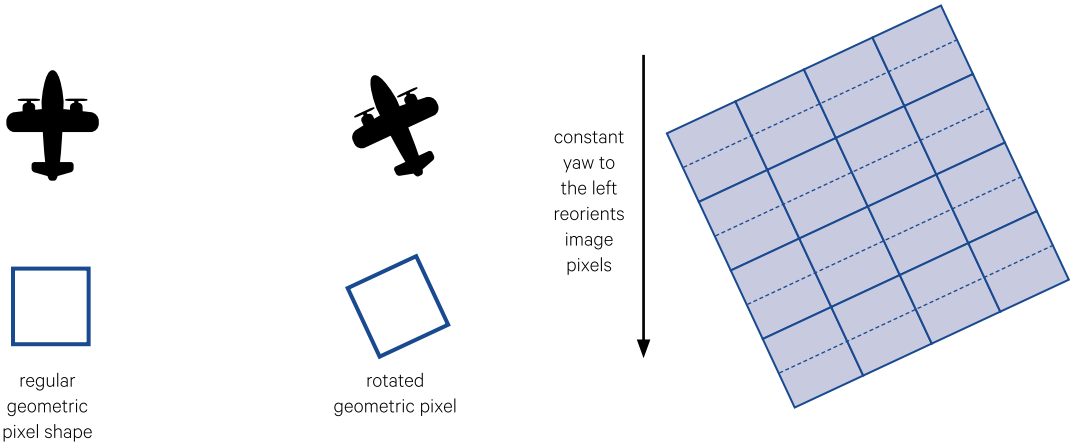
a. Pitch—variations in pitch affect pixel size and placement along track



b. Roll—variations in roll affect pixel shape and placement along scan lines



c. Yaw—variations in yaw affect pixel orientation



Source: Harrison and Jupp (1992) Figure 20. Airplane profiles: a. Shutterstock 509774308. b. Carl Davies. C. Fuerza Aerea.[2]

[2] Retrieved from: *https://www.flaticon.es/icono-gratis/avion-antiguo-con-dos-helices_84974*

Platform roll is an angular movement of the platform about the direction of travel (see Figure 2.12b). The effect of a constant roll is to shift the mid-scan line position away from nadir, increase the separation between pixels beyond this position and decrease the inter-pixel distance in the other half of the image line. Over a full image, this effect also tends to increase the separation between lines beyond the centre line position and reduce the inter-line distance in the 'nearest half' of the line. A variable roll rate skews the image in the direction of the roll, that is, the cumulative effect of line centre displacements, with their associated non-linear characteristics about this central point, result in a skewed grid of pixels on the ground. Again the effects of platform roll are most relevant to aircraft data.

The third platform attitude variable is described as yaw, that is, an angular rotation of the platform about its vertical axis. A constant yaw reorients scan lines in the direction of yaw so that they are no longer perpendicular to the direction of travel of the platform. A variation in yaw will alter the orientation of, and distance between, lines in the image as shown in Figure 2.12c.

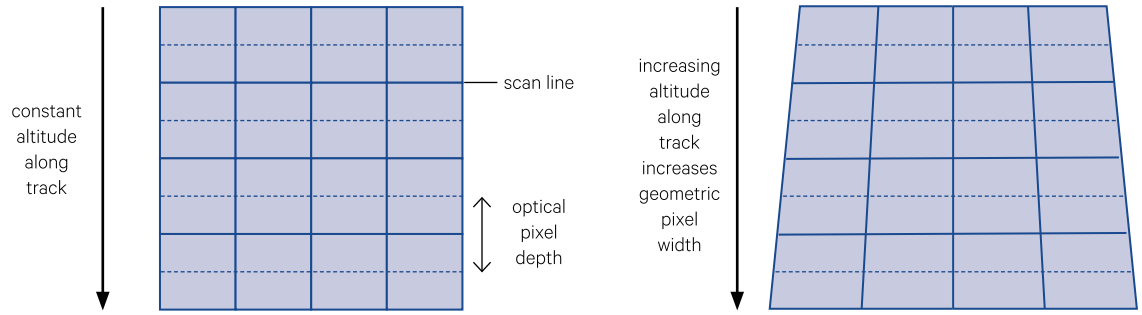
In addition to the effects of platform attitude, variations in platform altitude introduce distortions into EO imagery. The principle of this effect is illustrated in Figure 2.13. As the imaging system moves away from a target with increasing altitude, the IFOV is increased in size so the full image line covers a larger ground distance. For a constant number of pixels per line then, the geometric pixel width is increased. If the platform continues to travel at the same speed, the line spacing will not change significantly for the increased altitude so the geometric pixel depth can be considered to be constant. The overall effect of a gradual increase in altitude then is to gradually increase image width. Conversely, a decrease in altitude reduces the swath of ground being imaged. As with the attitude variations, this effect is more significant in imagery acquired by aircraft than by spacecraft (see Excursus 2.1).

Variations in the velocity of the platform affect the separation between image scan lines as shown in Figure 2.14. While the optical pixel depth remains the same, the spacing between rows of pixels in the image increases, thus increasing the geometric pixel size. This effect is also more noticeable in aircraft imagery than satellite imagery, due to the relatively tight controls on the satellite orbital parameters.

In terms of overall imaging stability, the scanning system being used by a sensor to form an image is also significant. In general, mirror scan systems are less stable than linear arrays, which are less stable that two-dimensional arrays with respect to relative orientation (RO; see Volume 1A—Section 14.2).

Figure 2.13 Platform altitude variation

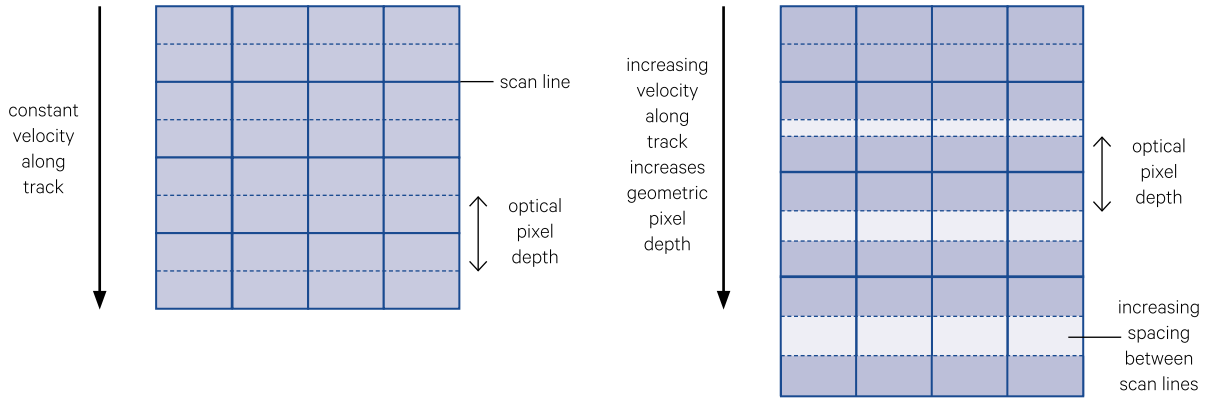
Increasing altitude along track of platform increases the image pixel width.



Source: Harrison and Jupp (1992) Figure 21

Figure 2.14 Platform velocity variation

Constant platform velocity along track results in constant pixel depth in the image, whereas increasing platform velocity along track increases the spacing between rows of optical pixels in the image, thus increasing the geometric pixel depth



Source: Harrison and Jupp (1992) Figure 22

Changes in platform heading alter the direction of the along-track skew distortion in the image. Again, such changes are largely constant within a satellite image and are relatively minor compared to the 'standard' skew due to the nominal heading.

If the attitude parameters, altitude above ground and geographic position of an imaging platform were precisely known, the external geometry would be known and a mapping function relating image pixels to location on the Earth's surface could be directly defined. While models for satellite positioning are increasingly accurate, that accuracy is not yet as high as orthorectification of imagery with precise ground positioning.

To accurately rectify an image containing these distortions (or to register an image to another coordinate system) control points are generally used (see Section 4). These are points which may be precisely located in the image and on another spatially 'accurate' data set, such as a map, and for which coordinates in the two geometries may be determined. The control points are then treated as a spatial sample from the two geometries and may be used to model the relationship between them in the same way that traditional photogrammetry uses tie points to mosaic aerial photographs. Most of the platform-related factors discussed above can be adequately modelled in this way, using low-order polynomial models for satellite imagery and possibly higher-order models for aircraft data. GNSS (Global Navigation Satellite Systems) and star-trackers are now able to supply some previously unavailable orbital information so that the use of control points in image rectification can be minimised. Additionally, image processing models for rectifying image data can remove the effects of quantifiable distortions in satellite and airborne scanner images such as sensor non-linear scan rate, Earth rotation and curvature, and panoramic distortion (see Section 3.1.2).

Rectified EO imagery will not always align well with various map projections, especially at smaller scales. This will largely depend on the projections inherent to both map and image. As discussed in Section 1.2.2, maps always involve some distortion so should not be treated as the 'geometric truth'. EO images, once rectified, present an accurate view of the Earth's surface from a particular perspective. It is not relevant to consider this perspective in terms of geometric accuracy relative to a particular map any more than the accuracy of one map projection is judged in terms of another, since all two-dimensional representations of the Earth's surface contain some geometric distortions. The differences are not so much a matter of accuracy as convenience with respect to accepted mapping standards but are significant when different data sources need to be registered.

Excursus 2.1—Generation of CASI-2 Image Mosaic

Source: Karen Joyce, James Cook University, and Peter Scarth, University of Queensland
Further Information: Joyce, K. (2004).

Airborne scanner imagery is often acquired via multiple flight lines, with each image strip potentially encountering different atmospheric conditions, which impact both the geometry and brightness of the acquired imagery. To create a seamless mosaic from such imagery requires careful geometric and radiometric calibration (see Volume 2A—Section 3.3).

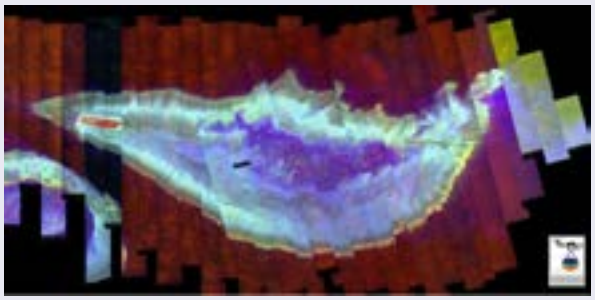
An example of CASI-2 (Compact Airborne Spectrographic Imager) imagery over Heron Island, Queensland is shown in Figure 2.15. These images were acquired in 45 flight lines over two flying days (1 June and 3 June 2002) with 1 m ground resolution. Heron Reef, in the southern Great Barrier Reef, spans approximately 11 km by 4 km, so requires high resolution imagery to achieve detailed coverage.

Airborne imagery can exhibit significant geometric distortion due to aircraft movement during image acquisition. The extent of this distortion can be seen in Figure 2.16 for a portion of one image strip.

To create a seamless mosaic, all image strips were processed to remove cross-track illumination effects (or limb brightening) and corrected for atmospheric and air-sea interface conditions at the tiime of acquisition (Brando and Dekker, 2003). Geometic correction to remove both systematic and non-systematic spatial distortions was based on in-flight inertial measurement unit data. The final mosaic was referenced to UTM WGS84. The corrected image mosaic, after both geometric and radiometric calibration, is shown in Figure 2.17.

Figure 2.15 Original image data showing individual flight lines

CASI imagery over Heron Island, Queensland, displayed as natural colour composite (bands 643.7 nm, 522.0 nm, 459.6 nm as RGB). The western portion was acquired on 1 June 2002 and the eastern portion was acquired on 3 June 2002.



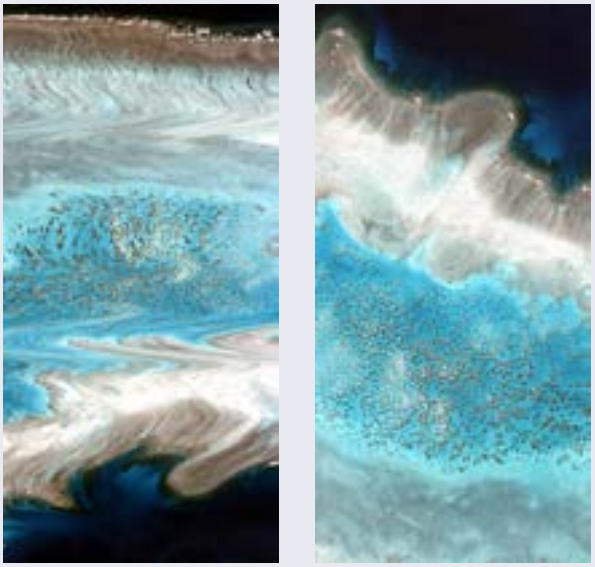
Source: Karen Joyce, James Cook University

Figure 2.16 Geometric correction of CASI imagery

Imagery acquired by airborne scanners can exhibit significant geometric distortion due to aircraft movement during image acquisition (see Figure 2.15 for image details). In this case the original image also needed to be flipped top to bottom during the correction process (see Volume 2A—Section 7.2.2).

a. Original image

b. Corrected image



Source: Karen Joyce, James Cook University

Figure 2.17 Corrected image mosaic

All original images were calibrated to surface reflectance, with cross track and deglint corrections applied. (Note: deglinting results in exposed features, such as the island, being obscured.) In this case, the image geometry has been rectified to a consistent grid comprising 1 m ground resolution elements.

Source: Karen Joyce, James Cook University

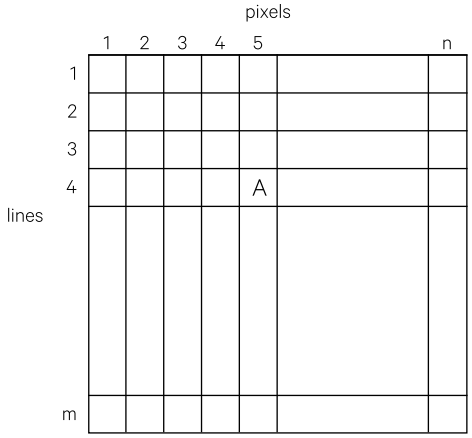
2.5 Image Grid Origin

Many sources of EO satellite data are scanned with an origin in the top-left (TL) of the image when it is oriented to north (see Figure 1.5). This relates to an image formed by a satellite travelling from north to south (descending mode) and scanning from west to east. Satellite scanners that form images this way include Landsat (MSS, TM, ETM+, OLI and TIRS), SPOT and MOS-1. Imagery that is scanned during the south to north travel of an imaging platform (ascending mode) will have a natural origin at the bottom of an image. For example, NOAA AVHRR daytime images are formed during the satellite's ascending mode and are scanned from west to east. This gives a natural origin of bottom-left (BL). NIMBUS CZCS also operated in an ascending mode and scanned from east to west to place the origin at bottom-right (BR).

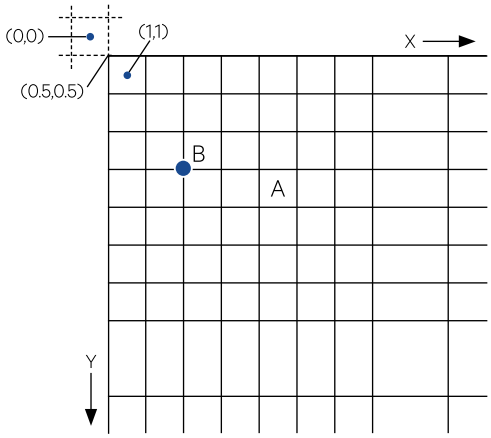
In the TL origin system, image coordinates of a particular pixel are given as the number of pixel columns to the right of the left-most edge of the image and the number of lines down from the top of the image. The centre of the first pixel on the first line of the image is considered to be (1,1) so that the actual origin (0,0) for this system is outside the image as illustrated in Figure 2.18. The location of the centre of pixel A in this system is (5,4). Fractional coordinates are also possible. For example, point B is referenced as (2.5,3.5).

Figure 2.18 Image coordinate system

a. Image coordinates are relative to an origin near the top left corner of the image, such that the precise location of a coordinate is the centre of its geometric pixel.



b. The centre of the top left pixel in an image is referenced as (1,1) relative to an origin (0,0) outside the image. Thus, the centre of pixel A is (5,4) and the location of point B is (2.5, 3.5).



Source: Harrison and Jupp (1992) Figure 23

2.6 Image Registration

Each pixel in an EO image corresponds to a measurement that is integrated over a 'volume' on the surface being imaged. The relationship between the geometry of the resulting image and the geometry of the scene being imaged may be described by modelling the optical and physical relationship between the scene and the sensor. Generally the scene geometry is represented by an independent mapping exercise so that the image is viewed as being registered with the scene when it is registered to a map of the scene (see Figure 1.1). To overlay the image onto the map it must be resampled to the same projection (see Section 6.1).

While image pixels are represented in image processing systems as a regular grid pattern, the distortions that occur in EO imagery mean that the pixels do not represent areas of equal size on the ground. To account for these distortions, the image rectification models often convert image pixel locations to Cartesian coordinates on a metre grid relative to a specified origin in the image (see Section 1.1).

To correct for geometric distortions and allow geo-referencing, EO image geometry has been traditionally re-projected to match a selected map coordinate system. While this improves image utility, it does not necessarily permit different image sources to be cross-referenced simply, especially for regional and global studies.

To address these limitations, a geographic grid, based directly on a graticule of latitude and longitude, can be used. This representation is also (erroneously) referred to as an un-projected grid as it is not, strictly speaking, a map projection (since it is neither equidistant, conformal or equivalent; see Excursus 1.1). In such grids, cells are defined in terms of degrees of latitude and longitude and, therefore, vary in size across an image, being larger near the equator and smaller near the poles. Simple projections, such as the Plate Carrée, are effectively based on a geographic grid (see Figure 1.17).

To improve comparison between EO data from different sources, an increasing number of EO products share a common underlying nested grid base with scalable cell sizes (see Volume 2A—Section 7.4). For compatibility with the major sources of EO imagery, the range of cell sizes includes 25 m, 250 m and 1000 m. For example, Landsat TM/ETM+ imagery, which was original acquired and supplied at 30 m pixel size, is now supplied using a 25 m grid cell. Nested grids allow EO data to become 'sensor agnostic' and readily comparable with spatial data from other sources. This feature is particularly relevant given the growing time series of EO imagery and will significantly simplify the process of monitoring long term changes in land and water resources.

The generation of 'actionable' data, that is, the automatic identification and extraction of significant features and patterns, from very high resolution, orthographic aerial imagery is introduced in Excursus 2.2.

Excursus 2.2—Geospatial Imagery Redefined

Source: Rohan Fernando, Spookfish
Further Information: *www.spookfish.com*

Actionable Data

It is extremely important that EO imaging technologies deliver data with quality, accuracy, currency, integrity, scalability, affordability and reliability. Research and development engineers at Spookfish have pushed past many commonly known technical limitations of camera sensors, optics, motion sensors and dynamic control systems, and photogrammetric mathematical transformations and image processing algorithms to develop leading edge aerial imaging technologies and image processing methodologies. Their image products are meeting the rapidly rising demand for high quality, cost-effective imagery with large ground coverage in both 2D and 3D. They are also enabling 'actionable' information.

The ability to provide actionable information is at the current frontier of imagery capture and processing technology. This is fundamentally about automatically identifying and extracting important visual features and patterns from imagery at very large scales. Automated artificial intelligence / machine learning (AI/ML) computational systems applied to very high quality image data are beginning to generate actionable data. Over time, it is anticipated that many new applications of imagery will be uncovered through the use of AI/ML (see Figure 2.19).

Image Capture and Processing

A wide range of aerial imagery technologies and methodologies are used operationally, or in research and development mode, including:

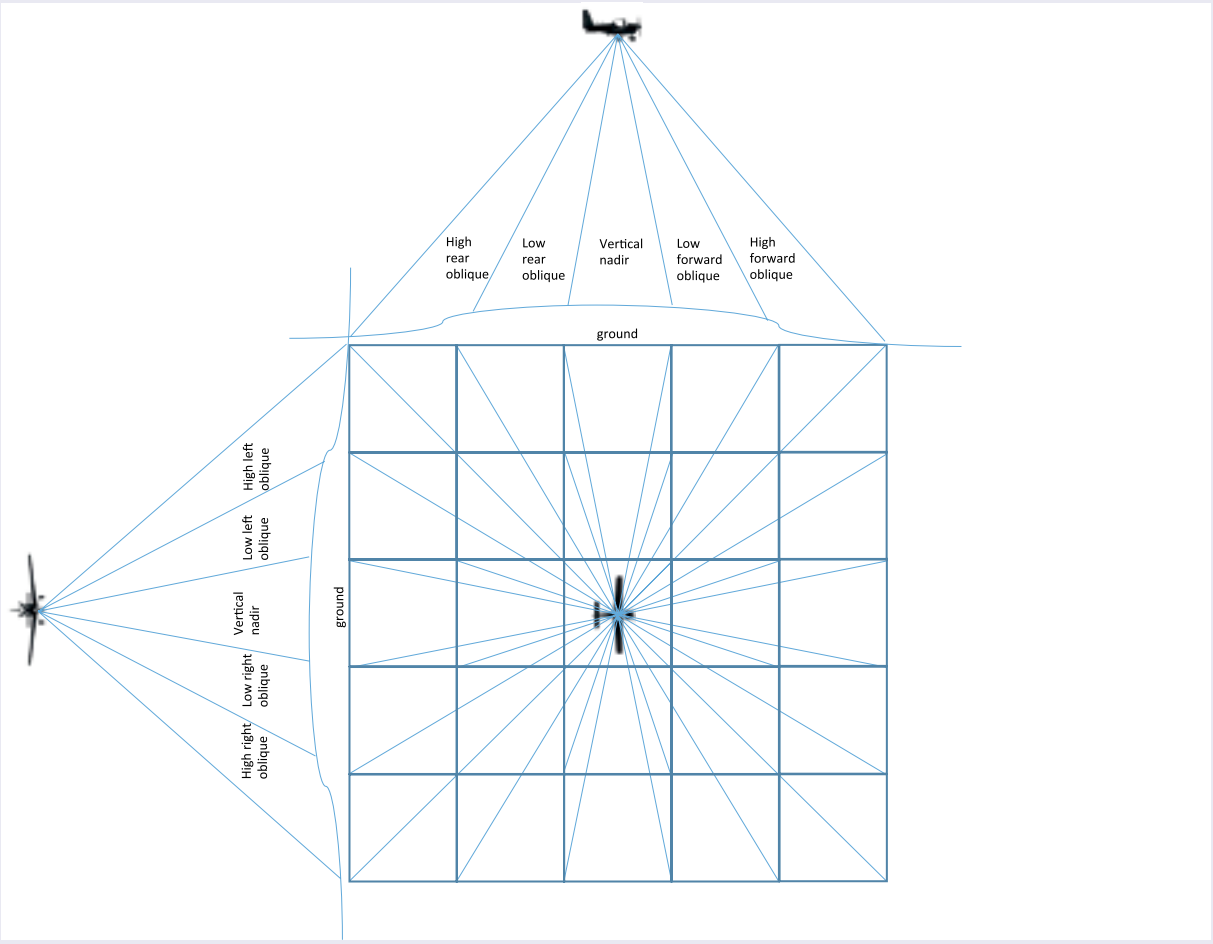
- multi-angle viewing for image capture (see Figure 2.20):
  - vertical nadir, low oblique and high oblique;
  - from left to right, and fore to aft, relative to flight direction; and
  - overlapped image frames;
- modular and transportable imagery capture pods that can be mounted underneath different classes of light aircraft (see Figure 2.21);
- precisely overlapped flight patterns for construction of:
  - orthographic images with precise spatial accuracy;
  - interpolated 3D images;
  - digital terrain models; and
  - digital surface models;
- complex camera sensor arrays and optical lens technologies;
- sensors for optical (visible and NIR) wavelengths with high dynamic range;
- GPS position and inertial measurement sensors; and
- automated AI/ML image feature recognition and extraction.

Figure 2.19 Actionable data

An AI/ML system automatically identifies roof tops, pools, trees and shrubs in Spookfish imagery.

Figure 2.20 Multi-angle viewing

Imagery can be acquired by viewing the ground at multiple angles, both along and across the flight track.



Some of the advantages of Spookfish aerial imagery capture and advanced analytical AI/ML information processing include:

- super-high resolution imagery, with a standard 5 cm ground sample distance, and ongoing developments to improve image resolution in the future (see Figure 2.22);
- image spatial accuracy to within +/-20 cm of true ground location (latitude/longitude), without the need for any ground survey markers;
- large image swath and progressively-assembled geographic extent of the imagery, without sacrificing the overall image spatial resolution;
- rapid image acquisition, enabling large areas to be captured efficiently, with flexible revisit times;
- exceptionally high data quality, which allows numerous AI/ML image processing functions to be performed to automatically identify and extract useful feature information;
- in-house processing with massively-scalable data centre infrastructure, delivering ready access via industry-standard, secure internet connections; and
- low overall operational costs, providing individuals or organisations affordable access to high quality imagery.

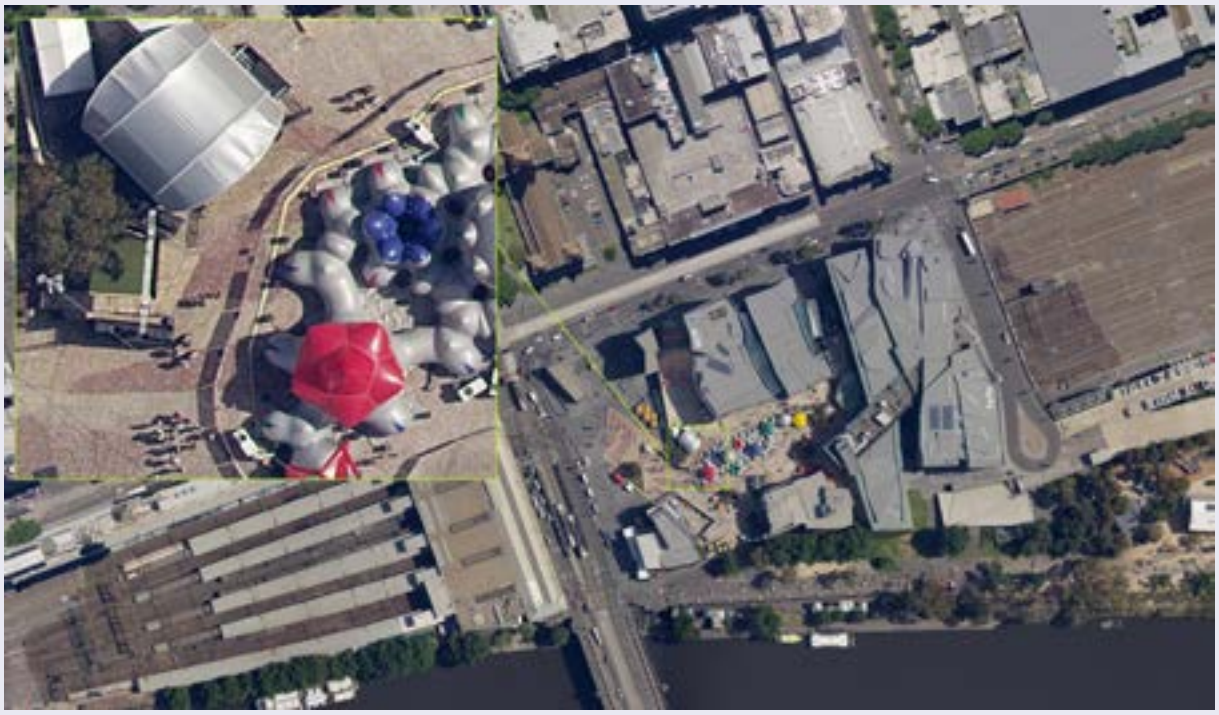
Figure 2.21 Camera system on aerial platform

Aerial image camera system is mounted underneath a Piper Aztec light aircraft.

a. Side view

b. Front view

Figure 2.22 Spookfish high resolution image of Federation Square, Melbourne



Overlapped, Multi-Angular Photogrammetry

Spookfish employs a complex process to assemble a high overlap, multi-angular photogrammetric image mosaic of the ground surface. This process is based on flying at high altitude, capturing imagery with extremely high resolution and spatial accuracy, and requires sophisticated timing and controls to manage the required flight paths and correctly assemble the continuous high flow of imagery data. Over the past few years the aerial capture and processing systems have been continually researched and developed to become what is possibly the most technically advanced system in the world. This system is not only used extensively across all the major cities and towns in Australia, but is also supplied as a modular platform under license to one of the largest aerial imagery companies in the USA, Eagle View Technologies, where it is used for similar private, commercial, and government applications.

There are several important issues associated with capturing aerial imagery that must be factored into every capture effort, including daily weather conditions, cloud cover, humidity and air pollution haze, time of day, seasonal Sun angles, and ground-based shadows. There are also additional factors that typically need to be considered such as approved flight plans, controlled air space, and air traffic controls.

A complex flight planning process is associated with each area where imagery is to be captured. Before and during the aerial imagery capture flight, the pilot is provided with a series of flight tracks graphically overlaid on an industry-standard aviation chart. The pilot uses this information to provide continuous navigation guidance of the aircraft and confirmation that the correct ground region is being captured at all times. During image acquisition, in addition to adjusting their work for various changing operational factors, such as weather conditions, pilots also need to strictly adhere to instructions from Air Traffic Control at all times.

Applications

This pioneering technology enables rapid imaging of vast areas in high resolution from a multitude of angles at a fraction of the cost of contemporary systems. As the costs and availability of high-quality aerial imagery data improve over time, the number of possible applications that can provide a useful benefit to society will continue to grow. All such applications are primarily focused on enabling users to make better decisions, drive more significant efficiency gains, and reduce operational costs. These capabilities will make it easy for a wide range of private and public organisations, of all sizes, to gain access to premium imagery content and pervasive 3D models, allowing concise, accurate and cost-effective decision-making.

2.7 Further Information

EO Image Distortions

PennState College of Earth and Mineral Sciences—Optical Sensors: *https://www.e-education.psu.edu/geog480/node/444*

Natural Resources Canada—Geometric Distortions: *http://www.nrcan.gc.ca/node/9401*

2.8 References

Brando, V. E., and Dekker, A. G. (2003). Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. *IEEE Transactions on Geoscience and Remote Sensing,* 41 (6), 1378–1387.

Colwell, R.N. (1983). *Manual of Remote Sensing*. 2nd edn. American Society of Photogrammetry and Remote Sensing. Falls Church, Virginia.

Harrison, B.A., and Jupp, D.L.B. (1989) *Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual*. CSIRO, Melbourne. 156pp.

Harrison, B.A., and Jupp, D.L.B. (1992). *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual*. MPA, Melbourne.

Joyce, K.E. (2004). *A Method for Mapping Live Coral Cover using Remote Sensing*. Ph.D Thesis, University of Queensland.

Slater, P.N. (1980). *Remote Sensing: Optics and Optical Systems*. Addison-Wesley, Reading, Mass.

Wolf, P.R., and Dewitt, B.A. (2000). *Elements of Photogrammetry: With Applications in GIS*. McGraw-Hill. 608pp.

Rectification

Image rectification involves modelling the geometry and topology of an image. Image registration models the relationship between the geometry of an image and the geometry of some other spatial representation of the imaged scene. Image resampling uses the modelled relationship between two geometries to rewrite one geometry to match the other, for example, to overlay an image onto another image or a map. This process changes the spatial relationship between objects in the image, and may be compared to selectively stretching sections of the image as if it were a rubber sheet.

Rectification is the process of accurately locating observations on the ground... Improved rectification methods are essential to make use of emerging and future EOS data sources and to deliver value to the Australian government and the community. (Lewis et al., 2011)

As outlined in Section 2, EO imagery contains a variety of geometric distortions that need to be accounted for before registering with different image sources or map data. Image processing software can implement a range of models to account for known geometric distortions in satellite and aircraft scanner imagery as well as map projection models, which compensate for the spatial inconsistencies between different map coordinate systems (see Section 3.1). These models can be used in conjunction with mathematical models derived from control points to achieve accurate registration between different coordinate systems (see Sections 3.2 and 3.3).

Contents

3 Transformation Models 53

4 Control Point Modelling 71

Background image on previous page: MODIS image of northeast Australia, acquired on 15 February 2009, while northern rivers were in flood and bushfires raged in southeastern Australia. **Source:** Norman Mueller, Geoscience Australia

3 Transformation Models

Mathematical models can be developed to allow locations on the Earth's surface to be transferred from one coordinate system to another, such as from:

- one map to another map;
- one image to another image;
- a map to an image; and
- an image to a map.

Nominal models can be defined to convert between well-defined map projections (see Section 3.1.1). Such models can also be developed to account for some of the known geometric distortions in EO imagery (see Section 3.1.2).

Alternatively, if the geometries of two different coordinate systems can be represented by a set of locations, a mathematical model can be derived to convert coordinates from one system to the other. In some image processing systems, such models are used in conjunction with nominal map and image models to account for arbitrary geometric effects in EO images and register two independent coordinate systems. These models are described in Section 3.2.

A common approach to image registration allows multiple nominal models to be used in conjunction with an arbitrary model (see Section 3.3). For example, when registering image and map datasets, the known image distortions can be adjusted using a nominal image model and map data can be reprojected onto the image acquisition path. Once the known differences between the image and map 'projections' have been 'aligned', an arbitrary geometric model can be determined to convert locations to and from the image and map. This three-stage approach has been shown to deliver robust and accurate results and is recommended for rectification and registration of EO datasets.

Equations are just the boring part of mathematics. I attempt to see things in terms of geometry. (Stephen Hawking)

3.1 Nominal Models

Nominal, or parametric, models can be used to describe the process of converting between coordinate systems on the basis of known distortions. These models exist for both map and image coordinate systems, such as:

- image (pixels) ↔ satellite or scanner model (metres);
- geographicals (decimal degrees) ↔ UTM (yards or metres);
- geographicals (decimal degrees) ↔ Local Transverse Mercator (LTM; metres);
- geographicals (decimal degrees) ↔ Hotine Oblique Mercator (HOM; metres);
- UTM (metres) ↔ UTM (yards);
- UTM (zone A) ↔ UTM (zone B);
- UTM (metres) ↔ LTM (metres);
- UTM (metres) ↔ HOM (metres); or
- LTM (metres) ↔ HOM (metres).

During image rectification, typically image points would be converted to satellite model coordinates and map points would be converted to either LTM or HOM projections.

3.1.1 Map

As introduced in Section 1, map projections attempt to represent regions on the Earth's surface as a two-dimensional coordinate system. If the mapping function is well specified, surface positions can be mathematically represented in any projection. Similarly, conversion from a projection to a surface position allows coordinates to be converted between different projections.

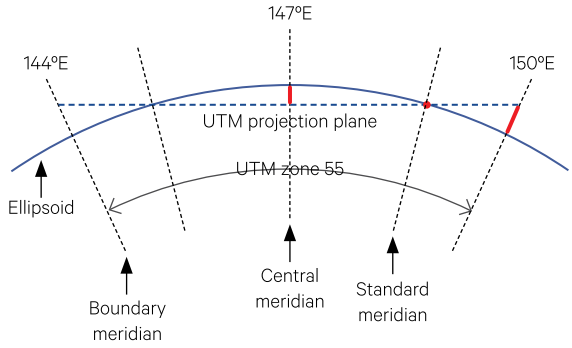
Background image: SPOT-1 image of Great Sandy Desert, central Australia, acquired in 1986. **Source:** © CNES 1986, Distribution Airbus DS. Retrieved from: http://www.intelligence-airbusds.com/en/5751-image-detail.php?img=1505#.Wq9UwmO8z8s

3.1.1.1 Local Transverse Mercator (LTM)

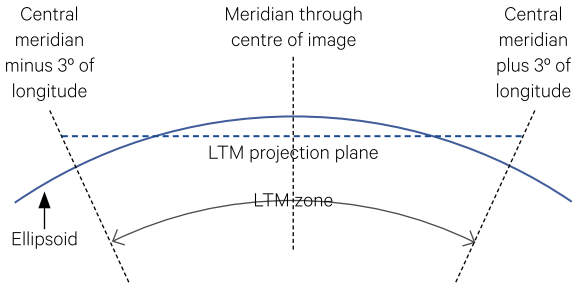
For a limited area on the Earth's surface, a reasonable two-dimensional representation of the area's geometry can be achieved using a suitable projection. The Universal Transverse Mercator (UTM) grid provides accurate representation of small ground areas, especially at the centre of a zone (see Section 1.3.2). When a study area is contained within a single UTM zone, and is located close to the centre of that zone, its map coordinates could be considered as a regular grid over the ground surface. Non-standard UTM zones, with a central meridian matching the image centre, can be used to simplify rectification of images that straddle standard UTM zones. These non-standard zones are sometimes referred to as Local Transverse Mercator (LTM; see Figure 3.1). Since LTM zones are usually centred on the image centre, zone edge distortions are also minimised (see Section 1.3.2). Even when the image is contained within a single UTM zone, projection distortions (convergence away from the conformal meridian) are minimised when the locations are converted to a coordinate system that is centred on the study site. All coordinates in the LTM projection are given in units of metres from the central origin.

Figure 3.1 Local Transverse Mercator projection

a. Cross-section of UTM grid zone 55, with scale factor (k) shown in red—at the central meridian, k=0.9996, at the standard meridian, k=1.0, at the zone boundary meridian, k=1.00098.



b. Cross-section of LTM grid zone, centred on an EO image, with zone boundaries 3° east and 3° west from the central meridian.



Adapted from: Stanaway (2011) Figures 1 and 3

Landsat scenes, for example, cover an area approximately 185 km square so can be easily contained within one UTM zone. Where a scene covers a zone boundary, UTM coordinates can be converted to Local Transverse Mercator (LTM) to ensure map coordinates are relative to a single origin. LTM defines a UTM-type zone which can be centred on any longitude. If the image scene centre is used as the central longitude, then some images (such as Landsat or SPOT) can be fully contained within one LTM zone so all map coordinates will be relative to the same origin. Even when the image is contained within a single UTM zone, projection distortions are minimised by converting the locations to a coordinate system that is centred on the image.

3.1.1.2 Hotine Oblique Mercator (HOM)

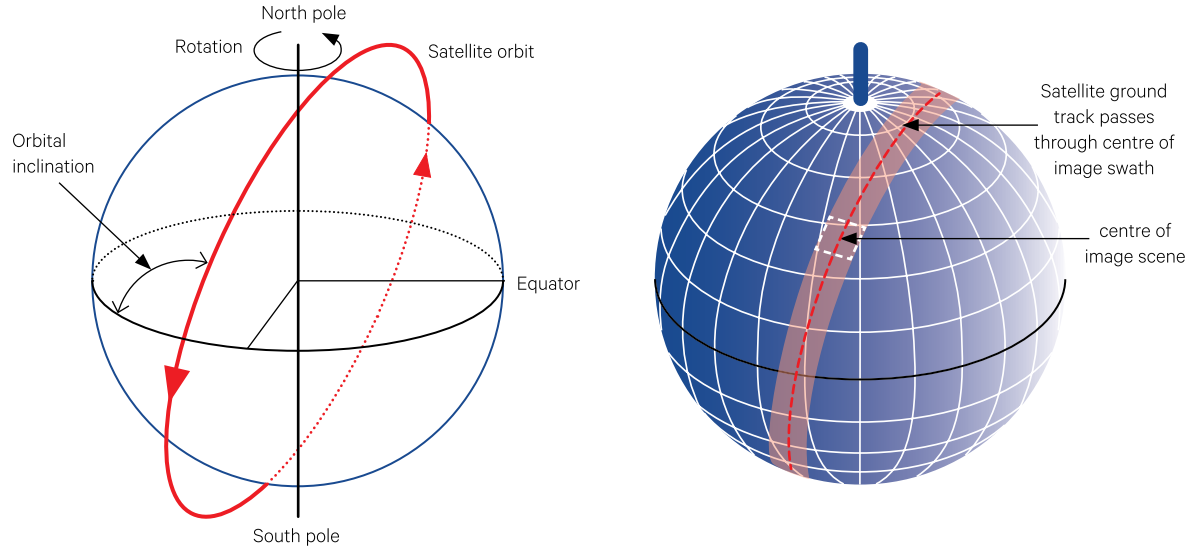
The perspective of the EO image needs to be considered when rectifying continental scale imagery such as MODIS and AVHRR. The satellite orbital path can be viewed as an Oblique Mercator Projection, with the orbit itself being considered as an expanded great circle encompassing the Earth. Full scenes of these large area images cover several UTM zones, so UTM coordinates cannot be used for image rectification. The distortions inherent to a map projection are particularly important for large areas. An image that is effectively produced as an Oblique Mercator Projection aligns most closely with an Oblique Mercator map. This is also relevant when using such images in conjunction with map overlays after rectification.

The Hotine Oblique Mercator projection (HOM) can be used during the rectification process of continental scale imagery. This cylindrical projection is conformal with the globe along a selected great circle (see Figure 3.2). The standard Mercator projection uses the equator as a conformal line and 'spreads out' the globe from that line. A Transverse Mercator Projection uses a meridian of longitude as the conformal line, whereas the HOM projection is conformal along a particular great circle (see Section 1.2.2). With satellite imagery, the satellite orbital path can be considered as a great circle and used as the conformal line for the projection. The great circle is defined by the angle at which the circle crosses the equator (the orbital inclination of the satellite) and the geographical coordinates of one point on the circle (the scene centre of the image). As is the case for LTM, a metre unit coordinate system is then defined with the great circle as the central Y-axis and the line perpendicular to this through the scene centre as the X-axis as illustrated in Figure 3.2.

Figure 3.2 Hotine Oblique Mercator projection

a. Satellite orbital configuration

b. The satellite ground track forms the centre of image scenes along this path.



Source: Harrison and Jupp (1992) Figure 13

Thus, the HOM projection allows map coordinates to be converted to a grid system that is relatively regular over a large study area and oriented in the same direction as the image swath. This projection is determined by the image scene centre and its orbital inclination, and measured in metres. When rectifying continental scale imagery, map control points should be read as latitude and longitude with nominal conversion to HOM to remove the rotation factor before fitting an arbitrary mathematical model (see Section 3.2).

3.1.2 Image

The systematic or 'quantifiable' distortions inherent in satellite imagery may be corrected using nominal image rectification models which account for geometric errors due to panoramic distortion, Earth rotation and Earth shape, or sensor characteristics such as non-linear scan rate and sensor delay offset (see Section 2). Such models typically rely on orbital and imaging parameters that are supplied with satellite image data, such as the latitude of the scene centre (used to determine satellite altitude and heading), the pixel ground dimensions and the number of pixels per scan line (for the full image scene).

Aircraft scanner models primarily account for pixel aspect ratio (that is, the difference between optical and geometric pixel sizes due to oversampling effects) and panoramic distortion. Input data required typically include aircraft velocity and altitude, and scan angle.

These satellite and scanner image models convert between image coordinates and model coordinates. The model coordinates represent pixel locations in metres from a specified origin in the image. The parameters for nominal satellite models are defined by specific procedures, which can be described in terms of three logical stages:

1. Compute nominal parameters:

- calculate the platform altitude for a given latitude using nominal orbit parameters for the satellite platform and an appropriate geoid model;
- compute nominal pixel width for given number of pixels per scan line using relevant parameters (scan time, scan angle, orbit radius) and an appropriate along-line correction model such as:
  - non-linear scan rate for Landsat MSS;
  - Earth curvature correction for AVHRR;
  - panoramic view distortion (all sensors);
- compute satellite heading using orbital inclination with geocentric latitude (calculated from given latitude and geoid model);
- compute nominal pixel depth using nominal angular velocity and sensors per swath parameters and platform heading; and
- compute Earth rotation skew per image line from platform heading, rotation factor, scan time and half scan angle.

2. Adjust for given parameters:

If pixel width and depth are supplied or can be estimated (see Excursus 3.1), these can be used to compute the actual parameters for platform movement.

- adjust nominal altitude computed in stage 1 by varying the pixel width; and
- adjust satellite heading by pixel depth and other parameters used to compute nominal value.

To verify model parameters, the expected pixel depth can be computed with the adjusted parameters. A model compatibility factor (MCF) can also be determined at this stage to check whether any difference between the actual and nominal pixel dimensions would be explained by a change in the platform's altitude. This factor is a useful way of checking that the specified parameters, such as the number of pixels per scan line, are within a valid range for a particular data source.

3. Apply the model:

- use models for along-line distortions to compute the X direction distance (in metres) of a given pixel location from a given origin, then adjust for the sensor delay offset and Earth rotation using the given line position in the image; and
- compute distance (in metres) of a given pixel location from a given origin in Y direction using the pixel depth.

Many nominal models are invertible so can also be used to compute the image pixel position of a location specified in satellite or scanner model coordinates.

Excursus 3.1—Estimating Image Pixel Width and Depth

Most satellite image data have good nominal rectification so the map and image (with the same projection and datum— see Section 6) should overlay easily. If there is significant error between the two, the fit can often be improved by adjusting the nominal pixel size by an appropriate factor. To determine this factor:

- mark a few pairs of features, which are visible on both the map and image, and are separated on the image in the vertical and/or horizontal directions as shown in Figure 3.3;
- measure the distance between each pair of points in the X direction on both the map and image; do the same for the Y direction;
- compute factors for stretching or contracting the pixel size for X and Y separately as the average of:

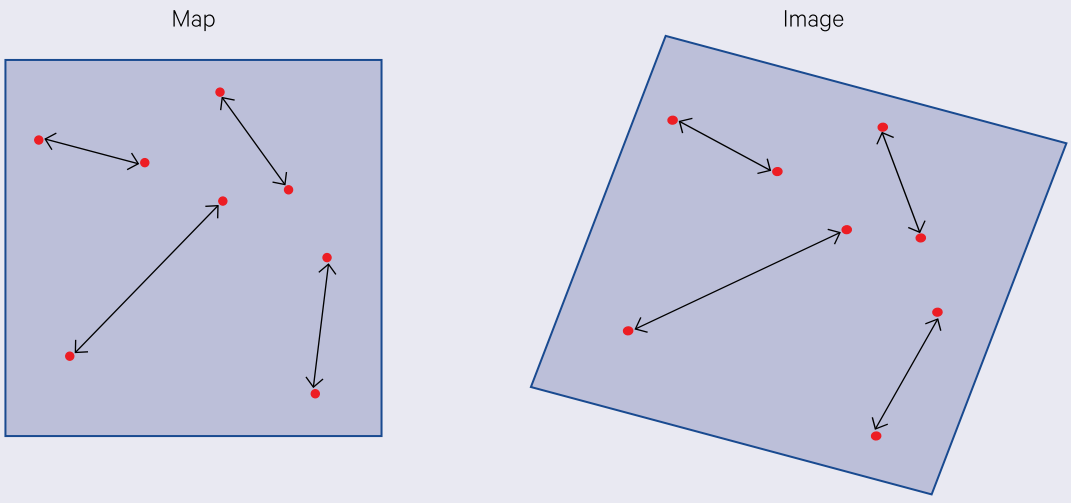
$$\frac{\text{map distance}}{\text{image distance}}$$

- redisplay the image (or reproduce the image hardcopy) with:
  - pixel width = X factor × nominal pixel width
  - pixel depth = Y factor × nominal pixel depth
- and all other factors the same as were used for the first set of hardcopy images.

This approximate rectification approach is often quite close to the final one.

Figure 3.3 Determining pixel size adjustment factor

The distances between corresponding pairs of locations can be measured in both the map and the image to determine an approximate value for image pixel width and depth.



Source: Harrison and Jupp (1992) Figure 36

3.2 Arbitrary Mathematical Models

Mathematical models allow a map or image to be treated mathematically like a rubber sheet, so that its shape may be stretched or contracted into a different geometry. These models are used to convert between any two coordinate systems for which a nominal model is not available (such as an image and a map, or two images), or to account for any arbitrary image distortions which are not considered in the nominal models. The arbitrary model can be used in conjunction with nominal models as the second step in the three-stage registration process (see Section 3.3).

These models are of the form:

$$X' = F_x\left(X, Y\right)$$
$$Y' = F_y\left(X, Y\right)$$

where

X' is the new X value after conversion;
Y' is the new Y value after conversion;
X is the original X value before conversion;
Y is the original Y value before conversion;
F_x is a mathematical function allowing the original X and Y values to be used to compute X'; and
F_y is a mathematical function allowing the original X and Y values to be used to compute Y'.

For polynomial models, F_x and F_y consist of mathematical equations which relate the new X' and Y' values to the original X and Y values using coefficients with linear and higher order terms of X and Y plus offsets. Typically, four types of polynomial equation may be used to compute X' and Y':

- affine;
- bilinear (or ruled or first order polynomial);
- quadratic (or second order polynomial); or
- cubic (or third order polynomial).

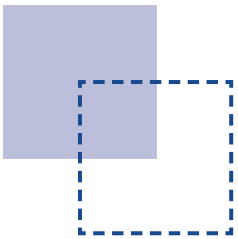
The geometric effects of each of these models are illustrated in Figure 3.4.

Wars of nations are fought to change maps.
But wars of poverty are fought to map change.
(Muhammad Ali)

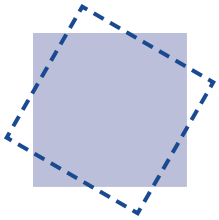
Figure 3.4 Geometric effects of polynomial models

a. An affine model changes origin (shift), orientation (rotation), scale (enlargement or reduction) and one-directional linear skew (in X or Y direction).

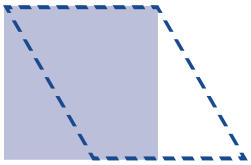
Shift of origin



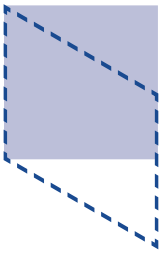
Rotation of axes



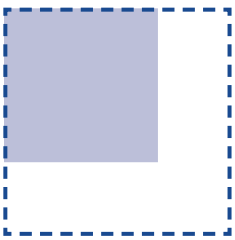
Skew in X direction



Skew in Y direction

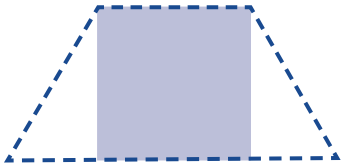


Scale change



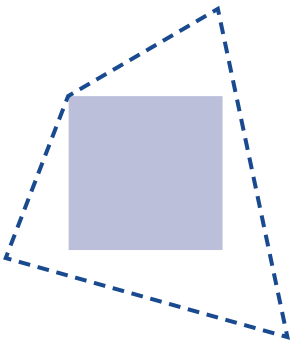
b. Bilinear model: changes origin, orientation, scale and two-directional, linear skew

Skew in X direction with changes in X and Y

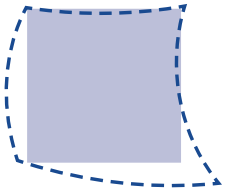


Skew in X direction with changes in X and Y

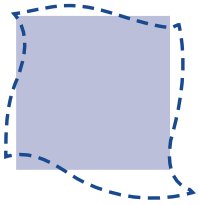
Skew in X and Y directions with changes in X and Y



c. Quadratic model: changes origin, orientation, scale and two-directional, non-linear skew with one point of inflection



d. Cubic model: changes origin, orientation, scale and two-directional, non-linear skew with two points of inflection



Source: Harrison and Jupp (1992) Figure 24b, 26, 27 and 28

3.2.1 Affine

An affine model is represented by equations of the form:

$$X' = A + BX + CY$$

$$Y' = K + LX + MY$$

or the two-dimensional matrix:

$$\begin{bmatrix} B & C \\ L & M \end{bmatrix}$$

with shift or offset factors:

$$\begin{bmatrix} A \\ K \end{bmatrix}$$

where

X' is the new X value after conversion;
Y' is the new Y value after conversion;
X is the original X value before conversion;
Y is the original Y value before conversion; and
A, B, C, K, L, M are coefficients of the affine equation.

These equations apply to the traditional planar coordinate system with an origin at bottom-left (BL) (see Figure 1.5). The operations can be used with a dataset whose origin is top-left (TL), such as the image coordinate system, by subtracting a constant from the appropriate input value(s). A conversion from TL to BL can be simply implemented by subtracting the number of lines plus 1 from each Y value before computing the affine model. The value being subtracted is referred to as the origin value. Origin values can be specified for both X and Y (denoted O_x and O_y respectively) to give equations of the form:

$$X' = A + B \times \left(X - O_x\right) + C \times \left(Y - O_y\right)$$

$$Y' = K + L \times \left(X - O_x\right) + M \times \left(Y - O_y\right)$$

Variations to the origin settings are given below.

The basic geometric operations that may be implemented using 2×2 matrix operations are skewing, rotating, rescaling and reflecting. With the additional offset term in the equation to allow shifting, an affine model can account for five types of difference between two coordinate systems:

- shift in X and/or Y direction (see Section 3.2.1.1);
- skew in X and/or Y direction (see Section 3.2.1.2);
- rotation (see Section 3.2.1.3);
- scale change (see Section 3.2.1.4); and
- reflection (see Section 3.2.1.5).

An affine transformation accounts for the major distortions in satellite imagery. A significant advantage of this transformation is that it can be implemented, up to a rotation, by the simple along-line resampling process (see Section 5.1). Along-line resampling allows the image to be rescaled in both directions and skewed in the x direction. The process is implemented in conjunction with the nominal satellite and scanner models so that the specific along-line distortions (such as non-linear scan rate, sensor delay offset, panoramic distortion and Earth curvature) may also be removed from the imagery.

Along-line resampling is well suited to EO imagery since the main resampling operations are applied along each image line. This form of resampling offers a useful means for producing nominally rectified imagery and provides a convenient starting point for developing an accurate rectification. When image distortions result in image lines no longer being parallel however, this process will not adequately correct for the geometric errors.

3.2.1.1 Shift in X and/or Y direction

This operation simply moves the reference origin of the original coordinate system (see Figure 3.5). A shift operation is represented mathematically by adding an offset value to each of the original coordinates. This may be modelled by the equations:

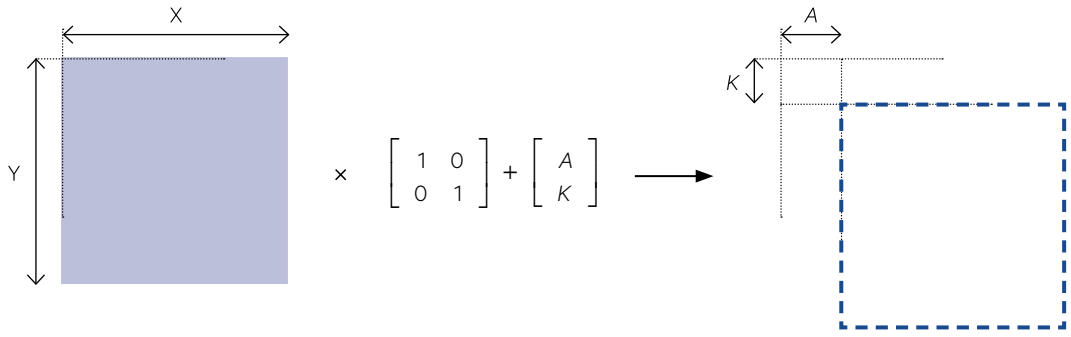
$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} A \\ K \end{bmatrix}$$

$$X' = A + X$$

$$Y' = K + Y$$

A shift accounts for differences in the origins of the two coordinate systems and/or displacement of the image due to platform pitch or roll.

Figure 3.5 Shift operation



3.2.1.2 Skew in X and/or Y direction

This form of skew is concerned with warping the image in a single direction at a constant rate (see Figure 3.6). Such a skew is implemented by modifying the X values by a factor related to the original Y values and/or modifying the Y values by a factor related to the original X values.

X values could be skewed with increasing line number using the matrix:

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

that is:

$$\begin{aligned} X' &= (1 \times X) + (k \times Y) \\ &= X + (k \times Y) \\ Y' &= (0 \times X) + (1 \times Y) \\ &= Y \end{aligned}$$

The affine model for this would be:

O_x	A	B	C	O_y	K	L	M
0	0	1	k	0	0	0	1

Or

$$\begin{aligned} X' &= 0 + 1 \times (X - 0) + k \times (Y - 0) \\ &= X + k \times Y \\ Y' &= 0 + 0 \times (X - 0) + 1 \times (Y - 0) \\ &= Y \end{aligned}$$

Similarly, Y values could be skewed with increasing pixel number by the matrix:

$$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

or

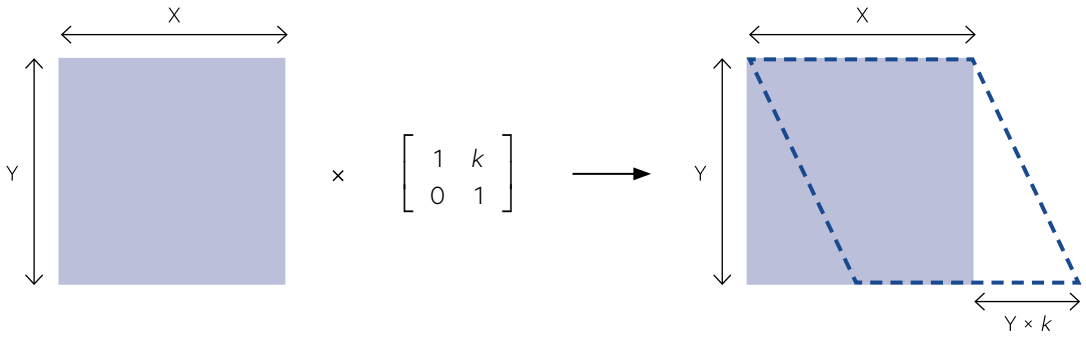
$$\begin{aligned} X' &= 1 \times X + 0 \times Y \\ &= X \\ Y' &= k \times X + 1 \times Y \\ &= kX + Y \end{aligned}$$

A skew in both directions can be implemented by a single skew in X or Y plus a rotation, or

$$\begin{aligned} X' &= X + CY \\ Y' &= LX + Y \end{aligned}$$

To skew by an angle θ, in the above equations C and L would be equal to $\tan\theta$. For skews in two directions, it is convenient to implement a skew in one direction and account for the second direction of skew in the rotation factor. This operation can be used to adjust for additional skew in EO imagery (in addition to the skew which is accounted for using the available nominal models), such as may be due to variations in the platform heading.

Figure 3.6 Skewing operation in X direction



3.2.1.3 Rotation

Rotation of θ° in an anti-clockwise direction may be expressed in matrix notation as:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

(see Figure 3.7). As a polynomial model in a standard planar coordinate system this matrix can be written as:

$$X' = X \times \cos\theta + Y \times -\sin\theta$$

$$Y' = X \times \sin\theta + Y \times \cos\theta$$

To allow for image coordinates, with reversed line numbering relative to the standard planar coordinate system (that is, origin TL), we need to offset the resulting Y values. For example, when θ is 90°, $\cos\theta$ equals 0 and $\sin\theta$ equals 1 so these equations become:

$$X' = Y$$

$$Y' = \#P - X$$

where *#P* is the number of pixels in an image line plus one (Figure 3.8a). This requires an affine model of:

O_x	A	B	C	O_y	K	L	M
0	0	0	*1*	*#P*	0	-1	0

When θ equals 180°, $\cos\theta$ equals -1 and $\sin\theta$ equals 0 so:

$$X' = \#L - Y$$

$$Y' = \#P - X$$

where *#L* is the number of lines in the image plus 1 (see Figure 3.8b). This rotation is implemented using the affine model:

O_x	A	B	C	O_y	K	L	M
#L	0	0	-1	*#P*	0	-1	0

When θ equals 270° anti-clockwise (or 90° clockwise), $\cos\theta$ is 0 and $\sin\theta$ equals –1 so we get:

$$X' = \#L - Y$$

$$Y' = X$$

(see Figure 3.8c). This transformation requires an affine model with the parameters:

O_x	A	B	C	O_y	K	L	M
#L	0	0	-1	0	0	1	0

Rotation of one coordinate system to match another accounts for small area differences in projection (for example, LTM is conformal along a meridian whereas Landsat satellite model coordinates are aligned with the satellite orbital path) and platform yaw.

Figure 3.7 Rotation operation

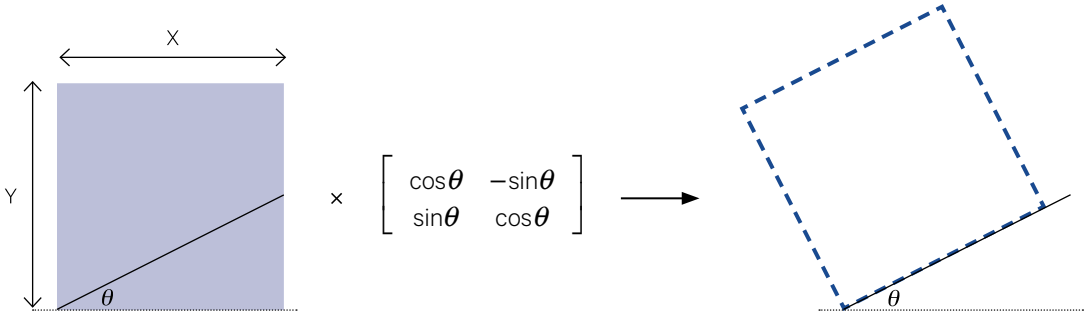
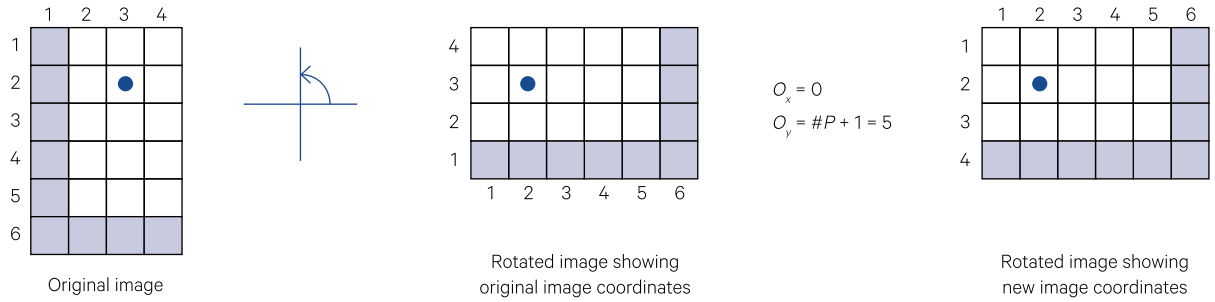
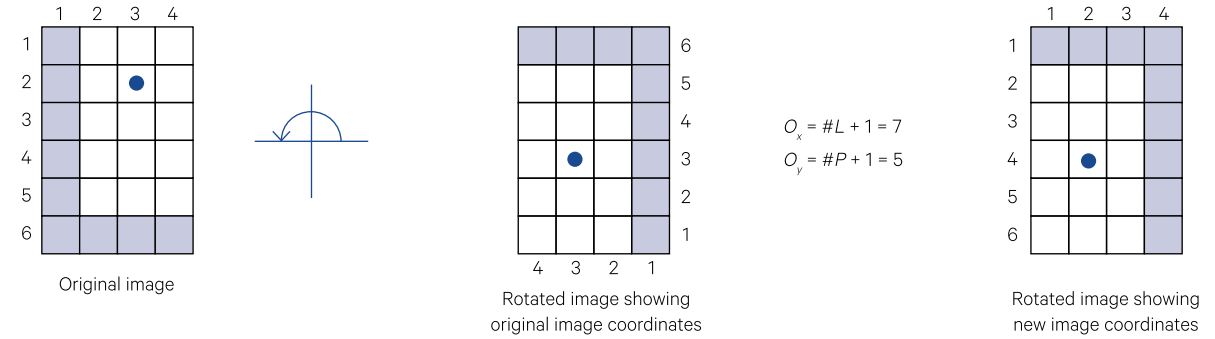


Figure 3.8 Origin adjustment for image rotation

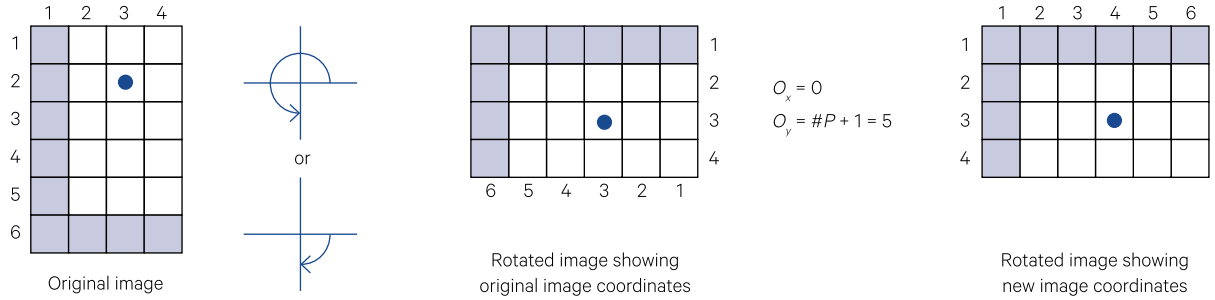
a. 90° anticlockwise rotation



b. 180° anticlockwise rotation



c. 270° anticlockwise or -90° clockwise rotation



Source: Harrison and Jupp (1990) Figure 25

3.2.1.4 Scale change

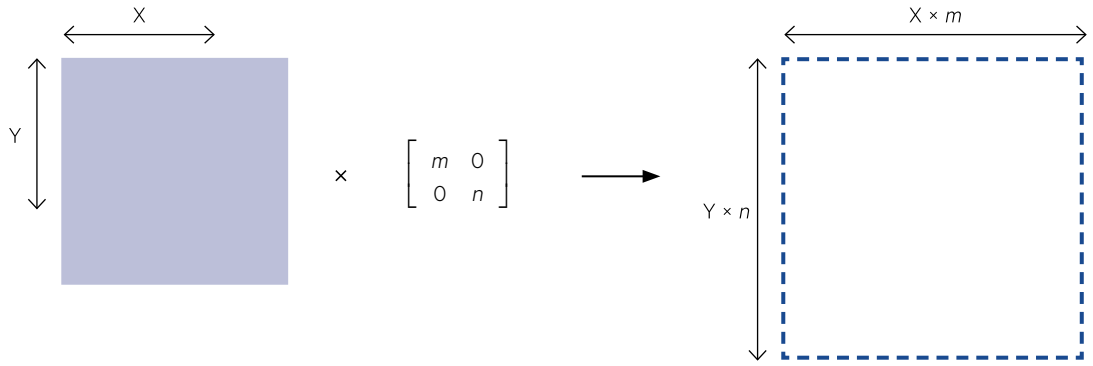
A change in scale is calculated simply by multiplying the coordinates of the original system by an appropriate factor (see Figure 3.9). Linear rescaling of coordinate axes is simply implemented by the matrix:

$$\begin{bmatrix} m & 0 \\ 0 & n \end{bmatrix}$$

or

$$X' = m \times X + 0 \times Y$$
$$= mX$$
$$Y' = 0 \times X + n \times Y$$
$$= nY$$

Figure 3.9 Rescaling operation



If *m=n*, both axes are rescaled by a constant factor. As an affine transformation model, this operation is specified as:

O_x	A	B	C	O_y	K	L	M
0	0	*m*	0	0	0	0	*n*

that is,

$$X' = 0 + m \times (X - 0) + 0 \times (Y - 0)$$
$$Y' = 0 + 0 \times (X - 0) + n \times (Y - 0)$$

or

$$X' = mX$$
$$Y' = nY$$

A scale change can be used to adjust imagery for a non-unity aspect ratio, as occurs when pixel width is not equal to pixel depth, or for a constant variation in platform altitude.

3.2.1.5 Reflection

The matrix for reflection on the line:

$$Y = X \times \tan\theta$$

Is

$$\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

or

$$X' = X \times \cos 2\theta + Y \times \sin 2\theta$$
$$Y' = X \times \sin 2\theta + Y \times -\cos 2\theta$$

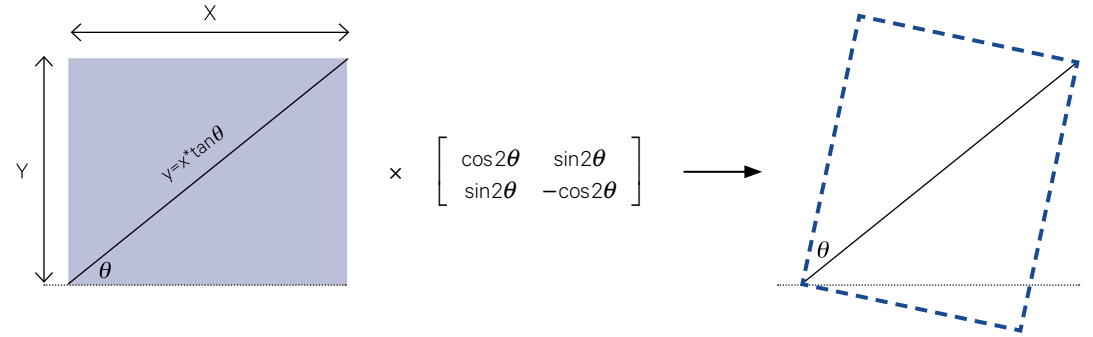
(see Figure 3.10). To reflect the coordinate system on the line Y=X (this is the same as transposing the image when *#P=#L*), $\theta = 45°$ so:

$$X' = Y$$
$$Y' = X$$

or the affine model:

O_x	A	B	C	O_y	K	L	M
0	0	0	1	0	0	1	0

Figure 3.10 Reflection operation



3.2.2 Bilinear

A bilinear transformation accounts for all the factors considered in the affine transformation, plus a bi-directional skew in X and/or Y (see Figure 3.11). This is represented by equations of the form:

$$X' = A + BX + CY + DXY$$

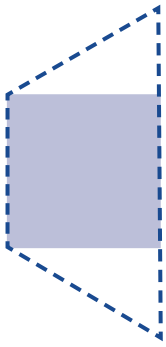
$$Y' = K + LX + MY + NXY$$

The XY term here allows a constant skew distortion to be applied in more than one direction along the lines and/or down the pixel columns of the image. This means that the image line length can change linearly with the line position in the image (or similarly vary the pixel column depth with the pixel position).

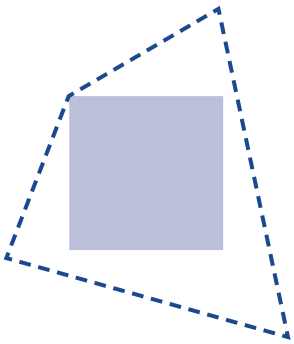
Figure 3.11 Bi-directional skew

a. Skew in X direction with changes in X and Y

b. Skew in Y direction with changes in X and Y



c. Skew in X and Y directions with changes in X an Y



Source: Harrison and Jupp (1992) Figure 26

This transformation accounts for variations due to a changing platform altitude during image acquisition or variable rates of platform pitch and roll. In satellite imagery, a bilinear model should account for most image distortions in the Y direction and, in conjunction with the satellite models, for many distortions in the X direction also. The extent of distortion in aircraft imagery however, largely due to the instability of that platform, usually requires that a higher order model be used to properly correct for the geometric error in such data.

3.2.3 Quadratic

The quadratic model accounts for all the factors considered in the bilinear transformation, plus 'curving' of grid lines in the X and/or Y directions (see Figure 3.8). This is represented by equations of the form

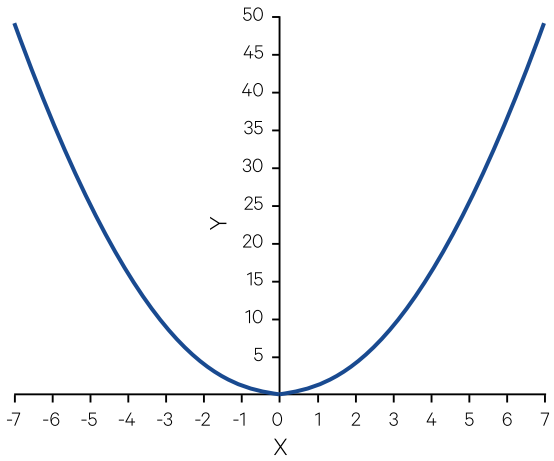
$$X' = A + BX + CY + DXY + EX^2 + FY^2$$

$$Y' = K + LX + MY + NXY + OX^2 + PY^2$$

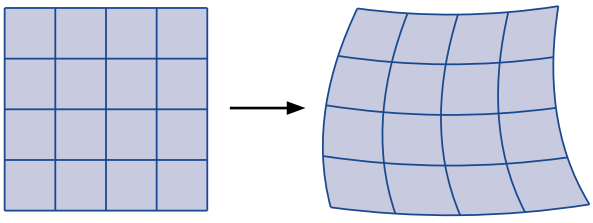
Thinking back to high school maths, a quadratic function is graphed as a parabola (see Figure 3.12a). This shape is characterised by a point at which the slope of the function line changes from being positive to negative. This characteristic is due to the squared terms (X^2 and/or Y^2) which vary the X' and Y' values in a non-linear way with respect to X and Y. The same principle is used here to model the distortions in the lines and/or pixel columns of the image in a way that allows the lines and/or columns to be curved. In a quadratic model, this curving of grid lines can only occur once. In conjunction with the other factors accounted for by the model, the severity of the curving may vary throughout the image, although again the rate of variation is in a single direction (see Figure 3.12b).

Figure 3.12 Quadratic transformation

a. Mathematical function: $Y = X^2$



b. Effect on image geometry



Source: Harrison and Jupp (1992) Figure 27

A quadratic model may be required for EO imagery to supplement the nominal along-line distortion models or account for 'single direction' variations in platform altitude, heading, velocity, pitch, roll or yaw within the imagery. That is, these factors may both increase and decrease during acquisition of the image. Some sensors (such as CZCS) also have the option of looking forward into the direction of travel to reduce the effects of sunglint on water surfaces. This characteristic results in complex distortions in the imagery, which would not accounted for by a satellite model but can be removed with a quadratic rectification model.

As higher order terms are introduced into the rectification model, it more closely approximates the minor geometric errors in the imagery but also becomes more specific to the sample data points from which it is derived. If these points are accurately located, and are representative of the spatial variations that occur throughout the image, then the model may be used reliably over the full image extent. However, if the locations contain error, the model will be fitted to these errors as well as to the data. Similarly, if the sample points are not representative of the whole image, the model will not produce reliable results away from control points (see Section 4). Higher order models can account for non-linear distortions in the image but, by being image-specific, they are also less stable away from the fitted data and this limitation must be considered when using such models for image rectification.

3.2.4 Cubic

The cubic model accounts for all the factors considered in a quadratic transformation but allows grid lines to be 'bent' at two places (see Figure 3.13). It is represented by equations of the form:

$$X' = A + BX + CY + DXY + EX^2 + FY^2 + GX^3 + HX^2Y + IXY^2 + JY^3$$

$$Y' = K + LX + MY + NXY + OX^2 + PY^2 + QX^3 + RX^2Y + SXY^2 + TY^3$$

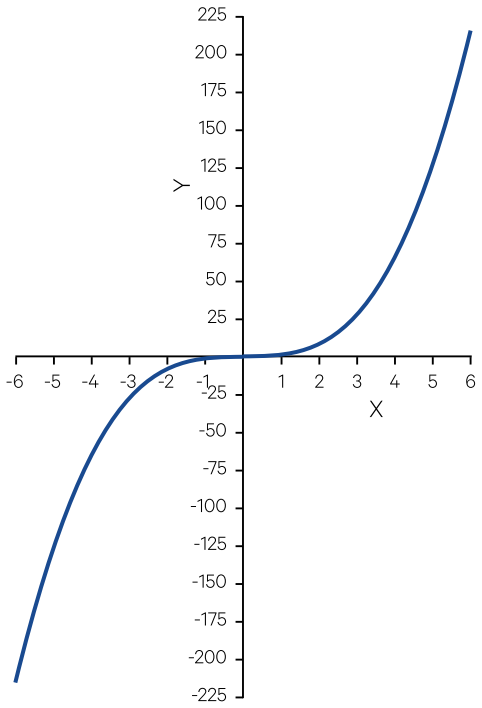
A cubic function is graphed as a line with two points of inflection, that is, points where the slope changes from positive to negative as illustrated in Figure 3.13a. This feature allows both the direction and rate of 'bending' of image lines and pixel columns to vary twice throughout the image (see Figure 3.13b). The four extra terms in the equations (X^3, X^2Y, XY^2 and Y^3) allow considerably more flexibility for warping, stretching or contracting the image to a new shape.

The cubic model accounts for platform altitude, heading, velocity or attitude changes in 'two directions' within the image. For example, whereas the quadratic function allowed a factor such as altitude to increase and decrease within an image, the cubic model allows this factor to increase, decrease and increase again (or decrease, increase and decrease again). Non-linear scan rate, for example, could be modelled by a cubic function, where the scan rate slows down, speeds up and slows again across an image line. For platform variations, this flexibility is generally only required for imagery acquired with aircraft.

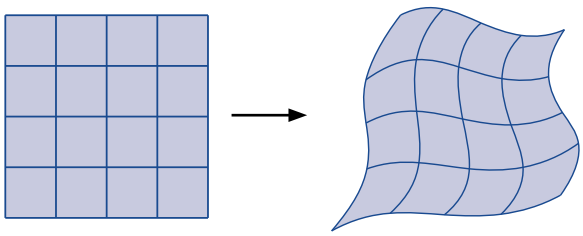
The cubic model can account for 'stronger' warps in the data but, by doing so, is also quite unstable and cannot be reliably applied away from the section of the image from which it is derived. The flexibility of higher order models is countered by their more limited applicability and, to achieve reliable rectification results with EO data, this must be appreciated when such models are used. The procedure for deriving polynomial rectification models using control points is described in Section 4.

Figure 3.13 Cubic transformation

a. Mathematical function: $Y = X^3$



b. Effect on image geometry



Source: Harrison and Jupp (1992) Figure 28

3.3 Multiple Models

Transformations between two different geometries, such as an EO image and a map, usually involve three stages. A three-stage registration process typically uses two nominal models and a polynomial model. The nominal models account for quantifiable image or map distortions and are used before and after a polynomial model. This composite transformation framework allows such distortions to be removed independently of a registration between the geometries.

The three stages of a full coordinate transformation are referred to as:

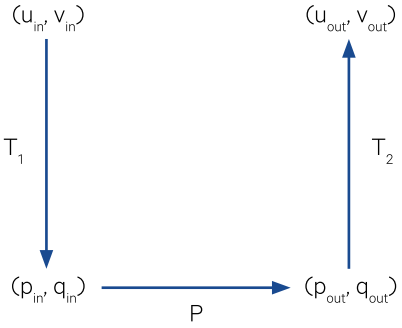
- T_1, a nominal coordinate transformation to convert (u_{in},v_{in}) to (p_{in},q_{in});
- P, a polynomial transformation to convert (p_{in},q_{in}) to (p_{out},q_{ou}t); and
- T_2, another nominal transformation to convert (pout,qout) to (uout,vout) (see Figure 3.14).

Figure 3.15 represents the standard coordinate transformation cases between image and map coordinates. Figure 3.15a illustrates these stages in the conversion of image to UTM coordinates, with the reverse conversion being shown in Figure 3.15b. Where the map coordinates are geographical (latitude/longitude), the T_2 in Figure 3.15a will be LTM to geographical (or HOM to geographical) and similarly T_1 in Figure 3.15b will be geographical to LTM or HOM. For example, Excursus 3.2 illustrates the geometric effects of each stage in the conversion between AVHRR image coordinates and geographical coordinates.

Some of the geometric distortions of EO imagery are not satisfactorily modelled by low order polynomials so this scheme provides much more accurate conversion results. Similarly, inconsistencies in map coordinates (such as a zone change in UTM coordinates) may not be accounted for by the mathematical model. The three-stage process allows image registration to be effected more accurately. The same transformation structure can be used to convert from one image system to another (see Figure 3.15c).

Figure 3.14 Generalisation of three-stage registration model

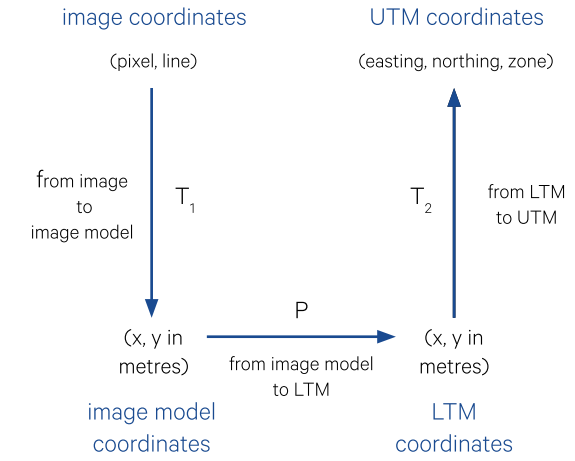
P is a polynomial transformation from (p_{in}, q_{in}) to (p_{out}, q_{out})
T_1 is a nominal model which converts from (u_{in}, v_{in}) to (p_{in}, q_{in})
T_2 is a nominal model which converts from (p_{out}, q_{out}) to (u_{out}, v_{out})



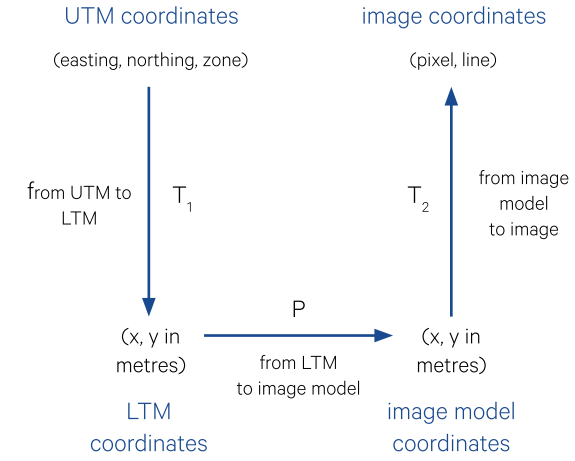
Source: Harrison and Jupp (1992) Figure 29

Figure 3.15 Three-stage registration model

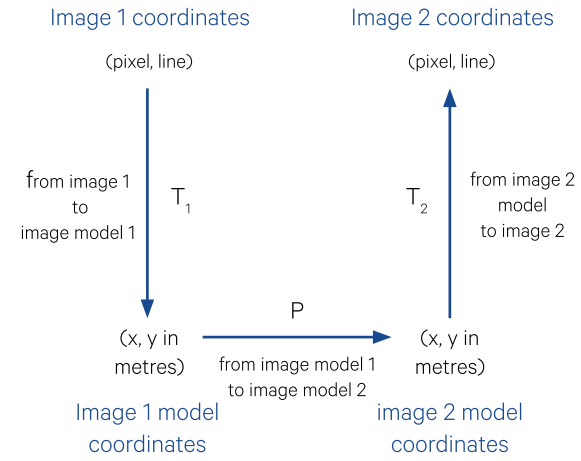
a. Converting from image to map



b. Converting from map to image



c. Converting from image to image



Source: Harrison and Jupp (1992) Figures 30 and 32

Excursus 3.2—Using Multiple Models for Image Rectification

This example uses a set control points from an AVHRR image to demonstrate the process of converting from image to map coordinates. Being continental scale data, AVHRR imagery is often acquired with significant spatial distortions, particularly panoramic distortion and Earth curvature (see Figure 2.7). When a regularly-spaced set of locations within an AVHRR image are selected, their positions on a map is not regular (see Figure 3.16). To simply convert directly from image to map coordinates would require a more complex, and less accurate, polynomial model (see Sections 3.2 and 4).

As detailed in Section 3.3, rectification of EO imagery is most stable if a sequence of three models are used to convert between two different coordinate systems (see Figure 3.14). Rather than just defining a polynomial model directly between these image and map coordinates, nominal models, which account for known geometric distortions in the satellite image and map projection, are first applied to convert both sets of coordinates into model coordinates. Once the 'quantifiable' distortions within the original coordinates are removed, a simpler polynomial model can be fitted to the two sets of model coordinates to describe the relationship between them.

In this case, to convert from AVHRR image coordinates to geographic coordinates (latitude and longitude), the three models would be:

- T_1—an AVHRR image model, which accounts for the systematic, geometric distortions known to exist in AVHRR imagery, converts from image coordinates (u_{in},v_{in}) in pixels to satellite model coordinates (p_{in},q_{in}) in metres;
- P—a tailored polynomial transformation converts from satellite model coordinates (p_{in},q_{in}) to Hotine Oblique Mercator (HOM) coordinates (p_{out},q_{out}), with both sets of coordinates expressed in metres relative to a defined origin; and
- T_2—a map transformation model converts from HOM coordinates (p_{out},q_{out}) in metres to longitude and latitude (u_{out},v_{out}) in decimal degrees.

This sequence is illustrated in Figure 3.17. The value of using this multi-model approach is clearly demonstrated in this illustration. The major differences between the image and map coordinates are removed when both sets of coordinates are expressed as coordinates of their respective nominal models. This means that a much lower order polynomial model is required to map from the satellite model projection to the HOM projection than would be required to directly map from AVHRR image coordinates to a geographic projection. Additionally, the transformations effected by nominal models precisely account for the known distortions in both map and image coordinates. This level of correction would be difficult to achieve using a polynomial model alone.

Figure 3.16 Locations of control points

A set of regularly-spaced locations within an AVHRR image are shown in image coordinates, but the same set of locations shown in geographic projection are spaced irregularly.

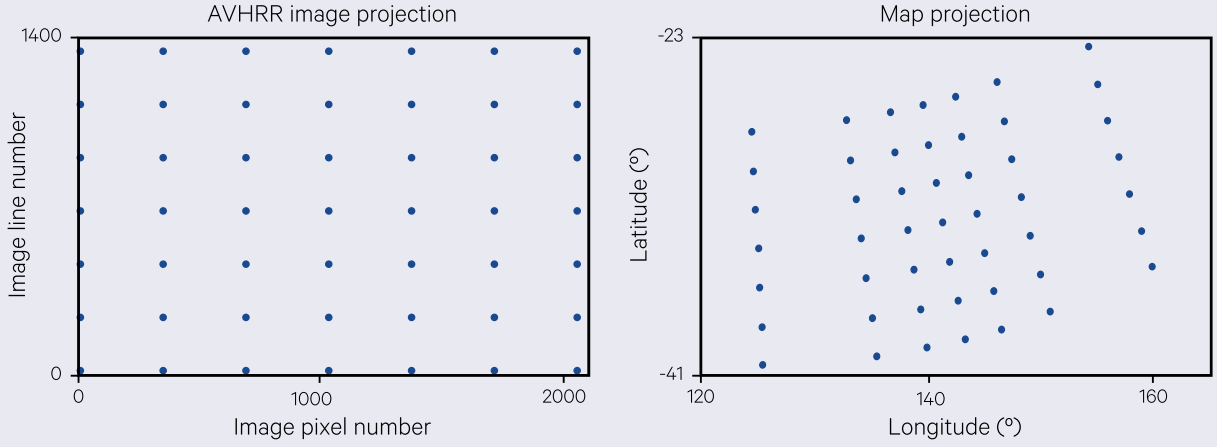
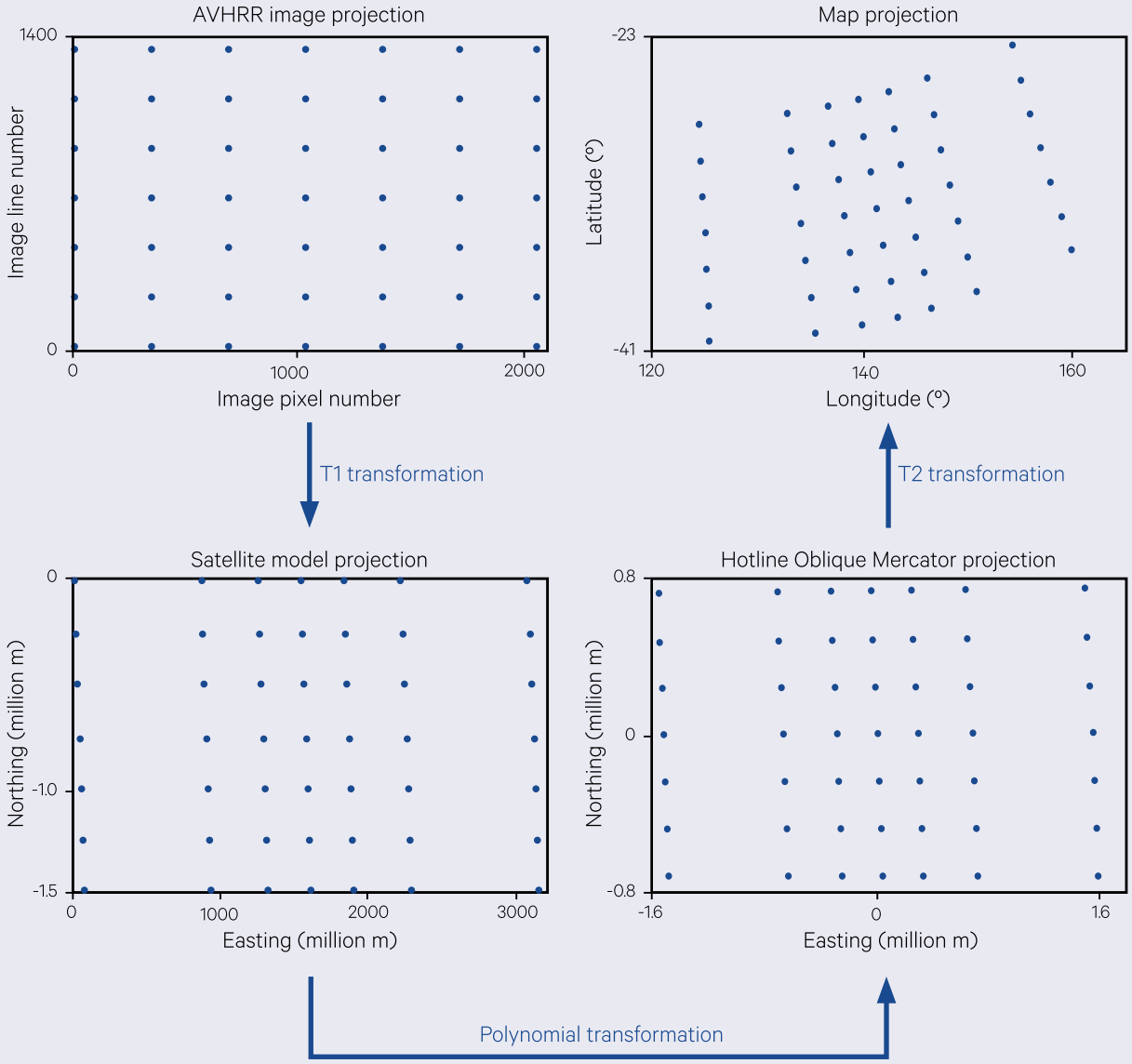


Figure 3.17 Three-stage conversion from AVHRR image to map coordinates



Source: Harrison and Jupp (1992) Figure 31

3.4 Further Information

Affine transformation:

MathWorks: *https://au.mathworks.com/discovery/affine-transformation.html*

Wolfram MathWorld: *http://mathworld.wolfram.com/AffineTransformation.html*

3.5 References

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Harrison, B.A., and Jupp, D.L.B. (1992). *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual*. MPA, Melbourne.

Lewis, A., Wang, L.-W., and Coghlan, R. (2011). *AGRI: The Australian Geographic Reference Image. A Technical Report*. GeoCat#72657. Geoscience Australia, Canberra. ISBN 978-1-921954-47-4.

Stanaway, R. (2011). *Aligning an LTM grid with UTM*. 45th ASPNG Congress, Madang, Papua New Guinea, 29–22 July 2011.

4 Control Point Modelling

Control points can be used to develop a mathematical model between two different coordinate systems, such as a map and an image (or two images). These points are precise locations that can be clearly identified in both coordinate systems as illustrated in Figure 4.1.

For each control point, precise coordinates in the two systems are obtained. Map points are typically recorded either as metres in UTM northing, easting and zone, or as degrees of longitude and latitude. The image points are generally recorded as pixel and line coordinates. Nominal models (discussed in Section 3.1) are recommended to remove the quantifiable distortions inherent to these original coordinate systems. The relationship between the resulting, more conformal, coordinate systems is then modelled on the basis of these sample points (seeExcursus 3.2).

The ultimate aim of this process is to relate a location in an image with a precise location on a map or another image, or vice versa. This may be done for individual locations, such as field sites or digitised strings, or for the overall image, such as to digitally resample the image to overlay it onto a map or another image (see Section 5).

Figure 4.1 Control point modelling

Ten common features have been selected as sample locations or 'control points' in map of image datasets over Parliament House, Canberra. For simplicity, this artificial example uses a street map, but such maps are rarely sufficiently accurate for EO image rectification. Note that the image has been rotated for this example so that there are no locations in this map corresponding to the lower left corner of the image.

a. Control points located on map as cyan crosses

b. Control points located on image as cyan crosses

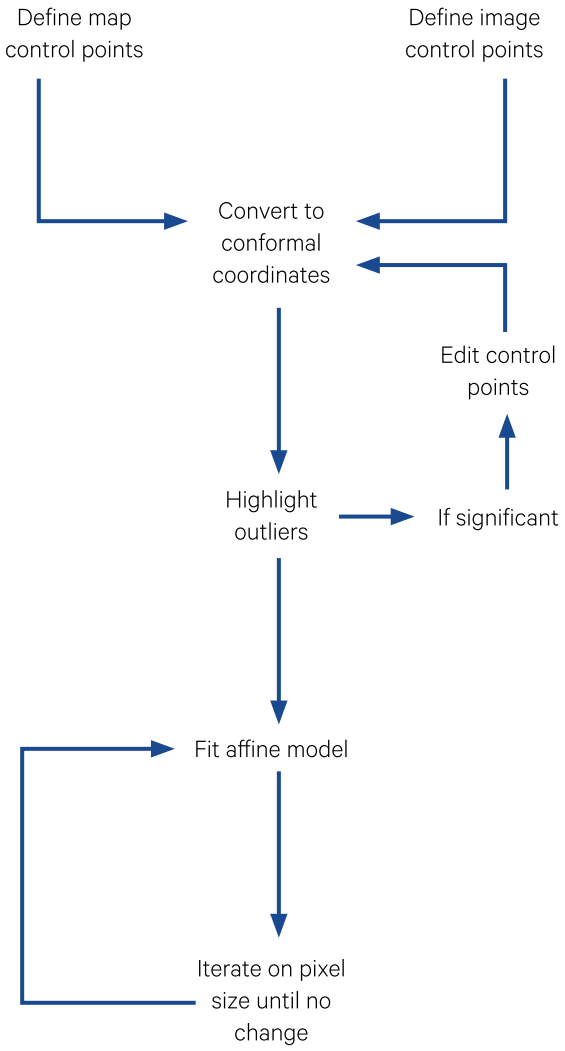
Source: a. Extracted from Whereis map of Canberra. b. Extracted from Google Maps in August 2018.

Background image: High resolution airborne image of Barrow Island, WA, acquired on 19 April 2016. **Source:** © Nearmap

In most image processing systems, image coordinates for GCP can be interactively located on a displayed image or entered numerically. In the three-stage rectification process described in Section 3.3, these image coordinates will be converted to nominal model coordinates before fitting with a polynomial model to match map coordinates. Some image processing systems allow sets of ground control points to be checked for outliers before fitting with a polynomial model. This modelling sequence is illustrated in the flow chart in Figure 4.2. The parameters that define the registration process can then be used to convert locations from one coordinate system to another or to resample an image to another geometry.

The need for control always comes from someone that has lost it.
(Shannon L. Alder)

Figure 4.2 Control point modelling sequence



Source: Harrison and Jupp (1992) Figure 35a

4.1 Selecting Control Points

Control points may be supplied with some sources of image data so that individual feature points do not need to be selected. However, usually this is not the case and these sample points must be located manually for each image being rectified.

4.1.1 Selection criteria

For image-to-map rectification, the maps used to assign coordinates to the control points must be accurate if the image needs to be rectified in an absolute sense. An inaccurate map may be used to give a relative rectification, that is, the image will be rectified relative to that map, but in an absolute sense the errors across the reference points will be larger. This may be valuable in cases where an accurate map is not available and multiple images need to be rectified relative to a common base. Provided a sufficient number of control points can be located on such maps (as described below) and a suitable coordinate system is available (or can be devised), they may be used as a rectification base.

Control points derived from suitable maps can and should be persistent, that is, they define features that will not move or disappear. This is most relevant to multi-image studies since the points can effectively be considered as a 'library' of accurately located features on the ground. For image registration, the features should also be located on imagery, with different sets of points being 'detectable' on different images.

Base maps for image rectification need to be at a scale appropriate to the image data (see Excursus 4.1). An important consideration with scale is the effect a reasonable error in map reading would have on the image in terms of pixel size. However, the scale of the imagery should not be so large as to distort the features that are to be used as control points—they should be clearly visible but not over-enlarged.

Examples of good control points would be road junctions, cultural boundaries, stable river junctions, and points of headlands. These points should be accurately identifiable and marked on both the image and the map. The points need to be well spread over the study area rather than clustered in a few areas. For example, the minimum requirements for control points used in the AGRI (Australian Geographic Reference Image) project (Lewis *et al.*, 2011; see Excursus 6.1) are described in Excursus 4.2.

Excursus 4.1—Map Accuracy Standards

Standards have been defined for map accuracy, which require that 90% of control points from an independent sample are actually located within a fixed distance of their locations on the map. Map Accuracy Standard 1 (MAS1) defines this distance as 0.5 mm; MAS2 uses 1 mm. For a 1:250,000 scale map, MAS1 relates to a ground distance of 125 m (0.5 mm × 250,000). A 90% probability in a normal distribution is associated with a standard deviation of 1.645, so that in the image rectification process the residual root mean square (RMS) error should be less than 125/1.645 or about 76 m. MAS2 at this scale considers a ground distance error of 250 m (1 mm × 250,000) or a residual RMS error of 250/1.645 (about 152 m).

At 1:100,000, MAS1 ground distance error is 50 m (0.5 mm × 100,000) with RMS error equal to about 30 m. At this scale, the 100 m ground distance error in MAS2 relates to an RMS error of about 60 m.

The potential accuracy of rectified EO images can also be assessed in terms of these accuracy standards. Landsat MSS data, for example, could easily be rectified to meet the MAS1 at 1:250,000. At 1:100,000 scale, rectified Landsat imagery readily meets the MAS2 standard and often satisfies MAS1. For a rectification base map, the MAS2 standard is adequate, so in this case a suitable base map scale would be 1:100,000 or even 1: 50,000. Continental scale data such as AVHRR can similarly be rectified using 1:1,000,000 or 1:500,000 scale base maps. It should be noted that the 'accuracy' here refers to the error between the grid coordinates predicted by a model for a given (fractional) pixel location and the surveyed grid coordinates of the corresponding point on a good base map. This differs from the accuracy of a photographic or paper product, which may have distortions due to the method of production.

Excursus 4.2—Control Point Selection in AGRI project

Source: Wang Lan-Wei, Geoscience Australia
Further Information: Lewis *et al.* (2011)

The *Australian Geographic Reference Image* (AGRI) is a consistent and accurate reference image, with 2.5 m spatial resolution, for rectification of imagery from multiple sources at spatial resolutions of 2.5 m or less. This product and its development are detailed in Excursus 6.1.

In this project, all control points were required to:

- be unambiguously identifiable at the sub-pixel level in the ALOS/PRISM image, and on the ground;
- be located either near the centre of, or near to the edge of, the image swath;
- have good radiometric contrast;
- be a 'suitable' type of feature as described below;
- where possible, be located in the overlap between adjacent swaths; and
- be at least 250 m from any other surveyed control point.

Ground features **suitable** for control points included:

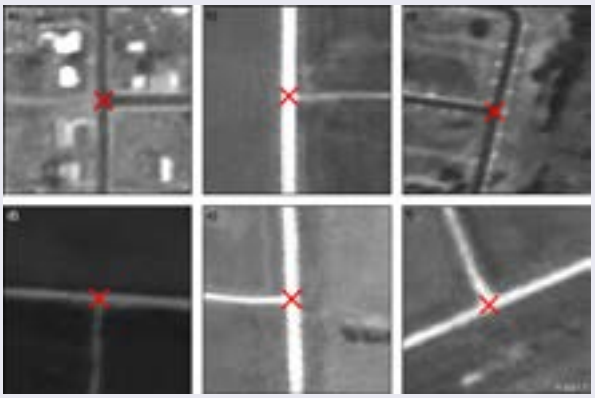
- narrow roads with intersection angles of at least 50° (see Figure 4.3 and Figure 4.4);
- centre points of quasi-circular features; and
- road-watercourse junctions.

Examples of **unsuitable** ground features included:

- the edges of inland waterbodies;
- the edges of roads (especially where the surface changes);
- coastal features subject to tidal variation;
- linear features intersecting at an angle of less than 50°;
- road intersections where the intersection cannot be identified precisely;
- features larger than about 5 pixels in extent; or
- features that are not clearly identifiable on the supplied ALOS/PRISM imagery.

Figure 4.3 Examples of suitable ground features for control point survey

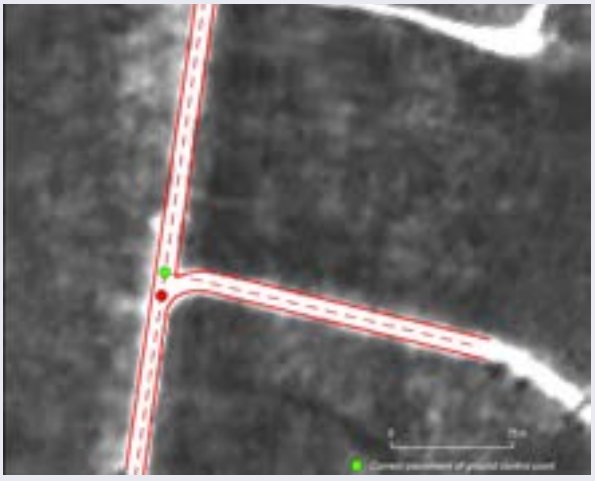
Some control point locations shown on PRISM imagery, which were used by surveyors to locate proposed features.



Source: Lewis *et al.* (2011) Figure 9

Figure 4.4 Precise placement of control point at road intersection

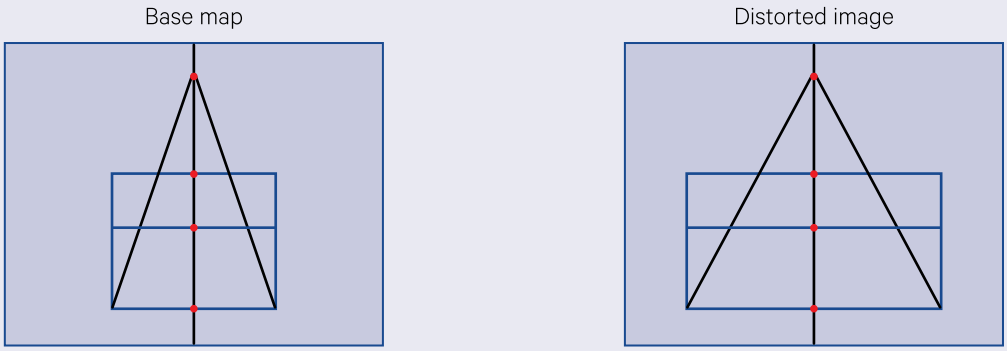
In this image solid red lines mark the edge of the linear features (roads) and dashed red lines mark the centrelines. The correct placement of a control point at this intersection is shown as a green dot. The red dot shows an incorrect placement of the control point.



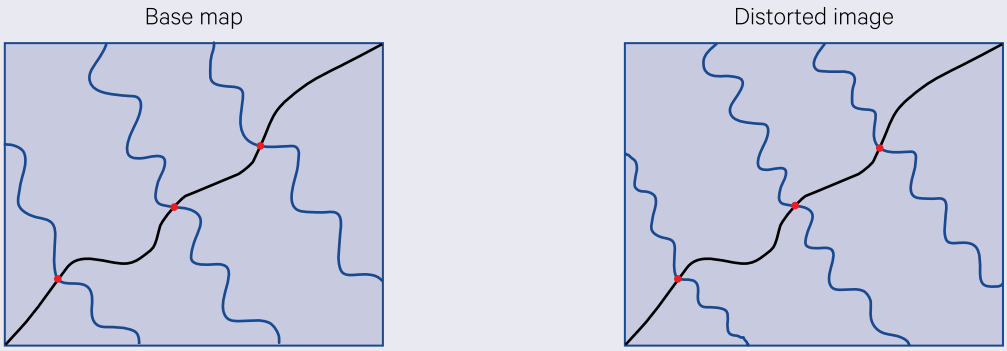
Source: Lewis *et al.* (2011) Figure 10

Figure 4.5 Control point distribution

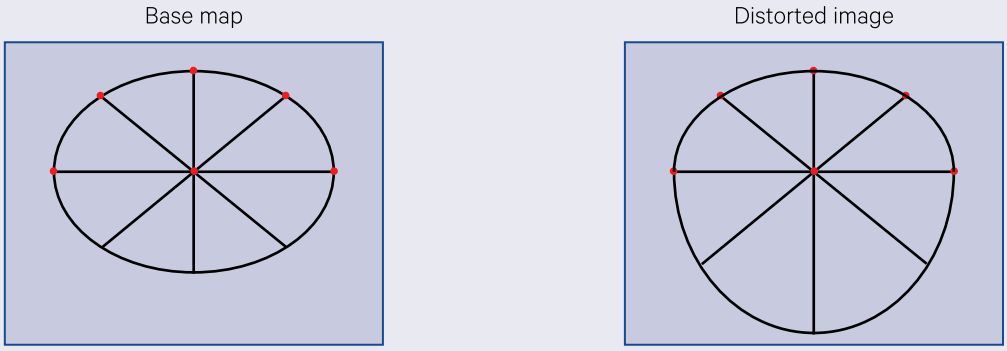
a. GCPs (marked •) are not well-distributed across the image, so cannot correct for the along-line distortions.



b. GCPs only occur along one diagonal so will not account for distortions in other corners of the image.



c. GCPs clustered on one side of a feature cannot account for different distortions in another part of the feature.



Source: Harrison and Jupp (1992) Figure 37

Since control points are being used to describe a statistical model between the image and base map (or two images), their spatial distribution is critical to successful image registration. Poorly spaced control points result in instability in the model away from these points as shown in Figure 4.5. It should be remembered that even when control point selection allows a stable model to be derived, the 'accuracy' of the rectified image is limited by the 'accuracy' of the base map. Any global scale characteristics of the base map—good or bad—are fitted onto the image by the rectification model. If a map, which contains systematic distortions, is being used as a base map, these distortions will also be evident in the rectified image.

However, it is not always possible to place control points where they are needed for sound model fitting. In the Great Barrier Reef, for example, the sparse distribution of reefs creates problems for selecting evenly distributed control points (Jupp *et al.*, 1982). In this case one approach involved locating a large number of control points (for example, 100 to 200 per Landsat MSS scene) and then relying on statistical averaging and a constrained registration model to iron out poor local control. Section 4.2.3 describes an outlier sieving method, which is invaluable for locating poor control points before model fitting.

4.1.2 Selection process

The following discussion will concentrate on the case of image-to-map rectification, however, the same procedures apply to image-to-image rectification with control points being selected from the two images rather than from an image and a map. For image-to-image rectifications, base maps are not required, but it is often useful to produce image hardcopies at appropriate scales.

4.1.2.1 Map control points

The coordinates of the map control points need to be carefully read or digitised from the map sheets and recorded digitally. Obviously, the points need to be labelled consistently on both map and image so that they can be paired correctly during the modelling process.

Map-based control points are usually recorded from maps in either geographical coordinates (latitude and longitude) or UTM (Universal Transverse Mercator) coordinates. In the UTM projection, coordinates within one zone are all relative to a common origin and are based on a distance unit of one metre. However, geographical coordinates (latitude and longitude) do not conform to these criteria. Longitude measures east-west distances in degrees along a particular latitude, however, since the circumference of each successive parallel of latitude decreases away from the equator, the actual distance on the ground being measured by a degree also decreases.

For the rectification of aircraft scanner imagery, which typically covers a very small area on the ground, any standard mapping coordinate system may be used. Pixel size estimates in metres can only be obtained with a coordinate system based on a one metre distance unit as discussed above.

A full Landsat image covers 185 km, which is approximately 15% of the width of a UTM zone at the equator. Regardless, a single image will frequently straddle two UTM zones, in which case it is necessary to convert the map points into a coordinate system where all points will be distanced relative to the same origin. Even when the image is contained within a single UTM zone, projection distortions (convergence away from the conformal meridian) are minimised by converting the locations to a coordinate system that is centred on the study site. If the map coordinates are recorded as latitude and longitude they will also need to be converted to a more conformal coordinate system. The LTM (Local Transverse Mercator) system described in Section 3.1.1.1 is recommended in both these cases.

Continental scale imagery, such as AVHRR, has a nominal swath width of thousands of kilometres. As such, a full scene may cover between two and five UTM zones. At this scale the LTM transformation does not produce a satisfactory coordinate base. The Hotine Oblique Mercator projection should be used to develop rectification models for such imagery (see Sections 3.1.1.2 and Excursus 3.2).

4.1.2.2 Image control points

Image coordinates for the selected control points are generally identified interactively on a displayed image. The image should be displayed so that the features that have been selected as control points are clearly visible. This requires that an appropriate linear stretch (and possibly a non-linear, or gamma, stretch) is determined for the channels being displayed. Non-standard channel combinations may also be useful to enhance particular features. In some images, the use of an edge enhancement filter transformation would sharpen roads and cultural boundaries thereby making such features easier to pinpoint on the image (see Volume 2C). The display zoom feature will make features easier to see but care should be taken to not over-enlarge the image on the screen since important contextual information will be lost.

Most image processing systems indicate the fractional image coordinates for the image pixel over which the cursor is positioned. The reported coordinate is often relative to the full image scene. Image channel values for the pixel may also be listed: this is a useful location check where the control point should have distinctive spectral values, such as land rather than water on a coastal boundary.

An example of registration of high resolution Pleiades-1 imagery is provided in Excursus 4.3.

Control is an illusion
(Melody Beattie)

Excursus 4.3—Registration of Pleiades-1 Satellite Imagery

Source: Tony Sparks, Icon Water

This excursus describes part of a Research and Development project undertaken by Icon Water[3] during 2012/2013, entitled 'Vegetation Structure and Biomass Estimation in the Area of Inundation of the Expanded Cotter Dam' (ACTEW Water, 2013). Immediately prior to the commencement of the filling phase of the Enlarged Cotter Dam (ECD), this vegetation survey was undertaken within the area of inundation utilising Pleiades imagery provided by Airbus Defence and Space[4]. Figure 4.6 shows the ECD wall on 24 December 2012, three days after image acquisition, when the wall is essentially at full height and matching design drawings for the completed dam.

Figure 4.6 Cotter Dam wall on 24 December 2012

The enlarged dam is 80 m and constructed using roller compacted concrete. This is the highest dam of its type in Australia.

ACTEW Water (2013) details methods used to prepare the imagery and derive the mapping, and discusses field information and final results. The vegetation structure and biomass results of this study were used by Icon Water to assist in further management and planning, specifically water quality issues and their likely impacts on native fish populations within the ECD reservoir and planting activities within the surrounding catchments. This summary focuses on the orthorectification of Pleiades-1 imagery.

Pleiades-1 Imagery

Pleiades-1 satellite imagery, acquired on the 21 December 2012, was used to map vegetation structure within the expanded Cotter reservoir area of inundation. Pleiades multispectral images comprise four spectral bands (blue: 430–550 nm; green: 490–610 nm; red: 600–720 nm and NIR: 750–950 nm) with a spatial resolution of 2 m, while panchromatic imagery (480–830 nm) has 0.5 m spatial resolution. Multispectral and panchromatic imagery for this study were acquired during moderate to dry environmental conditions, with limited rainfall in the preceding months. These conditions provided for the best chance to gain acceptable discrimination between woodland/forest and grassland land cover units. Imagery was cloud-free over the study area and indicated reasonable vegetative growth where the ground has not been cleared.

Image Registration

The Pleiades imagery required orthorectification to remove distortions resulting from changes in ground elevation and satellite viewing angle[5]. Orthorectification depends on accurate ground control and an appropriate digital elevation model (DEM), combined with specialised image processing software. Ground control points (GCP) had been collected by Icon Water since 2004, using both highly accurate survey information and differentially-corrected GPS data (DGPS) on assets visible in the imagery. The DEM, which covered the entire Cotter catchment (current to November 2012), was largely derived from lidar data and had been previously prepared by the Australian National University (ANU) and Icon Water using ANUDEM software (Hutchinson, 1989).

ENVI orthorectification software analysed satellite ephemeris information supplied with the Pleiades DIMAP formatted imagery and the DEM to generate an initial orbit and view geometry model, which was further optimised with multiple GCP. The orthorectification process used the DEM to correct the image pixel-by-pixel so that the resulting orthorectified image provided geolocated coordinates for each pixel.

[3] Then called Actew Water
[4] Then called Astrium Services
[5] Viewing angle = 14.5556588333929°

Ground Control

Icon Water retains records for a large number of GCP within and immediately surrounding the ACT that have been collected using:

- DGPS (Trimble GEO XT/XH);
- detailed on-ground surveys;
- Work-as-Executed drawings; or
- high resolution (e.g. 10 cm) imagery.

The post-processing of all differentially-corrected GPS survey data provides sub-metre accuracy (in X and Y only) for all GCP. This information is collected both as part of asset capture programs, and within environmental and consulting studies. These GCP are a mixture of road intersections and crossings, fence lines, drainage line confluences, water utilities infrastructure, and engineering structures.

Only a small number of the available GCP were used for the purpose of orthorectification of this imagery. Those selected were evenly-distributed, both across the entire image area and over a wide range of elevations. Each surveyed GCP included an actual elevation value. For those sites collected by DGPS, an elevation value was extracted from an underlying lidar-based DEM. The elevations of sites within the area of construction were based on design drawings.

Additional to these points, detailed road surveys were collected in the study area by the DGPS with the aid of an external antenna during June/July 2012 as part of a sediment source study. This information provides a visual confirmation of the quality of the final orthorectification.

Elevation Grid

The Elevation grid used for rectification was derived from the following source datasets:

- 2003 Cotter Catchment Lidar Survey (1 m resolution grid)—processed using ANUDEM to provide a continuous, drainage-enforced surface or grid over the entire catchment;
- 2006 Cotter Reservoir Bathymetric Survey—acquired by Ecowise Environmental at high resolution to provide an accurate volume assessment then merged with the 2003 lidar grid;
- March 2011 Murrumbidgee River: Cotter to Burrinjuck Lidar Survey—prepared by Icon Water using ANUDEM to ensure consistency between surveys;
- August 2011 Murrumbidgee River: Tantangarra to Cotter Lidar Survey—also prepared by Icon Water using ANUDEM; and
- Engineering drawings for the ECD (since the dam wall was not completed at the time of the lidar surveys)—extracted elevations for the ECD infrastructure (primarily the Roller Compacted Concrete wall) were used to build a three-dimensional representation of the dam wall and valve tower.

All elevation datasets were mosaicked as grids (rather than merged as raw point data and re-modelled), taking into account changes that had occurred since May 2003. The resulting elevation grid provides a 1 m representation of the elevations across the study area. Areas of lidar overlap (between 2003 and 2011 surveys) were consistent to within 10 cm where no changes had occurred, and known engineering structures were within 10 mm. All 'cliff' features within the landscape were maintained, particularly the vertical faces of the dam walls.

Implementation

The orthorectification process was implemented using eight GCP on the panchromatic image and 14 GCP on the multispectral images. The orthorectification used a bilinear resampling kernel and was projected onto MGA Zone 55. Though not specifically required for the study, the accuracy of the orthorectification was performed on the panchromatic image.

Accuracy Assessment

The resulting orthoimage was assessed for accuracy at a selection of independent GCP across the image and at differing altitudes (Figure 4.7). These sites were within and surrounding the primary study area of the project, specifically the area of inundation of the ECD. Additional visual checks on major infrastructure and along the road traces were also performed. These results are summarised in Table 4.1.

Figure 4.7 GCP used for accuracy assessment.

Locations of six control points were used to verify the accuracy of the orthorectified image.

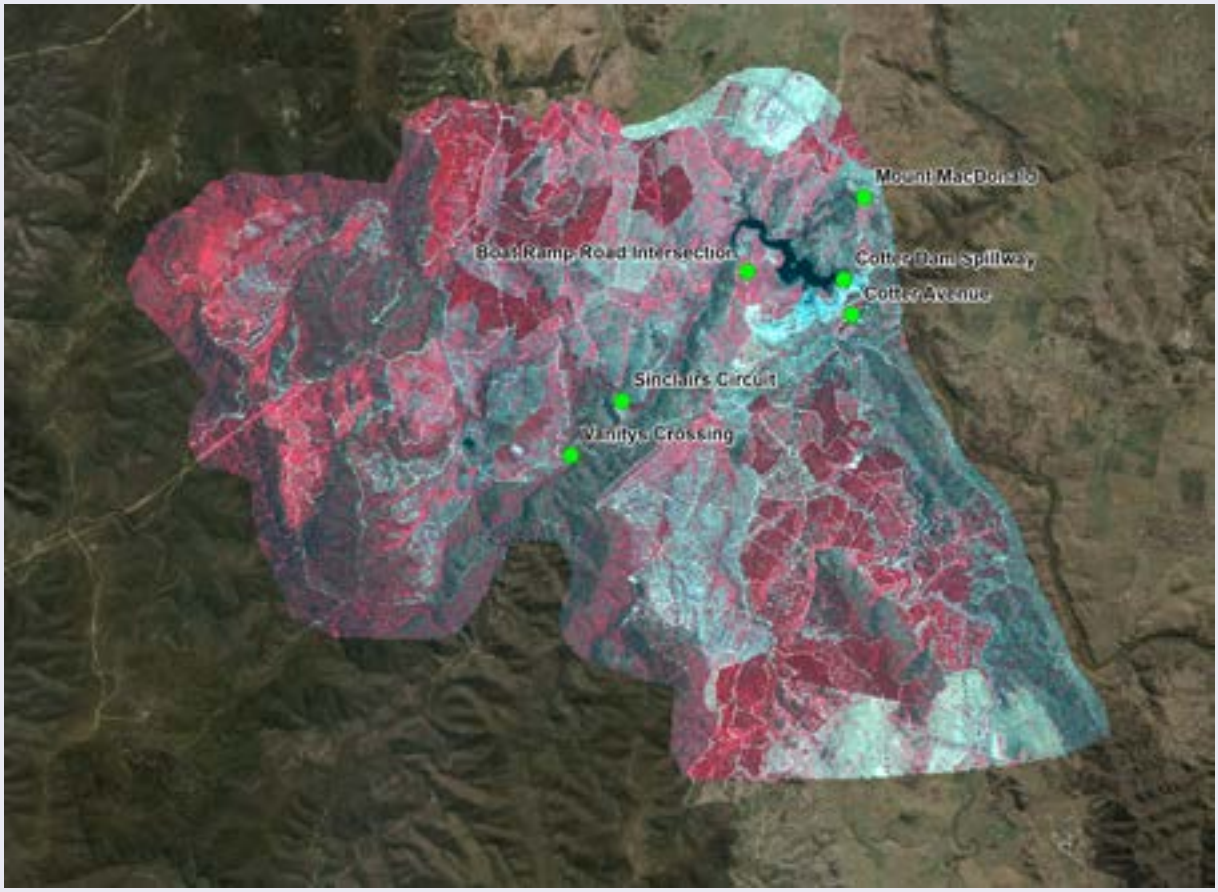


Table 4.1 Verification errors in selected GCP

Location	Elevation (m)	Error (m)
Cotter Dam spillway (pre 2013)	500.69	0.90
Vanity's Crossing	566.80	0.87
Road intersection near boat ramp	575.02	0.34
Sinclair Circuit	587.94	0.88
Mount MacDonald	786.50	0.86
Cotter Avenue	475.76	0.74

All errors within the imagery were within 1 m elevation and one pixel extent, and were within the anticipated accuracy of the ground control provided, with errors independent of elevation. Of particular interest is the expanded dam wall, which was successfully corrected (including the valve tower), confirming the quality of the DEM and GCP. The adverse effect of orthorectification on the cranes over the construction area is visible in Figure 4.8, but was not considered a concern for this study.

Figure 4.8 Orthorectified image

This portion of the orthorectified Pleiades-1 panchromatic image demonstrates the adverse effect of orthorectification on features above ground level as defined by elevation data, most noticeably the cranes.

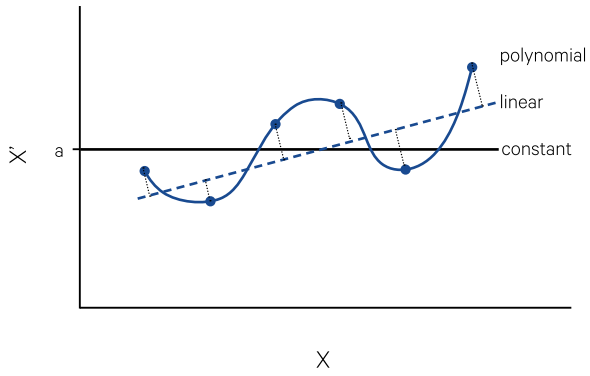
4.2 Fitting Arbitrary Models

As introduced in Section 2, the geometric distortions that can occur in EO imagery may be due to either 'systematic' or 'arbitrary' effects. Section 3.1.2 outlined the types of systematic or quantifiable distortions, such as Earth curvature or panoramic distortion, which can be corrected using satellite or scanner models. Arbitrary effects however, due to changes in the altitude or attitude of a sensor platform, can only be modelled using detailed orbit/flight information. Alternatively, an image containing such effects can be registered to another image or a map by fitting a mathematical model to sample points from the two geometries (see Section 3.2 above).

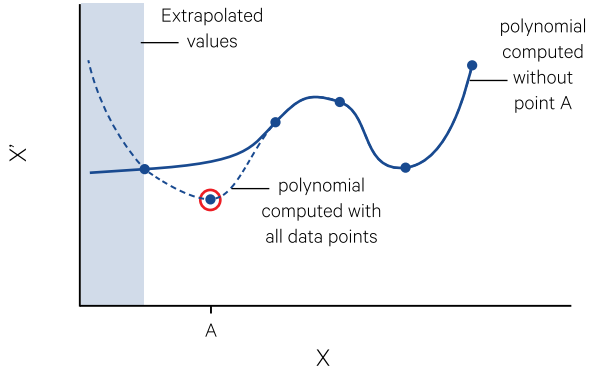
When fitting any mathematical model to a set of sample data points, the selected sample needs to be representative of the whole dataset before the model can be reliably applied to it. For modelling the geometric relationship between two sets of spatial data this requires that the sample locations be well distributed over the whole data area (see Section 4.1.1).

Figure 4.9 Model stability and complexity

a. Potential models that can be fitted to a set of points include a constant value (black line), a linear model (dashed blue line) and a higher order polynomial model (solid blue line).



b. The stability of the model to predict a particular value can be demonstrated when one or more data points are removed. In this example, if the second point (circled red) is removed, the polynomial model changes significantly, especially for extrapolated values.



Source: Harrison and Jupp (1992) Figure 38

Various criteria can be used to select the model which 'best' describes the relationship between two sets of points. The stability of a mathematical model decreases as its complexity increases so that, while a selected model may closely fit a limited number and range of data points, its behaviour away from those points—or without one of the points—may alter radically. Potential models that can be fitted to a set of points include a constant value, a linear model and a higher order polynomial model (see Figure 4.9a). In this example, the polynomial model changes signficantly when the second point is removed (see Figure 4.9) and would obtain quite different values if used to extrapolate X' from lower values of X.

The constant function:

$$X' = a$$

could be considered as the most stable model but, as illustrated in Figure 4.9a, this is rarely suitable. A linear model:

$$X' = b + cX$$

could be fitted to the data to approximate X' values for the given X values but in this case some residual errors are relatively large. Such low order models are relatively insensitive to small changes in the values of data points. A high order polynomial could be defined to precisely compute X' for each X value. Such a model however would be unstable for extrapolating away from the range of data points and is very sensitive to any changes in the data as illustrated in Figure 4.9b.

In this case the 'best' model depends on the number of model parameters the sample datasets can support and what the model will be used for. Three types of model selection criteria are relevant to image rectification:

- the minimum model is the order of polynomial which has minimum noise-to-signal ratio value (NSR). This statistic indicates the size of errors between fitted and actual data values relative to the range of the sample data points. The minimum model is recommended for extrapolating the model away from the range of sample points used to fit it.
- the optimum model is the polynomial with minimum predictive error value. The predictive error of a point indicates the importance of each data point to the fitted model and is further discussed in Sections 4.2.2 and 4.2.5. This model is recommended for interpolating within the range of the sample dataset. If the sample points are well distributed over the whole image, this model should be used for image resampling and point location.

- the maximum model is the highest order model that can be fitted to the data. In many image processing systems, this is the cubic polynomial (provided there are at least 11 control points). As illustrated in Figure 4.9, the maximum model is most sensitive to the range and spread of the sample data values used to fit it. The maximum model is useful for checking the extent of control provided by the sample points over the image.

Typically, image processing systems fit a polynomial model (affine, bilinear, quadratic or cubic—see Section 3.2) to sets of paired GCP using various statistical analyses to indicate the stability of the overall model and the individual points used to compute it (see Section 4.2.4). Before fitting a model, any outlier points can be detected as described in Section 4.2.3. The final registration parameters can used to convert point locations (see Excursus 4.4) or resample images (see Section 5).

Excursus 4.4—Image Rectification Example

Source: Tony Sparks, Icon Water

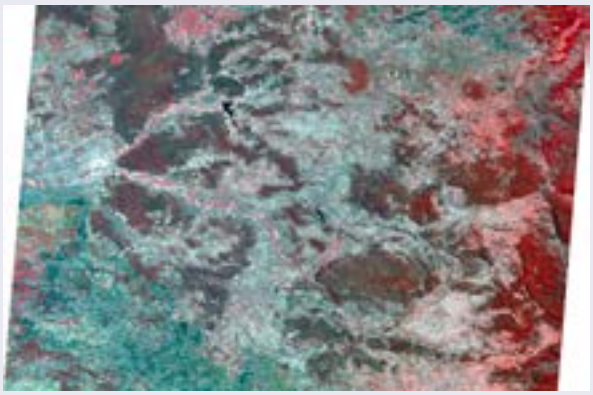
Two adjacent Landsat MSS scenes, acquired in September 1982 and shown in Figure 4.10, show the characteristic distortions due to Earth rotation skew that occur satellite imagery acquired in descending mode (see Section 2). They also demonstrate the 'squashed' vertical extent of unrectified Landsat MSS imagery that resulted from its rectangular geometric pixel dimensions (approximately 60 m horizontal width by 80 m vertical depth).

The locations of GCP used to rectify these image scenes are shown in Figure 4.11, and listed in Table 4.2 and Table 4.3 respectively. In this example, only the eastern portion of the western image was required for analysis, so most GCP were selected in the overlap area. As a result, the residual errors resulting from the rectification were higher for the path 90 image (see Table 4.4).

Figure 4.10 Landsat MSS imagery

These adjacent image scenes were acquired in September 1982 and are shown with correction for the along-line distortion due to Earth rotation skew. Both images appear 'squashed' vertically due to non-square geometric pixel dimensions of Landsat MSS imagery.

a. Path 90 Row 80



b. Path 89 Row 80

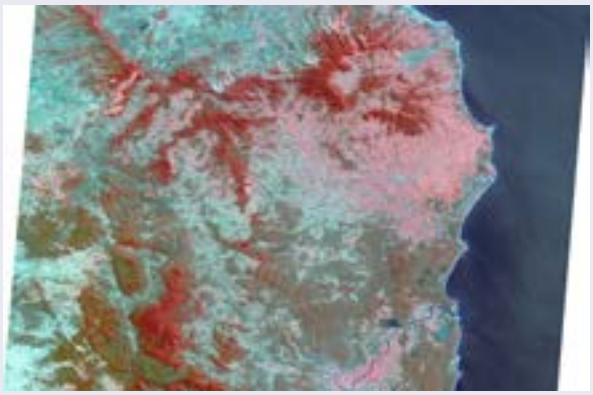
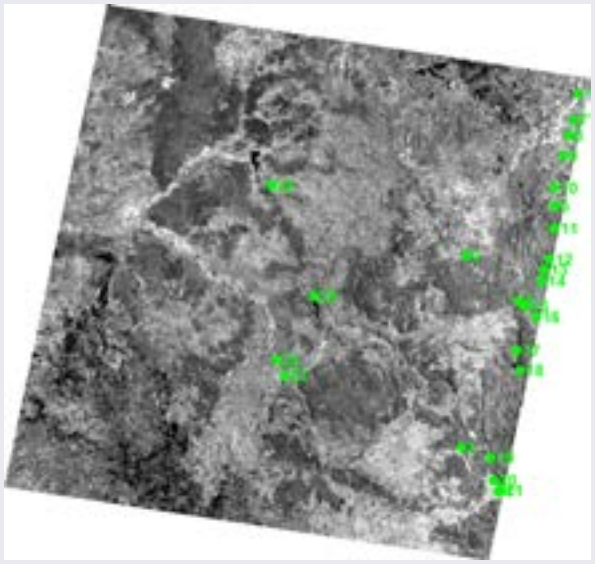


Figure 4.11 GCP locations shown on resampled, greyscale images

a. 25 GCP for path 90 scene



b. 16 GCP for path 89 scene

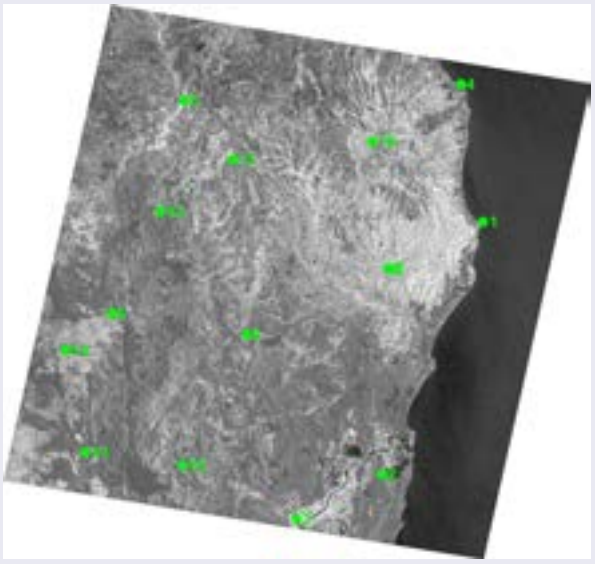


Table 4.2 15 GCP used to rectify path 89 scene

GCP ID	Image coordinates		Map coordinates								Residual (m)
	Column	Row	Easting (m)	Northing (m)	Longitude (°E)			Latitude (°S)			
					Degrees	Minutes	Seconds	Degrees	Minutes	Seconds	
1	2886.99	712.33	562067.31	6831256.23	153	38	6.25	28	38	39.15	21.96
2	2497.19	1944.41	523722.55	6735171.44	153	14	41.19	29	30	45.78	16.82
3	1981.76	2211.55	491251.11	6718308.71	152	54	34.53	29	39	54.35	30.57
4	2590.93	99.94	553628.13	6883574.61	153	32	46.66	28	10	20.43	31.02
5	786.91	384.09	448535.55	6877216.92	152	28	31.67	28	13	47.34	14.57
6	543.55	1411.41	421133.45	6796343.97	152	11	26.24	28	57	30.19	17.49
7	2325.64	998.56	526749.06	6813283.07	153	16	26.85	28	48	27.54	11.17
8	2315.52	997.64	526157.6	6813445.73	153	16	5.01	28	48	22.29	29.71
9	1451.73	1406.82	472241.93	6788349.1	152	42	53.72	29	1	57.65	36.63
10	1174.69	2053.34	448018.82	6738596.97	152	27	49.74	29	28	51.4	13.9
11	1172.32	618.13	467053.97	6854710.36	152	39	48.83	28	26	0.85	22.02
12	752.97	908.96	439616.98	6835085.92	152	22	56.51	28	36	34.99	9.42
13	520.29	2067.01	411064.5	6743540.03	152	4	59.05	29	26	3.32	19.6
14	279.26	1613.72	403583.36	6782439.67	152	0	33.61	29	4	57.59	17.45
15	2071.34	428.54	520117.74	6861770.43	153	12	19.13	28	22	12.37	45.01

Affine transformation models were fitted to rectify each image to UTM zone 56S (with central meridian 153°E) based on AGD66. In this example, nominal correction of the satellite image had already been applied (see Section 3.1.2), so the transformation models were computed directly between map and image coordinates. The resulting affine model parameters are detailed in Table 4.4. Transformation models for the path 89 image were computed as:

$$E = 56.15943155 \times C - 13.35866642 \times R + 409{,}473.11350178$$
$$N = -9.2434234 \times C - 80.91196505 \times R + 6{,}915{,}582.58314177$$
$$C = 0.01733537 \times E - 0.00286209 \times N + 12{,}694.65970755$$
$$R = -0.0019804 \times E - 0.01203214 \times N + 84{,}020.2121473$$

where

E is Easting map coordinate (in m);
N is Northing map coordinate (in m);
C is pixel column in image (in pixels); and
R is pixel row in image (in pixels).

The transformation models for the path 90 image using all 25 GCP were computed as:

$$E = 56.22224251 \times C - 12.31774692 \times R + 259{,}503.11469513$$
$$N = -8.57283482 \times C - 81.04980707 \times R + 6{,}912{,}558.30279933$$
$$C = 0.01738371 \times E - 0.00264193 \times N + 13{,}751.38715929$$
$$R = -0.00183872 \times E - 0.01205865 \times N + 83{,}833.26472743$$

These transformation models enable the rectification to be used to convert between image and map coordinates for specific locations and also to resample both images to a common base grid (see Figure 4.12).

Table 4.3 25 GCP used to rectify path 90 scene

GCP ID	Image coordinates		Map coordinates								Residual (m)
	Column	Row	Easting (m)	Northing (m)	Longitude (°E)			Latitude (°S)			
					Degrees	Minutes	Seconds	Degrees	Minutes	Seconds	
1	3380.92	78.98	448534.61	6877236.48	152	28	31.64	28	13	46.71	100.72
2	2808.98	891.02	406430.33	6816340.72	152	2	29.05	28	46	36.94	83.97
3	2964.01	1764.16	404351.69	6744223.32	152	0	50.14	29	25	39.34	88.07
4	3248.13	1939.02	418225.69	6727544.61	152	9	20.51	29	34	44.72	15.21
5	3335.94	603.01	439614.31	6835069.77	152	22	56.41	28	36	35.52	21.95
6	3190.01	1060	425858.01	6799268.53	152	14	21.47	28	55	56.19	68.85
7	3374.45	199.32	446760.51	6867461.37	152	27	24.94	28	19	4.1	14.75
8	3357.63	272.24	444905.75	6861687.5	152	26	15.84	28	22	11.44	27.39
9	3341.18	368.9	442805.26	6853994.39	152	24	57.3	28	26	21.1	21.39
10	3310.52	521.07	439228.23	6841962.05	152	22	43.51	28	32	51.5	25.22
11	3353.26	704.04	439388.25	6826687.32	152	22	46.49	28	41	7.85	68.17
12	3348.67	851.05	437368.67	6814895.08	152	21	29.73	28	47	30.67	81.54
13	3327.54	901.07	435528.33	6811019.13	152	20	21.06	28	49	36.28	46.33
14	3323.76	951.88	434627.69	6806831.09	152	19	46.95	28	51	52.19	85.40
15	3238.32	1079.06	428244.48	6797371.47	152	15	49.18	28	56	58.32	46.03
16	3325.28	1121.03	432630.76	6793234.13	152	18	30.31	28	59	13.6	46.22
17	3229.19	1286.12	425270.65	6780676.55	152	13	55.32	29	6	0.13	70.50
18	3269.49	1372.21	426486.23	6773215.43	152	14	38.51	29	10	2.79	118.31
19	3162.82	1789.08	415233.9	6740404.88	152	7	32.91	29	27	46.21	62.82
20	3209.02	1890.05	416649.76	6731881.92	152	8	23.13	29	32	23.44	24.49
21	3263.13	1930.19	419137.33	6728168.65	152	9	54.56	29	34	24.67	57.19
22	1457.52	704.89	332735.12	6842929.04	151	17	25.9	28	31	46.46	30.50
23	1736.35	1558.79	337987.21	6771268.13	151	20	2.08	29	10	36.43	91.03
24	1669.77	1494.74	334929.66	6777121.03	151	18	12.05	29	7	24.91	47.53
25	1849.78	1178.76	349017.24	6801167.88	151	27	4.82	28	54	30.15	35.35

Table 4.4 Model parameters

Note: Projection angle is the average of the applied projection to true north. Shear angle indicates the amount of angular distortion in the affine model, with higher angles generally being due to poorly spread GCP, but may also occur with off-nadir images. North angle is the average angle of the original image scene to true north.

Parameter		Path 89 scene	Path 90 scene		
			25 GCP	20 GCP	16 GCP
Cell size (m)	X	56.9150	56.8721	56.8838	56.7677
	Y	82.0073	81.9805	81.9636	81.9599
Angles (°)	Projection	-9.3467	-8.6697	-8.6716	-8.5648
	Shear	-0.0284	0.0282	0.0137	-0.0958
	North	-9.3104	-7.8974	-7.8992	-7.7930
RMS Residual (m)	X	20.31	43.58	37.34	36.88
	Y	13.71	44.49	27.59	28.16
	XY	24.51	62.28	46.43	46.40
Mean Absolute Residual (m)	X	16.17	37.10	32.91	31.90
	Y	12.28	36.53	21.92	23.68

Table 4.5 Residuals after removing five GCP for path 90 scene

Note: Disabled GCP are shaded grey.

GCP ID	Residual (m) with only 20 GCP enabled	GCP ID	Residual (m) with only 16 GCP enabled
1	129.91	1	123.75
2	82.36	2	158.95
3	64.05	3	113.83
4	33.21	4	29.45
5	29.79	5	29.33
6	69.07	6	52.24
7	28.4	7	24.98
8	38.45	8	37.05
9	22.79	9	22.64
10	24.72	10	26.26
11	60.63	11	65.66
12	74.11	12	76.08
13	40.14	13	40.49
14	88.18	14	89.48
15	44.72	15	55.33
16	42.85	16	39.75
17	67.96	17	66.63
18	128.7	18	124.43
19	66.08	19	64.29
20	24.54	20	20.03
21	35.97	21	37.97
22	20.9	22	301.37
23	122.25	23	123.99
24	17.41	24	270.99
25	50.92	25	195.41

To demonstrate the impact selected GCP have on rectification models, the set of GCP for the path 90 image was modified in two stages:

- five GCP with the highest residual errors were removed, then the affine model and residuals were recomputed with the 20 remaining GCP (see Table 4.5 column 2); and
- an additional four GCP with the highest residuals were removed before model parameters were recomputed with the 16 remaining GCP (see Table 4.5 column 4).

While the RMS residual and mean absolute residual statistics mostly decrease as GCP are removed (see Table 4.4), the residual errors for the disabled GCP increase significantly (see Table 4.5), demonstrating the significance of control at these locations. The mean absolute residual for Y does slightly increase when the GCP set is reduced from 20 to 16. This increase occurs because the removed points significantly change the overall spread of GCP across the image (particularly GCP 22, 24 and 25, which constitute the majority of points in the eastern portion of the image; see Figure 4.11a). The shear angle also degrades significantly since the removal of points effectively makes the georeference fail (see Table 4.4). The resulting affine transformation models also change slightly:

- with 20 GCP:

  E = 56.2334992 × C − 12.33834715 × R + 259,490.91117429

  N = −8.57644086 × C − 81.02962748 × R + 6,912,552.47139761

  C = 0.01737939 × E − 0.00264635 × N + 13,783.25317099

  R = −0.00183949 × E − 0.01206107 × N + 83,850.08579043

- with 16 GCP:

  E = 56.13459646 × C − 12.34154002 × R + 259,821.58553849

  N = −8.45425478 × C − 81.02540668 × R + 6,912,144.00392635

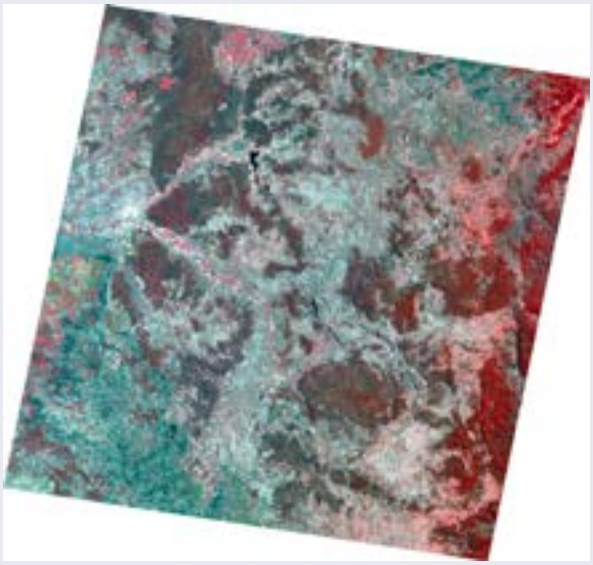
  C = 0.01741483 × E − 0.00265257 × N + 13,810.21933627

  R = −0.00181708 × E − 0.01206504 × N + 83,867.38393856

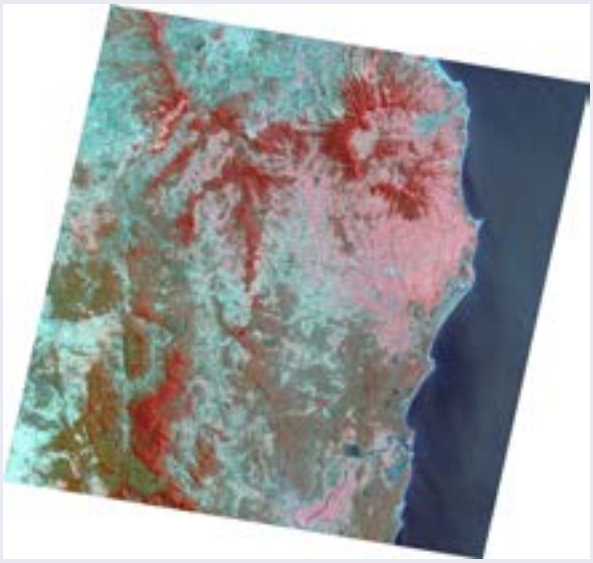
Figure 4.12 Resampled images

The affine transformation model compensates for rectangular geometric pixel dimensions in the original image and allows adjacent image paths to be resampled to a continuous grid.

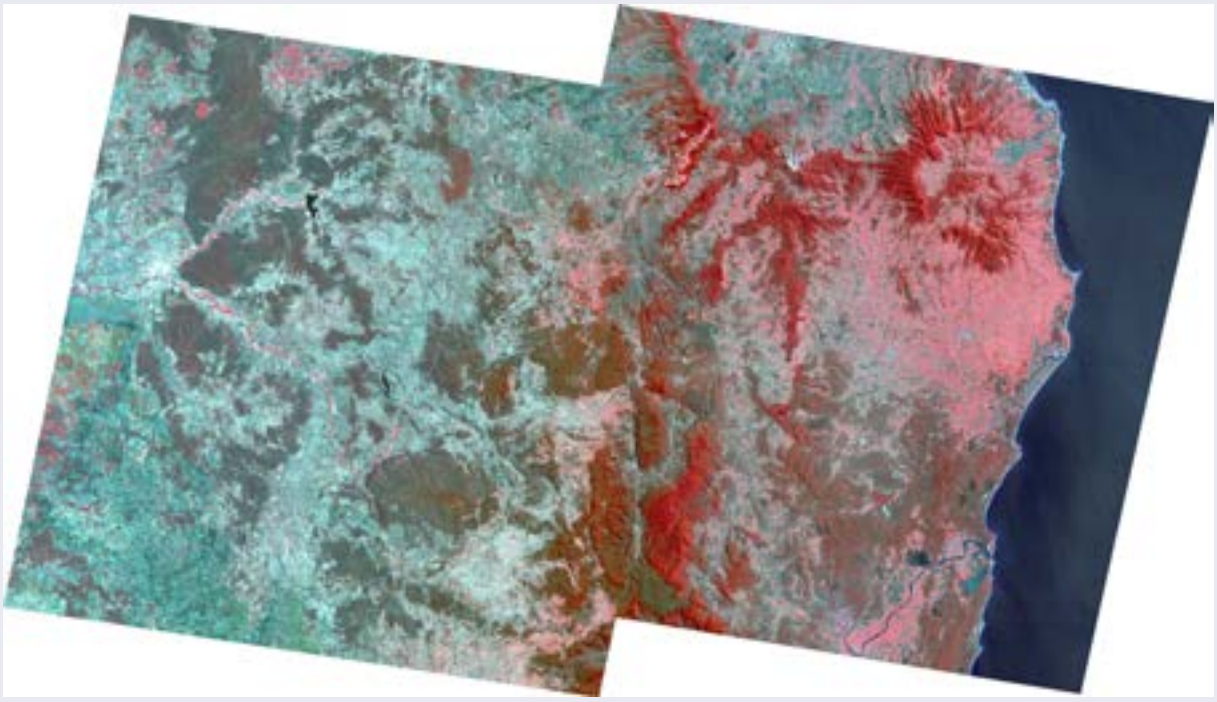
a. Path 90



b. Path 89



c. Mosaic of Paths 89 and 90



4.2.1 Method of least squares and polynomial models

If control points are well distributed over the image being rectified, and thus represent its geometry, the method of least squares can be used to compute a mathematical model between the two sets of coordinates. This model can then be used to convert any location represented in one geometric framework to a corresponding location in the other. In any distribution of samples, the fitted relationship will not exactly coincide with all points. The difference between a sample point and the fitted model is referred to as its residual error. The method of least squares allows a model to be derived in which the residual errors are minimised. For example, to convert from image coordinates (X,Y) to map coordinates (X',Y') we can develop the models F_x and F_y such that:

$$X' = F_x\left(X,Y\right) + e_x$$
$$Y' = F_y\left(X,Y\right) + e_y$$

where the residual errors e_x and e_y represent the difference between the fitted and actual values in the X and Y dimensions respectively. F_x and F_y can be conveniently computed using linear combinations of basic functions, such as polynomials, splines and other interpolation functions. In this case, the parameters of F_x and F_y can be fitted by least squares separately.

For a set of N ground control points, with image coordinates X_i,Y_i matching map coordinates, X_i', Y_i' for $i = 1, N$ then the least squares method computes the models F_x and F_y.

The least squares method computes the models F_x and F_y separately, so that if:

$$RMS_x = \left\|e_x\right\| = \sqrt{\frac{\sum_{i=1}^{N} e_{x_i}^{\ 2}}{N}}$$

$$RMS_y = \left\|e_y\right\| = \sqrt{\frac{\sum_{i=1}^{N} e_{y_i}^{\ 2}}{N}}$$

then the parameters can be found for F_x by minimising RMS_x^2 and the parameters for F_y can be found by minimising RMS_y^2 for N sample points.

A polynomial model of order n can be expressed as a linear combination of n basic functions:

$$h_j\left(X,Y\right)$$

where $j = 1, n$ with coefficients for these functions being a_j (see Section 3).

For example, in an affine model (order 3) the monomials would be:

$$h_1\left(X,Y\right) = 1$$
$$h_2\left(X,Y\right) = x$$
$$h_3\left(X,Y\right) = y$$

so the equation for an affine function can be expressed as:

$$F = a_1 \times h_1\left(X,Y\right) + a_2 \times h_2\left(X,Y\right) + a_3 \times h_3\left(X,Y\right)$$
$$= a_1 \times 1 + a_2 \times x + a_3 \times y$$

The generalised equations for F_x and F_y become:

$$F_x = \sum_{j=1}^{n} a_j \times h_j\left(X,Y\right)$$

$$F_y = \sum_{j=1}^{n} b_j \times h_j\left(X,Y\right)$$

where

a_j is the vector of coefficients for F_x, and
b_j is the vector of coefficients for F_y.

In vector notation, the solution to the least squares problem becomes:

$$\min\left\|X' - Aa\right\|$$

and

$$\min\left\|Y' - Ab\right\|$$

where

X' and Y' are the vectors of *TO* coordinates;
a and b are coefficient vectors for F_x and F_y; and
A is the scaled inverse covariance matrix, which is formed by evaluating the h_j functions over the (X,Y) GCPs. This is represented as:

$$\left[h_j\left(X_i,Y_i\right)\right]$$

where

j indicates the polynomial order; and
i indicates the GGP number.

The solution to the least squares problem then becomes:

$$\underline{c} = \left(A^T A\right)^{-1} A^T X'$$

and

$$\underline{d} = \left(A^T A\right)^{-1} A^T Y'$$

so the root mean square errors (RMS: that is the average internal errors of the model) are:

$$RMS_x = \left\| X' - Ac \right\|$$
$$RMS_y = \left\| Y' - Ad \right\|$$

At individual sample points the residual errors are the components of the vectors:

$$e_{x_i} = X'_i - Ac$$
$$e_{y_i} = Y'_i - Ad$$

and the inferred relationship between any pixel in the *FROM* coordinates (*X*,*Y*) and its corresponding value in the *TO* coordinates (*X'*,*Y'*) becomes:

$$X' = \sum_{j=1}^{N} c_j \times h_j\left(X,Y\right)$$
$$Y' = \sum_{j=1}^{N} d_j \times h_j\left(X,Y\right)$$

When all sample points are representative of the relationship between the two geometries, and this can be expressed as one of the available polynomial models, the residual errors should be relatively small and follow a normal distribution. However, if a point is not representative, that is, it may be incorrectly located, it will bias the model and have a large residual error. Such outlier points can thus be highlighted and checked by analysing the residual errors. A systematic error in all or a large proportion of the points will not be detected directly by the residual errors since the model is derived from the data points and if a large number are in error there is no reference against which this can be measured. Such errors can be discovered using other techniques as discussed in Sections 4.2.3 and 4.2.4.

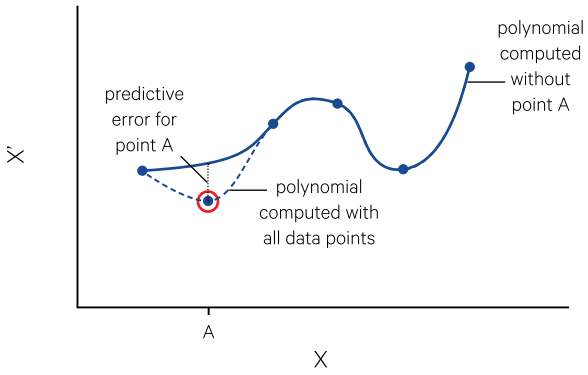
Conquest is easy. Control is not.
(William Shatner)

4.2.2 Predictive error

The predictive error statistic uses the variance in the fitted data to compute the root mean square (RMS) error expected at a given point if the model was fitted without that point. This value indicates the difference between a point's fitted value when it is used to fit the model and the fitted value it would have if the model is fitted without it. The predictive error value of a point highlights its importance to the current model as illustrated in Figure 4.13.

Figure 4.13 Predictive error

Plot of actual (*X*) and modelled (*X'*) RMS error values. The predictive error of a particular point (such as A) can be determined by comparing its estimated value when the model is based on all control points (solid line) and its modelled value when that point is omitted (dotted line).



Source: Harrison and Jupp (1992) Figure 39

The *X* variance at point *i* can be computed as:

$$VAR\left(X_i\right) = VAR\left(X\right)h^T\left(A^TA\right)^{-1}h$$

so the predictive error, *p* equals:

$$\frac{VAR\left(X_i\right)}{VAR\left(X\right)} = h^T\left(A^TA\right)^{-1}h$$

where

$VAR\left(X_i\right)$ is the variance of point *i*;
$VAR\left(X\right)$ is the variance of all points;
h is the vector of monomials defined in Section 4.2.1; and
A is the scaled inversed covariance matrix (see Volume 2A).

At any point, the predictive error, p_i, is:

$$p_i = p\left(X_i,Y_i\right)$$

The *A* matrix can be computed using X_i and Y_i so this equation produces a slightly less stable estimate of predictive error. The predictive error could also be estimated for every pixel in an image using a slightly different equation for predictive error:

$$p = 1 + h^T\left(A^TA\right)^{-1}h$$

where *A* is computed without the location for which the error is being computed.

The model predictive error over all points can be computed as:

$$\sum_{i=1}^{N} p_i e_i^2$$

where

p_i is the predictive error weight for point i;
e_i is the residual error at point i; and
N is the number of control points.

The optimum model is generally the one with minimum predictive error as detailed in Section 4.2.4. Another statistic that is sometimes used to determine the optimum model is the Generalised Cross-Validation (GCV), which is detailed in Volume 2X—Appendix 7.

4.2.3 Detecting outlier points

As discused above, for any pair of spatial datasets, such as EO images, transformations from the coordinate system of one dataset to the other can be derived, which enable any feature in one dataset to be located in the other. In the case where all transformations can be represented by a simple affine transformation, then if the transformation is exactly known in one direction (such as from dataset 1 to dataset 2), it could be inverted to obtain the corresponding transformation in the reverse direction (from dataset 2 to dataset 1).

However, transformations are never known exactly. Normally, the sets of GCP are used to estimate the coefficients by least squares or some other measure of goodness of fit between the original coordinates and those predicted by the transformation (see Section 4.2.1). GCP, however, are normally subject to error or 'noise' and in this process any outliers that are present will play a disturbing role. While there will always be some error or noise level that limits the locational accuracy of GCP, and therefore also limit the accuracy of the estimated coefficients of a transformation, 'outliers' are points that are inaccurate well beyond this base of error or 'noise' (see Volume 2X—Appendix 8 for more details).

One approach to identifying outlying points in a set of paired GCP assumes that an affine transformation exists between the two sets of coordinates. As detailed in Section 3.2.1 and Appendix 6, an affine transformation accounts for differences between the input coordinates that are due to shift (different origins), scale differences, rotation of coordinate axes, reflection and skew. When an affine model is fitted to a set of GCP pairs, there should only be two intrinsic components if each pair refers to the same basic set of X,Y locations—one component to identify the X location of each point, the other to identify its Y location.

For example, assume there exists a set of GCP in map coordinates and a corresponding set in image coordinates relating to the same locations. While this example involves two sets of coordinates (or four data columns), both sets describe the same set of locations. Accordingly, in terms of intrinsic dimensionality, they only describe two dimensions: X and Y.

Each control point can be considered to have n coordinate pairs where n is greater than or equal to 2. Thus, the input GCP coordinates can be treated as an $m \times 2n$ matrix, where m is the number of control points. However, $2n$ columns only represent two data dimensions since each pair is essentially a different measure of the same X,Y position on the Earth's surface. If each column of the matrix is normalised (the mean of the column is subtracted from each value so that the new mean equals zero then the set is rescaled so that the variance equals one) then the resulting n-dimensional matrix can be transformed by Principal Components Analysis (see Volume 2C) to reduce it to its two most significant dimensions—that is, the X and Y dimensions. If the input data contain more than two significant dimensions, as indicated by the eigenvalues of the correlation matrix, then the input data may contain erroneous points. It may be informative to analyse the coefficients of the PCA (or V) matrix and rotate the matrix of GCP coordinates into the PC space (U matrix—see Volume 2C).

If the data have been successfully reduced to two significant dimensions, a crossplot of these two dimensions should represent the spatial arrangement of the control points in a relative sense; the scaling and orientation of axes will be different and the plot may be a 'mirror' image of the map if one or both axes were reversed in the PCA. This is a good way to check for errors, especially systematic ones such as an incorrect starting position for a sub-image.

The lower dimensions (PC3+) should only represent random noise, so if any other dimensions are significant at this stage some points may be in error. PCs 3 and 4 can be crossplotted to give a diagrammatic representation of the error, or residual, at each point. The errors should fall inside a circle centred at zero with a radius of $\frac{2}{\sqrt{m}}$. This is a useful method for identifying regional or systematic errors.

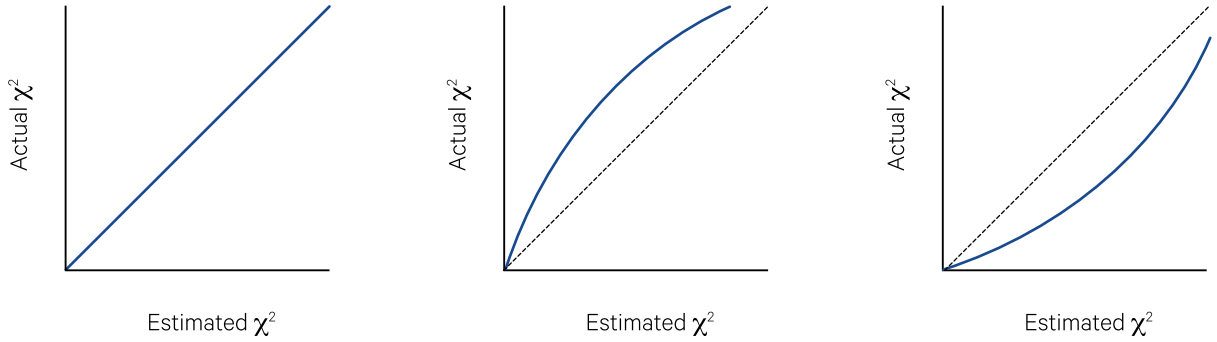
Error statistics can be determined for each point. For example, a predictive error weight value would indicates how important a point is in the current model—a new model computed without that point would potentially give an error of this value in predicting the point's position. A point is important to the model if the predictive error weight is high—this would indicate that it is probably the only GCP in that region of the image.

Figure 4.14 Q-Q plot

a. If actual and estimated chi-square values agree for all control points then the slope of Q-Q plot equals 1.

b. If the slope is greater than 1, then some outliers exist in the set of control points.

c. If the slope is less than 1, there are too few control points.



Source: Harrison and Jupp (1992) Figure 40

The number of points to delete can be determined by using a quantile-quantile (Q-Q) plot of the actual chi-square values against the predicted chi-square values (see Figure 4.14). The predicted chi-square value for a sample is determined from the percentage of samples that have errors less than that sample. If the error distribution is chi-square then the plot (with actual chi-square on the y axis) will have a slope close to one. If the plot kinks up from the diagonal, the high actual chi-square value(s) are much greater than predicted for that number of samples so the points with the largest chi-square values should be removed (subject to the size of the point's predictive error weight). If however the plot kinks down from the diagonal, this generally indicates that too many points have been removed. It is preferable to have the slope value greater than one, rather less than one. It should be noted though that the plot always has a slope less than one in the very low chi-square values. This is due to the non-random distribution of errors within the size of a pixel since positions are often truncated to a half or quarter pixel. The noise to signal ratio for the plot should also be low (less than 10%).

It is important that this process be used correctly to achieve a reliable and accurate rectification model. The Q-Q plot is an extremely useful indication for the number of points required. As noted earlier, a smaller number of well-chosen and well-distributed points will be more useful than a large number of points in a very localised area. This will also be indicated by the Q-Q plot since 100 points can produce a slope less than one if they are very clustered whereas 30 points can still contain outliers (that is, have a slope greater than one) if they are well-distributed. When points are highlighted as outliers, all possible sources of error should be investigated, and their predictive error weight checked, before they are removed. Only those pixels at the tail of the plot should be deleted—removal of other points may increase the slope of the line but will degrade the value of the model. If chi-square values are still high when the Q-Q plot slope is nearing one, more control points should be selected. An inadequate set of control points will only produce an inadequate rectification model, and deliver inadequate results when it is applied.

The points which are identified as outliers may be due to typographical errors in entering data, error in matching a point on both map and image, incorrect location of a pixel during interactive editing, or misreading coordinates on the map. Systematic errors may occur, for example, if an incorrect starting pixel and line is entered for one or more image subsets. Such errors can invalidate the assumptions of an affine transformation between the coordinate systems and may not be revealed until an attempt is made to fit a polynomial model.

4.2.4 Selecting the optimum model

Before computing a polynomial transformation between two sets of coordinates, outliers should have been removed as outlined in Section 4.2.3. Transformation models may be computed to convert from image coordinates to map coordinates, from map coordinates to image coordinates or from coordinates in one image to those in another. Polynomial models are generally one of:

- affine (3 coefficients);
- bilinear or ruled (4 coefficients);
- quadratic (6 coefficients); or
- cubic (10 coefficients).

The geometric factors accounted for by each of these models are described in Section 3.2.

If the GCP set has been checked for outliers, there is usually no need to delete additional control points at this stage. However, if large errors occur at a point when both forward and inverse models are fitted (that is models are fitted in both directions) in the model with minimum predictive error, the point may be deleted.

If systematic errors are involved in the points, these may not have been detected during outlier checking but will manifest as large error values in the modelling results. If all pixels have large (absolute) error values they can be checked for systematic error by annotating the error sign on a plot of their location. For example, when the starting pixel and line of a subset has been incorrectly used in a multi-subset rectification, a pattern of error will occur over the image area with the sign (that is positive or negative) of the errors changing systematically about the subset. In this case all points from that subset need to be adjusted to be relative to the correct starting pixel and line value.

If individual points have large error values they should be carefully checked with reference to the Q-Q plots and their predictive error values considered. While deleting control points at this stage may reduce the predictive error values and thus appear to produce a more accurate model, in reality the model will be less accurate for the points that were deleted, and only 'accurate' for those regions of the image represented by the remaining control points.

In the initial rectification of an image to a map base, the affine model should be used to verify that the assumed pixel size is valid. When the pixel size is satisfactory, each of the available models is selectively fitted to the input data. The optimal model is the one that produces the minimum predictive error value. Different order models can be selected as optimal for X and Y values using this criterion. With full scene Landsat data, a bilinear or quadratic model frequently gives minimum predictive error in the X values while affine or bilinear gives minimum predictive error in the Y values. For small area subsets however, an affine model should be adequate for both X and Y.

After a polynomial model has been fitted between two sets of control points, the fitted model may be tested on another pair of files. Obviously, the test 'from' points need to be the same coordinate type as the fitted 'from' points and the test 'to' points must be the same coordinate type as the fitted 'to' points. The testing option allows the model to be checked using additional points for which accurate map and image coordinates are available. Test control points should be selected in the same way as the original points, and nominally transformed using the same parameters (including the adjusted pixel size).

Control what you can,
acknowledge what you cannot.
(Jim Stovall)

This testing can also be used to compare predicted with actual values for points that were deleted during the rectification process. This is especially important for any points that were deleted after outlier detection. If this process has been used correctly, so that the slope of the Q-Q plot is close to one, there should be no need to delete additional points at this stage.

The polynomial model in the three-stage registration process is direction-specific. While the functions may be theoretically invertible, error assumptions are different in the two directions so it is more appropriate to compute the inverse transformation separately. Typically the rectification is initially modelled from image to map to produce an accurate estimation of pixel size. A map to image transformation may then be computed by using the (nominally transformed) map coordinates as the 'from' points and the (nominally transformed) image coordinates as the 'to' points. In this case the T_1 transformation parameters would be derived from the 'from' coordinates to convert original map coordinates to nominally transformed map coordinates (see Section 3.3). The polynomial model, P, converts from (nominally transformed) map to satellite or scanner model coordinates. The T_2 model parameters are obtained from the 'to' coordinates to convert from satellite or scanner model coordinates to image coordinates. The three-stage registration process can also be used to convert from image to image or map to map. The resulting stages can be similarly retained for subsequent use in image resampling (see Section 5).

4.2.5 Checking the extent of control

Predictive error has been introduced in Section 4.2.2. At any point associated with a rectification model, this is the root mean square (RMS) error expected at the point due to the variance in the fitted data. It can be used to assess the degree of control for existing control points or to plan further ground survey to establish better control.

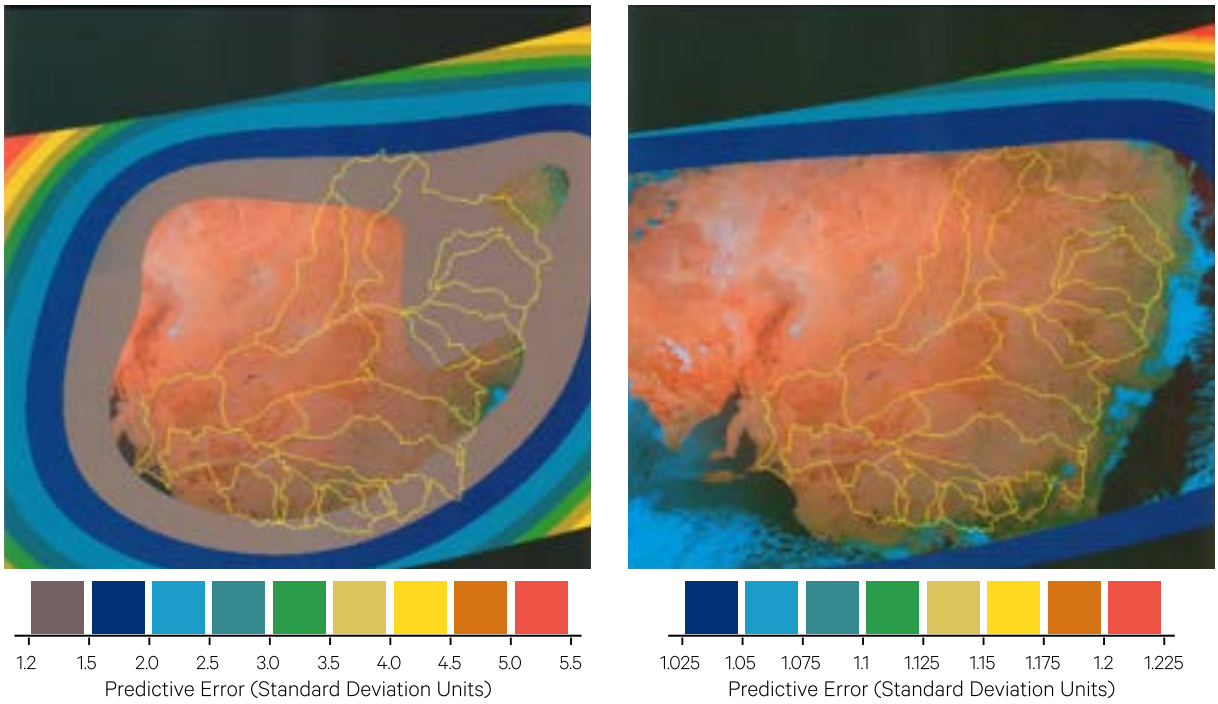
The predictive variance of a polynomial model, which converts from an image to some other geometry, can also be computed for each pixel, not just the GCP (see Volume 2X—Appendix 7). In this case, an output channel could be written with scaled values that indicate the predictive variance, in units of standard deviation, which is expected when applying the model at each pixel location in the image.

Figure 4.15 Predictive variance image

Predictive variance channel on AVHRR colour composite, acquired on 25 December 1987. Catchment boundaries are shown as yellow lines. Note that colour scales for predictive variance are different in each image.

a. Higher predictive variances result when the cubic transformation model is derived from 32 visually identified GCP that are poorly distributed, especially in northern catchments.

b. Low predictive variances over the Murray Darling Basin indicate good control for a cubic transformation model derived from 150 GCP. These points were systematically computed by the receiving station and well distributed across the image.



Source: McVicar and Mashford (1993) Plates 2 and 3

The predictive variance channel provides a useful tool for pictorially assessing the stability of a particular model. The output channel will contain low values in areas of good control, that is, in areas where an effective network of control points has been selected. Typically a predictive variance channel will have patterns of low values along coastlines, roads and similar features where control point selection is relatively easy. However, in parts of the image where there is poor control, the error involved in using the model is expected to be high and this will be indicated by high values in the predictive variance channel.

Figure 4.15 illustrates two predictive variance channels for an AVHRR image of Southeast Australia. Two sets of control points were used to develop two cubic registration models for the image. Figure 4.15a is based on a model developed from a poorly-distributed set of manually-selected control points, which were principally located along the coastline. The model for Figure 4.15b is based on a large number of points in a grid pattern across the image. These image channels graphically illustrate the areas that would have greater registration error if the models were used to extrapolate locations away from the control set.

The rectification statistics relating to each set of GCP are summarised in Table 4.6. and Table 4.7. Note that while the RMS errors indicates that the cubic model would be the most suitable for both sets of GCP, for the 32 visually-identified GCP set the predictive error and log (GCV)[6] results are minimised with the bilinear model. This example demonstrates the internal nature of the RMS statistics and highlights the value of other statistics (McVicar and Mashford, 1993).

Table 4.6 Error associated with polynomial models fitted to 32 GCP

Model order	RMS error		Predictive error		Log (GCV)
	X	Y	X	Y	
Affine	1829.333	1053.356	2088.768	1161.287	3.9176
Bilinear	1473.862	976.623	1793.894	1111.154	3.8244
Quadratic	1423.967	920.401	1961.649	1132.479	3.8886
Cubic	1333.328	867.728	2531.653	1285.987	4.0767

Source: McVicar and Mashford (1993) Table 5

[6] The GCV (Generalised Cross-Validation) statistics is detailed in Volume 2X—Appendix 7.

Table 4.7 Error associated with polynomial models fitted to 150 GCP

Model order	RMS error		Predictive error		Log (GCV)
	X	Y	X	Y	
Affine	4416.827	3162.452	4538.180	3261.620	3.2458
Bilinear	3426.231	633.530	3532.941	661.020	2.3252
Quadratic	1229.079	369.192	1295.138	397.831	1.5633
Cubic	332.798	33.146	375.287	35.824	-0.2529

Source: McVicar and Mashford (1993) Table 5

When computing a predictive variance channel, the parameters for the maximum model (that is, the highest order model available) will most clearly delineate the spatial variation in error. In most cases this will be the cubic model (provided more than 10 control points are being used to fit the model). Since the predictive variance surface is a function of the model, it indicates the variance that is inherent to the model due to the variance in the fitted parameters.

A low order polynomial is less sensitive to error in its coefficients than a high order polynomial. However, it is also dependent on control point distribution. If the cubic model is used to compute the predictive variance surface, the result indicates whether the control is good enough to fit the model stably even if it is cubic. If the optimal model is less than cubic, its stability will be greater.

The predictive variance values are computed in standard deviation units of pixels so need to be rescaled to form an image channel. As in the case of other transformation programs that require rescaling, this may require two iterations—once (using an arbitrary range) to determine the actual range and then secondly to use the actual range to optimally fill the output channel data range (see Volume 2C). For most rectifications, a reasonable initial range would be 1–10 units.

4.3 Further Information

Jensen (2016) Chapter 7

Richards (2013) Chapter 2

4.4 References

ACTEW Water (2013). *Research and Development—Vegetation Structure and Biomass Estimation in the Area of Inundation of the Expanded Cotter Dam*. Report for ASTRIUM Services on ACTEW Water Project Number 82808057, August 2013.

Harrison, B.A., and Jupp, D.L.B. (1992). *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual*. MPA, Melbourne.

Hutchinson, M.F. (1989). A new method for gridding elevation and streamline data with automatic removal of pits. *J. Hydrology* 106, 211–232.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Jupp, D.L.B., Guerin, P.R. and Lamond, W.D.D. (1982). Rectification of Landsat data to cartographic bases with application to the Great Barrier Reef region. *Proc. URPIS 10,* Aust. Urban and Regional Information Systems Assoc., Sydney, pp. 131–47.

Lewis, A., Wang, L.-W., and Coghlan, R. (2011). *AGRI: The Australian Geographic Reference Image. A Technical Report*. GeoCat#72657. Geoscience Australia, Canberra. ISBN 978-1-921954-47-4.

McVicar, T.R., and Mashford, K.E. (1993). *Establishment of Control Points for Accurate Co-registration of AVHRR Data for the Murray-Darling Basin*. Technical Memorandum 93/7, CSIRO Division of Water Resources.

Richards, J.A. (2013). *Remote Sensing Digital Image Analysis: An Introduction*. 5th edn. Springer-Verlag, Berlin. ISBN 978-3-642-30061-5

Resampling and Registration

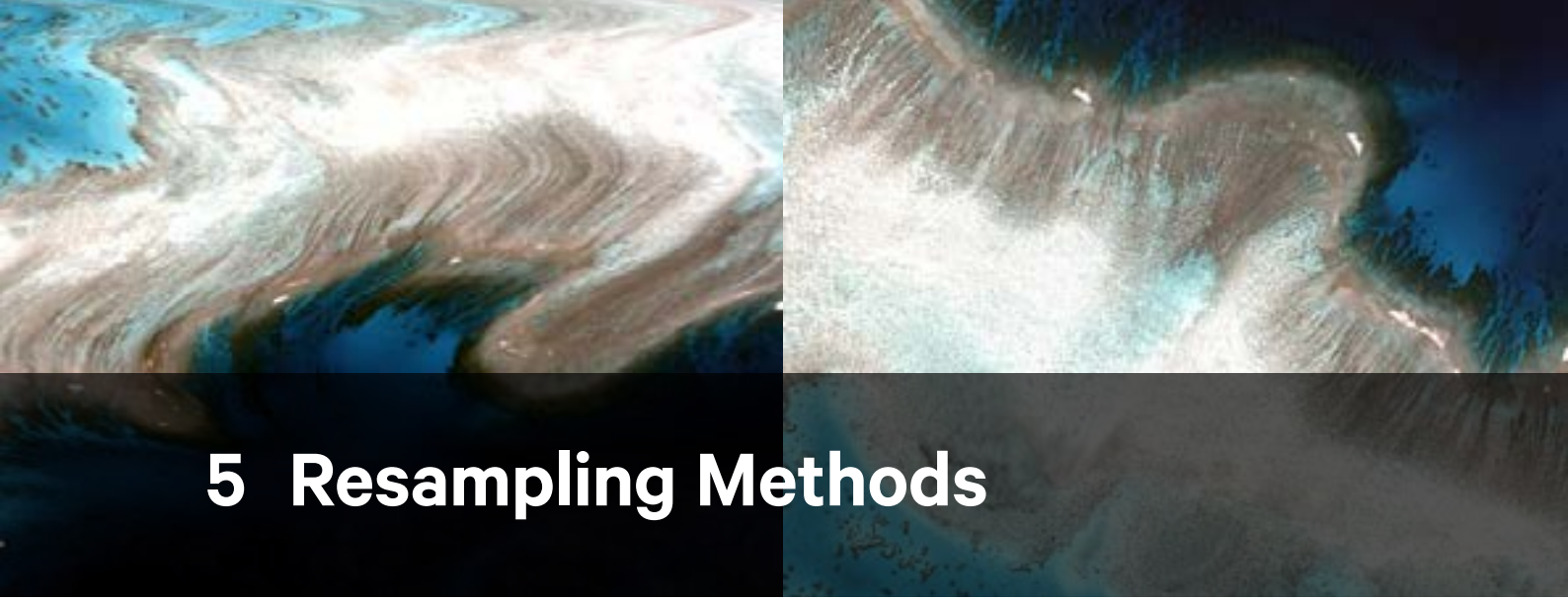
Once a precise registration model exists between two coordinate systems, a digital image can be resampled to match another coordinate system. Image resampling involves rearranging the spatial positions of pixels of an image, and usually changes the geometry of the image. This process involves the definition of a new coordinate system, which in EO usually matches an image to a map base or another image, and then assigns values to each pixel in this new system (or output image) using pixel values from the input image. To be able to accurately convert from locations in the new grid (or output image) to locations in the input image requires an accurate registration model between the new geometry base and the original image coordinates. Section 3.3 above describes the recommended three-stage registration process, with nominal models being used before and after a polynomial to convert from one coordinate space to another. For image resampling, then, these three transformations need to be sequenced to convert **from** the output grid **to** the input grid.

Contents

5 Resampling Methods 95

6 Registration 111

Background image on previous page: Landsat-5 image of Lake Eyre, South Australia, filling on 9 May, 2009, displayed using bands 5, 4, 2 as RGB and overlaid with contour lines (black), roads (red), and place names. **Source:** Norman Mueller and Erin Teifer, Geoscience Australia



5 Resampling Methods

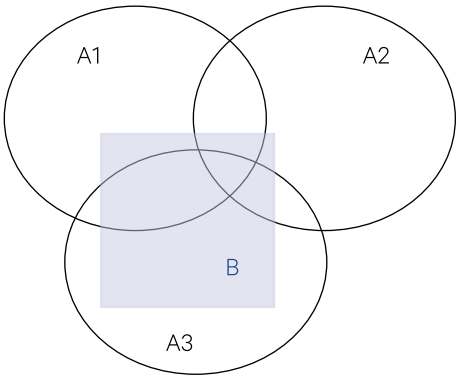
When an image is changed to have a different geometry, the resolutions of the input and output images need to be considered carefully. EO pixel data represent radiance integrated over a particular volume of the Earth's surface. As summarised in Section 2, different pixels in the image can represent volumes of different sizes depending on factors such as position in the scan line. The resampled pixels are normally associated with regions that have a different size and shape from the original pixels, that is, the locations of resampled pixels rarely coincide with the locations of the imaged pixels. The basic problem in resampling is to decide what data value(s) should be attached to the resampled pixels.

Figure 5.1a shows three optical pixels on the Earth's surface (see also Figure 2.2). Each pixel has a unique value, which represents the spatial integral over the corresponding ground area. In the context of this example, the problem of resampling then is to determine a value for B based on the values of A1, A2 and A3. The process of resampling involves magnification and/or 'minification' of the original image data range, as illustrated in Figure 5.1b, which necessitates either interpolation and/or subsampling of the image values. Whereas sub-sampling methods only use actual values of pixels in the original image, interpolation methods derive new values from the original image pixels.

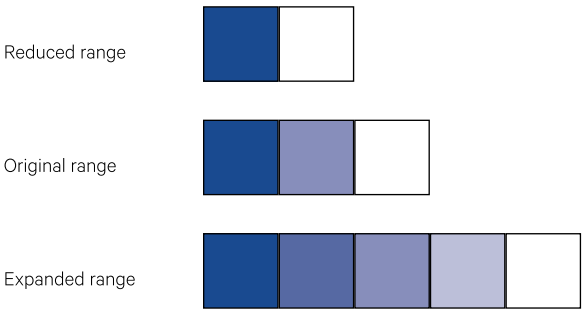
An example of interpolation is shown in Figure 5.2 using a one-dimensional dataset. Interpolation effectively attempts to convert a discrete sample set into a continuous function. In image resampling, this involves fitting a continuous function through selected sample pixels. This process implicitly assumes that the continuous function has been distorted by the sampling process and has the net effect of a low pass filter. In image resampling, the interpolation process determines the values at locations that lie between the observed pixel values. This implicitly assumes that a continuous function (the Earth's surface) has been sampled to derive the discrete values of image pixels.

Figure 5.1 The resampling problem

a. A new value for the resampled pixel, B, needs to be determined for the area of overlap between two or more optical pixels (labelled A1, A2 and A3).



b. Magnification and 'minification' of image scale require interpolation and/or sub-sampling of the original pixel values.



Source a. Harrison and Jupp (1992) Figure 42. b. Adapted from: *http://www.ldv.ei.tum.de/uploads/media/Vorlesung_3.4_Resampling.pdf*

Background image: CASI airborne imagery of Heron Island, Queensland, acquired in June 2002 (see Excursus 2.1). A portion of the original image is shown on the left, and the rectified image, resampled to 1 m grid and flipped vertically, is shown on the right. **Source:** Karen Joyce, James Cook University

Different resampling strategies can be used to define the way pixel values in the input image are transferred to appropriate locations in the output image. Commonly used methods of resampling the input image to a new geometry include:

- sub-sampling methods:
  - along-line resampling—rewrites the image to a new coordinate system, or grid, which has its image lines oriented the same way as the image lines of the original image. This mode of resampling is computationally simple but does not allow rotation of the image. Thus if two different images of an area are resampled this way they will not necessarily overlay. This mode of resampling is sometimes used by hardcopy algorithms to re-proportion imagery on-the-fly (see Section 5.1); and
  - nearest neighbour resampling—selects the closest pixel in the input image for each output pixel and transfers its value(s) to the output image. When resampling onto a finer mapping grid than required, the resampled image data can be blocked into larger pixels which represent an average of the original pixels weighted by their extent of overlap with the final larger pixel (see Section 5.2); or
- interpolation methods:
  - bilinear interpolation resampling—interpolates new output pixel value(s) by linear interpolation on the four closest pixel neighbours (within a 2×2 window) in the input image (see Section 5.3); and
  - cubic convolution—interpolates, using nested cubic polynomial models, new output value(s) from the closest 16 neighbouring pixels (within a 4×4 window) in the input image (see Section 5.4).

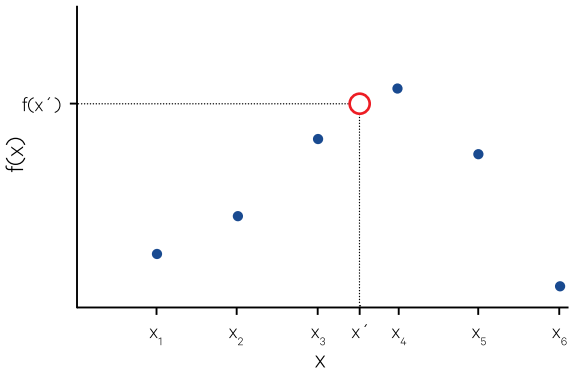
Imagery is powerful. Imagery is provocative—satellite imagery much more so because it is from space, and it allows us to get this perspective that we don't have to have otherwise.
(Sarah Parcak)

The last three of these resampling methods are compared in Figure 5.3 in terms of their interpolation functions.

The input image pixels involved in each output pixel computation are illustrated in Figure 5.4 for nearest neighbour, bilinear interpolation and cubic convolution resampling. The advantages and disadvantages of these methods are compared in Table 5.1. When the resampling operation essentially just changes the pixel size, all methods can deliver adequate results. However, when this rescaling process is combined with rotation, the results of nearest neighbour and bilinear resampling are far from ideal.

Figure 5.2 Interpolation example

Interpolation of an intermediate value (x') in a one-dimensional dataset (x) assumes that a continuous function (f(x)) can be fitted to the set of sample values.

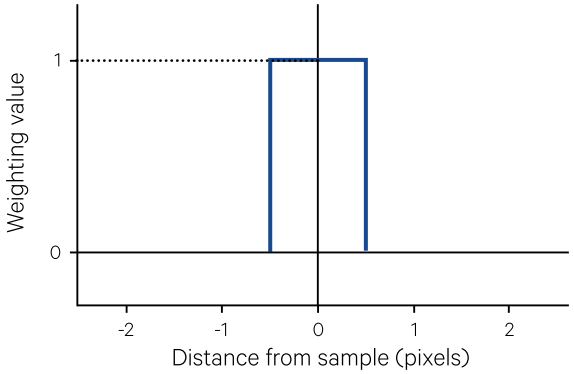


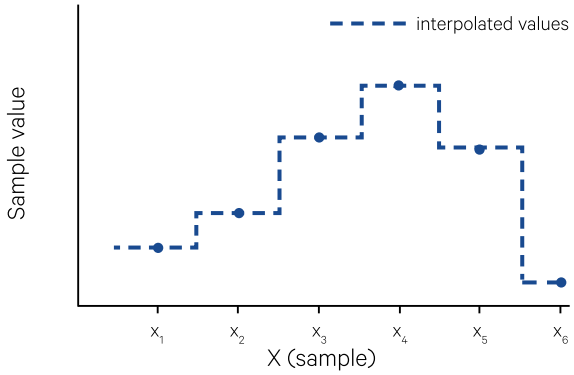
Adapted from: NIblack (1986) Figure 68

Figure 5.3 Interpolation functions for resampling methods

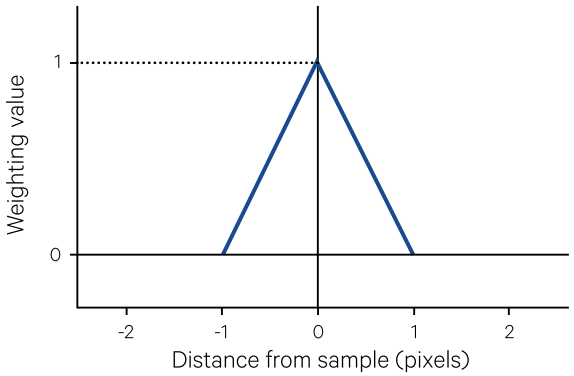
Using the one-dimensional dataset from Figure 5.2, the weighting functions (left) and interpolation results (right) are illustrated for three resampling methods.

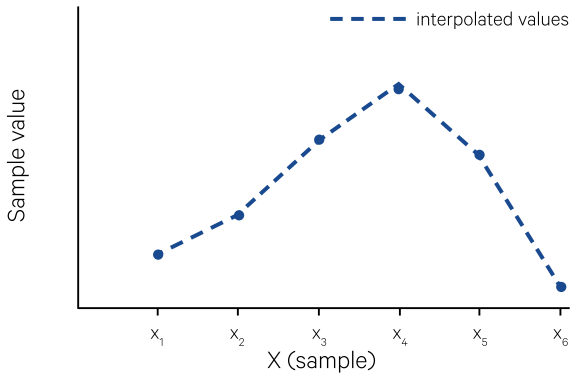
a. Nearest neighbour resampling



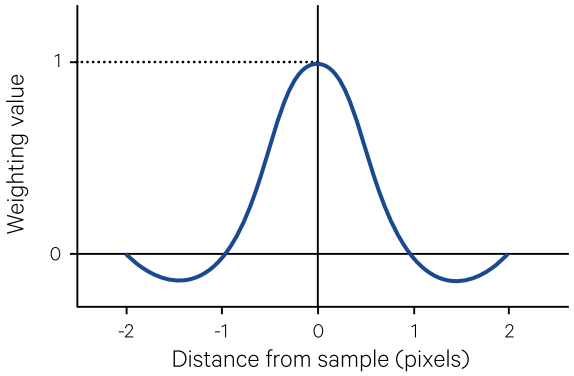


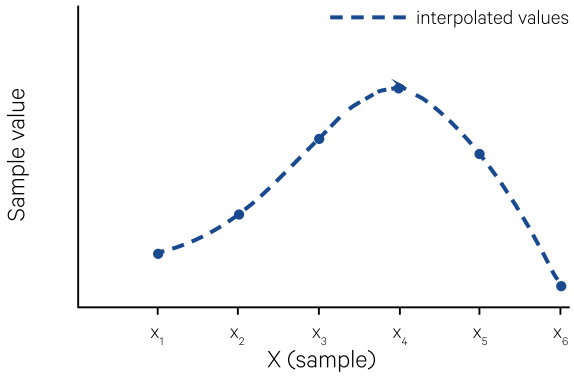
b. Bilinear interpolation resampling





c. Cubic convolution resampling



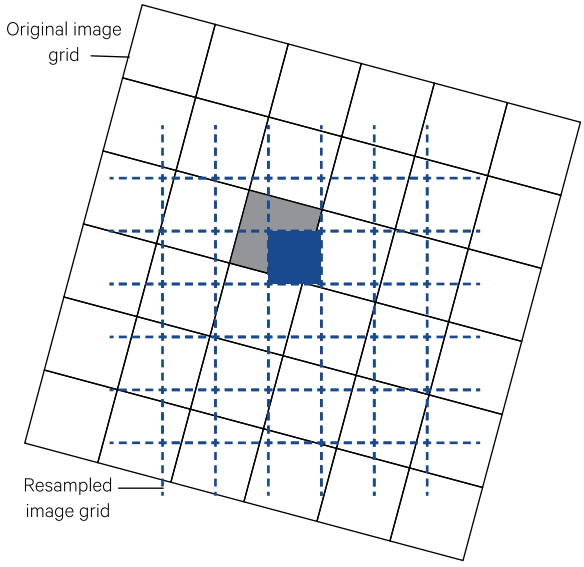


Adapted from: Niblack (1986) Figure 69

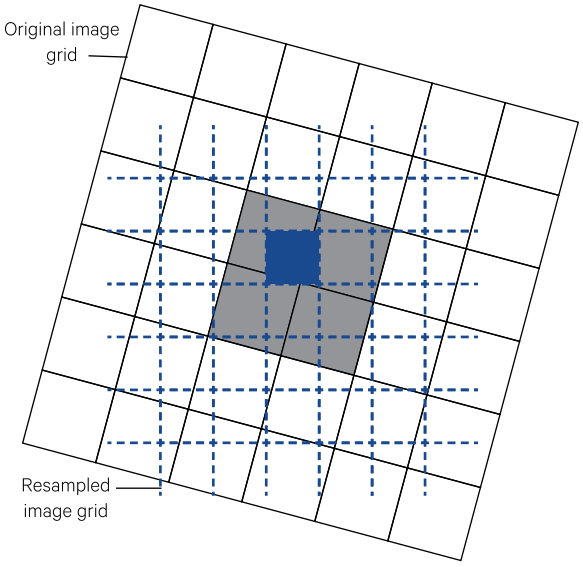
Figure 5.4 Resampling methods

This diagram shows the original image grid as black and the resampled image grid as blue. One example pixel in the resampled image is coloured blue. Those pixels in the original image whose values contributed to the example pixel are shaded grey.

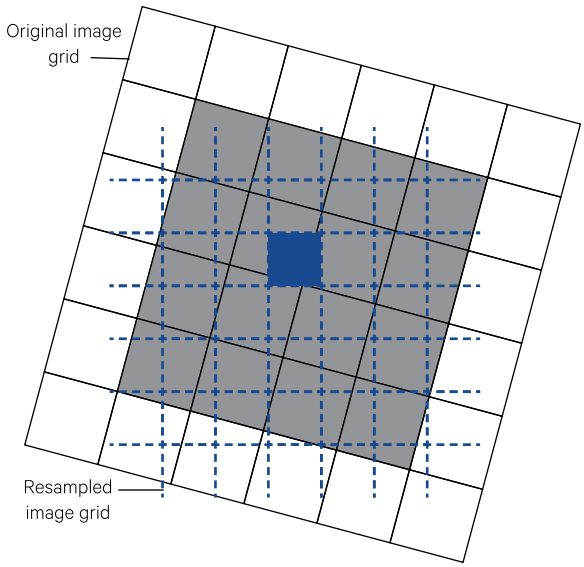
a. Nearest neighbour resampling uses the nearest input pixel value



b. Bilinear interpolation resampling interpolates from the four nearest input values in 2x2 window using a bilinear function



c. Cubic convolution resampling interpolates from 16 nearest input values in 4x4 window using a cubic function.



Adapted from: *http://www.tankonyvtar.hu/hu/tartalom/tamop425/0033_SCORM_MFGGT218-EN/sco_01_02.scorm*

Table 5.1 Resampling methods

Method	Process	Advantages	Disadvantages
Nearest neighbour	For each output pixel, uses value of pixel with closest proximity in the input image.	Computationally efficient. Useful for comparing field data with image values or processing categorical data.	Creates 'step' effect along curved edges and lines in image. Features may be mislocated by up to half a pixel.
Bilinear interpolation	For each output pixel, computes weighted value of four nearest neighbouring pixels in input image.	Generates spatially accurate result, especially when changing image spatial resolution.	Blurs sharp boundaries..
Cubic convolution	Uses cubic function to compute output value of each pixel from the 16 nearest neighbouring pixels in the input image.	Maintains mean and variance of original image data more closely than bilinear interpolation. Tends to sharpen image and smooth out 'noise'. Appropriate when resampling to a very different spatial resolution image.	Generates images with more natural appearance. Computationally expensive.

An EO image sensor samples the ground radiances of discrete areas (pixels) at often varying ground spacing and the geometry of the scanning process is different from that used in image analysis (see Section 2). The acquired image is therefore usually a distorted representation of the imaged scene with some added noise. Resampling aims to undo this distortion, but without adding noise. Interpolation methods reduce local image variance, which is sometimes interpreted as reducing the information content of the data. However, considering the effects of the sampling process involved in image acquisition (see Volume 1B—Section 2), it could also be argued that the sampling of the Earth's surface introduced local 'noise' variance that was reduced by the interpolation—that is, the interpolated values may more closely represent the imaged scene.

The effects of selected resampling methods are shown in Figure 5.5. In this set of example images, the differences between bilinear resampling and cubic convolution resampling are difficult to detect visually, but are highlighted in their difference image (see Figure 5.5f). Since nearest neighbour resampling retains the integrity of image radiometric values, it is often preferred when subsequent analyses are planned, especially when these involve validation with ground measurements or processing of categorical data. By contrast, bilinear and cubic convolution resampling necessarily modify the original image values. The latter are sometimes preferred to avoid the possibility of aliasing artefacts, however, these effects can be overcome with nearest neighbour by using the picking and binning approach in post-processing (see Section 5.2).

All resampling methods can degrade the input imagery in various ways. Discrete interpolation (nearest neighbour or along-line resampling) can introduce aliasing error, and an interpolation function (bilinear interpolation or cubic convolution resampling) can attenuate higher spatial frequencies that occur in the original image scene (Moik, 1980). Depending on the type of resampling method used, these degradations may appear as blurring and/or discontinuities in linear features or spatial boundaries.

He who seeks for methods without having a definite problem in mind seeks in the most part in vain.
(David Hilbert)

Figure 5.5 Effects of resampling methods

This Landsat-8 OLI image of Brisbane (acquired on 13 June 2018 and displayed using bands 6, 5, 3) was resampled from a UTM grid (WGS84) to the Albers Equal Area (AEA) projection (GDA94) using different resampling algorithms. The differences between bilinear and cubic convolution resampling are difficult to detect visually so have been highlighted in a difference image.

a. Original image (UTM grid; WGS84)

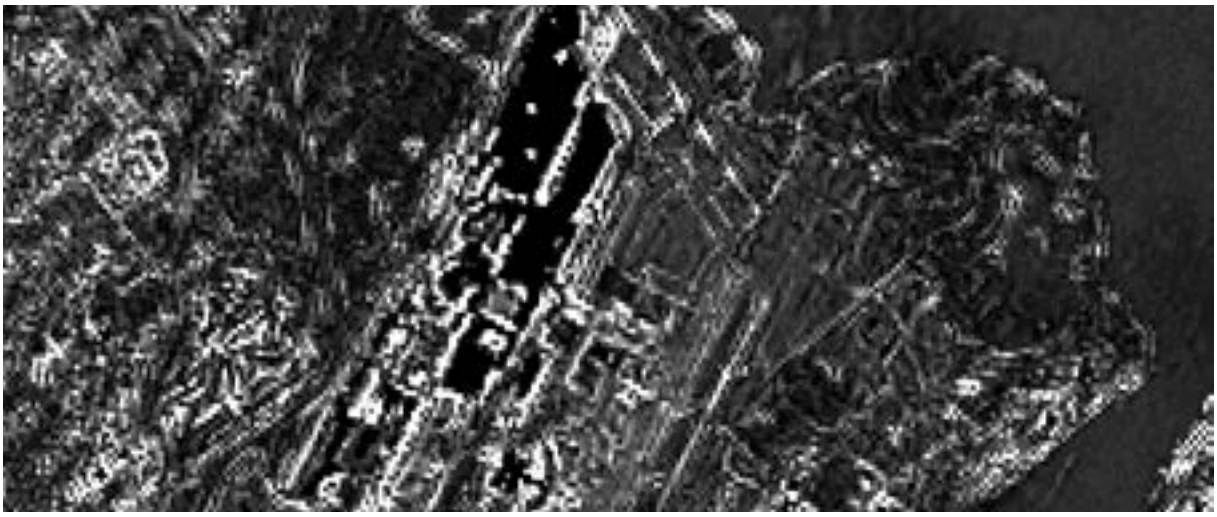
b. Nearest neighbour resampling (AEA projection; GDA94)

c. Nearest neighbouring resampling with picking and binning

d. Bilinear interpolation (AEA; GDA94)

e. Cubic convolution (AEA; GDA94)

f. Euclidean difference between bilinear and cubic convolution resampled images (black = minimum difference; white = maximum difference)



Source: Norman Mueller, Geoscience Australia

5.1 Along-line Resampling

This mode of resampling requires that the output grid be defined so that its lines are parallel to those of the current image. Along-line resampling does not allow the output image to be rotated relative to the input image, however it is a much less expensive method computationally than full image resampling. Since the along-line resampling approach does not rotate the image, hardcopy imagery produced this way may need to be manually rotated in order to overlay a map or another image.

Along-line resampling is performed by:

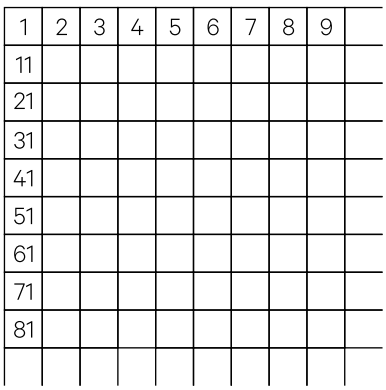
- omitting or duplicating lines;
- omitting or duplicating pixels along the line; and/or
- offsetting lines to account for image skew (see Figure 5.6).

In this way, each line of the output image is a contracted, expanded or shifted version of a line in the input image. However, by using whole image pixels as the basic adjustment factor for image rescaling and skewing, aliasing effects can occur in the resampled image; for example, straight lines may be stepped, small features may be distorted or superimposed boundary vectors could become discontinuous (or thickened) if an image line or pixel column containing two or more adjacent pixels is omitted (or duplicated).

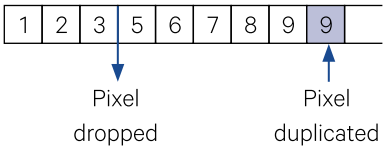
Many implementations of along-line resampling rely heavily on satellite or scanner models to remove along-line distortions (see Section 3.1.2). Additional parameters may also be required for scale adjustments and modelling specific image distortions, such as pixel size and image skew. These parameters may be derived from the results of an affine model that has been fitted to convert from image to map coordinates.

Figure 5.6 Along-line resampling

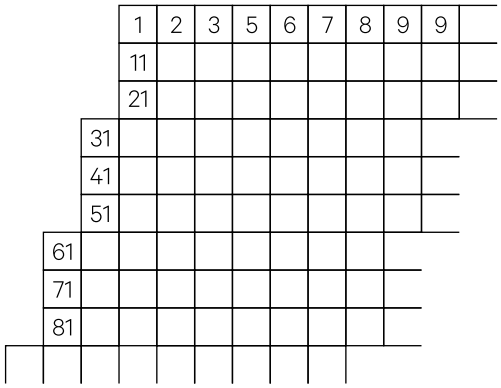
a. Original Image is labelled with sequential pixel and line numbers.



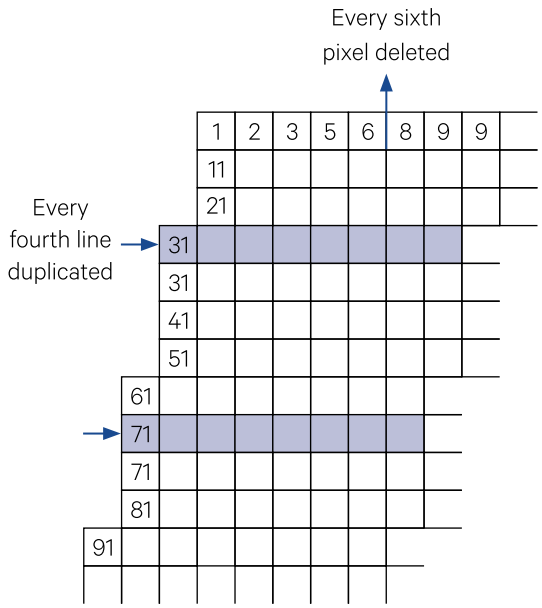
b. To rectify along-line distortions may require that pixel 4 is dropped out of the line and pixel 9 is duplicated.



c. To skew the image to the west by 1 pixel width every 3 lines, the first pixel in a line can be offset.



d. To rescale the rectified image to overlay with a map that has a smaller scale in the along-line direction and a larger scale in the along-track direction, the pixel size may be effectively narrowed and lengthened by respectively deleting pixels along a line and duplicating lines down the image. Here we delete every sixth pixel and duplicate every fourth line.



Source: Harrison and Jupp (1992) Figure 45

5.2 Nearest Neighbour Resampling

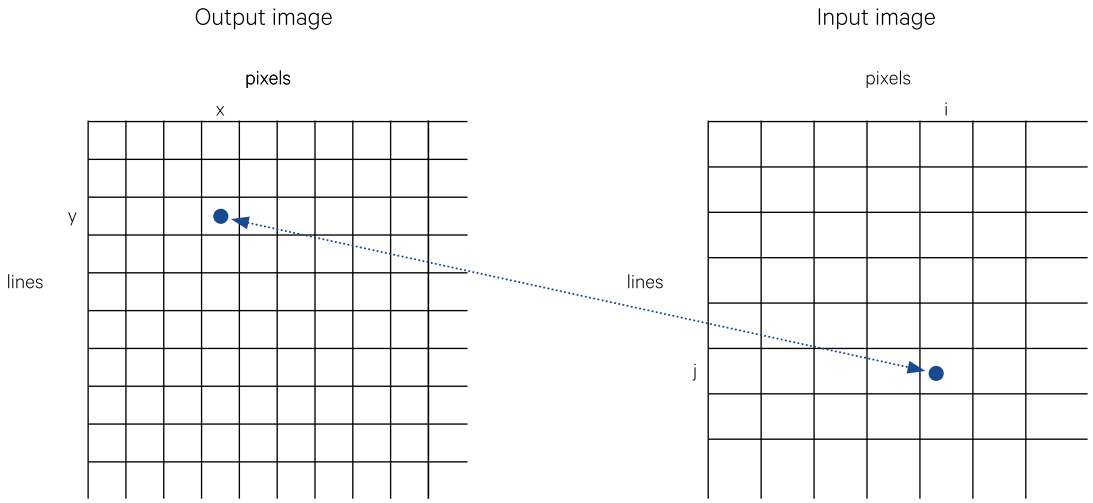
The nearest neighbour resampling method computes the location of the nearest input image pixel to each output image grid cell, then transfers the selected channel value(s) of the input pixel to the corresponding output pixel (see Figure 5.4a and Figure 5.7). This transfer process is also referred to as 'picking'. This method is much faster than other resampling methods and has the advantage of retaining the actual pixel values from the input images, which is especially important for attribute or non-numerical data. This sequence can be shown to be an 'optimal' approach to image resampling (see Excursus 5.2).

When pixels in the output image are larger than pixels in the input image, a picking process using only one sample per output pixel can cause aliasing (especially jagged lines) in the image since the spectral values of an output pixel may be 'mislocated' by up to half the input pixel size (see Figure 5.8). Accordingly, it is recommended that resampling be done to a smaller, intermediate pixel size than the input image, then the intermediate pixels are blocked or 'binned' to form larger pixels for a final output image (see Excursus 5.1).

It should be noted, however, that some implementations of nearest neighbour resampling have been observed to embed a grid artefact into the output image. This artefact is most obvious in resampled digital elevation data when it is displayed with relief shading (see Figure 5.9).

Figure 5.7 Nearest neighbour resampling

In this example, location x,y in the output image occurs in pixel i,j in the input image. The pixel value of this position in the input image is then used as the value of the cell at location x,y in the output image.



Source: Harrison and Jupp (1992) Figure 46

Figure 5.8 Nearest neighour interpolation

Resampling from smaller input image pixel size (solid black lines) to larger output image pixel size (dashed blue lines) can result in some input pixel values being mislocated or omitted in the output image.

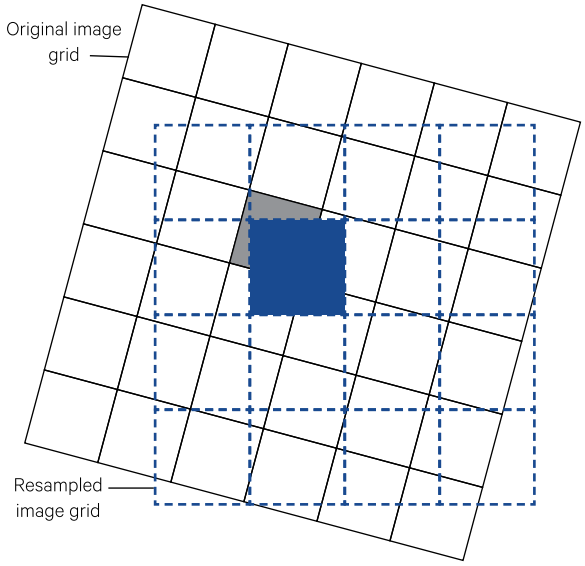
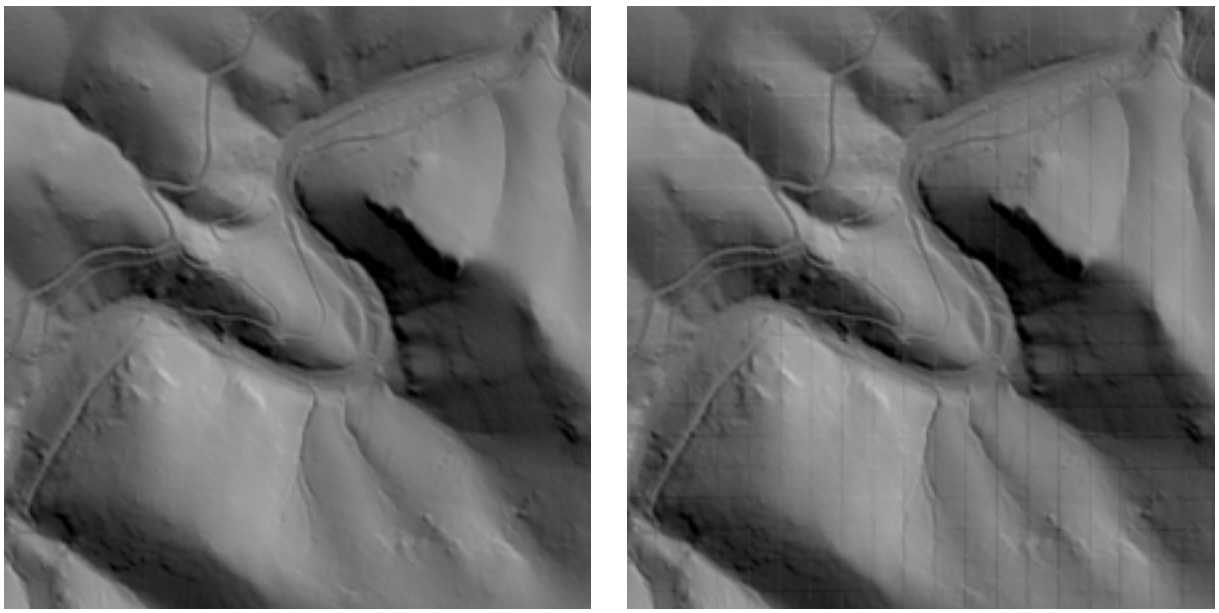


Figure 5.9 Potential artefacts from nearest neighbour resampling

a. Shaded relief image derived from DEM

b. Shaded relief image derived from same DEM after it was resampled using nearest neighbour resampling shows a grid artefact. This grid pattern is not visible in the input, resampled DEM



Source: Tony Sparks, Icon Water

Excursus 5.1—'Picking and Binning' Approach

To create smooth output images using the nearest neighbour resampling algorithm, it is recommended that the output image pixels be smaller than the input pixel size. A larger pixel size can be generated subsequently, if required, by 'blocking up' from a smaller pixel size (see Volume 2A—Section 7.2). For example, to resample Landsat TM imagery to an output pixel size of 50x50 metres, it is recommended that the image be first resampled to a pixel size of 25 m × 25 m then blocked using a 2×2 block size. In terms of resampling parameters, this merely requires that twice the number of output pixels and lines be specified for the base map. The basis of this approach can be explained using approximation theory (see Excursus 5.2).

This method is illustrated in Figure 5.10 using an EO image (A) that is to be resampled onto a map grid (C). The algorithm depends on defining an underlying grid or mesh (B), which is fine enough to represent the pixels of both the input and output grids. Data are 'picked' from A into the mesh B by nearest neighbour resampling then 'binned' into pixels of C to produce an approximation to the above algorithm.

Neat lines on a map are the boundaries at which it would be trimmed to join to adjacent maps. The resampled image can be considered as a base map whose boundaries form neat lines (that is the outer edges of the first and last pixels and lines in the image are like the limits of a map). In most image processing systems, the origin of the coordinates to define the neat lines can be anywhere inside or outside the base map. The number of grid cells across the grid, between the left and right neat lines, defines the width of a grid cell or pixel in the output image. Similarly the number of cells between the top and bottom neat lines determines the output pixel depth (see Figure 5.11), that is:

$$d_x = \frac{|xR - xL|}{NPIXELS}$$

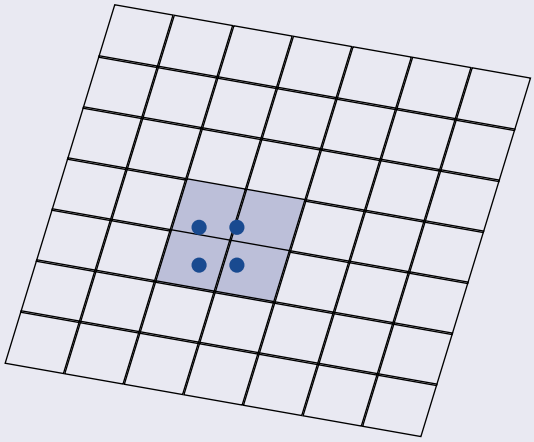
$$d_y = \frac{|yB - yT|}{NLINES}$$

in units of *xL*, *xR*, *yB* and *yT*, where

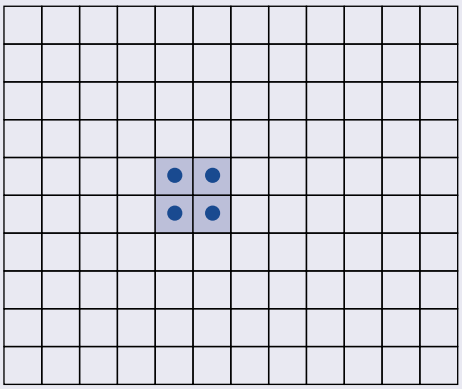
NPIXELS is the number of pixels in output line; and
NLINES is the number of lines in output grid.

Figure 5.10 Picking and binning approach

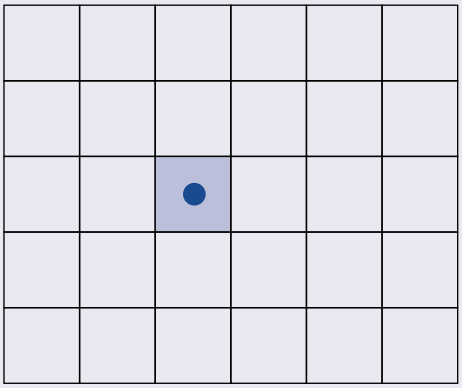
a. Original image grid showing pixels whose values will be used to create a single output pixel value



b. Placement of selected pixels on expanded grid corresponding to map scale



c. Output image with recomputed pixel value on final image grid

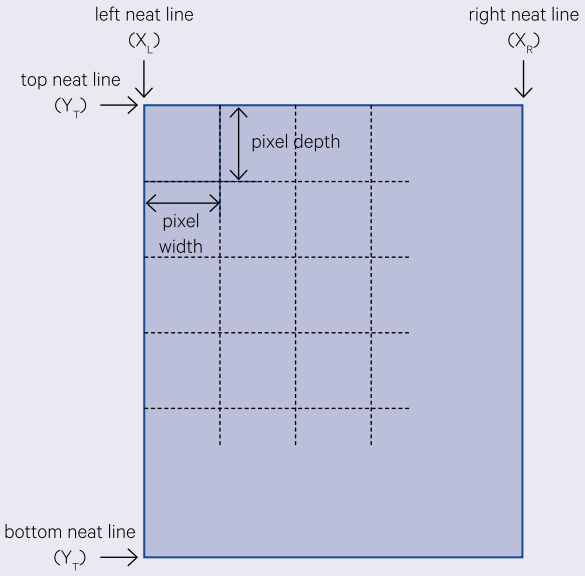


Source: Harrison and Jupp (1992) Figure 43

For example, to create an output UTM grid spanning 10 km from east to west and 15 km from north to south from a northwest origin at 341,000 m East and 8,065,000 m North with a pixel size of 100 m square, we would specify left (West) and right (East) neat lines as 341,000 and 352,000 (341,000 m + 10 km) and top (North) and bottom (South) neat lines as 8,065,000 m and 8,050,000 m (8,065,000 – 15 km; see Figure 5.12). Here, the pixel size is defined indirectly by neat lines and the number of pixels, so within this area there would be 100 pixels (10 km / 100 m) and 150 lines (15 km / 100 m) of 100 m square pixels. In this case, the output coordinates are based on a planar mapping grid in which distance units are consistently equal to 1 m of ground distance, so the output pixels represent areas of uniform size. Other coordinate systems, such as latitude/longitude can also be used to define an output grid, but in that case the output pixels will not represent uniformly sized areas.

Figure 5.11 Output image or base map

The pixel width is defined by the number of pixels between X_L and X_R, and the pixel depth is defined by the number of lines between Y_T and Y_B.

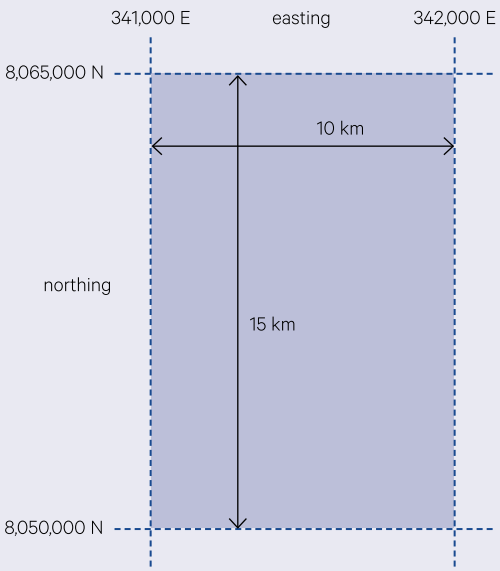


Source: Harrison and Jupp (1992) Figure 47

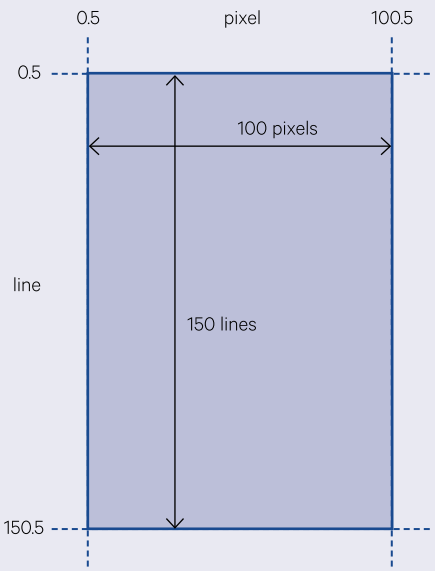
Figure 5.12 Example grid

This example creates an output UTM grid spanning 10 km from east to west and 15 km from north to south, relative to a northwest origin at 341,000 m East and 8,065,000 m North. The image pixel size is 100 m square.

a. Map coordinates



b. Coordinates of corresponding image grid



Source: Harrison and Jupp (1992) Figure 48

Excursus 5.2—Basis of the 'Picking and Binning' Approach

Source: David Jupp, CSIRO

The basis of the 'picking and binning' approach to resampling is explained by approximation theory. The general problem of resampling was illustrated in Figure 5.1 where a new region B was defined as overlapping three known areas A1, A2 and A3 in varying proportions. While each of the three original areas may represent some continuous underlying data function, in image format each would be represented as discrete pixel values. As illustrated in Figure 5.13, using subpixels followed by pixel aggregation more closely approximates regional values associated with vector boundaries (Tobler, 1979; Tapp, 2010). This subpixel approach is particularly advantageous when the resampling operation involves image rotation.

A newly defined data region such as B is commonly estimated as a linear combination of the original data:

$$\int\left[I_B\left(x\right)\right]f\left(x\right)dx=\sum_{j=1}^{3}a_j\int\left[I_{A_j}\left(x\right)\right]f\left(x\right)dx$$

where

$\int\left[Z\right]f\left(x\right)dx$ is the integral of function Z with respect to x;

$I_{A_j}\left(x\right)$ is the indicator function for A_j with value 1 when a location is in region A_j and 0 elsewhere;

$I_B\left(x\right)$ is the indicator function for region B;

a_j are coefficients to be determined; and

is the underlying data function.

Using this linear approximation, an 'error function' $E\left(j\right)$ can be written for the underlying data as:

$$E\left(j\right)=\int\left[I_B\left(x\right)\right]-\sum_{j=1}^{3}a_j\int\left[I_{A_j}\left(x\right)\right]f\left(x\right)dx$$

This statistic measures the extent to which the integral over one area can be estimated by knowing the integral over another area. Naturally if the underlying function is a constant, then there would be no error. This is assured if:

$$\sum_{j=1}^{3}a_j=1.0$$

A number of solutions could be derived for this type of general approximation theory problem, depending on how much is known about the underlying functions $f\left(x\right)$. The 'norm' of the error minimises the error functional such that:

$$\left\|E\right\|^2=\int\left[I_B\left(x\right)-\sum_{j=1}^{3}a_jI_{A_j}\left(x\right)\right]^2$$

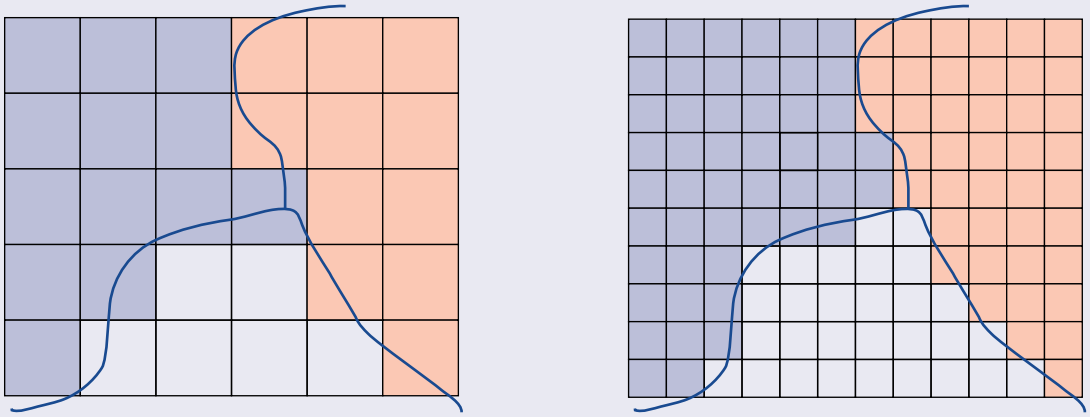
This gives minimum overall error when a_j is simply the area of A_j intersecting B (that is the proportion of B overlapping with A_j).

This result leads to the straightforward and computationally efficient method of resampling which can be implemented in most image processing systems (see Figure 5.10). The algorithm depends on defining an underlying mesh, which is fine enough to represent the pixels of both the input and output grids. Data are 'picked' from the input image into the intermediate mesh image using nearest neighbour selection then binned into pixels of the final output image to produce an approximation to the above algorithm. It is difficult to specify precisely how fine the intermediate mesh should be, but it needs to be less than one pixel of the input image and at least half the pixel width and/or depth of the output image.

Everything is related to everything else, but near things are more related than distant things.
(Waldo Tobler)

Figure 5.13 Sub-pixel behaviour

When vector data is represented in raster format, a finer pixel mesh renders better estimates of the original vector boundaries than a coarser mesh. Subpixels then allow the 'correct' pixel aggregate (or area average) solution.



5.3 Bilinear Interpolation

Resampling based on bilinear interpolation computes output image values from the closest four input pixel values within a 2×2 window surrounding each output pixel (see Figure 5.4b). This form of resampling generally involves weighting factors to emphasise the value(s) of the closest input pixel:

$$\text{output pixel} = \sum_{i=1}^{4} w_i \times \text{input pixel } i$$

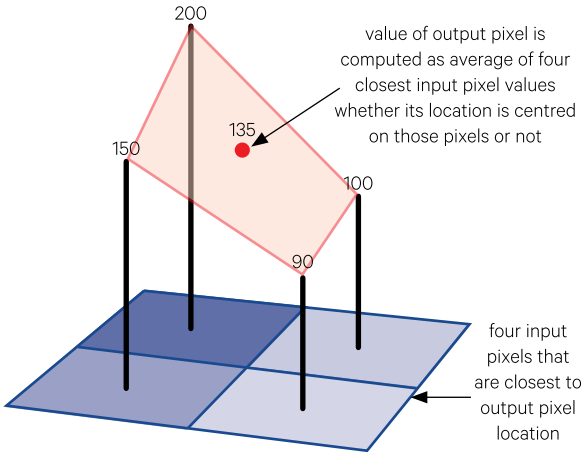
where

w_i are weighting factors derived from the relative distance between the input pixel and the output pixel, such that closer input pixels have higher weights (see Figure 5.14).

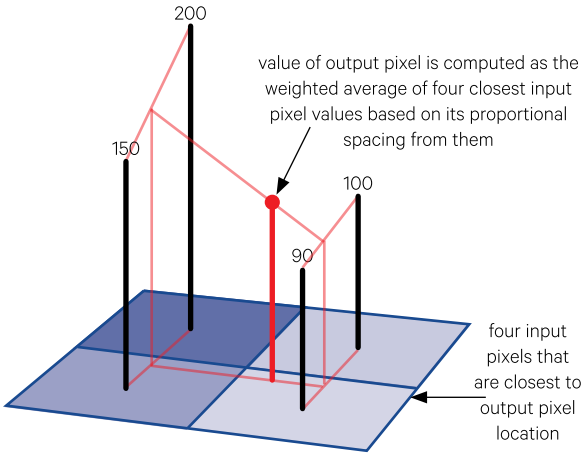
This process tends to create images with lowered contrast, in which edges are somewhat smoothed. It cannot be used with categorical data and is generally not recommended for EO imagery.

Figure 5.14 Bilinear interpolation

a. Example of unweighted interpolation, where the value of the output pixel is computed as the average value of the four closest input pixels: 135 = (200+150+90+100)/4.



b. To determine the interpolation weights to use for each output pixel, bilinear interpolation computes the proportional spacing from the four input pixels that are its closest neighbours.



5.4 Cubic Convolution

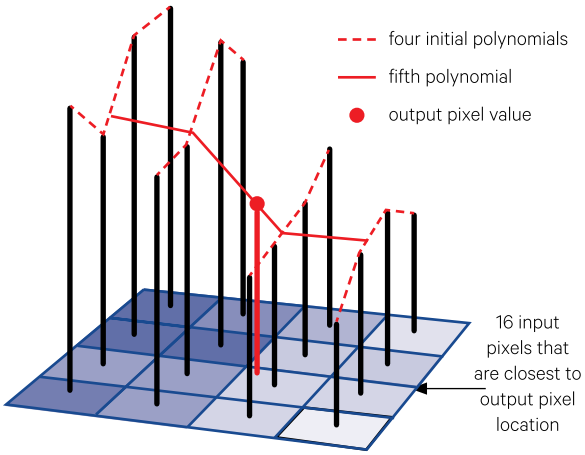
For each output pixel, cubic convolution resampling selects the closest 16 pixels within a 4×4 window in the input image and interpolates an output value by fitting polynomial models to the input values (see Figure 5.4c). For image data this is generally implemented by firstly fitting a cubic polynomial (see Figure 5.3) to each of the four 'horizontal' lines of four pixels surrounding the output pixel. The value of the output pixel is then determined by a fifth cubic polynomial fitted 'vertically' through these four functions and the output pixel (Richards, 2013; see Figure 5.15). Centring the fifth cubic function on the output pixel location ensures that the closest input pixels have exponentially more weight when calculating the output pixel value(s).

While the output image contains recomputed values, this method retains the input image mean and standard deviation more closely than bilinear interpolation. The cubic convolution interpolation function has been shown to be more accurate than nearest neighbour or bilinear resampling methods (Keys, 1981). This method is widely used for presentation of continuous EO image data, but cannot be used with categorical data. Most GA EO Products, for example, are resampled using Cubic Convolution and oriented to North Upwards. However, it is the most computationally expensive resampling method, being at least ten times slower to compute than nearest neighbour resampling.

To avoid aliasing and similar noise problems in resampled imagery, interpolation methods are generally preferred. When subsequent image processing requires that specific image values be preserved, however, such as when resampling categorical data, resampling methods that only use the original image pixel values are required.

Figure 5.15 Cubic convolution interpolation

The method is commonly implemented by fitting four individual cubic polynomials to the four lines of pixels surrounding the output pixel. A fifth polynomial is then fitted across these four these four functions and through the output pixel location.



5.5 Further Information

Resampling

Digital Light and Colour: *http://www.dl-c.com/Temp/downloads/Whitepapers/Resampling.pdf*

Digitales: *http://www.ldv.ei.tum.de/fileadmin/w00bfa/www/content_uploads/Vorlesung_3.4_Resampling.pdf*

Microimages: *http://www.microimages.com/documentation/TechGuides/77resampling.pdf*

5.6 References

Harrison, B.A., and Jupp, D.L.B. (1992). *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual*. MPA, Melbourne.

Keys, R.G. (1981). Cubic Convolution Interpolation for Digital Image Processing. *IEEE Trans. Acoustics, Speech and Signal Proc*. 29 (6), 1153–1160.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images*. NASA SP-431, NASA, Washington DC, USA.

Niblack, W. (1986). *An Introduction to Digital Image Processing*. Prentice-Hall, Denmark.

Richards, J.A. (2013). *Remote Sensing Digital Image Analysis: An Introduction*. 5th edn. Springer-Verlag, Berlin. ISBN 978-3-642-30061-5

Tapp, A.F. (2010). Areal Interpolation and Dasymetric Mapping Methods Using Local Ancillary Data Sources. *Cartography and Geographic Information Systems* 37(3), 215–228.

Tobler, W.R. (1979). Smooth Pycnophylactic Interpolation for Geographical Regions, *J. American Statistical Association* 74(367), 519–530. *https://doi.org/10.1080/01621459.1979.10481647*

6 Registration

Registration of EO imagery can take many forms, including:

- registering EO images with other spatial datasets, such as maps or other images (see Excursus 6.1 and Section 6.1);
- mosaicking two or more EO images to cover a common base grid (see Section 6.2); and
- locating sample sites in EO imagery, such as ground locations of field sites (see Section 6.3).

Excursus 6.1—The Australian Geographic Reference Image

Source: Wang Lan-Wei, Geoscience Australia
Further Information: Lewis *et al.* (2011)
Product Download: *https://data.gov.au/dataset/agri-the-australian-geographic-reference-image/resource/e66aa61a-5345-4141-be9f-3e7bf6c0a4a4*

The most reliable approach to orthorectification is to register all images to a single controlled image base—a reference image. Geoscience Australia (GA) has used this approach since 2002, rectifying images from the Landsat satellites to the national Landsat panchromatic mosaic. However, the Landsat panchromatic mosaic has an accuracy of no more than 15 m, and cannot meet the need to rectify higher resolution imagery available from more recent and future Earth Observation satellites. The Australian Geographic Reference Image (AGRI) is a national mosaic which provides a spatially correct reference image at a 2.5 m resolution across Australia. GA developed AGRI to address the need for a higher resolution reference image, of known accuracy, over the entire Australian continent. (Lewis et al., 2011)

The *Australian Geographic Reference Image* (AGRI) is a consistent and accurate reference image, with 2.5 m spatial resolution, for rectification of imagery from multiple sources at spatial resolutions of 2.5 m or less. Accurate and consistent rectification is essential to ensure that observations taken at different times, from different sources and in the field, can be compared. Image registration error should ideally be less than the image pixel size, because this allows each pixel to be compared 'with itself' through time. Whilst it is relatively simple to ensure high relative accuracy within a given type of imagery (by co-registration of all images to an initial base image of arbitrary accuracy) this is of limited benefit because it does not allow comparison with different data sources, including field observations. Accurate rectification is therefore essential for monitoring as well as for many other scientific and practical uses of EO data. In particular, the advantages of higher resolution imagery may be lost if accurate rectification is not possible.

Background image: Landsat-5 image of Lake Eyre, South Australia, acquired as it was filling on 9 May, 2009. This image is displayed using bands 5, 4, 2 as RGB and overlaid with registered cartographic detail showing contour lines (black) and roads (red). **Source:** Norman Mueller and Erin Teifer, Geoscience Australia

A simple and reliable approach to image rectification is to register all images to a single controlled image base, which is referred to as a reference image. GA has used this approach since 2002 to rectify Landsat images to the national Landsat panchromatic mosaic. However, the approach does require that the reference image be constructed in the first place, which is a non-trivial task. It also assumes that the image to be registered has a resolution less than or equal to the resolution of the reference image. The Landsat panchromatic mosaic has an accuracy (and resolution) of no more than 15 m and is therefore not suitable as a reference image to rectify higher resolution imagery from more recent—and forthcoming—EO satellites.

GA developed AGRI between July 2009 and June 2011. The project was made possible by a combination of:

- new data from Japan's Advanced Land Observing Satellite (ALOS) which produced panchromatic observations at 2.5 m resolution;
- new 'full pass' processing techniques for rectification of satellite imagery developed by the Cooperative Research Centre for Spatial Information (CRCSI) and included in *Barista* software (Weser *et al.*, 2008; Fraser *et al.*, 2009; Rottensteiner *et al.*, 2009);
- expertise in Geodesy and the Global Positioning System (GPS) at GA; and
- the capabilities of the Australian Spatial industry in GIS database design, field survey and image processing.

ALOS PRISM

The Advanced Land Observation Satellite (ALOS) was launched in January 2006 and operated at an altitude of 692 km in a sun-synchronous, sub-recurrent orbit from 2006 until April 2011. ALOS carried several instruments including the Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM). ALOS/PRISM imagery was acquired in 35 km or 70 km width swaths, referred to as 'OB1' or 'OB2' imaging modes respectively. OB1 images are nominally 35 km × 35 km, whereas OB2 images are nominally 35 km along-track and 70 km across-track. The ALOS/PRISM instrument included fixed forward and backward looking cameras to allow Digital Elevation Models (DEM) to be constructed from the triplet images (forward, nadir, backward), however only nadir images were used to develop AGRI. Technical details of the PRISM instrument are given in Table 6.1.

Table 6.1 ALOS/PRISM characteristics

Swath width (OB1)	35 km (nadir) scene footprint 35 km x 35 km
Swath width OB2)	70 km (nadir) scene footprint 70 km x 35 km
Spectral extent	0.52 μm to 0.77 μm
Spatial resolution	2.5 m (nadir)
Radiometric resolution	8 bit
Scanning method	Push broom

Barista Processing

Traditional scene-by-scene rectification methods are impractical for a project of this magnitude. Some 6,000 ALOS/PRISM scenes are required to cover Australia. Rectification of ALOS/PRISM imagery is therefore a serious problem because of the large numbers of images involved and because over much of the land surface, near-shore islands and reefs, there are few if any accurately mapped features to which an image can be registered with 2.5 m accuracy. Even allowing for maximal re-use of points by choosing Ground Control Points (GCP) within the overlap of adjacent swaths, and use of OB2 format imagery where available, up to 30,000 high-accuracy control points would be required to enable scene-based orthorectification across the entire continent. (For orthorectification purposes OB2 imaging mode is advantageous because the image size is effectively doubled, halving the number of ground control points required per unit area.) A practical approach to rectification of ALOS/PRISM imagery was found through the *Barista* software developed by the CRCSI.

Barista allows 'full pass' or 'strip' processing, that is, *Barista* rectifies a sequential strip of images taken during a single segment of a satellite's orbit (Fraser *et al.*, 2009). Data on the precise orientation of the satellite, a precise model of the sensor, a generic sensor orientation model, and GCPs at the end of each strip of imagery, are used in the correction. Importantly, control points are required only at the end of each strip, not for each image in the strip; this reduces the requirement for control points by a least an order of magnitude.

Using *Barista*, the metadata for each separate scene are merged to produce a single, continuous set of orbit and attitude parameters, such that the entire strip of tens of images can be treated as a single image, even though the separate scenes are not actually merged. The merging of orbit data results in a considerable reduction in both the number of unknown orientation parameters and the number of control points required in the sensor orientation adjustment (Barr *et al.*, 2010). In contrast to traditional rectification, which is completed scene by scene,

full pass processing can be expected to ensure consistency between sequential images, to require increased knowledge of the satellite orbit, and most importantly to require fewer control points.

The orthocorrection of images requires a number of optimally placed GCPs of high accuracy to refine the sensor orientation parameters. The metadata for each separate ALOS scene also are merged within *Barista* to produce a single, continuous set of orbit and attitude parameters, which results in a considerable reduction in both the number of unknown parameters and the number of control points required in the sensor orientation adjustment.

The reduced requirement for control points with full-pass processing makes rectification over large areas feasible. Given that a satellite pass over Australia may capture 50 to 100 images, the *Barista* methodology reduces the field work necessary to achieve a national coverage of rectified images by as much as 50 times. In fact, only around 500 control points were required to rectify ALOS/PRISM imagery over all of Australia using *Barista*, whereas 30,000 may have been necessary with conventional approaches. Initial studies applied to very long strips of ALOS/PRISM imagery indicated that pixel-level accuracy can be achieved over strip lengths of more than 50 images, or 1500 km, with as few as four GCPs.

AGRI development

Through a series of trials, the following were established:

- the number of GCPs required to rectify an image pass;
- the accuracy that could be expected in the corrected image;
- the design of field methods for GCP collection; and
- the project feasibility.

Once project feasiblity was confirmed, the project was undertaken as sequence of tasks:

- GIS database design for GCPs and associated images and data;
- design of field work campaign to capture control points for each ALOS data pass, concentrating on near-coastal areas for control points and central Australian locations for check points;
- development of control point field work specifications;
- industry contracts for completion of the control point field work, including quality assurance and ingest of data points to the GIS database;
- selection of the most fit for purpose ALOS/PRISM images from the GA data archive;
- preliminary processing of ALOS/PRISM images, from 'raw' data to 'radiometrically corrected' images;
- orthorectification, including accuracy assessment and quality assurance; and
- generation of image mosaics.

Ground control points

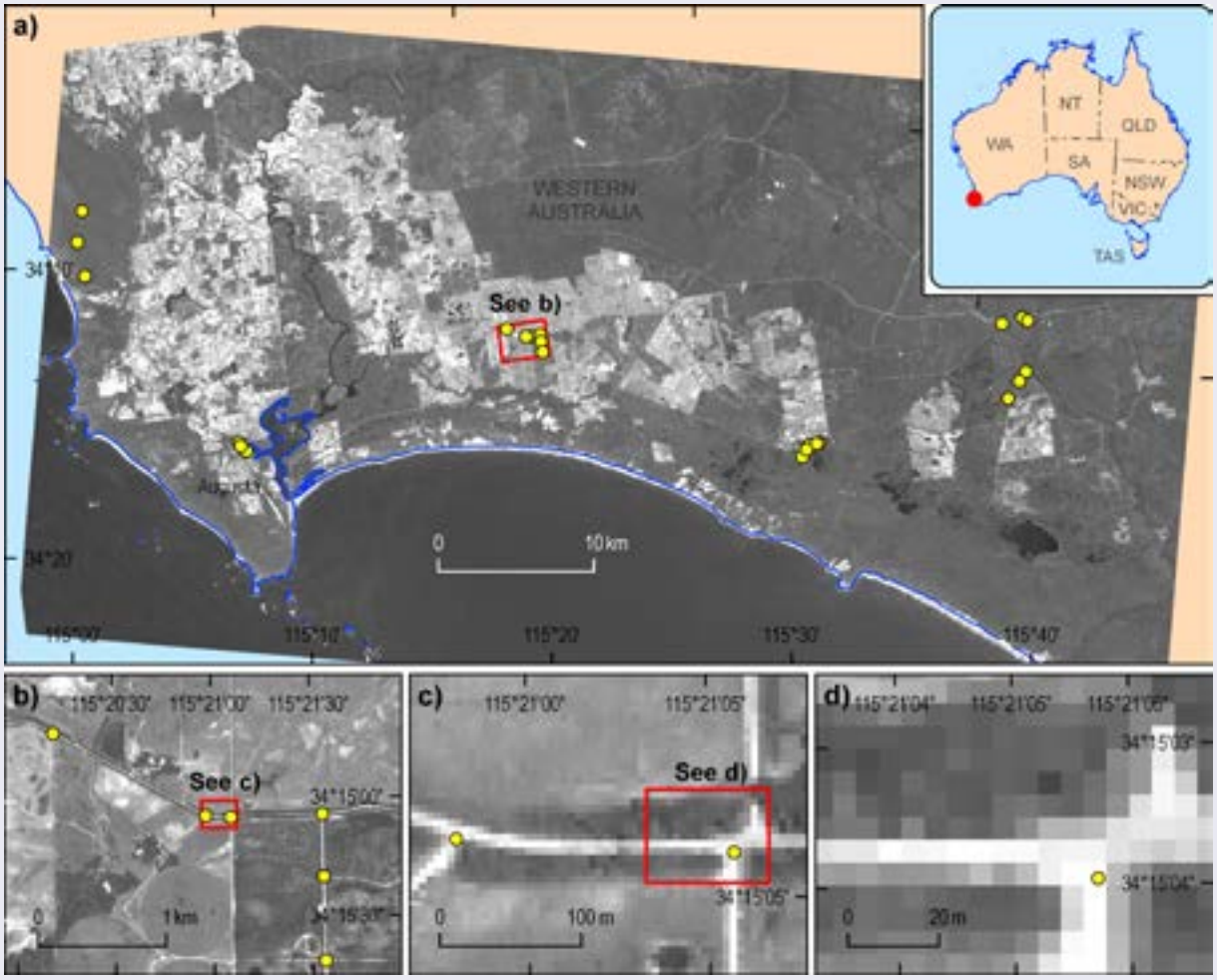
The control points were organised into 23 work packages, that included templates for data collection, corresponding imagery over the package, supporting GIS data, a set of specifications detailing the accuracy required, the type of data to be collected and the structure of the data to be returned (see Figure 6.1). Each survey package included a number of control point sites, where a minimum of two, and preferably three, features were to be surveyed (see. Figure 6.2). Each survey package also required the survey of a small number of existing Permanent Survey Marks (PSMs) using the same field techniques applied to the collection of control points. This served as a secondary assurance of accuracy. Criteria for selection of control points is detailed in Excursus 4.2.

Figure 6.1 GCP work packages

Source: Lewis *et al.* (2011) Figure 6

Figure 6.2 Example survey site

a. Examples of field survey sites in southwest WA. b. Control point features at one site. c/d. Precise location(s) of control point(s) shown on image.



Source: Lewis *et al.* (2011) Figure 7

To create the AGRI, a total of 2,885 control points at 737 sites in 23 packages were surveyed (see Figure 6.3). Surveyors were required to use GPS technology to obtain control point locations with accuracy of at least 0.25 m in the horizontal plane, and 0.5 m in elevation. Surveyors were also required to provide sketches in digital format of the surveyed features, digital photographs of the feature, and to indicate the location of the control point on a satellite image provided for use in the field. The submitted GIS points contained the precise coordinates of the surveyed point in three dimensions referenced to GDA94 geographical horizontal datum and GDA94 ellipsoid height vertical datum. In some parts of remote northern Australia, field access by vehicle was impractical and suitable ground control features are rare or absent. In the Kimberley region of WA, high accuracy aerial photography was therefore used to establish control. Aerial photography was commissioned in 5 km × 10 km blocks over fifteen sites identified as having suitable features. In the Kimberley, when features were not necessarily man-made, geological features were also considered.

Figure 6.3 Surveyed GCP for AGRI

Source: Lewis *et al.* (2011) Figure 11

Full-pass processing and accuracy assessment

Full-pass processing involved:

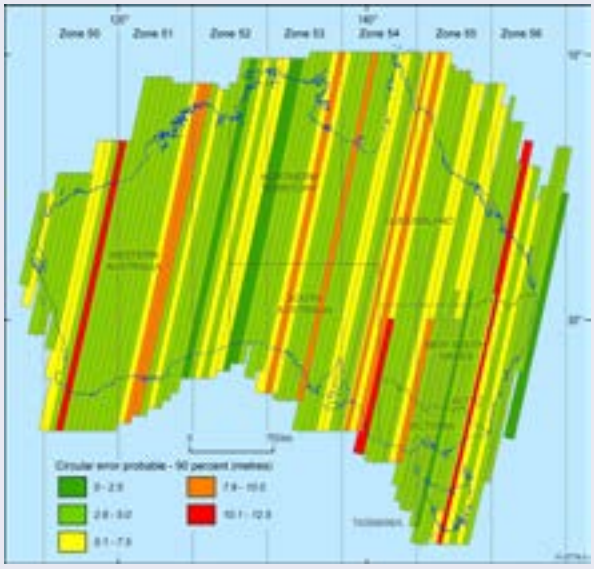
- analysis to filter out control point outliers (indicative of gross errors); then
- processing of the pass using a limited number of control points. The remaining GCPs served as check points to estimate the error in the correction.

Based on the field accuracy specifications and methods, the standard deviations of latitude, longitude and height were estimated as 0.25 m, 0.25 m and 0.5 m respectively. These data allow for the suitable weighting of the control point coordinates compared with the other observations and measurements used in the ALOS/PRISM full-pass model (Weser *et al.*, 2008).

Accuracy was assessed for each orbit segment against available check point residuals (see Figure 6.4a), by visually comparing rectified imagery with dGPS tracking data recorded in a moving vehicle (see Figure 6.4b), and by visual comparison with adjacent image paths (see Figure 6.4c).

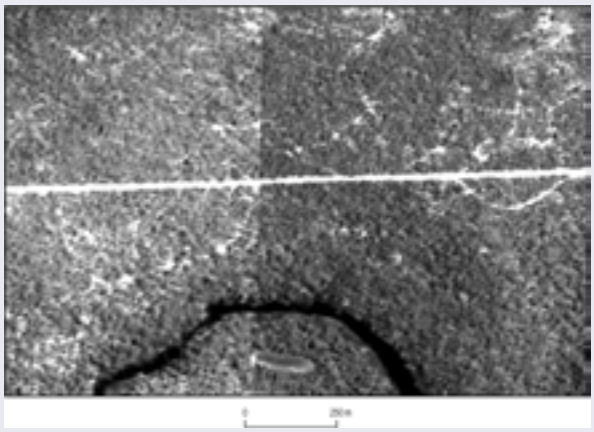
Figure 6.4 Accuracy assessment

a. Check point residuals



b. dGPS tracking data plotted on orthorectified ALOS/PRISM image

c. Visual assessment of adjoining features image pairs from adjacent paths



Source: a. Lewis *et al.* (2011) Figure 16. b. Lewis *et al.* (2011) Figure 17. c. Lewis *et al.* (2011) Figure 18

AGRI products

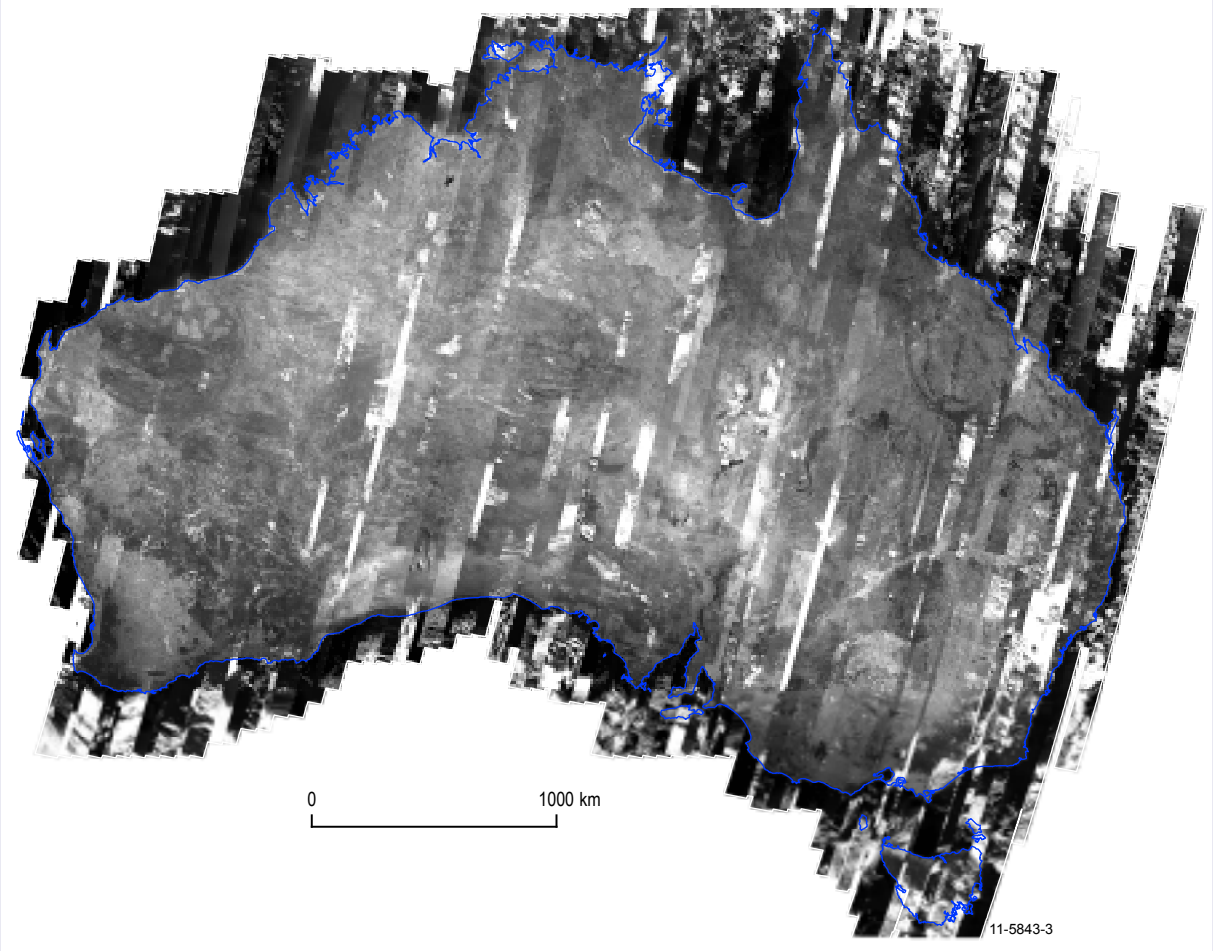
9,560 corrected ALOS/PRISM scenes, totally over 6 tb, were combined into a manageable set of image mosaics:

- Eight UTM zone-based contrast balanced mosaics at 2.5 m resolution covering Australia (Continental Australia, the Great Barrier Reef, other offshore reefs and islands); and
- A reduced resolution (0.0001º—approximately 10 m) national scale mosaic in geographical coordinates (Equirectangular projection; see Figure 1.17).

For visual consistency across each mosaic, intensity and contrast balancing was iteratively applied using linear contrast tables (see Volume 2D). These mosaics are freely downloadable from GA. The AGRI data package is also freely available from GA and includes:

- a control point database, which details precise survey data relating to each identified ground feature accepted from the field surveys, plus ancillary information; and
- a path metadata geodatabase, which identifies the approximate positions of each ALOS/PRISM scene used to create the mosaics, plus associated metadata.

Figure 6.5 AGRI national mosaic



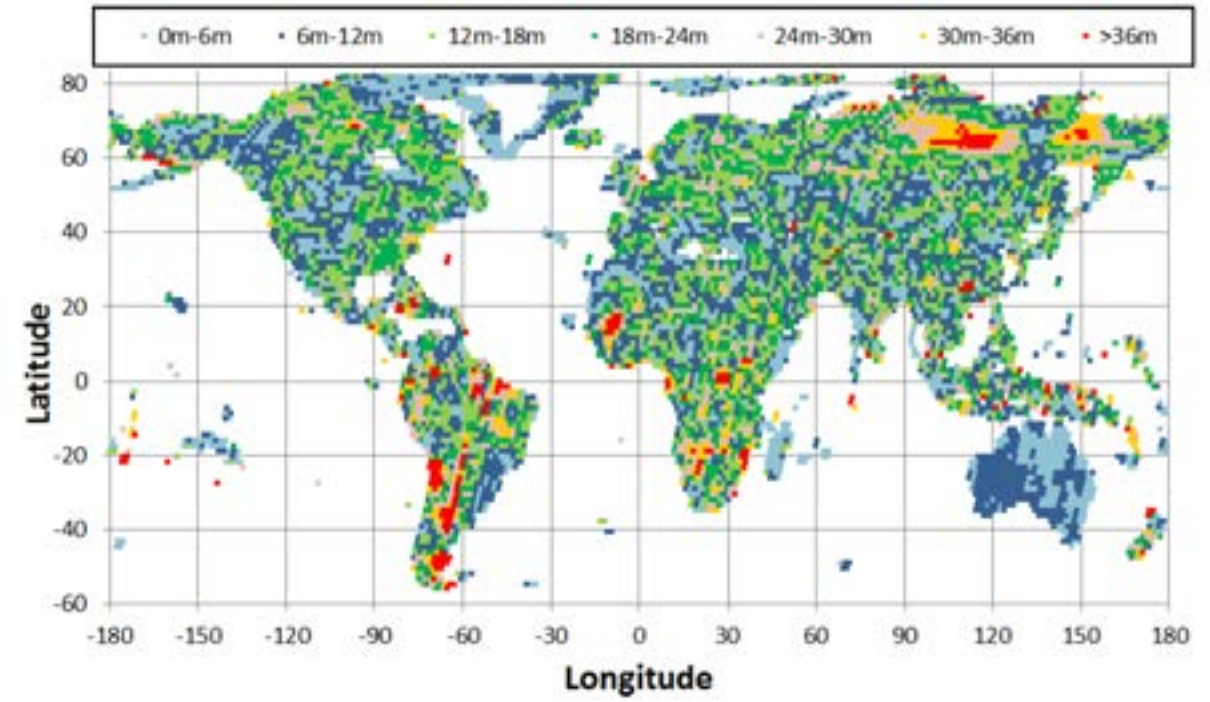
Source: Geoscience Australia

Global collaboration

Following the release of AGRI, GA collaborated with USGS and ESA to use this resource as the control source over Australia for rectifying imagery acquired by Landsat and Sentinel-2 respectively. USGS used AGRI to refine their block adjustment and derive new Landsat GCPs over Australia. When compared with similar data globally, the consistency and accuracy of horizontal error is impressive across the whole continent (see Figure 6.6). The Global Reference Image (GRI) being finalised by ESA will be used as a control source for future Sentinel-2 orthorectified products. It is expected that the resulting rectified products from Landsat and Sentinel-2 will have good geometric agreement with each other and therefore be particularly valuable for multitemporal analyses.

Figure 6.6 Landsat-8 estimate of current Global Land Survey (GLS) horizontal error

This estimate is based on data that was valid in September 2017.



Source: James Storey, USGS

6.1 Registering EO Images with Other Spatial Datasets

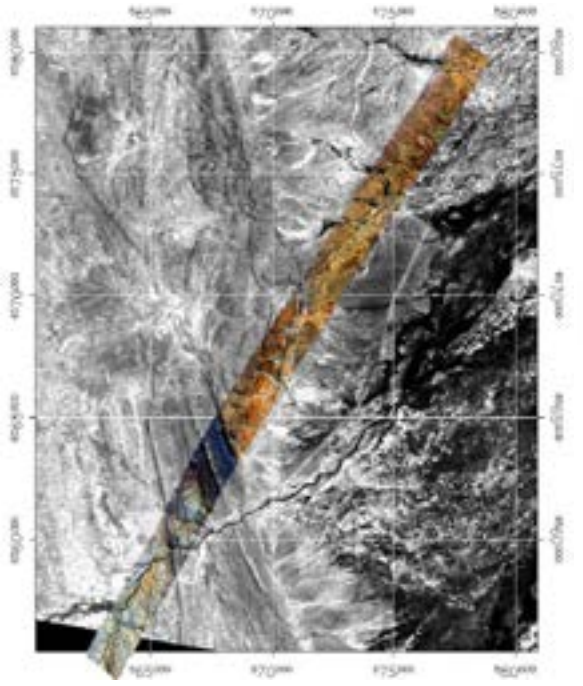
A significant volume of cartographical information is now available in digital format. Most commonly digital map data is stored as part of a Geographic Information System (GIS). EO and GIS datasets are becoming increasingly integrated for the purposes of analysis, interpretation and visual presentation (see Volume 2D). Similarly, the results of EO analyses can be recorded as additional layers in a GIS. Some uses of GIS data in EO analyses include:

- annotation imagery using map information;
- segmenting EO image using boundary of specific study area or for statistical comparisons (see Volume 2A—Section 10);
- resolving spectral ambiguities during image classification (see Volume 2A—Section 9 and Volume 2E);
- assessing the accuracy of image analyses relative to known map data; and
- modelling environmental processes, such as erosion and plant growth, using relevant spatial data variables, such as geology and elevation (see Volumes 2C and 2D).

When co-registering two sources of EO data, they both need to be resampled relative to the base grid, which is most often a standard map coordinate system (see Figure 6.7).

Figure 6.7 Georegistering two sources of EO data

CASI airborne scanner image strip georegistered to Landsat TM satellite image for Fowlers Gap Research Station, NSW, using UTM map coordinate system.



Source: Megan Lewis , University of Adelaide

Consideration needs to be given to scale when different sources of spatial data are being integrated. To determine the map scale corresponding to the spatial resolution of a raster grid, Tobler (1987, 1988) proposed:

$$\text{map scale} = \text{pixel size} \times 2 \times 1000$$

where pixel size is the image spatial resolution (in m). This rule is based on the assumption that one thousandth of the denominator of the map scale represents the size of 'detectable' features on the map, measured in metres (see Table 6.2).

However, to sensibly overlay an EO image onto a map, or map data onto an image, both the map and image data need to be viewed using the same projection and datum (see Figure 6.8). This can be done either by resampling the image data to the same projection as the map or, for visual presentation only, converting the map and/or image data to a matching projection during the image display process ('on the fly'). The latter option assumes that the map and image data are represented in projections that can be 'interconverted' during image display.

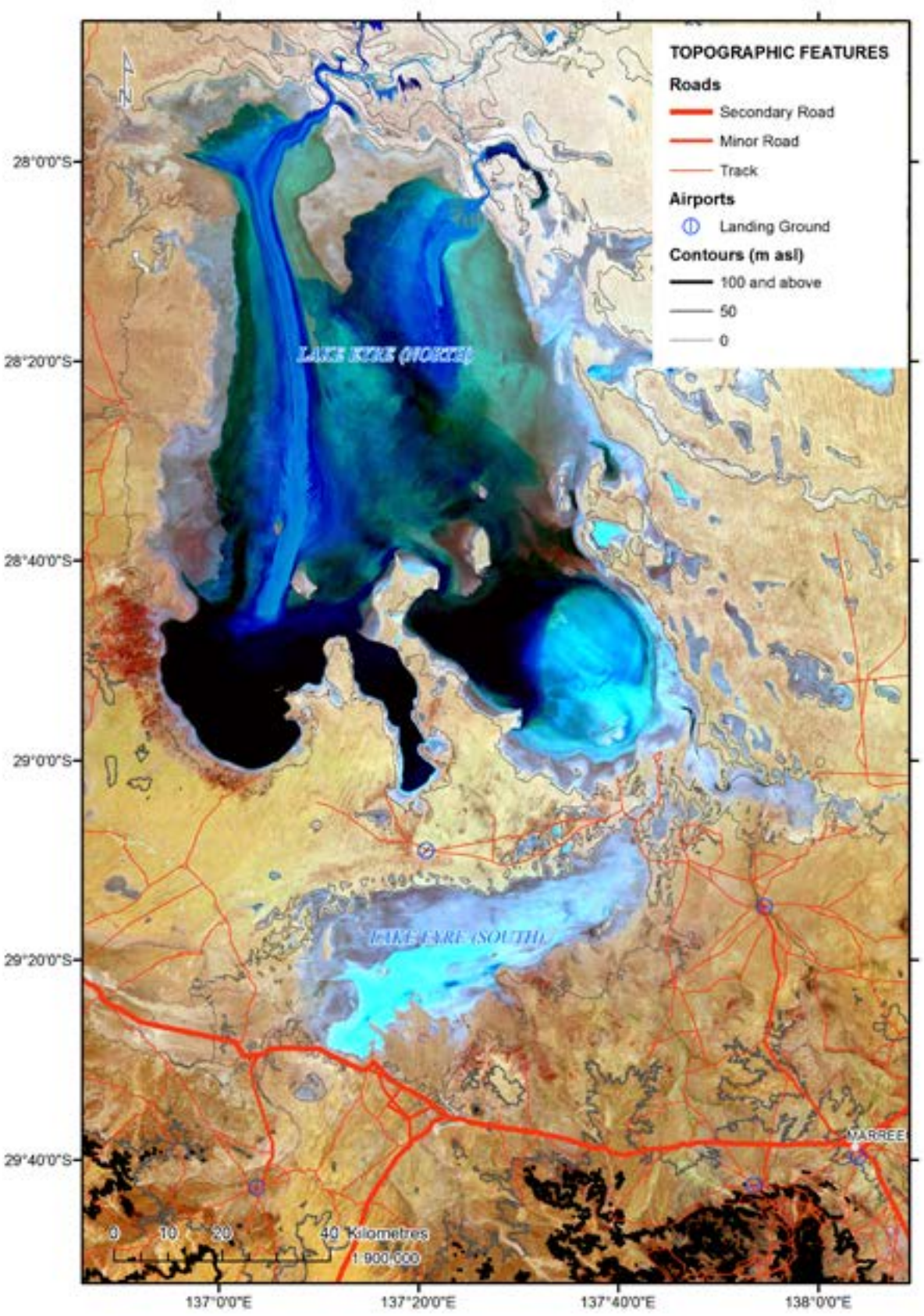
Table 6.2 Map scale and image resolution

Map Scale	Detectable size (m)	Image resolution (m)
1:1,000	1	0.5
1:5,000	5	2.5
1:10,000	10	5
1:50,000	50	25
1:100,000	100	50
1:250,000	250	125
1:500,000	500	250
1:1,000,000	1,000	500

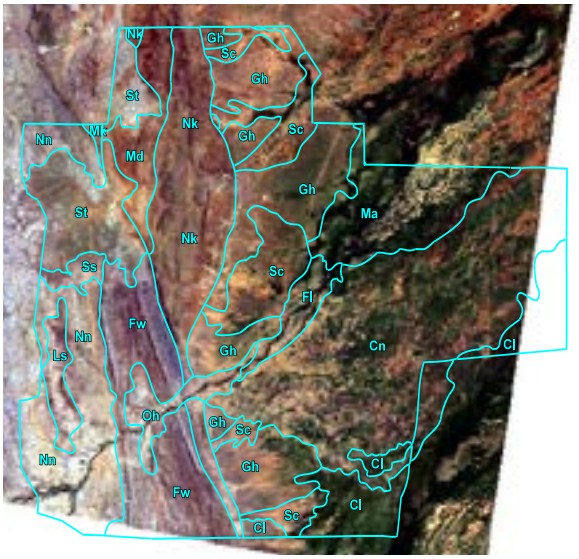
Adapted from: Nagi (2010)

Figure 6.8 Overlaying map information on EO imagery

a. Landsat-5 image of Lake Eyre, South Australia, filling on 9 May 2009, displayed using bands 5, 4, 2 as RGB and overlaid with relevant map information identifying location, scale and topographic features.



b. Land system boundaries were digitised from a pre-existing map and overlain on georegistered Landsat TM image to illustrate landscape features of Fowlers Gap Research Station (acquired on 30 January 1993 and displayed using bands 3, 2, 1 as RGB).



Source: a. Norman Mueller and Erin Teifer, Geoscience Australia. b. Megan Lewis, University of Adelaide

The importance of using the same projection and datum for all data sources when registering spatial data is illustrated in Figure 6.9. This example shows aerial imagery acquired in 2016, which has been orthorectified to the ACT Standard Grid coordinate system. This grid is derived from a Transverse Mercator projection based on the Australian Geodetic Datum 1966 (AGD66)[7]. In Figure 6.9a, both image and map data are transformed relative to the same datum (AGD66), so the map data overlays the EO image accurately. Figure 6.9b, however, shows the result of overlaying a map layer that is stored in Map Grid of Australia (MGA55) coordinates based on GDA94 over an image based on AGD66. As can be seen, due to differences between their projections and datums, when no transformation is applied the MGA coordinates of the map data are shifted approximately 200 m northeast of the ACT Standard Grid coordinates of the image data.

Figure 6.9 Datum differences

ADS40 aerial imagery acquired in 2016 and orthorectified to the ACT Standard Grid, a Transverse Mercator projection based on AGD66.

a. Imagery over central Canberra is overlain with a GIS layer that shows the actual extent of lakes. In this case both image and GIS layer are transformed relative to the same datum.

b. Overlay of Lakes GIS layer in MGA 55 coordinates relative to the Geodetic Datum of Australia 1994 (GDA94). The different datum results in an overlay shift to the northeast of approximately 200 m relative to image features.

Source: Tony Sparks, Icon Water, based on ACT Government orthorectified imagery supplied by AAM (Ref. No.: 26121A) CC BY 4.0, and GIS data courtesy of ACT Office of the Surveyor-General and Land Information.

[7] For further information on the ACT Grid, please visit: *https://www.planning.act.gov.au/tools_resources/survey-data-maps/surveying-data/surveyors_information/coordinate_system*

Excursus 6.2—Overlaying Map and Image Data

Source: Tony Sparks, Icon Water

When overlaying two spatial data sets, such as map and image data, it is essential that both data sources be appropriately managed to ensure that they can be projected onto the same projection and datum. The importance of this is illustrated in Figure 6.10. In all three example images, aerial imagery is projected using datum GDA94 and the map coordinate system of MGA zone 55, while the map data (roads and cadastre) are projected using the ACT Standard Grid coordinate system, based on the AGD66 datum.

The mismatch that results from directly overlaying these two different coordinate systems is shown in Figure 6.10a. However, when the map data are transformed ('on the fly') from AGD66 to GDA94 datum (using the AGD66 to GDA1994 11NTv2 transformation), these data sets overlay accurately (see Figure 6.10b). Similarly, when both are projected onto a geographic grid using the same datum, they overlay correctly even though the underlying imagery is also being reprojected 'on the fly' (see Figure 6.10c).

Figure 6.10 Projection differences

This example shows aerial imagery, acquired in March 2018 and centred on Collins Park, Forrest, ACT, overlaid with map information showing roads and cadastra in yellow. The imagery and map information remain in their original projections in all examples with reprojections and transformations being undertaken during the display process.

a. Imagery based on MGA55 coordinates (GDA94 datum), overlaid with map information based on the ACT Standard Grid (AGD66 datum) with no datum transformation applied.

b. Imagery based on MGA55 coordinates (GDA94 datum), overlaid with map information transformed to the same projection and datum.

c. Imagery transformed to geographic coordinates (GDA94 datum), overlaid with map information transformed to the same projection and datum.

Source: Tony Sparks, Icon Water, based on ACT Government orthorectified imagery acquired in 2018. The imagery was supplied by AAM (Ref. No.: 26121A) CC BY 4.0 and GIS data courtesy of ACT Office of the Surveyor-General and Land Information.

6.2 Mosaicking EO Images

In many cases, multiple EO images need to be joined to cover the full extent of a particular study area. Given the variations that can occur between EO imagery, even when images are acquired by the same scanner on the same day, many processing operations are required to produce a seamless mosaic. One example of this is given in Excursus 2.1.

Before a set of images can be sensibly mosaicked into a single image all input images must be corrected to a consistent standard in terms of geometric and radiometric properties. As detailed above, the processing steps involved in geometric correction generally involve three stages:

- rectification—the process of modelling geometric characteristics of EO imagery so that the image geometry accurately represents the geometric features of the Earth's surface;
- registration—the process of geometrically matching different spatial datasets, such as imagery and/or maps, so that positions in one dataset may be accurately located in others; and
- resampling—modifying the geometry of an image (which may be from either an EO image or a map data source, such as a GIS layer).

The objective of intensity and contrast balancing is to obtain visual consistency across the mosaic while maintaining the dynamic range of the image, especially over land areas, without any loss of geometric integrity of the image.
(Lewis et al., 2011)

As introduced in Volume 2A—Section 3.3, the logical components of radiometric correction include:

- correcting for atmospheric effects using calibration data (see Volume 2A—Section 3);
- minimising surface Bi-directional Reflectance Distribution Function (BRDF) to reduce differences due to changes in viewing and illumination positions within and between images (see Appendices 1 and 2).
- correcting for variations in terrain illumination due to topographic shading (see Appendix 1).
- using on-board and/or vicarious reference data to correct for detector differences. Correction of instrument errors may involve destriping algorithms to remove systematic patterns (see Volume 2C), or specialised algorithms for random distortions (see Volume 2A—Section 8.3).

In many cases EO data products are available as standard, calibrated images, which have effectively been corrected for geometric and radiometric differences. In order to mosaic multiple scenes, however, colour balancing is still required to minimise variations in ground features between overpass dates and to ensure that all images use the same range of values for display purposes (see Excursus 6.3). Methods for optimising colour balance and reducing the visual impact of boundaries in mosaicked imagery are detailed in Volumes 2C and 2D. An example of continental scale image mosaicking is given in Volume 2A—Excursus 8.3.

Excursus 6.3—Multi-scene Image Mosaic

Source: Norman Mueller, Geoscience Australia

An example of the colour variations that can exist in individual, adjacent Landsat-8 OLI image scenes is shown in Figure 6.11. These images over the Gulf of Carpentaria were acquired on two dates, with path 100 (Figure 6.11a and c) being acquired on 6 July 2016 and path 99 (Figure 6.11b and d) being acquired over three weeks later on 31 July 2016.

Using a common geometric and radiometric standard, the individual image scenes (Level 1 data) can be resampled to form the mosaic shown in Figure 6.12. This mosaic shows the full extent of the original image scenes with consistent colour balance along paths, but visible differences between image paths.

Figure 6.11 Uncorrected image scenes

These Landsat-8 OLI images (Level 1 data) show four adjacent image scenes in the Gulf of Carpentaria in northern Australia, displayed using bands 6, 5, 3 as RGB. While the basic geometric features have been corrected, colour variations between individual scenes are obvious.

a. Path 100, Row 72, acquired on 6 July 2016

b. Path 99, Row 72, acquired on 31 July 2016

c. Path 100, Row 73, acquired on 6 July 2016

d. Path 99, Row 73, acquired on 31 July 2016

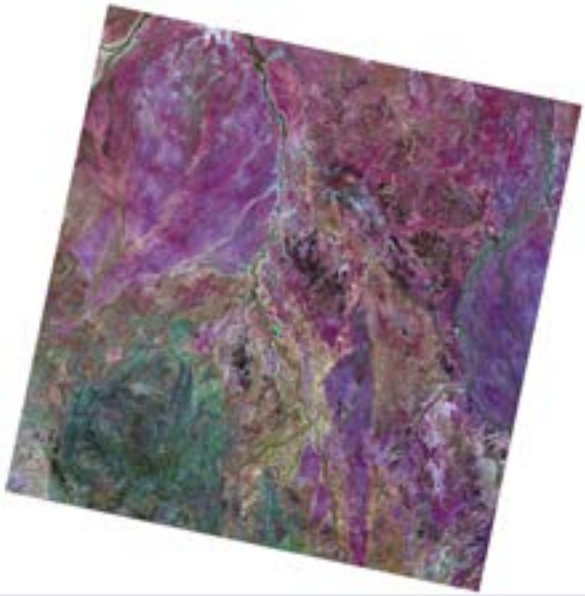


Figure 6.12 Scene mosaic

Once processed to a common geometric and radiometric standard, the individual image scenes in Figure 6.11 can be merged to create a single integrated mosaic. Since the two Landsat image paths used in this mosaic were acquired several weeks apart some variations in water colour—resulting from physical differences in water conditions on the two dates—cannot readily be removed from the mosaic. In this case the final colour balance masks the scene boundary in land features more effectively than in the water.

Colour balancing between images in a mosaic can be challenging, especially when the images contains both land and water features. Since these two Landsat image paths were acquired several weeks apart, some variations in water colour are due to physical differences in water conditions between these two dates. In this case the final colour balance was selected to mask the scene boundary in land features more effectively than in the water. A traditional mosaic based on surface reflectance data (rather than Level 1 data) is shown in Figure 6.13. The differences between these two mosaics highlight the consistency that can be achieved using high level radiometric correction.

Using pixel mosaicking techniques, however (based on the median value of each pixel over a six month window of acquisitions; Roberts *et al*., 2017), a seamless mosaic has been generated for this region (see Figure 6.14). While this image hides scene boundaries effectively, and therefore presents an integrated view of this region, its data values do not apply to a specific point in time. This mosaic should be considered as a 'synthetic' image since each pixel is not an observed value, but a statistical median derived from a set of observations. Such multi-image techniques are detailed in Volume 2D (see also Volume 1A—Excursus 1.1).

Figure 6.13 Surface reflectance mosaic

This scene mosaic over the Gulf of Carpentaria was created from the image scenes shown in Figure 6.11 after conversion of the Level 1 data to surface reflectance imagery.

Figure 6.14 Pixel mosaic

This 'synthetic', geometric median mosaic over the Gulf of Carpentaria was created by selecting the high-dimensional median value of all Landsat-8 OLI observations within a six month period in 2016 into a single image. It is displayed using bands 6, 5 and 3 as RGB.

6.3 Locating Sample Sites in EO imagery

To accurately identify image pixels that correspond to the location of sample sites on the ground requires registration of the geometries of the image and the positioning system used to describe ground locations. In most cases, ground locations will be recorded as mapping coordinates, in which case an accurate registration between the relevant image and map coordinate systems is a prerequisite (see Section 6.1). Trevithick (2015) advises that certain data elements should be recorded for all field sites, namely:

- geographic coordinates—latitude and longitude;
- data and time—including relevant time zone); and
- photographic records—numbered unambiguously, and preferably ordered such that a photograph of the relevant data sheet for a field site precedes the set of photos for that site.

Latitude and longitude offer an unambiguous record of location that is independent of map projection or datum. For geographic coordinates, field records needs to specify whether the coordinates are given as decimal degrees, degrees and decimal minutes or degress/minutes/seconds.

If position is recorded using other map coordinates then all essential map information, such as datum and zone, must also be recorded. When a bearing/distance method is used to record locations, the latitude/longitude of the reference point is essential (and, if known, an accuracy estimate of the positioning instrument should be included).

Field locations that are to be used as ground control points (GCP) for image registration purposes need to be clearly visible in the image (see Section 4.1 and Excursus 4.2) and recorded with the highest spatial accuracy available. Single point measurements, such as GCP, or transects (a series of points along a defined line) can be recorded using a standard GPS instrument (see Volume 1A—Section 10). A time stamp at the start and end of each transect is also recommended (Trevithick, 2015).

Field sites can represent areas of land rather than points or lines. Such plots are typically referred to as quadrats and used to sample the abundance and distribution of biota or biophysical attributes, including spectral characteristics (see Excursus 6.4). While traditionally quadrats have been square, other shapes are also used. For EO image analyses, a ground area corresponding to an image pixel should be the minimum extent for field measurements (Soto-Berelov *et al.*, 2015).

Field data is often collected to assist in labelling and verifying image analyses (Held *et al.*, 2015). Once an image has been accurately registered to map coordinates, samples pixels can be confidently located on the ground. Relevant attributes at these sites (such as vegetation greenness, soil characteristics, chlorophyll content of water or mineral composition) are then recorded and compared with the processed imagery (see Volume 2C and Volume 3). Processing procedures involved in image classification are introduced in Volume 2A—Section 9 and detailed in Volume 2E.

Excursus 6.4—Validation of Surface Reflectance Data

Source: Guy Byrne and Andrew Walsh, Geoscience Australia
Further Information: Malthus *et al.* (2018b)

The design and execution of field studies for the validation of EO data is directly linked to the data product being validated. The most basic validation exercise focuses on defining the *agreement and uncertainties* between a field-derived surface reflectance and an Analysis Ready Data (ARD) surface reflectance product (such as Digital Earth Australia: DEA; see Volume 2A—Section 3.4).

The radiance acquired by an EO sensor represents a complex interplay between the reflective and physical properties of the imaged surface and its structure, the viewing and illumination geometry between the sensor and the imaged scene, and the atmospheric conditions at the time of the sensor overpass (see Volume 1B—Section 3). For field measurements, there is a fundamental tension between what is conceptually understood about the physics of light and surface interactions—what can actually be measured in the field—and, although it is relatively easy to get good field spectra, it is *easier still* to get bad spectra (Milton *et al.*, 2009; Hueni *et al.*, 2017).

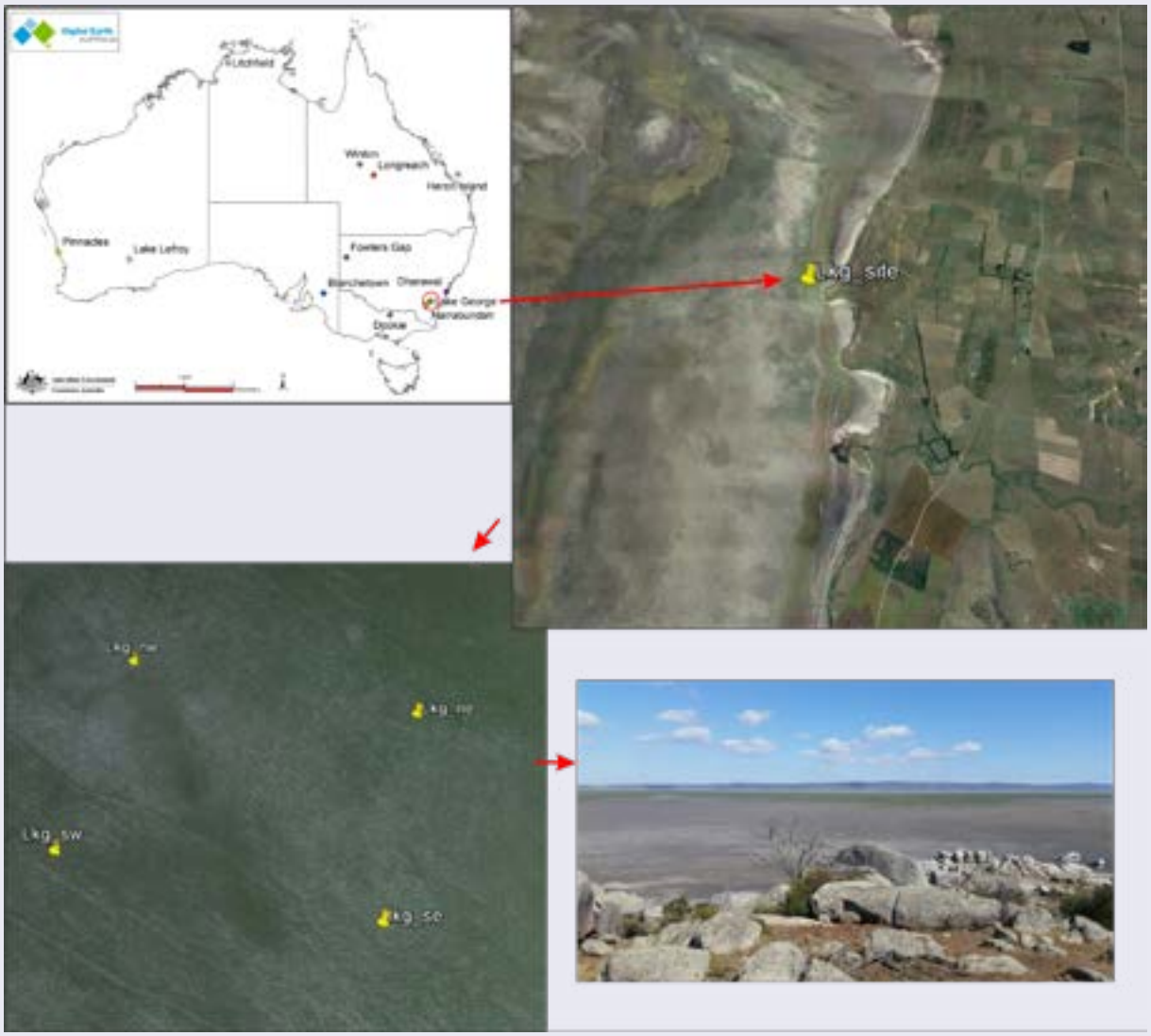
In December 2017, Geoscience Australia began a comprehensive, national validation program of the DEA 'foundation products' within the ARD set comprising Landsat-8 OLI and Sentinel-2 image time series. The first phase of this project has focused on validation of the Landsat-8 OLI and Sentinel-2a and -2b surface reflectance products. The validation project has developed CALVAL protocols for field spectroscopy as detailed in Malthus *et al.* (2018a).

The wider project selected 12 sites across the Australian continent that were deemed to be representative of the range of IBRA bioregions (Thackway and Cresswell, 1995; see also Volume 3A), and hence include diverse land covers with varying surface brightness. At each site, ground spectral data were captured on cloud-free days with a 'clear stable atmosphere' coincident with satellite overpasses. This excursus provides an overview of this project and reports interim results for one site located on Lake George, which was surveyed on six dates between February and July 2018.

The remote sensing community puts major efforts into calibration and validation of sensors, measurements and derived products to quantify and resolve uncertainties.
Schaepman-Strub et al. (2006)

Lake George is an ephemeral lake in NSW, about 50 km north of Canberra, ACT. This ancient closed lake basin, spanning 25 km long by 10 km wide, has a relatively small catchment area and high evaporation rates, so fills and dries over relatively short time intervals. It is the saltiest inland waterbody in NSW so only supports a sparse coverage of salt-tolerant vegetation on a substrate of fine dark silts and cracking clays. The Lake George site in this study is located at -35.01416667S, 149.3969444E, and is classified as IBRA bioregion South Eastern Highlands (SEH), with the IBRA description of 'lake bed cracking clays, dark' (see Figure 6.15). Being flat, homogeneous, and having a relative flat spectral response, it is an ideal spectral validation site.

Figure 6.15 Lake George validation site



The validation of surface reflectance at the pixel scale is relatively straightforward if good practice guidelines are followed. This requires a sampling strategy that 'characterises' variations in surface reflectance for a large, flat, 'homogeneous' patch of ground, by capturing a sufficient number of spectra to represent the pattern of variance in surface radiance. Such patches need to be easily identifiable in the imagery being validated, and not directly adjacent to a contrasting surface cover (Malthus *et al.*, 2018a).

The scale of sample areas on the ground is directly related to the pixel size of relevant imagery. In this case, since the study was validating Landsat-8 OLI with 30 m pixels and Sentinel-2a and -2b with 10–20 m pixels, sample areas of 100 m square were used. Within each sample area, between six and ten transect lines were sampled (see Figure 6.16). Each spectrometer captured a continuous series of spectra (350–2500 nm) as an operator walked along each line ('smear mode') and reflectance panel readings were taken at the beginning of each transect for calibration.

Figure 6.16 Spectral sampling protocol

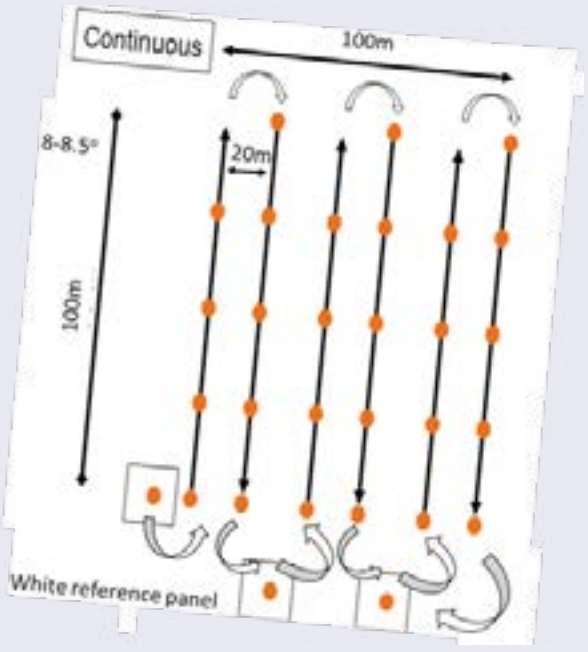


Figure 6.17 Lake George field sampling on 3 May 2018

Photographs of the field site and atmospheric conditions are an important part of any EO image validation exercise.

In theory, measuring surface reflectance requires simultaneous readings of:

- the total downwelling irradiance—the sum of direct sunlight and diffuse light scattered by the atmosphere; and
- the upwelling, reflected radiance from a target surface.

However, most field spectrometers are 'single field of view' sensors, so two separate measurements are taken:

- the total downwelling irradiance—measured as the radiance from a white reference panel; and then
- the upwelling radiance at ground level.

An important aspect of any EO validation exercise is taking comprehensive field records and detailed notes of the sampling protocol to accompany actual measurements, including photography to show weather and surface conditions (see Figure 6.17).

While validation of EO data requires field data to be accurately located in the imagery, some level of uncertainty is inevitable due to the effects of rasterisation and geocoding (see Volume 1B). Spectrally homogeneous field sites located within larger spectrally homogeneous patches tend to mitigate the impact of locational uncertainty (see Figure 6.18 and Figure 6.19). For the Lake George site, Figure 6.20 compares the locations of Landsat-8 OLI pixels (shown as black dots) and field site spectra captured on 3 May 2018 (shown as coloured dots).[8]

Figure 6.21 compare the average reflectances derived from field and image measurements for each image band, while Figure 6.22 offers a band-by-band comparison of reflectances for satellite and field data. These results will be used to validate the DEA surface reflectance products (Landsat-8 OLI and Sentinel-2 MSI) and characterise the uncertainty and fitness for purpose of these products. This analysis will also extend to the comparison of different ARD surface reflectance products based on the Landsat-8 OLI and Sentinel-2 MSI. The results are expected to lead to an improved understanding of the sensitivity and efficacy in applying particular processing models and corrections to these data, and thus support the ongoing development of ARD from a range of image sources.

[8] Note that the ground-based spectral geocoding is derived using a small Garmin 'navigational aid' Bluetooth GPS receiver and is not differentially (RTK) corrected. The lines appear to wander because of location 'solution' ambiguities, which are probably due to the GPS sensor being mounted on a field laptop, which was occasionally occluded by the operator. However, since the location uncertainty is well within one image pixel, this variation is not considered to be signficant.

Figure 6.18 Location of Lake George field site

This Landsat-8 OLI image was acquired on 3 May 2018, spans approximately 2 km by 2 km, and is displayed using bands 4, 3, 2 as RGB. The precise location of this field site is indicated by the central white box.

Figure 6.19 Landsat-8 OLI field site pixels

a. All pixels within field site (shown as white box in Figure 6.18)

b. Field site pixels that were sampled by spectral transects

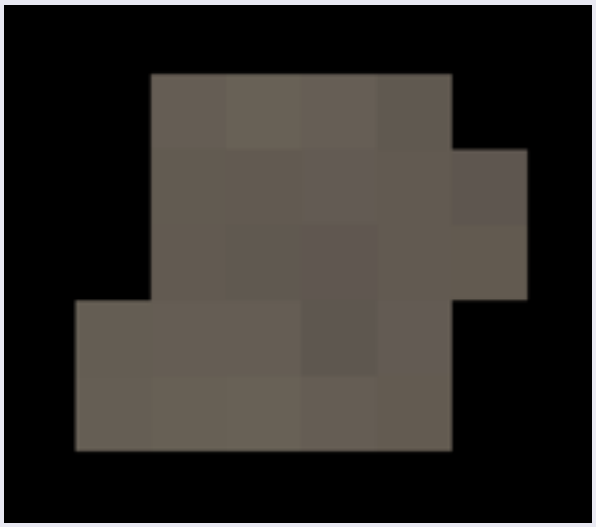


Figure 6.20 Lake George spectra transect lines overlaid on Landsat-8 OLI image grid

The centres of Landsat-8 OLI pixels are shown as black dots and the GPS locations of field measurements are shown as coloured dots, with different colours used for each transect line. Coordinates are relative to the central location of the field site (-35.01416667 S, 149.3969444 E) in Australian Albers projection (EPSG: 3577).

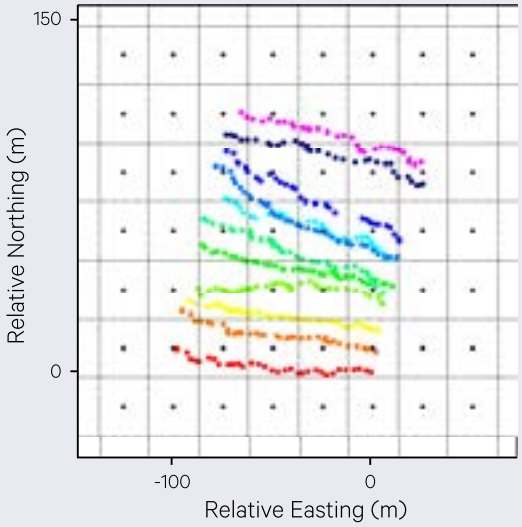


Figure 6.21 Field spectra versus Landat-8 OLI spectra on 3 May 2018

Y axis shows average reflectance measured by field spectra (red) and Landsat-8 OLI image (blue).

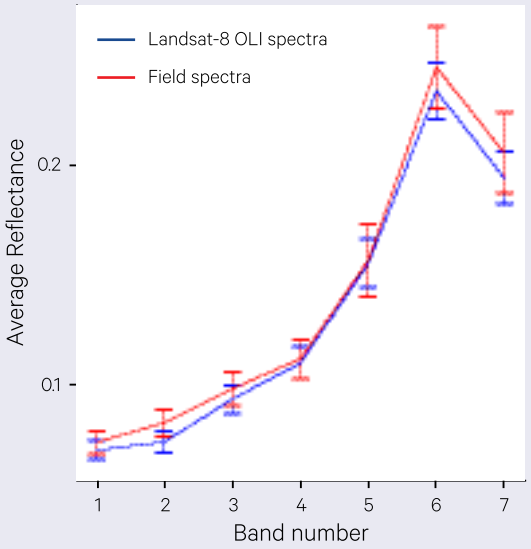
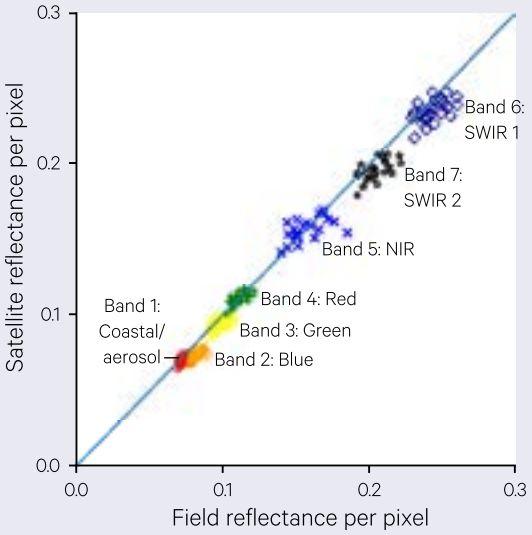
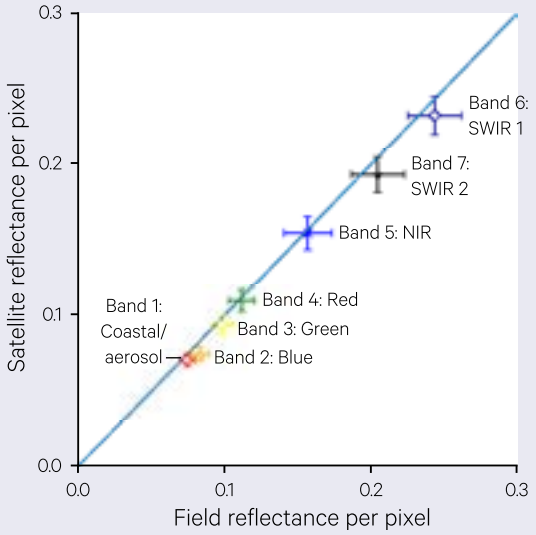


Figure 6.22 Comparison of field and satellite reflectances for each band

a. Based on field site pixels, reflectances measured by field spectra are compared with satellite data, with different colours representing different satellite bands. The line of equality between satellite and field data is shown as blue.



b. The average of field site spectra over the field site is compared with satellite data, with different colours representing different satellite bands. The line of equality between satellite and field data is shown as blue and error bars represent standard deviation of satelite/field pixels over the field site.



6.4 Further Information

Australian Geographic Reference Image (AGRI):

Technical Report: Lewis *et al*, (2011)

Data: *https://data.gov.au/dataset/agri-the-australian-geographic-reference-image*

Field work guidelines:

Held *et al*. (2015)

6.5 References

Barr, S., Wang, L.-W., Ravanbakhsh, M., Pasfield, M., and Lewis, A. (2010). *National Ground Control Point Database: a new national dataset for Australia*. Proc. 15th Australasian Remote Sensing and Photogrammetry Conference, Alice Springs, September 2010.

Fraser, C. S., Ravanbakhsh, M., and Awrangjeb, M. (2009). Precise georeferencing in the absence of ground control: a strip adjustment approach. *Int. Arch. Photogrammetry, Remote Sensing and Spatial Information Sciences*, Hannover, Germany, 38, I-4-7/W5.

Held, A., Phinn, S., Soto-Berelov, M., and Jones, S. (Eds.) (2015). *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products*. Version 1.1. TERN AusCover, ISBN 978-0-646-94137-0.

Hueni, A., Damm, A., Kneubuler, M., Schlapfer, D., and Schaepman, M.E. (2017). Field and Airborne Spectroscopy Cross Validation–Some Considerations. *IEEE J. Selected Topics in Applied Earth Observation and Remote Sensing* 10(3), 1117–1134. doi: 10.1109/JSTARS.2016.2593984

Lewis, A., Wang, L.-W., and Coghlan, R. (2011). *AGRI: The Australian Geographic Reference Image. A Technical Report*. GeoCat#72657. Geoscience Australia, Canberra. ISBN 978-1-921954-47-4.

Malthus, T.J., Ong , C., Lau, I., Fearns, P., Byrne, G.T., and Thankappan, M. (2018a). *A community approach to the standardised validation of surface reflectance data. A technical handbook to support the collection of field reflectance data. Release Version 1.0*. CSIRO, Australia. ISBN: 978-1-4863-0991-7.

Malthus, T.J., Byrne, G.T., Walsh, A.J., Thankappan, M., Ong, C., Lau, I., and Fearns, P. (2018b). *DEA Surface Reflectance Validation Project*. Geoscience Australia, Canberra.

Milton, E.J., Schaepman, M. E. Anderson, K., Kneubühler, M., and Fox, N. (2009). Progress in field spectroscopy. *Remote Sensing of Environment* 113 (Supplement 1), S92–S109.

Nagi, R. (2010). *On map scale and raster resolution*. ArcGIS Blog, ESRI. *https://www.esri.com/arcgis-blog/products/product/imagery/on-map-scale-and-raster-resolution/*

Roberts, D., Mueller, N., and Mcintyre, A. (2017). High-Dimensional Pixel Composites From Earth Observation Time Series. *IEEE Transactions on Geoscience and Remote Sensing* 55(11), 6254–6264. doi: 10.1109/TGRS.2017.2723896

Rottensteiner, F., Weser, T., Lewis, A. and Fraser, C.S. (2009). A Strip Adjustment Approach for Precise Georeferencing of ALOS Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 47 (12; Part I), 4083–4091.

Schaepman-Strub, Schaepman, M.E., Painter, T.H., Dangel, S., and Martonchik, J.V. (2006). Reflectance quantities in optical remote sensing–definitions and case studies. *Remote Sensing of Environment* 103, 27–42.

Soto-Berelov, M., Jones, S., Farmer, E., and Woodgate, W. (2015). Review of validation standards of biophysical Earth Observation products. Ch 2 in *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products*. Version 1.1. (Eds: Held, A., Phinn, S., Soto-Berelov, M., and Jones, S.). TERN AusCover, ISBN 978-0-646-94137-0.

Thackway, R., and Cresswell, I. (1995). *An Interim Biogeographic Regionalisation for Australia: a framework for setting priorities in the National Reserves System Cooperative Program Version 4*. Australian Nature Conservation Agency, Canberra. *http://www.environment.gov.au/land/nrs/science/ibra#ibra*

Tobler, W. (1987). Measuring Spatial Resolution. *Proc. Land Resources Information Systems Conference, Beijing*, pp. 12–16.

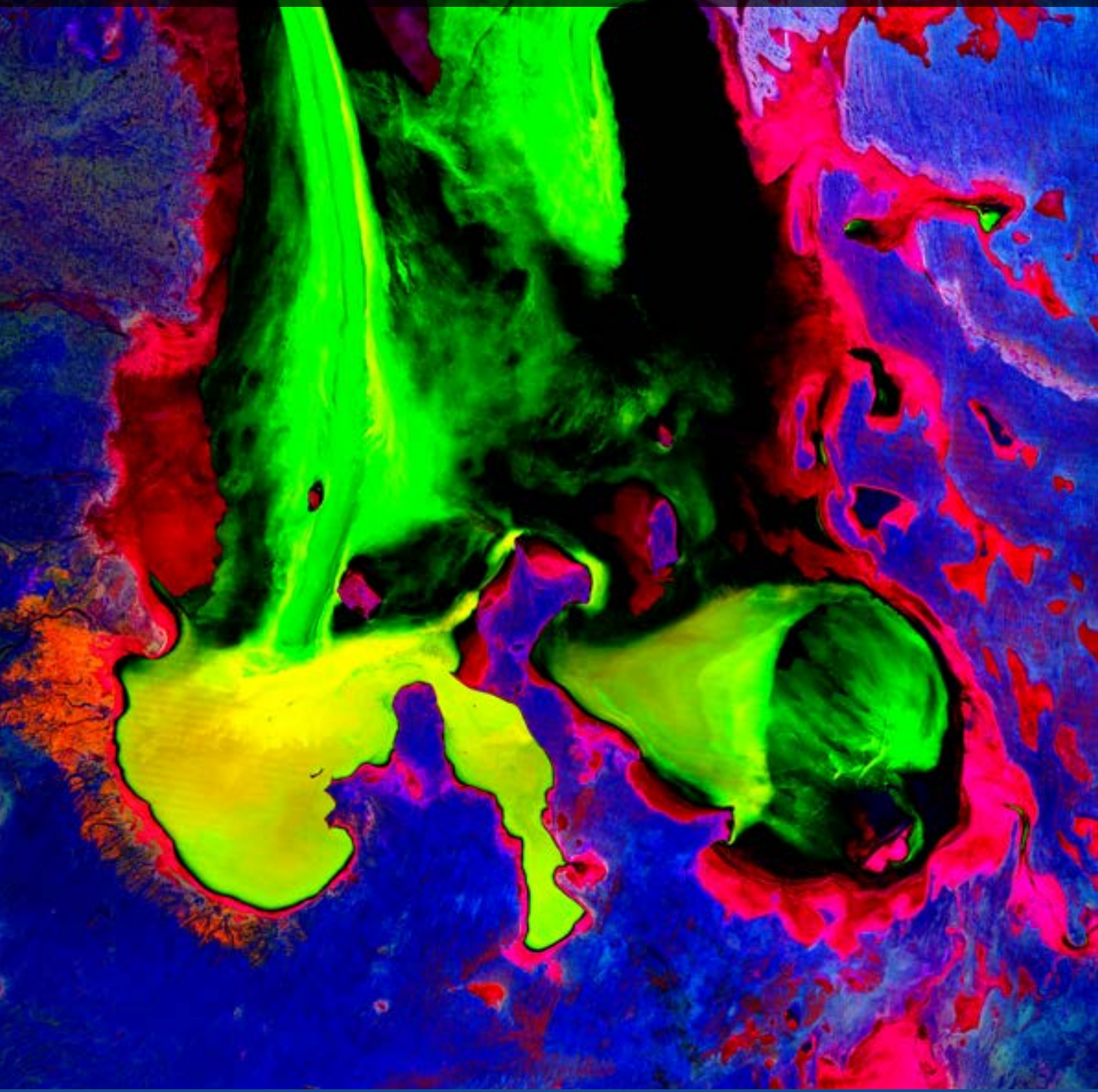
Tobler, W. (1988). Resolution, Resampling, and All That. In *Building Data Bases for Global Science* (Eds: H. Mounsey and R. Tomlinson) pp. 129–137. London, Taylor and Francis.

Trevithick, R. (2015). Field data collection and management for Earth Observation image validation. Ch 3 in *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products*. Version 1.1. (Eds: Held, A., Phinn, S., Soto-Berelov, M., and Jones, S.). TERN AusCover, ISBN 978-0-646-94137-0.

Weser, T., Rottensteiner, F., Willneff, J., Fraser, C.S. (2008). An Improved Pushbroom Scanner Model for Precise Georeferencing of ALOS PRISM Imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol XXXVII Part B1. Beijing.

Earth Observation:
Data, Processing and Applications

Volume 2C: Processing—Image Transformations



This publication is available in PDF format at *eopublications.au*.
We welcome your comments regarding the readability and usefulness of this report.
To provide feedback, please email *feedback@eopublications.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
978-0-6482278-1-6

Copyright:
All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, save for content supplied by third parties, and logos. Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. The full licence terms are available from *https://creativecommons.org/licenses/by/4.0/legalcode*. A summary of the licence terms is available from *https://creativecommons.org/licenses/by/4.0/*.

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Recommended Citation for Volume 2C:
CRCSI (2019) *Earth Observation: Data, Processing and Applications. Volume 2C: Processing—Image Transformations*. (Harrison, B.A., Jupp, D.L.B., Lewis, M.M., Sparks, T., Mueller, N., Byrne, G.) CRCSI, Melbourne.

Latest Revision:
October 2025

Background image on previous page: Principal Components Analysis of Landsat-5 image of Lake Eyre, South Australia, filling on 9 May, 2009, displayed using PCs 3, 4, 1 as RGB.
Source: Norman Mueller, Geoscience Australia

Acknowledgements

Production of this series of texts would not have been possible without the financial support of CSIRO, CRCSI, GA and BNHCRC, input from members of the editorial panels and direction from members of the various advisory panels. Administrative support from FrontierSI is also gratefully acknowledged.

Volumes 1 and 2 of this series are based on text originally published in Harrison and Jupp (1989, 1990, 1992 and 1993)[1]. Many illustrations and some text from these publications have been reproduced with permission from CSIRO.

Other contributors are gratefully acknowledged:

- reviewers: Stuart Phinn and Alfredo Huete, for comments on early drafts;
- illustrations: Norman Mueller kindly supplied most of the lovely images; other contributors of graphical material include: Paul Hutton, Tony Sparks, Megan Lewis, David Jupp, Craig Shepherd, and EagleView[2];
- excursus: Tony Sparks, Megan Lewis, and Norman Mueller;
- Daniel Rawson (Accessible Publication & Template Design) for layout and formatting; and
- Carl Davies (CMDphotographics) for selected graphical illustrations.

We thank those owners of copyrighted illustrative material for permission to reproduce their work. Credits for individual illustrations are provided below the relevant graphic.

All volumes in this series are covered by the copyright provisions of CC BY 4.0 AU.

1 Harrison, B.A., and Jupp, D.L.B. (1989) *Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual.* CSIRO, Melbourne. 156pp.
Harrison, B.A., and Jupp, D.L.B. (1990) *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.
Harrison, B.A., and Jupp, D.L.B. (1992) *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual.* MPA, Melbourne.
Harrison, B.A., and Jupp, D.L.B. (1993) *Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual.* MPA, Melbourne.

2 Formerly Spookfish

Table of Contents

Volume 2C: Processing—Image Transformations a

Acknowledgements i

Review Topics 1

1 Overview **3**
- **1.1** The Electromagnetic Spectrum 4
- **1.2** Image Data Characteristics 5
  - **1.2.1** Structure 5
  - **1.2.2** Scales of measurement 6
  - **1.2.3** Data scales 6
  - **1.2.4** Sampling dimensions 7
- **1.3** Spectral Statistics 8
  - **1.3.1** Single channel statistics 8
  - **1.3.2** Channel histogram 8
  - **1.3.3** Multi-channel statistics 9
  - **1.3.4** Channel crossplots 10
- **1.4** Processing Operations 11
- **1.5** Further Information 11
- **1.6** References 11

2 Rescaling **13**
- **2.1** Linear Rescaling 16
- **2.2** Non-linear Rescaling 17
- **2.3** Lookup Tables (LUT) 19
- **2.4** Destriping 19
- **2.5** Histogram Manipulation 22
- **2.6** Further Information 22
- **2.7** References 22

Single Channel Filtering 23

3 Filtering Operations **25**
- **3.1** Operation 26
- **3.2** Filter Weights 28
- **3.3** Filter Size 30
- **3.4** Thresholding 30
- **3.5** Adaptive Filtering 30
- **3.6** Homomorphic Filtering 31
- **3.7** Further Information 32
- **3.8** References 32

4 Smoothing (Low Pass) **33**
4.1 Average Filters 35
4.2 Median Filters 38
4.3 Modal Filters 40
4.4 Edge-Preserving Filters 41
4.5 Despiking (Noise Reduction) 43
4.6 Filling Missing Values (Interpolation) 45
4.7 Further Information 46
4.8 References 46

5 Highlighting Edges **47**
5.1 Edge Detection (Differential) 47
5.1.1 Image-based derivatives 50
5.1.2 Uses of edge detection 54
5.2 Edge Enhancement (High Pass) 55
5.3 Further Information 57
5.4 References 57

6 Highlighting Surface Variation **59**
6.1 Texture (Local Variance) 59
6.2 Relief Shading 62
6.3 Exposure (Directional) 64
6.4 Surface Curvature (Surface Shape) 67
6.5 Further Information 69
6.6 References 69

Linear Operations 71

7 Affine Transformations **73**
7.1 General Affine Transformation 75
7.2 Removing Spatial 'Noise' 78
7.2.1 Using a reference channel 78
7.2.2 Removing 'Limb Brightening' 79
7.3 Other Affine Transformations 80
7.4 Further Information 80
7.5 References 80

8 RGB-HSI Conversions **81**
8.1 RGB Colour Model 81
8.2 HSI Colour Model 83
8.3 Chromaticity Transformation 85
8.4 Pan-Sharpening 86
8.5 Adding Relief Shading 88
8.6 Further Information 91
8.7 References 91

9 Principal Components Analysis **93**
9.1 PCA Transformation 93
9.2 Interpretation of Principal Components 97
9.3 Transformation parameters 99
9.4 Decorrelation Stretching 100
9.5 Data Reduction 102
9.6 Highlighting Specific Features 102
9.7 Multi-scale Intensity Enhancement 103
9.8 Change Detection 104
9.9 Further Information 104
9.10 References 104

Non-Linear Operations 105

10 Channel Ratios **107**
10.1 Computation Methods 112
10.1.1 Simple ratios 112
10.1.2 Computing Products and Ratios using Log Channels 114
10.1.3 Smoothed ratios 114
10.1.4 Directed band ratioing 115
10.2 Processing Parameters 117
10.2.1 Dark values 117
10.2.2 Calibration impact 118
10.3 Further Information 119
10.4 References 119

11 Vegetation Indices **121**
11.1 Vegetation Greenness 121
11.1.1 Simple Ratio 126
11.1.2 Normalised Difference Vegetation Index (NDVI) and variants 126
11.1.3 Kauth-Thomas Greenness Transformation 128
11.2 Vegetation Water Content 131
11.3 Further Information 132
11.4 References 132

List of Figures

Figure 1.1 Electromagnetic spectrum 4

Figure 1.2 Comparison of spectral bands in different EO sensors 5

Figure 1.3 Structure of EO image 5

Figure 1.4 Data scales 6

Figure 1.5 Logarithmic and exponential functions 7

Figure 1.6 Image channel histograms 9

Figure 1.7 Format of variance/covariance matrix 9

Figure 1.8 Example image channels 10

Figure 1.9 Example crossplots 10

Figure 2.1 Examples of rescaling 14

Figure 2.2 Effect of rescaling operation on image data range 16

Figure 2.3 Logarithmic scaling 17

Figure 2.4 LUT operations 19

Figure 2.5 Rotated image striping pattern 20

Figure 2.6 Destriping example 21

Figure 3.1 Operation of an average filter 26

Figure 3.2 Convolution of two filters 27

Figure 3.3 Treatment of edge pixels in an image during filtering operation 28

Figure 3.4 Relationship between smoothing, edge enhancement and edge detection filters 29

Figure 3.5 Adaptive filtering 30

Figure 3.6 Wallis filter 31

Figure 4.1 Effect of different filter sizes and operations 34

Figure 4.2 Effect of smoothing filter on image data 35

Figure 4.3 Effect of different filter sizes and weights 36

Figure 4.4 Weights for gentle smoothing filter 36

Figure 4.5 Full average filters with differing weights 37

Figure 4.6 Example of destriping filter 37

Figure 4.7 Operation of a median filter 38

Figure 4.8 Average versus median smoothing 39

Figure 4.9 Example of moving average and moving median 39

Figure 4.10 Operation of a modal filter 40

Figure 4.11 Modal filtering of classified image 41

Figure 4.12 Effect of filter size on edge-preserving filters 42
Figure 4.13 Comparison of average and edge-preserving filters 42
Figure 4.14 Edge-preserving filter applied to EO image 43
Figure 4.15 Filter-based interpolation 45
Figure 4.16 Example of average filter 45
Figure 4.17 Average filter interpolation 46
Figure 5.1 First and second derivative for spectral transect 48
Figure 5.2 Examples of derivative filters 50
Figure 5.3 Effect of filter direction on differential filters 50
Figure 5.4 Combined smoothing and differentiating filter 51
Figure 5.5 Horizontal and diagonal derivative filters 51
Figure 5.6 Norm functions 52
Figure 5.7 Roberts cross operator 52
Figure 5.8 Sobel operator 52
Figure 5.9 Prewitt operator 52
Figure 5.10 Second difference filter 53
Figure 5.11 First and second differences 53
Figure 5.12 Laplacian edge detector 53
Figure 5.13 Bi-directional filter 53
Figure 5.14 Horizontal and vertical filters 54
Figure 5.15 Derivation of edge enhancement filter 55
Figure 5.16 Effect of filter size on edge enhancement filters 56
Figure 5.17 Operation of edge enhancement filter 56
Figure 5.18 Examples of edge enhancement filters 57
Figure 6.1 Different texture filters 60
Figure 6.2 Constant variance images 61
Figure 6.3 Cosine reflectance model 62
Figure 6.4 Parameters derived from elevation 63
Figure 6.5 Total insolation 64
Figure 6.6 Exposure transformation 64
Figure 6.7 Representing compass direction as X and Y components 65
Figure 6.8 Filter weights for exposure transformation 65
Figure 6.9 Exposure (directional) filtering 66
Figure 6.10 Parameters related to surface curvature 68
Figure 6.11 Curvature composite 68

Figure 7.1 General matrix operation 74
Figure 7.2 Reflecting image axes about the line *Y* = *X* 75
Figure 7.3 Matrix equation for channel inversion 77
Figure 7.4 Example matrix equation 77
Figure 7.5 Using channel crossplot to derive regression line with reference data 78
Figure 7.6 Filter to remove image noise 79
Figure 8.1 RGB colour cube 82
Figure 8.2 RGB triangular coordinates 82
Figure 8.3 HSI coordinates 83
Figure 8.4 Hue, Saturation and Intensity dimensions 84
Figure 8.5 Chromaticity transformation 85
Figure 8.6 Matrix equation to merge high resolution data 86
Figure 8.7 Integrating imagery with differing spatial resolution 87
Figure 8.8 Modifying colour brightness using relief shading 88
Figure 8.9 Matrix equation to add relief shading 88
Figure 8.10 Channel substitution using HSI transform 89
Figure 8.11 Single channel presented as colour composite 90
Figure 9.1 Redefining image axes 94
Figure 9.2 Principle of PCA transformation 94
Figure 9.3 Eigenvectors of a matrix 95
Figure 9.4 Principal components transformation 95
Figure 9.5 Original image channels 97
Figure 9.6 Colour composites 98
Figure 9.7 Principal component channels 99
Figure 9.8 Principal Component transformation 100
Figure 9.9 RGB colour cube showing original image data range 101
Figure 9.10 Decorrelation stretching 101
Figure 9.11 Software defoliant 103
Figure 10.1 Idealised spectral reflectance curves 108
Figure 10.2 Landsat-8 OLI image channels 109
Figure 10.3 Ratio channels 110
Figure 10.4 Ratios to highlight lithology 111
Figure 10.5 Colour composite of band ratios 111
Figure 10.6 Operation of simple channel ratio 112
Figure 10.7 Impact of ratioing on topographic shading 113

Figure 10.8 Smoothed ratio operation 115

Figure 10.9 Directed band ratioing 116

Figure 10.10 Effect of dark value adjustment in channel ratios 117

Figure 11.1 Reflectance characteristics of typical green leaf structure 122

Figure 11.2 Summer and winter example images 123

Figure 11.3 Individual channels for example images 124

Figure 11.4 Four vegetation indices for example images 125

Figure 11.5 Normalisation effect in 'Normalised Difference Vegetation Index' (NDVI) 126

Figure 11.6 The 'Tasselled Cap' feature in Landsat MSS image data 129

Figure 11.7 Kauth Thomas greenness transformation matrix 129

Figure 11.8 Typical green vegetation spectrum 131

List of Tables

Table 1.1 Sampling dimensions in EO imagery 7

Table 3.1 Relative processing times for different filter sizes 30

Table 6.1 Curvature image categories 68

Table 7.1 Matrix coefficients 74

Table 9.1 Principal component percentage variation 96

Table 9.2 Correlation matrix for original image 98

Table 9.3 PC transformation matrix 98

Table 10.1 Popular EO image channel ratios 108

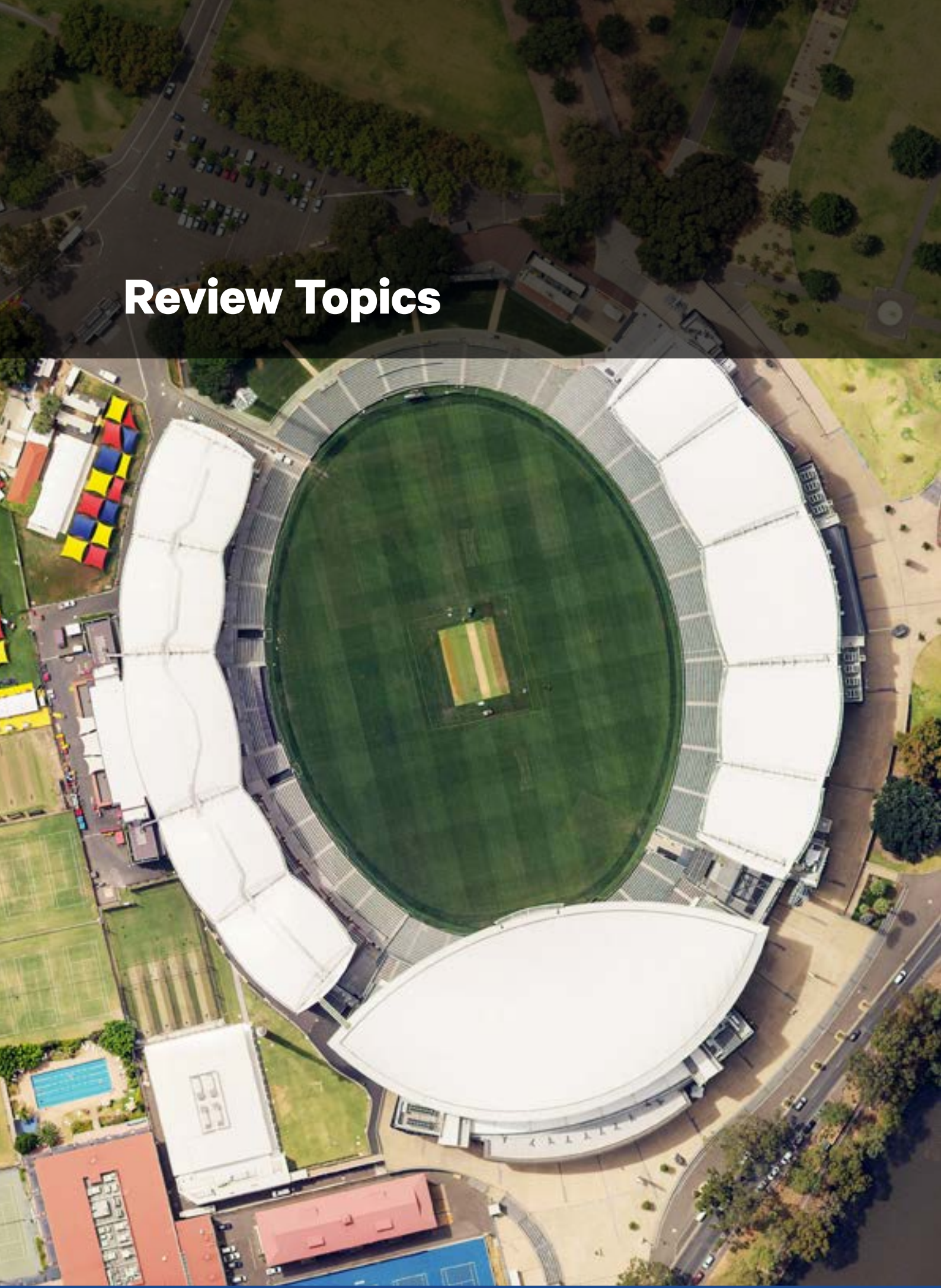
Table 11.1 NDVI variations 127

Table 11.2 Vegetation water content indices 131

List of Excurses

Excursus 2.1—Rescaling results 13
Excursus 3.1—Filter convolution 27
Excursus 3.2—Example of Filter Operations 29
Excursus 5.1—Differential calculus review 47
Excursus 5.2—Commonly used differential filters 52
Excursus 6.1—High Pass Filtering 61
Excursus 7.1—Review of Matrix Notation 74
Excursus 7.2—Using Affine Transformations 76
Excursus 8.1—Examples of HSI-based transformations 89
Excursus 9.1—PCA statistics 96
Excursus 9.2—PCA Example 97
Excursus 10.1—Commonly Used Ratios 108
Excursus 11.1—Comparison of Vegetation Indices 123

Review Topics



The following two sections review some principles that were introduced in earlier sub-volumes of this series of publications and have particular relevance to EO image transformations:

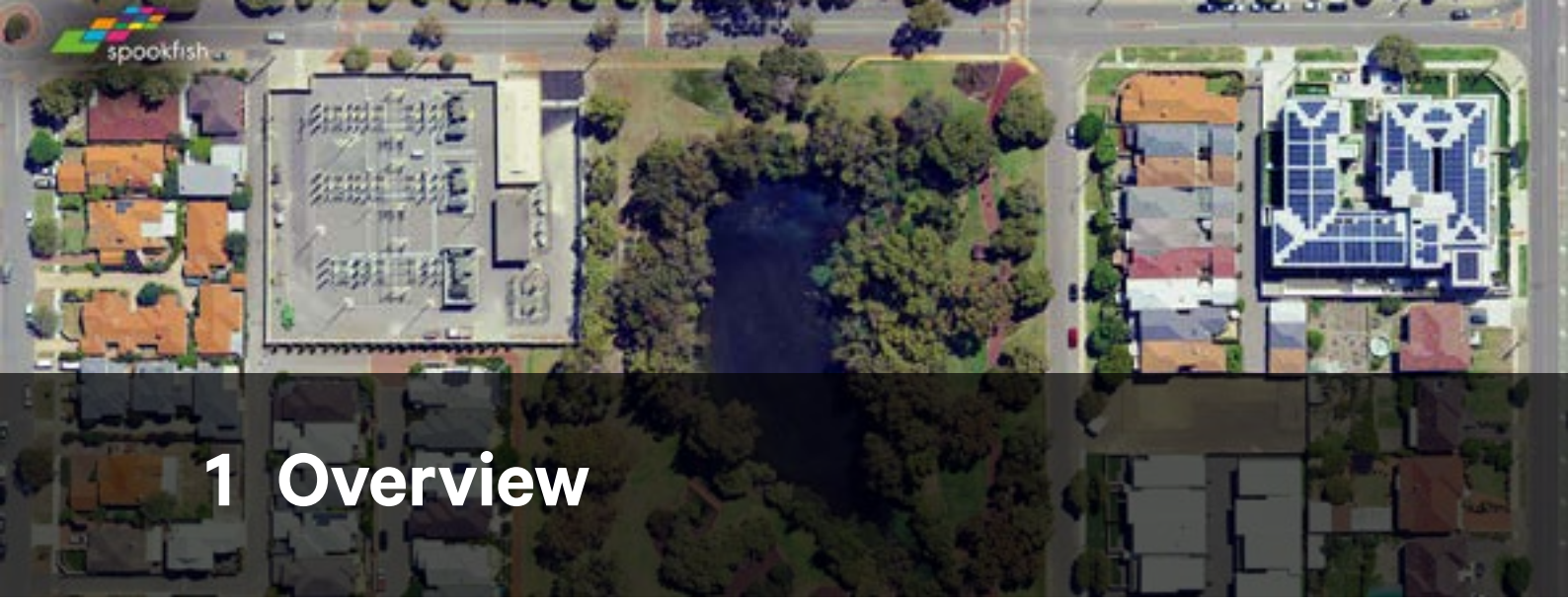
- overview—a reminder of some foundations when processing EO imagery (see Section 1); and
- rescaling—changing the data range for an image channel (see Section 2).

Contents

1 Overview 3

2 Rescaling 13

Background image on previous page: High resolution aerial image over Adelaide Oval, South Australian, acquired on 23 January 2018. **Source:** © EagleView



1 Overview

Basic image operations are introduced in Volume 2A of this series and specific analytical tools and methods are covered in Volumes 2B (Image Rectification), 2D (Image Integration) and 2E (Image Classification). This volume details various image transformations, which may be used for interpretation, analysis or modelling. The modelling stage relates image values to data from other sources and accesses a wide variety of transformations to determine the relationships between different data types, and is further discussed in the context of specific application areas in Volume 3: Applications.

Image transformations allow image data values to be changed by some consistent and defined process. Such processes are usually described in terms of modifying some or all of the existing pixel values in an image. While modern image processing systems allow transformations to be used without reference to the underlying algorithms, knowledge of their mathematical foundations does ensure that such transformations are applied appropriately. In the context of Earth Observation (EO), appropriate processing is more likely to result in consistent and repeatable results, and enables users to use transformed data with understanding and confidence.

In this introductory section we will review some of the basic concepts that underly image transformations commonly applied to EO data, namely:

- the electromagnetic spectrum—wavelength regions and commonly used EO sensor bands (see Section 1.1);
- image data characteristics—structure, scales of measurement, data scales and sampling dimensions (see Section 1.2);
- spectral statistics—for single channels and pairs of channels (see Section 1.3); and
- processing operations—linear versus non-linear (see Section 1.4).

The universe is transformation: life is opinion.
(Marcus Aurelius)

Background image: High resolution aerial image of residential area in North Perth, acquired on 18 September 2017. **Source:** © EagleView

1.1 The Electromagnetic Spectrum

One of the most commonly detected energy sources for EO is electromagnetic radiation (EMR; see Volume 1). For convenience, the electromagnetic spectrum is labelled in terms of wavelength regions (see Figure 1.1).

Commonly accepted ranges for the EMR spectral regions most relevant to EO are:

- visible: 0.38–0.7 μm;
- infrared (IR): 0.7–1,000 μm; and
- microwave: > 1 mm–1 m.

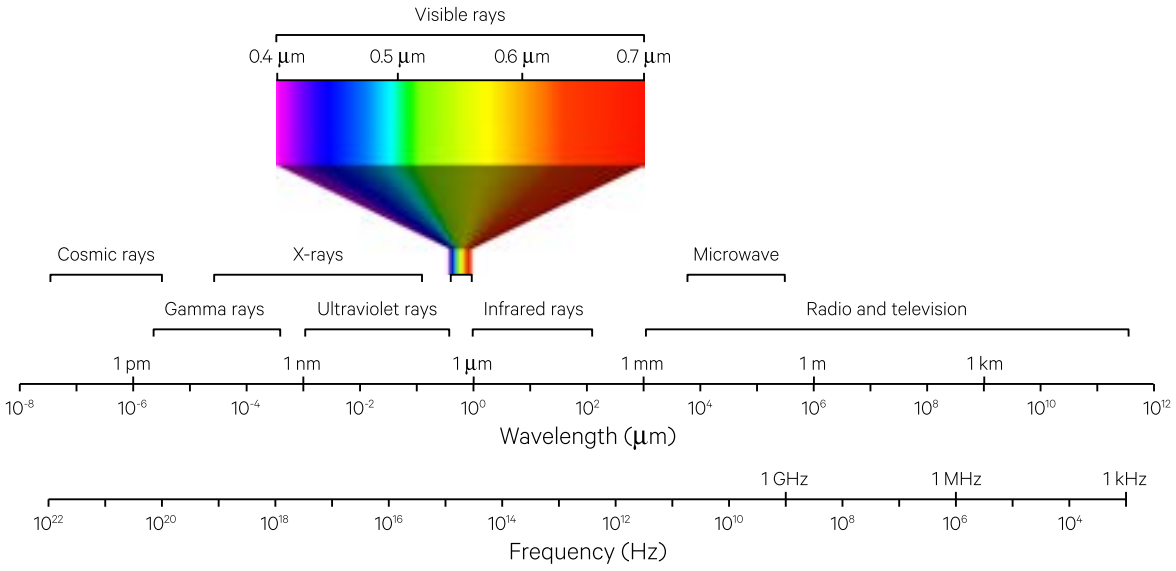
Infrared wavelengths are often further sub-divided into the regions:

- near infrared (NIR): 0.7–1.3 μm;
- short wave infrared (SWIR): 1.3–3 μm;
- middle infrared (MIR) or mid-wavelength (or medium wave) infrared (MWIR): 3–8 μm;
- thermal infrared (TIR) or long wavelength infrared (LWIR): 8–15 μm; and
- far infrared (FIR): 15–1,000 μm.

These terms will be used in the following text in reference to EO image data channels. Within these broad regions, multiple EO image bands can be defined. For example, the wavelength ranges detected by five different sensors carried on recent and current EO satellites are summarised in Figure 1.2.

Figure 1.1 Electromagnetic spectrum

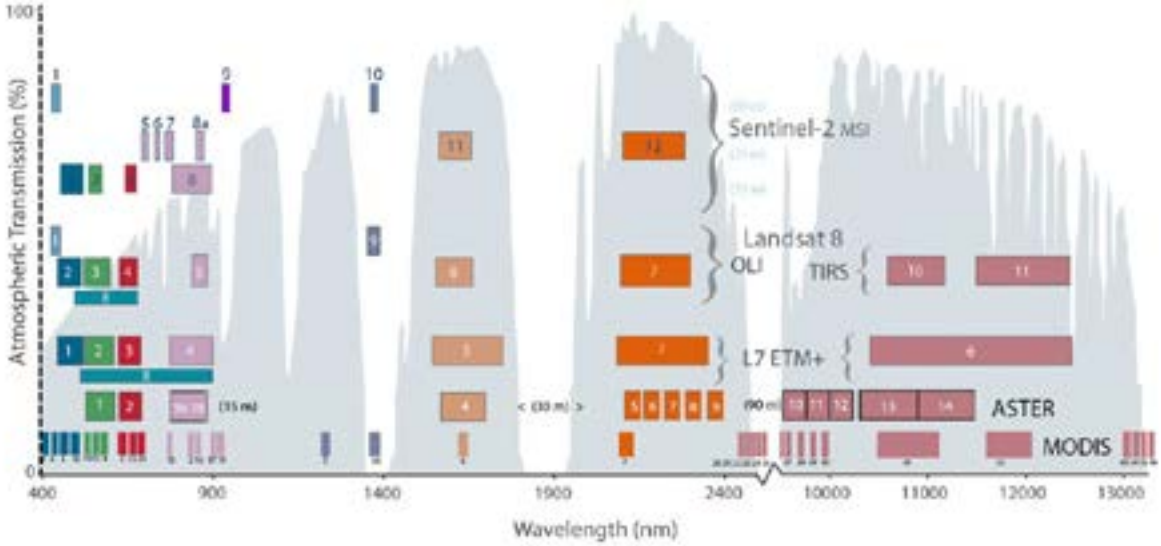
Note: Cosmic rays are high energy, charged particles that travel at close to the speed of light. They are not EMR but included for comparison. Scales are logarithmic



As Gregor Samsa awoke one morning from uneasy dreams
he found himself transformed in his bed into a gigantic insect.
(Franz Kafka, The Metamorphosis)

Figure 1.2 Comparison of spectral bands in different EO sensors

This diagram compares spectral bands from five commonly used, public good satellites by positioning their spectral bandwidths in the optical and thermal regions of the electromagnetic spectrum. The percentage of EMR transmission through the atmosphere in these ranges of wavelengths is shown in grey (see Volume 1B). Note that the wavelength axis is not continuous.



Source: NASA. Retrieved from: *https://landsat.gsfc.nasa.gov/wp-content/uploads/2016/10/all_Landsat_bands.png*

1.2 Image Data Characteristics

Characteristics of image data are reviewed below in terms of:

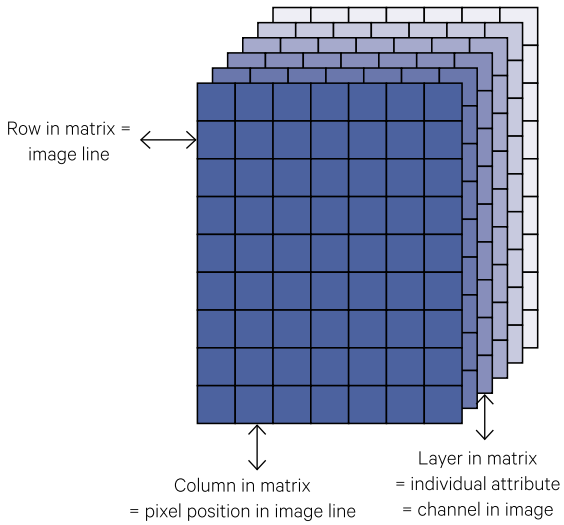
- structure (see Section 1.2.1);
- measurement scales (see Section 1.2.2);
- data scales (see Section 1.2.3); and
- sampling dimensions (see Section 1.2.4).

1.2.1 Structure

The structure of digital imagery is introduced in Volume 2A—Section 1. By assuming that the tessellation grid used to record EO image values is regular within and between all image channels (or bands), the resulting image dataset can be viewed as a three-dimensional structure (see Figure 1.3). This three-dimensional dataset can then be examined, analysed and transformed used a range of statistical tools. Transformed imagery may then be correlated with other datasets, such as crop yield or water depth (see Volume 3).

Figure 1.3 Structure of EO image

A multi-channel image comprises three data dimensions—pixels, lines and channels. Pixels (or 'elements' or 'samples') along the lines can be considered as columns in a matrix. Lines in the image are like rows of the matrix, with each channel being an attribute layer in the matrix.



1.2.2 Scales of measurement

Statisticians differentiate four scales of measurement for defining numeric variables (Anderberg, 1973):

- nominal—refers simply to categories which define different states (such as colour names for different hues, e.g. red, yellow, green);
- ordinal—involves an implied ordering between different states (such as categories of colour brightness, or ranges in topographic elevation);
- interval—measures the extent of difference between states (such as relative illumination for a given wavelength or bandwidth); and
- ratio—denotes a scaled measure with a meaningful zero point (such as absolute illumination for a given wavelength or bandwidth). The ratio of two ratio variables could be computed without any adjustment to their values.

EO image channels would generally be considered as interval variables. With appropriate processing, however, they could be treated as ratio variables. Classified data can also become image channels, which would be considered as nominal or possibly ordinal data (see Volume 2E). Much ancillary data, such as land cover categories, would most likely be nominal or ordinal, although some, such as DEM, may be interval or even ratio (see Volume 2D). These distinctions are important when considering the types of transformations that may be applied to EO imagery. The relevance of scales of measurement is further discussed in Volume 2D (in terms of integrating ancillary data) and in Volume 2E (in terms of the image classification process).

1.2.3 Data scales

Interval and ratio variables (see Section 1.2.2) can be referenced to a linear scale or a logartihmic scale. A linear scale has equal divisions for equal values (see Figure 1.4a). In a logarithmic (or log) scale, data points are represented by the logarithm of their value (see Figure 1.4b).

A logarithmic scale is generally used when there is a wide range of values and the distribution of data is skewed. While EO image values are acquired on a linear scale, it can be advantageous to transform image data to a logarithmic scale to enhance low values in one or more channels (see Section 2.2) or apply other transformations (see Section 10.2.2).

A logarithmic function represents the equation of:

$$y = \log_{10}\left(x\right)$$

where

$$x = 10^{y}$$

The exponential function is expressed as:

$$y = \exp_{10}\left(x\right)$$

where

$$\exp_{10}\left(x\right) = 10^{y}$$

These functions are often expressed more generally as $\log_a\left(x\right)$ or a^x where the choice of a affects the scaling and properties of the functions. When a is the natural constant e, the logarithmic function, $\log_e\left(x\right)$ or ln, is referred to as the Napierian logarithm and has special properties:

$$a^x = e^{\log\left(a^x\right)} = e^{x\log\left(a\right)}$$

The base 10 number system computes these functions with the simply remembered values of:

$$\log_{10}\left(1\right) = 0 \qquad \exp_{10}\left(0\right) = 1$$
$$\log_{10}\left(10\right) = 1 \qquad \exp_{10}\left(1\right) = 10$$
$$\log_{10}\left(100\right) = 2 \qquad \exp_{10}\left(2\right) = 100$$

as illustrated in the Figure 1.5.

Figure 1.4 Data scales

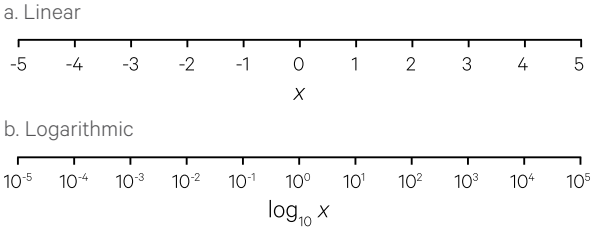
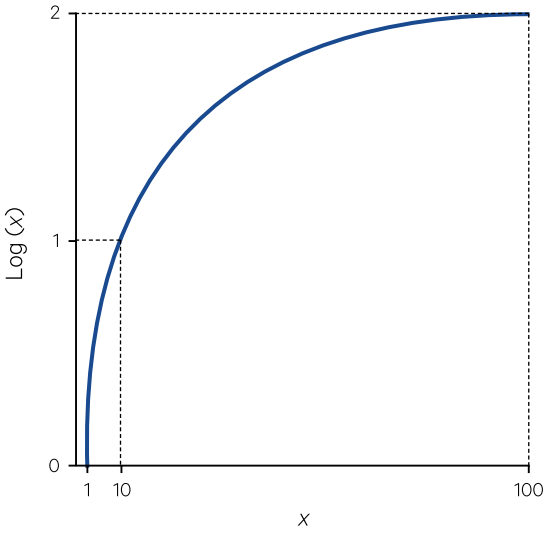
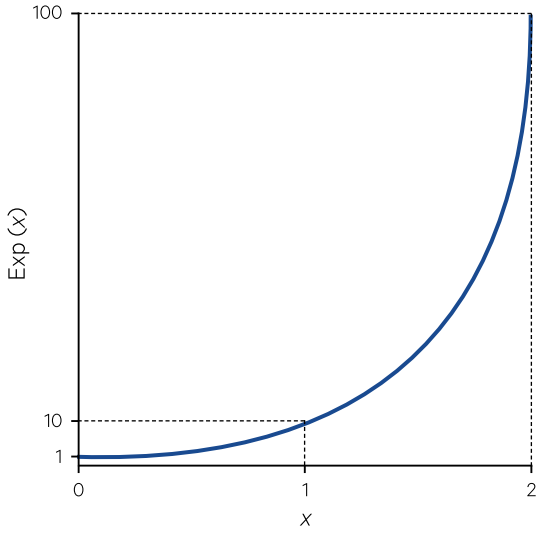


Figure 1.5 Logarithmic and exponential functions

a. Logarithmic function: $y = \log_{10}(x)$



b. Exponential function: $y = \exp_{10}(x)$



Source: Harrison and Jupp (1990) Figure 65

Since the exponential is the inverse function of the logarithm it can be used to convert logarithmically-scaled data to a linear scale. Similarly, the log of an exponent produces linear scaling, that is:

$$x = \log\left(\exp\left(x\right)\right) = \exp\left(\log\left(x\right)\right)$$

1.2.4 Sampling dimensions

The concepts of data resolution, density and extent were introduced in Volume 1A—Section 1 and expanded in Volume 1B—Section 1. These operate in four 'dimensions' of EO data acquisition:

- spectral—image channel values, which typically relate to spectral characteristics of the observed target;
- spatial—target area covered by image pixels;
- radiometric—increments of observed radiance represented by image pixel values; and
- temporal—time period for which an image or set of images is relevant.

The relationships between these sampling dimensions and the resolution, density and extent of EO datasets are summarised in Table 1.1.

Table 1.1 Sampling dimensions in EO imagery

Dimension	Characteristic		
	Resolution	Density	Extent
Spectral	Width of each wavelength channel	Number of channels detected by sensor	Range of wavelengths covered by all channels
Spatial	Ground area imaged per optical pixel	Number of pixels and lines in image	Area covered by image
Radiometric	Smallest change in detected energy that would be represented as a different image brightness level	Number of gradations (grey levels) used to represent full range of radiances that could be detected by sensor	Actual range of radiances detected in each channel
Temporal	Time period over which each image is acquired	Frequency of successive image acquisitions	Total time period for which this imagery is available

Adapted from: Emelyanova *et al.* (2012)

As introduced in Volume 2A—Section 8, a digital EO image can be analysed in terms of statistics derived from each of these four dimensions:

- spectral—techniques that effectively disregard the spatial properties of image data so that an image channel is treated as a one-dimensional array of numbers. These techniques derive statistics that describe the image in terms of the range and variation of data values they contain (see Section 1.3);
- spatial—describe the data values of an image as they relate to its spatial patterns. These statistics can be used to examine the spatial patterns within the image data and quantify the effects of changes in image spatial resolution;
- radiometric—methods to analyse components of image (or ground calibration) values that are not due to surface reflectance, such as atmospheric components, BRDF, terrain-related variations in illumination and scanner inconsistencies; and
- temporal—ways to standardise multi-date imagery to enable meaningful comparison and analysis of multiple images.

Most image transformations rely on spectral statistics but some also use spatial statistics (see Section 6). Radiometric statistics are most commonly derived during image pre-processing to correct for inconsistencies and artefacts of the imaging process (see Volume 2A—Section 3, and Sections 7.2 and 10.2.2 below). Temporal statistics are described in Volume 2D.

1.3 Spectral Statistics

Image spectral statistics summarise various aspects of the data values in each image channel and their correlation between channels (see Volume 2A—Sections 4 and 8 for details). Spectral statistics can be derived for individual channels (see Section 1.3.1) or selected pairs of channels (see Section 1.3.2). Those spectral statistics that are most commonly encountered when transforming EO imagery are summarised below.

1.3.1 Single channel statistics

For each individual image channel, the most commonly encountered spectral statistics are:

1.3.1.1 Measures of data extent

minimum: lowest value;
minimum: highest value; and
range: *minimum-maximum*.

1.3.1.2 Measures of central tendency

mode: most common value;
median: midpoint of range; and
mean: average of all pixel values.

1.3.1.3 Measures of dispersion from the mean

variance: summarises the differences between all pixel values and the mean value:

$$\sigma^2 = \frac{\sum\left(x-\mu\right)^2}{n}$$

where

x is pixel value in channel;
n is total number of pixels in channel (excluding null pixels); and
μ is true mean of the channel data.

and

standard deviation (σ): positive square root of the variance.

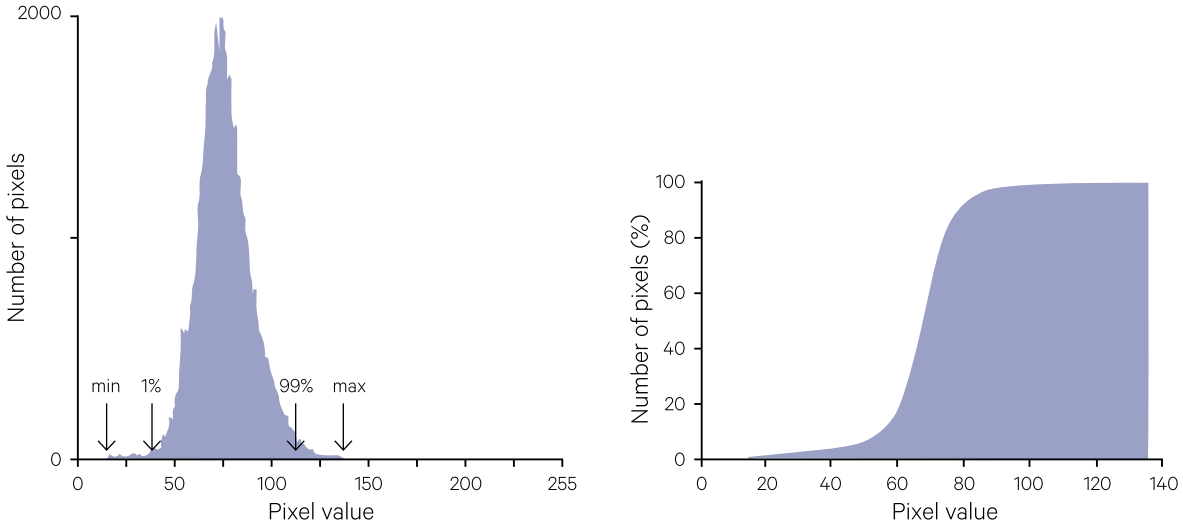
1.3.2 Channel histogram

A channel histogram is a statistical graph that summarises the range and distribution of values in an image channel, either in terms of the actual frequency of each pixel value (see Figure 1.6a) or the cumulative frequency of pixels with values equal to or less than a particular value (see Figure 1.6b). As transformations are applied to images to modify their values, these changes are reflected in the image histogram (Castleman, 1998; Gonzalez and Woods, 2018).

Figure 1.6 Image channel histograms

a. The frequency histogram is formed by plotting the number of pixels at each image value in a channel.

b. The (stylised) cumulative histogram is shown for the image channel in Figure 1.6a. In this example, 1% of pixels have values below or equal to 44, 20% have values below or equal to 66, while 95% of pixels have values below or equal to 89.



Source: Harrison and Jupp (1990) Figures 4 and 5

1.3.3 Multi-channel statistics

The degree of relationship beween the values in two image channels *i* and *j* is represented mathematically by the statistic:

covariance (v_{ij}):

$$v_{ij} = \frac{\sum_{k=1}^{n}\left(x_{ik} - \overline{x}_i\right) \times \left(x_{jk} - \overline{x}_j\right)}{n}$$

where

n is the total number of pixels in the image;
x_{ik} is value for pixel k in channel i ;
$\overline{x}_i$ is the mean of pixel values in channel i ;
x_{jk} is value for pixel k in channel j ; and
$\overline{x}_j$ is the mean of pixel values in channel j.

or its normalised form:

correlation (c_{ij}):

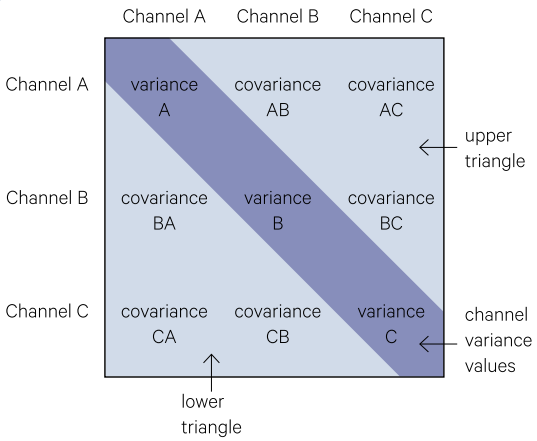
$$c_{ij} = \frac{v_{ij}}{\sqrt{v_{ii} \times v_{jj}}}$$

where

c_{ij} is the correlation between channels i and j;
v_{ij} is the covariance between channels i and j;
v_{ii} is the variance of channel i; and
v_{jj} is the variance of channel j.

Covariance or correlation values are typically summarised in a table for all channels in the image. This table is called a covariance or correlation matrix. The standard format of a covariance matrix is shown in Figure 1.7. Such matrices are used in a number of linear transformations as detailed in Sections 7, 8 and 9.

Figure 1.7 Format of variance/covariance matrix



1.3.4 Channel crossplots

The relationship between values in two image channels can be shown in a crossplot. For example, using the example image channels in Figure 1.8, crossplots can be constructed for each pair of channels as shown in Figure 1.9. On each crossplot, the process is shown for the values of pixel 1, line 5 in the image. The distribution of the points on the crossplots visualises the similarity, or correlation, between the two plotted channels and can be used to plot a regression line (see Figure 1.9b).

Figure 1.8 Example image channels

Channel A

	1	2	3	3	4
	3	3	5	6	3
	2	3	3	6	5
	5	4	2	2	1
pixel 1, line 5 →	7	6	3	2	1

Channel B

	3	6	7	8	7
	2	4	7	5	6
	1	1	3	4	5
	1	1	2	5	5
pixel 1, line 5 →	2	3	4	6	7

Channel C

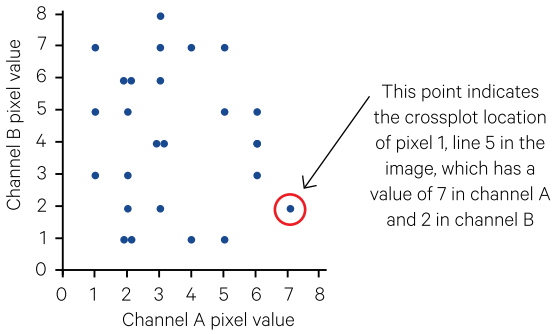
	1	3	4	3	3
	2	3	4	5	4
	2	2	3	6	5
	4	4	2	1	2
pixel 1, line 5 →	6	5	2	1	1

Source: Harrison and Jupp (1990) Figures 41

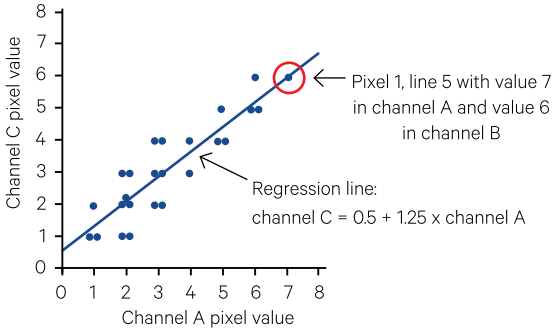
Figure 1.9 Example crossplots

The image channels in Figure 1.8 are crossplotted to show the extent of correlation between channels.

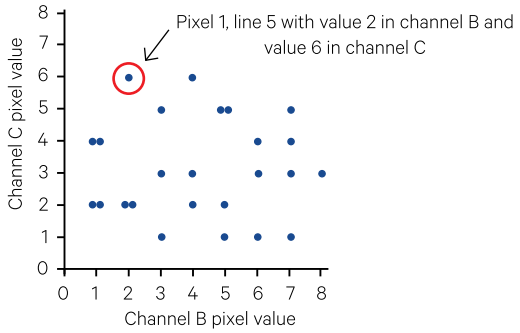
a. Channel B versus Channel A



b. Channel C versus Channel A



c. Channel C versus Channel B



Source: Harrison and Jupp (1990) Figures 42

1.4 Processing Operations

Image processing methods are often classified as either linear or non-linear operations. Linear operations have the important property of additivity (that is, the result of a linear operation on two separate inputs is the same as that operation applied to the sum of those two inputs) whereas non-linear operations do not have this property (Gonzalez and Woods, 2018). The distinction between linear and non-linear processes is particularly relevant to the sequencing of operations in a multi-stage processing methodology (see Volume 2D).

Linear operations involve addition and subtraction. In the context of image processing, this applies to adding and subtracting pixel values from different channels, often using fractions or multiples of different channel values. In mathematical terms, the simplest way to represent linear combinations of input values is to use matrices (see Section 7 and Volume 2X—Appendix 6).

As the name implies, non-linear operations are then defined as those that are not linear. These operations include multiplication, division and the selection of order-dependent statistics such as the minimum, maximum, median, or mode.

Transformations may be applied to either individual channels or multiple channels, based on the input values in one or more channels. Single channel transformations tend to involve either:

- rescaling of the spectral value range according to a predefined relationship as introduced in Volume 2A—Section 4 and further detailed in Section 2 below; or
- modifying each pixel value by considering the values of its neighbouring pixels, usually involving the use of filters. The mathematical foundations of image filtering is introduced in Section 3 below and detailed in terms of specific uses in Sections 4, 5 and 6.

Multiple channel transformations can be broadly grouped into two categories:

- combining input channels using addition and subtraction (linear operations) as introduced in Section 7 and further discussed in terms of applications in Sections 8 and 9; or
- dividing one image channel by another (non-linear operations; see Sections 10 and 11).

While some image processing systems do not differentiate between linear and non-linear operations (and simply refer to them by a generic title such as 'band maths'), most transformations are based on some combination of the processes listed above. Processes that analyse larger datasets formed from multiple EO images are specifically considered in Volume 2D.

1.5 Further Information

Image Processing

SEOS: eLearning Tutorials: *https://www.seos-project.eu*

LEOWorks—Image Processing Tutorials: *http://leoworks.terrasigna.com/tutorials*

Image Processing Systems

ERDAS Imagine: *http://www.hexagongeospatial.com*

ERMapper: *http://www.hexagongeospatial.com*

ENVI: *http://www.harrisgeospatial.com*

TNTmips: *http://www.microimages.com*

1.6 References

Anderberg, M.A. (1973). *Cluster Analysis for Applications*. Academic Press, New York.

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Emelyanova, I.V., McVicar, T.R., Van Niel, T.G., Li, L.T., and van Dijk, A.I.J.M (2012) *On blending Landsat-MODIS surface reflectances in two landscapes with contrasting spectral, spatial and temporal dynamics*. WIRADA Project 3.4: Technical Report. CSIRO Water for a Healthy Country Flagship, Australia. 72pp.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

2 Rescaling

A wide range of rescaling operations may be applied to image data for a variety of purposes. Some commonly encountered rescaling options that are available for displaying EO imagery are described in Volume 2A—Section 4. Rescaled image values can also be written to an image file to:

- match data values between different channels in an image;
- equalise values in imagery from different sources and thus allow them to be subsequently processed using similar parameters; or
- model the relationship between ground measurements and EO radiance (see Volume 2D).

All rescaling methods change the shape of the image histogram (see Section 1.3.2). Excursus 2.1 compares the impacts of some common rescaling methods when applied to an EO image channel and its histogram. This section describes the following single channel rescaling operations in terms of process and application:

- linear (see Section 2.1);
- non-linear (see Section 2.2);
- lookup tables (see Section 2.3);
- destriping (see Section 2.4); and
- histogram manipulation (see Section 2.5).

Excursus 2.1—Rescaling results

Source: Tony Sparks, Icon Water

This example uses an image acquired by Landsat-5 TM on 20 January 2010 over southeastern NSW. The town of Cooma is near the centre of this image, with Canberra in the northwest and the coastline visible in the east. Band 1 (green; see Figure 1.2) is used to differentiate spectral features on mountainous terrain, including forest cover, cleared pasture, water bodies, and cloud.

The selected image channel is shown below with six different stretches:

- none—shows the original data range for the example image channel. This presentation does not use all available grey levels for display and results in low contrast over most of the image (see Figure 2.1a);
- linear—spreads the full data range evenly over the full extent of available grey levels for display (see Figure 2.1b);
- logarithmic—this non-linear stretch redistributes the channel data range as a logarithmic function, which increases contrast in low data values (see Figure 2.1c);
- normalised—matches the image histogram to a normalised distribution (see Figure 2.1d);
- histogram equalised—this popular stretch aims at equalising the number of image pixels associated with each data value, and thus increases overall contrast (see Figure 2.1e); and
- user-defined—this non-statistical stretch is specified by a lookup table to achieve the desired spread of grey levels in the image (see Figure 2.1f).

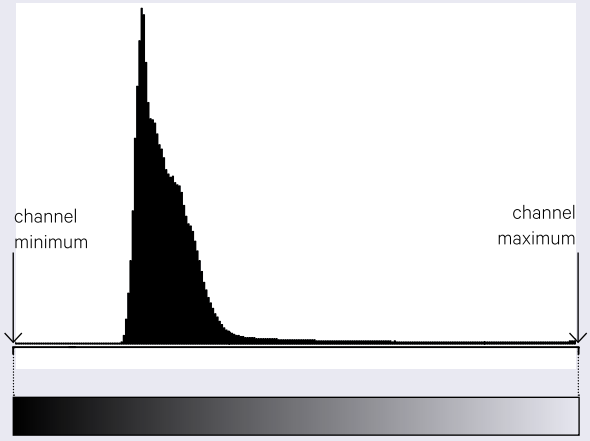
Each stretch presents a different balance of light and dark tones in the image channel by changing the shape of the image histogram. Selection of the most appropriate stretch will depend on the desired end result.

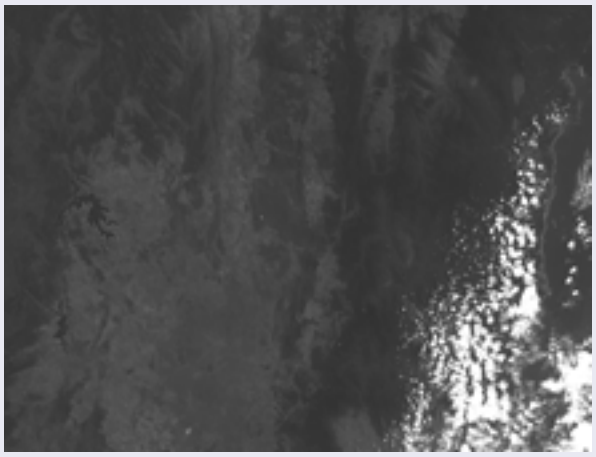
Background image: This Landsat-5 TM image was acquired on 6 October 2011 over the arid landscape of Munga-Thirri National Park in western Queensland. The image is displayed using bands 6, 4, 1 as RGB, contrasting active fires (red) and fire scars (dark red) with temporarily-filled waterways and parallel sand dunes. Note that the red linear streaks, oriented roughly east-west near the large active fire body, are image artefacts resulting from sensor oversaturation.
Source: Craig Shephard, DSITI

Figure 2.1 Examples of rescaling

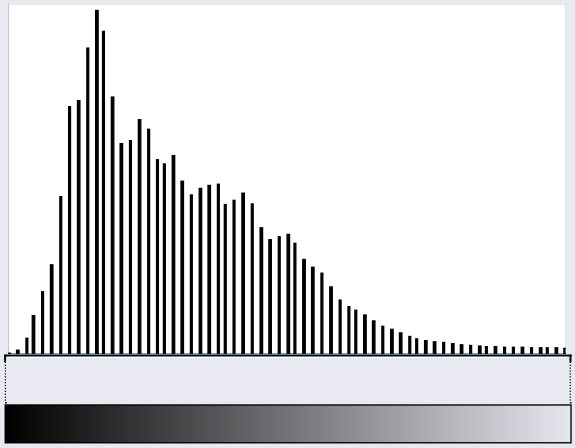
The example image channel is Band 1 (blue) from a Landsat-5 TM image acquired on 20 January 2010 over southeastern NSW, and features mountainous terrain with forest cover, cleared pasture, water bodies and cloud.

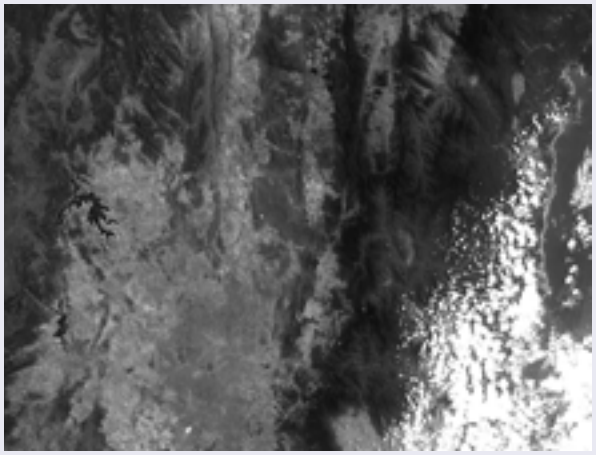
a. Unstretched image



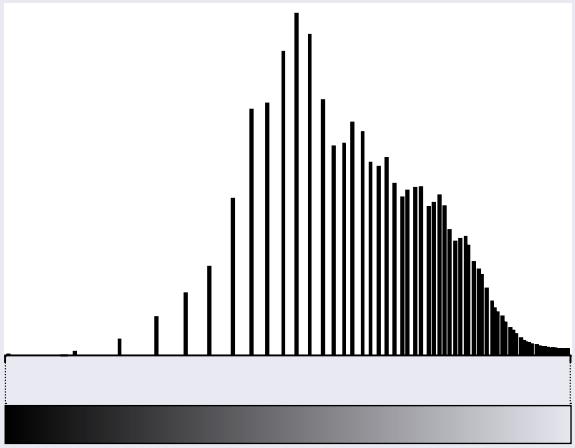


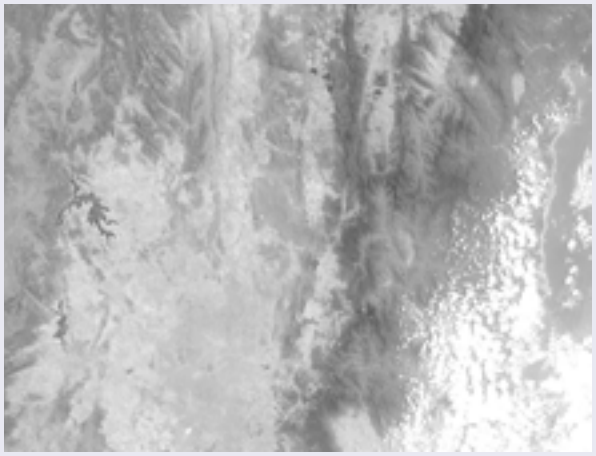
b. Linear stretch



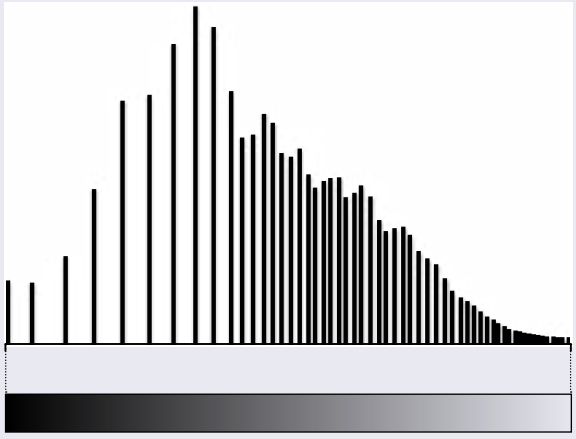


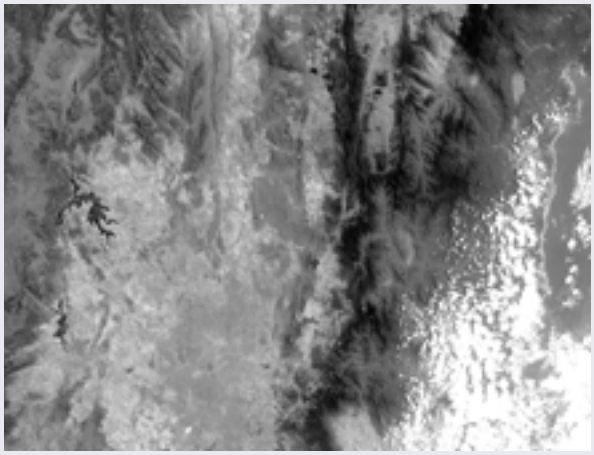
c. Logarithmic stretch



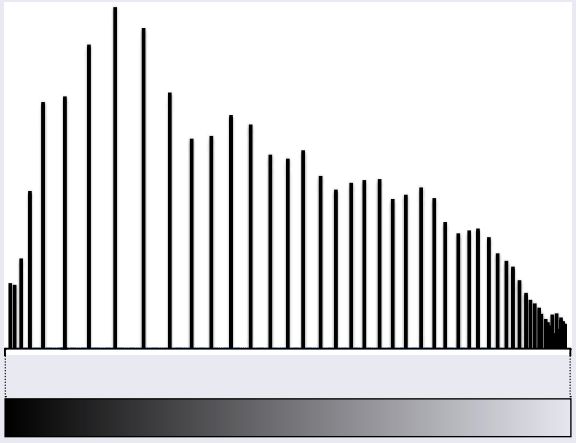


d. Normalised stretch

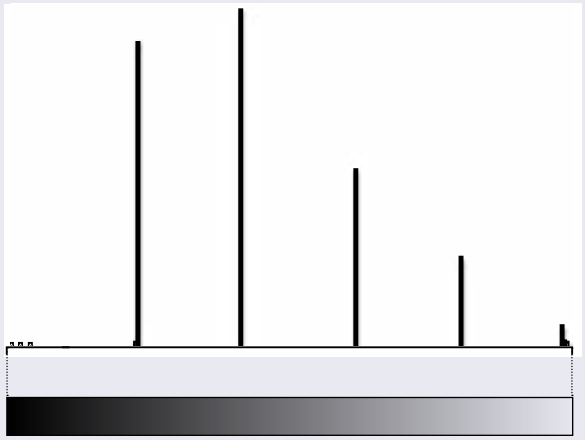


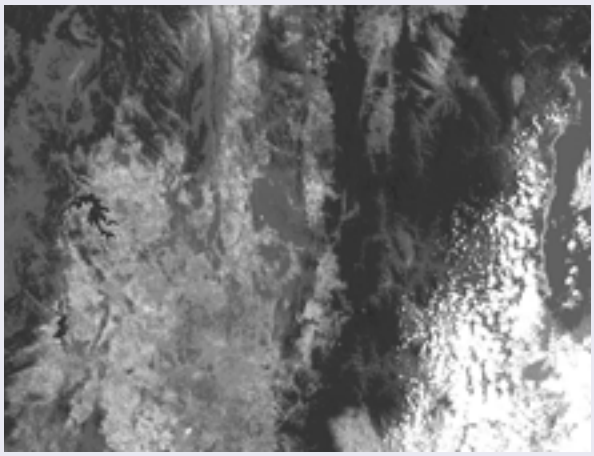


e. Histogram equalisation



f. User-defined stretch based on lookup table (LUT)





2.1 Linear Rescaling

Linear rescaling evenly matches the image data range to an output range, such as the available contrast in an image display device, while non-linear rescaling will selectively increase separation between low or high values in the image (see Volume 2A—Section 4.2). Linear rescaling is also useful for matching the contrast between images from different sources. Balancing of the data values between multi-source imagery, such as multiple images of a given ground location that were acquired on different overpass dates or by different sensors, is required before mosaicking or differencing such image data (see Volume 2D). Having the same minimum and maximum values in all channels of a single image is convenient for some operations, but such rescaling should be used with caution.

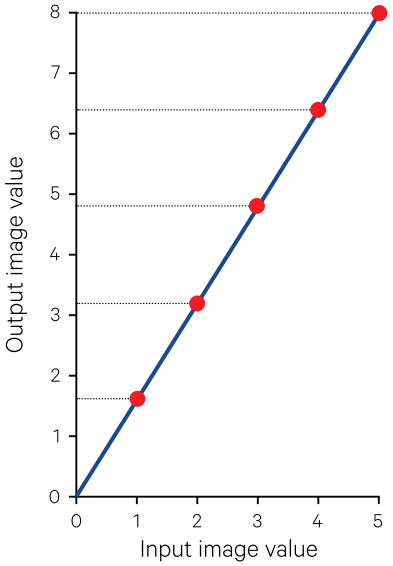
As discussed in Volume 2A—Sections 4 and 8, rescaling can produce discontinuities in the image data range (see Figure 2.1b). Given the discrete nature of image values, this discontinuity pattern can be exaggerated by rounding 'errors' during the rescaling process, which lead to a false representation of the differences between values in the output image. For example, in the rescaling illustrated in Figure 2.2, the input values 0, 1 and 2 become output values 0, 2 and 3 so that in the rescaled image the difference between the original values 0 and 1 is exaggerated relative to the original values 1 and 2. These potential problems are worth considering when selecting rescaling ranges, although in most cases the large potential range of data values in EO images means these problems do not seriously degrade image data. Similar discontinuities can occur when image data is converted to surface reflectance measurements.

The rescaling range is generally defined in terms of the minimum and maximum values in the input image, which are to be respectively mapped to the minimum and maximum values in the output image data range. When the selected rescaling range is within the absolute minimum and maximum values of the image, any differences in values greater than the rescaling maximum or less than the rescaling minimum are effectively lost in the output image and cannot be derived in any subsequent processing (unless the output data type is changed to avoid this problem).

Inversion of the data range in an image channel (that is reversing the relative order of pixel values to increase rather than decrease between the channel maximum and minimum values) can be implemented as a linear operation as described in Section 7.

Figure 2.2 Effect of rescaling operation on image data range

Original image values 0–5 are rescaled to the range 0–8. The rescaling falsely represents differences between values in the original image and the values 1, 4 and 7 do not occur in the rescaled image.



Input Value	Output Value
0	0
1	2
2	3
3	5
4	6
5	8

2.2 Non-linear Rescaling

Simple non-linear rescaling of image data values using logarithmic and exponential functions are introduced in Volume 2A—Section 4 and Section 1.2.3 above. These functions are often used to model the relationship between image radiance values and quantifiable ground parameters such as water depth (see Volume 3B). Transformations based on these functions then allow image values to be converted to or from a scaling which can be linearly related to ground measurements.

Some properties of logarithms make this transformation particularly useful as a pre-processing step for other transformations. For example, the product of two values in a linear data scale may be computed as the sum of their logarithmic values, that is:

$$x \times y = \exp\left(\log\left(x\right) + \log\left(y\right)\right)$$

Thus, filtering of log-transformed images effectively multiplies pixel neighbourhood values and can produce some useful enhancement effects (see Section 3.6). Similarly, the quotient of two linear values can be determined from the difference of their logarithmic values, that is:

$$\frac{x}{y} = \exp\left(\log\left(x\right) - \log\left(y\right)\right)$$

Another calculation that can be simply implemented in a logarithmic scale is exponentiation, since:

$$x^n = \exp\left(n \times \log\left(x\right)\right)$$

These relationships allow computation of complicated, non-linear combinations of channels such as:

$$\frac{ch1 \times ch2}{ch3^3 \times ch4}$$

If the logarithms of the four channels are computed, this combination can be produced from the expression:

$$\exp\left(\left(\log\left(ch1\right) + \log\left(ch2\right)\right) - \left(3 \times \log\left(ch3\right) + \log\left(ch4\right)\right)\right)$$

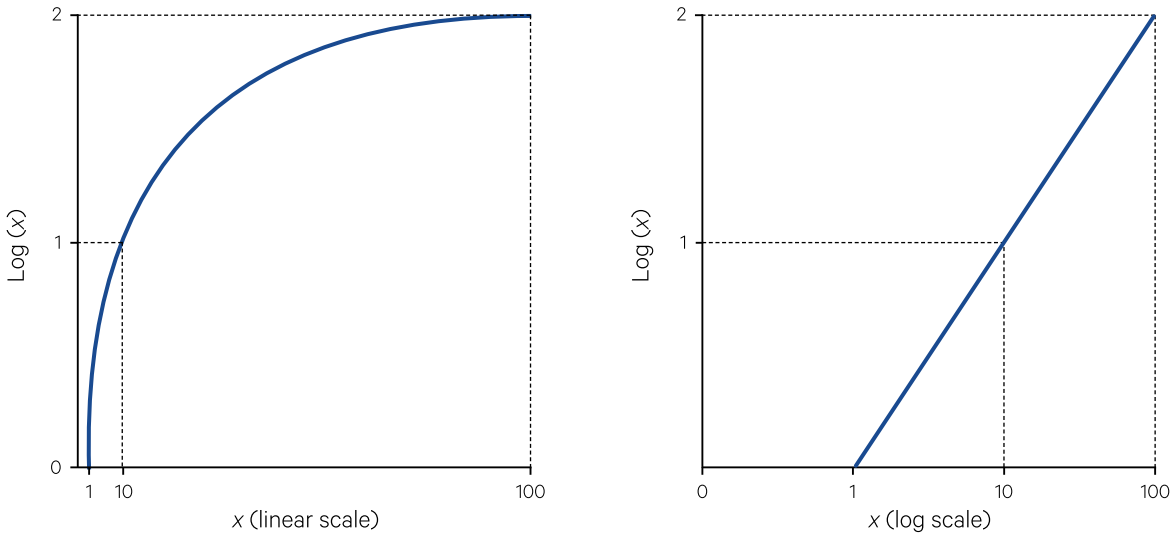
This computation can be easily implemented in an affine transformation as detailed in Section 7.1.

Extensive datasets are often plotted on a logarithmic scale. This scaling is effective when the full data range is large but only has significant variation near one end of the range. Logarithmic scaling does, however, affect the apparent relationship between the variables being plotted. For example, if a logarithmic function (see Figure 2.3a) is plotted using a logarithmic scale for the *x*-axis, the graph becomes linear as shown in Figure 2.3b.

Figure 2.3 Logarithmic scaling

a. The function $y = \log_{10}(x)$ is plotted using a linear scale.

b. The function $y = \log_{10}(x)$ is plotted here using a logarithmic scale for the *x* axis.



Source: Harrison and Jupp (1990) Figures 65a and 66

The logarithmic transformation has particular relevance to models relating EO image radiance to measurable ground features such as crop yield (see Volume 3A). In image processing systems, logarithmic and exponential transformations can be implemented by computing the functions:

$$\log\left(x + a\right)$$

and

$$\exp\left(\frac{x}{a}\right)$$

where

x is an image pixel value; and
a is the scaling factor.

As indicated in the graph of Figure 2.3a, the logarithm function for value 0 is undefined (said to be negative infinity). The scaling factor '*a*' allows an offset to be added to, or subtracted from, the image values during processing, that is:

$$a = \frac{\beta}{\alpha}$$

where

$$\log\left(\alpha \times x + \beta\right) = \log\left(\alpha\right) + \log\left(x + \frac{\beta}{\alpha}\right)$$

Since the $\log\left(\alpha\right)$ term is a constant offset, it can be ignored when rescaling image values. However, the term would become important for model development if image values were to be related to physical parameter measurements (see Volumes 2D and 3). When log-ratio images are being computed, dark values for the ratio can be specified as negative scaling factors during the logarithmic transformation (see Section 10.1.2).

Similarly a suitable divisor, *a*, for the exponential function avoids extremely large output values which would be clipped during the rescaling process:

$$\exp\left(\alpha \times x + \beta\right) = \exp\left(\alpha \times x\right) \times \exp\left(\beta\right)$$

In this case, the offset can again be ignored in image rescaling and:

$$a = \frac{1}{\alpha}$$

In addition to the scaling factor, a data range is specified for each image channel during rescaling so that all other values are clipped to this range and values within the range are adjusted to increment from a minimum value of zero. This sub-range has relevance when the log channels are to be used for other transformations, such as to form channel ratios (see Section 10.1.3). Generally the minimum range value should be one less than the absolute minimum of a channel. This value is scaled to the minimum value in the image data range of the output channel. Similarly, the maximum value specified in the processing range is scaled to maximum image data range value in the transformed channel.

The actual transformed minimum and maximum values are often reported during processing and are useful for subsequent rescaling operations and model fitting. The resulting image channels contain data ranges that have been redistributed using logarithmic and/or exponential functions as defined. However, in byte images, the pixel values are usually integer representations of these function values, which are subsequently rescaled to fill the maximum image data range. The final rescaling operation needs to be taken into account when using image pixels values in log (radiance) models or when computing the inverse transformation.

Nothing happens until the pain of remaining the same outweighs the pain of change.
(Arthur Burt)

2.3 Lookup Tables (LUT)

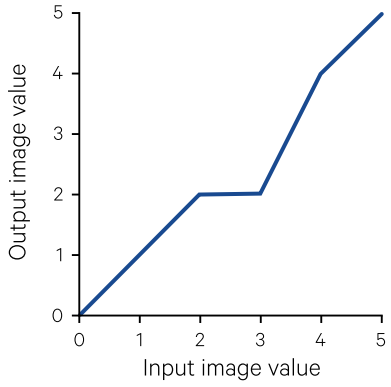
Image data values may be redefined using a Lookup Table (LUT). The use of LUT for image display enhancements is discussed in Volume 2A—Sections 4 and 8. The use of LUT for density slicing, or for aggregating basic feature classes into labelled categories in the mosaic model for image classification, is also discussed in Volume 2E.

In the context of image rescaling, an LUT performs a similar function by defining an output value for each possible input value in the image. An LUT is simply a single column table—the rows of the table represent input data values while the entries in the table indicate the output values to assign to each input value. The definitions may represent a wide range of contrast enhancements such as piecewise linear or sawtooth ('overlapping contrast enhancement') stretches as illustrated in Figure 2.4.

Other more sophisticated rescaling transformations, such as histogram equalisation (see Section 2.5), may be represented as an LUT (although this is more commonly done as an 'on the fly' statistical transformation). Indeed, for rapid computation, many other image processing functions such as non-linear functions are implemented using LUT in image processing programs. LUT are also a convenient means for representing hierarchies and other many-to-one mappings (see Volume 2E).

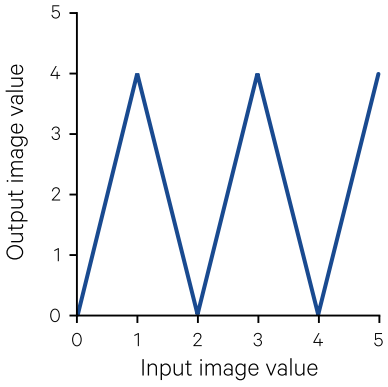
Figure 2.4 LUT operations

a. Piecewise linear rescaling



LUT entry	Output Value
0	0
1	1
2	2
3	2
4	4
5	5

b. Sawtooth stretch



LUT entry	Output Value
0	0
1	4
2	0
3	4
4	0
5	4

Source: Harrison and Jupp (1990) Figure 10

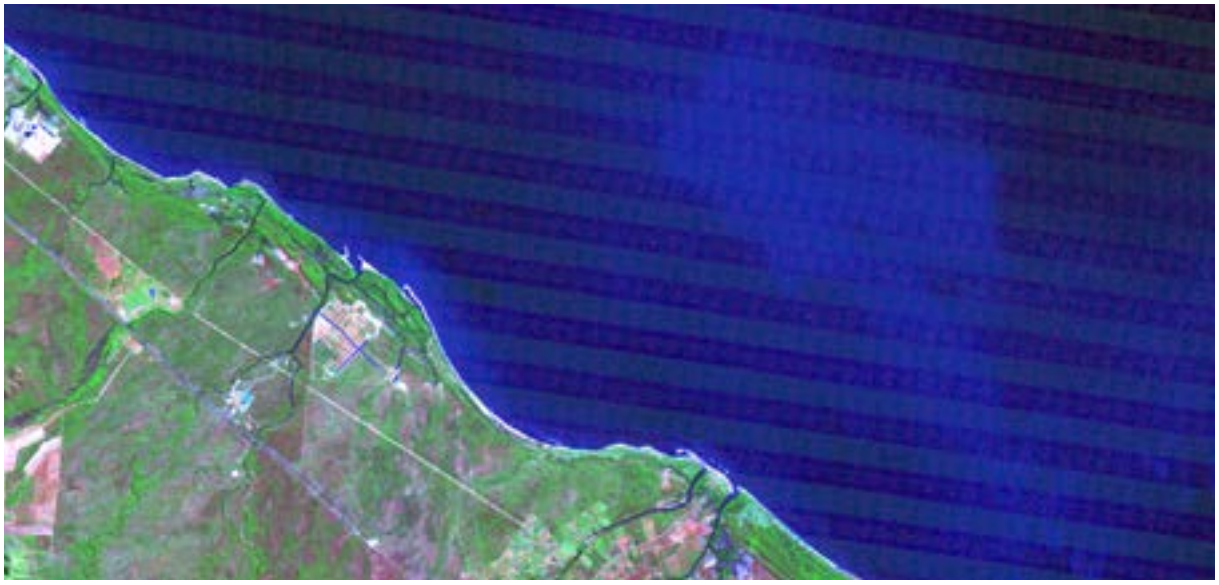
2.4 Destriping

EO imagery may contain spatial striping patterns due to the miscalibration between multiple detectors in the scanning system (see Volumes 1 and 2A). In scanners that record adjacent image lines using different detectors, a line striping is frequently observed in which the periodicity of the stripes matches the number of detectors in the instrument. Landsat TM, which senses 16 lines per scan, records during both forward and back scans across the swath so its imagery can have 16 or 32 line striping. When such data are geometrically corrected, the striping pattern can be observed to have 17 or 34 line periodicity. Additionally, in linear array or 'pushbroom' scanners, which operate banks of detectors across an image line, different pixel values in the line will be determined by different detectors. Imagery produced this way can show vertical striping, as was visible in some early SPOT data.

In an attempt to correct for miscalibration effects, various pre-processing operations are performed by ground receiving stations before image data are distributed. However, such variations are data-dependent so are difficult to remove effectively using whole scene statistics or instrument calibration readings. Consequently, striping patterns commonly occur in image data from EO sources and usually need to be reduced in some way before applying other image processing operations for enhancement, interpretation or classification. It must be emphasised that destriping algorithms can only be applied to imagery that has not been resampled since the resampling process generally rotates image lines, which invariably rotates the striping artefact (see Figure 2.5).

Figure 2.5 Rotated image striping pattern

This Landsat-5 TM image demonstrates a visible striping pattern in the water. As detailed in Volume 2A—Section 3, this striping results from miscalibration of the sensor's 16 detectors, which are designed to image a 16-line swath. After image resampling, however, the striping pattern has been rotated so would be difficult to remove. This image was acquired on 31 July 2013 over the northern fringe of Townsville and is displayed using bands 5, 4, 2 as RGB.



Source: Tony Sparks, Icon Water

Striping patterns are most evident in areas with relatively uniform image values, such as deep water, with different land covers showing different striping effects. In the scanning device, individual detectors are calibrated relative to a standard lamp, which radiates a standard range of wavelengths (see Volume 2A—Section 3). For this calibration target, all detectors may record, or can be 'calibrated' to record, the same radiance; however differences in response functions between the detectors can still cause miscalibration for other wavelengths. Thus, the striping pattern varies with the land cover being imaged. In destriping then, these patterns are best identified if the image is segmented into broad land cover regions (such as water, vegetation, bare soil and cloud) before applying destriping methods (see Volume 2A—Section 10 regarding image segmentation methods).

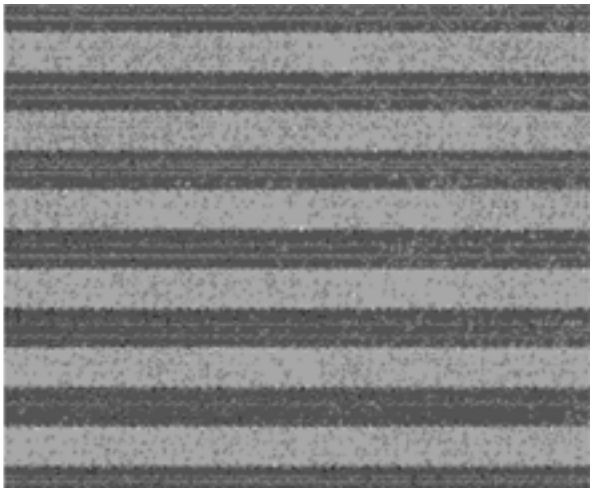
A number of destriping algorithms have been developed to try to minimise the effects of detector imbalance. These algorithms basically accumulate statistics relating to each detector. For example, for an image with six-detector striping, one set of statistics are gathered separately for lines 1, 7, 13, 19, etc., then another for lines 2, 8, 14, 20, etc., and so on. The destriping process then attempts to remove the differences by producing imagery in which selected statistics match between all detectors. A simple example of destriping based on this approach is shown in Figure 2.6. While some banding is still visible in the destriped image, partly due to the limited dynamic range of the original data (8 bit), the processed image has significantly reduced the striping pattern.

Destriping algorithms typically accumulate mean and variance statistics for a user-defined periodicity in each channel in a (stack of) image(s). Divergence and centrality values are generally reported, which indicate the extent to which each detector is similar to the other detectors in each channel (with low values indicating a close match between detectors). This algorithm is applied to the image(s) by matching all detectors to the mean detector values or to a selected reference detector. The detector reporting the smallest sum of centralities for each channel is recommended as the reference detector. This process should be applied to imagery before resampling as the original striping pattern is likely to be modified to some form of diagonal artefact in the geo-corrected image, which would be more difficult to correct (see Figure 2.5).

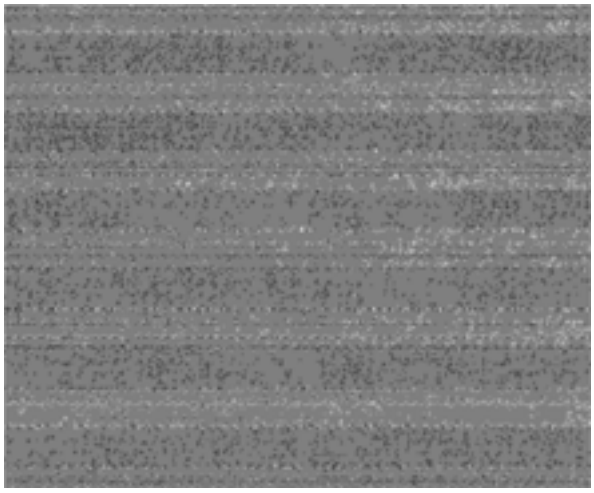
Figure 2.6 Destriping example

In this example, an ocean portion of band 4 (NIR) in an unresampled Landsat-5 TM image has been extracted and destriped (the resampled version of the full image is shown in Figure 2.5). While some banding is still visible in the destriped image, partly due to the limited dynamic range of the original data (8 bit), the processed image has significantly reduced the striping pattern.

a. Original image band



b. Destriped image band



Source: Tony Sparks, Icon Water

These detector statistics are used to compute a correction equation for each detector and channel combination. This equation can be expressed as:

$$x_{ij} + a_{ij} + b_{ij} \times x_{ij}$$

where

x_{ij} is pixel i in channel j; and
a_{ij} and b_{ij} are adjustment factors computed from the detector statistics.

Some destriping algorithms allow a theme (a range of channel values; see Volume 2A—Section 9.1.1) to be defined for the destriping operation, so that the process is only applied to pixels that satisfy the theme. This option effectively allows the image to be spectrally segmented into different land cover categories during destriping. It is highly recommended that appropriate theme values be used in conjunction with image destriping to account for data-dependent patterns.

The effects of destriping may be difficult to discern visually in some imagery but can be quantified by computing the difference between the original and destriped images (see Volume 2D). In some cases, a narrow data range may make the destriping operation very difficult since the striping effectively occurs about a single grey level value. Linear rescaling of the data channel before destriping may improve the performance of the destriping algorithm.

If a significant striping pattern cannot be removed adequately by using a destriping algorithm, image filtering may be used to reduce the visual impact of the pattern. This process is described in Section 4. In this context, average filtering using filter weights for a gentle smoothing filter, with a filter size that is at least as wide as the detector striping, would be most appropriate.

Very noisy imagery should be despiked before destriping (see Section 4.5). For example, isolated noise spikes (pixels with unrealistically high values) are commonly observed in imagery of deep water due to sensor saturation from sunglint off the water surface. The despiking process replaces these values with more reasonable ones and thus avoids biasing any statistics generated for the destriping algorithm.

2.5 Histogram Manipulation

As introduced in Section 1.3.2 and illustrated in Excursus 2.1, the 'appearance' of an image determines the shape and distribution of its histogram. Accordingly, transformations that change image values almost invariably change its histogram. Just as manipulating image values changes the image histogram, manipulating the image histogram changes the image appearance (Gondalez and Woods, 2018).

Commonly encountered approaches to histogram manipulation were introduced in Volume 2A:

- histogram normalisation (see Volume 2A—Section 4.2.3.1 and Figure 2.1d); and
- histogram equalisation (see Volume 2A—Section 4.2.3.2 and Figure 2.1e).

Some applications, however, require image histograms to be matched to a tailored shape (Gonzalez and Woods, 2018). Such transformations may be appropriate to compensate for 'over exposure' of image noise resulting from histogram equalisation, typically in low image values. In practice, they are often computed via the histogram equalisation process.

Most histogram manipulation methods effectively redistribute the histogram 'bins' along the intensity axis (see Figure 2.1). Accordingly, these methods only ever approximate the shape of the target histogram. In order to precisely match a specified histogram shape, the original image values need to be modified and redistributed. This process may be relevant to EO imagery to establish a reference image for sensor calibration. One approach to exact histogram matching is based on three stages (Coltuc *et al.*, 2006):

- order pixels in image by some predefined criteria;
- divide pixels into a set number of groups so that all groups are the same size; and
- allocate an appropriate intensity value to each group.

This method can still falter on images with large homogeneous features, in which case additional processing with image masks may be appropriate to enable detail in particular portions of the image to be enhanced. Further details on histogram matching methods are provided by Gonzalez and Woods (2018).

2.6 Further Information

Jensen (2016) Section 8

Gonzalez and Woods (2018)

2.7 References

Coltuc, D., Bolon, P., and Chassery, J.-M. (2006). Exact Histogram Specification. *IEEE Transactions on Image Processing,* 15(5), 1143–1152.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Single Channel Filtering

Filtering operations are generally applied to individual image channels to modify pixel values based on the values of each pixel neighbourhood. The following sections consider spatial filtering in terms of:

- filter operation, including weights, sizes, thresholds and variations (see Section 3);
- low pass or smoothing filters to blend pixel values (see Section 4);
- differential and high pass filters to detect and enhance edges respectively (see Section 5); and
- filters that highlight variations in a surface, including texture, relief shading, directional, and curvature (see Section 6).

Contents

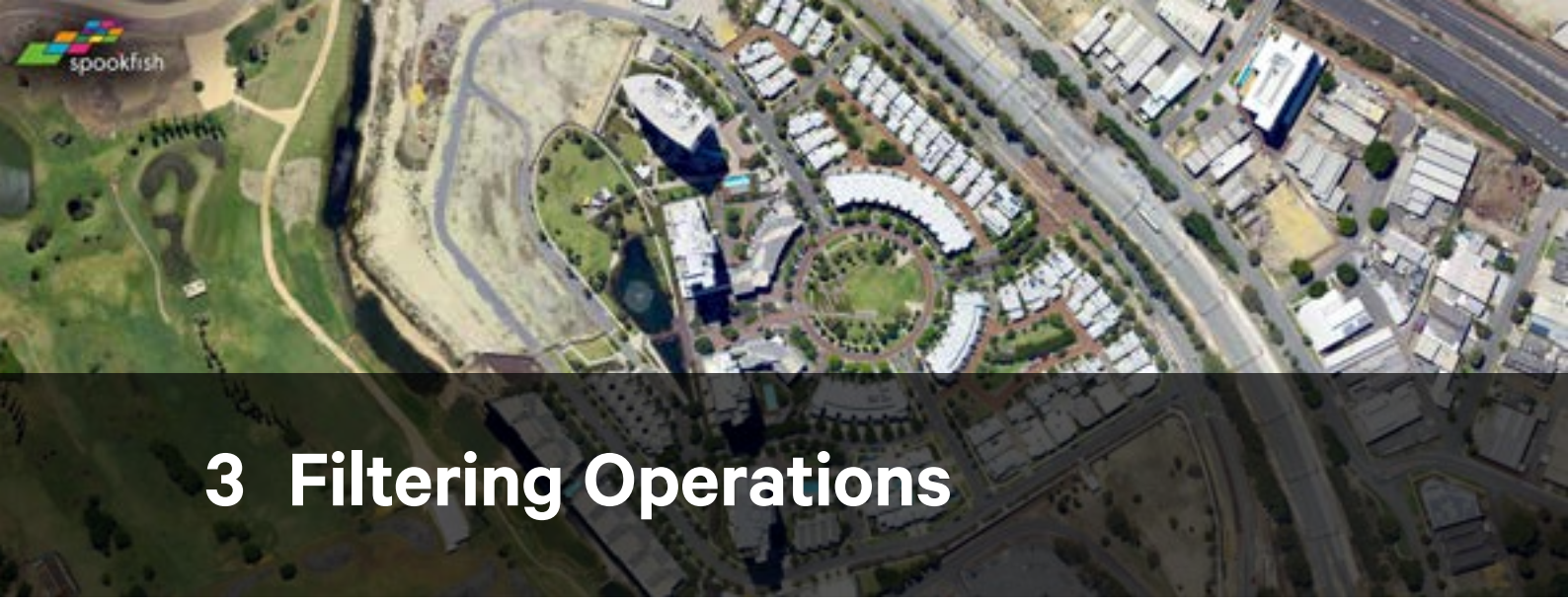
3 Filtering Operations 25

4 Smoothing (Low Pass) 33

5 Highlighting Edges 47

6 Highlighting Surface Variation 59

Background image on previous page: Panchromatic image acquired by Landsat-8 on 14 April 2014 over Norman River in northwest Queensland, which feeds into the Gulf of Carpentaria. **Source:** Craig Shephard, DSITI



3 Filtering Operations

Image data can be represented in terms of the spatial arrangement of pixels (the spatial domain) or in terms of the frequencies of their spatial patterns (the frequency domain). The frequency domain represents image spatial patterns as sinusoidal waves of differing frequency, amplitude and direction. In EO imagery, large spatial patterns (such as regional landscape patches) correspond to low frequency waves in the frequency domain while noise or feature edges are represented as high frequencies.

When image data are represented in the frequency domain, specific frequencies can be modified using frequency filtering techniques such as Fourier analysis. Fuller (1966), Castleman (1998) and Gonzalez and Woods (2018) discuss methods for designing filters to suppress, enhance or isolate specific ranges of frequencies. Discrete Fourier analysis allows image data to be transformed from the spatial to the frequency domain where they can be filtered in terms of specific frequencies then transformed back to the spatial domain using the inverse Fourier transform function. Fourier analysis is appropriate where data patterns are (semi-)periodic, such as in signal processing applications, but is less effective with data containing local variations, such as EO imagery.

Filtering can also be applied directly to spatial data by 'convolution'—a mathematical operation in which the value of an image pixel is weighted by the values of the surrounding pixels (see Section 3.1). Richards (2013), Castleman (1998) and Gonzalez and Woods (2018) discuss the relationship between spatial filter convolution and frequency filtering. Filters defined in the frequency domain can be represented as spatial filters and applied more efficiently (but in some cases less precisely) using convolution in the spatial domain. Conversely, simple and direct spatial filters are difficult to implement in the frequency domain. Spatial filtering is more applicable to EO data that exhibit local rather than periodic variations.

Filtering for EO imagery can involve:

- variable filter weighting values (see Section 3.2);
- different filter sizes, generally ranging from 3×3 to 11×11 (see Section 3.3);
- thresholds (defining the minimum difference between the input and filtered values of each pixel) which need to be reached before the input value is modified (see Section 3.4);
- adaptive or edge-preserving filters (see Section 3.5); and
- homomorphic filtering (see Section 3.6).

Most image processing systems offer a wide range of filter-based transformations. These can range from generic filtering to special purpose filters for particular data types, such as the computation of Sun shadowing or curvature from elevation data. The uses of specific filter-based transformations are detailed in Sections 4 to 6 below.

In the Information Age,
the first step to sanity is filtering.
Filter the information: extract for knowledge.
(Marc Stiegler)

Background image: High resolution aerial image of apartment development in South Perth, WA, acquired on 11 January 2018. **Source:** © EagleView

3.1 Operation

Spatial filter transformations described in this Volume use those image values that are covered by a filter centred on a pixel to change the values of that pixel. The simplest form of a filter is a (square) grid of pixels, which usually has an odd number of pixels along each side to ensure symmetry about the central pixel. In this context, the filter is also commonly referred to as a moving window or box car filter, or a template or a kernel.

In a discrete dataset such as an image, the convolution process at an individual pixel (*x,y*) is defined as:

$$\sum_{i=-m}^{m}\sum_{j=-n}^{n} f(x+i,y+j)\,g(i,j)$$

where

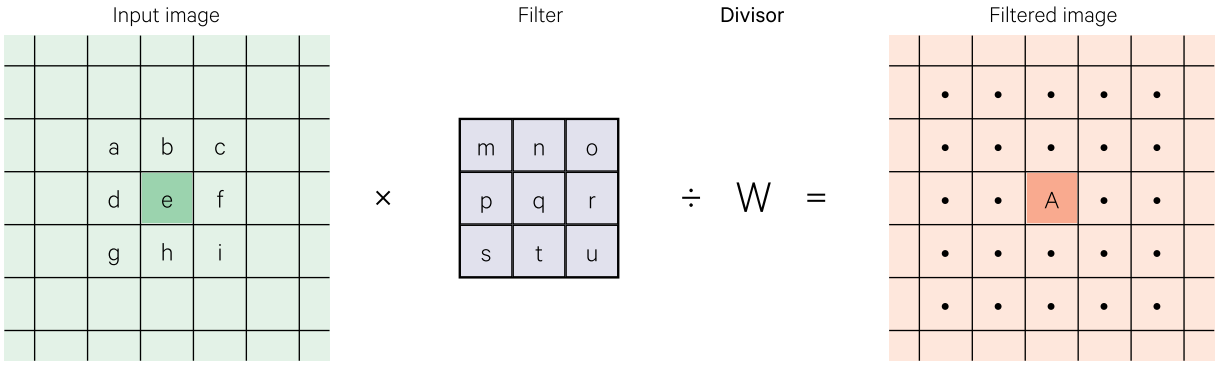
$2m+1$ is the number of pixels across the filter;
$2n+1$ is the number of lines down the filter;
f is the image;
g is the filter;
x is the pixel position in the image; and
y is the line position in the image.

Weighting values may also be associated with each cell of the filter.

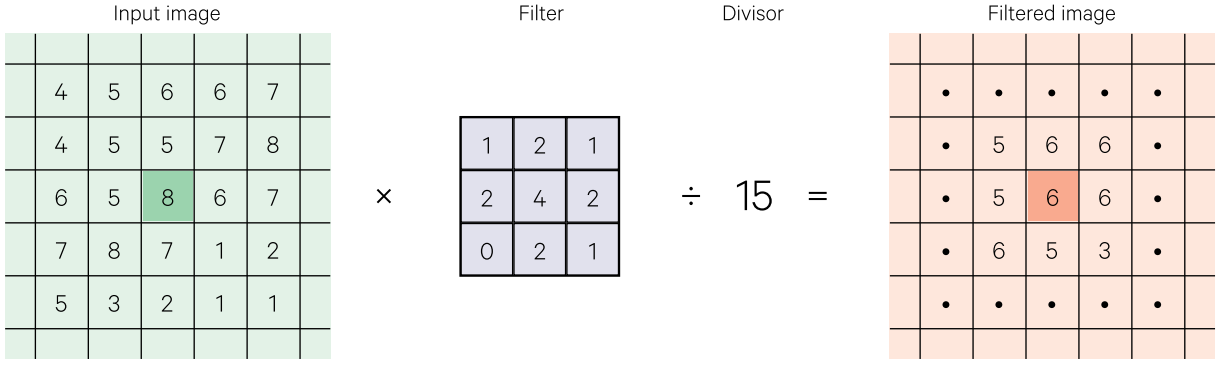
The operation of a 3×3 averaging filter is shown in Figure 3.1a. The filter is applied to each pixel in an image by multiplying it by the filter's central value and adding the products of the other filter elements and their corresponding neighbouring pixel values. Filters can operate by computing a (weighted) mean or average value as illustrated in Figure 3.1, or by determining the median or mode value of the pixels within the filter region (see Sections 4.2 and 4.3 below). In an average filter, a cell weight of zero effectively excludes the cell from the filter region (see Figure 3.1b).

Figure 3.1 Operation of an average filter

a. This operation is applied to every pixel in an image, that is, every pixel is treated as the central value in the filtering operation to determine its output value. To do this, the filter is systematically passed over the image and moved by one pixel after each operation. In this example, the filtered value for the central input pixel (e) will be: A = (a×m + b×n + c×o + d×p + e×q + f×r + g×s + h×t +i×u) / W



b. Numeric example. Note that edge pixels have not been filtered in this example.



Source: Harrison and Jupp (1990) Figure 67

The effect of the spatial filtering process on an image largely depends on filter type (and the weighting values used in the case of an average filter), while the extent of this effect depends on filter size. Multiple applications of the same or different average filters can be computed as a larger single filter by convolution of the original filters (see Excursus 3.1).

One variation for an average filter is called the Olympic filter, which excludes the minimum and maximum values within the filter region before computing the average value. Alternatives to computing the mean, median or mode of pixel values within the filter region include replacing the central value with the minimum or maximum value (Jensen, 2016).

Excursus 3.1—Filter convolution

Two filters A and B applied sequentially to an image would have the same effect as a single application of the filter C where C is the convolution product of A and B. This convolution process is shown in Figure 3.2. The value of a cell in the combined filter is computed as the product of the weights of the overlapping cells between the two filters when one filter is centred in the new filter region and the other filter is moved around it. This process is useful for determining higher order derivative filters for edge detection or surface shape delineation (see Sections 5 or 6).

Figure 3.2 Convolution of two filters

a. To compute the weights of a 5×5 filter, C = A×B, we sum the products of overlapping cells between A and B, when one filter is centred in the 5×5 region and the other filter is moved around it. The result becomes the weight of the cell at the centre of the moving filter. In this example, for the filter weight labelled X in Filter C, the overlap between Filters A (centred) and B (moving) includes only one cell, the top left cell in Filter A (value a) and the lower right cell in Filter B (value r). These values are multiplied to create the filter value a×r for cell X in Filter C.

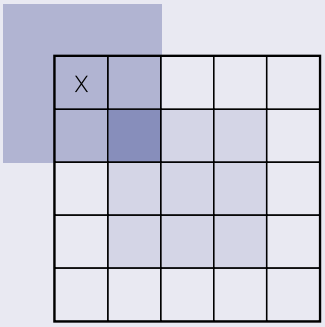
Filter A (3×3)

a	b	c
d	e	f
g	h	i

Filter B (3×3)

j	k	l
m	n	o
p	q	r

Filter C (5×5) is created by convolution of filters A and B



b. All resulting values for 5×5 Filter C

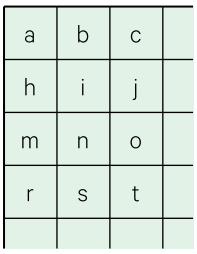
a×r	a×q + b×r	a×p + b×q + c×r	b×p + c×q	c×p
a×o + d×r	a×n + b×o + d×p + e×r	a×m + b×n + c×o + d×p + e×q + f×r	b×m + c×n + e×p + f×q	c×m + f×p
a×l + d×o + g×r	a×k + b×l + d×n + e×o + g×q + h×r	a×j + b×k + c×l + d×m + e×n + f×o + g×p + h×q + i×r	b×j + c×k + e×m + f×n + h×p + i×q	c×j + f×m + i×q
d×l + g×o	d×k + e×l g×n + h×o	d×j + e×k + f×l + g×m + h×n + i×o	e×j + f×k + h×n + i×n	f×i + i×m
g×l	g×k + h×l	g×p + h×q + i×r	h×k + i×k	i×j

Source: Harrison and Jupp (1990) Figure 68

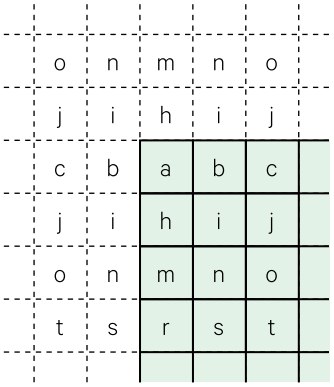
Figure 3.3 Treatment of edge pixels in an image during filtering operation

The values of pixels adjacent to an image edge are mirrored or reflected about the edge to provide a filtering neighbourhood for edge pixels.

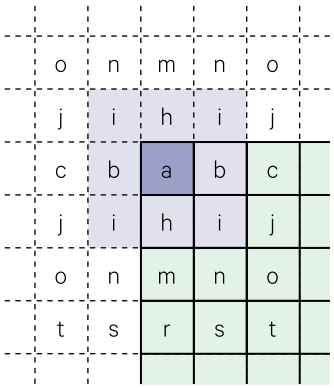
a. Input image values



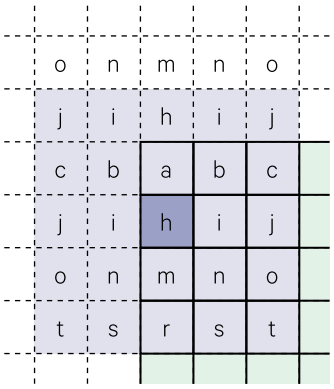
b. Reflected edge values



c. A 3×3 filter centred on the pixel with value 'a' uses the image values highlighted below.



d. A 5×5 filter centred on the pixel with value 'h' uses the image values highlighted below.



Source: Harrison and Jupp (1990) Figure 70

Pixels at the edge of an image (that is, in the first or last row or column) are typically processed by 'reflecting' the pixel values of the adjacent rows or columns to 'fill' the filter being used. This process is illustrated in Figure 3.3. The edge effects are not significant with smaller filters, but should be considered with larger filters especially if image width is relatively small. When significant, they can result is anomalous pixel values around the edges of the image (see Figure 4.14).

Null values in an image channel are generally ignored by the filtering process. In such cases, the filter divisor is appropriately decremented to compensate for the missing value.

3.2 Filter Weights

Image filters can be defined to:

- reduce spatial variation in the image—such as striping patterns or noise spikes (see Section 4.5); or to
- enhance selected spatial patterns—such as linear features or edges (see Section 5).

The impact of a particular filter is determined by the weighting values used. The relationship between smoothing, edge-enhancing and edge-detecting filters is illustrated in Excursus 3.2.

Excursus 3.2—Example of Filter Operations

This example uses a scanned photograph to illustrate the differences between smoothing, edge enhancing, and edge detection filters. The original image is shown in Figure 3.4a. The equivalent filter operation matrix is shown above it, which could be considered a null filter since each pixel remains unchanged.

The impact of a strong smoothing filter applied to our example image is illustrated in Figure 3.4b. This filter replaces each pixel in the image with the average of its original value and the values of its eight neighbouring pixels. The resulting image is clearly a blurred version of the original in which edges are less distinct. Smoothing filters are discussed in more detail in Section 4.

Edge-enhancement filters, by contrast, emphasise the difference between the central pixel and its neighbours, which results in clearer transitions between different tones (see Figure 3.4c). This operation enhances minor variations in the original image, such as image artefacts, as well as actual edges.

To highlight transition zones, edge detection filters effectively subtract the smoothed image from the original as illustrated in Figure 3.4d. This operation is evident in both the filter design and the resulting image. Filters to detect and enhance edges are detailed in Section 5.

Figure 3.4 Relationship between smoothing, edge enhancement and edge detection filters

a. Original image =null filter

0	0	0
0	9	0
0	0	0

÷ 9

b. Smoothed image = original image – edges

1	1	1
1	1	1
1	1	1

÷ 9

c. Edge enhanced image = original image + edges

-1	-1	-1
-1	17	-1
-1	-1	-1

÷ 9

d. Edge detection image = original – smoothed image

-1	-1	-1
-1	8	-1
-1	-1	-1

÷ 9

Source: Harrison and Jupp (1990) Plate 1

When applied to geophysical imagery (for example, magnetics or gravity data) edge-enhancement (or high-pass) filtering can allow regional trends to be removed, which simplifies analysis of small area variations (see Section 5.2). Alternatively, local variations may be removed to explore deeper or larger geological structures using smoothing (or low-pass) filters (see Section 4). Such filters are commonly applied to frequency data but may be defined in the spatial domain using an inverse Fourier transform on the desired frequency response and shortening to a practical filter size as detailed in Fuller (1966).

Standard filtering techniques derive the filtered value of the central pixel using all (active) pixels in the filter neighbourhood. When the filter lies across the edge of a feature in the image however, smoothing operators tend to blur feature boundaries and de-emphasise small features (see Section 4). Alternative filtering approaches, such as adaptive filtering (see Section 3.5), minimise this effect.

3.3 Filter Size

Larger filters obviously involve more computations and therefore require longer processing time. While these differences may not be significant for small data volumes, they are a consideration with large datasets, even with ongoing improvements in processing speeds. Table 3.1 shows the relative processing times required for different filter sizes. As demonstated in Excursus 3.1, larger filter sizes can be simulated by repeated filtering with smaller filters.

Table 3.1 Relative processing times for different filter sizes

Filter Size	Time Factor
3×3	1.0
5×5	1.7
7×7	2.7
9×9	4.2
11×11	6.0

3.4 Thresholding

Filtering can also be conducted in conjunction with thresholds such that the centre pixel is only replaced by a filtered value if the difference between the original and filtered values is greater than some specified magnitude. Thresholds are typically defined in terms of the image data range or as percentages of that range.

Percentage-based thresholds result in a greater filtering impact on pixels with low values and, given the skewed distribution that is typical for EO imagery, thresholds based on percentages are generally appropriate. In this case, the central pixel value is only modified if the percentage-based threshold, *x*, is greater than a user-defined value, where:

$$x = \frac{\text{original value} - \text{filtered value}}{\text{original value}} \times 100$$

Thresholding is useful for removal of noisy spikes in image data (see Section 4.5), without blurring the low frequency variations, or edge features with lower frequency than the noise. Image histogram information may be useful when selecting appropriate levels for filter thresholds (see Volume 2A—Section 8).

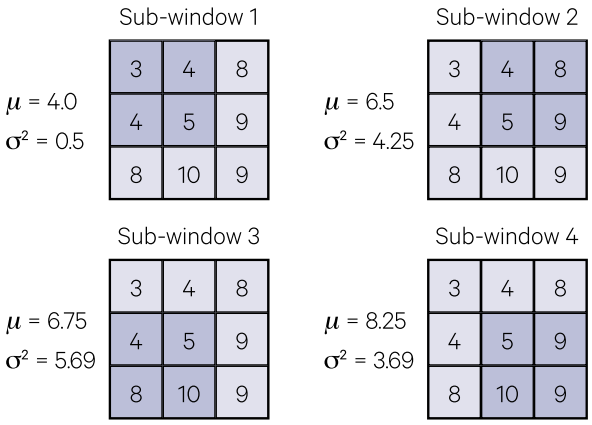
3.5 Adaptive Filtering

Adaptive filtering defines sub-windows within the filter region and uses some criterion, such as minimum variance, to select the 'best' sub-window to use to compute the filtered value for the central pixel. For example, four 2×2 sub-windows could be defined in a 3×3 smoothing filter as illustrated in Figure 3.5. In this case, sub-window 1 has minimum variance so the central value is replaced with the mean of that sub-window. This procedure tends to preserve, rather than blur, edges in smoothing filters and avoids emphasising edges in texture images (see Section 6.1).

Another implementation of adaptive filtering is known as the Wallis Filter (Wallis, 1976), which adjusts local image brightness to maintain or match specified local mean and standard deviation values within user-defined windows (Ballabeni *et al.*, 2015). This process results in more uniform tone, with improved local contrast in bright and dark regions (see Figure 3.6), so is particularly useful for highlighting local variations in imagery with uneven illumination. Adaptive filtering is further discussed in Section 4.4.

Figure 3.5 Adaptive filtering

A 3×3 filter can be divided into four sub-windows to allow selection of the filter region that is most compatible with the central pixel value. If the criterion of minimum variance were used to select a sub-window, the filtered value of the central pixel would be the mean of the top-left window, with value 4.0.



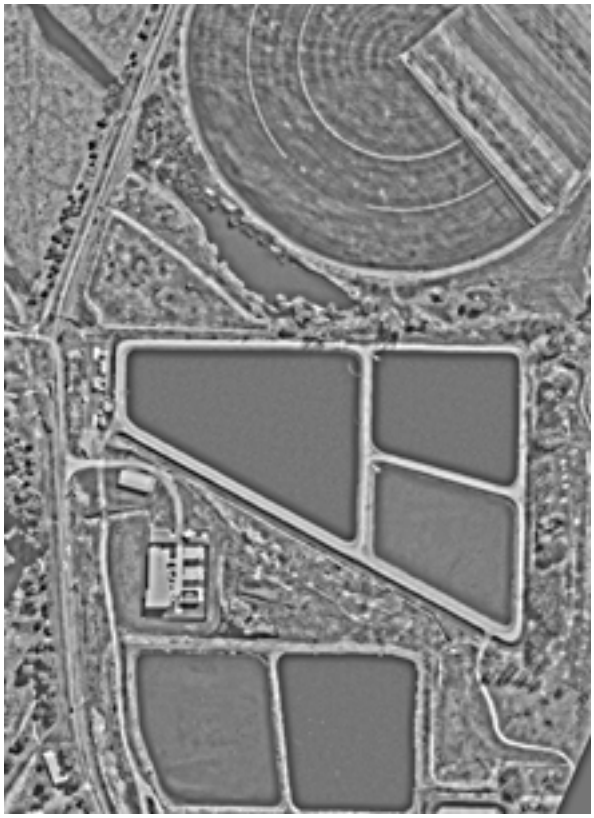
Source: Harrison and Jupp (1990) Figure 69

Figure 3.6 Wallis filter

This example, based on Pleiades pan-enhanced blue band, is centred over part of Fyshwick, ACT, and was acquired on 15 August 2016. The Wallis filter visibly enhances variations within features, especially the settlement ponds of the Fyshwick Sewerage Treatment Plant (located in central and lower portions of image) and the Canturf irrigated lawns (in top of image).

a. Pan-enhanced blue band

b. Result of Wallis filtering



Source: Tony Sparks, Icon Water

3.6 Homomorphic Filtering

The application of filtering techniques to non-linearly scaled image data can produce some interesting results. 'Homomorphic' filtering generally describes the application of a linear, high-pass filtering operation to a logarithmically-scaled image, then exponentially-transforms the result back to a linear scale (Oppenheim and Schafer, 2004; Pitas and Venetsanopoulos, 1990). The additive operation of the filter on the logarithmically-scaled image has a multiplicative effect on the original image values. This form of filtering has traditionally been used to remove multiplicative noise from an image, especially when due to non-uniform illumination (Moik, 1980).

As an image enhancement tool, the transformation sequence reduces the overall image illumination component and enhances local contrast related to feature-specific reflectance. For example, this type of processing would enhance the subtle variations in image radiance values of shallow water features. With EO imagery, homomorphic filtering has also been used to remove cloud cover (Mitchell *et al.*, 1977).

3.7 Further Information

Castleman (1998) Sections 9, 11 and 12

Gonzalez and Woods (2018) Section 3

Jensen (2016) Section 8

3.8 References

Ballabeni, A., Apollonio, F.I., Gaiani, M., and Remondino, F. (2015). Advances in Image Pre-Processing to Improve Automated 3D Reconstruction. Int. Arch. Photogram. Remote Sensing and Spatial Information Sciences, Volume XL-5/W4. *Proc. 3D Virtual Reconstruction and Visualization of Complex Architectures*. 25-27 February 2015, Avila, Spain.

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Fuller, B.D. (1966). Two-dimensional frequency analysis and design of Grid Operators. In *Society of Exploration Geophysicists Mining Geophysics. Vol II. Theory*. (Ed: D.A. Hansen).

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Mitchell, O.R., Delp, E.J., and Chen, P.L. (1977). Filtering to remove cloud cover in satellite imagery. *IEEE Transactions on Geoscience and Remote Sens*ing, 15, 137–141.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images*. NASA SP–431. Washington, DC. USA.

Oppenheim, A.V., and Schafer, R.W. (2004). From frequency to quefrency: A history of the cepstrum. *IEEE Signal Processing Magazine*, 21(5), 95–106.

Pitas, I., and Venetsanopoulos, A.N. (1990). Homomorphic Filters. Ch 7 in *Nonlinear Digital Filters*. The Springer International Series in Engineering and Computer Science (VLSI, Computer Architecture and Digital Signal Processing), Vol 84. Springer, Boston, MA

Richards, J.A. (2013). *Remote Sensing Digital Image Analysis: An Introduction*. 5th edn. Springer-Verlag, Berlin. ISBN 978-3-642-30061-5

Wallis, R. (1976). An approach to the space variant restoration and enhancement of images. *Proc. Symp. Current Mathematical Problems in Image Science*, Monterey, Naval Postgraduate School, pp. 329–340.

4 Smoothing (Low Pass)

The operation of spatial filters is introduced in Section 3.1. Various forms of smoothing filters are available, which replace the value of the central pixel in a pre-defined filter region with another value determined from the values of pixels within that filter region. Commonly available filters for image smoothing include:

- average or mean, which compute the (weighted) average of all pixels in the filter region (see Section 4.1);
- median, which select the median value of pixels within the filter region (see Section 4.2); and
- mode, which select the most common pixel value in the filter region (see Section 4.3).

The effects of these three types of filtering operation are illustrated for different filter sizes in Figure 4.1.

Other forms of smoothing filters include:

- edge-preserving filters (see Section 4.4), which are based on adaptive windows (see Section 3.5);
- despiking, which replace noise 'spikes' (pixels with markedly different values to their neighbours) with values derived from neighbouring pixels (see Section 4.5); and
- interpolation, which fill missing pixels with values derived from neighbouring pixels (see Section 4.6).

He was always smoothing and polishing himself, and in the end he became blunt before he was sharp.
(Georg Christoph Lichtenburg)

Background image: Landsat-8 OLI image of coastal waters just north of Brisbane, Queensland, acquired on 7 October 2014, displayed using bands 4, 3, 2 as RGB.
Source: Craig Shephard, DSITI

Figure 4.1 Effect of different filter sizes and operations

a. Average filtering (using weights of 1 with a divisor equal to the number of pixels in the filter)

Original Image | Filter size: 3×3 | Filter size: 7×7 | Filter size: 11×11

b. Median filtering

Original Image | Filter size: 3×3 | Filter size: 7×7 | Filter size: 11×11

c. Modal filtering

Original Image | Filter size: 3×3 | Filter size: 7×7 | Filter size: 11×11

Source: Harrison and Jupp (1990) Plates 2 and 3

4.1 Average Filters

Image smoothing is best applied to EO data using a mean or average filter. In most image processing systems, the calculation of this filter may be performed in integer or real arithmetic, the latter being slower but more precise. Similarly, average filters can often be defined as integer or real weights, with a divisor and offset being optionally applied to the filter product. The use of divisors simplifies the specification of the filter and enables scaling into the image data range. The filtering operation can generally be applied to a selection of, or all, channels in an image.

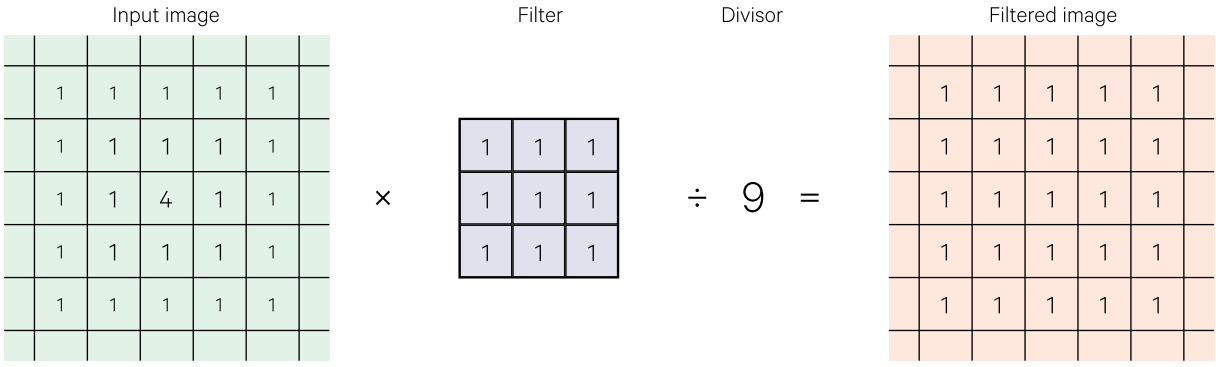
The operation of a simple average filter is introduced in Section 3.1 (see Figure 3.1). When all filter weights are defined using positive values, this process has a blurring or smoothing effect on image features with abrupt changes in value in the input image being softened into more gradual transitions in the filtered image (see Figure 4.2). Such filters are also referred to as low pass or regional variation filters as they effectively enhance the low frequency or regional information in an image by reducing higher frequencies.

Both the weighting values used in an average filter and the filter size combine to determine the type and extent of effect the filtering process will produce on an image. Smoothing average filters usually contain all positive weights with the central value being greater than or equal to other weights. The effect of a filter can be reduced by increasing its central value relative to the other values in the filter region. The filter size obviously determines the number of pixels that are involved in calculating the transformed value for each pixel, with larger filters producing a greater smoothing effect as illustrated in Figure 4.3.

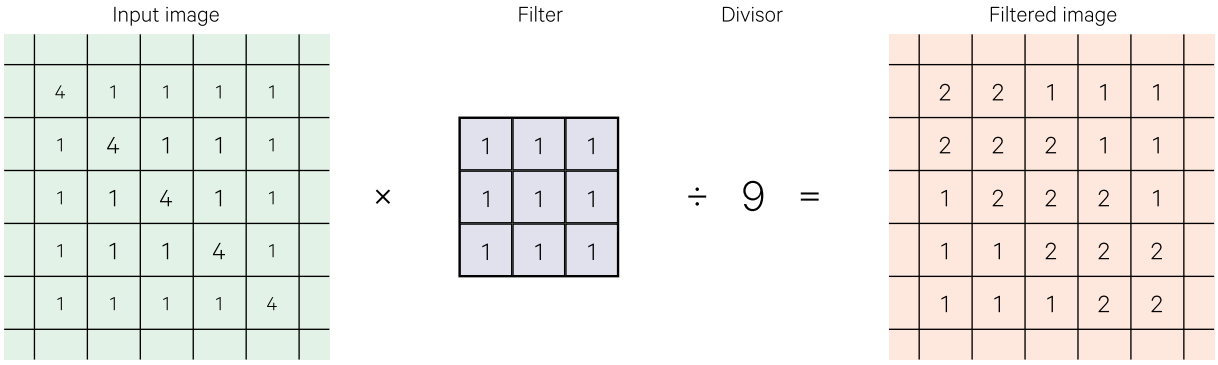
Most implementations of image filtering allow both filter size and weighting values to be varied (see Figure 4.1 and Figure 4.3). The size and weights of the selected filter will depend on the application of the operation. In many image processing systems the default filter is a 3×3 average filter, with filter weights similar to those shown in Figure 4.4a. This filter is formed from the outer product of the two one-dimensional smoothing filters shown in Figure 4.4b.

Figure 4.2 Effect of smoothing filter on image data

a. Image noise is reduced with a smoothing filter.



b. Linear features are broadened and reduced in intensity by smoothing. Note: in this example, the filtered values for edge pixels have been computed assuming that the pattern of the sub-image is continuous.



Source: Harrison and Jupp (1990) Figure 71

Figure 4.3 Effect of different filter sizes and weights

a. Weighted average filtering (using filter weights shown in Figure 4.4a)

Original Image | Filter size: 3×3 | Filter size: 7×7 | Filter size: 11×11

b. Full average filtering (using filter weights shown in Figure 4.5a)

Original Image | Filter size: 3×3 | Filter size: 7×7 | Filter size: 11×11

Source: Harrison and Jupp (1990) Plate 2

This is a relatively 'gentle' filter, that is, the value of the central pixel is largely retained, with the surrounding pixels contributing to the filtered value in proportion to their distance from the filter centre (the centres of pixels in adjacent rows and columns being closer to the centre of the filter than the centres of the four diagonal pixels). The default filter weights for larger filter sizes are derived by convolution of two smaller filters as detailed in Volume 2X—Appendix 5. The impact of these filter weights on image values is illustrated in Figure 4.3a using the default filter defined in Figure 4.4a.

Filtering operations such as smoothing generally do not rescale the data range of an image channel. Non-rescaling is ensured in an average filter if the sum of filter weights divided by the divisor equals 1. Figure 4.5 shows two simple full average filters with Figure 4.5b being a shaped window version of Figure 4.5a. Such filters would have a strong blurring effect on the image data values (see Figure 4.3b).

Figure 4.4 Weights for gentle smoothing filter

a. This gentle smoothing filter could be used to soften EO imagery.

1/36	4/36	1/36
4/36	16/36	4/36
1/36	4/36	1/36

or

1	4	1
4	16	4
1	4	1

with divisor = 16

b. This filter is formed from the outer products of two one-dimensional filters.

1
4
1

and

1	4	1

Figure 4.5 Full average filters with differing weights

a. A square filter, with all cells active, produces heavily smoothed image values (see Figure 4.1a and Figure 4.3a).

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

or

1	1	1
1	1	1
1	1	1

with divisor = 9

b. A filter with inactive corner cells does not consider the diagonally adjacent pixels when computing the filtered value. Smoothing filter weights are typically graduated to have higher values for those pixels that are closest to the central value. This graduation results in a 'softer' smoothing effect than can be achieved using a filter with equal weights in all cells.

0	1/5	0
1/5	1/5	1/5
0	1/5	0

or

0	1	0
1	1	1
0	1	0

with divisor = 5

Source: Harrison and Jupp (1990) Figure 72

In some implementations of average filtering, the effect of the overall filtering operation may also be varied by specifying a weighting value (λ), with values between 0 and 1. The weighting value is used to compute a weighted average of the filtered result and the original value for each pixel and is defined as:

$$\text{transformed value} = \lambda \times \text{original value} + \left(1 - \lambda\right) \times \text{filtered value}$$

so that a λ value of 0 results in full smoothing and a value of 1 retains the original pixel value.

In some image processing systems, threshold values may also be specified for each image channel to limit the smoothing operation to pixels where the percentage difference between the smoothed and original values (that is, *difference / original value* × 100) is greater than the threshold percentage value (see Section 3.4). Specification of the threshold as a percentage, rather than as a digital value, results in a greater filtering effect on low image values. This is usually desirable for EO data given its typically skewed distribution. Image histogram information may be useful when selecting appropriate levels for filter thresholds (see Volume 2A—Section 8.1.1).

The average filter operation can be used to reduce image noise but also blurs edge features. Within image patches such as vegetated paddocks, this process has an equalising effect so that value variations are reduced. However, definition of linear features such as edges or roads is also reduced by average filtering, although this can be advantageous for image destriping as discussed in Section 2.4. Noisy imagery often benefits from a smoothing operation before edge detection or enhancement filters are used since these operations will enhance both noise and edges.

An unbalanced filter design may be selected in an average filter to destripe an image. For example, to destripe imagery with 6 line (horizontal) striping, we could use a 7×7 filter (that is, greater than the six line striping pattern) with all values except those in the central column of the filter being zero so that only horizontal patterns in the image are modified (see Figure 4.6). Such weights should minimise the blurring effect of an average filter.

Image average smoothing effectively represents a process of removing edges from features within an image. The impact of smoothing can be seen by differencing the original and filtered images as shown in Figure 3.4 (see Volume 2D regarding image differencing). As such, this transformation performs a complementary process to edge enhancement (see Section 5.2).

Figure 4.6 Example of destriping filter

The divisor equals the sum of the coefficients to retain the same range of values in the image.

0	0	0	1	0	0	0
0	0	0	4	0	0	0
0	0	0	16	0	0	0
0	0	0	64	0	0	0
0	0	0	16	0	0	0
0	0	0	4	0	0	0
0	0	0	1	0	0	0

with divisor = 106

4.2 Median Filters

Image smoothing can also be effected by selecting the median of the image values within the filter region as the replacement value of the centre pixel. The median of a set of values is the middle value by count when the values are ordered (see Figure 4.7). Smoothing on this basis allows single pixels with anomalous values to be totally disregarded. By definition, the actual value of the median for the filter region already occurs in the image so it could be considered a 'reasonable' replacement value for the central pixel.

The difference between average and median filtering is demonstrated for a one-dimensional function (such as a line of image values in one channel) in Figure 4.8. The mean or average, a_m, of a set of points, x_i, would minimise the expression:

$$\sum \left(x_i - a_m \right)^2$$

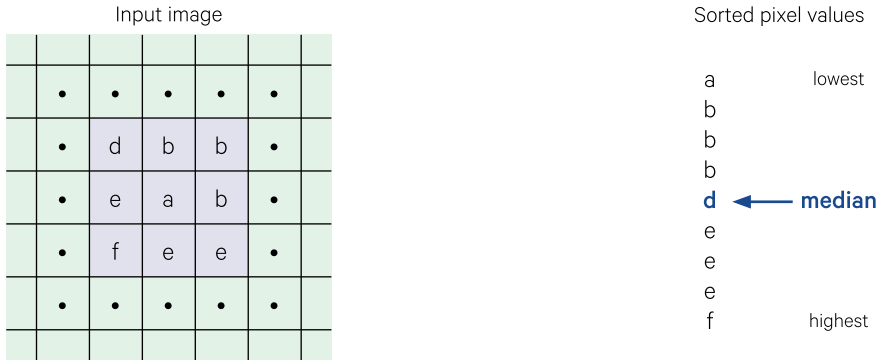
whereas the median, a_n, minimises the expression:

$$\sum \left| x_i - a_n \right|$$

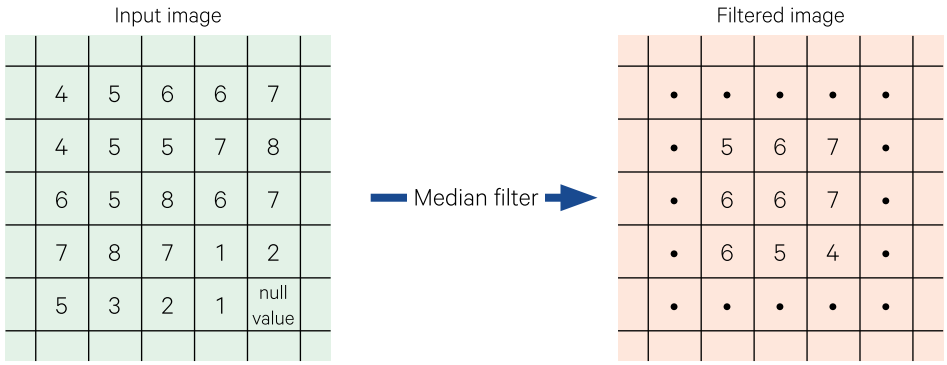
An example of results from moving average and moving median is given in Figure 4.9.

Figure 4.7 Operation of a median filter

a. The filtered value for the central pixel with value a in the input image is the median (or middle ranking value) of the values in the filter region—in this case, value d.



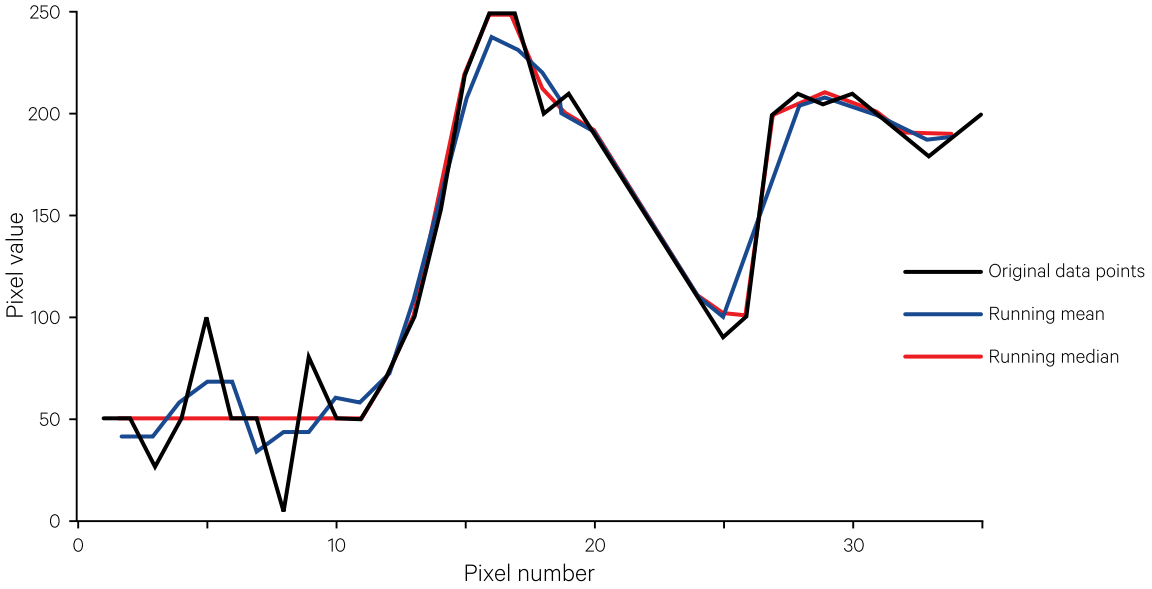
b. Numeric example. If an odd number of null values exist in the filter region, then two median values can occur. In this case, the two medians are generally averaged to generate the output value.



Source: Harrison and Jupp (1990) Figure 73

Figure 4.8 Average versus median smoothing

The effect of mean and median filtering is shown for a one-dimensional function. The median filter removes outlier values and preserves edges more effectively than the mean filter.



Source: Harrison and Jupp (1990) Figure 74

Figure 4.8 and Figure 4.9 demonstrate the way a median filter removes outlier points whereas averaging merely spreads their effect into adjacent data points. Sudden changes in data values however, such as feature boundaries, are also better preserved by median filtering as illustrated in Figure 4.1.

Repeated median filter operations tend to quickly reach a stable end point (Tukey, 1977) whereas average smoothing can cycle through a series of minor variations in pixel values. Median smoothing provides a powerful tool for noise reduction in image processing with less loss of edge information. This filtering mode is commonly used in conjunction with thresholds for despiking an image as described in Section 4.5.

Median filters can be applied in most image processing systems. The effect of filter size with a median filtering is illustrated in Figure 4.1 above. If an even number of non-null pixels occurs in a filter region (that is, some pixels may be equal to the null value so are excluded from the filtering calculation) the average of the two mid-range values is generally computed as the output median value.

Figure 4.9 Example of moving average and moving median

a. For the dataset below, the mean value is 15 / 6 = 2.5 while the median values is 1.0.

1.0	0.5	1.0	1.5	9.0	1.0	1.0

b. The moving average values (that is the average of each triplet of values) would be:

•	0.83	1.0	3.8	3.8	3.7	•

c. The moving median values are:

•	1.0	1.0	1.5	1.5	1.0	•

4.3 Modal Filters

Imagery can also be smoothed by selecting the mode, that is, the most 'popular' value, within the filter region as shown in Figure 4.10. This approach is particularly useful for nominal (categorical) data where the ordering of values does not imply a relationship between them since their average or median would be a meaningless result. The effect of variations in filter size in the modal filtering operation is illustrated in Figure 4.1 above.

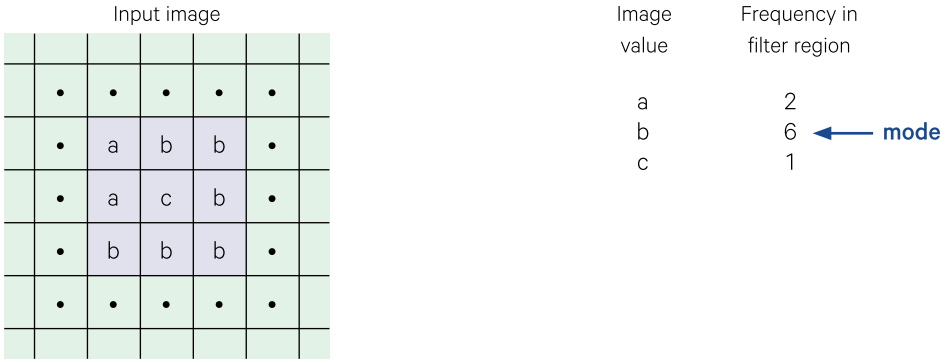
This operation performs a type of 'logical' smoothing by assuming that the central pixel should be related to its surrounding pixels. Typically, in an image classification of EO data for example, patch size will be larger than a single pixel so that patches consisting of only one pixel are likely to be noise. When the classification is represented as a single image channel, modal filtering can be used to replace isolated pixel patches with the most commonly occurring class in the filter region. Modal filtering can often be conducted in conjunction with thresholds, which only permit the centre pixel to be replaced by a filtered value if the mode occurs in a minimum number of pixels within the filter region (see Section 3.4).

This filtering operation is particularly useful for nominal data, such as an image classification channel, where the original image values can be reorganised but not interpolated (see Figure 4.11). When a filter region is bi-modal, that is there are two modes, a predefined rule is needed to determine whether the one with the lowest or highest image value is used. Any automatic selection of the lowest or highest value in bi-modal filter regions assumes that an implicit priority order exists in the nominal data, which should be considered when the data are coded.

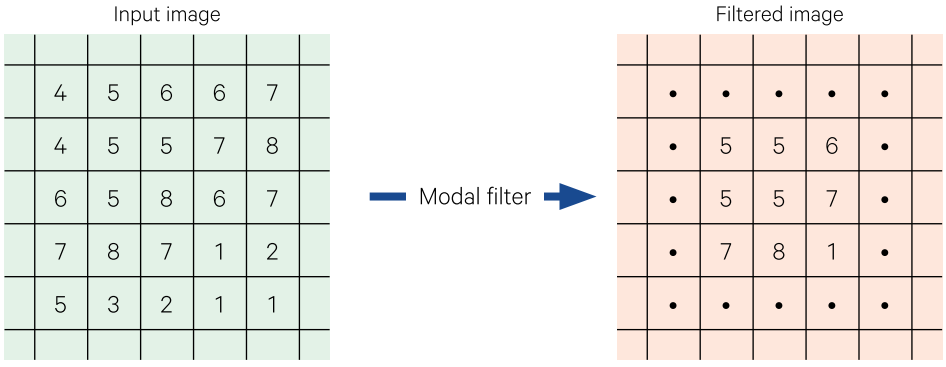
Nominal data may be derived from EO imagery, such as a classification channel, or created to represent ancillary data from map sources. The modal filter is a very effective image 'cleaning' tool after classification to remove small, disconnected clusters (see Volume 2E).

Figure 4.10 Operation of a modal filter

a. The filtered value for the central pixel (with value c in the input image) becomes the most commonly occurring value, or mode, in the filter region—in this case, value b.



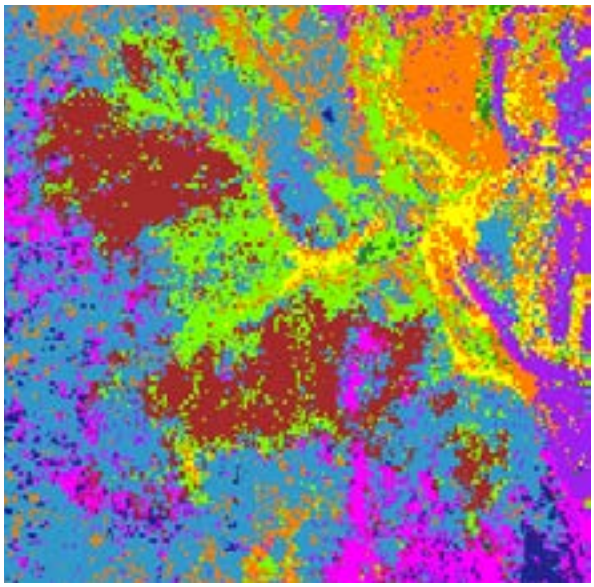
b. Numeric example. Where two or more values are equally 'popular' in the filter region, a predefined rule determines which value is used as the model. In this example, the lowest value is selected for the output pixel value.



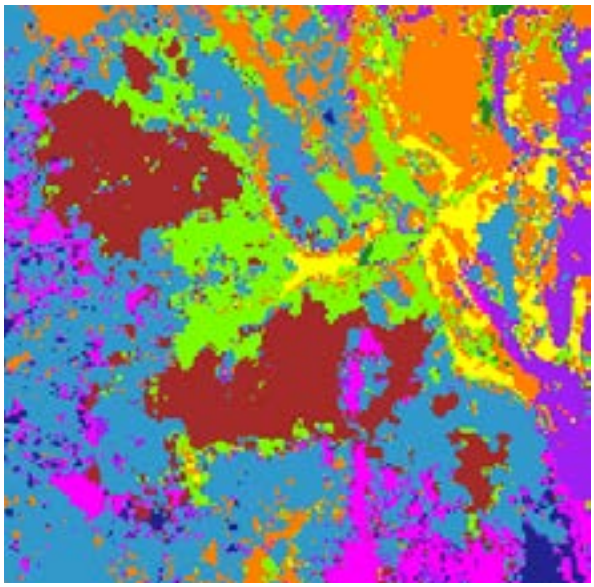
Source: Harrison and Jupp (1990) Figure 75

Figure 4.11 Modal filtering of classified image

a. Original classified image



b. Classified image after three passes of 3x3 modal filter



Source: Megan Lewis, University of Adelaide

4.4 Edge-Preserving Filters

As the name implies, edge-preserving smoothing filters are designed to smooth large, homogeneous patches in an image while preserving edges. These filters are particularly appropriate for pre-processing of noisy images (Cheng, 1989), including SAR data (Argenti *et al.*, 2013; Kupidura, 2016). Various algorithms have been proposed for edge-preserving filtering, not all of which are appropriate for EO imagery.

One type of edge-preserving filters, originally designed for medical imaging (Kuwahara *et al.*, 1976) uses adaptive windows within the filter (see Section 3.5). This technique involves selecting the best sub-window to use for the filtering process. The selection of sub-windows is based on some relevant criterion such as minimum variance (see Figure 3.5) or local standard deviation. This process removes patches smaller than the filter size and results in an image in which the boundaries of major patches are preserved rather than blurred, as typically occurs with smoothing operations. Other categories of edge-preserving filters include bilateral filters (which use differences between pixel values to determine the contribution of each pixel to the local average) and anisotropic diffusion filters (which iteratively smooth the image using the image gradient to prevent local 'diffusion' at edges; Perona and Malik, 1990; Burger and Burge, 2016).

The effect of filter size on edge-preserving filtering is illustrated in Figure 4.12. Two average filters are compared with an edge-preserving filter for a scanned photograph in Figure 4.13. The differences between the filters are evident in this example.

Alternative approaches to adaptive filtering rely on computing the difference between the central pixel and the mean of the filter region. For example, Eliason and McEwen (1990) use this difference in two ways:

- to remove random errors—pixel values are only replaced with the local mean when they differ from it by more than 1–2 standard deviations; and
- to smooth image noise—when computing the value for the central pixel, this approach only averages those pixels within the filter whose values are within 1–2 standard deviations of it (see Figure 4.14). This approach is similar to the sigma filter proposed by Lee (1983) and is often called a 'local sigma' filter.

Figure 4.12 Effect of filter size on edge-preserving filters

a. Original Image b. Filter size: 3×3 c. Filter size: 7×7 d. Filter size: 11×11

Source: Harrison and Jupp (1990) Plate 4

Figure 4.13 Comparison of average and edge-preserving filters

A scanned photograph has been filtered using three different 3×3 filters.

a. Original image b. Full average filter c. Weighted filter d. Edge-preserving filter

Source: Harrison and Jupp (1990) Plates 2 and 4

Figure 4.14 Edge-preserving filter applied to EO image

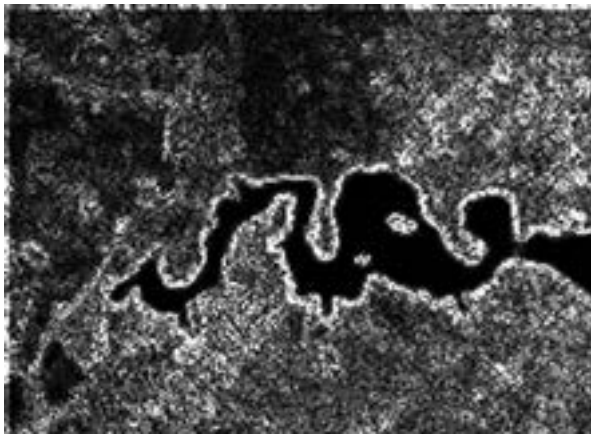
A Landsat-8 OLI image of Canberra, acquired on 10 August 2017 and shown with bands 6, 5, 3 as RGB, has been filtered in two ways. Large homogeneous patches are more even in both smoothed images but edges are visibly clearer after the edge-preserved filtering (Eliason and McEwen, 1990) compared with full average filtering. In this example, the active region of both filters was 3×3 pixels and the difference between them highlights edge features. The difference channel is based on Euclidean distance with brighter pixels indicating larger differences. Note that different treatment of edge pixels by the two filters are visible along the top and left edges of the difference image[3].

a. Original image

b. Edge-preserving filter (5x5)

c. Full average filter (3x3)

d. Difference between filtered images



Source: Norman Mueller, Geoscience Australia

4.5 Despiking (Noise Reduction)

A specialised use of image smoothing, called despiking, may be applied to imagery to remove outlier pixels, or noise 'spikes'. These pixels can occur in features of relatively uniform 'colour' and have unrealistically high or low values. For example, image spikes commonly occur in water features, so despiking is a standard pre-processing operation in various shallow water mapping analyses.

A measure of similarity between the central pixel and its neighbours can be used to identify image spikes. Then for every central pixel which is not 'similar' to its neighbours, its value can be replaced by a new estimate based on the neighbourhood values, or by a null value. This general function can be implemented to allow:

- selection of filter neighbourhood size;
- different measures to determine similarity;
- threshold values to identify spikes;
- replacement values based on neighbourhood values; and
- iteration of the despiking process.

3 In this example filtering was applied to a larger image so that the edge effect is only visible on two edges of the sub-image.

Measures of similarity can be defined to filter either by:

- outlier—based on some 'distance' from neighbouring values; or by
- rarity—based on the frequency of values occurring in the neighbourhood.

These two basic types of similarity generally apply to numerical and nominal data respectively (see Section 1.2.2 for definition of these terms).

A suitable measure for an outlier would be the numerical 'distance' between the central value and the values of the pixels in the selected neighbourhood (see Volume 2E). The distance function could be the sum of squares:

$$d_2\left(x_{ij},N\right)=\sum_{k,l\in N}\left(x_{ij}-x_{kl}\right)^2$$

or the sum of absolute values:

$$d_1\left(x_{ij},N\right)=\sum_{k,l\in N}\left|x_{ij}-x_{kl}\right|$$

where $d_p\left(x_{ij},N\right)$ (with p = 1, 2) represents the distance between the central pixel's values (x_{ij}) and those of the pixels in the filter region, or neighbourhood *N*. The central values that would minimise these distance functions are respectively the mean and median of the pixels in the filter region excluding the central pixel. That is, the neighbourhood mean and median values are most similar to the neighbourhood values as measured by the two distance functions where the neighbourhood can be selected from (usually square) patches of size 3×3 pixels and larger.

An outlier is defined as a pixel for which the distance function is greater than a selected threshold, which is commonly expressed as a percentage of the neighbourhood mean or median. Outliers are typically replaced by the mean, median or mode of the neighbourhood values or a null value. In this way, images can be despiked to remove isolated extreme values very efficiently. This process also alters the neighbourhoods so that a pixel that is not an outlier in the original image may become one if the process is iterated. Thus, iterative despiking can be used to remove higher densities of extreme values.

Rarity allows filtering with a similarity measure that is appropriate to the mode (or most common value) of the neighbourhood. A rarity filtering operation could be described as the number of pixels in the neighbourhood that are NOT the same as the central pixel in value:

$$d_R\left(x_{ij},N\right)=\sum_{k,l\in N}\left(1-\delta_{x_{ij}x_{kl}}\right)$$

where δ_{xy} is 1 if $x=y$, and 0 otherwise. The value of x_{ij} that minimises this distance measure is clearly the mode of the values in the filter region. This distance measure and the mode could be evaluated for any data type, including classified pixels and even null values.

A pixel is considered 'rare' if $d_R\left(x_{ij},N\right)$ is greater than some selected numerical threshold. If a pixel is rare, it can be replaced by the mean, median or mode of the values in the filter region or with a null value. The threshold could be any value between 1 and the number of pixel neighbours in the filter region. If the filter region has a radius *k* the threshold value could be between 1 and $((2k+1)^2-1)$. Using this definition, a large threshold value would indicate that the value of the central pixel is more rare, that is, its value differs from the values of most of its neighbouring pixels.

The key feature of such a rarity option is that it can be applied to nominal data and to nulls with very convenient results for image cleaning. Rare null values can be 'filled' by local means, medians or modes. If null values are scattered through a classification channel, for example, they can be 'filled' using the mode of the data within neighbourhood *N*. Alternatively, if null values are scattered through numerical image data, they may be 'filled' by the mean or median of the (non-null) data in the neighbourhood.

Despiking can be simply applied to imagery in two ways:

- removing the spike pixels with specialised algorithms then filling them with the neighbourhood mode value; or
- using a median filter with 70% thresholds.

Some specialised algorithms identify spikes using a median filter with selected thresholds then set those pixels to the null value in the output image. The use of 70% threshold values during median filtering ensures that a pixel is only altered if its value differs from the median of the filter by 70% or more. Using the median value to replace noisy pixels enforces the use of an existing image value, which is more reliable than the neighbourhood mean value for spikes covering more than one pixel.

Appropriate thresholds for filtering can be determined either by trial and error, or statistically (such as three standard deviation units; see Volume 2A—see Section 8.1.2). The null pixels can be subsequently filled using the mean, median or mode value of the surrounding pixels. This sequence is the preferred method for despiking since it uses the median to detect outliers then allows them to be replaced using some other criteria. The mode value is recommended in this sequence for classified channels.

4.6 Filling Missing Values (Interpolation)

Filtering techniques may be used to interpolate missing values in image data. Such pixels are usually given the image null value. Scanner or reception malfunctions can occasionally cause lines or partial lines in EO images to be 'lost' (see Volume 1B—Section 2). During processing, pixels may be deemed 'missing' after certain transformations (such as smoothed ratios with themes—see Section 10.1.3) or left unclassified after image analysis (see Volume 2E). The boundaries of segmented imagery may also be filled lightly to reduce jagged edges (see Volume 2A—Section 9).

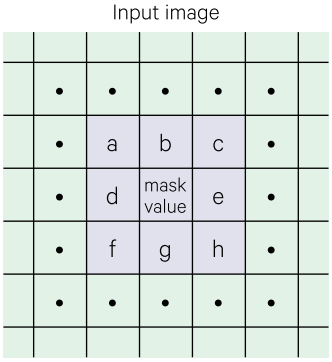
Alternatively, data from non-EO sources, such as water depth transects or terrain elevation, may be registered with an image base but form a sparse dataset relative to the image grid. Provided such data does not contain large contiguous areas of missing values, filter interpolation techniques may be used to fill the missing values (see Volume 2D).

Filter-based interpolation techniques operate in a similar fashion to image smoothing with only the values of the pixels in a 3×3 filter region surrounding the central 'missing' pixel being used to compute the filled value. This process is illustrated in Figure 4.15 using the filter shown in Figure 4.16 for the weighted mean.

Figure 4.15 Filter-based interpolation

Whenever the central pixel in a 3×3 neighbourhood is equal to a specified mask value (such as the null value), it can be replaced by the median, mode or weighted mean (as selected) of the surrounding 'active' pixels (that is, values not equal to the specified value). A minimum number of active pixels can be specified before the central pixel is filled. In this case, the output value for the central mask value could be the:

- median of (a, b, c, d, e, f, g); or
- mode of (a, b, c, d, e, f, g); or
- weighted mean of $\frac{a+4b+c+4d+4e+f+4g+h}{20}$



Source: Harrison and Jupp (1990) Figure 76

Filter interpolation is generally based on a 3×3 mean, median or modal filter to fill missing pixels. An average filter would use predefined weights, such as shown in Figure 4.16. The missing pixel value may generally be selected as any valid pixel value in the image data range (for example 0–255 for byte format images). To control the effect of this transformation, a minimum number of neighbouring non-missing pixels could be specified to ensure that the operation is only applied to isolated missing pixels. If more than one pixel is missing in the neighbourhood, the filter weights would need to be adjusted accordingly to preserve image scaling.

An example of average filter interpolation is shown in Figure 4.17. This example is based on an image acquired by Landsat-7 ETM+ after the Scan Line Corrector (SLC) had stopped operating (see Figure 4.17a). The SLC accounted for forward motion of the satellite and ensured that image lines represented contiguous ground pixels of constant width, without any duplication (USGS, 2018). Geometric pixels that do not have valid image values are indicated as null (missing) pixels in this imagery and shown as black in Figure 4.17a. When appropriate interpolation filtering is applied to such imagery, a continuous image can be created as shown in Figure 4.17b.

With median filtering, if an even number of 'active' (non-missing) pixels occurs in the filter region, a decision needs to be made as to which value to use for the replacement value in the output image. In some image processing systems, the lower of the two mid-range values is used as the replacement value. Modal interpolation filtering of a labelled classification channel provides a convenient method for assigning spatial labels to any unclassified pixels (see Volume 2E).

While large targets may be filled iteratively using this process, care must be taken to ensure that the results represent a reasonable interpolation of the original data (see Volume 2D). Filter-based interpolation is not appropriate for large area interpolation; it simply provides a mechanism for tidying spotty image data or creating a continuous data channel from thickly sampled ancillary data (see Figure 6.9).

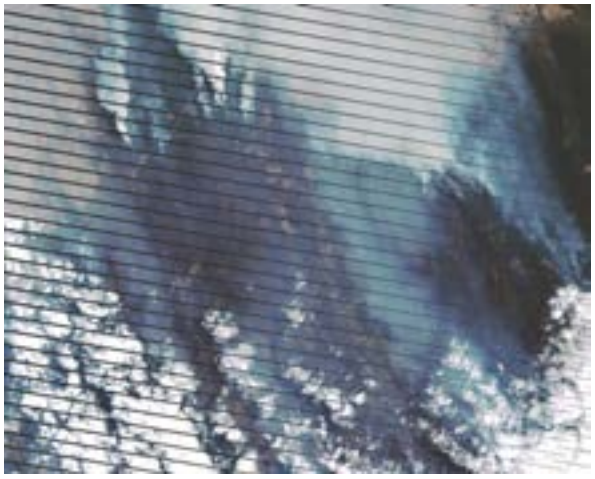
Figure 4.16 Example of average filter

1/20	4/20	1/20
4/20	•	4/20
1/20	4/20	1/20

Figure 4.17 Average filter interpolation

This Landsat-7 ETM+ image was acquired over the Black Saturday fires in central Victoria in 2009.

a. Original image showing missing pixels due to SLC failure.



b. Interpolation filtering creates a continuous image.

Source: Norman Mueller, Geoscience Australia

4.7 Further Information

Castleman (1988) Sections 9, 11 and 12

Gonzalez and Woods (2018) Section 3

Jensen (2016) Section 8

4.8 References

Argenti, F., Lapini, A., Bianchi, T., Alparone, L. (2013). A Tutorial on Speckle Reduction in Synthetic Aperture Radar Images. *IEEE Geoscience and Remote Sensing Magazine*, 1(3), 6–35.

Burger, W., and Burge, M.J. (2016). *Digital Image Processing*. Springer, London.

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Cheng, S. (1989). An adaptive noise smoothing filter for remotely sensed images with microphonic noise. *International Journal of Remote Sensing*, 10, 1015–1034.

Eliason, E.M., and McEwen, A.S. (1990). Adaptive Box Filters for Removal of Random Noise from Digital Images. *Photogrammetric Engineering and Remote Sensing*, 56(4), 453.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Kupidura, P. (2016). Comparison of filters dedicated to speckle suppression in SAR images. *International Archives of Photogrammetric, Remote Sensing and Spatial Information Sciences* XLI-B7. 12–19 July, Prague, Czech Republic. pp. 269–276. doi:*10.5194/isprsarchives-XLI-B7-269-2016*

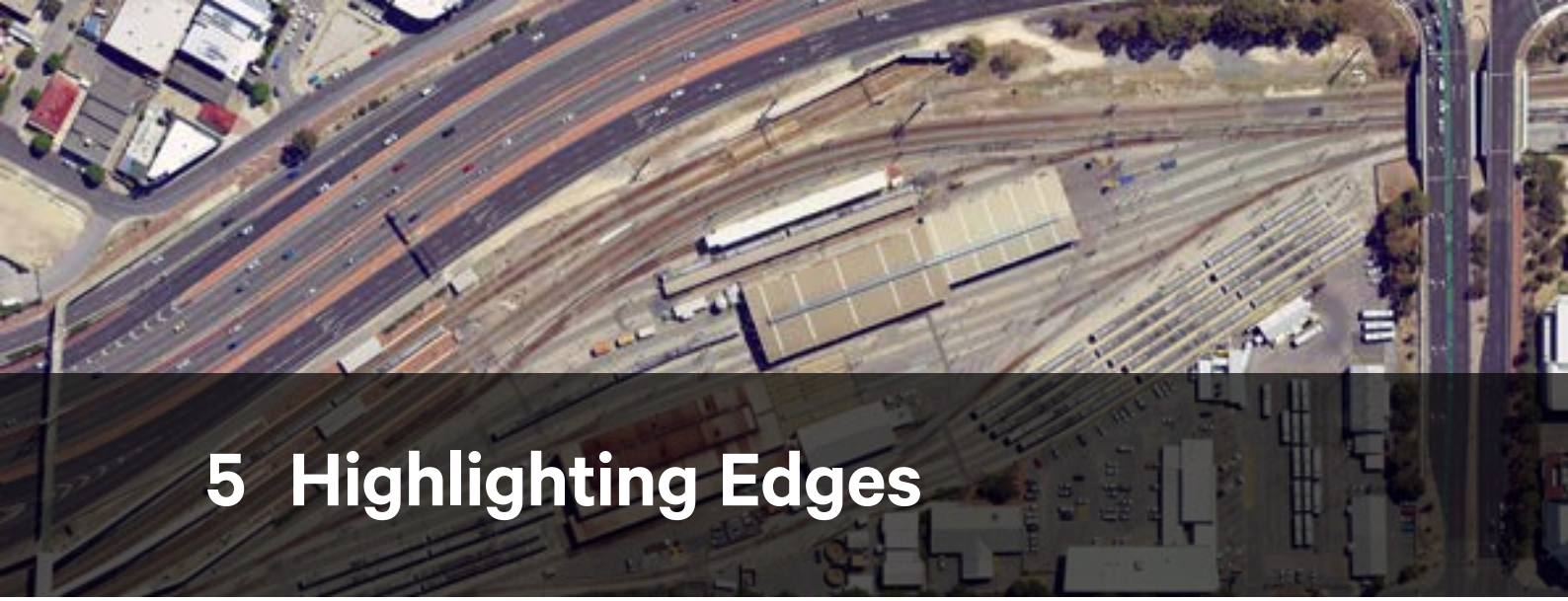
Kuwahara, M. Hachimura, K., Eiho, S., and Kinoshita, M. (1976). Processing of RI-angiocardiographic images. Ch 13 in *Digital Processing of Biomedical Images*. (Eds: K. Preston Jr. and M. Onoe). Springer, Boston. p. 187–202. ISBN 978-1-4684-0771-6

Lee, J.S. (1983). Digital image smoothing and the sigma filter. *Computer Vision, Graphics and Image Processing*, 24, 255–269.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7),* 629–639.

Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, Mass.

USGS (2018). *SLC-off Products*: *Background* webpage: *https://landsat.usgs.gov/slc-products-background*



5 Highlighting Edges

Edge features in EO imagery can be highlighted using differential filters. The following sub-sections consider filters that:

- detect edges (see Section 5.1); and
- enhance edges (see Section 5.2).

5.1 Edge Detection (Differential)

The filters described for image smoothing manipulate image values using balanced filter weights around the central pixel value (see Section 4). The effect of smoothing operations is to effectively reduce the strength of the central pixel value relative to those other pixels surrounding it.

In a 'balanced' filter, the same weighting values are used for all adjacent pixels at the same linear distance from the central pixel. Filters can also be designed as 'unbalanced' to highlight or reduce the effect of image features that have a directional component such as roads, lineaments or detector striping. Such filters basically rely on differential calculus, or derivatives, to achieve their result. Before considering the design of differential filters, we will review the meaning of the derivative in Excursus 5.1.

Excursus 5.1—Differential calculus review

From high school mathematics you may remember that the first derivative of a one-dimensional function, say

$$f(x)$$

is written as

$$f'(x) \text{ or } \frac{df}{dx}$$

In terms of image values, x could represent a row or column of pixels in an image that have the values $f(x)$ where these values can somehow be modelled in terms of pixel position.

For example, if the image values across an image line are as shown above the graph in Figure 5.1a, we could represent these as a graph (or a 'spectral transect'), then plot a continuous function line to describe the changes in image value in terms of position along the line. If we know the equation for this line, we could use differential calculus to determine the derivative equation, then compute a derivative value for each pixel along the line. For non-continuous functions, such as image data, these calculations can be approximated using filter-based processes

The initial function value, $f(x)$, describes the relationship between two variables—in the case of Figure 5.1a these are the pixel value and its position along a line. To determine the way values change along the line, we can look at the slope of the function line and identify where the values are maximum, minimum, increasing or decreasing. What we are really extracting from the graph is a visual measure of the change in the $f(x)$ value caused by a change in the x values. These observations may also be performed analytically using differential calculus.

Background image: High resolution aerial image of rail depot in Perth, WA, acquired on 5 March 2017. **Source:** © EagleView

Mathematically these changes can be expressed as the function:

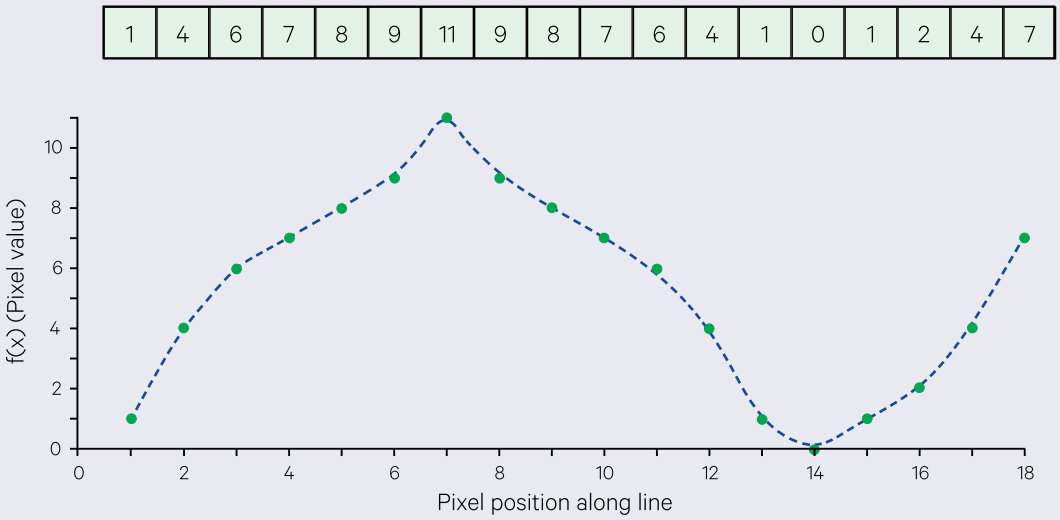
$$\frac{df}{dx} = f'(x) = \lim_{\delta x \to 0} \frac{\delta f}{\delta x}$$

where

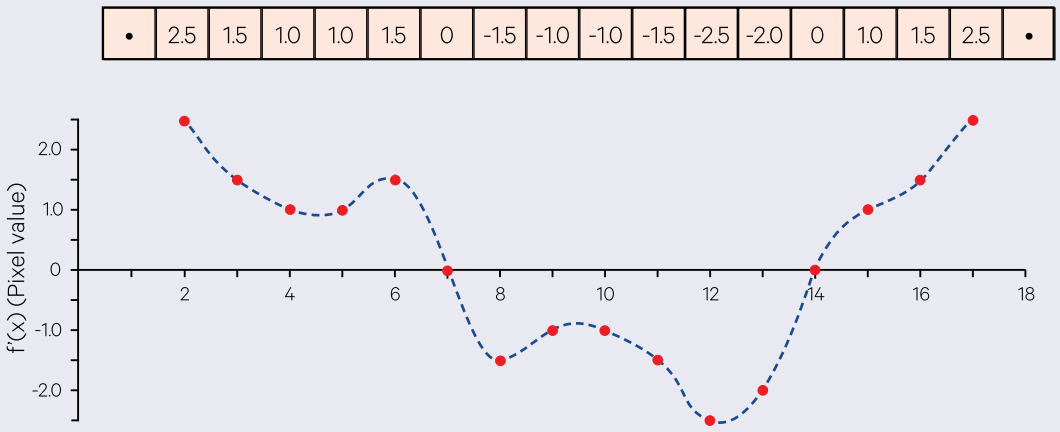
$\frac{df}{dx}$ is referred to as the derivative of x;
δx is the change in x values; and
δf is the change in $f(x)$ corresponding to δx.

Figure 5.1 First and second derivative for spectral transect

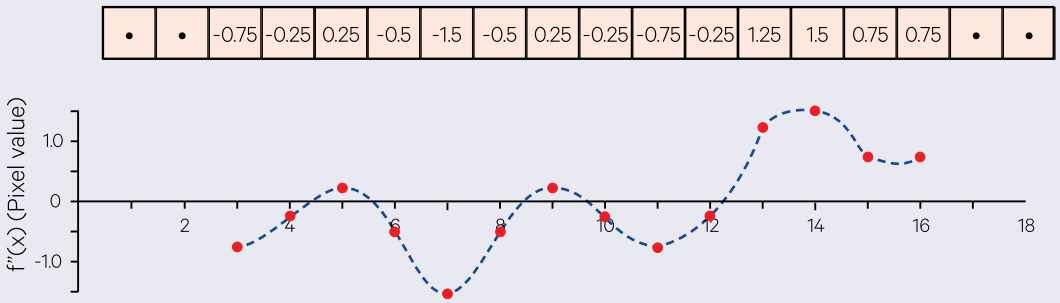
a. Pixel values across an image line are shown above the graph, in which the *X* axis represents pixel positions along an image line and the *Y* axis shows the pixel values.



b. Central difference values are shown and graphed below. The first derivative indicates the changes in slope of the original function.



c. Second derivative values indicate the direction and extent of curvature in the original function.



Source: Harrison and Jupp (1990) Figure 77

When the function is defined at points that are a constant distance apart, a number of operators may be used to measure its rate of change. For a forward difference, that is, the difference between values as x increases:

$$\Delta_F f(x) = f(x+h) - f(x)$$

where h is the δx, that is the number of pixel positions, or lag, between values. For a backward difference:

$$\Delta_B f(x) = f(x) - f(x-h)$$

or for the central difference (the average of the forward and backward differences):

$$\Delta_C f(x) = \frac{f(x+h) - f(x)}{2} + \frac{f(x) - f(x-h)}{2} = \frac{f(x+h) - f(x-h)}{2}$$

These difference formulae give different results where h is not equal to, or close to, 0. The derivative function defines the way the function changes as h approaches 0. Based on forward difference, the derivative function becomes:

$$f'(x) = \lim_{\delta x \to 0} \Delta_F \frac{f(x)}{h}$$

Similarly, the 'backward derivative' is:

$$f'(x) = \lim_{\delta x \to 0} \Delta_B \frac{f(x)}{h}$$

while the derivative computed from the central difference is defined as:

$$f'(x) = \lim_{\delta x \to 0} \Delta_C \frac{f(x)}{h} = \lim_{\delta x \to 0} \frac{f(x+h) - f(x-h)}{2h}$$

For a smooth (or 'differentiable') function all these operators give the same result.

Using the example in Figure 5.1a, we can compute and plot the derivative values for h=1 as illustrated in Figure 5.1b. This is done using the central difference filter to measure the slope of the function. Where we have increasing values in the original function, the derivative values are greater than zero and the slope is positive. Decreasing values in the original function give derivative values less than zero and so the slope is negative. Minimum and maximum values in the original function are where slope changes between positive and negative and so have zero values for the derivative. Thus the derivative function has quantified the relationship between slope and changes (increasing, decreasing, maximum and minimum) in the original function.

The second derivative can be obtained by applying the central difference filter again but to the derivative. The result is the slope of the derivative function, or 'slope of slope', and is closely related to the curvature of the original function. If line sections in the original function have constant slope (linear functions) over the section then the slope of slope is zero so there is no curvature. If sections in the original function have gradually increasing slope, then slope of slope is positive and sections with gradually decreasing slope in the original function have a negative slope of slope. Where the slope changes from gradually increasing to decreasing, or decreasing to increasing, the slope of slope passes through zero and slope (or derivative) is at a local minimum or maximum. Finally, when the slope is zero, the original function is at a local maximum or local minimum. In this case, if the slope of slope is positive it is a minimum and if the slope of slope is negative it is a maximum.

These observations about the value and slope of the derivative function are important for interpreting our original function as they highlight some of the more subtle changes. Since they all relate to the slope of the derivative function, they may be simply quantified by computing the second derivative or $f''(x)$.

For our example, the second derivative values and plot are as shown in Figure 5.1c. This function clearly identifies the direction and extent of curvature of the original function:

- positive values in $f''(x)$ show increasing slope in $f'(x)$ and increasing rate of change in $f(x)$;
- negative values in $f''(x)$ show decreasing slope in $f'(x)$ and decreasing rate of change in $f(x)$;
- stationary, or local maximum and minimum, points in f(x) have zero f'(x) and (respectively) negative and positive values for f''(x); and
- points of inflection in $f(x)$, that is where the sign of the curvature changes, are stationary points in $f'(x)$ and equal to zero in $f''(x)$. In Figure 5.1c, these occur between pixel values.

Just as the first derivative highlights positive and negative slopes and stationary points in the original function, more subtle changes, such as points of inflection, are significantly easier to detect in the second derivative.

5.1.1 Image-based derivatives

In image data, derivative functions may be simply derived along rows or columns in the image or along defined transects such as diagonals. These derivatives can then be used to indicate the rate of spatial variation in the defined direction. Where a dataset has more than one spatial dimension, the derivatives are computed with respect to one dimension at a time and are termed partial derivatives. For image data with two spatial dimensions, the terminology frequently used is X for across lines and Y down pixel columns in the image, with:

- f_x or $\frac{\partial f}{\partial X}$ being the first partial derivative function computed in the X direction for a given Y value (that is, an image line); and
- f_y or $\frac{\partial f}{\partial Y}$ being the first partial derivative in the Y direction for a given X value (that is, an image pixel column).

When the image data represent samples from a smooth underlying function (such as geophysical data), the derivative function may be approximated using filters. For example, the difference (or error) between the value for the central difference used as an approximation to the derivative and $f'(x_i)$ can be shown (Conte and de Boor, 1965) to be:

$$\left| f'(x_i) - \frac{\left(f(x_{i+1}) - f(x_{i-1})\right)}{2h} \right| = \frac{h^2}{6} \times \left| f'''(y) \right|$$

where $x_{i-1} \le y \le x_{i+1}$. That is, when the step h is small, and the underlying function is smooth (hence the third derivative is also small), the central difference can be a very close approximation to the actual derivative.

Figure 5.2 Examples of derivative filters

a. This filter computes the derivative across a line.

-1/2	0	1/2

b. This filter computes the derivative down a line of pixels.

-1/2
0
1/2

The derivative in X and Y can also be used to compute the derivative in a particular direction θ (where θ is measured anti-clockwise from the X axis) as:

$$f_s = f_x \times \cos\theta + f_y \times \sin\theta$$

$$s = x \times \cos\theta + y \times \sin\theta$$

(see Section 6.3). The formation of a directional difference uses the same equation with differences for derivatives.

To calculate the derivative values across a line in an image for h=1 pixels, we can compute the difference between pairs of pixels that are separated by one pixel. This spatial difference can be easily calculated using a filter with the weights shown in Figure 5.2a. Similarly, to compute derivative values down columns in the image, we could use the filter shown in Figure 5.2b. The results of applying some simple derivative filters are illustrated in Figure 5.3.

Figure 5.3 Effect of filter direction on differential filters

The original image is shown in Figure 5.3a. The vertical striping patterns visible in the horizontally and diagonally filtered images are due to slight scanner miscalibration which is not discernible in the orginal image. Differential filters accentuate this type of image 'noise' as well as feature boundaries.

a. Original image

b. Horizontal filter

c. Vertical filter

d. Diagonal filter

Source: Harrison and Jupp (1990) Plate 5

A filter with the weights shown in Figure 5.4a also uses the spatial variation of the two adjacent image columns to compute the central value. This filter is formed from the two one-dimensional filters shown in Figure 5.4b, so is equivalent to carrying out the two operations of smoothing and differentiating. Since the differential highlights the differences between pixel values, the use of adjacent pixel values to smooth the image values before differencing usually produces better results.

We can effectively vary the *h* value used for the derivative by varying the filter size. For example, for *h*=2 we could use the filter shown in Figure 5.5a. Similarly, the direction for the filter may be expressed in filter weights. For example, a 'diagonal' filter is shown in Figure 5.5b, where *a* is √2. The '*a*' term accounts for the longer distance to the centres of diagonally adjacent pixels from the centre pixel as compared with those that are in the same row or column. The differential filter weights are usually chosen so that the result is 'exact' for some function, such as a linear function. This implies that the weights sum to zero so that the value is zero when applied to constant data values and the filter produces constant values when applied to a linear function (that is, one with constant slope separately in *X* and *Y*). In practice though, if the '*a*' value equals 1 in this filter, the image scaling would be preserved satisfactorily.

Second derivative values may be computed by re-applying the same filter to a previously filtered channel (or by convolving the two filters as described in Section 3.1). Mixed derivatives may also be computed at this stage by applying the *Y* filter to the first partial derivative in *X* or the *X* filter to the first partial derivative in *Y* to incorporate two-dimensional information.

The second order partial derivatives are:

$$f_{xx} = \frac{\partial}{\partial X}\left(f_x\right) = \frac{\partial}{\partial X}\left(\frac{\partial f}{\partial X}\right) = \frac{\partial^2 f}{\partial X^2}$$

$$f_{xy} = \frac{\partial}{\partial X}\left(f_y\right) = \frac{\partial}{\partial X}\left(\frac{\partial f}{\partial Y}\right) = \frac{\partial^2 f}{\partial X \partial Y}$$

$$f_{yx} = \frac{\partial}{\partial Y}\left(f_x\right) = \frac{\partial}{\partial Y}\left(\frac{\partial f}{\partial X}\right) = \frac{\partial^2 f}{\partial Y \partial X}$$

$$f_{yy} = \frac{\partial}{\partial Y}\left(f_y\right) = \frac{\partial}{\partial Y}\left(\frac{\partial f}{\partial Y}\right) = \frac{\partial^2 f}{\partial Y^2}$$

This set of second partial derivative values can also be treated as a matrix and manipulated to determine the direction and magnitude of the maximum curvature within the filter region. A larger filter size, such as 7×7 should be used to compute surface curvature (see Section 6.4).

Figure 5.4 Combined smoothing and differentiating filter

a. This filter combines smoothing and differentiation.

-1/6	-1/6	-1/6
0	0	0
1/6	1/6	1/6

b. These two one-dimensional filters were used to create the filter above.

-1/2
0
1/2

×

1/3	1/3	1/3

Figure 5.5 Horizontal and diagonal derivative filters

a. Horizontal filter

-1/4	0	0	0	1/4

b. Diagonal filter (a=√2)

-1/2a	0	0
0	0	0
0	0	1/2a

c. First difference

1	-1

The application of these various formulae to edge detection has been widely implemented using expressions for edge magnitude (*M*) and edge angle (θ) (Nixon and Aguado, 2002), where in the most usual case, writing $diff\left(x\right)$ and $diff\left(y\right)$ for the derivatives or differences in *X* and *Y*:

$$M = \sqrt{diff\left(x\right)^2 + diff\left(y\right)^2}$$

$$\theta = \tan^{-1}\left(\frac{diff\left(y\right)}{diff\left(x\right)}\right)$$

By thresholding *M*, areas of potential edge are defined and these are usually post-processed by thinning and line joining to define edges and boundaries (Paine and Lodwick, 1989).

The edge magnitude can be viewed as a distance function. Various distance functions (or 'norms', denoted $\|\cdot\|$), can be defined. Common expressions for norms are (Moik 1980; see Figure 5.6):

$$\|\cdot\|_2 = \sqrt{diff(x)^2 + diff(y)^2}$$

$$\|\cdot\|_1 = |diff(x)| + |diff(y)|$$

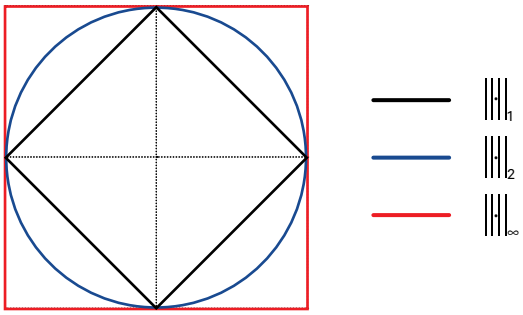
or

$$\|\cdot\|_\infty = \max\left[|diff(x)|, |diff(y)|\right]$$

Norms and distance functions are further discussed in Volume 2E. Some commonly used differential filters are detailed in Excursus 5.2

Figure 5.6 Norm functions

These can be used to estimate the maximum directional difference from two orthogonal distance measurements (see Volume 2E for details).



Source: Harrison and Jupp (1990) Figure 78

Imagination reveals the possibilities beyond the edges of our reality.
(L.R. Knost)

Excursus 5.2—Commonly used differential filters

The Roberts cross operator (Roberts, 1965) uses diagonal first differencing in two directions with the filters shown in Figure 5.7, and computes the magnitude and direction of the gradient using the equations given above.

Figure 5.7 Roberts cross operator

0	1
-1	0

and

0	1
-1	0

Figure 5.8 Sobel operator

-1	0	1
-2	0	2
-1	0	1

and

-1	-2	-1
0	0	0
1	2	1

Figure 5.9 Prewitt operator

-1	0	1
-1	0	1
-1	0	1

and

-1	-1	-1
0	0	0
1	1	1

The Sobel operator (Levine, 1985) uses a similar approach to the Roberts operator with the two differential filters shown in Figure 5.8. The Prewitt operator (Prewitt, 1970) varies these weights slightly as shown in Figure 5.9. These filters incorporate local averaging to reduce noise effects so tend to produce superior results to the Roberts operator for 'sharp' edges. The use of zero weights along the 'edge' is reported to reduce jitter (Levine, 1985).

A second difference, which approximates the second derivative, can also be computed as:

$$diff_2(x) = \left(f(x_{i+1}) - f(x_i)\right) - \left(f(x_i) - f(x_{i-1})\right)$$
$$= f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)$$

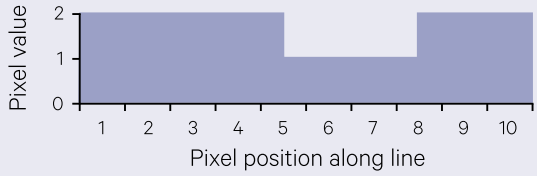
or with the horizontal filter shown in Figure 5.10. As illustrated in Figure 5.11, the first difference measures the change between two adjacent pixels whereas the second difference indicates the extent of change across a pixel. Images produced by second difference filters highlight edges as adjacent light and dark lines, whereas first difference filters result in single pixel edges.

Figure 5.10 Second difference filter

1	-2	1

Figure 5.11 First and second differences

a. Original image values ($f(x)$)



b. First forward difference ($x_i - x_{i+1}$)

c. First backward difference ($x_i - x_{i-1}$)

d. First difference backward across pixel i ($x_{i+1} - x_{i-1}$)

e. Second difference ($x_{i-1} + x_{i+1} - 2x_i$)

Source: Harrison and Jupp (1990) Figure 79

Figure 5.12 Laplacian edge detector

a. Omni-directional edge detector

-1	-1	-1
-1	8	-1
-1	-1	-1

b. Diagonal filter

a	0	a
0	-4a	0
a	0	a

Figure 5.13 Bi-directional filter

0	1	0
1	-4	1
0	1	0

=

1	-2	1

+

1
-2
1

An alternative edge detector is the Laplacian operator (an approximation to the mathematical Laplacian: $f_{xx} + f_{yy}$) which combines second differences in X and Y directions. This omni-directional operator can be implemented with the filter weights shown in Figure 5.12.

This operator is sometimes referred to as the centre-point 2-ring (or nine point) filter in geophysics. Another commonly used bi-directional filter, which is also referred to as a Laplacian operator (or the centre-point 1-ring filter), is shown in Figure 5.13. This basic filter can also be rotated to highlight diagonal edges with the weights shown in Figure 5.12b, where $a = \sqrt{2}$, or convolved to operate on a larger image neighbourhood (see Section 3.1). Some specific derivative filters which were developed for geophysical data and other directional filters are introduced in Volume 2X—Appendix 5.

There are no lines in nature, only areas of colour, one against another.
(Edouard Manet)

5.1.2 Uses of edge detection

Edge detection techniques are useful for a number of EO applications. For example, the effects of using horizontal and vertical filters on landscape surfaces are illustrated in Figure 5.14. Ton *et al*. (1989) outline a method for automatic road identification, which involves the use of edge detection filters to locate roads and categorise their 'status'. Paine and Lodwick (1989) report a multi-stage approach for extracting edges in an image to identify region boundaries that involves smoothing, edge detection, thresholding, thinning and linking. The ultimate result from this type of processing would be a set of polygons that are suitable for use in a Geographic Information System (GIS; see Volume 2D).

Jacobberger (1988) reports an interesting application of directional filtering to map abandoned river channels in Landsat TM imagery. Linear dunes in the study area obscured the morphology and trend of the channel system so directional filtering was used to selectively suppress edges parallel to the dune system. In this case the filter was based on a Laplacian operator (see Excursus 5.2) that was scaled to the width and orientation of the dune edges in the image. The first principal component (PC1—see Section 9) was filtered, then the residual between PC1 and the filtered channel was used for mapping. This residual image enhanced the high frequency detail that is not related to the dune system at the expense of detail that is parallel to it.

Filters that effectively compute higher partial derivatives also exponentially enhance image noise to the power of the order of the derivative. Noisy images should be appropriately pre-processed to reduce noise effects (as discussed in Sections 2.4 and 4.5) before edge detection filtering.

Differential or derivative filters may be applied in most image processing systems, possibly using selected weightings (defined as integer or real values) in an average filter for a selected filter size (see Section 4.1)

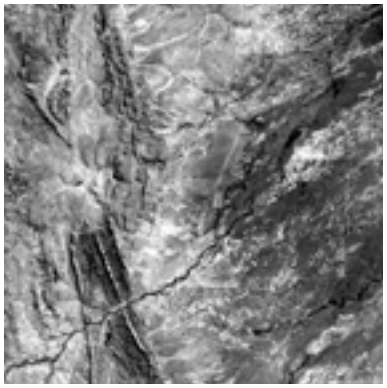
As described above, using smoother differences and thresholding an edge magnitude function are normally used to identify edges in noisy images. If the threshold is set too high, low-amplitude edges may not be detected but a low threshold will result in noise being highlighted. Edges in a multi-channel image may be summarised into a single channel by averaging the values across relevant channels (see Section 7) or by using Principal Components Analysis (PCA; see Section 9). Density slicing of edge channels can also be used to separate edge information from other image features (see Volume 2A—Section 9.2.1).

In earlier applications of edge detection, it was popular to use the magnitudes of the Prewitt or Sobel operators and thresholding edge magnitude to define edges. More recently, edge detectors based on more symmetric Gaussian functions and methods that align to the underlying edge more precisely have been used. The Canny detection operator (Canny, 1986) is probably the most popular of these. This operator and variants of it, as well as related applications to template matching and motion detection, can be found in Nixon and Aguado (2002).

Figure 5.14 Horizontal and vertical filters

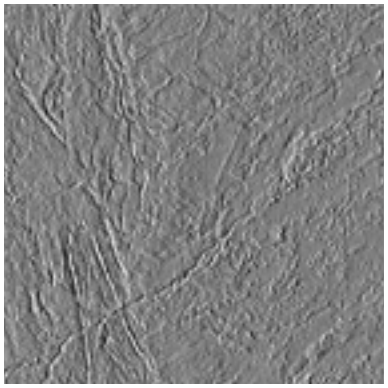
Landsat TM band 3 image of Fowlers Gap Research Station, NSW, acquired in January 2009.

a. Original image



b. After filtering with horizontal filter

c. After filtering with vertical filter



Source: Megan Lewis, University of Adelaide

The interpretation of derivative results depends on the image data being processed. For example, with elevation data the magnitude of the first derivative represents slope (or gradient) and the second derivative indicates surface convexity or curvature (see Excursus 5.1 and Section 6.4). Some image processing packages can compute slope and aspect for an image channel containing elevation data, and typically a 5×5 filter is used to estimate the local derivatives. A linearly scaled slope (or 'gradient') channel may be scaled with the minimum image value indicating a slope of 0° and the maximum value representing a slope of 90°. Aspect (which can also be interpreted as edge direction) may be linearly scaled from north at the minimum value through the mid-range value for an orientation of 180° to the maximum value corresponding to 359°. These results are further discussed in Section 6.4.

5.2 Edge Enhancement (High Pass)

Many image analysis tasks require that the features in an image be clearly defined in terms of spatial patterns and boundaries. For example, accurate positioning tasks require that linear features, such as roads, rivers or various boundaries, be clearly delineated in the image (see Volume 2B). Geological applications, in particular, frequently use edge enhancement methods to highlight lineaments. Image filtering transformations can be used to enhance the edge-related detail and also to identify edges at different scales. Such filters are also referred to as high pass or residual filters since they separate the high frequency variations in imagery from the lower frequency variations.

Section 3.1 and Figure 3.4 introduced the operation of image smoothing and showed that a smoothed image could effectively be defined as:

smoothed image = original image – edges

or

edges = original image – smoothed image

Conversely, an edge enhanced image could be described as:

edge enhanced image = original image + edges

or

edge enhanced image = original image + (original image – smoothed image)
= 2 × (original image) – smoothed image

Using some elementary arithmetic on the original and smoothed image values shown in Figure 5.15, we can use the relationship defined by the last equation to derive the weights for a simple edge enhancement filter.

The effect of edge enhancement using this filter, and varying its size, can be seen in Figure 5.16. This example filter has a typical pattern of weighting values for edge enhancing, with the central value being large and positive and surrounding values being small and negative. The effect of the filter is to exaggerate any difference between the central pixel and those pixels adjacent to it. The operation of this filter may be best understood using example image values as shown in Figure 5.17.

Figure 5.15 Derivation of edge enhancement filter

Original				Original				Smoothed				Edge enhanced							
0	0	0	+	0	0	0	–	1/9	1/9	1/9	=	-1/9	-1/9	-1/9	or	-1	-1	-1	with divisor = 9
0	1	0		0	1	0		1/9	1/9	1/9		-1/9	17/9	-1/9		-1	17	-1	
0	0	0		0	0	0		1/9	1/9	1/9		-1/9	-1/9	-1/9		-1	-1	-1	

Figure 5.16 Effect of filter size on edge enhancement filters

a. Original Image | b. Filter size: 3×3 | c. Filter size: 7×7 | d. Filter size: 11×11

Source: Harrison and Jupp (1990) Plate 4

Various edge detection filters such as the Laplacian operator are described in Section 5.1. As introduced above, an edge enhancement filter can be simply developed by adding an edge detection filter to the original image filter (see Figure 5.18a). An effective edge enhancement filter involves subtraction of the first difference Laplacian from the original (the weightings of this Laplacian required subtraction rather than addition to produce a 'positive' image; see Figure 5.18b).

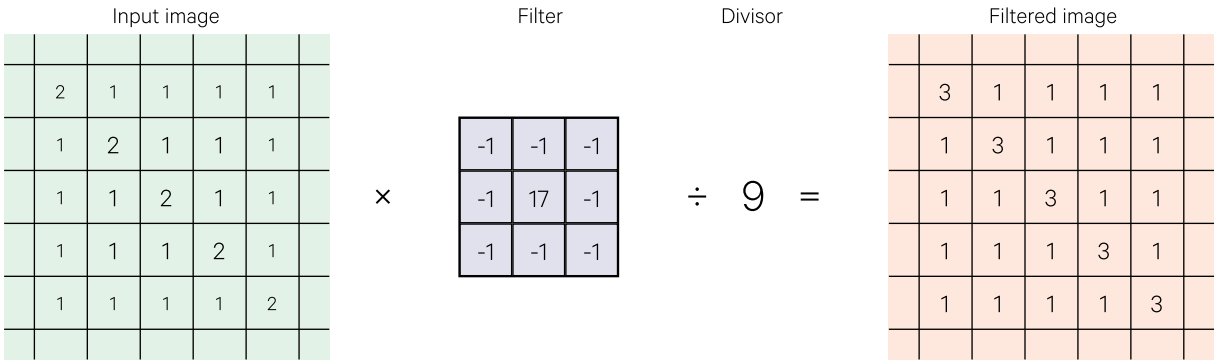
Niblack (1986) refers to this filter type as being similar to the photographic process of 'unsharp masking'.

Edge enhancement filters can similarly be derived from other edge detection filters, such as directional templates. Chavez and Bauer (1982) developed a method that uses the horizontal first difference to automatically select the best filter size to enhance edges in a particular image.

Edge enhancement transformations may be applied in most image processing systems with user-defined weights in an average filter. As with edge detection filters, thresholding or pre-processing to remove noise is often beneficial to avoid highlighting noise as well as edges.

Figure 5.17 Operation of edge enhancement filter

Differences between the central pixel and its neighbours are exaggerated using a filter that has a large central value with small edge values.



Source: Harrison and Jupp (1990) Figure 80

Figure 5.18 Examples of edge enhancement filters

a. Edge detection filter added to original image creates an edge enhancement filter.

Original				Edges				Edge enhanced		
0	0	0	+	-1/9	-1/9	-1/9	=	-1/9	-1/9	-1/9
0	1	0		-1/9	8/9	-1/9		-1/9	17/9	-1/9
0	0	0		-1/9	-1/9	-1/9		-1/9	-1/9	-1/9

b. Edge enhancement filter based on Laplacian filter requires subtraction from the original.

Original				Edges				Edge enhanced		
0	0	0	-	0	1	0	=	0	-1	0
0	1	0		1	-4	1		-1	5	-1
0	0	0		0	1	0		0	-1	0

5.3 Further Information

Castleman (1998) Sections 9, 11 and 12

Jensen (2016) Section 8

Gonzalez and Woods (2018) Section 3

Nixon and Aguado (2002) Section 4

5.4 References

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Chavez, P.S., and Bauer, B. (1982). An automatic optimum kernel size selection technique. *Remote Sensing of Environment,* 12, 23–38.

Conte, S.D., and de Boor, C. (1965). *Elementary Numerical Analysis*. McGraw-Hill Kogakusha Ltd, Japan.

Gonzalez, R.C., and Woods, R.E. (2018). *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jacobberger, P.A. (1988). Mapping abandoned river channels in Mali through directional filtering of TM data. *Remote Sensing of Environment,* 26, 161–170.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Levine, M.D. (1985). *Vision in Man and Machine*. McGraw-Hill Inc. USA.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images*. NASA SP–431. Washington, DC. USA.

Niblack, W. (1986). *An Introduction to Digital Image Processing*. Prentice-Hall International. New Jersey.

Nixon, M.S., and Aguado, A.S. (2002). *Feature extraction and image processing for computer vision*. Academic Press, Elsevier. 609 pp. *http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.375.6848&rep=rep1&type=pdf*

Paine, S.H., and Lodwick, G.D. (1989). Edge detection and processing of remotely sensed digital images. *Photogrammetria,* 43, 323–336.

Prewitt, J.M.S. (1970). Object Enhancement and Extraction. In *Picture Processing and Psychopictorics*. (Eds B.S. Lipkin and A. Rosenfeld). Academic Press, New York. pp 75–149.

Roberts, L.G. (1965). Machine perception of three-dimensional solids. In *Optical and Electro-optical Information Processing*. (Ed. J.T. Tippett). MIT Press, Cambridge, Mass. pp 159–197.

Ton, J., Jain, A.K., Enslin, W.R., and Hudson, W.D. (1989). Automatic road identification and labelling in Landsat TM images. *Photogrammetria,* 43, 257–276.

6 Highlighting Surface Variation

Local surface variations are called image 'texture' and can be characterised using a range of statistics (see Section 6.1). A variety of differential filters can also be adapted to accentuate surface variation. Such filters transform imagery to highlight:

- shaded relief (or Sun shadowing)—see Section 6.2;
- exposure to surface winds—see Section 6.3; or
- surface curvature—see Section 6.4.

6.1 Texture (Local Variance)

Image texture has been measured and characterised by a wide variety of methods (Haralick 1979; Levine 1985). Texture transformations use spatial statistics to compute the local standard deviation at each pixel. In this context, the square root of the local variance, is called 'texture' (see Volume 2A—Section 8.2.1). This statistic can be normalised by the mean of the pixel neighbourhood values to produce the 'coefficient of variation'.

A 3×3 filter is used to define the neighbourhood of each pixel. Within this filter region, the texture (T_{ij}) is computed by:

$$T_{ij} = \sqrt{\frac{1}{2n} \times \sum \left(x_{ij} - x_{kj} \right)^2}$$

where

x_{ij} is the central value of a 3×3 window centred at column *i* and row *j*;

x_{kl} is the pixel value at column *k* and row *l* where $k=i\pm1$ and $l=j\pm1$; and

n is the number of pixels involved in the calculation—usually 8 but sometimes less if there are null pixels within the 3×3 window.

Where $n=8$, this equation becomes:

$$T_{ij} = \sqrt{\frac{1}{16} \times \sum \left(x_{ij} - x_{kj} \right)^2}$$

This texture value can be normalised by $\overline{x}_{ij}$, the mean of the pixels in the filter neighbourhood, to give the Coefficient of Variation (C_{ij}) as:

$$C_{ij} = \frac{T_{ij}}{\overline{x}_{ij}}$$

The maximum texture value is generally difficult to estimate so is usually determined iteratively. (It can be estimated theoretically, using the minimum and maximum values for a given image channel, by computing the texture value of a 3×3 neighbourhood, where the central pixel has the minimum value and the surrounding pixels are equal to the maximum value.)

The texture transformation highlights areas of rapid spatial change, such as edges. In this sense, it is equivalent to the average slope of the image channel 'surface' over the filter region. It has also been used to categorise differences in the spatial structure of image features. For example, at the spatial resolution of a ~50 m pixel size, different forest canopies frequently radiate similar spectral values but can be separated by the difference in their spatial variability or 'texture'. To enhance vegetation texture, this transformation would be best applied to a NIR image channel or to a vegetation index channel (see Section 11 and Volume 3A).

Background image: Astronaut photograph of Gosses Bluff crater, in southern Northern Territory, acquired from the International Space Station.
Source: NASA ISS Photo retrieved from Wikimedia Commons: *https://commons.wikimedia.org/wiki/File:ISS007_Gosses_Bluff.jpg*

Figure 6.1 Different texture filters

a. Original image b. Texture c. Coefficient of variation d. Edge preserving texture

Source: Harrison and Jupp (1990) Plate 5d to 5f

This spatial analysis is similar to visually considering feature texture in aerial photography interpretation. The coefficient of variation effectively reduces the typically high texture values in bright areas of an image. Pickup and Foran (1987) observed that the coefficient of variation in multi-temporal imagery provides a measure of the effect of plant cover changes on spectral variability within a landscape. As with other spatial statistics, texture results are dependent on both the radiometric and spatial resolutions of an image (see Volume 1B—Section 1). The image texture values can also be used to create an image in which local variance (or standard deviation) is 'constant' (see Excursus 6.1).

As an edge enhancement transformation, texture analysis accentuates boundaries between any spectrally coherent patches (see Figure 6.1). The effect of this enhancement can highlight noise as well as edges. In most cases texture analysis on a smoothed image will reduce this effect.

Edge enhancement effects can be reduced with texture filtering by using adaptive windows (Woodcock and Ryherd, 1989—see Section 3.5). As illustrated in Figure 3.5, the adaptive window technique selects the 'best' sub-window to use within a filter region to derive the filtered value. In texture filtering, if the sub-window is selected on the basis of minimum variance, the resulting texture value is more likely to relate to a feature patch in the image than to an edge between two features.

Adaptive windows can be used to compute either smoothing or texture measures for an image. The texture results are generally selected from either the:

- minimum variance (based on the mean);
- deviation (associated with the median); or
- entropy (related to the mode)

of values within a sub-window of each filter region. The minimum variance result is illustrated in Figure 6.1, labelled as the 'Edge preserving texture'.

In conjunction with image classification, texture results are best used in a post-classification stage to split up specific, spectrally-defined classes (see Volume 2E). Unlike most classification algorithms, which are pixel specific, texture analysis considers the neighbouring pixel values so can add new information to the classification process. However, if texture channels are used to define the classification, a large number of classes is likely to be generated, as the boundaries themselves will tend to form separate classes. Gordon and Philipson (1986) suggest the use of ratioing texture-enhanced channels to reduce this characteristic edge brightness.

Each man is always in the middle of the surface of the earth
and under the zenith of his own hemisphere,
and over the centre of the earth.
(Leonardo da Vinci)

Excursus 6.1—High Pass Filtering

Source: David Jupp, CSIRO

Image texture values can be used to create an image in which the local variance (or standard deviation) is 'constant'. This type of enhancement has been variously referred to as 'constant variance' (Harris 1977), 'adaptive contrast' (Levine 1985) or 'statistical differencing' (Niblack 1986) and results in even contrast over the whole image. The general transformation is defined as:

$$\frac{x_{ij} - \overline{x}_{ij}}{\sigma_{ij}}$$

where

x_{ij} is the image value at pixel i and line j;
$\overline{x}_{ij}$ is the mean value of the neighbourhood of x_{ij}; and
σ_{ij} is the standard deviation of the neighbourhood of x_{ij}.

Figure 6.2a illustrates the results of applying this transformation to the image shown in Figure 6.1a using a range of filter sizes. To return this result to the original image data range, it can be multiplied by the mean of the local standard deviations, then added back to the local neighbourhood mean ($\overline{x}_{ij}$; see Figure 6.2b). These images illustrate the play off between high pass filtering and local variance equalising. Such differences will be data dependent, and in this example they are small when the high pass filtered and equalised image is added back to the original image. The high pass images lend themselves to further blur and stretching enhancements.

This enhancement is particularly effective for grey-scale imagery with a high dynamic range but little local variation. Medical imagery such as X-rays can be clarified using this technique.

Figure 6.2 Constant variance images

The scanned photograph shown in Figure 6.1a has been transformed using high pass filtering.

a. High pass filtering using equation above with four different filter sizes that are equivalent to 1%, 2.5%, 5% and 10% of the image width (from left to right)

b. High pass filtering using the same filter sizes with rescaling back to the original image data range.

Source: David Jupp, CSIRO

6.2 Relief Shading

As introduced in Volume 1, the term 'Sun shadowing' describes the relief shading produced on the Earth's surface by a particular Sun position. Relief shading is a useful transformation for elevation data as it highlights geomorphological patterns. When elevation data are registered with EO imagery, the relief shading corresponding to the time of image acquisition can be computed to help reduce topographic shading effects in the image.

Relief shading can be computed from relief data such as a Digital Elevation Model (DEM—see Volume 2D) if:

- a reflectance model is assumed for the terrain; and
- the Sun position can be defined in terms of:
  - altitude (or elevation)—angle above the horizon; and
  - azimuth (angle clockwise from north).

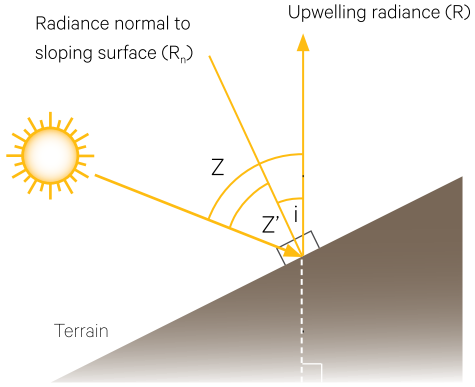
A range of surface reflectance models have been devised for various purposes (see Volume 1X—Appendix 5). For example, the simple cosine reflectance model defines the radiance (R) of a Lambertian surface (see Volume 1A—Section 5 or Volume 1B—Section 3) as:

$$R = R_n\left(\cos i\right)$$

where

R is the radiance of a target surface;
R_n is radiance viewed in a direction perpendicular to the target surface; and
i is the angle between upwelling radiance and the normal to the surface (see Figure 6.3).

Figure 6.3 Cosine reflectance model



Source: Harrison and Jupp (1990) Figure 83

This effectively models surface radiance for:

- a Sun position that is normal to the surface; and
- a sensor viewing the surface from a position directly overhead.

The modelled radiance can be modified for other Sun positions using the topographic modulation factor:

$$\frac{\max\left(\cos Z', 0\right)}{\cos\left(Z\right)}$$

where

Z is the zenith angle of the Sun (the angle from the directly vertical position), and
Z' is the angle between the Sun vector and the normal to the slope (Ahmad *et al.*, 1989).

As discussed in Section 5.1, the first partial derivatives of an elevation surface can be used to define the direction and magnitude of the slope. The 'downhill' direction of a slope vector is defined as the two-dimensional vector:

$$\begin{bmatrix} -f_x \\ -f_y \end{bmatrix}$$

In a three-dimensional space, a vector can be described in terms of the angles it makes with each of the X, Y and Z axes. This is commonly represented as the direction cosines (that is, the cosines of each of these angles). The vector which is normal (vertically orthogonal) to the slope is described by the direction cosines l, m and n where:

$$l = \frac{p}{\sqrt{1+p^2+q^2}}$$

$$m = \frac{q}{\sqrt{1+p^2+q^2}}$$

$$n = \frac{1}{\sqrt{1+p^2+q^2}}$$

and

$$p = f_x$$

$$q = f_y$$

The Sun vector can similarly be defined as:

$$l_s = \frac{p_s}{\sqrt{1+p_s^2+q_s^2}}$$

$$m_s = \frac{q_s}{\sqrt{1+p_s^2+q_s^2}}$$

$$n_s = \frac{1}{\sqrt{1+p_s^2+q_s^2}}$$

where

$$p_s = \tan Z \times \sin\theta$$

$$q_s = \tan Z \times \cos\theta$$

and

Z is the Sun zenith angle; and
θ is the Sun azimuth angle.

The cosine of the angle between two vectors can be computed as the scalar product of their direction cosines, that is:

$$\cos Z' = \begin{bmatrix} l & m & n \end{bmatrix} \begin{bmatrix} l_s \\ m_s \\ n_s \end{bmatrix}$$

$$= l \times l_s + m \times m_s + n \times n_s$$

The cosine of the angle between the Sun and the 'normal to the slope' vectors (that is Z') can thus be derived as:

$$\frac{p \times p_s + q \times q_s + 1}{\sqrt{1+p^2+q^2} \times \sqrt{1+p_s^2+q_s^2}}$$

Relief shading (or Sun shadowing) is then computed as:

$$\frac{R \times \cos Z'}{\cos Z}$$

This gives a relative value between 0 and 1 for $0° \leq Z \leq 90°$ and $Z' > 0$, or 0 for any other angles.

Some image processing systems compute relief shading, slope and aspect for an elevation channel using the cosine reflectance model for the terrain (see Figure 6.4). A 5×5 filter (detailed in Volume 2X—Appendix 5.2.2) may also be used to estimate local derivatives (see Section 5.1). The first partial derivative results can be used to define the principal gradient for the filter region at each pixel, that is, the vector that is perpendicular to the dominant direction of contour lines within the region. The slope of the gradient corresponds to the length of this vector and may be computed as:

$$\sqrt{f_x \times f_x + f_y \times f_y}$$

The direction of the gradient vector, or aspect, can also be computed from the partial derivatives as:

$$\tan^{-1}\left(\frac{-f_x}{-f_y}\right)$$

Relative relief shading can be computed as detailed above with f_x and f_y of the specified image channel defining the slope vector and user-selected Sun azimuth and zenith angles[4] defining the Sun vector. To compute relief shading parameters, the geometry of the input image needs to be defined in terms of scaling in the:

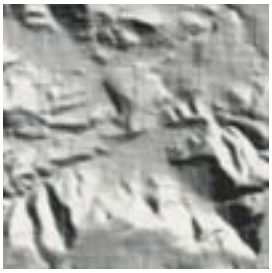
- X dimension (pixel width);
- Y dimension (pixel depth); and
- Z dimension (height scaling—the relationship between actual height and the values coded in the elevation channel); as well as the
- orientation relative to north.

Figure 6.4 Parameters derived from elevation

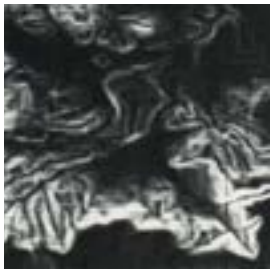
Based on a digitial elevation model, the relief shading, slope and aspect can be computed as image channels.

a. Elevation

b. Relief shading



c. Slope



d. Aspect

Source: Harrison and Jupp (1990) Plate 13

4 The zenith angle is 90° minus the Sun altitude angle.

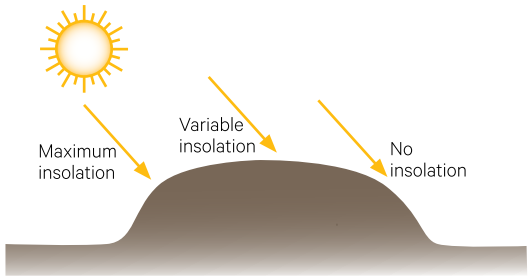
An atmospheric correction factor may also be specified to account for absorption and backscatter, which would reduce the measurement of surface radiance in conditions of haze or cloud.

Insolation, 'that amount of solar radiation received on a given surface in a given time period' (Landsberg and Sands, 2011), is a significant parameter in environmental physics and ecology (see Figure 6.5). Relief shading can serve as a surrogate measure for insolation to indicate the radiance that would be measured from a surface with uniform land cover for given parameters of Sun position and atmospheric transmittance conditions.

In conjunction with EO-based analyses, topographic data allow investigations of the relationship between spectral data and surface morphology (see Volume 2D). For example, registration between an image-based land cover classification and elevation data allows quantification of the occurrence and co-occurrence of classes at different topographic locations. Tabular summaries of such comparisons may be produced automatically in most image processing systems (see Volume 2E). Alternatively, land cover categories may be stratified by topography or morphology. The derivation of slope information also allows elevation categories to be stratified by slope classes for various management purposes.

Figure 6.5 Total insolation

Maximum insolation occurs when the surface is perpendicular to incoming radiation. For a flat surface, insolation depends on Sun position and atmospheric transmittance.



Source: Harrison and Jupp (1990) Figure 84

The best way to study Mars is with two hands, eyes and ears of a geologist, first at a moon orbiting Mars ... and then on the surface.
(Buzz Aldrin)

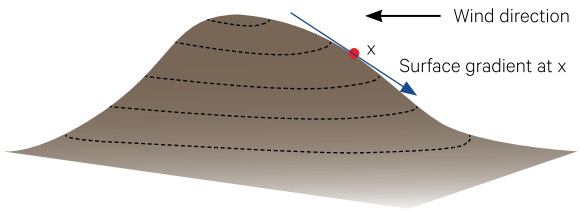
6.3 Exposure (Directional)

A special implementation of the differential filter is the exposure transformation. This transformation indicates the relative exposure for each pixel in an image channel 'surface' to a user-defined direction as illustrated in Figure 6.6. The 'exposure' result is similar to relief shading with the 'Sun position' effectively on the horizon.

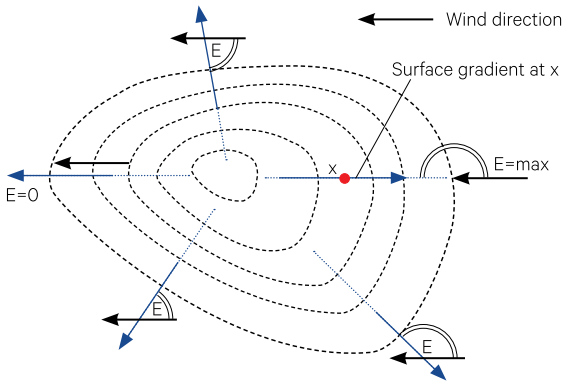
This process is relevant to shallow water mapping applications where parameters for water depth and prevailing wind direction can be used to derive a value indicating the degree of exposure of the sea floor at each pixel to that wind direction. In shallow water mapping exercises visible blue or green channels from EO imagery can be used as a surrogate for water depth (Jupp *et al.*, 1985). Directional parameters are typically expressed either in degrees, as a compass bearing clockwise from north, or in terms of *X* and *Y* components using polar coordinates, as illustrated in Figure 6.7 (see also Volume 2B).

Figure 6.6 Exposure transformation

a. Perspective view



b. Plan view (E=exposure value)



Source: Harrison and Jupp (1990) Figure 81

An exposure transformation uses two smoothing differential filters over the neighbourhood of each pixel. Typical weights for these filters are shown in Figure 6.8, with both filters having a divisor of 12. The exposure at each pixel is then derived from the equation:

$$E = -\left(\sin\theta \times f_x + \cos\theta \times f_y\right)$$

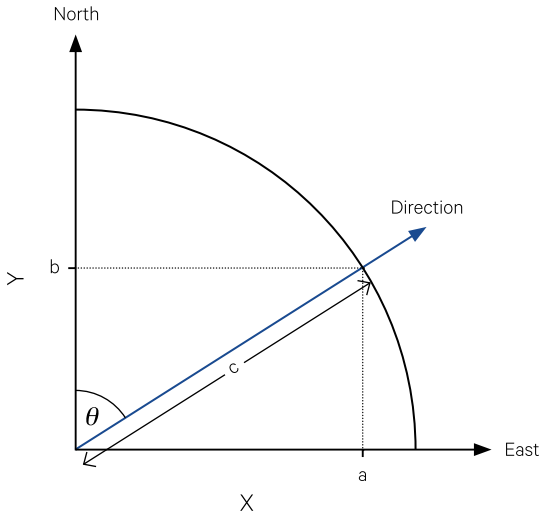
where

- $\sin\theta$ represents the *X* component of the compass direction;
- f_x is the first partial derivative in the *X* direction;
- $\cos\theta$ represents the *Y* component of the compass direction;
- f_y is the first partial derivative in the *Y* direction; and
- f is the image data, such as water depth (assumed to be oriented to north).

(Note: the form of this equation differs from the gradient direction equation given in Section 6.2. This difference is due to the way the compass direction is defined. Here the direction is given as clockwise from north whereas in the context of relief shading the direction is effectively measured as anti-clockwise from east.)

Figure 6.7 Representing compass direction as X and Y components

Since $\sin\theta$ = a/c and $\cos\theta$ = b/c, for c=1, $\sin\theta$ = a (X component) and $\cos\theta$ =b (Y component), where θ measures degrees clockwise from north.



Source: Harrison and Jupp (1990) Figure 82

A non-linear scaling of the resulting exposure value *E* may be used to increase contrast in low image values, in which case, the final exposure image values are derived from the product of the scaled exposure component and the input image value at each pixel. For best results, images should be destriped and despiked before computing the exposure transformation as differential filters will enhance image noise such as striping patterns or spikes (see Sections 2.4 and 4.5). Some algorithms incorporate a dark value to indicate the data level corresponding to atmospheric noise (see Section 10.2.1 for more details about dark value selection). An expected maximum value may also be requested for rescaling purposes. This is best determined iteratively, but the 99% maximum is usually satisfactory. The resulting exposure channel may be displayed as a grey-scale and contrast-enhanced, or used to modify the brightness values of a colour image to indicate surface relief (see Section 8.5).

As illustrated in Figure 6.9, the exposure transformation has proved to be valuable for mapping sub-surface morphology in coral reef studies. In this example, comparable exposure images have been derived from both Landsat MSS imagery (green band) and a sparse grid of water depth samples. In the exposure channel, bright areas indicate a higher local slope and aspect to the selected wind direction, and in the prevailing wind direction these areas indicate greater exposure. Darker areas infer sheltered positions relative to that wind direction. Thus, by using the relevant range of wind directions, this transformation allows a set of 'impact images' to be iteratively compiled for any suitably imaged location.

Figure 6.8 Filter weights for exposure transformation

These filters both have divisors of 12 to retain image scaling.

f_x

-1	0	1
-4	0	4
-1	0	1

f_y

1	4	1
0	0	0
-1	-4	-1

Figure 6.9 Exposure (directional) filtering

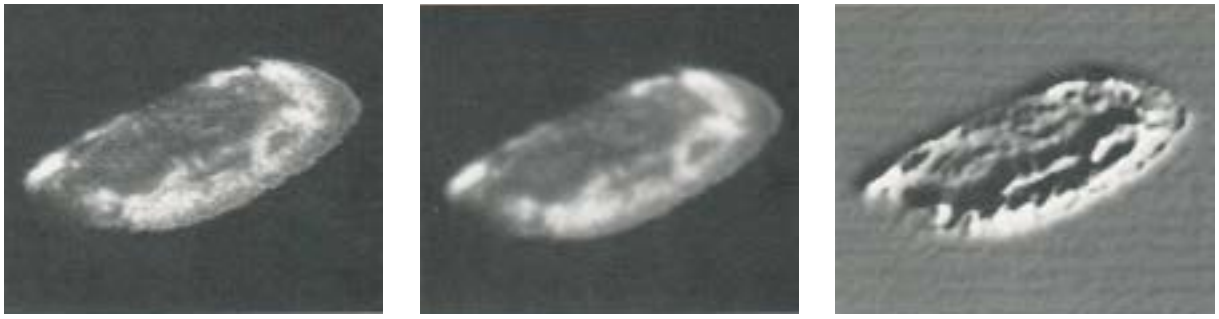
Comparable images have been derived for John Brewer Reef, Queensland, by directional filtering of Landsat MSS imagery and depth data. The resulting images highlight reef morphology and show exposure to the prevailing wind direction (southeast).

a. Transect samples of actual water depth formed a sparse grid of points when registered with Landsat MSS imagery. (Note: null pixels are shown as black.) This sparse dataset was interpolated (see Section 4.6) to create a continuous depth image before exposure (directional) filtering.

Real depth data | Interpolated depth data | Exposure on interpolated depth

b. A Landsat MSS4 (green band) image for John Brewer Reef was despiked and smoothed before directional filtering to show exposure to the prevailing wind direction. An exposure image derived from EO data offers an economical approximation to sourcing real depth data.

Landsat MSS4 image | Preprocessed MSS4 image | Exposure on preprocessed MSS4



Source: Harrison and Jupp (1990) Plate 6

The ability to predict the exposure of sub-surface relief to prevailing weather conditions provides valuable information to coastal zone managers in formulating management priorities for different sites. Spatial variations in the distribution of plant and animal species are usually strongly influenced by the wind and wave energy regimes that are operative in the littoral and sub-littoral zones (see Volume 3B).

As a general directional filter, the exposure transformation can be applied to topographic surfaces, magnetics and gravity data, or any data where a directional spatial pattern can be observed. The resulting channel is similar to relief shading (see Section 6.2) but does not produce any 'black' shadow areas on slopes away from the 'Sun'. The exposure transformation has also been used to enhance nighttime thermal imagery for structural interpretation.

Sometimes I enjoy just photographing the surface
because I think it can be as revealing as going to the heart of the matter.
(Annie Leibovitz)

6.4 Surface Curvature (Surface Shape)

The use of differential calculus to quantify the changes in curvature of a two-dimensional surface, such as a DEM channel, was introduced in Section 5.1, with:

- the first derivative representing the rate of change, or slope, of the surface; and
- the second derivative representing the change in slope, or curvature.

In two-dimensional data, these changes are expressed as partial derivatives, that is, two derivative values are computed separately indicating changes in the two dimensions. These second order statistics are invariant to rotation and slope, and describe intrinsic curvature properties of the data surface (Smith, 1983). Such mathematical analyses may be applied to image data, using filtering techniques with appropriately defined weights, to determine the curvature in both across line (X) and down column (Y) directions for each pixel in the image.

As detailed in Section 5.1, four second-order partial derivatives can be computed for two-dimensional data to quantify the changes in surface curvature:

$$f_{xx} = \frac{\partial^2 f}{\partial X^2}$$

$$f_{xy} = \frac{\partial^2 f}{\partial X \partial Y}$$

$$f_{yx} = \frac{\partial^2 f}{\partial Y \partial X}$$

$$f_{yy} = \frac{\partial^2 f}{\partial Y^2}$$

These derivatives can be treated as a matrix to determine the direction and magnitude of maximum curvature within the region. When such results are derived from elevation data, they indicate the relative size and orientation of the major landform features within an image filter region.

Volume 2X—Appendix 6 describes matrix algebra techniques which may be used to convert the 2×2 matrix of second partial derivatives of:

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

to the form:

$$\begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$$

where

k_1 is called the first eigenvalue and indicates the magnitude of a special vector in the direction of maximum data variation, called the first eigenvector; and

k_2 is the second eigenvalue, which indicates the magnitude of the second eigenvector, this being defined as orthogonal to the first eigenvector.

In terms of elevation data, the first eigenvector of the second partial derivative matrix is aligned with the direction of curvature of the major landform feature (within a filter region at an image pixel), while the corresponding eigenvalue indicates the relative magnitude of curvature of the feature. The second eigenvalue then indicates the extent of curvature in the direction at 90° to the first eigenvalue.

Such results can then be used to categorise different landform features in elevation data, or in fact the surface curvature of any two-dimensional dataset such as an image channel. In terms of image processing then, the curvature values of each pixel describe the surface changes within the image sub-area defined by the filter size used to derive them.

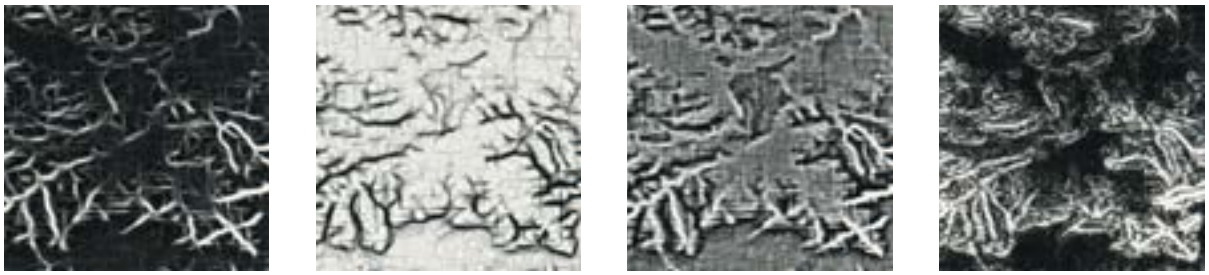
It is appropriate to compute the surface curvature of an image channel using a 7×7 filter region (see Volume 2X—Appendix 5.2.3). As with relief shading (see Section 6.2), the input image geometry needs to be specified in terms of pixel width (X), pixel depth (Y) and the height scaling used to convert actual height to image values (Z). Using appropriate filters weights (see Volume 2X—Appendix 5.2.3 for examples), the following statistics could be computed as new image channels (see Figure 6.10):

- the first eigenvalue of the second derivative matrix (also referred to as λ_1) indicates the extent of surface shape in the direction of maximum shape change;
- the second eigenvalue of the second derivative matrix (also referred to as λ_2) indicates the extent of the surface shape change in the direction that is orthogonal to the maximum change direction in the 7×7 filter region;
- the Laplacian: $\frac{\lambda_1 + \lambda_2}{2}$ indicates the average extent or magnitude of the surface curvature within the filter region; and
- the Discriminant: $\frac{\lambda_1 - \lambda_2}{2}$ summarises the directionality of surface curvature within the filter region (such as a ridge or gully).

Figure 6.10 Parameters related to surface curvature

Using the digitial elevation model shown in Figure 6.4, surface curvature parameters can be computed as image channels. The Laplacian is the average of the first and second eigenvalues, while the discriminant is their difference. To compute this example, a 7×7 filter region was used with the filter weights specified in Volume 2X—Appendix 5.2.3.

a. First eigenvalue b. Second eigenvalue c. Laplacian d. Discriminant



Source: Harrison and Jupp (1990) Plate 15

The scaling of these statistics depends on the ranges specified for the expected minimum and maximum values. As with many transformations, the actual range of transformed values that will be relevant to an image can not be predicted so these range values are best determined iteratively.

When applied to elevation data these four parameters clearly differentiate the major landform features of ridges, gullies, hills, basins, saddles and flat ground within each 7×7 pixel region. The relationship between these landform features and the output image values are given in Table 6.1.

These surface curvature parameter channels could be used to form a colour composite. If the first eigenvalue is displayed as blue, the Laplacian as green and the Discriminant as red, ridges would show as well-connected red lines, gullies as pink lines, undulating ground as aqua and flat ground as green (see Figure 6.11). It should be noted that the category 'flat' only applies to surface curvature within the filter region, not surface orientation, so that a uniformly sloping area with no change in surface shape within the filter would be described as 'flat'. A composite display of slope as blue, the Laplacian as green, and Discriminant as red provides enhanced contrast in the undulating ground. In geophysics, the Laplacian statistic is also called the second vertical derivative.

Table 6.1 Curvature image categories

λ_1 is first eigenvalue and λ_2 is second eigenvalue.

Surface Curvature Category	Curvature Image Values			
	λ_1	λ_2	Laplacian	Discriminant
Flat ground	~0	~0	~0	small
Basin (bowl shape)	>>0	>>0	>>0	small
Hill (cap shape)	<<0	<<0	<<0	small
Saddle	>>0	<<0	~0	large
Gully line	>>0	~0	>>0	large
Ridge line	~0	<<0	<<0	large

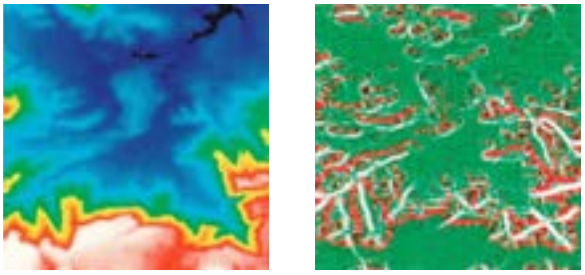
These statistics may be derived from either elevation data or appropriate EO imagery (Li *et al.*, 1986). Various methods have been proposed to isolate topographic shading (that is, effectively Sun shadowing) information from EO imagery (Ahmad *et al.*, 1989). Where land cover variation is relatively small, topographic shading is the major factor causing changes in image brightness and can be reasonably estimated as the first principal component (PC1; see Section 9). As with other analyses, Sun shadowing imagery that relates to a high Sun angle provides better discrimination of the landform features than is possible with low Sun angle imagery.

While the results of curvature analysis on Sun shadowing imagery will necessarily differ from those of elevation data, some interesting landform features are highlighted. At the very least, these results can be used to stratify the image into 'flat' and non-flat areas. Such stratifications then allow the analysis of land use and land cover features in the flat areas without influencing class definition by topographic brightness variations (see Volume 2E).

Figure 6.11 Curvature composite

This curvature composite image is derived from the digital elevation model shown in Figure 6.4 and used in Figure 6.10, here shown in pseudo-colour with blue indicating low elevation and red showing high elevation. The curvature composite image displays the first eigenvalue as blue, the Laplacian as green and the discriminant as red to highlight major landform categories.

a. Pseudo-coloured elevation b. Curvature composite



Source: Harrison and Jupp (1990) Plate 15

6.5 Further Information

Texture

Jensen (2016) Section 8

6.6 References

Ahmad, W., Jupp, D.L.B., and Nunez, M. (1989). Land cover mapping in a rugged terrain area using Landsat MSS data. *Int. J. Remote Sensing,* 13(4), 673–683. *https://doi.org/10.1080/01431169208904145*

Gordon, D.K., and Philipson, W.R. (1986). A texture-enhancement procedure for separating orchard from forest in Thematic Mapper data. *International Journal of Remote Sensing,* 7, 301–304.

Haralick, R.M. (1979). Statistical and Structural Approaches to Texture. *Proceedings of IEEE*, 67, 786–804.

Harris, J.L. Sr. (1977). Constant Variance Enhancement: A Digital Processing Technique. *Applied Optics,* 16, 1268–1271. *https://doi.org/10.1364/AO.16.001268*

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective.* 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Jupp, D.L.B., Mayo, K.K., Kuchler, D.A., Van R. Classen, D., Kenchington, R.A., and Guerin, P.R. (1985). Remote sensing for planning and managing the Great Barrier Reef of Australia. *Photogrammetria,* 40, 21–42.

Landsberg, J.J., and Sands, P.J. (2011). Weather and Energy Balance. Ch 2 in *Terrestrial Ecology Volume 4: Physiological Ecology of Forest Production.* Elsevier. doi: *10.1016/B978-0-12-374460-9.00002-0*

Levine, M.D. (1985). *Vision in Man and Machine.* McGraw-Hill Inc. USA.

Li, R., Jupp, D.L.B., Harrison,B.A., and Ahmad, W. (1986). Extraction of landform information from remotely sensed data: an initial study. *Proc*eedings of *Beijing International Symposium on Remote Sensing*, Beijing, China, 18–22 November, 67–73.

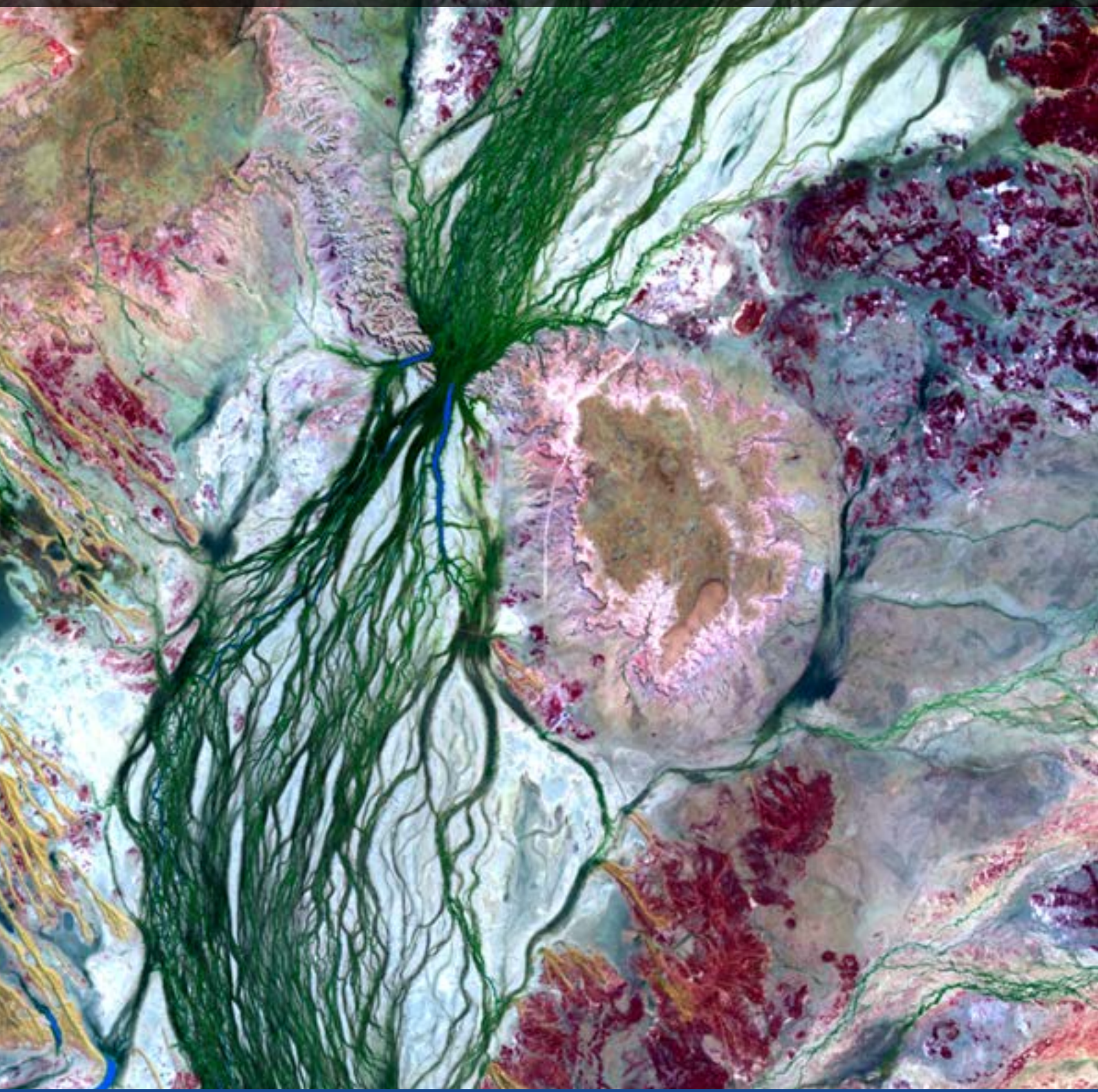
Niblack, W. (1986). *An Introduction to Digital Image Processing.* Prentice-Hall International. New Jersey.

Pickup, G., and Foran, B.D. (1987). The use of spectral and spatial variability to monitor cover changes on inert landscapes. *Remote Sensing of Environment,* 23, 351–363.

Smith, G.B. (1983). Shape from shading: an assessment. *Proc. NASA Symposium on Mathematical Pattern Recognition and Image Analysis*, June 1–3, 1983. Johnson Space Centre, Houston, Texas, 543–576.

Woodcock, C.E., and Ryherd, S.L. (1989). Generation of texture images using adaptive windows. *Tech. Papers, 55th Annual Meeting ASPRS*, 2, 11–22.

Linear Operations



The following three sections are concerned with techniques that apply linear operations to one or more image channels. Linear operations rely on matrix algebra and can implement simple operations (such as simple rescaling or inversion of individual channels) or powerful multidimensional manipulation of image data. Non-linear operations involve multiplication and division of channel values (see Section 1.4).

Most image processing systems offer a wide range of multi-channel transformations. Some of the most commonly available linear transformations involving multiple channels include:

- modifying image brightness (see Section 7);
- converting image colour space (see Section 8); and
- Principal Components Analysis (PCA; see Section 9);

Some image processing systems provide flexible 'band maths' options, whereby a sequence of linear and non-linear operations can be applied to selected sets of image channels. Such options offer significant flexibility, but need to be used intelligently with EO imagery to ensure that the transformed results are meaningful.

Contents

7 Affine Transformations 73

8 RGB-HSI Conversions 81

9 Principal Components Analysis 93

Background image on previous page: Landsat-8 OLI image of Diamantina Gates, southwestern Queensland, acquired in winter 2013. Due to constraints imposed by the adjacent Goyder and Hamilton Ranges, the Diamantina River narrows to about 1 km wide at Hunters Gorge, which is located near the centre of this image. **Source:** Craig Shephard, DSITI

7 Affine Transformations

An affine transformation is one that preserves points, straight lines and planes between the input and output datasets (see Volume 2B—Section 3.2.1). This linear transformation is defined by a multidimensional matrix plus an offset, or shift, and can be used to implement many image processing operations. Such techniques rely on matrix algebra, which is described in Volume 2X—Appendix 6.

In this section, we initially review matrix notation in Excursus 7.1, then Section 7.1 introduces the general affine transformation. Some simple uses of the affine transformation with EO imagery are described in Excursus 7.2, including:

- adding and subtracting image channels;
- averaging channels;
- inverting and rescaling channels;
- processing log-transformed channels; and
- implementing regression models.

These linear transformations can also be used for removing spatial 'noise' (see Section 7.2).

Affine transformations are useful for a range of image processing operations (see Section 7.3), including converting between colour spaces (see Section 8). Another commonly available linear transformation is Principal Components Analysis (PCA; see Section 9), possibly with the option to define selected transformation parameters and/or to rotate the transformed channels back to their original data space. A specific channel rotation operation, which is used to enhance vegetation greenness in EO imagery, is described in Section 11.1.3.

Life is a linear equation in which you can't cross multiply!
If you think you can do it, you can do it.
If you think you can't do it, you can't do it.
It's a simple formula!
(Isrealmore Ayivor)

Background image: Landsat-8 OLI image of Cubbie Station in southern Queensland, which is the largest irrigation property in the southern hemisphere.
Source: Craig Shephard, DSITI

Excursus 7.1—Review of Matrix Notation

Matrix operations allow new image channels to be defined as some linear combination of the original channels. For example, to compute the sum and difference of two channels, *X* and *Y*, we would use the equations:

$$X' = X + Y \text{ (or } X' = 1 \times X + 1 \times Y)$$

$$Y' = X - Y \text{ (or } Y' = 1 \times X + -1 \times Y)$$

In matrix notation, we assume that each new X' or Y' value can be computed as some weighted sum of the original *X* and *Y* values. This allows us to define a table of weighting values. For example, the general transformation equations:

$$X' = a \times X + b \times Y$$

$$Y' = c \times X + d \times Y$$

could be represented by their coefficients as summarised in Table 7.1. This tabulation is more generally represented as the matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Similarly, the matrix required for the initial sum and difference equations given above would be:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The general matrix operation in Table 7.1 is usually described as:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} X \\ Y \end{bmatrix}$$

where *X*, *Y*, X', Y', *a*, *b*, *c*, and *d* are defined above.

Table 7.1 Matrix coefficients

Output Value	Coefficients for Input Values	
	X	Y
X'	*a*	*b*
Y'	*c*	*d*

Such matrices can be used to transform imagery by computing new channels X' and/or Y' for each image pixel using its values in input channels *X* and *Y*. However, image data spaces usually have more than the two dimensions (channels) used in this example. For example, for a four channel image the general matrix operation is defined using a square four-dimensional matrix as shown in Figure 7.1.

These linear algebraic operations allow very flexible manipulation of the image data space. In algebraic terms, the operation is simply a linear combination of the image channel values for each pixel. However in geometric terms, a wide range of sophisticated manipulations of the data space may be performed. Guidelines for defining matrix coefficients for general reflection, rescaling, skewing and rotation operators on the image data space are detailed in Volume 2B—Section 3.2.1 and Volume 2X—Appendix 6.

For example, a simple geometric operation may be to reflect the data space about the line $Y = X$ as shown in Figure 7.2. In terms of coordinates referenced by the X' and Y' axes, the geometric operation can be summarised algebraically by the two simple equations:

$$X' = Y \text{ (or } X' = 0 \times X + 1 \times Y)$$

$$Y' = X \text{ (or } Y' = 1 \times X + 0 \times Y)$$

These equation can be represented by the matrix:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Figure 7.1 General matrix operation

$$\begin{bmatrix} \textit{output channel 1} \\ \textit{output channel 2} \\ \textit{output channel 3} \\ \textit{output channel 4} \end{bmatrix} = \begin{bmatrix} A & B & C & D \\ E & F & G & H \\ I & J & K & L \\ M & N & O & P \end{bmatrix} \times \begin{bmatrix} \textit{input channel 1} \\ \textit{input channel 2} \\ \textit{input channel 3} \\ \textit{input channel 4} \end{bmatrix}$$

or

output channel 1 = A × input channel 1 + B × input channel 2 + C × input channel 3 + D × input channel 4

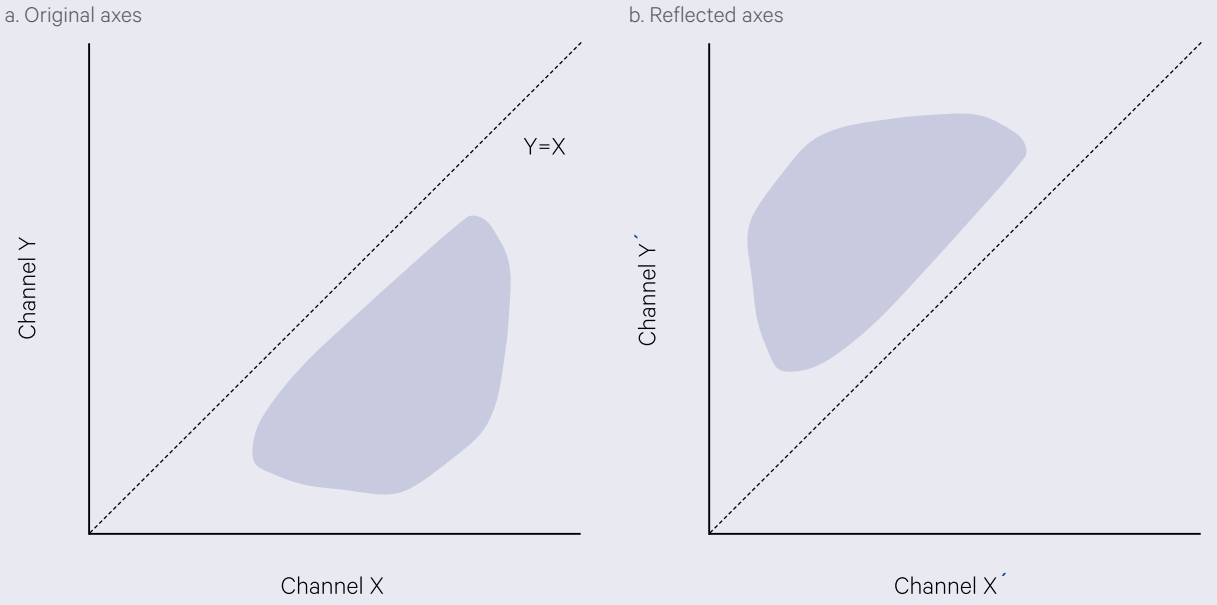
output channel 2 = E × input channel 1 + F × input channel 2 + G × input channel 3 + H × input channel 4

output channel 3 = I × input channel 1 + J × input channel 2 + K × input channel 3 + L × input channel 4

output channel 4 = M × input channel 1 + N × input channel 2 + O × input channel 3 + P × input channel 4

Figure 7.2 Reflecting image axes about the line *Y = X*

Simple matrix operations can be used to modify the geometry of an image data value space.



Source: Harrison and Jupp (1990) Figure 85

7.1 General Affine Transformation

Derivation of the linear transformation is introduced in Excursus 7.1 above. This operation allows a new channel to be created as a linear combination (or weighted sum) of the original image channels. Affine transformations are commonly used to define vegetation or soil brightness indices in EO datasets (see Section 11).

An affine transformation adds an offset value to the result of a linear transformation to form the matrix equation:

$$\begin{bmatrix} output \\ image \end{bmatrix} = \begin{bmatrix} transformation \\ matrix \end{bmatrix} \times \begin{bmatrix} input \\ image \end{bmatrix} + \begin{bmatrix} offset \end{bmatrix}$$

For a single channel image, this transformation computes the following for each pixel:

$$output\ value = matrix\ coefficient \times input\ value + offset\ value$$

In image processing, the offset allows the transformed values to be shifted back into the image data range. If a transformation is defined where one channel has a coefficient of one and all other coefficients are zero, the offset can be used to add or subtract a constant value to or from each pixel value in an image channel.

Weighted averaging of multiple channels in an image can be easily implemented using an affine transformation (see also Volume 2X—Appendix 6). For example, the equation:

$$output\ channel = a \times input\ channel\ 1 + b \times input\ channel\ 2 + c \times input\ channel\ 3$$

can be represented by the matrix equation:

$$\begin{bmatrix} output \\ channel \end{bmatrix} = \begin{bmatrix} a & b & c \end{bmatrix} \times \begin{bmatrix} input\ channel\ 1 \\ input\ channel\ 2 \\ input\ channel\ 3 \end{bmatrix}$$

The combined use of affine transformation, logarithmic or exponential transformation (see Section 2.2) and channel ratios (see Section 10) can produce most of the indices commonly encountered in EO literature (see Excursus 7.2).

Excursus 7.2—Using Affine Transformations

Adding and subtracting channels

An affine transformation can be used to perform simple channel arithmetic, such as addition or subtraction. (Addition and differencing of all channels in pairs of images are discussed in Volume 2D). These arithmetic operations are required to produce indices of particular image features. For example, a simple vegetation greenness index could be based on the difference between pixel values in NIR and red channels of an image (see Section 11):

$$Index\ value = NIR\ value - red\ value$$

This index could be computed using the affine transformation:

$$Index = \left(1 \times NIR\right) + \left(-1 \times red\right)$$

or

$$\left[index\right] = \begin{bmatrix} 1 & -1 \end{bmatrix} \times \begin{bmatrix} NIR \\ red \end{bmatrix}$$

with an offset equal to half the channel data range[5].

The popular Normalised Difference Vegetation Index (NDVI) involves ratioing this difference by the sum of the NIR and red channels:

$$NDVI = \frac{NIR - red}{NIR + red}$$

or

$$NDVI = \frac{\left(1 \times NIR\right) + \left(-1 \times red\right)}{\left(1 \times NIR\right) + \left(1 \times red\right)}$$

This can be implemented by computing the sum and difference of the two channels by affine transformation then ratioing the results (see Section 11.1.2). The affine transformation matrix required to compute this sum and difference would be:

$$\begin{bmatrix} difference \\ sum \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} NIR \\ red \end{bmatrix}$$

with offsets equal to half the data range for *difference* and 0 for *sum*[6]. The NDVI could then be computed as a ratio of the output difference channel (as the numerator) and the sum channel (as the denominator; see Section 11.1.2). NDVI has demonstrated effective and usable correlations with plant greenness metrics, such as biomass or leaf area index (see Volume 3A). It can saturate, however, in vegetation with high greenness, such as irrigated sports fields (Tony Sparks, *pers. comm.*).

Various channel combinations have been suggested to enhance and summarise image information relating to vegetation condition or soil brightness, such as the 'Tasselled Cap' transformations (see Section 11.1.3). Other specific transformations relating to vegetative cover and condition are considered in Section 11 and Volume 3A.

Averaging channels

The average of multiple channels may be required to summarise specific pattern information in an image, such as edges (see Section 5.1.2). For a three channel image, the average would be computed as:

$$average = \frac{channel\ a + channel\ b + channel\ c}{3}$$

and the required affine transformation matrix would be:

$$\left[average\right] = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \times \begin{bmatrix} channel\ a \\ channel\ b \\ channel\ c \end{bmatrix}$$

Inverting and rescaling channels

Channel data ranges can also be inverted and linearly rescaled using appropriate coefficients in affine transformations. For example, to invert the range of an image channel, for each pixel we can compute:

$$inverted\ channel = \left(channel\ value \times -1\right) + channel\ data\ range$$

This can be implemented as an affine transformation using -1 as the coefficient for the channel to be inverted plus a channel offset. For example, to invert channel *a* in a three channel image we would use the matrix equation shown in Figure 7.3. This transformation would only change values in channel a, leaving the values in channel b and channel c the same.

5 In many image processing systems, one image value is reserved to indicate 'null' pixels (see Volume 2A—Section 1.2), which effectively reduces the data range by one. With byte data, for example, the null value is often value 0 or 255, and the data range reduces to 255 instead of 256 values. In this case the offset would equal 127 rather than 128.

6 In many image processing systems, value 0 or 255 may be reserved in byte resolution imagery to indicate 'null' pixels, in which case the value 127 would be used as the offset.

Figure 7.3 Matrix equation for channel inversion

$$\begin{bmatrix} \text{inverted channel } a \\ \text{channel } b \\ \text{channel } c \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \text{channel } a \\ \text{channel } b \\ \text{channel } c \end{bmatrix} + \begin{bmatrix} \text{offset} \\ 0 \\ 0 \end{bmatrix}$$

Linear transformations on log-transformed channels

When an affine transformation is applied to log-transformed imagery (see Sections 1.2.3 and 2.2), channel addition and subtraction respectively represent multiplication and division of the original channels. For example,

$$\log\left(\text{channel } 1 \times \text{channel } 2\right) = \log\left(\text{channel } 1\right) + \log\left(\text{channel } 2\right)$$

$$\log\left(\frac{\text{channel } 2}{\text{channel } 3}\right) = \log\left(\text{channel } 2\right) - \log\left(\text{channel } 3\right)$$

$$\log\left(\left(\text{channel } 4\right)^3\right) = 3 \times \log\left(\text{channel } 4\right)$$

(In the case of channel ratios, dark values are implemented in this sequence during the logarithmic transformation; see Sections 2.2 and 10.2.1).

A complex non-linear combination such as:

$$\frac{\left(\text{channel } 1\right) \times \left(\text{channel } 2\right)}{\left(\text{channel } 3\right) \times \left(\text{channel } 4\right)^3}$$

would then be a linear concatenation of the previous operations. In terms of matrix coefficients, these transformations (with offsets equal to 0) would be represented as shown in Figure 7.4.

The exponentiation option, which should be available in most image processing systems (see Section 2.2), could then be used to convert the resulting channels back to a linear scaling.

Implementing regression models

Where a known linear relationship exists between image values and some ground parameter, such as crop yield, and can be expressed as a regression of the form:

$$\text{parameter} = a_i \times x_i + b_i$$

where

x_i is the image value in channel i;

a_i is the model coefficient for channel i; and

b_i is the model constant for channel i.

then a new channel, which represents a scaled value for the ground parameter, can be computed using the affine transformation:

$$\left[\text{parameter}\right] = \left[a_i\right] \times \left[\text{channel } i\right] + \left[b_i\right]$$

The regression technique described in Sections 1.3.4 and 7.2.1 can be used to determine the relationship between two image channels. One channel can then be rescaled to match the other using the regression coefficients and offsets as exampled in Section 7.2.1 (see also Volume 2D).

Figure 7.4 Example matrix equation

$$\begin{bmatrix} \log\left(ch1 \times ch2\right) \\ \log\left(ch2 / ch3\right) \\ \log\left(ch4^3\right) \\ \log\left(\frac{ch1 \times ch2}{ch3 \times ch4^3}\right) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 3 \\ 1 & 1 & -1 & -3 \end{bmatrix} \times \begin{bmatrix} \log ch1 \\ \log ch2 \\ \log ch3 \\ \log ch4 \end{bmatrix}$$

7.2 Removing Spatial 'Noise'

Spatial 'noise' can occur in EO imagery for a number of reasons, such as atmospheric haze, specular reflection or uneven illumination (due to topography or uneven lighting). These spatial variations are not constant over the image, so we cannot correct for them by simply subtracting a constant 'noise' level from each pixel.

Methods that may be used to adjust for these spatial variations include weighted channel subtraction (see Excursus 7.2) or channel ratioing (see Section 10). The choice of correction technique depends on the nature of the model that describes the interaction between surface reflectance and the image noise. Where the spatial noise is related to ground reflectance by a multiplicative model (that is the detected noise level is affected by target reflectance, such as with topographic shading) image ratios should be used. The models for atmospheric noise and sunglint (specular reflection from water) are linearly related to surface reflectance (that is the noise level is independent of target reflectance) so an affine transformation can be used to correct for these components in the image radiance values. That is:

$$x_{ij} = b_{ij} + d_{ij}$$

where

x_{ij} is image radiance value at pixel i in channel j ;
b_{ij} is the reflectance component of pixel i in channel j ; and
d_{ij} is the atmospheric noise and/or sensor calibration offset component.

7.2.1 Using a reference channel

This approach uses one image channel that contains the pattern as a 'reference' channel (in the case of scan-digitised imagery, this should always be specifically produced using a blank target) and requires that the reference channel be 'sacrificed' in any subsequent processing. Thus the reference channel is best selected from channels that would not normally be used. Assuming some constants α_j and β_j:

$$d_{ij} = \alpha_j \times d_{ir} + \beta_j$$
$$x_{ij} - \left(\alpha_j \times d_{ir} + \beta_j\right) = b_{ij} - \alpha_j \times b_{ir}$$

where

when $j=r$, $x_{ir} = d_{ir}$ is the image radiance value at pixel i in the reference channel r ; and
b_{ir} is the reflectance component of pixel i in the reference channel r. This reference channel is usually chosen so that this component is not significant for further processing.

In this equation, the result is free of d_{ij} and retrieves b_{ij} if $b_{ir} = 0$.

In an image to be used for water studies, for example, a haze pattern could be reduced by using an infrared (especially SWIR) channel as the reference channel, provided it contains the same haze pattern. Since the NIR and SWIR wavelengths are virtually totally absorbed by water, it is unlikely that they would be required for subsequent processing. Where several channels are available in this wavelength region, the first principal component of those channels could be used (see Section 8).

The relationship between each image channel and the reference channel can be determined by image crossplotting (see Section 1.3.4 and Volume 2A—Section 8.1.3). As illustrated in Figure 7.5, the regression line for each crossplot can then be computed as:

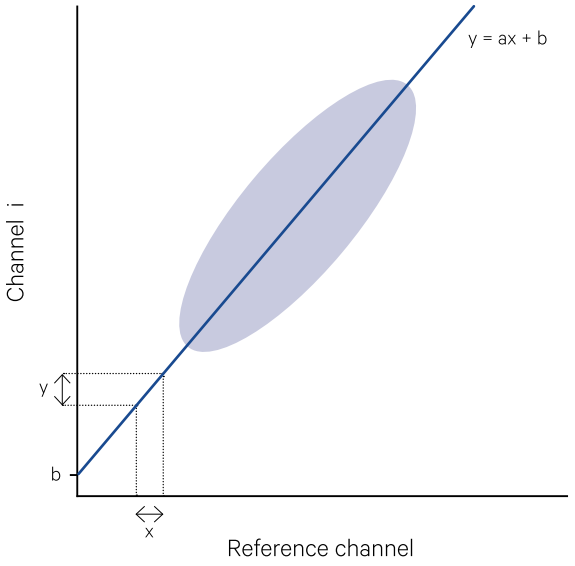
$$\text{channel i} = a_i \times \text{reference channel} + b_i$$

where

a_i is the regression coefficient
b_i is the regression offset.

Figure 7.5 Using channel crossplot to derive regression line with reference data

The regression equation summarises the relationship between two image channels as y = ax+b, where a = y/x.



Source: Harrison and Jupp (1990) Figure 86

Figure 7.6 Filter to remove image noise

$$\begin{bmatrix} \text{corrected ch 1} \\ \text{corrected ch 2} \\ \text{corrected ch 3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -a_1 \\ 0 & 1 & 0 & -a_2 \\ 0 & 0 & 1 & -a_3 \end{bmatrix} \times \begin{bmatrix} \text{input ch 1} \\ \text{input ch 2} \\ \text{input ch 3} \\ \text{reference ch} \end{bmatrix} + \begin{bmatrix} a_1 \times \overline{R} \\ a_2 \times \overline{R} \\ a_3 \times \overline{R} \end{bmatrix}$$

or

$$\text{corrected ch 1} = 1 \times \text{input ch 1} - a_1 \times \text{reference ch} + a_1 \times \overline{R}$$

$$\text{corrected ch 2} = 1 \times \text{input ch 2} - a_2 \times \text{reference ch} + a_2 \times \overline{R}$$

$$\text{corrected ch 3} = 1 \times \text{input ch 3} - a_3 \times \text{reference ch} + a_3 \times \overline{R}$$

The proportion of our image signal in channel *i* which is modelled as the spatial 'noise' can then be subtracted by the equation:

$$\text{channel } i' = \text{channel } i - a_i \times \text{reference channel}$$

To ensure that the scaling of channel i' is consistent with channel i (and that they have the same mean values), we usually add an offset of:

$$a_i \times \overline{R}$$

where

$\overline{R}$ is the mean of the reference channel.

In an affine transformation of a four channel input image, this model would be specified as shown in Figure 7.6 above.

The reference channel may need pre-processing for some image sources. For example, in scan-digitised imagery the blank target scan can have considerable local variation, so may need to be heavily smoothed (with a 5×5 or 7×7 average filter with all weights equal to 1—see Section 4 above) before being used for illumination correction. Other sources may require calibration to radiance units. To assess the impact of this correction, the transformed and original image pairs can be differenced (see Volume 2D).

7.2.2 Removing 'Limb Brightening'

'Limb brightening', a characteristic increase in brightness towards the edges of the image, commonly occurred in early airborne scanner imagery. This resulted from the wide scan angle, which caused edge pixels to be imaged through a significantly longer atmospheric path distance. This effect can be corrected using an affine transformation.

The brightness trend across the image can be estimated for each channel from the average brightness down each pixel column. The average brightness line may be represented in an image by duplicating the line to match the number of lines in the original image (see Volume 2A—Section 7.2.1.4). This is then subtracted from the original image (with coefficients of 1 and an offset equal to the mean of the original image) to produce a brightness corrected image :

$$\begin{aligned} \text{corrected channel} = {} & 1 \times \text{original channel} \\ & - 1 \times \text{average brightness line} \\ & + \text{original channel mean} \end{aligned}$$

To correct all channels in an image, this operation would be performed more efficiently using image differencing (see Volume 2D). The statistics associated with along-line brightness variation within an image can be accumulated in some image processing systems. For example, the mean intensity for each channel could be extracted to enable the correction described above.

During a transformation, it's understandable to be hesitant about moving too fast, especially at the outset.
(Irene Rosenfeld)

7.3 Other Affine Transformations

A range of affine transformations have been proposed for multivariate EO imagery to variously highlight data or noise components for particular purposes. These methods analyse the spectral and/or spatial correlation of one or more images and are particularly relevant to change detection studies (see Volume 2D). Examples of such transformations include:

- Principal Components Analysis (PCA; see Section 9)—rotates image channels such that the data are arranged along axes of decreasing variance;
- Canonical Variates Analysis (CVA) transformation matrix (see Volume 2E)—maximises the separation between class sample image values;
- Maximum Autocorrelation Factor (MAF)—derives new, orthogonal channels in which neighbouring pixels have maximum autocorrelation to highlight contiguous regions (Switzer and Green, 1984; Conradsen *et al.*, 1985; Nielsen *et al.*, 1998; Nielson, 2011). Autocorrelation statistics are derived from analysis of the dispersion matrix of the difference between each image channel and its spatially shifted self;
- Minimum/Maximum Noise Fraction (MNF)—orders transformed image components in terms of image quality (Green *et al.*, 1988; Nielsen, 2011); and
- Multivariate Alternation Detection (MAD)—analyses canonical correlations between pairs of image channels to highlight differences (Nielsen *et al.*, 1988; see Volume 2D). Nielsen (2007) describes an extension to this transformation, termed the Iteratively Reweighted MAD method (IR-MAD).

Factor analysis methods are further discussed in Volume 2E.

7.4 Further Information

Sawyer (1955)

7.5 References

Conradsen K., Ersboll B. K., Thyrsted T. (1985). A comparison of min/max autocorrelation factor analysis and ordinary factor analysis. *Nordic Symposium in Applied Statistics*, Lyngby. pp. 47-56.

Green, A.A., Berman, M., Switzer, P., and Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing,* 26(1), 65–74.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Nielsen, A.A., Conradsen, K. and Simpson, J.J. (1998): Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bi-temporal image data: new approaches to change detection studies. *Remote Sensing of Environment,* 64, pp. 1-19.

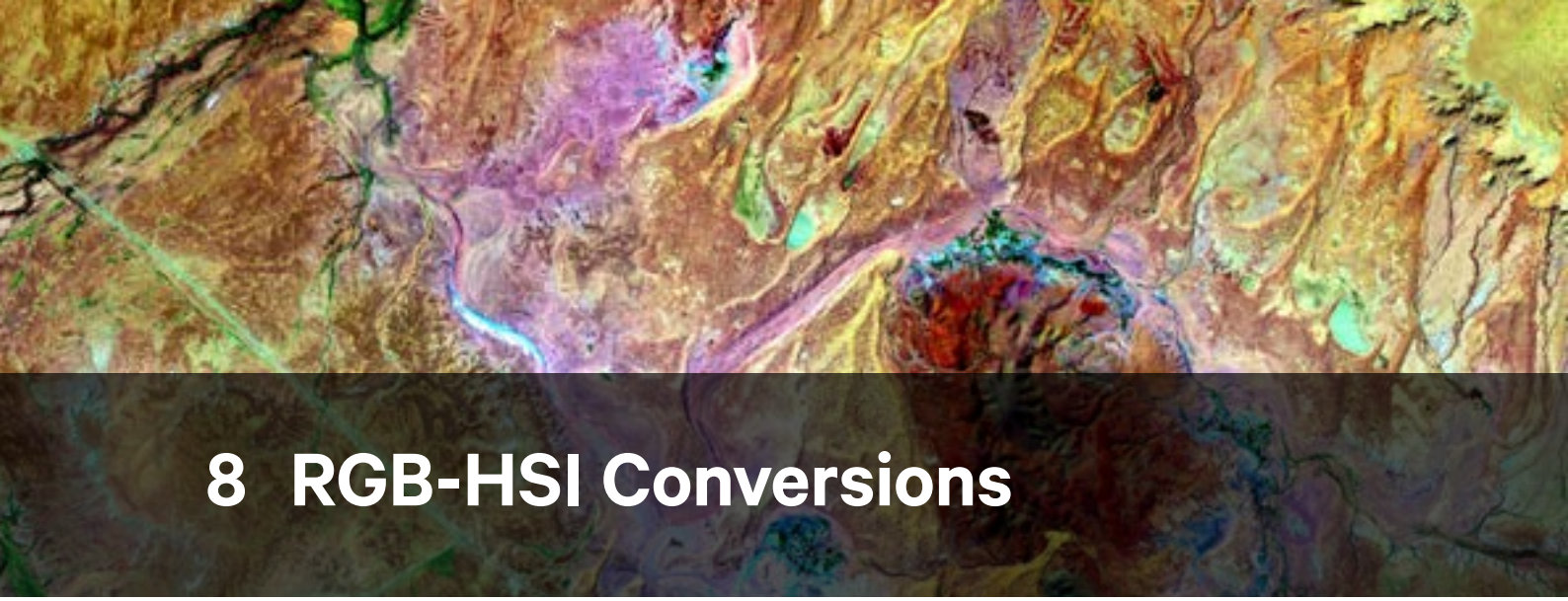
Nielson, A.A. (2007). The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. *IEEE Transactions on Image Proc*essing, 16(2), 463–478.

Nielson, A.A. (2011). Kernel Maximum Autocorrelation Factor and Minimum Noise Fraction Transformations. *IEEE Transactions on Image Proc*essing, 20(3), 612–624.

Sawyer, W.W. (1955). *Prelude to Mathematics*. Penguin Books Ltd. London.

Switzer, P., Kowalik, W.S., and Lyon, R.J.P. (1981). Estimation of atmospheric path-radiance by the covariance matrix method. *Photogrammetric Engineering and Remote Sensing,* 47, 1469–1476.

Switzer, P., and Green, A.A. (1984). Min/max autocorrelation factors for multivariate spatial imagery. *Technical Report 6,* April 1984, Dept. Statistics, Stanford.



8 RGB-HSI Conversions

The RGB (Red, Green, Blue) and HSI (Hue, Saturation, Intensity) models for colour specification were introduced in Volume 2A—Section 5.1.2. These models are related by their intensity axes and can be interconverted using affine transformations.

In this section, after reviewing the relationship between the RGB and HSI colour systems (see Sections 8.1 and 8.2), we consider some transformations that modify image colour by manipulating these colour models:

- enhancing chromaticity (see Section 8.3);
- merging higher spatial resolution panchromatic data with multispectral data (see Section 8.4); and
- combining topographic shading with colour imagery for relief effects (see Section 8.5).

By isolating the intensity information in the image to a single channel, various transformations can be used to modify image intensity before converting the HSI colour values back to the RGB colour space. Two examples of transformations that use the HSI colour system to enhance image appearance are given in Excursus 8.1.

8.1 RGB Colour Model

In the Red-Green Blue (RGB) colour cube, the intensity axis is defined by the line:

$$R=G=B$$

as illustrated in Figure 8.1. Colours in the RGB colour space can also be defined as triangular coordinates on equal intensity planes across the cube (see Figure 8.2a). These coordinates can be normalised within each plane such that the values along each axis from the triangle centre to an apex vary from 0–1 (see Figure 8.2b).

Cartesian coordinates in the RGB cube can be converted to normalised triangular coordinates by the equations:

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

By definition, these normalised coordinates sum to 1. Since, for any colour r+g+b=1, only two components need to be known to deduce the third (see Figure 8.2c). For example at the white point, $r=1/3$ and $g=1/3$, so:

$$b=1-r-g$$
$$=1/3$$

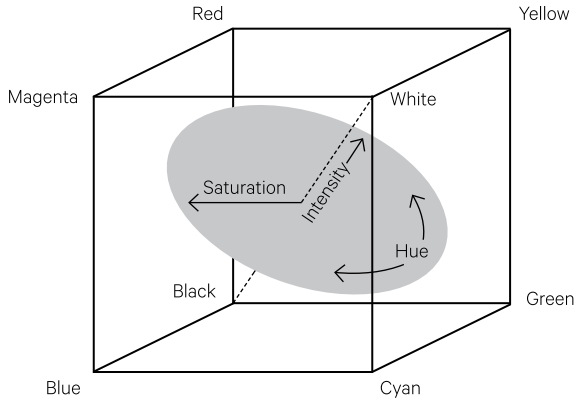
We are like chameleons, we take our hue and the color of our moral character from those who are around us.
(John Locke)

Background image: HyMap airborne hyperspectral image mosaic over Hermit Hill, South Australia, displayed using the 2.22 µm as red, the 0.833 µm band as green, and the 0.44 µm band as blue. **Source:** Megan Lewis, University of Adelaide

Figure 8.1 RGB colour cube

The colours formed by differing proportions of the additive primary colours, red green and blue, can be represented as a three-dimensional data space. The diagonal axis in this space, referred to as the intensity axis, represents the shades of grey, from black to white, which have equal proportions of each primary colour. On planes orthogonal to the intensity axis, all colours have equal intensity, with saturation increasing away from the axis and hue varying around it.

a. Orientation of RGB colour cube

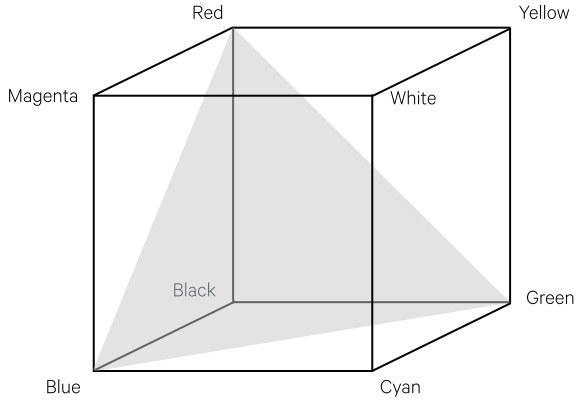


b. Colour planes on surface of RGB colour cube

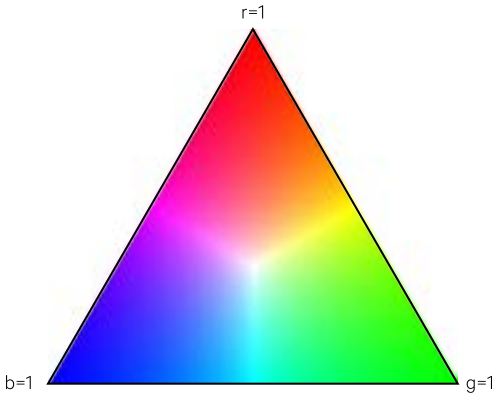
Source: a. Harrison and Jupp (1990) Figure 11; b. Frank Clark. Retrieved from: *http://www.frank-t-clark.com/Professional/Papers/ColorHCW/ColorHCW.html*)

Figure 8.2 RGB triangular coordinates

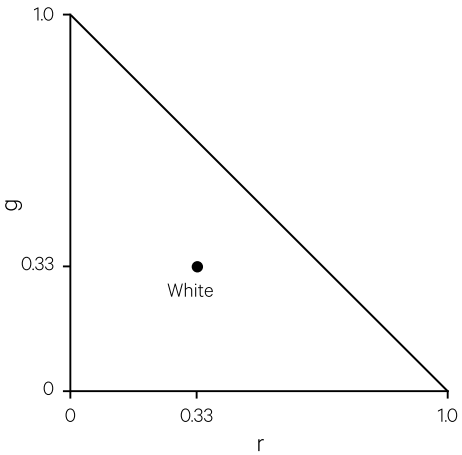
a. Definition of triangular planes in RGB cube



b. Normalised triangular coordinates



c. Normalised chromaticity coordinates. Since, for any colour r+g+b=1, only two components need to be known to deduce the third. For example at the white point, r=**1/3**, g=**1/3**, so b=1-r-g=**1/3**.



Source: Harrison and Jupp (1990) Figure 102

8.2 HSI Colour Model

The Hue-Saturation-Intensity (HSI or HSV: Hue-Saturation-Value or HSB: Hue-Saturation-Brightness[7]) model defines colours on the equal intensity planes as coordinates of hue and saturation, with hue being measured as an angle around the plane and saturation as the radial distance from the centre of the plane (see Figure 8.3). This colour model can be useful to vary the relationships between individual colours, and is a common feature in image processing systems for enhancing photographic images. In terms of EO imagery, it is also a valuable approach for substituting intensity information before visual interpretation.

To measure hue, we define a hue angle of 0° radiating from the centre to the blue apex. The hue angle can then be defined in degrees in an anti-clockwise direction from the blue radial as:

$$H = \cos^{-1}\left(\frac{2b-g-r}{\sqrt{6\left(b-\frac{1}{3}\right)^2+\left(g-\frac{1}{3}\right)^2+\left(r-\frac{1}{3}\right)^2}}\right)$$

or

$$H = \cos^{-1}\left(\frac{0.5\times\left((b-g)+(b-r)\right)}{\sqrt{(b-g)^2+(b-r)(g-r)}}\right)$$

On a triangular plane, colours of equal saturation are represented as concentric triangles and are measured as:

$$S = 1-3\times\min(r,g,b)$$

HSI cylindrical coordinates can also be computed from the RGB Cartesian coordinates (see Volume 2B) as:

$$\text{Intensity} = R+G+B$$

$$\text{Hue} = \tan^{-1}\left(\frac{v_2}{v_1}\right)$$

$$\text{Saturation} = \sqrt{v_1^2+v_2^2}$$

where

$$v_1 = \frac{1}{\sqrt{6}}R+\frac{1}{\sqrt{6}}G-\frac{2}{\sqrt{6}}B$$

$$v_2 = \frac{1}{\sqrt{2}}R-\frac{1}{\sqrt{2}}G$$

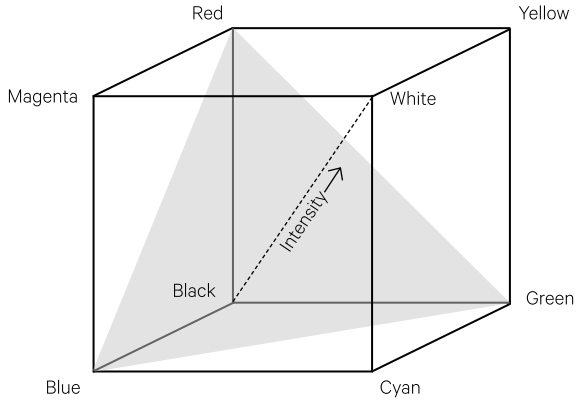
or, in matrix notation:

$$\begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

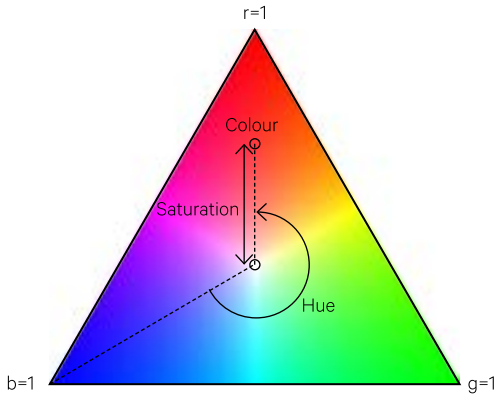
The intensity value is better divided by √3 to give a unit vector, that is the length of the vector from the origin to the point (√3, √3, √3) is 1. The HSI dimensions of an example image are shown in Figure 8.4.

Figure 8.3 HSI coordinates

a. Intensity is measured along the diagonal axis between black and white.



b. Saturation is the linear distance from the intensity axis to the selected colour on an equal intensity plane. Hue is the angular distance from a primary colour (usually blue) to the selected colour on an intensity plane.



Source: Harrison and Jupp (1990) Figure 104

7 Note that the HSL (Hue-Saturation-Luminosity) colour model is different to HSI.

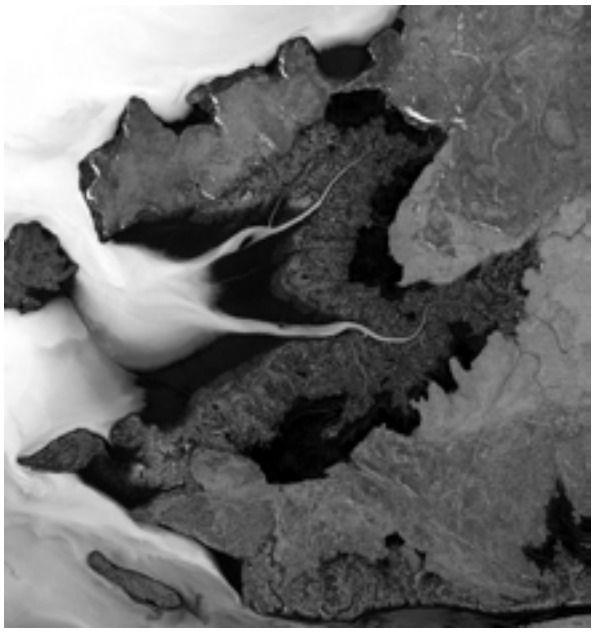
Figure 8.4 Hue, Saturation and Intensity dimensions

This example image was acquired by Sentinel-2a on 10 September 2018 in the Kimberley region, northwest WA.

a. Colour (bands 4, 3 and 2 as RGB)

b. Hue

c. Saturation



d. Intensity

Source Norman Mueller, Geoscience Australia

Colours are the smiles of nature.
(Leigh Hunt)

HSI coordinates can similarly be converted to RGB coordinates using the inverse transformation of:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \times \begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix}$$

where

v_1 = saturation × cos (hue); and

v_2 = saturation × sin (hue).

Since v_1 and v_2 are defined above, this transformation can also be expressed as:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{\sqrt{3}} & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{\sqrt{3}} & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \times \begin{bmatrix} I \\ R \\ G \\ B \end{bmatrix}$$

This single operation effectively replaces the intensity component of the RGB colour composite with another channel defined as the image intensity (I). The rotation from RGB to HSI can be varied to weight hues by their perceived brightness. For example, Niblack (1986) suggests a variation (derived from the luminosities of a NTSC (National Television System Committee: US) standard monitor), which effectively tilts the planes of constant intensity with respect to the intensity axis.

The HSI model has been developed to produce various colour spaces shaped as single and double cones, hexcones and the Munsell irregular shape (Munsell 1929). Spherical coordinates for HSI are also defined by Gillespie *et al.* (1987; see Section 8.3). Other systems for specifying colour are introduced in Volume 2X—Appendix 3. Examples of processes using the RGB-HSI transformation with EO imagery are described in Excursus 8.1.

8.3 Chromaticity Transformation

Gillespie *et al.* (1987) demonstrate the use of the rgb chromaticity coordinates to enhance colour composite imagery. Their method involves converting the channels displayed as red, green and blue to normalised rgb coordinate channels and computing an intensity channel as the sum of their values (see Section 8.1).

This transformation can be implemented using:

- an affine transformation to produce the intensity channel (I; see Section 7);

  $I=R+G+B$

then

- channel ratioing to divide each composite channel by the intensity channel (see Section 8.2):

  $r=\frac{R}{I}$

  $g=\frac{G}{I}$

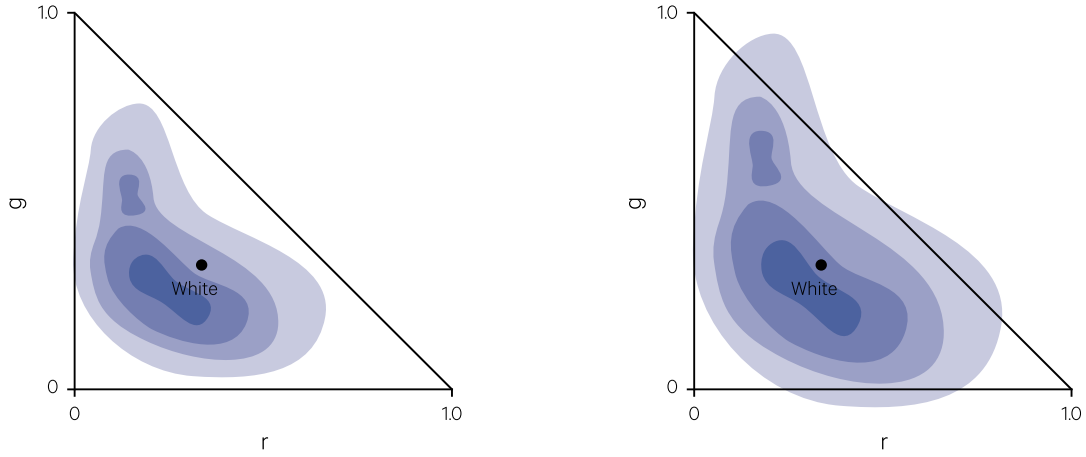
  $b=\frac{B}{I}$

The spectral geometry of this procedure is illustrated in Figure 8.5.

Figure 8.5 Chromaticity transformation

a. Original image data values are shown on a crossplot of r versus g.

b. Each colour component is expanded individually to fill the colour space.



Source: Harrison and Jupp (1990) Figure 103 [Adapted from Gillespie *et al.* (1987)]

The resulting channels can then be contrast-enhanced individually and transformed back to the original channels by multiplying the rgb channels by the intensity channel, I:

$$R = r \times I$$

$$G = g \times I$$

$$B = b \times I$$

Gillespie *et al.* (1987) suggest that colour pictures prepared this way are similar to those enhanced by the HSI and decorrelation methods (see Sections 8.2 and 9.6 respectively).

As further discussed in Section 9.7, alternative intensity data could also be used in the back transformation if required. For example, pan-sharpening describes the operation of merging finer scale panchromatic channel with registered, coarser scale multispectral imagery (see Section 8.4). Image fusion based on the chromaticity transformation is commonly called the Brovey transform (Hallada and Cox, 1983; Jensen, 2016). However, as with many other pan-sharpening techniques, colour distortions can occur in the merged image, especially when the spectral range of the intensity channel differs from the spectral ranges of the colour channels (Liu and Moore, 1998).

8.4 Pan-Sharpening

Pan-sharpening (or panchromatic sharpening) describes the process of merging finer scale panchromatic imagery with a coarser scale colour composite to improve the apparent spatial resolution of the colour image. As further detailed in Volume 2D a wide range of methods have been proposed for pan-sharpening EO imagery (Thomas *et al.*, 2008; Amro *et al.*, 2011; Jalan and Sokhi, 2012).

One approach to pan-sharpening uses an affine transformation to effectively convert between the RGB and HSI colour spaces (see Section 8.2). In this case, the colour composite image would need to be resampled to the same pixel size as the panchromatic imagery and both datasets would need to precisely overlay a common spatial grid (see Volume 2B). The affine transformation matrix shown in Figure 8.6 could then be used to replace the multispectral intensity information with the panchromatic channel.

An example of this transformation applied to Landsat-8 OLI and Panchromatic imagery is shown in Figure 8.7. Spectral distortions can occur in the merged image if the spectral range of the intensity channel differs from the spectral ranges of the colour channels (Liu and Moore, 1998; Thomas *et al.*, 2008; see Volume 2D). Other methods for pan-sharpening are described in Section 9.7 and Volume 2D.

Figure 8.6 Matrix equation to merge high resolution data

$$\begin{bmatrix} \text{high res ch 1} \\ \text{high res ch 2} \\ \text{high res ch 3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & \frac{1}{\sqrt{3}} \end{bmatrix} \times \begin{bmatrix} \text{rescaled ch 1} \\ \text{rescaled ch 2} \\ \text{rescaled ch 3} \\ \text{panchromatic ch} \end{bmatrix}$$

or

$$\text{high res ch 1} = \frac{2}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{panchromatic ch}$$

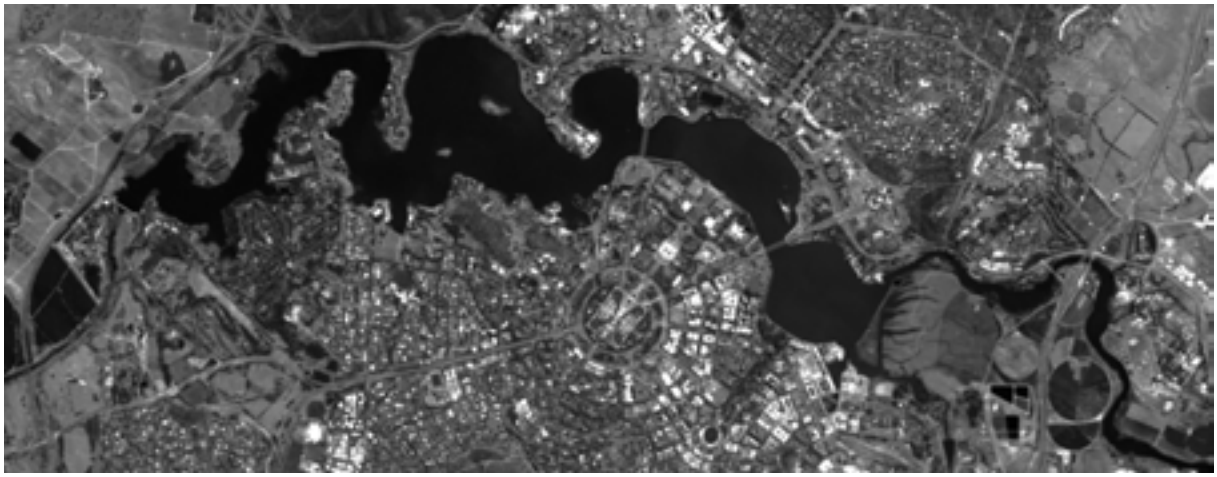
$$\text{high res ch 2} = -\frac{1}{3} \times \text{rescaled ch 1} + \frac{2}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{panchromatic ch}$$

$$\text{high res ch 3} = -\frac{1}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} + \frac{2}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{panchromatic ch}$$

Figure 8.7 Integrating imagery with differing spatial resolution

Landsat-8 panchromatic and OLI images were acquired over Canberra on 25 October 2018. These images have been merged to create a colour image with the spatial resolution of the panchromatic image.

a. Landsat-8 panchromatic image with 15 m pixels



b. Landsat-8 OLI image with 30 m pixels, displayed using bands 6, 5, 3 as RGB

c. Panchromatic and OLI images are merged to create a colour image with 15 m pixels.

Source: Norman Mueller, Geoscience Australia

8.5 Adding Relief Shading

A filter-based process for generating shaded relief imagery from topographic data was introduced in Section 6.2. Such relief imagery can be integrated with the brightness of the composite colours in a colour composite image to represent changes in elevation. For example, in Figure 8.8 a shaded relief channel for southeast Australia has been merged with a thematic image of the Murray-Darling basin to emphasis the role of topography in regional drainage.

The merging process requires that the data range of the relief channel matches those channels forming the composite image. If the data range of an elevation channel is much larger than the channels forming the colour composite, the image channels must first be rescaled to the same range (usually with the scaling parameters used to display the composite image) so that the relief shading will affect all channels equally.

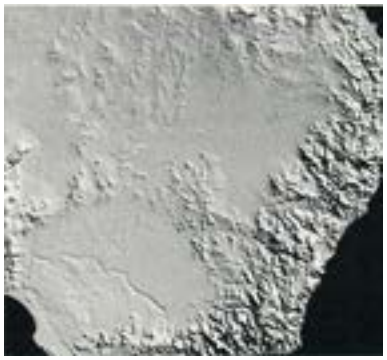
The relief shading can then be merged into the colour image using the same matrix coefficients given in Section 8.4 for pan-sharpening (see Figure 8.9). The resulting shaded composite will have modified brightness values but the same hues as illustrated in Figure 8.8. Weighting values can also be derived for this operation if required. The resulting shaded relief image is sometimes non-linearly scaled to enhance contrast in dark areas.

Figure 8.8 Modifying colour brightness using relief shading

An affine transformation can be used to replace the intensity axis of a colour composite image with some other intensity scale. In this example, a Sun shadowing channel (derived from elevation data for southeast Australia) has been used to highlight surface relief within colour-coded drainage divisions of the Murray-Darling basin. A segmentation mask has been used to restrict the effect of the transformation to within the basin.

a. Original colour image

b. Shaded relief image



c. Colour image with shaded relief

Source: Harrison and Jupp (1990) Plate 17

Figure 8.9 Matrix equation to add relief shading

$$\begin{bmatrix} \text{shaded ch 1} \\ \text{shaded ch 2} \\ \text{shaded ch 3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & \frac{1}{\sqrt{3}} \end{bmatrix} \times \begin{bmatrix} \text{rescaled ch 1} \\ \text{rescaled ch 2} \\ \text{rescaled ch 3} \\ \text{shaded relief ch} \end{bmatrix}$$

or

$$\text{shaded ch 1} = \frac{2}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{shaded relief ch}$$

$$\text{shaded ch 2} = -\frac{1}{3} \times \text{rescaled ch 1} + \frac{2}{3} \times \text{rescaled ch 2} - \frac{1}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{shaded relief ch}$$

$$\text{shaded ch 3} = -\frac{1}{3} \times \text{rescaled ch 1} - \frac{1}{3} \times \text{rescaled ch 2} + \frac{2}{3} \times \text{rescaled ch 3} + \frac{1}{\sqrt{3}} \times \text{shaded relief ch}$$

Excursus 8.1—Examples of HSI-based transformations

Various transformations of EO image channels can be informative for particular applications. To maximise the visual content of an image, selected transformations can be combined into a colour composite, either directly or via an HSI transformation (see Section 8.2). For example:

- information in multi-channel images may be enhanced by substituting appropriate transformations for the hue, saturation and/or intensity components; or
- multiple transformations applied to a single channel can be combined into a colour composite using the HSI transformation.

Two examples of these options are detailed below.

Substituting transformed channels for HSI components

Jaskolla and Henkel (1988) reported a transformation sequence that extends the colour variation in EO imagery compared with a standard false colour composite. Using three input channels containing green, red and NIR reflectance, this sequence computes three intermediate channels to represent the intensity, hue and saturation components:

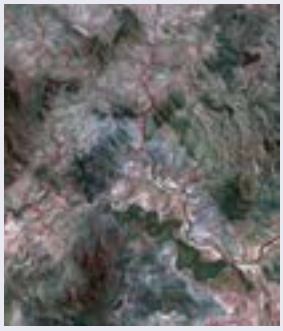
- intensity=high pass filter on the mean value of green, red and NIR bands;
- hue=NIR–(green+red)/2; and
- saturation=red–green.

These three channels can then be combined as an HSI composite and transformed back to RGB space (see Section 8.2). This transformation was developed for SPOT-MSS imagery and is shown for an example Landsat-5 image in Figure 8.10. This enhancement clearly improves colour contrast in the example image and highlights linear features. As such, it has been reported to be useful for geological interpretation (Jaskolla and Henkel, 1988)

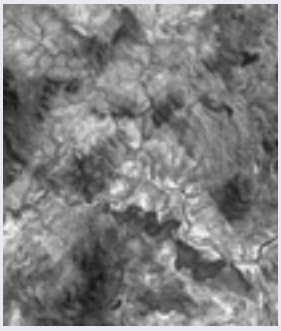
Figure 8.10 Channel substitution using HSI transform

Bands 2, 3 and 4 of a Landsat-5 image, acquired over Charters Towers region, Queensland, on 7 October 1987, are processed to create multiple channels that can be presented as a colour composite using the HSI transformation.

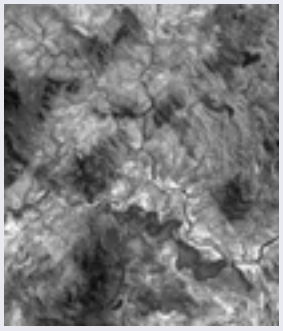
a. Original colour composite (bands 4, 3, 2 as RGB)



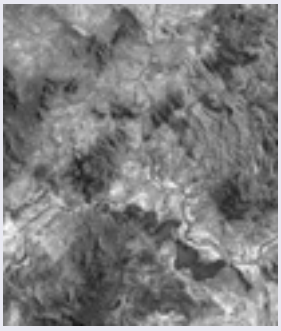
b. Band 2 (green)



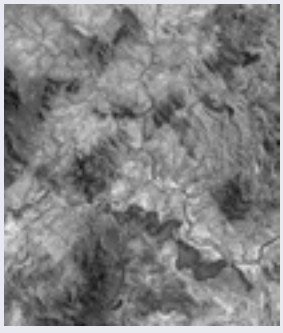
c. Band 3 (red)



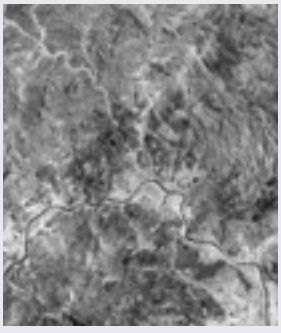
d. Band 4 (NIR)



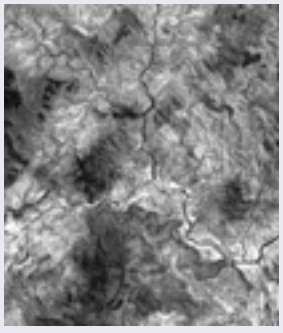
e. High pass filter on average of original channels represents intensity



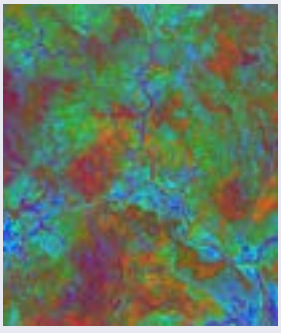
f. (NIR-(green+red) / 2) assumed to represent hue



g. (red-green) assumed to represent saturation



h. Colour composite of HSI transformed back to RGB space



Source: Tony Sparks, Icon Water

Presenting single channel data as a colour composite

Any relevant transformations of an image channel could be substituted for the hue, saturation and/or intensity components in a pseudo-HSI composite. For example, the HSI transformation may be used in conjunction with differential filtering of an elevation channel *k* to display an HSI pseudo-colour composite, which comprises:

intensity=elevation (channel *k*)

hue=aspect ($v_1 = \delta k / \delta x$)

saturation=slope ($v_2 = \delta k / \delta y$)

(Please refer to Sections 5.1.1 and 6 for derivative equations.)

These intensity, hue and saturation components can then be converted to RGB coordinates using the inverse HSI transformation (see Section 8.2):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix} \times \begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix}$$

or

$$R = \frac{1}{\sqrt{3}} \times I + \frac{1}{\sqrt{6}} \times v_1 + \frac{1}{\sqrt{2}} \times v_2$$

$$G = \frac{1}{\sqrt{3}} \times I + \frac{1}{\sqrt{6}} \times v_1 - \frac{1}{\sqrt{2}} \times v_2$$

$$B = \frac{1}{\sqrt{3}} \times I - \frac{2}{\sqrt{6}} \times v_1$$

This representation is applied to an elevation channel in Figure 8.11. It is also reported as being a useful enhancement sequence for visual interpretation of single channel geophysical data (Harrison and Jupp, 1990).

Figure 8.11 Single channel presented as colour composite

Slope and aspect channels have been computed from an elevation channel. When combined as an HSI composite image (in which elevation is represented as intensity, aspect is represented as hue, and slope represented as saturation), these channels can be back-transformed to form an RGB colour composite.

a. Elevation

b. Slope

c. Aspect

d. Colour composite

Source: Harrison and Jupp (1990) Plates 13 and 14

8.6 Further Information

Gonzalez and Woods (2018)

Jensen (2016) Section 8

Sawyer (1955)

8.7 References

Amro, I., Mateos, J., Vega, M., Molina, R., and Katsaggelos, A.K. (2011). A survey of classical methods and new trends in pansharpening of multispectral images. *EURASIP Journal of Advances in Signal Processing 2011,* 2011: 79. *https://doi.org/10.1186/1687-6180-2011-79*

Fung, T., and LeDrew, E. (1987). Application of Principal Components Analysis to change detection. *Photogrammetric Engineering and Remote Sensing,* 53, 1649–1658.

Gillespie, A.R., Kahle, A.B., and Walker, R.E. (1987). Colour enhancement of highly correlated images. II. Channel ratio and 'chromaticity' transformation techniques. *Remote Sensing of Environment,* 22(3), 343–365.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Hallada, W.A., and Cox, S. (1983). Image Sharpening for Mixed Spatial and Spectral Resolution Satellite Systems. *Proc 17th International Symposium on Remote Sensing of Environment*. 9–13 May 1983, Ann Arbor. pp 1023–1032.

Jalan, S., and Sokhi, B.S. (2012). Comparison of different pan-sharpening methods for spectral characteristic preparation: multi-temporal CARTOSAT-1 and IRS-P6 LISS-IV imagery. *International Journal of Remote Sensing*, 33(18), 5629–5643.

Jaskolla, F., and Henkel, J. (1988). Evaluation and digital processing of multispectral SPOT data. *International Journal of Remote Sensing*, 9, 1629–1637.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Liu, J.G., and Moore, J.Mc.M. (1998). Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery. *International Journal of Remote Sensing,* 19(13), 2477–2491. doi: *10.1080/014311698214578*

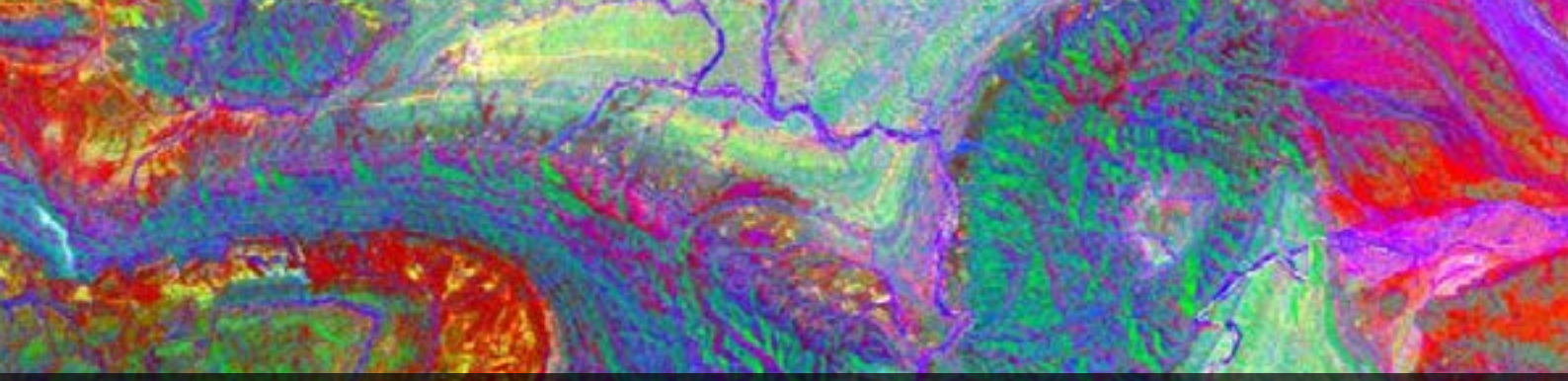
Munsell, A. H. (1929). *Munsell Book of Color*. Munsell Colour Co., Inc.

Niblack, W. (1986). *An Introduction to Digital Image Processing*. Prentice-Hall International. New Jersey.

Sawyer, W.W. (1955). *Prelude to Mathematics*. Penguin Books Ltd. London.

Singh, A., and Harrison, A. (1985). Standardised Principal Components. *International Journal of Remote Sensing*, 6, 883–896.

Thomas, C., Ranchin, T., Wald, L., and Chanussot, J. (2008). Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 1301–1312.



9 Principal Components Analysis

The Principal Components transformation (PCA, also referred to as eigenvalue, Hotelling or discrete Karhunen-Loeve (KL) transforms) uses image spectral statistics to define a rotation of the original image channels such that the data are arranged along axes of decreasing variance. The coordinates for the new axes are computed by an affine transformation (that is, a linear combination) of the original data coordinates (see Section 7).

The transformation parameters for PCA are derived from the covariance or correlation matrix (see Section 1.3.3 above) of an image (or sub-image). This transformation is particularly useful for imagery with many similar channels, such as hyperspectral imagery. The resulting Principal Components (PCs) are totally uncorrelated and can be computed by adding proportions of the original image channel values. Many image processing systems offer options for adjusting channel variance both when defining and applying the PC transformation matrix, and also to back transform from the PCA space to the original data channels.

In this section we introduce the PCA transformation (see Section 9.1), interpretation of principal components (PCs; see Section 9.2), and the transformation parameters (see Section 9.3), then present common applications of PCA in the context of EO imagery:

- decorrelation stretching (see Section 9.4);
- data reduction (see Section 9.5);
- highlighting specific features (see Section 9.6);
- multi-scale intensity enhancement (see Section 9.7); and
- change detection (see Section 9.8).

9.1 PCA Transformation

As discussed in Section 7.1, a multi-channel image can be considered as a multidimensional matrix, or multi-vector space. For a two channel image, the two-dimensional data space can be represented in a crossplot (see Section 1.3.4 above). It is possible to redefine pixel values in an image by geometrically changing the positions of, or relationship between, the axes of this data space. For example, in Figure 9.1, by rotating and shifting the channel axes, the point A becomes defined as position p' and q'. Data space axes can similarly be rescaled, reflected or rotated as detailed in Volume 2X—Appendix 6. Such changes can allow spectral patterns and relationships in the data to be enhanced and analysed more directly.

Mathematical matrices allow changes to the 'spectral' geometry of two and higher-dimensional data to be represented algebraically. For example, in Figure 9.1, channels X and Y show a high degree of correlation in the original data space. As detailed in Volume 2X—Appendix 6, the original axes can be rotated and each pixel assigned a new value relative to the new coordinate system. In Figure 9.1 the rotated channel X' is aligned with the direction of maximum variance so that the data channels are uncorrelated in the rotated space. This type of rotation is used in Principal Components Analysis for data reduction and enhancement.

The first rule of intelligent tinkering is to save all the parts. (Paul R. Ehrlich)

Background image: Landsat TM image of Flinders Ranges, South Australia, transformed to Principal Component image. PCs 1, 2 and 3 are displayed to highlight lithology, soils and vegetation. **Source:** Megan Lewis, University of Adelaide

The principle of a PCA transformation is illustrated in Figure 9.2 for a two-channel image. The first principal component, PC1, is defined as the axis that passes through the mean data value and is aligned with maximum variance in the image. The second principal component, PC2, is defined as being orthogonal (perpendicular) to PC1.

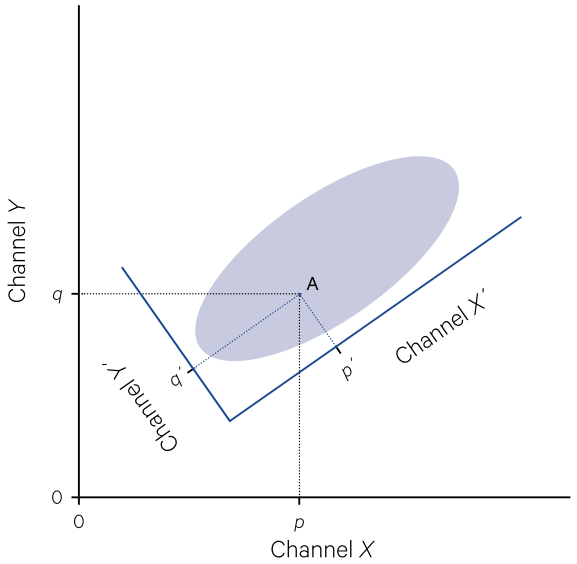
The same concept can be applied to images with more than two channels, with the rotation transformation being computed from the covariance matrix (see Section 1.3.3). This matrix would typically be derived from the whole image but, if PC1 was required to be aligned with the variation in a particular cover type to enhance a specific feature, it may also be based on a training set (see Volume 2A—Section 9.1.2) or spectral theme (see Volume 2A—Section 9.1.3).

The covariance or correlation matrices indicate the degree to which image channels are correlated (see Section 1.3.3 and Volume 2A—Section 8.1.4). The PCA transformation produces an image with totally uncorrelated channels. This is done by mathematically defining a rotation that creates an image with all zero covariance values. In terms of the covariance matrix, this would mean that all off-diagonal elements were equal to zero. Since only the diagonal elements are non-zero, such a matrix is referred to as a diagonal matrix.

As detailed in Section 7 above, a general linear transformation can be represented as a matrix operator. This transformation redefines each image pixel relative to a new coordinate system. Special vectors exist between the two coordinate systems such that points on these vectors are simply redefined by the transformation as a constant multiple of their original coordinates. These vectors are called eigenvectors of the matrix and the constant values are called eigenvalues. Mathematically this is expressed as shown in Figure 9.3. A simple image example of the computation of these statistics is presented in Excursus 9.1. The eigenvectors and eigenvalues of a matrix may be found using matrix algebra as detailed in Volume 2X—Appendix 6 (which gives the mathematical derivation of the PCA transformation; see also Volume 2E).

Figure 9.1 Redefining image axes

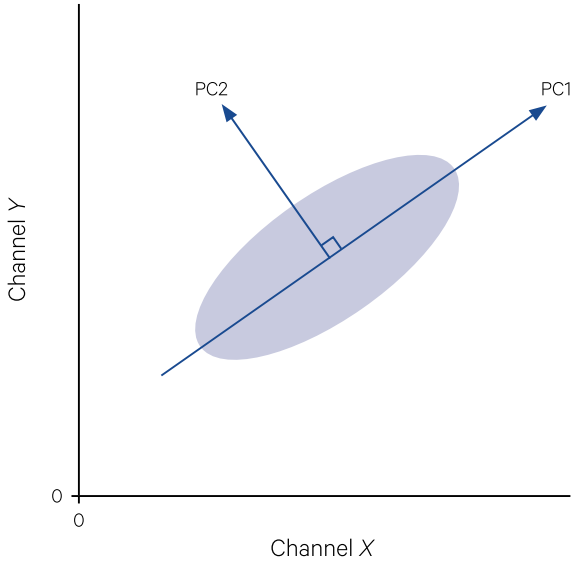
Image data values can be redefined relative to modified position and scaling of axes.



Source: Harrison and Jupp (1990) Figure 87

Figure 9.2 Principle of PCA transformation

A new primary data axis (PC1) is defined in the direction of maximum image variance (over all image channels) then other axes are defined to be orthogonal to this direction.

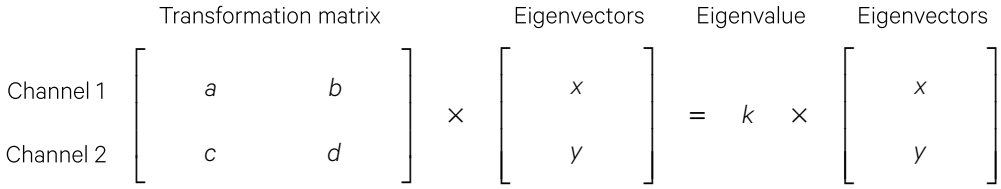


Source: Harrison and Jupp (1990) Figure 88

Nothing can be divided into more parts than it can be constituted of.
But matter (i.e. finite) cannot be constituted of infinite parts.
(Isaac Newton)

Figure 9.3 Eigenvectors of a matrix

Special vectors called eigenvectors exist between the original and transformed image coordinates, which can represent the transformation as a constant multiple (called the eigenvalue) of the original coordinates.



or

$$\text{Channel } 1 = a \times x + b \times y = k \times x$$

$$\text{Channel } 2 = c \times x + d \times y = k \times y$$

Source: Harrison and Jupp (1990) Figure 89

The major direction of data variation in multi-channel EO imagery rarely lies along one original channel axis. Similarly the variance gradient for specific image features is usually defined in at least two data dimensions. These statistical characteristics of EO data make it difficult to neatly analyse the component features or divide up the image brightness variations as channel themes (see Volume 2A—Section 9.1). By aligning the data axes with the variance of the whole image dataset, or just with a specific feature in it, we can access the changes within the major variance using a single channel.

Thus, the PCA transformation allows a greater proportion of the image variance to be represented in fewer uncorrelated channels. These PCs can be combined into a colour composite, which typically displays a larger range of colours than any composite images formed from the original channels. Colour composite images formed from PCs are frequently used to highlight linear features for geological analyses (see also Excursus 8.1).

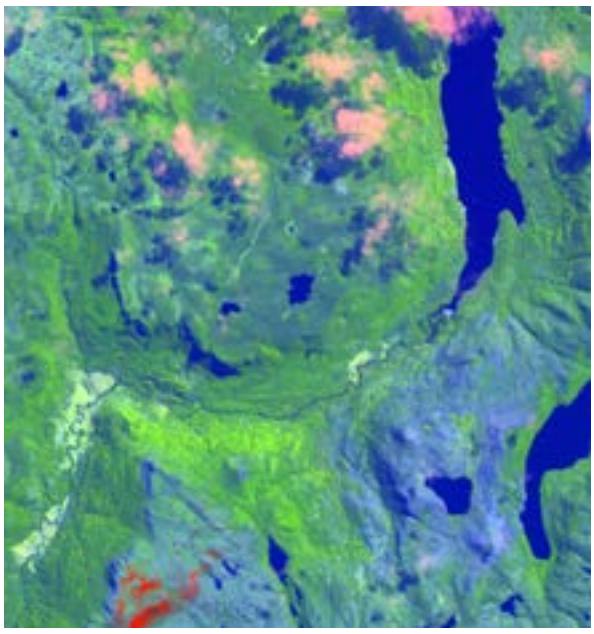
An example of the PCA transformation is shown in Figure 9.4. In this example, the contrast both between and within features in Figure 9.4a is visually enhanced in Figure 9.4b. In this case, PC1 is dominated by cloud (showing as red or pink in the PCA image) since these features have large contrasting values and create high multidimensional variance in the original data.

Figure 9.4 Principal components transformation

PCA was applied to a 10 band Sentinel-2a image over central Tasmania, acquired on 5 October 2018. This analysis is based on bands 1, 2, 3, 4, 5, 6, 7, 8, 11, and 12 (all visible, red-edge, and NIR bands except 8a; see Figure 1.2).

a., Original cloudy image is displayed as a natural colour composite using bands 4, 3 and 2 as RGB.

b. PCA image formed by displaying PC1 as red, PC2 as green and PC3 as blue.



Source: Norman Mueller, Geoscience Australia

Excursus 9.1—PCA statistics

To show the relationship between image statistics and PC equations, we will use the simple example image from Section 1.3.4 (see Figure 1.8). As detailed in Volume 2A—Section 8.1 and reviewed in Section 1.3.3, the covariance matrix for this three channel image can be computed as:

$$C_x = \begin{bmatrix} 2.9380 & -0.3840 & 2.3000 \\ -0.3840 & 4.7200 & -0.0144 \\ 2.3000 & -0.0144 & 2.2656 \end{bmatrix}$$

The eigenvalues of the matrix are derived as:

$$\begin{vmatrix} 2.938-k & -0.384 & 2.3 \\ -0.384 & 4.72-k & -0.0144 \\ 2.3 & -0.0144 & 2.2656-k \end{vmatrix} = 0$$

which gives eigenvalues of 5.1, 4.5 and 0.3 for the three channels (see Volume 2X—Appendix 6.3). The covariance matrix of the PCA image (which by definition has no correlation between channels) can then be written as:

$$\begin{bmatrix} 5.1 & 0 & 0 \\ 0 & 4.5 & 0 \\ 0 & 0 & 0.3 \end{bmatrix}$$

The eigenvectors of the example image covariance matrix can also be computed to form the rows of the PCA transformation matrix (see Volume 2X—Appendix 6.3 for details of derivation). For our example, the final transformation matrix is computed as:

$$\begin{bmatrix} 0.6306 & -0.5879 & 0.5068 \\ 0.4153 & 0.8072 & 0.4195 \\ 0.6556 & 0.0541 & -0.7531 \end{bmatrix}$$

This matrix can now be used to rotate the data space of original image to the uncorrelated PCA image space via the equations:

$$PC1 = 0.6306 \times \text{ch A} - 0.5879 \times \text{ch B} + 0.5068 \times \text{ch C}$$

$$PC2 = 0.4153 \times \text{ch A} + 0.8072 \times \text{ch B} + 0.4195 \times \text{ch C}$$

$$PC3 = 0.6556 \times \text{ch A} + 0.0541 \times \text{ch B} - 0.7531 \times \text{ch C}$$

Other statistics relevant to the PCA transformation are introduced in Volume 2X—Appendix 6.3. For example, the degree of correlation between each original image channel and each principal component can also be determined. For our example image, this is summarised by the normalised matrix:

$$\begin{bmatrix} 0.8307 & -0.6109 & 0.7604 \\ 0.5136 & 0.7875 & 0.5909 \\ 0.2096 & 0.0197 & -0.2742 \end{bmatrix}$$

The elements of this matrix indicate the degree of correlation between the original channels and each PC. For example, channels A and C are strongly correlated with PC1 (0.8307 and 0.7604 respectively) while channel B is negatively correlated with PC1 (-0.6109) and more strongly correlated with PC2 (0.7875).

The percentage variation accounted for by each PC can also be computed for our example image (see Table 9.1), indicating that PC1 and PC2 together represent 97.3% (51.8+45.5) of the variation contained in the three channels of the original image.

Table 9.1 Principal component percentage variation

PC	Variance	Percentage
1	5.1	51.8
2	4.5	45.5
3	0.3	2.7
Total	9.9	100

'Think simple' as my old master used to say -
meaning reduce the whole of its parts into the simplest terms,
getting back to first principles.
(Frank Lloyd Wright)

9.2 Interpretation of Principal Components

Interpretation of principal component channels depends on the features of the imagery being transformed. The matrix coefficients specify the contribution each original image channel is making to the transformed channels, so can indicate the 'meaning' of each PC. In Landsat TM imagery, for example, PC1 usually represents (at least) 90% of the data variation and is typically computed using all positive coefficients. Since the PC1 channel is a useful summary of image brightness, including topographic shading, it may be used in transformations that require a relief shading channel (see Section 8.4) or to split spectral classes derived from log-ratio channels into categories which indicate surface shading. In Landsat TM imagery containing extensive areas of vegetation, PC2 is usually related to vegetation greenness, while PC3 is generally indicative of vegetation 'yellowness' or soil colour (see also Section 11.1.3). In this case, the coefficients for PC2 would be positive in the infrared channels and negative in the visible channels (or vice versa), however these definitions vary widely depending on characteristics of the scene and the image channels being processed. Low order PCs are generally accepted to represent the noise level in the image and typically contain (much) less than 5% of the data variation.

An example of PCA results for a Landsat TM image is given in Excursus 9.2.

Excursus 9.2—PCA Example

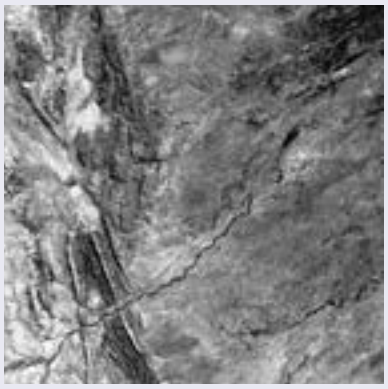
Source: Megan Lewis, University of Adelaide

This excursus presents an example of PCA transformation applied to the optical bands in Landsat TM imagery. The original image channels are shown in Figure 9.5 (see Section 1.1 for band definitions) and a colour composite of three original bands is shown in Figure 9.6.

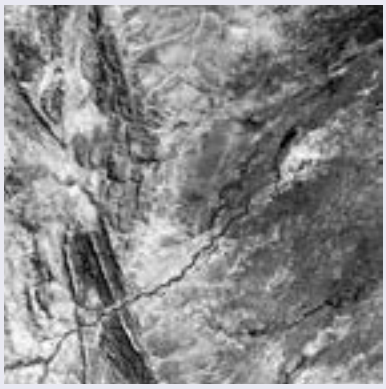
Figure 9.5 Original image channels

Landsat TM image channels over Fowlers Gap Research Station, NSW, acquired in June 1993.

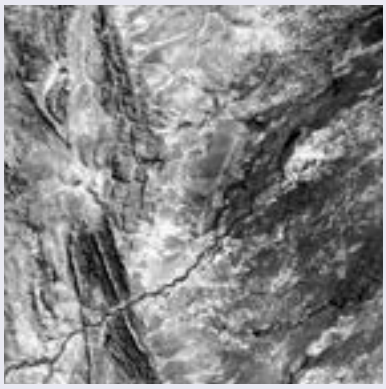
a. Band 1 (blue)



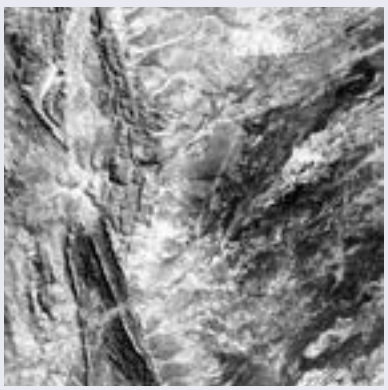
b. Band 2 (green)



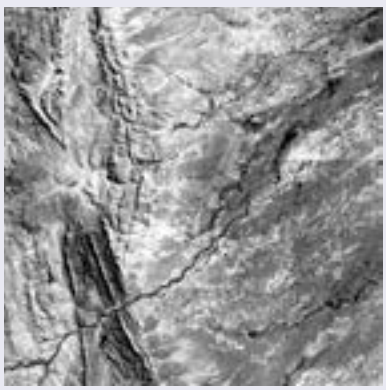
c. Band 3 (red)



d. Band 4 (NIR)



e. Band 5 (SWIR1)



f. Band 7 (SWIR2)

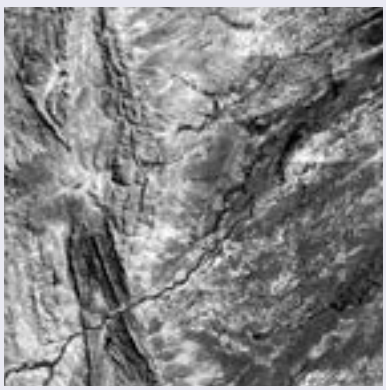
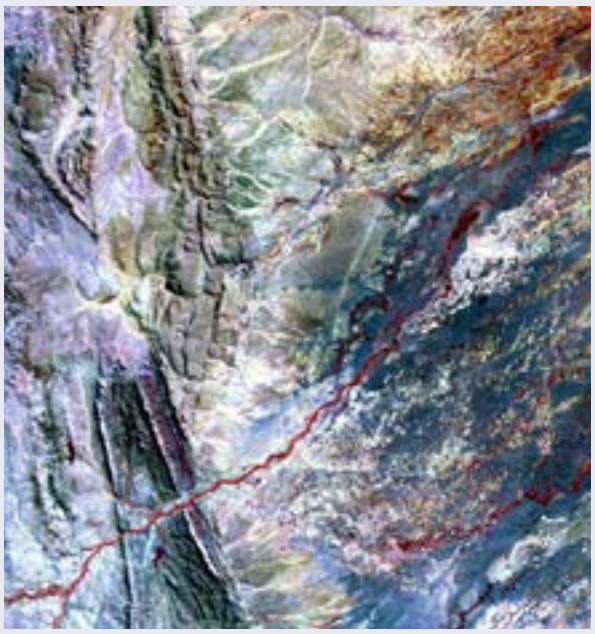
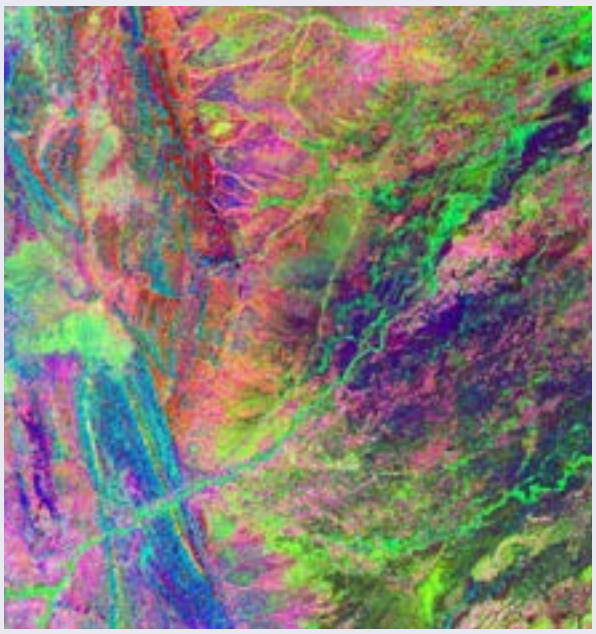


Figure 9.6 Colour composites

a. Landsat TM bands 4, 3, 2 as RGB



b. PC1, PC2 and PC3 as RGB



The correlation matrix for this image is shown in Table 9.2. As reviewed in Section 1.3.3 (and detailed in Volume 2A—Section 8.1), this table summarises the correlation between image channels. This example matrix indicates that the original TM bands are strongly correlated, with little spectral contrast between land surfaces. For example, the correlation between bands 3 (red) and 7 (SWIR) is 0.879 (~88%), meaning that 88% of the information in band 3 is also available from band 7.

The PC transformation is derived from the correlation matrix to give the transformation matrix shown in Table 9.3. This table summarises the PC equations to convert from the original channels to PCs (see Section 9.1). For example, the first row contains the coefficients required to compute PC1 from the original bands:

$$PC1 = 0.058 \times band1 + 0.059 \times band2 + 0.045 \times band3 + 0.375 \times band4 + 0.593 \times band5 + 0.706 \times band7$$

This transformation enables a new set of channels to be computed which isolate sensor noise to one component (PC6 in this example) and maximises the spectral information in the higher components (see Figure 9.7). The range of colours displayed in the false colour composite based on the original channels (see Figure 9.6a) clearly contrasts with the colour composite formed by PCs 1, 2 and 3 (Figure 9.6b), which demonstrates that the three highest PCs contain more 'unique' information than the three original bands displayed in the colour composite.

Table 9.2 Correlation matrix for original image

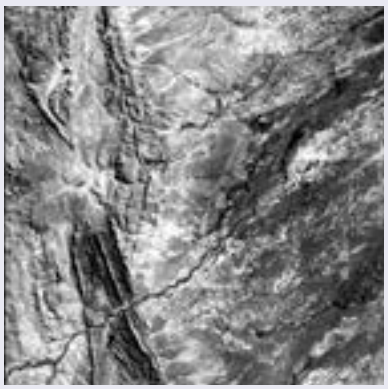
Band	Band					
	1	2	3	4	5	7
1	1.0	0.788	0.65	0.546	0.554	0.531
2	0.788	1.0	0.902	0.803	0.775	0.773
3	0.65	0.902	1.0	0.902	0.873	0.879
4	0.546	0.803	0.902	1.0	0.836	0.821
5	0.554	0.775	0.873	0.836	1.0	0.928
7	0.531	0.773	0.879	0.821	0.928	1.0

Table 9.3 PC transformation matrix

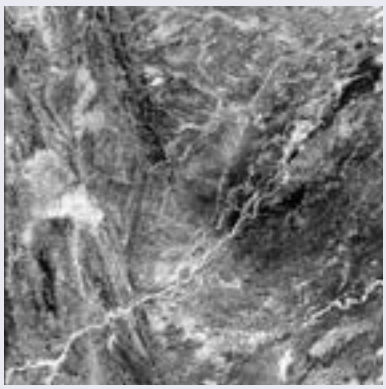
PC	Band					
	1	2	3	4	5	7
1	0.058	0.059	0.045	0.375	0.593	0.706
2	0.153	0.154	0.137	0.534	0.432	-0.680
3	0.336	0.271	0.243	0.540	-0.652	0.196
4	0.453	0.758	-0.131	-0.426	0.150	-.008
5	0.603	-0.408	-0.675	0.115	-0.021	-0.016
6	0.540	-0.398	0.669	-0.298	0.117	0.006

Figure 9.7 Principal component channels

a. PC1

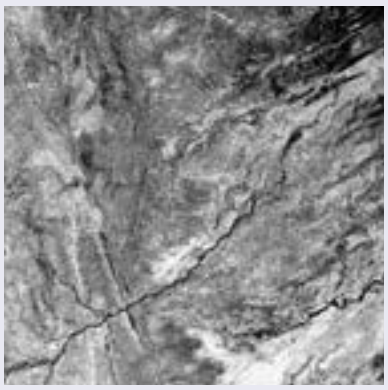


b. PC2

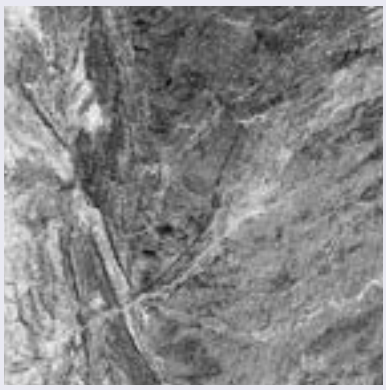


c. PC3

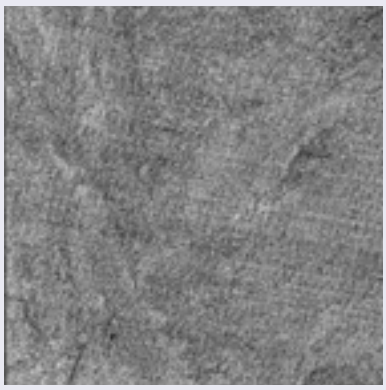
d. PC4



e. PC5



f. PC6



9.3 Transformation parameters

The PCA transformation matrix is generally computed using the specified image statistics (channel means and covariance or correlation matrix), possibly with adjustment for selected channel weights and a balance factor. This matrix can then be applied to selected images as a linear transformation. The resulting PC values are usually rescaled to the full image data range using user-defined expected minimum and maximum values, which are often best determined iteratively.

A principal component may also be removed in the 'back transformation' from the PCA axes to the original image space. This is particularly relevant to noisy imagery where the 'noise feature' can be isolated as one component. Back transformation without that component then effectively removes the noise from the original data space. In four channel Landsat MSS imagery, for example, the PC4 usually contains (much) less than 5% of the total image variance and visually appears to only summarise image noise.

The image statistics used to define the PC transformation may be derived from one or more images, or may be defined by previous analyses. The latter, for example, may apply to a specific image feature (see Volume 2A—Section 8) or an image classification (see Volume 2E). Once defined, to compute the PC transformation the channels of the input matrix may be weighted in different ways such as:

- equal—unweighted; this option would be applicable when channel scaling is correct and equal in all channels of the image;
- inversely proportional to channel standard deviation—this converts the covariance matrix to a correlation matrix so each channel is considered to have the same variance; this option may be useful when the image channels are from different sources (in terms of sensor or date) and/or different transformations, and have totally different scaling;
- inversely proportional to channel data range—this can be used for nominal data or when variance is not related to image distribution, for example in an image with a bi-modal histogram; or
- defined by the user—you may already know the relationship between channels and want to specify which one(s) make the greatest contribution to the PCA; this option is also useful to calibrate the data channels to radiance or some appropriate units before processing.

In general terms, unweighted PCA assumes the error component in the image data is uncorrelated and has the same variance in all channels whereas weighted PCA assumes that the error in all channels is uncorrelated and has variance proportional to the total image variance. Local variance statistics (see Volume 2A—Section 8.2.1) can be considered as an estimate of relative error levels between channels so, for each channel, weights for the PCA could be determined as:

$$\frac{1}{\text{local standard deviation}}$$

Singh and Harrison (1985) claim that a significant improvement in the signal-to-noise ratio of PC images can be achieved using the correlation rather than covariance matrix for PCA of EO data. They refer to the results of a transformation based on the correlation matrix as 'standardised PCs'. This standardisation is suggested to be especially useful in multi-temporal analyses to minimise differences due to atmospheric conditions or Sun angle between images (Fung and LeDrew, 1987).

User-defined channel weighting may be required for nominal data where the scaling in the original channels has been specifically adjusted for category definition, and the range and variance statistics are not comparable between channels. In most EO applications, however, channel weighting is not required since the scaling between channels of a single image is relatively consistent in current image data sources. It cannot be assumed, however, that the scaling will be consistent between different images (from multiple dates or sources; see Volume 2D).

The scaling of the PC channels can also be modified by a 'balance factor', which could range from a minimum value (often 0), where all PC scales are equal, to a maximum value (often 1), which scales each PC in proportion to variance. This parameter is most relevant to back rotation from the PCA space to the original image data space as it effectively allows contrast enhancement of the transformed data before conversion to the original data space. This produces similar results to a decorrelation stretch as described in Section 9.4.

9.4 Decorrelation Stretching

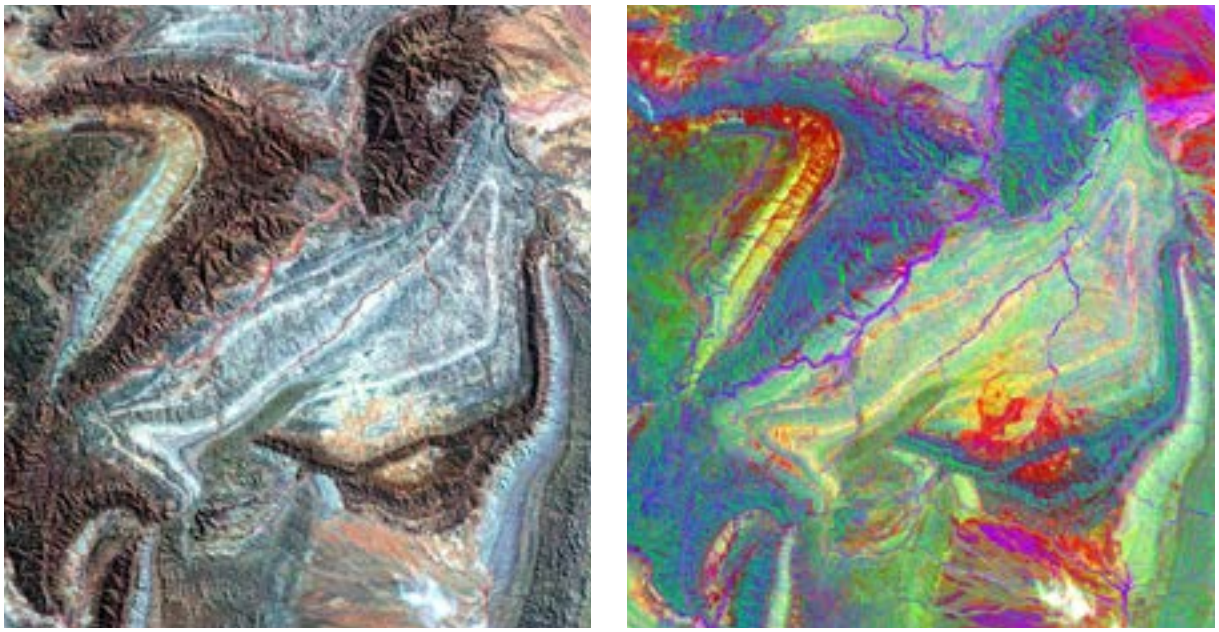
Imagery produced by PCA typically has much more saturated colours than standard colour composites as illustrated in Figure 9.8. This is due to the decorrelation between PC channels. In terms of the RGB colour cube described in Section 8.1, EO imagery generally displays little variation about the intensity axis, with no potential to form the subtractive primary colours at the corners of the colour cube (see Figure 9.9). A colour composite comprising the first three PC channels however, effectively fills the volume of this cube so that both additive and subtractive primaries are visible in the displayed imagery (see Excursus 9.2).

Figure 9.8 Principal Component transformation

Landsat TM image over Flinders Ranges, South Australia

a. Original false colour composite is displayed using bands 4, 3 and 2 as RGB.

b. Principal component image based on PCs 1, 2 and 3 highlights lithology, soils and vegetation.



Source: Megan Lewis, University of Adelaide

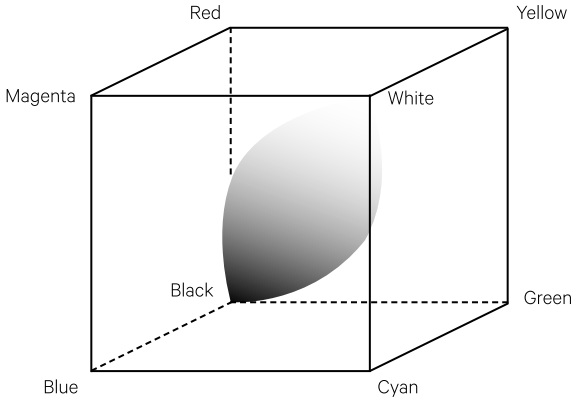
Decorrelation stretching provides a useful visual enhancement for increasing colour saturation in composite imagery (Gillespie *et al.*, 1986). As illustrated in Figure 9.10 this process involves:

- transforming image channels to the PC space;
- stretching the image data along the PC axes to take on a spherical distribution; then
- transforming the data back onto the original channel axes.

This enhancement is particularly useful for visual interpretation of EO imagery in a geological application (Rothery, 1987). The back transformation can also be applied with one component removed. This is useful for removing certain spatial noise patterns from imagery where the noise can be isolated as one component.

Figure 9.9 RGB colour cube showing original image data range

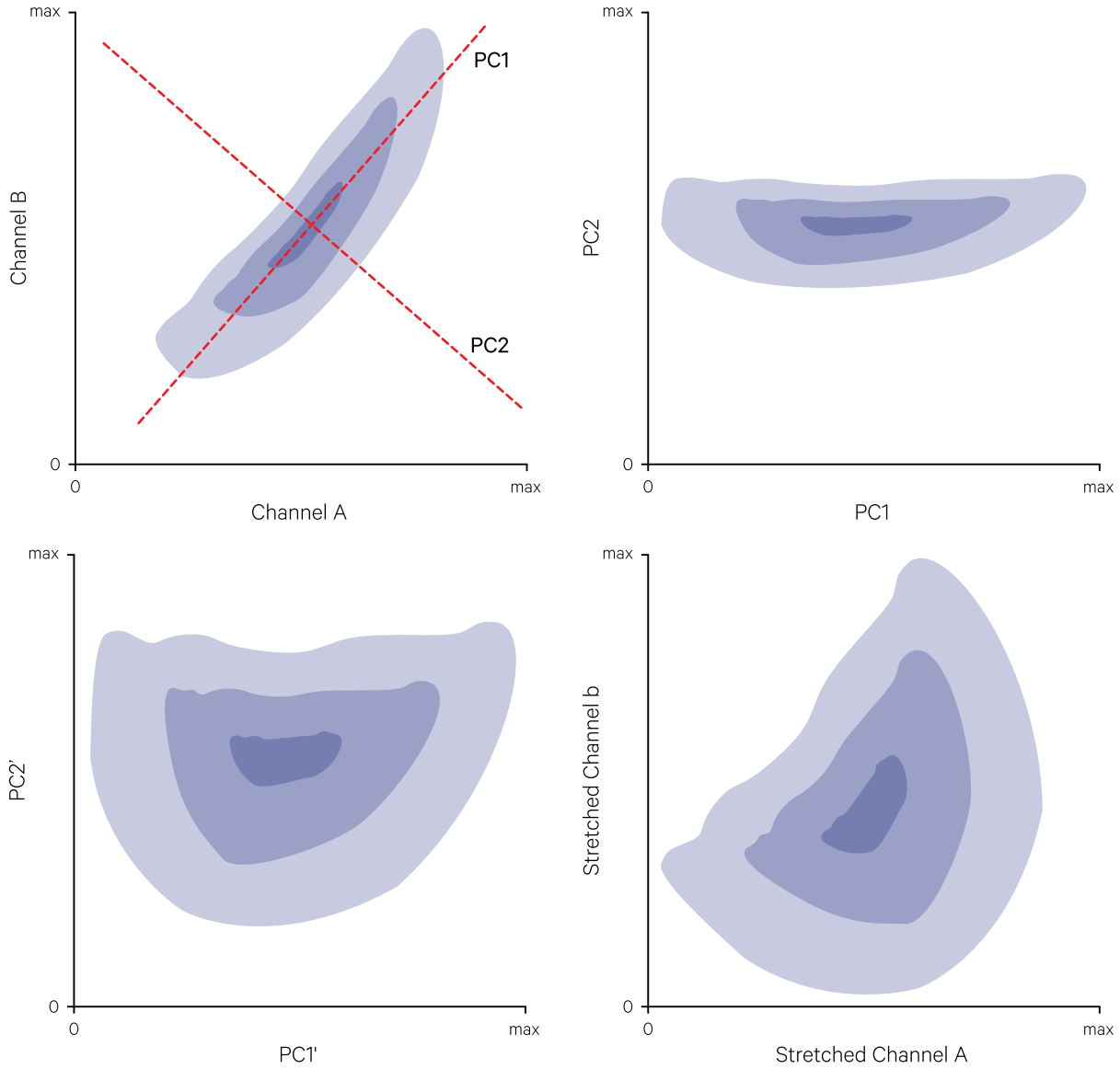
The correlation between channels in EO imagery generally means that the subtractive primary colours (yellow, magenta and cyan) do not occur in colour composite images.



Source: Harrison and Jupp (1990) Figure 90

Figure 9.10 Decorrelation stretching

Images can be transformed to the principal component space for contrast stretching then back transformed to the original image channels. By rescaling the data in an uncorrelated data space, the resulting composite image has enhanced colour contrast.



Source: Harrison and Jupp (1990) Figure 91 [Adapted from: Gillespie *et al.* (1986)]

9.5 Data Reduction

Most EO images contain some redundant information. For example, Landsat TM2 and TM3 bands contain similar data values since most ground cover types give similar responses in these two channels (that is, they are highly correlated). When redundancies do occur in an image, it is possible to 'remove' the duplicate information and represent the data in fewer channels. For many EO datasets, the vast majority of data variation contained within multiple image channels can be represented by their first few PCs (Chuvieco and Huete, 2010; Jensen, 2016).

PCA is principally used for data reduction, although frequently this is an intermediate step in other processes such as image classification, enhancement or integration. By representing the image data in fewer channels (that is, reducing the data dimensionality), various operations may be applied to the transformed channels before transforming back to the original data space. Such data reduction techniques are particularly relevant to hyperspectral imagery. Analysis of 128 channel HyMap airborne imagery, for example, showed that:

- PCs 1–4 contained 99% of the image variation;
- PCs 5–50 mostly contained local variation information and anomalies; and
- PCs 50–128 mostly contained instrument noise, although may also contain useful information about anomalies or small features (Megan Lewis, *pers. comm*; see Volume 2D)

This data reduction operation is particularly relevant to geological applications (where feature identification relies on visual interpretation of enhanced imagery) and hyperspectral imagery (with numerous, often correlated bands). In image classification, PC1 and PC2 provide a useful basis for autoseeding methods based on defining channel themes (see Volume 2A—Section 9.1). Volume 2E describes this procedure in more detail. Once the image has been classified using a set of basic classes defined this way, they may be transformed back to the original image data channels using the 'spectral transfer' operation described in Volumes 2D and 2E.

A sentence should contain no unnecessary words,
a paragraph no unnecessary sentences,
for the same reason that a drawing should have no unnecessary lines
and a machine no unnecessary parts.
(William Strunk, Jr.)

9.6 Highlighting Specific Features

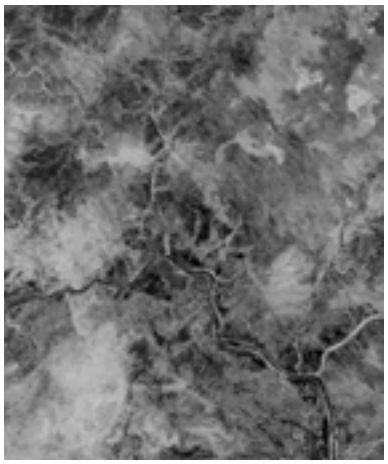
PCA can be advantageous for isolating specific image features as individual PCs. Some applications of PCA to EO imagery include mapping soils (Mitternicht and Zinck, 2003; Viscara-Rosel and Chen, 2011) and detecting cloud (Amato *et al*., 2008).

Fraser and Green (1987) describe an interesting technique called 'directed PCA' (DPCA), which is intended to map specific image information to a given PC. This technique involves calculating PCs on two input ratio channels. For geological analyses, these ratio channels are selected so that one is a geological discriminant that is often confused with vegetation (such as 1.65 μm/2.22 μm; see Figure 9.11a) and the other is primarily a vegetation index (such as NIR/red; see Figure 9.11b). PCA of these two channels produces two uncorrelated data axes: the first aligned with vegetation and second with non-vegetation (see Figure 9.11c). For geological analyses this technique is reported to substantially reduce the effects of vegetation in woodland environments. In this example, vegetation is visibly suppressed along drainage lines.

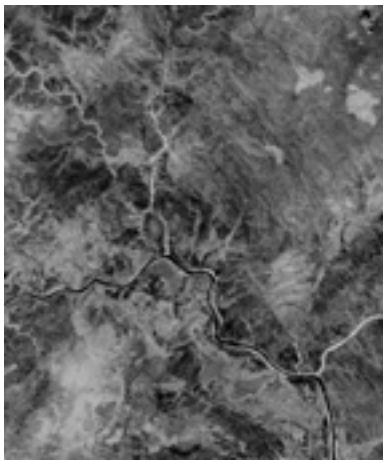
Figure 9.11 Software defoliant

A Landsat-5 image (path 95, row 73) over Charters Towers region, Queensland, acquired on 7 October 1987, was processed to produce two ratio channels, one highlighting clay soils and the other highlighting vegetation (see also Figure 8.10). When PCA is applied to these two ratio channels, PC2 effectively maps non-vegetation, so can be considered as a 'software defoliant'. In this case, suppression of vegetation is most evident in drainage channels.

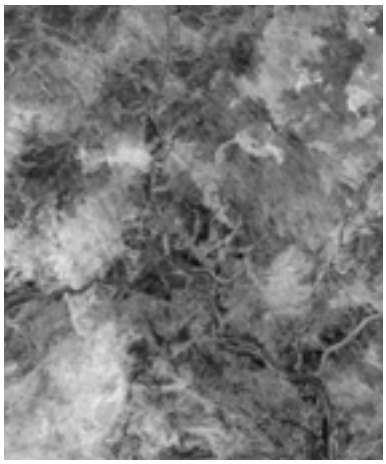
a. Ratio channel highlighting clay soils



b. Ratio channel highlighting vegetation



c. PC2 highlighting non-vegetation



Source: Tony Sparks, Icon Water

9.7 Multi-scale Intensity Enhancement

Panchromatic imagery with finer spatial resolution than a multispectral image can be integrated with the multichannel data to visually enhance its spatial detail. This pan-sharpening process is introduced in Section 8.4 in terms of the RGB-HSI transformation (see Figure 8.7).

Various methods have been suggested for integrating image pairs with differing spatial resolutions (such as Cliche *et al*., 1985 and Price, 1987a). PCA can also be used for pan-sharpening a spatially-registered panchromatic/multispectral image pair as part of the following processing sequence:

- resample multispectral image to the pixel size of the panchromatic image (see Volume 2A—Section 6, and Volume 2B);
- transform selected channels in the resampled multispectral image to PCs;
- replace PC1 with the registered panchromatic channel (since PC1 represents the brightness or intensity variations in EO data). The panchromatic channel should first be rescaled to match PC1 using the regression technique described in Section 7.2.1 (see also Volume 2D); and
- back-transform the resulting image to the original multispectral image channels to produce an image in which the spatial detail of the panchromatic channel is supplemented by the colours of the coarser composite image.

The back transformation matrix will need to be adjusted to account for the scaling between the actual PC ranges and their representation in the PC image. This matrix is implemented as an affine transformation (see Section 7.1).

This approach offers a very flexible method of merging new information with any spatial dataset, since the new 'vector' (or channel) can be used to replace any existing vector in the original data. However, for replacing the intensity component of a colour image, the direct method discussed in Section 8.4 is most appropriate.

Most pan-sharpening applications involve panchromatic and multispectral images from the same sensor (such as SPOT MSS and SPOT Pan). Procedures involving PCA and multivariate regression have also been suggested for substituting signals from one sensor for signals in another (such as Suits *et al*., 1988). If required, such methods can also be used to transfer spectral interpretation techniques from one sensor to another. Data fusion approaches are further discussed in Volume 2D in the context of image integration.

9.8 Change Detection

PCA is frequently used with multi-temporal imagery as a change detection method, by either:

- transforming two or more images separately to PCs, then selecting the most appropriate PCs for comparison; or
- registering multiple images and transforming them to PCs as a merged dataset.

The use of PCA for change detection studies is further discussed in Volume 2D.

9.9 Further Information

Chuvieco and Huete (2010) Section 7

Jensen (2016) Section 8

Gonzalez and Woods (2018) Section 12

9.10 References

Amato, U., Antoniadis, A., Cuomo, V., Cutillo, L., Franzese, M., Murino, L., and Serio, C. (2008). Statistical Cloud Detection from SEVIRI Multispectral Images. *Remote Sensing of Environment*, 112, 750–766.

Chuvieco, E., and Huete, A., (2010). *Fundamentals of Satellite Remote Sensing*. CRC Press, Boca Raton.

Cliché, G., Bonn, F., and Teillet, P. (1985). Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement. *Photogrammetric Engineering and Remote Sensing*, 51, 311–316.

Fraser, S.J., and Green, A.A. (1987). A software defoliant for geological analysis of band ratios. *International Journal of Remote Sensing*, 8, 525–532.

Gillespie, A.R., Kahle, A.B., and Walker, R.E. (1986). Colour enhancement of highly correlated images. I. Decorrelation and HSI contrast stretches. *Remote Sensing of Environment*, 20(3), 209–235.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Metternicht, G.I., and Zinck, J.A. (2003). Remote Sensing of Soil Salinity: Potentials and Constraints. *Remote Sensing of Environment*, 85, 1–20.

Price, J.C. (1987a). Combining Panchromatic and Multispectral Imagery from Dual Resolution Satellite Instruments. *Remote Sensing of Environment*, 21, 119–128.

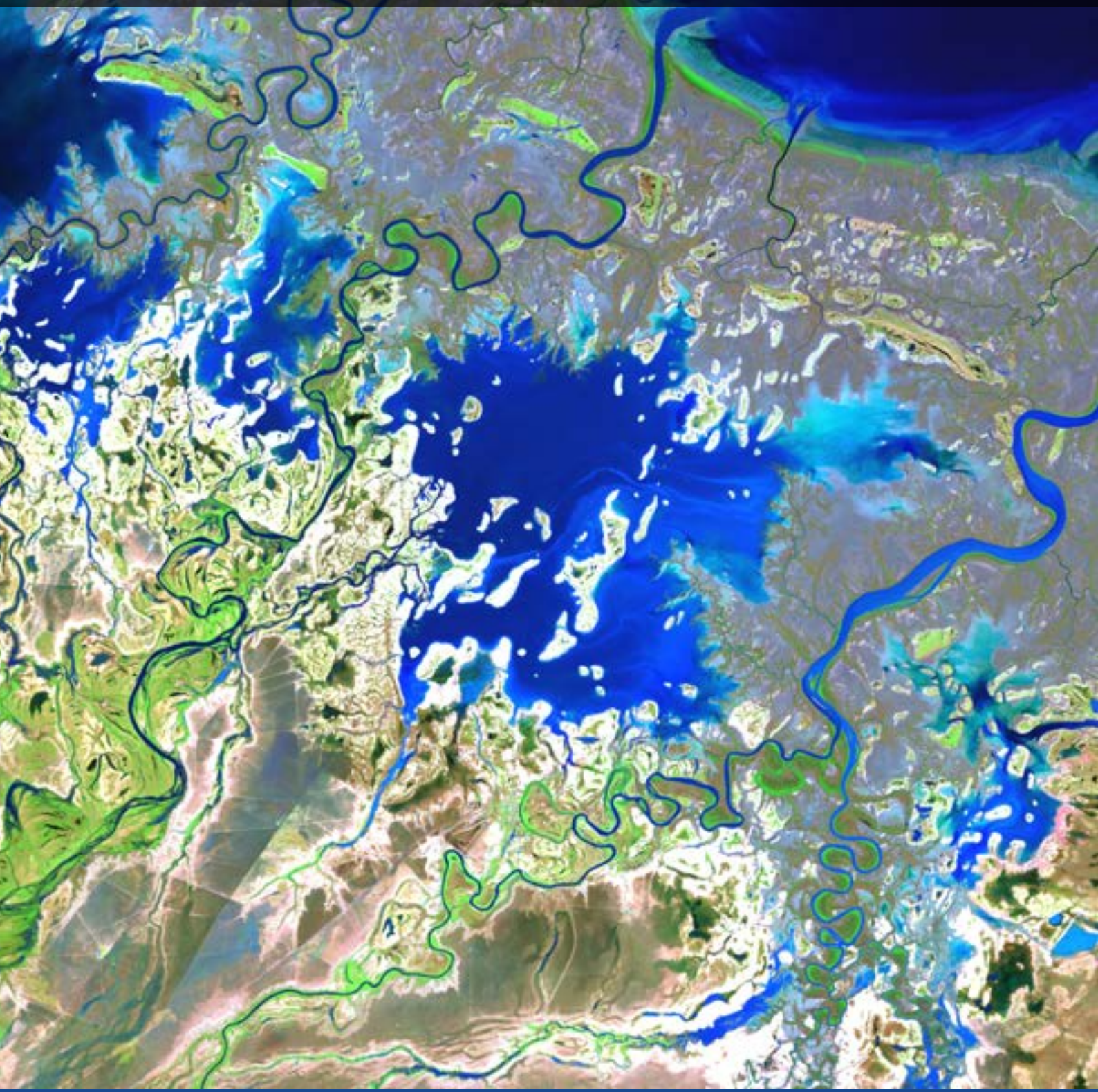
Price, J.C. (1987b). Calibration of satellite radiometers and the comparison of vegetation indices. *Remote Sensing of Environment*, 21, 15–27.

Rothery, D.A. (1987). Decorrelation stretching as an aid to image interpretation. *International Journal of Remote Sensing*, 8, 1253–1254.

Suits, G., Malila, W., and Weller, T. (1988). Procedures for using signals from one sensor as substitutes for signals of another. *Remote Sensing of Environment*, 25, 395–408.

Viscarra-Rossel, R.A., and Chen, C. (2011). Digitally Mapping the Information Content of Visible-Near Infrared Spectra of Surficial Australian Soils. *Remote Sensing of Environment*, 152, 1–14.

Non-Linear Operations



Sections 7, 8 and 9 discuss linear operations that are commonly applied to one to more EO image channels. In this context, addition and subtraction of channels are examples of simple linear operations. The following two sections are concerned with non-linear operations applied to pairs of channels in an image (see Section 1.4).

Non-linear operations include multiplication and division, with the most frequently used non-linear operation for image processing of EO imagery being ratioing of image bands (see Section 10). Multiplying image bands is not a commonly used operation, but may be appropriate for masking out regions of an image (Castleman, 1998) or for image modulation (Harrison and Jupp, 1990).

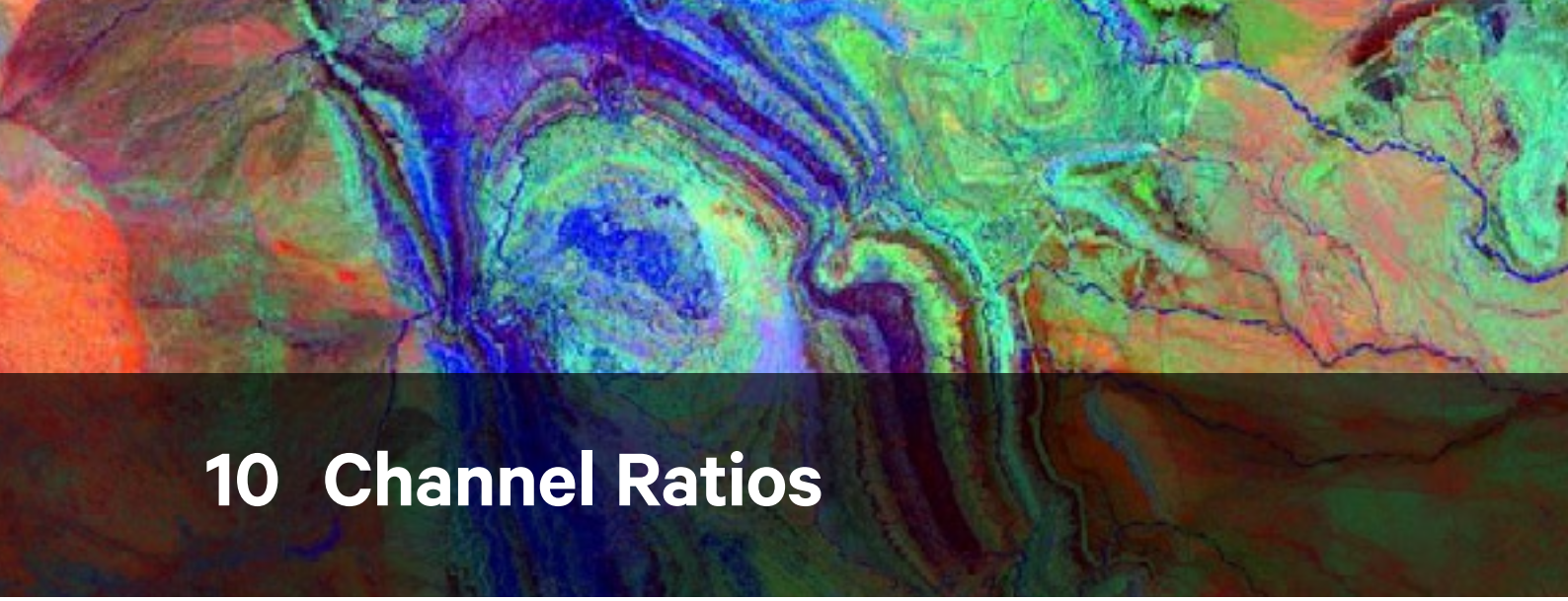
Many of the band indices that are popularly used with EO imagery rely on a combination of linear and non-linear operations. Some of the indices that have been developed to highlight and condense information in EO imagery relating to vegetation condition are introduced in Section 11. These indices are discussed in more detail, and in the context of specific application areas, in Volume 3A.

Contents

10 Channel Ratios 107

11 Vegetation Indices 121

Background image on previous page: Landsat-8 OLI image of Burketown, northern Queensland, acquired in on 5 January 2016.
Source: Craig Shephard, DSITI



10 Channel Ratios

A simple but effective transformation that is commonly used with EO imagery involves computing the ratio values between two image channels. The simplest implementation of this transformation involves dividing the value of a pixel in one channel by its value in a second channel and rescaling the result to the available image data range for all pixels in the image. The ratio value resulting from this calculation is generally rescaled according to user-defined parameters to fill the full image data range (see Section 2).

In this section ratioing of image channels will be discussed in terms of:

- computational methods—simple ratios, log ratios, smoothed ratios and directed ratios (see Section 10.1); and
- processing parameters—dark values and calibration impact (see Section 10.2).

Firstly, some ratios that are frequently used with EO imagery are introduced in Excursus 10.1.

Transformation literally means going beyond your form.
(Wayne Dyer)

Background image: Landsat TM colour ratio image over Flinders Ranges, SA, where the ferric iron (iron oxide) ratio (TM3/TM1) is shown as red, the ferrous minerals ratio (TM5/TM4) is shown as green, and the clay minerals ratio (TM5/TM7) is shown as blue. **Source:** Megan Lewis, University of Adelaide

Excursus 10.1—Commonly Used Ratios

While the ratioing of image channels is a very simple operation computationally, it is a very effective way of highlighting the differences between two channels. Ratios of various channels are used to highlight differences in particular features of the Earth's surface. As introduced in Volume 1, different land covers reflect and absorb EM radiation in different and often characteristic ways (see Figure 10.1 and Section 1.1). When image channels are ratioed, these differences are emphasised.

For example, healthy vegetation typically reflects NIR wavelengths strongly, but absorbs most wavelengths in the visible red region of the EM spectrum, whereas while NIR reflectance is relatively strong for most soils, the difference between their NIR and red reflectances is less dramatic. These differences are highlighted in a single channel when the ratio of NIR and red bands are computed. In this case a NIR/red ratio channel would allow soil and vegetation features to be clearly distinguished in the image.

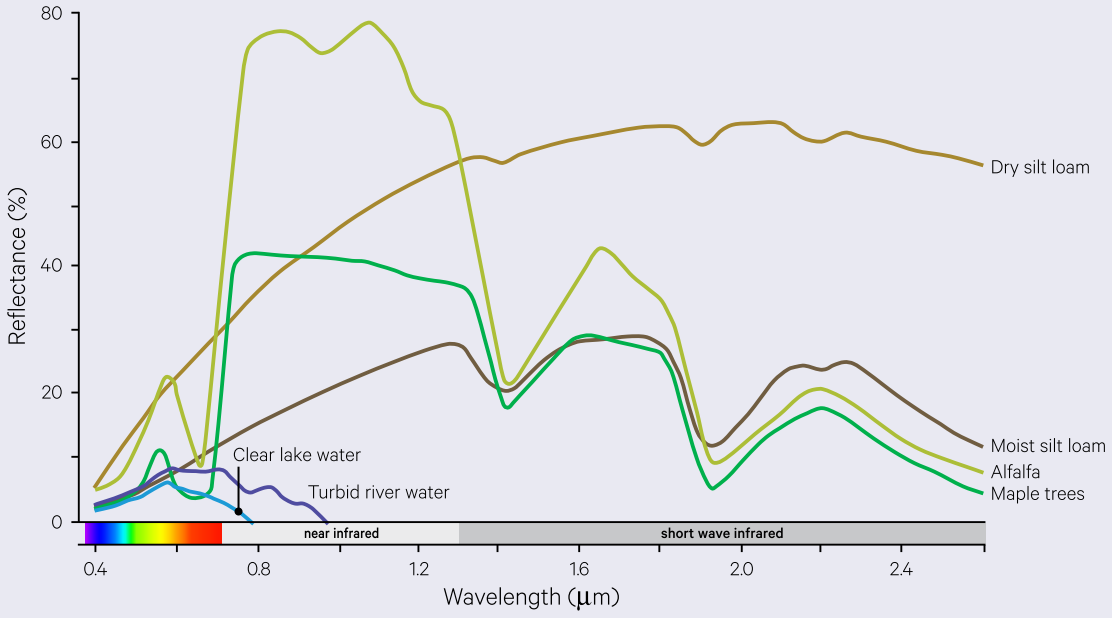
Some of the channel ratios that are commonly computed for EO imagery are summarised in Table 10.1. These ratios have been computed for a Landsat-8 OLI image of Lake Carnegie, WA. The original Landsat-8 OLI channels are shown in Figure 10.2 and the ratio channels are shown in Figure 10.3. Whereas several original pairs of channels (such as blue and green, and SWIR1 and SWIR2) are similar, the ratio channels highlight the differences between channels.

Table 10.1 Popular EO image channel ratios

Ratio channels		Highlights
Numerator	Denominator	
Green	Red	Soil colour Water colour
Green	Blue	Water colour
NIR	Red	Vegetation with green foliage
Red	Blue	Ferric iron
SWIR1 (~1.65 μm)	NIR	Ferrous minerals
SWIR1 (~1.65 μm)	SWIR2 (~2.22 μm)	Geological features

Figure 10.1 Idealised spectral reflectance curves

Reflectance from different Earth surface features varies for different wavelengths. Characteristic spectral reflectance curves can be created for different features to indicate a 'typical' shape over a particular range of wavelengths. This information can be used to differentiate between surface features on the basis of their observed reflectance characteristics.

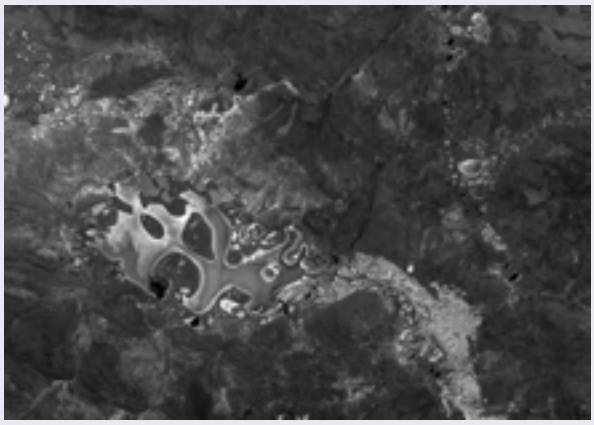


Adapted from: Harrison and Jupp (1989) Figure 6

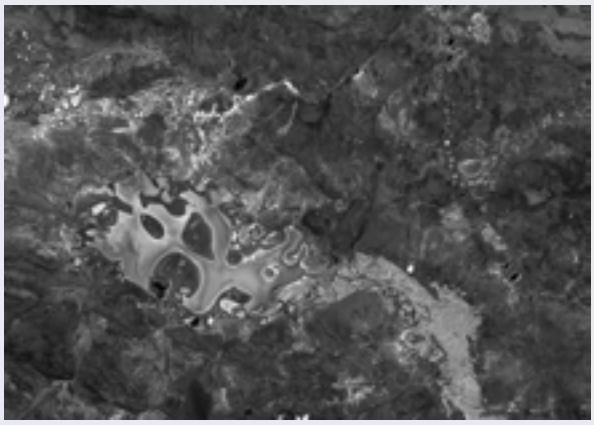
Figure 10.2 Landsat-8 OLI image channels

Six original image channels are shown for a Landsat-8 image over Lake Carnegie, WA, acquired on 8 July, 2017.

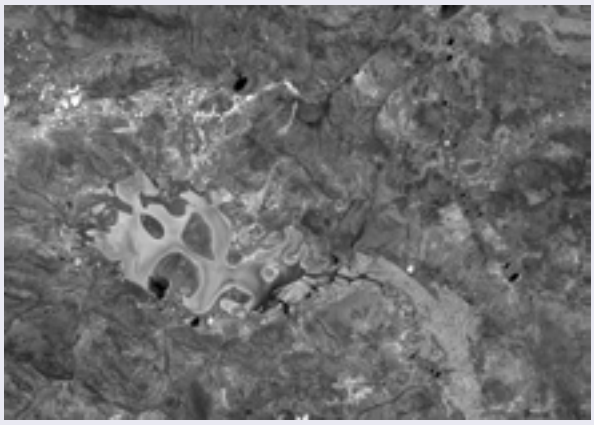
a. Blue channel (band 2: 0.452–0.512 μm)



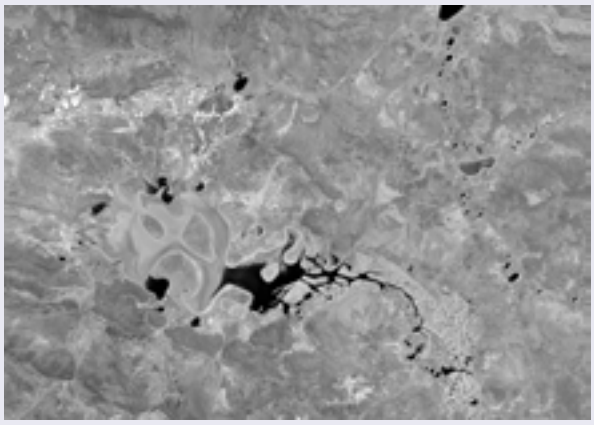
b. Green channel (band 3: 0.533–0.590 μm)



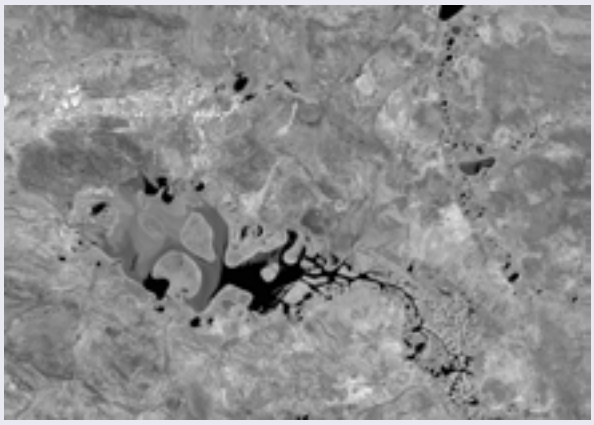
c. Red channel (band 4: 0.636–0.673 μm)



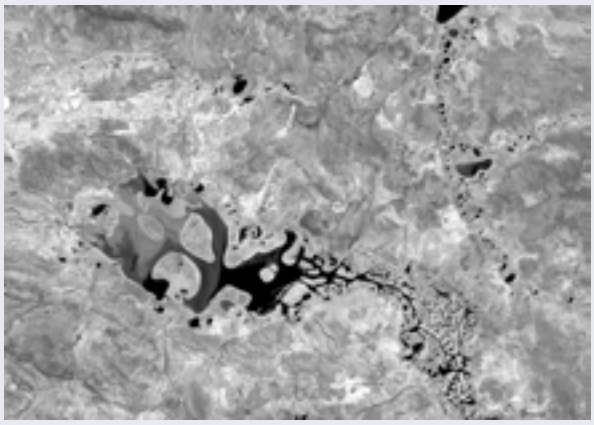
d. NIR channel (band 5: 0.851–0.879 μm)



e. SWIR1 channel (band 6: 1.566–1.651 μm)



f. SWIR2 channel (band 7: 2.107–2.294 μm)

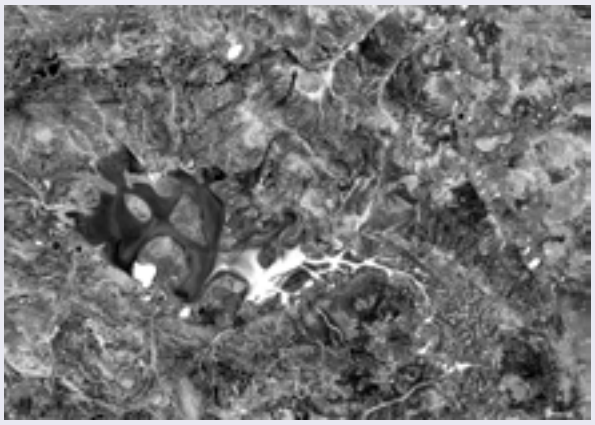


Source Norman Mueller, Geoscience Australia

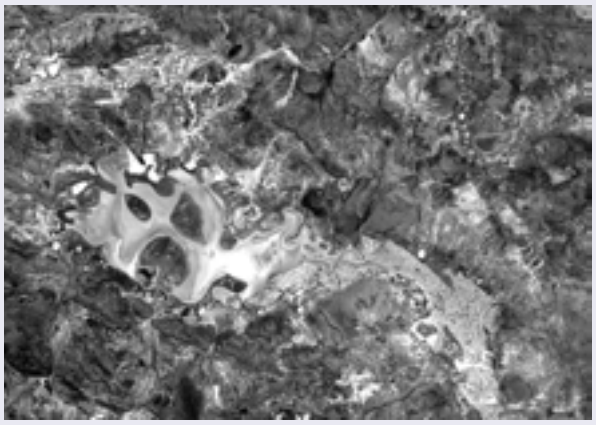
Figure 10.3 Ratio channels

Several ratio channels have been computed for the image channels shown in Figure 10.2.

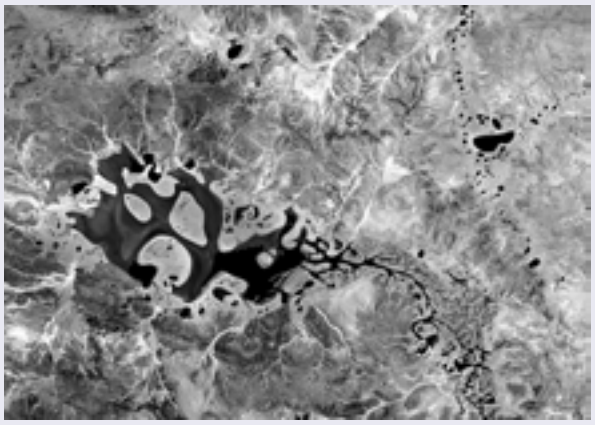
a. Ratio of green/red



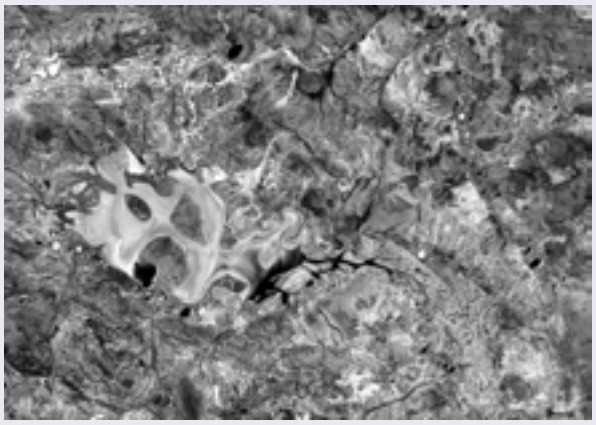
b. Ratio of green/ blue



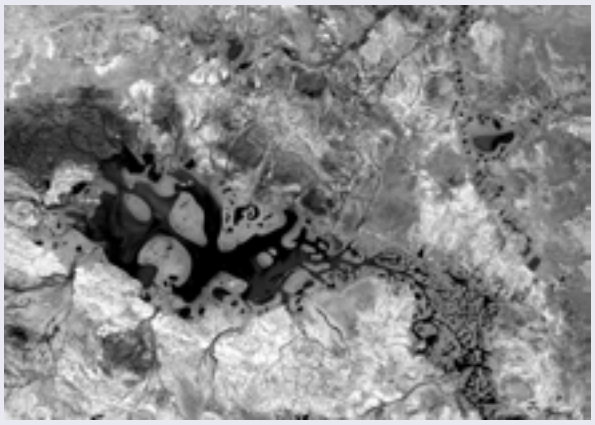
c. Ratio of NIR/red



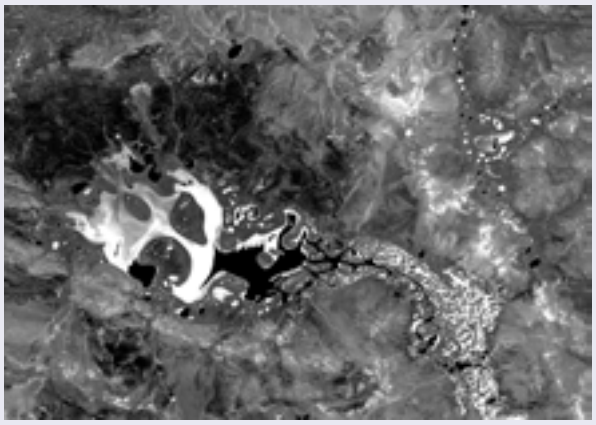
d. Ratio of red/ blue



e. Ratio of SWIR1/NIR



f. Ratio of SWIR1/SWIR2



Source Norman Mueller, Geoscience Australia

Colour composites of various ratio channels may be used to highlight particular surface features. For example, to maximise lithological spectral information in well-exposed terrain Crippen (1988) suggests a colour composite of selected Landsat TM ratio channels (see Figure 1.2 for band definitions):

- TM5/TM7 (SWIR1/SWIR2), which highlights clay (see Figure 10.4a), as red;
- TM5/TM4 (SWIR1/NIR), which highlights ferrous minerals, as green; and
- TM3/TM1 (red/ blue), which highlights ferric iron (iron oxide; see Figure 10.4b) as blue.

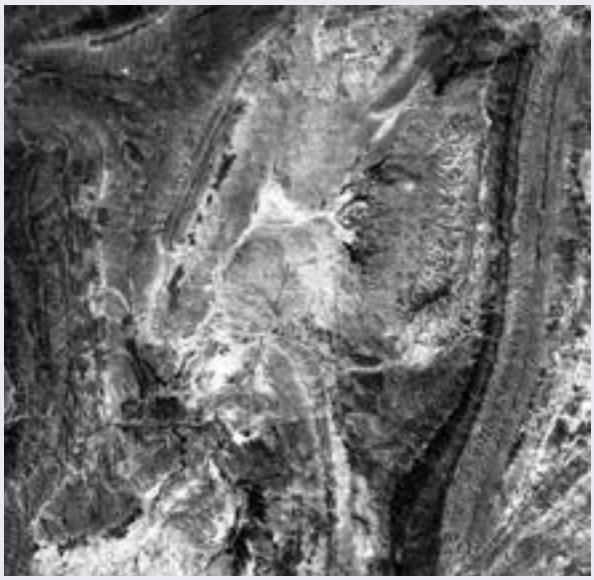
A variation on this combination is illustrated in Figure 10.5.

Figure 10.4 Ratios to highlight lithology

Two ratio channels are computed for a Landsat TM image of a portion of Flinders Ranges, SA.

a. Clay minerals (SWIR1/SWIR2)

b. Iron oxide (visible red/visible blue)



Source: Megan Lewis, University of Adelaide

Satterwhite and Henley (1987) evaluated a number of band ratioing techniques with Landsat TM data for discriminating spectral characteristics of arid region soils and vegetation conditions. Vogelmann and Rock (1988) found Landsat TM channel ratios of TM5/TM4 (SWIR1/NIR) and TM7/TM4 (SWIR2/NIR) correlated well with ground measurements of forest damage (% foliage loss in spruce-fir forests). A composite of the Landsat MSS bands MSS4/MSS5 (green/red) as blue, MSS5/MSS6 (red/NIR1) as green and MSS6/MSS7 (NIR1/NIR2) as red has been reported to enhance discrimination between altered rock types by many authors (such as Moik 1980; Drury 1987).

Chavez *et al.* (1982) discuss the selection of ratio channels to form colour composites for geological applications. These authors suggest an index of:

$$\frac{\sum \sigma}{\sum \text{absolute correlation values}}$$

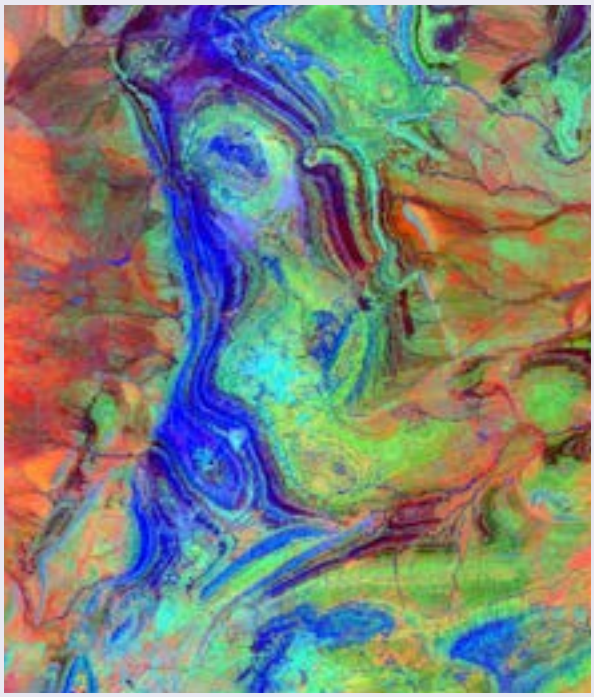
where

σ is the correlation coefficient (normalized covariance; see Section 1.3.3 and Volume 2A—Section 8.1.4)

be computed for each possible combination of channels, then the combination with the largest index value be selected. The resulting composite image would have the most variation within, and least correlation between, channels so should be the most visually informative (see Volume 2A—Section 5.2.2). This method confirmed the popular ratio composite suggested by Crippen (1988).

Figure 10.5 Colour composite of band ratios

Landsat TM colour ratio image over Flinders Ranges, SA, where the ferric iron (iron oxide) ratio (TM3/TM1) is shown as red, the ferrous minerals ratio (TM5/TM4) is shown as green, and the clay minerals ratio (TM5/TM7) is shown as blue.



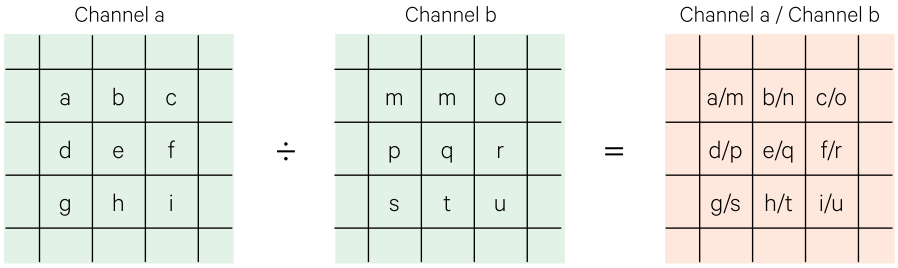
Source: Megan Lewis, University of Adelaide

10.1 Computation Methods

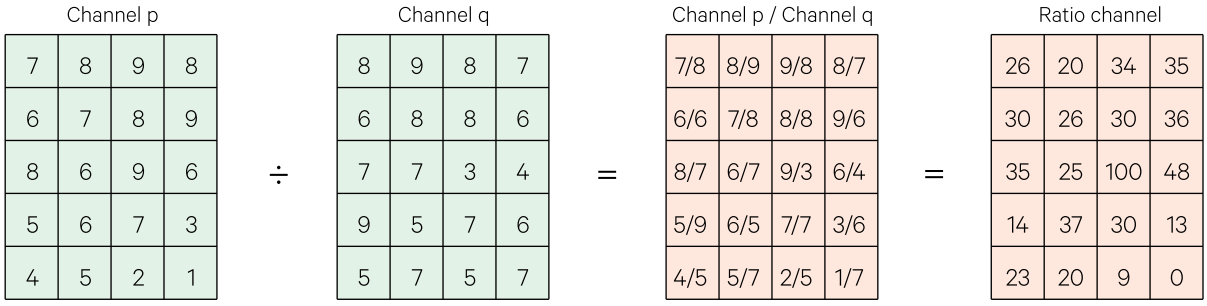
Channel ratios may be computed in a variety of ways. For example, a simple ratio may be computed between pairs of channels as shown in Figure 10.6 (see Section 10.1.1). Ratios of a pair of image channels may also be computed by differencing the logarithms of the channels (see Section 10.1.2). Alternatively a 'smoothed ratio' may be computed as detailed in Section 10.1.3. This produces a more 'stable' result by removing small variations in image values that would be accentuated by the ratioing process. Finally, a modified approach to computing channel ratios, called 'Directed Band Ratioing' (Crippen *et al.*, 1988), which enables the ratioed values to retain intensity information, is described in Section 10.1.4.

Figure 10.6 Operation of simple channel ratio

a. Process—each pixel value in channel a is divided by its corresponding pixel value in channel b then the resulting ratio value is rescaled in the output channel.



b. In this numeric example, the minimum ratio value ($\frac{1}{7}$) is rescaled to output channel value 0 and the maximum ratio value ($\frac{9}{3}$) is rescaled to value 100. In the output channel, values are computed as $\frac{\left(\text{ratio} - \frac{1}{7}\right) \times 100}{\left(\frac{9}{3} - \frac{1}{7}\right)}$



Source: Harrison and Jupp (1990) Figure 93

10.1.1 Simple ratios

Most image processing systems can construct multiple ratio channels from pairs of input channels. For each pair of corresponding pixels in the input images, this ratio is computed as:

$$\frac{x-a}{y-b}$$

where

- x and y are the paired pixel values in the input channels; and
- a and b are the respective 'dark' values for these channels.

The 'dark' value for an image channel represents the minimum 'data' response in that channel and corresponds to the darkest feature in the image. As detailed in Section 10.2.1, it provides an origin for computing the ratio of two channels such that values below the dark value are assumed to be due to atmospheric effects and/or sensor 'noise' (see Volume 1A—Section 13).

The ratio values resulting from this calculation are generally rescaled so that they fill the full image data range in the output image (see Figure 10.6b). The optimum rescaling parameters may need to be determined iteratively, and should be recorded if the ratio channel is to be subsequently used with models that relate ground measurements to image radiance (see Section 2). A wide distribution, but low frequency, of image values close to the minimum and maximum values in a channel can sometimes make this rescaling

operation difficult. This data-dependent problem can be resolved by reference to the ratio image histogram (see Section 1.3.2).

When a ratio channel has been computed to represent a larger range of ratios than is necessary for most pixels in the image, a suitable ratio range can be selected using the channel histogram then rescaled back to the actual ratio values for use in another processing run. This rescaling requires the expected ratio range (*a*,*b*) that was used to compute the ratio channel. Thus to rescale value *P* in the ratio channel to an actual ratio value we would compute:

$$\frac{P}{\text{image data range}} \times \left(b - a\right) + a$$

where

a is the minimum ratio value used to rescale the ratio channel; and
b is the maximum ratio value used to rescale the ratio channel.

The simple ratio transformation is computed quite rapidly and may be used to quickly determine ratio results for multiple channel pairs. The smoothed ratio, described in Section 10.1.3, is recommended for computation of the final ratio channel for feature analysis because the smoothed ratio process does not accentuate localised differences between channels as typically occurs in simple ratio channels.

Simple ratios are required however for processes which are attempting to reduce spatial 'noise' in an image. This noise may be due to uneven illumination effects, such as topographic shading. Since the latter has a similar effect on all image channels, shading differences are greatly reduced in ratioed channels (see Figure 10.7). This effectively means that shadow in one channel is being divided by shadow in another channel so that the resulting ratio channel retains significantly less 'shadow' information.

The ratio transformation effectively enhances differences between the channels and reduces similarities. Holben and Justice (1981) examined band ratioing as a technique to reduce topographic shading in Landsat MSS imagery, and reported a reduction in topographic effect by an average of 83%, with variations in this figure being due to Sun angle and surface slope.

Uneven illumination can be a problem with scan-digitised imagery. This may be compensated for in the same way as haze correction by using the scan of a blank target as the denominator or 'reference' channel (see Section 7.2.1). In this case, since the range of the reference channel is usually quite small, to preserve the data range of the image channels they should not be rescaled before ratioing and dark values should be given as zero.

Figure 10.7 Impact of ratioing on topographic shading

Topographic shading in EO imagery is significantly reduced in ratio channels. This Landsat TM image was acquired in June 1993 over Fowlers Gap Research Station, NSW.

a. Landsat TM band 4 shows marked topographic shadowing.

b. Ratio of Landsat TM4/TM7 reduces topographic shadowing in this image.

Source: Megan Lewis, University of Adelaide

10.1.2 Computing Products and Ratios using Log Channels

Ratioing of values in a linear scale is equivalent to the differencing of the corresponding values in a logarithmic scale (see Section 1.2.3), that is:

$$a \times b = \log(a) + \log(b)$$

and

$$\frac{a}{b} = \log(a) - \log(b)$$

Thus ratioing of image channels may be computed as the difference between log-transformed channels (see Section 2.2). The differencing operation could be computed using an affine transformation to produce log-ratio channels (see Section 7). When channel ratios are computed this way, dark values are specified as negative scaling (offset) values in the logarithmic transformation (see Section 10.2.1). The products of image channels can be derived in a similar way by adding logarithmically-transformed channels.

Log-ratio channels provide a useful enhancement for reducing topographic shading effects, since the log transformation reduces variation in the brightest areas of the image and provides greater contrast in the darker features. For example, Ahmad *et al.* (1989) used log-ratio channels to reduce the effect of topographic shading in Landsat MSS imagery of a mountainous area in northeast Tasmania without a digital terrain model.

10.1.3 Smoothed ratios

As well as computing the simple ratio result of individual pixel values in a pair of channels, it is possible to apply a smoothing process during the ratio calculation. This operation reduces the 'noise-enhancement' effects, which can occur in ratio channels due to small variations between channels. Gordon and Philipson (1986) suggest that ratioing selected texture channels would reduce their characteristic edge effects. In this case, a smoothed ratio would avoid increasing within-feature texture.

Some image processing systems can directly compute a 'smoothed ratio' channel from a pair of image channels. For example, the 3×3 neighbourhood of each pixel may be used to estimate the ratio as the solution to the linear equation:

$$c_i - d_i = r_{ij} \times (c_j - d_j)$$

where

c_i and c_j are the channels i and j being ratioed;
d_i and d_j are dark values (see Section 10.2.1) in channels i and j ; and
r_{ij} is the estimated ratio value.

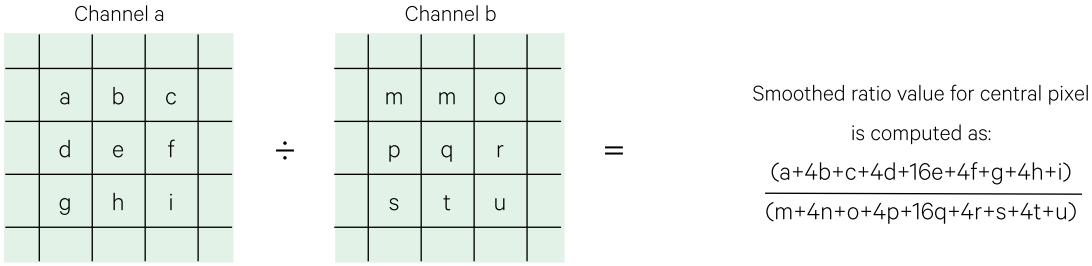
This operation is illustrated in Figure 10.8 (using the same minimum and maximum values as Figure 10.6 to enable comparison of results). The processing order, the non-linear nature of the ratioing operation, and the rescaling differences between processes, mean that this transformation produces slightly different results to either computing simple ratios of smoothed data channels or the smoothing of simple ratio channels using the same filter weights.

Some implementations of this process may also allow themes to be applied to the channels being ratioed. This option is a useful way of pre-segmenting the image into feature and non-feature areas and then computing a ratio value only for those pixels that are deemed to represent the selected feature. Pixels outside the specified theme are typically set to the default null value in the output image. These may be subsequently 'filled' using filter-based interpolation (see Section 4.6) if appropriate. However, it is usually recommended that broad, rather than narrow, themes be defined for this purpose to ensure all feature pixels are processed.

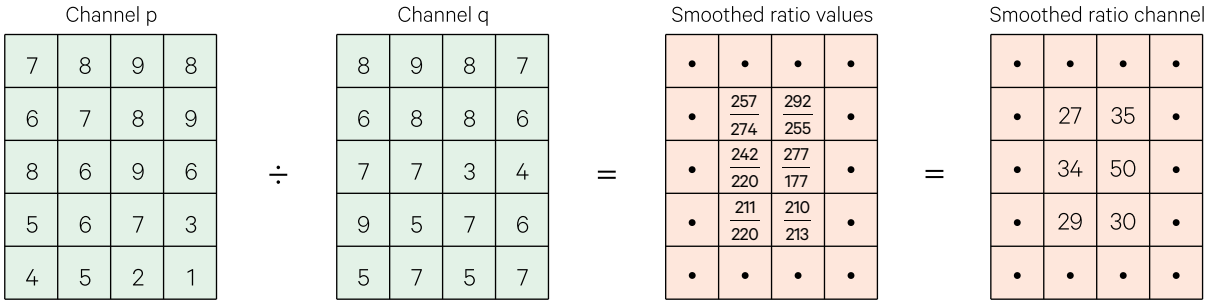
The additional computation required by the smoothing process means that this transformation is likely to be slower than the simple ratios described in Section 10.1.1. However, the results of this process provide a more 'stable' set of values, which highlight the major rather than minor variations between the channels. Thus, when ratio channels are being computed for density slicing or other interpretation purposes, it is highly recommended that they be produced as smoothed rather than simple ratios.

Figure 10.8 Smoothed ratio operation

a. Process—the 3x3 neighbourhood around each pixel pair is used to compute the smoothed ratio value, which is then rescaled in the output channel. Edge pixels have not been processed in this example.



b. For this numeric example, the minimum ratio value ($\frac{1}{7}$) is rescaled to output channel value 0 and the maximum ratio value ($\frac{9}{3}$) is rescaled to 100. In the output channel, values are computed as $\frac{\left(\text{smoothed ratio} - \frac{1}{7}\right) \times 100}{\left(\frac{9}{3} - \frac{1}{7}\right)}$



Source: Harrison and Jupp (1990) Figure 97

10.1.4 Directed band ratioing

Crippen *et al.* (1988) suggest a method called 'Directed Band Ratioing' (DBR) by which topographic and albedo information can be retained in colour composites formed from ratio channels. Various ratio channels are useful enhancements of geological information and the retention of geomorphic structure as relief shading can improve the interpretation of such imagery.

Directed band rationg involves retaining the illumination information as a consistent intensity variation in all ratio channels, so that, in a colour composite, the illumination components are correlated (channel invariant) and thus only vary as a grey-scale, while reflectance (possibly lithographic) variations are uncorrelated (channel variant) between ratios so are shown as colours. Thus, while topographic variations may mask reflectance information in individual channels, correlation of illumination variation between channels in a multi-channel composite separates topography as an achromatic feature from other coloured features. (This correlation could also be advantageous in multidimensional analyses such as classification or PCA to isolate topographic or albedo variations).

The procedure for directed band ratioing involves pre-processing the channels by:

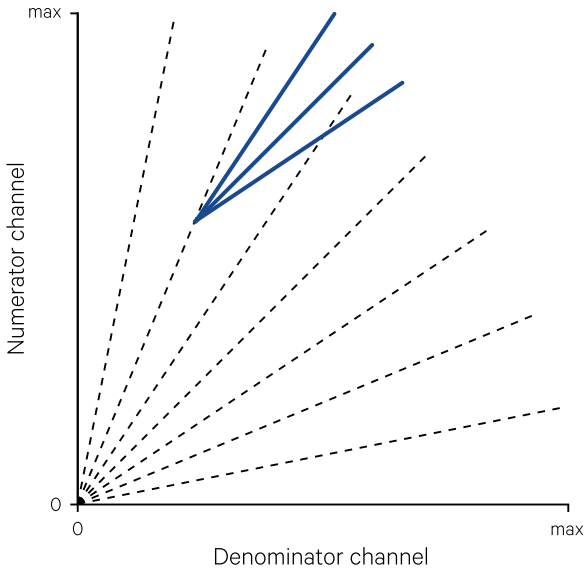
- subtracting appropriate dark values (see Section 10.2.1);
- multiplying the denominator channel by a constant;
- adding a constant to both channels; then
- ratioing the results.

The principle of this processing is to effectively increase the ratio 'angle' of reflectance variations relative to illumination variations in a channel crossplot (see Section 1.3.4 and Figure 10.9). The required pre-processing may be undertaken using an affine transformation (see Section 7.1) before ratioing.

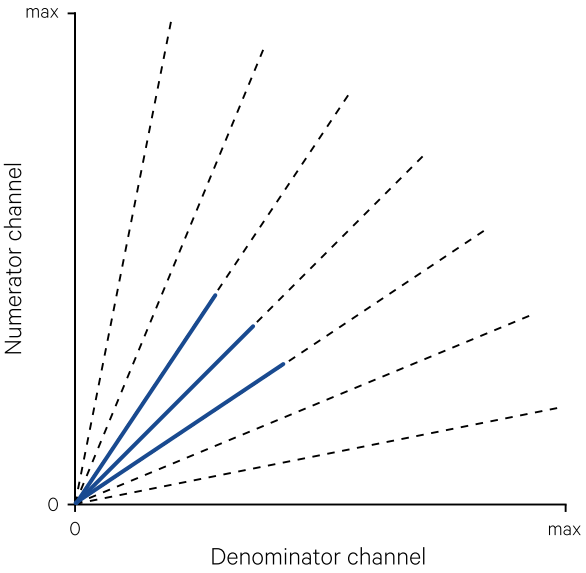
Figure 10.9 Directed band ratioing

This process increases the ratio 'angle' of reflectance variations in the image relative to the illumination variations. This preserves the topographic shading information, which is generally reduced by ratioing, but can provide useful information for visual interpretation, especially for geological applications.

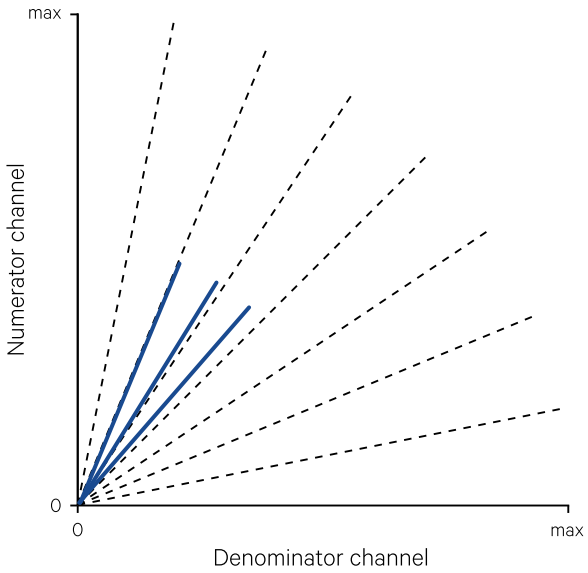
a. Crossplot of original data channels



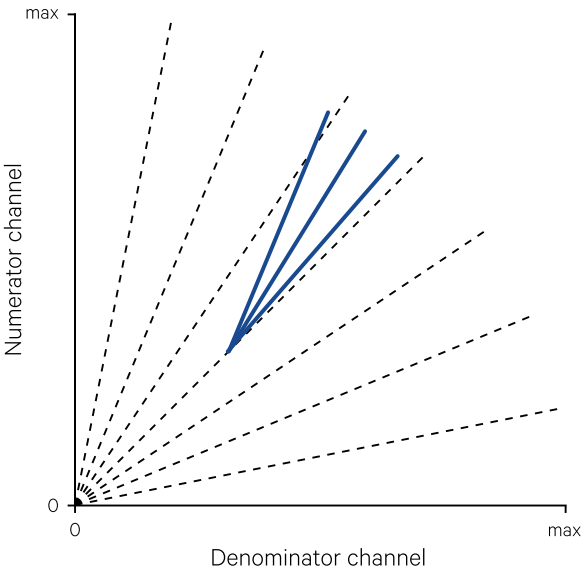
b. Ratios after dark value adjustment



c. The slope of data lines is increased by multiplying the denominator channel by a constant value.



d. Equal constants are added to each channel before ratioing



Source: Harrison and Jupp (1990) Figure 96 [Adapted from Crippen *et al.*, 1988]

10.2 Processing Parameters

Ratio variables are characterised by having a valid zero point (see Section 1.2.2). When ratioing a pair of channels in EO imagery, computations need to account for the true minimum value in each channel. This minimum value is commonly called the 'dark' value (see Section 10.2.1) as it effectively accounts for sensor 'dark current' (see Volume 1A—Section 13). When the channels in an image are not scaled consistently (due to variations in gains and/or calibration) the expected relationships between channels may not be observed in ratio results. Models to account for these variations are introduced in Section 10.2.2.

10.2.1 Dark values

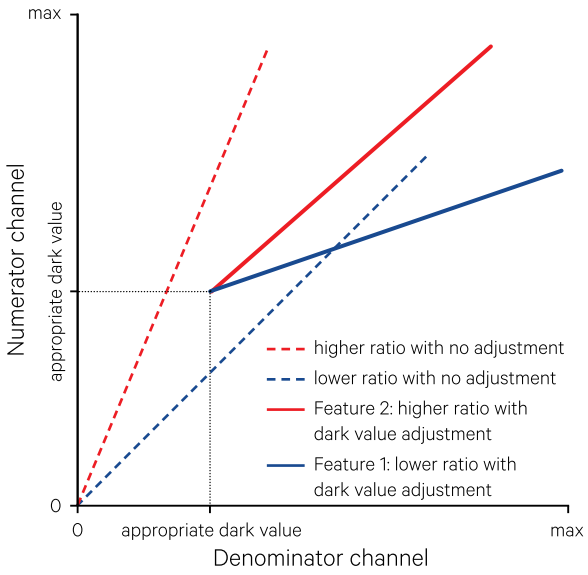
Ratio transformations generally subtract a constant 'dark value' (d_j) from each channel before the ratioing process. In EO data, the dark value level is assumed to be the upper threshold of the response in a channel that is due to atmospheric effects and sensor calibration offsets. This represents the minimum level of information in the channel (see Volume 1A—Section 13).

Dark values are so named as the values corresponding to the darkest features in the image, such as shadows (Crane, 1971), and may be derived by interactive training on such features (see Volume 2A—Section 9.1.2). A simple approximation of the dark value is one value less than the absolute minimum value of a channel. More sophisticated methods for determining dark values to correct for atmospheric path radiance were proposed by Switzer *et al*. (1981), Crippen (1987) and Chavez (1988).

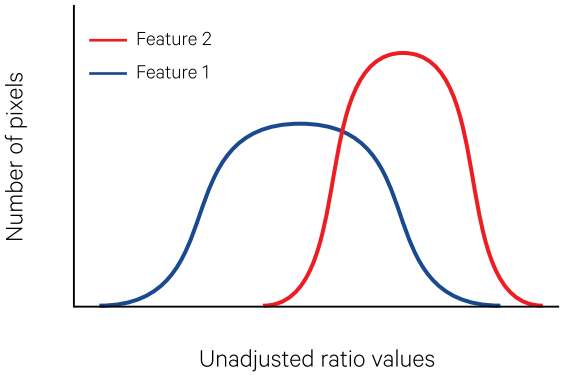
Crippen (1988) demonstrates the importance of appropriate selection of the dark value adjustment for producing useful ratio channels (see Figure 10.10). When image channels are not properly adjusted before ratioing, the resulting channel can contain undesirable overlap between the ratio ranges of different image features..

Figure 10.10 Effect of dark value adjustment in channel ratios

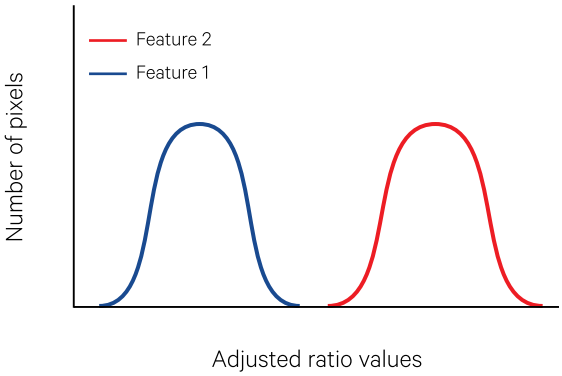
a. The origin of ratios, and hence their stratification of the image data space, is determined by dark value selection.



b. When the ratio range between features overlaps, they cannot be separated in the ratio channel.



c. Using appropriate dark values, the ratio channel can be computed so that the features have differing ratio ranges.



Source: Harrison and Jupp (1990) Figure 95 [Adapted from Crippen, 1988]

10.2.2 Calibration impact

The non-linear nature of the ratio operation accentuates any small variations in the spectral response between different channels so image channels are often smoothed before ratioing (see also Section 10.1.3). Conversely, the similarities between different channels are reduced by the ratioing operation. This effect may be used to correct for, or reduce, spatially variable 'noise' such as topographic shading in EO imagery or uneven illumination effects in scan-digitised imagery. These sources of noise (represented in the equation below by the term a_i) are related to ground reflectance by a multiplicative model of the form:

$$x_{ij} = a_i \times b_{ij} + d_j$$

where

- x_{ij} is the image radiance value at pixel i in channel j;
- b_{ij} is reflectance (or emissivity) of pixel i in channel j;
- a_i is illumination (or temperature) at pixel i; and
- d_j is a constant atmosphere or sensor calibration offset factor in channel j (or the 'dark value' as explained in Section 10.2.1 above).

These sources of noise can be removed by the ratio operation using an appropriate reference channel (x_{ir}) that represents the noise pattern:

$$\frac{x_{ij} - d_j}{x_{ir} - d_r} = \frac{a_i \times b_{ij}}{a_i \times b_{ir}} = \frac{b_{ij}}{b_{ir}}$$

where subscript r refers to the reference channel.

Radar backscatter and scanner angular effects can similarly be corrected using ratioing techniques. A multiplicative model applies when the detected noise level is affected by target reflectance, that is, the noise factor is not simply added to the reflectance value of the target. (Where the spatial noise is related to ground reflectance by a linear or additive model, as is the case with atmospheric noise, it should be removed using an affine transformation—see Section 7.) For scan-digitised imagery, a 'reference' channel is usually scanned from a blank target and this is used as the denominator channel when computing channel ratios. Pre-processing and selection of suitable reference (denominator) channels for noise correction is discussed in Section 7.2.

10.3 Further Information

Gonzalez and Woods (2018)

Jensen (2016) Section 8

10.4 References

Ahmad, W., Jupp, D.L.B., and Nunez, M. (1989). Land cover mapping in a rugged terrain area using Landsat MSS data. *International Journal of Remote Sensing*, 13(4), 673–683. *https://doi.org/10.1080/01431169208904145*

Castleman, K.R. (1998). *Digital Image Processing*. 2nd edn. Prentice-Hall, Inc. 667 pp.

Chavez, P.S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24, 459–479.

Chavez, P.S., Berlin, G.L., and Sower, L.B. (1982). Statistical method for selecting Landsat MSS ratios. *Journal of Applied Photographic Engineering*, 8, 23–30.

Crane, R.B. (1971). Preprocessing techniques to reduce atmospheric and sensor variability in multispectral scanner data. *Proceedings of 7th International Symposium on Remote Sensing of Environment*, Ann Arbor, Michigan, 17–21 May, 1345–1355.

Crippen, R.E. (1987). The regression intersection method of adjusting image data for band ratioing. *International Journal of Remote Sensing*, 8, 137–155.

Crippen, R.E. (1988). The dangers of underestimating the importance of data adjustments in band ratioing. *International Journal of Remote Sensing*, 9, 767–776.

Crippen, R.E., Blom, R.G., and Heyada, J.R. (1988). Directed band ratioing for the retention of perceptually-independent topographic expression in chromaticity-enhanced imagery. *International Journal of Remote Sensing*, 9, 749–765.

Drury, S.A. (1987). *Image Interpretation in Geology*. Allen and Unwin (Publ) Ltd, London.

Gonzalez, R.C., and Woods, R.E. (2018) *Digital Image Processing*. Pearson Educational Inc., New York.

Gordon, D.K., and Philipson, W.R. (1986). A texture-enhancement procedure for separating orchard from forest in Thematic Mapper data. *International Journal of Remote Sensing*, 7, 301–304.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Holben, B., and Justice, C. (1981). An examination of spectral band ratioing to reduce the topographic effect on remotely sensed data. *International Journal of Remote Sensing*, 2,115–133.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Levine, M.D. (1985). *Vision in Man and Machine*. McGraw-Hill Inc. USA.

Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images*. NASA SP–431. Washington, DC. USA.

Openheim, A.V., and Schafer, R.W. (1975). *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey.

Satterwhite, M.B., and Henley, J.P. (1987). Spectral characteristics of selected soils and vegetation in northern Nevada and their discrimination using band ratio techniques. *Remote Sensing of Environment*, 23, 155–175.

Switzer, P., Kowalik, W.S., and Lyon, R.J.P. (1981). Estimation of atmospheric path-radiance by the covariance matrix method. *Photogrammetric Engineering and Remote Sensing*, 47, 1469–1476.

Vogelmann, J.E., and Rock, B.N. (1988). Assessing forest damage in high elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment*, 24, 227–46.

11 Vegetation Indices

Various indices have been proposed to highlight and differentiate attributes of vegetation quantity and condition (Bannari *et al.*, 1995; Xue and Su, 2017). The following sections review the most commonly encountered indices used to measure the vegetation attributes of greenness (see Section 11.1) and water content (see Section 11.2). Other indices related to vegetation productivity, biochemistry or structure are described in Volume 3a.

11.1 Vegetation Greenness

Various indices have been proposed to highlight and differentiate vegetation 'greenness'. These vegetation indices have been well reviewed and compared by many authors, including Jackson (1983), Perry and Lautenschlager (1984), and Glenn *et al.* (2008).

The green colour of vegetation results from

- absorption of red and blue wavelengths by the chlorophyll pigment for photosynthesis; and
- reflection of green wavelengths.

These reflectance characteristics are illustrated in Figure 11.1 (see Volume 1B—Section 6.1 for details). However, it is the combined 'colour' resulting from the chlorophyll content of foliage, the area covered by leaves, and the density and structure of the canopy that determines the optical 'greenness' for a given patch of vegetation as measured by a remote sensor.

Healthy green vegetation typically has high reflectance in NIR and low reflectance in red wavelengths (see Figure 10.1 and Volume 1). Other image features may have either high NIR or low red values so it becomes difficult to identify vegetation pixels using only one of these channels. Most vegetation indices are based on highlighting the difference between pixel values in the red and NIR channels.

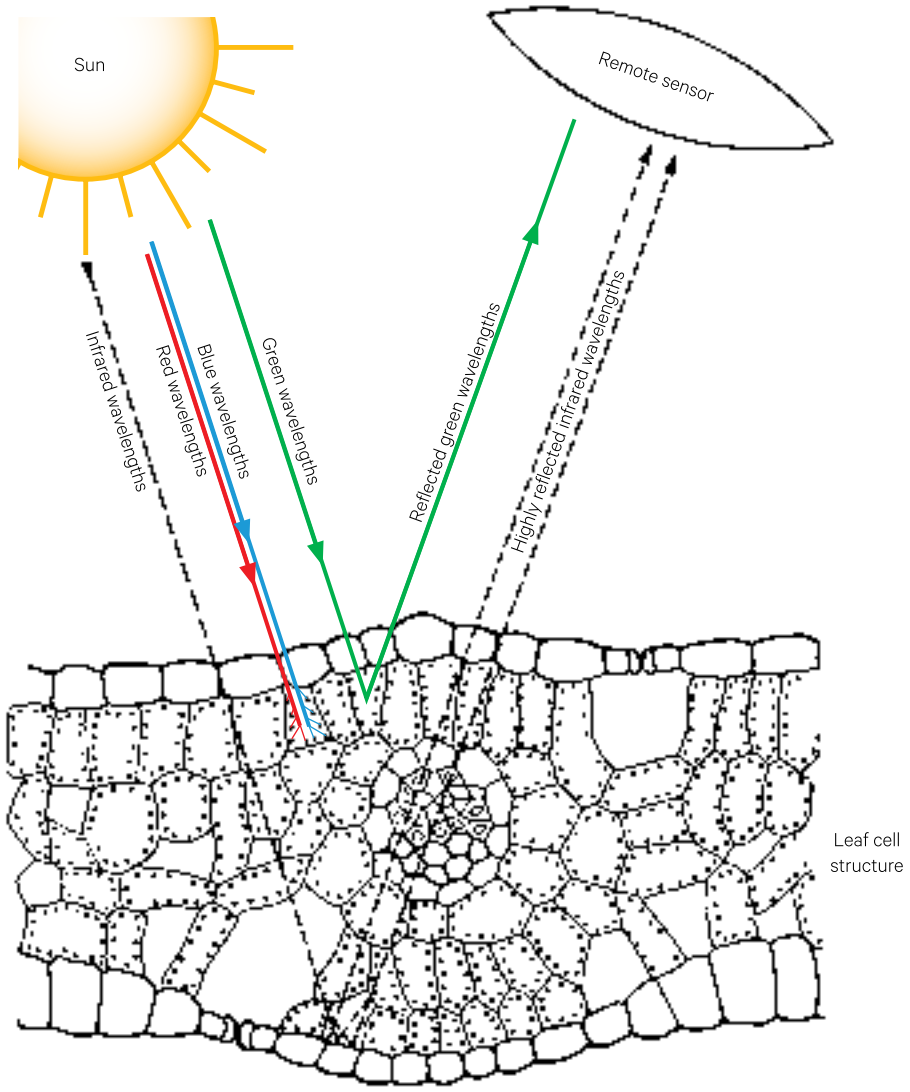
Reflectance values in NIR and red can be either differenced or ratioed to emphasise those pixels that are likely to contain a high proportion of photosynthetic material. Transformation results may be normalised for comparison between multiple images or differenced between two images to show temporal changes in greenness (see Volume 2D). Variations in vegetation index values over time have been used to monitor a range of changes in vegetation cover and condition (see Volume 3A), including:

- crop growth and senescence—greenness indices are commonly used to monitor crop phenology;
- seasonal changes in plant growth—the annual cycle of vegetation change can be tracked using greenness indices;
- land clearing and regeneration—greenness indices emphasise the difference between vegetated and cleared landscapes;
- fire impact—since fire reduces the amount of green vegetation, greenness indices have been used to map fire footprints, infer burn severity and monitor recovery; and
- fire risk—greenness indices are also used to determine fire risk by mapping the curing status of grasslands and the fire-prone nature of woody vegetation (Chuvieco *et al.*, 2003; Lasaponara, 2005).

Background image: High resolution aerial image over Werribee Farm (or Western Treatment Plant), Victoria, acquired on 16 January 2018. This sewage treatment farm has been processing effluent from Melbourne since 1897 and now contains a network of lagoons, wetlands, inter-tidal and shoreline areas that provide a haven for thousands of birds. **Source:** © EagleView

Figure 11.1 Reflectance characteristics of typical green leaf structure

Chloroplasts reflect green wavelengths but absorb blue and red wavelengths for use in photosynthesis. Spongy mesophyll cells strongly reflect infrared wavelengths.



Source: Harrison and Jupp (1989) Figure 5

The following three vegetation indices are commonly encountered in EO analyses:

- Simple Ratio (SR—see Section 11.1.1);
- Normalised Difference Vegetation Index (NDVI—see Section 11.1.2); and
- Kauth-Thomas greenness transformation (or 'Tasselled Cap' Greenness, TCG; see Section 11.1.3).

Applications using vegetation indices, including non-green leaf attributes, are further considered in Volume 3A. Some of the most commonly used vegetation indices are compared in Excursus 11.1.

Excursus 11.1—Comparison of Vegetation Indices

Source: Norman Mueller, Geoscience Australia

Two Landsat-8 OLI images over the diverse agricultural district surrounding the rural township of Kerang in northern Victoria are shown in Figure 11.2. These images contrast the variation in regional vegetation cover and condition between summer and winter in 2018 and will be used to compare four popular vegetation indices. The individual data channels for these example images are shown in Figure 11.3. In this climate, vegetation generally browns off in summer, but predictably the irrigated sections of this district show less seasonal difference than the non-irrigated.

Figure 11.2 Summer and winter example images

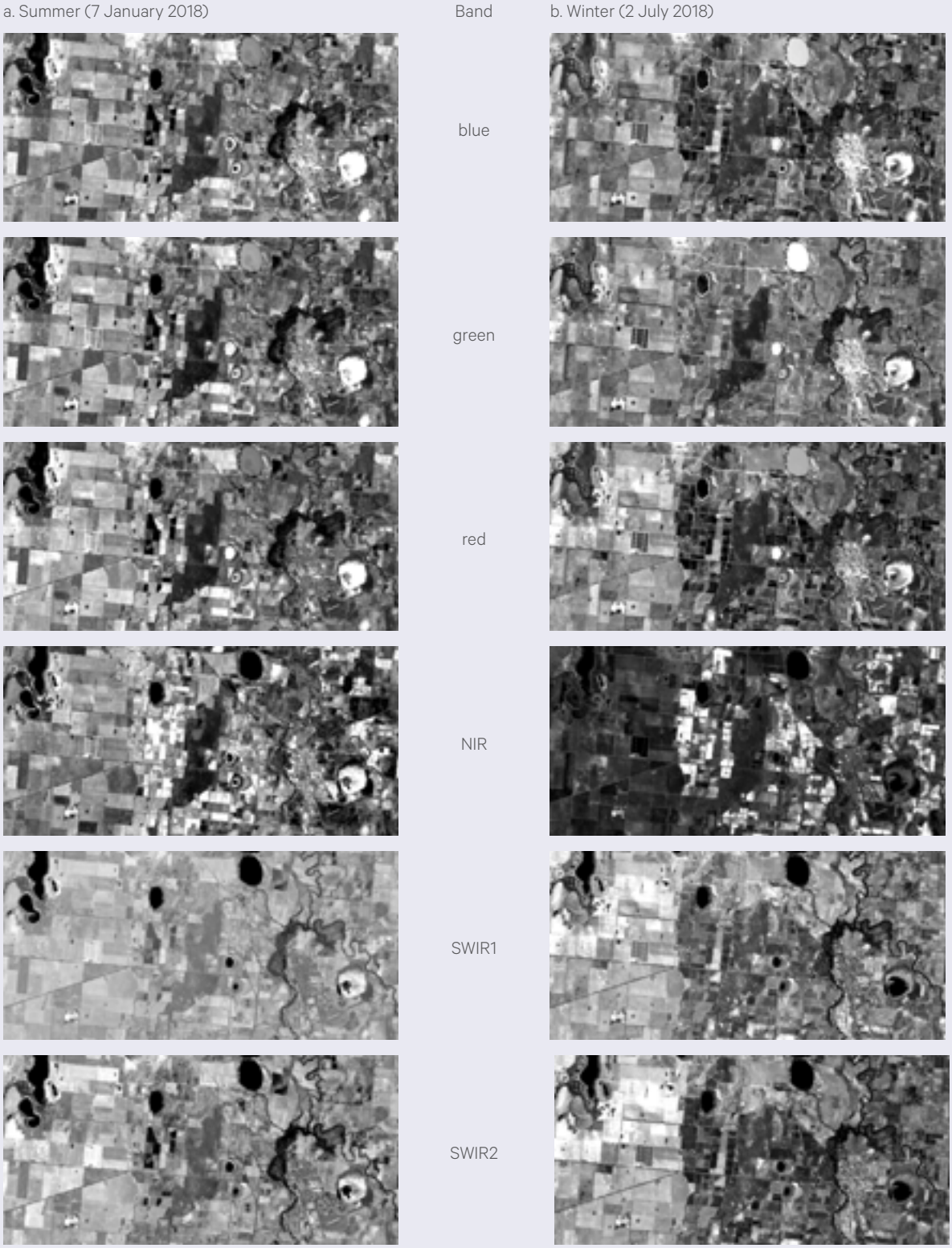
These example images were acquired by Landsat-8 OLI over Kerang, northern Victoria, and are displayed as natural colour composites using bands 4, 3, 2 as RGB. Some parts of this agricultural district are irrigated, so show less seasonal difference than others, but in this climate vegetation generally browns off in summer.

a. Summer (7 January 2018)

b. Winter (2 July 2018)

Figure 11.3 Individual channels for example images

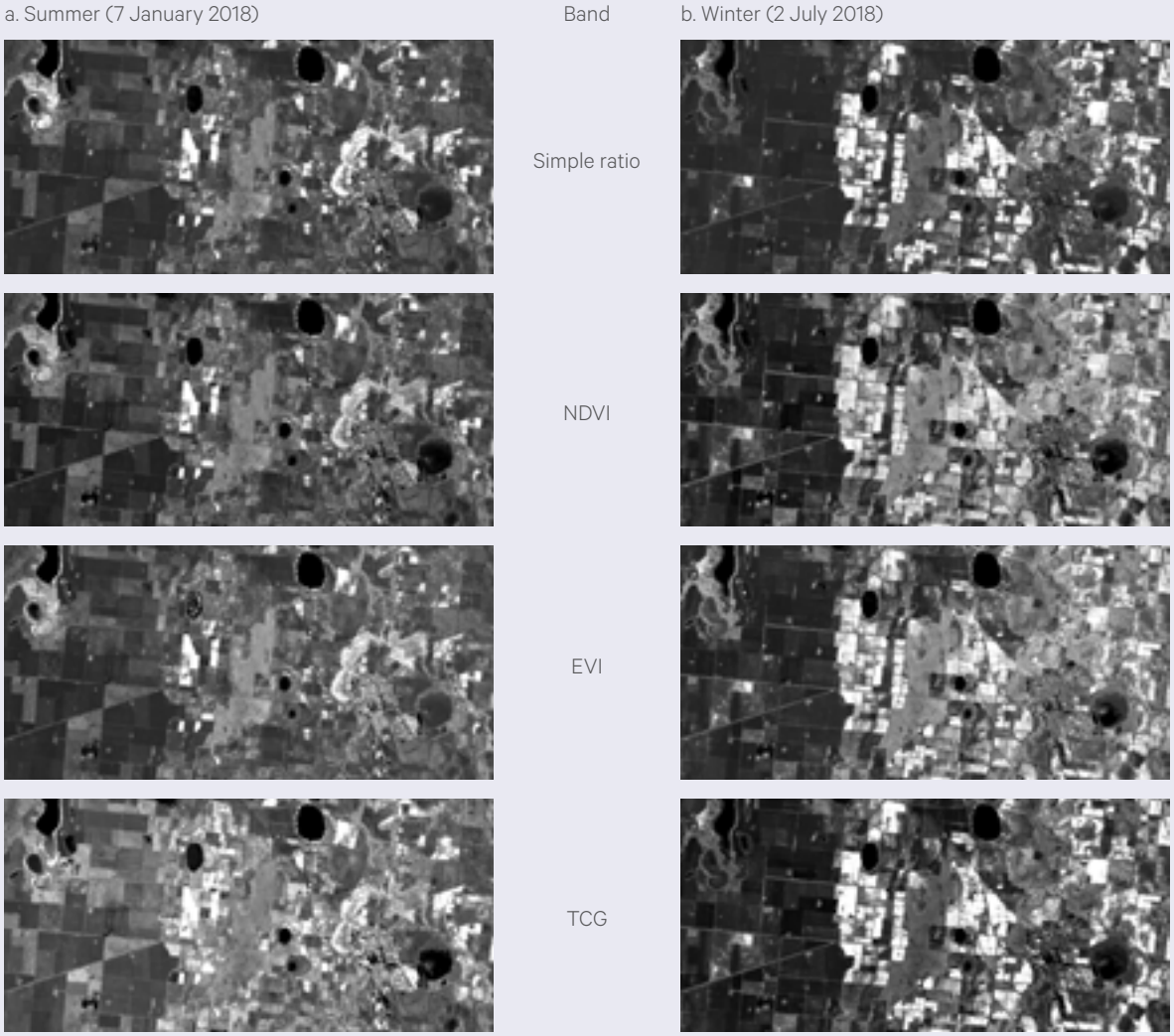
The most obvious differences between summer and winter images occur in the green and NIR channels, corresponding to changes in vegetative cover and condition.



Four different vegetation indices were computed for these two images. The resulting transformations are shown in Figure 11.4. All indices clearly highlight the differences between vegetated and non-vegetated pixels, but visible differences occur both within patches of green vegetation and between green vegetation and other land covers. The extent of green vegetation in the two seasons is clearly indicated by all indices.

Figure 11.4 Four vegetation indices for example images

Four vegetation indices are shown for the pair of example images in Figure 11.2: the simple ratio (NIR/red), Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and 'Tasselled Cap' Greenness (TCG). While all indices clearly highlight pixels containing green vegetation, there are differences in the way the index values are scaled.



Vegetation is the basic instrument the creator uses to set all of nature in motion.
(Antoine Lavoisier)

11.1.1 Simple Ratio

Dividing a NIR channel value by the red channel value for each pixel in the image gives us a ratio result in which high values only occur in pixels containing vegetation (see Figure 11.4). For vegetation, the strengths of both the NIR reflectance and the red absorption are indicative of plant vigour or type (see Volumes 1 and 3A). Consequently the ratioed values computed from these wavelength bands not only identify vegetative features but also allow us to stratify condition classes within those features. By summarising the ratio result as a single channel we can subsequently use interactive density slicing to define and highlight these classes (see Volume 2A—Section 9.2.1 and Volume 2E).

The direct ratio of NIR to red reflectance (NIR/red) is often referred to as the simple ratio (SR) vegetation index. It is commonly used as a measure of green biomass, but can also indicate vegetation water content (Tucker *et al*. 1980a) and crop condition (Tucker *et a*l. 1980b, Thompson and Wehmanen 1979). As well as vegetation 'greenness' or condition mapping, this index has been related to yield of agricultural crops using either logarithmic or linear regression analysis (see Volume 3A).

11.1.2 Normalised Difference Vegetation Index (NDVI) and variants

One of the most commonly used vegetation indices is the Normalised Difference Vegetation Index (NDVI), which is computed as:

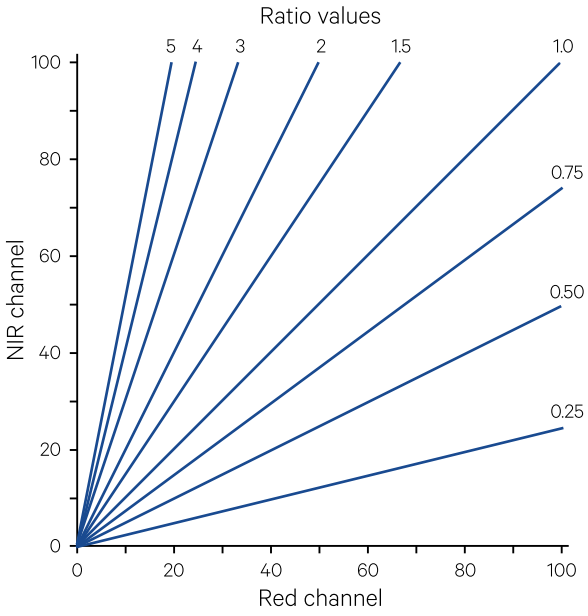
$$\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}$$

where ρ_{NIR} and ρ_{red} indicate the reflectance values for each pixel in the NIR and red bands respectively. As illustrated in Figure 11.5, the normalising of NDVI values results in a controlled range of raw ratio values between 1 and -1 and thus greatly simplifies the scaling of output data and comparison of multi-temporal imagery (Mather, 1987).

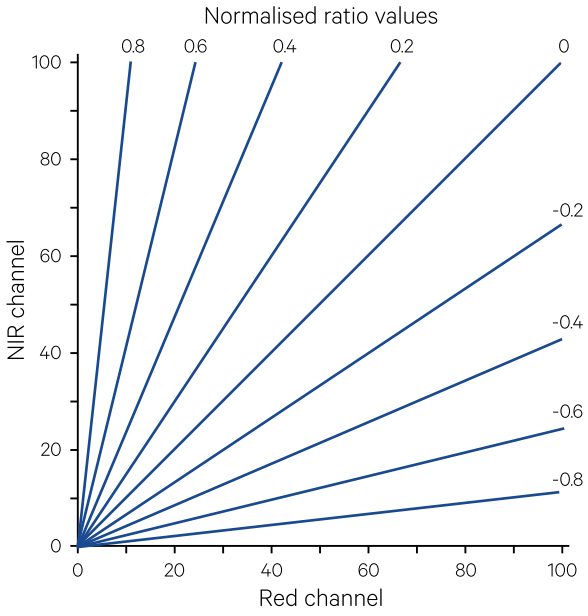
NDVI has been related to a range of vegetation characteristics, including cover and condition, which may indicate crop yield in agricultural systems (see Volume 3A). Studies have reported that NDVI saturates for dense vegetation (Carlson and Ripley, 1997; Huete *et al*, 1985). It has been observed to reach its maximum value for fractional covers as low as 60% (Jiang *et al*., 2006) and is adversely affected by bare soil backgrounds when mapping sparse vegetation in arid and semi-arid environments (see also Volume 3A).

Figure 11.5 Normalisation effect in 'Normalised Difference Vegetation Index' (NDVI)

a. Simple ratio (y/x)



b. Normalised ratio ((y-x)/(y+x))



Source: Harrison and Jupp (1990) Figure 94 [Adapted from Mather 1987]

Various modifications to the basic index have been proposed to correct for atmospheric or soil-induced variations (such as Huete 1988; Singh and Saull 1988). Some NDVI variants are summarised in Table 11.1 and detailed in Volume 3A. In particular, Enhanced Vegetation Index (EVI; Huete *et al.*, 1999; 2002; Liu and Huete, 1995; Jiang *et al.*, 2008) attempts to reduce saturation by accounting for effects due to atmosphere and background, though not necessarily topographic effects (Matsushita *et al.*, 2007):

$$EVI = G \times \left(\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \times \rho_{red} + C_2 \times \rho_{blue} + L} \right)$$

where

ρ_{red} and ρ_{NIR} are corrected values for MODIS bands 1 (red) and 2 (NIR) respectively;

ρ_{blue} is the corrected value for MODIS band 3 (blue);

G is a gain factor (G=2.5; Huete *et al.*, 1994; 1997);

L is an adjustment for the canopy background (L=1); and

C_1 and C_2 are coefficients for aerosol resistance (C_1=6, C_2=7.5).

Table 11.1 NDVI variations

Name	Equation	Advantages	Reference
Transformed Vegetation Index (TVI)	$TVI = \sqrt{\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} + 0.5}$	avoids operating with negative NDVI values and introduces normal distribution	Deering *et al.* (1975)
Enhanced Vegetation Index (EVI)	$EVI = G \times \left(\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \times \rho_{red} - C_2 \times \rho_{blue} + L} \right)$ where G is a gain factor; C_1 and C_2 are aerosol resistance coefficients; and L is the canopy background adjustment	reduces saturation by accounting for effects due to atmosphere and background	Huete *et al.* (2002)
Atmospherically Resistant Vegetation Index (ARVI)	$ARVI = \frac{\rho_{NIR} - \left(\rho_{red} + \gamma\left(\rho_{blue} - \rho_{red}\right)\right)}{\rho_{NIR} + \left(\rho_{red} + \gamma\left(\rho_{blue} - \rho_{red}\right)\right)}$ where usually $\gamma \sim 1$	reduces atmospheric effects	Kaufman and Tanre (1992)
Soil Adjusted Vegetation Index (SAVI)	$SAVI = \left(\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red} + L} \right)\left(1 + L\right)$ where L is a soil brightness correction factor (0 for high vegetation cover, 1 for low vegetation cover)	reduces the effects of soil background	Huete (1988)
Optimised Soil Adjusted Vegetation Index (OSAVI)	$OSAVI = \left(1 + L\right)\left(\frac{\rho_{800} - \rho_{670}}{\rho_{800} + \rho_{670} + L} \right)$ where L = 0.16	minimises soil effects	Rondeaux *et al.* (1996)
Visible Atmospherically Resistant Index (VARI)	$VARI = \left(\frac{\rho_{green} - \rho_{red}}{\rho_{green} + \rho_{red} - \rho_{blue}} \right)$	minimally sensitive to atmospheric effects	Gitelson *et al.* (2002)
Absolute Greenness index (Gabs)	$Gabs = \left(\frac{\text{observed NDVI} - \text{min NDVI}}{\text{max NDVI}} \right)$	distinguishes the weather-related variations in NDVI from longer term temporal variability (such as curing) within a selected time period	Eidenshink *et al.* (1990)
Relative Greenness index (RGI)	$RG = \left(\frac{\text{observed NDVI} - \text{min NDVI}}{\text{max NDVI} - \text{min NDVI}} \right) \times 100$	ranks the current NDVI of each pixel relative to its minimum and maximum NDVI over a specified period of time and better distinguishes the volume of dead fuel from live fuel	Burgan *et al.* (1998)
Photochemical Reflectance index (PRI)	$PRI = \left(\frac{\rho_{570} - \rho_{531}}{\rho_{570} + \rho_{531}} \right)$	assesses photosynthetic radiation use efficiency	Peñuelas *et al.* (1997)

EVI uses data from MODIS band 3 (500 nm) to remove the effects of smoke, aerosols and thin clouds, and masks out cloud, cloud shadows, marine (as opposed to inland) waters, and aerosols (Huete *et al.*, 2011). Accordingly, it is more sensitive to variations in canopy structure than NDVI (Gao *et al.*, 2000), and thus has been more responsive for land covers with high biomass, such as tropical forests (Huete *et al.*, 2002) and crops (Wardlow *et al.*, 2007). Both NDVI and EVI products are routinely and freely available at global and regional scales (see Volume 2D). Their use for phenological studies is discussed in Volume 3A.

The narrow bandwidths in hyperspectral imagery have enabled a range of indices that highlight subtle changes in vegetation type, health, and density, such as the Red Edge Normalised Difference Vegetation Index (RENDVI; Sims and Gamon, 2002). This index replaces the red and NIR reflectances in NDVI with the red edge peak and red edge vale respectively.

Vegetation indices have also been tailored for fuel modelling. The Fire Potential Index (FPI), for example, used RGI of AVHRR imagery for regional scale fire danger assessment (Burgan *et al.*, 1998) and demonstrated strong correlation with actual fire occurrence and predicting fire potential (see Volume 3A).

11.1.3 Kauth-Thomas Greenness Transformation

Another vegetation index, commonly known as the Kauth-Thomas or Tasselled Cap transformation, can be computed from a linear combination of visible and NIR channels. This transformation is sensor-specific and was originally defined for Landsat MSS (Kauth and Thomas, 1976). A range of similar transformations now exist for other sensors (Crist and Cicone 1984; Cicone and Metzler 1984; Roberts *et al.*, 2018) and can be used to analyse specific feature classes in imagery from multiple sensors. Jackson (1983) provides a detailed description for calculating the coefficients for '*n*-space' indices (that is, linear combinations of *n* spectral channels in *n*-dimensional space) using the Gram-Schmidt process with a minimum of data points.

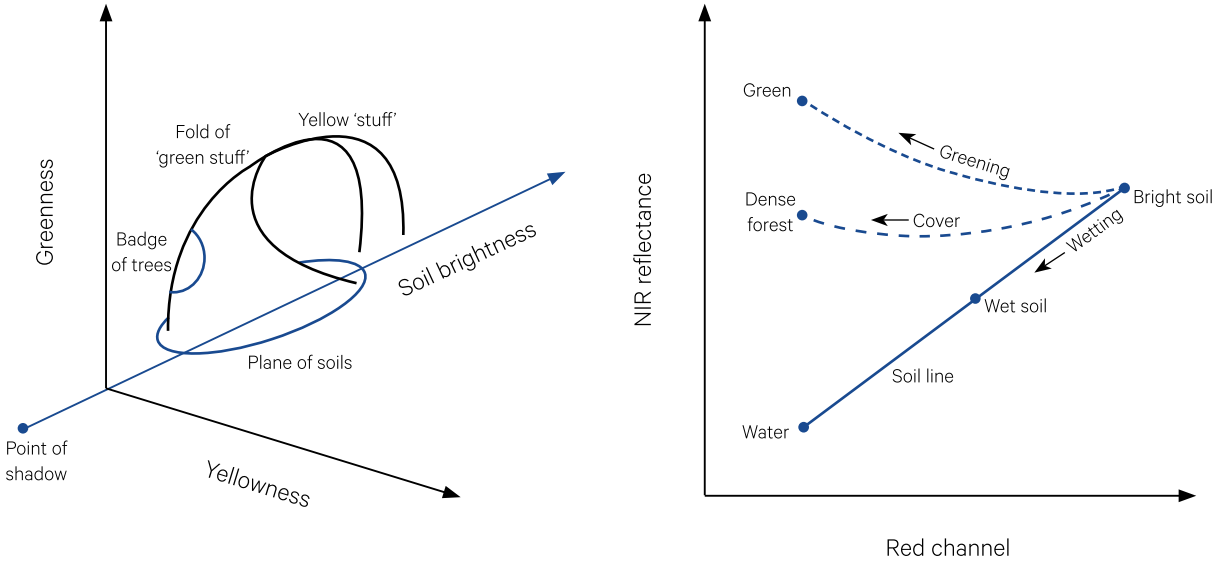
For mixed agricultural scenes, the Kauth-Thomas transformation describes the essential three dimensions of the Landsat MSS data space as forming the shape of a tasselled cap (see Figure 11.6a). This cap 'sits' on the principal data axis, which has been described as the 'plane of soils'. The initial change in crop reflectance during the crop growing cycle is described as moving down the soil line due to shadowing (though the extent of shadowing varies with the orientation of crop rows). Once maximum soil shadowing has been achieved, the trajectory of reflectance moves away from the soil line. This trajectory curves up to the region of 'green stuff' as the crop canopy becomes denser, then folds over to converge on the region of 'yellow stuff' as the vegetation hays off. The shadowed nature of tree canopies renders them a special position in this model, which is fancifully described as the 'badge' on the 'Tasselled Cap'. The point of shadow represents the minimum reflectance level and can be considered as the equivalent of dark values in image ratioing (see Section 10.2.1). Figure 11.6b illustrates the way changes in vegetation greenness affect its reflectance of red and NIR wavelengths. This dynamic crossplot represents one plane of the Tasselled Cap space.

When a wall is slowly covered over by earth,
the materials it's made from decay and become part of the soils around and above it,
sometimes causing vegetation above and next to the wall to grow faster or slower.
Satellite imagery helps archaeologists to pick up these subtle changes.
(Sarah Parcak)

Figure 11.6 The 'Tasselled Cap' feature in Landsat MSS image data

a. For vegetated landscapes, the major dimensions in Landsat MSS imagery can be described in terms of Brightness, Greenness and Yellowness. The Tasselled Cap feature graphically maps different stages of vegetation growth relative to these dimensions.

b. One plane through the Tasselled Cap shows the effect of changes in vegetation 'greenness' on red and NIR reflectance.



Source: Harrison and Jupp (1990) Figure 92

Using broad land cover categories such as vegetation and soils, a generalised affine transformation was defined to rotate Landsat MSS data channels into dimensions referred to as Brightness, Greenness, Yellowness and Non Such (Kauth and Thomas, 1976). The plane of Brightness versus Greenness contains at least 95% of the total variation for agricultural imagery, with Brightness defining the direction of soil reflectance variation and the perpendicular direction, Greenness, being associated with the reflectance characteristics of green vegetation (Crist and Kauth, 1986). The matrix defined for this rotation is shown in Figure 11.7.

Figure 11.7 Kauth Thomas greenness transformation matrix

$$\begin{bmatrix} \text{Brightness} \\ \text{Greenness} \\ \text{Yellowness} \\ \text{Non such} \end{bmatrix} = \begin{bmatrix} 0.433 & 0.632 & 0.586 & 0.264 \\ -0.290 & -0.562 & 0.600 & 0.491 \\ -0.829 & 0.522 & -0.039 & 0.194 \\ 0.223 & 0.012 & -0.543 & 0.810 \end{bmatrix} \times \begin{bmatrix} \text{MSS4} \\ \text{MSS5} \\ \text{MSS6} \\ \text{MSS7} \end{bmatrix} + \begin{bmatrix} 32 \\ 32 \\ 32 \\ 32 \end{bmatrix}$$

Here we see Brightness as a weighted average of the four original MSS channels while Greenness is essentially a weighted difference of NIR and red, which is typical for vegetation indices. The Greenness feature can also be normalised by the 'Green Brightness', that is:

$$\begin{aligned}\text{Unnormalised Greenness} = &-0.29 \times \text{MSS4} \\ &-0.562 \times \text{MSS5} \\ &+0.60 \times \text{MSS6} \\ &+0.491 \times \text{MSS7}\end{aligned}$$

and

$$\text{Normalised Greenness} = \frac{\text{Unnormalised Greenness}}{\text{Green Brightness}}$$

where

$$\begin{aligned}\text{Green Brightness} = &0.29 \times \text{MSS4} \\ &+0.562 \times \text{MSS5} \\ &+0.60 \times \text{MSS6} \\ &+0.491 \times \text{MSS7}\end{aligned}$$

This normalisation partially compensates for variations in solar elevation, so allows more valid comparisons of greenness changes in multi-temporal images.

A similar transformation was defined for six channels of Landsat TM data, which produced three significant features called Brightness, Greenness and Wetness (Crist and Cicone, 1984):

$$\begin{aligned}\text{Brightness} = &0.33183 \times \text{TM1} + 0.33121 \times \text{TM2} \\ &+0.55177 \times \text{TM3} + 0.42514 \times \text{TM4} \\ &+0.48087 \times \text{TM5} + 0.25252 \times \text{TM7} \\ \text{Greenness} = &-0.24717 \times \text{TM1} - 0.16263 \times \text{TM2} \\ &-0.40639 \times \text{TM3} + 0.85468 \times \text{TM4} \\ &+0.05493 \times \text{TM5} - 0.11749 \times \text{TM7} \\ \text{Wetness} = &0.13929 \times \text{TM1} + 0.22490 \times \text{TM2} \\ &+0.40359 \times \text{TM3} + 0.25178 \times \text{TM4} \\ &-0.70133 \times \text{TM5} - 0.45732 \times \text{TM7}\end{aligned}$$

Equivalent coefficients have been derived for Landsat-8 OLI imagery, calibrated to surface reflectance (Roberts *et al.*, 2018):

$$\begin{aligned}\text{Brightness} = &0.3567 \times \text{blue} + 0.3567 \times \text{green} \\ &+0.3567 \times \text{red} + 0.5350 \times \text{NIR} \\ &+0.5350 \times \text{SWIR1} + 0.2140 \times \text{SWIR2} \\ \text{Greenness} = &0.1987 \times \text{blue} - 0.2826 \times \text{green} \\ &-0.2724 \times \text{red} + 0.5357 \times \text{NIR} \\ &+0.2388 \times \text{SWIR1} - 0.6800 \times \text{SWIR2} \\ \text{Wetness} = &0.5702 \times \text{blue} + 0.1584 \times \text{green} \\ &+0.2627 \times \text{red} - 0.3959 \times \text{NIR} \\ &-0.0045 \times \text{SWIR1} - 0.6511 \times \text{SWIR2}\end{aligned}$$

Variations in these (land cover related) features are represented by changes in two or more channels in the original imagery. As with PCA, rotation of the original data space to axes that are aligned with the major data variation in land cover features allows them to be more simply analysed and related to ground-based measurements.

Dark values can generally be specified for each channel to represent the level of atmospheric or sensor noise. In most image processing systems, these values can be defined by interactive training on shadowed or deep water areas, or defaulted to image minimum values (see Section 10.2.1).

This transformation has been used in various agricultural models (see Volume 3A). As a single channel it both allows simpler relationships to be developed between image radiance and ground-based measurements, and permits vegetation categories to be analysed easily using interactive image painting (see Volume 2A—Section 9.2.1).

11.2 Vegetation Water Content

EO data can be used to quantify the water content of vegetation by relying on the liquid water absorption features in optical wavelengths (see Volume 3A for details). Both NIR and SWIR reflectance data are needed to determine water content since:

- NIR indicates the internal leaf structure and dry matter content (see Figure 11.1); and
- SWIR indicates the vegetation water content, internal structure and dry matter content (see Figure 11.8).

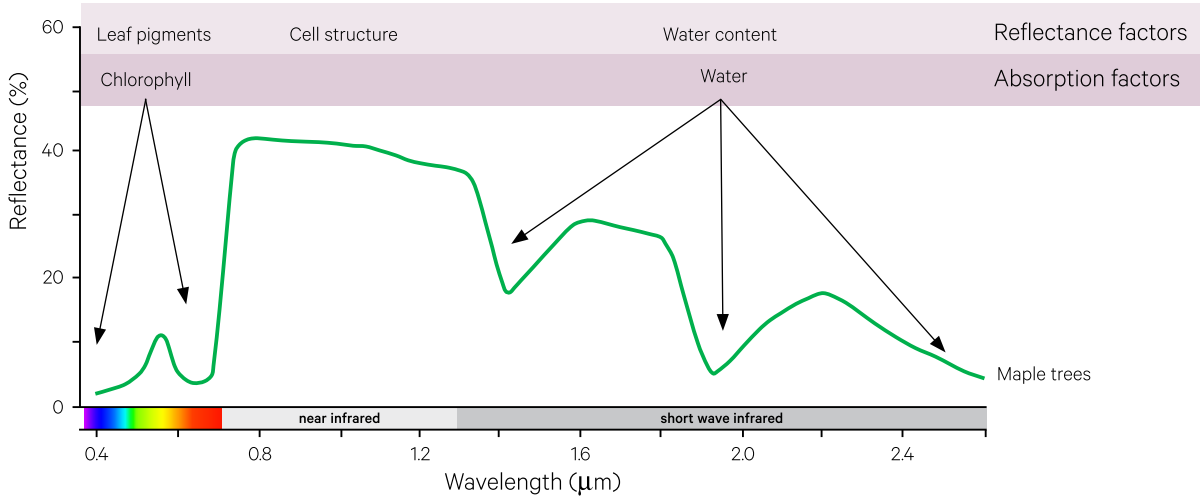
Some of the ratios and normalised ratios computed from these wavelengths are summarised in Table 11.2. These water content indices have shown good correlation with empirical measures of foliar water content, such as Equivalent Water Thickness (EWT) and have been used to infer fuel moisture for fire risk prediction (see Volume 3A).

Table 11.2 Vegetation water content indices

Name	Formulation	Reference
Water Index (WI)	$WI = \left(\frac{\rho_{900}}{\rho_{970}} \right)$	Peñuelas *et al.* (1997b)
Normalised Difference Water Index (NDWI)	$NDWI = \left(\frac{\rho_{860} - \rho_{1240}}{\rho_{860} + \rho_{1240}} \right)$	Gao (1996)
Simple Ratio Water Index (SRWI)	$SRWI = \left(\frac{\rho_{860}}{\rho_{1240}} \right)$	Zarco-Tejada *et al.* (2003)
Normalised Difference Infrared Index (NDII)	$NDII = \left(\frac{\rho_{819} - \rho_{1600}}{\rho_{819} + \rho_{1600}} \right)$	Hardinsky *et al.* (1983)
Moisture Stress Index (MSI)	$MSI = \left(\frac{\rho_{1600}}{\rho_{819}} \right)$	Hunt and Rock (1989)
Datt Water Index (DWI)	$DWI = \left(\frac{\rho_{895} - \rho_{675}}{\rho_{895} + \rho_{675}} \right)$	Datt (1999)

Figure 11.8 Typical green vegetation spectrum

Reflectance properties of green vegetation primarily derive from leaf pigments in visible wavelengths, cell structure in NIR wavelengths, and water content in SWIR wavelengths. Characteristic absorption features occur in visible wavelengths (due to chlorophyll) and SWIR wavelengths (due to water content).



Adapted from: Harrison and Jupp (1989) Figure 30

11.3 Further Information

Jensen (2016) Section 8

Index Database (a database for remote sensing indices): *https://www.indexdatabase.de*

11.4 References

Bannari, A., Morin, D., Bonn, F. and Huete, A. R.(1995). A review of vegetation indices. *Remote Sensing Reviews*, 13, 1, 95–120. doi: *10.1080/02757259509532298*

Burgan, R., Klaver, R., and Klaver, J. (1998). Fuel models and fire potential from satellite and surface observations. *International Journal of Wildland Fire*, 8, 159–170.

Carlson, T., and Ripley, D. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, 62, 241–252.

Chuvieco, E.,I. Aguado, D. Cocero, D. and, and Rian, O. (2003). Design of an empirical index to estimate fuel moisture content from NOAA-AVHRR images in forest fire danger studies. *International Journal of Remote Sensing*, 24, 1621–1637.

Cicone, B.C., and Metzler, M.D. (1984). Comparison of Landsat MSS, Nimbus–7, CZCS, and NOAA–7 AVHRR features for land-use analysis. *Remote Sensing of Environment*, 14, 257–265.

Crippen, R.E. (1987). The regression intersection method of adjusting image data for band ratioing. *International Journal of Remote Sensing,* 8, 137–155.

Crippen, R.E. (1988). The dangers of underestimating the importance of data adjustments in band ratioing. *International Journal of Remote Sensing*, 9, 767–776.

Crippen, R.E., Blom, R.G., and Heyada, J.R. (1988). Directed band ratioing for the retention of perceptually-independent topographic expression in chromaticity-enhanced imagery. *International Journal of Remote Sensing,* 9, 749–765.

Crist, E.P., and Cicone, R.C. (1984). Application of the Tasselled Cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, 50, 343–352.

Crist, E.P., and Kauth, R.J. (1986). The Tasselled Cap de-mystified. *Photogrammetric Engineering and Remote Sensing*, 52, 81–86.

Datt, B. (1999) Remote Sensing of Water Content in *Eucalyptus* Leaves. *Australian Journal of Bot*any, 47, 909–923.

Deering, D.W., Rouse, J.W., Haas, R.H., and Schell, J.A. (1975). Measuring Forage Production of Grazing Units from Landsat MSS Data. *10th International Symposium on Remote Sensing of Environment*, 2, 1169–1178.

Eidenshink J.C., Burgan, R.E., and Haas, R.H. (1990). Monitoring fire fuels condition by using time series composites of Advanced Very High Resolution Radiometer (AVHRR) data. *Proceedings of Resource Technology 90*, Washington, DC, ASPRS, 68–82.

Gao, B.C. (1996). NDWI—a normalised difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266.

Gao, X., Huete, A. R., Ni, W., and Miura, T. (2000). Optical–biophysical relationships of vegetation spectra without background contamination, *Remote Sensing of Environment*, 74, 609–620.

Gitelson, A.A., Kaufman, Y.J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment,* 80(1), 76–87.

Glenn, E.P., Huete, A.R., Nagler, P.L., and Nelson, S.G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8, 2136–2160.

Hardisky, M. A., Klemas, V., and Smart, R. M. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of *Spartina alterniflora* canopies. *Photogrammetric Engineering and Remote Sensing*, 49, 77–83.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.

Huete, A, Jackson, R., and Post, D. (1985). Spectral response of a plant canopy with different soil backgrounds. *Remote Sensing of Environment*, 17, 37–53.

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309.

Huete, A., Justice, C. and van Leeuwen, W. (1999). *MODIS vegetation index (MOD 13) algorithm theoretical basis document (ATBD)* Version 3. 0. EOS Project Office, NASA Goddard Space Flight Center, Greenbelt, MD.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213.

Huete, A., Didan, R., van Leeuwen, W., Miura, T., and Glenn, E. (2011). MODIS Vegetation Indices. Ch 26 in *Land Remote Sensing and Global Environmental Change, Remote Sensing and Digital Image Processing* Volume 11, Springer Science+Business Media, LLC. ISBN 978-1-4419-6748-0.

Hunt, E.R., Jr, and Rock, B.N. (1989). Detection of changes in leaf water content using near- and middle-infrared reflectances. *Remote Sensing of Environment*, 30, 43–54.

Jackson, R.D. (1983). Spectral Indices in n-Space. *Remote Sensing of Environment*, 13, 409–421.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Jiang, Z., Huete, A., Chen, J., Chen, Y., Li, J., Yan, G., and Zhang, X. (2006). Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment*, 101, 366–378.

Jiang, Z., Huete, A.R., Didan, K., and Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112, 3833–3845.

Kaufman, Y.J., and Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS- MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2), 261–270.

Kauth, R.J., and Thomas, G.S. (1976). The Tasselled Cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proc. Symposium on Machine Processing of Remotely Sensed Data*. Purdue University, West Lafayette, Indiana, 4B41–4B51.

Lasaponara, R. (2005). Inter-comparison of AHVRR-based fire susceptibility indicators for the Mediterranean ecosystems of Southern Italy. *International Journal of Remote Sensing*, 26, 853–870.

Liu, H.Q., and Huete, A.R. (1995). A feedback based modification of the NDV I to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 457–465.

Matsushita, B., Yang, W., Chen, J., Onda, Y., and Qiu, G. (2007). Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest. *Sensors*, 7, 2636–2651.

Peñuelas, J., Llusia, J., Piñol, J., and Filella, I. (1997a) Photochemical reflectance index and leaf photosynthetic radiation-use efficiency assessment in Mediterranean trees. *International Journal of Remote Sensing*, 18, 2863–2868.

Peñuelas, J. Pinol, J., Ogaya, R., and Filella, I. (1997b) Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). *International Journal of Remote Sensing,* 18(13), 2869–2875. doi: *10.1080/014311697217396*

Perry, C.R.Jr., and Lautenshlager, L.F. (1984). Functional equivalent of spectral vegetation indices. *Remote Sensing of Environment*, 14, 169–182.

Roberts, D., Dunn, B., and Mueller, N. (2018). Open Data Cube Products Using High-Dimensional Statistics of Time Series. *Proceedings 2018 IGARSS*, 22–27 July, Valencia Spain.

Rondeaux, G, Steven, M. and Baret, F. (1996). Optimisation of Soil-Adjusted Vegetation Indices. *Remote Sensing of Environment*, 55, 95–107.

Sims, DA., and Gamon, J.A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81, 337–354.

Singh, S.M., and Saull, R.J. (1988). The effect of atmospheric correction on the interpretation of multitemporal AVHRR-derived vegetation index dynamics. *Remote Sensing of Environment*, 25, 37–51.

Thompson, D.R., and Wehman, O.A. (1979). Using Landsat digital data to detect moisture stress. *Photogrammetric Engineering and Remote Sensing*, 45, 201–207.

Tucker, C.J., Elgin, J.H.Jr., and McMurtney, J.E.III. (1980a). Relationship of crop radiance to alfalfa agronomic values. *International Journal of Remote Sensing*, 1, 69–75.

Tucker, C.J., Holben, B.N., Elgin, J.H.Jr., and McMurtney, J.E.III. (1980b). Relationship of spectral data to grain yield variation. *Photogrammetric Engineering and Remote Sensing*, 46, 657–666.

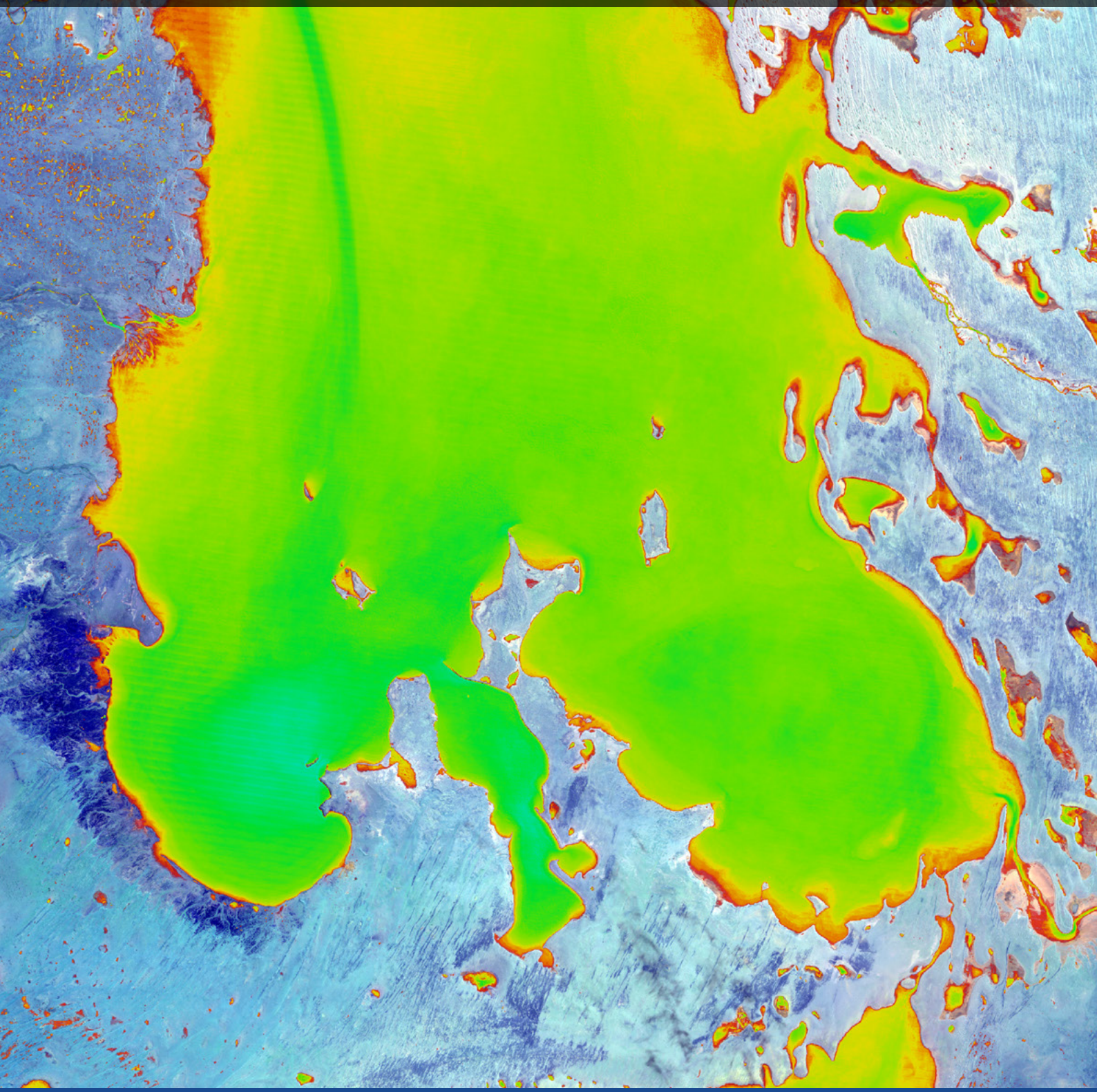
Wardlow, B.D., Egbert, S.L., and Kastens, J.H. (2007) Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*, 108, 290–310.

Xue, J., and Su, B. (2017). Review Article: Significant remote sensing vegetation indices: a review of developments and applications. *Journal of Sensors*, ID 1353691. *https://doi.org/10.1155/2017/1353691*

Zarco-Tejada, P.J., Rueda, C.A., and Ustin, S.L. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sensing of Environment*, 85, 109–124.

Earth Observation:
Data, Processing and Applications

Volume 2D: Processing—Image Integration



This publication is available in PDF format at *eopublications.au*.
We welcome your comments regarding the readability and usefulness of this report. To provide feedback, please email *feedback@eopublications.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
978-0-6482278-2-3

Copyright:
All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, save for content supplied by third parties, and logos. Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. The full licence terms are available from *https://creativecommons.org/licenses/by/4.0/legalcode*. A summary of the licence terms is available from *https://creativecommons.org/licenses/by/4.0/*.

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Recommended Citation for Volume 2D:
CRCSI (2019) *Earth Observation: Data, Processing and Applications. Volume 2D: Processing—Image Integration*. (Eds. Harrison, B.A., Jupp, D.L.B., Lewis, M.M., Mueller, N., Byrne, G., Hudson, D.A., Sparks, T.,) CRCSI, Melbourne.

Latest Revision:
October 2025

Background image on previous page: The Water Observations from Space (WOfS; see Volume 1A—Excursus 5.1) for Lake Eyre in South Australia, overlaid on a false colour Landsat-5 image acquired on 9 May 2009. WOfS shows the percentage of time that surface water was present in a location since 1987, with red areas having very low occurrence transitioning through yellow and green to blue which has the highest occurrence.
Source: Norman Mueller, Geoscience Australia

Acknowledgements

Production of this series of texts would not have been possible without the financial support of CSIRO, CRCSI, GA and BNHCRC, input from members of the editorial panels and direction from members of the various advisory panels. Administrative support from FrontierSI is also gratefully acknowledged.

Volumes 1 and 2 of this series are based on text originally published in Harrison and Jupp (1989, 1990, 1992 and 1993)[1]. Many illustrations and some text from these publications have been reproduced with permission from CSIRO.

Other contributors are gratefully acknowledged:

- reviewers: Alfredo Huete and Laurie Chisholm for valuable feedback on Volumes 1 and 2; Irina Emelyanova (Excursus 6.1); Mark Grant, Matthew Stenson and Matthew Paget (Excursus 12.2 and Excursus 14.1); Matthew Brooks (Excursus 13.2);
- illustrations: Norman Mueller kindly supplied most of the lovely images; other contributors of graphical material include: Tony Sparks, Megan Lewis, David Jupp, Craig Shephard, Michael Caccetta, Jan Verbesselt, Glenn Newnham, Tim McVicar, Randall Donohue, Steven Sagar, Fuqin Li, Nearmap, and EagleView;
- excursus: David Hudson, Megan Lewis, Randall Donohue, Richard Thackway, Marian Wiltshire, Janet Anstee, Albert van Dijk, Tim Danaher, Dan Tindall, Mark Baird; and
- Carl Davies (CMDphotographics) for selected graphical illustrations.

We thank those owners of copyrighted illustrative material for permission to reproduce their work. Credits for individual illustrations are provided below the relevant graphic.

All volumes in this series are covered by the copyright provisions of CC BY 4.0 AU.

1 Harrison, B.A., and Jupp, D.L.B. (1989) Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual. CSIRO, Melbourne. 156pp.
Harrison, B.A., and Jupp, D.L.B. (1990) Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual CSIRO, Melbourne. 256pp.
Harrison, B.A., and Jupp, D.L.B. (1992) Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual. MPA, Melbourne.
Harrison, B.A., and Jupp, D.L.B. (1993) Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual. MPA, Melbourne.

Table of Contents

Volume 2D: Processing—Image Integration a

Acknowledgements i

Getting Started 1

1 EO Analysis Framework **3**

1.1 Indirect Observations 3

1.2 Implicit Assumptions 4

1.3 Layers of Complexity 5

1.4 The Surface Mosaic 7

1.5 The Bigger Picture 10

1.6 Further Information 11

1.7 References 11

2 Multi-temporal EO Datasets **13**

2.1 Bi-temporal and Multi-temporal Imagery 14

2.1.1 Conceptual structure 14

2.1.2 Satellite-acquired datasets 15

2.2 Appropriate Data 17

2.2.1 Timing 17

2.2.2 Scaling 18

2.2.3 Spectral and radiometric differentiation 18

2.3 Appropriate Processes 18

2.4 Processing Methodologies 19

2.4.1 Processing order 19

2.4.2 Processing workflows 19

2.5 Accuracy, Stability, Validation and Trust 21

2.6 Further Information 25

2.7 References 25

3 Data Standardisation **27**

3.1 Image Calibration 27

3.1.1 Systematic effects 27

3.1.2 Local effects or 'noise' 30

3.2 Analysis Ready Data 30

3.3 Ancillary data 32

3.3.1 Provenance and caveats 32

3.3.2 Data formats 33

3.4 Further Information 34

3.5 References 34

4 Pre-processing **37**
4.1 Stratifying images 37
4.1.1 Spectral 37
4.1.2 Spatial 38
4.2 Reducing data volume 38
4.2.1 Spectral indices 38
4.2.2 Linear Transformations 39
4.3 Upscaling and Downscaling 40
4.3.1 Upscaling to a larger pixel size 40
4.3.2 Downscaling to a smaller pixel size 43
4.3.3 Resampling to a common pixel size 43
4.4 Further Information 44
4.5 References 44

Processing Image Pairs 45

5 Basic Comparisons **47**
5.1 Multi-date Colour Composites 47
5.2 Differencing and Addition 49
5.3 Manipulating Segmentation Masks 52
5.4 Transferring Category Statistics 54
5.5 Colour Balancing for Image Mosaicking 54
5.6 Further Information 56
5.7 References 56

6 Fusion Methods **57**
6.1 Spatial and Spectral Blending 58
6.2 Spatial and Temporal Blending 60
6.3 Further Information 66
6.4 References 66

7 Change Detection Methods **69**
7.1 Algebraic Operations 71
7.1.1 Channel subtraction 71
7.1.2 Channel ratioing 72
7.1.3 Channel regression 73
7.1.4 Change Vector Analysis 74
7.2 Linear Transformations 76
7.3 Classification Approaches 78
7.3.1 Comparing classifications 78
7.3.2 Classifying merged imagery 79
7.4 Post-processing 80
7.4.1 Thresholding and labelling 80
7.4.2 Filtering 80
7.4.3 Autocorrelation analysis 80
7.5 Accuracy Assessment 80
7.6 Further Information 81
7.7 References 81

Processing Time Series Datasets 85

8 Time Series Analysis **87**
- **8.1** Time Series Components and Processing 88
- **8.2** Understanding the Data 91
- **8.3** Creating a Data Cube 92
- **8.4** Detecting Anomalies 97
- **8.5** Further Information 98
- **8.6** References 98

9 Characterising Temporal Trends **101**
- **9.1** The Time Series Trajectory 101
- **9.2** Analysing Trends 104
- **9.3** Comparing Trajectories 105
- **9.4** Analytical Methods 106
- **9.5** Further Information 107
- **9.6** References 107

10 Pixel-based Composites **111**
- **10.1** Single Sensor Datasets 112
  - **10.1.1** Time period 112
  - **10.1.2** Composite values 114
- **10.2** Multiple Sensor Datasets 117
- **10.3** Further Information 119
- **10.4** References 119

11 Digital Earth **123**
- **11.1** The Concept 123
- **11.2** Digital Earth Australia 125
- **11.3** Further Information 129
- **11.4** References 129

Merging EO and Other Datasets 133

12 Sample-based Datasets **135**
- **12.1** Manually-Collected Field Site Data 135
- **12.2** Automatic Sensors 137
- **12.3** Spectroscopy datasets 141
- **12.4** Volunteered data 142
- **12.5** Further Information 144
- **12.6** References 144

13 Gridded Datasets **149**
- **13.1** Geographic Information Systems 149
- **13.2** Digital Elevation Models 150
- **13.3** Meteorological Records 154
- **13.4** Further Information 156
- **13.5** References 157

Integrated Solutions 159

14 Mapping, Monitoring and Modelling 161
14.1 Mapping 161
14.2 Monitoring 163
14.2.1 Environmental Accounting 163
14.2.2 Compliance Monitoring 169
14.3 Modelling 175
14.4 Further Information 179
14.5 References 180

List of Figures

Figure 1.1 Measurement and analysis of image features 4

Figure 1.2 Complexity levels in EO imagery 6

Figure 1.3 Upscaling and downscaling of soil data 7

Figure 1.4 Scales of EO 8

Figure 1.5 Mapping, monitoring and modelling for natural resource management 10

Figure 2.1 Structure of an EO image 14

Figure 2.2 Bi-temporal and multi-temporal image sets 15

Figure 2.3 Workflow for fractional cover product 20

Figure 2.4 Process flow chart example 20

Figure 2.5 Lifecycle model for a sensor data product 22

Figure 3.1 Five basic types of patches 33

Figure 4.1 Principal Component Analysis 39

Figure 4.2 Upscaling operations 40

Figure 4.3 Upscaling by picking and binning approach 41

Figure 4.4 Vector to raster conversion 41

Figure 4.5 Bilinear interpolation 42

Figure 4.6 Cubic convolution interpolation 42

Figure 4.7 Downscaling by duplicating pixels 43

Figure 5.1 Multi-date NDVI composite 48

Figure 5.2 Image differencing 49

Figure 5.3 Colour composite difference image 50

Figure 5.4 Residual image 51

Figure 5.5 Examples of Euclidean distance images 52

Figure 5.6 Recombining images 53

Figure 5.7 Colour balancing 55

Figure 6.1 STARFM approach 61

Figure 6.2 Temporal profiles of blending algorithms 62

Figure 6.3 Image blending algorithms 62

Figure 6.4 Generic overview of Landsat-MODIS blending. 65

Figure 7.1 Difference channel histogram 71

Figure 7.2 Change detection based on differencing 71

Figure 7.3 Ratio channel histogram 72

Figure 7.4 Change detection based on ratioing 73
Figure 7.5 Automatic Scattergram-Controlled Regression 73
Figure 7.6 Change Vector Analysis (CVA) 74
Figure 7.7 Direction of change 75
Figure 7.8 Magnitude of change 75
Figure 7.9 Original images 77
Figure 7.10 Principal components 78
Figure 8.1 Time series imagery of Australia 88
Figure 8.2 Components of time series variations 89
Figure 8.3 fPAR time series examples 90
Figure 8.4 Vegetation cover triangle 93
Figure 8.5 Land cover reflectances 94
Figure 8.6 Seasonal trends 1981–2006 95
Figure 8.7 fPAR decomposition 96
Figure 8.8 Dynamic Time Warping (DTW) 96
Figure 9.1 Temporal trajectory attributes 102
Figure 9.2 Time series distance measures 105
Figure 10.1 Brightness Temperature Composites from SSM/I 115
Figure 10.2 Workflow for Australian Landsat composite images 116
Figure 10.3 Flow chart for multisensor composites 118
Figure 11.1 Technological and cultural advances enabling Digital Earth 124
Figure 11.2 Global network of data cubes 124
Figure 11.3 DEA concept 125
Figure 11.4 National Intertidal Digital Elevation Model 126
Figure 11.5 Radiometric correction of LSR25 128
Figure 12.1 TERN sites 138
Figure 12.2 IMOS National mooring network 140
Figure 12.3 Inadvertent map of distance to earthquake epicentre 142
Figure 12.4 Using the EyeonWater Australia app 143
Figure 12.5 EyeonWater Australian observations 2017–2018 143
Figure 13.1 DTM versus DSM 150
Figure 13.2 Elevation data formats 151
Figure 13.3 Geofabric datasets 153
Figure 13.4 GPCC average precipitation 1982–2010 154
Figure 13.5 GPCC buffer mask 155

Figure 14.1 Australian environmental-economic account for agriculture 164

Figure 14.2 Australia's Environment Explorer 167

Figure 14.3 2018 OzWALD summary for ACT 168

Figure 14.4 Schematic representation of SLATS processing steps 171

Figure 14.5 Example of SLATS change index and likelihood images 172

Figure 14.6 Example of SLATS reporting outcome 173

Figure 14.7 eReefs modelling framework 177

Figure 14.8 Temporal and spatial scales of GBR processes and observations 177

Figure 14.9 eReefs model interactions 178

List of Tables

Table 1.1 Logical stages in EO data use 3

Table 1.2 Assumptions implicit in the analysis of EO data 5

Table 1.3 Scale hierarchy for soil and landscape data 8

Table 1.4 Landform models relating to mapping scales 9

Table 1.5 Complementary benefits of mapping, monitoring and modelling 11

Table 2.1 Sampling dimensions in EO imagery 14

Table 2.2 Global EO time series datasets 16

Table 2.3 CEOS WGCV validation hierarchy for moderate resolution global products 21

Table 2.4 Factors impacting the uptake of EO products 23

Table 2.5 EOPU framework for emergency management 24

Table 3.1 HLS processing overview 32

Table 5.1 Example histogram percentage points 55

Table 6.1 Terminology associated with data fusion 58

Table 6.2 Pan-sharpening methods 59

Table 6.3 Landsat and Terra/MODIS characteristics 60

Table 6.4 Notation for blending framework 63

Table 7.1 Bi-temporal change detection stages 70

Table 7.2 PC transformation matrix 77

Table 8.1 Time series processing sequence 89

Table 9.1 Time series sub-sequence differences 103

Table 9.2 Time series analysis methods 104

Table 9.3 Time series similarity measures 106

Table 9.4 Landsat-based change detection approaches 106

Table 10.1 MODIS land products 113

Table 10.2 Popular criteria for selecting 'best' pixel 114

Table 11.1 Foundations of DEA 126

Table 12.1 Information Hierarchy 136

Table 12.2 Spectral database processing levels 141

Table 13.1 Terrain variables 151

Table 13.2 Australian DEM 152

Table 14.1 TERN EO-based continental products 162

Table 14.2 OzWALD model components 166

List of Excurses

Excursus 1.1—Landform models 9

Excursus 2.1—Why is an EO Product Trusted? 22

Excursus 6.1—Theoretical Framework for Image Blending 63

Excursus 7.1—Change Detection using PCA 77

Excursus 8.1—Relative Calibration of Time Series Datasets 93

Excursus 10.1—MODIS Composites 113

Excursus 11.1—GA Analysis Ready Data 127

Excursus 12.1—National Vegetation Information System 136

Excursus 12.2—Terrestrial Ecosystem Research Network (TERN) 137

Excursus 12.3—Integrated Marine Observing System (IMOS) 139

Excursus 12.4—EyeonWater Australia 143

Excursus 13.1—NationalMap 150

Excursus 13.2—Australian Hydrological Geospatial Fabric 153

Excursus 14.1—TERN Continental Products 162

Excursus 14.2—OzWALD 165

Excursus 14.3—SLATS: Monitoring Woody Vegetation Clearing 170

Excursus 14.4—eReefs 176

Getting Started

In the context of Earth Observation (EO), data integration encompasses a range of image processing techniques that enable spatial datasets from different sources to be compared and combined. This sub-volume describes processing operations that are relevant to either:

- a pair of bi-temporal EO images—two images of the same ground area taken at different times (see Sections 5 to 7); or
- a multi-temporal EO image dataset—more than two images of the same ground area, each of which is acquired at different times (see Section 8 to 11).

Various transformations for merging or analysing data from multiple image channels are introduced in Volume 2C. While an awareness of the underlying EO framework is necessary for all image processing operations, as described in Section 1 it has particular significance when integrating EO image datasets acquired at different times. Similarly the constraints and complexities involved with analysing any EO image are multiplied when processing bi-temporal or multi-temporal EO datasets. Some of the factors that are relevant to selecting the most appropriate data and processes for multi-temporal analyses are discussed in Section 2.

One of the most valuable opportunities afforded by EO data is the regular monitoring of changes in features on the Earth's surface (see Volumes 1 and 3). When bi-temporal or multi-temporal datasets are involved in such analyses, appropriate radiometric and geometric calibration of the EO data is critical. Essential pre-processing operations for bi-temporal and multi-temporal EO datasets are reviewed in Section 3, and optional pre-processing operations are discussed in Section 4.

EO datasets can also be integrated with non-EO data sources. Some of the options and challenges involved with such integration are considered in Sections 12 and 13. Examples of some mapping, monitoring and modelling results based on EO analyses are given in Section 14 and further expanded in Volume 3.

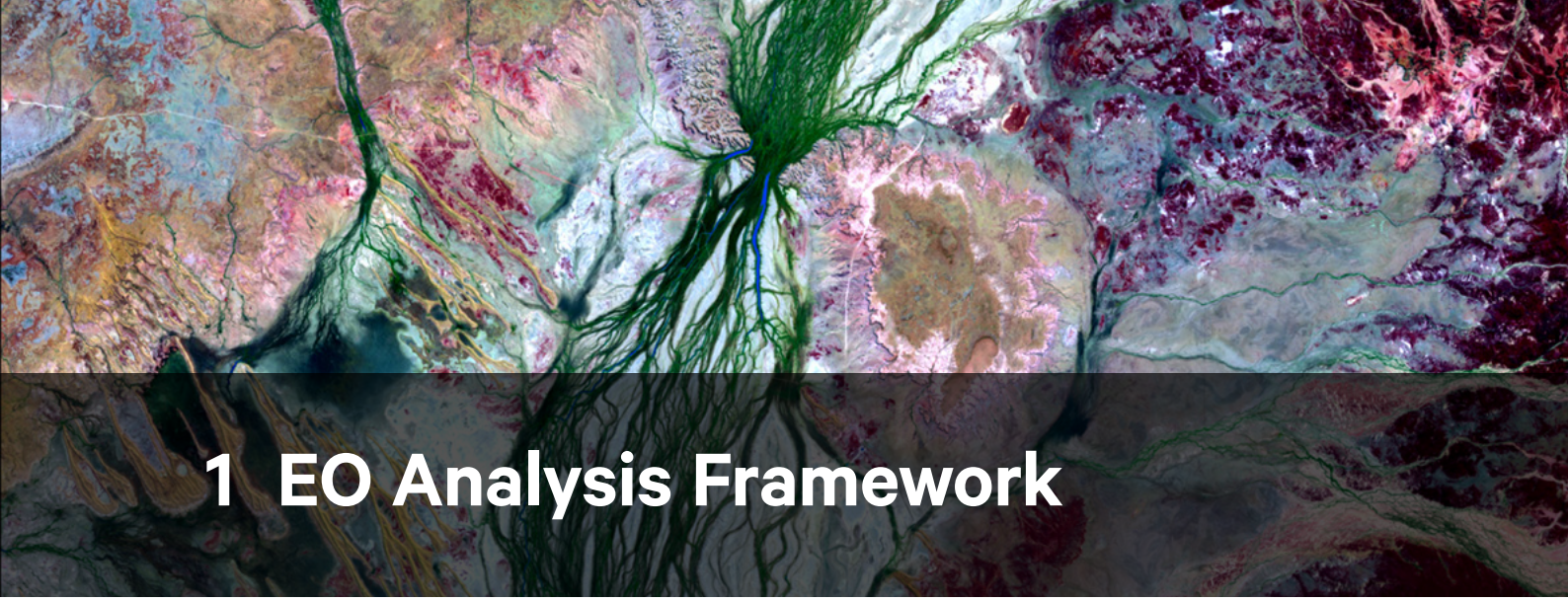
Contents

1 EO Analysis Framework 3

2 Multi-temporal EO Datasets 13

3 Data Standardisation 27

Background image on previous page: Sentinel-2a image over the Kimberly region, northwest WA, acquired on 10 September 2018. Bands 4, 3 and 2 are shown as RGB. **Source:** Norman Mueller, Geoscience Australia



1 EO Analysis Framework

Interpretation of any Earth Observation (EO) image requires consideration of the way 'real world' objects generate, or contribute to, its colours, patterns, and textures. While the relationships between the image and the imaged target are important in the analysis of individual EO images, they carry even greater importance when two or more images are being processed and/or when other data sources are being integrated with EO imagery. In this section we will review some basic principles that underlie interpretation and analysis of all EO datasets, namely the:

- indirect nature of EO observations (see Section 1.1);
- implicit assumptions inherent to EO (see Section 1.2);
- layers of complexity in every EO image (see Section 1.3); and the
- 'mosaic' formed by features on the surface being imaged (see Section 1.4).

Finally we consider the bigger picture of data integration in Section 1.5.

1.1 Indirect Observations

Volume 1A—Section 1 introduced the concept that EO observations are implicitly indirect and that their interpretation necessarily involves at least one level of inference, or indirection. In Table 1.1, the generalised process of using EO data is summarised in terms of three logical stages: acquisition, processing and interpretation.

The indirect relationship between features in an EO image and quantifiable attributes of the Earth's surface is summarised in Figure 1.1 (see Volume 1A—Section 1 for details). In this diagram, image processing operates on the 'data space', or observation domain, and implicitly develops or utilises various assumptions which relate the image data back to selected parameters of the imaged ground scene (see Section 1.2).

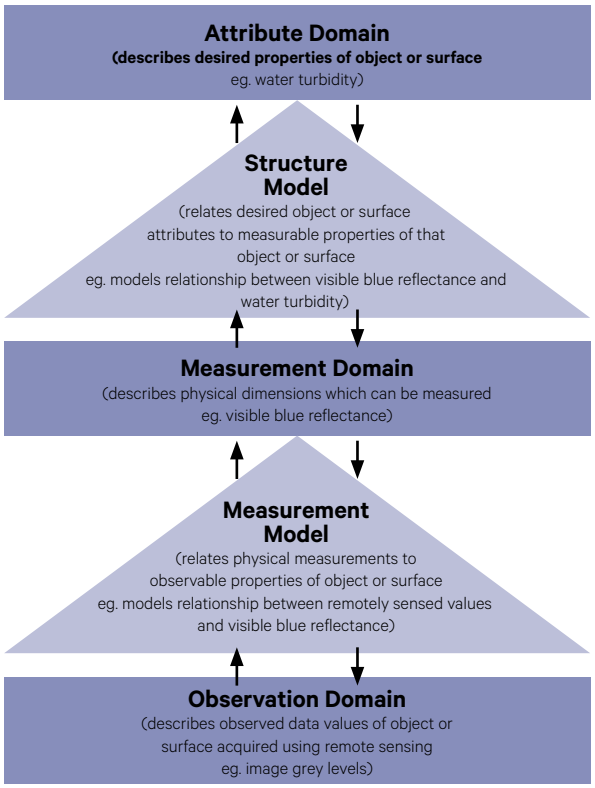
Table 1.1 Logical stages in EO data use

Stage	Description	Domain
Acquisition	Observe object or surface using remote sensing	Observation
Processing	Convert remote sensing observation to measurable property of object or surface using pre-defined measurement model	Measurement
Interpretation	Interpret measured properties of object or surface (that is, the converted remote sensing observation) in terms of a specific application using pre-defined structure model	Attribute

Background image: Landsat-8 OLI image of Diamantina Gates, southwestern Queensland, acquired in winter 2013. Due to constraints imposed by the adjacent Goyder and Hamilton Ranges, the Diamantina River narrows to about 1 km wide at Hunters Gorge, which is located near the centre of this image.
Source: Craig Shephard, DSITI

Figure 1.1 Measurement and analysis of image features

Interpreting imagery from EO sensors is an indirect process, whereby a measurement model is used to transform EO observations to measurements of some measurable property, and a structure model is used to relate those measurements to application-specific attributes.



Adapted from: Harrison and Jupp (1989) Figure 33

In the context of Figure 1.1, the structure model defines how the properties that may be measured remotely, such as reflectance or emission, represent the physical parameters we want to analyse. For example, various studies have established that remote measurements of thermal infrared emission of a surface can be directly related to its temperature (see Volume 1B—Section 7). However, other variables such as atmospheric conditions (water vapour and aerosols) and surface emissivity, can modify this relationship, so need to be considered when deriving temperature information from thermal infrared images (see Volume 3).

Further approximation to the image representation of the original parameters, in both the spectral and spatial dimensions, is added by the measurement model. For example, when the quantisation levels for a thermal scanner measure thermal infrared emission in steps that are equivalent to 0.1°C then the finest interval for temperature categories from such an image is automatically defined. Similarly, if the spatial resolution of one sample measurement (that is, the pixel size) is 5 m, we cannot use the resulting image to precisely locate objects that are smaller than this size (even if their presence can be detected). Such characteristics of the measurement model are critical for reliable and repeatible mapping, monitoring or modelling of Earth features and attributes using EO datasets (see Section 2.2.3).

1.2 Implicit Assumptions

The major assumptions which are implicit in acquiring and analysing EO data are summarised in Table 1.2. Each of these factors affects the correlation between features in the ground scene and results derived by analysis of digital EO data, and thus affect the accuracy of any image analyses. The five stages listed in Table 1.2 fundamentally address the following questions:

- has the best available imagery been selected (see Section 2.2)?
- is the data quality adequate to the intended purpose(s) of the analysis (see Volume 1)?
- has the data been processed appropriately (see Volume 2A and Sections 2.3 and 2.4 below)?
- has the accuracy of results been assessed and can the results be reproduced precisely (see Volume 2E)?
- have results been reported to honestly state known limitations and assumptions and clearly represent the final interpretation?

The answers to these questions directly impact the use of EO data for monitoring land cover/use changes (see Volume 3A). For example, differences between two image classifications of a ground scene may be due to actual changes on the ground, or to differences in either the data acquisition mode or the data analysis and interpretation methods that were used (see Volume 2E). Thus, before conclusions are deduced from the results of such analyses, an interpreter should investigate all possible causes for the observed differences.

Image processes that are commonly used with EO datasets are described in Volumes 2A, 2B and 2C and methods that can be used to assess the accuracy of EO analyses are detailed in Volume 2E. A further consideration relating to analysis of multi-temporal EO data is the delivery of results. Examples of this delivery stage may involve a wide range of presentation styles, including monitoring of specific environmental attributes and biophysical models. Australian examples of such outcomes are presented in Section 14 below and in Volume 3.

Table 1.2 Assumptions implicit in the analysis of EO data

Stage	Factor	Assumptions
Selection	Spatial resolution	Adequate to indicate shape, size and texture of target objects
	Spectral resolution	Discriminates target from background in terms of bandwidths and quantization
	Radiometric resolution	Detects differences in reflectance and/or emission required to differentiate target features
	Temporal resolution	Observes target at appropriate times of day, season or climatic event
Acquisition	Reflected and/or emitted radiance	Target objects are characterised by upwelling radiance
	Atmosphere	Atmospheric attenuation of upwelling radiance is negligible or can be corrected
	Analogue-to-digital conversion	Negligible A/D converter noise and required radiometric quantization level available from sensor
	Image transmission and reception	Data encoding and compression/decompression activities do not increase radiometric or geometric errors
Processing	Image calibration	Geometric correction approach is relevant and accurate
		Radiometric correction approach accounts for atmospheric and sensor 'noise'
	Pre-processing	Transformations applied to enhance image features do not misrepresent the relationship between image values and target features
	Analysis	Processes selected are appropriate to image dimensions and target attributes
Interpretation	Provenance	Data provenance is understood and retained
	Constraints	Limitations of selected processes are understood and acknowledged
	Validation	Results are verified against independent datasets
Delivery	Reporting	Provenance and limitations are clearly documented
	Accuracy	Accuracy assessment is reported in terms of methods and results
	Presentation	Analysis results are presented in a transparent and unambiguous manner

Adapted from Duggin and Robinove (1990)

1.3 Layers of Complexity

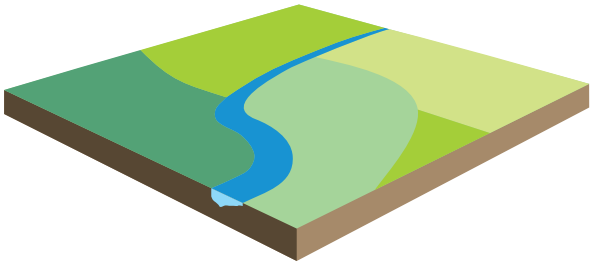
Some of the characteristics of EO data that complicate its analysis can be seen in another example. Let us start with imaging a very simple scene, which could be compared with a thematic map. In the simplest case, the scene would be a flat surface with flat features, which are identified by distinct, solid colours (see Figure 1.2a). We will assume that the scene is imaged with sufficient spatial resolution to clearly represent the boundaries between features and sufficient spectral resolution to differentiate all feature colours. Classification of this image would be able to reliably use the colour values to identify each feature in the image as a different class. In this case, assuming constant atmospheric conditions, images recorded with different illumination and viewing angles would be virtually identical.

If the same 'scene' is folded and bent before imaging, we would have to consider the effect of a non-flat surface on the imaged values. If the surface had considerable undulation, distinctly different images could result when the illumination and viewing positions are varied due to changes in shadowing. Surface topography alone could result in different surface colours being imaged with the same image values and/or the same surface colours being imaged with different image values (see Figure 1.2b).

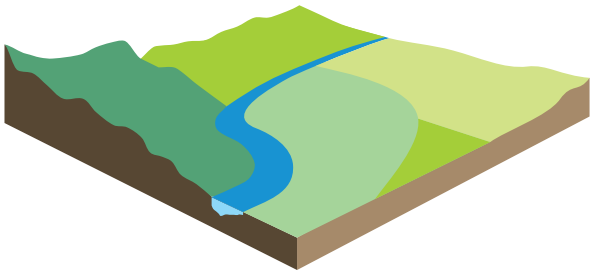
A further level of complexity could be introduced by giving the surface colours some three-dimensional form. For example, the dark green colour may comprise groups of tall, spheres (like trees) and the pale green patch may include boxes (like buildings). Variations in illumination and viewing will now be further complicated by shadowing within features as shown in Figure 1.2c.

Figure 1.2 Complexity levels in EO imagery

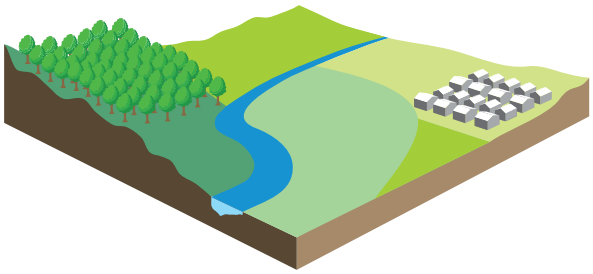
a. A simple scene would comprise a flat surface with flat features having distinct, constant colours.



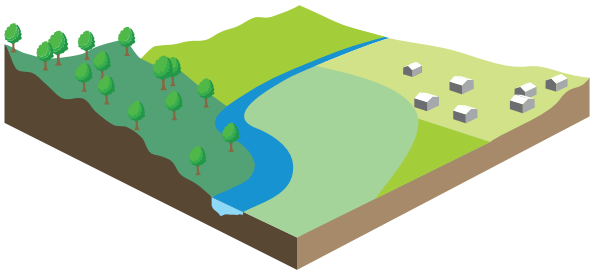
b. The simple scene with added topography introduces surface shading.



c. When three-dimensional features are imposed on the simple scene with topography shadowing between features is introduced.



d. By varying the spacing within the three-dimensional features, the interactions between topography, surface colour and feature separation become more complex.



Source: Harrison and Jupp (1993) Figure 4

As well as a three-dimensional surface containing three-dimensional features, we can allow variations of structure and colour within each surface feature. For example, the trees might be spaced so that some are adjacent while some are separated against a background of differing colour, and the buildings could be dispersed unevenly (see Figure 1.2d). These structural factors will vary not only the average colour of each surface feature but also the variance of each surface colour. This variance will also differ when the scene is imaged using different pixel sizes, and different illumination and viewing angles. A finer resolution image, such as Landsat TM rather than MSS, will show more detail about smaller objects, but this additional information will necessarily increase the variance of the image (see Volume 2A—Section 8.1). For example, some composite features, like woodland areas, are resolved into trees and background patches in fine resolution images (such as Landsat TM) but may conveniently blend into a single, distinct patch when imaged at coarser resolutions (such as Landsat MSS; see Volume 2A—Section 8.2).

The Earth's surface is comparable to the final scene described above and illustrated in Figure 1.2d, having variable relief and a range of three-dimensional structures (such as vegetation and buildings) with variable spacing. Representation of these features in terms of mean and variance in an image can also vary widely with different spatial and spectral resolutions. Sensor design and calibration characteristics can mean that the same reflectance level can sometimes be recorded as different values in the image (see Volume 1A). Remote imaging will be affected by illumination differences, since the Sun position varies with time of day and time of year, as well as view angle differences for different sensors and platforms (see Volume 1B). A further variable affecting EO data is atmospheric attenuation, since the condition of the Earth's atmosphere will modify the surface reflectance measured by sensors on aircraft and satellite platforms (see Volume 2A). Reliable and repeatible analysis of EO datasets requires a comprehensive understanding of these layers of complexity and their interrelationships with processing methods.

1.4 The Surface Mosaic

In addition to the layers of complexity inherent to EO imagery discussed above, it is instructive to view each landscape being imaged as 'a mosaic of tiles of odd shapes and sizes' (Speight, 2009). A similar view can be taken for features of the 'waterscape' (see Volume 3B).

Relating features and patterns in the surface mosaic to pixel values in a EO dataset relies on the:

- ability of the type of measurements being recorded in the image to resolve the surface features and patterns; and the
- effectiveness of the models that implicitly or explicitly relate surface features or physical processes to these EO-based measurements.

While the intrinsic scale of selected features and patterns on the Earth's surface may appear to be discernible in EO imagery, each image only captures those features and patterns of the surface mosaic at a single spatial scale. Different features and patterns, however, occur at different spatial scales, and can vary from place to place and from time to time. As such it is not possible for any single EO image, at one spatial resolution, to capture all features and patterns that comprise the imaged surface mosaic.

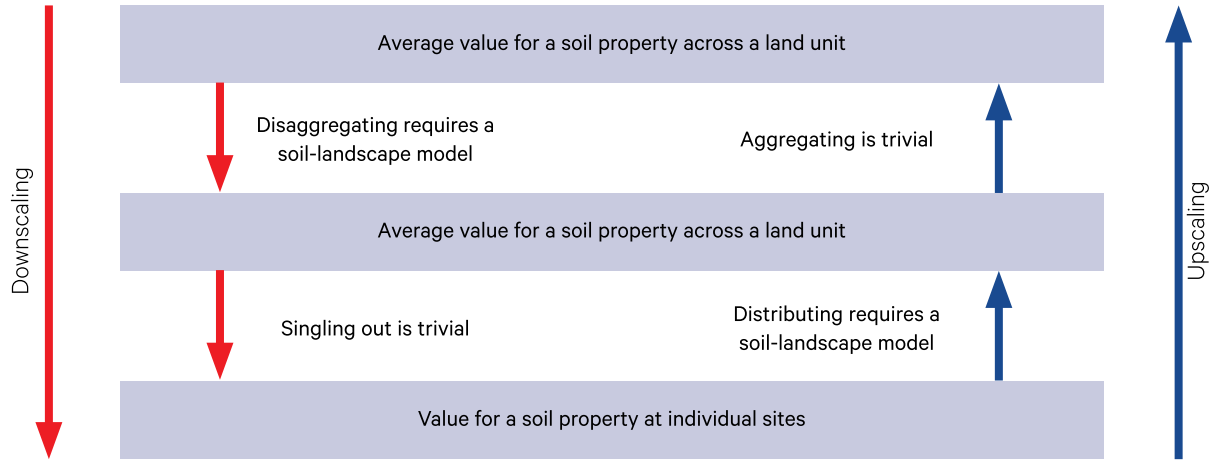
In order to sensibly interpret the features and patterns in a given EO image we need to consider whether their representation in the image adequately mirrors their characteristics on the Earth's surface.

This is not just a matter of spatial resolution, as some features and patterns with a larger extent are more easily distinguished in lower resolution imagery. When classifying some vegetation types, for example, a larger pixel size usefully integrates the variation within a vegetation canopy, making it more spatially coherent. The concept of a mosaic model for image classification is introduced in Volume 2A and expanded in Volume 2E.

Traditional landscape mapping approaches have long grappled with mismatches between the scales of measurement, estimation and prediction in terms of specific landscape features (Gallant *et al.*, 2008). Many landscape mapping methodologies embrace a multi-scale approach (see Excursus 1.1 and McKenzie *et al.*, 2008b), including recommended scales for mapping different surface features (see Table 1.4 below). For example, a hierarchy of spatial scales describing soils and landscapes is shown in Table 1.3. Within this hierarchy, however, the spatial and temporal scales for various landscape processes may not be correlated (Blöschl and Sivapalan, 1995). Further, while most landscape attributes can be represented as a nested hierarchy, with the associated levels being sequenced in time and space, the processes of moving either up or down the scale hierarchy (upscaling and downscaling respectively) present significant technical challenges (see Figure 1.3).

Figure 1.3 Upscaling and downscaling of soil data

While some forms of upscaling and downscaling are trivial, others are technically challenging.



Adapted from Gallant *et al.* (2008) Figure 3.7

Similiarly, hierarchies of imagery over a range of scales may be appropriate to resolve relevant surface features and patterns and their spatial interrelationships (see Volume 1A—Section 1). Since the design of EO sensors traditionally imposes a choice between higher frequency of imagery at lower spatial resolution or vice versa (see Volume 1A—Section 13), multi-scale hierarchies of images can allow the detailed information acquired from field sites and higher resolution imagery to be extended over a larger area (see Figure 1.4).

Selection of appropriate imagery for EO analyses is further discussed in Section 2.2. The landscape mosaic has particular relevance to image integration when datasets with differing spatial resolutions are being merged. While technological solutions exist to resample imagery to a range of scales (see Volume 2B—Section 5), the critical consideration here is whether the resulting scale is appropriate to the spatial dimensions of the features and patterns that are being studied in the underlying surface mosaic. Implications of the surface mosaic, in the context of upscaling and downscaling of EO images, are explored in Section 4.3 (see also Volume 1B—Section 2.4).

Table 1.3 Scale hierarchy for soil and landscape data

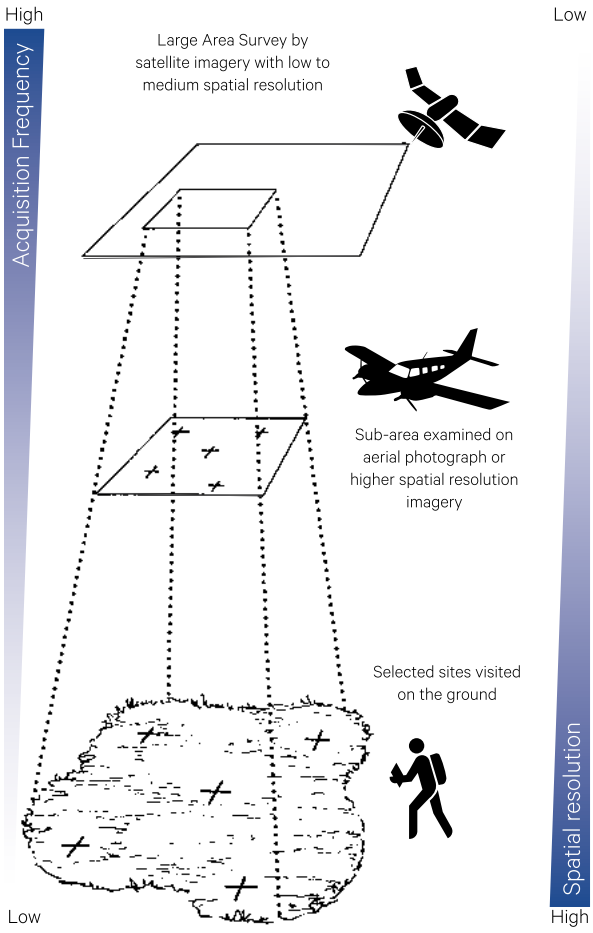
Levels of organisation in this hierarchy can be ordered in both space and time, with lower levels being characterised by smaller areas and shorter times, and higher levels being characterised by larger areas and longer time intervals. Upscaling involves moving up the scale hierarchy, while downscaling moves down.

Level	Region	Significance
I+6	World	Provide context
I+5	Continent	
I+4	Region	
I+3	Catchment	
I+2	Catena/farm	
I+1	Field	
i	Site	Obtain field measurements
i-1	Soil horizon	
i-2	Soil structure	
i-3	Basic structure	Understand mechanisms
i-4	Molecular interaction	

Adapted from: Gallant *et al.* (2008) Figure 3.1

Figure 1.4 Scales of EO

Analysis of low and medium resolution imagery can be more easily related to ground data through high resolution imagery. Low resolution imagery tends to be acquired more quickly and cover a larger extent, so is most appropriate for large area surveys. Selected locations within the surveyed area would typically be checked using data with higher spatial resolution, such as aerial photography. Some of these locations would then be visited in the field to verify the image analysis. This scaled approach enables the results of expensive and time-consuming field work to be extended to a larger area.



Adapted from: Harrison and Jupp (1989) Figure 39

Excursus 1.1—Landform Models

Source: Speight (2009)

The landform description methods recommended by Speight (2009) divide the landscape mosaic into two different scales, with the larger tiles comprising mosaics of the smaller tiles:

- landform patterns—larger tiles (>600 m across); and
- landform elements—smaller tiles (~40 m across).

The most appropriate scales for mapping these tiles as models of the landform are listed in Table 1.4. Thus, the landform pattern model is considered appropriate for parts of the landscape that span a complete toposequence, whereas the landform element model is more relevant to areas that contain only part of the toposequence.

Table 1.4 Landform models relating to mapping scales

Map scale	Minimum width of mapping units (m)	Appropriate landform model for mapping	Recommended uses
1:500,000	1500	Landform pattern	National/regional resource inventory
1:250,000	750	Landform pattern	Overview of land resources/status
1:100,000	300	Landform pattern	Land use suitability and strategic planning
1:50,000	150	Landform pattern	Major land uses and large catchment management
1:25,000	75	Landform pattern/ Landform element	Low intensity land use planning and medium catchment management
1:10,000	30	Landform element	Intensive land use developments and small catchment management
1:5,000	15	Landform element	Urban and farm planning

Source: Speight (2009) Table 1; Schoknecht *et al.* (2008) Table 14.1

In the Speight (2009) system, landform patterns are described in terms of around 40 predefined types. These types are based on the landform attributes that are observed within sample circles of radius of 300 m in the landform pattern. Examples of landform patterns include mountains, escarpment, hills, floodplains, dunefields and coral reefs. The attributes that determine landform patterns are:

- relief—difference in elevation between the highest and lowest points on the land surface;
- modal slope—most common slope class within the landform pattern;
- stream channel occurrence—described by spacing, development, depth-to-width proportions, pattern of streams, and integration and directionality of channel network;
- mode of geomorphological activity—gradational (eroded/aggragated) or anti-gradational (elevated/excavated/subsided);
- geomorphological agent—gravity, precipitation, stream flow, wind, ice, standing water, internal forces, biological agents or extraterrestrial forces;
- status of geomorphological activity—whether agent is active, episodic or relict; and
- component landform elements—as observed.

Landform elements are defined by the following attributes within circles of 20 m radius:

- slope—in terms of eight categories based on inclination;
- morphological type—flat, crest, slope and depression (open and closed), plus position in toposequence and relative slope inclination for non-flat surfaces;
- dimensions—length, width and height;
- mode of geomorphological activity—as above; and
- geomorphological agent—as above.

Examples of landform elements include cliff, footslope, swamp, swale and valley flat.

1.5 The Bigger Picture

As introduced in Section 1.2, an implicit goal of any image interpretation exercise should be that its results can be verified and reproduced independently. More than for any single image processing task, however, the integration of bi-temporal and multi-temporal EO datasets demands that the interpreter retains awareness of the 'bigger picture'. This can be condensed into two questions:

- what is the goal of this analysis?; and
- how can 'success' be assessed?

These questions need to be answered before the most appropriate data, processes and methods can be selected for any particular study (see Section 2).

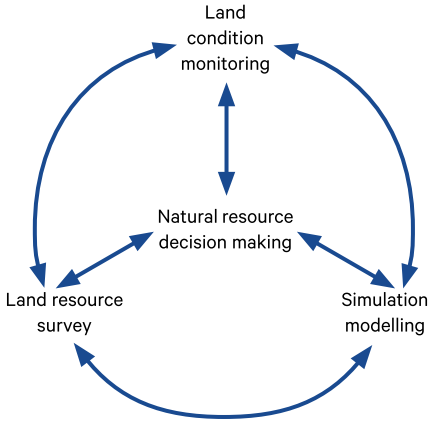
In the following sections we will consider established methods for calibrating and pre-processing multiple images (see Sections 3 and 4), processing bi-temporal EO image pairs (see Sections 5 to 7), and creating and analysing time series datasets (see Sections 8 to 11). The greater goal of most EO analyses involving bi-temporal and multi-temporal datasets, however, is to interpret measured properties of the target object (or surface) in terms of a specific application, using a pre-defined structure model (see Section 1.1). In this context the underlying EO framework must be explicitly acknowledged and considered in all stages of processing, interpretation and delivery.

It is also worth remembering at the outset of any EO-based study that mapping of Earth surface resources—whether derived directly from field observations or based on EO analyses—only describe biophysical attributes at one instance in time. To fully understand environmental change, however, mapping results must to be integrated with monitoring and modelling activites (see Figure 1.5). This integrated approach is required for competent land-use planning, assessing land management practices, understanding natural processes, forecasting future environmental conditions, and formulating appropriate policies for managing our natural resources (McKenzie *et al.*, 2008a).

Some of the current limitations with natural resource maps in Australia include incomplete and inconsistent coverage, inappropriate scaling, incompatible survey methods, insufficient rigour when measuring key attributes, lack of accuracy assessment, and lack of currency (McKenzie *et al.*, 2008a). Monitoring of natural resources requires a baseline from which observed changes over time can be measured (see Section 14.2 and Volume 3A), whereas modelling enables prediction of future conditions (see Section 14.3 and Volume 2A—Section 2.3). Modelling also provides an opportunity to explore the potential ramifications of changes in resource attributes, management regimes, land uses, or policies. The complementary relationships between mapping, monitoring and modelling are summarised in Table 1.5. Integrated mapping, monitoring and modelling systems that are based on EO datasets are introduced in Section 14, and further detailed in Volume 3.

Figure 1.5 Mapping, monitoring and modelling for natural resource management

To monitor environmental changes in a given landscape, these complementary activities need to focus on the key events and processes that drive those changes. The temporal scale for specific changes may span decades, centuries or millennia.



Adapted from: McKenzie *et al.* (2008a) Figure 1.1

Table 1.5 Complementary benefits of mapping, monitoring and modelling

Relationship	Benefits
Mapping → monitoring	Spatial framework for selecting representative sites; System for spatial extrapolation of monitoring results; and Broad assessment of resource condition.
Monitoring → mapping	Quantifies and defines important resource variables for mapping; and Assesses land suitability over time (including risk assessments for recommended land management).
Modelling → monitoring	Determines whether trends in specific land attributes can be successfully detected with monitoring; and Identifies key components of system behaviour that can be measured in a monitoring program.
Monitoring → modelling	Validates model results; and Provides data for modelling.
Modelling → mapping	Allows spatial and temporal prediction of landscape processes.
Mapping → modelling	Provides data for modelling; and Provides spatial association of input variables.

Source: McKenzie *et al.* (2008a) Table 1.1

1.6 Further Information

McKenzie *et al.* (2008b)

1.7 References

Blöschl, G., and Sivapalan, M. (1995). Scale issues in hydrological modelling: a review. *Hydrological Processes*, 9, 251–299.

Duggin, M.J., and Robinove, C.J. (1990). Assumptions implicit in remote sensing data acquisition and analysis. *International Journal of Remote Sensing*, 11, 1669–1694.

Gallant, J.C., McKenzie, N.J., and McBratney, A.B. (2008). Scale. Ch 3 in *Guidelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

McKenzie, N.J., Ringrose-Voase, A.J., and Grundy, M.J. (2008a). Rationale. In *Guidelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J. (Eds) (2008b). *Guidelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. CSIRO Publishing, Melbourne.

Schoknecht, N., Wilson, P.R., and Heiner, I. (2008). Survey specification and planning. Ch 14 in *Guidelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

Speight, J.G. (2009). Landform. In: *Australian Soil and Land Survey: Field Handbook* (3rd Edn). (Eds: National Committee on Soil and Terrain) pp. 15–72. CSIRO Publishing, Melbourne.

2 Multi-temporal EO Datasets

One of the strengths of EO is the provision of regularly updated imagery of the Earth, from which changes in surface features can be observed. While changes in surface features can occur for many reasons, most changes are driven by:

- climate variations, including seasonal cycles, short-term aberrations and long-term natural changes with consequent impact on soil moisture and phenology;
- landscape changes relating to erosion and deposition;
- natural disasters, including fire, floods, wild weather, tectonic activity and landslides;
- ecological succession in vegetation; and/or
- human activities, such as land cover changes, silviculture, agriculture, urban development and mining (see Volume 3).

All differences between bi-temporal and multi-temporal imagery, however, do not necessarily indicate real feature changes on the ground (see Section 1). As detailed in Volume 2A—Section 3, image differences can also result from changes in atmospheric condition, sensor characteristics, and viewing and illumination geometries. Accordingly, some form of image calibration is essential before surface changes can be reliably detected in bi-temporal or multi-temporal imagery (see Section 3).

When investigating changes in Earth surface features as viewed by EO sensors, it is tempting to think in terms of some 'equilibrium' state from which 'change' occurs. However, the Earth represents a dynamic system, and repeated observations of that system can vary for many reasons. Some of the temporal variations that can be observed in EO imagery derive from correlated changes—these image variations are correlated with natural environmental factors but do not necessarily indicate a change in the surface features. For example, terrain shading varies with Sun position and terrain characteristics, and is correlated with incident radiation. In EO imagery, variations in terrain shading may 'appear' as a change in land cover, but changes in shading are essentially ephemeral. Seasonally varying levels of solar radiation directly impact surface features but, on any given day, these are not necessarily represented by the current terrain shading. Rather, such seasonal patterns are inferred from a changing pattern that correlates with them. While imaging artefacts can be removed using physical models (see Section 3.1) and 'random noise' can be removed using statistical methods (see Volume 1B—Section 2), spuriously correlated interactions between the sensor 'view' and the Earth's surface are much harder to remove.

In this section we will firstly review some of the characteristics of multi-temporal imagery that directly impact its utility (see Sections 2.1 and 2.2), before considering appropriate processes (see Section 2.3) and methodologies (see Section 2.4) for such datasets.

Background image: ASTER image of Lake Mackay, Australia's fourth largest lake, acquired on 9 September 2010. This ephermal salt lake is located along the WA/NT border. **Source:** NASA ASTER gallery. Retrieved from *https://earthobservatory.nasa.gov/images/84984/australias-ephemeral-lake-mackay*

2.1 Bi-temporal and Multi-temporal Imagery

Time is but one dimension that is sampled by EO imagery. The four data dimensions inherent to EO datasets are introduced in Volume 1B—Section 1 and summarised in Table 2.1. Reliable processing of bi-temporal and multi-temporal image datasets requires that all images captured at different dates and/or times be comparable, that is, there is adequate consistency between their spectral, spatial and radiometric dimensions.

The time dimension, however, presents unique some unique challenges to an interpreter. As detailed in Section 1.3, each image represents a complex ground scene. While the spectral, spatial, radiometric and temporal characteristics of an EO sensor can be tailored to the idiosyncracies of particular ground features, the way those ground features change through time cannot be known or predicted with certainty. While the 'usual' range of spectral responses, for example, of a ground feature can be pre-determined, anomalies can occur. During extreme events, such as fires and floods, the designed sensitivity range of imaging sensors may be exceeded, to render saturated pixels in imagery at a time when reliable data is most needed (see Volume 1A—Section 13).

Table 2.1 Sampling dimensions in EO imagery

Dimension	Characteristic		
	Resolution	Density	Extent
Spectral	Width of each wavelength channel	Number of channels detected by sensor	Range of wavelengths covered by all channels
Spatial	Ground area imaged per optical pixel	Number of pixels and lines in image	Area covered by image
Radiometric	Smallest change in detected energy that would be represented as a different image brightness level	Number of gradations (grey levels) used to represent full range of radiances that could be detected by sensor	Actual range of radiances detected in each channel
Temporal	Time period over which each image is acquired	Frequency of successive image acquisitions	Total time period for which this imagery is available

Adapted from Emelyanova *et al.* (2013)

2.1.1 Conceptual structure

The conceptual structure of an EO image is introduced in Volume 2A—Section 1. For a single image, acquired on a particular date, this structure resembles a three-dimensional matrix as depicted in Figure 2.1. A bi-temporal image simply comprises two images acquired at different times (see Figure 2.2a). Sets of multi-temporal images can be considered as comprising multiple image matrices (see Figure 2.2b). The critical aspect of this structure is that any given pixel location in a set of multi-temporal images represents exactly the same ground area in all of the images in that set.

The number of images in a multi-temporal dataset can vary widely, potentially including tens of thousands of images. In terms of processing operations, it is convenient to distinguish between those that may be used to detect or highlight differences and similarities between two images and those that are relevant to a longer sequence of images.

Figure 2.1 Structure of an EO image

A multi-channel image comprises three data dimensions—pixels, lines and channels. Pixels (or 'elements' or 'samples') along the line can be considered as columns in a matrix. Lines in the image are like rows of a matrix, with each channel being an attribute layer in the matrix.

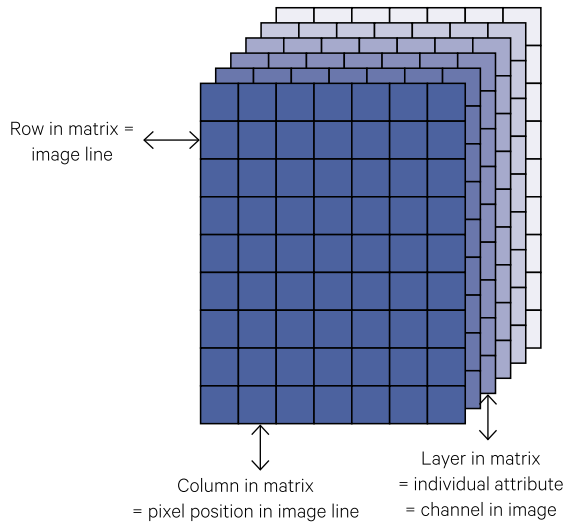
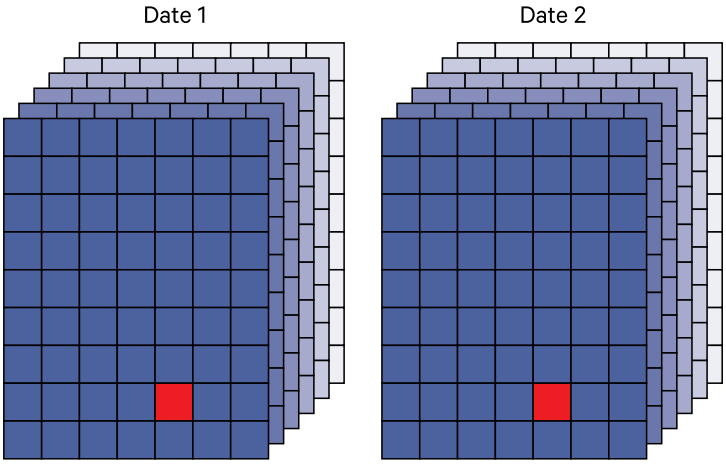


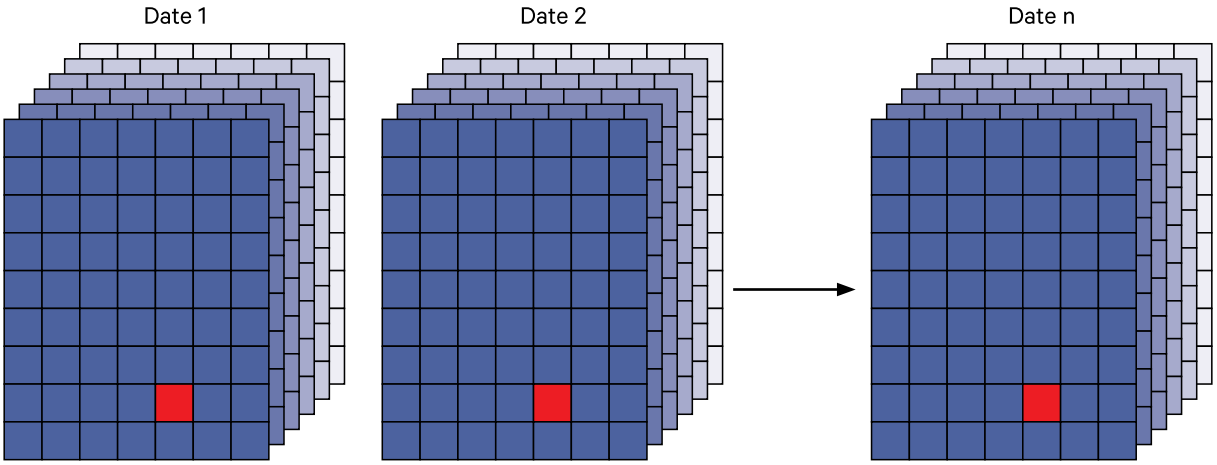
Figure 2.2 Bi-temporal and multi-temporal image sets

A given pixel location within a bi-temporal or multi-temporal image set (such as the red pixel below) is assumed to represent exactly the same target area on the ground.

a. A bi-temporal image set comprises two images acquired at different dates.



b. A multi-temporal image set comprises multiple images acquired at different dates, usually ordered chronologically.



Operations that are commonly used for a calibrated, bi-temporal pair of EO images include:

- basic operations such as differencing, segmenting, colour balancing (see Section 5);
- fusion methods such as pan-sharpening (see Section 6.1) and temporal interpolation (see Section 6.2); and
- processes for detecting changes (see Section 7).

Longer sequences of EO images can be processed as time series data if their acquisition dates are evenly spaced (see Section 8). Many, but not all, multi-temporal datasets of EO images are acquired at regular time intervals, so can be treated as time series archives. Relevant processes for these datasets include characterising temporal trends (see Section 9) and creating composite images (see Section 10).

2.1.2 Satellite-acquired datasets

Satellite imagery has been acquired for several decades and some global EO datasets now span a substantial time series. Most satellite sensors acquire images at standard times of the day on a predefined acquisition schedule from a nadir viewing position (see Volume 1A). This ensures that the image archives for a particular sensor are imaged with the same sensor and illumination geometry. While solar illumination changes with the seasons, such variations can be modelled with sufficient accuracy (see Volume 2A—Section 3). However, when imagery acquired by different sensors is being compared or merged, differences in the diurnal acquisition times can become significant. Similar challenges result when some images are acquired from non-nadir viewing positions, as is possible for some satellite and most airborne sensors.

Table 2.2 Global EO time series datasets

Satellite (Agency)	Sensor	Spectral bands (in latest sensor)	Spectral range (in latest sensor)	Radiometric Quantisation (bits)	Spatial Resolution (m, at nadir, in latest sensor)	Swath Width (km, at nadir, in latest sensor)	Revisit Time (days)	Image Archive
ALOS (JAXA, Japan)	AVNIR-2	4	Optical	8	10	70	14–46	2006–2011
	PALSAR	1	L-band SAR in 3 modes	5	1–100	25–490	14–46	2006–2011; 2014–present
	PRISM	1	Panchromatic/ stereo	8	2.5	70 (35 in triplet mode)	14–46	2006–2011
TIROS/ AVHRR (NOAA, USA)	AVHRR	6	Optical/ thermal	10	1,100	~2700	daily	1978–present
Landsat (NASA/ USGS, USA)	MSS	4/ 1	Optical/ thermal (L-3)	8	68 x 83	185	16–18	1972–1997
	TM/ETM+	7	Optical/ thermal	8	30/ 120	185	16	1987–2011
	OLI/TIRS	11	Panchromatic/ optical/ thermal	12	15/ 30/ 100	185	16	2014–present
Terra (NASA, USA)	ASTER	9/ 5	Optical/ thermal	8/ 12	15/30/ 90	60	4–16	1999–present
Terra and Aqua (NASA, USA)	MODIS	36	Optical/ thermal	12	250/ 500/ 1,000	2330	1–2	2000–present
Quickbird (Digital Globe, USA)	Quickbird	5	Panchromatic/ optical	11	0.55/ 2.16 (at 400 km altitude)	14.9 (at 400 km altitude)	1–3.5	2001–2015
SPOT (Airbus DS, France)	SPOT PAN SPOT MS	5	Panchromatic/ optical/ stereo	8–12	1.56/ 6	60	1–3	1986–present
Geoimage (USA)	GeoEye	5	Panchromatic/ optical	11	0.41/ 1.65	15.2	£3	2008–present
MAXAR (formerly DigitalGlobe, USA)	WorldView	1/ 8–16	Panchromatic/ optical	11	0.31–0.46/ 1.24–2.0	17.6/ 13.1–16.4	<1	2007–present/ 2009–present

When using bi-temporal or multi-temporal imagery for change detection studies, it is clearly essential that the differences between images derive from changes in surface features rather than differences in their acquisition parameters (see Section 7). Accordingly, any imagery destined for such analyses need to be carefully calibrated to an absolute standard as described in Section 3.

Finally, some of the global EO archives available as time series datasets are listed in Table 2.2. Very high spatial resolution datasets with rapid revisit times, such as FORMOSAT-2 and IKONOS, can also be used to generate EO time series. Several EO data suppliers provide time series datasets for both the image reflectance (or albedo) data and for transformations, such as the Normalised Difference Vegetation Index (NDVI) or the Enhanced Vegetation Index (EVI), which have been derived from image reflectance (see Volume 2C). Examples of such composite products for the MODIS sensor are given in Excursus 10.1.

2.2 Appropriate Data

In Section 1 above, we reviewed some of the foundations of EO image datasets, including an appreciation for the underlying surface mosaic that was imaged. For any EO analysis exercise, a good understanding of the key features and patterns in this mosaic is a pre-requisite for reliable results. In this context, the following EO variables are particularly relevant to the selection of appropriate bi-temporal or multi-temporal imagery:

- acquisition dates and times (see Section 2.2.1);
- issues relating to image spatial scale (see Section 2.2.2); and
- spectral and radiometric differentiation (see Section 2.2.3).

2.2.1 Timing

Clearly, to detect a particular surface feature in an EO image, the timing of image acquisition must capture the differentiating characteristics of that feature in the surface mosaic. This requirement is just as, if not more, relevant when selecting bi-temporal or multi-temporal imagery, especially when monitoring multiple stages in some environmental process, such as crop phenology (see Volume 3A) or flooding (see Volume 3B). In this context, timing relates to both time of year (season) and time of day (see Volume 1B—Section 1).

The timing of image acquisition is of primary importance in change detection exercises. When comparing two images, familiarity with the 'initial state' image—insight that links its reflectance values to relevant properties of the surface features being imaged—is critical. Ideally, the timing of this image will maximise the contrast in surface features of interest and minimise 'contamination' from imaging artefacts (including atmospheric, illumination, and viewing angle effects). Similarly, the timing of the 'final state' image needs to highlight the anticipated changes in one or more particular surface feature(s). Anniversary dates are often recommended to highlight annual changes in surface features and minimise variations associated with the Sun and sensor geometries. For ecosystem studies, however, phenological cycles for different surface features can still vary intra-annually even when the features themselves have not changed (Coppin *et al.*, 2004). Variations in soil moisture levels can also significantly impact image radiance, especially in sparsely vegetated areas.

Thus, an understanding of the surface feature(s) being studied is essential for selecting appropriate image dates. This requires that the:

- spectral, spatial, radiometric and temporal resolution and extent of the images are suitable to observe and isolate the feature(s) of interest relative to other features in the surface mosaic (see Section 2.1); and the
- time interval between selected (set of) images is sufficient to allow the change(s) of interest to be detected.

In general, while vegetation differences will be highlighted during the growing season, the precise timing of this varies for different types of vegetation and between different geographic regions. When multiple features are being studied, it may be more relevant to select dates that maximise the contrast in their reflectance properties, rather than maximise the reflectance of any individual feature. Likewise, the interval between an 'initial state' image and the 'final state' image must allow sufficient time for the surface feature(s) to change, and this may vary from very short intervals (minutes or hours) for natural disasters, to longer intervals (days or months) for agricultural crops, to much longer intervals (years to decades) for forest vegetation or land rehabilitation studies.

The impact of changes in Sun position on EO radiances is detailed in Volume 2A—Section 3. Radiometric calibration attempts to correct for these seasonal and diurnal changes in bi-temporal and multi-temporal imagery (see Section 3.1.1.2), and renders reasonable success in datasets derived from a single sensor. However, when using processing methods that interpolate imagery from one sensor, based on the image values acquired from a different sensor (see Section 6.2), potential limitations of radiometric calibration need to be considered.

The optimal timing of images for bi-temporal or multi-temporal analyses will also be determined by climatic factors such as cloud cover and aerosol concentration. Unfortunately, however, EO imagery is not always available for the most appropriate dates. In many regions of the world, for example, cloud-free imagery are only available for limited seasons, which further restricts the selection of appropriate imagery for change detection studies. Similarly, available imagery may not coincide with maximum flood levels or peak fire activity. The final selection of image dates may become a compromise between the most suitable environmental time and the most convenient logistical time with respect to available and/or affordable imagery. Where sub-optimal pairs of images are used for change detection, their limitations need to be considered during the analysis, and acknowledged in reporting outcomes.

2.2.2 Scaling

A limiting factor for differentiating many surface features and processes is image scale (see Volume 1B—Section 2). Quite simply, does the pixel size allow the relevant feature(s)—or a representative characteristic thereof—to be discriminated? In some cases the size of surface features can change with time and that transition may be discernible in an appropriately-scaled sequence of images.

While image resampling methods are primarily used to correct geometric distortions in EO imagery, they can generate artefacts as well (see Volume 2B). Such potential problems may be more misleading when multi-resolution data from different sensors is processed to a common pixel size (see Section 4.3). Similarly, when fusion methods are used to interpolate a new temporal image from a lower spatial resolution dataset (see Section 6.2), the implications on the true scale of the resulting image need to be considered carefully. The importance of image scale in the classification process is further discussed in Volume 2E.

2.2.3 Spectral and radiometric differentiation

The importance of spectral and radiometric resolutions in EO imagery are introduced in Volume 1B—Section 1. Given the seasonality of changes in reflectance characteristics for many surface features, such as vegetation, selection of the most appropriate spectral bands across multiple image dates is important for any monitoring studies. While a particular spectral feature may be discernible in one season within a given radiometric range, it may not be discernible six months later. Likewise, in terms of diurnal changes in reflectance, different sensors that acquire imagery at different times of the day may not detect the same surface feature patterns. The interactions between spectral and spatial resolutions further complicate detection of some surface features when comparing imagery acquired by EO sensors that have different spatial resolutions (see Section 1.4).

Accordingly, when selecting image dates for a monitoring study, the likely discrimination of surface feature reflectance/emission across all dates needs to be determined. Pre-processing methods also need to consider the likely changes in all features across the dates of interest. For example, vegetation indices applied to sparse vegetation over wet soil may generate misleading results when compared with the same indices applied to vegetation over a dry soil background (see Volume 3A). Similarly, the limitations of all image processing algorithms need to be understood over the full temporal range of any bi-temporal or multi-temporal EO dataset (see Section 2.3).

2.3 Appropriate Processes

Just as a wide range of statistics can be computed for any set of numbers, a wide range of processes can be applied to EO images. The value of any particular process will depend on its relevance to the image being processed and the assumptions underlying both the data and the processing algorithm.

Unless supplied as Analysis Ready Data (ARD; see Section 3.2), all bi-temporal and multi-temporal EO image datasets require radiometric and geometric calibration (see Section 3.1 below; or Volume 2A—Section 3 and Volume 2B). This data preparation stage is essential to ensure that corresponding spectral radiance values are comparable across all image dates and that corresponding pixel locations reference the same ground positions.

A simple test for any image processing that involves image statistics is to ask the question: “Are the results reasonable?”. For example, PCA of vegetation imagery tends to generate ‘typical’ statistics for which spectral bands are mostly likely to contribute to which PC (see Volume 2C—Section 9). Should the results of this analysis differ markedly from the usual situation then there is good reason suspect errors have been introduced which warrant investigation.

Occam’s Razor: The more assumptions you have to make, the more unlikely the explanation.

The most appropriate processes to apply to a bi-temporal or multi-temporal image dataset will obviously depend on the purpose of the study. However, some simple guidelines are relevant to most image processing (see also Section 1.2):

- Is there a simpler way to do this? Simpler usually means more robust and repeatable, two characteristics that should underpin good science.
- How is the data provenance being maintained and updated through these processing operations (see Volume 2A—Section 2.1)?
- How can the results be validated? If you use all your data to create the model, how do you check it is correct? Approaches to validation in the context of image classification are detailed in Volume 2E.
- Do the results from this process acknowledge any known limitations in the input data?

Unfortunately, a nice-looking output image is rarely sufficient confirmation that the processing was appropriate.

2.4 Processing Methodologies

The concept of a processing methodology was introduced in Volume 2A—Section 2.2. Three broad stages of processing imagery are followed in many EO applications and involve:

- descriptive interpretation—delineating and identifying patterns in the image (such as identifying major land covers);
- exploratory data analysis—using statistical tools to detect and highlight features and patterns across the image (such as creating a land cover map); and
- modelling—relating images values for identified features with measured attributes of those features from another data source (such as characterising mapped land cover categories in terms of structural or spectral properties).

The following terms are used in these publications to distinguish between EO imagery that has been calibrated versus imagery that has been processed to deliver application-specific information (see Volume 2A—Section 2.1):

- data products—corrected image observations from EO sensors, such as surface reflectance products (see Section 3); and
- information products—derived from EO data products using defined processing methodologies for use in specific applications, such as vegetation greenness or land cover categories (see Section 14).

While some EO image processing exercises may not include all three processing stages, most EO information products include them as sequential steps required for product delivery. As well as selecting appropriate processes for a particular image dataset, those processes need to be applied in an appropriate sequence or order (see Section 2.4.1). The actual steps involved in processing are commonly standardised in a workflow (see Section 2.4.2).

2.4.1 Processing order

Just as computation in equations is governed by a pre-defined order of operations, the sequence of image processes applied to bi-temporal or multi-temporal datasets need to be ordered appropriately. General guidelines include:

- apply geometric corrections before radiometric corrections (see Section 3);
- segment irrelevant parts of the image before generating surface feature statistics;
- maintain consistency when processing each image in a large bi-temporal or multi-temporal dataset;
- apply data smoothing operations (such as for presentation purposes) after the primary processing; and
- document the details of all processing for future reference.

The processing steps for selected examples integrated systems are described in Section 14, and examples of EO information products are given in Section 14.1.

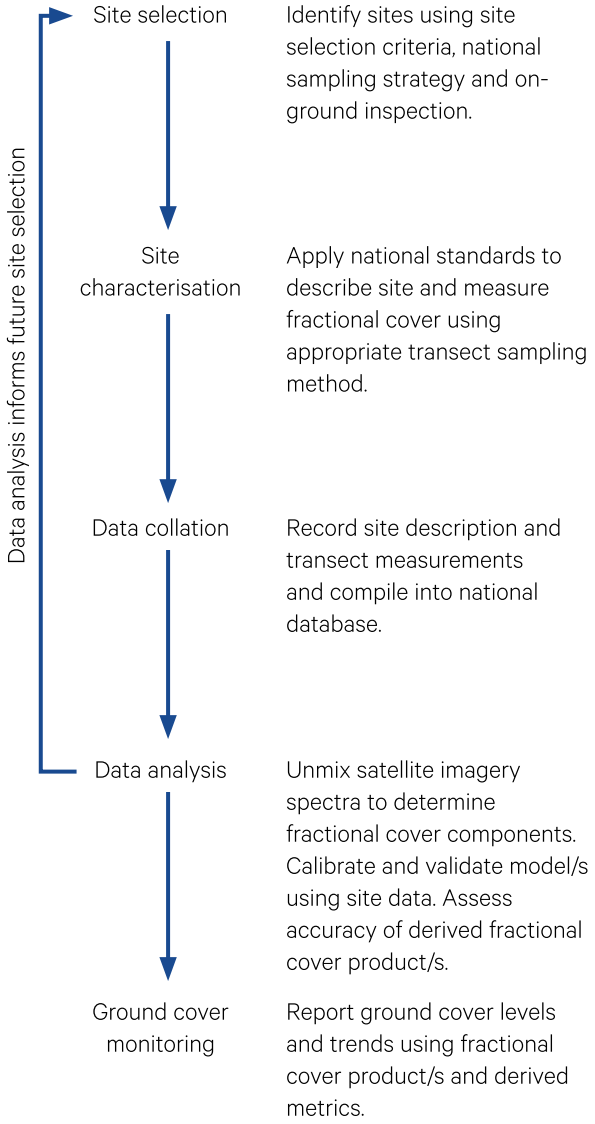
2.4.2 Processing workflows

Processing workflows (or process flows) are a convenient notation for defining the steps and conditions involved in a particular processing procedure. In a multi-step process, workflows are essential in some form to ensure that the overall method can be faithfully reproduced. Now that EO standard products are being used in a wide range of applications, including those involved with legislative monitoring such as compliance checking (see Section 14.2.2), their provenance and limitations must ensure scientific transparency and repeatability.

We have included examples of processing workflows in previous volumes (see Volume 2A—Sections 9 and 10), and will encounter several more in Volume 3. On example of a processing sequence is shown in Figure 2.3 for the Australian fractional cover product (see Volume 3A for details). This sequence summarises the major steps and cycles involved with generating this product.

Figure 2.3 Workflow for fractional cover product

This product Is shown as the banner image for Section 14 below and further described in Volume 3A.

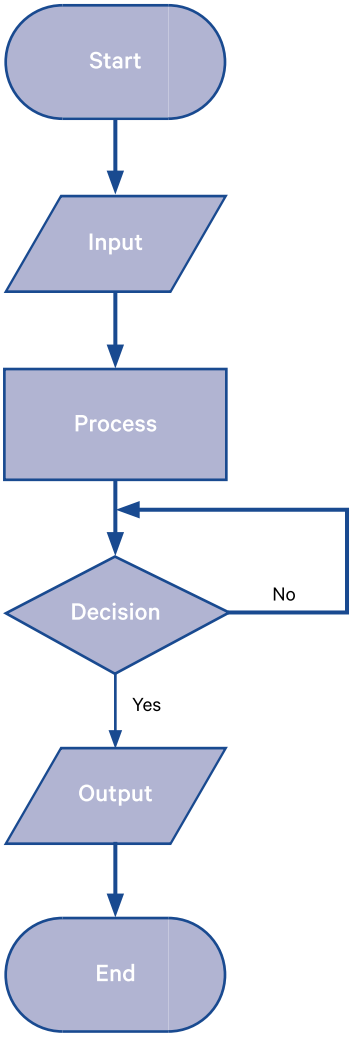


Adapted from: Stewart *et al.* (2014)

Standard symbols can be used in process flow charts to indicate the role of each segment, such as:

- oval—beginning or end of a process;
- rectangle—one step in the process;
- parallelogram—input or output data;
- diamond—decision points; and
- arrow—direction of connection between processes, data and decisions (see Figure 2.4).

Figure 2.4 Process flow chart example



2.5 Accuracy, Stability, Validation and Trust

Given the expanding range and volume of EO data and information products, there is growing awareness of the need to ensure that they achieve dependable standards of accuracy and stability. For a given EO product, these desirables could be defined as:

- accuracy—'the closeness of agreement between product values and true or reference values'; and
- stability—'the systematic error over a long period of time, typically a decade or more' (GCOS, 2011; Soto-Berelov *et al.*, 2018).

Validation, then can be defined as the process of determining the accuracy and stability of a product. For EO biophysical products, validation has been described as 'the process of assessing the uncertainty of higher level, satellite sensor derived products by analytical comparison to reference data, which is presumed to represent the true value of an attribute' (Soto-Berelov *et al.*, 2018).

In the context of EO products, it is increasingly important that an objective set of validation standards exist, against which product accuracy and stability can be measured. To this end, the Committee on Earth Observing Satellites (CEOS) established the Working Group on Calibration and Validation (WGCV) to develop international standards for calibration and validation of EO activities, including missions and data (Dowman, 2004; see Section 2.6). For example, four stages of validation have been identified for moderate resolution global products (see Table 2.3). In this validation hierarchy, each successive stage is more comprehensive than the previous one. Such systems will allow the accuracy and stability of EO products to be compared consistently.

Both data and information products need to be 'trusted' by users, however, before they are integrated into their decision-making processes. For example, the use of EO data and information products for emergency management was reviewed by Hudson (2015). This review yielded interesting insights into the attitudes of potential users towards EO products, and recommended a design framework that would increase product uptake by encouraging greater interactions between product designers and users (see Excursus 2.1).

Table 2.3 CEOS WGCV validation hierarchy for moderate resolution global products

Validation Stage	Description
1	Product accuracy has been estimated using a small number of independent measurements (typically < 30) obtained from selected locations and time periods and ground-truth/field program effort.
4	Product accuracy has been assessed over a widely distributed set of locations and time periods via several ground-truth and validation efforts. The spatial and temporal consistency of the product has been evaluated over globally representative locations and time periods. Results are published in peer-reviewed literature.
3	Product accuracy has been assessed over a globally distributed set of locations and time periods via several ground-truth and validation efforts. Product uncertainties have been well-established via independent measurements made in a systematic and statistically robust way that represents global conditions. Results are published in peer-reviewed literature.
4	Validation results for Stage 3 are systematically updated when new product versions are released and as the time-series expands.

Source: Soto-Berelov *et al.* (2018) Table 2.1

If you torture the data long enough, it will confess to anything.
(Ronald Coase)

Excursus 2.1—Why is an EO Product Trusted?

Source: David Hudson
Further Information: Hudson (2015)

Earth Observation (EO) aims to improve decision making processes by providing objective evidence. However, like all data and information products, the potential utility of any EO product is impacted by its perceived 'flexibility', that is, whether the design choices made by its creator(s) can be tailored to the desired purpose(s) of its users. The acceptance of any product can be described in terms of four sequential stages:

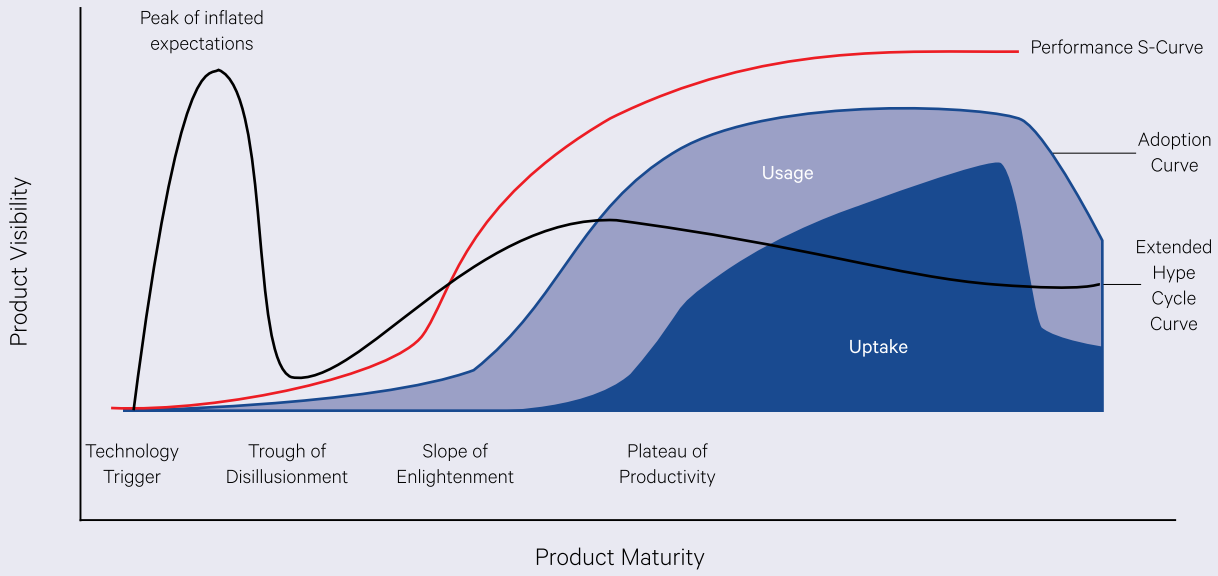
- awareness—user familiarity with the product;
- usage—frequency of access or use;
- uptake—inclusion of the product as evidence within a decision; and
- adoption—extent of uptake.

As illustrated in Figure 2.5, these responses vary during the life cycle of a product (see also Volume 1B—Section 10).

A recent study analysed the use of EO products (both data and information products) in emergency management (Hudson, 2015). As well as compiling an Australian disaster inquiry database, the study delivered a wide range of findings relating to user perceptions of EO products. Those findings that were directly relevant to the uptake of EO products are summarised in Table 2.4.

Figure 2.5 Lifecycle model for a sensor data product

The 'hype' cycle of a new product, when visibility to users is high, occurs early in its lifetime (see Volume 1B—Figure 10.4), but usage, uptake and adoption take much longer to establish. The S-curve illustrates the increasing performance of a product over time, while the adoption curve shows market acceptance of the product.



Source: Hudson (2015) Adapted from Linden and Fenn (2003)

Trust is a measure of the belief that someone or something can, has and will do something. More specifically, in the context of EO, trust is a collective measure of the design choices of a product.
Hudson (2015)

Table 2.4 Factors impacting the uptake of EO products

Finding number	Description
4	Not all data has sufficient quality to become an observation, variable or product (see Section 1.1 above)
10	The distinction between awareness, usage and uptake is critical in assessing the effectiveness of an EO product.
15	Open data and recent technological developments allow product creators to increasingly change and tailor EO products.
20	Data usage and/or uptake for Australian disaster enquiries are governed by at least 15 design choices (see Table 2.5).
22	EO products with greater maturity and integration have greater usage and uptake.
26	The most limiting factors for using EO products in the emergency management sector are availability of data, timeliness of data products, and skills to use the data.
28	Official data, or data from an organisation with an official mandate, is more trusted than data from an organisation without a mandate.
30	Emergency managers prefer EO products to be delivered to them by a public institution, even if the data was collected by the private sector.
31	Licence restrictions on the EO product or the original source data can decrease the trust in a dataset.
33	Australian emergency managers believe the most likely design choices to inform decision-making are accuracy, timeliness, reliability, simplicity, consistency and reputation.
40	The balance between accuracy and timeliness is complex and must be actively managed.
43	Reputation takes a long time to build but can be lost very quickly.
44	Trust is the overarching term to encompass all design choices.

Source: Hudson (2015)

Based on the findings of this study, an EO Product Uptake (EOPU) framework was defined to improve the uptake of sensor data products based on 'design choices' that data creators could consider during the product design and revision phases. The framework draws on the concept of a 'data adoption framework', which has been used in other fields (QA4EO Secretariat, 2007; Vanden Borre *et al.*, 2011). The EOPU framework has a hierarchical structure comprising four nested levels:

- the objective of the user for this data product;
- categories of design choices—such as quality, policy, reputation and maturity;
- design choices—specific features that the users want in this data product; and
- design levers—technical metrics and tools to help creators implement and track the design choices.

The EOPU framework is intended to be tailored to the requirements of specific outcomes or target user groups in order to encourage dialogue between the creators and users of sensor data products. The hierarchical structure of this framework enables traceability for designers to see the impact(s) of their design choices. The central proposition of this framework is that every aspect of a sensor data product involves a decision made by the creator of that product. These design choices are a mix of deliberate decisions, such as presentational choices, and inadvertent choices, such as the data policy or data delivery business model of the source dataset. Data product design choices cover many fields, including both traditional technical, scientific and engineering fields, as well as non-technical areas such as law, public policy and marketing. All of these fields must be considered to fully grasp the implications of uptake for a data product. Additionally, technological developments will inevitably create new design choices over time.

However, this framework can only be applied once certain technical criteria have been met. For example, EO datasets can only become products if the measurements used have sufficient accuracy and precision to resolve the desired surface attribute(s) and if the models being used to relate physical processes to the measurements are effective (see Section 1 above). It should also be noted that uptake takes longer and costs more to achieve than usage, which means that targeting design choices to improve uptake generally increases development costs and timeframes.

Gartner's Hype Cycle model adds another dimension to technology life cycle models: it characterizes the typical progression of an emerging technology from user and media overenthusiasm through a period of disillusionment to an eventual understanding of the technology's relevance and role in a market or domain. (Linden and Fenn, 2003)

Table 2.5 EOPU framework for emergency management

Objective	Design Choice Category	Design Choice	Design Lever
Trust	Maturity	Software readiness	Portability, numerical reproducibility, coding standards, code security
		Metadata	File level, collection level and standards
		Documentation	Algorithm Theoretical Basis Document, Operational Algorithm Description, Process Flow Chart, peer reviewed documents of algorithm and product
		Product validation	Independent validation, uncertainty, quality flag, operational monitoring
		Public access	Archive, updates to record, version control
		Utility	Data usage, Societal Sector Decision Support Systems, citations in peer-reviewed literature, user feedback
	Quality	Accuracy	Accuracy, consistency, completeness, errors of commission, omission, and human error
		Relevance	Tailored business model, coverage, frequency, specificity, absence of alternative resources, responsiveness, flexibility, level of value added
		Latency	Time from acquisition to delivery, efficiency of system
		Usability	Presentation, simplicity, consistency, accessibility, ease of use, consistency of terms, fitness for purpose
	Reputation	Credibility	Integrity, objectivity, provider familiarity, availability of experts, experience, user referrals, association with trusted resources, bias/perceived incentives
		Mandate	Officially recognised provider, absence of other alternative resources
		Reliability	Percentage uptime, availability of system, availability of communications, redundancy, scalability
		Security	Physical security, network security, authentication, authorisation
		Data source	Source, defence dual-use, third party source, sovereignty, ability to disclose source, confidentiality, privacy
		Popularity	Product visibility, user awareness
	Policy	Sensor data policy	Ownership, Intellectual property, pricing policy, archiving policy, licencing describing access to raw data, rights to distribute data, responsibilities, liabilities

Adapted from Hudson (2015)

Table 2.5 lists the design choices and levers that were identified as being most relevant to emergency managers. Design choices related to quality, data policy, reputation and maturity all contribute towards the likelihood of an emergency manager incorporating a data product in decision making. This likelihood can be summarised as a measure of *trust*. Thus, if a data product aligns all design choices with the requirements of a given user it is likely to be highly trusted, and therefore more likely to be considered when making decisions. Conversely, if a data product ignores some design choices, and does not tailor design choices to a given user, that user will be less likely to trust the product and unlikely to consider that product in decisions.

The EOPU framework can be implemented across two stages of the product cycle:

- before product launch—to ensure that the product will meet user requirements within the creator's funding and technical restrictions; and
- after product launch—to continuously improve the product based on its performance and feedback from users.

While products are traditionally launched as the hype cycle begins to recover and quality on the performance S-curve climbs (see Figure 2.5), this is not always the case as various studies have shown (Hudson, 2015). Uptake of sensor data products is not simply a function of the sensor, but is driven by a range of design choices. Quite simply, to maximise uptake of an EO product, the design choices and their levers are best determined by an ongoing dialogue between the product creators and its intended users.

2.6 Further Information

CEOS:

The Earth Observation Handbook: *http://www.eohandbook.com*

Committee on Earth Observation Satellites Working Group on Calibration and Validation Land Product Validation Subgroup: *https://lpvs.gsfc.nasa.gov/*

Global systems:

NASA Ocean Colour Web: *https://oceancolor.gsfc.nasa.gov/missions/*

Global Information and Early Warning System on Food and Agriculture (GIEWS, FAO): *http://www.fao.org/giews/earthobservation/index.jsp?lang=en* *http://www.fao.org/giews/earthobservation/asis/index_2.jsp?lang=en*

2.7 References

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing*, 25(9), 1565–1596.

Dowman, I. (2004). Foreword. In *Post-Launch Calibration of Satellite Sensors*. International Society for Photogrammetry and Remote Sensing (ISPRS) Book Series Vol. 2 (Eds: Morain, S.A., and Budge, A.M.) A. A. Balkema Publishers, Leiden, The Netherlands.

Emelyanova I. V., McVicar, T. R., Van Niel, T. G., Li, L. T., and van Dijk, A. I. J. M. (2013). Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sensing of Environment*, 133, 193–209.

GCOS (2011). *Systematic Observation Requirements for Satellite-based Data Products for Climate. 2011 Update*. GCOS-154. WMO. *http://cci.esa.int/sites/default/files/gcos-154.pdf*

Hudson, D.A. (2015). *Uptake of sensor data in emergency management*. PhD thesis, ANU. *https://doi.org/10.25911/5d78d486abc4d*

Linden, A., and Fenn, J. (2003). *Understanding Gartner's Hype Cycles*. Strategic Analysis Report, Gartner. *http://www.ask-force.org/web/Discourse/Linden-HypeCycle-2003.pdf*.

QA4EO Secretariat (2007). *QA4EO. Quality Assurance Framework for Earth Observation*. *http://qa4eo.org/*.

Soto-Berelov, M., Jones, S., Farmer, E., and Woodgate, W. (2018). Review of validation standards of biophysical Earth Observation products. Ch 2 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. Retrieved from *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Stewart, J.B., Rickards, J.E., Randall, L.A., McPhee, R.K., and Paplinska, J.Z. (2014). *Ground cover monitoring for Australia: Final report July 2012 to June 2013*, ABARES Technical report 14.1, Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra, May 2014.

Vanden Borre, J., Paelinckx, D., Mücher, C., Kooistra, L., Haest, B., De Blust, G., and Schmidt, A. (2011). Integrating Remote Sensing in Natura 2000 Habitat Monitoring: Prospects on the Way Forward. *Journal for Nature Conservation*, 19 (2), 116–125. *https://doi.org/10.1016/j.jnc.2010.07.003*.

3 Data Standardisation

In addition to the three functional stages described in Section 2.4, all EO imagery needs to be calibrated before processing (see Volume 2A—Section 3). EO data providers routinely correct for radiometric (see Volume 2A—Section 8.3) and geometric distortions (see Volume 2B) in all distributed imagery. These corrections are essential for compiling and analysing time series datasets to ensure that pixels are accurately colocated and have values that are comparable.

In the context of bi-temporal and multi-temporal datasets, the following sub-sections review:

- calibrating EO imagery (see Section 3.1);
- using pre-calibrated image datasets known as analysis ready data (see Section 3.2); and
- incorporating ancillary data sources with EO datasets (see Section 3.3).

As an example of rigorous pre-processing procedures, the process flows used by Geoscience Australia (GA) to create the Land Surface Reflectance (LSR) dataset (previously called the Australian Reflectance Grid, ARG) are described in Excursus 11.1 below.

3.1 Image Calibration

Before two or more images can be compared they need to be registered in terms of all image dimensions (see Section 2.1), that is:

- spectrally—corresponding pairs of image channels from each image should measure comparable spectral ranges in terms of reflected or emitted wavelengths;
- spatially—each geometric pixel needs to represent the same ground position and area in both images;
- radiometric—image values in both images need to relate to the same numeric range for all measured reflectances/emissions. Also, pixel values generally need to be represented by the same number of bits in both images (although some image processing systems can convert on the fly); and
- temporal—image differences resulting from changes in the illumination or sensor position, or atmospheric conditions, also need to be corrected (see Volume 1B—Section 1 for more details).

3.1.1 Systematic effects

All forms of EO imagery contain systematic variations which modify the local values of image attributes and can complicate comparison of attributes from different parts of the image (see Volume 2A—Section 3). Most of these effects arise from the:

- interactions between a scanner's viewing geometry and Earth surface layer(s) (see Volume 1B—Section 3);
- physics of radiative transfer through a variable atmosphere (see Volume 1B—Sections 4 and 5); and/or
- physical process of imaging (see Volume 1A—Section 13).

Background image: ASTER L1B scenes across the Australian continent before cross-calibration to create a seamless mosaic (see Volume 2A—Excursus 8.3).
Source: Michael Caccetta, CSIRO

Examples of systematic effects include:

- vignetting in early aerial photography which cause pixels at the edges of a scanned photograph to be brighter than those in the centre.
- surface radiance recorded by an airborne scanner is modified by the atmosphere and the relative positions of the Sun and the scanner view direction to produce systematic 'limb brightening' across scan lines (see Volume 2C—Section 7.2).
- over water, near to the specular reflection point between the Sun, the water surface and a scanning sensor, there will be brightening due to 'sunglint', which depends on both the view direction and wind speed over the surface.
- in some images, around the point where the shadow of the platform for a scanner (such as an aircraft) would be located, there is a general brightening called the land surface 'hotspot' which creates very large differences across images when the Sun and sensor geometry bring it into the image field of view (see Volume 2X—Appendix 1).
- more commonly, the changing angle between the sensor, the Sun and the orientation of the land surface creates systematic topographic brightness variations, especially in satellite images (see Volume 1B—Section 3).

Such variations are best handled is through physical models, which either allow the data to be interpreted in the context of the pixel position, or 'correct' or compensate for the variations. When the purpose is simply to 'balance' the effects over the extent of an image, it is sometimes possible to use a transformation based on a reference channel to extract the systematic component (see Volume 2C—Section 7.3.1). Alternatively, image transformation such as ratioing can reduce the systematic effects. For example, both topographic shading and sunglint appear to be reduced in ratio channels (see Volume 2C—Section 10). However, since some systematic variation is additive and some multiplicative, each type of variation should be considered individually.

Systematic variation can also arise from engineering aspects of detector calibration and data collection (see Volume 1B—Section 2). For example, any imbalance between detectors in a pushbroom scanner (a scanning device uses multiple detectors to form each line of image data) will result in systematic 'lines' running vertically through an image (see Volume 1A—Section 14.2 for details). Similarly, if the sensitivities of multiple detectors in a whiskbroom scanner vary during scanning, there will be an apparent horizontal (or line) striping in the resulting image. For example, the Landsat TM/ETM+/OLI whiskbroom sensors use multiple detectors to image groups of 16 lines per scan. While the detectors in these line groups are calibrated and balanced prior to launch, variations between on-board calibration of different detectors and their interactions with different surface features can result in a residual 16 line striping effect, which typically makes analysis difficult over low contrast targets such as forests and water (see Volume 2A—Section 3.2.3 and Volume 2C—Section 2.4). Geophysical data, such as airborne magnetics data, which is interpolated to an image format from separately-flown flightlines, may also need considerable balancing between lines to avoid strong striping effects in the image (see Volume 1A—Sections 7 and 8).

The logical pre-processing steps required before bi-temporal or multi-temporal image analysis are generally considered in terms of geometric correction (see Section 3.1.1.1) and radiometric correction (see Section 3.1.1.2).

3.1.1.1 Geometric correction

As detailed in Volume 2B, the mechanism of EO image acquistion necessarily results in geometric distortion in the imagery. Many sources of distortion can be modelled using known characteristics of the sensors, platforms and their modes of operation (see Volume 2B—Section 2). Accurately associating each pixel in one image with its corresponding pixel in another image, or with precise locations on a map or on the ground, requires the development of specific registration models that associate different coordinate systems (see Volume 2B—Sections 1 and 3). Registration models are generally based on sets of ground control points from each image (see Volume 2B—Section 4).

A rectification procedure should allow model accuracy to be iteratively tested and improved before further processing (see Volume 2X—Appendices 7 and 8). Adequate coverage of accurately located control points is essential for precise registration of bi-temporal or multi-temporal images. Various algorithms are available for resampling an image to a different geometry (see Volume 2B—Section 5). This process involves estimating the expected value for some new pixel, with a known size and location, from the values of those original pixels that overlap and/or are adjacent to it. The two most commonly used resampling approaches are:

- nearest neighbour resampling—simply selects the values of the pixel in the original image which is closest to the resampled pixel location, so produces an image which only contains pixel values that exist in the original image. This method can lead to significant aliasing errors, however, so it is generally recommended that the nearest neighbour algorithm

be used to resample to a smaller pixel size than is finally required (such as half the final pixel size), then blocks of pixels in the resulting image can be combined to determine pixel values in the resampled image (see Volume 2B—Section 5.2).

- cubic convolution—assumes the original pixel values represent some underlying surface that can be modelled by a smooth, continuous function, and effectively uses this function to derive the new pixel value. Such assumptions may not be valid for EO data (especially when the image scale changes during resampling), and images produced by this method can appear to be smoothed relative to the original data (see Volume 2B—Section 5.4).

However, the image processing procedures used to correct geometric errors and create information products can also introduce image artefacts into the resampled image. For example, the term 'gridding' is used to define the process of allocating the original sensor observations to cells in a predefined image grid (Wolfe *et al*., 1998). The gridding algorithms used for MODIS products resample the acquired observations into a grid based on the Sinusoidal projection. Inevitably, to translate the original image pixels into the resampled grid cells introduces 'pixel shift', that is, each original pixel does not have the same spatial dimensions or alignment as its corresponding grid cell in the resampled image. The MODIS gridding algorithms have been shown to produce imagery in which the locations of the original imaged observations only overlap with the final grid cells by less than 30% (Tan *et al*., 2006). A metric called 'obscov' (Wolfe *et al*., 1998; Yang and Wolfe, 2001) was been developed to indicate the proportion of each observation that is derived from the area of its corresponding grid cell. Given the range of potential mismatches and misalignments between observations and grid cells, the obscov metric can vary significantly between adjacent grid cells. This metric is routinely included with the geolocation information in MODIS level 2 products and needs to be considered when MODIS pixels are being compared with field sites or other spatial datasets. Band-to-band registration of MODIS products can also be degraded by gridding algorithms, especially when comparing MODIS products acquired at different spatial resolutions (Tan *et al*., 2006).

3.1.1.2 Radiometric registration

As introduced in Volume 2A—Section 3.3.2, radiometric registration corrects for:

- known sensor errors;
- atmospheric conditions such as cloud, haze and rain;
- topographic shadowing; and
- variations caused by changing illumination and viewing positions (see Volume 2A—Section 3 and Volume 2X—Appendix 1; Paolini *et al*., 2006).

Absolute radiometric registation aims to create a set of images that appear to have been imaged by the same sensor with standardised illumination and atmospheric conditions (see Volume 2A). Such correction methods convert pixel digital values to radiance at the sensor before using atmospheric/bidirectional reflectance distribution function (BRDF) models to estimate surface reflectance(s) (see Volume 2X—Appendices 1 and 2 for model details). The presence of cloud, cloud shadow and snow in imagery is commonly detected and corrected using the Fmask algorithm (Zhu and Woodcock, 2012, 2014; Zhu *et al*., 2015).

Relative radiometric registration of bi-temporal image pairs can be achieved by matching recorded channel radiance values for common features. A simple approach cross-calibrates and rescales channel histograms using crossplots of paired channels in each image (see Volume 2A—Section 8.1.3). This effectively fits a regression line through the crossplot, which can be used to derive an equation to convert channel values from one image to equivalent values in the other (see Volume 2C—Section 7.2). All regression-based techniques effectively assume values in paired channels can be related via a linear, or affine, transformation (see Volume 2C—Section 7).

3.1.2 Local effects or 'noise'

In contrast to the low frequency systematic effects, which make it difficult to compare pixel attributes in different parts of the image, there are often very local effects (or image 'noise'), which can make it difficult to compare the attributes of a pixel with those of its neighbour. Some of these effects are due to scaling issues (see Volume 1B—Section 2.3) and represent variations introduced into the image by spatial variation at a sub-pixel level. Sub-pixel effects may be considered either as useful data or 'noise'. In image classification, sub-pixel effects are generally assumed to represent noise, although once classes are formed the sub-pixel effects within the class may be investigated for their data (see Volume 2E). Uncorrelated variance at the pixel scale can also be introduced by physical processes operating in the atmosphere and land surface system, or in the process of imaging.

Just as "one man's meat is another man's poison" however, it is worth asking the question "What is noise?" in the context of any given analysis. For example, seasonal variations in a dynamic system may be considered as noise when the purpose of an analysis is to detect longer term changes in the system, but as information when season-to-season changes are being observed (see Section 2).

Whether uncorrelated variation, or variation with a very local range of autocorrelation, should be 'removed' will depend on the purpose of the analysis. When the data are to be 'cleaned', spatial filtering, adaptive filtering and thresholded 'despiking' can be used to reduce or remove high frequency variance from an image (see Volume 2C). In image classification, if the data are not modified prior to the generation of image classes, the effect of noise is often to create variation from pixel to pixel in any derived classification resulting in heterogeneous classes or small 'noise' classes. The spatial variance can sometimes be removed (or reduced) after image classification by filtering as described in Volume 2E.

3.2 Analysis Ready Data

Analysis Ready Data (ARD) was introduced in Volume 2A—Section 7.4. CEOS (2016) defines ARD for land (CARD4L) as: "*... satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets*".

In parallel with the growing volumes of EO time series data, there is an urgent need for efficient mechanisms to access and process these datasets. The potential of these already enormous archives can only be realised by appropriate organisational structures and analytical tools, to identify spatial and temporal patterns embedded in the imagery, and extract correlations with other biophysical parameters. The first step in this process is to remove the need for traditional image-by-image calibration to correct for geometric and radiometric distortions.

The traditional approach to calibration of EO imagery is outlined in Section 3.1. ARD, however, allows users to analyse different data sources at different resolutions and in differerent projections without this laborious pre-processing. The essential elements of ARD include:

- efficient data storage structures—to organise data for easy and rapid access;
- high-performance computing—to analyse data remotely;
- grid-based data format where pixels and locations are directly comparable—data-agnostic approach to storage and analysis;
- resampling and reprojection of spatial data 'on the fly';
- pixel values calibrated to surface reflectance and directly comparable;
- remote storage of data to avoid downloading large volumes for analyses; and
- quality assessment and provenance tracking for each pixel (Lewis *et al.*, 2017).

A CEOS Virtual Constellation is a set of space and ground segment capabilities operating together in a coordinated manner, in effect a virtual system that overlaps in coverage in order to meet a combined and common set of Earth observation requirements.
(CEOS, 2013)

A natural progression of analytical tools will involve deep learning (machine learning) techniques and high-dimensional statistical analyis to manipulate EO datasets. For example, Digital Earth Australia (DEA—detailed in Section 11.2) is a 'big data' platform that integrates a range of Australian geospatial datasets, including Australian EO time series datasets. Preparation and management of these time series datasets involve many stages of data processing and verification. To guarantee the veracity of this product, sophisticated protocols are required in three key areas:

- accurate spatial alignment of pixels—to be certain that they can be 'stacked' as time series observations;
- pixel values presented as normalised surface reflectance measurements—to ensure that they are comparable across the continent and through time; and
- dependable assessment of quality for each pixel—delivered as pixel quality flags and dataset metadata, which traces the provenance of each pixel (see Volume 2A—Section 1), so that users can decide which observations are suitable for particular uses (Lewis *et al*., 2017).

These protocols are detailed for one DEA dataset, the Land Surface Reflectance[2] product suite, in Excursus 11.1.

An extension to ARD sourced from a single sensor series (such as Landsat TM/ETM+/OLI) is to integrate imagery from different sensors into a single, consistent time series. The major advantage of such integration is to create a seamless time series populated with more frequent imagery. In this context, 'seamless' implies a smooth time series, with minimal temporal noise resulting from sensor-to-sensor difference, which accurately measures surface conditions (Claverie *et al*., 2018). More frequent temporal coverage is important for many applications, especially disaster management, water quality and phenological studies (Li and Roy, 2017; see Volume 3).

The Harmonized Landsat and Sentinel-2 (HLS) surface reflectance dataset (Claverie *et al*., 2018) is the first, publicly-available, multisensor ARD product. Helder *et al*. (2018) define 'harmonised' in this context to mean "that sensor-specific radiometric and geometric differences are adjusted and removed, such that it should be transparent to end users which sensor originated any specific reflectance observation within an HLS time series". This NASA initiative is designed to create a Virtual Constellation (VC) of the surface reflectance data acquired by Landsat's Operational Land Imager (OLI) and Sentinel's Multi-Spectral Imager (MSI). Since Landsat OLI and Sentinel-2 MSI both observe the Earth from a sun-synchronous orbit at a similar time of day using similar spectral, spatial and angular dimensions, they present an ideal pair of datasets for a multisensor ARD product. A VC has been defined as '*a coordinated set of space and/or ground segment capabilities from different partners that focuses on observing a particular parameter or set of parameters of the Earth system*' (CEOS, 2019). Claverie *et al*. (2018) describe harmonised products in terms of the following characteristics:

- "gridded to a common pixel resolution, map projection and spatial extent;
- atmospherically corrected and cloud masked to surface reflectance using a common radiative transfer algorithm;
- normalised to a common nadir view geometry via BRDF estimation; and
- adjusted to represent the response from common spectral band passes".

The additional characteristic when integrating imagery acquired by different sensors, compared with a single sensor, is the need to account for any differences in the sensors' spectral resolutions, densities and extents (see processing step 4 in Table 3.1). In this case, for the spectral bands common to both sensors, the radiometry of the finer resolution MSI data is adjusted to match the bandpasses of OLI. Some of the challenges being faced to calibrate and validate this product are reviewed by Helder *et al*. (2018).

Applications of the HLS product thus far are many and varied, including monitoring aquatic systems (Pahlevan *et al*., 2019), crop type inventories and projections (Torbick *et al*., 2018), and crop and land cover mapping (Griffiths *et al*., 2019). Composite imagery derived from multisensor ARD sources is described in Section 10.2.

2 Formerly referenced as the Australian Reflectance Grid (ARG)

Table 3.1 HLS processing overview

Processing stage		Sentinel-2 processing		Landsat-8 processing
Input		Sentinel-2 MSI (L1C)	Sentinel-2 MSI (L1C)	Landsat-8 OLI (L1T)
Processing step	1. Atmospheric correction (based on Landsat-8 Surface Reflectance Code (LaSRC) approach; Vermote *et al.*, 2016) and cloud masking (LaSRC output with Fmask algorithm; Zhu and Woodcock, 2012, 2014; Zhu *et al.*, 2015)	✓	✓	✓
	2. Geometric resampling and geographic registration to Sentinel-2 UTM tiling system at 30 m spatial resolution (Helder *et al.*, 2018)	✓	✓	✓
	3. BRDF adjustment to nadir view angle and constant solar elevation (Roy *et al.*, 2016, 2017)	×	✓	✓
	4. Band pass adjustment using a spectral band adjustment factor (Chander *et al.*, 2010; Helder et al., 2018; Barsi *et al.*, 2018)	×	✓	×
Output		S10 (MSI SR 10 m)	S30 (MSI NBAR 30 m)	L30 (OLI NBAR 30 m)

Source: Claverie *et al.* (2018) Figure 1

3.3 Ancillary data

In this context, ancillary data refers to any datasets that are not derived from EO. While processing of EO imagery is primarily based on image data, it is often useful (and sometimes necessary) to integrate data from different sources with the image data prior to further processing. This is especially true for other information, such as soil type, which cannot necessarily be inferred from the EO image, or topographic information, which is unlikely to change over time. The process of integrating the different data types involves bringing the geometry and scale of the non-image data sources to the geometry and scale of the image, and potentially integrating the data as additional image channels.

3.3.1 Provenance and caveats

The inherent framework that encapsulates EO data analyses is particularly relevant when using other data sources in conjunction with EO datasets (see Section 1). Familiarity with both datasets—their provenance and goals, and their limitations—is an important prerequisite to their successful integration.

Firstly, EO datasets comprise indirect observations (see Section 1.1). With appropriate processing (see Section 2.3) these observations can be treated as measurements. This distinction is critical for selecting ancillary data to calibrate or validate EO products (see Section 3.1).

Secondly, the ancillary data needs to be relevant to the EO dataset in terms of representing similar surface target features. The extent to which ancillary data must represent similar intrinsic attributes of the target will depend the type of analysis being undertaken. For example, to assess the accuracy of an image classification exercise, the classes could be presented as a set of labels describing the expected land cover categories at specific locations (see Volume 2E), whereas to validate a spectral index derived from EO data would require more detailed biophysical information (see Volumes 2C and 3A).

Finally, the importance of scale in all four dimensions of EO datasets needs to be considered when ancillary data sources are being integrated (see Section 2.1):

- spectral—are the wavelengths observed in the EO dataset relevant to the ancillary data?
- spatial—can the ancillary data be accurately located in the EO dataset?
- radiometric—is the level of differentiation between image values appropriate for the ancillary data?
- temporal—does the timing of acquisition for the EO dataset correspond to the currency of the ancillary data?

Potential sources of ancillary data that can be integrated with EO datasets, and considerations about their suitability, are further addressed in Sections 12 and 13.

3.3.2 Data formats

Ancillary data sources may be available in either raster (grid-based) or vector (point-based) formats. While image datasets, by definiton, are stored and processed in raster format (albeit sometimes using compression algorithms for storage and processing efficency; see Volume 2A—Section 1.3), vector data is stored as coordinate strings. In both data formats, target features can be described in terms of:

- points—specific locations defined by coordinates;
- lines—a straight line between two points, or a series of straight lines interpolated between consecutive points to form a string; or
- areas—rectangular boxes or polygons (see Figure 3.1).

Algorithms to convert vector to raster format are commonly available in image processing and Geographic Information Systems (GIS; see Section 13.1).

Figure 3.1 Five basic types of patches

In image formats, the precison of data points is limited by the size of each pixel, whereas in vector formats, data points are defined as map coordinates (relative or absolute) so can be precisely represent a given ground location

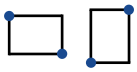
a. Point

single pixel recorded by column and row (referred to as pixel and line) position in the image

b. Line

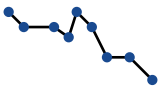
straight line interpolated between any two points

c. Box



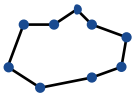
rectangle interpolated between two diagonally opposite corner points

d. String



series of straight lines interpolated between pairs of adjacent, irregularly spaced points to form a ragged line

e. Polygon



closed string which outlines a single region

Source: Harrison and Jupp (1995) Figure 45

Vector data can be used to spatially stratify an image to allow different parts of the image to be processed independently (see Section 4.1.2). Vector data from other sources, such as GIS, can be converted to a compatible format and rescaled to match the image coordinate system. Such data can then be used as patches, or referenced in categories, to stratify an image. Non-EO based spatial data, in vector or raster formats, can also be used to segment an image into multiple zones or classes. For example, patches or categories can be used to create a channel in which each pixel is given the value associated with the patch or category that it satisfies (see Volume 2E). This process allows vector data, derived independently of the image, to be represented in raster format and registered with the image. Ancillary data which is available in raster format can also be registered to the image geometry using a standard resampling process (see Section 3.1.1.1 and Volume 2B—Section 5).

A third data format, Triangular Irregular Networks (TIN), is sometimes used as an efficient method for recording elevations at defined locations. The density of surveying points can vary with relief so this approach means that the elevation at each location can be recorded precisely. To construct a Digital Elevation Model (DEM) from TIN data uses triangles between the each set of three closest points (or nodes), with the condition that the triangles do not intersect each other (see Section 13.2). Each triangle then represents a terrain surface (or facet) with uniform slope and aspect. An indexing system is required by the TIN data format to connect neighbouring points recursively (Bolstad, 2008).

3.4 Further Information

Analysis Ready Data:

Committee on Earth Observation Satellites (CEOS): *http://ceos.org/ard/*

Digital Earth Australia (DEA): *https://www.ga.gov.au/dea/products*

USGS:

US Analysis Ready Data (ARD): *https://www.usgs.gov/land-resources/nli/landsat/us-landsat-analysis-ready-data?qt-science_support_page_related_con=0#qt-science_support_page_related_con*

3.5 References

Barsi, J.A., Alhammoud, B., Czapla-Myers, J., Gascon, F., Haque, M.O., Kaewmanee, M., Leigh, L., and Markham, B.L. (2018). Sentinel-2A MSI and Landsat-8 OLI radiometric cross comparison over desert sites. *European Journal of Remote Sensing*, 51(1), 822–837. *https://doi.org/10.1080/22797254.2018.1507613*

Bolstad, P. (2008). *GIS Fundamentals: A First Text on Geographic Information Systems*. 3rd Edn. Elder Press, White Bear Lake, Minnesota. ISBN 978-0-9717647-2-9

CEOS (2016). *CEOS Analysis Ready Data* webpage, Committee on Earth Observation Satellites website: *http://ceos.org/ard/*

CEOS (2019). *Virtual Constellations* webpage, Committee on Earth Observation Satellites website: *http://ceos.org/ourwork/virtual-constellations/*

Chander, G., Mishra, N., Helder, D.L., Aaron, D., Choi, T., Angal, A., and Xiong, X. (2010). Use of EO-1 Hyperion data to calculate spectral band adjustment factors (SBAF) between the L7 ETM+ and Terra MODIS sensors. *Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium*, Honolulu, HI, USA, 25–30 July. IEEE, Piscataway, NJ, USA. pp. 1667–1670.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., and Justice, C. (2018). The Harmonised Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219, 145–161. *https://doi.org/10.1016/j.rse.2018.09.002*

Griffiths, P., Nendel, C., and Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment*, 220, 135–151. *https://doi.org/10.1016/j.rse.2018.10.031*

Harrison, B.A., and Jupp, D.L.B. (1995). *Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual*. MPA, Melbourne.

Helder, D., Markham, B., Morfitt, R., Storey, J. Barsi, J., Gascon, F., Clerc, S., LaFrance, B., Masek, J., Roy, D. (2018). Observations and recommendations for the calibration of Landsat 8 OLI and Sentinel 2 MSI for improved data interoperability. *Remote Sensing*, 10, 1340. *https://doi.org/10.3390/rs10091340*

Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J., Dekker, A., Dhu, T., Hicks, A., Ip, A., Purss, M., Richards, C., Sagar, S., Trenham, C., Wang, P., and Wang, L.-W. (2017). The Australian Geoscience Data Cube— Foundations and lessons learned. *Remote Sensing of Environment*, 202, 276–292. *https://doi.org/10.1016/j.rse.2017.03.015*

Li, J., and Roy, D.P. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, 9, 902. *https://doi.org/10.3390/rs9090902*

Paolini, L., Grings, F., Sobrino, J.A., Jiménez Muñoz, J.C., and Karszenbaum, H. (2006). Radiometric correction effects in Landsat multi-date/multi-sensor change detection studies. *International Journal of Remote Sensing*, 27(4), 684–704. *https://doi.org/10.1080/01431160500183057*

Pahlevan, N., Chittimalli, S.K., Balasubramanian, S.V., and Vellucci, V. (2019). Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems. *Remote Sensing of Environment*, 220, 19–29. *https://doi.org/10.1016/j.rse.2018.10.027*

Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., and Kovalskyy, V. (2016). A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sensing of Environment*, 176, 255–271.

Roy, D.P., Li, J., Zhang, H.K., Yan, L., Huang, H., and Li, Z. (2017). Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. *Remote Sensing of Environment*, 199, 25–38.

Tan, B., Woodcock, C.E., Hu, J., Zhang, P., Ozdogan, M., Huang, D., Yang, W., Knyazikhin, Y., and Myneni, R.B. (2006). The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions, *Remote Sensing Environment*, 105, 98–114.

Torbick, N., Huang, X., Ziniti, B., Johnson, D., Masek, J., and Reba, M. (2018). Fusion of Moderate Resolution Earth Observations for Operational Crop Type Mapping. *Remote Sensing*, 10, 1058. *https://doi.org/10.3390/rs10071058*

Vermote, E., Justice, C., Claverie, M., and Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, 185, 46–56.

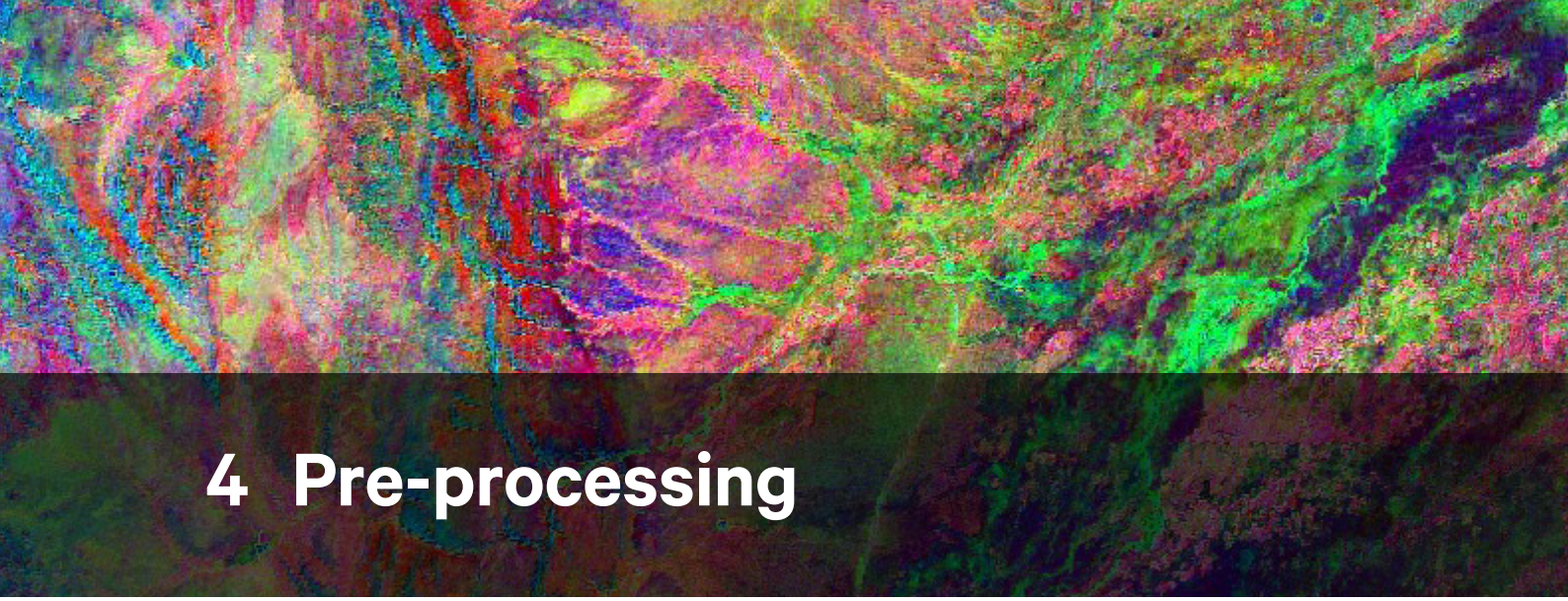
Wolfe, R.E., Roy, D.P., and Vermote, E. (1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1324–1338.

Yang, K., and Wolfe, R.E. (2001). MODIS level 2 grid with the ISIN map projection. *Proceedings of IGARSS 2001, 9–13 July, Sydney*, pp. 3291–3293. *https://doi.org/10.1109/IGARSS.2001.978332*

Zhu, Z., and Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83–94. *https://doi.org/10.1016/j.rse.2011.10.028*

Zhu, Z., and Woodcock, C.E. (2014). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152, 217–234. *https://doi.org/10.1016/j.rse.2014.06.012*

Zhu, Z., Wang, S., and Woodcock, C.E. (2015). Improvements and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8 and Sentinel 2 imags. *Remote Sensing of Environment*, 159, 269–277. *https://doi.org/10.1016/j.rse.2014.12.014*



4 Pre-processing

In addition to the image calibration processes recommended in Section 3.1, selected pre-processing of input images may be beneficial for bi-temporal and multi-temporal analyses. In this context, pre-processing refers to operations that prepare or modify an image for integration with other EO imagery and/or ancillary data. These operations can enhance certain features, develop new channels for specific tasks, reduce data volume, reduce noise, or remove features that are not of direct interest to a particular analysis.

Where image integration involves either two or more images with different spatial resolutions, or ancillary data with differing spatial resolution to the EO image dataset(s), changes to the pixel resolution in one or both images may be required in addition to geometric correction (see Section 3.1.1.1). The spatial implications of these upscaling or downscaling operations are introduced in Volume 1B—Section 2.

In this section, we will consider potential pre-processing methods in terms of:

- stratifying images to remove unwanted pixels (see Section 4.1);
- reducing data volume (see Section 4.2); and
- upscaling and downscaling (see Section 4.3).

4.1 Stratifying images

An image may contain features or regions that are not of interest to, or may disturb the statistics for, a particular study. Any unnecessary components in an image can be 'removed' before data integration to save computational time or avoid inaccurate statistics. Pixels may be 'removed' by setting their values to the null value (often reserved as the highest or lowest value in the data range) in selected channels. An alternative to embedding the null value would be to create an additional channel, usually referred to as a mask channel, in which the 'removed' pixels are set to a pre-defined (often binary) value.

An image may be stratified into regions by using spatial boundaries (defined interactively on a displayed image or from registered ancillary data) or by using spectral characteristics of different image features. Here, stratification refers to a binary, or two-state, decision: pixels either do or do not belong to some pre-defined spatial layer, whereas segmentation generally involves sub-dividing an image into more than two segments, that is pixels can belong to one of several layers (see Volume 2A—Section 10).

4.1.1 Spectral

Different image features are considered to be spectrally separable when they have distinctly different values in one or more attributes. In this case, ranges of values, or spectral categories, may be used to identify all pixels belonging to a particular feature. Training patches which represent a particular feature may be used to define these ranges (see Volumes 2A and 2E). When the statistics of one or more patches satisfactorily map the feature of interest, their spectral ranges can be represented as one or more categories (see Volume 2A—Section 9). The spectral stratification process then involves identifying those pixels whose attribute values satisfy these categories. Spectral stratification has been used to successfully remove 'unwanted' pixels from an image prior to further processing, such as water from an analysis of land features, land from a water-oriented analysis, or clouds and possibly cloud shadows from a ground-based study.

Background image: Principal Components transformation of Landsat TM image over Fowlers Gap Research Station, NSW, acquired in January 1993. PC1, PC2 and PC3 are shown as RGB. **Source:** Megan Lewis, University of Adelaide

A procedure for developing logical categories to map specific features is outlined in Volume 2E. It is important that any categories being used for stratification should cleanly include or exclude the feature(s) to be removed, since it is usually preferable to leave some irrelevant data in the image rather than remove some relevant pixels. The image attributes used to develop the stratification mask need not be the same ones used for subsequent process (provided they register precisely, in terms of image geometry), since a category can be defined using one set of attributes or channels and then used to stratify a different set.

4.1.2 Spatial

The spatial stratification process involves identifying those pixels that are located in pre-defined regions in the image. Region boundaries can be defined by image patterns, or by registering vector format data from some other spatial data set, such as a GIS.

Boundaries that are clearly visible in an image may be delineated interactively on a displayed image using the screen cursor (see Volume 2E). Examples of such boundaries are irrigation channels, which define the limits of irrigated crops in a particular image, or major roads or rivers, which may define cadastral boundaries. If necessary, edge enhancement techniques (see Volume 2C—Section 5.2) may be applied to the image prior to screen digitising.

Ancillary data may also be used to stratify the image where the boundary of a study area is not sufficiently visible to allow screen digitising (see Section 3.3). For example, cadastral or other boundaries may define the limits of a study area. Using non-image boundaries may involve registering and integrating the ancillary data (usually in vector format) with the image before stratification (see Volume 2A—Section 10).

4.2 Reducing data volume

Some pre-processing operations have the advantage of representing essential image data in a smaller data volume. For example, spectral indices can condense relevant image data into fewer channels to simplify subsequent image processing tasks (see Section 4.2.1). Selected linear pre-processing operations (such as Principal Component Analysis, PCA; see Section 7.2 and Volume 2A) can re-orient the image data space to both distinguish specific features and reduce data volume (see Section 4.2.2).

While reducing data volume is often useful for simplifying image processing operations, when appropriately implemented it can introduce valuable focus into the overall image analysis methodology. Methods for 'feature selection', to highlight the most informative image bands for classification, are introduced in Volume 2A—Section 9.1.3.1. Such methods and their rationale are further detailed in the context of understanding data dimensionality in Volume 2E.

Where appropriate, image data volume can also be reduced by changing the pixel 'bit depth' (see Volume 2A—Section 1.2). For example, when is no advantage in storing data with the precision of floating point, integer format would result in a smaller file size.

4.2.1 Spectral indices

Ratios of selected channels are useful for enhancing particular features, such as those used for mineralogy, or removing atmospheric or topographic effects. Various transformations may be suitable for this purpose as detailed in Volumes 2C and 3A, although vegetation indices are the most commonly used (Thonfeld *et al.*, 2010).

Spectral indices may also simplify change detection exercises (see Section 7). When appropriate to the change being observed, such indices both highlight selected features and reduce data volume. They also allow images acquired by different sensors to be compared more easily when slightly different spectral wavelengths are measured by corresponding image channels. For example, it is more reliable to compare NDVI (Normalised Difference Vegetation Index) channels computed from imagery acquired by different sensors (such as MODIS and AVHRR) than to compare image values in their red and NIR channels. For hyperspectral images, a very wide range of indices has been developed (Roberts *et al.*, 2011; Sonobe and Wang, 2017; Thenkabail *et al.*, 2018). In these data-rich images, selection of the optimal indices, ratios or channels simplifies subsequent image processing steps.

4.2.2 Linear Transformations

Linear transformations allow rotation, rescaling, skewing and/or reflection of the image data space (see Volumes 2C and 2X). The original channel axes in an image can be rotated to align with some other direction in the data. Principal Component Analysis (PCA) is a special type of affine transformation that uses the image variance to define a rotation in which the first principal component (PC1) axis is aligned with the direction of greatest variance in the *n*-dimensional data space (where *n* is the number of channels in the image) and each higher numbered axis is orthogonal to (or uncorrelated with) the previous one (see Figure 4.1). The uncorrelated nature of the PC channels is often as (or more) important than their alignment. The PCA transformation is typically computed from the image covariance matrix, but the correlation matrix can also be used to derive 'standardised PCs' (Singh and Harrison, 1985; see Volume 2C for details).

Since the most 'significant' information contained in the image is summarised in the highest PCs, it is sometimes advantageous, in terms of reducing data volume, to compute a PCA transformation prior to integrating an image and use only the highest PCs for subsequent processing.

PCA transformations can be especially useful for comparing images of different dates as discussed in Section 7.2. However, since PCA is a scene-dependent transformation, significant misregistration errors between a pair of images would increase the variation between channel pairs from different images and consequently inflate the covariance statistics. This bias may result in misregistered features being isolated as an individual PC and could mask actual changes. Therefore image pairs should be carefully registered geometrically before applying PCA (see Section 3.1.1.1).

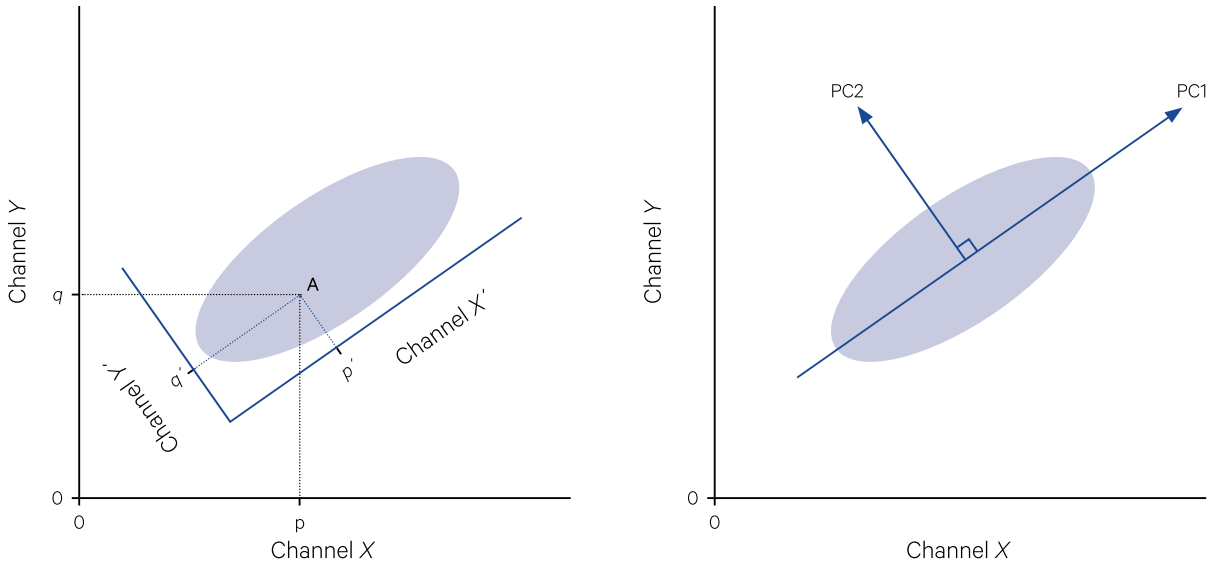
A PCA transformation can also be computed for specific features in the image. A training region encompassing the feature, or a logical category which spectrally describes it, can be used to derive a covariance matrix for a PCA. For example, if a study is particularly interested in forests, the covariance matrix from a forested area could be used to produce a PCA transformation which would optimally separate the variation contained in the forest vegetation type (see Volume 2C).

The value of using transformations which reduce the 'dimensionality' of an image (that is, the number of independent attributes used to describe a pixel) is further discussed in Volume 2E. Methods for optimum feature selection are also described in Volume 2E.

Figure 4.1 Principal Component Analysis

a. Image data values can be redefined relative to modified position and scaling of axes.

b. The PCA transformation defines a new primary data axis (PC1) in the direction of maximum image variance (over all image channels), then other axes are defined to be orthogonal to this direction.



Source: Harrison and Jupp (1990) Figures 87 and 88

4.3 Upscaling and Downscaling

As introduced Volume 1B—Section 2.4, and reviewed in Section 1.4 above, the terms upscaling and downscaling derive from cartographical scale, to indicate the processes of changing from:

- a finer scale to a coarser scale—upscaling (such as from 1:10,000 to 1:100,000); and
- a coarser scale to a finer scale—downscaling (such as from 1:100,000 to 1:10,000).

When a higher resolution dataset (from EO or ancillary data source) is being integrated with a lower resolution dataset, there are essentially three options to reconcile their scales:

- upscale the higher resolution dataset to match the spatial scale (pixel size) of the lower resolution dataset (see Section 4.3.1);
- downscale the lower resolution dataset to match the spatial scale (pixel size) of the higher resolution dataset (see Section 4.3.2); or
- convert the spatial scale of both datasets to a common pixel size (see Section 4.3.3).

As introduced in Volume 2C, filtering techniques allow the values of each pixel to be modified by the values of its neighbouring pixels. These techniques are often used to remove spatial noise and undesirable spatial variation in an image, or enhance regions of contrast such as edges. Image smoothing produces a more spatially coherent image at the expense of 'blurring' spatially detailed image features. In this context, depending on the 'natural' scale of ground target features (see Section 1), image smoothing may be appropriate in conjunction with image upscaling or downscaling.

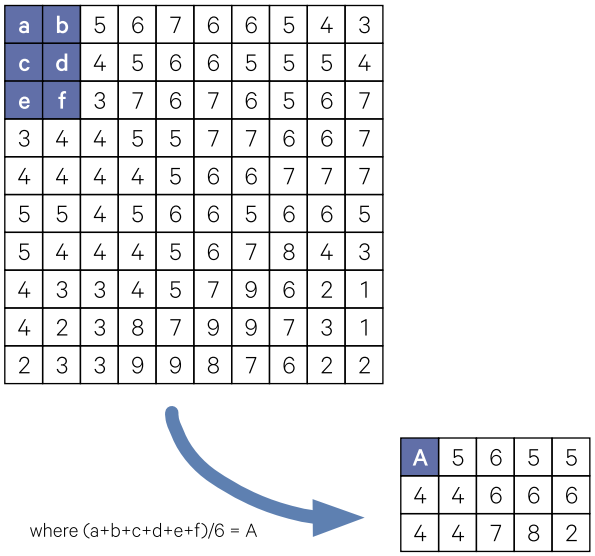
4.3.1 Upscaling to a larger pixel size

When applied to EO data, upscaling creates larger pixels, that is, ones that cover a larger ground area. This operation reduces the sample density in the higher resolution image to match the sample density in the lower resolution image. Computationally, upscaling typically involves:

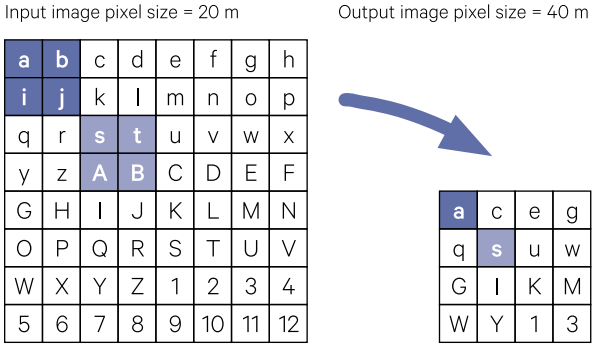
- blocking or aggregation of the pixels in the higher resolution image to effectively create larger ones with a larger pixel size that matches those in the lower resolution image (see Figure 4.2a). This process may introduce blurring (see Volume 2A—Section 7.2);
- sub-sampling of pixels in the higher resolution so that the selected pixels in the input image are assumed to represent a larger ground area, and effectively matching the pixel size in the lower resolution image (see Figure 4.2b). This process may introduce aliasing (see Volume 1B—Section 2.2); or
- resampling pixels in the higher resolution image to match the pixel size in the lower resolution image (see Volume 2B—Section 5).

Figure 4.2 Upscaling operations

a. Blocking: a higher resolution image channel image is blocked using a 2x3 blocking factor to produce lower resolution image.



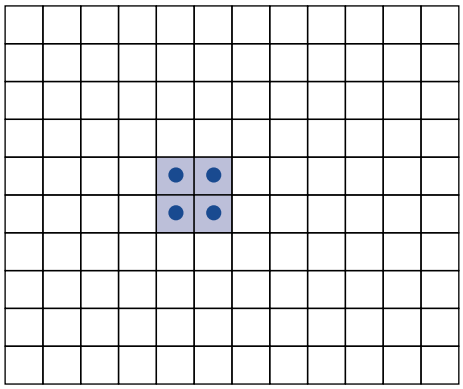
b. Sub-sampling: a higher resolution image is sub-sampled to simulate a lower resolution image by dropping alternate pixels and lines.



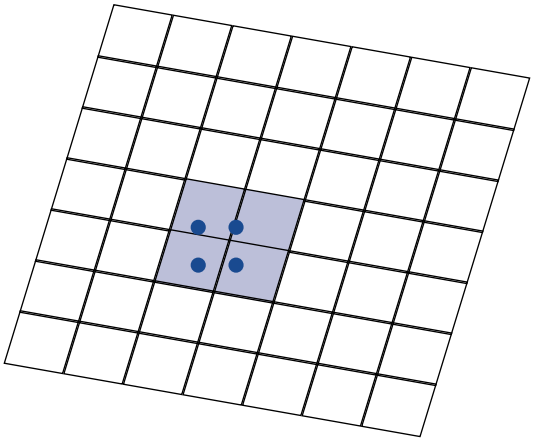
Source: Harrison and Jupp (1990) Figures 18 and 19

Figure 4.3 Upscaling by picking and binning approach

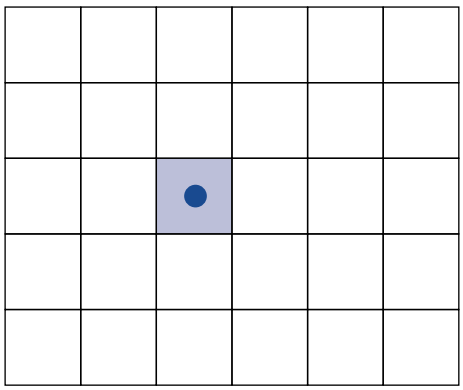
a. The underlying pixel grid of the lower resolution image is sub-divided to represent each pixel as four locations.



b. Higher resolution image grid showing location of the sub-divided pixel in the lower resolution image.



c. The pixel values in the higher resolution image at the four locations in Figure 4.3b are averaged to compute the upscaled pixel value.



Source: Harrison and Jupp (1992) Figure 43

As illustrated in Figure 4.2, the blocking operation generally averages the relevant input image values to create a single output pixel in the upscaled image, whereas sub-sampling selects one of the relevant input image values as the output value (see Volume 2A—Section 7.2.1.3). Both of these operations compromise the spatial detail represented in the higher resolution image.

Image resampling algorithms are detailed in Volume 2B—Section 5. These methods interpolate the pixel values in the higher resolution image to simulate a pixel size that matches the lower resolution image. For example, the picking and binning approach to resampling (see Figure 4.3) provides better estimates of the upscaled pixel values than blocking alone. As detailed in Volume 2B, it also provides a better estimate for representing vector data in raster format (see Figure 4.4).

Figure 4.4 Vector to raster conversion

When vector data is represented in raster format, a finer pixel mesh renders better estimates of the original vector boundaries than a coarser mesh. Subpixels then allow the 'correct' pixel aggregate (or area average) solution.

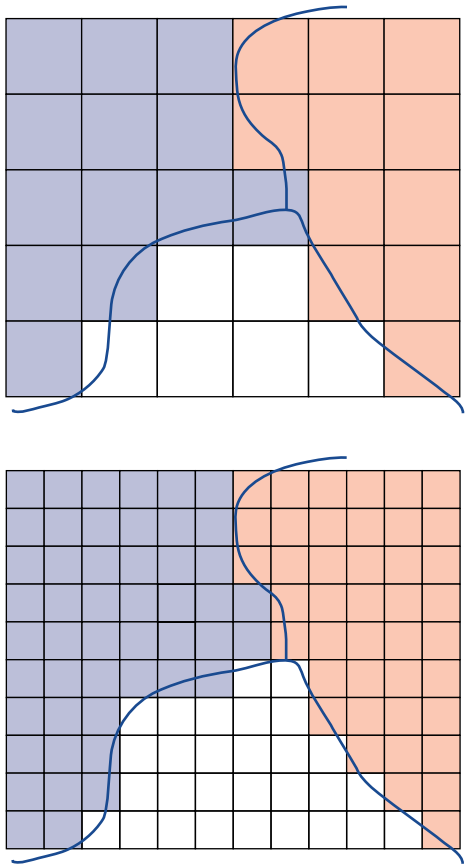
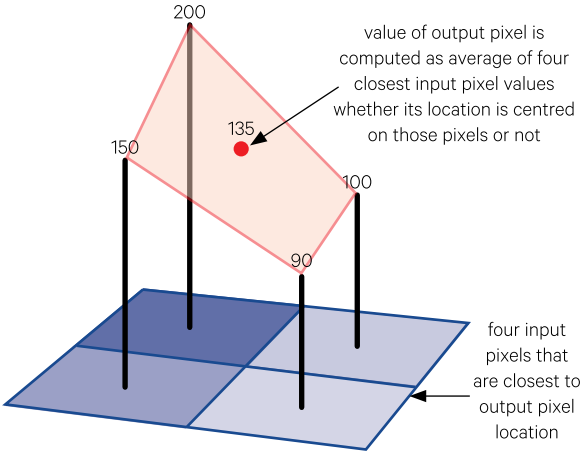
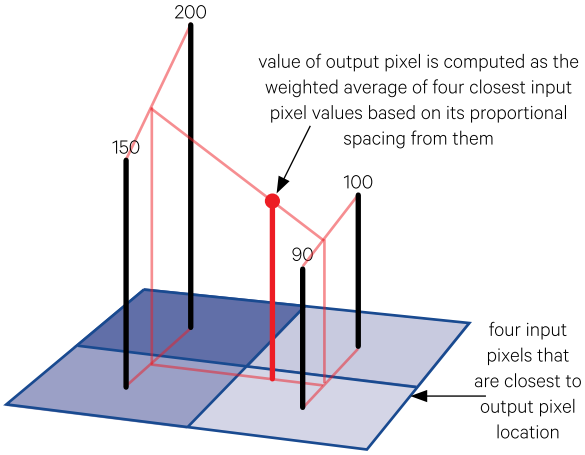


Figure 4.5 Bilinear interpolation

a. Example of unweighted interpolation, where the value of the output pixel is computed as the average value of the four closest input pixels: 135 = (200+150+90+100)/4



b. To determine the interpolation weights to use for each output pixel, bilinear interpolation computes the proportional spacing from the four input pixels that are its closest neighbours.

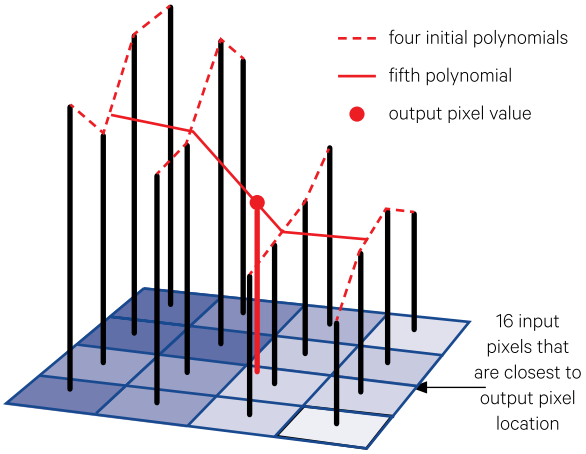


Bilinear and cubic convolution resampling can also be used to interpolate values from the spatial scale of the higher resolution image to the lower resolution image (see Figure 4.5 and Figure 4.6). Cubic convolution is the favoured resampling method for EO datasets (see Volume 2B—Section 5.4).

Image resampling attempts to consider the underlying spatial distribution of pixel values, so generally results in more continuous image values in the output image than blocking or sub-sampling. When the spatial detail of the original image values is retained (as occurs with selected compression and reconstruction algorithms), the upscaled image could be subsequently downscaled perfectly.

Figure 4.6 Cubic convolution interpolation

The method is commonly implemented by fitting four individual cubic polynomials to the four lines of pixels surrounding the output pixel. A fifth polynomial is then fitted across these four these four functions and through the output pixel location.



Scaling literally means to reduce or increase in size.
(Gallant et al., 2008)

4.3.2 Downscaling to a smaller pixel size

Since downscaling results in smaller pixels that cover a smaller ground area, the sample density of the image increases. This process can be likened to the display function of zooming in (see Volume 2A). Downscaling can be implemented by:

- enlarging the input image pixels, so that each input pixel becomes multiple pixels in the output image (see Figure 4.7); or
- interpolating the pixel values to create pseudo measurements positioned between the original ones using resampling methods (see Volume 2B—Section 5).

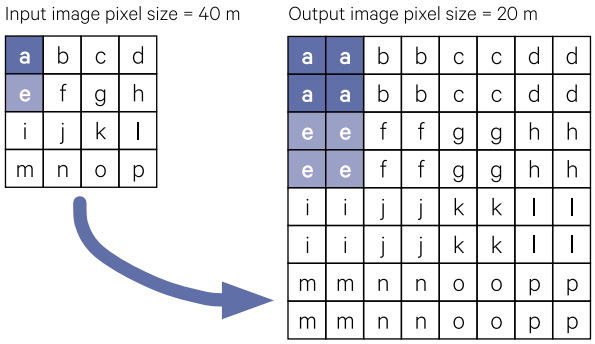
The process to enlarge input image pixels using pixel duplication is described in Volume 2A—Section 7.2.1.4. This process clearly misrepresents the spatial transitions between pixel values in the downscaled image, but offers an expedient method for representing the lower resolution image at the scale of the higher resolution image.

The resampling methods (described in Section 4.3.1 above and detailed in Volume 2B—Section 5) can also be used to downscale image data. However, unless care is taken with any of these methods, interpolation of EO data is ambiguous and can produce misleading information. The issues related to spatial resolution (see Section 1.2 above) and the underlying scale of the surface mosaic (see Section 1.4) are most relevant here.

Figure 4.7 Downscaling by duplicating pixels

Enlarging duplicate pixels in a lower resolution image to simulate a smaller pixel size.

a. Duplicate each pixel to represent a smaller pixel size



4.3.3 Resampling to a common pixel size

As introduced in Section 1.4, the surface of Earth, which is imaged in any EO dataset, can be considered as a mosaic of tiles. The tiles in this mosaic represent different surface features with varying dimensions and colours. Any single EO image, however, only captures these features and patterns at a single spatial scale. Given that the constituent tiles of the surface mosaic also vary in size spatially and temporally, it is not a simple matter to select the 'best' spatial resolution to suit all features and patterns being imaged. Rather, the resolution that is most appropriate for a particular application needs to be determined (see Section 2.2).

When resampling two (or more) EO images to a 'common grid' as defined by image pixel size and extent (spatial resolution and extent), both images generally need to be resampled (see Volume 2B—Section 5). Resampling will ensure that the original image data is more faithfully represented in the common grid compared with the simpler image rescaling methods of blocking, sub-sampling or duplicating pixels, as mentioned in Sections 4.3.1 and 4.3.2 above.

The most appropriate pixel size for this grid, however, will vary for different applications, and possibly at different acquisition times (see Section 2.2). There is no easy answer to this problem; it requires familarity with both the surface features and patterns of interest to a given study and the characteristics of available imagery. In many cases, the final selection of imagery may be governed by pragmatic factors, which will need to be addressed during analyses and reporting.

4.4 Further Information

Gonzalez and Woods (2018)

Jensen (2016) Chapter 8

4.5 References

Gallant, J.C., McKenzie, N.J., and McBratney, A.B. (2008). Scale. Ch 3 in *Guildelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

Gonzalez, R.C., and Woods, R.E. (2018). *Digital Image Processing*. 4th Edn, Pearson Education, Hoboken, NJ.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th Edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Roberts, D.A., Roth, K.L., and Perroy, R.L. (2011). Hyperspectral Vegetation Indices. Ch 14 in *Hyperspectral Remote Sensing of Vegetation*. CRC Press. pp 309–328.

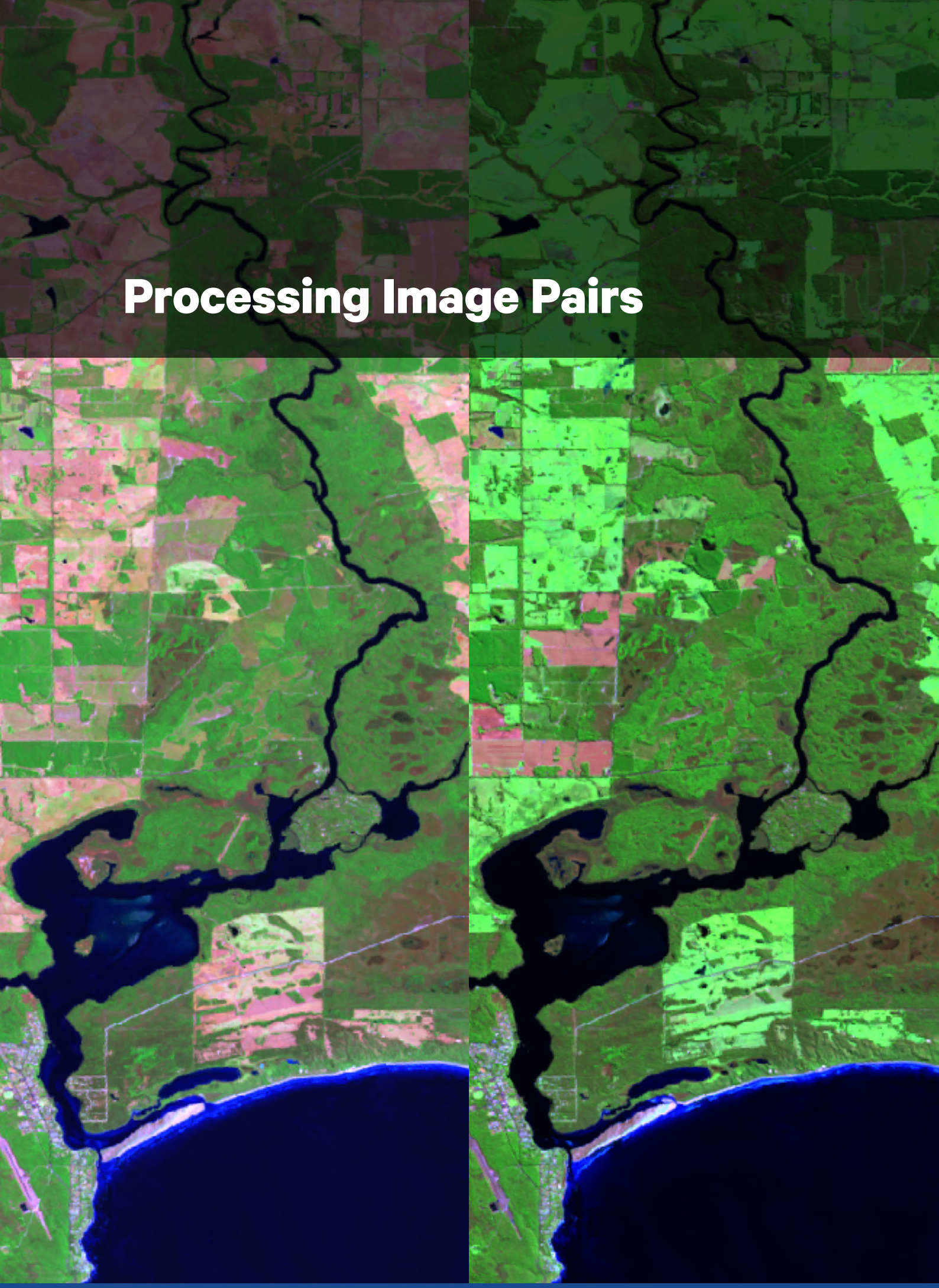
Singh, A., and Harrison, A. (1985). Standardised Principal Components. *International Journal of Remote Sensing*, 6, 883–896.

Sonobe, R., and Wang, Q. (2017). Hyperspectral indices for quantifying leaf chlorophyll concentrations performed differently with different leaf types in deciduous forests. *Ecological Informatics*, 37, 1–9. *https://doi.org/10.1016/j.ecoinf.2016.11.007*

Thenkabail, P.S., Lyon, J.G., and Huete, A. (2018). *Hyperspectral Indices and Image Classifications for Agriculture and Vegetation*. CRC Press. ISBN 9781138066038

Thonfeld, F., Hecheltjen, A., Braun, M., and Menz, G. (2010). From Algorithms to Processing Chains: A Review of Land Cover and Land Use Change Detection Methodologies. *Proceedings ESA Living Planet Symposium*. ESA SP-686. ESA, Bergen, Norway.

Processing Image Pairs



The following sections focus on image processing operations that can be used to analyse a pair of bi-temporal images. The processing procedures commonly used for comparing or combining image pairs are introduced in Section 5. Methods that merge and/or interpolate data from two different EO sources are considered in Section 6. Approaches to detecting changes between a pair of bi-temporal images are reviewed and evaluated in Section 7.

Contents

5 Basic Comparisons 47

6 Fusion Methods 57

7 Change Detection Methods 69

Background image on previous page: Landsat-8 images acquired over Augusta in summer (left) on 13 January 2015 and winter (right) on 25 August 2015. Both images are displayed as false colour composites, with band 6 as red, band 5 as green and band 3 as blue. **Source:** Norman Mueller, Geoscience Australia

5 Basic Comparisons

A simple method for comparing three EO images of the same ground area, which were acquired on three different dates, involves creating multi-date colour composites (see Section 5.1). The remainder of this section reviews some simple image processes for bi-temporal image pairs, namely:

- subtracting and adding images (see Section 5.2);
- manipulating image masks to separate, combine or complement images (see Section 5.3);
- transferring class statistics to another set of image channels (see Section 5.4); and
- balancing image colour for image mosaicking (see Section 5.5).

Other operations that can be used to merge multiple images include fusion techniques (see Section 6), affine transformations (see Volume 2C—Section 7); Principal Components Analysis (PCA; see Volume 2C—Section 9) and channel ratios (see Section 2C—Section 10). Specific processes for detecting changes in a pair of EO images are discussed in Section 7.

All operations described below assume that the bi-temporal pair of images has been registered to a common geometric grid (see Section 3.1.1.1). In particular, if the image pair was derived from different image sources, such as two different sensors or bi-temporal images from one sensor, the images must be well registered and radiometrically calibrated before they can be sensibly processed as a pair. This is also essential for the processes described in Section 7. The procedure for image registration is outlined in Volume 2B.

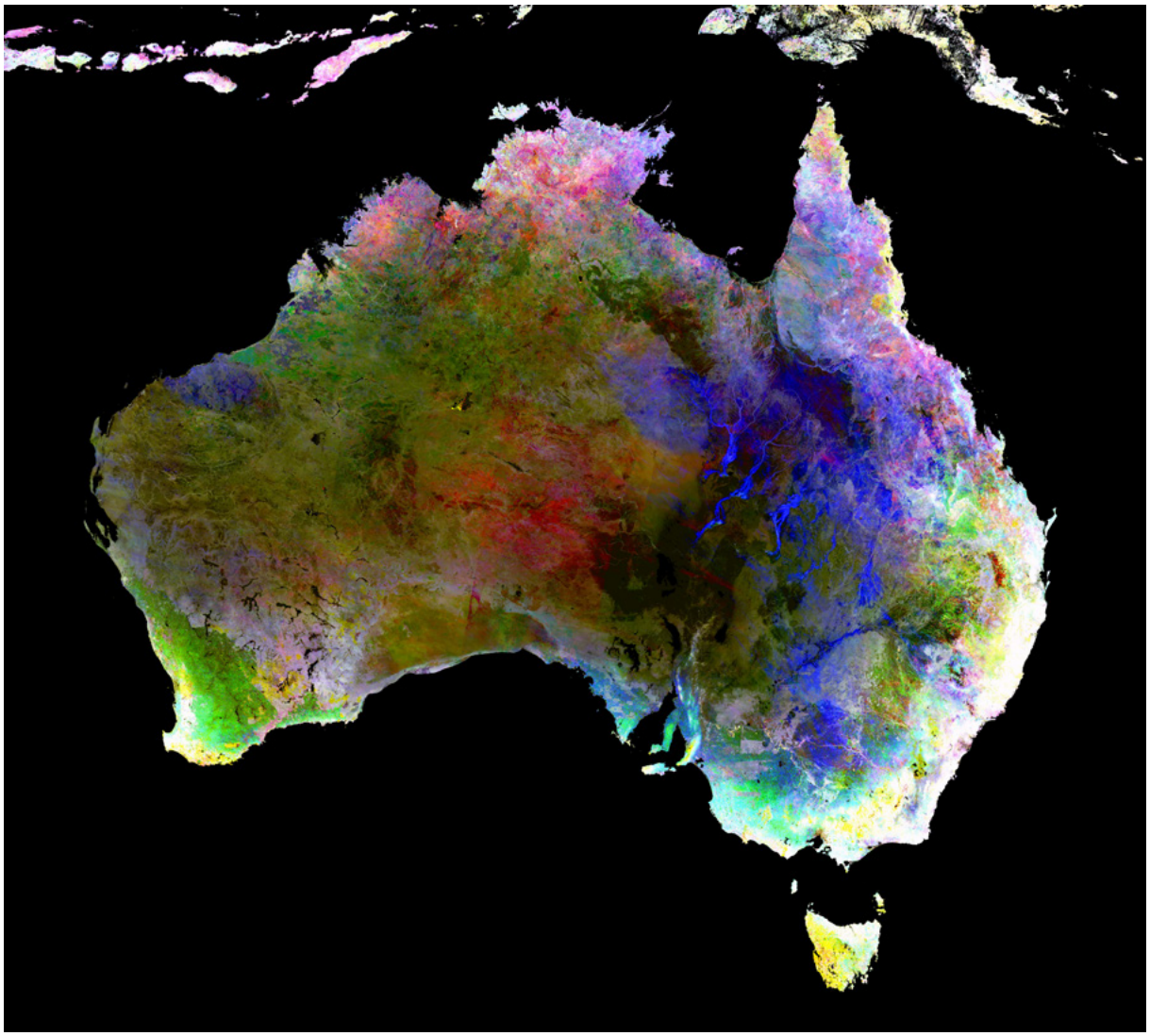
5.1 Multi-date Colour Composites

A simple way to compare registered, multi-date imagery of a given location is to display three channels representing three image dates as different primary colours in a colour composite. For example, in Figure 5.1 NDVI was derived for three contrasting seasons in 2018/19. When these greyscale images are combined into a colour composite, their combined colours highlight areas with consistently high or low vegetation cover. Similar composites could be generated using a wide range of spectral indices and other image transformations that summarise multispectral imagery into a single channel.

Background image: High resolution aerial image of Maroubra Beach, Sydney, acquired on 24 February 2016. **Source:** © Nearmap

Figure 5.1 Multi-date NDVI composite

In this colour composite, three NDVI composites derived from Sentinel-3 imagery are combined into a single colour image: December 2018 is shown as red, October 2018 is shown as green and June 2019 is shown as blue. These composite colours generate dark areas where vegetation greenness/cover was low for all three months and bright areas for 'greener' perennial vegetation. Interesting patterns are generated based on the seasonality of vegetation: vegetation resulting from early winter rainfall is highlighted as blue, monsoonal rains in summer invigorate vegetation in northern and central Australia (pink and red areas) whereas spring rains drive vegetation growth in temperate regions (brighter green patches). The less vegetated regions of Central and Western Australia show as darker colours. The NDVI channels in this composite have been enhanced using step-wise linear stretching to highlight lower values.



Source: Norman Mueller, Geoscience Australia

5.2 Differencing and Addition

Image differencing provides a simple assessment of the changes between two registered images. As illustrated in Figure 5.2, the process simply computes the difference between a co-registered pair of images on a pixel-by-pixel basis for each corresponding pair of channels. Naturally, for pixel values to be comparable between the two images, radiometric registration must be applied to the geometrically registered images before differencing (see Section 3.1). Also, for a difference image to consistently show the differences between an image pair, matching channels in the two images should have comparable data ranges. In the cases of bi-temporal or multi-source imagery, the channels should be rescaled to have matching data ranges before differencing. Methods for image colour 'balancing' are discussed in Section 5.5.

The differencing operation computes:

Channel i difference = Channel i in image A – Channel i in image B

The values in the difference channel, will be:

- < 0 when *channel i* values in *image A* are higher than in *image B*;
- = 0 when *channel i* values are the same in both images; and
- > 0 when *channel i* values in *image B* are higher than in *image A*.

Some image processing systems allow the difference result to be:

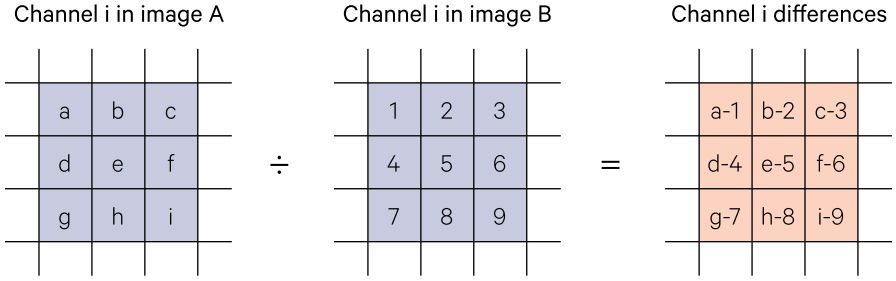
- compressed (divided by two);
- shifted (add an offset) to keep the output values within the image data range; or
- stretched within a defined output range.

An offset constant is usually added to the difference result to keep the output image values in the same range as the input image. For example, in a byte image (that is, with a data range from 0 to 255) a constant of 127 could be used to centre the range of difference values. If large differences (greater than half the data range) are expected between the images, the resulting difference values will also need to be rescaled (see Volume 2A). Unchanged pixels then have a value equal or close to the offset in the difference channel, and would appear as mid-grey when displayed as a greyscale.

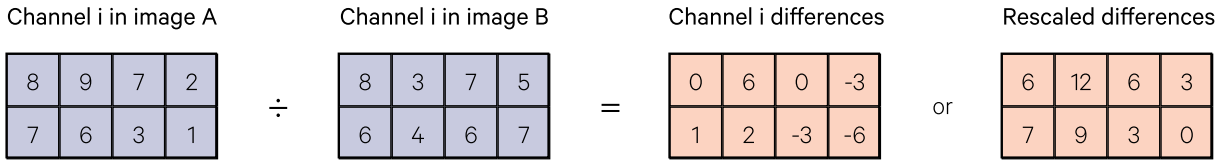
An example of two NIR channels that have been differenced from two registered images is shown in Figure 7.2 in Section 7 below. As illustrated, changed pixels fall into the tails of the histogram of the difference image, so that pixels with higher values in *Image A* appear brighter in the difference channel, while those with higher values in *Image B* appear darker. A similar distribution results from image ratioing (see Section 7.1.2).

Figure 5.2 Image differencing

a. Process: two channels are differenced by subtracting the value of a pixel in one channel from the corresponding pixel value in the other channel. When differencing bi-temporal image pairs the two channels typically represent the same spectral data (such as NIR).



b. Numeric example. The resulting difference values are rescaled to the image data range. In this example, the computed value -6 is rescaled to 0 and the computed value 6 is rescaled to 12.



Source: Harrison and Jupp (1990) Figure 60

This differencing operation can also be applied to all corresponding channels in a bi-temporal pair of images to create a colour composite difference image, although the colours in composite difference images can be harder to interpret (see Figure 5.3). They are, however, useful to assess particular image processing operations, such as image classification. For example, a residual image can be computed to highlight those pixels whose original image values are not close to the mean values of their allocated class (see Figure 5.4). In this case, the residual image indicates where additional classes may be required (see Volume 2E; Jupp and Mayo, 1982).

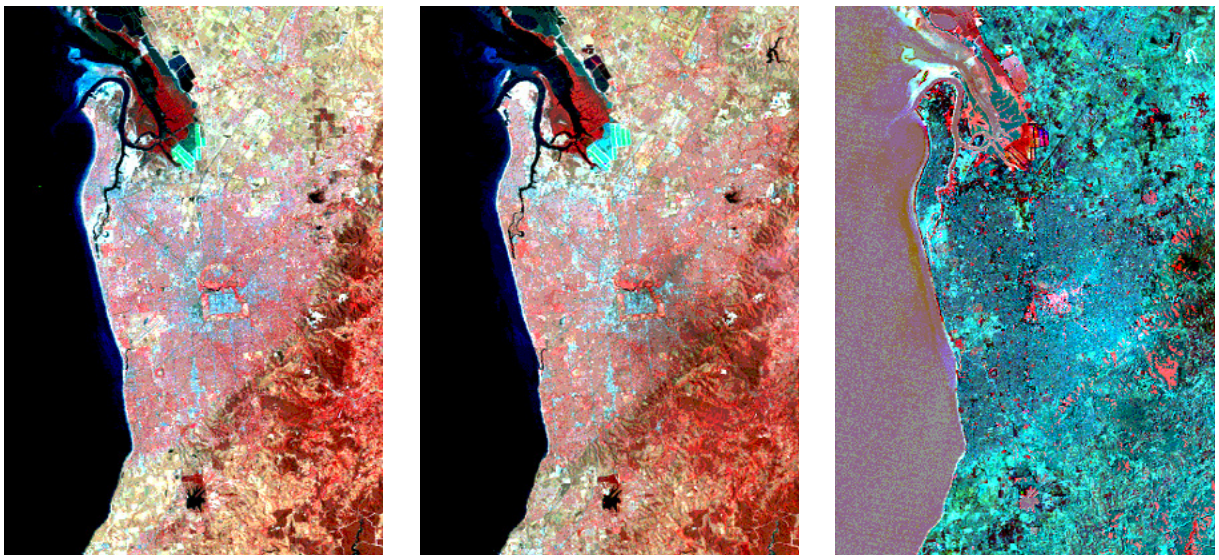
Channels can be processed before differencing, possibly to remove noisy pixels or to represent a land cover characteristic (such as vegetation greenness; see Volume 2C). When differencing pairs of transformed channels, rescaling may be required before differencing to ensure that channel values are comparable. Transformations that are scene-dependent, such as PCA (Principal Components Analysis) or MNF (Minimum Noise Fraction), are generally not recommended as a pre-processing operation before differencing images (see Volume 2E) although they are useful as tools to explore difference images. Differencing of ratioed or index channels offers the advantage of reducing any topographic, Sun angle and/or atmospheric differences between images. Ratios of two bands or indices can be more effective than differencing where terrain is significant as ratioing will subdue the effects of shadowing more than differencing (see Section 7.1.2 and Volume 2C—Section 10).

Multidimensional distance metrics that can be used to compute the 'direct' difference between multi-channel image values (such as the Euclidean distance) are discussed in Volume 2E. As illustrated in Figure 5.5, such metrics can be useful for comparing different image processes, such as filtering options or resampling methods, to show where the processed pixel values vary.

Figure 5.3 Colour composite difference image

This pair of Landsat MSS images of Adelaide have been differenced to show the impact of urban expansion between 1972 and 1989. In the difference image green indicates an increase in greenness and red indicates a decrease in greenness, white indicates an increase in cover and black indicates a decrease in cover.

a. 1972 b. 1989 c. Difference 1972-1989



Source: Megan Lewis, University of Adelaide

Image addition effectively 'strengthens' any pixels that have positive spectral correlation between two images. This feature can be useful for crop identification, for example, where an individual image may not clearly differentiate between different ages or types of agricultural crop on the basis of spectral values, but the spectral correlation through a time series of images is able to separate such differences. The addition operation could also be weighted to compute the weighted average of the two images rather than each image contributing equally to the result. The weighting option could be useful for adding an additional channel after two images have been added together. For example, to produce an image that is equivalent to an unweighted average of three channels *A*, *B* and *C*, we could add pairs of images in the sequence:

$$temp = \frac{A}{2} + \frac{B}{2}$$

$$A + B + C = \frac{2 \times temp}{3} + \frac{C}{2}$$

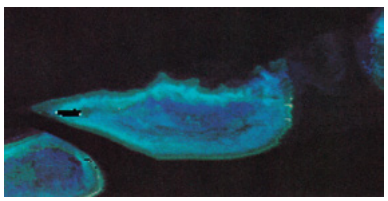
Both addition and differencing of image pairs require that the two images be precisely registered, that is, the registration model is able to accurately relate each pixel in the first image to its corresponding position in the second image (see Volume 2B). When the image pair is formed from an original image and a spectrally transformed version of it (as in Figure 5.4 or Figure 5.5a), the geometries will already be exactly registered. Figure 5.5b highlights the variations in geometric registration that can result from different resampling methods (see Volume 2B for details).

If locational errors exist when a resampled image is differenced with its bi-temporal paired image, artificial edges can appear in the difference image. For example, Townshend *et al.* (1992) used simulated MODIS imagery to measure the effects of geometric misregistration in NDVI difference images. They concluded that misregistration by one pixel can result in up to 50% increase in the semi-variance of the difference image for densely vegetated images and a 10% increase for sparse vegetation. (Semi-variance is a measure of the scale of spatial pattern and is related to the variogram; see Volume 2A—Section 8.2.) Thus, in densely vegetated images, registration errors would need to be less than 20% of a pixel to ensure that the ensuing misregistration effects were less than 10%.

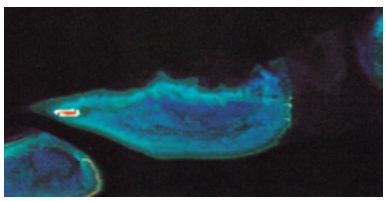
Figure 5.4 Residual image

This Landsat MSS image of Heron Island Reef, Queensland, was classified in terms of reef cover classes (Jupp *et al.*, 1985a; Jupp *et al.*, 1985b). The difference between the pixel values of the original image and the mean image (in which pixels are given the values of their class seeds) is shown as the residual image.

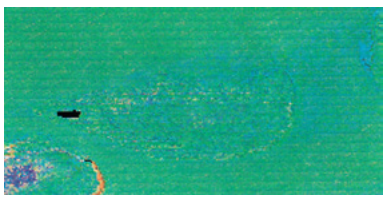
a. Original image



b. Mean image



c. Residual image



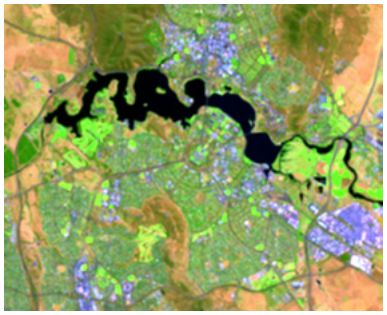
Source: Jupp *et al.* (1985a)

Figure 5.5 Examples of Euclidean distance images

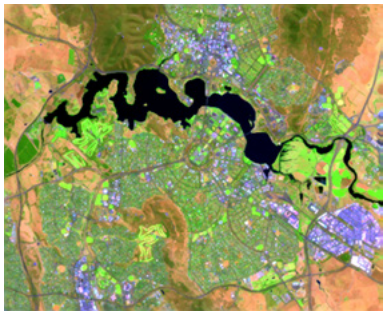
The Euclidean difference between corresponding pixels in these image pairs highlights where the transformed image values differ in the two input images in each pair.

a. A Landsat-8 OLI image of Canberra, acquired on 10 August 2017, was processed using two different filtering operations with an active filter region of 3x3 pixels.

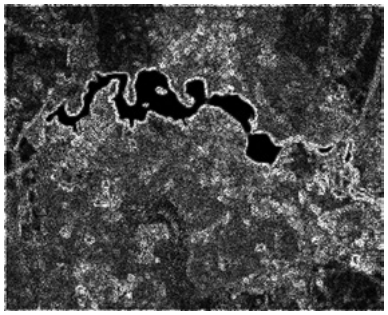
Full average filter



Edge-preserving filter



Euclidean difference

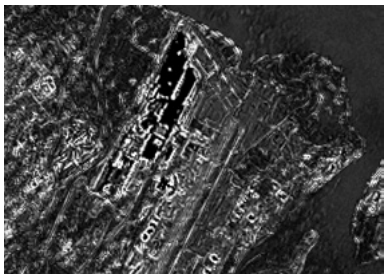


b. This pair of input images shows two resampling methods for a Landsat-8 OLI image of Brisbane airport, Queensland, acquired on 13 June 2018 and displayed using bands 6, 5, 3.

Bilinear resampling

Cubic Convolution resampling

Euclidean difference



Source: Norman Mueller, Geoscience Australia

5.3 Manipulating Segmentation Masks

An ancillary operation to image segmentation (see Volume 2A—Section 10) involves transferring masked regions (defined by a specified value or the image null value) between registered images. This process can take a number of forms:

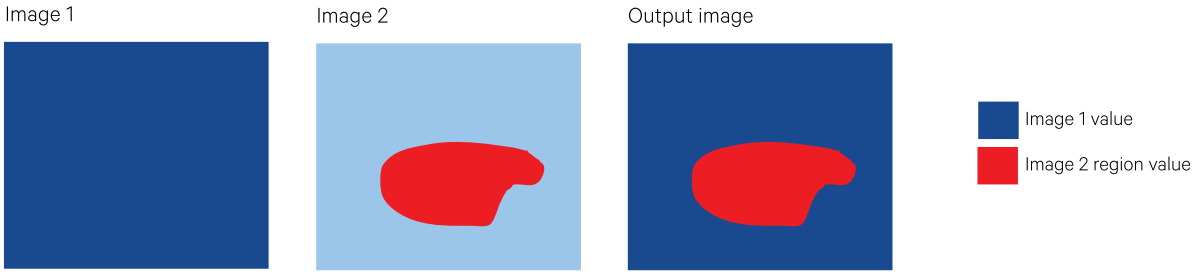
- paste a mask defined in one image into another image (see Figure 5.6a and Figure 5.6b);
- combine masks from two images (see Figure 5.6c);
- complement the masked and non-masked areas in an image (see Figure 5.6d); or
- recombine imagery to replace masked regions with data (see Figure 5.6e).

The complementing operation uses a mask to define the areas that will have non-null values in the output image but also sets all other values to the null value. The recombining function allows regions that have been masked out in one image can be given values from another image, possibly with any previously unmasked pixels becoming null (or masked) values. These operations may be useful during the classification process to extract unclassified pixels for further analyses (see Volume 2A—Section 9 and Volume 2E).

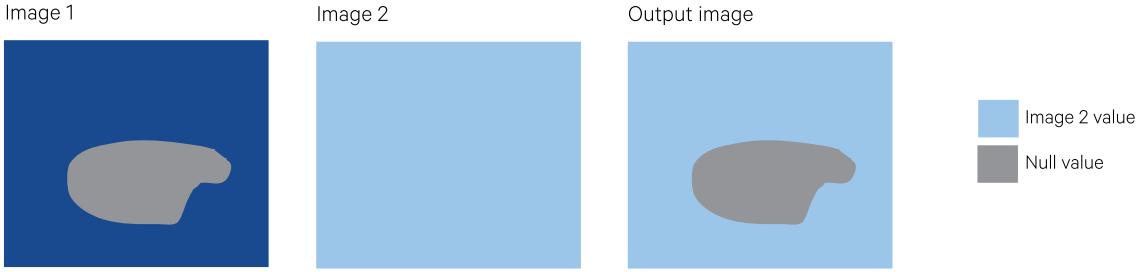
Figure 5.6 Recombining images

Masks can be transferred and combined between pairs of images in various ways.

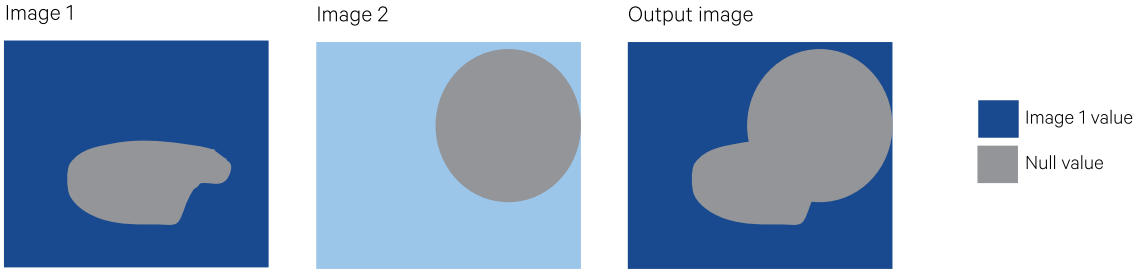
a. Paste a defined region from image 2 into image 1



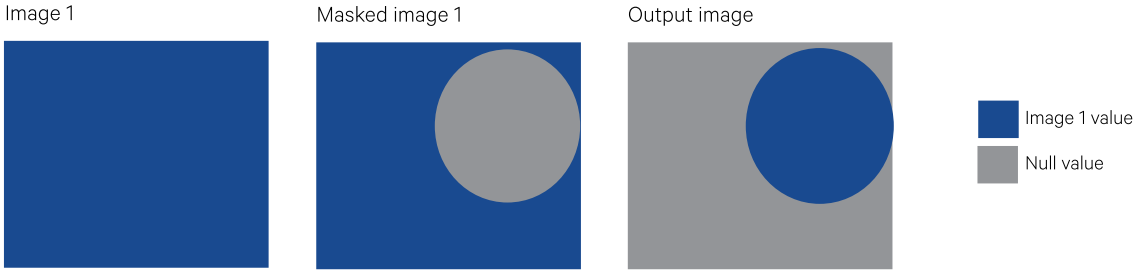
b. Transfer null values from image 1 to image 2



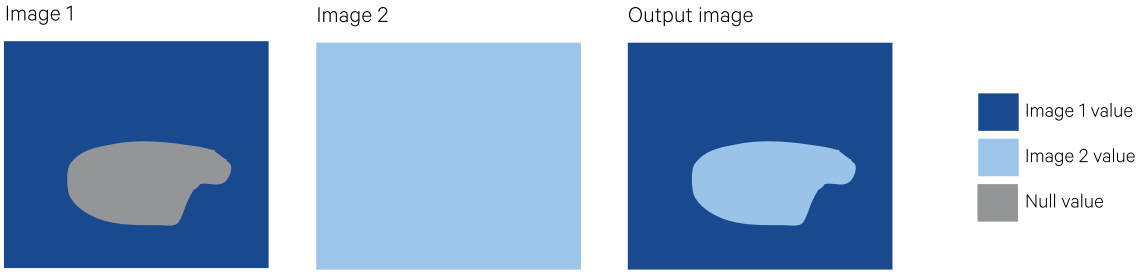
c. Combine null values from both images into image 1



d. Complement an image mask.



e. Recombine masked regions



Source: Harrison and Jupp (1990) Figures 61 and 62

5.4 Transferring Category Statistics

A unique operation in some image processing systems involves generating spectral statistics for sets of pixels that are grouped by a reference channel or image. Statistics relating to the mean, standard deviation, number of pixels, and optionally the covariance matrix could be generated based on values from input image 1 for each category defined in input image 2. Typically this process is used during image classification (see Volume 2A—Section 9.3.2 or Volume 2E). Alternatively, the reference channel may be an ancillary data channel that defines regions of interest. For example, this technique could be used to analyse the relationship between crop yield (derived from EO data) and soil type regions.

In the classification context, the statistics indicate the variability of some particular set of categories (or classes) in terms of the input image(s). The categories may have been formed from the input image(s), in which case the operation will simply be one of mean migration (see Volume 2A—Section 9.2.3 or Volume 2E). The input images could also be a different set of channels in the image (either before or after processing) or a different, but registered, data source. For example, a classification based on transformed channels such as Principal Components can be transferred back to the original image data for subsequent processing with reference to a familiar colour composite image. Alternatively, a classification may be performed using visible and NIR channels then transferred to SWIR and TIR channels to assess the variability of the classes in terms of those channels, and hence possibly which classes need to be split using those additional channels. This technique has also been used to split a set of land cover classes into sunlit and shaded categories by gathering statistics for the classes from a topographicallyenhanced channel (Ahmad *et al.*, 1989).

In a bi-temporal context, this spectral transfer process can be used to transfer a classification from an image at one date to a registered image at a different date. Major changes in image features will be highlighted in the 'residual' image (the difference between the original and classified images) of the second date (see Figure 5.4 and Volume 2E). Naturally, when different image sources are being used to transfer a classification, the two sources must be accurately registered (see Section 3.1 and Volume 2B).

5.5 Colour Balancing for Image Mosaicking

The geometric aspects of sewing and mosaicking multiple images are discussed in Volume 2A—Section 5.2.2 and Volume 2B. When the imagery being mosaicked is derived from different sources, a colour balancing or rescaling operation (see Volume 2C—Section 2) is usually required to ensure that the same features have similar values in the images being joined, so that the join position has low visual impact.

A number of methods may be used to rescale or 'balance' image colours. These usually involve matching the (normalised) histogram values of each image to obtain equivalent data ranges. A simple implementation of this is to use the channel histogram percentage points (see Volume 2A—Sections 4.1.1 and 8.1.1).

For example, histogram percentage points for a pair of channels A and B are shown in Table 5.1. These values can be compared graphically as shown in Figure 5.7. A regression line through the majority of points then defines a conversion scale from values in one image to the corresponding values in another image. The 1% and 99% points, and often the 5% and 95% points, are ignored for the regression if they depart significantly from a linear function. To rescale these channels, a good contrast stretch based on their actual histograms would have a minimum of 13 and a maximum of 56 in image B, which relate to the values 7 and 31 in image A. The selected scaling values should be checked visually, however, before rescaling images (Volume 2C—Section 2). However, if the mosaicked imagery is to be processed further, the maximum ranges should be used for rescaling to avoid truncation of values. In this example, to rescale channel A to match channel B, we would use a coefficient of 1/1.7 (= 0.588) with an offset of 2.

Table 5.1 Example histogram percentage points

Channel	Histogram percentage points												
	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
A	11	11	13	14	14	16	16	17	18	20	21	23	32
B	17	20	22	24	26	28	29	30	32	35	39	44	86

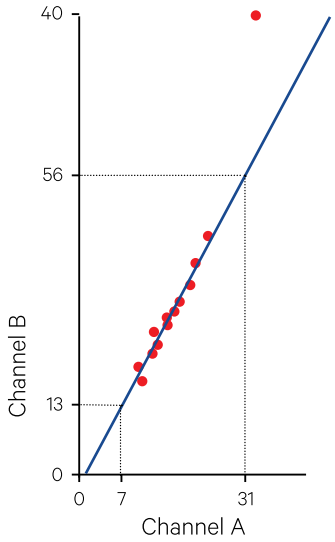
In a mosaicked image, the join lines between different input images are often visible even when care is taken to achieve a good colour balance before mosaicking (see Section 5.5). If sufficient overlap occurs between the input images after they have been registered to a common base, the join boundary can be manually defined along a linear feature in the image, such as a road or river, in the overlap region. This process can be considered as the digital equivalent of photo-mosaicking.

Typically, this operation would require resampling of the images to a common base (see Volume 2B) then digitising along a feature using the screen cursor (see Volume 2A—Section 10.1) to remove part of the overlap region in one image. The digitised image can be used as a mask on the second image to remove the common overlap pixels (see Section 5.3). This results in two registered images with no overlapping pixels, but which abut along an image feature. These two images can then be recombined using image pasting (see Section 5.3) to produce the final mosaic.

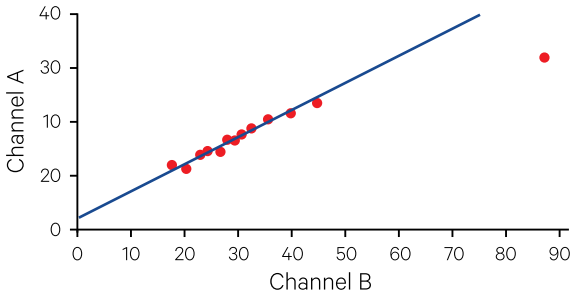
Figure 5.7 Colour balancing

Histogram statistics can be used to match contrast ranges between different images.

a. Channel B = 1.7 x Channel A- 2



b. Channel A = 1/1.7 x Channel B + 2



Source: Harrison and Jupp (1990) Figure 63

5.6 Further Information

Jensen (2016) Chapter 12

5.7 References

Ahmad, W., Jupp, D.L.B., and Nunez, M. (1989). Land cover mapping in a rugged terrain area using Landsat MSS data. *International Journal of Remote Sensing,* 13(4), 673–683. *https://doi.org/10.1080/01431169208904145*

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne.

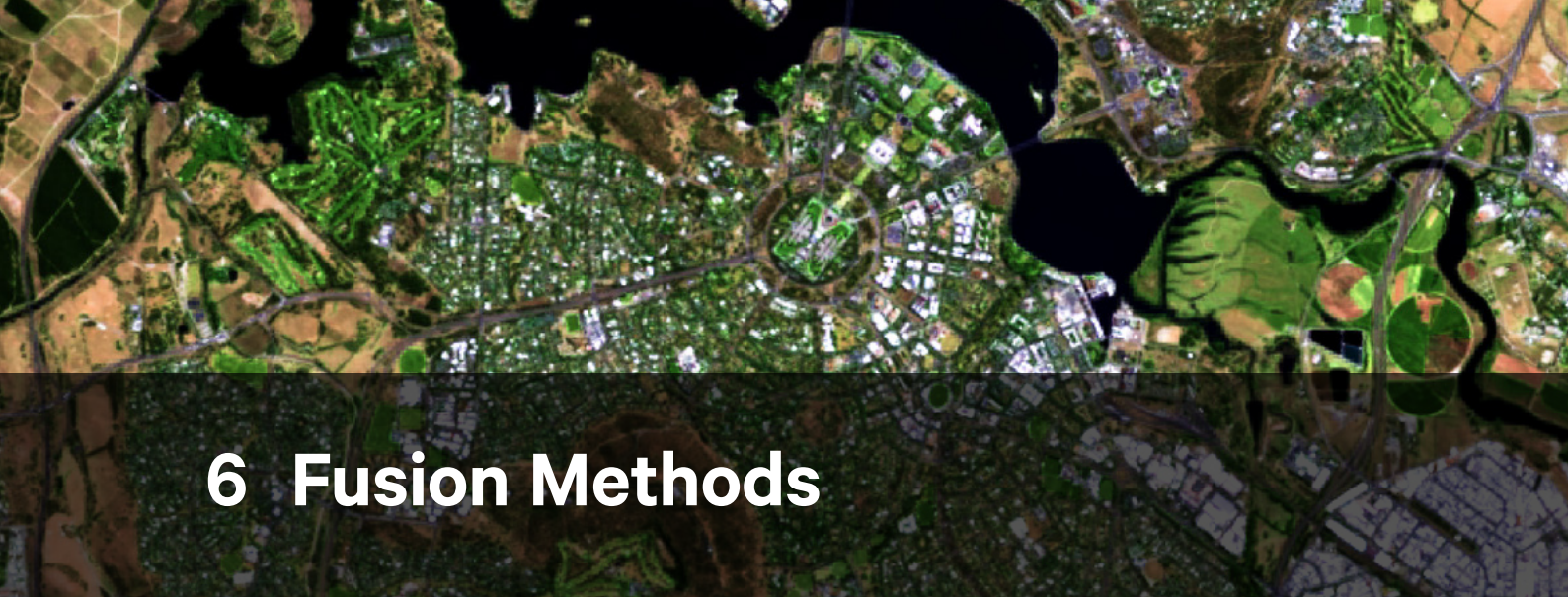
Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th Edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Jupp, D. L. B., and Mayo, K. K. (1982). The use of residual images in Landsat image analysis. *Photogrammetric Engineering and Remote Sensing*, 48(4), pp. 595–604.

Jupp, D. L. B., Heggen, S. J., Mayo, K. K., Kendall, S. W., Bolton, J. R., and Harrison, B. A. (1985a). The BRIAN Handbook. *An Introduction to Landsat and the BRIAN (Barrier Reef Image Analysis) System for Users, CSIRO Division of Water and Land Resources, Natural Resources Series No.3*. CSIRO, Canberra, Australia.

Jupp, D. L. B., Mayo, K. K., Kuchler, D. A., Heggen, S. J., Kendall, S. W., Haywood, M. J., and Ayling, T. (1985b). *A Landsat-based interpretation of the Cairns Section of the Great Barrier Reef Marine Park. Interpretation of Landsat data by Computer-based Classification and Labelling, CSIRO Division of Water and Land Resources, Natural Resources Series No.4*. CSIRO, Melbourne, Australia.

Townshend, J.R.G., Justice, C.O., Gurney, C., and McManus, J. (1992). The impact of misregistration on change detection. *IEEE Geoscence and Remote Sensing*, 30, 1054–60.



6 Fusion Methods

The four dimensions of EO dataset resolution, density and extent (namely spectral, spatial, radiometric and temporal) are introduced in Volume 1B—Section 1 and Volume 2C—Section 1, and summarised in Table 2.1 above. Sensor design constraints, such as detector efficiency in recording the incoming radiation and the data volume that can be be recorded or transmitted, make it difficult to deliver imagery with high resolution in more than one of these dimensions (see Volume 1A—Section 13).

A number of fusion methods have been proposed to combine different sources of EO imagery to achieve higher resolution in one dimension (Wald, 1999). Some of these methods use dynamic, physics-based models (Heinsch *et al*., 2006; Renzullo *et al*., 2008), while others rely on complementary EO-based datasets (Emelyanova *et al*., 2012, 2013). Methods solely relying on EO datasets are typically called fusion or blending algorithms and include:

- spatial and spectral blending, such as pan-sharpening—merging a higher spatial, but lower spectral, resolution image (generally panchromatic) with a compatible multispectral image that has lower spatial, but higher spectral, resolution to create a multispectral image with sharper spatial features (see Section 6.1); and
- spatial and temporal blending—merging imagery with higher temporal, but lower spatial, resolution with EO image datasets that have lower temporal, but higher spatial, resolutions to create a denser time series with higher spatial resolution (see Section 6.2).

Since terminology in this area can become confused, some working definitions of terms commonly used by the 'data fusion' research community are summarised in Table 6.1. Methods to create a composite image from time series datasets, in which the maximum number of pixels are unaffected by cloud, haze, smoke or sensor errors, are described in Section 10.

Data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources, in order to obtain information of greater quality.
Wald (1998)

Background image: Landsat-8 panchromatic and OLI images acquired over Canberra on 25 October 2018 have been merged to create a colour image with the spatial resolution of the panchromatic image. **Source:** Norman Mueller, Geoscience Australia.

Table 6.1 Terminology associated with data fusion

Term	Definition	Relevant Dimension(s)	Reference
Aggregating	Combining a scale of measurement		McVicar *et al.* (2002)
Blending	Merging images from different sources to generate information with high spatial and temporal resolution (see Section 6.2)	spectral spatial temporal	Hilker *et al.* (2009)
Compositing	Combining spatially overlapping images into a single image based on an aggregation function (see Volume 2B)	spatial temporal	Google Earth Engine
Disaggregation	Splitting a scale of measurement	spatial	McVicar *et al.* (2002)
Downscaling	Producing an image of finer resolution from an image of coarser resolution (see Section 4.3.2)	spatial	McVicar *et al.* (2002)
Prediction	Obtaining future image estimations in regular time step sequences by modelling future spatial-temporal states (behaviour) of an observed system (see Section 14.3 and Volume 3)	spatial temporal	Crespo *et al.* (2007)
Pan-sharpening	Integrating spatial detail in a higher resolution panchromatic image with the spectral information of a lower resolution multispectral image to produce a high resolution multispectral image (see Section 6.1)	spectral spatial	Du *et al.* (2007)
Reconstruction	Recovering detail in severely blurred images, the causes of whose imperfections are known a priori	spatial	Bates and McDonnel (1986)
Restoration	Recovering an image that has been recorded in the presence of one or more sources of degradation	spatial	Bates and McDonnel (1986)
Sharpening	Highlighting fine detail in an image or enhancing detail that has been blurred, either in error or as a consequence of the image acquisition method	spectral spatial	Gonzalez and Woods (1992)
Spectral unmixing	Classifying mixed pixels by deconvolving (unmixing) each pixel spectrum into fractional abundances of its surface constituents or endmember spectra (see Volume 2E)	spectral spatial	Rogge *et al.* (2006)
Upscaling	Producing an image of coarser resolution from an image of finer resolution (see Section 4.3.1)	Spatial	McVicar *et al.* (2002)

Adapted from Emelyanova *et al.* (2012)

6.1 Spatial and Spectral Blending

Some of the methods commonly used for panchromatic sharpening (or pan-sharpening)—blending the higher spatial resolution of a panchromatic image with the higher spectral resolution of a multispectral image—are introduced in Volume 2C. Pan-sharpening is most relevant for visual identification of detailed spatial features, such as intra-urban land cover and landslide scars (Fonseca *et al.*, 2011).

Prerequisites for quality results from image fusion methods include matching the panchromatic and multispectral images as closely as possible in terms of:

- acquisition date;
- spectral and radiometric extent;
- dynamic range within relevant spectral bands; and
- geometric registration (Fonseca *et al.*, 2011).

A wide range of methods has been proposed for pan-sharpening imagery (Amro *et al.*, 2011). All methods, and their numerous variations and hybrid versions, present different advantages and disadvantages (see Table 6.2).

Zhang (2004) lists the 'most popular and effective' pan-sharpening methods used for EO imagery as:

- HSI (Hue-Saturation-Intensity)-RGB (red-green-blue) conversion, where the intensity component of the MSS image is replaced with the panchromatic image (see Volume 2C—Section 8);
- PCA (Principal Components Analysis), where PC1 is replaced with the panchromatic image (see Volume 2C—Section 9);
- arithmetic combinations, such as the Brovey Transform (see Volume 2C—Section 8.3), Synthetic Variable Ratio (SVR) or Ratio Enhancement (RE) (see Volume 2C—Section 10); and
- wavelet fusion (Mallat, 1989), where the panchromatic image is decomposed into spatial details (wavelet coefficients), which are subsequently merged with each multispectral band by reverse wavelet transforms (Pajares and de la Cruz, 2004; Amolins *et al.*, 2007).

Table 6.2 Pan-sharpening methods

Category	Example methods	Advantages	Disadvantages	Reference
Substitute component derived from multispectral image with one derived from panchromatic image	Hue-Saturation-Intensity (HIS) Principal Components Analysis (PCA) Gram-Schmidt (GS)	Computationally simple	Colour distortions when substituted component does not contain same spectral information as original component	Dou *et al.* (2007)
Compute a linear combination of multispectral and panchromatic bands then replace relevant component in multispectral image	Brovey Transform	Computationally simple	Spectral distortion	Tu *et al.* (2005)
Inject high frequency information from panchromatic image into multispectral image	High pass filtering of panchromatic image then adding edge detail back into the multispectral image	Low spectral distortion	Can generate ripple effect	Ranchin and Wald (2000)
Exploit statistical characteristics of panchromatic and multispectral images	Price's method Spatially Adaptive Algorithm	Preserves low resolution radiometry Can use multiple panchromatic images for high frequency input	Can produce blocking artefact	Price (1999) Park and Kang (2004)
Use Bayesian framework to model images and estimate the pan-sharpened image	Markov Random Fields Stochastic Mixing Models	Defines pan-sharpening in clear probabilistic framework	Image noise is assumed to be modelled by Gaussian distribution	Fasbender *et al.* (2008)
Decompose multispectral and panchromatic images into different levels of spatial detail then merge higher frequency panchromatic detail into multispectral image	Laplacian pyramid Wavelet fusion	Good spectral representation	Quality of pan-sharpened image depends on the right number of decomposition levels	Burt and Adelson (1983) González-Audicána and Otazu (2005)

Source for data: Amro *et al.* (2011)

The first three methods listed above are sensitive to geometric registration errors (Jiang *et al.*, 2013) and are reliant on a skilled operator to minimise colour distortion and maximise fusion quality. This is particularly true for panchromatic imagery acquired by EO satellites launched after 1999 (Zhang, 2004), which have a wider spectral range to include both visible and IR wavelengths. These three methods also require prior upscaling of the multispectral image to match the spatial resolution of the panchromatic image (see Section 4.3).

By comparison, the wavelet transform reduces colour distortions and image noise (Thomas *et al.*, 2008; Jiang *et al.*, 2013). Variations of this fusion method include the Ridglet, Curvelet (Candès and Guo, 2002; Choi *et al.*, 2005) and Contourlet transforms (Donoho and Vetterli, 2002). Artificial neural networks (AAN; Shutao *et al.*, 2002) and the Dempster-Shafer theory of evidence (DS; Wu *et al.*, 2002; Le Hégarat-Mascle *et al.*, 2003) have also be used for image fusion of EO imagery.

Recommendations for assessing the spectral and spatial quality of pan-sharpened images (Wald *et al.*, 1997; Du *et al.*, 2007) can be summarised in terms of two properties:

- consistency—if the pan-sharpened image is downsampled to match the spatial resolution of the multispectral image, it should closely resemble the original multispectral image; and
- synthesis—the pan-sharpened image should resemble an image that would be acquired by an appropriate sensor with corresponding spectral and spatial resolutions (Amro *et al.*, 2011).

The ideal outcome of pan-sharpening is an image with enhanced spatial information from the panchromatic image that does not compromise spectral detail from the multispectral image. This is achieved by considering the physics of image acquisition and adapting methods appropriately (Thomas *et al.*, 2008).

6.2 Spatial and Temporal Blending

A number of methods have been proposed to merge a high temporal resolution dataset, such as MODIS, with higher spatial resolution imagery (such as Landsat TM/ETM+/OLI). Examples of such methods include:

- STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model; Gao *et al.*, 2006);
- STAARCH (Spatial Temporal Adaptive Algorithm for mapping Reflectance Change; Hilker *et al.*, 2009);
- downscaling unmixing models (Zurita-Milla *et al.*, 2009);
- STRUM (Spatial and Temporal Reflectance Unmixing Model; Gevaert and Garcia-Haro, 2015); and
- semi-physical BRDF spectral model (Roy *et al.*, 2008).

These methods attempt to simulate image datasets with higher spatial and temporal resolutions, which may be informative for monitoring rapidly changing biophysical processes. Since the Terra/MODIS and Landsat sensors have equivalent orbital parameters and acquire imagery at similar times of the day, their images have similar viewing and illumination geometries (see Volume 1B—Section 3). These sensors also share similar spectral resolutions for their red and NIR bands (see Table 6.3). Given these similarities, several algorithms have been proposed to synthesise daily Landsat images using coincident Terra/MODIS datasets. A generic theoretical framework for such methods is described in Excursus 6.1.

Table 6.3 Landsat and Terra/MODIS characteristics

Characteristic	Landsat TM/ ETM+/OLI	Terra/MODIS
Equatorial crossing time	~10:00 am	~10:30 am
Altitude	705 km (L-8)	713 km
Red bandwidth (nm)	630–690	620–670
NIR bandwidth (nm)	780–900	841–876
Spatial resolution (red and NIR)	30 m	250 m (red/NIR)
Revisit frequency	16 days	≤ 1 day

Adapted from Emelyanova *et al.* (2012) Table 4

One such method is the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) proposed by Gao *et al.* (2006), which uses the MODIS daily 500 m surface reflectance product (MOD91GHK; see Section 10) to effectively interpolate the Landsat time series. To ensure that both image datasets are spatially and radiometrically comparable, calibrated image products are used (see Section 3.1). This method is most simply described with reference to Figure 6.1, which shows how a higher resolution image at time T1 is processed with lower resolution imagery at times T1 and T0 to synthesise a higher resolution image at time T0 :

- step 1—within a moving window, pixels that are spectrally similar to the central pixel are identified in the fine resolution image to use as sample pixels;
- step 2—sample pixels are filtered;
- step 3—both fine and coarse resolution data are used to determine weights (darker shading indicating heavier weights) for each sample pixel based on the:
  - spectral difference between T1 Landsat and T1 MODIS surface reflectance values;
  - temporal difference between T1 and T0 MODIS surface reflectance values; and
  - spatial Euclidean difference between the sample pixel and the central pixel; then
- step 4—the weighted values are used to a predict a higher resolution surface reflectance value of the central pixel at time T1.

However, this sequence is applied to each spectral band separately, which can generate unrealistic spectra in the synthesised image (Gevaert and Garcia-Haro, 2015).

A significant advantage of blending methods based on unmixing is that they do not require the higher and lower resolution image datasets to have corresponding spectral bands (Gevaert and Garcia-Haro, 2015). This gives more flexibility on input datasets and allows inclusion of biophysical parameters and non-EO datasets. Such algorithms involve four steps:

- cluster higher resolution dataset to define end members;
- calculate fractions (abundances) of each end member in each lower resolution pixel;
- unmix lower resolution pixel; and
- assign reflectance spectra to higher resolution pixels (Gevaert and Garcia-Haro, 2015);

This approach delivers more reliable results for gap-filling and cloud masking, especially when the high resolution dataset is temporally sparse (Gevaert and Garcia-Haro, 2015). It is also computationally less expensive and creates synthesised images with more realistic spectra (see Figure 6.2). An approach that blends STARFM with the unmixing blending method is called STRUM, which computes simulated surface reflectance values that are more accurate than either single method, but requires corresponding spectral bands in both input datasets. Thus, for a pair of MODIS and Landsat images on the base date and a MODIS image on the prediction date, STRUM creates a synthesised Landsat image on the prediction date using these steps:

- cluster higher resolution dataset to define end members;
- calculate fractions (abundances) of each end member in each MODIS pixel;
- unmix the difference image between the two input MODIS images; and
- assign temporal change of the relevant endmember to each synthesised Landsat pixel (Gevaert and Garcia-Haro, 2015).

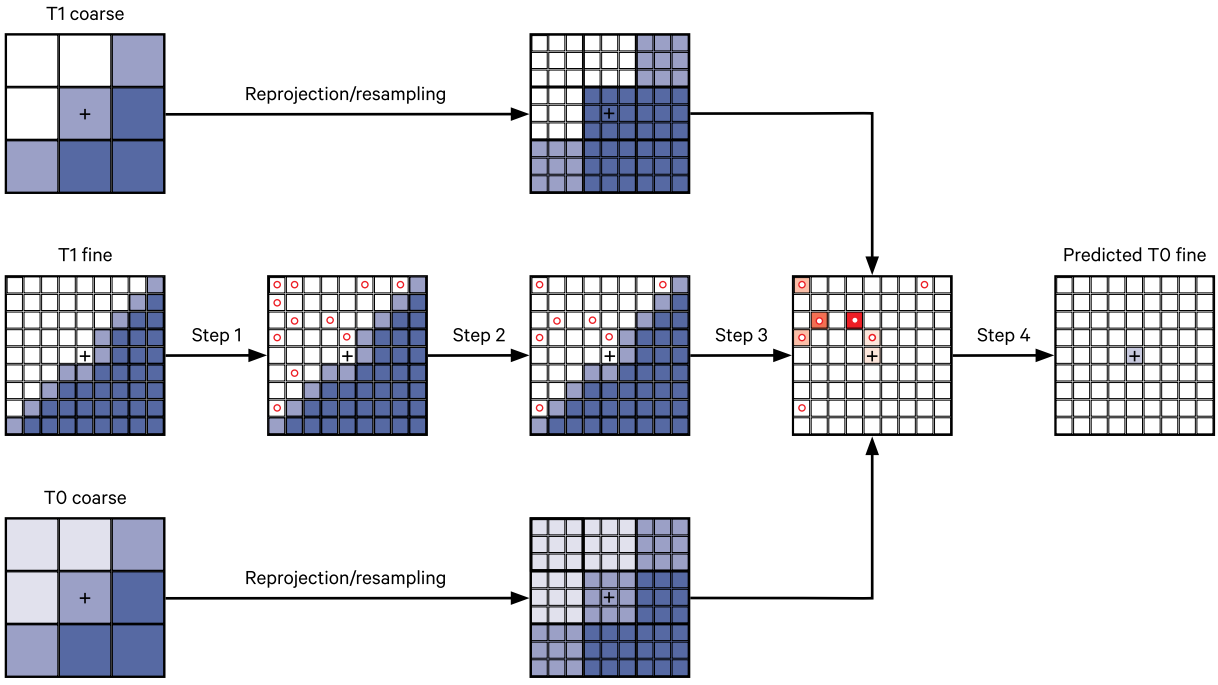
All blending methods are based on assumptions related to the rate of change in surface reflectance with time and the extent of spatial variance and autocorrelation within an image. Accordingly, different methods may generate more appropriate results for particular datasets. Since blending algorithms can generate errors in terms of changing spatial and temporal variance between the input images, their performance needs to be assessed in terms of accuracy and constraints. Emelyanova *et al.* (2013) compared four blending algorithms in two areas of NSW with contrasting land cover and inundation dynamics:

- LIM (Linear Interpolation Model; Emelyanova *et al.*, 2012);
- GEIFM (Global Empirical Image Fusion Model; Emelyanova *et al.*, 2012);
- STARFM; and
- ESTARFM (Enhanced STARFM; Zhu *et al.*, 2010).

The results of these analyses for two study sites are summarised in Figure 6.3. This study concluded that more sophisticated blending algorithms do not necessarily produce better results if the lower resolution input imagery does not sample the spatial variance of land cover changes adequately. It also observed that the site with the most dynamic spatial and temporal changes proved to be the most challenging for all algorithms.

Figure 6.1 STARFM approach

In this four step schematic, T1 is the base date for the input images and T0 is the image date being simulated at a higher resolution (or the prediction date). In step 1, within a moving window on finer resolution imagery, circles indicate pixels that are determined to be spectrally similar to the central pixel (marked with a cross). In step 2, these sample pixels are filtered. In step 3, both fine and coarse resolution data are used to determine sample weights, with darker shading indicating heavier weights. In step 4, the weighted values are used to compute the prediction value for the synthetic image.

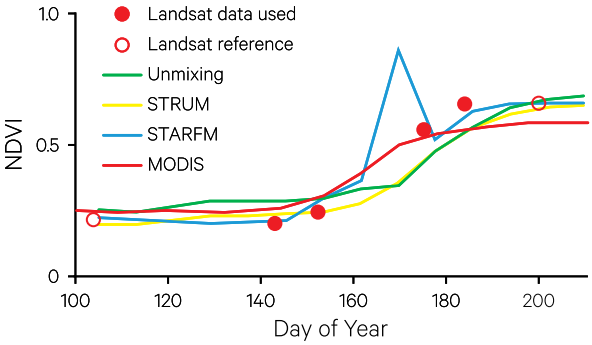


Adapted from Gao *et al.* (2006) Figure 1

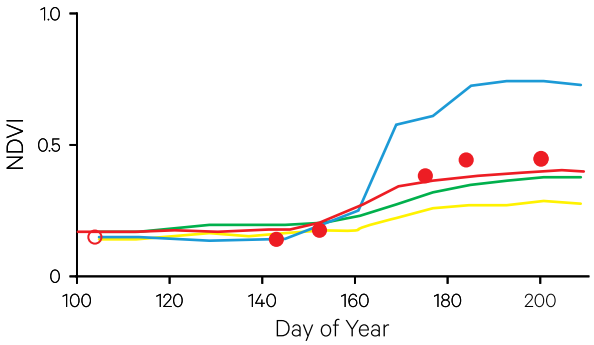
Figure 6.2 Temporal profiles of blending algorithms

For a predominantly agricultural study area, three temporal profiles from blending MODIS and Landsat imagery are shown in terms of the Normalised Difference Vegetation Index (NDVI).

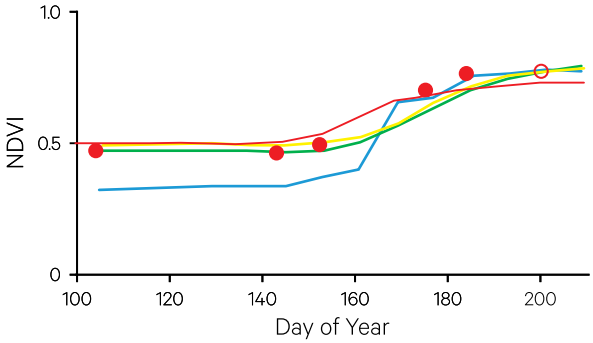
a. Based on all available Landsat images



b. Based based on first available Landsat image only



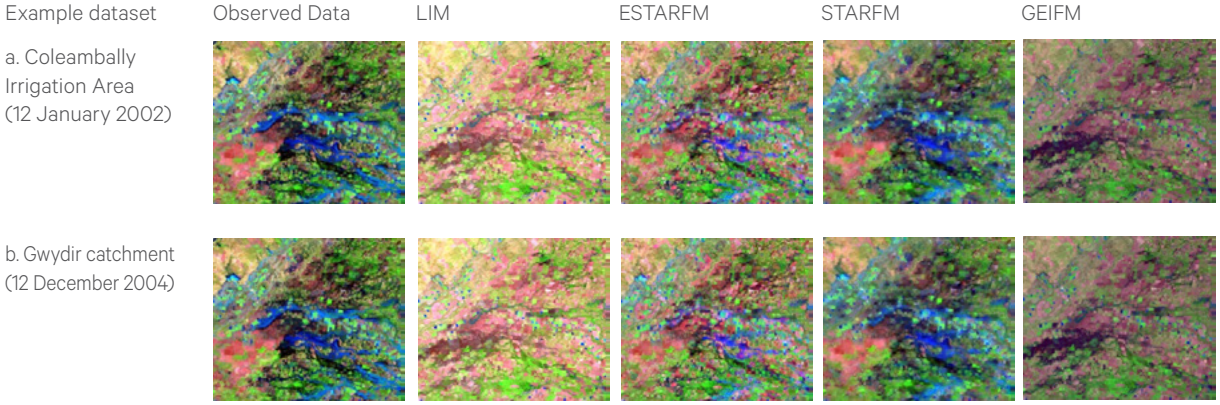
c. Based on last available Landsat image only



Source: Gevaert and Garcia-Haro (2015) Figure 6

Figure 6.3 Image blending algorithms

Observed Landsat data (bands 5, 4, 3 as RGB) compared with simulated Landsat-like images using four different blending algorithms.



Source: Emelyanova *et al.* (2013) Figure 3

Excursus 6.1—Theoretical Framework for Image Blending

Source: Emelyanova *et al.* (2012)
Further Information: Gao *et al.* (2006)

This Excursus presents a generic framework for simulating higher spatial resolution imagery at required dates (see Figure 6.4a). The example below blends Landsat and MODIS images, however the framework is adaptable to other image types. In this case:

- Landsat (l) is an example of high spatial resolution / low temporal density imagery; and
- MODIS (M) is an example of low spatial resolution / high temporal density imagery.

For clarity, the notation used below for different image datasets, variables and operators is summarised in Table 6.4.

Table 6.4 Notation for blending framework

Term	Definition
t_n	Image acquisition date n
t_s	Date of simulated image
l	Observed high spatial resolution/low temporal density image
L	l upscaled to the coarse resolution $l_t(X,Y)$
(x,y)	Locations in l
$l_t(x,y)$	Time series of l
$l_t(X,Y)$	Upscaled reflectance values of l
$\lambda_s(x,y)$	Simulated high spatial resolution/high temporal density image
M	Low spatial resolution / high temporal density image
(X,Y)	Locations in M
$M_t(X,Y)$	Time series of M
$M_t(x,y)$	Downscaled reflectance values of M
Ψ	Maps reflectance values from $M_t(X,Y)$ to $l_t(x,y)$
ε	Error between the observed $l_t(x,y)$ and modelled $\Psi(M_t(X,Y))$ data due to instrument and/or model discrepancies
h_t	Transformation of reflectance values from a finer to a coarser resolution
f_t	differences in sensor systems acquiring l and M

Time series of Landsat and MODIS images for a particular band can be described by the spatial and temporal functions $l_t(x,y)$ and $M_t(X,Y)$ respectively, where (x,y) and (X,Y) respectively denote spatial locations of fine and coarse spatial resolution pixels for images observed at acquisition date t. Since these spatial locations represent sample and line numbers of the pixels within the images, the geographical locations (longitude and latitude) of the centres of fine and coarse pixels with same spatial locations within the images are not necessarily the same. Notation $M_t(x,y)$ represents reflectance values of the MODIS image downscaled to the Landsat resolution. The simulated Landsat-like data is denoted $\lambda_s(x,y)$ to distinguish it from the observed (or measured) Landsat reflectances, $l_t(x,y)$.

Assuming that Landsat and MODIS sensors have similar spectral characteristics for three visible spectral bands (blue, green and red, which are B1, B2 and B3 Landsat bands respectively) and three IR spectral bands (NIR infrared and two SWIR which are B4, B5 and B7 Landsat bands respectively), a relationship between Landsat and MODIS surface reflectances for a particular band can be performed by an operator Ψ which maps a MODIS reflectance value $M_t(X,Y)$ into a Landsat reflectance value $l_t(x,y)$:

$$l_t(x,y) = \Psi\,(M_t(X,Y)) + \varepsilon$$

where

- ε represents the error between the observed $l_t(x,y)$ and modelled $\Psi\,(M_t(X,Y))$ data due to instrument and/or model discrepancies.

The operator $\Psi : M \xrightarrow{f_t^{-1}} L \xrightarrow{h_t^{-1}} l$ provides a two-step mapping from:

- set M of the MODIS image reflectance values at the coarse resolution observed at acquisition date t $\{M_t(X,Y)\}$ to
- set L of the Landsat reflectance values upscaled to the coarse resolution $l_t(X,Y)$ to
- set l of the observed fine resolution Landsat reflectance values $\{\, l_t(x,y) \,\}$.

If h_t describes the transformation of reflectance values from a finer to a coarser resolution:

$$l_t(X,Y) = h_t\,(l_t(x,y))$$

and f_t accounts for the differences in sensor systems (such as orbit parameters, bandwidths and spectral response functions) between the concurrent Landsat and MODIS surface reflectance measurement, then:

$$M_t(X,Y) = f_t\,(l_t(X,Y))$$

and we can define the operator Ψ as a composition of the inversed scale and reflectance sensor conversion functions:

$$\Psi_t = H_t \circ F_t$$

where $H_t = h_t^{-1}$ and $F_t = f_t^{-1}$.

We assume that these inverse functions exist. Please note, this composition means a consecutive application of function F_t and then function H_t.

Therefore, a change in Landsat reflectances at pixel (x,y) over a time interval $[t_1\,,\,t_2\,]$ can be described as a change in the transformed MODIS reflectance values:

$$l_2(x,y) - l_1(x,y) = \Psi_2\,(\,M_2(X,Y)) - \Psi_1\,(\,M_1(X,Y))$$

While the spatial and spectral resolution relationships h_1 and f_1 , respectively, can be modelled using the concurrent pair of observed Landsat and MODIS data recorded at time t_1 , it is impossible to explicitly describe functions h_s and f_s because the problem is ill-posed due to the absence of Landsat measurements of spectral reflectances $l_s(x,y)$ for the date of simulation. However, we can substitute h_s and f_s by the models derived from the available data to generate Landsat-like reflectances for time t_s .

From the definition of f_t above, the function f_s models the inter-sensor reflectance transformation and ideally is dependent on the inter-sensor spectral characteristics only (that is, theoretically f_s is spatially and temporally independent). The relationship described by function f_1 can be used to estimate MODIS reflectances from the Landsat reflectances observed at time t_s , that is, $t_1 \approx t_s$. We describe the MODIS temporal reflectance changes during the time period $[t_1\,,\,t_s\,]$ by the function:

$$g_s : M_s(X,Y) = g_s\,M_1\,(X,Y)$$

which is intra-sensor and thus independent of spatial resolution, so:

$$M_s(x,y) = g_s\,(M_1\,(x,y))$$

The MODIS reflectance grid recorded at the simulation date t_s can then be downscaled by:

$$H_s\,(M_s(X,Y)) = M_s(x,y)$$
$$= g_s\,(M_1\,(x,y)) = g_s\,H_1\,(M_1\,(X,Y))$$

If g_s and h_1 are linear functions that obey the distributive law (Kolmogorov and Fomin, 1999):

$$\lambda_s(x,y) = l_1(x,y) + F_1\,(H_s\,(M_s(X,Y))$$
$$- H_1\,(M_1(X,Y)))$$

they can be used for simulating Landsat-like reflectances at time t_s from available measurements of Landsat and MODIS reflectances observed at time t_1 and MODIS data acquired at time t_s . After downscaling, this can be rewritten:

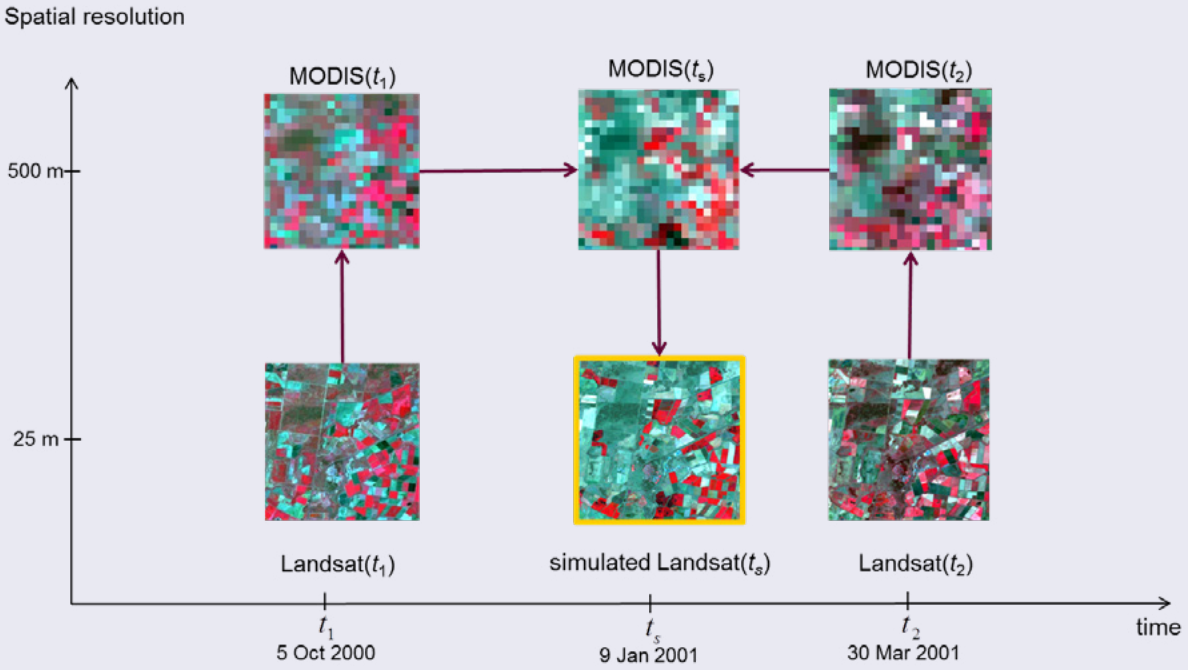
$$\lambda_s(x,y) = l_1\,(x,y) + F_1\,(M_s(X,Y) - M_1\,(X,Y))$$

This is a generic equation that describes Landsat-MODIS blending algorithms if the inter-sensor, spatial resolution and temporal relationships between Landsat and MODIS reflectances are assumed to be linear and have inverse functions (except for the temporal function as $t_1 < t_s < t_2$). A schematic of the generic blending process is shown in Figure 6.4b.

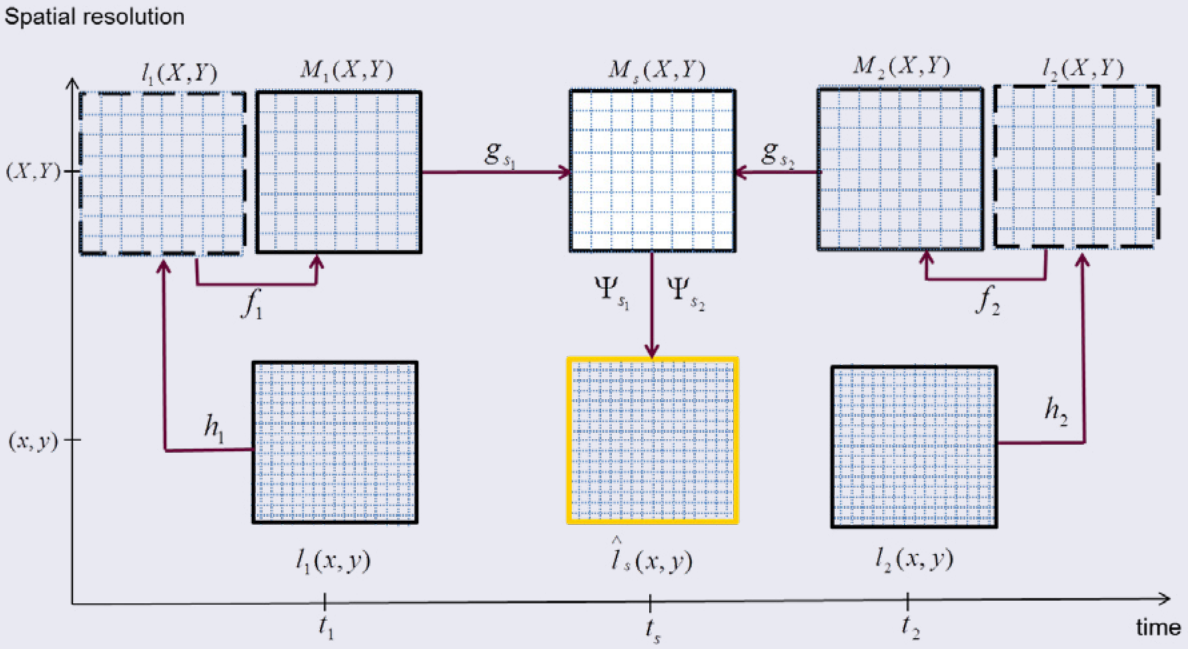
Figure 6.4 Generic overview of Landsat-MODIS blending.

The direction of information flow is represented by the arrows and the simulated image is indicated by the yellow line.

a. Simulation of a Landsat-like image at time t_s from two Landsat-MODIS image pairs (at times t_1 and t_2) and one MODIS image observed at t_s where $t_1 < t_s < t_2$.



b. Schematic representation of the blending procedure, see Table 6.4 for the description of the notation. The outer black solid and dashed lines represent observed and calculated data, respectively, and $L_1(x,y)$ and $L_2(x,y)$ are spatially upscaled Landsat images associated with 5 October 2000 and 30 March 2001, respectively.



6.3 Further Information

Chang and Bai (2018)

Perracchione *et al.* (2018)

Zheng *et al.* (2018)

6.4 References

Amolins, K., Zhang, Y., and Dare, P. (2007). Wavelet based image fusion techniques—An introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62 (4), 249–263. *https://doi.org/10.1016/j.isprsjprs.2007.05.009*

Amro, I., Mateos, J., Vega, M., Molina, R., and Katsaggelos, A.K. (2011). A survey of classical methods and new trends in pansharpening of multispectral images. *EURASIP Journal on Advances in Signal Processing*, 2011:79.

Bates, R.H.T., and McDonnel, M.J. (1986). *Image Restoration and Reconstruction*. Clarendon Press, Oxford.

Burt, P.J., and Adelson, E.H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, COM-3l(4), 532–540.

Crespo, J.L., Zorrilla, M., Bernardos, P., and Mora, E. (2007). A new image prediction model based on spatio- temporal techniques. *The Visual Computer*, 23(6), 419–431.

Candès, E.J., and Guo, F. (2002). New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction. *Signal Processing*, 82, 1519–1543.

Chang, N.-B., and Bai, K. (2018). *Multisensor Data Fusion and Machine Learning for Environmental Remote Sensing*. CRC Press, Boca Raton. ISBN 9781315154602

Choi, M., Kim, RY., Nam, MR., and Kim, H.O. (2005). Fusion of multi-spectral and panchromatic satellite images using the Curvelet transform. *IEEE Geoscience and Remote Sensing Letters*, 2(2), 136–140. *https://doi.org/10.1109/LGRS.2005.845313*

Donoho, M.N., and Vetterli, M. (2002). *Contourlets*. Academic Press, New York. ISSN 0890-5401

Dou, W., Chen, Y., Li, X., Sui, D.Z. (2007). A general framework for component substitution image fusion: An implementation using the fast image fusion method. *Computers and Geosciences*, 33, 219–228.

Du, Q., Younan, N.H., King, R., and Shah, V.P. (2007). On the Performance Evaluation of Pan-Sharpening Techniques. *IEEE Geoscience and Remote Sensing Letters*, 4(4), 518–522.

Emelyanova, I.V., McVicar, T.R., Van Niel, T.G., Li, L.T., and van Dijk, A.I.J.M. (2012). *On blending Landsat–MODIS surface reflectances in two landscapes with contrasting spectral, spatial and temporal dynamics*. WIRADA project 3.4: Technical report. Australia: CSIRO Water for a Healthy Country Flagship. *https://doi.org/10.4225/08/5852dc6879ab8*

Emelyanova I.V., McVicar, T.R., Van Niel, T.G., Li, L.T., and van Dijk, A.I.J.M. (2013). Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sensing of Environment*, 133, 193–209.

Fasbender, D., Radoux, J., Bogaert, P. (2008). Bayesian data fusion for adaptable image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 1847–1857.

Fonseca, L., Namikawa, L., Castejon, E., Carvalho, L., Pinho, C., and Pagamisse, A. (2011). Image Fusion for Remote Sensing Applications. Ch 9 in *New Advances in Image Fusion*. (Ed: Zheng, Y.) InTechOpen.

Gao, F., Masek, J., Schwaller, M., and Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8), 2207–2218. *https://doi.org/10.1109/TGRS.2006.872081*

Gevaert, C.M., and Carcía-Haro, F.J. (2015). A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment* 156, 34–44. *http://dx.doi.org/10.1016/j.rse.2014.09.012*

Gonzalez, R.C., and Woods, R.E. (1992). *Digital Image Processing*. 3rd Edn, Pearson/Prentice Hall, Harlow.

González-Audícana, M., and Otazu, X. (2005). Comparison between Mallat's and the à trous discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images. *International Journal of Remote Sensing*, 26(3), 595–614. *https://www.tandfonline.com/doi/abs/10.1080/01431160512331314056*

Heinsch, F.A., Zhao, M., Running, S.W., Kimball, J.S., and Nemani, R.R. (2006). Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. *IEEE Transactions on Geoscience and Remote Sensing*, 44(7), 1908–1925.

Hilker, T., Wulder, M.A., Coops, N.C., Seitz, N., White, J.C., Gao, F., Masek, J.G., and Stenhouse, G. (2009). Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sensing of Environment*, 113(9), 1988-1999.

Jiang, D., Zhuang, D., and Huang, Y. (2013). Investigation of Image Fusion for Remote Sensing. Ch 1 in *New Advances in Image Fusion*. (Ed: Zheng, Y.) InTechOpen.

Le Hégarat-Mascle, S., Richard, D., and Ottlé, C. (2003). Multi-scale data fusion using Dempster-Shafer evidence theory. *Integrated Computer-Aided Engineering*, 10(1), 9–22. *https://doi.org/10.3233/ICA-2003-10103*

Mallat, S.G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674–693,

McVicar, T.R., Davies, P.J., Qinke, Y., and Zhang, G. (2002). An Introduction to Temporal-Geographic Information Systems (TGIS) for Assessing, Monitoring and Modelling Regional Water and Soil Processes. In *Regional water and soil assessment for managing sustainable agriculture in China and Australia*. (Eds: McVicar, T.R., Rui, L., Walker, J., Fitzpatrick, R.W., and Changming, L.). ACIAR Monograph 84, Canberra, pp. 205–223, *http://www.eoc.csiro.au/aciar/book/PDF/Monograph_84_Chapter_16.pdf*

Pajares, G., and de la Cruz, J.M. (2004). A wavelet-based image fusion tutorial. *Pattern Recognition*, 37 (9), 1855–1872. *https://doi.org/10.1016/j.patcog.2004.03.010*

Park, J., and Kang, M. (2004). Spatially adaptive multi-resolution multispectral image fusion. *International Journal of Remote Sensing*, 25(23), 5491–5508. *https://doi.org/10.1080/01431160412331270830*

Perracchione, E., Polato, M., Tran, D., Piazzon, F., Aiolli, F., De Marchi, S., Kollet, S., Montzka, C., Sperduti, A., Vianello, M. Putti, M. (2018). *Data Fusion guidelines*. GEOEssential Deliverable 1.6, ERA-PLANET. *http://www.geoessential.eu/wp-content/uploads/2019/01/GEOEssential-D_1.6_final.pdf*

Price, J.C. (1999). Combining multispectral data of differing spatial resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1199–1203. *https://doi.org/10.1109/36.763272*

Ranchin, T., and Wald, L. (2000). Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66(1), 49–61.

Renzullo, L.J., Barrett, D.J., Marks, A.S., Hill, M.J., and Guerschman, J.P. (2008). Multi-sensor model–data fusion for estimation of hydrologic and energy flux parameters. *Remote Sensing of Environment*, 112(4), 1306–1319.

Rogge, D.M., Rivard, B., Zhang, J., and Feng, J. (2006). Iterative spectral unmixing for optimizing per-pixel endmember sets. *IEEE Transactions on Geoscience and Remote Sensing*, 44(12), 3725–3736.

Roy, D.P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., and Linquist, E. (2008). Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, 112(6), 3112–3130.

Shutao, L., Kwok, J.T., and Yaonan, W. (2002). Multifocus image fusion using artificial neural networks. *Pattern Recognition Letters*, 23, 985–997.

Thomas, C., Ranchin, T., Wald, L., and Chanussot, J. (2008). Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5), 1301–1312.

Tu, T.-M., Lee, Y.-C., Change, C.-P., and Huang, P.S. (2005). Adjustable intensity-hue-saturation and Brovey transform fusion technique for IKONOS/QuickBird imagery. *Optical Engineering,* 44(11), 116201.

Wald, L. (1998). Data fusion: a conceptual approach for an efficient exploitation of remote sensing images. *Proceedings of the 2nd International Conference on Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*. Sophia Antipolis, France. pp 17–24.

Wald, L. (1999). Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1190-1193.

Wald, L., Ranchin, T., and Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric Engineering and Remote Sensing,* 63, 691–699.

Wu, H., Siegel, M., Stiefelhagen, R., Yang, J. (2002). Sensor Fusion Using Dempster-Shafer Theory. IEEE Instrumentation and Measurement Technology Conference, 21-23 May 2002, Anchorage, AK, USA.

Zhang, Y. (2004). Understanding Image Fusion. *Photogrammetric Engineering and Remote Sensing,* 70(6), 657–661.

Zheng, Y., Blasch, E., and Lu, Z. (2018). *Multispectral Image Fusion and Colorization*. SPIE Press. ISBN 9781510619067 (Ch 1: *https://spie.org/samples/PM285.pdf*)

Zhu, X.L., Chen, J., Gao, F., Chen, X.H., and Masek, J.G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 114(11), 2610–2623.

Zurita-Milla, R., Kaiser, G., Clevers, J.G.P.W., Schneider, W., and Schaepman, M.E. (2009). Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sensing of Environment*, 113(9), 1874–1885.

7 Change Detection Methods

This Section considers bi-temporal methods for change detection, that is, detecting changes between two images acquired on different dates. Such methods compare an 'initial state' image (the first date) with a 'final state' image (the second date). Section 9 details time series ('temporal trajectory') methods, which isolate and analyse trends within a temporal sequence of multiple images.

The analysis of differences between pairs of EO images is relevant to a wide range of environmental applications, including assessment of:

- land cover changes in forests (Vogelmann, 1988), grasslands (Henebry, 1993), wetlands (Houhoulis and Michener, 2000; Munyati, 2000), and rangelands (Graetz *et al.*, 1988);
- land use changes in agricultural (Manavalan *et al.*, 1995) and urban areas (Jensen, 1983; Gupta and Munshi, 1985);
- defoliation (Muchoney and Haack, 1994), mortality (Macomber and Woodcock, 1994) and damage (Vogelmann and Rock, 1988; Vogelmann, 1990) in vegetated canopies;
- extent and severity of natural disasters, including fire (Jakubauskas *et al.*, 1990; Garcia-Haro *et al.*, 2001), flood (Zhou *et al.*, 2000), drought (Peters *et al.*, 2002) and landslides (Kimura and Yamaguchi, 2000); and
- glacial changes (Engeset *et al.*, 2002).

Change detection processing has traditionally required that one image be rewritten to precisely match the geometry of another (see Section 3.1 and Volume 2B—Section 5). As detailed in Volume 1, each EO image acquisition observes the reflectance/emissions from features on the Earth's surface. While these observations can be calibrated for internal consistency and comparison with reference standards, and associated with feature characteristics, they cannot be directly inverted into those features (see Section 1.1). Thus, changes detected between images are only indicative of surface feature changes.

However, some surface feature changes occur as part of natural cycles and may not be relevant to the type of changes being investigated. Accordingly, bi-temporal change detection methods simply attempt to highlight those pixels whose values differ between the 'initial state' image and the 'final state' image. Additional information is required to identify the type of change and suggest causal factors.

Prior to applying change detection methods, it may also be appropriate to remove irrelevant portions of the images, such as water in a land-based study, or pixels outside of a specified management region (see Section 4.1 and Volume 2A—Section 10). Some change detection studies have also segmented imagery into regions to assess changes within groups of pixels rather than on a per-pixel basis (Desclée *et al.*, 2006).

Most approaches to bi-temporal change detection comprise six fundamental stages, as summarised in Table 7.1). The results from these approaches can yield different levels of change information, such as:

- change or no change—locate the spatial distribution of differences;
- areal extent of change—use spatial distribution to compute the ground area of change;
- type of change—associate change patterns and extent with other information to determine the nature of change; and/or
- cause of change—correlate nature, area and distribution of change with other information to determine the reason(s) for changes.

Background image: Landsat-8 OLI over Lake Carnegie, WA, acquired on 8 July 2017 and displayed with bands 6, 5, 3 as RGB.
Source: Norman Mueller, Geoscience Australia

Table 7.1 Bi-temporal change detection stages

Stage	Description	Examples of Options
Image Selection	Select most appropriate image pair to highlight changes of interest	Anniversary dates
		Dates most relevant to change of interest
Pre-processing	Match geometries and radiance values of both images to a common base	Radiometric correction
		Geometric correction
		Spectral indices
Change Detection Method	Highlight pixels whose radiance values have changed between the two dates	Algebraic operations
		Linear transformations
		Classification approaches
Post-processing	Group changes into relevant categories	Thresholding to define change/no change
		Filtering to remove speckle
		Density slicing to define change classes
Labelling	Associate change categories with surface feature changes	Visual analysis
		Sample pixels
		GIS comparison
Accuracy Assessment	Determine accuracy of labelled change categories	Contingency table
		Kappa statistic

Adapted from Thonfeld *et al.* (2010)

For detected changes to be relevant to a particular purpose, it is paramount that an appropriate image pair is selected for analysis. Factors that impact this selection are considered in Section 2.2. Change detection approaches implicitly assume that the selected pair of images are comparable, and this generally involves several pre-processing steps as outlined in Sections 3.1 and 4 above.

Numerous authors have reviewed the wide range of image analysis methods that have been used to detect changes in EO imagery (including Lu *et al.*, 2003; Coppin *et al.*, 2004; Thonfeld *et al.*, 2010). In this section, we consider bi-temporal change detection methods in terms of three broad categories:

- algebraic operations between corresponding channels from each image—such as differencing or ratioing, (see Section 7.1);
- linear transformation of selected (or all) image channels from the two images—such as PCA, to highlight changes in fewer image dimensions (see Section 7.2); and
- classification of the two images, then per-pixel comparison of the resulting class allocations (see Section 7.3).

Various reviews of these methods suggest that no single approach can be recommended, but that different methods are appropriate for different cases (Lu *et al.*, 2003). As discussed in the introduction to Section 2, the analyst's underlying view of the landscape becomes evident when a change detection approach is selected: is the landscape (and the bi-temporal dataset(s) that observed it) an equilibrium state that changes or a dynamic system that is being observed? These concepts are further considered in Volumes 2E and 3. Post-processing of change detection results is discussed in Section 7.4 and approaches to accuracy assessment are introduced in Section 7.5 (see also Volume 2E).

7.1 Algebraic Operations

After suitable pre-processing, corresponding pairs of channels in the 'initial state' and 'final state' images can be subtracted or ratioed to highlight differences (see Sections 7.1.1 and 7.1.2 respectively). These operations can be computed for individual spectral bands or for spectral indices that highlight the most significant information in each image (see Volumes 2C and 3A). Results from these analyses indicate the location and extent of change, but do not generate detailed change matrices. Other algebraic operations that can be used to detect changes include channel regression (see Section 7.1.3) and Change Vector Analysis (see Section 7.1.4).

7.1.1 Channel subtraction

As detailed in Section 5.2 the differences between corresponding channels from two calibrated and registered image channels can be computed as their difference:

'change' image channel x = 'final state' image channel x – 'initial state' channel x

where

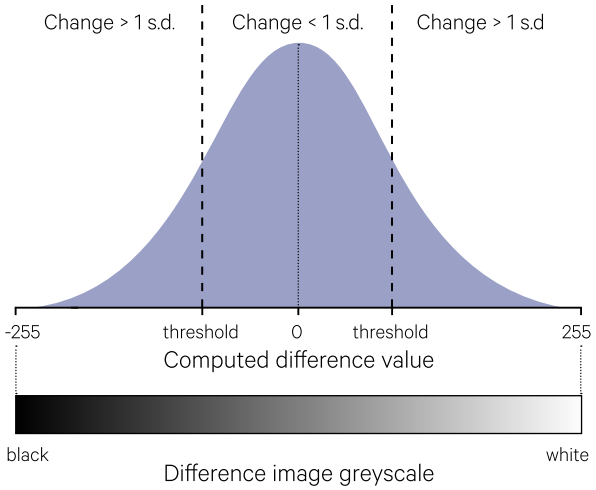
channel x represents the normalised reflectance measurements for a selected spectral attribute.

The resulting 'change' image is characterised by:

- those pixels whose reflectance values are brighter in the 'initial state' image forming one distribution peak (with values in 'change' image < 0); and
- those pixels whose reflectance values are brighter in the 'final state' image forming second distribution peak (with values in 'change' image > 0; see Figure 7.1).

Figure 7.1 Difference channel histogram

Change thresholds are equal to one standard deviation in this example. The data range of input images is assumed to be 0–255.

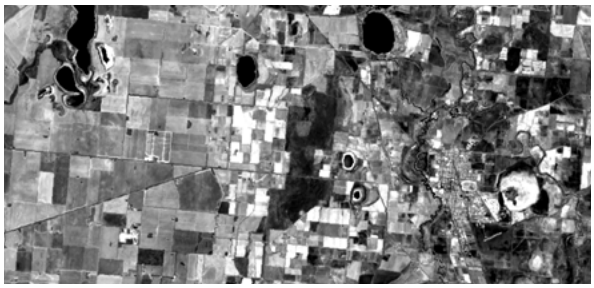


Before rescaling, those pixels with the same values in both images would have value zero in the 'change' image. An example of channel subtraction for a summer/winter pair of images in an agricultural district is shown in Figure 7.2.

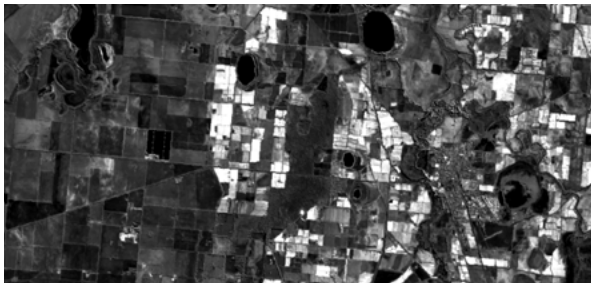
Figure 7.2 Change detection based on differencing

This example is based on a summer-winter pair of Landsat-8 OLI images over the mixed agricultural district of Kerang in northern Victoria. A difference band is computed by subtracting the NIR band in the winter image from the NIR band in the summer image. In the difference band, unchanged values appear as mid-grey, pixels with higher values in the summer image appear brighter and pixels with higher values in the winter image appear darker.

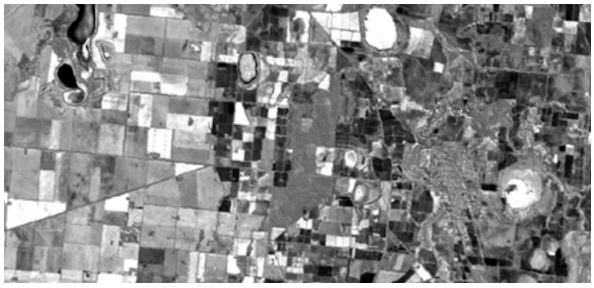
a. Summer NIR band (7 January 2018)



b. Winter NIR band (2 July 2018)



c. Difference of summer NIR band – winter NIR band



Source: Norman Mueller, Geoscience Australia

For example, to highlight changes associated with a fire event, the 'initial state' image would be acquired shortly before the fire and the 'final state' image would be acquired on an appropriate (smoke-free) date after the fire. If the image dates have been well selected, and no other major changes have occurred post-fire, the difference between these images would delineate the fire scar and thus provide information about the extent and severity of the fire.

Image differencing is computationally simple and offers a relatively straightforward summary of the distribution and extent of changes (see Section 5.1). Channel subtraction is the most commonly used method for bi-temporal change detection (Coppin *et al.*, 2004) and can be computed for individual spectral bands (albedo) or, more often, for derived spectral indices (Thonfeld *et al.*, 2010). Since the latter both emphasise the differences between relevant ground features and minimise the differences between images (see Volumes 2C and 3A), they are generally more efficient and robust for detecting image changes.

Various refinements of image differencing have been proposed to automate processing and/or to provide more robust results. For example, a moving window has been used to compute the minimum difference between a pixel in the 'initial state' image and the corresponding pixel values in its 3´3 neighbourhood in the 'final state' image, thereby compensating for imperfect geometric registration between the image pair (Castilla *et al.*, 2009).

7.1.2 Channel ratioing

Another relatively simple image processing option for comparing normalised image channels is to compute their ratio (see Volume 2C—Section 10):

$$\text{change image channel i} = \frac{\text{'final state' image channel i}}{\text{'initial state' image channel i}}$$

where

channel i represents normalised reflectance measurements for the same spectral attribute in both images.

This method produces a non-normal distribution so ratioed values are generally rescaled to permit further analyses (see Volume 2C—Section 2). Before the 'change' image results have been appropriate rescaled, its values will be:

- < 1 when reflectance values in the 'initial state' image are higher than in the 'final state' image;
- = 1 when reflectance values are the same in both images; and
- > 1 when reflectance values in the 'final state' image are higher than in the 'initial state' image (see Figure 7.3).

Channel ratioing reduces radiometric differences between images resulting from Sun position and topographic shading (see Volume 2C—Section 10). An example of channel ratioing for the same summer/ winter pair of images shown in Figure 7.2 is given in Figure 7.4.

Figure 7.3 Ratio channel histogram

This stylised histogram assumes input image data is in the range 0–255 and uses non-linear scaling to represent the computed ratio values.

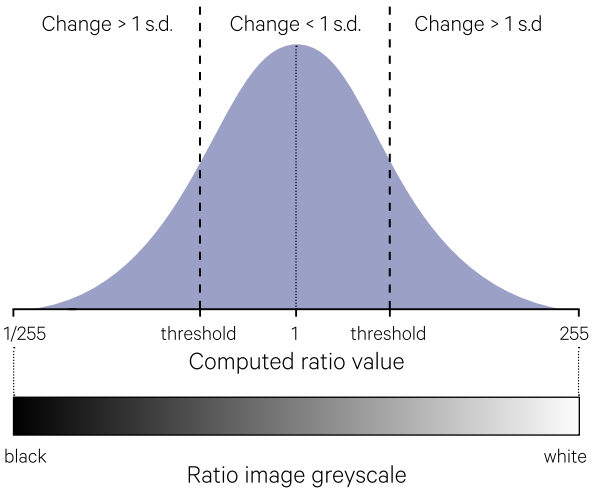
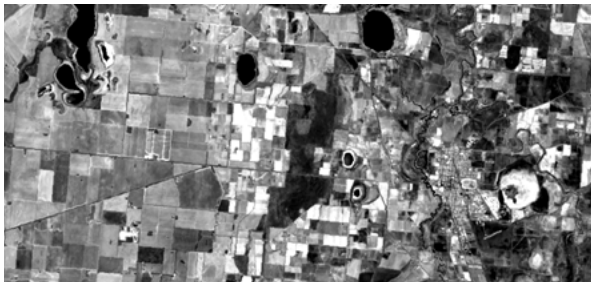


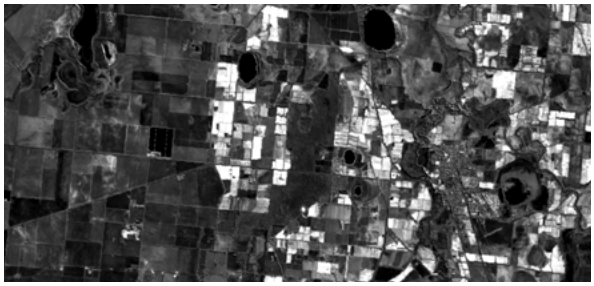
Figure 7.4 Change detection based on ratioing

This example is based on the same summer-winter pair of Landsat-8 OLI images over the mixed agricultural district of Kerang in northern Victoria that is used in Figure 7.2 above. To compute the ratio band, the NIR band in the summer image is divided by the NIR band in the winter image. In the ratio band, unchanged values appear as mid-grey, pixels with higher values in the summer image appear brighter and pixels with higher values in the winter image appear darker.

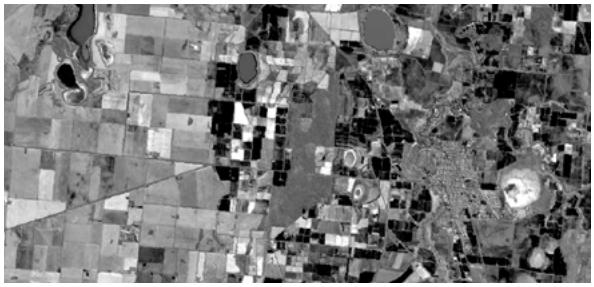
a. Summer NIR band (7 January 2018)



b. Winter NIR band (2 July 2018)



c. Difference of summer NIR band – winter NIR band



Source: Norman Mueller, Geoscience Australia

7.1.3 Channel regression

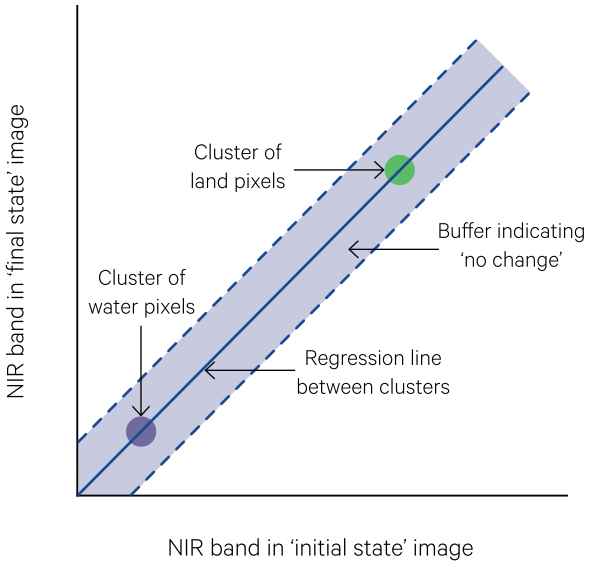
Regression techniques can be used to establish the relationship between pixels in the 'initial state' and 'final state' images for each pair of corresponding channels (see Volume 2A and Volume 2C—Section 7.2). In the context of change detection, the resulting regression relationship can be used to compute a theoretical 'final state' image from the 'initial state' image. The differences between the theoretical and actual 'final state' images highlight areas of change. This method also reduces radiometric differences between the two selected images resulting from atmospheric conditions, sensor variations and environmental factors.

A range of approaches can be used to determine appropriate regression parameters between channels including:

- Least squares regression (LS)—based on image statistics (see Volumes 2C and 2E);
- Pseudo-Invariant Feature (PIF)—based on the spectral values of features that are assumed to be invariant with time (Schott *et al.*, 1988);
- Radiometric Control Set (RCS)—based on the brightest and darkest features in the images (Hall *et al.*, 1991); and
- Automatic Scattergram-Controlled Regression (ASCR)—based on 'stable' clusters of land and water pixels (Elvidge *et al.*, 1995; Yang and Lo, 2000).

The ASCR method typically uses crossplots of NIR channels from the two images to identify stable clusters of land and water pixels then generates an initial regression line between the centres of the two clusters (see Figure 7.5). Thresholds can be placed either side of this regression line to demark 'no change', with adjacent regions representing positive and negative change.

Figure 7.5 Automatic Scattergram-Controlled Regression



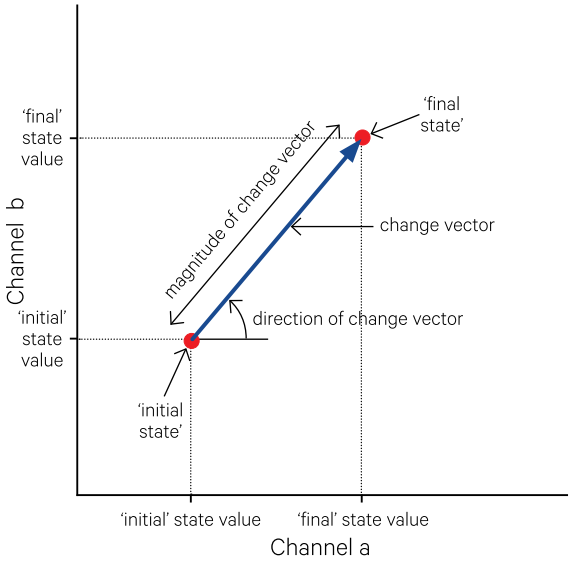
Adapted from: Yang and Lo (2000) Figure 1d

7.1.4 Change Vector Analysis

Change Vector Analysis (CVA) quantifies the per-pixel changes between a pair of bi-temporal images in terms of magnitude and direction (Malila, 1980). In the simplest implementation, two corresponding channels are selected from the 'initial state' and 'final' images then, for each pixel, the length and direction of the vector between these dates is determined (see Figure 7.6). These spectral change vector dimensions can also be computed for spectral indices and multi-channel datasets. The direction of the change vector is controlled by the type of change (see Figure 7.7), while the vector length is related to the severity of change (see Figure 7.8).

The greenness and brightness components from the Tasselled Cap transformation have typically been used with CVA to reduce data volume for processing (Kauth and Thomas, 1976; Crist and Cicone, 1984). In this transformation, brightness changes tend to be correlated with soil colour and moisture, while the greenness component tracks vegetation cover and vitality (see Volume 2C—Section 11). The total magnitude of change for each image pixel is computed as the Euclidean distance between the 'initial state' and 'final state' image values (see Figure 7.6 and Volume 2E). Changes in the direction of the brightness and greenness components can be grouped in quadrants to indicate four broad types of change (see Figure 7.7a). Possible land cover change classes corresponding to these four types of change direction are shown in Figure 7.7b. Similarly, changes in magnitude can be grouped into categories (see Figure 7.8a) and cross-tabulated against changes in direction to determine the extent of specific change classes (see Figure 7.8b). Thresholds can be defined to automatically determine the significance of detected changes (Chen *et al.*, 2003).

Figure 7.6 Change Vector Analysis (CVA)

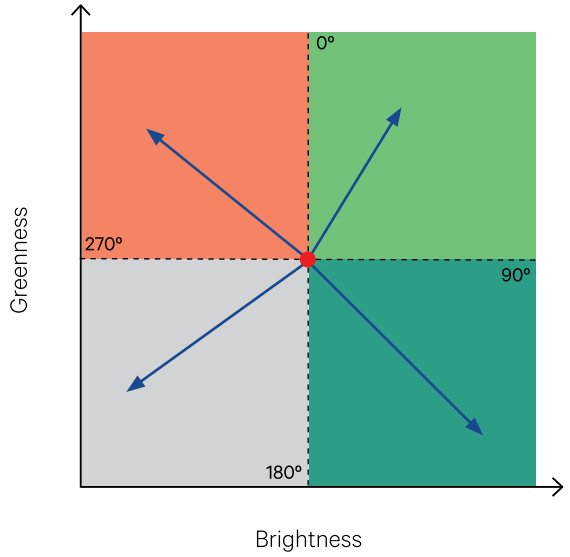


Adapted from Jensen (2016) Figure 12-35a

CVA has been used to delineate land cover changes (Lambin and Strahler, 1994a, 1994b; Kontoes, 2013), volcanic eruption damage (Kuzera *et al.*, 2005) and tsunami-related damage (Roemer *et al.*, 2010). Improved methods for computing CVA have been proposed by Chen *et al.* (2003), Carvalho *et al.* (2011) and Kontoes (2013).

Figure 7.7 Direction of change

a. Four potential directions of change between the 'initial state' image and the 'final state' image are shown using brightness and greenness axes. The 'initial state' values are shown as the central red dot and the change vector is shown as the blue arrow.



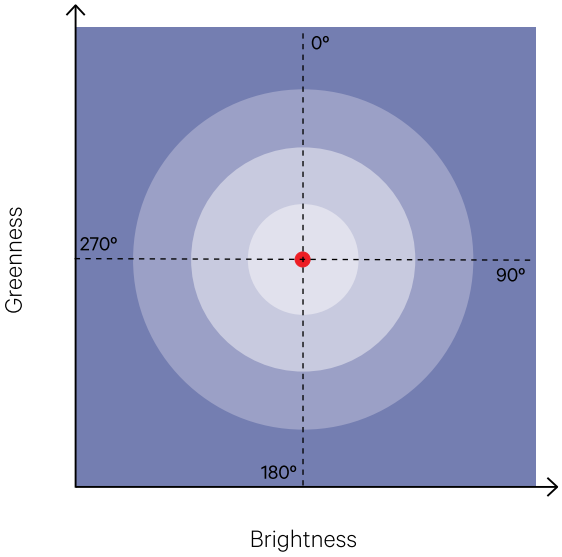
b. Interpretation of four potential change directions. These change classes were defined for a study to monitor land use/land cover dynamics in the Brazilian Amazon.

Change Angle (°)	Land Cover Change Class	Brightness Change	Greenness Change
0–89	Regeneration of woody vegetation	-	+
90–179	Deforestation	+	-
180–269	Conversion to pasture	+	+
270–359	Water or Fire Scar	-	-

Adapted from: a. Kuzera *et al.* (2005); b. Lorena *et al.* (2002)

Figure 7.8 Magnitude of change

a. The magnitude of change increases with distance from the 'initial stage' image values (shown as red dot below). Four magnitude categories (low, medium, high and extreme) are delineated in this example to indicate the change severity.



b. Land cover change classes defined in Figure 7.7a are cross-tabulated against the change severity categories delineated in Figure 7.8a to determine the extent of specific change classes.

Land Cover Change Class	Change Severity Categories			
	Low	Medium	High	Extreme
Regeneration of woody vegetation	*(area)*	*(area)*	*(area)*	*(area)*
Deforestation	*(area)*	*(area)*	*(area)*	*(area)*
Conversion to pasture	*(area)*	*(area)*	*(area)*	*(area)*
Water or Fire Scar	*(area)*	*(area)*	*(area)*	*(area)*

Adapted from Kuzera *et al.* (2005)

7.2 Linear Transformations

A wide range of linear transformations can be used to reduce data redundancy between image data channels, allowing relevant information to be represented in fewer channels (see Volume 2C—Sections 7, 8 and 9). Multiple images can also be compared using techniques such as:

- Principal Components Analysis (PCA; see Section 4.2.2 above and Volume 2C—Section 9);
- RGB to HSI Transformation (see Volume 2C—Section 8);
- Tasselled Cap or Kauth-Thomas Transformation (Kauth and Thomas, 1976; Crist and Cicone, 1984; see Volume 2C—Section 11);
- Multivariate Alternation Detection (MAD; Nielsen *et al*., 1998; see Volume 2C—Section 7.4); and
- Gram-Schmidt Orthogonalisation (Collins and Woodcock, 1994, 1996).

Such transformations may be applied separately to the individual images, then the resulting components compared, or applied to a merged dataset comprising channels from multiple images.

Principal Component Analysis (PCA) is reviewed in Section 4.2.2 above and detailed in Volumes 2C, 2E and 2X. The PCA transformation rotates the image data space to define uncorrelated channels, with the first component being aligned with the major axis of variance (see Figure 4.1 above). This results in a PC image in which covariance values are theoretically zero. The transformation is based on variance/covariance statistics from the image being processed, that is, its covariance or correlation matrix (Eklundh and Singh, 1993).

In terms of change detection, PCA can be applied to a merged set of geometrically-registered images. The rationale here is that the differences between the images will be highlighted in a small number of Principal Components (or PC channels). However, Fung and LeDrew (1987) suggest this approach has difficulties where major land cover changes have occurred since the transformations are not standardised between different images.

Alternatively, the two images (or selected channels thereof) are calibrated separately then combined into a single image. Transformations of the combined image are likely to highlight major changes as a single component (Richards, 1984; Ingebritsen and Lyon, 1985). Generally the higher order components of the merged image data are related to the brightness and greenness of surface features while the middle PCs are related to land cover changes. As with PCA of a single image, the lower PCs are insignificant in terms of the image variance they represent. Changes in specific features however may be summarised, and hence highlighted in a relative sense, in the lower PCs. An example of PCA applied to a merged two-date image is shown in Excursus 7.1.

Various authors have recommended the use of the image correlation matrix to produce standardised PCs (that is, with zero mean and unit variance; see Volume 2C—Section 9) since these minimise the effects due to atmospheric and illumination differences (Singh and Harrison, 1985; Elkundh and Singh, 1993; Fung and LeDrew, 1987). PCA for change detection should also be based on statistics derived from the whole image rather than apply only to a particular image feature (Fung and LeDrew, 1987).

PCA applied to a merged dataset is not the same as producing separate PC channels for each image then differencing them. Since PCA is necessarily scene-dependent, the components can represent different dimensions in different images. When the object of the exercise is to highlight changes, it is counter-productive to allow the pre-processing to mask differences. However, it is essential that image pairs be carefully registered geometrically (see Section 3.1.1.1) before applying PCA, as significant misregistration errors between a pair of images would increase the variation between channel pairs from different images and consequently inflate the covariance statistics.

Other linear transformations can be used to provide different representations of the image data. As detailed in Volume 2C—Section 11, the 'Tasselled Cap' transformation highlights stages of vegetation development and cover (Kauth and Thomas, 1976). MAD, based on canonical correlations analysis, sequentially extracts uncorrelated difference images to highlight changes (Nielsen and Conradsen, 1997). Unlike PCA, this transformation is invariant to linear and affine scaling, making it more relevant to change detection studies, especially when the MAF transformation is applied as post-processing (Nielsen *et al*., 1998; see Volume 2A—Section 7.4). The Gram-Schmidt Orthogonalisation produces multi-temporal analogues of brightness, greenness and wetness, plus a change component (Collins and Woodcock, 1994, 1996).

While capable of delivering good results, linear transformation methods are not readily transferable between datasets and tend to be used less frequently than arithmetic and classification methods for change detection (Thonfeld *et al*., 2010). Further, the results derived from these methods do not provide details of the type of change detected, and require thresholds to be defined to delineate the location and extent of changes.

Excursus 7.1—Change Detection using PCA

Source: Megan Lewis, University of Adelaide

Two Landsat MSS images of mallee vegetation, south of Renmark, SA, were selected to demonstrate change detection using PCA (see Figure 7.9). Land cover changes between these dates include scars from bushfires, clearing of native vegetation and changes in agricultural practices. The green, red and NIR bands from each image were combined into a single image then PCA was applied to the merged dataset.

The PCA matrix for the merged image indicates the contribution of each original band to each principal component (PC; see Table 7.2). Using the coefficients in the PCA matrix, we can deduce:

- PC1 represents the total brightness in both images as is typical of such analyses and essentially highlights the areas of no change;
- PC2 represents the differences between the corresponding bands in the two images so summarises all changes;
- PC4 represents the changes in contrast between vegetated areas in the Date 1 image and non-vegetated areas in the Date 2 image; and
- PC3, PC5 and PC6 represent less significant differences.

Table 7.2 PC transformation matrix

PC	Date 1 image bands			Date 2 image bands		
	Green	Red	NIR	Green	Red	NIR
1	0.390	0.602	0.469	0.207	0.346	0.321
2	0.218	0.355	0.301	-0.366	-0.622	-0.464
3	-0.309	-0.438	0.770	-0.096	-0.150	0.297
4	-0.116	-0.113	0.309	0.317	0.432	-0.046
5	0.825	--0.549	0.029	0.109	-0.059	-0.021
6	0.100	-0.070	0.016	-0.838	0.530	-0.046

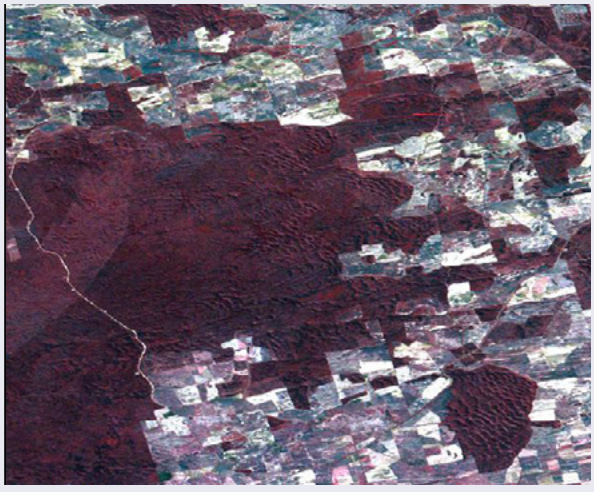
The resulting PC channels are shown in Figure 7.10. In this case, the most significant PCs for change detection are:

- PC2—which summarises the changes in vegetation cover; and
- PC4—which highlights the areas that have lost vegetation between the two image dates.

Figure 7.9 Original images

Landsat MSS images of mallee vegetation, south of Renmark, SA, are displayed using NIR, red and green bands as RGB. In these images dark red indicates mallee vegetation, while the large aqua region in the date 2 image highlights a fire scar.

a. Date 1 image



b. Date 2 image

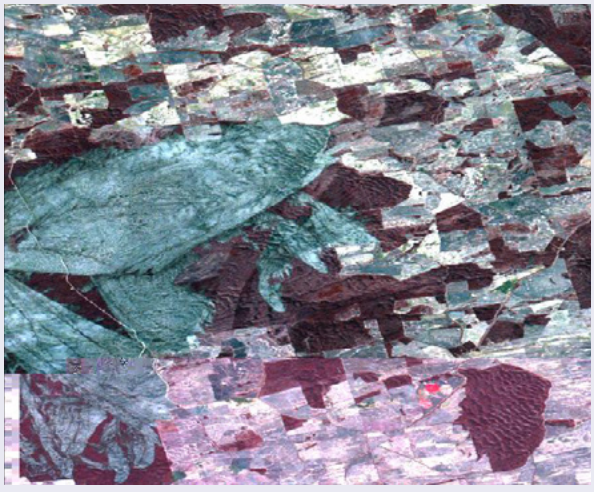
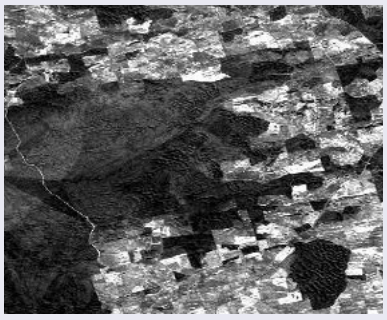


Figure 7.10 Principal components

Six principal component channels were derived for the merged image that contained the green, red and NIR bands of the images shown in Figure 7.9.

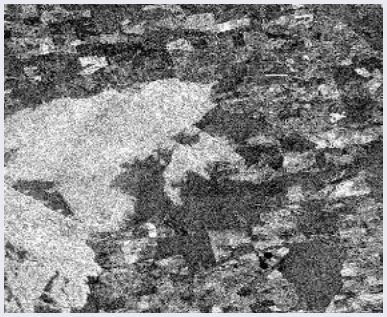
a. PC1



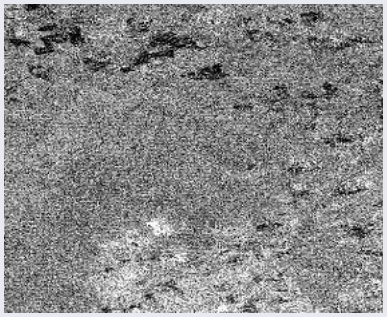
b. PC2

c. PC3

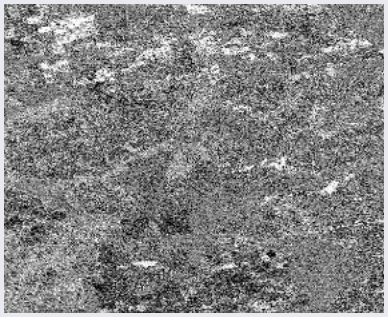
d. PC4



e. PC5



f. PC6



7.3 Classification Approaches

Classification can be used for change detection by:

- comparing separate classifications of the 'initial state' and 'final state' images (see Section 7.3.1);
- classifying a single composite image which contains corresponding channels from both the 'initial state' and 'final state' images (see Section 7.3.2); and
- classifying the differences between the bi-temporal images (see Section 7.1.1), which can either be based on the differences between corresponding pairs of original data bands, or between spectral indices derived from them (Sparks, 2005; see Volume 2C).

These approaches offer the opportunity to extract a complete matrix of change information as well as the location and area of each change class (Lu *et al.*, 2003). Data fusion methods have also been effective for integrating land cover change information from multiple classified images (Kiiveri and Campbell, 1992; Kiiveri and Caccetta, 1998 Kiiveri *et al.*, 2001; see Volume 2E). Image classification methods for EO imagery are summarised in Volume 2A—Section 9 and detailed in Volume 2E.

7.3.1 Comparing classifications

Rather than compare the reflectance values of the 'initial state' and 'final state' images directly, the changes between images can also be identified by first classifying each image separately and then comparing the classifications. This approach avoids standardising the two images to common radiometric values (Song *et al.*, 2001) and can be tailored to ignore irrelevant, transient changes. However, any pixel-based comparison between two images requires that both images be registered to a common geometric grid (Section 3.1 and Volume 2B).

This approach generally involves developing independent sets of classes in the two images that represent the categories of importance to a particular study. The resulting separate classifications can then be cross-compared using the sampling and analysis techniques detailed in Volume 2E.

Images can also be stratified before classification to reduce the possibility of misclassification (Xu and Young, 1990; see Section 4.1). Object-oriented methods distinguish bi-temporal changes for defined objects (often in conjunction with knowledge-based systems) rather than on a per-pixel basis (Hall and Hay, 2003; Desclée *et al*., 2006). The accuracy of any post-classification comparison exercise clearly depends on the accuracies of the two classifications being compared, and such comparisons invariably compound misclassification and misregistration errors.

Post-classification comparisons can produce summary statistics on the proportion of different land cover types that had changed between image dates (see Volume 2E). As described in Section 4.1.2, the classifications can be segmented using non-image boundaries (such as cadasta) to derive change statistics on the basis of some pre-defined regions. This method of change detection obviates the need for radiometric registration since each image is processed and interpreted individually. Provided sample points can be accurately located in each image (possibly using individual map-to-image registration models), geometric registration is not required either (see Volume 2B). However, two classifications and the comparison stage are time-consuming in terms of both computation and interpretation.

Singh (1989) considered post-classification comparison to be less accurate than difference or PC-based change detection methods and only slightly better than multi-date classification. A major problem with this approach is that the differences between classifications can be due to interpreter inconsistencies as well as image changes. To ensure compatibility between classifications, processing methodologies and labelling systems need to be as consistent as possible. Inaccuracies in either of the original classifications can also show as false changes in the post-classification comparison. Some authors have argued that the accuracy of change-mapping based on two classifications could equal the product of the accuracies of each classification (for example, if the original classifications were each deemed to be 80% accurate, the change analysis would be 64% accurate; Stow *et al*. 1980), though this will depend on the degree of spatial association between misclassified pixels in each image.

7.3.2 Classifying merged imagery

Some change detection studies have focussed on analysis of a merged image containing selected channels from the 'initial state' and 'final state' images, after appropriate pre-processing (see Section 3). Such analyses essentially assume that the statistics of changed patches of pixels will be significantly different from those that have not changed. To reduce the computational load of classifying a large number of channels, a subset or transformation, of channels is often used (see Section 4.2). Visual display of relevant channels from composite images can highlight prominent changes and isolate the most informative channels for subsequent analyses (see Volume 2A).

Classification methods based on this approach include 'multi-date clustering', 'spectral change pattern analysis' and 'spectral/temporal change classification'. The inherent heterogeneity of composite datasets is not well suited to segmentation and classification techniques, and the resulting classes have been difficult to label meaningfully (Coppin *et al*., 2004). Multi-temporal Spectral Mixture Analysis (MSMA) methods, however, have been reported as useful for detecting land cover changes (Adams *et al*., 1995; Roberts *et al*., 1998). Some other classification methods that have been used for bi-temporal change detection include artificial neural networks (Gopal and Woodcock, 1994), unsupervised change detection (Hame *et al*, 1998), and the Expectation-Maximisation algorithm (EM detection; Granville and Rasson, 1995; Bruzzone *et al*., 1999). These methods are reported to attain good accuracy, but involve greater processing complexity (see Volume 2E).

Given the diversity of possible land cover changes however, a large number of classes may be formed from multi-date classification. For example, in a given land cover type, pixels which have not changed should form one class and pixels which have changed will form one or more other classes depending on the types of change possible for that land cover. To avoid a complex interpretation task after classification, knowledge-based classifiers have been used to guide detection of class changes between images (see Volume 2E). Such solutions can be computationally intensive, however, and require prior knowledge of possible class changes. Singh (1989) observed that few authors reported accuracy testing of change detection methods, however his results showed multi-date classification to be the least accurate method of those described in this section.

7.4 Post-processing

Change detection results generally require some form of post-processing to assist interpretation. Commonly used methods include thresholding and labelling (see Section 7.4.1), filtering (see Section 7.4.2) and autocorrelation analysis (see Section 7.4.3).

7.4.1 Thresholding and labelling

Algebraic methods and linear transformations result in images that show regions of change and no change. The separation between these regions, however, requires thresholds to be defined, either by interactive testing (see Volume 2A—Section 9) or statistical measures (see Volume 2A—Section 8). Some thresholding techniques have also been proposed to determine thresholds automatically, based on attaining a pre-defined goal such as minimum noise (Bruzzone and Fernández Prieto, 2000a), minimum error (Bruzzone and Fernández Prieto, 2000b) or minimum 'cost' (Bruzzone and Fernández Prieto, 2000c). Obviously the accuracy of thresholds is as important as the accuracy of the change detection method in determining the value of the final product (Fung and LeDrew, 1988).

Once the thresholds have been selected to identify the areas of change and no change, some form of change labelling may be relevant to differentiate the change area in terms of change direction (positive or negative) and change severity (see Section 7.1.4). This is generally undertaken with reference to other datasets and can be implemented using density slicing (see Volume 2A—Section 9.2.1) or parallelepiped classification (see Volume 2A—Section 9.2.2). Image classification methods generally include a labelling stage, so implicitly involve this post-processing step as part of their processing sequence (see Volume 2E for details).

7.4.2 Filtering

A further post-processing method that may be relevant to some classified imagery is filtering to remove image speckle. Areas of change in difference images, for example, may be rendered more coherent using a modal filter. This technique is detailed in Volume 2C—Section 4.3 and Volume 2E. An image showing categorical changes could also be filtered using a modal, minimum or maximum filter, as appropriate for the data being processed (see Volume 2C—Sections 3 and 4).

7.4.3 Autocorrelation analysis

The Maximum Autocorrelation Factor (MAF) transformation can be used to highlight image regions where neighbouring pixels have similar values, or high spatial autocorrelation (see Volume 2C—Section 7.3). This process can be applied to difference images to highlight contiguous regions of change. For example, Nielsen *et al.* (1998) combined MAD and MAF transformations to retain spatial context in multi-temporal image analyses, and thus identify areas of spatially coherent change.

7.5 Accuracy Assessment

The accuracy of any change detection exercise needs to be determined before the results are used. As with any classification-style product, a range of statistical methods can be applied to verify the accuracy of the change categories.

Fundamentally, all accuracy assessment approaches compare the results from EO analyses with some reference dataset that is considered representative of the 'truth'. In terms of Earth surface features, ground 'truth' is generally derived from field sampling and/or visual analysis of higher resolution imagery. Relevant ground truth data for change detection implicitly requires two reference datasets: one to represent the 'initial state' image, and another to represent the 'final state' image. The differences between these two reference datasets are then assumed to indicate the type and extent of change in surface features.

Most accuracy assessment statistics are derived from an error matrix, which summarises the agreement between the EO results and the reference dataset for a set of sample pixels (Congalton and Plourde, 2002). For change detection studies, the error matrix tabulates the frequency with which sample pixels in the change image agree with the change indicated by the differences between the two reference datasets (Biging *et al.*, 1998, 1999). The most commonly derived statistics include overall accuracy, producer's accuracy, user's accuracy and Kappa coefficient (Congalton *et al.*, 1983; Hudson and Ramm, 1987; Congalton, 1991; Congalton and Green, 1999). Assessment methods tailored to change detection have also been proposed by Morisette and Khorram (2000) and Lowell (2001). Olofsson *et al.* (2014) recommend a set of good practice guidelines for assessing accuracy in change detection studies and estimating area of change. Accuracy assessment of EO analyses is considered in detail in Volume 2E.

7.6 Further Information

Land Cover Change:

CORINNE Land Cover (CLC): European Environment Agency (EEA) manages a hierarchical land cover inventory of Europe. This inventory defines 44 classes from EO imagery using a minimum mapping unit (MMU) of 25 ha. The original 1985 inventory (referenced as 1990) has been updated in 2000, 2006, 2012, and 2018, and is used for environmental modelling, agricultural monitoring and planning community infrastructure. Datasets and technical details can be downloaded from: *https://land.copernicus.eu/pan-european/corine-land-cover*

Global Vegetation Change:

NASA Land, Atmosphere Near real-time Capability for EOS (LANCE): *https://earthdata.nasa.gov/earth-observation-data/near-real-time*

7.7 References

Adams, J.B., Sabol, D., Kapos, V., Almeida Filho, R., Roberts, D.A., Smith, M.O., and Gillespie, A.R. (1995). Classification of Multispectral Images Based on Fractions Endmembers: Application to Land-Cover Change in the Brazilian Amazon. *Remote Sensing of Environment,* 52, 137–154.

Biging, G.S., Colby, D.R., and Congalton, R.G. (1998). Sampling systems for change detection accuracy assessment. In *Remote Sensing Change Detection—Environmental Monitoring Methods and Applications*. (Eds: Lunetta, R.S., and Elvidge, C.D.). Ann Arbor Press, Chelsea, MI. pp. 281–308.

Biging, G.S. Chrisman, N.R. Colby, D.R. Congalton, R.G. Dobson, J.E. Ferguson, R.L. Goodchild, M.F. Jensen, J.R. and Mace, T.H. (1999). *Accuracy assessment of remote sensing-detected change detection* Monograph Series. (Ed: Khorram, S.). American Society for Photogrammetry and Remote Sensing (ASPRS), Bethesda, MD, USA.

Bruzzone, L., and Fernandez Prieto, D. (2000a). An adaptive parcel-based technique for unsupervised change detection. *International Journal of Remote Sensing*, 21(4), 817–822, *https://doi.org/10.1080/014311600210614*

Bruzzone, L., and Fernández Prieto, D. (2000b). Automatic analysis of the difference image for unsupervised change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 38(3), 1171–1182.

Bruzzone, L., and Fernández Prieto, D. (2000c). A minimum-cost thresholding technique for unsupervised change detection. *International Journal of Remote Sensing*, 21, 3539–3544.

Bruzzone, L., Fernández Prieto, D., and Serpico, S.B. (1999). A Neural-Statistical Approach to Multitemporal and Multisource Remote-Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1350–1359.

Carvalho, O.A., Guimaraes, R.F., Gillespie, A.R., Silva, N.C., and Gomes, R.A.T. (2011). A new approach to Change Vector Analysis using distance and similarity measures. *Remote Sensing*, 3, 2473–2493.

Castilla, G., Guthrie, R.H., and Hay, G.J. (2009). The Land-cover Change Mapper (LCM) and its Application to Timber Harvest Monitoring in Western Canada. *Photogrammetric Engineering and Remote Sensing*, 75(8), 941–950.

Chen, J., Gong, P., He, C., Pu, R., and Shi, P. (2003). Land-Use/Land-Cover Change Detection using Improved Change Vector Analysis. *Photogrammetric Engineering and Remote Sensing*, 69(4), 369–379.

Collins, J.B., and Woodcock, C.E. (1994). Change detection using the Gramm–Schmidt transformation applied to mapping forest mortality. *Remote Sensing of Environment*, 50, 267–279.

Collins, J.B., and Woodcock, C.E. (1996). An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data. *Remote Sensing of Environment*, 56, 66–77.

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37, 35–46.

Congalton, R., and Green, K. (1999). Assessing the accuracy of remotely sensed data: Principles and practices. Lewis Publishers, New York.

Congalton, R.G., Oderwald, R.G. and Mead, R.A. (1983). Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing,* 49, 1671–1678. [see also Erratum: *Photogrammetric Engineering and Remote Sensing,* 50, 1477.]

Congalton, R.G., and Plourde, L. (2002). Quality assurance and accuracy assessment of information derived from remotely sensed data In *Manual of Geospatial Science and Technology.* (Ed: Bossler, J.) Taylor and Francis, London. pp. 349–361.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing*, 25(9), 1565–1596.

Crist, E.P., and Cicone, R.C. (1984). Application of the Tasselled Cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing,* 50, 343–52.

Desclée, B., Bogaert, P., and Defourny, P. (2006). Forest change detection by statistical object-based method. *Remote Sensing of Environment*, 102, 1–11.

Eklundh, L., and Singh, A. (1993). A comparative analysis of standardized and unstandardized principal component analysis in remote sensing. *International Journal of Remote Sensing*, 14, 1359–1370.

Elvidge, C.D., Yuan, D., Werrackoon, R.D., and Lunetta, R.S. (1995). Relative radiometric normalisation of Landsat Multispectral Scanner (MSS) data using an automatic scattergram-controlled regression. *Photogrammetric Engineering and Remote Sensing*, 61(10), 1255–1260.

Engeset, R.V., Kohler, J., Melvold, K., and Lunden, B. (2002). Change detection and monitoring of glacier mass balance and facies using ERS SAR winter images over Svalbard. *International Journal of Remote Sensing*, 23, 2023–2050. *https://doi.org/10.1080/01431160110075550*

Fung, T., and LeDrew, E. (1987). Application of PCA for change detection. *Photogrammetric Engineering and Remote Sensing*, 53, 1649–1658.

Fung, T., and Ledrew, E. (1988). The determination of optimal threshold levels for change detection using various accuracy indices. *Photogrammetric Engineering and Remote Sensing*, 54, 1449–1454.

Garcia-Haro, F.J, Gilabert, M.A., and Melia, J. (2001). Monitoring fire-affected areas using Thematic Mapper data. *International Journal of Remote Sensing*, 22, 533–549.

Gopal, S., and Woodcock, C.E. (1994).Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing,* 60, 181–188.

Graetz, R.D, Pech, R.P and Davis, A.W. (1988). The assessment and monitoring of sparsely vegetated rangelands using calibrated Landsat data. *International Journal of Remote Sensing*, 8, 1201–1222.

Granville, V., and Rasson, J.P. (1995). Multivariate discriminant analysis and maximum penalized likelihood density estimation. *Journal of the Royal Statistical Society*, 57(3), 501–517.

Gupta, D.M., and Munshi, M.K. (1985). Urban change detection and land-use mapping of Delhi. *International Journal of Remote Sensing*, 6, 529–534.

Hall, O., and Hay, G.J. (2003). A Multiscale Object-Specific Approach to Digital Change Detection. *International Journal of Applied Earth Observation*, 4(4), 311–327. *https://doi.org/10.1016/S0303-2434(03)00010-2*

Hall, F.G., Strebel, D.E., Nickeson, J.E., and Goetz, S.J. (1991). Radiometric rectification: Towards a common radiometric response among multidate, multisensor images. *Remote Sensing of Environment*, 35, 11–27.

Hame, T., Heiler, I., and San Miguel-Ayanz, J. (1998). An unsupervised change detection and recognition system for forestry. *International Journal of Remote Sensing*, 19(6), 1079–1099, *https://doi.org/10.1080/014311698215612*

Henebry, G.M. (1993). Detecting change in grasslands using measures of spatial dependence with Landsat TM data. *Remote Sensing of Environment*, 46(2), 223–234. *https://doi.org/10.1016/0034-4257(93)90097-H*

Houhoulis, P.F., and Michener, W.K. (2000). Detecting wetland change: a rule-based approach using NWI and SPOT-XS data. *Photogrammetric Engineering and Remote Sensing*, 66, 205–211.

Hudson, W.D., and Ramm, C.W. (1987). Correct formulation of the Kappa coefficient of agreement. *Photogrammetric Engineering and Remote Sensing,* 53, 421–422.

Ingebritson, S.E., and Lyon, R.J.P. (1985). Principal components analysis of multitemporal image pairs. *International Journal of Remote Sensing*, 6, 687–696.

Jakubauskas, M.E, Kamlesh, P.L., and Mausel, P.W. (1990). Assessment of vegetation change in a fire altered forest landscape. *Photogrammetric Engineering and Remote Sensing*, 56, 371–377.

Jensen, J.R. (Ed). (1983). Urban/suburban land use analysis. In *Manual of Remote Sensing. Vol. 2*. 2nd Edn. American Society of Photogrammetry, pp. 1571–1666.

Jensen, J.R. (2016). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th Edn. Pearson Education, Inc. ISBN 978-0-13-405816-0

Kauth, R.J., and Thomas, G.S. (1976). The Tasselled Cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*. Purdue University, West Lafayette, Indiana, 4B41–4B51.

Kiiveri, H., and Caccetta, P. (1998). Image Fusion with Conditional Probability Networks for Monitoring the Salinization of Farmland. *Digital Signal Processing*, 8(4), 225–230. *https://doi.org/10.1006/dspr.1998.0320*

Kliveri, H.T., and Campbell, N.A. (1992). Allocation of Remotely Sensed Data using Markov Models for Image Data and Pixel Labels. *Australian Journal of Statistics*, 34(3). *https://doi.org/10.1111/j.1467-842X.1992.tb01053.x*

Kiiveri, H., Caccetta, P., and Evans, F. (2001). Use of conditional probability networks for environmental monitoring. *International Journal of Remote Sensing*, 22(7), 1173–1190. *https://doi.org/10.1080/01431160151144305*

Kimura, H., and Yamaguchi, Y. (2000). Detection of landslide areas using satellite radar interferometry. *Photogrammetric Engineering and Remote Sensing*, 66, 337–343.

Kontoes, C.C. (2013). Operational Land Cover Change Detection using Change Vector Analysis. *International Journal of Remote Sensing*, 29(16), 4757–4779.

Kuzera, K., Rogan, J., and Eastman, R. (2005). Monitoring vegetation regeneration and deforestation using Change Vector Analysis: Mt. St. Helens Study Area. *ASPRS Annual Conference*, 7–11 March, Baltimore, Maryland.

Lambin, E.F., and Strahler, A.H. (1994a). Change-vector analysis in multitemporal space: a tool to detect and categorize land-cover change processes using high temporal resolution satellite data. *Remote Sensing of Environment*, 48, 231–244.

Lambin, E.F., and Strahler, A.H. (1994b). Indicators of land-cover change for change-vector analysis in multitemporal space at coarse spatial scales. *International Journal of Remote Sensing*, 15, 2099–2119.

Lorena, R.B., dos Santos, J. R., Shimabukuro, Y.E., Brown, I.F., and Kux, H.J.H. (2002). A change vector analysis techniques to monitor land use/land cover in SW Brazilian Amazon: ACRE state. In *PECORA–15 Integrating Remote Sensing at the Global, Regional and Local Scale*, Denver, Colorado, USA.

Lowell, K. (2001). An area-based accuracy assessment methodology for digital change maps. *International Journal of Remote Sensing*, 22, 3571–3596.

Lu, D., Mausel, P., Brondízio, E., and Moran, E. (2003). Change detection techniques. *International Journal of Remote Sensing*, 25(12), 2365–2407. *https://doi.org/10.1080/0143116031000139863*

Macomber, S.A., and Woodcock, C.E. (1994). Mapping and monitoring conifer mortality using remote sensing in Lake Tahoe Basin. *Remote Sensing of Environment*, 50(3), 255–266.

Malila, W.A. (1980). Change Vector Analysis: An Approach for Detecting Forest Changes with Landsat. *LARS Symposia*. Paper 385. *http://docs.lib.purdue.edu/lars_symp/385*

Manavalan, P., Kesavasamy, K., and Adiga, S (1995). Irrigated crops monitoring through seasons using digital change detection analysis of IRS-LISS 2 data. *International Journal of Remote Sensing*, 16(4), 633–340. *https://doi.org/10.1080/01431169508954430*

Morisette, J.T., and Khorram, S. (2000). Accuracy assessment curves for satellite-based change detection. *Photogrammetric Engineering and Remote Sensing*, 66(7), 875–880.

Muchoney, D.M., and Haack, B.N. (1994). Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, 60(10), 1243–1251.

Munyati, C. (2000). Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International Journal of Remote Sensing*, 21(9), 1787–1806, *https://doi.org/10.1080/014311600209742*

Nielsen, A.A., Conradsen, K., and Simpson. (1998). Multivariate Alternation Detection (MAD) and MAF Postprocessing in Multispectral Bitemporal Image Data: New Approaches to Change Detection Studies. *Remote Sensing of Environment*, 64, 1–19.

Nielsen, A.A., and Conradsen, K. (1997). *Multivariate alteration detection (MAD) in multispectral, bitemporal image data: a new approach to change detection studies*. Tech. Rep. 1997-11, IMM, Technical University of Denmark, Lynby.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., and Wulder, M.A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42–57.

Peters, A.J., Walter-Shea, E.A., Lei, J., Vina, A., Hayes, M., and Svoboda, M.D. (2002). Drought monitoring with NDVI-based standardized vegetation index. *Photogrammetric Engineering and Remote Sensing*, 68, 71–76.

Richards, J.A. (1984). Thematic mapping from multitemporal image data using the PC transformation. *Remote Sensing of Environment,* 16, 35-46.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., and Green, R.O. (1998). Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, 65, 267–279.

Roemer, H., Kaiser, G., Sterr, H., and Ludwig, R. (2010). Using remote sensing to assess tsunami-induced impacts on coastal forest ecosystems at the Andaman Sea coast of Thailand. Natural Hazards Earth Systems Science, 10, 729–745, *https://doi.org/10.5194/nhess-10-729-2010*

Schott, J.R., Salvaggio, C., and Volchok, W.J. (1988). Radiometric scene normalisation using pseudovarient features. *Remote Sensing of Environment*, 26, 1–6.

Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10, 989–1003.

Singh, A., and Harrison, A. (1985). Standardised Principal Components. *International Journal of Remote Sensing*, 6, 883–896.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., and Macomber, S.A. (2001). Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? *Remote Sensing of Environment,* 75, 230–244.

Sparks, T. (2005). Cotter Catchment Fire Recovery Mapping. *Proceedings of the SSC2005 Spatial Intelligence, Innovation and Praxis: The national biennial conference of the Spatial Sciences Institute,* Melbourne. ISBN 0-9581366-2-9

Stow, D.A., Tinney, L.R., and Estes, J.E. (1980). Revising land use/land cover change statistics from Landsat: A study of prime agricultural land. *Proceedings of 14th International Symposium on Remote Sensing of Environment*, ERIM, Ann Arbor, Michigan. pp 1227–1237.

Thonfeld, F., Hecheltjen, A., Braun, M., and Menz, G. (2010). From Algorithms to Processing Chains: A Review of Land Cover and Land Use Change Detection Methodologies. *Proceedings ESA Living Planet Symposium*. ESA SP-686. ESA, Bergen, Norway.

Vogelmann, J.E. (1988). Detection of forest change in the Green Mountains of Vermont using multi-spectral scanner data. *International Journal of Remote Sensing*, 9, 1187–1200.

Vogelmann, J.E. (1990). Comparison between two vegetation indices for measuring different types of forest damage in northeast United States. *International Journal of Remote Sensing*, 11, 2281–2297.

Vogelmann, J.E., and Rock, B.N. (1988). Assessing forest damage in high elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment,* 24, 227–246.

Xu, H., and Young, J.A. (1990). Monitoring changes in land use through integration of remote sensing and GIS. *Proceedings IGARSS'90 Symposium*, University of Maryland, College Park, Maryland, pp. 957–960.

Yang, X., and Lo, C.P. (2000). Relative Radiometric Normalisation Performance for Change Detection from Multi-date Satellite Images. *Photogrammetric Engineering and Remote Sensing,* 66(8), 967–980.

Zhou, G., Luo, J., Yang, C., Li, B., and Wang, S. (2000). Flood monitoring using multitemporal AVHRR and RADARSET imagery. *Photogrammetric Engineering and Remote Sensing*, 66, 633–638.

Processing Time Series Datasets

In the context of EO data, a time series dataset refers to a set of calibrated images acquired at regular time intervals. Some of the commonly available EO time series datasets are listed in Section 2.1 above. EO-based time series datasets have been used to derive:

- various phenological parameters for a range of studies relating to ecosystem dynamics (see Volume 3A); and
- water-related metrics for inundation and marine condition studies (see Volume 3B).

Section 8 describes how time series datasets are derived from EO imagery. Some methods to condense time series datasets into composite images are described in Section 10, while Section 9 introduces some of the basic tools that are used to analyse image time series. The Digital Earth concept is described in Section 10, including the Australian implementation of Digital Earth Australia (DEA; see Section 11.2).

Contents

8 Time Series Analysis 87

9 Characterising Temporal Trends 101

10 Pixel-based Composites 111

11 Digital Earth 123

Background image on previous page: Annual Landsat-8 OLI pixel composite of Tasmania and southern Victoria for 2014 displayed using bands SWIR1, NIR and Green as RGB. **Source**: Norman Mueller, Geoscience Australia

8 Time Series Analysis

A time series comprises a sequence of consistent measurements repeated through time (see Volume 2A—Section 8.4). Such data indicate variations in the feature being measured, which may be useful to characterise its behaviour, highlight anomalies and forecast future states. Time series are commonly collected to systematically monitor the condition of various indicators. Examples of common time series datasets include economic variables, such as financial stock values, consumer spending or employment levels, and environmental variables, such as river flow, precipitation trends or atmospheric changes.

EO imagery offers a readily accessible form of time series data for monitoring changes in Earth surface features. The archived data from various satellite sensors are now available for several decades (see Table 2.2), providing a unique record of planet Earth. For a given sensor, the time series represents a sequence of consistent observations with near-global coverage. A pictorial representation of an image time series for Australia, for example, is illustrated in Figure 8.1. Such datasets enable Earth surface changes to be monitored through major variations in climatic conditions and significant changes in land use.

The temporal aspect of time series imagery greatly enhances the opportunity to differentiate Earth surface features, but also adds another dimension to data volume, processing and interpretation. Classification of single date imagery, for example, generally involves allocating pixels to classes based on their spectral similarity within a small number of image bands. Given the significant data volumes associated with EO time series data, a number of algorithms have been proposed to simplify the process of identifying and characterising major trends. Much effort has been directed into developing methods to produce robust time series from the voluminous archives of EO imagery (Viovy *et al.*, 1992; Roerink *et al.*, 2000; Jönsson and Eklundh, 2002; Zhang *et al.*, 2003; Chen *et al.*, 2004; Bradley *et al.*, 2007). More recently, Analysis Ready Data (ARD) have simplified the process of accessing image time series (see Section 3.2).

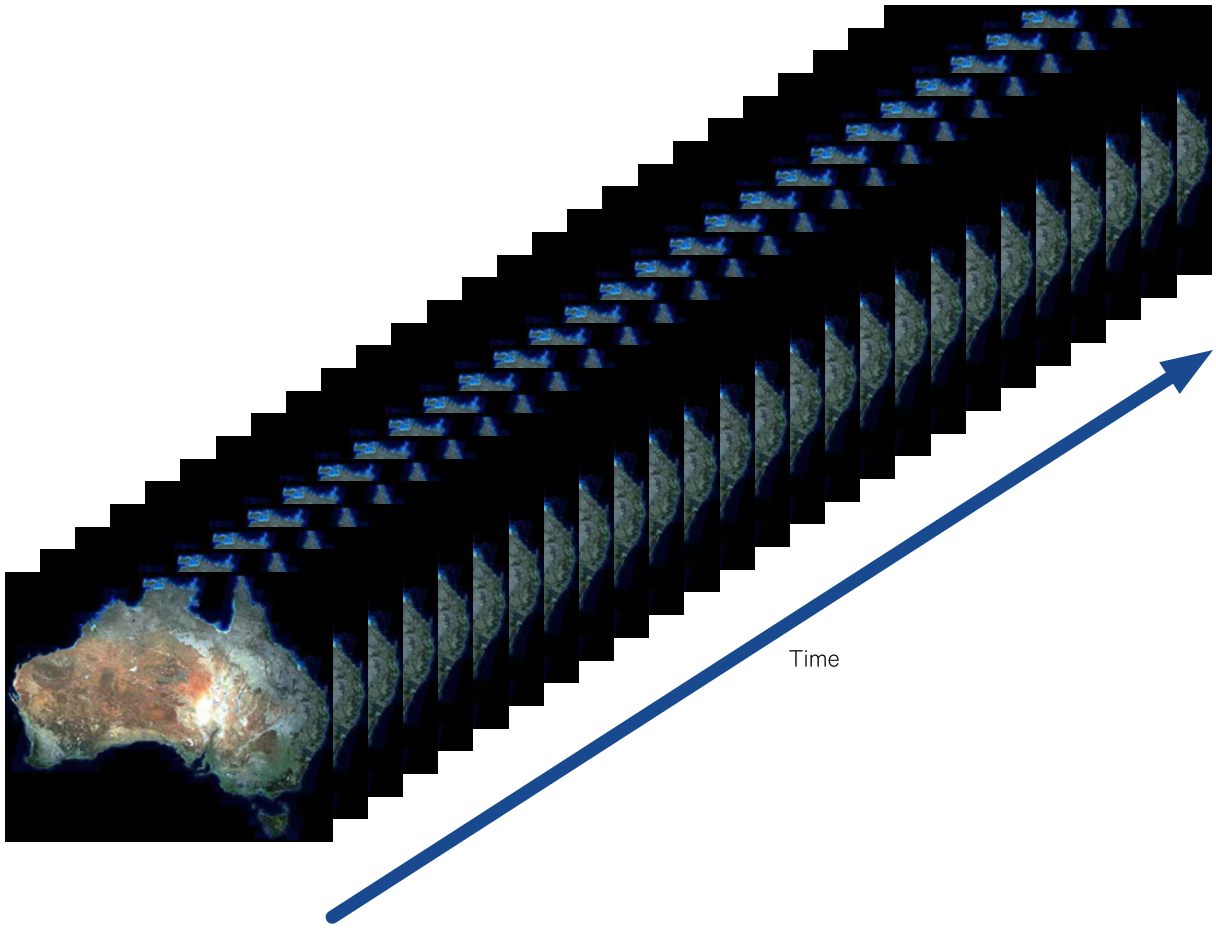
The basic characteristics of time series datasets are introduced in Volume 2A—Section 8.4. In the context of EO time series datasets, the following sub-sections consider:

- time series components and processing (see Section 8.1);
- the importance of understanding the dataset (see Section 8.2);
- methods to construct a data cube (see Section 8.3);
- detecting anomalies (see Section 8.4); and
- characterising data trends (see Section 9).

Background image: High resolution aerial image of Parliament House, Canberra., acquired on 25 April 2016. Source: © Nearmap.

Figure 8.1 Time series imagery of Australia

Successive EO images of Australia, acquired at regular intervals by the same sensor, provide a continental time series dataset.



8.1 Time Series Components and Processing

The variations within a time series can be separated into three components:

- long term trends—slow changes with time;
- seasonal trends—correlated with calendar patterns or diurnal cycles; and
- irregular, ephemeral trends—random noise (see Figure 8.2).

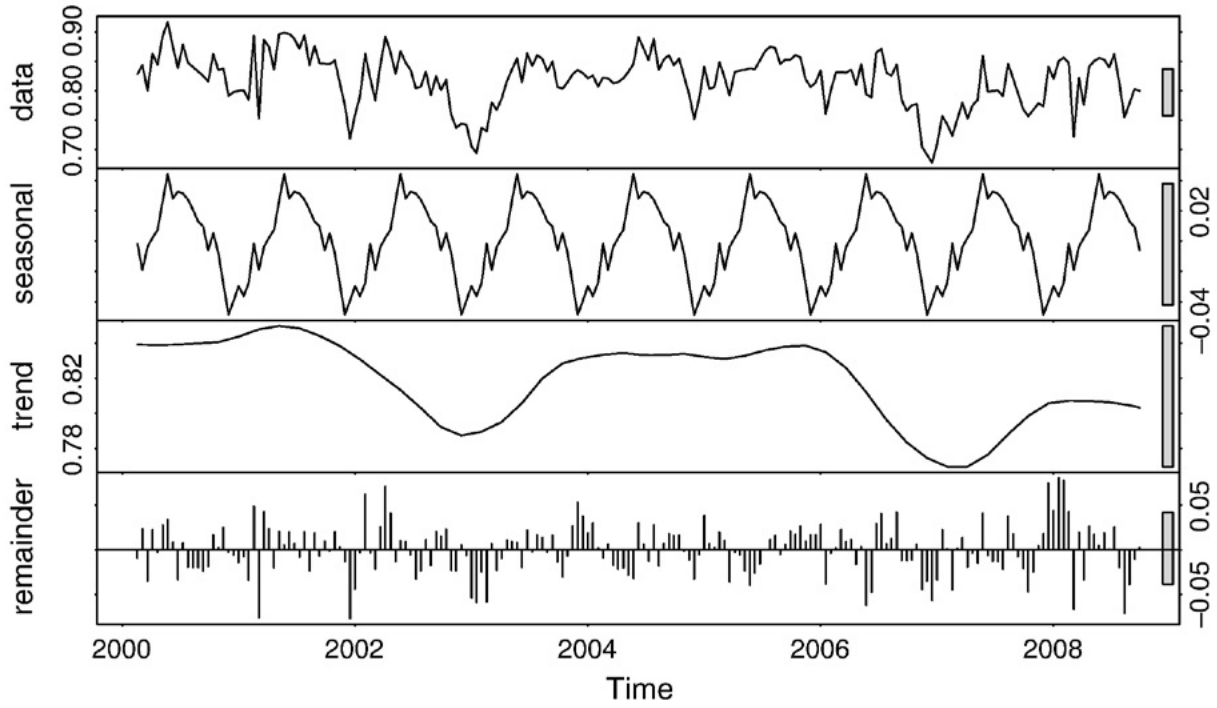
These components combine to create the total time series, but their separation (where possible) enables analysis of individual components. In terms of Earth surface features, these three components correspond to gradual, seasonal and abrupt changes. Seasonal trends, in particular, are generally more stable and contain valuable information on the timing, direction and magnitude of systematic changes. For many Earth surface features, seasonal trends are directly related to phenology (see Volume 3A). Examples of MODIS-derived fPAR (fraction of Photosynthetically Active Radiation; see Excursus 8.1) timeseries for selected TERN Ozflux sites (see Volume 3A), which have contrasting vegetation and precipitation characteristics, are shown in Figure 8.3.

Most pre-processing methods for time series analysis aim to reduce data volume and remove outliers, and many approaches modify the spatial and temporal image resolutions during this process. The analysis of time series data involves three conceptual processes (see Table 8.1):

- represent image archive as a standardised data cube, that is, a three-dimensional matrix in which the spatial and temporal separation between adjacent pixels is consistent (see Section 8.3);
- identify and replace outlying values that are likely to have resulted from imaging artefacts. This stage ensures that all pixel values are comparable over space and time (see Section 8.4); and
- characterise seasonal and annual trends within the prepared data cube (see Section 9).

Figure 8.2 Components of time series variations

A 16-day NDVI time series of a pine plantation is decomposed into seasonal, trend and irregular components. Right hand bars show comparable data ranges. Seasonal amplitude range is around 0.1 NDVI.



Source: Verbesselt *et al.* (2010) Figure 2

Table 8.1 Time series processing sequence

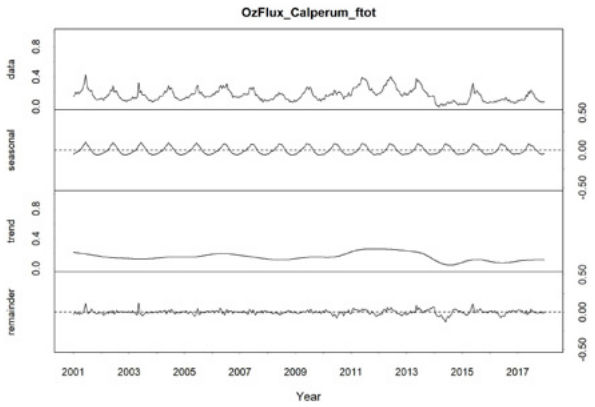
Stage	Options	Description
1. Create discrete data cube to reduce data volume	Reduce spectral density	Create spectral index
	Reduce temporal resolution	Create less frequent time steps between consecutive images by selecting 'best' pixel in time interval
	Reduce spatial resolution	Create larger pixel size
2. Identify and replace outliers	Compositing images acquired within selected time interval	Select 'best' pixel within time step
	Despiking of time series	Reset pixels with unusually high and low values to more likely levels using some filtering algorithm
3. Identify seasonal trends	Fit function to time series to model repeating patterns	See Table 9.2
	Break up time series into discrete stages based on rate of change	
	Analyse patterns within time series	

To steal ideas from one person is plagiarism but to steal from many is research.
(Wilson Mizner)

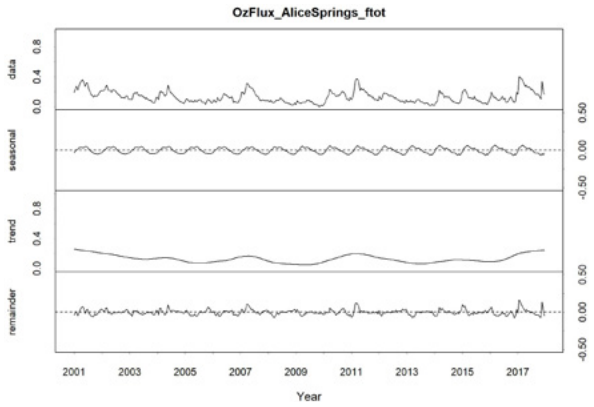
Figure 8.3 fPAR time series examples

These examples of MODIS-derived time series show fPAR (fraction of Photosynthetically Active Radiation) variations at selected TERN OzFlux sites.

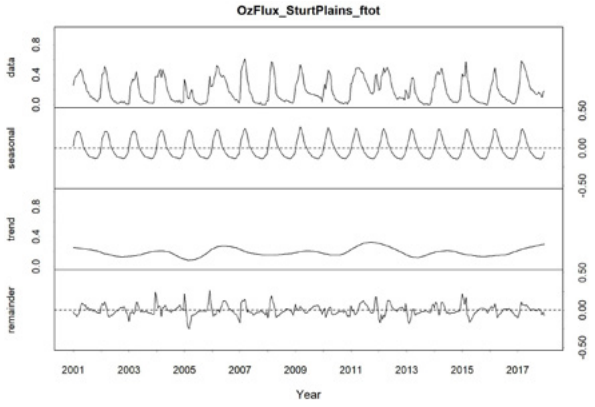
a. Recovering mallee woodland at Calperum Station near Renmark, South Australia (mean annual precipitation: 240 mm)



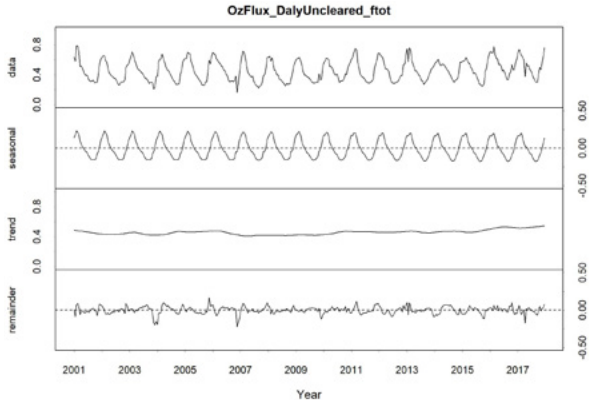
b. Semi-arid mulga at Pine Hill cattle station near Alice Springs, Northern Territory (mean annual rainfall: 306 mm)



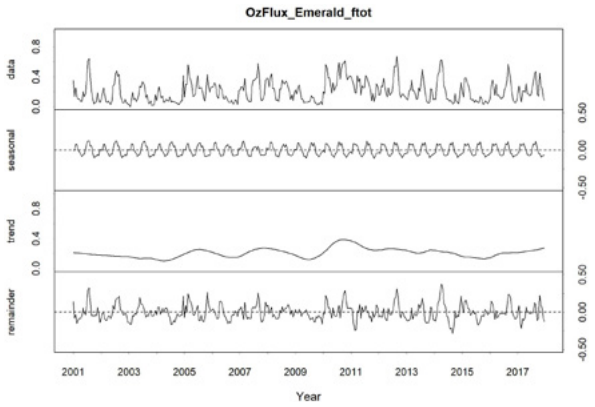
c. Predominantly Mitchell Grass in Sturt Plains, north of Tennant Creek, Northern Territory (mean annual precipitation: 640 mm)



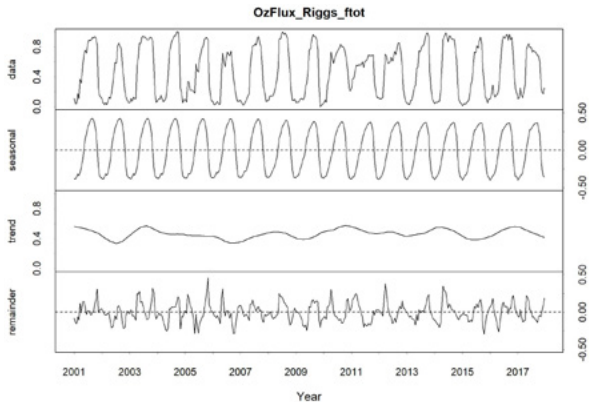
d. Uncleared woodland savanna in Douglas River Daly River Esplanade Conservation area, southwest of Pine Creek, Northern Territory (mean annual precipitation: 1170 mm)



e. Irrigated cropping and pasture at Arcturus, near Emerald, central Queensland (mean annual precipitation: 572 mm)



f. Dryland agriculture at Riggs Creek, near Shepparton, Victoria (mean annual precipitation: 650 mm)



Source: Glenn Newnham and Randall Donohue (CSIRO). For details of TERN OzFlux sites, see: *http://www.ozflux.org.au/monitoringsites/*

8.2 Understanding the Data

While EO time series datasets offer great potential for monitoring and modelling surface features, they also contain significant challenges. Before considering some of the methods used to analyse these datasets, it is worth considering the conceptual problems inherent in this potential.

Firstly, many time series analysis methods have been developed for one-dimensional sequences of 'independent' observations. EO observations, however, are not independent. While the spacing of observations can be manipulated in both spatial and temporal dimensions, the time sequence associated with each pixel will generally be correlated with, if not dependent upon, the time sequences of nearby pixels. Thus, correlations exist in both the spatial and temporal dimensions.

Secondly, most time series analyses aim to differentiate different types of surface features, but it is the cumulative reflectance of all surface features within each imaged pixel that is measured by a remote sensor. In many cases, the surface features have different temporal characteristics, but only their combined reflectance is recorded at each time step. The resulting time series then potentially represents multiple temporal patterns—one for each of the surface features, and these features may be most appropriately sampled at different intervals. For example, annual plants such as grasses follow a clear annual cycle of relatively rapid greening and browning whereas evergreen trees change colour less dramatically between seasons. The time series for a pixel that covers both open trees and grass will be some combination of both cycles, and separating the cycles for each component is not necessarily a simple task, especially when those surface features have not been identified.

Thirdly, remote sensors are not perfect. They can and do record anomalous pixels, their measurements are thwarted by adverse atmospheric conditions and Sun positions, and, over time, their calibration systems can change so that measurements are not consistent (see Volume 1). Many of these problems are well defined and can be solved adequately, but long-term sensor drift remains a vexatious issue for time series datasets. This concern is exacerbated by the second problem mentioned above.

Thus, the trends in EO time series datasets result from the interactions between multiple factors relating to the:

- observed reflectance of Earth surface features as impacted by atmospheric conditions (see Volume 1B—Section 4);
- data resolution in all imaging dimensions (see Table 2.1 above and Volume 1B—Section 1) ; and
- sensor calibration (see Volume 1A—Section 13, Volume 1B—Section 2 and Volume 2A—Section 3).

In addition to the usual concerns about whether the image resolution is suitable to map a particular surface feature, the sampling rate, in conjunction with the actual timing of image planes, needs to be appropriate to the observable reflectance changes in the surface features (see Section 2.2). The analysis of such datasets effectively involves multiple unknowns:

- What is the contribution of the long term trend?
- What is the contribution of the seasonal pattern?
- What is the contribution of the error component?
- What is the relationship between the long term trend, the seasonal pattern and the error component—is it additive, multiplicative, or more complex?
- Is the time interval between image planes appropriate for the surface features being mapped?

In essence, EO time series datasets provide multidimensional data that are indirectly related to multidimensional surface features (see Section 1). Appropriate analysis approaches will recognise the inherent limitations of the data and minimise the number of unknowns (see Section 2.3). Some studies of EO time series data have tended to view the image archive as a series of discrete observations, such that the temporal sequence of values for each geometric pixel have been pre-processed and analysed independently as pixel trajectories. Given the intrinsic interrelationships between adjacent pixels in any EO image, there is some justification to question whether it is appropriate to analyse such datasets this way, especially for real time monitoring and/or short-term forecasting applications (White and Nemani, 2006).

While the fundamental goal of time series analysis of EO datasets may be an improved understanding of ecosystem dynamics, a successful outcome requires some initial familiarity with the relevant ecosystem(s). For example, the extent of serial correlation is related to the temporal resolution of the time series, that is, the time interval between image layers (or time slices). However, the extent of actual correlation between consecutive observations in a particular study will depend on the dynamics of the relevant ecosystem (Zoffoli *et al*., 2008). In general, to characterise changes within a rapidly changing ecosystem, such as a grassland, would require a time series with higher temporal resolution than would be required to characterise changes in a more stable one, such as an old-growth forest.

The following preliminary analyses are recommended before time series datasets are analysed to derive trend information (Lhermitte *et al.*, 2011):

- mean time series—shows phenological cycle and differences over time;
- variance over time—shows temporal variability in amplitude and timing effects;
- serial correlation—shows the most appropriate temporal sampling interval for particular vegetation type (Alexandridis *et al.*, 2008);
- Principal Components Analysis (PCA; see Section 7.2 and Volume 2A—Section 9) and/or Fourier Transform (FT) components—significance of individual components needs to be understood; and
- noise estimation—must be able to distinguish signals from background noise.

8.3 Creating a Data Cube

It should be obvious that the primary dimension being analysed in time series studies is the temporal one. Accordingly, prior to that analysis, the image data needs to be standardised through time to ensure that data values are comparable in every sense. Many satellite sensors acquire imagery on a regular timetable, delivering observations that are theoretically consistent in space and time. When detecting trends in time series data, however, it is imperative that each temporal sequence of values for a given pixel be precisely:

- co-located, that is, represent the same ground area (see Section 3.1.1.1);
- comparable through time, that is, all pixel values in the dataset belong to a standardised distribution (see Section 3.1.1.2); and
- separated by equal time intervals.

The end result of a standardised time series dataset will be a three-dimensional data cube. For EO imagery, one dimension typically spans from north to south, a second spans from east to west, and the third spans from the first to the last image date (see Figure 8.1 above). Each two-dimensional image in the time series sequence can be viewed as an image plane, that is, a single image corresponding to a particular date. Where multiple image bands are included in a time series dataset, there are effectively multiple image planes for each date.

To convert the acquired image archive into a data cube requires that some of the intrinsic precision of the archive be sacrificed to increase consistency. Image resolution may need to be reduced in four dimensions:

- spectral—represent the image spectral extent in a single channel that is most appropriate for the study, such as a spectral index (see Volume 2C—Section 11);
- spatial—enlarge the original geometric pixel area to reduce data volume and/or improve locational accuracy (see Volume 2B);
- temporal—from a larger sequence of image acquisitions, select a smaller sequence of images separated by equal time intervals; and
- radiometric—represent the actual data range for the selected spectral channel in fewer bits.

These changes not only produce a consistent data cube, but also reduce the data volume that needs to be analysed. Given the inherently huge data volumes associated with EO time series datasets, this is a significant advantage, especially for global studies.

Reducing the spectral dimensionality of multispectral imagery is also advantageous for identifying ecologically significant factors in time series data. A common starting point for time series analyses of EO data is a spectral index, which summarises the changes in the most informative spectral channels for a particular surface feature (see Volume 2C—Section 11). For example, in terrestrial studies, an index that represents vegetation greenness, such as NDVI, is commonly used, particularly for phenological studies (see Excursus 8.1 and Volume 3A).

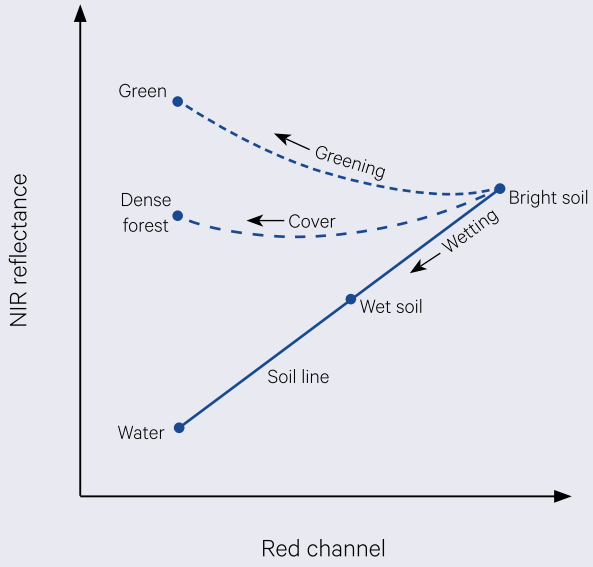
Excursus 8.1—Relative Calibration of Time Series Datasets

Source: Randall Donohue, CSIRO
Further Information: Donohue *et al.* (2007, 2008, 2009)
Calibrated fPAR dataset: *https://data.csiro.au/collections/#collection/CIcsiro:6084v1*

Some analyses have adopted a relative approach to calibrating EO time series datasets. For example, Donohue *et al.* (2008) developed an 'invariant-cover-triangle' method for relative calibration of AVHRR imagery over Australia. This method essentially standardises vegetation index values through time, and thus compensates for non-target variations in EO data, such as those resulting from atmospheric effects, sensor inconsistencies and changes in illumination geometry (see Section 3.1.1.2). As detailed below, the assumptions underlying this approach limit its application to regions with a wide range of bare soil types, which are spectrally distinct in the selected EO dataset.

Vegetation indices highlight the reflectance differences between surface features in order to differentiate levels of vegetation greenness (see Volume 2C—Section 11). When the spectral characteristics of surface features are plotted on a crossplot of red versus NIR reflectance, green vegetation forms a characteristic triangular shape, sometimes called the 'Tasselled Cap' (Kauth and Thomas, 1976). Bare soil typically plots in a line along the base of the vegetation triangle while vegetation plots above the soil line, with its greenness and cover determining distance from the line (see Figure 8.4).

Figure 8.4 Vegetation cover triangle



The invariant-cover-triangle approach assumes that the vegetation triangle does not vary in terms of red and NIR reflectances across all images. To achieve this, each image is modified so that its soil line has a slope of one and passes through the origin of the red/NIR crossplot, and the location of the dark value in the vegetation triangle is standardised as 2% of the red reflectance. These adjustments are consistent with the assumptions underlying the Normalised Difference Vegetation Index (NDVI). NDVI is then computed from the modified image values and converted to fraction of Photosynthetically Active Radiation (fPAR; see Volume 3A) for comparison with other biophysical parameters.

To illustrate the red/NIR reflectance ranges that are typical across Australia, six contrasting land cover types (based on the Interim Biogeographic Regionalisation for Australia (IBRA) 5.1; Environment Australia, 2000) are shown in Figure 8.5:

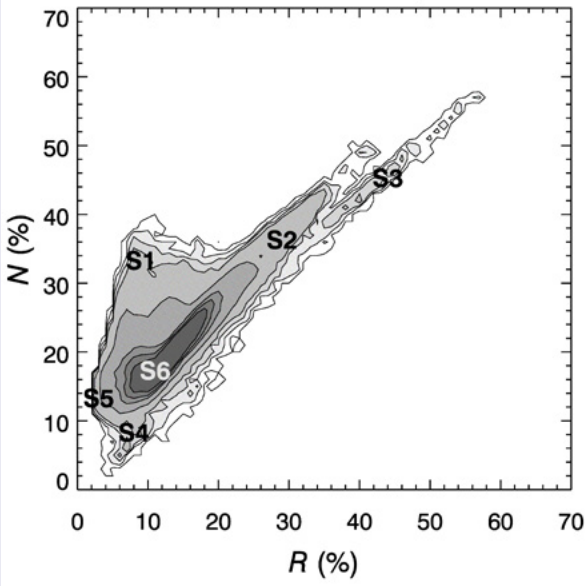
- dense, bright vegetation (1)—either closed forest (>70% cover) in summer or closed grasslands (cereal crops or improved pasture) in winter;
- bright soils with sparse vegetation in arid deserts (2)—open grasslands and shrublands with <30% cover;
- bright, bare surfaces (3)—typically dry salt lakes or winter snow;
- bare, dark surfaces (4)—water bodies or wet, bare soil;
- dense, dark vegetation (5)—open evergreen forests (30–70% cover) with dense, moist understoreys; and
- dark, moderately sparse vegetation on moderately dark soils (6)—woodlands (<30% cover), shrublands (<70% cover) or grasslands (<70% cover), plus improved pasture in southern regions during summer.

Figure 8.5 Land cover reflectances

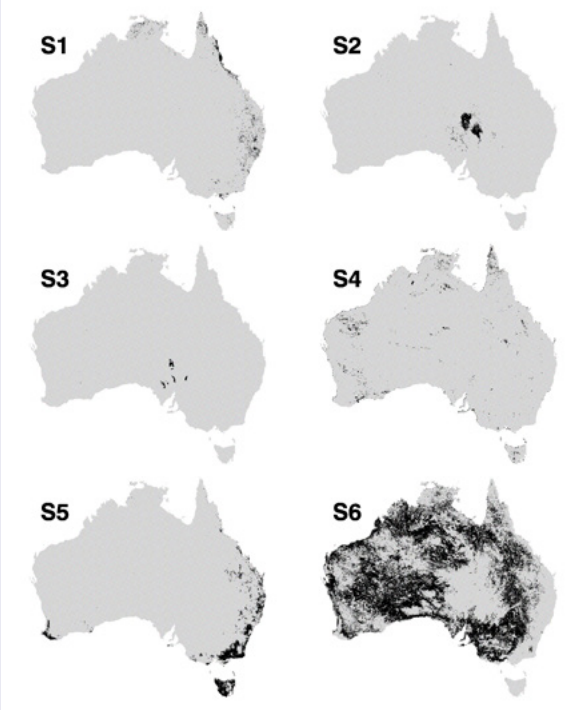
AVHRR reflectance characteristics of six types of Australian land cover are shown in terms of their spectral characteristics and their spatial distribution. These land covers show equivalent reflectance characteristics in both summer and winter.

a. Summer: February 1996

Crossplot of red versus NIR reflectance

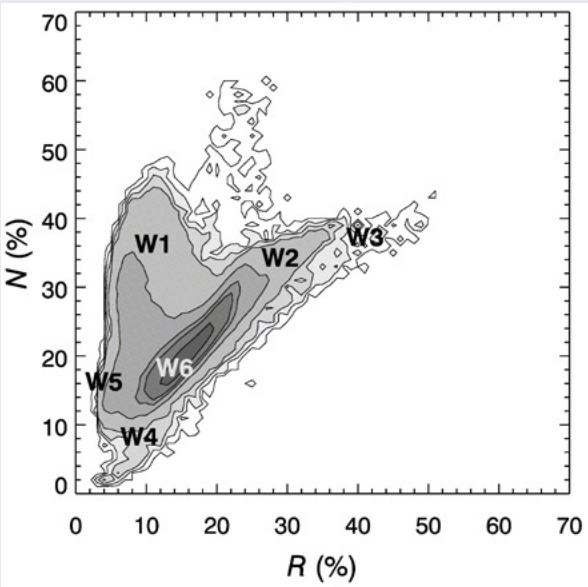


Spatial distribution of land cover types

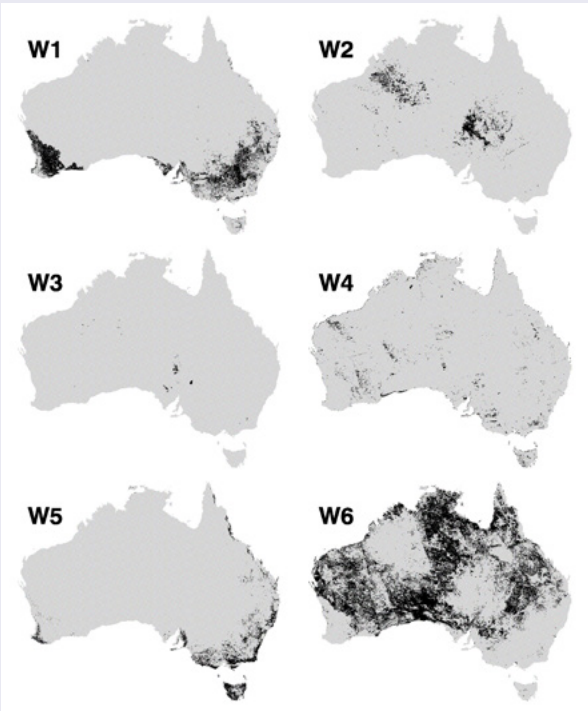


b. Winter: August 2003

Crossplot of red versus NIR reflectance



Spatial distribution of land cover types



Source: Donohue *et al.* (2008) Figure 1 © Elsevier: Used with permission.

The 'cover invariant triangle' method was applied to a 25-year AVHRR image archive, derived from two different sources:

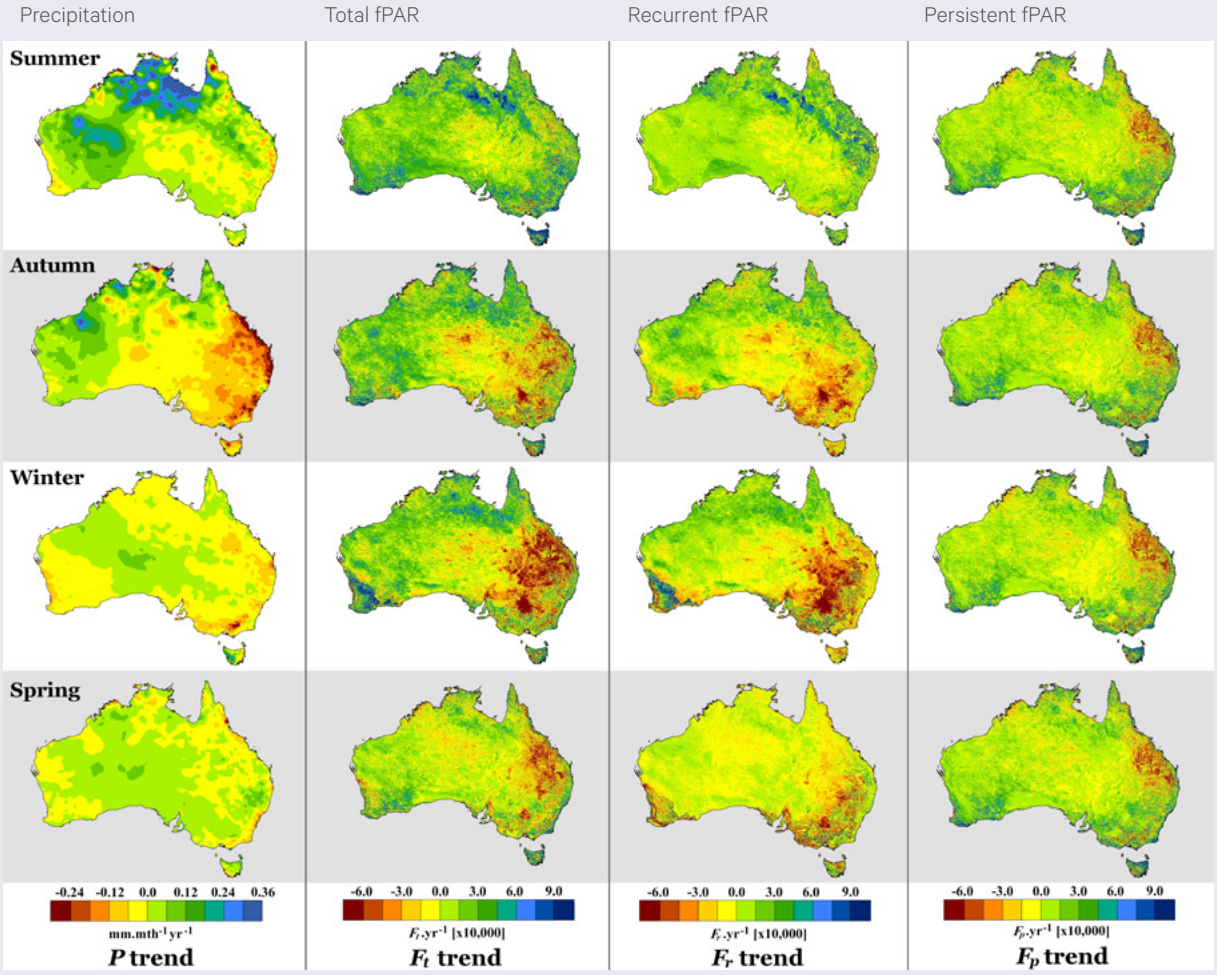
- PAL (Kidwell, 1998)—coarse resolution, fully calibrated, partially atmospherically corrected for 1981–1994; and
- CATS3.0 (King, 2003)—fine resolution, fully calibrated, non-atmospherically corrected for 1992–2004.

The resulting monthly fPAR dataset essentially harmonised these two datasets into a single time series, which was subsequently extended and used to model vegetation dynamics in Australia over two decades (Donohue *et al.* 2009; see Figure 8.6). The total fPAR time series was separated into two seasonally variable vegetation types:

- persistent—perennial vegetation that is actively photosynthesising all year; and
- recurrent—vegetation that grows in (annual) cycles of activity and dormancy (see Figure 8.7).

By comparing these three fPAR time series datasets with precipitation records for these years, this study demonstrated that during a period of increasing precipitation (7.2% between 1981–2006) across Australia, persistent vegetation increased by 21.3% while recurrent vegetation decreased by 7%. Site-based analyses showed that total cover increased even where precipitation decreased, with persistent vegetation being favoured over recurrent vegetation. The overall increase in vegetation cover was estimated as 7.8%, which may be attributable to the CO_2 fertilisation effect.

Figure 8.6 Seasonal trends 1981–2006



Source: Donohue *et al.* (2009) Figure 5 © Wiley: Used with permission.

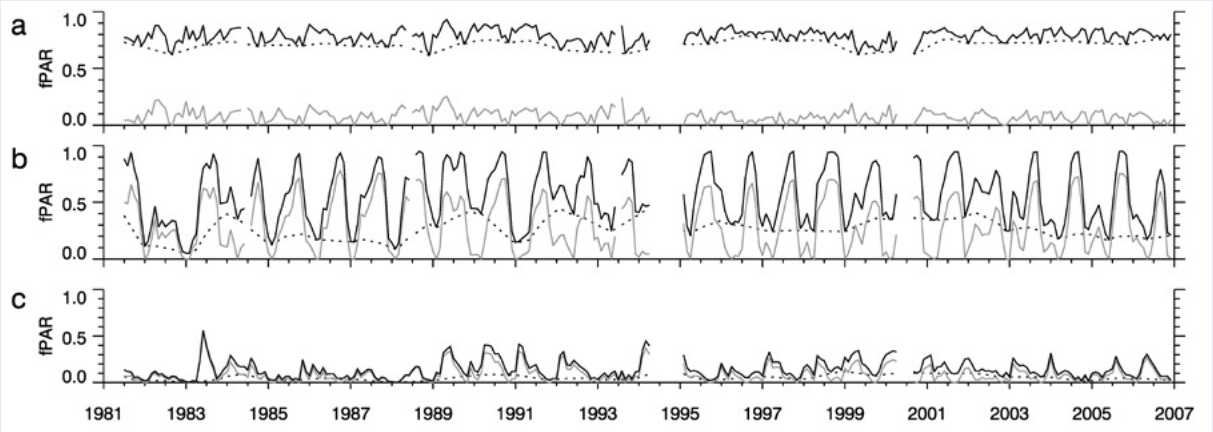
Figure 8.7 fPAR decomposition

These graphs demonstrate the temporal decomposition of the total fPAR time series (black line) into persistent (dotted line) and recurrent (grey line) vegetation for three sites:

a. Coffs Harbour, coastal NSW, with grazed pastures and open *Eucalyptus* forest (50–80% cover);

b. Cowra, central NSW, with grazed pastures and winter crops; and

c. Longreach, central Queensland, with low open *Astrebla* tussock grassland (20–50% cover).

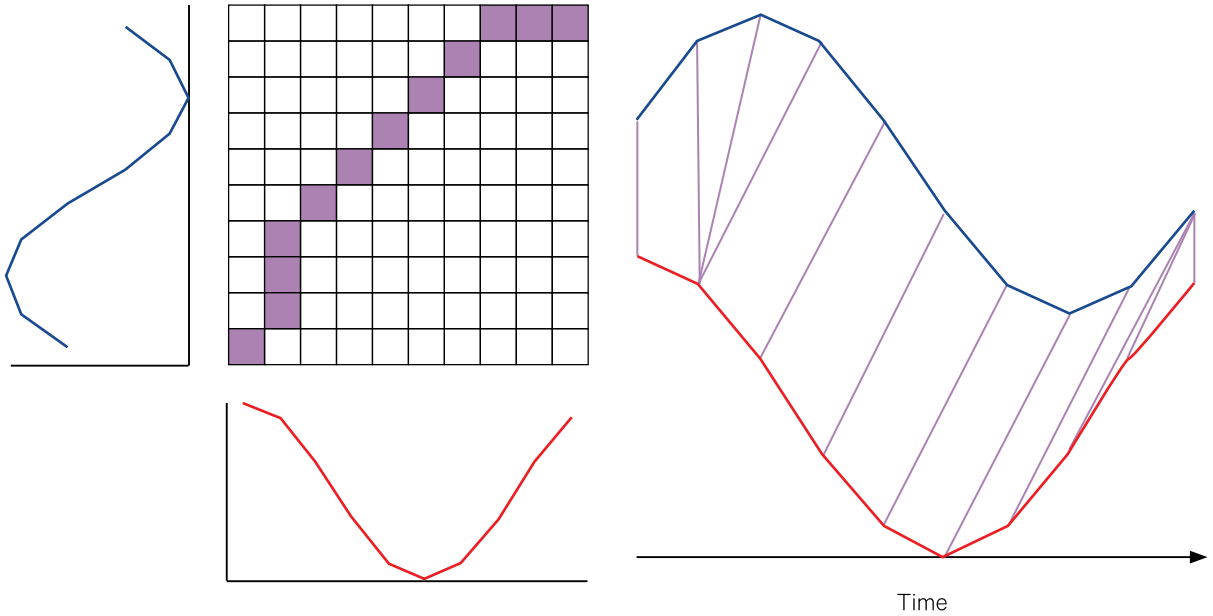


Source: Donohue *et al.* (2009) Figure 1 © Wiley: Used with permission.

Figure 8.8 Dynamic Time Warping (DTW)

a. Warping path matrix for two sequences

b. Resulting alignment



Source: Petitjean and Weber (2012b) Figure 3

To create equal time intervals between time series images, a composite image can be created from a number of acquisitions (see Section 10). While the goal of the compositing process is to convert a sequence of images into a manageable data cube, in which pixels have comparable values through regular time steps, it must be recognised that, by design, these products violate the criteria of equal time steps for each pixel.

An alternative approach to analysing EO time series datasets accepts irregular sampling in the temporal dimension and uses similarity measures such as Dynamic Time Warping (DTW) to 'match' seasonal trends (Petitjean *et al.*, 2011, 2012a, 2012b; Petitjean and Weber, 2014). As illustrated in Figure 8.8, DTW analyses the trend information in image sequences to synchronise their patterns. Given that many surface features do not follow a strictly synchronous cycle for seasonal changes, this flexibility is appropriate for analysing EO time series, especially when coupled with irregular coverage due to atmospheric conditions.

The advent of ARD has resulted in the availability of regularised time series datasets for selected EO image sources, principally Landsat data (see Section 3.2). Data cubes have been created for several locations, including Australia (Lewis *et al.*, 2016, 2017; see Section 11.2), USA (Dwyer *et al.*, 2018) and Switzerland (Giuliani *et al.*, 2017).

8.4 Detecting Anomalies

Apart from real changes in surface features, inter-annual variability in EO data can also result from data artefacts and atmospheric distortions, which confound trend detection methods. As introduced in Section 1, inconsistencies between image acquisitions can be introduced by instrumentation problems, changes in viewing and illumination geometry, surface anisotropy, and the presence of clouds, haze, and snow (see also Volume 1).

Outliers in images can be identified using a number of approaches. In EO imagery, outliers generally appear as unusually high values, but when spectral indices, such as NDVI, are used, they can manifest as unusually low values (see Volume 2C). Most outlier correction algorithms tend to consider each spatial pixel as a continuum of values in time. Given the statistical distribution underlying EO image acquisition, however, outlier analysis methods need to be appropriate to the data being processed (White and Nemani, 2006).

One of the difficulties in outlier detection in EO datasets is that cloud contamination does not occur as a random event, but will be related to seasonal cycles. This additional periodicity in the 'error' component confounds the assumption of stationarity, and requires careful consideration (see Volume 2E). Many cloud detection algorithms have been proposed and such artefacts should be identified in each image plane independently of the time series analysis (Zhu and Woodcock, 2012, 2014; Zhu *et al.*, 2015). Some algorithms refer to a cloud mask to determine which pixels should be included in time series analyses (see Section 3.1.1.2).

Once outliers are identified, they are generally replaced with values interpolated from valid data pixels, rather than leave data gaps that complicate trend detection algorithms (see Volume 2C). This process also potentially introduces uncertainty into the dataset, as interpolation routines will be forced to rely on composite data trends before the seasonal and longer-term components have been separated.

In time series data, however, the concept of an outlying value can occur in the dimensions of both space and time. Whereas in a single image an outlying pixel value would be dramatically and inexplicably different to its spatially-adjacent neighbours, in a time series an outlying value can also occur in the temporal dimension, that is, the value of a pixel may differ markedly from its value at the preceding and succeeding time intervals. For example, the temporal consistency of the HLS ARD product (see Section 3.2) is analysed using a time series smoothness index, which effectively estimates the noise component (Claverie *et al.*, 2018). In conjunction with a spatial filter to remove speckle, temporal outliers (often residual cloud pixels) are subsequently extracted using the Hampel (median) filter (Pearson, 2002), based on the simple ratio vegetation index (see Volume 2C). Other sources of noise in EO time series datasets could include long-term trends that are due to sensor shift over time (see Volume 1A—Section 13 and Volume 1B—Section 2) or the coincidence of geometric errors and outliers (see Volume 2A—Section 3 and Volume 2C).

8.5 Further Information

Data cubes:

Australia: *http://www.ga.gov.au/dea*

ESA Earth System Data Lab: *https://www.earthsystemdatalab.net*

8.6 References

Alexandridis, T.K., Gitas, I.Z., and Silleos, N.G. (2008). An estimation of the optimum temporal resolution for monitoring vegetation condition on a nationwide scale using MODIS/Terra data. *International Journal of Remote Sensing*, 29(12), 3589–3607.

Bradley, B.A., Jacob, R.W., Hermance, J.F., and Mustard, J.F. (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment*, 106(2), 137–145.

Chen, J.M., Jonsson, P., Tamura, M., Gu, Z. H., Matsushita, B., and Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, 91(3–4), 332–344.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., and Justice, C. (2018). The Harmonised Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219, 145–161. *https://doi.org/10.1016/j.rse.2018.09.002*

Donohue, R.J., Roderick, M.L., and McVicar, T.R. (2007). *Correcting long-term AVHRR reflectance data using the vegetation cover triangle*. CSIRO Land and Water Science Report 26/07. Canberra: CSIRO Land and Water. *http://www.clw.csiro.au/publications/science/2007/sr26-07.pdf*

Donohue, R.J., Roderick, M.L., McVicar, T.R. (2008) Deriving consistent long-term vegetation information from AVHRR reflectance data using a cover-triangle-based framework. *Remote Sensing of Environment*, 112, 2938–2949. *https://doi.org/10.1016/j.rse.2008.02.008*

Donohue, R.J., McVicar, T.R., and Roderick, M.L. (2009). Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. *Global Change Biology*, 15, 1025–1039. *https://doi.org/10.1111/j.1365-2486.2008.01746.x*

Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., Zhang, H.K., and Lymburner, L. (2018). Analysis Ready Data: Enabling Analysis of the Landsat Archive. *Remote Sensing*, 10, 1363. *https://doi.org/10.3390/rs10091363*

Environment Australia (2000). *Revision of the Interim Biogeographic Regionalisation for Australia (IBRA) and development of version 5.1*. Environment Australia, Canberra.

Giuliani, G., Chatenoux, B., De Bono, A., Rodila, D., Richard, J.-P., Allenbach, K., Dao, H., and Peduzzi, P. (2017). Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data*, 1(1–2), 100–117, *https://doi.org/10.1080/20964471.2017.1398903*

Jönsson, P., and Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8), 1824–1832.

Kauth, R.J., and Thomas, G.S. (1976). The Tasselled Cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*. Purdue University, West Lafayette, Indiana, 4B41–4B51.

Kidwell, K.B. (1998). *NOAA Polar Orbiter Data User's Guide (TIROS- N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10, NOAA- 11, NOAA-12, NOAA-13 and NOAA-14)*. National Environmental Satellite Data and Information Service, NOAA, Suitland.

King, E.A. (2003). The Australian AVHRR Data Set at CSIRO/EOC: Origins, Processes, Holdings and Prospects. *CSIRO Earth Observation Centre Report 2003/04*. CSIRO Earth Observation Centre, Canberra. *http://www.eoc.csiro.au/tech_reps/2003/tr2003_04.pdf*

Lewis, A., Lymburner, L., Purss, M.B.J., Brooke, B., Evans, B., Ip, A., Dekker, A.G., Irons, J.R., Minchin, S., Mueller, N., Oliver, S., Roberts, D., Ryan, B., Thankappan, M., Woodcock, R., and Wyborn, L. (2016). Rapid, high-resolution detection of environmental change over continental scales from satellite data—the Earth Observation Data Cube. *International Journal of Digital Earth*, 9(1), 106–111. *https://doi.org/10.1080/17538947.2015.1111952*

Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J., Dekker, A., Dhu, T., Hicks, A., Ip, A., Purss, M., Richards, C., Sagar, S., Trenham, C., Wang, P., and Wang, L.-W. (2017). The Australian geoscience data cube—Foundations and lessons learned. *Remote Sensing of Environment*, 202, 276–292. *https://doi.org/10.1016/j.rse.2017.03.015*

Lhermitte, S., Verbesselt, J., Verstraeten, W.W., and Coppin, P. (2011). A comparison of time series similarity measures for classification and change detection of ecosystem dynamics. *Remote Sensing of Environment*, 115, 3129–3152.

Pearson, R.K., (2002). Outliers in process modeling and identification. *IEEE Transactions on Control Systems Technology*, 10, 55–63. *https://doi.org/10.1109/87.974338*

Petitjean, F., Ketterlin, A. and Gançarski, P. (2011). A global averaging method for Dynamic Time Warping, with applications to clustering. *Pattern Recognition*, 44, 678–693.

Petitjean, F., Kurtz, C., Passat, N., and Gançarski, P. (2012a). Spatio-Temporal Reasoning for the Classification of Satellite Image Time Series. *Pattern Recognition Letters*, 33, 1805–1815. *https://doi.org/10.1016/j.patrec.2012.06.009*

Petitjean, F., Inglada, J., and Gançarski, P. (2012b). Satellite Image Time Series Analysis under Time Warping. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 3081–3095. *https://doi.org/10.1109/TGRS.2011.2179050*

Petitjean, F., and Weber, J. (2014). Efficient satellite image time series analysis under time warping. *IEEE Transactions on Geoscience and Remote Sensing*, 11(6), 1143–1147. *https://doi.org/10.1109/LGRS.2013.2288358*

Roerink, G.J., Menenti, M., and Verhoef, W. (2000). Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing*, 21(9), 1911–1917.

Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114 (1), 106–115.

Viovy, N., Arinoa, O., and Belward, A.S. (1992). The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing*, 13(8), 1585–1590.

White, M.A., and Nemani, R.R. (2006). Real-time monitoring and short-term forecasting of land surface phenology. *Remote Sensing of Environment*, 104, 43–49. *https://doi.org/10.1016/j.rse.2006.04.014*

Zhang, X., Friedl, M., Schaaf, C., Strahler, A., Hodges, J., Gao, F., Reed, B., and Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84(3), 471–475

Zhu, Z., and Woodcock, C.E. (2012). Object based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83–94. *https://doi.org/10.1016/j.rse.2011.10.028*

Zhu, Z., and Woodcock, C.E. (2014). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152, 217–234. *https://doi.org/10.1016/j.rse.2014.06.012*

Zhu, Z., Wang, S., and Woodcock, C.E. (2015). Improvements and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8 and Sentinel 2 imags. *Remote Sensing of Environment*, 159, 269–277. *https://doi.org/10.1016/j.rse.2014.12.014*

Zoffoli, M.L., Kandus, P., Madanes, N., and Calvo, D.H. (2008). Seasonal and interannual analysis of wetlands in South America using NOAA-AVHRR NDVI time series: The case of the Parana Delta Region. *Landscape Ecology*, 23(7), 833–848.

9 Characterising Temporal Trends

Once we have a data cube with consistent values as described in Section 8, temporal trends in the data can be characterised. Various approaches are used to identify relevant cycles from time series of EO imagery. Some of these methods combine treatment of outliers with trend detection. The methods proposed to analyse EO time series datasets can be differentiated in terms of:

- image trajectory being processed (see Section 9.1):
  - individual pixels, for the whole sequence of time series or selected sub-sequence(s); or
  - groups of pixels, based on whole sequence or selected sub-sequence(s);
- processing of trajectory (see Section 9.2):
  - original data values (outliers removed);
  - smoothed data values;
  - sub-sampled data values;
  - transformed data values;
  - modelled data values; or
  - parameters extracted from original, smoothed, sub-sampled or transformed data values;
- comparison of processed trajectories (see Section 9.3):
  - classifying trajectories using similarity or distance metrics (see Volume 2E); or
  - clustering trajectories using hierarchical or partitioning algorithms (see Volume 2E).

Advanced processing methods for time series datasets include high-dimensional statistical analysis, visualisation (e.g. Small, 2012) and machine learning techniques (see Section 9.4). Applications of time series datasets are further detailed in Volume 3.

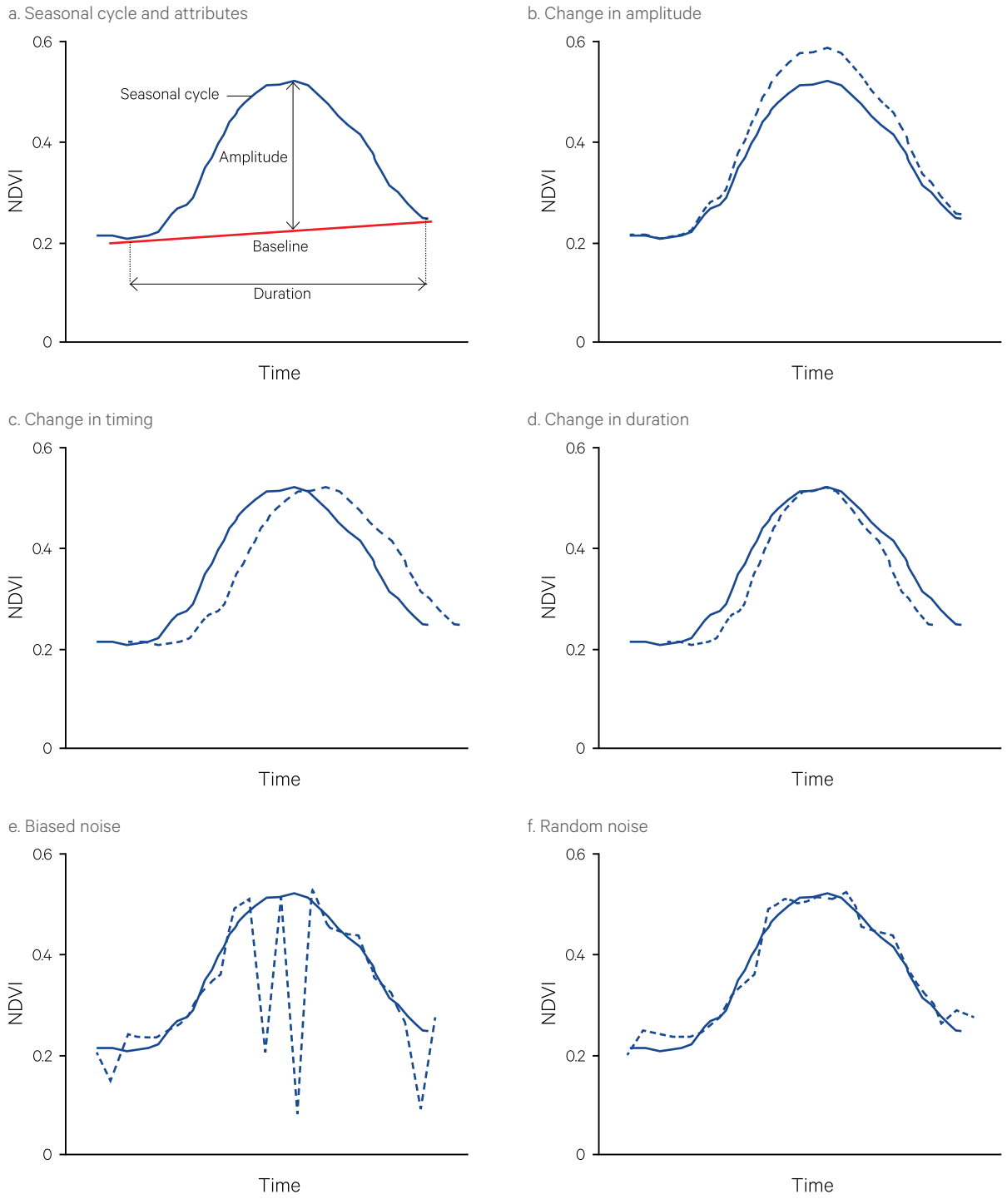
9.1 The Time Series Trajectory

The changing values of each pixel through time can be considered as its 'movement', which plots a trajectory. The temporal trajectories associated with Earth surface features can be characterised in terms of a number of discrete attributes, such as:

- baseline—background value of trajectory, without seasonal cycles;
- amplitude—difference between seasonal peak and baseline;
- timing—location of seasonal peak in time sequence;
- duration—length of seasonal cycle; and
- shape—rate of seasonal changes (see Figure 9.1).

Background image: International Space Station photograph of Albert and Leichhardt Rivers, Queensland, acquired on 21 June 2006 (ISS013-E39183).
Source: NASA. Retrieved from *https://eol.jsc.nasa.gov/searchphotos/photo.pl?mission=ISS013&roll=E&frame=39183*

Figure 9.1 Temporal trajectory attributes



Adapted from Lhermitte *et al.* (2011) Figure 1

Time series that follow defined annual or seasonal cycles may be sub-divided into sub-sequences to allow tracking algorithms to focus on the most relevant changes. Within a sub-sequence, various metrics may be appropriate to help identify noise or cyclic patterns. In some situations, for example, while the shape of the seasonal change may be similar for two different locations, the baseline, amplitude and/or timing of the seasonal change may differ due to differences in temperature or precipitation between sites. Comparison of sub-sequences, either to detect noise or analyse trends, may highlight one or more of the following differences:

- rescaling of:
  - spectral values, that is, change in amplitude; and/or
  - time dimension, that is, change in season length; and/or
- shifting of:
  - spectral values, that is, change in baseline; and/or
  - time dimension, that it, change in timing (Lhermitte *et al.*, 2011).

The relevance of these differences to interpreting EO imagery is summarised in Table 9.1.

Processing options can be applied to the trajectories of individual pixels or the time series of groups of pixels. In either case, the whole time series can be analysed or sub-sequences can be extracted based on some criteria. External datasets, such as climate data or land cover maps may be used to stratify the image time series into regions of phenologically and climatically similar pixels (White and Nemani, 2006; Bradley and Mustard, 2008; Petitjean *et al.*, 2012a). Analysing trajectories for pixels groups rather than for individual pixels increases radiometric homogeneity and reduces data processing volumes (Petitjean and Weber, 2014). It also provides a more robust basis for forecasting trends and highlighting temporal changes within those groups (White and Nemani, 2006).

Table 9.1 Time series sub-sequence differences

Subsequence Attribute Change	Possible Reason for Difference in		
	Spectral Dimension	Temporal Dimension	Example
Shift Baseline	Change in background colour of soil or understorey vegetation		NDVI saturates in wetter regions with higher vegetation density (Geerken *et al.*, 2005a; Heumann *et al.*, 2007)
Rescale Amplitude	Change related to vegetation cover or condition		Vegetation changes could change height of NDVI curve (Eklundh and Olsson, 2003)
Shift Timing		Change in timing of peak growing season	Timing of growing season can change with timing of Spring rains (Reed *et al.*, 1994; Hill and Donald, 2003; Wardlow *et al.*, 2007)
Rescale Shape		Changes in length of growing season (Myneni et al., 1997)	
Change Shape		Major changes in land cover or use	

9.2 Analysing Trends

Some time series analysis methods are applied directly to EO datasets, but generally there is good reason to smooth, sub-sample, transform or model the data in some way prior to analysis. Transformation processes generally involve some form of data compression, which simplifies the time series, such as:

- Fourier decomposition—decompose time series into sum of sine waves;
- Piecewise approximations—represent time series as sequence of segments, possibly characterised by their slope;
- Singular value decomposition (SVD)—optimise dimensionality reduction; or
- Symbolic approximation—label each value range and represent as a sequence.

Various parameters can also be extracted from the time series and these parameters can be the basis for trend analysis. Alternatively, the trajectory can be sub-sampled or represented as a mathematical function (linear or non-linear). Relevant parameters could include the seasonal mean, minimum, maximum, variance, or integral, the duration of a growing season, the slope of a selected sub-sequence, or its variance within a specified time interval.

Table 9.2 Time series analysis methods

Method		Description	Case Study
Group	Example		
Step-wise Linear Regression	Weighted Least Squares	Piecewise linear regression approximates time series, weighted for local maxima within a moving window	Verbesselt *et al.* (2010a, 2010b)
Non-linear Functions	High-order Spline	Function fitted to time series data to model inter-annual phenology	Bradley *et al.* (2007); Bradley and Mustard, 2008); Hermance *et al.* (2007)
	Asymmetric Gaussian Function	Least squares fit to skewed Gaussian model	TIMESAT: Jönsson and Eklundh (2002; 2004);
	Savitsky-Golay Filter	Local polynomial least squares fit within a moving window	TIMESAT: Jönsson and Eklundh (2002; 2004); Hird and McDermid (2009)
	Asymmetric Double-Sigmoid Function	Inflection point of function fitted to NDVI related to deciduous greening onset	Soudani *et al.* (2008)
Step-wise Non-linear Regression	Series of piecewise logistic functions	Represent intra-annual vegetation dynamics, often based on four key transition dates for annual crops (see Volume 3A)	Zhang *et al.* (2003); Zhu *et al.* (2011)
Best Slope	Best Index Slope Extraction (BISE)	The NDVI change rate is used as a benchmark to discard outliers	White *et al.* (1997); Viovy *et al.* (1992)
Harmonic Analysis	Harmonic Analysis of Time Series (HANTS)	Iterative FFT to identify and remove cloud contaminated pixels, then replace with interpolated values from imagery acquired in close temporal proximity	Azzali and Menenti (1996)
	Fourier Function	Represent time series as set of sine/cosine curves	Olsson and Eklundh (1994); Jakubauskas *et al.* (2001, 2002)
	Discrete Fourier Transform (DFT)	First and second harmonics summarise amplitude and phase of annual and biannual cycles in AVHRR NDVI	Moody and Johnson (2001)
Affine Transformations for Data Reduction	Principal Components Analysis (PCA)	Standardised PCA of AVHRR imagery for Africa separate greenness, seasonal greenness changes, and variations related to sensor and climatic factors	Eastman and Fulk (1993)
	Minimum Noise Fraction (MNF)	Cascaded PCA transform with new components ordered by increasing signal-to-noise ratio using maximum autocorrelation factors	Nielsen (2011)

Data reduction transformations are relevant in this context to isolate the noise component as a single band, which can then be analysed or discarded (see Volume 2C). Some of the operations that might be used to process time series trajectories are summarised in Table 9.2. The Minimum Noise Fraction (MNF) transformation effectively orders image planes by data quality as estimated by their signal-to-noise ratio (Green *et al.*, 1988). The Maximum Autocorrelation Factor (MAF) procedure is often used to estimate the noise covariance matrix by assuming that the noise values are unlikely to be spatially correlated (Switzer and Green, 1984; Nielsen *et al.*, 1998). The transformed components then comprise a linear combination of spatially-correlated, orthogonal signal component(s) and an uncorrelated noise component. Once isolated, the noise component may be filtered then recombined with the signal and back-transformed to the original image space (see Volume 2C).

9.3 Comparing Trajectories

Once trajectories have been appropriately selected and prepared, they can be compared on the basis of some measure of separation and either classified into pre-defined groups or clustered in order of similarity (see Volume 2E). Clustering methods that are used with time series datasets are reviewed by Liao (2005). Distance functions can be:

- metric, that is based on positivity, symmetry and triangle inequality (see Volume 2E), such as Euclidean distance or correlation methods (see Volume 2E); or
- non-metric, such as Dynamic Time Warping (DTW) or Longest Common Sub-Sequence (LCSS).

Euclidean distance and DTW are compared in Figure 9.2. Non-metric distance functions, such as DTW, have particular advantages when time series datasets contain missing values or irregular time intervals (Petitjean *et al.*, 2012a, 2012b). Various relevant classification and clustering metrics, algorithms and approaches are detailed in Volume 2E.

Several time series similarity measures used with EO data were reviewed by Lhermitte *et al.* (2011) in terms of three categories, namely:

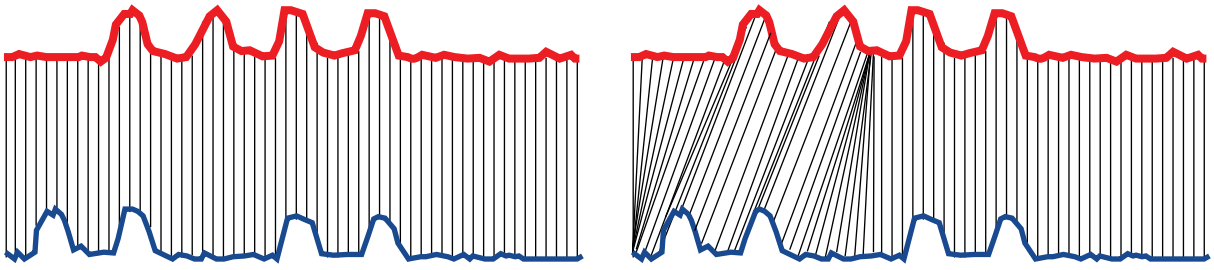
- direct use of the original time series data;
- indirect use via transformations extracted from the data; or
- indirect use via metrics derived from the data (see Table 9.3).

Their comparisons of different time series similarity measures stressed that both the time series characteristics and the ecosystem dynamics need to be understood before an analysis method is selected.

Figure 9.2 Time series distance measures

a. Euclidean Distance matches two datasets based on equal time intervals between observations.

b. Dynamic Time Warping allows two datasets to include missing values and irregular time intervals.



Adaped from: Mike Izbicki, Claremont McKenna College (*http://izbicki.me/blog/converting-images-into-time-series-for-data-mining*)

Table 9.3 Time series similarity measures

Data Used	Type of Approach	Example Approach	Features	Limitations	Example
Original Data	Distance measures	Minkowski distance	Easy to compute and interpret in vector analysis and land cover classification; quantifies difference between time series	Sensitive to rescaling/ shifting of amplitude and timing, and introduced noise	Bayarjargal *et al.* (2006); Viovy (2000); Loveland *et al.* (2000); Ehrlich and Lambin (1996); Bergen *et al.* (2005)
	Correlation measures	Pearson's cross-correlation coefficient	Accounts for temporal correlation	Insensitive to amplitude changes but sensitive to timing changes and noise	Geerken *et al.* (2005a, 2005b); Wang *et al.* (2009)
Transformed Data	Data Reduction	Principal Component Analysis	Higher order components capture most information; quantifies difference between time series	Sensitive to rescaling/ shifting of amplitude and timing, but not noise	Lobo and Maisongrande (2008)
	Harmonic or Frequency Analysis	Fourier Transform	Isolates periodic signals with specific frequency; quantifies difference between time series; derived metrics can quantify shape similarity (Evans and Geerken, 2006)	More sensitive to amplitude and timing rescaling and noise, but less sensitive to amplitude shifting	Azzali and Menenti (2000)
Indirect Metrics derived from original or transformed data	Growing Season Parameters related to photosynthetic activity	e.g. Beginning of Growing Season; End of Growing Season	Time series represented as simple statistics	Not inter-comparable	Reed *et al.* (1994); Zhang *et al.* (2003)
		Adapted Scale-Invariant Similarity Measures	Can be designed to user requirements	External knowledge required to interpret results	Defries *et al.* (1995); Borak *et al.* (2000); Hansen *et al.* (2003)

Adapted from Lhermitte *et al.* (2011)

9.4 Analytical Methods

While a wide range of analysis methods has been applied to EO time series datasets for many different environmental applications, hitherto there has been no definitive approach. For example, some of the integrated change detection approaches developed for multi-temporal Landsat imagery are summarised in Table 9.4. A greater emphasis on near real time products is anticipated in the future (Zhu *et al.*, 2017).

To date, most EO time series analyses have focussed on the temporal characteristics, especially for medium resolution datasets such as the Landsat archive (Zhu *et al.*, 2017). However, it is important when analysing EO time series datasets that the spatial dependence of pixels is considered in conjunction with the temporal dimension. Militino *et al.* (2018) consider this non-trivial problem, and discuss suitable approaches such as spatio-temporal stochastic models.

Table 9.4 Landsat-based change detection approaches

Name	Description	Method category	Applications
Landsat-based detection of Trends in Disturbance and Recovery (LandTrendr)	Segmentation methods are used to detect abrupt change in the Normailsed Burn Ration (NBR) value. Gradual changes are described by slope of segments between abrupt changes.	Segmentation	Forest disturbance and recovery (Kennedy *et al.*, 2010, 2012) Land cover change (Franklin *et al.*, 2015)
Vegetation Change Tracker (VCT)	Landsat images are normalised into a probability index called the integrated forest z-score (IFZ) then thresholding is used to detect disturbance.	Thresholding	Forest disturbance (Huang *et al.*, 2010) Wetland change (Kayastha *et al.*, 2012)
Breaks for Additive Season and Trend (BFAST Monitor)	The NDVI time series is decomposed in trend, season and noise components to detect gradual and abrupt changes	Statistical boundary method	Drought-related vegetation disturbance (Verbesselt *et al.*, 2012)
Continuous Change Detection and Classification (CCDC)	All spectral bands are used to detect changes in surface features.	Statistical boundary method	Land cover change (Zhu and Woodcock, 2014; Zhu et al., 2016)

Source: Zhu *et al.* (2017)

9.5 Further Information

TIMESAT: A software package for analysing time-series of satellite sensor data

Jönsson and Eklundh (2004)

Eklundh and Jönsson (2016)

https://www.nateko.lu.se/research/remote-sensing-and-earth-observation/lund-earth-observation-research-group/time-series-analysis-in-remote-sensing

http://web.nateko.lu.se/timesat/timesat.asp

LandTrendr: Temporal segmentation algorithms

Kennedy *et al.* (2010)

TimeSync: Time-series visualization and data collection tool

Cohen *et al.* (2010)

9.6 References

Azzali, S., and Menenti, M. (Eds) (1996). *Fourier analysis of temporal NDVI in the Southern African and American continents*. DLO Winand Staring Centre, Wageningen, The Netherlands. *http://edepot.wur.nl/363147*

Azzali, S., and Menenti, M. (2000). Mapping vegetation–soil–climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. *International Journal of Remote Sensing*, 21(5), 973–996.

Bayarjargal, Y., Karnieli, A., Bayasgalan, M., Khudulmur, S., Gandush, C., and Tucker, C.J. (2006). A comparative study of NOAA-AVHRR derived drought indices using change vector analysis. *International Journal of Remote Sensing*, 105(1), 9–22.

Bergen, K.M., Brown, D.G., Rutherford, J.F., and Gustafson, E.J. (2005). Change detection with heterogeneous data using ecoregional stratification, statistical summaries and a land allocation algorithm. *Remote Sensing of Environment*, 97(4), 434–446.

Borak, J.S., Lambin, E.F., and Strahler, A.H. (2000). The use of temporal metrics for land cover change detection at coarse spatial scales. *International Journal of Remote Sensing*, 21(6–7), 1415–1432.

Bradley, B.A., and Mustard, J.F. (2008). Comparison of phenology trends by land cover class: A case study in the Great Basin, USA. *Global Change Biology*, 14(2), 334–346.

Bradley, B.A., Jacob, R.W., Hermance, J.F., and Mustard, J.F. (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment*, 106(2), 137–145.

Cohen, W.B., Yang, Z., and Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync—Tools for calibration and validation. *Remote Sensing of Environment*, 114, 2911–2924. *https://doi.org/10.1016/j.rse.2010.07.010*

Defries, R.S., Hansen, M.C., and Townshend, J.R.G. (1995). Global discrimination of land cover types from metrics derived from AVHRR path finder data. *Remote Sensing of Environment*, 54(3), 209–222.

Eastman, J.R., and Fulk, M. (1993). Long sequence time series evaluation using standardized principal components. *Photogrammetric Engineering and Remote Sensing*, 59(8), 1307–1312.

Ehrlich, D., and Lambin, E.F. (1996). Broad-scale land-cover classification and interannual climatic variability. *International Journal of Remote Sensing*, 17(5), 845–862.

Eklundh, L., and Olsson, L. (2003). Vegetation index trends for the African Sahel 1982–1999. *Geophysical Research Letters*, 30(8), 1430. *https://doi.org/10.1029/2002GL016772*

Eklundh, L. and Jönsson, P (2016). TIMESAT for processing time-series data from satellite sensors for land surface monitoring. In *Multitemporal Remote Sensing*, (Ed. Y. Ban) Springer International Publishing, pp. 177-194.

Franklin, S.E., Ahmed, O.S., Wulder, M.A., White, J.C., Hermosilla, T., and Coops, N.C. (2015). Large Area Mapping of Annual Land Cover Dynamics Using Multitemporal Change Detection and Classification of Landsat Time Series Data. *Canadian Journal of Remote Sensing*, 41, 293–314. *https://doi.org/10.1080/07038992.2015.1089401*

Geerken, R., Batikha, N., Celis, D., and Depauw, E. (2005a). Differentiation of rangeland vegetation and assessment of its status: Field investigations and MODIS and SPOT VEGETATION data analyses. *International Journal of Remote Sensing*, 26(20), 4499–4526.

Geerken, R., Zaitchik, B., and Evans, J.P. (2005b). Classifying rangeland vegetation type and fractional cover of semi-arid and arid vegetation covers from NDVI time-series. *International Journal of Remote Sensing*, 26(24), 5535–5554.

Green, A.A., Berman, M., Switzer, P., and Craig, M.D. (1988). A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1), 65–74.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Carroll, M., Dimiceli, C., and Sohlberg, R.A. (2003). Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous elds algorithm. *Earth Interactions*, 7(10), 1–15.

Hermance, J.F., Jacob, R.W., Bradley, B.A., and Mustard, J.F. (2007). Extracting phenological signals from multiyear AVHRR NDVI time series: Framework for applying high-order annual splines with roughness damping. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), 3264–3276.

Heumann, B.W., Seaquist, J.W., Eklundh, L. and Jönsson, P. (2007). AVHRR Derived Phenological Change in the Sahel and Soudan, Africa, 1982-2005. *Remote Sensing of Environment*, 108, 385–392.

Hill, M.J., and Donald, G.E. (2003). Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. *Remote Sensing of Environment*, 84(3), 367–384.

Hird, J.N., and McDermid, G.J. (2009). Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment*, 113(1), 248–258.

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., and Vogelmann, J.E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, 114, 183–198. *https://doi.org/10.1016/j.rse.2009.08.017*

Jakubauskas, M.E., Legates, D.R., and Kastens, J.H. (2001). Harmonic Analysis of Time-Series AVHRR NDVI Data. Photogrammetric Engineering and Remote Sensing, 67(4), 461–470.

Jakubauskas, M.E., Peterson, D.L., Kastens, J.H., and Legates, D.R. (2002). Time Series Remote Sensing of Landscape-Vegetation Interactions in the Southern Great Plains. *Photogrammetric Engineering and Remote Sensing*, 68(10), 1021–1030.

Jönsson, P., and Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8), 1824–1832.

Jönsson, P., and Eklundh, L. (2004). TIMESAT—a program for analysing time-series of satellite sensor data. *Computers and Geosciences*, 30, 833–845.

Kayastha, N., Thomas, V., Galbraith, J., and Banskota, A. (2012). Monitoring wetland change using inter-annual landsat time-series data. *Wetlands*, 32, 1149–1162. *https://doi.org/10.1007/s13157-012-0345-1*

Kennedy, R.E., Yang, Z., and Cohen, W.B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—temporal segmentation algorithms. *Remote Sensing of Environment*, 114, 2897–2910. *https://doi.org/10.1016/j.rse.2010.07.008*

Kennedy, R.E., Yang, Z., Cohen, W.B., Pfaff, E., Braaten, J., and Nelson, P. (2012). Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sensing of Environment*, 122, 117–133. *https://doi.org/10.1016/j.rse.2011.09.024*

Liao, T. (2005). Clustering of time series data—A survey. *Pattern Recognition*, 38(11), 1857–1874.

Lobo, A., and Maisongrande, P. (2008). Searching for trends of change through exploratory data analysis of time series of remotely sensed images of SW Europe and NW Africa. *International Journal of Remote Sensing,* 29(17) 5237–5245.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., and Merchant, J.W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing,* 21(6–7), 1303–1330.

Lhermitte, S., Verbesselt, J., Verstraeten, W.W., and Coppin, P. (2011). A comparison of time series similarity measures for classification and change detection of ecosystem dynamics. *Remote Sensing of Environment*, 115, 3129–3152.

Militino, A.F., Ugarte, M.D., and Pérez-Goya, U. (2018). An Introduction to the Spatio-Temporal Analysis of Satellite Remote Sensing Data for Geostatisticians. Ch 13 in *Handbook of Mathematical Geosciences: Fifty Years of IAMG*. (Eds: Sagar, B.S.D., Cheng, Q., and Agterberg, F.). Springer Open. ISBN 978-3-319-78999-6 *https://link.springer.com/content/pdf/10.1007%2F978-3-319-78999-6.pdf*

Moody, A., and Johnson, D.M. (2001). Land-surface phenologies from AVHRR using the discrete Fourier transform. *Remote Sensing of Environment*, 75(3), 305–323.

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G., and Nemani, R.R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, 386(6626), 698–702.

Nielsen, A.A. (2011). Kernel Maximum Autocorrelation Factor and Minimum Noise Fraction Transformations. *IEEE Transactions on Geoscience and Remote Sensing,* 20(3), 612–624.

Nielsen, A.A., Conradsen, K., and Simpson, J.J. (1998). Multivariate Alteration Detection (MAD) and MAF Postprocessing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies. *Remote Sensing of Environment*, 64, 1–19.

Olsson, L., and Eklundh, L. (1994). Fourier Series for analysis of temporal sequences of satellite sensor imagery. *International Journal of Remote Sensing*, 15(18), 3735–3741. *https://doi.org/10.1080/01431169408954355*

Petitjean, F., Kurtz, C., Passat, N., and Gançarski, P. (2012a). Spatio-Temporal Reasoning for the Classification of Satellite Image Time Series. *Pattern Recognition Letters*, 33, 1805–1815. *https://doi.org/10.1016/j.patrec.2012.06.009*

Petitjean, F., Inglada, J., and Gançarski, P. (2012b). Satellite Image Time Series Analysis under Time Warping. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 3081–3095. *https://doi.org/10.1109/TGRS.2011.2179050*

Petitjean, F., and Weber, J. (2014). Efficient satellite image time series analysis under time warping. *IEEE Transactions on Geoscience and Remote Sensing*, 11(6), 1143–1147. *https://doi.org/10.1109/LGRS.2013.2288358*

Reed, B.C., Brown, J.F., Vanderzee, D., Loveland, T.S., Merchant, J.W., and Ohlen, D.O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5(5), 703–714.

Small, C. (2012). Spatiotemporal dimensionality and Time-Space characterization of multitemporal imagery. *Remote Sensing of Environment*, 124, 793–809. *https://doi.org/10.1016/j.rse.2012.05.031*

Soudani, K., le Maire, G., Dufrene, E., Francois, C., Delpierre, N., Ulrich, E., and Cecchini, S. (2008). Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sensing of Environment*, 112(5), 2643–2655.

Switzer, P., and Green, A.A. (1984). Min/max autocorrelation factors for multivariate spatial imagery. *Tech. Report 6,* April 1984, Dept. Statistics, Stanford.

Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D. (2010a). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114 (1), 106–115.

Verbesselt, J., Hyndman, R., Zeileis, A., and Culvenor, D. (2010b). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sensing of Environment*, 114(12), 2970–2980.

Verbesselt, J., Zeileis, A., and Herold, M. (2012). Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment,* 123, 98–108. *https://doi.org/10.1016/j.rse.2012.02.022*

Viovy, N. (2000). Automatic Classification of Time Series (ACTS): A new clustering method for remote sensing time series. *International Journal of Remote Sensing*, 21 (6–7), 1537–1560.

Viovy, N., Arinoa, O., and Belward, A.S. (1992). The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing*, 13(8), 1585–1590.

Wang, L., Chen, J., Gong, P., Shimazaki, H., and Tamura, M. (2009). Land cover change detection with a cross-correlogram spectral matching algorithm. *International Journal of Remote Sensing*, 30(12), 3259–3273.

Wardlow, B. D., Egbert, S. L., and Kastens, J. H. (2007). Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. *Remote Sensing of Environment*, 108(3), 290–310.

White, M.A., and Nemani, R.R. (2006). Real-time monitoring and short-term forecasting of land surface phenology. Remote Sensing of Environment, 104, 43–49. *https://doi.org/10.1016/j.rse.2006.04.014*

White, M.A., Thornton, P E., and Running, S.W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, 11, 217–234.

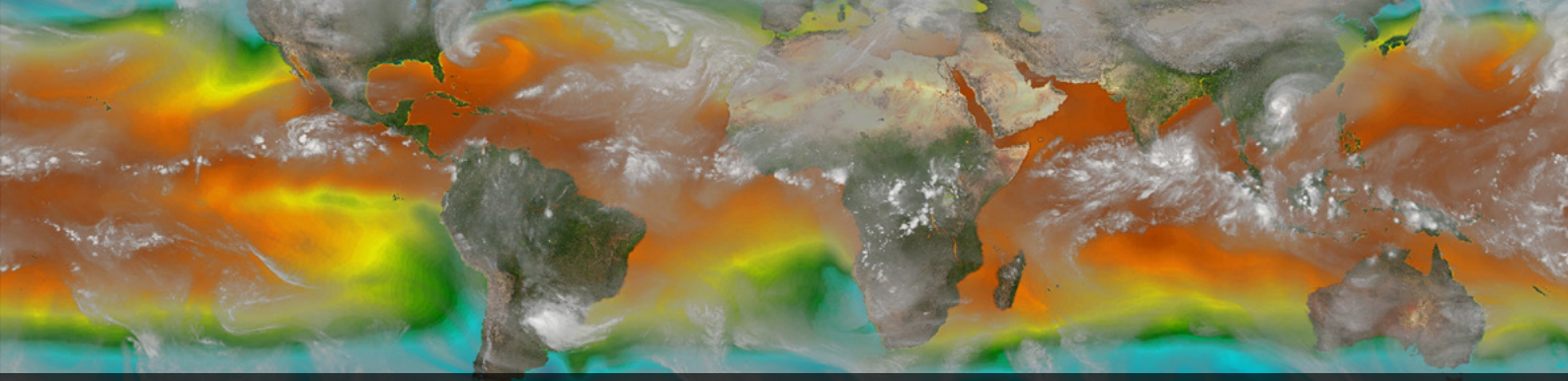
Zhang, X., Friedl, M., Schaaf, C., Strahler, A., Hodges, J., Gao, F., Reed, B., and Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84(3), 471–475.

Zhu, W., Tian, H., Xu, X., Pan, Y., Chen, G., and Lin, W. (2011). Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982–2006. *Global Ecology and Biogeography*, 21(2), 260–271. *https://doi.org/10.1111/j.1466-8238.2011.00675.x*

Zhu, Z. (2017). Change detection using landsat time series: A review of frequencies, preprocessing, algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 370–384. *https://doi.org/10.1016/j.isprsjprs.2017.06.013*

Zhu, Z., and Woodcock, C.E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152–171. *https://doi.org/10.1016/j.rse.2014.01.011*

Zhu, Z., Gallant, A.L., Woodcock, C.E., Pengra, B., Olofsson, P., Loveland, T.R., Jin, S., Dahal, D., Yang, L., and Auch, R.F. (2016). Optimizing selection of training and auxiliary data for operational land cover classification for the LCMAP initiative. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 206–221. *https://doi.org/10.1016/j.isprsjprs.2016.11.004*



10 Pixel-based Composites

Fusion methods that merged image values from a pair of spatially-coincident images into a single image were described In Section 6. Below we will consider pixel-based methods for compositing multiple, spatially-coincident images from a time series dataset acquired by a particular sensor into a single image (see Section 10.1). Pixel-based composite images are also being generated using pixels acquired by two or more different sensors (see Section 10.2).

A pixel-based composite image comprises pixels that were acquired in multiple, overlapping images within a given time period. A number of satellite image products are now routinely available from various EO data suppliers that use compositing methods to provide a single, summary image representing that time period. Such pixel-based, composite images are useful for large area coverage, rapid assessment of surface condition and change, and long term monitoring of trends. Pixel-based composites enable all images—even those that are partially clouded—to be potentially included in analyses (Griffiths *et al.*, 2019). They also tend to display better colour balance than is generally achievable from mosaicking of multiple image scenes (Roberts *et al.*, 2017).

Compositing methods generally involve selecting the 'best' pixel from a set of images acquired over a specified time period, where 'best' is judged by some criteria, such cloud-free, optimum Sun angle, and/or proximity to regular time steps. The resulting composite image thus includes pixels acquired on different dates with potentially different atmospheric conditions and/or different imaging geometries. While these images are particularly useful for global and regional studies, the compositing process may introduce artefacts that need to be considered in subsequent processing (see Section 3.1.1.1). To compare composited data through time, especially for global monitoring, Tan *et al.* (2006) recommend consistency in the:

- geometry and frequency of sensor observations;
- methods used for compositing; and
- cloud frequency.

It is worth noting that different gridding algorithms are used across the range of MODIS composite products (Tan *et al.*, 2006).

The values in composite images may be based on calibrated spectral reflectance values (see Section 3.1), or a selected spectral index that emphasises the trend in features of interest through time (see Section 4.2.1 or Volume 2C). Compositing methods reduce the impact of cloud, aerosols and differences in imaging geometry within a set of temporal images, and also signficantly reduce data volume, which is an important advantage when processing very large time series datasets. As such, they have been used to monitor meteorological conditions for several decades (Kohrs *et al.*, 2014) and are increasingly useful for generating regional to global maps of land cover and land use (Cihlar, 2000; Hansen and Loveland, 2012; Griffiths *et al.*, 2013).

Background image: Global composite image for 28 October 2012 (12:00 UTC) showing water vapour data (derived from geostationary satellite datasets and MODIS imagery) overlaid on sea surface temperature (Reynolds, 1988; Reynolds and Marsico, 1993). Land features are superimposed for context from NASA's 'Blue Marble' global image, which was mosaiked from MODIS imagery acquired during 2001 with topographic shading based on the GTOPO30 elevation dataset (see Section 13.4). **Source:** Kohrs *et al.* (2014) Figure 36

10.1 Single Sensor Datasets

Rapid EO imaging from the GOES geostationary satellites has been available since 1975 (Purdom and Menzel, 2016) and animations made from these images have provided ongoing insights into the drivers and mechanisms of weather events (Purdom, 1976). Composite images from polar orbiting satellites emerged early in the 1970s, with composites derived from geostationary satellite images appearing during the Global Weather Experiment (GWE)[3] at the end of that decade.

More recently, image composites for land-based studies have been generated from publicly-available, image archives for various satellite sensors with near-daily, global coverage, including:

- AVHRR (Holben 1986; Cihlar *et al.*, 1994);
- MODIS (Roy, 2000; Justice *et al.*, 2002; Ju *et al.*, 2010);
- Landsat (White *et al.*, 2014; Roberts *et al.*, 2017; Griffiths *et al.*, 2019); and
- Sentinel-2 (Griffiths *et al.*, 2019).

Pixel compositing techniques were first proposed to create regularly-updated, cloud-free images of specific features at regional and global scales. Sensors with large area coverage and rapid revisit frequency, such as AVHRR and MODIS, were primarily used to create such composites, especially since these sensors were carried by multiple satellites and thus delivered a higher temporal density of images. Such sensors achieve large area coverage by using a wide scan angle, although this viewing geometry also presents challenges relating to pixel size and placement in the composited images (see Section 3.1.1.1). From these dense time series, frequent image composites have been generated on a routine basis, to provide daily, weekly and/or monthly coverages (see Table 10.1). These global composite products provide valuable input to a wide range of studies of the Earth's surface and its atmosphere, including weather forecasting, climate modelling, pollution monitoring, vegetation productivity, water usage, carbon cycling and oceanography (see Volume 3).

More recently, with the availability of free Landsat images, more detailed composites have been generated using increasingly sophisticated metrics and corrections (Woodcock *et al.*, 2008; Zhu *et al.*, 2019). The frequency of Landsat Images is most appropriate for seasonal, annual or epoch composites, based on either surface reflectance or a vegetation index (Roy *et al.*, 2010; White *et al.*, 2014).

Below we will review some of the most commonly used methods for generating pixel-based composites from imagery acquired by a single sensor. Features of these methods are introduced in terms of selecting the:

- time period for compositing (see Section 10.1.1); and the
- values of pixels included in the composite image (see Section 10.1.2).

10.1.1 Time period

Compositing techniques are applied to time series of calibrated image data to generate a single representative dataset (Wolfe *et al.*, 1998). These techniques either apply to multiple orbits of imagery that were sensed on the same day, or multiple orbits sensed over several hours, days, weeks or months. A sequence of composite images can also be used to assess change on a seasonal or annual basis (White *et al.*, 2014).

The goal of the image compositing process is to populate the final composite with the 'best available pixel' (BAP) values that are located within each grid cell. The sequence of images from which BAP values are selected can include overlap regions that occur between acquistion paths for satellite sensors. Given the increased overlap of satellite paths closer to the poles, higher latitude regions typically have a larger number of images available for compositing and would potentially be able to produce more frequent image composites (Long *et al.*, 1999).

The time period selected for creating an image composite needs to be appropriate for the feature being investigated. For example, a wide range of time periods are used to composite the various land products derived from MODIS imagery (see Excursus 10.1). While some products, such as active fires, need to be up-to-date, land cover varies more slowly so can be described in annual, or longer, intervals (see Table 10.1). Meteorological products, such as brightness temperature, are typically composited within intervals less than one hour (Kohrs *et al.*, 2014).

The selected time period also needs to make sense in terms of the revisit frequency of the selected sensor and seasonal cycles that might impact image utility. Basically, to achieve a useful composite, there needs to be a sufficient number of 'good' pixels in the set of candidate images acquired during the compositing period.

3 Originally known as the First Global Atmospheric Research Program (GARP) Experiment (FGGE)

Excursus 10.1—MODIS Composites

Source: *https://lpdaac.usgs.gov/dataset_discovery/modis*

MODIS (Moderate Resolution Imaging Spectroradiometer) is a sensor carried by the Terra and Aqua satellites. Its swath width of 2,330 km enables the full globe to be imaged each day or two in 36 spectral bands at 250 m, 500 m or 1 km pixel resolutions (see Volume 1). In addition to the raw image data, a range of image products are generated by USGS to simplify access to the MODIS archive.

Since 2000, up to four MODIS images have been acquired daily over most of the globe. A wide range of image composite products has been derived from the MODIS time series (see Table 10.1). The MOD13Q1 package, for example, summarises this archive into a series of composite images for two spectral indices (Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI); see Volumes 2C and 3A) by selecting the 'best' pixels from all MODIS Terra images that were acquired within a 16-day window, after correction for atmospheric and Sun angle effects. In this case, the 'best' pixels are determined using the criteria of high NDVI and close-to-zenith sensor view angle. Spatially adjacent pixels from a single image are used where possible. As a result of the compositing process, however, the MOD13Q1 images can contain pixels that have been imaged from different view angles. The impact of view angle differences within composite datasets has been the subject of recent and ongoing research in Australia (e.g. Gill *et al.*, 2009 and Bhardari *et al.*, 2011).

Table 10.1 MODIS land products

MODIS product codes denote the satellite that acquired the image data by the letters 'OD' for Terra, 'YD' for Aqua and 'CD' forTerra or Aqua. In the product codes listed in column 2 below, these codes should be substituted for 'xx'.

Product	Product code	Spatial resolution (m)	Composites available					
			Daily	4-day	8-day	16-day	Monthly	Annual
Reflectance	Mxx09/19	250/ 500/ 1000/ 5600	✓	×	✓	×	×	×
Temperature, Emissivity	Mxx11/21	1000/ 5600	✓	×	✓	×	✓	×
Land cover	Mxx12C1/Q1	500/ 5600	×	×	×	×	×	✓
Vegetation Indices (NDVI and EVI)	Mxx13	250/ 500/ 1000/ 5600	×	×	×	✓	✓	×
Thermal Anomalies and Fire	Mxx14	1000	✓	×	✓	×	×	×
Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR)	Mxx15A2H/3H	500	×	✓	✓	×	×	×
Evapotranspiration	Mxx16	500	×	×	✓	×	×	✓
Gross/Net Primary Productivity (GPP/NPP)	Mxx17A2H/3H	500	×	×	✓	×	×	✓
Radiance	Mxx18A1/2	5600	✓	×	×	×	×	×
Aerosol Optical Depth	Mxx19A2	1000	✓	×	×	×	×	×
Bidirectional Reflectance Distribution Function (BRDF) and Albedo	Mxx43	500	✓	×	×	×	×	×

Source: NASA MODIS Data Products: *https://modis.gsfc.nasa.gov/data/dataprod/index.php*

MODIS EVI was developed to be less sensitive to the colour of soil background, atmospheric path radiance and saturation in high biomass vegetation than NDVI (Huete *et al.*, 2002). Variations within vegetation index (VI) time series reflect both the seasonal growth phases (or phenology) and irregular perturbations, such as changes in plant health or the impact of fire. Bhardari *et al.* (2011) observed both amplitude and phase differences between the seasonal response cycles in NDVI and EVI time series, with EVI being characterised by an earlier response. Being computed from both blue and red reflectance data, it was proposed that EVI was impacted by both albedo and greenness. The suitability of NDVI and EVI composites for vegetation studies are further discussed in Volume 3A. MODIS vegetation index products have been correlated with a range of ecological parameters, including photosynthetic potential and ecosystem structure and function in Australian evergreen forests (Restrepo-Coupe *et al.*, 2016; see Volume 3A).

The Lands Processes Distributed Active Archive Centre (LP DAAC), a partnership between USGS and NASA, has processed, archived and freely distributed NASA global EO products for land applications since 1990. Some commonly used MODIS Land Products are summarised in Table 10.1. Other MODIS products include:

- Level-1 and atmosphere products—distributed by LAADS (L1 and Atmosphere Archive and Distribution System);
- cryosphere data products—distributed by NSIDC DAAC (National Snow and Ice Data Centre Distributed Active Archive Centre); and
- ocean colour and sea surface temperature products—distributed by the Ocean Color Web (see Section 10.3).

10.1.2 Composite values

For each pixel in the time series grid, the 'composite' value is either based on a 'rule' or a statistic (Roberts *et al.*, 2017). A rule-based approach defines criteria that select the 'best' value from those acquired in the nominated time period (see Section 10.1.1), whereas a statistics-based approach computes a summary statistics (such as the mean) from the multiple pixel observations available for each grid cell. The goal here is to produce a single image which contains the 'best selection of clear-sky pixels' available within the time series (Latifovic *et al.*, 2015) and, to avoid potential imaging artefacts, preferring those observations that were acquired with near-nadir viewing angle (Wolfe *et al.*, 1998). The compositing methods, however, may inadvertently bias selection of pixels towards those that are not necessarily representative, which can impact subsequent analyses (Park *et al.*, 2005).

10.1.2.1 Rule-based criteria

A range of rule-based criteria have been proposed over the past few decades to select the 'best' pixel to use as the composite value (see Table 10.2). One or more rules are applied individually, or in ordered sets, to the time series of pixel values in each grid cell, in order to select those pixel observations to use as the composite values.

Table 10.2 Popular criteria for selecting 'best' pixel

Factor	Criteria	Reason	Reference
Atmospheric conditions	Distance to cloud and cloud shadow	Avoid using pixels that are close to cloud or cloud shadow	Potapov *et al.* (2011) Griffiths *et al.* (2013)
	Atmospheric conditions	Avoid pixels acquired in hazy conditions	Masek *et al.* (2006)
Surface cover	Maximum greenness	Greenness values are lower for cloud, poor atmospheric conditions and off-nadir viewing angles, but can be problematic over water and non-vegetated, or sparsely-vegetated, surfaces	Cihlar *et al.* (1994) Roy *et al.* (2011)
	Maximum brightness temperature	Reduces limb darkening effects and useful for non-vegetated pixels	Kohrs *et al* (2014) Roy *et al.* (2011)
Sensor characteristics	Minimum scan angle	Selects pixels acquired closer to nadir viewing angle	Cihlar *et al.* (1994) Huete *et al.* (2002)
	Preferred sensor	Can avoid known imaging artefacts (such as Landsat-7 ETM+ SLC-off data)	White *et al.* (2014)
Timing	Preferred time or date	Favours pixels acquired closer to the middle of compositing period or in a particular season	White *et al.* (2014) Griffiths *et al.* (2013)

In conjunction with using Analysis Ready Data (ARD; see Section 3.2), most current, rule-based criteria include atmospheric corrections and cloud-masking (see Section 3.1.1.2), and define preferences for sensor characteristics and timing of image acquisition. For example, to generate annual BAP composites over Canada from the Landsat image time series, White *et al.* (2014) used rules that allocated scores for:

- proximity of image to target date-of-year (DOY; Griffiths *et al.*, 2013);
- distance from pixel to cloud or cloud shadow (Griffiths *et al.*, 2013);
- atmospheric opacity (Kaufman *et al.*, 1997; Masek *et al.*, 2006); and
- preferred sensor.

The sum of these scores was then used to select the best pixel value for each grid cell.

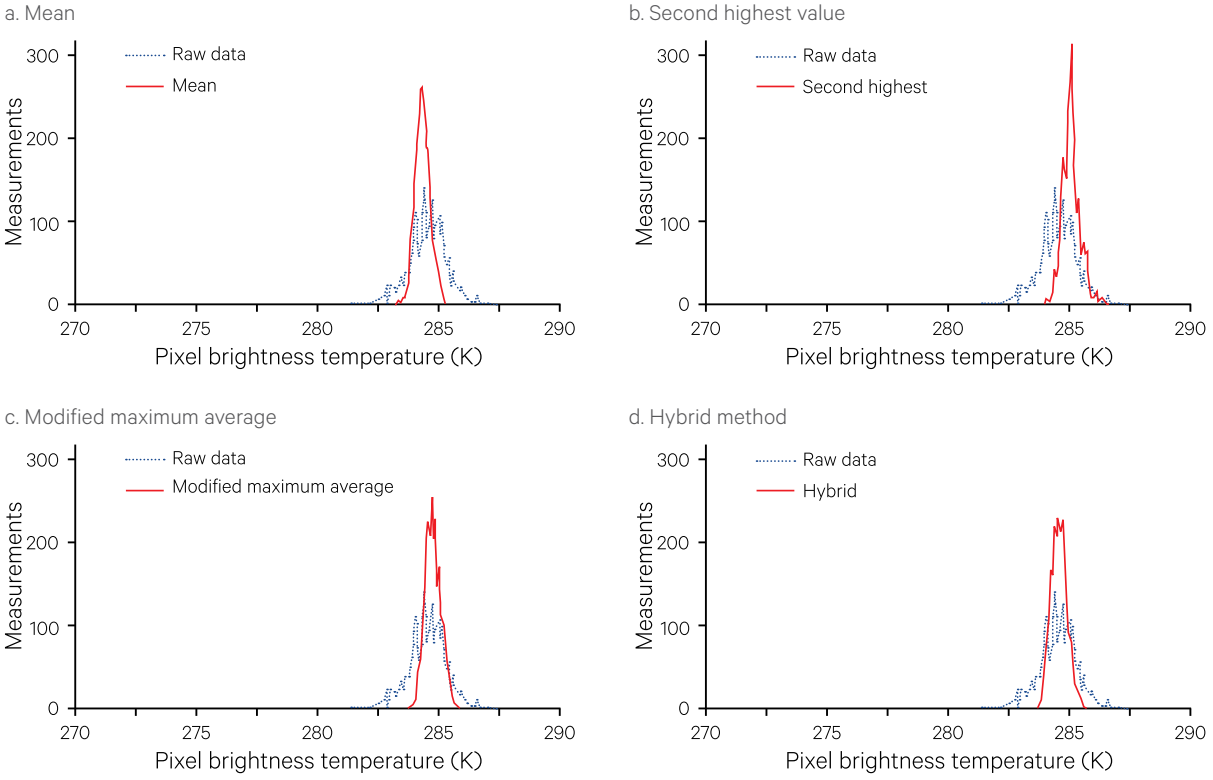
While many rule-based criteria for selecting the best pixel for image composites are based on spectral values (see Table 10.2), Tan *et al.* (2006) caution that such criteria can introduce bias, especially with spectral indices such as NDVI, potentially resulting in the preferential selection of vegetated pixels (Wolfe *et al.*, 1998). Some of the implications of gridding artefacts in MODIS image composites are introduced in Section 3.1.1.1 above, and such artefacts have also been reported to bias certain selection criteria, especially for data from sensors with wide scan angles (Tan *et al.*, 2006). Similarly, the maximum brightness temperature algorithm, developed to reduce limb darkening in geostationary images, results in a warm bias in the composited image due to parallax differences between overlapping input images (Kohrs *et al.*, 2014). When accurate cloud masking methods are available (such as Fmask: Zhu and Woodcock, 2012, 2014; Zhu *et al.*, 2015), spectrally independent criteria such as minimum view zenith angle (to select pixels closest to nadir) render least bias in the resulting image composite (Tan *et al.*, 2006).

10.1.2.2 Statistical criteria

A range of summary statistics can be generated from the available pixel observations to use as the composite value (Vancutsem *et al.*, 2015; Griffiths *et al.*, 2019). For example, Long *et al.* (1999) compared surface brightness temperature composites derived from SSM/I imagery, which were based on three statistics (mean, second-highest value, and a modified maximum average) with a hybrid compositing algorithm, and concluded that while each statistic offered both advantages and disadvantages in different situations, the hybrid algorithm combined the strengths of the best statistics (see Figure 10.1). Synthetic Landsat composites have also been derived by fitting harmonic models to all available data (Zhu *et al.*, 2015).

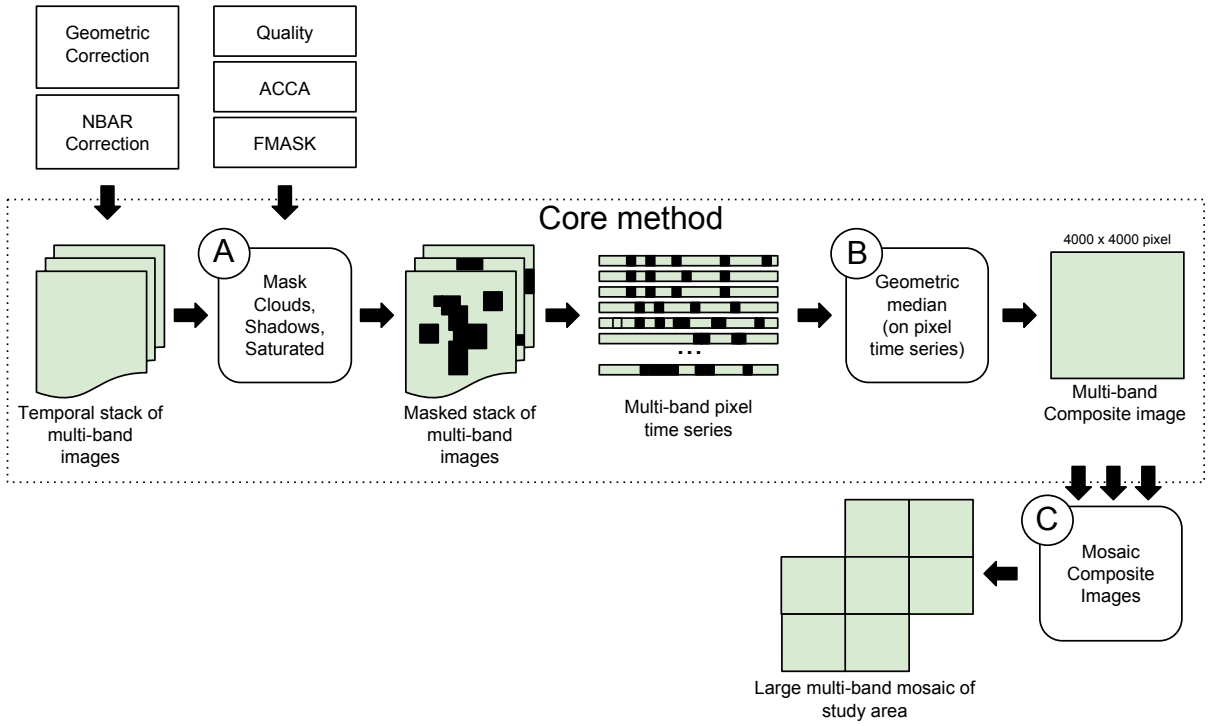
Figure 10.1 Brightness Temperature Composites from SSM/I

SSM/I 85 GHz vertical channel clear region spatial brightness temperature distributions are shown for four compositing statistics.



Source: Long *et al.* (1999) Figure 13

Figure 10.2 Workflow for Australian Landsat composite images



Source: Roberts *et al.* (2017) Figure 2

Using statistical criteria for multichannel imagery, however, requires consideration of the relationships between spectral bands. Although it is a simple process to compute such statistics for individual image bands, and the derived composite bands may recombine into a visually appealing colour composite image, the relationships between the resulting composite bands will not correctly represent those of the input images (Roberts *et al.*, 2017). This is particularly important where subsequent processing will be applied to the composite image. Examples of high-dimensional summary statistics that have been developed to avoid this problem include the:

- medoid (Flood, 2013), which calculates the median then selects the nearest actual observation to create a seasonal composite; and the
- geometric (or spatial) median (Roberts *et al.*, 2017), which is being used to generate annual continental composites from the Landsat archive for Australia (see Figure 10.2).

However, while statistics-based algorithms for compositing produce images that appear homogeneous, the pixel values in the composite image do not correspond with actual observations acquired by the sensor, which may undesirable for some applications.

10.1.2.3 Proxy values

A valid pixel value may be not available for a given grid cell within the images included in the compositing period, due to either missing or anomalous data. Missing data can result from sensor acquisition malfunction or cloud-masking, while anomalous data includes unexpected values, possibly caused by atmospheric haze, sunglint or sensor aberrations. Anomalous pixels may be identified by tracking the temporal trajectory of each pixel over time to highlight those values that fall outside of a pre-defined range (see Section 8.4).

In some compositing algorithms, a 'proxy' value may be assigned to missing and anomalous pixels. For pixels with 'stable' temporal trajectories, a suitable value may be determined using appropriate statistics derived from the pixel values in that trajectory. For example, the MODIS annual composite of Land Cover Dynamics (MCD12Q2 v006) product specifies each data gap is filled using a 'good quality' value of that cell in the preceding or subsequent year (USGS, 2019). However, for pixels deemed to be not temporally stable, the proxy value would be based on an understanding of change events in the trajectory of that pixel (see Section 9), and possibly the trajectories of neighbouring pixels, and it may be appropriate to derive the proxy from the pixel values near the missing grid cell location (White *et al.*, 2014). At higher latitudes with cloudy skies, however, this approach can still be problematic if 'good' data is persistently elusive over several years.

10.2 Multiple Sensor Datasets

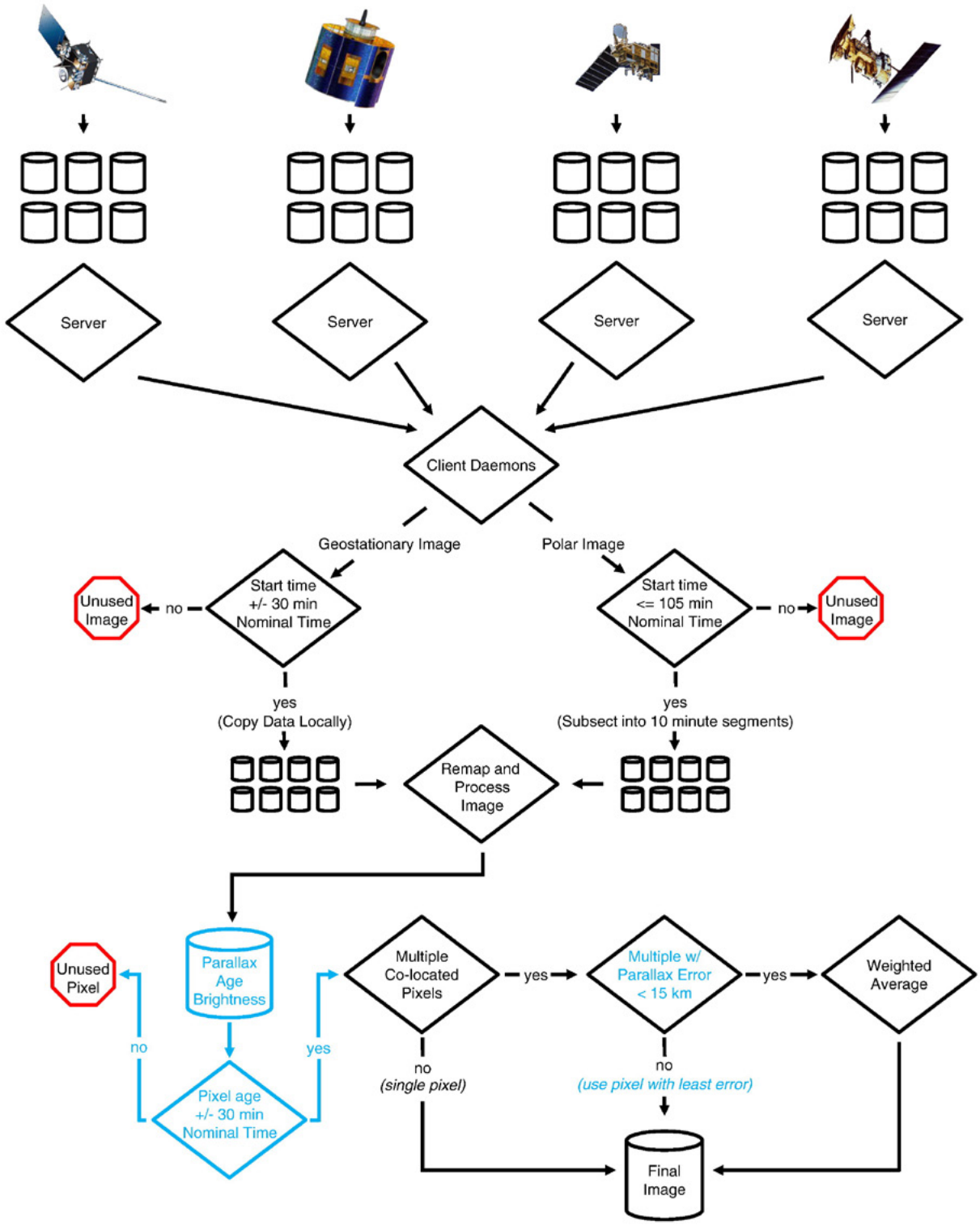
Technological advances have enabled faster and more extensive multisensor composites from a greater range of meteorological datasets with higher resolutions in all imaging dimensions (see Table 2.1). Datasets are operationally sourced from both geostationary and polar orbiting satellites to create composites with continental, hemispheric, polar and global coverage. For example, GMGSI (Global Mosaic of Geostationary Satellite Imagery) remaps GOES, METEOSAT and Himawari-8 datasets to 8 km spatial resolution every three hours (NOAA, 2016).

Multisensor composites have been routinely produced over the Southern Ocean and Antarctic continent since 1992 (Lazzara *et al*., 2003; see University of Wisconsin entry in Section 10.3). With increasing computing power, advanced algorithms for mosaiking and compositing have been developed to select between the available co-located pixels (see Section 10.1.2). In the early 1990s, three hours of processing was required to create global composites from six satellite datasets at 20 km resolution, whereas by 2014, composites were generated every 15 minutes from 14 satellite datasets at 1 km resolution for five spectral bands (Kohrs *et al*., 2014). An overview of the compositing process in 2014 is shown in Figure 10.3.

Multisensor ARD products are introduced in Section 3.2. These advanced time series enable denser temporal coverage that is suitable for image compositing of medium resolution EO imagery. For example, the HLS dataset (Claverie *et al*., 2018; Skakun *et al*., 2018) derived from Sentinel-2 MSI and Landsat-8 OLI has been composited to create the M30 product. M30 delivers surface reflectance data in Landsat OLI spectral bands with 5 day temporal frequency at 30 m spatial resolution over the Sentinel-2 tiling system. Another compositing approach for the HLS dataset has been proposed by Griffiths *et al*. (2019) for crop and land cover mapping. This area of EO image processing is changing rapidly and will, no doubt, generate a range of streamlined products, delivered in near real time, in the coming years.

Figure 10.3 Flow chart for multisensor composites

The University of Wisconsin Space Science and Engineering Centre (SSEC) routinely generates multisensor composites from five satellite sensors: GOES-W (NOAA-15), GOES-E (NOAA-16), MSG (NOAA-19), FY2E (Metop B) and MTSAT. This diagram summarises the data flow from ingest on separate workstations for each satellite, through a series of tests to select the best pixels for each composite image.



Source: Kohrs *et al.* (2014) Figure 22

10.3 Further Information

Image Products:

NOAA Composite Satellite Imagery: *https://www.ospo.noaa.gov/Products/imagery/composite.html*

MODIS: *https://modis.gsfc.nasa.gov/data/dataprod/*
Land: *https://lpdaac.usgs.gov/product_search/?collections=Combined+MODIS&collections=Terra+MODIS&collections=Aqua+MODIS&view=list*
Ocean Colour Web: *https://oceancolor.gsfc.nasa.gov*
Atmosphere: *https://ladsweb.modaps.eosdis.nasa.gov/search/order/1/MODIS:Terra,MODIS:Aqua*

ASTER: *https://asterweb.jpl.nasa.gov*

Landsat ARD:
Australia (DEA): *https://www.ga.gov.au/dea/products*
USA: *https://www.usgs.gov/land-resources/nli/landsat/us-landsat-analysis-ready-data?qt-science_support_page_related_con=0#qt-science_support_page_related_con*

Harmonised Landsat Sentinel-2 (HLS): *https://hls.gsfc.nasa.gov*

University of Wisconsin: *https://www.ssec.wisc.edu/data/composites/*

UN FAO Global Vegetation/Precipitation Indicators: *http://www.fao.org/giews/earthobservation/asis/index_2.jsp?lang=en*

ISS Image Composites: *https://eol.jsc.nasa.gov/Collections/Composites/*

10.4 References

Bhandari, S., Phinn, S., and Gill, T. (2011). Assessing viewing and illumination geometry effects on the MODIS vegetation index (MOD13Q1) time series: Implications for monitoring phenology and disturbances in forest communities in Queensland, Australia. *International Journal of Remote Sensing*, 32, 7513–7538.

Cihlar, J. (2000). Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing*, 21(6), 1093–1114.

Cihlar, J., Manak, D., and D'Iorio, M. (1994). Evaluation of compositing algorithms for AVHRR data over land. *IEEE Transactions on Geoscience and Remote Sensing*, 32, 427–437.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., and Justice, C. (2018). The Harmonised Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219, 145–161. *https://doi.org/10.1016/j.rse.2018.09.002*

Flood, N. (2013). Seasonal composite Landsat TM/ETM+ images using the medoid (a multi-dimensional median). *Remote Sensing*, 5(12), 6481–6500.

Gill, T.K., Phinn, S.R., Armston, J.D., and Pailthorpe, B.A. (2009) Estimating tree cover change in Australia: challenges of using the MODIS vegetation index product. *International Journal of Remote Sensing*, 30(6), 1547–1565, *https://doi.org/10.1080/01431160802509066*

Griffiths, P., van der Linden, S., Kuemmerle, T., and Hostert, P. (2013). A pixel based Landsat compositing algorithm for large area land cover mapping. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(5), 2088–2101.

Griffiths, P., Nendel, C., and Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment*, 220, 135–151. *https://doi.org/10.1016/j.rse.2018.10.031*

Hansen, M.C.and Loveland, T.R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122, 66–74.

Holben, B. (1986). Characteristics of maximum value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7, 1417–1434.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213.

Ju, J., Roy, D.P., Shuai, Y., and Schaaf, C. (2010). Development of an approach for generation of temporally complete daily nadir MODIS reflectance time series. *Remote Sensing of Environment*, 114, 1–20. *https://doi.org/10.1016/j.rse.2009.05.022*

Justice, C.O., Townshend, J.R.G., Vermote, E.F., Masuoka, E., Wolfe, R.E., Saleous, N., Roy, D.P., and Morisette, J.T. (2002). An overview of MODIS land data processing and product status. *Remote Sensing of Environment,* 83, 3–15.

Kaufman, Y.J., Tanre, D., Remer, L.A., Vermote, E.F., Chu, A., and Holben, B.N. (1997). Operational remote sensing of tropospheric aerosol over the land from EOS MODIS. *Journal of Geophysical Research Atmosphere*, 102, 17051–17068

Kohrs, R.A., Lazzara, M.A., Robaidek, J.O., Santek, D.A., and Knuth, S.L. (2014). Global satellite composites—20 years of evolution. *Atmospheric Research*, 135–136, 8–34. *http://dx.doi.org/10.1016/j.atmosres.2013.07.023*

Latifovic, R., Pouliot, D., Sun, L., Schwarz, J., and Parkinson, W. (2015). *Moderate Resolution Time Series Data Management and Analysis: Automated Large Area Mosaicking and Quality Control*. Geomatics Canada, Open File 6, 25 p. *https://doi.org/10.4095/296204*

Lazzara, M.A., Keller, L.M., Stearns, C.R., Thorn, J.E., and Weidner, G.A. (2003). Antarctic Satellite Meteorology: Applications for Weather Forecasting. *Monthly Weather Review*, 131, 371–383. *https://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%282003%29131%3C0371%3AASMAFW%3E2.0.CO%3B2*

Long, D.G., Remund, Q.P., and Daum, D.L. (1999). A Cloud-Removal Algorithm for SSM/I Data. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1), 54–62.

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., and Lim, T.K. (2006). A Landsat surface reflectance data set for North America, 1990 2000. *IEEE Geoscience and Remote Sensing Letters*, 3 (1), 68–72.

NOAA (2016). Composite Satellite Imagery webpage. NOAA Office of Satellite and Product Operations webpage: *https://www.ospo.noaa.gov/Products/imagery/composite.html*

Park, S., Feddema, J.J., and Egbert, S.L. (2005). MODIS land surface temperature composite data and their relationships with climatic water budget factors in the central Great Plains. *International Journal of Remote Sensing*, 26(6), 1127–1144.

Potapov, P., Turubanova, S., and Hansen, M.C. (2011). Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia, *Remote Sensing of Environment*, 115(2), 548–561.

Purdom, J.F.W. (1976). Some Uses of High-Resolution GOES Imagery in the Mesoscale Forecasting of Convection and Its Behaviour. *Monthly Weather Review*, 104, 1474–1983. *https://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281976%29104%3C1474%3ASUOHRG%3E2.0.CO%3B2*

Purdom, J.F.W., and Menzel, W.P. (2016). Evolution of Satellite Observations in the United States and Their Use in Meteorology. Ch 5 in *Historical Essays on Meteorology, 1919–1995*. (Ed: Fleming, J.) American Meteorology Society. 617 pp.

Restrepo-Coupe, N., Huete, A., Davies, K., Cleverly, J., Beringer, J., Eamus, D., van Gorsel, E., Hutley, L.B., and Meyer, W. (2016). MODIS vegetation products as proxies of photosynthetic potential along a gradient of meteorologically and biologically driven ecosystem productivity, *Biogeosciences*, 13, 5587–5608, *https://doi.org/10.5194/bg-13-5587-2016*

Reynolds, R.W. (1988). A real-time global sea-surface temperature analysis. *Journal of Climate*, 1, 75–86.

Reynolds, R.W., and Marsico, D.C. (1993). An improved real-time global temperature analysis. *Journal of Climate*, 6, 114–119.

Roberts, D., Mueller, N., and McIntyre, A. (2017). High-dimensional Pixel Composites From Earth Observation Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, 35(11), 6254–6264. *https://doi.org/10.1109/TGRS.2017.2723896*

Roy, D.P. (2000). The impact of misregistration upon composited wide field of view satellite data and implications for change detetcion. *IEEE Transactions on Geoscience and Remote Sensing*, 38, 2017–2032.

Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., Zhang, C. (2010). Web enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, 114, 35–49. *http://qa4eo.org/docs/case-studies/NASA_WELD_115890_v2.pdf*

Roy, D., Ju, J., Kommareddy, I., Hansen, M., Vermote, E., Zhang, C., and Kommareddy. A. (2011). Web Enabled Landsat Data (WELD) Products—Algorithm Theoretical Basis Document, February, 2011. *https://globalmonitoring.sdstate.edu/projects/weld/WELD_ATBD.pdf*

Skakun, S., Ju, J., Claverie, M., Roger, J.-C., Vermote, E., Franch, B., Dungan, J.L., and Masek, J.G. (2018). Harmonized Landsat Sentinel-2 (HLS) Product User's Guide. Version 1.4. NASA. *https://hls.gsfc.nasa.gov/wp-content/uploads/2019/01/HLS.v1.4.UserGuide_draft_ver3.1.pdf*

Tan, B., Woodcock, C.E., Hu, J., Zhang, P., Ozdogan, M., Huang, D., Yang, W., Knyazikhin, Y., and Myneni, R.B. (2006). The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions, *Remote Sensing of Environment*, 105, 98–114.

USGS (2019). *MCD12Q2 v002* webpage: *https://lpdaac.usgs.gov/products/mcd12q2v006/*

Vancutsem, C., Pekel, J.F., Bogaert, P., and Defourny, P. (2007). Mean compositing, an alternative strategy for producing temporal syntheses. Concepts and performance assessment for SPOT VEGETATION time series. *International Journal of Remote Sensing*, 28, 5123–5141. *https://doi.org/10.1080/01431160701253212*

White, J.C., Wulder, M.A., Hobart, G,W., Luther, J.E., Hermosilla, T., Griffiths, P., Coops, N.C., Hall, R.J., Hostert, P., Dyk, A., Guindon. L. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, 40(3), 192–212. *http://dx.doi.org/10.1080/07038992.2014.945827*

Wolfe, R.E., Roy, D.P., and Vermote, E. (1998). MODIS Land Data Storage, Gridding, and Compositing Methodology: Level 2 Grid. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4), 1324–1338.

Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., Wynne, R. (2008). Free access to Landsat imagery. *Science,* 320, 1011.

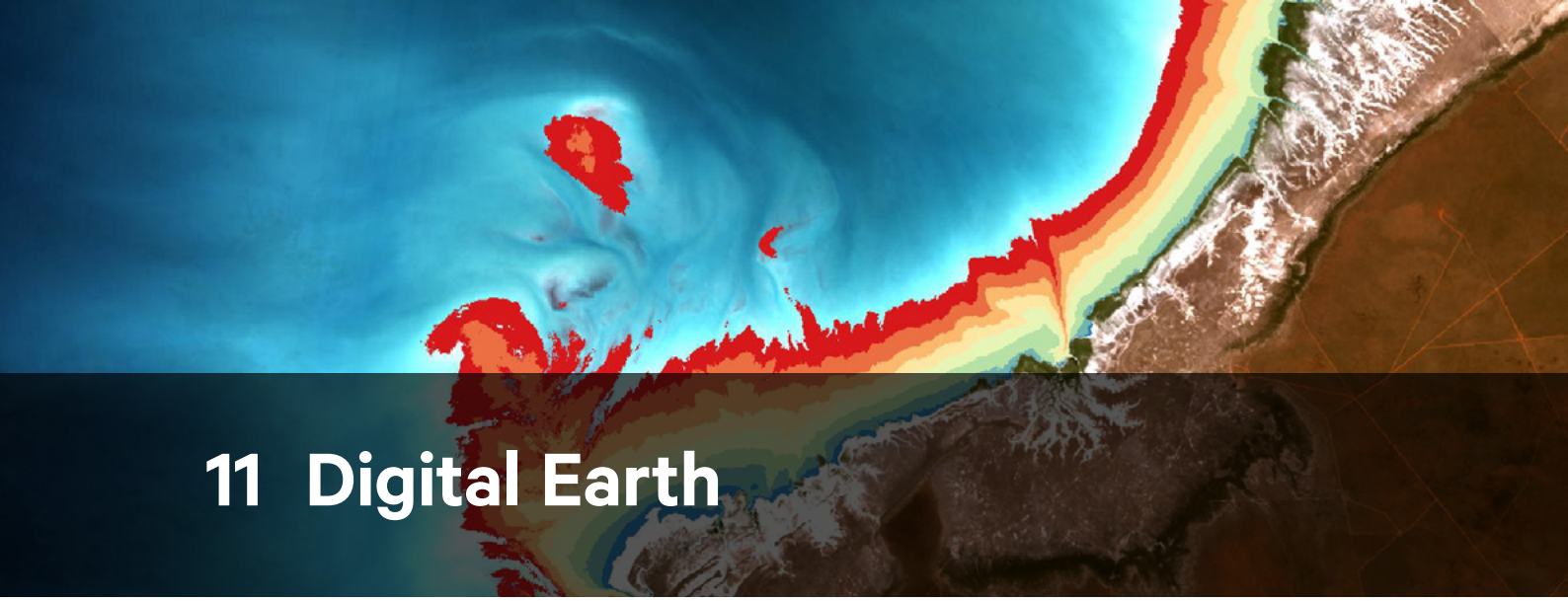
Zhu, Z., and Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83–94. *https://doi.org/10.1016/j.rse.2011.10.028*

Zhu, Z., and Woodcock, C.E. (2014). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152, 217–234. *https://doi.org/10.1016/j.rse.2014.06.012*

Zhu, Z., Wang, S., and Woodcock, C.E. (2015). Improvements and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8 and Sentinel 2 imags. *Remote Sensing of Environment*, 159, 269–277. *https://doi.org/10.1016/j.rse.2014.12.014*

Zhu, Z., Woodcock, C.E., Holden, C., and Yang, Z. (2015). Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sensing of Environment*, 162, 67–83. *https://doi.org/10.1016/j.rse.2015.02.009*

Zhu, Z., Wulder, M.A., Roy, D.P., Woodcock, C.E., Hansen, M.C., Radeloff, V.C., Healey, S.P., Schaaf, C., Hosteret, P., Strobl, P., Pekel, J.-F., Lymburner, L., Pahlevan, N., and Scambos, T.A. (2019). Benefits of the Free and Open Landsat Data Policy. *Remote Sensing of Environment*, 224, 382–385. *https://doi.org/10.1016/j.rse.2019.02.016*



11 Digital Earth

11.1 The Concept

The term 'Digital Earth' (DE) was coined in 1998 as part of a vision for the digital future, in which all citizens of the world could freely interact with volumes of geo-referenced information presented on a virtual globe (Gore, 1998). While technological advances have changed the original vision of this concept, the value of real time access to current and historical information about environmental, economic and social conditions on our planet is unprecedented. Such information can now be derived from myriad data sources, including EO.

A succession of technological and global advances have enabled development of the DE (see Figure 11.1). To realise the goal of easy accessibility by ordinary citizens, a range of technical challenges needed to be negiotated, including visualisation techniques, intuitive user interfaces, global data standards, fast data transfer rates and efficient storage mechanisms, as well as selection of the most relevant and reliable sources of information. The integration of a diverse range of technological and organisational challenges has required collaboration and coordination at global levels between governments and private enterprises (Foresman, 2008; Craglia *et al*., 2012; Goodchild *et al*., 2012; Mattmann, 2013).

As the world's population and resource consumption both continue to grow, sustainable development is a major concern. Our finite resources of energy, food and water need to be shared equitably and efficiently. To support sustainable development around the globe, the DE framework needs to integrate data and models describing economic, social and environmental systems and conditions. In 2009–2010, the International Council for Science (ICSU) canvassed the views of over one thousand scientists to identify five priorities, or Grand Challenges facing research into global sus.tainability:

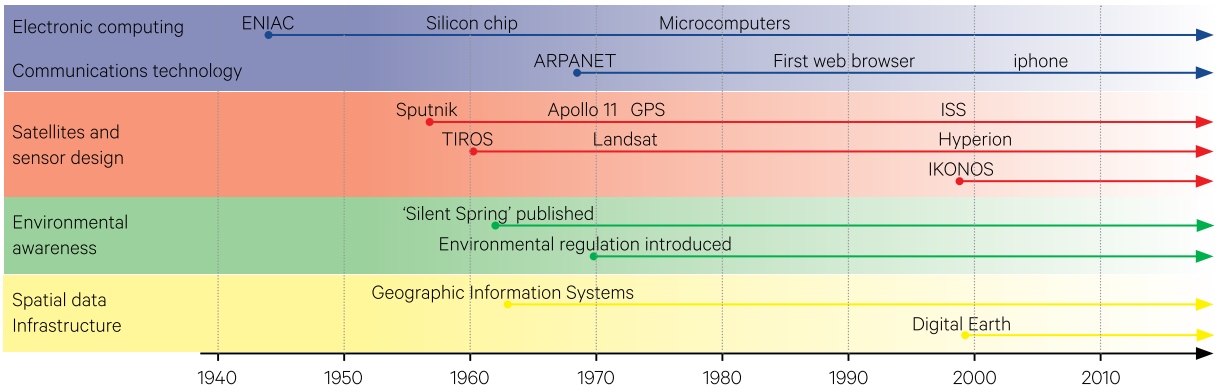
- developing the ***observation systems*** needed to manage global and regional environmental change;
- improving the usefulness of ***forecasts*** of future environmental conditions and their consequences for people;
- recognising key ***thresholds*** or non-linear changes to improve our ability to anticipate, recognise, avoid and adapt to abrupt global environmental change;
- determining what institutional, economic and behavioural ***responses*** can enable effective steps towards global sustainability; and
- encouraging ***innovation*** (coupled with sound mechanisms for evaluation) in developing technological, policy and social responses to achieve global sustainability (Craglia *et al*., 2012).

With greater awareness of pollution and environmental degradation, the DE provides an integrated platform for identifying key drivers of changes, and monitoring those changes at a range of spatial and temporal scales, with a view to minimising the adverse impacts of human civilisation on our planet into the future.

The ICSU also advocated the need for more effective communication of science to both citizens and policy makers, and emphasised the advantages of involving citizens in the scientific process of environmental monitoring (Craglia *et al*., 2012). One interesting outcome of DE is the two-way exchange of information with citizens. Not only are the Earth's residents able to download huge volumes of data for professional and personal use, but the geo-referenced observations and opinions of citizens can be uploaded to complement professional surveys related to natural disasters, environmental conditions, and a wide range of other concerns. This opportunity has particular relevance to EO for involving citizen scientists and crowdsourcing with ground truthing exercises in remote regions (see Section 12.4; also Volume 1A—Section 10.3 and Volume 3B).

Background image: Landsat-8 OLI image of Roebuck Bay, WA, overlaid with tidal regions defined by the Intertidal Extents Model (ITEM; see Volume 2A—Excursus 2.1). **Source:** Steven Sagar, Geoscience Australia

Figure 11.1 Technological and cultural advances enabling Digital Earth

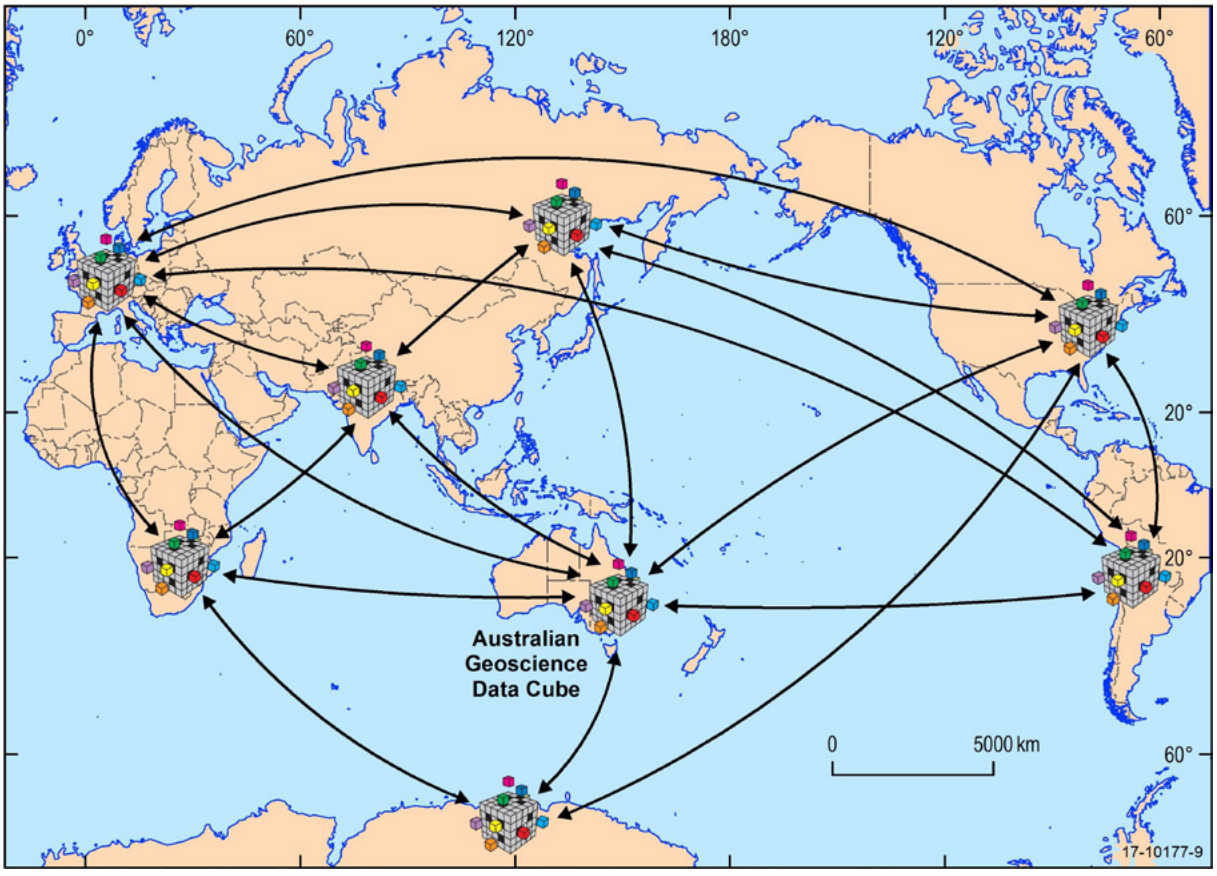


Adapted from Foresman (2008) Figure 1

As the array of sensors observing our planet—and our lives—increases (see Volume 1B—Section 10), so does the need for international cooperation, and better communication between data custodians and data users. The Sensor Web, including Volunteered Geographic Information (VGI) from citizens, has been likened to the DE nervous system—a mechanism to actively inform residents on Earth about current (and potential) events (De Longueville *et al.*, 2010). While the technological challenges of the sensor web increase with the volume of sensors, so do the expectations of users—data is expected to be reliable and relevant and its delivery is expected to be dynamic, interactive and engaging.

While originally envisioned as a virtual globe, and thus simplifying the interpretation of projected data (see Volume 2B), current realisations of the DE concept comprise a global network of data cubes (see Figure 11.2 and Section 8.3). The Australian implementation of a national data cube is introduced in Section 11.2.

Figure 11.2 Global network of data cubes



Source: Lewis *et al.* (2017) Figure 14

11.2 Digital Earth Australia

Originally called the Australian Geoscience Data Cube (AGDC) and rebadged as Digital Earth Australia (DEA) in 2016, this 'Big Data' infrastructure is a joint initiative between Geoscience Australia (GA), the National Computational Infrastructure (NCI) and CSIRO that 'aims to realise the full potential of Earth Observation data holdings for Australia' (Lewis *et al.*, 2017). The current Landsat archive managed by GA (dating from 1986), for example, comprises around 300,000 raw data images—or 10^{16} pixels of data. In addition to the raw data images, several continental-scale datasets have been derived from these Landsat images (see Section 14.1 and Volume 3), and Landsat is but one of many satellite series acquiring data over Australia (see Volume 1).

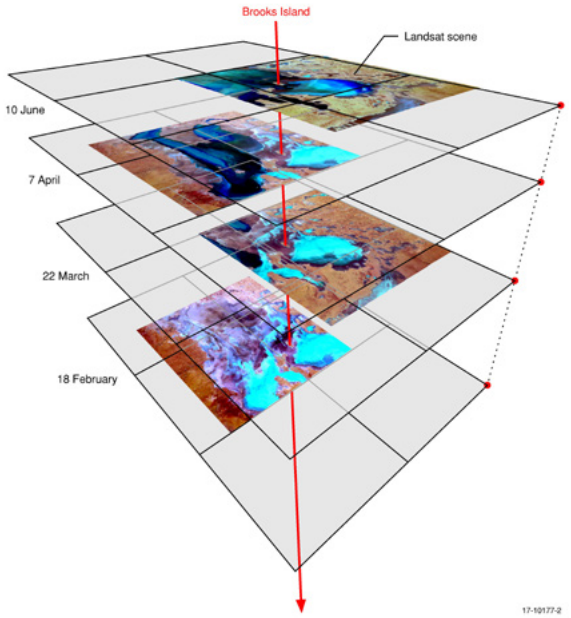
The existing archives of EO data clearly qualify the label of 'Big Data', having volume, velocity and variety (Laney, 2001), however, given the predicted increased volume of future sensors, particularly with higher revisit frequency, this data volume will increase significantly with time. As these volumes grow, so do the challenges for archiving and accessing the data efficiently. To meet these challenges, new methods of data analysis have been proposed that utilise the wealth of information contained within existing EO archives, including (Lewis *et al.*, 2017; Dhu *et al.*, 2017):

- detecting surface changes using time-series analysis techniques (Griffiths *et al.*, 2014; Kennedy *et al.*, 2010; Masek *et al.*, 2013; Zhu and Woodcock, 2014; see Section 9);
- systematically characterising land cover type(s) and climate indicators across multiple decades (Masek *et al.*, 2008; Mueller *et al.*, 2016; Sexton *et al.*, 2013; Hollman *et al.*, 2013; Kim *et al.*, 2015; see Volume 3A); and
- using 'best available pixel' composites, especially in cloudy locations (Hermosilla *et al.*, 2015; Thompson *et al.*, 2015; Zald *et al.*, 2016; see Section 10).

The DEA concept is illustrated in Figure 11.3 and comprises spatially and radiometrically consistent tiles. The original implementation was structured as regular, non-overlapping tiles of 4000´4000 pixels, with each tile spanning 1º by 1º, while in Version 2 the tile and pixel sizes are parameters that are set during data ingest. Version 2 also includes an indexing system for infrequently used data that directs processing functions to operate on the original data configuration, and thus avoids replication and associated problems of version control (Lewis *et al.*, 2016, 2017). The foundations for DEA products are summarised in Table 11.1.

Figure 11.3 DEA concept

Landsat scenes are reformatted as spatially-consistent tiles of data. While the spatial footprint of Landsat scenes changes over time, these tiles can maintain a constant footprint. This graphic shows four overlapping Landsat scenes over Brooks Island, Lake Eyre, acquired during 2009.



Source: Lewis *et al.* (2017) Figure 2

Digital Earth is a global initiative to construct a comprehensive virtual representation of the planet. It is a collaborative effort between Earth sciences, space sciences and information sciences to monitor and forecast natural and human phenomena.
(International Society for Digital Earth)

Table 11.1 Foundations of DEA

Level 1	Level 2
Protocols to prepare data and manage archives	Accurate geometric alignment so that pixels are comparable in space (see Section 3.1.1.1 and Volume 2B)
	Accurate radiometric correction so that pixel values are comparable in time and space (see Figure 11.5 and Volume 2A—Section 3)
	Pixel quality flags to indicate known shortcomings in pixel values (see Volume 2A—Section 1.5)
	Tracking provenance and data and products
Software environment to manage and access data	Acquisition and inflow of data to 'analysis ready' level (see Section 3.2)
	Data cube infrastructure to index available data (see Section 8.3)
	Data and application platform to generate routine products and explore new products
	User interface and application layer to enable a variety of applications
High performance computing (HPC)/high performance data (HPD) environment to store and access enormous data volumes	Infrastructure based on National Environmental Research Data Interoperability Platform (NERDIP; Evans et al., 2015)
	NetCDF4/HDF5 data formats suitable for transdisciplinary access and high performance access Lewis et al. (2017)

Some of the operational products DEA has enabled include:

- Land Surface Reflectance—for Landsat imagery from 1987 to present (25 m spatial resolution with 16 day revisit time; see Excursus 11.1);
- Water Observations from Space (WOfS)—a web service displaying historical surface water observations for all of Australia, derived from Landsat imagery from 1987 to the present day (Mueller *et al.*, 2016). WOfS aims is to allow a better understanding of:
  - where water is usually present;
  - where it is seldom observed; and
  - where inundation of the surface has been occasionally observed by satellite.

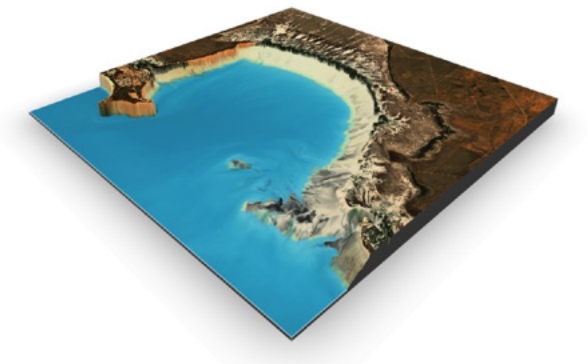
  This information is particularly valuable for flood modelling and assessment of environmental water flows (see Volume 1A—Excursus 5.1 and Volume 3B for details).
- Intertidal Extents Model (ITEM)—maps the extents and topography of Australia's intertidal mudflats, beaches and reefs, that is, the area exposed between high and low tide (Sagar *et al.*, 2017, 2018). This information contributes to coastal management, habitat mapping and modelling applications (see Volume 2A—Excursus 2.1 for details).
- National Intertidal Digital Elevation Model (NIDEM)—provides intertidal elevation data for Australia's coastline, which is informative for mapping coastal habitats, coastal erosion and storm and flood mitigation efforts (see Figure 11.4);
- High and Low Tide Composites (HLTC)—visualise Australia's coasts, estuaries and reefs at low and high tide, without the influence of noise features such as clouds, breaking water and sunglint (Sagar *et al.*, 2017). These datasets deliver a useful snapshot of Australia's coastline at different biophysical states.
- Fractional Cover (FC)—also derived from Landsat imagery from 1987 to the present day, this product maps the Australian landscape in terms of:
  - green (leaves, grass and growing crops);
  - brown (branches, dry grass or hay, and dead leaf litter); and
  - bare ground (soil or rock).

  This spatially detailed dataset monitors vegetation condition and extent through time, providing valuable insights for land managers (Scarth *et al.*, 2010; Stewart *et al.*, 2014; see Volume 3A); and
- Mangrove Dynamics—national maps of mangrove extent by canopy cover type have been generated annually from 1986 to 2016 for Australia (with 25 m spatial resolution; Lymburner *et al.*, 2018).

Combinations of these products provide fascinating insights into the dynamics of land and water environments (see Volume 3).

Figure 11.4 National Intertidal Digital Elevation Model

Tidal flats of Roebuck Bay, WA, visualised in three dimensions.



Excursus 11.1—GA Analysis Ready Data

Source: Lewis *et al.* (2017); GA (2015)

GA has acquired Landsat TM/ETM+ imagery over Australia for several decades and Landsat OLI imagery for several years. While this data has been used extensively for numerous land and coastal mapping studies, its utility for accurate monitoring of environmental resources has been limited by the processing methods that have been traditionally used to correct for the inherent geometric and radiometric distortions in EO imagery. To improve access to Australia's archive of Landsat TM/ETM+/OLI data, several collaborative projects have been undertaken (in conjunction with industry, government and academic partners), which have enabled implementation of a more integrated approach to image data correction. This approach is based on using physical models to correct for atmospheric effects, BRDF (Li *et al.*, 2010) and topographic shading (Li *et al.*, 2012) using DEM information (Li *et al.*, 2015) and has been applied to the complete Landsat archive to create the Land Surface Reflectance (LSR)[4] suite of products.

Across the Australian landmass and its coastal fringes, the LSR product suite provides standardised optical surface reflectance datasets, using robust physical models to correct for variations in image radiance values due to atmospheric properties, and Sun and sensor geometry. This suite of products aims to be available within a more general framework of nested grids with scalable cell sizes, for easy comparison between different products acquired by different sensors at varying spatial scales. The resulting stack of surface reflectance grids will effectively be consistent over space and time, and thus appropriate for identifying and quantifying environmental change.

The first product in this suite is the LSR25, a medium resolution (~25 m) grid based on the Landsat TM/ETM+/OLI archive and presents surface reflectance data in 25 m^2 grid cells. This product can be viewed as a data cube, with time as the third dimension (see Section 11.2). The surface reflectance data for each imaged date and time are stacked in chronological order, with Landsat TM/ETM+/OLI sources being interleaved as appropriate to their acquisition time. Interrogation of this data cube allows the time series for individual grid cells (as well as their heritage and data quality) to be extracted and compared.

To create a reliable data cube, all imagery in the time series need to be carefully calibrated to correct for radiometric and geometric distortions. The basic measurement in each LSR25 spectral band is Nadir BRDF Adjusted Reflectance (NBAR)—surface reflectance normalised for a nadir view with illumination at 45° elevation (Li *et al.*, 2010, 2012). The detailed procedures that are followed to create the LSR product are described below.

1. Geometric correction

Given that ancillary data sources are required for radiometric procedures (such as a digital surface model, atmospheric condition parameters, and BRDF parameters), geometric correction of imagery is a prerequisite to ensure that the ancillary data is correctly located in the image (Lewis *et al.*, 2017). This is most critical for terrain-illumination corrections and BRDF effect correction, which require sub-pixel locational accuracy (Li *et al.*, 2012, 2015).

The LSR25 product includes a Geometric Quality Assessment (GQA), which ranks pixels by an image matching algorithm. This information can be used to discard pixels with low spatial accuracy.

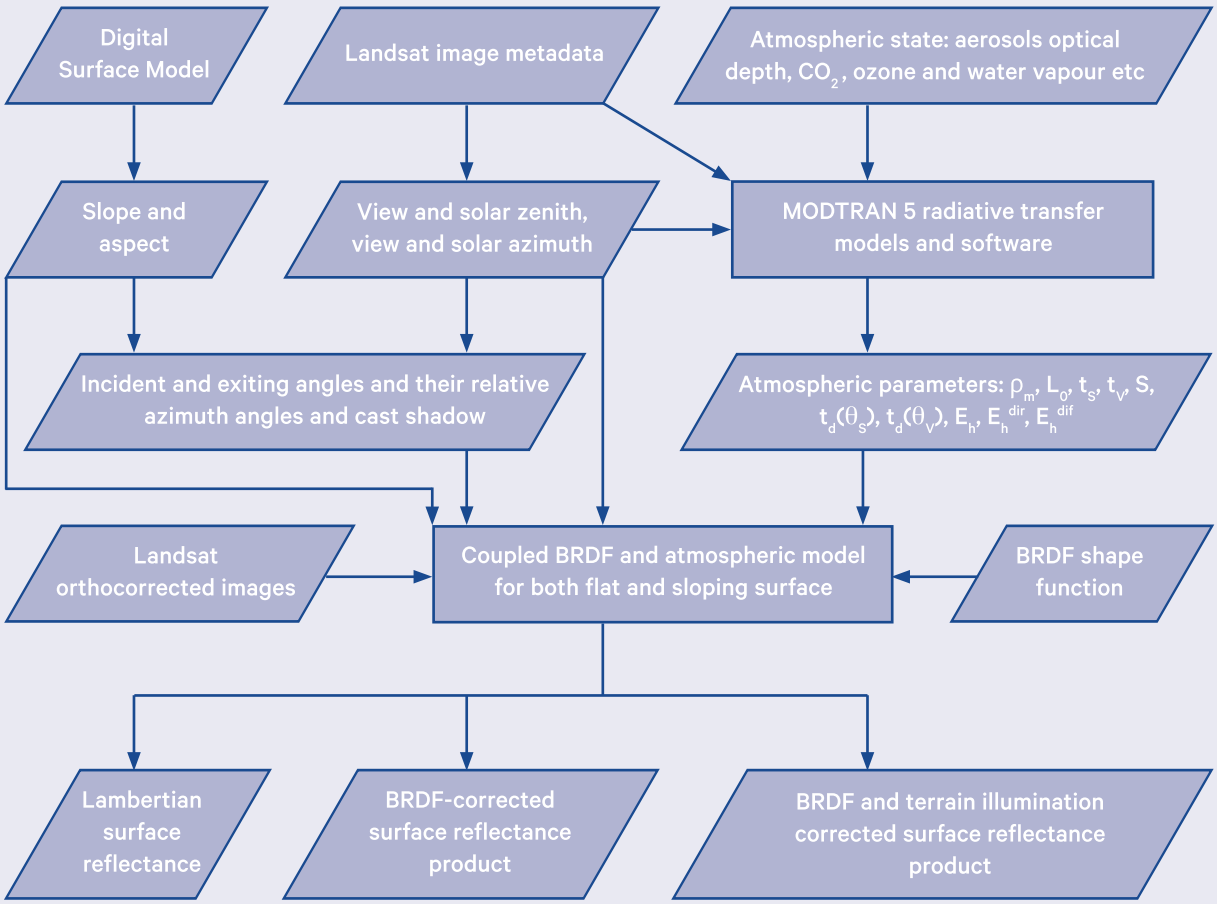
2. Radiometric correction

It has been traditionally assumed that Landsat imagery display negligible variation in Sun and sensor view angles, however these can vary significantly both within and between scenes, especially in different seasons and geographic regions (Li *et al.*, 2010, 2012). The LSR25 product delivers an estimate of surface reflectance from Landsat TM/ETM+/OLI/ data using physical rather than empirical models. This physical model couples BRDF and atmospheric correction models that can be applied to both flat and inclined surfaces (Li *et al.*, 2010, 2012; 2015), with the BRDF shape functions being derived from concurrent MODIS data (Lewis *et al.*, 2017). The resulting surface reflectance values are comparable both within individual images and between images acquired at different times and/or with different sensors (notwithstanding differences due to variations in spectral bandwidth between sensors). Accordingly, this product ensures that reflective value differences between imagery acquired at different times will be primarily due to on-ground changes in biophysical parameters rather than artefacts of the imaging environment (see Figure 11.5).

4 Formerly referenced as the Australian Reflectance Grid (ARG)

Figure 11.5 Radiometric correction of LSR25

Ancillary data inputs include digital surface model (1 arc second SRTM) to model elevation; concurrent MODIS imagery to derive BRDF shape function; aerosol data from Aeronet, AATSR, MISSR, MODIS and climatology sources; and field spectra collected at overpass times.



Source: Fuqin Li, Geoscience Australia

3. Pixel quality flags and metadata

To ensure that users understand the limitations of each pixel, quality flags are included in the LSR product. These are based on a series of tests applied during data preparation, including assessments of cloud, cloud shadow, topographic shadow, land/sea boundary, null values and sensor saturation (see Volume 2A—Section 1.5 for details). This approach allows all pixels in all available image scenes to be included in the product suite so that users can decide which pixels should be included in their analyses (Lewis *et al.*, 2017).

For complete transparency of EO products, the provenance of each pixel needs to be traceable back to its original observation values. The metadata included with the LSR products also assists dataset managers to determine when improvements in processing parameters, software or hardware warrants reprocessing of the archive or notifications to users. This includes maintaining hierarchical version numbers for software, ancillary data, and products, to indicate the significance of any changes.

11.3 Further Information

Digital Earth

International Society of Digital Earth (ISDE) (*www.digitalearth-isde.org*)

Global Spatial Data Infrastructure Association (*www.gsdi.org*),

International Council for Science and its Committee on Data for Science and Technology (*www.codata.org*),

Group on Earth Observations (GEO; and its geoportal, the GEO System of Systems (see Volume 1A—Section 1.5), *www.earthobservations.org*),

UN Committee of Experts on Global Geospatial Information Management (*ggim.un.org*),

Digital Earth Australia:

https://www.ga.gov.au/dea

11.4 References

Craglia, M., de Bie, K., Jackson, D., Pesaresi, M., Remetey-Fülöpp, G., Wang, C., Annoni, A., Bian, L., Campbel, F., Ehlers, M., van Genderen, J., Goodchild, M., Guo, H., Lewis, A., Simpson, R., Skidmore, A., and Woodgate, P. (2012). Digital Earth 2020: towards a vision for the next decade. *International Journal of Digital Earth*, 5, 4–21.

De Longueville, B., Annoni, A., Schade, S., Ostlander, N., and Whitmore, C. (2010). Digital Earth's nervous system for crisis events: real-time sensor Web enablement of volunteered geographic information. *International Journal of Digital Earth,* 3, 242–259.

Dhu, T., Dunn, B., Lewis, B., Lymburner, L., Mueller, N., Telfer, E., Lewis, A., McIntyre, A., Minchin, S., and Phillips, C. (2017) Digital earth Australia—unlocking new value from earth observation data. *Big Earth Data*, 1, (1–2), 64–74. *https://doi.org/10.1080/20964471.2017.1402490*

Evans, B., Wyborn, L., Pugh, T., Allen, C., Antony, J., Gohar, K., Porter, D., Smillie, J., Trenham, C., Wang, J., Ip, A., Bell, G. (2015). The NCI high performance computing and high performance data platform to support the analysis of petascale environ-mental data collections. In: *Environmental Software Systems. Infrastructures, Services and Applications* (Eds: Denzer, R., Argent, R.M., Schimak, G., Hrebicek, J.). IFIP Advances in Information and Communication Technology. Springer International Publishing, pp. 569–577.

Foresman, T.W. (2008). Evolution and implementation of the Digital Earth vision, technology and society, *International Journal of Digital Earth*, 1(1), 4–16, *https://doi.org/10.1080/17538940701782502*

GA (2015). *Australian Reflectance Grid (ARG25) Product Description V1.1*. Geoscience Australia, Canberra. *https://d28rz98at9flks.cloudfront.net/75062/Australian_Reflectance_Grid_ARG25_Product_Description_V1_1.PDF*

Goodchild, M.F., Guo, H., Annoni, A., Bian, L., de Bie, K., Campbell, F., Craglia, M., Ehlers, M., van Genderen, J., Jackson, D., Lewis, A.J., Pesaresi, M., Remetey-Fülöpp, G., Simpson, R., Skidmore, A., Wang, C., and Woodgate, P. (2012). Next-generation Digital Earth. *Proceedings of National Academy of Sciences of the USA*, 109 (28), 11088–11094.*https://doi.org/10.1073/pnas.1202383109*

Gore, A. (1998). *The Digital Earth: Understanding our planet in the 21st Century*. Speech given at the California Science Center, Los Angeles, on 31 January 1998. *http://www.crcsi.com.au/assets/Resources/c9fed807-0a8b-44ac-b167-690c32f6200a.pdf*

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V.C., Abrudan, I.V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., and Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010based on Landsat image composites. *Remote Sensing of Environment*, 151, 72–88. *http://dx. doi.org/10.1016/j.rse.2013.04.022*

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., and Hobart, G.W. (2015). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, 158, 220–234. *https://doi.org/10.1016/j.rse.2014.11.005*

Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendael, M., and Wagner, W. (2013). The ESA climate change initiative: satellite data records for essential climate variables. *Bulletin of the American Meteorological Society*, 94, 1541–1552. *http://dx.doi.org/10.1175/BAMS-D-11-00254.1*

Kennedy, R.E., Yang, Z., and Cohen, W.B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr temporal segmentation algorithms. *Remote Sensing of Environment*, 114, 2897–2910. *https://doi.org/10.1016/j.rse.2010.07.008*

Kim, Y., Kimball, J.S., Robinson, D.A., and Derksen, C. (2015). New satellite climate data records indicate strong coupling between recent frozen season changes and snow cover over high northern latitudes. *Environmental Research Letters*, 10, 84004.

Laney, D. (2001). 3D data management: controlling data volume, velocity and variety. *META Group Research Note. 6: p. 70. https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf*

Lewis, A., Lymburner, L., Purss, M.B.J., Brooke, B., Evans, B., Ip, A., Dekker, A.G., Irons, J.R., Minchin, S., Mueller, N., Oliver, S., Roberts, D., Ryan, B., Thankappan, M., Woodcock, R., and Wyborn, L. (2016). Rapid, high-resolution detection of environmental change over continental scales from satellite data—the Earth Observation Data Cube. *International Journal of Digital Earth*, 9(1), 106–111. *https://doi.org/10.1080/17538947.2015.1111952*

Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J., Dekker, A., Dhu, T., Hicks, A., Ip, A., Purss, M., Richards, C., Sagar, S., Trenham, C., Wang, P., and Wang, L.-W. (2017). The Australian geoscience data cube—Foundations and lessons learned. *Remote Sensing of Environment*, 202, 276–292. *https://doi.org/10.1016/j.rse.2017.03.015*

Li, F., D.L.B. Jupp, S. Reddy, L. Lymburner, N. Mueller, P. Tan, and A. Islam (2010). An Evaluation of the Use of Atmospheric and BRDF Correction to Standardize Landsat Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(3), 257–270. *https://doi.org/10.1109/JSTARS.2010.2042281*

Li, F., Jupp, D.L.B., Thankappan, M., Lymburner, L., Mueller, N., Lewis, A., and Held, A. (2012). A physics-based atmospheric and BRDF correction for Landsat data over mountainous terrain. *Remote Sensing of Environment*, 124, 756–770. *https://doi.org/10.1016/j.rse.2012.06.018*

Li, F., Jupp, D.L.B., and Thankappan, M. (2015). Issues in the application of Digital Surface Model data to correct the terrain illumination effects in Landsat images, *International Journal of Digital Earth*, 8(3), 235–257. *https://doi.org/10.1080/17538947.2013.866701*

Li, F., Jupp, D.L.B., Paget, M., Briggs, P.R., Thankappan, M., Lewis, A., and Held, A. (2017). Improving BRDF normalisation for Landsat data using statistical relationships between MODIS BRDF shape and vegetation structure in the Australian continent. *Remote Sensing of Environment*, 195, 275–296. *http://dx.doi.org/10.1016/j.rse.2017.03.032*

Lymburner, L., Bunting, P., Lucas, R., Scarth, P., Alam, I., Phillips, C., Ticehurst, C. and Held, A. (2019). Mapping the multi-decadal mangrove dynamics of the Australian coastline. *Remote Sensing of Environment,* 111185. *https://doi.org/10.1016/j.rse.2019.05.004*

Masek, J.G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., and Nelson, P. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, 112, 2914–2926. *http://dx.doi.org/10.1016/j.rse.2008.02.010*

Masek, J.G., Goward, S.N., Kennedy, R.E., Cohen, W.B., Moisen, G.G., Schleeweis, K., and Huang, C. (2013). United States Forest disturbance trends observed using Landsat time series. *Ecosystems*, 16, 1087–1104. *http://dx.doi.org/10.1007/s10021-013-9669-9*.

Mattmann, C.A. (2013). Computing: A vision for data science. *Nature*, 493:473–475. *http://dx.doi.org/10.1038/493473a*

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., and Ip, A. (2016). Water observations from space: mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341–352. *http://dx.doi.org/10.1016/j.rse.2015.11.003*

Sagar, S., Roberts, D., Bala, B., and Lymburner, L. (2017). Extracting the intertidal extent and topography of the Australian coastline from a 28 year time series of Landsat observations. *Remote Sensing of Environment,* 195, 153–169. *https://doi.org/10.1016/j.rse.2017.04.009*

Sagar, S., Phillips, C., Bala, B., Roberts, D., and Lymburner, L. (2018). Generating Continental Scale Pixel-Based Surface Reflectance Composites in Coastal Regions with the Use of a Multi-Resolution Tidal Model. *Remote Sensing*, 10, 480. *https://doi.org/10.3390/rs10030480*

Scarth, P., Roder, A., and Schmidt, M. (2010). Tracking grazing pressure and climate interaction—the role of Landsat fractional cover in time series analysis. *Proceedings of the 15th Australian Remote Sensing and Photogrammetry Conference*. 13-17 September, Alice Springs.

Sexton, J.O., Song, X.-P., Huang, C., Channan, S., Baker, M.E., and Townshend, J.R. (2013). Urban growth of the Washington, D.C.–Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover. *Remote Sensing of Environment*, 129, 42–53. *http://dx.doi.org/10.1016/j.rse.2012.10.025*

Stewart, J.B., Rickards, J.E., Randall, L.A., McPhee, R.K., and Paplinska, J.Z. (2014). *Ground cover monitoring for Australia: Final report July 2012 to June 2013*, ABARES Technical report 14.1, Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra, May 2014.

Thompson, S.D., Nelson, T.A., White, J.C., and Wulder, M.A. (2015). Mapping dominant tree species over large forested areas using Landsat best-available-pixel image composites. *Canadian Journal of Remote Sensing*, 41, 203–218. *http://dx.doi.org/10.1080/07038992.2015.1065708*

Zald, H.S.J., Wulder, M.A., White, J.C., Hilker, T., Hermosilla, T., Hobart, G.W., and Coops, N.C. (2016). Integrating Landsat pixel composites and change metrics with lidar plots to predictively map forest structure and aboveground biomass in Saskatchewan. *Canadian Journal of Remote Sensing*, 176, 188–201. *http://dx.doi.org/10.1016/j.rse.2016.01.015*

Zhu, Z., and Woodcock, C.E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152–171. *https://doi.org/10.1016/j.rse.2014.01.011*

Merging EO and Other Datasets

Ancillary datasets may be merged with EO imagery for a range of purposes and applications. Relevant attributes of ancillary data are introduced in Section 3.3 above. Assuming that any selected ancillary data sources are thematically appropriate to the EO dataset, the most critical aspects of scale for their integration are generally space and time.

Clearly, to compare two sets of data for a location they need to describe the exact same position on the ground. While the advent of portable GPS has simplified the process of accurately locating ground sites, locational accuracy alone does not ensure spatial compatibility between different sources of spatial data. If the characteristics of resolution, density and extent of the EO dataset cannot be related to those characteristics in the ancillary dataset then they will not integrate sensibly. For example, it is difficult to relate a pixel from an AVHRR image (1.1 km x 1.1 km) with point observations based on an area of 1 m x 1 m. The concept of a surface mosaic is introduced in Section 1.4 above, and the issue of scale is further discussed in Volume 2E.

In addition to spatial scale, the two datasets need to be coincident in time, at least in terms of relevant attributes of the surface features being identified. Timing of data acquisition can matter in terms of regular natural rhythms in target features, such as diurnal, seasonal and annual cycles. It is also relevant to variations in local conditions, such as changes in soil moisture, extent of cloud cover, and water surface conditions. The precise requirements for coincident timing will differ for different applications but, yet again, an understanding of the way the EO and ancillary datasets have been acquired is a prerequisite for credible results.

The following two sections present examples of ancillary datasets that may be integrated with EO data, including sample-based data, such as manual field sites, automatic sensors, spectral libraries and volunteered data (see Section 12), and gridded datasets, such as GIS, Digital Elevation Models (DEM) and meteorological records (see Section 13).

Contents

12 Sample-based Datasets 135

13 Gridded Datasets 149

Background image on previous page: Aerial imagery, acquired in March 2018 and centred on Collins Park, Forrest, ACT, overlaid with map information showing roads and cadastre in yellow. **Source:** Tony Sparks, Icon Water, based on ACT Government orthorectified imagery (MGA55 coordinates, GDA94 datum) acquired in 2018. The imagery was supplied by AAM (Ref. No.: 26121A) CC BY 4.0 and GIS data courtesy of ACT Office of the Surveyor-General and Land Information

12 Sample-based Datasets

The concept of scale is reviewed in Sections 1 and 2 above (see also Volume 1B—Section 1) and has particular significance when using ancillary datasets with EO imagery. In this context, scale is most relevant to the spatial and temporal dimensions. These spatial and temporal scales, however, may not be correlated, that is, the most appropriate scale for measuring a given attribute may not be the same as the best scale to either:

- estimate its values from population statistics (such as mean or variance) from a sample; or
- predict its values by computing the expected value at a particular location/time (Gallant *et al.*, 2008).

The following sub-sections will review different types of ancillary data that are available in Australia. Some of the caveats for integration with EO datasets are also considered:

- manually-collected field site data (see Section 12.1);
- automatic sensors (see Section 12.2);
- spectral libraries (see Section 12.3); and
- volunteered data (see Section 12.4).

12.1 Manually-Collected Field Site Data

Recommended procedures and protocols for the manual collection of soil and vegetation information at Australian field sites are detailed in Hnatiuk *et al.* (2009) and McKenzie *et al.* (2008). Hnatiuk *et al.* (2009) defines 'attributes needed to describe systematically the site and soil conditions related to landform, vegetation, land surface, soil profile and substrate materials' as well as relevant terminology.

TERN Australia (2018) presents detailed information for the calibration and validation (Cal/Val) of terrestrial environmental variables derived from EO datasets and models, including best practice guidelines for collection of field datasets in the Australian environment for this purpose. This reference recommends standards for:

- collection and management of field data (see Trevithick, 2018), including ground biomass (Schaefer, 2018b), vegetation spectroscopy (Suarez *et al.*, 2018; Chisholm and Hueni, 2018), ground cover (Stewart and Howorth, 2018), and terrestrial lidar (Calders *et al.*, 2018);
- calibration of EO sensors (see Malthus, 2018); and
- calibration and/or validation of EO datasets and derived biophysical products (Soto-Berelov *et al.*, 2018), including SAR (Mitchell and Thankappan, 2018), LAI (Schaefer *et al.*, 2018), ground cover monitoring (Stewart *et al.*, 2018), fractional cover (Scarth *et al.*, 2018), persistent green vegetation fraction (Gill *et al.*, 2018), phenology (Restrepo-Coupe *et al.*, 2018), foliar chemistry (Youngentob, 2018), tree crown delineation (Cabello-Leblic, 2018), hyperspectral data (Broomhall *et al.*, 2018), airborne lidar (Quadros and Keysers, 2018), and airborne imagery (Johansen *et al.*, 2018).

Excursus 12.1 introduces an Australian system, the National Vegetation Information System (NVIS), that has been adopted by state and federal government agencies for collecting and classifying information related to vegetation type. This system is further detailed in Volume 3A. An EO-based system for mapping vegetation type in NSW, which is based on the NVIS framework, is called the NSW Stage Vegetation Type Map (SVTM). This product is described in Volume 3A.

Background image: Fisheye photograph of *Eucalyptus* canopy at TERN Warra Tall Eucalypt Supersite in Tasmania[5]. **Source:** Tim Wardlaw, University of Tasmania

5 *https://supersites.tern.org.au/supersites/wsbe*

Excursus 12.1—National Vegetation Information System

Source: Richard Thackway, Australian National University
Further Information: NVIS Technical Working Group (2017); NVIS (2019)

NVIS (National Vegetation Information System) is now accepted as the national method for compiling and translating State and Territory mapped vegetation data on the extent and distribution of vegetation types in Australia (see Volume 3A for details). Much of the vegetation information available in the national NVIS database was collected and classified using the CSIRO guidelines prior to 2003 (such as Specht, 1970; Walker and Hopkins, 1990). In those states and territories that used different attribute and mapping frameworks, the NVIS framework was used to reclassify and remap those datasets. Today all states and territories generate and supply vegetation mapping and attribute data to the Australian Government's NVIS database that is consistent with the NVIS framework (NVIS, 2019).

As well as providing a framework for translating and compiling available vegetation mapping, NVIS is a comprehensive system that:

- specifies guidelines for standardising the collection, compilation and monitoring of Australia's vegetation;
- stores data on the type and extent of native vegetation with grid cell resolution of 100 m × 100 m;
- holds standardised geographic and attribute data for native vegetation, non-native vegetation and non-vegetated cover types across Australia that complements and facilitates analysis and reporting; and
- provides, monitors and maintains the technical infrastructure to support these activities (NVIS, 2019).

The approach adopted for describing and classifying vegetation derives from Hnatiuk *et al.* (2009). This hierarchical approach can be used to classify native, agricultural and other vegetation on the basis of its:

- structure—both vertical and horizontal—growth form, height, cover and strata; and
- floristics—the dominant genera or species in different strata, plus characteristic plant species (Hnatiuk *et al.*, 2009).

Basically, at each mapped polygon or grid cell (native vegetation, non-native vegetation and non-vegetated) is described systematically (see Table 12.1). Initially the dominant stratum and, where present, the mid-stratum and/or ground stratum, are distinguished for specific locations. Key attributes of relevant strata are then measured at those locations and the NVIS framework is used to classify them. For any given location, the dominant stratum is generally determined by analysing available EO imagery prior to assessing the other levels by field survey. Additional levels may be included to describe broad florisitic subformations, and other associations if required for particular types of vegetation, such as wetlands or rainforests. Nomenclature and methods for using the NVIS classification system are detailed in Thackway *et al.* (2008), Hnatiuk *et al.* (2009) and NVIS Technical Working Group (2017).

Table 12.1 Information Hierarchy

Level 5 (association) is recommended as the minimum level of input for input data to NVIS. Hnatiuk *et al.* (2009) refer to Level 1 as Formation.

Level	Category	Description
1	Class	Dominant growth form for the structurally dominant stratum
2	Structural formation	Dominant growth form, cover and height for the structurally dominant stratum
3	Broad floristic formation	Dominant genus (or genera) plus growth form, cover and height for the structurally dominant stratum
4	Sub-formation	Dominant genus (or genera) plus growth form, cover and height for each of the three main strata. (Upper, Mid and Ground)
5	Assocation	Dominant growth form, height, cover and species (to a maximum of 3 species) for each of the three main strata. (Upper, Mid and Ground)
6	Sub-association	Dominant growth form, height, cover and species (to a maximum of 5 species) for each of the substrata

Source: NVIS Technical Working Group (2017)

To populate the national NVIS database, detailed mapped polygon or grid cell (native vegetation, non-native vegetation and non-vegetated) datasets from each State and Territory have been translated and compiled to conform to the NVIS framework. The national database is updated as new data becomes available. NVIS Version 5.1 products include:

- detailed data from state and territory custodians, which is appropriate for use at a regional scale; and
- for national scale analyses, 33 Major Vegetation Groups (MVG) and 85 Major Vegetation Subgroups (MVS) defined by generalising the detailed data and adding non-NVIS gap-fill data (NVIS, 2019).

For the national-scale (Version 5.1) products, analysis tools are also provided, such as raster uncertainty layers to indicate data quality. MVG and MVS represent the dominant native vegetation type in each grid cell. These maps are produced for both current (extant) vegetation and estimated (pre-European) vegetation (see Volume 3A). For example, the NVIS methodology has also been adopted within the workflows used to create the NSW State Vegetation Type Map (SVTM; see Volume 3A).

12.2 Automatic Sensors

A variety of unmanned sensors may be used to record biophysical data at defined field locations on or near the ground, either on a temporary or permanent basis. Examples of such sensors include cameras, acoustic recorders, meteorological sensors, water sensors, and flux towers (see Volume 1A). While webcam photography has been used for monitoring environmental conditons for some time, webcam networks are being established for continual monitoring of ecosystems (Tanis *et al*., 2018).

Two of the key integrated observation networks that have been established in Australia are the:

- Terrestrial Ecosystem Research Network (TERN; see Excursus 12.2); and the
- Integrated Marine Observing System (IMOS; see Excursus 12.3).

As introduced below, these networks deliver a range of biophysical parameters that can be used to calibrate and validate EO datasets.

Excursus 12.2—Terrestrial Ecosystem Research Network (TERN)

Source: *http://www.tern.org.au*

The Terrestrial Ecosystem Research Network (TERN) observes, measures and records critical terrestrial ecosystem parameters and conditions for Australia over time, from continental scale to field sites at hundreds of representative locations. This information is standardised, integrated and transformed into model-ready data, enabling researchers to discern and interpret changes in land ecosystems.

The goal of TERN is to provide open access for researchers to Australia's land-based ecosystem monitoring infrastructure, data and research tools. This involves using nationally-consistent and standardised ecosystem measures over time to observe and measure ecosystem change across the continent in three measurement themes:

- biodiversity—monitoring patterns of change and their drivers at scales from genes to ecosystems;
- carbon and water—monitoring carbon stocks/flows and water resources; and
- land and terrain—tracking key processes and monitoring soils and vegetation.

As Australia's land ecosystem observatory, TERN is structured around three scales of observation:

- Landscape assessment monitoring—mostly through EO techniques based on satellite data, with the use of airborne data from autonomous vehicles (UAVs). Modelling and synthesis activities are also undertaken to extrapolate and interpolate from observational data to produce modelled data products.
- Ecosystem surveillance monitoring—which enables detection and monitoring of biodiversity change across a wider spatial extent of environments. TERN uses an extensive network of monitoring plots distributed along environmental gradients and in key biomes (see Figure 12.1).
- Targeted ecosystem process monitoring—which gives a high level of detail at a small number of sites. This is done through intensive field stations or 'SuperSites' which combine instrumented or sensor measurements.

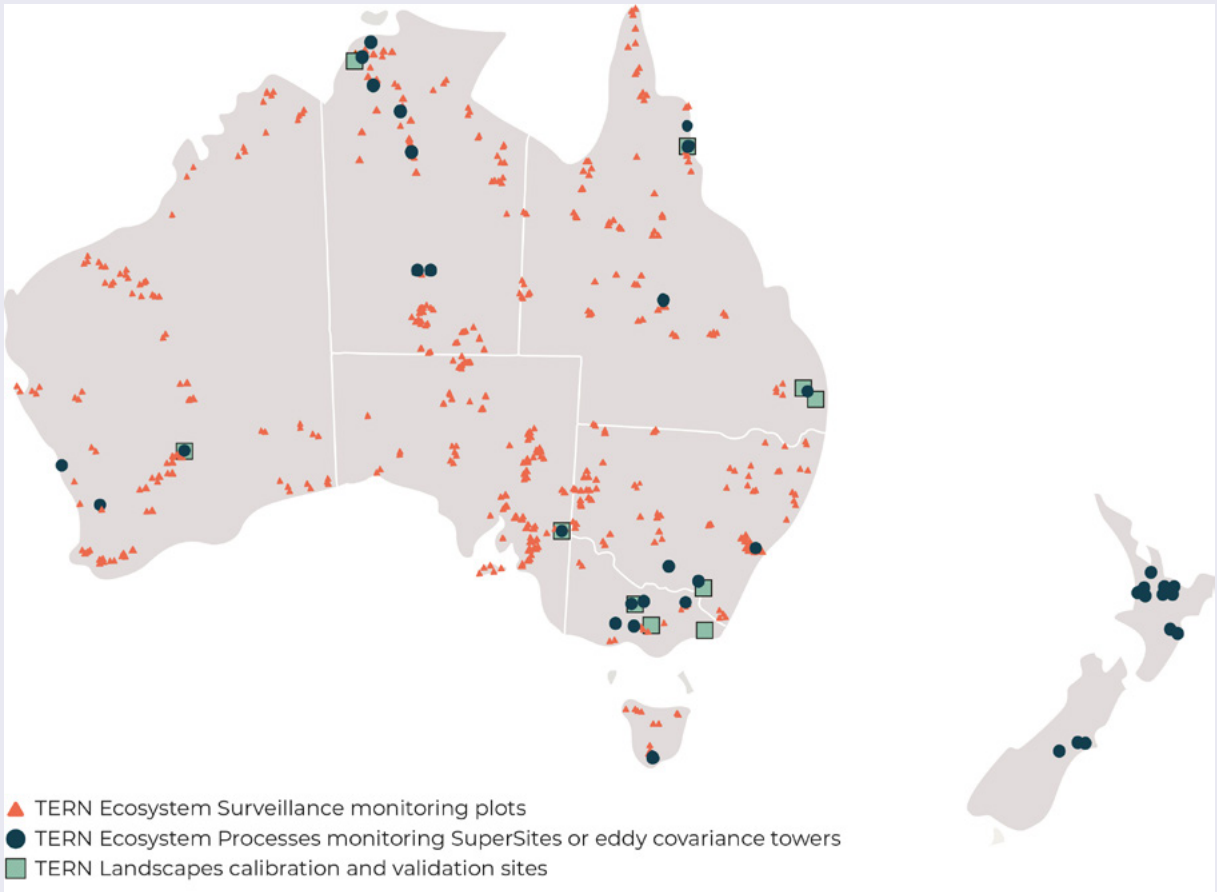
TERN delivers:

- Data—soil and vegetation measurements; gas, energy, and water exchange measurements; standardised EO datasets; and modelled data products about soil, climate, and landscape attributes;
- Tools—nationally consistent field methods for data collection and validation; vegetation and soil samples for physical analysis; data collection apps; and data publishing systems; and
- Infrastructure—supporting the discovery, access, and re-use of TERN data and other ecosystem data housed in our systems.

In particular, the EO products delivered via TERN include land cover dynamics and phenology, vegetation composition and diversity, fire dynamics and impacts, vegetation structural properties and biomass, field survey datasets, airborne datasets, corrected surface reflectance products and other environmental data such as solar radiation, rainfall, and water vapour pressure (see Excursus 14.1).

Figure 12.1 TERN sites

Field sites around Australia being monitored by TERN (current in September 2019).



A site is a small area of land considered representative of the landform, vegetation, land surface and other land features associated with the soil observation. (Speight and McDonald, 2009)

Excursus 12.3—Integrated Marine Observing System (IMOS)

Source: *http://imos.org.au*

Since 2006, the Integrated Marine Observing System (IMOS) has been routinely operating a wide range of observing equipment throughout Australia's coastal and open oceans, making all of its data accessible to the marine and climate science community, other stakeholders and users, and international collaborators.

IMOS is a national collaborative research infrastructure, supported by Australian Government. It is operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania as Lead Agent.

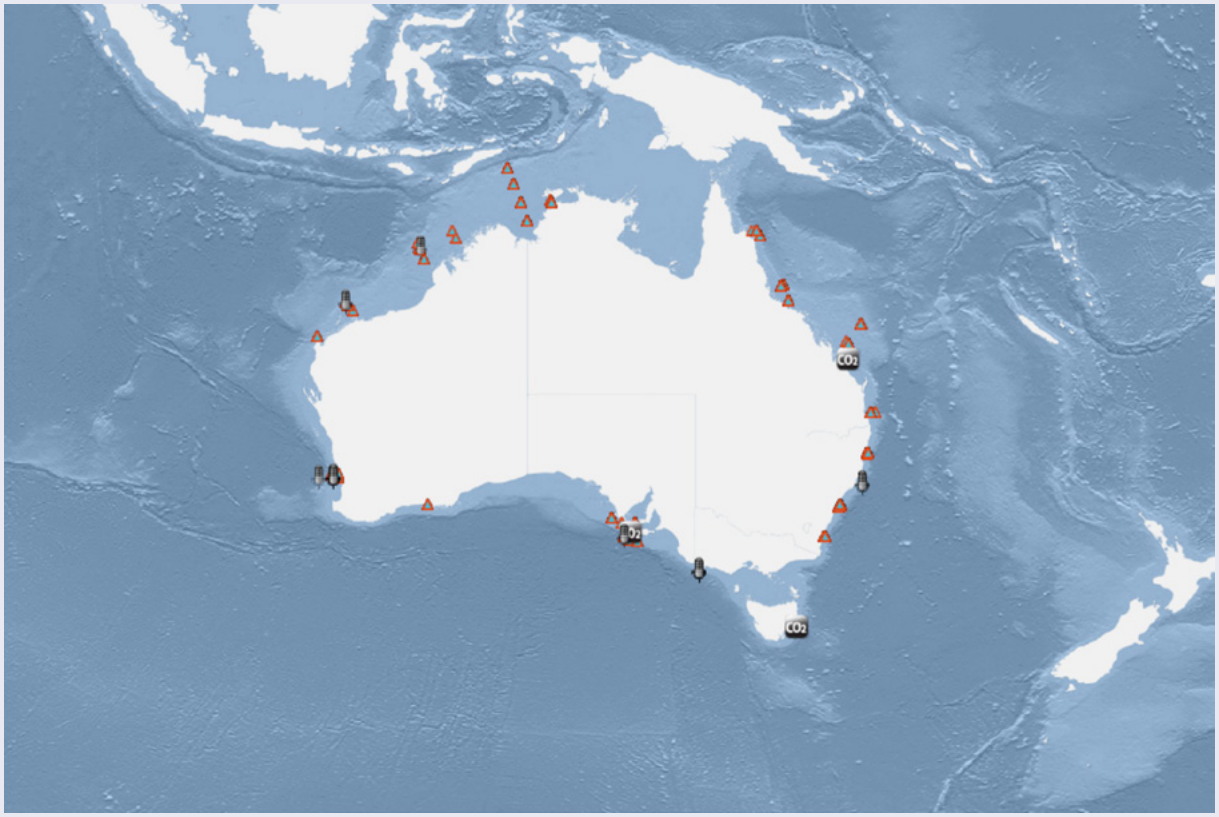
IMOS currently has a portfolio of Facilities that undertake systematic and sustained observing of Australia's marine environment, across scales (from open ocean, onto the continental shelf, and into the coast), and across disciplines (physics, biogeochemistry, and biology and ecosystems), including:

- Argo Floats—autonomous profiling floats that measure both temperature and salinity in the upper 2000 m of the ocean every 10 days;
- Ships of Opportunity—volunteer commercial and research vessels to collect data relating to physical, chemical and biological oceanography and ecology;
- Deep Water Moorings—located in Antarctica, sub-Antarctic, sub-tropical and tropical open waters around Australia, these moorings provide long-term information about temporal patterns in oceanographic and biological conditions;
- Ocean Gliders—autonomous underwater gliders that measure temperature, salinity, dissolved oxygen, chlorophyll and turbidity in sub-surface waters;
- Autonomous Underwater Vehicles—provide rapid and cost-effective high resolution, accurately georeferenced and targeted acoustic imagery of the seafloor;
- National Mooring Network—located around the Australian coastline (see Figure 12.2), this network delivers long term, time series observations from a range of instruments including Acoustic Doppler Current Profilers (ADCP), WetLabs Water Quality Meters (WQM), fluorometers, Conductivity-Temperature-Depth (CTD) with turbidity and Dissolved Oxygen (DO) sensors;
- Ocean Radar—map ocean surface currents over meso-scale areas with high spatial and temporal resolution;
- Animal Tracking—uses acoustic technology, CTD satellite trackers and bio-loggers to monitor coastal and oceanic movements of marine animals around Australia and Antarctica;
- Wireless Sensor Networks—provide real time, spatially dense measurements of biophysical environmental variables on the Great Barrier Reef;
- Satellite Remote Sensing—acquire and process EO imagery for sea surface temperature (SST), ocean colour, oceanic surface wave and sea surface altimetry applications;
- Marine Microbiome Initiative—monthly water samples collected around Australia will build a time series of microbial data; and
- New Technology Proving—capability allowing for the development and establishment of new technologies, methods, and approaches in collecting sustained ocean observations.

IMOS observations are turned into data that can be discovered, accessed, downloaded, used and reused in perpetuity by our data Facility, the Australian Ocean Data Network (AODN). All of the IMOS data is freely accessible via the AODN Portal: *http://portal.aodn.org.au*

Figure 12.2 IMOS National mooring network

Location of the national reference station and shelf moorings (red triangles), acidification moorings (CO_2 icon) and passive acoustic moorings (hydrophone icon) around the Australian continent.



Know thy site.
(Ray Leuning)

12.3 Spectroscopy datasets

Spectroscopy datasets are commonly acquired in the field and/or laboratory to augment EO analyses. Some of the sensors that can be used to acquire spectroscopy data included spectrometers/ spectroradiometers, spectrophotometers, spectrographs or spectral analyzers (Chisholm and Hueni, 2018; see Volume 1A—Section 13). Recommended procedures for collecting spectral datasets for different applications in an Australian context are detailed in TERN Australia (2018).

Spectroscopy datasets have been used to calibrate data from EO sensors, validate EO analyses and/or investigate spectral reflectance characteristics of specific ground materials, such as particular plants or minerals (Chisholm and Hueni, 2018). A range of systems has been proposed to organise spectroscopy datasets, including:

- spectral libraries—static collections of reference spectra, such as the USGS spectral library or the ASTER spectral library;
- spectral databases—store spectral data in an organised manner, with appropriate metadata (Hueni *et al*., 2011); and
- spectral information systems (SIS)—provide tools to access and process spectral databases (Chisholm and Hueni, 2018).

Chisholm and Hueni (2018) describe the lifecycle of spectroscopy datasets in terms of six generic stages:

- planning of sampling experiments, including the definition of sampling protocols adhering to a metadata standard;
- actual data acquisition, where data are acquired according to predefined sampling protocols;
- ingestion of the acquired spectral data into the SIS;
- augmentation of the automatically generated metadata by manually or semi-automatically adding further metadata parameters to the spectral data collection;
- building further information by applying algorithms to spectral data and metadata; and
- retrieval of information for a particular purpose.

Unlike simple spectral libraries, SIS allow multiple reflectance measurements to be dynamically stored for a given target, with subsequent additional measurements being added over time and space. As with EO datasets, processing levels can be defined for spectroscopic data and its metadata (see Table 12.2), which enable the provenance of specific data to be traced and/or the most appropriate data for a particular task to be selected. SIS analysis tools are designed to use the full potential of combined metadata spaces (Wason and Wiley, 2000) and spectral spaces (Hueni *et al* 2012).

For example, SPECCHIO (Bojinski *et al*., 2003; Hueni *et al*., 2006, 2009) is an open source project that enables storage and analysis of spectral databases, including generating higher-level products and correcting spectral data for sampling equipment or sensor artefacts (Hueni *et al*., 2012). This Swiss product is freely available for ingesting, managing, and processing of spectral datasets (Hueni *et al*., 2009; SPECCHIO, 2019). The Australian Spectral Database, AUS-SPECCHIO (Chisholm and Hueni, 2018) will be hosted by Geoscience Australia in the near future (see Volume 3A).

Table 12.2 Spectral database processing levels

Level	Description
Raw	Raw, sensor generated files, stored as binary objects on a file system or in the database system. This forms the first tier of the DIKW hierarchy and allows regeneration of data/information at the following tiers (see Volume 1A—Section 1).
Level 0	Spectral measurements as digital number (DN), described by auto-generated metadata augmented by user-defined metadata parameters.
Level 1	Spectral measurements as radiances traceable to an international standard. Metadata as in level 0 but includes information related to the data calibration process.
Level 2	Spectral measurements as factors (reflectance factors, transmittance, absorbance), corrected for reference panel deficiencies where needed (non-ideal reflective and Lambertian properties). Metadata as in level 1 but includes information related to the data calibration process.
Higher level products	Products derived from the lower levels, similar to products generated in imaging spectrometer processing systems, such as estimated bio-geophysical properties.

Source: Chisholm and Hueni (2018) Table 14.2

12.4 Volunteered data

Data that has been volunteered by community members is becoming a valuable source of timely, local observations for a range of EO-based applications, including natural disaster management, faunal and floral surveys, compliance monitoring and assessment of water conditions. Volunteered data can be broadly divided into two categories:

- citizen science—collection of pre-defined measurements, photos and observations at identified locations, as part of a coordinated, scientific study; and
- crowdsourcing—*ad hoc* collection of potentially relevant observations.

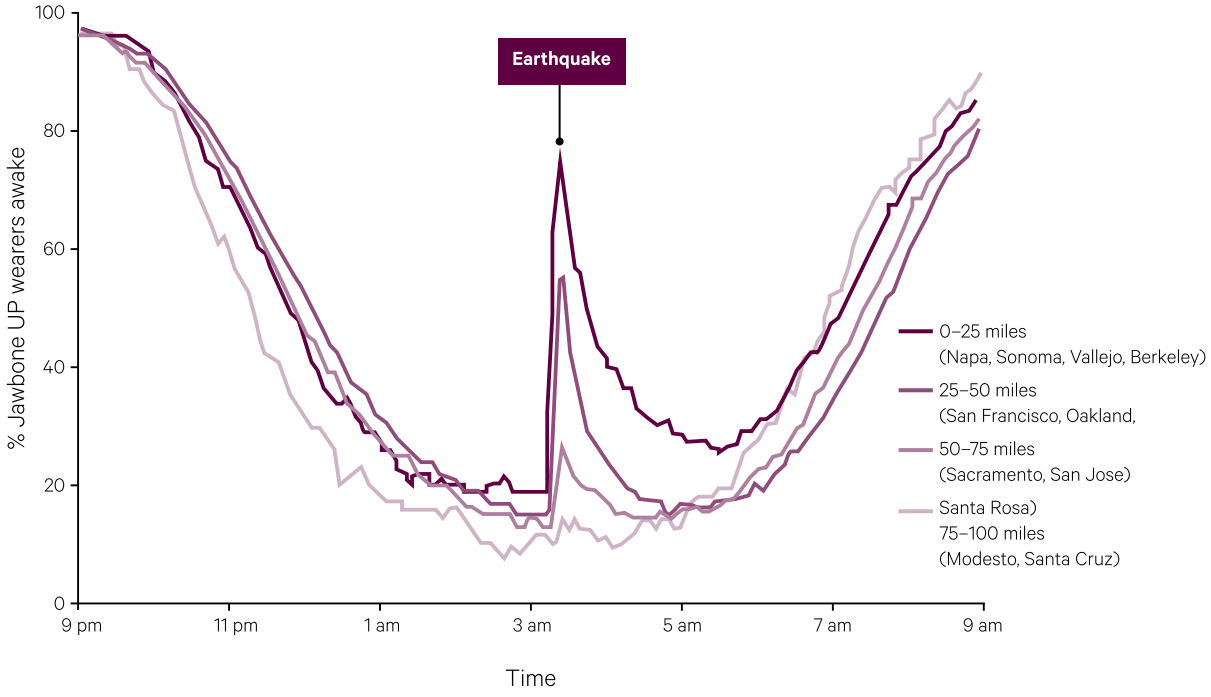
The locational accuracy of volunteered information has improved with easy access to GPS technology in smart phones. Similarly, phone-based cameras allow volunteers to acquire and transmit local perspectives of various environmental events and conditions. The value of such information is further improved by independent time and date stamping.

Crowdsourcing can involve unplanned data collection by informal sensors. For example, as described in Volume 1B—Section 10, in California on 24 August 2014, fitness monitors tracked the time their wearers were woken by an earthquake. The resulting compilation of data by the manufacturer provided a map of the epicentre, based on the relationship between each wearer's location and the time they were woken from sleep (see Figure 12.3).

The key to acquiring consistently valuable data from volunteers is the same as for most other ancillary data sources: relevant timing, compatible spatial scaling and training of volunteers. One successful example of citizen scientists assisting EO studies is the global monitoring program 'EyeonWater'. The Australian implementation of this initiative is described in Excursus 12.4 (see Volume 3B—Section 11 for details).

Figure 12.3 Inadvertent map of distance to earthquake epicentre

Wearable devices assist seismic observations after a 6.0 earthquake occurred in Souther Napa Valley, California at 3:20 am on August 2014. As distance from the epicentre increased, fewer wearers of the Jawbone Up fitness monitor were woken from sleep.



Adapted from Mandel (2014)

Excursus 12.4—EyeonWater Australia

Source: Janet Anstee, CSIRO
Further Information: *https://www.eyeonwater.org*

Another approach to EO of surface waters involves directly photographing the water surface and inferring water quality from image colour. A smartphone application, Eye on Water, which is used worldwide by different citizen science projects (Wernand *et al*. 2013; Novoa *et al*. 2015; Busch *et al*. 2016), has been adapted by CSIRO for Australian inland and coastal waters. The EyeonWater project was spawned from the European Union funded Citizens' Observatory for Coast and Ocean Optical Monitoring project (Citclops). As part of that project a smart phone app was developed to enable citizen scientists to easily add their own observations of water colour to the global database.

The EyeonWater Australia application was developed to assist validation of water colour from satellite imagery and to improve knowledge of aquatic systems, particularly in remote inland and coastal regions of Australia. Once installed on a smartphone, the app quantifies water colour. This involves guiding the citizen scientist through:

- reliably photographing the water surface; then
- comparing the water colour to the Forel-Ule colour scale (see Figure 12.4).

Figure 12.4 Using the EyeonWater Australia app

Part of the EyeonWater Australia project involves training students from the Bush Ranger program at the Broome Senior High School as citizen scientists to contribute water colour observations.

Source: a. Janet Anstee, CSIRO b. Broome Senior High School

Citizen scientists are needed for the project because water quality information is lacking in many places. As well as being part of a research project that is improving satellite-derived water quality products., community members gain a better understanding of water quality and its impact on our ecosystems. Scientists benefit from an increased number of observations, particularly in remote regions, as well as knowledge transfer from people who are well-connected to their localities.

The database currently stores all the water quality information that has been uploaded and this will help scientists determine seasonal influences on water quality and potentially determine the environmental drivers for detected changes (such as after a cyclone). Observations collated during 2017 and 2018 are shown in Figure 12.5. Over time, these observations will enable long-term monitoring of water colour, which otherwise would not be possible without expensive scientific surveys and water quality analyses.

Figure 12.5 EyeonWater Australian observations 2017–2018

Colours at sites relate to water colour on the Forel-Ule Scale (see Figure 12.4a and Volume 3B).

12.5 Further Information

Citizen Science:

EyeOnWater: *https://www.eyeonwater.org*

EyeOnWater Australia: *https://www.eyeonwater.org/apps/eyeonwater-australia*

Atlas of Living Australia: *https://www.ala.org.au*

Automatic Sensor Networks:

TERN: *http://www.tern.org.au*

TERN Ozflux: *http://www.ozflux.org.au*

IMOS: *http://imos.org.au*

Australian Phenocam Network: *https://phenocam.org.au/*

National Vegetation Information System (NVIS):

NVIS: *http://www.environment.gov.au/land/native-vegetation/national-vegetation-information-system*

Spectral Libraries and Databases:

National Spectral Database: planned for launch in 2020 by Geoscience Australia

Specchio: *https://specchio.ch*

USGS Spectroscopy Lab: *https://www.usgs.gov/labs/spec-lab*

USGS High Resolution Spectral Library: *https://www.usgs.gov/energy-and-minerals/mineral-resources-program/science/usgs-high-resolution-spectral-library?qt-science_center_objects=0#qt-science_center_objects*

12.6 References

Bojinski, S., Schaepman, M., Schlaepfer, D. and Itten, K. (2003). SPECCHIO: a spectrum database for remote sensing applications. *Computers and Geosciences*, 29, 27–38.

Broomhall, M., Johansen, K., and Wu, D. (2018). Quality assurance steps for TERN landscape assessment hyperspectral data. Ch 15 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Busch, J.A., Price, I., Jeansou, E., Zielinski, O., and van der Woerd, H.J. (2016). Citizens and satellites: Assessment of phytoplankton dynamics in a NW Mediterranean aquaculture zone. *International Journal of Applied Earth Observation Geoinformation,* 47, 40–49.

Cabello-Leblic, A. (2018). Tree crown delineation. Ch 11 in in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Calders, K., Wilkes, P., Disney, M.I., Armston, J., Schaefer, M., and Woodgate, W. (2018). Terrestrial lidar for measuring above-ground biomass and forest structure. Ch 19 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Chisholm, L.A., and Hueni, A. (2018). The Spectroscopy Dataset Lifecycle: Best Practice for Exchange and Dissemination. Ch 14 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Gallant, J.C., McKenzie, N.J., and McBratney, A.B. (2008). Scale. Ch 3 in *Guildelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

Gill, T., Johansen, K., Scarth, P., Armston, J., Trevithick, R., and Flood, N. (2018). Persistent green vegetation fraction. Ch 8 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Hnatiuk R.J., Thackway R. and Walker J. (2009). *Vegetation*. In: *Australian Soil and Land Survey: Field Handbook* (3rd Edition). (Eds National Committee on Soil and Terrain) pp. 73–125. CSIRO Publishing, Melbourne.

Hueni, A., and Tuohy, M. (2006). Spectroradiometer Data Structuring, Pre-Processing and Analysis—An IT Based Approach. *Journal of Spatial Science*, 51(2), 93–102.

Hueni, A., Nieke, J., Schopfer, J., Kneubühler, M., Itten, K. (2009). The spectral database SPECCHIO for improved long term usability and data sharing. *Computers and Geosciences*. *https://doi.org/10.1016/j.cageo.2008.03.015*

Hueni, A., Malthus, T., Kneubuehler, M. and Schaepman, M. (2011). Data Exchange between distributed Spectral Databases. *Computers and Geosciences,* 37: 861–873. *https://doi.org/10.1016/j.cageo.2010.12.009*.

Hueni, A., Chisholm, L., Suarez, L., Ong, C., Wyatt, M. (2012). Spectral information system development for Australia, in *Proceedings of the Geospatial Science Research Symposium 2*, RMIT, Melbourne, December 2012.

Johansen, K., Trevithick, R., Bradford, M., Hacker, J., McGrath, A., and Lieff, W. (2018). Australian examples of field and airborne TERN Landscape Assessment campaigns. Ch 17 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Malthus, T. (2018). Calibration of optical satellite and airborne sensors. Ch 4 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Mandel, E. (2014). *How the Napa earthquake affected bay area sleepers*. See *https://www.forbes.com/sites/dandiamond/2014/08/25/how-an-earthquake-will-mess-up-your-sleep-a-jawbone-chart/#7344fc2e3b3a*

McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J. (Eds) (2008). *Guildelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. CSIRO Publishing, Melbourne.

Mitchell, A., and Thankappan, M. (2018). Effective guidelines for calibration and validation of SAR data and derived biophysical products. Ch 5 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Novoa, S., Wernand, M.R., and van der Woerd, H.J. (2015). WACODI: A generic algorithm to derive the intrinsic color of natural waters from digital images. *Limnological and Oceanographic Methods*, 13, 697–711.

NVIS Technical Working Group (2017). *Australian Vegetation Attribute Manual: National Vegetation Information System, Version 7.0*. (Eds: Bolton, M.P., deLacey, C. and Bossard, K.B.). Department of the Environment and Energy, Canberra.

NVIS (2019). *NVIS data products* webpage: *http://www.environment.gov.au/land/native-vegetation/national-vegetation-information-system/data-products*

Quadros, N., and Keysers, J. (2018). Airborne lidar acquisition and validation. Ch 16 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Restrepo-Coupe, N., Huete, A., and Davies, K. (2018). Satellite phenology validation. Ch 9 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Scarth, P., Guerschman, J.P., Clarke, K., and Phinn, S. (2018). Validation of Australian Fractional Cover Products form MODIS and Landsat Data. Ch 7 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Schaefer, M.T., Farmer, E., Soto-Berelov, M., Woodgate, W., and Jones, S. (2018). Ch 6 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Schaefer, M.T. (2018). Measurement of above ground biomass. Ch 12 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Soto-Berelov, M., Jones, S., Farmer, E., and Woodgate, W. (2018). Review of validation standards of biophysical Earth Observation products. Ch 2 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

SPECCHIO (2019). *SPECCHIO* website: *https://specchio.ch*

Specht, R.L. (1970). Vegetation. In *The Australian Environment* (4th Edn). (Ed: Leeper, G.W.). CSIRO, Melbourne.

Speight, J.G., and McDonald, R.C. (2009). *The Site Concept*. In: *Australian Soil and Land Survey: Field Handbook* (3rd Edn). (Eds National Committee on Soil and Terrain) pp. 73–125. CSIRO Publishing, Melbourne.

Stewart, J.B., and Howorth, J.E. (2018). A calibration and validation framework to support ground cover monitoring for Australia. Ch 18 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Suarez, L., Restrepo-Coupe, N., Hueni, A., and Chisholm, L.A. (2018). Vegetation spectroscopy. Ch 13 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Tanis, C.M., Peltoniemi, M., Linkosalmi, M., Aurela, M., Böttcher, K., Manninen, T., and Arslan, A.N. (2018). A System for Acquisition, Processing and Visualization of Image Time Series from Multiple Camera Networks. *Data*, 3, 23. *https://doi.org/10.3390/data3030023*

TERN Australia (2018). *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Thackway, R., Neldner, J., and Bolton, M. (2008). Vegetation. In: *Australian Soil and Land Survey Handbook Guidelines for Conducting Surveys* (2nd Edition). (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and. Ringrose-Voase, A.J.). CSIRO Publishing, Melbourne.

Trevithick, R. (2018). Field data collection and management for Earth Observation image validation. Ch 3 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

Walker, J., and Hopkins, M.S. (1990). Vegetation. In *Australian Soil and Land Survey Handbook: Guidelines for Conducting Surveys*. (Eds: Gunn, R.H., Beattie, J.A., Reid, R.E., and van der Graaff, R.H.M.). Inkata Press, Melbourne. pp 58–86.

Wason, T.D., and Wiley, D. (2000). Structured Metadata Spaces. *Journal of Internet Cataloging*, 3(2/3), 263–277.

Wernand, M.R., van der Woerd, and H.J., Gieskes, W.W.C., (2013). Trends in ocean colour and chlorophyll concentration from 1889 to 2000, worldwide. *PLoS One*, 8, e63766.

Youngentob, K.N. (2018). Estimating foliar chemistry of individual tree crowns with imaging spectroscopy. Ch 10 in *Effective Field Calibration and Validation Practices: A practical handbook for calibration and validation satellie and model-derived terrestrial environmental variables for research and management*. A TERN Landscape Assessment Initiative, NCRIS. ISBN: ISBN 978-0-646-94137-0. *https://www.tern.org.au/NEW-CalVal-handbook-for-remote-sensing-bgp4370.html*

13 Gridded Datasets

This section considers three categories of gridded, spatial data that can be integrated with EO datasets:

- GIS attributes (see Section 13.1);
- digital elevation models (see Section 13.2); and
- meteorological records (see Section 13.3).

13.1 Geographic Information Systems

Geographic Information Systems (GIS) compile, access and interrogate spatial data in digital format. The unique characteristic of spatial data is the component that indicates location, either explicitly or implicitly. As introduced in Section 3.3.2, most digital geographic data is represented in either raster or vector format, but it is more appropriate to code some types of data (such as terrain heights, see Section 13.2) in a format known as Triangular Irregular Network (TIN).

The strength of GIS is their ability to store and analyse large volumes of spatial data for environmental, commercial and legislative purposes. Analyses typically include some of the raster operations described in Volumes 2A, 2B and 2C, such as combining attributes, neighbourhood operations, and segmentation/classification, plus specific vector-based analyses, such as proximity functions, network modelling, spatial prediction and terrain analysis (Bolstad, 2008). Spatial and spatio-temporal modelling tools are increasingly available in GIS.

A wide range of environmental and socio-economic attributes are now available as spatial coverages in Geographic Information Systems (GIS). For example, the online mapping tool, NationalMap, enables access to a growing range of geospatial datasets for Australia, many of which have been derived using EO data (see Excursus 13.1). As with other forms of ancillary data, to be sensibly integrated with EO datasets, GIS attributes need to be spatially accurate, geometrically coincident and temporally relevant (see Section 3.3 above and Volume 2B).

A GIS is a computer-based system to aid in the collection, maintenance, storage, analysis, output and distribution of spatial data and information.
(Bolstad, 2008)

Background image: Digital Elevation Model over alpine region in Victoria, displayed with black for lowest altitudes to white for highest altitudes. **Source:** Fuqin Li, Geoscience Australia

Excursus 13.1—NationalMap

Source: *https://nationalmap.gov.au/about.html*

NationalMap is an online map-based tool to allow easy access to geospatial data from Australian federal, state and territory government agencies. In this open architecture implementation, data is generally accessed directly from the custodians. It is intended for information purposes only. In March 2019, base maps for Australia included topography (Geoscience Australia), Bing maps (Microsoft), Natural Earth II (public domain), Black Marble (NASA), and Positron (Light) and Dark Matter (Carto). These underlays can be integrated with terrain information from Cesium Word Terrain (Cesium Ion) and geocoded information from G-NAF (PSMA geocoded national address file), Bing Maps (Microsoft) and the Gazetter of Australia.

The NationalMap tool enables multiple geospatial datasets to be integrated on a selected base map for a given location. Categories of national datasets, for example, include communications, elevation, environment, framework, groundwater, habitation, health, infrastructure, land, national boundaries, satellite imagery, social and economic, statistical boundaries, surface water and marine, terrain, transport, utility and vegetation. Resulting compilations can be downloaded and shared, or integrated with personal datasets, and data values for displayed locations can be queried.

National spatial websites that can be accessed via NationalMap include:

- AREMI (Australian Renewable Energy Mapping Infrastructure;
- NEII viewer (National Environmental Information Infrastructure);
- AURIN (Australian Urban Research Infrastructure Network); and
- SOE map (State of the Environment).

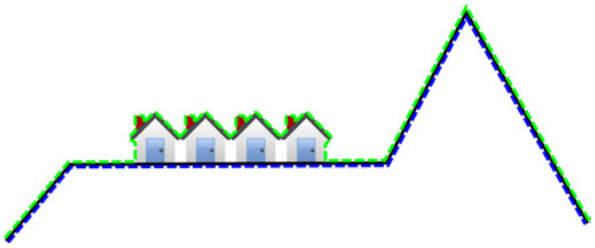
13.2 Digital Elevation Models

A digital elevation model (DEM; or digital terrain model, DTM) represents a terrain surface as a three-dimensional, digital model. A digital surface model (DSM), represents the height of the surface, including the heights of any objects that are above the terrain (see Figure 13.1). Traditionally, elevation has been recorded as topographic contours, that is, lines of uniform elevation that are orthogonal to the local slope (Bolstad, 2008), but other common formats for elevation data include a raster format of grid cells and a triangulated irregular network (TIN; see Section 3.3.2 and Figure 13.2). Most DEM used for resource assessment are now based on a raster format (Gallant and Hutchinson, 2008).

DEM have also been derived from 236 lidar surveys over Australia that were acquired between 2001 and 2015 (GA, 2019). Covering over 245,000 sq. km, these datasets have been merged into 5 m and 25 m grids, which focus on populated regions, coastal zones, and floodplains within the Murray-Darling Basin.

Figure 13.1 DTM versus DSM

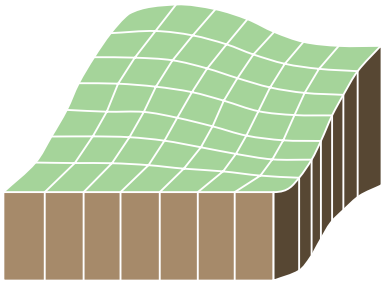
A digital surface model (DSM) represents the heights of all objects on the terrain (shown as green) whereas a digital terrain model only represents the elevation of the terrain surface (shown as blue).



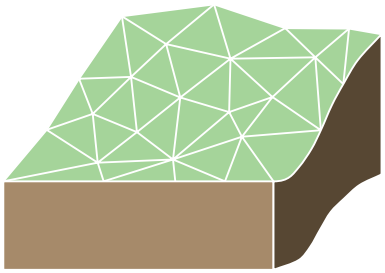
Adapted from: Wikimedia Commons (Arbeck). Retrieved from *https://commons.wikimedia.org/wiki/File:The_difference_between_Digital_Surface_Model_(DSM)_and_Digital_Terrain_Models_(DTM)_when_talking_about_Digital_Elevation_models_(DEM).svg*

Figure 13.2 Elevation data formats

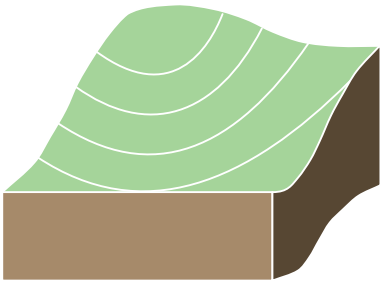
a. Raster



b. Vector



c. Triangulated Irregular Network (TIN)



Since terrain height directly determines surface water distribution and biogeography (including human settlement), accurate mapping of elevation is essential for many environmental and demographic analyses. In addition to elevation, several terrain variables have been identified, which quantify information about landform (see Table 13.1). By summarising the relationship between landform and other landscape attributes, such as soil type and vegetation (Gallant and Hutchinson, 2008), these terrain variables have become important for a range of activities related to landscape planning and resource management including stratifying sampling sites (Hewitt *et al.*, 2008), providing explanatory variables in soil-landscape models (McKenzie *et al.*, 2000), and visualising landscapes (Gallant and Hutchinson, 2008).

Table 13.1 Terrain variables

Group	Variable	Description	Importance
Exposure	Height	Elevation above base	Temperature, vegetation, visibility
	Slope	Change in elevation relative to change in horizontal distance	Water flow, flooding, erosion, transport cost, construction suitability, geology, insolation, soil depth
	Aspect	Downhill direction of steepest slope	Insolation, temperature, vegetation, soil characteristics and moisture, hazards
Curvature	Profile curvature	Curvature parallel to slope direction	Erosion, water flow acceleration
	Plan curvature	Curvature perpendicular to slope direction	Water flow convergence, soil water, erosion
Hydrologic functions	Flow direction	Direction of flow of excess surface water	Water runoff, erosion
	Flow length	Longest upstream flow path to a point	Sediment and erosion rates
	Upslope area	Watershed area above a point	Soil moisture, volume and timing of water runoff, pollution or erosion hazards
	Upslope length	Mean upstream flow path length to a point	Sediment and erosion rates
Landscape aesthetics	Landscape position	Location of site within context of surrounding landscape	Spatial prediction of landscape patterns
	Visiblilty	Site obstruction from given viewpoints	Utility location, viewshed preservation
	Viewshed	Collection of areas visible from a point	Utility location,
Non-terrain	Topographic wetness index	Ratio of specific catchment area to slope	Measure of relative soil wetness
	Stream power index	Product of specific catchment area to slope	Measure of ability of surface flow to transport sediment
	Shortwave radiation	Solar radiation reaching land surface	Insolation, Sun shading

Source: Bolstad (2008) Table 11-1 and Gallant and Hutchinson (2008)

As the word 'model' implies, a DEM only offers an approximate 'representation' of the actual surface, and does not attempt to mimic all local variations in terrain height or necessarily represent the actual elevation at specific locations (Gallant and Hutchinson, 2008). Construction of a reliable DEM requires consideration of numerous constraints relating to drainage, scaling, and interpolation, as well as addressing any underlying data errors. The ANUDEM software suite, for example, computes DEM for a range of scales with logical shape and drainage structure. This software has been used to develop the national DEM-9S and improve DEM products derived from the Shuttle Radar Topography Mission (SRTM; ANU/GA, 2008).

The national DEM available for Australia are listed in Table 13.2. Versions of the SRTM DEM include:

- DSM (3-second DEM only): ground surface topography plus features above the ground, such as vegetation and man-made structures;
- DEM: ground surface topography with vegetation features removed;
- DEM-S: ground surface topography with vegetation features removed and smoothed to improve surface shape and reduce noise; and
- DEM-H (1-second DEM only): hydrologically-enforced version of DEM-S.

The Australian Hydrological Geospatial Fabric is a specialised GIS based on the 1-second DEM, which defines logical hydrological features in the Australian landscape (see Excursus 13.2).

Table 13.2 Australian DEM

Name	Spatial resolution (m)	Based on
GEODATA 9-second DEM and D8 Flow Direction Grid 2008 (DEM-9S) v3	250	Elevation data at 1:100,000 and 1:250,000 cartographic scales
SRTM 3-second DEM v1.0	90	Shuttle Radar Topographic Mission, 2000
SRTM, 1-second DEM v10	30	Shuttle Radar Topographic Mission, 2000
Lidar 25 m grid	25	236 Lidar surveys between 2001 and 2015
Lidar 5 m grid	5	236 Lidar surveys between 2001 and 2015

Source: GA (2019)

Excursus 13.2—Australian Hydrological Geospatial Fabric

Source: BoM (2019a)
Further Information: *http://www.bom.gov.au/water/geofabric/documents/Geofabric_Info_Sheet_online.pdf*
http://www.bom.gov.au/water/geofabric/documentation.shtml

A collaborative development between the Bureau of Meteorology (BoM), Geoscience Australia (GA), ANU[6] and CSIRO[7] has produced the Australian Hydrological Geospatial Fabric (Geofabric). This specialised GIS registers the spatial relationships between important hydrological features such as rivers, water bodies, aquifers and monitoring points (BoM, 2019a). The spatial dimensions of these hydrofeatures indicate how they are connected, and how water is stored, transported and used through the landscape.

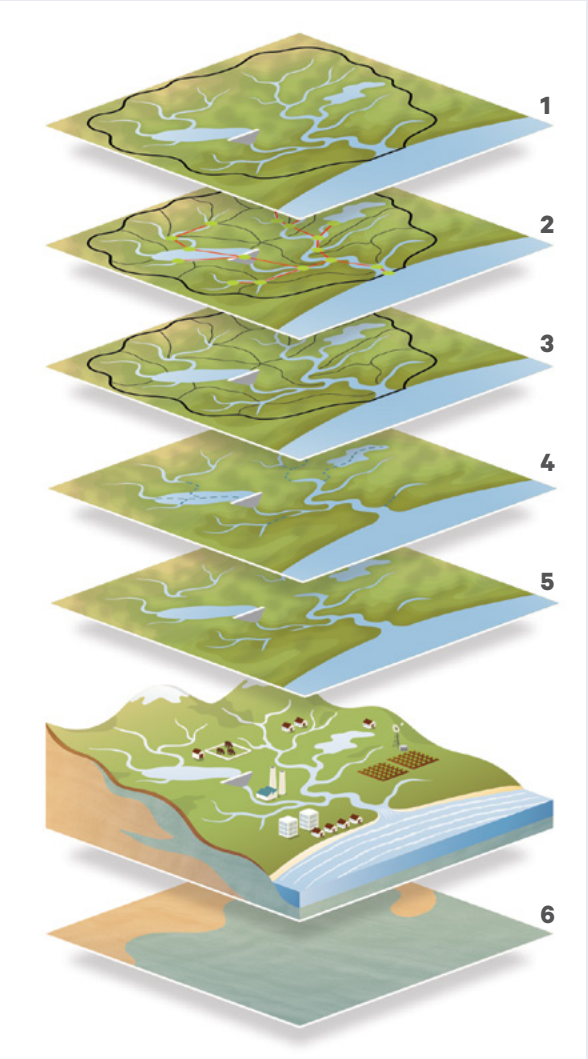
The Geofabric comprises six product datasets:

- (1) Hydrology Reporting Regions—define drainage divisions and river regions across Australia, for consistent reporting by governments and other organisations;
- (2) Hydrology Reporting Catchments—building blocks of reporting regions, providing increased detail for smaller rivers, including a simplfied stream network;
- (3) Surface Catchments—define the base level of catchments for stream segments, sinks and coastal draining areas. Universities, governments and consultants use these to identify contributing catchment areas;
- (4) Surface Network—provides a detailed, fully-connected and directed stream network. This enables stream flow paths to be traced, then linked to Surface Catchments;
- (5) Surface Cartography—allows surface water features, such as dams, canals and bridges, to be visualised. This context is useful for water managers and emergency services, for example, to anticipate how downstream communities may be affected by floods; and
- (6) Groundwater Cartography—shows groundwater resources and their features, such as aquifer boundaries, salinity, and rocks and sediments at different levels below the surface (see Figure 13.3).

The Geofabric is based on a set of contracted nodes (points), which represent important hydrological features in the landscape. These points have a permanent identifier included in all versions of the Geofabric, and provide a persistent framework through space and time. A subset of the contracted nodes is used to create a simplified node-link network with associated contracted catchments. Phase 2 of the Geofabric is based on 1:250,000 scale map features and the 9-second DEM. Phase 3 (released in 2019) is based on 1-second DEM and finer-scaled map features. The ANUDEM software (see Section 13.2) was used in development of data for the Geofabric.

Figure 13.3 Geofabric datasets

The Geofabric comprises six dataset products: (1) Hydrology Reporting Regions; (2) Hydrology Reporting Catchments; (3) Surface Catchments; (4) Surface Network; (5) Surface Cartography; (6) Groundwater Cartography.



Source: BoM (2015)

6 Australian National University Fenner School of Environment and Society
7 CSIRO Water for a Healthy Country Flagship

13.3 Meteorological Records

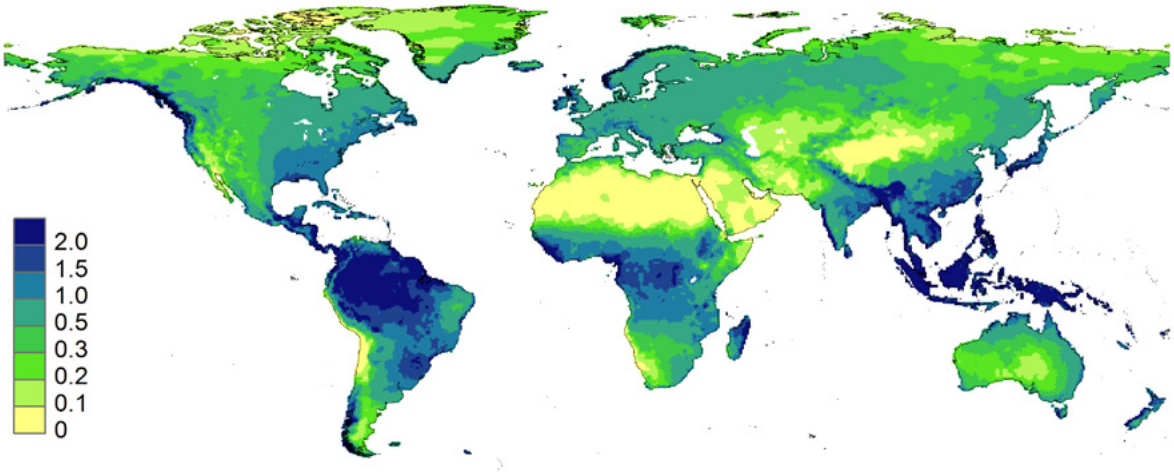
Meteorological records are another source of ancillary data that may be analysed in conjunction with EO datasets. Instrumental weather records in Australia started within months of the First Fleet and continued in various locations using available equipment and expertise. Standardised equipment was gradually introduced by the Bureau of Meteorology since 1908 to create the current Australian network of 752 temperature recording sites, nearly 6,000 rain gauges and 610 automatic weather stations (AWS; BoM, 2019b). The AWS observe air temperature, humidity, wind speed, wind direction and precipitation every 30 minutes (Isaac *et al.*, 2017).

These meteorological records have been calibrated and validated over time and space to ensure their accuracy and consistency, and spatially interpolated to create continuous, climate datasets. Some of the climate datasets now available for Australia include:

- Australian Climate Observations Reference Network (ACORN-SAT version 2)—Surface Air Temperature 1910–2018 (BoM, 2019c);
- daily and monthly estimates of precipitation, maximum daily temperature, minimum daily temperature and vapour pressure humidity (9 am and 3 pm) on a 0.05° grid (Jones *et al.*, 2009) derived from the national climate databank, Australian Data Archive for Meteorology (ADAM; CLIMARC, 2020);
- Australian Water Availability Project (AWAP; Raupach *et al.*, 2008); and
- daily and monthly observations and statistics for rainfall, temperature and solar exposure at weather stations (BoM, 2019c).

Figure 13.4 GPCC average precipitation 1982–2010

Precipitation units are shown in m/a.



Source: Donohue *et al.* (2013) Figure 1 © Wiley: Used with permission. Data source: Rudolf *et al.* (2010)

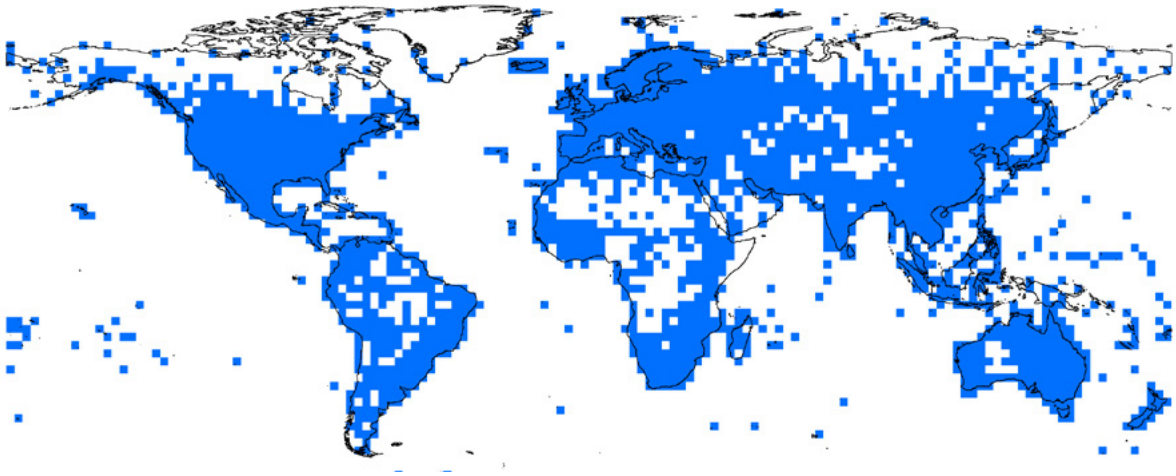
Some of the geostatistical techniques used to generate climate analyses for Australia are reviewed by Seaman and Hutchinson (1985) and Jones and Trewin (2000). For example, an anomaly-based approach, which decomposes each meteorological variable into its long-term average and an associated 'anomaly', was used by Jones *et al.* (2009) to create the 0.05° gridded datasets. While such datasets appear to provide continuous coverage, the spatial interpolation methods can generate data for large areas at considerable distance from a real weather station. Use of such data requires that consideration be given to the underlying observations, which may mean that data cells with significant temporal discontinuities generally need to be excluded from time series analyses. For example, when examining climate-related trends in Australian vegetation cover using EO imagery, Donohue *et al.* (2009) extracted site-based measures of montly precipitation and pan evaporation from the Monthly Australian Data Archive for Meteorology (MADAM) database (BoM, 2006) but only included sites with near-complete records, where complete was defined as:

- at least 25 days of data each month;
- at least 9 complete months of data each year; and
- at least 20 complete years of data for each station included (see Excursus 8.1).

Similarly, as part of an EO-based study to assess the impact of CO_2 fertilisation on global vegetation, Donohue *et al.* (2013) sourced global precipitation data from the Global Precipitation Climatology Centre (GPCC) Full Data Reanalysis Version 6 dataset (Rudolf *et al.*, 2010; see Figure 13.4). This spatially interpolated dataset spanned 29 years but included data-poor regions. Spurious cells were excluded by defining a buffer around those with less than one rainfall station for 90% of included months. The remaining data stations are shown in Figure 13.5.

Figure 13.5 GPCC buffer mask

Blue cells indicate the locations where near-continuous temporal coverage exists for the data shown in Figure 13.4.



Source: Donohue *et al.* (2013) Figure 2 © Wiley: Used with permission. Data source: Rudolf *et al.* (2010)

13.4 Further Information

Geographic Information Systems:

NationalMap: *https://nationalmap.gov.au/about.html*

Digital Elevation Models:

GA (2019)

CHARIM (Caribbean Handbook on Risk Information Management): *http://www.charim.net/datamanagement/32*

Australian Spatial Data Directory:

A national metadata hub for searching other national, state and territory directories to facilitate the discovery of published geospatial datasets throughout Australia:
http://www.nccarf.jcu.edu.au/terrestrialbiodiversity/index.php/41-general/272-australian-spatial-data-directory.html
https://www.australia.gov.au/information-and-services/business-and-industry/science-and-technology/spatial-information

Geoscience Australia:

Digital Elevation Data: *https://www.ga.gov.au/scientific-topics/national-location-information/digital-elevation-data*

Bureau of Meteorology:

National Environmental Information Infrastructure (NEII): *http://www.neii.gov.au/*

Climate online: *http://www.bom.gov.au/climate/data/*

Other climate data: *http://www.bom.gov.au/climate/data-services/*

Australian Landscape Water Balance: *http://www.bom.gov.au/water/landscape/#/sm/Actual/day/-28.4/130.4/3/Point////2019/8/11/*

Weather Stations: *http://www.bom.gov.au/climate/data/stations/about-weather-station-data.shtml*

CSIRO:

Soil and Landscape Grid of Australia: *http://www.clw.csiro.au/aclep/soilandlandscapegrid/About-FactSheet.html*

Global Meteorological Data:

Global Precipitation Climatology Centre (GPCC) Full Data Reanalysis Version 6: Rudolf *et al.* (2010)

Global vapor pressure deficit (based on European Centre for Medium-Range Weather Forecasts' ERA-interim reanalysis monthly near-surface ddew point and near-surface air temperature products: Dee *et al.* (2011)

Global DEM:

Intermap:
NEXTMap Elevation Data Suite: *https://www.intermap.com/nextmap*
Airbus Defence and Space: *https://www.intelligence-airbusds.com/elevation-models/*

NASA:
ASTER GDEM2: *https://asterweb.jpl.nasa.gov/gdem.asp*

USGS:
GTOPO30: *https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-30-arc-second-elevation-gtopo30?qt-science_center_objects=0#qt-science_center_objects*
SRTM: *https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-srtm-mission-summary?qt-science_center_objects=0#qt-science_center_objects*

JAXA:
ALOS World 3D: *https://www.eorc.jaxa.jp/ALOS/en/aw3d/index_e.htm*

13.5 References

ANU/GA (2008). *GEODATA 9 Second DEM and D8 Digital Elevation Model and Flow Direction Grid, User Guide*. Geoscience Australia.

Bolstad, P. (2008). *GIS Fundamentals: A First Text on Geographic Information Systems*. 3rd Edn. Elder Press, White Bear Lake, Minnesota. ISBN 978-0-9717647-2-9

BoM (2006). *Climate data: Australia CD-ROM*, Version 2.2, December 2006. Australian Government Bureau of Meteorology.

BoM (2015). *Australian Hydrological Geospatial Fabric factsheet*. Bureau of Meteorology, Melbourne. *http://www.bom.gov.au/water/geofabric/documents/Geofabric_Info_Sheet_online.pdf*

BoM (2019a). Bureau of Meteorology website. *About the Geofabric* webpage: *http://www.bom.gov.au/water/geofabric/about.shtml*

BoM (2019b). Bureau of Meteorology website. *Climate Data Online* webpage: *http://www.bom.gov.au/climate/data/*

BoM (2019c). Bureau of Meteorology website. *Long-term temperature record* webpage: *http://www.bom.gov.au/climate/data/acorn-sat/*

CLIMARC (2000). *Australian Data Archive for Meteorology. A draft document prepared for Cli-Manage 2000*. Bureau of Meteorology, Melbourne.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. *https://doi.org/10.1002/qj.828*.

Donohue, R.J., McVicar, T.R., and Roderick, M.L. (2009). Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006, *Global Change Biology*, 15(4), 1025–1039, *https://doi.org/10.1111/j.1365-2486.2008.01746.x*.

Donohue, R.J., Roderick, M.L., McVicar, T.R., and Farquhar, G.D. (2013). Impact of CO_2 fertilisation on maximum foliage cover across the globe's warm, arid environments. *Geophysical Research Letters*, 40, 1–5. *https://doi.org/10.1002/grl.50563*

GA (2019). Geoscience Australian website, *Digital Elevation Data* webpage: *http://www.ga.gov.au/scientific-topics/national-location-information/digital-elevation-data*

Gallant, J.C., and Hutchinson, M.F. (2008). Digital terrain analysis. Ch. 6 in *Guildelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

Hewitt, A.E., McKenzie, N.J., Grundy, M.J., and Slater, B.K. (2008). Qualitative Survey. Ch 18 in *Guildelines for Surveying Soil and Land Resources*. Australian Soil and Land Survey Handbook Series. (Eds: McKenzie, N.J., Grundy, M.J., Webster, R., and Ringrose-Voase, A.J.) CSIRO Publishing, Melbourne.

Isaac, P., Cleverly, J., McHugh, I., van Gorsel, E., Ewenz, C., and Beringer, J. (2017). OzFlux data: network integration from collection to curation, *Biogeosciences*, 14, 2903–2928, *https://doi.org/10.5194/bg-14-2903-2017*

Jones, D.A., and Trewin, B.C. (2000). On the description of monthly temperature anomalies over Australia. *Australian Meteorological Magazine*, 49, 261–76.

Jones, D.A., Wang, W., and Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, 58(4), 233–248.

McKenzie, N.J., Gessler, P.E., Ryan, P.J., and O'Connell, D. (2000). The role of terrain analysis in soil mapping. In *Terrain Analysis: Principles and Applications*. (Eds: Wilson, J.P., and. Gallant, J.C.). Wiley, New York.

Raupach, M.R., Briggs, P.R., Haverd, V., King, E.A., Paget, M., and Trudinger, C.M. (2008). *Australian Water Availability Project*. CSIRO Marine and Atmospheric Research Component. Final Report for Phase 3. Available from CSIRO, Canberra, Australia.

Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A.,and Ziese, M. (2010). *GPCC Status Report*. 7 pp, Global Precipitation Climatology Centre.

Seaman, R.S., and Hutchinson, M.F. (1985). Comparative real data tests of some objective analysis methods by withholding observations. *Australian Meteorological Magazine*, 33, 37–46.

Speight, J.G., and McDonald, R.C. (2009). *The Site Concept*. In: *Australian Soil and Land Survey: Field Handbook* (3rd Edn). (Eds: National Committee on Soil and Terrain) pp. 73–125. CSIRO Publishing, Melbourne.

Integrated Solutions

Environmental conditions are notoriously difficult to measure, let alone monitor, and considerable ecological research has not produced standard, dependable methods for measuring environmental 'condition'. However, parameters that are believed to be indicative of particular conditions can be measured (see Figure 1.1). Repeated measurements over time have been used to model the 'normal' cycles for selected environmental attributes and such structure models are now being used operationally to 'calibrate' EO-based measurements to highlight significant surface changes.

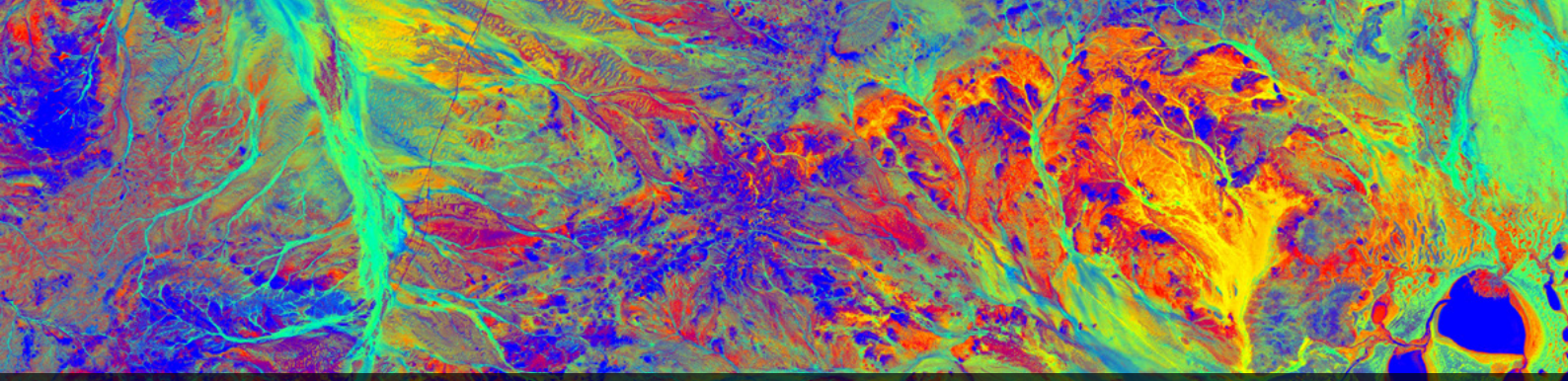
These final sections present examples of integrating EO information products with environmental process models to deliver current, regional and continental scale information about the Australian environment. As introduced in Section 1.4, EO analyses are generally part of a 'bigger picture', involving other datasets and producing pre-defined and reproducible metrics for a specific application. Once reliable datasets are available, the next step in making EO data accessible for a wide range of applications is the generation of information products that reliably condense the relevant information contained within EO datasets into attributes that directly relate to measurable surface features (see Section 14.1).

Information products can then be integrated into specific environmental monitoring and modelling applications, such as environmental accounting (see Section 14.2.1), compliance monitoring (see Section 14.2.2) and modelling of ecosystem dynamics (see Section 14.3). A wider range of application areas reliant on EO datasets are detailed in Volume 3.

Contents

14 Mapping, Monitoring and Modelling 161

Background image on previous page: The High Tide – Low Tide (HTLC) low tide composite mosaic showing the Kimberley Coastline and Montgomery Reef, Western Australia at low tide, with the overlaid InterTidal Elevation Model (ITEM v2.0) products illustrating the topography of the exposed intertidal mudflats and reefs in the region (see Section 11.2). **Source:** Stephen Sagar, Geoscience Australia



14 Mapping, Monitoring and Modelling

EO provides an opportunity to obtain regular, objective observations of energy levels being reflected or emitted by the Earth's surface (see Volume 1). Provided that sensor calibrations are reasonably accurate, the corrected measurements can usually be considered consistent within a given image and comparable between images (see Volumes 2A and 2B), with variations in atmospheric and illumination conditions over time being accounted for by suitable measurement models (see Figure 1.1 and Section 3). The importance of data provenance for EO imagery was introduced in Section 2.2 above and Volume 2A—Section 2.1. This is especially important when processing multiple images or time series datasets. Relevant caveats and processing methodologies for time series datasets are described in Sections 8 to 10 above.

Analysis of a single, calibated EO image enables mapping of surface features. This inventory approach describes surface characteristics at one point in time. A variety of image processing techniques can be applied to a single image to map specific surface features, including spectral indices (see Volume 2C), classification (see Volume 2E) and sub-pixel analyses (see Volume 3A). Some of the EO-based mapping products that are now routinely produced for Australia are introduced in Section 14.1.

The change detection methods described in Section 7 offer a mechanism for highlighting significantly different pixels in a pair of bi-temporal images. However, by definition, monitoring involves regularly-repeated measurements—this is not simply a case of detecting the changes that have occurred between two points in time. Change detection can be seen as one stage of monitoring only. To be effective, changes need to be monitored regularly and for a sufficient time period, which is greater than known natural cycles. Time series datasets offer the potential to achieve regular environmental monitoring. Section 14.2 considers some of the challenges for environmental monitoring using EO datasets, then provides two examples of environmental monitoring approaches in Australia: environmental accounting (see Section 14.2.1) and compliance monitoring (see Section 14.2.2). Environmental monitoring is further discussed in Volume 3 in the context of specific application areas.

Finally the integration of EO datasets with environmental modelling applications is introduced in Section 14.3.

14.1 Mapping

A range of image processing techniques have been used to map specific features on the Earth's surface using EO image datasets (see Volumes 2A, 2C and 2E). In particular, satellite-derived imagery offers a unique perspective for regional and global mapping. Potential mapping applications include:

- vegetation type, condition and structure (see Volume 3A); and
- water colour, temperature, extent and quality (see Volume 3B).

The distinction between EO data and information products was introduced in Volume 2A—Section 2.1 and reviewed in Section 2.4 above. In the context of Table 1.1, EO data products can be seen as examples of 'measurements' rather than observations, and information products represent application-specific attributes. For these relationships to hold, however, requires that the generation of both data and information products follows well-defined and validated procedures (see Section 2.5).

Background image: Fractional cover image derived from Landsat-8 OLI image over Three Rivers Station, near Kumarina in Western Australia. Fractional cover splits the landscape into fractions of green (leaves, grass, and growing crops), brown (branches, dry grass or hay, and dead leaf litter), and bare ground (soil or rock) and is used to monitor vegetation extent and grazing pressure (see Figure 2.3 and Volume 3A) Source: Norman Mueller, Geoscience Australia

Excursus 14.1—TERN Continental Products

Source: *https://www.tern.org.au/*
Further Information: *https://portal.tern.org.au/#/8063c731; https://maps.tern.org.au/#/*

TERN is Australia's land-based ecosystem observatory, delivering a wide range of data streams for environmental research and management (see Excursus 12.2). This critical research infrastructure enables Australians to track, understand and respond to changes in ecosystems—both through time and across the breadth of the continent, by providing ready access—for community members, policy-makers, industry leaders and scientists—to national information about the ecosystems on which our livelihoods, lifestyles and identity depends.

The TERN Data Discovery portal is a national data delivery service and expert network that provides a range of EO information about Australian environments. The major products include:

- time-series of key environmental variables—derived from AVHRR, MODIS, Landsat, and SPOT imagery;
- high-resolution airborne imagery and field datasets for particular sites; and
- essential ground calibration and validation datasets for airborne and satellite image data.

Products listed in Table 14.1 demonstrate the range of validated EO-based continental products that TERN has delivered over the past 10 years to address landscape monitoring, change and diversity. The list of products that TERN develops with its partners and delivers through the TERN Portal changes over time due to changing priorities and the availability of source EO or validation data.

Table 14.1 TERN EO-based continental products

Product	Derived from	Spatial Resolution	Temporal Resolution	Temporal Extent
Normalised Difference Vegetation Index	AVHRR (NOAA)	1 km/ 5 km	10 day composite/ 1 month composite	1992–2014
Fraction of Photosynthetically Active Radiation	AVHRR (NOAA)	1 km/ 5 km/ 8 km	1 month composite	1995–2011 (1/5 km) 1981–2011 (8 km)
Grassland curing	MODIS (Terra/Aqua)	500 m	8 day composite	2000+
Enhanced Vegetation Index (EVI) and Land Surface Temperature (spatially and temporally integrated)	MOD13A1/ MOD11A2	500 m	8 day composite	2000–2014
Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR)	MOD15A2(c5)	1000 m	8 day composite	2000+
Gross Primary Productivity (GPP)	MOD17A2(c5)	1000 m	8 day composite	2000+
Fractional Cover (FC)	MCD43A4	500 m	8 day composite / 1 month composite	2000+
Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI)	MOD13Q1(c5)	250 m	16 day composite	2000+
Land Condition Index (LCI)	MCD43A4	500 m	16 day composite	2000–2011
Fraction of Photosynthetic Radiation absorbed by Chlorophyll	MOD13C2	5 km	1 month composite	2000–2014
Ecosystem Disturbance Index	MOD13A1/ MYD11A2	500 m	12 month composite	2000–2013
Phenology	MOD13C1 EVI	5 km	12 month composite	2000–2015
Land Cover Type	MCD12Q1(c5.1)	500 m	12 month composite	2001+
Land Cover Dynamics	MCD12Q2(c5)	500 m	2 year composite	2001+
Seasonal Fractional Cover	Landsat TM/ETM+/ OLI	30 m	Seasonal composite	1986+
Persistent Green-Vegetation Fraction and Wooded Mask	Landsat TM/ETM+	30 m	Single product	2000–2010
Vegetation Height and Structure	GLAS (ICESat), PALSAR, Landsat	30 m	Single product	2003–2009

An increasing number of EO information products are becoming available for a wide variety of applications. For example, some of the continental-scale, information products that have been derived from EO imagery over Australia by TERN are listed in Excursus 14.1. A comparable set of information products relevant to water applications are available from the Australian Ocean Data Network (AODN; see Section 14.4). Examples of other information products are detailed in Volume 3.

14.2 Monitoring

Monitoring generally implies continuous, or regular, checks or measurements on some condition or attribute. The dynamics, that is the extent and type of change, of the condition indicate its stability, and possibly its independence relative to other known conditions. To interpret monitoring information sensibly requires that the normal amplitude and frequency of variation for the condition are known. Extraordinary changes can then be identified when the monitored measurements indicate a greater or lesser variation than expected for a given time interval. The emphasis here is on continuous measurements with some expectation of the range of values which would result from a 'normal' situation. Abnormalities are then identified by their deviation from the presumed norm.

Environmental monitoring is concerned with observing and quantifying changes in the environment. Topical environmental issues concerned with monitoring change include deforestation, pollution levels, sedimentation in water bodies, urban development, and the severity and extent of natural disasters, such as fire, flood, subsidence, volcanic activity and storms. The relationship between the environmental parameter (or attribute) domain and the image data (or observation) domain is summarised in Figure 1.1. Starting with a physical parameter, such as water depth, a structure model relates the parameter space to a measurement space, such as reflectance in a given wavelength range. A measurement model then relates the measurement space to the image data space, namely pixels, channels and calibrated values (see Section 3.1).

Effective monitoring using EO data must rely on this framework of models, which relates image values to physical, independently-measurable parameters. A variety of physically-based models are being developed in a diverse range of applications that offer a potential basis for using EO to monitor environmental change (see Section 14.3 and Volume 3). Traditional image analysis techniques that highlight changes over ad hoc time periods can offer constructive information pertaining to environmental changes but should not be viewed as monitoring.

Multi-temporal EO datasets provide valuable information for monitoring a wide range of environmental attributes. As detailed in Section 2.1.2, extensive archives are now available for several satellite sensors, including AVHRR, Landsat, MODIS and SPOT-Vegetation. These repositories offer a unique set of observations to explore phenological and other cyclic changes in biosphere properties (see Section 14.2.1). Environmental legislation has also prompted monitoring the status of specific natural resources, such as native vegetation, inland water quality and carbon sequestration (see Section 14.2.2 and Volume 3).

Finally, detecting reflectance changes between two image dates is not the same as identifying the changes on the ground, and is at least one step removed from linking actual changes to causal agents. Some detected changes may be more significant than others, both depending on management objectives and the severity of change. Additional information is generally required to differentiate between short-term, cyclic changes resulting from natural events, and longer-term changes resulting from anthropogenic activities or climatic variations.

14.2.1 Environmental Accounting

Environmental accounting broadly refers to tracking relevant environmental attributes to monitor the condition of natural resources. In order to manage these resources sustainably, environmental accounts need to be detailed, frequent and comprehensive. In common with the more traditional forms of accounting, environmental accounts essentially aim to 'track the transfer of value through time, between locations, and between owners (including households, businesses, and governments)' (BoM, 2013). By contrast, ecosystem accounting is defined as an '*integrated statistical framework for organising biophysical data, measuring ecosystem services, tracking changes in ecosystem assets and linking this information to economic and other human activity.*' (UN, 2014).

Environmental accounting is promoted in Australia and around the world as a way of providing regular information on changes in the state of the environment to support policy development and decision-making. (Vardon et al., 2018)

Sustainability, by definition, involves preserving the capital base. To make informed decisions, policy makers need more comprehensive information about Australia's natural capital base upon which to judge the sustainability of the systems they manage and the best mix of economic, social, human and environmental policies. (BoM, 2013)

There is a growing awareness that economic growth and living standards directly depend on environmental assets, such as land, soil, minerals, rivers, oceans and biodiversity (DEE, 2018). Accordingly, sustainable management of these resources is essential for the future livelihood of all socio-economic sectors of our community. Environmental-economic accounting aims to understand both the condition of the environment and its relationship with the economy. It can be used to define and demonstrate the relationships between different economic sectors and specific environmental resources, and increasingly underpins the development and implementation of government policy and legislation (see Section 14.2.2). EO data, which is spatially explicit with global coverage and standardized over space and time, provides a unique input for environmental accounting.

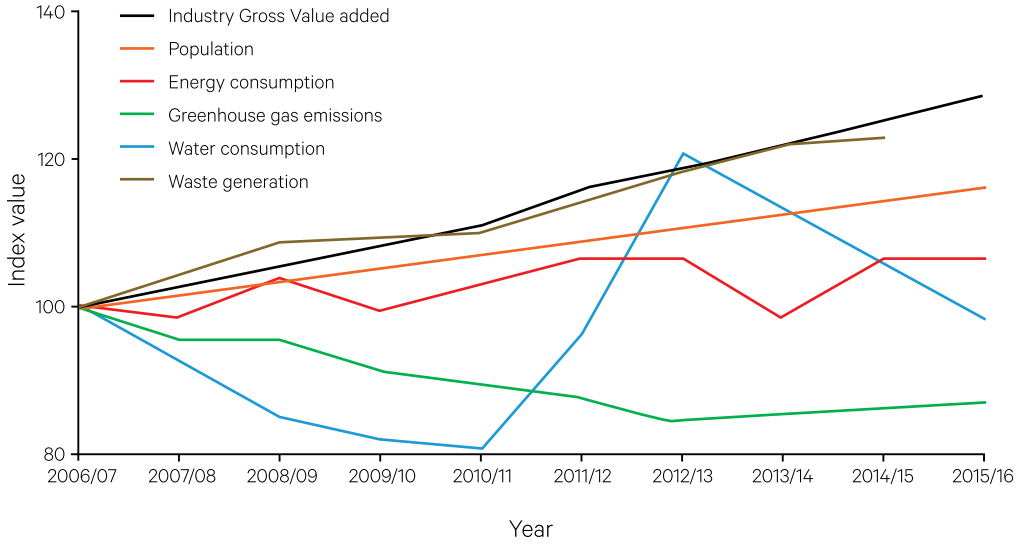
For example, ABS generates an annual set of environment-economic accounts in accordance with the System of Environmental-Economic Accounting (SEEA[8]; an international statistical standard that is consistent with the System of National Accounts). From 2002/03 to 2015/16, Australia's consumption of water and energy, emissions of greenhouse gases and generation of waste are compared with the value of economic production and change in population in Figure 14.1. These comparisons confirm that, since 2006/07, the economy has grown faster than both population and the selected environmental consumption measures, which indicates that resources are being used efficiently.

An introduction to environmental accounting in Australia is presented in BoM (2013). While ground measurements are an essential part of environmental accounting, they are not sufficient to capture the regional and continental perspective. EO imagery, however, with contiguous, calibrated observations, provides a unique view of the Earth (see Volume 1). When integrated with ground measurements and biophysical models, EO datasets can be used to monitor a wide range of environmental attributes, such as vegetation cover and condition, water resources, carbon cycling, and natural disasters.

One example of environmental accounting is the Australian model-data fusion system, OzWALD (see Excursus 14.2). This system assimilates various EO datasets into a water balance model (a variant of AWRA-L; Van Dijk, 2010; Frost *et al*., 2016) to estimate vegetation properties and energy, water and carbon balance parameters across Australia. OzWALD is used to generate annual updates on the condition, change and trajectory of the Australian environment (ANU, 2019).

Figure 14.1 Australian environmental-economic account for agriculture

Selected socio-economic and environmental measures for Australia are shown from 2006/07 to 2015/16.



Source: ABS (2018)

8 The SEEA defines a flexible accounting framework that outlines methods to account for ecosystems in non-monetary terms (BoM, 2013).

Excursus 14.2—OzWALD

Source: Albert van Dijk, Australian National University
Further Information: van Dijk *et al.* (2017); van Dijk and Rahman (2019); *www.ausenv.online*

The Australian Water and Landscape Dynamics (OzWALD) system is a modular, data production workflow that is implemented using a combination of MatLab and Python scripts and libraries and executed on the National Computational Infrastructure (NCI)[9]. The system is largely automated with processing requirements amounting to around 30,000 processor-hours on the NCI's High Performance Computing Infrastructure. The volume of input data processed by OzWALD totals over 30 TB daily, which is condensed into annual summaries that total around 1 GB.

Development of OzWALD followed a review of national environmental reporting and accounting in Australia by van Dijk *et al.* (2014) which concluded 'Despite strong demand for information [and] considerable effort and investment, nation-wide environmental data collection and analysis remains a substantially unmet challenge'. Subsequent research demonstrated that environmental data, which is already being collected in Australia, can be usefully synthesised, interpreted and reported as an annual summary provided the technical challenges of routine data processing, web-based data access and effective communication of results are addressed.

The OzWALD system currently comprises five basic components as detailed below. Its statistical summaries can be visualised for selected time periods and regions via the Australia's Environment Explorer (AEE) and are reported in annual factsheets called Australia's Environment.

Data acquisition

Data from several gridded climate and satellite sources, most of which are stored in the NCI, are integrated in OzWALD, including:

- satellite-derived imagery produced by NASA:
  - surface reflectance and albedo (MCD43A4.006 and MCD43A3.006; Schaaf and Wang, 2015);
  - vegetation leaf area index (LAI, MCD15A3H.006; Myneni *et al.*, 2015); and
  - land surface temperature (LST, MYD11A1.006; Wan *et al.*, 2015);
- estimates of the fractions of bare soil, photosynthetic vegetation and non-photosynthetic vegetation derived from MCD43A4.006 (Guerschman and Hill, 2018) and produced by CSIRO;
- Landsat-based water presence mapping data (Water Observations from Space (WoFS; see Volume 1A—Excursus 5.1); Mueller *et al.*, 2016) available from Digital Earth Australia (DEA; see Section 11.2);
- precipitation and short-wave incoming radiation produced as part of the Australian Gridded Climate Data (AGDC; Jones et al., 2009) by Bureau of Meteorology (BoM); and
- hourly temperature, humidity, pressure and wind speed produced by the European Centre for Medium-range Weather Forecasting (ERA5, ECMWF; Copernicus Climate Change Service, 2017). These data are available and accessible to NCI users, while most can also be accessed via NCI THREDDS[10].

In addition, the following data are downloaded as a scheduled daily process or on demand:

- satellite-derived rainfall estimates (IMERG, Huffman *et al.*, 2015) from NASA[11];
- satellite-detected fire occurrence and intensity from the Sentinel Hotspots system (Geoscience Australia, 2014)[12]; and
- fire carbon emission estimates (Kaiser *et al.*, 2012) from ECMWF[13].

The stability, reliability, latency and continuity of data services are important factors that constrain workflow outcomes. For example, replacing MODIS Collection 5 with Collection 6 introduced a series of technical challenges.

Data reformatting

All original (raster and vector) data are resampled and reformatted to a single gridded standard in geographic coordinates and stored in NetCDF file format (see Volume 2A—Section 1.6). Data are temporally aggregated to a minimum daily resolution where necessary.

9 *www.nci.org.au*
10 *http://dapds00.nci.org.au/thredds/*
11 *https://pmm.nasa.gov/data-access/downloads/gpm*
12 *https://sentinel.ga.gov.au/*
13 *https://apps.ecmwf.int/datasets/data/cams-gfas/*

Data fusion and downscaling

Data fusion, blending of like data into a single consistent best-estimate data set, is applied to produce observation-based estimates of:

- precipitation—the AGCD and IMERG data grids are merged to estimate daily accumulated rainfall. Since the AGCD are derived by interpolating point data from weather stations, they are more accurate than the IMERG estimates where gauge density is greater than 1 /~500 km2 but become gradually less accurate at lower densities (Renzullo *et al.*, 2011). To select the most accurate rainfall estimates, an inverse-distance weighting scheme is used to blend the AGCD and IMERG datasets; and
- surface water extent—an internally-consistent data set of surface feature type (at 500-m and 8-day resolution) is derived from the fractions of surface water, snow, bare soil and living and dead vegetation, as well as LAI and albedo using the Water And Landscape Dynamics Multi-Observation Reanalysis and Filtering (WALDMORF) algorithm. This process involves various checks and balances, including data fusion methods (see Section 6.2), interpolation of missing values (see Volume 2A), and detecting outliers (see Volume 2E).

Climate data is downscaled to 500 m spatial resolution using the DEM-H product to enforce topographic and elevation consistency (Gallant *et al.*, 2011; see Section 13.2). Thus far, downscaling has been applied to terrain illumination from incoming shortwave radiation, air temperature, and vapour pressure.

Model-data assimilation

Biophysical modelling is used to estimate environmental variables for which spatial data are not directly available, such as root zone soil moisture, streamflow generation, and carbon uptake by the vegetation. The OzWALD model evolved from AWRA-L v0.5 (Van Dijk, 2010), which formed the core of BoM's operational daily landscape water balance modelling system (Frost *et al.*, 2016), the associated spatial information service[14], and the AWRA Community Modelling System[15]. The workflow steps in this multi-layer model are summarised in Table 14.2. Workflows for both the spin-up and final models comprise:

- retrieving and transforming climate forcing and satellite-derived surface properties;
- updating model parameters using satellite dataset with a simple nudging scheme;
- evolving model states one day forward using a daily time step model; then
- concatenating and storing the requested data.

Table 14.2 OzWALD model components

Module	Workflow description
HPC job generation	deploy workflow for specified model domains
Model workflow management	set up system environment
	initialise static model parameters and model states
	optional spin-up model run for a specified period to produce 'hot' initial states if needed
	produce and store the requested variables
Post-processing	collate and store outputs into annual national grids

Statistical summary

The daily or 8-daily, continental, 500-m resolution data files produced as described above are published through NCI THREDDS. For many users these results may be challenging to interpret, so temporal and regional summary statistics are calculated in a final post-processing step as these are what users are most likely to visualise through web applications:

- Annual summary maps—most commonly minimum, maximum and average and/or accumulated daily values are calculated for the year, and these are stored at the original spatial data resolution in gridded NETCDF format.
- Regional-average annual time series—annual summary maps are intersected with different region definitions derived from thematic vector maps. Because of their small data volume, these summary data are stored as CSV files on an external web server.
- Regional-average sub-annual time series—are calculated at the original temporal and stored as NetCDF files on NCI THREDDS with the two spatial dimensions being replaced by the region identifier.
- Annual regional summaries by broad land use type—as computed for the regional-average annual time-series maps described above with further stratifying by broad land-use type, and as cumulative spatial sums rather than averages where appropriate. The land-use classes are an amalgamation of the many classes in the 50 m resolution Catchment-scale Land Use Data (ABARES, 2017) into 20 broad classes. The resulting data are quite comparable to that used to construct environmental accounts.

14 *http://www.bom.gov.au/water/landscape/*
15 *https://github.com/awracms/awra_cms*

Visualisation and Analysis

OzWALD statistical summaries can be visualised via the Australia's Environment Explorer (AEE) website[16]. This interactive facility enables users to explore environmental changes by region, location or land use type which are presented as maps and charts (see Figure 14.2). Data can be queried for a particular location, or as the average across a predefined region, where region can be defined as:

- political and administrative units (states and territories, local government areas, federal electorates, natural resource management regions, national parks, Ramsar wetlands);
- biophysical units (bioregions, river regions and drainage divisions); and
- statistical accounting regions.

Regional summary data can be directly downloaded via the AEE website or edited online using the plotly visualization app (*http://plot.ly/*). The original and summary gridded time series can be accessed and downloaded from the NCI Data Collection, which provides an API to the data. All data is available under a Creative Commons license (CC BY 3.0 AU). Sub-annual and annual data produced by OzWALD are also available through the ANU-TERN Landscape Data Visualiser (*http://maps.tern.org.au*), a web app designed to help researchers and other interested users to visualise and explore data collected or produced by the Terrestrial Ecosystem Research Network (see Excursus 12.2) or ANU-WALD. Unlike the AEE, this website provides tools to visualise sub-annual gridded data, and to compare different datasets (such as flux tower measurements and gridded data) for a location of interest (see Volume 3A).

Regional and national Environmental Condition Scores (ECS) are also produced for seven variables by ranking annual values of each variable within the time series since the year 2000 (see Figure 14.3). Interpretation of the ECS component scores confirms the scientific consensus that environmental condition is directly dependent on water availability for most of Australia, which partially mitigates its limitations.

Figure 14.2 Australia's Environment Explorer

This example screenshot from Australia's Environment Explorer shows national persistent vegetation cover (see Volume 3A).

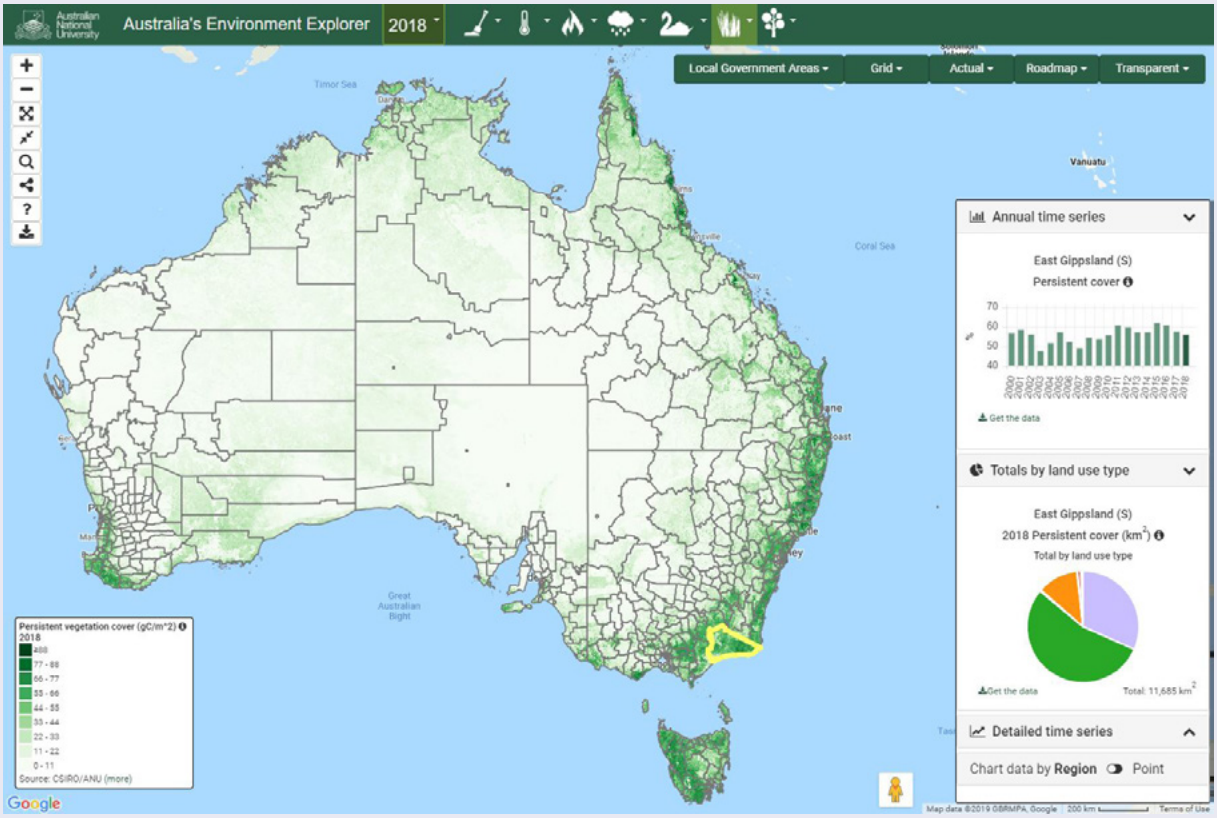
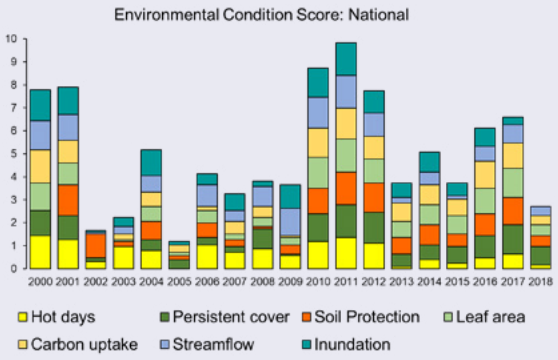
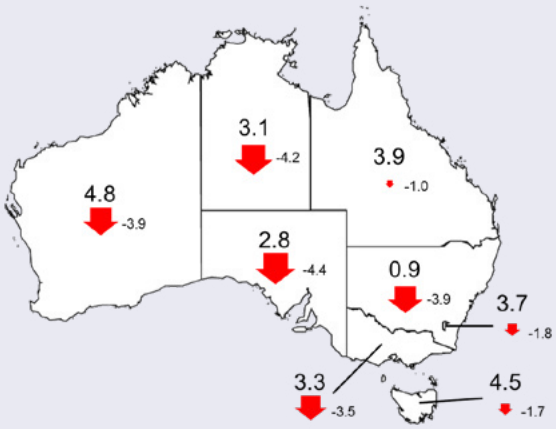


Figure 14.3 2018 OzWALD summary for ACT

a. Annual environmental condition scores have been computed since 2000 as the average of seven components.



b. National map showing overall change from 2017 to 2018



Source: van Dijk and Rahman (2019) Figure 3

Australia's Environment Report

The Australia's Environment (AE) report[17], which synthesises information from other sources and interprets specific events and temporal trends in global or non-spatial data (including ABS, BoM, IMOS, TERN, and Atlas of Living Australia, and the WMO), is produced annually. This provides additional opportunities to interpret information and report on time series representing smaller or unspecified areas within Australia, specific events and their impacts, different parts of the environment, and the global context.

Status Overview

The OzWALD system and AEE demonstrate that it is feasible to produce useful, observation-based annual environmental reports. Insights from this research and development include the:

- challenges to operational workflows presented by an unstable, and continually evolving, spatial data services environment;
- efficiencies provided by the rapid development of open source and cloud technologies;
- necessity of regular, detailed and accurate land cover and land use mapping to achieve successful environmental accounting;
- opportunities for environmental reporting enabled by the abundance of past, current and future satellite mission datasets; and
- ongoing difficulties in obtaining regular and reliable biodiversity data.

The goal of OzWALD is to continue—and as much as possible improve—the annual, environmental modelling and reporting process until a similar service is available from another source.

17 *http://wald.anu.edu.au/australias-environment/*

Problems cannot be solved with the same mind set that created them.
(Albert Einstein)

14.2.2 Compliance Monitoring

The responsibility for a range of regulatory functions under legislation rests with various government departments and agencies at federal, state and local levels. In order to achieve the specific objectives of each piece of legislation, regulated communities need to comply with defined rules or standards, which are frequently presented in terms of conditions associated with approvals, permits or licences (DEE, 2016). Compliance monitoring describes the bureaucratic process of ensuring that the community obligations to legislation are met.

An increasing volume of legislation relates to environmental protection in some form. For example, at a federal level, compliance monitoring activities are undertaken under the legal frameworks of four pieces of legislation:

- Environment Protection and Biodiversity Conservation Act 1999 (Parts 7, 9, and 10)—protects and manages Australia's land and marine biodiversity, threatened species, ecosystems, environment and heritage;
- Environment Protection (Sea Dumping) Act 1981—regulates the dumping of waste at sea, incineration at sea, and the placement of artificial reefs;
- Fuel Quality Standards Act 2000— places obligations on the fuel industry to ensure that supplied fuel meets strict environmental and human health requirements; and
- Ozone Protection and Synthetic Greenhouse Gas Management Act (1989)—helps to ensure that Australia meets its legal obligations under the Montreal Protocol (1989) and the United Nations Framework Convention on Climate Change.

Given the extent of Australia's land and maritime resources, EO-based datasets are playing an increasing role in compliance monitoring. For example, several Australian state governments have implemented legislation to protect natural biodiversity and restrict clearing of native vegetation. Regulation of these activities requires a benchmark from which changes in vegetation—and by inference, biodiversity—can be assessed. The benchmark that has been adopted for Australia is the estimated condition of vegetation in 1750, prior to European settlement (see Volume 3A).

To ensure that land owners are complying with these legislation, sophisticated systems based on EO datasets have been developed by several state governments. An overview of the well-established SLATS (State-wide Landcover and Trees Study) system used in Queensland and NSW is presented in Excursus 14.3.

Compliance monitoring ensures that the regulated community undertakes actions that are in accordance with the respective legislation. By ensuring that the regulated community does this, impacts to the environment and human health can be managed and minimised.
(DEE, 2016)

Excursus 14.3—SLATS: Monitoring Woody Vegetation Clearing

Source: Tim Danaher, NSW Department of Planning, Industry and Environment and Dan Tindall, Queenland Department of Environment and Science
Further Information: SLATS Queensland: *https://www.qld.gov.au/environment/land/management/mapping/statewide-monitoring/slats*
SLATS NSW: *https://www.environment.nsw.gov.au/topics/animals-and-plants/native-vegetation/reports-and-resources*

Background

Vegetation types within Australia are diverse due to many factors including climate, geology, soils, presence or absence of fire, and land management practices (see Volume 3A). Historic patterns of development and land management have altered, and in some cases degraded, many of these landscapes through land clearing, altered fire regimes and other land management activities, such as grazing.

In the past 20-30 years, there has been greater recognition of the legal and environmental responsibilities of state governments and the community to balance the production and conservation interests in these often fragile landscapes. This includes allowing appropriate clearing of vegetation but avoiding further degradation of these systems. The debate about how to achieve that balance is often polarised by different stakeholder interests and may be informed by data that is either incomplete or has limited spatial extent. Objective information about land cover changes, which is both spatially and temporally comprehensive, is needed to:

- inform policy decisions;
- support legislation regulation and compliance; and
- assist landholders and non-government organizations with managing these lands.

Satellite imagery, particularly imagery derived from systematic acquisition programs such as Landsat, provides an opportunity to monitor land cover change in an objective, repeatable and comprehensive way.

In 1995, the Queensland Government established a monitoring initiative known as the Statewide Landcover and Trees Study (SLATS) to provide monitoring of vegetation based on Landsat TM/ETM+/OLI satellite imagery across the entire state. The Landsat imagery delivered a suitable spatial resolution and extent for informing policy and regulatory compliance under relevant legislation, as well as restoration and other land management actions across the state. With a spatial resolution of 30 m, this imagery enabled the reliable delineation of woody vegetation changes greater than 1 ha, and direct calibration, validation, and interpretation of those changes in the field. The SLATS continues to monitor land cover change in Queensland annually.

The NSW government has also used the SLATS method for monitoring vegetation since 2006. The data generated by SLATS in Queensland and NSW has provided the foundation for a number of changes to vegetation and biodiversity management legislation and other land management and monitoring programs. This has included, for example, globally significant programs relating to the management of the Great Barrier Reef (see Volume 3B) and the Murray-Darling Basin (see Volume 3A).

In recent years, with the advent of higher resolution satellite missions (such as SPOT-5 and the European Space Agency's Sentinel-2A and -2B as part of their Copernicus programme), the SLATS methodology has been adapted to make use of these higher spatial resolution sensors. Firstly, NSW changed to using higher resolution 5 m resolution SPOT-5 imagery in 2009 and more recently Queensland and NSW have both transitioned to 10 m resolution Sentinel-2 imagery.

SLATS model

The primary objective of the SLATS methodology is to map the location and extent of woody vegetation clearing that results from anthropogenic (human-induced) removal of vegetation across an entire Australian state. The reduction in woody vegetation cover due to clearing needs to be mapped at time intervals of no longer than 1-2 years. Longer periods may miss some clearing events, since regrowth (which can have significant canopy density) and additional pasture or other herbaceous vegetation may establish within the intervening period, making it difficult to reliably detect and map changes in woody vegetation.

SLATS monitors woody vegetation clearing using a combination of automated and manual mapping techniques, primarily based on medium resolution satellite imagery acquired from the Landsat, SPOT-5 and Sentinel-2 satellites. The manual editing stage is supported by ancillary data sources, including higher resolution satellite imagery, and is an essential component for ensuring an accurate account of woody vegetation clearing across each state.

The SLATS methodology relies on efficient data management and processing systems that have been purpose-built on open source software foundations.

These systems enable largely automated processing of satellite imagery workflows in a high performance computing environment. Data management and software version control systems enable data provenance tracking and code management, and ensure efficiency and traceability in all stages of the mapping and analysis processes. A schematic overview of the SLATS methodology is shown in Figure 14.4 and an example of a clearing event mapped by SLATS is shown in Figure 14.5. The key processing and analysis steps can be summarised as follows:

1. Satellite imagery is acquired, then corrected for sensor calibration, atmospheric, topographic effects and Sun and sensor viewing angles (Flood *et al.*, 2013; see Section 3.1). The most cloud-free images from the dry season period are selected for each year (see Figure 14.5a and Figure 14.5b).
2. A woody vegetation clearing index is calculated to detect areas of change that represent possible clearing of woody vegetation. This model has been calibrated using historic mapping of cleared areas (Scarth *et al.*, 2008; Danaher *et al.*, 2011). It highlights most of the possible clearing and omits areas that are almost certain not to represent clearing (see Figure 14.5c and Figure 14.5d).
3. This initial clearing index is visually inspected, and manually edited by trained remote sensing scientists to confirm areas that have been cleared and omit areas that have not been cleared. The reason for the extensive manual editing is because naturally occurring events can affect vegetation in ways that appear similar to woody vegetation clearing in terms of the spectral and temporal responses observed by the satellite sensor (and used to calculate the clearing index). For example, damage by storms, fire and drought can all cause a reduction in canopy health or cover that can appear similar to a clearing event, and are often detected by the automated clearing index as possible clearing. This manual process makes use of any additional information available to aid decisions.

   During the manual editing step, each change area is allocated an appropriate class that indicates the replacement land cover or land use. The assignment or coding of these classes is primarily based on visual interpretation, with reference to ancillary data sources. In areas where there are many different forms of land use, it can be difficult to interpret the final replacement class and therefore this classification is indicative only. Broadly these classes distinguish clearing for pasture, agriculture, forestry, infrastructure and natural causes. There is a specific set of more detailed classes that varies for each jurisdiction.
4. Senior remote sensing scientists review the manual editing, so that mapped clearing has been visually checked and independently verified by another scientist. Further edits and quality control checks are undertaken to finalise the woody vegetation clearing mapping.
5. The mapping is compiled, and a statewide mosaic (see Volume 2B) of woody vegetation clearing is created.
6. Since the time period between annual images varies (depending on the availability of cloud-free satellite images), the mapping is converted to an annualised clearing rate to enable inter-annual comparisons.

Figure 14.4 Schematic representation of SLATS processing steps

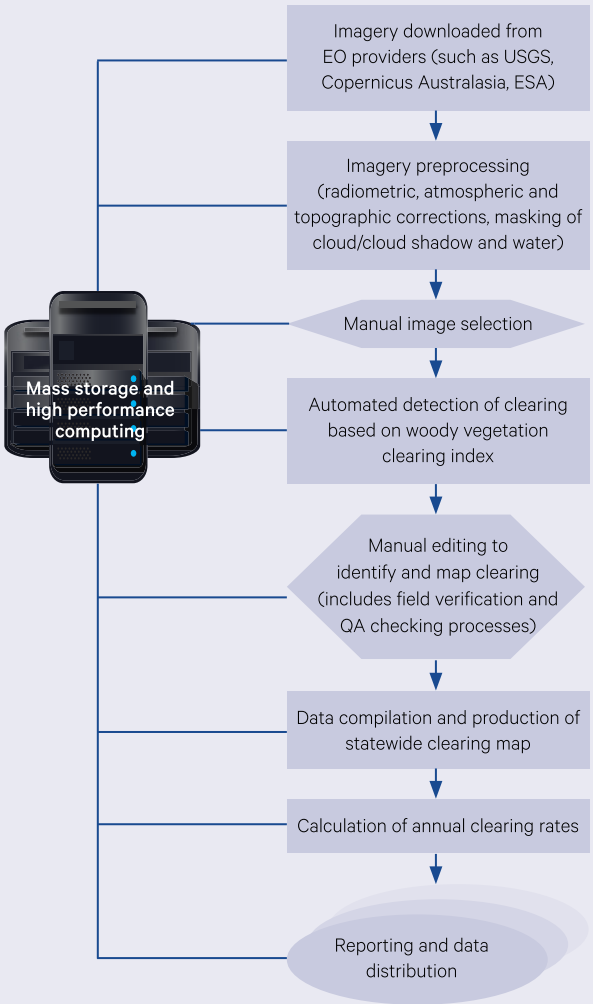


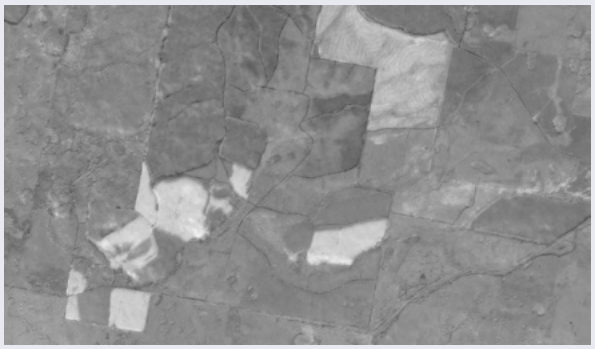
Figure 14.5 Example of SLATS change index and likelihood images

Two SPOT-5 images in the central tablelands of NSW were acquired on 18 April 2011 and 8 July 2012, before and after woody vegetation clearing. These images were used to create a change index image, in which the white areas are likely to indicate clearing. The change index image can be thresholded to create a change likelihood image. In this example, red highlights areas that are most likely to have been cleared, grey levels from dark grey to white are progressively less likely to have been cleared and yellow areas have not been cleared.

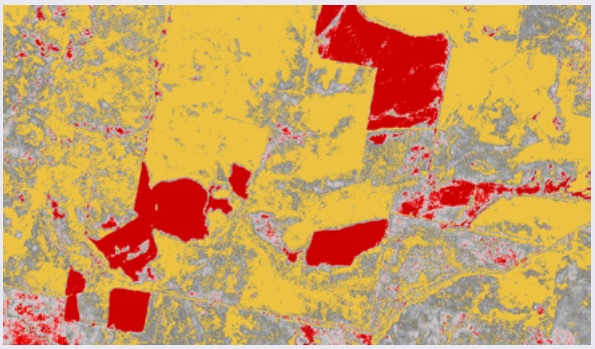
a. SPOT-5 image acquired before woody vegetation clearing

b. SPOT-5 image acquired after woody vegetation clearing

c. SLATS change index image



d. SLATS change likelihood image



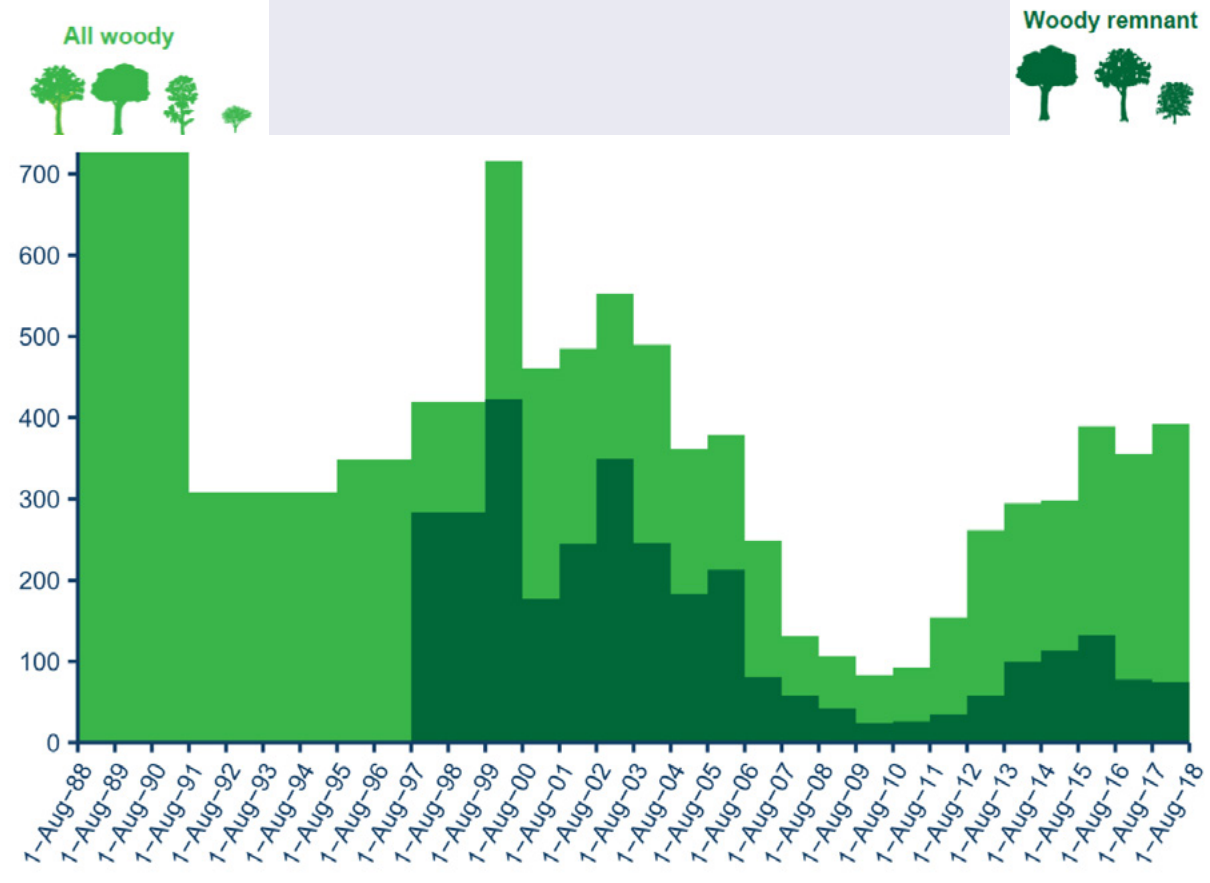
SLATS products and use

The SLATS information on vegetation change is generally published on an annual basis as a report with accompanying spreadsheets showing woody vegetation change by different regions, such as catchments, natural resource management regions, tenure etc. Additionally, the vegetation clearing data are analysed and compared with permits and development approvals to identify potential compliance issues that may require further investigation, and also to inform policy effectiveness and implementation of the legislation and regulatory frameworks. An example of a SLATS reporting outcome is shown in Figure 14.6 below.The SLATS change data is also used to update other forms of regulatory mapping, such as the Queensland Regional Ecosystems mapping or the NSW Native Vegetation Regulatory maps.

After annual reports are released the spatial vegetation change data is also released as open data and can be used by other practitioners and researchers for purposes such as biodiversity and land management, informing carbon and greenhouse gas emissions monitoring and sequestration programs, or for understanding the ecological impacts of vegetation change.

Figure 14.6 Example of SLATS reporting outcome

Annualised woody vegetation clearing rates in Queensland between 1988 and 2018 as reported by SLATS. In Queensland, the clearing rates are further analysed by a remnant and non-remnant vegetation status. This provides one indicator of the effectiveness of the vegetation management legislation in achieving its purpose, which includes conserving regional ecosystems.



Applying SLATS in near real time

As governments have become more accustomed to the use of satellite imagery to inform monitoring programs, there has been increased interest in its ability to be used for regular monitoring for proactive compliance and regulatory actions relating to the management of various natural resources, such as vegetation and water. This interest, and the ability to develop methods for regular monitoring, has also been facilitated by free and open data policies, such as those of the Landsat and Copernicus programmes, coupled with significant developments in high performance computing capabilities.

In relatively recent developments, both Queensland and NSW have adapted the SLATS model to build proactive compliance programs that detect recent clearing activity on weekly to fortnightly time scales. Queensland's system was originally developed using Landsat, but more recently both states have adopted Sentinel-2 satellite imagery, due to its higher spatial resolution and temporal frequency. The adaptation of an annual monitoring program to a more regular and frequent time step is not without its challenges:

- imagery and data management systems need be revised to track and manage the increased and regular data and workflows;
- the vegetation change algorithm can be less effective, or suffer significant commission errors, in wetter or cloudier periods, due to landscape green-up, inadequate cloud masking or inundation; and
- a more regular flow of information can overload available resources to address any potential compliance issues.

In order to manage for these challenges, the systems which have been developed in Queensland and NSW apply a number of imagery and spatial filters (such as cloud cover thresholds, regulatory mapping filters) and a manual checking component, which is still essential to confirm clearing events. However the manual checking component is aimed at identifying or detecting clearing events, rather than mapping them, as is the case in the annual SLATS monitoring programs.

These programs are being successfully implemented in both states, and are aimed at early intervention compliance action, thus raising public awareness with regards to vegetation management responsibilities, and reducing the amount of serious, significant clearing cases due to the ability to use timely intervention methods that deter further planned clearing from occurring. In Queensland, for example, it is estimated that this system has accounted for approximately 90–95% of all of the potential compliance actions that would previously have resulted from an analysis of the annual SLATS data, with the advantage that action has been taken within weeks to months (sometimes within days) of the clearing activity, rather than a year or more later as previously occurred.

Collaborative research program and complementary products

Ongoing maintenance of, and improvements to, the SLATS methods are undertaken collaboratively through the Joint Remote Sensing Research Program (JRSRP). This research includes:

- building and maintaining a software environment to enable efficient processing and data management tools for large archives of satellite imagery;
- pre-processing corrections including methods for image calibration, radiometric correction and change analysis;
- transitioning the SLATS methods to work with different medium resolution imagery sources such as SPOT-5 and Sentinel-2; and
- developing related systems such as the continuous early detection of vegetation clearing events using Sentinel-2 imagery.

This long-standing collaboration has enabled sharing of capacity and scientific processes and extended the benefits of the SLATS model for the Queensland and NSW governments, and their partners in other jurisdictions. These benefits include the progressive development of SLATS, and other complementary EO products, which have been built on its processes and systems. These products include:

- fractional cover (Guerschman *et al.*, 2015; Scarth *et al.*, 2010);
- foliage projective cover and woody extent data (Armston *et al.*, 2009; Kitchen *et al.*, 2010);
- fire scar and fire severity mapping (Goodwin and Collett, 2014);
- crop monitoring (Pringle *et al.*, 2018; Schmidt *et al.*, 2016); and
- biomass mapping (Scarth *et al.*, 2019).

Current and future developments for SLATS include improving woody extent mapping using computer vision approaches (Flood *et al.*, 2019) and developing regrowth and biomass monitoring approaches using Landsat and Sentinel time series, radar (such as ALOS-2) and spaceborne lidar.

Accounts are a structured, systematic way of organising data into information for a clearly defined decision-making purpose. They track flows and stores of value (stocks) over a set period of time. Accounts present comparable information in a systematic fashion, using standard definitions based on a sound conceptual framework. They encourage the development of comprehensive and consistent data and provide a platform for producing a range of accounting reports and analyses.
(BoM, 2013)

14.3 Modelling

People don't understand the Earth, but they want to, so they build a model, and then they have two things they don't understand.
(Gerard Roe in 'The Whale and the Supercomputer' by C. Wohlforth)

A complex range of dynamic environmental attributes constitute planet Earth (see Volume 1A—Section 3). The structure, composition, biology and energy of Earth continue to change in variable time frames. The processes that drive these changes are the focus of modelling activities. By identifying the driving forces and the mechanisms involved in environmental processes, we can start to understand their interrelationships and rate of change, then predict the likely result of variations in current conditions. For example, EO-based models for climate and weather are the backbone of current weather forecasting systems (see Volume 1A—Section 16). Similarly, modelling indicators of photosynthetic activity derived from EO datasets now underpin a wide range of environmental and agricultural activities (see Volume 3A).

EO-based modelling was introduced in Volume 1A—Section 1.4.3 and some types of models that are used with EO datasets were described in Volume 2A—Section 2. In the context of EO, forward models are used to compute the likely EO values from known environmental attributes, whereas inverse models use the EO attributes to compute the likely environmental attributes. Broad categories of models between EO values and environmental attributes include:

- empirical or analytical, which are based on simple relationships such as regression analysis;
- semi-empirical or semi-analytical, which use prior knowledge to determine the relationship; and
- physics-based, which rely on their known physical relationship(s).

Modelling of carbon, water and energy fluxes between the Earth's surface and the atmosphere is increasingly important for hydrological and climate studies. Such models can focus on particular spatial scales then, with appropriate prerequisites and constraints, be scaled up to increasingly larger scenarios (Nemani *et al.*, 2009; Pasetto *et al.*, 2018). The extensive spatial and temporal coverage of many EO datasets offer particular advantages for multi-scale modelling (see Section 1.4). In this section, we will consider one Australian example of modelling ecosystem dynamics based on EO datasets; further examples are detailed in Volume 3.

The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge.
(Stephen Hawking)

Excursus 14.4—eReefs

Source: Mark Baird, CSIRO
Further Information: *https://ereefs.org.au/ereefs*

Despite being one of the world's best managed reef systems, Australia's iconic Great Barrier Reef (GBR) has, with other reefs worldwide, suffered a decline in coral cover. In order to understand the factors that are influencing this decline, and to inform management strategies, a sophisticated information system, eReefs, has been developed to deliver past and near real time observations, modelling and analysis (Steven *et al.*, 2019). The eReefs information system is built upon an integrated system of data, catchment and marine models, visualisation, reporting and decision support tools that span the entire GBR area.

The eReefs marine models includes a range of process models that simulate hydrodynamic state (Herzfeld *et al.*, 2016), sediment transport (Margvelashvili *et al.*, 2018), water quality (Baird *et al.*, 2016), carbon chemistry (Mongin *et al.*, 2016) and basal ecology (Skerratt *et al*, 2019; Baird *et al.*, 2018) of the GBR lagoon and reef matrix. This collaborative project[18], started in 2012, is viewed as the first step in building a comprehensive coastal information system for Australia.

Simulations of the eReefs models enable the effectiveness of current management strategies to be assessed and provide valuable insights for regulatory authorities on a range of emergent and potential reef management challenges (Brodie *et al.*, 2017). The eReefs Platform integrates EO tools with near real time models to quantify significant coastal management issues, including:

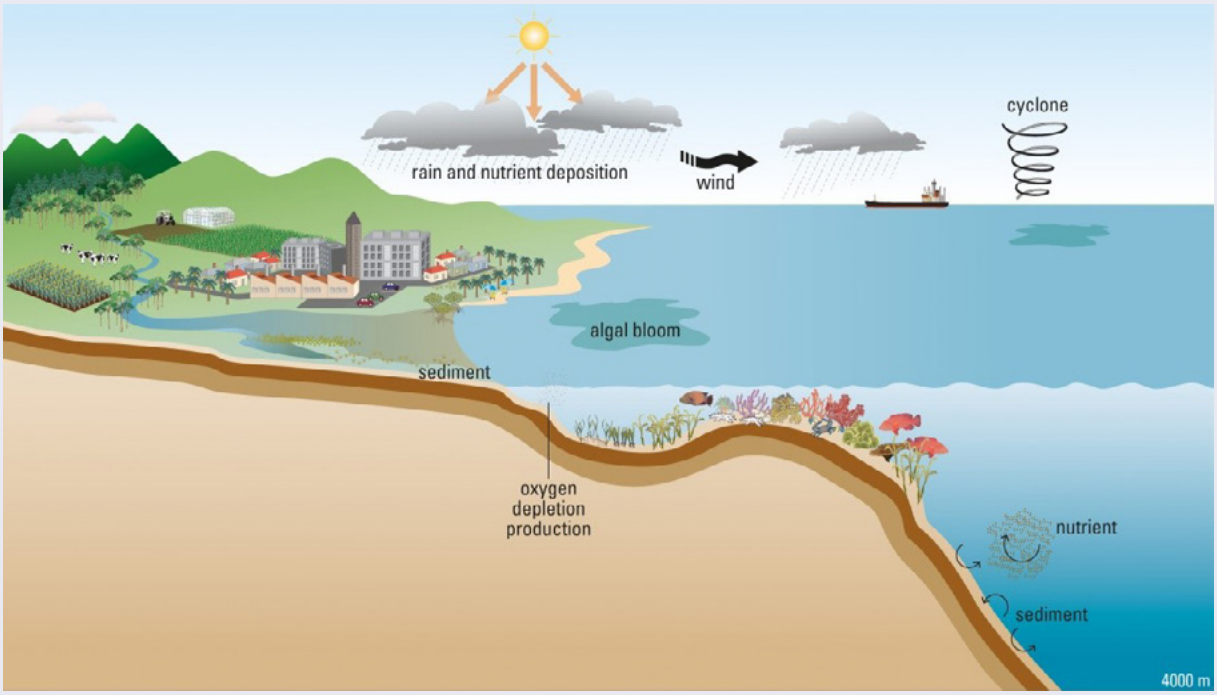
- regional, three-dimensional hydrodynamic and biogeochemical models to map, monitor and forecast coastal water quality (such as sediment and nutrient loads) and ecological responses for the GBR region in relation to agricultural and land management activities in adjoining catchments (see Figure 14.7). In conjunction with EO and meteorological observations, these nested models consider land use, catchment flows, circulation of mesoscale oceanographic processes, and sediment transport to simulate their impact on reef health in a range of timeframes; and
- operational hydrodynamic coastal ocean forecasting based on meteorological data to analyse, predict and archive data relating to currents, temperature, salinity, and sea level within the waters of the GBR region, plus river tracer concentrations from major rivers. This information has relevance to a number of marine management activities, including ship routing, port management, modelling plumes, monitoring spills, fishing, recreation, and search and rescue operations.

The temporal and spatial scales involved with processes and observations in the GBR are illustrated in Figure 14.8. EO datasets, from both satellite and aerial platforms, are relevant to a range of these scales, in both space and time, and are used in conjunction with a variety of data from other sensors to derive essential inputs to eReefs models. One strength of the biogeochemical modelling is the assimilation of remotely sensed ocean colour data. The eReefs model became the first published biogeochemical model to use EO reflectance (at multiple wavelengths) as the model variable for which the mismatch drives the assimilation procedure (Jones *et al.*, 2016). The use of EO reflectance required the development of a complex optical model in eReefs, and replaced the use of ocean colour algorithms in other biogeochemical data assimilation systems. In coastal waters where ocean colour algorithms have large errors, this was a breakthrough.

18 Collaborators include the Science Industry Endowment Fund, CSIRO, The Australian Institute of Marine Science, the Bureau of Meteorology, and the Great Barrier Reef Foundation, with support from BHP Billiton Mitsubishi Alliance and the Australian and Queensland governments, and observations obtained through IMOS.

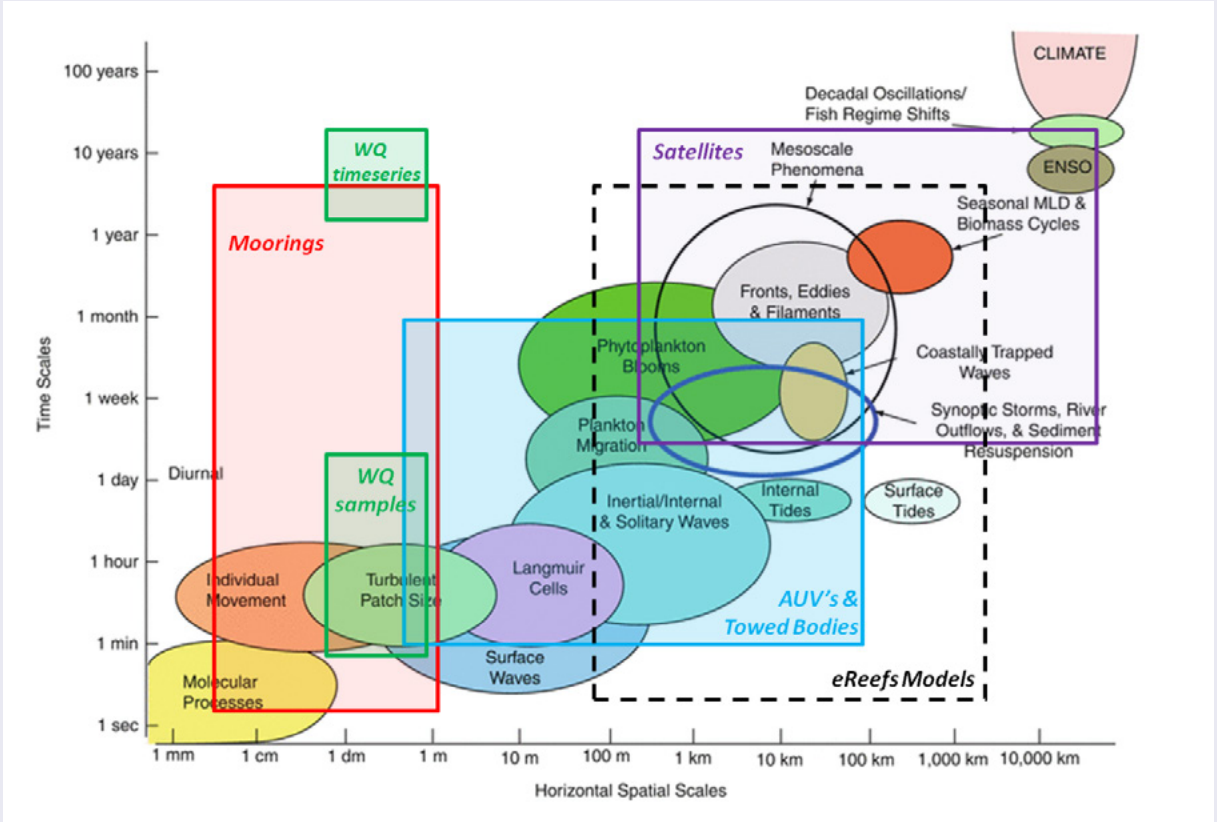
Figure 14.7 eReefs modelling framework

a. Conceptual model of GBR shelf ecosystem



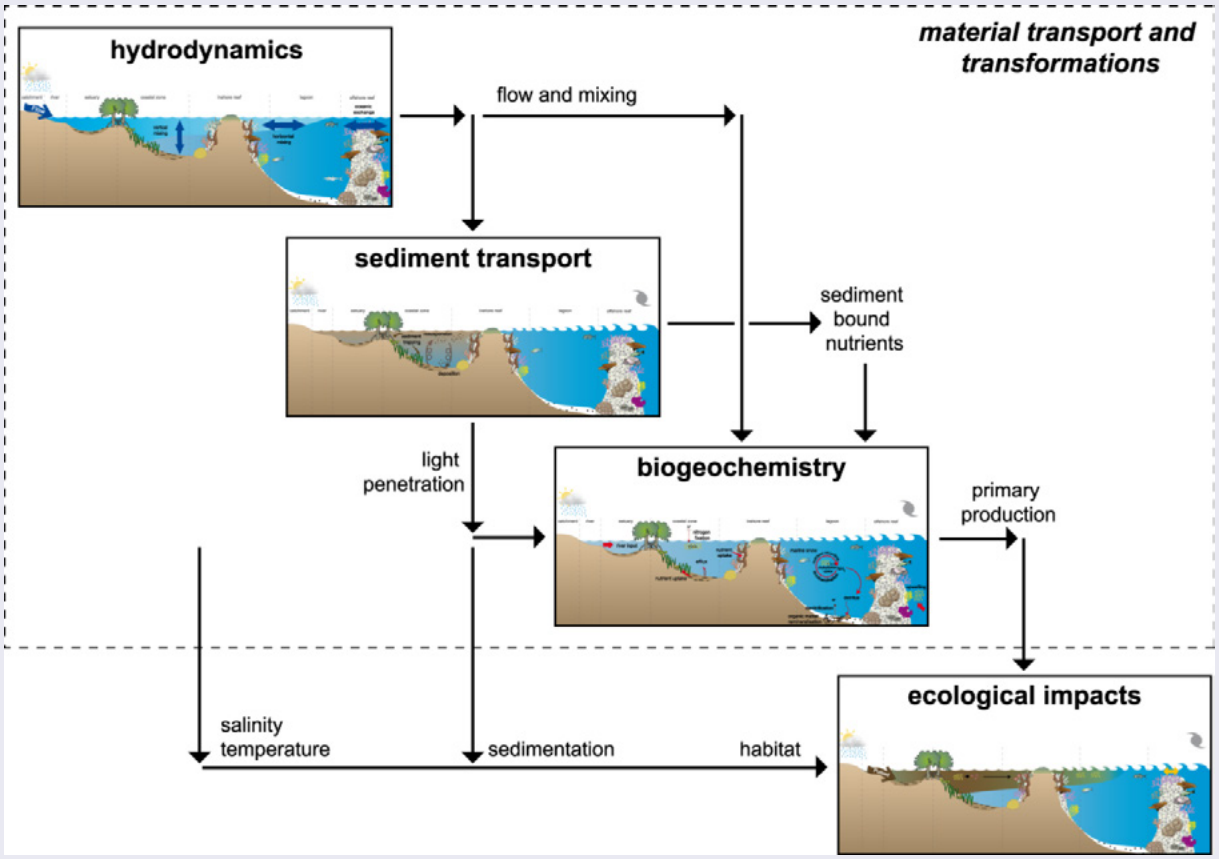
Source: Herzfeld *et al.* (2016) Figure 4.7.3 (page 155). Retrieved from *http://www.marine.csiro.au/cem/gbr4/eReefs_Marine_Modelling.pdf*

Figure 14.8 Temporal and spatial scales of GBR processes and observations



Source: Herzfeld *et al.* (2016) Figure 5.5.7 (page 354). Retrieved from *http://www.marine.csiro.au/cem/gbr4/eReefs_Marine_Modelling.pdf*

Figure 14.9 eReefs model interactions



Source: Webster *et al.* (2008) Figure 1.1-2. Retrieved from *https://ereefs.org.au/ereefs/platform/regional-hydrodynamic-sediment-and-wq-modeling*

The three major process-based marine models described above that underpin eReefs to determine ecological impacts (see Figure 14.9) are:

- hydrodynamic model—predict physical state of the GBR regional system based on meteorological data and models;
- sediment transport model—predict movement of suspended fine sediments and biogeochemical variables; and
- biogeochemical model—simulates water column and benthic production, water quality and nutrient cycling; with supporting models that account for water optics and wave action.

Model results have been calibrated and validated against independent observations, and model uncertainty estimated (Skerratt *et al.*, 2016). Model outputs can be accessed as raw data, viewed online as animated loops of key variables, or interactively explored in web applications. This integrated modelling approach enables early detection and precise location of potential problems, such as infestations of crown of thorns starfish, hotspots for coral bleaching, and decreasing oxygen levels.

Some of the reef management challenges that can be simulated and predicted by eReefs include the impact of:

- ocean acidity on coral calcification;
- catchment-based activities that reduce water quality and undermine reef health;
- combined effect of extreme light and temperature on coral bleaching;
- extreme weather events and floods on marine systems; and
- activities of water users, including shipping, fishing and dredging, on reefal ecosystems.

The eReefs project has now been providing near real time and past reconstructions of since the beginning of 2016, and is expected to be an ongoing tool to support researchers and managers in understanding the fragile GBR environment and enhancing reef health for future generations.

14.4 Further Information

Global Data Products:

United Nations Environment Programme (UNEP) Environmental Data Explorer: *http://geodata.grid.unep.ch*

LandScan Population Distribution Model: *https://landscan.ornl.gov*

Global EO Images and Image Products:

Global Vegetation Index: *https://www.ospo.noaa.gov/Products/land/gvi/*

Nighttime 'city lights' DMSP OLS: *https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html*

NASA Visible Earth: *http://visibleearth.nasa.gov*

NASA Ocean Colour: *https://oceandata.sci.gsfc.nasa.gov*

USGS Earth Explorer: *https://earthexplorer.usgs.gov*

USGS Global Visualisation Viewer (GLOVIS): *https://glovis.usgs.gov*

USGS EROS Archive for AVHRR Data and Products: *https://www.usgs.gov/centers/eros/science/usgs-eros-archive-advanced-very-high-resolution-radiometer-avhrr?qt-science_center_objects=0#qt-science_center_objects*

Environmental Accounting Framework:

System of Environmental-Economic Accounting (SEEA): *https://seea.un.org/*

Australian EO Information Products:

Terrestrial Ecosystem Research Network (TERN): *https://www.tern.org.au/*

TERN-ANU Landscape Data Visualiser: *https://maps.tern.org.au*

Australian Ocean Data Network (AODN): This interoperable online network of marine and climate data resources includes data that covers a wide range of parameters in different ocean environments collected from ocean-going ships, autonomous vehicles, moorings and other platforms: *https://portal.aodn.org.au*

OzWALD:

Australia's Environment: *http://wald.anu.science/data_services/data/continuous-and-comprehensive-national-environmental-reporting-australias-environment/*

Australia's Environmental Explorer: *www.ausenv.online*

State of Environment:

SOE 2016: *https://www.environment.gov.au/science/soe*

State-based Environmental Monitoring:

Statewide Landcover and Trees Study (SLATS , Queensland): *https://www.qld.gov.au/environment/land/management/mapping/statewide-monitoring/slats*

Ground cover monitoring (Queensland): *https://www.qld.gov.au/environment/land/management/mapping/statewide-monitoring/groundcover*

SLATS NSW: *https://www.environment.nsw.gov.au/topics/animals-and-plants/native-vegetation/reports-and-resources*

Habitat Hectares (Victoria): *https://www.environment.vic.gov.au/native-vegetation/native-vegetation/biodiversity-information-and-site-assessment*

Spatial data for Victoria: *https://www2.delwp.vic.gov.au/maps?_ga=2.6531511.420315026.1549698228-873342548.1549698228*

eReefs:

https://ereefs.org.au/ereefs

14.5 References

ABS (2018). *Australian Environmental-Economic Accounts, 2018*, webpage: *http://www.abs.gov.au/ausstats/abs@.nsf/mf/4655.0*

ABARES (2017). *Catchment Scale Land Use of Australia—Update September 2017*. ABARES, Canberra.

Armston, J.D., Denham, R.J., Danaher, T.J., Scarth, P.F., and Moffiet, T. (2009). Prediction and validation of foliage projective cover from Landsat-5TM and Landsat-7TM ETM+ imagery for Queensland, Australia. *Journal of Applied Remote Sensing*, 3(1), 033540-033567.

ANU (2019). Australian National University, *Australia's Environment in 2018* webpage: *http://wald.anu.science/australias-environment/*

Baird, M.E., Cherukuru, N., Jones, E., Margvelashvili, N., Mongin, M., Oubelkheir, K., Ralph, P.J., Rizwi, F., Robson, B.J., Schroeder, T., Skerratt, J., Steven, A.D.L., and Wild-Allen, K.A. (2016). Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with satellite data. *Environmental Modelling Software*, 78, 79–96.

Baird, M.E., Mongin, M., Rizwi, F., Bay, L.K., Cantin, N.E., Soja-Wozniak, M., and Skerratt, J. (2018). A mechanistic model of coral bleaching due to temperature-mediated light-driven reactive oxygen build-up in zooxanthellae. *Ecological Modelling*, 386, 20–37.

Brodie, J., Baird, M. Mongin, M., Skerratt, J., Robillot, C., and Waterhouse, J. (2017). Pollutant target setting for the Great Barrier Reef: Using the eReefs framework. In *MODSIM2017, 22nd International Congress on Modelling and Simulation*. (Eds: Syme, G., Hatton, T., MacDonald, D., Fulton, B. and Piantadosi, J.) Modelling and Simulation Society of Australia and New Zealand, December 2017, pp. 1913-1919. ISBN: 978-0-9872143-7-9. *https://www.mssanz.org.au/modsim2017/L22/brodie.pdf*

BoM (2013). *Guide to environmental accounting in Australia, Environmental Information Programme Publication Series no. 3,* Bureau of Meteorology, Canberra, Australia.

Copernicus Climate Change Service (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. In: *Copernicus Climate Change Service Climate Data Store (CDS)*, (Ed: ECMWF) *https://cds.climate.copernicus.eu/cdsapp#!/home*.

Danaher, T., Scarth, P., Armston, J., Collett, L., Kitchen, J., and Gillingham, S. (2011). Remote Sensing of Tree-Grass Systems: The Eastern Australian Woodlands. In *Ecosystem Function in Savannas: Measurement and Modelling at Landscape to Global Scales*. (Eds: Hill, M.J., and Hanan, N.P.) CRC Press, Boca Raton, pp. 175–194.

DEE (2016). *Environment Protection and Biodiversity Conservation Act 1999 (Parts 7, 9 and 10), Environment Protection (Sea Dumping) Act 1981, Fuel Quality Standards Act 2000 and Ozone Protection and Synthetic Greenhouse Gas Management Act 1989 Compliance Monitoring Program 2016–2017*. Department of Environment and Energy, Canberra.

DEE (2018). Department of Environment and Energy. *Environmental-economic accounting* webpage: *https://www.environment.gov.au/science/environmental-economic-accounting*

Flood, N., Danaher, T., Gill, T., and Gillingham, S. (2013). An Operational Scheme for Deriving Standardised Surface Reflectance from Landsat TM/ETM+ and SPOT HRG Imagery for Eastern Australia. *Remote Sensing*, 5, 83–109.

Frost, A.J., Ramchurn, A. and Smith, A.B. (2016). *The Bureau's Operational AWRA Landscape (AWRA-L) Model*. Bureau of Meteorology, Melbourne.

Gallant, J.C., Wilson, N., Dowling, T., Read, A., and Inskeep, C. (2011). *1 second SRTM Derived Digital Elevation Models User Guide*. Geoscience Australia, Canberra. *www.ga.gov.au/topographic-mapping/digital-elevation-data.html*

Geoscience Australia (2014). *Sentinel Hotspots Product Description Document V1.2*, code D2014-145826, Geocat Reference 70869.

Goodwin, N.R., and Collett, L.J. (2014). Development of an automated method for mapping fire history captured in Landsat TM and ETM+ time series across Queensland, Australia. *Remote Sensing of Environment*, 148, 206–221.

Guerschman, J.P., and Hill, M.J. (2018). Calibration and validation of the Australian fractional cover product for MODIS collection 6. *Remote Sensing Letters*, 9(7), 696–705.

Guerschman, J., Scarth, P., McVicar, T., Renzullo, L., Malthus, T., Stewart, J., Rickards, J., and Trevithick, R. (2015). Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. *Remote Sensing of Environment*, 161, 12–26.

Herzfeld, M., Andrewartha, J., Baird, M., Brinkman, R., Furnas, M., Gillibrand, P., Hener, M., Joehnk, K.D., Jones, E., McKinnon, D., Margevlashvili, N., Mongin, M., Oke, P., Rizwi, F., Robson, B., Seaton, S., Skerratt, J., Tonin, H. and Wild-Allen, K. (2016). *eReefs Marine Modelling: Final Report*. CSIRO, Hobart, Tasmania.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., and Xie, P. (2015). *NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG).* Algorithm theoretical basis document, version, 4: 30.

Jones, D., Wang, W. and Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, 58, 233–248.

Jones, E., Baird, M., Mongin, M., Parslow, J., Skerratt, J., Lovell, J., Margvelashvili, J.N., Matear, R., Wild-Allen, K., Robson, B., Rizwi, F., Oke, P., King, E., Schroeder, T., Steven, A., and J. Taylor, J. (2016). Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the Great Barrier Reef. *Biogeosciences*, 13, 6441–6469.

Kaiser, J.W. Heil, A., Andreae, M.O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M.G., Suttie, M., and van der Werf, G.R. (2012). Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9(1), 527–554.

Kitchen, J., Armston, J., Clark, A., Danaher, T., and Scarth, P. (2010). Operational use of annual Landsat-5 TM and Landsat-7 ETM+ image time-series for mapping wooded extent and foliage projective cover in north-eastern Australia. In *Proceedings of the 15th Australasian Remote Sensing and Photogrammetry Conference,* Alice Springs, Australia.

Margvelashvili, N.M., Andrewartha, J., Baird, M. Herzfeld, M.,Jones, E., Mongin, M., Rizwi, F., Robson, B., Skerratt, J., Wild-Allen, K., and Steven, A. (2018). Simulated fate of catchment-derived sediment on the Great Barrier Reef shelf. *Marine Pollution Bulletin*, 135, 954–962.

Mongin, M., Baird, M.E., Tilbrook, B., Matear, R.J., Lenton, A., Herzfeld, M., Wild-Allen, K.A., Skerratt, J. Margvelashvili, N., Robson, B.J., Duarte, C.M., Gustafsson, M.S.M., Ralph, P.J., Steven, A.D.L. (2016). The exposure of the Great Barrier Reef to ocean acidification. *Nature Communications*, 7, 10732.

Mueller, N., Lewis, A., Roberts, D., Rign, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., and Ip, A. (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341–352.

Myneni, R., Knyazikhin, Y., and Park, T. (2015). *MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006. 2015.* NASA EOSDIS Land Processes DAAC. *https://doi.org/10.5067/MODIS/MCD15A3H.006*.

Nemani, R, Hasimoto, H., Votava, P., Melton, F., Wang, W., Michaelis, A., Mutch, L., MIlesi, C., Hiatt, S., and White, M. (2009). Monitoring and forecasting ecosystem dynamics using the Terrestrial Observation and Prediction System (TOPS). *Remote Sensing of Environment*, 113, 1497–1509. *https://doi.org/10.1016/j.rse.2008.06.017*

Pasetto, D., Arenas-Castro, S., Bustamante, J., Casagrandi, R., Chrysoulakis, N., Cord, A.F., Dittrich, A., Domingo-Marimon, C., El Serafy, G., Karnieli, A., Kordelas, G.A., Manakos, I., Mari, L., Monteiro, A., Palazzi, E., Poursanidis, D., Rinaldo, A., Terzago, S., Ziemba, A., and Ziv, G. (2018). Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends. *Methods in Ecology and Evolution*, 2018:9, 1810–1821. *https://doi.org/10.1111/2041-210X.13018*

Pringle, M.J., Schmidt, M., and Tindall, D.R. (2018). Multi-decade, multi-sensor time-series modelling based on geostatistical concepts to predict broad groups of crops. *Remote Sensing of Environment*, 216, 183–200. *https://doi.org/10.1016/j.rse.2018.06.046*

Renzullo, L.J. Chappell, A., Raupach, T., Dyce, P., Li, M., and Shao, Q. (2011). An Assessment of Statistically Blended Satellite-Gauge Precipitation Data for Daily Rainfall Analysis in Australia, *34th International Symposium of Remote Sensing of Environment*. 10–15 April, Sydney.

Scarth, P., Gillingham, S., and Muir, J. (2008). Assimilation of spectral information and temporal history into a statewide woody cover change classification. *Proceedings of the 14th Australasian Remote Sensing and Photogrammetry Conference*, Darwin, Australia.

Scarth, P., Roder, A., and Schmidt, M. (2010). Tracking grazing pressure and climate interaction – the role of Landsat fractional cover in time series analysis. *Proceedings of the 15th Australasian Remote Sensing and Photogrammetry Conference,* Alice Springs, Australia, pp. 13–17.

Scarth, P., Armston, J., Lucas, R., and Bunting, P. (2019). A structural classification of Australian vegetation using ICESat/GLAS, ALOS PALSAR, and landsat sensor data. *Remote Sensing*, 11, 147. *https://doi.org/10.3390/rs11020147*

Schaaf, C., and Wang, Z., (2015). MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted RefDaily L3 Global-500m V006. NASA EOSDIS Land Processes DAAC. *https://doi.org/10.5067/MODIS/MCD43A4.006*

Schmidt, M., Pringle, M., Devadas, R., Denham, J., and Tindall, D. (2016). A framework for large-area mapping of past and present cropping activity using seasonal Landsat images and time series metrics. *Remote Sensing*, 8(9), 312. *https://doi.org/10.3390/rs8040312*.

Skerratt, J.H., Mongin, M., Wild-Allen, K.A., Baird, M.E., Robson, B.J., Schaffelke, B., Soja-Wozniak, M., Margvelashvili, N., Davies, C.H., Richardson, A.J., Steven, A.D.L. (2019). Simulated nutrient and plankton dynamics in the Great Barrier Reef (2011-2016). *Journal of Marine Systems*, 192, 51–74.

Steven., A.D.L., Baird, M.E., Brinkman, R., Car, N.J., Cox, S.J., Herzfeld, M., Hodge, J., Jones, E., King, E., Margvelashvili, N., Robillot, C., Robson, B., Schroeder, T., Skerratt, J., Tuteja, N., Wild-Allen, K., and Yu, J. (2019). eReefs: An operational information system for managing the Great Barrier Reef. *Journal of Operational Oceanography*, 12:sup2, S12–S28. *https://doi.org/10.1080/1755876X.2019.1650589*.

UN (2014). *System of Environmental-Economic Accounting 2012: Experimental Ecosystem Accounting,* European Commission, Organisation for Economic Co-Operation and Development, United Nations and World Bank. eISBN: 978-92-1-056285-0. *https://unstats.un.org/unsd/envaccounting/seearev/eea_final_en.pdf*

van Dijk, A.I.J.M. (2010). *AWRA Technical Report 3. Landscape Model (version 0.5) Technical Description*. WIRADA/CSIRO Water for a Healthy Country Flagship, Canberra.

van Dijk, A., Mount, R., Gibbons, P., Vardon, M. and Canadell, P. (2014). Environmental reporting and accounting in Australia: Progress, prospects and research priorities. *Science of The Total Environment*, 473–474, 338–349.

van Dijk, A., Summers, D., Renzullo, L., and Yebra, M. (2017). *OzWALD Gridded Environmental Data for Australia. v2*. CSIRO. Data Collection. *https://doi.org/10.4225/08/5a3c52da083b8*

van Dijk, A., and Rahman, J. (2019). Synthesising multiple observations into annual environmental condition reports: the OzWALD system and Australia's Environment Explorer. *MODSIM2019*, Canberra.

Vardon, M., Ovington, J., Juskevics, V., Purcell, J., and Eigenraam, M. (2018). Down Payments on National Environmental Accounts. Ch 14 in *Land Use in Australia: Past, Present and Future*. (Ed: Thackway, R.) ANU eVIEW, Canberra. ISBN: 9781921934421

Wan, Z., Hook, S. and Hulley, G. (2015). *MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006*. NASA EOSDIS Land Processes DAAC.

Webster, I.T., Atminskin, I., Bostock, H., Brooke, B., Douglas, G., Ford, P., Hancock, G., Herzfeld, M.J., Leeming, R., Lemckert, C., Margvelashvili, N., Noble, B., Oubelkheir, K., Radke, L., Revill, A., Robson, B.J., Ryan, D., Schacht, C., Smith, C., Smith, J., Vicente-Beckett, V. and Wild-Allen, K. (2006). *The Fitzroy Contaminants Project - A study of the nutrient and fine-sediment dynamics of the Fitzroy Estuary and Keppel Bay*, CRC for Coastal Zone, Estuary and Waterway Managment.

Earth Observation:
Data, Processing and Applications

Volume 2X: Processing—Appendices

This publication is available in PDF format at *eopublications.au*.
We welcome your comments regarding the readability and usefulness of this report.
To provide feedback, please email *feedback@eopublications.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
978-0-6482278-4-7

Copyright:
All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, save for content supplied by third parties, and logos. Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. The full licence terms are available from *https://creativecommons.org/licenses/by/4.0/legalcode*. A summary of the licence terms is available from *https://creativecommons.org/licenses/by/4.0/*.

Disclaimer:
While every effort has been made to ensure its accuracy, the CRCSI does not offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. The CRCSI and its employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

Recommended Citation for Volume 2X:
CRCSI (2018). *Earth Observation: Data, Processing and Applications. Volume 2X: Processing—Appendices*. (Eds. Jupp, D.L.B., and Harrison, B.A.) CRCSI, Melbourne.

Latest Revision:
October 2025

Background image on previous page: Landsat-5 image of Lake Eyre, South Australia, filling on 9 May, 2009, displayed using bands 2 4, 5 as RGB.
Source: Norman Mueller, Geoscience Australia

Acknowledgements

Production of this series of texts would not have been possible without the financial support of CSIRO, CRCSI, GA and Digital Concepts, input from members of the editorial panels and direction from members of the various advisory panels.

Volumes 1 and 2 of this series are based on text originally published in Harrison and Jupp (1989, 1990, 1992 and 1993)[1]. Many illustrations and some text from these publications have been reproduced with permission from CSIRO.

Other contributors are gratefully acknowledged:

- reviewers: Megan Lewis, Stuart Finn and Tony Sparks provided comments on early drafts of selected appendices;
- illustrations: Norman Mueller, Andreas Hüni, Linda Skoog (EagleView);
- Daniel Rawson (Accessible Publication & Template Design) for layout and formatting; and
- Carl Davies (CMDphotographics) for selected graphical illustrations.

We thank those owners of copyrighted illustrative material for permission to reproduce their work. Credits for individual illustrations are provided below the relevant graphic.

All volumes in this series are covered by the copyright provisions of CC BY 4.0 AU.

1 Harrison, B.A., and Jupp, D.L.B. (1989) *Introduction to Remotely Sensed Data: Part ONE of the microBRIAN Resource Manual.* CSIRO, Melbourne. 156pp.
Harrison, B.A., and Jupp, D.L.B. (1990) *Introduction to Image Processing: Part TWO of the microBRIAN Resource Manual.* CSIRO, Melbourne. 256pp.
Harrison, B.A., and Jupp, D.L.B. (1992) *Image Rectification and Registration: Part FOUR of the microBRIAN Resource Manual.* MPA, Melbourne.
Harrison, B.A., and Jupp, D.L.B. (1993) *Image Classification and Analysis: Part THREE of the microBRIAN Resource Manual.* MPA, Melbourne.

Table of Contents

Acknowledgements i

Appendix 1 Brightness Correction Models **1**

A1.1 Empirical Methods 1

A1.2 Measurement Model Approach 3

A1.2.1 Atmospheric correction and reflectance normalisation 3

A1.2.2 Analytic BRDF models 4

A1.3 Semi-empirical Methods 7

A1.3.1 Theoretical description 7

A1.3.2 Kernels 7

A1.3.3 Kernel-driven models 9

A1.3.4 Modified Walthall model 10

A1.3.5 Advantages of linear models 10

A1.4 The Issue of Scale 10

A1.5 References 11

Appendix 2 BRDF Models **15**

A2.1 Kernel Models 15

A2.2 Primary AMBRALS Kernel Models 16

A2.2.1 Ross-thin (vol) model 16

A2.2.2 Ross-thick (vol) model 16

A2.2.3 Roujean (geo) model 16

A2.2.4 Li-sparse (geo) model 16

A2.2.5 Li-dense (geo) model 16

A2.2.6 Cox-Munk (geo) model 17

A2.2.7 Walthall (vol) model 17

A2.2.8 MODIS products 17

A2.3 Extensions and Alternatives 17

A2.3.1 Reciprocal Li models 17

A2.3.2 RPV model 18

A2.3.3 MRPV model 18

A2.3.4 Staylor and Suttles model 18

A2.3.5 Shibayama and Weigand model 18

A2.3.6 Dymond and Qi model 18

A2.3.7 Chen modification to Roujean model 18

A2.3.8 Liang modification to Walthall model 19

A2.3.9 Pickup and Chewings model 19

A2.4 Slightly More Complex Models: Not in 'Kernel' Form 19

A2.4.1 Otterman model 19

A2.4.2 Verstraete, Pinty and Dickinson (VPD) model 20

A2.4.3 Strahler and Jupp simple model 20

A2.4.4 Li's 'top easy seen' modification 21
A2.4.5 Hapke BRDF model for soils 21
A2.4.6 Ross' simplified vegetation canopy formula 22
A2.4.7 Minnaert's original model 22
A2.5 References 23

Appendix 3 Colour Systems and Sequences **25**
A3.1 Colour Measurement 25
A3.2 Colour Systems 26
A3.2.1 RGB and HSI 26
A3.2.2 XYZ and UVW 27
A3.3 Colour Coordinates 28
A3.4 References 28

Appendix 4 Dithering **29**
A4.1 Dither Matrices 29
A4.2 Designed Order 30
A4.3 References 32

Appendix 5 Filters **33**
A5.1 Averaging Filters 33
A5.2 Directional Filters 35
A5.2.1 Exposure 36
A5.2.2 Relief shading 36
A5.2.3 Curvature 37
A5.3 Geophysical Filters 38
A5.4 References 38

Appendix 6 Matrix Algebra **39**
A6.1 Geometric Effect of Matrix Operations 39
A6.2 Matrix Arithmetic and Properties 43
A6.2.1 Trace 43
A6.2.2 Matrix equality 43
A6.2.3 Matrix addition 43
A6.2.4 Scalar multiplication 43
A6.2.5 Matrix multiplication 43
A6.2.6 Transpose of a matrix 43
A6.2.7 Additive inverse of a matrix 43
A6.2.8 Multiplicative inverse of a matrix 43
A6.2.9 Determinant of a matrix 44
A6.2.10 Characteristic equation 45
A6.2.11 Eigenvalues of a matrix 45
A6.2.12 Eigenvectors of a Matrix 46
A6.3 Derivation of the Principal Components Transformation 46
A6.4 References 48

Appendix 7 Predictive Error and Variance **49**
A7.1 Predictive Error for Least Squares 49
A7.2 Generalised Cross-Validation (GCV) 51
A7.3 Singular Value Decomposition 51
A7.4 Predictive Variance 52
A7.5 References 52

Appendix 8 'Sieving' Outliers **53**
A8.1 Basic Assumptions and Framework 53
A8.2 Fitting Transformation Models 53
A8.3 The SIEVER Model 54
A8.4 Solving for Underlying Model and Errors 55
A8.5 Analysing the SIEVER Solution 56
A8.5.1 Basic 'PCA' summary 56
A8.5.2 Row or feature-based error analysis 57
A8.6 Using the Q-Q Plot 57
A8.6.1 Quantile-Quantile plot 57
A8.6.2 Gamma probability plotting 58
A8.6.3 Interpreting the Q-Q plot 58
A8.7 Data Plots to Indicate Outliers 59
A8.7.1 Listing and plotting the Matrix *US* *59*
A8.7.2 Listing and plotting the Error Matrix 59
A8.8 Using SIEVER to Locate Outliers 59
A8.9 References 60

List of Figures

Figure A2.1 Geometry of remote observer 15

Figure A3.1 RGB colour coordinates 26

Figure A3.2 HSI coordinates 27

Figure A3.3 CIE rg chromaticity space 27

Figure A3.4 Perceptual colour differences 28

Figure A4.1 Dithering example 29

Figure A4.2 Dithering sequence 29

Figure A4.3 Example dither matrices 30

Figure A4.4 Formation of a balanced dither matrix 31

Figure A4.5 Colour representation using dithering techniques with a 2×2 dither. 31

Figure A5.1 Averaging filters 34

Figure A5.2 Derivative filters 35

Figure A5.3 Directional, edge-detecting filters 35

Figure A5.4 Directional filters 36

Figure A5.5 Kirsch directional filters 36

Figure A6.1 Effect of matrix operation on data geometry 40

Figure A6.2 Reflecting axes 40

Figure A6.3 Reflecting axes about the line Y=X 41

Figure A6.4 Rescaling axes 41

Figure A6.5 Rotating axes 42

Figure A6.6 Skewing axes 42

Figure A6.7 Effect of matrix operation on unit area 44

Figure A6.8 Effect of singular matrix operation 45

Figure A6.9 Example image channels 46

Figure A8.1 Q-Q plot 58

List of Tables

Table A1.1 Semi-empirical weight formulas 9

Table A6.1 Principal component percentage variation 47

List of Excurses

Excursus A1.1—Geometric Optical Models 5

Appendix 1
Brightness Correction Models

The need for image calibration is introduced in Volume 2A—Section 3 and is particularly important for multi-date or multi-sensor image datasets (see Volume 2D). The general approach to image brightness variation is to 'correct' it in some way. Correction means changing a pixel spectrum so that it is what would have been measured if the Sun and sensor geometry were a fixed standard pair rather than a varying pair of angles over the image extent. As discussed later, this action may be valid as an average over a spatial area but can be a difficult concept at the scale of a single tree crown or soil clump.

We will outline three main approaches to scene brightness modelling and 'correction', namely:

- empirical methods (see Appendix A1.1);
- measurement modelling methods (see Appendix A1.2); and
- semi-empirical methods (see Appendix A1.3)[2].

A1.1 Empirical Methods

Empirical methods involve fitting the variation in general brightness ascribed to the atmosphere and Bidirectional Reflectance Distribution Function (BRDF) effects by a parametric function (using least squares or some other criterion) and then normalising by some combination of additive and multiplicative image transformation. The simplest methods involve fitting polynomials to average variations over scan lines or frames and subtracting or dividing out the general level to 'flatten' the data. These methods are well known and have been used with single frame photographs or scanner runs prior to other image processing operations.

For mosaics, brightness normalisation has a strict analogy with geometric mosaicking where images are geometrically matched to the ground by Ground Control Points (GCP) and to other images by tie points (see Volume 2B). At tie points, in geometric mosaicking, the constraint is that the same feature must be located at the same point in the geometric frame model. In the same way, in brightness normalisation, there may be targets in some frames of known reflectance. These are sometimes called 'invariant targets'. When their reflectance is only assumed from previous work they are sometimes called 'pseudo-invariant' targets. These are analogous to GCP. In the overlap between frames, the analogy with tie points is patches or features that are assumed to have the same corrected spectra. For a given functional form of brightness model, the estimation reduces to a linear 'bundle adjustment'.

The greatest problems with the simple empirical approach are image variance and heterogeneity. It is often extremely difficult to determine which components of the variation are atmospheric or BRDF brightening and which are differences in surface type. The high image spatial variance that is most obvious in high resolution images also makes fitting functions very difficult. Another problem is to decide which effects are additive and which are multiplicative. As long as the variation is not too great, the additive approximation is not too bad but when raw, uncalibrated data are used, the interaction between additive and multiplicative effects can be a problem.

2 Most text and equations in this appendix have been sourced from Jupp and Strahler (2000).

Background image: This 180° panorama near William Creek, SA, shows the hotspot effect around the focal point of the camera, caused by the predominant backscatter of vegetation. This image is shown here with vertical exaggeration (see also Volume 1B—Figure 3.22). **Source:** Andreas Hüni, University of Zurich (taken on iPhone6s)

For example, calibration and atmospheric correction are linear effects but the atmospheric terms are angle dependent and vary with incident (μ_i) and view (μ_v) directions. That is, even if the surface were Lambertian (with a flat BRDF) the digital data recorded by a scanner in waveband *j* (dn_j) will be related to the reflectance (ρ_j) by:

$$L_j = \alpha_j \, dn_j + \beta_j$$

$$L_j = \frac{1}{\pi} E_T^*(\mu_i) \, T(\mu_v) \rho_j + L_p(\mu_i, \mu_v)$$

hence

$$dn_j = \gamma_j(\mu_i, \mu_v) \, \rho_j + \delta_j(\mu_i, \mu_v)$$

where

L_j is the radiance recorded in waveband *j*;
μ_i is the Sun or 'incident' radiation direction;
E_T^* is the total irradiance at the target; and
L_p is the radiation scattered into the sensor from the atmosphere.

The brightness variation described here (which has no surface BRDF effect in that ρ_j is not dependent on Sun or view angles) obviously has both additive and multiplicative terms, which are complex functions of the incident and view angles. When the surface target itself has a strong BRDF effect the variability will compound.

In the face of this, the empirical methods are generally constrained to be of low order functional forms that still have the structure of the variation being fitted. In this way, it is hoped that variance and image heterogeneity are 'orthogonal' to the function being fitted. In addition, the fitting is best done in the physical framework of the brightness variation, that is, in terms of the phase angles between Sun and sensor view angle and between sensor view angle and the specular vector. There have been many functions of this type developed such as Walthall, Hapke, Roujean and others. These will be addressed below in Appendix A1.3.

For the purely empirical approach and for small frames such as occur with aerial photography and video data (or even single scanner runs) the main problem is scene heterogeneity. There will be scene variations that occur at spatial scales near to or greater than the image and it is very difficult to separate these effects from atmospheric and broad BRDF effects. To illustrate this you can consider an image to be roughly approximated by:

$$I = I_H + I_L + I_A$$

where

I_H is the high spatial frequency component;
I_L is the low frequency component; and
I_A is the angular brightness variation.

In many images, over the extent of a single frame, the low frequency and angular components are not in any way orthogonal. Hence it is near impossible to estimate the angular function from the single frame data. Moving to models based on view geometry and semi-empirical functions can help (Pickup *et al.*, 1995) but essentially some way to bring in the wider spatial context is needed.

Among empirical methods are also 'referencing' methods. In the case of image digitising, a reference approach is to scan an image and scan a reference (white Lambertian) standard sheet with the same illumination. The image is then 'normalised' by subtracting or dividing the reference from the image to remove lighting based view angle effects. In reconnaissance image data this is not possible unless a large reference target were to exist and scene BRDF effects would still remain. Hick and Ong (Ong *et al.*, 1995) have used a referencing method to extract brightness variations from video frames by referencing against the background of a Landsat TM image. It may be possible to fit empirical functions to the residual between the individual frame and the reference image. In this way, the problem of heterogeneity and correlated spatial frequencies described above could be reduced significantly.

It is worth mentioning here that as soon as the empirical methods extend from simple functional forms to ones involving image and Sun geometry, the processing methods start to become more costly. The cost is in extra work registering images and/or in extra costs associated with collecting attitude and position information (such as INS and GPS) for the platform. If the end result is worth this extra cost then it will be worth some extra modelling as discussed below.

A1.2 Measurement Model Approach

A1.2.1 Atmospheric correction and reflectance normalisation

It would clearly be an advantage to model the atmospheric and BRDF effects by physical models and estimate the corrected image in terms of a normalised reflectance factor. The added value would be the link this gives to physical models of the relationship between Earth parameters of interest and the measured reflectance. Obviously, a reflectance is closer to a material property and therefore resolves the problem of monitoring.

If reference targets in the image have measured reflectance (that is, the invariant or pseudo-invariant targets) it is possible to get close to reflectance from the empirical approach. However, it will never be clear how much residual atmosphere and BDRF effect remains. Obviously, most natural targets that are chosen as 'invariant' will have some BRDF that may not be known and that effect will be folded into the final image data.

Assuming the data are accurately calibrated to radiances, there are different ways to describe the process of correction. One is as follows:

An equation relating the recorded radiance sensed at altitude *h* above the target to the target reflectance factor is:

$$L_t(\mu_v,\mu_s,h,\lambda) = \tfrac{1}{\pi}E_T^*(\lambda)\,t(\mu_v,h,\lambda)(\rho_t+\rho_{env}) + L_p(\mu_v,\mu_s,h,\lambda)+L_g(\mu_v,\mu_s,h,\lambda)$$

where

$L_t(\mu_v,\mu_s,h,\lambda)$ is the radiance observed by the instrument from altitude *h*, with look (or view) direction μ_v and Sun direction μ_s at wavelength λ;

$t(\mu_v,h,\lambda)$ is the beam transmittance through the layer between the surface and altitude *h* in direction μ_v;

$T(\mu_v,h,\lambda)$ is the diffuse transmittance for a layer of thickness *h* and for initial beam direction μ_v;

$L_p(\mu_v,\mu_s,h,\lambda)$ is the path radiance of light which did not interact with the surface; and

$L_g(\mu_v,\mu_s,h,\lambda)$ is the glint term that is most significantly present over water-covered targets and is sometimes present over land targets.

E_T^* is the effective irradiance at the target, or

$$E_T^*(\lambda) = \frac{E_T(\lambda)}{1-s\,\rho^*}$$

$E_T(\lambda)$ is the irradiance at the target for a 'black' earth;

s is the sky hemispherical albedo;

ρ^* is the background earth albedo;

ρ_t is the target directional reflectance factor; and

ρ_{env} is the environmental reflectance due to the background albedo ρ^* or

$$\rho_{env} = \rho^*\left(\frac{T(\mu_v,h,\lambda)}{t(\mu_v,h,\lambda)}-1\right)$$

If the atmosphere is characterised then it is possible to retrieve the directional reflectance factor (ρ_t) for each pixel. This term needs to be defined carefully as there are many different types of 'reflectance' used.

The directional reflectance factor (ρ_t) as used here is defined as:

$$\rho_t(\mu_s,\mu_v,f_d,\lambda) = \frac{\pi L(\mu_v,\lambda)}{E_T^*(\mu_s,f_d,\lambda)}$$

in which the irradiance (E_T^*) is the sum of diffuse and direct terms and the fraction of diffuse (f_d) is included as a parameter. The assumption that the irradiance can be characterised in this context by the Sun position and the fraction of diffuse radiation is one that needs evaluation. The value of using this form of reflectance is that it corresponds to what is measured in the field using an irradiance radiometer or a reference standard.

The physical approach depends on two steps. The first is to determine this reflectance factor for a surface by atmospheric correction. In atmospheric correction, the atmospheric terms are modelled and measured from image and ancillary data. The reflectance factor in an image may be obtained iteratively if the atmospheric turbidity makes the adjacency and other atmosphere/surface interactions significant. The second step is to normalise the reflectance factor in some way to account for its BRDF variation.

In order to go from the reflectance factor to such a corrected value, however, we effectively need to assume that:

$$\rho_t(\mu_s,\mu_v,f_d,\lambda) = \rho_t^0(\lambda)\,\iota(\mu_s,\mu_v,f_d,\lambda)$$

where $\iota(\mu_s,\mu_v,f_d,\lambda)$ is the angular variation function that is assumed to characterise the land surface type and be normalised to 1.0 at a reference pair of Sun and view angles and standard atmosphere. Then 'corrected' data are reachable as:

$$\rho_t^0(\lambda) = \frac{\rho_t(\mu_s,\mu_v,f_d,\lambda)}{\iota(\mu_s,\mu_v,f_d,\lambda)}$$

The value of getting to this point is that both $\rho_t^0(\lambda)$ and $\iota(\mu_s, \mu_v, f_d, \lambda)$ can be interfaced with radiative transfer models to obtain parameters for the Earth's surface by inversion. The inversion may be simple (such as end member methods) or sophisticated (such as complete nonlinear modelling).

This physical approach is obviously highly sophisticated but demands a level of data quality and time in value adding that sometimes cannot economically be committed. However, if the end products are of high information value and valued by the client it is clearly an approach that could be pursued.

A1.2.2 Analytic BRDF models

Assuming the atmospheric correction can be done (and the surface interaction needs to be flagged here as the coupling means the atmospheric correction and surface BRDF estimation are not independent) the surface BRDF can sometimes be defined by an analytic model.

The broad brightness variations in all EO images can be considered as the composite of the spectral properties of component materials combined with the structural effects that result in angular variations (see Volume 1B). Angular effects can be viewed in terms of three factors:

- volume effect—which depends on surface structure;
- occlusion (or hotspot) effect—shadow cast by the Sun occludes parts of the surface; and
- specular (or glint) effect—when surface properties strongly reflect illumination, such as water glint (see Volume 2A—Section 3.2.1).

Among the many models for the Volume effect are the Suits (1972) and SAIL (Scattering by Arbitrary Inclined Leaves; Verhoef, 1984, 1985) models as well as more advanced ones such as the hotspot-based model described (as examples) in Qin and Jupp (1993) and Jupp and Strahler (1991). The field of Geo-optical modelling (GO Models) is now very extensive but not as well served in its entirety yet by any major books or fully inclusive reviews.

The models used for the Volume effect range from the plane parallel layer models (Suits, 1972; Verhoef, 1984, 1985) to sophisticated three-dimensional ones such as the hotspot-based Geometric Optical-Radiative Transfer (GORT) models (Li and Strahler, 1995). Myneni and Ross (1990) provides a formative early overview with many of the original aspirations well presented. Liang *et al.* (2000) reviews the rapid progress during the following decade (see also Chen *et al.*, 2000, and Qin and Liang, 2000). More recently, the Radiation Transfer Model Intercomparison (RAMI) experiment has provided a major influence for stabilizing model development and ensuring maturity (Pinty *et al.*, 2004; Widlowski *et al.*, 2007).

The hotspot effect is a geometric or occlusion effect. From among the many papers that describe the hotspot effect, Excursus A1.1 provides an overview of one implementation of Geometric-Optical (GO) models (extracted from Jupp and Walker, 1996). The GO model approach has been extended into models that combine geometric and radiative transfer components into effective and complete models for canopies. Other GO models include:

- Nilson and Kuusk models (Nilson and Peterson, 1991; Kuusk and Nilson, 2000; and Kuusk, 2001);
- Chen and White models (Chen and Leblanc, 1997, 2001; White *et al.*, 2001, 2002; Chen *et al.*, 2003, 2005);
- Lacaze and Roujean models (Roujean, 2000); Lacaze and Roujean, 2001);
- Li, Strahler and Ni-Meister (GORT) models (Li and Strahler, 1995; Ni *et al.*, 1997, 1999; Yang and Friedl (2003); and
- Huemmrick and Goward (1997), Huemmrich (2001), Hanan (2001) and Haverd *et al.* (2012).

Among these, the GORT model (Li *et al.*, 1995; Ni *et al.*, 1997) has had extensive application, including lidar data inversion, as described (for example) in Ni-Meister *et al.* (2010). However, despite the enthusiasm of the early authors, more pragmatic approaches have been taken since that time based on semi-empirical BRDF models, as outlined below, and on mixture models for vegetation canopies.

Excursus A1.1—Geometric Optical Models

Source: Jupp and Walker (1996)

A simple model for the remote sensing of a canopy is the Geometric Optical (GO) model. The simple GO (or hotspot) model for scenes which describe open forest or woodland areas is based on the one described in Jupp *et al.* (1986), Strahler and Jupp (1991a, 1991b) and Li and Strahler (1992). In this model, there are four kinds of ground cover 'visible' from a given direction. These are referred to as scene components and comprise:

- sunlit canopy (*sc*);
- shaded canopy (*shc*);
- sunlit background (*sb*); and
- shaded background (*shb*).

Each component is assumed to have a characteristic radiance and the radiance of a pixel is modelled as the area weighted combination (or linear mixture) of the characteristic component radiances. That is, the observed radiance of a single pixel (r_s) is modelled as:

$$r_S = k_{sc}R_{sc} + k_{shc}R_{shc} + k_{sb}R_{sb} + k_{shb}R_{shb}$$

where

subscripts *sc*, *shc*, *sb*, and *shb* indicate the radiances of the four components as named above;
R_j represents the (mean) radiance of component *j*; and
k indicates the sensed proportion of each component within the pixel from the given view direction.

Assuming the view and Sun directions are constant, the mean radiance over the scene (R_s) can be written as:

$$R_s = K_{sc}R_{sc} + K_{shc}R_{shc} + K_{sb}R_{sb} + K_{shb}R_{shb}$$

where

capital K_j represents the mean or expected value of the varying proportions k_j over the scene for *j* as the components *sc*, *shc*, *sb* or *shb*; and
mean value (R_s), as a function of Sun and observer position, defines the BRDF of the scene.

In order for the scene BRDF model to be computed, a description of the size and shapes of the objects, their density and how they are distributed over the background is needed and the geometrical relationships between the objects and the expected values of the four components must be established. Jupp *et al.* (1986), Strahler and Jupp (1991a; 1991b) and Li and Strahler (1992) describe such a model for spheroidal crown (not necessarily opaque) volumes, which is valid for any view or illumination angles using the 'Boolean' model of Serra (1982). In the Boolean model, the object 'centres' are assumed to be randomly distributed in a 'Poisson' distribution. By defining the geometry and the distributions, expressions for K_j may be derived. Strahler and Jupp (1991a, 1991b) use a simple model for spheroids, which is adequate for moderate Sun and view zenith angles, while Li and Strahler (1992) provide some more general alternative models for resolving the K_j. These basic scene BRDF models are quite simple and are easily implemented in various forms such as mathematical packages or spreadsheets.

In the woodlands and open forest areas typical of the area of Australia where the model studies have been made, the pixel to pixel behaviour of the image is conveniently (if not as accurately) described by a simpler form of the model in which the shaded background, sunlit (but still relatively dark) tree and shaded tree components are combined into one so that:

$$r_s = R_X + k_{sb}(R_{sb} - R_X)$$

where *X* is a composite component combining sunlit and shaded tree and shaded background and R_X is computed as:

$$R_X = \frac{K_{sc}R_{sc} + K_{shc}R_{shc} + K_{shb}R_{shb}}{1 - K_{sb}}$$

The simpler model has the advantage for this discussion that it shows clearly how, in many woodlands, the image pixel to pixel variation is driven primarily by the variation in the proportion of sunlit background that is visible in the pixels and the contrast between this sunlit background and the other components. It also provides a simple estimate for k_{sb} from images where R_{sb} and R_X are known for an appropriate image channel, or channel combination, as:

$$k_{sb} = \frac{r_s - R_X}{R_{sb} - R_X}$$

For such a model, the mean radiance (that is, BRDF) over all pixels in a patch with the same basic underlying type of cover and structure is therefore:

$$E(r_s) = R_s = R_X + (R_{sb} - R_X)K_{sb}$$

where K_{sb} is the mean value of k_{sb}, or the expected proportion of visible sunlit background for the particular Sun and view positions. This simple model has been found to be adequate for describing the data obtained by a Daedalus scanner over woodlands.

Linear end-member analysis is similar to the estimation of components described above. It has been the subject of useful research and application in Australia (Pech *et al*., 1986, Pickup and Foran, 1987) and at regional scales where all pixels are mixtures of land covers of interest (Cross *et al*., 1991). End-member analysis assumes each pixel to be a composition, or mixing, of a few base components or 'end-members'. The pixel signature is assumed to be a linear sum of reflectances from each of *n* end-members weighted in proportion to its cover (k_j) in the pixel:

$$r_s = \sum_{j=1}^{n} k_j R_j$$

where $k_j \geq 0$ for $j=1,n$

$$\sum_{j=1}^{n} k_j = 1$$

End-member analysis seeks to invert this mixing by deriving the proportions (k_j) of each component in the pixel signature. This can feasibly be derived from EO data provided that if there are *n* components (trees, shrubs, grass etc.) then there are at least (*n*-1) channels of data that separate the end-members spectrally. The key assumptions built into the end-member method are that:

- end-members (pure examples of total cover by trees, shrubs, grass and background) are spectrally consistent between sites; and
- reflectance values for end members (R_j) are available from EO data or can be accurately derived by other means (such as field spectral measurements).

There has been considerable work aimed at deriving end-members from the data (a form of Principal Components Analysis; see Boardman, 1990; Appendix A6.3 below; and Volume 2C) and employing high spectral resolution data to effect separation of more than a few components (Adams *et al*., 1989). However, with a lack of available high resolution spectral data, this linear approach suffers from several significant limitations to its applicability in Australia:

- available broad band signatures of the tree and shrub crowns over much of Australia whilst different, are not markedly spectrally distinct;
- even if spectrally distinct crowns did exist for the available bands, their distinction is confounded by the effects of shadowing within crowns and cast shadow on the background (with bigger plants shading smaller plants). This makes the signature of the end-members difficult to estimate as the signature depends on the proportions of crowns and shadows present and variations in Sun and viewing angles; and
- relatively low covers of trees and shrubs, together with shadowing, introduce such high spectral variance into the data relative to the spectral contrasts between end-members that the numerical methods used in the end-member analysis become highly unstable.

Shadow effects obviously depend primarily on the Sun angle. Although the crown cover is the same, lower Sun angles clearly decrease image brightness. Differences due to shadowing can be taken into account in end-member analysis, provided the end-member values are recalculated for each temporal image and one or more components labelled 'shade' are added to the list. However, its successful application still depends on an assumption of linear scaling along cover gradients due to Sun positional and sensor view angle changes. These assumptions in practice are erroneous in structured vegetation (for example, vegetation with discontinuous cover of trees or shrubs), and this limits the application of such methods to general synoptic estimates of change in cover.

In structured vegetation, that is where the cover of trees and/or shrubs is discontinuous, it is appropriate to model vegetation cover directly as an assemblage of various sizes and shapes of three-dimensional objects (trees, shrubs, grass tussocks, herbs, etc.) scattered on a background that may be uniform or heterogeneous (Li and Strahler, 1985; Jupp *et al*., 1986). The GO model may then be used to model the bidirectional reflectance of the canopies. In this approach, the effects due to shadowing on the overall reflectance (or infrared temperature) from a scene become important and useful features and the correlated interactions between shaded and sunlit components are built into the analysis—although it now becomes nonlinear. The directional radiance of the vegetation is then a mixture of four components (sunlit and shaded tree crowns, and sunlit and shaded backgrounds) that is seen from a given viewing angle. The areal proportions of these four components, for given illumination and viewing directions (which can be off-nadir), will be a function of the sizes, shapes, orientations and placements of the objects (that is, individual plants) within the scenes.

A GO model is most appropriate to woodlands or vegetation in which the vegetative cover is discontinuous, that is, where tree and shadowed background interactions account for a large proportion of the variance in the image. The further advantage of such models would be that they could also potentially be invertible and provide structural as well as cover information. The invertibility of GO models was demonstrated in principle by Franklin and Strahler (1988) and Wu and Strahler (1993) in which

tree size and density were estimated from reflectance data. When size, shape and orientation are fixed or characterised by distributions of known parameters, and the object centres are randomly distributed, then the proportions of the four components can be estimated using the Boolean model of Serra (1982). This GO model is termed the Boolean version (Strahler and Jupp, 1991a; 1991b; Li and Strahler, 1992). It accounts for the changes in proportions that occur with random overlapping objects as the density of objects increases and can easily model scale effects and changing Sun and view directions. The GO aspect of the model implies that multiple scattering of radiation in the vegetation layer is neglected. While the evidence of our eyes supports this, there are wavelengths (in particular the near infrared) where multiple scattering is very significant (Li *et al.*, 1995).

A1.3 Semi-empirical Methods

Semi-empirical BRDF models have also been developed which are empirical but are based on physical models and often contain parameters that relate to surface parameters. They offer the means of using the mosaic approach and 'empirical' model fitting but also enable surface parameters to be extracted. The methods are called 'semi-empirical' because they generally assume atmospheric correction can be done and also often have forms that involve Earth surface parameters. The main class of these models is that of the 'kernel' models.

The following text describing the kernel approach was extracted from the MODIS-Land ATBD (Strahler *et al.*, 1995b).

A1.3.1 Theoretical description

Kernel-driven models for the bidirectional reflectance distribution function of vegetated land surfaces attempt to describe the BRDF as a linear superposition of a set of kernels that describe basic BRDF shapes, with the coefficients or weights chosen to adapt the sum of the kernels to the given case. Typically, semi-empirical kernels are based either on one of several possible approximations to a radiative transfer scenario of light scattering in a horizontally homogeneous plant canopy (such as a crop canopy), or on one of several approximations feasible in a geometric-optical model of light scattering from a surface covered with vertical projections that cast shadows (such as a forest canopy). Deriving a kernel of this nature requires simplifying and manipulating a model for the BRDF until it reaches the form:

$$R = c_1 k + c_2$$

in which k is a function only of view and illumination geometry, c_1 and c_2 are constants containing physical parameters, and R is the modelled value of the true BRDF, ρ.

Semi-empirical kernels can be of two types. First, they may contain only geometric terms, but no physical parameters. The complete model then is linear, and may be scaled to arbitrary scales even for mixed scenes; neglecting adjacency effects, the weights of the kernels will be linear functions of the areal proportions of the sub-pixel weights. The so-called Ross-kernels, which are approximations to the radiative transfer theory in plant canopies of Ross (1981) described below, belong to this class, as does the so-called Roujean geometric-optical kernel (Roujean *et al.*, 1992). In the second case, kernels contain one or very few physical parameters and thus instead of having one kernel, provide a family of kernels depending on these parameters. The geometric-optical Li-kernels (Wanner *et al.*, 1995) belong to this type. In order to reduce them to the form given in the previous equation, the kernel-internal parameters are fixed to a selection of a few values each. The following discussion presents each of the kernels used in the BRDF/Albedo algorithm. For more complete information on the theory and derivation of the kernels, see Wanner *et al.* (1995).

A1.3.2 Kernels

The Ross kernels are derived from a formula presented by Ross (1981) for the directional reflectance above a horizontally homogeneous plant canopy calculated from radiative transfer theory in a single scattering approximation. The Ross-thick kernel was derived and described by Roujean *et al.* (1992). It is based on an approximation for large LAI values:

$$k_{thick} = \frac{\left(\frac{\pi}{2} - \xi\right)\cos\xi + \sin\xi}{\cos\theta_i + \cos\theta_v} - \frac{\pi}{4}$$

$$c_1 = \frac{4s}{3\pi}\left(1 - e^{-LAIB}\right)$$

$$c_2 = \frac{s}{3} + e^{-LAIB}\left(\rho_s - \frac{s}{3}\right)$$

In the kernel, θ_i and θ_v are zenith angles for illumination and view, respectively; ϕ is the relative azimuth of illumination and view directions; and ξ is the phase angle of scattering:

$$\cos\xi = \cos\theta_i \cos\theta_v + \sin\theta_i \sin\theta_v \cos\phi$$

In the constants:

s is leaf reflectance (= leaf transmittance);

ρ_s is the (assumed isotropic) surface reflectance of the soil or understory;

LAI is the leaf area index; and

B is the average of secants of possible view and illumination zenith angles. For this formulation, a spherical leaf angle distribution is assumed.

The Ross-thin kernel simplifies Ross's equation based on an approximation for small LAI values:

$$k_{thin} = \frac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\cos\theta_i \cos\theta_v} - \frac{\pi}{2}$$

$$c_1 = \frac{2sLAI}{3\pi}$$

$$c_2 = \frac{sLAI}{3} + \rho_s$$

Although this kernel applies primarily to the case of a thin canopy of scatterers over a uniform background, it can also be appropriate for a very dense, uniform canopy of high leaf area, since in that case the leaf layers below the uppermost can act like a uniform background (Strahler *et al.*, 1995a).

The derivation of the Roujean geometric-optical kernel is presented in the appendix to the paper of Roujean *et al.* (1992). The kernel may be regarded as accounting for the scattering of a random arrangement of three-dimensional rectangular solids ('bricks') with isotropic scattering surfaces. Shadows are taken as perfectly black. Mutual shadowing, in which the shadow of one object falls on another object, is not taken into account. The bricks are long with respect to their height *h* and width *b*. The kernel is:

$$k_{brick} = \frac{1}{2\pi}\left[(\pi - \phi)\cos\phi + \sin\phi\right]\tan\theta_i \tan\theta_v - \frac{1}{\pi}\left(\tan\theta_i + \tan\theta_v + \sqrt{\tan^2\theta_i + \tan^2\theta_v - 2\tan\theta_i \tan\theta_v \cos\phi}\right)$$

$$c_1 = \rho_s \frac{h}{b}$$

$$c_2 = \rho_s$$

The Li kernels are derived from the modelling approach of Li and Strahler (1986, 1992). In this approach, the surface is taken as covered by randomly-placed projections (such as tree crowns) that are taken to be spheroidal in shape and centred randomly within a layer above the surface. The BRDF is modelled as a function of the relative areas of sunlit and shaded, crown and background that are visible from the viewing position in the hemisphere. For the Li-sparse kernel, it is assumed that shaded crown and shaded background are black, and that sunlit crown and background are equally bright. Under these circumstances, and with some further approximations in the way that view and illumination shadows overlap, the Li-sparse kernel is:

$$k_{sparse} = (O - 1)(\sec\theta'_i + \sec\theta'_v) + \frac{1}{2}(1 + \cos\xi')\sec\theta'_v$$

where

$$O = \frac{1}{\pi}(t - \sin t \cos t)$$

$$\cos t = \frac{h}{b}\frac{\sqrt{D^2 + (\tan\theta'_i \tan\theta'_v \sin\phi)^2}}{\sec\theta'_i + \sec\theta'_v}$$

$$D = \sqrt{\tan^2\theta'_i + \tan^2\theta'_v - 2\tan\theta'_i \tan\theta'_v \cos\phi}$$

$$\cos\xi' = \cos\theta'_i \cos\theta'_v + \sin\theta'_i \sin\theta'_v \cos\phi$$

$$\theta' = \tan^{-1}\left(\frac{b}{r}\tan\theta\right)$$

In these expressions, *b* is the vertical radius of the spheroid; *r* is the horizontal radius of the spheroid; and *h* is the height of the centre of the spheroid. For this model,

$$c_1 = C\lambda\pi r^2$$

$$c_2 = C$$

Here, *C* is the brightness of sunlit surface, and λ is the count density of spheroids (number of spheroids per unit area).

The Li-dense kernel differs from the Li-sparse kernel in that it accommodates mutual shadowing. It assumes a random distribution of crown heights to maximize the geometric-optical effect in a dense ensemble of canopies.

$$k_{dense} = \frac{\left(1+\cos\xi'\right)\sec\theta'_v}{\left(1-O\right)\left(\sec\theta'_i+\sec\theta'_v\right)} - 2$$

$$c_1 = \frac{C}{2}$$

$$c_2 = C$$

These kernels are not yet linear in that they still contain two parameters, namely the ratios *b/r* and *h/b*, describing crown shape and relative height. The kernels therefore actually represent two families of kernels, governed by the values of these two internal parameters. For the present, we provide two choices for each parameter, thus providing four kernels within each family. For the *b/r* ratio, the values 2.5 (prolate shape) and 0.75 (oblate shape) are used; for the *h/b* ratio, we use values of 2.5 (tall) and 1.5 (short).

A1.3.3 Kernel-driven models

A complete kernel-driven semi-empirical model is formulated as a linear combination of kernels. Most suitably it has the form:

$$R = f_{iso} + f_{geo}k_{geo} + f_{vol}k_{vol}$$

which is derived from adding appropriate choices of geometric-optical surface-scattering and radiative-transfer volume-scattering kernels, each multiplied by a proportion α or (1- α) that weights the contribution of each model. These proportions may be regarded as the areal proportions of land cover types exhibiting each type of scattering (neglecting multiple scattering between the two components), or as mixing proportions for land cover types that display both a volume-scattering and a geometric-optical contribution to the BRDF. The quantities k_{geo} and k_{vol} are the respective kernels; the factors f_{geo} and f_{vol} are their respective weights; and the term f_{iso} is the isotropic contribution. Using the kernels presented above, there will be two choices for k_{vol} and three for k_{geo}, providing six possible models (disregarding here the four sets of internal parameters used for the Li kernels). The formulae for f_{iso}, f_{geo}, and f_{vol} are shown in Table A1.1.

In the inversion and fitting of a semi-empirical model to data, estimates of the weights *f* are retrieved from bidirectional reflectances and specification of viewing and illumination positions. Although this objective satisfies many of the goals of the BRDF/Albedo product, the existence of formulae for the weights *f* in terms of physical parameters could provide for direct inference of physical parameters from the weight values fitted.

Table A1.1 Semi-empirical weight formulas

Model	Weight Formulae
Ross-thin + Roujean	$f_{iso} = \alpha\rho_s = \left(1-\alpha\right)\left(\frac{sLAI}{3}+\rho_s\right)$ $f_{geo} = \alpha\rho_s\frac{h}{b}$ $f_{vol} = \left(1-\alpha\right)\frac{2sLAI}{3\pi}$
Ross-thick + Roujean	$f_{iso} = \alpha\rho_s = \left(1-\alpha\right)\left(\frac{s}{3}+e^{-LAIB}\left(\rho_s-\frac{s}{3}\right)\right)$ $f_{geo} = \alpha\rho_s\frac{h}{b}$ $f_{vol} = \left(1-\alpha\right)\frac{4s}{3\pi}\left(1-e^{-LAIB}\right)$
Ross-thin + Li-sparse	$f_{iso} = \alpha C = \left(1-\alpha\right)\left(\frac{sLAI}{3}+\rho_s\right)$ $f_{geo} = \alpha C\lambda\pi r^2$ $f_{vol} = \left(1-\alpha\right)\frac{2sLAI}{3\pi}$
Ross-thick + Li-sparse	$f_{iso} = \alpha C = \left(1-\alpha\right)\left(\frac{s}{3}+e^{-LAIB}\left(\rho_s-\frac{s}{3}\right)\right)$ $f_{geo} = \alpha C\lambda\pi r^2$ $f_{vol} = \left(1-\alpha\right)\frac{4s}{3\pi}\left(1-e^{-LAIB}\right)$
Ross-thin + Li-dense	$f_{iso} = \alpha C = \left(1-\alpha\right)\left(\frac{sLAI}{3}+\rho_s\right)$ $f_{geo} = \alpha\frac{C}{2}$ $f_{vol} = \left(1-\alpha\right)\frac{2sLAI}{3\pi}$
Ross-thick + Li-sparse	$f_{iso} = \alpha C = \left(1-\alpha\right)\left(\frac{s}{3}+e^{-LAIB}\left(\rho_s-\frac{s}{3}\right)\right)$ $f_{geo} = \alpha\frac{C}{2}$ $f_{vol} = \left(1-\alpha\right)\frac{4s}{3\pi}\left(1-e^{-LAIB}\right)$

A1.3.4 Modified Walthall model

Empirical models can be understood as being of the kernel-driven model type as well, where the kernels are empirical functions. An example is the modified Walthall model, derived by Walthall *et al.* (1985) and improved by Nilson and Kuusk (1989). It has the form:

$$R = p_0\left(\theta_i^2 + \theta_v^2\right) + p_1\theta_i^2\theta_v^2 + p_2\theta_i\theta v\cos\phi + p_3$$

Note that this is the same form as the semi-empirical models discussed above—it comprises a weighted superposition of functions of angles, and the weights p_0K3 are the parameters of the model. As a consequence, models like the modified Walthall model can be processed along with linear semi-empirical models by the same linear inversion scheme.

A1.3.5 Advantages of linear models

Linear models have a number of advantages in the context of global data processing. The opportunities arising from linearity in BRDF models have been discussed by Barnsley *et al.* (1990) and Strahler *et al.* (1995b). A first advantage is that, provided sufficient observations of BRDF are available, any linear model can be inverted analytically through matrix inversion for the system of equations obtained by setting the derivative of the error function to zero. This provides direct estimates of the parameters f_{iso}, f_{geo} and f_{vol} while avoiding numerical inversion problems associated with nonlinear models.

Second, both the directional-hemispherical and bi-hemispherical integrals of the BRDF (black-sky and white-sky albedos) may be pre-calculated for each kernel individually. The albedo of a model then is simply the sum of the kernel albedos, weighted by *f* values. By using a look-up table, numerical integration of the models can thus be avoided.

Third, linear BRDF models scale linearly in space if adjacency effects are assumed to be small. This allows for mixed pixel cases, as indicated by the areal proportion parameter α in the model factors listed above. This feature also allows scaling BRDF and albedo from one spatial resolution up to a coarser one, such as to a particular resolution needed for a climate model. Finally, since some of the parameters driving the models are dependent on wavelength while others (such as structural parameters) are not, it may be possible to extract all or some of them from multiband analysis, making assumptions about the others.

A1.4 The Issue of Scale

The BRDF models mentioned above and used to develop the kernels are general 'average' scene models that apply to spatially averaged data over homogeneous target areas. They do not model the image variance (which is also angle dependent) nor do they apply to a single object in a scene. The BRDF of a single tree is different from that of a stand of trees and the BRDF of a stand of trees is different from that of an open soil area.

Scaling affects the basis for image stratification. The tree and gap structure at one scale provides different functions for a stand and bare area delineation. At a broad scale, the composite of tree and gap merges into stands and at another stands and clearings merge in a 'woodland' category. At each level, the BRDF of the separated components will be different from each other and the composite land cover.

This leads to the consideration of some issues most important in an approach to correcting high resolution images for BRDF effects. The atmospheric effect is very smooth having been 'filtered' by an MTF with a size of about 200 m in most cases. However, the BRDF is highly variable. Consider a photograph with a wide field of view and the solar hotspot near to its centre. A tree near the nadir view and a similar one near the edge will have different brightness but the difference will not be as great as the overall brightening due to Volume and hotspot effects. In particular, there will be much more shadow at the edges reached away from the hotspot direction and path lengths for Volume scattering will be longer through the canopy than at the geometric centre of the frame. The 'colour' of shadow near the edges will also be slightly darker than at the nadir but not very different. In this case, applying an overall BRDF correction function will tend to change the colour of shadows and tree crowns at the edge in a way that may be quite wrong—even though the overall area average may be 'balanced'. Obviously, what works at one scale may not help at another.

Even if an H-resolution (high resolution) approach is taken in which the BRDF of the objects is modelled (trees for example), the fact will still be that the image contains different geometric proportions of scene objects and components in a way that depends on the phase angles between the Sun, the target object(s) and the view. Perhaps the only perfect way to resolve the issue is by 3D stereo modelling. This is unlikely to be cost-effective with video or digitised aerial photograph data.

BRDF models are further detailed in Appendix 2.

A1.5 References

Adams, J.B., Smith, M.O., and Gillespie, A.R. (1989). Simple models for complex natural surfaces: a strategy for the hyperspectral era of remote sensing. *Proc. IGARSS'89*, Vancouver, B.C., Canada, 16–21.

Barnsley, M., Morris, K., and Reid, A. (1990). Preliminary analysis of a multiple view angle image data set. *Proc. NERC 1989 Airborne Remote Sensing Campaign Symposium*, Keyworth, England, 18–19 December 1990, 49–68.

Boardman, J. (1990). Inversion of high spectral resolution data. *Proc. SPIE, vol 1298, Imaging Spectroscopy and Terrestrial Environment*, Orlando, FL, 222–223.

Chen, J.M., and Leblanc, S.G. (1997). A four-scale bidirectional reflectance model based on canopy architecture. *IEEE Transactions on Geoscience and Remote Sensing,* 35(5), 1316–1337.

Chen, J.M., and Leblanc, S.G. (2001). Multiple-scattering scheme useful for geometric optical modeling. *IEEE Transactions on Geoscience and Remote Sensing,* 39(5), 1061–1071.

Chen, J.M., J. Liu, Leblanc, S.G., Lacaze, R., and Roujean, J.-L. (2003). Multi-angular optical remote sensing for assessing vegetation structure and carbon absorption. *Remote Sensing of Environment,* 84(4), 516–525.

Chen, J.M., Menges, C.H., and Leblanc, S.G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sensing of Environment*, 97(4), 447–457. doi: *10.1016/j.rse.2005.05.003*

Chen, J.M., Li, X., Nilson. T., and Strahler, A. (2000). Recent advances in geometrical optical modelling and its applications. *Remote Sensing Reviews,* 18 (2–4), 227–262, doi: *10.1080/02757250009532391*

Cross, A.M., Settle, J.J. Drake, N.A., and Paivinen, R.T.W. (1991). Subpixel measurement of tropical forest cover using AVHRR data. *International Journal of Remote Sensing,* 12, 1119–1129.

Franklin, J., and Strahler, A.H. (1988). Invertible canopy reflectance modeling of vegetation structure in semi-arid woodland. *IEEE Transactions on Geoscience and Remote Sensing*, 26, 809–825.

Hanan, N.P. (2001). Enhanced two-layer radiative transfer scheme for a land surface model with a discontinuous upper canopy. *Agricultural and Forest Meteorology,* 109(4), 265–281.

Haverd, V., Lovell, J.L., Cuntz, M., Jupp, D., and Newnham, G. (2012). The Canopy Semi-analytic P_{gap} and Radiative Transfer (CanSPART) model: Formulation and application. *Agricultural and Forest Meteorology*, 160, 14–35. doi: *10.1016/j.agrformet.2012.01.018*

Huemmrich, K.F., and Goward, S.N. (1997). Vegetation Canopy PAR Absorptance and NDVI: An Assessment for Ten Tree Species with the SAIL Model. *Remote Sensing of Environment,* 61(2), 254–269.

Huemmrich, K.F. (2001). The GeoSail model: a simple addition to the SAIL model to describe discontinuous canopy reflectance. *Remote Sensing of Environment,* 75(3), 423–431.

Jupp, D.L.B., and Walker, J. (1996). Detecting structural and growth changes in woodlands and forests: the challenge for remote sensing and the role of geometric optical modelling. In: Gholz, H.L., Nakane, K. and Shimoda, H. (eds). *The Use of Remote Sensing in the Modeling of Forest Productivity at Scales from the Stand to the Globe*. Kluwer Academic Publishers, Dordrecht.

Jupp, D.L.B., and Strahler, A. (2000). *Image Brightness and BRDF Workshop Issues*. CSIRO Earth Observation Centre, Canberra.

Jupp, D.L.B., Walker, J., and Penridge, L.K. (1986). Interpretation of vegetation structure in Landsat MSS imagery: a case study in disturbed semi-arid eucalypt woodland. Part 2. Model based analysis. *Journal of Environmental Management*, 23, 35–57.

Kuusk, A., and Nilson, T. (2000). A directional multispectral forest reflectance model. *Remote Sensing of Environment*, 72(2), 244–252.

Kuusk, A. (2001). A two-layer canopy reflectance model. *Journal of Quantitative Spectroscopy and Radiative Transfer,* 71, 1–9.

Lacaze, R., and Roujean, J.L. (2001). G-function and HOt SpoT (GHOST) reflectance model—Application to multi-scale airborne POLDER measurements. *Remote Sensing of Environment,* 76(1), 67–80.

Li, X., and Strahler, A.H. (1986). Geometric-optical bidirectional reflectance modeling of a coniferous forest canopy. *IEEE Transactions on Geoscience and Remote Sensing*, GE24(6), 906–919.

Li, X., and Strahler, A.H. (1992). Geometric-optical bidirectional reflectance modeling of the discrete-crown vegetation canopy: Effect of crown shape and mutual shadowing, *IEEE Transactions on Geoscience and Remote Sensing*, 30, 276–292.

Li, X., Strahler, A.H., and Woodcock, C.E. (1995). A hybrid geometric optical-radiative transfer approach for modeling albedo and directional reflectance of discontinuous canopies, *IEEE Transactions on Geoscience and Remote Sensing*, 33, 466–480.

Liang, S., Strahler, A.H., Barnsley, J., Borel, C.C., Gerstl, S.A.W., Diner, D.J., Prata, A.J., and Walthall, C.L. (2000). Multiangle Remote Sensing: Past, Present and Future. *Remote Sensing Reviews,* 18, 83–102.

Myneni, R.B., and Ross, J. (1990). *Applications in Optical Remote Sensing and Plant Ecology*. Heidelberg, FRG: (Springer-Verlag).

Ni, W., Li, X., Woodcock, C.E., Roujean, J.-L., and Davis, R.E. (1997). Transmission of solar radiation in boreal conifer forests: Measurements and models. *Journal of Geophysical Research*, 102(D24), 29555–29566.

Ni, W., Li, X., Woodcock, C.E., Caetano, M.R., and Strahler, A.H. (1999). An analytical hybrid GORT model for bidirectional reflectance over discontinuous plant canopies. *IEEE Transactions on Geoscience and Remote Sensing,* 37(2), 987–999.

Ni-Meister, W., Lee, S., Strahler, A.H., Woodcock, C., Schaaf, C., Yao, T., Ranson, K.J., Sun, G., Bryan, B.J. (2010). Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. *Journal of Geophysical Research*, 115. doi: *10.1029/2009JG000936*.

Nilson, T., and Kuusk, A. (1989). A reflectance model for the homogeneous plant canopy and its inversion. *Remote Sensing of Environment*, 27, 157–167.

Nilson, T., and U. Peterson (1991). A forest canopy reflectance model and a test case. *Remote Sensing of Environment,* 37(3), 131–142.

Ong, C., Hick, P., Craig, M., Warren, P., and Newman, C. (1995). A correlative technique for correction of shading effects in digital multispectral video imagery. *Proc. ISSSR*, Melbourne, November 1995.

Pech, R.P., Davis, A.W., and Graetz, R.D. (1986). Reflectance modelling and the derivation of vegetation indices for an Australian semi-arid shrubland. *International Journal of Remote Sensing*, 7, 389–412.

Pickup, G., and Foran, B.D. (1987). The use of spectral and spatial variability to monitor cover change on inert landscapes. *Remote Sensing of Environment*, 23, 351–63.

Pickup, G., Chewings, V.H., and Pearce, G. (1995). Procedures for correcting high resolution airborne video imagery. *International Journal of Remote Sensing*, 16, 1647–1662.

Pinty, B., Widlowski, J.-L., Taberner, M., Gobron, N., Verstraete, M. M., Disney, M., Gascon, F., Gastellu, J.-P., Jiang, L., Kuusk, A., Lewis, P., Li, X., Ni-Meister, W., Nilson, T., North, P., Qin, W., Su, L., Tang, S., Thompson, R., Verhoef, W., Wang, H., Wang, J, Yan, G., Zang, H. (2004). Radiation Transfer Model Intercomparison (RAMI) exercise: Results from the second phase. *Journal of Geophysical Research*, 109, D06210, doi: *10.1029/2003JD004252*.

Qin, W., and Jupp, D.L.B. (1993). An Analytical and Computationally Efficient Reflectance Model for Leaf Canopies. *Agricultural and Forest Meteorology*, 66, 31–64.

Qin, W., and Liang, S. (2000). Plane parallel canopy radiation transfer modeling: Recent advances and future directions, *Remote Sensing Reviews*, 18(2-4), 281–305. doi: *10.1080/02757250009532393*

Ross, J.K. (1981). *The radiation regime and architecture of plant stands*. Junk Publ., The Hague. 391pp.

Roujean, J.L. (2000). A parametric hot spot model for optical remote sensing applications. *Remote Sensing of Environment,* 71(2), 197–206.

Roujean, J.L., Latoy, M., and Deschamps, P.Y. (1992). A bidirectional reflectance model of the earth's surface for the correction of remote sensing data. *Journal of Geophysical Research*, 97, 20455–20468.

Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London, New York.

Strahler, A.H., and Jupp, D.L.B. (1991a). Modeling bidirectional reflectance of forests and woodlands using boolean models and geometric optics. *Remote Sensing of Environment*, 34, 153–166.

Strahler, A.H., and Jupp, D.L.B. (1991b). Geometric-Optical Modeling of Forests as Scenes Composed of Three-Dimensional Discrete Objects. In *Photon-Vegetation Interactions: Applications in Optical Remote Sensing and Plant Ecology*, (Eds: R.B Myneni and J. Ross), Springer-Verlag, Heidelberg, FRG, 415–440.

Strahler, A., Wanner, W., Zhu, Q., and Jin, X. (1995a). Bidirectional reflectance modeling of data from vegetation obtained in the Changchun solar simulation laboratory, *Proc. 15th Int. Geosci. and Remote Sensing Symp.,* Florence, Italy, July 10-14, 1995, vol. 3. pp. 1965–1967.

Strahler, A.H., Barnsley, M.J., d'Entremont, R., Hu, B., Lewis, P., Li, X., Muller, J-P., Barker Schaaf, Wanner, W., and Zhang, B. (1995b). *MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document Version 3.2*. NASA EOS, May 1995.

Suits, G.H. (1972). The calculation of the directional reflectance of a vegetative canopy. *Remote Sensing of Environment,* 2, 117–125.

Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sensing of Environment,* 16, 125–141.

Verhoef, W. (1985). Earth observation modeling based on layer scattering matrices. *Remote Sensing of Environment*, 17, 165–178.

Walker, J., Jupp, D.L.B., Penridge, L.K., and Tian, G. (1986). Interpretation of vegetation structure in Landsat MSS imagery: a case study in disturbed semi-arid eucalypt woodland. Part 1. Field data analysis. *Journal of Environmental Management*, 23, 19–33.

Walthall, C.L., Norman, J.M., Welles, J.M., Campbell, G., and Blad, B.L. (1985). Simple equation to approximate the bidirectional reflectance from vegetation canopies and bare soil surfaces. *Applied Optics*, 24, 383–387.

Wanner, W., Li, X., and Strahler, A.H. (1995). On the derivation of kernels for kernel-driven models of bidirectional reflectance, *Journal of Geophysical Research*, 21077–21090.

White, H.P., Miller, J.R., and Chen, J.M. (2001). Four-Scale Linear Model for Anisotropic Reflectance (FLAIR) for plant canopies - Part I: Model description and partial validation. *IEEE Transactions on Geoscience and Remote Sensing,* 39(5), 1072–1083.

White, H.P., Miller, J.R., and Chen, J.M. (2002). Four-Scale Linear Model for Anisotropic Reflectance (FLAIR) for plant canopies - Part II: Validation and inversion with CASI, POLDER, and PARABOLA data at BOREAS. *IEEE Transactions on Geoscience and Remote Sensing*, 40(5), 1038–1046.

Widlowski, J. L., Taberner, M., Pinty, B., Bruniquel-Pinel, V., Disney, M., Fernandes, R., Gastellu-Etchegorry, J.-P., Gobron, N., Kuusk, A., Lavergne, T., Leblanc, S., Lewis, P.E., Martin, E., Mõttus, M., North, P.R.J., Qin, W., Robustelli, M., Rochdi, N., Ruiloba, R., Soler, C., Thompson, R., Verhoef, W., Verstraete, M.M., Xie, D. (2007). Third Radiation Transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance models, *Journal of Geophysical Research,* 112, D09111. doi: *10.1029/2006JD007821*.

Wu., Y., and Strahler, A.H. (1993). Remote estimation of crown size, stand density and foliage biomass on the Oregan transect. *Ecological Applications*, 4, 299–312.

Yang, R.Q., and Friedl, M.A. (2003). Modeling the effects of three-dimensional vegetation structure on surface radiation and energy balance in boreal forests. *Journal of Geophysical Research*, 108(D16).

Appendix 2
BRDF Models

This Appendix reviews a number of Bidirectional Reflectance Distribution Functions (BRDF) that are commonly used in conjunction with EO imagery[3].

A2.1 Kernel Models

These are (semi-)empirical models based on linear combinations of 'kernels':

$$\rho = f_{iso} + f_{geo} k_{geo} + f_{vol} k_{vol}$$

which represent surface reflectance (ρ) as a function of component reflectances (f_x) and the kernels (k_x), which are mathematical functions that depend on Sun (or incident) and view (or observer) angles θ_i and θ_v. The subscripts 'geo' and 'vol' refer to the physical bases for some kernels in which there is an identification of a 'geometric' or hotspot factor and a 'volume' or path length and scattering factor.

By convention, kernel models are expressed in way is that when Sun and observer are at zenith:

$$k_{geo}(0,0) = k_{vol}(0,0) = 0$$

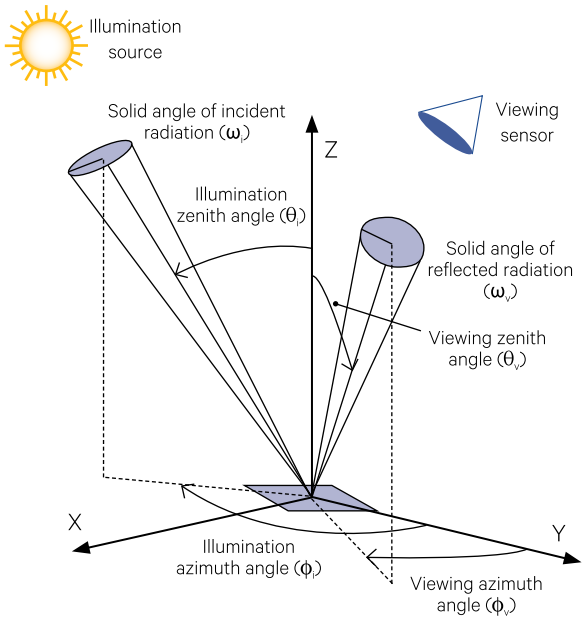
so that $\rho(0,0) = f_{iso}$. Some published models do not enforce this convention but all can if needed.

To compute the kernel models, we need to specify which choice of the angular description is being used for Sun and view directions. Sometimes this is 'target based' and sometimes it is 'sensor based'. Relative to an airborne scanner, the position of the hotspot is dependent on flight direction, aircraft attitude, scan read-out order, etc.

In an aircraft and scanner situation, the pointing vector for a pixel may be defined pointing 'up' or 'down' and there are important issues of whether the scanner scans left to right or right to left! In atmospheric correction routines, the conventions of the scanner and aircraft model frame must be very carefully specified. Any data on angles that came with images being investigated with kernel models must follow a single convention or be carefully and fully specified.

For the following, the consistent convention being used is the 'target based' one. That is, the geometry is as illustrated in Figure A2.1, where θ_i is the Sun or incident ray zenith angle, θ_v is the view or observed zenith angle and the azimuth (ϕ) is relative to the Sun position. This is the convention used by the MODIS MODLAND team in their product ATBD (Strahler *et al.*, 1996) and also in AVHRR pathfinder.

Figure A2.1 Geometry of remote observer



3 In this appendix, most text and equations have been sourced from Jupp (2000).

Background image: The appearance of identical metal cylinders, illuminated from a fixed light source, changes with their position relative to the viewer. This rendered image is based on the Ward Duer model, an anisotropic BRDF model for brushed metal effects (Ward, 1992; Walter, 2005).

Two important functions of the view and Sun positions, which are used often in the following, are the angle between the Sun and view vectors (ξ) and the related 'distance' between the Sun and view vectors (D) as defined by:

$$\cos\xi = \cos\theta_i \cos\theta_v + \sin\theta_i \sin\theta_v \cos\phi$$

and

$$D = \left[\tan^2\theta_i + \tan^2\theta_v - 2\tan\theta_i \tan\theta_v \cos\phi \right]^{1/2}$$

Clearly, if the Sun and view positions align (so that $\theta_v = \theta_i$ and ϕ=0) then both ξ and D equal zero. To see how they operate away from this 'hotspot' point it is useful to consider the 'Principal azimuth' (or almucantar) for which $\theta_v = \theta_I$. That is left to you.

A2.2 Primary AMBRALS Kernel Models

The base components of the AMBRALS kernel models (Wanner *et al.*, 1995) can either be used as individual kernels of combinations or as described in Appendix A2.2.8. They are as follows:

A2.2.1 Ross-thin (vol) model

Derived in Wanner *et al.* (1995) based on Ross (1981):

$$\cos\xi = \cos\theta_i \cos\theta_v + \sin\theta_i \sin\theta_v \cos\phi$$

$$k_{thin} = \frac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\cos\theta_i \cos\theta_v} - \frac{\pi}{2}$$

A2.2.2 Ross-thick (vol) model

Derived in Roujean *et al.* (1992) based on Ross (1981):

$$k_{thick} = \frac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\cos\theta_i + \cos\theta_v} - \frac{\pi}{4}$$

A2.2.3 Roujean (geo) model

Original Roujean model was 'Roujean' + Ross-thick kernel model. Here 'Roujean' refers to the geometric kernel:

$$D = \left[\tan^2\theta_i + \tan^2\theta_v - 2\tan\theta_i \tan\theta_v \cos\phi \right]^{1/2}$$

$$k_{brick} = \frac{1}{2\pi}\left[(\pi - \phi)\cos\phi + \sin\phi \right]\tan\theta_i \tan\theta_v - \frac{1}{\pi}\left(\tan\theta_i + \tan\theta_v + D \right)$$

A2.2.4 Li-sparse (geo) model

The Li-sparse and Li-dense kernels are derived in Wanner *et al.* (1995). A detailed mathematical outline of the computation (the derivation is discussed in Appendix A2.4.3) is as follows:

- sequence of steps for calculations:

$$\theta_i' = \tan^{-1}\left(\frac{b}{r}\tan\theta_i \right)$$

$$\cos\xi' = \cos\theta_i'\cos\theta_v' + \sin\theta_i'\sin\theta_v'\cos\phi$$

$$D' = \left[\tan^2\theta_i' + \tan^2\theta_v' - 2\tan\theta_i'\tan\theta_v'\cos\phi \right]^{1/2}$$

$$\cos t = \min\left\{ 1, \frac{h}{b}\frac{\left[D'^2 + (\tan\theta_i'\tan\theta_v'\sin\phi)^2 \right]^{1/2}}{\sec\theta_i' + \sec\theta_v'} \right\}$$

$$t = \cos^{-1}(\cos t)$$

$$O(\tfrac{b}{r}, \tfrac{h}{b}, \theta_i, \theta_v, \phi) = \frac{1}{\pi}(t - \sin t\cos t)(\sec\theta_i' + \sec\theta_v')$$

- The Li-sparse kernel follows from these definitions as:

$$Li_sp = O - \sec\theta_i' - \sec\theta_v' + \frac{1}{2}(1 + \cos\xi')\sec\theta_v'$$

A2.2.5 Li-dense (geo) model

The Li-dense approximation refers to dense vegetation canopies. Using the same definitions as for the Li-sparse kernel it is:

$$Li_dn = \frac{(1 + \cos\xi')\sec\theta_v'}{\sec\theta_i' + \sec\theta_v' - O} - 2$$

A2.2.6 Cox-Munk (geo) model

Based on the Cox-Munk (1954) model for sunglint and presented in Strahler *et al.* (1996):

$$k_{spec} = \begin{cases} \frac{1}{\cos\theta_i}\left(1-\frac{\tan^2\theta_n}{\sigma^2}\right)-1 & if \ \frac{\tan^2\theta_n}{\sigma^2} \le 1 \\ -1 & else \end{cases}$$

where σ^2 is the wave slope variance (nonlinear parameter):

$$\sigma^2 = 0.003 + 0.00512\,w$$

w is the wind speed and the parameter θ_n is defined as:

$$\cos^2\theta_n = \frac{1}{2}\left(\frac{(\cos\theta_i + \cos\theta_v)^2}{1+\cos\phi}\right)$$

A2.2.7 Walthall (vol) model

This empirical model (also called the 'limaçon' function) was due to Walthall *et al.* (1985) and improved by Nilson and Kuusk (1989) to include reciprocity. It has four linear parameters:

$$\rho(\theta_i,\theta_v,\phi) = p_0(\theta_i^2+\theta_v^2) + p_1\theta_i^2\theta_v^2 + p_2\theta_i\theta_v\cos\phi + p_3$$

A2.2.8 MODIS products

Best fitting model from a selection of kernels (AMBRALS) include:

- Ross-thin + Li-sparse;
- Ross thin + Li-dense;
- Ross-thick + Li-sparse;
- Ross-thick + Li-dense;
- Cox-Munk + Li-sparse; and
- Walthall.

The first four of these have parameters that can be converted to 'physical' parameters.

A2.3 Extensions and Alternatives

A2.3.1 Reciprocal Li models

The proposed MODIS AMBRALS approach and the later choice of a single pair of Ross and Li models for operational MODIS BRDF products also used modified definitions of the kernels. The change was to make the models 'reciprocal' in that if the Sun and observer positions are interchanged the resulting BRDF value is unchanged. Reciprocity of the BRDF models is a fundamental property of the models (see Appendices A2.4.3 and A2.4.4), from which the Li kernels derive. Although the changes do not have a physical basis, the models are thought to be 'better' for having reciprocity. The changes were (for the same definitions of terms as before):

$$k_{Sparse,\,Recip} = (O-1)\left(\sec\theta_i' + \sec\theta_v'\right) + \frac{1}{2}\left(1+\cos\xi'\right)\sec\theta_i'\sec\theta_v'$$

$$k_{Dense,\,Recip} = \frac{\left(1+\cos\xi'\right)\sec\theta_i'\sec\theta_v'}{(1-O)\left(\sec\theta_i' + \sec\theta_v'\right)} - 2$$

A2.3.2 RPV model

Rahman, Pinty and Verstraete (RPV) model derived in Rahman *et al.* (1993) based on Pinty *et al.* (1990) and Pinty and Verstraete (1991):

$$\rho(\theta_i,\theta_v,\phi) = \rho_0\, M(\theta_i,\theta_v,k)\, P(g,\xi)\, H(\rho_0,D)$$

$$M(\theta_i,\theta_v,k) = \frac{\cos^{k-1}\theta_i \cos^{k-1}\theta_v}{(\cos\theta_i + \cos\theta_v)^{1-k}}$$

$$P(g,\xi) = \frac{1-g^2}{\left[1+g^2-2g\cos(\pi-\xi)\right]^{1.5}}$$

$$H(\rho_0,D) = 1+\frac{1-\rho_0}{\delta+D}$$

where the linear parameter is ρ_0 and g, k [and δ] are nonlinear parameters (generally $\delta \approx 1$). M is the Minnaert model, P is the Henyey Greenstein phase function and H is the hotspot function.

A2.3.3 MRPV model

Multiangle Imaging Spectroradiometer (MISR) version of RPV model as suggested by Martonchik *et al.* (1998):

$$P(g,\xi) = P_M(\xi) = e^{-b_M\cos\xi}$$

$$H(\overline{\rho},D) = 1+\frac{1-\overline{\rho}}{1+D}$$

where $\overline{\rho}$ is an average reflectance assumed known if the algorithm is 'linearised' by taking logarithms; otherwise a simple nonlinear iteration can be used.

A2.3.4 Staylor and Suttles model

As given in Cosnefroy *et al.* (1996), following Staylor and Suttles (1986):

$$\rho(\theta_i,\theta_v,\phi) = \frac{1}{\cos\theta_i\cos\theta_v}\left[Y_0 + Y_1\left(\frac{\cos\theta_i\cos\theta_v}{\cos\theta_i+\cos\theta_v}\right)^N\right] \times \frac{1+C_{sw}\cos^2\xi}{1+C_{sw}\left[\cos^2\theta_i\cos^2\theta_v + (\sin^2\theta_i\sin^2\theta_v)/2\right]}$$

where Y_0 and Y_1 are two linear parameters and C_{sw} and N are two nonlinear parameters. This model is very flexible and quite useful. The nonlinear parameters seem to converge quite quickly.

A2.3.5 Shibayama and Weigand model

Developed in Shibayama and Weigand (1985, 1986) and used in Qi *et al.* (1995):

$$\rho(\theta_i,\theta_v,\phi) = \rho_0(1+\left[\beta_0+\beta_1\sin(\phi/2)+\beta_2(1/\cos\theta_i)\right]\sin\theta_v)$$

A2.3.6 Dymond and Qi model

Derived by Dymond and Qi (1997):

$$S(\theta_i,\theta_v) = \frac{\cos\theta_i}{\cos\theta_i + (\sigma_{\theta_i}/\sigma_{\theta_v})\cos\theta_v}$$

$$H(\xi,\theta_i) = \begin{cases} 2\,e^{-\tan(\xi/2)/h\theta_i} & \xi < \pi/2 \\ 2\,e^{-\tan(\pi/4)/h\theta_i} & \xi \geq \pi/2 \end{cases}$$

$$B(\xi) = \frac{4\rho_0}{3\pi^2}\left(\sin\xi + (\pi/2-\xi)\cos\xi\right)$$

$$\rho(\theta_i,\theta_v,\xi) = S(\theta_i,\theta_v)\, H(\xi,\theta_i)\, B(\xi)$$

Parameters are one linear (ρ_0) and two nonlinear (h and R) where $R = \sigma_i/\sigma_v$

A2.3.7 Chen modification to Roujean model

In Chen and Cihlar (1997) a modification of the Roujean model is proposed based on a simplification of Chen's more complex canopy model. The simple model reduces to:

$$\rho(\theta_i,\theta_v,\phi) = (f_{iso} + f_{geo}k_{brick} + f_{vol}k_{thick})(1+C_1 e^{-C_2\frac{\xi}{\pi}})$$

where the constants C_1 and C_2 must be determined from the data. This added flexibility allows the Roujean model to approximate the hotspot effect much more successfully than before. Perhaps choosing a better model would be even better!

The Chen modification could, of course, be applied to any of the Kernel models to 'sharpen' the hotspot. However, there are two extra parameters to estimate.

A2.3.8 Liang modification to Walthall model

In a similar way to the modification of the Roujean model by Chen to improve modelling of the hotspot, Liang and Strahler (1994) modified the Walthall model to include a hotspot effect. They increased the number of parameters from four to six by proposing a linear sum of the two models in the form:

$$\rho(\theta_i,\theta_v,\phi) = \rho_1(\theta_i,\theta_v,\phi) + \rho_2(\theta_i,\theta_v,\phi)$$

where the first is the Walthall model described above and:

$$\rho_2(\theta_i,\theta_v,\phi) = C_1\, e^{-C_2 \tan(\pi-\xi)}$$

where, as above, C_1 and C_2 are adjustable parameters. In this case, one is non-linear and the hotspot shape is added rather than used multiplicatively.

Jupp (1997) described how a Walthall model plus a Li-sparse kernel could also effectively model a woodland BRDF with the Walthall accommodating the Volume effect and the Li-sparse modelling the hotspot.

A2.3.9 Pickup and Chewings model

Pickup *et al.* (1995a,b) developed and extended an approach similar to one originally reported by Royer *et al.* (1985) to correct Video images for BRDF and other angular effects.

A simple linear 'kernel' model of the form:

$$\rho(\theta_i,\theta_v,\phi) = p_0 + p_1\xi + p_2\xi^2 + p_3\cos^4\theta_v$$

Was fitted to a line of Video data frames and then used to normalise the individual frames in the line. The $\cos^4$ term was introduced to account for lens effects but can also model a Volume effect. The authors claim the power is insensitive to choices between 2 and 4.

A2.4 Slightly More Complex Models: Not in 'Kernel' Form

A2.4.1 Otterman model

Used in Deering *et al.* (1990) and based on previous work by Otterman (such as Otterman and Weiss, 1984), where the geometry is based on thin vertical cylinders:

$$\rho = R_g e^{-s(\tan\theta_i+\tan\theta_v)} + R_p(1-e^{-s(\tan\theta_i+\tan\theta_v)})$$

$$R_g = (1-f) \times \frac{r(\sin\phi-\phi\cos\phi)+t((\phi-\pi)\cos(\phi-\pi)-\sin(\phi-\pi))}{4(\cot\theta_i+\cot\theta_v)} + fr_0$$

$$R_p = \frac{r_p(\sin\phi-\phi\cos\phi)}{4(\cot\theta_i+\cot\theta_v)}$$

where the parameters to be modelled are:

s or 'cylinder area index'; and
r_p or plant reflectance.

and three soil reflectance parameters:

$L = fr_0$

$B = (1-f)r$

$T = (1-f)t$

where

f is the Lambertian fraction;
r_0 is the Lambertian reflectance;
r is facet reflectance; and
t is facet transmittance.

That is, five parameters in all. Four of the parameters are linear but s is nonlinear.

A2.4.2 Verstraete, Pinty and Dickinson (VPD) model

Derived in Verstraete *et al.* (1990) and Pinty *et al.* (1990). This is a more complex form than the RPV model given above.

$$\rho(\theta_i,\theta_v,\phi)=\frac{\omega}{4}\frac{\kappa_i}{\kappa_i\mu_v+\kappa_v\mu_i}\times\left[\begin{array}{l}P_v(D)P(g,\xi)+\\H\left(\frac{\mu_i}{\kappa_i},\omega\right)H\left(\frac{\mu_v}{\kappa_v},\omega\right)-1\end{array}\right]$$

where

$\mu_i=\cos\theta_i$
$\mu_v=\cos\theta_v$

Following Dickinson *et al.* (1990) after Goudriaan (1977):

$$\kappa_x=\Psi_1+\Psi_2\mu_x$$
$$\Psi_1=0.5-0.489\chi_l-0.33\chi_l^2$$
$$\Psi_2=1-2\Psi_1$$

where

χ_l is a function of the leaf angle distribution of the canopy and varies from -0.4 for an erectophile canopy to 0.6 for a planophile canopy. Random orientation is zero.

$$P_v(D)=\frac{1}{1+V_P(D)}$$
$$V_P(D)=4\left(1-\frac{4}{3\pi}\right)\frac{D}{2r\Lambda}\frac{\mu_v}{\kappa_v}$$

For Isotropic: $P(g,\xi)=1$

For Henyey and Greenstein:

$$P(g,\xi)=\frac{1-g^2}{\left[1+g^2-2g\cos(\pi-\xi)\right]^{1.5}}$$

For Legendre:

$$P(g,\xi)=1+g\cos\xi+\frac{L_2}{2}(3\cos^2\xi-1)$$

$$H(x,\omega)=\frac{1+x}{1+x\sqrt{(1-\omega)}}$$

The parameters for this model that need to be fitted in inversion or supplied for a given land surface are:

ω is the single scattering albedo;
g is the asymmetry of the phase function (and L_2 if Legendre);
χ_l is the scatterer orientation parameter (used to obtain κ_i and κ_v); and
$2r\Lambda$ is the structural parameter (r is the sunfleck radius and Λ is the scatterer area density).

A2.4.3 Strahler and Jupp simple model

This is derived as in Strahler and Jupp (1991) but here using overlap function derived for the Li kernels (see Appendices A2.2.4 and A2.4.4). In this model, there are four kinds of ground cover 'visible' from a given direction. These are referred to as scene components and consist of sunlit canopy (*C*), shaded canopy (*T*), sunlit background (*G*), and shaded background (*Z*). Each component is assumed to have a characteristic reflectance and the reflectance of a pixel is modelled as the area weighted combination (or linear mixture) of the characteristic component reflectances. That is, the observed reflectance of a single pixel is modelled as:

$$\rho(\theta_i,\theta_v,\phi)=k_CR_C+k_TR_T+k_GR_G+k_ZR_Z$$

where *C*, *T*, *G*, and *Z* indicate the reflectances of the four components as named above, R_x represents the (mean) radiance of component '*x*' and *k* indicates the sensed proportion of each component within the pixel from the given view direction.

Obviously, with these definitions:

$$k_x\geq 0$$
$$\sum_{x=C,T,G,Z}k_x=1$$

If the proportions are replaced by their expectations then the various components can be derived from the equations:

$$k_G=e^{-\lambda\bar{A}(\sec\theta_i'+\sec\theta_v'-O)}$$
$$k_C+k_T=1-e^{-\lambda\bar{A}\sec\theta_v'}$$

which, given the closure above, means there is one extra condition to complete the model. This was done one way in Strahler and Jupp (1991) and in a second way in Li and Strahler (1992).

In Strahler and Jupp (1991) it was assumed that the proportion of visible sunlit tree was the same as for a single tree. That is:

$$\frac{k_C}{k_C+k_T}=\frac{1}{2}(1+\cos\xi')$$

This model is the basis for the simplified Li-sparse formulation. In the full model, the four signatures must be determined as well as the ratios *b*/*r* and *h*/*b* and the vertical projected 'crown area index', or cover, λA.

There are some simplifications. For example, in cases where the background is bright and trees dark it is possible to write:

$$\rho(\theta_i,\theta_v,\phi) \approx R_X + k_G(R_G - R_X)$$

where *X* refers to a composite of tree plus shade, which is assumed dark. Obviously,

$$R_X = \frac{k_C R_C + k_T R_T + k_Z R_Z}{1-k_G}$$

This simple two-component model can often do a very good job. If, however, the variation is dependent on the differences between sunlit and shaded tree then the model will not do so well. It is often better just to assume that shaded tree and shaded background have the same colour. This reduces the number of parameters and is usually a reasonable assumption.

In the Li-sparse kernel given above and derived in Wanner *et al.* (1995), it is assumed that R_T=R_Z=0 and R_C=R_G. Then:

$$\begin{aligned}\rho(\theta_i,\theta_v,\phi) &\approx (k_G + k_C)R_C \\ &= R_C\left[\begin{array}{l}\frac{1}{2}(1+\cos\xi')(1-e^{-\lambda\bar{A}\sec\theta_v'}) \\ +e^{-\lambda\bar{A}(\sec\theta_i'+\sec\theta_v'-O)}\end{array}\right] \\ &\approx R_C\lambda\bar{A}\left[\begin{array}{l}O-\sec\theta_i'-\sec\theta' \\ +\frac{1}{2}(1+\cos\xi')\sec\theta_v'\end{array}\right]+R_C \\ &= c_1 k_{sparse} + c_2\end{aligned}$$

which is the form previously given.

A2.4.4 Li's 'top easy seen' modification

Li and Strahler (1992) modified the simple model to take account of the fact that when the density of trees increases then the shaded crown tends to be 'hidden' and the view becomes dominated by the sunlit crown tops.

Of a number of choices, the one used for the simple dense canopy models assumes that:

$$f = \frac{k_C}{1-k_G} = F$$

where *F* is the ratio for a single crown or rather the very dense case. It follows that:

$$F = \frac{\frac{1}{2}(1+\cos\xi')\sec\theta_v'}{\sec\theta_v' + \sec\theta_i' - O}$$

The model may now be resolved and even inverted as it stands or it may be approximated in various ways with simpler forms. This model is used in dense vegetation and the previous one in sparse vegetation.

Note that this is equivalent to:

$$\begin{aligned}\frac{k_c}{k_c+k_t} &= F\frac{1-k_g}{k_c+k_t} \\ &= \frac{\frac{1}{2}(1+\cos\xi')\sec\theta_v'}{\sec\theta_v'+\sec\theta_i'-O}\,\frac{1-e^{-\lambda\bar{A}(\sec\theta_v'+\sec\theta_i'-O)}}{1-e^{-\lambda\bar{A}\sec\theta_v'}}\end{aligned}$$

which asymptotes to the same result as before when the density is low ($1-e^{-x} \approx x$).

To derive the Li-dense kernel (Wanner *et al.*, 1995), suppose R_G is neglected since there is so little sunlit soil showing in dense vegetation and also the forest floor may be dark material. Assume all shadow and R_G is zero. Then:

$$\begin{aligned}\rho(\theta_i,\theta_v,\phi) &= k_C R_C \\ &\approx f\,R_C \\ &= \left[\frac{(1+\cos\xi')\sec\theta_v'}{\sec\theta_v'+\sec\theta_i'-O} - 2\right]\frac{R_C}{2} + R_C \\ &= c_1 k_{dense} + c_2\end{aligned}$$

which is again in the form of the Li-dense kernel approximation. But again, it is possible also to solve the complete model or a model with some less restrictive assumptions if the linear kernel structure is not enforced.

A2.4.5 Hapke BRDF model for soils

Hapke (1981, 1986) derived a model for dimensionless particles, which has principally been applied to Soil data and is used also in the RPV and VPD models. The model was:

$$\begin{aligned}\rho(\theta_i,\theta_v,\phi) = \frac{\omega}{4}\frac{1}{\mu_v+\mu_i} \\ \times\left[\begin{array}{l}\left[1+B(\xi)\right]P(g,\xi) \\ +H\left(\mu_i,\omega\right)H\left(\mu_v,\omega\right)-1\end{array}\right]\end{aligned}$$

where

μ_i=$\cos\theta_i$
μ_v=$\cos\theta_v$

$B(\xi)$ is a backscattering function that accounts for the hotspot effect:

$$B(\xi) = \frac{S(0)}{\omega\, P(g,0)} \frac{1}{\left[1 + (1/h)\tan(\xi/2)\right]}$$

Where $S(0)$ defines the magnitude of the 'hotspot' and h defines the width.

$P(g,\xi)$ is the phase function for the particle collection with asymmetry g:

$$P(g,\xi) = \frac{1-g^2}{\left[1+g^2-2g\cos(\pi-\xi)\right]^{1.5}}$$

(Henyey and Greenstein)

$$P(g,\xi) = 1 + g\cos\xi + \frac{L_2}{2}(3\cos^2\xi - 1)$$

(Legendre)

In the Legendre case there is an extra parameter L_2.

$H(x,\omega)$ is a function to account for multiple scattering:

$$H(x,\omega) = \frac{1+2x}{1+2x\sqrt{(1-\omega)}}$$

The parameters for this model that need to be fitted in inversion or supplied for a given land surface are:

- ω is the single scattering albedo;
- g is the asymmetry of the phase function (plus L_2 for the Legendre phase function);
- $S(0)$ is a parameter defining the height of the hotspot function at the hotspot; and
- h is a parameter that controls the width of the hotspot function.

A2.4.6 Ross' simplified vegetation canopy formula

The 'turbid medium' formula by Ross (1981) which was used by Roujean *et al.* (1992) to derive a Volume kernel (the Ross-Thick) and by Wanner *et al.* (1995) to derive another (the Ross-Thin) kernel was simplified by both groups of writers to the following form as an initial step in their formulation:

$$\rho(\theta_i,\theta_v,\phi) = \frac{4\rho_L}{3\pi}\frac{(\pi/2-\xi)\cos\xi + \sin\xi}{\cos\theta_i + \cos\theta_v} \times\left(1 - e^{-\frac{LAI}{2}(\sec\theta_i+\sec\theta_v)}\right) + \rho_s e^{-\frac{LAI}{2}(\sec\theta_i+\sec\theta_v)}$$

where

- ρ_L is the reflectance of a single leaf;
- ρ_s is the background soil reflectance; and
- LAI is Leaf Area Index

If these three parameters are estimated subject to the reflectances being in the range (0,1) and $LAI>0$ then this provides a model that includes both the Ross-thin and the Ross-thick (Wanner *et al.*, 1995). As with all these non-kernel models, however, the fit will need to be nonlinear. However, this often presents little problem.

A2.4.7 Minnaert's original model

Minnaert (1941) proposed a model that has been used for modelling topographic shading and has been used as a component in a number of semi-empirical models. Its form is:

$$\rho(\theta_i,\theta_v) = \rho_L \frac{k+1}{2}\mu_i^{k-1}\mu_v^{k-1}$$

This function is a reciprocal 'volume' scattering kernel in its behaviour with k=1 being Lambertian, k=0 being a 'bowl' with highest reflectances away from nadir and k=2 being an inverted 'bowl' with darkening away from nadir.

Parameters to estimate are ρ_L and k.

Topographic shading and large pixels, although not described in detail here, can also be modelled as a BRDF effect. In addition to the Minnaert model, the model by Hapke (Hapke, 1984) has been used successfully in this case.

A2.5 References

Cosnefroy, H., Leroy, M., and Briottet, X. (1996). Selection and characterisation of Saharan and Arabian Desert sites for the calibration of optical satellite sensors. *Remote Sensing of Environment*, 58, 101–114.

Chen, J.M., and Cihlar, J. (1997). A hotspot function in a simple bidirectional reflectance model for satellite applications. *Journal of Geophysical Research*, 102(D22), 25907–25913. doi: *10.1029/97JD02010*.

Cox, C., and Munk, W. (1954). Measurement of the roughness of the sea surface from photographs of the sun's glitter. *Journal of the Optical Society of America*, 44, 838–850.

Deering, D.W., Eck, T.F., and Otterman, J. (1990). Bidirectional reflectances of selected desert surfaces and their three-parameter soil characterisation. *Agricultural and Forest Meteorology*, 52, 71–93.

Dickinson, R.E., Pinty, B., And Verstraete, M. (1990). Relating surface albedos in GCM to remotely sensed data. *Agricultural and Forest Meteorology*, 52, 109–131.

Dymond, J.R., and Qi, J. (1997). Reflection of visible light from a dense vegetation canopy - a physical model. *Agricultural and Forest Meteorology,* 86, 143–155.

Goudriaan, J. (1977). *Crop micrometeorology: a simulation study*. Wageningen: Wageningen Centre for Agricultural Publishing and Documentation.

Hapke, B.W. (1981). Bidirectional reflectance spectroscopy. 1. Theory. *Journal of Geophysical Research*, 86, 3039–3054.

Hapke, B.W. (1984). Bidirectional reflectance spectroscopy. 3. Correction for macroscopic roughness. *Icarus*, 59, 41–59.

Hapke, B.W. (1986). Bidirectional reflectance spectroscopy. 4. The extinction coefficient and the opposition effect. *Icarus*, 67, 264–280.

Jupp, D.L.B. (1997). Some Research and Applications in the CSIRO (Australia) Earth Observation Centre on Scene Brightness due to BRDF. *ENAMORS'97*, Tuusula, Finland.

Jupp (2000). *A Compendium of Kernel and Other (Semi-) Empirical BRDF Models*. CSIRO Earth Observation Centre, Canberra.

Li, X., and Strahler, A.H. (1992). Geometrical-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. *IEEE Transactions on Geoscience and Remote Sensing*, 30, 276–292.

Liang, S., and Strahler, A.H. (1994). Retrieval of surface BRDF from multiangle remotely sensed data. *Remote Sensing of Environment*, 50, 18–30.

Martonchik, J.V., Diner, D.J., Kahn, R., Verstraete, M.M., Pinty, B., Gordon, H.R., and Ackerman, T.P. (1998). Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1212–1227.

Minnaert, M. (1941). The reciprocity principle in lunar photometry. *Astrophysical Journal*, 93, 403–410.

Nilson, T., and Kuusk, A. (1989). A reflectance model for the homogeneous plant canopy and its inversion. *Remote Sensing of Environment*, 27, 157–167.

Otterman, J., And Weiss, G.H. (1984). Reflections from a field of randomly located vertically protrusions. *Applied Optics*, 23, 1931–1936.

Pickup, G., Chewings, V.H. and Pearce, G. (1995a). Procedures for correcting high resolution airborne video imagery. *International Journal of Remote Sensing*, 16(9), 1647–1662.

Pickup, G., Bastin, G.N., Chewings, V.H. and Pearce, G. (1995b). Correction and classification procedures for assessing rangeland vegetation cover with airborne video data. *Proc. 15th Biennial Workshop on Color Photography and Video Graphy in Resource Assessment*, 2–3 May 1995, USA.

Pinty, B., Verstraete, M.M. and Dickinson, R.E. (1990). A physical model of the bidirectional reflectance of vegetation canopies. 2. Inversion and validation. *Journal of Geophysical Research*, 95, 11767–11775.

Pinty, B., And Verstraete, M.M. (1991). Extracting information on surface properties from bidirectional reflectance measurements. *Journal of Geophysical Research*, 96, 2865–2874.

Qi, J., Moran, M.S., Cabot, F., and Dedieu, G. (1995). Normalization of Sun/View Angle Effects Using Spectral Albedo-based Vegetation Indices. *Remote Sensing of Environment* 52, 207–217.

Rahman, H., Pinty, B., and Verstraete, M.M. (1993). Coupled surface-atmosphere reflectance (CSAR) model. 2. Semiempirical surface model usable with NOAA Advanced Very High resolution Radiometer data. *Journal of Geophysical Research*, 98, 20791–20801.

Ross, J.K. (1981). *The radiation regime and architecture of plant stands*. Junk Publ., The Hague. 391pp.

Roujean, J.L., Latoy, M., and Deschamps, P.Y. (1992). A bidirectional reflectance model of the earth's surface for the correction of remote sensing data. *Journal of Geophysical Research*, 97, 20455–20468.

Royer, A., Vincent, P., and Bonn, F. (1985). Evaluation and correction of viewing angle effects on satellite measurements of bi-directional reflectance. *Photogrammetric Engineering and Remote Sensing*, 51, 1899–1914.

Shibayama, M., and Wiegand, C.L. (1985). View azimuth and zenith, and solar angle effects using spectral albedo-based vegetation indices. *Remote Sensing of Environment*, 18, 91–103.

Shibayama, M., and Wiegand, C.L. (1986). Diurnal patterns of bidirectional vegetation indices for wheat canopies. *International Journal of Remote Sensing,* 7, 233–246.

Staylor, W.F., and Suttles, J.T. (1986). Reflection and emission models for deserts derived from Nimbus-7 ERB scanner measurements. *Journal of Climate and Applied Meteorology*, 25, 196–202.

Strahler, A.H., and Jupp, D.L.B. (1991). Modeling bidirectional reflectance of forests and woodlands using boolean models and geometric optics. *Remote Sensing of Environment*, 34, 153–166.

Strahler, A.H., Wanner, W., Barker Schaaf, C., Li, X., Hu, B., Muller, J-P., Lewis, P., and Barnsley, M.J. (1996). *MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document Version 4.0. MODIS Product ID: MOD43*, 252pp.

Verstraete, M.M., Pinty, B., and Dickinson, R.E. (1990). A physical model of the bidirectional reflectance of vegetation canopies. 1. Theory. *Journal of Geophysical Research*, 95, 11755-11765.

Walthall, C.L., Norman, J.M., Welles, J.M., Campbell, G., and Blad, B.L. (1985). Simple equation to approximate the bidirectional reflectance from vegetation canopies and bare soil surfaces. *Applied Optics*, 24, 383–387.

Wanner, W., Li, X., and Strahler. A.H. (1995). On the derivation of kernels for kernel-driven models of bidirectional reflectance, *Journal of Geophysical Research*, 100, 21077–21089.

Appendix 3
Colour Systems and Sequences

Colour science is introduced in Volume 2A—Section 5. The following sub-sections elaborate on additional aspects of colour systems and sequences, namely:

- colour measurement (see Appendix A3.1);
- colour systems (see Appendix A3.2); and
- colour coordinates (see Appendix A3.3).

A3.1 Colour Measurement

An interesting aspect of human vision is that it is the mixture of wavelengths that are perceived and not the individual components. This means that a given same visual sensation could be produced by different physical stimuli. (Colours which appear to be identical but have different spectral composition are called metamers). This phenomenon is used as the basis for psychophysical colour matching experiments in which a colour is defined as a mixture of three primary colours. An infinite number of sets of primary colours may be defined subject to the condition that any one may not be produced by a mixture of the other two. The red, green and blue lights used in the additive system were selected to obtain the maximum range of mixture colours. It should be noted however that this range is still less than the full range of colours that can be perceived by humans (see Volume 2A—Section 5).

In colour matching experiments, a reference colour is defined by the proportions of the primary lights required to reproduce it. Due to the limited mixture range of the primaries however, some colours cannot be reproduced this way. These colours are matched by adding one or two of the primaries to the reference colour and defining the reference by subtracting the proportion of those primaries added from the mixture used to match it. These proportions can then be plotted on a chromaticity diagram.

To avoid negative terms in these definitions, three artificial primaries (that is imaginary colours) were determined by the CIE (Commission Internationale d'Eclairage) which are referred to as X, Y and Z (see Volume 2A—Section 5.1.2 and Appendix A3.3). While real primaries are used to perform the actual colour matching then the proportions are transformed algebraically to X, Y, and Z values for plotting on the standardised CIE diagram, this system has the advantage of allowing the measurement to be made with different sets of primaries.

Even in modern art, artists have used methods based on calculation, inasmuch as these elements, alongside those of a more personal and emotional nature, give balance and harmony to any work of art. (Max Bill)

Background image: The range of electromagnetic (EM) wavelengths that are visible to the human eye spans from 380 nm to 700 nm.

A3.2 Colour Systems

A3.2.1 RGB and HSI

The RGB (Red,Green,Blue) and HSI (Hue,Saturation,Intensity) models for colour specification were introduced in Volume 2A—Section 5.1.2. These models are related by their intensity axes. For the RGB cube, the intensity axis is defined by the line:

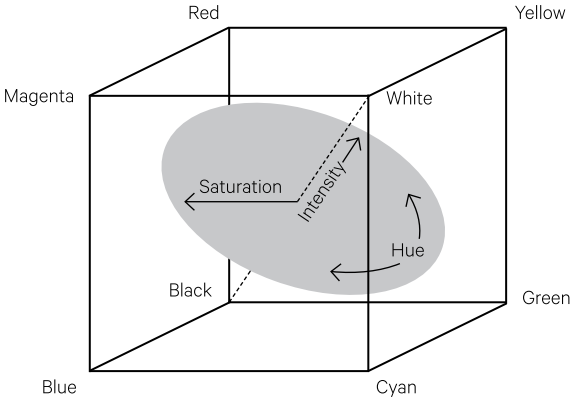
R = G = B

as illustrated in Figure A3.1a. Colours in the RGB colour space can also be defined as triangular coordinates on equal intensity planes across the cube (see Figure A3.1a). These coordinates can be normalised within each plane such that the values along each axis from the triangle centre to an apex vary from 0 to 1 (see Figure A3.1b).

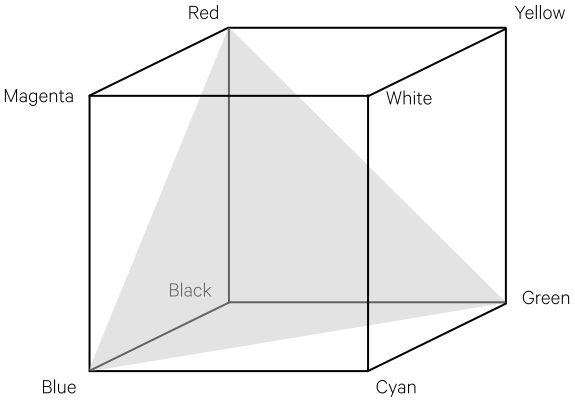
Figure A3.1 RGB colour coordinates

The colours formed by differing proportions of the additive primary colours, red green and blue, can be represented as a three-dimensional data space. The diagonal axis in this space, referred to as the intensity axis, represents the shades of grey from black to white that have equal proportions of each primary colour. The subtractive primaries occur at the diagonal 'corners' on each two-dimensional plane through two axes. An infinite number of planes that are orthogonal to the intensity axis can be defined. Colours on each plane have equal intensity, with saturation increasing away from the axis and hue varying around it.

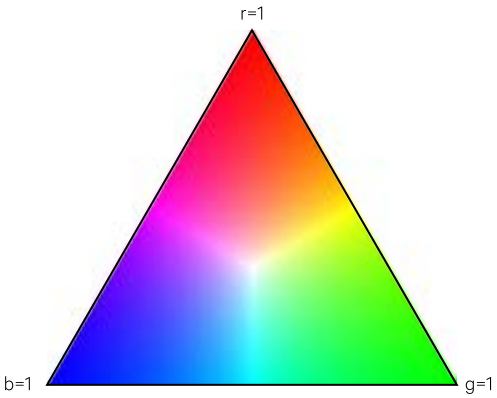
a. Orientation of RGB colour cube



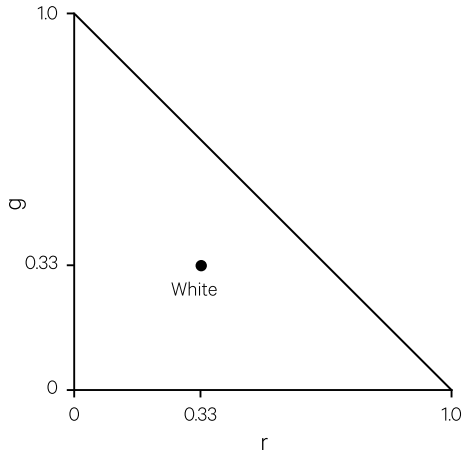
b. Definition of triangular planes in RGB cube



c. Normalised triangular coordinates



d. Normalised chromaticity coordinates. Since, for any colour r+g+b=1, only two components need to be known to deduce a third. For example at the white point, r=1/3, g=1/3, so b=1-r-g=1/3.



Source: a. Harrison and Jupp (1990) Figures 11 and 102

Cartesian coordinates in the RGB cube can be converted to normalised triangular coordinates by the equations:

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

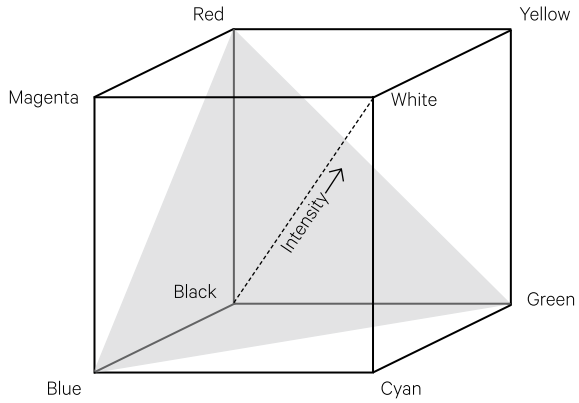
By definition, these normalised coordinates (r, g, b) sum to one, with the intensity of the plane being equal to the sum of the RGB values:

$$I = R + G + B$$

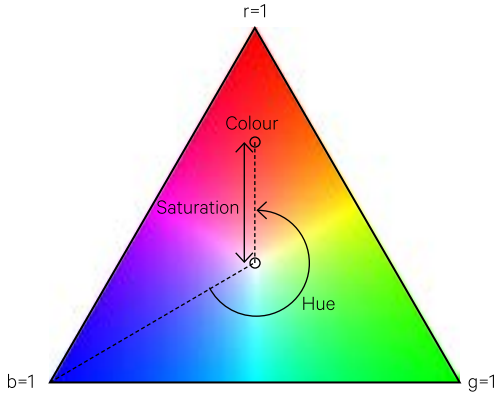
The HSI model defines colours on the equal intensity planes as coordinates of hue and saturation, with hue being measured as an angle around the plane and saturation as the radial distance from the centre of the plane (see Figure A3.2). Measurement of these parameters is detailed in Volume 2C—Section 9.3.2.

Figure A3.2 HSI coordinates

a. Intensity is measured along the diagonal axis between black and white.



b. Saturation is the linear distance from the intensity axis to the selected colour on an equal intensity plane. Hue is the angular distance from a primary colour (usually blue) to the colour on an intensity plane.



Source: Harrison and Jupp (1990) Figure 104

A3.2.2 XYZ and UVW

The rgb chromaticity coordinates were defined above as normalised proportions of R, G and B values. Since the normalised coordinates sum to one, only two are required to define a given colour—the third coordinate can be deduced from these two, for example:

$$b = 1 - r - g$$

The values of any two coordinates (usually r and g) can then be usefully represented as a chromaticity diagram as shown in Figure A3.3.

CIE have defined colour-matching functions for monochromatic stimuli of wavelengths 700 nm (R_o), 546.1 nm (G_o) and 435.8 nm (B_o), which can be used to match the spectral power distribution curve of any colour. These stimuli were defined to enclose the chromaticity diagram as illustrated in Figure A3.3.

For 2° standard observer data (that is, standard colour measurement data based on a 2° field of view) the CIE tristimulus values, X, Y and Z (see Volume 2A—Section 5.1.2), are related to R_o, G_o and B_o (Hunt, 1987) as:

$$X = 0.49 \times R_o + 0.31 \times G_o + 0.20 \times B_o$$

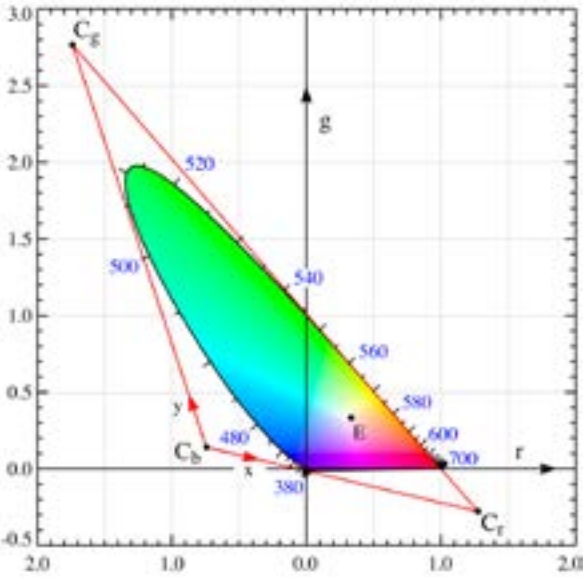
$$Y = 0.17697 \times R_o + 0.81240 \times G_o + 0.01063 \times B_o$$

$$Z = 0.0 \times R_o + 0.01 \times G_o + 0.99 \times B_o$$

For diagrammatic simplicity, the XYZ values are usually normalised in the same way as the RGB coordinates and expressed as xyz values. A chromaticity diagram of x and y can then be used to represent the colour gamut on Cartesian coordinates.

Figure A3.3 CIE rg chromaticity space

The triangle C_b-C_g-C_r is equivalent to the triangle (0,0)-(0,1)-(1,0) in the CIE xy chromaticity space.

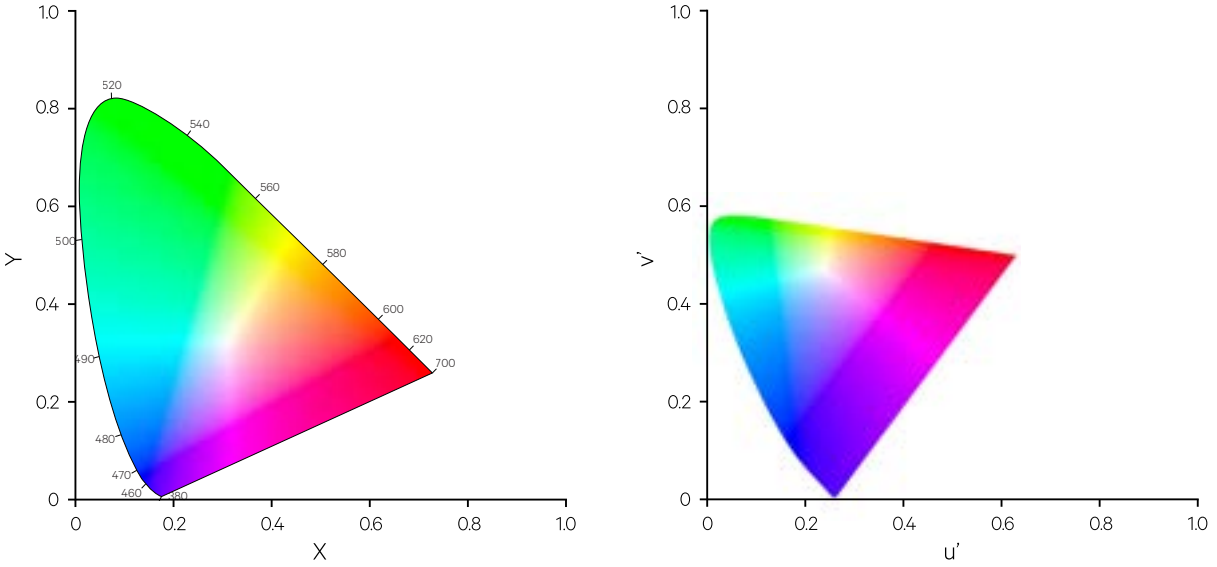


Retrieved from: *https://upload.wikimedia.org/wikipedia/commons/1/16/CIE1931_rgxy.png*

Figure A3.4 Perceptual colour differences

a. CIE 1931 xyz colour space

b. CIE 1976 uvw colour space. The linear distance between colours in this coordinate system can be used to indicate the perceptual difference between them.



Source: Colour space retrieved from *https://upload.wikimedia.org/wikipedia/commons/8/83/CIE_1976_UCS.png*

As illustrated in Figure A3.4a, this xyz space is not uniform in terms of perceived colour differences. MacAdam (1937) proposed a 'uniform chromaticity' triangle using the coordinates U, V and W (see Figure A3.4b). The normalised coordinates u and v of this system are related to the xy coordinates by:

$$u = 2x / (6y - x + 1.5)$$

$$v = 3y / (6y - x + 1.5)$$

Conversion from xy to uv coordinates can be simplified by use of a nomogram (Hunt, 1975). The uv chromaticity diagram approximates a uniform colour space, so allows relative perceptual differences between two colours to be computed as the Euclidean distance between the points that represent the colours on this diagram.

A3.3 Colour Coordinates

In many image processing systems, colours for image painting are specified as numeric values, such as in the range 0–511. Such numbers generally represent the proportion of the three additive primaries used in each colour. Using the standard CIE equations given in Appendix A3.2, these numbers can be transformed into uvw coordinates to show their distribution within the perceptible colour space. This conversion assumed that the proportions of each primary are directly related to the rgb coordinate scale, so do not account for the effect of device-specific gamuts.

To achieve maximum visual separation within a set of colours, the CIE uvw space needs to sampled evenly before transforming the sample coordinates back to their RGB components. Given the extensive spectral resolutions of many colour display devices, the resulting range of colours should be well-represented. The reproduction of the display colour on hardcopy devices, however, can still be challenging (see Volume 2A—Sections 5.1.3 and 6).

A3.4 References

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Hunt, R.W.G. (1975). *The Reproduction of Colour*. Fountain Press, London.

Hunt, R.W.G. (1987). *Measuring Colour*. Ellis-Horwood Ltd. Chichester, UK.

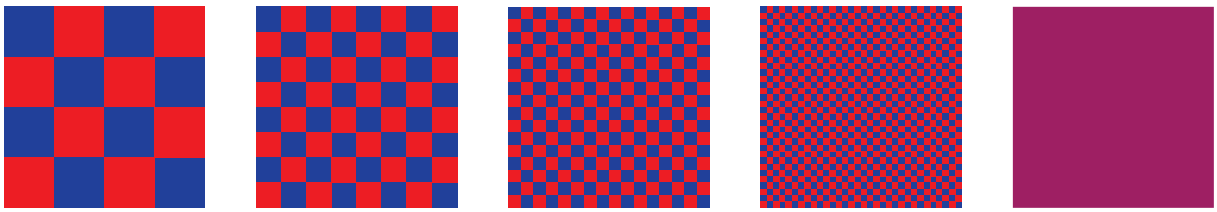
MacAdam, D.L. (1937). Projective Transformations of I.C.I. Color Specifications. *Journal of the Optical Society of America,* 27(8), 294–299. doi: *10.1364/JOSA.27.000294*

Appendix 4
Dithering

In computer graphics, to extend the perceived range of colours that is actually available from a limited colour palette on an output device, a technique known as 'dithering' can be used. This approach relies on the human eye blurring adjacent colours to perceive a mixed colour (see Figure A4.1). Dithering methods are comparable to halftone printing techniques and are commonly used to expand the range of colours represented by display hardware.

Figure A4.1 Dithering example

This sequence of image use just two colours, red and blue, but as the area of each colour is reduced, the human eye blurs them to appear as the single colour purple.

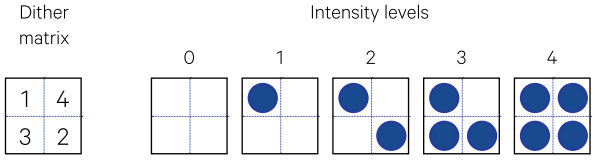


A4.1 Dither Matrices

One form of dithering relies on using a 'dither matrix'. The dithering technique used for inkjet printers, for example, is implemented by grouping together square or rectangular blocks of dots (called 'nibs'), then using the blended colour of each group of dots to represent a single pixel in the image. The size of a dither block can vary and affects the number of colours that can be represented. A non-square image pixel may be more closely represented using a rectangular dither matrix such as 3×4 dots.

Figure A4.2 Dithering sequence

For example, to print value 3 with an output value range 0–4 (such as in a 2×2 dither), dots are printed at positions 1, 2 and 3 in the dither matrix.



Source: Harrison and Jupp (1990) Figure 35

If we use a 2×2 dither block (or 'printer pixel') for each image pixel, we have increased the intensity variations for the pixel from 2 (one dot ON or OFF) to 5 (all 4 dots OFF, one dot ON, two dots ON, three dots ON or all 4 dots ON) as shown in Figure A4.2. The three primary inks allow $2^3 = 8$ different colour possibilities for a single dot. In a 2×2 dither we can represent $5^3 = 125$ different colours. Other dither sizes include:

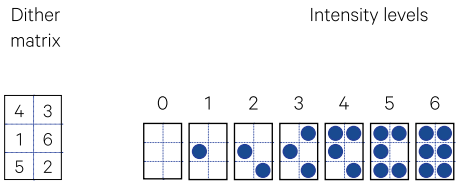
- 2×3 (7 intensity levels: $7^3 = 343$ colours);
- 3×3 (10 intensity levels: $10^3 = 1{,}000$ colours);
- 3×4 ($13^3 = 2{,}197$ colours); and
- 4×4 ($17^3 = 4{,}913$ colours).

For these colour ranges to be effectively used in a printed image however, the digital image must have a correspondingly diverse range of data values.

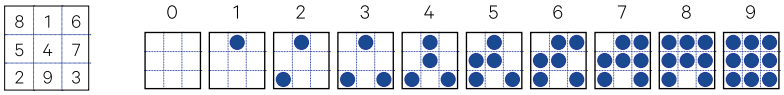
Background image: High resolution aerial image of coastal environment near Broome, WA, acquired on 27 July 2017. **Source:** © EagleView

Figure A4.3 Example dither matrices

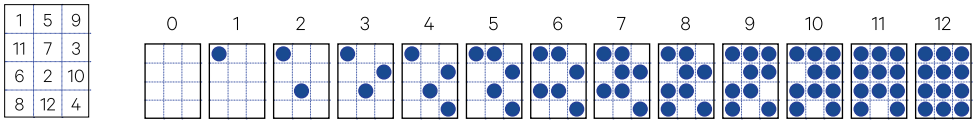
a. 2×3



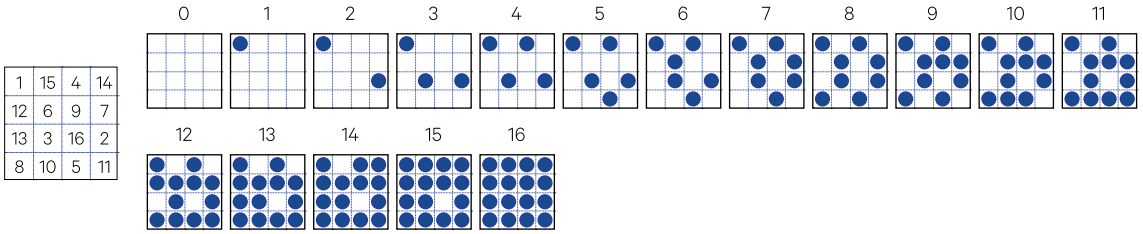
b. 3×3



c. 3×4



d. 4×4



Source: Harrison and Jupp (1990) Figure 36

The 'cost' of this increased colour range is the number of pixels that may be plotted across the page. With a 2×2 dither we are using two dots across the page and two dots down to represent one printed pixel. If the printer allows only 1,024 dots across the page, then using a 2×2 or 2×3 dither we can only print 512 pixels. Similarly, a 3×3 or 3×4 dither would only allow 1024/3 or 341 pixels across the page.

The order of intensity changes within a dither block are usually coded in matrix format (hence the term dither matrix). The values at each dot position in the matrix indicate the intensity at which that dot will become activated. Any other dots with values below this value are also used. For example, in Figure A4.2, to print intensity level 3 we use the dots at positions 1, 2 and 3 in the dither. Some other dither matrices that have been used in image processing software are given in Figure A4.3. As discussed, the dither size determines the intensity range of a printed image. To make best use of the available contrast, enhancement techniques (as described in Volume 2A—Section 4.2) are usually applied to image data before mapping to this range.

A4.2 Designed Order

The ordering within a dither matrix needs to be carefully selected to avoid geometric patterns (such as basket-weave, twill, or herringbone tweed) in areas of uniform intensity or low frequency blotchiness. Colour transitions in an image need to be smooth to give the impression of continuous variations in tones. One method for achieving this effect is to use dither matrices in which the rows and columns sum to approximately the same value. This criterion can only be achieved for square matrices but satisfactory results may still be obtained in other shapes.

The 4×4 dither matrix shown in Figure A4.3 is based on a pan-diagonal magic square or 'nasik' in which any pair of adjacent cells (horizontal, vertical or diagonal) do not contain sequential numbers. In this case all 32 four-element groups within the matrix (that is, four columns and rows, eight diagonals—including broken diagonals such as [3+5+12]+[10]) and sixteen 2×2 squares—all sum to 34 (Lippel and Kurland, 1971). This pattern avoids horizontal, vertical and diagonal line structures within the dither or across two adjacent dithers as illustrated in Figure A4.4.

The final result of applying dithering and inverse mapping to the complementary primaries results in the ability to plot apparently continuous tone imagery on inkjet technology. This operation is demonstrated in Figure A4.5.

Figure A4.4 Formation of a balanced dither matrix

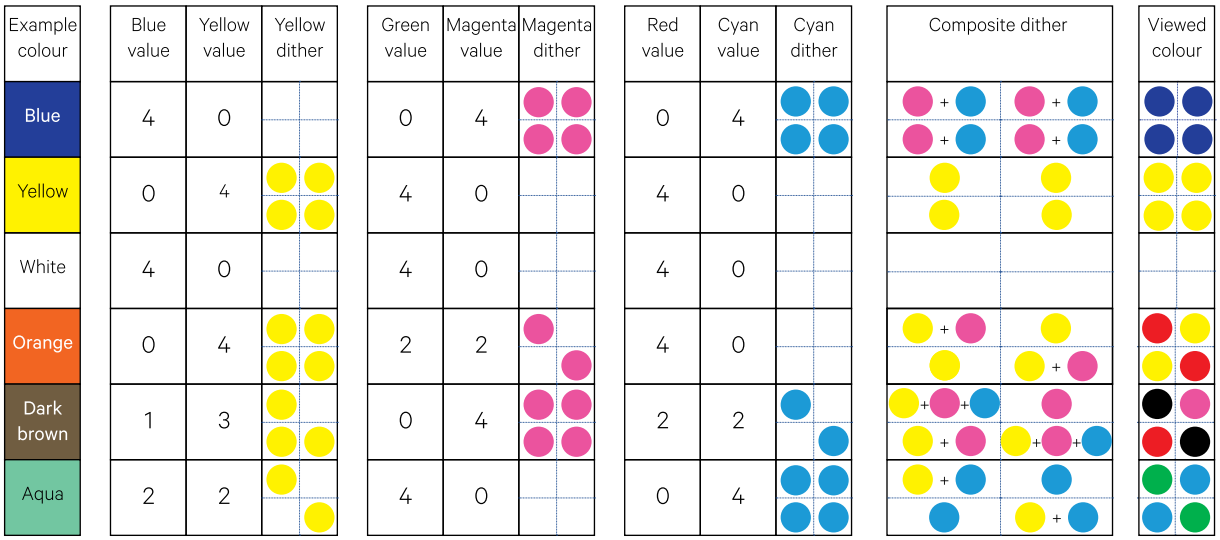
The 4×4 matrix above is based on a 'nasik': all groups of four adjacent cells, both within one dither and across adjacent dithers, sum to the same value.

16	2	13	3	16	2	13	3
5	11	8	10	5	11	8	10
4	14	1	15	4	14	1	15
9	7	12	6	9	7	12	6
16	2	13	3	16	2	13	3
5	11	8	10	5	11	8	10
4	14	1	15	4	14	1	15
9	7	12	6	9	7	12	6

Source: Harrison and Jupp (1990) Figure 37

Figure A4.5 Colour representation using dithering techniques with a 2×2 dither.

The inverse intensity values of three channels are assigned to the subtractive primaries then plotted using dither matrices.



A4.3 References

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Lippel, B., and Kurland, M. (1971). The effect of dither on luminance quantisation of pictures. *IEEE Transactions on Communication Technology*, 19, 879–888.

Appendix 5
Filters

The use of spatial filters is introduced in Volume 2C. The following sub-sections detail a range of filters that have been used effectively with EO imagery:

- averaging filters—used to smooth imagery (see Appendix A5.1);
- directional filters—for edge enhancement, including exposure, insolation and curvature (see Appendix A5.2); and
- derivature filters—commonly used for geophysical applications (see Appendix A5.3)

A5.1 Averaging Filters

Balanced average filters are formed from B-spline formulae to ensure a smooth graduation of weighting values (Jupp 1976). For example, a good 3×3 filter is generated from the function:

$$\begin{bmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{bmatrix}$$

as the outer product with its transpose. Larger filters are convolution products of the smaller ones (see Volume 2C—Section 3). For all sizes, the sum of all weighting values is used as a divisor to preserve scaling. A set of example filters is shown in Figure A5.1.

Background image: High resolution aerial image of soak wells at Maddington, WA, acquired on 1 November 2017. Source: © EagleView

Figure A5.1 Averaging filters

a. 3×3 filter with divisor 36

1	4	1
4	16	4
1	4	1

b. 5×5 filter with divisor 256

1	4	6	4	1
4	16	24	16	4
6	24	36	24	6
4	16	24	16	4
1	4	6	4	1

c. 7×7 filter with divisor 9,216

1	8	23	32	23	8	1
8	64	184	256	184	64	8
23	184	529	736	529	184	23
32	256	736	1024	736	256	32
23	184	529	736	529	184	23
8	64	184	256	184	64	8
1	8	23	32	23	8	1

d 9×9 filter with divisor 589,824

1	16	76	176	230	176	76	16	1
16	256	1216	2816	3680	2816	1216	256	16
76	1216	5776	13376	17480	13376	5776	1216	76
176	2816	13376	30976	40480	30976	13376	2816	176
230	3680	17480	40480	52900	40480	17480	3680	230
176	2816	13376	30976	40480	30976	13376	2816	176
76	1216	5776	13376	17480	13376	5776	1216	76
16	256	1216	2816	3680	2816	1216	256	16
1	16	76	176	230	176	76	16	1

e. 11×11 filter with divisor 58,982,400

1	32	237	832	1682	2112	1682	832	237	32	1
32	1024	7584	26624	53824	67584	53824	26624	7584	1024	32
237	7584	56169	197184	398634	500544	398634	197184	56169	7584	237
832	26624	197184	692224	1399424	1757184	1399424	692224	197184	26624	832
1682	53824	398634	1399424	2829124	3552384	2829124	1399424	398634	53824	1682
2112	67584	500544	1757184	3552384	4460544	3552384	1757184	500544	67584	2112
1682	53824	398634	1399424	2829124	3552384	2829124	1399424	398634	53824	1682
832	26624	197184	692224	1399424	1757184	1399424	692224	197184	26624	832
237	7584	56169	197184	398634	500544	398634	197184	56169	7584	237
32	1024	7584	26624	53824	67584	53824	26624	7584	1024	32
1	32	237	832	1682	2112	1682	832	237	32	1

A5.2 Directional Filters

Various edge enhancement filters are presented in Volume 2C. Some additional directional filters are given below.

Prewitt (1970) suggests a bi-Laplacian operator (a fourth derivative bidirectional filter: $\frac{\partial^4}{\partial x^2 \partial y^2}$) as:

$$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$$

This 3×3 implementation of the operator is not recommended as its own spatial frequency is quite abrupt and would cause 'ringing' in the processed image. A set of second derivative filters with directional properties is shown in Figure A5.2.

Various directionally selective filters have been used to detect edges. These generally rely on the eight compass directions (North, North-east, East, South-east, South, South-west, West, North-west), which can be easily aligned with the image grid. These filters can be applied in sets of eight with the maximum (absolute) filtered value being selected as the output value and the direction associated with the selected filter being the edge direction.

Frei and Chen (1977) defined a set of directional, edge-detecting filters with the weights shown in Figure A5.3, where a=√2 (listed as anti-clockwise from north).

The Sobel and Prewitt operators could similarly be implemented in these eight directions by using a value of '*a*' equal to 2 or 1 respectively. Another set of directional filter weights proposed by Prewitt (1970) are shown in Figure A5.4.

Kirsch (1971) defined a slightly different set of directional filters (see Figure A5.5).

Any of these filters could be used individually to enhance the edge information towards a particular direction. The exposure transformation (see Volume 2C and Appendix A5.2.1) implements directional filtering of imagery from a selected compass direction.

Figure A5.2 Derivative filters

-1	-1	-1
2	2	2
-1	-1	-1

-1	-1	2
-1	2	-1
2	-1	-1

-1	2	-1
-1	2	-1
-1	2	-1

2	-1	-1
-1	2	-1
-1	-1	2

Figure A5.3 Directional, edge-detecting filters

a=√2

1	a	1
0	0	0
-1	-a	-1

a	1	0
1	0	-1
0	-1	-a

1	0	-1
a	0	-a
1	0	-1

0	-1	-a
1	0	-1
a	1	0

-1	-a	-1
0	0	0
1	a	1

-a	-1	0
-1	0	1
0	1	a

-1	0	1
-a	0	a
-1	0	1

0	1	a
-1	0	1
-a	-1	0

Figure A5.4 Directional filters

1	1	1
1	-2	1
-1	-1	-1

1	1	1
1	-2	-1
1	-1	-1

1	1	-1
1	-2	-1
1	1	-1

1	-1	-1
1	-2	-1
1	1	1

-1	-1	-1
1	-2	1
1	1	1

-1	-1	1
-1	-2	1
1	1	1

-1	1	1
-1	-2	1
-1	1	1

1	1	1
-1	-2	1
-1	-1	1

Figure A5.5 Kirsch directional filters

5	5	5
-3	0	-3
-3	-3	-3

5	5	-3
5	0	-3
-3	-3	-3

5	-3	-3
5	0	-3
5	-3	-3

-3	-3	-3
5	0	-3
5	5	-3

-3	-3	-3
-3	0	-3
5	5	5

-3	-3	-3
-3	0	5
-3	5	5

-3	-3	5
-3	0	5
-3	-3	5

-3	5	5
-3	0	5
-3	-3	-3

A5.2.1 Exposure

As detailed in Volume 2C—Section 6.3, the exposure transformation uses two smoothing, differentiating filters to compute the exposure of each 'pixel' to a selected direction. The 3×3 smoothing derivative filters, with divisor 12 and offset 127, are defined as:

$$\frac{\partial}{\partial X} = \begin{bmatrix} -1 & 0 & 1 \\ -4 & 0 & 4 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\frac{\partial}{\partial Y} = \begin{bmatrix} 1 & 4 & 1 \\ 0 & 0 & 0 \\ -1 & -4 & -1 \end{bmatrix}$$

These are the product of a B-spline filter:

$$\begin{bmatrix} 1 & 4 & 1 \end{bmatrix}$$

and a central difference filter:

$$\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

A5.2.2 Relief shading

To produce an relief shading filter, two differential filters are used to compute the first derivatives in two directions: along-line (X) and down pixel columns (Y) (see Volume 2C). These filters are the product of a B-spline filter:

1	4	6	4	1

(divisor 16), with a central difference filter:

-1	-2	0	2	1

(divisor 8).

The 5×5 average filter shown in Figure A5.1b may be applied to the elevation channel before using differential filters.

$\frac{\partial}{\partial X}$ filter with divisor 128:

-1	-4	-6	-4	-1
-2	-8	-12	-8	-2
0	0	0	0	0
2	8	12	8	2
1	4	6	4	1

$\frac{\partial}{\partial Y}$ filter with divisor 128:

1	2	0	-2	-1
4	8	0	-8	-4
6	12	0	-12	-6
4	8	0	-8	-4
1	2	0	-2	-1

A5.2.3 Curvature

Three differential filters are used in the microBRIAN Curvature transformation to compute the second derivatives in three directions: along-line (*X*) and down pixel columns (*Y*) and along the diagonal (*XY*) (see Section 12.9). The *X* and *Y* filters are the product of a B-spline filter (divisor 96):

1	8	23	32	23	8	1

with a second derivative filter (divisor 4):

1	2	-1	-4	-1	2	1

The *XY* filter is formed as the outer product of the first derivative filter (divisor 4):

-1	-4	-5	0	5	4	1

with its transpose.

$\frac{\partial^2}{\partial X^2}$ filter with divisor 1536:

1	8	23	32	23	8	1
2	16	46	64	46	16	2
-1	-8	-23	-32	-23	-8	-1
-4	-32	-92	-128	-92	-32	-4
-1	-8	-23	-32	-23	-8	-1
2	16	46	64	46	16	2
1	8	23	32	23	8	1

$\frac{\partial^2}{\partial Y^2}$ filter with divisor 1536:

1	2	-1	-4	1	2	1
8	16	-8	-32	46	16	8
23	46	-23	-92	-23	46	23
32	64	-32	-128	-92	64	32
23	46	-23	-92	-23	46	23
8	16	-8	-32	46	16	8
1	2	-1	-4	23	2	1

$\frac{\partial^2}{\partial X \partial Y}$ filter with divisor 1024:

-1	-4	-5	0	5	4	1
-4	-16	-20	0	20	16	4
-5	-20	-25	0	25	20	5
0	0	0	0	0	0	0
5	20	25	0	-25	-20	-5
4	16	20	0	-20	-16	-4
1	4	5	0	-5	-4	-1

The Laplacian can be computed in the curvature transformation as the average of the second derivatives in *X* and *Y*. A Laplacian filter that is equivalent to this operation can be derived from the $\frac{\partial^2}{\partial X^2}$ and $\frac{\partial^2}{\partial Y^2}$ filters given above:

1	5	11	14	11	5	1
5	16	19	16	19	16	5
11	19	-23	-62	-23	19	11
14	16	-62	-128	-62	16	14
11	19	-23	-62	-23	19	11
5	16	19	16	19	16	5
1	5	11	14	11	5	1

A5.3 Geophysical Filters

For geophysical applications, Fuller (1966) reports a number of derivative and other filters (including second and fourth vertical derivations, residual and analytic continuation), which implement useful frequency filters in the spatial domain. Several derivatives (or 'residuals') of geomagnetic fields were proposed by Henderson and Zeitz (1949) to highlight magnetic anomalies. Their equations 10, 13 and 15 were represented as spatial filters by Fuller (1966).

Peters (1949) computed second and fourth derivatives of magnetic data. Second derivative equations for interpreting gravity data were reported by Elkins (1951). Rosenbach (1953) also proposed equations for computing the second vertical derivative from gravity data.

A5.4 References

Elkins, T.A. (1951). The second derivative method of gravity interpretation. *Geophysics,* 16, 29–50.

Frei, W., and Chen, C.C. (1977). Fast boundary detection: A generalisation and a new algorithm. *IEEE Transactions on Computers* C–26, 988–98.

Fuller, B.D. (1966). Two-dimensional frequency analysis and design of Grid Operators. In *Society of Exploration Geophysicists Mining Geophysics. Vol II. Theory*. (Ed: D.A. Hansen), p. 658–708.

Henderson, R.G., and Zeitz, I. (1949). The computation of second vertical derivatives of geomagnetic fields. *Geophysics* 14, 508–16.

Jupp, D.L. (1976). B-Splines for smoothing and differentiating data sequences. *Mathematical Geol*ogy, 8, 243–66.

Kirsch, R.A. (1971). Computer determination of the constituent structure of biological images. *Computers and Biomedical Research* 4, 315–28.

Peters, L.J. (1949). The direct approach to magnetic interpretation and its practical application. *Geophysics* 14, 290–319.

Prewitt, J.M.S. (1970*). Object Enhancement and Extraction. In Picture Processing and Psychopictorics*. (Eds: B.S. Lipkin and A. Rosenfeld.) Academic Press, New York. pp 75–149.

Rosenbach, O. (1953). A contribution to the computation of the 'second derivative' from gravity data. *Geophysics* 18, 894–909.

Appendix 6
Matrix Algebra

An excellent explanation of various mathematical analyses, including matrix algebra, is given in Sawyer (1955). Unlike most mathematical texts, this unique account was specifically written to interest people in mathematics and exposes various branches of the subject using lucid, commonplace examples and only assumes 'a confused recollection of School Certificate mathematics'.

Using material from this source, Appendix A6.1 describes the operation of matrices in terms of their geometric effects on image spaces. Some characteristics of matrix algebra which may be used for representing specific geometrical transformations are detailed in Appendices A6.2 and A6.3.

A6.1 Geometric Effect of Matrix Operations

The geometrical effect of the general matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

on the image data space can be understood using example points along the X and Y axes. In Figure A6.1, we use three equally spaced points on each axis: P, Q and R on the X axis and U, V and W on the Y axis. The matrix operation then creates two new axes X′, where the three X axis points have been moved to P′, Q′ and R′, and Y′, where the Y axis points are now located at U′, V′ and W′. Because this is a linear operation, straight lines remain straight and parallel lines remain parallel. For geometrically representable operations, the matrix coefficients can be simply determined as the transformed values of the points P (that is a,c at P′) and U (namely b,d at U′), that is:

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} X \\ Y \end{bmatrix}$$

For example in Figure A6.1 the original coordinates of P (1,0) are transformed to the values (a,c) at P′.

Image data axes may be reflected about any line L by the matrix coefficients:

$$\begin{bmatrix} \cos 2a & \sin 2a \\ \sin 2a & -\cos 2a \end{bmatrix}$$

where a is the angle (in degrees) between L and the X axis (see Figure A6.2). For:

$Y=X$
$a=45°$

then

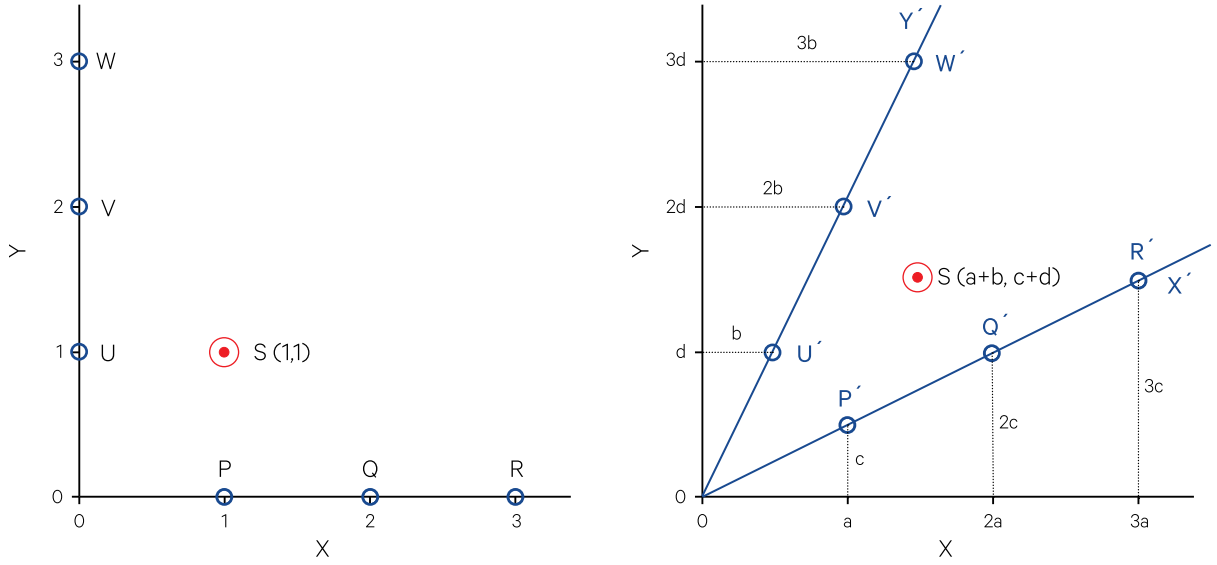
$2a=90°$
$\sin 2a=1$
$\cos 2a=0$

to produce the coefficients used in the example of Figure A6.3.

Background image: High resolution aerial image of aluminium works at Kwinana, WA, acquired on 11 January 2018. Source: © EagleView

Figure A6.1 Effect of matrix operation on data geometry

The transformation from X,Y to X´,Y´can be described by the matrix equations: X´= aX + bY and Y´= cX + dY.

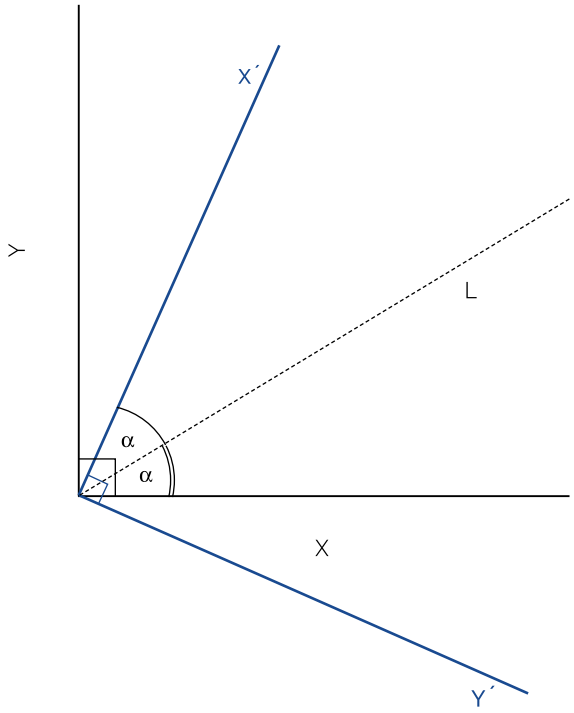


Source: Harrison and Jupp (1990) Figure 108

Figure A6.2 Reflecting axes

X´= cos 2αX + sin 2αY

Y´= sin 2αX – cos 2αY



Source: Harrison and Jupp (1990) Figure 109

Linear rescaling of image values (see Volume 2A and Figure A6.4) can be simply implemented by multiplying X and/or Y values by some constant value(s). This can be represented by the matrix:

$$\begin{bmatrix} m & 0 \\ 0 & n \end{bmatrix}$$

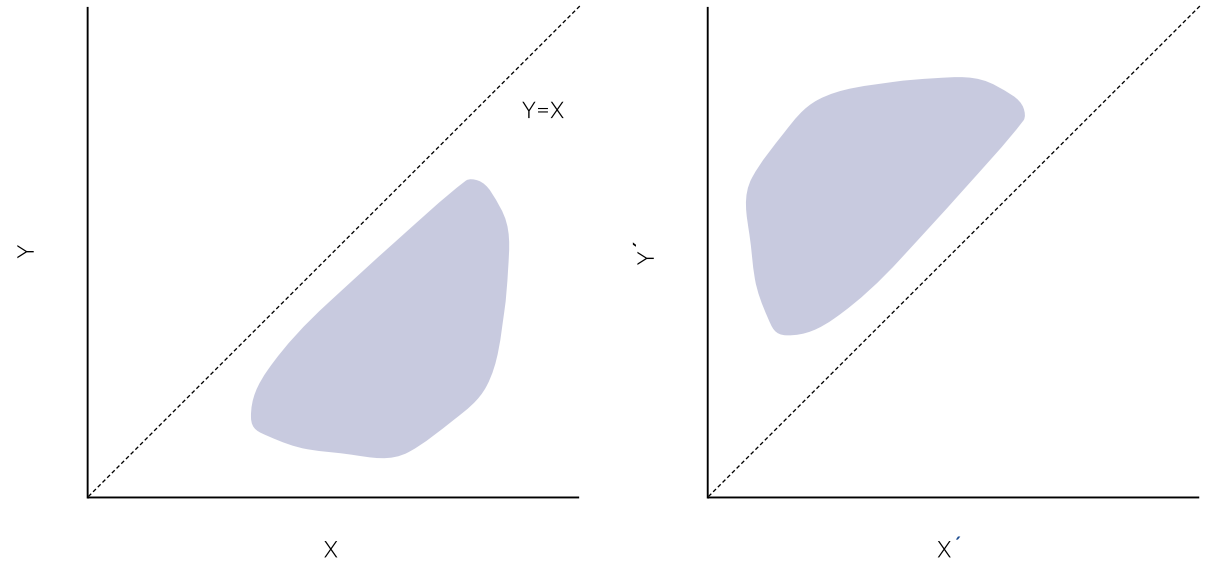
where, if m = n, then both X and Y channels are being rescaled by the same amount. A special case of channel rescaling involves inverting the channel values (that is, changing a positive increasing order to a negative decreasing order). This can be effected by using m and n values equal to -1. In image processing, an offset value would usually need to be applied to this result to preserve positive pixel values.

By three methods we may learn wisdom:
First, by reflection, which is noblest;
Second, by imitation, which is easiest; and
Third by experience, which is the bitterest.
(Confucius)

Figure A6.3 Reflecting axes about the line Y=X

a. Original axes

b. Reflected axes

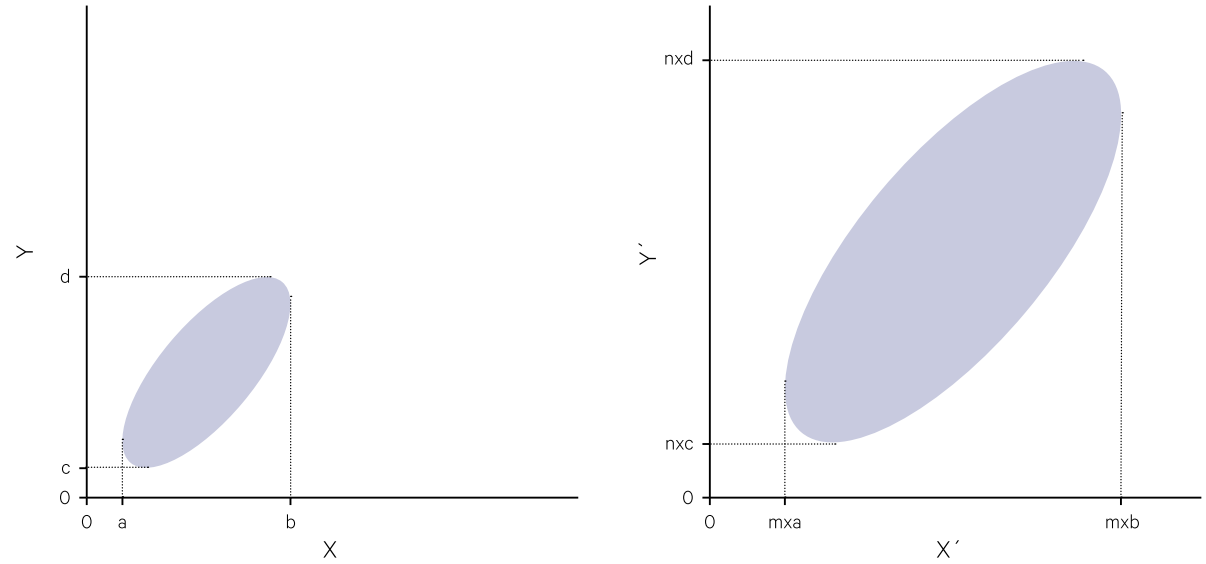


Source: Harrison and Jupp (1990) Figure 85

Figure A6.4 Rescaling axes

X´= mX

Y´= nY



Source: Harrison and Jupp (1990) Figure 110

Rotation of the image data space (actually a combination of skewing and scaling) can be represented by the matrix:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

where θ is defined as the clockwise angle between X and X´(see Figure A6.5). Some special cases of this operation are:

- θ =90°, cos θ =0 and sin θ=1, which produces a matrix of reflection about Y=X´with inverted scaling in X´; and
- θ =180°, cos θ =-1 and sin θ =0, so the rotation matrix simply inverts scaling on both the X´and Y´ axes.

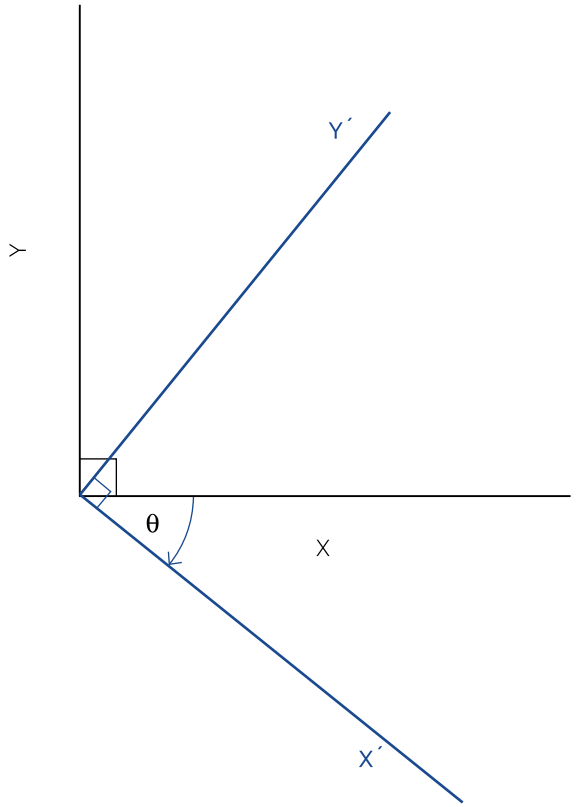
A skewing operation in the X direction as illustrated in Figure A6.6 can be represented as the matrix:

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

Figure A6.5 Rotating axes

X´ = cos θX – sin θY

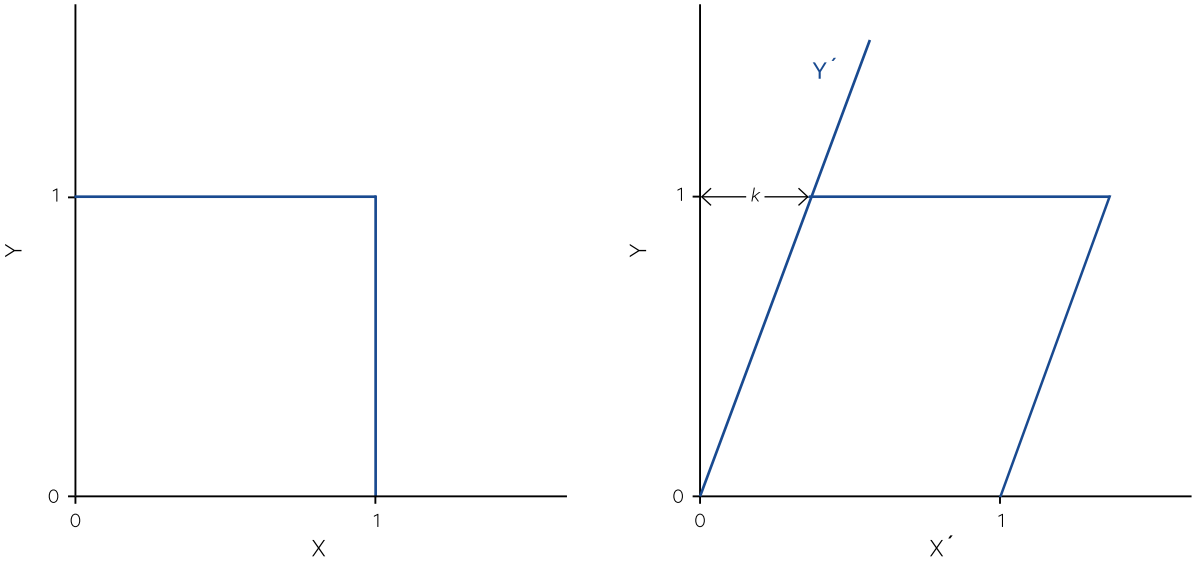
Y´ = sin θX + cos θY



Source: Harrison and Jupp (1990) Figure 111

Figure A6.6 Skewing axes

X´ = X + *k*Y



Source: Harrison and Jupp (1990) Figure 112

A6.2 Matrix Arithmetic and Properties

A Zero matrix (all elements equal to 0) is defined as:

$$\begin{bmatrix} 0 \end{bmatrix}$$

For a 2×2 matrix, this would be:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

A Unit or Identity matrix (all diagonal elements are 1 and off-diagonals equal 0) is represented as:

$$\begin{bmatrix} 1 \end{bmatrix}$$

For a 2×2 matrix, this would be:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

For two 2×2 matrices:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$B = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$$

A6.2.1 Trace

The Trace of a matrix equals the sum of its diagonal elements:

*Trace (**A**) = a+d*

A6.2.2 Matrix equality

A=B

only if:

a=p; and
b=q; and
c=r; and
d=s.

A6.2.3 Matrix addition

$$A + B = \begin{bmatrix} a+p & b+q \\ c+r & d+s \end{bmatrix} = B + A$$

A6.2.4 Scalar multiplication

For a constant *n*:

$$n \times A = \begin{bmatrix} n \times a & n \times b \\ n \times c & n \times d \end{bmatrix} = \begin{bmatrix} n & 0 \\ 0 & n \end{bmatrix} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

A6.2.5 Matrix multiplication

$$A \times B = \begin{bmatrix} a \times p + b \times r & a \times q + b \times s \\ c \times p + d \times r & c \times q + d \times r \end{bmatrix}$$

$$B \times A = \begin{bmatrix} p \times a + q \times c & p \times b + q \times d \\ r \times a + s \times c & r \times b + s \times d \end{bmatrix}$$

Note: matrix multiplication is not commutative: ***AB*** is not equal to ***BA***.

A6.2.6 Transpose of a matrix

$$A^{T} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

A6.2.7 Additive inverse of a matrix

$$A + (-A) = (-A) + A = 0$$

$$-A = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$$

A6.2.8 Multiplicative inverse of a matrix

$$A \times A^{-1} = A^{-1} \times A = 1$$

If (*ad–bc*) is not equal to 0, then ***A*** is called an orthogonal matrix and

$$A^{-1} = \begin{bmatrix} \frac{d}{a \times d - b \times c} & \frac{-b}{a \times d - b \times c} \\ \frac{-c}{a \times d - b \times c} & \frac{a}{a \times d - b \times c} \end{bmatrix} = \frac{d}{a \times d - b \times c} \times \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If *ad–bc* equals 0, ***A*** is called a singular matrix and A^{-1} does not exist.

(*ad–bc*)

is called the determinant of ***A****, and*

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

is called the adjoint of ***A***.

A6.2.9 Determinant of a matrix

In Figure A6.1, we applied a matrix to a set of points located on the X and Y axes. The effect of moving points P, U and S changed the shape of a unit square to that of a parallelogram. The dimensions of this parallelogram are described in Figure A6.7. The area of this parallelogram equals the determinant of ***A*** and can be computed as:

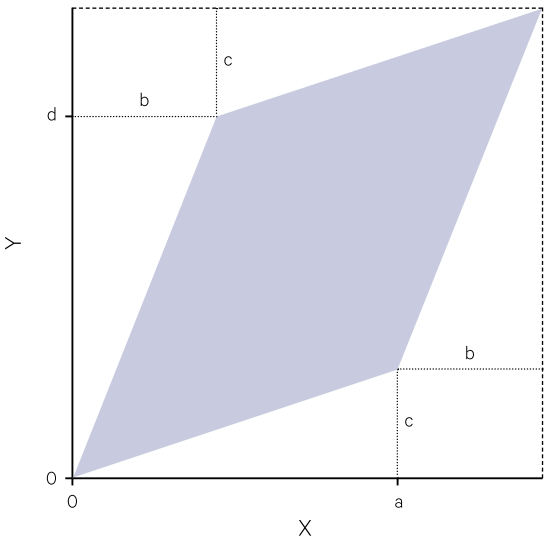
$$(a+b)\times(c+d)-2(b\times c)-(a\times c)-(b\times d)$$
$$=a\times d-b\times c$$

Since this Figure originated from a unit square (area 1), the determinant represents the change in area caused by the matrix operation. The determinant is usually written as:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Figure A6.7 Effect of matrix operation on unit area

The area of a parallelogram resulting from a matrix operation is described by the determinant of the matrix. In this case the determinant equals ad–bc. For a 2×2 matrix this is defined as the ratio in which the matrix changes areas (see Figure A6.1).



Source: Harrison and Jupp (1990) Figure 113

For a 2×2 matrix, the is defined as the ratio in which the matrix changes area. The determinant of a 3×3 matrix is defined as:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = a\begin{vmatrix} e & f \\ h & k \end{vmatrix} - b\begin{vmatrix} d & f \\ g & k \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$
$$=a\times e\times k+b\times f\times g+c\times d\times h$$
$$-a\times h\times f-b\times d\times k-c\times e\times g$$

and represents the constant ratio in which the matrix changes volume in a three-dimensional data space. The concept can be used with higher-dimensional dataset although the geometric process cannot be visualised.

The determinant is a useful summary of many matrix characteristics such as:

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\begin{vmatrix} a+k\times c & b+k\times d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \times \begin{vmatrix} p & q \\ r & s \end{vmatrix} = \begin{vmatrix} a\times p+b\times r & a\times q+b\times s \\ c\times p+d\times r & c\times q+d\times s \end{vmatrix}$$

These operations can be used to determine the overall data scaling action of a sequence of matrix transformations on a dataset. The reflection, rotation and skewing operations described above do not alter the area of the dataset so have determinants equal to one in magnitude. The reflection transformations, however, all have negative determinants which characteristically indicates that there is a reversal of axes involved with such matrices. For example, the skewing operation can be represented as:

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+k\times c & b+k\times d \\ c & d \end{bmatrix}$$

and

$$\begin{vmatrix} a+k\times c & b+k\times d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

or

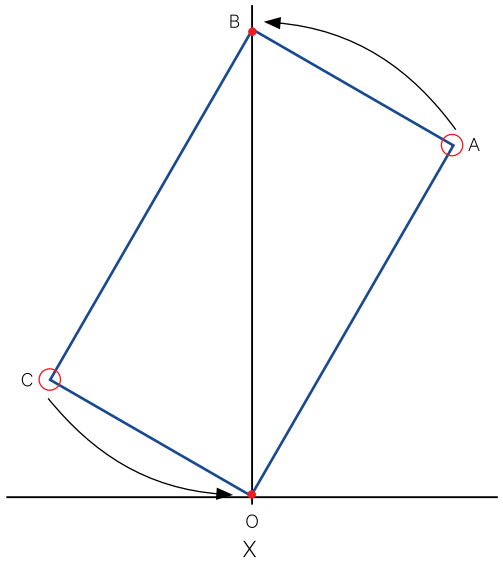
$$\begin{vmatrix} 1 & k \\ 0 & 1 \end{vmatrix} = 1\times 1-k\times 0=1$$

so does not change area.

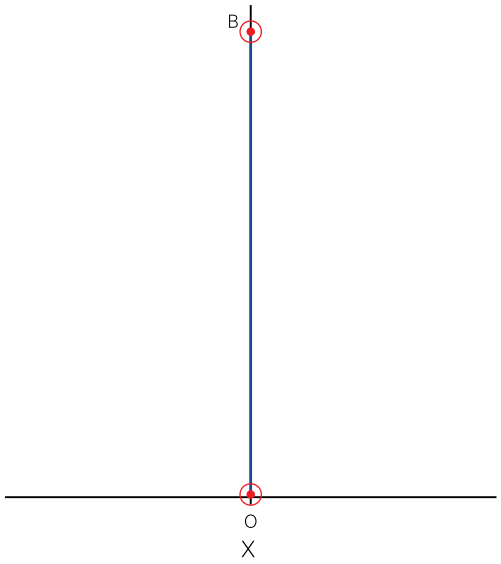
Figure A6.8 Effect of singular matrix operation

A matrix whose determinant equals zero represents an operation which collapses a two-dimensional area to one dimension (see Figure A6.7). In this example, point C is collapsed onto O and point A is collapsed onto point B.

a. Original area



b. Two dimensional area collapsed to one dimensional line



Source: Harrison and Jupp (1990) Figure 114

Finally a special property of some matrices is that their determinant value is zero. Such matrices are called singular and operate by collapsing areas in the data space into lines (of zero area) as illustrated in Figure A6.8. In image processing, this matrix operation loses a dimension so effectively 'destroys' data.

A6.2.10 Characteristic equation

The Characteristic Equation of any square matrix A is:

$$\left|A-k\times 1\right|=0$$

or

$$\left|\begin{bmatrix} a & b \\ c & d \end{bmatrix}-\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}\right|=0$$

or

$$\begin{vmatrix} a-k & b \\ c & d-k \end{vmatrix}=0$$

Thus

$$\left(a-k\right)\times\left(d-k\right)-b\times c=0$$

A6.2.11 Eigenvalues of a matrix

The eigenvalues of a matrix are defined as the solutions of its characteristic equation. The eigenvalues k of matrix $\boldsymbol{A}$ may be written as:

$$A\times\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}=k\times\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

where

r_1 and r_2 are points representing a vector (called the eigenvector) from the origin of the original coordinate system, then

$$A\times\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}-k\times 1\times\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}=0$$

or

$$\left(A-k\times 1\right)\times\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}=0$$

that is

$$\begin{bmatrix} a-k & b \\ c & d-k \end{bmatrix}\times\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solutions of this equation are:

$$\left(a-k\right)\times r_1+b\times r_2=0$$

and

$$c\times r_1+\left(d-k\right)\times r_2=0$$

A6.2.12 Eigenvectors of a matrix

The eigenvectors r of matrix $\boldsymbol{A}$ are defined by:

$$\left(\boldsymbol{A}-k_1\times 1\right)\times \boldsymbol{r}=0$$

and

$$\left(\boldsymbol{A}-k_2\times 1\right)\times \boldsymbol{r}=0$$

where

k_1 and k_2 are the roots of the characteristic equation (also called the eigenvalues).

A6.3 Derivation of the Principal Components Transformation

In an image $\boldsymbol{X}$, the covariance of channels j and k is defined as:

$$v_{ij}=\frac{\sum_{k=1}^{n}\left(x_{ik}-\overline{x}_i\right)\times\left(x_{jk}-\overline{x}_j\right)}{n}$$

where

n is the total number of pixels in the image
x_{ik} is value for pixel k in channel i;
$\overline{x}_i$ is the mean of pixel values in channel i;
x_{jk} is value for pixel k in channel j; and
$\overline{x}_j$ is the mean of pixel values in channel j.

see Volume 2A—Section 8.1.4). The covariance matrix, $\boldsymbol{C}_x$, can be defined as:

$$\frac{1}{n}\times\sum_{i=1,n}\left(\boldsymbol{x}_i-\overline{\boldsymbol{x}}\right)\times\left(\boldsymbol{x}_i-\overline{\boldsymbol{x}}\right)^T$$

or

$$E\left\{\left(\boldsymbol{x}-\overline{\boldsymbol{x}}\right)\times\left(\boldsymbol{x}-\overline{\boldsymbol{x}}\right)^T\right\}$$

As defined in Volume 2C, a PCA (Principal Components Analysis) image has a diagonal covariance matrix, that is all off-diagonal elements are zero. The PCA image $\boldsymbol{Y}$ can be computed from the image $\boldsymbol{X}$ by a rotation transformation P where:

$$\boldsymbol{y}=\boldsymbol{P}\times\left(\boldsymbol{x}_i-\overline{\boldsymbol{x}}\right)$$

The covariance matrix of a PCA image $\boldsymbol{Y}$ can be similarly defined as:

$$\boldsymbol{C}_x=E\left\{\left(\boldsymbol{y}-\overline{\boldsymbol{y}}\right)\times\left(\boldsymbol{y}-\overline{\boldsymbol{y}}\right)^T\right\}$$

By substituting the equation relating $\boldsymbol{Y}$ and $\boldsymbol{X}$ we get:

$$\begin{aligned}\boldsymbol{C}_y&=E\left\{\left(\boldsymbol{P}\times\boldsymbol{x}-\boldsymbol{P}\times\overline{\boldsymbol{x}}\right)\times\left(\boldsymbol{P}\times\boldsymbol{x}-\boldsymbol{P}\times\overline{\boldsymbol{x}}\right)^T\right\}\\&=\boldsymbol{P}\times E\left\{\left(\boldsymbol{x}-\overline{\boldsymbol{x}}\right)\times\left(\boldsymbol{x}-\overline{\boldsymbol{x}}\right)^T\right\}\times\boldsymbol{P}^T\\&=\boldsymbol{P}\times\boldsymbol{C}_x\times\boldsymbol{P}^T\end{aligned}$$

Since $\boldsymbol{C}_y$ is diagonal, its diagonal elements can be defined as the eigenvalues of $\boldsymbol{C}_x$ and the rows in the orthogonal matrix P are the eigenvectors of $\boldsymbol{C}_x$ (Gonzalez and Wintz, 1977). The trace of $\boldsymbol{C}_x$ also equals the trace of $\boldsymbol{C}_y$.

The eigenvalues k can be computed as defined in Section A6.2.10 above, that is:

$$\left|\boldsymbol{C}_x-k\times 1\right|=0$$

For the example image from Volume 2A (see Figure A6.9), the covariance matrix is:

$$\boldsymbol{C}_x=\begin{bmatrix}2.9380 & -0.3840 & 2.3000\\ -0.3840 & 4.7200 & -0.0144\\ 2.3000 & -0.0144 & 2.2656\end{bmatrix}$$

Figure A6.9 Example image channels

Channel A

	1	2	3	3	4
	3	3	5	6	3
	2	3	3	6	5
	5	4	2	2	1
pixel 1, line 5 →	7	6	3	2	1

Channel B

	3	6	7	8	7
	2	4	7	5	6
	1	1	3	4	5
	1	1	2	5	5
pixel 1, line 5 →	2	3	4	6	7

Channel C

	1	3	4	3	3
	2	3	4	5	4
	2	2	3	6	5
	4	4	2	1	2
pixel 1, line 5 →	6	5	2	1	1

Source: Harrison and Jupp (1990) Figures 41

The eigenvalues of the matrix are derived as:

$$\begin{vmatrix} 2.938-k & -0.384 & 2.3 \\ -0.384 & 4.72-k & -0.0144 \\ 2.3 & -0.0144 & 2.2656-k \end{vmatrix} = 0$$

or

$$\begin{aligned} &(2.938-k)\times(4.72-k)\times(2.2656-k) \\ &+(-0.384\times-0.0144\times2.3)+(2.3\times-0.384\times-0.0144) \\ &-((2.938-k)\times-0.0144^2)-(-0.384^2\times(2.2656-k)) \\ &-(2.3\times(4.72-k)\times2.3) \end{aligned}$$

which gives eigenvalues of 5.1, 4.5 and 0.3. The covariance matrix of $\boldsymbol{Y}$ can then be written as:

$$\begin{bmatrix} 5.1 & 0 & 0 \\ 0 & 4.5 & 0 \\ 0 & 0 & 0.3 \end{bmatrix}$$

The trace of this matrix (5.1+4.5+0.3) equals the trace of $\boldsymbol{C}_x$ (2.938+4.72+ 2.2656). The percentage variation accounted for by each principal component (PC) can be determined using the equation:

$$\frac{\sigma^2\times100}{trace}$$

The variance and percent variation results are shown in Table A6.1.

A noise-to-signal ratio (NSR) percentage summarises the variance of a component *j* relative to those components above it:

$$NSR_j\% = 100\times\sqrt{\frac{\sum_{i=j+1,n}\left(\frac{\sigma_i^2}{n-j}\right)}{\sum_{i=1,j}\left(\frac{\sigma_i^2}{j}\right)}}$$

where

n is the number of components;
j is the selected component for NSR; and
i is the component number.

Table A6.1 Principal component percentage variation

PC	Variance	Percentage
1	5.1	51.8
2	4.5	45.5
3	0.3	2.7
Total	9.9	100

This can be used to identify the number of data dimensions, or rank, of the image matrix. Typically useful dimensions have a high variance value and 'noise' components have low variance so the number of data dimensions can be determined from the component with the minimum noise to signal ratio value. (This ratio is generally observed as a trend of decreasing values through the data components then marginally increasing in the noise components.) For example, in an eight channel image we have eight components *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h* in which the variances of (*a*, *b*, *c*, *d*, *e*) are much greater than the variances of components *f*, *g* and *h*. The NSR % at component *e* can be computed as:

$$NSR_e\% = 100\times\sqrt{\frac{\left(\sigma_f^2+\sigma_g^2+\sigma_h^2\right)/3}{\left(\sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_d^2+\sigma_e^2\right)/5}}$$

Since the variance of *f* is much smaller than that of *e*, this percentage will be larger than the NSR % at component *f*:

$$NSR_e\% = 100\times\sqrt{\frac{\left(\sigma_g^2+\sigma_h^2\right)/2}{\left(\sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_d^2+\sigma_e^2+\sigma_f^2\right)/6}}$$

If the variances of *f*, *g* and *h* are all small relative to components *a* to *e*, the NSR % at *g* and *h* will be marginally larger than *f* since a similarly small numerator is being divided by a slightly smaller denominator (the additional noise channel reduces the arithmetic mean).

The eigenvectors of $\boldsymbol{C}_x$ can also be computed to form the rows of the transformation matrix $\boldsymbol{P}$:

$$\begin{bmatrix} p_{ii} & p_{ij} & p_{ik} \\ p_{ji} & p_{jj} & p_{jk} \\ p_{ki} & p_{kj} & p_{kk} \end{bmatrix}$$

using the equations:

$$\begin{aligned} &(\boldsymbol{C}_x-k_1\times1)\times p_{.i}=0 \\ &(\boldsymbol{C}_x-k_2\times1)\times p_{.j}=0 \\ &(\boldsymbol{C}_x-k_3\times1)\times p_{.k}=0 \end{aligned}$$

For k_1 this expands to the three simultaneous equations:

$$\begin{aligned} &(2.938-5.1)\times p_{ii}-0.384\times p_{ji}+2.3\times p_{ki}=0 \\ &0.384\times p_{ii}(4.72-5.1)\times p_{ji}-0.0144\times p_{ki}=0 \\ &2.3\times p_{ii}-0.0144\times p_{ji}+(2.2656-5.1)\times p_{ki}=0 \end{aligned}$$

Similarly, two more sets of three equations can be generated by substituting the values of k_2 and k_3. Since matrix $\boldsymbol{A}$ must be orthogonal ($\boldsymbol{A}^T = \boldsymbol{A}^{-1}$), the eigenvectors are also normalised such that:

$$p_{ii}^{\ 2} + p_{ji}^{\ 2} + p_{ki}^{\ 2} = 1$$
$$p_{ij}^{\ 2} + p_{jj}^{\ 2} + p_{kj}^{\ 2} = 1$$
$$p_{ik}^{\ 2} + p_{jk}^{\ 2} + p_{kk}^{\ 2} = 1$$

For our example, the final transformation matrix is computed as:

$$\begin{bmatrix} 0.6306 & -5.879 & 0.5068 \\ 0.4153 & 0.8072 & 0.4195 \\ 0.6556 & 0.0541 & -0.7531 \end{bmatrix}$$

As detailed in Volume 2C, this matrix can now be used to rotate the data space of image $\boldsymbol{X}$ to the uncorrelated PCA image space. This is implemented by the equations:

$$PC1 = 0.6306 \times \text{ch a} - 0.5879 \times \text{ch b} + 0.5068 \times \text{ch c}$$
$$PC2 = 0.4153 \times \text{ch a} + 0.8072 \times \text{ch b} + 0.4195 \times \text{ch c}$$
$$PC3 = 0.6556 \times \text{ch a} + 0.0541 \times \text{ch b} - 0.7531 \times \text{ch c}$$

The determinant of this matrix can be computed using the equation given in Appendix A6.2.9:

$$\begin{aligned} &\left(0.6306 \times 0.8072 \times -0.7531\right) + \left(-0.5879 \times 0.4195 \times 0.6556\right) \\ &+ \left(0.5068 \times 0.4153 \times 0.0541\right) - \left(0.6306 \times 0.0541 \times 0.4195\right) \\ &+ \left(-0.5879 \times 0.4153 \times 0.7531\right) - \left(0.5068 \times 0.8072 \times 0.6556\right) \\ &= -1 \end{aligned}$$

which indicates that there is a reversal of axes involved with the transformation but no change in the 'area' occupied by the data.

The degree of correlation of each channel i with each principal component j can also be determined using the equation (Short, 1982):

$$p_{ij} \times \frac{\sqrt{k_j}}{\sigma_{ii}}$$

to give the normalised matrix:

$$\begin{bmatrix} 0.8307 & -0.6109 & 0.7604 \\ 0.5136 & 0.7875 & 0.5909 \\ 0.2096 & 0.0197 & -0.2742 \end{bmatrix}$$

The elements of this matrix indicate the degree of correlation between the original channels and each PC. For example, channels *a* and *c* are strongly correlated with PC1 (0.8307 and 0.7604 respectively) while channel *b* is negatively correlated with PC1 (-0.6109) and more strongly correlated with PC2 (0.7875).

A6.4 References

Gonzalez, R.C., and Wintz, P. (1977). *Digital Image Processing*. Addison-Wesley Pub. Co. Reading, Mass.

Harrison, B.A., and Jupp, D.L.B. (1990). *Introduction to Image Processing. Part TWO of the microBRIAN Resource Manual*. CSIRO, Melbourne. 256pp.

Sawyer, W.W. (1955). *Prelude to Mathematics*. Penguin Books Ltd. London.

Short, N.M. (1982). *The Landsat Tutorial Workbook. Basics of Satellite Remote Sensing*. NASA Reference Publication 1078.

Appendix 7
Predictive Error and Variance

A7.1 Predictive Error for Least Squares

The predictive error (or Predictive Residual Error Sum of Squares, PRESS) described below is the one defined by Allen (1974) and is a variation on the 'Jacknife' statistic discussed in Tuckey (1967).

The Least Squares (LS) situation for fitting transformation models by LS can be written by defining the error at Ground Control Point (GCP) '*i*' as:

$$e_i = \underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha} - z_i$$

where $\underset{\sim}{\varphi}_i$ is the vector of 'monomials' evaluated at the *i*'th GCP (x_i, y_i) in the 'from' coordinate system, $\underset{\sim}{\alpha}$ is the vector of coefficients and z_i is the coordinate (x or y) being predicted in the 'to' coordinate system (see Volume 2B).

If there are *M* GCPs then the complete expression for the residuals at the GCP points is:

$$\underset{\sim}{e} = \underset{\sim}{z} - A\underset{\sim}{\alpha}$$

The LS solution vector $\underset{\sim}{\alpha}^*$ minimises the sum of squared residuals or vector 'norm':

$$\underset{\alpha}{Min} \left\| \underset{\sim}{e} \right\|^2 = \underset{\sim}{e}^T \underset{\sim}{e} = \sum_{i=1}^{M} e_i^2$$

The solution is well known and can be written in this notation as:

$$\underset{\sim}{\alpha}^* = \left(A^T A\right)^{-1} A^T \underset{\sim}{z}$$

$$\underset{\sim}{e}^* = \underset{\sim}{z} - A\underset{\sim}{\alpha}^*$$

$$= \left[I - A\left(A^T A\right)^{-1} A^T \right] \underset{\sim}{z}$$

The consequent estimate for the error called the Residual Mean Square error (RMS) error is:

$$RMS = \frac{1}{M} \sum_{i=1}^{M} \left| e_i^* \right|^2$$

$$= \frac{1}{M} RSS$$

where RSS is termed the Residual Sum of Squares, which is not normalised by the number of observations *M*.

For reasons that will become clearer in the following we will denote the matrix:

$$A\left(A^T A\right)^{-1} A^T = H$$

$$= \left[h_{ij} \right]$$

$$h_{ij} = \underset{\sim}{\varphi}_i^T \left(A^T A\right)^{-1} \underset{\sim}{\varphi}_j$$

The (symmetric) matrix ***H*** is a special matrix in that powers of ***H*** simply result in ***H*** again as is the case with (***I-H***). It follows that the RSS can be written:

$$RSS = \underset{\sim}{z}^T (I - H) \underset{\sim}{z}$$

$$= \underset{\sim}{z}^T \underset{\sim}{z} - \underset{\sim}{z}^T H \underset{\sim}{z}$$

$$= TSS - MSS$$

which simply expresses the 'analysis of variance' associated with LS of the total Sum of Squares (TSS) into the sum of Model Sum of Squares (MSS) and Residual sum of Squares (RSS).

The 'Predictive' error is developed by noting that the errors at the GCPs after the model is fitted are too small as estimates for the actual statistical error. This comes about as they are used to fit the model. The consequent bias in the RMS is more than the normal allowance for the model degrees of freedom to obtain the estimated variance (*EVAR*) as in:

$$EVAR = \frac{1}{M - p} \sum_{i=1}^{M} \left| e_i^* \right|^2$$

(where *p* is the order of the model) will allow.

As an alternative and unbiased estimate of the error, if each GCP is removed from the model fit in turn and the error between the point removed and its prediction by the fitted model without its influence is used in place of the LS error at that point then the Predictive Error (PRESS) is obtained.

Background image: High resolution aerial image of Perth airport, WA, acquired on 3 May 2017. **Source:** © EagleView

The method is really a variation of retaining some points as 'test points' and some as 'training points'. However, the existence of the test points is always a problem in such a scheme in that people wish to use all of the information they have to train and to test. The Predictive Error is a good compromise.

Let us write for the predictive error at the *i*'th GCP:

$$e_i^{(i)} = z_i - \underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^{(i)}$$

where $\underset{\sim}{\alpha}^{(i)}$ is the set of model parameters obtained if the point z_i is not used to fit the model. Then:

$$\text{PRESS} = \frac{1}{M}\sum_{i=1}^{M}\left|e_i^{(i)}\right|^2$$

It may seem that this quantity involves *M* LS solutions and is quite messy. However, it turns out that the result is immediately available following the LS solution since:

$$\text{PRESS} = \frac{1}{M}\sum_{i=1}^{M} w_i \left|e_i^*\right|^2$$

$$w_i = \frac{1}{\left(1-h_{ii}\right)^2}$$

where e_i^* is the full LS residual obtained above and h_{ii} is the *i*'th diagonal element of the matrix ***H*** above.

The proof of this result can be found in Golub *et al.* (1979) but is easily and usefully derived as follows:

- noting that we can write the components of the LS solution as:

$$A^T A = \sum_{i=1}^{M} \underset{\sim}{\varphi}_i \underset{\sim}{\varphi}_i^T$$

$$A^T \underset{\sim}{z} = \sum_{i=1}^{M} z_i \underset{\sim}{\varphi}_i$$

- it follows that we can express the *i*'th component of the PRESS as:

$$e_i^{(i)} = z_i - \underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^{(i)}$$

where:

$$\underset{\sim}{\alpha}^{(i)} = \left(A^T A - \underset{\sim}{\varphi}_i \underset{\sim}{\varphi}_i^T\right)^{-1}\left(A^T \underset{\sim}{z} - z_i \underset{\sim}{\varphi}_i\right)$$

- with the help of the well-known Sherman-Morrison formula (Sherman and Morrison, 1949) for (in this case) a non-singular and symmetric matrix, ***H***:

$$(H + \lambda \underset{\sim}{a}\underset{\sim}{a}^T)^{-1} = H^{-1} - \frac{\lambda (H^{-1}\underset{\sim}{a})(H^{-1}\underset{\sim}{a})^T}{1+\lambda \underset{\sim}{a}^T H^{-1}\underset{\sim}{a}}$$

- it follows with a little manipulation that:

$$\underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^{(i)} = \left(1+\frac{h_{ii}}{1-h_{ii}}\right)\left(\underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^* - h_{ii} z_i\right)$$

$$= \frac{\underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^* - h_{ii} z_i}{1-h_{ii}}$$

so that

$$e_i^{(i)} = z_i - \underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^{(i)}$$

$$= z_i - \frac{\underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^* - h_{ii} z_i}{1-h_{ii}}$$

$$= \frac{z_i - \underset{\sim}{\varphi}_i^T \underset{\sim}{\alpha}^*}{1-h_{ii}}$$

$$= \frac{1}{1-h_{ii}} e_i^*$$

QED

A7.2 Generalised Cross-Validation (GCV)

Golub *et al.* (1979) favoured the Generalised Cross-Validation or GCV statistic over PRESS because they say that PRESS is not 'rotation invariant'. Perhaps this is not a vital characteristic of a statistic to have as I believe PRESS is an excellent tool that is sensitive to the distribution of the GCPs. This is the main reason for its use.

However, for completeness and because GCV is used to resolve 'ties' in the modelling described here, the GCV needs to be derived as well.

We have seen above that PRESS can be written as:

$$PRESS = \frac{1}{M}\sum_{i=1}^{M} w_i \left| e_i^* \right|^2$$

$$w_i = \frac{1}{\left(1-h_{ii}\right)^2}$$

The GCV obtained by writing:

$$GCV = \frac{\frac{1}{M}\sum_{i=1}^{M}\left| e_i^* \right|^2}{\left(1-\frac{1}{M}\sum_{i=1}^{M} h_{ii}\right)^2}$$

It has a similar nature to PRESS but does not separately weight the points. This can be a disadvantage when the location of sensitive points and their stability to outlier data is being assessed as well as simply fitting data.

Both PRESS and GCV have the property that as the model order increases the fit becomes less stable and the statistics (as expressed by how close $(1-h_{ii})$ is to zero) increase the errors. As the model order increases, RMS and RSS will both continue to decrease. PRESS and GCV are therefore commonly used to choose the 'best' model order as described in this report.

[NOTE: The microBRIAN program Model used PRESS to help select the 'best' order of model to use. The SIEVER model (see Appendix 8 below) also has a calculation for the value of w_i to help identify points where errors will have greatest effect on the LS solution. These points can be the best to use if GCPs are accurate but also the worst to use if GCPs are inaccurate.]

A7.3 Singular Value Decomposition

The Singular Value Decomposition (SVD; Lanczos, 1958) provides a convenient method to solve the LS problem and also derive PRESS or GCV.

The SVD of the M row by N column matrix $\boldsymbol{A}$ is the (unique) decomposition of the matrix into factors such that:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are ortho-normal and $\boldsymbol{S}$ is a diagonal matrix such that:

$$\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}_M$$

$$\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{I}_p$$

$$\boldsymbol{S} = diag[s_1, s_2, ..., s_p]$$

$$s_1 \geq s_2 \geq ... \geq s_p \geq 0$$

The $p \times p$ matrix $\boldsymbol{V}$ is the eigenvector matrix for $\boldsymbol{A}^T\boldsymbol{A}$ and the squares of the singular values (s_j) are the eigenvalues. The $M \times p$ matrix $\boldsymbol{U}$ consists of the first p columns of the eigenvector matrix for $\boldsymbol{A}\boldsymbol{A}^T$.

If the matrix has Rank 'q' ($<p$) then:

$$s_q > 0$$

$$s_{q+1} = ... = s_p = 0$$

Also, if $\underset{\sim}{u}_j$ is the j'th column of $\boldsymbol{U}$ and $\underset{\sim}{v}_j$ is the j'th column of $\boldsymbol{V}$ then the SVD can also be written as:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T = \sum_{j=1}^{p} s_j \underset{\sim}{u}_j \underset{\sim}{v}_j^T$$

The famous theorem of Eckert and Young (1936) concerning the approximation of a matrix $\boldsymbol{A}$ by a matrix $\boldsymbol{B}$ of lower rank is provided with a solution by the SVD in that the best approximation to $\boldsymbol{A}$ by a matrix $\boldsymbol{B}$ of rank $q<p$ in the sense of minimum square norm is simply:

$$\boldsymbol{B} = \sum_{j=1}^{q} s_j \underset{\sim}{u}_j \underset{\sim}{v}_j^T$$

$$\left|\left| \boldsymbol{A}-\boldsymbol{B} \right|\right|^2 = \sum_{i=1}^{M}\sum_{j=1}^{p}\left| a_{ij} - b_{ij} \right|^2 = \sum_{j=q+1}^{p} s_j^2$$

The SVD can be used to define a 'generalised inverse' for the matrix $\boldsymbol{A}$ such that:

$$\boldsymbol{A}^+ = \boldsymbol{V}\boldsymbol{S}^+\boldsymbol{U}^T$$

$$s_j^+ = \begin{cases} \dfrac{1}{s_j} & if\ s_j > 0 \\ 0 & if\ s_j = 0 \end{cases}$$

The relationship with LS is that when the matrix $\boldsymbol{A}$ is the LS matrix above:

$$\underset{\sim}{\alpha}^{*} = A^{+}\underset{\sim}{z}$$
$$\underset{\sim}{e}^{*} = (I - AA^{+})\underset{\sim}{z}$$
$$H = AA^{+} = UU^{T}$$

where in this case the columns of the matrix $\boldsymbol{U}$ are reduced to those corresponding to non-zero singular values. It follows that h_{ii} may be easily derived if the SVD is used to solve the LS problem since:

$$h_{ii} = \sum_{j=1}^{p} u_{ij}^{2}$$

$$w_{i} = \frac{1}{\left(1 - \sum_{j=1}^{p} u_{ij}^{2}\right)^{2}}$$

The SVD is also used to check that LS solutions are 'well posed' in that singular values are either zero or not 'close' to zero. Near-zero singular values correspond to unstable parameters. Various methods exist to modify the LS equations (often called 'ridge' regression) to provide more stable (but biased) solutions. These will not be pursued here.

A7.4 Predictive Variance

Predictive Variance (PV) is a related but different idea of the variation away from control. In this case, the resulting equation expresses how the model may vary at places away from the control due to the variation we know is in our estimate of the model parameters.

We can consider that the data vector values $\underset{\sim}{z}$ may have a variance, which we can estimate from the residual errors or (better) the PRESS. We will assume (although mainly because it is rare to have enough information to do otherwise—the expressions can be derived under different knowledge) that the error model is:

$$\underset{\sim}{z} = \underset{\sim}{z}' + \delta\underset{\sim}{z}$$
$$E(\delta\underset{\sim}{z}) = \underset{\sim}{0}$$
$$Var(\delta\underset{\sim}{z}) = E(\delta\underset{\sim}{z}\delta\underset{\sim}{z}^{T}) = \sigma^{2}I$$

Even accepting that the LS model predicted data values are an estimate for $\underset{\sim}{z}'$ there will be an expected variation around the model parameters ($\delta\underset{\sim}{\alpha}$) with mean zero and variance:

$$\delta\underset{\sim}{\alpha} = \left(A^{T}A\right)^{-1}A^{T}\delta\underset{\sim}{z}$$
$$E(\delta\underset{\sim}{\alpha}) = \underset{\sim}{0}$$
$$Var(\delta\underset{\sim}{\alpha}) = E(\delta\underset{\sim}{\alpha}\delta\underset{\sim}{\alpha}^{T}) = \sigma^{2}\left(A^{T}A\right)^{-1}$$

The variation of the model in the 'to' coordinate system at a point predicted from the point (x,y) in the 'from' system is therefore:

$$\delta z(x,y) = \underset{\sim}{\varphi}(x,y)^{T}\delta\underset{\sim}{\alpha}$$
$$E(\delta z) = 0$$
$$Var(\delta z) = \sigma^{2}\underset{\sim}{\varphi}(x,y)^{T}\left(A^{T}A\right)^{-1}\underset{\sim}{\varphi}(x,y)$$
$$= \sigma^{2}h(x,y)$$

The function $h(x,y)$ is related to the Predictive Error above since at the GCP points:

$$h(x_{i},y_{i}) = h_{ii}$$

The function $h(x,y)$ may be plotted over the 'from' space coordinate system to indicate how variations may increase away from control. To allow for the possibility of higher order models being present this is often done using PRESS for the data variance and a higher order of model than the one used for the fitting. Areas of high predictive variation are ones where more control should be obtained if possible.

A7.5 References

Allen, D.M. (1974). The relationship between variable selection and data augmentation and a method of prediction. *Technometrics*, 16, 125–127.

Eckert, C., and Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211–218.

Golub, G.H., Heath, M., and Wahba, G. (1979). Generalised cross validation as a method for choosing a good ridge parameter. *Technometrics*, 21, 215–223.

Lanczos, C. (1958). Linear systems in self-adjoint form. *American Mathematical Monthly*, 65, 665–679.

Sherman, J.. and Morrison, W.J. (1949). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *Annals of Mathematical Statistics*, 21, 124–127.

Tuckey, J.W. (1967). Discussion of Anscome (1967). *Journal of the Royal Statistical Society B*, 29, 47–48.

Appendix 8
'Sieving' Outliers

A8.1 Basic Assumptions and Framework

This Appendix explains the basis for an approach to identifying outliers in spatial datasets used for EO image registration (see Volume 2B). Typically such spatial datasets are either a pair of EO images to be coregistered, or an EO image and its corresponding map data, but may also involve more than two datasets. In the microBRIAN image processing system (Harrison and Jupp, 1992), the program SIEVER implemented this approach and the following sub-sections describe it in terms of the functionality of the SIEVER model.

The context of this approach is that there exists an image or a number of images (perhaps overlapping frames or the same area at different times) and a map base or coordinate system within which they will be referenced. Clearly identified points (GCP or Ground Control Points) are selected pairwise between the spatial datasets as accurately as possible. This may be done manually or by computer (such using a correlation program). In the case of the identification of positions from a map base it is generally necessary to manually identify the GCP (see Volume 2B).

For this approach, it is necessary that each identified feature have a unique identifier so that a given feature can be located in all of the images to be coregistered, or to the map base frame. This may be accomplished simply by associating a separate list of GCP with each image or map base and associating each record in the file with a unique spatial feature. In this way, if a feature is not present in one dataset, its location in the associated file is empty. If this or a similar method of identifying features across multiple images is available then the GCP that are common to all of a group of images can be identified quickly.

This method also assumes that major geometric distortions have been 'nominally' removed from EO imagery, that is, panoramic distortion and other artefacts created by the geometry of the EO sensor viewing perspective and acquisition method have been taken into account in the coordinates of the image GCP. The remaining geometric differences between the spatial datasets should then be accounted for (or well described) by an 'affine' transformation. An affine transformation is linear such that if one image has coordinates (x,y) and the other (x',y') then the relationship between the coordinates has the form:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_{01} \\ a_{02} \end{bmatrix}$$

This six parameter transformation can be thought of as some combination of shift of origin, rotation, skew and possibly a separate 'stretch' in the x and y directions (see Appendix 6 above and Volume 2B).

A8.2 Fitting Transformation Models

For any pair of spatial datasets there can be transformations from each coordinate system to the other. If all of the transformations have been determined then any feature in one dataset can be located in each of the others. In the case where all of the transformations can be represented by a simple affine transformation then if the transformation is exactly known in one direction it can be inverted to obtain the transformation in the other direction.

However, transformations are never known exactly. Normally, the GCP are used to estimate the coefficients by least squares or some other measure of goodness of fit between the original coordinates and those predicted by the transformation. The GCP are normally subject to error or 'noise' and in this process any outliers present will play a disturbing role.

Background image: High resolution aerial image of sluice dam at Karratha, WA, acquired on 19 July 2018. Source: © EagleView

There will always be some error or noise level that limits the accuracy of location of the GCPs and therefore also limit the accuracy of the estimated coefficients of a transformation. However, 'outliers' are points that are inaccurate well beyond this base of error or 'noise'.

When the transformation is being estimated it is usual to suppose that errors in GCP are only in those of the 'TO' side of the transformation. That is in the affine transformation above we are calling the (x,y) data the FROM side of the transformation (or the coordinates the transformation acts on) and the (x',y') data the TO side of the transformation (or the coordinates produced by the action of the transformation). In this case, assuming the errors are uncorrelated, it is possible to separate the fit to the x and y components of the TO side as:

$$x'_i = a_{10} + a_{11}x_i + a_{12}y_i + \varepsilon_i$$

and

$$y'_i = a_{20} + a_{21}x_i + a_{22}y_i + \eta_i$$

where ε_i and η_i and noise variances and the coefficients can be efficiently found by two separate least squares solutions.

However, the errors usually occur in both the FROM and the TO coordinate systems—a fact that is utilised when the inverse transformation is estimated—where the errors are then assumed to be only in the FROM side of the transformation. In the case of an affine transformation it is possible to solve for the coefficients assuming errors in both variables—but only with a simplified error model. This solution is a by-product of the work described here and will be described in passing.

The more significant issue, however, relates to outliers. In this case, if the outlier is in a GCP that is on the FROM side of the system then it is difficult if not impossible to actually identify the outlier. Its presence is often only indicated by a very poor fit to the data.

Moreover, if the outlier is on the TO side of the system then because least squares balances all errors it is likely the fit to the outlier will be better than to other data.

The situation where the outliers are strongly fitted by the model—and hence very difficult to identify as outliers—is closely related to a concept called 'predictive error' (see Appendix 7 above). The predictive error is the error at a point between the data and the model prediction based on all data *except* the point in question (see Volume 2B—Section 4). Conversely, it measures the sensitivity of the model to the data at the point or the degree of control that the data point exerts on the model. The sensitivity can be expressed as a 'predictive error multiplier' which will be large if the model is very sensitive to the presence of a data point.

High predictive error multipliers are both good and bad. A high value indicates a very important point in the modelling—or one that exerts a lot of influence over the model. However, such points are also the ones that introduce the greatest effects of error and when they are an outlier the effect can be very great but the actual error at the point is usually small—since the model fits the outlier in preference to other points!

The SIEVER algorithm implements three analyses:

- solves for the affine model without assuming the errors are all on one of the FROM and TO sides of the transformation model;
- estimates the predictive errors of the residuals as one means of identifying outliers; and
- plots the squares of the errors (which would have a Chi-square distribution if they were from a normal population or errors) against the Gamma distribution in what is called a Q-Q (or Quantile-Quantile) plot (see Volume 2B—Section 4). This also allows errors that are too large to be explained as simply large but still possible error values to be identified.

A8.3 The SIEVER Model

The SIEVER model assumes there is a set of N_v spatial datasets with M common GCP as identified by their individual GCP files. These GCP sets can be represented as a matrix with M rows and $2N_v$ columns:

$$A = \left[\underset{\sim}{x}_1, \underset{\sim}{y}_1, \underset{\sim}{x}_2, \underset{\sim}{y}_2, \ldots \underset{\sim}{x}_N, \underset{\sim}{y}_N \right]$$

(where the curled underscore indicates a column vector), which is the starting point for the analysis.

Because we have assumed that nominal geometric transformations have been carried out on the data to remove the major image distortions (see Volume 2B—Sections 2 and 3) and that the relationships between the images can (in the absense of noise) be modelled as affine transformations, the situation is one of the Generalised Linear Model (GLM) or Factor Analysis. That is, we are assuming that there is an underlying coordinate system (μ,ν) such that for each of the image coordinate systems we have:

$$x_{jk} = a_{j0} + a_{j1}\mu_k + a_{j2}\nu_k + \varepsilon$$

and

$$y_{jk} = b_{j0} + b_{j1}\mu_k + b_{j2}\nu_k + \eta$$

where (ε,η) are the errors in the coordinates and (μ_k,ν_k) are the M common (and transformed) coordinates.

We are going to assume for convenience that the images coordinates have had the means extracted from them so that the data have 'zero mean'. The underlying coordinates will also be assumed to have 'zero mean'. In this case it is possible to express the form of the underlying transformation as:

$$\begin{bmatrix} x_{jk} \\ y_{jk} \end{bmatrix} = \begin{bmatrix} a_{j1} & a_{j2} \\ b_{j1} & b_{j2} \end{bmatrix} \begin{bmatrix} \mu_k \\ \nu_k \end{bmatrix} + \begin{bmatrix} \varepsilon \\ \eta \end{bmatrix}$$

$$= B_j \begin{bmatrix} \mu_k \\ \nu_k \end{bmatrix} + \begin{bmatrix} \varepsilon \\ \eta \end{bmatrix}$$

It follows that we could express the matrix A in the form:

$$A = \begin{bmatrix} \mu_1 & \nu_1 \\ \mu_2 & \nu_2 \\ \dots & \dots \\ \mu_M & \nu_M \end{bmatrix} \begin{bmatrix} B_1^T & B_2^T & \dots & B_N^T \end{bmatrix} + E$$

where E is the matrix of random errors or 'noise'. That is, The underlying model is one where A has basic rank 2 or A is the product of a matrix of size $M\times 2$ and one of size $2\times 2N_v$.

A8.4 Solving for Underlying Model and Errors

To separate the underlying model from the errors we will use the theorem of Eckart and Young (1936) which states that the best approximation to a matrix by one of lower rank can be conveniently found using the Singular Value Transformation (SVD) of the matrix. The SVD (introduced by Lancos, 1958) provides an eigenvalue decomposition for an arbitrary matrix. That is, if A is any $M\times N$ matrix then there exist matrices U ($M\times N$), S ($N\times N$ and diagonal) and V ($N\times N$) that satisfy the conditions:

$$S = diag\left[s_1, s_2, \dots, s_N \right]$$

$$s_1 \geq s_2 \geq \dots \geq s_N \geq 0$$

$$U^T U = I_N$$

$$V^T V = V V^T = I_M$$

$$A = USV^T$$

Since, using these rules we have that:

$$A^T A = VS^2 V^T$$

it follows that the squares of the singular values are the eigenvalues of $A^T A$ and the matrix V is the matrix of its eigenvectors. An alternative way to write the SVD is in terms of the column vectors of U and V in the form:

$$A = \sum_{j=1}^{N} s_j \underset{\sim}{u}_j \underset{\sim}{v}_j^T$$

The Eckert-Young theorem can be stated in the way used here in the following form. If B is the best approximation to A by a matrix of rank $p<N$ where 'best' means in the sense of minimising the sums of squares of differences between all elements of the matrices then:

$$A = B + E$$

$$B = \sum_{j=1}^{p} s_j \underset{\sim}{u}_j \underset{\sim}{v}_j^T$$

$$E = \sum_{j=p+1}^{N} s_j \underset{\sim}{u}_j \underset{\sim}{v}_j^T$$

$$\left\| A - B \right\|^2 = \sum_{i=1}^{M} \sum_{j=1}^{N} (a_{ij} - b_{ij})^2$$

$$= \left\| E \right\|^2 = \sum_{j=p+1}^{N} s_j^2$$

As with many cases of least squares there is an Analysis of Variance (AOV) for this generalised system in the form:

$$TSS = MSS + RSS$$

$$\sum_{i=1}^{M} \sum_{j=1}^{N} a_{ij}^2 = \sum_{i=1}^{M} \sum_{j=1}^{N} b_{ij}^2 + \sum_{i=1}^{M} \sum_{j=1}^{N} e_{ij}^2$$

$$\left\| A \right\|^2 = \left\| B \right\|^2 + \left\| E \right\|^2$$

or

$$\sum_{k=1}^{N} s_k^2 = \sum_{k=1}^{p} s_k^2 + \sum_{k=p+1}^{N} s_k^2$$

where TSS stands for the Total Sum of Squares, MSS the Model Sum of Squares and RSS the Residual Sum of Squares.

Of special interest to us is the AOV for a single row of each matrix. Let the *i*'th row of a matrix be denoted by $\underset{\sim}{r}_i(\boldsymbol{A})$. It is a row vector that will, in this case, have *N* components. Then:

$$TSS_i = MSS_i + RSS_i$$

$$\left\|\underset{\sim}{r}_i(A)\right\|^2 = \left\|\underset{\sim}{r}_i(B)\right\|^2 + \left\|\underset{\sim}{r}_i(E)\right\|^2$$

or

$$\sum_{k=1}^{N} s_k^2 u_{ik}^2 = \sum_{k=1}^{p} s_k^2 u_{ik}^2 + \sum_{k=p+1}^{N} s_k^2 u_{ik}^2$$

From this (or directly) it can be seen that

$$\left\|\underset{\sim}{r}_i(A)\right\|^2 = \left\|\underset{\sim}{r}_i(US)\right\|^2$$

It is easy to see that in our specific case, where p=2 and N as used above is actually $2N_v$ then the corresponding solution to the SIEVER model is:

$$\begin{bmatrix} \mu_1 & \nu_1 \\ \mu_2 & \nu_2 \\ \ldots & \ldots \\ \mu_M & \nu_M \end{bmatrix} = \begin{bmatrix} s_1 \underset{\sim}{u}_1 & s_2 \underset{\sim}{u}_2 \end{bmatrix}$$

$$\begin{bmatrix} B_1^T & B_2^T & \ldots & B_N^T \end{bmatrix} = \begin{bmatrix} \underset{\sim}{v}_1^T \\ \underset{\sim}{v}_2^T \end{bmatrix}$$

Because of the orthonormal properties of ***V*** it therefore follows that the AOV above on ***US*** is a direct analysis of the errors between the underying two-factor model of the GCP and the selected GCP. Moreover, the row AOV above is an analysis of each row or identifiable spatial feature and our objective is to isolate the outliers to the features and images in which they occur.

Note that these steps have overcome the very serious problems that can occur if it is assumed that the errors are all in one 'side' of the equations. However, we still have a difficulty in that least squares 'spreads' error across all the points and can be dominated by a few points with strong 'control'. If these points are also outliers then their identification can be very difficult. In many cases the best tool to use is the human eye and so in the following there will be a number of suggested plots that are sometimes much more decisive than the accompanying statistics!

A8.5 Analysing the SIEVER Solution

The first action in SIEVER is to normalise the matrix ***A*** so that the columns have zero mean and the sum of squares of the columns is 1.0. In this case, the matrix $\boldsymbol{A}^T\boldsymbol{A}$ is in 'correlation' matrix form.

The value of the normalisations is twofold:

- firstly, as noted above, the formulation is more convenient when the columns have zero mean; but
- secondly, the normalisation frees the data of scale changes that different coordinate systems may have.

The SVD of the matrix is then formed to obtain the components discussed above.

A8.5.1 Basic 'PCA' summary

The SVD of the normalised matrix ***A*** can be interpreted as a Principal Component Analysis or PCA of ***A*** in correlation form (see Appendix A6.3 above and Volume 2C).

For each of the *N* components, a simple set of summary statistics comprises

- the Principal Variance (P_{var});
- the percent variance (*Var*%);
- the Total Variance (T_{va_r}%); and
- the Noise to Signal Ratio (as %; NSR%);

each of which is defined by:

$$P_{\text{var},k} = s_k^2$$

$$Var_k(\%) = 100 \times \frac{s_k^2}{\sum_{j=1}^{N} s_j^2}$$

$$T_{\text{var},k}(\%) = 100 \times \frac{\sum_{j=1}^{k} s_j^2}{\sum_{j=1}^{N} s_j^2}$$

$$NSR_k(\%) = 100 \times \frac{\sum_{j=k+1}^{N} s_j^2 / (N-k)}{\sum_{j=1}^{k} s_j^2 / k}$$

These can be calculated for any PCA or SVD exercise. In our case, if the rank of the system is 2, we should have the Total variance in the first two components as very high and the contribution of the remaining components as very low and NSR should reach its minimum at k=2.

It may be shown that these statistics are not very reliable for only two images and it is better in every way if SIEVER is used on a number of images in combinations—such as up to five in an initial investigation and down to studies of pairs of images in final outlier searches.

A8.5.2 Row or feature-based error analysis

The issue, of course, is how to measure what is 'high' and 'low' with regard to the model fit. To help with this we can also generate some statistics for the rows of ***A***—that is for the spatial features that are being mapped as GCP in the images being registered.

For this we will work assuming that the assumption of the rank of ***A*** (p) being 2 is confirmed by the PCA analysis. The consequences of its not being true are discussed in subsequent sub-sections. The assumption we have from this is that the estimated variance of the residual can be written:

$$\tilde{\sigma}^2 = \frac{1}{N-p}\sum_{j=p+1}^{N} s_j^2$$

For Feature i, the row or feature statistics we can generate to help to go on from here are:

- the Normalised Error (e_i^2);
- the feature Chi-square (X_i^2); and
- the Predictive Error Weight (PE_i(%)).

as defined by:

$$e_i^2 = \frac{1}{N-p}\sum_{k=p+1}^{N} s_k^2 u_{ik}^2$$

$$X_i^2 = M\frac{e_i^2}{\tilde{\sigma}^2}$$

$$PE_i = 100\times\left(\frac{1}{1-\sum_{k=1}^{p} u_{ik}^2}-1\right)$$

The derivation of the Predictive Error formula and its relationship with 'jackknifing' is described fully in Appendix 7 above.

A8.6 Using the Q-Q Plot

To investigate the distribution of the errors remaining after fitting the general affine model, SIEVER uses a technique that is very useful from exploratory data analysis called Quantile-Quantile plotting (see also Volume 2B—Section 4.2.3).

A8.6.1 Quantile-Quantile plot

A Quantile-Quantile (Q-Q) plot is constructed to test the statistical distribution of data against a known model. A 'quantile' is a fraction (which when expressed as a percentage is called a 'percentile') between zero and one. Consider that you have M data values (y_j) which are sorted into increasing order such that:

$$y_1 \le y_2 \le \ldots \le y_M$$

Associated with each y_i is a estimate of the fraction (or quantile) of data values less than or equal to y_i. This is simply:

$$q_i = \frac{i-1/2}{M}$$

Suppose $F(x)$ is the cumulative distribution function for a probability distribution function $P(x)$ such that:

$$F(x) = \int_{-\infty}^{x} P(x')dx'$$

The set of x values that satisfy:

$$F(x_i) = q_i \quad i = 1,M$$

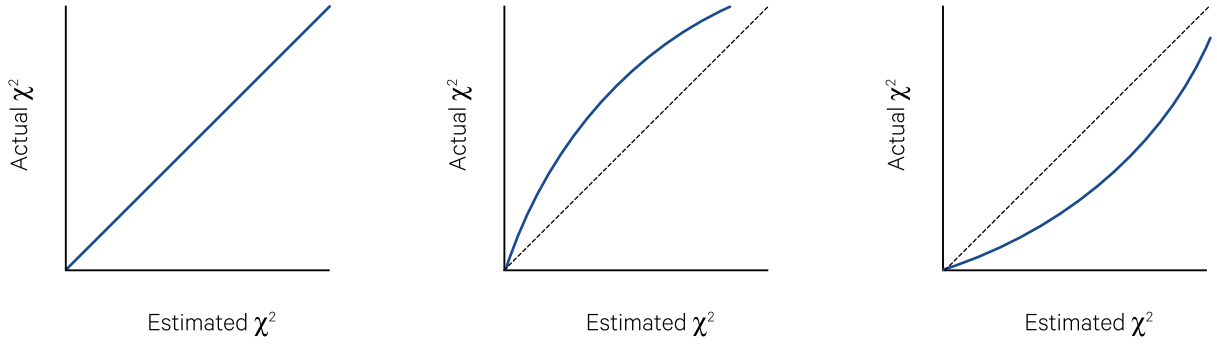
are the set of values for which the expected fraction of values from a sample of M 'drawings' which are less than or equal to x_i is the quantile q_i. It follows that if the data y_i are drawings from the distribution function $P(x)$ then the plot of y_i against x_i for i=1,M will be close to a straight line with slope one and intercept zero. This plot is the Q-Q plot (see Figure A8.1) and tests whether the data values are samples from the distribution $P(x)$.

Figure A8.1 Q-Q plot

a. If actual and estimated chi-square values agree for all control points then the slope of Q-Q plot equals 1

b. If the slope is greater than 1, then some outliers exist in the set of control points.

c. If the slope is less than 1, there are too few control points.



Source: Harrison and Jupp (1992) Figure 40

A8.6.2 Gamma probability plotting

To test the distribution of the residual errors from the SIEVER model, SIEVER uses a general method called Gamma probability plotting. This allows the distribution to be tested against a family of possible distributions, one of which is χ^2. The hypothesis is that if the residuals are normal or near normal then the statistic X_i^2 defined in the previous section will be distributed as Chi-square with (N–p) degrees of freedom, or χ^2_{N-p}. The cumulative distribution for the incomplete Gamma distribution can be defined as:

$$F(x;\alpha,\lambda,\eta) = \int_{\alpha}^{x} P(x';\alpha,\lambda,\eta)dx'$$

$$P(x;\alpha,\lambda,\eta) = \frac{\lambda^{\eta}}{\Gamma(\eta)}(x-\alpha)^{\eta-1}e^{-\lambda(x-\alpha)}$$

If x is χ^2_r then $\alpha = 0$, $\lambda = 1/2$ and $\eta = r/2$. It follows that if the y_i values above are the sorted, X_i^2 defined as in the previous section, and you obtain the values x_i by solving:

$$F(x_i;0,1/2,(N-p)/2) = q_i$$

then the Q-Q Plot of the x_i values against the y_i values should be a straight line with slope one and intercept zero (see Figure A8.1a). In SIEVER this is tested by a linear regression and associated statistics.

A8.6.3 Interpreting the Q-Q plot

The plot can depart from the ideal for a number of reasons. One is that the errors are not normal so the X_i^2 are not distributed as χ^2. It is known that for the general incomplete Gamma distribution the fitted linear regression will have slope $\lambda/2$ and intercept α. Also, if the value (η) for which the data become well approximated by a straight line is different from the theoretical value, it may mean the effective degrees of freedom are different from (N–p).

However, any of these reasons simply relates to the distributions and not the extreme events that represent outliers. As discussed further in Appendix A8.7, outliers are often indicated by very high values of the y_i data for the high quantiles. In this case, the slope is usually much greater than one and the intercept large and negative.

Once outliers are detected they should be traced to the images where the features have been located and checked very closely. Removing points suspected of being outliers is not always justified as in any well-defined distribution there can be large values. SIEVER actually provides a good indication of when there are too few large errors (which might happen if points with large errors are arbitrarily deleted), since in this case the Q-Q plot will often 'flatten' at the higher quantiles indicating there has been truncation of larger but still statistically possible errors (see Figure A8.1b).

A8.7 Data Plots to Indicate Outliers

In addition to the Q-Q Plot, SIEVER provides for some other plots that can give some assurance that the analysis is working as expected or for the location of outliers. These involve listing and plotting columns of the matrix ***US*** and/or the Error Matrix.

A8.7.1 Listing and plotting the matrix *US*

As noted above, the matrix first two columns ***US*** constructed from the SVD can be identified with the modelled 'underlying' set of feature coordinates:

$$\begin{bmatrix} \mu_1 & \nu_1 \\ \mu_2 & \nu_2 \\ \cdots & \cdots \\ \mu_M & \nu_M \end{bmatrix} = \begin{bmatrix} s_1 \underset{\sim}{u}_1 & s_2 \underset{\sim}{u}_2 \end{bmatrix}$$

The last *N*-2 columns can be identified as orthogonal transformations of the errors between the affine model and the data.

If only two spatial datasets are being considered, there will be two columns of the underlying coordinates and two columns of orthogonalised errors. If more than two datasets have been considered then there will be more possible error plots.

An XY plot of the first two columns will indicate if the transformation is working well as the coordinates should be recognisable as linear transformations (such as rotations and scaling) of the coordinates in any one of the datasets. An XY plot of pairs of the columns of orthogonalised errors can locate outliers or large errors and confirm any findings from the Q-Q plot. These alternative analyses of the data are important since, in many cases, outliers so affect the SIEVER model that the actual location is hidden, apart from the general lack of fit or improbability of the model result.

A8.7.2 Listing and plotting the Error Matrix

Another way to search for errors in the location of features, and to identify the image from which they are likely to have come, involves listing and plotting of the Error Matrix. The Error Matrix (***E***) is obtained by 'rotating the errors back' into the original system. That is, it is a reconstruction of the data matrix but with the model components left out:

$$\boldsymbol{E} = \sum_{k=p+1}^{N} s_k \underset{\sim}{u}_k \underset{\sim}{v}_k^T$$

It is the estimated error matrix in the model for the data matrix ***A*** into model and error as:

$$\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{E}$$

XY plots of the columns of ***E*** are sometimes very helpful, but are also often either hard to interpret or not much more informative than the Q-Q plot and plots of the columns of ***US***.

A8.8 Using SIEVER to Locate Outliers

The SIEVER approach provides a set of tools that need to be applied skillfully. The main problem is that outliers, especially when they occur at features with a high Predictive Error weight, can affect the model to a degree where the outlier or outliers may well be fitted better than the majority of the features. Guarding against this possibiity is important and is best done by ensuring that the features SIEVER identifies as having a large predictive error (PE) weight are as accurate and well positioned as possible. A high PE weight indicates that a point has a lot of control over the model. This can be good and often some of the most important features are those with high PE weight. However, if a point with high PE weight is in error it can be very bad for the model.

The existence of a number of outliers to the point where there is not a clear distinction between the fitted errors and the outliers is usually indicated by a very poor fitting model and a lack of clear support for the rank of the ***A*** matrix as 2. These situations indicate the need to look carefully among the located features and images for the problem points.

What is more common is for the outliers to be clearly indicated as large errors after the model is fitted. These unusually large errors will show up in the Q-Q plot and in the XY plots of the columns of ***US*** and/or the columns of the estimated error matrix ***E***.

Because of the possibility that eliminating points with large, but still statistically feasible, errors from the dataset will create a truncated distribution, it is important to improve the locations of features rather than deleting them. The Q-Q plot allows you to see if this has occurred. It is important not to over-SIEVE datasets before models are fitted. The objective is to find outliers and not remove points with large but statistically feasible error values.

The objective of the SIEVER approach is to identify and advise. The feature identifications in the spatial datasets should be revisited before action is taken with a specific point. If this is kept in mind SIEVER provides a very powerful means for screening collections of GCP and preparing the corresponding datasets for resampling.

A8.9 References

Eckart, C., and Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika,* 1(3), 211–218. doi: *10.1007/BF02288367*.

Harrison, B.A., and Jupp, D.L.B. (1992). *Image Rectification and Registration. Part FOUR of the microBRIAN Resource Manual*. CSIRO and MPA, Melbourne.

Lanczos, C. (1958). Linear systems in self-adjoint form. *American Mathematical Monthly*, 65, 665–679.

crc•si

CSIRO

Australian Government
Geoscience Australia

bushfire&natural
HAZARDSCRC