Digital field data acquisition: towards increased quantification of uncertainty during geological mapping

RICHARD R. JONES [*],
CognIT a.s, Postboks 610, 1754 Halden, Norway
KENNETH J.W. MCCAFFREY, ROBERT W. WILSON, and ROBERT E. HOLDSWORTH
Department of Earth Sciences, University of Durham, Durham, DH1 3LE, U.K.

Corresponding author: r.r.jones@durham.ac.uk

Number of words: 7701 (including all text, references, captions, tables).
Abbreviated title: "Digital mapping: quantifying uncertainty" (max. 40 chars)

Abstract: Traditional methods of geological mapping were developed within the inherent constraints imposed by paper-based publishing. These methods are still dominant in the Earth Sciences, despite recent advances in digital technology in a range of fields that include Geographical Positioning Systems, Geographical Information Systems, 3D computer visualisation, portable computer devices, knowledge engineering and Artificial Intelligence. Digital geological mapping has the potential to overcome some serious limitations of paper-based maps. Although geological maps are usually highly interpretive, traditional maps show little of the raw field data collected or the reasoning used during interpretation. In geological mapping, interpretation typically relies on the prior experience and prior knowledge of the mapper, but this input is rarely published explicitly with the final printed map. Digital mapping techniques open up new possibilities for publishing maps digitally in a GIS format, together with spatially referenced raw field data, field photos, explanation of the interpretation process, and background information relevant to the map area. Having field data in a digital form allows the use of interpolation methods based on Fuzzy Logic to quantify some types of uncertainty associated with subsurface interpretation, and using this uncertainty to evaluate the validity of competing interpretations.

__________________ (end of abstract) __________________

---

[*] Current addresses: e-Science Research Institute, University of Durham, Durham, DH1 3LE, U.K. and Geospatial Research Ltd., Department of Earth Sciences, University of Durham, Durham, DH1 3LE, U.K.

Although the methodology of geological mapping has remained largely unchanged since William Smith's pioneering work two hundred years ago (Smith 1801, 1815), recent developments in digital technology have the potential to revolutionise the way in which geological field data is gathered, stored, processed, displayed and distributed. The digitalisation of field mapping is occurring through advances in GPS ("Geographical Positioning Systems"), GIS ("Geographical Information Systems"), highly portable hand-held PDA devices ("Personal Digital Assistants"), high-powered 3D computer graphics, and satellite communication equipment. In this paper we discuss the fundamental importance of how this new technology can help to remove some of the inherent limitations of traditional methods of geological mapping.

Prior Information

In common usage the concept of "Prior Information" conveys a variety of meanings, largely dependent upon context, ranging from philosophical to mathematical. This paper is not intended as an exhaustive discussion on the semantics of prior information, but rather to focus pragmatically on specific issues that have direct relevance to the process of geological mapping.

Relationship between information and knowledge

In relation to geosciences it is convenient for us to consider "prior *information*" in a broad way that also encompasses related concepts such as "*data*" and "*knowledge*". During the last decade many business sectors, including the hydrocarbon and mining industries have focused heavily on maximising re-use of corporate knowledge and expertise. This highly profit-driven process has given rise to very pragmatic research in areas of applied knowledge engineering. Workers in these fields (e.g. van der Spek & Spijkervet 1997; Liebowitz & Beckman 1998) generally view "information" as just one part of a knowledge hierarchy that ranges from basic data input to sophisticated interpretation by an expert (Table 1).

	Level	Example
5	Expertise (= ability to apply knowledge effectively to perform tasks accurately and efficiently within existing constraints)	the finished geological map! a geological summary of the mapped area scientific paper describing new fossil found in area
4	Knowledge (= application of information to be able to make decisions, solve problems, or perform tasks)	boundary between rock types A and B is in valley rock A is granite mineral P is biotite
3	Information (= data in a specified context, data with meaning)	rock A is observed at location GR 234 789 rock A contains minerals P, Q, R mineral P has colour that is black
2	Data (= inputs that can be represented explicitly in symbolic form)	colour is black lustre is shiny GPS location at time T is 123456 from satellite 1 GPS location at time T is 123678 from satellite 2
1	Inputs (= sensory signals, machine measurements)	geologist receives visual sensory signals GPS unit receives satellite signals

Table 1. The knowledge hierarchy with examples from geological mapping

Interdependence of data, information and knowledge

In traditional views concerning the way in which scientific progress occurs, heavy emphasis is placed on the role of induction to derive scientific theories based on observed phenomena. Because there was generally little consideration that sensory inputs might be misleading, objective “scientific” observations could be treated as facts. This view of scientific methodology was largely dominant until early in the twentieth Century, and still remains influential in much of geoscience education today:

$$\text{observation = fact} \xrightarrow{\text{induction}} \text{scientific theory} \xrightarrow{\text{deduction}} \text{prediction.}$$

The inductivist view of scientific methodology has been refuted in several ways, and individual scientific theories are now more generally viewed as belonging to larger knowledge structures (the “paradigm” of Kuhn 1962, or the “research programme” of Lakatos 1974). Of central importance in the rejection of induction is that observations cannot be made independently of prevalent theory, and that the formation of individual theories builds upon prior scientific knowledge. Thus there is a solid philosophical foundation for stating that prior information (*sensu lato*) is always an influential factor during scientific research, including geological mapping and other disciplines within the Earth Sciences. Put more simply, this discussion merely emphasises something that most geologists will take for granted: the end product of a geological mapping process is a non-unique, subjective interpretation that is markedly influenced by the previous background, prior experience, and expertise of the geologist, and is an interpretation made within the context of current geological understanding.

Maps created using traditional methods of production and distribution often emphasise the knowledge and expertise (i.e. levels 4 & 5 in Table 1) of the mapper, whilst raw data and information (i.e. levels 2 & 3 in Table 1) gathered during the mapping process usually remains unpublished in field maps and notebooks. This situation is particularly prevalent with regional maps typically produced by national surveys. Digital methods of production and distribution allow the possibility of presenting not only the final interpretation but also additional data/information to the user, so that individual observations become reproducible, assumptions or inferences made can be questioned, and overall interpretations more readily tested.

Accessibility of knowledge

Researchers who have studied the transfer of knowledge between workers in modern industry have recognised that there is great variation in the accessibility of an individual's knowledge and expertise (Nonaka 1994; Nonaka & Takeuchi 1995). Whilst some knowledge often exists in an *explicit* format that is readily available and that can be communicated and understood by other colleagues, there is usually also a large amount of *implicit* knowledge that is not (yet) in a format that is easily accessible to others (Table 2). Implicit knowledge exists within the head of a person, but can readily be made available to others (i.e. made explicit) through query and discussion, or through a conscious decision to document what one knows about a subject. In contrast, a large amount knowledge possessed by humans is usually *tacit* (Polanyi 1958, 1966). Tacit knowledge exists within the head of a person, but is generally in a form that is not easily accessible to other people, either because its owner is not aware that they possess the knowledge, or will not be able to express it in a useful, understandable format. An example is the knowledge of how to ride a bicycle: although you may know how to cycle without having to think, it is very difficult to give an explicit description of how to perform the task to someone who has never tried themselves. Similarly, the way in which a skilled map reader navigates by converting the symbols shown on a map into a mental 3D picture of the terrain is a process that is done automatically at a subconscious level, and is difficult to describe explicitly in a meaningful way. A further example is the way in which a geologist interprets structural symbols on a map to form an image of the 3D architecture that the map depicts. Traditionally the transfer of tacit knowledge of this type is achieved by repeated learning through "trial and error", often in a teacher/pupil or master/apprentice situation. More recently, specialised knowledge elicitation techniques have been devised that extract tacit knowledge and represent it in more explicit forms that can be manipulated by a computer (e.g. Boose 1986; Kidd 1987; Ford & Sterman 1998; White & Sleeman 1999).

Workers who are considered to be "experts" within a particular domain typically possess high levels of tacit knowledge. This is often particularly acute within geological mapping, where the quality of output will be greatly dependent upon the skill and expertise of the mapper, although their proficiency is rarely immediately obvious simply by looking at the final result.

Knowledge accessibility	General examples	Examples from mapping
Explicit	scientific papers published in journals. textbooks. instruction manuals.	detailed outcrop map ("green-line" mapping).
Implicit	unpublished observations. working hypotheses. undocumented troubleshooting fixes.	observations not recorded explicitly on final map. general geological theory that has influenced specific interpretation.
Tacit	instant recognition of minerals based on "look and feel" rather than explicit physical tests. interpretation of seismic sections.	expertise of the mapper. the mapper's preconceived bias, insight, gut-feeling and intuition.

Table 2. Different levels of accessibility of knowledge

Conditional probability

Ideally, when faced with uncertainty associated with interpretation of individual field observations, the geologist should attempt to make further field observations in order to resolve any outstanding issues. In practice, this process will almost always be restricted by limitations imposed through lack of exposure and limited resources (manpower, time, money), and the geologist is almost always obliged to supplement direct observation with a mixture of insight, guile, guesswork, and "gut-feeling". These aspects reflect the geologist's personal bias, and represent a prior belief which if it can be expressed explicitly, can be used to predict the likelihood that a particular interpretation is true. This forms the basis of Bayesian statistics, in which the probability that hypothesis H is true, given the occurrence of event E (in context I) is given by:

$$p(H \mid EI) = (p(H \mid I) * p(E \mid HI)) / p(E \mid I).$$

Thus Bayesian probability can provide a mathematical framework for expressing the uncertainty associated with geological interpretation, and offers a possible way to extend and improve traditional methods of geological mapping.

Traditional Geological Mapping

The ability to produce accurate field maps and to record associated observational data in a notebook lies at the core of Earth Science activities (e.g. Barnes 1981; McClay 1987), and forms the basis by which geological maps are constructed. Whilst a geological map is a two-dimensional (2D) representation of the distribution of rock formations in an area, it also conveys through symbols and graphics the three-dimensional (3D) geometry and form of the rocks and structures in the area. The gathering of field data occurs in a broad range of natural environments and is typically carried out by individuals or small teams of geoscientists working on foot or using various forms of transport: this requires that any equipment or techniques used to gather information are highly mobile and easy to maintain. For these reasons, it perhaps not surprising that field mapping has since its inception remained as a paper-based activity using maps, field notebooks and compass-clinometers.

The scientific aims of a study and the time available for fieldwork – which is often dictated by funding – will determine the type of geological mapping to be carried out. Barnes (1981) identifies four main types of geological mapping activities:

a) *Reconnaissance mapping* typically covers a large area and is carried out to find out as much as possible about a poorly-known region over a short period of field study. Significant amounts of work may be done using remote sensing techniques or interpretation of aerial photographs.

b) *Regional mapping* typically results in geological maps at 1:50,000 scale recorded on an accurate topographic base map. Such mapping is generally the result of systematic programmes of field-based data gathering, fully supported by photogeological interpretation and integration of other sub-surface geological or geophysical datasets.

c) *Detailed mapping* generally refers to maps made at 1:10,000 or larger scales and in many cases are produced to document key geological relationships in detail. Many require the field geologist to produce their own base map simultaneously using planetable-, chain- or cairn-mapping techniques (see Barnes 1981 and references therein).

d) *Specialized mapping* where maps are constructed for specific purposes and do not necessarily include all aspects of the observed geology. These include mine plans and maps showing geotechnical, geophysical or geochemical data.

The following information is particularly critical to all field-based data-gathering and observational activities: accurate location, geological context, and the spatial/temporal relationship to other data or observations gathered at that or other location(s). In addition, all field maps and notebooks must be legible, be readable by another geologist and must clearly distinguish between observed facts and inferences drawn from those facts (Ramsay & Huber 1987). Generally speaking, observations and data are gathered at a series of localities, the locations of which are marked by hand onto a topographic or aerial photographic base map, with all data measurements and observations being recorded simultaneously in a field notebook. Ideally, the extent of visible outcrop ("green-line mapping") and location of mappable geological boundaries (including some indication of how well constrained these boundaries are in terms of the available exposure, using solid, dashed or pecked lines), will be added to the base map by the field geologist as they move between localities. Most geologists are encouraged to interpret their observations and measurements as they map and to modify these interpretations as more information is acquired. Thus field mapping is an iterative exercise in which both data-gathering and interpretation occur simultaneously. This also means, however, that field-based data gathering presents a number of very significant challenges viewed from a perspective of prior information. In particular, we highlight four main problems:

- Field mapping involves extensive use of tacit knowledge, in which the *a priori* assumptions made both when making interpretations or even when gathering data are either not stated or may not even be considered – thus 'facts' and 'inferences' are often not clearly separated in many, if not all studies.
- The workflow from field-mapping to published map is generally a complex process involving data collection, interpretation, data reduction and final map drafting. The map is an abstraction at one specific scale of a large amount of data collected at the outcrop scale. Therefore the vast majority of 'inputs', 'data' and 'information' are typically excluded from the final result as they generally cannot all be incorporated into the final paper map. Most maps are therefore dominated by interpretations ('knowledge' + 'expertise'). In many cases, some of the interpretation is made in locations far removed from the field either before or after the actual data-gathering was carried out.
- All paper maps are inherently limited in terms of what they show by their scale. In many cases this means that they lack precise spatial accuracy, meaning that *reproducible* observations or measurements are often difficult or impossible.

- The final map generally shows little expression of the uncertainties involved in its production, and where uncertainty is depicted it is primitive, ad-hoc, and qualitative. Thus traditional geological mapping remains a highly interpretative, subjective art form in which uncertainty is difficult to quantify in any statistically meaningful way.

Digital Geological Mapping

GIS has evolved from its early use as a computer mapping system and is now defined as ‘an information management system for organizing, visualizing and analysing spatially-orientated data’ (Coburn & Yarus 2000 p.1). Since GIS became commercially available in the 1980s, GIS products are now used in a large number of applications that deal with spatial data, including social and economic planning, marketing, facilities management, environmental and resource assessment (Rhind 1992; Longley *et al.* 2001). Bonham-Carter (2000) describes the core GIS activities in a geoscience project as being: (1) data organisation, (2) data visualization, (3) data search, (4) data combination, (5) data analysis, and (6) data prediction and decision support. The combination of these capabilities and the ability to handle large databases (up to a terabyte) indicate the power of the GIS approach for handling spatial data and its attraction for geoscience users such as the petroleum and mining industries. A generalised work flow for digital geological mapping is shown in Fig. 1.

In its original guise, GIS largely dealt with 2D data that was mapped onto the earth’s surface (Rhind 1992). It was recognised that in order to deal with volumetric spatial information or 3D geometries from sub-surface data, a 3D GIS or a “GSIS” (Geo-Scientific Information System) was required, and such systems have now been developed for commercial purposes (Turner 1992, 2000). gOcad™ is one example of a powerful software system that is capable of displaying and analysing complex 3D geological sub-surface architectures (Mallet 1992).

Digital Geological Mapping (DGM) is a methodology by which a geologist collects GPS-located field data in a digital format. The method has been adapted from digital mapping and surveying

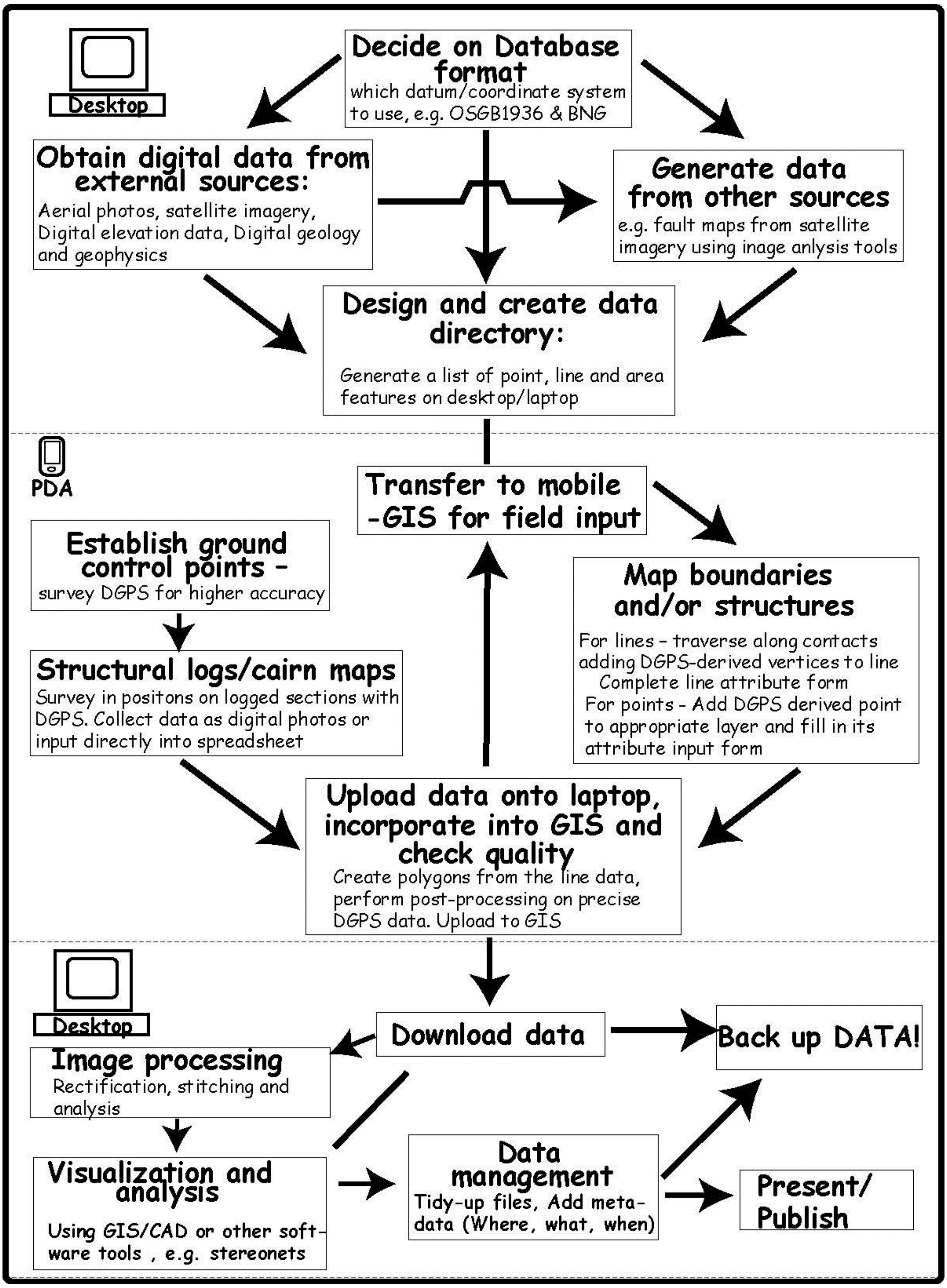


Fig. 1. Generalised workflow for Digital Geological Mapping.

techniques which are now widely used in construction, engineering and environmental industries. Early pioneers that have customized these techniques for use in Earth Science fieldwork include Struik *et al.* 1991, Schetselaar 1995, Brodaric 1997, Briner *et al.* 1999, Pundt

& Brinkkotter-Runde 2000, Bryant *et al.* 2000, Maerten *et al.* 2001 and Xu *et al.* 2001. Digital acquisition is gradually becoming more commonplace, particularly in North America, although adoption by European national surveys has been slow (Jackson & Asch 2002), and teaching of digital data capture is still not widespread in Europe.

The DGM we describe here is a mapping system that would be suitable for most geological purposes. The system involves the integration of three key technological components: (1) a GPS receiver usually capable of obtaining differential correction data that enable sub-metre positional accuracy; (2) a PDA or other digital data-logger, and (3) mobile GIS software. Mobile GIS is a specialised version of PDA software that can exchange information with more general purpose desktop GIS. When used in 3D mode we suggest that DGM provides an onshore equivalent to a wider definition of digital mapping as used by petroleum industry to make 3D structural interpretations in sub-surface data.

In a GIS, information is usually displayed as a series of layers that can be superimposed with each comprising a single type of data. Typically this may comprise features or objects that have distinct shape and size or field data that vary continuously over a surface (Longley *et al.* 2001) as summarised in *Table 3*. The advantage of a GIS-based mapping system is that any number of different types of data may be georeferenced and included as a separate layer in the database. These can then be displayed and analysed in conjunction with newly acquired field data (Fig. 2). Examples of data that may be included are field photographs, regional geophysical maps, aerial photography, satellite imagery, topographic data, previously digitised geological information, sample catalogues, geochronological data, geochemical data etc. In this way the GIS graphical user-interface represents a single entry point to a wide range of spatially related relevant data that can be easily accessed in seconds. By comparison, such disparate types of data would traditionally be spread widely between field notebook, paper maps, isolated files on a computer, boxes of photographic slides or prints, library journals, and loose papers.

GIS data	Type	Geological data	Specific example
Point	Object	Locations of structural stations, anywhere a measurement is made or a sample is taken	Bedding strike & dip Gravity measurement Geochemical sample
Line	Object	Boundaries between areas or linear objects	Contact between rock units Fault trace
Polygon	Object	Areas of rock units	Formation extent Area of igneous intrusion
Raster	Field	data sampled on a matrix of equally sized squares	Elevation (Digital Elevation Model) LandSat™ image

Table 3. Data types in a typical GIS system

Other advantages of DGM over traditional mapping include: (a) improved time efficiency, especially regarding data management, analysis and output; (b) because digitally captured data have high spatial precision, a significant reduction in uncertainty regarding location errors can be achieved; (c) when elevation data is acquired DGM is inherently 2.5D or 3D (see below), so that 3D geometries are easier to visualize (McCaffrey *et al.* 2003). At present the disadvantages of DGM include the relatively high cost of robust equipment, the potential for data loss in the event of equipment failure, and the reluctance of more conservative geologists to explore the increased possibilities offered by new technology.

2D, 3D and 2.5D data

In DGM, data are collected in either 2D or 3D modes. If the objective is purely to produce a map of the region then point, line and polygon data may be stored using only x- and y- coordinates (i.e. longitude and latitude). Real-time Differential GPS systems regularly give precision to approximately 1 metre in the horizontal plane and survey systems that post-process positional data can attain cm-scale accuracy. The positional precision and accuracy that may be achieved using GPS is dependant on variations in the input satellite configuration (an error summarised by

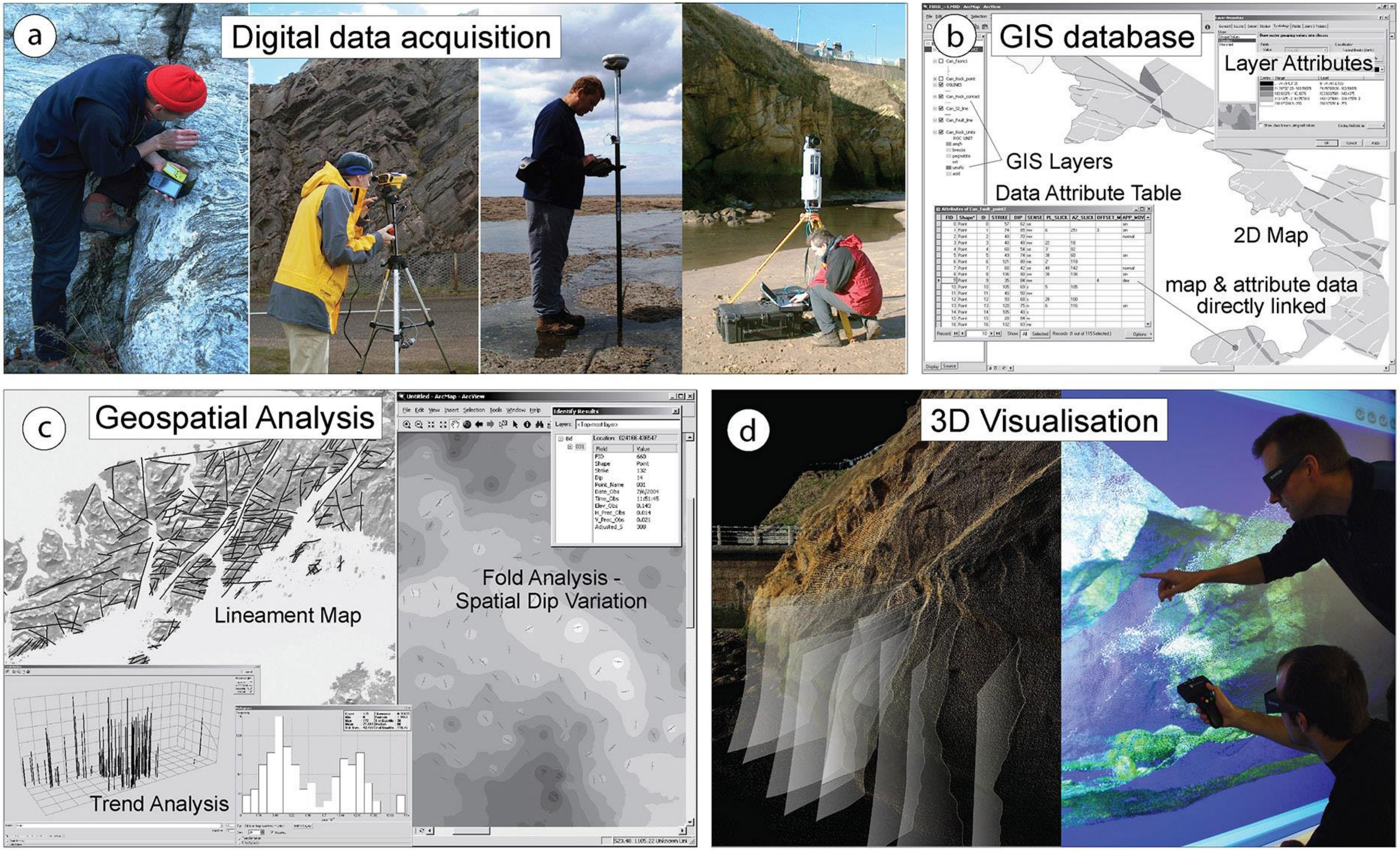


Fig. 2. Illustrative examples of aspects of digital geological mapping: (a) four methods of digital data capture (left to right); hand-held GPS connected to PDA (Østfold, Norway); laser ranging device (Kirtomy, Scotland); high precision Real Time Kinematic GPS (Yorkshire, England); ground-based Lidar laser scanner (Tyneside, England); (b) example screen-shot from ArcGIS showing the close association between cartographic representation, GIS data-layers, and data atttibute tables stored in the underlying database (Achmelvich Bay, NW Scotland); (c) examples of geospatial analysis; (left) 2-D trend analysis (Lofoten, Norway); (right) 3-D analysis of small-scale open folding using kriging (Tyneside, England); (d) 3D visualisation; (left) interpolated fault planes in a 3-D point cloud acquired by laser scanning (Tyneside, England); (right) collaborative data analysis point cloud data using immersive 3-D visualization.

the Dilution of Precision statistic calculated continuously by GPS receivers). Topography or buildings can limit the number of input satellites available to a GPS receiver and thus accurate positioning near a cliff or a in a deep valley may be difficult to achieve.

Geological information gathered by traditional geological mapping has been displayed on 2D representations such as geological maps, however this format has disadvantages as described above. Whilst GIS software products are often used to produce traditional geological maps the systems also allow more flexible methods of visualisation that can be easily tailored to individual requirements. For example, on screen data can be viewed at different scales using the zoom and pan functions with different combinations of data layers visible as required.

GIS data may be overlain or ‘draped’ onto a digital elevation model, in the form of a surface fitted to a raster map of elevation values, to produce a display that has been referred to as a 2.5D representation (Longley *et al.* 2001). These data may then be displayed using a 3D viewer that allows rotation to different vantage points as well as zoom and pan. One particularly useful geological application of 2.5D displays is to study how geological formations and structure are related to topography (Fig. 3).

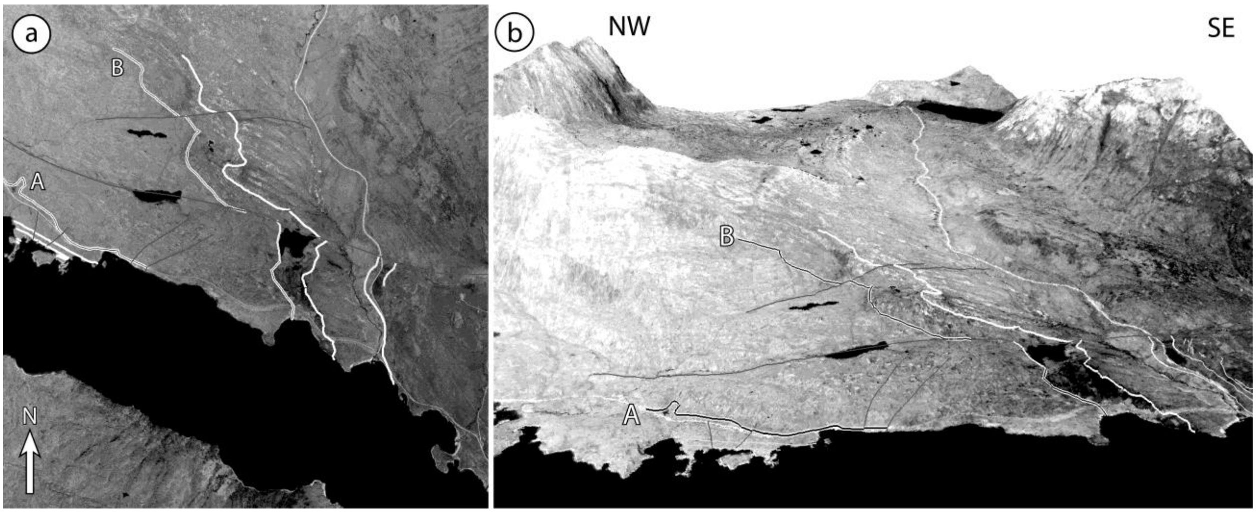


Fig. 3. Digital Geological Mapping of part of the foreland to the Moine Thrust at Loch Assynt, N. Scotland; (a) map view using GIS, showing geological boundaries (upper GIS layers) and aerial photograph (lowest GIS layer); (b) oblique perspective (2.5D view) of the same data, from the south, showing that boundary A is sub-horizontal and boundary B dips to the east. Display has 1.5 times vertical exaggeration.

For accurate 3D reconstructions of geological architectures, the z-coordinate (i.e. elevation) for all positions is essential. Most mobile-GIS applications allow this to be incorporated in the data table. Despite GPS having poorer resolution in the z direction, in good conditions, a differential GPS can give a vertical precision of approximately 1m. Alternatively, 2D data may be converted to 3D by locating the positions onto a digital elevation model, however the resolution is limited by the horizontal spacing of the grid nodes (typically 20-50m) and the precision of the values at each node (typically ± 3-10m).

DGM and Prior Information

DGM is capable of incorporating prior information for the following reasons:

- The powerful data handling capability and scaling functionality of GIS means that an *a priori* framework, newly collected data, and interpretation can all be maintained in a single digital model;
- Several types of data can be stored together, all tied to their geospatial position within the GIS model: attribute data, metadata, photos, sketches, ideas, notes, video, speech etc;
- Having data in digital format is the starting point for quantification of uncertainty (as discussed below).

In order to improve the inclusion of prior information in DGM workflows, more work needs to be done to integrate the various steps involved in the process of data acquisition, interpretation, and final model. User-friendly data gathering methods need to be developed to make it possible for geologists to capture information in ways that are more intuitive (e.g. using more rapid ways of data entry for PDAs such as speech recognition), rather than ways dictated by the non-flexibility of existing hardware and software.

Quantification of Uncertainty in Geological Mapping

Irrespective of whether a mapper uses traditional or digital methods, the acquisition and interpretation of field data inevitably involves a wide range of different types of uncertainty

(*Table 4*). Uncertainties accrue with the onset of data acquisition, and accumulate and propagate throughout the overall interpretation process. Some sources of uncertainty can readily be expressed in terms of a quantitative assessment of the precision of measurement for a piece of equipment (e.g. error tolerance of GPS or clinometer measurements). Other uncertainties can be reduced and quantified by repeating observations so that a measure of variance can be ascertained (e.g. variation of dip and strike at a single exposure). However, much uncertainty typically associated with geological interpretation is less easy to quantify. This is uncertainty that is not associated with accuracy of individual measurements, but rather with the non-uniqueness of multiple solutions, each of which seems to be a viable interpretation to the problem, based on available data.

Level	Type of uncertainty	Examples
Data acquisition	Positional	"how sure am I of my current location?" "how reliable is my base map?" "what is the precision of my GPS measurements?" "is the borehole straight, or has it deviated without me knowing?"
	Measurement	"what is the precision of my clinometer?" "what is the accuracy of my dip/strike readings?"
	Scale-dependant variability	"how much does the dip and strike vary over the scale of the outcrop?" "is my reading representative of the surrounding area?"
	Observational	"is this rock best described as a granite?" "is this fossil the brachiopod *pentamerus*?" "is that a stretching lineation or an intersection lineation?"
	Temporal	"how reliable is this way-up criteria?" "is the relative age of these structures identified correctly?"
	Sampling bias	"is my data biased by the natural predominance of sub-horizontal exposures?" "has my sampling been skewed by me focusing only on the zones of high strain?"
Primary interpretation	Correlation	"is this limestone the same unit as the limestone at the last outcrop?" "is it valid to correlate the S2 fabric here with the S2 fabric observed on the other side of the area?"
	Interpolation	"how likely is it that all the ground between these two outcrops of slate also consists of slate?" "how much control do I have over the geometry of this fold trace?"
	Inference from topography	"is there really a fault running along the unexposed valley floor?" "does this sharp change in slope correspond to a lithological boundary?"
Compound interpretation	Finished 2D map	"how can I quantify the uncertainty associated with this sophisticated interpretive model that I have slowly built up through a long iterative process of data collection and individual primary interpretations?!"
	Geological cross-section	
	3D structural model	

Table 4. Examples of different types of uncertainty arising during geological mapping

Interpolation

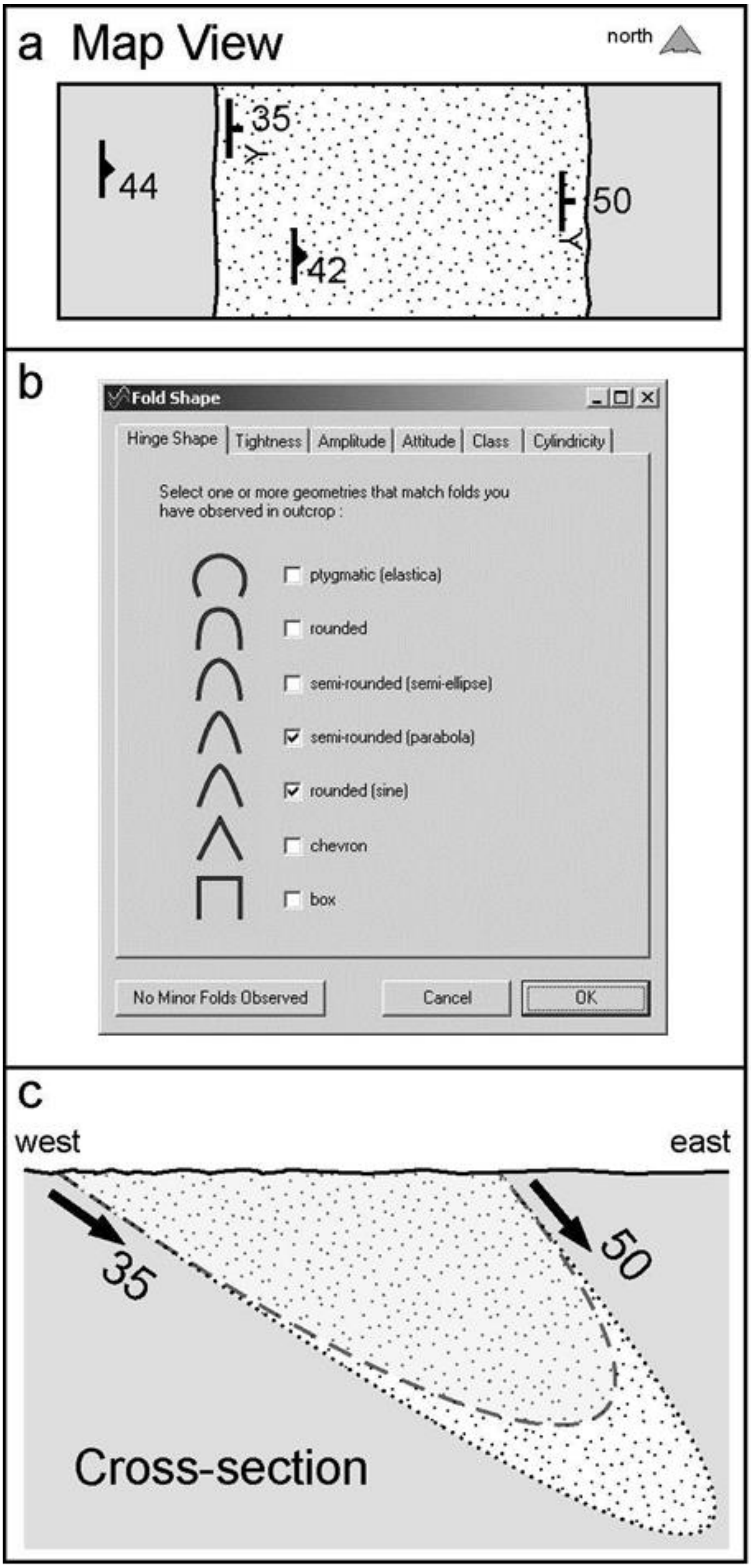
Geological mapping tends to produce sparse datasets. This is usually because the amount of exposure is limited, but even when there are high levels of exposure it is generally impractical to study all exposed rock in detail. Therefore one of the most important aspects of creating a geological map involves the interpolation of data to fill the areas between the intermittent data points actually measured. Interpolation below (and above) the surface of the earth is of course also central to producing 2D cross-sections and 3D models. Most GIS systems have in-built analysis tools for the interpolation of geospatial point data across a topographical surface. These include deterministic methods (based on curve fitting of mathematical functions), and more advanced geostatistical methods ("kriging") that combine statistical analysis with mathematical curve fitting and which also provide a statistical measure of uncertainty across the whole surface. The most basic approaches to kriging incorporate simplistic probability distributions, and the values of uncertainty derived are simply based on the sparseness of data. Kriging has a tendency to smooth out sharp variations between adjacent data points, so care is needed when showing spatial variation in parameters that might change abruptly across discontinuities such as faults.

Geostatistical methods of interpolation rely upon the basic assumption of spatial autocorrelation: i.e. that points that are spatially near to each another tend to be more similar than those farther away. This assumption is often acceptable for many types of geoscientific data (e.g. elevation values, mineral concentrations, pollution levels etc.), which can generally be mapped and interpolated using standard GIS functionality.

Whilst geostatistical kriging methods usually work well for interpolation of geospatial point data distributed on the earth's surface, they are generally less well suited to interpretation of subsurface structure unless a reasonable amount of data is available at depth (e.g. dense borehole data, seismic grid, data from mine workings, or high topographic relief). Although geological surfaces (bedding, foliation, fractures) do tend to have high spatial autocorrelation, for subsurface interpretation based on outcrop data alone the availability of data is typically much too sparse in the z-direction to allow meaningful interpolation at depth. Structural measurements at outcrop include vector data (strike, dip, plunge, azimuth) that describe the 3D orientation of geological surfaces, as well as crucial supporting information concerning structural polarity (younging, facing, vergence) and temporal relationships (relative age of beds, structures, cross-cutting relationships, multiple generations of structures). Therefore, structural measurements typically encapsulate important additional information that should be used as a prior input to the interpolation process (Fig. 4). Simply using a pre-defined mathematical curve (c.f. deterministic interpolation methods) or statistical probability distribution (c.f. kriging) disregards the extra structural information gathered by the geologist, and is therefore much less likely to produce a realistic interpolation.

Qualitative uncertainties

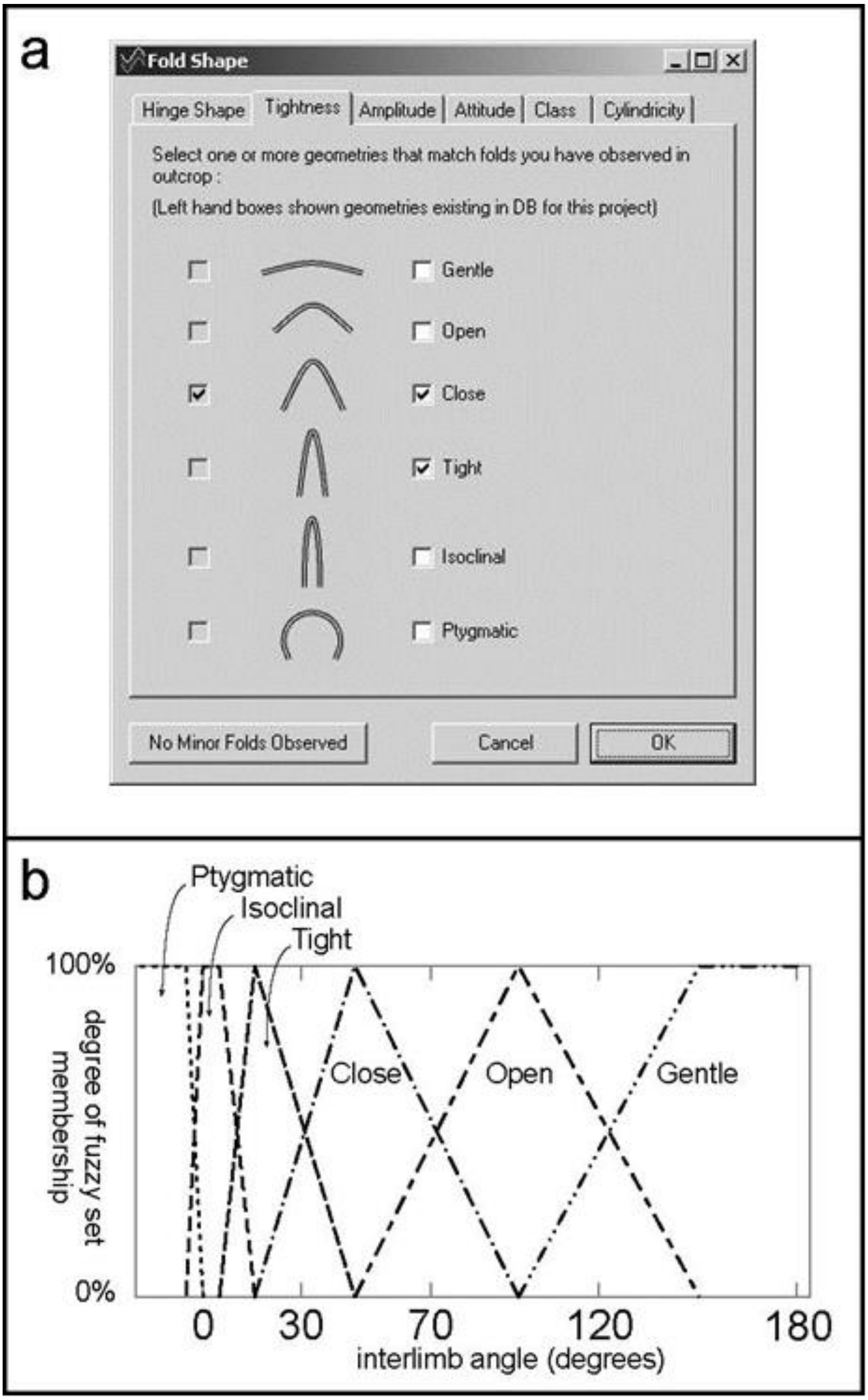
Many types of uncertainty are difficult to express quantitatively, and are more suited to a *qualitative* evaluation by the geologist. Although this may be abhorrent to inductivists that believe that science consists only of quantitative, objective measurement, a subjective statement such as “this rock looks sheared and I am reasonably confident that it is a mylonite” is a more useful and representative observation statement than having to make a binary choice between “this *is* a mylonite” and “this is *not* a mylonite”. An obvious strategy to tackle this situation would be for field geologists to specifically record an estimate of confidence with every observation, as a matter of routine. However, most geologists will perceive this data as superfluous, and gathering it as an additional, unnecessary burden, because there is a general lack of methodology developed within geological mapping which allows such information to be used in a systematic way.



Select one or more geometries that match folds you have observed in outcrop :

☐ ptygmatic (elastica)
☐ rounded
☐ semi-rounded (semi-ellipse)
☑ semi-rounded (parabola)
☑ rounded (sine)
☐ chevron
☐ box

Fig. 4. Use of structural field observations as prior input for subsurface interpolation; (a) example sketch map showing dip and strike of bedding and cleavage with younging data; (b) data entry dialog box for field observations of fold hinge geometry, which are used to influence subsurface interpolation; (c) cross-section showing example of subsurface interpolation. A range of large-scale fold geometries are consistent with evidence from minor folds observed at the surface. The dotted fold profile traces a fold surface that is 33% longer than the dashed profile, and forms a fold in the stippled rocks with 70% additional area.



Select one or more geometries that match folds you have observed in outcrop :

(Left hand boxes shown geometries existing in DB for this project)

☐ ☐ Gentle
☐ ☐ Open
☑ ☑ Close
☐ ☑ Tight
☐ ☐ Isoclinal
☐ ☐ Ptygmatic

Fig. 5. Further use of structural field observations as prior input for subsurface interpolation; (a) data entry dialog box in which the user can specify the typical inter-limb angles of minor folds seen in outcrop; (b) fuzzy set definition showing mapping of inter-limb angle to fuzzy descriptors (‘open’, ‘tight’ etc.). The range of inter-limb values indicated by the geologist is used to constrain the range of realistic sub-surface interpolations, and also to test for internal inconsistencies between field data and other user inputs.

The potential now exists for this situation to change, following recent advances in mathematics and computer technology, especially within several branches of artificial intelligence (AI). In particular, developments in Fuzzy Logic provide a formal framework in which relative terms (such as “*quite* sheared”, “*very* fractured”) can be transformed to discrete numerical values and represented within binary computer code. Figure 5 is an example of the use of fuzzy descriptors to represent the tightness of inter-limb angles in folds. The value(s) of fold tightness chosen by the user as representative of minor folds observed in outcrop are combined with other user input for fold data and used to constrain subsurface interpolation. Checks are also made to identify potential areas where the data might not be internally consistent using a fuzzy rule base with rules such as: “if hinge_shape is ‘chevron’ or ‘box’ and interlimb_angle is ‘close’ then conflict_risk is moderate”; “if hinge_shape is ‘chevron’ or ‘box’ and interlimb_angle is less than or equal to ‘tight’ then conflict_risk is significant”.

Other branches of AI (e.g. Bayesian networks, neural networks, genetic algorithms, constraint satisfaction techniques) also have potential for finding solutions for complicated non-linear models involving very many variables. All these methodologies have a proven track record of finding good solutions to real problems with a much shorter amount of computer processing than traditional approaches (e.g. Buttenfield & Mark 1991; Braunschweig & Day 1995; Hamburger 1995; Murnion 1996; Jones *et al.* 1998; Nordlund 1999; Ouenes 2000; Peddle 2002; Luo & Dimitrakopoulos 2003).

Future Trends in Geological Mapping

For the last two decades the growth in Information Technology (IT) has generally been so great that geologists have struggled to keep abreast of technological advances. Within geological mapping Earth Scientists have been slow to improve workflows and methods of interpretation that exploit the newly developing technologies. There is no indication that the current rate of growth within IT is set to diminish, and the following trends are likely to provide improved opportunities for geoscientists to improve the process of geological mapping:

- portable equipment will continue to become lighter, cheaper, more robust, more powerful, more intuitive and user-friendly, and more integrated with the user
- an increased number of 2D and 3D analytical tools will be incorporated into existing GIS software to provide an single integrated tool for geological mapping and interpretation.
- interpretation tools will propagate information about uncertainties through the modelling process so that various interpretations can be tested in parallel and an indication of overall uncertainty can be given for each interpretation.
- satellite communications technology combined with GRID facilities will bring super-computing power to field geologists (a “PersonalGRID”). This will increase the possibilities for ongoing iterative interpretation of field data whilst still in the field.
- speech recognition software combined with semantic based search technology can help to encourage the geologist to verbalise (= make explicit) more of the decision-making processes involved in mapping.

CONCLUSIONS

Although traditional processes of geological mapping have a proven track record established over a period of two hundred years, there are nevertheless important methodological shortcomings seen from a scientific perspective:

- paper-based published maps generally show only a fraction of the field data that have been collected and used as the basis for the map’s creation. Other data remain hidden in the field notebook and in the head of the mapper.

- published maps rarely make reference to the reasoning used during interpretation of the basic field data. Reasoning typically relies heavily on prior information and knowledge that represents the experience of the geologist before the onset of the mapping project.
- paper-based maps are by necessity published at a fixed scale. The skill of the cartographer is to present as much relevant information as possible whilst maintaining legibility at the chosen scale, but inevitably there is a loss of precision especially with respect to geospatial positioning.
- with traditional maps there are generally only very limited possibilities for expressing any uncertainty concerning the given interpretation, and these are not quantitative.

Whilst the above limitations have always been acceptable as long as there were no viable alternatives to paper-based publishing, today's information technology makes it possible to store all the necessary data for a mapping project on a compact disc that costs just a few pence. Digital mapping techniques have the potential to improve the scientific validity of the mapping process in the following ways:

- collected field data can be stored and distributed together with the interpreted map in a single digital model within a GIS. In the future it should be possible to capture and store an even wider range of multimedia datatypes in a seamless way (including metadata, photos, sketches, ideas and notes, video, speech etc.) all tied to the appropriate geospatial position within the GIS model;
- prior inputs used as the basis for interpretations can be stated explicitly within the same GIS mode. Although this alone does not involve the quantification of uncertainty associated with interpretation, it does represent a radical improvement to geological mapping practice as it shows the user of the map not only the mapper's interpretation but also the raw data upon which the interpretation is based;

- traditional paper-based maps can only display a finite amount of information for a given scale of map without loss of legibility. The inherent scalability of a GIS model makes it possible to store a huge amount of data for any geospatial locality, so that the amount of data available to the map user through the user-interface is not restricted in the same way;
- as progressively more analytical tools are incorporated into GIS software, geospatially referenced data can be analysed and interpreted within a single software environment;
- many types of uncertainty that arise during the mapping process can either be quantified or can be estimated qualitatively in a way that can be represented digitally (using AI techniques).

Digital geological mapping is still in its infancy. Future work should concentrate on the following challenges:

- the digital workflow should be continuously improved to make it more intuitive and quicker to capture field data in digital format;
- further integration of analysis tools and GIS is needed;
- methodologies should be further developed that use the uncertainties associated with individual data or interpretations as the input to produce an overall estimate of uncertainty associated with a given model. Alternative interpretations can be modelled simultaneously, with cumulative uncertainty calculated for each model;
- efforts should be made to combine portable GIS with GRID technology, to provide the field geologist with powerful analysis tools whilst in the field. This can increase the ability of the geologist to analyse their data whilst still on the outcrop, thereby helping to optimise strategy for further data collection.

Acknowledgements — Thanks to Nick Drinkwater (Schlumberger) and Dave Jeanette (Bureau of Mineral Resources, Texas) for constructive reviews. Johnny Imber, Nicola De Paola, Phil Clegg and other members of Durham RRG and Bernt Bremdal at CognIT contributed with useful debate. Immeasurable inspiration has been gathered from the pioneering efforts of Professor Oleg McNoleg (McNoleg 1996, 1998). Steve Freeman and others at Leeds RDR made early progress with DGM. Much of this work was part of the Ocean Margins Link project (NER/T/S/2000/01018), funded by Statoil UK Ltd, BP and NERC.

REFERENCES

Barnes, J.W. 1981. *Basic Geological Mapping*. Geological Society of London Handbook Series, Open University Press, 112 pp.

Boose, J H. 1986. *Expertise Transfer for Expert System Design*. Elsevier, Amsterdam, 312 pp.

Bonham-Carter, G. 2000. An overview of GIS in the geosciences. In: Coburn, T.C. & Yarus, J.M. (eds) *Geographic information systems in petroleum exploration and development.* AAPG Computer Applications in Geology, **4**, 17-26.

Braunschweig, B. & Day, R. 1995. An overview of AI Techniques and of their use in the Petroleum Industry. In: Braunschweig, B. & Day, R. (eds) *Artificial Intelligence in the Petroleum Industry: Symbolic and Computational Applications*. Imprimerie La source d'Or, France, 3-56.

Briner, A.P., Kronenberg, H., Mazurek, M., Horn, H., Engi, M. & Peters, T. 1999. FieldBook and GeoDatabase - tools for field data acquisition and analysis. Computers & Geosciences, **25**, 1101-1111.

Brodaric, B. 1997. Field data capture and manipulation using GSC FIELDLOG v3.0. In: Soller, D.R. (ed.) *Proceedings of a workshop on Digital Mapping Techniques: Methods for Geologic Map Capture, Management and Publication*. US Geological Survey, Open File Report, 77-81.

Bryant I., Carr D., Cirilli P., Drinkwater N., McCormick D., Tilke P. & Thurmond J. 2000. Use of 3D digital analogues as templates in reservoir modelling. Petroleum Geoscience, **6**, 195-201.

Buttenfield, B. P., and Mark, D. M., 1991, Expert systems in cartographic design. In Taylor, D. R. F. (ed) *Geographic Information Systems: The Microcomputer and Modern Cartography*. Pergamon Press, Oxford, 129-150.

Coburn, T.C. & Yarus, J.M. 2000. *Geographic information systems in petroleum exploration and development*. AAPG Computer Applications in Geology, **4**, 315 pp.

Ford, D. & Sterman, J. 1998. Expert Knowledge Elicitation for Improving Mental and Formal Models. System Dynamics Review, **14**, 309-340.

Hamburger, J. 1995. GeoCos 3D: A diagnosis system, based on the FROG formalism, to help geologists set up valid interpretations. In: Braunschweig, B. & Day, R. (eds) *Artificial Intelligence in the Petroleum Industry: Symbolic and Computational Applications*. Imprimerie La source d'Or, France, 57-85.

Jackson, I. & Asch, K. 2002. The status of digital geological mapping in Europe: The results of a census of the digital mapping coverage, approaches and standards of 29 European geological survey organisations in the year 2000. Computers & Geosciences, **28**, 783-788.

Jones, R.R., Larsen, S.J. & Hansen, G.I. 1998. An Expert System for Building Virtual Copies of Real World Environments. In: Cantu, F.J., Soto, R., Liebowitz, J. & Sucar, E. (eds) *Expert Systems: Application of Advanced Information Technologies*, Cognizant, New York, 248-255.

Kidd, A.L. (ed.) 1987. *Knowledge Elicitation for Expert Systems: A Practical Handbook*. Plenum Press, New York, 194 pp.

Kuhn, T. S. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 172 pp.

Lakatos, I. 1974. Falsification and the Methodology of Scientific Research Programmes. In: Lakatos, I & Musgrave, A. (eds) *Criticism and the Growth of Knowledge*. Cambridge University Press, Cambridge, 91-196.

Liebowitz, J. & Beckman, T. 1998. *Knowledge Organizations*. CRC Press, Florida, 197 pp.

Longley, P.A., Goodchild, M.F., Maguire, D.J. & Rhind, D.W. 2001. *Geographic Information Systems and Science*. Wiley & Sons Ltd, Chichester, 454 pp.

Luo, X. & Dimitrakopoulos, R. 2003. Data-driven fuzzy analysis in quantitative resource assessment. Computers & Geosciences, **29**, 3-13.

Maerten, L., Pollard, D.D. & Maerten, F. 2001. Digital Mapping of three dimensional structures of the Chimney Rock fault system, central Utah. Journal of Structural Geology, **23**, 585-592.

Mallet, J.L. 1992. GOCAD: A computer aided design program for geological applications. In: Turner, A.K. (ed) *Three-Dimensional Modeling with Geoscientific Information Systems*. NATO ASI Series C: Mathematical and Physical Sciences, **354**. Kluwer Academic Publishers, Dordrecht, Netherlands, 123-141.

McCaffrey, K.J.W, Holdsworth, R.E., Clegg, P., Jones, R.R. & Wilson, R. 2003. Using digital mapping and 3D visualization to enhance and modernise undergraduate fieldwork experience. Planet Magazine, Special Edition 5.

McClay, K.R. 1987. *The Mapping of Geological Structures*. Geological Society of London Handbook Series, Open University Press, 161 pp.

McNoleg, O. 1996. The integration of GIS, remote sensing, expert systems and adaptive co-kriging for environmental habitat modelling of the Highland Haggis using object oriented, fuzzy logic and neural network techniques. Computers and Geosciences, **22**, 585-588.

McNoleg, O. 1998. Professor Oleg McNoleg's guide to the successful use of Geographical Information Systems: (Ten ways to say nothing with GIS). International Journal of Geographical Information Science, **12**, 429-430.

Murnion, S. 1996. Spatial analysis using unsupervised neural networks. Computers & Geosciences, **22**, 1027-1031.

Nonaka, I. 1994. A Dynamic Theory of Organizational Knowledge Creation. Organization Science, **5**, 14-37.

Nonaka, I. & Takeuchi, H. 1995. *The knowledge-creating company: how Japanese companies create the dynamics of innovation*. Oxford University Press, 284 pp.

Nordlund, U. 1999. FUZZIM: forward stratigraphic modeling made simple. Computers & Geosciences, **25**, 449-456.

Ouenes, A. 2000. Practical application of fuzzy logic and neural networks to fractured reservoir characterization. Computers & Geosciences, **26**, 953-962.

Peddle, D.R. 2002. Optimisation of multisource data analysis: An example using evidential reasoning for GIS data classification. Computers & Geosciences, **28**, 45-52.

Polanyi, M. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. Routledge & Kegan Paul, London, 428 pp.

Polanyi, M. 1966. *The Tacit Dimension*. Routledge & Kegan Paul, London, 108 pp.

Pundt, H. & Brinkkotter-Runde, K. 2000. Visualization of spatial data for field based GIS. Computers & Geosciences, **26**, 51-56.

Ramsay, J.G. & Huber, M.I. 1987. *The Techniques of Modern Structural Geology. Volume 2: Folds and Fractures*. Academic Press, 700 pp.

Rhind, D.W. 1992. Spatial data handling in the Geosciences. In: Turner, A.K. (ed) *Three-Dimensional Modeling with Geoscientific Information Systems*. NATO ASI Series C: Mathematical and Physical Sciences, **354**. Kluwer Academic Publishers, Dordrecht, Netherlands, 13-27.

Schetselaar, E.M. 1995. Computerized field-data capture and GIS analysis for generation of cross sections in 3-D perspective views. Computers & Geosciences, **21**, 687-701.

Smith, W. 1801. General Map of Strata in England and Wales.

Smith, W. 1815. A Delineation of the Strata of England and Wales, with part of Scotland.

Struik, L.C., Atrens, A. & Haynes, A. 1991. Hand-held computer as a field notebook and its integration with the Ontario Geological Survey's ‘FIELDLOG’ program. In: Geological Survey of Canada, Current Research, Part A, 279-284.

Turner, A.K. 1992. *Three-Dimensional Modeling with Geoscientific Information Systems*. NATO ASI Series C: Mathematical and Physical Sciences, **354**. Kluwer Academic Publishers, Dordrecht, Netherlands, 443 pp.

Turner, A.K. 2000. Geoscientific modeling: past, present and future. In: Coburn, T.C. & Yarus, J.M. (eds) *Geographic information systems in petroleum exploration and development.* AAPG Computer Applications in Geology, **4**, 27-36.

White, S. & Sleeman, D. 1999. A Constraint-Based Approach to the Description of Competence. In: Fensel, D. & Studer R. (eds) *Knowledge Acquisition, Modeling and Management*: 11th European Workshop, EKAW 99. Lecture Notes in Computer Science, **1621**. Springer Verlag, 291-308.

van der Spek, R. & Spijkervet, A. 1997. *Knowledge Management: Dealing Intelligently with Knowledge Management Network* (CIBIT/CSC), Utrecht. 55 pp.

Xu, X, Battacharya, J.A., Davies, R.K. & Aiken, C.L.V. 2000. Digital Geologic Mapping of the Ferron sandstone, Muddy Creek, Utah, with GPS and reflectorless Laser Rangefinders. GPS Solutions, **5**, 15-23.