Volume 2B:
Processing—Image Rectification

The report is available in PDF format at *http://www.crcsi.com.au/earth-observation-series*
We welcome your comments regarding the readability and usefulness of this report. To provide feedback, please contact us at *info@crcsi.com.au*.

Publisher:
Australia and New Zealand CRC for Spatial Information

ISBN [ONLINE]:
978-0-6482278-0-9

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.

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.

[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.

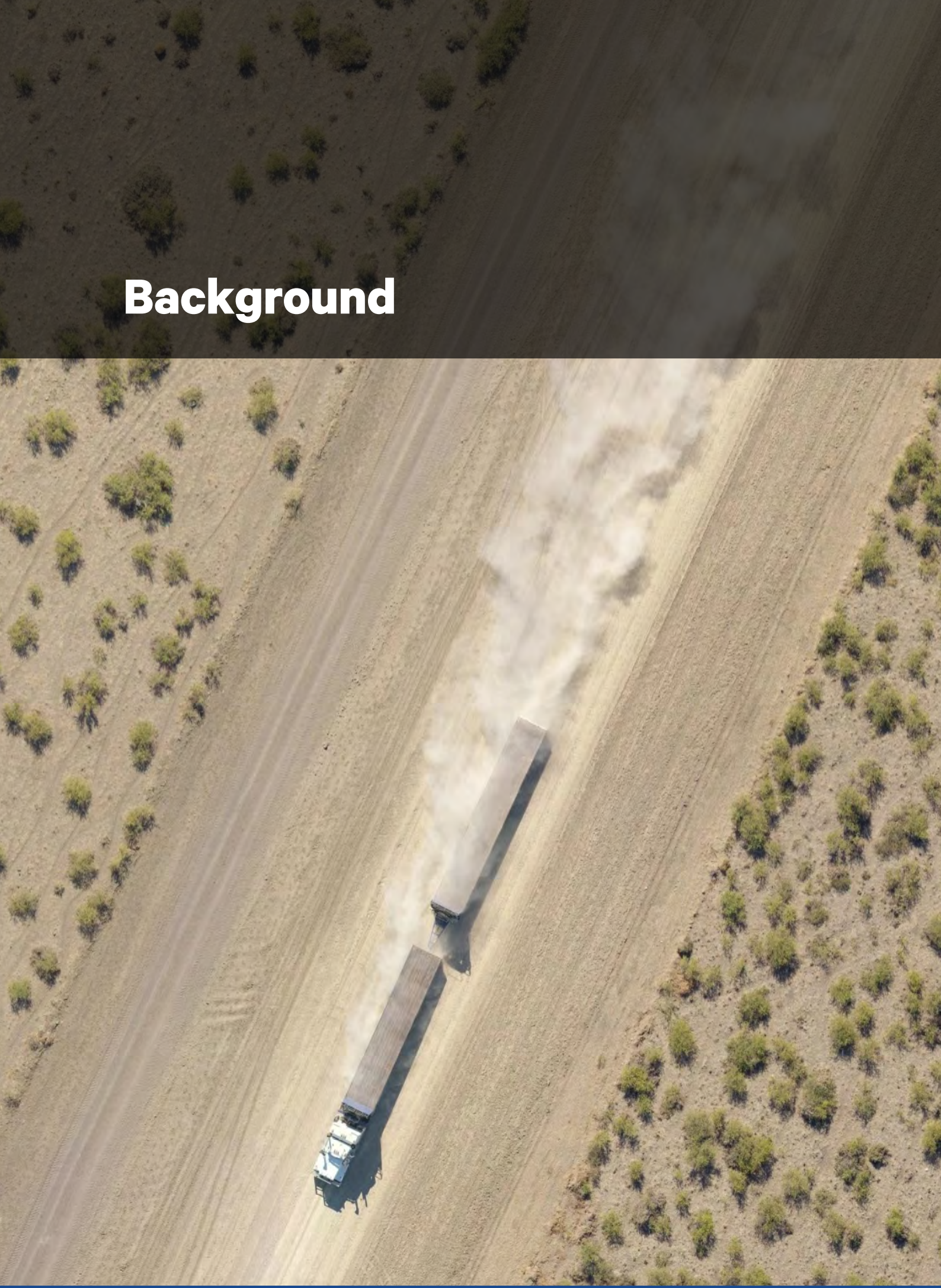
Table of Contents

List of Figures

List of Tables

List of Excurses

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

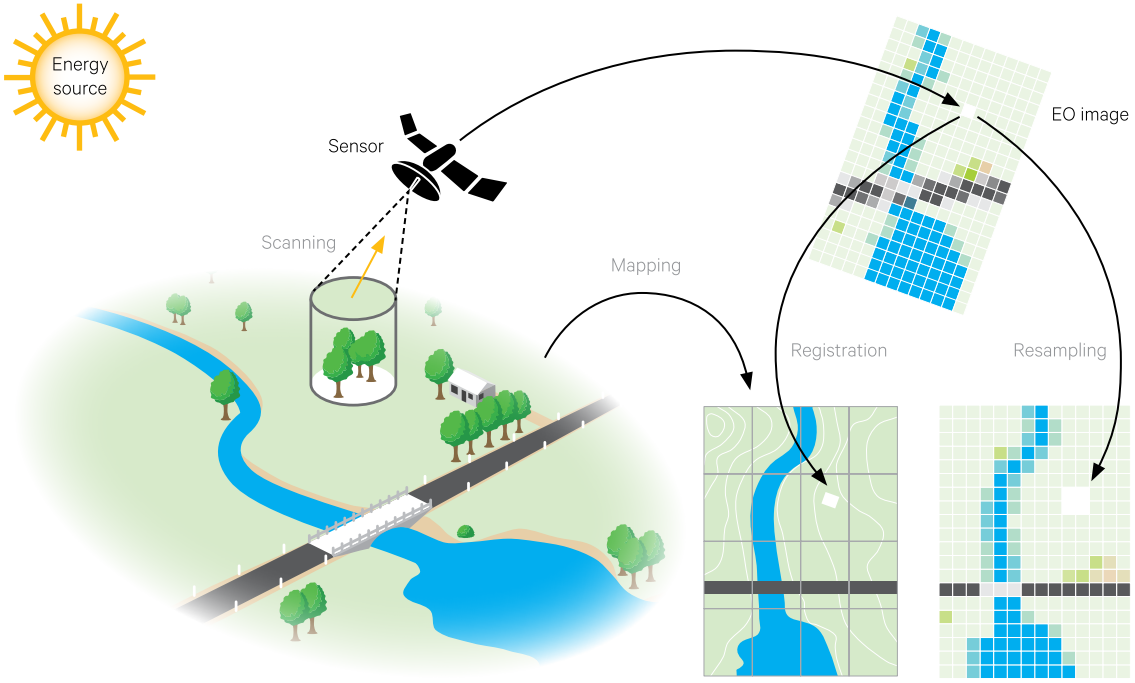
Background image on previous page: High resolution aerial image of a road train near Gurley, NSW, acquired on 25 June 2017. This image has a pixel size of 1 cm. **Source:** Spookfish

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:

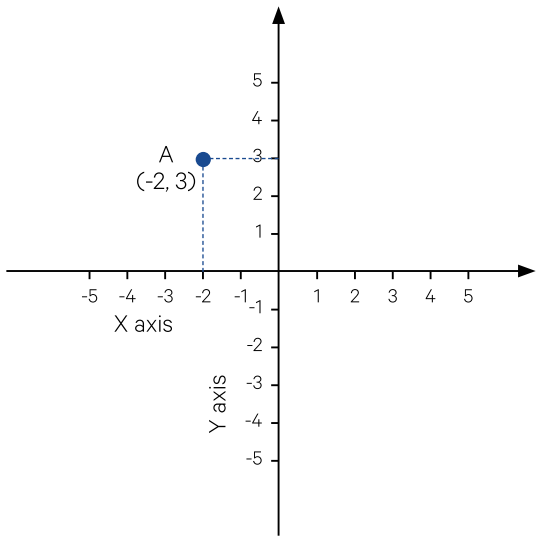
- Cartesian (or rectangular) coordinates (see Section 1.1.1); and
- polar coordinates (see Section 1.1.2).

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



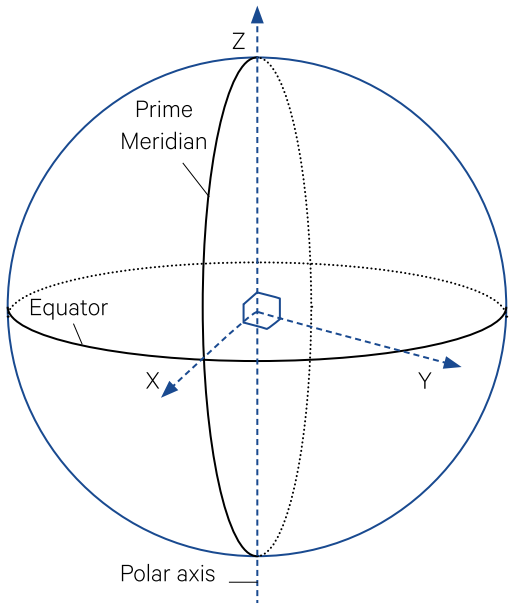
Source: Harrison and Jupp (1992) Figure 2

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

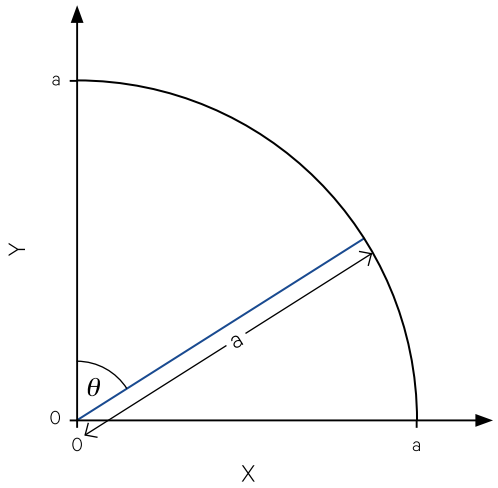
1.1.2 Polar

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

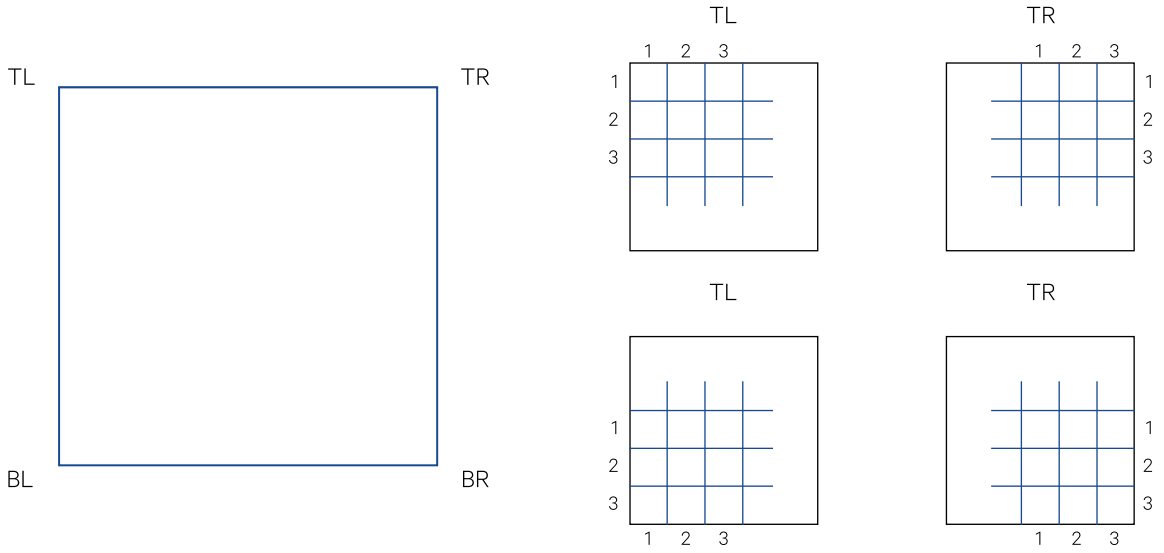
A common difference between coordinate systems is the origin position used for coordinate referencing. Four fundamental grid-addressing schemes can be used as the basis of a planar coordinate system. In image processing these are commonly referred to by the origin positions of:

- 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:

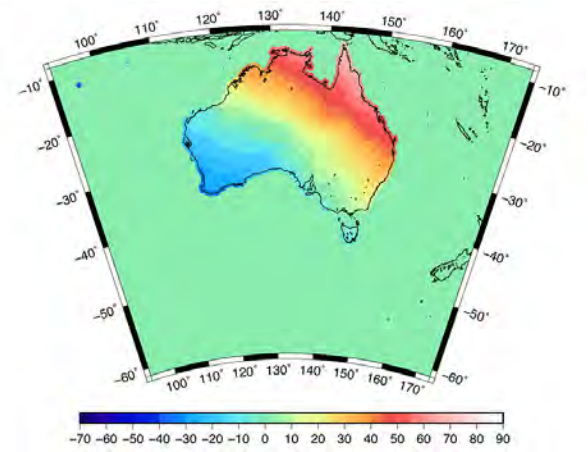
- geodesy—mathematical models of the Earth's shape and size (see Section 1.2.1);
- projections—methods to project from three-dimensions to two-dimensions (see Section 1.2.2); and
- scale—ratio of distances on the map to those it represents on the ground (see Section 1.2.3).

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.

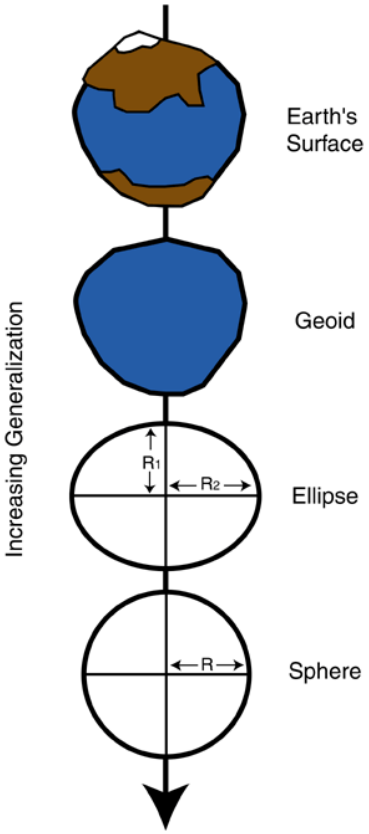
For simplicity in mapping, a mathematical surface called the ellipsoid (or, in older texts, a spheroid) is defined that closely matches the shape of the geoid for a selected portion of the surface (see Figure 1.7). The shape of the ellipsoid is defined by its major and minor axes, which respectively 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

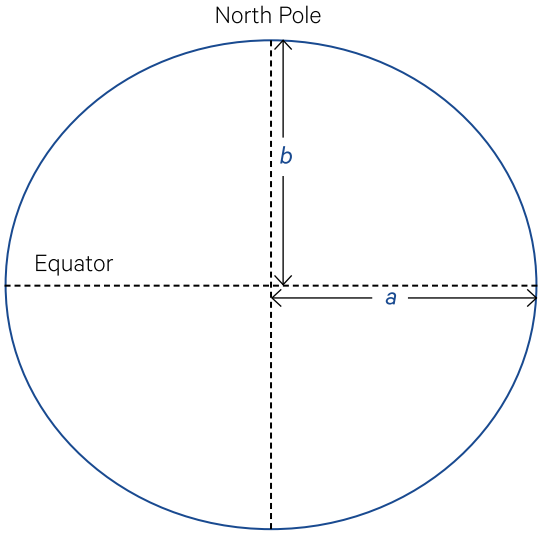
The shape of Earth is irregular. With increasing accuracy, it can be modelled by a sphere, an ellipse or a geoid.



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.

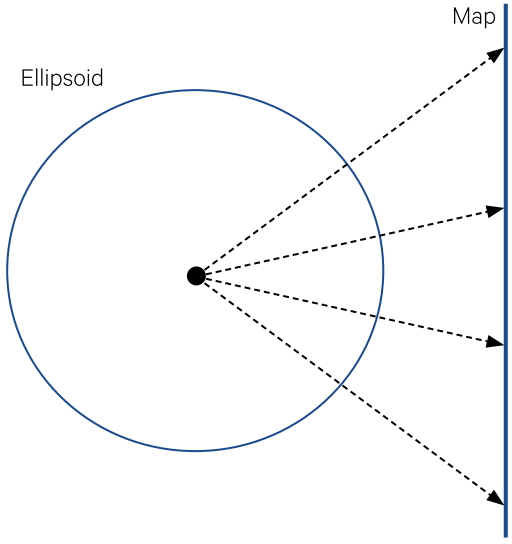
1.2.2 Projections

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.



Source: 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).

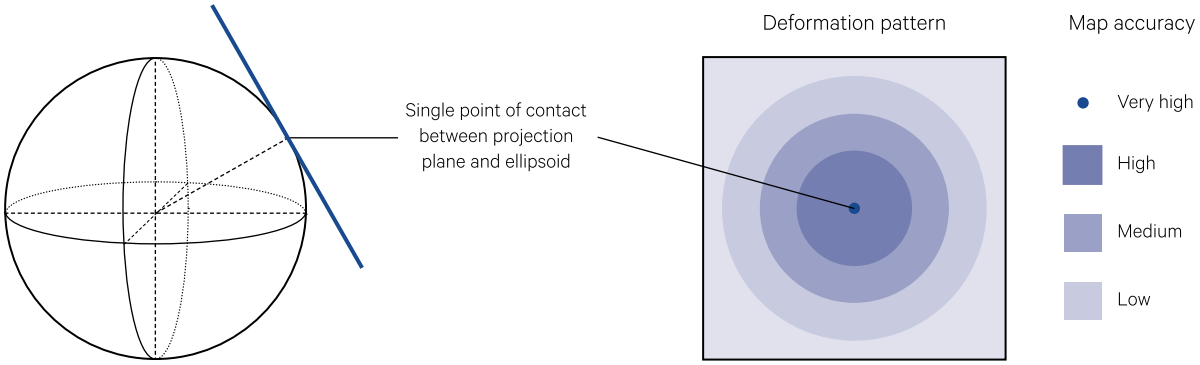
Various map projections have been developed for different purposes. For example, some applications, such as navigation, require that angular information be accurately represented on a map, while others need distances or areas to be truthfully represented.

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).

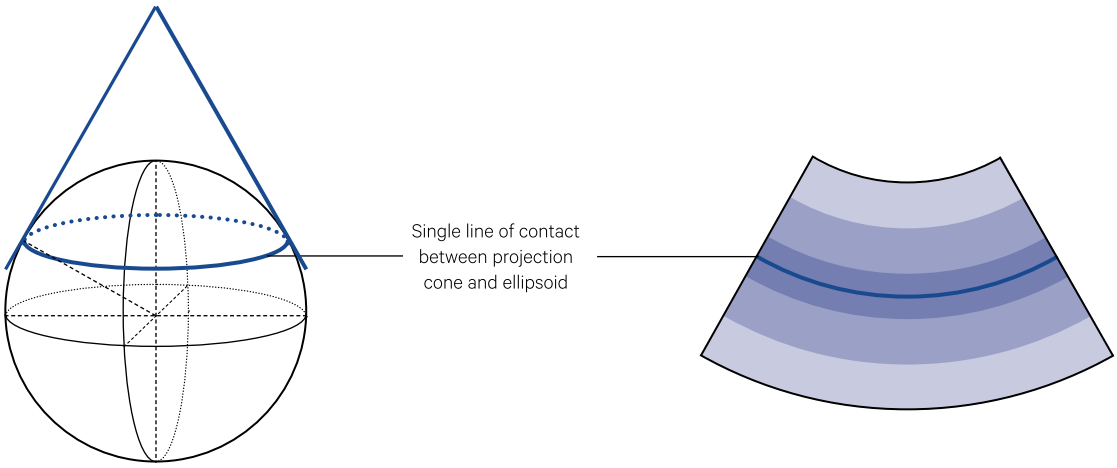
Figure 1.10 Major map projection surfaces

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.

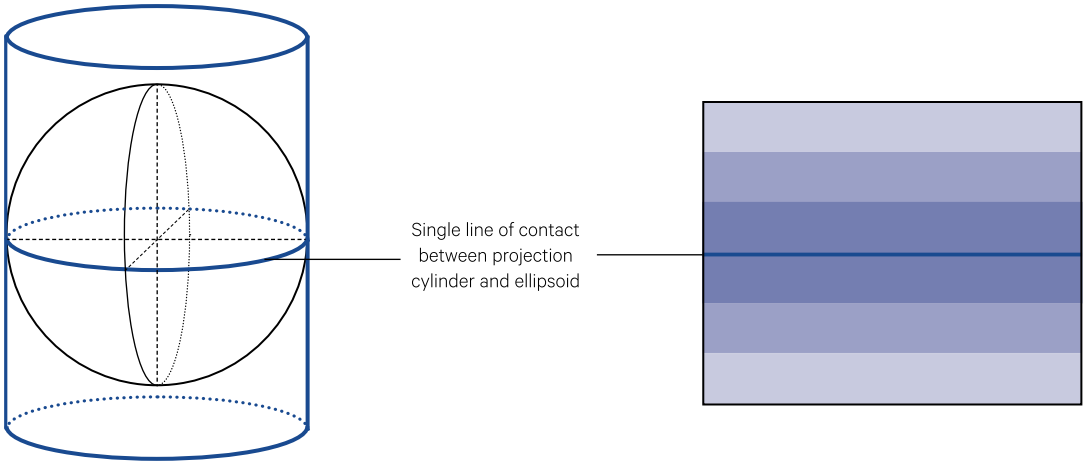
a. Planar projection



b. Conical projection



c. Cylindrical projection



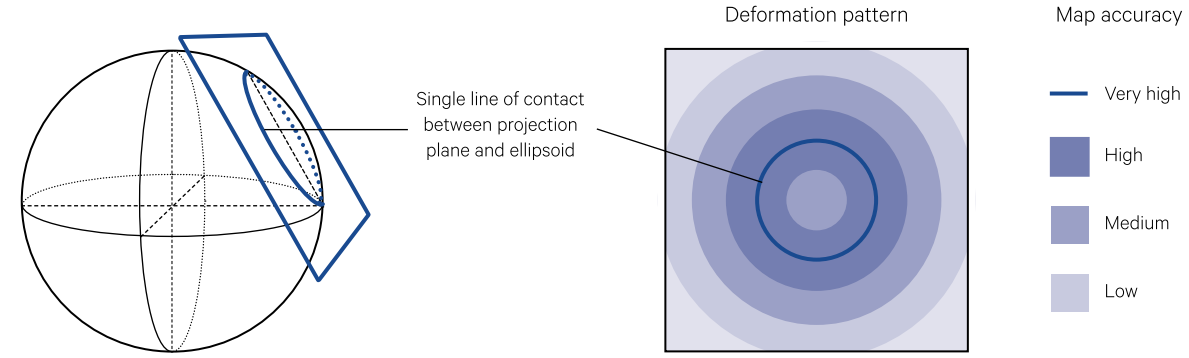
Adapted from: Richardus and Adler (1972) and Robinson *et al.* (1978)

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).

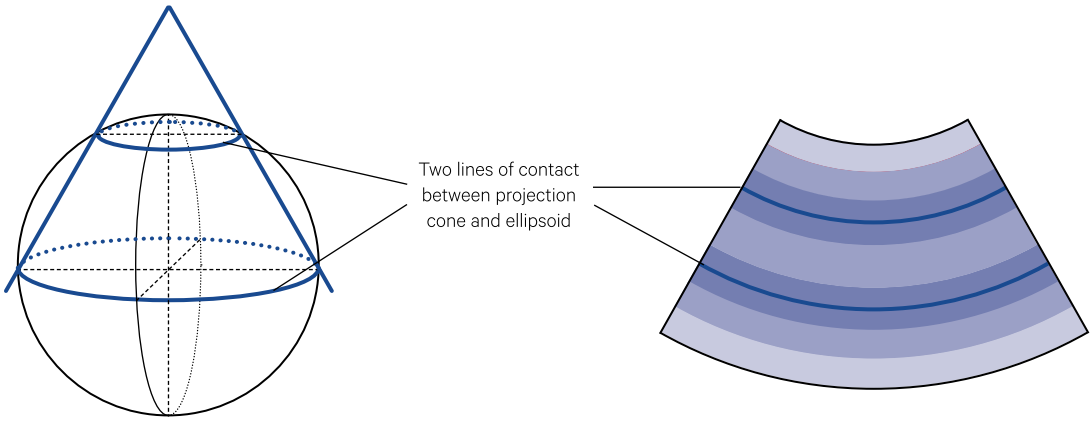
Figure 1.11 Secant contact between projection surface and sphere

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).

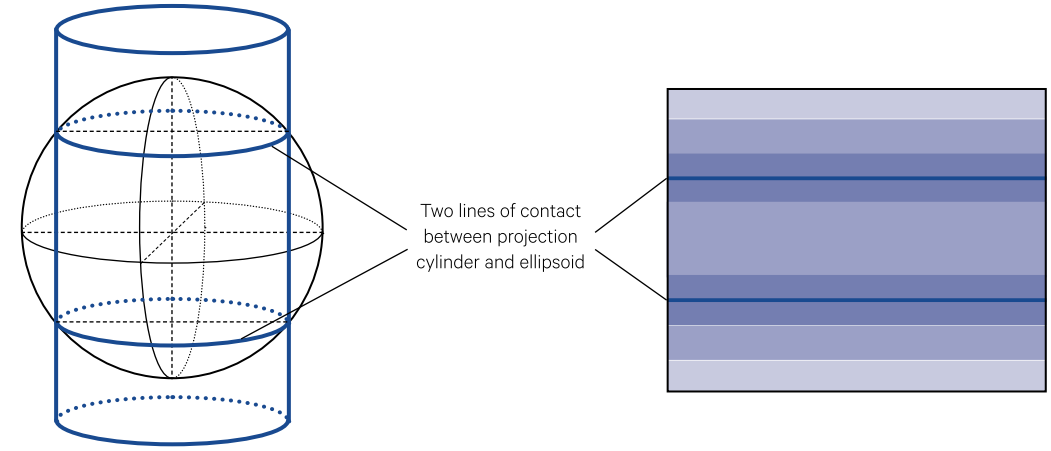
a. Secant projection plane with one circle of contact between projection plane and sphere



b. Conical projection with two circles of contact between projection cone and sphere



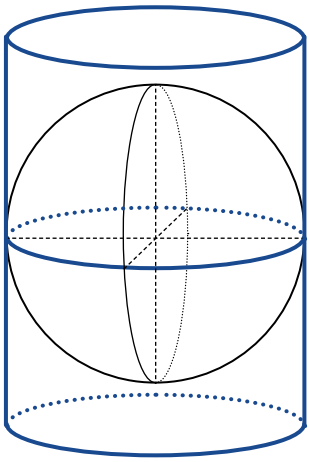
c. Cylindrical projection with two circles of contact between projection cylinder and sphere



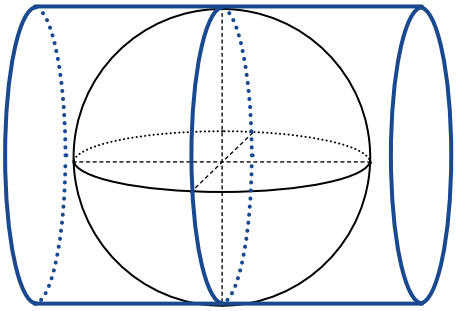
Adapted from: Richardus and Adler (1972) and Robinson *et al.* (1978)

Figure 1.12 Aspects for cylindrical projection

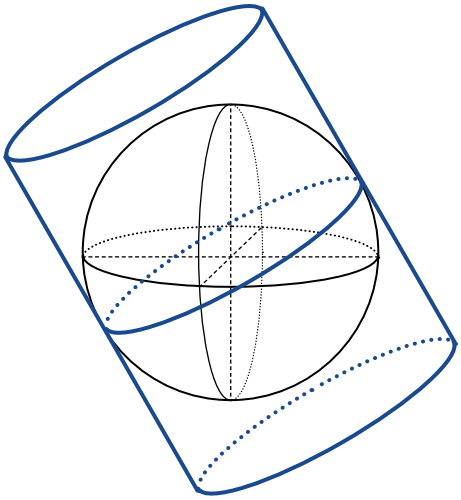
a. Normal cylindrical projection—axis of symmetry for cylinder is Earth's axis of rotation



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).

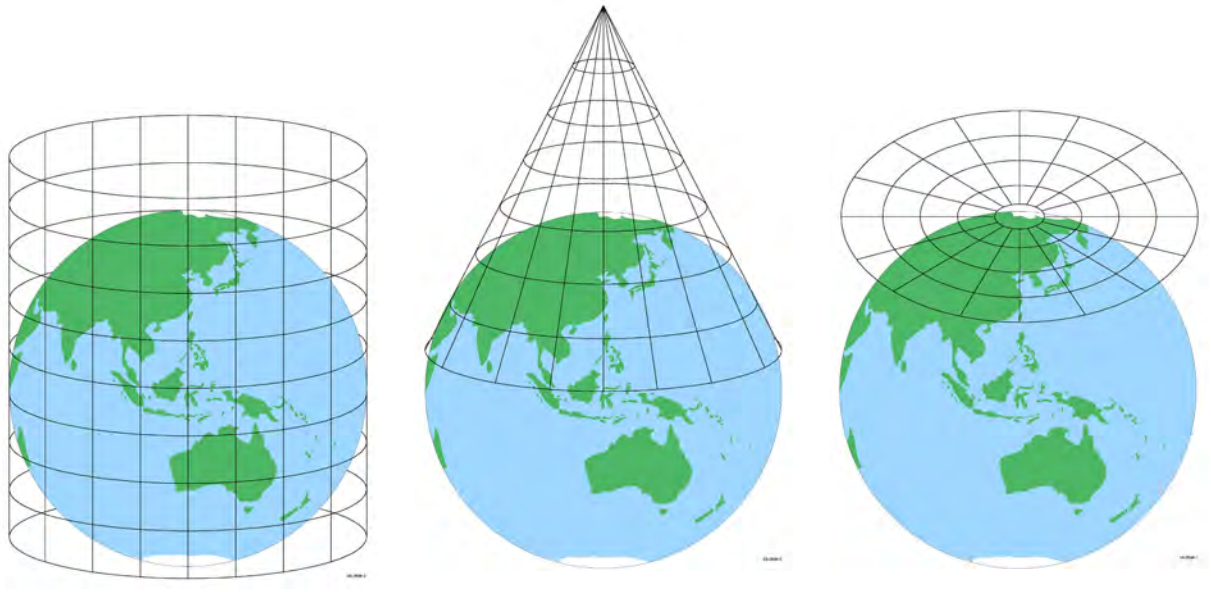
Figure 1.13 Graticules

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.

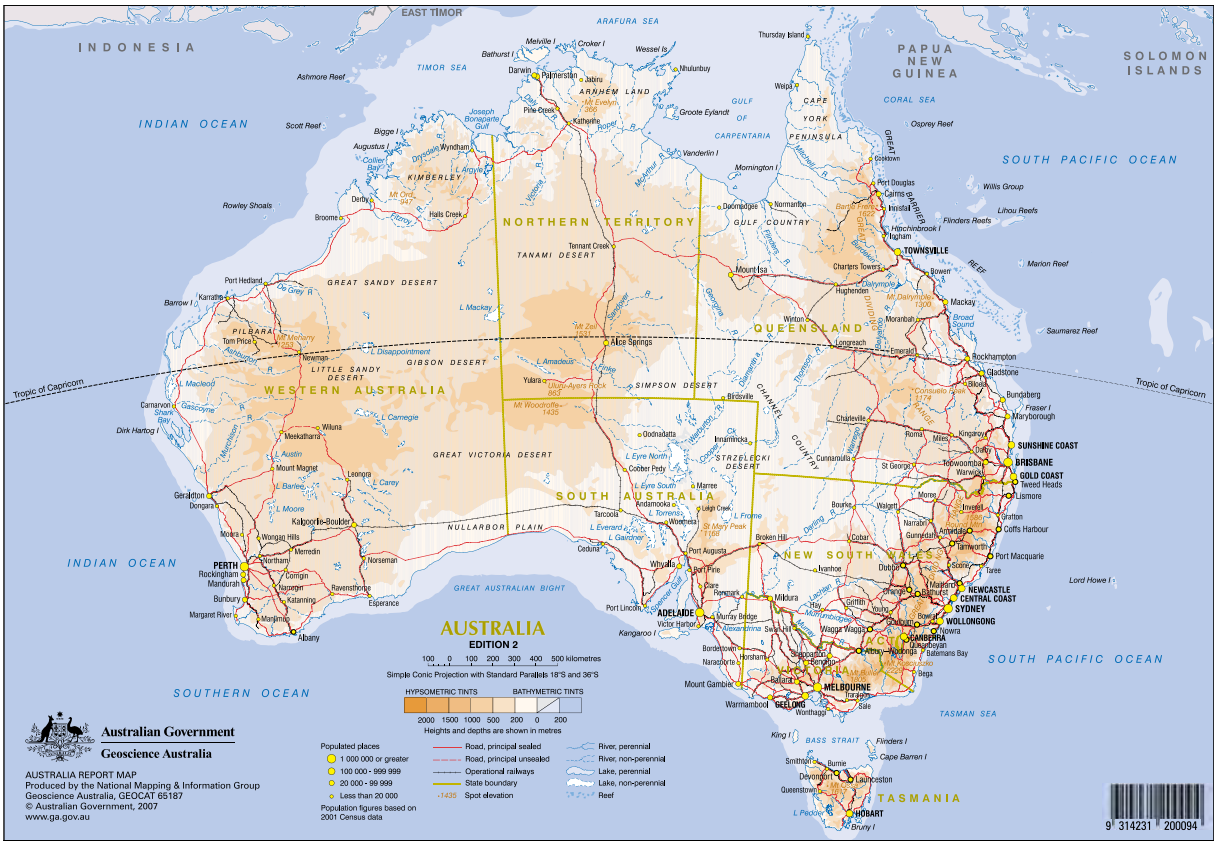
b. Conic projection with normal aspect shows meridians of longitude as evenly-spaced straight lines radiating from the apex of the cone, with parallels of latitude as equally-spaced arcs.

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

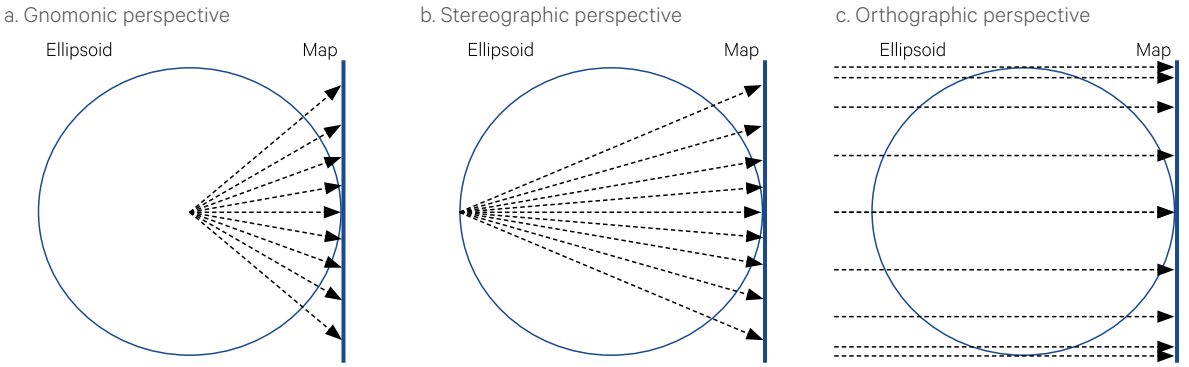
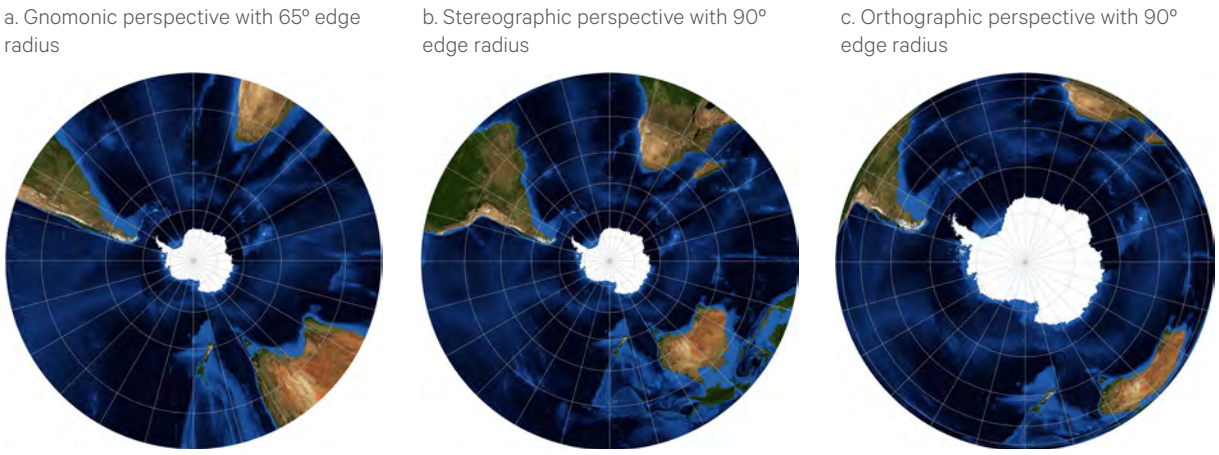


Figure 1.16 Azimuthal maps of South Pole



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

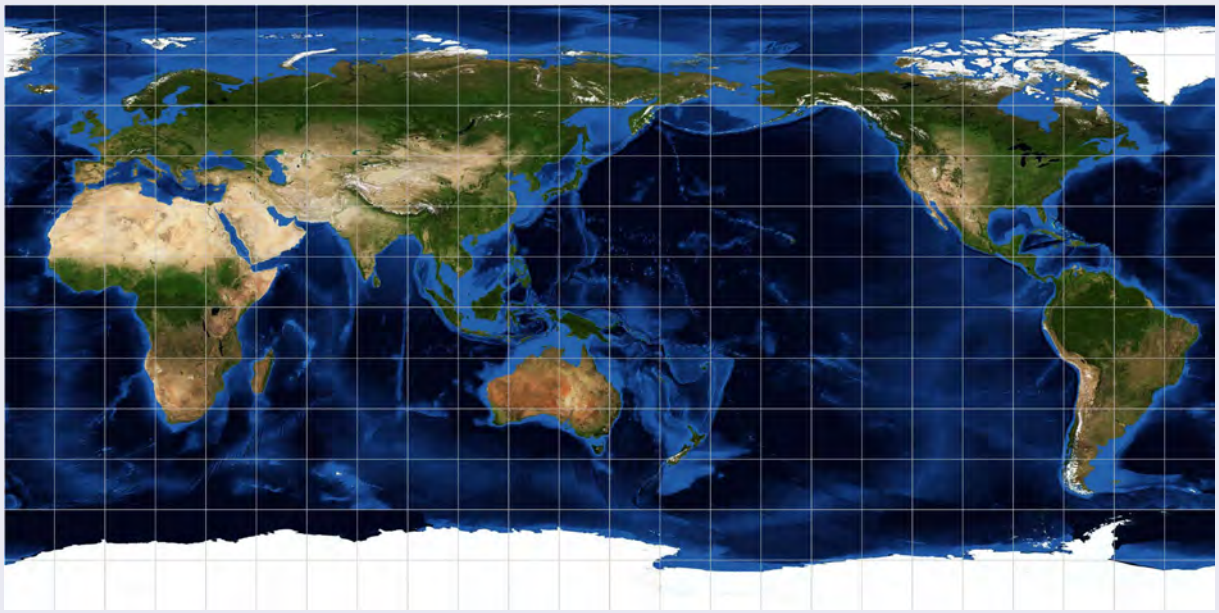
Source: Adapted from ICSM (2018)

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.



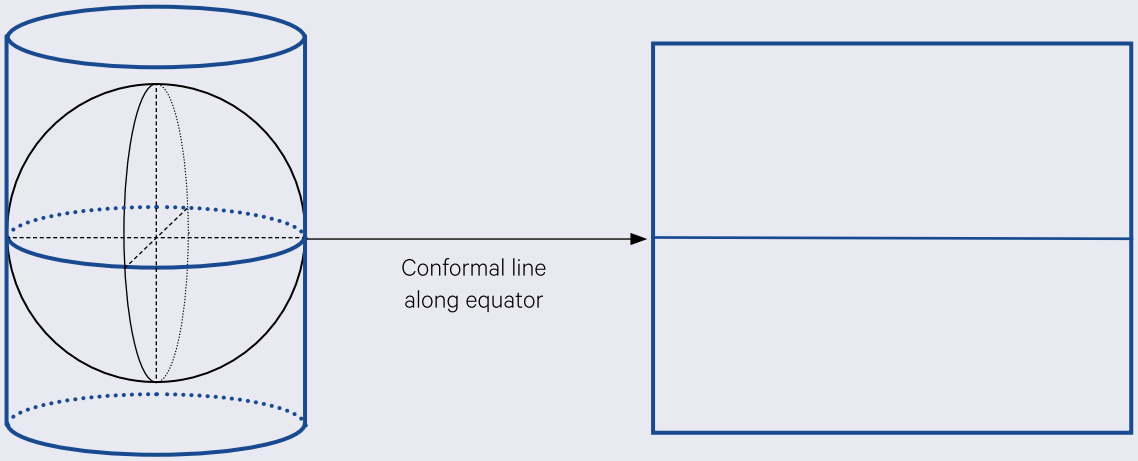
Created using G.Projector 2.3.0 (Schmunk, 2018)

Mercator

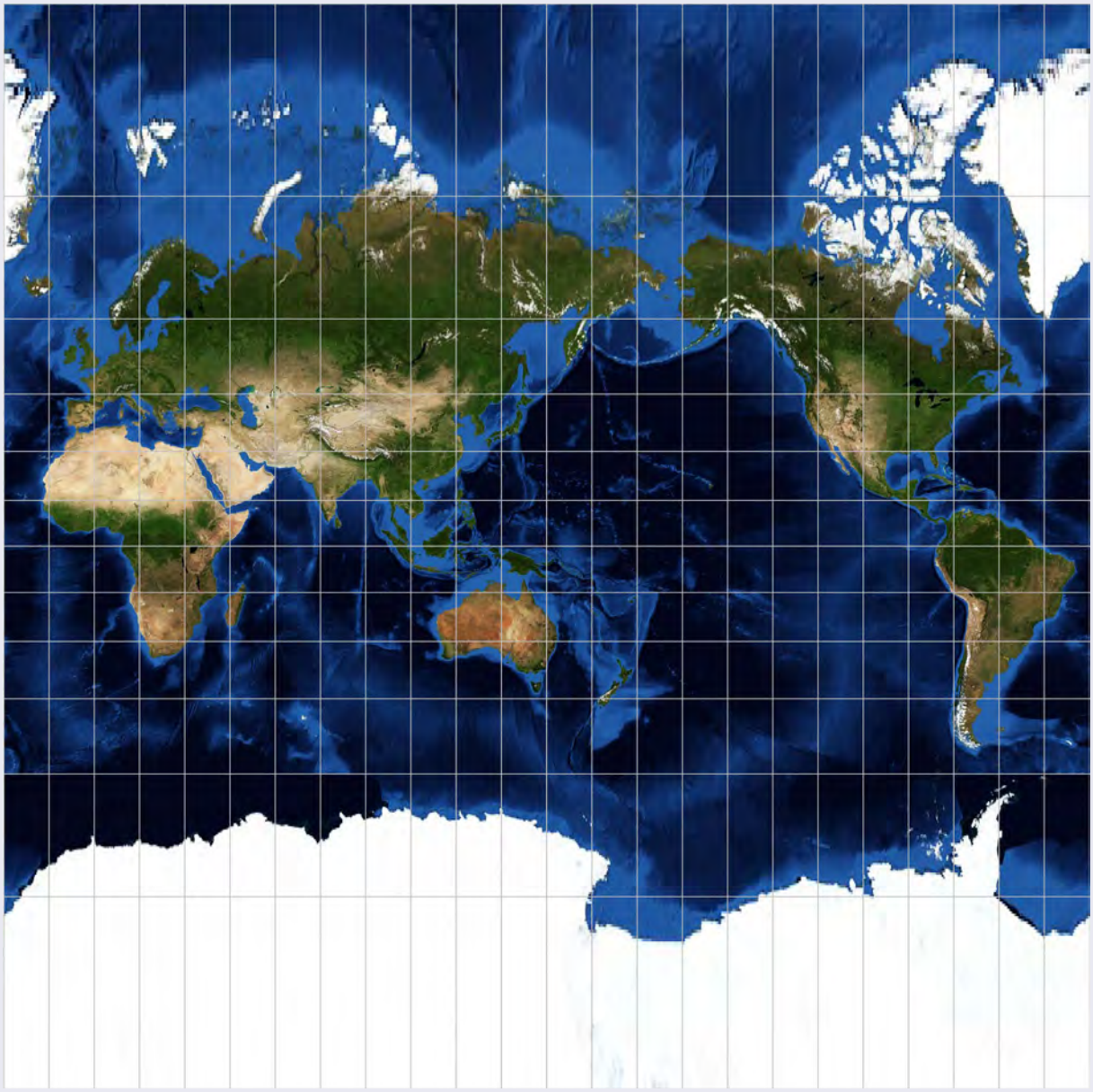
The Mercator projection is a cylindrical projection with a normal aspect and the equator as the standard parallel (see Figure 1.18a). This projection accurately represents the globe's features at the equator, that is, it is said to be 'conformal' at the equator only. Scale, and hence distance measurements, 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



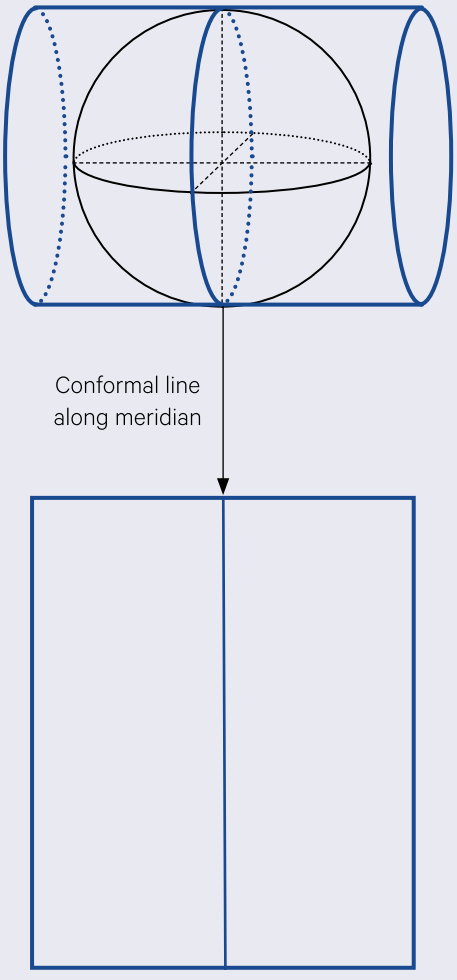
Created using G.Projector 2.3.0 (Schmunk, 2018)

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.

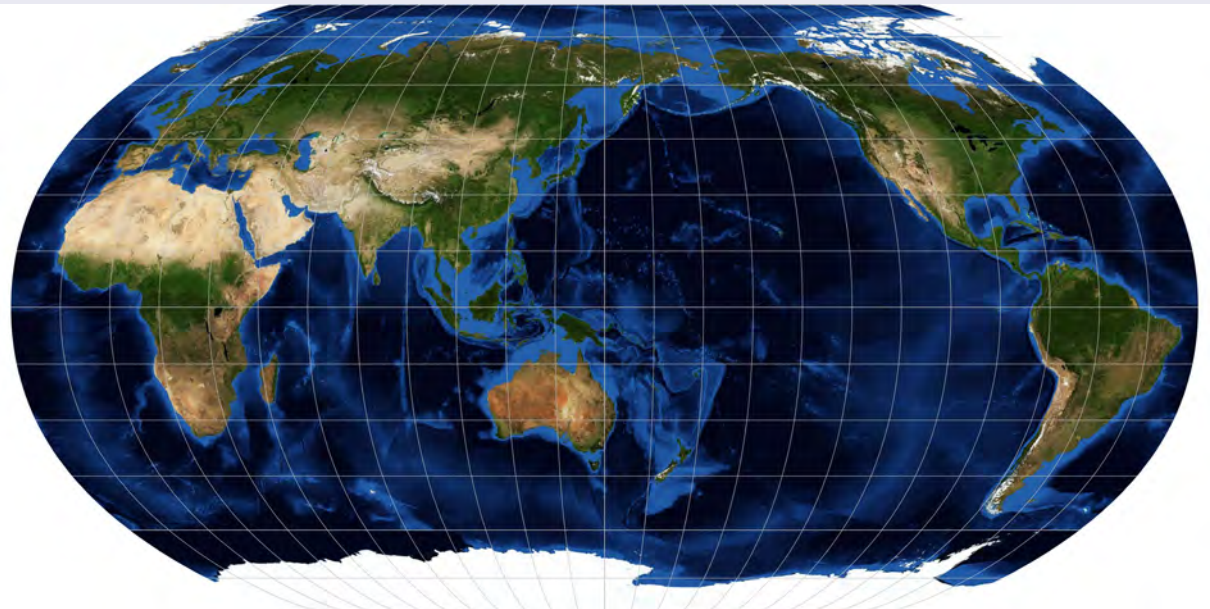


Robinson

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.

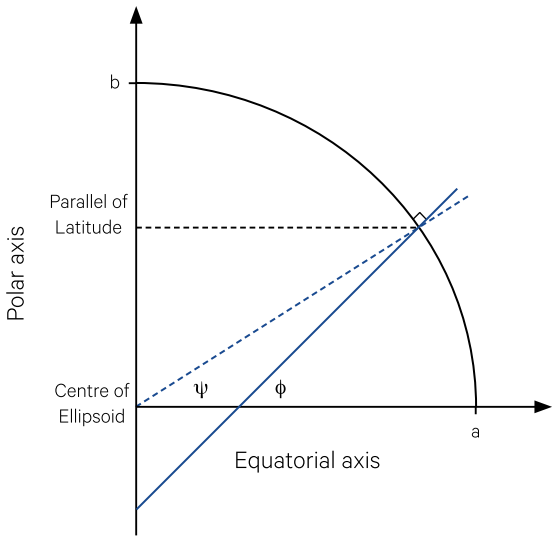
The non-spherical shape of the Earth allows latitude to be measured in a number of ways. The most commonly used method for mapping is geodetic latitude, which measures the angle between the equatorial plane and the normal to the spheroid at a selected point (see Figure 1.23). If the Earth were a perfect 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

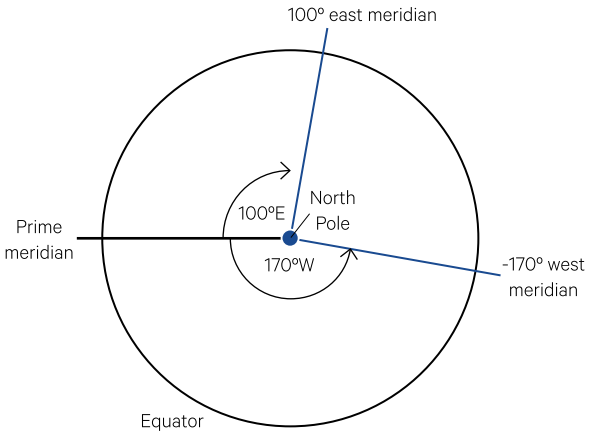
Reference framework for geodetic latitude (ϕ) compared with geocentric latitude (ψ).



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

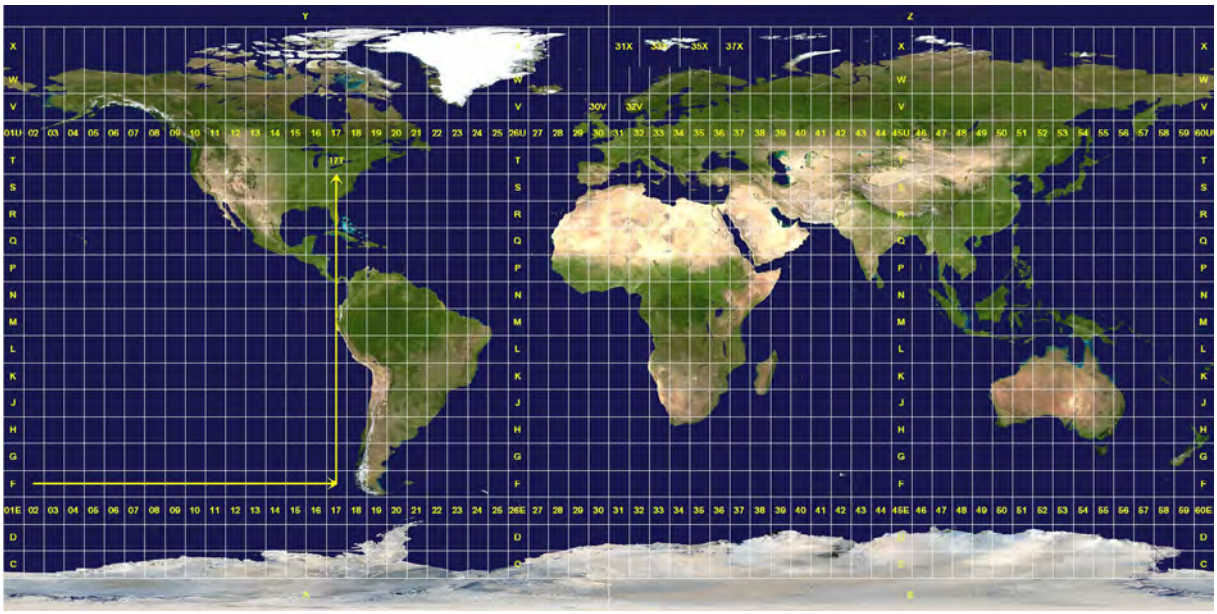
The Universal Transverse Mercator (UTM) system uses a series of 60 Transverse Mercator projections around the globe as the basis of a rectangular grid coordinate system with each UTM zone spanning 6° of longitude (see Figure 1.25). To minimize the scale variations that occur away from the central meridian, 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.



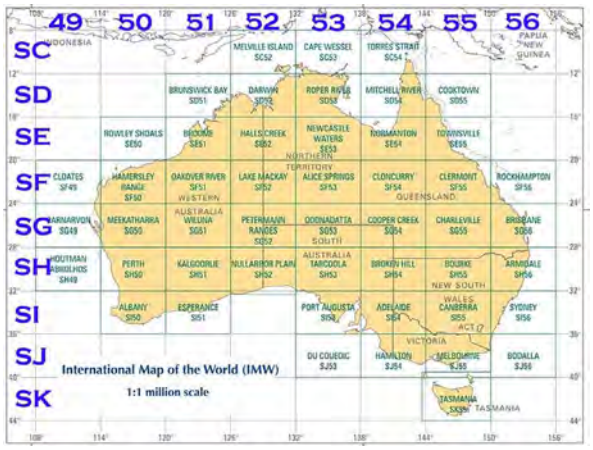
Source: Jan Krymmel (Wikimedia Commons). Retrieved from: https://commons.wikimedia.org/wiki/File:Utm-zones.jpg#file

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

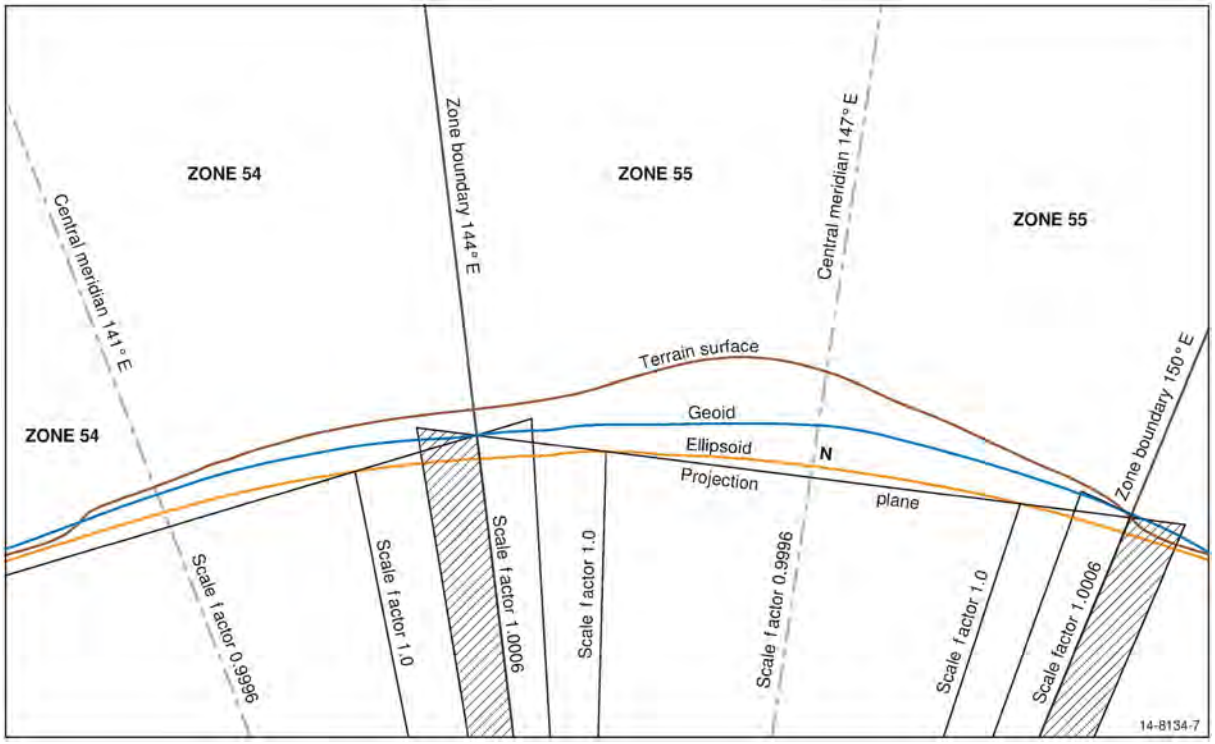
Based on the International Map of the World (IMW), the 1:1,000,000 scale map series indicative index shows the 50 map sheets covering Australia, with the alphanumeric identification grid in blue characters. In this series, the two letter combination identifies the hemisphere and a 4° band of latitude, 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 two reflectance channels (see also Figure 2.7). **Source:** Edward King, CSIRO

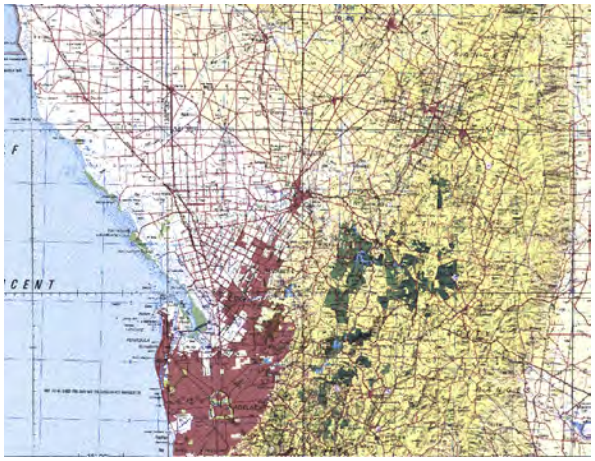
Figure 2.1 Image rectification

Landsat TM images over Adelaide, SA, displayed using bands 1, 2 and 3 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:

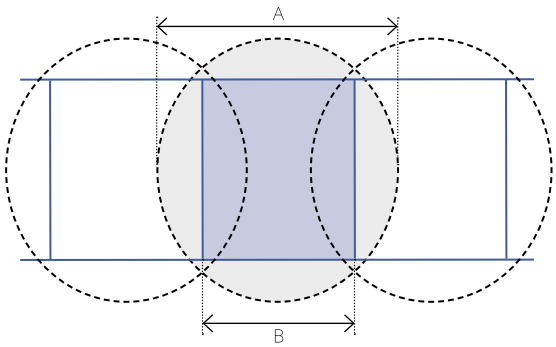
$$2h \times \tan\left(\frac{\theta}{2N}\right)$$

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.

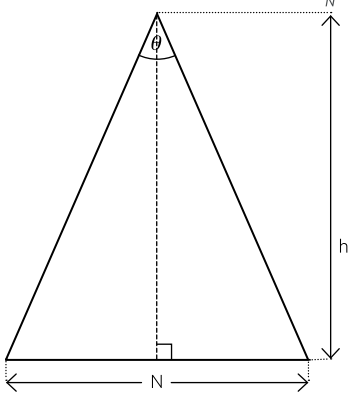


Source: Harrison and Jupp (1989) Figure 35

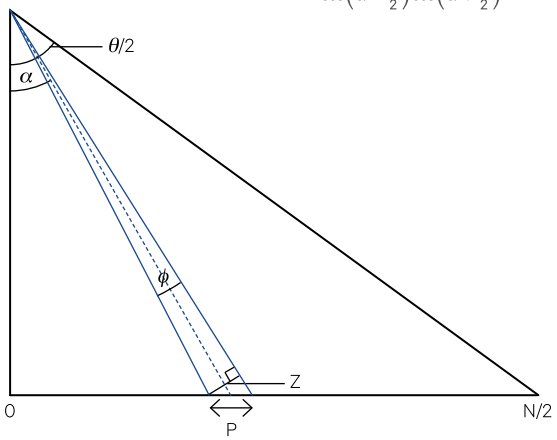
Geometry is the most complete science.
(David Hilbert)

Figure 2.3 Pixel size calculations

a. Average geometric width of a single pixel = $\frac{2h\left(\tan\frac{\theta}{2}\right)}{N}$

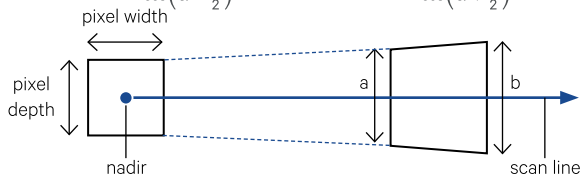


b. Ground width of a single pixel, $P = \frac{h\sin\phi}{\cos\left(\alpha-\frac{\phi}{2}\right)\cos\left(\alpha+\frac{\phi}{2}\right)}$



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 $b = \frac{2h\tan\frac{\phi}{2}}{\cos\left(\alpha+\frac{\phi}{2}\right)}$

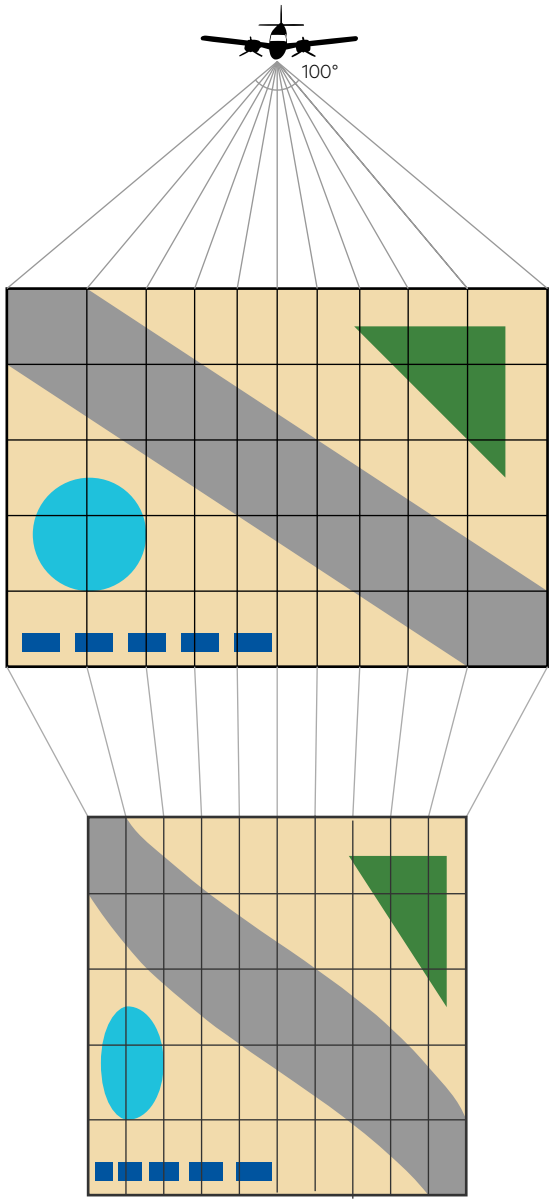


Source: Harrison and Jupp (1992) Figure 15

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

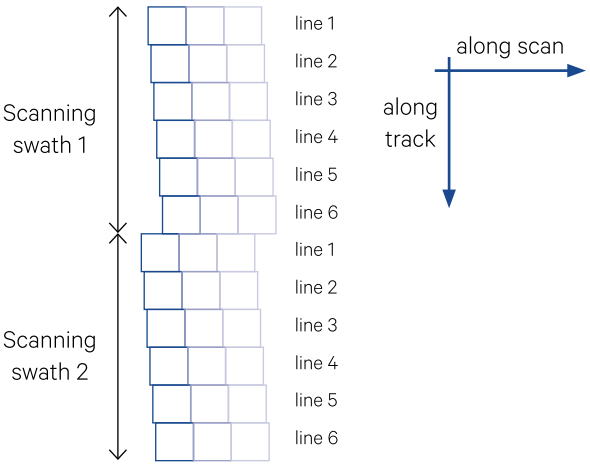
Inconsistencies in the rate of pixel sampling along an image line can cause geometric distortions of sub-areas in the acquired image. If the scanning mechanism used by the sensor does not operate at a constant speed, the ground areas being imaged as individual pixels along a line will not be of equal 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

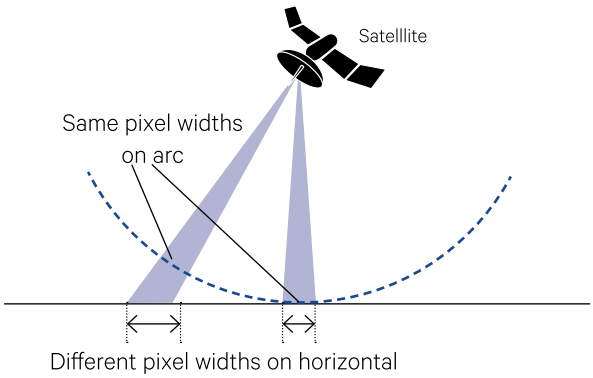
2.3 Viewing Perspective

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.

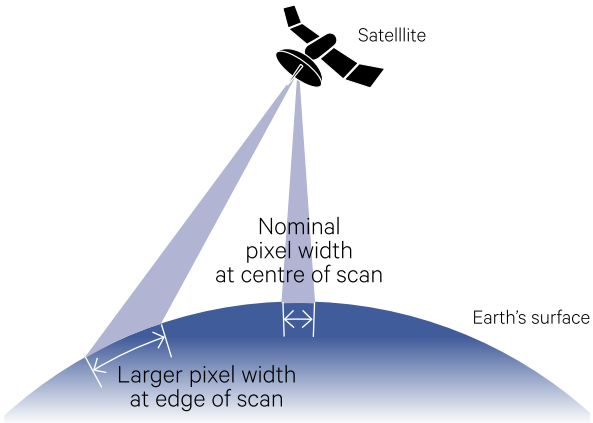
The three-dimensional nature of the Earth's surface becomes relevant to image geometry in the same way that it affects map 'accuracy'. This effect is most significant in continental scale imagery when it combines with panoramic view distortion as shown in Figure 2.6b and Figure 2.7.

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.

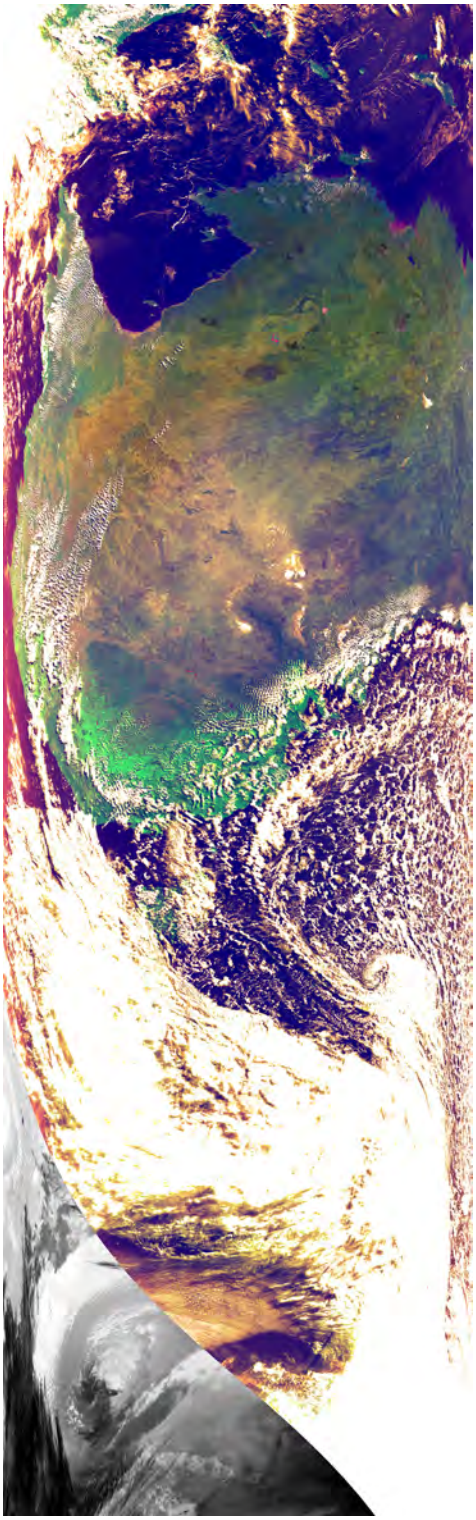


Source: Harrison and Jupp (1992) Figure 17

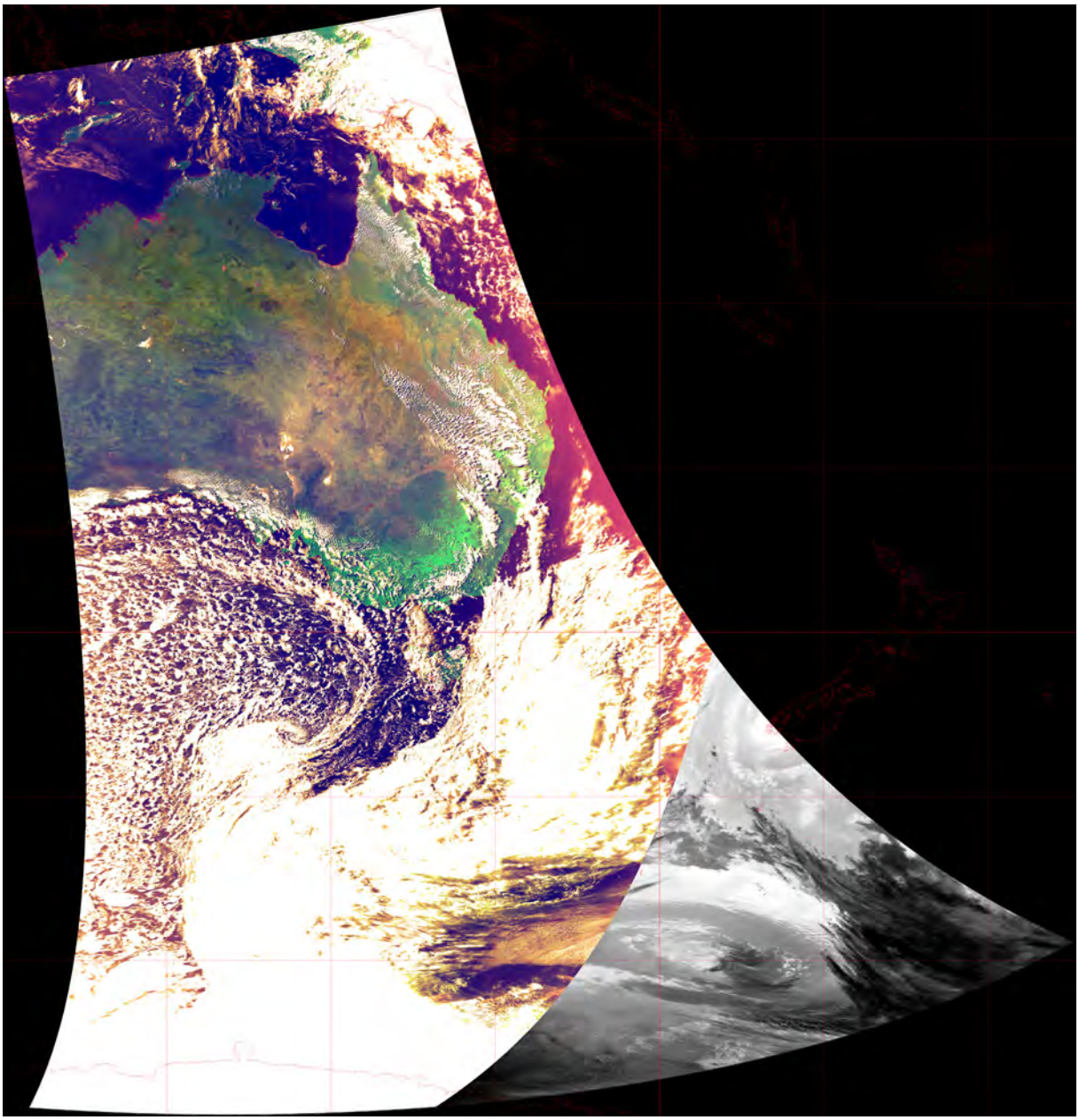
Figure 2.7 Spatial distortions in continental scale imagery

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. 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.



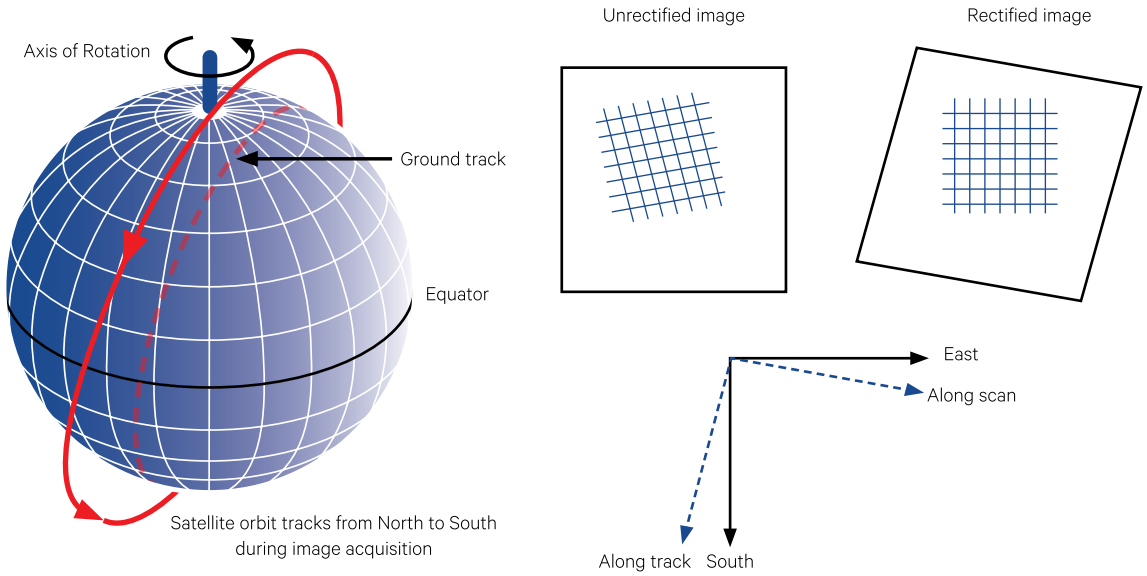
Source: Edward King, CSIRO

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).

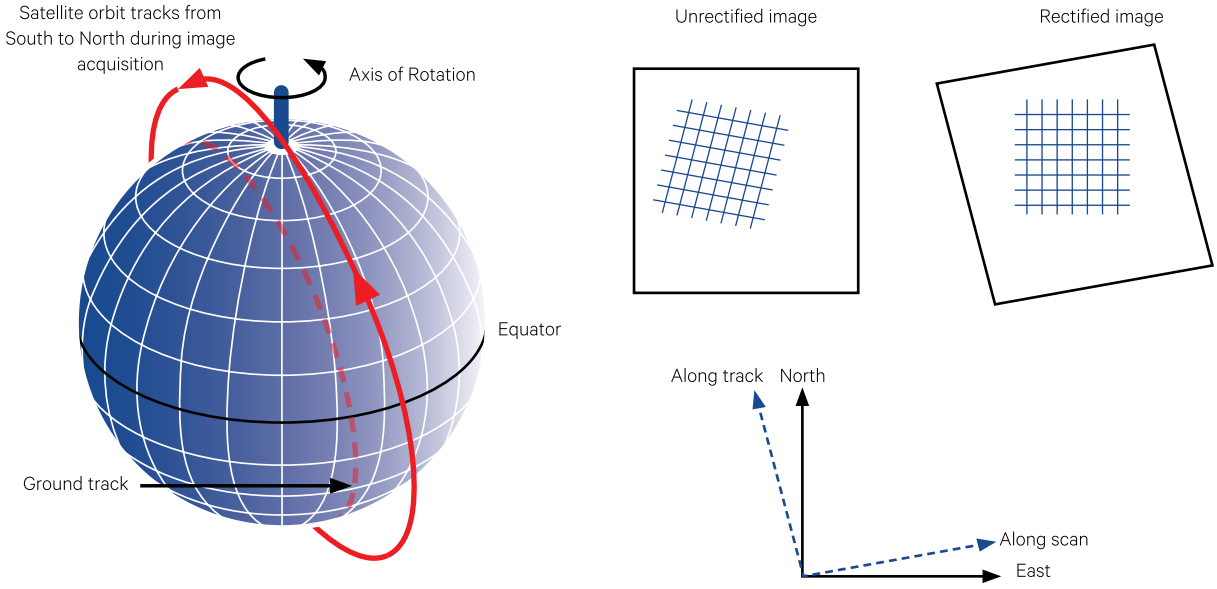
The only thing you sometimes have
control over is perspective.
You don't have control over your situation.
But you have a choice about how you view it.
(Chris Pine)

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.

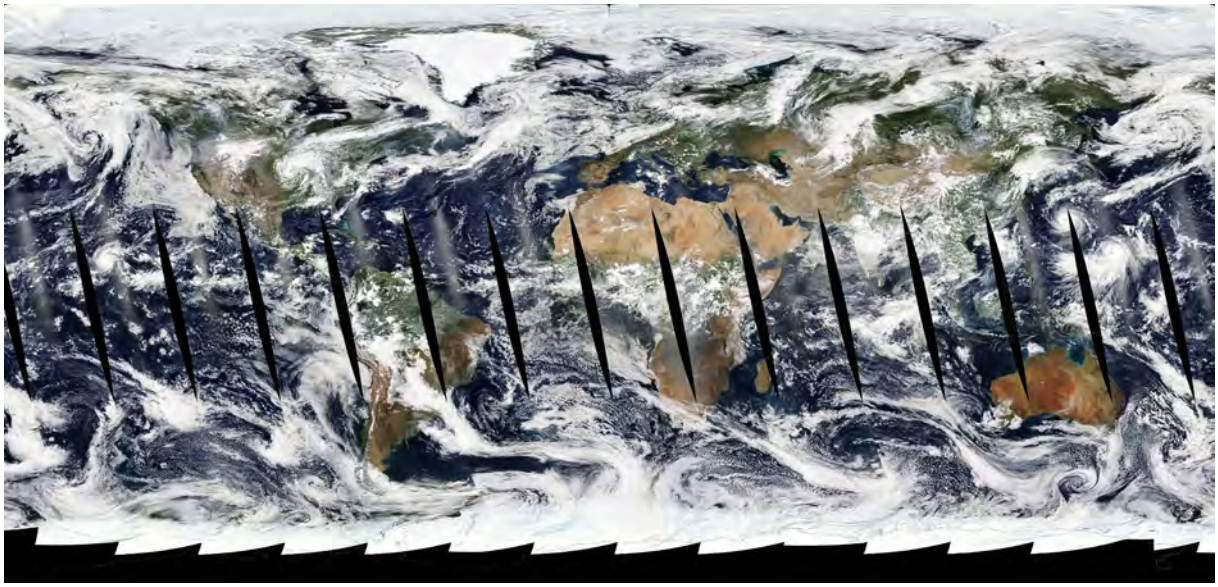
Scanning devices use the continuous forward movement of the platform during image acquisition to determine the separation between lines in an image. However, this movement also results in another skew distortion. As each line is being imaged, the platform moves forward causing the end of the line to 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.

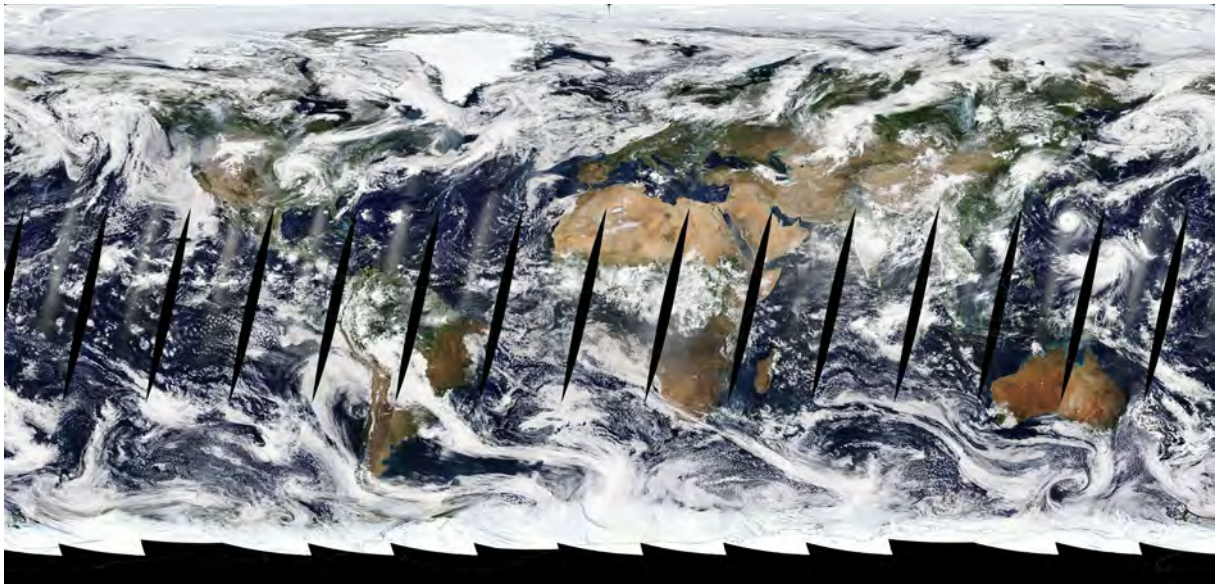
Figure 2.9 Daily acquisitions for MODIS/Aqua and MODIS/Terra

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/Aqua collects imagery during its ascending mode so image strips are skewed in northwest/southeast direction.



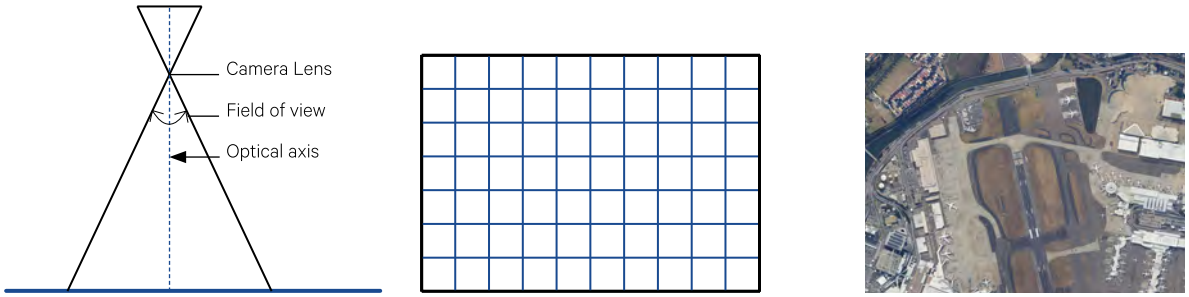
b. MODIS/Terra collects imagery during its descending mode so image strips are skewed in northeast/southwest direction.



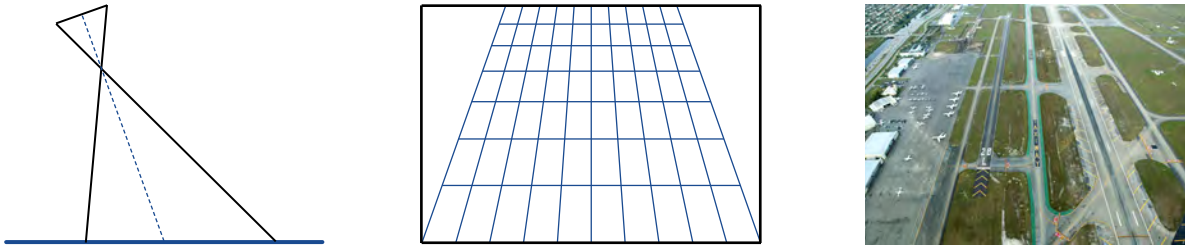
Source: NASA Worldview. Retrieved from: *https://worldview.earthdata.nasa.gov*

Figure 2.10 Effect of camera orientation in aerial photography

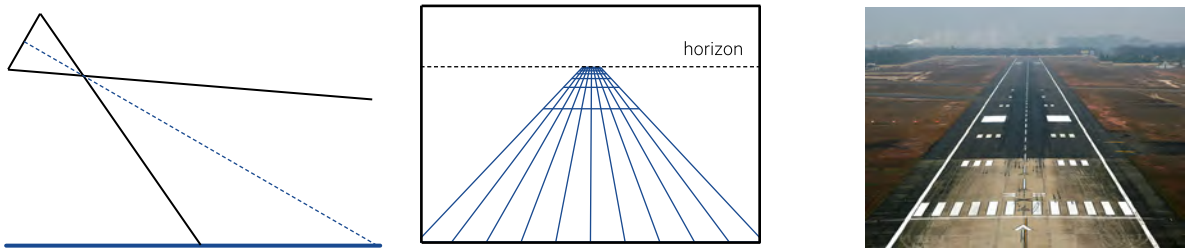
a. Vertical camera orientation yields relatively regular image grid.



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.



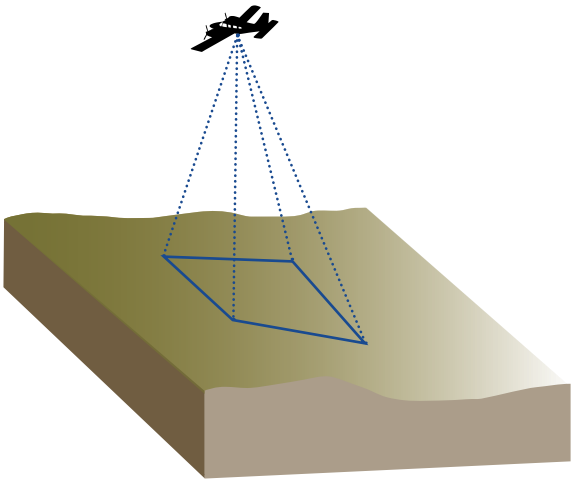
Adapted from: Wolf and Dewitt (2000)
Image sources: a. Perth Airport Runway 16R acquired on 3 May 2017, ©Spookfish.
b. Palm Beach International Airport Runway 28L, ©Dunn's Aerial Photography (Wellington, Florida, USA).
c. Trivandrum International Airport runway 32R photographed on 31 January 2007, © 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

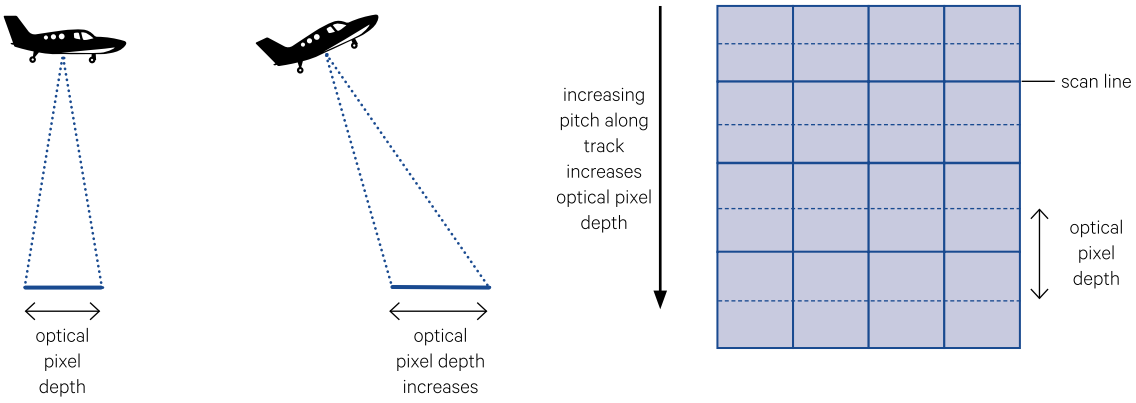
Both platform stabiity and surface relief impact the ground area imaged in each image pixel.



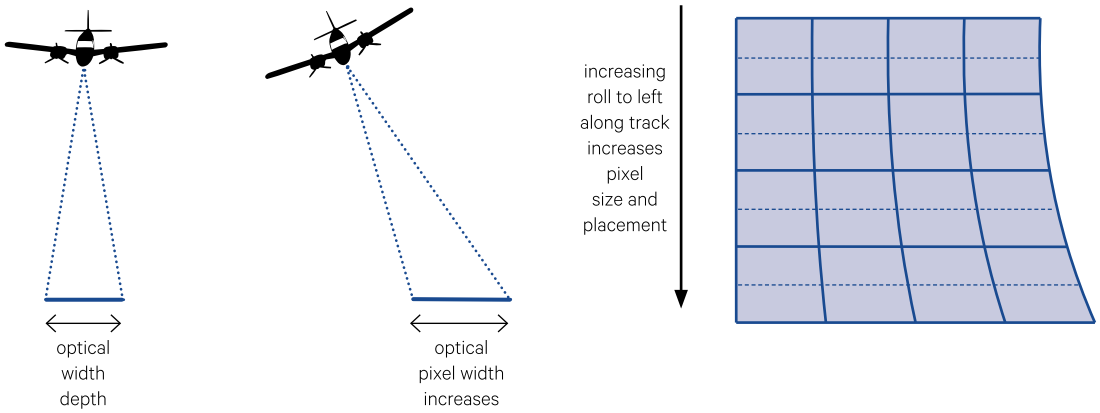
The pitch of a platform is defined as an angular movement about the scan line direction (that is, perpendicular to the direction of platform movement) as shown in Figure 2.12a. A constant pitch has the effect of displacing the complete image scan line. If the pitch is forward (that is, lowering the front 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

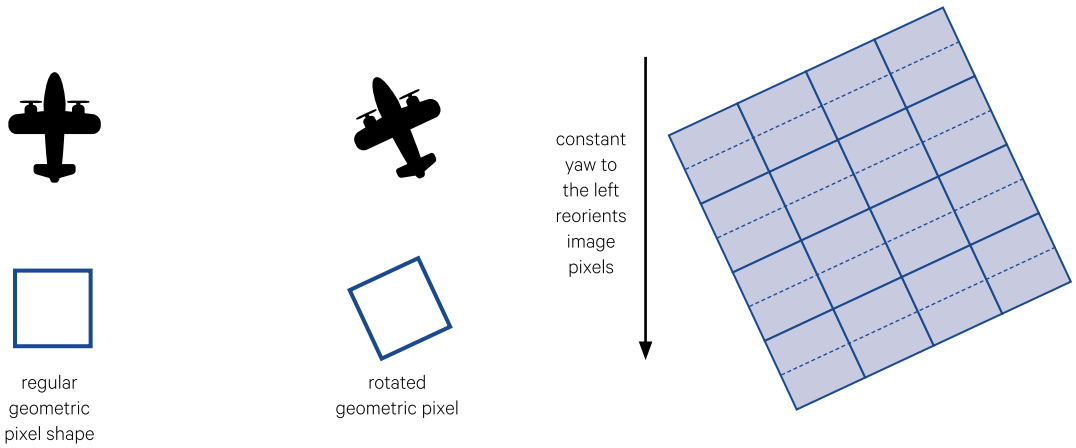
a. Pitch—variations in pitch affect pixel size and placement along track



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.

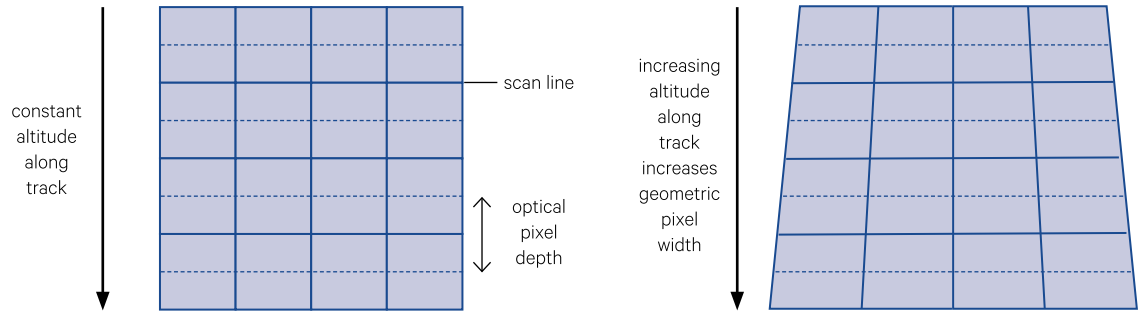
In addition to the effects of platform attitude, variations in platform altitude introduce distortions into EO imagery. The principle of this effect is illustrated in Figure 2.13. As the imaging system moves away from a target with increasing altitude, the IFOV is increased in size so the full image 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

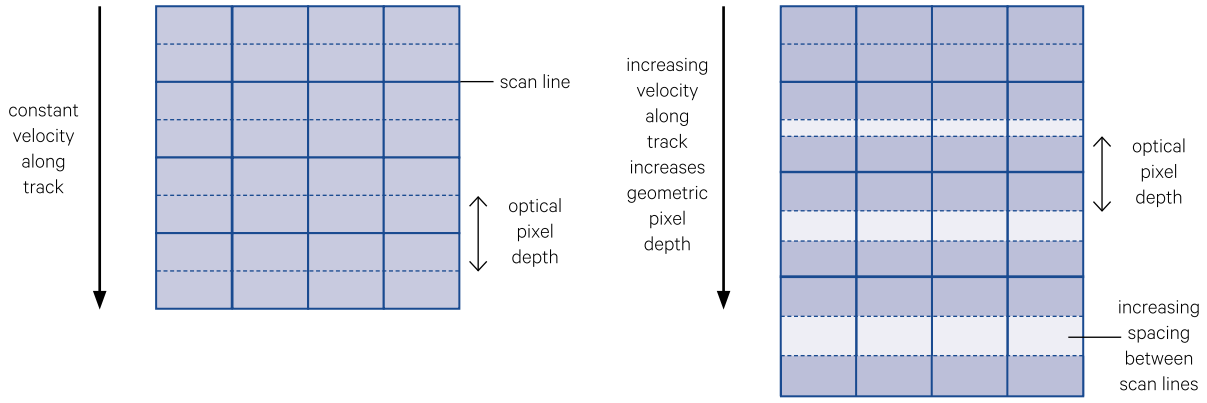
Increasing altitude along track of platform increases the image pixel width.



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.

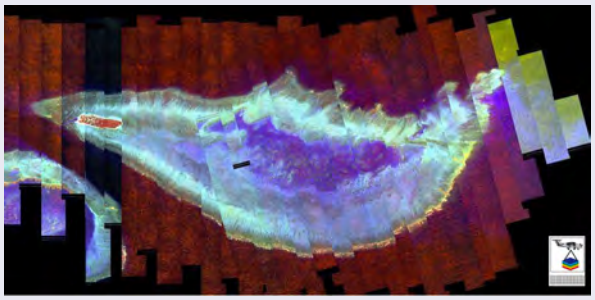
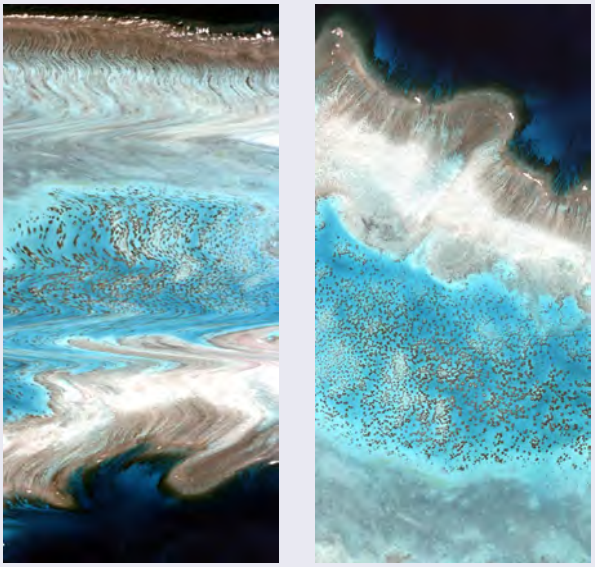


Figure 2.16 Geometric correction of CASI imagery

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

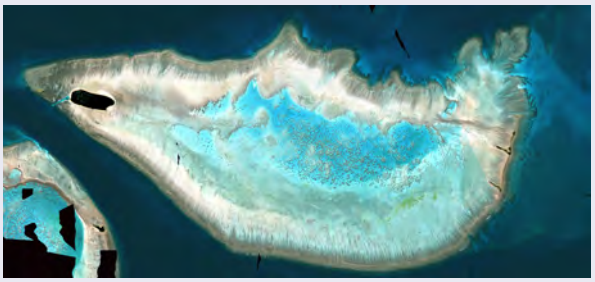
b. Corrected image



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

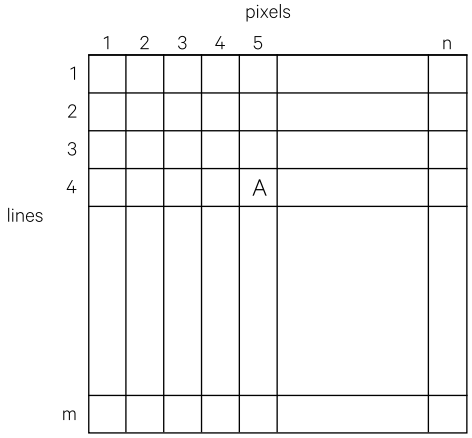
2.5 Image Grid Origin

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).

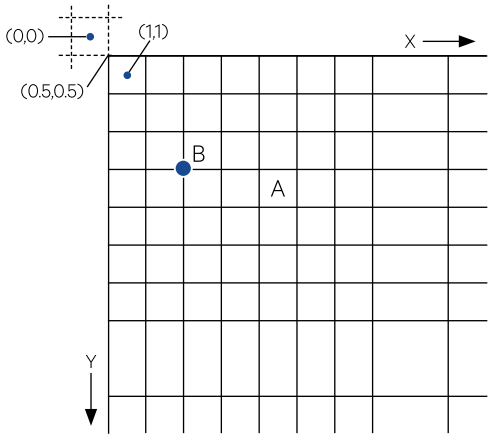
In the TL origin system, image coordinates of a particular pixel are given as the number of pixel columns to the right of the left-most edge of the image and the number of lines down from the top of the image. The centre of the first pixel on the first line of the image is considered to be (1,1) so that 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:

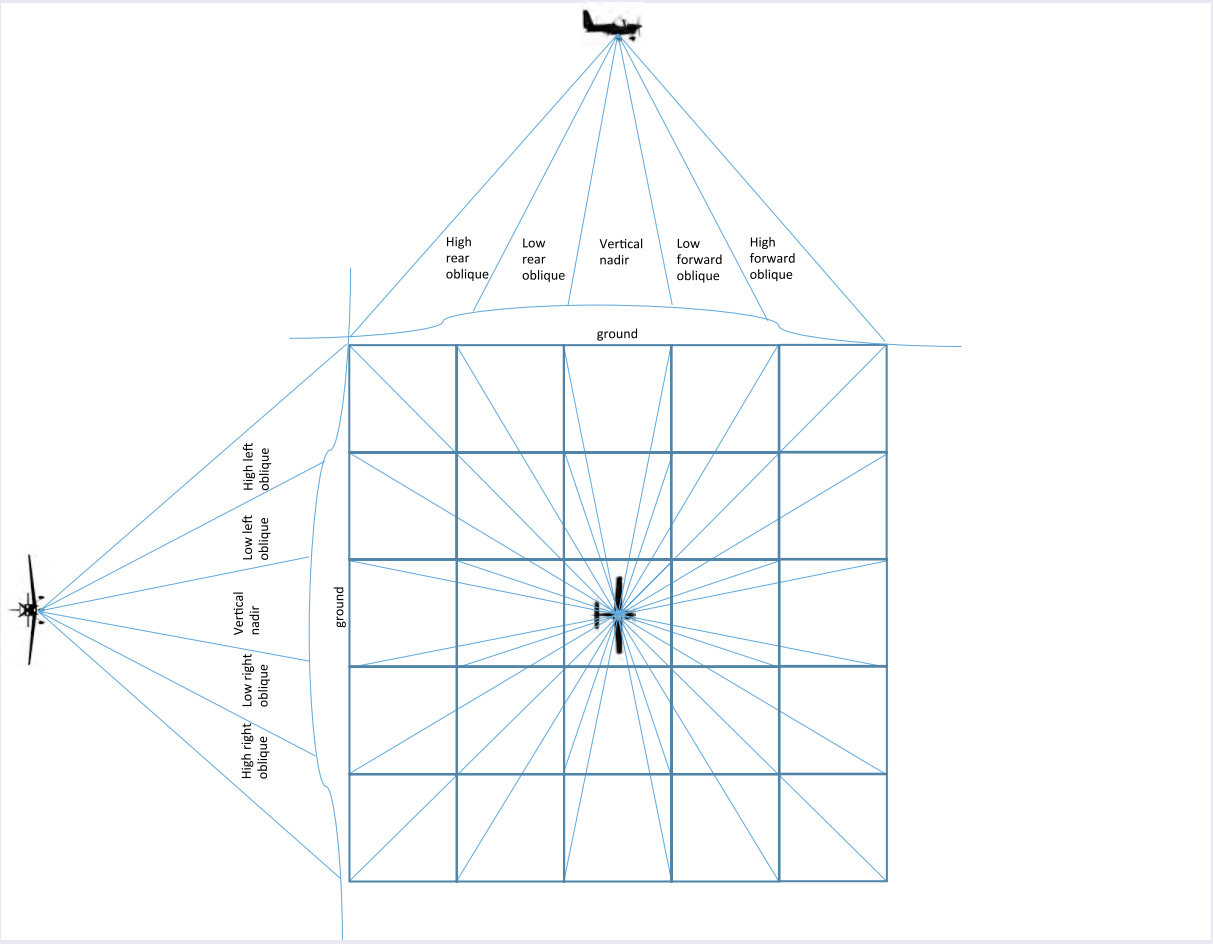
- multi-angle viewing for image capture (see Figure 2.20):
  - vertical nadir, low oblique and high oblique;
  - from left to right, and fore to aft, relative to flight direction; and
  - overlapped image frames;
- modular and transportable imagery capture pods that can be mounted underneath different 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 revolutionary 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.

Spookfish aims to use these capabilities to 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

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.

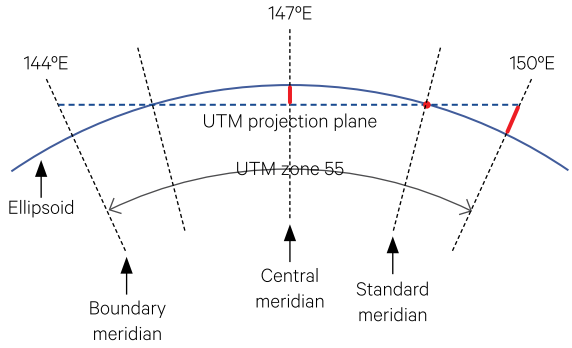
Background image: SPOT-1 image of Great Sandy Desert, central Australia, acquired in 1986. **Source:** © CNES 1986, Distribution Airbus DS. Retrieved from: http://www.intelligence-airbusds.com/en/5751-image-detail.php?img=1505#.Wq9UwmO8z8s

3.1.1.1 Local Transverse Mercator (LTM)

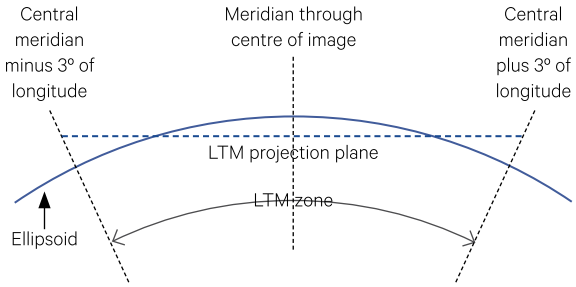
For a limited area on the Earth's surface, a reasonable two-dimensional representation of the area's geometry can be achieved using a suitable projection. The Universal Transverse Mercator (UTM) grid provides accurate representation of small ground areas, especially at the centre of a zone (see Section 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)

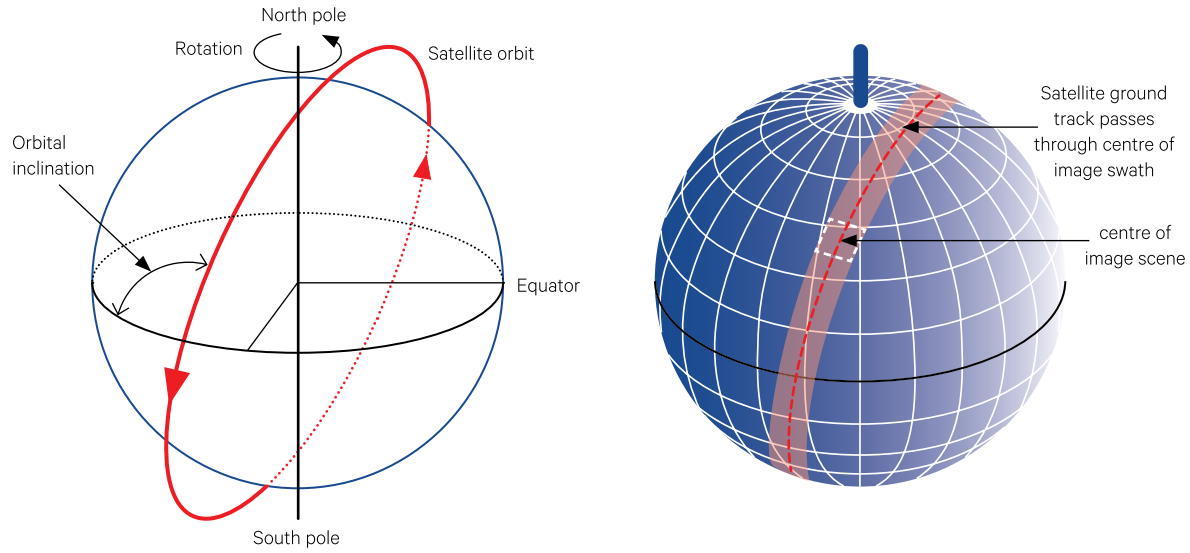
The perspective of the EO image needs to be considered when rectifying continental scale imagery such as MODIS and AVHRR. The satellite orbital path can be viewed as an Oblique Mercator Projection, with the orbit itself being considered as an expanded great circle encompassing the Earth. Full scenes 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:

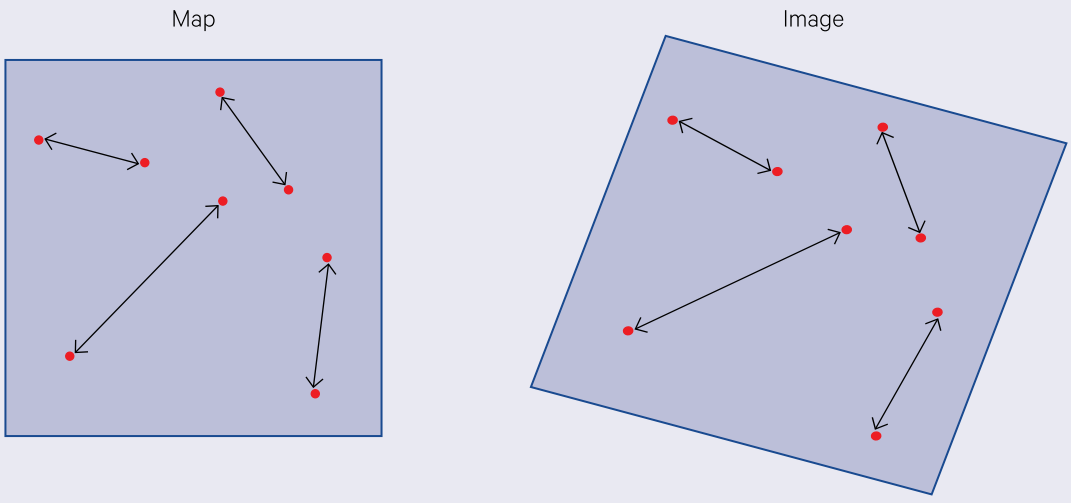
$$\frac{\text{map distance}}{\text{image distance}}$$

- 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 and :

- affine;
- bilinear (or ruled or first order polynomial);
- quadratic (or second order polynomial); or
- cubic (or third order polynomial).

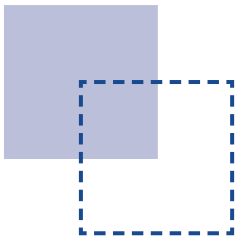
The geometric effects of each of these models are illustrated in Figure 3.4.

Wars of nations are fought to change maps.
But wars of poverty are fought to map change.
(Muhammad Ali)

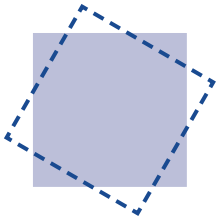
Figure 3.4 Geometric effects of polynomial models

a. An affine model changes origin (shift), orientation (rotation), scale (enlargement or reduction) and one-directional linear skew (in X or Y direction).

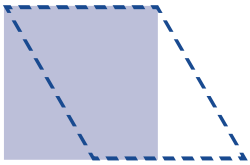
Shift of origin



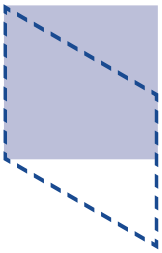
Rotation of axes



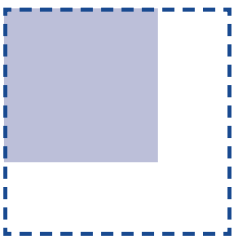
Skew in X direction



Skew in Y direction

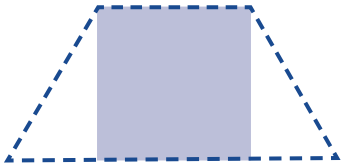


Scale change



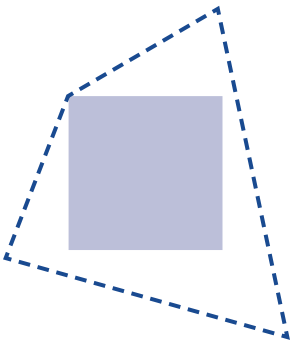
b. Bilinear model: changes origin, orientation, scale and two-directional, linear skew

Skew in X direction with changes in X and Y

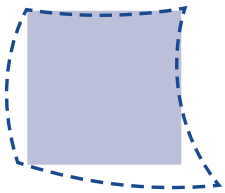


Skew in X direction with changes in X and Y

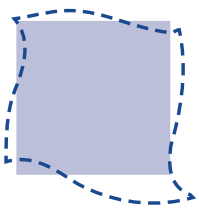
Skew in X and Y directions with changes in X and Y



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. As such, this form of resampling is often used for image hardcopy. Once an image is in hardcopy form, the other affine transformation functions of shift and rotation can be implemented by manually registering the image over an appropriately scaled map. Along-line 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.

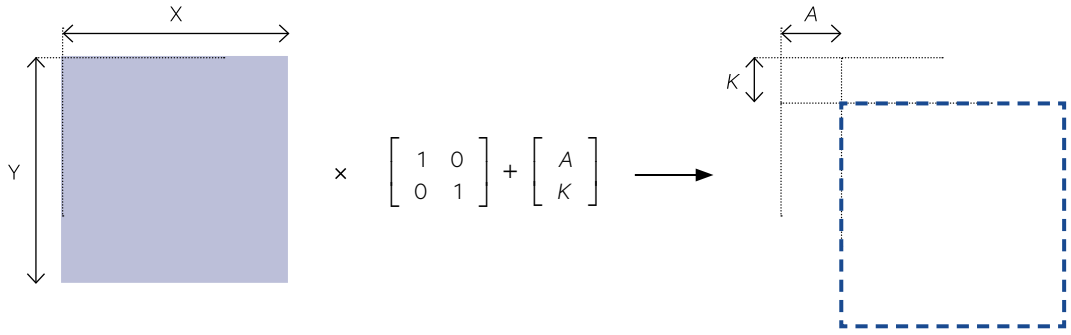
3.2.1.1 Shift in X and/or Y direction

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

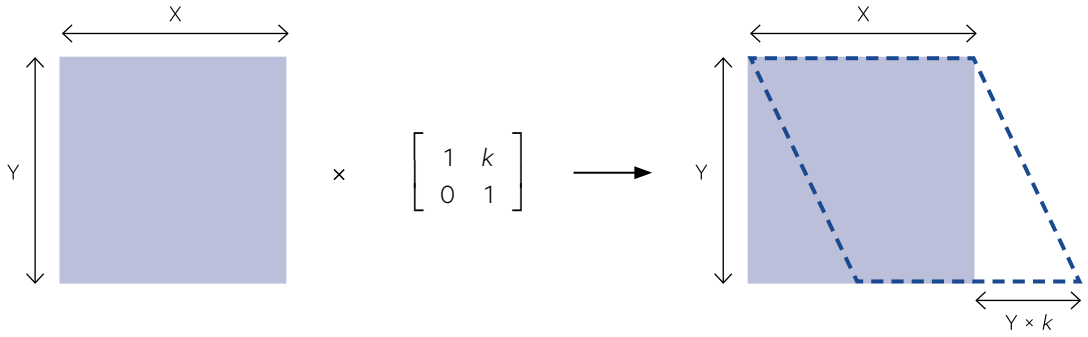
$$\begin{aligned} X' &= 1 \times X + 0 \times Y \\ &= X \\ Y' &= k \times X + 1 \times Y \\ &= kX + Y \end{aligned}$$

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

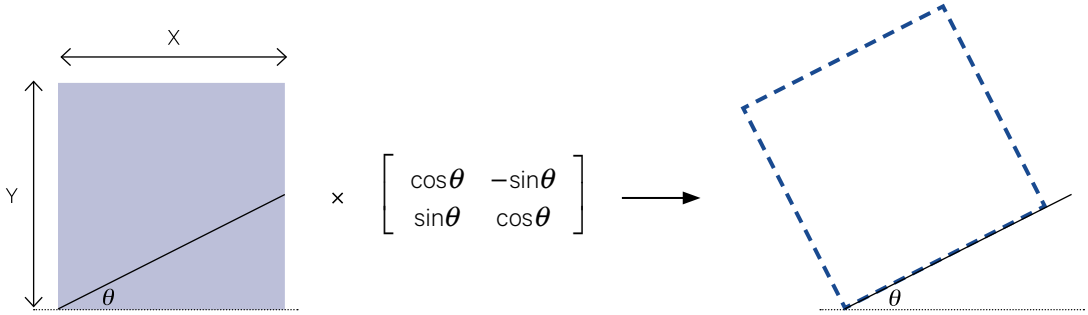
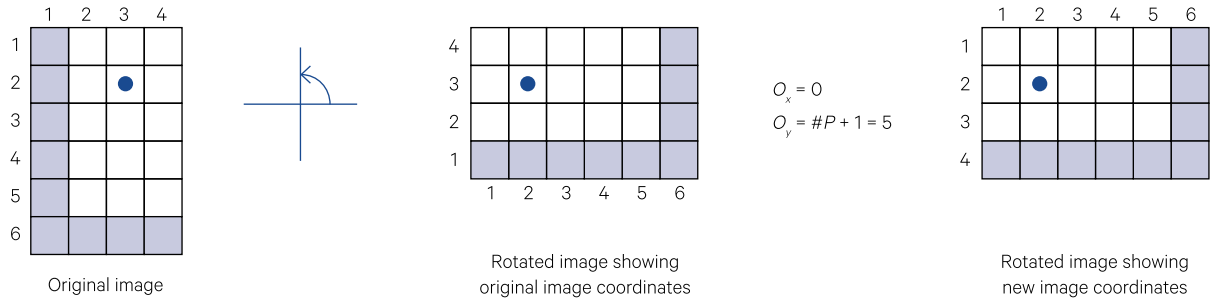
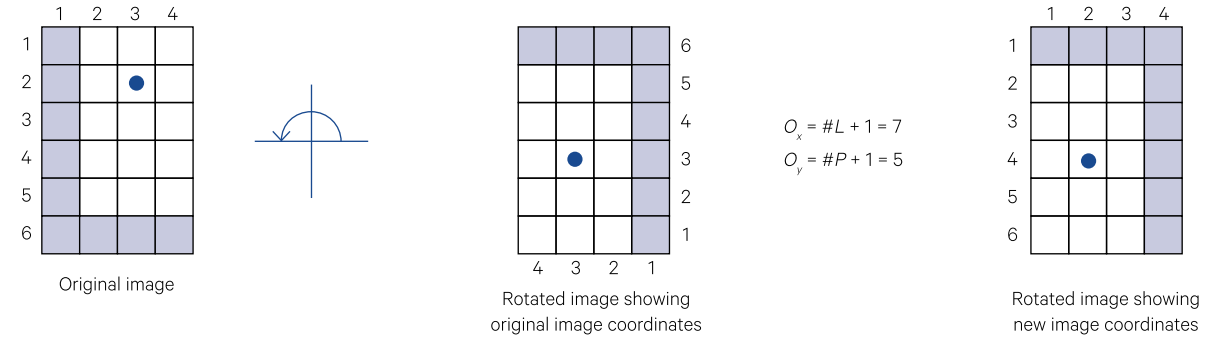


Figure 3.8 Origin adjustment for image rotation

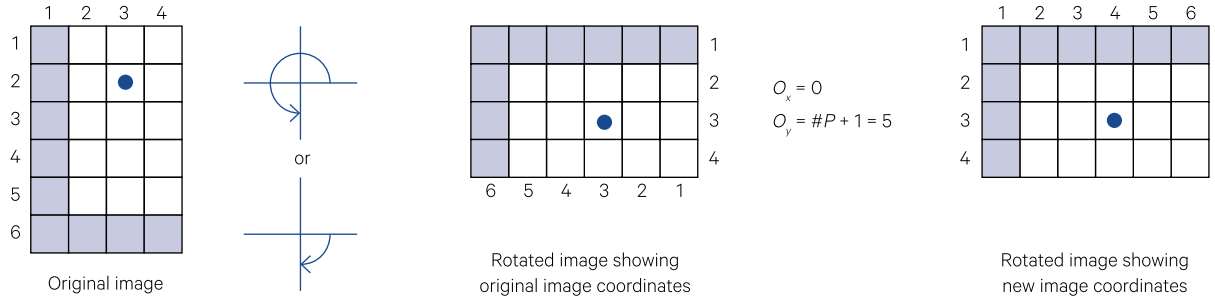
a. 90° anticlockwise rotation



b. 180° anticlockwise rotation



c. 270° anticlockwise or -90° clockwise rotation



Source: Harrison and Jupp (1990) Figure 25

3.2.1.4 Scale change

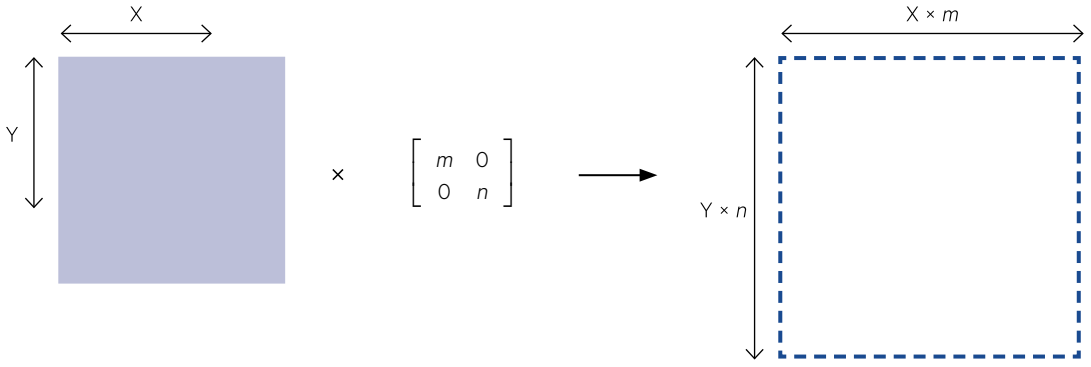
A change in scale is calculated simply by multiplying the coordinates of the original system by an appropriate factor (see Figure 3.9). Linear rescaling of coordinate axes is simply implemented by the matrix:

$$\begin{bmatrix} m & 0 \\ 0 & n \end{bmatrix}$$

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,

$$X' = 0 + m \times \left(X - 0\right) + 0 \times \left(Y - 0\right)$$
$$Y' = 0 + 0 \times \left(X - 0\right) + n \times \left(Y - 0\right)$$

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

$$X' = X \times \cos 2\theta + Y \times \sin 2\theta$$
$$Y' = X \times \sin 2\theta + Y \times -\cos 2\theta$$

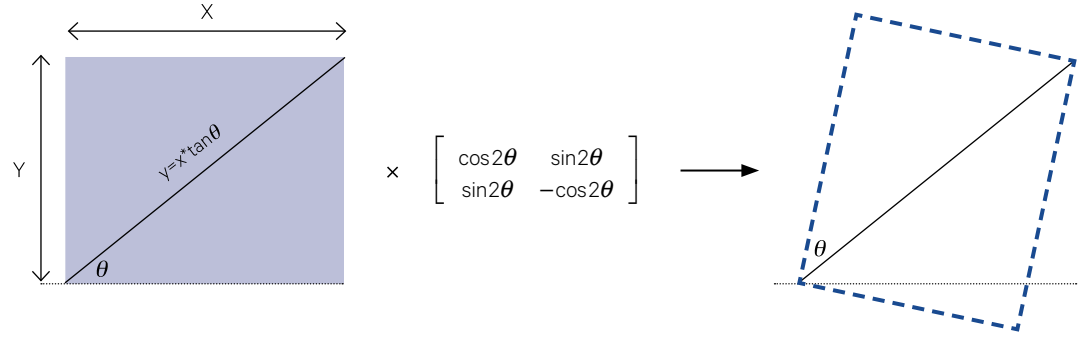
(see Figure 3.10). To reflect the coordinate system on the line Y=X (this is the same as transposing the image when *#P=#L*), $\theta = 45°$ so:

$$X' = Y$$
$$Y' = X$$

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:

$$X' = A + BX + CY + DXY$$

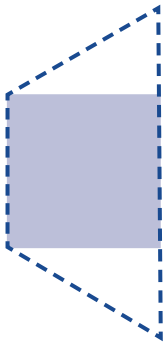
$$Y' = K + LX + MY + NXY$$

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).

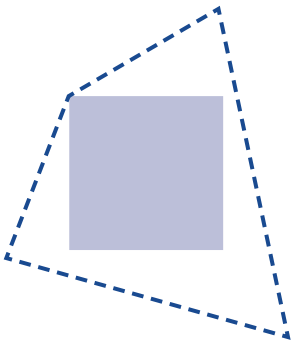
Figure 3.11 Bi-directional skew

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

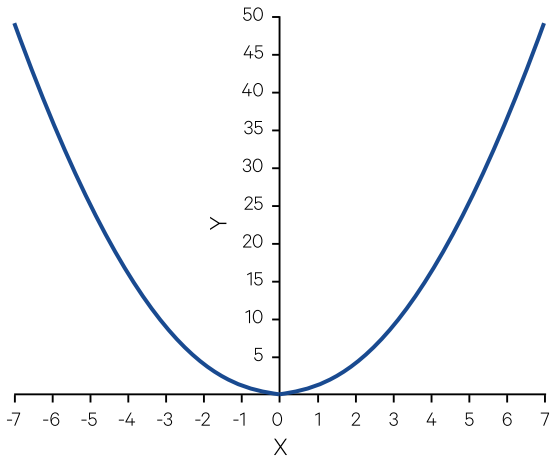
$$X' = A + BX + CY + DXY + EX^2 + FY^2$$

$$Y' = K + LX + MY + NXY + OX^2 + PY^2$$

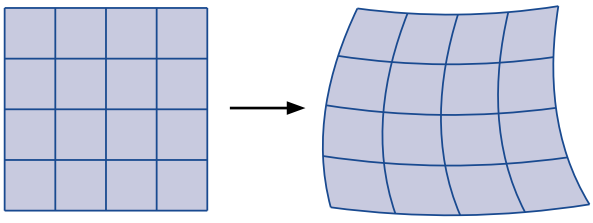
Thinking back to high school maths, a quadratic function is graphed as a parabola (see Figure 3.12a). This shape is characterised by a point at which the slope of the function line changes from being positive to negative. This characteristic is due to the squared terms (X^2 and/or Y^2) which vary the 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:

$$X' = A + BX + CY + DXY + EX^2 + FY^2 + GX^3 + HX^2Y + IXY^2 + JY^3$$

$$Y' = K + LX + MY + NXY + OX^2 + PY^2 + QX^3 + RX^2Y + SXY^2 + TY^3$$

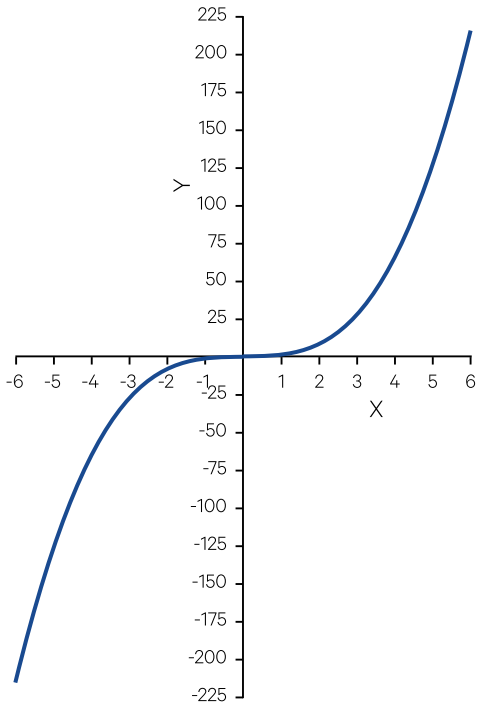
A cubic function is graphed as a line with two points of inflection, that is, points where the slope changes from positive to negative as illustrated in Figure 3.13a. This feature allows both the direction and rate of 'bending' of image lines and pixel columns to vary twice throughout the image (see 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.

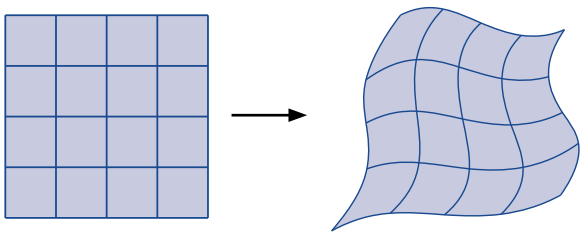
The cubic model can account for 'stronger' warps in the data but, by doing so, is also quite unstable and cannot be reliably applied away from the section of the image from which it is derived. The flexibility of higher order models is countered by their more limited applicability and, to achieve reliable 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:

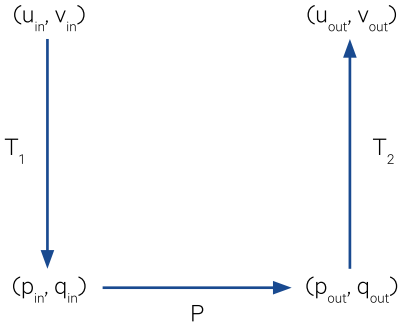
- T_1, a nominal coordinate transformation to convert (u_{in},v_{in}) to (p_{in},q_{in});
- P, a polynomial transformation to convert (p_{in},q_{in}) to (p_{out},q_{ou}t); and
- T_2, another nominal transformation to convert (pout,qout) to (uout,vout) (see Figure 3.14).

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

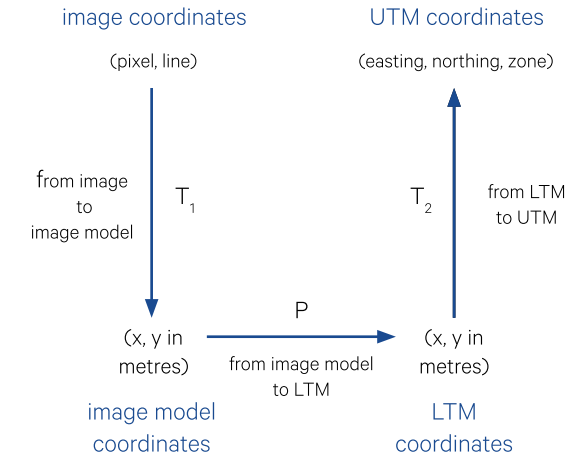
P is a polynomial transformation from (p_{in}, q_{in}) to (p_{out}, q_{out})
T_1 is a nominal model which converts from (u_{in}, v_{in}) to (p_{in}, q_{in})
T_2 is a nominal model which converts from (p_{out}, q_{out}) to (u_{out}, v_{out})



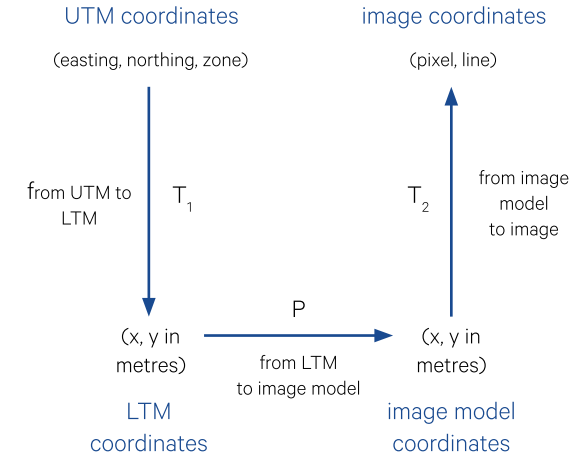
Source: Harrison and Jupp (1992) Figure 29

Figure 3.15 Three-stage registration model

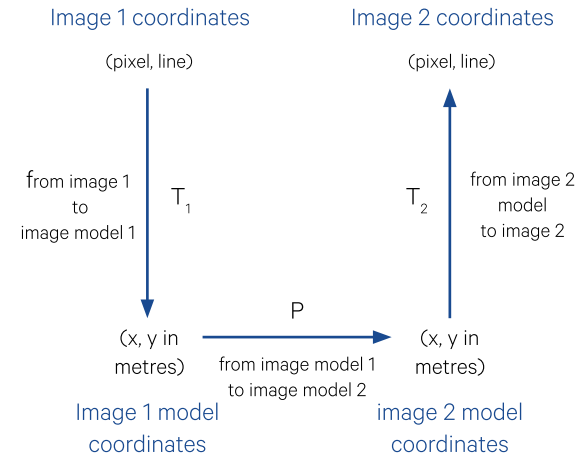
a. Converting from image to map



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 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 Hotine Oblique Mercator coordinates (pout,qout) in metres to longitude and latitude (uout,vout) 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 Hotine Oblique Mercator 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.

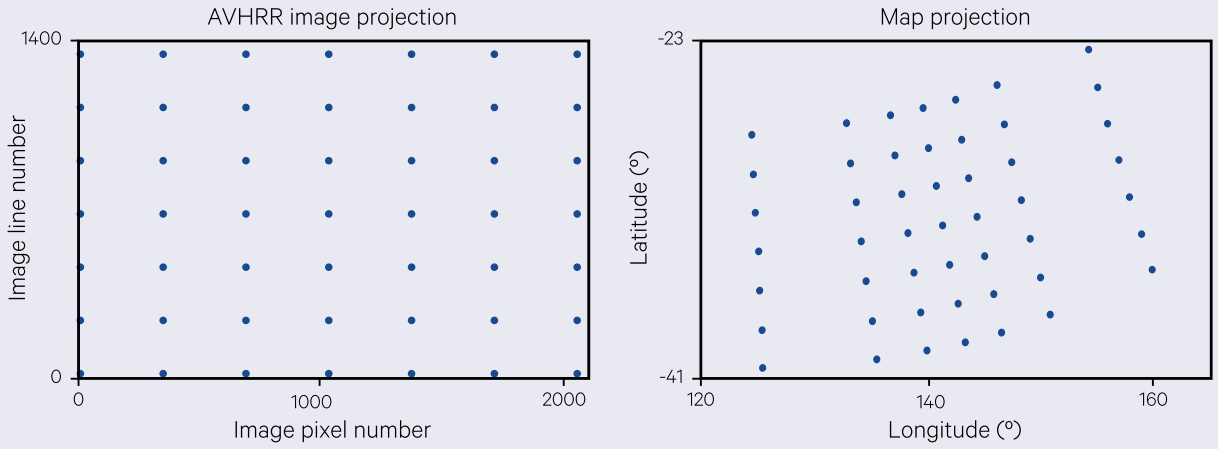
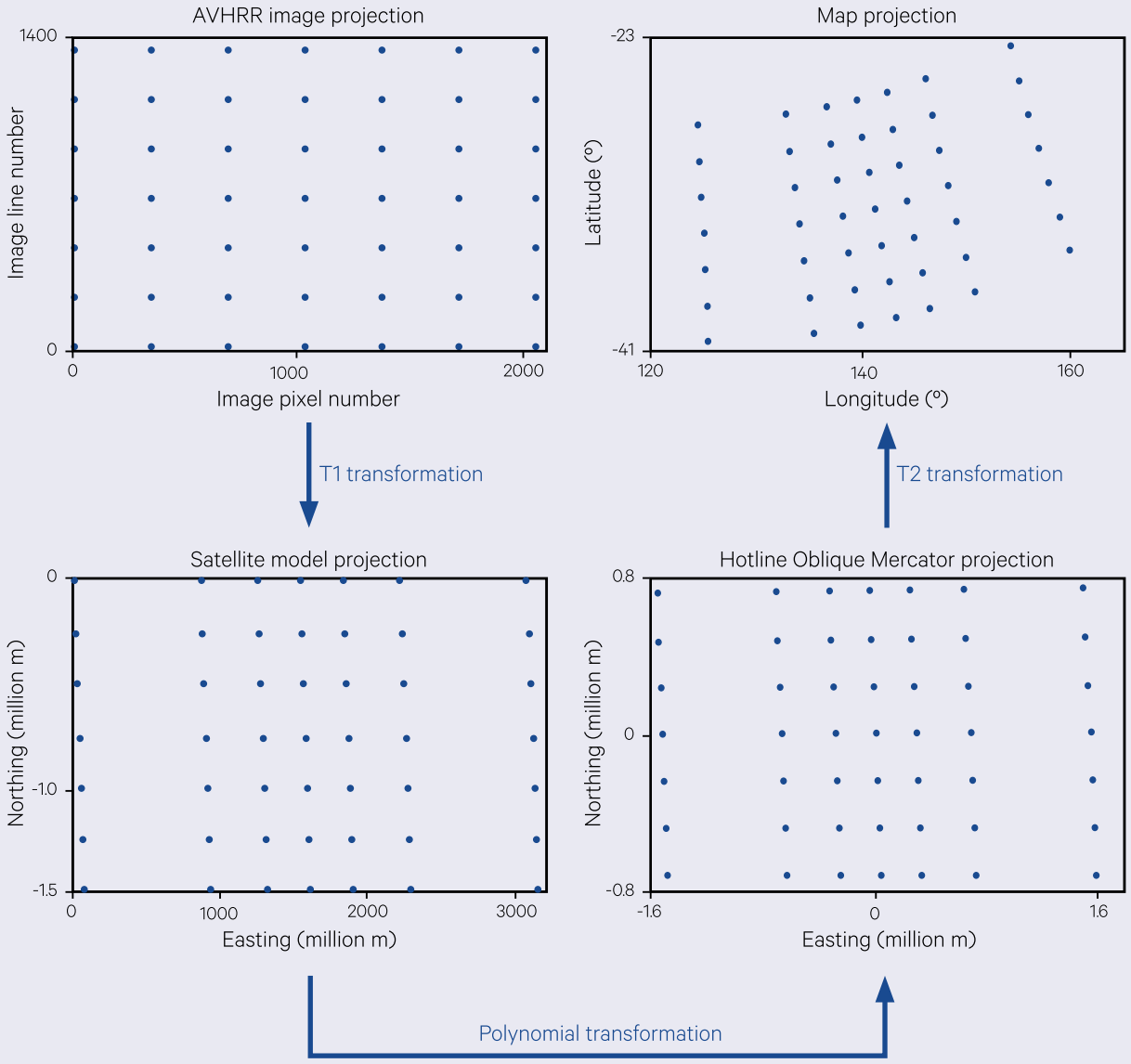


Figure 3.17 Three-stage conversion from AVHRR image to map coordinates



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*

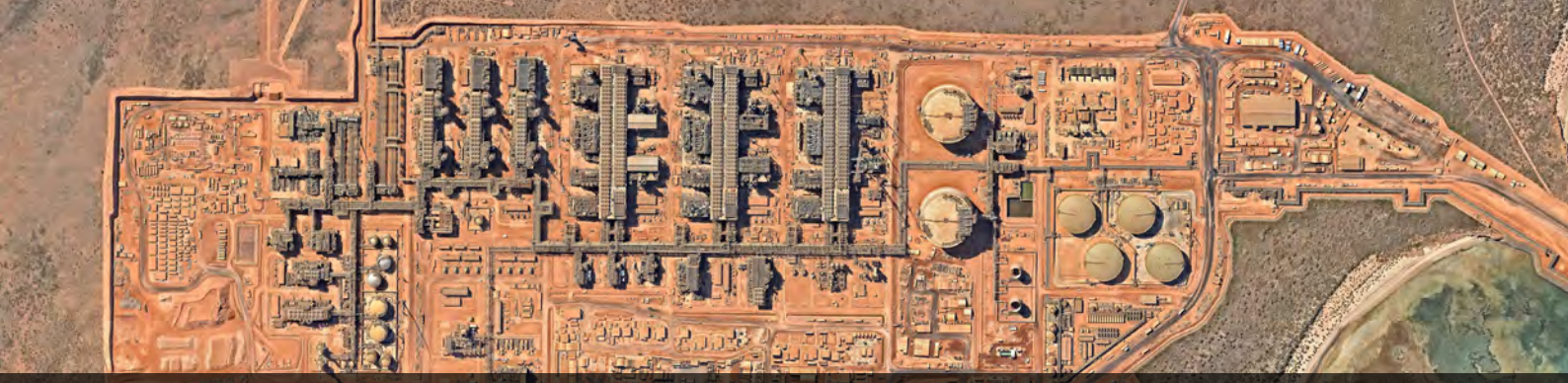
3.5 References

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 such 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

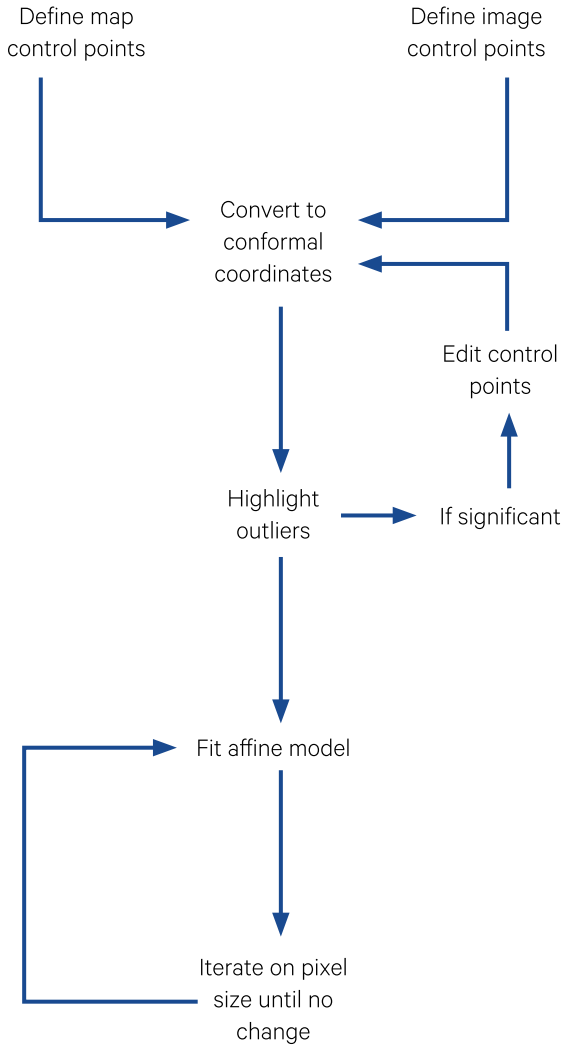
Source: a. Extracted from Whereis map of Canberra. b. Extracted from Google Maps in August 2018.

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. 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. Also particular information that needs to be compared with, or in the case of digitising to be incorporated with, the image data may not be available on an accurate mapping base. Provided a sufficient number of control points may 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:

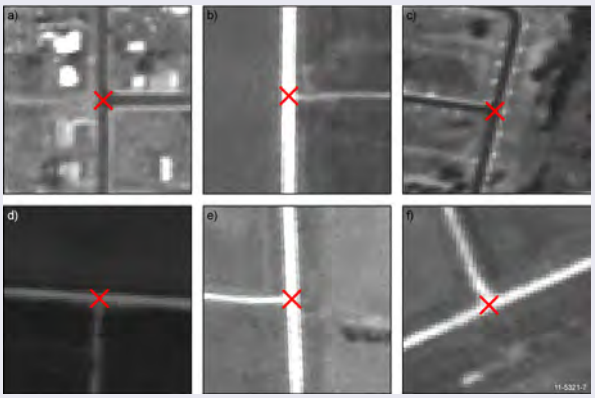
- narrow roads with intersection angles of at least 50° (see Figure 4.3 and Figure 4.4);
- centre points of quasi-circular features; and
- road-watercourse junctions.

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

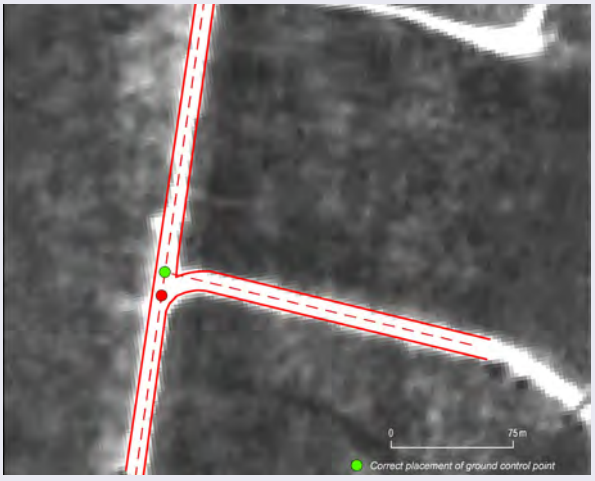
Some control point locations shown on PRISM imagery, which were used by surveyors to locate proposed features.



Source: Lewis *et al.* (2011) Figure 9

Figure 4.4 Precise placement of control point at road intersection

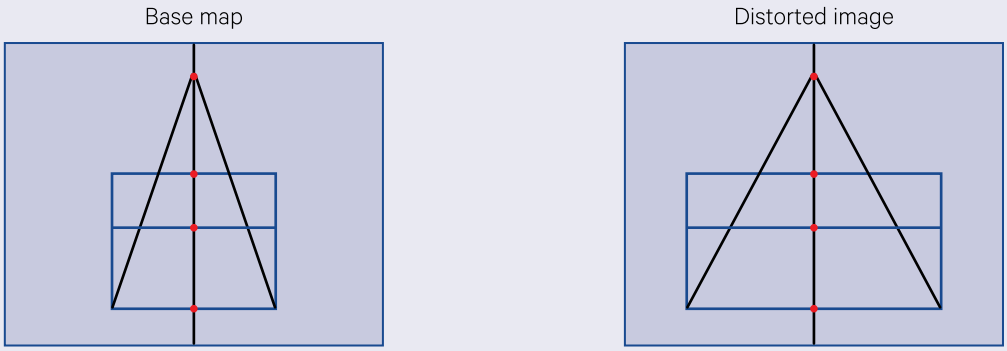
In this image solid red lines mark the edge of the linear features (roads) and dashed red lines mark the centrelines. The correct placement of a control point at this intersection is shown as a green dot. The red dot shows an incorrect placement of the control point.



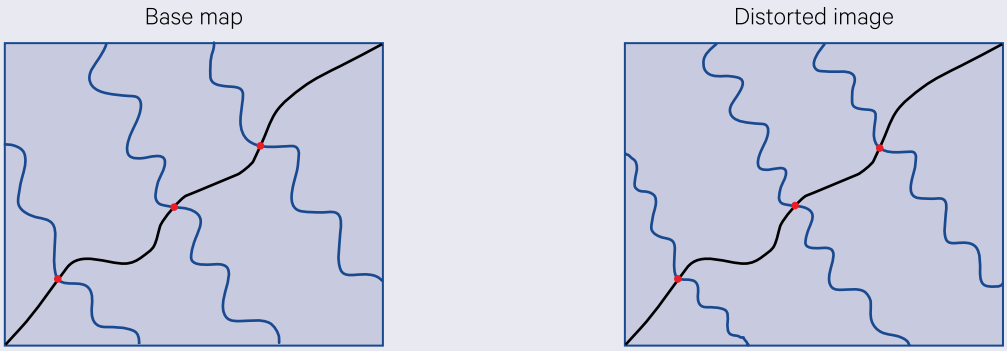
Source: Lewis *et al.* (2011) Figure 10

Figure 4.5 Control point distribution

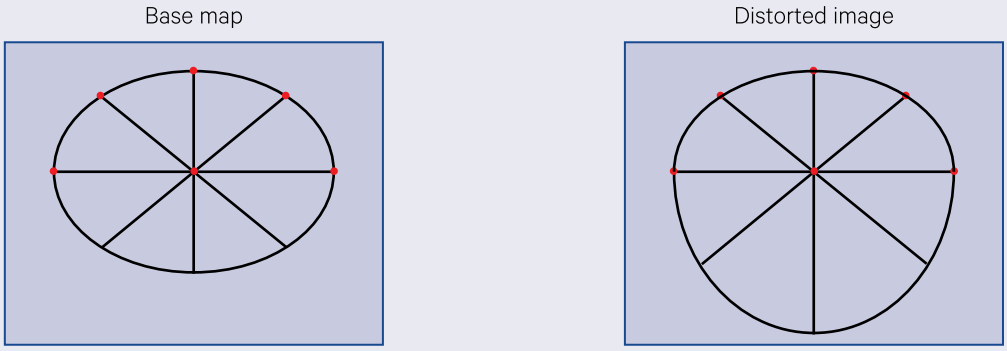
a. GCPs (marked •) are not well-distributed across the image, so cannot correct for the along-line distortions.



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 numbered in the same order 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 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.

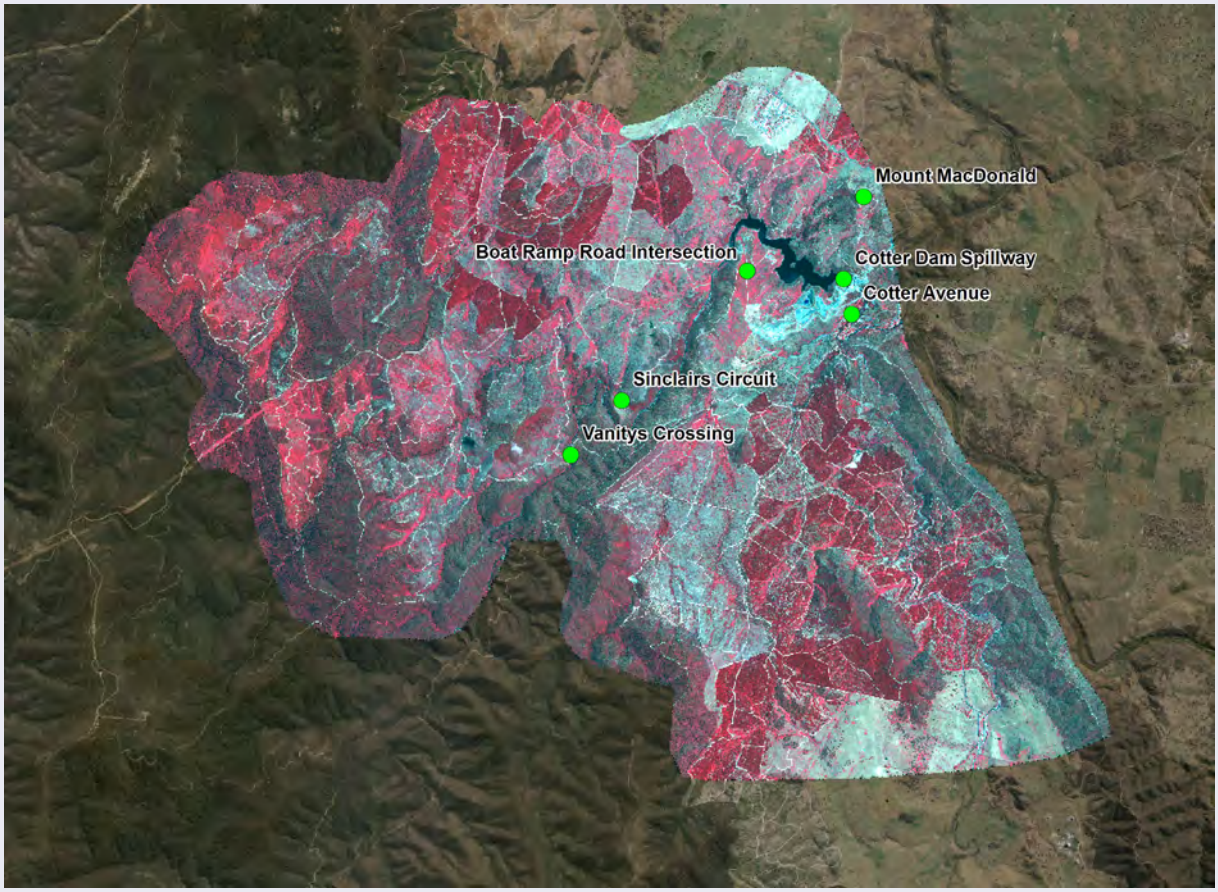


Table 4.1 Verification errors in selected GCP

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.

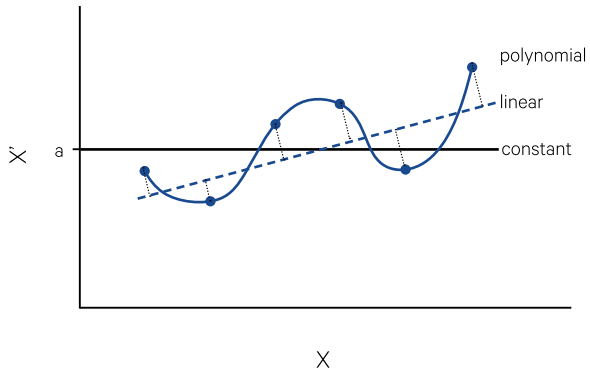
4.2 Fitting Arbitrary Models

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).

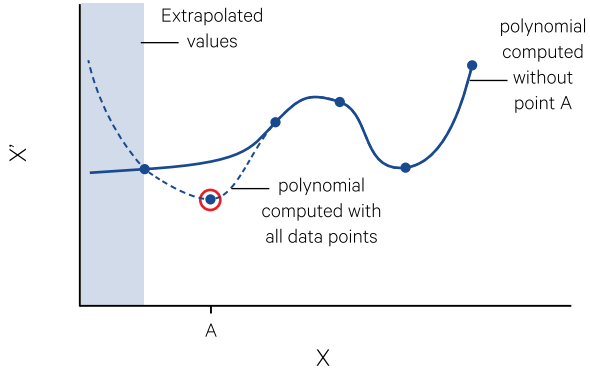
When fitting any mathematical model to a set of sample data points, the selected sample needs to be representative of the whole dataset before the model can be reliably applied to it. For modelling the geometric relationship between two sets of spatial data this requires that the sample locations be 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

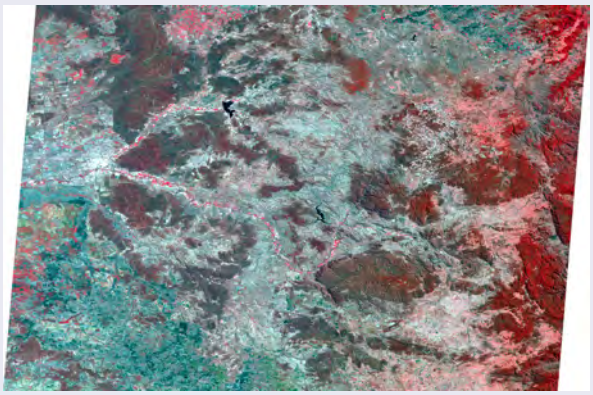
Two adjacent Landsat MSS scenes, acquired in September 1982 and shown in Figure 4.10, show the characteristic distortions due to Earth rotation skew that occur satellite imagery acquired in descending mode (see Section 2). They also demonstrate the 'squashed' vertical extent of unrectified Landsat MSS 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.

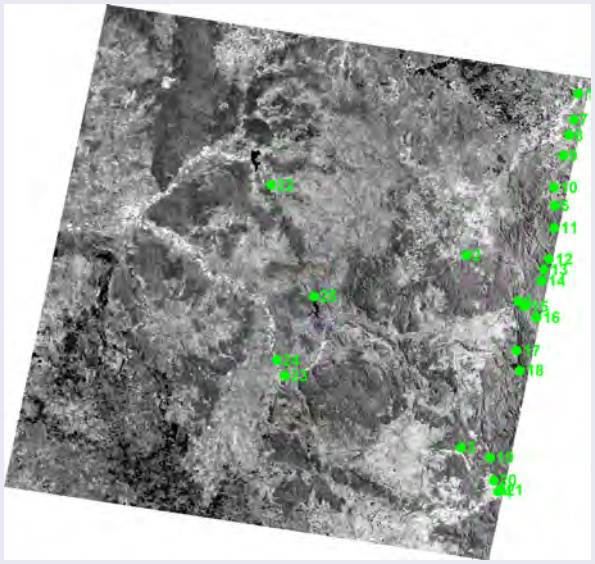
a. Path 90 Row 80



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

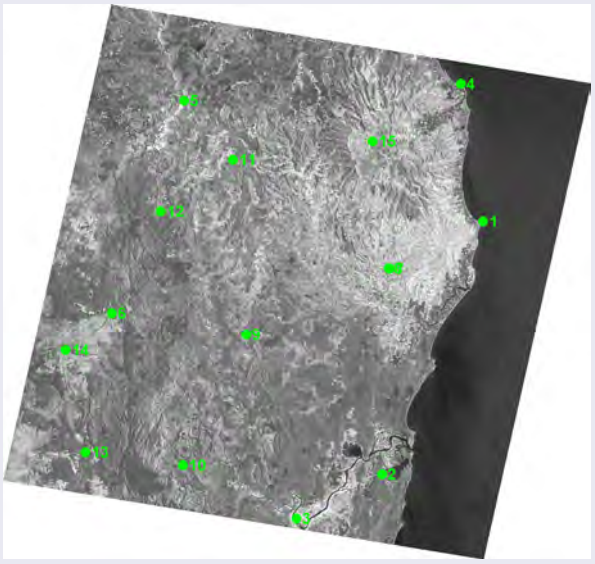


Table 4.2 15 GCP used to rectify path 89 scene

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

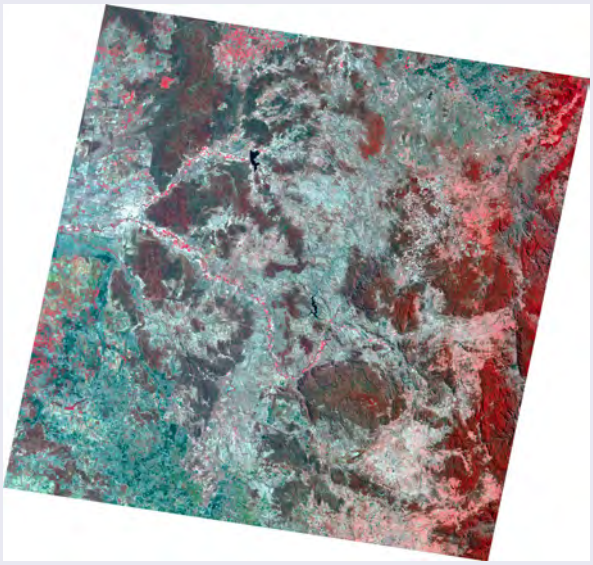
  C = 0.01741483 × E − 0.00265257 × N + 13,810.21933627

  R = −0.00181708 × E − 0.01206504 × N + 83,867.38393856

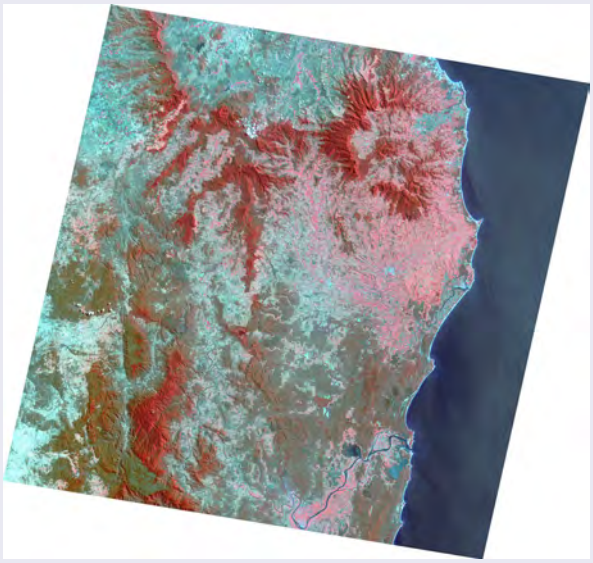
Figure 4.12 Resampled images

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.

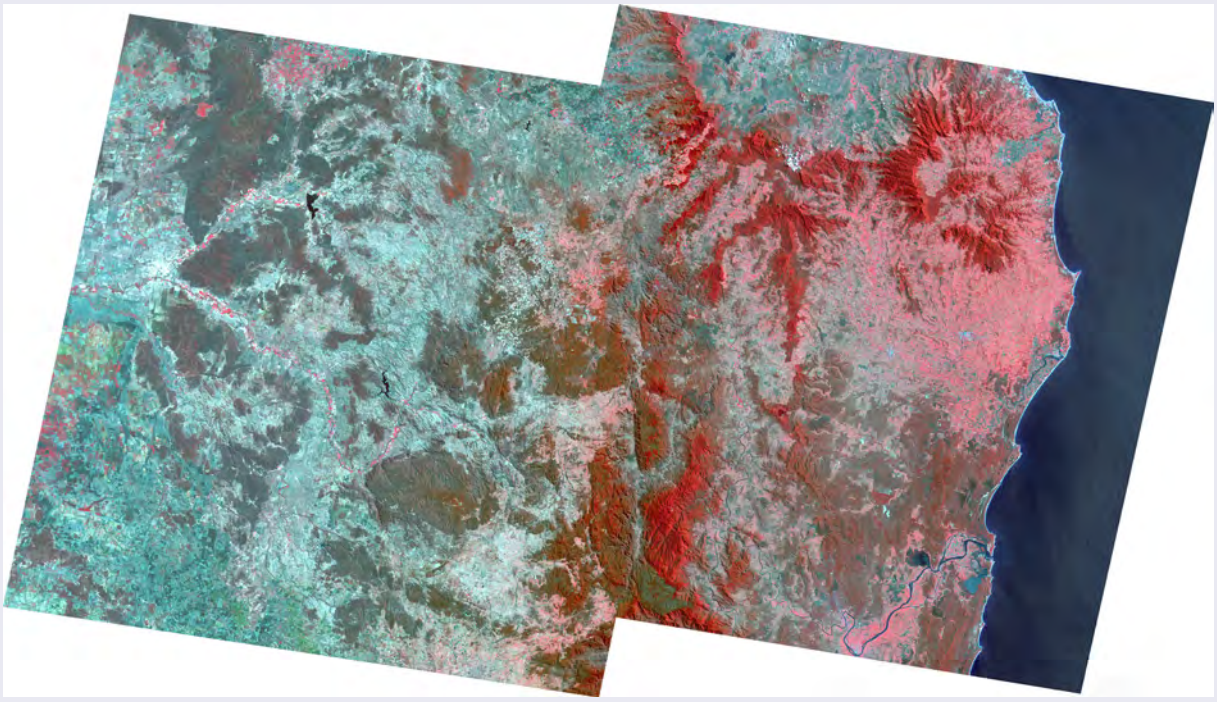
a. Path 90



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.

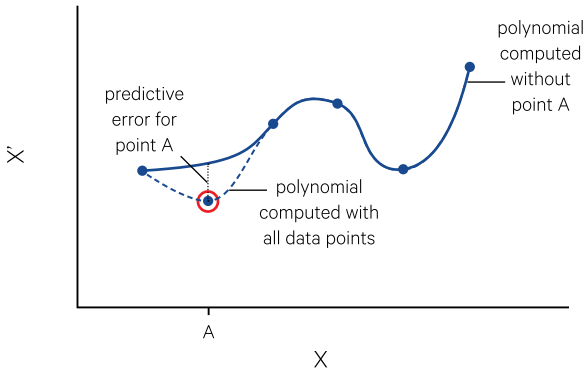
Conquest is easy. Control is not.
(William Shatner)

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. When all transformations can be represented by a simple affine transformation then, were transformation exactly known in one direction (such as from dataset 1 to dataset 2), it could be inverted to obtain the corresponding transformation in the other direction (namely, 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.

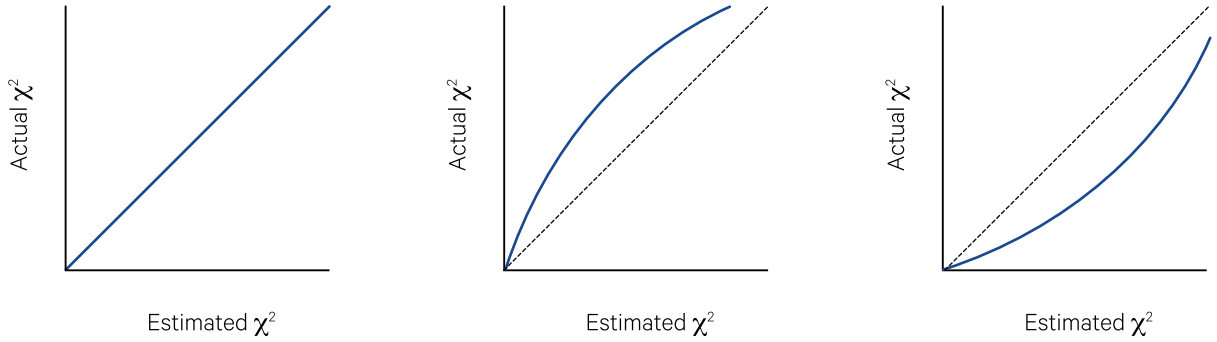
Error statistics can be determined for each point. For example, a predictive error weight value would indicates how important a point is in the current model—a new model computed without that point would potentially give an error of this value in predicting the point's position. A point is important 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 by 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.

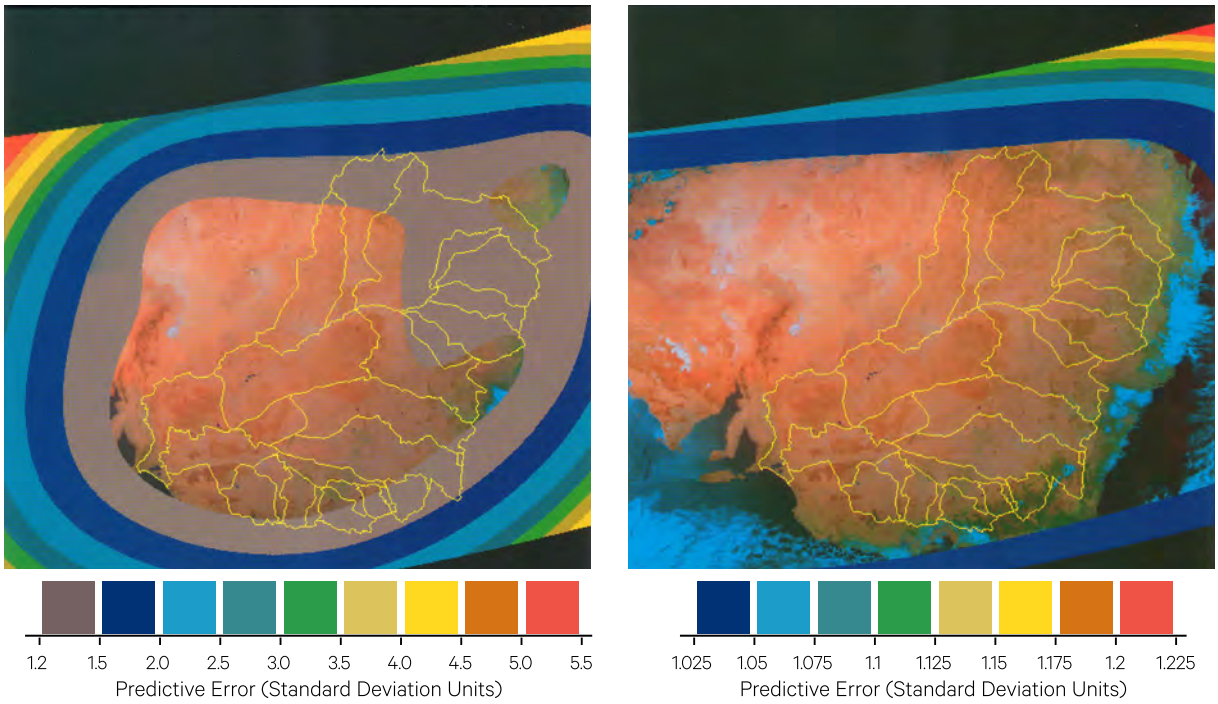
The predictive variance of a polynomial model, which converts from an image to some other geometry, can also be computed for each pixel, not just the GCP (see Volume 2X—Appendix 7). In this case, an output channel could be written with scaled values that indicate the predictive variance, in units of 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. 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.

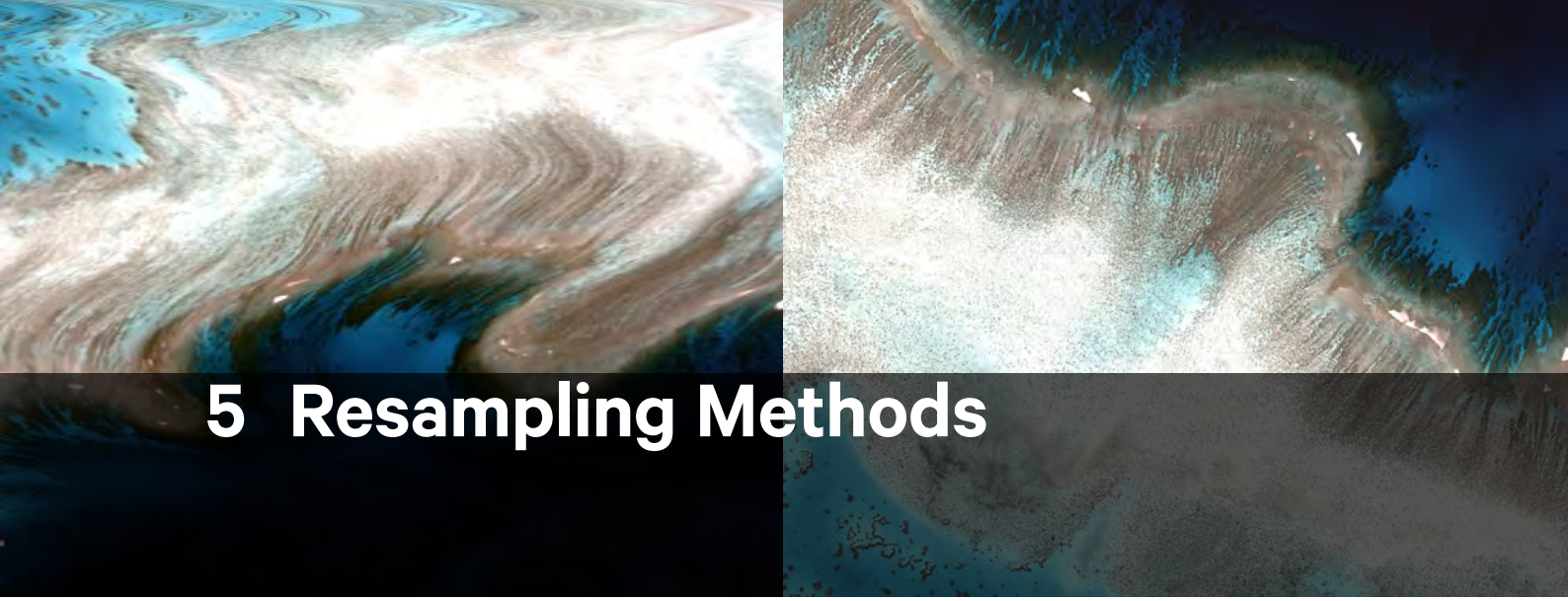
Richards, J.A. (2013). *Remote Sensing Digital Image Analysis: An Introduction*. 5th edn. Springer-Verlag, Berlin. ISBN 978-3-642-30061-5

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

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

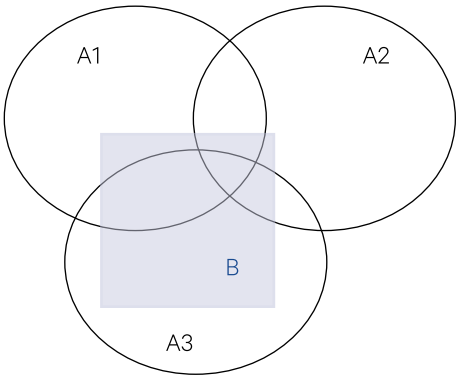
When an image is changed to have a different geometry, the resolutions of the input and output images need to be considered carefully. EO pixel data represent radiance integrated over a particular volume of the Earth's surface. As summarised in Section 2, different pixels in the image can represent volumes 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.

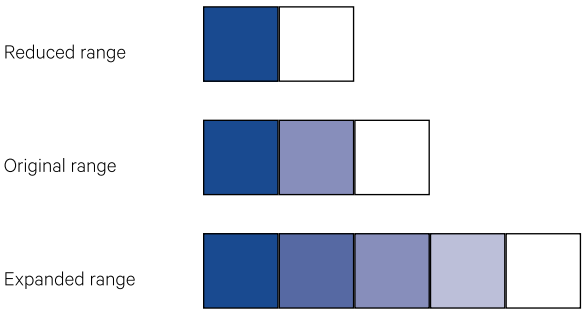
An example of interpolation is shown in Figure 5.2 using a one-dimensional dataset. Interpolation effectively attempts to convert a discrete sample set into a continuous function. In image resampling, this involves fitting a continuous function through selected sample pixels. This process implicitly 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. *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).

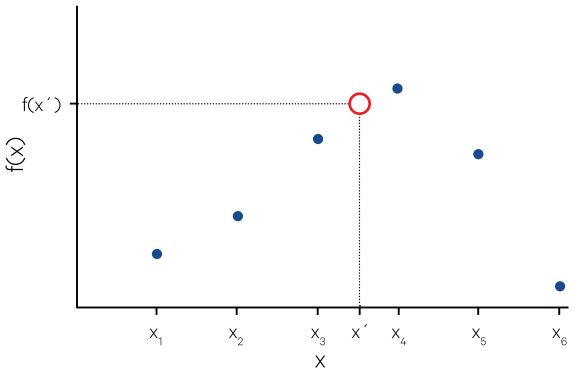
Imagery is powerful. Imagery is provocative—satellite imagery much more so because it is from space, and it allows us to get this perspective that we don't have to have otherwise.
(Sarah Parcak)

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.

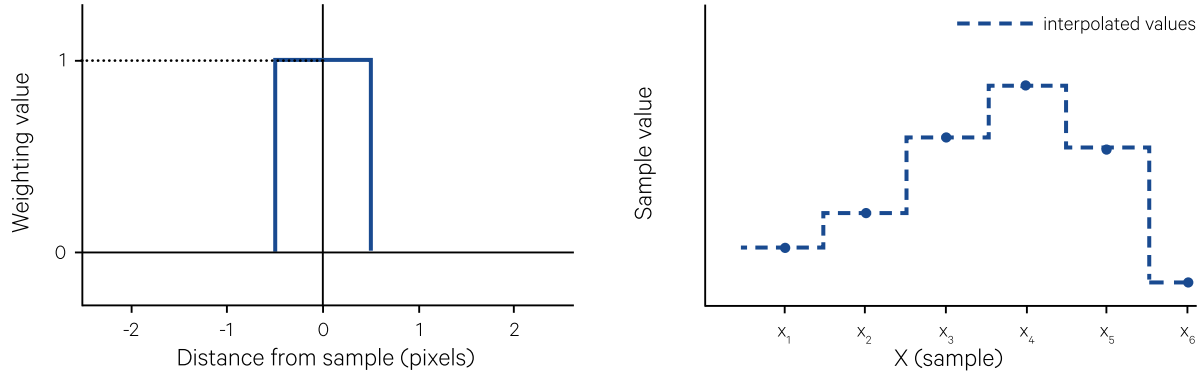


Adapted from: NIblack (1986) Figure 68

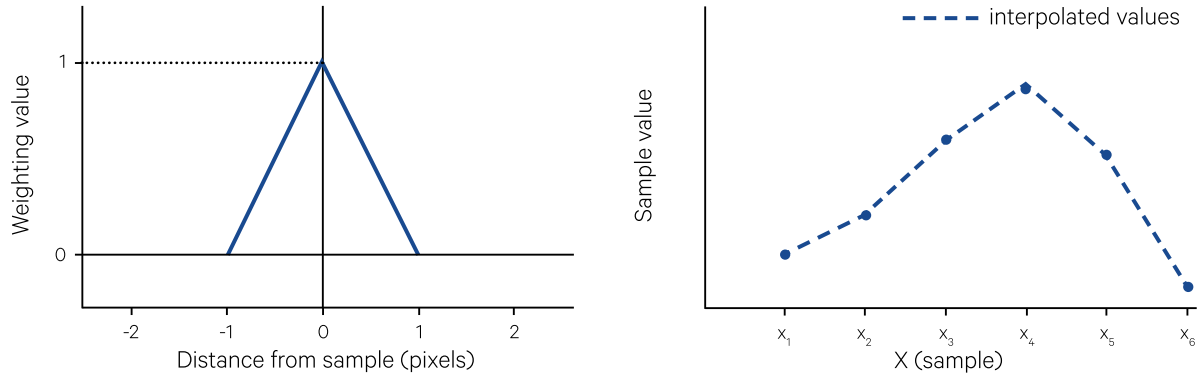
Figure 5.3 Interpolation functions for resampling methods

Using the one-dimensional dataset from Figure 5.2, the weighting functions (left) and interpolation results (right) are illustrated for three resampling methods.

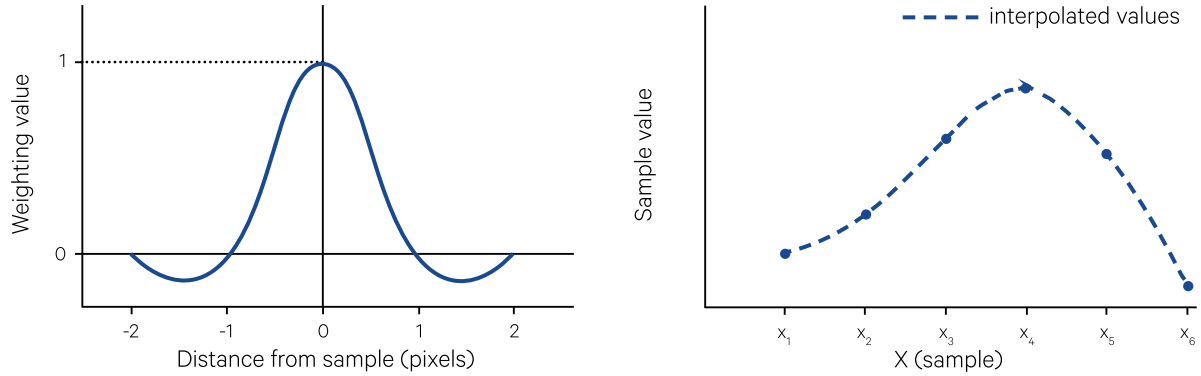
a. Nearest neighbour resampling



b. Bilinear interpolation resampling



c. Cubic convolution resampling

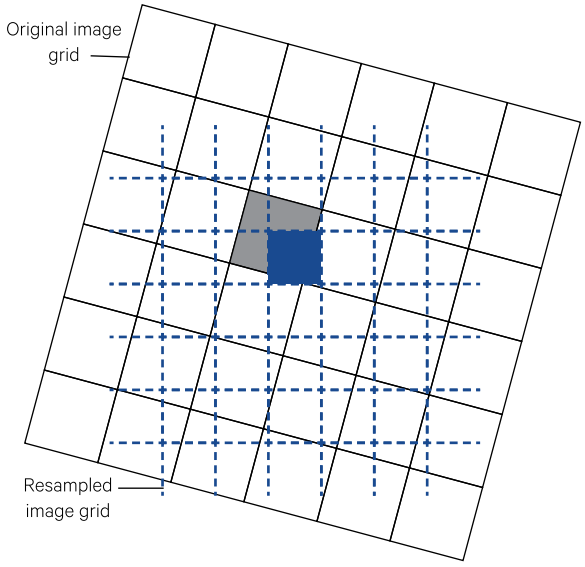


Adapted from: Niblack (1986) Figure 69

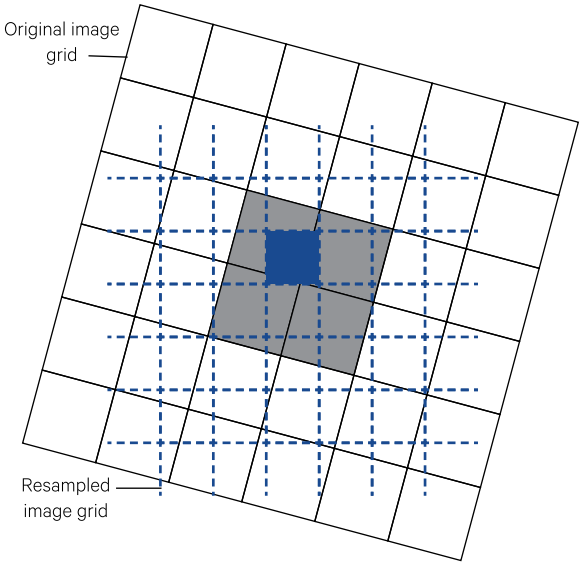
Figure 5.4 Resampling methods

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.

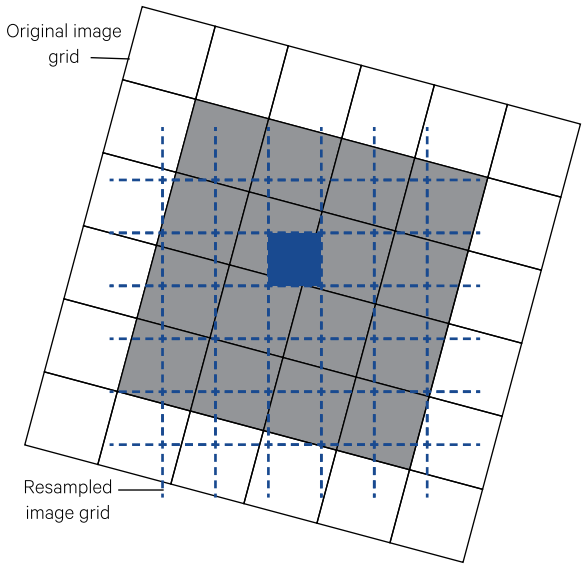
a. Nearest neighbour resampling uses the nearest input pixel value



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.



Source: *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. Difference between bilinear and cubic convolution resampled images

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:

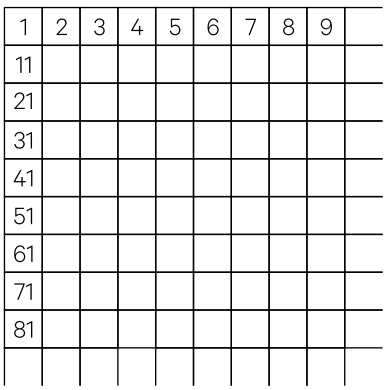
- omitting or duplicating lines;
- omitting or duplicating pixels along the line; and/or
- offsetting lines to account for image skew (see Figure 5.6).

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).

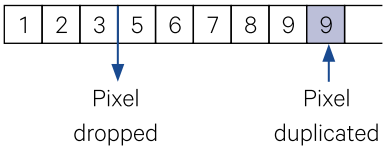
Many implementations of along-line resampling rely heavily on satellite or scanner models to remove along-line distortions (see Section 3.1.2). Additional parameters may also be required for scale adjustments and modelling specific image distortions, such as pixel size and image skew. These parameters 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

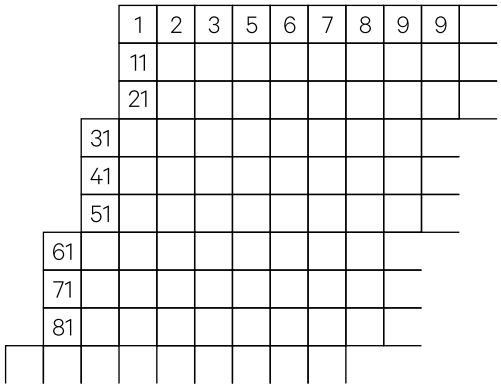
a. Original Image is labelled with sequential pixel and line numbers.



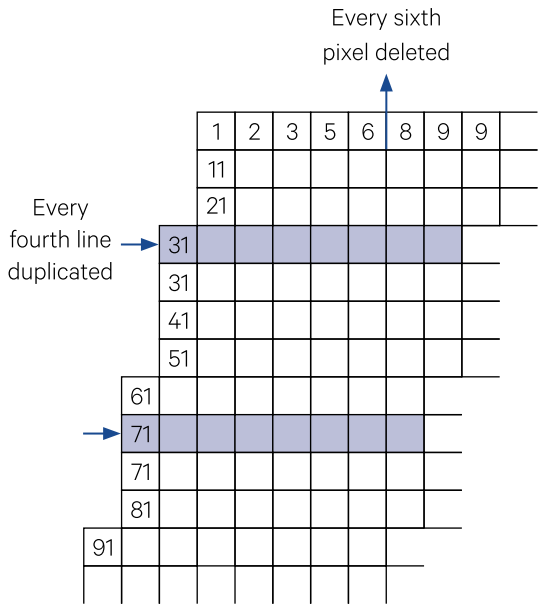
b. To rectify along-line distortions may require that pixel 4 is dropped out of the line and pixel 9 is duplicated.



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

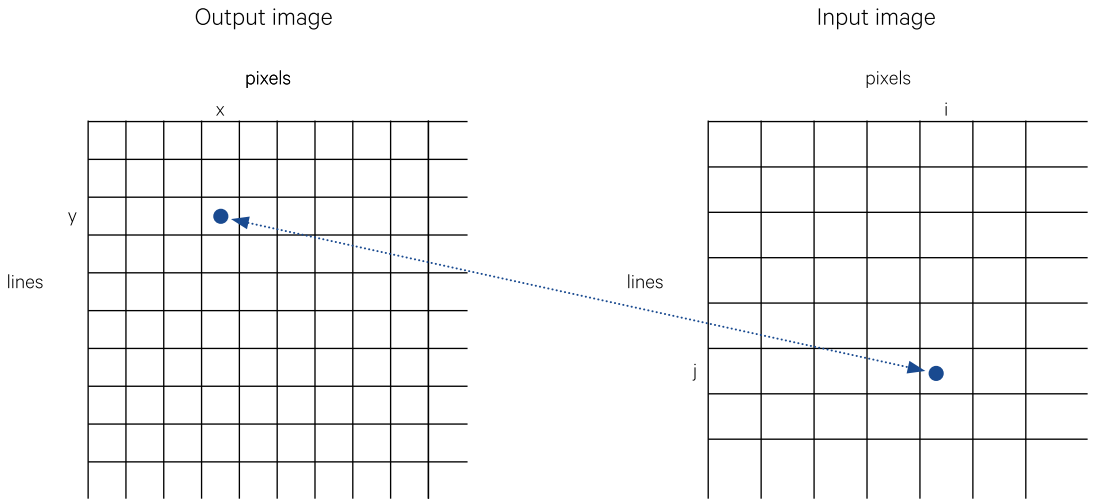
The nearest neighbour resampling method computes the location of the nearest input image pixel to each output image grid cell, then transfers the selected channel value(s) of the input pixel to the corresponding output pixel (see Figure 5.4a and Figure 5.7). This transfer process is also referred to 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.

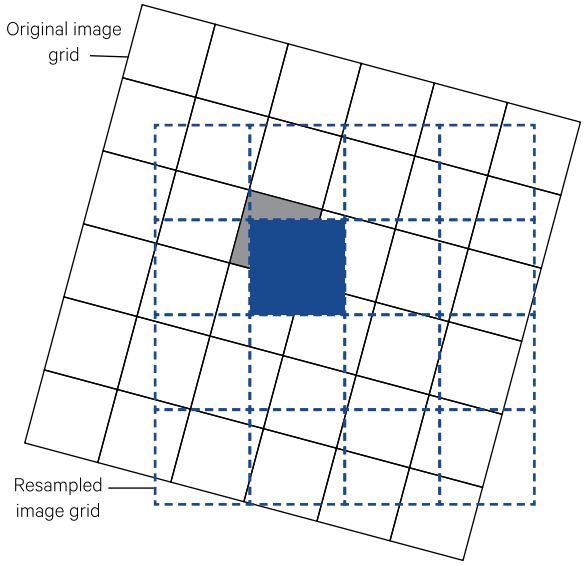
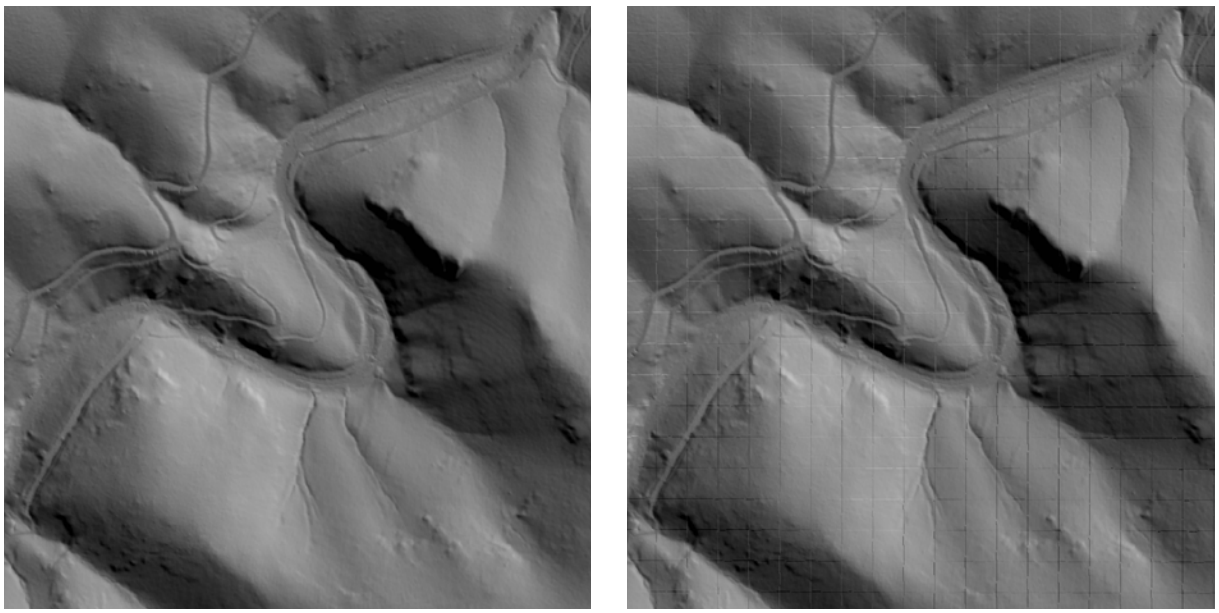


Figure 5.9 Potential artefacts from nearest neighbour resampling

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 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:

$$d_x = \frac{|xR - xL|}{NPIXELS}$$

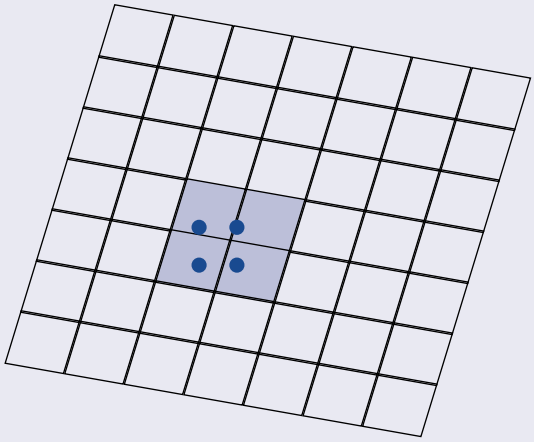
$$d_y = \frac{|yB - yT|}{NLINES}$$

in units of xL, xR, yB and yT, where

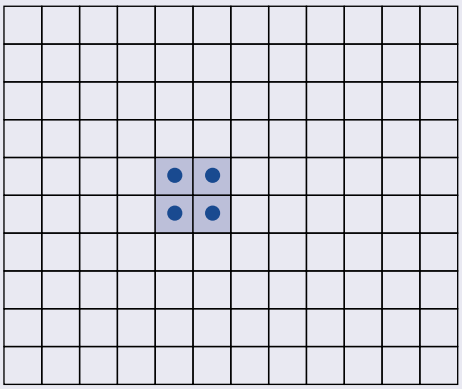
NPIXELS is the number of pixels in output line; and
NLINES is the number of lines in output grid.

Figure 5.10 Picking and binning approach

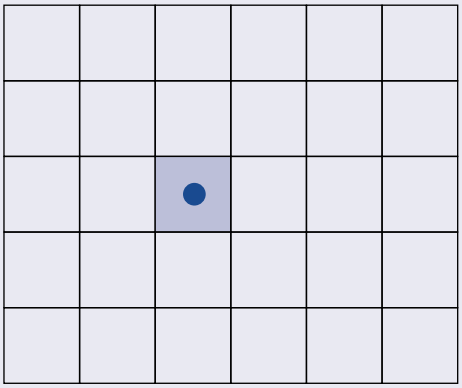
a. Original image grid showing pixels whose values will be used to create a single output pixel value



b. Placement of selected pixels on expanded grid corresponding to map scale



c. Output image with recomputed pixel value on final image grid

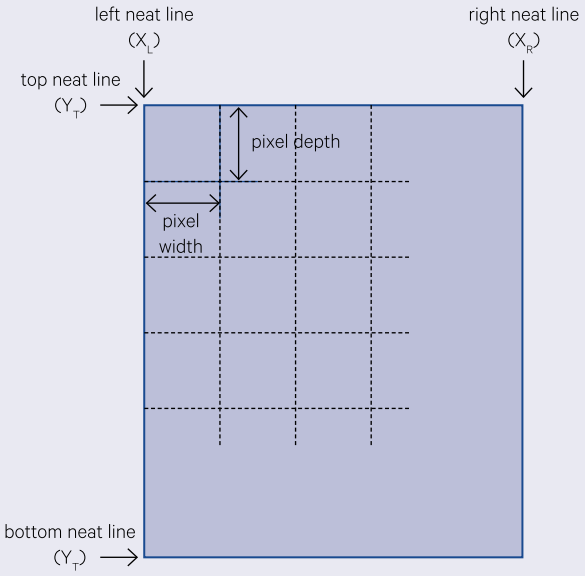


Source: Harrison and Jupp (1992) Figure 43

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.

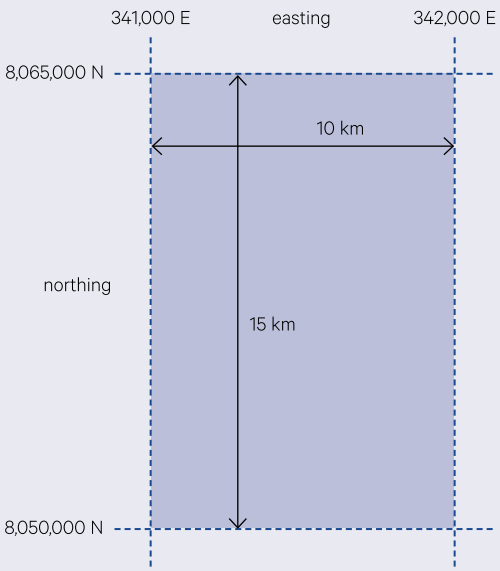


Source: Harrison and Jupp (1992) Figure 47

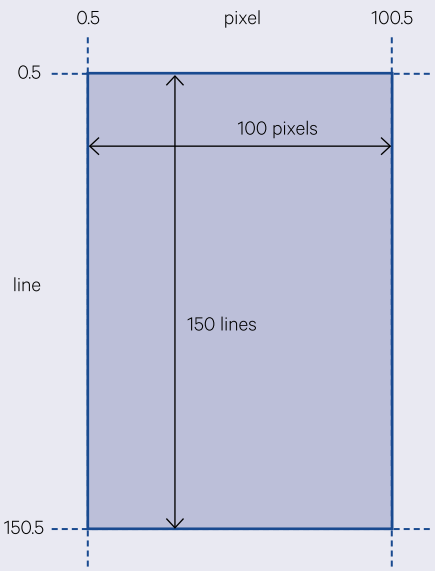
Figure 5.12 Example grid

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:

$$\left\|E\right\|^2=\int\left[I_B\left(x\right)-\sum_{j=1}^{3}a_jI_{A_j}\left(x\right)\right]^2$$

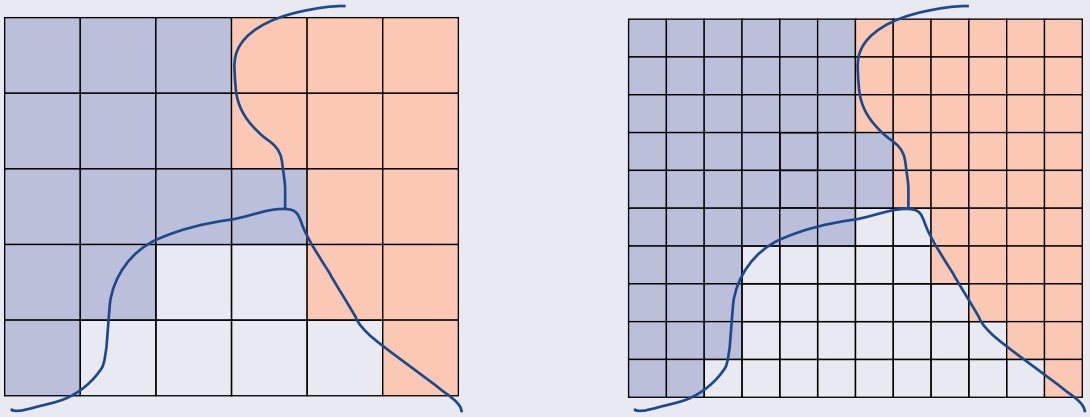
This gives minimum overall error when a_j is simply the area of A_j intersecting B (that is the proportion of B overlapping with A_j).

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:

$$\text{output pixel} = \sum_{i=1}^{4} w_i \times \text{input pixel } i$$

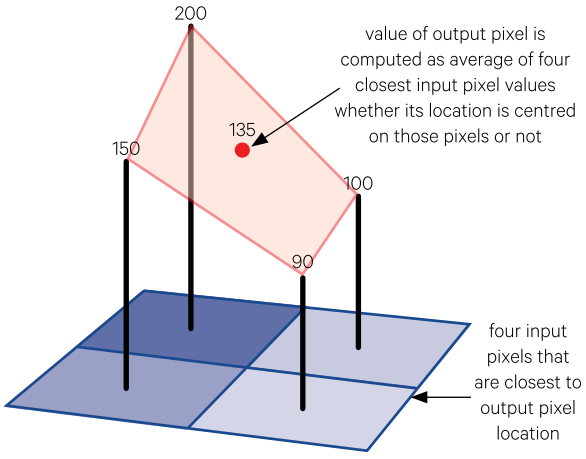
where

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).

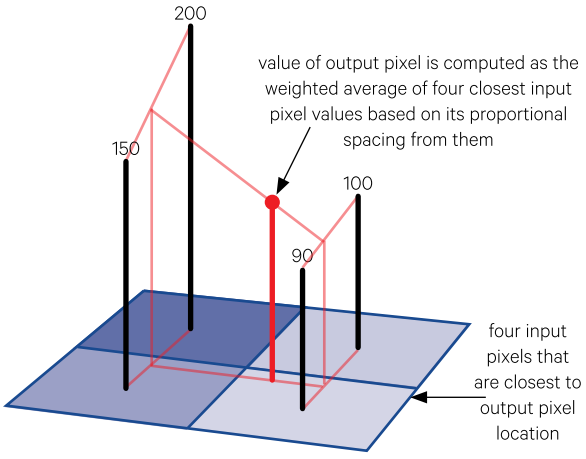
This process tends to create images with lowered contrast, in which edges are somewhat smoothed. It cannot be used with categorical data and is generally not recommended for EO imagery.

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

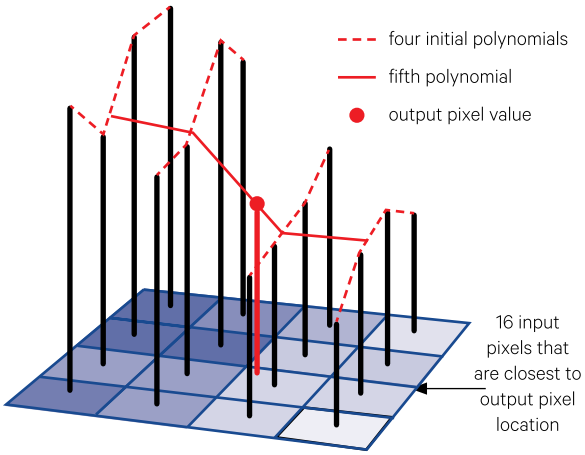
For each output pixel, cubic convolution resampling selects the closest 16 pixels within a 4×4 window in the input image and interpolates an output value by fitting polynomial models to the input values (see Figure 5.4c). For image data this is generally implemented by firstly fitting a cubic polynomial (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.

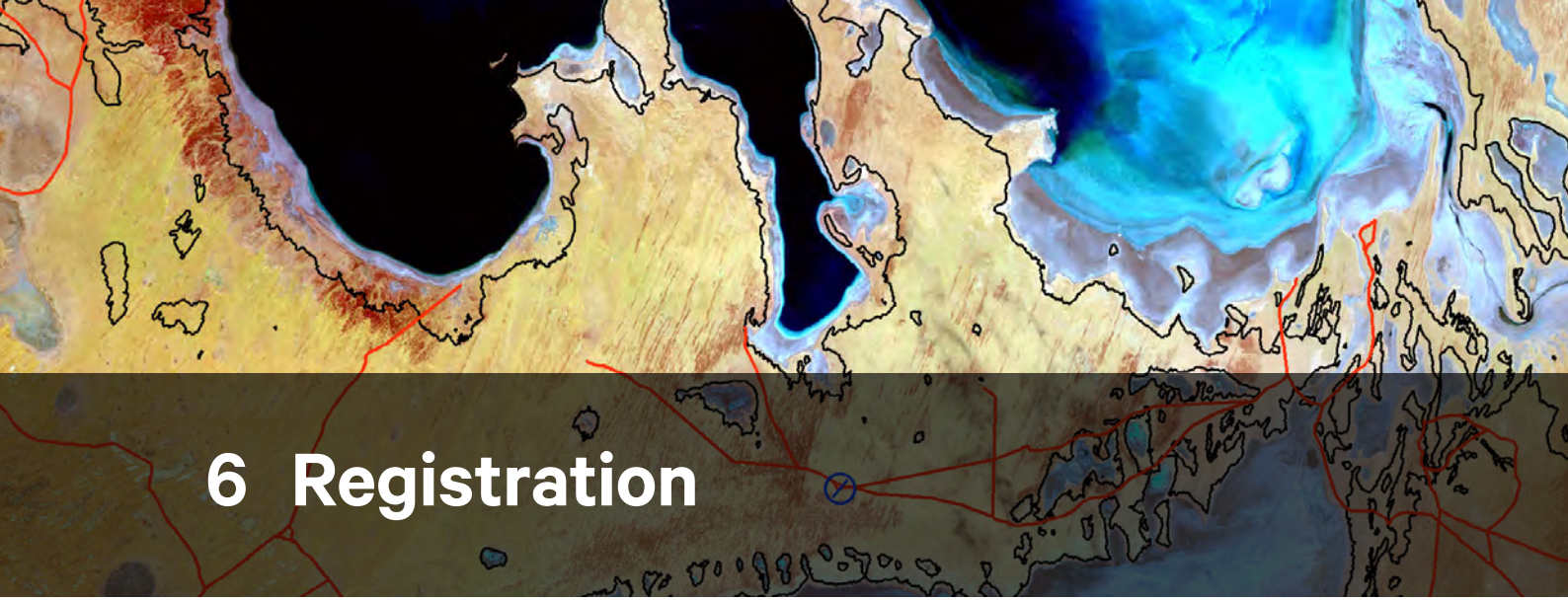
Moik, J.G. (1980). *Digital Processing of Remotely Sensed Images*. NASA SP-431, NASA, Washington DC, USA.

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.

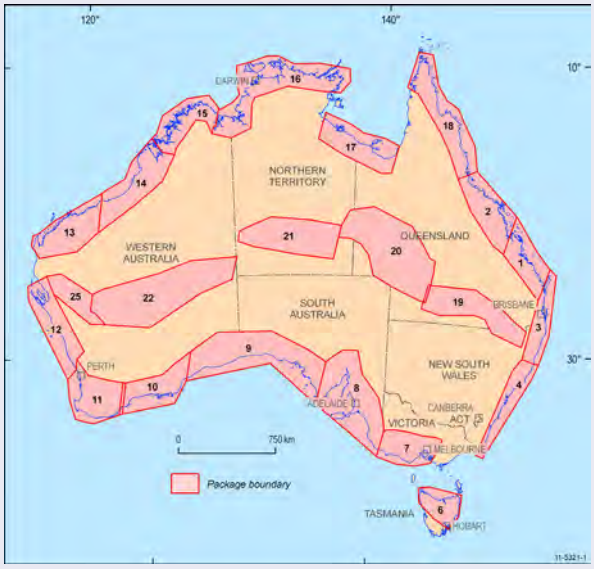
Once project feasiblity was confirmed, the project was undertaken as sequence of tasks:

- 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.

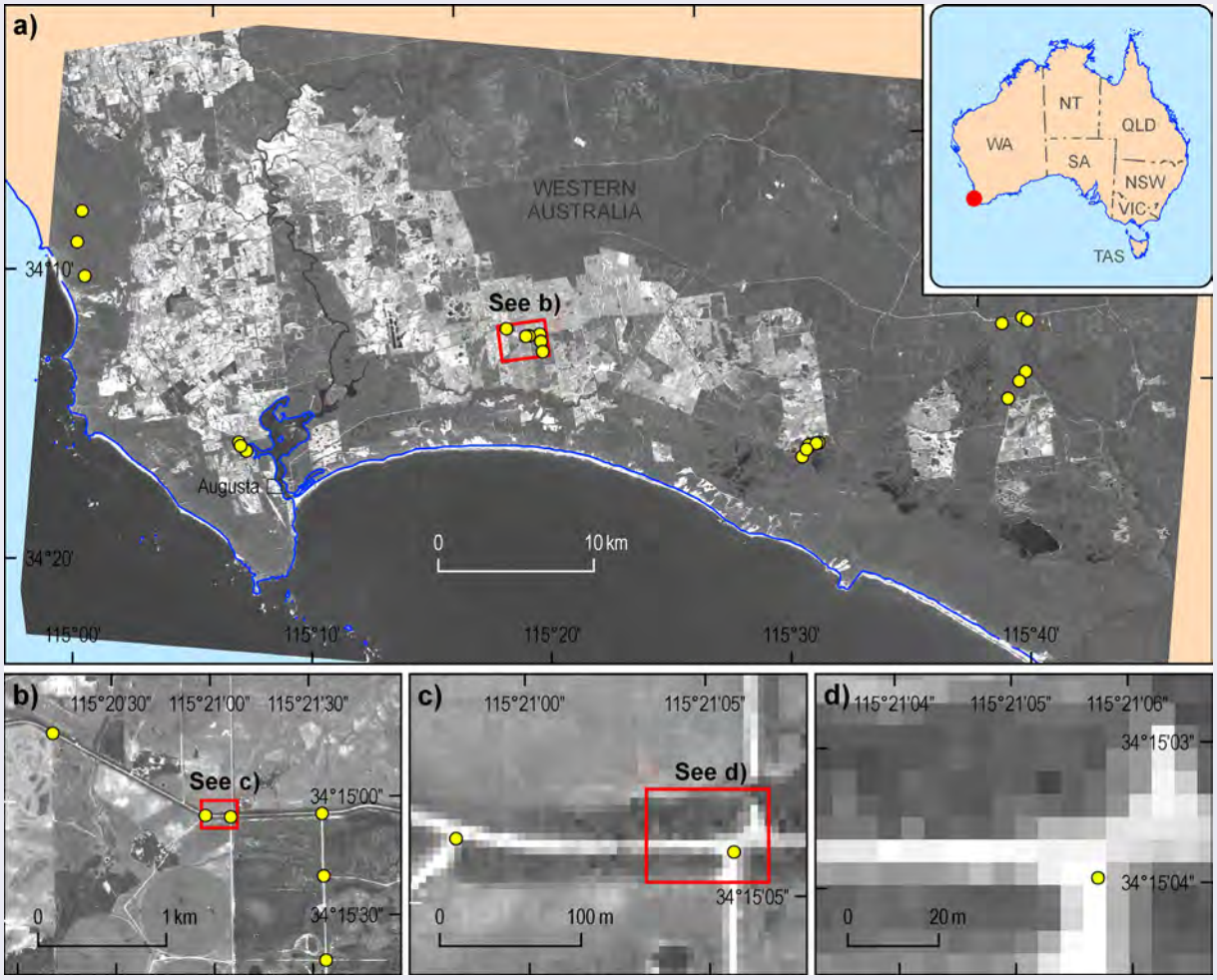
Figure 6.1 GCP work packages



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:

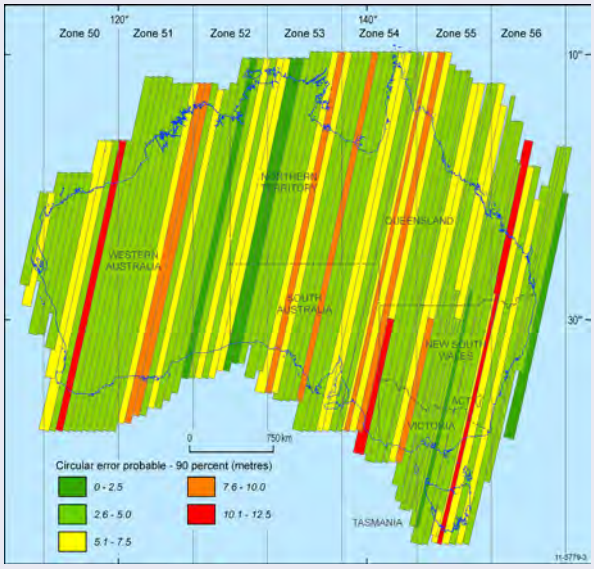
- analysis to filter out control point outliers (indicative of gross errors); then
- processing of the pass using a limited number of control points. The remaining GCPs served as check points to estimate the error in the correction.

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).

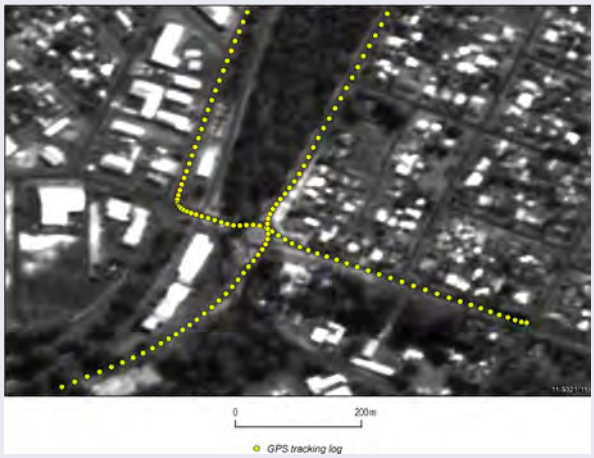
Accuracy was assessed for each orbit segment against available check point residuals (see Figure 6.4a), by visually comparing rectified imagery with dGPS tracking data recorded in a moving vehicle (see Figure 6.4b), and by visual comparison with adjacent image paths (see Figure 6.4c).

Figure 6.4 Accuracy assessment

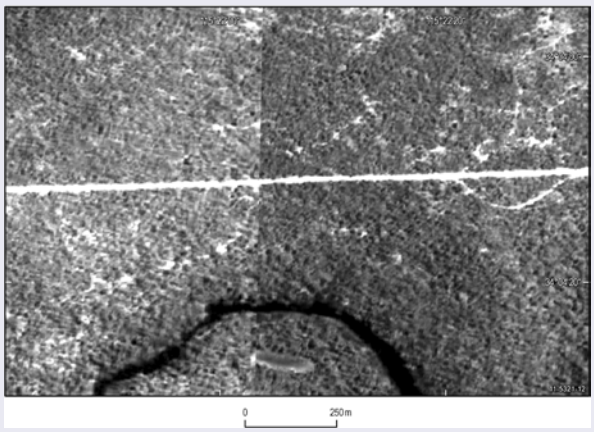
a. Check point residuals



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

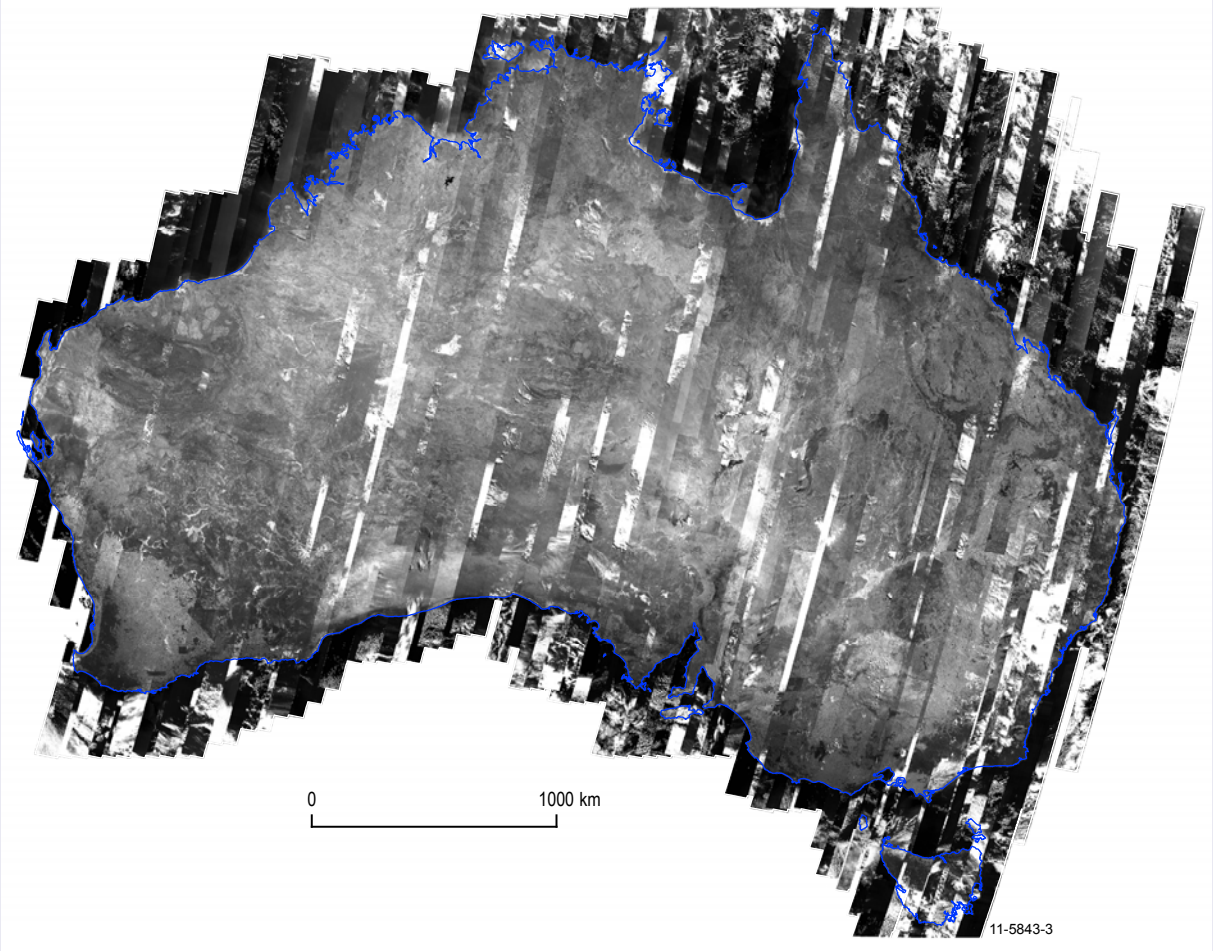
9,560 corrected ALOS/PRISM scenes, totally over 6 tb, were combined into a manageable set of image mosaics:

- 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



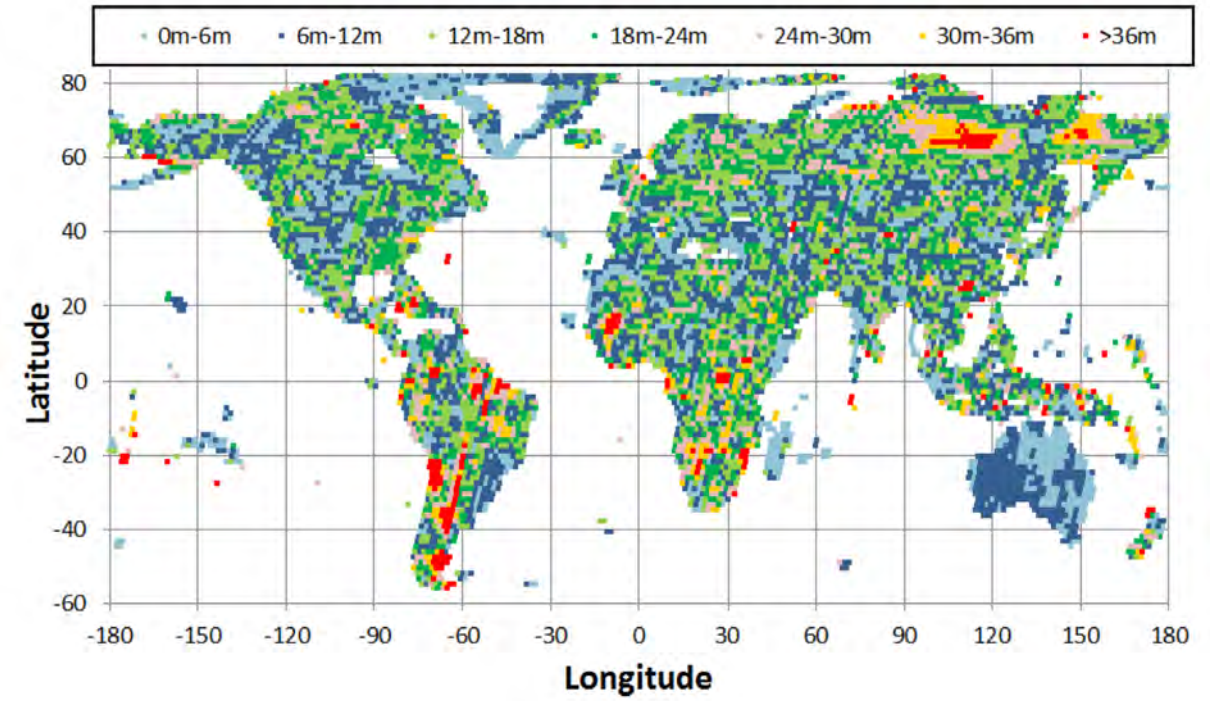
Source: Geoscience Australia

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

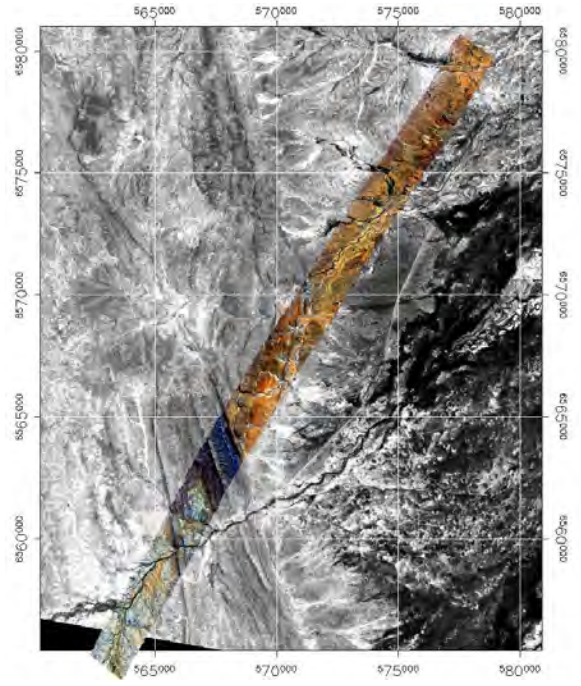
A significant volume of cartographical information is now available in digital format. Most commonly digital map data is stored as part of a Geographic Information System (GIS). EO and GIS datasets are becoming increasingly integrated for the purposes of analysis, interpretation and visual presentation (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:

map scale = pixel size × 2 × 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.

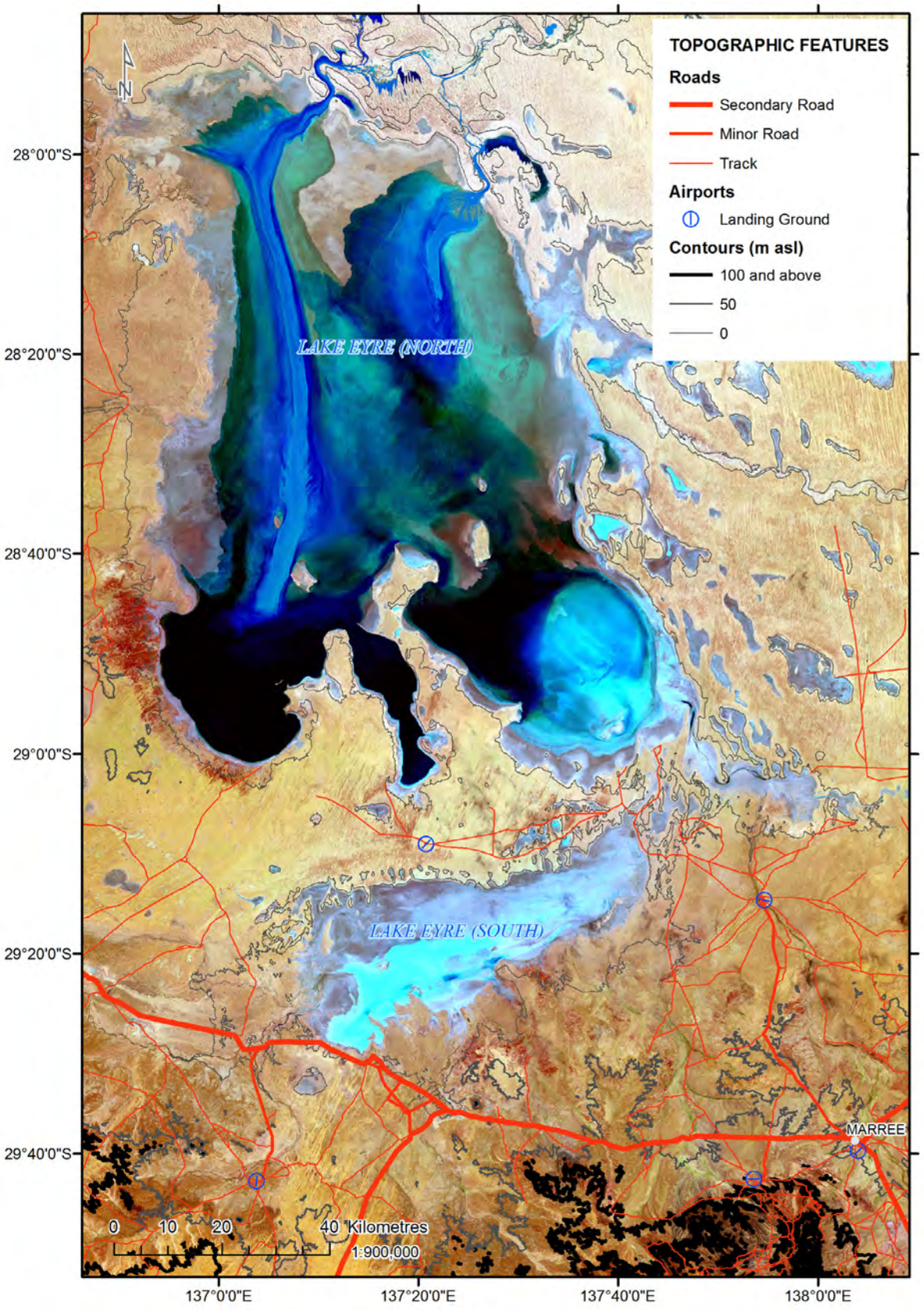
Table 6.2 Map scale and image resolution

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

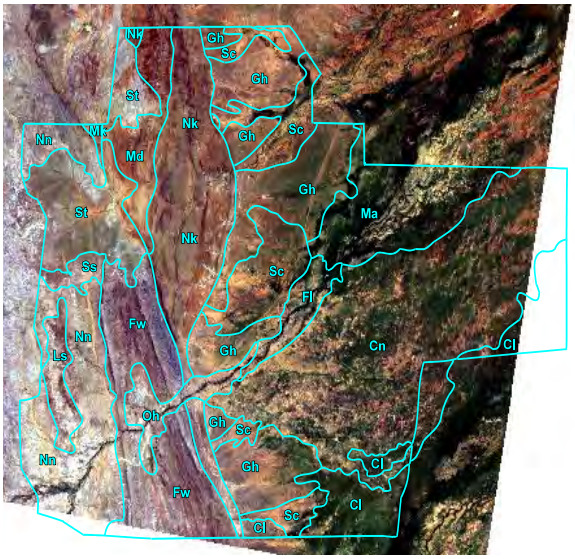
Adapted from: Nagi (2010)

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: 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 projection, please visit: *https://www.planning.act.gov.au/tools_resources/survey-data-maps/surveying-data/surveyors_information/coordinate_system/act_standard_grid_parameters*

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 cadastra) 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.

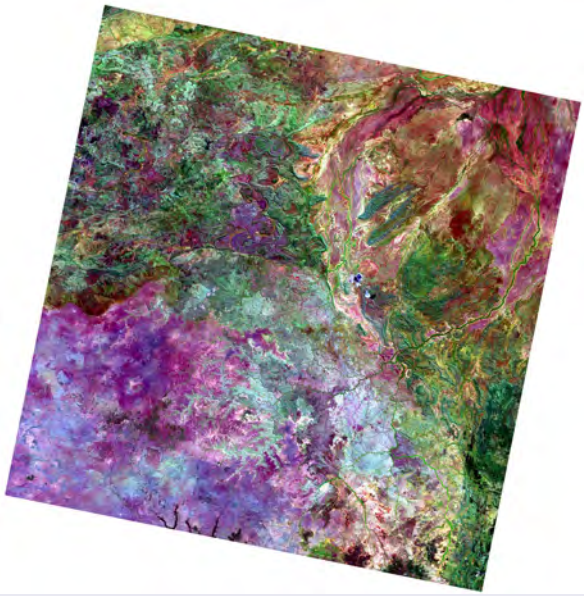
Figure 6.11 Uncorrected image scenes

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

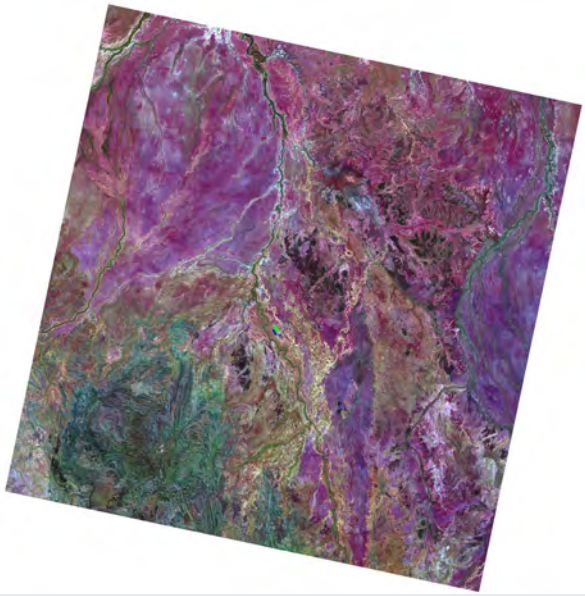
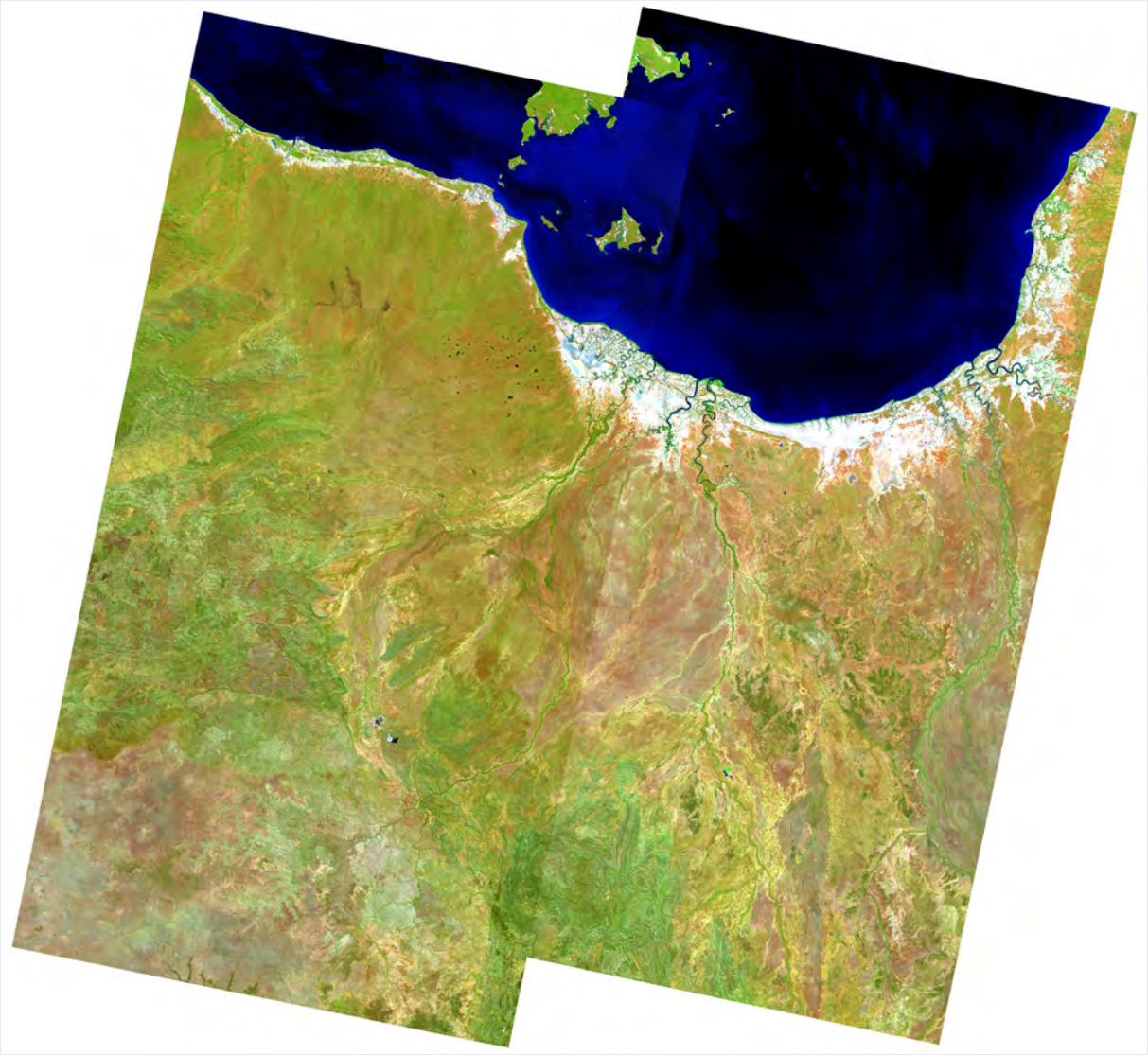


Figure 6.12 Scene mosaic

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.

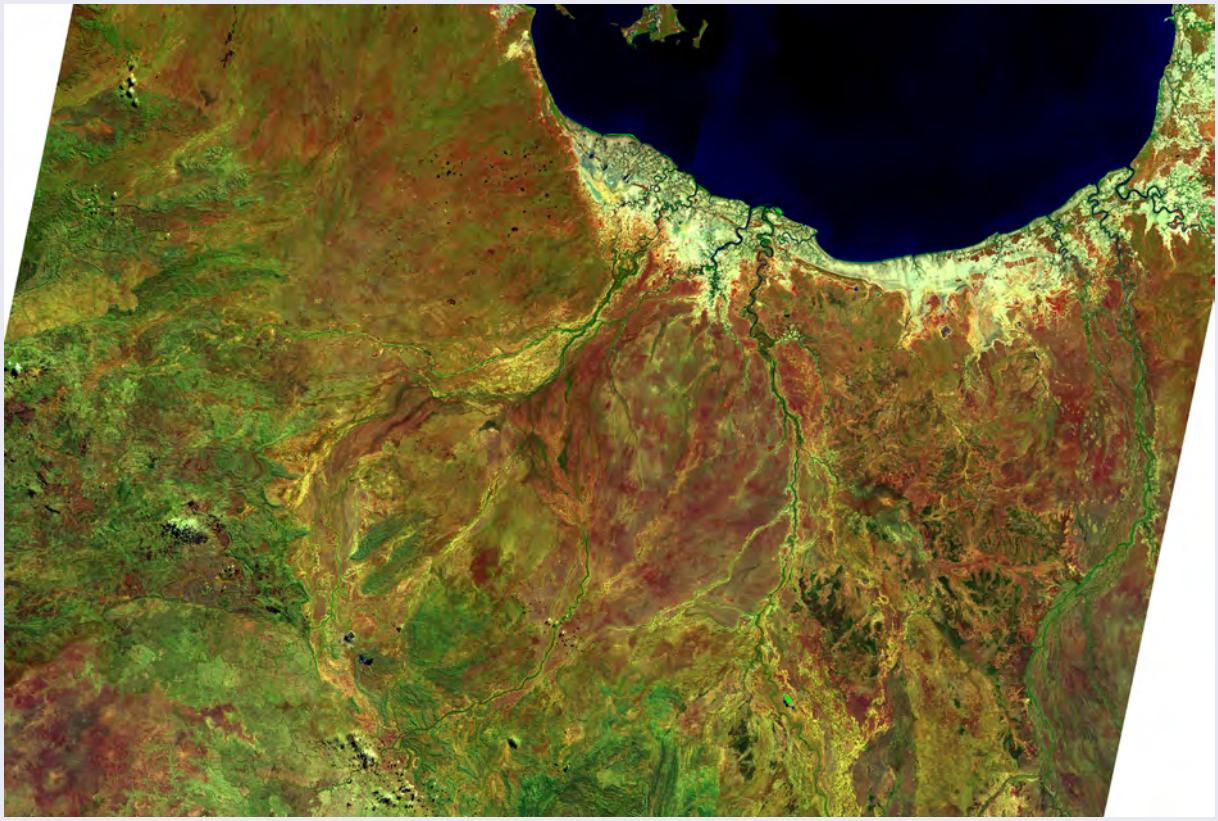
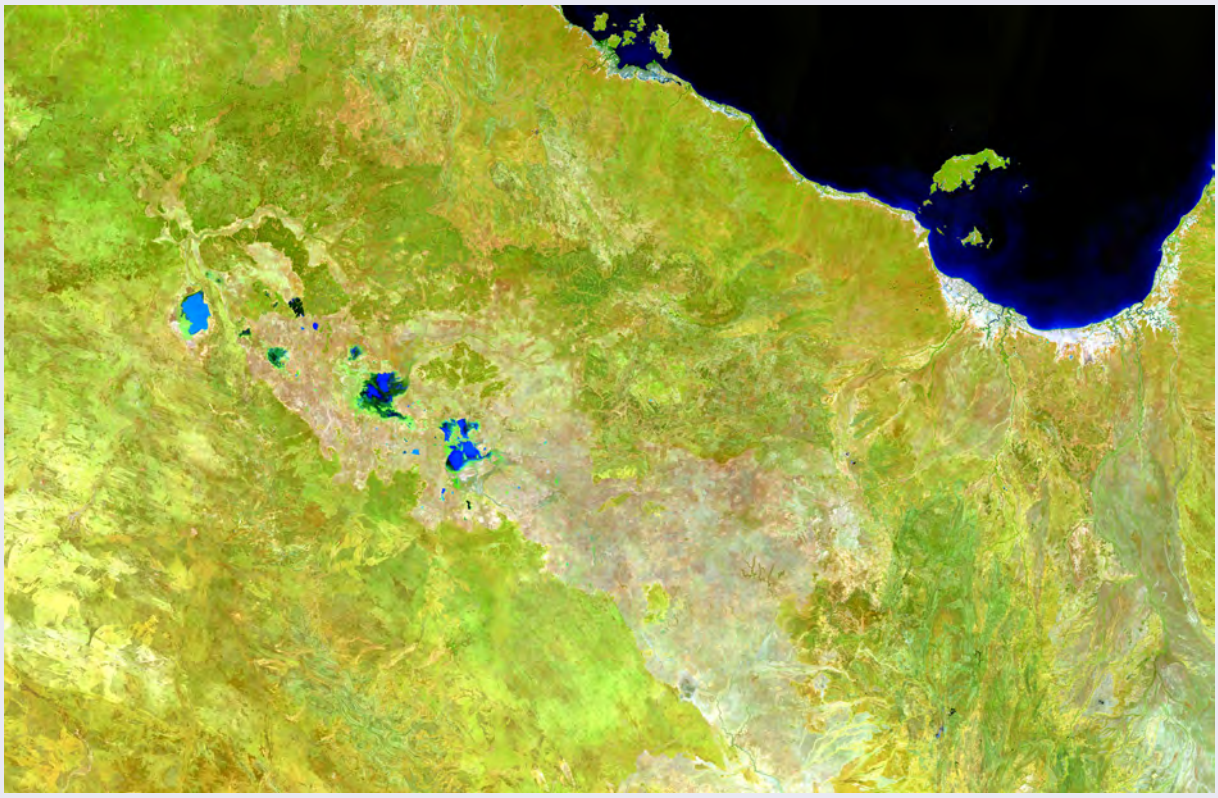


Figure 6.14 Pixel mosaic

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 during 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 (Malthus *et al.*, 2018a) and was framed around a workflow that is summarised in Figure 6.15.

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 6.16). Being flat, homogeneous, and having a relative flat spectral response, it is an ideal spectral validation site.

Figure 6.15 DEA validation workflow

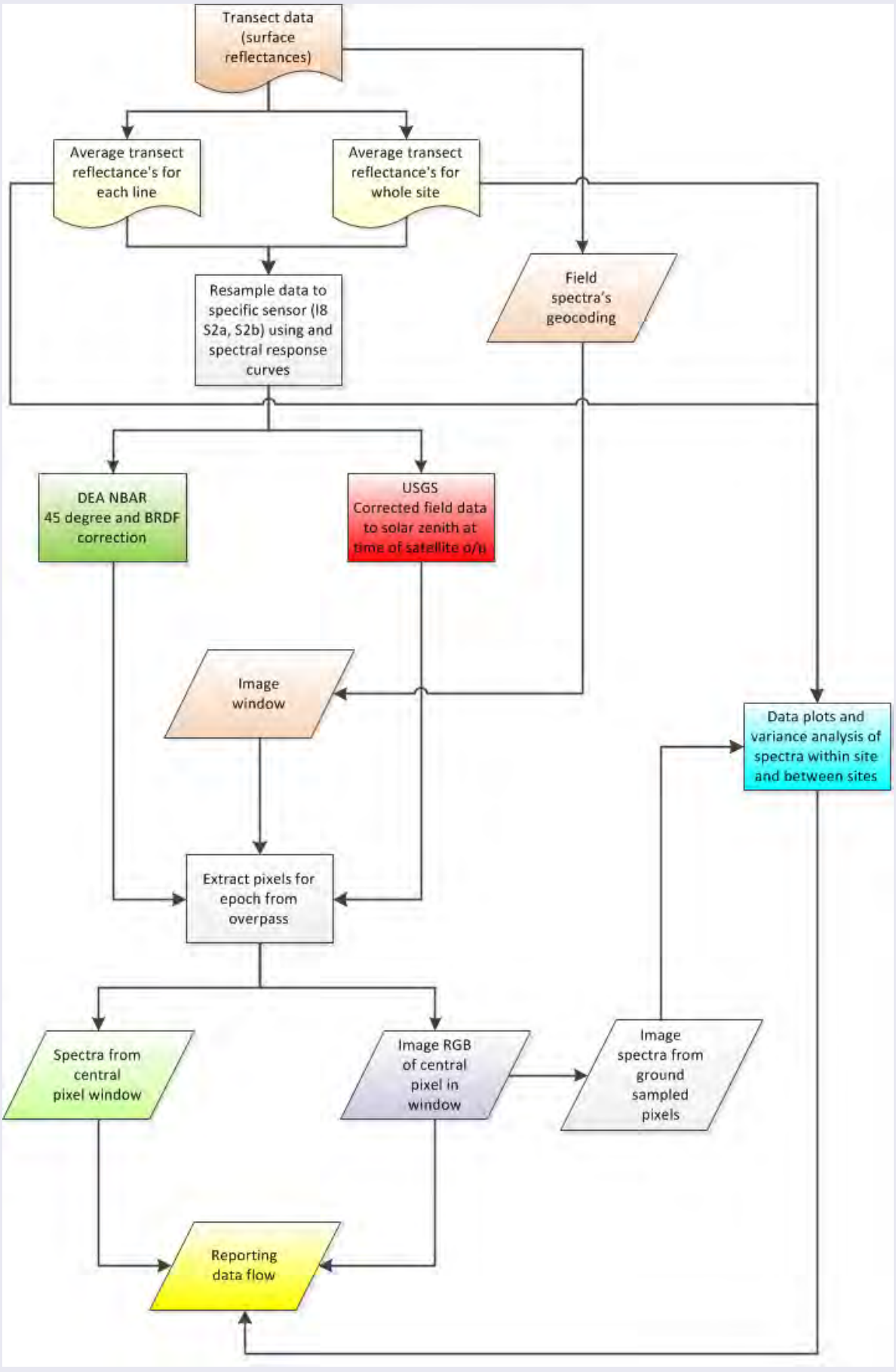
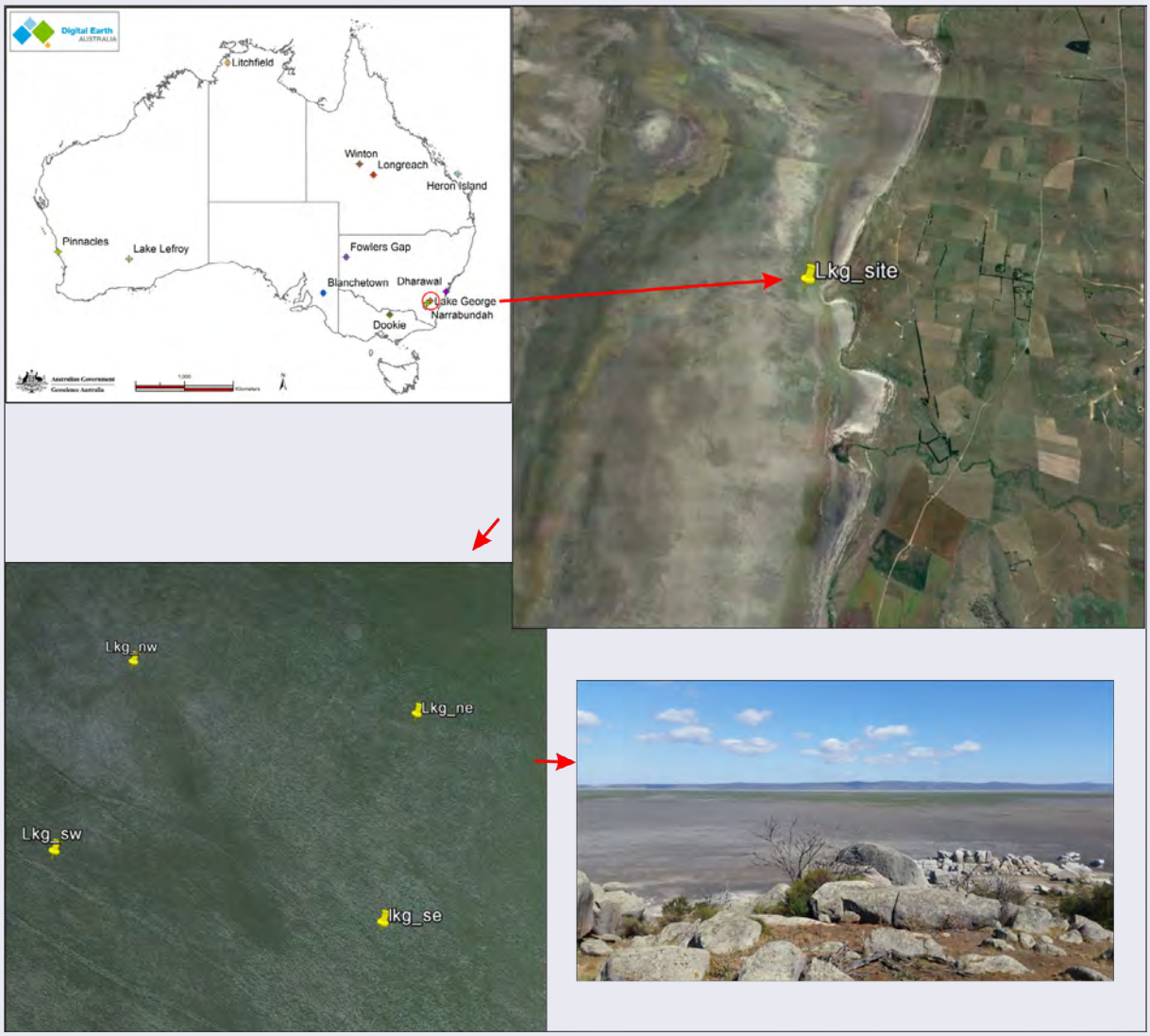


Figure 6.16 Lake George validation site



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 6.17). 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 6.17 Spectral sampling protocol

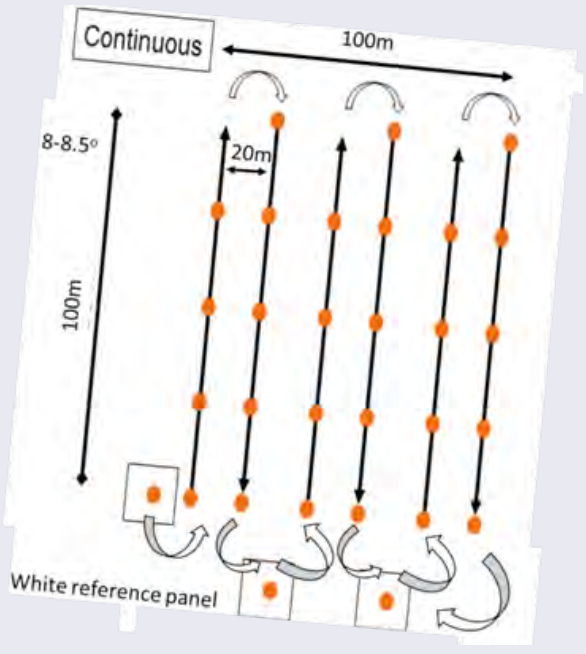


Figure 6.18 Lake George field sampling on 3 May 2018

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 6.18).

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 filed sites located within larger spectrally homogeneous patches tend to mitigate the impact of locational uncertainty (see 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]

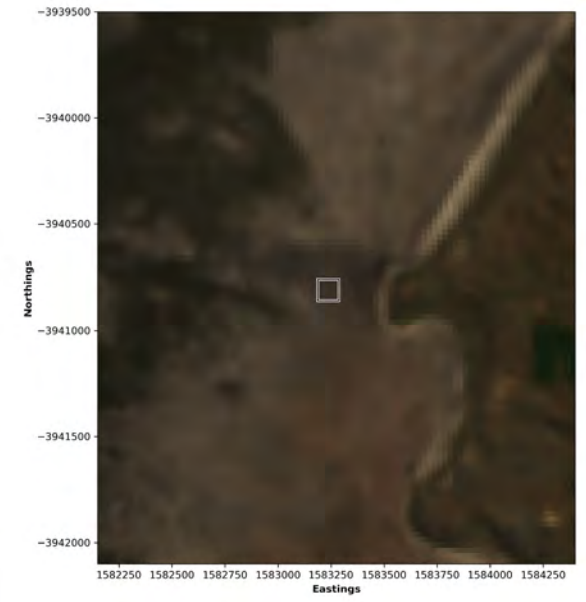
Figure 6.21 compare the average reflectances derived from field and image measurements for each image band, while Figure 6.22 offers a band-by-band comparison of reflectances for satellite and field data. These results will be used to validate the DEA surface reflectance products (Landsat-8 OLI and Sentinel-2 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.

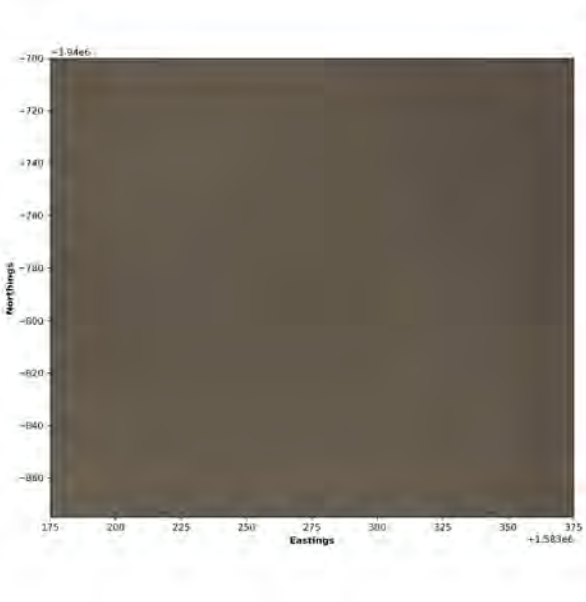
Figure 6.19 Landsat-8 OLI imagery for Lake George site on 3 May 2018

Coordinates for all images are in Australian Albers projection. In images b and c, axes are labelled relative to offsets of 1,583,000m E and -3,940,000 N.

a. Landsat-8 OLI image spanning 2 km by 2 km with white box showing the precise location of this field site



b. Landsat-8 OLI image showing field site pixels



c. Landsat-8 OLI image showing only those field site pixels that were sampled by spectral transects

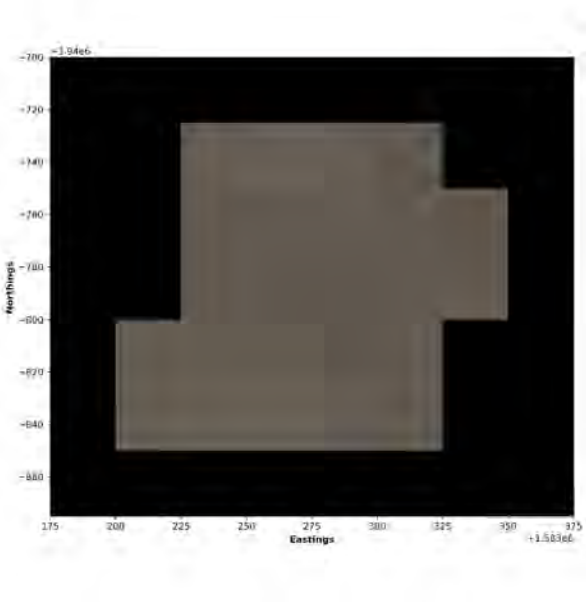


Figure 6.20 Lake George spectra transect lines overlaid on Landsat-8 OLI image grid

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).

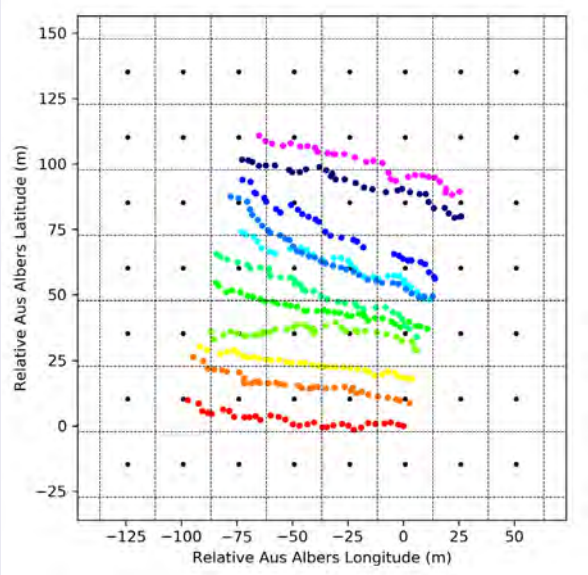


Figure 6.21 Field spectra versus Landat-8 OLI spectra on 3 May 2018

Y axis shows average reflectance measured by field spectra (red) and Landsat-8 OLI image (blue).

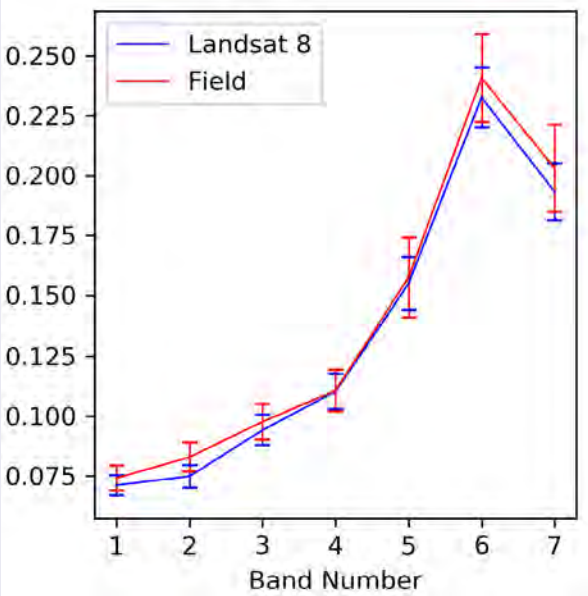
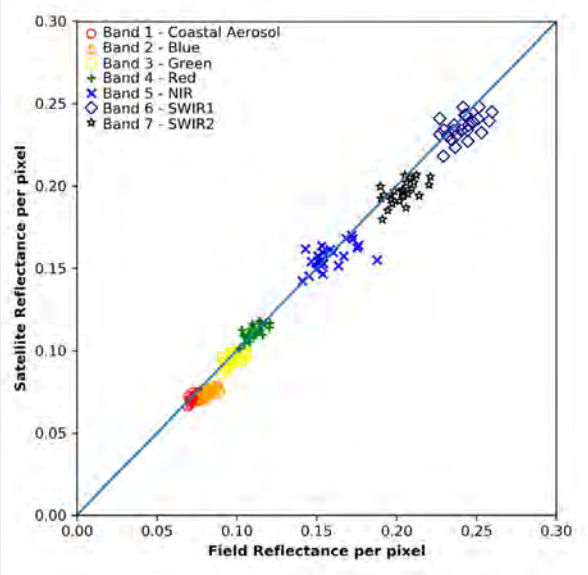
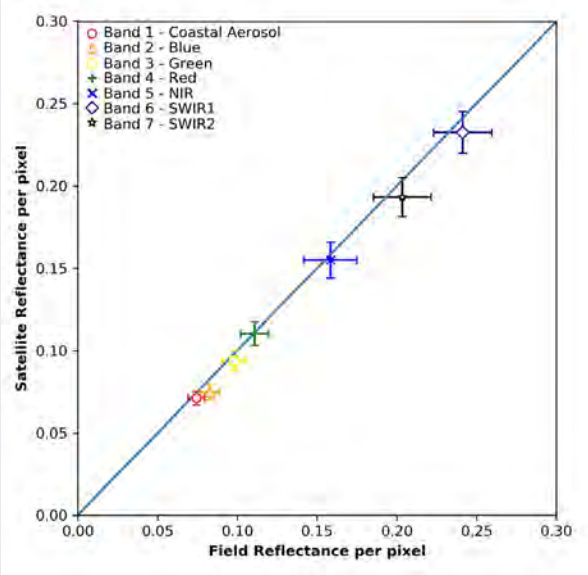


Figure 6.22 Comparison of field and satellite reflectances for each band

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.


crc•si

CSIRO

Australian Government
Geoscience Australia

bushfire&natural
HAZARDSCRC