Field Techniques:
GIS, GPS and Remote Sensing

Royal Geographical Society with IBG

Advancing geography and geographical learning

- **Geography Outdoors:**
  the centre supporting field research, exploration and outdoor learning

RGS-IBG Geography Outdoors: the centre supporting field research, exploration and outdoor learning

The Royal Geographical Society (with The Institute of British Geographers) is the UK's main organisation for screening and funding small independent research expeditions. These assist in furthering geographical knowledge and the encouragement of life-long learning, leadership and team skills.

Geography Outdoors: the centre supporting field research, exploration and outdoor learning The provides information, training and advice to anyone planning an expedition overseas through a range of training seminars and workshops, publications and information resources. For more details including information on projects and fieldwork supported in the past please visit the Geography Outdoors website: **www.rgs.org/go**

Geography Outdoors receives core support from Shell International Limited. Shell has provided sponsorship for over a decade so that the Centre can maintain and improve its services to schools, universities and the scientific and academic communities in general, and so promote interest in, and research into, geographical and environmental concerns worldwide. Further information on the Royal Dutch/Shell Group of Companies and their oil, natural gas, chemical and renewable energy businesses, can be found on the Shell website: www.shell.com

Field Techniques:
GIS, GPS and Remote Sensing

Authors:
Richard Teeuw, Martin Whiteside, Nicholas McWilliam
and Paul Zukowskyj

With contributions from:
Daniel Hourigan, Nick Mount and Jonathan Raper

Published by Geography Outdoors
Royal Geographical Society (with IBG)
1 Kensington Gore
London SW7 2AR

Tel +44 (0)20 7591 3030
Fax +44 (0)20 7591 3031
Email go@rgs.org
Website www.rgs.org/go

November 2005

ISBN **978-0-907649-88-5**

Cover illustration: Where the barren sands of the Western Sahara meet the fertile floodplain of the Senegal River: the Mauritania-Senegal border zone, West Africa. This image was produced from Landsat Thematic Mapper imagery, using ER Mapper image processing software and a laptop computer. The image was produced in Mauritania, as part of a programme of natural resource mapping, technology transfer and training, funded by the Japanese International Cooperation Agency.

About the Authors

Richard Teeuw *richard.teeuw@port.ac.uk*
Senior lecturer in applied geomorphology and remote sensing at the School of Earth and Environmental Sciences, Portsmouth University. His first expeditionary fieldwork was with Brathay in Iceland and Kenya. Richard's PhD focused on diamond prospecting in Sierra Leone. Since then he has used remote sensing for many projects, mostly in the tropics: water exploration in Ghana; biodiversity and deforestation surveys in Guyana; soil erosion assessment in Zambia; mineral exploration in Borneo, Burkina Faso, Ivory Coast, Guinea and Mauritania. His enthusiasm for technology transfer and low-cost remote sensing applications has led to consultancy work for the overseas aid agencies of Canada, Germany and Japan. Richard is a founder member of the EAC Mapping Unit at the RGS-IBG - volunteers who offer advice on how to best use GIS, GPS and remote sensing for exploration projects.

Nick McWilliam *n.mcwilliam@apu.ac.uk*
GIS coordinator for MapAction, a charity that provides emergency mapping for aid agencies following disasters, and a geography lecturer at Anglia Polytechnic University in Cambridge. Previous GIS work and expeditions include mapping water provision in Delhi slums in 2004, food insecurity in Lesotho 2003, British Antarctic Survey Mapping Centre 1999-2002, Rungwa Game Reserve, Tanzania 2000, RGS Mkomazi Ecological Research Programme, Tanzania 1995-99, RGS Expedition Advisory Centre 1989-92, field research in Sumatra, Siberia and the Alps, and trips to the mountains of Ladakh, Tien Shan, Norway and various parts of East Africa.

Martin Whiteside *whiteside_martin@hotmail.com*
Having studied combined geology and astronomy at university, Martin turned these skills to remote sensing. Using satellite data he was involved in diamond / jade mineral exploration along the front of the Kunlun Shan and co-organised an expedition to Bogda Shan, Xinjiang, China. After returning from China he completed a DIC at Imperial College London. Subsequently he has led mapping projects in the Sierra Nevada and the Sorbas Basin testing the benefits of declassified US military data for geomorphological analysis. His current research interests include GPS techniques, image interpretation and photogrammetry. Martin has used GIS for many years and is a contributor to various EAC mapping unit activities and became a fellow of the RGS-IBG in 2003. Martin currently works for a major mining company and is studying a PhD at University College London.

Paul Zukowskyj *p.m.zukowskyj@herts.ac.uk*
A lecturer in GIS and Remote Sensing at the University of Hertfordshire. Paul is interested in applications of remote sensing (RS) and GIS and in algorithm development for RS and GIS. Paul’s particular interest is the integration of digital elevation models with RS. Research projects include vegetation in southern Spain, glaciological and sedimentary studies in Iceland and fluvial systems in the UK. A recent interest of his is spatial patterns in health, particularly infectious disease epidemiology. Paul began as a lecturer at the University of Greenwich in 1996. In 1998 he joined Huntings Technical services as a photogrammetry consultant in Poland before joining the University of Hertfordshire as senior technical officer for GIS in 1999. In September 2003 he began his current post as lecturer in remote sensing and GIS. Currently working on a Tidal Thames landcover change project for the Thames Region Environment Agency.

Dan Hourigan *danielhourigan@hotmail.com*
Currently an IT Engineer in South London, Dan trained as a geologist at Imperial College in the early 1990’s and followed this up with a Remote Sensing MSc from University College London in 1996. He then worked in exploration geology in the deserts of South Australia for several months as well as building numerous GIS projects for a mining company in Perth, Western Australia. After travelling New Zealand and Australia Dan rejoined the remote sensing department of Imperial College and became involved in the Bogda Shan expedition to Xinjiang, China with Martin Whiteside.

Preface

For many years now Geography Outdoors have run an autumn training weekend for people planning fieldwork and exploration projects. The authors of this handbook first met at these training weekends in the early 1990s, where they were making presentations on the uses of Geographical Information Systems (GIS), the Global Positioning System (GPS) and remote sensing – collectively known as Geographical Information Sciences (GISci). By the mid-1990s we realised that our collective experience could be developed into a manual for GISci usage in fieldwork. It has taken a decade to go from ideas developed in the aftermath of a training weekend (over a few beers in the Student Union bar of Imperial College), to the first edition of this handbook. That delay has nothing to do with too much beer, and everything to do with this project being carried out on a voluntary basis, amongst the many demands of full-time jobs.

Every year we hold a question-and-answers session at the RGS-IBG *Explore* training weekend. A brief review of the types of question that dominated over the past decade shows how rapidly the capabilities of geospatial technologies have changed. In the early 1990s, most of the questions were on GIS applications: GPS and remote sensing were of minor interest, due to their high costs and limited detail. By 2000 all that had changed: GPS accuracies had increased from +/- 100 m to +/-10 m and most of the interest was in GPS applications for navigation and mapping. Five years on, with low-cost or free satellite images and 3-D elevation models of the earth readily available over the Internet - plus computers getting more powerful, yet cheaper – the main interest is in how remote sensing can help exploration projects. Ironically, today, the main limitation on the more widespread use of geospatial technologies in exploration fieldwork is the relatively high cost of GIS software.

This handbook is designed to help explorers with their exploring; particularly the bits where you have to (delete as appropriate): decide on what you are going to survey, work out where you are, figure out where the feature to be surveyed is (and how to get there); sample and produce maps of said feature.

We welcome your comments and suggestions on how to improve the manual. As scientists with geological, geomorphological and ecological backgrounds, we are well aware of the lack of socio-economic and anthropological coverage in the handbook: we are keen to see more examples and case studies from those sectors of exploration – indeed, if there is anyone out there who would like to contribute a chapter on such applications for the next edition, please contact us.

Should you have any questions about GISci applications in exploration fieldwork that are not answered by this manual, please contact the RGS-IBG Geography Outdoors (*www.rgs.org/go*) – even if they cannot provide an answer to your question, they will probably be able to put you in touch with someone who can.

Finally, a big THANK YOU to all the people who have offered advice and examples over the years. Without the constant support and encouragement of Mrs Shane Winser, the Expeditions Advisory Centre manager, this handbook would have remained as no more than a set of chapter headings on beer-stained notepaper. Many assistants in Geography

Outdoors have helped with earlier drafts, Christine Eriksen doing an excellent job formatting this edition. Professor Jonathan Raper (Information Science Department, City University, London) was very involved with the early stages of this manual and provided useful material for the GIS and surveying chapters. The later stages of the manual benefited greatly from the inputs of Dan Hourigan, who provided useful inputs on data sources, ancillary equipment and software, as well as Dr Nick Mount (Geography Department, Birkbeck College, London), who produced the 'virtual reality' conquest of Everest case study. Thanks also to Dr Steve Drury (Earth Sciences, the Open University), Dr Hazel Faulkner (Flood Hazard Research Centre, Middlesex University) and Dr Tim Stott (Geography Department, Liverpool John Moores University), for their helpful comments on an earlier draft of this handbook. Last, but not least, thanks to ESRI UK for generously providing sponsorship to help us include the colour illustrations at the centre of this handbook.

All feedback on this manual is most welcome: please email the RGS-IBG Mapping Unit with your comments (*go@rgs.org*). Particularly welcome are further references for the Bibliography, on GIS, remote sensing or GPS applications in fieldwork-based projects; two to three page illustrated case studies, for possible inclusion in future editions; and weblinks to useful internet sites.

Good luck with your endeavours!

Richard Teeuw,
School of Earth & Environmental Sciences, University of Portsmouth.

Nick McWilliam,
Geography Department, Anglia Polytechnic University & Oryx Mapping, Cambridge.

Martin Whiteside,
Inco Europe Ltd, Acton London.

Paul Zukowskyj,
Geography Division, University of Hertfordshire, Hatfield.

Contents

- **SECTION A: INTRODUCTION 1**

1 GIS, GPS, REMOTE SENSING AND FIELDWORK 3
  1.1 Organisation of this manual 4
2 THE GEOGRAPHICAL FRAMEWORK 11
  2.1 Geographical reference systems 12
  2.2 The Universal Transverse Mercator (UTM) system 22
  2.3 Choosing a co-ordinate system for your GIS 26
  2.4 Conclusion 27
3 GEOGRAPHICAL INFORMATION SYSTEMS (GIS) 31
  3.1 Introduction 31
  3.2 Data input and organisation 31
  3.3 How can GIS help? 34
  3.4 GIS and fieldwork 35
  3.5 Limitations of GIS 37
  3.6 GIS data types and sources 39

- **SECTION B: DATA 43**

4 GIS DATABASE MECHANICS 45
  4.1 The underlying information in the GIS 45
  4.2 Selecting the project database 46
  4.3 Database configuration 48
  4.4 Example of database structure designs 49
  4.5 Using the database in the field 52
  4.6 Integrating existing data 54
  4.7 Data import/export and cleaning 55
  4.8 Adding raw expedition data 56
  4.9 Accessing project data 58
5 REMOTE SENSING 67
  5.1 Introduction 67
  5.2 Satellite imagery 68
  5.3 Resolution 69
  5.4 Aerial photography 72
  5.5 Radar Imagery 74
  5.6 Recent developments 74
  5.7 Sources of remotely sensed data 75
6 THE GLOBAL POSITIONING SYSTEM (GPS): PRINCIPLES & CONCEPTS 79
  6.1 GPS and field navigation 79
  6.2 Introduction to GPS functions/features 80
  6.3 GNS history 84
  6.4 How GPS works 85
  6.5 GPS accuracy 88
  6.6 Correct GPS handling 92
  6.7 Assessing data quality 95
  6.8 How GPS calculates and stores positional data 98
  6.9 NMEA sentences and stored information 99
  6.10 Understanding precision and improving accuracy 102
  6.11 Future developments of the Global Positioning System 108

- **SECTION C: TECHNIQUES 111**

7 GISci Analysis 113
  - *7.1 Introduction 113*
  - *7.2 Mapping and symbolising 113*
  - *7.3 Spatial selections and queries 114*
  - *7.4 Creating 'buffers' 115*
  - *7.5 Testing relationships between data layers 116*
  - *7.6 Spatial overlay operations 116*
  - *7.7 Spatial analysis 118*
  - *7.8 Digital elevation data 126*

8 Image interpretation and processing 135
  - *8.1 Image interpretation 135*
  - *8.2 Geomorphological mapping 137*
  - *8.3 Mapping geo-ecological features 139*
  - *8.4 Human population estimates 140*
  - *8.5 Digital image processing 141*

9 Geocorrection and photogrammetry 153
  - *9.1 Data from space-borne sensors 153*
  - *9.2 Photogrammetry 154*
  - *9.3 Geocorrection 161*
  - *9.4 Digital photogrammetry 169*
  - *9.5 Practicalities of accurate mapping using airphotos 174*
  - *9.6 Summary 175*

10 Traditional Surveying 179
  - *10.1 Introduction 179*
  - *10.2 Surveying context 179*
  - *10.3 Surveying instruments/devices 180*
  - *10.4 Survey techniques 184*
  - *10.5 Checking, logging and downloading survey data 190*

11 Using GPS for Fieldwork 193
  - *11.1 GPS applications 193*
  - *11.2 GPS care & power requirements 194*
  - *11.3 Using GPS in non-ideal conditions 195*
  - *11.4 Integrating GPS readings and historical maps 196*
  - *11.5 Integrating GPS and other GISci data 197*
  - *11.6 GPS receivers for creating expedition maps 201*
  - *11.7 GPS receiver models 204*

- **SECTION D: PLANNING & PRACTICALITIES 219**

12 Project Planning and Management 221
  - *12.1 Project design 221*
  - *12.2 Sampling strategies 226*
  - *12.3 Data 231*
  - *12.4 Equipment, software and fieldwork practicalities 233*

13 Field Equipment 239
  - *13.1 Field-based applications 239*
  - *13.2 Field computing overview 239*
  - *13.3 Data storage 240*
  - *13.4 Understanding device protection 245*
  - *13.5 Powering field equipment. 248*

  13.6 *Caring for GPS receivers* 259
  13.7 *Field communications & remote data access* 260
  13.8 *Field concerns with cellular phones* 261
  13.9 *Photographic equipment and geo-tagging* 262
  13.10 *Selecting the expedition hardware* 265
  13.11 *Purchasing equipment* 267
  13.12 *Disaster recovery planning* 269
14 GISci Software 277
  14.1 *General principles in software selection* 277
  14.2 *Software for GPS data retrieval* 281
  14.3 *Software for terrain visualisation* 282
  14.4 *Remote sensing / image processing software* 282
  14.5 *GIS software* 284
  14.6 *Very low cost software* 289
  14.7 *Summary* 292
15 Completing the Project 297
  15.1 *GIS analysis* 297
  15.2 *Useful GIS functions* 297
  15.3 *Results and 'deliverables'* 298
  15.4 *Continuation and project sustainability* 298
16 GISci Applications 303
  16.1 *Archaeology* 303
  16.2 *Ecology* 303
  16.3 *Geology* 304
  16.4 *Geomorphology* 305
  16.5 *Natural resource management* 306
  16.6 *Wildlife* 307
  16.7 *Socio-economic applications* 307
17 Case Studies 313
  17.1 *Bogda Shan Expedition 2000* 313
  17.2 *GISci helps in conquest of Everest* 315
  17.3 *Jarlhettur Ridge, Iceland* 318
  17.4 *A semi-desert soil erosion and landslide site* 319

- **SECTION E: APPENDIX 329**

Appendix 1: Sample field data collection form 331
Appendix 2: Worked example using GPS. 332
  Using GPS to Rectify Images 332
  Using HyperTerminal to Check a GPS signal 334
  Using GPS software to Download Data 335
  Editing GPS data in text format 338
  Rectifying Data using ER Mapper 341
Appendix 3: World Register of Field Centres 344
Appendix 4: Buyer's guide to GPS units 346
References & bibliography 355
Web addresses for remote sensing 363
GISci Glossary 366

Figures and Plates

The nature of Geographical Information Sciences requires the use of colour images. To help the reader visualise many of the concepts discussed, within, important colour images from the chapters have been moved into two sections of colour plates.

- The first section of colour plates (Plate 1 to Plate 20) can be found between pages 207 and 218.
- The second section of colour plates (Plate 21 to Plate 32) can be found between pages 322 and 328.

About GISci Acronyms & Terminology

A common problem for novice GISci expeditioners consulting a textbook such as this one, is the ever present and growing use of acronyms. Hopefully each section of this manual should ease the user into the subject matter with clearly defined terms, however, if the reader encounters a term that has not been defined adequately we have included a brief acronym list / glossary at the rear of the manual. Obviously if after consulting this list the reader is still unable to find the required definition then they are welcome to refer to the RGS-IBG Expedition Advisory Centre. The EAC will then be able to direct the reader to an appropriate specialist who can offer further help. Also if a reader believes there is a problem with the manual or if terms are not well enough explained then they are welcome to contact the EAC and hopefully we can address the omission in a later edition of this manual.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section A: Introduction

Chapter 1: GIS, GPS, Remote Sensing and Fieldwork

1 GIS, GPS, Remote Sensing and Fieldwork

The widespread use of computers has led to the development of new technologies, collectively known as geographical information sciences (GISci), for mapping and monitoring features on the surface of the Earth. Foremost for exploration and fieldwork among these technologies are: *geographical information systems* (GIS), which can take digital datasets and produce maps showing features of interest in matter of seconds; the *global positioning system* (GPS), which allows positions to be determined to ±10 m anywhere on the Earth's surface; and methods of observing features from a distance, such as photography or infra-red scanning, known as *remote sensing*. These GISci techniques complement the surveys and sampling that are at the heart of scientific exploration (Figure 1-1): they greatly enhance the types of fieldwork that can be carried out, reduce the amount of time needed for many tasks and improve the quality of results.

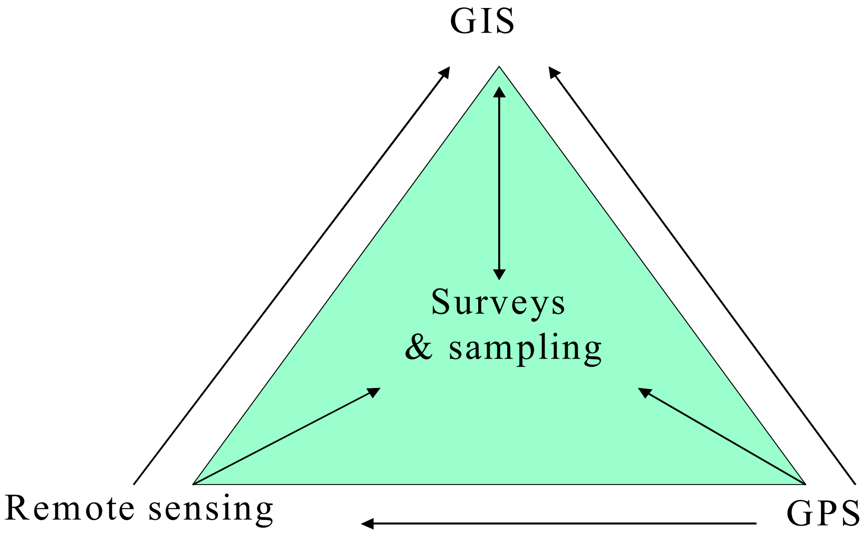


Figure 1-1 Geographical information sciences and expedition fieldwork.

A fundamental objective of most exploration is to observe and record information about the part of the world being studied, for instance by field surveys, photography, or questionnaires. The development of ever-cheaper and more powerful computers, GIS software and GPS kit, along with low-cost satellite pictures of the Earth, has greatly improved the potential of expeditionary fieldwork to record, analyse and present data that may help us to improve conditions on this beleaguered planet.

Remote sensing provides us with a means of recording the distribution of features on the surface of the Earth and changes in those features over time: it is often the only source of new data about a region that will be available to you, prior to you going there to collect field data. Your GPS will tell you where you are in your study region and allows you to input your sample locations into a GIS. A GIS is a means of combining existing data and new data from fieldwork or the interpretation of remotely sensed images. GIS-generated maps greatly reduce the original amounts of data and can be designed to focus on specific themes of interest to your research.

This manual aims to provide expeditions with details of fieldwork techniques, from ‘traditional’ compass-based surveying, through to the use of GIS to show GPS-located sites on satellite images displayed on a laptop screen in, say, Amazonia or the Himalayas. There are many ways in which geographical information sciences can help with fieldwork projects, these are just a few of the possible applications:

- *Logistics:* planning routes and navigation
- *Research:* mapping vegetation, wildlife, urbanisation, soils and geological features
- *Monitoring:* data logging of fire extents, forest loss, river channel changes
- *Conservation applications*: assessing biodiversity, park zonation, impact assessment
- *Technology transfer:* training local technical staff, donating hardware and shareware
- *Education:* maps for displays, involving schoolchildren with fieldwork.

1.1 Organisation of this manual

There are four parts to this manual: data, techniques, planning and results. The first set of chapters examines the various types of *geographical data* and the *geographical frameworks* fundamental to mapping. The main types of geographical data used on a fieldwork-based project are illustrated in Figure 1-2 and fall into two categories: (i) externally produced maps, images, photos and databases; and (ii) data collected by the expedition, using surveys and sampling. Data stored in a GIS can be manipulated, integrated or combined with other data sets for research or management purposes. The merging and statistical analysing of spatial (map-based) data sets provides new opportunities to formulate and test hypothesis on the organisation and functioning of the socio-economic or geo-ecological system being examined.

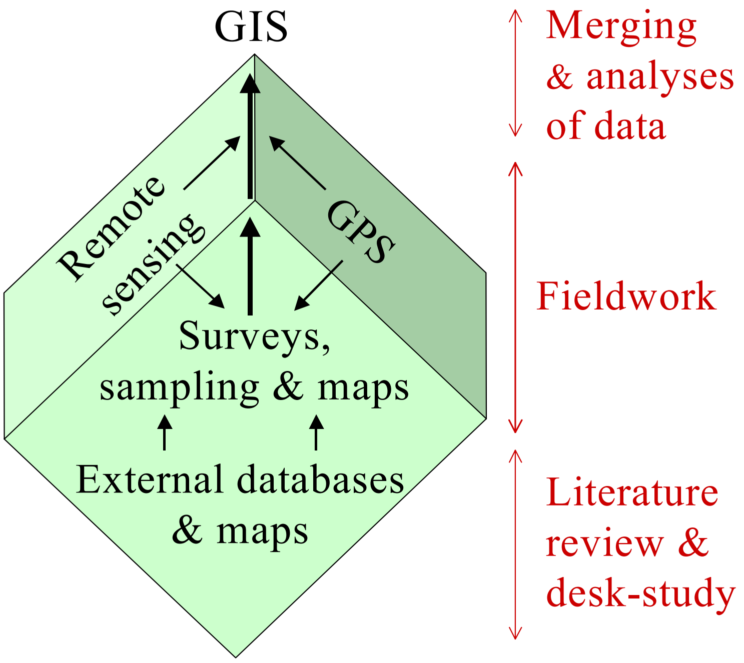


Figure 1-2 Data sources for expedition fieldwork. The arrows in the diamond indicate how data from one source can feed into another.

For data to be used in a GIS, they first has to be geo-referenced or ‘fixed’ in geographical space. The Geographical Framework chapter examines the various ways in which this can be done, and associated problems. Geo-referencing uses co-ordinates to locate points, with longitude and latitude being the fundamental system. Certain assumptions need to be made about the shape of the earth and these must be made clear in any GPS or GIS co-ordinate system, in the form of a spheroid and a datum. For most purposes in GIS and on paper maps, a flat 'x,y' co-ordinate system is far easier to deal with than a spheroidal 'longitude, latitude' system, so a further step is introduced to the geographical framework, that of map projections.

The next set of chapters provides detailed reviews of *GISci techniques* that are of use to expeditions, moving from geographical information systems (GIS) to remote sensing and photogrammetry, before considering surveying and the global positioning system (GPS). GIS has been in use for several decades and is having a pervasive impact on the conduct of expedition fieldwork. Burrough & McDonnel (1998) define a GIS as "a powerful set of tools for collecting, storing, retrieving at will, transformation and displaying spatial data from the real world for a particular set of purposes". Some outstanding textbooks on the application of GIS have recently been published (Burrough & McDonnell 1998, Johnston 1998, Heywood *et al.* 1998, Wadsworth & Treweek 1999, Longley *et al.* 2001). Expedition fieldwork will almost always involve the production of some form of map, both from field surveys and from the interpretation of features on images from aircraft and/or satellites, often using geomorphological and ecological mapping. GIS allows the integration and analysis of spatial databases, as well as providing a means of producing high-quality cartographic outputs.

Collecting images of Earth surface features, using aircraft or satellites, is a key aspect of *remote sensing*. Multi-spectral scanning instruments work in the visible and infrared part of the spectrum, whereas radar techniques use the microwave part, providing us with a unique capability to 'see' through cloud cover. Remote sensing has seen a rapid increase in computing power and software for image processing, allowing users to deal with growing volumes of data and more sophisticated image processing (Drury 2001, Gibson & Power 2000, Lillesand & Kiefer 2000, Mather 1999). An important stage of processing remotely sensed data for mapping is *classification*, which allows images to be more readily related to ecological and geomorphological features. Supervised classification requires the user to specify what classes are present and to identify homogeneous 'training' patches of each class: pixels (picture elements) are then allocated to one class or another, based on the unique *spectral signature* of each class. Surface areas of various land cover classes can be calculated by counting the number of pixels per class over the area of interest. A specialist aspect of remote sensing concerns collecting precise and accurate measurements from photographs, termed *photogrammetry*. With aerial photography being the most widely used form of remote sensing (Petrie 1999), a basic knowledge of photogrammetry is useful for most expedition fieldwork. The widespread use of flatbed scanners, allowing digital copies to be made of airphoto prints, and the advent of relatively low cost digital photogrammetry software, has allowed changes in land cover types and landforms to be mapped in detail, using airphoto archives that may cover the past 50 years. The costs of remote sensing vary considerably, depending on the technique and image processing, and are in the order of US$ 60 and US$ 1 per kilometre square for airborne and satellite images, respectively. Some archive imagery is now being made available free.

The *global positioning system* (GPS) is an American military navigation system, parts of which are available to the general public, that uses a network of satellites to locate GPS receivers positioned anywhere on the Earth's surface. As most GPS receivers are lightweight, portable and cheap (some costing less than $100), GPS usage has become a key aspect of expeditionary fieldwork. Accurate ground control is essential when mapping using remotely sensed data, as remotely sensed measurements can only be as reliable as the ground truth on which they are based. In 2000 the USA improved the accuracy of standard GPS signals, from ±100 m to ±10 m, greatly improving the ground-truthing of medium-scale satellite imagery, such as the widely-used Landsat data. Differential GPS (dGPS) can now readily provide centimetre to metre detail in field survey locations: this has helped to improve the accuracy of land cover mapping. Many GPS receivers are also data-loggers: you can devise a code system for a given set of features, use the GPS to locate those features, then plot and analyse the resulting features' locations by exporting the GPS data into a GIS.

The *expedition planning and management* chapter highlights how organising a successful scientific expedition takes a lot of planning, which in turn requires a lot of time: allow at least one year, preferably two, to get your ideas from the drawing board to fruition. Each stage in the 'life' of an expedition involves different types of data, and associated problems, many of which are highlighted in the questions below:

Planning:

- what is the overall aim or goal?
- what objectives do you hope to achieve?
- are there any maps or previous studies of your study area?
- which surveying and sampling techniques are appropriate?
- does the project have appropriate staff and equipment?
- how much time (and money) will the project require?

Fieldwork:

- where are your study areas located?
- how precise are your maps?
- how should field data be collected and stored?
- how will laptops, GPS units, etc, be powered in remote settings?

This handbook, plus the listings of websites given on the handbook's CD, will help you to find the answers to many of the questions posed above. A number of expedition training courses are also available, foremost among which are those run by the RGS-IBG's Expedition Advisory Centre (EAC).

The final set of chapters focuses on the fruits of your expeditionary labour, the *results* of your fieldwork. An expedition GIS allows you to start data entry, analysis and even presentation - in the form of maps - during your fieldwork, speeding up a process usually reserved (often with some dread) for after returning home. The main benefit, however, can be for local people involved in data collection, data analysis and the useful application of expedition findings.

The last part of the manual gives *examples* of the sorts of exploration fieldwork that you could follow using geographical information sciences. A wide range of applications are covered: from archaeology, ecology and geoscience, to natural resource management and socio-economic studies. Some more detailed case studies are also included, illustrating ways in which GIS, GPS and remote sensing can aid fieldwork in remote settings.

There is an ever-increasing range of topics that you could examine, especially with regard to risks from natural hazards, threats to biodiversity and the depletion of natural resources. Good luck with your endeavours!

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section A: Introduction

Chapter 2: The Geographical Framework

2 The Geographical Framework

One key feature of a GIS allows different types of data – whether from paper maps, GPS, satellite images or aerial photographs – to be integrated and overlaid: and that is the common geographical reference system, or the ‘geographical framework’. Once a piece of information (say the species of a tree or the height of a mountain) has been given a co-ordinate within a commonly recognised reference system (say from a GPS or from a map), it can be added to a GIS. This ability to integrate geographical data underlies all GIS functions, whether for mapping, navigation or more complex analysis. Therefore GIS users need an understanding of how geographical reference systems work. This perhaps applies particularly to expedition GISers, as they deal with data from a variety of sources. From being a traditionally specialist subject, geographical reference systems are becoming common currency for anyone using a GPS or GIS. Expeditioners often come face to face with co-ordinate issues only when things go wrong – layers don’t overlay, or GPS readings don’t match the map – so it is best to be prepared with an understanding of the underlying principles set out here.

In practice, as this chapter explains, there is no one single reference system always used world-wide, but a multitude, according to variations in the shape of the earth and the purposes of the GIS. However, provided that co-ordinates can be transformed between different systems, data can for most purposes be regarded as having a common referencing system. Most current GIS software allows near-instantaneous transformation between referencing systems, but it remains essential to know which system each of your data layers uses. Also note that a co-ordinate reference system is entirely independent of the nature of the data using it; raster, vector and any another spatial data types all ‘inhabit’ the same framework. (Raster is a grid-based data structure, while vector data use co-ordinate geometry; these are explored fully in Chapter 3.)

As examples, here are some situations where you will need to know which system to use:

- Telling a GPS how to display its co-ordinates;
- Saving data downloaded from a GPS;
- Digitizing features from paper maps;
- Fitting an aerial photograph or satellite image to a co-ordinate system;
- Importing data provided by others;
- Deciding on a suitable projection for cartographic output;
- Documenting data you share with others.

This chapter is in three sections: (i) the principles of how geographical reference systems are constructed, the longest section; (ii) details of one particular co-ordinate system, UTM, both as an example and because it is commonly used worldwide; and (iii) how to choose a suitable system for your GIS.

2.1 Geographical reference systems

This section covers the following four stages in turn:

- Find a regular mathematical shape that approximates to the shape of the earth, in the form of an *ellipsoid*. Once its position is defined in relation to the 'actual' earth, it provides a reference surface against on which measure positions – this is a *geodetic datum*.
- Measure positions of our points of interest – whatever they might be – using two angles, *latitude* and *longitude*, in relation to the datum.
- Transform our positions from the curved surface of the ellipsoid onto the flat surface of a map by means of a *map projection*.
- Finally, create a grid on the projected map to provide a convenient rectangular *co-ordinate system*.

Some aspects are considerably simplified here, but the aim is to provide enough detail to start working with GIS and GPS co-ordinate systems. Iliffe (2000) and Jones (1999) are recommended as being more comprehensive yet easy to read, while Maling (1992) is a definitive and very detailed text.

2.1.1 A reference model of the earth (the geodetic datum)

A ruler measures one-dimensional lengths from its zero mark: in other words, it provides a known reference point, also called a *datum*. In just the same way, when we want to establish positions of features on the earth's surface, we need a three-dimensional reference system against which to make measurements. In this case, the model is referred to as a *geodetic datum*.

It is desirable to have a model that fits well to the shape and size of the earth's surface, as that is where positions are measured. The science of determining the shape and size of the earth is known as *geodesy*, and the whole topic of geodesy would be a great deal simpler if the earth was exactly spherical; we would then be able to use a sphere of a certain radius as the geodetic datum for accurate positioning anywhere in the world. (Indeed, a spherical datum is used for some mapping at a global scale, but for expedition fieldwork and most other GIS applications we need greater accuracy.) Two factors mean that we need to refine the simple spherical model: the flattening of the earth at the poles, and irregularities in the shape of the earth.

2.1.1.1 Ellipsoidal datums for a flattened earth

The spin of the earth gives our planet a slight flattening at the poles and a bulge around the equator. Its radius at the equator is approximately 6,378 km, and at the poles is about 6,357 km, a difference of 21 km, or 0.33%. This shape is still relatively simple to model mathematically, using an ellipsoid. This is the figure formed by rotating an ellipse around its short axis (Figure 2-1).

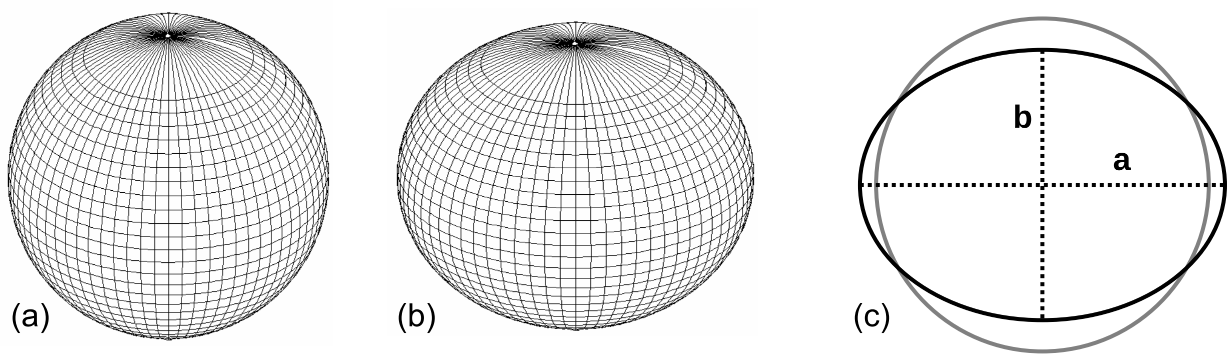


Figure 2-1 (a) A sphere, a relatively inaccurate approximation to the shape of the earth; (b) an ellipsoid, a better approximation; and (c) a cross-section of an ellipsoid, with a circle for comparison, showing the dimensions commonly used to describe an ellipsoid: a = longer radius, or semi-major axis, b = shorter radius, or semi-minor axis.

An optional note on terminology:

'Ellipsoid' and 'spheroid' are usually used interchangeably in geodesy. We have used 'ellipsoid' here as it more readily suggests the shape. However, in the field of geometry, an ellipsoid is a sphere that has been squashed both downwards and sideways – each of its three axes can be of different lengths. In a spheroid, by contrast, only two axes are different (Figure 2-1c), while in a sphere, all axes are of course the same. So while it is not wrong to use the term ellipsoid in this context, it would arguably be more accurate to use spheroid.

2.1.1.2 The 'lumpy' earth, or geoid

As well as being slightly flattened, the earth's 'average' surface level is slightly irregular: even if the planet was entirely sea-covered, its shape would be a slightly lumpy ellipsoid, rather than a smooth mathematical one. Again, the variations are slight, but in terms of positioning, they are significant. This is mainly due to variations in the density of different parts of the earth's mantle and crust. Instead of acting uniformly towards the earth's centre of mass, there are local variations in the direction of the earth's gravitational force, and the result is a slight deformation of surface features.

This irregular surface is called the *geoid*. More specifically, the geoid is a surface of equal gravitational potential. There are of course an infinite number of such surfaces closer or further to the earth, with higher or lower gravitational potential, but the earth's geoid is taken as the surface that coincides on average with the surface of the oceans (Figure 2-2).

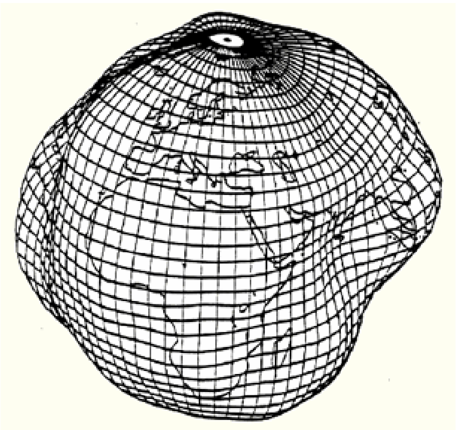


Figure 2-2 An exaggerated view of the geoid, a hypothetical sea-level surface covering the globe. In this figure, the separation between an ellipsoid surface and the geoid surface has been exaggerated by a factor of 15,000. The actual separation ranges from c. -100 m just below India to c. +70 m in the North Atlantic and near New Guinea.

2.1.1.3 Geodetic datums

Because of the irregularities in the shape of the earth, surveyors since the 18th Century have found that while one ellipsoid can provide a good fit to the geoid in one region of the world, a different ellipsoid will be needed in another region. As a result, a multitude of different geodetic datums has been developed. Each datum comprises two elements: (1) an ellipsoid of given shape and size, and (2) a definition of how the ellipsoid is positioned and oriented in relation to the geoid.

Historically, a local datum would be defined first by intensive trigonometric survey work on the ground then by complex calculations to find an optimal solution to fitting an ellipsoid – an astonishing achievement in pre-computer days. For local datums, the positioning of the ellipsoid is typically specified in relation to a particular marked survey point on the earth's surface, known as the initial point, whose position was accurately established by astronomical observations (Figure 2-3). Each of these local datums gives a good result locally, but is unlikely to produce a close fit in other parts of the world. Because of this, a plethora of datums has emerged, each one suited to a certain region, and often revised as surveying of the area improved.

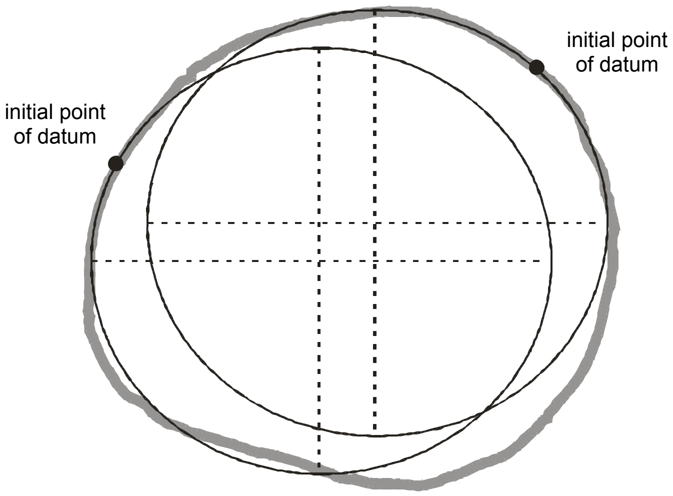


Figure 2-3 A cross section through the earth along the 0º/180º meridian. The grey line shows the mean sea-level surface, or geoid, overlain with two ellipsoids used by two different datums. Each datum is designed to provide a good local fit to the shape of the geoid, with its position defined by its initial point.

In many parts of the world frequented by expeditioners, the best maps are those produced from the late 19th Century up until the 1960s – before the advent of satellite surveying – and being of this era, are based on local datums. In much of eastern and southern Africa, for example, the 'Arc 1960' datum is used (itself a revision of Arc 1950); this is based on the 'Clarke 1880' ellipsoid and is fixed to an initial point in South Africa. The early dates that occur in many ellipsoid names are evidence of the success of the survey and calculation techniques developed in the 19th Century.

More recently, satellite-based measurements of the earth have made it possible to create *geocentric* ('earth-centred') datums. Instead of relating the ellipsoid to a point on the surface, a geocentric datum uses an ellipsoid whose centre coincides exactly with the centre of mass of the earth. The result is a reference system that provides a relatively good fit for all parts of the earth's geoidal surface (Figure 2-4).

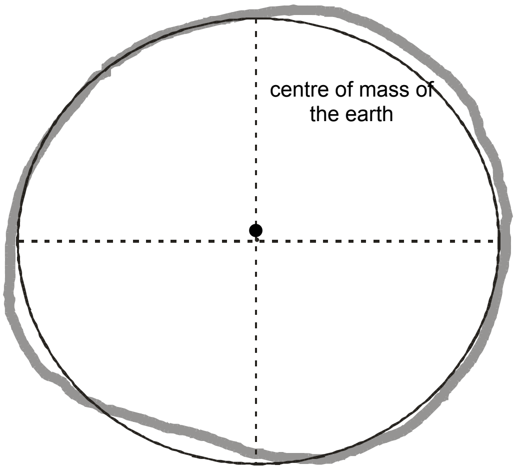


Figure 2-4 The geoid with the WGS84 ellipsoid. Undulations in the geoid have been exaggerated by a factor of 10,000 to make them conspicuous, and the flattening of the earth has also been exaggerated.

Such a datum comprises details of the shape and size of the ellipsoid and its orientation in relation to the earth's centre. This is of particular relevance to fieldworkers, as the GPS system uses a geocentric datum, the World Geodetic System 1984, or WGS84. Indeed, the separations between the ellipsoid and the geoid surfaces given in Figure 2-2 are derived from WGS84: the fit varies from -100 m to +70 m, a relatively tiny fraction of the size of the earth. Chapter 11 gives further details about how this difference is accounted for when measuring elevations with a GPS, as well as the difference between the geoid and the actual earth's surface (Figure 2-5).

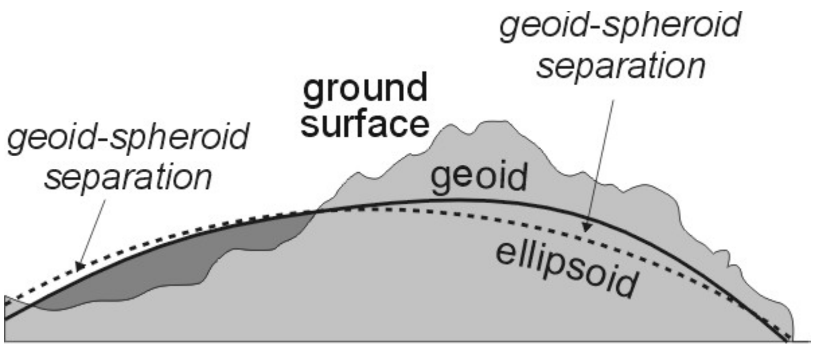


Figure 2-5 How a mathematically-defined ellipsoid surface might fit to the geoid; and the relationship between the geoid (defined as mean sea level around the world) and actual ground topography.

Because of its accuracy, global applicability and use by GPS, WGS84 has become widely used as the 'standard' datum for much geographical information. If all mapping efforts were to re-start from scratch now, almost all GIS and spatially-referenced data would use a system such as WGS84. However, the huge number of older maps and datasets means that many different datums are still in use, and you may decide to use one of the older ones for your GIS – for example, to match local maps. More advice is given later in this chapter. However, it is still vital to be able to transform co-ordinates between datums – for example, from GPS readings that use WGS84 to a GIS of a national park in Tanzania that uses Arc 1960. To do this, conversion software uses a table of known mathematical relationships between WGS84 and other datums. These tables specify the size and shape of the ellipsoid being used, and its position and angle – in three dimensions – in relation to the WGS84 datum. Knowing these parameters, datum conversions can be performed.

One confusion sometimes arises in relation to datums. If you stand at one point, without moving, you have a different co-ordinate depending on which datum you use. On reflection, this is hardly surprising: your measured position depends on where you measure from. By switching to a different datum, your frame of reference is different. The likely errors are significant in relation to the accuracy often needed for fieldwork, ranging from 10s up to 100s of metres. So, a rule of GIS is to make sure that you always (1) know which datum you are using for measuring positions, and (2) state which datum you are using when using, publishing or distributing positional data.

2.1.2 Positions on the ellipsoid: latitude and longitude

Positions on an ellipsoid can be specified with two angles: how far round (longitude) and how far up or down (latitude). Both angles are defined with reference to ‘zero’ positions. Conventionally, a circle passing through Greenwich, London, defines zero longitude, with positive values starting at 0° and increasing eastwards to 180°, and negative values ranging from 0° westwards to 180°. The circle on the plane of the earth’s rotation – the equator, also the widest section of the ellipsoid – defines zero latitude, with values increasing northwards to 90°N and decreasing southwards to 90°S.

For example, the RGS-IBG in London is approximately at 0°10’W, 51°31’N, using WGS84 datum. Simply to say that “the RGS-IBG is at 0°10’W, 51°31’N” is incomplete; for accuracy, we need to know the datum too.

The way in which latitudes and longitudes are written on paper maps and documents differs in several ways from their digital representations. These can easily create confusion, so are explained here.

2.1.2.1 Units: degrees, minutes and seconds

Latitude and longitude, being angular measures, are normally expressed in degrees. To provide greater precision, several different sub-divisions of a degree have historically been used, and you are likely to come across various formats (Table 2-1). Degrees are conventionally divided into 60 minutes (also called minutes of arc, or arc-minutes), and each minute further sub-divided into 60 seconds (seconds of arc, or arc-seconds). These units are clearly not convenient for most digital applications, so the ‘decimal degrees’ format is usually used in GIS software.

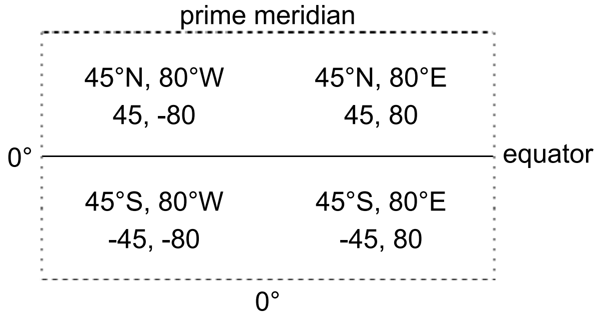
Table 2-1 Different units and formats for expressing latitudes and longitudes.

Degrees	D	52°N, 0°W
Decimal degrees	DD	51.521°N, 0.178°W
Degrees, decimal minutes	DDM	51°31’N, 0°10’W
Degrees, minutes, seconds	DMS	51°31’14”N, 0°10’39”W

2.1.2.2 *Directions of angles: north – south, east – west*

There is no standard for expressing the direction of angles. Printed references normally use ‘N’/‘S’ and ‘E’/‘W’, along with conventional signs for degrees (°), minutes (’) and seconds (”). When co-ordinates are stored digitally, however, these signs are not normally used, and directions are indicated by positive or negative numbers. Table 2-2 shows examples.

Table 2-2 Ways of expressing a co-ordinate in each of the four quadrants, showing a conventional printed format above and a digital format below.



45°N, 80°W 45, -80	45°N, 80°E 45, 80
45°S, 80°W -45, -80	45°S, 80°E -45, 80

When adding co-ordinates to your GIS from sources such as atlas, gazetteers or printed references, it is often necessary to convert from ‘DMS’ to ‘DD’. This is easily done in a spreadsheet, using a formula such as: DD = degrees + (minutes / 60) + (seconds / 3600). Care is needed to ensure that the right sign – positive or negative – is used.

2.1.2.3 *Order: ‘latitude, longitude’ or ‘longitude, latitude’*

A final difference is in the order. Co-ordinates are conventionally given with their latitude first, followed by their longitude: the shorthand “lat-long” is often heard. By contrast, most GIS software assumes, if not told otherwise, that co-ordinates follow the ‘x, y’ convention. In other words, longitude first, then latitude. If you are storing co-ordinates in a spreadsheet, make sure that columns are labelled clearly and if necessary change the order of the columns. I have overlooked this point more than once, only to find that a set of co-ordinates imported from a spreadsheet is mapped out ‘on its side’ in a GIS. This is particularly liable to happen when using geographic co-ordinates downloaded from a GPS (see Section 4.7).

What does a degree look like?

Having an idea of the ground distance represented by a degree helps to visualise latitudes and longitudes. The original definition of the metre is one ten-millionth of the distance between the equator and the poles: 90° of latitude = 10,000,000 m. Hence 1° of latitude = 111,111 m, or about 111 km. This is a useful rule of thumb when estimating the *precision* of a co-ordinate: the first decimal place corresponds to about 11 km, the second decimal place to 1.1 km, and so on. So, without doing any complicated calculations, we know that the latitude figure “51.52°N” is precise to about 1 km. Clearly, this applies only to degrees of *latitude*, which are always equally spaced (hence the name ‘parallel’ for lines of latitude). Degrees of *longitude* are also about 111 km at the equator, but from there become smaller as they converge towards the poles.

To summarise so far, we now have a reference system from which to measure positions on the earth’s surface. This gives us a precise way to specify the longitude and latitude of any feature, anywhere, on the earth, with reference to an ellipsoid; and we know how these positions relate to the actual shape of the earth – the geoid – by means of the datum.

2.1.3 From curved surface to flat map: map projections

Having established the locations of features on a rounded surface, the next step is to transfer them onto a flat surface – the process of map projection. Projections originated as a way of creating paper maps, but have an additional role in a GIS context: even if data are never plotted to a sheet of paper, much geographical analysis in a GIS relies on measurements of distances, lengths and areas. Such analysis is impractical or impossible using latitude/longitude co-ordinates, because they are angular measurements. Once these data have been projected onto a plane surface, the co-ordinates typically use metres or feet as units, so they become far more amenable to measurement and analysis, as well as being more intuitive and comprehensible to their human users.

The best known and ‘original’ projection is the Mercator projection (Figure 2-6), named after its inventor Gerardus Mercator, 1512-1594. Although it is well known now for its distortions in distances and areas, it has a unique application: a straight line drawn on a Mercator map corresponds to a compass bearing on the ground (or at sea), a property which was vital to the marine navigators who were (and still are) its main users. It is also worth noting that distances and areas are relatively accurate close to the equator, where the projection cylinder is in contact with the earth – this is known as the line of tangency.

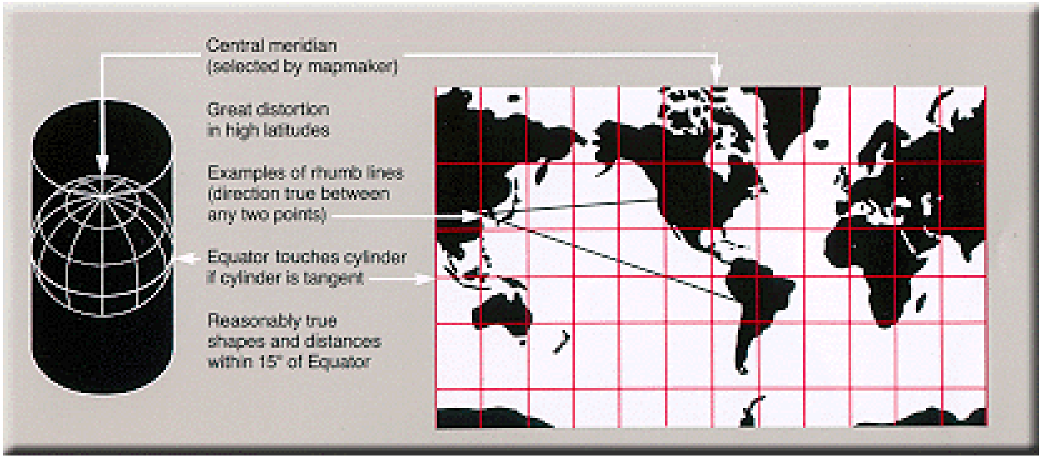


Figure 2-6 Construction of the Mercator projection
(from http://erg.usgs.gov/isb/pubs/MapProjections/projections.html).

The example of the Mercator projection shows us that some compromise is inevitable when making a projection: while a globe can accurately portray directions, distances, areas and shapes, one or more of these properties becomes distorted when projected onto a plane. As a result, a huge variety of projections have been devised, each suitable for particular applications, different parts of the world and different areas of interest. Large scale topographic mapping at 1:50,000, for example, will use a different projection to a map showing population distribution over an entire continent.

Some projections commonly used for fieldwork and research are shown below, but for more details readers are referred to a US Geological Survey webpage:
erg.usgs.gov/isb/pubs/MapProjections/projections.html

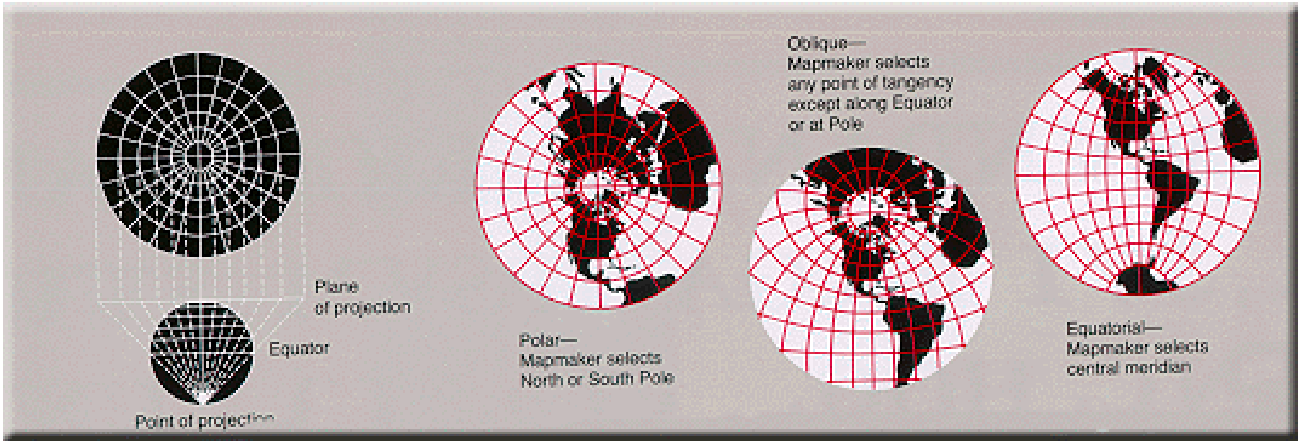


Figure 2-7 The stereographic projection, often used for mapping Arctic areas, and by the USGS and British Antarctic Survey for Antarctic mapping. Note that unlike the Mercator projection which uses a cylinder, it is projected onto a plane – this is known as an azimuthal *projection.*

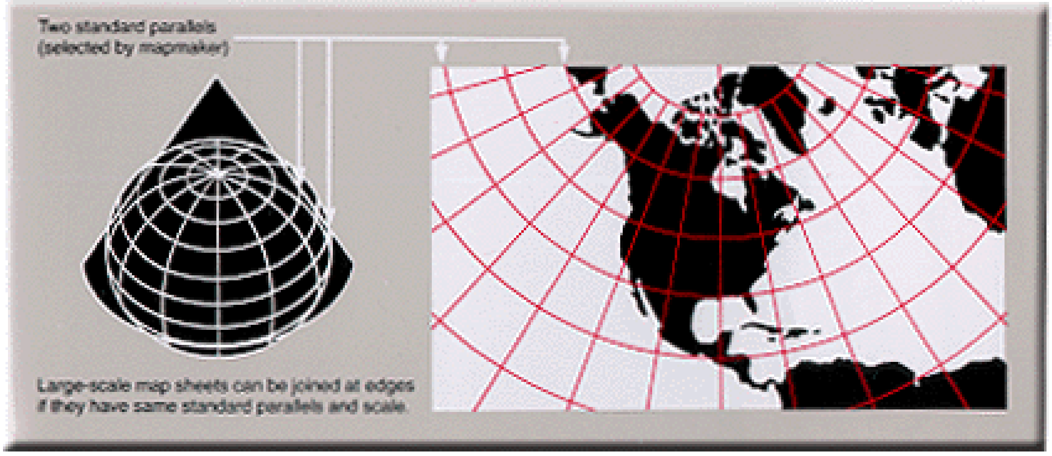


Figure 2-8 Lambert Conformal Conic projection, used for large-scale topographic mapping of the USA. Also suitable for topographic mapping of small areas or larger areas that run east-west. Following the previous two figures, this projection uses a third construction method, being projected onto a cone, hence the name conic.

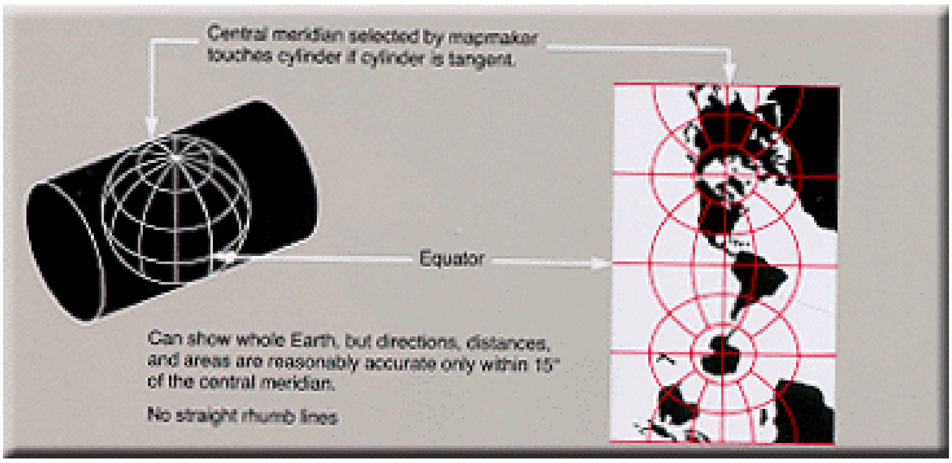


Figure 2-9 Transverse Mercator: identical in construction to the Mercator, except that the cylinder is oriented transversely. The line of tangency, where scale distortion is at its minimum, therefore follows a meridian, i.e. a line of longitude; this property is used in the Universal Transverse Mercator co-ordinate system, described in a following section.

2.1.4 A rectangular co-ordinate grid

We now come to the final stage: once projected, a map is usually given a Cartesian co-ordinate system – one based on a flat, square grid – to refer to locations and to make measurements. This can be visualised as a grid overlaid on the map. Distances along the *x* axis are often referred to as *eastings* (i.e. measurements made in an eastward direction), while *y* axis distances are *northings* (measurements made northwards). The units are usually 'real world' ones: metres or (particularly in the USA) feet. An example of grid co-ordinates is shown in Figure 2-10, or look at any Ordnance Survey map to see similar examples.

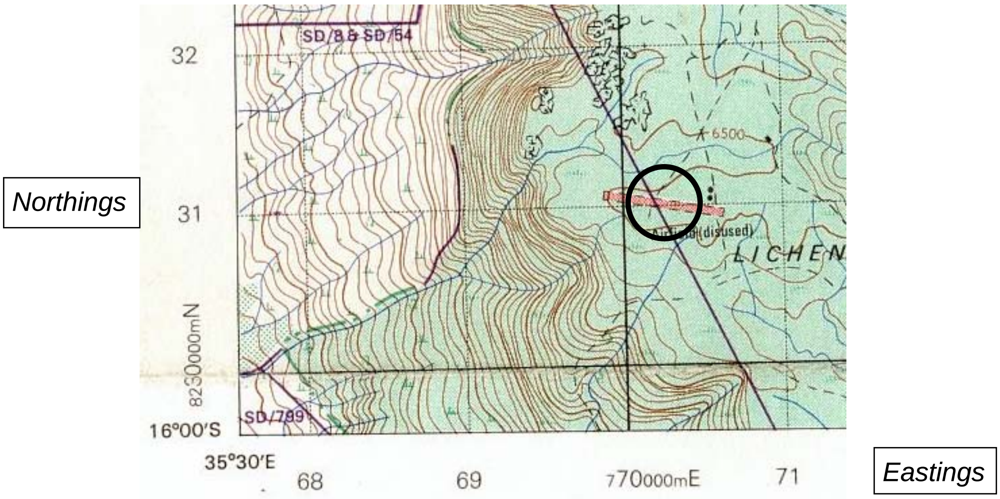


Figure 2-10 Scanned portion of a 1:50,000 map of Malawi showing its co-ordinate grid and the labels. Grid lines are 1,000 m (1 km) apart. Note how most of the labels comprise just two digits taken from the full co-ordinate. This is done to save space and to make map-reading easier. On this map, the complete co-ordinates are given every 10 km (in this case, 770000 east and 8230000 north). Note that the complete co-ordinate shows the unit (metres) and the direction (E/N). For example, the eastern end of the disused airfield (circled) is at: 770625mE, 8230950mN.

Most countries of the world have defined their own national grid systems, while an international system (UTM, detailed in the next section) is also widely used around the world. Expedition GISers therefore almost always use an existing national grid, for compatibility with local maps and with standard GPS/GIS settings. As an example, Section 11.4 shows the settings needed to make GPS readings correspond with co-ordinates on a 1960s map series of Vietnam. The following section now explains how such a co-ordinate grid is put together. There might be cases where a field project defines its own local grid system; this is unlikely, but might be done, for example, on a sub-Antarctic island, or where a local grid sampling system is needed. In such cases, if a grid is defined using the concepts shown here, it will be possible to transform from the local non-standard co-ordinate system to a standard system. A neat – if rather unusual – example was the creation of a new national grid for the Maldive islands; because they straddle the equator (just), UTM would have been inappropriate, so a modified UTM grid was defined, with an origin 100 km south of the equator (Hobbs 2003).

How is the co-ordinate grid positioned in relation to a map that has been projected? Horizontally, it is usually aligned with the central meridian (line of longitude) of the map projection. This means that the grid is upright – north-aligned – at the centre of the mapped area. Vertically, it is usually based on a particular parallel (line of latitude), often the equator, although other parallels may be used. The British national grid, for example, is aligned horizontally with the central meridian of the underlying map projection (2°W) and vertically with a parallel just to the south of the British land mass (49°N).

The grid also needs an origin, the point defining the 0,0 co-ordinate. The 'natural' origin for x co-ordinates (eastings) is the central meridian of the map projection. However, to avoid negative values occurring to the west of this line, a 'false easting' is typically used, as shown in Figure 2-11. This is a value added to all *x* co-ordinates, in effect creating an origin to the west of the area being mapped. Thus all *x* co-ordinates within the mapped area will be positive. In the British national grid, for example, the false easting is 400,000, which means that eastings are measured from a line lying offshore to the west of Britain.

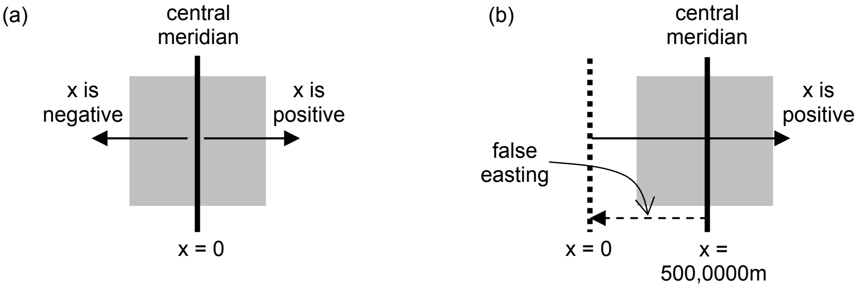


Figure 2-11 Using a false easting to create positive co-ordinates within the mapped area (marked in grey). (a) Using the central meridian as the origin results in negative values to the west. (b) Adding a false easting of 500,000m ensures positive values throughout the mapped area.

The origin for *y* co-ordinates (northings) is usually a particular parallel (line of latitude). In many cases the equator is used; this provides a 'natural' latitude of origin. Alternatively, a parallel closer to the area being mapped may be used: in the British example, it is 49ºN. Again, an offset – the false northing – may be applied to avoid negative numbers. The false northing value is added to all *y* co-ordinates to make them positive. Figure 2-12 shows a common case, where a large false northing is applied in order to make southern hemisphere co-ordinates positive when measured relative to the equator. The case of the British national grid is slightly unusual: negative northings do not occur in any case, as the latitude of origin has been chosen to lie south of the British landmass. However, a false northing of 100,000m is still used, in order to bring the origin closer to the mapped area.

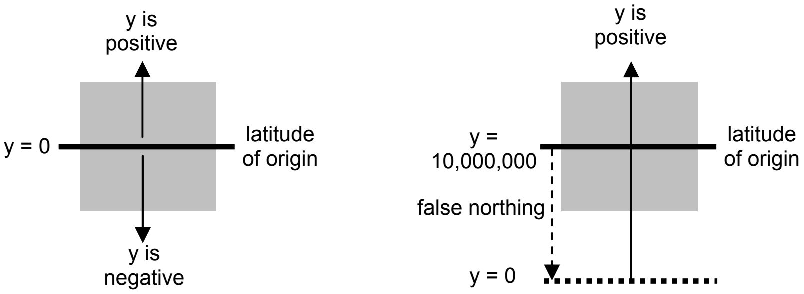
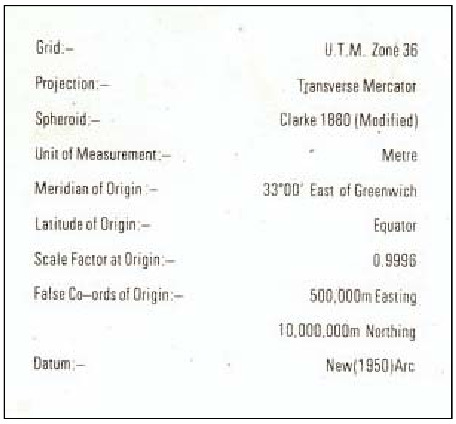


Figure 2-12 Using a false northing to create positive co-ordinates within an area being mapped (shown in grey). In this case, the false northing is 10,000,000m.

The South African grid system is an exception: it does not use a false northing, so co-ordinates have an increasingly large negative value as one moves southwards from the equator. This system is shared by Lesotho and Swaziland (Chief Directorate: Surveys & Mapping 2003).

We now have all of the information necessary to describe how the position of a point – the mountain we are climbing or the tree we are studying – can be (1) measured (using a datum), (2) mapped (using a projection) and (3) given a co-ordinate (using a grid). At best, maps show all of these details (as in Figure 2-13), allowing you to configure your GPS and GIS correctly. Similarly, all GIS data should have this information specified, either within a particular file format or in an associated document (metadata). Often, however, some further detective work may be required to find all of the parameters, for example by asking a national mapping agency.



Grid:–	U.T.M. Zone 36
Projection:–	Transverse Mercator
Spheroid:–	Clarke 1880 (Modified)
Unit of Measurement:–	Metre
Meridian of Origin:–	33°00′ East of Greenwich
Latitude of Origin:–	Equator
Scale Factor at Origin:–	0.9996
False Co–ords of Origin:–	500,000m Easting
	10,000,000m Northing
Datum:–	New(1950)Arc

Figure 2-13 Full details of the projection and co-ordinate system printed in the margin of the Malawi map shown in Figure 2-10.

2.2 The Universal Transverse Mercator (UTM) system

UTM is not a particular projection or co-ordinate system, but rather defines a set of map projections and co-ordinate systems designed for large scale mapping in all parts of the world. Many national mapping agencies use UTM for their topographic map series; Landsat images provided free by the Global Land Cover Facility (University of Maryland; see Chapter 5) use UTM; and it is supported by almost all GIS software and GPS receivers. Thus it is a common choice for expeditions, whether researchers, adventurers, or both.

2.2.1 Map projection and UTM zones

As the name suggests, UTM uses the Transverse Mercator projection. In fact, it comprises 60 different Transverse Mercator projections, each one with a central meridian 6º greater than the previous one. Visualise the cylinder in Figure 2-9 being rotated in 6º increments, each time forming a new central meridian where it is tangent with (touches) the earth's surface. This results in 60 'zones' around the world, each 6º wide, within which map distortions are insignificant. In the northern hemisphere these zones are numbered 1N to 60N, while in the south they are 1S to 60S (Figure 2-14). Thus, for example, most of Madagascar lies in Zone 38S, which spans from 42ºE to 48ºE, with a central meridian of 45ºE.

2.2.2 UTM co-ordinates

The unit for all UTM co-ordinates is metres. Northings (*y* co-ordinates) are measured with reference to the equator, increasing northwards from 0 m at the equator. In the southern hemisphere, a false northing of 10,000,000 m is applied, to ensure positive co-ordinate values (see Section 2.1.4).

Note that all *y* co-ordinates, whether north or south of the equator, therefore lie within the same range of values (0 to 10 million): a *y* value by itself gives no indication of which hemisphere it is in. It is therefore vital to specify whether a co-ordinates is in a southern or a northern UTM zone.

UTM eastings (*x* co-ordinates) are measured with reference to the central meridian of the zone, with a false easting of 500,000 m. In this case, note that there is no indication in the co-ordinate itself of its zone; a given easting could exist within any one of the 60 zones, so be sure to specify the zone number.

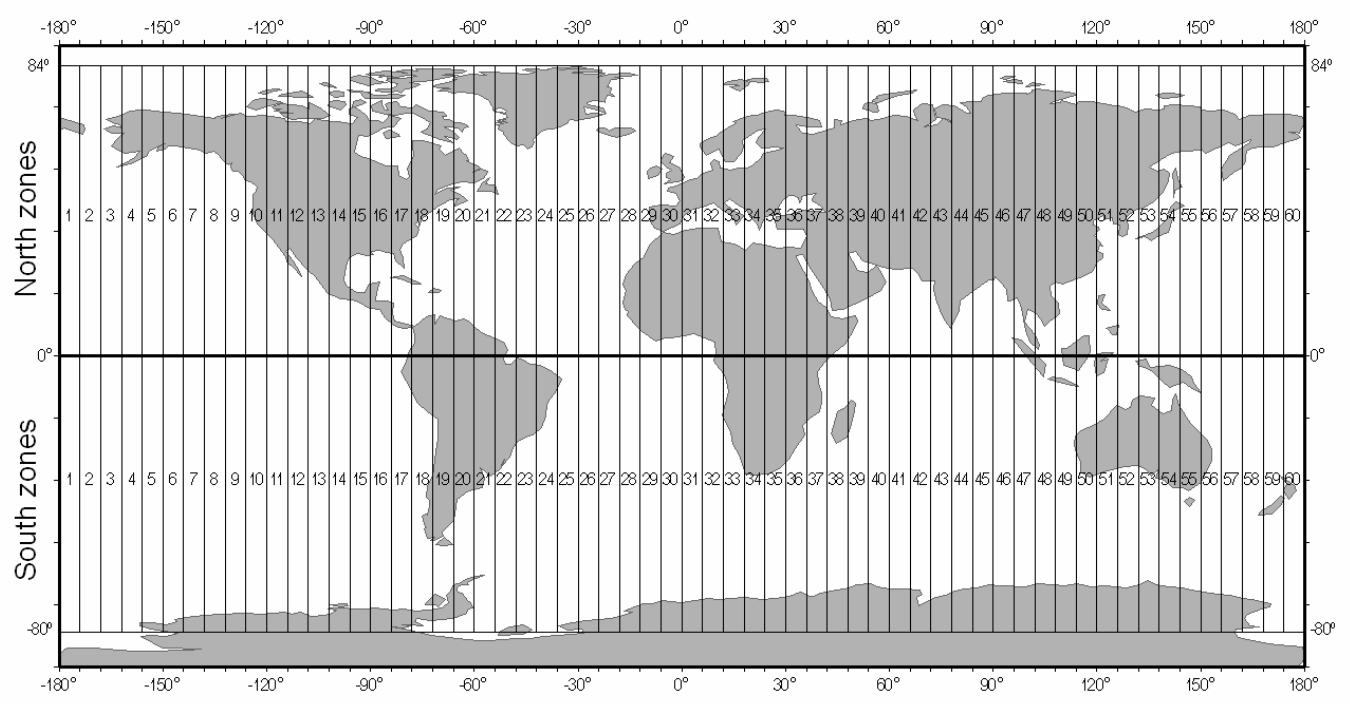


Figure 2-14 The UTM zones, numbered 1-60. Northern ones are denoted as Zone 1N, 2N, etc, and southern ones as 1S, 2S, etc. Most of Madagascar, for example, is in Zone 38S.

2.2.3 Datum

While UTM specifies a projection and grid co-ordinate system, it does not specify a particular datum, leaving the choice of datum to be appropriate to the local mapping area. Indeed, the UTM definition has existed since the 1940s, well before global geocentric datums such as WGS84 came into existence. Figure 2-13 for example shows the details of a UTM map that uses the Arc 1950 datum. More recent UTM maps and GIS data sets, however, tend to use WGS84, and this is the choice of the Landsat images available on the GCLF website (http://glcf.umiacs.umd.edu/data/).

2.2.4 Scale factor in UTM

Within each zone, distortion in scale is minimised by the use of a scale factor. Normally, the scale of a map is 'true' along the central meridian of a Transverse Mercator projection; the scale factor is said to be 1. To the east and west, however, scale is increasingly distorted: the scale factor increases above 1, and the effect is to increase distances between projected points.

While this effect is negligible close to the central meridian, the distortion increases steeply away from centre towards the edges of the UTM zone. To minimise this effect, the east-west scale across the zone is reduced by a given scale factor, applied from the central meridian outward. In the UTM system (as with British OS maps), the scale factor is 0.9996. This means that the scale along the central meridian is reduced by this factor, but the advantage is that there are then two lines, one to either side of the central meridian, along which the scale factor is 1 – no scale distortion. Further out towards the edges of the zone, scale distortion is significantly reduced. Figure 2-13 shows the scale factor quoted on the Malawi map example.

What happens at the edge of a UTM zone

Reducing scale distortion at the edges of UTM zones has a useful effect for fieldworkers: if your study area happens to stray across the defined limit of a UTM zone (we cannot expect national parks or species ranges to stay within UTM zones!), then there is little harm in keeping to the same UTM parameters beyond the zone limits. Figure 2-16 shows the distortion that occurs when the same projection is used to map areas far outside a zone: distortion is large far away from the central zone, but immediately to each side the distortion is, for most purposes, insignificant.

Indeed, some single-sheet national maps use precisely this method. Although Tanzania spans three UTM zones (35S, 36S, 37S), Harms-Verlag based its recent 1:1,400,000 map of Tanzania on the projection for the UTM zone at the centre of the map, i.e. zone 36S. Note that at this scale, distortions are not significant, but using the same UTM system across such a wide area would not be acceptable on 1:50,000 maps, as distortions would become far more apparent. However, it does demonstrate that this method can be applied, with care, in appropriate situations.

This method will work when storing data and creating maps in a GIS. However, GPS units will always display the ‘right’ UTM zone, calculated on the basis of their known longitude; they cannot normally be forced to display co-ordinates for a UTM zone that they are not in.

2.2.5 Example of a UTM co-ordinate

The following two figures show how the location of a mountain peak in southern Ethiopia is mapped to a UTM co-ordinate. In this case, the full co-ordinate is:

700600mE, 987700mN, UTM zone 36N, WGS84 datum.

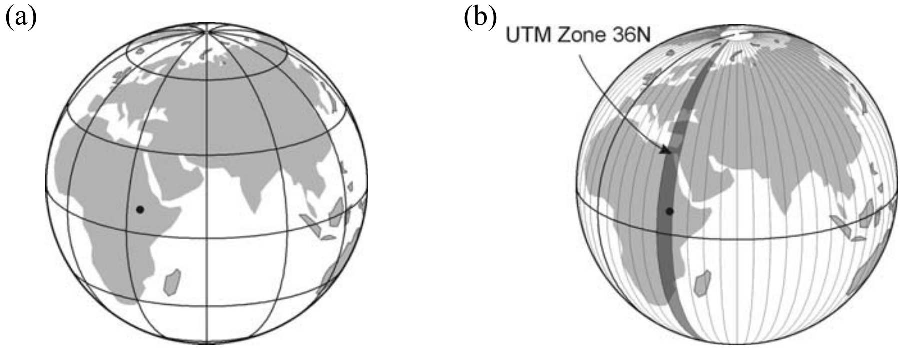


Figure 2-15 (a) Location of the point on the globe. (b) The relevant UTM zone is 36N.

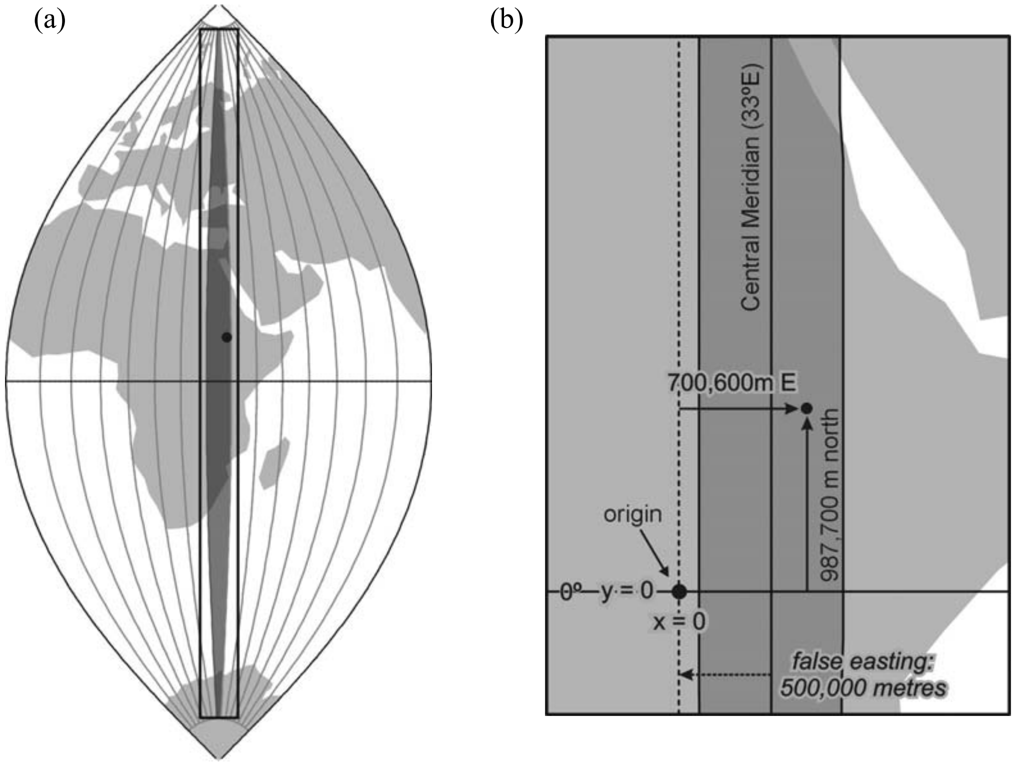
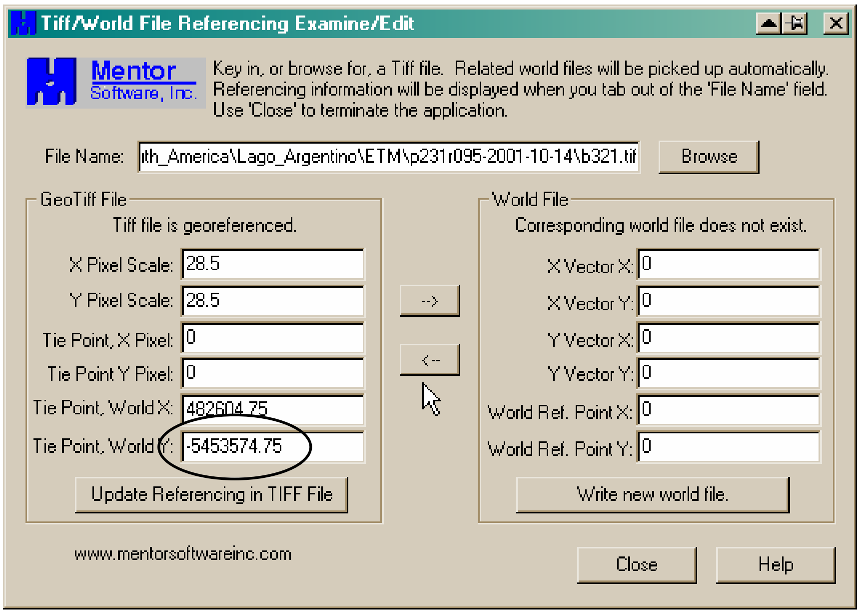


Figure 2-16 (a) UTM zone 36 after being projected onto a plane using Transverse Mercator. Adjacent UTM zones are also displayed to show the effect of scale distortion away from the central meridian. The black box shows the rectangular extent of the co-ordinate grid. (b) Detail showing the x *co-ordinate in relation to the central meridian, with a false easting of 500,000m, and the* y *co-ordinate in relation to the equator. Being in the northern hemisphere, there is no false northing.*

Landsat images from the Global Land Cover Facility website: which UTM zone?

There appears to be a UTM zone problem with some of the free satellite images available from the GLCF website, affecting Landsat ETM images in the southern hemisphere that use the TIFF file format. The co-ordinates of these images are given as negative values; in other words, no false northing has been applied, and the images are incorrectly referenced to UTM north zones.

This can be corrected with a free utility programme called GeoTiffExaminer, which can read and modify the geo-referencing information held in GeoTIFF files. GeoTIFFs are the same as normal TIFF bitmap files, but with added information about their geographical co-ordinates contained in their headers. Use GeoTiffExaminer to read the header information of a TIFF file (Figure 2-17), and note the negative value of the ‘Tie Point, World Y’. Add the false northing (10,000,000) to this number and click ‘Update Referencing in TIFF File’. This corrects the co-ordinate referencing error. Then when you open the image in a GIS programme, you also need to modify the UTM zone designation. For example, if the zone is read by the software as 18N (or if it is not read at all), then change it to 18S. The image should now be correctly referenced, and other data layers should overlay it in the right positions. This is a fine example of a problem whose solution requires an understanding of co-ordinate systems!



Tiff/World File Referencing Examine/Edit

Key in, or browse for, a Tiff file. Related world files will be picked up automatically.
Referencing information will be displayed when you tab out of the 'File Name' field.
Use 'Close' to terminate the application.

File Name: ith_America\Lago_Argentino\ETM\p231r095-2001-10-14\b321.tif

GeoTiff File
Tiff file is georeferenced.
X Pixel Scale: 28.5
Y Pixel Scale: 28.5
Tie Point, X Pixel: 0
Tie Point Y Pixel: 0
Tie Point, World X: 482604.75
Tie Point, World Y: -5453574.75

World File
Corresponding world file does not exist.
X Vector X: 0
X Vector Y: 0
Y Vector X: 0
Y Vector Y: 0
World Ref. Point X: 0
World Ref. Point Y: 0

www.mentorsoftwareinc.com

Figure 2-17 Using GeoTiffExaminer to correct an error that occurs in some southern hemisphere Landsat ETM images. Increase the circled value by 10,000,000, in this case to 4,546,425.25. Then in a GIS programme, specify a southern rather than northern UTM zone.

2.3 Choosing a co-ordinate system for your GIS

2.3.1 Are you working with an existing GIS?

If someone else has already established a GIS of your study area, it is worth considering the same co-ordinate system as theirs: you are able to use their data to get started, and later on will be able to share your data with them. Check that your GIS software and your GPS unit support the system.

Example: in 2002 a small team helped map the road network and establish a field GIS in Mkomazi Game Reserve, Tanzania. They based the GIS on the same co-ordinate system as the one used by an RGS project in Mkomazi during the 1990s. The team was thus able to use existing GIS data layers, which were available for download on the internet. In this case, the system was: UTM, zone 37 south, Arc 1960 datum.

2.3.2 Will you be working with existing maps?

If detailed local maps already exist, for example at a scale of 1:50,000 or 1:100,000, then using the same co-ordinate system as the maps has several advantages: (1) a suitable projection and datum has been chosen for the area; (2) features can be digitised from the maps without the need for re-projection; (3) and you will be able to read the maps using the same co-ordinate system as your GIS, making for easier planning and navigation.

Example: the same team was asked to produce a road map of Mikumi National Park in Tanzania, for use by both researchers and tourists. 1:50,000 topographic maps were available, so they based the GIS on the same co-ordinate system as the maps (UTM / Arc 1960 datum). GPS units were also set to the same system. Thus park staff and researchers are now using the maps, GPS and GIS, all with the same co-ordinate system and without need for conversions.

2.3.3 Do you have existing data already geo-rectified and geo-registered?

It might be that one existing data set will determine the co-ordinate system you use. If you start with a series of large raster images, for example, it would be time-consuming to re-project them, and you would loose some accuracy in the process. The issue of accuracy is particularly important if you are undertaking any numerical analysis of the raster data sets: the actual values of each pixel are not very important if the image is just being displayed or printed as a map, but if it is being used for change analysis or correlation, then pixel values are critical. If the images are already in an appropriate projection for the study area, it might be best to keep that projection, and re-project other vector data to match it. In most cases, vector data can be re-projected with little or no degradation.

Example: several sets of environmental data were used for a GIS model of mammal distribution in Mkomazi Game Reserve, Tanzania. The main set comprised satellite imagery (AVHRR), downloaded from the internet. This comprised images taken every 10 days over two years, each with several different wavebands, making hundreds of images in all. They were already geo-registered and projected in Sinusoidal projection. All my other data, however, used the same projection as the Tanzanian 1:50,000 maps. It was far easier to re-project the vector data to the Sinusoidal projection than to re-project hundreds of satellite images.

2.3.4 Starting from scratch?

If there is no existing mapping or GIS data, what system to use? If you find yourself in the field with no idea which co-ordinate system to use, the best option is to save data in unprojected co-ordinates – in other words, latitude and longitude – using WGS84 datum. This provides the maximum flexibility for transforming the data to other systems should different needs arise in the future.

Where a paper map is to be printed or analysis to be undertaken, then UTM typically provides a good answer, being designed for large scale use anywhere in the world. It is also commonly recognised and accurately handled by most GIS software and GPS units.

Other considerations may apply to more specialist applications, particularly in large study areas at country or continental scale, although these tend to be outside the scope of most expeditions.

Example: Bogda Shan expedition in NW China was mapping a relatively small mountainous area with no existing detailed maps. They chose the UTM co-ordinate system, with the WGS84 datum.

2.4 Conclusion

Perhaps most importantly is to (i) know which projection, datum and co-ordinate system you are using and (ii) always record it, in fieldnotes, the GIS, and reports and publications. Why? You can re-project the data if you decide another system would be better; and others will be able to use your valuable geographical information.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section A: Introduction

Chapter 3: Geographical Information Systems (GIS)

3 Geographical Information Systems (GIS)

3.1 Introduction

The Chorley Report for the UK Department of the Environment (1987) defined GIS as:

> *"…a system for capturing, storing, checking, integrating, manipulating, analysing and displaying data which are spatially referenced to the Earth."*

The primary type of GIS output is usually a map of some sort, together with a statistical analysis of the features of interest. To obtain the optimal return on time invested in a GIS, new users should examine how integration and analyses were achieved in other GIS applications. This should be followed by reflection on how GIS can be used within a given project.

The development of GIS can be traced back to three types of computer application, developed over the 1960s: computer-aided design (CAD), automated mapping and facilities management. CAD systems are key features of architectural and engineering design; however they are limited to graphics data, rather than map-based information. Automated mapping systems link external files to electronic maps: they are of particular use for mapping and managing utilities and complex facilities, such as airports and hospitals. The Canadian Geographic Information System (CGIS) is widely regarded as the first fully-functional GIS, developed over 1964-1971, to map and analyse land inventory data. CGIS demonstrated how GIS could produce huge savings in time, with the automation of map outputs, digital storage and retrieval, and computer-based analysis of map-linked databases. A comparative analysis of the 6,000 maps in the CGIS archive would have taken 1650 years to carry out manually, but only took a few days using CGIS. As computers have become faster and better at both processing and storing large volumes of data, so GIS software has evolved from slow-running and complex systems, to user-friendly 'desk-top mapping' systems with many users. For more detailed coverage of GIS, the reader is referred to books by Burrough and McDonnell (1998), DeMers (2000) and Heywood *et al.* (1998). Wadsworth and Treweek (1999) focus on GIS for ecologists.

3.2 Data input and organisation

Data organisation and structure is key to successful GIS implementation. Geographical data are available at widely differing scales, so the first decision a user has to make is, *at what scale does the project data need to be captured?* Appropriate scale is determined by various factors including the size of the project area, the size of the features to be analysed and the scale of available data, such as topographic maps.

Data can be input into a GIS by various methods, notably scanning, digitizing and databases, as well as direct importing of pre-existing digital data, such as satellite imagery or commercially produced digital maps. The type of data input will determine which data model is the most appropriate for a particular project. Three data models are widely available: *raster, vector* and *hybrid*. The initial decision on which to use can have profound and far-reaching consequences to a project if the data model is not suitable.

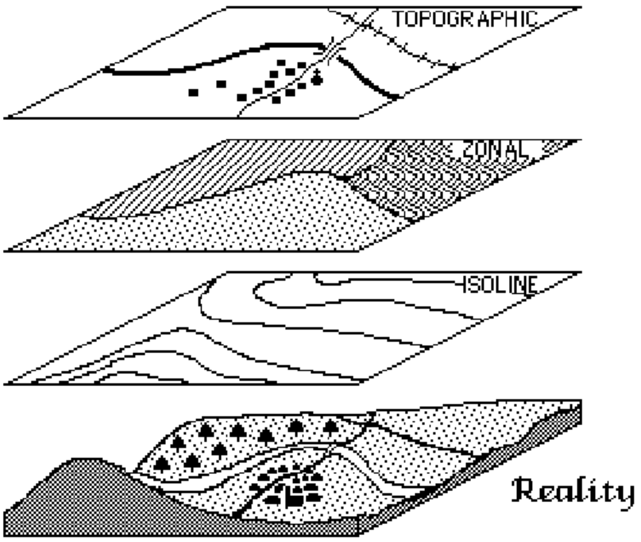


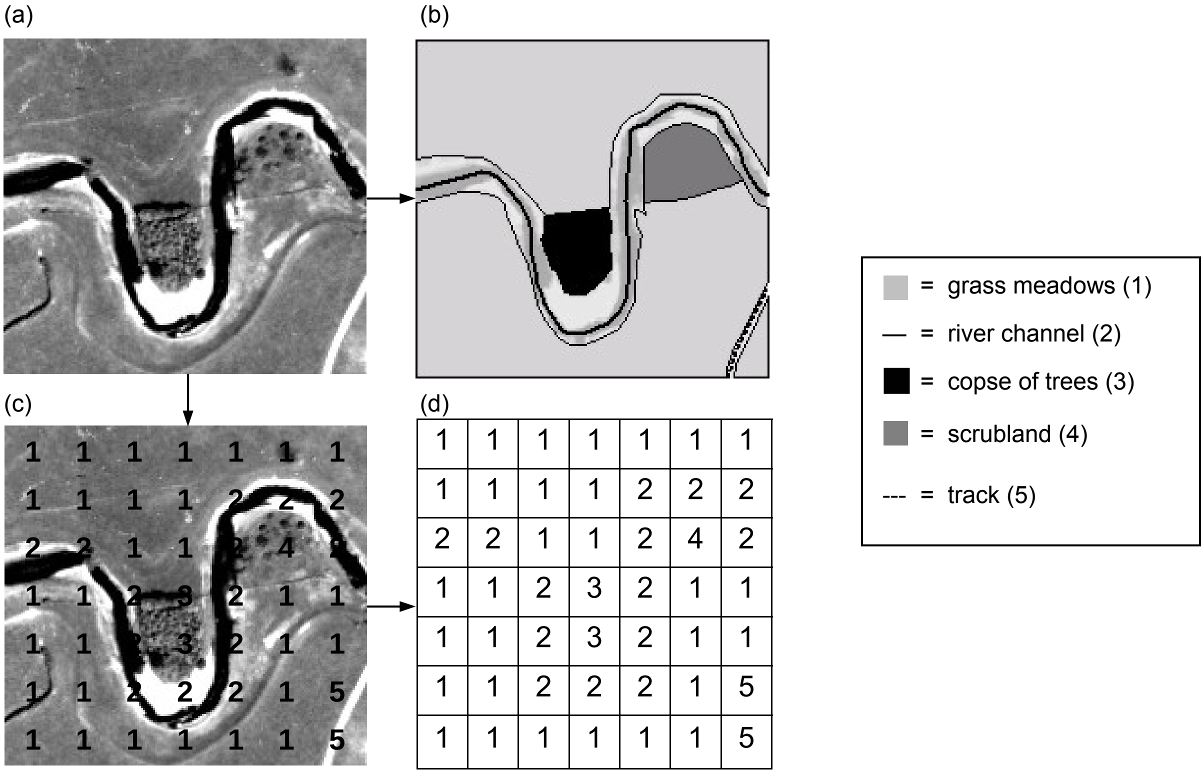
Figure 3-1 An example of GIS data layers (source: Jonathan Raper).

3.2.1 Raster data

The raster model uses a large number of (usually) square ‘cells’ to represent a geographical surface. These cells are also referred to as *pixels* (picture elements). Each pixel covers a set proportion of the Earth’s surface, the size of which is determined by the operator when creating the dataset, or by the data input system (e.g. the spatial resolution of a satellite scanner). Each cell will contain a single value: this could represent surface reflectance (e.g. of a given Landsat band); class data (1 = vegetation, 2 = urban, etc); the average elevation of the cell (forming a digital elevation model); or any other feature of the Earth’s surface within that cell. Multiple layers can be used to represent different features, classes, concentrations, etc. Analysis can then take place using arithmetic and mathematical manipulation of the numbers in the various layers and cells.

3.2.2 Vector data

The vector model represents features of the Earth’s surface using points, lines and polygons. Points are features that can be represented by a single co-ordinate pair. Points have no length or area, position being their only parameter. Lines can represent linear features such as roads and footpaths that can be represented by a series of connected points. Lines have position and length, but no area. Polygons are features that can be represented by closed lines where the line start and end points are at the same position. Information on the feature is associated with the enclosed area, polygons having position, length (perimeter) and area parameters. Most data types can be represented by a single category, although there may be exceptions to this: elevation data can sometimes consist of contours (lines) and spot heights (points), requiring a combination to fully represent the surface. As with the raster system, multiple layers can be used to build up a map of the Earth’s surface and the specific features of interest, furthermore, analysis between the layers can be undertaken. Each created feature can have a significant amount of information attached to it and will have an entry in a data table, called an attribute table. This approach to storing data makes analysis a complex, but powerful, process and also allows easy data retrieval for individual or multiple features.



1	1	1	1	1	1	1
1	1	1	1	2	2	2
2	2	1	1	2	4	2
1	1	2	3	2	1	1
1	1	2	3	2	1	1
1	1	2	2	2	1	5
1	1	1	1	1	1	5

Figure 3-2 Vector and raster data models. Figure (a) is a subset of an aerial photograph showing river meanders. Figure (b) illustrates how this area could be represented as a vector layer. Figures (c) and (d) illustrate how this area could be represented as a raster layer.

3.2.3 Hybrid systems

The hybrid model combines both raster and vector models and allows relatively easy transfer of data, and even combined analysis, in the two formats. The advantage with using a hybrid model is that all potentially useful data can be utilised easily, regardless of the model in which it was encoded. For larger projects with ambitious goals this is invaluable. The disadvantage for the smaller project is the cost of the software and the complexity of the system. Hybrid systems are usually at the higher end of the cost scale for GIS software and are inherently more complex than single model systems.

Once the data types and their respective models have been ascertained, data input can be undertaken. Most current GIS software can import data from various existing sources, and all will have features allowing data to be manually entered. Datasets in a GIS are configured in layers, usually containing a single feature-type, perhaps divided into a number of classes. The feature-type can be broad or narrow, depending on its intended use. A land use layer, for example, may typically contain thirty classes, e.g. from a suburban class to an upland moorland class; whereas a water layer may only contain two classes: ‘water’ and ‘not water’. Speed, accuracy and intended further processing will determine which features should be input into which layers. Deciding which layers are necessary for a project, deciding how they should be input (scale, detail, etc.) and deciding what level of accuracy is necessary, are all factors that should be determined *before* starting a project. Once the project is underway, those parameters may be difficult - if not impossible - to modify. Some idea of how the data can be analysed is therefore necessary: *what can a GIS do to get the most from your ‘raw’ dataset?*

3.3 How can GIS help?

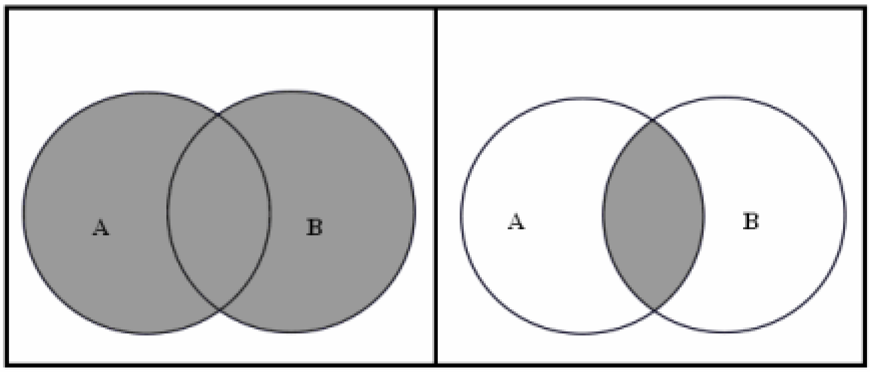
Geographical Information Systems are designed to answer questions about *where* something is, and *what* that feature is (its shape, size, name, etc). Frequently-used simple GIS queries include: "*Where is / Where are…*", "*How long / How big…*". Questions such as "Where is the scrubland on the aerial photo in Figure 3-2?" would fall in this category. The answer is that any pixel in category 4 (raster) or coloured dark grey (vector) highlights where the scrubland is. More complex queries would be:

"*What patterns exist…*" and "*Where have changes occurred…*". The ability of a GIS to rapidly sift through numerous datasets, select relevant information, merge datasets, carry out required analyses, and produce colour summary maps, has been a boon to decision makers: they can much more effectively consider a range of "*What if…*" scenarios.

A GIS allows users to highlight answers to questions about the location and distribution of a particular feature or class. Information about the feature/class can be extracted by the GIS, provided each data-layer shares the same geographical referencing system. This 'geo-referencing' is usually undertaken at the data input stage (see Chapter 2 for more details). As the co-ordinates are known, it is easy to derive statistical information about a class from the data layer through simple arithmetic. Thus when we get the answer to our question '*Where is the scrubland?*' we can also answer the question '*How much scrubland is there?*'. With the vector model, trigonometry provides the answer. With the raster model, the pixel size can be determined, and simple arithmetic can provide the areas covered by the relevant pixels.

Inter-relationships between layers can be determined within a GIS by breaking down each question into its component parts. Each factor can then be assessed individually and the results combined to produce an answer. A complex question that a suitably constructed GIS could answer may be '*Where is the optimum site to build a new airport?*' This can be broken down into specific criteria (or contributing factors), each of which can be addressed by manipulation and querying of a single layer. One criterion which might have to be met would be '*The airport should be more than five kilometres from the nearest town*'. By determining urban areas from a land use map, then creating a new layer where all urban perimeters are 'grown' by five kilometres, the new layer will contain areas that meet the specific criteria. The extraction of urban areas is very straight-forward, simply a case of deriving a new layer where all non-urban areas are assigned to a 'non-urban' class and all urban areas are retained. The 'distance from' operation is usually called a 'buffering' operation, where the operator assigns a new value to anything within a specific distance from a feature or features. The data-layer that results from this manipulation therefore only contains areas that do, or do not, meet the criterion.

Once a complex question has been broken down into specific criteria and data-layers have been produced which define those criteria, simple mathematics can be used to produce a final data-layer which answers the original question. This approach to combining data-layers, based on 'Boolean logic', is central to GIS analysis, and is shown schematically in Figure 3-3; further details are given in Chapter 7. Burrough and MacDonnell (1998), DeMers (2000) and Heywood *et al.* (1998) provide readable, yet comprehensive, descriptions of more complex GIS analysis techniques.

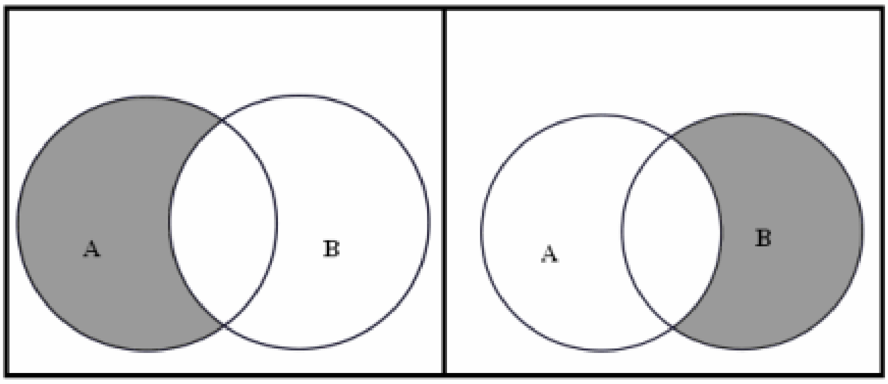


A OR B
Spatial area covered by either A or B is shaded

A AND B
Spatial area covered by both A and B is shaded

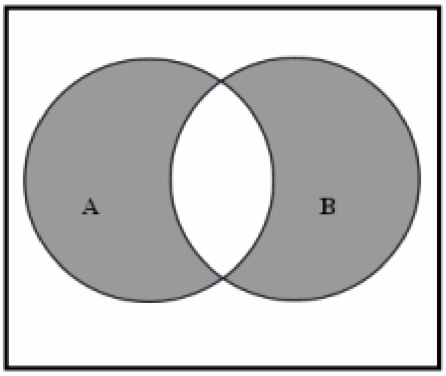
A NOT B

B NOT A



Spatial area covered by only A and not B is shaded

Spatial area covered by only B and not A is shaded



A XOR B
Exclusive OR: only A or B but not both

Figure 3-3 Logical operators in spatial queries and operations.

3.4 GIS and fieldwork

Empirical approaches to complex and dynamic systems necessitate the handling of very large sets of data. Multivariate analyses are often used with GIS, either to pre-treat data before GIS input or to process data after GIS analysis. Comparison of remotely sensed data from two or more points in time, using archive aerial photography or multispectral imaging, allows the quantification of dynamic processes, such as floodplain erosion (Winterbottom & Gilvaer 1997, Mount *et al.* 2003). Jorge & Garcia (1997) used GIS and Landsat-derived land cover maps to assess the landscape ecology of a floodplain forest-savanna region. The degree of habitat forest fragmentation was derived from (1) the mean patch area and perimeter, (2) the patch number and density, (3) the perimeter-area ratio, fractal dimension (D), and shape diversity index (SI), and (4) the distance between patches and dispersion index (R). Other types of application are listed in Table 3.1.

Table 3-1 Possible uses of GIS for expeditions and fieldwork.

- Expedition planning: site selection, sampling design, navigation
- Simple maps: data collection sites, roads, contours,
- Data recording and verification:
- Thematic maps: vegetation, animal distribution, soils, land use
- Map layers: overlay and compare different data sets
- Spatial operations: intersecting/overlaying layers, distance buffering
- Spatial analysis: correlation, interpolation
- Modelling: using some spatial data sets to predict another e.g. plant distribution
- Monitoring: record, map and analyze spatial phenomena at time intervals: change analysis
- Training: team members from all countries learning GISci field and lab techniques
- Education: data collection and maps provide a good way to communicate and involve people
- Results: maps are an effective means of presenting results

Apart from the increasing presence on the Internet, the next step in the evolution of GIS is to merge 3-D GIS applications with map and remote sensing archives, allowing the visualisation of changes over time. These techniques offer opportunities for animations, showing complex environmental processes, such as landscape dynamics or nutrient flows, varying in space and time (Reichhardt 1996, Mount *et al.* 2003).

Perhaps the most challenging advance lies in GIS-based modelling, offering a means to test hypotheses, through experimental designs, at different spatial and temporal scales. Models allow scientists to study the response of simplified systems (reduced to a set of variables deemed to be most important) to a restricted set of variations. GIS provides effective tools to run spatially explicit models and offers opportunities to extend models with cartographic input and output (Burrough & McDonnell 1998, Wadsworth & Treeweek 1999). GIS-based models have been used for fundamental research (Carpenter *et al.* 1999, Ormerod & Watkinson 2000, Nemani & Running 1996). Several GIS-based models for mapping the Habitat Suitability Index (HSI) of various species have recently been developed. GIS databases provide a source of habitat information for developing spatial HSI models. By extending the use of HSI models from simple habitat mapping to scenario testing, GIS can assist in the prediction of the outcomes of alternative resource-use strategies in ecosystem studies (Kliskey *et al.* 1999, Jochem *et al*. 2002, Leuven *et al*. 2000).

Table 3-2 gives an example of how several geographical datasets have been used in a GIS to predict the likely distributions of large mammals in a Tanzanian Game Reserve. Data were assembled from several sources: conducting regular surveys during an expedition, digitizing from existing maps, and making use of satellite imagery. The results of this study provided information for Mkomazi's management plan; expeditions using GIS can have a useful role in supporting environmental decision-makers and policy-makers.

Table 3-2 Geographical data inputs, analyses and outputs used in a study in Mkomazi Game Reserve, Tanzania. The aim was to model the distribution of large mammals in relation to their observed habitat preferences.

data requirements ↓	***mammal distribution in sample areas***	***habitat preferences, indicated by environmental variables across the entire study area***			
		seasonal changes in vegetation cover	vegetation categories	proximity to open water	terrain steepness
date source ↓	regular field surveys using GPS over 1 year	satellite imagery: AVHRR, 1km resolution, 1- day composites	satellite imagery: Landsat, 30m resolution	waterhole locations, from GPS survey	digitized 1:50,000 contour map
data processing ↓	create raster grid showing presence / absence of each species	temporal Fourier processing to give indices of seasonality	classification of images to give categories	calculate distance to closest water hole for each cell in a taster grid	convert to raster elevation model and calculate slope angles
analysis	mathematical model relating locations of mammal observations to environmental variables (using logistic regression)				
output	probability of each species' occurrence in each 1km grid cell in the study area -> map of the predicted distribution of each species (elephant, giraffe, impala, etc.)				

3.5 Limitations of GIS

Table 3-3 summarises common drawbacks when attempting to use GIS. With the multiplicity of GIS software, problems frequently arise from format compatibility when importing data. For GIS databases to be comparable or compatible, a preliminary standardisation of the data (e.g. time and spatial scale, variable definition) must be operated. A GIS is useless without good data ('GIGO': garbage in, garbage out). With reliable data, there is still a risk that inappropriate data processing will produce unreliable results, even though the GIS-generated output might look very impressive. Moreover, there is always the danger of over-interpreting GIS results. Users need to know the assumptions that went into remotely sensed data and creating a map or GIS-based model, including the scale and quality of the original data (i.e. meta-data records) (Reichhardt 1996). Understanding the errors and uncertainties in the input and output data are central concerns. Table 3-4 gives some general advice to anyone considering a GIS-based approach.

Table 3-3 Common drawbacks when attempting to use GIS (Wadsworth & Treeweek 1999).

- Lack of digital data
- Lack of time for data collection and entry
- Lack of experience and familiarity with software
- False precision (obscuring sources of error)
- A technology-led approach
- Over-investment in data irrelevant to research goals or decision-making

Table 3-4 General advice to anyone considering a GIS-based approach (Johnston 1998).

- Keep it simple
- Ask whether a GIS is necessary to tackle key questions
- Use existing data where possible;
- Plan ahead or conceptualise (for example, use data management systems or flow-charts to guide GIS development)
- Keep good records particularly source data and analyses performed at each step in the GIS process
- Always check results to see if the GIS output is logical
- Consult with experienced GIS users for advice on database management, data needs and procedures

At a more fundamental level, several objections have been made to the approach of a ‘sliced environment’ introduced by remote sensing and GIS tools. According to Muller *et al.* (1993), there is no evidence that each discipline can be reduced to a set of layers of spatial information. For most botanists, for example, the quality and hence the value, of a vegetation map is highly dependant on the selected classification system. A vegetation layer in GIS can never show the vegetation in all its aspects.

Over recent decades scientists have battled with the problem of how to present spatial information within the limits of physiology and the psychology of perception (Wadsworth & Treweek 1999). Tufte (1983) described five principles of ‘graphical excellence’:

- Well-designed presentation of interesting data (substance, statistics and design)
- Complex ideas communicated with clarity, precision and efficiency
- Giving the viewer the greatest number of ideas in the shortest time
- Data is nearly always multivariate
- Telling the truth about the data

While there are always going to be limitations as to what can be done with *GISci* technologies, there have been tremendous advances in recent years, with the quality of data going up, as the costs of data and data processing have been falling. Hatton (2003) produced a review of development and possible future applications of geoscience information systems, which included some quotes that highlight how rapidly information technologies have developed. Pause for a moment and reflect on how difficult it would be to carry out fieldwork using *GISci* techniques, if the following expert predictions had come true….

"*Computers in the future may weigh no more than 1.5 tons*." Popular Mechanics, 1949.

"*I have travelled the length and breadth of this country and talked with the best people, and I can assure you that data processing is a fad that won't last out the year*." Senior editor for Prentice Hall publishers, 1957.

"*640K ought to be enough for anybody*." Bill Gates, 1981.

3.6 GIS data types and sources

This section provides an overview of the sorts of data that might be of use to fieldwork projects where GIS usage is planned. Aspects of remote sensing, photogrammetry and GPS data are briefly covered here, but are dealt with in detail in later chapters. The contact details for data suppliers have been kept to minimal levels, as full details are given in the Appendix and as web links on the Manual's CD; further information can also be readily obtained, in most cases, via Internet searches. Allow adequate time when ordering data: digital data such as satellite images can now be delivered within days, thanks to advances in data compression technology and Internet FTP (File Transfer Protocol) links; but paper maps, and prints of aerial photographs in particular, can take many weeks to arrive.

3.6.1 Types of data for fieldwork

Analogue data are all forms of paper record or 'hard copy': maps, sketches, photographs (graphical representations), plus tables and statistics (quantitative summaries). Graphical representations can be transferred onto a GIS either by digitising, using a vector system; or by scanning, using a raster system. Quantitative analogue data has to be typed into a database or spreadsheet, such as Access or Excel, before it can be loaded onto a GIS. *Digital data* can be tabular (database/spreadsheet) or graphical (digitised or scanned) and are in a computer-coded format, mostly stored on CD-ROMs, ZIP disks or DVDs, all of which can hold 100s of megabytes of data. Satellite remote sensing images, which are of particular value to expeditions looking at poorly-mapped regions, require larger-volume data storage systems. For instance, a Landsat TM scene covering c. 180 km x 180 km in seven parts of the electro-magnetic spectrum takes up c. 260 megabytes.

3.6.2 Analogue maps

Published maps for topography, geology, vegetation cover and soils will probably exist at national and regional levels.....*whether they are still in stock in the host country is another thing.* Try to purchase regional and local maps from a UK map supplier, such as Stanfords or GeoPubs, before going. *Tactical Pilotage Charts* (TPCs) show elevation, topography and basic vegetation cover, at 1:500,000 scale, and are available for most parts of the world. Regional surveys of soils and vegetation cover have been carried out for many countries by the Food & Agriculture Organisation of the United Nations (FAO, Rome). You may be able to view maps in the map collection of the RGS-IBG, or review the coverage of a region using the CARTO-NET database of the British Library. A worldwide directory of national earth-science agencies and related international organisations is published by the US Geological Survey. It gives details of overseas agencies that might hold maps or sets of airphotos. It may also be worthwhile to enquire at the National Cartographic Information Service (NCIS) of the US Geological Survey, or the Cartographic Section of the US National Archive. More detailed maps, at 1:50,000 or 1:25,000, often only cover a small percentage of developing countries and may date from colonial times. If you can obtain detailed maps, think yourself lucky, but bear in mind that the map data may be over 20 years old and accuracy may not be as high as UK Ordnance Survey maps.

3.6.3 Analogue data sources

Maps, reports and data on features in your study area (such as climate, soils and vegetation cover) should be available at the national, and probably regional, level in major encyclopaedias, atlases and textbooks. Useful information for overseas projects might be obtained from the following sources:

- For former UK colonies: the Natural Resources Institute at Greenwich University, the British Geological Survey, the Ordnance Survey, the Natural History Museum, Royal Geographical Society (with IBG) Foyle Reading Room.
- French overseas dependencies: Institut Geographique National (IGN), Bureau de Recherche Geologique et Miniere (BRGM), the Centre Nationale de Recherche Scientifique (CNRS) or the Institut de Recherche pour le Development (IRD, formerly known as ORSTOM).
- Libraries of learned societies (e.g. the RGS-IBG, the Geological Society) or research institutes, such as the Scott Polar Research Institute or the Oxford Forestry Institute.
- University libraries, both in the UK and in host countries, which may have archives of reports and maps.
- The British Council or Association of Commonwealth Universities, which may have records of existing links between UK Universities and the expedition's host country.
- UK aid agencies (e.g. OXFAM, VSO, ITDG, WaterAid).
- The US National Archive and the Smithsonian Institute hold geographical data from around the world that may be of use.
- Relevant ministries/agencies in the host country (agriculture, forestry, land survey).
- Commercial map suppliers, notably Stanfords and Geopubs in the UK, or Omni Resources in the USA.

If your fieldwork is overseas, you should be aiming to carry out your fieldwork jointly with a team from the host country: this will facilitate access to data. Furthermore, this facilitates the transfer of skills to places where they are most needed and helps to reduce bureaucratic and logistical problems.

3.6.4 Sources of digital data

More and more data summarising the features of the world are being digitised and are available over the Internet, from word-processed documents and spreadsheet tables, to scanned airphotos and satellite images of entire continents. Digital map data can be obtained in three ways: inputting your own data, purchasing ‘off-the-shelf’, and downloading from the Internet.

(a) *Inputting your own data* can be time consuming and tedious, but analogue data is more useful in a GIS-usable format. Data can be input as follows:

- Typed-in tables, spreadsheets or databases, with an object's grid reference followed by its attribute data.
- With a vector-based system data can be digitised in as points, lines and areas, each with an identifying code that can be linked to a database management system (DBMS). The required digitising board and stylus will cost at least US$300.
- Graphical data can be scanned in, saving a lot of time, and provides an excellent means of comparing or merging different maps or airphotos of the study area. However, the scanned image is only a picture made up of cells with varying values, not an inter-related network of points lines and areas: vector overlays of points, lines and area boundaries may have to be created by on-screen digitising using your pc mouse.

(b*) Surfing the web* A vast array of information, including digital maps and images can now be found on the Internet: a summary of useful Internet websites is given in the Appendix. At the expedition planning stage, the sites of the RGS-IBG (*www.rgs.org/mapping*), the US National Geographic Society and the US Geological Survey (USGS) are particularly useful.

(c) *Digital maps* are now easily obtainable at broad continental and national levels, usually on CDs or increasingly for download from websites. These digital maps of developing countries are unlikely to be more detailed than 1:1,000,000, and are therefore only of use at the expedition's preliminary planning stage or where regional studies are important. A summary of some digital map coverage that might fit into an expedition budget is given in Table 3-5.

Table 3-5 Some examples of digital map data sources.

package	datasets	regions	scale	format	approx. cost
Bartholomew Digital Maps	Similar to DCW (*licensed to some universities - check for availability)*	Global Europe	1:1M 1:250K	CD or Internet	£360/year
Digital Chart of the World (DCW)	Altitude, coasts, rivers, vegetation, borders, roads/rail, populated places. Data source: Tactical Pilotage Charts	Global	1:1M	CD / Internet	£245 / FREE
Eurographics	Various, e.g. soils, vegetation, geology, topography, relief.	European nations	various	CD / Internet	enquire for details
Global Land Cover Facility (GLCF),	Online access to archive Landsat data hosted by University of Maryland	Global (ex. Antarctica)	various	CD / Internet	FREE via FTP, c. £30 per CD
GRASS (datasets for this popular GIS software)	Soils, altitude, aspect, vegetation, threatened species, marine biology	Global	1:1M	CD / Internet	FREE
MacArthur Project Cambridge University	Vegetation, land use, soils, admin boundaries, elevation, climate, people	Inner Asia			enquire for details
National Geographic	Various, notably key biodiversity sites, national parks, relief, politics	Global	various	Internet	enquire for details
ODDENS, University of Utrecht, NL	Extensive online access to atlases, catalogues and maps	Global	various	CD / Internet	enquire for details
Shuttle Radar Topography Mission (SRTM)	Digital elevation model (DEM), good horizontal and vertical accuracy	60ºN to 60ºS	90 m cell size	Internet	FREE
United Nations (UNEP / FAO)	Global Resource Information Database (GRID) *NB many GRID datasets (e.g. FAO Africa soils map) are included with IDRISI GIS software*	Many: environment data sets	1:1M, others	CD or Internet	£15 each
US Geological Survey (USGS)	Topography, relief, geology, aerial photography – mostly US coverage Satellite imagery – worldwide cover.	Mostly USA and the Americas	various	CD / Internet	enquire for details
World Conservation Monitoring Centre	Digital atlases of tropical forests, coral reefs, wetlands, protected areas	Global	various	CD / Internet	enquire for details
Vector Map of the World	Supersedes DCW	Global	1:1 M	Internet	Free

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section B: Data

Chapter 4: GIS Database Mechanics

4 GIS Database Mechanics

4.1 The underlying information in the GIS

A GIS is simply a database with graphical capabilities that displays its information within a given co-ordinate space. As with any database due attention has to be given to the way GIS data are formatted and stored. This configuration is referred to as the database structure or schema. Designing effective database structures can be very difficult but some of the more common concepts will be discussed here. A more in-depth look at database structure pertaining to GIS can be found in Burrough & McDonnell (1998) or for more on database structure itself the reader is referred to Roman (1999). The GIS can display co-ordinate information relative to other data or to digital images or maps. The information displayed, often referred to as a theme (event theme in ArcView), consists of data taken from the database (these data points are referred to as features) and each point has many details or attributes. Figure 4-1 shows a project consisting of the countries of Europe. The collection of countries is a theme, as would be a collection of cities, rivers or other geographic features. Each theme can have various details such as a country's size, population, currency etc. These descriptions about the theme are called attributes.

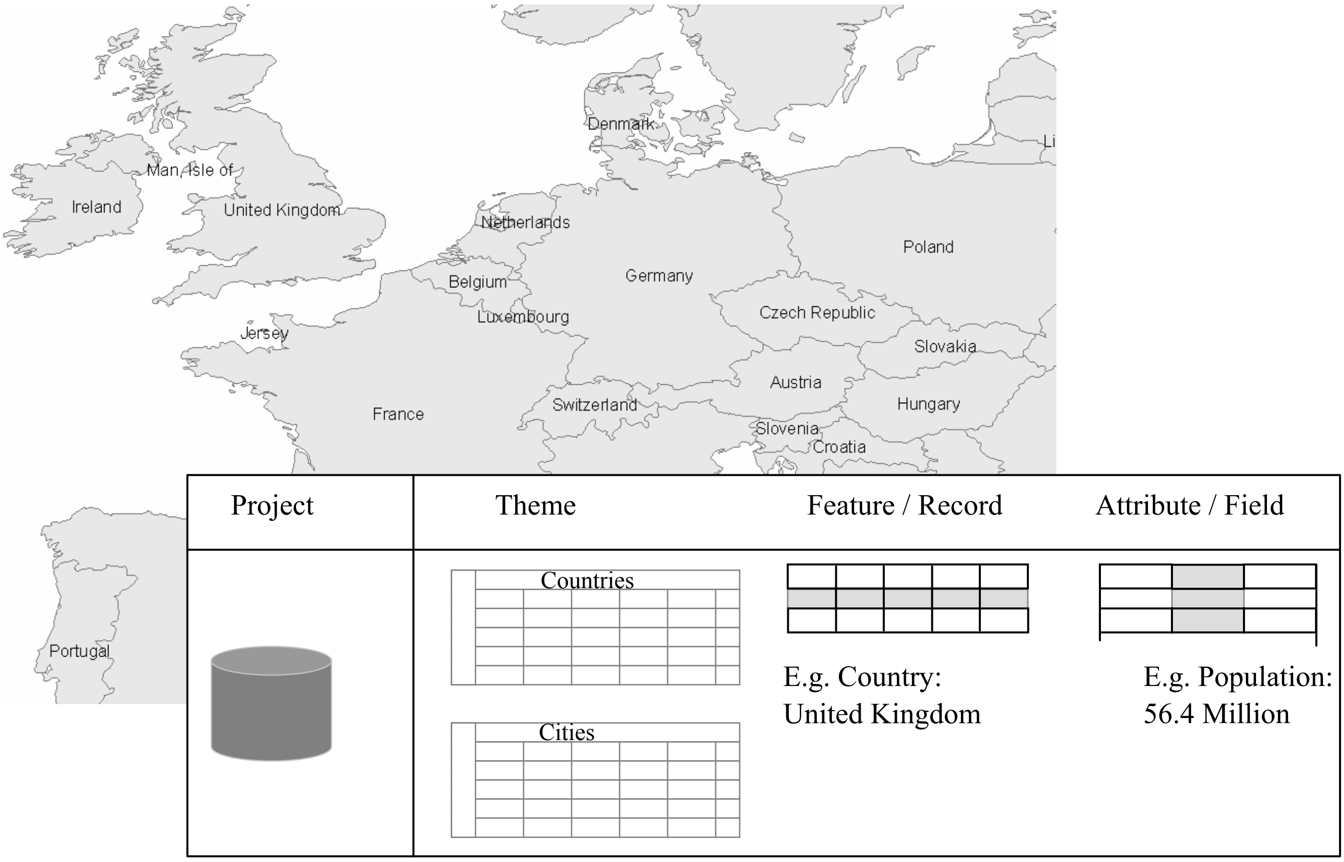
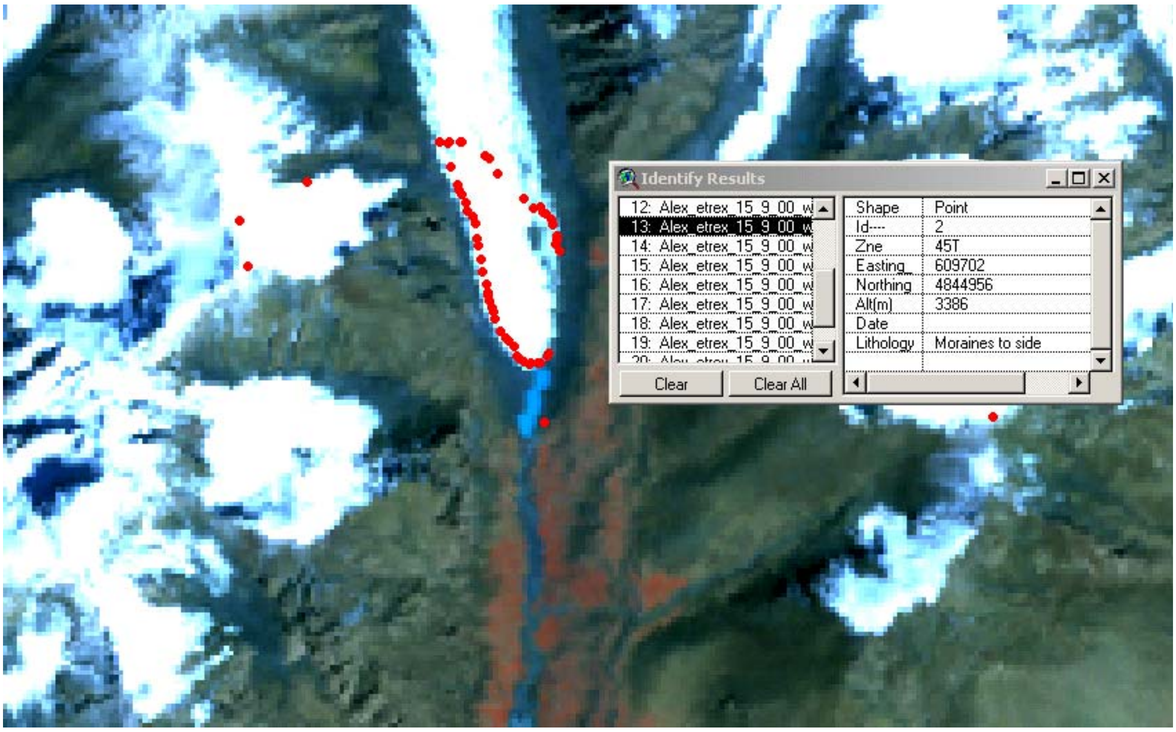


Figure 4-1 Relationship between themes and attributes. Data from ESRI ArcView example data.

These attributes will often come from an external database, which could be as simple as a Microsoft Excel spreadsheet containing GPS positional data, or it could be a large ORACLE database with hundreds of tables and gigabytes of data. A GIS can access these data directly or the data can be exported into the GIS in a tabular format. What the database is and how the GIS accesses it is vitally important to any project. All themes require attributes for latitude and longitude (or equivalent) so the GIS can plot them as X Y co-ordinates. In Figure 4-2 GPS points from the Bogda Shan dataset are plotted above a Landsat ETM+ sub-scene. In the dialogue box various attributes associated with the localities theme.



12: Alex_etrex_15_9_00_w
13: Alex_etrex_15_9_00_w
14: Alex_etrex_15_9_00_w
15: Alex_etrex_15_9_00_w
16: Alex_etrex_15_9_00_w
17: Alex_etrex_15_9_00_w
18: Alex_etrex_15_9_00_w
19: Alex_etrex_15_9_00_w

Shape	Point
Id---	2
Zne	45T
Easting	609702
Northing	4844956
Alt(m)	3386
Date	
Lithology	Moraines to side

Figure 4-2 Figure showing Bogda Shan data and GPS co-ordinates with associated data.

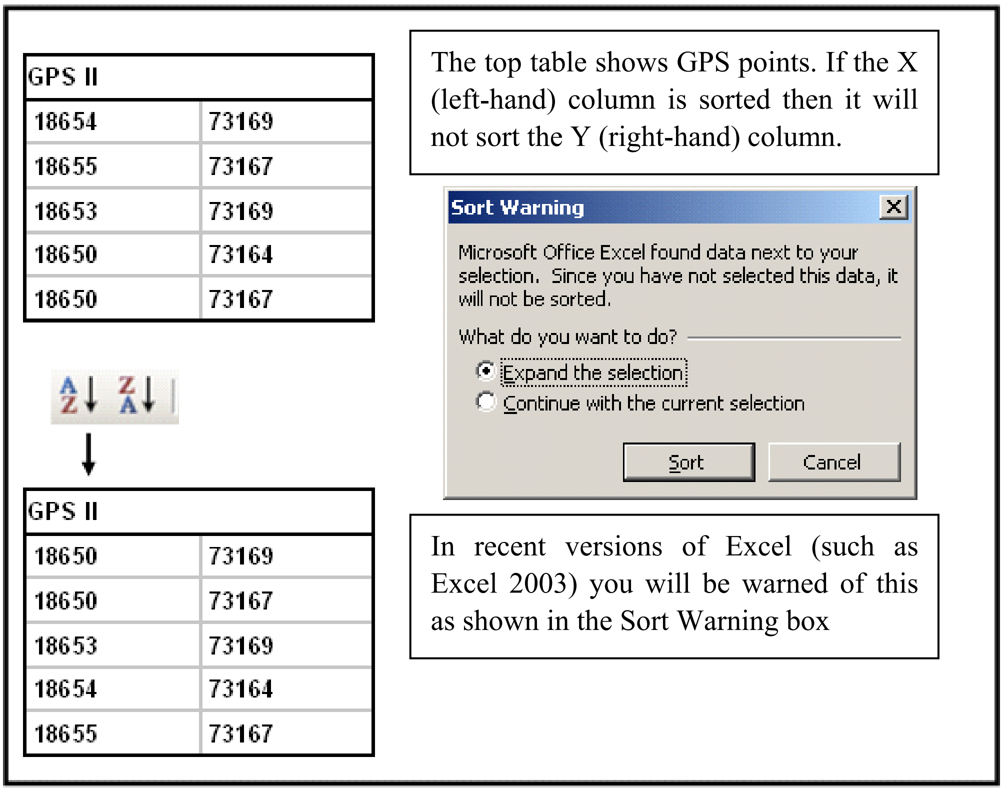
The number and complexity of the attributes associated with the theme will influence how the team decides to store them.

4.2 Selecting the project database

The first decision to be addressed is the type of database to be used. There are two types of databases: file/server and server based solutions. *File/server databases* have a database file on a computer hard drive. This could be held locally on the expedition PC or on a shared network drive of a PC. Examples of this type of database include Microsoft Access *.mdb and *.mde files, Microsoft Fox Pro *.dbf or Borland Dbase *.dbf files. Microsoft Excel spreadsheets also take this format but, as discussed below, are not the same as databases. The other type of database is a *server database* where a computer is used as a dedicated database server that processes instructions and sends them back to a PC over a network. Examples of this type include ORACLE databases and Microsoft SQL Server. Server based solutions are more expensive, generally require more hardware and require a network connection. They are useful for high-end databases where large quantities of data need to be accessed by large numbers of people (>50 users). Server side solutions also benefit from automatic data backup and file security. File/server solutions are cheaper and generally better suited to small databases or where data access across network is unavailable. It is possible that the main database could be a server solution but that certain tables are output as Dbase files onto a file/server connection for accessing by the GIS. This, however, runs the risk of losing data integrity within the GIS and underlying database as the two may rapidly become very different and require synchronising on a regular basis.

The second decision with databases is whether the system is to be relational or not. This question addresses the database engine, the system that handles all queries. The engines can either be a Database Management System (DBMS) or a Relational Database Management System (RDBMS). Relational databases automatically connect data in various tables and keep records in the correct order when sorting or querying. Relational databases take longer to set up and are often considered less flexible than non-relational databases but benefit from improved data integrity and substantial speed improvements.

In a non-relational database such as Microsoft Excel data entered has no connection with any other information. This means if a column is sorted it will destroy the integrity of your information. This is shown below in Figure 4-3. The left hand column is sorted but the right hand column stays the same. This means when plotted the data will no longer make sense. If this was in a relational database then both columns would automatically be sorted. The dangers of having your data accidentally scrambled should demonstrate the need for a relational system.



GPS II	
18654	73169
18655	73167
18653	73169
18650	73164
18650	73167

GPS II	
18650	73169
18650	73167
18653	73169
18654	73164
18655	73167

The top table shows GPS points. If the X (left-hand) column is sorted then it will not sort the Y (right-hand) column.

Microsoft Office Excel found data next to your selection. Since you have not selected this data, it will not be sorted.

What do you want to do?

● Expand the selection

○ Continue with the current selection

In recent versions of Excel (such as Excel 2003) you will be warned of this as shown in the Sort Warning box

Figure 4-3 Disadvantages of non-relational systems.

What this means in practise is that many decisions have to be made in advance. The structure of a relational database is not very flexible but the way data is accessed is. A good example is that once data is entered into a non-relational database it can be very slow and time consuming to alter. If a relational database is set up correctly altering information on the fly is simple, rapid and accurate. For example, on the first few days in the field assumptions may be made in the sampling and recording of data that later turn out to be incorrect. A rock type might be misidentified or a bias might be introduced into some other study. A relational database will allow one change to be made to the data that cascades through all relevant features instantly without searching through and potentially correcting the wrong data. For these reasons it is a very good idea to store your data relationally if possible.

4.3 Database configuration

When the type of database has been chosen it must be correctly configured. This process defines the database structure. The database structure or schema describes the number of tables, fields and the relationship of each one to another. The database structure has the biggest impact on the speed, flexibility and accuracy of the information. Databases are efficient systems for storing and retrieving data but each record has to be recorded in a set manner. Each individual field must be given a format that the database expects such as a number, string or date. When this format has been selected no record for that field can deviate from this selection. Careful consideration must therefore be given to the type of data required. Standard databases store data in one of six ways: integers, double numbers, strings, binaries, dates or memos. These data types are described below in Table 4-1.

Table 4-1 Data storage types for typical databases.

Data Types	**Restrictions**	**Uses**	**Speed / Flexibility**
Integers	2^{16} possible values. **Unsigned** approx ± 32,768. **Signed** + 32,768 to -32,767 inc 0). No decimals.	Used for integer attributes and to create internal database references.	Require small amounts of memory to store (2 bytes). A computer can sort & process integers quickly
Doubles	Allows up to 15 decimal places and up to 324 sig. figures.	A double number is the longest number a database can store	Requires 8 bytes of storage space and are larger and slow to sort.
Strings	A line of text containing up to 255 characters.	Descriptive fields	String fields can't be used for arithmetic processes.
Binaries	Binary (or Boolean) consists of a 1 or 0.	These fields are very quick to process & store.	Access uses Y or N boxes. ArcView uses True or False.
Dates	Commonly converted to a Julian Date format. Each day is given a number and the time is given as a fraction of this day.	Dates and times.	Very slow to query.
Memos (Blobs)	Memos are large files that can be stored externally and can be of any size.	Used for supporting notes and data.	Very slow to process in a database and should be used carefully

Some forms of data are easier to store and faster to process than other forms of data. Integer numbers are very useful to use, as they are both small and fast. In many cases this will not be acceptable because the information recorded will be of a more complex nature. Good database design allows the use of integers even when descriptions are more complex. An example that can demonstrate this is habitat mapping. Habitat mapping is an important fieldwork exercise being conducted by many expeditions. The example in Section 4.4 will show the difference between a good database structure for animal sightings and a poor structure.

4.4 Example of database structure designs

Habitat mapping is an important fieldwork exercise being conducted by increasing numbers of expeditions. This example will show the difference between a good database structure for animal sightings and a poor structure. While recording animal sightings in a game reserve using a GPS, commonly a description of an animal will be recorded with its corresponding GPS location as shown in Table 4-2.

Table 4-2 An example fieldslip in digital format.

species	habitat	waypoint	x co-ordinate	y co-ordinate	weather	altitude
impala	grassland	12	55234	233444	clear	500
buffalo	grassland	12	55234	233444	clear	500
cheetah	grassland	14	55520	232500	cloudy	480
impala	grasland	15	55420	232418	clear	450
buffalo	grassland	15	55420	232418	clear	450
giraffe	grassland	16	55435	232480	clear	411
buffalo	grassland	17	55412	235212	cloudy	422
cheetah	grassland	17	55412	235212	cloudy	422
impala	forest	18	55214	235444	cloudy	455
impala	forest	13	55280	233501	raining	200

Table 4-2 is poorly laid out in database terms: there have been several animal sightings at any one location but the information for that location is replicated. For example, the grid reference and altitude are all identical for any one GPS location so it is inefficient to write them multiple times, it is time consuming in the field and there is a significant potential for error. Putting one number wrong could seriously compromise a team’s work. Also text fields are slow to sort; therefore we should find a method for replacing some text with integer numbers. We can do this by creating a table for all the variables.

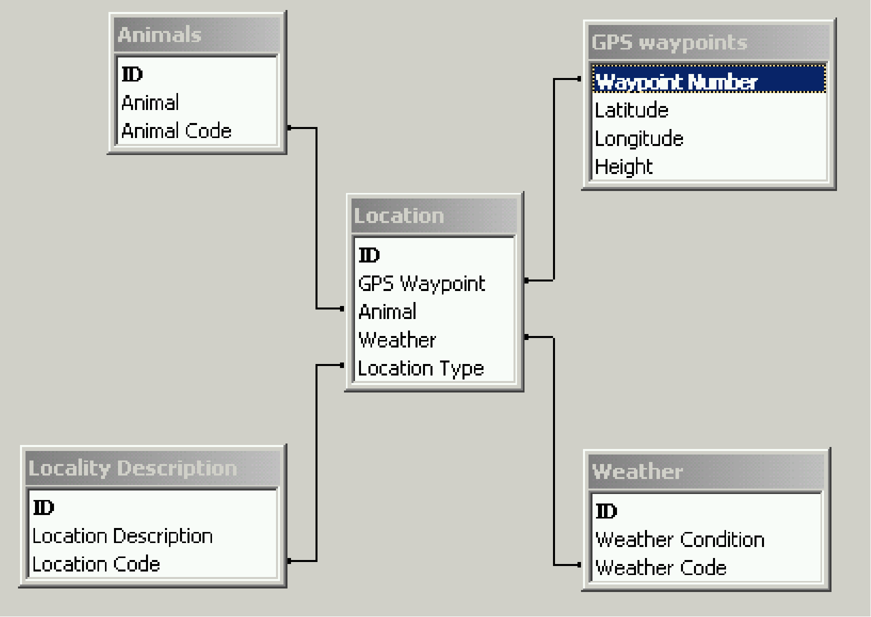


Figure 4-4 Relationship of data within the habitat mapping GIS.

In Figure 4-4 the GPS waypoints have been stored separately. This way data is not replicated. The first table contains 462 characters most of which are strings of text. By relating it into a database we can simplify it to less than 100 characters all of which are integer numbers as shown below in Table 4-3. This means that as the database grows in size its speed compared to the first method will grow exponentially.

Table 4-3 The database equivalent of the fieldslip. It is no longer necessary to record the latitude and longitude of each waypoint individually. By saying waypoint 12, the GIS can look up the co-ordinates for waypoint 12 from a separate table. Similarly for animal species, instead of recording the name each time, we can list all the animals in a separate table and reference them by a number.

id	waypoint	animal	weather	habitat
1	12	1	2	2
2	12	3	2	2
3	14	2	1	2
4	15	1	2	2
5	15	3	2	2
6	16	4	2	2
7	17	3	1	2
8	17	2	1	2
9	18	1	1	1
10	13	1	2	1

The species' numbers etc. can then be looked up in a database to create the same result as seen in the first table. The resultant query would look like Table 4-4, which is identical to Table 4-2.

Table 4-4 The result of a query on the re-structured database.

species	habitat	waypoint	x	y	weather	height
impala	Grassland	12	55234	233444	clear	500
buffalo	Grassland	12	55234	233444	clear	500
cheetah	Grassland	14	55520	232500	cloudy	480
impala	Grassland	15	55420	232418	clear	450
buffalo	Grassland	15	55420	232418	clear	450
giraffe	Grassland	16	55435	232480	clear	411
buffalo	Grassland	17	55412	235212	cloudy	422
cheetah	Grassland	17	55412	235212	cloudy	422
impala	Forest	18	55214	235444	cloudy	455
impala	Forest	13	55280	233501	raining	200

As discussed above a relational system reduces accidental errors. All the numbers are entered from a computer form where the animal type is selected from a drop-down list showing the relevant species. Only the number is entered into the table reducing the possibility of accidental typos and errors. In Table 4-2, the Impala at location 15 has its habitat spelt incorrectly. This is likely to result in incorrect analysis of data. If the typo were in the GPS location field, then the problems may go undetected and result in a spurious interpretation when animal sightings are plotted in the GIS.

The speed of a database when accessing data or performing queries is related to the amount of data in any field and the type of data used. In the previous example the size of the fields was reduced to a fraction by using integer numbers. This reduction in size is one way to speed up the database. The second speed increase is gained from the fact that the numbers were easier to process than the previous text string. All these help to make the database more efficient. There is also a major speed increase in entering data because fields with repeated data do not need to be re-typed. This is among the most important factors when establishing where data will be kept in the database. The process of removing redundant data is called normalisation. Normalisation works by separating tables into groups of data that are dependant on each other (a relationship called Functional Dependency) (Roman 1999). In the worked example above, the original field slip contained many fields that were not dependant on one another. The animal spotted is unlikely to be dependent on date, or exact grid location. For example impala could be spotted at many grid locations on many days. Normalisation requires that these tables be separated. Similarly, if we had details about an animal such as its colour etc. this should also be excluded into a separate table. There are six levels of normalisation - these place the database into the first normalized form up to the third normalised form, followed by the Boyce-Codd normalised form and finally the fifth and sixth normalised form. Each form represents a more efficient table design than the previous method. The increase in performance slows with each step in normalisation form and anything above third is generally not relevant.

The most common way to increase speed (usually covered by second normalisation form) is to give every record a unique identification. This is commonly referred to as an ID or Primary Key (this is usually required even before normalisation can begin). When a database stores data it puts it in a sequential order according to the key. When an ordered list of sequential data is queried, the database examines the middle record first. If this is the correct record it stops, if not the database assess if the number it requires is greater than or less than the current value. The database then moves to the middle of the next section and repeats. In an unstructured data table the database starts from the beginning and reads through until the end, this method is very inefficient. The Bogda Shan Expedition discussed in Section 17.1, collected large amounts of field data. A GPS transect across a mountain range and an intermontane basin was constructed using GPS receivers. The expedition collected 12,855 waypoints and how that data was stored had a significant effect on the query speed. This is demonstrated in the Table 4-5. The reader is referred to Burrough and McDonnell (1998) for a more thorough explanation.

Table 4-5 Structured and unstructured queries.

	unstructured list	**structured list**
average number of searches to find one of *n* records in a database	$(n+1)/2$	$\log_2(n+1)$
average number of searches to find a record from the Bogda Shan Expedition database, containing 12,855 records	$(12,855 + 1) / 2$ $= 6,428$ searches	$\log_2(12,855+1)$ $= 13.65$ ≈ 14 searches

The advantage of a relational database to the expedition should now be obvious. It is worthy of note that for large projects a hybrid system between relational and GIS spatial data has been created. These spatial databases such as ARC SDE and ORACLE Spatial can be used to great effect but are generally too involved and expensive for an expedition. If the reader requires more information on the most effective method for storing and querying very large sets of complex spatial geographic data then they should consult a specific text book such as Rigaux *et al.* (2002). For most purposes a relational database combined with spatial queries within the database will suffice.

4.5 Using the database in the field

During fieldwork, the structure of the database will be critical to what is recorded. There is little point in configuring all fields as integers if this is not valid for the data being collected. On many occasions, decimal data will be essential for a project. In recent years the size of a database used on a desktop PC has become less of a concern, and hard drives are now commonly over 60 gigabytes in size. The size of the GIS on a high specification machine is not as relevant as in previous years. In the field, however, high specification PCs are not generally available and so smaller data loggers are important. These might include the memory of a GPS receiver or a handheld PDA or similar electronic device (see Chapter 13 on Field Equipment). These machines have comparatively low storage capabilities and make numerical annotation as important as ever. Complex processing in a GIS is not a significant concern on high-end computers but on low specification field computers, the processing has to be streamlined to work within the computer's capacity and often within the length of the battery time.

It is important that the setup of the database has been given sufficient thought. When a database has been configured it can be very difficult or even impossible to change the format of the data e.g. to change a string field to an integer field. In the habitat mapping example (Section 4.4), if a different species of animal were observed, such as seeing a lion, it would be easy to add, as this animal could become animal number 5 in the animal table. Changing the fundamental data structure is much more difficult than adding new records to a database. These decisions need to be made in advance of the expedition and adhered to in the field, changing the waypoint data from integer to decimal would be significantly more difficult.

Before the data can be imported into the database it needs to be collected in the field. Commonly data is collected in a field notebook and retyped into the database. This is a good method as it retains a hardcopy of the data in case the database becomes lost or corrupted. During the Bogda Shan Expedition data was written in notebooks and referenced to selected waypoints stored in the GPS. At base camp information from the notebooks was transferred to an Access database and linked to the GPS waypoints by joined fields. In Figure 4-5 the table on the right contains all the GPS data, the table on the left contains all the locality descriptions.

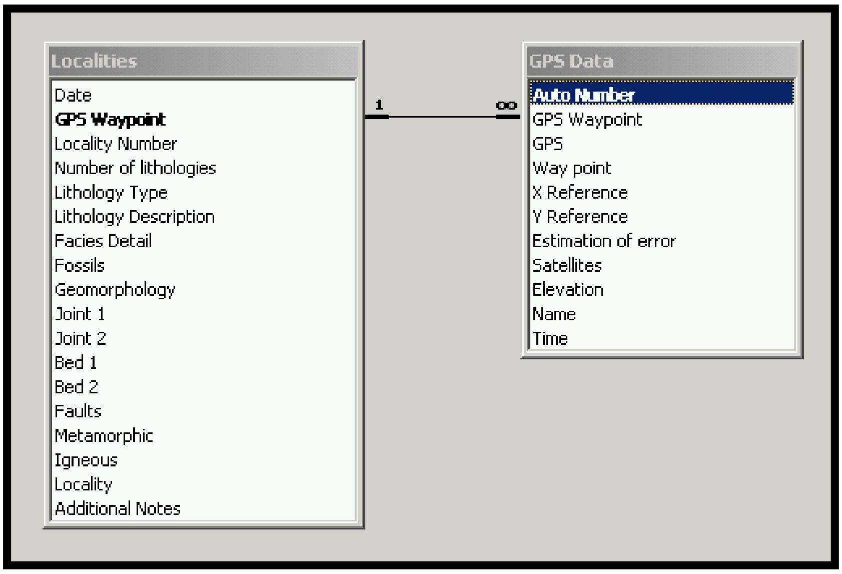


Figure 4-5 The relationship of GPS data to locality data in the Bogda Shan Access database. Note the relationship type one to many.

Figure 4-5 shows not just the relationship joining two numbering systems but the type of relationship. Joins can only exist between data of the same type, in this case two integer values. The data in the table on the right [GPS Data] has the symbol '∞' next to its join while the data in the [Localities] Table has the symbol '1' next to its join. This is referred to as a 'one to many' link. This means that there could be lots of GPS points but they can all relate to only one field locality. In the field four GPS receivers were commonly used. So at any one location four GPS waypoints could have been collected, though they all refer to the same location (*due to errors in the GPS receiver the latitude and longitude co-ordinates would be slightly different between them*). The GPS waypoints table could have up to four times as many records as the localities table. Also one locality description may describe a field outcrop that may be many meters in length. There could be multiple GPS waypoints along the length of the exposure. All of these describe just one location, so we need to tell the database that there are lots of readings at one point and we do this with the one-to-many join. Other joins include the one-to-one join where there are an equal number of records in each table or a many-to-many link where the tables could both contain different quantities of data. Many-to-many links are more difficult and in most databases require a linking table between the joined tables.

4.6 Integrating existing data

In many cases the GIS being constructed for an expedition will build on some previous work. There might be an existing database of previous field results that need to be compared or a list of waypoints that need to be visited. There may even be an existing GIS that needs incorporating into the current version. The use of historical data can be essential to a good field work project, however, getting this existing data into the GIS can be very problematic. The way each database stores data is different and they are not all compatible. Very few standards exist but the format that the data is accessed in has been given some uniformity.

The ANSI and ISO standards have supported the use of the Structured Query Language (SQL) for data retrieval, which is important for data export and integration. Unfortunately, over time several slightly different versions of SQL have appeared. For example, Microsoft SQL Server, Visual FoxPro and ORACLE all have their own versions. Microsoft has made significant progress in the field of database integration and there are many techniques for getting disparate data sources to work together. More recent versions of Microsoft products (Access 2000 and SQL Server 2000 upwards) all integrate transparently. It is very easy now to have a small front end Access database referencing a larger GIS database behind the scenes without the need for separating the data and running export and import queries.

The second, and more useful, method of data integration is the Open Database Connectivity (ODBC) interface. ODBC is an Application Programming Interface (API) developed by Microsoft in 1992. It allows disparate data sources to talk to one another through a common syntax. Microsoft software can use an ODBC connection to retrieve data from a host of different sources. Once the data has been read by the ODBC API it can be exported to a Dbase *.dbf table which is perfect for adding to a new ArcView GIS project. ArcView can also reference a variety of other sources such as *.mdb files directly via a link. Alternatively if the data can be exported to a text file (ASCII) it can be re-imported to the database.

The problem with exporting and importing data is that referential integrity (the links between the data) can be lost. Also the data types (see Section 4.2) will not be retained. To restore data types a parsing filter can be used. Microsoft Excel has a very good parsing filter called Text to Columns found under the Data menu. This can reconvert data back to its original format if it has been lost. For example an integer data field will be changed to a string data field when exported as text. Selecting that column in Microsoft Excel and running the Text to Columns parsing filter over the data will return the data to its original form. When the data is in an Excel or Dbase format it can readily be imported into the GIS.

The biggest concern with historical data is the geographical co-ordinates. The GIS is concerned with the spatial relationship of data. With existing data it is not sufficient to import an X and Y co-ordinate without some knowledge of the datum and projection of the data. These concepts are discussed more thoroughly in the chapter on co-ordinate systems. It is important to be aware of them as they can considerably alter the quality of the data. Specifying an incorrect datum or importing data to a different projection can cause considerable problems.

4.7 Data import/export and cleaning

As discussed above, a GIS often does not exist in isolation. Data will have to be brought into the project and exported out of the project for reports, graphs and data sharing with other projects. Most databases have proprietary data formats. Problems can be avoided by exporting data to a common format, readable by many systems. Unfortunately, there are no GIS data standards. ASCII text files are useful but they are only raw data and they have no method for describing the information. The procedure of defining the data types is left to the automatic parsing filter or to the GIS administrator. Neither of these are ideal situations. A better solution is to use a data format that understands the types of information encoded but is not proprietary to a single system. Traditionally Dbase files have been very useful for this purpose. Most database engines will import a *.dbf file and they are a good method if nothing else is available. Dbase files are, however, far from perfect and a better solution is to use a totally format free import and export method. A common solution is to use HTML (Hyper-Text Markup Language), a text format that can be read by almost any computer, regardless of operating system and hardware. HTML script is structured into a series of blocks that are described by tags which tell the computer how to format the text.

HTML data format

<html>	*tells the computer the data format*
<head>	*tells the computer to format the following text as a heading type (usually bold)*
<body>	*tells the computer to format the rest of the text as the default text style.*

The HTML format is a text store in the same way as ASCII format. HTML can be used for data storage and many packages, such as Microsoft Office, support data exports to HTML format. HTML is good for text but often handles data such as spreadsheets and databases very badly and file size can inflate dramatically. For the complexities of GIS data we require a storage format that is none-proprietary, like HTML, but has the same advantages of the *.mdb, *.dbf etc. The data equivalent of the HTML text is XML (eXtensible Mark-up Language). XML replaces the default HTML tags (head, body, etc.) with user defined data types: it is the export method used by the most powerful data storage solutions. An example of an XML layout for a GIS database is shown below.

XML Data Format

```
</TABLEDEF>
        <COLDEF name = “Latitude” datatype = “Long” />
        <COLDEF name = “longitude” datatype = “Long” />
        <COLDEF name = “altitude” datatype = “double” />
```

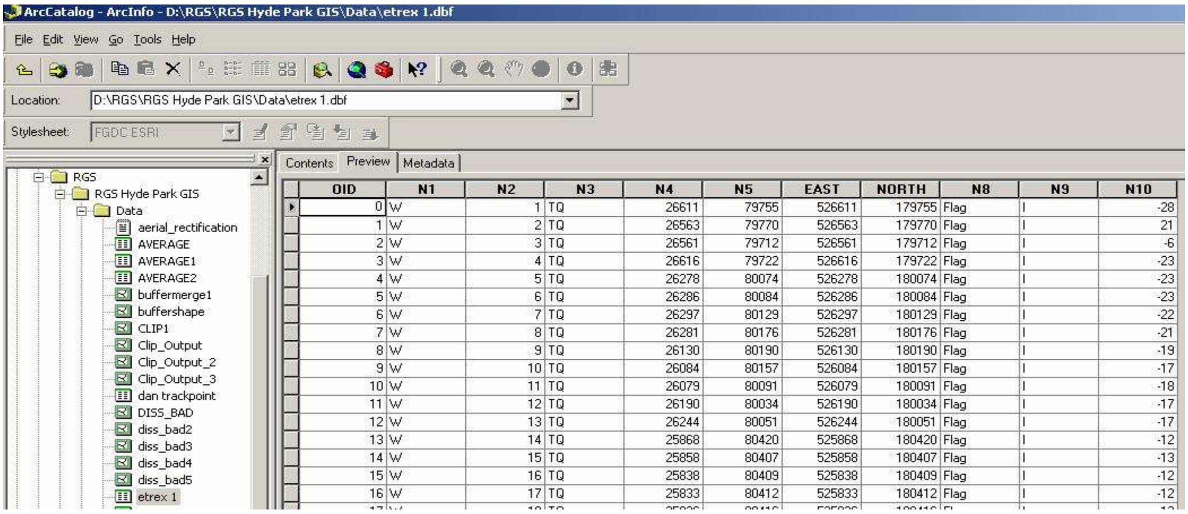
XML will become increasingly common as it is integrated into the more common Microsoft Office programmes. A basic XML generator is included with Microsoft Office XP. This is a limited XML solution and a far superior one is bundled in Office 2003. XML will by 2006 become the default method of saving data in Microsoft applications, with proprietary file formats becoming redundant. For some time ESRI have published ArcXML, a protocol to control data exchange between the ArcIMS Spatial Server, the Application Server and the Server connectors. There is a GISci specific mark-up language in development called GML (Geographic Markup Language) that has been developed by the Open GIS Consortium. More details about the development of XML style solutions for GISci can be found at *gislounge.com/ll/xml.shtml*.

One perennial problem associated with GISci data is the order in which latitude and longitude are referenced. As already mentioned in Chapter 2 the convention used in speech of saying latitude and longitude (Y values followed by X values) is rarely used in written terms. More often we write longitudinal values before latitude values XY. In Excel spreadsheets XY is the common format for referencing cells (A1, B1 etc.). However, when delving more deeply into Excel with any form of programming the convention of R1C1 is used (YX). It is imperative you understand how the data is being output from the team's GPS and how it is used in the GIS. Failure to ensure the correct format can cause the data to be plotted in a skewed manner.

Regardless of the import/export formats used, there can still be a problem with the format or quality of the data. The process of preparing information to be placed into the database is called cleaning. Most existing data will have fields that are not required in the final database. Stripping out redundant data and re-ordering columns is important, as is reformatting or reprojecting data. Cleaning can take long periods of time and is essential to building a good GIS. Software for cleaning and parts of the cleaning practise are discussed in the chapter on GIS software.

4.8 Adding raw expedition data

A full relational database may be too complex for some expeditions. In this case the database may be a flat format series of tables. These tables are usually *.dbf format. ARC GIS has a good utility for managing these files called ArcCatalog. ArcCatalog allows the user to make connections to various data sources to pull them into the GIS. In Figure 4-6 the data available as .dbf is listed in the left hand pane and a preview of the data is shown in the right hand pane. There are three different symbols shown on the left hand pane. The symbols show tabular data, point shape files and polygon shape files. For shape files, it is possible to not only see the raw data as shown in Figure 4-6 but also a preview of the data in a spatial framework. This is done by changing the drop down at the bottom of the screen from Table to Geographical.

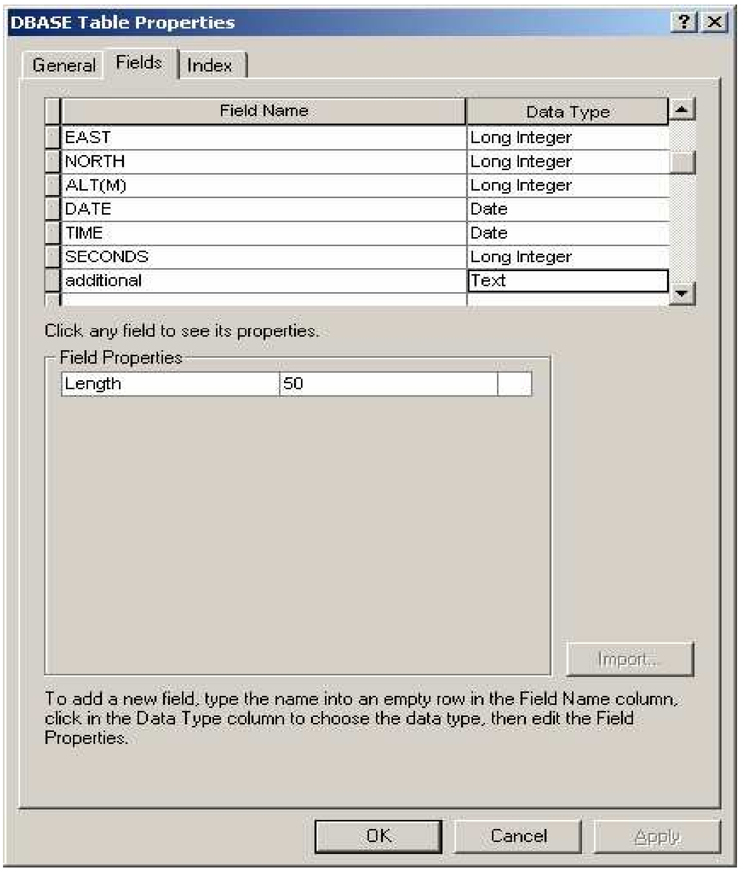


Location: D:\RGS\RGS Hyde Park GIS\Data\etrex 1.dbf

OID	N1	N2	N3	N4	N5	EAST	NORTH	N8	N9	N10
0	W	1	TQ	26611	79755	526611	179755	Flag	I	-28
1	W	2	TQ	26563	79770	526563	179770	Flag	I	21
2	W	3	TQ	26561	79712	526561	179712	Flag	I	-6
3	W	4	TQ	26616	79722	526616	179722	Flag	I	-23
4	W	5	TQ	26278	80074	526278	180074	Flag	I	-23
5	W	6	TQ	26286	80084	526286	180084	Flag	I	-23
6	W	7	TQ	26297	80129	526297	180129	Flag	I	-22
7	W	8	TQ	26281	80176	526281	180176	Flag	I	-21
8	W	9	TQ	26130	80190	526130	180190	Flag	I	-19
9	W	10	TQ	26084	80157	526084	180157	Flag	I	-17
10	W	11	TQ	26079	80091	526079	180091	Flag	I	-18
11	W	12	TQ	26190	80034	526190	180034	Flag	I	-17
12	W	13	TQ	26244	80051	526244	180051	Flag	I	-17
13	W	14	TQ	25868	80420	525868	180420	Flag	I	-12
14	W	15	TQ	25858	80407	525858	180407	Flag	I	-13
15	W	16	TQ	25838	80409	525838	180409	Flag	I	-12
16	W	17	TQ	25833	80412	525833	180412	Flag	I	-12

Figure 4-6 Example screen from ArcCatalog showing the available data for a project.

Right clicking on any of the tables listed in the left hand pane and selecting properties brings up a list of all fields associated with this table. The fields can then be customised in a manner similar to those discussed above in Section 4.2.



Field Name	Data Type
EAST	Long Integer
NORTH	Long Integer
ALT(M)	Long Integer
DATE	Date
TIME	Date
SECONDS	Long Integer
additional	Text

Click any field to see its properties.

Field Properties

Length: 50

To add a new field, type the name into an empty row in the Field Name column, click in the Data Type column to choose the data type, then edit the Field Properties.

Figure 4-7 ArcCatalog data management window. Data can be managed from ArcCatalog. Additional fields can be added and their data types can be selected. Adding data fields like this can be useful to merge GPS data and data collected on field slips without the need for a more complex relational database. If there are only a few annotations to be made then this method is quicker. If there are a large number of GPS points and annotations then it is better to prepare a full database solution.

By using the method shown in Figure 4-7 details about the GPS waypoint can quickly be added to the data. These details might include animal sightings or waypoint descriptions. Managing your data with ArcCatalog is a simple way of bringing lots of data together.

A good companion to ArcCatalog is Microsoft Excel. It is in Excel that most of the actual data will be added. Excel is a good tool for this as it can open and export *.dbf files. A problem with Excel, as with all flat file systems, is that the data is nit relational and it is very easy to corrupt the information. Care must be taken to ensure that the fields or columns are not sorted incorrectly and that the data is held together and not mixed. It cannot be stressed strongly enough the benefits of keeping data in a relational database over individual *.dbf files.

4.9 Accessing project data

As discussed in Section 4.3 it is important to structure the data correctly. Previously this section has looked purely at storing data and using normalisation for improving data accuracy. Table 4-5 Structured and unstructured queriesshowed how structuring data can be important for the speed of large databases. This section will look at the need for good data structures for accessing the data in the GIS. It is important that the data entered into the database can be extracted and plotted into the GIS effectively. In a habitat mapping exercise it may be necessary to note the type and number of animals seen at a location. There is often a temptation to record this in one long line of information as shown below in Table 4-6:

Table 4-6 Basic record structure on a fieldslip.

id	waypoint	animal	habitat
1	12	2 x impala 1 x buffalo	Grass land

Though this is easy for a human to read it becomes very difficult for a computer to extract the animals from the animal column. In fact plotting this in the GIS would be very difficult indeed. A better example might be to put the animals into the column header. It becomes clear how this is a better structure because the GIS can then plot the data for waypoint 12 as two different attributes.

Table 4-7 A basic record structure favouring plotting data in a GIS.

id	waypoint	impala	buffalo	Habitat
1	12	2	1	Grass land

The disadvantage of this structure is that the GIS requires a column for every possible animal. This becomes problematic especially if the slip needs printing because the page could be very long with dozens of animals and therefore not very practical. Though the second example is better for the computer and the GIS, it is not very practical for a human. A hybrid of the two is needed for the database schema. It is very common to record data manually in the field using the original format but to transcribe it into the database in a different manner. The best method to compromise between the two is to split the data into different tables; one containing data about the waypoint and one containing data about the animals, this is shown in Table 4-8.

Table 4-8 A relational data structure.

id	waypoint	habitat
1	12	grassland

waypoint	animal	number
12	impala	2
12	buffalo	1

A relational database tool such as Microsoft Access can be used to create this type of data. In the strictest terms this structure is not as far normalised as might be required because the animals are still listed as string values. As was shown in Table 4-2 string values are not ideal as they introduce error into any database. The best method for storing values is shown in Table 4-9. In some occasions it is perfectly acceptable to reformat Table 4-9 so that there is a tally against each animal id instead of a separate row for it.

Table 4-9 Many to many data schema.

id	waypoint	Habitat
1	12	Grassland
2	13	Grassland

waypoint	Animal_id
12	1
12	1
12	2

id	animal
1	impala
2	buffalo

The advantage of keeping data in a relational structure is that the data can be re-structured on demand to fit any of the tables shown in this section. Simple SQL queries can re-arrange the fields to form a structure suited to human reading or computer plotting as was shown in Table 4-8.

4.9.1 Understanding query operators

Now that we have our data as a table that can be easily read by the database, we can display it in the GIS. There are some important commands that need to be included in understanding selecting of data. The exact methods for implementing these commands (or operators) are shown below in Section4.9.2.

In the case of habitat study we might want to preferentially select a specific animal or exclude certain locations. These criteria are called the logical operators. The basic logical operators are AND, OR, NOT and XOR. They all have specific functions in choosing the correct data. The effect of the logical operators on the data presented in Table 4-4 is shown below.

If only the locations where impala and buffalo were seen we would use an AND query. This would leave only locations 12 and 15.

Table 4-10 Results of an AND query on the basic data.

species	habitat	waypoint	x	y	weather	height
impala	Grassland	12	55234	233444	clear	500
buffalo	Grassland	12	55234	233444	clear	500
impala	Grassland	15	55420	232418	clear	450
buffalo	Grassland	15	55420	232418	clear	450

Often this is a difficult concept for new users who expect the query to return locations 12,13,15,17 and 18. This is because the OR and AND operators are often confused because of their use in everyday language. Users often expect AND to return all impala and all buffalo sightings but it only returns the areas where both of them are found together. The OR query selects those where one or the other or both were observed.

Table 4-11 Results of an OR query on the basic data.

species	habitat	waypoint	x	y	weather	height
impala	Grassland	12	55234	233444	clear	500
buffalo	Grassland	12	55234	233444	clear	500
impala	Grassland	15	55420	232418	clear	450
buffalo	Grassland	15	55420	232418	clear	450
buffalo	Grassland	17	55412	235212	cloudy	422
impala	Forest	18	55214	235444	cloudy	455
impala	Forest	13	55280	233501	raining	200

The next logical operator is NOT and is used to exclude data. For example, sightings that were not conducted during clear conditions would result in the table shown below.

Table 4-12 Results of a NOT query on the basic data.

species	habitat	waypoint	x	y	weather	height
cheetah	grassland	14	55520	232500	cloudy	480
buffalo	Grassland	17	55412	235212	cloudy	422
cheetah	Grassland	17	55412	235212	cloudy	422
impala	Forest	18	55214	235444	cloudy	455
impala	Forest	13	55280	233501	raining	200

The final operator to be discussed here is the XOR query. XOR stands for Exclusive OR. When using an OR query the data returned included both the locations where both impala and buffalo were seen such as 12 and 15 as seen in Table 4.10 and the locations where one or the other is present such as 17, 18 and 13. XOR only returns 17, 18 and 13 but not 12 and 15.

Table 4-13 Results of an XOR query on the basic data.

species	habitat	waypoint	x	y	weather	height
buffalo	Grassland	17	55412	235212	cloudy	422
impala	Forest	18	55214	235444	cloudy	455
impala	Forest	13	55280	233501	raining	200

The basic logical operators can obviously be combined to return data in a more complex form by excluding some data and combining other data. Almost all operations in GIS spatial analysis involve the use of the simple Boolean operations, whose use in spatial analysis is shown in Chapters 3 and 7.

4.9.2 Implementing queries in a database

The rest of this chapter will concentrate on how the above described clauses can actually be used in SQL statements to select the required data. All the following examples use industry standard code that can be lifted directly into almost any GIS.

To plot the data onto the GIS it is important to understand the methods for selecting only the correct data. The Bogda Shan Expedition collected large quantities of geographical data. The biggest files generated were from downloaded GPS waypoints and trackpoints. The expedition collected 12,855 GPS points in a transect across the mountains. 11,233 of these had recorded height or elevation data. If the team was only interested in plotting the data that had accompanying height information, then a simple SQL query can be used. SQL queries usually contain at least two pieces of information, a command and a clause. A command tells the database what is required, this would most commonly be a command to select information without affecting the rest of the table, though it could be a command to delete or change the data.

The query shown below is selecting information about the GPS data according to a specific criterion; i.e. where the Elevation field contains data. This is referred to as the clause part of the query. The clause tells the database when to use its command. We are interested in the *select* command, when the Elevation field has a record of the elevation. The database assesses whether a record has data or not by describing it as Null (*empty*) or Not Null (*contains information*). Therefore our simple query would look like:

```
SELECT [GPS Data].Elevation
FROM [GPS Data]
WHERE ((([GPS Data].Elevation) Is Not Null));
```

The first line is the command section, in this case a simple select command, and we are interested in the *Elevation* field from the table *GPS Data*. The second line tells the database specifically the table to use and the third line contains the criterion. This is a very simple example; to be more useful to us we also want the database to return all the other fields in the table. The current query would simply return a list of elevation data. To return all fields where the GPS has recorded elevation data we need to specify them in the query.

```
SELECT [GPS Data].[Index], [GPS Data].[Sequential Numbering System], [GPS Data].GPS,
[GPS Data].[Way point], [GPS Data].[X Reference], [GPS Data].[Y Reference], [GPS
Data].[Estimation of error], [GPS Data].Satellites, [GPS Data].Elevation, [GPS Data].Name, [GPS
Data].Time
FROM [GPS Data]
WHERE ((([GPS Data].Elevation) Is Not Null));
```

Most often if all fields are required a wildcard character can be used to specify select all. It is the use of the wildcard character that makes the FROM statement important.

```
SELECT *,
FROM [GPS Data]
WHERE ((([GPS Data].Elevation) Is Not Null));
```

This now returns a replica of the original GPS Data table but with only 11,233 records. A query is transient; it does not make a hardcopy of the data and must recreate the list every time the list is required. This is useful if the table may change regularly but is slower than accessing an actual data table. To analyse the data after the Bogda Shan Expedition it would be more useful to have the fields containing GPS elevation data available as an actual table. To do this we can modify the SQL command slightly. We are still interested in GPS data where the receiver has recorded elevation information, but we want the database to put this data somewhere permanent. Consider the simple example again:

```
SELECT *, INTO [GPS Height Table]
FROM [GPS Data]
WHERE ((([GPS Data].Elevation) Is Not Null));
```

In the new example the command line has been altered to include the additional instruction INTO [GPS Height Table]. This instruction tells the database to take the selected fields and make a new permanent table called GPS Height Table. If the INTO table already exists this query will always destroy the old INTO table and make a new one of the same name. If we had a table set up in advance or if we were adding to a GIS that already existed we could add our data to that table without altering it by using an INSERT INTO command. An INSERT INTO command is often referred to as an append query.

```
INSERT INTO [GPS Data] (GPS Height Table)
SELECT *,
FROM [GPS Data]
WHERE ((([GPS Data].Elevation) Is Not Null));
```

The other type of query that can be used in analysing data is the UPDATE command. This changes data from one value to another. For example to change all co-ordinate from one system to another you might apply a formula to all the latitude and longitude values.

```
UPDATE [Waypoints] SET [Waypoints].[X ref] = ([Waypoints].[X ref]-10000)*1.41
```

These are the basic clauses and commands used in database structures. They allow data to be prepared and analysed by the GIS even if the information collected in the field was not in a structure the database is initially set up to use. The methods for doing this querying are detailed below in Table 4-14.

Table 4-14 Summary of commands and clauses for queries.

SQL Queries	**Logical Operators**		**Commands**	
	OR	Selects A or B (includes both)	SELECT	Selects data for a query but does not write the data anywhere.
	AND	Selects only where A and B are both equal to a criteria	INTO	Creates a new table for the SELECT query.
	NOT	Selects only where A and B are not both equal to a criteria.	INSERT INTO	Puts the results of the SELECT query into an existing table.
	XOR	Selects only where A or B is equal to a criteria but not where both of them are equal to the criteria.	UPDATE	Updates data in an existing table.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section B: Data

Chapter 5: Remote Sensing

5 Remote Sensing

5.1 Introduction

Remote sensing involves the use of aircraft or satellites to collect photographs or scanned images of the Earth’s surface. Remotely sensed imagery is just one of many types of geographically-referenced datasets that can be processed using a GIS. The origins of remote sensing date back to a photograph taken from a balloon in 1858. By World War I, the aeroplane had become the main platform from which aerial photography was collected. During the inter-war period, film chemicals were developed that allowed colour and infra-red photography: the latter was of particular interest to the military, as it highlighted camouflaged features. Since the 1950s, black and white aerial photography has been the basis of most Earth surface mapping: it still accounts for 99% of all topographic mapping (Petrie 2000). Landsat Earth-observation satellites, operative since 1972, produce digital images in the visible and infra-red parts of the spectrum: each image covers thousands of km^2, with sufficient detail to map many geo-ecological features.

Digital images have many advantages over photographs: digital data can be easily stored and processed by computers, perfect copies of an image can be made in seconds, and digital imagery is readily available for GIS analysis. Over 1999-2000 a new generation of satellites, utilising ‘spy-satellite’ technology, began to provide commercially-available multi-spectral digital images that could rival the detail obtained by aerial photography. Other recent developments, such as millimetre-precision Global Positioning System (GPS) receivers, improved digital data compression software and Internet transfers of remotely sensed data, have facilitated the use of remotely sensed data in GIS-based projects. The reader should refer to Lillesand & Kiefer (2000) or Drury (2001) for detailed coverage of remote sensing. Techniques for processing and interpreting remotely sensed images are reviewed in Chapter 8.

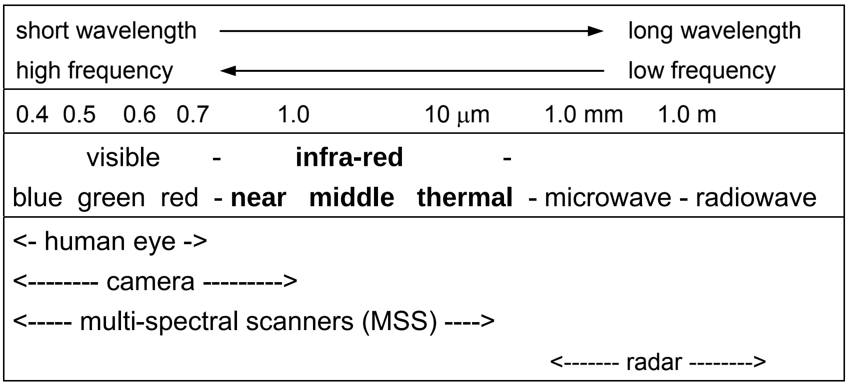


Figure 5-1 The electro-magnetic spectrum and associated sensors (1 μm = 0.001 mm).

The part of the electro-magnetic spectrum that is most widely used in remote sensing extends from the visible wavelengths, through progressively longer wavelengths, to the microwave and radio wavelengths used by radar systems (Figure 5-1). The limited ranges of human vision and conventional photography are apparent. Many features, particularly vegetation and water, show unique variations in the infra-red parts of the electro-magnetic spectrum. Thermal infra-red can be used to detect areas with high rates of evaporation or evapo-transpiration, due to their lower temperatures relative to their surroundings, as well as zones of thermal pollution in water, which could be caused by chemicals or hot water discharges from power stations. Finally, at wavelengths of millimetres to metres, the microwave or radio pulses utilised by radar systems are particularly useful for mapping soil moisture contents and areas of inundation, with the added advantage of being able to 'see' through cloud cover. Radar systems are particularly effecting at measuring surface roughness and thereby mapping the texture and shape of features on the surface of the Earth.

5.2 Satellite imagery

Our ability to produce maps of features on the Earth's surface, using remote sensing, is based on the characteristic ways in which different features reflect or emit electro-magnetic radiation. Figure 5-2 illustrates the reflectance responses of various types of Earth surface features: note how each has a unique response, known as its 'spectral signature'. Vegetation absorbs the red part of the spectrum for photosynthesis, but reflects the green and near infra-red wavelengths for which it has no use, the amount of reflectance varying with, for instance, variations in leaf moisture content, leaf shape and stress due to disease. Bare rock and soil exhibit a great variety of spectral variation, which can be used to determine their mineral composition. The near infra-red part of the spectrum is absorbed by water: this is particularly useful when mapping wetlands, as boundaries between dry land and wet areas are highlighted.

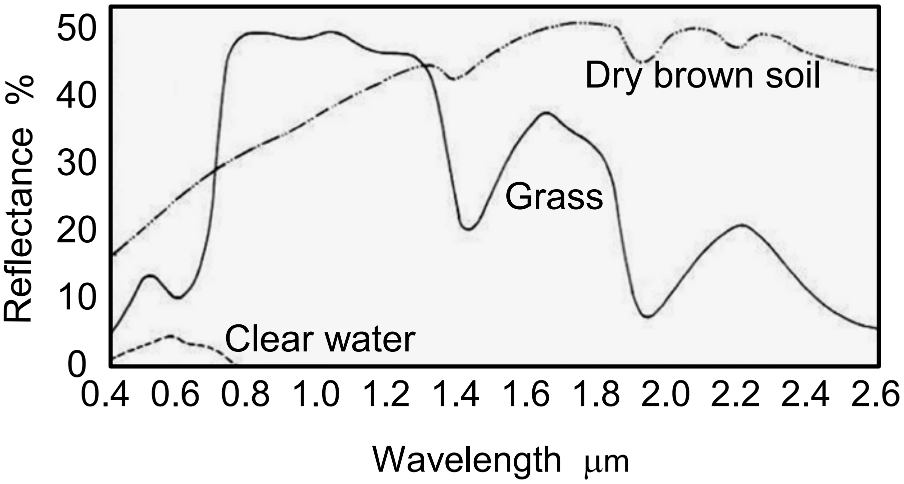


Figure 5-2 Electro-magnetic reflectance curves ('spectral signatures') for various surface features.

Multi-spectral scanners detect in the visible and infra-red parts of the spectrum. As we have seen above, each type of land cover feature has a characteristic spectral signature, with peaks and troughs of reflectance dependant on the wavelengths at which electro-magnetic radiation is absorbed and reflected or emitted by that type of feature. An illustration of the sorts of crude spectral signature that can be obtained from a 5-band multi-spectral scanner (MSS) is given in Figure 5-3. Note the general low reflectances with water, the high reflectances with beach sand, the low red / high near infra-red shift with photo-synthesizing vegetation and the corresponding low near-infra-red values in vegetation-depleted urban areas.

Sensors that can detect in hundreds of slices of the electro-magnetic spectrum, known as hyperspectral scanners, allow much more detailed spectral signatures to be collected, greatly improving the identification of different surface features. Most hyperspectral scanners are experimental and are mounted on aircraft (e.g. AVIRIS, CASI, HyMAP), resulting in only limited global coverage, but some space-born systems have recently been launched, notably MODIS (European Space Agency) and HYPERION (NASA).

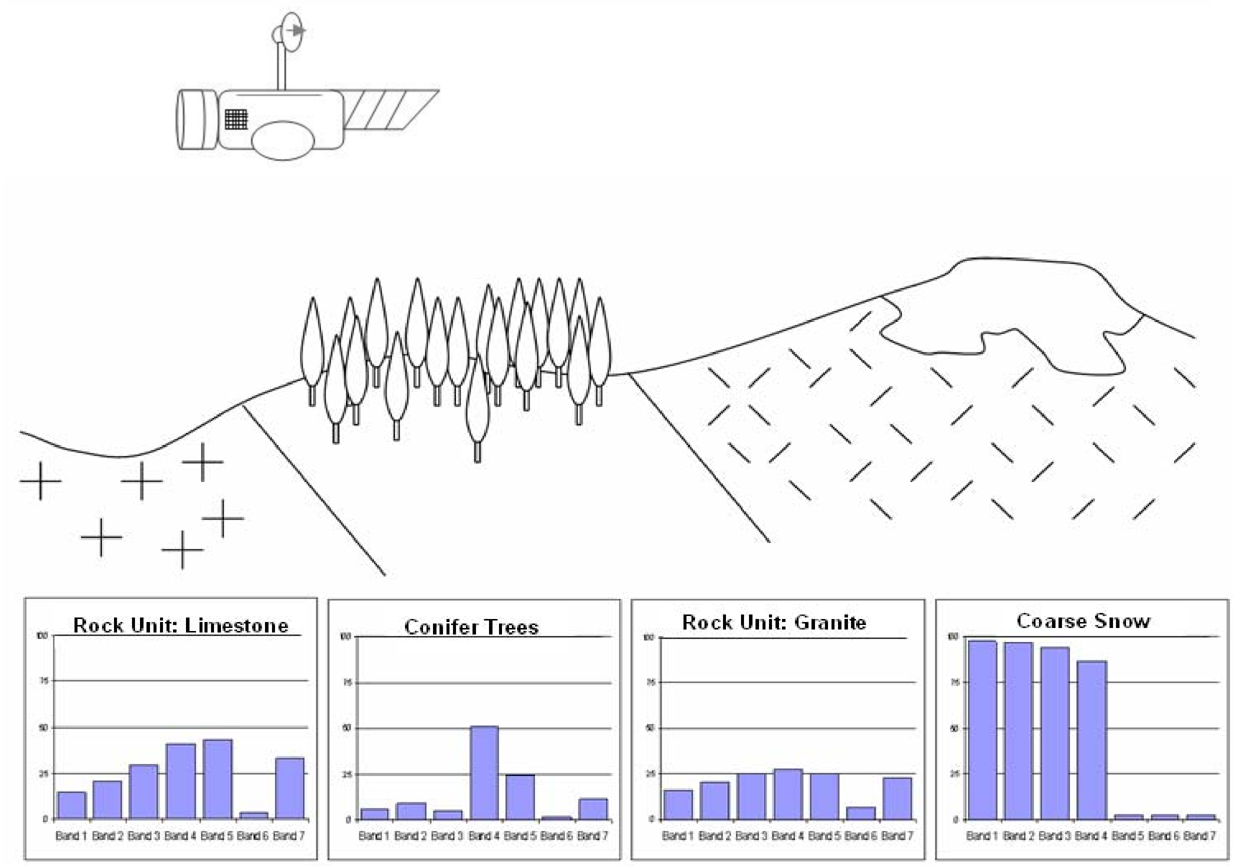


Figure 5-3 Spectral variations, approximating spectral signatures, from a multi-spectral sensor. Band 1 = blue; 2 = green; 3 = red; 4 = near infra-red; 5, 6 = middle infra-red; 7 = thermal infra-red.

5.3 Resolution

When selecting remotely sensed imagery for a given project, the various types of resolution are key considerations (Table 5-1).

Spatial resolution depends on the proximity of a given sensor to the surface of the Earth and the width of each image scene. For instance, the NOAA weather satellites orbit at 833 km: their images are 2400 km wide, with 1.1 km pixels. Landsat TM, orbiting at 705 km, is limited to images that are 185 km wide: consequently it has 30 m pixels. An even better spatial resolution comes from the Ikonos satellite, orbiting at 350 km: its images are only 40 km wide, but they have 1 m pixels. Figure 5-4 shows examples. Remote sensing from aircraft, usually as aerial photography, can provide imagery with resolutions in the 10 cm to 1 m range, allowing detailed mapping.

Temporal resolution is the frequency with which the orbit of a given satellite passes over a location on the Earth's surface. This is an important consideration with features that need frequent monitoring, such as forest fires or algal blooms. The NOAA satellites have a twice-daily temporal resolution, ideal for monitoring, whereas Landsat TM only has a 16-day return period.

Spectral resolution concerns the number of sample-slices of the electro-magnetic spectrum that a given sensor can take: radar satellites currently only sample one slice, or "band"; Landsat TM samples seven bands; and hyper-spectral sensors sample hundreds of bands. The greater the number of spectral bands, the easier it is to automatically map Earth surface features, as their resulting *spectral signatures* will contain more distinguishing details.

Finally, *radiometric resolution* controls the greyscale tonal range of the sensor, i.e. the number of bits into which the recorded energy is divided. Landsat TM produced 8-bit data (i.e. 2x2x2x2x2x2x2x2 = 256 values) with Digital Numbers (DNs) ranging from 0 to 255; while the earlier Landsat MSS only produced 7-bit data, with DNs ranging from 0 to 127.

Table 5-1 Varying spatial, temporal and spectral resolutions of satellite imagery. (Continued over page)

sensor system		**resolution**		
		spatial	**temporal**	**spectral**
AVHRR weather satellite	visible / IR thermal	1100 m	twice daily	5 or 6 bands: 0.4 - 12.5 μm
Landsat MSS	visible / IR	80 m	16-18 days	5 bands: 0.5-1.1 μm
	thermal	240 m	16-18 days	1 band: 10.4-12.6 μm
Landsat TM & ETM+	visible / IR	30 m	16 days	7 bands: 0.45-2.35 μm
	thermal	120 m	16 days	1 band: 10.4-12.5 μm
Landsat ETM+	panchromatic	15 m	16 days	1 band: 0.52-0.90 μm
SPOT	Visible/near-IR	20 m	26 days, or up to 5 consecutive days	3 bands: 0.5-0.89 μm
	panchromatic	3 m		1 band: 0.49-0.73 μm
	Visible/near-IR	20 m		5 bands: 0.43-1.75 μm
IRS	LISS	23 m	24 days	4 bands: 0.52-0.86 μm
	mid-IR	70 m	24 days	1 band: 1.55-1.7 μm
	panchromatic	6 m	up to 5 days	1 band: 0.5-0.75 μm
	wide field sensor	188 m	5 days	2 bands: 0.62-0.86 μm
ASTER	near-IR	15 m	16 days	3 bands: 0.52-0.86 μm
	mid-IR	30 m	16 days	6 bands: 1.6-2.43 μm
	thermal IR	90 m	16 days	5 bands: 8.12-11.6 μm

sensor system		resolution		
		spatial	temporal	spectral
IKONOS	Visible/near-IR	4 m	11 days: can move to new targets	4 bands: 0.45-0.9 μm
	panchromatic	1 m		1 band: 0.45-0.9 μm
Quickbird	Visible/near-IR	2.5 m	11 days: or up to 3 consecutive days	4 bands: 0.45-0.9 μm
	panchromatic	0.6 m		1 band: 0.45-0.9 μm
ERS-1 & 2	SAR / raw	12 m	16 to 18 days	C-band radar: 3.8-7.8 cm
	SAR enhanced	20 m	16 to 18 days	C-band radar: 3.8-7.8 cm
RADARSAT	SAR	8 m	24 days, or up to 3 consecutive days	C-band radar, with 4 polarization options
	wide field SAR	100 m		

The satellite systems summarised in Table 5-1 are very useful when mapping large-area regional features, such as major river systems, forests, or deserts. Using satellite imagery to map regional features is also very cost-effective in terms of the areas mapped with relative speed and the cost of the imagery per km². Complimenting the regional satellite imagery of the Earth is the Space Shuttle Radar Topography Mission. The SRTM has produced a global set of digital elevation data, based on 90 m pixels, which can be downloaded free from the University of Maryland website (see the weblinks on the CD).

For detailed observations, such as field systems and buildings, use can be made of the panchromatic (grey scale or 'black and white') imagery of IRS, SPOT and Landsat, with 5-15 m pixels. The Japanese ASTER satellite is becoming widely used, particularly in studies of geo-hazards, because of (i) its useful spatial and spectral resolutions (see Table 1) allowing effective mapping of soil and rock types; (ii) its ability to produce a 3-D Digital Elevation Model (DEM) of the ground surface; and (iii) its relatively low cost, at US$ 55 for a 60 km x 60 km scene. Another new development is the new generation of modified 'spy satellites', such as IKONOS and Quickbird, which have 0.5-2 m panchromatic pixels and 2-5 m multi-spectral pixels, producing images that cover hundreds of square kilometres, with detail that is beginning to rival aerial photography.

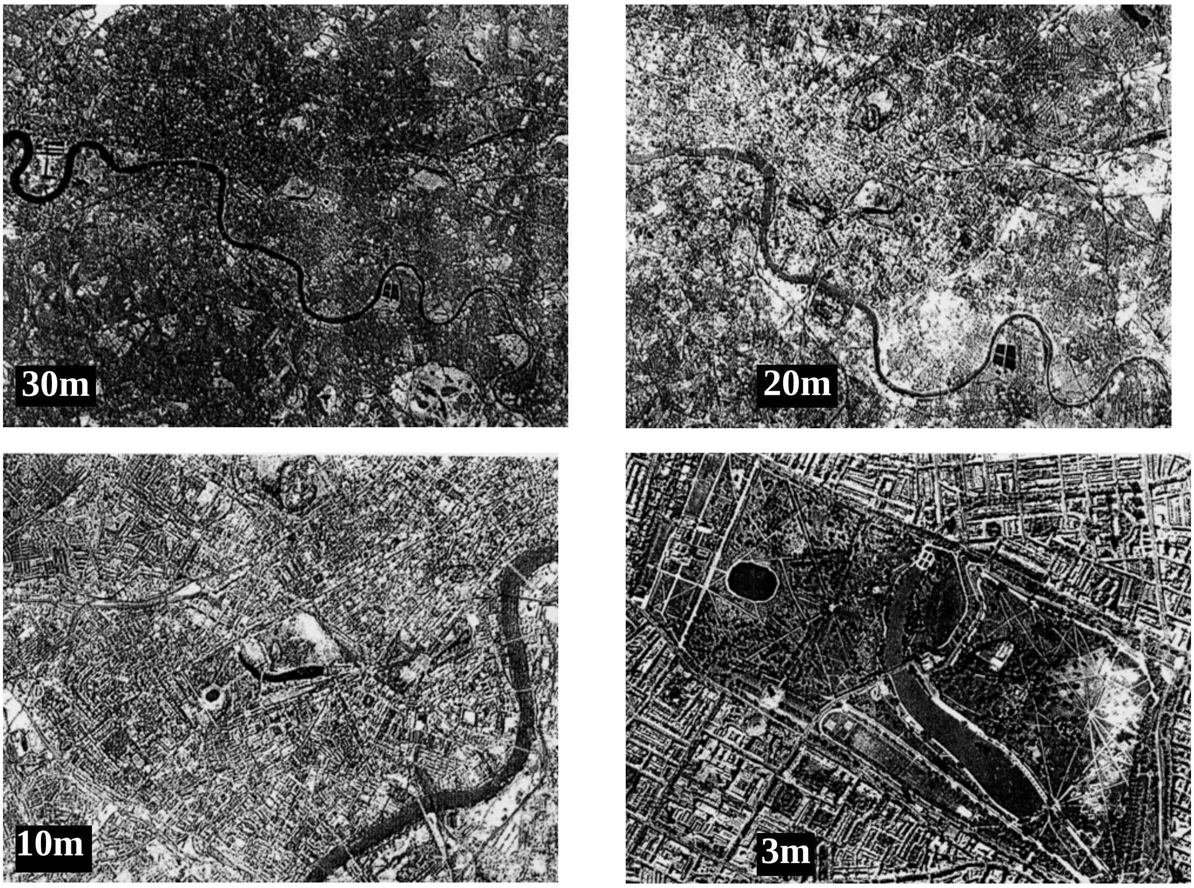


Figure 5-4 Satellite images of varying resolutions: Landsat Thematic Mapper (30 m pixels), SPOT multispectral (20 m pixels), SPOT panchromatic (10 m pixels), and Soviet KVR photo of Hyde Park (3 m). Images courtesy of NPA Group.

5.4 Aerial photography

For more detailed observations and mapping, aerial photography remains the most widely-used method for mapping geo-ecological features and the generation of contour maps via photogrammetry. Identification of vegetation types is easiest using infra-red film: false-colour prints are expensive, but black and white infra-red prints cost little more than panchromatic prints. Conventional aerial photography can be augmented by digital airborne sensing systems (e.g. Milton *et al* 1995). Using digital photography eliminates the loss of time and loss of detail associated with scanning aerial photographs, plus it can be directly utilised by most GIS software. For detailed mapping at scales of 1:10,000 or more, digital photogrammetry looks set to replace conventional photogrammetry, with the arrival of powerful - but affordable - computers and software, notably ERDAS Imagine's Orthomax module. Mount *et al.* (2000, 2003) review digital photogrammetric techniques, in a case study involving river channel changes (see Chapter 9 for detailed coverage of photogrammetry). Also see Plate 1.

With the arrival of relatively low-cost desk-top digital photogrammetry systems, studies of land cover change and dynamic geomorphological processes, such as erosion and landslide activity, have become much easier to carry out. The value of many airphoto archives, some covering the past 50 years or so, is now becoming evident. Many parts of the world are covered by American spy satellite photography, from the late 1950s onwards, notably the

KH series of satellites (available from the US Geological Survey). World-wide coverage of regional-scale Landsat multi-spectral digital imagery, useful for mapping changes in land cover types, dates back to 1972. Bear in mind that large areas of both airborne and satellite imagery may be obscured by clouds, so check cloud-cover details before purchasing. Amongst the negative aspects of aerial photography are the limited aerial coverage of each airphoto and the corresponding high costs (£2-£30/km^2). Airborne remote sensing is only feasible when relatively small areas, up to a few hundred km^2 in size, are to be mapped. In many developing countries, suitable aircraft may not be available and the areas to be mapped may be vast. Such problems have largely been overcome by satellite remote sensing; for instance, each image scene from the most widely used multi-spectral sensor, Landsat TM, covers 31,450 km^2, with a spatial resolution of 30 m. On a natural resources project in Ghana, re-prints of the only available airphotos, a panchromatic set flown in the 1950s, cost £2/km^2; in contrast, digital multispectral Landsat imagery from 1988 cost £0.02/km^2 and covered the entire region, rather than a few specific sites (Teeuw 1995). A summary of key aspects of remote sensing is given in Table 5.2 and a key to some of the terminology can be found on the following page.

Table 5-2 Key aspects of remote sensing systems.

	AIRBORNE SENSORS					**SATELLITE SENSORS**				
	spectral wavelength									
spectral resolution	**visible**	**NIR**	**MIR**	**TIR**	**micro wave**	**visible**	**NIR**	**MIR**	**TIR**	**micro wave**
• photography	✓	✓				✓	✓			
• laser / LiDAR	✓									
• RADAR					✓					✓
• multispectral	✓	✓	✓	✓		✓	✓	✓	✓	
• hyperspectral	✓	✓								
	limited by cloud cover				all-weather	limited by cloud cover				all-weather
spatial resolution	cm to m (depends on flight altitude)					m to km (depends on orbit altitude)				
temporal resolution	flexible (may be weather-dependent)					0.5 to 35 days (orbit-dependant)				
operational time period	some 1930s archive airphotos; US National Archive WW2 airphotos; most national / regional airphoto surveys are post-1947. 1950+ US spyplane airphotos 1967+ radar surveys (SLAR) 1980+ ATM multispectral 1990+ hyperspectral scanners 1995+ multi-band radar 2000+ LiDAR: laser altimetry (10-50cm contours)					1957-70 CORONA photography 1960+ weather satellites 1972+ Landsat multispectral 1978 Seasat radar 1980+ SPOT visible+NIR 1990 ERS radar, IRS multispectral 2000+ IKONOS, ASTER 2001+ ENVISAT (multi-sensor) 2002+ Hyperion hyperspectral 3m Radarsat 0.6m Quickbird				
approximate cost	varies, depending on flight altitude and processing costs: generally a minimum of US$50 / km^2					varies considerably: mostly less than US$ 0.10 / km^2, but generally with a minimum order of US$100				

Notes for Table 5-2.

remote sensing wavelengths (1µm = 0.001 mm)			
visible	0.4 - 0.7 µm	**NIR**, near infra-red	0.7 - 1 µm
MIR, middle infra-red	1 – 3 µm	**TIR**, thermal infra-red	3 - 15 µm
microwave	1,000-1,000,000 µm (1 mm - 1 m)		

5.5 Radar Imagery

Radar (Radio Detection and Ranging) initially used radiowaves to determine the distance of an object from the sensor. Radiowaves emitted by the sensor bounce back off objects: the greater the time taken for the radiowaves to return, the further away the objects are from the sensor. Modern radar systems use microwaves instead of radiowaves. Radar has three main advantages over other forms of remote sensing: (1) it can 'see' through cloud cover; (2) it generates its own source of electro-magnetic radiation, so it is not dependant on reflected sunlight and can operate at night; and (3) long wavelength (c. 1 m) radar can penetrate dry sand to a depth of about 5 m. There are two major radar applications for expeditions:

- *Rainforests:* previously deprived of remote sensing data because of excessive cloud cover, can now obtain radar images showing terrain, vegetation cover, water bodies, roads and settlements;
- *Deserts:* using long-wavelength radar, images have detected structures buried under sand dunes. Such radar images have proved to be useful aids to mapping underground water supplies and highlighting archaeological sites.

Side-Looking Airborne Radar (SLAR) was developed in the 1960s. The pioneering RADAM-BRASIL survey of 1971-76 revealed the previously un-mapped Amazon Basin. Many governments in the humid tropics (e.g. Indonesia, Malaysia, Nicaragua, Nigeria, and Venezuela) have commissioned radar survey of large regions. Radar sensors have to generate their own electro-magnetic radiation, consequently they are larger than most other remote sensing systems (which passively record solar radiation reflected or emitted by Earth surface features): this delayed radar's deployment on satellites until 1978. Since then, many more radar satellites have been launched and applications developed in monitoring slope stability, subsidence, tectonic and volcanic activity, as well as land cover mapping.

5.6 Recent developments

A recent development related to radar has been Light Detection and Ranging (LiDAR), also known as laser altimetry because it uses a laser beam to measure the elevations of surface features. Airborne LiDAR has been very useful in mapping the heights of trees and generating detailed maps of floodplains, producing contour maps with 15-50 cm intervals, however, it is expensive data and therefore often limited to specific sites.

Integration of remotely sensed data with GIS-based mapping and analysis has become a major growth area, due to five coincidental developments:

- Relatively low-cost Earth observation imagery, with 15 m pixels and good multispectral capabilities, such as Landsat ETM (about US$ 600 per 180 km x 180 km scene) and ASTER (US$ 60 per 60 km x 60 km scene);
- The launch of very high resolution (VHR) satellites, yielding images with 0.5 m – 3 m panchromatic pixels and 3 m – 5 m multi-spectral pixels;
- Hyperspectral satellites, allowing the automatic mapping of Earth surface features, based on computer-generated spectral signatures derived from hundreds of detected wavelengths;
- The ending of the US Government's GPS Selective Availability, improving the accuracy of standard GPS readings from +/- 100 m to +/- 10 m (+/- 3 m with readings averaged over a few minutes);
- Improved integration of GIS and Internet technology should greatly improve the speed and accuracy of fieldwork mapping. The use of data compression techniques, such as the Multi-resolution Seamless Image Database compressor (MrSID), allows rapid Internet transfers of digital imagery.

5.7 Sources of remotely sensed data

5.7.1 Airborne or satellite analogue photography

Airphoto coverage of former British colonies may be held by the Natural Resources Institute (NRI, formerly the ODA) or the Ordnance Survey and three major survey companies: Hunting Surveys, Clyde Surveys, and Aerofilms. For French dependencies or ex-colonies, contact the Phototheque Nationale of the Institute Geographique Nationale. Details of overseas airphoto surveys carried out by the USA can be obtained from the US Geological Survey. To obtain copies of airphotos covering overseas countries, you will almost certainly require an official letter of permission from the relevant land survey department or associated ministry.

Satellite remote sensing began with the USA's CORONA spy satellite missions in the late 1950s. Panchromatic photographs with spatial resolutions of up to 2 m, collected from the KH series of satellites, are available from the US Geological Survey. Photographs from the later Apollo missions can be obtained from NASA's Goddard Space Flight Centre. High quality photographs were obtained using Large Format cameras on the Skylab space station and on the Space Shuttle missions; these can be obtained from the US Geological Survey. 10 m-resolution photographs were taken using Metric Cameras on Space Shuttle missions, these can be obtained from NASA or the German space centre (DLR). Due to the relatively narrow fields of view and the limited numbers of orbits, the chances of getting satellite photos of your study area are slim - but it is well worth enquiring, as their prices are generally low.

5.7.2 Digital images

Scanners that detect variations in visible, infrared and microwave radiation over the Earth's surface can be carried by both aircraft and satellites. Aircraft Multi-spectral data is usually of high spatial resolution (1-10 m) and very high cost. Free, or low cost airborne imagery, derived from test flights, is available, but largely limited to the arid regions of the USA and Europe. The websites of Infoterra and the UK's Natural Environment Research Council (NERC) and the German aerospace agency (DLR) contain useful information on the various types of airborne remote sensing systems.

The main contractors for airborne radar surveys are Aero Service (USA) and Interra Technologies Ltd (Canada). HTS Ltd (UK) (formerly Hunting Technical Services) has also carried out a number of surveys in developing countries. Useful, though limited, coverage of the world was obtained by the US Space Shuttle missions, using the Shuttle Imaging Radar (SIR-A, SIR-B, SIR-C). Although a swath only 50 km wide was viewed during most orbits, many cloud-covered tropical regions were observed for the first time. Another plus point is that the SIR data are *free* from the US National Space Science Data Centre.

A vast array of information on digital remotely sensed images can be found on the Internet: a summary of useful Internet websites is given in the Appendix. The US Geological Survey website is particularly useful for checking global coverage of satellite imagery. It is also worthwhile checking the websites of commercial suppliers of satellite imagery, such as Infoterra (formerly NRSC) and Nigel Press Associates (NPA Group), as they may have archive imagery of your study area. A useful summary of data sources for geoscience fieldwork projects has been produced by Malcolm Whitworth of the School of Earth and Environmental Sciences at Portsmouth University and can be downloaded from *http://web.port.ac.uk/departments/sees/staff/whitworth/dataguide/*

5.7.3 How much will it all cost?

Taken as a single purchase, even Landsat MSS or ETM data (US$ 100 to US$ 500 respectively) can seem to be a luxury item for an expedition; but that price is very reasonable when seen as a cost of *less than one penny per square kilometre*. For comparison, the cost of traditional ground survey techniques to map 1 km^2 has been estimated at £12.50, with only 4 km being mapped in a day (Cooke & Doornkamp 1990). Aerial photography is generally much more expensive per square kilometre, than satellite imagery. Individual airphoto prints tend to cost at least US$50, each print typically covering less than 100 km^2. The market for remotely sensed satellite imagery is very competitive with some dramatic reductions in the price of Landsat data recently and occasional ‘special offers’ from suppliers. As prices can vary so much over just a few weeks, the best approach is to check the websites of major supplies given in the Appendix.

One final tip: discuss your expedition research programme with local natural resource organisations, be they government departments, conservation groups or mining companies – they may give you free access to aerial photography or satellite images and could also provide valuable logistical assistance, while giving you a chance to undertake projects and produce results that will be locally valuable.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section B: Data

Chapter 6: The Global Positioning System (GPS): Principles & Concepts

6 The Global Positioning System (GPS): Principles & Concepts

Over the last five years Global Positioning Systems (GPS) have changed the way fieldwork is conducted. There are two principal reasons for using GPS in the field; these are navigation and determining co-ordinates for points in the GIS. This manual will not deal in depth with navigation, as this topic is described well elsewhere (for a good introduction see Simmonds (2004), which is included on the CD accompanying this book). Navigation is touched on briefly in Section 6.1 and the reader should note that, even though GPS are excellent tools for field navigation, their very nature as electrical equipment means they are fallible. As such, a more traditional backup including a map and compass are essential.

This chapter discusses GPS in detail. It does not aim to describe all manufacturers' units and does not replace the unit's user manual. Even though many aspects of GPS are described, not all of them will be relevant to any given expedition. There are many different types of GPS and different methods for using them. These differences give accuracies ranging from several centimetres to tens of metres. Chapter 11 discusses the appropriate use of GPS for various expeditions and teams should not always be concerned with obtaining the most accurate sets with the most features if this is not appropriate for their studies. An informed decision cannot be made without a thorough understanding of all the aspects of GPS so this chapter describes as much relevant GPS information as possible. Some of the techniques will be too involved for smaller expeditions and expeditions should study this chapter in conjunction with Chapter 11 to select the most practical and appropriate methodologies. Expeditions should not select expensive, time consuming and difficult to use navigation solutions if they are not required. Although there is always a push towards more accurate and precise methods, they should not be used if not required. Studying this chapter should help you to make an informed choice.

6.1 GPS and field navigation

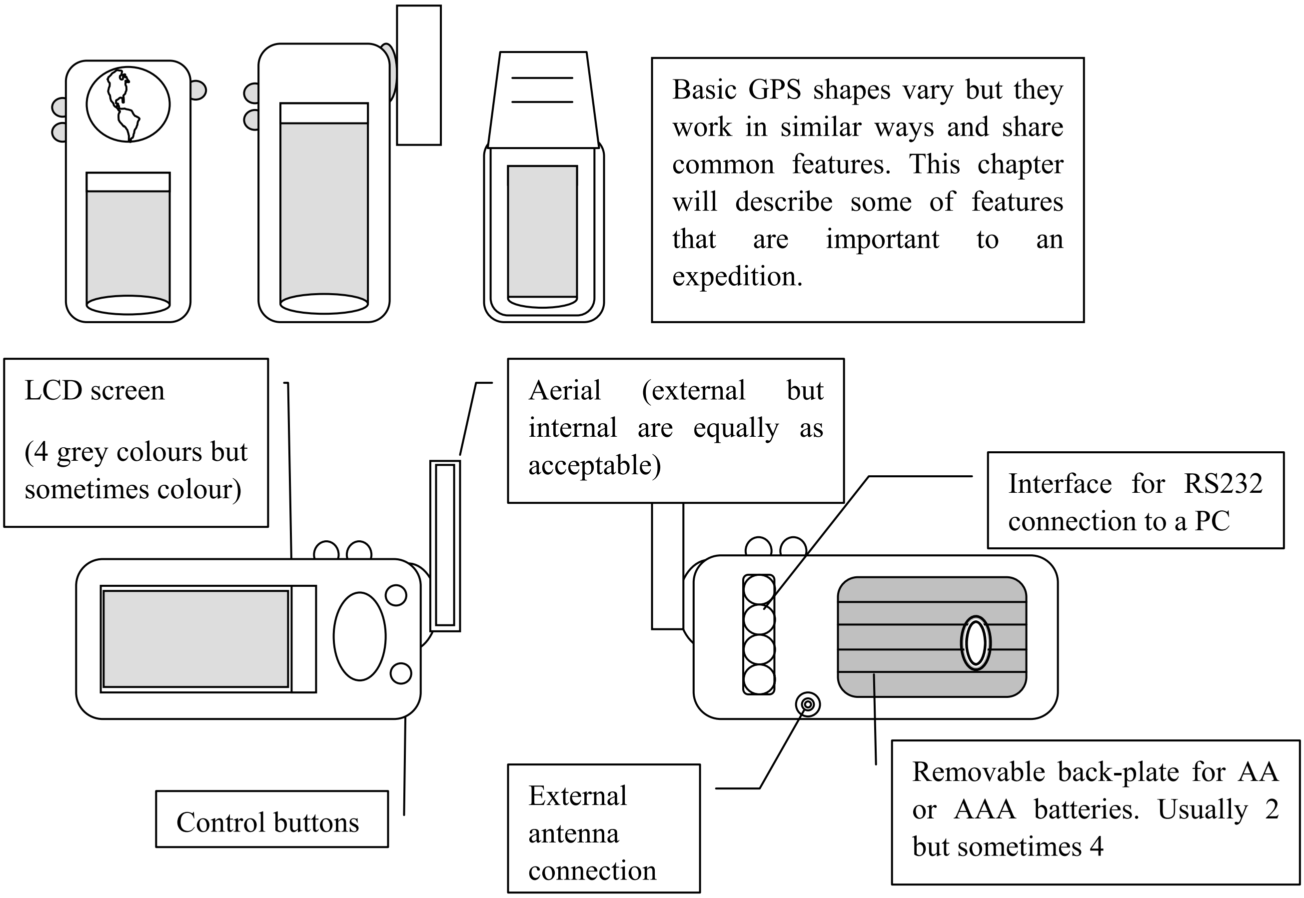
Navigation is vital to the safety of any field expedition. When combined with the necessity of fixing a location's co-ordinates for scientific research, the need for accurate, rapid and cost-effective navigation tools becomes paramount. Increasingly GPS receivers are becoming a standard – some would say essential – item of expedition equipment. Determining the co-ordinates of a point in the field can be achieved in a number of ways. The most common traditional approach involves triangulation with a map and magnetic compass. Triangulation (see Chapter 10) is often very accurate but relies on accurate maps and navigable objects. The Ordnance Survey of Great Britain produces very reliable maps but even they admit:

> *"On top of the nationwide errors in OSGB36, individual features on the map may only have been surveyed to a local accuracy of 7 m (for 1:25,000 scale maps) and some features such as boulders may only be shown schematically.*"
>
> Ordnance Survey Website

The result is that any triangulation achieved is relative to the map, which may in fact be quite inaccurate. Lines on navigation charts have accuracy on paper of ±1.5 mm. On a 1:10,000 chart that could be an error of 75 m. In addition, when drafting, the tools used may introduce additional errors. Triangulation is also time consuming and of limited use outside of areas of human influence i.e. those areas with man made objects surveyed to an acceptable accuracy. Other methods have been employed to determine location but they are either difficult in the field or rely on expensive equipment, examples include sextants for astronomical positioning and various types of theodolites for astronomical triangulation. There has for some time been a move to establish Global Navigation Systems (GNS) that are quick, cost effective and reliable. GPS has been the most successful of these systems.

6.2 Introduction to GPS functions/features

GPS use satellite data to calculate an accurate position on the earth. These calculations can relate the user's position to almost any map projection within milli-seconds. All GPS work in a similar manner but they often look very different and have different software. The most significant difference between GPS receivers is the number of satellites they can simultaneously communicate with. Most receivers are described as 12 channel meaning they can communicate with 12 satellites. Older models may be 8 or even 5 channel with more modern receivers capable of communicating with 14 – 20. Given the current (2005) makeup of the GPS satellite's constellation 12 channel is more than adequate.



Basic GPS shapes vary but they work in similar ways and share common features. This chapter will describe some of features that are important to an expedition.

Figure 6-1 A basic guide to GPS models.

Almost all units have an LCD screen or at least software that links to a PC/PDA with an output screen. The unit might have several different pages that can be displayed on screen but usually the default page is very similar. Commonly on starting a receiver you will be presented with a map of the satellites in view. The GPS receiver shows a view of the sky split into four quadrants. These represent the NE, SE, SW, NW parts of the sky, with the concentric circles representing the horizon at 90° from the zenith, with the inner circles representing 60° and 30°. The cross at the centre represents the zenith. The dots/circles represent the satellites and the bars at the bottom represent satellite signal strength. The higher the bar the stronger the signal. This display is typical of a 12 channel set. The dots and bars will commonly be labelled with a number to represent the identity of the satellite. The bars are commonly either hollow or solid (usually white or black on a monochrome display). Hollow lines represent a satellite for which the Ephemeris data is not known. It is therefore not being used to calculate a position. Black bars represent "Fixed" satellites whose ephemeris data has been collected successfully. These satellites are thus available for calculating a position. This is not consistent across all models and some may use grey bars as well as hollow bars to represent satellites not yet fixed.

The number, position and strength of signal from the satellites allows the GPS to calculate a rough estimate of the error in its reported position. This error or dilution of precision is a good guide to how accurate any reading would be. It should be closely monitored and readings should only be taken when this is below 10 m (ideally below 5 m).

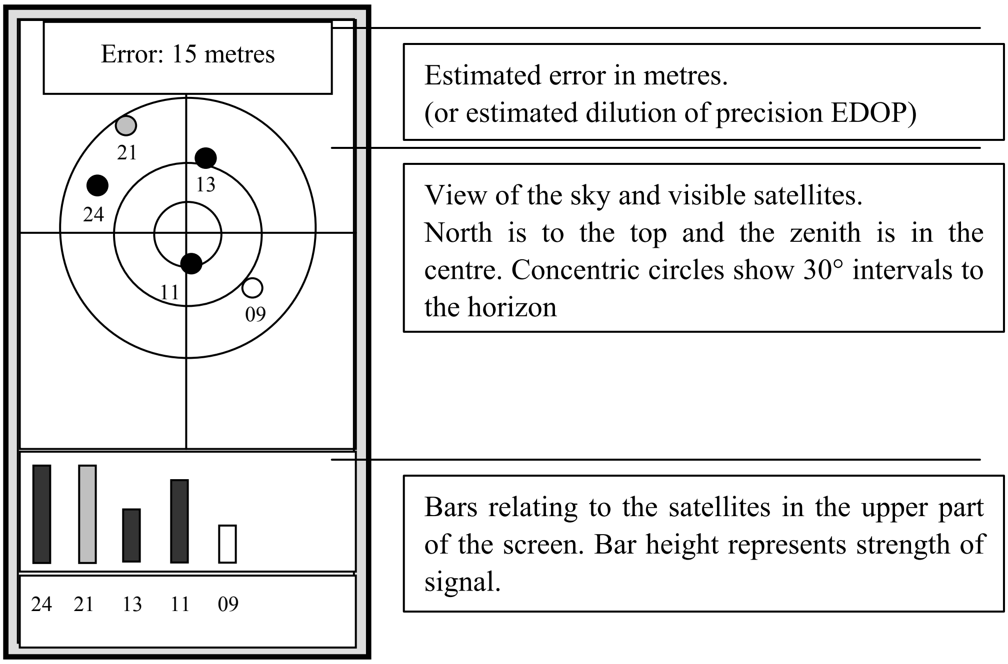


Figure 6-2 Basic GPS screen layout.

The way the GPS records data is generally the same across all units. GPS receivers automatically records data into their memory according to elapsed time or distance moved. These points are called trackpoints. The device can be forced to record additional data, generally with additional information, at user discretion. These user recorded points are called waypoints. Some of the common pages used for viewing this data are shown below. The more expensive sets have more detailed screens.

The GPS receiver can display the entire tracklog 'all trackpoints collected' or a selected route 'specific associated waypoints' on its display. This can show the direction travelled or plot a course to follow. GPS receivers can receive data uploads, either through a COM port or from user input. These uploaded co-ordinates might describe a route to follow or mark locations of scientific interest. The GPS set can then be used to navigate to these areas. In the diagram opposite, two symbols are used for the start and end of a route and waypoints along the route are marked with the locational Ids 1-4. In between these a route has been drawn from automated trackpoints.

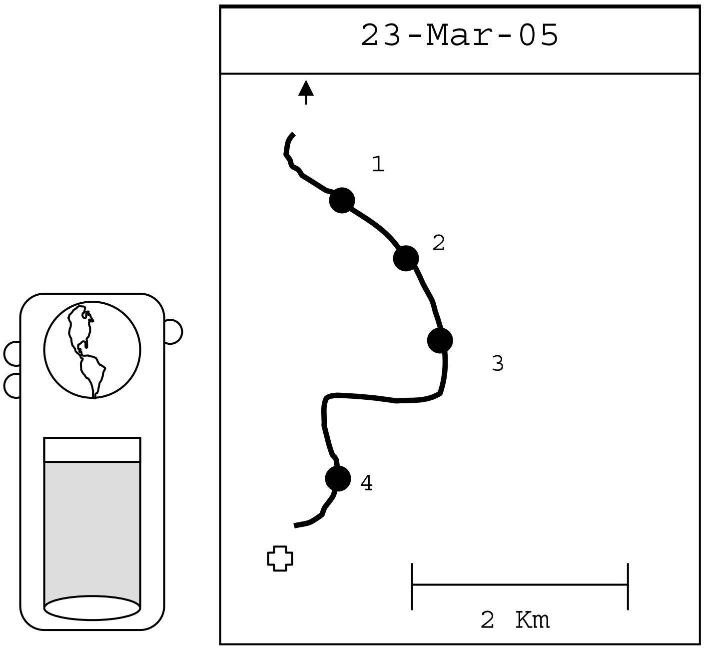


Figure 6-3 Basic GPS Tracklog screen showing track and way points.

In some newer models trackpoints can be used for another purpose. Those sets with inbuilt '*area calculations'* can use tracklogs to calculate the area of an object. By selecting the beginning and end of a log, the GPS receiver will attempt to calculate the area enclosed.

The trackpoints on the diagram to the right have been used to create an area calculation tracklog of a lake. This method can be quickly used to calculate areas of larger objects. However, the ~15 m GPS positional errors mean that anything smaller than around 2500 m^2 will be subject to very large errors:

E.g.: Area of a 50 m wide object

50m ±15m x 50m ±15m =

1225m^2 – 4225m^2

A large percentage difference (*345%*)
Whereas: Area of a 1km wide object
1000m ±15m x 1000 ± 15m =
970225m^2 - 1030225m^2
Which is a significantly smaller percentage difference (*6%*)

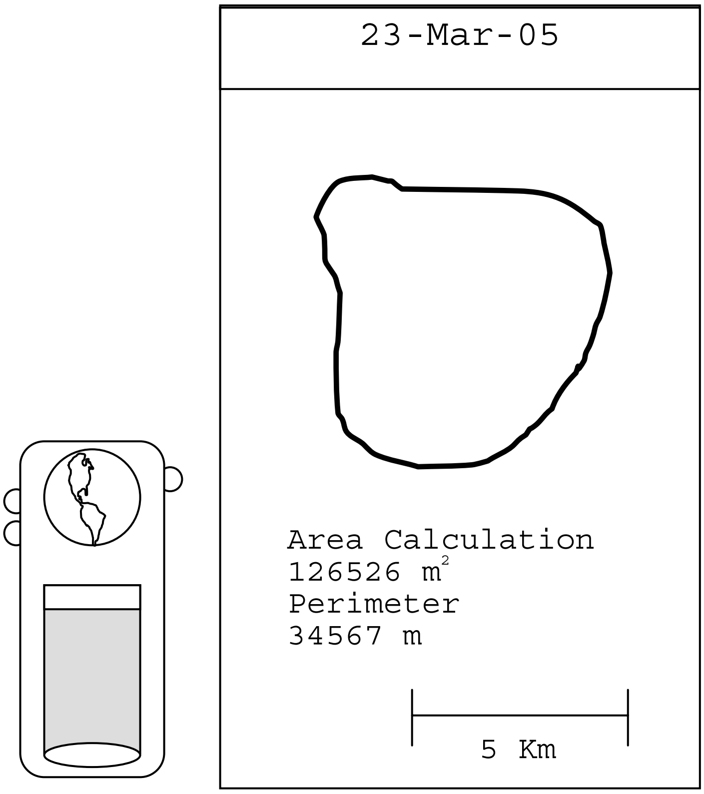


Figure 6-4 GPS area calculations showing when they should and should not be used.

Care must always be exercised before leaving for the field that the characteristics of a given GPS receiver are understood. For example, some receivers will not record elevation within trackpoints while others will not record elevation below sea level. Either of these might be crucial to the aims and objectives of a team. Other common GPS features include support for a ‘Man Over Board’ (MOB) alert. MOB automatically records details of the current location and immediately instructs the user how to return to that point. This is commonly used at sea to return to a lost crewmember but can equally be used to return to any expedition location or the expedition vehicle.

GPS receivers can often be used as a complete navigation tool, not only offering directions and location details but also navigation tools to move between locations. Most receivers come with an inbuilt digital compass. The digital compass is based on data from satellites and is not a magnetic compass. The digital compass will only work when moving and will not re-orientate if the set is rotated. More expensive models can come equipped with a magnetic compass and an inbuilt barometric altitude calculator.

The navigation pages often include a digital compass, an odometer showing distance travelled since the counter was reset, a current speed indicator, a maximum speed and average speed for a trip. The navigation page can also be tailored to show a variety of other statistics e.g. to show how far is remaining until an objective (waypoint, MOB etc) is reached. Some high-end sets include a magnetic compass and a barometric altimeter. These are generally more accurate than the satellite determined heading and elevation. GPS receivers can also display the time of sunrise and sunset at a given location.

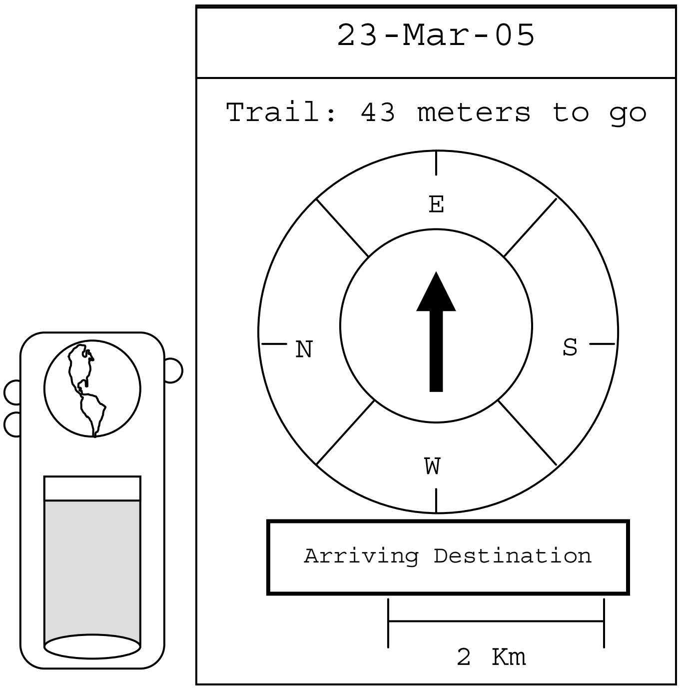
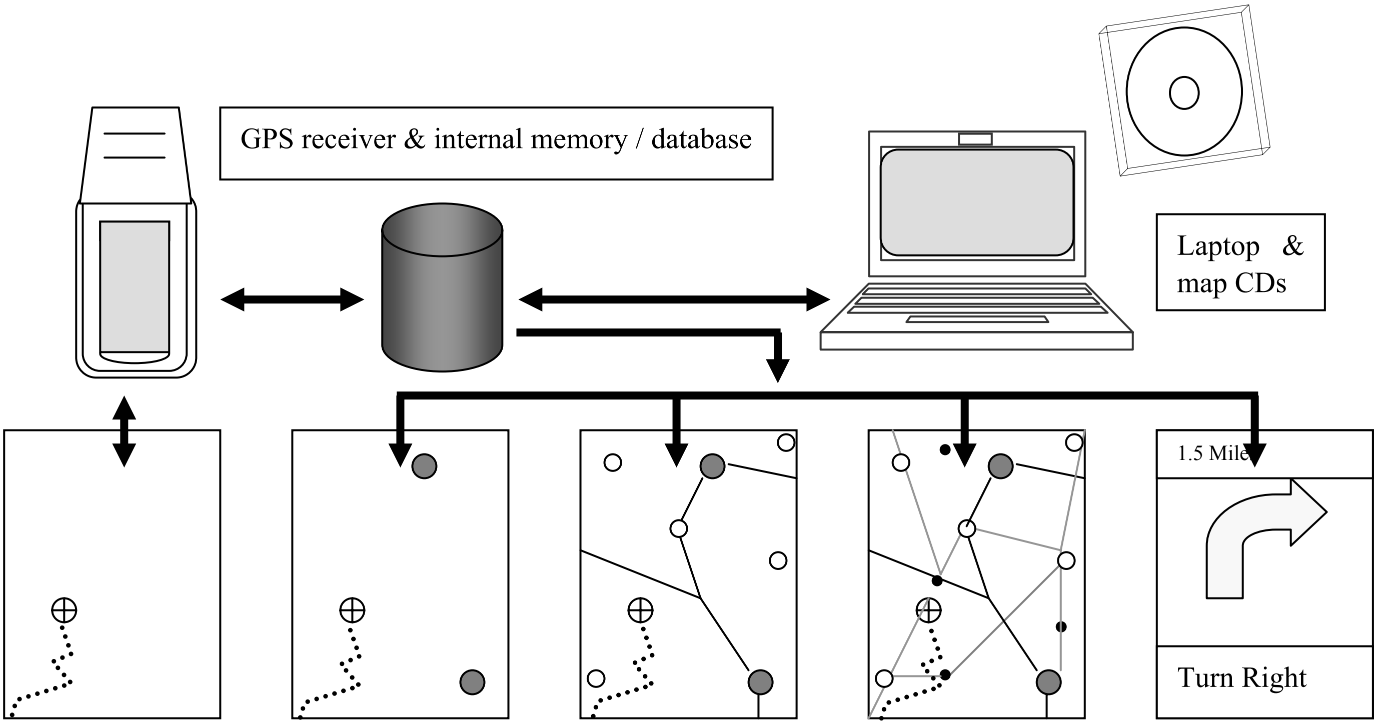


Figure 6-5 Basic GPS Man Over Board or Track Back features.

Standard GPS receivers range in price from under £100 to over £400. There are specialised high accuracy models that can cost up to £40,000 but the common models that an expedition might use are typically under £200. The difference between the models is not related to their accuracy but the number of additional screens and levels of data in the receiver’s database as shown in Figure 6-6. This database can usually be expanded by purchasing CDs of street-level data. These might cost around £100 on a per country or area basis but are of little use to an expedition. Most commonly fieldwork will take place in areas where road maps for GPS may not be available and printed maps would probably be better.



Basic GPS	Intermediate GPS	High-End GPS	Street-Level Maps	Navigation Aid
Display shows current position a track of where the user has been and any user entered locations.	Display shows a database of major cities and towns.	Display shows major cities down to minor towns and most roads.	All roads and streets, but GPS memory means cities or areas may need to be loaded from a PC.	GPS can calculate routes and offer real time accurate directions to a driver.
ETREX	ETREX Venture	ETREX Legend	ETREX Vista	Tom Tom Nav’

Figure 6-6 The difference between GPS models and their databases.

For most expedition work, a basic or intermediate GPS would be sufficient. The rest of this chapter describes in more detail how GPS actually calculates its position and how the data can be manipulated in a GISci context. Street level mapping and in car navigation aids are rarely useful for GISci fieldwork. It is better to concentrate on basic units that have features that can benefit the team. The buyer’s guide in Appendix 4 describes these features in more detail.

6.3 GNS history

Though there are various land based navigational services such as DECA and Loran, this chapter looks purely at satellite GNS. In the 1950s the US Navy began a programme to study navigation from artificial satellites. The first satellite navigational aid, TRANSIT, was accurate to approximately 160 m for stationary receivers. Moving receivers introduced additional errors of around 1 km per 1 m per second speed. These initial tests were generally accurate for a ship at sea but were of limited use for navigating into ports or shallow waters. The TIMATION I satellites launched in 1967 allowed comparatively slow moving receivers to calculate positions via embedded atomic clocks. This was much more accurate for ships but soon the aviation industry became interested and a system was developed for faster objects. The current Global Position System suitable for aircraft and high-speed navigation, NAVSTAR, was initiated by the US Air Force in 1978. Further details about this can be found at: *www.britannica.com* using a search for GPS.

NAVSTAR theoretically gives global coverage with accurate positional information down to sub metre levels with elevation at sub 10 m accuracy. Four years later the USSR launched a similar system, GLONASS. GLONASS offers global coverage like the NAVSTAR system but has less than half the satellites (9 as of 2004 but with more launches planned to take this back to around 14 by late 2005). The system has a strong bias towards the Northern Hemisphere. GLONASS results obtained below the equator are less accurate and the system has suffered from poor maintenance in recent years. Northern hemisphere results are considered slightly more accurate than standard NAVSTAR results. In December 2004 the US and Russian Federation agreed to co-operate on the development of their systems.

6.3.1 The NAVSTAR system

The NAVSTAR system is managed by the Interagency GPS Executive Board (IGEB) of the US government. Details of this can be found on their website at *www.igeb.gov*. A statement by President G.W. Bush in December 2004 indicates that this may change in the future but the IGEB website remains an excellent source of up to date information. The current specification requires satellites to orbit the Earth in one of six orbits inclined at different degrees to the equator, between –55° +55° at an altitude of 20,200 km. The DoD maintains 4 satellites in each orbital plane, giving a total constellation of 24 satellites, currently supported by up to 5 spares. Satellites are being replaced over time and the newest satellites are referred to as GPS IIR SVs. The design of the satellites allows navigation at all latitudes during all weather conditions. The L-band radio wave used to communicate to Earth from the satellites is effectively immune to local atmospheric conditions such as rain, storms etc. The satellites broadcast two L-band signals (L1 and L2) operating at the following frequencies, L1 = 1575.42 MHz and L2 = 1227.6 MHz. The NAVSTAR system operates two services, standard and precise. The Standard Positioning System (SPS) is available worldwide at no charge and operates the L1 frequency. The Precise Positioning Service (PPS) broadcasts on the L2 band only and accurate data are attained by correlating the two bands. PPS receivers are used solely by the U.S. Military and allies, as well as by the U.S. Federal Government. Applications for access to PPS by non-Federal Government organisations, both domestic US and foreign, can be made and are considered on a case-by-case basis. As a result, all expedition GPS use will probably involve only the SPS signal. Further details about the SPS signal used by most expeditions can be found at *www.igeb.gov/SPS-2001-final.pdf*.

6.4 How GPS works

GPS signals do not contain positional data. The position reported by the receiver on the ground is a calculated position based on range-finding triangulation. GPS positioning is achieved by measuring the time taken for a signal to reach a receiver. Almost one million times a second the satellite transmits a one or a zero in a complex string of digits that appears random. In actuality this code is not random and repeats every 266 days. The receiver knows that the portion of the signal received from the satellite matches exactly with a portion it generated a set number of seconds ago. When the receiver has determined this time, the distance to the satellite can be calculated using simple trigonometry where:

Distance to the satellite = speed x *(t_r - t_{to})* (where speed is c, the speed of light, in a vacuum (299792.5 x 10^3 ms^{-1}). t_{to} is the time at the origin and t_r is the time at the receiver).

The DoD maintains very accurate telemetry data on the satellites and their positions are known to a high level of precision. This simple operation allows the distance to a satellite to be calculated accurately. When the distance to three satellites is known then there is only one point at which the user can be standing. This principle is demonstrated in the diagrams on the following pages.

From one measurement we know the receiver can be anywhere at a uniform distance from the satellite with a radius equal to $r = c \times (t_r - t_{to})$. This defines the outer surface of a sphere of radius r.

Where:
r = radius
c = speed of light
t_{to} is the time at the origin
t_r is the time at the receiver

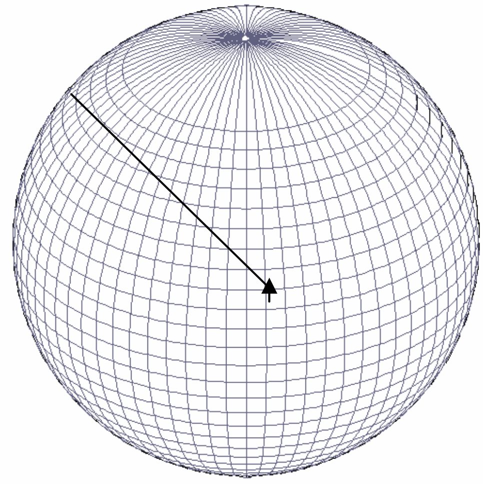


Figure 6-7 Basic Trigonometry - Single Satellite.

From two measurements we know the receiver must be anywhere on the line of the outer edge of a circle of intersection between the two spheres shown as a shaded ellipse below:

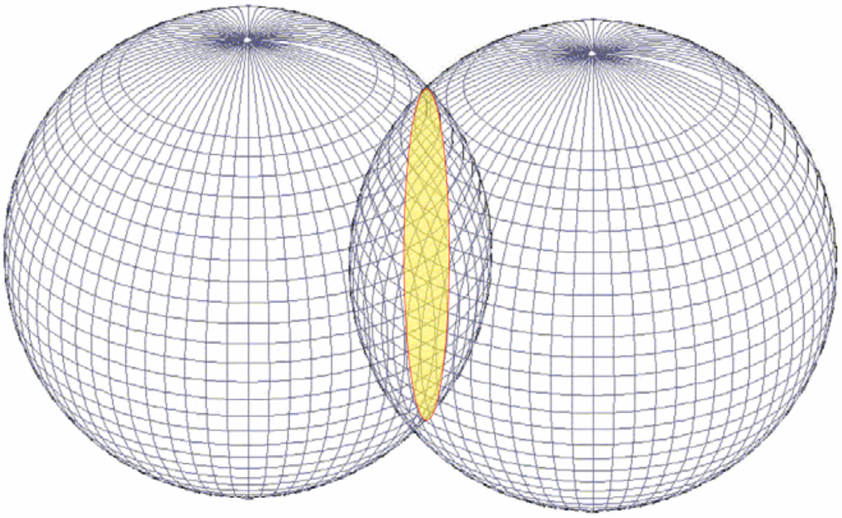
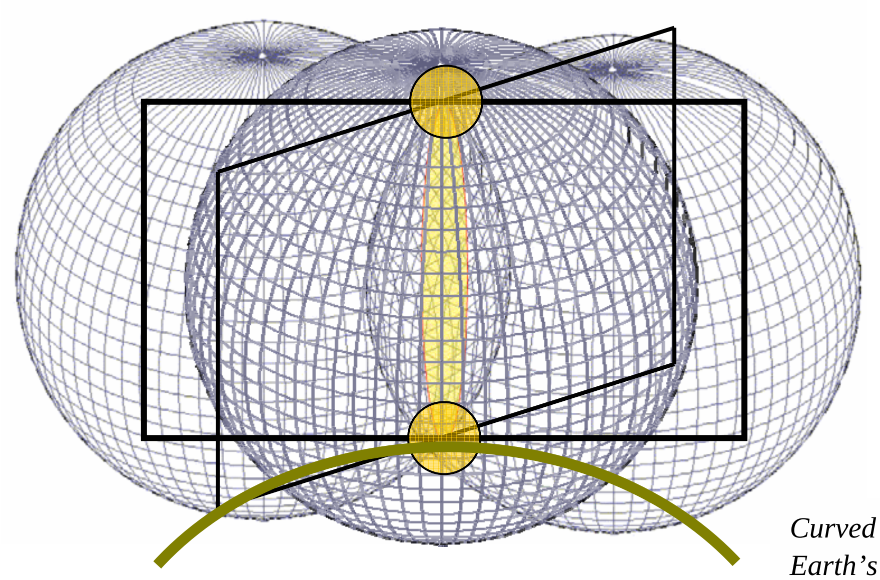


Figure 6-8 Basic Trigonometry - Two Satellites.

A third measurement reduces this to the intersection of a plane with the circle. This reduces the possible location to two points. Only one of these can be on the Earth's surface.



Of the two possible points of intersection only one point will lie on the Earth's surface

Curved hemisphere represents Earth's surface

Figure 6-9 Basic Trigonometry - Three Satellites.

Unfortunately, the above description is an oversimplification. This method of triangulation requires the receiver to know the precise time that the signal was transmitted and received. Even though time at the satellite (t_{to}) is known precisely because it is time stamped by the atomic clock on board the satellite, time at the receiver (t_r) is not known because this is generated by the internal receiver clock. To determine positional fixes to metre accuracy requires the GPS receiver to measure time accurately to 10^{-10} of a second. To keep the cost of GPS receivers below several thousand dollars per unit, atomic clocks are not used in the handsets. Due to these inaccuracies in timing the margins of error in calculated positions are very large. The way GPS receivers circumvent this problem is by using an additional measurement. The internal clock of the receiver will measure t_r incorrectly for all satellites. Therefore, because the offset is the same for all satellites, the receiver can use an additional satellite to bring all the points to one location.

The number of satellites a GPS receiver can talk to at one time affects the accuracy and the speed at which the system can function. 12 channel are the most commonly used receivers today, and are both quicker and more accurate than older models.

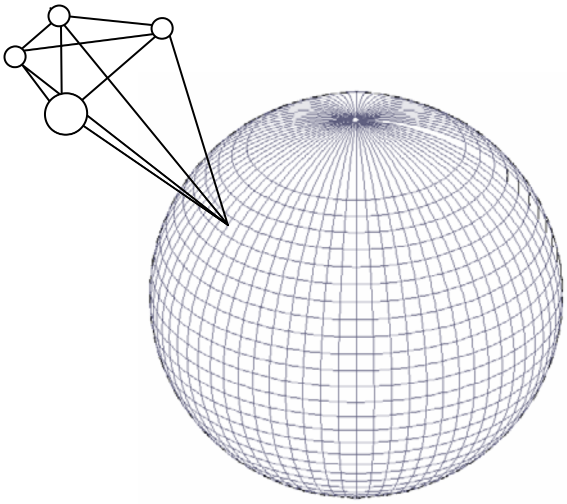


Figure 6-10 Basic Trigonometry - Four Satellites.

6.5 GPS accuracy

The signal transmitted by the satellites has a potential accuracy of <1 m but several factors influence this and reduce the actual resolution. The US military designed the end user of the SPS to be able to resolve a position 95.4% of the time (two standard deviations) to an accuracy of 100 m in X and Y (longitude and latitude) and 156 m in Z. Using the PPS service the end user should be able to resolve 22 m in X and Y and 27 m in Z. These are very conservative estimations and actual accuracy will lie between the theoretical resolution and these design schematics.

6.5.1 Factors affecting GPS accuracy

The reason why the actual locational position is significantly less accurate than the data transmitted by the satellite is due to various influences on the signal. These can be collectively termed *local* and *atmospheric* effects. Local effects are detrimental conditions on the ground near the receiver or in the receiver's software while atmospheric effects are problems with the medium through which the signal passes.

Table 6-1 Common factors that affect GPS accuracy.

Local Effects	**Atmospheric Effects**
Receiver Clock Error	Ionospheric Effects
Percentage Sky Visible	Tropospheric Effects
Satellite Geometry	
Multipath Error	
Ellipsoid	

6.5.2 Local effects

Receiver Clock Error: This is the error in the offset of the GPS measurement of the pseudo random code and the time recorded by the satellite for the data. The receiver attempts to compensate for this with additional measurements but it remains the single largest error that affects positional accuracy.

<u>Potential Error</u>: 0-10,000 m <u>Common Error:</u> 3-10 m

Percentage Sky Visible: This is of concern when getting an initial fix and generally causes the second largest error in calculated positions. It is linked to *satellite geometry* (below) and is a measure of how obscured the sky is. In areas where large parts of the sky are out of sight to the receiver, such as beneath a cliff or when surrounded by buildings, the error in the calculated position will be very large. This is also an issue in areas where the receiver antenna is beneath a thick forest canopy when the signal can be lost altogether.

<u>Potential Error</u>: 0-100 m <u>Common Error:</u> 5 m

Satellite Geometry: GPS receivers are only accurate when the quality of the data they receive is of a high standard. When the satellites being used for determining position are clustered together or all within one hemisphere the quality of the data will be poor. For accurate positions GPS receivers require satellite coverage from across the sky.

<u>Potential Error</u>: 0-20 m <u>Common Error:</u> 5 m

Multipath Error: When the receiver calculates the length of time the signal has taken to travel from the satellite to use in determining distance to the satellite it assumes the signal has taken the shortest path i.e., a geometric straight line. In actuality the signal may have bounced off a surface before reaching the receiver and the travel time could be slightly longer because of this. In these occasions the receiver will overestimate the distance to the satellite.

Potential Error: 0-10 m Common Error: <5 m

Ellipsoid: As discussed below, GPS receivers are designed to function in the WGS84 ellipsoid. Any other datum displayed by the receiver is a product of applying the Molodensky formula to the data. This gives a good approximation to the resultant datum but is not perfect.

Potential Error: ~5-10 m Common Error: 5 m

6.5.3 Atmospheric effects

Ionospheric Effects: All GPS signals travel through the charged plasma of the ionosphere. This can cause the signal to be attenuated (slowed down). Any changes in the signal involve changes in the travel time and thus affect calculated positions similar to multipath errors.

Potential Error: 2-30 m Common Error: 5-10 m

Tropospheric Effects: The water particles in the upper atmosphere cause very slight changes to the signal. These are very small but can affect minor changes.

Potential Error: 0-5 m Common Error: <2 m

Many of these errors can be quite easily compensated for and the section below will deal with the correct use of GPS receivers. Best practise with receivers involves using them in areas where their view of the sky is unobstructed, buildings or other corner reflectors are not present and that data is only recorded when the satellite geometry is of an acceptably high standard.

Table 6-2 Magnitude of errors in calculated GPS position.

Error Type	Compensation	Typical error (m)	Max. error (m)
Atmospheric / Ionospheric	WAAS or Differential	5	30
Receiver clock error	None (Differential)	5	10,000
Percentage full sky visible	Averaging	5*	100
Satellite geometry	Averaging	5*	20
Multipath	None (Covariance)	<5	10
Ellipsoid	Manual / use WGS84	~5	10
(Selective Availability)	*None Applicable*	*20-60 (23 RMS)*	*100*

** Percent sky visible will affect geometry and these two errors are not necessarily cumulative.*

The final error in Table 6-2, Selective Availability, has not been discussed and is a historical feature of only limited significance. When the GPS project was announced, as well as encrypting the PPS signal the US DoD applied a signal scrambling code to the SPS signal. This was to eliminate foreign powers using the signal to plan and orchestrate military attacks and to safeguard the US from precision attacks. Selective Availability was designed to ensure civilian grade GPS receivers were never more than accurate to 100 m. In May 2000, US President Bill Clinton signed a decree ending SA. However, the US DoD reserves the right to reactivate the system in times of war. This can be done over specific regions leaving other areas unaffected.

6.5.4 Real world accuracy

Most manufacturers quote receiver accuracy as <15 m. The total effect of the typical errors shown in Table 6-2 is closer to 30 m but repeated tests show that under good conditions the accuracy of a standard civilian set using the SPS signal on L1 should be considerably better than this.

The distance of any given point from the actual location is called the dilution of precision. Sometimes data is quoted in circular error probability (CEP). The CEP describes a circle of a radius containing 50% of the data. A typical GPS might have a CEP of 3 metres. More commonly the 2σ dilution of precision is quoted. The symbol σ is equal to the standard deviation of the data set. This is equal to the square root of the sum of the values of a data set minus the average value of the data squared divided by the number of points in the series. The 2σ dilution of a 12 channel GPS receiver is often assumed to be a circle with a radius of ± 7.5 m (a 15 m diameter circle around a point's true location).

To see the significance of the standard deviation of a dataset, imagine plotting all values in a dataset against the frequency with which each value occurs. Figure 6-11 gives an example, using the error found in many GPS readings taken at a point. In this case, the mean value is 0, and the other values are distributed to either side in a characteristic bell-shaped curve. This shape indicates a 'normal' data distribution (also called Gaussian); many statistical measures and tests assume that data are normally distributed in this way. In the case of the standard deviation, if data are distributed normally, then we can say that:

- 68.2% of all values lie within $\pm 1\ \sigma$ of the mean
- 95.4% of all values lie within $\pm 2\ \sigma$ of the mean
- 99.7% of all values lie within $\pm 3\ \sigma$ of the mean

So, for example, when GPS are described as having 7.5 m accuracy to 2σ this means that 95.4% of readings are within error margin.

In many cases data will not be distributed normally. For example, if the data in the distribution is skewed to one side, has more than one peak, or if the number of values is relatively small then the curve may not be Gaussian. In such cases, the standard deviation must be used more conservatively, e.g. 2σ contain 75% of the data, rather than 95.4%, or may not be applicable at all. The nature of the distribution can be assessed graphically, as in Figure 6-11, although statistics text books give more rigorous tests.

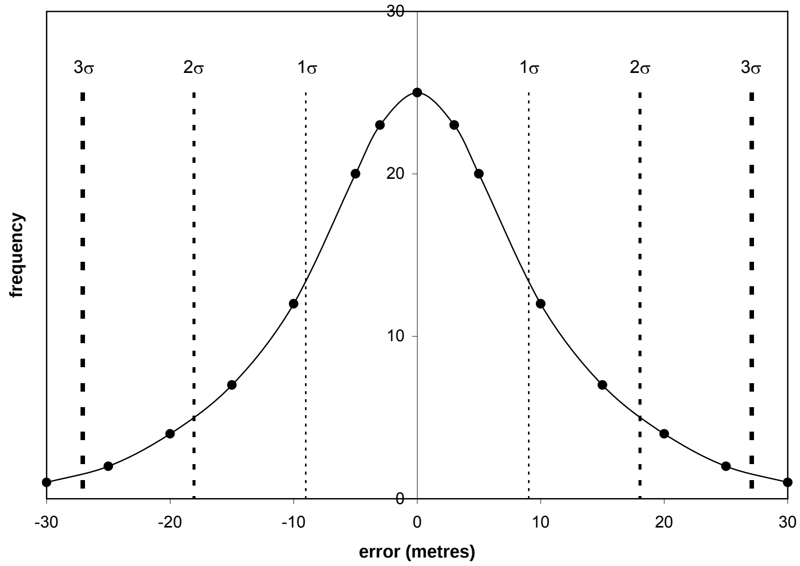
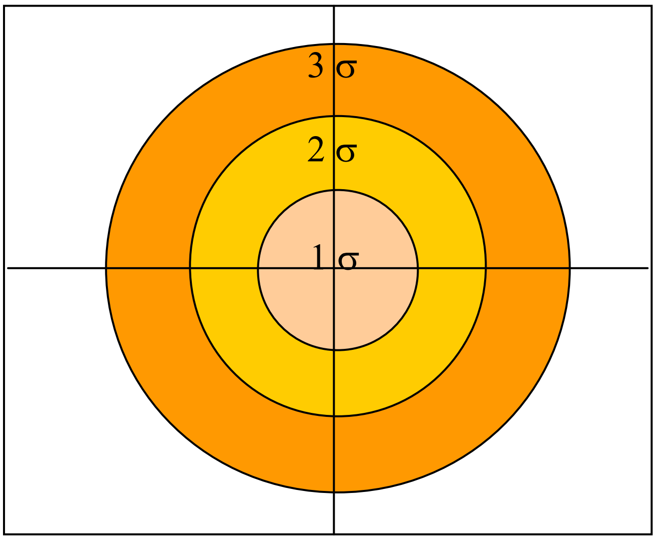


Figure 6-11 Distribution of errors in a set of GPS readings taken at one point. The line drawn through the points has a bell-shape, indicating a Gaussian or 'normal' distribution of the data. The vertical bars show the number of GPS readings lying within different standard deviations of the mean: 1σ (68.2% of all values), 2σ (95.4%) and 3σ (99.7%).

The scatter distribution of a GPS receiver is shown below in Figure 6-12.



The circles in the diagram are concentrically centred about the intersection of two lines. If this intersection is the actual location of a point, then a Gaussian distribution says that:

68.2% of values lie within the inner circle

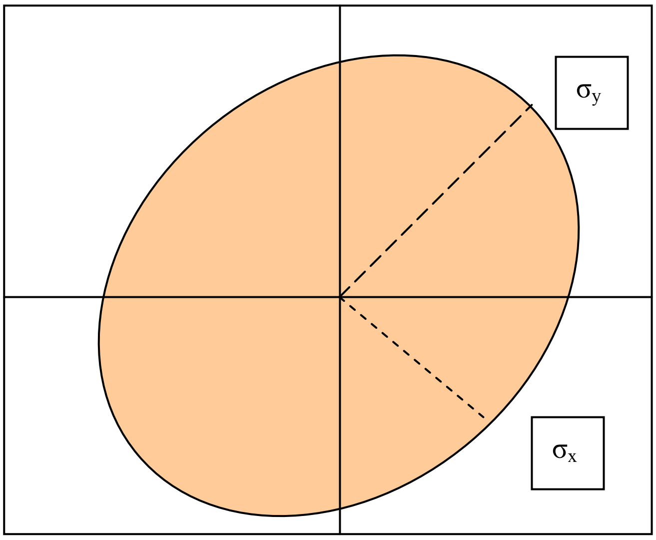
95.4% of values lie within the middle circle

99.7% of values lie within the outer circle

The size of these circles allows a determination of GPS accuracy. The design of the Standard Positioning System was for the 2σ circle to have a diameter of 100m but in reality the system is significantly better.

Figure 6-12 Error circles for SPS signals.

A 2σ circle means 95.4% of the readings are in the circle as shown above. Extensive tests on GPS receivers show that the data actually scatters with a slightly elongated shape. This 'error ellipse' is normally orientated with the semi-major-axis directed northeast - southwest with the following dimensions:



Error ellipses are different for the different types of GPS receivers. The statistics below are for generic receivers for standard deviations 1-3 (2σ - 6σ)

Twelve Channel Sets (%, 2σx, 2σy)

68.2% 3 m / 4 m

95.4% 6 m / 8 m

99.7% 9 metres /12 metres

Figure 6-13 Error ellipse for 12 channel GPS.

Standard deviation can be shown in a clearer format using either of the following equations:

$$\sigma = \sqrt{\frac{\sum (\mathrm{x} - \bar{\mathrm{x}})^2}{n}}$$

However,

Standard Deviation is more commonly expressed in the form :

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

For GPS analysis where the data has a positional nature, standard distance is commonly used instead of standard deviation. The equation looks the same but is a description of the spread of points around the mean centre. In standard distance, *x* is the *x* co-ordinate (*longitude*) of any individual point and *x*-bar is the mean centre of the distribution. An identical equation for *y* (latitude) is also required.

$$\text{Standard Distance} = \sqrt{\frac{\sum (x - \bar{x})^2}{\mathrm{n}}} \text{ where } x \text{ is x co-ordinate or longitude.}$$

Elevation is a more complicated variable because it requires more satellites and because the ellipsoids are more difficult to calculate. As such elevation data will be considered separately later on but is normally quoted with a dilution of precision 1.5 to 2 times that of the *x, y* value.

6.6 Correct GPS handling

As discussed above GPS receivers are only accurate when used correctly. Improper use or failure to consider environmental factors such as canopy cover or urbanisation will result in severely degraded data. This section discusses how the GPS should be used in the field and important considerations when using them on expeditions.

To calculate its position, a receiver needs to know which satellites it is talking to and where they are. The time taken to do this can vary considerably and is an important consideration for fieldwork. The time taken from switching the receiver on to the time of the first fix is referred to as start-up time. This start-up time is controlled by the capability of the receiver (how many satellites it can ‘talk’ to, commonly 8 or 12) and the accuracy of the GPS almanac. The GPS almanac is a digital record kept by the receiver of where satellites are in the sky and what satellites it should be receiving data from. It is essential the receiver knows the satellites it is using in order for it to accurately calculate its position. Determining this from fresh takes a significant length of time (up to 12 minutes for a complete download from a satellite), so the GPS keeps a record of the almanac data it has previously collected to speed up the start-up time. Almanac data is not very precise; it is accurate for geographic regions up to ~300 km diameter and for up to two months. If the receiver moves further than a hundred kilometres from the last time it was operational or if it is left inactive for a period of longer than two months then the almanac will not be valid. In these cases the receiver will have to determine this data from fresh. This is known as a ‘cold start’ and can take up to 12 minutes. If the Almanac is accurate the GPS receiver can initialise a ‘warm start’.

The GPS receiver collects data for its almanac from information transmitted by the satellites. Each satellite transmits its own locational data. This data is referred to as ephemeris data and is very precise. The ephemeris information is broadcast every 30 seconds but it can take up to 24 transmissions to completely describe the orbit in detail for the GPS. The difference between warm and cold starts is therefore quite significant. Before a GPS can use a satellite, it must have a complete ‘packet’ of ephemeris information. Any glitches while acquiring this information will cause the GPS receiver to start over again for that satellite. This means a satellite that is in a difficult position or affected by multipath or similar errors may take a very long time to send an uninterrupted stream of information for the receiver to use. This is why some satellites remain inaccessible even though the GPS can see that they are in the sky. The time taken for modern sets to acquire a lock is usually considerably less than for older models. These time differences are summarised in Table 6-3.

Table 6-3 Realistic acquisition times.

Receiver	**Almanac is correct (‘warm start’)**	**Almanac is incorrect (‘cold start’)**
8 channel receiver	6 minutes	12 minutes
12 channel receiver	1 minute	4 minutes

After acquisition, the data from the receiver will fluctuate for a period of time. This fluctuation will generally continue for 5 minutes after the GPS is first activated and will then settle to give a better fix more inline with the data quoted in Section 6.5.2. The satellites being tracked, and the quality of information coming from them is commonly shown as a sky view. This view is important in determining both the quality of data and whether the receiver has acquired a satellite lock.

When a GPS receiver has a satellite fix it can be used in one of two ways. Location points can be recorded at a user's discretion by clicking a button on the receiver (usually Mark or similar) or automatically at given time or distance intervals. These two methods are referred to as Waypoints and Trackpoints respectively.

Trackpoints: The behaviour of a receiver when recording a trackpoint is different for each model. Typically the receiver will store a co-ordinate value and location ID either at time intervals or when it detects a significant change in user direction. This option is often configurable on the more expensive receivers and this is a useful option because it avoids filling the memory with unnecessary points. Less commonly the receiver will record altitude with the trackpoints. The receiver will often not quote how many points it can store but with increasing memory capabilities over 5000 points is not uncommon. Once this limit has been reached the receiver will begin to write over its initial track (often with no warning). To avoid this, the receiver can store the track in a more stable form as a tracklog. Most receivers can store around 10 tracks as tracklogs but these are simplified descriptions of the original data used to save memory. They are not as accurate as the original data and may only use 30 points from the original ~1000 points. Modern Magellan models such as the Sportrack and Meridian ranges overcome this and have a detailed tracklog option. This can be used to create a fully detailed tracklog referred to as a 'backtrack'. The major disadvantage with a trackpoint is that the receiver will take a reading regardless of the satellite constellation at the time. If the receiver is set up to take a reading every minute and the receiver is moved under dense canopy for a period of time, then the readings it records may be very inaccurate (>50 m). Trackpoints can be useful to a team because they allow large amounts of data to be collected very quickly. If the team is visiting an area where the road network is poorly mapped or not known, then by driving the roads for a day the GPS can accurately map the road network. Even with the problems associated with trackpoints such as taking readings in non-ideal conditions, the map would still be c. 1:5,000 to 1:10,000 scale.

Waypoints: Waypoints are recordings of a location's co-ordinates, commonly with user descriptions and elevation data. Waypoints can usually use an associated symbol for displaying on map views. This is sometimes downloaded as label that can be re-associated with a graphic at a later time. Waypoints can be given a text string to accompany them of between 8 and 12 characters that can be used to annotate the waypoint.

Waypoints are only recorded with user interaction and are not automatic. This is a severe disadvantage, as user interaction with the unit may be difficult due to local conditions or difficult terrain. This often means relying on trackpoints is safer and easier. User interaction does, however, mean that only accurate positions are recorded. The major disadvantage of trackpoints is that they record data whether or not the GPS constellation is good. Waypoints will always be of a higher standard. Waypoints have many other advantages and disadvantages. Receivers can commonly store 500 waypoints or more, with newer sets with greater memory capabilities becoming available. However, the storage capacity of waypoints will only be 1/10 that of the trackpoint storage. The collected data can be used individually or can be collated into a route. Routes are selected waypoints that define a path to follow. Around 50 waypoints can be selectively chosen to add to a route and the GPS can instruct the user on how to travel between each one with direction and distance supplied.

It is important to be able to associate a waypoint location with any field description. This should involve noting the waypoint in the field log, however, the GPS receiver can also be used to take notes. Though the waypoint description in a GPS is often limited to 6 (ETREX) 8 (Magellan 310) or 12 characters, new Magellan models such as the Sportrack Pro offer the ability to add 'messages' to the waypoint locations. The Sportrack can store 500 waypoints at an 8 character limit but 204 of these can have an associated 30 character message.

Trackpoints and waypoints have distinct advantages and disadvantages. The most significant difference is that trackpoints are automated whereas waypoints require user interaction with the set. User interaction can be difficult in certain terrains where handling a set can be hazardous. However, user interaction means that only data of a high standard is acquired and the overall data set is of a high quality. Trackpoints record points automatically and give no consideration to the quality of the data. This means that points can be recorded at times of poor satellite geometry. When studying these points later, they may have a very large error but because this error is unknown interpretations may be made based on data points that are not correct.

Table 6-4 Relative advantages and disadvantages of waypoints and trackpoints.

	TrackPoints	**WayPoints**
Advantages	No interaction with the set is necessary Large numbers can be stored to create accurate routes.	Waypoints can be recorded when accuracy is high. Additional data such as a name and comment can be put with the point.
Disadvantages	Although you can set a time or distance interval for the tracklog it automatically records position regardless of satellite geometry and visibility.	The set must be handled repeatedly. Only a small number of points are likely to be recorded making routes less detailed.

As with any part of expedition kit GPS receivers need to be treated with respect and due care. GPS should never be the only form of navigation tool taken into the field but if treated correctly they can be very reliable complimentary tools. GPS receivers require battery power and this is the most important factor that needs to be considered in remote areas. In addition, most GPS receivers have LCD screens that are susceptible to cold conditions. Temperatures below 0°C may irreparably damage a receiver's screen. How to combat these problems is discussed in the Field Technologies chapter. A GPS should never be left exposed at night during cold conditions and should always be packed into a rucksack or similar, to protect its screen.

6.7 Assessing data quality

GPS is a valuable tool for expeditions but the ease with which buttons can be clicked and data collected can lead to poor scientific practise. The team should always be conscious of what the GPS is recording and whether the data is of a high or low standard.

When using waypoints, the user has the opportunity to select readings of a high quality. The quality of data can be very important to an expedition. If a position is to be recorded in the field for future visits then poor quality data may make the location ambiguous. The difference between a good fix (±5 m) and a poor fix (±20 m) may place the waypoint on the wrong side of a river or crevasse making relocating the point or making interpretations very difficult. The GPS receiver will show the quality of data within the sky view as an estimated positional error (EPE).

The EPE is based on the satellite geometry and should not be considered an accurate gauge of the actual error. The EPE is a good indication of data quality and care should be taken to record waypoints only when it is low and data is of a high standard. Ideally, the reported position and its associated estimated error should also be recorded in a notebook or handheld PDA computer as a backup precaution. Figure 6-14 below shows how the geometry of the satellites affects the EPE. As skilled users of GPS the team members should be able to approximate the EPE just from the satellite geometry.

The receiver display to the left has very poor satellite constellation geometry and correspondingly, a very high estimated positional error.

Even though the receiver on the right has fewer satellites, they are positioned better and therefore the error would be correspondingly lower.

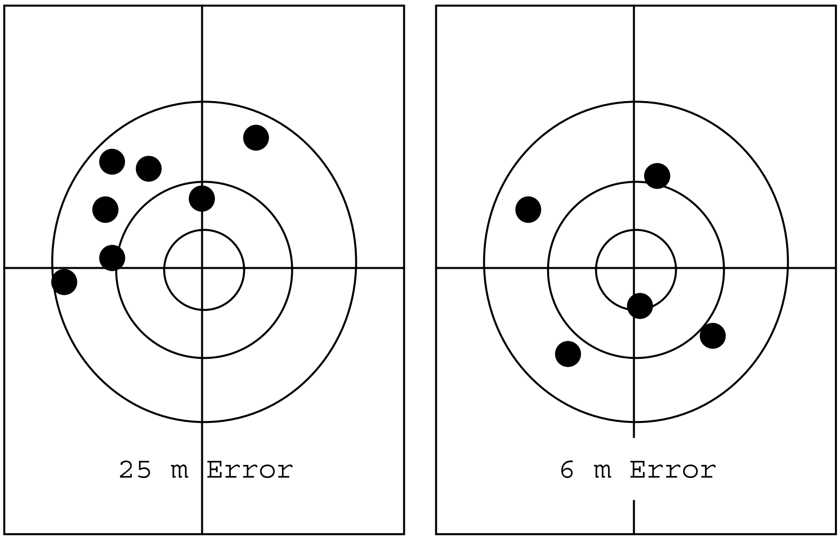


Figure 6-14 The difference between good and bad constellations.

When using GPS it is important to check that the receiver is giving the data that is required. The receiver has to be set to the correct projection and datum and it is always important to be conscious of the data it is producing. The nature of the GPS signal means that it is possible for any one reading to be wrong. As can be seen in Plate 4, some readings can be a long way out due to the nature of the errors associated with GPS. As a user in the field it is important to be aware of the previous data to ensure that a reading does not diverge too much from the expected value. Obviously, detecting small fluctuations is very difficult but it can be a good warning that something has gone wrong with your settings or the GPS system as a whole.

On 1 January 2004 one of the older satellites PRN 23 experienced a system failure error. Most satellites operate with triple redundancy to mitigate the problem of a system failure but some of the older satellites are now operating without redundancy. This lack of redundancy means that errors can occur in the system. As the GPS IIR satellites replace them this will not be an issue because they are both newer and they are configured to use their back up systems as a real time check on the satellite status. This should mean problem satellites are taken off line much more quickly but it is always worth checking readings to make sure they are consistent and that there is no sudden change in the calculated position. In the case of PRN 23 the error was around 200-300 m in Dover Straits and up to 40 km in Scotland. The problem was mitigated within three hours by setting the satellite to ‘unhealthy’ but any results collected at such a time would be useless for scientific work. The satellite was brought back on line by 20 January 2004 running on its backup caesium clock. The SPS and PPS performance standards require ≤ 3 failures per year. This incident was the first major failure since 2001, so for most purposes the GPS signals are incredibly reliable. More detailed reports can be found at the US National Coast Guard Navigation Centre *www.navcen.uscg.gov*.

The receiver should always be used in accordance with the instructions here and in the receiver manual to ensure the best quality of data. The receiver should be held away from the body with the antenna (either in the head of the unit or in the aerial) held towards the sky. This is shown below in Figure 6-15.

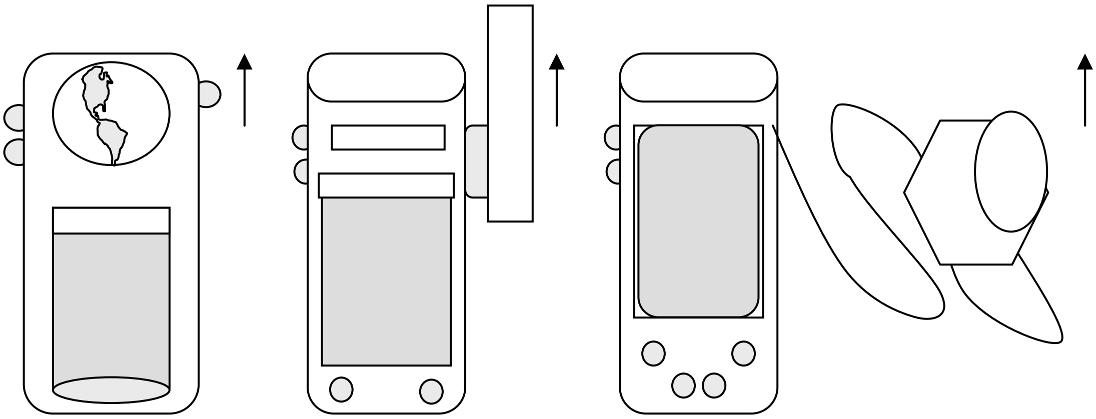


Figure 6-15 Location of GPS antennas. All GPS antennas should be directed towards the sky even when housed internally. Far left GPS is typical of modern units that house the antenna internally in the device head. Middle and right devices have different forms of external antenna.

While discussing antennas and their positioning, it is worthwhile mentioning the specific type of antenna in the unit. There are essentially two different types of antenna used in common GPS models. These are the patch antenna and the quad helix antenna. Though there are a number of reports and publications stressing the advantages of quad helix over patch in areas of weak signal strength, the actual differences are negligible. Many newer receivers such as the Magellan Meridian and Garmin 76 use quad helix as opposed to the older Garmins and Magellans that use a patch antenna. There is a slight performance gain with these newer units under canopy cover but not enough in itself to warrant purchasing new units. Attaching an external antenna is a better solution than using either forms of internal antenna.

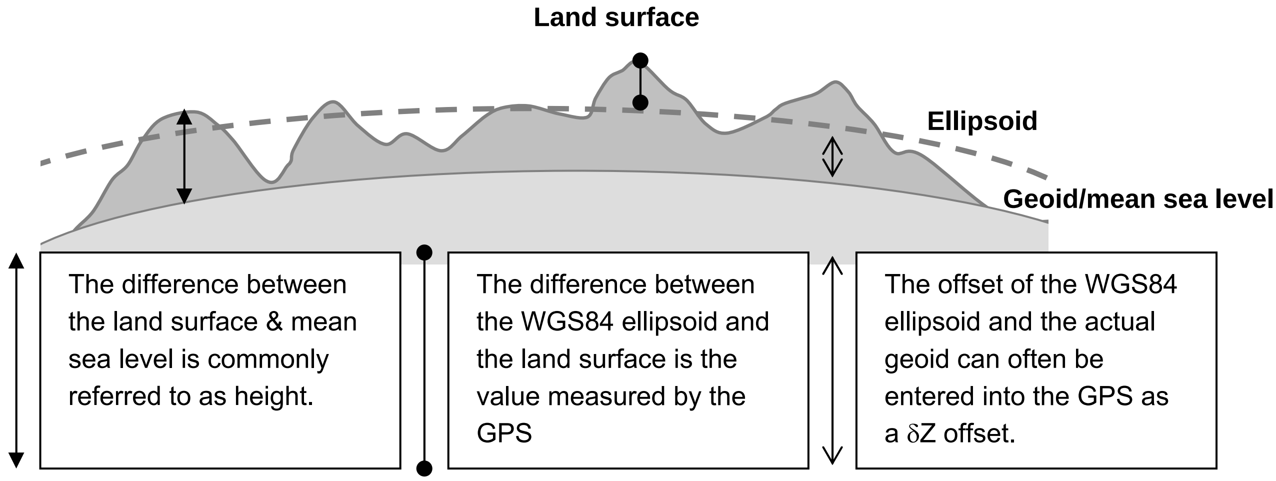
The most sensitive receivers available use the SIRF Star III hardware. This is a combination of sensitive patch antenna combined with additional signal processing firmware located inside the GPS. The SIRF Star III can lock onto 20 satellites at one time and can reprocess signals usually discarded by receivers as being too noisy or too weak. SIRF Star III can reprocess these and can work indoors. The receiver has only just come onto the market (mid 2005) and the practical ‘real-world’ results with the set are still not fully known. The improved sensitivity is believed to offer improved accuracy. Whereas normal 12 channel sets usually have an error of around 15 m, 20 channel Star III sets claim an accuracy of between 5 – 25 cm CEP.

Most importantly when assessing data quality is to make sure the GPS is being used in an appropriate manner. Even the most sensitive GPS will record poor data when used incorrectly. The constellation of satellites should also be checked as described above to make sure the data is being collected at the appropriate time. Using a GPS blindly without any appreciation for the values recorded is poor practice and should be avoided in all but extreme cases.

6.8 How GPS calculates and stores positional data

When assessing data quality it is important to understand how the GPS is arriving at its results. The NAVSTAR system was designed to work with a mathematical model of the Earth. This model was the shape of an ellipsoid (a three dimensional ellipse flattened at the poles and elongated at the equator) and the GPS calculates positions relative to this model. The GRS80 (Geodetic Reference System 1980) ellipsoid is the model that was taken for use with NAVSTAR. This is a good approximation of the shape of the Earth. It was modified slightly to be the World Geodetic Survey 1984 (WGS84) and this is the common reference system that all receivers use. This best fit of the whole Earth is not accurate for all areas and most countries use their own ellipsoid. In Britain the ellipsoid used is the Airy 1830 ellipsoid. When displaying data the GPS will by default display co-ordinate information according to the WGS84 ellipsoid. This is acceptable for latitude and longitude but would often give unexpected and erroneous height data. The height of topography above or below a hypothetical ellipsoid is often of limited use and a more conventional description is often required. Height is most commonly expressed as elevation above mean sea level. Mean sea level can be approximated by a geoid. As described in Chapter 2 a geoid is a model of the earth defined as a surface where the lines of gravitational force are perpendicular. Mean sea level itself is not constant across the globe; it can alter by as much as 2 m, depending on where it is measured. Heights in Britain are measured relative to the tide gauge at Newlyn, Cornwall. Even the best geoid available will still not tally with every country's maps and heights, because there is often a discrepancy in the zero altitude used. The geoid and mean sea level are commonly very close to one another, so heights against the geoid are an acceptable measure of heights against mean sea level. These two measurements do not diverge by large amounts (up to a maximum of 2 m but most commonly sub-metre) though they can both diverge by up to +85 m or –102 m against the WGS84 ellipsoid.

The GPS receiver always measures heights relative to the WGS84 ellipsoid. It can apply a transformation to get to a better fitting ellipsoid for a given area, but always works internally to WGS84. When working outside of WGS84 the GPS needs to know the conversion factors defined by the variables δX, δY and δZ. These will be built in for most typical datums so all the user needs to do is select the correct setting and the map and GPS will correlate almost perfectly. Section 11.4 describes how this can be done when mapping in an area with maps that are not compatible with the GPS. These can be used for calculating better approximations of the height at a given area. Some models also allow correct height to be input at start up to aid in calculating elevation.



The difference between the land surface & mean sea level is commonly referred to as height.

The difference between the WGS84 ellipsoid and the land surface is the value measured by the GPS

The offset of the WGS84 ellipsoid and the actual geoid can often be entered into the GPS as a δZ offset.

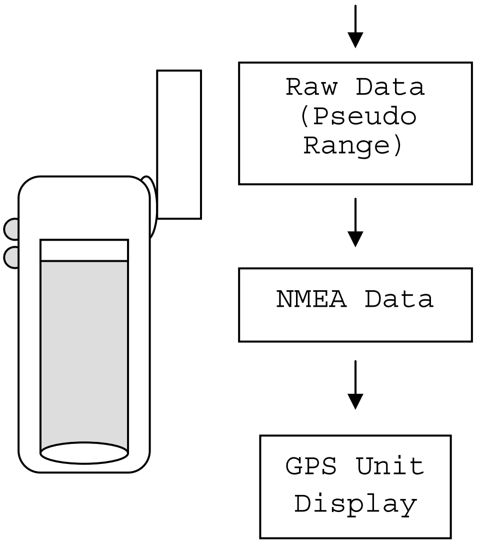
Figure 6-16 Offsets between geoid, ellipsoidal and land surface.

The height above mean sea level is called orthometric height. To calculate orthometric height the GPS requires an elevation above the WGS84 ellipsoid and a knowledge of how much the WGS84 ellipsoid and the geoid differ at that location. Unlike the WGS84 ellipsoid, which is a comparatively simple shape, the WGS84 geoid is very complicated requiring many megabytes of data to store it. This is beyond most GPS receivers so a crude approximation of the WGS84 geoid is included with spot heights at a number of locations in a look-up-table. This method is how all modern GPS receivers' measure height. Though height is described as metres above sea level, it is in fact metres above the low-resolution geoid look-up-table that approximates sea level. A true height above sea-level is therefore very difficult for the GPS receiver to calculate. To gain accurate height the receiver needs the standard constellation of satellites (usually 4) plus an additional satellite to fix elevation. Commonly, this means a good view of five satellites. Because of this, the error in elevation is generally greater than the error in X or Y by at least 1.5-2 times. Realistically manufacturers quote an error of at least 25 m in Z. The conversion from the WGS84 ellipsoid to the WGS84 geoid downgrades the accuracy of the height, but this is generally of limited concern because the inherent errors in generating height are so large.

6.9 NMEA sentences and stored information

NMEA is a standard format that can be used to download information from a GPS unit in real-time (NMEA = National Marine Electronics Association, which defined the standard). Almost all receivers generate NMEA information called sentences. These are a middle layer of information that is less processed than the onscreen data but more refined than the internal GPS calculations. This hierarchy of layers is shown schematically in Figure 6-17.

Processing levels in GPS output



Lowest processed layer: raw data. Initial calculations based on data received from the GPS satellites. These data cannot be downloaded easily. A GPS has to be forced to release this code by sending a correct data string into the receiver.

Products such as Async logger can help with this.

Second Level NMEA Data, these data have been processed to assess the strengths of signals from satellites and to calculate positions relative to the WGS84 ellipsoid.

Figure 6-17 Processing tiers in the GPS positioning architecture.

If a GPS uses NMEA, a sentence is available that shows the current ellipsoid and how it differs from the WGS84 ellipsoid. This can be of vital importance if elevation is required as seen in Figure 6-16. Downloading NMEA data requires a computer to be connected to the receiver. The data streams as ASCII information at a standard bit rate as discussed below. Each NMEA sentence is prefixed with a $ symbol and a 4-5 letter code denoting the information contained within the sentence and a comma separates each value. If a value is unavailable it is left blank and a comma prefixes the next variable. NMEA data is very valuable because it can be used as a step back from the standard GPS interface as shown in the figure over the page. In standard WGS-84 mode, the corrections and transformations the unit would normally do to display the data have not been done. The data is therefore provided in a ‘raw’ or ‘native’ state.

Downloading NMEA data is comparatively easy; it requires the GPS to be connected to a computer with its interface set to NMEA data. The following example shows how to do this using a typical Garmin ETREX. On the ETREX models go to the Menu page ⇒ Setup ⇒ Interface set the input output (I/O) type to NMEA and the baud rate to 4800 (4800 is the default NMEA standard). A software interface with the RS232 connection must then be established on the computer. The computer interface must be set to recognise 8 bit data at the correct baud rate (4800), with no parity and 1 byte checksum stop values. This can be achieved comparatively easily on PCs using the MSComm control. More recently MSComm.ocx, an ActiveX component, has become available and can be easily distributed amongst PCs for this purpose.

Most GPS manufacturers add additional sentences to the NMEA output. These are prefixed differently. Garmin use a ‘P’ for Prefix. These sentences give additional information applicable to the manufacturer’s model. One of the most important NMEA sentences is the GGA sentence. This sentence is examined in Figure 6-18.

One of the most important sentences is GGA (sometimes) GPGGA. This sentence describes the X,Y and Z co-ordinates of the calculated fix. It is also very useful because it shows the difference in the height values. An example sentence is shown below.

Sentence Structure
Time
$GPGGA,**172704**,5126.7759,N,00020.6673,W,1,05,1.3,13.8,M,47.4,M,,*67
Northing
$GPGGA,172704,**5126.7759,N**,00020.6673,W,1,05,1.3,13.8,M,47.4,M,,*67
Easting
$GPGGA,172704,5126.7759,N,**00020.6673,W**,1,05,1.3,13.8,M,47.4,M,,*67
Fix Type
$GPGGA,172704,5126.7759,N,00020.6673,W,**1**,05,1.3,13.8,M,47.4,M,,*67
Number of SVs tracked
$GPGGA,172704,5126.7759,N,00020.6673,W,1,**05**,1.3,13.8,M,47.4,M,,*67
Approximate dilution of precision (horizontal)
$GPGGA,172704,5126.7759,N,00020.6673,W,1,05,**1.3**,13.8,M,47.4,M,,*67
Height above sea level (geoid)
$GPGGA,172704,5126.7759,N,00020.6673,W,1,05,1.3,13.8,M,**47.4**,M,,*67
Orthometric correction
$GPGGA,172704,5126.7759,N,00020.6673,W,1,05,1.3,13.8,M,47.4,M,,***67**
End of sentence checksum.

Figure 6-18 Analysis of an NMEA sentence.

The data in the sentence in Figure 6-18 shows the information output from the GPS in NMEA mode during a test using OSGB, British Grid. Changing the projection and co-ordinate system has no effect on the output data, as can be seen in the example below, using UTM degrees, minutes and seconds:

British Grid: $GPGGA,172704,5126.7759,N,00020.6673,W,1,05,1.3,13.8,M,47.4,M,,*67

OSGB: $GPGGA,172706,5126.7759,N,00020.6673,W,1,05,1.3,13.9,M,47.4,M,,*64

The GPS makes all calculations relative to the WGS84 ellipsoid, but stores and outputs height relative to the mean sea level, as calculated by the receiver. NMEA data is vital for conducting user-defined transformations and for ascertaining accurate height information from the GPS. NMEA data records both the orthometric and ellipsoid height information. By taking the raw ellipsoidal height and correcting it with a known MSL-Ellipsoid offset (this can be found on any good map) yields better height data. Though NMEA data is very useful, it cannot be used alone to give cheap differential corrections. Post processing NMEA is not possible because the data from each SV is not stored in the sentence. For individual SV data, a pseudo range output is required. Pseudo range is the name given to the pre-processed data and is discussed more fully in Section 6.10.2.

NMEA data can be very useful for making some manual corrections to a GPS. Though GPS receivers are capable of applying an equation to warp their results to any system co-ordinate system, they do this by using the Molodensky formula. This technique is a simplified three co-ordinate shift that has at best an accuracy of 5 m. The current advice from the Hydrographic Office and the Admiralty (*Admiralty List of Radio Signals Vol. 8*) is to keep GPS receivers in WGS84 mode and make any shifts manually. This is significant for height data, as the only way an accurate height measurement can be obtained is by stripping out NMEA data to a PC and applying manual calculations. Accurate height calculations require a measurement against the WGS84 ellipsoid and then a calculation of offset between the ellipsoid mean sea level. It must be noted that the output NMEA data are relative to the geoid model selected in the receiver. A GPS receiver will always store data in WGS84 but the NMEA output will have a co-ordinate transformation. As such, the receiver should always be set back to WGS84 before NMEA data logging. If a simple and convenient method of communicating with a GPS is required without the need for complex data filtering, then the communications programme HyperTerminal, supplied as part of the Microsoft Windows operating system, should suffice (see Appendix 2).

6.10 Understanding precision and improving accuracy

GPS accuracy is better than 15 m (around 6 m by 8 m). GPS receivers are often called upon to report data more precisely than this. The UTM system is a measurement of metres from a datum in 1 m intervals. The common degrees minutes and seconds system is also often more accurate than the resolution of the GPS receiver. The Earth is approximately 6,378,200 m in radius at the equator giving an equatorial circumference of 40,075,413 m. If this is divided into degrees minutes and seconds each degree is 111,320.6 m, each minute is 1,855.34 m and each second is 30.92 m. When divided decimally each fraction of a second is 3.1 m. This means that the data is recorded to 3 m intervals, which is still more accurate than the GPS is actually capable of. This level of precision (how many decimal points quoted) is often incorrect and caution should be taken when using seemingly accurate information.

6.10.1 Improving accuracy (standard methods)

There are a number of techniques for reducing dilution and improving data quality. The signals sent from the NAVSTAR satellites are accurate to the centimetre scale. This accuracy is downgraded by the various factors discussed in Section 6.5.1. In October 2001 the US Military released their first indication of the post-SA quality of the data. Their findings indicate that though the signal is now accurate to at least 13 m, local errors will often reduce this. If a GPS receiver can be used in an area of unobstructed sky the position should be as accurate as the diagrams shown in Section 6.5.4. If a reading has to be taken in an area of obscured sky, then a GPS can be combined with more traditional surveying techniques. If a point of clear sky can be found, then the expedition can mark this point using the GPS as normal. The co-ordinates of the desired point can then be calculated by surveying a line back from the known point to the required location. To do this the expedition would have to use a compass to measure the angle from the known point to the desired location then accurately measure the distance between them. The exact method of transforming this information to the GIS will depend on the co-ordinate system used. This process is shown in Plate 5.

Combining basic surveying and GPS work can be an effective method for surveying points and entering them back in the GIS later. It is, however, a very time consuming method and the accuracy reported on the GPS should be considered to see if the additional work is required.

As well as using the GPS in areas of unobstructed sky, there are additional methods for improving the SPS accuracy without the need for the PPS information. The most common methods for improving data are listed below and discussed in turn in the following section.

Methods for Improving GPS (Standard Positioning Service – SPS)	
Averaging	A simple method using standard receivers
Differential Corrections	A more involved process using specialised equipment
WAAS	A cheaper more accessible version of differential
Carrier Wave	A more complex version of differential signals

Averaging: When a GPS records a location it will lie a certain distance from the true location. This 'dilution of precision' is random but multiple readings will plot within concentric ellipses forming Gaussian distributions along each axis as shown above in Section 6.5.4. The data plotted in Plate 4 shows real world examples of this. In the left hand plot there are outliers of information shaded dark red with inner lighter areas. The dark areas show low concentrations of data and the light areas show high concentrations. Because this data forms a Gaussian distribution it is probable that the actual location lies closer to one of these light areas. This is shown on the plot on the right. The error in the GPS data is plotted along the abscissas and the frequency of this measurement is shown on the ordinate. The vertical line shows zero error. It is clear that the most frequently reported co-ordinate is very close to the line of zero error.

The most commonly occurring value in a distribution is referred to as the modal. By taking the modal value of the distribution, a point's most likely true location should be significantly more accurate. This requires a substantial amount of data to be collected at each point for a significant statistical analysis to take place. Table 6-5 compares an averaged GPS reading, computed from taking readings every 30 seconds for 10 minutes after successfully acquiring 4 satellites, to a standard single recorded waypoint value.

Table 6-5 Effects of averaging on GPS accuracy.

GPS Type	**Single Plot Error in X**	**Single Plot Error in Y**
8 Channel	**20 m**	**40 m**
12 Channel	**6 m**	**8 m**
GPS Type	**Averaged Error in X**	**Averaged Error in Y**
8 Channel	3 m	14 m
12 Channel	3 m	4 m

Averaging improves the accuracy considerably but requires more time to be spent at a site to record data than standard single point data collections.

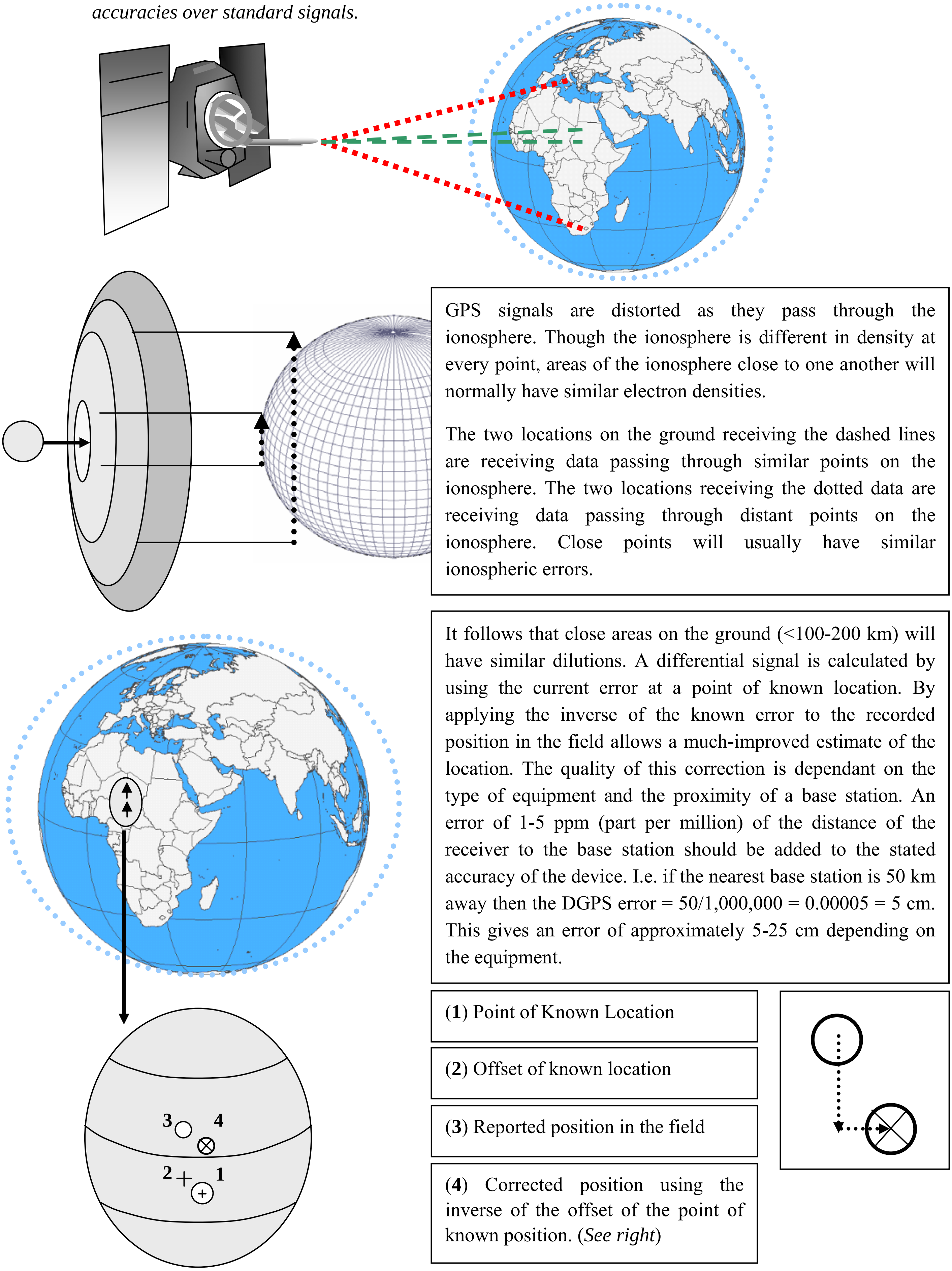
Differential GPS (DGPS): The signal from a satellite passes through a section of the ionosphere that attenuates it and introduces an error of between 2 and 30 m. The actual effect of the atmosphere is to increase the travel time of the signal. For receivers located in geographically small areas, the signal to each receiver will have passed through essentially the same part of the ionosphere and be attenuated in the same way. If a point whose location is known very accurately can be used for a GPS base-station, then the travel time to any satellite should be known. Any deviations from this will be because of errors generated by the ionosphere. These errors can then be transmitted to properly equipped GPS receivers in the local area and these errors can be removed.

The station records all the travel delays for all the satellites and re-transmits them. The receiver in the field can then make corrections for the delay. This reduces the GPS error to between <1-5 m, depending on the sophistication of the hardware used. Differential corrections can be conducted in real-time, i.e. the correction signals from the station are used by the receiver in determining its location. Alternatively the receiver data can be post processed. This means logging all the GPS data then downloading the timing errors from a station and applying corrections later.

The final method sometimes cited is inverted differential where the receiver takes a standard reading and transmits it to the station that corrects the data without sending information back to the receiver. Many standard GPS units are differential ready but adding the additional hardware is more expensive than using the GPS in its standard mode. Differential equipment ranges in price from £500 - £40,000 and is dependent on the accurately positioned 'base stations'. The differential kit can also be cumbersome in the field and politically sensitive, as some countries may object to the use of such high accuracy surveying equipment being used in the field. However, the accuracy and reliability of DGPS is substantially better than any other form of GPS technique. The main principles of differential GPS are shown in Figure 6-19.

Wide Area Augmentation System: WAAS is the American name for a system similar to differential GPS with corrections based on deviations of GPS estimations compared to known locations. Geo-stationary communication satellites are used to transmit this correctional information to GPS receivers. The European system is known as EGNOS and the Japanese system currently under development is called MSAS. An EGNOS prototype system (ESTN) is currently broadcasting signals as part of its testing phase. Sometimes the satellite will stamp them with a '*Do not use*' signal and GPS receivers will ignore these signals when making calculations. Many modern 12 channel sets include the capability to receive WAAS corrections and they dedicate one of their channels to receiving these. WAAS is of particular interest to the US Federal Aviation Administration (FAA) who developed the system to aid in landing planes. The system is designed to offer a guaranteed position of better than 7 m and down to 3 m near airports (see *http://gps.faa.gov/*).

Figure 6-19 Fundamentals of differential GPS. Differential GPS offer significantly improved accuracies over standard signals.



GPS signals are distorted as they pass through the ionosphere. Though the ionosphere is different in density at every point, areas of the ionosphere close to one another will normally have similar electron densities.

The two locations on the ground receiving the dashed lines are receiving data passing through similar points on the ionosphere. The two locations receiving the dotted data are receiving data passing through distant points on the ionosphere. Close points will usually have similar ionospheric errors.

It follows that close areas on the ground (<100-200 km) will have similar dilutions. A differential signal is calculated by using the current error at a point of known location. By applying the inverse of the known error to the recorded position in the field allows a much-improved estimate of the location. The quality of this correction is dependant on the type of equipment and the proximity of a base station. An error of 1-5 ppm (part per million) of the distance of the receiver to the base station should be added to the stated accuracy of the device. I.e. if the nearest base station is 50 km away then the DGPS error = 50/1,000,000 = 0.00005 = 5 cm. This gives an error of approximately 5-25 cm depending on the equipment.

(1) Point of Known Location

(2) Offset of known location

(3) Reported position in the field

(4) Corrected position using the inverse of the offset of the point of known position. (*See right*)

Carrier Wave Post Processing: Surveying grade differential GPS stations use this feature. It is an augmentation that analyses the actual radio signal sent from the satellite and not the information it carries. Because the pseudo random code only emits a 1 or 0 every microsecond it is very difficult for a receiver to transform this into a precise time. Within a microsecond light travels 300 m, which is a very large error. High quality receivers can measure to within 1 or 2% of this but that error is still very large (3-6 m). The carrier wave itself has a frequency of 1,570 MHz, meaning it transmits about a billion times a second. In this length of time light travels less than 30 cm. The carrier wave alone is useless for timing because each wave looks essentially identical but carrier wave processing combines the pseudo random code and the carrier wave to determine a more precise location. Carrier wave processing can get positional resolutions down to 1-3 % of the frequency of the signal (1-2% of 30 cm = 3-6 mm) but this requires significant post processing. The realistic maximum resolution is the 20-30 cm wavelength of the signal, but is commonly around 1 m. Carrier wave post processing requires a laptop or PDA to be connected to the GPS at the time of data collection and requires a lot of processing to determine a location.

For inexpensive use carrier wave processing can sometimes be achieved by using a standard set of output files from the receiver (known as RINEX) but this requires access to unprocessed, poorly-accessible pseudo-range data held within the GPS. This method is explored in more detail in Section 6.10.2. The various techniques and their accuracy levels are summarised in Table 6-6.

Table 6-6 Accuracy improvements for GPS.

Type	Error	Cost
Standard quoted GPS Accuracy	15 m	Free
Typical GPS Accuracy (post SA)	6-8 m	Free
Averaging	~3-4 m	Free
Differential	<1-5 m	£500-£40,000
WAAS	~3-7 m	Free on compatible GPS receivers
Carrier Wave	<1 m	Free or Professionally done (>£500)

6.10.2 Improving accuracy (manual post processing)

As this chapter has stressed on numerous occasions high accuracy GPS data is often no real benefit to an expedition. However, if this accuracy is needed and the party cannot borrow equipment or purchase a several thousand pound DGPS system then there are alternatives.

Many methods for improving GPS accuracy rely on accurately sited base-stations in close proximity to the expedition and or a degree of expense. In remote expeditions on a tight budget these two facets make many of the techniques inaccessible. As we have seen from Section 6.10 a differential GPS can be corrected by using signals from a point of known location. The section on averaging also suggests that if a GPS left for a length of time any point can be surveyed to a high degree of accuracy. Based on these two facts it might seem logical to assume that a GPS could be setup at a base-camp and used to calibrate a roving receiver in the field. If all the readings of the base-camp GPS were recorded for a day then at the end of the day the roving GPS could have all of its readings post-processed to remove the errors at all the waypoints. Unfortunately, this is not possible. Because not even NMEA data (the complete output of all the calculated data performed by a GPS as shown in Section 6.9) contains the details about the SV used when making the calculation you can not know if the calculated error from the base-camp GPS has any relevance to the waypoints recorded by the GPS. Proper differential GPS record the pseudo-range data. This pseudo-range contains the information about the individual satellites and can be used to correct the signals appropriately. If a user tries this without the complete data then the processed information will be meaningless and less accurate than a standard waypoint. If the pseudo-range data can be obtained then there is the possibility of post-processing the data.

Until recently the most common inexpensive software for pulling out pseudo-range data was Async Logger. Async Logger outputs a RINEX (Receiver Independent Exchange) file; this is a data standard used in many applications. RINEX dumps generate large amounts of data and care should be taken if only small amounts of storage space are available. A 5 min dump will create somewhere in the region of 250 Kb of data. Where several GPS are being used in the field these dumps will be creating several Mb files per day. This compares against the few Kb generated by normal Waypoints and track points. The entire Bogda Shan GPS file is about 988 Kb. This was over a period of several weeks using up to five GPS receivers. It is easy to see how RINEX dumps can get out of hand if storage space is a premium.

Async Logger is a difficult product to use that has not been updated for a while. A much easier to use modern equivalent is the Delorme GPSPostPro 2.0 software. This software comes bundled with a roving GPS and a base-station GPS. The base-station can be used to calibrate the roving unit. The only disadvantage of this system is that the base station unit is a self contained box that has no user configuration features on the outside and no I/O functions except Bluetooth. This means that to use this you will need either a PDA with Bluetooth or an expedition laptop similarly configured.

The Delorme system costs around $300 which is very cheap for a differential solution. This system claims sub-meter accuracy and is both very cost effective and easy to use. Both the base-station and the roving receiver are complemented by WAAS and can search several FTP sites for Continuously Operating Reference Stations (CORS) or Orbit and Permanent Array Center (SOPAC) signals in the vicinity of the expedition base-camp. CORS and SOPAC are differential signal stations that post their data onto FTP sites. These can be accessed from the Internet and post-process expedition data. An idea for how this hardware can be used in the field is shown below in Figure 6-20. The advantage of using the equipment shown is that it can all be purchased very cheaply. The laptop can be acquired for £199.74 through SterlingXS. A yellow Garmin EXTREX retails for £84.84 through GPS Warehouse (*www.gpsw.co.uk*). The Delorme differential GPS set including the roving antenna and the calibration set costs $349.95, the external antenna costs $32.95 and additional PDA software costs $39.95. Shipping from the US for the Delorme set is $40.00, import duty comes to £11.00 and VAT comes to £36.74. Using an exchange rate of 1.91 (current at time of writing) gives a total cost of £279.07 for the GPS equipment. The total amount for the equipment is £563.65 for sub metre accuracy and a field hardened laptop.

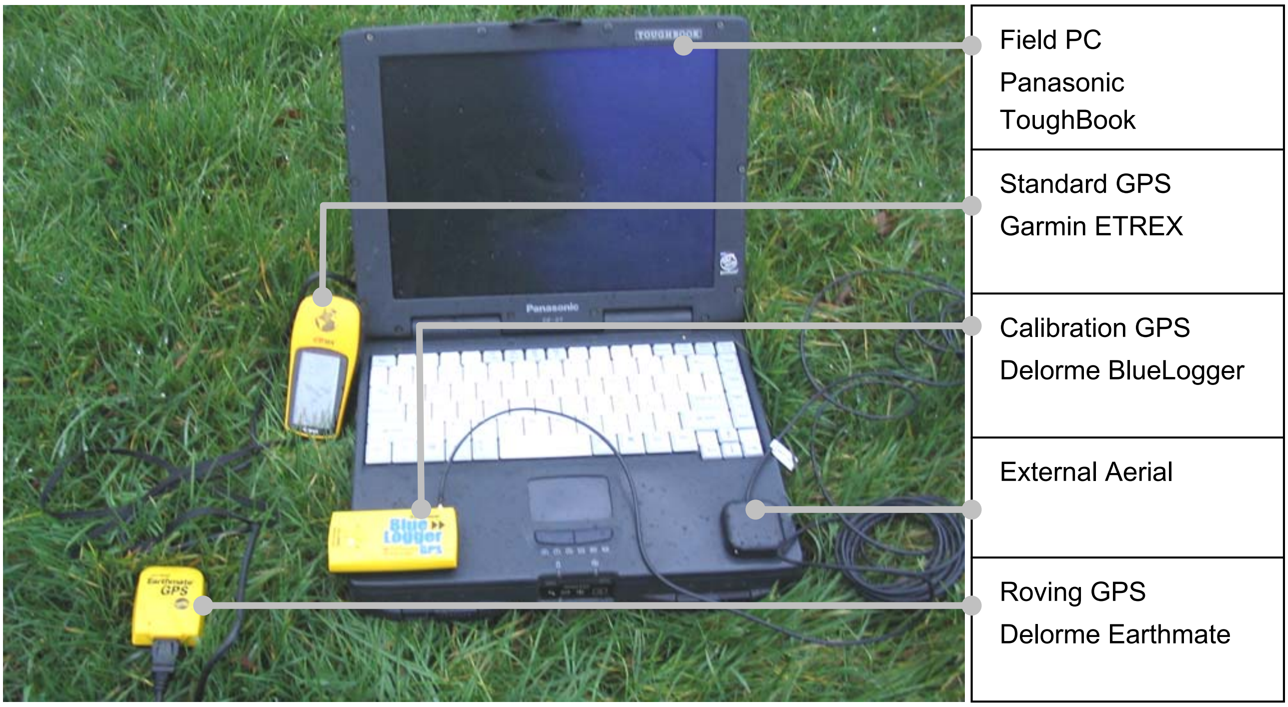


Figure 6-20 Delorme Earthmate hardware with GPS PostPro on a field laptop.

Five hundred pounds is a lot of money for an expedition but it is the only realistic way of obtaining a full hardware solution capable of delivering sub-metre accuracy GPS data.

6.11 Future developments of the Global Positioning System

The NAVSTAR system is continually evolving and it is worthwhile to be aware of the new systems that are coming online over the next few years so that any expedition is best placed to take advantage of them.

The biggest current development in GPS accuracy has been the satellite based WAAS. Though these systems were due to come online in 2003-2004 their current (2005) coverage remains poor (*www.esa.int/EGNOS*). The US is further ahead with their plans while the European system is still undergoing testing. There have been issues with the European WAAS satellites (EGNOS) picking up US based corrections. As can be seen from Section 6.5.3, a differential correction for America will have passed through a significantly different part of the ionosphere and received different levels of attenuation. A firmware update for Garmin units was issued in 2003 to stop the unit confusing US and European corrections. There are plans to launch WAAS satellites for Japan, China, Australia, India and Brazil. This will significantly improve the NAVSTAR system throughout the world. The timings of these launches are not yet officially known and the testing phases of the system are often longer than stated. Any expeditions to these territories should check in advance on the status of the corrections.

In 2005 the satellites should begin broadcasting extra signals that will allow a form of basic differential correction to be conducted using a single receiver. Two military signals will be added to the L1 band and a civilian signal will be added to the L2 band. The specifics of this system are not yet available but in future even standard non-WAAS GPS signals will be more accurate and better able to cope with an obscured sky, such as under dense tree cover.

Another significant development is a third GPS constellation due to be launched within five years; this is a European GNS called Galileo aimed at commercial use. To encourage take up, Galileo will not be susceptible to SA. Should the US rescramble any signals the Galileo system will guarantee global coverage. Though the system is no more accurate than NAVSTAR, having access to more satellites will speed up acquisition time and improve accuracy.

Eventually, by 2008, a new generation of NAVSTAR satellites will begin broadcasting a third wavelength in the L5 band. The principle behind the inclusion of the L5 band is that the difference in retardation between L1 and L5 will tell the receiver the electron density in the ionosphere. If the receiver knows the status of the ionosphere it can eliminate this error. Eventually the only errors in the GPS system will be from local effects. This will offer real-time inexpensive centimetre accuracy.

The expedition can keep up to date with recent advances by checking the websites of the various service providers.

Table 6-7 GPS Service Providers.

System	**Internet address**
GPS	*www.gps.losangeles.af.mil*
GLONASS	*www.glonass-center.ru*
Galileo	*www.europa.eu.int/comm/dgs/energy_transport/galileo/*
EGNOS	*www.esa.int/EGNOS*

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section C: Techniques

Chapter 7: GISci Analysis

7 GISci Analysis

7.1 Introduction

A GIS is not simply a tool for storing data from different sources (images, maps, points), nor just a method for relating points on a map to a database. The real advantage of a GIS is that it allows the display, analysis and output of spatially-related data. This is made possible by the geographical framework under-pinning all of the data layers. GISci analysis is when all the hard work of integrating spatially-referenced data starts to pay off! In this chapter we start by looking at straightforward GIS functions: *mapping* layers of data and graphically *symbolizing* their attributes so that patterns may become apparent. *Spatial selections and queries* allow features from one layer to be extracted from an area defined by another layer. Similarly, *spatial overlay functions* provide a useful way of manipulating data, using the extent of one layer to add or subtract from another or to combine the two. *Buffering* and *proximity analysis* are useful when looking at phenomena within (or beyond) a certain distance from another. More advanced techniques include *correlation*, to look at the relationship between different layers.

Several techniques are used when a 'third dimension' (such as height, rainfall, temperature) is being considered. *Interpolation* is used to estimate values between known data points, while *contouring* is used to draw lines around existing values, as on a normal height contour map. *Triangular irregular networks* provide an efficient way to represent and analyse complex '3D' surfaces. Zukowskyj and Teeuw (2002) compare the performances of various GIS software in carrying out interpolation along a river valley.

This chapter also looks at data *topology,* which defines the relationships between points, lines and areas in geographic data. This seemingly obscure topic becomes relevant when, for example, mapping roads or analysing area data such as vegetation types. When data contained in several layers determine the pattern in another layer, for example, vegetation and soil determining the distribution of a plant species, then *modelling* techniques can help to predict or understand the distribution of that plant. Finally, there is a description of statistical techniques for characterising and analysing data, useful for both spatial and non-spatial data sets.

Most of the techniques described here relate mainly to vector data - points, lines and areas. The next chapter looks at the analysis of raster data and digital image processing. However, some techniques here concern 'surfaces' of values being calculated or predicted, and the raster model proves useful in these cases. A summary of suitable software for all of these functions is given in Chapter 1.

7.2 Mapping and symbolizing

A visual assessment of field data can be made using different symbolisation techniques: Figure 7-1 shows a variety of styles available in the MapInfo GIS package.

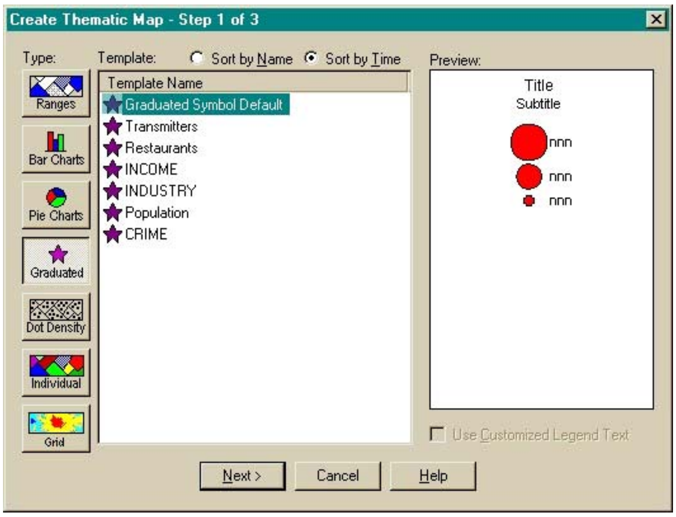


Figure 7-1 Defining plot types based on attributes.

For example, records of zebra sightings in the Mkomazi Game Reserve, Tanzania, were mapped out using circles whose diameter is proportional to the size of each group. This is shown below in Figure 7-2. Simple mapping techniques like this can summarise a large quantity of field data in a readily accessible way, as well as providing a useful overview prior to further analysis.

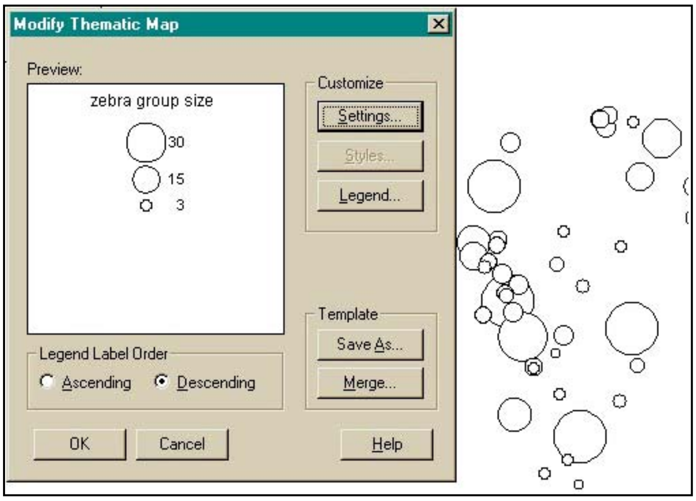


Figure 7-2 Defining actual plot sizes.

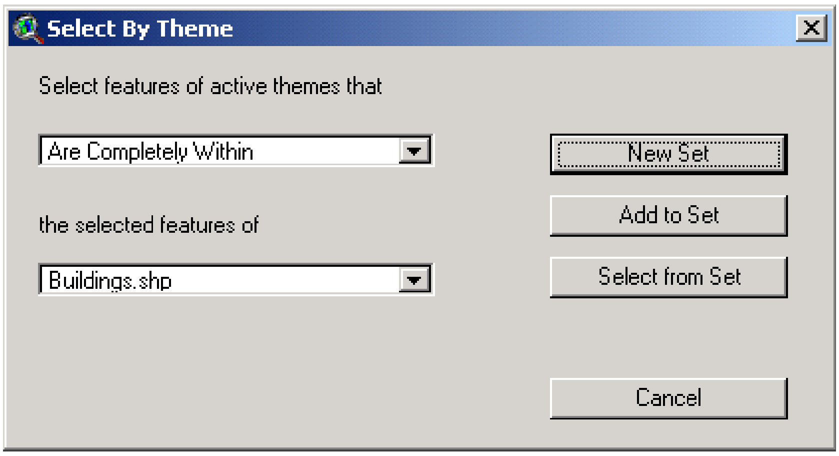
7.3 Spatial selections and queries

7.3.1 Simple data queries

A simple 'conventional' database query might select records from a GIS dataset based on some attribute or value within the data. For example, in the Bogda Shan set of GPS readings, we might select all point with an altitude above 1,000 m. Making such a selection is a function available in most GIS programmes, or in database terms a function that can be entered with a Structured Query Language (SQL) statement (see Chapter 4). The resulting selection can be mapped, as in Plate 6.

However, a key benefit of using a GIS is being able to query data not only according to its attributes, but according to its geographic position. This might take one of several forms:

- Selecting all features in one layer that lie within another layer (see *Figure 7-3*), e.g. all zebra sightings made within an area of grassland, or all occurrences of a plant within a certain country.
- Selecting all features that lie within (or beyond) a certain distance of another feature e.g. all woodland that lies within 1 km of waterholes, or all houses within 10 km of a volcano (See Plate 7).
- Selecting all features within an area drawn on the map – this can be a regular area such as a square or circle, or an irregular shape drawn with a 'lasso'.



Select features of active themes that

Are Completely Within

the selected features of

Buildings.shp

Figure 7-3 Using a 'Select By Theme' query in ArcView 3.2 to select the points in one data layer that lie within areas covered by buildings.

In all of these cases, point, line or area features can be selected. Such spatial and attribute queries can be combined: for example, selecting all trees within 1 km of a waterhole (spatial query) and above 5 m in height (attribute query).

7.4 Creating 'buffers'

GIS can create new layers comprising areas within a specified distance of existing features – these are known as buffers. *Figure 7-4* gives an example, where different levels of error are shown as buffers around a series of GPS readings. In fact, in this case three buffers have been created, showing the distribution of likely GPS positional errors within one, two and three standard deviations respectively. In this case, buffers have been used to visualise error analysis but there are many other expedition applications: areas prone to flooding within a given distance of a river; an 'exclusion zone' around breeding bird colonies; or areas of radio coverage, for example.

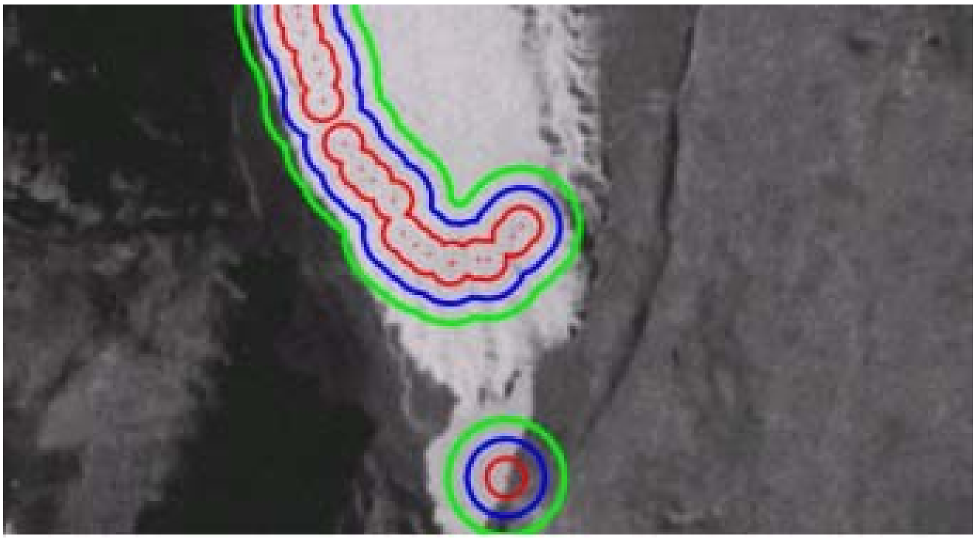


Figure 7-4 Buffers around GPS points, at 15 metre intervals, indicating the likely positional error.

7.5 Testing relationships between data layers

In some cases the relationships between data will not be easily measurable and some level of uncertainty may exist. This section touches upon testing hypotheses and the confidence that can be placed in an analysis. It is common for an expedition's project to involve testing some theory or hypotheses. In a GISci context, this typically involves an investigation of whether the distribution of one phenomena (e.g. vegetation) is significantly associated with another (e.g. elephant distribution). There may or may not be a causative relationship between the two: does the vegetation determine where elephants are seen; do elephants determine the vegetation through their feeding; or does some other factor influence them both at the same time?

This manual does not aim to describe the statistical methods needed for such tests. Rogerson's *Statistical Methods for Geography* (2001) is recommended, as it presents statistical techniques and issues with a strong spatial emphasis, as well as setting them in a GIS context: "GIS do not reach their full potential without the ability to carry out methods of statistical and spatial analysis".

The chi-squared test is an example. It measures the probability that the differences between two datasets have occurred by chance – or put another way, it indicates the similarity between two datasets. It is often used in GIS and geographical studies, as it makes few assumptions about the data being used: they do not need to be normally distributed (see the final section of this chapter), small sample sizes can be sufficient, and categorical data (such as soil types) and ranked data (such as a list of villages ordered by population size) can be used as well as numerical data. Many other methods are used to test whether sets of variables are associated and to quantify the relationships between them. The chi-squared test considers just two sets of data, but other methods exist for the analysis of many sets of variables – see Rogerson (2001) or other statistics textbooks.

7.6 Spatial overlay operations

We have seen how one data layer can be used to *select* data from another. We now look at overlay operations, where one layer is used to *change* another. The general principles are shown in *Figure 7-5*: layers can be combined together (top), overlapped (middle), or cut (bottom). Different GIS programmes use different terminology for these operations; for example in ArcView and ArcInfo, these three functions would be called union, intersect and erase. Generally, one layer must be a polygon (for example, vegetation type), while other layers can be point, line or area/polygon (for example, animal sightings, roads or protected areas).

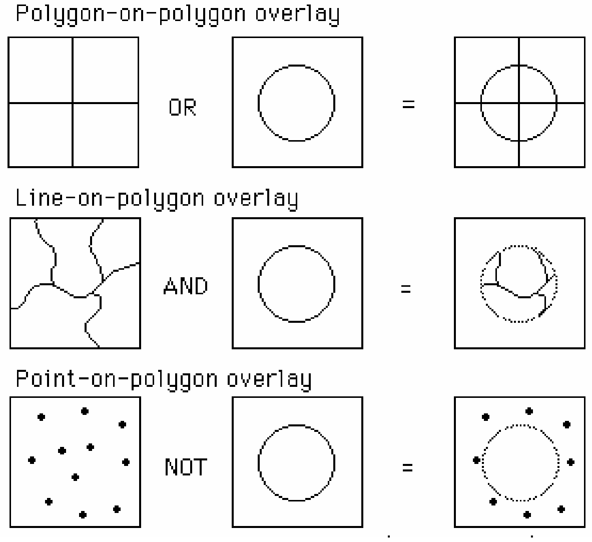


Figure 7-5 The principles of some spatial overlays using 'or', 'and' and 'not' logic. In these examples, point, line and area/polygon data types are all used. These three functions correspond generally to ArcView's 'union', 'intersect' and 'erase'.

These operations are commonly used in expedition GIS. In describing an animal habitat we might wish to find all rivers (represented as lines) flowing through open grassland habitats (polygons): this is an 'intersect' operation. We could also find all grassland lying more than 2 km from roads by first creating a buffer around the roads and using it for an 'erase' operation on the grassland layer. And we might use 'union' to combine two habitats, for example to measure their total surface area or to map a site for protection.

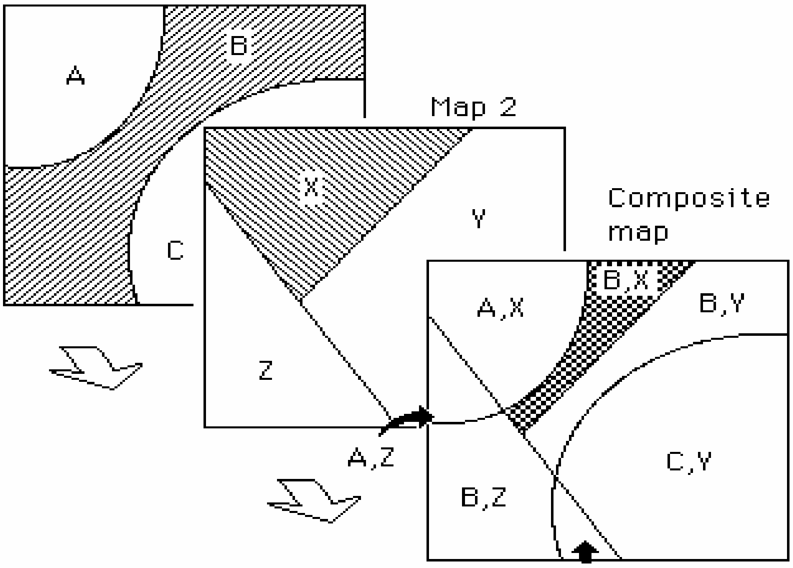


Figure 7-6 Example of the union of two polygon layers. Note that the attributes from each of the two original layers are all retained in the resulting output layer.

7.7 Spatial analysis

GIS analysis can be used to combine various datasets and query them effectively. Unfortunately, GIS are often seen as ‘bins’ for depositing lots of layers of information without delving into and manipulating / modelling the data. It is not until the team starts querying the relationships of these themes the true benefits of a GIS will be realised. This section of the manual uses two case studies to show how layers can be built up and used together to predict results. In the first example, the location of semi-precious minerals are modelled using limited field data and remote sensing techniques. In the second example similar techniques are used to predict the nesting sites of animals based on certain criteria.

In the first example from Moore *et al*. 2000, a GIS was used to determine the most likely occurrence of diamonds and jade from the Kunlun Shan, NW China. Alluvial deposits had been found in the Yurunkash River and a larger source lithology was logically located upstream. The project used Landsat TM data, previous field observations and hardcopy topographic and geological maps to identify the source rock for the minerals. The work concentrated on locating large source localities further back in the mountains.

The comparatively coarse pixel size of TM data (28.5 m) meant small-scale lithological changes or detrital jade localities were unresolvable. Also, there was not enough information to determine exactly where jade would occur, only enough to locate lithologies that had similar metamorphic histories. Only knowing the general metamorphic units from the maps and the Landsat data revealed hundreds of square kilometres of potential source rocks. This was too great an area to ground truth for jade and additional queries were required in the GIS to find more definite locations. The geological map was digitised by hand into a GIS database for use in querying the relationships between rock units and potential mineral distribution. Landsat data was used in an effort to corroborate the geological map and locate areas of medium grade metamorphic units. The image was enhanced to remove the dominant brightness of the red and blue bands and then processed to make a lithological map. The techniques for removing colour bias and processing images for specific minerals using band ratios is described in more detail in Chapter 1.

The exploration model used in the preliminary stages of this investigation was to define fluvial activity that crossed known jade localities. A vectorised overlay of the rivers was produced and projected onto the digitised geological map as shown in Plate 8. Once data is in a GIS it can be extracted and analysed in many ways. Certain parts of a theme can be extracted based on their relationship with another theme or based on their own attribute data. The Hotien case study required both of these techniques. To find the source of the jade deposits required examination of the area’s drainage systems. The rivers in the area were digitised by hand from the satellite data and then imported into the GIS. From the stream data, vector overlays of the drainage patterns could be constructed. From those overlays, streams could be identified that intersect lithologies of potential economic significance. The GIS was then used to locate lithologies most likely to have deposited the jade fragments.

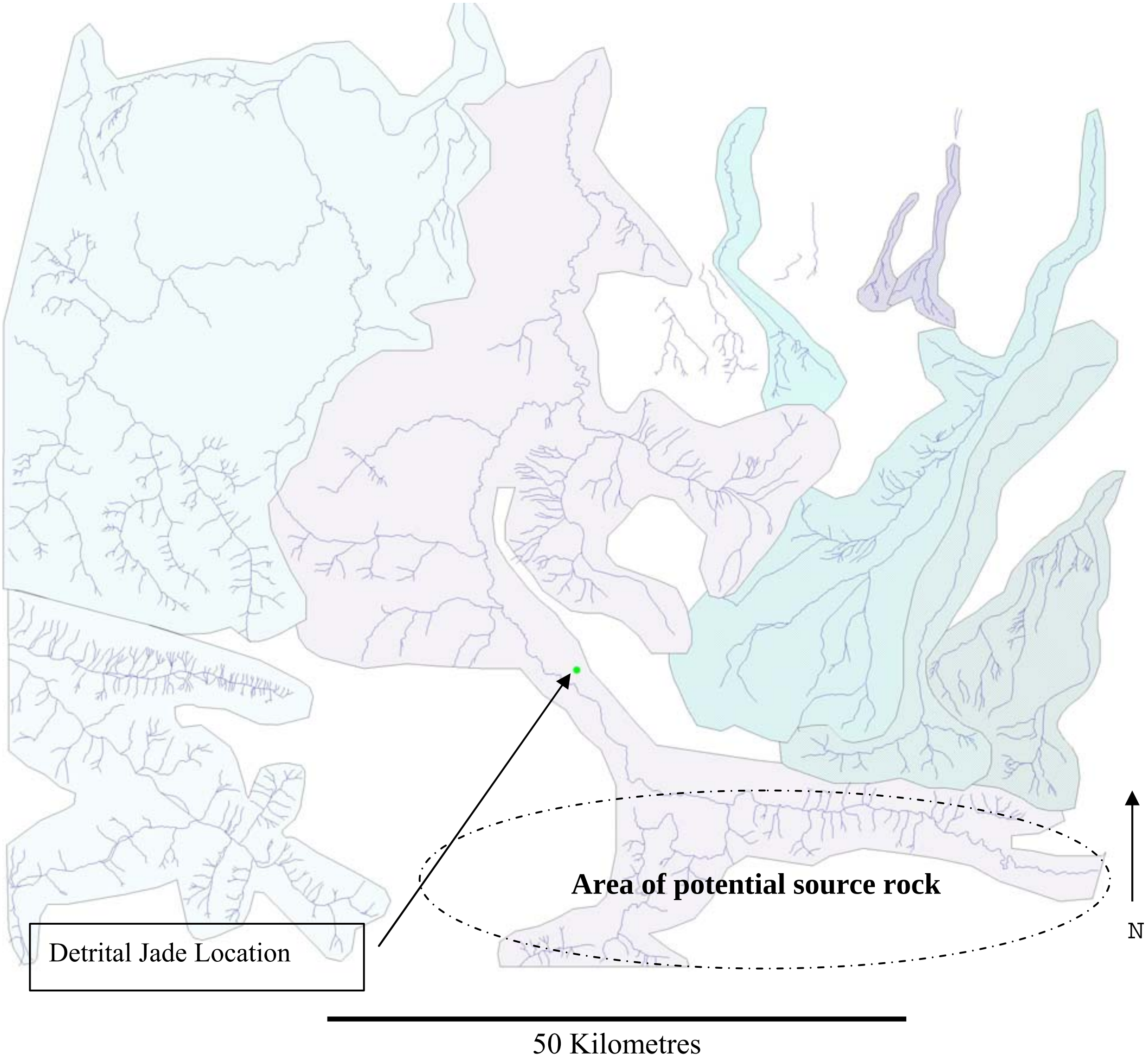
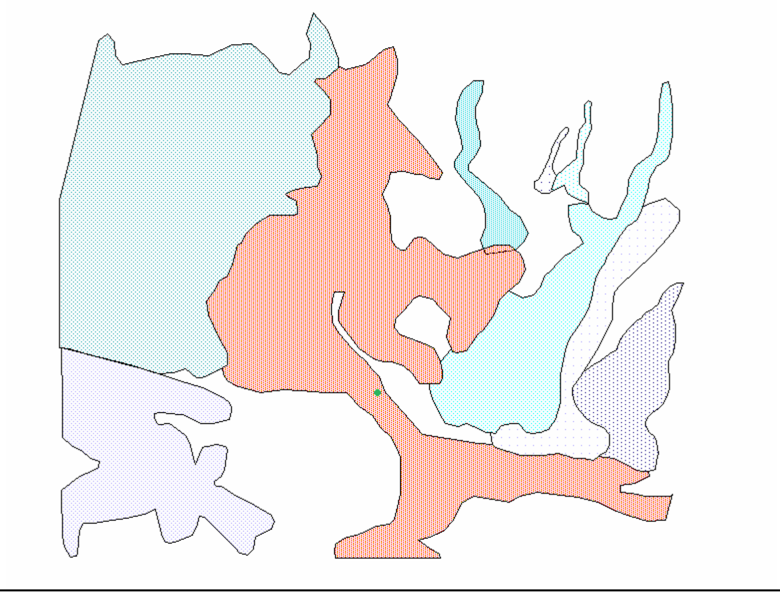
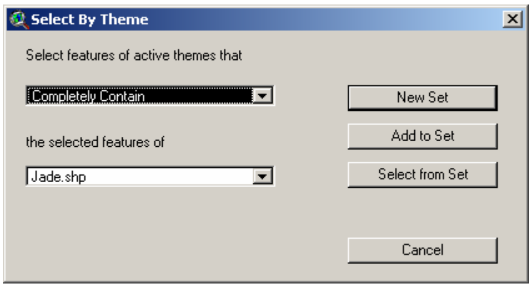


Figure 7-7 Extracted drainage patterns from GIS.

Using the GIS we can merge drainage patterns that intersect the known jade localities and exclude those where no jade was found. We can then highlight other potential units further back in the mountains to reveal larger sources of the semi-precious minerals. This step-by-step procedure is shown in the series of figures over the page.



The rivers have been simplified into catchment areas. The relevant catchment areas are selected using a spatial Select by Theme query. This highlights the drainage areas that cut through jade bearing lithologies.

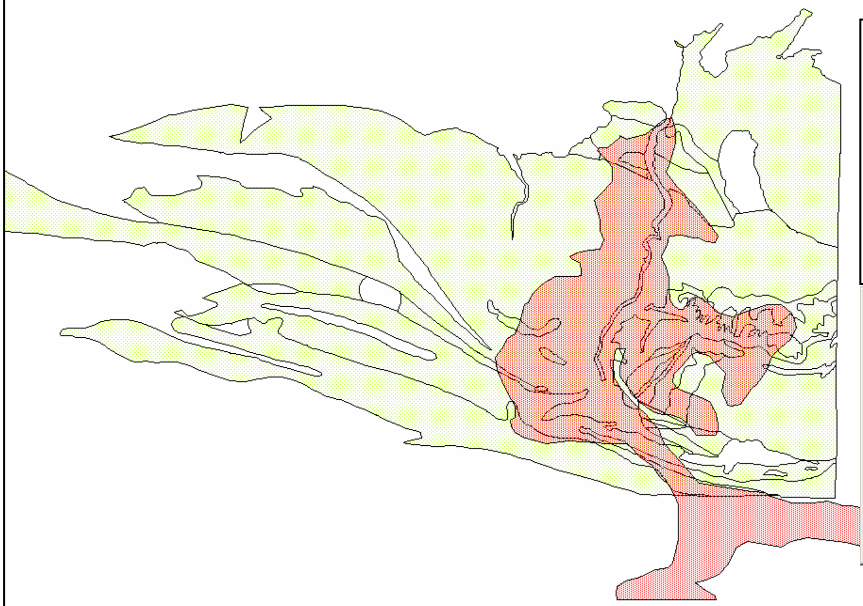


Select features of active themes that

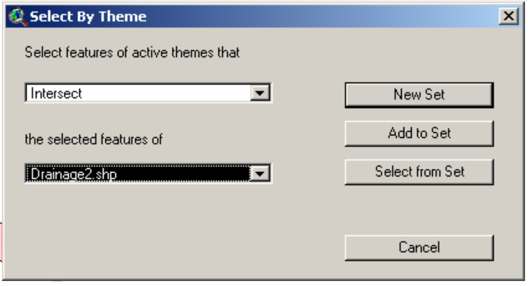
Completely Contain

the selected features of

Jade.shp



The selected catchment area can be used to filter the lithologies that are underneath it. Only rock units that intersect the river are highlighted.



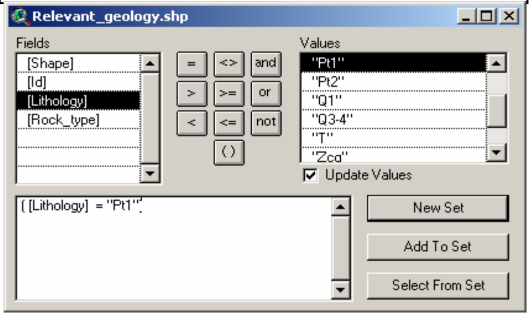
Select features of active themes that

Intersect

the selected features of

Drainage2.shp

Each lithology has a set of attributes including a code and other information. Knowing the potential source units of jade (i.e. code ‘Pt1’ from key in Plate 8) we can write a standard query to extract the relevant source materials.



Fields

[Shape]
[Id]
[Lithology]
[Rock_type]

Values

"Pt1"
"Pt2"
"Q1"
"Q3-4"
"T"
"Zcq"

☑ Update Values

([Lithology] = "Pt1")

Using a GIS to quickly interrogate data using spatial queries and the relationship of different themes restricts the need for prolonged work in the field and maximises the chances of getting the best results from a project. In the Hotien project, hundreds of square kilometres of potential jade bearing rocks were reduced to small areas believed to have jade bearing lithologies Plate 9 shows the project conclusions and attempts to localise the jade source rocks.

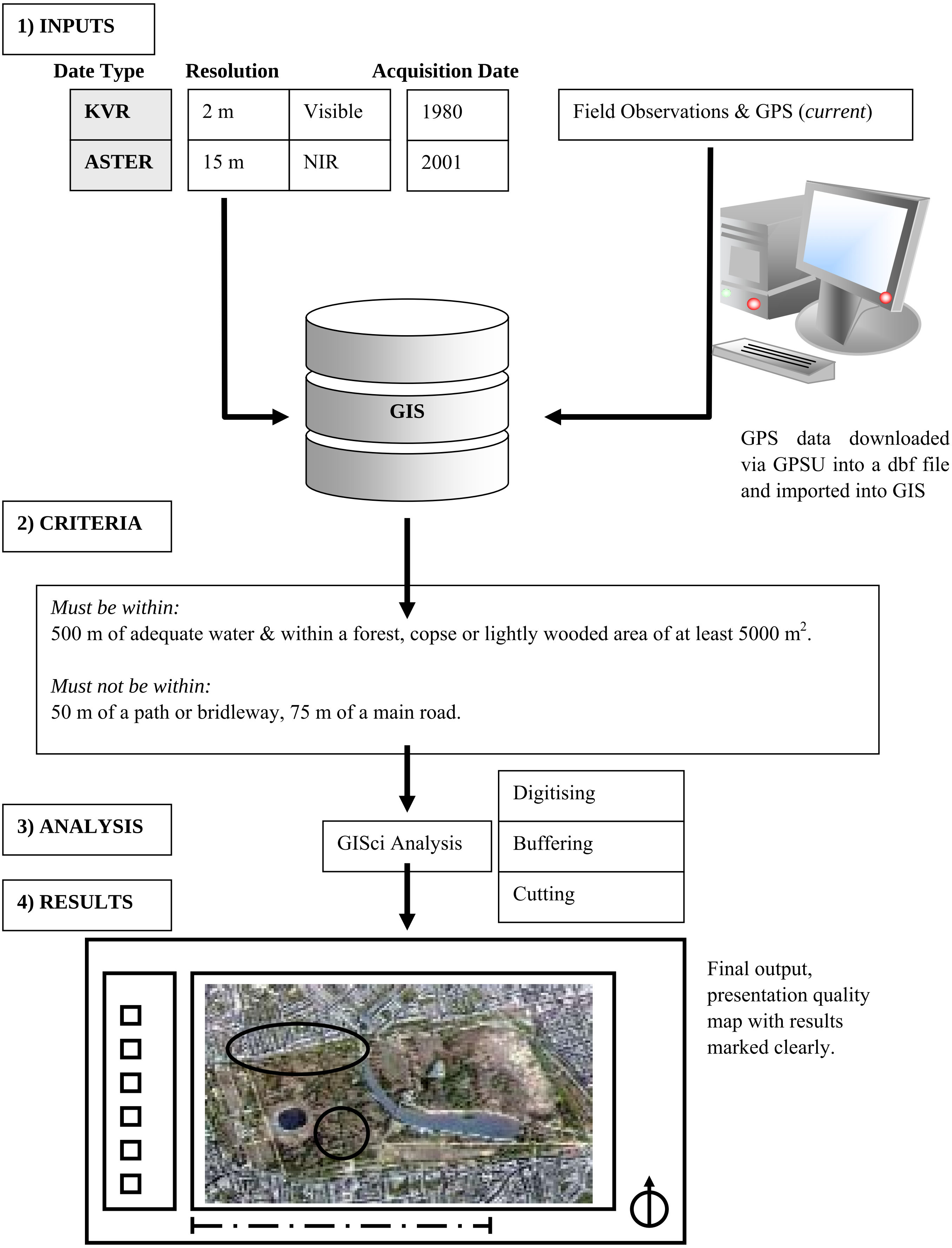
This sort of analysis can be used in many other disciplines, not just in geological sciences. Selecting polygon layers based on other themes is a common tool in habitat mapping, biogeography and various social/political geography studies. To visualise how these techniques might be used in a habitat mapping exercise, a simplified example is shown below using a similar methodology to the above geological mapping model.

In this example, potential habitats for squirrels are predicted using ASTER commercial satellite data, KVR declassified spy-satellite data and GPS track points. This example is a theoretical study used as a training exercise at the RGS-IBG EAC Mapping Unit training weekends. It is overly simplified with very basic criteria, but the principles can be scaled up to apply to even large scale projects. The example considers parts of Hyde Park in London but the reader should be able to see how these techniques can be extended to large, complex mapping exercises. We can input data into the GIS define certain criteria and get the GIS to output the most likely result. The schematic flow is shown over the page in Figure 7-8.

The data sets supplied for the exercise include high spatial resolution KVR panchromatic data and medium resolution ASTER multispectral data. When conducting any kind of analysis it is important to select the dataset that best suits the task. Both ASTER and KVR have strengths and weaknesses.

The criteria required for the mapping exercise are to find nesting grounds that are within 500 m of adequate water and within a forest, copse or lightly wooded area of at least 5000 m^2. An additional set of restraints are defined as follows; the habitat must not be within 50 m of a path or bridleway or 75 m of a main road. We can use the imagery to define these areas and then ask the GIS to display the most likely habitat.

Figure 7-8 Schematic Flow Showing Inputs and Outputs from a GISci Habitat Analysis Exercise.



Date Type	Resolution		Acquisition Date
KVR	2 m	Visible	1980
ASTER	15 m	NIR	2001

Field Observations & GPS (*current*)

GPS data downloaded via GPSU into a dbf file and imported into GIS

Must be within:
500 m of adequate water & within a forest, copse or lightly wooded area of at least 5000 m^2.

Must not be within:
50 m of a path or bridleway, 75 m of a main road.

Final output, presentation quality map with results marked clearly.

To begin with we will define the areas that are good potential habitat areas. Though the supplied KVR data is high resolution and could be used to determine tree locations, its age makes it unsuitable. It is before the storms of 1987 and many other factors may have affected tree growth over a 25-year period. Decisions on the quality and relevance of data to a given task must always be given as much thought as possible. In this example, what seems to be the best data for the task is not immediately suitable. The NIR data from ASTER is good at delineating areas of dense, healthy tree growth. These areas have been hand digitised using ArcGIS and can be seen below in Figure 7-9.

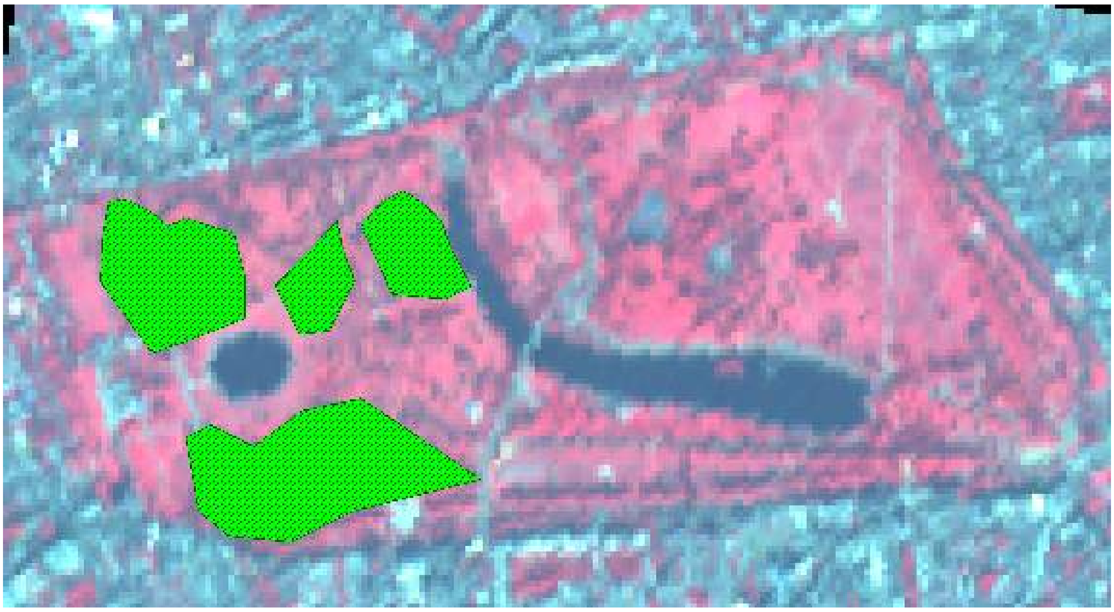
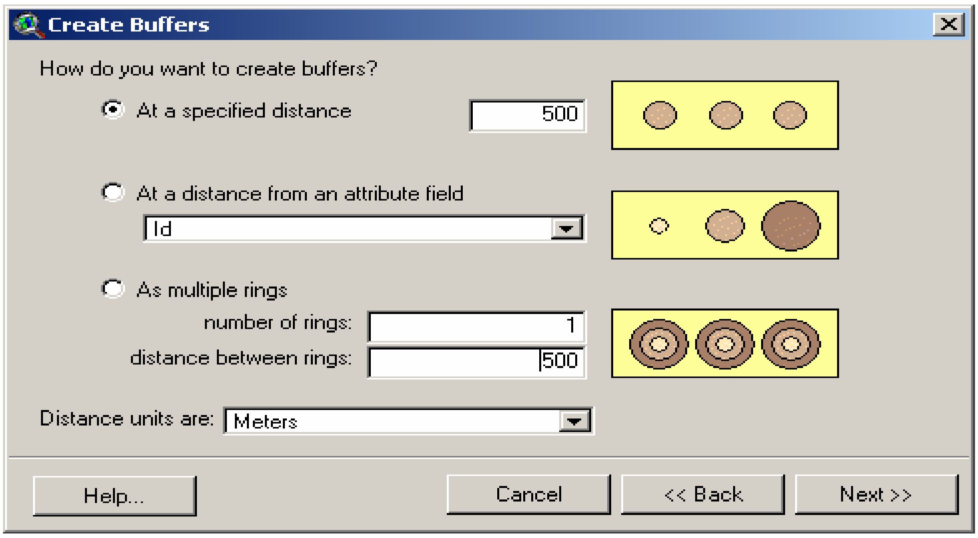


Figure 7-9 Areas of dense copses of trees in Hyde Park.

The squirrels are given an arbitrary requirement of needing a nesting site within 500 m of water. Considering just the areas in Kensington Gardens, the most notable water supply is the round pond. The pond is digitised from the KVR data because it is less likely a pond will have changed in 20 years and the higher spatial resolution of KVR data means the digitising will be more accurate. After digitising the pond we can project a buffer around it using ArcGIS's buffer wizard. This works in the same way as the GPS points were buffered around the glacial snout in the diagram above.



Create Buffers

How do you want to create buffers?

● At a specified distance 500

○ At a distance from an attribute field Id

○ As multiple rings

number of rings: 1

distance between rings: 500

Distance units are: Meters

The process of buffering creates a ring around the outside of the pond. When buffering, the final step of the wizard asks whether to include areas within the original shapefile. In this case, this area must be excluded as the squirrel habitat can not be within the water. However, if this was a boundary around a forest or wildlife reserve then the area within the shapefile would be a valid habitat. In these cases the area would need to be included.

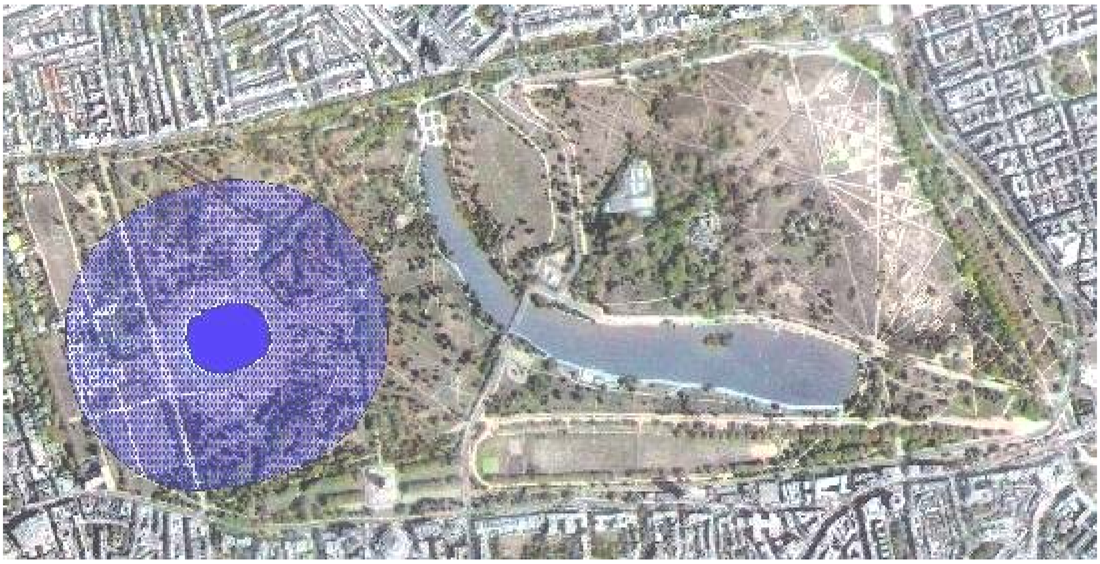


Figure 7-10 500 m buffer around round pond (excluding areas within the pond).

The pond buffer can then be used to 'cookie-cut' the forest polygons. In the Hotien lithology example in Plate 9 the whole lithological unit was selected if it intersected with a river. This was because any part of that lithology might contain jade and knowing the full extent of the jade bearing rocks would be critical to its extraction. In the habitat example we do not just want the forests that intersect the pond buffer, we specifically need the parts of that forest that are within 500 m of the water. This cuts off the outer extent of the forests and eliminates one polygon. The result of this query can be seen overlaid on the ASTER data in Figure 7-11.

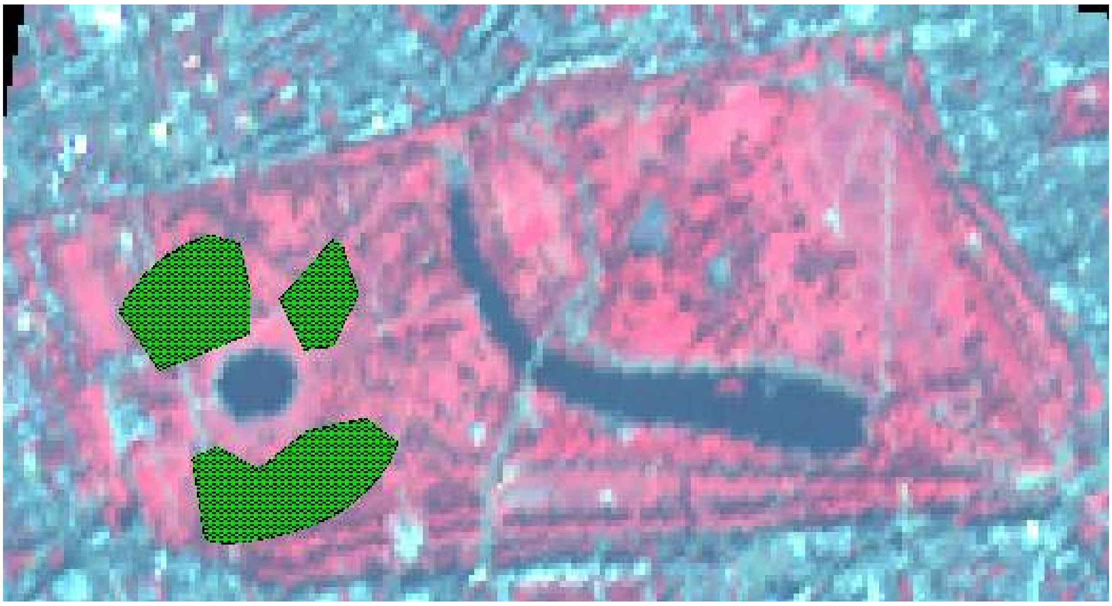


Figure 7-11 Selected relevant areas of forest.

Finally we want to eliminate any areas along footpaths where the squirrels might be disturbed. To do this we take point data from the GPS and buffer this by 50 m. This creates another polygon overlay that can cut the remaining three forest areas.

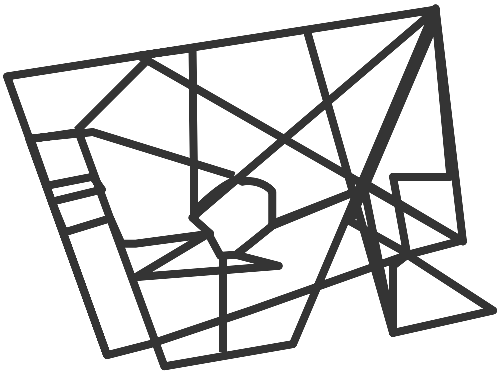


Figure 7-12 Shapefile from GPS tracklogs shown without background raster data.

We can also use GPS waypoints to map any tracks that don't appear in the raster data. We can bring those in as points and buffer them. This combination of good sites overlain with exclusion zones is shown in Figure 7-13. The GIS can union these features together creating one shapefile of all the exclusion zones. We can then 'cookie-cut' the forested zones to leave only the acceptable habitats.

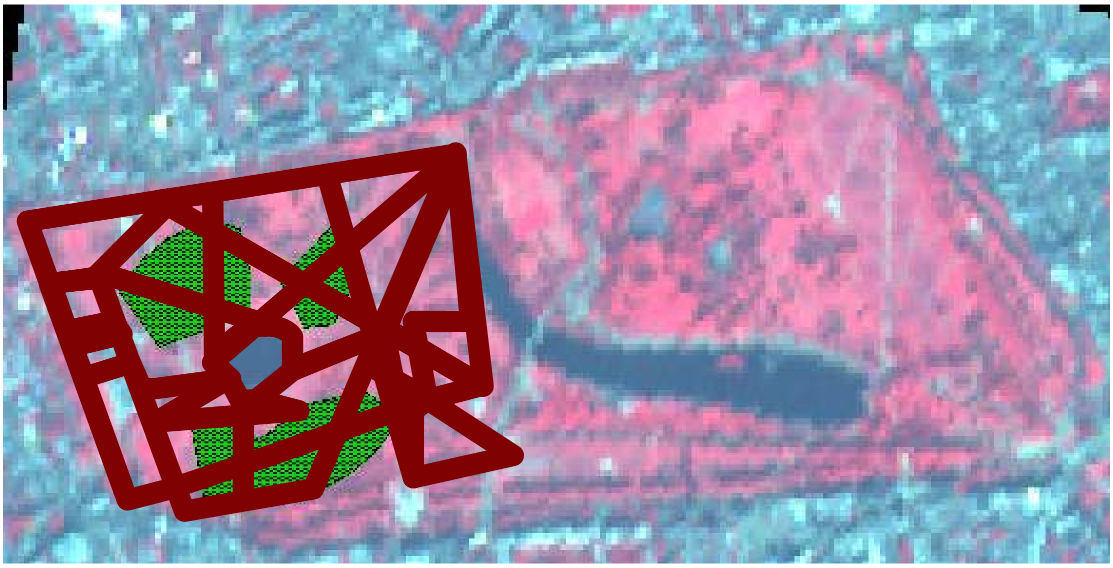


Figure 7-13 Potential habitat areas with exclusion zones.

There are approximately 10 potential sites. However, we might want to restrict our potential habitats to those above a certain size. For example, we might decide that the area must be greater than 5000 m^2. In ArcView a simple query can select the relevant polygons. Because our units are already defined in the GIS we can simply use:

[Area] > 5000.

The result is shown in Figure 7-14 as a yellow polygon.

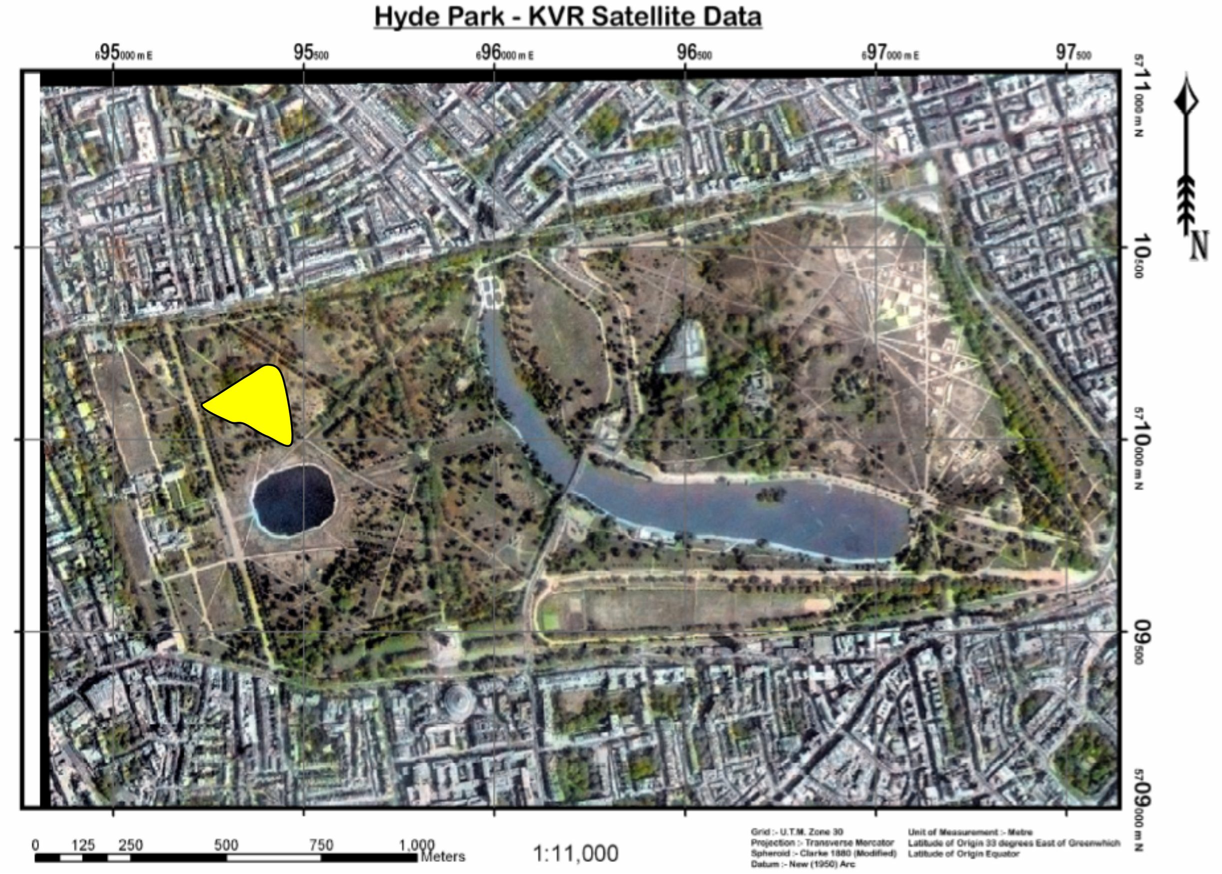


Figure 7-14 Most likely nesting site based on specified criteria.

This kind of analysis can be applied to many study types. The expedition might look at how many people are dependant on a source of water such as a well. It would be simple to query houses within a set distance of the well. If each of those houses had an attribute for number of people living there, it would be simple to total the number of people using the resource.

7.8 Digital elevation data

All GISci data requires two dimensions (X,Y) to be plotted in the GIS. Logically this data will require additional attribute information about the point. This information could be considered a third dimension. The most typical third dimension and the easiest to visualise is altitude. If a series of GPS points were available their height data could be plotted on a Z-axis. This can be seen in below in Figure 7-15.

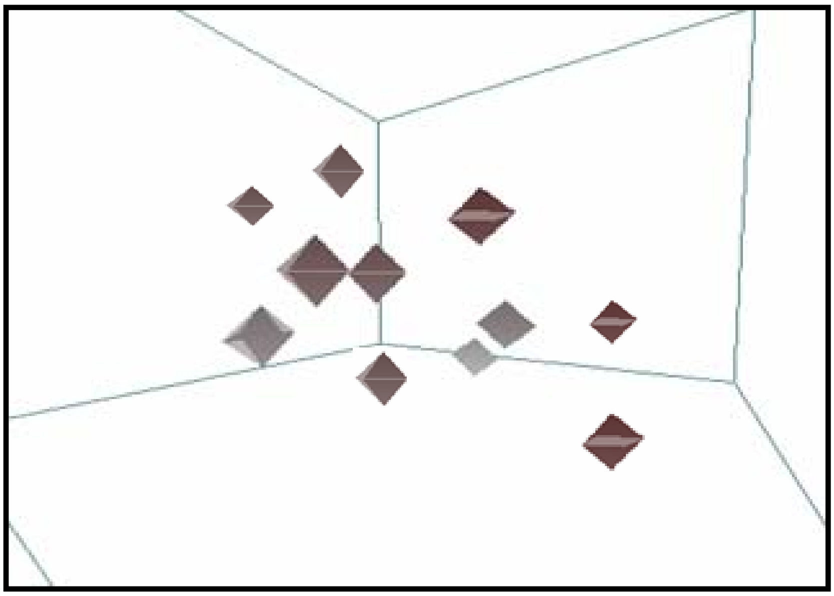


Figure 7-15 Individual GPS points plotted in 3D space.

There are three ways of inputting height data into the GIS. Height data can be random points as seen in Figure 7-15, or a regular grid with a height value at every intersection or a raster surface where the height is expressed as a DN Value. A raster image is essentially the same as a regular grid but expressed as pixels not discrete values. For the GIS to use it effectively the data must be extracted and converted to a grid.

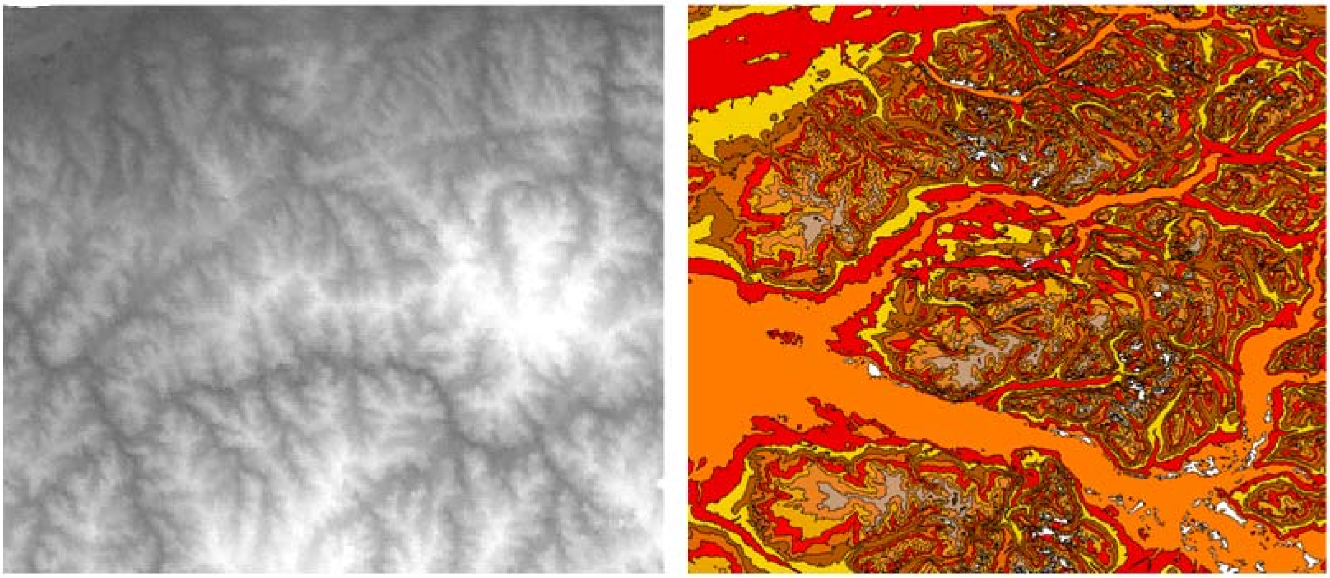
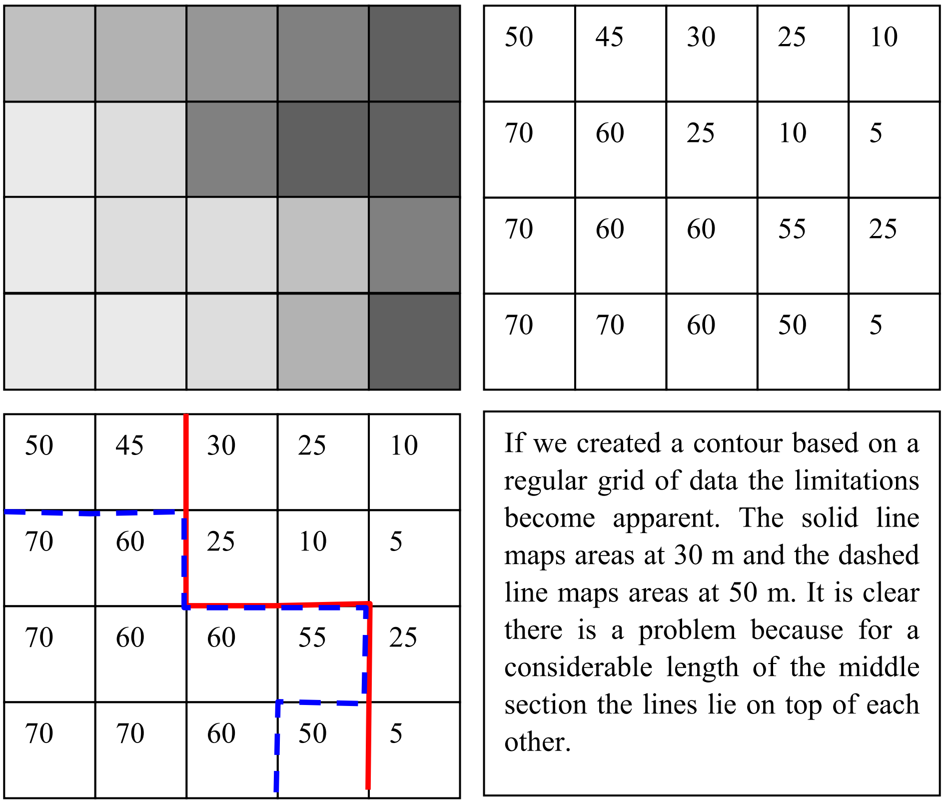


Figure 7-16 Examples of Raster Data (left) and Grid Data (right) (actual areas do not correspond).

When expressed as individual points either at random or in a regular grid some GIS analysis can be conducted. However, no matter how densely the points are plotted they are still individual points and the GIS needs to know more about their relationship. For example, one typical use of altitude data is to generate contours for a map. The weaknesses of individual points or raster data are shown below in Figure 7-17.

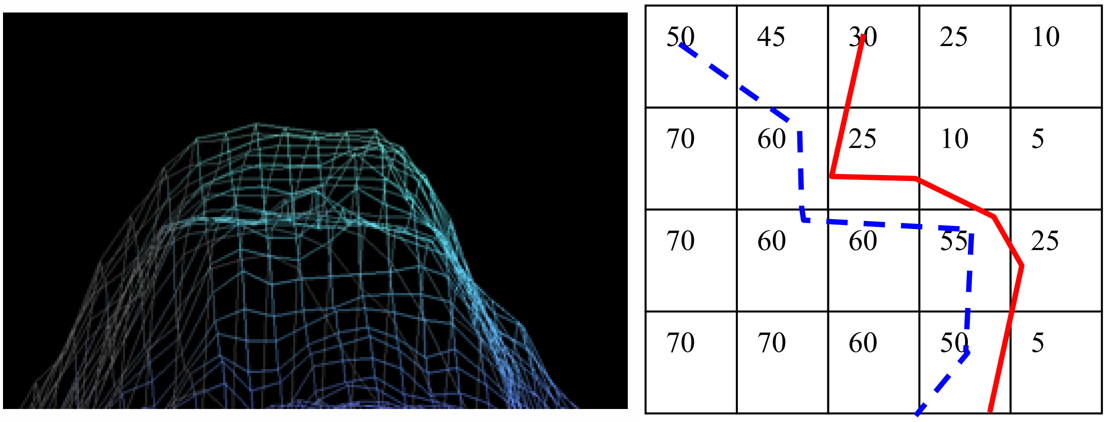


50	45	30	25	10
70	60	25	10	5
70	60	60	55	25
70	70	60	50	5

If we created a contour based on a regular grid of data the limitations become apparent. The solid line maps areas at 30 m and the dashed line maps areas at 50 m. It is clear there is a problem because for a considerable length of the middle section the lines lie on top of each other.

Figure 7-17 A raster is converted to a grid of points but the GIS can not understand where unknown / un-sampled points would lie.

The reason the contouring failed was because of the limitation of the data. The GIS was unable to calculate what happened in between the points. Cleary, in between two points such as 55 m and 25 m the topography must pass through 30 m and it is probable that the 30 m line would be closer to the 25 m point than the 55 m. The individual points can not tell us where that would be. The GIS is capable of turning these discrete points into a surface by interpolation. This process allows the GIS to predict the values of un-sampled points based on the surrounding points. The GIS interpolates the points into a series of triangles with the diagonals of the triangle used to calculate points that are not specifically mapped. These triangles are often referred to as a Triangular Irregular Network (TIN). A TIN allows the contours to be plotted more realistically as shown in Figure 7-18.



50	45	30	25	10
70	60	25	10	5
70	60	60	55	25
70	70	60	50	5

Figure 7-18 Grid data converted to a TIN and contours generated. Each node in the network is a point of value 'z' and from those nodes lines can be drawn to predict values at unknown points.

In the right hand image of Figure 7-18 the GIS is using a TIN to predict where the unsampled 30 m and 50 m lines should be. The transition of raster data to a TIN allows for a lot more analysis. The TIN allows the user to see and model aspects of the geography. Areas at risk from flooding can easily be viewed by putting in a flood depth and seeing which areas would be above or below the water. It is also important to remember that the TIN need not be referring to actual elevation data. It could easily be expressing variables such as temperature or rainfall.

The example in Plate 10 shows a real life use of this data. By using stereo imagery first in an image processing suite and then in a GIS application we can see how the data can be used to calculate changes in volume. The example uses data from the ER Mapper 6.3 examples tutorial. Even though image processing software allows a view of the data it does not allow an analysis of the data. For this we need to transfer the data into a GIS such as ArcView. Unfortunately, when raster data comes into the GIS the software does not necessarily know any details about the image. It simply represents the z-axis as unknown DN values. These DN values could be height, rainfall or just standard reflectance returns. First of all the GIS needs to be told that the data should represent individual points. This creates a grid as shown below in Figure 7-19.

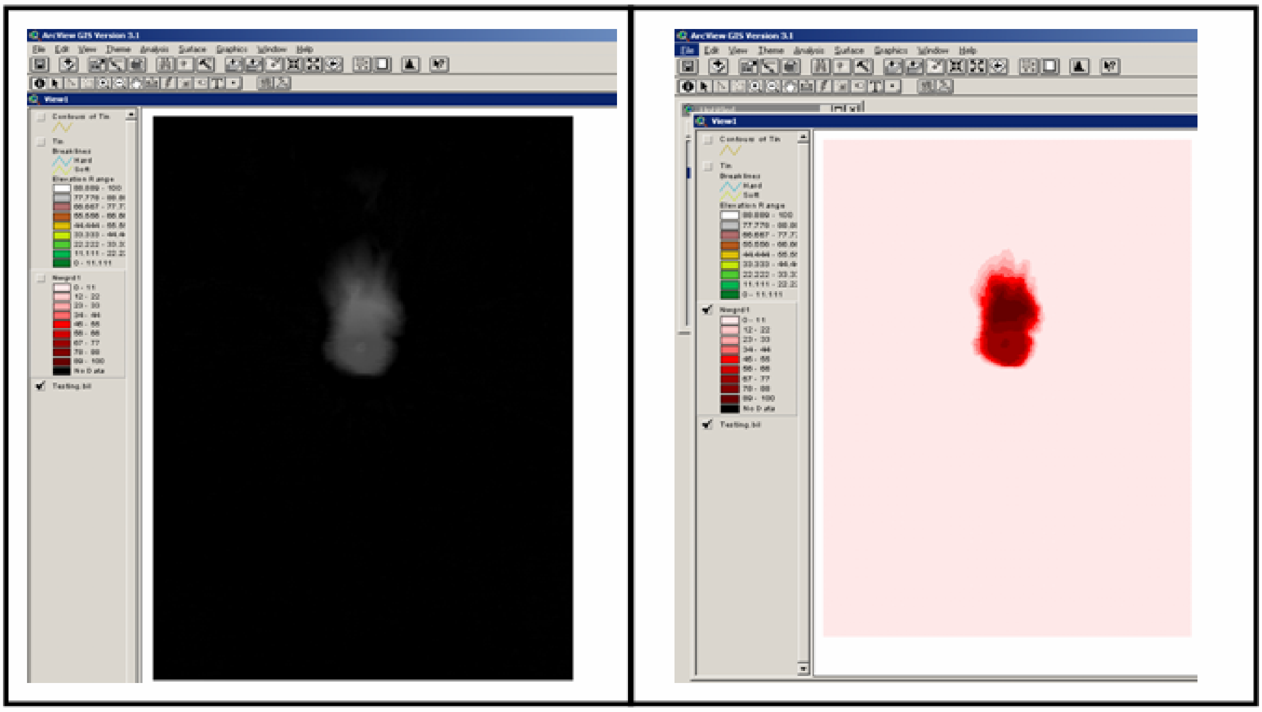
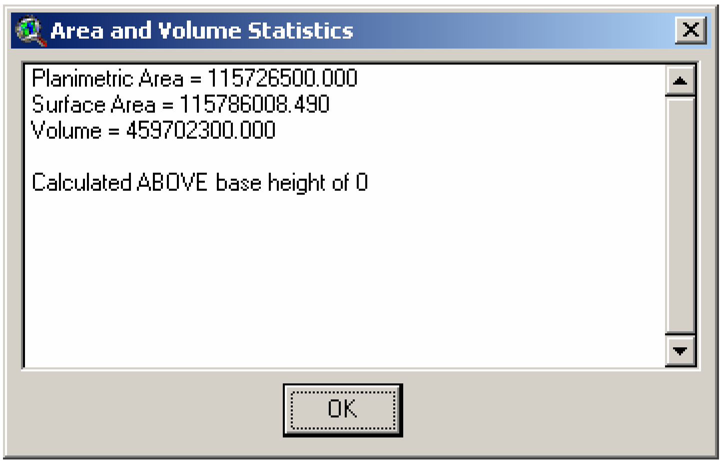


Figure 7-19 The raster image shown of the left is converted to a regular grid of points as shown on the right. This grid is then shaded at known intervals. This can then be used to generate a TIN for GIS analysis.

In the left-hand image the GIS is simply displaying a raster image of unknown DN values. By gridding the data we tell the GIS the data has intrinsic values that are important. A TIN can be built to connect the points and create a surface the GIS can query. We can ask a GIS to analyse a TIN in many ways. One interesting process in this example would be to find the volume of rock destroyed in the eruption. We can do that by using the area and volume statistics options from ArcGIS surface analysis tools. This generates the following dialogue box shown in Figure 7-20.



```
Planimetric Area = 115726500.000
Surface Area = 115786008.490
Volume = 459702300.000

Calculated ABOVE base height of 0
```

Figure 7-20 ArcView report based on Mt St Helens TIN.

The result of our analysis is 459,702,300 cubic metres of material or 0.46 cubic kilometres of rock. This shows the use of this type of data and hopefully the reader can see how these processes can be applied to many other applications.

Another similar application is measuring the volume of ice lost from a glacier. This can be conducted from a series of GPS points over time but is more easily done using stereo imagery. The glacial heights can be subtracted in a similar manner to the volcanic dome to give a volume of ice lost.

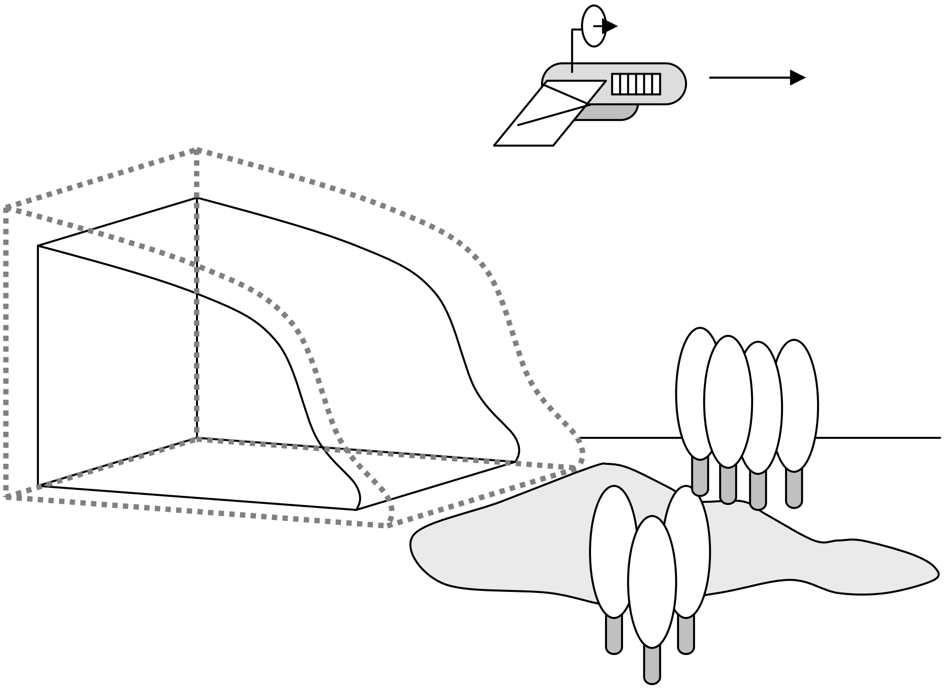


Figure 7-21 Schematic diagram showing glacial retreat in 3D.

The dotted snout in Figure 7-21 represents the level of the ice at some point in the past. The inner solid line represents the position of the ice today. The GIS can take points across the glacier, grid them and then calculate a TIN from points.

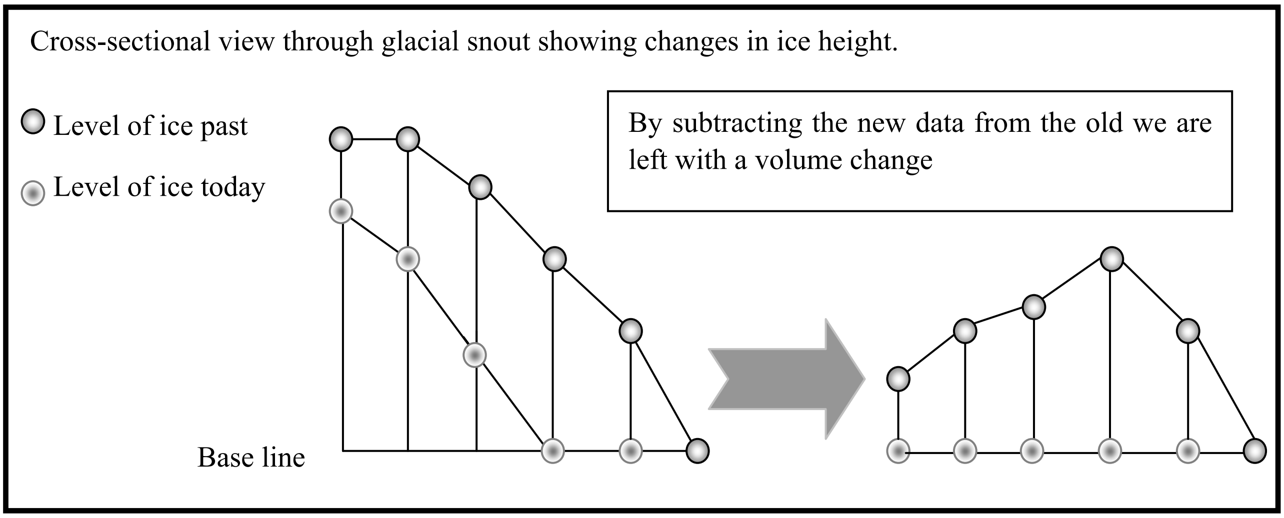


Figure 7-22 Subtracting new data from old calculates the volume change.

Some of the oldest available stereo satellite data is from the US Corona series (known as KH for Key Hole). High resolution KH4 data has a resolution better than 6 foot. This level of accuracy from the early sixties gives a temporal resolution of over 40 years and combined with modern GPS data or equally high quality stereo imagery such as IKONOS gives excellent results. Any stereo data can be used but using either ASTER 15 m pixels or SRTM 90 m data will degrade the results when compared to KH4 or aerial photography. An example of glacier data used in a study of this type is shown below in Figure 7-23.

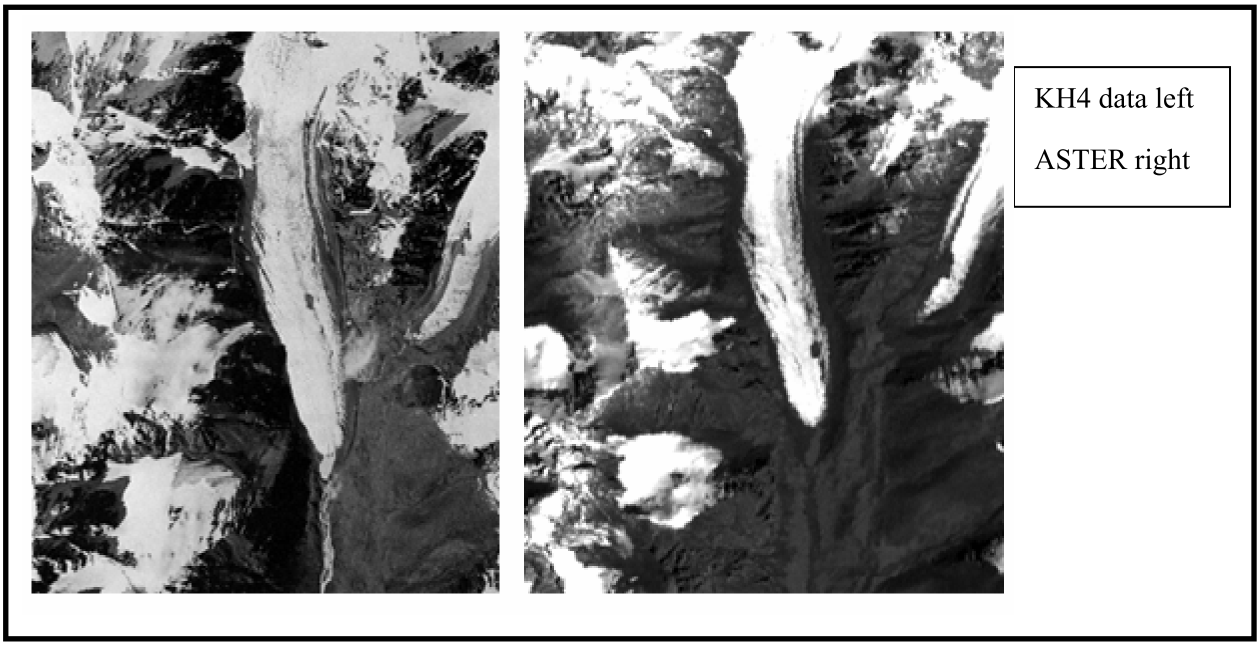


Figure 7-23 The imagery above clearly shows a change in the snout of the imagery. Example from Whiteside et al 2001 (b).

Using the highest possible resolution will always be beneficial to the expedition but the team should note that even theoretically accurate data like KH4 suffers from errors when rectified. Rectifying hardcopy data introduces considerable errors and the team should not count on getting results as accurate as modern IKONOS data from older hardcopy information.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section C: Techniques

Chapter 8: Image Interpretation and Processing

8 Image Interpretation and Processing

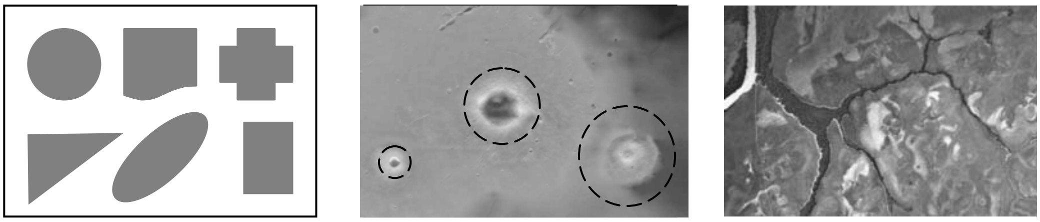
This chapter examines the two main approaches that we can use to add geographical information to a fieldwork-based survey: manual interpretation and automatic feature classification. The first to be considered is the manual delineation of features using image interpretation and conventional cartographic techniques. That is followed by a brief explanation of digital image processing techniques that can use variations in the spectral response of features to produce computer-generated maps.

8.1 Image interpretation

The features that our brains use when we interpret an image can be grouped into six main types, summarised below and in Figure 8-1:

1. **Tone**: variations in relative brightness or colour.
2. **Texture**: areas of an image with varying degrees of ‘smoothness’ or ‘roughness’.
3. **Pattern**: the arrangement of different tones and textures; may indicate certain types of geology or land use.
4. **Shape**: distinct patterns may be due to natural landforms or human shaping of the land.
5. **Size**: recognition of familiar objects allows size estimation of other features; size is an important aspect of association: for instance, a 20 km-wide circular surface depression is unlikely to be a sinkhole, but might be a volcanic caldera.
6. **Association**: the context of features in an image, e.g. a drainage pattern.

(1) tone (2) texture (3) pattern



(4) shape (5) size (6) association

Figure 8-1 Features used in image interpretation.

Tone, texture, pattern and shape are all influenced by illumination conditions and vegetation cover: they can therefore vary with time. Subtle features, such as buried archaeological structures, are best detected with the low sun angles and long shadows found at the start and end of the day, as well as during the winter season. Conversely, when mapping rugged terrain, try to use mid-day or summer imagery to reduce the loss of detail caused by valley-side shadows. Wet season and dry season variations in vegetation cover may be linked to variations in hydrogeology. The orientation, size and density of shadows can give clues about an area’s relief and degree of dissection, as illustrated in Figure 8-2.

The following features can be identified:

- *lava flows:* recent flows have the darkest tone; the oldest flow has a grey tone and the most fragmented texture (top right)
- *volcanic cones and craters*: discrete circular shapes and tonal contrasts with surrounding terrain
- *a geological fault zone*, possibly part of a fissure volcano (bottom left)

Figure 8-2 Interpretation of an airphoto from volcanic terrain.

8.1.1 Drainage patterns

The patterns produced by drainage networks are a useful guide to underlying soils and geology, as illustrated in Figure 8-3 and the subsequent summary table.

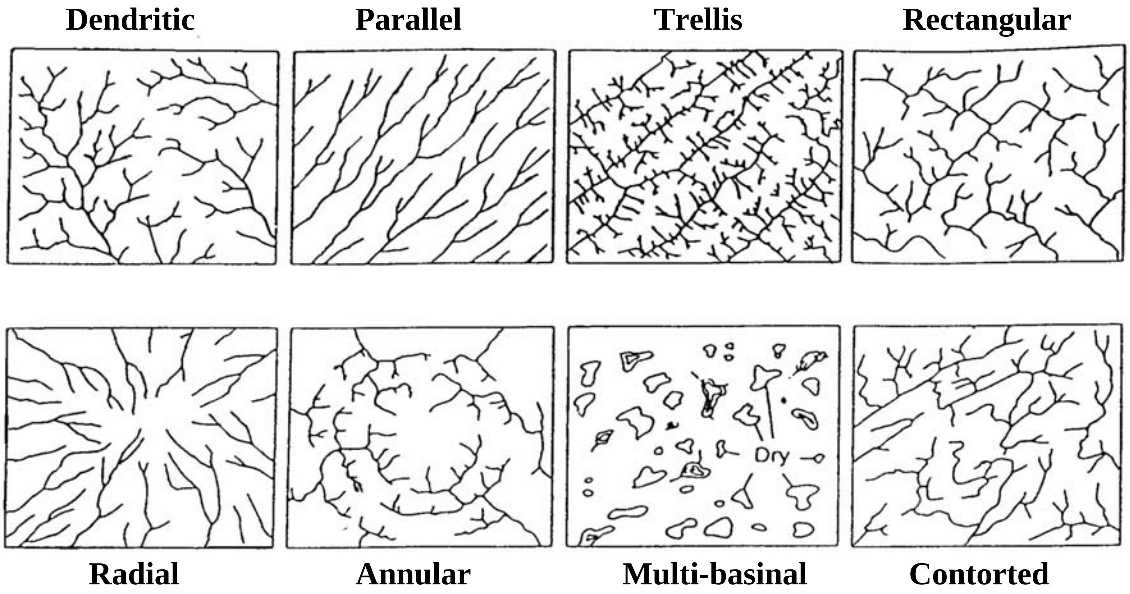


Figure 8-3 Common drainage patterns and their interpretation (after Howard 1967).

Pattern	Interpretation
Dendritic	Horizontal sediments or uniformly resistant crystalline rocks. Gentle regional slope at time of drainage inception.
Parallel	Moderate to steep slopes but also found in areas of parallel, elongate landforms.
Trellis	Dipping or folded sedimentary, volcanic, or low-grade metasedimentary rocks; areas of parallel fractures, exposed lake or seafloors with beach ridges
Rectangular	Joints and/or faults/fractures at right-angles. Lacks orderly repetitive quality of trellis pattern: streams and divides lack regional continuity.
Radial	Volcanoes, domes, and residual hills/inselbergs. A complex of radial patterns in a volcanic field might be called multi-radial.
Annular	Structural domes and basins, diatremes and possibly igneous stocks
Multi-basinal	Hummocky surface deposits; differentially scoured bedrock; areas of recent volcanism, limestone solution, and/or permafrost.
Contorted	Coarsely layered metamorphic rocks. Dikes, veins, and migmatized bands provide the resistant layers in some areas.

The most common drainage pattern is *Dendritic*, typical of relatively uniform, moderately well-drained soils and rocks: a variant pattern; and *Pinnate,* forms in easily-erodible silty deposits, such as wind-blown glacial loess. Dry valleys and sink-holes dominate in limestone landscapes, with most of the water flowing through cave systems, producing a *Dislocated* drainage pattern. Formerly glaciated terrain may have a *Deranged* drainage pattern, due to the melting of ice blocks within the glacial till, producing ‘kettle-hole’ lakes and a landscape with a poorly developed drainage network. Most of the other drainage patterns are strongly influenced by geological structures, such as alternate beds of gently-dipping soft and hard sedimentary rocks (e.g. *Trellis*); bedding and joint structures in metamorphic rocks or hard sandstones (*Rectangular*); joint structures in plutonic rock masses, such as granite (*Annular*); fault zones (*Elongate, Parallel*) and volcanoes (*Radial*). Detailed reviews of drainage pattern and image interpretation are given in Avery & Berlin (1992), Lawrance *et al* (1993) and Drury (2001).

Drainage texture or drainage density (the total length of channel in a given square kilometre) is a good indicator of the permeability of the soil and rock of the study area. Highly permeable substrates, such as limestone bedrock or sandy soils, have a low drainage density, as most of the surface water ends up underground. Conversely, relatively impermeable clay-rich soils and soft rocks such as shale or mudstone may have very high drainage densities, due to extensive overland flow and severe gully erosion. This produces highly dissected landscapes known as ‘badlands’ because of their unsuitability for farming. River rejuvenation or tectonic uplift can trigger headward fluvial erosion and enhanced gullying, leading to a higher drainage density.

8.2 Geomorphological mapping

The interaction between Earth-surface processes (i.e. weathering, erosion, deposition) and a few sub-surface processes (notably surface movements caused by earthquakes) tends to produce distinctive sets of landforms made up of distinctive materials. Examples are fluvial processes producing terraces made up of alluvium or sub-glacial processes forming drumlins made of boulder clay. Thus from our knowledge of geomorphology we can make a fair estimate of the formative processes and component materials of a given landform (Figure 8-4). This can be very useful in hazard assessment, as landforms may provide clues to the types of hazardous processes occurring in the study area, as well as the frequency and magnitude of hazardous events. Similarly, landforms may also give useful indications of earth resources, notably various sizes of aggregate associated with coastal, fluvial and fluvio-glacial deposits.

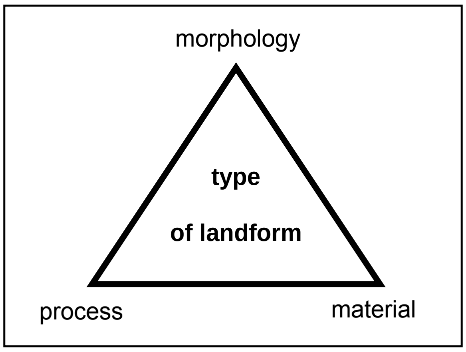


Figure 8-4 The 'Holy Trinity' of geomorphology: process, material and morphology.

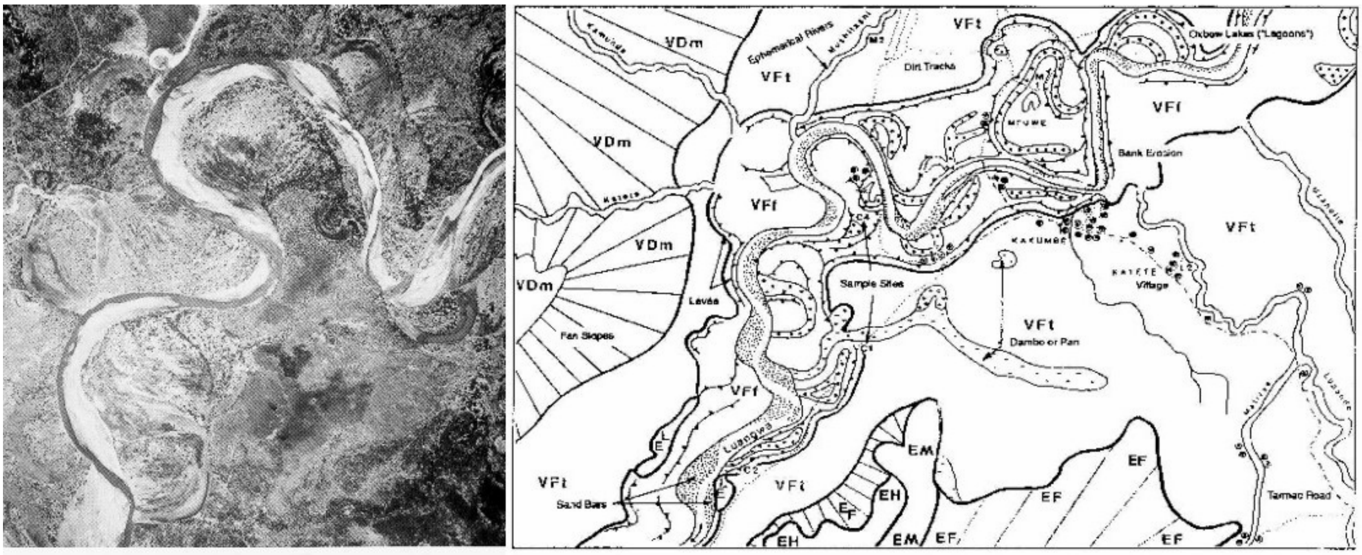
Geomorphological mapping is based on the identification of landforms or assemblages of landforms. This involves subjective assessment by the mapper, with the most reliable maps being produced by the most experienced geomorphologists (e.g., Brunsden *et al* 1975; Cooke & Doornkamp 1990; Fookes 1997). The best type of remote sensing data for detailed geomorphological mapping is vertical aerial photography, because of the general high degree of detail and possible stereoscopic (3-D) viewing. Image interpretation and fieldwork are iterative tasks: preliminary satellite or airphoto mapping precedes the initial field reconnaissance survey, with each stage of field survey producing additions or corrections to the image interpretation scheme.

Some satellite data, notably SPOT and IKONOS panchromatic imagery (with 3 m and 1 m pixels respectively), can be viewed and interpreted stereoscopically. Digital Elevation Models (DEMs) can be generated from the SPOT and IKONOS data, allowing geomorphological mapping, 3-D visualisation and 'virtual reality' fly-overs of study areas. The only down-side of the SPOT and IKONOS imagery is its relatively high cost. However, for c. £60 the ASTER sensor can provide a DEM covering 60 km x 60 km with 15 m pixels and 15 m contours, equivalent to a 1:50,000-scale map. For regional-scale studies, free 1:250,000 DEM data are available from the Shuttle Radar Topography Mission (SRTM).

One way to reduce to subjective/interpretive element is to limit a survey to *morphological mapping*: breaks of slope, amounts of slope and directions of slope are mapped using airphoto stereoscopy, but no attempt is made to interpret the origin and composition of mapped features. Taking things a step further, *morphometric mapping* relies entirely on the field mapping of slope breaks, steepness and aspect. Although this produces a map with a high degree of objectivity (and replicability), such 'walk-over' surveys take far longer than airphoto interpretation.

8.3 Mapping geo-ecological features

Geo-ecological mapping has been developed since the 1940s, when initial interest focused on using aerial photography to map the suitability of terrain for military purposes. By the 1950s, the Land Systems methodology had been developed to map vast areas of Australia and assess their suitability for agriculture (Christian and Stewart 1952). The different textures, tones and patterns displayed by different vegetation types are utilised in Land Systems mapping to map geo-ecological zones, each with a distinctive type of vegetation cover, soil, geology and hydrology. This process of 'mapping by proxy' allows large tracts to be mapped from remotely sensed images with just a few visits to check the 'ground-truth' at representative sites. The case studies feature an example of Land Systems mapping along the Luangwa River, Zambia, based on the interpretation of Landsat imagery and aerial photography (Figure 8-5). A number of other mapping systems are based on identifying geo-ecological features, notably that of the ITC (Verstappen & Van Zuidam 1975; Meijerink 1988) and Brunsden *et al*'s (1975) rapid geomorphological mapping techniques for civil engineering projects. Methods of aerial mapping and monitoring, from an ecological perspective, are reviewed by Clarke (1986).



LAND UNIT	LANDFORM	LAND ELEMENT	RELIEF m	SLOPE %	MATERIAL	VEGETATION	DRAINAGE	ERODIBILITY	U.S. SOIL TAXONOMY
Valley Unit V SubUnit F valley floor alluvium	Floodplain & Terraces. (VFf)	1. Channel	1-2	0	Silty	Water, shrub phragmites	Excessive	Very high	Non-soil/ Fluvaquents.
		2. Point bar		0-3	Sandy	Acacia thicket			Fluvaquants/Ustipsamments
		3. Levee		0-3	Stratified sand/loam	Munga woodland	Good	High	Ustipsamments Eutopepts
		4. Oxbow		0-1	Silt/clay	Water/grass	Poor	Low	Ustifluvents
		5. Backplain		0-1	Clayey	Short grasses			Vertic Eutropepts Pellusterts
	Higher terraces (VFt)	1. Channel	2-5	0	Stratified silt/clay	Grassland	Poor	High	Fluvaquents
		2. Interfluve		1-3	Clayey	Mopane & Short grass	Good	Very high	Pellusterts/Natraqualfs Haplustolls
					Sandy	Thicket			Ustipsamments/Natrustolls
		3. Backplain		0-1	Loam/clay	Mopane woodland	Poor	High	Pellusterts/Natraqualfs Haplustolls
		4. Pan			Clayey	Grassland			Fluvaquents/Pellusterts Haplaquolls/Natraqualfs
SubUnit D Dissected valley floor	Slightly dissected surface (VDs)	1. Backplain	2-5	0-1	Karoo sediments: sandstones, mudstones, conglomerates	Short Grasses	Imperfect	Moderate	Fluvaquents/Pellusterts Haplaquolls/Natrustalfs
		2. Pan/Dambo							Fluvaquents/Pellusterts Haplaquolls
		3. Interfluve	2-10	1-10		Miombo	Good	Moderate	Haplustalfs/Haplustults

Figure 8-5 Extract from a Land Systems map and classification system, derived from 1:40,000 aerial photography (inset): Luangwa Valley, Zambia (Teeuw 1990).

A brief summary of geo-ecological features that can be detected by remote sensing is given in Table 8-1.

Table 8-1 Summary of geo-ecological remote sensing applications (NIR, MIR and TIR are near, middle and thermal infra-red). Black = very useful, grey = may be useful, blank = not useful. See Figure 5.1 for the spectral ranges of NIR, MIR and TIR (after Leuven et al. 2000).

	AIRBORNE (cm to m detail)						**SATELLITE** (m to km detail)				
Water	visible	lidar	NIR	MIR	TIR	radar	visible	NIR	MIR	TIR	Radar
Aquatic plants	▨	▨	■				▨	▨			
Bathymetry	▨	■	▨			▨	▨	▨			
Flood extent	▨	▨	▨		■	■		▨		■	■
Ice cover	▨	■			■	■	▨			■	■
Pollution: - oil	▨	▨			■	■	▨			■	■
- chemical	▨				■		▨			■	
Temperature		▨			■					■	
Turbidity: - algae	▨		■	▨			▨	■	▨		
- plankton	▨	▨	■				▨	■			
Landscape	visible	lidar	NIR	MIR	TIR	radar	visible	NIR	MIR	TIR	Radar
Animal counts	■				■						
Archaeology	▨	■	■		■	▨	▨	■			▨
Channel changes	■	■	■				▨	▨			
Drainage pattern	■	■	■			■	■	■			■
Dykes / walls	▨	■	▨		▨		▨	▨			
Floodplain	▨	■	■		▨	■	▨	■		▨	■
Land: - built-on	■	■	■		▨		▨	■			▨
- contaminated	▨		■	▨	■		▨	▨	▨	▨	
Relief	■	■				▨	▨				▨
- surface change	■	■				▨	▨				■
- mm subsidence											■
Rock or soil type	▨			■	▨	▨	▨		■	▨	▨
Soil moisture						■					■
Soil / sed texture		■		▨		■			▨		■
Snow / ice cover	▨	■	▨		▨	■	▨	▨		▨	■
Seepage / spring	▨		■		■	▨		■		■	▨
Vegetation - type	▨	▨	■	▨		▨	▨	■	▨		▨
- woody biomass	▨	■	▨			■	▨	▨			■
- height	▨	■	▨								
- stress	▨		■				▨	■			
	visible	lidar	NIR	MIR	TIR	radar	visible	NIR	MIR	TIR	Radar

8.4 Human population estimates

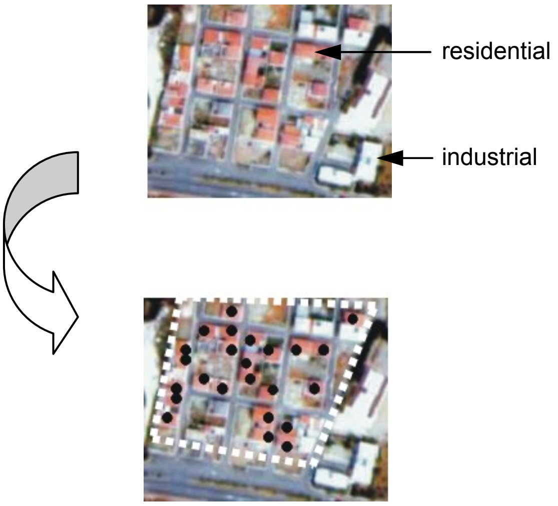
Population numbers and population densities are essential data for planners both in urban and rural settings. Comparison with earlier records allows an assessment of increasing or decreasing population trends. Areas suitable for development can be identified, along with areas that are already over-crowded. In civil emergency planning, airphotos are an effective way of identifying vulnerable populations, as well as evacuation routes. In many developing countries, airphotos may offer the best means of gaining population information (Figure 8-6).

Figure 8-6 How aerial photography can be used in estimating urban population densities.

1. Delimit homogenous housing area
2. Calculate area
3. Determine the number of buildings
4. Determine number of storeys/building

Built-up area = 1.1 ha
Number of buildings = 24.5
Number of flats/storey = 2
Average number of storeys = 2
Average family size = 5

5 x 2 x 2 = 20 people/building
20 people x 24.5 buildings = 490 people
490 / 1.1 = 445 people / hectare



8.5 Digital image processing

Scanned images are made up of Digital Numbers (DNs): the lower the DN values, the darker the resulting greyscale image, as illustrated in Figure 8-7.

2050	2044	2047	2061	2076	2087
2060	2059	2057	2060	2067	2074
2073	2070	2066	2064	2065	2070
2085	2084	2078	2073	2072	2073
2097	2098	2094	2088	2085	2084
2110	2113	2110	2105	2101	2100
2122	2127	2127	2122	2119	2118
2133	2140	2142	2139	2137	2137
2142	2151	2155	2155	2155	2156
2148	2158	2165	2168	2170	2173
2149	2160	2170	2176	2181	2187
2147	2159	2170	2180	2189	2198
2141	2154	2167	2180	2192	2205
2133	2147	2161	2176	2192	2207
2123	2137	2153	2169	2186	2203
2114	2128	2145	2162	2180	2198

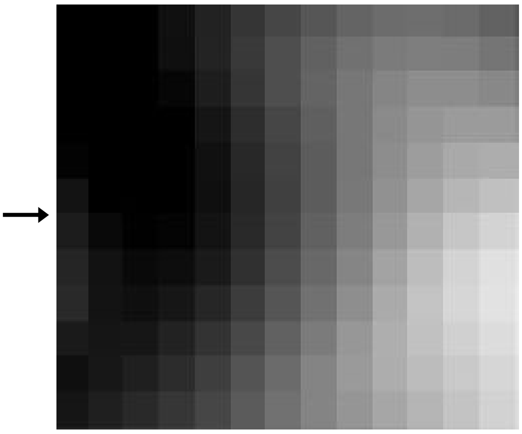


Figure 8-7 An example of DN values in a raster matrix, forming the greyscale image on the right.

Digital images have some major advantages over paper or film (analogue) images: they take up less storage space, perfect copies can be created time and time again, they can be reduced or enlarged at the push of a button, cartographic errors can easily be removed, and - most important of all – digital images can be processed using statistics, to enhance, analyse and classify their features. There are three main stages in the digital image processing of remotely sensed data: pre-processing, removing distortions from the imagery; image enhancement, highlighting features in the imagery; and image classification, producing maps of the features in the imagery.

8.5.1 Pre-processing

Distortions are inherent in remotely sensed imagery, whether from satellite or airborne platforms. Dust and water vapour in the atmosphere will absorb and reflect certain wavelengths of electromagnetic radiation, especially at shorter wavelengths and particularly in the blue part of the spectrum. Such *atmospheric distortions*, or attenuation, can be corrected by the 'black body' approach: clear water absorbs near infra-red and middle infra-red radiation, so its DN values in those wavelengths should be at or close to zero. If, say, the DN values for a middle infra-red band over water range from 18 to 27, then 18 could be subtracted from all of the DN values forming the raster matrix for that band, giving features over water 'truer' DN values ranging from 0 to 9.

Geometric distortions are either random or systematic. The former are difficult to correct and typically occur on aircraft buffeted by winds. The best solution is to 'rubber-stretch' the image, using landmarks distributed around the imagery that have exact locations (Ground Control Points or GCPs), known by map co-ordinates or GPS readings, onto which the distorted GCP locations on the imagery are stretched. Systematic distortions are relatively easy to correct, the commonest being the 'parallelogram effect' produced by the rotation of the Earth in the seconds that it takes an satellite sensor to scan 100-200 km of a given scene.

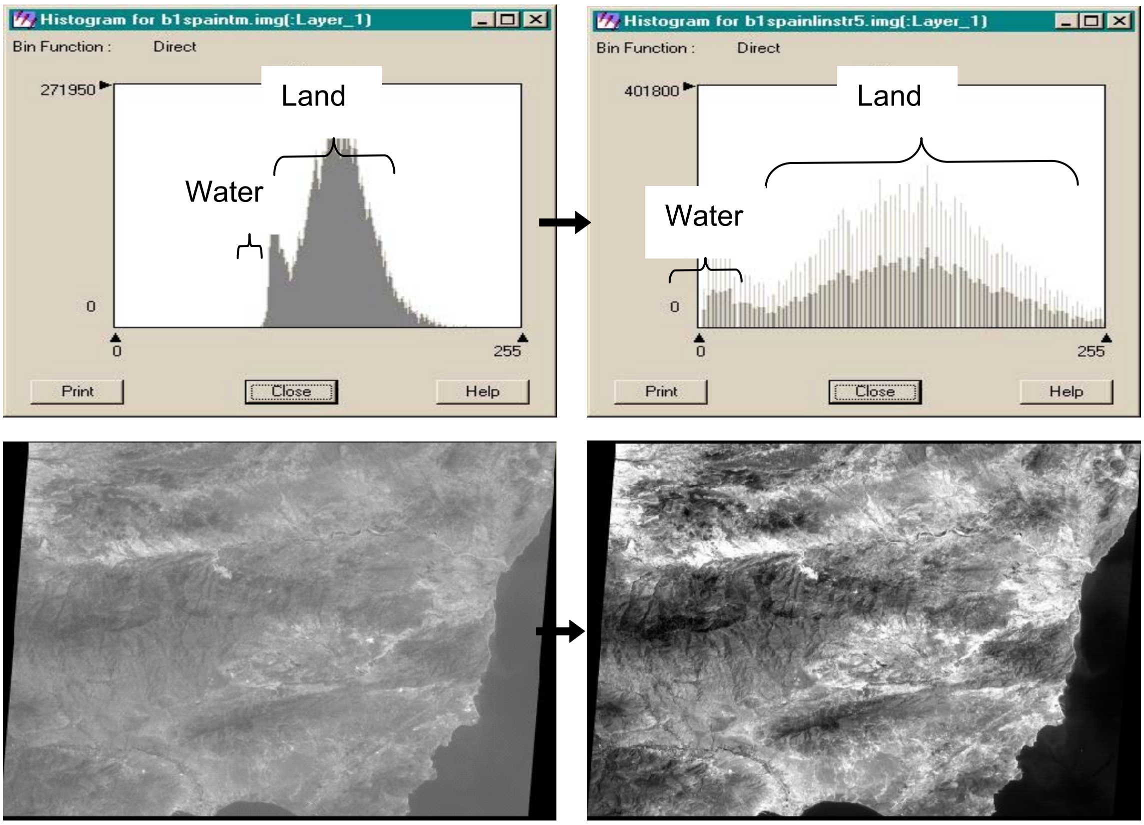
Sensor noise may be caused by faulty scanner heads or overloading of a given sensor, such as Landsat's thermal band, often resulting in a line of DN values with values of either 0 or 255. This can be corrected by substituting an erroneous value with the mean value from the pixels above and below it. Most modern imagery has already been pre-processed, removing geometric errors and geo-correcting the imagery to fit regional map projections. For more details, the reader is referred to the textbooks of Lillesand & Kiefer (2000), Drury (2001) and Mather (1999).

8.5.2 Image Enhancement

There are many ways to enhance digital imagery: contrast stretching, spatial filtering, colour composites, band-ratio images and advanced statistical techniques, such as Principal Components Analysis. The digital imagery used in the following examples is of Almeria province, along the Mediterranean coast of SE Spain.

8.5.3 Contrast stretching

This is a relatively straight-forward way of enhancing differences in the tone and texture of digital imagery. The process is illustrated by the graphs in Land Figure 8-8, each of which is underlain by the corresponding greyscale image. The illustrations on the left are of raw, unprocessed Landsat data: note that (i) the range of DN values in the histogram is narrow, ranging from about 100 to 180, and (ii) the corresponding image is relatively dark. The histogram on the right consists of the same dataset, re-sampled to cover the entire 0-255 range of DN values, resulting in an image with much better contrast. There are many types of contrast stretch: that used in Land Figure 8-8 is the Histogram Equalisation type, but various types of Linear Stretch can be used to mask or enhance features with distinct DN ranges, such as the low DNs associated with water.

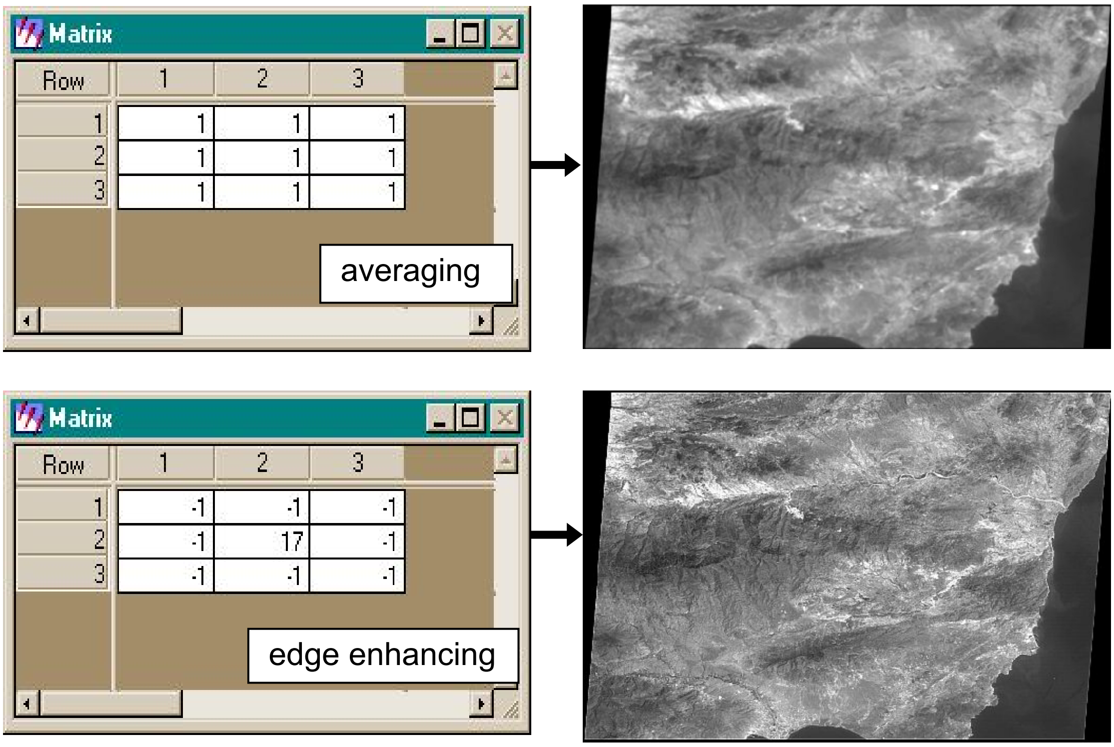

Histogram for b1spaintm.img(:Layer_1)

Bin Function : Direct

Histogram for b1spainlinstr5.img(:Layer_1)

Bin Function : Direct

Land Figure 8-8 Raw Landsat data (left) and stretched data (right).


Matrix

Row	1	2	3
1	1	1	1
2	1	1	1
3	1	1	1

Matrix

Row	1	2	3
1	-1	-1	-1
2	-1	17	-1
3	-1	-1	-1

Figure 8-9 Two types of spatial filter (left) and the resulting images (right).

8.5.4 Spatial filtering

A digital image can be modified by moving a spatial filter, or kernel, through the matrix of DN values, rather like using a moving average. The top pair of illustrations in Figure 8-9 shows a smoothing filter and the resulting image (compare with the original contrast stretched image in Land Figure 8-8), in which each pixel has been given the average DN value from all eight adjacent pixels. The lower set of illustrations in Figure 8-9 shows the results of the opposite process, in which differences between adjacent pixels are enhanced, sharpening the image.

8.5.5 Colour composite images

False-colour images are the most widely used product of image processing: the saying goes that "a picture can tell a thousand stories" – even more so if it is in colour. By allocating three of the bands (i.e. wavelengths) from a scanned image to the blue, green and red colour-guns of a computer screen, a false-colour image is produced, as illustrated in Plate 11. A list of composite image band combinations and the resulting highlighted features, using Landsat TM and ETM imagery, is given in Table 8-2 and Table 8-3.

A 'true colour' composite image is produced by allocating the R, G, and B colour guns to the R, G and B wavebands of a given set of multi-spectral imagery: using Landsat TM that would correspond to bands 3, 2 and 1, with the resulting image described as 'RGB321' in remote sensing short-hand. Creating 'false colour' images is very useful, as it allows us to view images captured in parts of the spectrum that would otherwise be invisible to our eyes, such as infra-red or ultra-violet. 'False Colour Infra-Red' (FCIR) images are particularly useful, as they allow us to view the pronounced variations in Near Infra-Red (NIR) reflectance associated with photosynthesizing vegetation. Using Landsat TM, FCIR images are created by allocating the R, G and B colour-guns to bands 4 (NIR), 3 (R) and 2 (G) respectively (i.e. RGB432). FCIR images look a bit odd at first, as water (which absorbs NIR radiation) is black, vegetated areas are shades of red and non-vegetated areas are shades of light grey, but an inspection of the composite images in Plate 11 illustrates how relatively subtle variations in vegetation cover are enhanced when using FCIR. Table 8-2 and Table 8-3 below summarise the uses of various colour composite images - using Landsat and ASTER imagery - for detecting variations in types of land cover, soils, rocks and minerals.

Table 8-2 Some useful waveband display combinations for Landsat TM and ETM+ (view as RGB colour images).

R G B	Highlighted Features
3 2 1	'True colour' image
4 3 2	False-colour infra-red: vegetation chlorophyll shows as red
1 3 5 2 4 5 7 3 1	Vegetation and soils
4 5 3 4 5 6 4 5 7	Soils and rock types
5 4 1 5 3 1 7 4 1	Hydrothermal alteration of volcanic rocks
7 5 1	Discrimination between Fe-rich soil and rock
5 3 1 6 7 2 2 3 4 6 5 4	Useful with various vegetation and rock types
6 7 5	Urban/rural boundaries
1 7 4	Coastal sediment plumes and variations in land cover types
7 4 2	Coastal features (- best with linear stretching)

Table 8-3 Some useful waveband ratios for Landsat TM and ETM (best viewed in greyscale)

Band ratios	Highlighted features
4/3	Vegetation Index: responds to green biomass, chlorophyll content and leaf water stress
(4-3) / (4+3)	Normalised Difference Vegetation Index (NDVI): responds to green biomass, chlorophyll content and leaf water stress. NB not suitable for regions with <30% vegetation cover.
5/4	Infrared Index: responds to changes in green biomass and water stress better than NDVI.
5/7	Responds to soil moisture content; gypsum and different types of clay
3/1	Responds to Fe-rich soil or rock
4/5	Separates hydrous rocks from Fe-rich rocks

Table 8-4 ASTER band combinations and their uses in geological mapping. Band ratios are best viewed as greyscale; band composites in RGB colour.

ASTER bands	Highlighted features
468 or 631	Useful for general geological mapping
321	False Colour InfraRed (photo-synthesising vegetation shows as red)
2/1 or **((2/1) 3 1)** **4/1 or 4/3**	Ferric oxides
5/3 or **((1/2)+(5/3))** **7/4**	Ferrous oxides (7/4 applies to ferrous oxides in carbonate or silicate rocks)
456 **(4+6) / (5x2)** **5/6 7/6 7/5**	Hydrothermal (argillic) alteration
4/3 2/1 6/4	Regolith mapping (ferricrete / laterite, saprolite)
(7+9) / (8x2) **13/14** **(6+8) / (7x2)**	Carbonates (nb. the third combination highlights dolomitised carbonates)
11/10, 13/12 **14/12** **11/(10+12)**	Quartz
13 12 10 **14 12 10**	Quartz, carbonate & mafic minerals Quartz, basic minerals & saline soils
12/13	'Basic minerals': garnet, pyroxene, epidote, chlorite

8.5.6 Band ratios

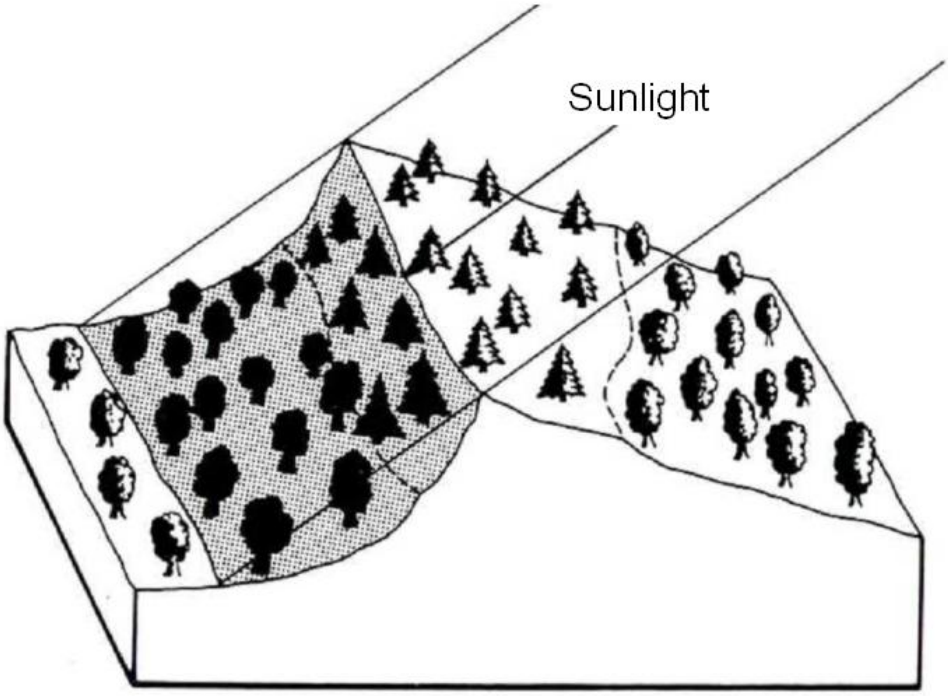
With a given multi-spectral dataset, if you divide all of the DN values in one band by all of the DN values from corresponding pixels in another band, certain features that either strongly absorb or strongly reflect radiation in the two bands will be highlighted. For instance, dividing Landsat TM band 5 by band 3 highlights areas of gypsum-rich soil and certain clay minerals associated with hydrothermal activity in volcanic terrain. Table 8-4 summarises the various types of land cover, rocks and minerals that can be detected using band ratios with Landsat and ASTER imagery. Another useful aspect of band ratio images is that the effects of shadows are eliminated, as illustrated in Figure 8-10. The pronounced difference between the relative levels of red and near infra-red radiation reflected by vegetation can be used to create an index of vegetation abundance, the most widely used example being the Normalised Difference Vegetation Index (NDVI):

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$$

With Landsat TM or ETM+:

$$\text{NDVI} = (\text{Band 4} - \text{Band 3}) / (\text{Band 4} + \text{Band 3})$$

NDVI has been used very effectively with regional weather satellite imagery of Amazonia to quantify and monitor the loss of rainforest since the 1970s. The use of NDVI in semi-arid regions, as in Figure 8-11, should be treated with caution, due to large amounts of bare soil and rock that will influence the reflectance values. Specially modified vegetation indices for use over arid regions have been developed and are reviewed by Lillesand & Kiefer (2000).



Land cover	**Illumination**	**Digital number**		
		Band A	**Band B**	**Ratio (A / B)**
Deciduous	Sunlit	65	72	0.90
	Shadow	19	21	0.90
Coniferous	Sunlit	32	48	0.67
	Shadow	10	15	0.67

Figure 8-10 Use of band ratio processing to remove the effects of shadow (after Lillesand & Kiefer 2000).

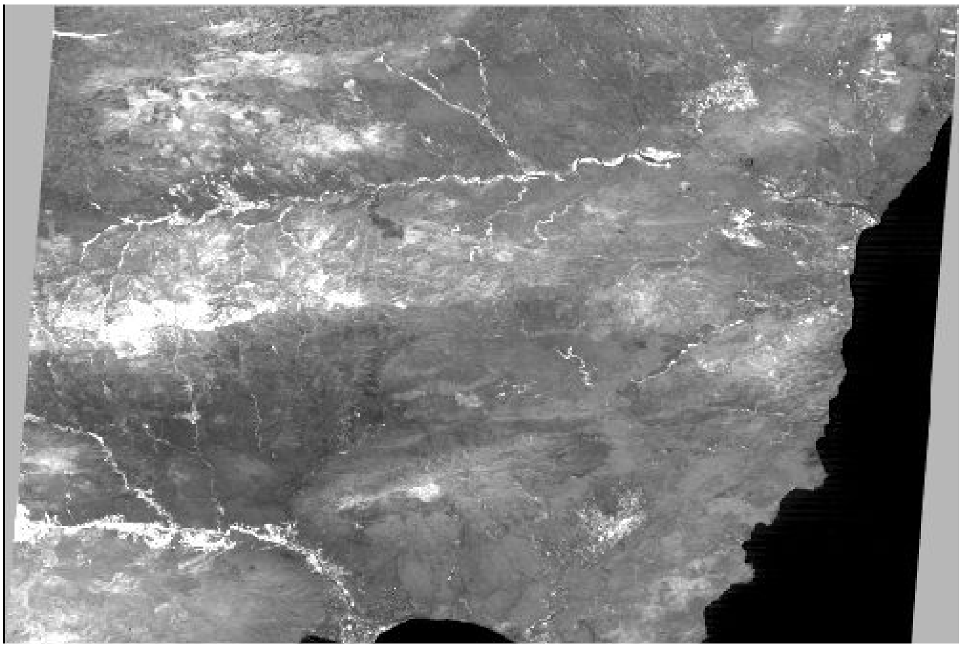


Figure 8-11 NDVI image of Almeria province, SE Spain. Areas with abundant vegetation show as white. This region contains the largest expanses of semi-arid and desert terrain in Europe. Most of the vegetation is limited to river valleys (bottom left and upper half of the image) and the summits of mountain ranges, notably in the top left quarter of the image.

8.5.7 Image classification

The subdivision of Earth surface features into different types on the basis of their spectral responses is known as classification. There are two types of classification: unsupervised and supervised, with the latter being more complicated but also potentially more accurate.

Unsupervised classification is carried out by the image processing software without any initial input, or 'training', from the user. The process can be illustrated with reference to Figure 8-12. Diagram A shows four different land cover types, each with a distinct spectral signature. The differing spectral responses can be more effectively distinguished by plotting the DN values from the bands of the original Landsat imagery against each other, as illustrated in Figure 8-12B. This produces distinct clusters of DN values, along with a modified image on which each cluster is colour-coded. The user then has to allocate each cluster to its corresponding land cover type, providing a legend, thereby turning the colour-coded cluster analysis image into a land cover map.

Supervised classification requires the user to select training areas containing about 100 pixels of each land cover type: these pixels are used to 'teach' a computer to recognise the spectral responses of each land cover type. A legend for the ensuing land cover map is built up as the user inputs the training areas for each land cover type. The software then uses the training areas to derive statistical summaries of each land cover type's spectral response, from which it goes on to classify all of the remaining pixels in the image. The purer the sample of pixels in the training area for each land cover type, the better the accuracy of the ensuing classification.

Classification based on a spectral reflectance from a single band and then from two bands is illustrated graphically in Figure 8-12. Including a third band in the classification routine will improve the discrimination between land cover types, as illustrated in the 3-D graph in Plate 13. Using multispectral or hyperspectral imagery, classification software can utilise *n* bands in *n* dimensions, giving increasingly better separations between land cover types – in some cases allowing automated mapping based solely on the varying spectral responses along scan lines. The maths behind image classification can be complex: the reader is referred to Lillesand & Kiefer (2000), Drury (2001), or Mather (1999) for useful summaries

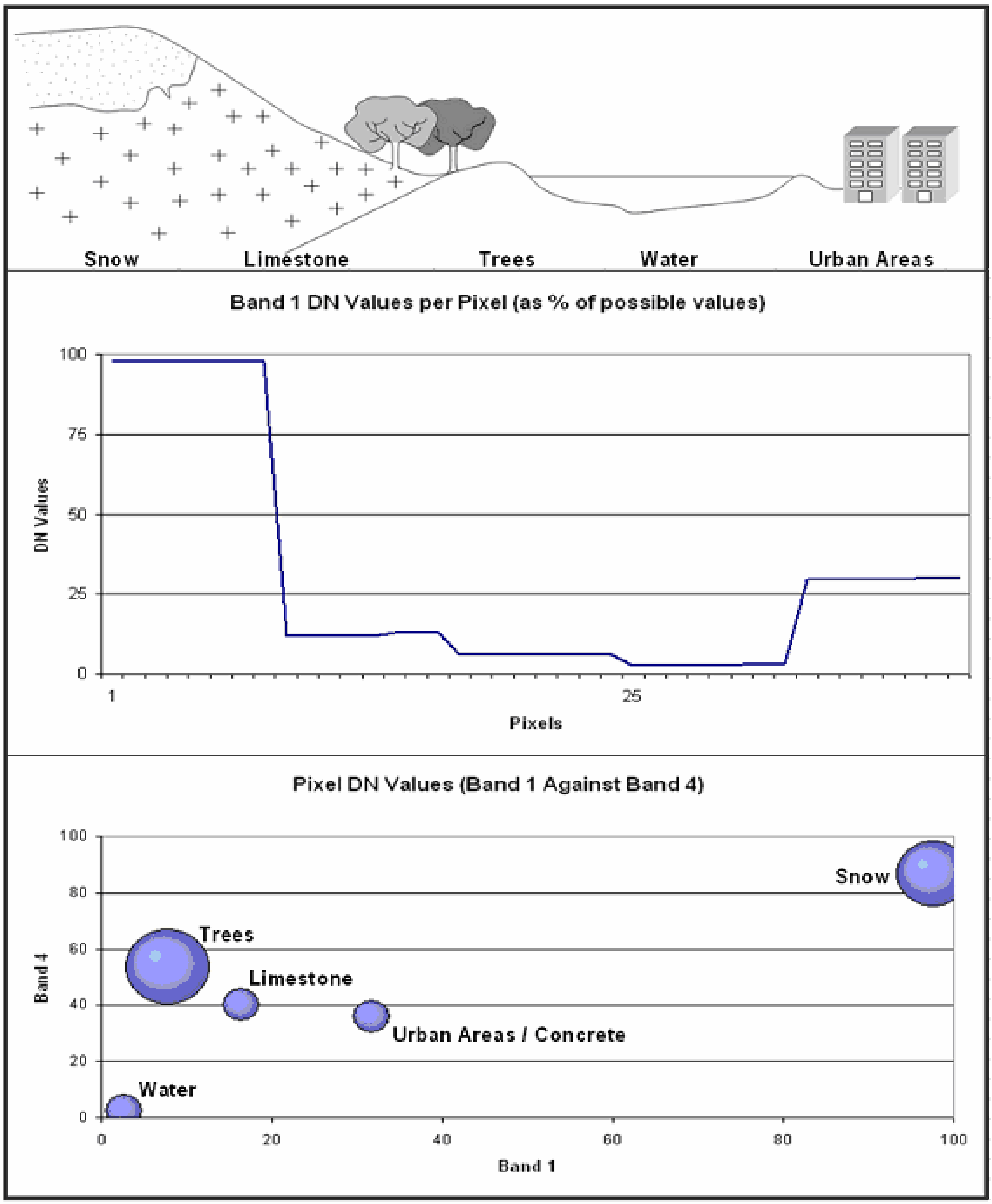


Figure 8-12 Digital image classification, for an idealised landscape (top). Upper graph shows variations in Band 1 reflectance along a scan line: note that the high values for snow and low values for water. The lower graph is a plot of reflectance values for Band 1 against those for Band 4: note the distinct clustering of values for each of the land cover types.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- ## Section C: Techniques

Chapter 9: Geocorrection and photogrammetry

9 Geocorrection and Photogrammetry

Aerial photographs are, arguably, the most widely used form of remote sensing. Airphotos have been used in the production and update of almost all topographic maps since the 1950s in almost every country in the world. It is therefore highly likely that photographs already exist of the area you intend to visit, wherever that happens to be. Unfortunately, you may not be able to access airphotos of your study area, most commonly because of military sensitivity, but perhaps due to other factors, such as local government policy, copyright issues and depleted airphoto archives. Further problems encountered by the editorial team on expeditions in developing countries include a lack of chemicals to develop prints and termite damage to airphoto archives. The scale, quality and age (date acquired) of the photography could also be very variable, again dependent on why the photographs were taken.

Measurement from aerial photographs is called *photogrammetry*. The term is derived from three Greek root words and literally means *light-writing-measurement*. The science of photogrammetry is solely concerned with collecting precise and accurate measurements from photographs. Photogrammetry is one of the more complex GISci subjects, so this chapter will focus on practical aspects that are of use in field work. For more in-depth coverage of photogrammetry theory and techniques, the reader is referred to Avery & Berlin (1992), Lillesand & Kiefer (2000), Mikhail *et al.* (2001) and Wolf & Dewitt (2000). A review of software with photogrammetric capabilities is given in the GISci software chapter.

Most aerial photographs are acquired using specialised mapping cameras and are acquired with the camera pointing vertically down from the aircraft. Other photography may exist, notably oblique photos, such as ‘tourist snaps’ from aircraft windows or non-vertical specialist mapping-camera shots, but these are not as useful for mapping as vertical photography. An example of an oblique database is that held by English Heritage – these were primarily acquired for archaeological investigations. The reason why most photographs are acquired vertically and with specialised cameras is so that measurements of objects and areas on the Earth’s surface can be taken with the minimum of calculation and correction for distortion. Distortion is still present in the photography - as it is in all photography, terrestrial or airborne - but the vertical viewpoint reduces it to manageable proportions.

Aerial photography is typically acquired with significant overlap between images. This is done to produce views of the Earth’s surface from two different perspectives to allow ‘stereo-viewing’. When one image is viewed with your left eye and the other with your right eye, your brain integrates the images and the effect is a 3-D view of the area.

9.1 Data from space-borne sensors

A number of satellite systems exist that can produce imagery suitable for stereo-viewing. Notable systems are Corona (1960s-1970s US spy satellite images, many now declassified), SPOT (French satellite series, collecting data since the mid 1980s), IKONOS (since 1999, 1 m resolution commercial satellite series from US) and ASTER (since 2000, Japan/NASA collaboration). NASA have an archive of colour photographs from the Metric Camera (10-20 m pixels) taken during orbits of the Space Shuttle. ASTER and SPOT data can also be processed digitally to derive elevation information (DEMs) using photogrammetry; however this is a complicated and time-consuming process that requires access to expensive software for the best results. A much cheaper and quicker means of getting DEMs of your study area is to use data collected by the USA's Shuttle Radar Topographic Mission (SRTM). The SRTM DEM coverage is world-wide and can yield contours at 8-10 m intervals with 90 m pixels (30 m over the USA – hopefully this more detailed coverage will eventually be available beyond North America). One of the best things about the SRTM data is that it can be downloaded free via the internet (either from the GLCF site or from the US Geological Survey the latter giving better accuracy and being easier to mosaic together – see the Appendix and Weblinks section for contact details). Furthermore, software to process the SRTM DEMs, producing 3-D views and virtual fly-overs is also freely available via the Internet. A tutorial detailing how to download SRTM data from the GLCF website, plus how to download the *3-DEM* software and use it to generate DEMs from the SRTM data, is included in the CD that accompanies this handbook.

Many satellite data suppliers provide images that are geo-located, but the accuracy of this positional information should always be verified independently - the author has personal experience of corrected radar data with positional errors of over 300 m - and other sources have found positional errors of up to 1 km with ASTER imagery. Landsat ETM+ imagery, derived from the ESDI site (see appendix for web link and further information), is generally accurately located and can actually be used to correct similar data (such as ASTER, satellite radar, etc) to within 50 m, using image-to-image geocorrection with the Landsat image treated as the 'true position' map reference dataset. Although focused on aerial photography, this chapter provides an introduction to the techniques that you could use to make precise measurements, or generate DEMs, from all types of stereoscopic imagery, be they airphotos, spy satellite photos or satellite scanner images.

9.2 Photogrammetry

9.2.1 Perspective distortions

As mentioned earlier, all photography contains distortions. Unfortunately, these distortions, which are caused by the camera lens, the topography of the imaged surface (the Earth's surface) and the single-point instantaneous collection nature of photographs, mean that aerial photos cannot be used directly as maps. Attempts to derive distances, areas or angular relationships directly from photography will result in serious errors. Photogrammetric analysis attempts to account for these distortions to allow distances, areas and angles to be derived accurately and precisely. Most of the distortion present in photography is the result of the perspective viewpoint from which the photograph was taken. This perspective is essential, however, for stereo-viewing and is necessary for calculating elevation from photography. The perspective distortion in a vertical aerial photograph increases away from the centre of the photograph. This is because the centre of the photo (correctly called the principal point) is the only truly vertical part of the image. Figure 9-1, Figure 9-2 and Figure 9-3 illustrate this.

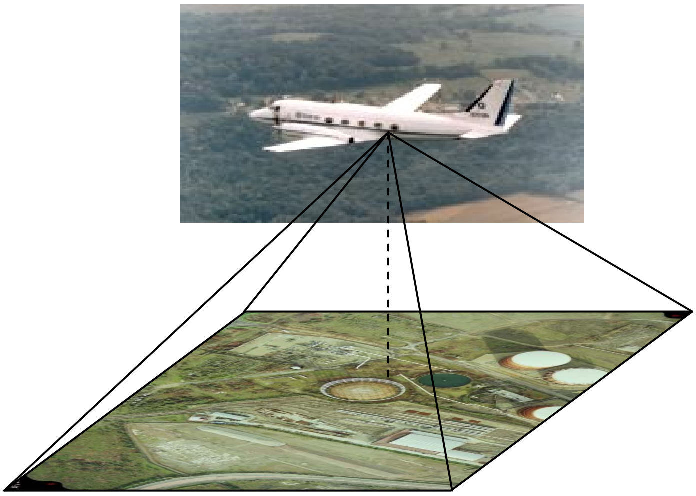


Figure 9-1 Aerial photo perspective and principal point (the only truly vertical point on the photograph is shown by the dashed line).

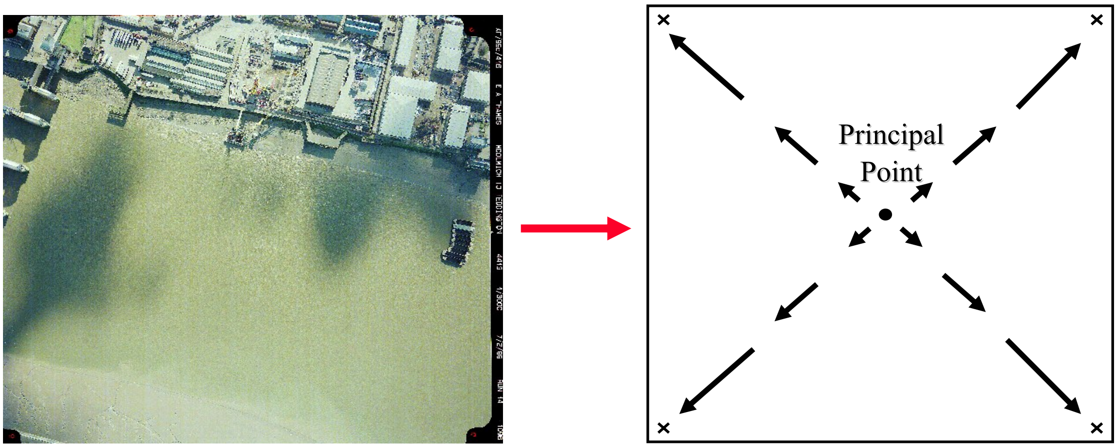


Figure 9-2 Radial distortion in a photograph caused by photographic perspective. Crosses in the corners are the fiducial marks (see Figure 9-7).

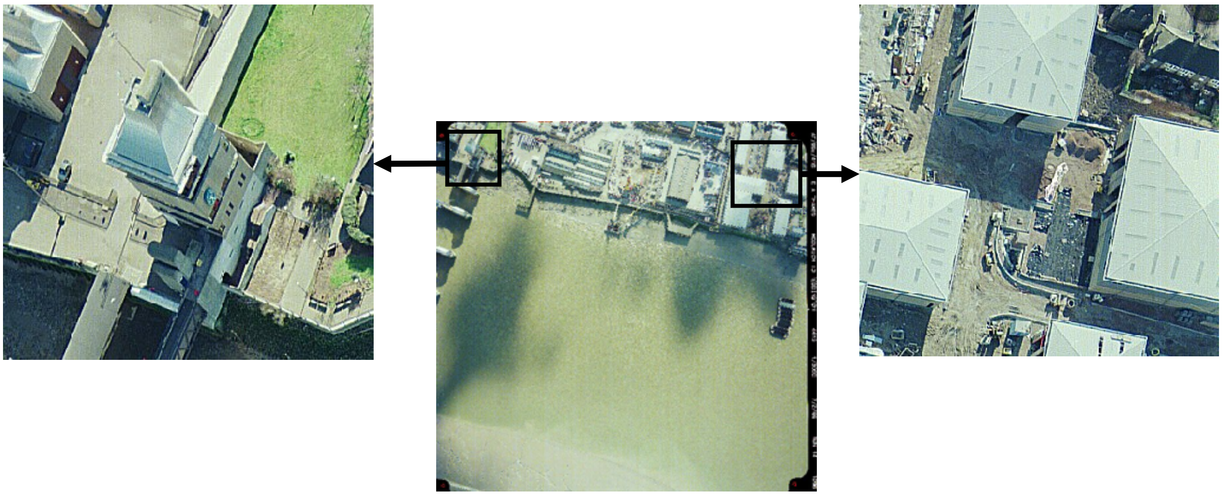


Figure 9-3 Radial distortion effects on buildings. Note how the buildings appear to lean away from the principal point (the photos are of the Thames Barrier, the river Thames and surrounding areas)

The effect of perspective distortion is that vertical buildings, hills and trees appear to 'lay-over' on their side as shown in Figure 9-3. At the principal point of a vertical air photo, the only part of a building that is visible is the roof. The roofline is actually what is used to define a building on a planimetrically correct map. The further away from the principal point a feature is, the larger the ‘layover’ effect gets as the photo becomes progressively closer to horizontal (Figure 9-2, Figure 9-3). An interesting result of this is that tall buildings, trees and other structures will appear to radiate away from the principal point of a vertical air photo.

The effects of terrain and radial distortion are illustrated in Plate 14, using a home experiment set-up. Three of the glasses in Plate 14 (upper left, lower left and lower right – 3(b)) are equally spaced, with 20 cm between their centres and the centre glass. The fourth glass (upper right – 3(b)) is 15 cm from the centre glass. In the vertical photo, the glasses appear to radiate outwards from the centre of the image, with the closer glass apparently less tilted and the top of the raised glass apparently further away. All these effects are systematic and can be used to help measure features and objects in three dimensions from vertical photography.

9.2.2 Parallax explained

On successive photographs, different sides of a particular feature or building may be visible. A tall feature may also ‘layover’ in a completely different direction on two successive photos. This effect is due to the change in location of the aeroplane between photo exposures. The apparent change in the relative positions of the bottom and top of a non-horizontal feature is called parallax. Parallax between two images is illustrated in Figure 9-4.

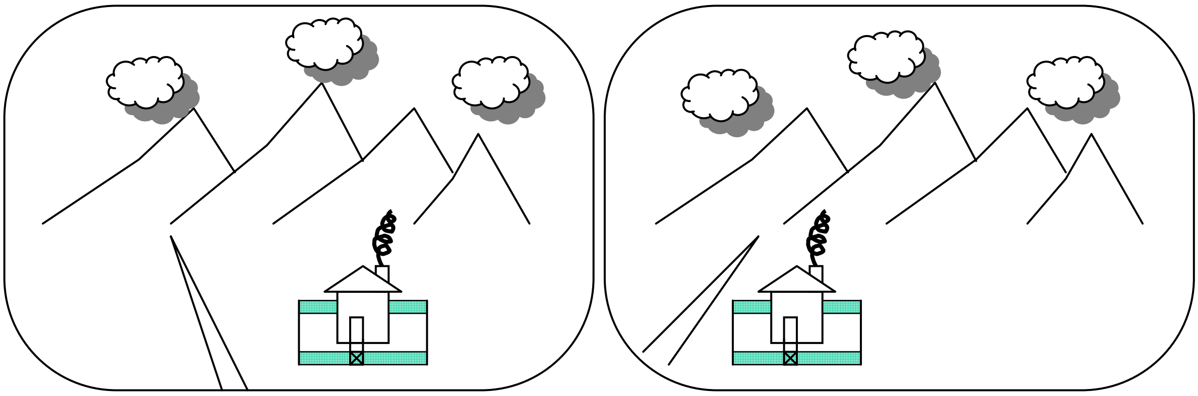


Figure 9-4 Parallax in action: Relative position change of objects from a moving train.

Imagine you are looking out of the window of a moving train as it traverses a plain with mountains visible in the distance. Visible in the right foreground is a house on the plain. You close your eyes for a few seconds and then open them again. The view is similar, but the farm appears to have changed from right to left of the tallest mountain in the distance, because of the movement of the train that you are travelling on. This is a change in parallax. Objects closer to the observer move through the observer’s field of view faster than more distant objects. This same effect occurs with aerial photographs, where the tops of buildings, trees or hills appear to move further across successive photos than their bases do. This difference in apparent movement is extremely useful, as it is directly correlated with the height difference between top and bottom. The higher an object is, the closer its top is to the aeroplane and the further it will appear to move, relative to its base, on successive photos.

The outcome of this effect is that by determining the difference in distance between the top and bottom of a feature, we can calculate how high the feature is. We need some other information to do this accurately, notably how high the camera was when the photos were taken, how high the base of a feature is above sea level and the focal length of the camera, but this is all relatively easy to collect, from field survey, mapping and/or the photo title strip (viz. Figure 9-6).

Figure 9-5 Example of a Black and White (panchromatic) aerial photo with title strip information.

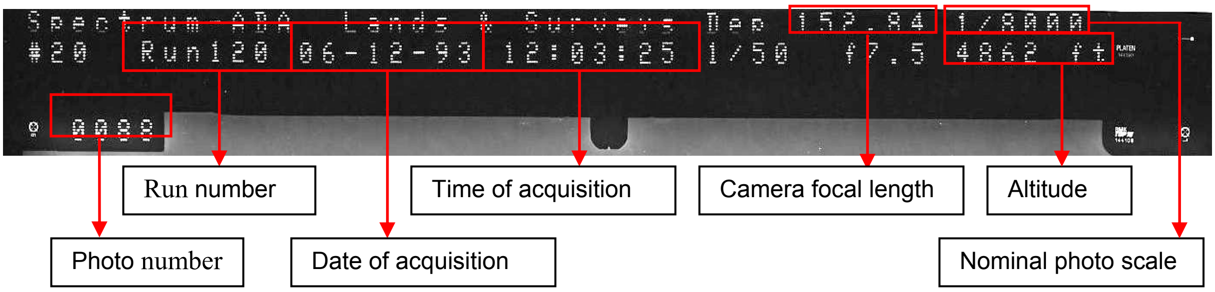


Figure 9-6 Zoom of the titling strip for the photo shown in Figure 9-5.

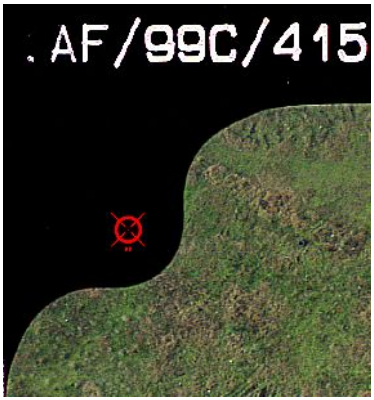


Figure 9-7 Example of a fiducial mark at the corner of an aerial photograph. There can be four or eight of these marks at the corners and in the centre of edges of photos, dependent on the camera make and model.

9.2.3 Scale in aerial photography

Perspective distortion means that photographs cannot have an accurate scale across an entire photo. The scale of an imaged object is dependent on how far away it is from the camera. Nevertheless, scale is widely quoted for aerial photography, but a user should be aware that a quoted scale is a nominal value – it gives an indication of the general level of detail and coverage of a photo, but should not be treated as a measurement ratio, as it can be with a map scale. Measurements derived using photo scales are likely to contain significant errors. If a scale for the photography is not given on the photo, or provided by the data supplier, a nominal scale is easy to calculate. There are a number of methods that can be used, but the most commonly used are:

- Direct comparison of the ground distance between objects visible in the imagery, usually calculated using a topographic map
- Calculation based on camera focal length and aeroplane elevation above the land surface

The first approach uses information from the airphoto and relates it either to a topographic map or to known real-world distances. The technique uses the following equation:

Equation 1:

$$S = \frac{Dp}{Drw}$$

Where: S = Scale of airphoto (nominal only)

Dp = Distance between objects on the photo

Drw = Distance between objects in the real world

When using the equation above, care must be taken to ensure that the numbers used for Dp and Drw are in the same units. As an example of how this equation might be used, imagine a photo where two path junctions are visible and also shown on a topographic map, the distance between the path junctions on the photo being 17.5 cm. The distance between the path junctions, as measured on our 1:50,000 map, is 2.4 cm. The equation then becomes:

$$S = \frac{17.5cm}{120{,}000cm} = 6{,}857$$

So our air photo has a nominal scale of 1:7,000, assuming the path junctions are at about the same topographic height (this could be checked from the map).

The second approach will give a nominal scale derived from three pieces of information, using equation 2:

Equation 2:

$$S = \frac{f}{H - h}$$

Where: S = Scale of the airphoto (nominal only)

f = Focal length of the camera

H = Aeroplane altitude, usually above Mean Sea Level (MSL)

h = Average altitude of terrain, above MSL

The aeroplane altitude and focal length of the camera are usually shown on the title strip of airphotos taken with a mapping camera. Because of the nature of these cameras, the focal length is not generally adjustable (as it is with most personal cameras), because the target is usually far enough away for the focal length to be set to infinity. The altitude should be recorded directly from an altimeter – if it is hand written, it may well be wrong! The average altitude of the terrain is sometimes difficult to decide, but an educated guess (based on a map or local knowledge) will allow photography of the most remote areas to be given at least a nominal scale. A great deal of care must be taken with units when using this equation as there are typically many conversions necessary. An example of this would be if (i) the focal length of a camera was given as 152.4 mm (a typical mapping camera, used for professional surveys, has a focal length close to this), (ii) the aircraft altitude was 4000 feet (4000 x 3.25 = 1231 m), and (iii) the average surface elevation was 150 m. Below is equation 2 with all values converted to metres:

$$Scale = \frac{0.1524}{1231 - 150} = \frac{0.1524}{1081} = 7093$$

So, the nominal scale of this photograph is also 1:7,000. In fact, it is entirely possible that the two values would come from the same photograph, as the scales are both nominal. There is every possibility that the differences are due to rounding error, unaccounted-for topography, error in the altitude reading, or a host of other error sources. Plate 15 shows the effect of topographic changes on an aerial photographic survey. If the aeroplane flies at a nominal altitude above sea level and the surface elevation changes, the scale of the photography will change significantly, too.

9.2.4 Further manual photogrammetry

Air photo surveys typically acquire photos with at least 60% overlap in the flying direction (Plate 16: upper) and at least 20% overlap (sidelap) between flying runs (Plate 16: lower). This ensures total stereo-coverage of an area so that photogrammetry can be undertaken between any points within the area. An entire area survey is usually termed a photo ‘block’.

Many other equations exist for undertaking a host of calculations from aerial photographs, including parallax equations to allow an interpreter to calculate the heights of objects. Hardware to speed up this kind of calculation also exists, such as a parallax bar, which can reduce the complexity of each calculation between objects on a pair of air photos and is usually included with mirror stereoscopes. Stereoscopes allow an interpreter to view airphotos in such a way that the interpreter’s brain is fooled into thinking a real surface is being seen in three dimensions. A stereopair of airphotos is viewed through a set of angled mirrors: when positioned correctly, the operator will perceive a three-dimensional view of the photographic subject. In this way, complex volumetric relationships can be assessed and evaluated. An example of a lab-spec mirror stereoscope is shown in Figure 9-8.

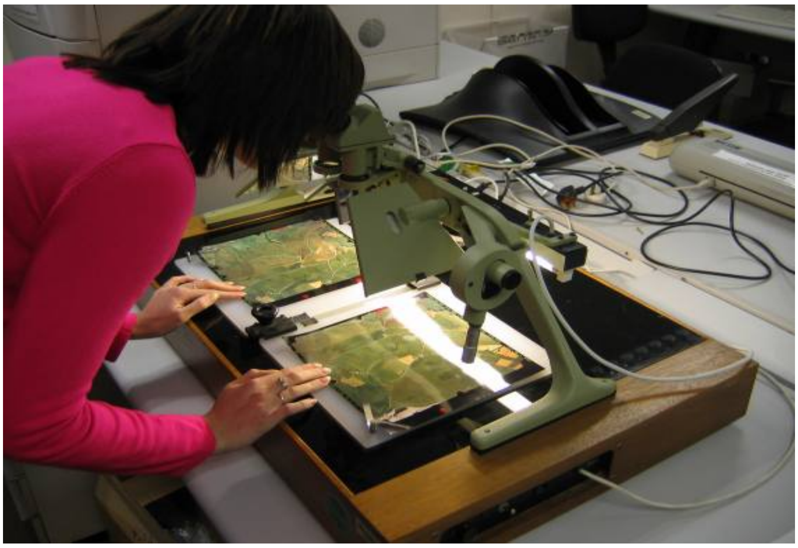


Figure 9-8 Air photo interpretation with a mirror stereoscope mounted on a light table (useful for viewing transparent air photos, diapositives, as well as traditional photographic prints).

Air photos can also be extremely useful in the field, especially if their 3-D viewpoint can be exploited. There are a number of ways this can be done, perhaps the most straightforward being the use of a pocket stereoscope (Figure 9-9). This equipment is small enough to carry into the field and will allow basic 3-D viewing of aerial photo stereo-pairs. Other approaches are use of virtual 3-D software such as ArcScene in ArcGIS and VGIS in Erdas Imagine, or the viewing of anaglyph imagery (two slightly different perspective images overlain in two different colours) in an appropriate software package on a laptop PC with red-blue glasses (Plate 17).

Figure 9-9 Pocket stereoscopes and aerial photo stereo-pair, one stereoscope set up for use, the other folded for transport. Photos are 9" per side for scale.

More complex photogrammetric machinery attempts to model the camera system used to acquire the photos, in an attempt to remove the distortions from the photography. The two main categories of these are analogue and analytical stereoplotters. *Analogue stereoplotters* are entirely mechanical systems for modelling the location, orientation and distortion of the camera system. Using a number of known survey points visible in the imagery, along with detailed information on the camera system, these devices allow an operator to construct an accurate, planimetrically correct map, directly from aerial photos.

Analytical stereo-plotters automate some of the above tasks, speeding up repetitive processes. Analytical stereoplotters are still sold today and are widely used for creating and updating maps. Both systems require an experienced operator and detailed photogrammetric knowledge to achieve good results - they are also very expensive to buy when new. Fortunately, software is now available that will allow photogrammetric processes to be undertaken without specialised equipment.

9.3 Geocorrection

Map co-ordinates are central to what makes a GIS different from a conventional information system – the spatial element to the data. A GIS allows us to spatially integrate data and interrogate disparate datasets based on the location of the features of interest. To undertake this with remotely sensed imagery, an image needs to be spatially located. Geocorrection is the process whereby imagery is assigned spatial map co-ordinates, transforming it into a valuable information source. Remotely sensed imagery and aerial photography are traditionally supplied without a digital co-ordinate system, so they need to be geo-corrected before they can be fully used in a GIS. This is changing, however, and some data, such as Landsat ETM, are now supplied as standard in a geo-corrected form.

The first stage in Geo-correcting an image is to select an appropriate map reference system for the area of interest. Correction of all data sources to a single map projection will allow accurate and easy integration, so choosing the correct system is critical to success. Refer to Chapter 2 for a more in-depth discussion of map projections and reference systems. The process of geo-correcting an image can be thought of as two-stage process, with transformation followed by resampling.

9.3.1 Transformation

To see how this process works, an understanding of file co-ordinates is necessary.

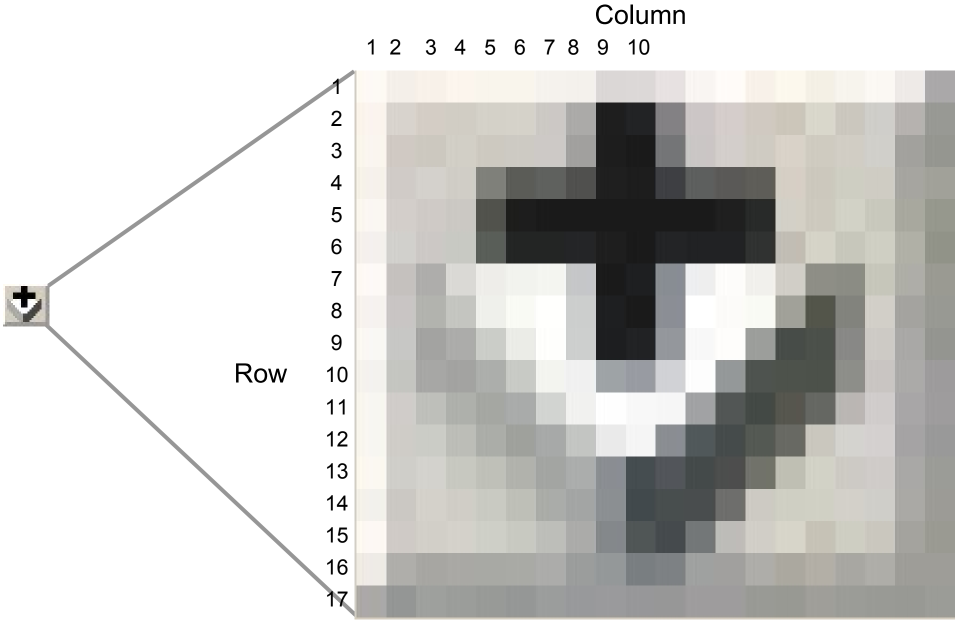


Figure 9-10 File co-ordinates of an image, usually given as column and row numbers. Note that pixel ‘1,1’ is at the top left, rather than bottom left.

File co-ordinates (as in Figure 9-10) reference each pixel in an image individually. If the image is a remotely sensed image of the Earth's surface, the pixel also represents an area on the ground for which map co-ordinates can be found. The process of transformation determines the mathematical relationship between a pixel's file co-ordinates and the map co-ordinates of the area that pixel represents. This is represented graphically in Figure 9-11.

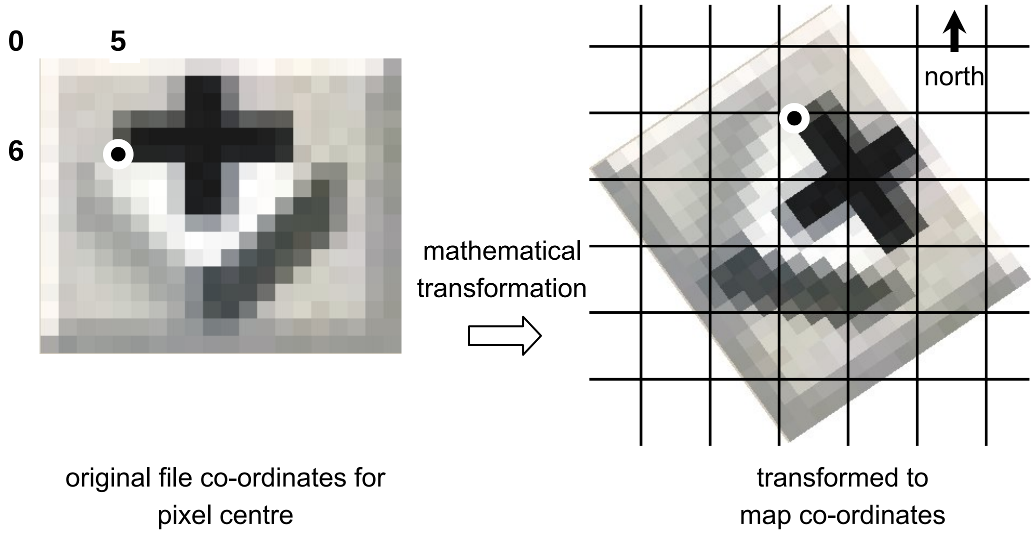


Figure 9-11 File co-ordinates transformed mathematically to map co-ordinates.

The choice of which mathematical transformation method to use is critical to successfully geo-correcting imagery. A wide range is available, but a comprehensive description is outside the scope of this chapter. For further information on the range available, please refer to the background reading suggested elsewhere in this manual and the relevant software's tutorials and manual. The two most commonly used transformations are affine and polynomial.

Affine transformations apply simple scale, rotation or translation corrections. They require no ground control, but require the user to know what the shift required is in terms of degrees rotation, direction and distance to translate, and difference in scale. Each package which implements affine transformations does so differently, so the user will need to consult the help files of the particular package used.

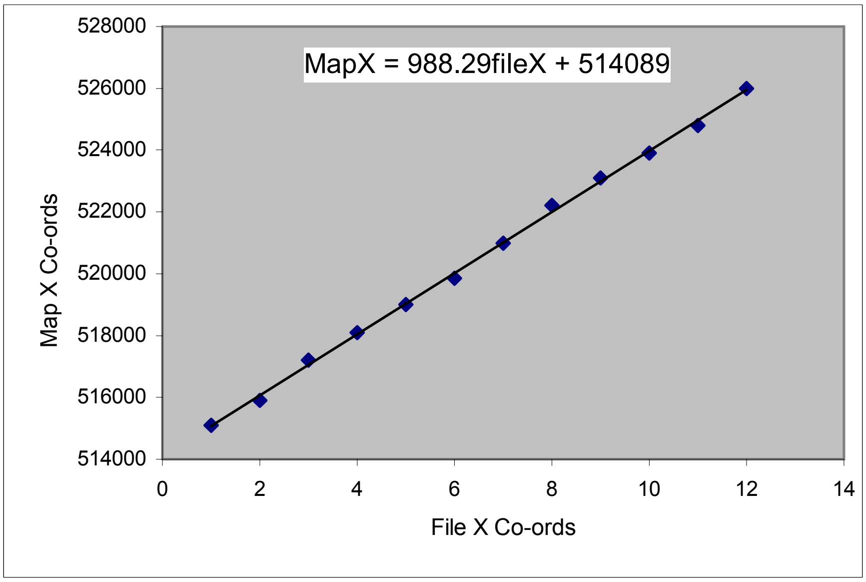


Figure 9-12 Map X co-ordinates (Easting) plotted against file X-axis co-ordinates with a polynomial best-fit line. Unknown co-ordinates can be estimated from the line position.

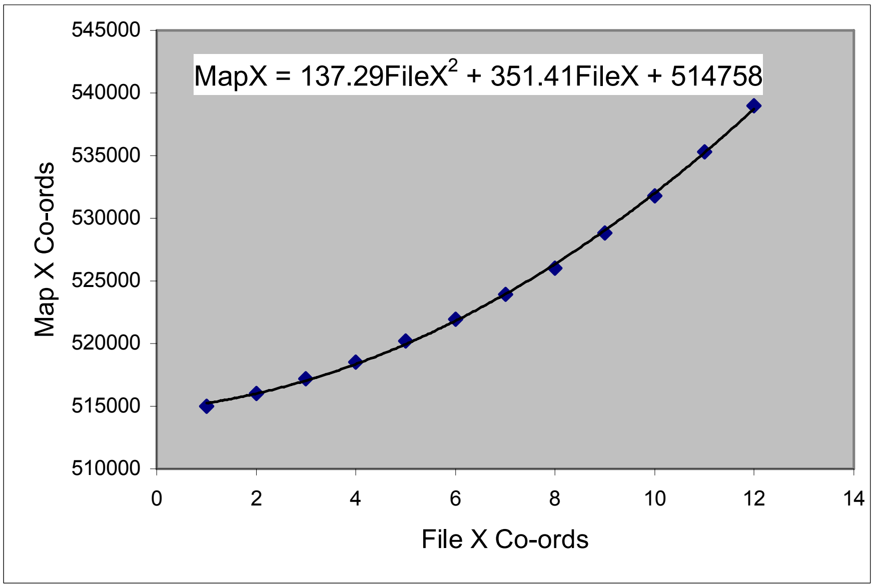


Figure 9-13 Map X co-ordinates (Easting) plotted against file X-axis co-ordinates with a second-order polynomial best-fit line.

Polynomial transformations are the most widely used method of mathematically transforming file co-ordinates to map co-ordinates. A polynomial transformation allows constants and variables to be applied to the original file co-ordinates, to obtain a logical desired result: the map co-ordinates. Conceptually, they can be thought of as a 'best fit' line through a series of points plotted on a graph. The defined constants and variables mathematically define this best-fit line, as illustrated in Figure 9-12 and Figure 9-13. Two equations are defined for each transformation, one for the X, Easting or Longitude co-ordinate and a second for the Y, Northing or Latitude co-ordinate. These equations are based on user-entered data and are used to estimate map co-ordinates for pixels not pre-defined by the user. Polynomial equations are defined by their 'order'. This is the number of variables allowable in the equation. The higher the order, the more complex the equation becomes (see Figure 9-14, Figure 9-15 and Figure 9-16). A higher order polynomial equation can remove more complex distortions from imagery than a lower order polynomial. Unfortunately, the higher the order, the more ground control is needed, and the greater the error as you move away from these points. Users are advised to use the lowest polynomial order that yields acceptable error levels to avoid these issues where possible. Which order to use depends on how the image needs to be distorted and in how many directions simultaneously.

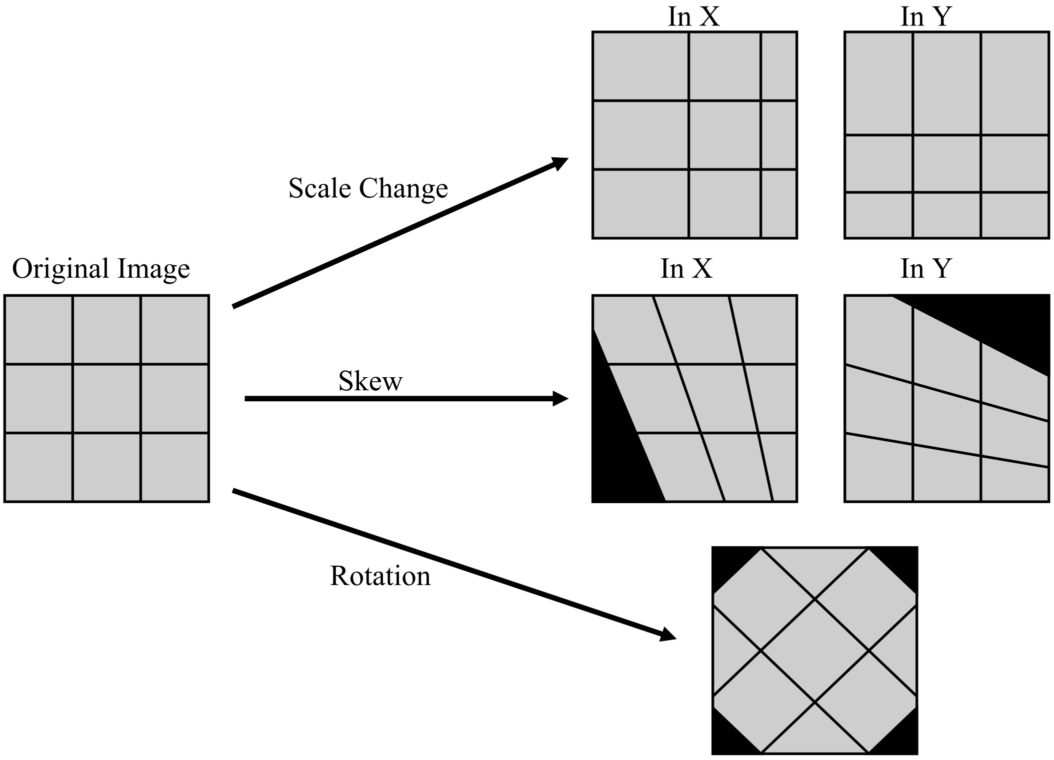


Figure 9-14 Distortions correctable with polynomial order 1 transformations.

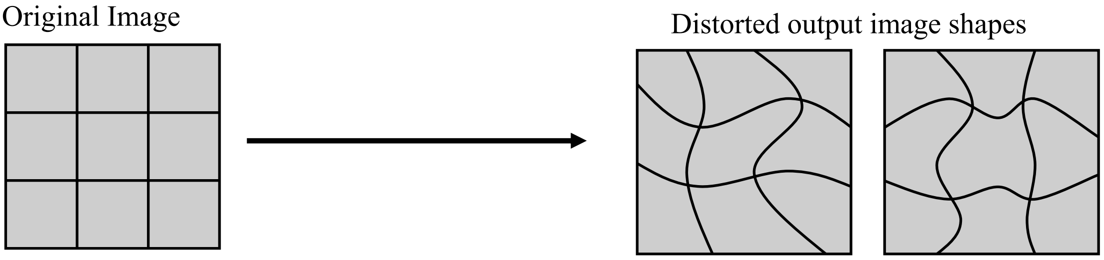


Figure 9-15 Distortions correctable with polynomial order 2 transformations.

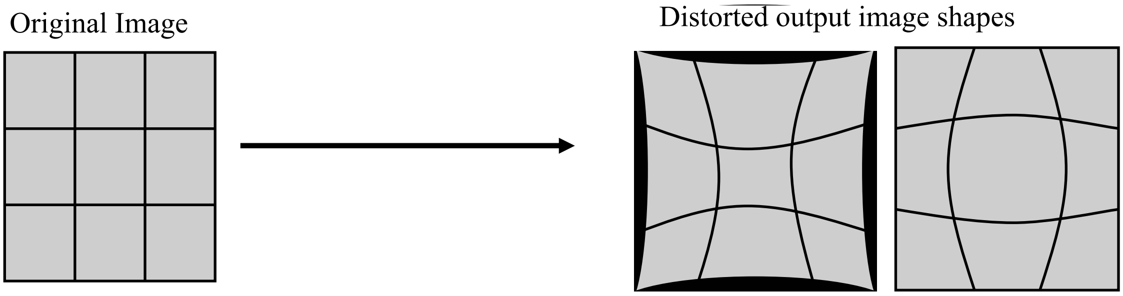


Figure 9-16 Distortions correctable with polynomial order 3 transformations.

9.3.2 Error measurement

The standard error measure in geocorrection is Root Mean Square Error (RMS error or RMSE). RMSE takes the difference between observed value and estimated value for each user-defined point, squares it, finds the mean squared value, and then finds the root of that value.

The equation for Root Mean Square Error is:

$$\text{RMS value} = \sqrt{}\ \Sigma(X_S - X_c)^2 + (Y_S - Y_c)^2$$

(where Xs and Ys are the user-derived source co-ordinates and Xc and Yc are the estimated ‘best-fit’ co-ordinates (see Figure 9-17))

RMSE is used in geocorrection because it gives a more sensitive and accurate measure of actual error in the result than simpler methods tend to. A modified calculation can also estimate individual Ground Control Point (GCP) error allowing the editing of points to improve the correction and reduce the error. RMSE is also widely used in scientific circles and, when correctly used, accurately describes the level of error we can expect to find in the corrected image.

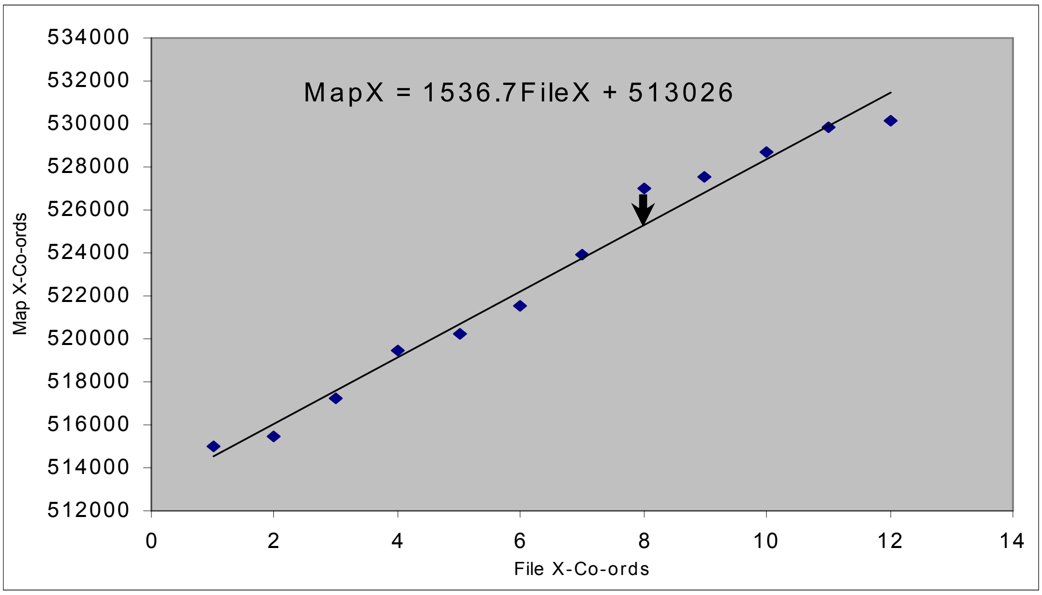


Figure 9-17 Arrow indicates the calculated error in the X-dimension. Total error is calculated by combining with Y-dimension error calculated in the same way.

Overall RMS error is calculated based on all the data entered by a user regarding the location of an image on the Earth’s surface and the transformation method selected. Deciding what is an acceptable level of error should be based on an understanding of the process alongside a knowledge of the project for which the data is destined. RMS error values are defined in terms of *input image pixel size.*

The amount of error is related to the nominal pixel size of the image. For example, when correcting Landsat TM, the general rule is that the total RMS should be <1, but as a pixel is 30 m x 30 m, the actual error is up to 30 m. When correcting air-photos, the nominal pixel size may only be 25 cm. Even with an RMS of 7, the total error is still sub 2 m. Deciding on an acceptable error level therefore depends very much on the imagery, the quality of the data used to correct it and its anticipated use.

However, RMS error calculations are not foolproof: they are reliant on a good distribution of GCPs within an image. This may well be difficult, if not impossible, to achieve in some areas of the world. Poor distribution of GCPs can lead to huge errors that only become obvious when integration with other datasets is undertaken (Figure 9-18). Selection of appropriate and well distributed GCPs and choosing the correct transformation are therefore critical to limiting error in the output image.

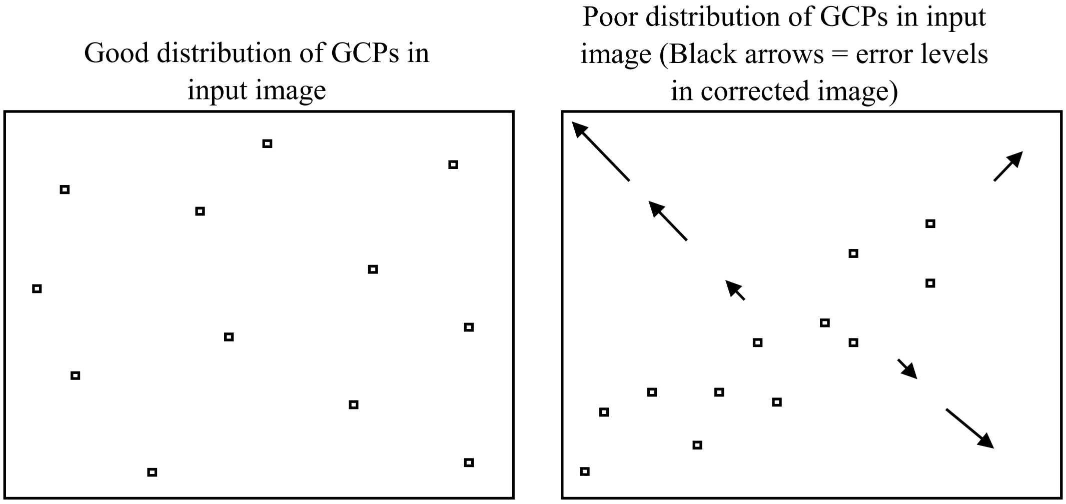


Figure 9-18 Effect on error in geocorrected image of GCP distribution. RMSE levels are the same for both images.

9.3.3 Resampling

The second stage of geocorrection is resampling, a process that creates a new image from the original, based on the information that you have entered. This stage is not essential for using the image, but computers deal very well with regular, orderly grids of square pixels, arranged to match a 'graph/origin' structure. Computers struggle to deal with data arranged in other ways, particularly with distorted grids and irregularly shaped pixels. Resampling allows an image that is distorted (relative to a map grid) to be stored in a regular, orderly grid of pixels, as illustrated in Figure 9-19.

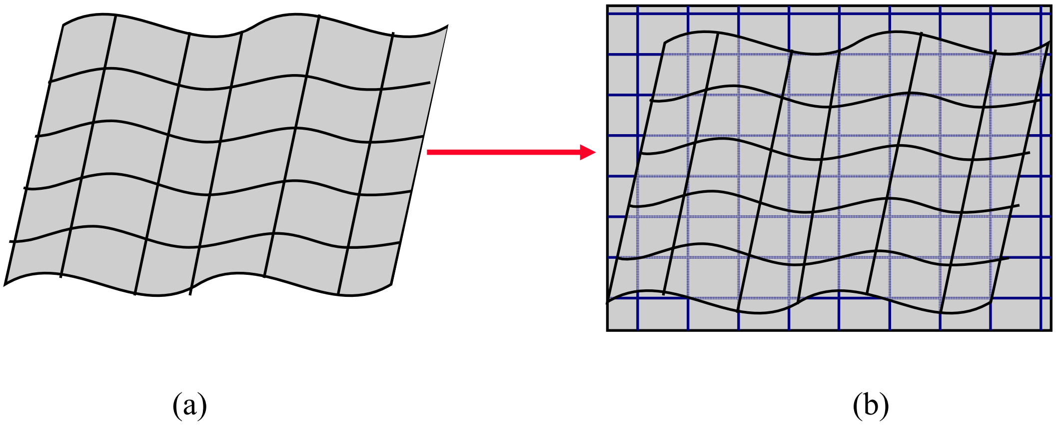


Figure 9-19 Creation of undistorted image grid via mathematical transformation and resampling of original image. (a) Distorted original image pixels. (b) New, regular, ordered image pixels.

The problem that resampling solves is that our transformed pixel centre co-ordinates might not fall exactly in the centre of a pixel in our new image. How do we therefore choose a value to represent our new pixel? Resampling provides the answer. There are three widely used resampling strategies available in most geocorrection-capable software: (i) Nearest-Neighbour, (ii) Bilinear Interpolation and (iii) Cubic Convolution. These three approaches for deciding what digital number to use for our new pixel center all have advantages and disadvantages. How these approaches work and issues with using them are discussed below.

The *nearest neighbour* approach uses the closest old pixel centre value as the new pixel centre value. The advantage with this is that the pixel values remain unchanged, which can be critical if further processing, such as image classification, is to be undertaken. A further advantage is that the decision rule used is simple to understand. The primary disadvantage is that a given old pixel centre might be closest to more than one new pixel, so the values are used twice, whilst another centre may not be used at all. This can result in ‘steps’ in images, irregular breaks in image features that are clearly artifacts of processing. The artifacts this process can produce can detract significantly from an image used for manual interpretation.

Bilinear interpolation calculates the new pixel value based on a weighted average of the four closest old pixel centers. The advantage with this technique is that processing artifacts, particularly ‘steps’, do not occur, giving an improved visual appearance. A second advantage is that the technique is computationally fast compared to cubic convolution. The main disadvantage is the original pixel values are lost, so further automatic image analysis is unsound.

The *cubic convolution* approach calculates a new pixel value, using the closest sixteen surrounding old pixel values. The main advantage with this approach is that it produces the most visually acceptable image, so is optimal for further manual processing. The main disadvantage, as with bilinear interpolation, is that the original pixel values are lost, making further automatic image processing unsound. The technique is also very complex to understand and computationally intensive.

Resampling is a fundamental process required to geocorrect remotely sensed imagery, however, the approach used can limit the utility of the corrected imagery. Even if imagery is supplied in a geocorrected form, knowledge of the processing applied to achieve this can be useful, as geo-corrected imagery will have to have been resampled at some stage. Care should be taken to ensure a user fully appreciates the pre-processing that an image has undergone, so that errors are not made.

9.3.4 Choosing Ground Control Points

Typically, to geocorrect imagery, a minimum number of points must be identified in the image, for which map co-ordinates are known and which are identifiable in the image (ground control points, GCPs). The number necessary will depend on many factors, but the main criteria are the transformation technique used, the type of imagery to be corrected and the accuracy required. Choosing optimal ground control points will aid in ensuring high accuracy and may also potentially reduce the number of points needed to correct an image. Points which are as precisely located as possible should be chosen. Road/path junctions should be selected where the angular relationship between the features is as close to 90° as possible. Figure 9-20 illustrates the issues with angular relationships when identifying points.

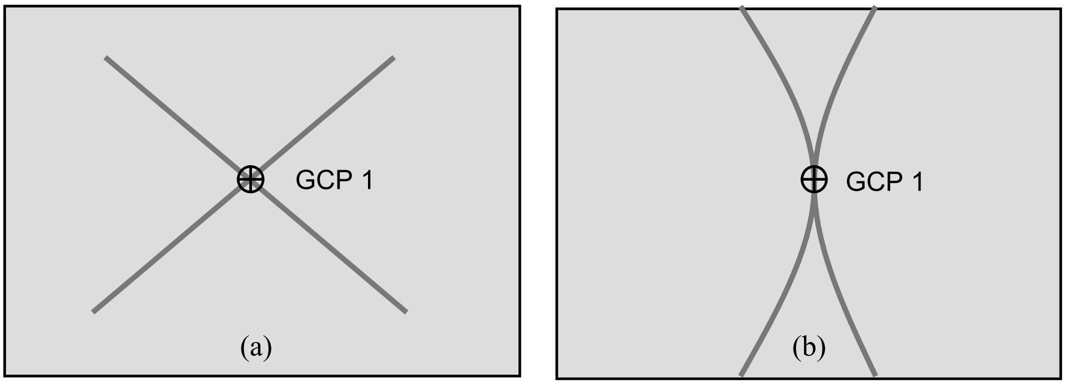


Figure 9-20 (a) Optimal and (b) sub-optimal location of a ground control point. (b) is less precise because the region in which the point could be correct is larger.

When using the polynomial transformation, there are theoretical limits on the number of points required to correct an image. With order one, at least three points are required. With order two, six points are needed and with order three, ten points are necessary. Many more should be used, however, to ensure reasonably accurate error reporting and redundancy in points. At least *four times* the minimum number should be defined before any faith should be put in the accuracy of the result. These should be well distributed throughout the image to ensure any error is minimised. When using ground control derived from GPS data, care must be taken to account for and assess the error inherent in the GPS derived co-ordinates. See the chapter on GPS in this manual for further information on GPS error.

9.3.5 Practicalities of geocorrecting images

The procedure for geocorrecting images is: first define which image you want to correct. This can be problematic, especially if the image is of an area of high relief, as geocorrection may give very poor results and orthocorrection becomes the only accurate option. In areas of low relief or where relatively large errors are acceptable, geocorrection will yield relatively reasonable results in most cases. The choice of which correction process to use is with the user and the choice will depend somewhat on their experience and the use to which the image will be put. Even the best quality satellite data may contain uncorrected distortions so a user must always be wary.

Second, decide what data to use for ground control: typically this will be a map, but could also consist of GPS co-ordinates and a description file, or other imagery, if it is already geocorrected. Digital map data can be used if available, although this may be difficult to find and purchase. Care should be taken that the data used for ground control is at an appropriate scale for geocorrecting your image. In many cases, small scale mapping will be all that is available. Features on this kind of map are usually highly generalized, so care is needed in selecting features that can be accurately located, and the error levels of using this kind of data should be kept in mind. Remember, 90 m on the ground is less than 1 mm on a 1:100,000 map, but represents three Landsat pixels, a significant error margin. Geocorrection is usually possible, regardless of the data quality, but the location error must be appreciated when using the output if serious mistakes are to be avoided.

The third step is to choose a transformation type. As already mentioned, numerous types exist, but the most appropriate will depend on the individual circumstances. Choosing a polynomial transformation is a good place to start unless an alternative, offered in the software, has been researched in some depth.

The fourth stage is the most time consuming, identifying and defining the ground control points. These need to be individually identified in the image, then their co-ordinates entered, from whatever source they have been derived. Control can usually be entered from another image, a digital map, from a text file (e.g., GPS co-ords) or from the keyboard, among others. Usually, entry from the keyboard is most prone to error. A single figure entered incorrectly can place your point thousands of kilometers out. Real care needs to be taken that every value entered is correct. Real time error reporting, shown in Plate 18 as a column for individual point RMS Error, is a great help in identifying the problem.

Point filtering is the final step in the process. The points entered should be reviewed to identify any serious errors. This could be due to operator error, but could also be due to map generalization. Experience in this area is essential, so the new user is advised to refer to a more experienced colleague, but in the absence of this advice, common sense and careful contemplation can suffice. An example of point errors can be drawn from a personal experience with Spanish 1:25,000 mapping when correcting a Landsat scene. A road over a stream marked on the study area map was identified as a likely GCP. Having placed the point on the image and entered the co-ordinates for the crossing, the error level was >3, which was unacceptable for the intended use. When the map was re-examined, it was found that the stream marked on the map did not actually cross the road where the bridge was marked. This led to a re-examination of other areas on the map. It turned out that the map had been mis-printed and consequently the colour layers were not precisely registered. All inter-relationships between roads and waterways were mis-located. When points using this kind of relationship were removed, the error fell to <0.7.

When maps are produced, more care is usually taken to place anthropogenic features accurately than natural ones. River junctions should therefore be treated with caution, whereas major road junctions can usually be used with confidence. The final, but possibly most important point about GCP filtering is to resist the temptation to automatically remove the points with the highest error. This can very easily lead to poor point distribution, which can produce a less accurate output than a high error level with a good distribution. This is especially true of isolated points towards the edge of the image, these can be essential to maintaining the overall positional accuracy of an image.

9.4 Digital photogrammetry

Digital photogrammetry uses the power of computers, namely rapidly repeating complex calculations, to reduce the difficulty of producing ortho-images and surface elevation models. By using software for this kind of processing, the requirement for expensive and complex machinery is substantially reduced. A little explanation of the terms ortho-image and surface elevation models is, perhaps, appropriate here. Ortho-images are images that are planimetrically correct and so can be used directly as maps. All distortions have been removed, so scale is constant, angular relationships are correct and areas can be measured directly. Surface elevation models are representations of (usually, although not always) the Earth’s surface. They can consist of line data (contours), regularly spaced spot heights

(Digital Elevation Model or DEM) or irregularly spaced linked spot heights (Triangulated Irregular Network or TIN).

Production of ortho-photos and surface models from aerial photography involves modelling the location, orientation and optics of the camera in relation to a map co-ordinate system. Knowing precisely how the camera is located in three dimensions and how light is transmitted through the optics onto the photographic film, allows the software to retrace the light that formed the photographic image. From multiple viewpoints and from a few ground control points, the Earth's surface relative to the camera (using parallax – see last section) can be determined. Using this information, the full path of light, which is reflected from the Earth's surface, passes through the camera system and is recorded by the photographic film, can be modelled. Retracing this path allows the software to correct the location of each point in the image, resulting in a planimetrically correct image (Figure 9-21 and Figure 9-22).

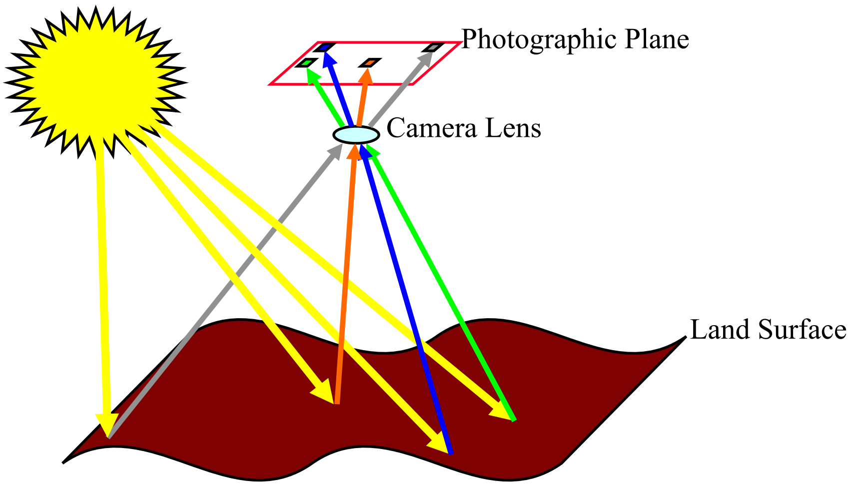


Figure 9-21 Conceptual diagram of an aerial photographic system. Land surface is sun illuminated, the reflected light passes through the lens, which focuses an image of the surface onto a 2-D photographic plane.

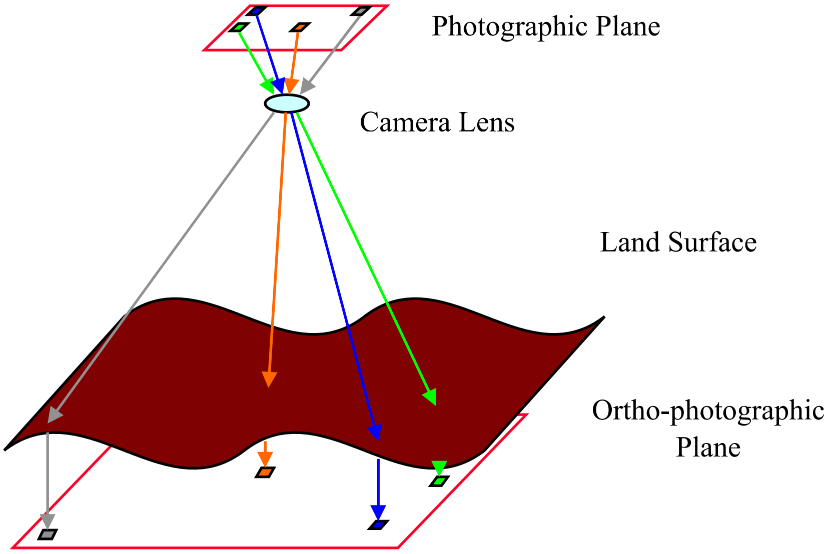


Figure 9-22 Ray-tracing approach to ortho-correction of aerial photographic distortion. The image is re-projected, via a surface model onto a planimetrically correct ortho-photo plane. All distortion is removed to produce an 'image-map', usually called an ortho-photo.

The result of this modelling procedure is extremely useful, giving all the information of a topographic map, but also including information usually excluded from traditional mapping, such as vegetation density, vegetation type and geomorphological context. The surface model created can also be extremely useful, giving information on a much finer scale than most maps, so that, for example, floodplain features are visible and can be analysed. The level of detail is only limited by the original photography and information not visible on standard maps is usually easily found (Figure 9-23).

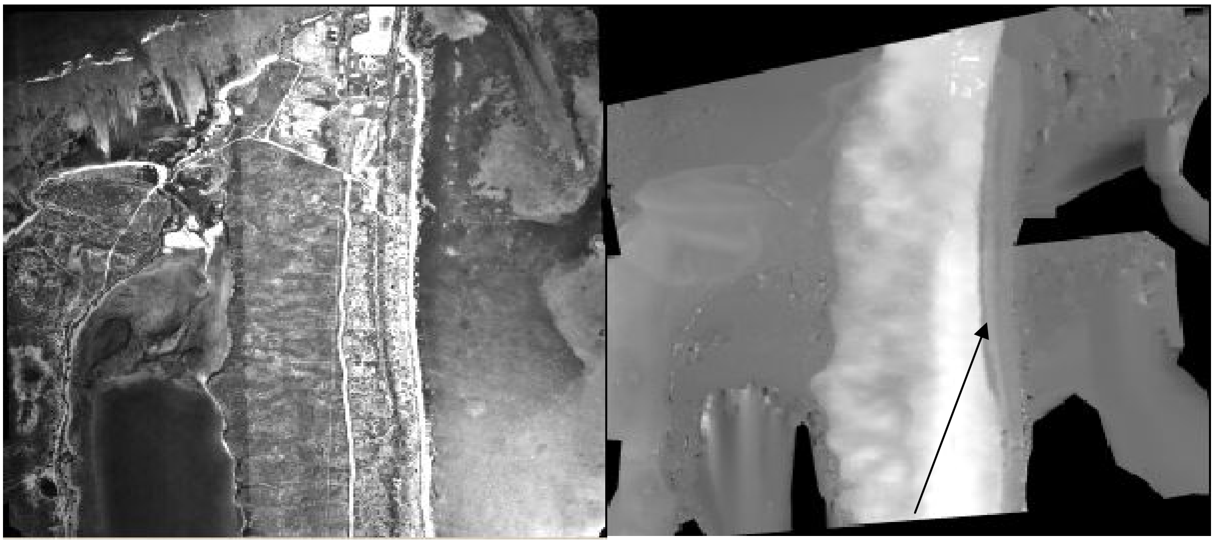


Figure 9-23 Orthophoto and DEM of the northern part of Grand Turk island, part of the Turks & Caicos island group. Note the beach forms on the eastern side of the island. The lighter band (arrowed) is a raised bench, probably caused by storm processes.

As all the output from digital photogrammetry is, by definition, digital, the image and the elevation model can be combined to produce pseudo-3D views and fly-throughs of the area of interest. Figure 9-24 shows a 3D view of Grand Turk, constructed from the images shown in Figure 9-23, illustrating the information present in the model and its usefulness in understanding the morphology of the Earth’s surface.

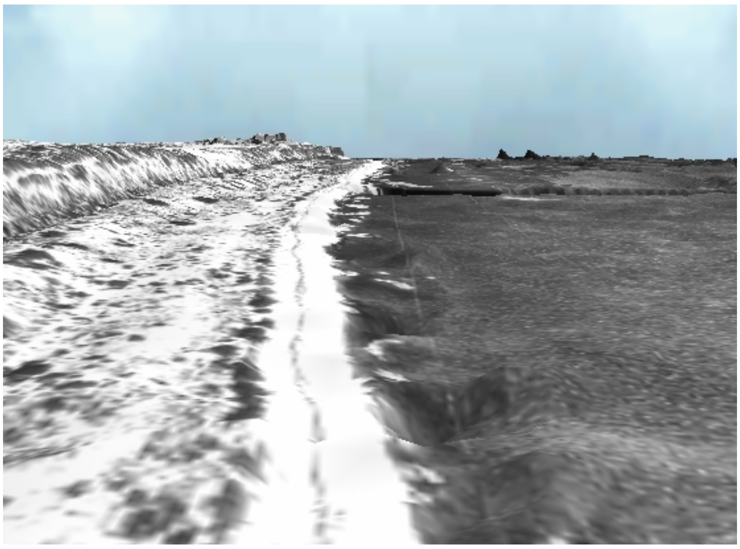


Figure 9-24 Pseudo-3D view of shoreline (arrowed in Figure 9-23) from the south.

This kind of information, which can be extracted from the photography without the need for field data of any kind (some ground information is necessary – see below) can be invaluable in locating features or areas of particular interest, finding accessible routes into remote areas, locating base camps in appropriate areas and for adding value and understanding to the fieldwork proposed. For conceptualising the terrain of an area, this approach is unsurpassed.

9.4.1 Undertaking Digital Photogrammetry

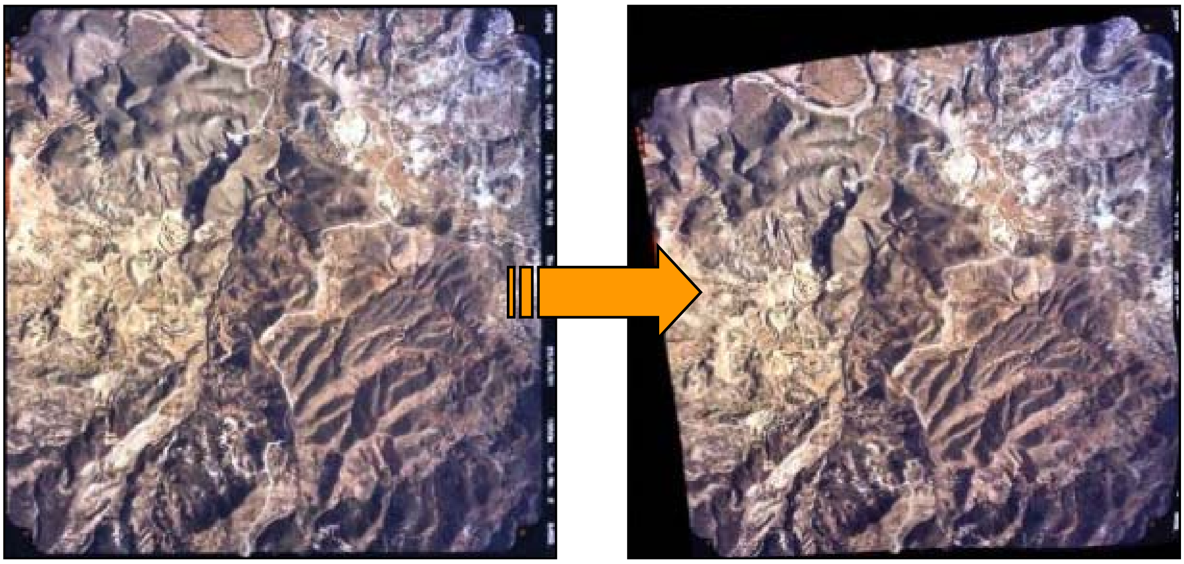


Figure 9-25 The image on the left is an uncorrected scanned air photo, the image on the right is an orthophoto produced from digital photogrammetry of the same area. Note the distortions in the shape of the photo caused by removal of camera distortions, topographic distortion and perspective distortions.

Digital photogrammetry software is quite specialised and complex to use. Detailed descriptions of the processing stages are outside the scope of this manual and you will need expert advice and assistance to undertake the process yourself. If you are aware of the necessary stages, however, it is possible for a committed novice to achieve a successful result using information from many sources, particularly software manuals and the internet.

The first step in the process is usually to get the photography into digital form and a number of issues need to be addressed at this point in the process.

Most aerial photography is analogue, i.e. the photo is recorded on film using chemical reactions. Analogue photographs must therefore be scanned before they can be used digitally. To achieve the best results, particularly in terms of subtle elevation changes, a photogrammetric-quality scanner should be used and film-based diapositives should be scanned rather than the traditional paper prints. If a high accuracy (usually <50 cm vertical error from 1:10,000 photos) is not essential, paper prints scanned in a reasonable quality desktop scanner are more than adequate. The photos should be scanned at an appropriate resolution for their intended use – generally 300-600 dpi is appropriate. The more detailed the result required, the higher the resolution should be, but care needs to be taken that the optical resolution of the scanner is not exceeded as this adds significantly to the data file size without increasing the actual detail visible in the image.

The next stage would be to import the scanned photos into the software package. This is typically done as a photo ‘block’. Usually this block would consist of all the photos that are necessary to provide stereo-coverage of the area of interest. The photo block may consist of multiple photos per run and multiple runs to cover a wider area. All should be processed simultaneously to achieve the best results. In practise, this may not be feasible due to storage constraints, particularly if the images are scanned at a high resolution and in colour (a single colour photo, scanned at 600 dpi, requires around 80 MB of storage. Processing it will take around three times this space. To process 20 photos would therefore need at least 5 GB of storage).

At this point you would need to define the internal parameters of the camera used to acquire the photos. This information can be found on the Camera Calibration Certificate, which should be available from the same source as the photos. Most software has default camera specifications that can be used, but this will mean some camera-induced distortions may not be properly removed. At the very least, the information about the camera focal length should be correct. This information is generally included on the photography itself on the titling strip.

Once the camera has been defined, the next stage is to define the internal orientation. This means matching the camera model to the digital images and simply involves defining where on the digital photos the fiducial marks appear. These marks are precisely defined in the camera calibration certificate, so can be used to determine where any part of the image was located on the photographic plane when the photo was taken.

The next stage is to define the exterior orientation. In this process, the position of the camera in relation to the chosen map co-ordinate system is determined through the use of ground control points. Further points can be defined, called tie points, which are not linked to the map co-ordinates, but help to ‘tie’ a stereo-pair of photos together so that ground control on one can be used to correct both. Once a series of ground control points (GCPs) and tie points has been defined for a whole photo block, triangulation can be undertaken. As a guide, a minimum of one GCP and four tie points per photo is required, though clearly more points would be useful in reducing error. Triangulation is an attempt to reconcile the camera model, the ground control points and the point locations to an acceptable model. If errors in data entry have occurred, triangulation is likely to fail. Error reporting at this stage is a good indicator of the error present in the final datasets. Time spent reducing error levels at this point is usually well spent, in terms of improved accuracy in the DEM and ortho-photo output.

The penultimate stage is DEM extraction. The DEM for each stereo-pair in turn can be derived for the area of stereo coverage through reference to the triangulation information and pattern matching between the images. The pattern-matching algorithm attempts to match the two photos together, based on the nearest GCP or tie point to a particular pixel. When the algorithm finds a match, a parallax measurement is undertaken which calculates the topographic height of that point, based on the GCP data. This height is recorded and the algorithm begins to try to match the next point on the image. This continues across the whole area of stereo-coverage, building up a digital elevation model. Some software is capable of directly deriving a TIN via a similar method, but the process is fundamentally the same.

The last stage is *ortho-correction* of the photography. Having produced an elevation model, all the necessary components for the ray-tracing approach are defined, so the process re-projects the photo image onto an ortho-photo plane, removing all distortions and producing a planimetrically correct image. Ortho-correction provides accurate image-maps that can be used on the expedition in the same way as traditional topographic maps.

9.5 Practicalities of accurate mapping using airphotos

Aerial photos in remote regions can always be useful, even if they are only used to orient yourself relative to ground features. Sources of aerial photography are discussed in the Remote Sensing chapter and contact details for major suppliers are given in the Appendix. For many projects, air photos will not be available, or ground control will be insufficient to fully gain the benefits they can provide. For other projects, a reconnaissance trip would be needed but the financial limitations mean this is not an option. If air photos are available, they can provide invaluable information regardless of the level at which they are processed, so the effort to acquire them is rarely wasted.

Airphotos show land cover features that are missing from traditional topographic maps, and are usually much more detailed than available mapping. As long as the distortions inherent in photography are understood, photos can be used in a similar way to maps, helping find trails, locate potential campsites, navigate around areas and regions, etc, especially if used in conjunction with a stereoscope, whether pocket or lab-based. For more complicated tasks, undertaking ortho-correction and DEM extraction are likely to be necessary, especially if combining the data with other GIS data layers. This requires ground control, however, and getting this information can be difficult or even impossible. To fully correct a series of aerial photos, a significant number of features must be identified on the photos, each feature being assigned accurate co-ordinates and altitudes. On average, three ground control points per photo are required for an acceptable result, although these do not need to be evenly distributed over a sequence of photos. Some of the photos can be included with no ground control at all, as long as the other photos in the run have sufficient points. Collecting or finding accurate location information can be difficult: the best approach is to make a reconnaissance trip. Copies of the photos can be taken on the recce trip, features visible on the photos identified on the ground and GPS readings of the location taken to provide ground control. These ground control points can then be used to accurately ortho-correct the images.

This optimum of at least two trips to a study area may not, of course, be possible with all expeditions. With this scenario, obtaining ground control must be achieved through other means. GPS co-ordinates from other researchers who have visited the area could be used if there is sufficient information about where a reading was taken, although care must be taken to exclude ambiguous points. However, such data are always suspect, so caution must be used when using data not collected personally. Mapping, if it exists, can also be used, but much mapping around the world is very dated where it exists at all. Small scale maps (1:100,000, 1:500,000 etc) are insufficiently detailed to provide accurate control for most aerial photography, so 1:50,000 or larger scale maps are needed. Features derived from this type of mapping also need to be interpreted and their potential accuracy weighed. A good example of this is a recent project looking at a rainforest area in Guyana. The only ground control available was 1:50,000 maps, originally surveyed about fifteen years after the photography was taken. Very few features on these maps related to features on the photos, though some path junctions and river features were eventually used. One point used for ground control was a river confluence with a very distinctive sediment bar: this eventually turned out to be located around 300 m from its location on the map, probably because of channel migration or mapping error. As with all aspects of an expedition, advance planning can be critical to success, so plan your data production and use of aerial photography carefully.

9.6 Summary

The key with using all photography or imagery, whether aerial or satellite derived, is to only use the data in a sensible and meaningful way, understand the error levels and their potential impact, and plan a project in such a way that possible problems are avoided or mitigated where possible. *Geocorrection* can be a complicated process, full of pitfalls for the unknowledgeable and unwary. However, when undertaken by a knowledgeable user with care and attention it transforms a ‘pretty picture’ into really useful GIS information. Geocorrection is an essential stage in using remotely sensed data, and an understanding of the process helps the user to appreciate the limitations and advantages of the images they are using, even if they did not undertake the geocorrection themselves. *Photogrammetry* is a useful tool that everyone involved in understanding spatially distributed processes, objects or features should be aware of. For the explorer, understanding the terrain of the area to be explored is clearly critical, so photogrammetry can help them gain that understanding from photography prior to any visit to the region. For most projects, it is likely that having an understanding of photographic distortion would be sufficient to prevent any disastrous mistakes, but for more detailed survey work, more in-depth understanding and knowledge may well be necessary.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section C: Techniques

Chapter 10: Traditional Surveying

10 Traditional Surveying

10.1 Introduction

Traditional surveying is at least 3000 years old, having been used by the ancient Egyptians to mark out fields after the annual Nile floods and refined by the ancient Greek and Romans in the Classical era. Survey techniques and instruments were then re-invented in the Italian Renaissance with reference to the Arabic scholars of the 1st Millennium. Survey methods were considerably enhanced by inventions in optics, metallurgy and mechanical engineering during the 17th and 18th Centuries, which allowed the development of instruments such as theodolites, levels, compasses and sextants. This generation of optical/mechanical instruments were used to measure the 19th Century primary triangulations on which the mapping of many nations is based. In the late 20th Century digital electronic techniques such as laser distance measuring, electronic logging, GPS and handheld computers revolutionised traditional surveying, although it introduced power demands that had not existed previously.

This chapter covers traditional surveying of two kinds:

- Surveying with optical/mechanical devices requiring some experience and training in data capture and logging;
- Surveying with digital electronic techniques requiring knowledge of devices, memory and data integration with computers.

No attempt is made to provide advice on how to operate particular surveying instruments and devices: this must be obtained from the instrument owner/device manuals. This chapter simply explains the way such devices may be used to survey features in field situations and the generic principles of their use. For more detailed coverage of traditional surveying the reader is referred to a field techniques booklet produced by Keith Miller (1993), published by the Expedition Advisory Centre. Details of GPS usage for surveying are given in Chapter 11 of this handbook and an example of a ‘hybrid’ surveying project, mapping a semi-desert site in Spain using optical/mechanical and digital equipment, is given in the case studies chapter at the end of this book.

10.2 Surveying context

Traditional surveying techniques generally require a line-of-sight between the instrument and a roving surveyor, or they require ground access across areas to be surveyed, or they require you to be able to see over the whole area to be surveyed. The first step in any traditional survey is to check on access permissions and the physical practicality of visiting the locations from which the surveying is to take place. This may involve testing the instrument at a possible location, for example, to see if the range of an electronic distance finder is sufficient to cover the area of interest, or if it is possible to set up the theodolite precisely level. Locations from which surveys can take place are often places where a surveyor may spend a considerable time, so ensure that you will not cause erosion, disturb birds, affect the living of anyone usually working there or put yourself at risk. In cold climates you can get hypothermic standing around making nothing but small hand movements at a survey location, while the roving surveyor can be tired out by traversing rough terrain. In hot climates you can be shaded at a survey location and protected from the sun, while the roving surveyor can get sunstroke. Hence sensible rotation of jobs and appropriate equipment and protection is required.

10.3 Surveying instruments/devices

10.3.1 Optical/mechanical instruments

10.3.1.1 Theodolite

An optical/mechanical theodolite is a device for measuring angles horizontally and vertically from a survey 'station' (a known location) to a survey target at a location whose position is unknown. These instruments *do not* measure distances directly. Basic principles of trigonometry are used to work out the internal angles of triangles that are measured, and then co-ordinate geometry techniques are used to establish positions within a national or regional reference framework for the unknown points. Most pre-1985 optical/mechanical instruments have no internal storage and no power needs, but they are still sensitive to knocks and extreme changes of temperature. They are set up on tripods with adjustable legs. Surveying cannot be carried out without tripods, so avoid damaging them or using them to hang out washing!

In the case of a horizontal survey the theodolite will first be set up at a ‘station’ with known co-ordinates: this will often be a physical benchmark marked with a metal spike or nail. The station needs to be set precisely level using two bubbles arranged at right angles to each other. This is often quite time consuming, as when the theodolite is set level it is no longer directly over the benchmark; and when it is directly over the benchmark it is no longer level! A controlled iteration of these processes needs to be carried out: if one or other of these settings remains wrong, then picking up the tripod and starting again is usually better. At this point the height of the theodolite over the benchmark needs to be read precisely with a measuring tape. The horizontal angles are usually measured relative to a baseline connecting the benchmark station with another station at a visible point (a 'backsight' point) whose co-ordinates are known (usually set up for the purpose). Since the entire survey depends on the fidelity of the baseline, backsight points are usually survey targets set up on a stable tripod. The angle between the benchmark station and the backsight point will have an angle of zero degrees; all other measured angles will have values that should increase clockwise from the baseline. In the case of the vertical survey the reference datum is the horizontal, which is recorded as zero. Observed points above the horizontal range from 0 to +90 (vertically up); below the horizontal, from 0 to -90 (vertically down).

Optical/mechanical theodolites can measure distances approximately using the techniques of tacheometry. The viewfinder of the instrument has three horizontal crosshairs (or ‘stadia’): the central one is used for pinpointing the centre of a survey target. The upper and lower stadia can be read off on a graduated survey pole: the distance between the two stadia when multiplied by a scale factor that varies by instrument (often 100 or 200) is equal to the distance between the instrument and the survey pole. Clearly, the accuracy with which you can read a graduated survey pole defines the accuracy of the distance derived.

Once set up and a baseline observed, a roving surveyor with a survey pole or target can be tracked by the theodolite operator who will record horizontal and vertical angles to the appropriate points as well as upper and lower stadia values for tacheometry. For precise work and as a basic check on accuracy, points can be read normally and then the instrument can be rotated 180° and the telescope 'turned over' so that the point can be read again (‘face left’ and ‘face right’ measurements). Stationary is available to structure the booking down of this kind of data. Examples are given in the EAC Simple Surveying booklet (Miller 1993).

10.3.1.2 Level

To establish heights relative to a benchmark a device with a sighting telescope called a level is used (Figure 10-1), which is set up on a tripod and can be set to be precisely horizontal using a bubble. Once level, the cross hairs seen through the telescope will sight points with the same elevation, regardless of position, within the range of the instrument (typically a few hundred metres). We can now measure deviations above and below this precisely level line using a graduated measuring staff: the level operator records the values read off on a graduated survey pole through the telescope. Values read off the pole that are greater than the height of the level instrument above the benchmark indicate that the ground on which the pole stands is below the benchmark height and vice versa.

Figure 10-1 Sighting through a level mounted on a tripod.

10.3.1.3 Plane table

Plane table surveys are a form of scale drawing. They establish the location of points on the ground relative to a baseline by viewing them from different locations and finding the intersection of lines of sight. The sightings are made from a small level table (hence ‘plane’ table) from two fixed points called 'stations' at either end of a measured baseline. Use a surveyor’s tape to measure the length of the baseline and to keep it as straight as possible. The sights, or ‘rays’, are made using a ruler with fold up sighting hairs called an ‘alidade’, and the rays are drawn on paper fixed to the plane table. You can also sight along the edge of a standard ruler, but the results will not be as accurate. Under extreme circumstances you can ‘sight’ and record a ray using a laptop screen as the ‘plane’ and draw the lines in a drawing package.

10.3.1.4 Barometer

A barometer measures height above sea level using the (known) decline in atmospheric pressure with altitude. These need to be calibrated at known height, especially after being taken on a plane. Small lightweight devices are available that will give reasonable assessments of height on any given day, but absolute height values may vary between days if atmospheric pressure systems change with weather systems. They are usually rather volatile and cannot be used to measure small variations in height (less than ~20 m). Some GPS units can now be purchased, which have a barometric altimeter built into the unit. This can provide a good second source of altitude data to confirm the accuracy of the GPS location.

10.3.2 Digital electronic devices

10.3.2.1 Total station

A total station is a theodolite that incorporates distance measuring using a laser or infrared beam, along with internal/external electronic data logging (Figure 10-2). These devices save the surveyor from collecting multiple measurements for each point location whose position needs to be known, as triangles can be constructed trigonometrically once you know one internal angle (between baseline and unknown point) and two edge lengths (the baseline and measured distance to unknown point). Total stations have LCD displays where you can set up the kind of survey task you want to carry out and it will prompt you for set up data (co-ordinates of the benchmark, height of the instrument over the benchmark etc.). The total station will then log the observations that you make, which can save on recording errors (unless you lose the data!). Rigorous checks need to be carried out on how the data is being stored to ensure its integrity and it should be downloaded and backed up as soon as the survey is complete. To prevent data loss, it is advisable to manually record data whilst surveying. If using a prism based system, the station operator will typically have to wait whilst the roving team member moves to the next location, giving the opportunity to record data.

Newer total stations sometimes do not require a prism and second operator. The laser or infra-red beam is of sufficient intensity and the receiver sufficiently sensitive, that reflections from typical Earth surface materials (rocks, tree trunks, etc) are sufficient to provide a measurement. These units are especially useful when the area of interest is inaccessible due to hazards such as overhanging unstable rocks or slopes, deep or fast moving water or other limiting factors, such as disturbance to the feature to be surveyed. The range of the beam is somewhat limited (typically around 100-200 m) compared to traditional prism based stations (up to 4 km), but their advantages in close-range surveys is significant, particularly in reducing necessary manpower and increased speed of data collection (not having to wait for the prism holder to move to the next point). The biggest disadvantage with these units is their cost. As with all newer technology, they tend to cost more to hire or buy than traditional total stations.

Figure 10-2 Total station mounted and ready to use on tripod. Note the use of a prism with this system. LCD display shows Easting and Northing of prism position.

Integration of total stations and GPS, either differential or stand-alone, can enable field data to be integrated with other map information in real time. Surveying the baseline with a GPS system, or deriving the accurate location of the line in map co-ordinates in some other way, allows this information to be used in surveying subsequent points. The baseline information, when input into a total station with the correct functionality, will enable the total station software to convert all subsequent measurements to map co-ordinates in real time without the need for post-processing. This can aid in collecting and processing data, and allow the effectiveness of data collection to be gauged on site. This functionality is common in total stations and can provide so many advantages that potential users are encouraged to ensure this option is available to them. Total stations are relatively bulky items and are also easily damaged, particularly the internal sensors and optics. Careful packing and transport are essential, and vehicle transport is recommended wherever possible.

10.3.2.2 Ground-based laser scanners

A number of companies now offer tripod mounted automatic laser scanning survey instruments. These effectively scan a horizontal and vertical field from a static location. The technology is similar to a total station, but the system scans a measurement beam at regular intervals across a field of view to build up a ‘cloud’ of points that can be used to produce a 3D surface of the imaged area. These systems are complex and delicate pieces of equipment, and although necessarily relatively rugged, may be too bulky and delicate to transport to a very remote location. Power requirements for these systems are likely to be significant too in remote areas. As this is relatively new technology, the costs of this kind of kit are also high, the retail price of a Cyra scanner being over £50,000. The systems can be hired, but insurance may be impossible to arrange. Some University departments do own these types of scanner, so an enquiry may be worthwhile.

10.3.2.3 Handheld computers

Hand held computers, personal digital assistants (PDAs) or smart-phones can be used with total stations and other equipment as external data loggers. Which units are compatible with which system is very variable and the method of data logging should be decided upon when considering which system to use. The use of data loggers raises more issues about power supply, however, as well as data security, so backups should always be in place. Some total stations and other surveying systems have limited ports for linking, a good example being the Leica TCS1600 series. These stations have a single power/data port, meaning that *either* an external battery can be used *or* a data logger, but not both simultaneously. As the internal battery is more prone to damage/discharge, an external battery and data recording in a field notebook would be recommended.

10.4 Survey techniques

10.4.1 Tacheometry

Full scale mapping can be carried out using survey instruments to position features and connect them into a network of measured triangles using the procedures of tacheometry. This is beyond the scope of this chapter but is covered in full in the EAC Simple Surveying guide (Miller 1993).

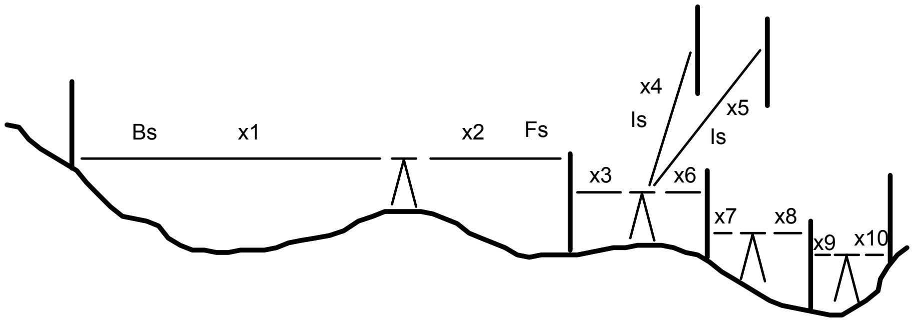
10.4.2 Traversing with a level

Firstly, unscrew the leg clamps on the tripod and extend the legs to their full length, reclamping them and setting the tripod upright. Now bed the legs in the ground firmly such that the apex of the tripod is roughly level. Taking the level instrument itself, place it on top of the tripod and using the screw rod under the apex screw it into the tripod, but not fully tight until after finalising the levelling process. To make the level sight horizontally the bubble in the oil filled capsule must be made to rest in the graduated circle. This is achieved by moving the level around slowly on top of the tripod until the bubble brought to rest in the circle. Once level look through the telescope and bring some distant object into focus. Now turn the eyepiece itself until the cross hairs are sharply in focus at the same time as the distant object. The crossing point of the two long lines at the centre is the horizontal level defined by the instrument. Practice focussing on the levelling staff and learning to read the graduations which are marked in 5 cm high 'E' shapes: first read the large number (e.g. 1.9) *below* the cross hairs, then estimate the two-figure number in mm *up* the 'E' shape (e.g. 33) and finally add together to get a reading (e.g. 1.933 m). Remember that the staff must be held absolutely vertically to reduce errors.

Since the level can only sight horizontally, we use the graduated staff to measure heights relative to this horizontal line. Hence if you read 1 m on the staff through the level in location A and 3 m on the staff in location B then location A is 2 m higher than location B. If location A is a benchmark then we know the height of location B. Note that the height of the level above the ground does not matter since you are establishing relative heights.

The normal practice for levelling profiles (Figure 10-3) is to make sights both backwards onto a benchmark (a ‘backsight’) and forwards onto new target locations ('foresights'). Although you can measure many relative heights with the level located in one spot, to make a traverse you need to move the level position from time to time. In this case the staff must remain *exactly* in its last position (i.e. a foresight location) whilst the level is moved and re-levelled. Now you can measure a backsight onto the staff still in its last position, and work out the difference in height between the horizontal lines measured by the level. It follows from this that there is always only ONE backsight for each level location (onto a benchmark or previous staff location), whereas there may be many foresights, although one of them (the ‘change point’) is used to link to the next set of observations.

The levelling readings should be recorded as shown in the table in Figure 10-3, indicating which measurements are backsights and which are foresights. Start with a backsight on a separate line and place the first foresight on the next line in the appropriate column. Further foresights are known as ‘intermediate sights’ and can be recorded one per line in that column, along with a full commentary on the location being heighted. Finally you will need to work out the 'rise' or ‘fall’ from one point to another and established the 'reduced level' of the location by relating these to the benchmark height. Once you have carried out the calculations you can plot the results up on graph paper to showing your long profile. An example of a levelling survey table and profile over a soil erosion plot in Zambia is shown in Figure 10-3.



Bs	Is	Fs	Rise (F<B)	Fall (F>B)	Level	Comments
x1						Benchmark height
x3		x2				
	x4					
	x5					
x7		x6				
x9		x8				
		x10				

Figure 10-3 Profile cross-section and data recording table for levelling, showing 'backsights' and 'foresights' and method of data recording in the field.

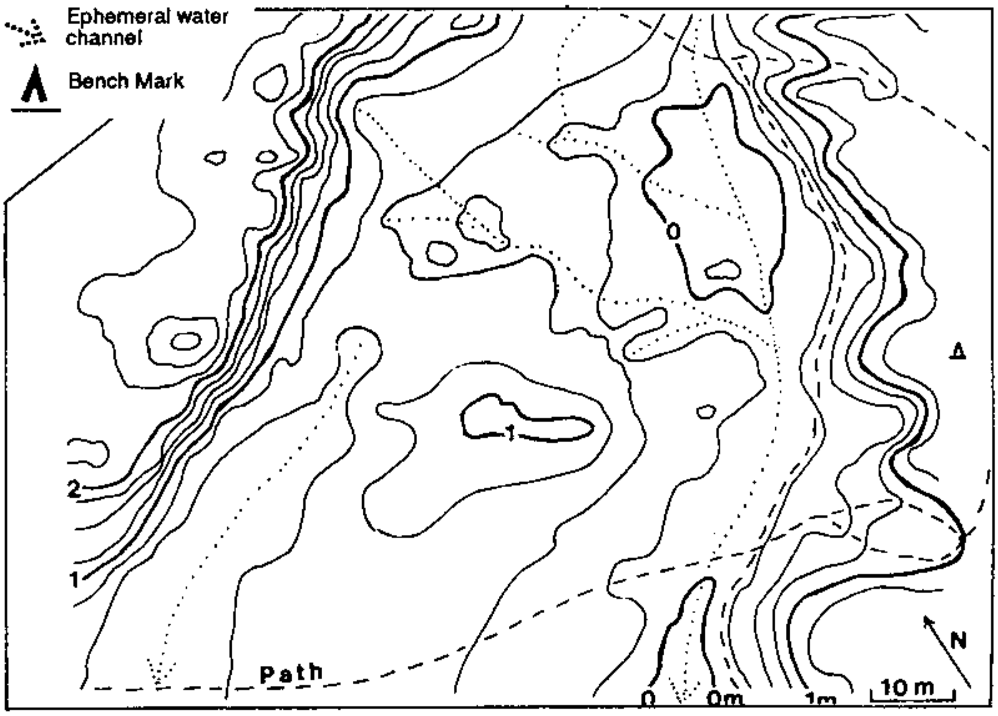


Figure 10-4 Detailed contour map of a soil erosion plot, produced using a surveyor's level and tachometry, from a geomorphology and land use map of river terraces in the Luangwa valley, Zambia (Teeuw, 1990).

10.4.3 Plane tabling

First you need to establish a base line on the ground of an appropriate length (decide what is appropriate) using a measuring tape. This should be at right angles to the feature to be mapped, e.g. a salt marsh creek (see Figure 10-5). Place a peg in the ground at each end of the base line: these are the 'stations'. Now move your plane table to one of the stations. Screw the tripod into the table and adjust the legs so that the table stands level and stable. Second, fix the paper squarely to the table using the masking tape; then get out the alidade and unfold the sights.

Next you need to draw a line along the long edge of the paper as in Figure 10-5, making the length of the line on the paper a scaled representation of the base line ground length, and marking the stations (use the maximum length of the paper). Now adjust the location of the plane table so that the station marked on the paper is *exactly* over the peg in the ground marking the location of the station. Place the alidade along the base line marked on the paper and *rotate the table until you sight the position of the other station marked by a peg.* Since you will not be able to see the peg someone will need to stand at the location with an arm raised or holding a pole. Now use the compass provided to mark the direction of magnetic north once your table is 'oriented'. Finally, clamp everything, since nothing must now move until you finish work at this station.

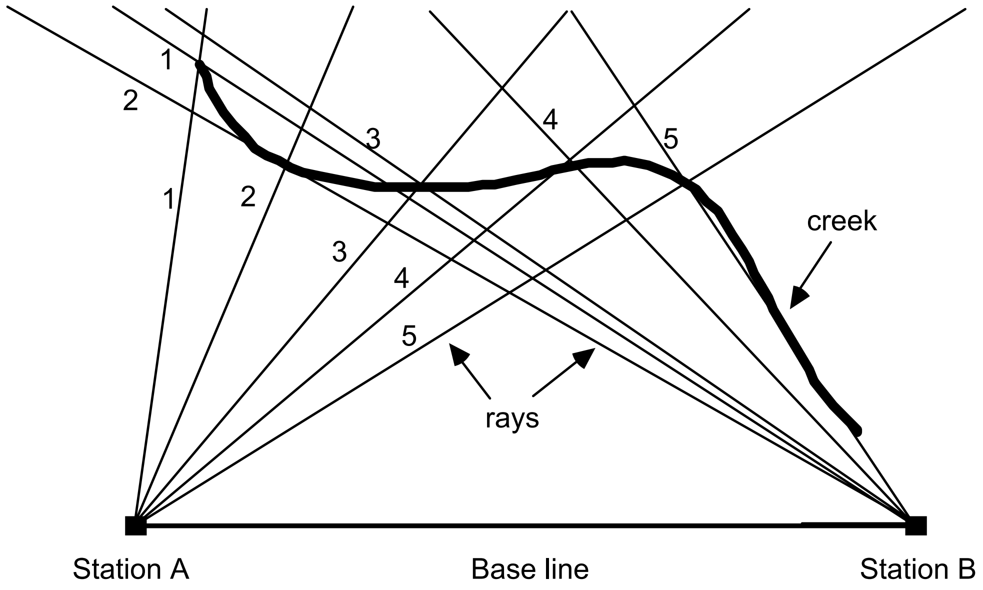


Figure 10-5 Example of plane table surveying.

You now need to send out group members to locate 'significant' points along the creek length which control its shape. The aim of this exercise is to capture the shape of the (say) salt marsh creek as effectively as possible using a conservative number of points. This process of data capture on the ground is a crucial step in any study and so you should consider how best to represent the actual shape of the landform in the necessarily simplified form of a map, either by choosing significant features you see or using other sampling techniques (such as mapping every 10 m along linear features).

As illustrated in Figure 10-5, the location of each position must be tightly controlled (usually through placement of ground pegs or nails) as each location must be visited twice. Hence, start by sighting the first peg location using the alidade: place the edge of the alidade against the location on paper of the station from which you are sighting and swing it round until you sight the person stood at the peg location. Now draw a pencil line from the station to the edge of the paper and label it with a number. Continue sighting peg locations and marking rays from the first station until you have captured what you consider to be enough points.

When you have finished at one station, move the table to the other and set up once again, remembering to rotate the table so that with the alidade placed along the paper base line you sight the location of the other station exactly. Now simply re-visit the peg locations drawing a second set of rays, which will now intersect the first set of rays. Mark each intersection with a circle when you make it. Make sure you collect all the pegs you have placed. Remember, if you miss a peg when measuring the second set of rays your locations will be completely wrong!

The results of this exercise are all on the map that you have made as you carry out the survey. The big advantage of this method is that you can see any major mistakes during the survey as you plot them, allowing rapid corrective re-surveying. At the crossing of the rays sighted from each station to the same peg location you should have marked the intersections. You can now join up the intersections with an inked line: this should give you the plan form of the creek to scale. One point to ponder - should you use a straight or curved line?

10.4.4 Resectioning

You can find your current position if you can see at least three conspicuous landmarks that also appear on a map or whose positions you know. Simply measure a compass bearing as accurately as you can to each of the landmarks and correct them for magnetic variation to true north, preferably using a sighting compass. Then draw a line from each landmark on the map at the angle of the bearing. They should meet at a point, which is your current position. If there are errors in the bearing, then the lines may form a small triangle, which is a zone of uncertainty, known as a ‘triangle of error’. When checking or adjusting bearings to reduce these errors, bear in mind that the amount of each error – the amount of ‘swing’ from the true bearing - will be proportional to the length of that line.

10.4.5 Tape and compass traverse

Basic surveys of small areas can be carried out using a surveyor’s tape and a compass to measure the length and orientation of features you can visit and delineate e.g. archaeological remains. Clearly, distances will only be true if the tape is straight and unknotted: this can be hard to ensure in windy conditions, in flowing water or where there is thick vegetation. Compass ‘bearings’ are measured relative to magnetic north and need to be corrected to true north by a small ‘magnetic variation’ angle to account for the fact that the magnetic north pole is in a slightly different place to the geographic north pole and is moving slowly (!), in a process known as ‘magnetic declination’. Information on this is usually given in the margins of local large scale maps, with a rate of change per year: however, you only need to correct bearings if you want to plot them on a map. Magnetic or metal objects worn on the body or carried with you like pendants, belt buckles and mobile phones can easily alter compass measurements unless you keep the compass clear of them. The compass traverse technique is particularly useful when surveying relatively flat terrain where there is poor inter-visibility of features: e.g. there are few prominent hills and/or the vegetation is too high and dense to make plane table or resection bearings.

10.4.6 Contouring

Using a barometer (or GPS receiver), and where it is physically practical to do so, you can attempt to walk ‘around’ a hill or ‘across’ a slope while remaining at the same height. Given a suitable map you can plot your route (and that of the contour) directly. However, it is probably best to carry out a contour survey in conjunction with a plane table survey or a survey traverse.

10.4.7 Slope steepness surveys

An inclinometer, such as those produced for foresters to estimate the heights of trees (e.g. by Suunto of Finland), or an Abney Level, can be used to produce very rapid surveys of slope steepness variations. Contours are not involved; instead, breaks of slope – whether sharp or gentle, convex or concave – are plotted to produce a *morphometric* map of the study area. This type of survey is often used in rapid geomorphological and geo-ecological surveys, and is particularly useful where details of soil type, drainage and vegetation cover are recorded at the same time as each slope steepness measurement. The slope steepness survey could follow existing footpaths as part of a compass traverse survey, but is most effective when following survey lines that run perpendicular to the main breaks of slope.

An ideal setting for this technique is a set of parallel survey lines, 100 m apart and running perpendicular to a baseline along the floodplain of a river, up to the crests of adjacent hillslopes, where another cross-line allows access to the survey lines and provides a check that they are still 100 m apart.

The first stage of this technique involves the observer standing as straight as possible, sighting the inclinometer against a nearby survey staff (or suitable piece of local timber!), while an assistant marks the height on that staff at which the inclinometer reads ‘0’ – i.e. the height of the observer’s eyes above the land surface. Inclinometer readings on the survey are typically taken at 25 m intervals, with the observer sighting the inclinometer on the ‘eye-height’ mark of the staff held by the assistant: the angle given by the inclinometer is the slope over the 25 m sample interval (see Plate 19). Small geomorphological features, such as gullies, can still be recorded using tape and compass measurements. An example of a slope morphometry survey sheet from a single survey line of a mineral exploration project in Indonesia is given in Figure 10-6.

Many of the major advances in land surveying techniques over the past 400 years resulted from the demands of marine navigation, as the nations of Europe set about world-wide exploration and trade. A key aspect of determining the course travelled by a ship was based on ‘dead-reckoning’, with distance estimated from lengths of knotted rope – hence the use of ‘knots’ for nautical speed – and direction calculated from the positions of the sun and stars. In essence this is similar to the tape and compass traverse surveying technique discussed above. A modern variant of this technique uses the odometer (‘mileometer’) readings of vehicles, coupled with compass bearings or GPS readings to navigate. This is a particularly effective navigation technique in remote, featureless and relatively flat regions, such as the Sahara and Namib deserts. Tom Sheppard (1998) provides a useful summary of dead-reckoning navigation techniques, including use of remote sensing images and GPS, in his *Vehicle-dependent expedition guide*. A variant of dead-reckoning vehicle navigation can be used as a sampling technique when carrying out a reconnaissance survey: the odometer is used to determine the distance travelled, with samples collected every kilometre.

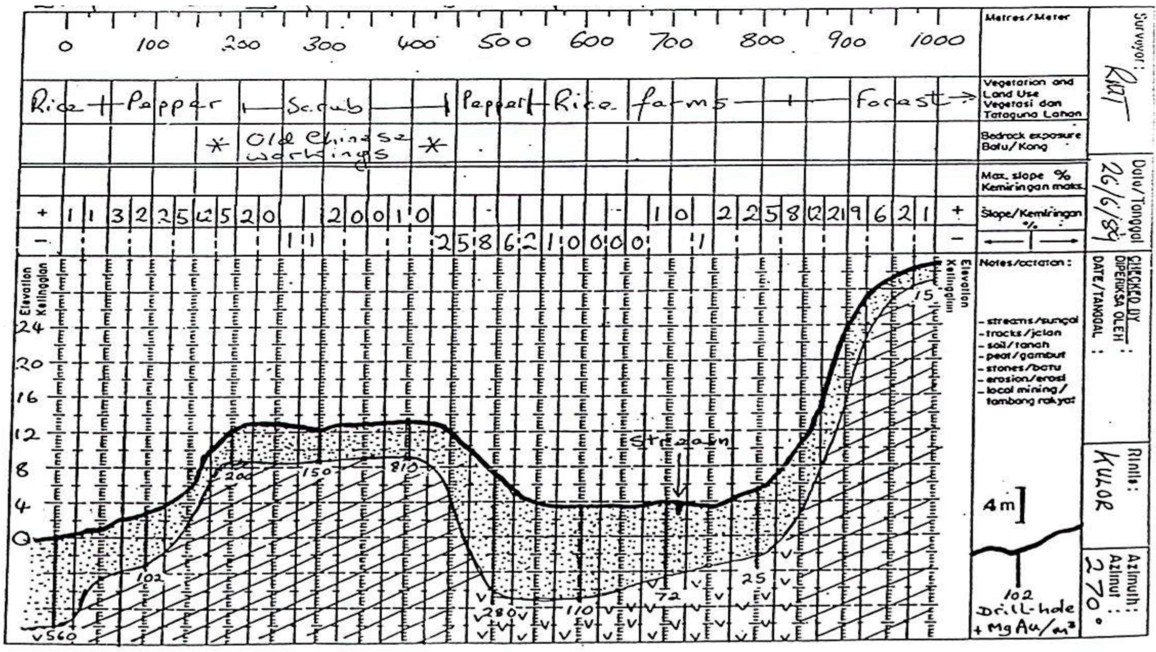


Figure 10-6 Slope steepness survey: transect line cross-section from an alluvial gold deposit, Indonesian Borneo.

10.5 Checking, logging and downloading survey data

Surveys are only as good as the data you record. Sometimes data is wrongly recorded, sometimes it is lost and sometimes you do not know what it means when you get back home. To avoid such tragedies, check the work as it is done, record the essential survey metadata and backup your results. Do not throw away data that you cannot work out: an expert may be able to deconstruct it later, and do not dispose of paper records once you have the data in digital form once back home, office PCs can crash just as easily as field-based computers!

A good survey can provide a framework for your project and can leave a record that future researchers can use. It is worth doing well and recording for posterity, you (or others) may want to expand your research in the future and your current survey may form the basis that underpins that future project.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section C: Techniques

Chapter 11: Using GPS for Fieldwork

11 Using GPS for Fieldwork

11.1 GPS applications

The NAVSTAR system has great potential for high-resolution surveying. The way the expedition uses the system will need to be decided before leaving for the field. This decision will involve weighing the need for accuracy against the cost and encumbrance of a very accurate system. Though there is often a perception that only the highest accuracy possible should be considered, there are strong arguments for using cheaper, quicker and less accurate systems. The basic types of GPS available include: standard sets, standard sets with WAAS augmentation, differential and carrier wave differential. The cost increases for each of these technologies but the benefit to the expedition may be slight. When used correctly a standard waypoint fix is often enough for most types of fieldwork. The accuracy required can be determined by the type of project and the quality of the data it is to be referenced with.

The first decision to be made is whether the GPS reading is simply to locate the object of study so that it can be re-found by another expedition or whether the recorded point is scientifically important. If the reading is simply to guide a team back to a location, then a single GPS reading will normally suffice. If the reading is significant then the accuracy is usually determined by the accuracy of the data it is to be related to. Taking a differential reading at 1 m accuracy to compare to a Landsat scene at 30 m accuracy is not a valid scientific exercise. In this case a standard GPS reading is sufficient; this is shown in *Figure 11-3*. Similarly, if the GPS is being used in conjunction with a map then standard GPS readings are valid down to 1:10,000 scale. Only maps of 1:5,000 require any form of averaging or WAAS and maps of 1:2,500 require differential readings. Mapping to a higher precision than 1:5,000 is rare and the area covered would be small. Most expeditions cover larger areas at 1:25,000 to 1:100,000.

11.1.1 Types of fieldwork & surveying using GPS

The nature of an expedition's data will dictate the accuracy required. The information contained in Chapter 6 should help you achieve these results. In most cases the observations the expedition makes as part of the standard field work will take more than a few minutes. This gives time to collect ample GPS data for averaging. However, averaging is always better performed by hand not using the averaging features of the set which are notoriously poor. Achieving GPS accuracy should not be pursued to the cost of the expedition's field measurements. The field measurements should always be paramount and GPS accuracy should only be pursued if the situation dictates it.

Environmental change analysis: Measuring changes in the physical conditions of the ground often requires very detailed analysis. For these surveys, averaging, WAAS or even differential measurements are required. Averaging is a good method because it does not require the use of cumbersome equipment, especially when mapping in difficult terrain but it takes much longer. If time in the field costs a lot of money then hiring differential equipment may offset this. There are often political sensitivities to consider when using high resolution mapping equipment in some countries.

Geological mapping: To record the location of lithological units for future reference it is only necessary to record a location to standard waypoint accuracy. From this data the outcrop could be easily found. For recording individual rock units in a GIS the standard GPS accuracy is generally slightly too coarse and the time required for averaging makes this unsuitable. There may be many units at one location. For detailed work of this nature it is best to use a combination of GPS and traditional surveying techniques. The resolution required will be closely linked to the resolution of the accompanying data in the GIS. If a Landsat or ASTER scene is being used in combination with the collected data, then accurate marking of units may not be necessary, as they will not be shown on the image.

Habitat mapping: Mapping the habitat of animals or the location of observed animals is only usually necessary to the accuracy of standard GPS waypoints. The location of an animal in the field rarely needs to be more accurate than tens of metres because by their very nature they are mobile. For boundary mapping where a vehicle is used then an external antenna and a standard GPS is the best solution. Using a GPS in a vehicle without an external aerial is not ideal even when the receiver has a reasonable lock through the windscreen, because its view of the sky will be severely restricted. Some units do not have external antennas such as the Geko and ETREX range from Garmin. Care must be taken to read the specification of GPS units before selecting one. This research should not rely on the schematics in this manual that serve only as a basic guide. Check with the manufacturer's websites listed in Appendix 4.

Image rectification: Digital images often need to be given co-ordinates. This can be achieved from a map, by referencing common points or can be achieved in the field using GPS. The accuracy of the rectification will be affected by the accuracy of the data collected and commonly high precision data is required. However, the maximum rectification possible in imagery is related to the pixel size. For data sets such as Landsat TM and Landsat ETM+ multispectral, a good quality single waypoint will be sufficient. For more accurate data such as Landsat Pan or SPOT then averaging or WAAS might be more appropriate. For high-resolution satellite data, such as KVR, CORONA, or the new high-resolution IKONOS or Quickbird sensors, differential or carrier wave may be a consideration. These choices depend on the confidence that can be given to finding similar objects in the field as on the images. Where confidence in the control points is low there is no benefit in high precision or expensive data collection. This process is discussed in detail in Chapter 9 and outlined below in Section 11.6.

Detailed Boundary Mapping: Mapping political boundaries is often not substantially more accurate than a standard GPS fix. Small areas will on occasion require a more precise fix but this depends on the type of work conducted. For wildlife habitats it can be useful to know locations of reserves accurately but if the boundary is large, then averaging is not viable because of the length of time required to get a statistically meaningful fix. WAAS may help in these circumstances or differential GPS techniques. Modern surveying often requires high-resolution data and so are best suited to differential GPS work.

11.2 GPS care & power requirements

GPS receivers are generally robust pieces of equipment but the difficulty in repairing or replacing them in field locations means care must be exercised. GPS sets commonly work between –15°C and +70°C. This is as much a product of the operating conditions of liquid crystal used in the display as the operating conditions of the machinery. Battery life in these lower conditions can be lower but the GPS should function without ill effect. Below –20°C liquid crystal stops functioning and the display on the GPS may become irreparably damaged if held at this temperature for long periods of time. Specially designed GPS receivers equipped to function below these temperatures can be bought; examples include models from Silva that can extend the working to temperature to –25°C but much below this GPS receivers are of limited use. While working outdoors the GPS screen is prone to damage and plastic faced holders can be purchased to protect the receiver screen. These wallets can also give some protection to the receiver from poor weather conditions. Many GPS sets are splash proof but only newer models conform to waterproof standards (see Chapter 13). Modern sets are waterproof to 1 m for 30 minutes. Many of the newer Magellan models are waterproof and are designed to float.

Power requirements are discussed in depth in Section 13.6. However, it is touched on here for reader convenience. Most GPS receivers are powered by 1.5 volt AA batteries though Garmin Gekos use AAA batteries. Modern sets require two batteries though some older sets require four or even six. GPS receivers can be run in one of two modes, standard (continuous positional updates) and battery save (updates position once a second). For expedition use battery save mode is ideal and will yield around two days of field time per set of batteries. Generally standard rechargeables are not well suited for GPS units because of their voltage fall off and their lower overall voltage. Battery life might be less than a day with these sets, however, modern NiMh rechargeables with over 2000 mAh charges might last over three days.

11.3 Using GPS in non-ideal conditions

Section 6.9 describes the standard methods for improving GPS accuracy. These include WAAS, differential and averaging. This section looks at how the selected GPS should be used in the field to better ensure the chosen unit functions adequately. GPS receivers need clear views of the sky to operate. If placed in a pocket close to the body their view of the sky is impeded and the data collected will be of poorer quality. Care is also needed when using GPS in forested areas where canopy cover can disrupt the signal. These problems can be circumvented to some extent by using an external antenna. Many GPS receivers can take an external antenna via a small input socket on the rear or side of the unit. The external antenna can be attached to the receiver and extended to give a better view of the sky. Before selecting a GPS it is vital to decide if this is important for your expedition. Some popular models do not support external antennas. Units such as the Garmin ETREX range do not support this functionality. One use of an external antenna is to get the antenna off the ground and thus give it a clearer view of the sky by attaching it to a pole. If an expedition is planning to use a receiver under tree cover then getting the aerial just a few metres off the ground helps the signal enormously and cuts down on reflected signals from the forest floor. An aerial held up into the canopy or simply mounted on a backpack so signals are not obscured by an expedition member's body will benefit signal acquisition. One expedition-proven solution is to (1) place a bottle top (or other small metal object) on your head; (2) put on a baseball cap; and (3) put the magnetic antenna on top of the cap – it stays firmly in position, providing a low-cost, mobile mast. Another advantage of this system is that you can pocket the GPS while walking between sites, without losing a signal. For an example of how GPS was used in African rainforest see Dominy and

Duncan (2001). Signal reception is also a concern in vehicles, as GPS signals cannot travel though a metal body. To use a GPS receiver in a vehicle also requires an external antenna. If you are using GPS mostly in a vehicle, make sure you have an antenna-compatible model (e.g. Garmin 12XL or GPS 76 not an ETREX). More details on the use of vehicles in expedition work can be found in Sheppard 1998.

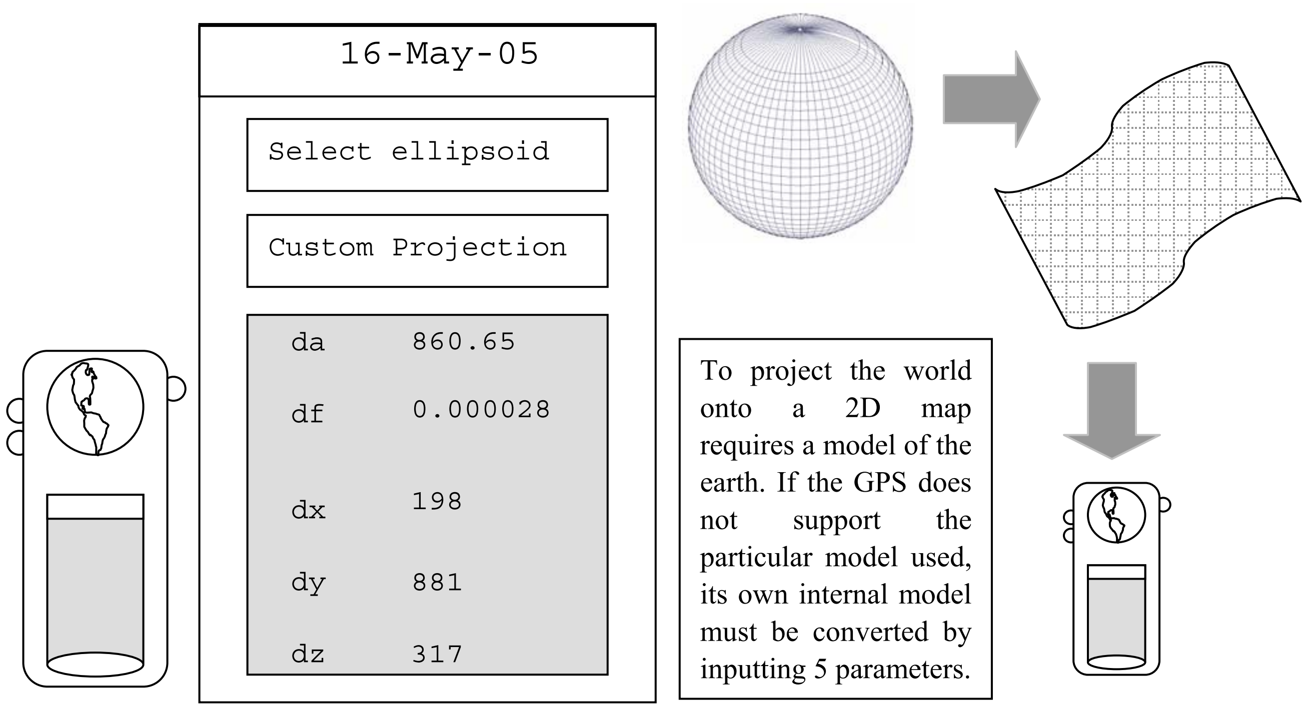
GPS receivers offer global coverage but the accuracy of NAVSTAR is greatly reduced at the poles. For northern polar expeditions, the GLONASS satellites may be a more viable alternative. GLONASS satellites orbit at up to 65° north/south, and give better polar coverage than NAVSTAR. Take care when switching between systems, as (1) most receivers are not dual constellation compatible, and (2) GLONASS operates in the PZ 90 datum, rather than WGS84. If this is not an option, NAVSTAR systems work at polar latitudes but their time to acquire and dilute will be greater.

11.4 Integrating GPS readings and historical maps

All NAVSTAR GPS work internally according to the WGS84 ellipsoid and datum. Because the expedition will no doubt visit areas with existing maps, these settings will need to be changed so that the data shown on the GPS screen is compatible with the map. Modern GPS support an increasing number of projections and datums but invariably at some stage the team will come across an unsupported map. In these circumstances the user can input a user defined datum into the GPS.

To transform the WGS84 co-ordinates to a custom set of parameters the GPS needs to know at least 5 pieces of information. These are the difference between the radius of the WGS84 and the custom ellipsoid (da), the amount by which the ellipsoid is flattened (df) and a three dimensional co-ordinate for referencing the ellipsoids together (dx,dy,dz). Usually df will be a very small number, typically < 0.00005. Therefore some GPS units will require reciprocal flattening. For example, if the expedition was working in SE Asia using a GPS in conjunction with maps using an Indian datum such as Indian Vietnam 1960 on the Everest 1830 datum then the GPS may not have all the necessary parameters in its database. In this case it is comparatively simple to research the required parameters and input them. A search on Google for example will invariably point you in the correct direction. An example of how the WGS84 readings can be converted for a SE Asia expedition by the GPS is shown below in Figure 11-1.

The Everest 1830 ellipsoid has a radius of 6,377,276.345 m against the WGS84 measurement of 6,378,137 m. The difference between the two is 860.65 metres so this is the da reading. The flattening of the Everest model is a factor of 0.003324449297 compared to the WGS84 figure of 0.003352810665. This gives a difference of 0.000028. Because this is such a small figure sometimes df is calculated as 1f-1/f (300.8107 - 298.2572 = 2.5445). The Cartesian co-ordinates dx, dy and dz are fixed figures and for referencing WGS84 and Indian Vietnam 1960 are 198, 881 and 317 respectively. To input these into the GPS consult your manual, navigate to the settings page and select the option for projections. Enter them as shown below. Once this has been done all the GPS readings will match up with your map. As mentioned in Chapter 6, the GPS is not perfect at its calculations and when working outside of WGS84, it can add another 5 m to the inherent errors.



To project the world onto a 2D map requires a model of the earth. If the GPS does not support the particular model used, its own internal model must be converted by inputting 5 parameters.

Figure 11-1: Inputting user datums into a GPS.

The five parameters above may not fully describe the map projection. Ideally the GPS also requires a meridian or origin (central meridian), a scale factor and information about any false northings or eastings.

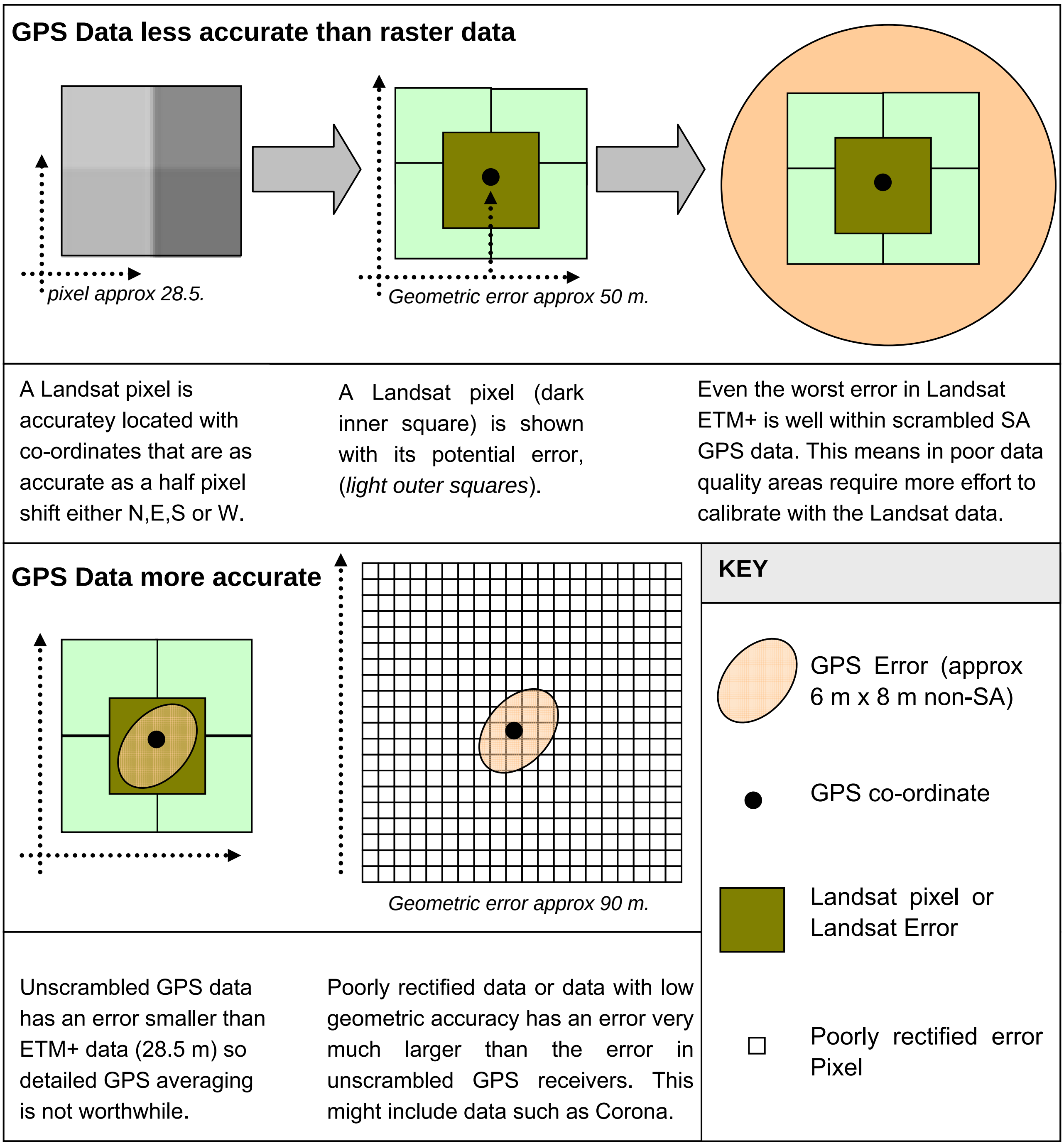
11.5 Integrating GPS and other GISci data

GPS data is often useful for navigation purposes but often it will need to be stored and referenced later. If these data are stored in the GPS memory, it will often need to be accessed or downloaded to a computer. Most GPS receivers come with a cable interface to connect between a computer and the receiver, but the cable is often an additional product and can be quite expensive. Some of the more expensive GPS receivers come with free cables making them a more economical solution. It is very important that the cables are compatible with the hardware the expedition is using in the field. Most GPS receivers use a unique connector that plugs into a serial (9 pin male RS232 port) on the computer. It is important to check that the computer has a relevant port because many new PCs and laptops are only shipping with USB connectors. A USB to GPS cable can be purchased in some cases but this is rarely free with a receiver. A serial to USB converter can be obtained but these often need drivers to emulate a virtual COM port. These cables cost between £6 and £25 GBP depending on length and whether a driver is supplied. An example of this type of converter is shown below in Figure 11-2. Converters can be obtained to change 9 pin serial into 25 pin serial, which at first glance appear to be compatible with a parallel port. Parallel ports are faster connectors used commonly with printers and are available on most PCs. The converter cannot change the serial signal into a parallel one so these do not usually work. It is vitally important to check that all equipment communicates successfully before leaving for the field. There is also an advantage in keeping expedition kit as similar as possible between all members. This is especially true of GPS where manufacturers use different proprietary connectors on the back of all units. If a cable is lost or damaged and the expedition kit is similar between all members then another member's cable can be substituted in to continue downloading and using a receiver. In addition, with the complexity in setting up a receiver, there is a time benefit in not having to learn each receiver's start-up process. When using a more exotic form of GPS receiver it is vitally important to apply stringent tests with all field equipment. Units that communicate via Blue Tooth do not always send their data stream out in a standard NMEA 4800 bps signal. Units such as the AnyComm GPS 600 shown in Chapter 6 are ideal for expedition work in that they are field hardened, but communicate using 38,400 bps. Some software will not cope with this form of data possibly rendering the equipment useless to the expedition.

Figure 11-2 Serial - RS232 9 pin male to USB converter. The RS23 serial connector is shown on the left and the USB connector is shown on the right. The RS232 connector receives the female GPS connector and transfers the data through the USB port of the PC.

Software is also not included and this is an additional purchase. Several shareware or freeware products exist to download the data into a PC and these are discussed in detail in Chapter 14. When a GPS outputs its data to a PC, usually as an ASCII text file, it can be imported into a spreadsheet or database file. This data can then be input into a GIS for spatial analysis. The data held in the spreadsheet or database is usually in a series of columns consisting of latitude, longitude, elevation and additional information. What data the GPS makes available for downloading is not a constant across all units and models. Many units do not download elevation data for tracklogs and this can seriously compromise an expedition that requires this data. Most modern receivers will download elevation for way points but only X and Y can be taken for granted. The type of data output from the unit can be vital to an expedition so it is important to check the output strings with the manufacturer before purchasing the unit. A good tool for downloading data is GPS Utility. This can export the data as *.dbf or as Arc Shape Files that can be uploaded very quickly into a GIS.

Once the data has been output from the unit it can be overlain onto other data in the GIS. Though this is a comparatively simple task and usually quite accurate, understanding the exact relationship between GPS data with other data stored in the GIS is more complex. The GIS system used for collating the field data can be configured to be as accurate as the co-ordinate system used (i.e. metres for UTM, decimal degrees for latitude and longitude etc.). However, the correlation of data within this system will not always be of this accuracy as shown in Figure 11-3 and described in the text below.



GPS Data less accurate than raster data

A Landsat pixel is accuratey located with co-ordinates that are as accurate as a half pixel shift either N,E,S or W.

A Landsat pixel (dark inner square) is shown with its potential error, (*light outer squares*).

Even the worst error in Landsat ETM+ is well within scrambled SA GPS data. This means in poor data quality areas require more effort to calibrate with the Landsat data.

GPS Data more accurate

Unscrambled GPS data has an error smaller than ETM+ data (28.5 m) so detailed GPS averaging is not worthwhile.

Poorly rectified data or data with low geometric accuracy has an error very much larger than the error in unscrambled GPS receivers. This might include data such as Corona.

Figure 11-3 Approximate error ellipses when correlating data in a GIS (shown to scale).

The GPS data will have an inherent error, and the digital data (maps or images) will also have an error. The error in placing a location on an image is determined by the geometric accuracy of the data. This is a measure of the difference between the co-ordinate given to a point on an image and its actual co-ordinate location. Modern digital images generally have a very high geometric accuracy with only small amounts of shift. Landsat ETM+ imagery has a geometric accuracy better than 50 m, commonly quoted to be about one pixel (Landsat website). The Landsat ETM+ data is commonly supplied georeferenced with corner co-ordinates given in a separate text header file. The high quality geometric accuracy of modern digital images means the errors in overlaying the GPS points are controlled mainly by the error of the GPS technology used. When Selective Availability was an issue, this error would have been of the order of 100 m, considerably larger than the error in the projected Landsat data. The disabling of this scrambling has increased the accuracy by an order of magnitude. Current accuracy when using 12 channel sets is of the order of one pixel (10-15 m) in Landsat ETM+ panchromatic and is similar to the inherent error in Landsat data. For older data such as analogue data that has been scanned into a computer, the error is controlled by the quality of rectification. In the Bogda Shan study the CORONA data has an error of ± 90 m. This was compounded by the 15 m accuracy of the GPS sets giving an overall error of over 200 m.

In the top two images in Figure 11-3, the dark green central square represents the Landsat pixel. This pixel has a geometric error meaning it could be shifted to any of the light green outer pixels. When using scrambled GPS data the Landsat error is less important than the GPS error. When using unscrambled data (any GPS reading after May 2000) the Landsat error is more important than the GPS error. This has a bearing when selecting the type of GPS to use in the field (i.e. standard, WAAS or differential). Even high resolution data, such as the Corona, shown in the bottom left, which apparently looks like very high resolution, quite often has an error much larger than the GPS error, meaning standard sets are still compatible with high resolution data.

During the Bogda Shan Expedition (see Section 17.1), which used receivers after the end of SA scrambling, the correlation of the point and image data was most significantly affected by the inherent image error with Landsat, or the limits of rectification with Corona. The maximum offset for a series of points measured in the field was measured in metres and found to be a function of GPS error and image error. The GPS error is calculated as the sum of atmospheric and local effects as discussed in Chapter 6. This error is referred to as dilution of precision or DOP.

$$\text{Correlation Error} = \frac{\text{GPS DOP} = (\text{Atmospheric} + \text{Local Effects})}{2} + \frac{\text{Image Error}}{2}$$

This error describes the error when plotting a point onto the image. It does not take into account the error of locating an object in the image and giving this a location. When trying to overlay data about a specific point in an image, the spatial resolution of the image must be taken into account. For example, the Bogda Shan Expedition tried to map the movement of the front of a glacier over a 40 year period using various datasets. The Expedition recorded co-ordinates for the front of the glacier and tried to calculate the movement based on the co-ordinates of the GPS against the image. Because the pixel size of the image dictates the accuracy, this can be combined with the error in correlating the two datasets. The overall error is given by:

$$\text{Correlation Error} = \frac{\text{GPS DOP} = (\text{Atmospheric} + \text{Local Effects})}{2} + \frac{\text{Image Error}}{2} + \frac{\text{Image IFOV}}{2}$$

For GPS to Landsat that might mean a correlation error of:

$$\text{Correlation Error} = \frac{15}{2} + \frac{50}{2} + \frac{28}{2} = 7.5 + 25 + 14 = 46.5$$

This means an approximate error of 45 m was introduced when correlating the data. Any GPS reading of an object on the Landsat image would have a 45 m error so for the Bogda Shan Expedition the glacier would have had an error of ± 22.5 m for the Landsat data. The Corona data was substantially worse because the image error was so much higher. Although the spatial resolution of Corona is so much better, the actual measurements taken are significantly worse due to the poor control on rectification. The accuracy of the data in Corona would have been of the order of ± 60-100 m depending on the rectification process. This is described in the following section.

11.6 GPS receivers for creating expedition maps

Digital images can be used as accurate and useful expedition maps. In some instances, the purpose of an expedition might be to create a map. GPS are excellent tools for either making maps on their own or making maps in conjunction with digital images. When used on their own GPS receivers can be used to map out linear features using their tracklogs to a very high degree of accuracy ~1:10,000. This might be useful when visiting an area whose road network is poorly mapped or not known. Using the GPS while doing a day's reconnaissance would give a quick and accurate map of the road networks.

Figure 11-4 The use of vehicles in rapid reconnaissance and map making. When first arriving in the field a vehicle can be used to quickly assess the area and produce a rough basemap.

If the aim of the expedition is to create small maps and use them in the GPS then it may be in the team's interest to acquire a GPS with a larger screen and better resolution than some of the smaller units. If maps, such as the one shown in Figure 11-5, are used in the GPS, then a unit such as the GPS76 from Garmin may be preferable to a unit such as an ETREX.

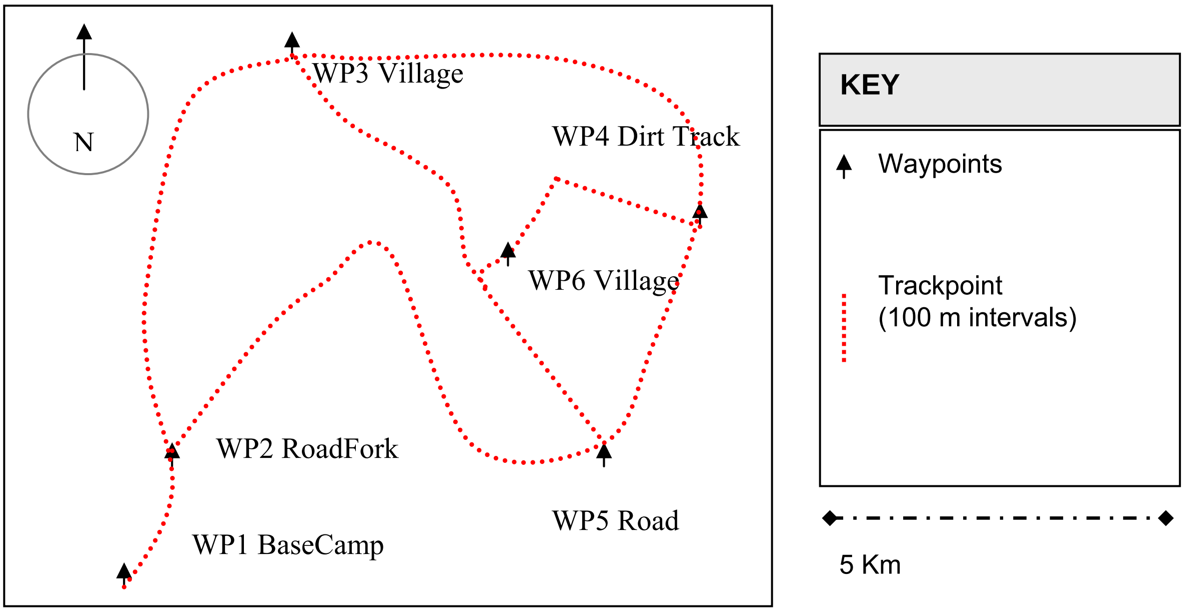


Figure 11-5 Schematic point plot of GPS Data showing road and tracks in a study area annotated with way points. This type of information can be displayed from the GPS screen or it can be downloaded into a Utility such as GPS Utility or downloaded and plotted into a GIS.

Figure 11-5 shows how GPS data can be used on its own but GPS readings are most useful when they can be combined with some form of digital data. Alternatively, analogue data may exist that needs to be corrected to turn into a map. The process of georectification is described in detail in Chapter 9. The information here is purely to illustrate how a GPS should be used when collecting field data to be fed into a GIS or image processing program for rectification purposes. Commonly digital imagery that has no co-ordinates is rectified off high quality maps. However, in many of the areas an expedition team may aim to visit, maps might be poor quality, restricted or simply non-existent. Three results of a rectification exercise are shown below to illustrate the power of using GPS in the field for image rectification. The Bogda Shan expedition rectified data using image-to-image rectification. The Landsat ETM+ data was used to correct the Corona data. This was the best that could be done before leaving for the field because Chinese maps were not available. Each pixel had an error of ~± 80 – 90 m in X and Y. This is shown schematically in Figure 11.6 for errors in the X co-ordinate.

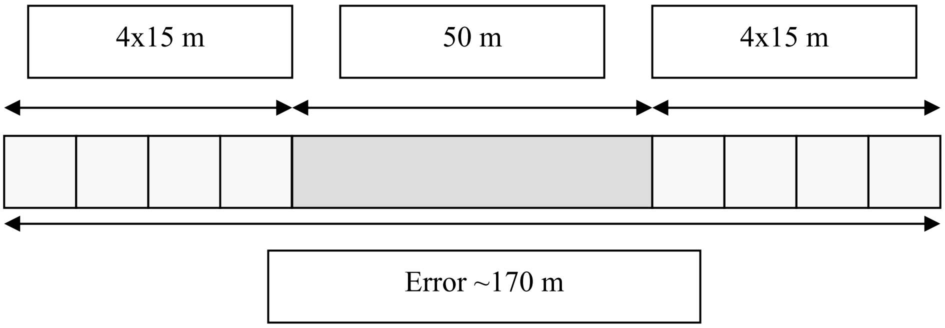


Figure 11.6a Errors generated from rectification (in X dimension using Landsat ETM+ data).

A far better method is to take a GPS unit and map the co-ordinates reported from the receiver to the scanned data (this can be seen in Appendix 2). GPS receivers can be used to rectify data to a much greater degree of precision. By locating areas in the field from the CORONA data and recording the co-ordinates for that location allows georeferencing in the field. If the team could find a visible outcrop in the field to within a 10 pixel square on the CORONA imagery, the error would be equal to the GPS error ~20 m (3σ error) + error of 10x10 pixels locating pixel in the image.

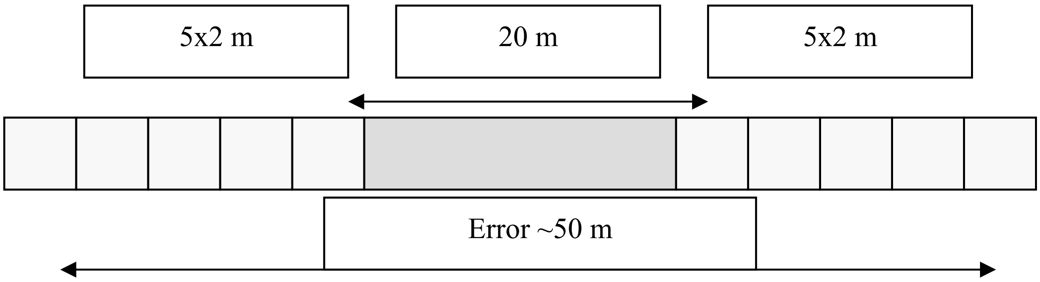


Figure 11-6b Errors generated from GPS rectification (in X dimension)

This method gives an error of ± 25 m, which is an improvement of three times on the current method. This is already a conservative estimate of GPS resolution and it is possible to improve GPS confidence to significantly better than 20 m with averaging, pseudo range, dGPS, or WAAS. By using one of these methods the error in rectified digital data could be reduced to about 25 m (± 12 m).

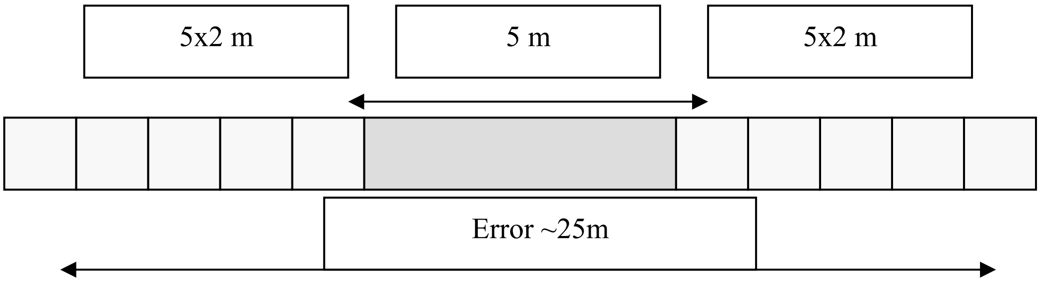


Figure 11.6c Errors generated from processed GPS rectification (in X dimension).

The better the correction used, the better the final product. If an accurate map is being constructed then the only realistic method is an averaged GPS position or a WAAS enabled GPS receiver.

Table 11-1 Characteristics of satellite image data.

Type of data	Resolution of image	Comparable Scale
Landsat	28.8 m	1:100,000
Landsat ETM+	15 m (panchromatic)	1:75,000
Corona	2 m – 10 m	None (no co-ordinates)
Rectified Corona (Landsat)	2 m – 10 m	1:200,000
Rectified Corona (GPS)	2 m – 10 m	1:50,000
Rectified Corona (Averaged GPS)	2 m – 10 m	1:30,000
Rectified Corona (dGPS or WAAS)	2 m – 10 m	1:25,000

These examples show that GPS receivers are an excellent method for rectifying raw images and making base maps for expeditions. An example of how to use GPS data for rectification purposes is shown in Appendix 2.

11.7 GPS receiver models

Several companies manufacture GPS receivers and a 'catalogue of common sets' is shown in Appendix 4. The accuracy of similar channel models is not substantially different between the companies. The two major manufacturers of entry-level sets are Magellan and Garmin. Traditionally Magellan is more commonly associated with marine navigation and Garmin with recreational hiking or aviation. Both companies provide basic, cheap sets that would be ideally suited to an expedition. Other companies such as Silva, Trimble, MLR and Lowrance also produce receivers but many of these are for more specialised pursuits. In general, all models are created to the same standards and more expensive models will not give better results. Magellan has made some significant steps forward with the Meridian and Sportrack ranges. These offer a very reliable WAAS service (where WAAS coverage is available) and a more powerful active antenna that offers the best civilian performance available under canopy cover.

The list in Appendix 4 is by no means exhaustive and may not include recent firmware updates. When considering purchasing a GPS unit consult the manufacturer's website to determine the exact specifications or cheap mail order companies such as GPS Warehouse (*www.gpsw.co.uk*) or Expansys (*www.expansys.co.uk*). The following costs are accurate as at June 2004 from mail order companies. High Street stores may be more expensive. Costs of GPS receivers rose slightly in 2002 and stayed fairly static throughout 2003. A price drop of around 10-15% came in mid 2004. For larger orders it may be worth considering purchasing abroad and paying import duty on units, as this is often cheaper than buying in the UK (for a fuller description of this see Section 13.11). The features list gives an idea of the features of each GPS but is not exhaustive. If a feature is not listed it is not an automatic indication that the GPS does not have it.

All the units listed in Appendix 4 are 'standard' types of receiver meaning they have an antenna, locally based software for calculating a position and some form of LCD screen for displaying the result. There are other types of receivers available including Bluetooth, PDA and CF cards and AGPS receivers. These more exotic units are often built for in car navigation or other specific tasks. When purchasing these units for an expedition, make sure they are not sold with included navigation software because this will double the cost of the unit. The first type of GPS specifically built for a PDA was the NavMan GPS500 built for the original Compaq IPAQ. This used a jacket to slot over the IPAQ. Subsequent PDA GPS units have used more generic interfaces and the current most popular method is Bluetooth. Bluetooth GPS include the Navman GPS 4460 and AnyCom GP 600. The GPS 4460 has a 30 hour battery life and the AnyCom lasts for around 16 hours. Some PDAs are shipping with integrated GPS antennas. The Mitac Mio-Digi-Walker has a GPS antenna built into it. This saves some money by not having to purchase the individual units. The Mitac unit runs a 300MHz PXA255 with 64 Mb of RAM. These units may be useful to an expedition but PDAs are expensive compared to standard GPS units. Also GPS are finding their way into mobile phones with units such as the Motorola A835 having chips integrated into them and Nokia 5140 supporting GPS clip on antennas. Care must be taken when looking at a mobile phone GPS because often the phone lacks the processing power to calculate a position. The data is actually received from the satellite and sent via GSM to a central computer for processing. The data is then sent back to the phone, sometimes with a map of the current location. This means the phone GPS will only work where there is GSM coverage and the service provider often charges for the position calculation (approx. 10p per use). An example of this is shown below.

Figure 11-7 The use of Assisted GPS on a smart phone. This can give reasonable accuracy even in areas where normal GPS signals do not reach. However, it is expensive and poorly suited to most expeditions. A better solution for areas of poor signal strength is an external antenna connected to a standard receiver or a SIRF III 20 channel receiver.

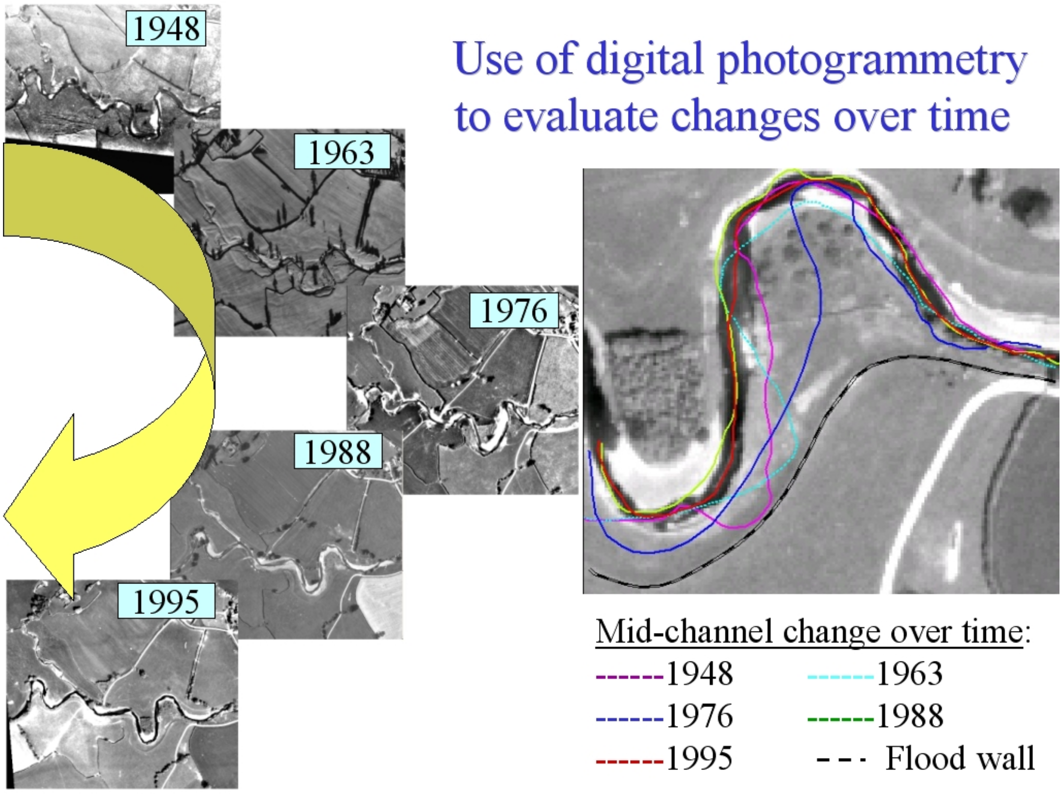


Plate 1 Use of scanned archive aerial photographs, digital photogrammetry and GIS to plot river channel erosion along the Afon Trannon, Wales (part of the study by Mount et al 2000, 2003).

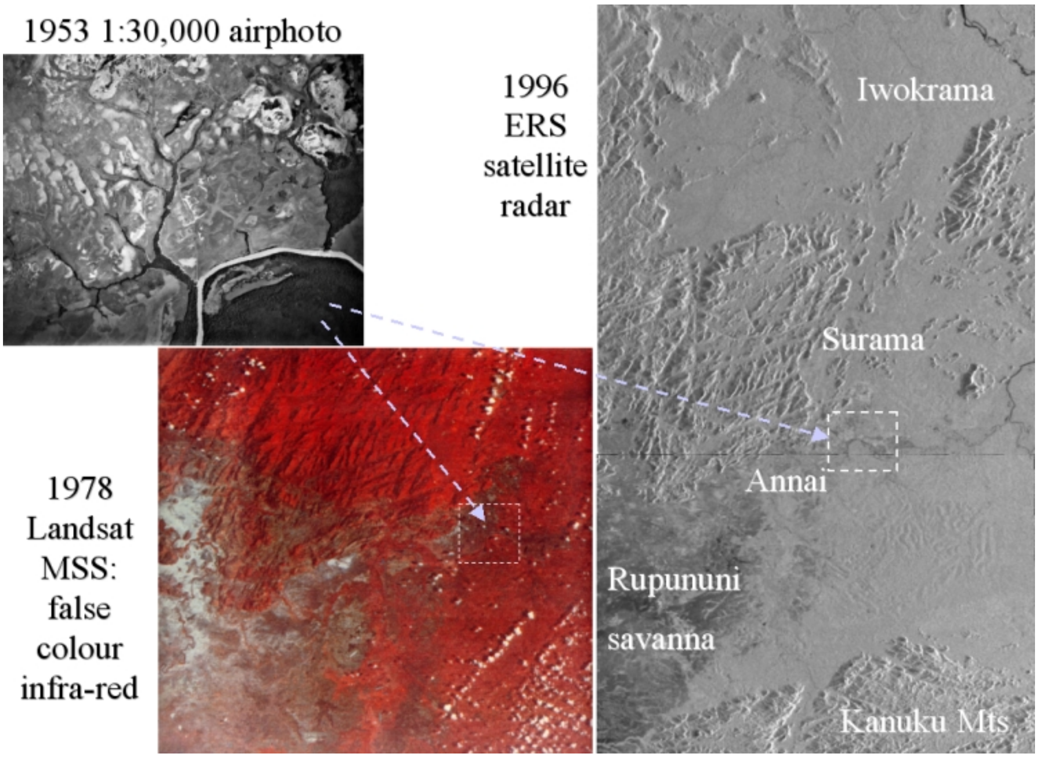


Plate 2 Part of the rainforest-savanna boundary in Guyana, viewed from three different sensors: panchromatic aerial photography, false-colour infra-red Landsat MSS and radar from the ERS-1 satellite (courtesy of the European Space Agency).

Plate 3 Do-it-yourself aerial photography: shot from the window of a passenger aircraft in West Africa, using Kodak Ektachrome infra-red film. Chlorophyll-rich vegetation shows as dark tones, distinguishing lush, often forested, valley swamps from dry grasslands (grey) and bare soil (white).

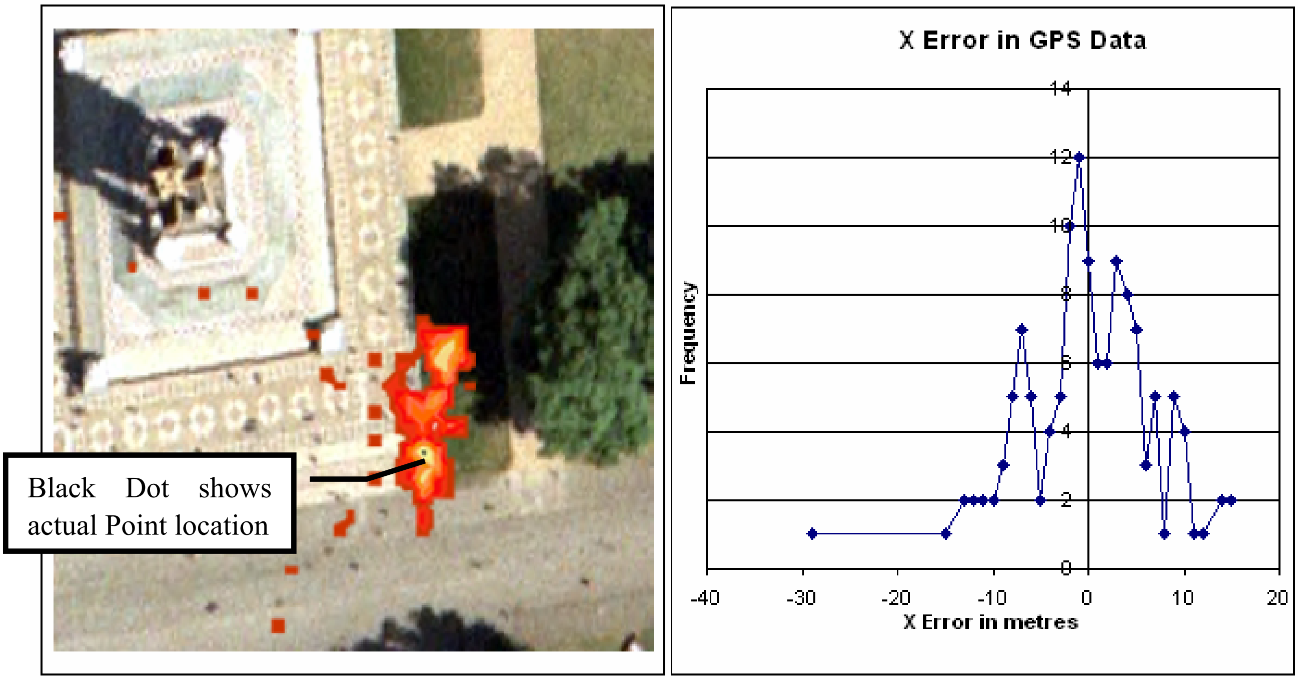


Plate 4 GPS error. Left hand diagram shows a colour coded frequency distribution of GPS results around a location. The lighter the red colour the more times the GPS reported that location. Spread approximately 35 m X and 25 m Y. Apart from two tails of data in the X dimension, X values are grouped very tightly together (within 5 m). Y values fluctuate more (15 m). The diagram to the right shows this more clearly with a cross section through the X axis. The cross section shows an almost Gaussian distribution. Dropouts are most likely due to insufficient data points.

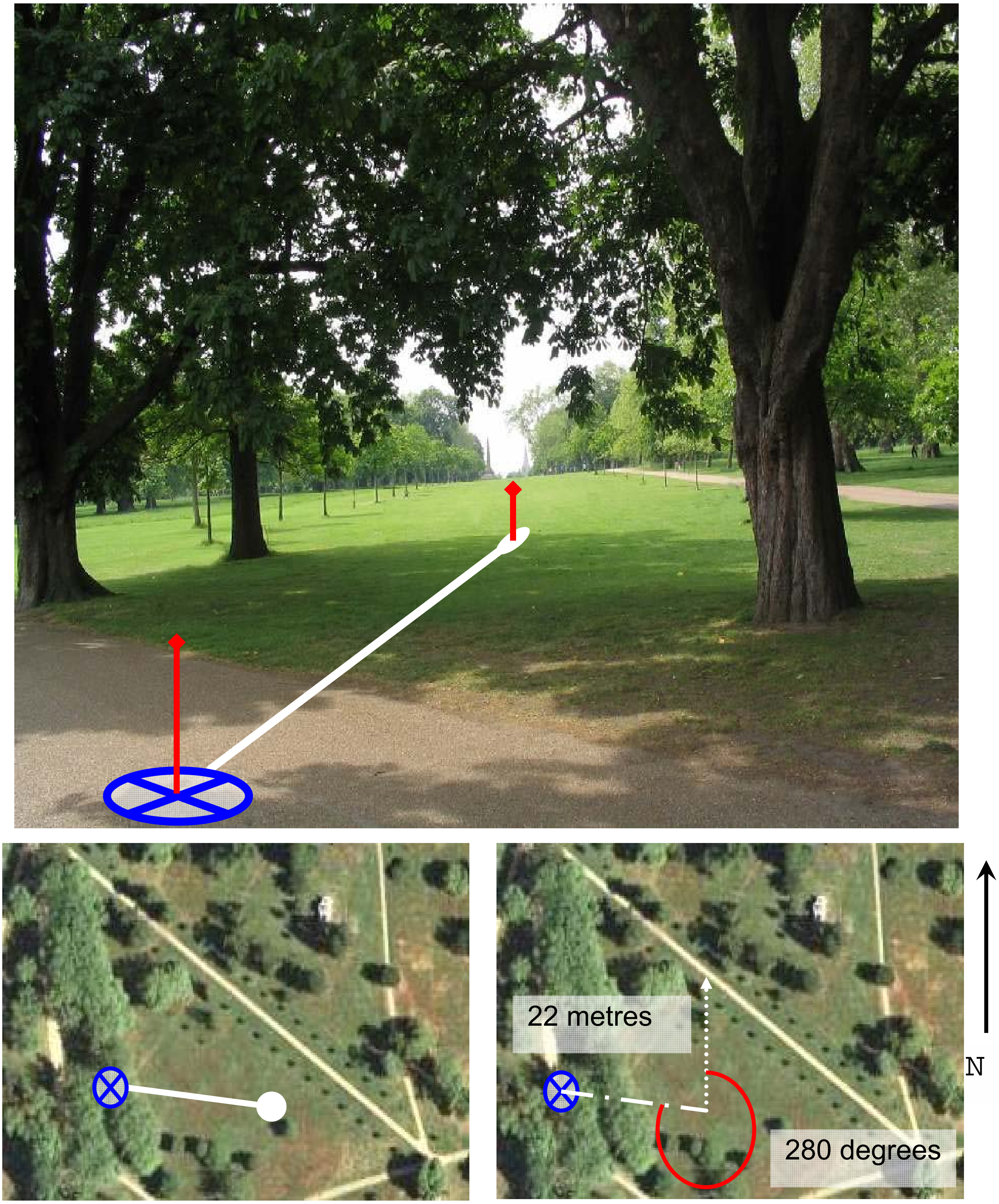


Plate 5 Using surveying techniques with a GPS to improve fixes in difficult locations. Top: a GPS reading is required at the blue cross under the tree cover. However, a few metres to the east the sky is not obscured by trees: the accuracy of the fix here will be an order of magnitude better. Lower diagrams: the receiver has been taken into the open and a reading taken. A compass has then been used to site the actual location required and a measurement has been taken along the line between the two points; this can then be plotted back into the GIS later.

Plate 6 A query on GPS fixes from the Bogda Shan expedition: locations above 1,000 m are selected in a lighter colour (top half of the image); locations below 1,000 m are displayed but not selected (lower right of image).

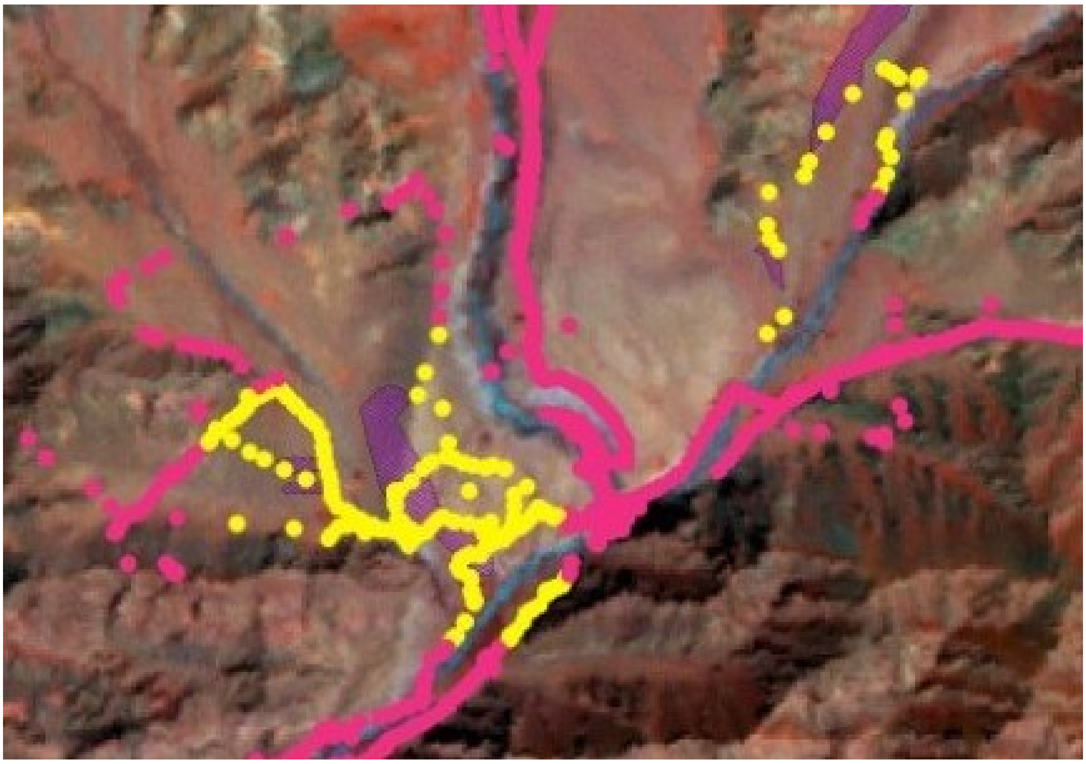


Plate 7 Results of a selection using a spatial query in the Bogda Shan GIS: GPS locations within 0.5 km of moraines are highlighted; moraine areas are shaded in grey.

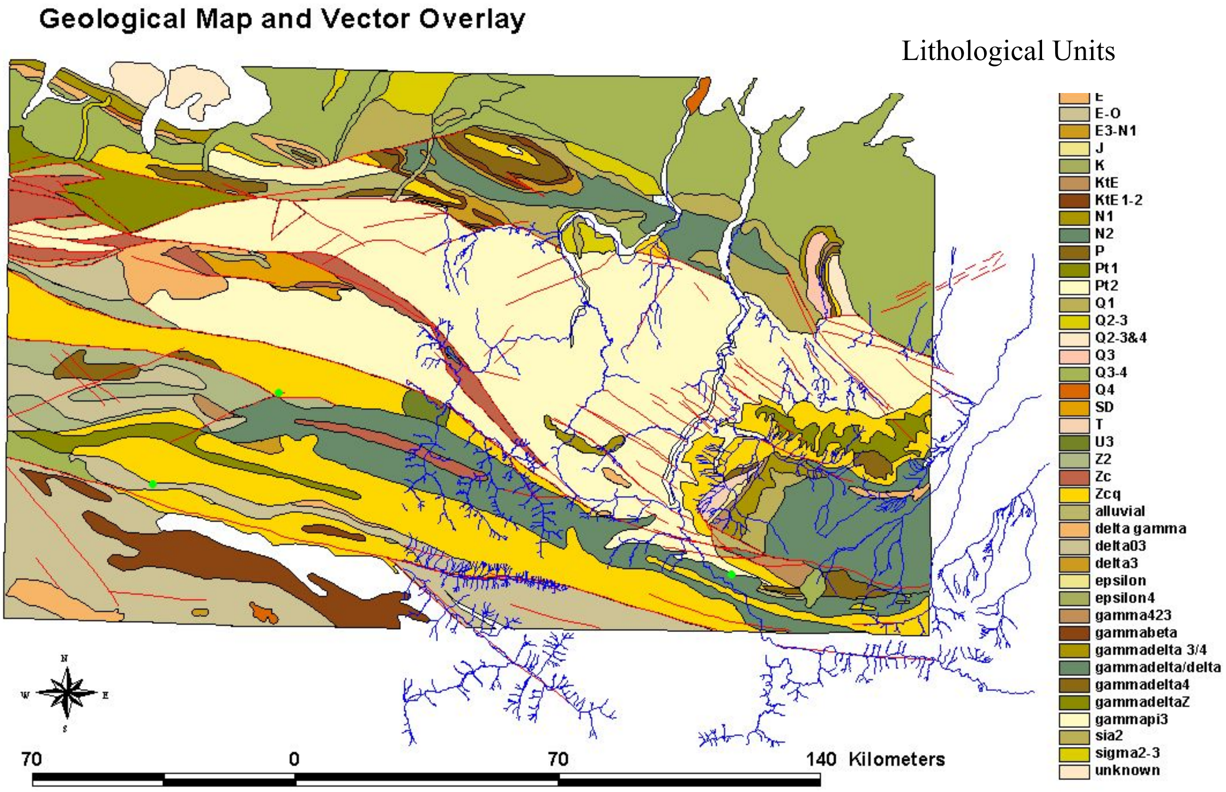
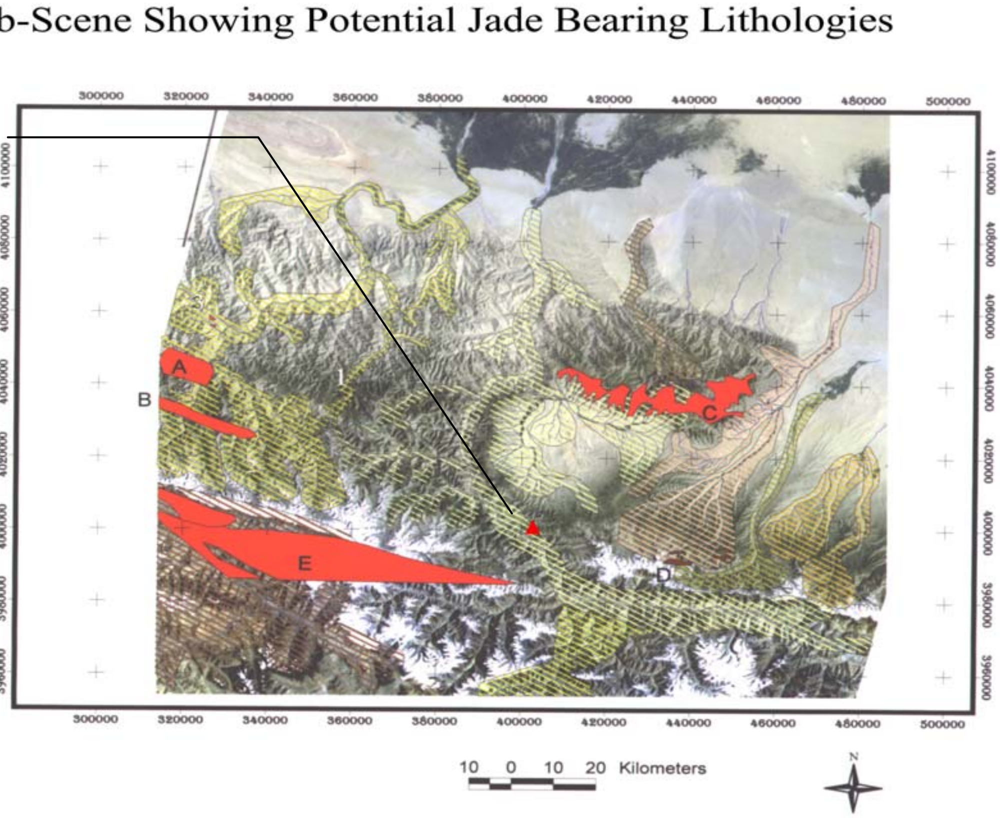


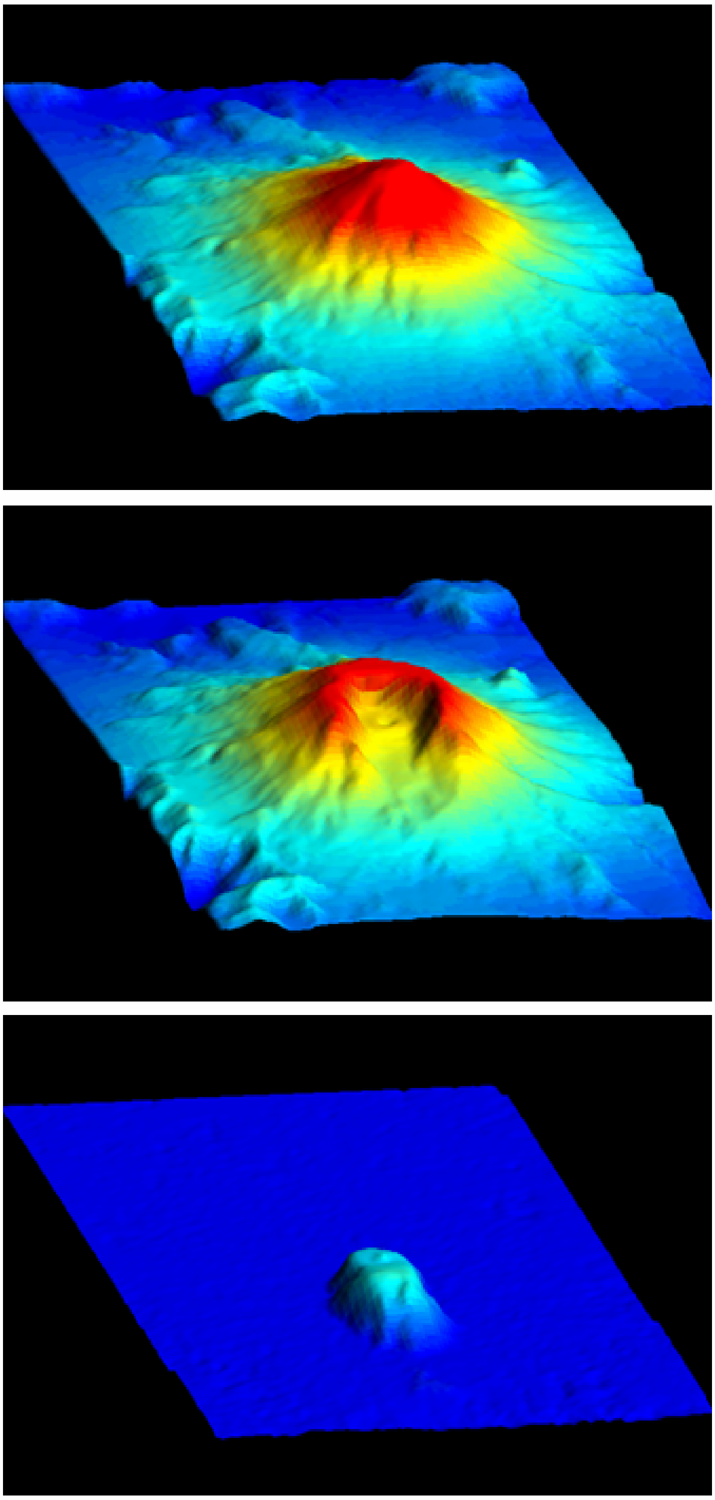
Plate 8 Vector layers showing rock units and drainage. Diagram from Whiteside, 2002.



Known Jade Locality

Units labelled A-E show possible jade localities. Eastern tip of Unit E is believed to be the most likely jade bearing parent lithology for the detrital jade.

Plate 9 Finished GIS showing Likely Host Units (Moore et al. 2000).

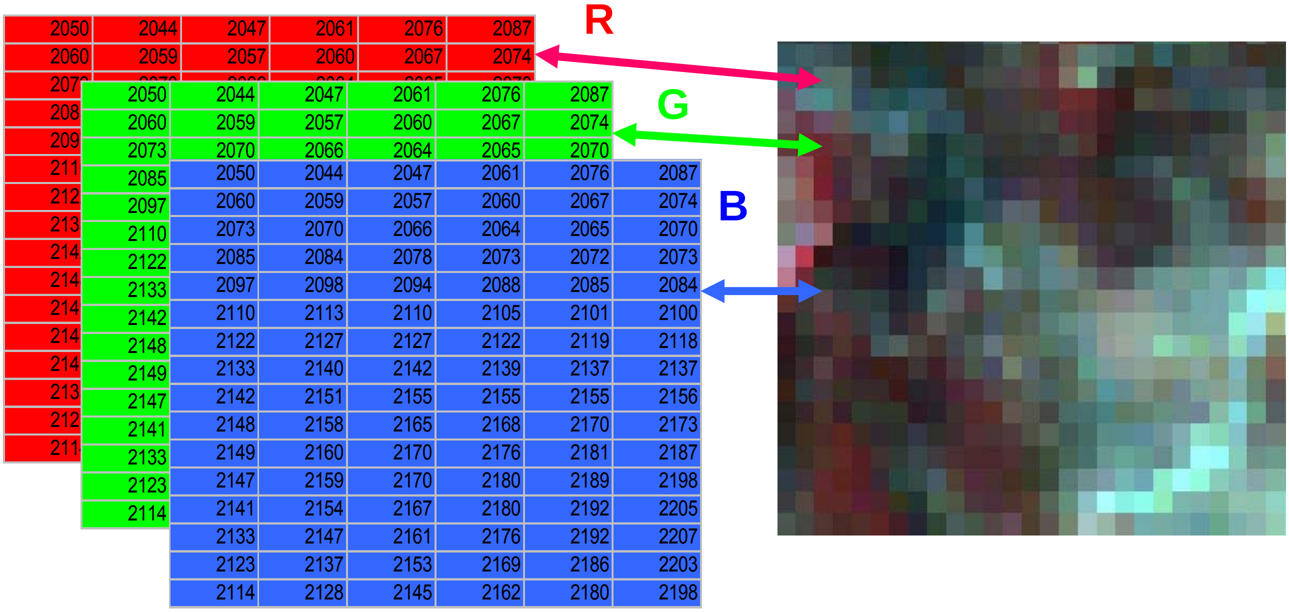


In the first image we can see a 3D reconstruction of Mt St Helens before the May 1980 eruption. Note that in this image there has been no image overlay, that is all the colours and textures in the image are derived purely from height data. The colours represent height and north is to the bottom of the image.

In the second image we can see Mt St Helens after the blast. The northern flank of the volcano has collapsed removing large quantities of rock.

The amount of rock removed in the eruption can be very easily viewed in image processing software by using simple arithmetic processes (see Chapter 8) to subtract one volume from the other.

Plate 10 ER Mapper example data allows a 3D analysis of the Mt St Helens disaster of 1980.



2050	2044	2047	2061	2076	2087
2060	2059	2057	2060	2067	2074
2073	2070	2066	2064	2065	2070
2085	2084	2078	2073	2072	2073
2097	2098	2094	2088	2085	2084
2110	2113	2110	2105	2101	2100
2122	2127	2127	2122	2119	2118
2133	2140	2142	2139	2137	2137
2142	2151	2155	2155	2155	2156
2148	2158	2165	2168	2170	2173
2149	2160	2170	2176	2181	2187
2147	2159	2170	2180	2189	2198
2141	2154	2167	2180	2192	2205
2133	2147	2161	2176	2192	2207
2123	2137	2153	2169	2186	2203
2114	2128	2145	2162	2180	2198

Plate 11 A colour digital image (right), produced by allocating three bands of an image to the red (R), green (G) and blue (B) colour-guns of a computer screen.

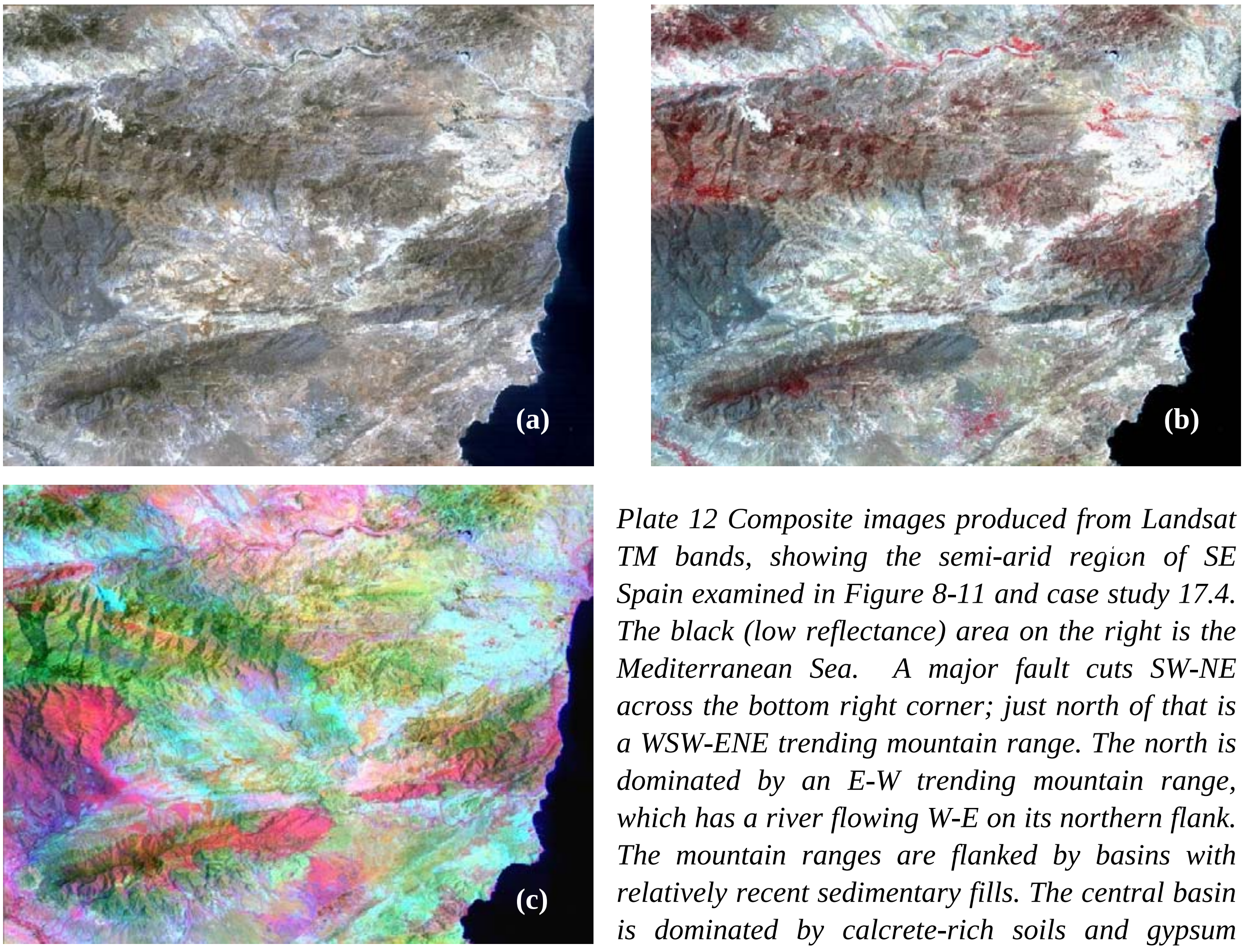
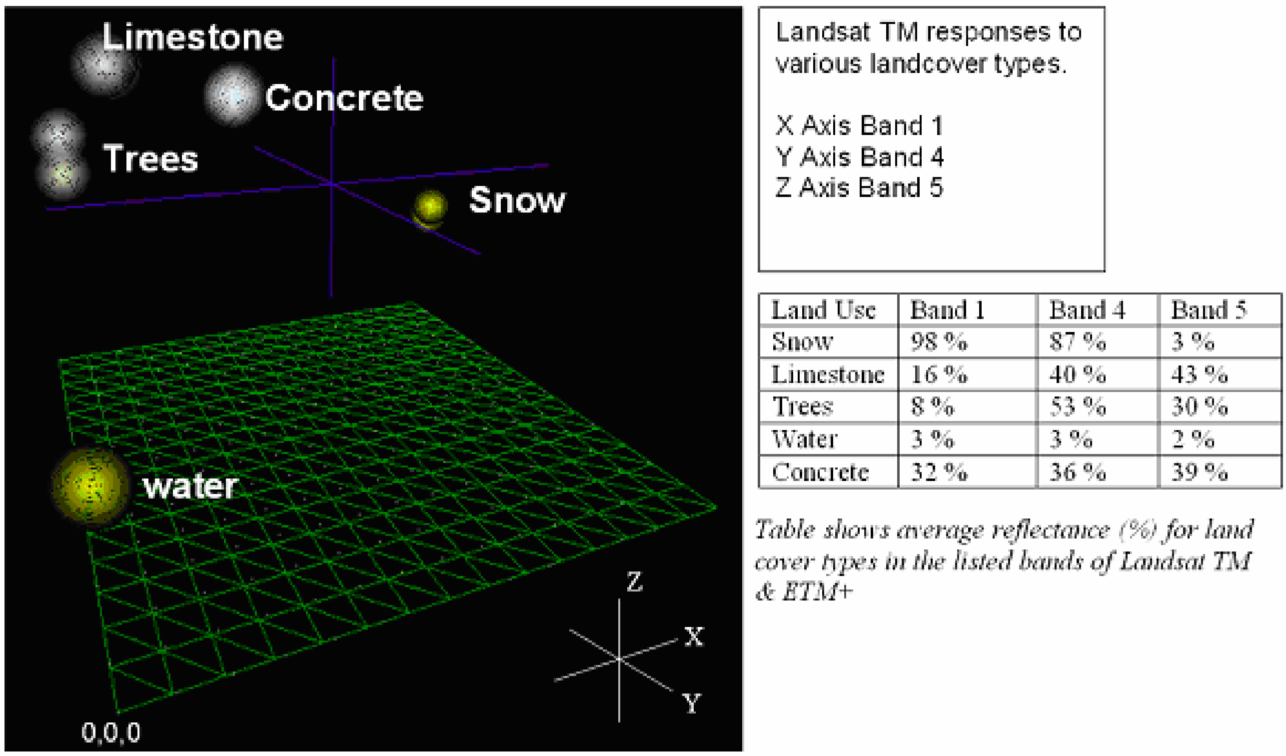


Plate 12 Composite images produced from Landsat TM bands, showing the semi-arid region of SE Spain examined in Figure 8-11 and case study 17.4. The black (low reflectance) area on the right is the Mediterranean Sea. A major fault cuts SW-NE across the bottom right corner; just north of that is a WSW-ENE trending mountain range. The north is dominated by an E-W trending mountain range, which has a river flowing W-E on its northern flank. The mountain ranges are flanked by basins with relatively recent sedimentary fills. The central basin is dominated by calcrete-rich soils and gypsum deposits, hence its dominant white colour (high reflectance); the western side of the central basin has been severely eroded, creating 'badland' terrain dominated by dark-coloured debris from adjacent hills. (a) 'True colour' (RGB321): the white areas are bare ground, mostly limestone, plus quarries of gypsum or marble; (b) 'False-colour infra-red': vegetation shows as red (irrigated agriculture or dense tree cover on the wetter mountains) or pink (sparse 'maquis' shrub cover); (c) RGB651: band 6, thermal infra-red, has been allocated the red colour gun – dark rocks and debris with a high thermal inertia show as shades of red, the hotter they are the stronger the red.



Landsat TM responses to various landcover types.

X Axis Band 1
Y Axis Band 4
Z Axis Band 5

Land Use	Band 1	Band 4	Band 5
Snow	98 %	87 %	3 %
Limestone	16 %	40 %	43 %
Trees	8 %	53 %	30 %
Water	3 %	3 %	2 %
Concrete	32 %	36 %	39 %

Table shows average reflectance (%) for land cover types in the listed bands of Landsat TM & ETM+

Plate 13 Schematic representation of trivariate classification, showing plots of reflectance values from Band 1 (blue), Band 4 (near infra-red) and Band 5 (mid infra-red) of Landsat Thematic Mapper. The various land cover types are now clearly distinguished.

Plate 14 A home experiment to recreate the perspective distortion of an air photo.

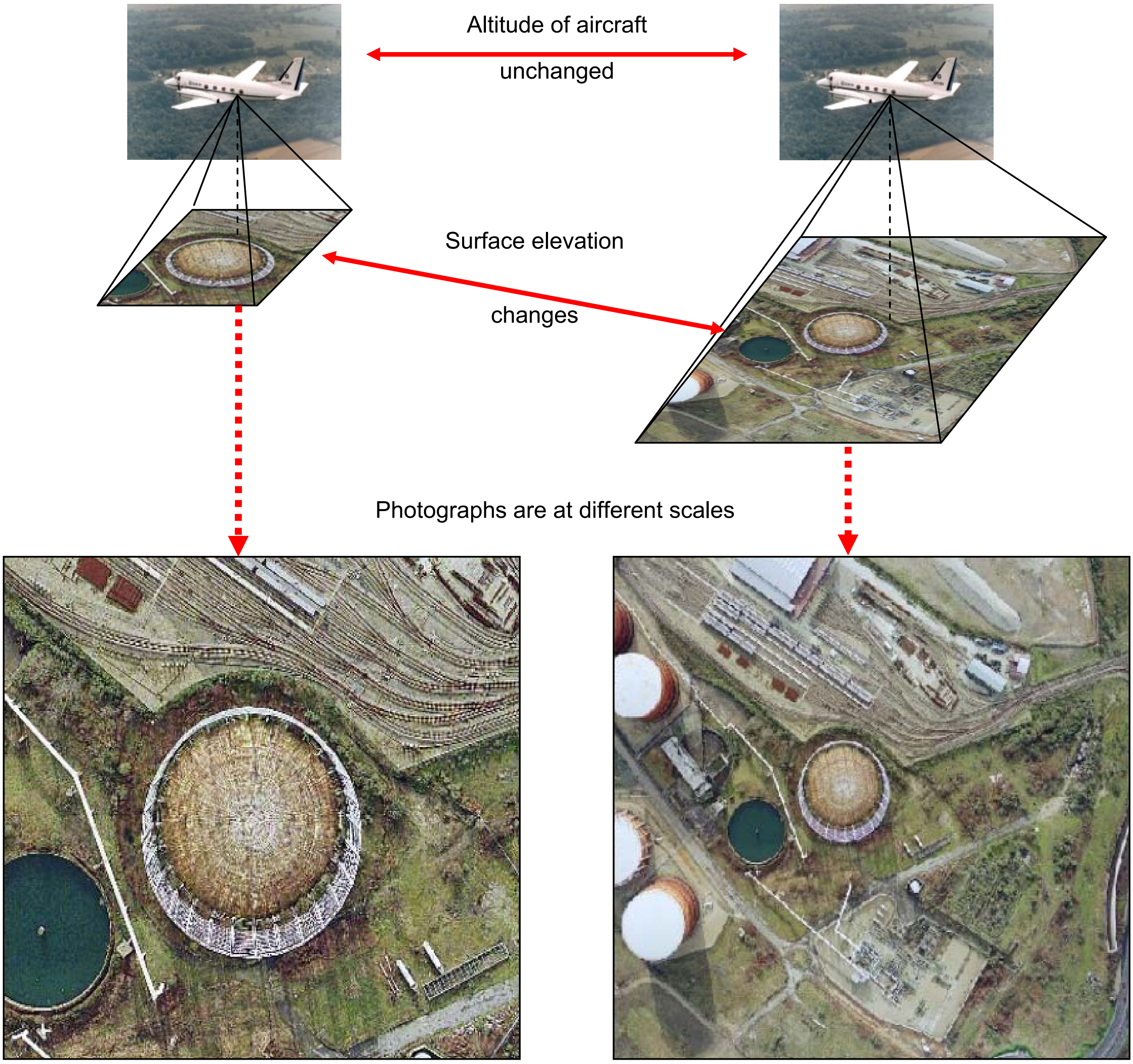


Plate 15 Illustration of the effect of terrain elevation on air photo scale.

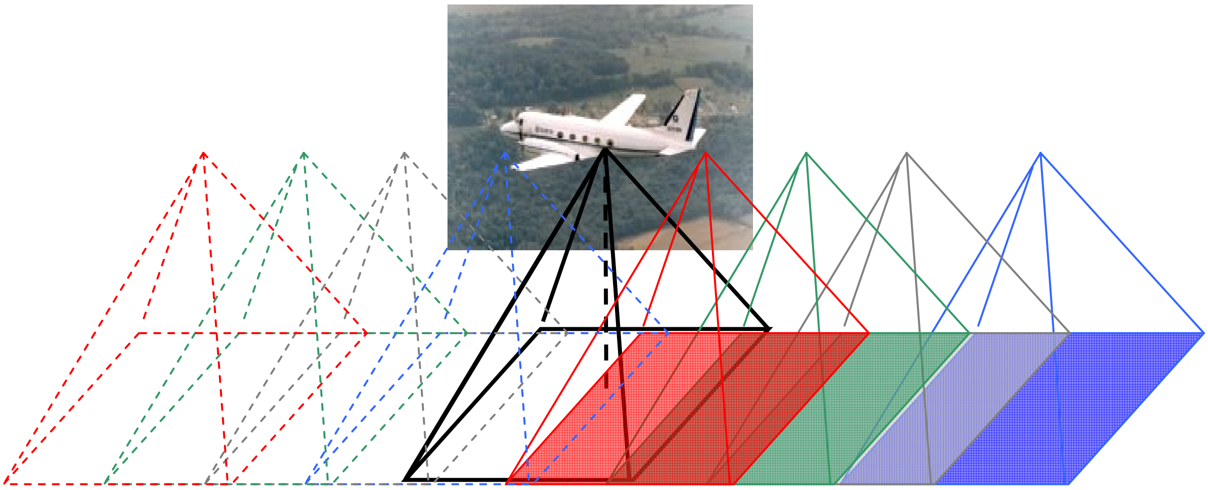
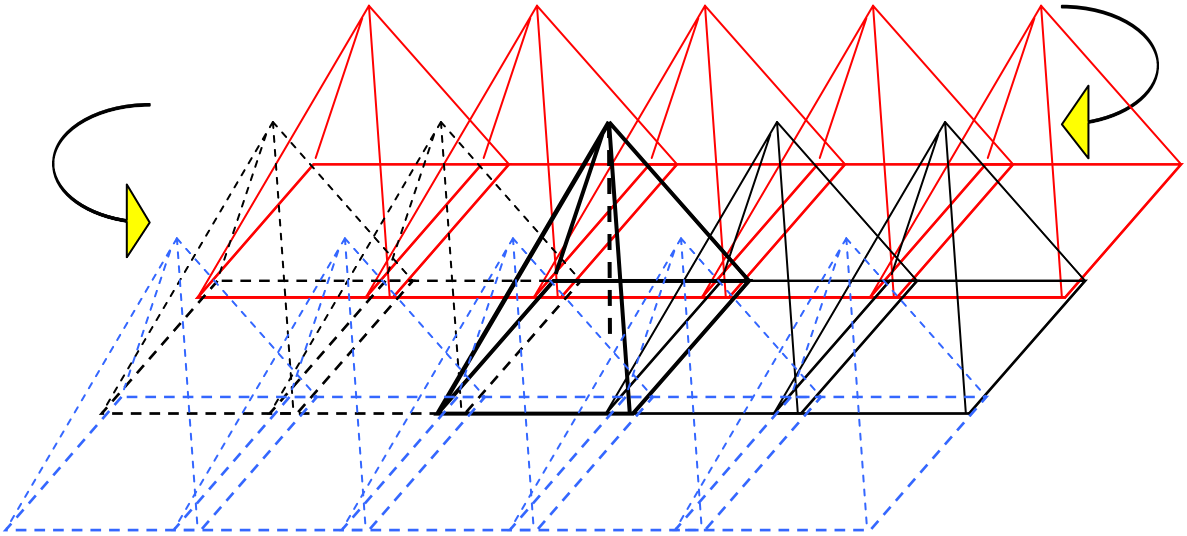


Plate 16

(Upper) Airphoto survey flight run showing coverage of each photo taken when acquiring coverage with 60% overlap between photos (solid lines show photos taken, dashed lines photos to be taken).



(Lower) Airphoto block of three runs showing coverage of each run and sidelap between runs. Solid black lines show photos taken, dashed lines show photos to be taken. Arrows show direction of flight (every other photo removed for clarity).

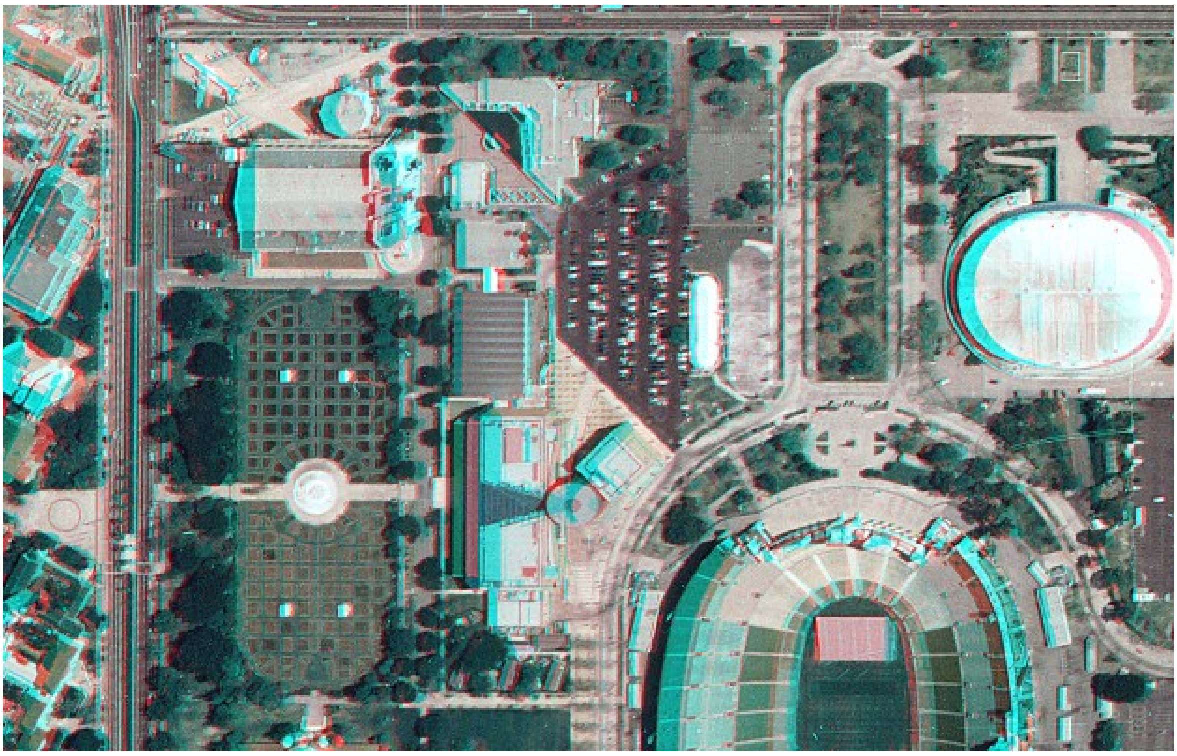
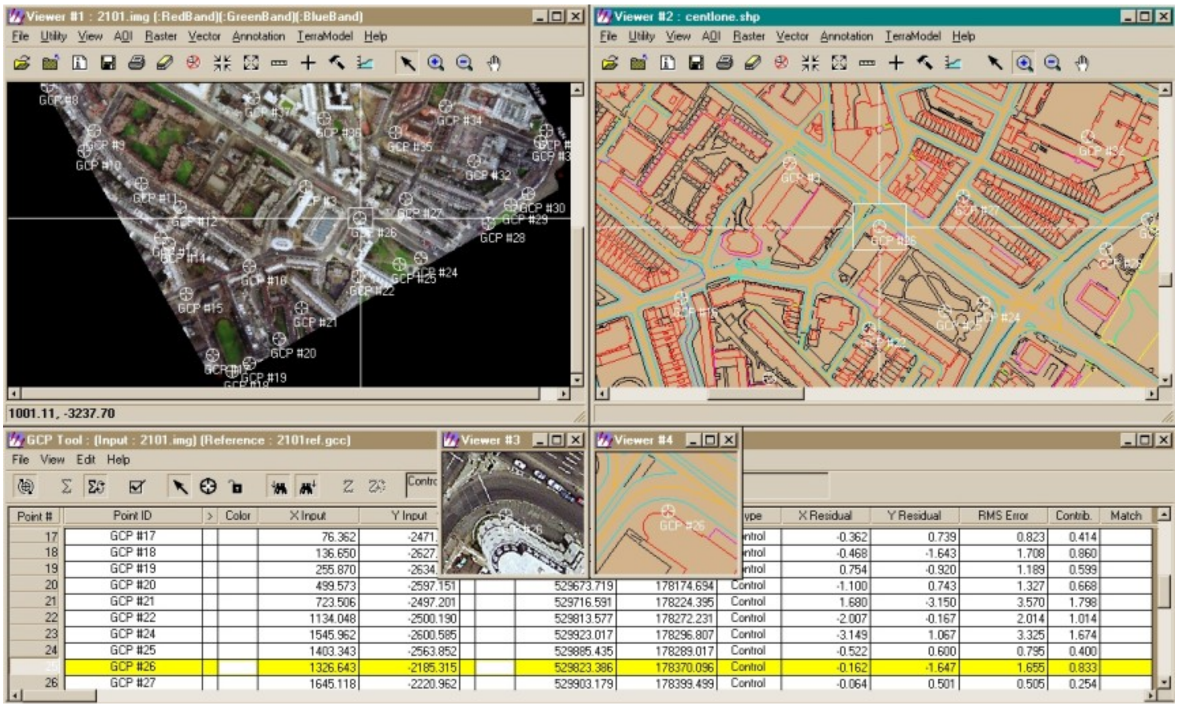


Plate 17 Anaglyph image showing two perspectives of an American football stadium and surrounding areas. Blue and red show the perspective differences between two co-located aerial photos. When viewed through blue/red glasses, the effect is a 3-D surface.



Point #	Point ID	>	Color	X Input	Y Input					ype	X Residual	Y Residual	RMS Error	Contrib.	Match
17	GCP #17			76.362	-2471.					ntrol	-0.362	0.739	0.823	0.414	
18	GCP #18			136.650	-2627.					ntrol	-0.468	-1.643	1.708	0.860	
19	GCP #19			255.870	-2634.					ntrol	0.754	-0.920	1.189	0.599	
20	GCP #20			499.573	-2597.151			529673.719	178174.694	Control	-1.100	0.743	1.327	0.668	
21	GCP #21			723.506	-2497.201			529716.591	178224.395	Control	1.680	-3.150	3.570	1.798	
22	GCP #22			1134.048	-2500.190			529813.577	178272.231	Control	-2.007	-0.167	2.014	1.014	
23	GCP #24			1545.962	-2600.585			529923.017	178296.807	Control	-3.149	1.067	3.325	1.674	
24	GCP #25			1403.343	-2563.852			529885.435	178289.017	Control	-0.522	0.600	0.795	0.400	
25	GCP #26			1326.643	-2185.315			529823.386	178370.096	Control	-0.162	-1.647	1.655	0.833	
26	GCP #27			1645.118	-2220.962			529903.179	178399.499	Control	-0.064	0.501	0.505	0.254	

Plate 18 Creating Ground Control Points (GCPs) for an aerial photo from digital maps of central London. Note the columns for file co-ordinates and map co-ordinates.

Plate 19 Compass traverse and slope inclinometer survey in dense jungle, Guyana. Note the problem of inter-visibility: photo (b) is looking back along the survey line towards the surveyor in the white T-shirt (circled) of photo (a).

Plate 20 Integration of field notes, PC data and GPS points can be used to rectify imagery and create reliable and accurate expedition maps (photograph by Daniel Hourigan, 2001). Rectification points are best taken on easily identifiable points such as tight bends or junctions in roads.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- <u>Section D: Planning & Practicalities</u>

Chapter 12: Project Planning and Management

12 Project Planning and Management

Unless you have a sound way of utilising GISci tools and techniques during fieldwork, a lot of time and effort could be wasted. The remaining chapters therefore focus on the practical aspects of GISci usage on expeditions. For general information on expedition planning, the reader should refer to the Expedition Advisory Centre's *Expedition Handbook* (Winser 2004). A brief set of expedition planning guidelines was published in a special supplement of the November 2003 *Geographical Magazine,* while Nash (2000a, 2000b) has produced a useful set of general guidelines on carrying out overseas fieldwork projects. Most of the field technique examples given in this handbook are from geoscience, ecology or conservation projects – such are the backgrounds of the authors – however, if you want to carry out anthropological, socio-economic or cultural studies, the Expedition Advisory Centre has published a very useful booklet on people-orientated research (Kapila & Lyon 1994). Advice on planning terrain surveys is given by Mitchell (1991); while Clarke (1986) produced a useful handbook for those planning an ecological monitoring project. Clifford & Valentine's (2003) textbook on *Key methods in geography* contains general advice on fieldwork techniques, sampling strategies and data analysis, with a particularly useful chapter by Professor Ian Reid, on making field observations and measurements.

The origins of 'planning' go back to the Latin word *planus*: a plain, an area where all is visible. When you plan you expedition, you should strive to have a clear vision of what you will be studying and how you will carry out those studies. Professor Robert Allison (2003) states that there are four key ingredients in making an expedition fieldwork project as practicable, useful and focussed as possible:

- Well-developed questions or hypotheses
- Realistic goals within the given constraints and circumstances (time, people, skills, equipment, climate, logistics, money)
- Neither too much nor too little work
- Results that relate to a wider research framework and/or have useful application

Try going through these four points, relating each one to your own ideas. You should regard Geographical Information Science as a tool, rather than an end in itself. For example, if the aim is to produce a vegetation map, a GIS may be used to store field observations of vegetation structure, relate them to satellite images, and produce a map of different vegetation classes. Furthermore, ask yourself if you are choosing 'appropriate' technology, in terms of cost, time, technical skills, maintenance, access to technology, and most important, in relation to local agendas for research, monitoring, training and education (see Dunn *et al.* 1997).

12.1 Project design

There are a number of distinct stages that most fieldwork-based projects follow, regardless of whether they are using GISci technologies. These are listed below in the order in which they would be carried out and with an indication of the time required for each stage of the project.

At least one year before departing:

1) Desk study 1: deciding on the scope of the study
2) Deciding the project aim and objectives
3) Desk study 2: getting as much information as possible on the project area
4) Selecting appropriate fieldwork techniques, possible modifications to the project aim and objectives
5) Site selection: seek advice from contacts in host country, use existing maps or remotely sensed imagery to find optimal survey sites

About six months before departing:

1) Preparation of work schedule, obtaining necessary permits, insurance and medical cover

In the host country (allow a minimum of one month):

1) In the study area: recce visit to selected sites
2) Detailed fieldwork at optimal sites, possible field laboratory analyses
3) On-site data processing: allows you to check and correct erroneous field data
4) Presentation of preliminary results: a summary will probably be appreciated by relevant researchers and managers in the host country

Back home:

1) Further data processing: here the full power of GIS analyses can be very useful
2) Presentation of results: try to complete your final report within one year of returning, otherwise it may never get finished!

12.1.1 Aim and objectives

"…a principal aim of most geographical research is to make useful generalisations - that is, to seek out and explain patterns, relations, and fluxes that might help to model, predict, postdict or otherwise understand better the human and physical worlds around us." Rice (2003)

A crucial initial stage is deciding what the overall goal or aim of the project will be. Once the aim of the project is decided, a set of objectives can be drawn-up which will need to be met if you are to achieve that aim. Some projects will have relatively simple sample-and-record goals, such as compiling inventories of species for a conservation area. On inventory projects, using GPS receivers as data-loggers is a very effective sampling technique, especially if GIS can be used to examine the distribution of mapped features. Other projects might be driven by the need to find solutions to geo-ecological and/or socio-economic problems affecting a study area: this often involves many factors and complex systems of cause and effect. When examining such complex situations the best approach is to develop concepts or hypotheses that can be used to test a set of research questions that seek to identify the causes and effects affecting a given site. Most field sites are multifactoral – such is the inherent complexity of Nature – and hypothesis testing will therefore involve statistical statements of probability about targeted causes and effects. GIS can be a very useful tool at this stage of the project, as it allows the visual comparison of existing map-based datasets (including remotely sensed imagery), assisting with the targeting of key factors for testing.

When formulating your project aim, objectives and activities (or 'deliverables'), there are some management techniques that can help to try to clarify your thoughts. A comprehensive approach to this is the Logical Framework: this is used by the UK Darwin Initiative for Biodiversity Conservation when assessing projects bidding for funding. The logical structure that links the project activities takes this form: IF <activities> AND <assumptions> THEN <outputs>; IF <outputs> AND <assumptions> THEN <purpose>, and so on. It can get a tad tedious working through these logical steps, but the plus side is that you are forced to review the practical aspects of each step of your project plan and may find inconsistencies or invalid assumptions that would otherwise have caused severe problems later. The example given in Table 12-1 is from a project that aimed to map areas of enhanced habitat loss on the coastal plain of the Republic of Georgia, with an associated programme of GIS training for Georgian scientists to continue the sampling after the UK team returned home.

Table 12-1 Example of a completed logical framework proforma.

Project summary	**Measurable indicators**	**Means of verification**	**Important assumptions**
Goal To assist countries rich in biodiversity, but poor in resources, with the conservation of biological diversity & implementation of the Biodiversity Convention.	Provision of training materials for GIS mapping of land cover types, habitat types, habitat loss and target areas for biodiversity conservation. Case study of Georgia coastal plain.	Production of GIS training materials (in English, Georgian and Russian). Publication of the cited database and maps; access to the information via a dedicated website.	Will the products be understandable ? *- yes: text will be in English, Georgian and Russian.* Does the region have adequate provision for website access ? *- yes.*
Purpose To provide a database and 1:60,000 scale maps of: (i) land cover types, (ii) habitat types, (iii) habitat loss (1991-2001), (iv) biodiversity target areas.	The cited database and three sets of maps, will be available in Georgia by the end of December 2002.	i) Publication of the Project Report, database and cited maps; posting of findings on the Project website.	Is there a need for the cited database and cited maps in Georgia ? - *yes, checked and confirmed by our Georgian partners.*
Outputs 1) GIS-generated database and 1:60,000 scale maps of Georgia coastal lowlands for: a) Land Cover Types; b) Habitat Types; c) Habitat Loss, 1991-2001; d) Biodiversity Target Areas. 2) Project Report. 3) A Project website,	Land cover types, habitat types, habitat loss and targeted areas for bio-diversity conservation, mapped by July 2002. 'Ground-truth' of GIS maps checked by end Sept 2002.	Weekly email discussions, with monthly progress reports posted on the Project website. Fieldwork in Georgia to check groundtruth of the GIS maps. Publish a Report; establish a Project website. Disseminate findings in Georgia and Black Sea countries; number of visits	Is access to lowland Georgian fieldsites straight-forward ? - *yes.*

allowing public access to database, and digital maps. 4) Press releases, TV and radio cover, Georgia & UK. 5) Methodology and results published in peer-reviewed journals. 6) Batumi Biodiversity Conservation Workshop.	Report and website completed by end Dec 2002. Media coverage over Dec 2002, with scientific publications submitted by March 2003. Biodiversity Conservation Workshop March 2003, Batumi, Georgia.	to Website; distribute Report to libraries and govt agencies; number of radio, TV & newspaper features. Two scientific publications by mid-2002. List of delegates attending the Biodiversity Conservation Workshop and the GIS Training course.	
Activities i) Process Landsat data; GIS maps of Land Cover Types, Habitat Loss & Biodiversity Target Areas. ii) 1-week UK visit by Georgian partners: train to edit digital maps; review methodology and outputs. iii) Fieldwork by Georgians, checking map accuracies. iv) Up-dating of the data-base and maps; write-up Report. v) Place Report and maps on website*. vi) Press releases, radio & TV features: Georgia & UK. vii) Black Sea Workshop on Biological Conservation / GIS training course. viii) Submit scientific papers.	Land cover types, habitat types, habitat loss and biodiversity target areas mapped using Landsat images & GIS by July 2001. 1-week UK training course for Georgian ‘ground-truth’ surveyors, July 2001. ‘Ground-truth’ fieldwork & editing of maps over August – September 2002. Report written, translated, placed on website and publicised by end-Dec 2002.Two scientific papers submitted by March 2003.	Public access to the processed Landsat images and provisional GIS-generated maps, via the Project website. Informal assessment of skills, before and after the training. Posting of the Georgia ‘ground truth ‘ field data on the Project website. Publication of Project Report; posting of findings on Project website. Number of peer reviewed papers, international conference presentations: UK/EU and Georgia/Russia.	Is Landsat cover available for the relevant areas and required periods ? - *yes.* Can the Georgia team travel to the UK then ? - *yes.* Do the Georgia team have basic GIS skills and fieldwork survey skills ? *- training to be provided during their week in the UK.* * = Translate into Georgian and Russian

Another approach to project planning is SWOT analysis: expedition participants and its advisors are asked to list the **S**trengths, **W**eaknesses, **O**pportunities and **T**hreats associated with the proposed project, or a particular aspect of that project. This technique allows all concerned to have their viewpoints considered and it may well highlight adverse features that were missed in the original project formulation. When carrying out a SWOT analysis of your project, try to answer the questions in Table 12-2, from your own perspective and that of others. Bear in mind that new Opportunities may be opened (a) by some of your project’s strengths and (b) by eliminating some of its weaknesses.

Table 12-2 A framework for SWOT analysis.

Strengths: • what are your skills? • what resources do you have?	**Weaknesses**: • what do you do badly? • what could you improve?
Opportunities: • what are the mapping and analytical opportunities facing the project? • how can you utilise your GISci capabilities?	**Threats**: • what obstacles does the project face? • can any of your weaknesses seriously affect the project?

12.1.2 The GISci KISS Principle

During your expedition planning, try to follow the KISS Principle ("*Keep It Simple & Scientific*"). There is no doubt that remote sensing, image processing and GIS-based map production can be of great value to planners and natural resource managers in developing countries. The transfer of GISci technologies is certainly appropriate, but it must also be affordable. If your expedition is using GISci techniques in association with partners in a developing country, then the use of low-cost software – supported by user-friendly training materials – will improve the long-term sustainability of the project. Ideally, workers in the host country should be able to continue to collect samples, analyse data and produce summary maps, after you have returned to the UK. In this respect, expeditionary fieldwork is ideal for providing sustainable GISci-based projects in developing countries, especially if the software and hardware are extensively used during fieldwork, as most expeditions simply cannot afford expensive kit. A modified version of the KISS Principle, "*Keep I.T. Simple & Sustainable*", operates with GISci on most expeditions because they cannot afford to use the sophisticated, top-of-the-range software and hardware. In contrast, some projects initiated in developing countries by major governmental aid agencies have opted for the most sophisticated software available, requiring the powerful computers and highly-trained support staff. The results are all too predictable: after the initial training period there are often problems with computer maintenance; furthermore, the highly-trained staff look for better-paid work overseas, leaving behind some very expensive, sophisticated, but useless computer kit.

Landsat or ASTER imagery provides a low cost means of both navigating and mapping study areas. In Plate 21, from a geological mapping project in Burkina Faso, a geocorrected Landsat TM image has been used in conjunction with GPS to guide the expedition Land Rovers around areas of active sand dunes and river valleys prone to flooding at the onset of the rainy season.

12.2 Sampling strategies

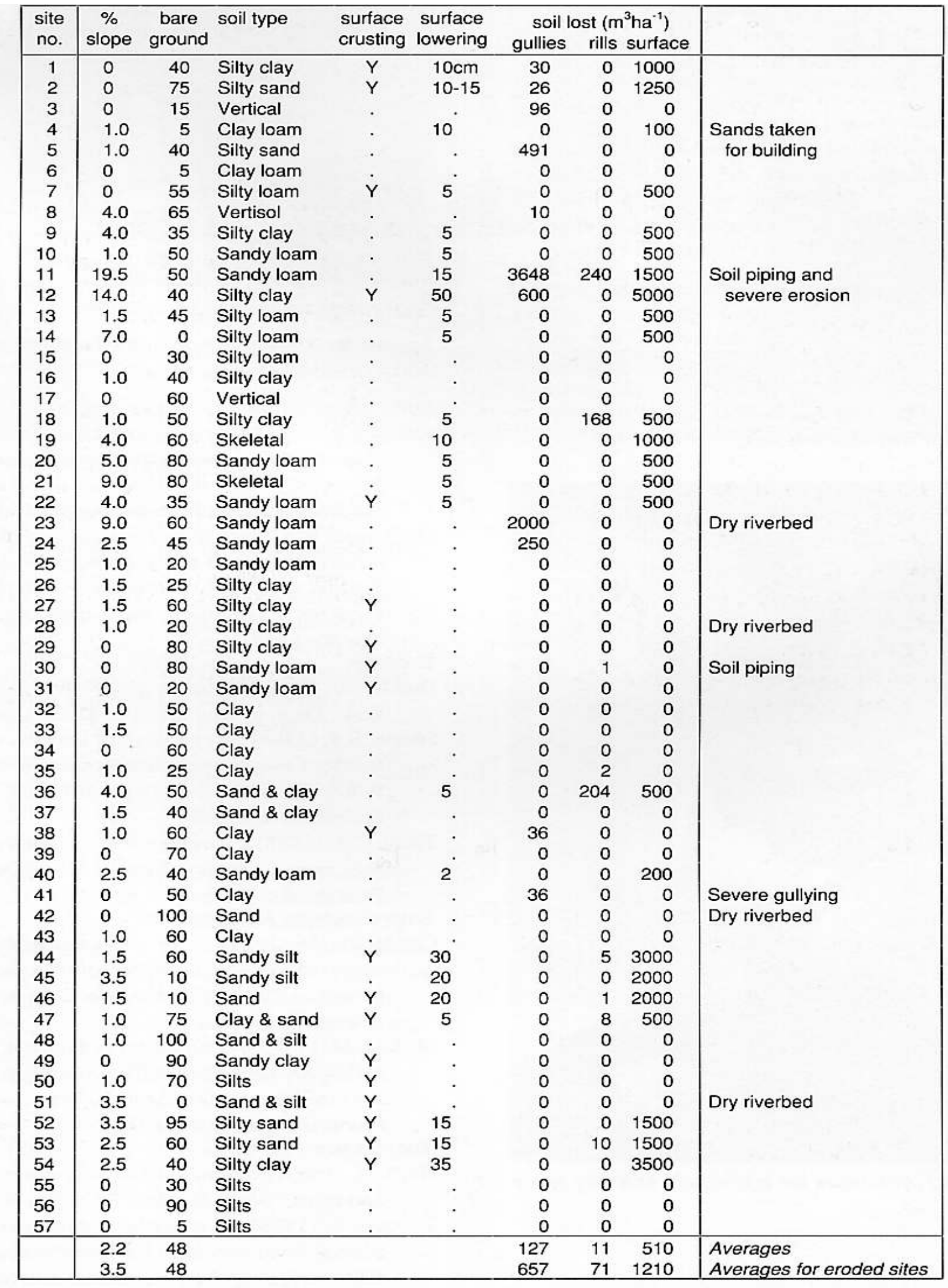
"*Making field observations is not difficult. Making appropriate field observations is a challenge*." Reid (2003)

There are basically two approaches to sampling on fieldwork projects: generalist or specialist. The specialist, or 'case study', approach will examine a few sites in great detail, studying the mechanisms that form general pattern: for example, the trapping of rainforest butterflies at set locations and set times of the day – dawn, noon and dusk – over periods of many weeks. Another example is illustrated in *Figure 12-1* a very detailed study of gully erosion (see Faulkner *et al*. 2004 for more details). Problems with the specialist approach are that (i) the few selected sites may be inappropriately located and (ii) the sampling intervals might miss the impacts of key events, such as thunderstorms or seasonal climatic changes. Expeditions to the tropics and sub-tropics should bear the latter in mind: for most expeditions the easiest and most comfortable time to carry out fieldwork is the dry season; unfortunately in many cases most of the geomorphological activity occurs during the wet season and there are correspondingly great seasonal variations in tropical ecosystems.

Figure 12-1 Detailed surveying and sampling of a gully erosion site near Almeria, Spain: an example of the 'one site / many samples' approach (courtesy of Hazel Faulkner).

The generalist approach involves sampling many locations a few times, allowing the mapping of general patterns. Most expeditions are faced with problems of limited time and money, with study areas in remote and poorly-mapped regions, so the generalist approach to fieldwork is often the most appropriate. You could spend most of your time and budget examining in great detail the erosion processes along a single gully, when a more useful task would be to produce a geomorphological map for the whole region. An effective way of rapidly collecting a lot of samples from a study region is to select roads or motorable tracks that traverse the main geo-ecological units of the area and then to use the odometer ('mileometer') of your vehicle to stop every kilometre along those routes to record features and/or collect samples. This can be done very simply using compass bearings alongside the kilometre readings (though keep the compass well away from the vehicle when you take a reading!). If you use a GPS, this type of 'many sites/few samples' survey becomes even more versatile: the vehicle odometer is not needed, as most GPS sets can be set to beep every kilometre along a given bearing – allowing surveys on foot (or on horse-back, or on yak-back…). Furthermore, the GPS unit can double-up as a data-logger that can input your field records directly into a GIS – though remember to take back-up paper records too. An example of a regional soil erosion survey in Zambia's Luangwa Valley, carried out using a Land Rover to traverse almost 60 km of tracks with samples collected every kilometre, is shown in Table 12-3.

Table 12-3 Results from a regional soil erosion survey, Luangwa Valley, Zambia: an example of the 'many sites / few samples' approach (Teeuw 1990).



site no.	% slope	bare ground	soil type	surface crusting	surface lowering	soil lost (m^3ha^{-1}) gullies	rills	surface	
1	0	40	Silty clay	Y	10cm	30	0	1000	
2	0	75	Silty sand	Y	10-15	26	0	1250	
3	0	15	Vertical	.	.	96	0	0	
4	1.0	5	Clay loam	.	10	0	0	100	Sands taken for building
5	1.0	40	Silty sand	.	.	491	0	0	
6	0	5	Clay loam	.	.	0	0	0	
7	0	55	Silty loam	Y	5	0	0	500	
8	4.0	65	Vertisol	.	.	10	0	0	
9	4.0	35	Silty clay	.	5	0	0	500	
10	1.0	50	Sandy loam	.	5	0	0	500	
11	19.5	50	Sandy loam	.	15	3648	240	1500	Soil piping and severe erosion
12	14.0	40	Silty clay	Y	50	600	0	5000	
13	1.5	45	Silty loam	.	5	0	0	500	
14	7.0	0	Silty loam	.	5	0	0	500	
15	0	30	Silty loam	.	.	0	0	0	
16	1.0	40	Silty clay	.	.	0	0	0	
17	0	60	Vertical	.	.	0	0	0	
18	1.0	50	Silty clay	.	5	0	168	500	
19	4.0	60	Skeletal	.	10	0	0	1000	
20	5.0	80	Sandy loam	.	5	0	0	500	
21	9.0	80	Skeletal	.	5	0	0	500	
22	4.0	35	Sandy loam	Y	5	0	0	500	
23	9.0	60	Sandy loam	.	.	2000	0	0	Dry riverbed
24	2.5	45	Sandy loam	.	.	250	0	0	
25	1.0	20	Sandy loam	.	.	0	0	0	
26	1.5	25	Silty clay	.	.	0	0	0	
27	1.5	60	Silty clay	Y	.	0	0	0	
28	1.0	20	Silty clay	.	.	0	0	0	Dry riverbed
29	0	80	Silty clay	Y	.	0	0	0	
30	0	80	Sandy loam	Y	.	0	1	0	Soil piping
31	0	20	Sandy loam	Y	.	0	0	0	
32	1.0	50	Clay	.	.	0	0	0	
33	1.5	50	Clay	.	.	0	0	0	
34	0	60	Clay	.	.	0	0	0	
35	1.0	25	Clay	.	.	0	2	0	
36	4.0	50	Sand & clay	.	5	0	204	500	
37	1.5	40	Sand & clay	.	.	0	0	0	
38	1.0	60	Clay	Y	.	36	0	0	
39	0	70	Clay	.	.	0	0	0	
40	2.5	40	Sandy loam	.	2	0	0	200	
41	0	50	Clay	.	.	36	0	0	Severe gullying
42	0	100	Sand	.	.	0	0	0	Dry riverbed
43	1.0	60	Clay	.	.	0	0	0	
44	1.5	60	Sandy silt	Y	30	0	5	3000	
45	3.5	10	Sandy silt	.	20	0	0	2000	
46	1.5	10	Sand	Y	20	0	1	2000	
47	1.0	75	Clay & sand	Y	5	0	8	500	
48	1.0	100	Sand & silt	.	.	0	0	0	
49	0	90	Sandy clay	Y	.	0	0	0	
50	1.0	70	Silts	Y	.	0	0	0	
51	3.5	0	Sand & silt	Y	.	0	0	0	Dry riverbed
52	3.5	95	Silty sand	Y	15	0	0	1500	
53	2.5	60	Silty sand	Y	15	0	10	1500	
54	2.5	40	Silty clay	Y	35	0	0	3500	
55	0	30	Silts	.	.	0	0	0	
56	0	90	Silts	.	.	0	0	0	
57	0	5	Silts	.	.	0	0	0	
	2.2	48				127	11	510	*Averages*
	3.5	48				657	71	1210	*Averages for eroded sites*

Another example of a generalist fieldwork methodology is the Land Systems technique of geo-ecological mapping discussed in the *Image processing and interpretation* chapter (viz. Figure 8-5). This was originally devised in the 1950s to map land suitability for agriculture across the whole of Australia. Once regional geomorphological or geo-ecological maps have been produced, more specific studies can be attempted, such as detailed sampling of soil and slope properties, which can then be used in a GIS to generate soil erosion hazard maps. The Land Systems approach has been used extensively in the Luangwa Valley of Zambia as the basis for erosion hazard mapping and examples from the planning of some of those expeditions are given below. Mapping the geo-ecological characteristics using Land Systems is an example of *Nested or Stratified* sampling, moving from regional levels to more detailed scales of study. A disadvantage of this strategy is that it is dependant on image interpretation, a subjective process that may introduce bias into the dataset. *Systematic* sampling involves the collection of samples at regular intervals, such as the corners of grid squares on a 1:50,000 topographic map, or at 1 km intervals along a track: the main disadvantage of this method is that the sampling interval may introduce bias into the dataset. Adopting a *Random* sampling strategy will remove operator bias and sampling interval bias, but it may leave you with no samples from what are clearly important parts of a data population. An effective compromise, widely used in field surveys, is to use a *Stratified Random* sampling strategy: all of the major groupings in a data population are recognised and sampled, but the sampling of each data grouping is done randomly. One of the most rapid types of geo-ecological surveys is to follow a straight transect line from a hilltop to an adjacent valley bottom, measuring breaks of slope along the way and marking them with survey tape or blazes on trees, then returning back up the transect line, sampling soil and vegetation types in each of the similar-slope units.

Whether you opt to follow the specialist or generalist mode of sampling you will probably be faced with more potential samples than you have time to collect. Sampling methods have been devised that allow you to collect data from a relatively small part of a much larger data population, often with a means of inferring generalisations about the larger group. The basis of most of these sampling methods is probability and the assumption that most data populations follow a normal distribution (giving a bell-shaped curve when plotted as a graph) allowing the validation of inferences by statistics. See Rice (2003) for an easy to follow summary of geographical sampling and inferential statistics. In addition to statistical analyses, a GIS can carry out spatial and temporal analyses, allowing you to examine distribution patterns and changes over time.

12.2.1 Project management

Expeditions will vary in their complexity depending on the number of objectives that they are aiming to complete. Most will involve a number of often inter-linked fieldwork projects based in one study area, although another common variant involves sampling just one thing at many remote locations. Whatever the case, careful management of time, money, equipment and expertise is needed, if all of the project objectives are to be met. Reconnaissance ('recce') surveys are an essential aspect of successful field project management. A common practice is for one or two expedition members to make a one-week recce visit to the study region, so that access to study areas can be checked, permissions obtained and goodwill visits made to key associates in the host country. Plans can then be modified accordingly after the recce visit and the whole expedition then decamps to the host country for at least a month's fieldwork. On many expeditions costs will preclude a recce visit, however, this can be compensated by having expedition partners in the host country who can check the field conditions and associated logistical factors.

Project budgeting is not just limited to financial considerations. Most expeditions are faced with limited time and limited manpower, so careful consideration is needed of who does what and when, with which bit of equipment. With so many eventualities on any given expedition, the best approach to tackling these logistical problems is for you to go through the following check-list, considering how the questions might apply to your field project.

People

- Who will have the time & skills for GIS work? Who can help?
- Who will be involved? (students, rangers, wardens, villagers, politicians)
- Who will have training? When, how and where?

Time-tabling

- Before fieldwork (1/3): liaison, learning, research, datasourcing, digitising, testing
- During fieldwork (1/3): data entry, possible problems
- Afterwards (1/3): data cleaning, analysis, mapping, reporting, distribution

Equipment

- What hardware and software is needed for the project?
- Should you use existing kit or purchase new items? Check GIS software licenses for multiple usage conditions
- Can the GISci kit be integrated with existing facilities? e.g. university Departments, National Parks, Government agencies, maintenance, insurance, power supplies

Finance

- Have you budgeted thoroughly, e.g. for equipment, provisions, accommodation, transport, insurance?

Flow diagrams and *Gantt Charts* (Figure 12-2 and Table 12-4*)* are widely-used and useful techniques for ensuring the effective use of time, manpower and equipment during the life of the field project. Gantt charts usually have a horizontal time axis and a vertical list of project tasks: more complex variants can include the names of expedition members who have been allocated to each task. In Table 12-4 the shaded boxes show the time in the life of the project when a given task is to be carried out. Colour coding of time lines can be used to illustrated whether tasks are completed, in progress or overdue, and marker symbols can be used to highlight when key stages in the project have been completed.

Another technique to consider is 'C*ritical Path Analysis'* (CPA), a type of flow diagram that divides a work programme into its components and displays them in a chronological network, detailing the number of days needed for given tasks, using flag symbols to highlight critical events or activities along the time-line of the project. Mitchell (1991) discusses CPA and provides an example from a terrain analysis project. The advantage of

CPA is that it identifies critical tasks that have to be completed on time for the whole project to be completed on time, as well as those tasks that you could afford to delay if a resource needs to be reallocated to help speed-up a delayed critical task. The disadvantages of CPA are that it is most effective when prepared by someone with experience of expedition planning, who will best be able to judge how much time tasks will take – and CPA can be very complex! Useful websites to find out about and download these planning aids are: *www.smartdraw*.com and *www.mindtools.com*.

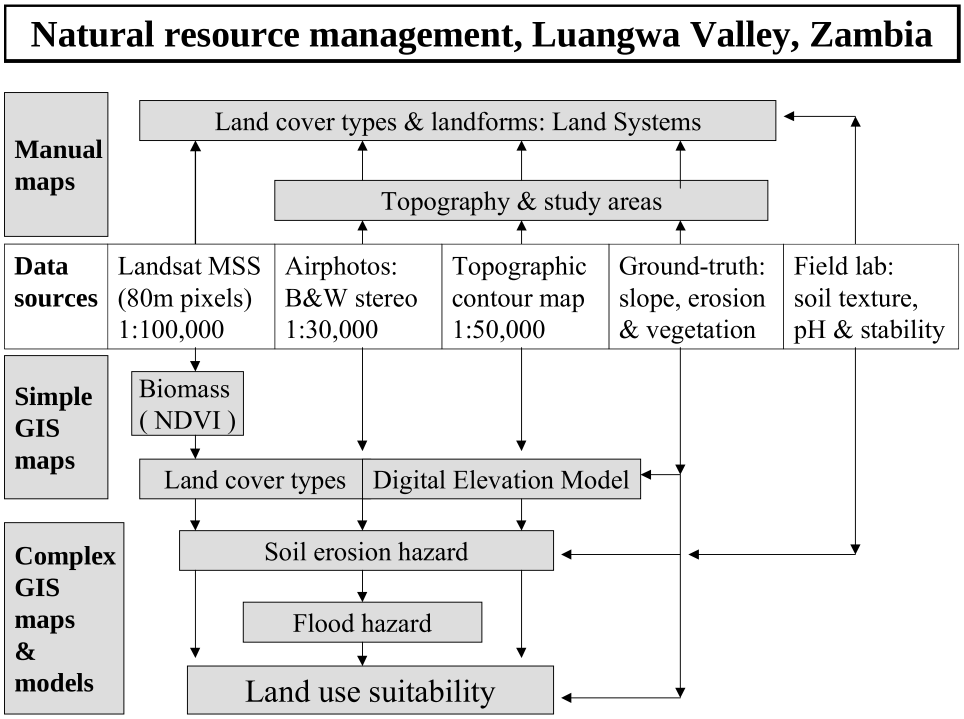


Figure 12-2 Flow diagram illustrating the scales of study, types of data and analyses carried out on a soil conservation expedition in Zambia.

Table 12-4 Gantt chart for a project examining urban water quality (original example courtesy of Hazel Faulkner).

Activity	**Year 1**			**Year 2**			**Year 3**		
Work plan & model calibration Desk study & field site recce visits Field site sediment granulometry Cultures of field-associated bacteria	▨								
Flood event data collection Site hydrology Field variations in bacteria		▨		▨		▨			
Hydraulic flow model development		▨	▨	▨	▨	▨			
Field and laboratory data analysis						▨	▨		
Model testing							▨	▨	
Write-up and dissemination of results								▨	▨

12.3 Data

12.3.1 Data requirements

Data requirements are determined primarily by the project's GIS aim. Having formulated your expedition GIS question or aim, what data will be required to answer the question or achieve the goal? There is something of a ‘chicken and egg situation’ with the project aim and the available data. Ideally the aim, or research question, will lead to the data *requirements*; but it also likely that data *availability* will to some extent determine the questions being asked. The availability of a particular dataset should not in itself lead your GIS operations. It is tempting, for example, to use a good cloud-free satellite image to produce attractive image maps, but make sure these are being produced with a use or user in mind. An example of the various types of data, scales of data and degrees of data analysis that were needed for a soil conservation project, is given in Table 12-5.

Table 12-5 Data sources used by the Luangwa Soil Conservation Project, Zambia.

Data sources	**Details: type of source data used**	**Manual maps produced**	**Simple GIS maps produced**	**Complex GIS maps of models**
1. Digital maps & databases	CIA: Zambia FAO: soils			
2. Satellite imagery	Landsat MSS	Geo-ecology (land systems)		
3. Aerial photography	1:30,000 panchromatic	Geo-ecology (land systems)	River bank erosion hazard	Land suitability map
4. Fieldwork: surveying	Compass / tape / inclinometer; surveyor’s level	Geo-ecology (land systems)	Erosion plot micro-relief	Change analysis: net erosion & deposition
5. Fieldwork: sampling	Soil pits; gully sizes; signs of surface erosion	Geo-ecology (land systems)	Soil erosion hazard map	Land suitability map
6. Laboratory: (a) Field (b) UK-based	Soil texture, pH, aggreg. Stability (b) % organic, % clay / silt / sand	Geo-ecology (land systems)	Soil erosion hazard map	Land suitability map

Key considerations when gathering data before the expedition concern the relevance, accessibility and cost of the datasets. A dataset may be available for your study region, but does it actually cover your study sites and is it an appropriate scale for your research objectives? The date when the data were collected should also be checked, as should the amount of cloud cover or haze if you are seeking remotely sensed imagery of your study sites. Some ideal datasets may exist in the host country, but you may not get access to them unless you have a partnership agreement with an organisation in that country - and even then you may experience bureaucratic delays of many months. Data costs have always been a major concern with scientific expeditions: the good news is that international agencies such as the United Nations Environment Programme (UNEP) the UN Food & Agriculture Organisation (FAO) and various major universities – notably the University of Maryland - have built-up numerous useful datasets that are open to free public access.

A major boost for overseas fieldwork, especially at local to regional scales of study, are the improved capabilities – but lower prices – of Earth observation satellite imagery. For instance, 60 km x 60 km scenes from the Japanese-American ASTER satellite give you multispectral coverage of just about anywhere on the planet with 15 m pixel detail, for just US$60 (UK £35). Details on data types and data sources are given in the GIS chapter near the start of this handbook and details of data suppliers are given in the Appendix.

There are a number of questions that can guide you in selecting the most appropriate satellite imagery for a fieldwork project:

- *Do you need data from as near a time as possible to your fieldwork period?* If so, the available imagery will probably be expensive and may well have significant amounts of cloud cover. If having very up-to-date imagery is not an issue, get free Landsat ETM+ imagery from the University of Maryland website.
- *Is the season of the imagery important (e.g. start or end of dry season, end of rainy season)?* Archives of satellite imagery spanning many years are available that should allow you to select imagery from suitable seasons.
- *Is the mapping of vegetation, soils and rock types a key aspect of your project?* For vegetation mapping, opt for imagery from the start of the dry season; for soils and rocks, select imagery from the end of the dry season. If dry season dust storms affect the area, opt for early rainy season (and pray that you can find a cloud-free image).
- *Do you need a wide range of spectral bands across the visible and infra-red (e.g. for mapping soils and rock types)?* If yes, opt for the relatively low cost ASTER imagery
- *Would a 3-D digital elevation model be of use to your project?* If so, ASTER can provide a DEM with 15 m pixels and 15 m contours, while the SRTM DEM (available free from the University of Maryland website) can provide 90 m pixels with 10 m contours.
- *Do you know of someone who has suitable data?* If so, start practicing your begging skills!

12.3.2 Scale and resolution

What scale or resolution is appropriate for the collection of data and the presentation of results? In choosing a suitable spatial resolution, a key criterion is the size of the spatial variations you are investigating, e.g. AVHRR data with 1 km pixels provide useful environmental information, but will not help if you are mapping habitat variations over a few hundred metres. Likewise, if you are scanning or digitizing airphotos or paper maps, choose a resolution appropriate to the task. Too much detail makes the data sets large, expensive and slow; but too little detail will not answer your questions. The same consideration applies to temporal and spectral resolutions. If you are using a series of satellite images to monitor seasonal changes in vegetation, they need to be frequent enough to represent significant temporal changes, yet not so frequent as to be impractical or too expensive. Particular care is needed when using remotely sensed images to map tidal regions, due to the daily variations in inundation extent and soil/sediment moisture content, both of which can have notable effects on radar and multispectral imagery.

Bear in mind that the optimum scales of remotely sensed data required for most geo-ecological surveys will change as the project progresses. During the desk study and recce visit stage you will probably be using satellite imagery at scales of 1:500,000 to 1: 25,000 (e.g., AVHRR or MODIS for regional mapping, with Landsat, ASTER or SPOT for mapping smaller areas of interest). Once areas for detailed study have been selected, more detailed imagery will be needed, typically at 1:25,000 to 1:10,000 scale, consisting of aerial photographs or 0.6 m-3 m pixel imagery from IKONOS or Quickbird. Finally, detailed surveys and mapping will be carried out, ideally aided by airphoto basemaps at scales of 1:10,000 to 1:3,000.

12.4 Equipment, software and fieldwork practicalities

Although guidelines on choosing appropriate expeditionary hardware and software are given in dedicated chapters, an overview of key points is given here, in the form of checklists that you can compare against your own project plans.

Hardware (and to some degree, software) considerations:

- Suitability for project's needs
- Ease of use / expertise required
- Compatibility with systems already used, for example at your own institution or the one you will be working with, such as a local university or national park office
- Support in the host country
- Costs
- Robustness for fieldwork: heat, humidity, dust, cold….

Maintenance, especially for printers:

- Consumables (ink, toner, paper, etc.)
- Connections (computers - data recorders - GPS - power supplies)
- Power supplies, in the field and smaller towns: solar recharging panels, lightweight generators (500W), vehicle batteries
- Insurance / packaging / import duties (or tax-free status?)
- Regulations about technology import / export

Training and testing – crucial for a successful expedition:

- Attend the annual EAC Expedition and Fieldwork Planning Seminar - EXPLORE (held in mid-November at the RGS-IBG, London)
- Attend specialist GPS and GIS courses (low-cost workshops are run by the EAC)
- Assemble your equipment, check how it works
- Organise a 'dry run' over a weekend to test all equipment in the field and become familiar with its usage
- Check how you are going to collect, analyse and present data on the 'dry run' to simulate expedition conditions

How will field data be recorded?

- Pencil and paper, then type into a computer in the evenings. Pro-forma systems for data entry can save a great deal of time in the field and help ensure a degree of consistency in observations; an example is shown in the Appendix.
- Enter data in the field, via laptop, data-organiser, GPS with data-logging capabilities
- A ‘belt and braces’ combination: use of both paper and digital data collection
- Other methods: photographic, specialist instruments, data loggers

How will locations be recorded?

- GPS data loggers, computer connections

How will you minimise the chances of errors occurring during data collection?

- Try to establish a clean, secure, and tidy field laboratory
- Check data as you work: revise data collection methods if necessary; download digital field data at the end of each day
- Have back-up equipment & procedures, *use paper records, as well as digital data*
- Prepare a data-loss disaster plan e.g. if your computer or GPS crashes, theft etc.

Document details about what you sample!

In summary, the key to a successful expedition with fieldwork projects using GISci technologies, is to have a well thought-out aim and set of objectives; a project team who are well-versed at using the GISci kit and techniques under field conditions; a sound sampling strategy, preferably building on a recce visit or advice from co-workers in the host country; and a tried-and-tested data back-up system to avoid data loss. And finally, remember that all of this logistical planning involves a lot of thought, a lot of people and a lot of time: allow at least one year, preferably two, to get your ideas from the drawing board to fruition. A little bit of planning can avoid a lot of confusion in the field.

Figure 12-3 Geoscientists working on the western edge of the Sahara ponder the next move…..

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section D: Planning & Practicalities

Chapter 13: Field Equipment

13 Field Equipment

13.1 Field-based applications

Until recently, much of the GIS component of expedition work may often have been conducted separately to the fieldwork component. The two were often seen as different skill sets, with different team members proficient in different disciplines. There is, however, a considerable advantage in being able to study GISci data *in-situ* and work with technologies in the field. At the very least in any modern fieldwork project, there will be digital equipment used for collecting data. This equipment will be vital to the expedition and due to the nature of most fieldwork will be irreplaceable when the work has started. Care and consideration must be given to making sure the apparatus performs to an optimum level for as long as possible. The ability to take GIS into the field means the expedition can respond more effectively to changes in the project plan and can be very cost effective. Failure to treat equipment with care and respect can be very costly to an expedition. The following chapter discusses how equipment can be used safely and reliably in often-hostile conditions. The Bogda Shan Expedition 2000 is used as an example through most of the chapter. The Expedition is described in detail in the appendix.

13.2 Field computing overview

Expedition computers are commonly used to store collected data and to plan a day's fieldwork. Fieldwork planning commonly involves the use of the GIS to view digital images and maps to help define and target fieldwork objectives. Having this capability is very useful, allowing large amounts of data to be collected more effectively, and facilitating changes if a fieldwork plan does not run smoothly.

The GIS can be used to conduct in-field analysis, but most of the computer modelling and interpretation is usually conducted after a fieldwork session has been completed. The ability to have access to computer facilities while in the field can be critical to many projects, but the field PC does not need to be a high specification machine. The field PC would normally be a relatively cheap low to medium specification laptop PC. Details for obtaining field equipment can be found in Section 13.11. The problem with field computing is that modern laptops are generally poorly equipped to deal with field use. Factors to note when considering using a PC in the field are those of protecting it from damage (either by drops, falls or poor weather conditions) and powering it. Section 13.4 describes the considerations for protecting equipment in inclement conditions and Section 13-5 describes considerations for powering the device when no mains AC power is available.

There are many other pit-falls that the team should avoid when relying on laptops PCs. One common problem is the ability to view the screen in direct sunlight as most screens are not suited for this task. Laptops are rarely used in very bright natural light so people are unaware of the limitations. In bright sunlight most laptop screens are unreadable. The team should check this before deciding to rely on the PC in the field. There are a number of technologies that help viewing in bright light. Many field hardened computers are available with transflective screens that are better suited to outdoor use. Other screen technologies include double bulb screens that are far brighter (but consume more amps) and high gloss screens such as the Sony X-Black screen that are brighter but in some conditions suffer from glare. Another important check is that the screen is not prone to damage. If a laptop screen is damaged in the field it can be a major issue for the expedition. An easy test for the reliability of a screen is whether the top of the computer touches the TFT screen when load is placed on it. To check, place your thumb on the plastic rim of the screen and apply a firm pressure between the thumb and forefinger. If the screen discolours around the pressure then the screen is prone to damage in the field. If there is no obvious discolouration then the PC is ideally suited to expedition work. The Bogda Shan Expedition 2000 used a Panasonic CF-25 Toughbook with an Intel 200 MHz MMX processor, 32 MB of RAM and a 2.1 GB hard drive. This PC was designed for outdoor use and was better at displaying data in sunlight but still not perfect. Though the laptop was technologically very dated its field design more than compensated for its technological limitations. More modern models are still behind the cutting edge; the CF-R1s use Intel M class chips running at 800 MHz and CF-50s use a 1.7 Ghz P4. The low specification of the CF-25 was acceptable for viewing images and planning workloads but slower machines should not be considered for modern GISci. The significant advantage of the Panasonic CF-25 despite its low processing power was the magnesium alloy weatherproof casing, which is an important consideration in cold or wet environments. The modern CF range is not the most protected laptop available but offer some degree of protection (see Section 13.3.1 for more details).

13.3 Data storage

The data the expedition collects will need to be stored. There is nothing wrong with recording information in a field book and, in many cases, that is where the majority of the notes will be made. The disadvantage of hand written information is that it's static, can not be manipulated, shared easily or backed up. To use the GIS to the best of its ability requires digital data to be available in the field. This data needs some form of storage.

13.3.1 Field PC for data storage

Digital images such as satellite scenes and scanned aerial photographs can create very large files. Though they can be printed and taken as hardcopies in the field, this limits their use. Hardcopies have a fixed scale, which restricts the data that can be seen and they cannot be digitally analysed. It is useful to have the data in digital form so that it can be integrated into the GIS. The size of the images can be up to several hundred megabytes, so storing and retrieving the data can be difficult and time consuming. Digital images can be successfully reduced in size, with acceptable loss of information, using tools such as ER Mapper's ECW compression algorithms (freeware data compression for images <500 MB), or free JPEG software (though most JPEG software will lose any co-ordinate information vital to a GIS).

Collected data is traditionally very small in size. Retyped field notes will generally be less than a few megabytes in size and downloaded GPS co-ordinates will also be very small. A recent mapping project conducted by the author used real time spooling of data from GPS receivers that generated very large volumes of raw GPS data but this still came to less than 20 MB of data after several days of collection.

These forms of data can be stored with ease on the hard drive of any laptop PC. Larger storage space is only required when storing digital images. These might be satellite or scanned air photographs, or photographs taken in the field with a digital camera. A Landsat scene is typically between 300 and 600 MB depending on the version of Landsat used and field photos from a digital camera may take 1 MB each (approx 2048 x 1536 pixel high-quality JPEG). Aerial photographs scanned into a PC will often create large files. A 600 dpi scan of an aerial photograph with a 20 cm pixel size resolution creates a 625 MB scene for each 5 x 5 km square, comparable in size to a Landsat 180 x 180 km scene. The Bogda Shan field PC was used to store around 900 MB of data and 750 MB of program files. On a 2 GB field PC this is a significant consideration. It does not take into account extra data acquired in the field, digital photographs or swap files generated by the PC while processing for which there should be a minimum of 200 MB.

Table 13-1 Storage space required for Bogda Shan Expedition.

Digital imagery	
1 full Landsat scene ETM+	590 Mb
2 processed Landsat subscenes	165 Mb each
2 high resolution Corona scenes	120 Mb

Programs	
Windows 98	200 Mb
Microsoft Office	200 Mb
Arc View	92 Mb
ER Mapper	271 Mb

This was a data storage total of about 1.6 GB, which still left ample room for storing GPS data gathered in the field. The data stored during the day in the GPS receiver memory was downloaded nightly onto the field PC and stored in Excel spreadsheets for import into the GIS. Over a period of four weeks fieldwork the GPS data set generated was 1.5 MB in size with an accompanying 1 MB Access database. The entire GIS project was described by a 1 MB file referencing all the stored image data and data tables. It is clear that the actual computer need not be of a high specification, and field protection is a much greater concern than processing power or storage space.

Table 13-2 Some typical sizes of geographical data.

Data	Unit	Storage space
Aerial photography	1 x 1 km	25 MB
Corona	1 x 1 km	8 MB
Landsat ETM+ subscene	1 x 1 km	0.01 MB
Landsat ETM+ full scene	1 x 1 km	0.2 MB

As well as the PC hard drive, the laptop can be used to store data on removable media such as floppy disks, optical disks (CDs and DVDs) and solid-state storage chips. The relative advantages of these devices are discussed in Section 13.10. Alternatively, there are several other methods for storing field data when a full-blown PC or laptop cannot be used.

13.3.2 GPS memory for data storage

Modern GPS units store 500 or more named waypoints with or without a graphical symbol. The name is often limited to less than 12 characters and is not useful in recording any significant quantity of information. For some purposes, however, it is entirely possible to generate a series of keys to accompany a waypoint. These keys could be referenced via a table in a notebook and a letter or number could be used to denote a photo or sample being taken at a particular point. A waypoint numbered as 23 at which photo 16 was taken along with sample 4 could be described by the waypoint name as w23p16s4.

Trackpoints cannot have names or symbols attached to them and must be backed up more frequently than waypoints. Trackpoints may be overwritten when the trackpoint memory has been filled making them a poor method for storing data. Trackpoints can be grouped into tracklogs (simplified trackpoints sometimes called routes) but this reduces the resolution of the track log to <50 points. To be certain that the data is safe requires backing up to a PC, laptop PC, PDA, data logger or by hand to a notebook. The sheer quantity of data stored can make written methods of backing up data inappropriate and so a digital medium is often required.

13.3.3 Data loggers for data storage

Data loggers are devices used to store large amounts of information in compact digital units. Several have been developed for GPS receivers and make convenient and portable methods for the storage of large amounts of GPS data (130,000 waypoints compared with 500 in a GPS unit’s memory). Data loggers spool data from a GPS continually for 10 days and run on a 9V battery. They come with download software to transfer the data to a PC and make very useful pieces of equipment when downloading the GPS directly to a PC every night is not possible. They can be programmed to take NMEA sentence structure and time stamped location data. The Delorme Earthmate BlueLogger GPS can also store pseudo-range data for post processing.

13.3.4 PDA for data storage

PDAs (Personal Digital Assistants) are handheld computer units that have capabilities similar to desktop PCs but with reduced processing power and functionality. The market leader is Microsoft with the Pocket PC range accounting for just over 50% of the PDA market. The previous leader in the market was Palm One which still has a strong market share with the Palm series of handhelds.

The first Palm was released in 1996 and competed against PSION handhelds, which have subsequently become obsolete. Currently the two major types of PDA on the market today are Palm OS PDAs (those made by Palm-One or compatible with Palm OS software) and Pocket PCs powered by later versions of Microsoft's Windows CE called PocketPC (a scaled down version of the Windows 2000 kernel). Many third parties make models of each, with IBM, Qualcomm, Symbol Technologies and Sony making Palm compatible devices and Toshiba, Fujitsu-Siemens and Hewlett Packard/Compaq making Pocket PCs. In general, the Pocket PC is the more powerful of the two types but with this extra power and colour screen comes a reduction in battery time and robustness. Palm OS 3.5 supports only 256 colours models but newer models supporting OS 4 and OS 5 are on a par with Pocket PCs. Pocket PCs have up to 16,000 colours, making them more useful for displaying images. Palms have been around longer than Pocket PCs and have more shareware software developed for them. Both types can be integrated with GPS either through cables or via Bluetooth.

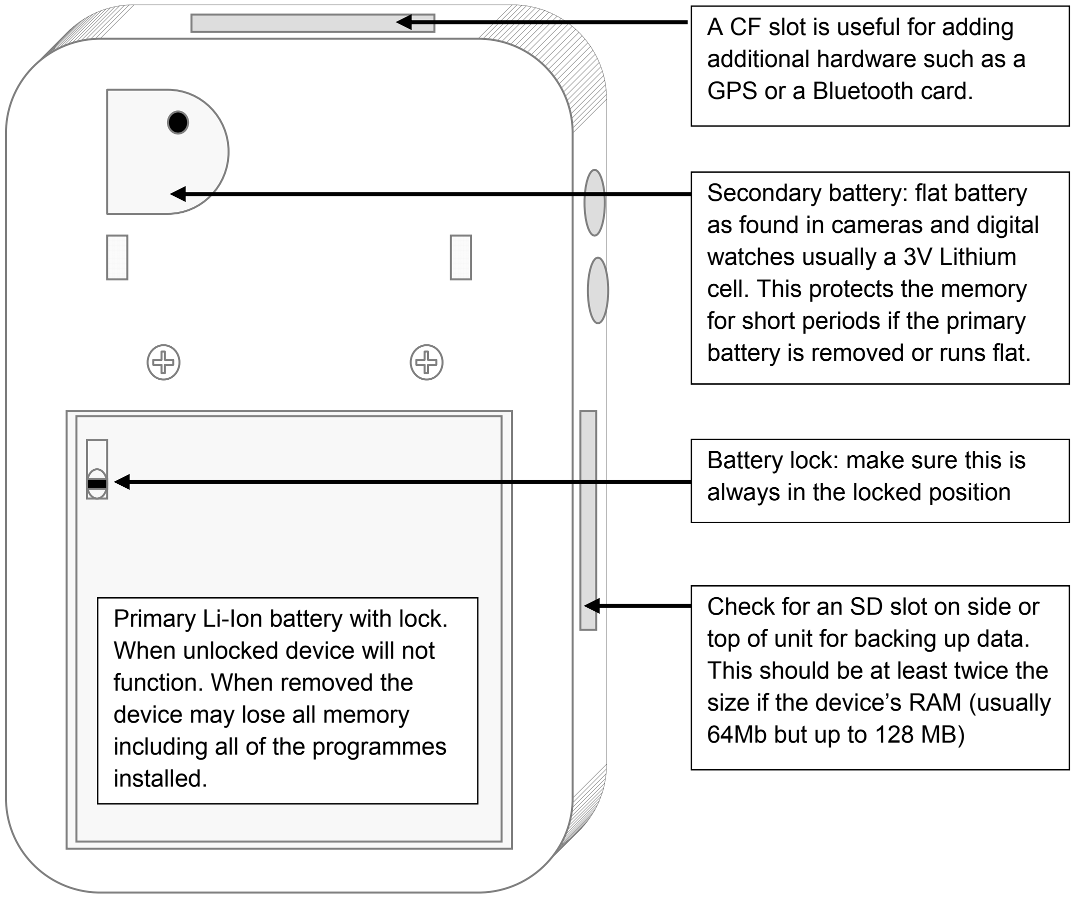
The basic memory on the units can be small: typically Pocket PCs will come with around 48 to 128 MB of which ~20 MB is used for applications. To ensure data is not lost when a battery expires and to allow more data to be collected, an SD card, CF chip, Memory Stick or MicroDrive should be attached. These have storage capacities from anywhere between 16 MB and 4 GB. They often represent very good storage methods if you can communicate between your hardware and the PDA. Bear in mind that many new PDAs do not have RS232 ports neither on themselves nor on their docking stations. This limits how effectively data can be transferred from a GPS to the unit.

Table 13-3 Approximate costs of data storage (discussed in detail in Section 13.12, Disaster recovery planning).

Device	**Capacity**	**Cost ***	**Limitation**	**Notes**
CF Card **	16 MB	-	Not widely available	-
CF Card	64 MB	£15	Requires Card Reader for use with PCs	Rewriteable
CF Card	128 MB	£20	Requires Card Reader for use with PCs	Rewriteable
CF Card	256 MB	£30	Requires Card Reader for use with PCs	Rewriteable
CF Card	512 MB	£50	Requires Card Reader for use with PCs	Rewriteable
SD Card	16 MB	-	Not widely available	-
SD Card	64 MB	£20	Requires Card Reader for use with PCs	Rewriteable
SD Card	128 MB	£30	Requires Card Reader for use with PCs	Rewriteable
SD Card	256 MB	£35	Requires Card Reader for use with PCs	Rewriteable
SD Card	512 MB	£65	Requires Card Reader for use with PCs	Rewriteable
USB Stick [#]	32 MB	£30	PC requires USB Port (not compatible with PDA)	Rewriteable
USB Stick [#]	64 MB	£15	PC requires USB Port (not compatible with PDA)	Rewriteable
USB Stick [#]	128 MB	£25	PC requires USB Port (not compatible with PDA)	Rewriteable
USB Stick [#]	512 MB	£55	PC requires USB Port (not compatible with PDA)	Rewriteable
CDR	600 MB +	~£5 for 10	Requires CD recorder / rewriter	Record Once (WORM)
CD RW	600 MB +	~£10 for 10	Requires CD recorder / rewriter	Rewriteable
DVD (various)	4700 MB	£3 - £5 each	Different formats. Ensure expedition DVD equipment can write to disc	Rewriteable

** Prices are approximate, early 2005. Some companies will sell cheaper than this, others will be more expensive. Mail order companies such as those listed in the popular monthly computer magazines will typically offer better value than high street stores. Prices will continue to fall and this list will become out of date. However, the list should show the relative costs of the media e.g. CF will always be ~75% of the price of SD and approximately the same price as USB data sticks.*
*** Using memory cards in a PC requires a card reader costing £15-£20 connecting through USB.*
[#] Most PDAs do not have mass driver support for USB memory, even if they have a USB connector.

The main disadvantage of PDAs is that they lose the information in their memory if their battery discharges. Information saved to a backup card is safe but data saved to the device will be lost. Given that battery life only gives 6-10 hours of continual use on some models, this is a major concern for expeditions. Pocket PCs are usually more susceptible to this than Palm devices, as the battery on Palms usually lasts longer. It is important to note how long the device can be kept on standby. Some devices will lose >10% of battery life per day even if they are not used. Other models will lose only 1 or 2% per day. This problem is discussed in Section 13.5.2 in relation to the lithium-ion battery supplied with the unit. As important as the primary battery life is the presence of a secondary battery. The secondary battery cannot run the unit but will protect data if the primary battery is removed or discharged. Without a secondary battery if the primary one is removed, the device will lose all data stored on the unit. Though the primary battery is rechargeable the secondary one is not and after discharging it will need replacing.



A CF slot is useful for adding additional hardware such as a GPS or a Bluetooth card.

Secondary battery: flat battery as found in cameras and digital watches usually a 3V Lithium cell. This protects the memory for short periods if the primary battery is removed or runs flat.

Battery lock: make sure this is always in the locked position

Check for an SD slot on side or top of unit for backing up data. This should be at least twice the size if the device’s RAM (usually 64Mb but up to 128 MB)

Primary Li-Ion battery with lock. When unlocked device will not function. When removed the device may lose all memory including all of the programmes installed.

Figure 13-1 Key features at the rear of a Pocket PC.

A PDA is very useful for recording digital data and very quickly transfers it to a computer GIS. The real advantage of a PDA over a field slip is in the amount of time eliminated from the procedure of transcribing notes and uploading them. If digital data is required to be viewed in the field or if the GISci project has to be delivered very quickly on return from the field, then a PDA is an excellent way of speeding up this process. To be certain the data is protected the PDA needs to be equipped with some form of backup card like an SD or CF card. It is important to check the slots available in the unit before leaving for the field. If the unit only has one expansion slot and the expedition requires a Bluetooth card to be added then there will be no space left for a backup card. These factors must be considered before leaving for the field. The latest Pocket PCs also do not have an RS232 connection so it important that the PC or GPS is compatible with the USB connector on the PDA. Good units for the field include the Dell Axim X5 which is sturdy and features rubber casings that will protect the unit to a certain degree. It is also cheap, features SD and CF expansion and has a secondary battery. Battery life on the X5 is also excellent and it retains its charge well during periods of none use.

An important factor when selecting a PDA or field PC is the screen. Most PDAs and most laptops have screens that are designed to be seen in office environments or out of direct sunlight. In the field, where ambient brightness can be very high, many laptop screens are unreadable. This can be mitigated by using the device within a vehicle but using a computer outside is often impossible. This must be checked on the unit selected in advance. Not being able to read the screen renders any of these devices pointless and the expedition would be better off using field slips and paper notes.

13.3.5 The Internet & FTP sites

The previous examples have all included information for storing data locally. Though Flash cards and CD-ROMs are excellent storage mediums, they are still kept with the party in the field. If luggage is mislaid or equipment stolen all the data collected will be lost. By far the best and most re-assuring method for data storage is a periodic upload to the Internet.

Section 13.7 on field communications and Section 13.12 on disaster planning both go into this in more detail. The team should be aware of the ease of backing up data from the field via a mobile phone or sat phone or from the nearest town or city using an Internet café. This either uses a simple email to send the data to an email account or a file transfer protocol programme (FTP) to put the data into a folder on a remote computer or network. There are many cheap file stores available to the party and some of the common ones are listed in Section 13.12.1. The expedition should take the security of its data very seriously because this often underpins the goals of the expedition. As well as risk assessments for the tasks undertaken in the field the team must adequately consider the risks to their data and look for best practise methods to mitigate any loss.

13.4 Understanding device protection

Electronic devices, including GPS receivers, field computers and communication equipment are becoming essential for fieldwork. To maintain safe and reliable use in all conditions, both battery life and extreme weather performance are an essential consideration. The expedition must be prepared to supply either a sufficient number of standard batteries or use some other form of power such as deep-cycle lead acid batteries, solar power or a car battery.

Equally important as powering the devices is consideration for their protection. If the equipment is used for expedition critical information then it must be able to work in inhospitable conditions. The degree of protection offered by a device is often based on the type of equipment it uses. In general the fewer moving parts involved, the less likely the machine is to break from drops of falls. For this reason PDAs can sometimes be more rugged than laptops, due to the solid-state hardware employed. Most commonly, as discussed in Section 13.3.4, PDAs have severe limitations and their use should be carefully considered before selecting them for the field. Also, if in the fall the battery is jolted loose or the pins momentarily separate in the battery housing then the PDA will require a secondary battery to kick in. If there is not a secondary battery or if the battery has already been discharged due to a long period in the field then the PDA will lose all data in its internal memory. Storage cards such as SD and CF cards will not be affected but programs and data that can not be stored on to the removable medium, such as contacts, calendar information and some other file types will be lost.

When purchasing equipment it is important to ensure the equipment is complaint with a relevant hardware specification. Some units quote the US military specification MIL STD 810F which is a standard for field equipment. Another more common test is the IP rating (Ingress Protection) that takes into account the level of protection from moisture and solid contaminants. This chapter will look at the meaning of IP numbers, as these are the most common standards used for expedition hardware. The IP number is a two-digit value. The first digit is between 0 and 6 and represents the resistance to solid contaminants such as dust and other particles. The higher the rating the more protection the device yields. The second digit is for liquids and ranges between 0 and 8. A value of 0 represents no special protection (standard electronics) and the higher the number the better the protection. The highest levels 6 and 8 respectively represent complete protection. Table 13-4 describes the device protection afforded by each digit.

Table 13-4 Ingress Protection (IP) rating of electrical equipment.

1st digit	Description	2nd digit	Description
0	No protection	0	No protection
1	Protected from 50 mm or larger	1	Vertical dripping water
2	Protected from 12 mm or larger	2	Inclined and vertical drips
3	Protected from 2.5 mm or larger	3	Protection from spray from 60°
4	Protected from 1 mm or larger	4	Protected from splashes
5	Dust Protected	5	Protected from higher pressure water
6	Rarely used – complete device protection	6	Protected from heavy seas or high pressure hoses
N/A	N/A	7	Withstand total submersion at a stated pressure for a stated duration.

GPS receivers often quote the second digit to show how comprehensively they deal with water protection. Level 7 is usually quoted as 1 m depth for a brief period, which is what most new GPS receivers can withstand. This number is often found at the end of the GPS specification. For example, the Garmin ETREX is IEC 529 IPX7 where X7 represents the IP level 7-submersion factor. Level 8 is sometimes quoted and this would be protection from a sustained submersion. A common score for hardened laptops is IP54. This offers near complete protection against dust and particulate contaminants and complete protection against particles over 1 mm in size with protection from sustained water from splashes or torrential rain from any angle.

If a greater degree of protection is required than the device inherently supports, then there are various companies specialising in added protection for field equipment. Several digital camera manufacturers make plastic housings for their cameras so they can be used underwater. The housings are usually device specific do that the buttons on the casing line up with the buttons on the camera. These housings are designed for diving but equally protect in areas where splashes to the camera are likely, such as water based expeditions or those conducting coastal or river research. Digital cameras are very susceptible to water damage. Similarly, older GPS units may not be IPX7 compliant and greater protection may be sought for these units. PRO Sports produces AquaPac equipment for housing digital equipment that may be damaged by water or other climate conditions. A heavy duty water proof covering for a GPS can be bought for around £30 and will protect a GPS, mobile phone or similar device. The AquaPac will also allow the GPS to be used below the typical temperatures where the screen would fail. The unit can still be accessed through the covering and this could be a solution for expeditions needing GPS readings in cold climates (Figure 13-2).

Figure 13-2 PDA protection casing. PDA supplied by MapAction.

Water is a major concern for any electrical device and even very small quantities of water can be enough to disable a device. Fortunately, water damage is often temporary and even when damage appears more serious, long term problems can often be avoided through rapid and carefully applied mitigation procedures. One of the common problems associated with water damage is the short-circuiting of contacts within the device. Some of the most vulnerable contacts are those associated with any external buttons. One device that is both susceptible to damage and critical to expedition safety is a mobile phone. Mobile phones are discussed in detail in Section 13.8 but they function as a good generic case study for devices that are susceptible to damage. When damaged by water a mobile phone will usually stop working almost immediately. The expedition must act quickly to reduce the risk of permanent, irreparable damage. The damaged kit should be disassembled into its constituent parts and dried separately. When re-assembled the kit will normally work well. The expedition should make sure that it has a small set of screw drivers to do this.

Figure 13-3 Disassembling a mobile phone to prevent water damage.

13.5 Powering field equipment.

The most important factor when considering the care and use of electronic equipment in the field is battery life. Most equipment used in modern GISci will require some form of electricity. If the expedition is operating from a basecamp where a stable electricity supply is available then this is less of a problem but powering the devices becomes more difficult in the field. Before entering a discussion on specific batteries it is worth understanding the terminology of power requirements.

Batteries supply a current (amps or A) at a specified potential difference (voltage, volts or V). The battery is rated to describe how much charge is stored within it. This stored charge is called the amp hour (Ah) rating and is sometimes expressed as milli-amp hours (mAh) where 1000 milli-amps is equal to 1 Amp hour. Amp hours are a measure of how many hours the battery can deliver the required charge. For instance, a 3 Ah battery can deliver 3 A for 1 hour or 6 A for ½ an hour. This does not hold for all currents as a 3 Ah battery will not supply 12 A for ¼ of an hour as this would likely be too high a current for a small battery. Some batteries are designed for high currents over short periods and the most common type of this battery is the lead/acid car battery. Car batteries are designed to deliver very high currents for short periods to start the car but are unsuited to longer continuous use.

A device requires a fixed voltage to work. The voltage supplied by a battery should remain constant or ‘flat’ for the length of the batteries life; in reality it will never behave like this. In all cases, the voltage will fall off when the amps get lower. Though this happens in all batteries, the gradient of fall off should be as shallow as possible. This is why some batteries stop powering a device well before the estimated amp-hour limit. Cheaper batteries and rechargeable batteries suffer from this more than expensive batteries do and this compounds their often low Ah rating.

Often more than one battery is used at a time in the device. In the case of GPS receivers, the unit may require two, four or even six batteries. When multiple AA batteries are used they are usually configured in series to increase the voltage not in parallel to increase the charge. For GPS receivers the amp-hours of the batteries are not added only the voltage. Therefore, four AA batteries rated at 1.5V with 2 Ah gives 6V at 2 Ah not 6V at 8Ah.

On occasions, it is essential to know the power consumption of the device (also called its wattage). This is a description of the current used at a specified voltage and is found by multiplying volts and amps (Power = current x volts, or P=IV).

Most field devices run from some form of battery, either standard 1.5V AA batteries in the case of GPS receivers or from rechargeable Li-Ion batteries in the case of field computers, PDAs and phones. The following sections look at the different types of battery and how they should be used to maximise their performance.

13.5.1 Battery life from standard AA batteries

For field use there are five different types of standard AA battery that can be used with the field equipment. These different types are shown in Table 13-5 and discussed in more detail below. Occasionally some devices will require a different type of battery such as AAA batteries (slightly smaller than AA and thinner) or 9V rectangular batteries.

Table 13-5 Different battery types.

Battery type	Common name	Size	Voltage	Approximate (mAh) *	Approximate power
Lead / acid	Lead / Acid	AA	1.5 V	600	0.9 watts
Nickel cadmium	NiCad	AA	1.2 V	1000	1.2 watts
Alkaline	Alkaline	AA	1.5 V	1000 – 2000	2.3 watts
Lithium	Lithium	AA#	15 V or 3 V#	2000	3 watts
Nickel metal hydride	NiMh	AA	1.2 to 1.5V	2300	2.75 watts
Alkaline	Alkaline	AAA	1.5 V	1120	1.68 watts
9V alkaline	9V	9V	9 V	560	5 watts

** Amp Hour ratings for batteries are highly variable across different manufacturers. These are rough guidelines. Also, bear in mind that the voltage drop off means that the full capacity will never be discharged because the device will stop registering the voltage before the battery is fully exhausted.*

*** The Amp Hour rating for standard alkaline batteries is never quoted, generally because of the voltage drop off. The expedition should budget for a value in the range of 1500 mAh.*

 # Lithium cells are sometimes supplied as compact units designed to replace 2 AA batteries. If this is the case, then the configuration of the batteries in the device is an important issue, i.e. the batteries must lie side by side and cannot be staggered. Not all lithium cells are supplied like this and some are standard AA shape.

Lead/acid style batteries are cheap but are not suited to the high-energy drain of equipment such as GPS sets and will last less than a third the time of alkaline batteries. Similarly NiCad, though having the advantage of being rechargeable, last for even shorter periods. NiCad batteries also run at lower voltages and are often not powerful enough to run high drain equipment. Alkaline batteries are the standard type used in field equipment but their life is substantially reduced at temperatures below 5°C. In conditions approaching or below freezing lithium batteries are recommended. Modern Nickel Metal Hydride (NiMh) batteries are on the market that have acceptable voltage levels and very high Amp Hours. Modern NiMh batteries are rated at up to 2300 mAh compared with between 1000 and 2000 mAh for alkaline AAs (commonly around 1500 mAh). The disadvantage of NiMh batteries is that they are very expensive at around £25 GBP for a pack. NiMh 2300 batteries will last 1 ½ times as long as a standard alkaline battery. Similar to NiCads the voltage of NiMh is still usually lower than standard alkaline batteries so care must be taken to test any 4 battery device receiver thoroughly before equipping the expedition. Some expedition devices may require AAA batteries: these are rated at 1.5 V like AAs but have a smaller mAh rating. It is often desirable to keep the type of battery consistent through all expedition devices. This means that batteries can be shared between equipment in the case of battery failure.

A GPS receiver is a typical device used in the field on AA batteries. GPS manuals recommend alkaline batteries and often quote their receivers at around 22-24 hours continuous use in battery save mode. This is generally on the high side and realistic use and handling of the set may see this drop to a little over 16 hours (approx. two days field time). The rechargeable NiCad batteries give about three-quarters of a day's use in the field (~5 hours) but obviously have the advantage of being rechargeable off a lead/acid car battery or solar panel. With the use of a lead/acid battery or solar cell these batteries can be recharged daily, however, the voltage fall off is worse with rechargeable batteries than with other batteries. This means that towards the end of the charge even though they have enough amperage they do not have enough voltage to power the device. This is compounded by the number of batteries a GPS requires. Four battery receivers fair worse on rechargeables than two battery ones. This can be used to reduce weight for a team but is often a false economy. Most rechargeable do not deliver as high voltage as standard alkaline AAs. For a two battery GPS this is not a problem but some GPS receivers require four AA batteries. Four batteries are more common in older models and the difference in voltage between four rechargeables and four AAs is often too great for the GPS. This is shown in Table 13-6.

Table 13-6 Effect of differing voltages on electrical equipment.

Battery requirement	Standard AA batteries	Standard rechargeables	Voltage explanation	Acceptable Performance
1 battery	1.5 V	1.3 V	0.2 V - no real difference	✓
2 battery	3 V	2.6 V	0.4 V – no real difference	✓
4 battery	6 V	5.2 V	0.8 V – significant difference	×

To calculate the battery life of the device used requires knowledge of the current it pulls against the current stored in the battery. Standard alkaline batteries are rated around 1000 to 1600 mAh so it is possible to get an approximation of the length of time the batteries will last in the field. A GPS unit pulls about 100 mA. This is increased by around 10-15 mA if an external antenna is attached. Many GPS receivers have a back-lit screen that can be enabled, sometimes by pressing the power key for a short length of time. This reduces battery time significantly and should only be used for very short lengths of time. The current pulled depends on the unit, the size of the screen and the exact settings used but adds around a further 40-50 mA to the current. If an electronic flux-gate compass is enabled then the current is increased by another 40 mA. WAAS also increases battery drain considerably and both the electronic compass and WAAS should be disabled if they are not needed. Therefore, a GPS would last for around 10 hours in standard mode using 1000 mAh batteries reduced by around 1 hour by adding each option above. Realistically, using good quality 1600 mAh AAs battery life could be between 16 hours and 8 hours depending on the options selected. Modern 2300 mAh NiMh batteries would last for closer to 23 hours with most options turned off which would give around three days field time for a GPS in battery save mode. This is shown in Table 13-7.

For subzero temperatures, lithium AA batteries are better than alkaline AAs. Lithium batteries carry up to three times the charge of alkaline batteries and weigh less. Their charge is affected far less by the cold, making them a good choice for cold weather environments. Turning the set off for periods can also extend battery life, however, this does carry the caveat of waiting for a warm fix and for the data to settle.

Table 13-7 Current pulled by differing devices.

Device	Standard AA batteries	Approximate current	Approximate battery life using 1600 mAh AAs
GPS in battery save mode	2 * 1.5 V	100 mA	16 hours
GPS as above with aerial	2 * 1.5 V	110 mA	14 hours 30 minutes
GPS with aerial and backlight	2 * 1.5 V	150 mA	10 hours 30 minutes
GPS with electronic compass	2 * 1.5 V	140 mA	11 hours 30 minutes
GPS with Bluetooth	2 * 1.5 V or 3.3 V Li-Ion	110 mA	14 hours 30 minutes

If the GPS is to be used in a vehicle then powering it through the 12V adaptor inside the vehicle is a good method for cutting down on batteries. A power adaptor is usually sold separate to the GPS for around £20 GBP. They connect via the PC connector on the back of the GPS. If this is important to the expedition then check the GPS has an adaptor for this. Basic models such as the Garmin Geko 101 do not have PC connectors and can not be used in this way. It is also worth checking that the vehicle has a 12V adaptor because some older models that may be used on expeditions do not.

13.5.2 Battery life from standard lithium ion battery

Standard laptops come with a single rechargeable Li-Ion (lithium ion) battery rated at around 3600 mAh. This gives an operating time of 1 hour at 3.6 Amps running at 10.8V. The actual current pulled by the computer varies with the application. A typical test of amperage pulled by different computer functions while running off the battery is shown in Table 13-8.

Table 13-8 Tests conducted on field computer. Figures show mean A at 10.8V for Panasonic CF25.

Status	Amps
Idle	1.7 Amps
Low level processing	1.9 Amps
Booting	2 Amps
High disk usage	2.05 Amps

For normal usage, downloading, processing, and uploading data, the PC will average a current of 1.9 Amps. On a 3600 mAh battery this will give a performance life for a standard Li-Ion cell of:

Run time = Amp hours/Ampere = 3.6 / 1.9
= Approx. 1.9 hours, about 1 hour 55 minutes

Due to inaccuracies and potential idle time, we can round this to two hours of continuous use. Various battery designs exist but the optimum performance is achieved with Li-Ion. Older NiCad (nickel cadmium) batteries are now rarely used due to memory effects. Memory effects involve a reduction of battery capacity when the battery is not discharged fully before recharging. A second alternative, NiMh (nickel metal-hydride) has no memory problems but exhibits inferior performance such as shorter battery life compared to new Li-Ion batteries. The advantage of NiMh over Li-Ion is its resilience to low temperatures. Li-Ion batteries have poor reliability in cold conditions and are prone to failure at temperatures below 0°C. Therefore for the mountain expeditions or those in cold climates a NiMh would have been preferable. It is recommended that to protect against failure and to extend battery life, additional battery power is provided by either a second battery or other rechargeable mechanism. As MHz clock speeds have increased in laptops their battery requirements have also increased. In some instances, it is worth the expedition looking at the processor requirements and opting for a lower speed to increase the battery life. This may involve buying older second hand PCs. The disadvantage with second hand laptops is that batteries can sometimes have a short life span (2-4 years) and therefore second hand units may have damaged batteries that require replacing. Intel produces Centrino and Dothan processors that are designed to pull smaller amperages than standard chipsets. Some Centrino processors can run for 3 to 5 hours on a single charge, significantly more than the 1.5 to 2 hours supplied by a laptop running a P4 processor. Lithium Polymer batteries (Li-Pol) give better performance by size and weight than Li-Ion but the difference is not significant. Li-Pol are becoming popular on the new Palm-One PDA units including the modern Zire and Tungsten range.

Ensuring a good power supply for the expedition PDA is vital due to the device's volatile RAM. As discussed in Section 13.3.4, the PDA will lose all its memory if the battery discharges. A typical Pocket PC PDA is supplied with a Li-Ion battery rated around 1000 mAh to 1400 mAh but they can often be fitted with high performance batteries rated around 3500 mAh. These batteries are usually expensive (~£100 GBP) and add significantly to the weight and size of the unit. Palm-One units are usually supplied with smaller batteries rated around 840 mAh or 900 mAh. However, their battery drain is significantly less than that of a Pocket PC and the battery will usually last longer. A rough rule of thumb for PDA battery drain is that in normal operation the unit will pull around 100 mA per 100 MHz clock speed. Therefore, a 126 MHz Texas processor Tungsten E from Palm-One would pull a current of around 126 mA and last just under 7 hours continual usage on its 840 mAh battery. A 400 MHz PXA 250 Pocket PC running on a 1400 mAh battery would run for less than 4 hours. The X-Scale processors from Intel used in many PDAs are designed to vary the processor speed (and therefore battery consumption) depending on the task being performed. Pocket PCs running Pocket Windows 2003 can take advantage of this but older versions of Windows cannot. In those cases, software such as Pocket Hack Master can be used to force a change in clock speed. This can reduce the clock speed to conserve the battery or increase speeds to increase performance. Over clocking is not recommended as it significantly reduces battery life and may damage the unit. Also very few applications need high processor speeds.

The above rules are not hard and fast rules and depend on the way the device is used and, most importantly, the level of brightness on the screen. A screen set to maximum brightness will almost double the charge pulled. Using a wireless communication like

Bluetooth will add another 50-60 mA. These rules give a good approximation for most PDAs but the expedition equipment should be fully tested before leaving for the field. In addition, in conditions below 15°C the battery life will fall away and perform significantly worse than in the optimum 15°C to 40°C range.

PDAs should also be fitted with a secondary battery. It is better if this is a separate cell but some PDAs simply reserve part of the primary battery for backup purposes. This is of little use if the primary battery is physically removed either by accident, for cleaning or replacing. A secondary battery is easy to check for, as shown in Figure 13-1. The secondary battery is usually a CR flat coin type battery (CR2032). These deliver a nominal 3V and hold 220 mAh charge. The length of time the secondary battery will protect the device for depends on the amount of charge required by the unit to keep the volatile SD RAM active. This can be anywhere from 1-10 mA and so could be from 4 to 40 days from the primary battery or between 1 and 10 days from the secondary. If the secondary battery is a CR flat battery then it will not be rechargeable. The Dell Axim X5 discussed in Section 13.3.4 takes just over 1 mA whereas some older units such as the Toshiba E740 take closer to 10 mA. In the case of the Dell Axim X5 the primary battery will last around 30-40 days without charge, with an additional 8 or 9 days supplied by the secondary battery before all memory is lost. The Toshiba E740 has no secondary battery and so would only last around 1 week before the memory is lost. Calculating or estimating this in advance is vital because the device will hard reset if left for longer than this period, a real problem if a lengthy travel to and from the field is required.

Larger Li-Ion batteries can be purchased that can be used to power multiple devices. The Socket Mobile Power Supply available from *www.socketcom.com/product/AC4009-541.asp* can be used to power multiple devices and has a 7.2 Ah rating. This would be useful for charging equipment or running PDAs for a longer period of time. A Dell Axim x5 would run for a total of 18 hours using this equipment. The device is bulky and heavy but can be clipped to a member's belt for continuous use or kept at basecamp to charge equipment. The specifications claim the battery can recharge a PDA 10 times per charge or recharge a mobile phone 15 times.

13.5.3 Battery life from solar panels

One method of alternative power is the use of a solar panel. A general solar panel design for powering laptops uses two photovoltaic cells. The solar panel apparatus is usually housed in a strong waterproof casing with hardened specifications inline with 'ruggedised' field PCs. The panels themselves are coated to be weatherproof and withstand small impacts. These panels are connected to a voltage controller, which comes as a standard part of the equipment. Calculations suggest that the recharge of a laptop PC during the course of one full day of bright sun light would allow just over one hour of PC use a day.

Table 13-9 Battery life generated by a full day of recharging from a solar panel that delivers 0.29 A per hour of direct sunlight.

Solar panel performance	Approximate current	Results
Over an 8 hour period	0.29 A * 8 hours	2.32 A @ 18 volts
Run time of PC if charged from a solar panel	2.32 / 1.9	1 hour 13 minutes

A modern solar cell designed for expedition work is made by ICP Global Technologies Inc. Their iSun Portable charger is a solar power device capable of delivering up to 2.2 amps that can be used for powering GPS units and PDAs. The current is generally too small to directly power a laptop computer, which requires a minimum of 2 amps because the generated amperage is dependent on the brightness of the sun. However, the unit could be used to trickle charge a laptop battery, as shown in the battery life worked example above. The units can also be daisy-chained together and two units working together could be used to run a laptop. The units are weatherproof and can be used outdoors, though they are not certified as being waterproof and should be protected from extreme conditions or submersion. Their temperature range is between -40°C and +80°C and this should be sufficient for most expeditions. More information can be found at *www.isunpower.com*. In September 2003 the units were retailing for $80 US and available from some retailer in the UK for around £35-£40. They have subsequently become slightly more difficult to locate and at the time of printing (early 2005) are retailing for closer to £50. The expedition should try to shop around to get the best price.

For more heavy duty use and for powering more equipment Maplin produce a 45 watt solar panel costing £349.99 from *www.maplin.co.uk*. This unit consists of three 15 watt panels. This is a much larger more powerful power source but is consequently far more expensive. The iSun is a more practical solution and is shown below in Figure 13-4.

Figure 13-4 Garmin ETREX GPS and I-SUN charger.

Other useful tools for conserving battery power include windup torches. Windup torches can be useful pieces of field kit and offer 16 minutes of light after a 60 second turn. They can be further wound to offer up to five hours of battery life. Windups can alternatively be charged from the mains or from a car battery.

13.5.4 Battery life from lead acid batteries

Alternative recharge methods include lead-acid car batteries, which perform better than solar panels but are not designed for deep cycle operation (i.e. draining to empty and recharging). These batteries are rated between ~45-50 Ah and are designed to generate high amperage over very short periods but not to drain. If a vehicle is available to the expedition then its battery can be used for powering equipment. If a vehicle engine is switched off the 12V adapter can supply power to run a field PC for over two hours but running for too long only on the battery can flatten it. Car batteries do not perform well if they are discharged (deep cycled) and starting the vehicle will be problematic; flat batteries require jump starting from a second vehicle. To negate this, the vehicle's ignition should be started for 10 minutes every hour.

An alternative to lead/acid vehicle batteries are deep cycle lead/acid 'leisure batteries'. Deep cycle leisure batteries do not offer the same current, operating at a maximum of 10-20 amps, but for similar size and weight may be of higher Ah rating. Equivalent deep cycle batteries may offer performances of 80 Ah with deep cycle discharge performance. The Bogda Shan Expedition used a 100 Ah deep cycle, sealed, lead-acid battery purchased in China from the Xinjiang Seismological Bureau. Battery life was never fully tested but gave over 14 hours of continuous use on the Southern side of the Bogda Shan. Calculations suggest that for the operations used in the field the battery should have yielded over 40 days of use. The main factors to be considered when selecting between lead-acid batteries and solar cells are their weight and dimensions. Deep cycle batteries are sealed and easier to transport than car batteries but are correspondingly heavier and more awkward. Battery weight can be over 10 kilograms and they can measure over two feet in length. This makes them significantly more difficult to transport than boxed solar panels and two people may be required to carry them. The Bogda Shan Expedition used camels and horses to transport the equipment to campsites, so battery weight was not significant concern, but where animal or vehicle transport is not available, large deep cycle batteries may be impractical. The following example shows how long a 100 Ah battery could sustain PC given typical expedition power requirements.

Table 13-10 Battery life from a lead acid battery.

Co-ordinate download from GPS:
Download co-ordinate information from GPS: 10 mins.
Process co-ordinate information in GIS: 25 mins.
= (1.9 A x 10 minutes) + (1.9 x 25 minutes)
= (1.9 A x 0.16 hours) + (1.9 x 0.417 hours)
= 1.10 Ah per day

Other field equipment needs: phone charging to maintain communications:
5 hour charge every 3 days @ 0.255 A
= 5 x 0.255 = 1.275 Ah per 3 days
= 0.425 Ah per day

Expedition planning using satellite imagery on laptop:
30 minutes computer time each evening
= 1.9 A x 0.5 hour
= 0.95 Ah per day

Number of expedition days using a 100 Ah battery life:
1.10 Ah + 0.425 Ah + 0.95 Ah = 2.475 Ah per day
100 Ah / 2.473 A per day = approx. 40 days field time

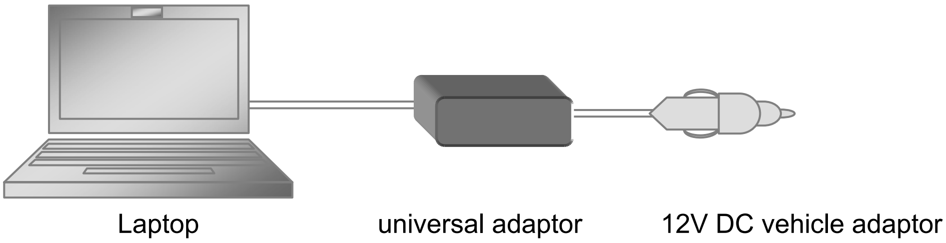
The conclusion is that a deep cycle lead acid battery should give about 40 days of operating time in a remote location. If a car can be reached during this period the battery can be recharged. Completely discharging a standard car battery can be detrimental to its operation and therefore deep cycle cells capable of complete discharge are recommended.

13.5.5 Power adaptors and inverters

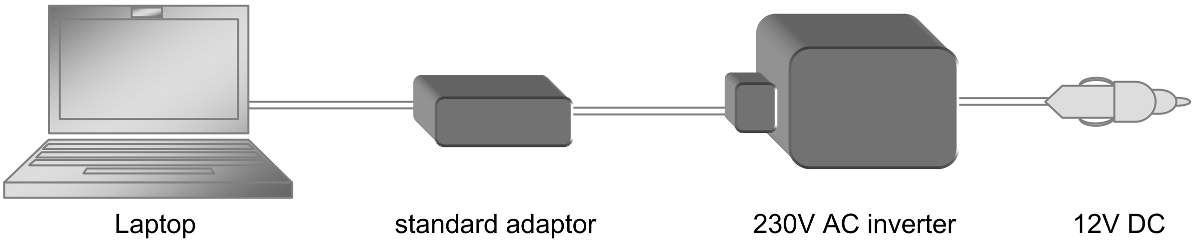
Field PCs bought in Europe can be run from 230-240 V mains power supply. They usually come with an external power supply to transform the 230 V AC to a smaller DC voltage (perhaps 10 – 20 V). Commonly field equipment will need to be used in areas away from mains power. In these cases, power can be supplied from a vehicle that can output 12 V DC. 12 V DC straight from the vehicle is not usually practical for directly running a device. Therefore, the best solution is to keep an inverter available for powering the equipment. An inverter takes the 12 V DC and changes it up to 230 V AC. This standard ‘household’ current can be transformed back to DC using the PCs external power supply.

There are methods for taking the 12V DC straight into the PC without using the step to AC. If the expedition can be certain of the electrical needs of the equipment then this is an ideal solution. However, taking power from the vehicle straight to the device, even through a transformer, can irreparably damage the equipment. Universal adaptors for changing the 12V DC to a laptop’s required DC only work if the output is exactly what the laptop requires. If the voltages are different, the field laptop will be damaged. Due to this risk, it is sometimes preferable to go through an inverter. The other advantage of an inverter is that they can then be used to power any AC device, such as a battery charger, cellular phone, sat phone or PDA. Both inverters and universal adaptors cost around £50-£70. These two methods are shown in Figure 13-5.

Figure 13-5 Methods for powering a field computer from a vehicle:



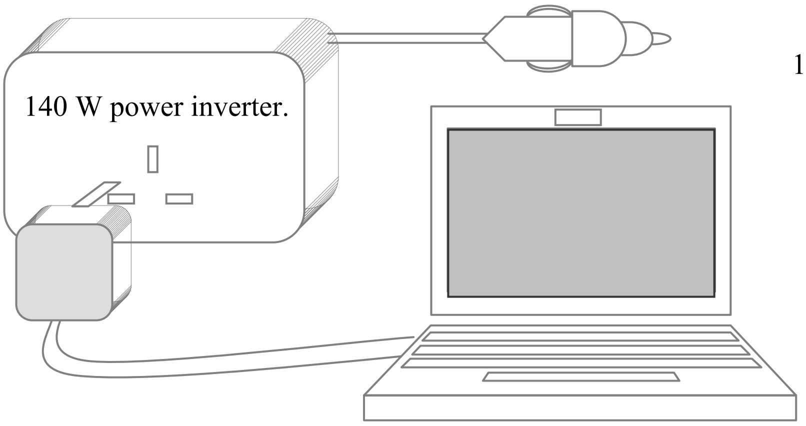
Method 1: Universal adaptor replaces computer transformer and changes 12V DC into 15V DC required for a laptop. Method reduces cables and is easier to use in the field. This method requires that the adaptor is completely compatible with the field PC. It is not recommended because the field PC can be damaged by this method. Other forms of universal adaptor are discussed later.



Method 2: Power inverter is used to change the 12V DC into mains 230V which can then be used with the laptop power cable supplied with the computer. This method has more cables but is safer for an expedition.

When selecting an inverter the unit must be capable of handling the load required by the equipment. This needs to be calculated before leaving for the field. Load or power consumption is measured in wattage. This is sometimes printed on equipment; however, occasionally the power consumption is only listed as amps and volts. If wattage is not present, it can be calculated by multiplying the amps and volts: Power = (Current * Voltage) or P=IV.

Laptops draw currents in the region of 0.4 to 0.7 A at 230 V giving a power consumption between 90 W and 160 W. The current is different to the current listed in Table 13-8, as that showed current from the internal battery not from a 230 V electricity supply. To run a standard laptop from a 12V battery usually requires a 140 W power inverter. Some more powerful modern laptop PCs may draw more power than a standard 140W inverter can easily supply. In these cases, a 300 W unit may be required but that is unlikely and 140 W should suffice. The disadvantage of using a higher wattage inverter is the cost (which can nearly double from approx. £50-£70 for a 140 W to over £120 for a 300 W) and the weight and dimensions (over 2 cm and 100 grams heavier for a 300 W). The other problem with the units is that even when they are not plugged into an appliance, they draw power from the car battery. A 140 W unit draws 0.3 A and a 300 W unit draws 0.4 A. This could quickly flatten a car battery if left plugged in without starting the engine.



12 V from vehicle adaptor.

Laptop pulls around 90 W meaning a 140W inverter should suffice. Modern units may require higher power consumption levels.

Figure 13-6 Schematic layout of equipment required to power a laptop in the field

Adapters can be purchased for GPS receivers that change the 12 V DC vehicle supply into 3 V DC for the unit without the requirement for an inverter. These are often not interchangeable between units so care must be taken to ensure that the unit is compatible with the power supply lead.

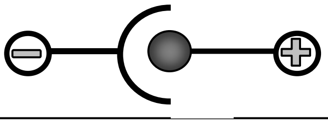
When on an expedition, being in the field limits the repairs you can do and the replacement kit you can obtain. A prudent measure when using equipment with DC adaptors is to purchase a universal adaptor plug. Though these plugs should never replace the manufacturer’s own plug and transformer supplied with the electrical equipment, they are a useful backup tool if things go wrong. Universal adaptors come with a selection of different sized input jacks to plug into various computers, mobile phones etc. Adaptors also have a switch to select the most applicable output voltage. They then change a 240 V electrical current into the current matching the device. You may also need a travel adaptor to change a 3-pin UK plug into the configuration used in your host country. Figure 13-7 shows the details on the back of a Toshiba standard transformer for a UK E800 PDA.

Figure 13-7 Reading electrical equipment requirements.

Input 100-240V @ 50-60 Hz

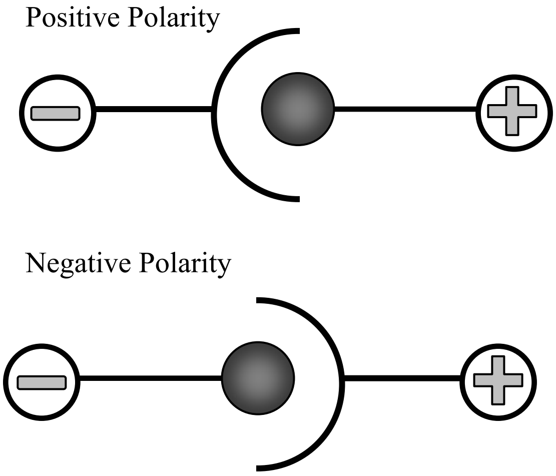
Output 5V @ 3A

The diagram in the bottom right shows the polarity of the input jack, either positive or negative: in this case it is positive.



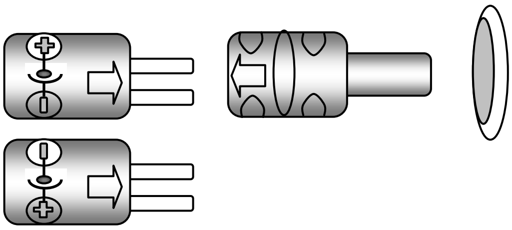
Ensuring you set your adaptor correctly and safely is of paramount importance. On the end of the adaptor where the multi-sized input jacks clip in, there are usually two thin metal prongs. The way the prongs are attached to the jack determines its polarity.

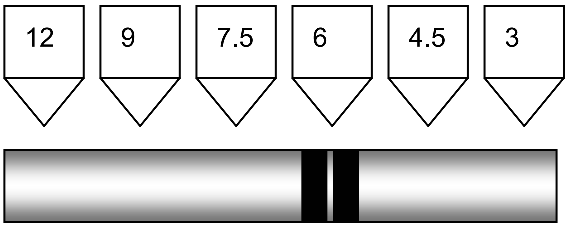
Determining the polarity of equipment



An electrical device connects to a power supply via a male-female adaptor plug. The male tip comes from the power supply and is either positively or negatively charged. This plugs into the DC input jack of the equipment

The polarity of the adaptor must match that of the equipment. Use the key on the left for determining the polarity. Connect the universal adaptor with the correct setting.





The Universal Adaptor is set here to 6 Volts

Power

Figure 13-8 Configuring a universal adaptor.

When using electrical equipment there are some basic standby kit you should have with you. Items that are common in towns may be unavailable in rural areas or in the field. You should not let a simple electrical failure end your GISci work and a little planning should help mitigate these incidents. The tables below show some of the more important back up equipment required.

Table 13-11 Basic field equipment requirements.

Equipment	Use	Cost
AC DC Inverter	Run AC equipment directly from a vehicle by scaling 12V up to 230/240 AC Voltage.	£70 - £150
Universal DC Adaptor	Run any DC device by scaling a 230/240 AC Voltage to any DC voltage.	£5
Fuses (5 amp)	Fuses are a very common item, but can be difficult to obtain in the field. Because of this an ample supply should be kept close to hand.	> £0.50
UK ⇒ Country AC adaptor	Convert UK manufactured equipment (3 pin plug) to the country of the expedition's type.	£1-£5
Non UK ⇒ UK AC adaptor	Convert items manufactured outside of the UK to a 3 pin UK plug type. Useful for purchasing items outside of the UK and using them on return from the expedition.	

Power Supply	Use	Cost
NiMh Rechargeable Batteries	Use 2100 or 2300 NiMh batteries to power GPS, Digital Cameras etc.	£20 (for four)
Standard AA alkaline batteries	Use if the NiMh cannot be recharged. £4 (for four) AA. Use branded makes such as Duracell or Energizer. Avoid Zinc.	
Standard AAA alkaline.		

Standard Kit	Use
Black electrical tape	Useful for patching cables and offering some degree of waterproof protection to a device.
Set of screw drivers (check compatibility)	Check you can take apart any equipment and service any parts you need to. This can save a lot of time in the field.

13.6 Caring for GPS receivers

GPS receivers are most commonly powered using 1.5 V AA batteries. Some units such as the Garmin Gekos use AAA batteries. Battery life can be increased by the use of battery save mode. Battery save mode reduces the refresh time of the receiver and only receives satellite data once a second. There should be no reduction in performance by using this option. Continuous operation for two field days (total +16 hours) per set of alkaline batteries with the receiver in battery save mode will be typical for most expeditions. Tests on rechargeable batteries show that they give half to three-quarters of a day's use (~5 hours).

GPS units being used in vehicles can usually be powered directly from the vehicle battery, ideally via a 12 V adapter and a suitable power cable (available for most GPSs). Older vehicles in some parts of the world will not have modern adaptors, so if you know the car battery output and the required GPS input, then connection can be made by direct wiring with suitable fuse protection.

Operating conditions for GPS receivers are rated from -15°C to 70°C. If the GPS is to be used at the lower extremities of these temperatures such as subzero temperatures, lithium batteries are better than alkaline ones. Lithium batteries carry up to three times the charge of alkaline batteries and weigh less. Their charge is also affected far less by the cold, making them a good choice for cold weather environments.

13.7 Field communications & remote data access

In remote locations if communication from the field is required, expeditions are limited to satellite phones or cellular phones compliant with the Global System for Mobile Communications (GSM) standard. Current satellite phones are usually low bandwidth, high cost devices but offer access from most latitudes. The most popular type of satellite phone works on the Iridium network. The Iridium network uses 66 low earth orbit (LEO) satellites to allow communications anywhere in the world. GSM is cheaper but works within given cells, rarely more than twenty to fifty kilometres from a transmitter depending on line of sight and other factors. Such transmitters are usually only located in highly populous areas. Maps of estimated transmitter range for most countries are easily obtainable from network service providers. The Bogda Shan Expedition relied solely on GSM compliant communication. Signal strength in the valleys was most commonly zero, but was improved at higher elevations, even over 100 km from populous areas. Even though GSM is becoming more widely distributed no expedition should rely on its use and careful consideration should be given to field communications.

For a GSM phone to be able to '*talk*' to a cell it must be compatible with the frequency of the cell. There are essentially three frequency bands broadcast according to GSM standards. These frequencies are GSM 900 (900 MHz), GSM 1800 (1800MHz) and GSM 1900 (1900MHz). Any one phone would originally have been configured to communicate with one frequency. Since the mid 1990s handsets have been developed that can switch between frequencies. These are referred to as dual or tri-band phones. In the UK both GSM 900 and GSM 1800 are used by the main service providers. GSM 900 is used throughout mainland Europe and the Asia-Pacific region. GSM 1800 is also used throughout Europe and in Australia. The third band frequency, GSM 1900, is used in North America and will be introduced in Latin America and Africa.

If coverage is available, access to online databases and datasets is possible but limited to the capability of the handset. Most handsets operating on the GSM network offer 9.6 Kbs connectivity (note Kbs means Kilobits per second compared to KBs which is kilobytes per second, the same notation applies to Mb and MB). GSM speeds compare poorly with a standard V.90 landline connection that offers up to 56 Kbs. Some high-speed WAP enabled phones extend to 14.4 Kbs. GPRS (General Packet Radio Service) uses GSM to combine channels to provide an upper limit of 144 Kbs. This technology is sometimes referred to as G2.5. There are two different types of GSM 2.5 technology. The highest implemented G2.5 technology is the High Speed Circuit Switched Data (HSCSD), which allows communication speeds up to around 40 Kbs both up stream and downstream. In areas of limited coverage, these speeds will fall below quoted maximums. The other form of GSM is an always-on data connection where data volume rather than connection time governs the cost of the service. This form of GSM is much more localised in availability and ranges in speed from 28.8 Kbs to 36 Kbs and is sometimes, erroneously, called GPRS.

Where higher speed access is required, the most significant increase will be the introduction of Universal Mobile Telecommunications System (UMTS), also known as G3 or 3G (Third Generation). G3 will offer maximum throughput of 2 Mbs but in most common installations runs at 384 Kbs in areas of good coverage; it is still unclear how many countries will implement the service. Many phones that are 3G enabled are currently being sold with the data communication parts disabled. Very few service providers currently allow data to be transferred through the UMTS network. As of the time of writing the only data implementation of UMTS in the UK is the Vodafone 3G Mobile Connect PCMCIA card for laptops. This is rated at up to 384 Kbs in areas of good signal strength but rapidly falls off towards modem levels in poorer coverage. Obviously, this type of data is only available in a few major cities worldwide but could be useful for a team to upload their data during periods when the expedition passes through a city. Remote access via satellite phones does not rely on a country's adoption of a given technology. Satellite services of significance to expeditions include Teledesic (*www.teledesic.com*), and EUTELSAT's Hotbird satellite (*www.eutelsat.org/home/index.html*). These are moderately fast services operating at around broadband speeds.

A third alternative, SkyBridge (*www.skybridgesatellite.com*), will use 80 satellites offering 20 Mbs downstream and 2 Mbs upstream connectivity with 30 millisecond (fibre-like) latency, contrasting geostationary earth orbit's 500 millisecond latency, due to their 913 mile low earth orbit. SkyBridge's 435 mile radius footprint gives continual line of sight to at least one satellite for latitudes –68° to +68°.

Any of the systems described would offer high-speed access in remote areas but the future implementation of the technology is still uncertain and expedition access will be controlled by their cost. Iridium currently offers two data packages with global access but they are rated at 2.4 Kbs and 9.6 Kbs significantly below the specification of either GSM or the newer mobile data rates. The take up of sat phones in Africa is currently being encouraged. An initiative called "Go Africa" from Telenor Satellite Services is the cheapest method for getting access to satellite communications in Africa. The system uses Motorola 9505 handsets on the Iridium Network and offers a 40% discount over any other sitcom solution.

13.8 Field concerns with cellular phones

Standard cellular phones are powered by rechargeable Li-Ion battery operating at around 3.6 volts similar to those discussed in Section 13.5.2. Battery time between charges varies according to the handset used. Typically, you can expect between 7 and 14 days of standby in areas of good signal strength. In remote locations signal strength will often drop to near zero or zero. In these environments, cellular phones increase signal strength in an effort to boost reception from transmitters: in such conditions battery life will drop below three days. Actual usage times vary but talk time of three hours in good signal areas, will drop to two in areas of poor to medium signal strength. Calculating the projected battery life of a mobile phone is more difficult than with a PDA. The phone's battery will be listed in the same format (usually sub 1000 mAh) but battery drain is not constant. Older monochrome phones pull less than ½ mA when in standby mode and somewhere nearer 150-200 mA when transmitting (dependant on the reception quality). Modern colour phones pull closer to 1 mA in standby and similar levels when transmitting. Some large screen phones have higher currents with smart phones and UMTS (3G) phones pulling closer to 10 mA. Obviously additional features such as GPS and Bluetooth built into the phone will increase battery drain. GPS often adds around 5 mA and Bluetooth adds another 5mA. Most phones do not lose their memory when discharged but modern phones can be difficult to keep well charged in the field. Phones such as the Motorola A835 with AGPS (Assisted GPS, a method of using additional radio signals to keep the GPS working indoors) will pull over 15 mA on standby, which on a 980 mAh battery only gives just over two days stand by.

Small solar panels are readily available that are designed for mobile phones. These units are often more compact than those shown in Section 13.5.3 and fold in two making them light weight methods for powering the device. They charge both 3.6 V NiMh and Li-Ion phones. However, the units are often manufacturer specific and only suited to certain mobile phones. As such, if many electrical items are being taken into the field such as GPS receivers, PDAs and battery chargers then it is better to take a multi-purpose solar panel rather than a device specific one.

It is important to check that the phone used on the project is compatible with the cellular frequency of the host country visited for the GISci project. Even if the phone is compatible it will only work if your service provider has enabled the handset to talk to different cells. This has to be done before leaving for the field by contacting the service provider and requesting roaming to be enabled. There is usually no cost for this service but it must be switched on before leaving for the field. Enabling this after you have left is very difficult. Calls, however, can be very costly when phoning from abroad and bear in mind that you will also be charged a portion of the call cost for receiving calls to a handset while abroad. These costs should be investigated thoroughly before leaving for the field, as they can cause severe problems when returning home.

As with most electrical equipment, phones are equipped with LCD screens. The usual problems persist with these when used in inhospitable climates. The screen will cease to function if kept below 0ºC for any period. In addition, units are not generally waterproof though some more rugged units are available. Rugged units include the Nokia 5140. The Nokia 5140 is compatible with a clip on GPS attachment. It should be noted that the colour screen on most new phones is usually not suited to expedition work. Modern colour screens are very difficult to read in bright daylight and monochrome screens are often a better choice. The more additional features a phone uses the more it will drain the Li-Ion battery.

13.9 Photographic equipment and geo-tagging

Photographs add immense value to an expedition. A brief note is included here on their role in a GISci focused expedition. For inclusion in a GIS the images must be in a digital format. This could be film that is scanned into a computer or an image from a digital camera. Digital cameras are becoming more powerful and better suited to expedition work all the time and have many uses that an expedition might find valuable.

Digital cameras can be configured on the fly with the ability to change ISO settings, exposures etc. quickly between shots. Some of these settings would require complete film changes in standard film cameras especially where film speed need altering. Digital cameras also allow the photographer to view the image immediately to be certain the picture has come out. The author has been on an expedition where an entire six-week field project's photographs were rendered unusable because the ISO dial on the expedition camera had been knocked onto the wrong setting. All the developed pictures came out bleached white, an unfortunate event that would have been avoided with a digital camera.

A digital camera records data onto an array of CCDs. This is similar to the remote sensing platforms discussed in Chapter 5. Cameras are rated with a Mega Pixel (MP) value to describe their resolution. Basic entry-level models are rated at 2 MP but 3 MP make good all round cameras. Higher resolution 4-6 MP models make excellent cameras but may be too expensive with little appreciable gain for the expedition. Higher resolution models exist beyond 10 MP that are reserved for specialised professional applications. The expedition should be aware that a MP rating does not fully describe the quality of a camera. Other features such as lens quality and size of CCD etc. will also affect the final image. It is worth noting that CCDs are significantly less responsive to light than film. As a result digital images are good in bright light but require a powerful flash in dim conditions or appear 'noisy'. In addition, unless the camera has a large CCD array then resolutions above 3 MP can be very noisy because the same amount of light is being used on a higher number of CCDs. In this case, a 3 MP camera is often preferable for low light conditions. Table 13-12 shows the X and Y resolutions of these camera types.

Table 13-12 Resolution of typical camera types.

Camera	Resolution	File size (compressed)
2 MP	1600 x 1200	600 KB
3 MP	2048 x 1536	900 KB
5 MP	2560 x 1920	2 MB
7 MP	3072 x 2304	3 MB

The sizes of print shown in Table 13-12 are at high print quality and sizes can be increased with slight loss in clarity as discussed below and shown in Table 13-13. The file sizes quoted are for a camera employing JPEG compression, a feature all cameras employ. A 3 MP image stored as a TIFF or BMP would take up around 9 MB but saved as a JPEG it would only take around 1 MB. JPEG compression does reduce the quality of the print but the difference is barely detectable in all but the highest levels of JPEG compression. It is therefore generally not worth using a RAW or TIFF format even if the camera has one. Lowering the level of compression increases the print size slightly but vastly increase the file size. This trade off is rarely worth it because the team may find uncompressed images are too large for most purposes and take too long to backup and store on the web. Certainly, as far as presentation graphics on a website are concerned, images should rarely be more than 100 KB. The team should check all these factors before leaving for the field.

Larger sizes should be reserved for printouts while images on the web and in presentations should be smaller. Printing out high quality images can be difficult and time consuming so it is important to understand how the quality of output is affected by the image size. Printers are not rated in Pixels but in dots per inch (DPI). A DPI rating for a good photographic quality inkjet printer might be 4800 x 1200 DPI. Because a printer cannot print every possible shade of colour that can appear in a photograph (16.3 Million), it improvises this colour by mixing dots of coloured ink from its printer heads. Typically the colours magenta, cyan and yellow are used for the primary colours, with a second head supplying either black or a three colour mix for flesh tones including pink, pale blue and black. The actual DPI rating of a printer is the quoted resolution divided by the number of colours in the colour head. In the case of most printers this would be (4800 / 3) x (1200 / 3) = 1600 x 400. The required DPI to fool the human eye is considered to be 300 x 300. So modern printers easily achieve this but older printers may struggle. There are also many issues regarding ink quality and paper quality that influence the final product. To equate the MP rating of cameras to a final image size on paper a factor of 300 DPI should be used. Unfortunately, because of JPEG compression the actual acceptable output size is slightly smaller and Table 13-13 summarises the resultant outputs.

Table 13-13 Output sizes of typical cameras.

Camera	Output size (perfect)	Output paper size (perfect)	Acceptable size
2 MP	4” x 3”	small photograph	large photograph
3 MP	6” x 4”	standard photograph	large photograph - A4 page
5 MP	7” x 5”	large photograph	> A4 page
7 MP	8” x 6”	large photograph	< A3 page

All data for a GISci project needs some form of co-ordinates. Typically for photography the co-ordinates will be transcribed in a field notebook. More recently, the text strings from a GPS memory have been used to reference where a photograph was taken as shown in Section 13.3.2. These text strings can encode photograph numbers but they are cumbersome and better methods exist for encoding co-ordinates. Some modern cameras are entering the market with connections for a GPS or soon with GPS modules. The method these modules use is to encode the co-ordinates in the embedded header file of the digital photograph. The digital pictures, even when stored as JPEGs, have data referred to as an EXIF format embedded in them. The EXIF format has a string reserved for co-ordinates and the GPS modules populate them. Using a GPS module is an excellent way of attaching co-ordinates to a picture but it is expensive and there are cheaper methods. All digital cameras record the time the image was taken. Similarly, GPS receivers record the time that a waypoint was taken (it should confirm that either the trackpoint or waypoint downloads time from the expedition receiver and although this is common, it should not be assumed). Software can be used to match the GPS co-ordinate to the photograph using matching times. This can then assign the photographs a co-ordinate to relate them into the GPS. One software package for this method is Robo-Geo found at *www.robogeo.com/home*.

13.10 Selecting the expedition hardware

In advising expeditions on the hardware required, it is often difficult and unhelpful to talk in detail about requirements without mentioning manufacturers, models and suppliers. This discussion should, however, not be taken as a complete summary of every product on the market and where a company name is mentioned it should be assumed that other similar companies exist that supply equally capable systems. In all cases, research should be conducted to look into the current situation before purchasing any equipment. This information is only accurate at the time of writing (2004-2005).

Hardware must be purchased that will survive in the climate and conditions that the expedition is visiting. For most expeditions, cost will be a significant factor in selecting project kit. However, if mission critical expedition equipment fails in the field the cost to the team could be substantial. A device failure may mean vital data cannot be collected and may even cause the project to fail or have to be temporarily abandoned. Insuring against this by using the correct equipment is a valuable and cost effective strategy. Field hardened equipment is very expensive but worth the cost. It is better to try to obtain this equipment second-hand or through a sponsor, such as a University, than to use hardware that is prone to failure.

If an inhospitable climate is a significant consideration, either with temperature extremes or with issues of water or dust near the computer, then hardened laptops are required. For extreme temperatures Itronix laptops can operate between –20 °C to 60 °C and can be extended to work at lower temperatures down to ~-30°C with additional battery save hardware. Itronix PCs significantly lag behind standard PCs and laptops in processing power but are becoming more powerful all the time. There are several laptops available from Itronix, older models have included the Itronix GoBook Pro and GoBook II and their most up to date model available in late 2004 is the Go Book III. The Go Book Pro featured a 700 Mhz PIII-M processor, 256 MB RAM and 20 GB IBM hard disk. The GoBookII featured a 12.1-inch screen and a faster processor. The current model meets and exceeds the MIL-STD 810F standard and comes equipped with a 1.8 Ghz processor and up to 2 Gb of RAM. The screen is touch sensitive and uses ‘Outdoor Viewable Transmissive’ to make it more readable in direct sunlight (a major problem with most laptops). These specifications will continue to become obsolete but field hardened laptops will typically stay 18 months behind the most powerful laptops available. These Itronix PCs are very strong and will survive metre drops onto concrete and have IP54 ratings. The drawback of these models is the price, which is around £4,500 for the current models. Older Itronix kit features solid-state storage but both GoBooks use a single spindle drive encased in shockproof gels at the heart of the PC. More details can be found at *www.gobookmax.co.uk*. The current models use 5,400mAh lithium ion batteries with inbuilt battery conservation software that allows the CPU to run at 25, 50, 75 and 100 percent power. On its lowest setting, the GoBook lasts for up to 3 hours 20 minutes of continual heavy usage or 5 hours at lower usage settings. The Itronixs PCs also now have wireless 802.11b and Bluetooth for increased connectivity options while in the field as well as a GSM and GPRS card for field communications.

The Panasonic range of Toughbooks has a long pedigree in the area of hardened laptops but more recent models are not as sturdy as some of the older models. They are, however, significantly faster than other hardened laptops with speed ratings that are more inline with standard laptop PCs. The modern Panasonic range is also much cheaper and depending on the model retail for between £1500 - £2000. A recent test in Computer Shopper (November 2003) reviews a range of hardened PCs and the reader can find more information on different manufacturer's equipment from this review (see *www.computershopper.co.uk*).

Figure 13-9 Field PC and wireless GPS unit. The GPS shown here has no LCD screen and can communicate wirelessly via blue-tooth. The advantage of this is that the GPS unit can be used in conditions below 0°C. The Panasonic CF-27 can also be used in inhospitable conditions without risk of damage.

When selecting a PC it is vital to make sure it is compatible with the equipment you will be taking into the field. If the PC is required to run from a car it must have an inverter to take the 12V DC and change it to a power supply compatible with the laptop. If it is to be used with a GPS unit it must be compatible with the connector on a GPS, usually an RS232 connector.

While all modern laptops feature similar battery save settings most are generally unsuitable for fieldwork due to either their fragile nature or the power overheads from most high mega-hertz chipsets. Lowering power consumption helps, using techniques such as the GoBook's speed reduction or Intel's Speedstep, but it is unrealistic to budget for over 2 to 2½ hours usage from one battery on a standard PC. This combined with the expense of new laptops means older models are generally more suited to expedition work.

In general, slower processors will draw less current from the battery. If battery life is a significant consideration then the expedition should look for products powered by lower consuming processors such as those using the Transmeta Crusoe or Intel Centrino chips. Lower specification machines may last between four and six hours. The IP rating of these machines should be checked because they are not designed for use in inhospitable regions. In a hot dry environment a low IP rating for moisture might not be a major concern. Try to match the kit you obtain to the conditions of your expedition.

GPS receivers must also be targeted to the conditions on the expedition. If the expedition will be under tree cover an external antenna is vital. The capabilities of this antenna should always be tested in a local forested environment before leaving for the field. If the conditions will be cold (around 0°C) then a receiver such as the Silva Multi Navigator should be considered. If the conditions are below this then special care needs to be taken. A GPS without an LCD screen is often preferable with a link to a device that can be kept sheltered and protected. Three common models are shown below in Figure 13-10. A better method still, is to keep the GPS entirely sheltered and only use an external antenna. This would take the GPS down well below zero.

Figure 13-10 GPS Receivers: Basic ETREX, Silva MulitNavigator, AnyComm GP600. The ETREX has an operating range of 50℃ to 0℃, The Silva 50° to -15℃ and the AnyComm 60℃ to -20℃. For temperatures beyond this a GPS using an external antenna is the best solution but that requires the unit to be stowed somewhere protected from the elements and generally negates the use of waypoints.

The Bogda Shan Expedition used many of the technologies listed above. Because vehicles could not be driven to the basecamp a deep-cycle lead-acid battery was carried to the camp. This battery provided power for the entire time in the field. The GPS sets performed adequately in all conditions and the use of the field hardened Panasonic Toughbook meant that data was less vulnerable to loss or corruption. The GPS sets all performed well, but for the high altitude mountaineering sections, the ETREX was the preferred device. The small size and internal antenna make it the easiest to handle. The casing is also ruggedised giving good shock resistance and waterproofing. These factors help make new models like the Garmin ETREX range or Magellan meridian series the most suitable devices for difficult terrain.

13.11 Purchasing equipment

As explained in Section 13.10 there are many alternatives to buying expedition equipment, including borrowing from a university department or renting equipment. In some cases, purchasing equipment it the only valid option. In this case, the expedition will no doubt have to put a significant effort into obtaining equipment as cheaply as possible. This section briefly looks at some methods for reducing the cost of expedition hardware.

A complete set of GPS receivers for all members, combined with field hardened computers, professional photography equipment and PDAs for field recording is an ideal scenario but is made impossible by the immense cost (several thousand pounds). Instead, there are some methods that can be used to reduce the cost. The expedition laptop might be the most expensive part of the equipment required. Computers get faster and more capable all the time; however, there is not a need to use the fastest PC when in the field. The Bogda Shan expedition used a Panasonic CF-25 laptop running at around 200 MHz (see Section 13.2 and Figure 13-9). Though this was not ideal in all case it was acceptable for most uses. Recently CF-27 have been appearing as refurbished laptops in some mail order catalogue listing for around £200 GBP + VAT (SterlingXS June 2004). These are excellent computers for field use even though they have limited processing power by today's standard. The CF-27 is about 4 years old but would have cost £3,500 in mid 2000. The CF-27 runs an Intel 266 MHz MMX processor with 168 MB Ram and a 4 GB hard drive. These computers make excellent tools for expedition use. Second hand machines can often be picked up on auction sites from the Internet such as EBAY (*www.ebay.com*). These sites are useful but bear in mind that Li-Ion batteries have a very limited life so do your best to check before purchasing that the batteries still work.

A very tempting plan for obtaining equipment is to purchase it over the Internet from a non-UK company. Field equipment in the US often costs a similar dollar value to the sterling value, i.e. a £1000 PC might cost around $1000 in the US. Depending on the current exchange rate this can look very tempting. At the time of writing the exchange rate is 1.9059 dollars to the pound. This appears to nearly half the cost of equipment. However, there are some important considerations. Equipment brought into the UK is subject to three possible taxes: customs import duty, excise tax and value added tax (VAT). This coupled with the added expense of shipping across the Atlantic means the equipment is not as good value as might have been thought. Import duty is a complicated issue. Import duties are variable and range from 0% through to over 10%. For example, a PDA has no import tax but a PDA with calculator functions incurs a 3.7% tax. The expedition can check possible import taxes by visiting *www.hmce.gov.uk*. To determine the actual level of tax requires an 8-digit TARIC code. This can be found on the website and allows a full quote to be calculated. For expedition surveying equipment the code takes the form 90 – 15 – XX – XX and general taxes are around 3.7%. This duty is imposed on all goods over £7. Excise duty is not applicable for expedition imports and applies to alcohol, tobacco and some other closely monitored imports. VAT is calculated at 17.5% on top of the cost of the goods + shipping + import duty and is charged on goods over £18. As an example, an ETREX Venture GPS retails in the UK for around £114.38 but only $124.95 in the US. This appears at first glace to convert to a price of $124.95 / 1.9059 (current exchange rate at time of writing) = £65.56. The actual difference in price is described in Table 13-14.

As can be see in Table 13-14 there is still a benefit from importing goods but the time delays in obtaining them (some companies wait an additional five days before dispatching international orders) and the problems in returning goods may make the process unsuitable. Some retailers may be difficult to contact in the case of the goods needing to be returned and UK laws will be of little help if the equipment is not fit for its purpose. US goods also come with US power adaptors; converters will need to be bought to run them in the UK.

Table 13-14 Import taxes on imported GPS equipment.

Item	**cost from UK**	**cost from US ($ / £)**
Device	£ 114.38	$ 124.95 = £ 65.56
Shipping	£ 5.00	$ 30.00 = £ 15.74
Subtotal	£ 119.38	$ 154.95 = £ 81.30
Import Duty @ 3.7%		£ 3.01
Excise (N/A)		£ 0.00
VAT		£ 14.76
Total	£ 119.38	£ 99.06

An important note in a discussion on importing GISci equipment is the need to declare the equipment correctly to customs. There is a possible temptation to use a slightly different code to avoid the import duty. This is illegal and should not be done. Some exporters may state on their website that goods will be labelled to pass favourably through UK customs (either through a change to the goods description or by undervaluing the goods). You should be aware that as the designated importer, under British law you are responsible for the labelling and declaring of goods even if all this is handled by the shipper in the US. If a declaration is found to be incorrect you may be liable to financial penalties and criminal prosecution and the goods can be forfeited. It is essential any imports are declared accurately.

13.12 Disaster recovery planning

13.12.1 Data backups

All systems are prone to failure and a recovery plan is essential in any field operation. The first and most critical factor in data-disaster recovery planning and mitigation is to keep a backup of your data. A backup should consist of all the files in the GIS and be kept separate from the hard drive storing the GIS. Ideally, it should be located away from the basecamp. Many of the data storage devices discussed below are listed in Section 13.3. A rough price guide can also be found in Table 13-3.

A simple backup procedure is to use a floppy disk to copy key data from the host PC to a temporary storage medium. The disadvantage of a floppy disk is its limited storage space (<1.44 MB) and its susceptibility to corruption. Floppy disks are not robust and should not be recommended for use in the field. Alternative storage medium include recordable and re-writable CDs (CD-R and CD-RW). The advantage of CD based technology is that they can store between 550 and 700 MB depending on the type used. A CD is generally more robust than a floppy disk but the disadvantage is that they require a more specialised drive. Also if the disk is not 'closed' after a recording phase a standard CD-ROM will not be able to read the data on the disk.

The storage space for a GIS is often actually quite small. The point and shape data files are often only a few KB in size. The Bogda Shan Expedition discussed in Section 17.1, generated very large quantities of GPS data. However, the total size of the database was only 988 KB. This would quite easily fit onto a floppy disk. The image files were much larger at just over 1 GB. Image files could easily be backed up before the expedition so making additional backups as the expedition progresses is commonly not necessary. Other storage mediums include DVD RAM, DVDR and DVD Recordable, all of which are different standards of DVD recording. These allow about 4.7 GB of storage onto a single sided, single layered disk.

Other types of data storage include USB key ring datasticks. These very small devices clip into the USB slot of a PC and act as a removable hard disk. When used on machines running Windows 98 SE, Windows ME, Windows 2000, Windows XP, or any Apple Mac using OS9 or OS X they are auto detected and require no software. The sticks come in sizes from 16 MB up to 2 GB. Larger sticks are becoming available all the time. They form a robust and convenient way of storing data. USB devices are powered through the USB port by the computer they are attached to. They commonly draw a 94 milli-amp charge from the battery.

The problem with floppy disks, CDs, DVDs and memory sticks is that they are still generally storing data locally. A much better method is to use a none-local storage medium. As described in Section 13.3.5, the best none-local storage area in general, is the Internet. If an area of the Internet is prepared in advance then data can be uploaded and protected. The most common way of doing this is to use FTP (File Transfer Protocol). FTP is built into Windows by default, and the common programme used is Telnet. More user-friendly programmes can be downloaded. One example is Cute FTP. The main disadvantage of using non-local storage is that it requires connection to the Internet, which requires some form of telecommunication. The most common form would be a phone line. In the field, where this is not available, a mobile phone may suffice or an Internet café might be located in a nearby town.

Different types of connections are discussed in Section 13.7. A basic field configuration for lightweight FTP can be seen below in Figure 13-101. If the party intends to use FTP on a Pocket PC they should ensure the device has an FTP client. This is not always available as default and software may need to be downloaded or purchased to make it operational. There are a number of free FTP clients, including Robust FTP 3.0, available from many of the download sites such as *www.pocketpccity.com*.

Figure 13-11 A possible FTP / HTTP field configuration using small apparatus.

There are many cheap file stores available and not all can be listed here. One company of note is the US online storage specialist Streamload (*www.streamload.com*), which will allocate 10 GB of free storage space to a user. This can be accessed when the expedition wants but with a monthly limit of 100 MB of transactions. Alternatively, unlimited storage with a 1 GB transaction limit starts at $4.95 a month. If such large allocations are not required, the team should check whether they have free webspace from an Internet Service Provider (ISP). ISPs often allocate between 5 and 30 MB to a user for personal webpages. This might be an acceptable size for storing GPS data and field notes. All FTP links require an upload via a phone line. Some Internet cafés might not support FTP and the cost for dialling into an ISP from abroad may be prohibitive. In these cases, it might be easier to look at using an email account. Email accounts can usually be accessed from anywhere, especially ones with Internet portals. A good example of free email through the Internet is Hotmail. Hotmail offers free email storage of 250 MB or a paid for service of 2 GB for £14.95 a year. Hotmail cannot be used as a standard FTP site but users may feel more comfortable emailing data to a secure location rather than relying on FTP. There is a 30 MB per email limit on Hotmail’s paid account but there is no limit on the number of emails that can be received in a month. There are many other similar email services to Hotmail. These include Google’s GMail (not publicly available at time of going to press) offering 1 GB and Lycos’ email services offering between 1 and 10 GB of space. These can be found at *gmail.google.com* and *www.lycos.co.uk* respectively.

13.12.2 Data recovery

If data on a disk becomes corrupted then there are some basic techniques that can be used to salvage information. A hard disk and floppy disk work in similar ways. The sectors on the magnetic disk store data. If these are corrupted, it is often possible to move the data to recover most of it.

The first step is to scan the disk for errors. Scanning the disk can often show that there are problems. Microsoft has a scandisk tool built into all versions of Windows and DOS (typically accessed as: My Computer ⇒ C:\ ⇒ Properties ⇒ Tools ⇒ Check Now).

The recovery option will attempt to take the data on those sectors and save them to a text file. Though this may result in a loss of the exact data type it should allow the basic information to be put back together. Another effective method, especially with floppy disks, is a disk defragment. This too is included in Windows and DOS (My Computer ⇒ C:\ ⇒ Properties ⇒ Tools ⇒ Defragment). Defragmenting moves the information on a disk into continuous sectors; the intention is to speed up disk access times by making the information contiguous. However, it also has the effect of moving information away from corrupted sectors. This can be an effective method for recovering data from a floppy disk. If a hard disk is very badly damaged by water or fire, specialist companies can normally still recover the data. These services can be expensive but save the expedition having to re-do work to recover lost information.

13.12.3 System recovery / repairing catastrophic system failures in the field

Sometimes data cannot be recovered form a hard disk. If the disk is in a stable and undamaged state (i.e. there are no bad sectors detected by scandisk or the number of bad sectors is not increasing with time) then a restore should be possible. It is always advisable to take a copy of the Operating System disk with you, as well as a boot disk containing basic boot files such as command.com, config.sys, and autoexec.bat. Ensure these files contain generic CD drivers such as atapi to get the PC started. The bootdisk should also include basic DOS programmes such as scandisk, format and fdisk. If the hard disk is damaged and contains bad sectors then replacing it is preferable. When replacing the hard disk, not just the GISci data but all operating system files will be lost requiring the system to be re-loaded. This is also true if the field PC has to be completely replaced by another unit.

A complete restore is only possible if all the GISci data has been taken out into the field on disks. This will require creating disks for all the imagery etc. before venturing into the field. It is also advisable to have all manuals either printed in hardcopy or stored as documents (PDF) to help get the various parts of the GIS project up and running. There are a number of standard programmes that should be taken into the field to help you with this task. The Operating System disks are essential; even if a restore is not needed, you may require additional programmes/drivers from the original disks. Other software of vital importance is a copy of an Office Suite (such as Microsoft Office/Works, Star Office etc.), Adobe Acrobat Reader, WinZip and a FTP client. There may be other software specific to your expedition that you should carry with you. These additional programmes will include the disks for the GIS software and any supporting software such as GPS Utility. These will be essential to ensure that data can be restored or for the field PC to be replaced. A simple method for restoring the system is to use a utility such as Norton Ghost to ghost the main disk of the computer. Ghosting creates an exact copy including the boot record. If a ghost is taken at regular intervals it should be possible to very quickly get the expedition PC back to a workable state.

When using a PDA in the field it is possible to install some of the programs to the removable media cards. After a battery discharge the unit will hard reset back to its factory standard state. This is a state where only the operating system and key operating system applications are installed. To re-install the additional expedition software such as ArcPad or GPS Tuner requires the PDA to be reconnected to a computer and have the files transferred back across. To mitigate this some software can be installed to a CF or SD card. These cards retain their memory after a battery discharge. This is one method for ensuring the PDA can be brought back up to speed as quickly as possible in the case of a system failure.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section D: Planning & Practicalities

Chapter 14: GISci Software

14 GISci Software

Geographical Information Sciences (GISci) encompass several disciplines, from surveying to database manipulation. Each discipline uses its own systems and software. Usually the software that links all these together is the Geographical Information System (GIS). There are several different GIS software products on the market and there are a number of different pieces of software that can be used for GISci data collection and data analysis.

In the space available here, only a small number of products can be discussed. Preference has been given in these discussions to software used in the RGS-IBG Expedition Advisory Centre's Mapping Unit teaching exercises. This is not to suggest that these tools are the best on the market or designed to exclude any other suppliers, but the authors have found that they work well for expedition tasks. Internet addresses of major GISci software suppliers are provided in the Appendix. Bear in mind too that GIS magazines, such as *Geo-Informatics* (www.geoinformatics.com) often have useful reviews of the latest software.

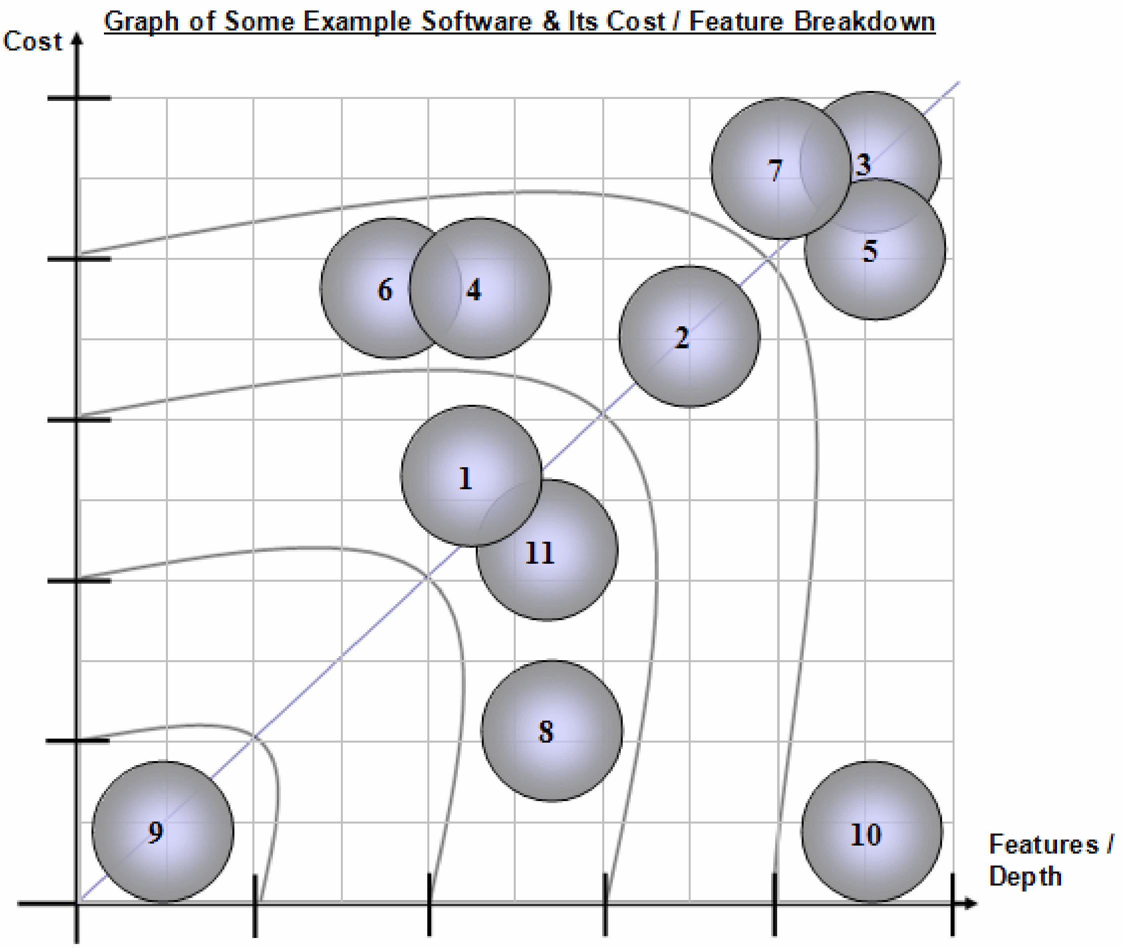
The software discussed in this chapter cover the core needs of an expedition. There is an immense range of specialist, often more costly, software available for use in GISci, which might be needed for particular expedition needs. Generally, however, it has been found that the majority of typical expedition GIS needs can be met with a relatively small set of software combined with ingenuity.

The core subjects covered are GPS data, data manipulation, remote sensing analysis and GIS functions. Many GIS and image processing packages tend to be very expensive, so there is a section on low cost alternatives. If you are based in, or have links with, a university check whether a low-cost student license or university-usage arrangement exists. These include the UK CHEST agreement for educational usage of software (*www.chest.ac.uk*). For example, ESRI's ArcGIS retails for many thousands of pounds but via CHEST it is available for about £110 in universities that have agreements. If your association with the educational institution is sufficient the software may even be provided free of charge.

14.1 General principles in software selection

In general, the more features a software product has the more it will cost. Also the more specialised and technical a product is the more it will cost. Logically, a simple product with a single basic function would be cheap (examples include GPS Utility); whereas a product with a full suite of functions such as ArcGIS would tend to be expensive. However the expedition may require software with highly specialised and intricate functions, which may be very expensive (examples include various photogrammetry packages).

The basic relationships of cost and features are shown in Figure 14-1. Any software that plots below the central line generally represents good value for money. Software that plots above the line does not appear to be good value.



Product		Type	Features / Depth	Cost	*Ease of Use Support*
1	ArcView	GIS	Medium	Medium	*High*
2	ArcGIS	GIS	High	High	*Medium*
3	TNT MIPS	GIS + Image Processing	V. High	V. High	*Medium*
4	ERDAS	Image Processing	Medium	High	*Low*
5	ERDAS (Ortho Base)	Image Processing	V. High	V. High	*Low*
6	ER Mapper	Image Processing	Medium	High	*Medium*
7	Socet Set	Photogrammetry	V. High	V. High	*Low*
8	Manifold	GIS + Image Processing	Medium	Low	*Medium*
9	GPS Utility (GPSU)	GPS Download / Upload	Low	Low	*High*
10	GRASS	GIS + Image Processing	High	V. Low	*V. Low*
11	IDRISI	GIS + Image Processing	Medium	Medium	*Medium*

Software that plots on the centre line or below it can be considered good value. Software above the line is generally offers less value. Software such as GRASS (10) shows very good value for money. However, because of the difficulty in using GRASS it may not be applicable for every expedition

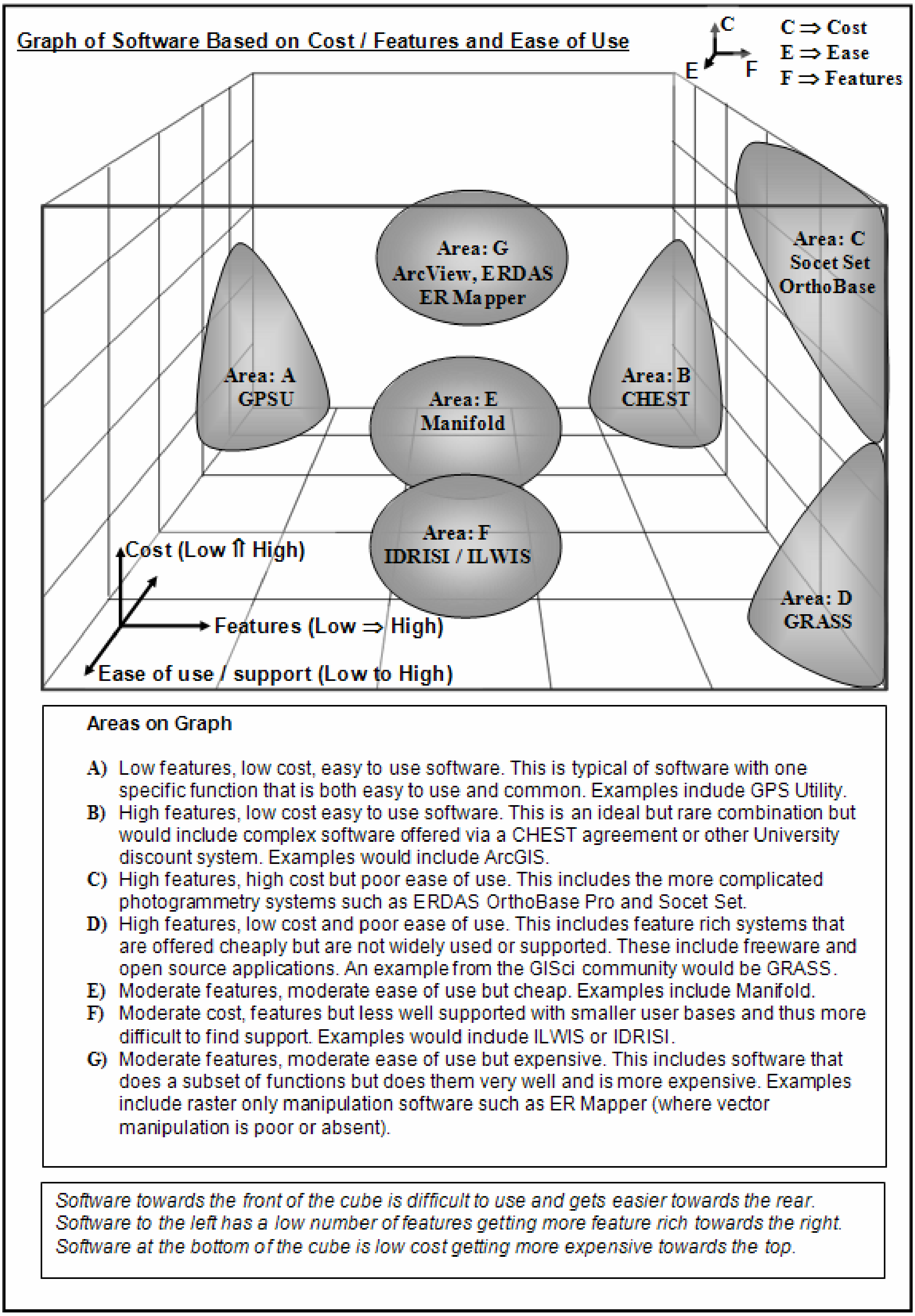
Figure 14-1 The relationship between features & cost of GISci software.

This relationship is only the most basic expression of the value of software to an expedition. Software that is very inaccessible to an expedition or requires considerable time and effort to learn may not be suitable. As many team members as possible should be capable of manipulating the software. It is not acceptable to run an expedition with a single user proficient in the use of the software. Therefore, software selection should take into account the usability of the program. As an example an area where this principal has been used and refined is in industry.

In industry, software is selected on two basic foundations; initial cost and total cost of ownership. Initial cost represents the cost to purchase software. Total cost of ownership (TCO) reflects the cost to train staff and implement and run the software over a long period of time. TCO is often much more costly than the upfront payment for the program. TCO includes possible lost productivity due to software downtime and the man-hours involved in setting up the system. An expedition should be run with a similar thought process. The significant difference between industry and expedition software requirements is the time available. Though it may be a clichéd statement, in industry 'time is money'. The time taken to set up software and the delay in getting a return on the investment is a major cause for concern. Because staff time is expensive and delays can cause problems in a business environment, it is generally more practical for a business to invest more heavily in the initial cost than suffer increased overall TCO. In an expedition where money is usually very tight but timeframes are longer, personnel time is unlikely to be as great a contributor to cost of the project as it is in industry, where staff on training courses still have to be paid a salary. An expedition can spend longer deciphering software in advance and spend time that would be un-economical in a standard business model, however, the expedition should be careful when employing this methodology.

Industry selects software that is accessible and quick to deploy and reap benefits from. This explains the continued proliferation of Windows software in a market that would seem to favour a LINUX deployment. In an expedition environment it is possible to stray from the standard model to a degree. However, the expedition should not lose sight of its primary goal which should never be software led. The expedition will have many goals and objectives and learning software should not be one of them. For that reason very difficult to learn software should not be selected purely based on cost.

Figure 14-2 shows this relationship in a basic 3-axis cube. In this diagram software such as GRASS, which plots very favourably in Figure 14-1, no longer looks as attractive. Software in Figure 14-2 that plots towards the front of the cube is inaccessible. Software that plots towards the right of the cube has many features and a complexity to its functions. Software that plots towards the top of the cube is expensive. Many expeditions spend an inordinate amount of time attempting to obtain software either for free or for as little as possible. This approach often fails to work and the expedition is left without the means to fulfil its goals. Though expeditions have time to learn and obtain software; the business model should never be forgotten. Often it is worth investing some money in software to alleviate the problems of complexity or unsupportable systems.



Areas on Graph

A) Low features, low cost, easy to use software. This is typical of software with one specific function that is both easy to use and common. Examples include GPS Utility.
B) High features, low cost easy to use software. This is an ideal but rare combination but would include complex software offered via a CHEST agreement or other University discount system. Examples would include ArcGIS.
C) High features, high cost but poor ease of use. This includes the more complicated photogrammetry systems such as ERDAS OrthoBase Pro and Socet Set.
D) High features, low cost and poor ease of use. This includes feature rich systems that are offered cheaply but are not widely used or supported. These include freeware and open source applications. An example from the GISci community would be GRASS.
E) Moderate features, moderate ease of use but cheap. Examples include Manifold.
F) Moderate cost, features but less well supported with smaller user bases and thus more difficult to find support. Examples would include ILWIS or IDRISI.
G) Moderate features, moderate ease of use but expensive. This includes software that does a subset of functions but does them very well and is more expensive. Examples include raster only manipulation software such as ER Mapper (where vector manipulation is poor or absent).

Software towards the front of the cube is difficult to use and gets easier towards the rear.
Software to the left has a low number of features getting more feature rich towards the right.
Software at the bottom of the cube is low cost getting more expensive towards the top.

Figure 14-2 Relationship between cost, features and ease of use of GISci software.

Before discussing GIS software, this chapter will describe software for collecting GISci data and formatting it for inclusion in GIS.

14.2 Software for GPS data retrieval

When the GPS field data have been collected, it is a good precaution to make back-up copies. Back-up by hand is usually unfeasible, so a digital connection to the GPS or surveying equipment is the standard method. To download GPS co-ordinates into any computer for use in a GIS requires a cable to connect the two together. This usually consists of a 9 pin serial cable (RS232) on the PC side and a special connection on the GPS. It is wise to ensure that the expedition laptop, if there is one, has a serial port because many newer laptops no longer have these ports as standard. The following websites offer information on connecting to various GPS models.

- Wolfe's GPS cables: *www.gpscables.com*
- GPS-PC connectors: *pfranc.com*
- Joe Mehaffey's GPS information page: *www.gpsinformation.net*

If a laptop computer is available in the field then GPS data can be accessed directly. A very cheap way to do this is to use software bundled with Microsoft Windows, such as HyperTerminal. Though this is probably the cheapest way for an expedition to collect data, it is inefficient and requires a constant connection to a PC. It may, however, be a good solution if the expedition is spending a long time travelling and the PC can be powered from the vehicle adapter. A description of how to use HyperTerminal for GPS connections is included in the appendix.

There are several products on the market to connect to a GPS and download stored data. These range from freeware utilities, easy to code software solutions, inexpensive software and software plug-ins for major software. For expedition use, this chapter will concentrate on inexpensive software designed to upload and download GPS waypoints and trackpoints. Two of the most popular download programmes are OziExplorer and GPS Utility. Both of these are for the PC but a popular Mac alternative is GPSy. GPSy has a large user base and is popular. A standard user license of GPSy retails for US $50 but student licenses can be obtained for US $30. More information can be found at *www.gpsy.com*. GPS Utility is used in many of the examples in this manual and is the software used in the RGS-IBG training weekends. The most recent version released on 20 December 2004 is Version 4.15. This is a 1.57Mb download from *www.gpsu.co.uk*. This version is a 32-bit application for Win 9x, NT and XP. There is also a 16-bit version called GPS Utility-Lite (GPSU-LT) for use on Windows 3.1.

OziExplorer is a more in-depth package with several modules for use on laptops, PDAs and for generating 3D terrains. The basic package costs $85.00 US (+VAT). OziExplorer is a 4.7 Mb download from *www.oziexplorer.com*. The OziExplorer3D add-on to the OziExplorer software costs US$30.00 (+VAT) while the handheld version, OziExplorerCE retails for US $30. OziExplorerCE is an add-on to the major product. To get the full use of OziExplorerCE you must have a full purchased copy of OziExplorerCE and a full purchased copy of PC OziExplorer. OziExplorerCE runs on PocketPC and Windows CE PDA's but relies on map calibrations from the PC version.

Though OziExplorer is a more advanced software package than GPS Utility, it is often better to select packages that perform single operations well, rather than those that attempt to span several functions. All of OziExplorer's functions can be replicated better in the expedition's main GIS. If, however, a main GIS is not being used then a product like OziExplorer is an excellent choice. Selecting GPS integration software depends on the larger picture of all the expedition's software. If a fully featured GIS is used then select cheap and simple GPS software. If a full GIS is not available select a GPS tool with added functionality.

14.3 Software for terrain visualisation

Many GIS packages come with advanced 3D terrain visualisation modules. Some suppliers will charge extra for this and the cost might put it beyond many expeditions. In these cases there are an increasing number of free or cheap packages that can work in partnership with the GIS to produce 3D views.

3DEM is a freeware programme for visualising 3 dimensional datasets such as NASA's SRTM data. The most recent version of the software, 18.9, can be downloaded from *www.visualizationsoftware.com*. The software can even take in the raw unpatched SRTM and fill the holes. The download size is 5.8 Mb. GPS data can be used on the maps and fly-throughs can be created. 3DEM uses the OpenGL libraries for high speed 3D rendering and will render 24 bit colour three-dimensional projections. These can be displayed either in perspective views, red-blue projections requiring red-blue 3D glasses, or colour 3D projections requiring Liquid Crystal Shutter (LCS) electronic shutter glasses for viewing. Another 3D visualisation package is Manifold 3D View Studio. Manifold is a full-blown GIS software package reviewed in Section 14.5. Though modern versions (5+) have 3D capabilities, earlier versions lacked this capability. To rectify this a separate module was developed and released as a standalone product. The capabilities of 3D View are beyond those of 3DEM but the software carries a US$75 price tag. The one advantage of 3D View over 3DEM beyond the advanced textures and animations is that 3D View can plot none surface data. This means it can be used as a 3D graphing tool for any dataset.

14.4 Remote sensing / image processing software

There are many software packages that can handle image processing. Anything from a basic image editor, such as the free Microsoft PhotoEditor, can enhance brightness, contrast and gamma but dedicated remote sensing software is the only real solution for scientific analysis. If cost is a real consideration and no access is available to high-end packages, then any professional graphics-processing package that allows histogram manipulation is an acceptable, though not perfect, option. Examples of this may include Corel Photo-Paint or PaintShop Pro.

There are two or three problems with this form of image processing. The first is that the image formats supported by these packages are different to the standard GIS and remote sensing standards. Though TIF files are supported in most professional applications, GIS sometimes require more complex data types, often including meta-data. There are also more efficient data storage solutions that are not supported, such as ERMapper's .ECW and .ERS formats and MrSid compression routines. Many GIS packages will only work with .BIL and other binary formats not supported in graphics processing suites. The second problem is that some of these graphics packages are not designed to deal with very large images. A programme such as ERMapper or ERDAS Imagine could open a 1 Gb image almost instantly, a task that would render a PC running a 'Paint' application unresponsive for a significant period of time. This is because remote sensing packages are designed to deal with large data sets by building 'pyramids'. Pyramids are views of an image at various scales and help a programme switch views very efficiently. They also segment large image files into subsets which are indexed, allowing rapid transit to the pertinent section of files, an approach that cheap or free graphics software does not use.

Finally, even though the professional drawing and image editing software can perform contrast stretches, these processes are relatively basic. To undertake GISci work, a package capable of various specialised processes is required. These processes include proper colour composite production, calculating band ratios, and image classification (these techniques are described in detail in Chapter 8). A summary of the relative costs and image processing capabilities of various software packages is given in Table 14-1.

Table 14-1 Functionality and costs of image processing software. Shading indicates cost ranges of software: unshaded = <US$100; pale grey = $1,000 - 3,000; dark grey = $3,000 - 15,000. Ticks indicate the presence of a given function; tick size indicates the ease of use and effectiveness of that function.

	Function										
software	import export	screen digitise	contrast stretch	spatial filters	composite images	band ratios	PCA	unsupervised classification	supervised classification	spatial analysis	change analysis
CorelDraw	✓	✓	✓	✓							
Paintshop Pro	✓	✓	✓	✓							
ERDAS Viewer	✔	✓	✓	✓						✔	
DMAP	✓	✓	✓	✓						✔	
ArcGIS	✔	✔	✓	✓	✓	✓				✔	✔
MapInfo	✔	✔	✓	✓	✓	✓				✔	✔
Idrisi	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ILWIS	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ERDAS Imagine	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ER Mapper	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
TNT MIPS	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
PCI Geomatics	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ENVI	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔

14.4.1 ERDAS Imagine 8.6

ERDAS Imagine is a very popular image-processing suite, which is becoming increasingly more integrated with ESRI's ArcGIS software, as they now have the same parent company, Leica Geosystems. ERDAS Imagine is a major competitor to ERMapper but offers more functionality through a series of additional programmes and plugins. Imagine has plugins for DEM generation whereas ERMapper can simply display 3D images. The package required for 3D viewing is VGIS. Imagine is expensive, but is available at much lower cost to universities through the CHEST agreement. There are many different licenses for ERDAS Imagine, from a basic package called Essentials, up to the more complete Professional version. A further very expensive plugin for Imagine, called OrthoBase Pro, enables full photogrammetric functionality, including DEM generation from aerial photography and suitable remotely sensed datasets.

14.4.2 ER Mapper 6.4

This is a popular piece of software for image processing, produced by Earth Resource Mapping. It became popular in the mid 1990s due to its revolutionary storage system. Unlike other image processing packages of the time, ERMapper stored ‘algorithms’, which were very small text files that told the programme how to process the original image. By doing this, it meant that the system did not have to store a second re-processed image and stored only the original and a series of small text files to re-create large images. By contrast, competitors’ software would need to store the original image as well as a number of copies all at the same size. This made saving very slow and back when hard drive space was at a premium it meant that very few images could be created. When considering the smaller size of hard drives on expedition PCs (see Chapter 13) this method may be an advantage. The main disadvantage to this approach is that a complex process may have to be run every time an image is called up, making image display slow, especially on low specification processor PC’s. ER Mapper also has a proprietary, very powerful wavelet compression algorithm that can compress images very rapidly with an acceptable level of detail loss, assuming manual interpretation is the only use to which the image is put. This format is known as ECW. Older versions of ArcView support .ECW as does Manifold. ERMapper also has a sophisticated webserver for displaying very large images across the Internet. The software is very user-friendly and CHEST licences exist through universities.

14.5 GIS software

When selecting GIS software, the most important consideration for an expedition will be the cost of the package. With professional GIS solutions costing up to £5,000 and image processing suites at around the same price, an expedition may find affording these solutions very difficult. There are a number of ways of mitigating these costs, such as academic licenses and the UK CHEST agreement. However, it is important to understand some of the costs that may normally be overlooked in expedition planning. The cost of purchasing software is only one cost. There are time and money costs in learning the software and time costs in uploading data. Due to the high cost of time in the field it is better to have a reliable and easy to use solution so that unnecessary time is not spent working with the system instead of doing fieldwork. It is better to use a system you are familiar with or can easily get support for, than to try and use a cheap but unsuitable tool. The payback in terms of field time will be well worth the investment. Having said this, expeditions generally have a long time to plan the work before leaving for the field. The weeks before the field trip could be spent getting familiar with the software. Either way, it is imperative that once in the field, the team can work with the software as efficiently as possible.

14.5.1 ArcView 3.x and ArcGIS 8.x/9.x

This very popular GIS software is provided by ESRI and is as close to an industry standard as it is possible to get in a very competitive market. ESRI used to be compatible with ERMapper and supported both .ERS and .ECW files, but this is becoming less valid as ESRI moves more to support *IMG files and MrSid compression. That said, ERMapper can export other more standard file formats, which can be loaded by all GIS packages, meaning that matching GIS and image processing software is still not critical.

ESRI's software is now known as the ArcGIS suite. Until recently ESRI's software came in two distinct packages, ArcView and ArcInfo. ArcView was a desktop package with limited raster and analytical capabilities. ArcInfo was the more powerful of the two but was predominantly command-line based. In 2001, these two packages were replaced by a single product called ArcGIS8, which has three levels of functionality: ArcView, ArcEditor and ArcInfo. The user interface is the same across all these levels, the difference being the functionality increases from ArcView to ArcInfo, as does the cost. The software interface for all levels consists of three components: ArcCatalog, ArcToolbox and ArcMap. ArcCatalog is designed for managing geo-datasets and various metadata types either locally or from networked databases. It works in a similar manner to Windows Explorer but is significantly more sophisticated. ArcToolbox is an analytical interface designed for various geoprocessing applications. ArcMap is the interface for displaying the spatial data. ArcMap allows spatial queries and the application of various analytical techniques to both raster and vector datasets, although an add-in may need to be purchased to fully manipulate raster data. ArcGIS can also deal in a basic way with imagery, allowing histogram manipulation and application of some basic stretches.

ArcView 3.3 is still supported and is a good start for expeditions requiring a good GIS solution, as it is sold at a significantly lower price. Older versions of ArcView are available for free download (see *www.esri.com*). ArcView 3.x is the GIS solution seen in many of the screenshots of this manual and is the package taught at the EAC workshops hosted by the authors. ESRI run a web-based GIS training system: the 'virtual-campus'. Access to introductory tutorials is free, but a charge is made for more advanced tutorials. Many of these can be accessed free for CHEST users.

The disadvantage of ESRI software for expeditions is the expense. A complete ArcGIS installation would cost many thousands of pounds depending on the level of functionality (ArcView, ArcEditor, ArcInfo) and a basic copy costs £1,500. Fortunately, if the expedition is associated with a University then a complete ArcGIS ArcInfo install can be obtained for around £130 through the CHEST agreement (depending on the package opted for by the University). The Chest version can be installed on a single PC but requires a license code file from ESRI and may need a hardware dongle on the back of the PC. This severely restricts the expedition's ability to re-install onto a second PC if there is a problem in the field. ArcView 3.x is still sold and supported and can usually still be obtained, though most Universities have switched to ArcGIS 8.x or 9. The advantage of ArcView 3.x is its single interface and development environment, and many free add-ins being available from ESRI's website (written in Avenue, a proprietary development language).

ArcGIS add-ins are also available as free downloads from ESRI; numbers are limited, mainly because the development environment for ArcGIS is Visual Basic and many users still have to get to grips with such a fundamental change, however, this is changing quickly. ArcGIS cannot be installed on Windows 95 or 98 and can take up to 2 Gb (although a typical installation including sample datasets is ~500 Mb), which is a major consideration if an older PC is used as the expedition computer. ArcView 3.x requires significantly less space, with a complete install (not including sample data) requiring 100 Mb, and runs on older versions of Windows. The system requirements are also substantially less for ArcView 3.x, which was designed in the late 1990s (when a high specification PC was a Pentium II 300MHz).

If you are considering purchasing ArcGIS, but would like to see what it is capable of before spending the capital, ESRI produce a tutorial book, *Getting to Know ArcGIS*. This includes a fully functional demonstration copy of the software, licensed for six months, and costs about £55 - an excellent way to gain access to commercial software for a real evaluation of its usefulness.

14.5.2 GRASS (Geographic Resources Analysis Support System)

GRASS is described as the world’s leading free GIS solution. It is an open source code download and compile only solution. The system was developed in 1982 by the US Army Construction Engineering Research Laboratories (USA-CERL). It was designed as a land planning and management tool. Subsequently it has emerged as a competent and reliable GIS package. A network of worldwide developers continually updates the software as part of an open source development arrangement. The current release at the time of writing is GRASS 5 (5.0.2 released 10 April 2003). Version 5 has increased the Windows PC compatibility, but a number of features have yet to be ported from LINUX and UNIX. The software is only available for NT and similar versions of Windows (NT, Win2000 and XP) but not all modules of the software are guaranteed to work under a Windows environment. The software is free to download and is provided under the GNU General Public License (GPL). GRASS is very popular in the academic world, but also has many uses in both commercial and government arenas. The relative complexity of setting up a GRASS system has limited its further spread.

GRASS is a raster/vector GIS with image processing facilities. It can manage and store many types of spatial data. GRASS combines both an intuitive Windows interface as well as the more powerful command line interface. This flexibility makes it a very popular solution, even though some commercial GIS solutions have moved away from this approach. The main programme has over 350 programmes and tools to render maps and images and has a continuous influx of new features developed by the open source community. The software can be downloaded from *www.grass.itc.it* and is available for a variety of operating systems in one of two formats, Binary and Source Code. Binaries are zipped compiled programmes that are easier to begin using for the novice, but lack the expandability and flexibility of pure source code. Source code on the other hand requires compiling, but is required to develop GRASS extensions or to run GRASS on less standard platforms. Precompiled binaries exist for GNU/Linux, Mac OSX and Windows NT. The windows download is a 40Mb zipped download. The Windows PC version is still in development and 9x and ME are not supported. GRASS is a powerful package, but very difficult for new users to get started in. There are many downloadable manuals for using the system and for getting it started, but the software is not widely used.

14.5.3 IDRISI

IDRISI is a popular GISci package sold by Clark Labs: it has a very long pedigree and is used in many universities. IDRISI is not an anagram; it is named after a pioneering 12th Century Muslim map-maker, Abu Abd Allah Muhammed al-Idrisi. The latest version is the 14th edition of the software package, which was originally released in 1987, part of a joint initiative between Clark University, USA, and the United Nations Institute for Training and Research (UNITAR). The software became popular due to its low pricing and well-rounded functionality: it has the most extensive set of GIS and Image Processing tools in a single, low-cost package. The latest release, named *Kilimanjaro*, has over 200 modules and offers a large set of tools. Clark Labs also publish CartaLinx. CarataLinx is a 32-bit spatial database tool that can capture data from input devices, as well as from existing GIS and databases. This makes IDRISI very compatible with existing GISs that the expedition may have access to, such as ArcView, ArcInfo and MapInfo.

IDRISI is popular because of the complete solution offered by a single package reducing the need for an expensive dedicated image-processing suite. IDRISI was among the first to offer this multi-featured approach, but is no longer unique, with both GRASS and Manifold offering similar solutions. Another major advantage of IDRISI is in the set of easy to use tutorials that come with the package, as well as an extensive collection of case-study based tutorials that can be purchased separately. Modern versions of IDRISI are competitively priced in the GISci market, but probably too expensive for most expeditions. The pricing has increased in recent times, as the company has attempted to manoeuvre into the commercial world as well as the academic. Clark Laboratories offer special prices for students and non-profit organisations (Table 14-2); further details can be found at their website (*www.clarklabs.org*). In the UK IDRISI is available via the CHEST agreement, so more competitive prices may be available through a local university.

Table 14-2 IDRISI pricing system.

license	cost	notes
General	$995	Standard install
Student	$250	For full time students
Student Starter	$95	One year time limited release

14.5.4 ILWIS (Integrated Land Water Information System)

ILWIS began as a government-funded research project at the Dutch international centre for applied remote sensing, ITC, in 1984. The original DOS-based GIS software was superseded by a Windows version in 1996. In 1998 PCI Geomatics (a large photogrammetry company) became a partner with ITC. ILWIS 3.2 is the latest version, released in 2004. This release marked a change from the more commercialised direction the company had been pursing in recent years to an academic one. Previously there were two different licenses but these were merged in 2004 into a single license costing just €100.

The company boasts that over 5000 licences have been sold, but this is a very small market share compared to either ArcGIS or IDRISI. The ILWIS website at *www.itc.nl* offers some good tutorials, though finding further support for the product maybe difficult, due to its limited user base. The academic licenses can be obtained through ITC via their website or through PCI for corporate licences.

The advantage of ILWIS is that it supports both image processing and GIS applications. ILWIS 3.11 onwards is particularly useful, as it allows the software to pull in both Level 1a and 1b ASTER data. This will be of particular interest to expeditions who may be struggling to import ASTER *.hdf files into other applications. The ability to combine both remote sensing and GIS analysis for a very low price (€100 via ITC) makes ILWIS a very attractive GIS solution. ILWIS is considered a cheap and easy to use GIS solution for expeditionary fieldwork.

14.5.5 Manifold

This is a professional GIS package that is trying to change the way GIS is sold and distributed. By developing code from the ground up with no legacy software Manifold.Net, the company behind the GIS, believes it can offer a more efficient and substantially cheaper GIS solution than any competitor. The software is developed purely for Microsoft Windows with no Unix, Linux or MacOS support. Though this is a disadvantage to some users, it means the company can focus on developing a package cost effectively for the majority of users.

Manifold is a well-rounded package with a very good selection of projections and image converters. The system also features very powerful image processing facilities beyond those normally found in a GIS package. Manifold retails at US$245 but there is no student agreement to reduce this cost. It is, however, the cheapest commercially oriented GIS available. A large source of Manifold.Net’s income comes from its support lines. Whereas ESRI offer 1 year’s free support with a copy of ArcGIS, Manifold’s support can be expensive. Any contact with Manifold.Net requiring an answer will be charged for and even asking for confirmation of a bug in the Manifold system costs around US$50. This is before any attempt is made by the company to address the problem. Ordering the software outside the US is done via the website *www.manifold.net*.

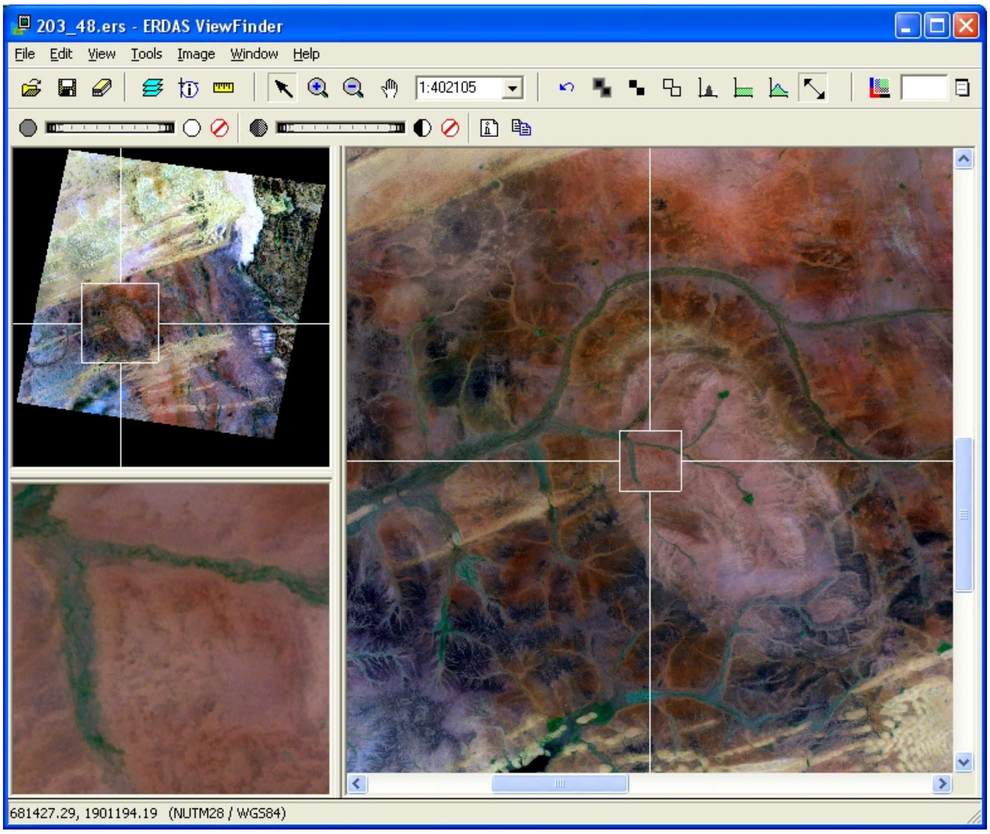
Manifold uses ECW compression as its standard imaging format over competitors such as MrSid. The latest version is 5.50 and Service Pack 1 for 5.50 was released in June 2003. One advantage of Manifold is its support for COM and Active X allowing standard languages such as C++ and Visual Basic to be used to develop it. This is a significant advantage over proprietary languages such as MapBasic, AML and Avenue. Though Manifold is not a client/server solution like some of its competitors, it can support accessing files located on remote servers. Manifold is becoming a popular GIS package and its functionality is continually growing.

14.5.6 MapInfo

This is a well-developed professional package, designed for offering business solutions to corporate clients. It features all the standard GIS tools but its primary focus is towards business not fieldwork. The main product offered by the company is MapInfo 7.8 but various other GIS tools are available to help analyse location-based data. The website is very professional and the company has continued to develop its presence and GIS software since inception in 1986. Quoting a price for the product is difficult because of the number of tailored solutions the company offers. There is, however, a version available via the CHEST agreement.

14.6 Very low cost software

Cost is most commonly the deciding factor in selecting software but ease of use and the level of assistance available should also be considered. Being forced to use a difficult system with no help available because of economic decisions should be avoided. If the expedition cannot acquire any software for economic viability reasons, there are some very cheap alternatives that offer some of the benefits of a full-blown GIS. There is a cut-down version of ER Mapper called ER Viewer, which allows users to open .ERS and .ECW images and do very basic analysis. A better solution is ERDAS View Finder, which allows the import of many image formats and, notably, the re-projection of image data. Figure 14-3 shows just how effective ERDAS View Finder can be when examining Landsat data.



681427.29, 1901194.19 (NUTM28 / WGS84)

Figure 14-3 Use of ERDAS Viewfinder to examine Landsat TM imagery of the Western Sahara, SW Mauritania. Top left, the full Landsat scene (c. 180 km x 170 km), geocorrected to the local UTM map projection. The selected TM band combinations highlight sand dunes (North) and evaporite deposits (South), as well as NNW-SSE trending rock formations. Right: detail of the window in the previous image, showing a granitic intrusion and its associated annular drainage system. Bottom left: detail of the window on the granitic terrain, showing vegetation along the valley floors. (Imagery provided by the Office Mauritanien de Recherches Geologiques and the Japanese International Cooperation Agency).

Other inexpensive software is the freeware MultiSpec downloadable from Purdue University at *http://dynamo.ecn.purdue.edu/~biehl/MultiSpec* or available on the CD accompanying this manual. A tutorial for using MultiSpec for Landsat combined with 3-DEM for SRTM is also included on the CD.

A recent development in applied GIS software is the introduction of rich content Internet based data servers. These servers offer both global raster and vector data. The principal versions include Google Earth (*http://earth.google.com*), MSN Virtual Earth (*http://virtualearth.msn.com*) and NASA World Winds (*http://worldwind.arc.nasa.gov*). These are both exciting and interesting developments in the GISci world and have come about since mid 2005. They are, however, currently extremely limited and the reader should read through the caveats listed below with care.

The first significant advantage of these data sources is that they are free. Google Earth is perhaps the most sophisticated and is based around a product called KeyHole acquired by Google in late 2004. Google Earth featured heavily in the press due to impressive images of urban areas such as the one shown in Figure 14-4. As can be seen the high resolution 2D data raster layer has a 3D vector layer of buildings. These views are currently limited to 38 major US cities and are of limited use to expeditions, but they show the potential power of an ever developing system.

Figure 14-4 3D View of New York Manhattan Island

Of more use to expeditions is the raster data available for more remote areas. This is often combined with DEM data to give a good feeling for the terrain. The mountains of southern Patagonia are shown below in Figure 14-5. This data is medium resolution and adequate for some expedition needs; Google Earth also links to historical Digital Globe (very high resolution data) that can be previewed in Quick Looks and purchased separately.

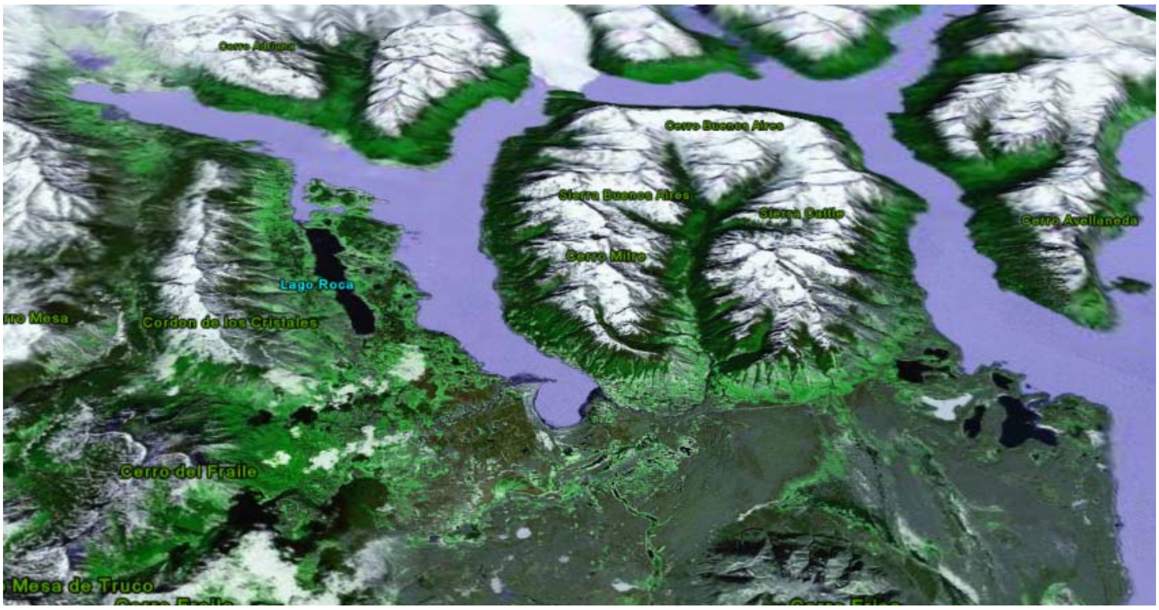


Figure 14-5 Glaciers in the South American Andes from Google Earth shown in 3D.

Additional benefits of Google Earth include locating urban regions near to the expedition's chosen study areas. By selecting the airport layer the party can very quickly locate the easiest way into and out of the area. GPS points can be uploaded into Google Earth but to take full advantage of this system requires a nominal annual subscription of US$20.

The major competitor to Google Earth in the online data hosting market is Microsoft's MSN Virtual Earth product. This integrates the existing Microsoft Terra Server data with new global data and Microsoft's mapping technology (marketed separately as Map Point GIS or AutoRoute 2006). This product is currently heavily biased towards both business and the US. It is designed to be used for locating businesses and plotting routes between them. As can be seen from the screenshot below in Figure 14-6 the same view of Patagonia is nowhere near as effective as the Google Earth view shown in Figure 14-5.

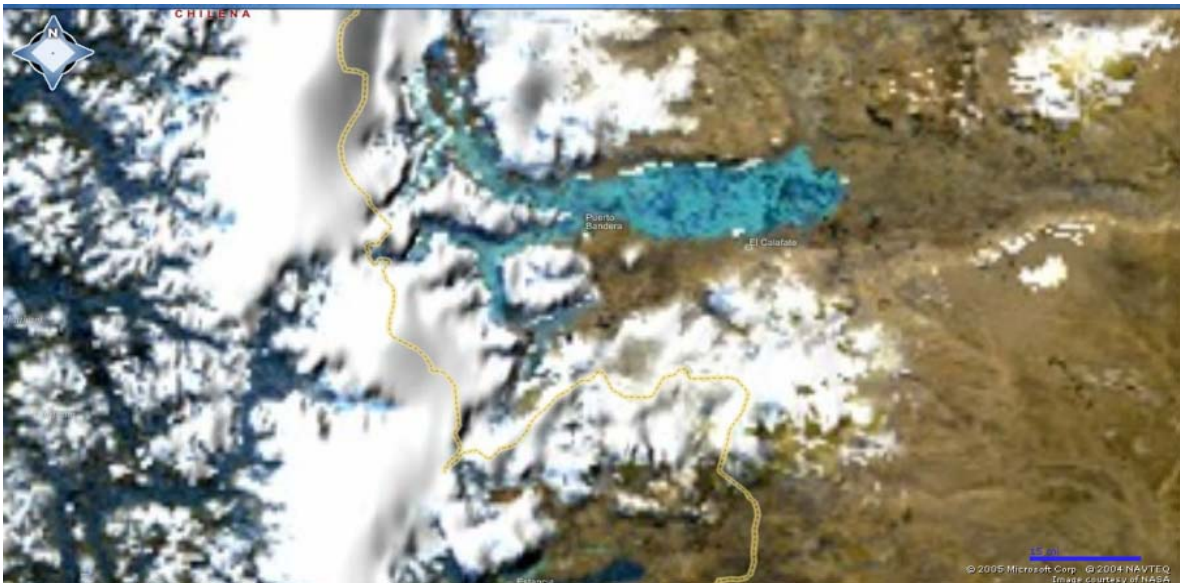


Figure 14-6 Glaciers in the South American Andes from Microsoft Virtual Earth shown in 3D.

The significant differences between Virtual Earth and Google Earth are the manner in which the user interfaces with the data. Virtual Earth is a thin client zero-footprint solution whereas Google Earth is a rich client system. This means Virtual Earth can be used straight through a web browser such as Internet Explorer while Google Earth requires a moderate download (10.5 Mb) and installation even before the data can be accessed. The actual data needs to be streamed across the Internet and requires a broadband link and a powerful PC with a 3D accelerated graphics card. There are a number of other issues with the software and more importantly the data served by these applications. There is no information in Google Earth about the age of the data. The Google Earth homepage says the data should be no more than three years old but this is not guaranteed. There is also no information on the seasonal aspects of the data. This means that an expedition could plan around water supplies that turn out to be ephemeral and not present during the actual fieldwork. The expedition would be far better served by traditional data suppliers who can offer superior information on the temporal nature of the data. Also, the data is only presented in true colour with no scope for multispectral analysis. Used carefully on a powerful PC with a broadband connection Google Earth makes a useful first stage planning tool. If in the future the data provided is of a higher resolution, across more of the world and more meta-data is supplied then it could be useful enough to replace some of the more traditional GISci software.

NASA World Winds is slightly different to either Google Earth or Virtual Earth. It too is a rich client solution and requires a 53 Mb download for the browser. Like Google Earth the system requirements are high and for the SRTM 3D views a modern 3D accelerated graphics card is essential. The current version 1.3.3.1 (25 October 2005) has the advantage over Google Earth in that it has data for every month of the year so seasonal changes can be assessed. It also has the ability to upload daily temperature and environmental information into its database. The data is a base of MODIS data with Landsat 7 ETM+. Areas of the US are available with 0.25 m resolutions with other high resolution data available where licensing restrictions allow.

14.7 Summary

A field project can succeed or fail to achieve its aims depending on choices made very early in the planning process and a critical area for success may well be software used for data collection and analysis in the field. Many projects fail to devote sufficient attention to this subject and suffer accordingly. A little time spent making informed decisions, underpinned by testing and careful review of the options available, can mean the difference between achieving more than your aims and failing to meet any of them. Software is no exception to this. If you are not sure of the issues involved with a specific piece of software that seems ideal for your project, then asking the Expedition Advisory Centre for advice.

Table 14-3 Summary of low-cost GIS packages.

	ArcView 3.x	**ArcGIS**	**MapInfo**	**IDRISI**	**Manifold**	**ILWIS**	**GRASS**
Typical UK student cost (approx.)	£100-£300	£100-£300	variable	$95-$200	$250	€100	Free
Full cost (approx.)	+$1,195	+$1,500	variable	$995	$250	€100	Free
UK CHEST agreement *	✓	✓	✓	✓	✗	✗	✗
Windows 9x	✓	✗	✓	✗	✓	✓	✗
Windows NT	✓	✓	✓	✓	✓	✓	✓
Windows XP	✗	✗	✓	✗	✓	✗	✗
System (CPU) **	Pentium	PII	Pentium	PII	PII	Pentium	PII
System (RAM) Mb	64	128	32	128	64	64	128
System (Disk) Mb	100	1 GB	103	600	100	100	100
Development language	Avenue	Visual Basic	Map Basic	COM	COM	COM	Various
Image processing	✗	✗	✗	✓	✓	✓	✓
Raster reprojection	✓	✓	✓	✓	✓	✓	✓
Raster geo-correction	✗ #	✓	✗	✓	✓	✓	✓
Latest version	3.3	9	7.8	'Kilimanjaro'	6	3.2	5.0.2

** CHEST is a UK higher education software licensing agreement; ask your university for details. CHEST agreement costs may be independent of the student licenses and could differ in price.*

*** System requirements are for lowest acceptable performance. Many will run on lower specification machines but this is not recommended. Where an NT release is available the basic system requirements will generally be higher but NT will make better use of higher specifications than Windows 9.x or ME. Requirements are for the latest version, older versions may be available with lower requirements but will lack the functionality of more modern releases.*

 # ArcView has many extensions including a free utility called Rift that can be used for geocorrection. Rift is compatible with ArcView 3.2 but requires Spatial Analyst, which may be an extra purchase.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section D: Planning & Practicalities

Chapter 15: Completing the Project

15 Completing the Project

15.1 GIS analysis

You should try to carry out as much GIS analysis as possible during your fieldwork campaign, as this will allow you to: (i) rapidly detect, correct or resample any erroneous field data; (ii) modify your field sampling to take into account unexpected findings; (iii) give a summary of your preliminary findings to interested parties in the host country. The latter is an important consideration, not just because it builds up good will with your hosts, but also because it a guaranteed way of transferring some of your new-found knowledge back to the people of the host country.

15.2 Useful GIS functions

GIS and geographical analysis techniques are covered in Chapter 7. Once you have collected your field data, it has to be checked and 'cleaned' (edited) before GIS analyses can be carried out. The 'data cleaning' is aimed at finding 'messy' data that might affect map outputs and statistical work, e.g. inconsistent names and spelling, GPS tracking data errors, or incomplete topology. You may have to modify the file formats of your datasets to make them compatible with your GIS or statistical software, although most modern GIS software comes with extensive data import-export capabilities. Once the datasets have been pre-processed, you can use the various analytical techniques available with your GIS. As with all aspects of an expedition, to be sure of a successful outcome you need to plan ahead: consider the types of data that you will be processing, the needs of the project, which analytical techniques are most appropriate, and suitable types of output for the results. Some of the more useful analytical GIS functions are summarised below, with some examples of possible applications:

- *overlaying* point records with a vegetation map
- *classifying* satellite images based on field data
- *correlating* soil and vegetation
- *gap analysis*, useful in deforestation and biodiversity studies
- *modelling* animal habitat distributions
- *change detection and analysis*, e.g. land cover, river migration
- *buffering,* useful when assessing the possible impact of new roads
- *proximity analysis*, e.g. how far do villagers have to walk to collect firewood?
- *error estimation* should always be carried out: how reliable are your findings?

With regard to the last point, information presented in the form of glossy colour maps often carries much authority and important decisions may be made on this (see Pickles 1995). However, uncertainties and errors are not necessarily apparent in maps, so make sure that any limitations are made very clear. State such limitations on the map itself, as the map is liable to be copied or shown without its accompanying text.

15.3 Results and ‘deliverables’

Results are the reason for the whole expedition! The presentation of your results should be a key consideration during your project planning: without a report detailing your findings, your expedition becomes little more than a holiday, no matter how exotic the setting. In addition to assisting with the project planning and fieldwork stages, GIS can help you to assess spatial relationships between your datasets, examine changes over time and produce professional-quality summary maps. The digital nature of the GIS-generated maps facilitates their incorporation in reports, as well as the dissemination of findings via an expedition website. Before the fieldwork, establish:

- What sort of information is needed?
- Who will benefit?
- What form will be most appropriate?

15.3.1 Data dissemination

The potential benefits that can be gained from using GIS include: summaries of your findings in easy to follow maps; spatial databases and summary statistics; monitoring systems and infrastructure development; GIS training and technology transfer; raised levels of awareness and education material based on your findings. That said, unless you carefully consider how you are going to publicise and apply your findings, the only people in the world to know what you have discovered will be the small number of people directly associated with the expedition. So the dissemination and application of your findings should be carefully considered during your project planning. It is particularly important that copies of your final report, ideally with a CD containing your data and maps, goes to agencies and institutes that you liaised with in the host country, as well as sponsoring agencies in your home country. There are various types of ‘deliverables’ that could be produced using the GIS-based part of a fieldwork-based project (Table 12-1); key ones are summarised below:

- maps, reports, databases, models
- a dedicated website, allowing relatively easy public access to your findings
- integrating your data with existing GIS databases (local, national, global)
- technology transfer via training courses that use your data and findings
- publicising your findings via the media (TV, radio, press) in the host country, the UK and internationally
- adding your findings to the RGS-IBG expedition reports archive

Almost all fieldwork projects will find a report, even if only short, a useful product, not only as a record for themselves and others undertaking similar work or visiting the same location, but so others can benefit from the expedition’s findings – whether in research, management, educational, social or other contexts. The chapter on report-writing in Winser (2004) suggests how to go about this concluding stage of the expedition.

15.4 Continuation and project sustainability

As establishing a GIS is essentially a once-only effort, continuing the GIS is an efficient undertaking that will not require a lot of time, so try to arrange that field studies initiated by your expedition can be continued by your local associates. To ensure that maximum benefit is gained from your establishing a GIS for mapping or monitoring a given set of features, ask yourself the following questions:

- How might the GIS contribute to local on-going strategies for management, research, monitoring, training and education?
- Who would be potential future users and collaborators? (e.g. universities, NGOs, local schools, Government Departments, National Parks, future expeditions, UK school projects and annual exchanges)
- Can you produce training programmes to: (a) help to build links with local / host groups, (b) help the long-term sustainability of the project? (ideally, your hosts should be able to continue your good work after you have returned home).
- Continuation is unlikely to be easy: how will it be financed, staffed, equipped and used? Would its results justify its continuation?

Figure 15-1 Water supply technicians in Bolgatanga, northern Ghana, discussing the merits of the IDRISI GIS and image processing system during a training course.

If you can train your host country associates in the use of GPS, computers, image interpretation and GIS techniques that may be the most valuable result of your expedition. Training in GISci techniques can enable people to produce some very useful material at very low cost, such as 1:30,000 scale maps of agricultural land use based on US$60 ASTER satellite imagery. Leake (1995) working in Paraguay, and the Tumucumaque Mapping Project in north east Brazil (see website at *www.ethnobotany.org/actnew/BrazilTumucumaqueMapping.html*) show how training can bring benefits not so much for research or monitoring programmes, but as a tool for indigenous communities to map their lands, defend their land claims, plan resource use and strengthen their own skills and institutions.

Expertise in GISci survey techniques is a very valuable skill: once trained to use GPS and even the most basic of GIS software packages, your local associates can potentially earn considerable amounts of money as mapping specialists. GISci-based revenues fund many

UK university mapping centres and could be just as effective in the generally far more cash-strapped institutes of developing countries, so giving GISci training to staff at host-country universities or research centres could produce long-term benefits.

Finally, related to the points raised here, is *project sustainability*: if people that you worked with in the study region can continue to collect, archive and analyse data, then you have achieved a significant success. Long-term records of features might then be collected: these are valuable to science and resource management, especially in the remote, environmentally-sensitive regions that are often frequented by expeditions.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section D: Planning & Practicalities

Chapter 16: GISci Applications

16 GISci Applications

16.1 Archaeology

One of the earliest applications of remote sensing was in the use of aerial photography to map buried archaeological features, which can often be detected because of variations in the growth of overlying vegetation: infra-red aerial photography and imagery from multispectral scanners is therefore particularly useful (e.g. Anon. 2001). The regional overview provided by satellite remote sensing is useful for mapping large features, such as ancient city walls or road systems, a famous example being the re-discovery of the 'lost city of Ubar' in the desert of southern Oman. Multi-spectral and radar remote sensing have been used on various projects trying to detect lost cities in the jungles of Central and South America. McCauley *et al.* (1982) used free Space Shuttle radar imagery to map buried valleys and geoarchaeological features in the Sahara. More recently, both GIS and remote sensing have been used to model and visualise ancient landscapes. The Internet *Archaeology Journal* provides a good source of contemporary applications.

16.2 Ecology

Check recent editions of the *Journal of Landscape Ecology*: this frequently has papers summarising projects that have used GIS-based analyses for map-based ecological research – summaries of the papers can be downloaded from the journal's website. Jorgensen and Nohr (1996) used satellite images and GIS to map landscape and biodiversity in the Sahel, whilst Gillman and Teeuw (1996) used satellite radar and GIS to examine rainforest biodiversity, using associated ground-based surveys of butterflies, the selected biodiversity indicator species. On a regional scale, AVHRR weather satellite images have been used to map habitat types across the whole of the Mkomazi National Park, Tanzania (Coe *et al* 1999): the habitat zones were then compared with the locations of animal species, using field observations and locations derived from a Land Rover-mounted GPS receiver.

16.2.1 Land cover and land use

Maps showing land cover types are essential for effective environmental management. Landsat TM imagery has been used to create a digital land cover map of Britain, using imagery from summer and winter months (Fuller *et al.* 1994). The USA, with such a large surface area to map, has used GIS to derive 'ecological resource' maps from Landsat TM images (Cain *et al.* 1997). Adinarayana *et al* (1994) used GIS and remote sensing to map land use patterns. Since the 30th anniversary of the Landsat family of multi-spectral satellites in 2002, there have been many studies of land cover changes based on 10-year observation intervals – these are too numerous to cite here, but examples can be found via an internet search under the key words 'Landsat' 'land cover' and 'change analysis'.

16.2.2 Species mapping

Satellite-derived land-cover maps and GIS have been used on many projects that have mapped habitat types and indirectly assess species distributions, a classic example of which is summarised in a paper by Loffler & Margulis (1980): Wombats detected from space. Avery & Haines-Young (1990) used GIS and Landsat TM images to map dunlin

(*Caliadris alpina*) habitats in the remote and boggy 'flow country' of northern Scotland: GIS was then used to estimate the breeding population. Lauver & Whistler (1993) and Lauver (1997) used a similar approach to map species diversity patterns in Kansas grasslands. A useful summery of new advances in species distribution modelling is given by Vaughan & Ormerod (2005), who also cite several recent case studies. Some studies have attempted to monitor key biodiversity indicator species, such as certain butterflies or birds, using trap-record-and-release methods, with a GIS used to plot and monitor their distribution. This approach was extended by Gillman & Teeuw (1996) who used satellite radar to map rainforest habitat types, particularly forest canopy gaps.

16.2.3 Pollution

Jensen *et al* (1995) used Landsat MSS images from 1973, 1976 and 1982 to quantify changes in the distribution of Sawgrass (C*ladium samaicense)* and Caltail (*Typha samaicense*) in the Florida Everglades: further GIS analysis, involving water quality data, showed that community changes were related to high phosphorous concentrations.

16.3 Geology

16.3.1 General geological mapping

A good introductory text to geological fieldwork methods, including the use of airphotos is given by Moseley (1981). Lawrance *et al* (1993) clearly illustrate the use of remote sensing in terrain mapping. There are numerous textbooks that focus on geological remote sensing, notably Drury (1987, 2001), Beaumont & Foster (1992), Vincent (1997) and Berger (1996). Some 'classic' papers that may give you ideas on how to carry out geological mapping using remote sensing are: Baker & Baldwin (1981), mapping gossan duricrusts in Chile; Crosta & Moore (1989), mapping hydrothermally altered volcanic terrain; Loughlin (1990), using thermal scanning to detect buried features in Nevada; Rowan *et al.* (2003) assessing the mineral mapping potential of ASTER imagery; and finally, De Souza Filho & Drury (1998) and Derion *et al.* (1998), who have assessed the effectiveness of satellite radar and multispectral sensors for mapping in desert regions. There is also a Special Interest Group of the Geological Society of London and the UK Remote Sensing & Photogrammetric Society: the Geological Remote Sensing Group (GRSG) - for details of their activities (and bargain student membership rates) see *www.grsg.org*.

16.3.2 Groundwater

Lattman and Parizek (1964) used airphotos to map water-bearing rock fractures. Some excellent examples of satellite data and GIS used as tools for groundwater exploration in semi-arid regions have been produced by Gustafsson (1993) and Sander (1996). Teeuw (1995) used Landsat and airphotos, fieldwork and a low cost GIS in a groundwater exploration project in Ghana. A useful example of methodologies used to evaluate water resources in Jordan, using Landsat TM and SPOT imagery to map hydrogeomorphological and hydrogeological features, is given by Allison (2001).

16.3.3 Hazards

Van Western (1992) and co-workers at the ITC (Netherlands) have produced some excellent tutorials on the use of remote sensing and GIS in mapping slope instability features and associated hazard assessments. Further examples, involving the mapping of hazardous terrain for civil engineering projects, are given in Griffiths (2001).

16.4 Geomorphology

16.4.1 Landform mapping

A summary of landscape geomorphology with remote sensing examples is given in Drury (1987, 2002) and Foster & Beaumont (1992). Reviews of geomorphological sampling and mapping techniques are given in Goudie (1990), Brunsden *et al* (1975), Verstappen & van Zuidam (1975) and Lawrance *et al* (1993) detail the use of airphoto interpretation and geomorphological mapping for rapid surveys of terrain. Further examples with a civil engineering and hazard mapping emphasis are given in Cooke & Doornkamp (1990), Fookes *et al.* (1991) and Griffiths (2001). Carroll *et al.* (1976) review the use of airphotos for soil mapping. In recent years, a whole new field of 'geomorphometric' research has developed out of the integration of multispectral satellite images, Digital Elevation Models and GIS-based analyses. A good example of this is the project of Hung *et al.* (2002), who showed how Landsat-based mapping can be integrated with ground-based GIS analyses to predict the occurrence of limestone karst cave systems in Vietnam.

16.4.2 Glaciers and mountains

Cornelius *et al.* (1994) and Carver *et al* (1995) used remote sensing, fieldwork and GIS to carry out a baseline environmental survey for a proposed national park in the glaciated landscape of central Asia's Altai Mountains: A survey of the snowline and snow patches was carried out, based on Landsat imagery and GPS-aided fieldwork. A recent survey of glacier retreat, utilizing archive CORONA photos, Landsat ETM and GPS-aided fieldwork is given by Martin Whiteside in the Case Studies section of this manual (Chapter 17). For a useful summary of glacier mapping methods using Landsat TM imagery see Paul (2000). For those of you aiming to climb high altitude peaks, Mount and Rumsbachs (2002) have ingeniously used GIS to model climbing conditions for mountaineers (headline: 'Everest climbed by computer nerds') – this is written up in Chapter 17.

16.4.3 Rivers and wetlands

Mertes *et al.* (1995) and Mertes 2002, focusing on the Amazon, have produced a useful review of remote sensing for mapping patterns of hydrology, geomorphology and vegetation. Jensen *et al* (1995) – mapped changes in the Florida Everglades using change analysis on normalized Landsat TM images. Middelkoop (2000) ingeniously sampled suspended sediment loads in the Rhine during major floods and was then able to calibrate Landsat TM images of those flood events, allowing better estimations of suspended sediment loads, as well as zones of deposition, along major Dutch rivers. Mount *et al* (2000, 2003) used archive aerial photography to quantify upland river erosion and assessed the effectiveness of the ERDAS Imagine GIS against a low-cost approach based on the Paintshop Pro graphics software.

16.4.4 Marine and coastal environments

Blue, green and red light will penetrate clear water to depths of approximately 15 m, 10 m and 5 m respectively, allowing bathymetric maps to be produced from multi-spectral scanners data (Lyzenga 1981; Van Hengel & Spitzer 1991). This has been useful for mapping and monitoring tidal habitats (e.g. Khan *et al* 1992), particularly coral reefs (e.g. Michalek *et al* 1993). By contrast, Schweizer *et al.* (2005) attempted to remove the effects of variable depth when mapping the shallow-water habitats and biomass in a marine park off the coast of Venezuela. Suspended sediment is detectable by Landsat's infra-red bands: in the UK this has been used to illustrate tidal sediment fluxes during environmental impact studies for proposed barrages on the Mersey and Severn estuaries. Phinn *et al.* (2000) give examples of remote sensing used in the mapping, monitoring, modelling and management of coastal wetlands, using Californian and Australian examples. The uses of GIS in mapping coral reefs and as a tool for integrated coastal zone mapping have been summarised by Mumby *et al* (1995). Further offshore, Wright (1996) described the various uses of GIS in an expedition mapping ocean floor features.

16.4.5 Desert and semi-arid environments

The remote sensing textbooks of Sabins (1987) and Drury (2001) have considerable coverage of geological and geomorphological applications in desert and semi-arid regions. The 1940s saw the development of airphoto-based geomorphological mapping, focused on the suitability of terrain for troop movements and tank trafficability in the deserts of north Africa (Mitchell 1991). In Australia those techniques evolved into the Land Systems approach for mapping land suitability for agriculture (e.g. Christian & Stewart 1952; see also the summaries of Lawrence *et al.* 1993). By the 1970s, geomorphological surveys were developed to map hazardous terrain and potential construction materials for civil engineering projects, primarily in the desert environments of rapidly-developing petroleum-rich countries (see summaries in Brunsden *et al.* 1975; Cooke & Doornkamp 1990, Griffiths 2001).

16.5 Natural resource management

16.5.1 Forestry

The use of conventional aerial photographs to produce 'photo-maps' for community forestry has been reviewed by Mather *et al.* (1998). Hewitt (1990) used Landsat to map riverine forests. Satellite radar has been particularly useful in mapping rainforest regions, though discrimination between tree species can be complex, involving multi-temporal composite images or composite images produced from the use of different filters (Kuntz 1995).

16.5.2 Soils / land degradation

Aerial photography has been the basis of many surveys of soil types and soil erosion, from detailed site mapping through to national surveys, though since the 1980s many regional or national surveys have used multi-spectral satellite imagery and GIS (Clarke 1986; Landon 1991). Landsat images have been used to map UK peat bogs and areas of associated erosion using Principal Components Analysis (e.g. Reid *et al* 1993) and more recently using Artificial Neural Networks. The McArthur Project based at Cambridge University, used regional weather satellite images, census data and GIS to map vegetation cover and grazing pressures in Mongolia. Teeuw (1990) reviews a soil conservation expedition that used Landsat, airphotos and field surveys in Zambia's Luangwa Valley.

16.5.3 Flood defence and catchment management

Radar's ability to both operate at night and to 'see' through cloud makes it particularly useful for mapping and monitoring major floods. Agencies involved in flood defence have benefited from satellite radar imagery, for instance in Mozambique and China plus along the Mississippi and Rhine. Satellite radar allows a GIS-based assessment of antecedent soil moisture contents, river level variations, flooded areas, flood-water volumes and the probable rate at which the flood pulse moves down-river, enabling early warnings to be given to vulnerable sites (e.g. Goodrich *et al.* 1994; Koblinsky *et al.* 1993; McKim *et al.* 1993, Schultz 1993). Townsend & Walsh (1998) used GIS to integrate radar and optical remote sensing to map and model flooding. Hamilton *et al.* (1996) and Kasischke *et al.* (1997) have used microwave remote sensing to map wetland habitats. Clark *et al* (1991) reviewed the uses of GIS for river management, examining both catchment and channel-specific applications.

16.5.4 Vegetation changes

With sufficient records over many years, GIS allows changes in distribution over time to be assessed. AVHRR weather satellite images have enabled us to monitor deforestation in Amazonia, using the Normalised Difference Vegetation Index (NDVI) discussed in Section 2.4 (Myers 1988; Sader *et al.* 1990; Smith *et al.* 1990). AVHRR imagery has also been used to map and monitor desert vegetation changes as part of an early-warning system for locust activity (by Tucker *et al.* 1985). Along the coastal plain of Venezuela, Sebastini *et al.* (1989) used past Landsat imagery for a GIS-based assessment of wetland loss. Munyati (2000) used a multi-temporal set of Landsat images to classify and monitor wetland vegetation in Zambia's Kafue Flats.

16.6 Wildlife

An interesting case study is provided by Scotland's Dee Catchment Management Planning GIS, which assessed the optimum conditions for salmon, using hydrological, geomorphological, geological, land cover and water quality data (see *www.nmw.ac.uk/ite/banc/deecamp.html*).

16.6.1 Nature conservation

The US Nature Conservancy have used aerial photography and video filming and Landsat images for rapid recce surveys designed to guide conservation planning and fieldwork (Muchoney *et al.* 1991). Furley *et al* (1994) used Landsat images, aerial photography, field surveys and GIS to devise management guidelines for a nature reserve on Maraca Island in Amazonia.

16.7 Socio-economic applications

16.7.1 Health and epidemiology

Maps indicating malaria risk along the Pacific coastal plain of Mexico were generated by Pope *et al* (1994): the optimum habitat types of *Anopheles albimanus* mosquito larvae were determined from fieldwork and were then mapped across the whole region, using GIS, with Landsat TM images from the winter (wet) and summer (dry) seasons. Monitoring of mosquito occurrences at villages in the study region confirmed the GIS predictions. GISci analysis of environmental factors and population statistics helped Simon Brooker advice on delivery of a national *helminth* (parasitic worm) control programme in Chad (Brooker *et al* 2002). Satellite data was used to determine ecological zones, epidemiological information was used to identify eco-zones where *helminth* transmission was active, and this information was used to develop a national control strategy for the country, targeting resources where they were most needed.

16.7.2 Emergency mapping in disasters

The trends that allow small expeditions to make increasing use of GIS – accessible software, hardware and data – are now being used to help humanitarian aid agencies in the aftermath of disasters. A UK charity - MapAction (*www.mapaction.org*), sent a mapping team to Sri Lanka within two days of the Tsunami that struck on 26 December 2004. Basic infrastructure maps were the first need, showing roads, railways, ports and airfields; this helped assessment teams to access affected areas. The presence of a field teams with GPS and GIS allowed information to be incorporated in 'real time'; collapsed bridges and impassable roads were shown on the initial maps, and as the infrastructure was repaired, these maps were updated and re-issued. At this stage, Landsat imagery, available free for almost all parts of the world, was also helpful in providing an overview of terrain and land use. More detailed mapping resulted from further GPS fieldwork, showing the locations of affected populations, with their number and needs, and showing the resources at the disposal of local authorities and aid agencies in the response effort. Missions such as these can readily create close links with local mapping agencies and NGOs, providing an opportunity for collaboration and training.

16.7.3 Urbanisation

GIS can model the potential environmental impacts of proposed new developments. Pereira and Moura (1999) used GIS and SPOT multi-spectral imagery to model the impact of a new bridge over the Tagus estuary in Portugal, with particular reference to wading birds. Field observations of sediment types and vegetation covers were used to classify the SPOT imagery via a supervised classification and a Maximum Likelihood algorithm. The impact of the bridge on wading bird populations was assessed at 200 m, 400 m and 600 m intervals from the bridge axis, using probability analysis.

16.7.4 Land claims, participatory mapping and institution-building

Leake (1995 and *www.herts.ac.uk/natsci/Env/Research/Leake/resource_use.htm*), working in Paraguay (see Figure 16-1) and the Tumucumaque Mapping Project in north east Brazil (*www.ethnobotany.org/actnew/BrazilTumucumaqueMapping.html*), have shown how training in GI techniques can provide an effective tool for indigenous communities to map their lands, defend land claims, plan resource use and strengthen their own skills and institutions. In Zimbabwe, participatory mapping using GIS has been effective in investigating deforestation and highlighting the differing perceptions of local farmers and the forest authorities (Mapedza *et al.* 2003).

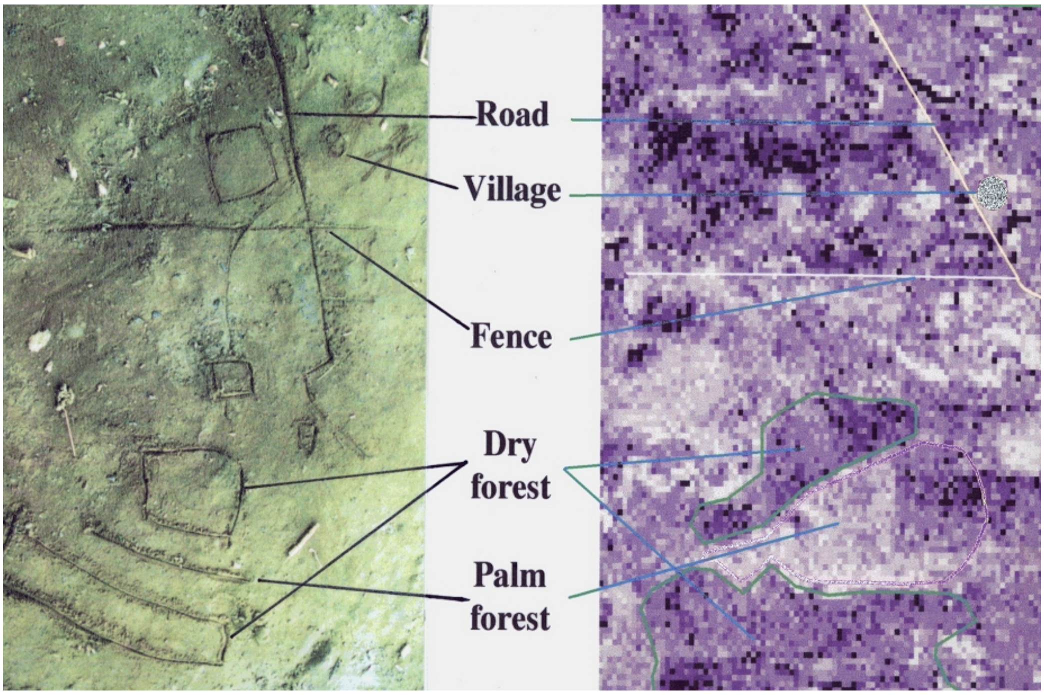


Figure 16-1 Linking indigenous people's knowledge with remotely sensed images, GPS locations and GIS analyses: on the left, a sketch map of hunting grounds scratched on the ground by Amerindians in the Gran Chaco of Paraguay; on the right, the same features plotted on a Landsat MSS image (photos courtesy of Andrew Leake).

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section D: Planning & Practicalities

Chapter 17: Case Studies

17 Case Studies

17.1 Bogda Shan Expedition 2000

The Bogda Shan 2000 Expedition was a multi-disciplined research project, jointly funded by the RGS-IBG, Imperial College Expedition Fund and the University College London Convocation Fund. The expedition was led by Alex Atkinson with assistance from co-organiser Martin Whiteside. The project studied environmental change, basic geology and other physical disciplines in the Bogda Shan range of mountains in Xinjiang Autonomous Region of the People's Republic of China.

Many GISci techniques were used, including large volumes of GPS data, targeted remote sensing and GIS analysis. One of the most important aspects of the expedition that benefited from GISci was the pre-expedition planning. Due to the political sensitivity of the region, the best maps available before the expedition were 1:1,000,000 ONC charts and 1:500,000 Tactical Pilot Charts. These charts showed a basic layout of the area but could not be used for placing base camps or detailed project planning.

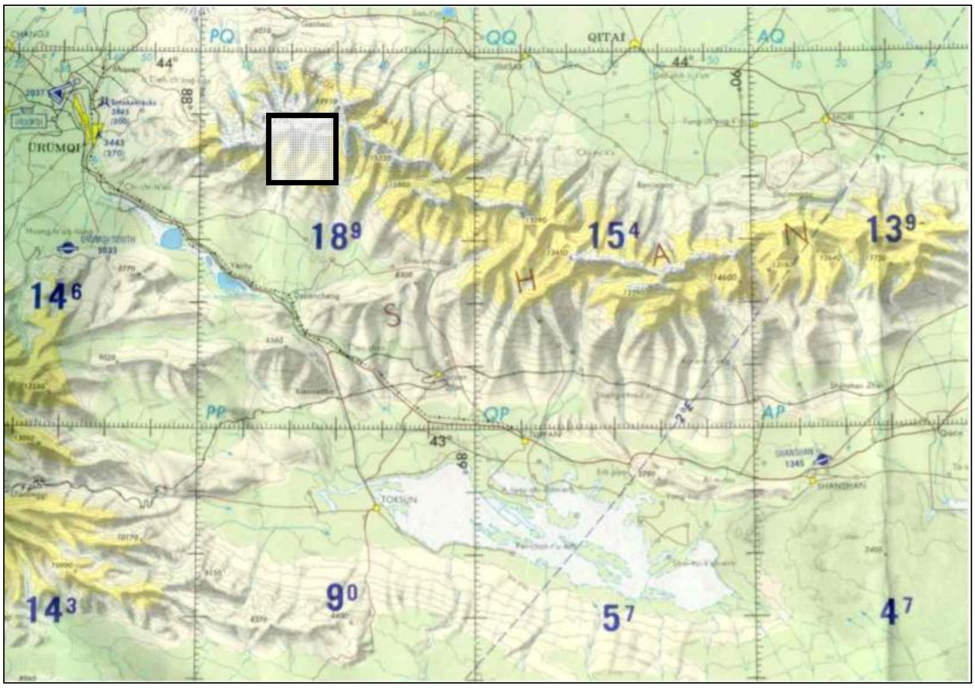


Figure 17-1 1:1,000,000 scale ONC map of the Western Tian Shan and Bogda Shan. The map was one of the most detailed documents available in the UK prior to the expedition. The black square shows the approx. 10 km x 10 km area studied by the team.

As an alternative to standard maps, the expedition used Landsat images. Landsat ETM+ multi-spectral imagery is approximately equivalent to a map of 1:75,000 scale. The advantage of multi-spectral imagery is that it can be analysed to show specific features of the terrain. The expedition needed to find a location to place the base-camp that was close enough to the mountains for easy access to study the glaciers, without being too high in the mountains where conditions would be too harsh. A Landsat image (bands 4, 3, 2) shows vegetation in a bright red colour. By studying where the red colour on the image ended the team could select an area within the vegetation line but close to the study area. The next image shows the study area. Basecamp was selected to be at the confluence of the two tributary streams. The data suggested the area was vegetated, and as the photograph inset shows, this was the case when the team arrived. The results of this pre-field study can be seen in Plate 22.

A more detailed examination of the streams size was possible by using band 8 of Landsat ETM+ data. Band 8 is a panchromatic band with four times the detail of the Landsat multi-spectral data. The team also used Corona spy satellite data that had nearly 50 times the detail of Landsat Band 8. These high-resolution black and white datasets allowed excellent planning at approximately 1:25,000 scale.

The datasets were also used as accurate base maps for the geological mapping. High resolution Corona data is so accurate that individual boulders and outcrops can be located on it. Corona data was used for field sketches that were referenced against GPS co-ordinates and digitised into Arc View 3.1. When in the GIS the data can be overlaid back onto any of the raster data. Plate 22 shows an overlay of various lithological units on the Landsat ETM+ panchromatic data.

The Bogda Shan Expedition could not have taken place without using GISci techniques to aid in the planning of the work. Without remote sensing data there would have been no way to plan where basecamp was to be situated or what areas would be accessible. Significant time would have been wasted in the field and the work would have been of a much lower quality had these steps been missed out. Some of the main GISci tools and techniques are shown below in Table 17-1.

Table 17-1 GISci techniques used by the Bogda Shan Expedition.

		Type	planning	navigation	analysis
GISci data	GPS	8/12 channel	×	✓	✓
	remote sensing data	Corona	✓	×	✓
	remote sensing data	Landsat ETM+	✓	✓	✓
GISci tools	GIS	ArcView 3.2	✓	✓	✓
	image processing	ER Mapper 6.0	✓	×	✓
	GPS download	GPS Utility 4	×	✓	✓

The Imperial College Bogda Shan Expedition obtained Landsat imagery through EurImage and the data was processed using the Intercollegiate Remote Sensing Facilities at Imperial College London. The work was supported and approved by the Royal Geographical Society (with IBG) ‘Expedition Research Grant’, with accompanying grants from the Imperial College Exploration Board and the University of London Convocation Award.

17.2 GISci helps in conquest of Everest

No prior altitude experience, no mountain training, no climbing equipment, and success at the world's highest peak – climbing will never be the same again…

I have spent the last 5 years with two big dreams. The first is to do a hobby for a job (in my case this involved becoming an academic). The second is to have a chance of a stab at Everest. To most of us such dreams are just that – dreams. But on 24 August 2001 they became reality for me: at exactly 5pm, myself and 15 Environmental Science degree students made it, without oxygen, to the summit of Everest. Between us we had climbed nine different routes to the top, including previously unclimbed routes or new variations on existing lines. Even the East Face was attempted with success gained via the central buttress and a new route tracing a line north of the trinity gullies. The whole expedition had taken just four weeks from start to finish and my employers (Charles Sturt University in Australia) had generously covered the entire cost of the trip. Only one student had ever done any climbing before, yet despite our novice status, everyone summited via their chosen route. There was not a single fatality, nobody suffered any frostbite, only occasionally did anyone break sweat and by 6pm we were all back in the pub with a well earned beer in our undamaged hands. Surely this will go down in history as one of the greatest climbing achievements of the 21st Century?

All this probably seems totally unbelievable and I expect that you are wondering how we did it. Well, maybe this article has been a little misleading so far, because the truth is that no one went anywhere near the mountain - not the real mountain anyway. You will not find Charles Sturt University recorded in the Everest role of honour because the students were studying and climbing a virtual Everest from the comfort of an air-conditioned computer laboratory in Wagga Wagga, Australia; the flattest continent on the planet. Instead of the usual ice axes, crampons, down suits and extreme physical effort we climbed the highest peak on Earth with nothing more than some basic knowledge about mountain land-forming processes, a PC, mouse and keyboard and some sophisticated geographic information system (GIS) software.

The idea for this daring mission had come partly from my love of mountaineering and partly from a belief that, with the right research and development, geographic information systems could become a useful arsenal in the climbers rack. Put very simply, GIS can be thought of as mapping and analysing the world around us on a computer. It has been successfully used for some years to assess risks in the physical environment. I wondered if it could be used to deliver climbing information to high-altitude expeditions by using digital maps of mountains to quickly identify and map mountain hazard information such as avalanche zones, provide information on air temperature, air pressure, wind chill and identify locations suitable for establishing high altitude camps. If it were possible, combining the experience of the climber with such maps could produce a powerful resource to the climber prospecting for new mountain routes.

Figure 17-2 One of the students looks for a suitable route up the Lhotse face.

So, the mountaineering challenge was offered to the students together with a highly accurate computer map of Mt. Everest and some information about the basic causes of slab avalanches. On top of this they researched the logistics of establishing a high altitude camp and what is currently possible for the human body at altitude, such as the maximum elevation gain a climber can be expected to make per day at altitude. Armed with this knowledge and fully equipped with state of the art GIS packages, the students set about finding a new, logistically and physiologically-possible route to the top. The GIS was used to identify and map areas where avalanches might occur, locate suitable camp locations, trace the ascent route and provide route data such as total elevation gain per day and the gradient at any point on the route. Within four weeks the new routes were born and 15 young environmental scientists were on top of the world. The 3D maps produced are shown in Plate 24, Plate 25, Plate 26 and Plate 27.

Whilst the technology is impressive it certainly represents a rather clinical and simplified version of the high-drama reality of high altitude ascents. However, the research and development has only just begun. Imagine a computer program where you could sit at base camp with your laptop computer and draw your mountain route onto an on-screen map of the mountain. Now put in your age, some information about previous high altitude experience, some physiological information about yourself and the latest weather report. Sit back whilst the computer undertakes complex, physically-based atmospheric modelling to predict the air pressure and ambient air temperature at any point on the mountain. Relax whilst the computer processes these data and wind speed estimations and uses these to estimates wind chill and predict a climber’s likelihood of frostbite. Let the computer integrate this data and the current knowledge of high altitude physiological relationships to calculate information about the stress your body will be under at any point on the climb and the calorific intake your body will need to meet the stress and make a successful ascent.

This may all sound like pie in the sky but it is, in theory at least, possible. As a result of the advances in computing power, the accurate on-screen 3D maps of the mountains, which are capable of telling you information such as the altitude, slope orientation and gradient at any point on your route, are already available. There is a growing pool of information about physiological stresses experienced by high altitude. These have been collected by numerous medical expeditions to Everest and others of the world's high peaks. These data can be fed into GIS-based models of physical processes (such as avalanching) and human physiology. The atmospheric physics is well understood and the academic expertise is also there to put the data together and allow predictions about the physical and physiological stresses on a climber during a climb. With such a tool in the climber's rack would high altitude climbing ever be the same again?

Contributed by Dr Nick Mount, School of Geography, Birkbeck College, University of London, Malet Street, London, WC1E 7HX.

17.3 Jarlhettur Ridge, Iceland

The Jarlhettur ridge in central Iceland is of significant environmental interest, as it lies on the edge of the Langjokull ice cap. The ridge is a geologically recent volcanic structure, composed primarily of loosely consolidated volcanic ash and lava. This ridge has naturally dammed a meltwater outflow from the Langjokull ice cap and the glacial lake, which has formed as a result, now covers over two square kilometres and contains a huge volume of water.

The geological materials that form this dam are relatively weak and some are porous. There is evidence that previous breaches, which significantly lowered lake levels, have occurred very recently and a further breach could release a catastrophic flood into the outflow river. As human developments in Iceland include habitation on the floodplain of this outflow, a breach may result in significant loss of life.

Clearly, gaining a good understanding of the detailed geological makeup and structures of this ridge are key to ensuring that hazards are assessed appropriately. This can be difficult in a location as remote as this, and fieldwork is time consuming, arduous and expensive. Detailed mapping of the area would be key to optimising the amount and level of detail possible in a field campaign.

Mapping in Iceland is developing fast. The area has recently been mapped commercially for the first time by the Icelandic mapping agency ‘Landmaelingar Islands’ at 1:50000 scale. Unfortunately this scale of map was too small and more detail was needed to properly plan a field survey. NERC agreed to collect aerial photography and airborne remotely sensed data of the area in 2001.

The aerial photos were very detailed and 57 were needed to cover the area of interest. These were photocopied to enlarge them and the copies used to collect and identify ground control on a reconnaissance visit to the area in 2002. GPS was used to identify features on the photos and the co-ordinates were recorded directly on the copies. The photos were then scanned, processed photogrammetrically using the ERDAS OrthoBase Pro software and the GPS co-ordinates for ground control and mosaiced together to produce both a seamless Digital Elevation Model and a seamless orthophoto.

The resulting orthophoto mosaics were then printed with a grid system matching the Icelandic national grid, as used on their 1:50,000 map sheets, titles, north arrows etc. The prints were produced at various scales from 1:10,000 to 1:2,500. These were then laminated and were used to plan and carry out a more detailed field survey, which took place in August 2003. An example of the data produced for use in Iceland, including a section of the glacial meltwater lake, is shown in Plate 28.

17.4 A semi-desert soil erosion and landslide site

Areas of severe soil erosion and gullying that are relatively unsuitable for farming are known as 'badlands': their very irregular, often poorly-accessible terrain makes them a challenge for surveyors. High levels of Ca or Na in silty/clayey soils – quite a common feature over sedimentary deposits in arid and semi-arid regions – lead to the development of soil pipes, which in turn facilitates the formation of gullies. This was one of the erosion problems being examined in the Luangwa Valley fieldwork cited elsewhere in this book, but it is also a problem in many Mediterranean landscapes, such as SE Spain (viz Figure 12-1). The example given here, the Mocatan Valley near Sorbas, Almeria, is further complicated by the badland terrain developing on and around a large landslide (Mather *et al.* 2003; Faulkner *et al.* 2004). The area can be seen in Plate 12.

17.4.1 The Mocatan catchment – the problem

The Mocatan catchment study area is of interest, as the knick point of a rejuvenation event is currently progressing upstream into the tributary valleys and gullys of the system. The rejuvenation event was initiated by a river capture event, dated at around 100,000 years ago, which lowered the base level for the local river system by around 90 m. This knick point is providing local river systems with substantially increased stream power, leading to rapid erosion and over-steepening of valley sides. To compound this, the region is currently undergoing tectonic uplift, >15 m since the capture event.

The geo-materials in the area are very prone to piping. This process occurs when a geo-material that is low in certain ions and also low in clay, is exposed to a hydrologic gradient. The material will contain micro-fractures, which present little resistance to significant water pressure. The flow of water through these cavities rapidly erodes them, producing underground pipes that can be up to a metre or so in diameter. These channels frequently collapse, and pose a hazard to people and vehicles. The speed at which the material is eroded poses a more significant threat to farming, as substantial volumes of sediment are released by the process.

The climate in Almeria exacerbates the soil piping problem, in that the area is semi-desert, with the characteristic flash-flood and drought sequence of such areas. Flash floods precipitate very rapid erosion, causing pipes to develop very rapidly. Anthropogenic disturbance of the landscape also exacerbates the problem, as agricultural terraces provide perfect hydrologic gradients for soil pipe development. Road cuttings and other surface modification also precipitate piping.

The concept of pipe development is well established, but the distribution of susceptible materials, of anthropogenic disturbance and the pipe relationship with the topographic surface and consequent distribution of vegetation was not clear in southern Spain. There was a clear need to map accurately where pipes occurred topographically, where on landscape surfaces they tended to be initiated and what management strategy for these landscapes was advisable to minimise their impact. The study described here attempted to help address these issues.

17.4.2 The approach

To fully analyse the instance of piping in the Mocatan catchment, a detailed three dimensional model of the area was needed. NERC air photos from two dates were available, but investigation of the accuracy possible from extract of elevation data via digital photogrammetry yielded too much error. A field survey using a total station was therefore considered the best option.

To geo-locate the total station points, a network of base stations were established using a differential GPS system. Using a Leica total station system yielded an accuracy of ~10 cm in X, Y and Z. Seven stations were surveyed, selected for their line-of-sight coverage of the study area. Once located, these stations were used to survey-in the study area using the total station. Topographic breaks of slope were surveyed, a series of points collected from each allowing reconstruction of the landscape. An example of the data recorded for each survey point is shown in Table 17-2.

Table 17-2 Data collected from the total station survey of the landslide study area. Only 7 of the 260 surveyed points are shown: note that the heights in this subset range from 514.83 m to 561.04 m. EASTINGSCO and NORTHINGSCO are the x *(EASTINGS) and* y *(NORTHINGS) co-ordinates, corrected to Spanish national grid (UTM zone 30s, international 1950 datum).*

POINT_NUMB	EASTINGS	NORTHINGS	HEIGHT	EASTINGSCO	NORTHINGSCO
2,1	6696.9	3859.2	545.80	586696.9	4103859.2
2,2	6689.5	3841.3	551.10	586689.5	4103841.3
2,3	6673.5	3830.5	552.63	586673.5	4103830.5
2,4	6672.6	3817.2	559.07	586672.6	4103817.2
2,5	6645.1	3808.4	561.04	586645.1	4103808.4
2,6	6629.5	3823.2	549.96	586629.5	4103823.2
2,7	6617.5	3825.0	549.93	586617.5	4103825.0

The points collected were then imported into ArcGIS GIS software and a Triangulated Irregular Network (TIN) constructed from them. This was done on a laptop in the field at the end of each day so that subsequent field days could be used to position points to fill in perceived gaps in the network. Plate 30 shows the final TIN from field data only.

During data collection, field sketches were also drawn, locating accurately the pipe entrances and exits. The field sketches could be drawn to scale, as tracing paper was used with graph paper, with the map grid drawn on below. Various elements were also included in the sketches and topographic features sketched in. These field sketches were subsequently scanned in and geocorrected, allowing updating of the TIN model with the information from the sketches. That allowed a more detailed model to be constructed than the survey points alone. Plate 29 shows the updated TIN.

The sketch maps were also valuable, after scanning and geocorrecting, for producing a pipe map of the area. This allowed the estimation of connectivity between pipe entrances and exits to be made (otherwise a difficult task in such an arid environment), as well as providing a clear delineation of the topographic positions of pipe entrances and exits. Plate 31 shows a part of the sketch map: the points in red are the locations of points surveyed by the total station where detailed observations were made. Geocorrection of the sketch maps allowed them to be integrated with the total station elevation model and viewed in 3-D, as shown in Plate 32. This process aided the visualisation of the landscape, particularly where the areas where pipes were located.

Using co-ordinate data collected from the field, a set of aerial photos were ortho-rectified to provide an image layer for analysis. This was an invaluable supplement to the field data, in that it clearly defined the density variations of vegetation in the area. Vegetation is clearly an important factor in soil piping, as the data from the study area indicated a very high correlation between slopes undergoing piping and dense vegetation.

It is possible that the vegetation is preventing a proportion of rainwater from running off the slope and therefore water is being retained on these surfaces. The retention means more water percolates into whatever natural pores or passages are available, therefore pipes are more likely to form. It is also possible that the substantial interweaved shallow root system of the semi-desert plants hold the surface together and prevent collapse of underground channels until much more substantial structures are formed. However, the plant root systems could also cause the opening and development of soil pipes in the first place!

Vegetation density is clearly therefore a critical component of soil piping. The air photo datasets provided information about the vegetation density in the area. Geo-correction of the photos allowed comparison of vegetation in entrance-rich and entrance-poor areas to help define more accurately on what slope, in what position, below what type and thickness of vegetation, soil pipes develop. Figure 17-3 shows an ortho-corrected colour aerial photo draped over the elevation data to produce a 3-D model to help us understand and interpret the inter-relationships between topography, vegetation and soil pipes.

GIS, remote sensing and field survey have all been combined in this study to produce information of critical importance to the aim of the research project. Much of this work was undertaken prior to the field visit, which allowed more effective use to be made of the limited time available for fieldwork. A good understanding of all the techniques, their advantages and their limitations, was critical to achieving the project goals.

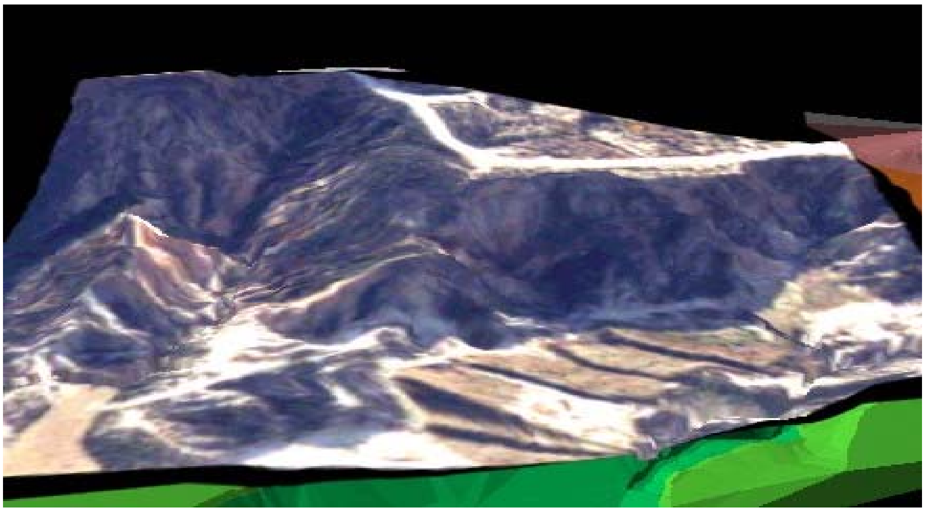


Figure 17-3 An airphoto draped over the geomorphology-modified TIN, aiding visualisation of the landscape. Airphoto courtesy of the UK Natural Environment Research Council.

Plate 21 Annotated hard-copy Landsat TM image, used alongside GPS locations to plot the course of an expedition in the Sahel.

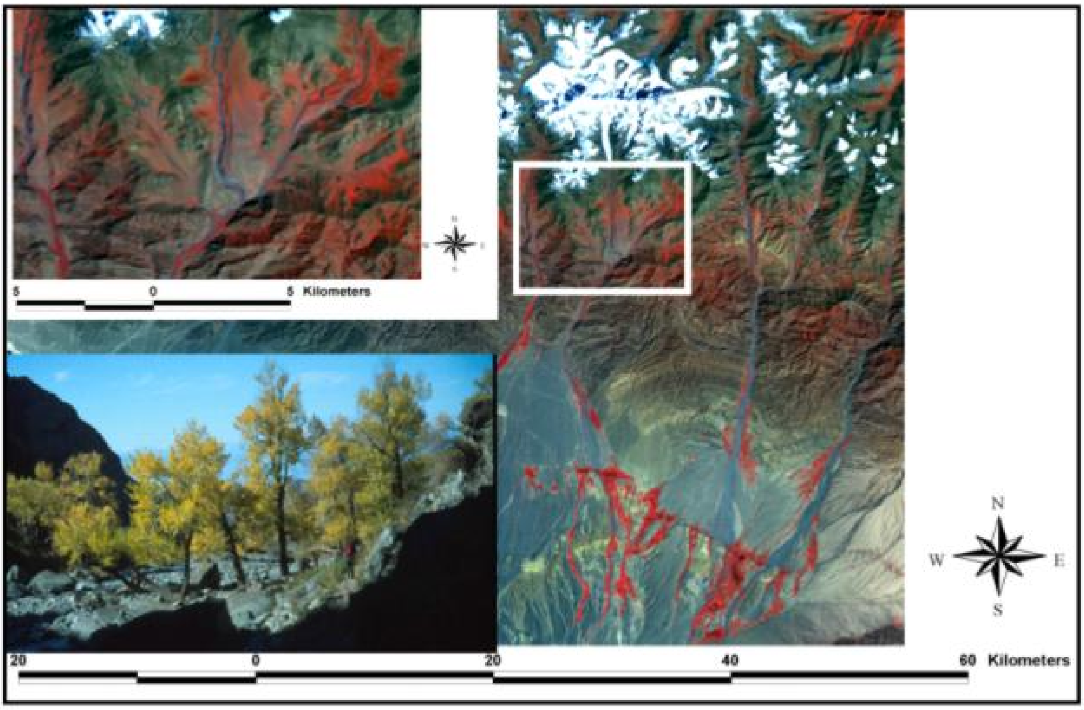


Plate 22 Landsat ETM+ Processed multi-spectral data. Red colours represent vegetation. Bright reds indicate heavy, healthy growth. The subscene (top right) shows the base camp at the stream confluence, indicating vegetation was still present at those altitudes. The photographic inset shows the trees near the basecamp (photograph: Alex Atkinson 2000).

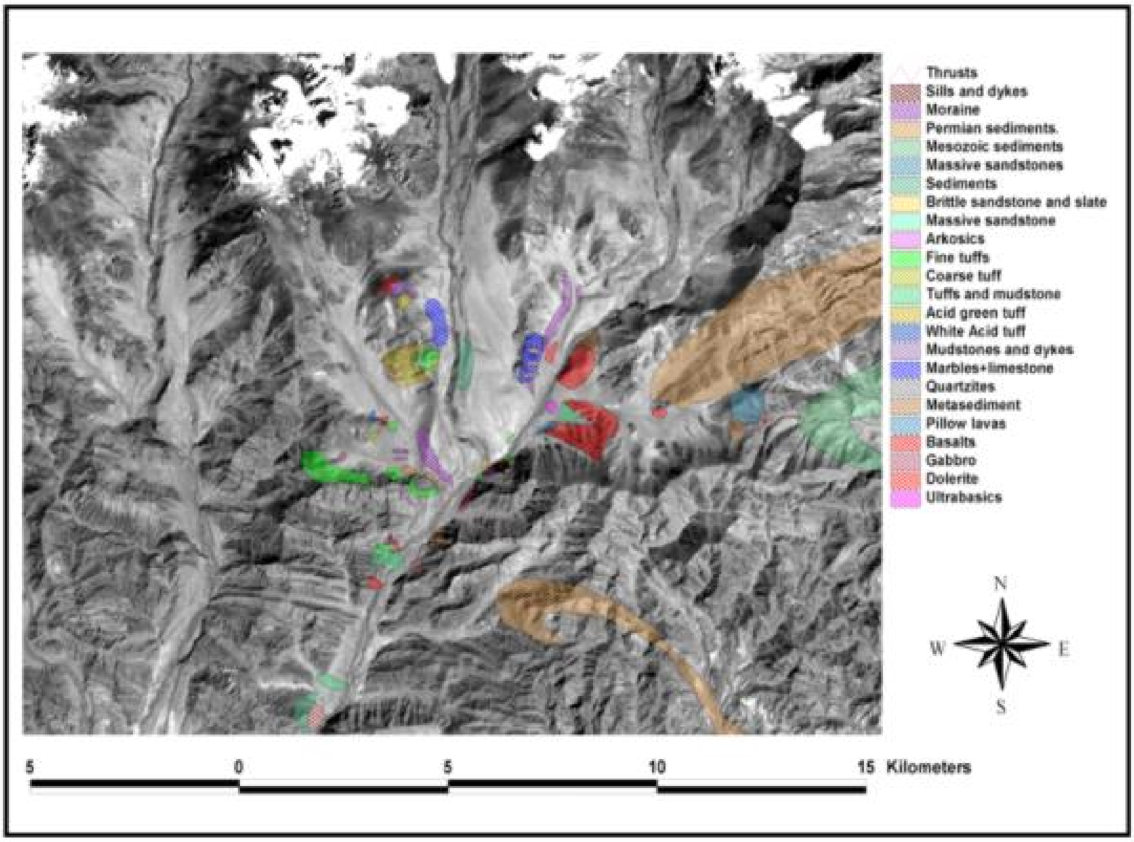


Plate 23 Panchromatic high-resolution data overlaid with geological units mapped in the field using standard GPS receivers.

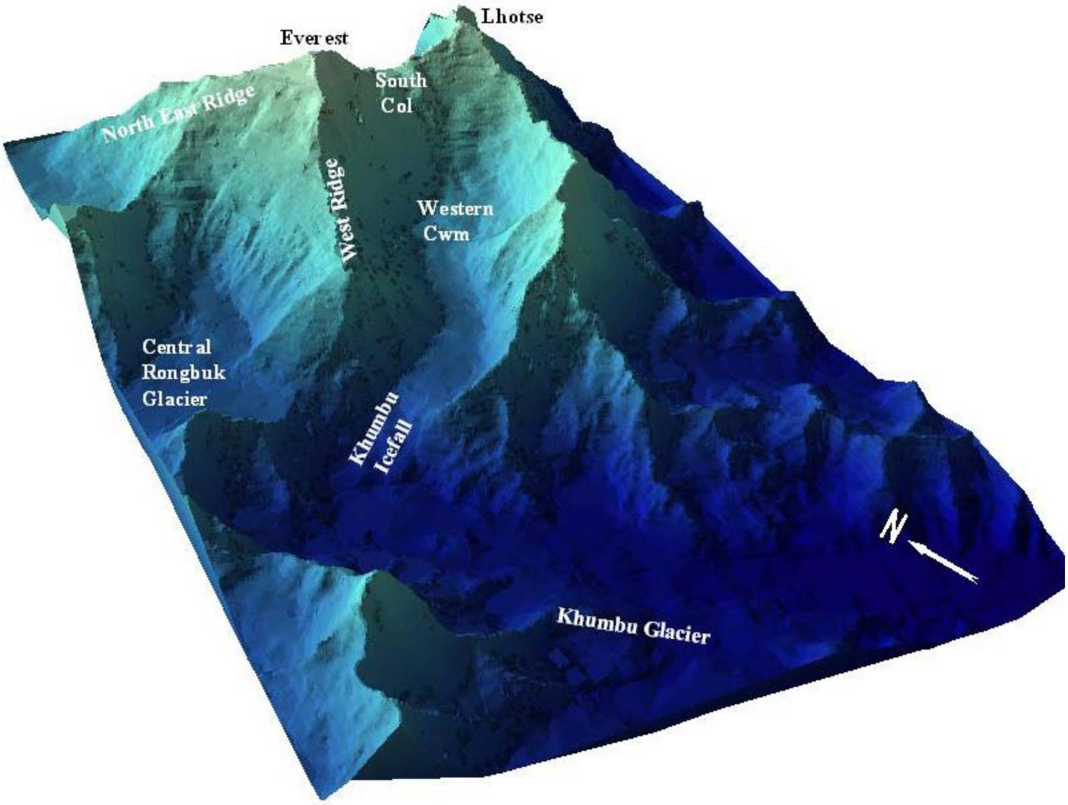


Plate 24 The computer model of Mt Everest is the basis of the analysis. The model was generated from the 1988, 1:50,000 National Geographic map of Mt. Everest by students at the Centre of Geographic Studies (COGS), Nova Scotia, Canada. The model is of sufficient detail to locate the first and second steps on the north east ridge.

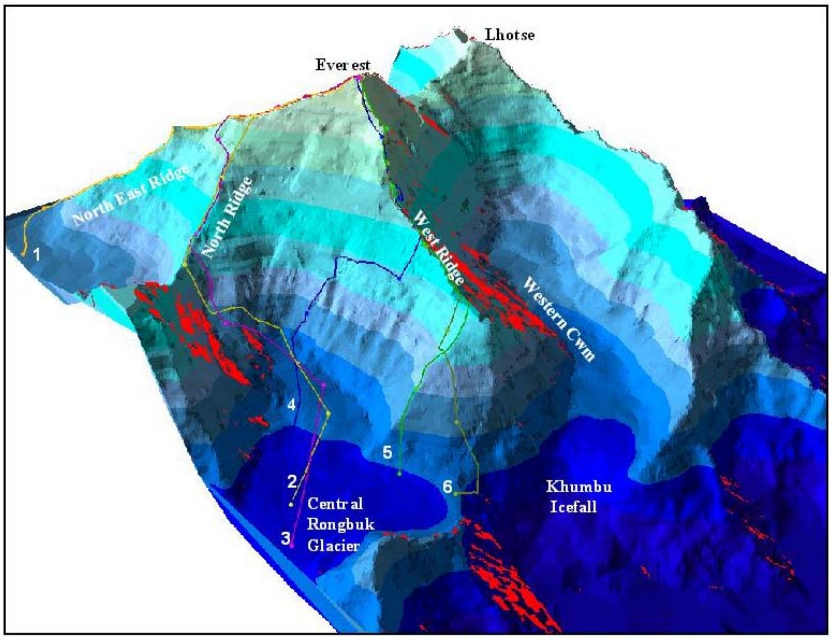


Plate 25 A view of the west ridge with student's routes 1-6 and high risk avalanche zones displayed in red. Despite a lack of climbing experience, many of the student's routes were comparable to those already climbed suggesting that the technique is capable of locating suitable climbing routes.

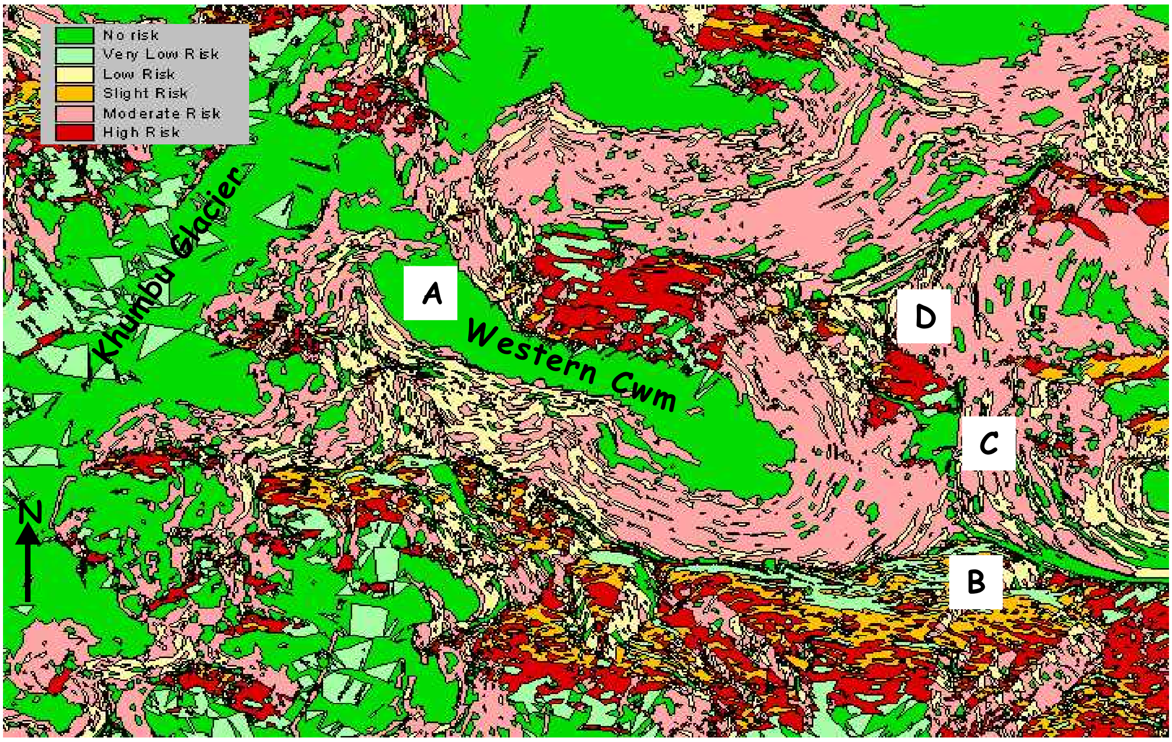


Plate 26 An example of a student's Everest physical hazards risk map showing the Khumbu Icefall (A), Lhotse summit (B), the south col (C) and Everest summit (D). The Khumbu Glacier and Western Cwm are shown relatively safe areas, but the Khumbu Icefall is clearly identified as a moderate risk zone. The final climb to the summit from the south col is identified as being of particularly high risk.

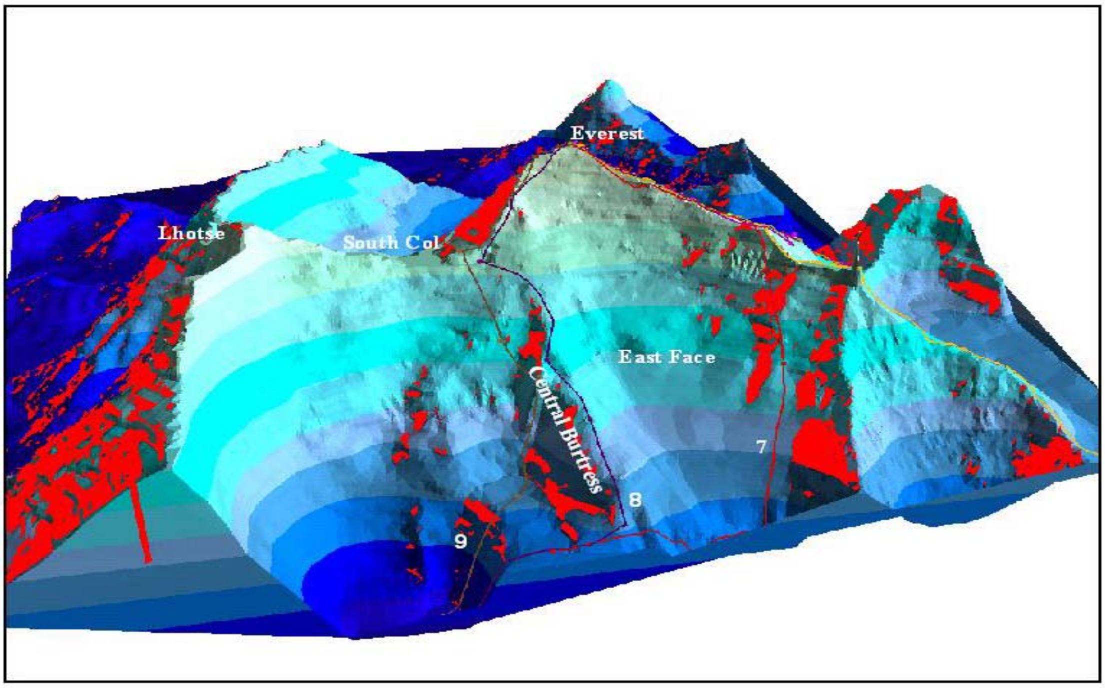
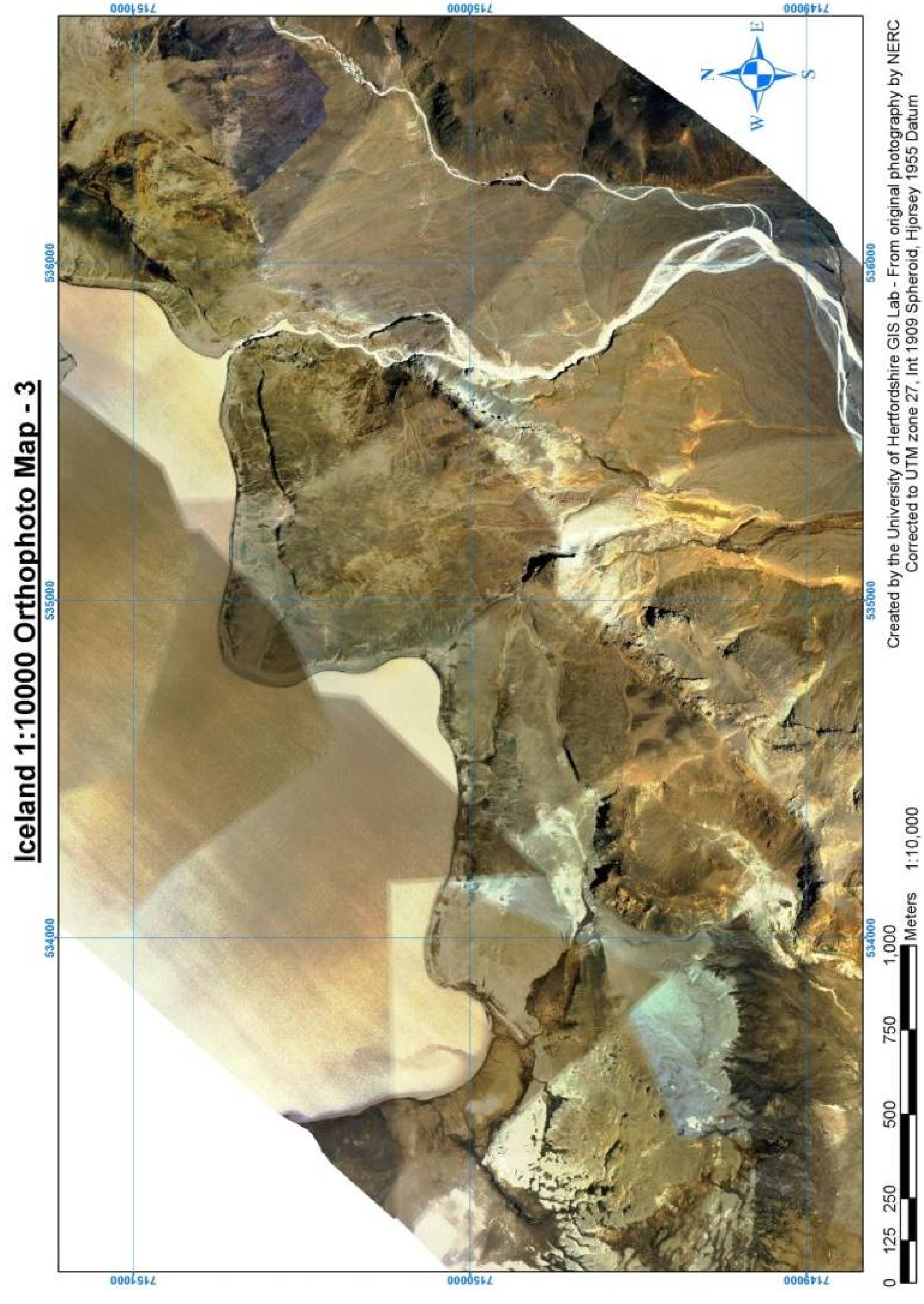


Plate 27 The extremely serious nature of the east face of Everest is clear with an increase in the red high risk zones showing an increase in snow slides and avalanches. Despite the seriousness, routes were attempted, with route 8 approximating the 1983 Central Butress route of Riechardt, Momb and Buhler.



Created by the University of Hertfordshire GIS Lab - From original photography by NERC
Corrected to UTM zone 27, Int 1909 Spheroid, Hjorsey 1955 Datum

Plate 28 Orthophoto mosaic map of Jarlhettur ridge, central Iceland.

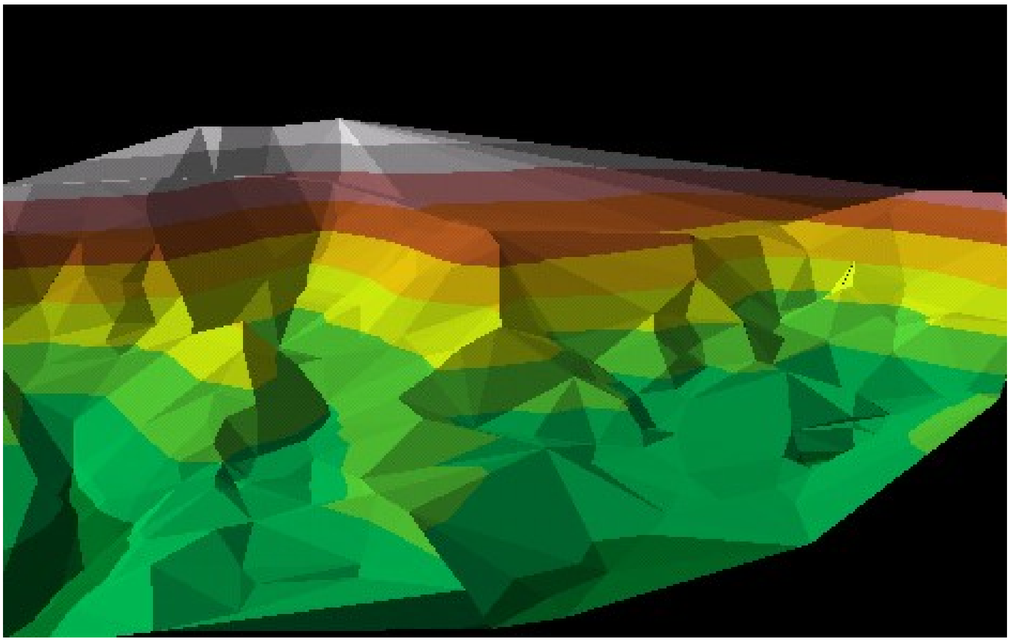


Plate 29 A Triangulation Irregular Network (TIN) produced from total station points and shaded according to height.

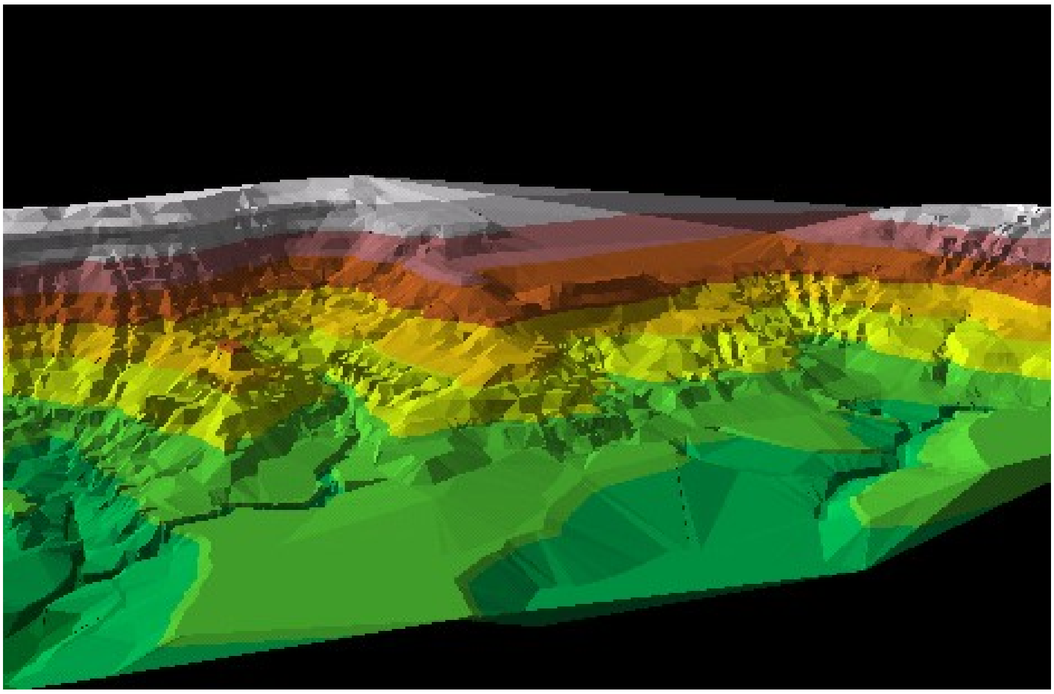


Plate 30 The same TIN as above, modified to include field observations of geomorphological features (such as zones of intense gullying and visible soil pipes) occurring between each survey point.

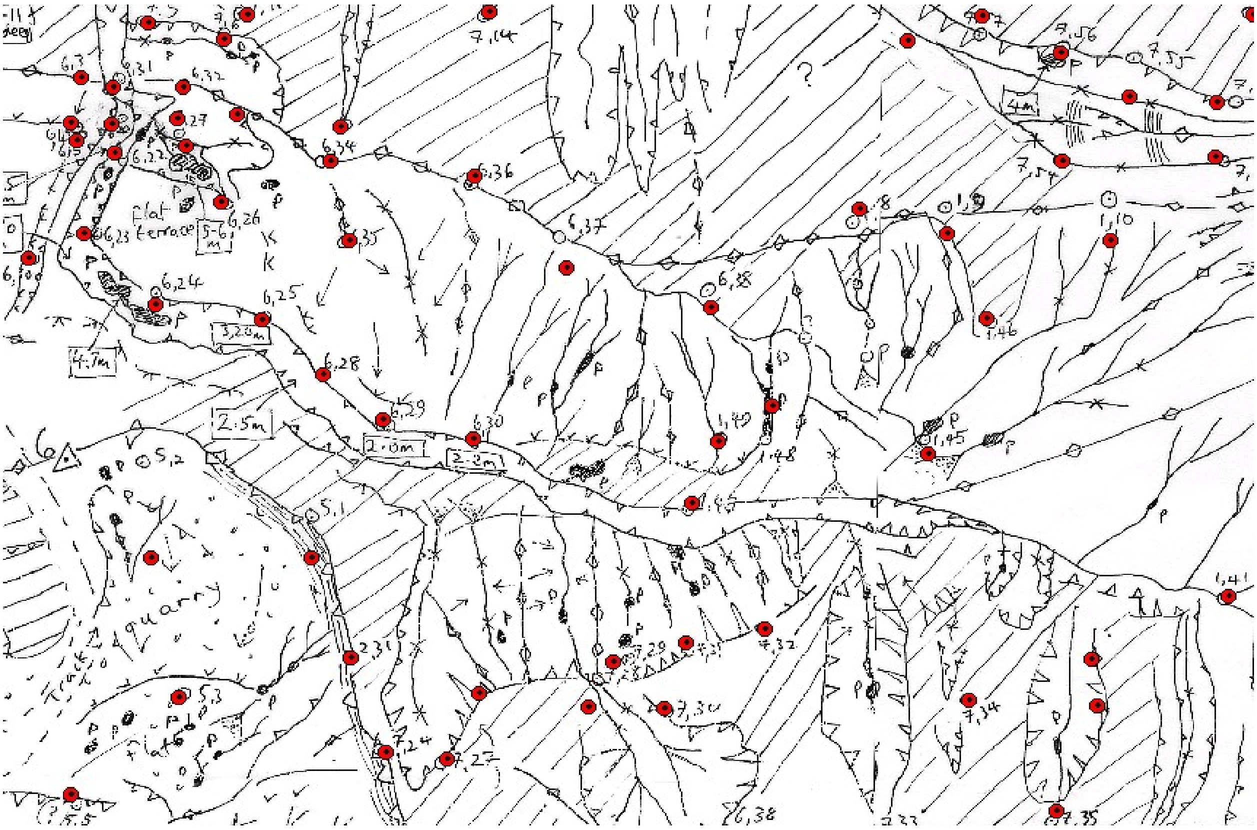


Plate 31 A geomorphological map showing directions of slope and breaks of slope, gullies and soil pipes (red dots). Map courtesy of Dr. Hazel Faulkner.

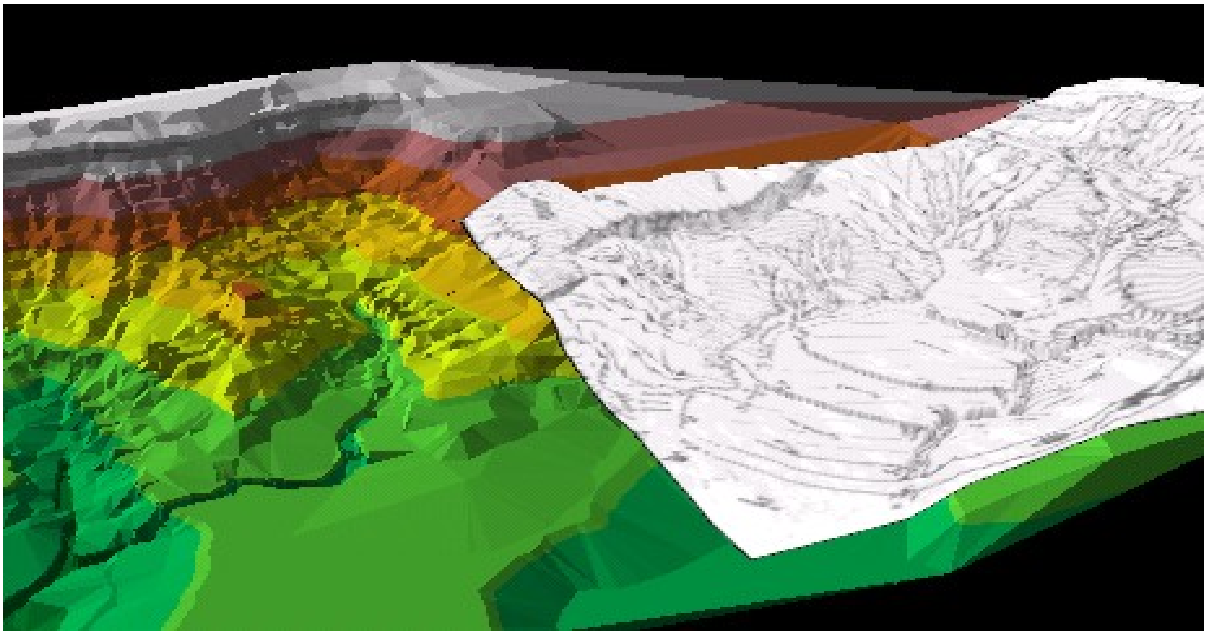


Plate 32 The original TIN, now overlain by the geomorphological sketch-map. The value of a detailed geomorphological map when interpreting landscapes is readily apparent – even if it is only based on the sketching-in of features between survey points.

Field Techniques Manual:
GIS, GPS and Remote Sensing

- Section E: Appendix

Appendix 1: Sample field data collection form

Form used by MapAction (www.mapaction.org) during fieldwork to prepare maps in disaster situations. Ready-made forms like this not only save time, but help to lower errors and ensure consistency of data. The 'Feature' columns, for example, can be filled in with agreed headings before starting fieldwork and are simply ticked when appropriate. In this case, examples were hospitals, airfields, collapsed bridges and refugee camps, but these could as well be bird species, rock types, vegetation types, etc. In this form there is an option to send information by radio for rapid incorporation into maps, but for most expedition fieldwork there will not be so much urgency and relating each record to a GPS waypoint number for download later, is simple and effective. Note the 'GIS team' heading: once data have been typed into a PC, this section is completed as confirmation.

before fieldwork: fill in the shaded boxes
during fieldwork: fill in 'waypoint' ***or*** 'radio report' section, and 'details'

date	
team name	
team initials	

GPS unit #	
co-ord system	
datum	

GIS team	
initials	
date	

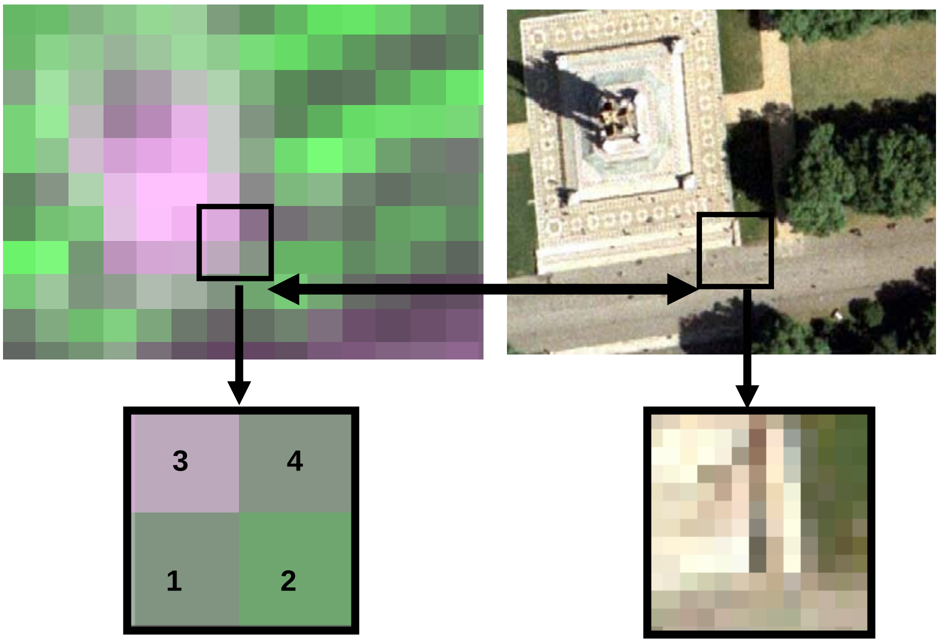
WAYPOINT	RADIO REPORT				DETAILS											
wpt num	**ref num**	**X**	**Y**	**sent?**	**feature**											**notes**
from GPS 001, 002, etc.	R1, R2, etc.	('easting') use all digits	('northing') use all digits												functional (Y/N)?	

Appendix 2: Worked example using GPS

This appendix will talk the reader through how to use a GPS, get data out of the unit and use that data in the expedition GIS. The specific example used is a simple image rectification process. This chapter will describe all the principles behind using GPS data and show how quickly it can be used to create an expedition map. To understand all the terms and processes in this appendix the user will have to read and understand the basic GPS chapter (Chapter 6) and the Photogrammetry and rectification chapter (Chapter 9).

Using GPS to Rectify Images

When collecting GPS points for rectification they must be taken as accurately as possible. This is done by locating a visible object in an image and determining the co-ordinate for that object either from a previous map, a rectified image or from ground control in the field. Taking these rectification points in the field is very difficult. An example of how difficult it is to locate point features on raster imagery is shown below. There is a clear difficulty in mapping the co-ordinates for the SE corner of the Royal Albert Memorial for ASTER (left) and even in an aerial photograph (right). Locating an exact pixel in the ASTER data to match a GPS co-ordinate is difficult even though the pixels for the statue are very different from the surrounding park. This example is substantially easier than the type of exercise expeditions will encounter in the field.

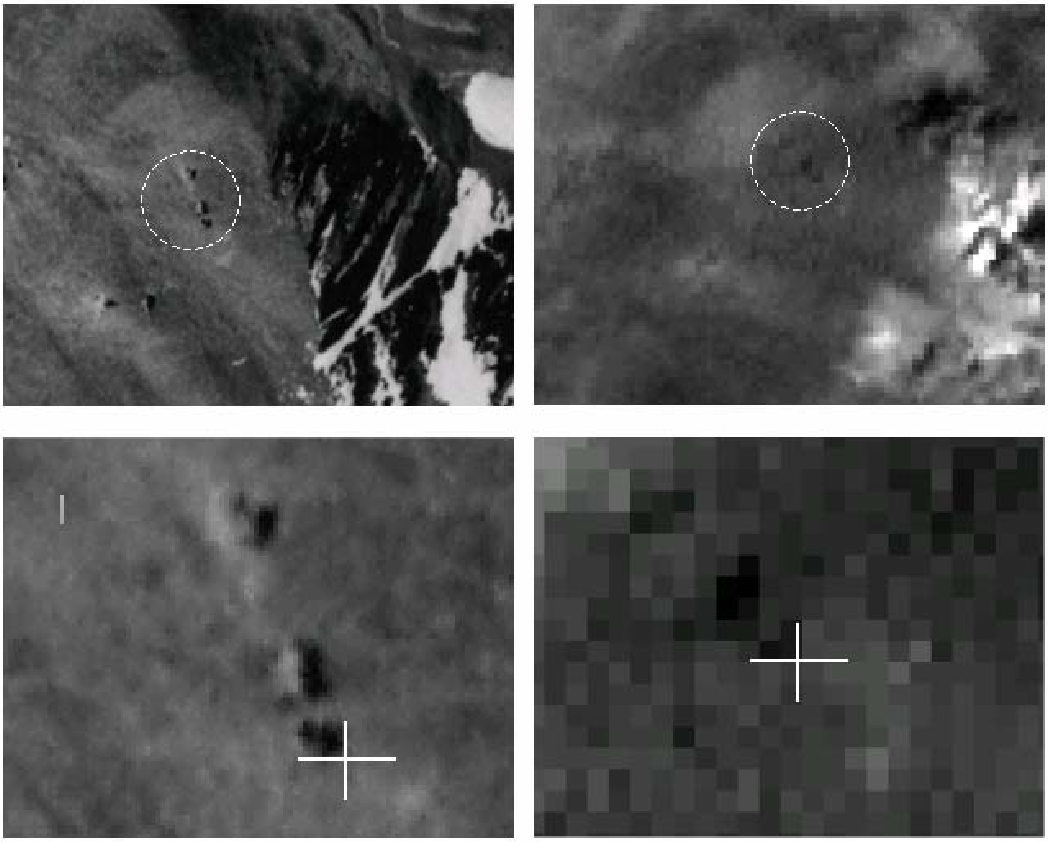


Selecting Control Points. Difficulty in accurately mapping pixels for different data sets.

The pixels labelled 1, 2, 3 and 4 in the ASTER enlargement show the problem of signal returns in areas of mixed ground types. Pixel 4 is clearly grass and pixel 1 appears to be urban and therefore part of the statue. It is therefore tempting to place the rectification mark at the edge of pixel 1 and pixel 3. Pixel 1 is not, however, as bright as some of the other pixels in the image above and therefore may be part of the path not the statue. The aerial photograph clearly shows that the statue is not the only urban object in the area and the returns might be those of the pathway so the actual point may be further back. Pixels 2 and 3 are even more difficult to interpret because their returns are definitely a mix of urban and grass. The edge of the statue might be somewhere in the middle of pixels 2 and 3 or back to the NW of pixel 1. Even on this simplest of exercises there is already a 2 pixel margin of error which in ASTER is at least 30 m.

The Bogda Shan expedition discussed in Section 17.1 relied heavily on field rectified imagery. This was because maps better than 1:500,000 were considered sensitive and not made readily available to the team. Reading information off a map of 1:500,000 scale gives an error of 500 m for every millimetre of error. Lines and points plotted on this map may be plotted with an error of more than 2 mm making all the information less accurate than 1 km. The Corona data was rectified by identifying common features on both the scanned imagery and the Landsat ETM+ panchromatic image. Even though the images were taken at a comparable time of the year, the different sun angle, differences in the terrain due to the temporal differences in the data, made the task very time consuming and inaccurate.

Image to image rectification can be very difficult and some example screen grabs of rectifying a Corona image from a Landsat pan band is shown below:



Example of Corona image (left) and Landsat pan band (right) showing three boulders used for rectification. Even with precise boulders and outcrops this exercise is difficult.

The greater the difference in image formats the more difficult this exercise becomes. A common data source for expeditions will be aerial photography. Aerial photography has a much greater spatial resolution than satellite data and can be more difficult to rectify.

Because of these differences in the Corona and Landsat imagery, identifying features was difficult and no confidence could be given in anything other than a 6x6, 36-pixel grid. This introduced large errors, as shown in the table below.

Errors generated from GPS rectification.

Error Type	**Degree of Error**	**Real World Error**
Landsat Geometric Error	~ ± 1 pixel (TM)	~50 m total (± 25 m)
Point Identification	~ ± 3 pixels (Pan)	~90 m total (± 45 m)
Error Propagation	Unknown (~±1 pixel Pan)	~30 m Total (± 15 m)

A much better method would have been to quickly use a GPS in the field to rectify the data. Once an expedition has collected GPS points they will need to be accessed to process out the data. To get the data out of the GPS needs some form of communication between the device and the expedition PC.

Using HyperTerminal to Check a GPS signal

When using a GPS and a PC together you first need to connect your GPS to a computer's serial port. Most manufacturers have a proprietary cable connector so make sure you obtain the correct one. It is often a good idea to keep GPS units the same for all members, so that if one cable is lost or fails it can be substituted with an identical one.

When physically connected a piece of software will be needed to bridge between the unit and the PC. This software is very rarely supplied with either the GPS or the cable. HyperTerminal is Microsoft's software for communicating with connected devices. Though not very useful for expedition work, it is an excellent place to start and a good method for checking a GPS is working and that you understand GPS data. It can be used to test your settings and make sure everything is working before connecting to more in-depth software. Check HyperTerminal is installed: it should be located under Programmes ⇒ Accessories ⇒ Communications. If it is not present, it can be loaded by going to Control Panel ⇒ Add Remove Programmes ⇒ Windows Settings.

When installed and located, open HyperTerminal and create a new connection. These instructions assume you are using a standard RS-232 connection but they can easily be adapted for either Bluetooth or USB connections. HyperTerminal must be told which port the GPS is connected to. Commonly this would be COM 1. Set the software to listen on Com port 1 unless instructed to do otherwise (a modem could be using COM 1 and the GPS may be assigned COM 2). A Bluetooth connection might be using a virtual port on COM 16, a USB – Serial adaptor commonly uses either COM 4 or 6. Set the Com port settings to the NMEA standard. Baud Rate 4800, Data Bits 8, No Parity, 1 stop bit as shown below. The GPS will then start sending the data to the screen. To log this to a text file select Transfer ⇒ Send to text file.

GPS NMEA settings.	
Baud Rate	4800 bits per second
Bits	8 bit
Parity	None
Stop Bits	1
Flow Control	Hardware Controlled

This should demonstrate if the GPS is working and you should see a text stream like the following.

```
$GPGGA,172704,5126.7759,N,00020.6673,W,1,05,1.3,13.8,M,47.4,M,,*67
```

It is common if one of the settings is wrong, to see a string of random characters, usually using symbols. This suggests that the baud rate has been set incorrectly. Bear in mind that some GPS units do not broadcast a standard NMEA string and some Bluetooth devices commonly broadcast at 38400 bits. This will allow you to test all possibilities and see the actual data before using a more automated solution.

Using GPS software to Download Data

A better method than streaming the data into HyperTerminal is to use a dedicated piece of software to extract stored data from the memory of a GPS. The advantage of this approach is that the GPS can be used separately from the PC for a sustained period before coming back to have its data downloaded. The other advantage of dedicate software is that it can overcome problems with projections. The X and Y data in the GPS is always stored in WGS84, so if the GPS is required to process an image into any other format, the software will have to reconvert the data back to the display information. There are many packages on the market that can do this and some GIS packages have built in scripts to handle data downloads.

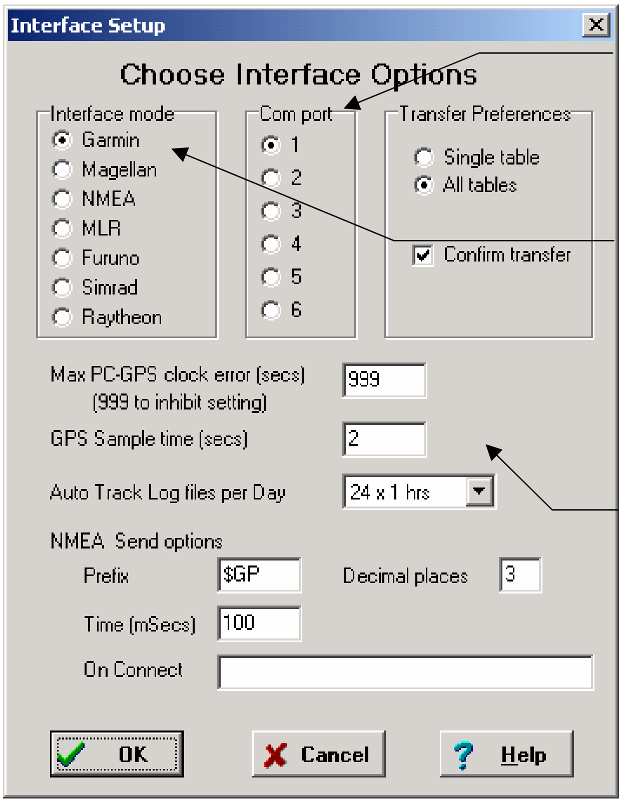
A good product for converting co-ordinates and for accessing the GPS memory is GPS Utility sometimes referred to as GPSU. GPS Utility can be downloaded free from *www.gpsu.co.uk*. To use the full product requires a registered licence. On registration (£25 GBP), the programme can be used for downloading large files of data. The programme is predominantly configured for downloading stored GPS waypoint, trackpoints and routes. It is not designed for real-time access to NMEA data or for pseudo range extraction.

Using GPS Utility to download data

GPS Utility's strong points are its ability to interface with many models of GPS and to correct/reproject co-ordinates. It can also be used to generate ArcView and MapInfo files from the GPS data. The three steps in using GPSU to download data from GPS are:

Step 1: Tell the Software about the GPS type you are using
Step 2: Define datum and projection
Step 3: Connect to the GPS

Step 1: the first step is to tell the software about the receiver being used. Go to the menu item GPS ⇒ Setup and set the parameters as shown below.



Interface mode: ● Garmin ○ Magellan ○ NMEA ○ MLR ○ Furuno ○ Simrad ○ Raytheon

Com port: ● 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6

Transfer Preferences: ○ Single table ● All tables ☑ Confirm transfer

Max PC-GPS clock error (secs) (999 to inhibit setting): 999

GPS Sample time (secs): 2

Auto Track Log files per Day: 24 x 1 hrs

NMEA Send options

Prefix: $GP Decimal places: 3

Time (mSecs): 100

On Connect:

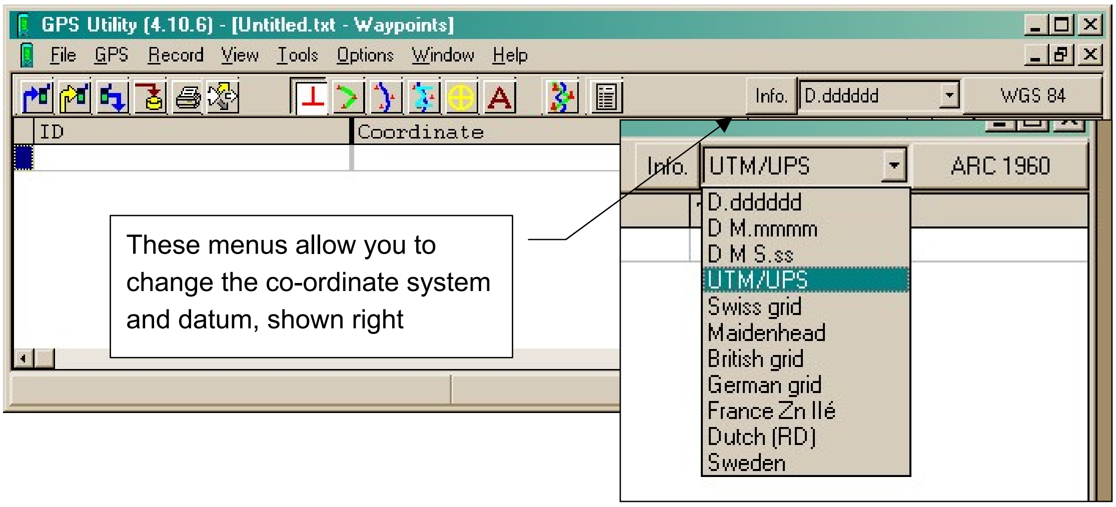
Com Port (usually 1) - this is the name of the 9-pin serial port on the computer

Select the type of GPS you are using GPS Type normally either Garmin or Magellan

The lower half of the screen can be left as the defaults. More details on the NMEA strings can be found in Chapter 11.

It is also worth checking Options ⇒ Field Properties to ensure that any altitude data is being downloaded correctly and that stored symbols can be output to GPS Utility. It is important to be aware what altitude data is stored in both waypoints and trackpoints by the expedition GPS receivers.

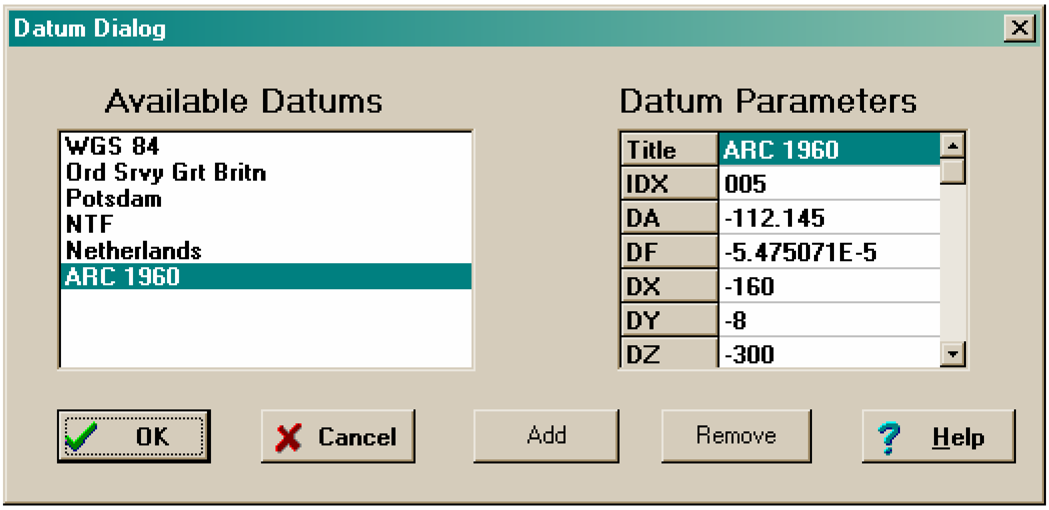
Step 2: define the co-ordinate system and datum, using File ⇒ New (or CTRL-N). First set the co-ordinate system being used (see Chapter 2), then select the appropriate datum to its right. These settings do not affect the format in which data is downloaded; rather, it determines the format that GPSU uses for displaying and saving data.



Info. D.dddddd WGS 84

Info. UTM/UPS ARC 1960

These menus allow you to change the co-ordinate system and datum, shown right



WGS 84
Ord Srvy Grt Britn
Potsdam
NTF
Netherlands
ARC 1960

Title	ARC 1960
IDX	005
DA	-112.145
DF	-5.475071E-5
DX	-160
DY	-8
DZ	-300

The datum is selected from a pre-defined list. If a required datum is not already held in GPSU, then a custom datum can be defined (in terms of angular and distance offsets from the WGS84 datum). A far greater selection is available than listed in this dialog: use ‘Add’ to include further datums. In this example, ‘Arc 1960’ has been added; it is used in much mapping of East and Central Africa.

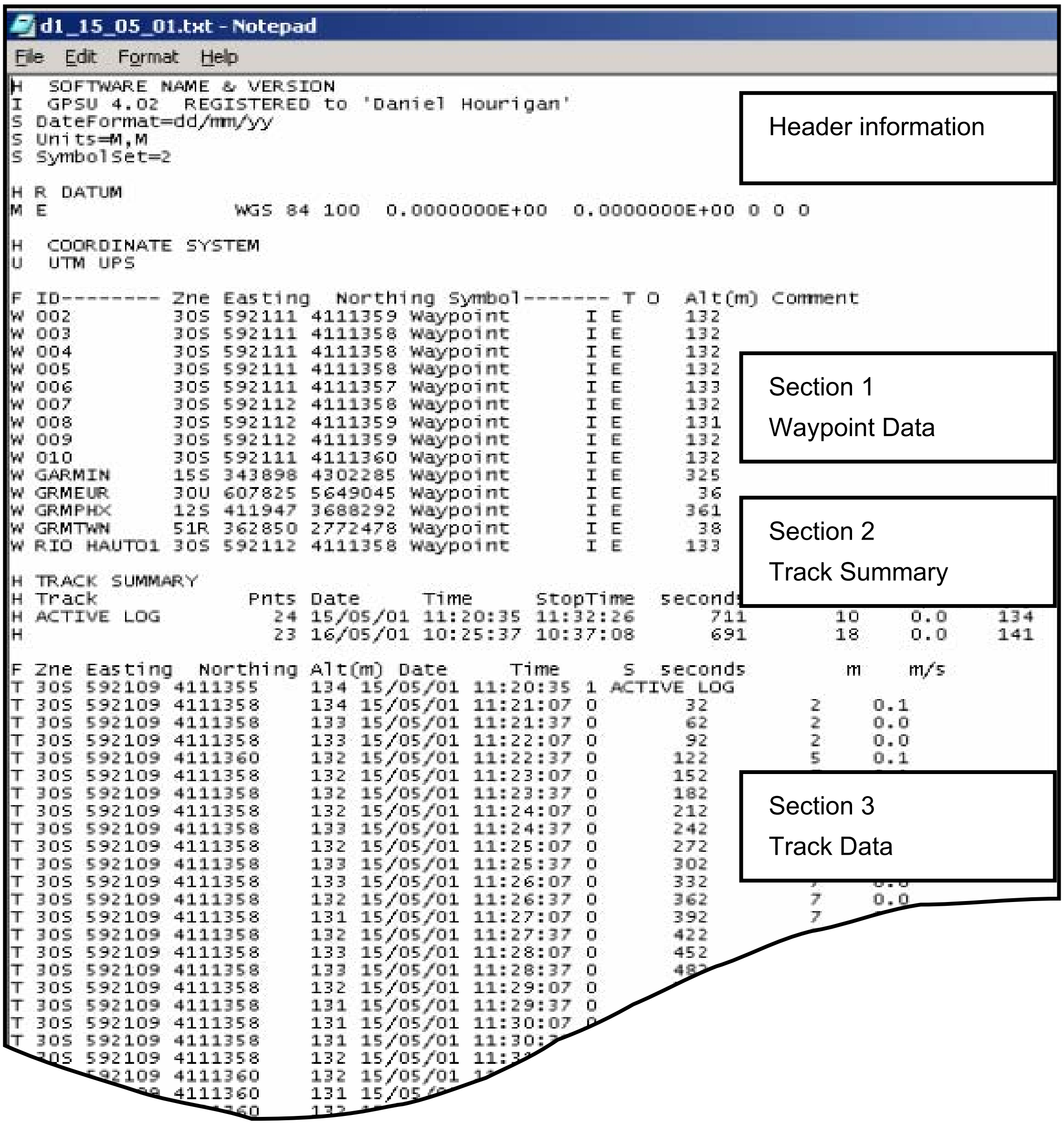
Step 3: connect to the GPS using GPS ⇒ Connect. This will automatically find the GPS model attached to the PC and display its details down in the status bar at the bottom of the screen. Next go to GPS ⇒ Download all. This gives the opportunity to select all the GPS point types that are required for the study (i.e. tracks waypoints etc.). Click OK and the programme downloads all the data. To save the data for import into a GIS click File ⇒ Save. There are many file format options here, making GPSU a useful adjunct to most GIS programmes.

Three commonly used formats are:

Shapefile: these can be read directly into ArcView (versions 3.x and 8.x), as well as many other GIS programmes. There is the option of saving Shapefiles as points, or in the case of track-logs, as lines. Shapefiles comprise several separate files: GPSU typically saves three (.dbf, .shp, .shx). The .dbf file can also be read by database and spreadsheet applications such as Excel, Access and FoxPro.

MapInfo Interchange File (.MIF): for import into MapInfo. Waypoints are saved as individual points, while tracklogs are saved as lines.

Text file: this can be useful for several reasons. Proprietary file formats change and may be error-prone, so it is always advisable to make a plain-text copy of all downloaded data. It also allows greatest flexibility if you are unsure which format to use: it can later be read back into GPSU for re-export in a particular format. Finally, plain text files can readily be edited in any text editor, such as Notepad or Wordpad, explained in the next section.

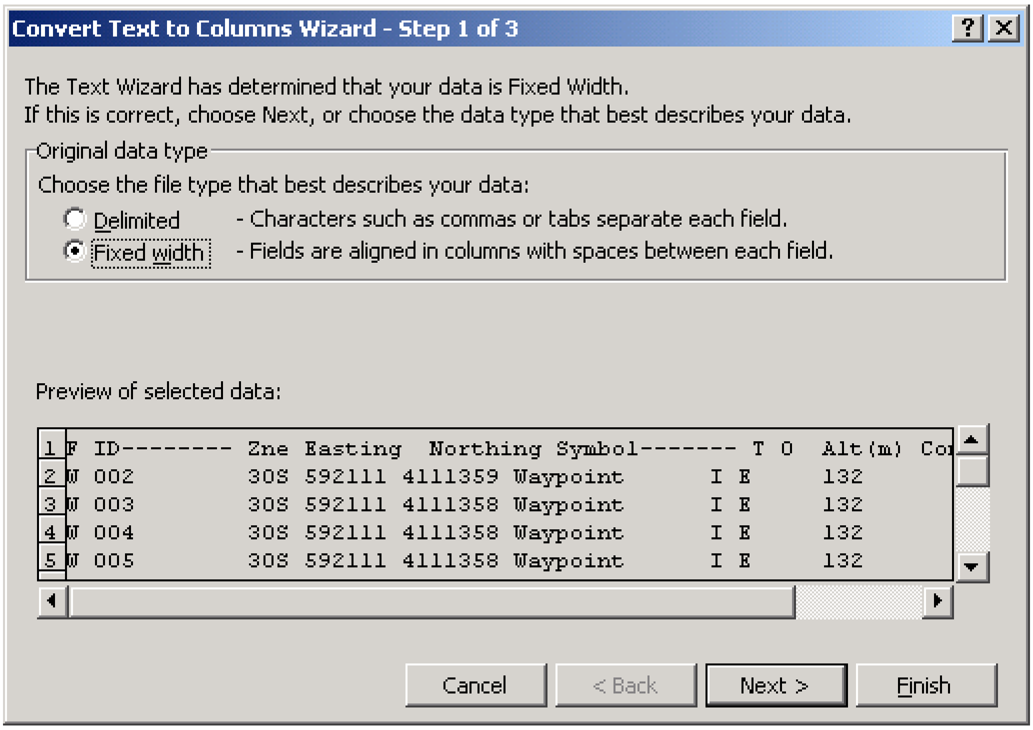


```
H   SOFTWARE NAME & VERSION
I   GPSU 4.02  REGISTERED to 'Daniel Hourigan'
S DateFormat=dd/mm/yy
S Units=M,M
S SymbolSet=2

H R DATUM
M E                WGS 84 100   0.0000000E+00   0.0000000E+00 0 0 0

H   COORDINATE SYSTEM
U   UTM UPS

F ID-------- Zne Easting  Northing Symbol------- T O  Alt(m) Comment
W 002        30S 592111 4111359 Waypoint      I E     132
W 003        30S 592111 4111358 Waypoint      I E     132
W 004        30S 592111 4111358 Waypoint      I E     132
W 005        30S 592111 4111358 Waypoint      I E     132
W 006        30S 592111 4111357 Waypoint      I E     133
W 007        30S 592112 4111358 Waypoint      I E     132
W 008        30S 592112 4111359 Waypoint      I E     131
W 009        30S 592112 4111359 Waypoint      I E     132
W 010        30S 592111 4111360 Waypoint      I E     132
W GARMIN     15S 343898 4302285 Waypoint      I E     325
W GRMEUR     30U 607825 5649045 Waypoint      I E      36
W GRMPHX     12S 411947 3688292 Waypoint      I E     361
W GRMTWN     51R 362850 2772478 Waypoint      I E      38
W RIO HAUTO1 30S 592112 4111358 Waypoint      I E     133

H TRACK SUMMARY
H Track               Pnts Date      Time      StopTime  seconds
H ACTIVE LOG            24 15/05/01 11:20:35 11:32:26       711        10     0.0     134
H                       23 16/05/01 10:25:37 10:37:08       691        18     0.0     141

F Zne Easting  Northing Alt(m) Date      Time      S  seconds        m     m/s
T 30S 592109 4111355     134 15/05/01 11:20:35 1 ACTIVE LOG
T 30S 592109 4111358     134 15/05/01 11:21:07 0       32        2     0.1
T 30S 592109 4111358     133 15/05/01 11:21:37 0       62        2     0.0
T 30S 592109 4111358     133 15/05/01 11:22:07 0       92        2     0.0
T 30S 592109 4111360     132 15/05/01 11:22:37 0      122        5     0.1
T 30S 592109 4111358     132 15/05/01 11:23:07 0      152
T 30S 592109 4111358     132 15/05/01 11:23:37 0      182
T 30S 592109 4111358     132 15/05/01 11:24:07 0      212
T 30S 592109 4111358     133 15/05/01 11:24:37 0      242
T 30S 592109 4111358     132 15/05/01 11:25:07 0      272
T 30S 592109 4111358     133 15/05/01 11:25:37 0      302
T 30S 592109 4111358     133 15/05/01 11:26:07 0      332
T 30S 592109 4111358     132 15/05/01 11:26:37 0      362        7     0.0
T 30S 592109 4111358     131 15/05/01 11:27:07 0      392        7
T 30S 592109 4111358     132 15/05/01 11:27:37 0      422
T 30S 592109 4111358     133 15/05/01 11:28:07 0      452
T 30S 592109 4111358     133 15/05/01 11:28:37 0      48
T 30S 592109 4111358     132 15/05/01 11:29:07 0
T 30S 592109 4111358     131 15/05/01 11:29:37 0
T 30S 592109 4111358     131 15/05/01 11:30:07
T 30S 592109 4111358     131 15/05/01 11:30:
  30S 592109 4111358     132 15/05/01 11:3
      592109 4111360     132 15/05/01
             4111360     131 15/05
```

Editing GPS data in text format

Though there are ways of bringing data directly into a GIS, converting it to a text file first can be useful because it allows initial checks and editing to be carried out and un-needed data to be weeded out. Because we need the data for Ground Control Points (GCPs) we require the data in a text format, so we will look at methods for exporting the data. The figure above shows how a text file saved by GPSU is structured. The header specifies the co-ordinate system and datum, followed by the sections containing waypoint data, a track summary (which can usually be ignored), and track point data.

Before data can be read by a GIS or database it requires cleaning and formatting. If an application such as HyperTerminal or GPSU has been used to export a text file, then the text will have to be converted to its correct data type before being analysed (see Chapter 4 for more details on data types). The best way to achieve this formatting is by using a parsing filter. Microsoft Excel has a good parsing filter built in as standard and is a reasonably common piece of software. This section will concentrate on how Excel can be used as part of the GISci software suite.



The Text Wizard has determined that your data is Fixed Width.
If this is correct, choose Next, or choose the data type that best describes your data.

Original data type
Choose the file type that best describes your data:
○ Delimited - Characters such as commas or tabs separate each field.
● Fixed width - Fields are aligned in columns with spaces between each field.

Preview of selected data:

```
1 F ID--------  Zne Easting  Northing Symbol-------  T O  Alt(m)  Co
2 W 002         30S 592111 4111359 Waypoint       I E    132
3 W 003         30S 592111 4111358 Waypoint       I E    132
4 W 004         30S 592111 4111358 Waypoint       I E    132
5 W 005         30S 592111 4111358 Waypoint       I E    132
```

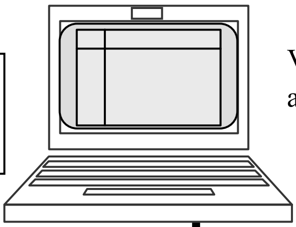
Microsoft Excel's parsing filter.

Copy the waypoints from the text file and paste them into Excel (remembering to delete any spurious header data). The data will be pasted into Excel as text, usually all together in one column. To divide the data into the correct columns and restore the appropriate data types to each field, use the parsing filter 'Text to Columns': highlight the column that the data has been exported to (usually column A) and choose Data > Text to Columns. Select 'Fixed Width', then 'Next'. Check the new columns look right (a line should separate each column, and no data should have been chopped from one column into another), then click 'Finish'. The data will now be in columns for the import. Instead of pasting, you can also open the text file using 'File ⇒ Open', then follow the 'Import' dialog, similar to 'Convert Text'.

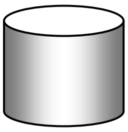
Save as a Dbase IV file and then repeat for the track data. Clean out any data that looks spurious. Garmin GPS units often have strings of test data in the downloads; these will need to be deleted before loading the information into the GIS. When the data is formatted and cleaned, it can be imported into the GIS. The GPS data can also be used to rectify raw satellite data. This process can be seen schematically over the page.

For many expeditions raw, high resolution data such as Corona may be ideal. Depending on how Corona data is imported into the computer, the imagery can have a spatial resolution of down to six feet. When in the PC the GPS can be used to make a map from the raw data. The four steps in this process are shown below.

1

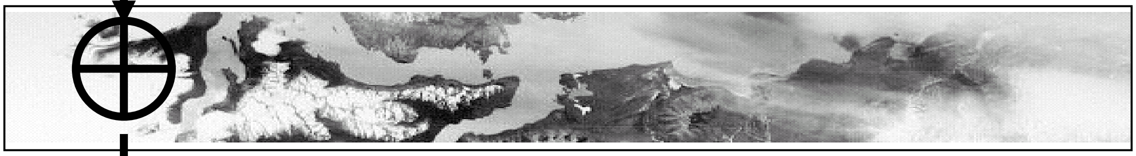


Visit USGS EROS Data centre to acquire data http://eros.usgs.gov/



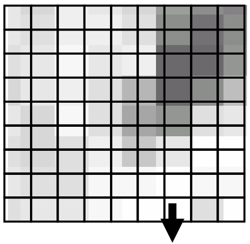
Corona data is in strips of film measuring approx 2 inches by 2 foot covering a swath on the ground of 8.6 miles by 117 miles.

117 miles convert to feet = (5280*117) = 61,7760 feet
61,7760 / 29 = 20730 ∴ each inch on the negative = 20730 feet



2

Acquire negatives and have them scanned or select the scanning method straight from the USGS. Scan type affects the pixel size in the raster images.

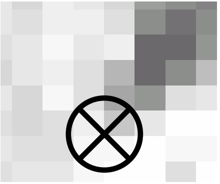


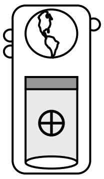
Scan types available from USGS and resultant pixel resolutions:

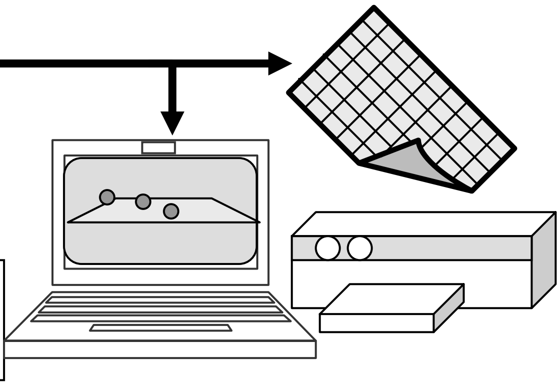
21 micron scan 1200 dots per inch = 20730.2/1200 = **17.28 feet**
14 micron scan 1800 dots per inch = 20730.2/1800 = **11.52 feet**
7 micron scan 3600 dots per inch = 20730.2/1200 = **5.76 feet**

3

Acquire ground control points using a GPS in the field or other data such as maps or rectified imagery.





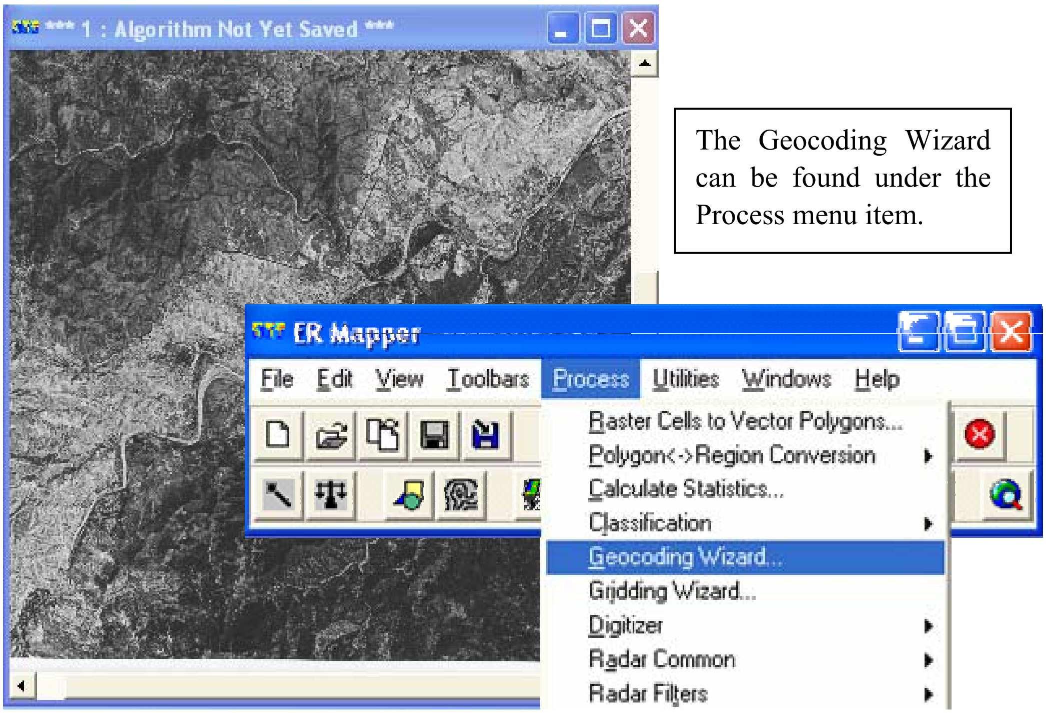


4

Rectify the data using the expedition PC. This can then be overlain with a grid to create a map.

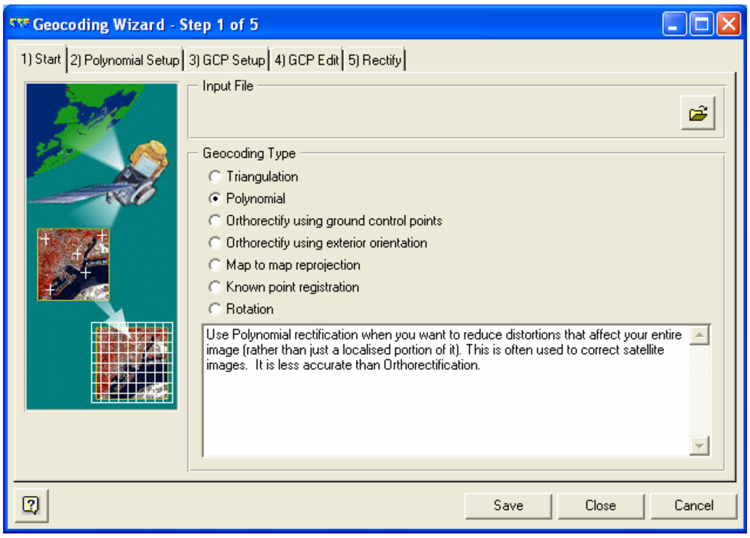
Rectifying Data using ER Mapper

To demonstrate the power of accurate GPS co-ordinates this section will demonstrate the rectification of a Corona scene of Southern Spain. ER Mapper (described in Section 14.4.2) has a good geocoding wizard that is very simple to follow. It might not be the best suited to expeditions but it should allow a good understanding of how GPS data can be used when extracted from the unit. The geocoding wizard has a series of tabs and this tutorial will work through each of the five steps in turn. This is a very rough and ready walk through for using GPS data with your imagery and the team should understand the principles of rectification outlined in Chapter 9.



The Geocoding Wizard can be found under the Process menu item.

On Tab 1 select the input file. ER Mapper will not detect this from the open window and you must specify it separately. Then you must select the geocoding type. The critical choices are either polynomial or orthorectify. These options are displayed below.



Input File

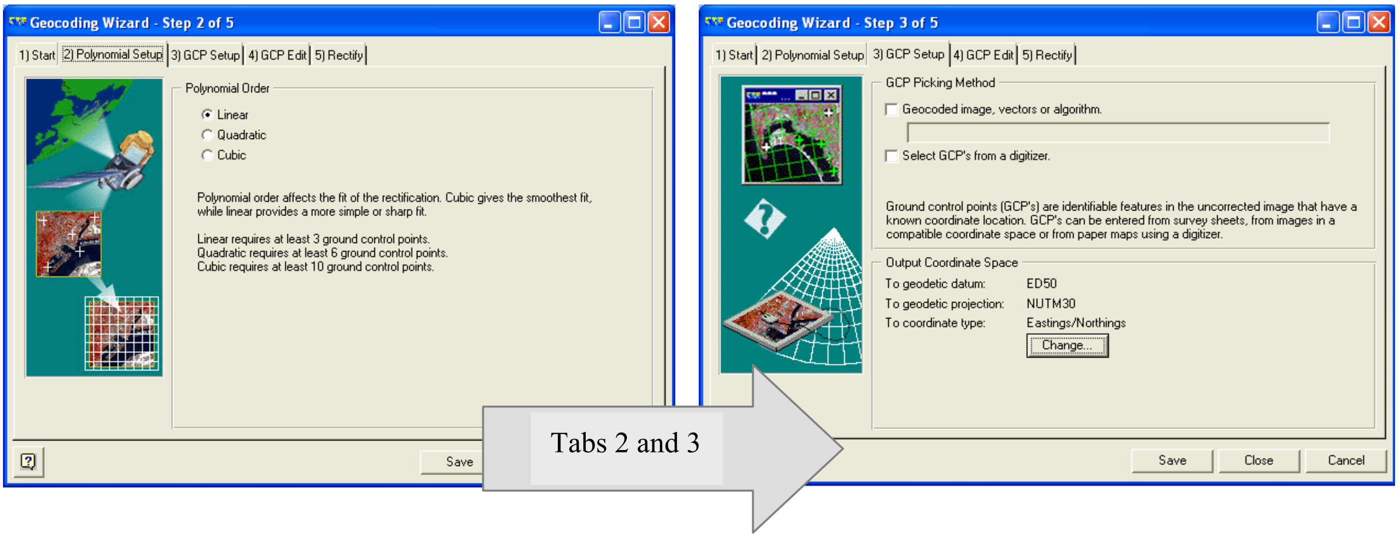
Geocoding Type

○ Triangulation
● Polynomial
○ Orthorectify using ground control points
○ Orthorectify using exterior orientation
○ Map to map reprojection
○ Known point registration
○ Rotation

Use Polynomial rectification when you want to reduce distortions that affect your entire image (rather than just a localised portion of it). This is often used to correct satellite images. It is less accurate than Orthorectification.

Polynomial rectification is used for individually recorded points such as GPS data or points read off a map; whereas orthorectification requires a 3D model.

The most common choice when using GPS data is Polynomial. When this option is selected the wizard guides moves onto Tab 2.



Polynomial Order

● Linear
○ Quadratic
○ Cubic

Polynomial order affects the fit of the rectification. Cubic gives the smoothest fit, while linear provides a more simple or sharp fit.

Linear requires at least 3 ground control points.
Quadratic requires at least 6 ground control points.
Cubic requires at least 10 ground control points.

GCP Picking Method

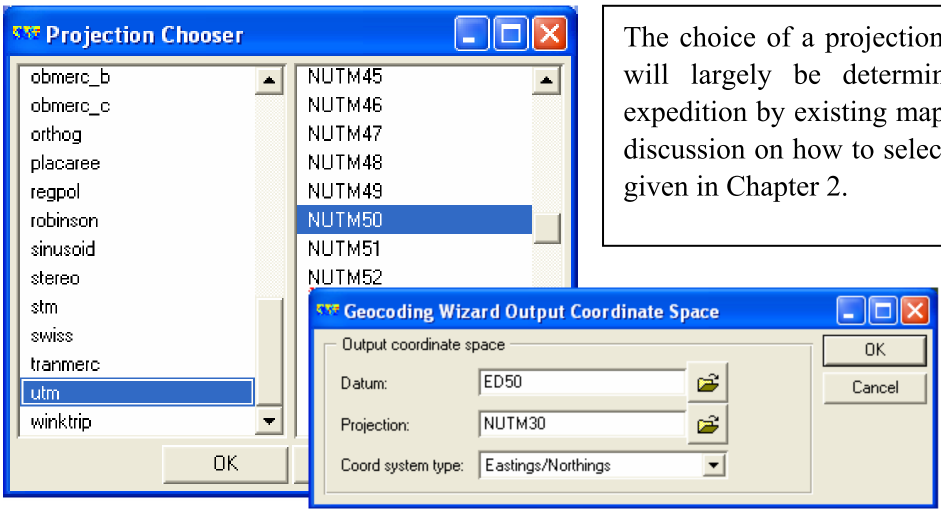
☐ Geocoded image, vectors or algorithm.
☐ Select GCP's from a digitizer.

Ground control points (GCP's) are identifiable features in the uncorrected image that have a known coordinate location. GCP's can be entered from survey sheets, from images in a compatible coordinate space or from paper maps using a digitizer.

Output Coordinate Space

To geodetic datum: ED50
To geodetic projection: NUTM30
To coordinate type: Eastings/Northings

On Tab 2 the polynomial order is selected. A general equation for polynomial order is GCP+1*GCP+2 / 2. The reader is advised to read Chapter 9 to understand the different options available when rectifying data. This is only a brief walk-through of how one program handles the operations and the expedition needs to understand the concepts in detail so they can apply them across the board to any software they encounter. ER Mapper offers three degrees of transformation. A linear transformation requires 3 GCP, Quadratic requires 6, points and cubic requires 10 GCP. Select the transformation that matches the number of GCP recorded. Tab 3 determines the output co-ordinate types for the finished data. Though this might seem complicated at first, there are various options to help you. The projection defines the method of laying the 3D surface onto a 2D plane. The system defines how these co-ordinates are entered and the choices will be dictated by the maps or imagery the rectified data needs to be referenced against. In this case, a UTM projection for Spain is used. Tab 3 allows you to select the method for transferring your co-ordinates from your corrected data to your raw data. Select the Geocoded image, vector or algorithm dialogue and in the bottom half of the screen click ‘change’. Then input the correct datum, projection and co-ordinate type. A co-ordinate space is defined by a datum, a projection and a system. The datum defines where the ellipsoid intersects the geoid. The projection defines the method of laying the 3D surface onto a 2D plane. The software delineates where and how these co-ordinates should be entered. Existing maps and imagery will largely dictate the choices made. In this case, a UTM projection for Spain is used. When an image is first brought into the system it has no co-ordinates and is described as being “RAW” data.

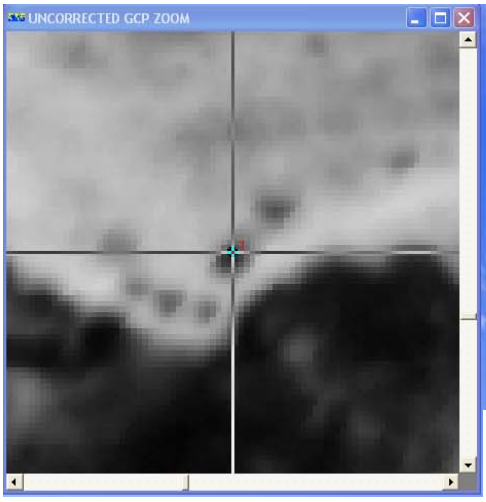


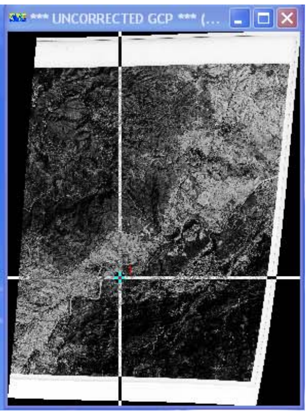
Output coordinate space

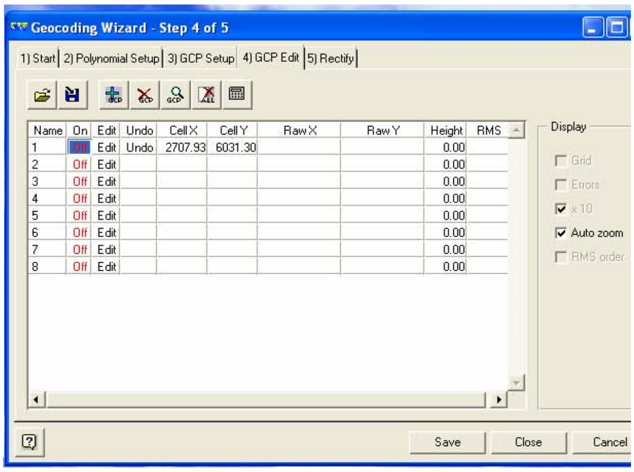
Datum: ED50
Projection: NUTM30
Coord system type: Eastings/Northings

The choice of a projection, datum etc. will largely be determined for the expedition by existing maps etc. A full discussion on how to select this data is given in Chapter 2.

Once these decisions have been made the Ground Control Points can be added to the image and the GPS co-ordinates can be entered. GCPs can be added and manipulated at will. Each GCP will have an estimated RMS error. Try to keep these errors as low as possible but refer to Chapter 9 (especially Figure 9-18) to see the pitfalls when only relying on RMS. The error should preferably be around the error of your collected data, though this is not always possible. In this case, the GPS data has an error of around 10-20 m so the RMS errors should not be more than twice this level.



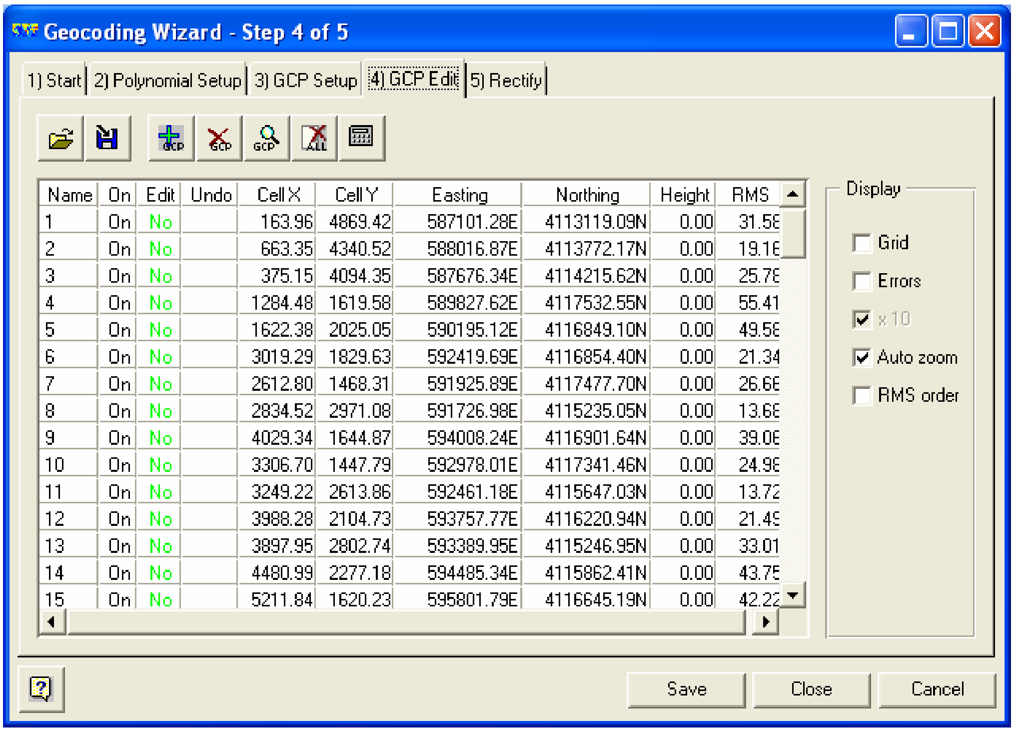




Name	On	Edit	Undo	Cell X	Cell Y	Raw X	Raw Y	Height	RMS
1	Off	Edit	Undo	2707.93	6031.30			0.00	
2	Off	Edit						0.00	
3	Off	Edit						0.00	
4	Off	Edit						0.00	
5	Off	Edit						0.00	
6	Off	Edit						0.00	
7	Off	Edit						0.00	
8	Off	Edit						0.00	

☐ Grid
☐ Errors
☑ x 10
☑ Auto zoom
☐ RMS order

The finished project should look something similar to the table below:



Name	On	Edit	Undo	Cell X	Cell Y	Easting	Northing	Height	RMS
1	On	No		163.96	4869.42	587101.28E	4113119.09N	0.00	31.5
2	On	No		663.35	4340.52	588016.87E	4113772.17N	0.00	19.1
3	On	No		375.15	4094.35	587676.34E	4114215.62N	0.00	25.7
4	On	No		1284.48	1619.58	589827.62E	4117532.55N	0.00	55.41
5	On	No		1622.38	2025.05	590195.12E	4116849.10N	0.00	49.5
6	On	No		3019.29	1829.63	592419.69E	4116854.40N	0.00	21.34
7	On	No		2612.80	1468.31	591925.89E	4117477.70N	0.00	26.6
8	On	No		2834.52	2971.08	591726.98E	4115235.05N	0.00	13.6
9	On	No		4029.34	1644.87	594008.24E	4116901.64N	0.00	39.0
10	On	No		3306.70	1447.79	592978.01E	4117341.46N	0.00	24.9
11	On	No		3249.22	2613.86	592461.18E	4115647.03N	0.00	13.72
12	On	No		3988.28	2104.73	593757.77E	4116220.94N	0.00	21.4
13	On	No		3897.95	2802.74	593389.95E	4115246.95N	0.00	33.01
14	On	No		4480.99	2277.18	594485.34E	4115862.41N	0.00	43.7
15	On	No		5211.84	1620.23	595801.79E	4116645.19N	0.00	42.22

☐ Grid
☐ Errors
☑ x 10
☑ Auto zoom
☐ RMS order

The data can then be used with other datasets such as SRTM to give perspective views or used with other 2D datasets to give change images.

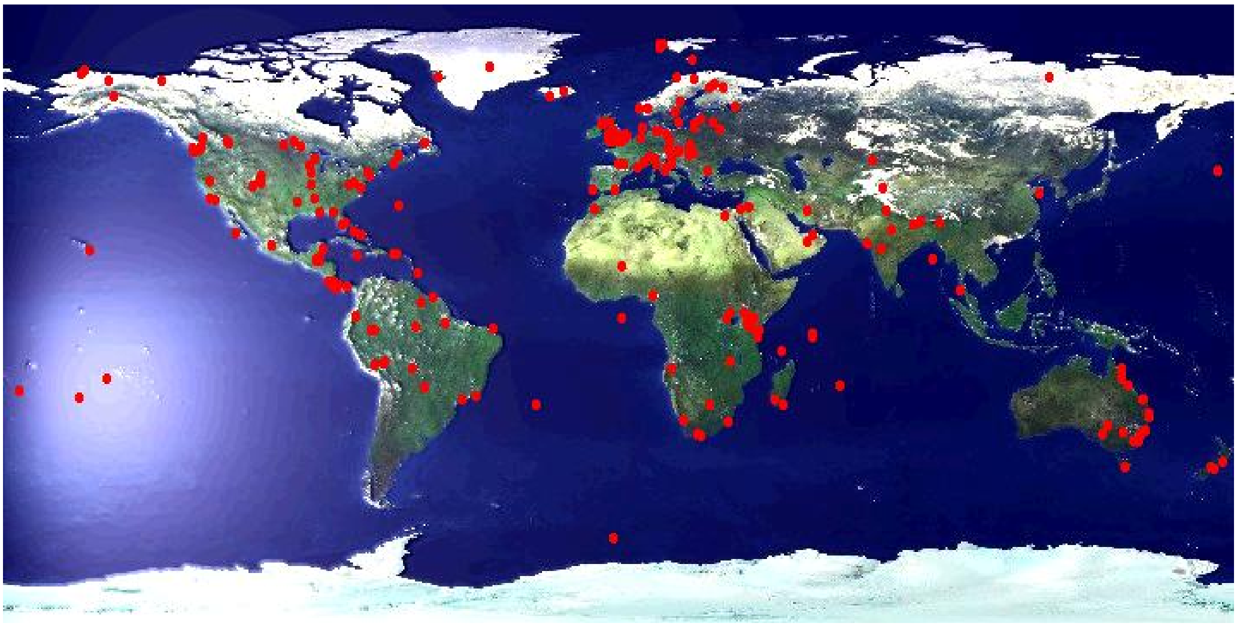
Appendix 3: World Register of Field Centres

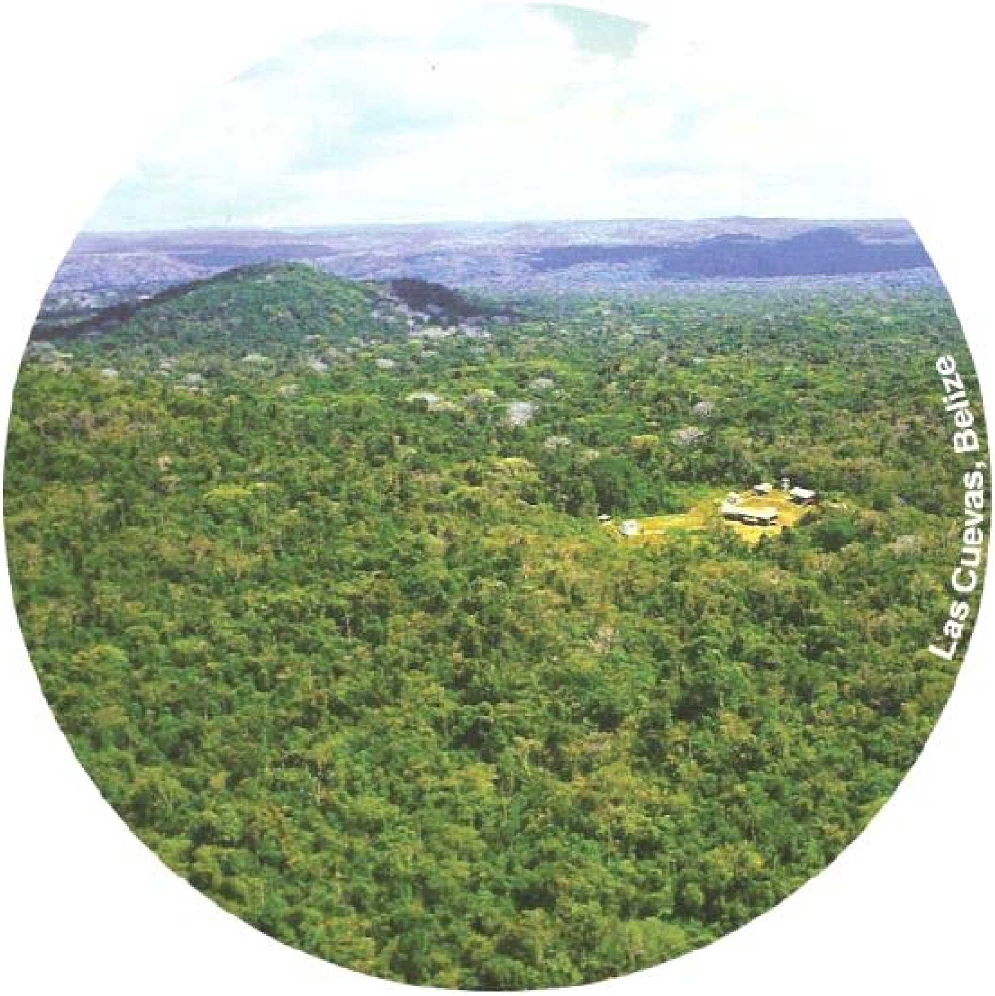
A helping hand in field based GISci.

The World Register of Field Centres (WRFC) is a project run by the Royal Geographical Society with IBG managed by Martin Whiteside in collaboration with the EAC and developed with funds from Professor Keith Miller of Sheffield University. The WRFC is located on the RGS website at *www.rgs.org/fieldcentres*.

The WRFC links both small independent field centres and long-term international facilities throughout the world that have the capability to welcome international visitors who wish to undertake fieldwork at the centre, whether scientists, students, teachers, school pupils or others. The WRFC centres do not necessarily have the facilities to host in-depth GISci research but the centres can be useful basecamps where analysis can be done and research co-ordinated.

The register currently has 377 centres online but is growing all the time. The database can be queried online by continent, country, biome and soon research interests. The figure below shows where some of the currently listed centres are located.





Las Cuevas, Belize

Appendix 4: Buyer’s guide to GPS units

This guide is designed for expeditions and is an attempt to summarise some of the key decisions a team should make when assessing the need for GPS. There are a wide variety of manufacturers producing several differing models of GPS receiver. Models are designed for specific budgets and applications. For information on individual units and their capabilities and applications, see the following manufacturer websites: *www.garmin.com*, *www.magellangps.com*, *www.lowrance.com*, and *www.brunton.com*. This guide has been complied as a checklist of everything required in an expedition GPS. A more in depth description of the features and technical terms used can be located in the main body of the text in either this chapter or in Chapter 6.

An excellent site filled with reviews of units which often gives you additional information not found in manuals is *www.gpsinformation.net*. Further queries if not answered there can be placed on the newsgroup *sci.geo.satellite-nav*. The newsgroup postings can be found by using the Google newsgroup search (*www.google.com*). It has an extremely comprehensive set of answers on a very wide range of GPS topics.

General work in the field with a variety of receiver types has highlighted a series of features that are very useful when collecting GPS co-ordinate information. For expeditions, we believe the following features ought to be considered.

Waypoint storage

As discussed in Chapter 6, waypoints are consciously gathered GPS co-ordinate points. Manufacturers differ in the amount of detail that they allow to be kept about the waypoint in internal memory storage. As an example some units may not store the altitude of a waypoint in the unit, merely display it on screen. This means that when you download the information into a GIS that you will have no height data. A further complexity may be that the GPS unit has an altimeter. Determining whether the unit downloads the height from the altimeter or the GPS calculation is very important.

Another potentially important point is the names and symbols associated with a waypoint. Units vary in the length of the name they allow for a waypoint. Waypoint names can be used as a code to relate information you are gathering in a field notebook. The length and type of waypoint name allowed can therefore be very important. Symbols are usually used to differentiate point features within a map or route display on a GPS unit. This symbol can also be downloaded with the associated waypoint. Using the GPS Utility software as an example, the downloaded symbol is given a name. This again can be used for helping add extra information about a waypoint.

The actual number of waypoints a unit can store also varies. Most modern units store 500 waypoints and the latest more expensive models can store up to 1000. The extra storage is handy if you will be gathering data for long periods of time or collecting large numbers of waypoints between downloads to a computer.

It cannot be emphasised enough that you must download and backup your data as often as you can. Do not use the excuse of having a large waypoint memory for not doing this.

Some units automatically average your position once you stop. This generally means that the longer that you stand at a location the closer to the true co-ordinate location you will be. With such units it is worth waiting several minutes before taking a waypoint fix. Units which do not average will create a meandering co-ordinate location, as the geometry of the satellites above changes over time, as well as being affected by the other atmospheric factors which can affect GPS accuracy. The only way to average a location with these units is by collecting a series of waypoints, which are then processed after download from the unit.

Trackpoints and tracklogs

Initially just designed for basic automatic tracking functions the tracklog is becoming increasing large and easy to control. Tracklogs essentially contain a breadcrumb trail of trackpoints generated as you walk along with the GPS receiver on. They are used in features such as track routes and are displayed on screen with units having map page screens to show where you have been. The tracklog is one area that has a wide range of differentiation between receiver types. Some allow control over how the tracklog is gathered but some do not. Those that do allow you to specify how often a trackpoint is generated generally allow either a time or distance setting. The time setting collects a trackpoint every x number of minutes and the distance setting collects a trackpoint every x number of meters. The amount of data stored in the trackpoint itself also varies. All units gather latitude and longitude, however, only some units gather height as well. This is a very important point if collection of height data is important to your project and is never normally stated on a manufacturer website.

Routes

Waypoints can be added to routes. Routes are paths you can create by linking selected waypoint. Properly created routes are useful in getting about in difficult or potentially dangerous terrains but their value is affected by how they are constructed. The number of waypoints allowed in each route can vary between units and this will affect how closely you can follow a generated route. A similar route or trackback feature can be generated from the active tracklog. A saved tracklog will again vary in the number of trackpoints it can hold between units. It is rarely as many points as found in the active log itself.

Area calculations from routes

A very useful feature found in some newer units is the ability to calculate the area of an object by generating a route around it. You tell the unit the start of the route, navigate about the object and then tell the unit to finish the route. The unit then calculates the area enclosed. Because of the positioning errors that accompany a GPS co-ordinate it is really only worth measuring the areas of fairly large objects. Area calculations from routes can also be accomplished in software once the route has been downloaded to a PC.

External antennae

This is often a less obvious consideration when you purchase a GPS unit but can be very important. Many of the budget GPS units that are well suited to expeditions do not allow connection to external antennae. This is an issue if the GPS does not have a clear sky view at all times. Using the GPS unit in a car is one example. Some windscreens can shield the GPS unit from receiving a signal and the signal itself can be weakened as it passes through a standard windscreen. Additionally, the shell of the car will stop GPS signals and the windscreen will only offer a small view of the sky. Larger positional fix errors are generated because of these issues. The best way to resolve this is to fit an external antenna to the vehicle. This can then receive a clear signal from all satellites available about the vehicle. If you still want a unit without an external antenna connection, it is now possible to buy a signal re-radiating antenna, which has a bolt on external section and an active retransmitting section to sit in the vehicle.

Maps

One of the big differences between units is the ability to store proprietary maps created by the GPS unit manufacturers. This is one of the big reasons for cost differences between units. The general problems with the maps provided or that can be purchased is that they tend to be country and city specific. America is obviously well covered but other parts of the world may well have no map information available at all. In nearly all units the extra memory areas of the units are only available for uploaded map data. That means that a unit with 1 Mb of memory can hold the same numbers of trackpoints/ waypoints etc as a unit with 24 Mb of memory. The only exception to this currently is the Lowrence ifinder GPS unit which allows trackpoints to be stored on its removable mmc card memory. This gives this unit the ability to store 10 tracks with up to 10,000 trackpoints each.

There is currently no official software product available to allow you to generate your own maps to upload to a GPS unit. Several third party packages are available to do just this but their use may invalidate the warranty of your GPS unit. Currently the only normal way to generate map style features is by uploading waypoints and trackpoints of features important to you as field markers with software such as GPS Utility.

Operating temperature range

Most units have similar operating temperature ranges between about 5°C and 50°C; this is generally due to the operating ranges of the LCD screens used in the unit. The operating range is fine for hot climates (apart from leaving the unit on any bare hot surfaces or out directly in the sun for several hours) but can be an issue in cold or mountainous environments. A GPS unit which operates at several degrees lower than most units is the Silva Multinavigator. It has more basic display screen than most units which functions below freezing and has an insulated case allowing operation down to several degrees below 0.

Altimeters/barometers and electronic compasses

A number of higher spec units come with an altimeter and electronic compass. The electronic compass is useful for gaining bearings when you are not moving. The compass in an ordinary GPS unit can only function when you move, as it relies on the change in co-ordinate position to generate a bearing. The altimeter/barometer is a very useful feature if accurate height determination is important. With accurate and regular calibrations at locations of known altitude the unit can be accurate to 3 m or so. This contrasts with average GPS calculated height errors of more than 20 m or so. As mentioned before it is important to know how this altimeter information is formatted when downloading height information into a GIS from the GPS unit. As an example, altitude derived from a single waypoint fix in a Garmin GPS76s model will use the barometric altitude. However, if you create your waypoint from an averaged position the unit uses the GPS calculated altitude, which is also averaged.

Battery life

Most models use 2 or 4 AA batteries. The actual battery life between models does vary by several hours with some units being much higher drainers of energy than others. Features such as the electronic compass on some units also consume power. The Silva multinavigator can get many days usage from its batteries if the GPS unit is used sparingly (navigation being mainly undertaken by bearings from the electronic compass). An easy way to conserve battery power is to use the GPS unit in a battery save mode so that the unit is not actively tracking satellites continuously, usually the tracking is then done every second or so. At low speeds this does not lead to positional errors.

External power connection

The ability to use a DC current other than the internal batteries is very useful, especially if the unit is to be used in a vehicle a lot of the time. Units vary in the range of voltages required. External power connection is also useful if a solar powered cell is connected to the unit. This will extend the life of batteries considerably.

Quality of signal reception by units

Another area of variability of units is their performance in areas of poor reception. The website *www.gpsinformation.net* carries a number of articles looking at antenna sensitivities and the variability can be important if your work is to be undertaken in areas of heavy plant canopies. Currently the most sensitive antenna that the website has found is in the Magellan Sportrak models. Signal reception can be improved for units that accept external antenna by placing an antenna up above the unit for an improved view of the sky.

Differential GPS ready

Some units can connect to differential GPS Beacons. Ideal if unit is being used on a survey with a local differential GPS base-station. As discussed earlier the use of a differential base-station system reduces positional errors to 3 m or so. The WAAS system means that this will not be strictly necessary in Europe or America but a lot of the world will be outside of any satellite based augmentation systems for some time to come. Also note that because the satellites broadcasting corrections are lower on the horizon at higher latitudes, it may not be feasible to rely on WAAS signals for differential corrections.

Firmware

Most GPS manufacturers now put the latest firmware updates for their units on their websites. It is often worth checking back to a manufactures website to see what fixes / extra functionality have been added to a unit’s software. Be wary of units that do not have much support available.

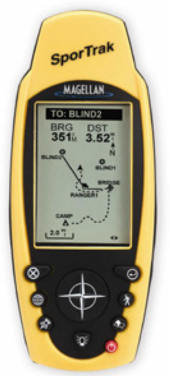
Common GPS Models

Over the next few pages examples of common GPS models are discussed. The prices are only accurate up to early 2005. New models and new specifications are becoming available all the time and the reader should check the various websites discussed here and in Chapter 11.

Older Magellan Models

Unit	**Magellan 310**	**Magellan 315**	**Magellan 320**	**Magellan 330**
Cost	£90 + VAT	£100 + VAT	£150 + VAT	£215 + VAT
Range		Magellan 300 Range		
Manufacturer		Magellan		
Manufacturer website		www.magellangps.com		
Status		Replaced by new Magellan models		
Key Components		Models 315 upwards are designed to float		
		Downloadable co-ordinates from all models *(though altitude not commonly downloaded on trackpoints)*		
		All units are 12 channel but not WAAS enabled		
Drawbacks		Very basic LCD interface that can be confusing		
		Model 310 does not report altitude when below sea level		

Newer Magellan Models



Unit	**SporTrak**	**SporTrak Map**	**SporTrak Pro**	**SporTrak Colour**
Cost	£125 + VAT	£150 + VAT	£190 +VAT	£250 + VAT
Range		Magellan SporTrak Range		
Manufacturer		Magellan		
Manufacturer website		www.magellangps.com		
Status		Current		
Key Components		North finder: electronic compass finds north even when GPS is switched off conserving battery time		
		Downloadable co-ordinates from all models		
		All units are 12 channel and WAAS enabled		
		Ruggedised construction offers protection and units floats		
Drawbacks		No external antenna socket		

Unit	**Meridian**	**Meridian Gold**	**Meridian Platinum**
Cost	£150 + VAT	£215 + VAT	£250 + VAT
Range		Magellan Meridian Range	
Manufacturer		Magellan	
Manufacturer website		www.magellangps.com	
Status		Current	
Key Components		North finder: electronic compass finds north even when GPS is switched off conserving battery time	
		Downloadable co-ordinates from all models	
		All units are 12 channel and WAAS enabled	
		Excellent screens for mapping	
		Magellan's top of the range receivers. Large memory for trackpoints	

Other Manufacturers

There are many different manufacturers of GPS units. Some more modern units are coming on the market with more powerful processors able to re-process weak signals and work in areas that some standard units would fail. The reader will have to investigate these for themselves. Many of these are CF cards for PDAs. A good unit manufactured by Silva is the MultiNavigator. This is described below.

Unit	**Silva Multi Navigator**	
Cost	£225 + VAT	
Range		Silva Multi Navigator
Manufacturer		Silva / Brunton
Manufacturer website		www.brunton.com
Status		Current
Key Components		Electronic compass finds north even when GPS is switched off conserving battery time
		Screen withstands cold temperatures down to -15°C.
		12 channel but not WAAS enabled
Known Problems		Very basic screen makes using the device more difficult.

Common Garmin Models

Unit	ETREX	ETREX Venture	ETREX Legend
Cost	£75 + VAT	£120 + VAT	£135 + VAT
Range		Garmin ETREX	
Manufacturer		Garmin	
Manufacturer website		http://www.garmin.com	
Status		Current	
Key Components		Water proof & rugged construction	
		Downloadable co-ordinates from all models	
		All units are 12 channel and Venture upwards (excluding Summit) WAAS enabled. Venture upwards (excluding Summit) ships with free data cable	
		Small and light weight and easily operated by one hand	
Known Problems		No external antenna port, not ideal for vehicle based expeditions	

Unit	GPS 12	GPS 12 XL
Cost	£100 + VAT	£150 + VAT
Range	Garmin GPS 12	
Manufacturer	Garmin	
Manufacturer website	http://www.garmin.com	
Status	Obsolete	
Key Components	Splash proof casing	
	Highly configurable for user defined datums etc.	
	Units are 12 channel but not WAAS enabled and can host an external antenna	

Unit	**GPS 72**	**GPS 76**	**GPS 76 Map**
Cost	£120 +VAT	£150 + VAT	£200 + VAT
Range		GPS 70	
Manufacturer		Garmin	
Manufacturer website		http://www.garmin.com	
Status		Current :- Garmin replacements for GPS 12 series	
Key Components		Water proof but not particularly rugged	
		All units 12 channel & WAAS enabled	
		Large screen and user configurable datums. Very useful if maps are being used within the GPS, either purchased maps or maps created by the user from trackpoints and waypoints.	
		Small and light weight and easily operated by one hand	
Known Problems		No external antenna port on the GPS 72	

Unit	**GPS 3**
Cost	£250 + VAT
Range	Garmin GPS 3
Manufacturer	Garmin
Manufacturer website	www.garmin.com
Status	Older model
Key Components	Mapping software and large screen
	Large screen and user configurable datums. Very useful if maps are being used within the GPS, either purchased maps or maps created by the user from trackpoints and waypoints.
	12 channel but not WAAS enabled
Known Problems	No WAAS and unit is a 4 battery receiver

Unit	**Geko 101**	**Geko 201**	**Geko 301**
Cost	£75 +VAT	£90 + VAT	£250 + VAT
Range		Garmin Geko	
Manufacturer		Garmin	
Manufacturer website		http://www.garmin.com	
Status		Current	
Key Components		Light weight and small units (smaller than ETREX range)	
		All units 12 channel & WAAS enabled	
		Very large number of trackpoints, highest of any GPS receiver	
Known Problems		Lower models not ideal for expedition work	
		101 unit can not download and does not have an external antenna	
		Run on AAA batteries and battery life is shorter than on other units	

GPS images were downloaded from GPS Warehouse website at *www.gpsw.co.uk*

The reader should check online to ensure the receivers have the configuration you require. The above list is not a definitive list of all models or all of their features. There are other GPS receivers available and new models are becoming available all the time. Magellan/Thales released a small unit similar to the Garmin Geko in July 2004 called the eXplorist. The eXplorist will come in three varieties, the 100, 200 and 300 also similar to the Geko 101, 201 and 301. The eXplorist were the first commercially available 14-channel receiver and are fully WAAS enabled. However, similar to the Geko they are probably not best suited to expedition work and better for walking, hiking and less professional activities.

References & bibliography

Adinarayana, J., Flach, J.D. & Collins, W.G. 1994 Mapping land use patterns in a river catchment using GIS. *Journal of Environmental Management*, 42(1), 55-61.

Allison, R.J. 2001 Surface and groundwater resources survey in Jordan. In: J.S.Griffiths (ed.) *Land surface evaluation for engineering practice*. Geological Society, London, Engineering Group Special Publication, 18, 65-71.

Allum, J.A.E. 1966 *Photogeology and regional mapping*. Pergammon Press, Oxford.

Anon. 1971 *Handbook for expeditions.* Brathay Exploration Group / Geographical Magazine, 137pp.

Anon. 1985 *Airphoto interpretation for urban studies and physical planning*. GTZ Sonderpublikation No. 230, GTZ GmbH, Eschborn, Germany, 71pp.

Anon. 2001 Looking for lost cities with infra-red. *GeoEurope*, 10 (11), 35.

Avery, M.I., & Haines-Young, R.H. 1990 Population estimates for the dunlin *Caliadris alpina* derived from remotely sensed satellite imagery of the Flow Country, northern Scotland. *Nature*, 344, 860-862.

Avery, T.E. & Berlin, G.L. 1992 *Fundamentals of Remote Sensing and Airphoto Interpretation*. 5th Edition. Prentice Hall.

Bajjouk, T., Populus, J., & Guillaumont, B. 1998 Quantification of subpixel cover fractions using Principal Component Analysis and a linear programming method: application to the coastal zone of Roscoff (France). *Remote Sensing Environment,* 64, 153-165.

Baker, M.C.W. & Baldwin, J.A. 1981 Application of Landsat multispectral classification for locating gossans in north Chile. In: J.A. Allan & M. Bradshaw (eds) *Remote sensing in geological and terrain studies*. Remote Sensing Society, London, 25-33.

Barnes, J. 1995 *Basic geological mapping*. 3rd edition.

Beaumont, E.A. & Foster, N.H. (compilers) 1992. *Remote Sensing* (Treatise of Petroleum Geology Reprint Series), American Association of Petroleum Geology

Berger, Z. 1996 *Remote sensing and petroleum exploration*.

Breimer, R.F., van Kekem, A.J. & van Reuler, H. 1986 *Guidelines for soil survey and land evaluation in ecological research.* MAB Tech. Notes 17, UNESCO, Paris.

Brooker, S., Beasley, M., Ndinaromtan, M., Madjiouroum, E.M., Baboguel, M., Djenguinabe, E., Hay, S.I., & Bundy, D.A.P. 2002 Use of remote sensing and a geographical information system in a national helminth control programme in Chad. *Bulletin of the World Health Organisation*, 80 (10), 783-789.

Brunsden, D., Doornkamp, J.C., Fookes, P.G., Jones, D.K.C., & Kelly, J.M.H. 1975 Large scale geomorphological mapping and highway engineering design. *Quarterly Journal of Engineering Geology*, 8, 227-253.

Burrough, P.A. (1986) *Principles of GIS for land resources assessment*. Clarendon Press, Oxford, 194pp.

Burrough, P.A. & MacDonnell, R.A. 1998 *Principles of Geographical Information Systems* Oxford University Press, 333pp.

Cain, D.H., Ritters, K., & Orvis, K. 1997 A multi-scale analysis of landscape statistics. *Landscape Ecology*, 12, 199-212.

Carpenter, S.R., Brock, W.A., & Hansen, J. 1999 Ecological and social dynamics in simple models of ecosystem management. *Conservation Ecology*, 3, 121-128.

Carrol, D.M., Evans, R. & Bendelow, V.C. 1976 Air photo-interpretation for soil mapping. *Soils Survey Technical Monograph* No.8, Harpenden, 85pp.

Carver, S.J., Cornelius, S.C., Heywood, D.I., & Sear, D.A. 1995 Using computers and GIS for expedition fieldwork. *The Geographical Journal*, 161(2), 167-176.

Chambers. R. 1994 The origins and practice of participatory rural appraisal. *World Development*, 7, 953-969.

Chaussier, J-B., & Morer, J. 1992 *Manual du prospecteur minier*. Editions de BRGM, Orleans, France, 272pp.

Chief Directorate: Surveys & Mapping 2003 *Map Projections used in South Africa*. Published at: *w3sli.wcape.gov.za/Surveys/Mapping/mapprojn.htm*

Chorley, R. 1987 *Handling Geographic Information*. UK Department of the Environment, HMSO, London.

Christian, C.S., and Stewart, G.A. 1952 Summary of general report on survey of Katherine-Darwin Region, 1946. *Land Research Series*, 1, CSIRO, Australia, 24pp.

Clarke, R. 1986 *The Handbook of Ecological Monitoring*. Oxford University Press, 298pp.

Clarke, S. E., White, D. and Schaedel, A. L. 1991. Ecological regions and subregions for water quality management. *Environ. Manag*. 15: 847–856.

Clifford, N.J., & Valentine, G. (Eds.) 2003 *Key methods in geography*. Sage Publications, London, 572pp.

Coe, M.J, McWilliam, N.C., Stone, G.N., & Packer, M.J. (eds.) 1999 *Mkomazi: the ecology, biodiversity and conservation of a Tanzanian savanna.* RGS-IBG, London, 608pp.

Cooke, R.U. & Doornkamp, J.C. 1990 *Geomorphology in environmental management*. Cambs. Univ. Press, 410pp.

Cornelius, S.C., Sear, D.A., Carver, S.J. & Heywood, D.I. 1994 GPS, GIS and geomorphological fieldwork. *Earth Surface Process and Landforms*, 19, 777-787.

Crosta, A.P. & Moore, J.McM. 1989 Geological mapping using Landsat TM imagery in Almeria Province, SE Spain. *International Journal of Remote Sensing,* 10, 505-514.

DeMers, M.N., 2000 *Fundamentals of GIS*, Wiley, New York.

Derion, JP., Motti, E., & Simonin, A. 1998 A comparison of the potential for using optical and SAR data for geological mapping in an arid region: the Atar site, western Sahara, Mauritania. *International Journal of Remote Sensing*, 19, 1115-1132.

De Souza Filho, C.R., & Drury, S.A. 1998 Evaluation of JERS-1 OPS and Landsat TM images for mapping of gneissic terrain in arid areas. *International Journal of Remote Sensing*, 19, 3569-3594.

Dominy, N. J. & Duncan, B. 2001 GPS and GIS methods in an African rain forest: applications to tropical ecology and conservation. *Conservation Ecology* 5(2), 6. Available at: *www.consecol.org/vol5/iss2/art6*

Drury, S.A. 1987 *Image interpretation in geology*. Allen & Unwin, London (1st edition)

Drury, S.A. 2001 *Image interpretation in geology*. Allen & Unwin, London, 243pp. (3rd edition)

Dunn, C.E., Atkins, P.J. and Townsend, J.G. 1997 GIS for development: a contradiction in terms? *Area* 29, 151-159.

Ebdon, C. 1985 *Statistics for geographers.*

Engman, E.T., and Gurney, R.J. 1991 *Remote sensing in Hydrology*. Chapman & Hall, London.

Faulkner, H, Alexander, R , Teeuw R.M. and Zukowskyj, P.M. (2004) Variations in soil dispersivity across a gully head displaying shallow subsurface pipes. *Earth Surface Processes and Landforms,* 29 (9), 1143-1160.

Fookes, P.G. 1997 The First Glossop Lecture: Geology for Engineers: the Geological Model, prediction and Performance. *Quarterly Journal of Engineering Geology*, 30, 293-424.

Fookes, P.G., Dale, S.G., & Land, J.M. 1991 Some observations on a comparative aerial photography interpretation of a landslipped area. *Quarterly Journal of Engineering Geology*, 24, 249-265.

Fuller, R.M., Groom, G.B., & Jones, A.R. 1994 The land cover map of Great Britain: an automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, 60: 553-562.

Furley, P.A., Dargie, T.C.D., & Place, C.J. 1994 Remote sensing and the establishment of a GIS for resource management on and around Maraca Island. In: J.Hemming (ed.) *The rainforest edge*. Manchester University Press, 115-133.

Gibson, P. & Power, C. 2000. *Introductory remote sensing: digital image processing and applications*. Routledge, London.

Gillespie, M., and Culling, A. 1998 Is the pencil mightier than the computer ? *Mapping Awareness*, 12(3), 28-30.

Gillman, M.J. & Teeuw, R.M. 1996 Use of butterflies and GIS to assess biodiversity. *In*: Jermy, Long, Sands, Stork and Winser (eds.) *Guidelines on Biodiversity Assessment*. UK Department of the Environment, HMSO, London, 255-6.

Goodrich, D.C., Schmugge, T.J., Jackson, C.L., Unkrich, T.O., Parry, R., Bach, L.B., & Amer, S.A. 1994 Runoff simulation sensitivity to remotely sensed initial soil water content. *Water Resources Research,* 30 (5), 1393-1405.

Goudie, A. (ed.) 1990 *Geomorphological techniques.* Unwin Hyman, 570pp.

Griffiths, J.S. (ed.) 2001 *Land surface evaluation for engineering practice*. Geological Society, London, Engineering Geology Special Publications, 18, 248pp.

Gustafsson, P., 1993 High resolution satellite imagery and GIS as a dynamic tool in groundwater exploration in a semi-arid area. In: *HydroGIS 93 Conference Proceedings*, IAHS Publ. No. 211, 93-100.

Gustafsson, P., 1993 *Satellite data and GIS as a tool in groundwater exploration in a semi-arid area*. Geology Dept Publ A 74, Chalmers Technical University, Goteborg, Sweden (ISSN 0348-2367)

Hamilton, S.K., Sippel, S.J. & Melack, J.M. 1996. Inundation patterns in the Pantanal wetland of South America determined from passive microwave remote sensing. *Archiv für Hydrobiologie* 137/1: 1-23.

Hatton, W. 2003 Geoscience Information Systems – technologies and techniques for solving geoscientific problems – "Quo vadunt – quo vadis? *Proceedings Annual Conference of the International Association for Mathematical Geology*, Portsmouth, Sept 2003

Hewitt, M.J. 1990 Synoptic inventory of riparian ecosystems: the utility of Landsat TM data. *Forest Ecology & Management*, 33/34, 605-620.

Heywood, I., Cornelius, S., & Carver, S. 1998 *An introduction to Geographical Information Systems*. Longman, London.

Hockey, B., Richards, T., *et al.* 1990 Prospects for operational remote sensing of surface water. *Remote Sensing Reviews*, 4, 265-283.

Hobbs, F. 2003 Positioning The Maldives by handheld GPS. *Geomatics World* 11(4). Full text available at: *www.pvpubs.com/read_article.asp?ID=13& article_id=102*

Howard, A.D., 1967, Drainage analysis in geologic interpretation: a summation: *The Amer. Assoc. of Petr. Geol.*, v. 51, n. 11, p. 2246-2259.

Hung,L.Q., Dinh, N.Q.,Batelaan, O., Tam, V.T., & Lagrou, D. 2002 Remote sensing and GIS-based analysis of cave development in the Suoimuoi Catchment (Son La - NW Vietnam). *Journal of Cave and Karst Studies 64*(1): 23-33.

Iliffe J.C. 2000 *Datums and Map Projections*. Whittles Publishing, Caithness. Ordering details at: http://www.whittlespublishing.com/engineer/datums.html.

Jensen, J.R., Rutchey, K., Koch, M.S. & Narumalani, S. 1995 Inland wetland change detection in the Everglades Water Conservation Area using a time series of normalized remotely sensed data. *Photogrammetric Engineering and Remote Sensing* 61:199-209.

Jermy, C., Long, Sands, Stork & Winser (eds.) 1996 *Guidelines on Biodiversity Assessment*. UK Department of the Environment, HMSO, London, 255-6.

Jin, Y-Q. 1999 A flooding index and its regional threshold value for monitoring floods in China from SSM/I data. *International Journal of Remote Sensing* 20: 1025-1031.

Jochem, R., Kooistra, B.J.H., & Chardon, P. 2002 LARCH: an ecological application of GIS in river studies. *In*: R.S.E.W Leuven, I.Poudevigne and R.M.Teeuw (eds.) *Application of Geographic Information Systems and remote sensing in river studies.* Backhuys Publishers, Leiden.

Johnson, C.A. 1998 *GIS in ecology*. Blackwell, Oxford.

Jones, A. 1999 *Where in the World are We?* Government of South Australia, Adelaide. Full text available at: *www.environment.sa.gov.au/mapland/sicom/sicom/publications.html*.

Jorge, L.A.B., & Garcia, G.J. 1997 A study of habitat fragmentation in South-eastern Brazil using remote sensing and GIS. *Forest Ecology & Management*, 98(1), 35-47.

Jorgensen, A.F. & Nohr, H. 1996 The use of satellite images for mapping of landscape and biological diversity in the Sahel. *International Journal of Remote Sensing* 17: 91-109.

Kapila, S., & Lyon, F. 1994 Expedition field techniques: people orientated research., EAC, Royal Geographical Society (with IBG), London, 70pp.

Kasischke, E.S., Melack, J.M., & Dobson, M.C. 1997. The use of imaging radars for ecological applications – a review. Remote Sensing of Environment, 59: 141-156.

Keay, W. 2000 *Expedition Guide* (Duke of Edinburgh's Award), Sterling Press, Wellingborough, UK.

Khan, M.A., Fadlallah, Y.H., & Al-Hinai, K.G., 1992 Thematic mapping of subtidal coastal habitats in the Western Arabian Gulf using Landsat TM data. *International Journal of Remote Sensin*g, 13(4), 605-614.

Kliskey, A.D., Lofroth, E.C., Thompson, W.A., Brown, S., & Schriever, H. 1999 Simulating and evaluating alternative resource-use strategies using GIS-based habitat suitability indices. *Landscape & Urban Planning*, 45, 162-175.

Knippers, R. (2000) *Geometric Aspects of Mapping: Reference Surfaces for Mapping*. Published online at: *kartoweb.itc.nl/geometrics/index.asp*

Koblinsky, C.J., R.T. Clarke, A.C. Brenner, & Frey, H., 1993 Measurement of river level variations with satellite altimetry, *Water Resources Research,* 29(6), 1839–1848.

Komp, K-U., 1991 Remote sensing: tools for development (Fernerkundung: instrument fur die entwicklungszusammenarbeit), GTZ GmbH, Eschborn, Germany, 63pp.

Kuntz, 1995 ERS SAR for mapping rainforest in Indonesia. ESA Journal.

Landon, J.R. (ed.) (1991) *Booker tropical soil manual*. Longmans, Harlow, 474pp.

Lattman, L.H. & Parizek, R.R., 1964 Relationship between fracture traces and the occurrence of ground water. *Journal of Hydrology*, 2, 73-91.

Lauver, C.L. 1997 Mapping species diversity patterns in the Kansas shortgrass region by integrating remote sensing and vegetation analysis. *Journal of Vegetation Science* 8: 387-394.

Lauver, C.L. & Whistler. J.L. 1993 A hierarchical classification of LANDSAT TM imagery to identify natural grassland areas and rare species habitat. *Photogrammetric Engineering &. Remote Sensing* 59: 627-634.

Lawrance, C.J., Bayard, R.J. & Beaven, P.J. (1993) Terrain Evaluation Manual State of the Art Review 7, Transport Research Laboratory, HMSO, 285pp.

Leake, A. 1995 *Mapping Amerindian lands in the Gran Chaco, Paraguay*. PhD. (unpublished), University of Hertfordshire.

Leuven, R.S.E.W., Poudevigne, I. & Teeuw, R.M. (eds.) 2000 *Application of Geographic Information Systems and remote sensing in river studies.* Bakhuys Publishers, Leiden.

Lillesand, T.M., and Kiefer, R.W. 2000 *Remote sensing and image interpretation*. J Wiley & Sons, New York, 750pp.

Loffler, E. & Margules, C. 1980 Wombats detected from space. *Remote Sensing of Environment,* 9, 47-56.

Longley, P., Goodchild, M., Maguire,D. & Rhind, D. 2001 *Geographic Information Systems and Science*, Wiley.

Loughlin, W.P. 1990 Geological exploration in the western United States by use of airborne scanner imagery. In: *Remote sensing: an operational technology for the mineral and petroleum industries*. Institution of Mining & Metallurgy, London, 223-245.

Lyzenga, D.R. 1981 Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. *International Journal of Remote Sensing*, 2(1), 71-82.

Maling, D.H. 1992 *Co-ordinate Systems and Map Projections*. Pergamon Press Ltd., Oxford.

Mapedza, E., Wright, J. & Fawcett, R. 2003 An investigation of land cover change in Mafungausti Forest, Zimbabwe, using GIS and participatory planning. *Applied Geography*, 23(1), 1-21.

Mather, A.E., Griffiths, J.S. & Stokes, M. 2003 Anatomy of a ‘fossil’ landslide from the Pleistocene of SE Spain. *Geomorphology*, 50, 135-149.

Mather, P.M. 1999 *Computer processing of remotely-sensed images*. Wiley, Chichester, 292pp.

Mather, R, de Boer, M., Gurung, M. & Roche, N. 1998 *Aerial photographs and ‘photo-maps’ for community forestry*. Rural Development Forestry Network, London, Paper 2e.

McCauley, J.F., Schaber, G.G., Breed, C.S., Grolier, M.J., Haynes, C.V., Issawi, B., Elachi, C., & Blom, R.G. 1982 Subsurface valleys and geoarcheology of the eastern Sahara revealed by Shuttle radar. *Science*, 218, 1004-1020.

McKim, H.L., Cassell, E.A and LaPontin, P.J. 1993 Water resource modelling using remote sensing and object-orientated simulation. *Hydrological Processes,* 7, 153-165.

Meadon, G.J. & Kapetsky, J.M. 1991 *Geographical Information Systems and remote sensing in inland fisheries and aquaculture*. FAO Tech. Paper 318, Rome, 262pp.

Meijerink, A. 1988 Data acquisition and data capture through terrain mapping units. *ITC Journal*, 1988-1, 23-44.

Mertes, L.A.K., Daniel, D.L., Melack, J.M., Nelson, B., Martinelli, L.A, & Forsberg, B.R. 1995. Spatial patterns of hydrology, geomorphology and vegetation on the floodplain of the Amazon River in Brazil, from a remote sensing perspective. *Geomorphology*, 13: 215-232.

Mertes, L.A.K. 2002 Remote sensing of freshwater riverine landscapes: an update. *Freshwater Biology*, 47 (4), 799-816.

Michalek, J.L, Wagner, T.W., Luczkovich, J.J., & Stoffle, R.W. 1993 Multispectral change vector analysis for monitoring coastal marine environments. *Photogrammetric Engineering & Remote Sensing.*

Mikhail, E.M., Bethel, J.S., McGlone, J.C. & McGlone, C. 2001. *Introduction to Modern Photogrammetry.* John Wiley & Sons

Middelkoop, H. 2000 Application of remote sensing and GIS-based modelling in the analysis of floodplain sedimentation. *In:* R.S.E.W.Leuven, I. Poudevigne and R.M.Teeuw (eds.) 2000 *Application of Geographic Information Systems and remote sensing in river studies.* Bakhuys Publishers, Leiden.

Mikhail, E.M., Bethel, J.S., McGlone, J.C. & McGlone, C. 2001. *Introduction to Modern Photogrammetry.* John Wiley & Sons

Miller, K. 1993 *Simple Surveying.* Expedition Field Techniques, Expedition Advisory Centre, Royal Geographical Society (with IBG), London, 50pp.

Millington, A.C., Walsh, S.J., & Osborne, P.E. 2001 GIS and remote sensing applications in biogeography and ecology. Kluwer.

Milton, E.J., Gilvear, D.J., & Hopper, I.D. 1995 Investigating change in fluvial systems using remotely sensed data. *In:* A.M.Gurnell and G.E.Petts (eds.) *Changing river channels*. John Wiley & Sons, 277-298.

Mitchell, C. 1991 *Terrain evaluation*. Longman, Harlow, 141pp.

Moore *et al*, 2000, *Applied Tectonic Geomorphology for Diamond Prospecting in the Tarim Basin, Xinjiang China Using Combined DEM and enhanced Imagery*, 14th International Conference on Applied Geological Remote Sensing. November 2000

Moseley, F. 1981 *Methods in field geology*. Freeman & Co., Oxford, 211pp.

Mott, P. 1978. The use of Landsat imagery by small expeditions. *The Geographical Journal*, 144 (2) 271-278.

Mount, N.J. & Rumbachs, R. (2002) High Altitude GIS. *GIS User* 51: 48-51.

Mount, N.P, Zukowskyj, P.M, Teeuw, R.M., & Stott, T.A. 2000 Assessing river channel destabilisation, using aerial photography, digital photogrammetry and GIS. *In*: R.S.E.W Leuven, I.Poudevigne and R.M.Teeuw (eds.) *Application of Geographic Information Systems and remote sensing in river studies.* Backhuys Publishers, Leiden.

Mount, N.J., Louis, J., Teeuw, R.M., Zukowskyj, P.M. & Stott, T. 2003 Estimation of bankfull width comparisons from temporally sequenced raw and corrected aerial photographs. *Geomorphology*, 56, 65-77.

Muchoney, D.M., Grossman, D.H., & Solomon, R., 1991 Rapid ecological assessment for conservation planning. *Procs. ACSM / ASPRS annual convention*, Baltimore, Vol. 4: GIS, 141-145.

Muller, E., Decamps, H., & Dobson, M.K. 1993 Contribution of space remote sensing to river studies. *Freshwater Biology,* 29, 301-312.

Mumby, P.J., D.A. Gray, J.P. Gibson & P.S. Raines. 1995. Geographic Information Systems: a tool for integrated coastal zone management in Belize. *Coastal Management* 23(2): 111-121.

Myers, N. 1988 Tropical deforestation and remote sensing. *Forest Ecology & Management,* 23, 215-225.

Munyati, C. 2000 Wetland change detection on the Kafue Flats, Zambia, by classification of multitemporal remote sensing image dataset. *International Journal of Remote Sensing* 21: 1798-1807.

Nash, D.J. 2000 Doing independent overseas fieldwork 1: practicalities and pitfalls. *Journal of Geography in Higher Education,* 24(1), 139-149.

Nash, D.J. 2000 Doing independent overseas fieldwork 2: getting funding. *Journal of Geography in Higher Education,* 24(3), 425-433.

Nemani, R., & Running, S.W. 1996 Implementation of a hierarchial global vegetation classification in ecosystem function models. *Journal of Vegetation Science*, 7, 337-346.

Nunny, R, 1996 Belize sediment in Web surf. In: Teeuw, R.M. (ed.) *Procs. Workshop on Digital Fieldwork*, EAC/RGS, London, 49pp.

OAS, 1991 *Primer in natural hazard management in integrated regional development planning.* Organisation of American States, Washington, DC.

Ormerod, S.J., & Watkinson, A.R. 2000 Large-scale ecology and hydrology: an introductory perspective. *Journal of Applied Ecology*, 37, 1-5.

Paul, F. 2000 Evaluation of different methods for glacier mapping using Landsat TM. *Proceedings of EARSeL-SIG-Workshop Land Ice and Snow, Dresden, June 2000*. EARSeL Proceedings No.1, 239.

Pearson, D.M. 2002 The application of local measures of spatial autocorrelation for describing pattern in north Australia landscapes. *Journal of Environmental Management*, 64(1) 85-95.

Pereira, M.J. and Moura, D. 1999 Mapping the spatial distribution of waders on south Tagus riverside. *Remote Sensing Society Newsletter*, 100 (4), 18-22.

Petrie, G. 1999 Characteristics and applications of high-resolution space imagery. *Mapping Awareness,* 13 (10): 33-37.

Petrie, G. 2000 Transportable ground receiving stations. *GeoInformatics*, 3, 18-21.

Phinn, S.R., Menges, C., Hill, G.J.E., & Stanford, M. 2000 Optimising remotely sensed solutions for monitoring, modeling and managing coastal environments. *Remote Sensing of Environment.* 73: 117-132.

Pickles, J. (ed.) 1995 Ground truth: the social implications of geographic information systems. Guilford Publications, New York.

Pope, K., Rejmankova, E., and Savage, H. 1994 Remote sensing of tropical wetlands for malaria control in Chiapas, Mexico. *Ecological Applications*, Vol.1, No.4, Department of Preventive Medicine and Biometrics, University of Health Services, Bethesda, 81-90.

Reichhardt, T. 1996 Environmental GIS: the world in a computer. *Environmental Science & Technology*, 30 (8), 340-343.

Reid, E., Mortimer, G.N., Lindsay, R.A., & Thompson, D.B.A. 1993 Blanket bogs of Great Britain: an assessment of large scale pattern and distribution using remote sensing and GIS. In: P.J.Edwards *et al*. (eds.) Large scale ecology and conservation biology. Procs. 35th BES symposium, Southampton, 1993. Blackwell Scientific, 229-246.

Reid, I. 2003 Making observations and measurements in the field: an overview. In: Clifford & Valentine (Eds.) *Key methods in geography*, Sage Publications, London, 209-222.

Rengers, N. & Soeters, R. 1980 Regional geological mapping from aerial photographs. *Bulletin International Association of Engineering Geologists*, 21, 103-111.

Rengers, N., Soeters, R. & Van Western, C.J., 1992 Remote sensing and GIS applied to mountain hazard mapping. *Episodes,* 15(1), 36-45.

Rigaux, Scholl, & Voisard, 2002, *Spatial Databases, With Applications to GIS*. Morgan Kaufmann Publishers.

Rice, S. 2003 Sampling in geography. In: Clifford & Valentine (Eds.) *Key methods in geography*, Sage Publications, London, 223-247.

Ritchie, W., Tait, D.A., Wood, M., & Wright, R. 1977 *Mapping for field scientists.* David & Charles, Newton Abbot, 327pp.

Rogerson, P.A. 2001 *Statistical methods for geography.* Sage Publications.

Roman, S. 1999 *Access database design and programming (2 ed.).* O'Reilly and Associates, USA.

Rowan, L., S. Hook, M. Abrams & J. Mars 2003 Mapping hydrothermally altered rocks at Cuprite, Nevada using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *Economic Geology*, 98(5), 1019-1027.

Sabins, F. F. 1987 *Remote sensing principles and applications*. Freeman & Co., New York, 449pp.

Sader, S.H., Stone, T.A., & Joyce, A.T. 1990 Remote sensing of tropical forests: an overview of research and applications using non-photographic sensors. *Photogrammetric Engineering & Remote Sensing,* 56(10), 1343-1351.

Sander, P. 1996 *Remote sensing and GIS for groundwater assessment in hard-rock areas*. Geologiska Institutionen, Chalmers University of Technology, Goteborg, Publ.A80,

Schultz, G.A. 1993 Hydrological modelling based on remote sensing information. *Advances in Space Research*, 13 (5), 149-166.

Schweizer, D., Armstrong, R. & Posada, J.M. (2005) Remote sensing characterization of benthic habitats and submerged vegetation biomass in Los Roques Archipelago National Park, Venezuela. *Internaltional Journal of Remote Sensing* 26(12): pp. 2657-2667

Sebastini, M., Sambrano, A., Villamizar, A., & Villalba, C. 1989 Cumulative impact and sequential geographical analysis as tools for land use planning: a case study: Laguna La Reina, Miranda State, Venezuela. *Journal of Environmental Management*, 29, 237-248.

Sheppard, T. 1998 *Vehicle-dependent expedition guide*. Desert Winds.

Simmonds, P. 2004 Maps, navigation and GPS, in: *Expedition Handbook 2004.* Expedition Advisory Centre, RGS-IBG, London.

Smith, M.O., Ustin, S.L. Adams, J.B., & Gillespie, A.R. 1990 Vegetation in deserts: a regional measure of abundance from Multispectral images. *Remote sensing of Environment*, 31, 1-26.

Soller, D.R. 1999 *Digital mapping techniques '99 – workshop proceedings*. US Geological Survey Open Report 99-386, 214pp.

Stoms, D.M. & Estes, J.E. 1993 A remote sensing research agenda for mapping and monitoring biodiversity. *International Journal of Remote Sensing* 14: 1839-1860.

Teeuw, R.M. (ed.) 1990 The Luangwa Soil Conservation Project. *Shell Environmental Papers*, I, Royal Geographical Society, London, 25pp

Teeuw, R.M. 1995 Low cost remote sensing and GIS for groundwater exploration in northern Ghana. *Hydrogeology Journal,* 3 (3), 21-30.

Townsend, P.A, & Walsh, S.J., 1998. Modelling floodplain inundation using an integrated GIS with radar and optical remote sensing. *Geomorphology,* 21: 295-312.

Tucker, C.J., Hielkma, J.U., & Roffley, J. 1985 The potential of satellite remote sensing of ecological conditions for survey and forecasting of desert locust activity. *International Journal of Remote Sensing,* 6, 127-138.

Tufte, E.R. 1983 *The visual display of quantitative information*. Graphics Press. Cheshire.

Van Hengel, W., & Spitzer, D., 1991 Multi-temporal water depth mapping by means of Landsat TM. *International Journal of Remote Sensing,* 12(4), 703-712.

Van Western, C.J. 1992 *Application of GIS to landslide hazard zones.* ITC Publication No. 15, Part 1, 245pp.

Vaughan, I.P. & Ormerod, S.J. (2005) The continuing challenges of testing species distribution models. *Journal of Applied Ecology* 42(4): pp. 720-730

Verstappen, H.T., & van Zuidam, R.A. 1975 *ITC system of geomorphological survey*. ITC Textbook VII-2, ITC, Enschede, 52pp.

Vincent, R.K. 1997 *Fundamentals of Geological and Environmental Remote Sensing.* Prentice Hall, 366pp.

Wadsworth, R. & Treweek, J. 1999 *Geographical Information Systems for ecology: an introduction*. Longman, Harlow.

Walsh, S.J., Butler, D.R., & Malanson, G.P. 1998. An overview of scale, pattern, process relationships in geomorphology: a remote sensing and GIS perspective. *Geomorphology*, 21: 183-205.

Whiteside et al (a) 2001 Report to the RGS-IBG, *Bogda Shan Expedition, Volume 1.*

Whiteside et al (b) 2001 Report to the RGS-IBG, *Bogda Shan Expedition, Volume 2.*

Whiteside 2002 unpublished DIC Thesis: *The Integration and Analysis of Digital Images and Field Data: 2 North China Case Studies*. Imperial College London, Royal School of Mines.

Winser, S. (2004) (ed) *Expedition Handbook*. Expedition Advisory Centre, RGS-IBG, London.

Winterbottom, S.J., & Gilvear, D.J. 1997 Quantification of channel bed morphology in gravel-bed rivers using airborne multispectral imagery and aerial photography. *Regulated Rivers research & Management*, 13, 489-499.

Wright, J.D. 1996 Rumblings on the Ocean Floor: GIS Supports Deep-Sea Research, *Geo Info Systems* (now called *Geospatial Solutions*) 6(1): 22-29, 1996.

Wolf, P.A. & Dewitt, B.A. 2000. *Elements of Photogrammetry (with Applications in GIS).* 3rd Edition. McGraw Hill College Division.

Young, T.Y. & Fu, KS (Eds). 1986 *Handbook of Pattern Recognition and Image Processing*. Academic, Orlando, FL.

Zukowskyj, P.M., & Teeuw, R.M. 2002 A comparison of interpolation procedures for deriving river long profiles from digital contour data. *In*: R.S.E.W Leuven, I.Poudevigne and R.M.Teeuw (eds.) *Application of Geographic Information Systems and remote sensing in river studies.* Backhuys Publishers, Leiden.

Web addresses for remote sensing

Please note that websites, unlike published books, are frequently changing, so our apologies if some of these web-links are no longer active. This list covers only a fraction of the *GISci* information that is available via the Internet – hopefully these sites will save you some web-search time. *If you find any particularly good sites, please email the details to the Expedition Advisory Centre so that we can include them in future editions of this handbook.* For a list covering more GISci topics see www.rgs.org/mapping.

General

web.port.ac.uk/departments/sees/staff/whitworth/dataguide/
A useful summary of data sources for fieldwork-based geoscience projects.

mercator.upc.es/tutorial/nicktutor_I-2.html
A tutorial about remote sensing and image interpretation, produced by the Goddard Space Centre.

www.infoterra-global.com/
Infoterra plc (formerly UK National Remote Sensing Centre): major vendors of remote sensing data.

www.npa.co.uk
Nigel Press Associates, suppliers of remote sensing data: lots of useful case studies.

terraweb.wr.usgs.gov
Terrestrial remote sensing.

rst.gsfc.nasa.gov/Front/tofc.html
rst.gsfc.nasa.gov/Sect1/examq.html#1-5
An excellent remote sensing introduction from NASA, with an accompanying test.

www.gsfc.nasa.gov/IAS/handbook/handbook_toc.html
NASA’s online handbook for the latest generation of Landsat earth observation satellites.

www.ccrs/learn/tutorials
The Canadian Centre for Remote Sensing: an excellent site.

Aerial photography

naplib@aerialarchaeology.freeserve.co.uk.
Email address for the UK National Air Photo Library (NAPLIB).

www.crworld.co.uk
Ortho-rectified, digitised airphoto cover of London and the UK.

www.ordnancesurvey.co.uk
UK and overseas (mostly Commonwealth) airphoto archives. To check coverage, email details of study are locations to: customerservices@ordnancesurvey.co.uk

www.ukperspectives.com
Ortho-rectified, digitised airphoto cover of London and the UK.

www.simmonsaerofilms.com/library.html
Aerofilms have an extensive archive of airphotos from around the world, particularly former British colonies.

www.uflm.cam.ac.uk
Online catalogue of extensive UK airphoto archive.

Airborne multispectral and hyperspectral sensors

www.dlr.de/HRSC-A

Details of the multi-spectral High Resolution Stereo Camera.

www.bathspa.ac.uk

Digital aerial photography: UK and overseas coastal case studies.

www.intspec.com/hymap.htm

Integrated Spectronics, manufacturers of hyperspectral scanners: HyMap and the hand-held PIMA device.

www.dlr.de

The German Aerospace Research Establishment (DLR) website has extensive coverage of airborne multi-spectral and hyperspectral sensor systems.

Radar remote sensing

www.ee.ubc.ca

University of British Columbia, extensive web links on radar remote sensing.

www.vexel.com/radar/index.html

Downloadable published papers on radar technologies.

www.infoserv.kp.dlr.de/NE-HF/projects/ESAR/igars96_scheiber.html

The German Aerospace Research Establishment (DLR): details of the airborne radar (E-SAR) now used by the UK Defence Research Agency (DERA).

www.rsi.ca

Radarsat International.

LiDAR remote sensing

www.optech.on.ca

The main manufacturer of LiDAR systems.

www.airbornelasermapping.com

An excellent industry and scientific site on LiDAR.

Remote sensing & archaeology

www.informatics.org/france/france.html

An award-winning website.

www.ghcc.msfc.nasa.gov/archeology/remote_sensing.html

Various world-wide examples from NASA.

naplib@aerialarchaeology.freeserve.co.uk.

Email address for the UK National Air Photo Library (NAPLIB).

GIS applications

www.ex.ac.uk/~yszhang/erosion.htm

Many useful links to websites involved with modelling soil erosion/deposition.

www.engineering.usu.edu/dtarb

Terrain modelling using Digital Elevation Models (DEMs), for hydrology and erosion see studies.www.nmw.ac.uk/ite/banc/deecamp.html - River Dee integrated catchment management plan, NE Scotland.

www.pobonline.com/CDA/ArticleInformation/features/BNPFeaturesItem/0,2338,117907,00.html

Summary of an expedition that used GIS and RS in mapping a glacier in Ecuador.

www.scgis.org/1999_Conference/abstracts/palhares.html

Briefly describes the use of ArcView GIS as a Decision Support Tool for expeditions examining the distribution of genetic resources in Brazil.

The Global Positioning System (GPS)

www.trimble.com/gps/fsections/aa_f1.htm

www.spatial.maine.edu/~leick/gpshome.htm

www.aero.org/education/primers/gps/

www.colorado.Edu/geography/gcraft/notes/gps/gps_f.html

www.control.auc.dk/~tb/gps_view_graph/

everest.hunter.cuny.edu/mp/

Applied remote sensing

www.earth1.esrin.esa.it

European Space Agency: ERS satellite radar images of major floods, as well as earthquake and volcanic hazards.

www.calmit.unl.edu/cohyst/scope.html

Use of Landsat TM imagery in a study of hydrology and agriculture.

www.ais.sai.jrc.it/environmental/lacoast.html

European space Agency examples of coastal remote sensing applications.

www.shef.ac.uk/geography/staff/bryant_rob.html

http://glcf.umiacs.umd.edu/data/

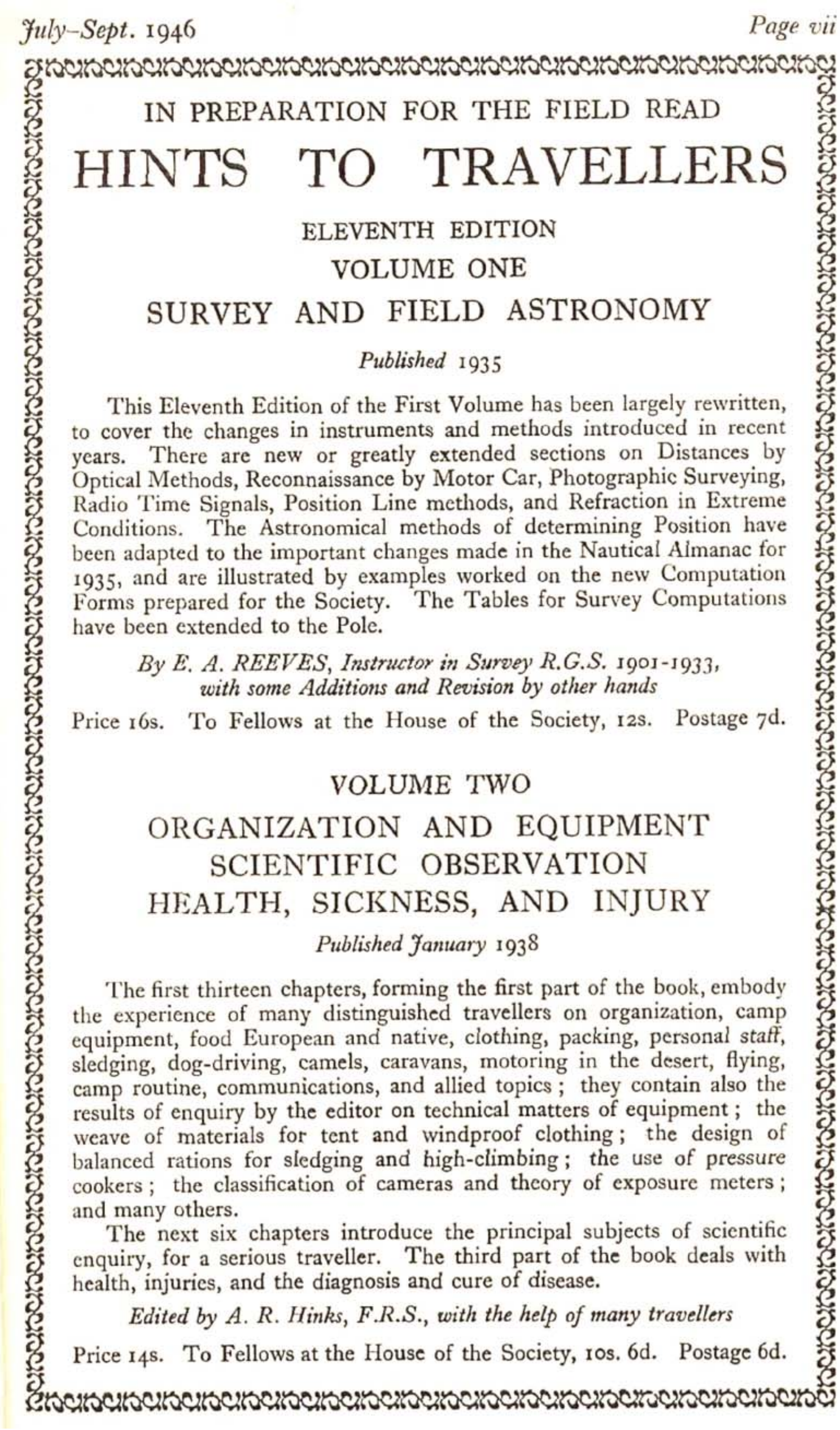
Global Land Cover Facility provides earth science data and products to help everyone to better understand global environmental systems.

GISci Glossary

ASCII	American Standard Code for Information Interchange	Standard for displaying character based data on a computer
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer	Japanese / American multispectral sensor (launched 2000 similar to Landsat ETM+)
CAD	Computer Aided Design	Popular method for scale drawings in a computer similar to GIS but without the necessity for spatial referencing
CEP	Circular Error Probability	A circle containing a location and 50% of the data from a device. Sometimes used to assess GPS accuracy (see Chapter 6)
CHEST		UK Agreement between education establishments and companies to supply students with low cost access to expensive commercial products (does not apply to all software see Chapter 14)
DEM	Digital Elevation Model	Three dimensional representation of the Earth
DGPS	Differential GPS	High accuracy GPS capable of accuracies from 1 to 5 m
ETM	Enhanced Thematic Mapper	A sensor aboard the Landsat 7 satellite, launched in 1999 (see Chapter 5).
FC	False Colour	Type of image from a multispectral sensor where bands have been combined to create an image that does not correspond to the visible spectrum (see Chapter 5)
FTP	File Transfer Protocol	Method for moving data on the Internet
GCP	Ground Control Point	A grid reference from the field referenced to a point on a map or image (commonly used in rectification) (see Chapter 9)
GIS	Geographical Information System	Software for storing, integrating, manipulating and analysing data that has a spatial reference (see Chapter 3)
GISci	Geographical Information Sciences	A term for the combined use or GIS, GPS and Remote Sensing
GIT	Geographical Information Technologies or Techniques	A term used interchangeably with GISci
GLONASS	Global Orbiting Navigational Satellite System	Russian equivalent to US GPS (rarely used in modern GISci)
GNS	Global Navigation System	Navigation system that operates worldwide (such as a GPS)
GPS	Global Positioning System	A system using satellites and small (usually low cost) handheld receivers to determine position to within about 10 m (see Chapter 6)
IR	Infra Red	The part of the electromagnetic spectrum with wavelengths longer than visible light (approx. 1 to 10 micrometres)
Landsat		American earth observation multispectral sensor (three varieties in increasing resolution MSS, TM, ETM+)
Multispectral		Sensor capable of splitting radiation into discrete bands allowing objects to be viewed at different spectral frequencies
MSS	Multi Spectral Scanner	A sensor abroad numbers 1 to 4 of the Landsat satellite series (see Chapter 5).
NDVI or VI	Normalised Difference Vegetation Index	Image processing technique for removing superfluous information from a scene to only leave highlighted vegetation
Panchromatic		Sensor where all radiation is focused into one detector creating a grey-scale image but usually of very high spatial resolution
PDA	Personal Desktop	Handheld computer with capabilities similar to a normal PC but

	Assistant or Personal Digital Assistant	usually with reduced power and functionality
RDBMS or DBMS	Relational Database Management System	Popular method for storing data, where the data has relationships and dependencies to other data in the system (see Chapter 4)
RGB	Red Green Blue	A popular method for creating colour images by defining the amount red, green and blue present. Commonly used in image processing to determine the order of the bands in a false colour image (e.g. 531 RGB) (see Chapter 5 and 8)
RS 232		Common adaptor on rear of older computers and on most GPS units (also called serial adaptor)
SQL	Structured Query Language	The accepted method for querying information in a database
TC	True Colour	Type of image from a multispectral sensor where bands have been combined to create an image that corresponds to the visible spectrum (see Chapter 5)
TIN	Triangular Irregular Network	Three dimensional skeleton used to join points in 3D space to act as the support for a DEM
TIR	Thermal Infra Red	A subset of the infra red portion of the spectrum furthest from visible light (approx. 7-10 micrometres)
TM	Thematic Mapper	A sensor aboard numbers 4 and 5 of the Landsat satellite series (see Chapter 5).
UTM	Universal Transverse Mercator	A set of map projections and co-ordinate systems designed for large scale mapping (see Chapter 2)
VNIR or NIR	Very Near Infra Red	A subset of the infra red portion of the spectrum nearest visible light (approx. 1-2 micrometres)
WAAS	Wide Area Augmentation Service	Methods for improving GPS accuracy (not as accurate as DGPS and requires a modern receiver)
WGS84	World Geodetic Survey 1984	Standard model of the earth used frequently in GISci and the model on which most GPS operate (see Chapter 2)

Field survey work techniques did not change greatly until the 1980s, but with the progressive advances in computer-based technologies there have been tremendous advances over the past 25 years. This advert for a field survey manual produced in the 1940s gives a taste of how much things have changed since then.



July–Sept. 1946 *Page vii*

IN PREPARATION FOR THE FIELD READ

HINTS TO TRAVELLERS

ELEVENTH EDITION

VOLUME ONE

SURVEY AND FIELD ASTRONOMY

Published 1935

This Eleventh Edition of the First Volume has been largely rewritten, to cover the changes in instruments and methods introduced in recent years. There are new or greatly extended sections on Distances by Optical Methods, Reconnaissance by Motor Car, Photographic Surveying, Radio Time Signals, Position Line methods, and Refraction in Extreme Conditions. The Astronomical methods of determining Position have been adapted to the important changes made in the Nautical Almanac for 1935, and are illustrated by examples worked on the new Computation Forms prepared for the Society. The Tables for Survey Computations have been extended to the Pole.

By E. A. REEVES, Instructor in Survey R.G.S. 1901-1933,
with some Additions and Revision by other hands

Price 16s. To Fellows at the House of the Society, 12s. Postage 7d.

VOLUME TWO

ORGANIZATION AND EQUIPMENT
SCIENTIFIC OBSERVATION
HEALTH, SICKNESS, AND INJURY

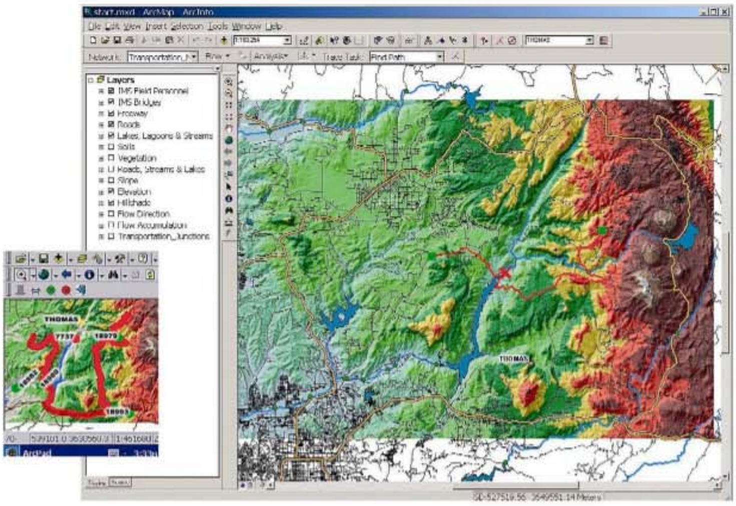
Published January 1938

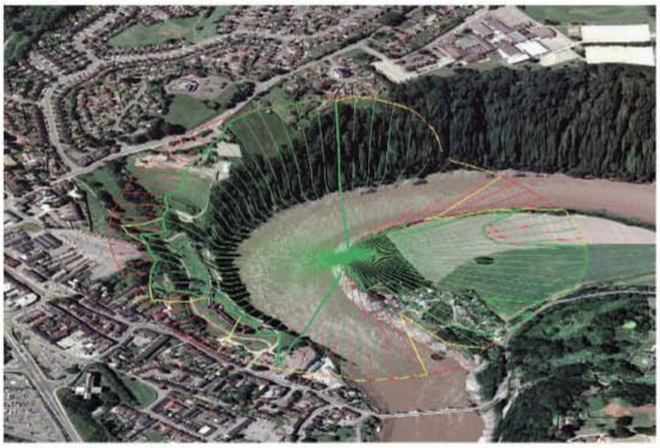
The first thirteen chapters, forming the first part of the book, embody the experience of many distinguished travellers on organization, camp equipment, food European and native, clothing, packing, personal staff, sledging, dog-driving, camels, caravans, motoring in the desert, flying, camp routine, communications, and allied topics; they contain also the results of enquiry by the editor on technical matters of equipment; the weave of materials for tent and windproof clothing; the design of balanced rations for sledging and high-climbing; the use of pressure cookers; the classification of cameras and theory of exposure meters; and many others.

The next six chapters introduce the principal subjects of scientific enquiry, for a serious traveller. The third part of the book deals with health, injuries, and the diagnosis and cure of disease.

Edited by A. R. Hinks, F.R.S., with the help of many travellers

Price 14s. To Fellows at the House of the Society, 10s. 6d. Postage 6d.


Layers
IMS Field Personnel
IMS Bridges
Freeway
Roads
Lakes, Lagoons & Streams
Soils
Vegetation
Roads, Streams & Lakes
Slope
Elevation
Hillshade
Flow Direction
Flow Accumulation
Transportation_Junctions
THOMAS



Expedition Advisory Centre
Royal Geographical Society with IBG
1 Kensington Gore London SW7 2AR

T +44 (0)20 7591 3030
F +44 (0)20 7591 3031
E eac@rgs.org
W www.rgs.org/eac