BMR
GEOLOGY AND GEOPHYSICS
AUSTRALIA

BMR RECORD 1992/27

GEOGRAPHIC INFORMATION SYSTEMS, CARTOGRAPHIC AND GEOSCIENCE DATA STANDARDS

BUREAU OF MINERAL RESOURCES, GEOLOGY AND GEOPHYSICS

18–20 MARCH 1992

WORKSHOP PROCEEDINGS

GEOGRAPHIC INFORMATION SYSTEMS, CARTOGRAPHIC AND GEOSCIENCE DATA STANDARDS

BUREAU OF MINERAL RESOURCES, GEOLOGY AND GEOPHYSICS

18-20 MARCH 1992

WORKSHOP PROCEEDINGS

BMR RECORD 1992/27

ISSN 0811-062X
ISBN 0 642 17585 3

Page

interpretation of the geology of a region, without providing the background observational data. A variety of proprietary data sets available for analysis with basic geological information may well change the interpretation. The availability of digital geological data is the pre-requisite for digital spatial analysis.

John Hillier - Bureau of Mineral Resources, Canberra 93
GIS and digital cartography: output options, printing on demand and relative pricing. The production of geoscientific maps can be a long and costly process. Advances in computer graphics technology has impacted upon cartography so that it is now possible to undertake all cartographic processes from initial design and compilation to colour proofing and production of composite overlays with the new technology. Repetitive and tedious time consuming tasks are eliminated to let the cartographer concentrate on creative aspects of map publishing and integration of disparate geoscientific data sets through the wealth of opportunities presented in Integrated Geoscientific Spatial Databases and Information Systems. An exciting development amongst these is the opportunity to produce customised colour maps on demand from the digital database, especially colour electrostatic product, both paper and film based - which can be kept maintained and updated quickly as new information is added to the knowledge base. The session describes current and future technology and new applications which these new and exciting systems may discover.

In the year 2092 will it be written "I am told there are people who have never seen a geological map?", or will maps become merely "a regressive parent of GIS"?

Session 3: Integration of Spatial Data Sets. Chairman: Professor Colin Reeves

Prame Chopra and John Creasey- Bureau of Mineral Resources, Canberra --- ***Integrating raster data into a spatial data system.*** 103
In this session we will discuss the raster data domain and its role in integrated geoscience studies. We will highlight the strengths and weaknesses of raster data as it is currently handled and we will demonstrate the major ways in which raster data can be usefully integrated into an overall spatial data system.

Robyn Gallagher - Bureau of Mineral Resources, Canberra 121
The unexpected benefits and problems of data integration. A discussion of the less obvious benefits of data integration using GIS: flexibility of data storage and display; combination of layers; data integrity checking. Also a consideration of some of the problems which may be encountered: scale and projection; image display mechanisms; data input; 3-dimensional data sets.

Brian Lees - Australian National University, Canberra 125
Current progress in the modelling of integrated spatial data using artificial intelligence techniques. Integrating data with very disparate characteristics requires analytical techniques which make only minimal assumptions about data structures. Decision tree and rule-induction will be discussed and the potential advantages of neural networks introduced.

Opening Remarks to the Workshop

Professor R W R Rutland

Executive Director
Bureau of Mineral Resources,
Geology and Geophysics

Welcome to the workshop. This very good response to the workshop is heartening, and evidence of a developing maturity of application that is growing among the users of the new technology.

The Bureau of Mineral Resources' involvement with spatial data management technology commenced with the introduction, in 1985, of the Intergraph system for its cartographic needs. In 1989 NRIC was established to serve the need for ministers to have access to a common, consistent information base when debating resource and environmental policy issues.

The Bureau is currently embarking on a major initiative, in concert with the State and Territory Geological Surveys, to produce a new series of geological map products to meet Australia's requirements into the 21st century. These new map products are to be based on a new program of geophysical data acquisition and satellite and aircraft borne passive remote sensing.

A major challenge facing project scientists is discovering new methods to analyse these data sets, new techniques for their presentation, and efficient ways of transferring the information to our client groups.

This workshop has been convened as a joint initiative sponsored by BMR, GGDPAC (the Government Geologists Database Policy Advisory Committee) and the State/Territory Surveys, to address four major issues facing the National Geoscience Mapping Accord:

> Firstly, the use of computer-based methods for handling the large volumes of spatial data generated by the new mapping programs. Geographic Information Systems - perhaps better thought of as spatial data management systems - are providing a new means of both managing and analysing geoscience map data. However, the use of these systems has raised the issue of how best to structure the data so that its information content is best realised.
>
> Secondly, the move away from manual map production to digital data collection, machine-assisted compilation and computer aided map production is leading to changes in the way work is organised and the way scientists and cartographers interact.

Thirdly, to ensure that data can be freely exchanged between the various state and federal bodies, client groups and even internally within organisations, a common language - or standards for digital maps - must be developed and promoted. It is also important to ensure that there are no impediments to data exchange imposed at the government level. Public good data sets should be made as freely available for exchange between agencies as is possible. A focus on cost recovery is missing the point - cost recovery can only ever be a tiny fraction of the cost of collection.

Fourthly, the range of products and data layers that could be incorporated in an information system is vast. The essential datasets for geological interpretation must be made available for analysis. The complex links between spatial data sets, raster-based data and the more traditional relational databases need investigation to help improve the reliability of interpretation, improve the efficiency of map production, and develop new products. Strategies are also required to cope with continual rapid developments in technology, to maximise the value of purchases and avoid redundancy.

The range of topics you are covering is indeed large, and the recommendations from the workshop will provide us with the basis for on-going discussions and experimentation with these new methods of data management.

Before closing, it is instructive to reflect on one's early training in observation - sketching formations as a student. We must take care not to treat the systems we build as solutions. They are but a tool, offering great potential and providing discipline, but not a substitute for careful observation and scientific thought. The field geologist will always need to observe, think and interpret. But now there is the promise of access to data which provide an objective base against which to judge field observations.

We hope that the workshop will lead to a better understanding of the way our various organisations are approaching the issues at hand, and lead to a common approach within the Mapping Accord, aimed at a major advancement of the National Geoscience knowledge base.

As well, it is important to remember that we are also part of a wider community grappling with similar issues. Communication on these matters is occurring internationally, via IUGS subcommissions involved with geoscience information technology, documentation and standards.

SECOND GENERATION MAPS AND THE NATIONAL GEOSCIENCE MAPPING ACCORD

A.L. Jaques

Bureau of Mineral Resources, Geology and Geophysics
GPO Box 378,
Canberra City, ACT 2601.

Abstract - The National Geoscience Mapping Accord is a joint Commonwealth-State/Territory initiative to produce a new generation of geoscientific maps, datasets and other information using modern technology of strategically important regions of Australia over the next 20 years.

Second generation mapping under the NGMA encompasses both the preparation of new generation geological maps based on high resolution airborne geophysics and other remotely sensed data and the development of new map products through a multidisciplinary approach as derivative maps through the use of Geographic Information Systems.

These new maps and datasets will provide a much more comprehensive understanding of the geology of Australia and its mineral and petroleum resources and contribute to strategies for sustainable development.

Introduction

The National Geoscience Mapping Accord, endorsed by the Australian (now Australian and New Zealand) Minerals and Energy Council in August 1990, is a joint Commonwealth-State/Territory initiative to produce using modern technology a new generation of geoscientific maps, datasets and other information of strategically important regions of Australia over the next 20 years.

Under the Accord BMR and State and Territory geological surveys have identified areas of high priority for mineral and petroleum exploration and/or where significant issues of land use exist. Collaborative 5 year work programs have commenced on 6 NGMA projects in Mineral Provinces (North Queensland, Eastern Goldfields, Arnhem Land, Kimberley-Arunta, Lachlan-Kanmantoo Fold Belts and Musgrave), and 3 in Petroleum Provinces (the Canning, East Australian, and Otway Basins) and work will shortly commence in the Officer Basin. The distribution of current NGMA projects is shown in Figure 1.

The objective of the National Geoscience Mapping Accord is to produce a new generation of geoscience maps and datasets which will underpin mineral and petroleum exploration and at the same time provide a sound basis for resource assessment and for the development of sound land use management strategies (Rutland, 1989).

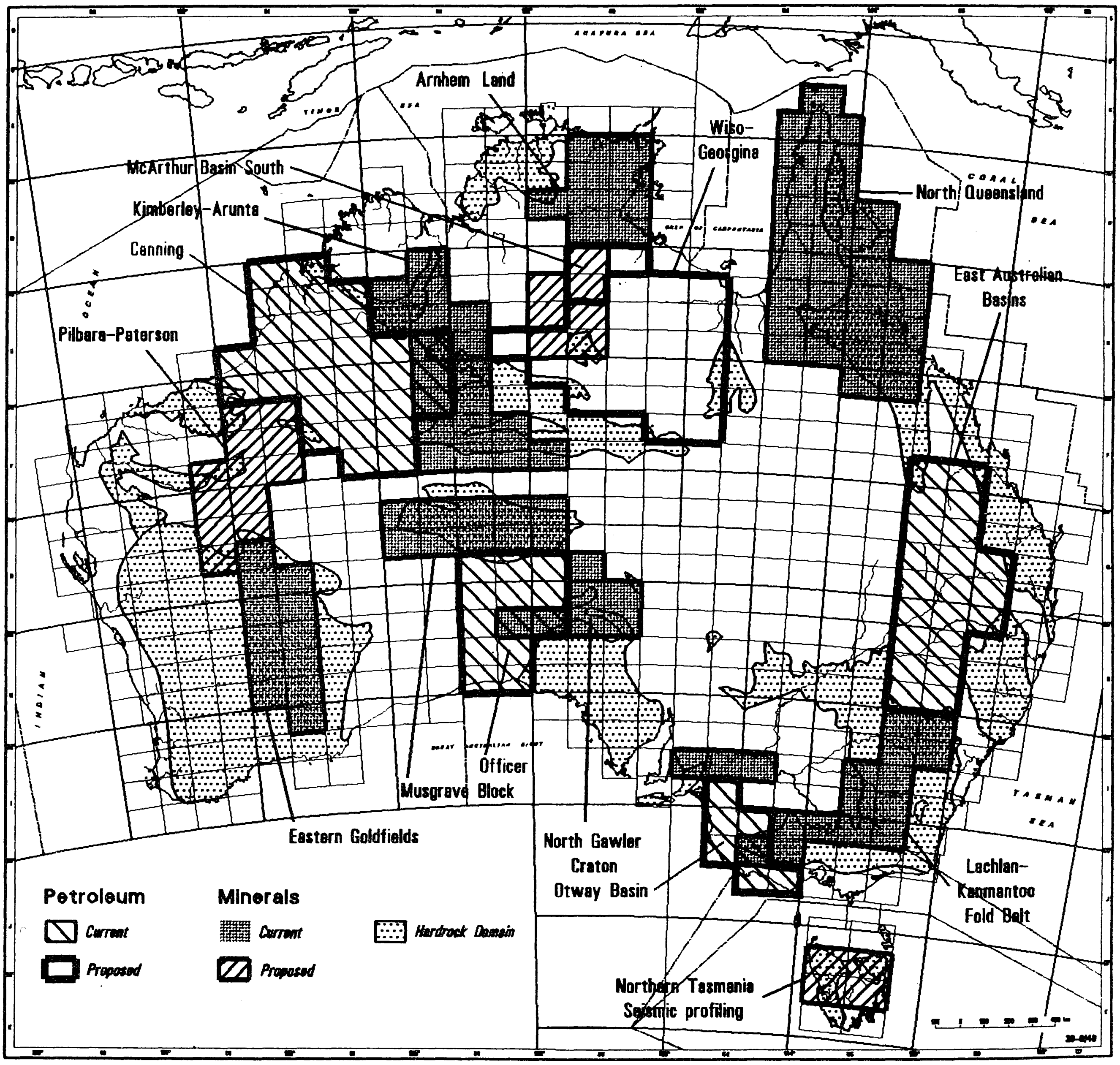


Figure 1. Distribution of current and planned Minerals and Petroleum projects under the National Geoscience Mapping Accord.

The purpose of this Workshop is to bring together the different disciplines involved in the provision of "second generation" geoscientific maps. As an introduction to the Workshop this paper will briefly explore the concept of "Second Generation Maps" within the context of the National Geoscience Mapping Accord.

First Generation Maps

Within the context of the National Geoscience Mapping Accord, "first generation" geoscientific maps are generally considered to be the 1:250 000 scale geological maps produced by the BMR and State Geological Surveys in the 1950s, 60s and 70s.

This mapping program provided a knowledge base, admittedly at reconnaissance level, of the geology of Australia and has been of immense value to mineral and petroleum exploration. The mining and energy resource industries, of course, are vital components of our economy, and currently account for approximately half of Australia's export income.

The first generation mapping was based mainly on field traverses using 4-wheel drive vehicles and aerial photography - mostly black and white RC9 photography at approximately 1:80 000 scale but later colour at 1:25 000 and larger scales. The reconnaissance mapping was supported, particularly in later years, by laboratory investigations such as palaeontology, geochemistry and geochronology.

Second Generation Mapping - a multidisciplinary approach

A multidisciplinary approach has been adopted for the current generation of mapping under the NGMA. Underpinning the "second generation" mapping of Mineral provinces are new airborne magnetic and radiometric datasets (flown at the NGMA standard of 400m line spacing and 100m flight height) supported by remotely sensed (satellite) imagery, regolith terrain mapping, and geochronological, geochemical, structural geology, and metallogenic studies.

The new aeromagnetic and radiometric data are being acquired over the mineral provinces using either the BMR survey aircraft or by geophysical companies under contract to the BMR and/or State and Northern Territory Surveys. In petroleum provinces the objective is to bring the regional 1.5km line spacing airborne magnetic and radiometric data up to modern standards. Since August 1990 more than 540 000 line km of new airborne magnetic and radiometric data have been acquired in support of the NGMA.

Major advances were made in the application of high resolution aeromagnetic surveys coupled with image processing techniques to mineral exploration in Australia in the mid-1980s (e.g. Isles et al., 1989). High resolution aeromagnetic surveys first employed in Precambrian gold exploration programs have now become an integral component of most mineral exploration in Australia. The value of integrated geophysical/geological mapping to the NGMA was recognised early (Cook et al., 1989; Webster et al., 1989) and the contribution of airborne magnetic data in providing more detailed structural and lithological information, particularly in areas of poor outcrop or thin cover, has been reinforced in each of the NGMA projects in which BMR is involved. One specific example of the improved mapping provided by the integration of new high resolution airborne magnetic and radiometric data is on the Bathurst 1:250 000 sheet in the Lachlan-Kanmantoo Fold Belt Mapping Accord project where the geophysics has not only enabled a better definition of lithologic boundaries than indicated on the 1st edition 1:250 000 geological map sheet but also provided important new discrimination of lithology. A greyscale image of the magnetic data clearly distinguishes discrete plutons within the Oberon granite, and their relative timing, which was not previously recognized. Similarly, radiometric data over the same area provide a powerful discriminant of Silurian shales (higher intensity) from the Ordovician sediments and permit more precise location, and in some cases re-definition, of geological boundaries.

High resolution radiometric data gathered under the NGMA is also providing a powerful new tool for differentiating regolith types based on their Th, K and U signatures. On the Ebagoola 1:250 000 sheet in the North Queensland NGMA project the ability of the radiometric signal to discriminate, even in the presence of a thick vegetation cover, has provided a considerable advantage. Image processing techniques have been used to enhance the radiometric data and integrate the radiometrics with the high frequency component (band 5) of Landsat TM imagery (Wilford et al., in prep.). This has allowed definition of broad lithological units and structural domains, with the radiometric data providing information on the weathering and geomorphic history of the region. Based on this success NGMA standard radiometric data are being routinely employed in BMR's regolith mapping in other NGMA projects and are also being used in a pilot application to soils mapping in the joint project on sustainable land use in the Murray-Darling Basin centred on the Wagga 1:100 000 sheet involving BMR, CSIRO Division of Soils, NSW Conservation and Land Management and the Centre for Resource and Environment Studies Australian National University.

As well as the airborne geophysics other remotely sensed datasets, particularly satellite imagery such as Landsat Thematic Mapper (TM) and Spot, also provide fundamental datasets to support the NGMA mapping (Simpson, 1989). Important as the satellite imagery is in providing a synoptic view, trials in North Queensland of high resolution (10m pixel) panchromatic SPOT imagery indicate that, at least in that area where there is a lack of dynamic range due to dense rain forest, satellite imagery is not a surrogate for aerial photography (BMR 91).

Other new technology being applied in the collection of geoscientific field data are:
- the use of hand-held Global Positioning System (GPS) units for accurate location in the field (BMR 91)
- the recording of field rock property information such as magnetic susceptibility and scintillometer measurements
- the use of computers to capture data in the field such as REGMAP, the PC-based field data management system (Withnall, this volume; BMR 91).

This focus on digital technology, a multidisciplinary approach, and the underpinning of the mapping by remotely sensed data from aircraft and satellite platforms leads directly to the real issue of this Workshop.

Geographic Information Systems, Databases and Computer Cartography

GIS provides the means by which the new remotely sensed data can be integrated with field and other multidisciplinary data held in databases and with digital cartography to generate maps, interpretations and models. Although a relatively new technology GIS is already revolutionizing the way spatially referenced data are handled and used. In Australia the main application of GIS to date has been to land management in the public sector through the development of Land Information Systems (LIS), particularly at the state level, although significant progress has also been made in natural resources and environmental applications (O'Callaghan and Garner, 1991). In the geosciences the major LIS/GIS application has been in the development of computerised State mining and exploration tenements.

It is because of its capacity to integrate spatially referenced data and produce products as maps that GIS technology was identified early as fundamental to the production of second generation maps under the NGMA (Cook, 1989; Etheridge, 1989). In the context of the NGMA, GIS provides the interface between the geological map and other spatially referenced datasets such as remotely sensed data with computerised cartography, database management and computer modelling (Fig. 2).

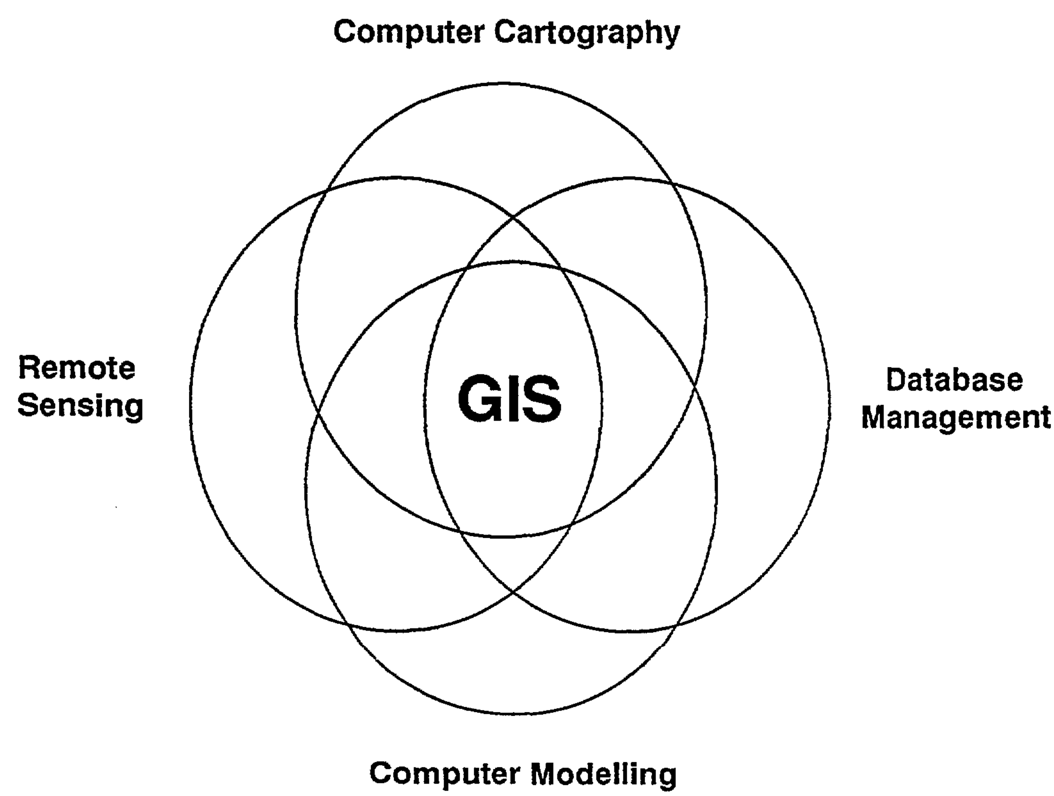


Figure 2. Schematic representation of the relationship between Geographic Information Systems and remotely sensed data (aircraft and satellite), computer modelling, database management and computer cartography (modified after Maguire, 1991).

Rapid advances are being made in each of these areas. The purpose of this workshop is to examine the methodology being developed at both the Commonwealth and State/Territory level to ensure that developments in the acquisition, storage and handling of data acquired under the NGMA will permit the maximum benefit to be obtained from it. The multidisciplinary approach to the NGMA means that large amounts of relatively complex data are being acquired at considerable expense. Hence there is a need to ensure that the methods and data structures being employed are as far as possible consistent and compatible to permit the effective interchange and application of all available data.

Databases

Database Management Systems (DBMS) are an integrated and crucial component of any successful GIS. DBMS are used to store, manipulate and retrieve data from a database. A range of DBMS structures are in use in GIS but relational databases currently dominate the field (e.g. Healey, 1991; Ryburn, this volume). In terms of proprietary systems the most common relational DBMS currently in use in geological surveys and the major mineral exploration companies in Australia appear to be ORACLE, INGRES, and DB2.

BMR has devoted considerable resources to establishing corporate geoscientific relational databases (Ryburn and Lenz, 1991), and this has required the development of new and in some cases the redesign of older databases to support an effective GIS (BMR 91). Similarly, the South Australian Department of Minerals and Energy (SADME) are in the process of implementing a comprehensive corporate geoscience information system (Belperio, this volume). These and similar commitments by other government agencies and mineral and petroleum exploration companies are a recognition of the value of primary geoscientific data as a long-term corporate resource.

Papers presented in this Workshop address many of the key issues concerning the integration of databases in the GIS. These include capturing geological field data in digital format (Mitchell, this volume) including the use of the REGMAP field data management system developed by the Geological Survey of Queensland (Withnall, this volume) and the structuring of ORACLE databases to standardise, store and manipulate geological field data (Blewett and Ryburn, this volume; Lenz and Pain, this volume).

Of fundamental importance in establishing corporate and especially national databases is the question of data standards. The Government Geological Database Policy Advisory Committee (GGDPAC) is charged with providing advice on the management and exchange of geoscientific data. GGDPAC has established a Geoscience Data Standards Working Group to identify core geosciences and related cartographic data types and to suggest Australian standards for their representation and definition, based where possible on existing standards, authorities and common practice (Schmidt, this volume). The development of national data standards will result in improved data consistency, integrity and availability.

The sharing of diverse geoscientific datasets is particularly important under the NGMA but complicated by the diversity of hardware, system software and application software now being used by both Commonwealth and State agencies and by industry. The proposed Australian Spatial Data Transfer Standard (ASDTS), which is based on the US Spatial Data Transfer Standard and replaces the pre-GIS Australian mapping standard AS 2482, will provide an important basis for the interchange of digital geoscientific data (Wood and Foley, this volume).

Digital Cartography

Advances in computer cartography now enable the complete design and production of geological maps by digital methods. The advantages offered by computer cartography - greater efficiency, the ability to easily update maps, and the capacity to integrate data in ways not previously possible - have been widely recognised and digital cartographic methods are being employed in the production of maps under the NGMA (Hillier, this volume).

Papers given at this workshop address a number of important issues in digital cartography including:
- digital map compilation methodology
- digital cartographic standards
- digital topographic data - standards and availability
- standard maps series - i.e. the 1:250 000 series
- customised maps - opportunities and limitations
- exchange standards

Spatial Information Systems for the Geosciences

The application of GIS in the geosciences in Australia is relatively new. However, it is clear from experience overseas and from pilot projects that it is proving to be a powerful and versatile tool for the integration and analysis of spatially referenced geological data, especially at the regional scale.

Substantial progress has already been made in the preparation and release of digital geological maps. SADME have released the state geological map SA_GEOLOGY as a GIS which is used to supplement the 1:250 000 series and to provide specialist and thematic geological maps for a range of applications in exploration and land use planning (Parker, this volume). BMR have released 5 1:100 000 geological maps as ARC/INFO coverages (Williams and Gallagher, this volume).

Successful implementation of GIS technology as spatial data management systems has been identified as a major objective of the NGMA. Much of this workshop is aimed exploring the ways in which the different organisations are developing GIS with a view to developing compatible and consistent methodology.

Key issues examined in the workshop include:
- integration of raster data and the raster-vector dichotomy (Chopra and Creasey, this volume)
- problems associated with handling the large size of remotely sensed datasets (for example, 262Mb for a Landsat Thematic Mapper (TM) scene

comprising a full 7 bands, 150Mb for 400m line spacing airborne magnetic and radiometric coverage of a 1:250 000 sheet - random access ungridded; Reeves, this volume)
- the integrity of spatial data (Brodie and Tucker, this volume)
- the building of key layers
- spatial data standards (Nanninga, this volume)
- definition of data items (Gallagher, this volume)
- interpolation of datasets
- exchange of datasets
- links between digital cartography and GIS
- the availability of spatial data

The multidisciplinary approach being adopted in the NGMA is generating new geoscientific datasets which can form multiple layers in a GIS (Rattenbury et al., this volume). This is shown schematically in Fig. 3 where multiple layers of remotely sensed data (airborne geophysics, satellite data) are combined with geological, geochemical, regolith, and mineral deposit data to produce an assessment of resource potential.

Incorporation of data on land tenure, vegetation and other cultural layers will provide a powerful knowledge base which can contribute substantially to the environmental debate and sustainable development strategies.

GIS offers the capacity to generate second generation maps using the other definition of "second generation" - i.e. derivative maps. In this context GIS is now and will in future derivative maps for:
- land use planning; e.g. the Cape York Peninsula Land Use Strategy - CYPLUS - involves the integration of multiple datasets in a GIS
- hazard mitigation; flood, fire, drought and volcanic hazards
- development planning
- environmental planning and management; environmental impacts
- metallogenic and resource assessment studies; the gold resource studies of Nova Scotia and Saskatchewan by Agterberg and Bonham-Carter (1990)
- geological modelling; for example the USGS Quebec-Gulf of Maine lithosphere transect project (Stewart, 1989)
- modelling of sedimentary basins
- modelling of hydrocarbon reservoirs
- tectonic analysis
- modelling of landscape development - regolith modelling
- hydrogeology

and many other applications.

Most current GIS have only 2-D capability but permit 2.5-D visualisation in applications such as digital elevation models. However, attention is now focussing on the design and implementation of 3-D GIS in a range of geoscientific application areas (Raper and Kelk, 1991). The impetus has been the development of new processor architecture (especially RISC technology) and the development of new data structures. Recent developments have been in oil exploration, mining, hydrogeology and geological modelling as well as environmental and other applications (Raper and Kelk, 1991) and further developments in the field of geosciences in particular can be anticipated.

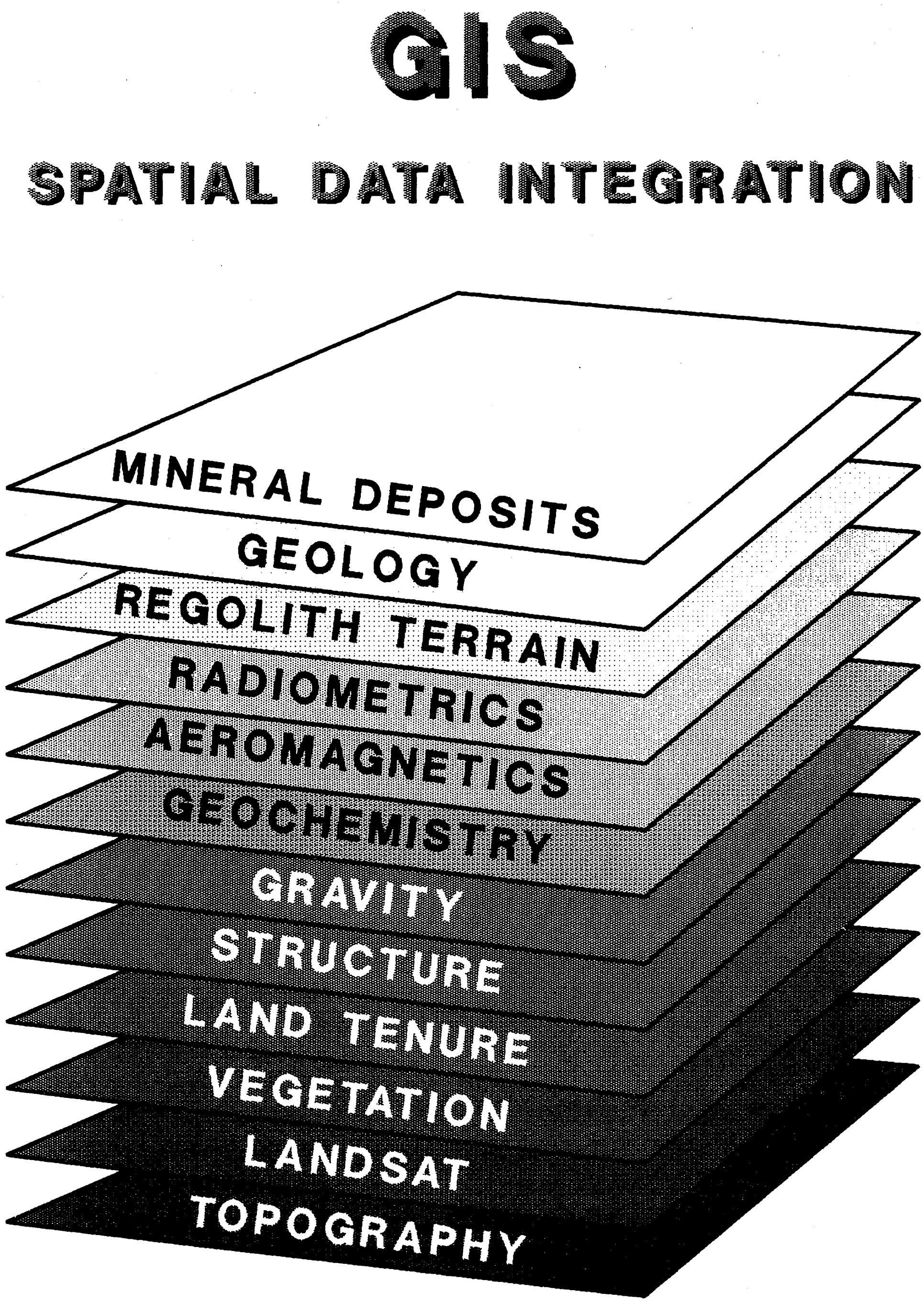


Figure 3. Schematic representation of the application of GIS to integrate multiple datasets as 'layers' in a GIS.

Data Exchange and Availability

A fundamental limitation to the widespread application of GIS is the availability of appropriate digital data. In a major contribution to establishing an inventory of available resource data NRIC has established the FINDAR directory system to provide up-to-date information on what resource data is available, and its location, custodian, and accessibility (Shelley, this volume). Another potential limitation on the availability of data is the price of digital data in Australia which reflects the pressures on many Government agencies to achieve cost recovery through sale of products. In contrast, there has been a greater recognition by some countries, notably the USA, of the emerging need for large volumes of spatial data, particularly geographic, natural resource and environmental data, and a growing interdependence among data users and data producers (Bradbury, this volume).

These issues are also addressed in the workshop within the context of the scope for the advent of GIS to transform land and resources planning and management in this country and a proposal made for a multi-disciplinary, multi-agency data collection strategy (Kelly, this volume).

Conclusion

Second generation mapping under the NGMA encompasses both definitions of "Second Generation". It is both the repainting of the geological maps of Australia using a new brush in the form of new technology - particularly high resolution airborne geophysics and other remotely sensed data - and the development of new map products in the form of second generation or derivative maps through the use of GIS.

The development of second generation maps under the NGMA presents us with a major and exciting challenge. The new technology opens up new horizons and presents us with tremendous opportunities to develop a much more comprehensive understanding of the geology of Australia and its mineral and petroleum resources and to contribute to strategies for sustainable development.

References

Bonham-Carter, G.F., Agterberg, F.P. and Wright, D.F., 1990. Weights of evidence modelling: a new approach to mapping mineral potential. Geological Survey of Canada Paper 89-9, pp. 171-183.

BMR 91 - Yearbook of the Bureau of Mineral Resources, Geology and Geophysics, Australia.

Cook, A., Pridmore, D.F., Hallberg, J.A., Isles, D.J., and Ross, J.R., 1989. Advances in geological mapping and interpretation through the high-resolution aeromagnetic method. BMR Research Symposium 1989.

Cook, P.J., 1989. Possible directions for Australian geoscience maps and databases - an introduction. BMR Research Symposium 1989.

Etheridge, M.A., 1989. Recent and future trends in geoscience mapping. BMR Research Newsletter 11, 1-3.

Healey, R.G., 1991. Database management systems. In Maguire, D.J. et al. (eds), Geographic Information Systems: Principles and Applications; Vol. 1, Principles, pp. 251-267. Longman Scientific and Technical, New York.

Hillier, J., 1989. Current and future computer-aided drafting in BMR. BMR Research Symposium 1989.

Isles, D.J., Harman, P.G., and Cunneen, J.P., 1989. The contribution of high resolution aeromagnetics to Archaean gold exploration in the Kalgoolie region, Western Australia. Economic Geology, Monograph 6, 389-397.

Maguire, D.J., 1991. An overview and definition of GIS. In Maguire, D.J. et al. (eds), Geographic Information Systems: Principles and Applications; Vol. 1, Principles, pp. 9-20. Longman Scientific and Technical, New York.

O'Callaghan, J.F. and Garner, B.J., 1991. Land and geographic information systems in Australia. In Maguire, D.J. et al. (eds), Geographic Information Systems: Principles and Applications; Vol. 2, Applications, pp. 57-70. Longman Scientific and Technical, New York.

Raper, J.F. and Kelk, B., 1991. Three-dimensional GIS. In Maguire, D.J. et al. (eds), Geographic Information Systems: Principles and Applications; Vol. 1, Principles, pp. 299-317. Longman Scientific and Technical, New York.

Rutland, R.W.R., 1989. Future directions for BMR geoscience mapping: the National Geoscience Mapping Accord. BMR Research Symposium 1989.

Ryburn, R.J., and Lenz, S.L., 1991. Geoscientific relational databases in BMR and the 'client-server' method. Proceedings of the National Conference on the Management of Geoscientific Information and Data, Adelaide, Session 7, 1-15.

Simpson, C., 1989. Remote sensing. Current capabilities and limitations. BMR Research Symposium 1989.

Stewart, D., 1989. 3-D GIS - a new approach to data integration for geoscience mapping. BMR Research Symposium 1989.

Webster, S.S., Woods, K.T., and Wyborn, D., 1989. Integrated geophysical/geological mapping interpretation: case histories from the Charters Towers region and elsewhere. BMR Research Symposium 1989.

GEOGRAPHIC INFORMATION SYSTEMS TOWARDS A BETTER TOOLBOX

Ian P Musto
National Resource Information Centre
Canberra

Abstract - Major developments in geographic information systems in recent years have substantially increased their capacity for data integration in addition to their traditional functions of spatial database management, analysis and cartographic output.

These new facilities, together with improvements in usability, create opportunities for improved data analysis and output under programmes such as the Nation Geoscience Mapping accord. These same improvements in the support technologies place increased emphasis on management of spatial data and a common understanding of opportunities and problems.

Introduction

In the past three years there have been major increases in the capability of commercial geographic information system (GIS) software, particularly in their ability to integrate a wide range of spatial data. In addition to their traditional functions of spatial database management, vector data analysis using topological geometry and high-quality output of map products, GIS can now provide major benefits in the integration and cross-analysis of a wide range of spatially located data.

It is now widely claimed, and frequently demanded, that information systems support a 'corporate' view of our data. Though loosely used, we define this to mean that the system should provide information appropriate to the decision-making process at all levels from the individual section, through a wider programme view to institutional, state and national perspectives. While the level of complexity of issue obviously increases dramatically as we pass through this continuum, the basic data and systems requirements differ only in size and in the requirements for data integration.

With this increased capability, GIS has become much more relevant to programmes such as the Geosciences Mapping Accord, and indeed to other national data supply programmes. The purpose of this paper is to critically examine the current state of the commercially available GIS software in light of the capability required to support this particular 'corporate' view.

Data Integration

In general terms, we can identify five major groups of spatial or spatially-referenced data relevant to geosciences applications. These data are:

- vector data,
- raster data,
- image data,
- database systems including a spatial component, and,
- document and textural data.

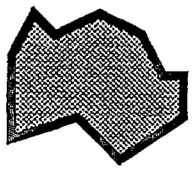
Integration of Vector and Raster Data

Traditionally GIS software has utilised one of two major conceptual or data models, either a vector-topological model (sometimes referred to as a spatial-relational model) or a raster model. The vector-topological model has been employed in most large, database-oriented GIS systems while the raster or grid model has been employed in image analysis systems or GISs oriented towards spatial analysis.

The vector model stores geographic entities according to their 'natural' form, i.e.,

Points stored as single x,y coordinate pair

Lines stored as series of x,y coordinate pairs



Polygons (irregular-shaped areas) stored as a series of lines making a 'closed' homogenous geographic entity.

Each geographic entity feature has an unique identifier which forms the link to descriptive data stored in relational database tables. The descriptive data may be of any form or amount relevant to the feature.

The raster or grid model divides the world into discrete, regular units called cells or pixels. Each cell is given a value which corresponds to a measurement of the phenomena dominant for that whole cell.

Because of the regular nature of the raster, it is unnecessary to store every geographic coordinate but rather the raster s accessed by row and column from a known geographic origin.

Data stored in rasters may form discrete classes such as soil type or may form a continuously varying surface such as elevation or aero-magnetics and gravity.

Image analysis systems often utilize rasters corresponding to individual bands of reflectance from remotely sensed platforms.

The various merits of vector versus raster data models have been hotly debated for a number of years in a quite pointless dichotomy - both data models have advantages for certain purposes and the only relevant measure is efficiency for the job in hand.

GIS software may now employ both models in an integrated spatial environment allowing for the following functionality:

- import and export of proprietary and non-proprietary image and raster formats
- geographic registration, rectification and projection of raster data,
- geographic overlay (graphic) of vector and raster data,
- cross-analysis of raster data,
- hardcopy output of combined vector and raster data.

A number of GIS vendors are now moving to a common syntax for modelling languages associated with raster data thus allowing for a consistent implementation of modelling algorithms.

Integration of Surface Data and Analysis

Surface formation and analysis occupies much of our attention in the natural resources application of spatial information systems, whether because we sample data from a continuously varying phenomena, as in the case of aero-magnetic data, or because surfaces such as elevation are a primary datasets in environmental analysis and prediction.

Facilities for surface formation and analysis in GIS software have progressed rapidly in recent years in response to demand from users. Major GISs now support advanced surface formation and analysis functionality, including:

- preparation of surfaces from regular or irregular sample data,
- preparation of surfaces from large datasets (greater than 500,000 sample points),
- surface construction models appropriate to the form of the sample data and to the nature of the surface, e.g. Triangulated Irregular Networks (TINs), Inverse Distance Weighting, Kriging, Trend Fitting, etc.,
- integration of surfaces with raster and vector data,
- advanced surface analysis facilities including; slope, aspect, viewshed, stacked profiling, volume calculation, etc.,
- 3-dimensional visualisation of surfaces with raster and vector data, e.g. draping a Landsat TM image over a 3D elevation surface or draping a vector geology map over an TMI surface.

The use of 3-dimensional visualisation of spatial datasets has proven to be an invaluable tool in investigating relationships between datasets. As the tools in this area improve, and we assemble a greater pool of experience this trend is expected to accelerate.

Integration of Corporate Database Systems

All GIS systems utilising a vector-topological model use a relational database management structure for the storage of attribute data associated with geographic entities. Frequently these internal database systems have been highly optimised to support storage and retrieval of geographic coordinates.

In early versions of GIS software, the user was restricted to the use of the internal database system. This was seen as a disadvantage as it acted against a corporate database structure which included non-spatial databases or where spatial interrogation was not the primary database access mode.

The adoption of the Structured Query Language (SQL) as an industry standard has made it possible to developed a standardised interface between the internal database structure and other relational database systems. Thus modern GIS systems include SQL database interfaces which provide the following functionality:

- database of choice for the storage of attribute data,
- a client-server model providing for use of SQL interrogation syntax in addition to the systems own syntax,
- multiple-level table relates,
- incorporation of higher level tools as they become available under the industry standard, e.g. forms interfaces and the Binary Large OBject (BLOB) for the storage of documents in image format within databases.

Database design to provide optimum retrieval in a GIS context may not be the same as that for non-spatial retrieval and careful attention, including the use of prototypes, should be given to database design to suit a range of data retrieval functions.

The use of GISs to manage databases that include spatial data but have not traditionally been incorporated into the GIS can provide efficiencies in error tracking and updating.

Integration of Document and Text Data

In the scientific use of GIS, we have a requirement to incorporate the accumulated knowledge contained in the literature into our total spatial information framework.

While requirements for text integration are broad, they can be considered in two parts. Firstly to associate a scanned image of a document by assigning to a specific spatial feature and secondly to include a full document in text format so that the internal text structure may be interrogated as part of the total spatial interrogation strategy.

It is early days in the integration of text storage and retrieval systems and as yet no single industry standard has emerged for the storage of scanned documents or for full text. This has hampered the development of facilities within the GIS environment but the following facilities are available for scanned documents:

- linking of scanned documents to geographic entities,
- linking of scanned documents to attributes or attribute classes,
- pageing or scrolling of documents within the GIS display environment,
- hardcopy of documents form within display environment.

In the area of text retrieval, the following functionality is desirable if not fully implemented at present:

- linking of text database to geographic entities,
- linking of text database to attributes or attribute classes,
- ability to perform text or key word search, or as part of total spatial database interrogation,
- links to existing bibliographic systems.

It should be noted that these facilities exist within the FINDAR software which supports the National Directory of Australian Resources developed and maintained by NRIC, though acting at the metadata level.

Trends in Software

Predicting trends in software development is always difficult and in GIS, which has a market growth rate of 25 - 40% per annum, is particularly difficult. However, in terms of new capability to support natural resource applications there are a number of clear trends, viz:

- implementation of a 2.5 D data model,
- enhancement of facilities for 3 D visualisation,
- faster, more efficient spatial intersection algorithms,
- more flexible attribute structure structures for spatial entities, including dynamic segmentation,
- more powerful modelling languages and high-order modelling tools in the raster environment.

The question of 3 dimensional structures within GIS remains problematic with no definitive list of requirements available. Most GIS vendors have committed to supporting 2.5 D coordinate structures, which provide for the storage of a **Z** value for every x,y coordinate pair. This structure provides the facility to store and analyse elevations at points, depth on linear structures such as faults and to define a thickness at polygon boundaries. It does not define irregular polygons as true solids for which we must include the shape of top and bottom surfaces and total volume.

Including z values into the vector model will incur a major computing overhead as z values must be interpolated for all intersections as well as calculating the x and y intercepts.

Even given a satisfactory conclusion to the debate as to the most appropriate interpolation method (some favour linear or quadratic functions, while others favour the smoothing function of a high-order polynomial interpolation), computing effort would be increased by at least an order of magnitude over the existing 2 D data model. This is a heavy overhead if not required.

Collaborative work between vendors and user agencies, particularly in the geosciences should provide us with a clearer picture of requirements and the balance between 3 D data models and the need for 3 D visualisation in the near future.

What is now called dynamic segmentation has been implemented recently by a number of major GIS vendors. Dynamic segmentation provides a facility to assign descriptive data to overlapping segments of line networks to points associated with linear features. Multiple layers or points can be assigned to any one location on the feature thus providing more flexibility in the attribute structure.

While dynamic segmentation was originally developed to meet the requirements of the transportation industry, as in other GIS facilities, it has wide application in natural resource application. Multiple descriptors for fault structures is one obvious application in the geosciences.

The other major trend that will substantially influence the way we facilitate access to natural resources data, is the recent release by at least three vendors of a new class of spatial information software. This software which is probably most appropriately called 'viewing software' is sold separately from and at far-reduced cost from GIS software, yet provides users with only minimal training with sophisticated tools for visualisation and analysis of data in GIS form.

This software, while still in its infancy, provides facilities for viewing and analysis of both vector and raster data in proprietary and non-proprietary formats and already is being bundled with datasets by overseas data supply agencies.

When the availability of this software is considered with the development of new non-proprietary, executable data formats, such as the Vector Product Format used by the Digital Chart of the World, we can expect an increasing demand from users not currently in the GIS 'community' for direct, low-cost access to spatial databases.

Trends in Hardware

To try to accurately predict developments in hardware would be sheer folly but a number of general trends are clear as they relate to the GIS environment:

- huge increase in CPU and floating-point performance,
- multiple processor technologies in servers and desktop workstations,
- increase in memory capacity and decrease in memory cost,
- decrease in disk storage costs,
- availability of a wide range of storage devices,
- increase in backup media capacity.

The other major trend in the GIS community is the almost total dominance of the UNIX workstation as the platform for GIS at the agency or corporate level. GIS vendors now support their software on a wide range of UNIX workstations providing increased choice, particularly in regard to corporate computing strategies.

The major decrease in disk storage cost, availability of large capacity magneto-optical disks and large capacity backup tape devices has provided us with the ability to properly maintain large on-line GIS databases such as those necessary to support the Geosciences Mapping Accord.

It is quite clear that hardware is not currently a limiting factor in our application of GIS technology and is likely to influence our decisions less in future.

Current Status

In general terms, modern GIS technology provides most, if not all, facilities required to integrate large and complex spatial datasets to meet 'corporate' information requirements.

The UNIX workstation hardware platforms can provide cost-effective processing and data storage appropriate to very large spatial databases, with options for distributed or centralised processing as part of an integrated computing strategy.

Corporate databases structured for limited spatial interrogation can be incorporated within the spatial information framework.

Thus, we may consider that GIS technology provides a mature information support framework for spatial data which in most situations will provide significant benefits at reasonable cost. We are no longer primarily limited by shortcoming in the software and hardware.

Issues

The issues in the application of GIS technology towards the National Geosciences Mapping Accord are complex and will be addressed in detail by other authors within this workspace but a number of general issues are clear:

- the need to properly analyse the data and information requirments to support the programme and to ensure that these functions are supported logistically,
- the need to discuss and agree to the appropriate 'corporate' view of the data,
- the urgent need to ensure rigorous evaluation of current data integrity (including accuracy, precision and consistency) and to ensure that data uptake and processing procedures are in place to maintain this integrity,
- the need to support efforts to remove impediments to data availability, and,
- the need to recognise the need for training and education programmes to support staff at all levels with involvement in spatial information technologies.

The common thread to all these issues is that they are all 'human'. They are functions of our perceptions or the way we currently perform our functions. While spatial information systems, and specifically GIS, can provide us with a better toolbox it is up to us to determine the manner in which they can be applied to our best advantage.

Resourcing for a GIS : A data users perspective

By: Graeme Mitchell
CRA Exploration Pty Ltd
Western / Central Region
18th March, 1992

Abstract

The successful implementation of a GIS in an exploration company requires the consideration of a wide range of issues. What is required from the GIS, who are to be its users and what systems, support and training needs will they have. Remotely sensed image data, geochemical information, geological & structural maps, mineral occurrence databases and other datasets all need to be integrated with minimal duplication of storage. Commitment on personnel issues from management together with an "easy to use" computer interface are both essential to the success of any GIS project.

Introduction

When planning the investment of time and money in a Geographic Information System, it is extremely important to consider all of the planning, personnel, software and hardware issues well in advance of making any purchases. The decision to proceed with the development of a GIS is often made without proper thought being given to these issues, often resulting in less than satisfactory advancement being made or even total system failure.

Operational Issues

Who are the Users?

The most important issue to address is who will be the main users of the GIS, which narrows the choice of system. If it was to be used in a "contract data processing" type of environment, with staff dedicated to its operation, the complexity of a system is not as critical as with casual users. To gain the most from a GIS in a geological environment it is preferable that users familiar with the geological problems can use the GIS as a tool, rather than rely on "white coat" specialists.

What Are the Aims of the GIS

Many problems occur when users of a GIS become disillusioned in the system because of its failure to meet their initial unrealistic expectations. Careful analysis of their needs is required to ensure they are aware of any limitations. It can be difficult however, to determine system limitations before purchasing a product.

If users only require data integration facilities, then GIS may be an overkill, and a CAD system could suffice. However, beware of this approach as this initial requirement often grows into more ambitious demands of database queries and analysis, way beyond the scope of a CAD system.

Data Requirements

Data availability / capture issues need to be examined to judge if the cost of accumulating the data required for a task can be justified. Whilst there is a vast amount of data available in the public domain, the unrealistic prices charged by groups like AUSLIG make it very difficult to justify its purchase. The Land tenure data is a case in point. Thankfully some of the State government agencies have a more realistic approach to digital data pricing, which encourages the use and development of digital maps.

Users should be made aware of the time consuming nature of data capture, as it is generally months, sometimes years before a GIS is fully operational.

Output Mechanisms

If the GIS integration capabilities are fully exploited with the use of magnetic / TM image backdrops, a reasonable plotter must be purchased. Pen plotters obviously cannot handle rastor output and Thermal, Ink Jet or Electrostatic types are preferred. The most effective type for large format colour is the electrostatic plotters available from Versatec, Calcomp and others. These are very expensive to purchase and maintain, but are invaluable if any reasonable volume of plots are to be generated.

User Interface

Simple & Intuitive

The critical factors in gaining user acceptance of a system are its ease of use and functionality. Most geotechnical staff fall into the "casual user" category, and hence it is important that the learning curve is not too arduous. They need to easily pick up the program again after several weeks / months in the field.

The most acceptable type of interface is the "point and click" style made popular by Macintosh. Most GIS developers are heading in this direction, but they still have a long way to go.

User Friendly with a Geological Focus

The continuing trend is for computers to become more acceptable to the new computer user, rather than the old perception of people having to become computer literate. The advent of pen based "tablet" computers is further evidence of this trend towards a format that people are more comfortable with. Following on this theme it would be more acceptable to geotechnical staff if they could make queries of a system in geological terms, rather than computer jargon. To achieve this it is necessary for geological "intelligence" to be built into the data structures to allow such queries as discussed by (Currie & Ady 1991)

Personnel Management Issues

Staff issues

At least two people in an organisation should have a good overall understanding in all aspects of the GIS operation. This is necessary to provide development expertise and the product skills required to support new users. It is also a foil to the unhealthy situation of having the total system reliant on one person. It is very important to select the right type of person as "Champion" of the GIS, to ensure the many problems encountered aFe resolved. The person should have some computing expertise together with a good knowledge of the GIS application areas. Having the tenacity of a bulldog would also help.

Management Commitment

Sometimes the most difficult issue is convincing management of the need to allocate several months of the GIS "champions" time for training and initial setup of the GIS. It generally results that little useful output occurs during this phase, which tests the patience of management. A well documented schedule needs to be put in place outlining to management the timing of the implementation stages, and the resources required to meet them.

Data Structures

It is important to have a well defined data model of all the data in the GIS and that to be accessed from external data bases. Definition of Geological data structures is a very difficult task trying to cover the varied geological environments at a range of scales. Whilst it is preferable to have geology stored at the greatest detail available, time, data avilability and cost pressures usually result in data being captured at a more regional scale. The capture scale also effectively controls the detail of the geological description recorded, which may result in problems at a later date if a greater level of detail is required. Geological "intelligence" needs to be added to enable geotechnical staff to present queries to the system couched in geological terms rather than computer jargon. Whilst some GIS systems provide sophisticated tools in this area, currently no GIS handles inheritance where the subgroups of a class retain the properties of the parent.(Currie & Ady 1991) This makes it difficult to submit a query that will cover the whole range of possible rocktypes included in a particular classification. These problems can be partly addressed via relational attributes stored within the host database, but it is difficult to build an acceptable "user friendly" front end to this type of system.

Data Management

Storage Location / data availability

The physical storage location of data files needs to be considered, to ensure high availability to users. This is largely dependent on the geographical spread of the potential users, and the computer network facilities in place. If the network cannot support data transfer at sufficient speeds, publishing of the data on some form of optical storage could be considered. It is essential that the data is easily available to the user, and the best form of access is via a workstation on the users desk.

Formats

No single GIS system available today is capable of providing all of the solutions required in a geotechnical environment. Other systems therefore need access to the underlying base data and hence file and data formats are very important. The translation of data from one format to another is often required to provide input to various systems. This translation generally results in a lot of time wasted for little apparent gain, and a lot of storage duplication. This also raises the issue of currency problems - which is the latest version of this file ??

Rastor Data

The GIS should be able to handle the generic file formats of existing in-house and publicly available systems to minimise duplication of storage. If this is not supported it can cause major image storage and handling problems with the voluminous nature of remotely sensed data.

Point

Point data sets vary widely in their structure, and their attributes need to be managed by a flexible DBMS. Such databases need to provide the user with the ability to easily create a new database, modify or add data to an existing one or link to other external databases or programs for reporting or other purposes. Data standards are important to enable the easy integration of this data.

Vector

The vector data can generally be stored within the GIS because of the easy export facilities available. The attributes however may be better stored in an external database, as other systems may need to have easy access to them for operational purposes such as tenement maintenance. This depends entirely on the type of data captured and the external system needs.

Textual

The storage of textual data is one of the more difficult areas of data storage. Relational databases have some difficulty in handling the long strings involved and also providing querying facilities on those strings. There is a lot of research currently into document scanning and Optical Text Recognition to find the most effective means for storing the huge volumes of textual data stored in reports and technical documents. Effective affordable solutions are still a number of years away. The best solution available today is to avoid attempting to store the full body of the report, but store a summary and a direction of where to find the full report.

File Management

Most GIS systems generate a large number of files during the theme generation, and a flexible file naming convention is required to enable easy management. This should be able to differentiate between data types, project, geographic projection, version number and the stage of processing the file was used for. Organisation of files into "Global", "Project" and "Personal" locations is desirable for general access control and verification purposes. The Global and Project categories have access restrictions placed on them to ensure no corruption of the verified data files.

Systems Integration

No "Black Box" technology

To encourage widespread use of a GIS, the hardware must be readily available, without expensive specialist boards and CPU's.

Integrated with existing systems

The ability to easily transfer data between the GIS and other Image Processing, CAD and Desktop Publishing systems is important. Whilst a GIS might be able to execute a task, it may not always be the most efficient route to take. Easy transfer between applications allows the user to take advantage of the best tools for the task at hand, improving productivity.

Training

Focus the Training at the User Level

Most GIS systems have many tools available that are rarely needed by the general user. Training should be focused at the users requirements, taking common problems they face and demonstrating those facets of the GIS they can use to arrive at the required solution. This means several levels of training are needed for casual and intensive users, focusing on their different requirements.

Computer Based Training

A major problem faced by geotechnical staff is the casual requirement for use of systems like GIS. This can mean retraining every couple of months when the need to use the system arises. One solution to this is the development of a Computer Based training / reminder system that uses "multi-media" type tools. These can provide the casual user with "on the job" assistance as they attempt to run the GIS. New developments are progressing fast in this area, particularly with the new "QuickTime" video functionality recently released by Apple on the Macintosh.

Conclusions

The resourcing considerations for a GIS vary dependent on the application area and scope of the GIS application. This document however outlines some of the general issues that cross application areas, while focusing on the mineral exploration environment. Whilst deliberation of these points will certainly not guarantee a problem free introduction to GIS, it will highlight some of the issues that need to be addressed to ensure a smoother implementation.

References

Alaster Currie & Bridget Ady 1991: Geosis Project: Knowledge representation and data structures for geoscientific data.Geosis Newsletler No 9 Summer 1991.

GIS, CARTOGRAPHIC AND DATA STANDARDS WORKSHOP
BMR, CANBERRA 18 - 20 MARCH 1992

GIS: a data providers perspective

by

P C Kelly, Deputy Surveyor General of NSW
Land Information Centre
Department of Conservation and Land Management

In 1985 I prepared a paper for the International Geoscience Information Conference in Adelaide. It was entitled "Better Data for Decisionmaking - Implications for the Geosciences". It outlined what GIS was, and what it could do for geoscientists. There is no need to go over that background today because the utility of GIS in the geosciences is now well recognised.

At the conference, I posed three challenges to those involved in geoscience information:

(a) they must assess the impact of GIS concepts and technology on the collection, management and presentation of geoscience and related information;

That has been largely done, the process will continue during this workshop.

(b) geoscientists must assess their need for data derived from outside their discipline, and make these needs known.

Geoscience information people have become more active in issues affecting data availability, typically through Federal and State co-ordination of data collection activities.

(c) a harder, but analogous task is for geoscience information specialists to ascertain general needs of users outside the geosciences for their data.

This is a major issue which needs to be considered in your planning of data collection and development of products in the context of the National Geoscience Mapping Accord.

In NSW, key land information agencies have developed strategies to prioritise needs for data, and to co-ordinate the collection of the data necessary to meet these needs.

Let us have no allusions. The days of individual agencies pursuing large scale data collections, without recourse to external data sources or oblivious to potential external users, are over. Governments cannot afford duplication of effort. Micro-economic reform agendas being followed by all Australian governments, mean that single agency, Commonwealth only, State only, local government only, data capture programs are being relegated to the scrap heaps of history.

In all States and Territories, initial emphasis for common jurisdiction wide data collections has been on land administration applications. Land ownership, values and boundaries have predominated. For example, in NSW full State coverage of data on all parcels of land will be essentially complete within two years. Other jurisdictions have similar timeframes.

When I spoke in 1985, the massive cadastral databases covering Australia were still a dream. Soon they will be a reality. All the talk about corporate data systems, better decisionmaking processes, data as a major resource, have come to pass. Governments now recognise this, an example is the $30m provided to the Land Information Centre to accelerate capture of the Digital Cadastral Database in NSW.

Cadastral data is only a beginning. In NSW we are proceeding with the concept of a Natural Resource Inventory (NRI). The Inventory will contain the layers of information used by many agencies to address issues related to natural resources use and management. Layers will contain information giving the location and basic attributes of land form, economic infrastructure and resource occurrence.

The layers to be collected follow closely the list compiled at a national workshop held in Adelaide two years ago, convened by the Australian Land Information Council. The layers will need to be collected by a number of agencies. The keys for success are to stick to the basic information needs of a broad user community, to work to predefined standards so that the data collected by the various agencies can be integrated (i.e. the layers fit together), and to make a realistic assessment in deciding what will constitute the "first edition" of the NRI data without compromising longer term data collection strategies. You will note that geology and soils are in the list of basic layers.

What do data collections like the NRI mean for geoscience users? It means that they do not need to compile basic layers of data such as land surface, transportation, hydrology or cadastre. These will be provided by others who have primary responsibility to collect, and keep up to date, this data. It means they can concentrate their resources on compiling the data directly related to their field. It means concentrating expenditure on smaller systems geared to using data, rather than collecting data.

We are only in the early stages of planning for the NRI. Steps taken so far include:

- An economic appraisal of the costs and benefits of collecting the basic "framework" data - relief, transportation network, streams and administrative boundaries - used to "register" other data sets. The outcome was a benefit/cost ratio of 7:1, mainly based on avoided duplication of collection.

- Discussions on transfer and pricing issues.

- Sorting out who the key players are in each layer.

The challenge now is to decide on priority of collection programs and data standards. Pilot studies testing data integration have already begun.

In developing the NRI, we welcome input from Commonwealth agencies, and are actively seeking input from bodies such as the Murray Darling Basin Commission. At the end of the day, high quality data sources compiled by the States will need to satisfy the priorities and requirements of regional and national users.

The challenge for the NRI, and for all data collection programs, is to ensure user requirements are met. Often this is difficult. Enunciation of need implies perfect knowledge of future directions in corporate function, new technology, community expectations and funding. I have never yet seen an occasion when users had perfect knowledge, and that as a data provider, I was 100% sure that user expectations were being met.

I don't mean to inject a religious note in the proceedings, but at the end it does come down to faith. Faith that users will move away from dependence on static maps and embrace the flexibility provided by digital data; that they will assess their basic information needs and want new products (and not rely on standardised products forced upon them in the past); and they will recognise that it is not just about drawing lines over maps or measuring something off a map, but using the power of information systems to combine, analyse, assess and present information in entirely new ways.

That is the market that is emerging, and that is the market we, as information professionals, should be addressing.

In the Land Information Centre, putting this strategy in place has involved:

- putting under one manager existing photogrammetric, topographic mapping and remote sensing functions;
- merging single technology programs into one, aimed at meeting user needs for basic spatial data, particularly in natural resource applications using the right mixture of technologies;
- ensuring a balanced investment in people, equipment and systems to provide efficient data collection;
- ensuring the data produced is of high quality, integrity and timeliness, but most importantly, is capable of providing a flexible product range to meet specific user needs; and
- building the electronic models of the real world using GIS technology, and at the same time investing in the infrastructure that will improve the model and enable people to navigate around the model, e.g. GPS differential network.

In conclusion, I renew my challenges for those involved in geoscience information.

- Continue to embrace GIS technology and techniques.
- Take a multi-disciplinary, multi-agency approach to data collection and use.
- Do not forget the wider audience for geoscience data.

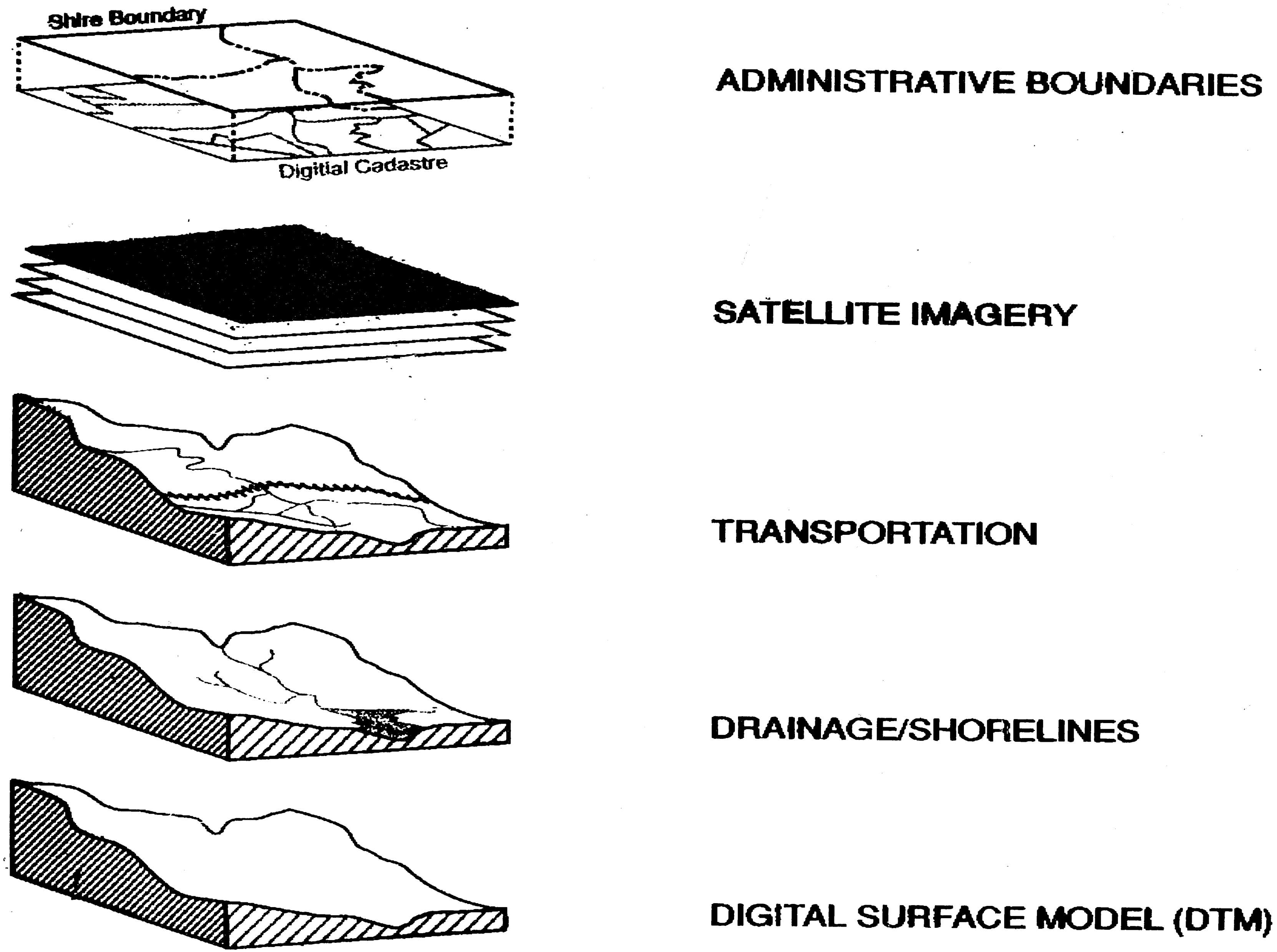
THREE CHALLENGES FOR GEOSCIENCE INFORMATION

1. ASSESS THE IMPACT OF GIS

2. ASSESS NEEDS FOR DATA FROM OUTSIDE BODIES, MAKE THESE NEEDS KNOWN

3. ASCERTAIN NEEDS OF NON-GEOSCIENCE PEOPLE FOR GEOSCIENCE DATA

NSW PRIMARY SPATIAL DATABASE



THREE NEW CHALLENGES FOR GEOSCIENCE INFORMATION

1. CONTINUE TO EMBRACE GIS TECHNOLOGY AND TECHNIQUES

2. TAKE A MULTI-DISCIPLINARY, MULTI-AGENCY APPROACH TO DATA COLLECTION AND USE

3. DO NOT FORGET THE WIDER AUDIENCE FOR GEOSCIENCE DATA

LAND INFORMATION CENTRE
DEPARTMENT OF LANDS

Digital Natural Resource Information

* Provide digital data for individual basic spatial layers of natural resource information (e.g. relief, roads, drainage, boundaries) on a regional or whole State scale in as short a time frame as possible

* Put updating processes in place at the beginning of the process.

* This data will usually be "unintelligent", non-topologically structured back ground data initially, but capable of further development.

* Further enhance these digital data to introduce more structure and therefore the ability to support more "intelligent" usage of the data

* Provide an enhanced range of standard and customised products in both digital and hard copy

* Help individual agencies develop in-house information systems which will accept these basic spatial data and combine with their own application specific data to meet their operational and policy decision-making needs.

Precis of
a Cost Benefit Study

by

Price Waterhouse Urwick

into the provision of

a State-Wide Digital
Primary Spatial Database

October 1992

LAYER	KEY DATA SUPPLIERS	
	SPATIAL LOCATION	ATTRIBUTES
Relief*	LIC	LIC
Streams/Drainage/ Catchments*	LIC/DWR LIC/Water Boards	Water Boards DWR
Coastline*	LIC, MSB	LIC MSB
Transportation* (Roads)	LIC	RTA
Transportation* (Railways)	LIC	SRA
Other major economic infrastructure	LIC/Utility Services	Utility Services
Land Cover/Vegetation	Forestry/Ag/NPWS/LIC others	Forestry/Ag/NPWS others
Restrictive Sites (Public Lands, special tenures)	LIC, Mineral Resources NPWS, C&LM, others	C&LM, NPWS, Mineral Resources, others
Landuse	Needs to be defined	Needs to be defined
Soils	SCS	SCS
Geology/Mining	Mineral Resources	Mineral Resources
Climate	?	?
Administrative Boundaries*	LIC/others	LIC/others
Statistical Boundaries	ABS/LIC	ABS
Cadastre	LIC	C&LM, WLC, LTO, VG, Mineral Resources, etc.
Feature Names	LIC	LIC (GNB)
Other Culture/Details	Needs to be defined	Needs to be defined

Covered in part by Primary Spatial Database proposal or the Digital Cadastral Database

Note: Source data is at indicated scale unless otherwise specified.

GEOGRAPHIC INFORMATION SYSTEMS: KEY TO DATA MANAGEMENT FOR 2ND GENERATION MAP PRODUCTS

Robyn Gallagher

GISolutions / Bureau of Mineral Resources, Geology & Geophysics

Abstract - A summary of the system in place in the Minerals and Land Use Program at BMR for using GIS to store, export and create cartographic products from geotechnical data sets.

Keywords - Geographic Information Systems / data dictionary / data export

Introduction

Based on initial GIS work by Peter Williams, Mark Rattenbury and other geologists of the Eastern Goldfields Group at BMR a system for managing geological data has been created with the following major components:
- a data dictionary embodying a detailed description of each data set
- a collection of standard data sets for a number of 1:100 000 map sheets
- GIS utilities to export the standard data sets and create cartographic products using them

This has been implemented using Arc/Info, and the menu utilities were created using Arc Macro Language. There is currently no RDBIO link to existing Oracle data bases: some auxiliary data sets are imported into Info from Oracle, but the longer term goal is to establish SQL*NET links. The Info tables which will soon disappear are therefore kept as simple as possible.

Data Dictionary

The following data layers (coverages) have been defined:
- geology polygons
- linear structures ... faults, folds, linears, veins&dykes, horizons
- small (point) structures ... bedding, cleavage, foliation, igneous layering, axial surfaces, minor faults, minor veins, joints, minor folds, lineation
- BMR sites
- mineral deposits from MINLOC
- geochemistry samples from ROCKCHEM
- cartographic coverages ... map frame, sheet border, AMG grid, map graticule, annotation, legend, mapsurrounds

The data dictionary for a coverage lists its spatial components, defines each of its attributes, and describes its related or expansion tables. An example of the data dictionary entry for the geological boundary polygons is in Appendix A.

Coverages for which the mapping groups are custodians also have (Info) descriptor files which stay with the coverage throughout its life. Each coverage has two tables which together describe the contents of the data set and its attributes in a detailed manner. The entries in the tables correspond to the information required for the FINDAR system supported by NRIC. An example of the descriptor files for the geological boundary polygons is in Appendix B.

Standard Data Sets

Several data sets now exist in the format specified in the data dictionary. These are mainly for 1:100 000 maps in the Eastern Goldfields area. Because work began on mapping those sheets before the data dictionary was established, the coverages were individually modified to conform to the standard.

In the meantime a menu-driven utility has been developed to help create standard coverages. This ensures that the data sets are exactly in the data dictionary format: in particular the attributes and expansion tables are defined correctly; and the descriptor tables are created and filled with minimal stress on the geologists.

Export and Hardcopy

The existence of standard data sets allows routine data export and map creation. A utility has been developed to convert a set of standard coverages to any of the supported export formats: Arc/Info export, Autocad DXF, Intergraph design file and DLG3. Use of a non-GIS format naturally causes loss of information. However a great deal of effort was spent in ensuring as much descriptive information as possible is included with those formats, even if it is only as extra ASCII files.

A utility will be developed to create in-house cartographic products via Arc/Info using the standard data sets, based on macros currently used by individual geologists. As well, a scheme is being implemented for transferring data layers as standard Intergraph design files to BMR's Cartographic Services Unit for quality map production: transfer of design files is already possible in a general sense.

Future Directions

Obviously what has been achieved so far is only a beginning. The data dictionary will be amended as experience with the data sets grows, and its applicability to other data types such as regolith or to data from other mapping areas is already being investigated.

The direct link to Oracle must be implemented as soon as the appropriate hardware and software components are in place. The upgrade of Arc/Info to version 6 and the arrival of ArcView will also influence future work.

Finally, the use of the data sets together with raster images opens up a whole area of analysis and modelling using GIS.

Appendix A

Data Dictionary Entry for Geological Boundary Polygons

Coverage Name :	GEOLPG_n
Features :	polygons
	arcs

Polygon attributes :

geolpg_n-id	4,5,B	
polylabel	4,5,B	unique identifier used instead of -id
map_symbol	6,6,C	

Related expansion table :

map_symbol	6,6,C	used for relates
rock_unit_name	64,64,C	
legend_text	256,256,C	
rock_unit_descn	1024,1024,C	

Arc attributes :

geolpg_n-id	4,5,B	
bmr_code	6,6,I	see notes
class	1,1,I	optional qualifier for BMR symbols code
*** redefined item for plotting		
big_code	7,7,I	concatenation of bmr_code & class

Because the coverage may be built as arcs and/or polygons, there is no need for a second coverage GEOLLN, which no-one was using. Arc attributes are expressed in terms of BMR symbols codes, if available.

It is considered imperative to keep an early copy of the polygon linework, before cleaning, attributing etc., as a backup.

Appendix B
Descriptor Files for Geological Boundaries Coverage

A.1.NAME	=GEOLOGICAL BOUNDARIES MT MASON 2939
A.2.REFERENCE	=GEOPLG
A.3.OTHER.NAME	=
A.4.ABSTRACT	=GEOLOGICAL BOUNDARIES FOR THE MT MASON 100000 SHEET
A.5.OWNER	=BUREAU OF MINERAL RESOURCES, GEOLOGY & GEOPHYSICS
A.6.OTHER.ORG	=
A.7.OTH.ORG.DESC	=
A.8.AUTHORS	=M DUGGAN
A.9.OTH.DIR.REF	=
A.10.COPY	=NO
A.11.M.D.DIR.ID	=
A.12.DIR.PUBLIC	=YES
A.13.TRANSFER	=NO
C.1.DIR.SPAT.REF	=COMPLETELY
C.2.PROJECT.TYPE	=AMG
C.3.COORD.UNITS	=MTR
C.4.SPAT.FEAT.T	=LIN POL
C.5.TOPOL.STRUCT	=YES
C.6.TOPOL.DESC	=NONE
D.1.GENERAL.AREA	=W A
D.2.ON/OFF.CODE	=ONSHORE
D.3.SPAT.DESC	=MAPSHEET MT MASON 2939
E.1.FED/ST.AREA	=W A
E.2.1.MAP.SHEETS	=100,000 2939 MT MASON
E.2.2.MAP.SHEETS	=
E.2.3.MAP.SHEETS	=
E.2.4.MAP.SHEETS	=
E.5.1.MIN.LONG	=
E.5.2.MIN.LAT	=
E.5.3.MAX.LONG	=
E.5.4.MAX.LAT	=
E.6.1.SW.ZONE	=
E.6.2.SW.EAST	=
E.6.3.SW.NORTH	=
E.6.4.NE.ZONE	=
E.6.5.NE.EAST	=
E.6.6.NE.NORTH	=
F.1.DATA.FORM	=GIS VEC
F.2.WORK.MEDIUM	=MDI
F.3.SUPP.DOC	=
F.4.DIG.DATA	=COMPLETELY
F.5.SOFTWARE	=ARC/INFO
F.6.HARDWARE	=SUN

F.7.INT.FORMATS	=ASCII EXPORT DXF DLG3 IGDS
F.8.SIZE	=100 KBYTES
G.1.CUST.ORG	=BUREAU OF MINERAL RESOURCES, GEOLOGY & GEOPHYSICS
G.2.CUST.POS	=HEAD OF PROGRAM, MLUP
G.3.STATUS	=INPROGRESS
H.1.PRIM.DATA	=COMPLETELY
H.2.DATA.COLL	=RSS RSA FME FNO FSP LME API
H.3.SOURCE.MAT	=
H.4.SOURCE.DESC	=
H.5.DP.DET	=
H.6.POS.ACC	=10 M to 100 M
H.7.POS.ACC.DESC	=GPS & AUSLIG BASE MAPS
H.9.LOG.CONS	=GRAPHICAL COMPARISON WITH OTHER DIGITAL DATA SETS AND IMAGES
H.10.COMPLETE	=BASED ON REGIONAL TRAVERSES AND IMAGE INTERPRETATION
J.1.ACC.RES	=NO RESTRICTIONS
J.2.1.OUT.PROD	=DIGITAL DATA ON MAGNETIC TAPE
J.2.2.OUT.PROD	=
J.2.3.OUT.PROD	=
J.2.4.OUT.PROD	=
J.2.5.OUT.PROD	=
J.3.1.PROD.CH	=$300
J.3.2.PROD.CH	=
J.3.3.PROD.CH	=
J.3.4.PROD.CH	=
J.3.5.PROD.CH	=
J.4.SUPP.ORG	=BUREAU OF MINERAL RESOURCES, GEOLOGY & GEOPHYSICS
J.5.SUPP.POS	=SALES CENTRE
J.6.ORD.PROC	=THROUGH BMR SALES CENTRE
J.7.ORDER.TIME	=2 WEEKS
K.1.1.KEYWORD	=GEOLOGY
K.1.2.KEYWORD	=GEOLOGICAL MAP
K.1.3.KEYWORD	=GEOLOGICAL STRUCTURES
K.1.4.KEYWORD	=
K.1.5.KEYWORD	=
O.1.ORG.NAME	=BUREAU OF MINERAL RESOURCES, GEOLOGY & GEOPHYSICS
O.2.ORG.TYPE	=FEDERAL GOVERNMENT
O.3.ORG.ACC	=BMR
O.4.POS.TITLE	=SALES CENTRE
O.5.POS.ACC	=
O.7.RES.ADD	=ANZAC PARK EAST BDG, CNR CONSTITUTION AVE & ANZAC PDE, PARKES ACT 2600
O.8.POST.ADD	=GPO BOX 378, CANBERRA ACT 2601
O.9.PHONE.NO	=06-249-9519
O.10.FAX.NO	=06-257-6466

```
O.11.EL.MAIL.ADD        =PETERW@BMR.GOV.AU
O.12.NET.NAME   =INTERNET


B.1.1.ITEM.NAME =POLYLABEL
B.2.1.ITEM.DESC         =4,5,B  UNIQUE IDENTIFIER FOR POLYGON
B.3.1.SPAT.RES          =1:25000
B.1.2.ITEM.NAME =MAP_SYMBOL
B.2.2.ITEM.DESC         =6,6,C
B.3.2.SPAT.RES          =1:25000
B.1.3.ITEM.NAME =ROCK_UNIT_NAME
B.2.3.ITEM.DESC         =64 C  ROCK UNIT NAME IN OPTIONAL
                               RELATED TABLE
B.3.3.SPAT.RES          =1:25000
B.1.4.ITEM.NAME =LEGEND_TEXT
B.2.4.ITEM.DESC         =256 C  MAP LEGEND TEXT IN OPTIONAL
                               RELATED TABLE
B.3.4.SPAT.RES          =1:25000
B.1.5.ITEM.NAME =ROCK_UNIT_DESCN
B.2.5.ITEM.DESC         =1024 C  DESCRIPTION OF ROCK UNIT IN
                               OPTIONAL TABLE
B.3.5.SPAT.RES          =1:25000
B.1.6.ITEM.NAME =BMR_CODE
B.2.6.ITEM.DESC         =6,6,I  BMR SYMBOLS CODE FOR ARC
B.3.6.SPAT.RES          =1:25000
B.1.7.ITEM.NAME =CLASS
B.2.7.ITEM.DESC         =1,1,I CLASSIFIER ADDED TO BMR CODE FOR
                        ARC
B.3.7.SPAT.RES          =1:25000
```

SA_GEOLOGY and the future direction of 1:250 000 atlas-series geological maps and explanatory notes

A John Parker
Geological Survey of South Australia
PO Box 151, Eastwood SA 5063

INTRODUCTION

During the late 1950's and early 1960's, Australia adopted 1:250 000 scale (and to a lesser extent 1:100 000 scale) as the basis for a national geological map series. In the 1990's with the growing use of computerised GIS and CAD, the question is do we continue to maintain and update that series (including explanatory notes) or do we restrict our output to digital datasets and/or colour plots.

Printed, particularly published full-colour geological maps form the basis for <u>all</u> geological investigations including mineral and petroleum exploration, groundwater investigations, environmental management, geological research, engineering site investigations, road construction and many other forms of geological and natural resources planning and management.

Printed maps and explanatory notes will never be replaced..
..they will be enhanced by GIS

Although other printed maps and digital datasets including GIS have made substantial contributions to such investigations, printed geological maps are still the fundamental database upon which most exploration and management decisions are made; they are convenient to carry around in the field or on site inspections, they are abstractions of several layers in the geoscientific database containing both spatial and critical aspatial information, and they present an overview of the geology of any particular region. Printed paper maps and accompanying explanatory notes will never be replaced; they will be enhanced by GIS datasets but as a means of summarising and conveying geological information, they are still by far the best medium. In fact, recent surveys among geoscientists throughout Australia have conclusively identified published geological maps, particularly at 1:250 000 and 1:100 000 scales, as the first priority required of state geological surveys and the BMR (Woods, 1989; Parker & Felstead, 1990; Table 1). However, the surveys also identified geological field compilation, structure/tectonic, metallogenic, geochemical and geophysical maps as high priorities. 67% of all respondents to the SA mapping questionnaire also required explanatory notes for 1:250 000 geological maps.

The problem is not one of replacing printed maps, but how to produce them more efficiently at scales and formats which serve all user requirements. Most geological surveys and the BMR can only publish a few full-colour geological maps a year even though geological staff compile many more.

An additional problem with current map compilation and publication procedures is related to timeliness and accuracy. From the time a geologist completes field work etc to the time that map is published, it is often several years; rarely is a map published within two years of completion of field work. Furthermore, once maps have been published, they cannot be easily updated and it is often more than 20 years before 2nd-edition maps are produced. Therefore, given that the "half life" of a geological map is about 10 years, many maps are out of date and do not contain all the available information that could or should be portrayed.

The Geological Survey of South Australia has partly overcome these problems in the past by producing detailed field compilation maps (often at 1:100 000 scale) as progress reports during the mapping process. However, this still does not overcome the problem of ageing maps and these maps are only available as black and white dyeline prints which are not always easily comprehended.

Table 1: Demand for geoscience maps in South Australia by mineral and petroleum industry geologists.

	1st Priority	Scale
Published Geological Maps	59%	
1:1 000 000 Scale		15%
1:250 000 Scale		37%
1:100 000 Scale		32%
1:50 000 Scale		16%
Field Compilation Maps	20%	
Hydrogeology	1%	
Structure/Tectonic	12%	
Metallogenic	3%	
Geochemical	2%	
Drillhole Locations	3%	

The problem of a long shelf life for geological maps is not only related to demand but also to volume of sales. As Table 1 indicates, published 1:100 000 geological maps are in higher demand than 1:1 000 000 maps. However, sales of 1:1 000 000 maps are much larger. Taking South Australia as an example, there is only one 1:1 000 000 geological map covering the entire State, yet there are sixty-eight 1:250 000 maps and four hundred 1:100 000 maps. Therefore, because sales volumes for 1:100 000 geological maps are spread over 400 maps, the number of copies sold for any individual 1:100 000 map is generally very low and leads to an unacceptable long shelf life. The exceptions to this conclusion would be maps of key areas such as Broken Hill or around Adelaide etc.

..to encourage exploration, development and responsible land use
....South Australia introduced SA_GEOLOGY....
to supplement 1:250 000 geological maps and explanatory notes..

To satisfy user needs for high quality, coloured geological maps and hence to further encourage exploration, development and responsible land use, the South Australian Geological Survey resolved that it must either increase map publication or introduce a new system for producing near-publication-quality maps, providing greater flexibility with regard to scale and format, and improving the timeliness and updating of map publication. To achieve a much higher production of full-colour geological maps without increasing personnel and without having maps sitting on the shelf for long periods of time, the South Australian Geological Survey introduced SA_GEOLOGY, a digital geological map system to supplement (not replace) publication of 1:250 000 geological maps and explanatory notes (Parker, 1990).

SA_GEOLOGY GIS

During the last few years, the use of GIS has grown dramatically and they are now commonplace in many government departments, universities and businesses throughout the world. However, the use of GIS in geoscience has not developed at the same rate as in many other disciplines such as environmental science, agricultural science (particularly soil mapping) and local government.

In Australia, most state geological surveys now have GIS in various forms and they are now developing substantial spatial databases which contain large volumes of essential and critical data to mineral and petroleum exploration companies, environmentalists, government and inter-government agencies, land-use planners, and educational and research organisations. These databases are valuable resources not only in what they can offer to potential users but also in terms of the cost of hardware, software and personnel resources that have been used to compile them. These costs and resource requirements are substantial so it is imperative to understand what a GIS is, what its benefits are and how it can be most effectively utilised before embarking on establishing such a system.

To most GIS users, GIS is more than just a software system; it is: *a complete system which comprises hardware, software, data, personnel and procedures designed to support the capture, management, manipulation, analysis, modelling, display and presentation of spatially referenced data for solving complex problems.*

Geological maps are a form of GIS in that they contain much of the essential data that might comprise a typical geological GIS. For example, they contain both spatial and descriptive information, and they provide facilities for analysing and getting more information from the map. However, there are fundamental differences, advantages and disadvantages of maps versus GIS.

..SA_GEOLOGY offers....digital geological map data....
and timely provision of full-colour, near-publication-quality maps..

In South Australia, following an extensive pilot project commencing in 1986, SA_GEOLOGY was developed and is now undergoing final implementation using Arc/Info software on Data General Aviion workstations. The system offers a number of advantages over traditional manual mapping procedures:

- Provision of digital geological map data and associated datasets (and hence enable all GIS analytical functions to be applied to those data and also enable data to be incorporated into image processsing applications)
- Greater flexibility in choice of scale and format for map plots (eg maps can be produced at any reasonable scale and for any area not necessarily restricted to traditional map boundaries)
- Greater flexibility in choice of map style or theme and attributes for output (eg can selectively plot geochemistry on Precambrian geology)
- Provision of relatively cheap (less than $50 per normal-sized map plot) full-colour geological maps of near publication quality (ie negating the need to publish maps for which there might be a high demand yet only low sales volume)
- Timely provision of field compilation maps (and in colour)
- Full edit and update facilities enabling maps to be updated easily whenever new data become available
- Potential time and financial savings in map compilation and publication for those maps that might still be published (eg 1:250 000 series)
- Potential to produce a full range of thematic maps and, in particular, metallogenic, geochemical and drillhole maps

- Potential for development of new data capture and compilation techniques (eg digitising directly onto scanned and rectified aerial photos; input of GPS data).

By contrast, the advantages of a paper geological map are: it is transportable, easy to distribute and use (external users/clients do not require special hardware/software), it is easier to standardise, it is familiar and requires low technology for operational use. All these advantages are essential for field use. Other disadvantages of GIS are: its high initial acquisition cost (hardware/software), continuing maintenance costs, high start-up cost for data capture and conversion of existing data, paucity of expertise and difficulties in transferring data between agencies due to incompatible hardware, software and standards.

During the SA_GEOLOGY pilot project, a number of important conclusions were reached relating to some of these advantages and disadvantages:

- High-resolution graphics workstations are essential for efficient and effective digitising and processing of large datasets (eg a complex 1:100 000 geological map)
- Effective, tailored training courses are essential - these should be tailored to the needs of the user (drafting personnel have different needs to those of geologists)
- Existing branch responsibilities and map compilation procedures do not need to be changed dramatically (Fig. 1) - it is only the methods which change
- There are few (if any) time savings during initial map compilation (Cowley & Freeman, 1991) - the advantages (as above) come mainly down the track
- Detailed geological map data are generally not suitable for small scale (eg 1:1 million) plots - it is still necessary to produce two or more datasets each tailored to the final output/analysis requirements
- Data volumes for spatial geological map data can become quite considerable: in Arc/Info format, geological boundaries, faults, polygons and labels for a standard 1:250 000 map sheet area (eg ADELAIDE) digitised from 1:63 360, 1:50 000 and 1:100 000 compilation sheets is approximately 10 Mb, while the State 1:2 million/1:1 million geological map dataset is approximately 9 Mb (or approximately 20 Mb in Arc/Info export format)
- There are substantial cost benefits compared to traditional publication procedures - for SA, the cost of the proposed system over 5 years will be approximately $500 000 yet to publish an equivalent number of 1:100 000 scale maps would cost at least $1.6 million. There are also substantial savings in the publication process - traditional manual map publication would cost about $20 000 per map yet using computerised methods these costs could be reduced to less than $10 000
- Data security is important - only qualified drafting personnel under the direction of the geologist responsible for a particular map sheet (the "custodian") can modify that map dataset
- Customised user's manuals are essential for all users
- Research and development are an essential component to examine new techniques for data capture etc.

Clearly, two fundamental requirements of a GIS are suitable computer hardware and software. SA_GEOLOGY was originally developed using EASINET or DG/GEO GIS software running on a Data General MV20000 mini-computer. However, that software was not fully supported to the level required and was dropped by Data General leading to our subsequent acquisition of Arc/Info and DG Aviion workstations. While EASINET (an Australian product) was originally an extremely good state-of-the-art 3D GIS, it did not have the user base or international recognition to become a major player in the GIS marketplace. This ultimately contributed to its unfortunate demise. Transferring data and procedures etc to a new GIS has been a major exercise, taking nearly 12 months and setting back the entire development programme.

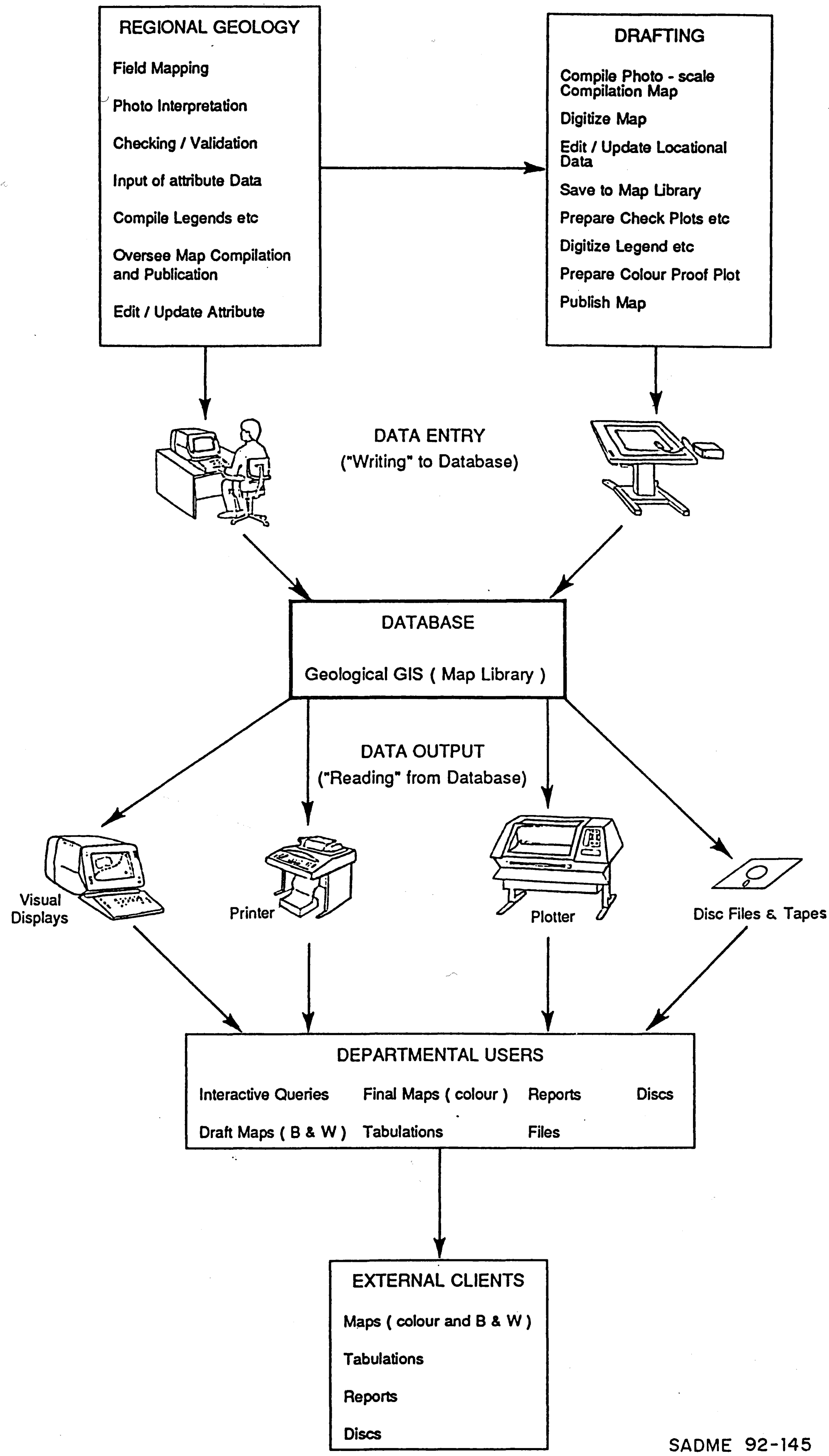


Figure 1: Overview of geological map compilation and GIS procedures.

MAP ELEMENTS AND LAYER STRUCTURE

Geological maps are graphic representations of a geologist's perception of a particular area and contain a number of different components:

- Geographic or locational features (points, lines, polygons) - eg topography, cultural features, hydrography, geological boundaries, faults, fold axes etc, drillhole locations, mine locations etc.
- Descriptive information about those geographic features (eg stratigraphic unit, type of geological boundary, drillhole name etc)
- Titles, scale bars etc.
- Legends
- Rock relation diagrams and cross sections etc.
- Tectonic and/or palaeogeographic sketches.

On geological maps, spatial relationships are graphically represented in legends, cross sections etc (as interpreted by the map compiler) or are derived through interpretation by the map user (eg distance between two points, rock units surrounding a mine, proximity of water wells to faults etc). In SA_GEOLOGY as in most GIS, some spatial relationships (eg relationship of boundaries to polygons) are maintained by the software system as topology and various functions such as proximity analysis can be performed automatically. However, other spatial relationships are maintained either by specific lookup tables or by other layers in the database.

..power of SA_GEOLOGY is....organizing map information into layers....
geological boundaries, faults, drillholes, mines etc....
and linking attribute data with map features..

The organizing of map information into logical categories or 'layers' is an important concept of GIS and is fundamental to most GIS software. The concept is based on traditional map categories such as topography, gravity, drillhole locations etc. GIS layer structures separate or organize map features into logical dataclasses or layers similar in concept to map overlays. For example, in a GIS database, a geological map can be organized into geological boundaries, faults, stratigraphic units (polygons), streams, drillholes, mines etc (Fig. 2; Parker, 1990). These are physically separated in different files, but can be selectively overlaid to produce a composite or thematic map as required.

Part of the power of SA_GEOLOGY is the linking of attribute data with specific map features. This data integration provides a whole new dimension to accessing, retrieving and analysing information. It allows users to access attribute information in an attribute database simply by pointing to a map feature on a graphics display (eg to display stratigraphic data pertaining to a specific drillhole). Alternatively, it permits users to create special 'thematic' maps based on information in the attribute database (eg to create a map of mineral exploration drillholes with copper > 100 ppm, to selectively display certain rock types, or to progressively "strip off" surficial units).

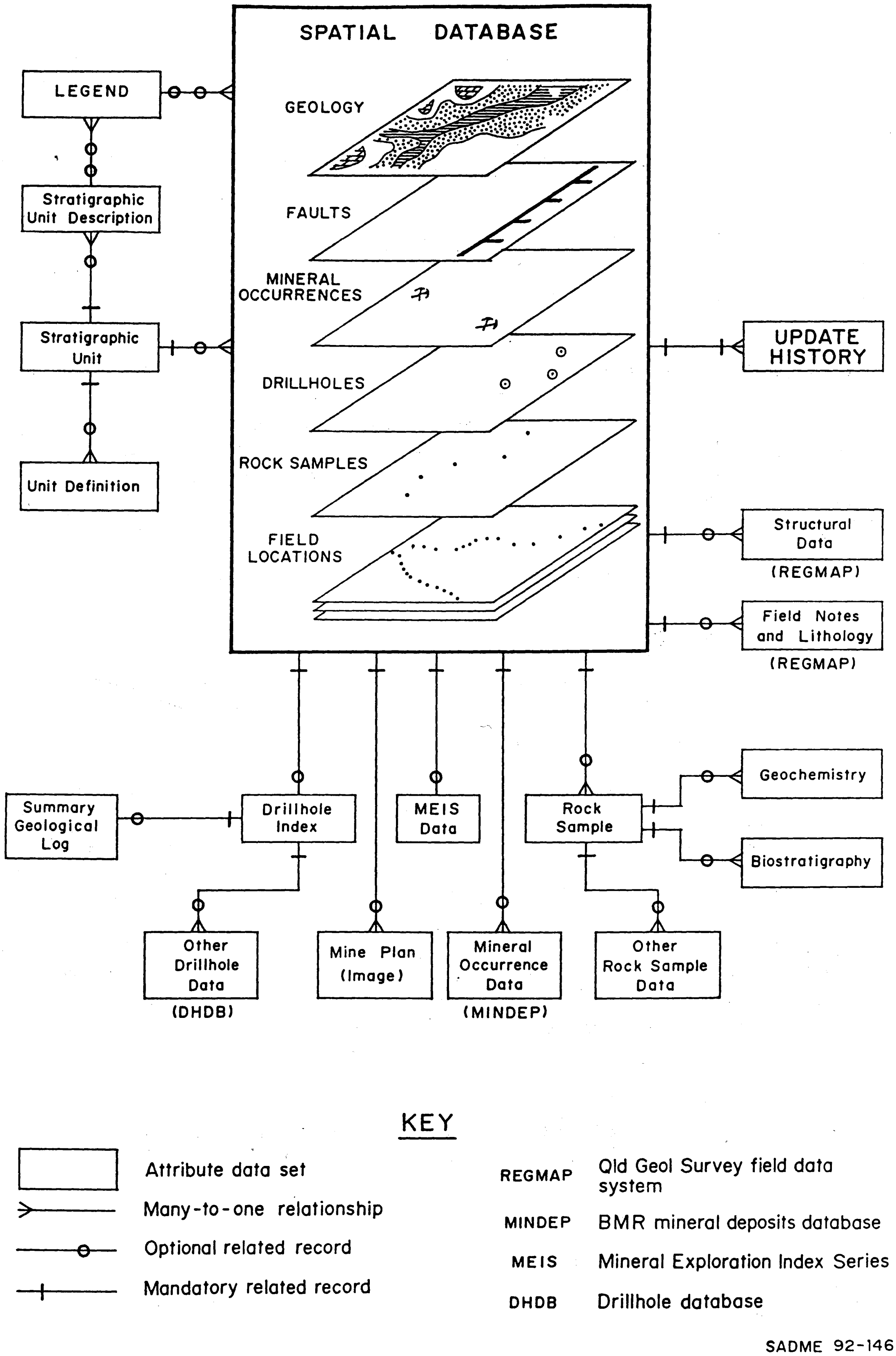


Figure 2: SA_GEOLOGY layer structure and related attribute tables.

Currently, SA_GEOLOGY attribute data are stored in a variety of databases; some, vital to production of geological maps or querying the database, are stored in Arc/Info as related tables but other data reside in SAS datasets. In the long term, most attribute data will be stored in a relational database system (eg ORACLE; Belperio et al., 1991; Belperio & Parker, 1991) with a relational database interface linking them to the spatial data in Arc/Info. SA_GEOLOGY will store all spatial data (including geological boundaries, drillhole locations, mine locations etc) in Arc/Info. Storing spatial locations in a RDBMS (eg ORACLE) is inefficient and requires special retrieval and analysis macros/programs to be written (programs which come with GIS software). Since approximately 75% of all geological database queries are geographically based or require some form of map output, it is more practical and more efficient to establish the GIS graphical user interface as the primary window to the data.

SA_GEOLOGY MAP LIBRARY

Just as a set of encyclopedias is comprised of manageable-sized volumes systematically arranged on shelves or an atlas of maps is sytematically stored in a Vertiplan for easy access and retrieval, it is necessary in large corporate or governmental GIS to partition digital geographic data into manageable-sized parcels, partitions, coverages or tiles which are stored in a map "library" (ESRI, 1990).

Map libraries are used whenever:

- There are large volumes of spatially continuous data
- Data resolution is consistent (eg data digitized from maps of similar scale)
- Data consistency is important across a large corporate database
- Different security levels apply to users and individual datasets or areas
- There are multiple users of the data, each with different functional requirements
- Data is required to be kept either permanently or for long periods of time.

Map libraries not only organize data into manageable partitions (both spatial and thematic) for efficient storage and retrieval, but also control access and data flow (particularly in relation to preventing concurrent updates etc), maintain a transaction history, maintain spatial indexing, and provide management tools for queries, modifying and deleting etc.

SA_GEOLOGY....organised into....libraries....
SA_GEOLOGY/State and SA_GEOLOGY/Geodetail

In SA_GEOLOGY, libraries are organised into a hierarchical structure closely related to the directory structure. At the uppermost level, there is a library definition file which keeps track of all map libraries, then the next directory level below contains the individual map libraries - SA_GEOLOGY/State for 1:1 million scale data and SA_GEOLOGY/Geodetail (- Zone52, Zone53 and Zone54 -) for photo-scale or 1:100 000 data. Partitions or tiles are stored as lower level directories within each library, and then each tile contains further subdirectories containing individual layers (Fig. 3).

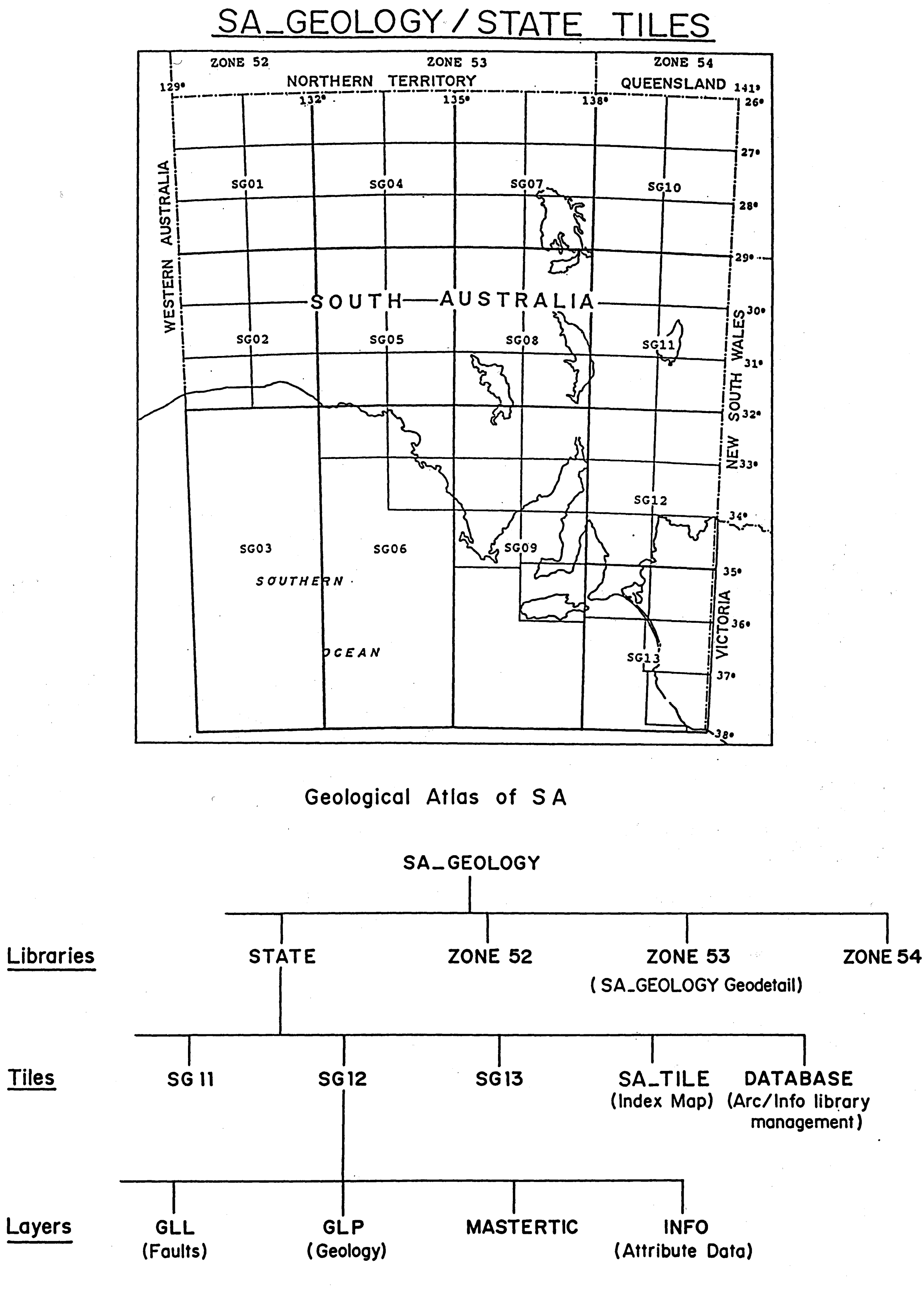


Figure 3: Hierarchical directory and tile structure of SA_GEOLOGY/State map library (based on Arc/Info's Librarian).

Partitions, coverages or tiles are the digital equivalents of map sheets and subdivide the data into component subareas. These subareas may be of irregular shape and size or may correspond to normal map-sheet borders (eg 1:50 000 map sheets). In SA_GEOLOGY/State, tiles have been defined by groups of 1:250 000 map sheets (nominally six to a tile); in SA_GEOLOGY/Geodetail, tiles are nominally defined by 1:100 000 map sheets but there is provision for smaller or larger tiles depending on data density.

Data standards are critical for ensuring data quality and reliability. As indicated above, data in a map library should be consistent with respect to resolution and accuracy, but it is also important to use a single co-ordinate system appropriate to the scale and resolution. For example, in SA_GEOLOGY the Australian Map Grid is used for 1:50 000 and 1:100 000 scale maps and locational data but is inappropriate for state-wide datasets which use decimal degrees of latitude and longitude.

FUTURE GEOLOGICAL MAPS AND RELATED PRODUCTS

The South Australian Geological Survey will continue to produce and publish 1:250 000 scale geological maps and explanatory notes in essentially the same format as it currently does. Minor changes or additions will be made according to the economic potential or nature of the area and this has already begun; for example, aeromagnetic and gravity images were incorporated on the recently-published STREAKY BAY map and several important structure-contour and isopach maps of the Cooper Basin are being incorporated on the STRZELECKI map (in press).

Explanatory notes will still be published to provide a regional overview or summary of the geology of a map sheet. However, with the availability of a comprehensive GIS, there is an opportunity to produce digital datasets of drillhole, rock sample, mineral occurrence and geochemical data on floppy disc to accompany explanatory notes. This has recently been done for KINGOONYA Explanatory Notes and will be continued on a regular basis for other maps/explanatory notes. While the KINGOONYA data package was a one-off compilation, the ultimate aim of SA_GEOLOGY is to be able to produce such packages for any specified area, at any time and containing all the up-to-date digital data available.

While State geological maps will also be published (eg at 1:2 000 000 scale), full-colour computer-generated plots are and will be available at other scales and digital data on magnetic cartridge tapes can also be acquired in Arc/Info export or DXF format.

The same also applies to 1:100 000 scale geological maps; full-colour, publication-quality plots and/or digital datasets are currently available for selected areas and there is an ongoing programme to digitise all existing geological maps of SA at 1:100 000 or photo scale. These datasets will ultimately form the basis of metallogenic and geochemical maps (plots). Selected maps may be published if there is a perceived market (eg a geological map of the Flinders Ranges National Park).

REFERENCES

Belperio, A.P., Molineux, R., Ivic, D., Sambura, A., Sandercock, N., Daly, S.J., Morton, J. & Slade, O., 1991. Strategic framework for the development of a geoscience information system. South Australia. Department of Mines & Energy. Report Book 91/4.

Belperio, A.P. & Parker, A.J., 1991. Towards a corporate geological mapping and geoscience information system for Soutrh Australia. National Conference on the Management of Geoscience Information and Data. Proceedings Volume, Session 3:43-60. (Australian Mineral Foundation, Adelaide).

A J Parker - March 93

Cowley, W.M. & Freeman, P.J., 1991. Digital geological map compilation and production. National Conference on the Management of Geoscience Information and Data. Proceedings Volume, Session 2:15-16. (Australian Mineral Foundation, Adelaide).

Environmental Systems Research Institute, Inc., 1990. Understanding GIS: The ARC/INFO Method. (ESRI, California).

Parker, A.J., 1990. Design and implementation of SA_GEOLOGY: A computerised geological map database system. South Australian Dept of Mines and Energy, Report Book 90/11.

Parker, A.J., 1990. National Geoscience Mapping Accord: Results of 1989 questionnaire. South Australian Dept of Mines and Energy, Report Book 90/20.

Applications and Data Standards for Digital Mapping at the Murray-Darling Basin Commission

Paul Nanninga, Haikai Tané and Peter Golding

Murray-Darling Basin Commission
Canberra

Abstract - Since acquiring GIS facilities in 1990 the Murray-Darling Basin Commission (MDBC) has undertaken a number of mapping projects to support planning and resource management activities. An orthophotographic map base for River Murray floodplain mapping and the 20 year profile pilot study of the Murray-Darling Basin are two examples. In both projects, a lack of uniform, consistent data standards between State governments and their agencies was one of the main problems which had to be addressed.

Keywords - Data Standards/Digital Mapping/Murray-Darling Basin/River Murray

1. Introduction

The role of the Murray-Darling Basin Commission is to:

- Conserve and regulate the waters of the River Murray in accordance with the Murray-Darling Basin Agreement.

- Protect, and where necessary improve, the quality of water resources of the River Murray and its tributaries throughout the Basin.

- Provide advice to the Murray-Darling Basin Ministerial Council to promote and coordinate effective planning and management for the sustainable use of the water, land and environmental resources of the Basin.

To provide policy advice to the Ministerial Council and manage the River Murray, officers of the Commission often need to assemble, at short notice, the best available information on a range of issues. Planning and management must also be based on sound information. Acquisition of Geographical Information System (GIS) facilities and significant amounts of environmental and socio-economic data from State and Commonwealth government agencies are needed to fulfil these roles.

GPO Box 409 Canberra ACT 2601 ◇ Telephone 06 279 0100

A major cause for concern when aggregating data from disparate sources is compatibility of the generalised theme maps and statistical information normally transferred. The role of data standards in relation to these potential causes of incompatibility is discussed in this paper.

The discussion of data standards is followed by examples from two MDBC projects. The Murray-Darling Basin 20 Year Profile project investigated the current state of natural resources compared with scenarios for 20 years in the future, in three study areas of the Basin. The short time scale and limited funding of this project implied making use of as much existing data as possible. This raised many issues but one of the most serious was that of data compatibility. The River Murray Mapping project involved planned acquisition of data across three States using consistent methods. The two projects contrast in their approaches to data acquisition, but both emphasise the need for consistency and standardisation.

2. Data Standards

A Standards Australia subcommittee is currently working towards the introduction of a Spatial Data Transfer Standard (SDTS) to replace the current Australian Standard (AS2482) which does not provide for topologically structured GIS data. SDTS provides a framework to describe the technical format and topological structure of any data set but also requires a commitment to truth in labelling and full documentation of the history and quality of the data set. The main categories for documentation are:

- *Lineage*, includes information about the source material used to generate the data set
- *Positional Accuracy*, defines the degree of compliance with geographic positional controls
- *Attribute Accuracy*, refers to the accuracy of the spatial representation of the data
- *Logical Consistency*, relates to the internal consistency of the data
- *Completeness*, refers to information about the selection criteria used to decide what features are included in the data set.

This initiative aims to alleviate many of the technical problems confronting the transfer of data between agencies but it does not address the problem of compiling together and comparing data sets from disparate sources. This problem is demonstrated by examples from the Murray-Darling Basin 20 Year Profile project.

It is often presumed that the most efficient and cost effective way to generate information about an issue is to make use of existing data which has been collected or generated for other purposes. However, experience has shown that this is only true if the primary data, types of classification used and mapping scales are compatible.

Agreement on uniform classification systems is not a practical solution because there is no such thing as a general purpose classification meeting all user requirements. However, primary attribute data collected in a uniform and consistent manner, including documentation of the methods and transformations applied, can be remapped for other applications if all essential attribute data is available at an appropriate scale.

The Commission, therefore, set up a working group with the objectives of defining a framework for:

- consistent and uniform terminology, attribute definitions, and methods for measuring and recording primary attribute data, and

- ensuring the attributes measured, and the collection and recording methods, are consistent with those accepted by State and Commonwealth agencies.

A fundamental aim of this initiative is to encourage cooperation between agencies and create an environment in which data collected for specific purposes can be more effectively utilised for other purposes in the future. This will not help solve the problem where there is a need to use historical data as in the 20 Year Profile project.

3. The Murray-Darling Basin 20 Year Profile Project

The main objectives, to be completed within a period of 6 months, were:

- assessment of the current condition of the natural resources in the study areas;

- development of 20 year forward projections of conditions;

- assessment of methods used for extension to the whole Basin.

Three study areas were selected: the West Berriquin Irrigation District in New South Wales; the Horsham area in Victoria and the Renmark area in South Australia. The areas were selected because of the ready availability of data.

The success of the project depended upon being able to make comparisons of conditions, trends and spatial variations across three States.

There was a heavy reliance on existing data. Since data originated from three States and several different agencies, data compatibility was a major issue. There is no guarantee that data gathered by different State agencies are measured, classified and mapped in a consistent manner. The following examples from this project illustrate some of the problems of compatibility that were encountered.

Example 1: Groundwater Levels

One of the major problems in irrigation districts such as Berriquin is the rise of saline groundwater tables and the consequent impact on life in and above the soil. Water resource agencies have sought to monitor groundwater through the use of boreholes. Time series data for each borehole then provides the basis for trend analysis and mapping of the extent of shallow watertables now and in the future. This information was used to help fulfil the project objectives.

Primary data about groundwater levels can be specified in several ways: metres above a specified datum (e.g. AHD), metres below the ground surface, metres below the top of the bore casing. It is also common to derive secondary data by classifying groundwater levels such as: within one metre of the surface, between one and two metres deep, and greater than two metres deep.

Conversions are necessary between these forms before valid comparisons can be made. This often requires additional data and the process is made difficult and error-prone if all that is received is data on "depth of groundwater". The lesson to be learned is that good documentation of data is essential when comparisons with other data sets are to be made.

Example 2: Groundwater Salinity

Groundwater salinity is also measured in boreholes by water resource agencies. The Australian Standard of electrical conductivity (and hence salinity) measurement is the deciSiemen per metre (dS/m for short). It is also possible to measure salinity as the amount of total dissolved salts (TDS), e.g. in milligrams per litre (mg/L) or parts per million (ppm).

The salinity map provided for one study area was in TDS units and the other was in EC units. Transformation between the two units requires a conversion factor which varies with the relative quantities of different salts present. This factor can vary from 530 to 900. Therefore, without knowing the exact composition of the sample water, the conversion can be very inaccurate. This kind of information needs to be documented (if available) to accurately convert the salinity units.

Salinity data was provided in the form of contour maps. Since the contour intervals were specified in different units, the salinity classes were completely different, even when converted to the Australian Standard EC units. Conversion into comparable data sets can only be achieved by using the primary salinity measurements.

Example 3: Landform Patterns

Landform patterns are classifications of the landscape based on landform features. Vegetation and soils can be considered attributes of the landform patterns. In this way it is possible to produce soil-landscape maps. These can give information on the distribution of soil textures, e.g. where the sandy soils are located. The soil texture determines how much irrigation water will diffuse through the soil into the groundwater. Such a soil-landscape map was produced for Berriquin by the CSIRO in 1946, the most up-to-date soil data that could be found for the area.

A field survey of 86 sites was conducted in the Berriquin study area by the MDBC to supplement existing data. Soil samples were taken and analysed in a laboratory for the relative fractions of sand, clay and silt. These were classified into texture classes such as sand, clay loam, clay, etc. The data was mapped by constructing Thiessen polygons for each site tagged with the soil texture information. The relative area of each Thiessen polygon gives the statistical contribution of the corresponding field survey site to the whole.

However, when mapped, the Thiessen texture distribution did not compare at all to that derived by CSIRO. One reason for the difference is that the field survey sites were not chosen to lie within, or to represent, landform patterns. Another reason is that the soil texture descriptions from the CSIRO data did not match those of the MDBC field survey.

4. River Murray Mapping

Background

In 1989, at the request of NSW, SA and Victoria, the MDBC established the Tri-State Floodplain Planning Working Group to coordinate planning for the River Murray and its floodplain. Traditionally the main thrust of floodplain planning in the States has been solving problems associated with flooding hazards. However, flood mitigation is only one of several issues which need to be addressed by the MDBC.

The Floodplain Planning Working Group is required to ensure the protection and where necessary, enhancement of the riverine environment consistent with the Natural Resources Management Strategy. Various initiatives are being undertaken to ensure uniform and consistent policies and guidelines for planning, development and management of the River Murray and its floodplain.

Among the terms of reference for the River Murray Floodplain Planning Working Group, a Land Use Planning Spatial Data Base for the River Murray, was identified as a priority task. This data base is being called River Murray Mapping.

The main objectives of River Murray Mapping are to:

- Meet spatial information requirements for floodplain planning, including flood mapping, resource inventories and assessments, plus tenure and zoning information.
- Revise and upgrade existing resource information to consistent standards across the River Murray, by developing uniform base mapping, orthophotography and remote sensing imagery
- Develop prototypes of the River Murray Mapping system to demonstrate the benefits of floodplain mapping and modelling in coordinating planning for the River Murray between the Contracting States.

Spatial Information Standards

As a result of tri-State deliberations of user requirements, the following scale standards were specified for River Murray Mapping:

Function	**Urban**	**Rural**	**Regional**
Data Capture	1: 5,000	1:25,000	1:50,000
Work Sheet	1:10,000	1:50,000	1:100,000
Publication	1:25,000	1:100,000	1:250,000

A systematic survey of existing topographic mapping was undertaken to derive a suitable geographic base meeting these mapping requirements. The largest uniform scale mapping for the whole River Murray is the national 1:100,000 series. While there are parts of the Murray floodplain mapped at larger scales, (ranging through 1:10,000, 1:25,000 and 1:50,000), they were found to be inconsistent both along and across the River Murray.

The mistake of enlarging a smaller scale topographic map base for larger scale mapping is not uncommon among people and agencies with inadequate geographic or cartographic training. This type of mistake was highlighted by the problems generated in mapping riparian vegetation for the River Murray at 1:50,000 on enlarged 1:100,000 map bases. In developing topographic models for a prescribed scale, a modelling strategy is defined which involves significant generalisation and aggregation, specific to the scale of the finished product. The standards of accuracy for locating spatial data for a 1:100,000 product, for example, usually allow variations from true location of the order of plus or minus 50 metres. The corresponding standard for 1:50,000 mapping is plus or minus 25 metres.

When scale anomalies in the riparian mapping project were multiplied by inaccuracies inherent in the use of outdated base maps and aerial photographs for data capture, the result was to diminish the scale standards for the riparian vegetation mapping from the published scale of 1:50,000 to between 1:250,000 - 500,000: that is the information displayed is only correct to within plus or minus 150 - 250 metres. This is unacceptable for floodplain planning and natural resource management purposes, where decisions are made on the basis of the reliable location of riparian vegetation in relation to property boundaries. Scale standards of 1:25,000 worksheets are required, though final publication at 1:50,000 is acceptable for display purposes.

Consequently, the assessment concluded that the Riparian Vegetation maps were compiled on unreliable and inadequate base maps. Furthermore it was concluded that the existing topographic mapping sets for the River Murray do not provide a uniform, consistent and reliable geographic base for mapping and managing the River Murray.

Geographic Bases

A new geographic base meeting the requirements of River Murray planning and management was identified as essential. Orthophoto maps were recommended as the cheapest and quickest way of providing the base map. A number of planning projects undertaken by the Commission had demonstrated the value of orthophoto maps in River Murray management. Consequently, a decision was made to generate an orthophotographic base for the River Murray to 1:25,000 topographic map standards. Flying commenced in May 1991 and concluded in June 1991. Processing the draft orthophoto maps (over 170 A2 size maps) took another six months. Final processing and formatting is now underway with a release date for the River Murray orthophoto maps anticipated to be less than a year from the commencement of flying.

Thematic Mapping

In 1991 the Commission granted funds under the Natural Resources Management Strategy to prepare consistent, uniform and reliable flood mapping for the River Murray. The River Murray Floodplain Planning Working Group emphasised that reliable floodplain mapping is an essential prerequisite to achieving coordinated floodplain planning between the Contracting States. This task is now underway with the Rural Water Commission of Victoria managing the project and working closely with NSW and SA agencies.

The next step is to rescue the riparian vegetation and other resource mapping information, now stored on unreliable base maps, by transferring these data to the new orthophoto map base. This task is being undertaken on contract, with the data being captured digitally afresh, after various trials demonstrated that rubber sheeting techniques were more time consuming and less reliable.

5. Conclusions

The technical capability to transfer large amounts of data between agencies and systems is already possible and should be facilitated with the introduction of SDTS. However, for data to be effectively used for other purposes it will be necessary for all agencies to work towards the:

- adoption of cartographic conventions and standards;
- adoption of uniform, consistent terminology and attribute definitions for primary data;
- standard methods for the collection and recording of primary attribute data, including units of measurement;
- full documentation provided with digital data sets of the scale, accuracy and precision;
- full description of the assumptions, generalisations, limitations and methods used to derive data sets from primary attribute data, and
- full documentation of the coordinate transformations applied.

Digital mapping is opening opportunities for a wider range of people to undertake mapping. However, the technology is open to abuse and misuse when cartographic conventions and standards are ignored. Establishing scale and geographic base map standards at the design stage made it possible to proceed with uniform development of the spatial data base for the River Murray.

Experience gained from the Murray-Darling Basin 20 Year Profile project emphasised difficulties in using existing data sets. Funds available for projects, the scale of operations, and the variety of purposes for which data are collected generally results in data differences. Our experience has demonstrated that while standards can be applied to primary data collection for particular purposes, there is no such thing as general data standards for mapping and modelling spatial systems.

ooo0ooo

HYDROGEOLOGICAL MAPPING AND MODELLING USING A GIS

Ross Brodie and Andrew Tucker

Bureau of Mineral Resources, Geology and Geophysics
Canberra, ACT 2601

Abstract - The compilation of datasets for hydrogeological maps and models is based on the available borehole record. In a GIS environment, borehole data is integrated with traditional geological datasets and information on topography, meteorology, vegetation, land use and surface drainage. The typical outcomes are continuous data surfaces of hydrogeological parameters such as yield, salinity and aquifer geometry. These data surfaces are represented as contours or range-coded colour fills on a map or reformatted as averaged cell values in a groundwater model. Data standards and documentation are fundamental in optimising this process of data storage, integration, interpretation and presentation.

1.0 Introduction:

The BMR Groundwater work program is focused on understanding the hydrodynamics and hydrochemistry of regional groundwater systems and assessing their resource potential, principally for the Murray-Darling Basin and the Great Artesian Basin. Output is distributed as hydrogeological maps and numerical models - the former being the traditional hardcopy depiction of multi-layer aquifer systems while the latter are used to predict the groundwater response to anthropogenic change.

Borehole data is the main source of direct observations of groundwater systems. This data is supplied from the agencies responsible for water resource management within each state, and is augmented with data from BMR's corporate databases and field programs. The compiled point data may be sparse, both temporally and spatially.

The production of groundwater maps and models requires that data surfaces such as head, yield or salinity trends be interpolated from point data. Generally, these surfaces are interpreted by integrating the borehole data with surface geology/geomorphology and geophysical data as well as topographic, meteorologic, vegetation, land use and surface hydrologic data (Table 1).

Our work in integrating, compiling and interpreting hydrogeologic datasets has highlighted the need for uniform documentation and data standards.

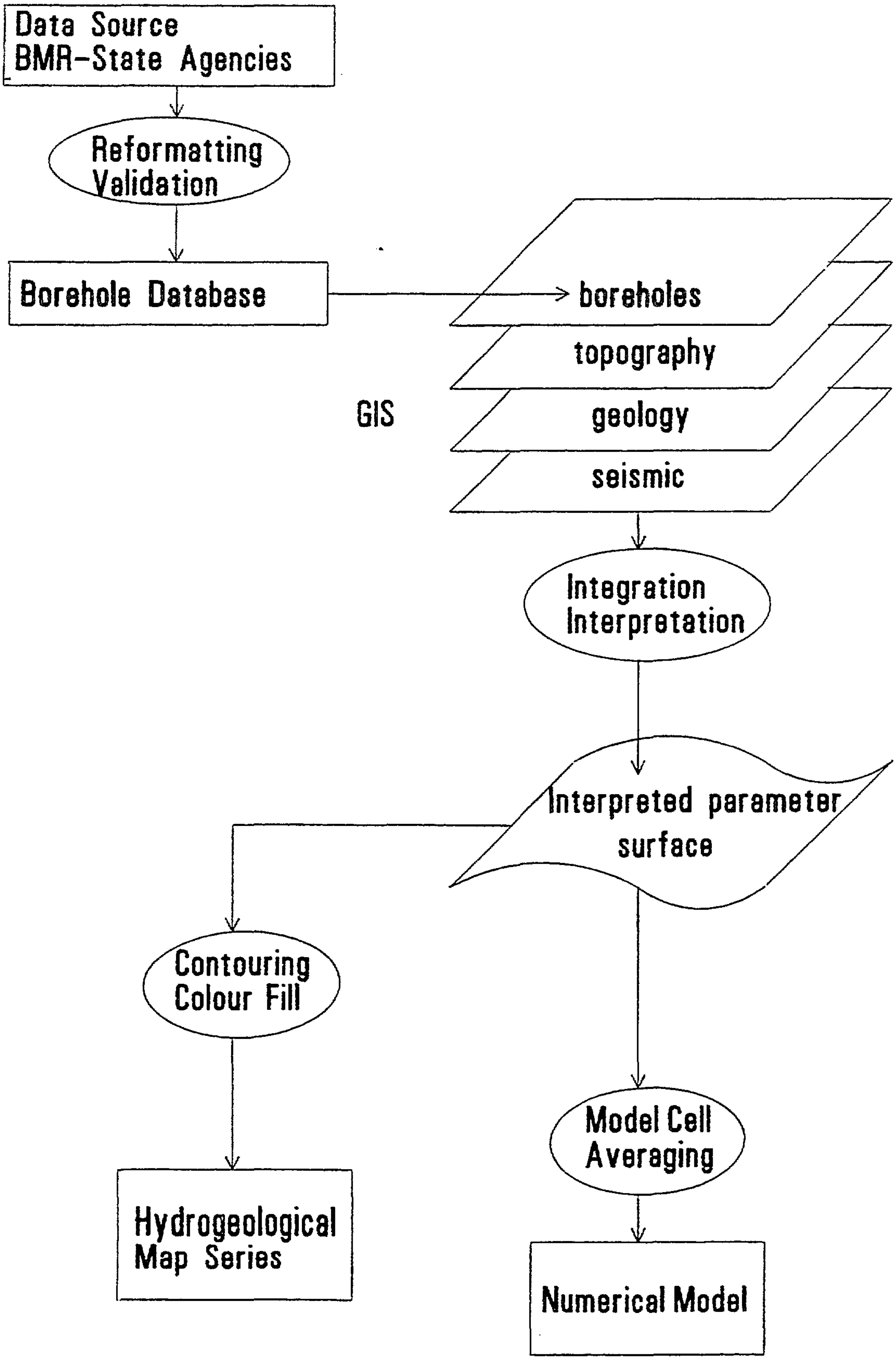


Figure 1: Compilation of parameters for Groundwater Maps and Models

Table 1: Available data used in the Compilation of BMR Groundwater Maps and Models

Data	Format	Source	Application	Comments
Borehole	Oracle database	State water agencies PEDIN	Principal source of hydrogeological data	Logs of registered water bores, petroleum exploration wells and groundwater reconnaissance drilling, variable reliability.
Seismic	Hardcopy	BMR archives State water agencies	Basement features Aquifer geometry Water table	Source of original data varies, typically petroleum companies or government agencies. Highly interpretive with accuracy dependant on technology used.
Aeromagnetics	Digital	BMR	Basement features	1.5 km line spacing, 17m sampling interval. Computer generated Total Magnetic Intensity contours at 2 nT interval.
Gravity	Digital	BMR	Basement features	Accuracy ± 5μms^{-2} at stations. Computer generated Bouguer anomaly contours at 50μms^{-2} intervals derived from 1.5 minute grided values.
Radiometrics	Digital	BMR	Recharge/Discharge Soil Types	3 km line spacing, 60m sampling interval. Total Count contours at 10 counts/s interval.
Surface Geology	Hardcopy	BMR State Geological Surveys	Groundwater features Stratigraphy	Geological interpretation based on airphoto interpretation, mapping traverses and available drillhole data. Presented as 1:250 000 compilations.
Topography	Hardcopy	AUSLIG CMA NSW	DTM	Currently digitised into .dgn files for Murray Basin from 1:50 000 or 1:100 000 scale maps. 20m contour interval conforming to National Mapping Standards, point elevations to nearest metre on AHD.
Drainage	Hardcopy	AUSLIG CMA NSW	Surface-groundwater interaction	Digitised as above.
Culture	Hardcopy	AUSLIG CMA NSW	Land use Location	Digitised as above.
Rainfall Evaporation	Digital	Bureau Meteorology	Recharge	Historical data from rainfall stations
Surface Hydrology	Hardcopy	State water agencies	Surface water-groundwater interaction	Routine monitoring of major rivers - discharge ratings, weir heights, flood plans, gauge records.
Landsat TM AVHRR	Digital	BMR	Vegetation-land use Recharge-discharge	Landsat coverage over Murray Basin available inhouse

2.0 Data Sources:

2.1 Borehole Data:

A large amount of historical borehole data has been made available to us by state water authorities. As water rights are vested in the states, the data was gathered and stored to meet the regulatory requirements of these authorities. In some areas, the borehole record may extend from the 1880's to the present day.

The sedimentary basins for which BMR Groundwater are undertaking regional synthesis, are crossed by a number of state borders. Borehole data must therefore be integrated from several state-managed databases, as well as BMR borehole databases. For example, construction of the GABMOD borehole database designed to support the modelling effort within the Great Artesian Basin involved reformatting eight pre-existing databases. As well, some borehole data is accessible only from hardcopies or microfiche.

This data integration process was complex and several data management issues were highlighted:

1. The source databases have been designed to service different objectives to the requirements of the BMR Groundwater work program as state databases perform a regulatory role as well as a store of scientific data. This may result in the incomplete recording of field data.

2. Different approaches have been taken in terms of table structure, codes and abbreviations, data representations of borehole coordinates and dates, units of measurement and the handling of null values. This has necessitated the development of conversion procedures unique to each of the source databases. Various conventions have also been used in database documentation.

3. Stratigraphic nomenclature for geological units may not be consistent between the databases. Continent-wide registers of stratigraphic names, such as GEODX, are useful in providing such consistency.

4. The relative balance between storing measured field data and any subsequent interpretation is different for each database. For example, recording a descriptive lithological log as opposed to a stratigraphic interpretation or pump test measurements instead of derived storage and transmissivity values. Ideally, source data and methodology should be referenced to any data interpretations.

Invalid data may enter a database through incorrect measurement methodologies, data entry errors, corruption during data transfer, or incorrect database manipulation commands. We may also be the first database users to access and use the data, and so we must attempt to validate the data during the data integration phase. Money and time constraints preclude us from validating the data by field measurements. Also, field data such as historical head measurements simply cannot be remeasured.

Typically, data validation during database compilation consists of:

1. Exploiting data redundancy. For example, if a mapsheet is listed along with the borehole co-ordinates, the co-ordinates can be compared with the maximum and minimum co-ordinates of the map sheet.

2. Ensuring that data values are physically possible. For example, dates must be before the current date, hydraulic conductivity values and ion concentrations must be positive.

3. Ensuring that data values fall within estimated maximum and minimum values. For example, bore diameters must be within realistic constraints. Calculating general statistics gives an indication of the range of values.

4. Ensuring that related data items are logically consistent. For example, the top of a unit must be above the bottom, and the date a bore was constructed will generally precede the date of any hydrogeological measurements.

5. Checking referential integrity between tables, with foreign keys corresponding to primary keys and data fields validated by look up tables.

Our approach has been to take a "snapshot" of each of the data bases - unless data updates are significant or vital to the project outcome, the work involved in integrating and validating the databases is too great to allow us to continually update the compiled databases. The data compilation and validation tasks would have been made considerably easier if consistent standards were used in the original datasets.

2.2 Support Data:

Data additional to the borehole records is used to assist in the interpolation of hydrogeological parameter surfaces (Figure 1). GIS coverages of topographic contours, spot heights, drainage, surface geology, vegetation, seismic lines and cultural data (roads, towns, homesteads) have been constructed.

To date, the principal source of digital data has been the corporate Intergraph CAD facility. As an example the compilation of the Murray Basin 1:250 000 Hydrogeological Map Series has required the routine digitising of base map detail such as topography, drainage, culture and geology over the entire Murray Basin. Some of this dataset has subsequently been imported and converted to Arc/Info coverages to support mapsheet compilation.

This conversion process has involved a new perspective on the Intergraph design file, different from the traditional cartographic view. In streamlining Intergraph design file to Arc/Info conversions, the importance of a number of cartographic procedures has been highlighted:

1. Documentation of original data used including the bibliographic reference, data source, publication date, scale and accuracy, map projection parameters and type & condition of medium.

2. Documentation of the digitising process and tolerances. The proper setting of snap locks ensures polygon closures, considerably reducing the time required in cleaning the subsequent Arc/Info coverage. Details of any subsequent processes such as generalisation, reprojection or rescaling should also be recorded.

3. Documentation and standardisation of design file levels within the mapping project (Cooper, 1990). The routine separating of thematic data into specified design levels allows the orderly and complete transferral into thematic GIS coverages.

4. The coordinates of data digitised within the design plane should match exactly the coordinates as defined by the map coordinate system and projection. This omits the need to transform the data from an arbitrary design plane to the real world. At the very least, four registration points, such as at map corners, should be provided to allow such a transformation.

Historical data , derived from both external sources and BMR archives, is typically in hardcopy form. This has resulted in a relatively expensive, time consuming and sometimes inaccurate exercise in converting to digital form. Valuable information such as map projection, grid coordinates, data source and accuracy may not be depicted.

Interpreted data surfaces have historically been depicted as hardcopy contour plots. The factual point data, or information on how the contours were interpolated is commonly omitted. Converting contour lines to a digital surface requires time and operator skill.

3.0 Data Interpretation:

The main goal of data interpretation is to produce a parameter surface from point data and subsidiary surfaces (Figure 1). As these surfaces are the primary store of our interpretation, details on its origin and status are critical. Even if a surface is available as a grid or TIN, we need to know more about the surface than just the data values at grid points or triangle vertices.

We need to know from which data set the surface was generated. If the surface was generated manually, then it can be important to know the interpretation rationale. If the surface was generated algorithmically, then it can be important to know what algorithm was used, and what parameter values were used to control the algorithm. A history of any subsequent data processing (eg. editing, filtering, smoothing) should be supplied. It can also be very important to have an estimate of the accuracy and reliability of the surface.

Appropriate use of map projections should be made, especially for small scale maps. For example, if areal analysis will be performed on the data, an equal area projection should be used (Snyder, Maling). There will always be some loss of accuracy if an ad hoc "projection" is used for a surface, and rubbersheeting is required to project the surface to a conventional map projection. Projection details should always be supplied with surfaces.

It would be advantageous if all the information related to the surface could be integrated and permanently linked with the surface. This data package should be convertible to a format independent of hardware and software environments.

4.0 Products:

4.1 Mapping:

Hydrogeological maps are the cartographic representation of the interpreted data surfaces (Figure 1). For example the POONCARIE sheet of the Murray Basin 1:250 000 map series depicts, mostly as contours, some twenty one data surfaces. Maps are a useful and succinct medium for displaying hydrogeological data, but require a large amount of time and resources to prepare. They are also only a snapshot of the level of understanding of a groundwater system at the time of publication. Parameters such as depth to watertable may actually vary over time, while the hydrogeological interpretation may change with the advent of new data.

We must be able to supply data to our clients in a digital form, in a format that is supported by or transferable to their hardware/software environment. This digital format also allows maps to be interactive, allowing updates to be easily made.

4.2 Modelling:

The data surfaces also fulfil another role as input to and output from numerical groundwater models (Figure 1). The issues of proper data management and documentation continue to prevail. Data standards are part of the wider issues of quality assurance in groundwater modelling, which has been defined as :

> "a procedural and operational framework put in place by to assure technically and scientifically adequate execution of all project tasks included in the study, and to assure that all modelling-based analysis is verifiable and defensible." (Taylor, 1985)

Therefore, the model parameter sets, along with the data from which they were derived must be:

1. documented and accessible to both modellers and reviewers.

2. portable to different hardware and software environments.

3. easily updatable as new information becomes available.

4. recreatable at some point in the future - model predictions are often required only when new resource management strategies are being developed to meet changing resource utilisation patterns.

5.0 Conclusion:

Wherever possible, all source data for an integrated data set should be validated. Data validation should continue during all phases of data set generation.

Appropriate use of data standards during data compilation and integration increases the reliability of the dataset, and eases the updating of the data set as new information becomes available.

The use of data standards during modelling and mapping aids in the review of these products, and play a vital part in maintaining quality assurance.

It is appropriate and beneficial to use data standards during integration, interpretation, presentation and data archiving phases of hydrogeological dataset generation and distribution.

References:

Cooper, R., 1990-- 1:250 000 Series Murray Basin Mapping Project Intergraph Design File Specifications. *BMR Groundwater (unpubl).*

Maling, D.H., 1976-- Coordinate Systems and Map Projections, *George Philip and Sons*, pp 255.

Snyder, J.P., 1987-- Map Projections - A Working Manual, *U.S. Geological Survey Professional Paper 1395.*

Taylor, J.K., 1985-- What is Quality Assurance? In: J.K. Taylor and T.W. Stanley (eds.), *Quality Assurance for Environmental Measurements*, pp.5-11. ASTM Special Technical Publication 867, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania.

ADDING LAYERS TOWARDS A GIS DATA PACKAGE.

Mark Rattenbury, Michael Craig, Brian Oversby, Alan Whitaker

BMR Geology and Geophysics,
P.O. Box 378,
Canberra, ACT 2601.

Abstract - The Eastern Goldfields Mapping Accord Project is establishing a comprehensive digital GIS geological dataset package consisting of a number of geological and geophysical datasets, including geological map units, regolith map units, faults, folds, structural measurements, airborne magnetics/radiometrics interpretation and remote sensing interpretation. Other available datasets which can be integrated with the Estern Goldfields Project geological data include gravity, cultural, hydrological, topographical, airborne magnetics and radiometrics imagery and remote sensing imagery. Relational database-stored sites, samples, geochemical, lithological, mineral deposits, geochronological and mineralogical data are also available for the project area. The Eastern Goldfields Project aims to prove the value of a GIS in the effective storage, manipulation and analysis of spatially linked but otherwise disparate data. Problems of scale and map projection can be easily overcome, but data sampling densities effectively limits the range of operational scales for the GIS, between about 1:25,000 and 1:500,000. Most data layers, apart from proprietary data, are available to the mineral exploration industry to assist in their project formulation and regional prospectivity evaluations. In addition, the Eastern Goldfields Project is using the digital data through the GIS medium to aid understanding of structural/tectonic, geochemical, regolith and other geological processes in the Eastern Goldfields Province.

Keywords - GIS/geology/geophysics/imagery/Eastern Goldfields/Archaean

Introduction

The National Geoscience Mapping Accord is centred around data collection from a number of key mineral provinces in Australia. Data management has been recognised from the outset as a critical facet of each mapping accord project, given the range and sheer quantity of data becoming available. Consequently new techniques of geological data storage and handling are being developed within the BMR which have so far focussed around relational database storage linked to a Geographic Information System. The Eastern Goldfields Mapping Accord Project in Western Australia is targetting the rich gold and nickel-bearing Archaean rocks of the Yilgarn craton, with the BMR concentrating on a transect across the northern sector of the Eastern Goldfields Province, and the Geological Survey of Western Australia currently concentrating on the southern sector. The BMR component of the Eastern Goldfields Project is producing and/or utilising a number of datasets, which include geological mapping, regolith-terrain mapping, aeromagnetic, radiometrics and gravity data, satellite-based remote sensing and airborne multispectral scanning, digital terrain models, mineral deposits, rock geochemistry, and geochronology. The project aims to focus the mineral exploration of the region directly through marketing of the products to the exploration industry, and indirectly through development of integrated geological and geophysical models to further the understanding of the evolution and metallogenesis of the Eastern Goldfields Province.

The datasets vary in their degree of influence and dependence from other datasets. For instance, straight measurement datasets such as aeromagnetics, gravity, and remote sensing imagery are neither influenced nor are dependent on other datasets. Observational datasets such as geological and regolith-terrain mapping are neccessarily more interpretative and can draw on other datasets particularly where initial mapping or interpretation is equivocal. For the most part, however, these datasets remain distinct, and optimum analytical power is achieved when these datasets are integrated with intelligent deliberation.

Data availability at this relatively early stage in the project life, varies considerably for any given area within the project limits. For instance, Landsat TM satellite imagery exists throughout, although mostly in an unrectified or unprocessed state. Airborne magnetics and radiometrics because of contractual obligations and financial restrictions covers the Edjudina 1:250,000 sheet area with impending plans to obtain at least two more sheets of coverage. Reprocessed wide line spacing BMR aeromagnetic data is available throughout. Gravity data has been collected for the Edjudina 1:250,000 sheet area. Geological mapping coverage is ongoing with five 1:100,000 sheets essentially completed, and regolith-terrain mapping is also ongoing with two 1:250,000 near completion. Additions to the corporate ORACLE-hosted BMR databases such as LOCALITIES/SITES, RTMAP, STRUCTURES, LITHOLOGY, PETROGRAPHY, ROCKCHEM, SAMPLES, OZCHRON, MINLOC AND MINDEP are continuing.

Geological Mapping

At the time of publication, five 1:100,000 scale geological maps have been or are nearly completed from the BMR component of the Eastern Goldfields Project. All these maps have been captured digitally, that is, field compilation maps produced from aerial photograph overlays have been digitised. The digital data were then topologically structured and assigned attributes where required using the ARC/INFO GIS. The initially piecemeal and *ad hoc* approach to data file design has by neccessity been standardised (see Gallagher, this volume). The geological mapping data have been subdivided into a number of layers, which include geological boundaries, faults, folds, lineaments, bed horizons, veins and dykes, and structural measurements. Presentation of the geological mapping data as a hard-copy geological map is an essential product of the GIS, but because the hard-copy geological map has a predetermined scale, data may have to be culled for legibility and clarity appropriate to that scale. Therefore the data represented in geological mapping layers intended for hard-copy map output is often a subset of the available data.

The geological boundaries layer is the fundamental component of standard geological map. This layer areally defines the extent of map units (rock types and/or associations) across the mapping area. Discrete map units occur within enclosed polygons each of which has an internal label point which is assigned a distinguishing attribute pertaining to that polygon. In addition, the type of the geological boundary on the polygon perimeter can be assigned attributes on one level such as definite, inferred, or concealed, and on another level whether depositional, intrusive or faulted.

Most of the other geological layers are discontinuous linear features or planar features with a linear surface expression; faults, folds, lineaments, bed horizons, veins and dykes. For GIS and cartographic purposes these linear features have no assigned width, which is an approximation based on the scale of intended output. For instance, a vein does have a finite thickness but at standard map scales most veins could not be drawn as discrete polygons. Since the map output of the Eastern Goldfields Project is primarily 1:100,000 scale, any geological unit or structure less than about 100 metres wide is better represented as a linear feature, except for special purpose detailed presentations.

Integral to the interpretation of geological map data is the legend. The legend exists in two forms; the more conventional and familiar box layout with accompanying text suitable for hard-copy presentation, and as an expansion table relating the legend text and map unit description via the rock code to the various geological polygons.

Sites

Locational data reside in the corporate BMR ORACLE-hosted database SITES, which consist of Site ID numbers, collection dates, originator numbers and names, method of determining location (aerial photo or GPS), positional accuracy, latitude/longitude and AMG easting/northing, and a number of other fields (see Blewett & Ryburn, this volume). Site ID numbers provide the fundamental relating item to other corporate BMR ORACLE-hosted databases. The neccessary hardware links between the GIS and the ORACLE-hosted databases have not yet been installed in the BMR but the transfer of ORACLE-hosted data into GIS is routine elsewhere. SITES data will ultimately be extracted and displayed within the GIS, for example, to show the distribution of observational sites upon which the extent of the geological map units has been based.

Structural Measurements

Structural measurements are numerical representations of rock structure orientations, stored in a database and linked through locality or site numbers to a spatial position in a GIS. Types of structural measurements include bedding, cleavage, foliation, igneous layering, axial surfaces, minor faults, minor veins, joints, minor folds, and lineations. The Eastern Goldfields Project typically collects far more structural measurements than can be presented on any one 1:100,000 scale map. The structural measurements layer used for hard-copy output is neccessarily culled, which is best achieved by prioritising each measurement at the database level so that only measurements of a specified priority will be plotted. The initial data, however, can be plotted at any scale, for example, more structural data may be useful on a mineral exploration tenement scale. A recently proposed argument that gold mineralisation occurred during early subhorizontal deformation (rather than late strike-slip) could be pursued by analysing mineral deposits within regions where rock foliation is inclined less than 30' to horizontal.

Lithology/Petrography

Site-specific lithological and petrographical information is currently being built up on the corporate ORACLE-hosted databases LITHOLOGY and PETROGRAPHY (see Blewett & Ryburn, this volume). This ability to tap into relational databases and isolate specific lithologies, minerals, textures etc spatially into a map context adds a further dimension to the functionality of a GIS. For example, displaying mapped occurrences of the ultramafic cumulate and komatiitic rocks over a total magnetic intensity image could be used to aid future targetting of prospective nickel deposits.

Mineral Deposits

The location and description of mineral deposits in the Eastern Goldfields is an essential component of any exploration-oriented data package. Data are being compiled from two sources; through digitising of known deposits from existing geological maps into the corporate ORACLE-hosted MINLOC/MINDEP database, and by field location and description by the mapping geologist. Mineral deposits are considered as points for GIS and cartographic purposes, despite their often large areal (and volumetric) extent. The MINLOC/MINDEP data have many useful attributes such as commodity, tonnage, grade, alteration type etc which are ideal for spatial integration in a GIS, for example, displaying occurrences of gold grades greater than 5 g/t occurring in sercite altered rocks within 200 metres of a mapped fault.

Airborne Magnetics and Radiometrics

Aeromagnetic data for the Eastern Goldfields Project is aimed at two basic levels of detail; broadly spaced regional 1500 metre line spacing, and semi-detailed 400 metre line spacing data sets. The regional data set covers twelve 1:250,000 map sheet areas acquired by the BMR. Four hundred metre line spacing aeromagnetic and U, Th, K radiometric data will be acquired for up to six of these sheets, however commercial interest in data and data products of the semi-detailed data sets will be retained by the contractor for three to five years from the date of purchase. The incorporation of derived interpretations in the Eastern Goldfields Project data package is unrestricted, however, and the GIS package will eventually include selected images. The 400 metre line spacing airborne geophysics yield satisfactory images with a pixel resolution of 100 metres. These images support interpretation and compilation at 1:100,000 scale for map production at 1:250,000 scale. The aeromagnetics data provide constraints on a model of the basement geology unimpeded by surficial cover, although in areas of good outcrop geological mapping offers greater resolution.

Gravity

The project area is covered by the Australia wide, 11km gravity grid. Up to four 1:250,000 map sheets in the Eastern Goldfields Province will be infilled with the station spacing reduced to approximately four kilometres (4km). These data sets provide a complementary constraints to the subsurface geology model to those derived from the aeromagnetic data. An interpretation of the gravity, with either the contoured or imaged data will eventually be included as layers in the GIS package.

Regolith Landform Mapping

Regolith-landform mapping is part of a nationwide programme to determine the distribution, properties and nature of the regolith. Regolith-landform mapping began in the Yilgarn with BMR mapping regolith at a scale of 1:1,000,000. BMR has now undertaken a programme of reconnaissance regolith mapping in the Yilgarn with the aim of capturing data at 1:100,000 scale with the view to presenting a final maps at 1:250,000 scale. Utilising satellite imagery, radiometric data, computerised data manipulation and storage as well as map production via a GIS sets aside these new approaches from earlier mapping practices. To maximise the time and returns whilst field work is in progress, the second generation mapping approach requires that Landsat TM and radiometric imagery as well as the more traditional techniques such as photogeology interpretation be used before field works begins. Field data are entered into formatted note books which replicate screen entry menus for the corporate regolith ORACLE-hosted database RTMAP. RTMAP stores comprehensive attribute data pertaining to regolith and landform development, for example, regolith site and mapping unit descriptions. Spatial associations from RTMAP data can be integrated and analysed using a GIS. Digital GIS layers available include a polygon layer of regolith landform units, and a layer of geomorphic features. Now, for example we can point to the Venn type intersection between known mineral occurrences and their host regolith types. The union of regolith polygons and the hard rock polygon layers can lead to a better understanding of the extent of rock types beyond just their outcrop boundaries.

Satellite-based Remote Sensing

Landsat TM (thematic mapper) imagery using both processed and simple colour composites are now used successfully to assist with the identification of rock types, materials and landforms both in pre-field and field assessments during geological and regolith mapping. Satellite imagery is relatively cheap and complete coverage exists over the entire Eastern

Goldfields project area. These data can require considerable computer memory storage, and a GIS "library" of remote sensing imagery can be difficult to manipulate on existing hardware systems. Landsat TM images convey discrete visible and near-infrared spectral wavelengths reflected to varying degrees by different minerals and vegetation types. For example, the spectral response of some clay minerals indicative of alteration and mineralisation can be spatially integrated and analysed within a GIS with other layers such as known mineral deposit locations or regolith types.

Aerial Photography

Aerial photography can scanned and warped within the GIS environment to provide useful imagery for locational and photogeological purposes. The Eastern Goldfields Project has no plans for including aerial photography within the data package but recognises the value of aerial photography in an exploration company's GIS.

Geochemistry

Within the Eastern Goldfields Province there exists a large number of whole-rock geochemical analyses which reside within the corporate ORACLE-hosted ROCKCHEM database. This dataset, given the neccessary hardware links, can be utilised spatially through a GIS, for example, displaying basaltic rocks (from the geological boundaries layer) where Cr concentrations exceed 1000 ppm. Where detailed geochemical data are available, gridded images could be generated for the GIS, although the Eastern Goldfields Project is not producing any datasets suitable for incorporation the data package.

Geochronology

The growing component of Eastern Goldfields Province geochronology in the corporate ORACLE-hosted database OZCHRON can be tapped spatially, for example, displaying mineral deposits occurring within 200 metres of 2650 to 2655 Ma granites.

Topography and Digital Terrain Models

Topographic contours and spot heights are increasingly available in digital form from the Australian Surveying and Land Information Group (AUSLIG), and the Department of Lands Administration (DOLA) for the Eastern Goldfields Project area. These data may be available only as a scanned image initially, which is useful for locational purposes. These data, however, remain the custodial property of their respective source and are not released as part of the data package. With existing vector topographic contours or spot heights of suitable density, digital terrain models can be constructed. Pseudo three-dimensional display programs within GIS can drape various layers such as regolith or geological map units, faults etc over the model topography to examine the effects of geology on landform, or landform on geology.

Map Grid and Graticule

Spatial reference to a particular map projection is essential in any map construction, and most GIS incorporate facilities to convert data from one projection to another. The Eastern Goldfields Project is using the Australian Map Grid (AMG) coordinate reference system (Universal Transverse Mercator projection) which is appropriate for the 1:100,000 map scale. Should a different projection be needed, to display data at a different scale or projection such as 1:5,000,000, then automatic conversion can be done relatively easily. Some GPS units record spatial position in latitude and longitude which need to be converted to a map projection for plotting, and many older maps containing valuable data in a different map projection can be readily transformed. The Eastern Goldfields Project has automatic routines for generating the graticule and AMG grid for any given 1:100,000 sheet area.

Cultural Features

Roads, wells, buildings, towns and other cultural features are essential pieces of information for the user to locate or spatially reference geological information on screen or on map. The Eastern Goldfields Project uses digital data (where available) supplied by AUSLIG or from DOLA, but do not release the layers as part of the data package.

Hydrological Features

Streams, lakes, and other hydrological features are increasingly available as digital data supplied by AUSLIG and DOLA. These data are particularly important for recent geological processes which affect alluvial mineral deposits, stream sediment geochemistry, and regolith-terrain mapping, but being proprietry data are not available in the Eastern Goldfields data package.

Other Potential Datasets

Any spatially referenced data can be incorporated and integrated within a GIS. The primary goal or objective, however, must not be lost in the process of incorporating possibly irrelevant datasets. Agricultural, forestry, census, and electorate spatial databases, amongst others, exist in many if not all parts of Australia and could be incorporated with the geological data in a GIS. Other, possibly more relevant datasets to the exploration industry, are digital land tenure and cadastral maps, and digital mineral exploration tenement maps which are increasingly available through State Mines Departments and Geological Surveys. The FINDAR system maintained by the National Resource Information Centre (NRIC) is a spatially oriented directory of the digital databases available for any given area in Australia (see Shelley, this volume).

Map Production

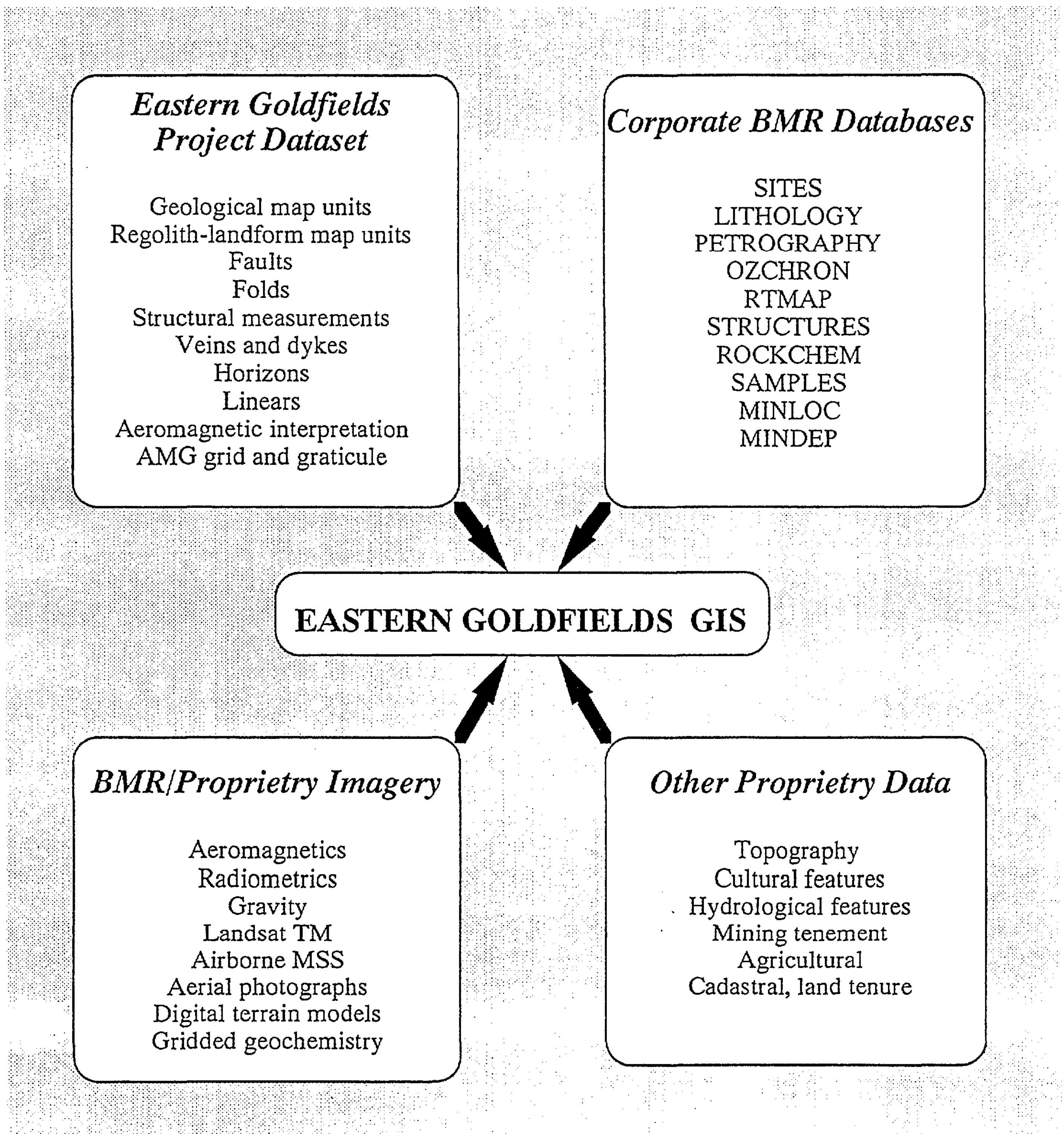
Historically, geological maps have been the primary product of the BMR's mineral province programmes. A digital GIS dataset could ultimately be viewed as a replacement for a hard-copy geological map, but at this stage the Eastern Goldfields Project recognises the value and demand for a hard-copy map product. A GIS system complete with useful quantities of data is presently a very expensive outlay and beyond the resources of many small- to medium-scale mineral exploration companies. Geological maps therefore will continue to be a primary product of the Eastern Goldfields Project, although the technical and philosophical problems of incorporating the results of GIS data integration and analysis are currently being addressed. For instance, the integration of aeromagnetics, gravity, regolith and geology to produce a basement geology model could enhance information presented on a geological map, but presents a cartographic challenge to maximise that information whilst retaining clarity.

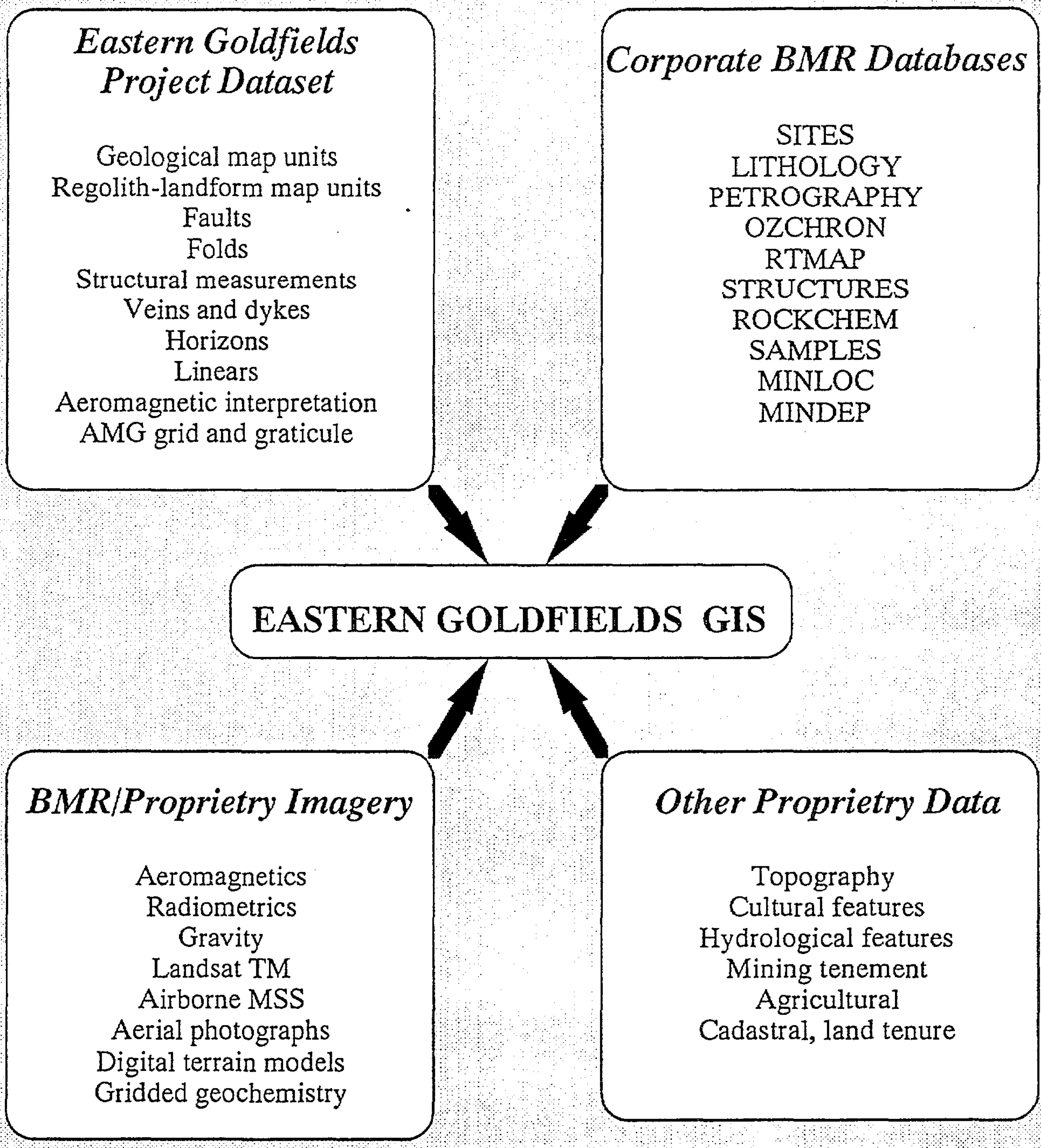
Discussion

GISs provide the means and the tools to manipulate, integrate, display and analyse any type of spatially defined digital data. For the Eastern Goldfields Project, the number and wide variety of data layers available enables many useful combinations of that data to assist the mineral industry in future exploration. GISs effectively removes the tyrannies of scale and map projection which dogged non-digital methods of data integration, such that spatial data at any initial compilation scale or projection can be compared with other spatial data with a different scale or projection. GISs are much more than overlaying facilities, however. The relational database storing the data behind the GIS can be manipulated to isolate key data or parameters, allowing the user to quickly assess possible relationships between data groups. The true power of GISs lies in their ability to spatially analyse the data groups. Relating point-source data (e.g., mineral deposits) within polygon-source data (e.g., geological boundaries), or within specified distances surrounding other point, line or polygon data is potentially a very powerful tool for the mineral exploration industry. Much of the mineral exploration process involves formulating a mineralisation model, possibly based on existing deposits or theoretical considerations, then identifying geological, geophysical or geochemical criteria for recognising the model in a variety of terrains or environments. Once the criteria have been established a GIS can quickly focus on the areas of best potential, limited only by the amount and quality of data available to analyse. The Eastern Goldfields Project aims to make available as much spatially defined digital data as possible. The clients with their own complementary data, their mineralising models, and their search criteria can harness the power of the GIS to target their exploration more effectively.

GIS data are still scale-dependent in the sense that the data's usefulness and significance are less at different, particularly larger scales than the original sampling density. The Eastern Goldfields Project GIS data package has been developed at 1:100,000 scale for the geology, and 1:250,000 scale for the regolith-terrain mapping and geophysics. Using these datasets at more than an order of magnitude scale greater or smaller than their original development scale will run into serious data sampling problems (too little or too much data respectively). The Eastern Goldfields Project GIS data package therefore is envisaged to be more useful for the exploration industry at a regional project evaluation scale rather than a detailed tenement analysis scale.

Much of the value of data integration in a GIS is yet to be realised because, firstly, the datasets need to be more fully established to maximise the benefit and, secondly, the matter of what combinations might be worth considering needs time to gestate. This gestation process is now just beginning to happen. As the datasets become more comprehensive it will become more inviting to explore previously untried data combinations.




Eastern Goldfields Project Dataset
Geological map units
Regolith-landform map units
Faults
Folds
Structural measurements
Veins and dykes
Horizons
Linears
Aeromagnetic interpretation
AMG grid and graticule
Corporate BMR Databases
SITES
LITHOLOGY
PETROGRAPHY
OZCHRON
RTMAP
STRUCTURES
ROCKCHEM
SAMPLES
MINLOC
MINDEP
EASTERN GOLDFIELDS GIS
BMR/Proprietry Imagery
Aeromagnetics
Radiometrics
Gravity
Landsat TM
Airborne MSS
Aerial photographs
Digital terrain models
Gridded geochemistry
Other Proprietry Data
Topography
Cultural features
Hydrological features
Mining tenement
Agricultural
Cadastral, land tenure

DIGITAL MAPS - MAKING BETTER USE OF SPATIAL DATA

Peter Williams and Robyn Gallagher

Bureau of Mineral Resources, Geology and Geophysics

GPO Box 378 Canberra, ACT

Abstract

A digital map is the digital database containing the spatial data for an area along with related numeric, descriptive and point-located information for the same region. It may also contain pointers to other datasets which hold data for the map area. The geological map is a pictorial representation of the information held in the digital map database, and is only one product which may be generated from the database. A range of information presented on traditional geological maps is not a primary part of the digital map database, and is thus not derivable from that source. Although some of this information can be generated automatically, its absence from the database presents constraints on the quality of the derived map products. Despite this constraint, rapid data releases from digital map databases are possible and the development of map databases provides the best means of improving the quality of data provided as early data-releases from projects.

Introduction

One major component of the planning for the National Geoscience Mapping Accord was the method of geological data handling which would need to be in place in order both to analyse and to disseminate the results of the new mapping and research programmes. With the decision to base new generation geological mapping on detailed interpretation of geophysical information, another dimension was added to the data problem, that of coping with large volumes of raster image data in a way which would allow rapid comparison with map and "attribute" data. In essence, all the data of immediate interest to the Accord could be classed as spatial data, and a highly generalised work-flow model was prepared to illustrate the progress of data analysis through the life of Accord projects (Figure 1).

However, this model did not predict many of the complexities actually encountered when the model was put into practice. The relationship between the spatial data management system and the generation of products from data has proved to be complicated. The method of analysis of raster data has improved, but is still not properly aligned to the geological map databases, and the output products are still not adequately defined to allow routine product generation.

Despite the problems however, there are clear gains emerging from following the initial approach outlined in the crude work-flow model. Centred around the concept of a digital geological map, numerous changes in data collection techniques and work practices have resulted in a rapid turn-around in data-release, enhanced interpretation of geophysical data and the development of a new product range.

The digital geological map

One might argue that there is no such thing as a "digital" geological map. A map is a picture, a representation of abstract concepts interpreted by the user through visual acumen. Thus an immediate problem arises when considering "digital" maps: - the digital representation of data is essentially useless to a geologist, there must be a conversion to visual representation to allow the user to glean useful geological information from the map.

However, the linking of geological map patterns to patterns in other data set, subdividing areas into domains, subtracting subsets of related information and widening the area of analysis through attaching adjacent maps with their related data are facilities not allowed for in the hardcopy geological map. These maps present a fixed interpretation which may not be valid through time. Thus to expand the range of data which can be used to constrain geological interpretation, some type of digital representation of the map is needed, and a certain level of analytical functionality is essential to manipulate the stored data. Consequently, a "digital" map is not a luxury for a few computer "buffs", but an essential element in the national strategy to greatly improve the geoscience knowledge base and the utility and availability of that knowledge base. "Digital" maps are the databases containing the spatial data for an area in computer-readable form, along with the related numeric, descriptive and point located information for the same area.

The digital map databases and the geological map

As the geological map is still, and will remain, an essential tool in geological interpretation, the relationship between the digital and graphical representation of the map data is a crucial question. Is it possible to take a snapshot of the database from which a "traditional" geological map can be generated? Before answering this question, it is necessary to determine what the differences are between a geological map and a digital geological map database.

It is important to accept that a digital geological map database is not designed to be a replica of the geological map produced by the geologist during field and image interpretation. Gallagher (this volume) has discussed the structure of the geological map database developed to describe the geological data behind maps from the Eastern Goldfields Project. Most of the issues discussed by Gallagher are not issues which are addressed during traditional map compilation. One way of appreciating the difference in approach is to compare the work flow necessary in producing the digital map database compared to the traditional geological map (Figure 2). Whereas the final product of the traditional approach was the geological map plus report, the new product can provide a much more comprehensive primary data set than the traditional product.

However, the geological map contains a wealth of information not depicted on the map face. The legend, rock relationship diagram, structural sketch, tectonic element diagram and reliability diagram usually added to a geological map, are not intrinsic parts of the collected geological data; however they are a part of the information conveyed to the reader. These data can, of course, be included in a digital geological map database, but are certainly not a necessary part of it. Similarly, the geological map commentary or explanatory notes are not *a priori* part of the digital database.

Frank (1991) discusses the general differences between maps and spatial databases as a spectrum of possibilities. At one extreme is the representation of the map in digital form. At the other extreme is the spatial database designed solely for spatial analysis. Between these extremes is the database from which maps of varying scale and content can be produced and which will allow sufficient versatility for spatial analysis (Figure 3). With the current database designs and cartographic systems, there is very little hope that a high quality map can be constructed from a database designed for spatial analysis. This is largely because the research necessary to develop generalization procedures, priority procedures for symbolisation, and decision support interfaces has not progressed very far.

Geological maps from the digital map database

The production of the geological map "snap shot" from the BMR digital map database has not been a simple task. A large number of cartographic elements essential for the traditional geological map are not present within the digital map database. Some of these elements are used to provide a better understanding of the geology of map area, others are designed to make the map a more readable and easy-to-use product. The example here (Figure 4) is typical of the product developed to replace the traditional compilation sheet release and the preliminary map release. The map can be released in this format within 6 months of initial data collection. However, there are several problems with the product which require discussion

Map surrounds

Examples of the information in the map surrounds are map authorship, responsibility, the means of quoting grid references, copyright information, declination and grid convergence data, map name, sheet number and the legend. This information is usually conveyed in a standard way, but is different in wording from map to map. Thus there is still a significant amount of work involved in generating the map surrounds related to any map, which is difficult to automate because the variability of design. However, most of this information is in essence non-graphical information. It is therefore questionable whether this sort of data is necessary to add to a preliminary, rapid-release map. For example, instead of the grid referencing diagram, words will adequately convey the same information.

"GRID REFERENCES:
Grid Zone designation 51H
100 000 metre square identification UH
Ignore the first two digits of northing
Ignore the first digit of easting
EXAMPLE: GRID REFERENCE TO NEAREST 100 METRES
Mount Ross: Full grid reference 6709400N,323100E
Quote as UH094231.
If reporting beyond 18° in any direction
Quote as 51UH094231.

The declination information is also easily described in writing, as is authorship and copyright. However, as copyright wording and logos do not change, they can be added automatically to any map products produced from the database.

Map Legends

The geological map legend is a complex graphical element designed to convey both descriptive and inferential information. The legend is usually arranged in stratigraphic order and may include complications such as facies equivalences, intrusive events and level of unconformities. Because the digital map database may not contain this type of information, it is necessary for the author to add this as non-primary graphical data. Diagrammatic (ie not spatially referenced) coverages may be needed to assist in representing this type of inferential data. The development of simple legend formats is necessary if the time needed to produce maps from the digital database is to be minimised. The development of a comprehensive, traditional legend should be left to the cartographic process. The necessary information for understanding the primary observational data is contained in the digital map database, and thus can be produced automatically. The label points for geological polygons can also be linked to additional external databases which can contain greatly expanded unit descriptions or other data (eg geochemical). In this way, the map legend can become quite complex, but far removed from the graphical geological reference of standard geological maps.

The symbols reference is also an essential component for understanding the geological map. The reference contains three major components, the structural symbols, geological line symbols and cartographic symbols. Because of the non-spatial and non-topological character of the symbols, there is no automatic way of generating the reference from database. As a consequence, a significant time input is necessary to generate the symbols reference. At BMR we have included the BMR Symbols Book as a digital database, but have yet to develop a suitable software link to allow automatic generation of symbols and text suitable for any particular geological map.

Point-located data

Plotting point-located data on the geological map is ideally done automatically from the digital databases holding the spatial information. For example, structural geology data, such as bedding and cleavage orientations, can be plotted automatically with appropriate software in the correct standard symbol at the point of observation. However, there is not a great deal of point in plotting multiple symbols on a single geographic location as the result will become unreadable. There must be a decision as to which symbol will be plotted and which of a group of closely located symbols will be dropped during plotting. It is relatively easy to encode the first decision in the database, by including an item such as "plot-priority", but the second decision is scale dependent and not easily encoded.

Symbols, labels and cultural features

Symbols and labels have no location in reality, but are necessarily drawn at a particular place on the geological map. This presents a significant problem in producing a readable map, because the spatial data held in the digital map database has to be checked against the position in which cartographic symbols and labels will plot. As these data are rarely topologically structured, automatic checking is not possible and a large amount of graphical editing is required to avoid conflicts. This task is highly time-consuming, and will usually need repeating if the geological compilation is later to be released as a published map. The digital map database does not store such graphical alterations back to the primary database, because to do so would compromise the integrity of the original data.

Quality of early-release map product

Given the time-consuming tasks involved in developing non-spatial graphical elements, graphical editing of point-located information and editing of cartographic symbols, the question of the overall cartographic quality of early-release products versus information accuracy and availability needs to be addressed. The problem only arises because of the availability of the geological information in a spatial database and the facilities now available to produce a product which mimics the previous first-edition maps. Early data releases in the past have been in the form of annotated, hand-drawn photo-scale compilations (or reductions there from). The products currently being prepared by the BMR Eastern Goldfields project are of much better quality. All photo joins and map joins are resolved, the annotation information has been converted into a map unit description, as an entry in the map database, and all structural symbols are accurately plotted from the actual readings, which are also stored in a digital database. Much of the map surround information (such as grids, graticule, scale bar) is generated automatically by ancillary programs. The products are available in colour as electrostatic plots or as line plots using a pen plotter. All the data which has gone into producing the map is also available in digital topologically structured form.

However the maps produced, although they look like standard geological maps, have not been intensively edited, they have numerous cartographic faults, and several data types are not well represented on the map. This particularly applies to areas which would normally carry a stipple overlay or other cartographically complex structure. Graphical elements such as the map legend, cross section and any rock relationship diagrams need to be considered separately also. These are generally not presented with a compilation sheet, and were a significant cause for delay in the release of the Eastern Goldfields map products. As much of this information as possible should be presented as text files and made available to purchasers to plot or print as required. Thus despite appearances, the products produced from the digital databases are best considered as "snapshots" of the digital map database and not equivalent to or replacements of the "first edition" geological map. Despite the availability of map products, the early-release product is really a digital map rapidly converted to pictorial form.

Conclusion

The task of converting information in the digital database to a geological map is not trivial, and there is a clear difference in the products currently generated from a digital database compared to those generated from a cartographic database. We believe that the aims of rapid data release are best served by collecting map data and point-located data in digital format, developing a digital map database from the information, and releasing the products in digital format, together with a map "pictorial representation" of the map database. The development of a first edition map and cartographic database from the primary digital map database is another issue. The advantages gained by making the digital data available early, in terms of enhanced usability of the data, improved interaction of the geological data with the geophysical data in allowing incorporation of the latter into the base for the "first edition" map and providing the basis for second-generation map products and databases far outweighs the disadvantages incurred by needing to develop a companion cartographic database.

Reference

Frank, A.U. 1991. Design of cartographic databases. In Advances in Cartography, J.C. Muller (ed) Elsevier Applied Science, London and New York, 15-44.

Gallagher, R. 1992. Geographic Information Systems: key to data management for 2nd generation map products. Record, Bureau of Mineral Resources, Geology and Geophysics.

BMR
GEOLOGY AND
GEOPHYSICS
AUSTRALIA

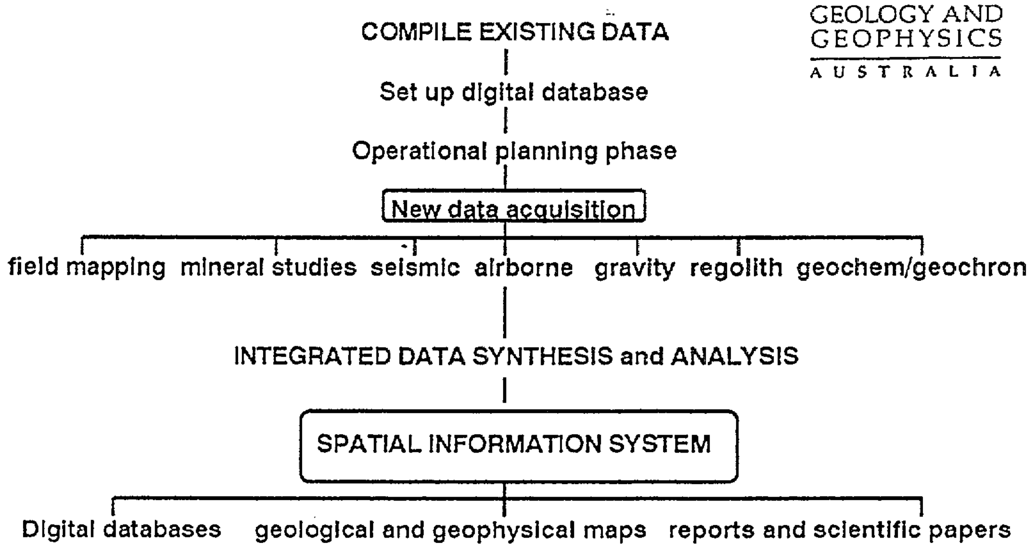


Figure 1: Idealised data flow for producing digital maps and datasets for National Geoscience Mapping Accord Projects.

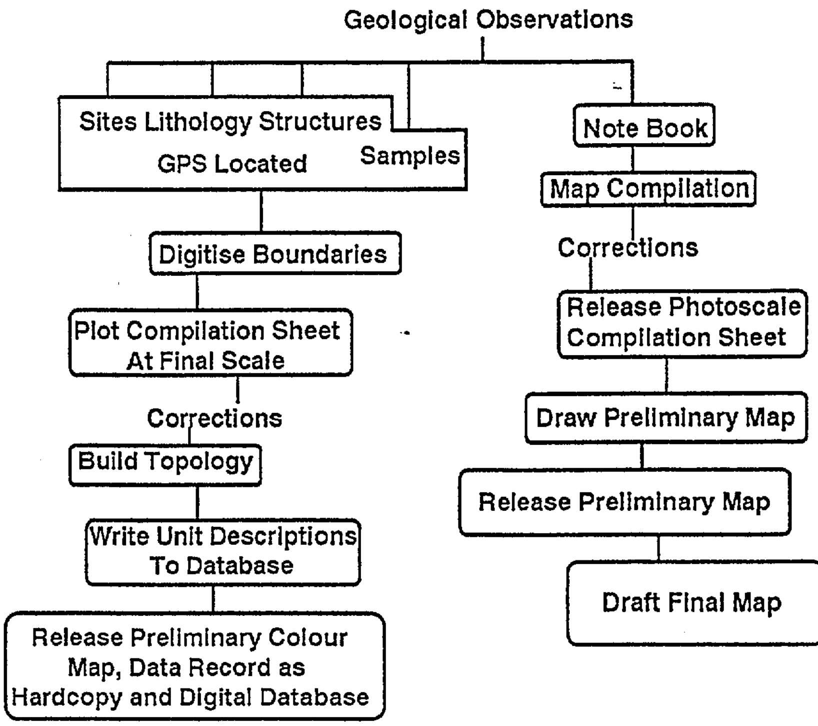


Figure 2: Comparison of the steps necessary to produce geological maps using computer-based data management and paper based data management.

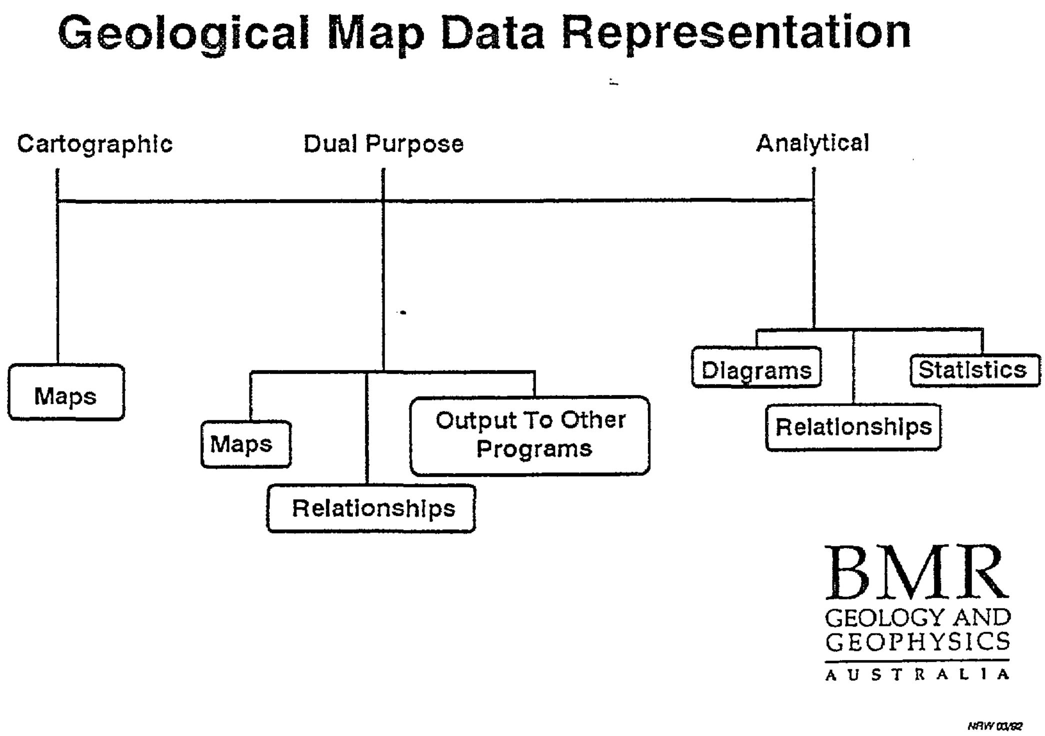


Figure 3: Schematic differences between spatial databases designed for cartographic purposes as opposed to analytical purposes.

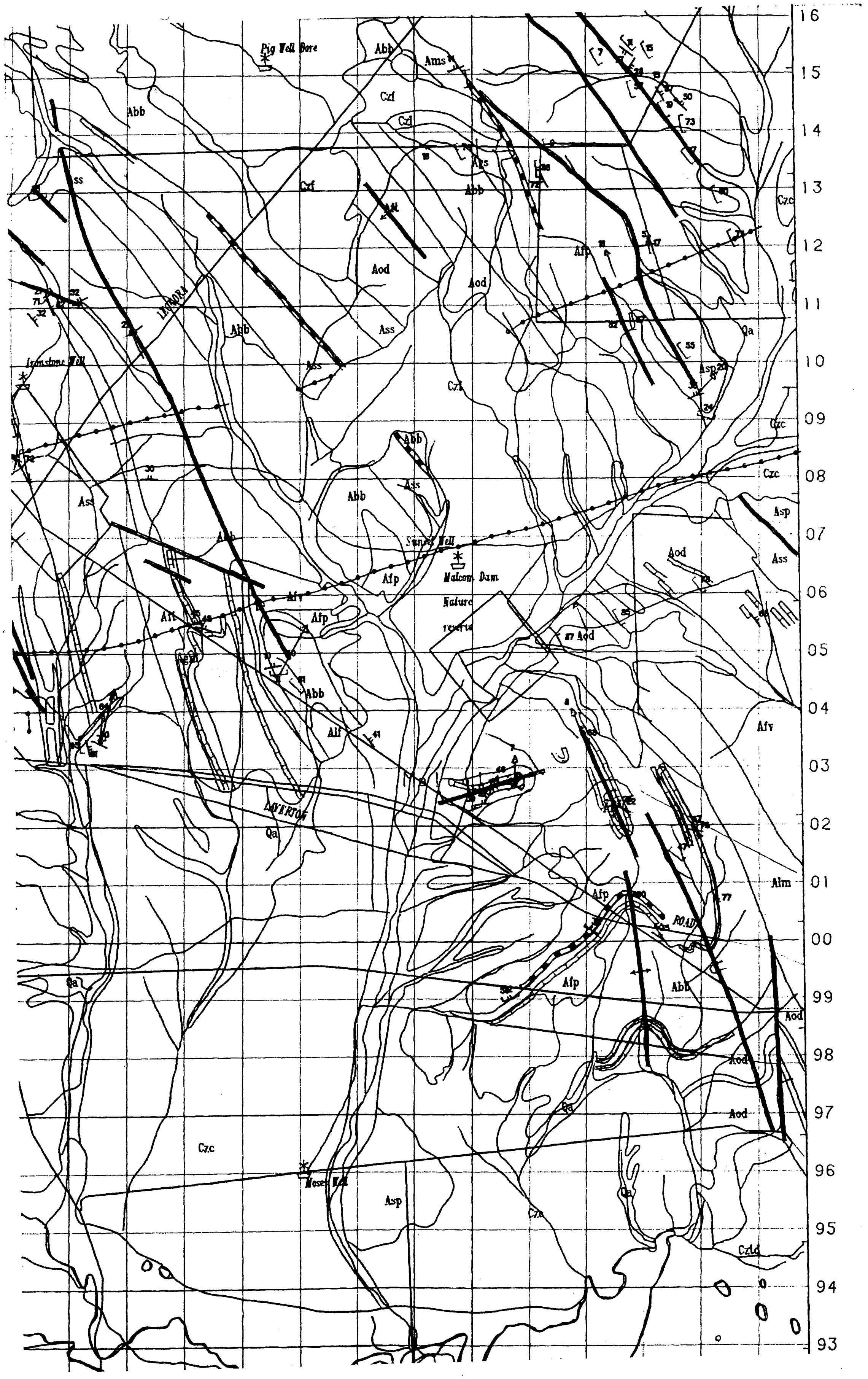


Figure 4: An example of automated output of a geological map from a spatial database designed for dual cartographic and analytical purposes.

GIS and DIGITAL CARTOGRAPHY OUTPUT OPTIONS, PRINTING ON DEMAND AND RELATIVE PRICING

John Hillier

Bureau of Mineral Resources, Geology and Geophysics

Canberra, ACT, 2600

Abstract - The production of geoscientific maps can be a long and costly process. Advances in computer graphics technology has impacted upon cartography so that it is now possible to undertake all cartographic processes from initial design and compilation to colour proofing and production of composite overlays with the new technology. Repetitive and tedious time consuming tasks are eliminated to let the cartographer concentrate on creative aspects of map publishing and integration of disparate geoscientific data sets through the wealth of opportunities presented in Integrated Geoscientific Spatial Databases and Information Systems. An exciting development amongst these is the opportunity to produce customised colour maps on demand from the digital database, especially colour electrostatic products, both paper and film based - which can be kept maintained and updated quickly as new information is added to the knowledge base. The session describes current and future technology and new applications which these new and exciting systems may discover.

In the year 2092 will it be written "I am told there are people who have never seen a geological map?", or will maps become merely "a regressive parent of GIS"?

Keywords - Geographic Information Systems/Digital Cartography

Introduction

Will there be someone to say for future generations, as Robert Louis Stevenson did for his, "I am told there are people who do not care for maps, and find it hard to believe. The names, the shapes of the woodlands, the courses of the roads and rivers, the prehistoric footsteps of man still distinctly traceable up hill and down dale, the mills and the ruins, the ponds and the ferries ... here is an inexhaustible fund of interest for any man with eyes to see or twopence-worth of imagination to understand with."

Background Reflections

There is no better way to start than to return in time and reflect on the experiences of the past two decades. Geological mapping is the key to the past and to the future.

Scenario One - Reality?

The year is 1972 and celebrations have been in full flight all around the Bureau of Mineral Resources (BMR), then within the Department of National Development. But within BMR there is not much to celebrate. After 25 years of toiling away with traditional methods of classical cartography trying to produce the first ever national geological map coverage at 1:250,000 scale there are still insufficient resources to guarantee success. Shortages of base mapping, aerial photography and experienced highly skilled cartographers bedevil progress.

Scenario Two - A Vision?

"But wait, here's another scenario. Again it is the year 1972 AD (After Digits). At the last meeting of the Expenditure Review Committee it was decreed that Australia really was the "lucky country", and the Minister for National Development was going to enrich the quality of life and standard of living of everyone. The BMR was going to show the way by completing publication of an enhanced national coverage of geological mapping at a scale of 1:250,000 so that exploration companies could develop the endowment of national resources of minerals, petroleum and hydrocarbons for the well being of the community in general, as a priority Community Service Obligation (CSO). To do this approval was given to the BMR to obtain the best computer hosted equipment available so that digital data bases could be created using new age technology. The first step was the delivery and installation of a high precision cartographic quality flat bed plotter to cope with these demands; in late 1972.

'Twas the beginning of two decades of exciting progress which "just growed", like the surges of economic prosperity, development and advancement that have taken place since the first European settlement.

PERFORMANCE ASSESSMENT

Well it is now March 1992 and twenty years down the track towards the 21st century. So how are we currently measuring up?

Is it Doom and Gloom?

Some may say we haven't changed much from the doom and gloom of the first scenario of 1971. A "banana republic" recession with unhealthy negative budgets growth, growth in deficits, unemployment, lost opportunities and diminishing education and training courses across the nation.

Renaissance!

But wait again, even with changes of governments, structures, deregulation and diminishing resources, there has been a renaissance in geological mapping through "new age" technology and a National Geoscience Mapping Accord between the BMR and the States and Territory Geological Surveys. GIS, remote sensing and integration of disparate data sets provides the catalyst. And a new P.M. has put in place a vision "One Nation" to kick start the ailing economy.

MAPS ARE EXCITING

A well drawn map is exciting; it presents information with such efficiency that with a single glance of the eye you can appreciate a fundamental truth about a city or a whole country and do this with a precision that cannot be matched by words. And the geological map marks the pinnacle of data presentation by displaying events which cover millions of years and showing clearly how they are related to each other.

MAPS PLUS INFORMATION

We have all heard the expression, "a picture is worth a thousand words", and how true it is. I would like to add, "and a picture plus a thousand words is better!". A geological map with its associated scientific commentary is information plus, plus, plus.

A geological map calls for the ability to convey the ideas of the concepts of geology by means of symbols patterns and colours. It combines the various tools to communicate geological information (language, mathematical and abstract symbols and graphic presentation). Gondwanaland, the Palaeographic Atlas and the Bowen Basin map are just three examples among many.

Map areas in geological mapping vary greatly in complexity and a high quality map product requires very close liaison between geologist, cartographer and printer. Modern regional geological maps are highly conceptual syntheses, as they embrace not only the three dimensions of space but also to an increasing degree, the fourth dimension of time.

VALUE ADDED EXCELLENCE - "MADE IN AUSTRALIA"

It is generally accepted that the standard of geological maps produced by Australia is equal to or better than anything elsewhere in the world. To make such maps is a major scientific achievement, but we take it all for granted. By studying geological maps the geoscientist selects the target of research and by using them to discover ore deposits and oil and gas reservoirs contributes to a major source of Australia's income from overseas. These maps have many of the attributes of works of art and one of no less value to the community than many better publicised paintings.

LET'S MOVE THE EARTH

Archimedes, when he was extolling the advantages of levers said, "Give me a place to stand and I will move the earth". And the same is true in geoscientists trying to understand and develop the knowledge base essential for government and the wider community for responsible development of our natural resources and to protect our environment. Let's have a look at how changes have moved the earth in the last twenty years.

"SLOWLY; SLOWLY CATCHEE MONKEY"

Four years rolled by and in 1976 a drum plotter and digitising system were added with only limited advances in automated cartographic methods, even though advances in technology had taken place, and so it continued into the 1980's with almost exclusive emphasis on production of geophysical maps.

But patience was rewarded in 1981 when the suitability of "turn-key" interactive graphics systems for the compilation and drawing of geological maps was assessed using test data comprising

- Air-photo overlap prepared by the geologist,
- an accurate topographic base, and
- control point data marked on each overlay.

It worked. The test showed that it is possible to create a distortion free map from the air photo overlays without the need for manual plotting and the digital data base was suitable for geoscientific data.

EUREKA!

At the end of 1982 a local drafting service bureau installed an "Intergraph" facility and a pilot test was made of an area from the HATCHES 1:100,000 region in the Northern Territory which showed

- enhanced productivity over manual methods,
- overall reduction in project costs,
- potential reduction in erroneous data,
- reduction in people requirements,
- standardisation between maps,
- reduction in turn around time between editing and correction of data.

A GIANT STEP FORWARD

Rejuvenation of BMR

Responding to the Challenge

National Geoscience Mapping Accord (NGMA)

As well as these activities was a comprehensive review of BMR by the late Alan Woods in 1988, endorsed by the Minister for Resources in June 1989.

NEW DIRECTIONS

The report emphasized the importance of geoscientific mapping to Australia's economic and environmental well being, as an essential prerequisite for Governments to formulate mineral, energy and land-use policies, stimulate resource exploration and development and administer the efficient use of discovered resources.

It also pointed out that there had been a serious decline in geoscientific mapping throughout Australia in recent years (since 1978).

To rectify this the Woods Report recommended that the State/NT Geological Surveys and BMR should develop a NGMA. Projects proposed were adopted in 1990 following extensive discussions complemented by a series of seminars for industry and the public. The Accord aims to produce a new generation of geological maps for Australia based on digital data bases and modern data collection and processing technologies, with the main emphasis on updating 1:250,000 or in some cases larger scale, maps of key regions. Nine projects have been identified in four different geological regions covering mineral provinces. Five further projects cover major sedimentary basins, supported by continent wide and regional studies of geological events transcending individual basin boundaries. High resolution magnetic and radiometric data is a pre-requisite for these new generation products, as well as gravity and remote sensing (satellite imagery and airborne multispectral scanning).

"Geoscience for Australia's Future"

In June 1990 the Minister for Resources launched a new program strategy statement for BMR - "Geoscience for Australia's Future".

Environmental Geoscience

The statement very clearly showed the recognition of the need to be concerned with geoscience relevant to the consequences of resource development and use, in response to the broader needs of society for sustainable development, to the extent that the Bureau established an Environmental Geoscience Unit to commence work in 1990/91. The challenge before us is to develop and manage our natural resources in a way "which meets the needs of this generation without making it difficult or impossible for future generations to fulfil their own needs (UN definition of sustainable development).

New Structure

The Woods Review recommendation that the bureau be restructured to assist it to meet contemporary challenges was accepted by Government.

Developing the Tools

As well as people development we need to acquire methods, procedures and applications development.

Experimental Development

The Australian Bureau of Statistics defines this as "Experimental development" - "the systematic work, using existing knowledge gained from research or practical experience for the purpose of creating new or improved products/processes."

We are certainly creating new and improved products involving new and improved processes to meet the insatiable demands of new initiatives arising from programs such as the National Geoscience Mapping Accord, the Groundwater program and the Environmental Geoscience program.

Unlike cartographic data bases, which model maps and record such map features as the widths of lines or the positions of symbols, geographic information systems are models of reality about which information may be extracted and presented in various formats. The pure storage of map graphics is often inappropriate, because the exact rendering depends on the scale, area of display, and content of a drawing, and must be recomputed for each actual display. Users may commit serious errors if, for instance, labels are placed outside of the visible screen area. A GIS stores objects with their properties and records the (spatial) relations among objects and therefore is a repository for a wide range of graphical presentations (eg, map-like sketches, blue-prints, high-quality topographic maps, and thematic maps).

The multipurpose character of a GIS requires that geographic information about these objects be displayed in a variety of formats. A GIS has to be more flexible and show data in the form most appropriate for users' tasks. The style of the graphical presentation, therefore, cannot be determined ahead of time and users must be able to do this themselves.

The digital encoding of map features and subsequent computer representation at many levels of resolution raises challenging questions. Is it possible to produce multiple graphical representations from a single digital file, and preserve both accuracy and recognizability? Over what range of scales may a single data base be applied? How may digital encoding be designed to accommodate those features whose structure changes with scale, while also constraining other features in the data base whose structure does not change? And finally, without attempting to formalize the rules for generalization, can we at least formalize the types of geometric structure one can expect to encounter in representing features on maps?

LETS KEEP TRYING

Comparison Established and Well Practised Cartographic Processes

Lets be realistic, there's no doubt nor uncertainty that with technological reform traditional cartography merging art and science to create accurate hand drawn maps is undergoing historic change. Now cartography has been revolutionised by computer and changed forever by advances made in:

- CAD/CAM and GIS systems
- Electronic publishing
- Geodesy/geopositioning
- Data acquisition systems

However, there is a need to be aware of subtle differences. A digital cartographic object is not the same as the geographic feature it represents. That is to say, the graphic representation may progress through structural and visual variation as it is depicted at finer levels of resolution. In the past when base maps were compiled manually this was not a problem, because a different compilation was produced for each required map scale. However, use of a single digital base map file for depicting features at several map scales brings to attention both consistency and recognizability of base map information as it is generalized.

Traditional cartography utilises fairly standardised series of tasks divided into discrete work areas.

Drawing Office

- collecting, processing data
- compiling manuscript - integrating data sources
- generalisation and classification
- symbolisation
- scribing
- type order and placement, legends and margins
- colour design

Visual cues are evident in a graphic depiction that map readers use to identify one feature as a river, another as a road, or a third as a political boundary. Some cues, such as color do not have a direct geometric component. Other visual cues may be measurable as the geometric details evident along the extent of a feature. These details may also be used to identify the scale of graphic depiction, in the context of geomorphic applications (Goodchild 1980). Cartographically speaking, it is essential to retain both the details required for geographical accuracy and required for recognizability within a digital data base. This is the basis of map generalization, and the task is accomplished intuitively, when done by hand, and is based upon visual logic ('this solution looks right').

Reprographics

- expose separates
- make masks, peel coats
- expose composites
- make proofs
- repromat to printers for plate making

Digital Cartographic Processes

The digital cartographic work environment changes from traditional drawing tables and vacuum frames to workstations and film writing.

Workstation

- Input vector source or extract from data base
- Transform co-ordinates
- Generalise features
- Place text, generate legend, margins and so on
- Define output symbology
- Process data, review on screen

For computer mapping or analysis of GIS map information the issue becomes more complex, as analytic computations are performed on the digital version stored on disk, and not necessarily on the version displayed on the CRT screen. The graphic depiction seen by a GIS user may appear at a different scale than the digital version, and may contain only a subset of the coordinates stored on disk. Rules to determine which subset of detail to incorporate into a given scale of feature depiction must be explicitly defined, and incorporated into our display software automatically, to preserve consistency and accuracy for GIS displays.

To preserve accuracy and recognizability automatically during map generalization, one must be able to describe digitally the details that must be preserved. Digital encoding of the type of details stored in a coordinate file may provide more consistent guidelines for both statistical and cartographic generalization, and reduce problems in subsequent analytical tasks, such as edge matching across map sheets. The difficulty is that graphic details may vary distinctly from one scale of depiction to the next, and from one line feature to the next. It is probable that a single digital description model cannot appropriately describe all types of cartographic details that may be encoded for map representation.

Film Writer

- Generate composite films

Reprographics

- Make proof or preferably use screen based "soft-copy"
- Repromat to printer for film making

Wait - Customised Maps on Demand

Benefits of Automation in Cartography

A number of factors lead to automation:

- Fragile BMR repromat collection. This vast collection of graphic map separates at various scales and projections is becoming fragile with age and as a unique resource represents a huge investment in labour; but keeping it up to date by hand revision of each map separate is becoming prohibitively resource intensive.
- Increased high cost of materials. A single map requires many pieces of film becoming more expensive each year.
- Archive costs. There is no economical way to archive hard copy maps repromat. Fire and natural disaster are serious hazards.
- Revision. Maintenance and updating has not been undertaken. There is a practical limit to the revision of boundaries, name changes, additions and corrections that can be made to aging sets of repromat film.
- Labour Costs. Re-compiling regional maps on new projections and datums takes a lot of skilled cartographic effort, a computer can do it in a speedy and trivial process.

Key factors in digital map publishing.

Availability of high quality film writers.

- Demand for multi-purpose, multi-user spatially referenced data bases to support analysis and integrated products - paper maps cannot compete
- Availability of digital data - purchase or exchange
- Print quality text now available in digital fonts.

Advantages of digital map publishing

- Near perfect registration, screen tint and colour control - enhanced quality
- Increased throughput during revision and maintenance
- Increased capabilities of a digital data base to produce multiple products
- Increased flexibility in map design - easy to customise product for the user
- Streamlined quality control/assurance cycle using softcopy review and hard copy check plots from colour electrostatic process on demand
- Increased resource control - few processes required (scribing, type-setting, proofing)

Digital map publishing implementation

A totally automated map publishing system from source data digitisation, compilation and finally cartographic quality repromat is in place and available for the cartographer now.

New customised software replaces entire darkroom processes.

New tools such as cursors are used instead of a scribe cutter and repetitive and labour intensive jobs such as duffing and peel coating can be cast aside, while the cartographer can concentrate on the creative aspects of cartographic design to develop new products.

The power of computer technology presents new opportunities to benefit not only from being able to manipulate information in the data base for the production of final maps but also in the wealth of opportunities in the integration and analysis of data sets presented in geoscientific information systems.

Video Discs and Mapping

A number of mapping agencies already use video discs as mapping mediums. The introduction of integrated microcomputer/video systems (especially those containing video discs) enables the production of highly flexible and very usable "video atlases" which will enable the portrayal of dynamic optical data in a format acceptable to a variety of users, from novice to expert from administrator to engineer and scientist.

The contemporary classic example is the US Defence Mapping Agency Digital Chart of the World (DCW) and "Desert Storm" applications.

The Value of Maps in Contemporary Society

Since World War 11 a great effort has been made in Australia to make sure that information is available to those who wish to use it, and this is important because although maps are things of beauty they are also made to be used.

Most of us if we are going on a journey, will consult a map. If it is a good map we will probably reach our destination without difficulty. The geological maps show the way for the organisations which are developing the resources of the country; a good geological map will not remove all the risks involved in the enterprise any more than a good topographic map will keep you safe from all the dangers of the road, but it will provide the additional information required to guide the explorer away from the more barren areas or provide the information needed to recognise the potential of other regions.

Discoveries do not come directly from published data. They are the result of the combination of the results of long term systematic investigations by the government departments with the special insight provided by the exploration groups; success usually depends on the contribution from both parties.

The future is not likely to be so different from the past; the mineral and the oil discoveries which we need to carry us into and through the next century will depend on the work which has been done over the last three decades and which must be continued; for the time between the basic mapping and related studies and the discoveries is often long and to neglect the basic work now will hazard the success of the next generation.

The present is not only the key to the past, it is also the key to the future. The way will be found by having the best possible geological maps at our disposal. If they are good maps we will probably reach our destination successfully.

Conclusion

Leonardo de Vinci, the Renaissance genius at the end of the 15th century, was in a sense a father of modern geology. As his superb drawings of rock outcrops show, he had a clear concept of rock layering and folding.

What is more remarkable is that, in one of his capacities as a civil and hydraulics engineer, he observed fossils within rock structures of his native Tuscany and correctly deduced their time significance.

The modern science of geology, which depended on Da Vinci's wider approach had to wait for some hundreds of years until the work of the pioneer Babbage introduced the "modern computer".

This new technology eases the burden of repetitive and labour intensive techniques. Embracing the technology helps decision makers to solve problems by having all the needed information "on hand" when and where required.

As society becomes increasingly complex there are more and more conflicting natural resource, land management and environmental issues. There are greater demands to more readily capture, manipulate and provide analytical capabilities from spatial data. We live in a rich world of visual experience, a world of complexity and intricacy. We must capture that world within the bounds of flatland, the two-dimensional page. "Escaping this flatland and enriching the density of data displays are the essential tasks of information design". Escape from flatland depends upon two strategies: enhancing dimensionality (increasing the number of represented dimensions) and increasing data density (increasing the amount of information per unit area.) In bursting the bounds of flatland, we struggle with the constraint of dimensional compression that stems from compromise and trade-offs, and we encounter difficulty as the "... ties of data to our familiar three-space world weaken (with more abstract measures) and as the number of dimensions increases (with more complex data)".

Cartographers might be pleased with the assertion that "(s)tandards of excellence for information design are set by **high quality maps** ..."

What exactly is information? Is there a difference between information and data? To what extent are these concepts found in the eyes of the beholder and/or the producer? The question of the function(s) of a particular graphic rarely is addressed satisfactorily. Is the map the acme of information design because it contains so much information? Are the best maps those that contain the most information? "...it is not how much empty space there is, but rather how it is used. It is not how much information there is, but rather how effectively it is arranged".

"Knowing where things are and how they relate to one another is crucial for management, planning and investment decisions taken within both public and private sectors".[1]

Reference

[1] HMSO (1987)., Handling Geographic Information. Report to the Secretary of State for the Environment of the Committee of Enquiry into the Handling of Geographic Information.

Integrating Raster Data into a Spatial Data System

Prame N. Chopra and John W. Creasey

Bureau of Mineral Resources, Geology & Geophysics
Canberra, ACT, 2601

Abstract - In this paper we discuss the raster data domain and its role in integrated geoscience studies. The strengths and weaknesses of raster data as it is currently handled are presented and the major ways in which raster data can be usefully integrated into an overall spatial data system are discussed.

Three levels of integration of raster data with other spatial types are described: "Separate but Equal", "Seamless Integration", and "Total Integration" (Ehlers et al, 1989). The first level of integration is more or less possible now using existing GIS and image processing systems together with file transfers in standard interchange formats. The second and third levels of integration are still someway off in the future.

Keywords - Raster, vector, data integration, image processing, GIS

Introduction

In this paper we are concerned with raster images of the Earth, not with rasters derived by scanning maps, overlays and documents.

A spatial data system can be comprised of three types of spatial data: point, vector and raster. The former data types share many features and are thus readily integrated. Both point and vector data are composed of discrete spatial entities which do not have a spatial extent. Points have no size and arcs, while they do have a length, have no area. Polygons (i.e. sets of linked arcs) have a spatial extent but because they are comprised of arcs they can still be integrated easily with points and other arcs.

Raster data on the other hand, are more difficult to integrate with point and vector data because of their innately different format. Raster data such as a satellite-derived image or a grid constructed from geophysical observations, consist of a regular array of non-overlapping cells of a particular size and shape. Each cell in this array has associated with it a number which represents the average value over that cell of the parameter that the raster depicts. Multiple-band rasters such as Landsat TM contain measurements by different instruments over the same set of cells.

While the integration of raster data with the other data types can be difficult, the rewards can be great. Raster data offer a different way of looking at the Earth when compared with discrete ground-based observations and provide unique processing and modelling opportunities. In the case of satellite-derived raster data, there are additional benefits such as the availability of regular repeat coverages, and access to data adjacent to the main area of interest.

The integration of raster data with point and vector data into an overall spatial data system is normally an iterative process. The point and vector data are used to interpret processed rasters (i.e. images) and to guide further processing. The interpretations that result from this phase can in turn be used to update the point and vector data holdings themselves. This iterative process can result in considerable synergy.

Vector-Raster Data Structures

Perhaps the fundamental difference between the raster data model and those for point and vector data as epitomised in a modern geographic information system (GIS), concerns topology. In a modern vector GIS (e.g. Arc/Info) the points, lines and polygons in a GIS dataset are topologically related. Thus the spatial relationships between the various entities are defined. The spatial location of an arc, for example, includes its direction and where this is meaningful, the identity of the polygons it abuts.

Raster datasets, however, offer a fundamentally different representation of the Earth than is provided in a vector-based GIS. A distinction can be made between object-based and domain-based information (Ehlers et al, 1989). Vector data in GIS and cartographic systems on the one hand fill an empty 2D space with objects. Images on the other hand separate the space into gradational domains which carry with them more information. Raster data in their raw form are a lower level of information than that which is held in a GIS.

The raster data model is not topological. The cells or pixels in the raster can be stored sequentially (e.g. band interleaved by line (BIL) and band sequential (BSQ)) but this need not be the case (e.g. band interleaved by pixel (BIP) or tile/brick format). For the purposes of processing and display, these rasters are made up of individual pixels, each of which is an independent entity.

The pixels in a raster are also usually constrained in a way that other spatial elements are not. With the exception of those raster systems which work with quadtrees (e.g. SPANS) raster systems generally require each pixel in a raster to be the same size and shape as all the other pixels. This requirement makes the combination of raster datasets from different sources difficult when the pixel sizes differ. This issue has been addressed in recent systems such as ER-Mapper.

Because the pixels in a given raster need to be of a fixed size, inaccuracies can arise in regions of an image where the gradient in the pixel values is high. In such instances, boundaries will be imprecisely located. Thus raster data are not scale independent in the way that point data are for example.

Lastly, raster datasets normally cannot be linked with auxilliary attribute data held in external database tables. Thus it is generally not possible to, for example, select pixels from a Landsat scene on the basis of a landuse code stored in an accompanying attribute table. However, the Arc/Info Rev. 6 module GRID does allow this type of data manipulation for rasters in which the pixel values are integers.

Raster Data Characteristics

Raster datasets fall into two broad types: those which are directly recorded in a raster format (e.g. satellite data such as SPOT) and those which are computed from some other primary data source (e.g. a digital terrain model constructed from topographic height data). Of these two types, the first tend to be uncalibrated data (i.e. the data values carried by the pixels are only in relative rather than real physical units). The second type are generally calibrated data.

Calibrated Data

Calibrated data refer to measurements of physical properties such as total magnetic intensity or gamma radiation. Geoscience-related raster data of this type include:

- airborne geophysics (e.g. magnetics, radiometrics)
- ground-based geophysics (e.g. gravity, transient electromagnetism)
- geochemical analyses of rocks and stream sediments
- topographic elevation

For these datasets, considerable processing is required in order to generate a raster dataset that can then be further processed. Generally the processing requires interpolation and extrapolation of a regular array of points from what can often be a very inhomogeneous set of observations.

The raster produced can contain artifacts introduced by the gridding procedure and these can, in the worst case, overwhelm the actual signal. Even in cases where the processing artifacts are relatively minor however, there is still a risk that the information preserved in the raster will be significantly degraded when compared with that in the orginal measurements. The spatial definition of features can for example become smeared. Striping of the image can also arise from the levelling problems inherent in the gridding of line data such as airborne geophysics.
The primary data itself serves as an essential input to a GIS along with various filtered images that emphasise and better resolve local geological features.

Uncalibrated Data

In the majority of cases, raster data sets are uncalibrated. By this we mean that the raw values represented by the raster are not related to the physical properties of the materials being studied. For example, the detectors on Landsat and SPOT satellites measure the reflectance of light from the earth's surface and record the data in byte format (i.e. integers in the range 0 to 255). These "reflectance" data can then be further processed to derive 'thematic' images that estimate proportions of materials at the earth's surface. How well this is done is a function of the spectral and spatial resolution of the measurement system and the processing methodologies used.

Thematic images represent a higher level of abstraction of raster data which is more akin to the data structures normally handled in current GIS systems. However these thematic images are still domain-based and thus contain a continuum of data values rather than the conventional structure of categorical data.

Spatial Referencing
Raster data can be used to complement and enhance the other data in a GIS by providing up-to-date and even more accurate definition of spatial objects. Using digital matching techniques, stereo satellite imagery can also be used to develop digital terrain models. Accurate representation of the terrain is at least as important as the accurate representation of land cover (Trotter, 1991).

Incomplete Data Coverage
Many, but not all, remotely sensed raster datasets are adversely affected by cloud cover (examples of datasets that are unaffected include airborne magnetics and radiometrics and air- and satellite-borne radar). When raster datasets that have been affected by cloud cover are to be integrated with other data, some thought must be given as to how the obscured areas are to be displayed. Should these areas be left as they are, or should image processing be used to set them to a null value. Similar thought must also be given to shadows associated with the cloud cover.

The Size of Raster Datasets
Raster datasets tend to be very large. For example, a typical Landsat TM scene requires on the order of 250 Mbyte of storage space when it is stored in BIL format. Image processing usually requires the creation of a number of intermediate datasets before a final processed version is produced. Thus virtually any processing of whole Landsat TM scenes will require upwards of 1Gbyte of disc storage per scene.

Given the size of these datasets it is obviously very important to minimise the duplication of image holdings within an overall spatial data system. Thus it is very inefficient to maintain the same raster data in one format for the image processing system and in a separate format for use as a GIS backdrop. This has however been necessary in BMR with the I^2S and Arc/Info systems because the former's FLAB format is incompatible with the latter's supported formats.

Repeat Datasets
In the case of satellite-derived raster datasets, it is possible to obtain repeat datasets at a frequency of between a few hours and a few weeks. This succession of datasets allows monitoring of changes in the area of interest and also provides the opportunity to fill in gaps in the data arising from factors such as shadows and cloud cover.

Processing of repeat datasets can be fairly straightforward once a processing scheme has been established. This is especially so when using dynamic algorithm processing such as is employed in ER Mapper. The data volume issue rapidly becomes important however.

Multi-band datasets
Most of the instruments used in remote sensing produce multi-band raster datasets. Landsat TM for example has 7 bands, each of which records data for a range of electromagnetic wavelengths. Processing this data to a higher level can produce more multi-band datasets. To display these datasets properly for interpretation requires a full 24-bit colour display capability such as is usually found in powerful image processing systems. Current GIS hardware generally only includes 8-bit displays.

Thus hardware limitations can jeopardise the early stages of vector-raster integration in a GIS since it may be impossible to display the raster data in its final processed form. One solution may be to do the integration within the image processing system.

Georectification of Raster Data Sets

Whether a raster has been directly recorded (as in the case of a remotely sensed data set such as a SPOT or Landsat scene) or has been computed from other data, it will inevitably need to be warped into a common projection before it can be integrated with other data. This warping procedure has a far more dramatic effect on raster data than it does on point and vector data. The latter types of data can be warped in a GIS without any loss of information. The spatial relationships between entities are preserved during the warp and the attributes associated with each entity are unchanged. Thus, for example, a fault with a throw of 250 metre separating two polygons representing different lithologies will still have the same throw and the same relationship to the polygons, irrespective of whether the data are displayed in a simple conic projection or in a UTM projection.

Raster data however must be resampled when they are warped from one projection to another. This resampling is necessary because each pixel in the raster is remapped to a different coordinate system and hence usually ends up representing an at least slightly different area of the image. This effect is illustrated in Figure 1 for a hypothetical raster of six Landsat pixels.

In Figure 1, the square pixels represent the initial coordinate projection. The non-rectilinear pixels have been derived by warping the original six pixels into a new coordinate projection. These warped pixels have then been drawn back over their original projection. The non-rectilinearity arises because the two projections are not linearly related.

Note in particular that the warped pixel in the upper left corner of the raster incorporates portions of three original pixels. Hence the reflectance value that the new pixel carries should be a composite of that borne by the three pixels it overlaps. This is the essence of resampling and it should be apparent that the process will lead to degradation of both the spatial resolution and the signal carried by the raster.

Because of this degrading of the image by resampling, it is generally recommended that image interpretation be carried out before warping (Harrison and Jupp, 1990; p. 32). This processing arrangement however presents a problem when images are to be integrated with other spatial data, a point that we will return to later.

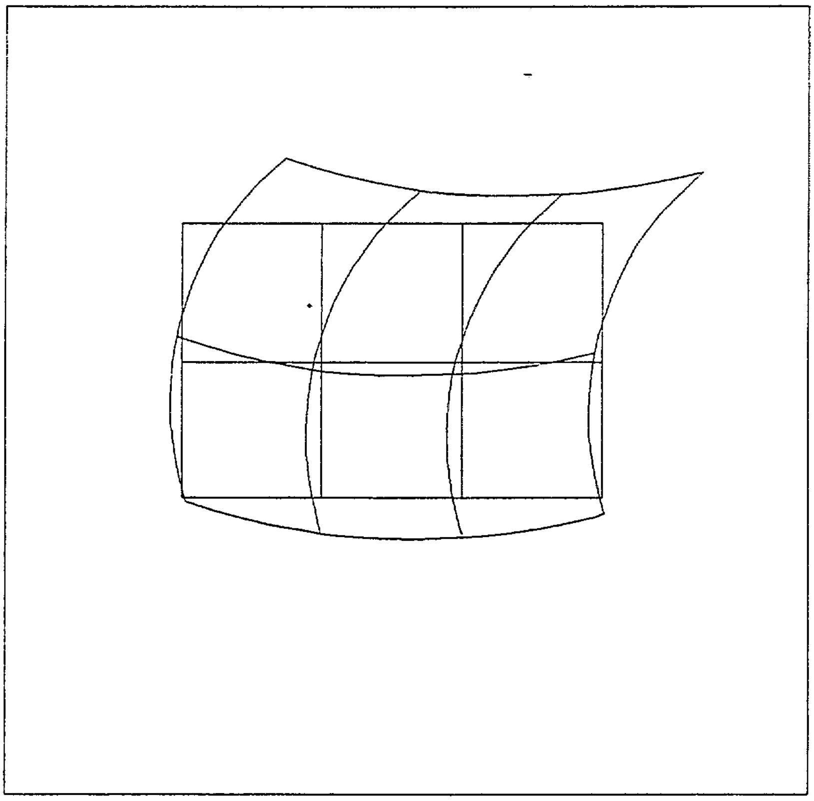


Figure 1 A schematic diagram illustrating how the warping of a raster from one coordinate projection to another causes the pixels to represent different areas of the image. This change in pixel location necessitates resampling of the raster with concomitant degrading of its content.

Modelling with Raster Data

Raster data are particularly well suited to modelling tasks and for this reason most early GISs were raster rather than vector based. In particular, it is fairly simple to construct new composite raster layers by overlaying and combining a number of source rasters. In the vector domain, these composite layers are much more difficult to produce (Bracken and Webster, 1989).

Types of Integration

Three levels of integration of spatial data and spatial data systems can be defined according to Ehlers et al, (1989). These levels are:

1. "<u>Separate but Equal</u>"

   - At this level, the simultaneous display of raster and vector data is possible.
   - The results of low-level image processing can be moved to the GIS to permit on the one hand, identification of "themes" and, on the other hand, refinement of the topology.
   - The results of integrating vector and raster data in the GIS can be transferred, in the form of vector overlays, to the image processing system for further analysis of imagery.
   - The results of GIS modelling can be moved to the image processing system for support and validation of the image.

As remarked previously, these types of low level of integration are often used in an iterative processing scheme. The ultimate goal of this scheme is to enhance both types of data. For example, the spatial accuracy of the vector data and the level of interpretation of the raster data can be increased.

2. "<u>Seamless Integration</u>"

   - At this level, the two technologies, GIS and image processing still exist in their own right, but tandem processing is possible.
   - It is possible to exercise entity-like control over image components such as themes.
   - GIS vector data can be directly incorporated into images (i.e. by automatic rasterisation).
   - Errors, both random and systematic, can be analysed.
   - Simulations involving the integration of vector and image data and temporal considerations can be used to model "what if" questions.

3. "<u>Total Integration</u>"

   - At this final level, a new hybrid data model has been developed which allows the development of a single technology to replace GIS and Image Processing.
   - An integrated spatial query language has also been developed and this includes image-oriented query functions.

Conclusions

In this paper we have discussed the raster data domain and its role in integrated geoscience studies. We have emphasised the strengths and weaknesses of raster data as it is currently handled and we have discussed the major ways in which raster data can be usefully integrated into an overall spatial data system.

The vector-raster dichotomy has been a serious impediment to the effective integration of spatial data. Advances in technology, improvements in data models and a better understanding of the role of raster data are reducing this obstacle to data integration.

Three levels of integration of raster data with other spatial types have been described. The first level of integration is more or less possible now using existing GIS and image processing systems together with file transfers in standard interchange formats. Thus in BMR for example, GIS-derived vectors can be displayed over imagery in both GIS and image processing systems.

The second and third levels of integration are still someway off in the future.

References

Bracken I. and Webster C., 1989 -- Towards a typology of geographic information systems. *International Journal of Geographic Information Systems,* 3, 137-152.

Ehlers M. , Edwards G. and Bedard Y., 1989 -- Integration of remote sensing with geographic information systems: a necessary evolution. *Photogrammetric Engineering & Remote Sensing,* 55, 1619-1627.

Harrison B.A. & Jupp D.L.B., 1990. *Introduction to Image Processing,* CSIRO Publications, Melbourne , Australia, pp. 256.

Trotter C. M., 1991 -- Remotely sensed data as an information source for geographic information systems in natural resource management: a review. *International Journal of Geographical Information Systems,* 5, 225-240.

GIS DATA STRUCTURES

VECTOR OR RASTER BASED GIS? - *DILEMMA*

VECTOR AND RASTER - *INCOMPATIBILITY*

VECTOR DATA: POINTS, ARCS, POLYGONS -
- -*TOPOLOGY*

RASTER DATA: CATEGORICAL OR CONTINUOUS
- *NO TOPOLOGY*

VECTOR DATA - OBJECT BASED
- ABSTRACTION OF REAL WORLD

RASTER DATA - DOMAIN OR FIELD BASED
- MEASURE OF REAL WORLD
- LOWER LEVEL OF INFORMATION

IMAGE PROCESSING INCREASES THE INFORMATION LEVEL OF RASTER DATA

DATA RESOLUTION AND EXTENT

SPECTRAL	- *RESOLUTION:* THE WIDTH OF WAVELENGTH CHANNELS
	- *EXTENT:* THE NUMBER AND SPECTRAL RANGE OF CHANNELS
SPATIAL	- *RESOLUTION*: PIXEL SIZE
	- *EXTENT:* THE OVERALL IMAGE COVERAGE
RADIOMETRIC	- *RESOLUTION*: THE ENERGY REQUIRED TO INCREASE THE SIGNAL BY ONE VALUE
	- *EXTENT*: THE NUMBER OF LEVELS DETECTED
TEMPORAL	- *RESOLUTION:* THE REPEAT CYCLE BETWEEN SUCCESSIVE ACQUISITIONS
	- *EXTENT:* TOTAL PERIOD OVER WHICH DATA IS AVAILABLE

RASTER DATA CHARACTERISTICS

TWO TYPES:

1. *CALIBRATED* - AEROMAGNETICS, RADIOMETRICS

- DIRECT MEASUREMENTS OF PHYSICAL PROPERTIES (EG. TMI)

- DATA ARE IRREGULARLY SPACED POINTS: PRIMARY DATA

- GRIDDED TO FORM RASTER DATA: INTERPOLATED PRIMARY DATA

2. *UNCALIBRATED* - LANDSAT TM, SPOT

- DATA RECORDED IN RASTER FORMAT

- PRIMARY DATA NOT DIRECTLY RELATED TO PHYSICAL PROPERTIES OF MATERIALS

- PRIMARY DATA IS A MEASURE OF RADIANCE FROM SURFACE OF EARTH

- A MULTICOMPONENT SIGNAL

PROCESSING RASTER DATA - IMAGE PROCESSING

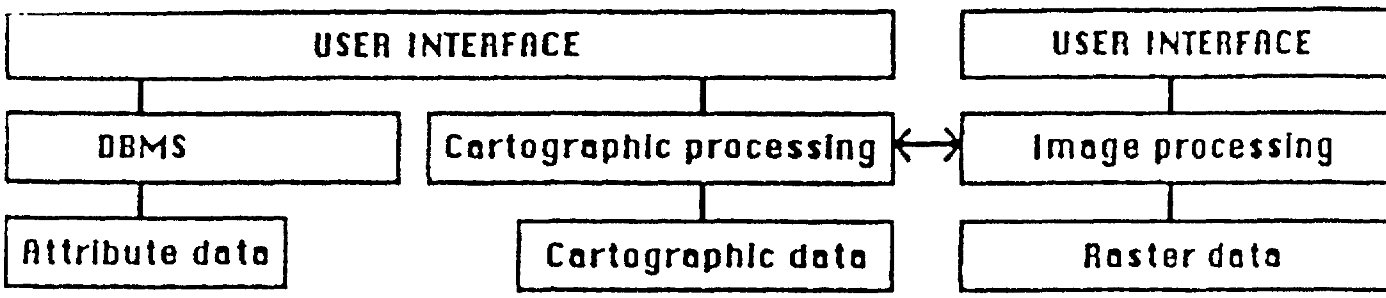
- GEORECTIFICATION OF RASTER DATA: MANDATORY BUT REQUIRES FURTHER RESAMPLING OR INTERPOLATION

- RASTER DATA NEED TO BE PROCESSED FURTHER TO PRODUCE THEMATIC IMAGES OR MAPS
EG. LANDSAT TM: CLAY, IRON OXIDE, VEGETATION
AEROMAGNETICS - SPATIAL FILTERING FOR HIGH FREQUENCY MAGNETIC EFFECTS

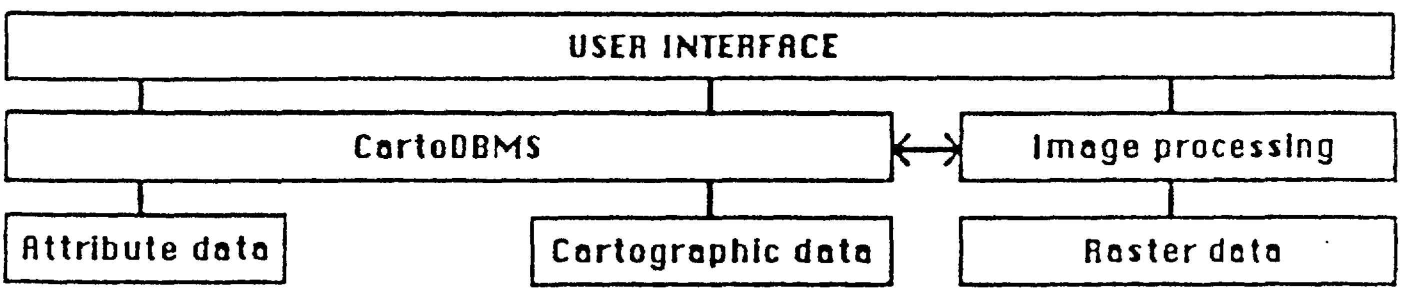
- HOW WELL THIS IS DONE IS A FUNCTION OF RESOLUTION OF THE MEASUREMENT SYSTEM AND PROCESSING METHODOLOGIES USED

- THEMATIC IMAGES REPRESENT A HIGHER LEVEL OF ABSTRACTION OF THE PRIMARY RASTER DATA

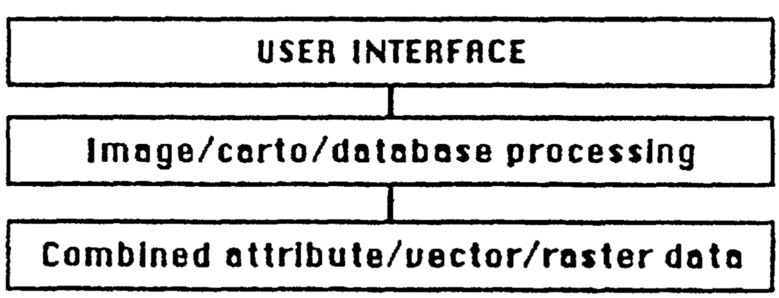
- THEMATIC IMAGE DATA STILL CONTINUOUS DATA, BUT ARE THEY DOMAIN OR OBJECT BASED?


USER INTERFACE
USER INTERFACE
DBMS
Cartographic processing
Image processing
Attribute data
Cartographic data
Raster data

(a)


USER INTERFACE
CartoDBMS
Image processing
Attribute data
Cartographic data
Raster data

(b)


USER INTERFACE
Image/carto/database processing
Combined attribute/vector/raster data

(c)

RASTER - VECTOR INTEGRATION

3 LEVELS OF INTEGRATION

1. FIRST LEVEL - " SEPARATE BUT EQUAL"

- MOVE RASTER AND VECTOR DATA BETWEEN GIS AND IMAGE PROCESSING SYSTEM

- ITERATIVE PROCESS

RASTER DATA:

- PROVIDE BACKDROP FOR VISUALISATION
- VERIFY AND/OR MODIFY VECTOR OVERLAYS

VECTOR DATA:

- IMAGE STATISTICS FROM REGIONS OF INTEREST

INTERPRETATION:

- PRODUCE HIGH LEVEL VECTOR OVERLAY FROM RASTER DATA

2. SECOND LEVEL - "SEAMLESS INTEGRATION"

- *DYNAMIC LINKS BETWEEN GIS AND IMAGE PROCESSING*

- *RASTER DATA RESOLUTION AND ERRORS ACCOUNTED FOR*

VECTOR DATA:

- TOPOLOGY REMAINS INTACT IN IMAGE PROCESSING SYSTEM

RASTER DATA:

- ATTRIBUTE LINKED IN GIS

INTERPRETATION:

- ANALYSIS AND MODELLING USING BOTH DATA SETS IN GIS

3. TOTAL INTEGRATION

- NEW HYBRID DATA MODEL DEVELOPED WHICH ALLOWS A SINGLE TECHNOLOGY TO REPLACE GIS AND IMAGE PROCESSING

- INTEGRATED SPATIAL QUERY LANGUAGE DEVELOPED INCLUDING IMAGE-ORIENTED QUERY FUNCTIONS

CONCLUSION

VECTOR - RASTER DATA STRUCTURE

A DILEMMA OR A FALLACY?

WHAT IS NEEDED?

1. BETTER UNDERSTANDING OF HANDLING RASTER DATA SETS FOR GIS ENVIRONMENT

- THEIR PRIMARY DATA VALUE

- RASTER DATA INTEGRITY

- PROCESSING CONSIDERATIONS WHEN MOVING FROM LOW LEVEL INFORMATION TO HIGHER LEVEL OF ABSTRACTION

SPACE REPRESENTATION - OBJECT VERSUS FIELD IS A DEEPER CHARACTERISATION OF GIS THAN DATA STRUCTURE (VECTOR VERSUS RASTER)

The Unexpected Benefits and Problems of Data Integration

Robyn Gallagher

GISolutions / Bureau of Mineral Resources, Geology & Geophysics

Abstract - A discussion of the less obvious benefits of data integration using GIS: flexibility of data storage; combination of layers; data integrity checking. Also a consideration of some of the problems which may be encountered: scale and projection; image display mechanisms; data input; 3-dimensional data sets.

Keywords - integration / raster & vector data

Introduction

There is no question about the desirability of data integration: each data set is much more valuable in combination with others than it is in isolation. However, the process involves more than just collecting together all available data sets.

Defining Data Integration

At its simplest, data integration is the ability to display different types of vector data, often with a georeferenced image as a backdrop. The images are not raster GIS or classified images: they are functions of a continuous variable such as potential field, radiometrics or satellite data; they are often multi-channelled; and pixels may be stored in a byte or floating-point number.

The ability to display data sets together leads quickly to a desire to query the vector data attributes while the data sets are being displayed. Slightly more complicated is the ability to process the image using the vector data sets: for example to interrogate pixel values or calculate image statistics within a polygon.

Finally, the combination of data via raster GIS functions provides an interesting tool for modelling and resource assessment.

Implicit in these simple-sounding operations are some assumptions:
- it must be possible to do all the work within a single system, at least from the user's point of view. If different pieces of software are being combined to give greatest functionality, this fact should be invisible.
- the user interface must be simple, with easy-to-learn menus, appropriate defaults, and fast point-and-click functions. We are after tools to increase productivity.
- the tools should be widely available, in a network of locations and ultimately in field camps.

Old Issues

Fortunately the old question of raster vs vector GIS has ceased to be an issue, with most vendors accepting the need for a mix of data structures.

Several problems arose directly from the need to display images as part of the data integration process. It must be stressed that these are multi-channelled, continuous images, often derived by gridding floating-point data. There is currently limited use for schemes which are designed to handle quantised grids.

Historically it has been considered necessary to have 32-bit (rather than 8-bit) display capabilities for coloured images: even the excellent new display algorithms which compress RGB into 8 bits lose the subtleties the geologists are searching for. Such screens are still quite rare and expensive.

Even with good display devices it has been necessary to scale or shape the image so that its pixels are byte-valued for display. This means there must be two types of links back from the display if information is not to be lost: to the image it was derived from; and to the survey data used to create the image. The former is no longer an issue with some commercial software, but it is unfortunate that survey lines do not fit very satisfactorily into the GIS data model since they are very dense and do not have 'attributes' which are constant along line segments.

Happily there is one area which has ceased to present difficulties. Most new software packages can combine images of different pixel sizes without the time-consuming resampling required by the old image-processing systems.

It should not be forgotten that even the best display devices only allow a portion of the available data to be viewed, and the place of large format hardcopy (both maps and images) has not been completely taken by interactive displays.

Current Issues

Simply putting together the data sets for a given project can require an unexpected level of expertise: the problem being not the lack of standards but their multiplicity. Images can usually be obtained in one of half a dozen formats (TIFF, PICT, ERDAS, DISIMP, I2S, LSOWG) but vector data formats are less restricted, and many of them are totally unsuited to GIS data. Even the emerging Australian standard will offer little relief in the short term: we will be dealing with historical data for a long time to come.

The issue of using data at different scales is a recognised problem, and it is necessary to be aware of the limits of use of each data set. However, the problem is quite complex: for example, when does a mine site stop being point data and become a polygon; or at what scale do some geological rock units become irrelevant on a map. We do not want to store multiple copies of the 'same' data set, but rules need to be formulated for users.

It does not take very long before the problems of combining data sets from different sources becomes apparent. One of the strengths of GIS is that it allows us to relate one data set to another and provides a mechanism for editting and cleaning up data sets if we have the expertise. Images play a very important role in this process.

Because of the size of Australia the issue of map projection is a recurring one. The cleanest solution is to have all vector data stored in lat/longs and project on-the-fly, rather than assembling a local data set in the appropriate projection. This strategy is particularly desirable if a concept like map libraries is to be implemented.However, this introduces significant complications in the interface, since the process must be made as invisible as possible to the user, and compels the user to provide any of his own data in lat/longs. If an image is used as a display backdrop it is easy to pick up the projection information from the image header, but in other situations a method must be provided to set up the necessary parameters.

Future Issues

How do we integrate 3-dimensional data: where the third dimension may be either space or time ?

The current group of general-purpose GISes is not addressing the problem, and users are forced to buy special-purpose software and hardware to handle particular applications. We have at least two routine data sets which are spatially 3-dimensional: well logs and geology. The former is handled by attaching attributes to the point location of the collar; the latter is considered only 2-dimensionally within the GIS. However, in reality the restrictions are far greater, since there are data sets such as geochemistry which at a project rather than regional level should be treated as 3-dimensional.

When the third dimension is time, the situation is little better. Data collected over time, such as changing tenement boundaries, fit badly into the topological model and can only be handled satisfactorily if the user writes his own code extensions. Similarly, any time-varying modelling processes must be managed by the modeller.

Now that the integration of raster and vector data is a reality, our demands extend to inclusion of other information.

In particular, there is a need to have easy access to reports and other textual data. While scanning and use of such data as raster images is a first step, the ideal situation is the ability to search and summarise 'intelligent' documents.

Finally, links to other processing packages such as statistical analysis and geophysical modelling will further increase productivity. It goes without saying that these links must be as unobtrusive as possible, without the need for constant data conversion and with easy transfer mechanisms between packages.

CURRENT PROGRESS IN THE MODELLING OF INTEGRATED SPATIAL DATA USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Brian G. Lees

Department of Geography, Australian National University,
Box 4 Canberra, ACT 2601. Australia

Abstract - The search for a data selection algorithm suited to disparate spatial data types has led to a detailed examination of the behaviour of inductive reasoning and neural network algorithms. Both make demands on both sample size and quality and it is clear that our current view of data standards is inadequate if these tools are to be used effectively.

Introduction

The integration of different spatial data types is taking place across a wide spectrum of applications from meteorology, oceanography, mineral exploration and vegetation mapping. The underlying problems are common to all of these fields, and although my own interests encompass all of these fields we have found that vegetation mapping provides a particularly useful test bed to examine in detail. Field checking of results is comparatively cheap and easy.

Our interest in this area of research began when the initial specifications of the NASA EOS platforms were published. It was clear that, if the program went ahead, very large amounts of remotely sensed data with very disparate characteristics would be available. Many of the proposed uses were to be supported by standard products developed over the next decade. However, given past experience, many users, particularly in Australia, would require special products. It seemed unlikely that many of these potential users of this data would be free of cost or hardware constraints and that a data selection algorithm would be useful. I envisaged this a a tool which would allow users to select the most appropriate suite of data sets for any particular problem. After some searching, we found that the characteristics of decision tree classifiers appeared to be an option which would allow a selectivity of techniques and act as an aid in defining the most appropriate techniques and data sets.

Decision Trees

Decision trees are an efficient mode of separating observations into classes, or for predicting the best response to a given situation. The structure of the decision tree classifier may be derived from an understanding of the phenomena or through rule induction techniques using both observations and known information (Pearson & Collin, 1989). As an alternative to gathering rules from experts, and interrelating them to produce an expert system, verification data can be used with artificial intelligence techniques that allow the rules to be induced from the occurrences themselves. This is done by recursive partitioning of the data sets. The selection of the most appropriate data set at each split is an optimisation procedure.
In many ways the phrase 'Artificial Intelligence' is not useful, the procedure is best described as 'inductive' learning, rather than 'deductive' reasoning.

Building and using a decision tree can be separated into three phases. In the first phase a learning set (of quadrat data, in vegetation studies) is used with a heuristic to obtain a decision tree. In the second phase, a test set is passed down through the tree and the rules simplified. In the third, the tree is used to classify any new examples. It is in the second phase that the tree structure is being used as an efficient data selection algorithm.

In our early work we used CART software (Breiman et al., 1984), in a similar configuration to that described by Walker and Moore (1988), to integrate remotely sensed data with environmental GIS data sets (Lees and Ritman, 1991). Inherent in this project was the use of the decision tree algorithm to select the optimum data set for analyses. We have since discovered that a similar approach allows us to refine the data sets by examining the effects of changing scale, pre-processing operations and resolution. In many cases we have found that simpler pre-processing operations do not significantly degrade the final outcome. This has been an important use of the algorithm.

One of the unexpected side effects of using this approach has been the removal of the shadowing effects so persistent in remote sensing data. The selection of a training set which includes cases with all aspect situations, and the inclusion of terrain variables such as slope and aspect in the GIS data, allows the system to build rules which can cope with the variations in illumination affecting the remote sensing part of the data set.

There are some major drawbacks to this approach, one is that process knowledge can only be introduced into the analysis through the initial choice of candidate correlates. In order to overcome this we moved to an interactive decision tree algorithm which allows the analyst to review the alternates at each split in the tree as it is grown and to make a choice, on the basis of process knowledge, which may not be the optimum. In other words, the analyst has the option of overriding the recursive partitioning process where the statistically optimum split does not fit our process model, and forcing the system to examine alternate splits. This requires that the analyst is both expert in the field of the analysis and statistically sophisticated.

Another of the problems we have encountered is that the Decision Tree approach assumes that the data set is spatially uniform in its data quality and data density. One of the characteristics of decision tree rule-induction is the need for a comparatively large learning sample. This is necessary so that the level of confidence at each splitting rule remains at an acceptable level. In general terms the lower limit is where n = 30. This requirement means that the number of splits is directly limited by the size of the learning sample, an important restriction. In environmental analyses using decision tree rule-induction we typically cascade through a series of deterministic variables. The order in which these are addressed in forest mapping exercises is surprisingly constant and usually takes the form;

CLIMATE

GEOLOGY

TOPOGRAPHY (Geometric)

REMOTE SENSING

TOPOGRAPHIC (Hydrological)

Under these broad headings are very specific data sets, for example the climatic indices, in their fullest form (Mackey, 1991), comprise;

Radiation Regime (kj/m2/day)

1. annual mean monthly daily radiation
2. monthly mean daily radiation of minimum month
3. monthly mean daily radiation of maximum month
4. range (3-2)
5. co-efficient of variation
6. monthly mean daily radiation of wettest quarter
7. monthly mean daily radiation of driest quarter

Thermal Regime (degrees celcius)

8. annual mean temperature
9. mean daily maximum temperature of hottest month
10. mean daily minimum temperature of coldest month
11. range (9 /10)
12. co-efficient of variation
13. mean temperature of wettest quarter
14. mean temperature of driest quarter

Moisture Regime (mm of rainfall)

15. total annual precipitation
16. wettest month precipitation
17. driest month precipitation
18. range (16-17)
19. co-efficient of variation
20. wettest quarter precipitation
21. driest quarter precipitation

Clearly all such deterministic variables cannot be included in decision tree-type analyses with even quite large learning samples. While our research has shown that decision tree rule-induction algorithms provide an excellent data selection tool identifying the most informative attributes, or indices, at each split, the limit to the number of variables which can be included in the analyses is a significant disadvantage.

Interactive decision tree algorithms provide us with the statistical power to build very precise expert systems. These two uses of decision-tree rule-induction algorithms, as data selection tools, and as prompts in the building of expert systems, mean that they will have increasingly wide use. Their use in the integration of GIS and remote sensing data is fairly new but almost certain to become conventional.

Neural Networks

We are currently investigating the use of Neural Networks in a similar configuration in the expectation that these may be more 'forgiving' of real world data variability and make more effective use of learning samples. Our initial results suggest that they do perform better, but they are considerably more difficult to use in their present state of development.

Neural Networks provide an interesting alternative to Decision trees but lack one of the most useful attributes of Decision tree analysis, the ability to simply examine the internal linkages of the classifier. It is possible to examine the internal weightings of a neural network, but only with great difficulty. They do provide a useful alternate inductive reasoning system which will also take its place in our growing armoury of non-parametric statistical tools. Like Decision trees they demand accurate and precise learning samples in order to give reasonable results.

Data Quality

Our research into data modelling using integrated GIS and remote sensing data sets is forcing us to consider, very carefully, the characteristics of the data we are using. The development of new techniques for spatial analysis, particularly those based on Artificial Intelligence inductive reasoning systems, has meant a complete re-assessment of the data standards required.

The most obvious requirement for the integration of spatial data is accurate spatial co-registration. This is a considerably more complex problem than might, at first, appear. Despite careful selection of control points, RMS errors during co-registration of raster data can rarely be reduced much below 1.5 times the relevant cell dimensions. This fundamental process, the resampling of all the data to a common co-ordinate system is often seen as the major source of error in this type of analysis. This is probably not true in most cases. The individual data layers have varying degrees of spatial error. Data sets where pre-processing has been carried out and the original data is no longer available are the most problematic. The classic case of this type is thematic soils data. The variation down a soil catena (within class variation) may often be more significant than the difference between two soil classes depicted on a soils map (between class variation). Conceptually, soils do not lend themselves to being depicted as discrete, internally homogeneous, polygons. Similarly, rainfall is not well represented by the traditional Theissen Polygon structure and is better seen as a continuously varying attribute distributed in the landscape as a function of terrain, elevation and land cover characteristics. In order to minimise these sources of inaccuracy the data structure used for each attribute should be chosen to reflect the natural spatial characteristics of that attribute. This is, of course, true for all spatial data and is not a problem unique to the analysis of integrated data.

Conclusion

Integration of disparate spatial data requires non-parametric procedures that make few, if any, assumptions about the structure and characteristics of the data. In parametric statistics the assumptions about data distribution mean that comparatively small samples can be used. Where such assumptions cannot be made, larger samples are required. Decision trees, in particular, make significant demands on sample size even where there are only minimal numbers of deterministic variables. Moving to neural networks as an alternative approach, with many of the attractive characteristics of decision trees, means that the simplicity of interpretation of decision trees is lost. However they are more efficient users of sample data and appear to offer a more practical, if considerably more complex, set of tools for integrating spatial data.

References

Breiman, L.H., Friedman, R., Olshen, A. and C.J. Stone. 1984. Classification and regression trees. Wadsworth International, Belmont, California, 358p.

Lees, B.L. and Ritman, K. 1991. Decision-Tree and Rule-induction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environments. Environmental Management 15/5, 823-832.

Mackey, B.G. 1991. The spatial extension of vegetation site data- a case study in the rainforests of the wet tropics of Queensland, Australia. unpublished PhD thesis, Australian National University, Canberra, 316p.

Pearson, R.A. and G.J. Collin. 1989. Vector evaluation of decision tree structures. Tech.Rep. CS89/18. Dept. Comp. Sci., U.C., U.N.S.W., Australian Defence Force Academy, Canberra.

Walker, P.A. and D.M. Moore. 1988. SIMPLE - An inductive modelling tool for spatially distributed data. International Journal for Geographic Information Systems. 2: 347-364.

GEOGRAPHIC INFORMATION SYSTEMS,

CARTOGRAPHIC & GEOSCIENCE DATA STANDARDS

BMR WORKSHOP 18 - 20 MARCH 1992

PANEL PRESENTATION:

'RELIABILITY & INTEGRITY OF SPATIAL DATA'

THURSDAY 19 MARCH 1992 11.00 - 12.30

Peter James
Manager, Computer Services
Department of Mines
Hobart
Tasmania

This Paper discusses 'Reliability & Integrity' from a technical point of view and a management perspective.

It only touches on single topics within each aspect, namely digitising standards and data licence agreements.

Technical Perspective

The SDTS adopts a 'Truth in Labelling' approach to data quality which requires users to report what is known about the quality of the data being transferred.

The report has five modules: Lineage; Position Accuracy; Attribute Accuracy; Logical Consistency; Completeness.

Details of these modules can be found in any description of SDTS.

The advantage of this descriptive 'Truth in Labelling' quality standard is that users can determine the data's 'Fitness for Use'.

But, how do we, as data capturers, ensure consistency in our data, and in particular, in its positional accuracy?

What I would like to discuss now is the data capture aspects of positional accuracy.

After doing a quick survey of organisations present at the workshop the following results were compiled:

1. A large number of organisations digitise spatial data;
2. About six of these organisations use some type of digitising standard;
3. Only two of these organisations have common digitising standards.

The point I am trying to make here is that the different digitising techniques are resulting in the creation of differing 'quality' data sets.

This difference is highlighted when bringing together common data sets, such as the 1:250 000 geology, from different agencies. These sets will not marry correctly and will contain areas of differing quality or resolution.

Inconsistent results may be produced when this differing quality data is amalgamated with constant quality data. The real fear is that this may occur without the user's knowledge.

This should not be the case. We should be working together to minimise or eliminate these differences. Common digitising standards and techniques are essential to achieving this.

The sorts of things that need agreement include:

- the distance between points on lines at particular scales;
  - a small distance adds unnecessary storage overheads and adds nothing to the information content;
  - too large a distance reduces the data quality and therefore its usefulness;
- methods for verifying digitising accuracy and completeness, e.g. overlaying transparent plots on the original and comparing;
- allowable map sheet registration error based on a minimum of say four registration points;
- fuzzy tolerance, or the distance points are allowed to move when processed;
- the different feature layers for each theme.

It may prove simpler and more expedient to agree on a range of scales to which the above list can be applied.

Some digitising techniques that should be common:

- Prepare manuscripts before digitising:
  - at least four registration marks placed accurately;
  - use a stable base;
  - use complete and up-to-date data;
  - edge match adjoining sheets.
- Remove any unnecessary nodes from arcs. This helps reduce the complexity and size of data sets.
- Ensure all features are digitised completely, e.g. roads under bridges should not be broken.
- Ensure continuity across map sheets by placing adjoining map sheets on the digitiser.
- Use snapping techniques to ensure polygon closure and coincident features.
- Avoid complex features such as curves, splines and circles as they are difficult to transfer, or transform them into straight line segments.

A key component of effective data sharing is high quality digitising at the originating site and a proper concern for the potential that this data may be put to other uses than those specified.

It is hoped that this workshop provides the forum for the geoscience community to come together and collaborate on the creation of a set of digitising standards and techniques.

Managerial Perspective

From the management perspective I would like to discuss the Tasmanian Government standard data licensing agreement currently in use and point out some aspects of the agreement which address the reliability and integrity of spatial data. A copy of the agreement is included in the proceedings. This agreement is completed to ensure that the Crown's interest in the data is protected, and the rights of the user are clearly defined.

Format of the data licensing agreement

The basic data licensing agreement comprises a set of standard terms and conditions, and a schedule which details those aspects of the agreement specific to the particular case. Under certain circumstances, the standard document may also be varied to suit specific requirements.

The basic form of the agreement covers a right to use the data set available at the time the agreement is completed, and restricts its use for internal purposes only. Should the user require an extension on this facility, three special cases are covered by completion of attachments to the main agreement. These attachments may be completed at any time after entering into the initial agreement. The three cases are as follows:

Updates: Completion of this attachment will extend the agreement to include regular updates on the original data set;

Commercial Exploitation of Outputs: A user who wishes to produce hard copy paper printouts, any film based product, or any product visible to the human eye beyond those required for their own use, will need to extend the basic agreement by completing this attachment;

On-selling of Raw Data: Where a user wishes to provide a digital copy of all or part of the data to a third party, the basic agreement will need to be extended by completing this attachment, and this in turn requires the third party to enter into a similar agreement with the Licensee.

What this agreement allows us to do is control the use of data sets to ensure they are not used in situations where their information content can be abused or misused, intentionally or unintentionally.

It is the Schedule that contains full details of each agreement. These descriptions must be precise and cover all matters that have been agreed, not an easy task sometimes.

Some of the Schedule items include:

Item 1: The Data
Full description of data including format and storage media if applicable.

Item 2: Approved Purpose
The Licensee's use of the data will be restricted to the internal conduct of the Licensee's business. The type of internal use needs to be detailed with specific reference to the data. Refer to Clause 2 of the agreement.

Item 8: Performance Criteria

This is where such things as lineage, positional accuracy, attribute accuracy, logical consistency and completeness are reported. Refer to Clause 10 of the agreement.

Item 13: Modifications

Any permissible modifications to the data are detailed here. Refer to Clause 2.1.6 of the agreement.

This agreement was drafted by the Crown Solicitor's Office in consultation with various Government GIS users. Although it appears somewhat complex and daunting it more than adequately protects the State, the State's data and the client's interests.

A simplified agreement exists for use between Tasmanian Government Agencies in the form of a Memorandum of Understanding (MOU).

WORKSHOP.GIS 25 March 1992

Accuracy and Precision

Dave Richards

With the increasing application of Geographic Information Systems (GIS), there is a growing requirement for information which includes location information. The information needs, however, to be reliable, and location information requires a statement of both its accuracy and its precision. There is often no clear distinction made about the difference between accuracy and precision, and it is apparent that many people who should know better are unaware that there is a difference. It is worth elaborating on the difference.

The accuracy of a location is a measure of how close the stated position is to the real position. Precision is a measure of the degree of approximation we use in recording the location. Information about accuracy and precision is particularly important for GIS systems which allow scale changes such that inaccuracy and imprecision may become appreciable. Ideally, accuracy and precision should be commensurate.

As an example, the position of a mineral occurrence may be known from historic records with an accuracy of about lOOOm and might be plotted by the mapper in what he/she considers to be the best position on a 1:100 000 scale map. There would be little justification for quoting the AMG location of this point with a precision of 50m, as might be provided by an electronic digitisation procedure applied to the map. In this situation accuracy and precision are obviously not commensurate, and this needs to be flagged.

Possible situations arising from lack of accuracy/precision information include:

- difficulty in reconciling two versions of the position of the same point, and
- difficulty in determining whether or not two points from different sources are in fact the same point.

Attribution (ie noting the source of the information) is undoubtedly a key factor in resolving some of these dilemmas, and it is pleasing to see this feature beginning to be addressed in some datasets.

How should we specify accuracy and precision quantitatively?

Modern global positioning system (GPS) technology permits levels of precision to be gauged quantitatively and these will often be better than can be obtained by more traditional mapping methods. Accuracy is not really an issue for GPS since you have to be at the spot to take the reading. Nonetheless you may not be at the spot you think you are, and although the GPS may tell you precisely where you are, it will not tell you that you are at the wrong copper occurrence.

There is a considerable heritage of valuable historic information which we can capture and use if we know about its accuracy and precision.

A point for discussion might be how we can gauge the accuracy of historic information.

Although we tend to think of mineral occurrences as points, they are in fact finite areas rather than points, and perhaps ultimately we need to consider this finite size in determining the accuracy and precision of location.

Do all points ultimately become polygons?

GIS CARTOGRAPHIC AND GEOSCIENCE DATA STANDARDS

BMR WORKSHOP - 18-20 MARCH 1992

PANEL PRESENTATION - RELIABILITY AND INTEGRITY OF SPATIAL DATA

PAPER PRESENTED BY COLONEL S W LEMON RASVY

I have chosen to talk today on the subject of topographic map data as an important sub-set of spatial data. It is important because many GIS developed over the next few years will rely heavily on digitital data derived from this source. In order to make a judgement on the quality of the GIS data it therefore makes sense to understand the reliability and integrity of the topographic map product itself as well as the impact on data quality that will result from the process of converting this source into a digital data base for a GIS.

The topographic line map is a product which at medium and small scales provides a uniform and consistent data source with coverage over large areas, indeed world wide at 1:1 000 000 scale. Consistency is achieved through strict adherance to well-defined and rigorous product specifications which by their nature restrict the flexibility of the end product. Uniformity results from a common specification throughout a complete map series. Reliability and integrity are qualities inherent not just in the map data itself but in the limitations placed upon the user of the end product. Extreme care must be exercised in the ad-hoc transfer of data derived from a map product data base into the less rigorous environment of a GIS. On the other hand a product such as the DCW (VPF) or the AGDB (AS SDTS) may call for more explicit information on data quality which will enhance the reliability and integrity of the spatial data.

Data reliability can be viewed as being directly related to data integrity. Their sum equates to overall data quality. In turn data quality is directly related to map accuracy which is expressed in both quantitative and qualitative terms. It is useful at this stage to refer to the seven aspects of data quality which the data exchange standard DIGEST employs in describing data integrity and reliability. These are as shown here: source, positional accuracy, attribute accuracy, attribute completeness, currency, logical consistency and feature completeness. Interestingly, only one descriptor of data quality is truly quantitative - that of positional accuracy. This is a particularly important quality descriptor if you are relying on topographic data to provide the spatial reference framework of your GIS. Positional accuracy is expressed as the upper bound on the deviation from true position. Its specification is not scale related and includes all errors introduced by source documents, data capture and processing.

Map accuracy statements are also quoted as an upper bound and are inclusive of source documents, data capture and processing but they <u>are</u> scale related. In some ways it is unfortunate that cartographers are so conservative in their estimation of positional accuracies of map data. Data integrity is sacrificed for the sake of uniformity. "90% of well defined features measured on the map will fall within 25 metres relative to their true planimetric positions referred to the graticule of the map" is in many instances understating the quality of data available from the topo map source. A recent analysis of the error budget of 1:50 000 scale maps produced by RASvy reveal that worst case planimtric accuracies for digital map data can be as good as 7.3 m / 17.0 m CE at 90% confidence level (depending on the method of capture) and as poor as 12.1 m / 19.5 m on the printed map product. Vertical accuracies at the digital data base for relief range from 4.0 m / 5.7m LE at 90% confidence level for contours to 5.1 m / 6.7 m for DTM. Relative accuracies within a single block of controlled photography arė of course substantially better.

In summary, topographic map data provides a valuable and very exploitable source of data for a GIS. Whilst there is a risk of quality degredation when importing map data into a GIS data base the use of topographic spatial data products such as DCW and AGDB can enhance or at least preserve the inherent quality of the source. There is a trade-off between data quality, integrity and reliabilty but the choice is there (or almost) for the discerning buyer.

DIGEST DATA QUALITY STATEMENTS

SOURCE	Origin or derivation of feature (entity), primitive (faces, nodes & edges) or attributes (properties of features), including processing techniques applied to data as well as source data.
POSITIONAL ACCURACY	Upper bound on deviation from true position, specified not scale related & includes all errors introduced by source documents, data capture and processing.
ATTRIBUTE ACCURACY	Accuracy or reliability of data for an attribute.
ATTRIBUTE COMPLETENESS	Degree to which all attributes of a feature have been included for that feature.
CURRENCY	Date of introduction/modification of data in data base.
LOGICAL CONSISTENCY	Fidelity of encoded relationships (topology).
FEATURE COMPLETENESS	Degree to which all features of a type for the area have been included.

RELATIONAL DATABASES FOR CONTINENT-WIDE DATA

Roderick J. Ryburn

Information Systems Branch, Bureau of Mineral Resources
GPO Box 386, Canberra, ACT 2601

Abstract - With interest focussed on geographic information systems (GIS) we must not loose sight of the integrating role of the corporate database system. GIS tends to emphasise 2.5 dimensions and specific map areas, whereas much geoscientific data has ***n*** dimensions and requires data models with continental perspectives. Automated data presentation and analysis require standards best achieved in an integrated system, and a relational database management system is currently the best tool. All types of data may ultimately be stored in an integrated corporate database system.

Keywords - Earth science information/Databases/Computer applications

Introduction

Increasing interest in geographic information systems by the mining and petroleum exploration industries raises a number of strategic questions that are far from being resolved in the minds of many observers. One of the most interesting current issues is to what extent GIS should be regarded as an end in itself, or merely an adjuct to other information systems. The answer, I feel, depends on the size of your organisation and the type and scope of your data.

I aim to convince you that in an organisation like BMR, that deals in diverse data covering the whole continent, the corporate database is the central issue, and that GIS and other data analysis and presentation systems should be regarded as data presentation and analysis tools, and as 'front ends' to the database. Furthermore, I hope to show that a relational database management system (RDBMS) is currently the best tool for managing the wide-ranging attribute data sets that geoscientific studies require.

Has GIS been Oversold?

As with any new technology we must be careful not to overstate the usefulness of GIS to the extent that users become disillusioned after experiencing the reality. Remote sensing, for example, has previously suffered from exaggerated claims as to the applicability of that technology. Mostly as a counterbalance to overly enthusiastic accounts what GIS can do for us, I wish to point out some current limitations to GIS -

» **GIS 2-3 Dimensional** - GIS tends to constrain the user to '2.5' dimensional data models - rarely 3. The user tends to think in terms of data layers and ignore complexities that go beyond this concept. Real data typically encompass more than 3 dimensions. Geochemical data, for example, have so many independent variables that they need simplifying projections for graphical portrayal. Time is also often a factor.

» **Non-Spatial Data** - Much data stored in databases are non-spatial, or have extensive non-spatial ramifications. Statistical analysis commonly involves non-spatial data. The GDA geochemical data analysis system used by BMR is a good example of a non-spatial application (Sheraton & Simons, 1992).

» **Coincident Data** - A series of observations that occupy a single point on the map can be difficult to handle in GIS - they require the services of a proper underlying DBMS. A good example is provided by the propensity of geologists to take multiple observations and samples at a single outcrop.

» **Few Innovative Geoscience GIS Applications** - To date, most geoscience GIS applications mimic operations previously done on paper. It seems to be difficult to find truly innovative geoscience applications - at least to the extent that this is possible in agriculture, forestry, and the environmental and life sciences. One would like to be able to ask questions such as "where are the best petroleum targets?", but GIS is currently suited to problems relating to the surface of the earth rather than its interior. Geological surface expression is often diffuse, imprecisely located, or displaced with respect to deeper geology.

» **GIS Databases Often Spatially Limited** - Databases generated specifically for GIS are commonly restricted to areas of immediate interest and the data model used is commonly inadequate for larger areas. For example, the letter symbols used for stratigraphic units on geological maps may provide unique handles in limited areas, but not on a continent-wide basis. The effort involved in assembling the data is often such that only the map areas of immediate interest are covered, leaving vast gaps on a continental scale.

» **GIS as a Sledgehammer** - GIS systems such as Arc/Info usually require big investments in time and money before satisfactory results can be obtained. Potential users may merely wish to plot diagrams for publication, not to make full use of the GIS's ability to interactively query and update the underlying database. All that may be required is a simple PC package that can select data from the database and plot the results on a map projection. The learning curve can be half an hour rather than weeks or months. Horses for courses.

I think the answer to the question in the heading of this section is that GIS systems may have been slightly oversold, but that the geoscience community is a canny lot who know that GIS stands for 'Get Involved Slowly'. In BMR there are currently a number of spatial information systems, including Spanz, Arc/Info, Intergraph, ER Mapper, I^{2}S (image processing), and Petroseis (basin analysis). There is considerable willingness to experiment with different systems. No one system has yet gained total ascendancy.

The Corporate Database

What the above is leading up to is the need for a comprehensive corporate database system that is independent of specific applications such as those provided by spatial information systems. We need to be sure that hard-won geoscientific data is stored in an accessible and standard system that is 'on-line' to a variety of application systems - not just GIS. We must ensure that the primary data is preserved in its original form, and is not thrown out in preference to interpreted, smoothed, contoured, or sampled subsets. Methods of massaging data can go in and out of fashion quite fast. We also need to avoid the situation of multiple versions of the same database, and the inevitable confusion as to which version is the most up-to-date. A single corporate database system solves this problem, as well as most security and archival worries.

Of course, there are limits to the volume of data that can currently be stored in on-line systems. Satellite images and air-magnetic data, for example, are a bit too voluminous to be handled in their original form. Nevertheless, with optical storage methods providing a further archival layer to the total computer memory system, even these types of data will soon be able to be stored in the corporate database system and made directly accessible. Given sufficient resources this can be done right now. Imaged documents and publications could also be included in an integrated corporate database.

All this is not to say that GIS is not an important client of the corporate database system, and in BMR we have taken steps to adapt the corporate database to the requirements of GIS. BMR's National Geoscience Mapping Accord (NGMA) databases are being linked to a central 'SITES' table that standardises methods of recording point location data (Blewett & Ryburn, 1992). The advent of GIS in BMR has highlighted the need for better point location data in the corporate database. Also, derived data sets have a place in the corporate database as long as they do not displace or disguise the original data.

Devolution and the 'Corporate Data Model'

An issue that most organisations like BMR must eventually face is to what extent individual project areas be encouraged to manage their own database requirements - *versus* the need to maintain corporate databases and 'enterprise-wide data models'. With so many research projects in BMR, and such a high proportion of 'rugged individuals' on the staff, the centrifugal forces are inevitably quite strong. As a result, there are many databases in BMR and a diverse set of hardware and software systems.

However, with the limited resources available under the current economic climate, there is mounting pressure to make do with the corporate facilities and an increasing appreciation of the economies inherent in a centralised system. For example, security and backup are best handled centrally. More important though, is the growing realisation that automation demands standards, and that corporate standards are best catered for in an integrated system. Such a system is necessary to share data in a timely manner and to make the best use of look-up tables and authority databases.

This does not mean that BMR should blindly adhere to a highly structured business concept of what constitutes an 'enterprise-wide data model'. To attempt an allencompassing plan in such a diverse and rapidly changing scene is to doom oneself to disappointment. The model would be well and truly out of date by the time it was finished, and it would meet with a prompt mutiny anyway. The approach we tend to follow in BMR is an evolutionary one. We try to establish authority databases that satisfy the most critical need for national and corporate standards, and to let other databases tap into these as required. In general, these standard databases are maintained by the Programs and Projects best qualified to do so. Custodianship within BMR is distributed, not centralised.

Some Standard Databases

The importance of standard databases and authority tables is best explained with the help of some examples from BMR's corporate database system. Single tables are generally called 'authority tables' if they contain many records, or 'look-up tables' if they are small. Standard or authority databases have a number of related tables, although they often appear to the user as a single screen form and the terms 'table' and 'database' tend to be used somewhat interchangeably.

A good example of an authority table is the **HMAPS** table, which contains the number, name, 1:250 000 sheet ID, minimum AMG reference, and minimum latitude and longitude for all 3000 1:100 000 maps covering Australia. Databases that refer to 1:100 000 sheet areas use this table to validate map numbers and names, and to check that entered latitudes and longitudes fall within the nominated sheet area. HMAPS can also be used to provide a map name given a latitude and longitude.

BMR MAPPING DATABASE - 1:100 000 MAPS

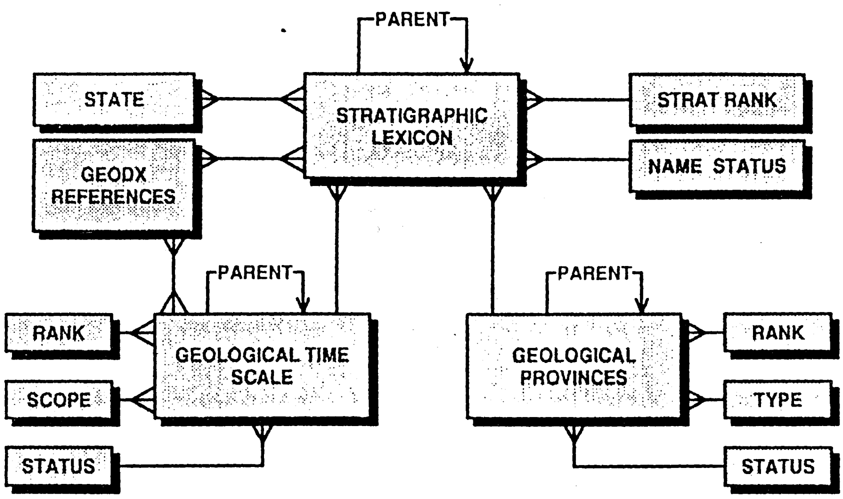
100K No.	1M ID	250K No.	100K Name	NW Corner Lat	NW Corner Long	AMG Ref. SW Cnr m East	AMG Ref. SW Cnr m North
7050	SF54	14	SPRINGVALE	23.5	140.5	449143	7345674
7051	SF54	14	ELIZABETH SPRINGS	23	140.5	448949	7401030
7052	SF54	10	LUCKNOW	22.5	140.5	448758	7456384
7053	SF54	10	TOOLEBUC	22	140.5	448571	7511733
7054	SF54	6	SELWYN	21.5	140.5	448388	7567080
7055	SF54	6	MOUNT ANGELAY	21	140.5	448209	7622423
7056	SF54	2	CLONCURRY	20.5	140.5	448034	7677763
7057	SF54	2	CLONAGH	20	140.5	447863	7733099
7058	SE54	14	ALCALA	19.5	140.5	447696	7788433
7059	SE54	14	CANOBIE	19	140.5	447533	7843763
7060	SE54	10	DONORS HILL	18.5	140.5	447374	7899090
7061	SE54	10	MILGARRA	18	140.5	447219	7954414
7062	SE54	6	MAGOWRA	17.5	140.5	447067	8009736
7063	SE54	6	KARUMBA	17	140.5	446920	8065054
7121	SJ54	11	NELSON	38	141	500000	5738691
7122	SJ54	7	CASTERTON	37.5	141	500000	5794170
7123	SJ54	7	EDENHOPE	37	141	500000	5849644

v Char Mode: Replace Page 1 Count: 17

Press Ctrl-Break to exit the Terminal option.

A more complex example of an authority database is the recently developed **Stratigraphic Authority Database**, comprising an **Australian Stratigraphic Lexicon**, a **Geological Time Scale**, and a table of the **Geological Provinces of Australia.**

Because the Stratigraphic Lexicon depends on both the Time Scale and Provinces Tables, the various elements of the database had to be assembled at the same time. The Stratigraphic Lexicon is under the custodianship of BMR's Stratigraphic Index Group, while the Time Scale is looked after by Phanerozoic Time Scale Project, and the Provinces Table by David Palfreyman - based on his account of the tectonic provinces of Australia (Palfreyman, 1981). The structure of this database is as follows -



The Stratigraphic Lexicon is designed so that other databases can validate stratigraphic names at whatever rank is most appropriate. It allows other databases to have a single field that can accept names at the member, formation or group rank (amongst others). Having validated a name the Lexicon can supply the rank of the unit, its map symbol, age range, parent unit (e.g. group name, if a formation), location of type area and the overlying and underlying units there and its relationship to them. Users view the Stratigraphic Lexicon as a single screen form -

```
------------------- BMR STRATIGRAPHIC LEXICON -------------------
         No.   Name
Unit    *4132  Mount Norna Quartzite

State   >QLD  Rank 4>Formation   Status 1>defined     GEODX ID>13982

Age from 144 >Proterozoic               to     >

Geological Province 54 >Mount Isa Inlier
                                                        (240 chars--->)
Comments
(including synonymy & history)
---------------------------- TYPE AREA DATA ----------------------------

State   >QLD  Lat 20.9239 S  Long 140.731 E  Map Symbol P_on

Parent U.2327 >Soldiers Cap Group               Boundary Relations
Ovrlng U.2486 >Toole Creek Volcanics            3 >conformity
Undlng U.1525 >Llewellyn Creek Formation        3 >conformity

* Automatically generated - can only be entered in query mode
> Look-up field - Press 'List Field Values' to see associated form
Press 'List Field Values' for BMR Geological Time Scale Form
        Char Mode: Replace  Page 1                  Count: *1
Press Ctrl-Break to exit the Terminal option.
```

The Lexicon also allows terms of differing status. Terms with an informal status may be required where pre-existing databases are linked to the Lexicon. For example, if BMR's PEDIN Petroleum Database is linked to the Lexicon (and I hope it will be) stratigraphic names that are currently in PEDIN, but not in the Lexicon, can be added to the Lexicon as informal terms until someone has the time to establish their credentials.

In addition to the Lexicon itself, a comprehensive bibliography is provided by BMR's GEODX database. Given the number or name of a unit in the Lexicon one can then query this database with the help of a single References Form -

```
----------------- BMR STRATIGRAPHIC LEXICON - REFERENCES FORM -----------------
*GEODX ID 13982    Strat >Mount Norna Quartzite          Ref.ID 88/26192
 (id_     13982    Name   Mount Norna Quartzite          (cd_   88/26384
 strat    13982           Mount Norna Quartzite          art-   89/26476
 name)    13982           Mount Norna Quartzite          icle)  89/26488
          13982           Mount Norna Quartzite                 98/26980
          13982           Mount Norna Quartzite                 8225
 * Enter ID or Name, Execute Query, put cursor on row required, press Commit

 Ref. ID     89/26488   Author Ryburn R.J.                      Order 1
             89/26488          Wilson I.H.                            2
             89/26488          Grimes K.G.                            3
             89/26488          Hill R.M.                              4

 Ref. ID     89/26488     Year 1988

 Title       Cloncurry, Queensland, 1:100 000 geological map commentary.
 Publication Bureau Mineral Resources, Australia & Geological Survey of Qld
 Vol or Part 1v                         Page

        ^  Char Mode: Replace  Page 1                   Count: *22
Press Ctrl-Break to exit the Terminal option.
```

The Time Scale and Provinces tables are similar in concept to the Lexicon, and use the same recursive hierarchical structure, with several small associated look-up tables. The Stratigraphic Authority Database provides a good illustration of the effectiveness of standard databases. Not only do they allow you to validate data, but they also provide a comprehensive set of associated data. When designing a new database, one field can give access to a whole set of related data that is (hopefully) up-to-date and maintained by well qualified people. Relational database management systems permit the on-line access of a standard database from within another database or from other computers on the network.

If data is to be combined from several databases on one map (for example), then it is important that the databases all speak the same language. The examples given above show how data can be constrained to fit standards, so that data mismatches do not occur. BMR has a national as well as corporate role to perform in providing some of these standard databases, and we hope to throw some of them open to dial-up access very soon. On-line validation over long-distance links is also possible.

Relational *versus* Object-Oriented Database Management Systems

The current debate as to the merits of Relational *versus* Object Oriented schools of database management (e.g. Smith, 1991; Edelstein, 1991) cuts right to the core of the relationship between GIS and the corporate database. Object-oriented DBMSs, which stem from the concepts embodied in object-oriented programming languages like Smalltalk, appear to be well suited to the requirements of GIS and topological data. They deliver the performance required for GIS and CAD (Computer Aided Design), as well as the much-desired extensible complex data models. The GIS world is moving rapidly to embrace object-oriented database systems, such that it is unlikely that many alternatives will survive in a few years.

Unfortunately, it is not yet clear that object-oriented DBMSs, at least as they currently exist, are well suited to corporate database systems. The data navigation paths in object-oriented systems appear to be inherently predetermined and hierarchical, precluding arbitrary joins and ad-hoc queries. This may be of little consequence in GIS applications, but it raises questions for corporate databases that make use of standard tables and databases such as those described above. The philosophy behind a standard database like the Stratigraphic Authority Database is to provide properly maintained data standards that can be tapped into by other databases as required. Another example is BMR's PEDIN petroleum database, which is used to supply well information and validation to a number of other databases in BMR. I suppose these databases could become objects common to other databases, but the mechanisms are not well known.

Relational DBMSs (see Lenz & Pain, this volume) are currently the conventional wisdom for corporate databases as they embrace 30 years of experience in database management and have a vast investment in data and data models riding with them. The problems of data integrity, security, concurrent access, transaction processing, crash recovery, programmable interfaces, distributed databases and database evolution, to name a few, are all addressed by current relational systems, and it will probably take object-oriented systems some years to approach the same level of completeness and reliability. For the time being, RDBMSs will continue to dominate the corporate database scene, but one would be well advised to keep an eye on commercial offerings in the object-oriented world.

According to Edelstein (1991) we are doomed to two competing philosophies for quite some time. A likely outcome for GISs is they will all offer gateways to relational databases, but will become internally reliant on object-oriented database systems. This scenario is far from satisfactory, and one hopes the ultimate target will be a unified database management philosophy for all types of data and all applications.

Whither Databases?

Although modern computer-based database management systems may seem sophisticated and 'high-tech', they still have a long way to go to match concepts of what a DBMS should be - more like a steam roller than a Jumbo Jet. They are slow, limited as to data type and volume, unintelligent, unnecessarily complex and non-standard.

However, there seems to be a fair deal of agreement as to the direction of future developments. We can see the destination clearly but the means of getting there is not quite so obvious. Some likely developments are as follows -

» **Open Networks -** Data networks must be refined to match the simplicity and standardisation of the telephone system. The goal is to be able to plug your PC into a standard network system that gives access to data at the project, program, organization, city, state, nation and world levels, and at band widths sufficient to allow rapid transmission of images - even moving images. Telecom's ISDN system is just beginning to realise this ideal.

» **Standard Directories** - DBMSs need a standard directory system for navigation purposes. You should be able to log into an unfamiliar system and easily find your own way around - at least in those areas you are permitted. An ANSI standard for just such a system is being developed. Another way of stating this is that databases should be self advertising.

» **Merged Relational and Object-Oriented Technologies** - Relational DBMSs are business oriented and more suited to numbers and text than images and linework. Object-oriented databases are well suited to images and topological data, but lack the theoretical foundations and refinements of relational systems. They also constrain the user to predetermined data navigation paths. All types of data should reside in one logical DBMS, though a variety of techniques may be used behind the scenes.

» **Standard Security and Charging Methods** - There is a need for standard layered security methods in accessing data, and built-in methods of account keeping. Oracle, for example, cannot define a single view giving different access rights for select, insert, update and delete. Most DBMSs require a great deal of work to set up on-line automatic charging.

» **Open Separation of DBMSs and Applications** - A clear trend has emerged for separation of applications and database systems - the so-called 'client-server' concept. Open standards are needed for communication protocols between database clients and servers - an eighth OSI layer. Applications should be able to talk over a network to any DBMS using open protocols. GIS systems, for example, should be able to work directly with any accessible remote database.

Even if these objectives are realised there will undoubtedly be further targets at which to aim. Verbal querying - like HAL in '2001' - combined with browsing and heuristic query methods are distinct possibilities. Eventually the day may come when all types of data may be easily and transparently obtained from an integrated national database sytem - without the need for 'Big Brother'.

References

Blewett, R.S. & Ryburn, R.J., (this volume). Recent developments in field geoscience databases in BMR. *In* Williams, P.R. (Compiler). Proceedings of the Workshop on Geographic Information Systems, Cartographic and Geoscience Data Standards, Bureau of Mineral Resources, Geology and Geophysics, Canberra. *Bureau of Mineral Resources, Australia, Record* 1992/x.

Edelstein, H., 1991. Relational vs. object-oriented. DBMS, November 1991, 68-79.

Lenz, S.L. & Pain C., (this volume). Categorising descriptive data with reference to the regolith and environmental geoscience databases. *In* Williams, P.R. (Compiler). Proceedings of the Workshop on Geographic Information Systems, Cartographic and Geoscience Data Standards, Bureau of Mineral Resources, Geology and Geophysics, Canberra. *Bureau of Mineral Resources, Australia, Record* 1992/x.

Palfreyman, W.D., 1984. Guide to the geology of Australia. *Bureau of Mineral Resources, Australia, Bulletin* 181, 111p.

Sheraton, L.W. & Simons, L., 1992. Geochemical Data Analysis system (GDA) reference manual. *Bureau of Mineral Resources, Australia, Record* 1992/1, 169p.

Smith, J.L, 1991. Database management - vision for the future. *Proceedings of the Conference on Land Information Management, University of New South Wales, Sydney, July 1991.*

THE SADME DATA MODEL FOR A COMPREHENSIVE GEOSCIENCE INFORMATION SYSTEM

Antonio P. Belperio

Geological Survey of South Australia
PO Box 151 Eastwood SA 5063

Abstract - A State Geological Survey is a complex earth science information system that manages very large volumes of geoscientific data for a variety of purposes. GIS and RDBMS software are now essential tools for data management, analysis, integration and presentation. Successful integration of complex data sets requires a detailed understanding of the data, its use and custodianship. Data modelling forms the basis for the design and construction of databases and for rationalising data flow and data management. For a complex information system with many categories of inter-related data, both detailed and high-level data modelling are necessary to provide the overview of the information system and for appropriate modular system development.

Key words - Geographic Information System/Data Modelling/Earth Science Information

Introduction

The South Australian Department of Mines and Energy (SADME), like its interstate and Federal counterparts, is an information utility that specialises in earth science data. The Geoscientific Information System is driven by SADME scientists as they collect, capture, store, analyse and present geoscientific data, interpretations and recommendations. Because most of the data are inherently geographic, this geoscientific information system is essentially a GIS (Fig. 1), albeit one that has traditionally operated largely in manual mode with spatial locations stored on a variety of maps. Fundamental to most of SADME's goals is the efficient, flexible, rapid and accurate collection, management, interpretation and dissemination of geoscientific information. There is no question, therefore, that computerised GIS has a major role to play in a Geological Survey. Because of the breadth, depth and complexity of data, the efficient use of a computerised GIS requires particular attention to the inter-relationships of the data as well as to procedures, practices and flow of data within the organisation.

In ongoing moves to improve accessibility and deliverability of data, SADME is progressing towards a comprehensive digital geoscientific information system know as GEOSIS (Belperio et al., 1991; Belperio & Parker, 1991). It aims to provide a unified system for corporate data capture, storage, retrieval, analysis and presentation. The GEOSIS model proposes distributed processing and data management across the Department's functional work groups. A number of discrete digital databases will form the core of the GEOSIS system linked though a common user interface. The major work groups will act as the focus for database management (custodians of data) of selected data that together comprise the total information system. In particular it is proposed to maintain a strategic separation of data and applications through the use of client-server and RDBMS models, an open systems environment and a consistent application development environment. The common spatial and attribute data bases, together with GIS functionality, will allow geographic overlay and interactive analysis of a wide selection of data types (Fig. 2). At this stage, it is proposed to store spatial data in a GIS and attribute data in a RDBMS, although the possibility of storing both spatial and attribute data together on the RDBMS may be further evaluated.

The critical benefit of separating data from applications is that it will preserve and enhance the value of data archived in the Department. The economic lifespan of computer equipment is just 2-5 years, of software from 5-10 years, but the geoscientific data stored by a Geological Survey will continue to be of value indefinitely with proper management.

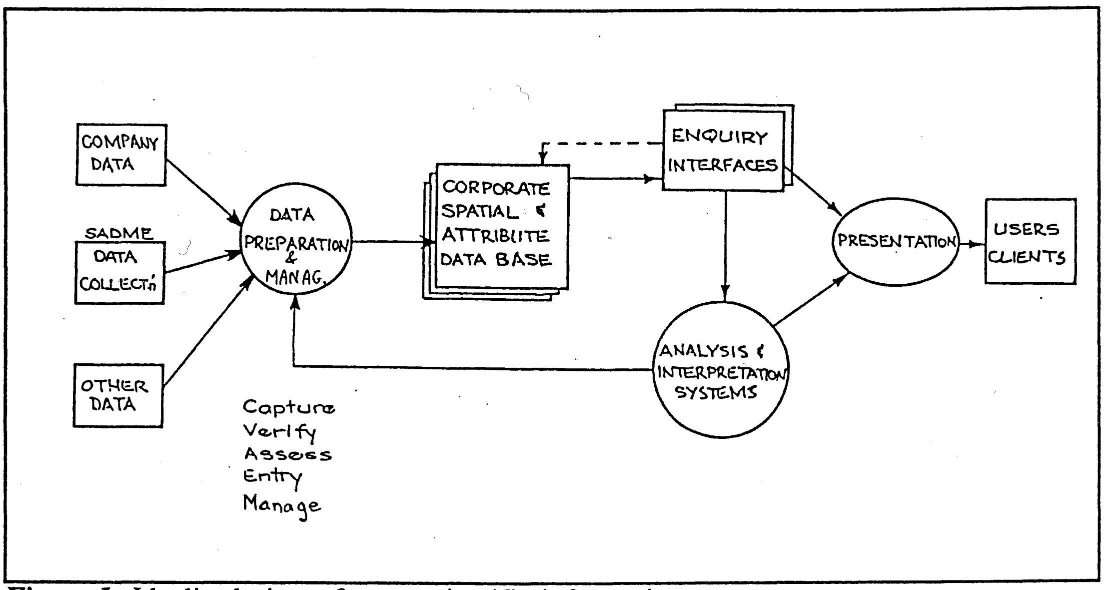


Figure 1 Idealised view of a geoscientific information system

Geoscience Data

Corporate Data requirements within SADME have been defined from user surveys and comprise more than 30 data classes in 10 broad categories (Fig. 3). The logical groupings reflect in part the inter-relationships of data types, and in part current database technology of relational, spatial, image and text systems. Although specific working groups are responsible for collecting or maintaining certain data, the data are used widely across the organisation by all professionals and management, and by external clients that include industry, other Government departments and the public. It is this common requirement for access to data that underlies the corporate GEOSIS concept. In the past, most of this data interchange has been by manual means, generally as maps of various forms. However, data are increasingly being stored in a multitude of digital databases created to serve the requirements of specific working groups.

The main problems from a data management view point are the large number of inter-related data classes and hence the complex physical data base design, the major integration requirements, and the necessary procedural rationalisation for efficient and effective data capture. If all that was required was simple spatial overlay of discrete data sets in a GIS, then the exercise would be relatively simple. Indeed this is why the application of GIS at the project level of some other disciplines has been so successful. However, in geoscience, integrating the various data sets requires data modelling and needs analysis at the broad level, as well as the detailed database design stage.

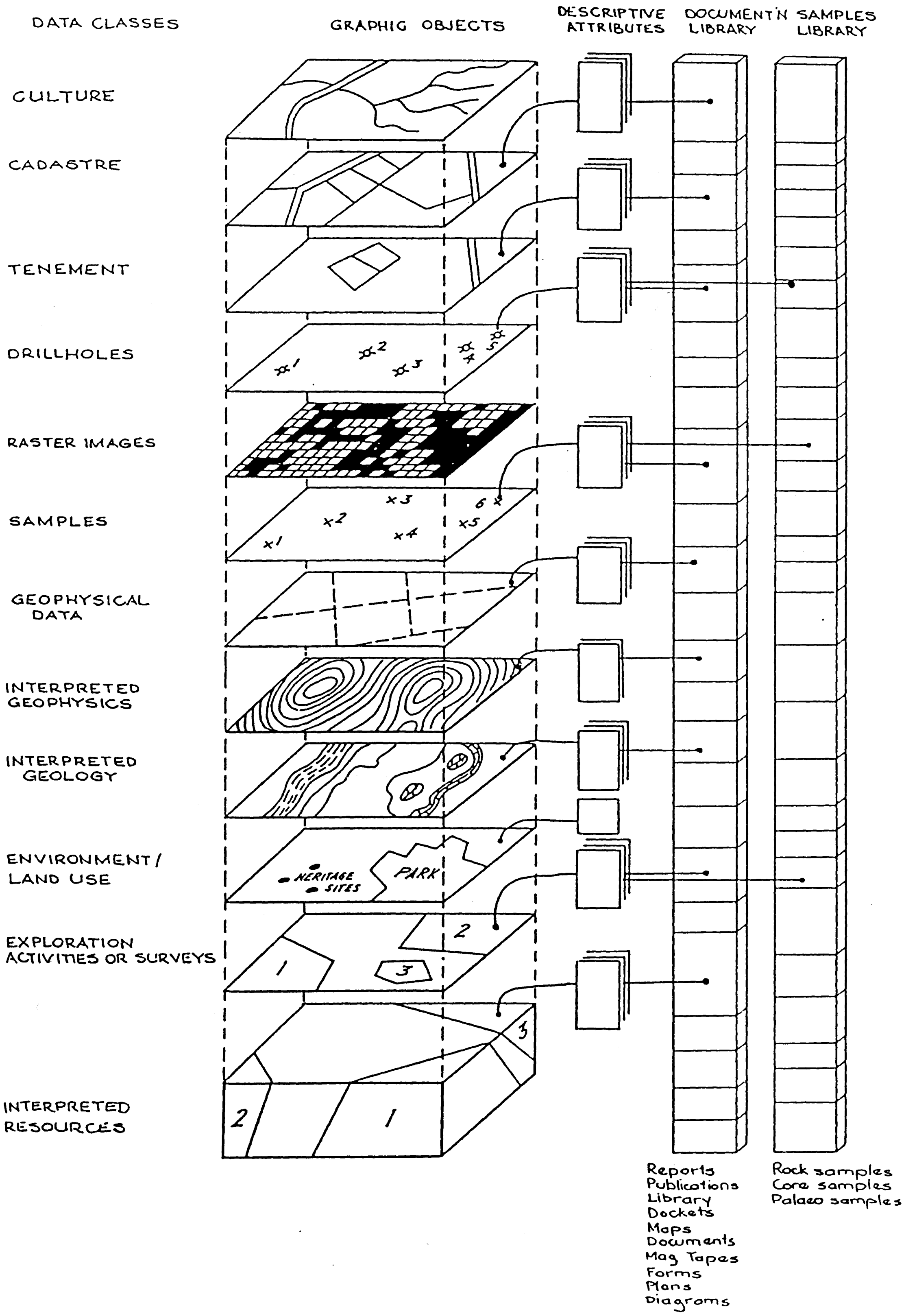


Figure 2 Conceptual integrated geoscience database

G03876.APB

LOGICAL DATABASE	DATA CLASS	EXAMPLES
External	Cultural	Track, Homestead, Fence, Evaporation pond, Pipeline, Well
	Hydrography	Creek, Lake, Coast, Spring
	Topography	Height, Depth, Elevation Contour,Bathymetry
	Cadastre	Hundred, Section, Allotment, Pastoral lease
	Environment	Landform, Soil Type, Vegetation density
Land Admin.	Land Use	National Parks, Zoning, Geology & Aboriginal Heritage Sites, Land access
	Tenements	Exploration, mining and production licences
Physical Item	Documents	References, Reports, Envelopes, Dockets, Magnetic tapes, Seismic sepias
	Physical Samples	Core, Rock, Thin section, Palaeo extract
Drillholes (DH)	Stratigraphy	Formation, R.L., Interval, Lithology
	Aquifer	Aquifer, R.L., Interval, Hydraulics
	Geophysical log	Well, R.L., Tool, Interval value
	Petroleum Well	Reservoir management, Production history, Hardware, Pressure test
Sample Information (SI)	Rock Samples	Lithology, Age, Formation
	Analytical	Geochemistry, Organic analyses, Mag Susc.
	Petrophysics	Core plug, Depth, Value
	Palaeo	Sample, Forams, Pollen, Palaeozone
	Fluid	Static level, Date, Chemistry, pH; Hydrocarbon analysis, Salinity
Field Information	Geology Field Notes	Section description, Structural data, Survey, Hydro data
	Mineral Occurrence	Metallic, Non-metallic, Industrial mineral, Petroleum seep, Gemstone
Resource Data	Resource Data	Petroleum, mineral and groundwater reserves and production data
Geophysical Data	Gravity	Station, Date, Time, Value, Elevation
	Magnetic	Survey, Line, Value, Elevation
	Radiometric	Survey, Line, Channel, Value, Elevation
	Seismic	Survey, Line, Shotpoint location, Elevation
	Electromagnetic	Survey, Type, Value
	Electrical	Survey, Type, Value
	Earthquake	Epicentre, Magnitude, Date, Time
Spatial Image	Raster Imagery	Landsat, Spot, Slar
	Air photography	Survey, Date, Scale, Photo number
Compiled Map	Geological	Formation, Fault, Legend, Cross-section
	Tectonic	Tectonic unit, Province boundary, Structure
	Structure	Isopachs, Structure contours, Formation
	Metallogenic	Mineral occurrence, Geology, Tectonics
	Exploration Index	Major activities and surveys
	Geophysical	Magnetic, Gravity, Radiometric
	Other created	Seismic risk, Coal quality, Hydrogeological

Figure 3 Logical databases and major data classes

Data Modelling

A data model is a graphical representation of the information requirements of an organisation expressed in terms of logical groups (or entities) and their inter relationships. Entities are real phenomena about which data are stored. Entities may correspond to a layer in a GIS, or a table in a RDBMS. Data or entity-relationship modelling is a database design technique that, together with CASE tools, is used to generate a database structure, and to analyse data flow and data management procedures.

Data modelling is an accepted technique for conventional attribute data but has only recently been applied to spatial systems (Firns, 1990) without any as yet accepted approach. It requires a clear knowledge of the data and its relationships and aims to create databases that are efficient to maintain, modify and update. High-level data models provide an overview of a corporate information system. They also facilitate the identification of relatively discrete modules that can be developed separately. This is particularly important for large and complex information systems such as GEOSIS. Detailed data models developed for individual modules must be consistent with the conceptual corporate model.

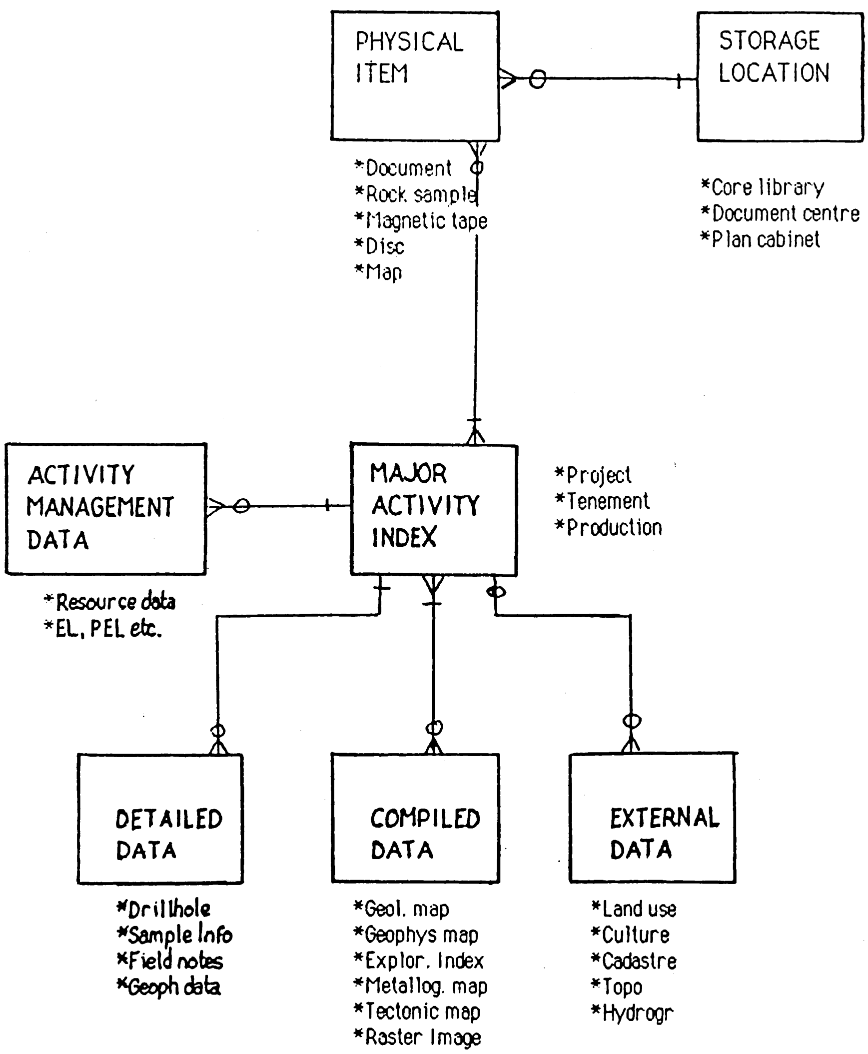


Figure 4 GEOSIS high-level data model

Early data modelling attempts within SADME were hampered by prejudiced perceived priorities and poor understanding of the Department's overall information requirements. Spatial location was either ignored, or was perceived as a separate entity, a questionable approach given that location should be an attribute of an entity (Firns, 1990). After more comprehensive considerations of geoscientific data flow, more complete conceptual corporate data models were generated. In these cases, entities were perceived to have some form of associated spatial connotation and would reside either in a tabular database or as a cartographic layer in a GIS. We have now reached the stage where most of the recognised data classes identified from the needs surveys can be accommodated in a corporate data schema (Fig. 4), and appropriate modules separated for detailed entity-relationship modelling.

At the detailed level, data modelling forms an integral part of a common systems development methodology in which more detailed surveys of user requirements and data flow are used to define and design the data base prior to software development (eg Clough, 1990). In this way, an overly complex corporate system can be successfully broken down into a number of manageable modules that can be implemented sequentially, providing the added advantage of prioritizing systems development where resources are limited. A data dictionary of entity descriptions ensures that duplication of data and functions are minimised.

It is well known that the major cost of a digital information system is in populating the data base. For a Geological Survey, there are two major considerations; capturing data as it enters the Department, and converting existing data that may extend back for more than 100 years. The rationalisation of procedures for the capture of geoscience data as it enters the Department is a critical issue that must be addressed at the detailed data modelling stage. This will not necessarily result in a reduction in personnel, but will almost certainly require new skills and training. Cartographic drafters, for example, will need to acquire new skills of digital cartography. Different strategies are required for converting the huge store of existing data held within the Department's archives. Decisions on the extent of conversion and the approach to be taken need to be made on a case by case basis.

References

Belperio, A.P. and Parker, A.J., 1991. Towards a corporate digital geological mapping and geoscience information system for South Australia. *Proceedings of the National Conference on the Management of Geoscience Information and Data*, Adelaide, pp. 43-60.

Belperio, A., Molyneaux, R., Ivic, D., Sandercock, N., Sambura, A., Daly, S., Morton, J. and Slade, O., 1991. A strategic framework for the development of a geoscientific information system. *Department of Mines and Energy, South Australia, Report Book* 91/4 (2 volumes).

Clough, B.J., 1990. Feasibility study and implementation plan for revising the Rock Sample and Geochemistry databases. *Department of Mines and Energy, South Australia, Report Book* 90/69.

Firns, P.G., 1990. Entity relationship modelling in GIS design: an efficacious approach or an exercise in futility? *AURISA90; Proceedings of the 1990 Annual Conference of the Australasian Urban and Regional Information Systems Association Inc.*, pp. 429-439.

RECENT DEVELOPMENTS IN FIELD-GEOSCIENCE DATABASES IN BMR

Richard S. Blewett & Roderick J. Ryburn

BUREAU OF MINERAL RESOURCES, GEOLOGY AND GEOPHYSICS

GPO BOX 378, CANBERRA CITY, ACT, 2601.

ABSTRACT -BMR has developed ORACLE databases for standardising, storing and manipulating geological field data. A common SITES table is connected to the field and other BMR databases. The OUTCROP table holds details specific to a site, while lithological and petrographical data are stored in the ROCKS and PETROGRAPHY tables. Authority tables and user-friendly screen forms enhance the ease of data entry and data quality.

Introduction

One of the results of the National Geoscience Mapping Accord (NGMA) is the increasing attention being paid to corporate relational databases and their relevance to Geographic Information Systems (GIS) and automated cartography. Minerals and Land Use Program, in conjunction with the Information Systems Branch, has evolved a strategy for the storage of basic field-geoscientific data in new ORACLE databases and the restructuring of existing ORACLE databases. This will help to standardise methods, attributes and terms, and facilitate automated methods of presentation.

The new NGMA databases are shown in Figure 1 and include the SITE, OUTCROPS and PETROGRAPHY databases. The links to the restructured OZCHRON, ROCKCHEM and STRUCTURE databases (see Ryburn & Lenz, 1991) are also indicated. A number of authority or look-up tables have been created and established ones adapted. Their usage reduces the need to key in long text strings, maintains data standards and minimises the number of codes that need to be remembered. A number of triggers have been designed to reduce the need for repetitive data entry.

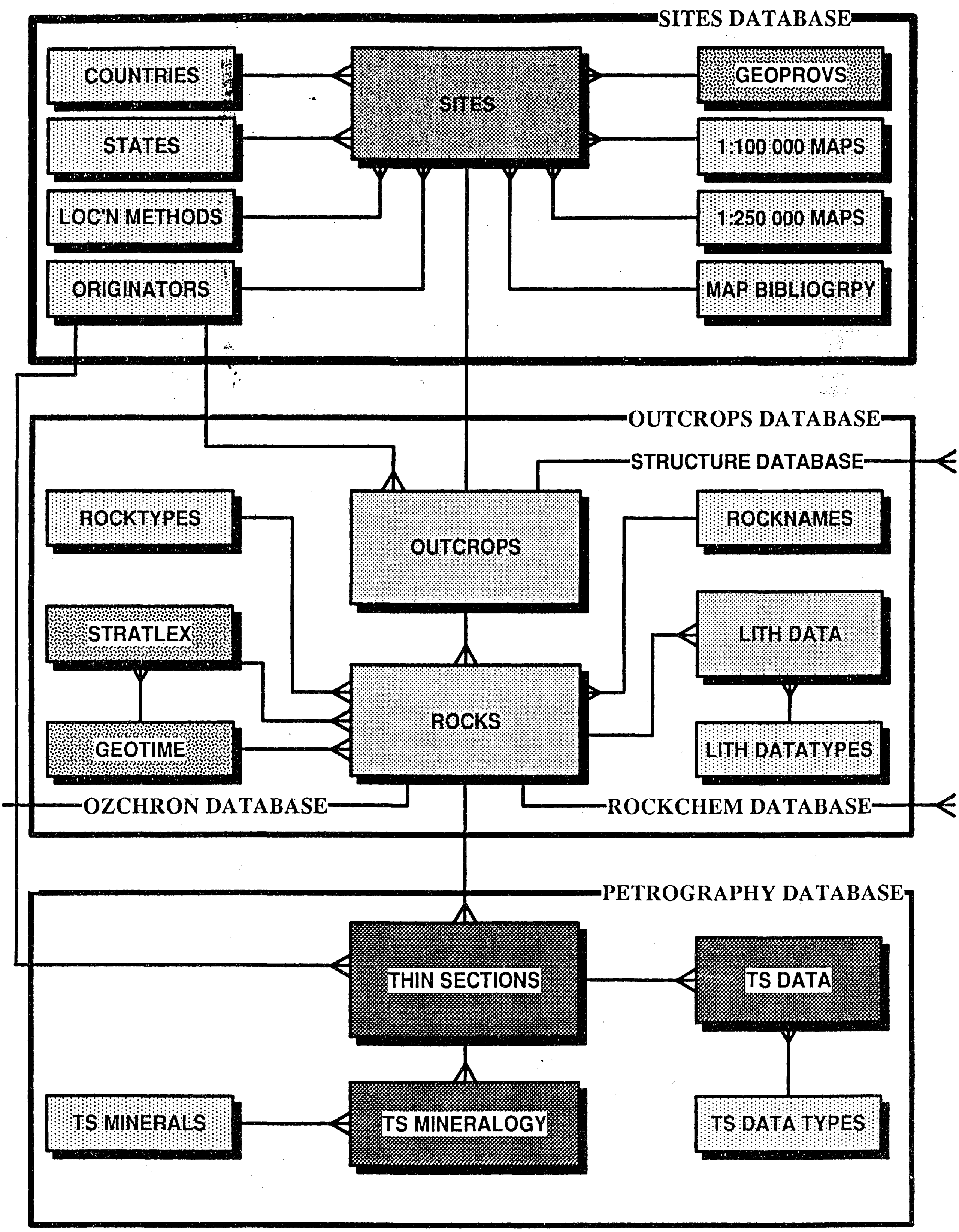


FIGURE 1

Sites Database

Central to the new system (Fig.1) is the SITES table (Fig. 2) which standardises the recording of the location of point data. SITES is designed to be applicable to non-geological data (e.g. gravity measurements) as well as outcrop and sample locations. The SITES table has been partly adapted from the existing SAMPLES table in the ROCKCHEM database (Ryburn, 1990) and partly from the Geological Survey of Queensland's REGMAP field-mapping database system (Withnall 1992 - this volume).

```
-------------------- BMR MAPPING DATABASE - SITES FORM --------------------
Originator    17 >Glikson, A.Y.            Obs. Date 23-JUL-69   Time
Site ID       69200079   Field ID AYG269     Country >AUS   State >QLD
Geol. Prov.   54 >Mount Isa Inlier         Subprov 96 >Eastern Fold Belt
Loc. Descr.   Sandy Creek
                                            Metres          Metres
1:250K Map    SF5402>CLONCURRY              East            North
                                            Decimal         Decimal
1:100K Map    7056>CLONCURRY               Lat 20.9239   S  Long 140.731   E
                                            Map             Abs. Accur.
Loc'n Meth.   2  >1:100 000 geological map  ID >            Metres 100
Airphoto      CLONCURRY 5/2345            Related   Oc St Pt RC Oz MD SC RT RP
                                          Data Sets X     X  X
Height (m)         +/- (m)                ('X' indicates related data exists)

> Lookup Field - Press 'List Field Values' to view associated form.

Press 'List Field Values' for the Originators Form
   v Char Mode: Replace  Page 1                         Count: 2
Press Ctrl-Break to exit the Terminal option.
```

FIGURE 2

The SITES table uses original location or sample numbering systems as part of its primary key. If all locations belonged to BMR and had unique BMR registered numbers, this number would be sufficient in itself. But, the SITES table is also intended for location data provided by other organisations and individuals, all of whom use their own numbering systems. An "Originator" is thus required to maintain the uniqueness of a location amongst diverse location numbering systems. The primary key for the SITES table is thus the combination of "Originator" and "Site ID". "Field ID" allows for alternative numbering systems. The "Originator" also provides an indication of data lineage. "Observation Date" and "Time" can also be noted.

"Absolute Accuracy in Metres" of a location is triggered by the appropriate "Location Method" which is controlled by a look-up table. For GPS located sites, the 'GPS' mnemonic is inserted. If a topographic or geological map is used as the location method, the field "Map ID" cross references to a Map References table (Fig. 1) containing bibliographic map data.

The "Geological Province" field is triggered by the "Subprovince" field number (Fig. 2), and both are linked to a Geological Provinces authority table (Fig. 1). "Location Description" is a free-text field to describe the field location.

The "1:100K map" is the field for the standard map-sheet number, and is referenced to an authority table (Fig. 1). The relevant "1:250K Map" field is automatically derived from the 1:100K map ID. The site position is stored as an AMG ("Metres East"and "Metres North") or as latitude and longitude ("Decimal Latitude" and "Decimal Longitude"). Only one co-ordinate system is required as a routine calculates and inserts the other.

Aerial photograph details are stored in the free text "Airphoto" field, while altitude above sea level (if required) is stored in "Height" with its accuracy in "+/- (m)". "Related Data Sets" provide a number of checkable boxes to quickly indicate what related data sets are available for a particular site.

Outcrops Database

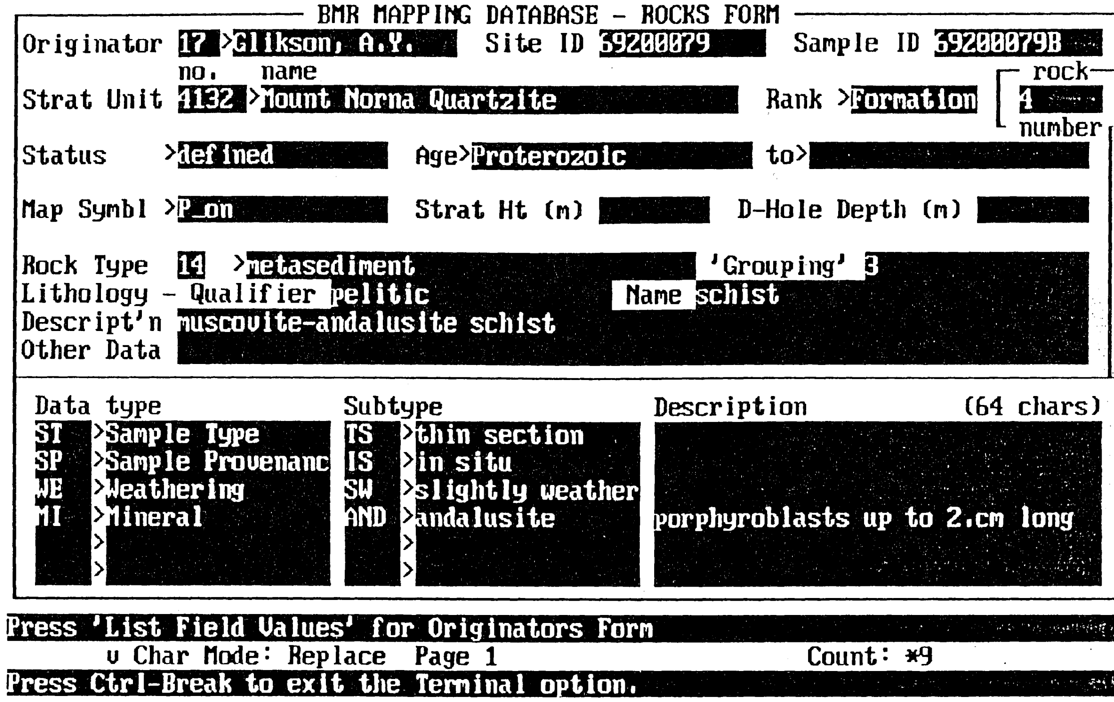
The OUTCROPS table has a one-to-one relationship with the SITES table through the "Originator" and "Site ID". The OUTCROPS form (Fig. 3) is designed to accommodate information relating to the OUTCROP as a whole - for example, stratigraphic contacts, intrusions, faulted contacts, unconformities etc. The OUTCROPS table is essentially a free-text entity, allowing observations about "Rock Relations", "Sketches", "Photographs", "Drill Hole Data", "Vegetation" and "Landform".

```
──────────── BMR MAPPING DATABASE - OUTCROPS FORM ────────────
Originator      17 >Glikson, A.Y.      Site ID 59200079
                                                        (128 chars)
Rock Relations  quartzo-feldspathic meta-arenite interbedded with andalusite

Sketches        refracted S-surfaces at boundary of schist & metarenite

Photographs     andalusite porphyroblasts in pelitic schist

Drill Hole   |  Company
Data         |
             |  Hole ID
             |
             |  Azimuth        Inclination

Vegetation      tussock, sorrel & scabweed

Landform        incised stream valley

 > Lookup Field                 Press 'Next Block' for the Lithologies Form

Press 'List Field Values' for Originators Form
      Char Mode: Replace  Page 1                    Count: *1
Press Ctrl-Break to exit the Terminal option.
```

FIGURE 3

The OUTCROPS table has a one-to-many relationship with the STRUCTURES database as many structures can be found at the one outcrop. STRUCTURES is an earlier database that stores mesoscopic structural orientation data.

The ROCKS table (Fig. 4) is essentially for information on lithologies and samples at an outcrop. The ROCKS form is not exclusively for samples - some records are for samples, others merely describe lithologies at an outcrop. But, where a sample has been taken the "Sample ID" is entered. The "Sample ID" links the ROCKS table with the PETROGRAPHY, ROCKCHEM (contains whole-rock geochemistry) and OZCHRON (contains isotopic dates) databases.



BMR MAPPING DATABASE - ROCKS FORM

Originator 17 >Glikson, A.Y. Site ID 69200079 Sample ID 69200079B
no. name
Strat Unit 4132 >Mount Norna Quartzite Rank >Formation rock 4 number

Status >defined Age>Proterozoic to>

Map Symbl >P_on Strat Ht (m) D-Hole Depth (m)

Rock Type 14 >metasediment 'Grouping' 3
Lithology - Qualifier pelitic Name schist
Descript'n muscovite-andalusite schist
Other Data

Data type Subtype Description (64 chars)
ST >Sample Type TS >thin section
SP >Sample Provenanc IS >in situ
WE >Weathering SW >slightly weather
MI >Mineral AND >andalusite porphyroblasts up to 2.cm long
>
>

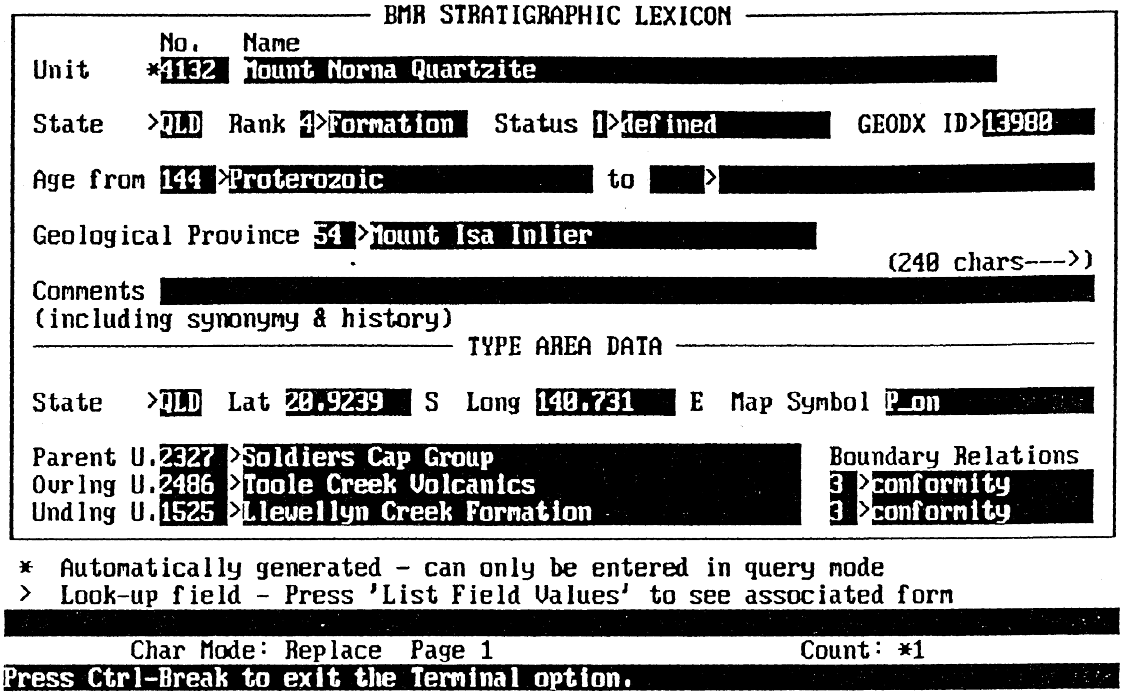
Press 'List Field Values' for Originators Form
v Char Mode: Replace Page 1 Count: *9
Press Ctrl-Break to exit the Terminal option.

FIGURE 4

The "Stratigraphic Unit Number" points to the name of the unit in the BMR Stratigraphic Lexicon (Fig. 5), which contains information on each stratigraphic unit. The unit can be of any "Rank" e.g., member, formation, group etc. The Stratigraphic Lexicon inserts "Rank", "State", "Status" (defined, informal etc.), "Age" (range) and "Map Symbol" into the appropriate fields in the ROCKS form. The Stratigraphic Lexicon also contains data on the location of the type area (as latitude and longitude), the "Parent Unit", "Overlying Unit" and "Underlying Unit" (referred to by "Stratigraphic Unit Number"), as well as a small look-up table with the "Boundary Relations".

The "Map Symbol" is important as it is the polygon map label used to link the attribute data with the graphic data in the GIS. This link can also be achieved by using the AMG or latitude/longitude fields from the SITES table.

The third dimension is accommodated for height data from logs of vertical sections in "Strat. Height" as well as drill hole data in "Drill Hole Depth".



BMR STRATIGRAPHIC LEXICON

	No.	Name
Unit	*4132	Mount Norna Quartzite

State >QLD Rank 4>Formation Status 1>defined GEODX ID>13988

Age from 144 >Proterozoic to >

Geological Province 54 >Mount Isa Inlier

(240 chars--->)

Comments
(including synonymy & history)

TYPE AREA DATA

State >QLD Lat 20.9239 S Long 140.731 E Map Symbol P_on

			Boundary Relations
Parent U.	2327	>Soldiers Cap Group	
Ovrlng U.	2486	>Toole Creek Volcanics	3 >conformity
Undlng U.	1525	>Llewellyn Creek Formation	3 >conformity

* Automatically generated - can only be entered in query mode
\> Look-up field - Press 'List Field Values' to see associated form

Char Mode: Replace Page 1 Count: *1
Press Ctrl-Break to exit the Terminal option.

FIGURE 5

The "Rock Type" field combines rocks into their broad lithological groups (e.g. felsic intrusive or clastic sediment), it was adapted from ROCKCHEM. "Rock Type" is analogous to 'Rock Class' in REGMAP (Withnall 1992 - this volume). "Rock Name", adapted from the REGMAP look-up tables is a more specific title for a sample (e.g. granite or greywacke). We suggest that the free-text field "Description" be used as an abstract, analogous to the 'ABST' code in REGMAP.

The following section is adapted from the "Data Types" code principle of REGMAP, it allows a variety of users or specialists to use the database and yet still share many common attributes. This is possible as "Data Types" can be designed to suit any need and thus has significant advantages over the 'tick-the-box' type of database. In the ROCKS database, common "Data Types" are grouped together with a variety of "Subtypes", the user inserts a code and the definition is supplied by the authority table. The user can also query the authority table for the list of codes and definitions. The "Description" field is free-text, no provision has been set for suggested abbreviations (unlike REGMAP).

PETROGRAPHY DATABASE

The "Originator" and "Site ID" numbers are common to the SITES, OUTCROPS and ROCKS databases, while "Site ID" number is the relate item for ROCKCHEM, ROCKS and OZCHRON databases. The "Thin Section" field accommodates multiple sections for any one sample. We suggest that "Description" is a free-text field of the basic mineralogy of the slide; "Comments" are self explanatory.

```
------------------ BMR MAPPING DATABASE  - PETROGRAPHY FORM ----------
Originator 17 >Glikson, A.Y.    Site ID 5920079B    Sample ID 59200079B.2
                                                                  TS
Thin Sect. AYG-79     Descript'n q-mu-bi-stau schist              1
                                                                  No
Comments   Oriented 130/84, cut in ZX plane

Data Type         Subtype           Description                   (64 chars)
TEXT>texture      TCFL>foliation    main foliation predates andalusite porph
TEXT>texture      CREN>crenulation  some evidence of S3 as crenulations
ALT >alteration   CHLT>choritised   biotite, partial
ALT >alteration   SERI>sericitised  some pinnite alteration of andalusite
    >                 >

Mineral               %vol Comments                               (64 chars)
QTZ >quartz           55
MS  >muscovite        28
BT  >biotite          12   some of it interleaved with muscovite
AND >andalusite       3    large porhpyroblasts riddled with inclusions
ST  >staurolite       1.5  small ammount in one part of section
TUR >tourmaline       -1   accessory
Press 'List Field Values' for Petrography Data Types Form
          Char Mode: Replace  Page 1                 Count: *4
Press Ctrl-Break to exit the Terminal option.
```

FIGURE 6

The "Data Type", "Subtype" and "Description" fields are similar in design to those in ROCKS. Many of the data types are common to both authority tables. This section is designed for descriptive thin section data (other than mineralogy). The "Mineral" field has an authority table for mineral codes, the "%vol" and "Comments" fields are free-text related to the mineral defining the row.

CONCLUSIONS

Although the number of tables involved in the NGMA field mapping database may at first sight appear daunting, many of the tables are small look-up tables used to validate data entries. The SITES database, for example can be viewed for most purposes as a single flat table despite the fact that it actually has nine tables.

The NGMA field mapping database system provides a significant saving in data entry, much control of data quality, the necessary data lineage, as well as providing a powerful and yet flexible, non-graphic attribute system for 'specialist' users in a GIS environment.

REFERENCES

Ryburn, R.J., 1990--Users' guide to the Petchem database. *Bureau of Mineral Resources, Australia*, Record 1990/19, 53pp.

Ryburn, R.J., this volume. Relational databases for continent-wide data.

Ryburn, R.J. & Lenz, S.L., 1991--Geoscientific relational databases in BMR and the 'client-server' method. *Proceedings of the National Conference on the Management of Geoscientce Information and Data, Adelaide*, Session 7, 1-15.

Withnall, I.W., this volume. REGMAP as a model for regional geological mapping.

THE ROLE OF STANDARDS IN ENSURING DATA INTEGRITY AND PROVIDING A BASIS FOR DATA EXCHANGE

Cherryl Schmidt

Department of Resource Industries
GPO Box 194
Brisbane, Queensland, 4001

Abstract:

There is an increasing appreciation of the value of data and information as a resource. This in turn is leading to an awareness of the need to share data, both internally and externally to our organisations.

The establishment of, and adherence to, data standards facilitates this communication by improving data consistency, thus ensuring data integrity.

Introduction

Most, if not all, of the people here today are involved in the collection, management, presentation or application of geoscientific data. Over the years, our organisations have invested vast amounts of resources in the collection of this data, but unfortunately, in the majority of cases, this important corporate asset runs a very real risk of being seriously devalued.

In some ways, it is difficult to put a price on the data collected, because it has no value until it is needed. However, if you consider what it would cost to regenerate the data from scratch, for example to re-run a seismic survey, or to re-analyse a rock sample, the concept of 'value' becomes clearer. It is probably true to say that in all of our organisations, the knowledge base, or the information resource, is the single most important asset. This is partly because of the cost of collecting the data initially, and therefore the implied cost of recapturing it if it were lost; and partly because of the time, effort and resources poured into analysing, massaging, and presenting the information.

Generally speaking, individual business units within an organisation gather data for their own purposes. This may be in connection with, for example, surface geology and mapping, petroleum or mineral exploration. In recent years, Information Technology has become more accessible to people who are not computing specialists, for example with the introduction of personal computers (PC's) and affordable database products such as DBase and FoxBase. As a result, these business units have often taken advantage of the new technology and developed their own databases to manage their data. The power of the computer has been brought to the desk of anyone who has the requirement, and the inclination, to take advantage of it.

This, in itself, is not a bad thing. I have been employed in the computing industry for many years and know that in the past there has often been an air of mystery surrounding 'the computer' and its application. Most people's perception of Information Technology today, however, is significantly different from that of only a few years ago. In most cases, the computer and the various software packages are now seen as tools, like any other, to assist in the performance of our normal business functions. And this is how it should be.

The problem however lies in the fact that many geoscience professionals and specialists are now pouring increasing amounts of their resources into performing computing tasks. It has become necessary for them to learn how to use the various software that is available to them, and to devise the most efficient way to manage their data. Unfortunately, knowing how to use a piece of software is only part of the Information Technology requirement. Unless the application or database is designed correctly, although it may well appear to 'manage' the data better than it has been managed before, it will often have considerable limitations. So often, a digital database merely mirrors a previously operated manual system.

These limitations become apparent when the data needs to be used for a previously unconsidered application, for example the use of petroleum exploration data to identify mineral deposits. Unless the database has been designed with the necessary 'key' fields identified, and with enough flexibility to allow using the data to perform a different function, it may be extremely time consuming and laborious to extract the required information. It can be particularly frustrating when you know that all the necessary information is in the database, but you are not able to access it in a useful form.

This problem stems from two basic causes. The first is that the database has been developed in a vacuum. Although the resulting product may well meet the original single requirement, its application is extremely limited. In order to resolve this, even the simplest requirement should be fully specified before attempting to develop the database. Databases have a nasty habit of growing and growing, rather like Pinocchio's nose, and in no time at all can become a liability rather than an asset. Detailed analysis of the data and functionality, and the corresponding design of the database are very important steps, which often get neglected.

The second cause brings me to the main point of this paper, and that is the absence of data standards. A database can be designed with all the flexibility in the world built into it, but for example, if each of three databases call a data item by a different name, or conversely if each have different data items with the same name, the advantage is lost.

Information Technology Evolution

I have observed within my own Department the effects of the Information Technology evolution, and I believe that it has probably been similar in most of your own organisations. Before the accessibility of computers, each business unit was responsible for gathering its data, analysing it, and managing the storage of it, whether in a card index, a filing cabinet, or a computer database. As a result, the business unit clearly 'owned' the data, and held on very tightly to it.

There were a number of reasons for this. One of these was data security; if someone else uses it, they may corrupt it somehow. Another reason was job security; 'if this information becomes freely available, I may be out of a job'. Yet another reason was data integrity; someone else may not understand the intricacies of the data, and may thus interpret it incorrectly. All of these are valid reasons in their own way, and I am sure that each of you can think of many more.

With the advent of personal computing and relational databases, people generally are more aware of the potential for combining data that was not previously possible. Vast quantities of data can be sorted and merged, analysed and massaged, at incredibly high speed, and answers can be provided to all sorts of questions that were never previously asked.

Requirement for Data Standards

That is only true, though, if standard terms and naming are in place. We all speak the same language, or at least, we think we do, and therein lies another problem. I believe it was Oscar Wilde who once said something like "the English and the Americans have everything in common, except the language" and for our purposes, that could be extended to include the Australians!

Although most terms are well documented in the geoscience field, each area has many specialists, and they often impose their own interpretation. Over time, this can result in two or more people having an entirely different understanding of what is meant by a particular term. On a person to person basis, this does not normally cause too many problems because any misunderstanding soon becomes apparent and can usually be resolved by discussion. In the case of data, however, it can be disastrous. Unless the data is adequately defined in unambiguous terms, there is always the possibility that it will be misinterpreted.

Corporatisation of Data

Many people seem concerned that by corporatising their data, they will lose control of it. In fact what happens is the value of that data is increased because it is no longer being viewed in isolation. Most organisations are now recognising the advantage in 'corporatising' their information resource. Some are actively promoting this concept by employing people like myself, whose role it is to ensure the efficient and effective management of data across the whole Department. Other organisations have a shared data management responsibility, where a person or group of people manage the data in a specific area, for example all petroleum-related or mineral-related exploration data. Yet other organisations agree whole-heartedly with the concept, but do nothing about it. No doubt they believe that "it will be all right on the night!", but unfortunately, it rarely is.

This is all fine as far as it goes, but we must go much further. The transition from 'hoarding' data, to sharing it between individual business units and recognising the corporate view, is no longer enough. The world continues to get smaller, and particularly in Government there is a growing appreciation of the advantage in sharing data, not only within an organisation, but also across other Government Agencies, and even with Industry. The pooling of information resources is vital, particularly in these days of shrinking budgets. If we make information available to Industry, they will have a better knowledge on which to base their future exploration activity, thus resulting in Royalties coming back into the State coffers.

So now instead of being concerned with organisational data standards, we need to seriously consider Australian Standards, so that we can communicate better with each other.

And of course, we should not stop there. It is essential that we recognise international standards because that is the next natural progression in the communication and sharing of our data resource.

There are a number of committees and working groups addressing various aspects of standards, some of which we have heard about during the course of this workshop, for example the Spatial Data Transfer Standard (SDTS) and the Australian Mining Industry Data Exchange Standard (AMDEX). All of these are equally important to the process of establishing a base line from which we can all work. The outcomes of these committees will contribute to the laying of ground rules for future development, and will provide significant benefits to all of us, as long as we actively participate in the process. It is essential that organisations recognise the investment placed in their information resource, and start to manage it both efficiently and effectively. Until this happens, they run the risk of being left behind.

Benefits of Standardisation

The long term benefits of standardisation in geoscience data management can be demonstrated in a number of ways. For example it will reduce duplication of effort, because if we are all observing similar rules, it will be much easier to share data across organisations. It will result in better quality data, because of improved consistency and discipline of documentation; by using common terms, definitions, and formats there will be less room for ambiguity. In addition, there will be the 'value-added' aspect of viewing one set of data in relation to others, both within our own organisations, and with external bodies. But the most important benefit, however, will be the promotion of exploration, and that, I believe, is our reason for being in this field.

GGDPAC Geoscience Data Standards Working Group

Last October, GGDPAC asked me to chair a Geoscience Data Standards Working Group with a representative from each State. Our brief was to identify core geoscience and related cartographic data types, and suggest Australian Standards for their representation and definition, based where possible, on existing standards, authorities or common practice. A draft report has recently been prepared which documents the most significant databases in each government geoscience agency, the standards being applied within each organisation, a comment on the relevant computing environments, and the policy with regard to data standards. This report is being discussed and revised during the course of this workshop, and will be made available to anyone who has an interest in our findings.

Suggested core geoscience and related cartographic data types will be identified in the report, and this will provide a basis for the next step, which will be to agree definitions and formats, and where appropriate, sets of values. This will no doubt be a long process, but the main point is that the need has been recognised, and is now being acted upon. Increasing the general awareness of the requirement for data standards is, in itself, a means of stimulating communication. More people are acknowledging the fact that there are others out there who may be interested in the data they are gathering or analysing, and I believe this is an important step forward.

Developments in the Queensland Department of Resource Industries

Within the Queensland Department of Resource Industries, the need to share data has been recognised for a number of years, and the concept of a corporate database combined with a geographic information system (GIS) is in the process of evolving from a vision to reality. The Mineral and Energy Locational Information Network (MERLIN) has been developed and refined over the past three years, and the first phase of this project is now firmly established as the Mining Tenures Data Base (MTDB). Strategic planning for the second phase of MERLIN has recently been completed, and development of the Geoscience and Resource Data Base (GRDB) should commence this year.

The second phase of MERLIN involves pulling together approximately eighteen diverse data sets into a cohesive corporate database, which will interface with the mining tenures data introduced in phase 1. This development provides an excellent opportunity for the Department to implement Departmental data standards. During the course of the conversion of these datasets to the Oracle and Arc/Info platform, standard lookup tables will be established where appropriate, to be shared by all of the datasets. Data Custodians have already been identified and they will each have the responsibility of ensuring that the data is correct and available to anyone who requires access.

Departmental Quality Manual

To facilitate this process, I am developing the data standards section of the Departmental Quality Manual. This will be a stand-alone document issued to each branch within the organisation. It will clearly describe the purpose of Departmental data standards, and the steps to be followed to access the data. It will include definitions of all the relevant terms and identify the physical location of the data, as well as the Data Custodian. The document will also contain a complete set of all Departmental standard codes and descriptions, and will thus be a useful source of reference. All future systems development within the Department will be required to adhere to these standards, and in this way, the Department has an opportunity to regain control of its data resource.

Corporate Data Dictionary

All of these terms and definitions are recorded in the corporate data dictionary. It is a computerised database holding information about the data and processes of interest to the Department, and is used for storing and handling data definitions. It is a central repository of meta-data, that is, data about data.

Each piece of data is defined only once, but can be linked to many applications. The advantage of using such a dictionary is that the information is consistent across all applications, and the impact of proposed enhancements can be assessed, based on the knowledge of where that data is used.

The corporate data dictionary is effectively the knowledge base of the organisation, and therefore has the potential to improve the quality of applications by ensuring that all relevant information is made available. It is the stepping stone to better quality systems design, based on knowledge and understanding of what the data is rather than what it is perceived to be. Any systems development can have a ripple effect, and will often impact on other areas. By knowing the potential overlaps before starting development, the analyst is in a much better position to develop a more efficient application.

Conclusion

To conclude my presentation today, I would like to summarise the main points again.

Organisations need to recognise the value of their information resource.

Data needs to be viewed from the corporate perspective rather from than the narrower view of an individual business unit.

We must continue to exploit advances in Information Technology, but it is important to ensure that databases are developed in a professional manner.

Organisational data standards need to be documented and made readily available. These standards should be developed based on existing international and Australian standards, where they already exist.

And finally, we should endeavour, where possible, to share our information, thus reducing duplication of effort and resources in gathering data, and effectively promoting exploration in Australia.

CATEGORISING DESCRIPTIVE DATA WITH REFERENCE TO THE REGOLITH AND ENVIRONMENTAL GEOSCIENCE DATABASES

Sonja L. Lenz and Colin F. Pain

Bureau of Mineral Resources, Geology and Geophysics

Abstract - Traditionally, information in regolith-related disciplines like geomorphology, pedology and Quaternary geology has been recorded in descriptive form containing subjective, even emotive, terminology. To be able to use relational database technology workers in these fields must agree on classifications and terminology to be used.

Keywords - Classification, Databases, Earth Science Information, Landforms, Regolith

INTRODUCTION

One of the main functions of the Bureau of Mineral Resources has always been the gathering, processing and dissemination of geoscientific information on the Australian continent and its surroundings. Over the years an enormous storehouse of knowledge about the geology and geophysics of Australia has been built up. The potential uses of this knowledge base are many, not only in the traditional field of exploitation of Australia's natural resources, but increasingly BMR's scientists are contributing to decisions on environmental planning, mitigation of natural hazards, land use and land management issues, to name some of the more recent developments. Hand in hand with this, the regolith and environmental geosciences are becoming acknowledged as important geoscience disciplines.

At the same time there have been fundamental technological changes, of which we will only mention computers, that have led to enormous changes in the way things are done at BMR. In the area of computerised databases we have taken the path along flat files, network and hierarchical databases to relational databases and geographic information systems (GIS), the major theme of this workshop. We would like to share our experiences in setting up a relational database for the regolith data with you. They might seem specific to this area of geoscience but can, in fact, be generalised for most work with relational database management systems (RDBMS).

RELATIONAL DATABASES

To begin with, just a short overview of relational database concepts: In the relational database model we talk about *data entities* and the *relationships* between these entities. The basic (and only) unit of data storage in a relational database is the *table*, a two-dimensional grid of columns and rows. An *entity* is any distinguishable object that is to be represented in the database; usually at least one table is set up to contain the *attributes* of each entity. Every table within the database is defined with a name and a set of columns. The attributes which characterise the entity are the columns in the table. Each column is given a name, a datatype, and a width. The distinct instances or occurrences of the entity, the so-called records or table rows, each have a certain set of attribute values.

Relational databases are firmly based on the mathematical set theory and relational algebra. Operators like UNION, INTERSECT and DIFFERENCE can be used to create new tables from existing ones. The *relationships* between entities are logical links between them which can be used to associate data in one table with data items in another. This is usually done by "joining" two tables through data values which are common to both tables.

If one record in a table relates to a single record in another table we talk of a *one-to-one* relationship. Similarly, if one record in the first table corresponds to more than one record in the second table, we are looking at a *one-to-many* relationship. Finally, a *many-to-many* relationship occurs when several records in the first table correspond to more than one record in the second table.

A *key* within a table is an attribute (column) or attributes whose values uniquely identify each record (row). *Indexes* can be used to speed up execution of transactions and, in the case of "unique indexes", to guarantee uniqueness of records.

There are no precompiled links between the tables in a relational database and therefore any retrieval, that is, any combination of fields from any number of tables, is theoretically possible. The user does not need to know anything about the physical storage of the data on the computer, all the user needs to know is the logical structure of the data. All data manipulations are non-procedural, which means the user tells the system what is needed, not how to go about selecting the data. In addition, the structure of the database can be changed relatively easily. Relational database management systems are widely accepted because of their conceptual simplicity and flexibility.

There are, of course, some disadvantages in using relational database systems, the major one being that their performance is often not as good as for instance databases built using a procedural language. As most BMR databases are not transaction intensive, that disadvantage would not justify the effort and time it would take to program them in a procedural language.

DATA ANALYSIS FOR THE REGOLITH DATABASE

When we started looking at the regolith data gathered in the field with the view of putting them into ORACLE on BMR's DG, we realised that ill defined terminology was being used. Moreover, it was being used inconsistently, and many of the attributes were being recorded in verbose descriptions that didn't really enable valid comparisons from unit to unit. In addition, relationships between attributes had not been properly established.

In the ensuing analysis process it was further realised that for some entities like geomorphic and weathering processes, drainage and landform patterns as well as tectonic structure elements there were only a relatively limited number of distinct values. Where available, authoritative references were used to derive lists of these distinct values and their definitions, in other cases the Regolith Group themselves devised lists after discussions with as many workers in the specific fields as possible (e.g. a successful regolith workshop was organised at BMR in June 1990). These authority lists form the basis of the lookup tables in the database which are used for the description of field sites and/or mapping units.

In this way considerable effort went into structuring regolith and terrain attributes into distinct, well-defined categories more suitable to a relational database.

RTMAP

RTMAP, the BMR Regolith Database, was developed jointly by the Regolith Group of the Minerals and Land Use Program and the Information Systems Branch (Pain *et al.* 1991, Lenz 1991). A number of basic guidelines were recognised.

1. The database must be constructed so it can be used with a geographic information system.

2. Standard terminology must be used. Moreover, in order for RTMAP to be accepted by users, the terminology used should be adopted from that already in use, rather than developing new terms.

3. Terminology must be "protected" in the sense that only agreed terms can be used. Individual users cannot be allowed to make up their own terms.

4. The terminology used must fit into a hierarchical classification, and be assigned codes accordingly. This allows more efficient use of the database, especially for searches.

Each of these guidelines is discussed in more detail below.

The Database and the GIS

Most earth science databases collect information about a number of attributes at sites. (Sites are small areas of land considered to be representative of the land features associated with the observation being made.) For example, soil scientists (McDonald *et al.* 1990) are mainly concerned about describing land and soil attributes at soil profile sites. Similarly, REGMAP as used by the Queensland Department of Resource Industry is concerned entirely with data collected at sites (Lang *et al.* 1990). Site information is also important for regolith studies, so site data is collected for RTMAP (Fig. 1).

Sites can be spatially referenced, and thus become point data in a GIS. However, point data do not give the GIS user general information relating to broader areas. For regolith, as for most land related attributes, it is not possible to use site data to characterise mapping units. There is too much spatial variability, which results in varying degrees of impurity of mapping unit, depending on mapping scale.

For this reason a second set of attributes is entered into RTMAP. These data relate to the whole mapping unit, or regolith landform unit (Fig. 2). They can incorporate site data, but they also incorporate observations of the whole mapping unit, made both in the field and from maps and images. Among other things, variability is described. These data become the polygon attributes in the GIS, and support the great potential for spatial analyses provided by the GIS.

Thus, both site and unit data are essential for full utilisation of the GIS.

Standard Terminology

The two most important groups of attributes in RTMAP relate to regolith and landforms. Regolith is important because that is what we are trying to understand. However, because regolith occurrence in the landscape is very complex, we use landforms as a base for our regional scale mapping. This is, of course, standard procedure for most types of landscape mapping.

After researching the various classifications available for landforms, we decided to adopt that presented by McDonald *et al.* (1990), with a few changes that reflected our needs and experience. Our landform classification and terminology is the landform pattern of McDonald *et al.* (1990), and is suited very well to our regional mapping scale. McDonald *et al.* was also used for terms relating to drainage, geomorphic processes and weathering processes.

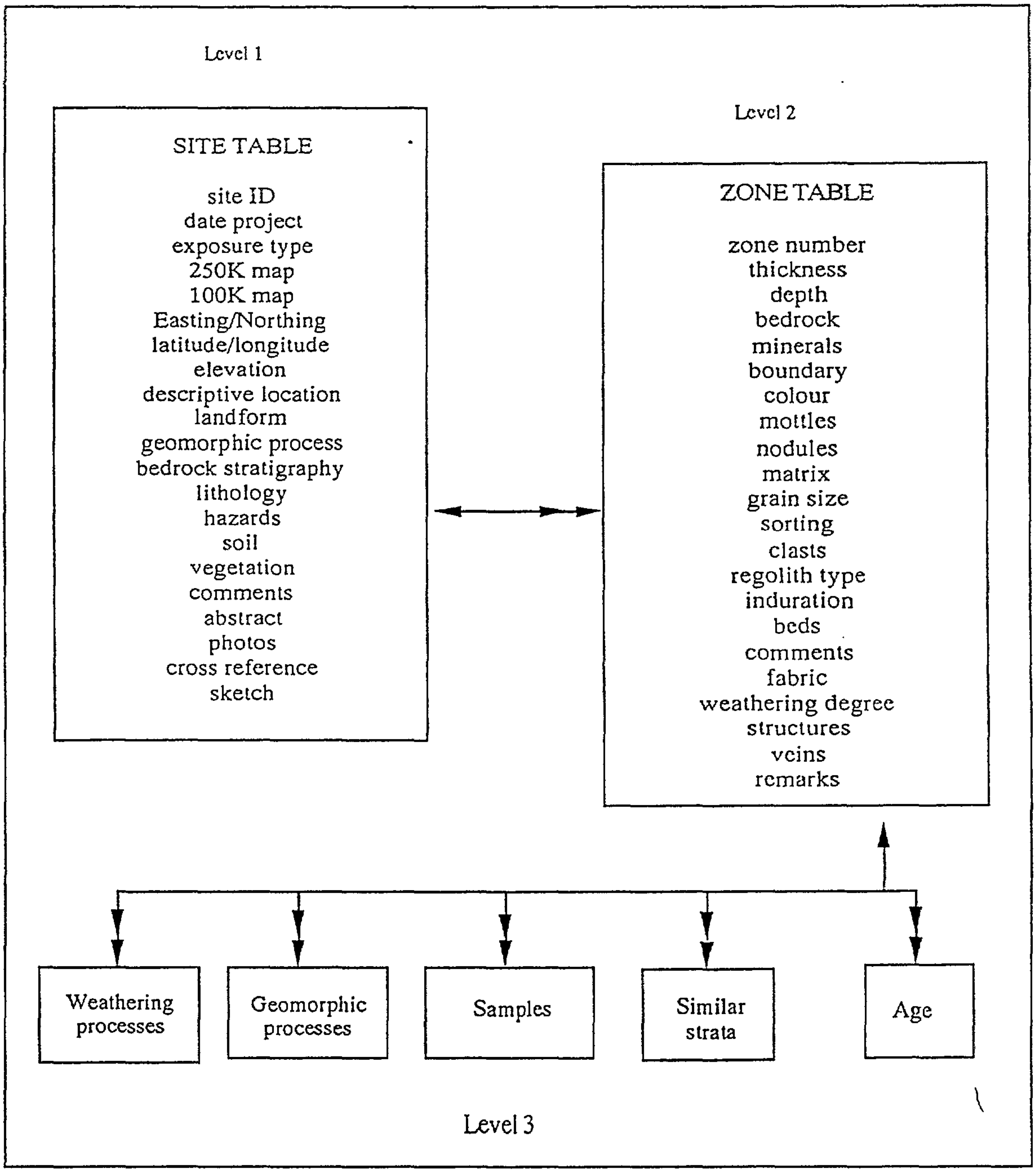


Figure 1. Simplified logical relationships of field site data in RTMAP (from Pain *et al.* 1991). Double headed arrows indicate many-to-one relationships between tables.

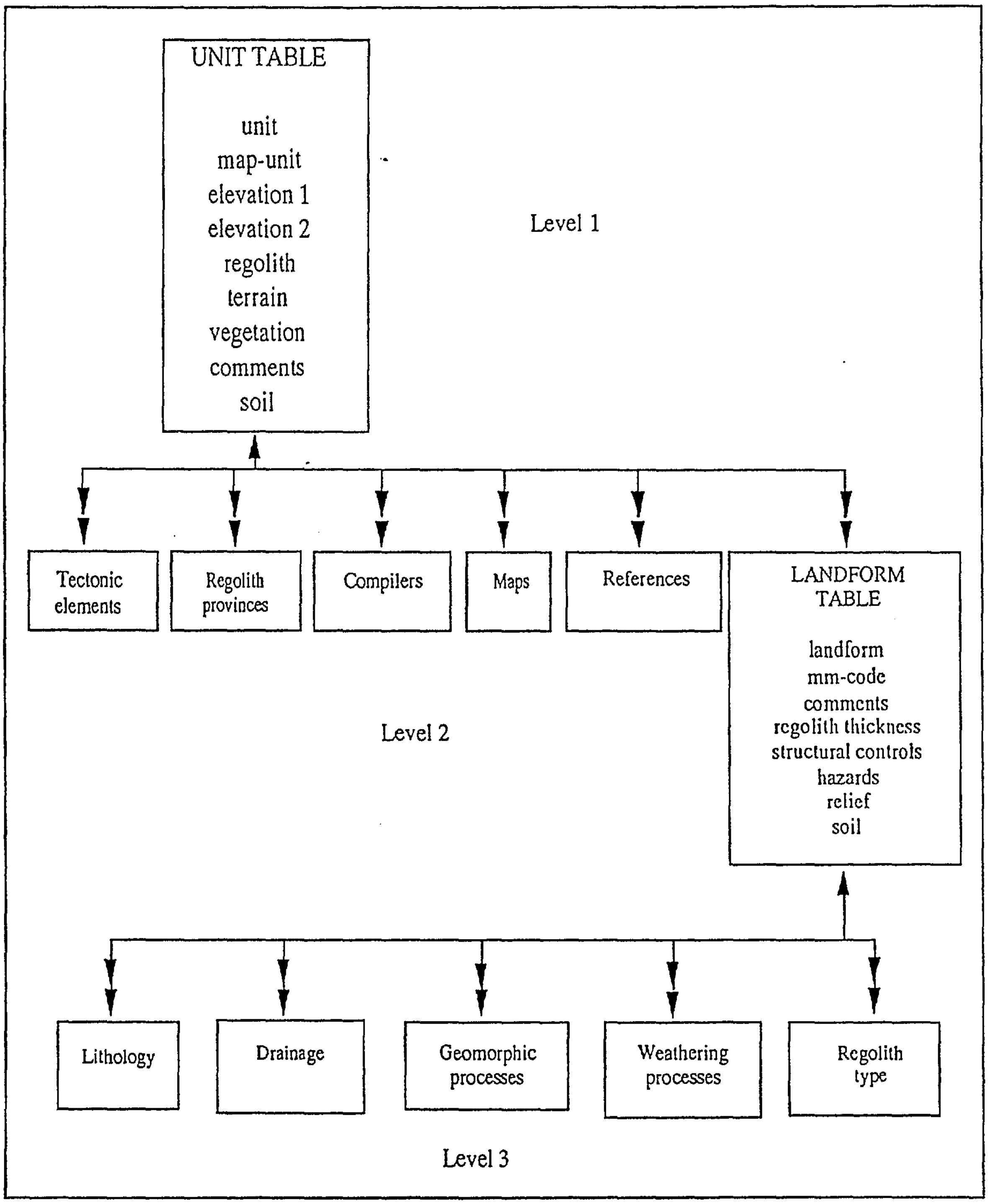


Figure 2. Simplified logical relationships of mapping unit data in RTMAP (from Pain *et al.* 1991). Double headed arrows indicate many-to-one relationships between tables.

Regolith terminology presented a problem. Unlike soil materials, regolith has not been classified in any systematic way, particularly for use at a regional mapping scale. Landform classification has reached the stage where, although there may be disagreements, most people generally agree on the major groupings. However, regolith classification is an entirely different matter. There is still disagreement about what regolith is, and this disagreement extends to the use of different terms to describe regolith types. A good example is the use of the words *laterite* and *ferricrete*.

In developing a classification of regolith we kept in mind its purpose. We are mapping regolith at a publication scale of 1:250,000, although field mapping is at 1:100,000. In most cases this rules out mapping specific regolith materials such as the mottled zone of a laterite profile. Regolith landform units will contain groupings of specific regolith materials. Often these groupings will be related both spatially and genetically, in the same way as the soils in a toposequence are related. The difference between a regolith "toposequence" and a soil toposequence is that the former is likely to be deeper and more complex. A group of regolith types will be a three dimensional entity which frequently contains a wide variety of specific materials.

The classification we use contains the basic regolith types, derived in part from McDonald *et al.* (1990). It will expand, particularly with the addition of categories of regolith profiles, as data come both from our field mapping program and from workers outside the BMR. We are also working with the CSIRO Division of Exploration Geoscience, with the aim of joining their detailed materials classification (unpublished) with our more general classification.

Protection of Terminology

Protection of terminology is a simple matter. Lookup tables in ORACLE can be changed only by those who have the right to do so. However, it is important that users agree on the classifications being used, and we are continually liaising with other workers in the general area of regolith studies to ensure that our classifications are acceptable to the broader geoscience community. For this reason RTMAP, and especially those parts related to regolith terminology, is evolving quite rapidly.

Hierarchical Coding

Although we have adopted large amounts of terminology from McDonald *et al.* (1990), we have not adopted their coding system. They coded their system by simply using the first two or three characters of each term. This means that their coding, and their classifications in general, are not hierarchical. This presents problems for database use, particularly interrogation. In most cases we have therefore changed their codes to allow grouping of attributes. For example, all erosional landforms have a code beginning ER, and all weathered *in situ* regolith types have a code beginning WI. To retrieve regolith landform units containing low hills, the specific code ER30 would be used. To retrieve regolith landform units with all kinds of erosional landforms, the code ER% would be used (% is a wildcard character in ORACLE).

Map legends

We are still dealing with the problem of how to display our information on maps. Unlike geology maps or even soil maps, there is no standard for regolith maps. As an interim measure we have adapted a system used in the New Zealand Land Resource Inventory. Symbols are used in the form of a fraction:

$$\frac{\text{regolith type + induration}}{\text{landform type + bedrock lithology}}$$

For example:

$$\frac{D1 + N1}{E0 + S}$$

is bauxitic nodules over saprolite on an erosional plain formed on sedimentary bedrock.

$$\frac{D2}{S + G}$$

is deep weathered structured saprolite on low hills formed on granite bedrock.

This system allows a great deal of information to be displayed on the map face in a flexible and easy-to-use way.

CONCLUSION

The use of standard terminology is obviously an advantage when results from one area must be compared and contrasted with those from another. This has long been recognised by those engaged in soil and land survey, and this recognition has been expressed recently in the *Australian Soil and Land Survey Handbook* (McDonald *et al.* 1990). The development of databases for regolith and environmental geoscience at BMR has meant the adoption of standard terminology for a number of attributes relating to landforms and regolith. We strongly suggest that, where relevant, RTMAP tables be used in other BMR databases.

REFERENCES

Lang, S.C., Withnall, I.W. and Grimes, K.G., 1990--Regional mapping field data management system (REGMAP) - guide to data collection and entry. *Queensland Department of Resource Industries, Geological Mapping Manual,* 1.

Lenz, S., 1991--RTMAP, BMR Regolith Database User's Manual. *Bureau of Mineral Resources Record* 1991/30.

McDonald, R.C., Isbell, R.F., Speight, J.G., Walker, J. and Hopkins, M.S., 1990-- *Australian Soil and Land Survey Field Handbook* (2nd Edition). Inkarta Press, Melbourne, 198pp.

Pain, C., Chan, R., Craig, M., Hazell, M., Kamprad, J. and Wilford, J., 1991--RTMAP BMR Regolith Database Field Handbook. *Bureau of Mineral Resources Record* 1991/29, 125pp.

Ryburn, R.J. and Lenz, S.L., 1991--Geoscientific relational databases in BMR and the "client-server" method, in *Proceedings of the National Conference on the Management of Geoscience Information and Data*, Adelaide, 22-25 July 1991.

COMPUTERISED GEOLOGICAL FIELD DATA MANAGEMENT IN QUEENSLAND - THE REGMAP SYSTEM

I.W. Withnall, K.G. Grimes, S.C. Lang & M. Thornton
Geological Survey of Queensland

Abstract - REGMAP has been used for recording field data by all GSQ regional mapping staff since 1986 and has recently been adopted by the BMR. It is a system for recording geological field notes in a structured format for entry into computerised database management systems and allows rapid, selective retrieval of data. The data can also be incorporated as a layer in the GIS environment.

Keywords - computers, data bases, information storage and retrieval

Reconnaissance mapping of the Queensland at 1:250 000 was completed in the early 1970's with the support of the Bureau of Mineral Resources. In 1984 the GSQ took over the responsibility of revising this geological mapping database. It was apparent that the field notes made during the 'first pass' mapping could be useful when revising the maps. If the notes were sufficiently detailed, it might not be necessary to revisit many areas, and field work could thus be better directed. However, it was found that these notes, if preserved and available, required considerable work to retrieve selected and meaningful data and this often precluded their use. It was a lot easier to start more or less from scratch, using only the published maps and synthesis reports as a basis. Another problem was how to take advantage of modern technology to rapidly retrieve, manipulate, and analyse the large amount of data collected as part of the more detailed second generation mapping, not only by individual geologists working on their own data, but groups of geologists being able to pool data. It was also realised that the field data should be compatible with geographic information systems (GIS) which were just starting to be developed.

After some experimentation and examining of systems used by other organisations, REGMAP was successfully trialled in 1986 (Lang & others, 1987), and thereafter adopted for use by all geological mapping teams. REGMAP uses standardised field notebooks. Data collected at each outcrop visited is divided into three basic types of information - site, structure, and lithology (Figure 1).

The site information records where the outcrop occurs spatially (map sheet, descriptive location, AMG coordinates etc.) and geologically (stratigraphic units, tectonic unit, and age). Province and rock unit names are coded, usually for simplicity by the first five letters of the name except where duplication occurs (Table 2); hierarchal codes are also used for geological ages. The orientations of various structural fabrics and palaeocurrents are recorded under the structural data. Non-numeric, descriptive structural information can also be recorded.

The major part of the database is the description of the rocks themselves, the lithological data. This data was the most difficult to adapt to a computerised system, because it is largely descriptive. For the system to overcome user resistance, it could not depend on too many codes, and had to be flexible enough to cope with the complete range of geological situations. The concept behind the system we adopted is that the description of each rock type at a field station or site can be broken down under a number of information types or data types. The data type field is the key to the flexibility of the system, avoiding the restrictions imposed by fixed fields or 'tick-the-box' forms. The geologist decides what types of information need to be recorded, aided by a prompt list of common 'data types' and their four-letter codes (Table 1). New data types can be added to the

Field no. SLD191 Date: 21/07/88

1:250 000 map: SE55/14 1:100 000 map: 8158 GR: 4349 78210

Descriptive location: 100 M SOUTH OF RD. TO 'DOTSWOOD' STN, IN SANDY CK. Geol. age: PZDL to PZCE

Province: BURDE Subprov: ______ Group: DOTSW Subgroup: ______

Formation: JULIA Member: ______ Informal unit: Duj

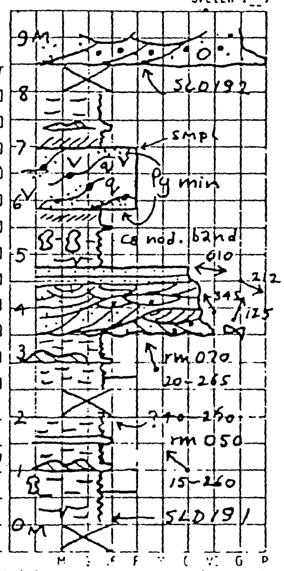
Exposure type: STREAM, GULLY, OUTCROP

Photo Name: DOTSW Run: 6 No.: 69 Scale: 25 Year: 86 Store: G

Sketch [✓]

Structural and palaeocurrent data

Rel	Data type	Dip/Azimuth Plunge/Azim	Young	Made vector
1	SØ	15/260		
?	SØ	40/250		
2	SØ	20/265		
	RM	15/260		050
	RM	20/265		070
	SSTR	20/265		125
	SSTR	20/265		080
	MSTR	20/265		212
	PL	20/265		010



Other structural data: DYKE TRENDS 148, ALONG GULLY, THEN UP RIDGE

General comments

[1] [DRFT] gyrd slst, pbbly Q c-vc sdst, rhlt

[ ] dyke with Py min. in Q vein

[VEG] SILVER LEAVED IRONBARK ON SEDIMENTS,

[ ] WATTLE SCRUB ON RHLT

[TOPO] DYKE FORMS RIDGE, TREND SE.

Rock class	Data type	Description
D	LITH	SLST
	COLR	GYRD (10R4/2)
	COMP	VOLCANICLASTIC ?
	GNSZ	SL-VF; OCC. F SAND STREAKS
	BEDG	TN-M
	IBED	LAM, GRADING IN PLACES
	XBED	RXL (WITH RM)
	SEDS	MUDCRACKS, CALCRETE NOD.
	REM	SLIGHTLY HORNFELSED BY DYKE
		AT 5·8 & 7M, SDST AT 3·5–4·75M.
	INTV	0–3·5M; 4·75–8·5M
	LFAC	Fm, Fl, minor P
D	LITH	SDST
	COLR	LTGY (N7)–PKGY (5YR 8/1)
	COMP	Q
	GNSZ	C-VC; PBBLY (UP TO 2CM)
	CLST	MAINLY QTZT, MINOR GRNT
	BEDG	TK-VTK
	IBED	FLAT LAM AT TOP, WITH PL
	XBED	SSTR & MSTR, MINOR MSPL
	SEDS	ERO. BASE, COSETS BOUNDED BY
		CONVEX-UP TOPS IN PLACES,
		MOSTLY EROSIVE. SHARP PLANAR TOP
	INTV	3·5–4·75M
	FOSV	RARE FISH BONE FRAGS ON XBEDS.
	SORT	GOOD
	LFAC	St, Sh, Minor Sp
B	LITH	RHLT
	COLR	PK
	GNSZ	F
	TEXT	PORPHYRITIC
	MIN	PY ASSOC. WITH Q VEINS
	VEIN	TN (1CM) Q VEINS, WITH PY
	SMPL	FOR T.S.

Carry over [✓] Sheet [1 of 2]

J 181—Govt Printer, Qld

Figure 1. Example of a completed REGMAP field sheet.

ABOR	ABORIGINAL RELICTS
ABST	ABSTRACT/SUMMARY
ALTN	ALTERATION
AMYG	AYMGDALES
AUGR	AUGER HOLE LOG
BEDG	BEDDING
BIBL	BIBLIOGRAPHIC REFERENCE
BMRR	BMR REG NO
BNDY	BOUNDARY DETAILS
CEMT	CEMENT
CHEM	CHEMICAL
CLEV	CLEAVAGES
CLST	CLASTS
CNTC	CONTACT DETAILS
COHR	COHERENCE
COLR	COLOUR
COMM	COMMODITIES
COMP	COMPOSITION
CREN	DESCRIPTION OF CRENULATIONS
DPDI	DEPOSIT DIMENSIONS
DPOR	DEPOSIT ORIENTATION
DRFT	DRAFTING ANNOTATIONS
DYKE	DETAILS OF DYKES
FAB	FABRIC
FACE	STRUCTURAL FACING
FACI	LITHOFACIES DESCRIPTION
FALT	FAULT DESCRIPTION
FOLD	TECTONIC FOLDS
FOSI	FOSSIL INVERTEBRATES
FOSM	MICROFOSSILS
FOSP	FOSSIL PLANTS
FOST	TRACE FOSSILS
FOSV	FOSSIL VERTEBRATES
FRCT	FRACTURE DESCRIPTION
GANG	GANGUE
GMAS	GROUNDMASS
GNSZ	GRAIN-SIZE
GOSN	GOSSAN
GSH	GRAIN SIZE HISTOGRAM
GSQD	GSQ RADIOMETRIC DATE NO.
GSQL	GSQ FOSSIL LOCALITY NO.
GSQR	GSQ REGISTERED ROCK NO.
GTYP	GRAINTYPE
HIST	HISTORICAL FEATURE
HORX	HOST ROCKS
HT	HEIGHT (OF CLIFF SECTION)
IBED	INTERNAL STRATIFICATION
IGFL	IGNEOUS FOLIATIONS
IGLN	IGNEOUS LINEATIONS
IGST	IGNEOUS STRUCTURES
INCL	INCLUSIONS
INTV	INTERVAL FOR MEASD SECTIONS
JOB	JOB DETAILS OR NO.
JTNG	JOINTS
KINK	DESCRIPTION OF KINK FOLDS
LAYR	LAYERING
LITH	LITHOLOGY
MAGN	MAGNETIC DATA
MAP	MAP DETAILS
MCOM	MINOR COMPONENTS
MEGX	MEGACRYSTS
META	METAMORPHIC GRADE
MIN	MINERALOGY
MINE	TYPE OF MINE WORKINGS
MINP	ORE MINERALS, PRIMARY
MINS	ORE MINERALS, SECONDARY
MTX	MATRIX
NAME	NAME OF MINE, TOWN ETC
NODS	NODULES
ORE	ORE TYPE
ORTX	ORE TEXTURES
OSTR	OTHER STRUCTURES
PERS	PERSONAL NOTES
PHNX	PHENOCRYSTS
PHOT	PHOTOGRAPH DETAILS
PORP	PORPHYROBLASTS
PORY	POROSITY
RAD	RADIOMETRIC VALUES
REGO	REGOLITH DETAILS
REM	REMARKS
RNAM	ROCK NAME (IN FULL)
ROND	ROUNDING
SECT	SECTION
SEDS	SEDIMENTARY STRUCTURES
SEQ	SEQUENCE
SIG	SIGNIFICANCE
SIZE	SIZE OF DEPOSIT
SMPL	SAMPLE DETAILS
SOIL	SOIL DETAILS
SORT	SORTING
SPHY	SPHERICITY
STAT	STATUS OF MINE/PROSPECT
STRU	STRUCTURAL AFILIATIONS
TCFL	TECTONIC FOLIATION
TCLN	TECTONIC LINEATION
TEXT	TEXTURE
TKNS	THICKNESS (OF SECTION)
TOPO	TOPOGRAPHIC DETAILS
VEG	VEGETATION DETAILS
VEIN	VEINS
VERG	VERGENCE DETAILS
VESI	DETAILS OF VESICLES
WETH	WEATHERING DETAILS
WIDE	WIDTH OF DYKE
WKMD	METHOD OF WORKING
WKXT	EXTENT OF WORKINGS
WSTR	WEATHERING STRUCTURE
XBED	CROSS BEDDING
XENL	XENOLITHS
XENX	XENOCRYSTS
XREF	CROSS-REFERENCE
XTAL	CRYSTALS IN TUFFS
YUNG	YOUNGING

Table 1. List of lithological data types used in REGMAP.

EXAMPLES OF PROVINCE CODES

CODES	DESC
ABERC	ABERCORN TROUGH
ADAVA	ADAVALE BASIN
ANAKI	ANAKIE INLIER
ATHER	ATHERTON BASALT PROVINCE
AUBUR	AUBURN ARCH
BARAM	BARAMBAH BASALT
BARNA	BARNARD SUBPROVINCE
BEAUD	BEAUDESERT BASIN
BEENL	BEENLEIGH BLOCK
BILOE	BILOELA BASIN
BOOVA	BOOVAL BASIN
BOWEN	BOWEN BASIN
BROKE	BROKEN RIVER EMBAYMENT
BULGO	BULGONUNNA VOLCANIC PROVINCE
BUNDA	BUNDABURG TROUGH
BUNDO	BUNDOCK BASIN
BURDE	BURDEKIN BASIN
BURKE	BURKE RIVER STRUCTURAL BELT
CALLI	CALLIOPE BLOCK
CAMEL	CAMEL CREEK SUBPROVINCE
CAMPW	CAMPWYN BLOCK
CAPEY	CAPE YORK-ORIOMO INLIER
CAPRI	CAPRICORN BASIN
CARPE	CARPENTARIA BASIN
CASUA	CASUARINA BASIN
CHILL	CHILLAGOE SUBPROVINCE
CHUDL	CHUDLEIGH PROVINCE
CLARK	CLARKE RIVER BASIN
CLERM	CLERMONT-SPRINGSURE BASALT PROVINCE
COAST	COASTAL BLOCK
COEN	COEN INLIER
COLLI	COLLINSVILLE SHELF
COMET	COMET PLATFORM
CONNO	CONNORS ARCH
COOPE	COOPER BASIN
CRAIG	CRAIGILLE BLOCK
CROYD	CROYDON SUBPROVINCE
CZCOV	CAINOZOIC COVER
CZVOL	CAINOZOIC VOLCANICS UNDIFFERENTIATED
DARGA	DARGALONG INLIER
DENIS	DENISON TROUGH
DRUMM	DRUMMOND BASIN
DUARI	DUARINGA BASIN
EROMA	EROMANGA BASIN
ESK T	ESK TROUGH
FORSA	FORSAYTH SUBPROVINCE
GALIL	GALILEE BASIN
GEORG	GEORGETOWN INLIER
GEORI	GEORGINA BASIN
GILBE	GILBERTON BASIN
GOGAN	GOGANGO OVERFOLDED ZONE
GRAVE	GRAVEYARD CREEK SUBPROVINCE

EXAMPLES OF ROCK-UNIT CODES

CODES	DESC
ABERD	ABERDARE CONGLOMERATE
ABERN	ABERNETHY BASALT
ADAVA	ADAVALE GROUP
ADDER	ADDER CREEK DACITE
ADORI	ADORI SANDSTONE
AGATD	AGATE DOWNS SILTSTONE
AGATE	AGATE CREEK VOLCANICS
AGE	AGE CREEK FORMATION
AGNES	AGNES WATER VOLCANICS
AIRLI	AIRLIE VOLCANICS
AIRPO	AIRPORT QUARTZ DIORITE
ALABA	ALABAMA DIORITE
ALBAN	ALBANY PASS BEDS
ALBER	ALBERT BASALT
ALDEB	ALDEBARAN SANDSTONE
ALLAN	ALLANDALE GRANITE
ALLAR	ALLARU MUDSTONE
ALLIN	ALLINGHAM FORMATION
ALLSO	ALLSORTS CREEK RHYOLITE
ALMAD	ALMADEN GRANITE
ALSAC	ALSACE QUARTZITE
ALTAN	ALTANMOUI GRANITE
ALUM	ALUM ROCK CONGLOMERATE
AMAMO	AMAMOOR BEDS
AMBER	AMBER
AMITY	AMITY APLITE
ANAKI	ANAKIE METAMORPHICS
ANNIN	ANNING GRANITE
ANSWE	ANSWER SLATE
AQUAR	AQUARIUM FORMATION
ARALB	ARALBA ADAMELLITE
ARAMA	ARAMAC COAL MEASURES
ARANB	ARANBANGA VOLCANIC GROUP
ARCAD	ARCADIA FORMATION
ARCH	ARCH CREEK LIMESTONE MEMBER
ARGEN	ARGENTINE METAMORPHICS
ARGYL	ARGYLLA FORMATION
ATLAN	ATLANTA GRANITE
AUSTR	AUSTRAL DOWNS LIMESTONE
BABAL	BABALANGEE AMPHIBOLITE
BACK	BACK CREEK GROUP
BADU	BADU GRANITE
BAGST	BAGSTOWE GRANITE
BAKER	BAKERVILLE GRANODIORITE
BALCO	BALCOOMA METAVOLCANICS
BALLA	BALLARA QUARTZITE
BALLY	BALLY KNOB VOLCANICS
BALNA	BALNAGOWAN VOLCANIC MEMBER
BAMFO	BAMFORD GRANITE
BANAN	BANANA FORMATION
BANCR	BANCROFT FORMATION
BANDA	BANDANNA FORMATION

AGE CODES

CODES	DESC
CZ	CAINOZOIC*
CZQ	CAINOZOIC*QUATERNARY
CZQH	CAINOZOIC*QUATERNARY*HOLOCENE
CZQP	CAINOZOIC*QUATERNARY*PLEISTOCENE
CZT	CAINOZOIC*TERTIARY
CZTL	CAINOZOIC*TERTIARY*LATE
CZTM	CAINOZOIC*TERTIARY*MIDDLE
CZTE	CAINOZOIC*TERTIARY*EARLY
CZP	CAINOZOIC*TERTIARY*PLIOCENE
CZM	CAINOZOIC*TERTIARY*MIOCENE
CZO	CAINOZOIC*TERTIARY*OLIGOCENE
CZE	CAINOZOIC*TERTIARY*EOCENE
CZA	CAINOZOIC*TERTIARY*PALEOCENE
MZ	MESOZOIC*
MZK	MESOZOIC*CRETACEOUS
MZKL	MESOZOIC*CRETACEOUS*LATE
MZKE	MESOZOIC*CRETACEOUS*EARLY
MZJ	MESOZOIC*JURASSIC
MZJL	MESOZOIC*JURASSIC*LATE
MZJM	MESOZOIC*JURASSIC*MIDDLE
MZJE	MESOZOIC*JURASSIC*EARLY
MZT	MESOZOIC*TRIASSIC
MZTL	MESOZOIC*TRIASSIC*LATE
MZTM	MESOZOIC*TRIASSIC*MIDDLE
MZTE	MESOZOIC*TRIASSIC*EARLY
PZ	PALAEOZOIC*
PZL	PALAEOZOIC*LATE
PZM	PALAEOZOIC*MIDDLE
PZE	PALAEOZOIC*EARLY
PZP	PALAEOZOIC*PERMIAN
PZPL	PALAEOZOIC*PERMIAN*LATE
PZPM	PALAEOZOIC*PERMIAN*MIDDLE
PZPE	PALAEOZOIC*PERMIAN*EARLY
PZC	PALAEOZOIC*CARBONIFEROUS
PZCL	PALAEOZOIC*CARBONIFEROUS*LATE
PZCE	PALAEOZOIC*CARBONIFEROUS*EARLY
PZD	PALAEOZOIC*DEVONIAN
PZDL	PALAEOZOIC*DEVONIAN*LATE
PZDM	PALAEOZOIC*DEVONIAN*MIDDLE
PZDE	PALAEOZOIC*DEVONIAN*EARLY
PZS	PALAEOZOIC*SILURIAN
PZSL	PALAEOZOIC*SILURIAN*LATE
PZSM	PALAEOZOIC*SILURIAN*MIDDLE
PZSE	PALAEOZOIC*SILURIAN*EARLY
PZO	PALAEOZOIC*ORDOVICIAN
PZOL	PALAEOZOIC*ORDOVICIAN*LATE
PZOE	PALAEOZOIC*ORDOVICIAN*EARLY
PZK	PALAEOZOIC*CAMBRIAN
PZKL	PALAEOZOIC*CAMBRIAN*LATE
PZKM	PALAEOZOIC*CAMBRIAN*MIDDLE
PZKE	PALAEOZOIC*CAMBRIAN*EARLY
PC	PRECAMBRIAN*
PCP	PRECAMBRIAN*PROTEROZOIC
PCPL	PRECAMBRIAN*PROTEROZOIC*LATE
PCPM	PRECAMBRIAN*PROTEROZOIC*MIDDLE
PCPE	PRECAMBRIAN*PROTEROZOIC*EARLY
PCA	PRECAMBRIAN*ARCHAEAN

Table 2. Examples of SITE codes used in REGMAP.

ADML	ADAMELLITE
AGLM	AGGLOMERATE
ALSK	ALASKITE
ALLV	ALLUVIUM
AMPH	AMPHIBOLITE
ANDS	ANDESITE
APLT	APLITE
ARNT	ARENITE
ARGL	ARGILLITE
ARKS	ARKOSE
BIF	BANDED-IRON-FORMATION
BSLT	BASALT
BDST	BOUNDSTONE
BREC	BRECCIA
BKFM	BROKEN FORMATION
CASI	CALC-SILICATE
CAAR	CALCARENITE
CALU	CALCILUTITE
CLRD	CALCIRUDITE
CACT	CALCRETE
CASI	CALCSILICATE
CACH	CALICHE
CATA	CATACLASITE
CHRT	CHERT
CRMT	CHROMITITE
CLST	CLAYSTONE
CPXT	CLINOPYROXENITE
CLLV	COLLUVIUM
CONC	CONCRETION
CGLM	CONGLOMERATE
COQT	COQUINITE
DACT	DACITE
DIAT	DIATOMITE
DIOR	DIORITE
DOLR	DOLERITE
DOLM	DOLOMITE
DUNT	DUNITE
DURI	DURICRUST
DYKE	DYKE DESCRIPTION
FSPO	FELDSPAR PORPHYRY
FLST	FELSITE
FECT	FERRICRETE
GBBR	GABBRO
GLSS	GLASS
GNSS	GNEISS
GOSS	GOSSAN
GOUG	GOUGE
GSTN	GRAINSTONE
GRNT	GRANITE
GRTD	GRANITOID
GRDI	GRANODIORITE
GRFL	GRANOFELS
GRNP	GRANOPHYRE
GRNL	GRANULITE

GRVL	GRAVEL
GNSC	GREENSCHIST
GNST	GREENSTONE
GREI	GREISEN
GYWK	GREYWACKE
GPSM	GYPSUM
HARZ	HARZBURGITE
HFLS	HORNFELS
IGNM	IGNIMBRITE
INTR	INTRUSIVE
FEST	IRONSTONE
JASP	JASPER
JSPT	JASPILITE
KPHR	KERATOPHYRE
LAMP	LAMPROPHYRE
LATR	LATERITE
LCGR	LEUCOGRANITE
LGRT	LEUCOGRANITOID
LGNT	LIGNITE
LMST	LIMESTONE
MRBL	MARBLE
MLNG	MELANGE
MAND	META-ANDESITE
MARN	META-ARENITE
MBAS	METABASALT
MDAC	METADACITE
MDIO	METADIORITE
MDLR	METADOLERITE
MGBR	METAGABBRO
MGYW	METAGREYWACKE
METS	METAMORPHICS
MPEL	METAPELITE
MPOR	METAPORPHYRY
MRHY	METARHYOLITE
MSED	METASEDIMENT
MSED	METASEDIMENTS
MTUF	METATUFF
MVOL	METAVOLCANIC
MCRT	MICRITE
MCDI	MICRODIORITE
MCGR	MICROGRANITE
MCGT	MICROGRANITOID
MCGD	MICROGRANODIORITE
MCMZ	MICROMONZONITE
MCSY	MICROSYENITE
MCTO	MICROTONALITE
MIGM	MIGMATITE
MODI	MONZODIORITE
MOGB	MONZOGABBRO
MOGR	MONZOGRANITE
MONZ	MONZONITE
MOSY	MONZOSYENITE
MDST	MUDSTONE
MYLN	MYLONITE
NEPH	NEPHELENITE

OBSD	OBSIDIAN
OSHL	OIL SHALE
OPXT	ORTHOPYROXENITE
OQTZ	ORTHOQUARTZITE
PKST	PACKSTONE
PAMP	PARA-AMPHIBOLITE
PEGM	PEGMATITE
PELT	PELITE
PRDT	PERIDOTITE
PERL	PERLITE
PHON	PHONOLITE
PHOS	PHOSPHORITE
PHYL	PHYLLITE
PCST	PITCHSTONE
PORT	PORPHYRITE
PORP	PORPHYRY
PSCG	PSEUDOCONGLOMERATE
PXNT	PYROXENITE
QDIO	QUARTZ-DIORITE
QFPO	QUARTZ-FELDSPAR PORPHYRY
QMDI	QUARTZ-MONZODIORITE
QMON	QUARTZ-MONZONITE
QZPO	QUARTZ-PORPHYRY
QSYN	QUARTZ-SYENITE
QTZT	QUARTZITE
RHDA	RHYODACITE
RHLT	RHYOLITE
SDST	SANDSTONE
SCHT	SCHIST
SERP	SERPENTINITE
SHLE	SHALE
SICT	SILCRETE
SLST	SILTSTONE
SKRN	SKARN
SLAT	SLATE
SPST	SOAPSTONE
SPIL	SPILITE
SGYW	SUBGREYWACKE
SYEN	SYENITE
SYGR	SYENOGRANITE
TILL	TILLITE
TONL	TONALITE
TAND	TRACHYANDESITE
TBAS	TRACHYBASALT
TRAC	TRACHYTE
TRAV	TRAVERTINE
TRON	TRONDHJEMITE
VEIN	VEIN DESCRIPTION
VOLC	VOLCANICS
WKST	WACKESTONE
XENL	XENOLITH

Table 3. Codes for rock types (LITHs) used in REGMAP.

system at any time to meet specialised needs. The description field for each 'data type' can contain as many lines as required and is essentially free text. The information is structured by the use of a special 'data type' called LITH, with a rock name in the description field. A set of standard codes is used for rock names (Table 3), or alternatively the name is spelt-out in full and subsequently coded by the computer. All subsequent data are taken to refer to that LITH until a further LITH is entered. The system is explained fully in the field manual (Lang & others, 1990).

Data is entered into IBM-compatible computers using a C program which produces an ASCII file which can be loaded into a database management system. Data entry is generally done on a PC at the field base camp or at head office, although laptops have also been used in the field. No attempt has been made to enter data directly at the outcrop, although this is feasible. We consider the hand-written field sheets to be the ultimate backup.

At present we use Foxbase as the database management system and have developed a range of programs to produce various kinds of selective retrieval and output. A User Manual explaining the use of these programs is available (Grimes & others, 1990). It is planned to transfer the database to ORACLE as part of MERLIN, the corporate GIS of the Department of Resource Industries.

Since 1986, data from all our field programs has been recorded using REGMAP and by the end of 1991, the database contained data from almost 30 000 field stations from about forty 1:100 000 map sheet areas. The system was quickly accepted by all geologists as a field recording system, and has resulted in major changes to work practises in the office. The system is now proving its worth as many of the projects are in the final write-up stage. Data describing attributes of particular rocks or units can be selectively retrieved and output as lists and in tabulated form without tedious manual searching of notebooks. The site information provides the main basis for selecting subsets of the data, but lithological data can also be searched using either the 'LITH', other 'data types', or keywords in the 'description' field. The system is particularly useful for searching the data of staff who have resigned before writing up their work. Output in graphical format is also aided by output of ASCII files in a variety of formats which can be read into in house or proprietary software to produce stereo and rose plots, scatter plots of palaeocurrent or structural data etc. Automatic plotting of structural data onto our final maps in ARC-INFO is also possible.

The most important feature lacking from REGMAP at present is an interactive graphical interface. We can output files for input into CAD packages such as AUTOCADD and Generic CaDD, but these are for display only and cannot be interrogated. One great advantage of the conversion to ORACLE will be that we will be able to link the database to graphical layers in ARC-INFO. Cheaper packages that interface graphics with a Foxbase-style database have not been evaluated because of the anticipated move to ORACLE and ARC-INFO.

REGMAP has great potential for 'training' in remote sensing interpretation. We have successfully trialled this on an ERDAS system on which we integrated REGMAP data and graphical map layers from ARC-INFO with the Aerodata multi-client geophysical data set from the Charters Towers area.

In addition to its use by GSQ field parties, the joint BMR-GSQ party in Cape York has used the system since the project began in 1990. The BMR now intends to use REGMAP in projects outside Queensland, initially in the Lachlan Fold Belt and Macarthur Basin. Interest has been shown by the Northern Territory Geological Survey, the New South Wales Geological Survey, and DSIR in New Zealand, and some universities.

Another advantage of REGMAP is that raw data, not just the final interpretation and syntheses as presented in maps and reports, can easily be distributed to other users, in particular the exploration industry. Access to such data should help reduce duplication of effort and minimise the environmental and cultural intrusion of exploration.

A marketing strategy for REGMAP data is being developed. The data from the Ebagoola 1:250 000 Sheet area in Cape York is being marketed as a complete package jointly with the BMR (von Gnielinski & Blewett, 1992). Other GSQ data may be sold on a 'per site' basis. The REGMAP software is provided free of charge, with a nominal charge for disks and manuals. The runtime package requires a minimum of 500 kb of hard disk space. An additional 600 kb of source code and documentation is also provided. The user would need to a copy of Foxbase or similar Xbase database software.

References

Lang, S.C., Murphy, P.R., Grimes, K.G. & Withnall, I.W., 1987. Regional Mapping Field Data Management System (REGMAP) - report on the trial period and the recommendations for full implementation. *Geological Survey of Queensland, Record*, 1987/13.

Lang, S.C., Withnall, I.W. & Grimes, K.G., 1990. Regional Mapping Field Data Management System (REGMAP) - Guide to data collection and entry. *Department of Resource Industries, Queensland, Geological Mapping Manual* 1.

Grimes, K.G., Withnall, I.W., Lang, S.C., Murphy, P.R., & Thornton, M.P., 1990. Regional Mapping Field Data Management System (REGMAP) - User's guide. *Department of Resource Industries, Queensland, Geological Mapping Manual* 2.

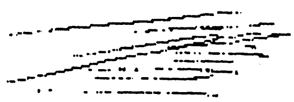
von Gnielinski, F.E. & Blewett, R.S., 1992. Regional Mapping Field Data Management System (REGMAP) data release for the Ebagoola 1:250 000 Sheet area, Coen Inlier, Cape York Peninsula, 1990-92. *Queensland Resource Industries, Record* 1992/3.

CR3/00/3
2 photos B&W of Bouma cycles
(a) ? | C B |

(b) C. B |

CR3/00/4
2 photos B&W of the bedded flysch (one general & one close-up)
also 2 colour.

Faults subparallel to bedding are present & give rise to slight irregular discordances



CR3/00/3
Mainly v thick (up to 5 m) m-c arenite, locally pebbly & qtz & v rare q-f-p pc. a few thinner interbeds of fine arenite : mudstone & lamina & ripple x-lamina (B C). Thick beds sometimes show grading & scoured bases. Abundant shale clasts in some arenite beds particularly at base.
S_0 024/78E - facing west

CR3/00/4 Marble Ck upstream of bridge.
Good exposure of thin-bedded flysch - alternating mudst & fine arenite c B C Bouma structures indicating W facing S_0 015/85W. C division is almost universal here so beds are either BC or C. Weak S_1 025/90. N: Sm
~~Ripple~~ Crests of ripples on bedding surfaces pitch 20°S : indicate currents from ~~+~~ west (current direction : pitches 70°N.
CR3/00/5
v thinly bedded, f. arenite & mudsilt
S_0 040/70E - no clear facing in cutting but in gutter about 100m W. thin-med bedded arenite has good X-lamina indicating E facing . current directions plunge → 120
V open folds & axial planes 160/45W ✓
CR3/00/6
Thick to v thick bedded massive poorly sorted arenite ?
& some pebbly zones of matrix supported (?)

Field no. SLD191 Date: 21/07/88

1:250 000 map: SE55/14 1:100 000 map: 8158 GR: 4349 78210

Descriptive location: 100 M SOUTH OF RD. TO 'DOTSWOOD' STN, IN SANDY CK.

Geol. age: PZDL to PZCE

Province: BURDE Subprov: ________ Group: DOTSW Subgroup: ________

Formation: JULIA Member: ________ Informal unit: Dqj

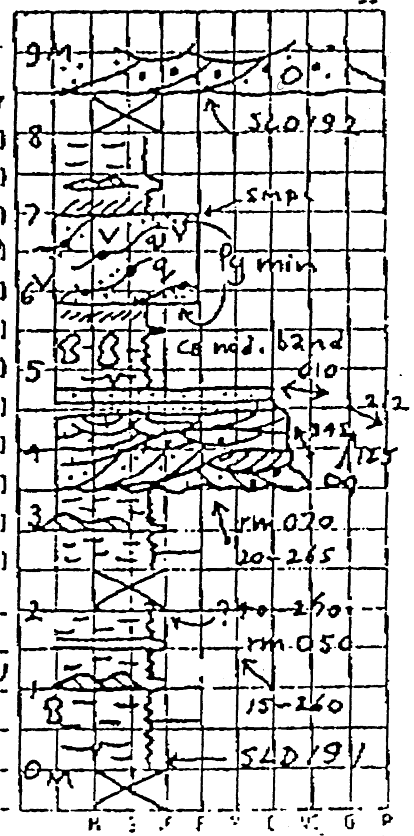
Exposure type: STREAM, GULLY, OUTCROP

Photo Name: DOTSW Run: 6 No.: 69 Scale: 25 Year: 86 Store: G

Sketch [4]

Structural and palaeocurrent data

Rel	Data type	Dip/Azimuth Plunge/Azim	Young	Made vector
1	SØ	15/260		
?	SØ	40/250		
2	SØ	20/265		
	RM	15/260		050
	RM	20/265		070
	SSTR	20/265		125
	SSTR	20/265		080
	MSTR	20/265		212
	PL	20/265		010
		/		
		/		
	DYKE	/148		



Other structural data: DYKE TRENDS 148, ALONG GULLY, THEN UP RIDGE

General comments

(1) DRFT	gyrd slst, pbbly Q c-vc sdst, rhlt
	dyke with Py min. in Q vein
VEG	SILVER LEAVED IRONBARK ON SEDIMENTS,
	WATTLE SCRUB ON RHLT
TOPO	DYKE FORMS RIDGE, TREND SE.

J 161—Govt. Printer, Qld.

Rock class	Data type	Description
D	LITH	SLST
	COLR	GYRD (10 R4/2)
	COMP	VOLCANICLASTIC ?
	GNSZ	SL-VF; OCC. F SAND STREAKS
	BEDG	TN-M
	IBED	LAM, GRADING IN PLACES
	XBED	RXL (WITH RM)
	SEDS	MUDCRACKS, CALCRETE NOD.
	REM	SLIGHTLY HORNFELSED BY DYKE
		AT 5.8 & 7M, SDST AT 3.5-4.75M
	INTV	0-3.5M; 4.75-8.5M
	LFAC	Fm, Fl, minor P
D	LITH	SDST
	COLR	LT GY (N7)-PK GY (5YR 8/1)
	COMP	Q
	GNSZ	C-VC; PBBLY (UP TO 2CM)
	CLST	MAINLY QTZT, MINOR GRNT
	BEDG	TK-VTK
	IBED	FLAT LAM AT TOP, WITH PL
	XBED	SSTR & MSTR, MINOR MSPL
	SEDS	ERO. BASE, COSETS BOUNDED BY
		CONVEX-UP TOPS IN PLACES,
		MOSTLY EROSIVE. SHARP PLANAR TOP
	INTV	3.5-4.75M
	FOSV	RARE FISH BONE FRAGS ON XBEDS.
	SORT	GOOD
	LFAC	St, Sh, minor Sp
B	LITH	RHLT
	COLR	PK
	GNSZ	F
	TEXT	PORPHYRITIC
	MIN	PY ASSOC. WITH Q VEINS
	VEIN	TN (1CM) Q VEINS, WITH PY
	SMPL	FOR T.S.

Carry over [✓] Sheet [1] of 2

Field no. S L D 1 9 1   Date: 21/07/88

1:250 000 map: SE55/14   1:100 000 map: 8158   GR: 4349 78210

Descriptive location: 100 M SOUTH OF RD. TO 'DOTSWOOD' STN, IN SANDY CK.   Geol. age: PZDL to PZCE

Province: BURDE   Subprov: ________   Group: DOTSW   Subgroup: ______

Formation: JULIA   Member: ________   Informal unit: Duj

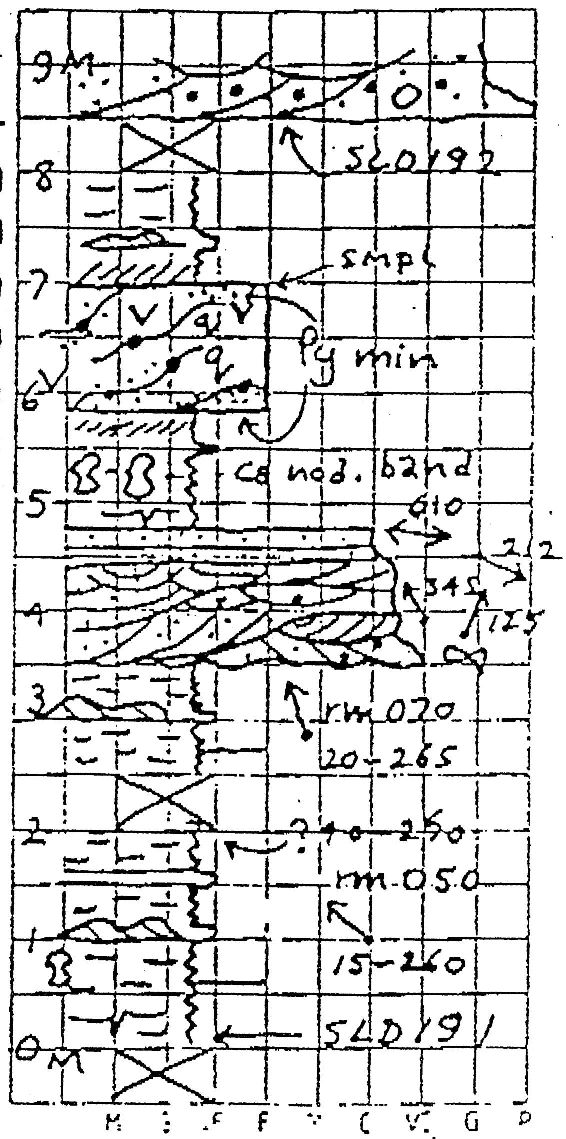
Exposure type: STREAM, GULLY, OUTCROP

Photo Name: DOTSW   Run: 6   No.: 69   Scale: 25   Year: 86   Store: G

Sketch [Y]

Structural and palaeocurrent data

Rel	Data type	Dip/Azimuth Plunge/Azim	Young	Made vector
1	SØ	15/260		
?	SØ	40/250		
2	SØ	20/265		
	RM	15/260		050
	RM	20/265		070
	SSTR	20/265		125
	SSTR	20/265		080
	MSTR	20/265		212
	PL	20/265		010
		/		
		/		



Other structural data:

DYKE TRENDS 148, ALONG GULLY, THEN UP RIDGE

General comments

[ ]	[DRFT]	gyrd slst, pbbly Q c-vc sdst, rhlt dyke with Py min. in Q vein
	[VEG]	SILVER LEAVED IRONBARK ON SEDIMENTS, WATTLE SCRUB ON RHLT
	[TOPO]	DYKE FORMS RIDGE, TREND SE.

EXAMPLES OF PROVINCE CODES

CODES	DESC
ABERC	ABERCORN TROUGH
ADAVA	ADAVALE BASIN
ANAKI	ANAKIE INLIER
ATHER	ATHERTON BASALT PROVINCE
AUBUR	AUBURN ARCH
BARAM	BARAMBAH BASALT
BARNA	BARNARD SUBPROVINCE
BEAUD	BEAUDESERT BASIN
BEENL	BEENLEIGH BLOCK
BILOE	BILOELA BASIN
BOOVA	BOOVAL BASIN
BOWEN	BOWEN BASIN
BROKE	BROKEN RIVER EMBAYMENT
BULGO	BULGONUNNA VOLCANIC PROVINCE
BUNDA	BUNDABURG TROUGH
BUNDO	BUNDOCK BASIN
BURDE	BURDEKIN BASIN
BURKE	BURKE RIVER STRUCTURAL BELT
CALLI	CALLIOPE BLOCK
CAMEL	CAMEL CREEK SUBPROVINCE
CAMPW	CAMPWYN BLOCK
CAPEY	CAPE YORK-ORIOMO INLIER
CAPRI	CAPRICORN BASIN
CARPE	CARPENTARIA BASIN
CASUA	CASUARINA BASIN
CHILL	CHILLAGOE SUBPROVINCE
CHUDL	CHUDLEIGH PROVINCE
CLARK	CLARKE RIVER BASIN
CLERM	CLERMONT-SPRINGSURE BASALT PROVINCE
COAST	COASTAL BLOCK
COEN	COEN INLIER
COLLI	COLLINSVILLE SHELF
COMET	COMET PLATFORM
CONNO	CONNORS ARCH
COOPE	COOPER BASIN
CRAIG	CRAIGILLE BLOCK
CROYD	CROYDON SUBPROVINCE
CZCOV	CAINOZOIC COVER
CZVOL	CAINOZOIC VOLCANICS UNDIFFERENTIATED
DARGA	DARGALONG INLIER
DENIS	DENISON TROUGH
DRUMM	DRUMMOND BASIN
DUARI	DUARINGA BASIN
EROMA	EROMANGA BASIN
ESK T	ESK TROUGH
FORSA	FORSAYTH SUBPROVINCE
GALIL	GALILEE BASIN
GEORG	GEORGETOWN INLIER
GEORI	GEORGINA BASIN
GILBE	GILBERTON BASIN
GOGAN	GOGANGO OVERFOLDED ZONE
GRAVE	GRAVEYARD CREEK SUBPROVINCE

EXAMPLES OF ROCK-UNIT CODES

CODES	DESC
ABERD	ABERDARE CONGLOMERATE
ABERN	ABERNETHY BASALT
ADAVA	ADAVALE GROUP
ADDER	ADDER CREEK DACITE
ADORI	ADORI SANDSTONE
AGATD	AGATE DOWNS SILTSTONE
AGATE	AGATE CREEK VOLCANICS
AGE	AGE CREEK FORMATION
AGNES	AGNES WATER VOLCANICS
AIRLI	AIRLIE VOLCANICS
AIRPO	AIRPORT QUARTZ DIORITE
ALABA	ALABAMA DIORITE
ALBAN	ALBANY PASS BEDS
ALBER	ALBERT BASALT
ALDEB	ALDEBARAN SANDSTONE
ALLAN	ALLANDALE GRANITE
ALLAR	ALLARU MUDSTONE
ALLIN	ALLINGHAM FORMATION
ALLSO	ALLSORTS CREEK RHYOLITE
ALMAD	ALMADEN GRANITE
ALSAC	ALSACE QUARTZITE
ALTAN	ALTANMOUI GRANITE
ALUM	ALUM ROCK CONGLOMERATE
AMAMO	AMAMOOR BEDS
AMBER	AMBER
AMITY	AMITY APLITE
ANAKI	ANAKIE METAMORPHICS
ANNIN	ANNING GRANITE
ANSWE	ANSWER SLATE
AQUAR	AQUARIUM FORMATION
ARALB	ARALBA ADAMELLITE
ARAMA	ARAMAC COAL MEASURES
ARANB	ARANBANGA VOLCANIC GROUP
ARCAD	ARCADIA FORMATION
ARCH	ARCH CREEK LIMESTONE MEMBER
ARGEN	ARGENTINE METAMORPHICS
ARGYL	ARGYLLA FORMATION
ATLAN	ATLANTA GRANITE
AUSTR	AUSTRAL DOWNS LIMESTONE
BABAL	BABALANGEE AMPHIBOLITE
BACK	BACK CREEK GROUP
BADU	BADU GRANITE
BAGST	BAGSTOWE GRANITE
BAKER	BAKERVILLE GRANODIORITE
BALCO	BALCOOMA METAVOLCANICS
BALLA	BALLARA QUARTZITE
BALLY	BALLY KNOB VOLCANICS
BALNA	BALNAGOWAN VOLCANIC MEMBER
BAMFO	BAMFORD GRANITE
BANAN	BANANA FORMATION
BANCR	BANCROFT FORMATION
BANDA	BANDANNA FORMATION

AGE CODES

CODES	DESC
CZ	CAINOZOIC*
CZQ	CAINOZOIC*QUATERNARY
CZQH	CAINOZOIC*QUATERNARY*HOLOCENE
CZQP	CAINOZOIC*QUATERNARY*PLEISTOCENE
CZT	CAINOZOIC*TERTIARY
CZTL	CAINOZOIC*TERTIARY*LATE
CZTM	CAINOZOIC*TERTIARY*MIDDLE
CZTE	CAINOZOIC*TERTIARY*EARLY
CZP	CAINOZOIC*TERTIARY*PLIOCENE
CZM	CAINOZOIC*TERTIARY*MIOCENE
CZO	CAINOZOIC*TERTIARY*OLIGOCENE
CZE	CAINOZOIC*TERTIARY*EOCENE
CZA	CAINOZOIC*TERTIARY*PALEOCENE
MZ	MESOZOIC*
MZK	MESOZOIC*CRETACEOUS
MZKL	MESOZOIC*CRETACEOUS*LATE
MZKE	MESOZOIC*CRETACEOUS*EARLY
MZJ	MESOZOIC*JURASSIC
MZJL	MESOZOIC*JURASSIC*LATE
MZJM	MESOZOIC*JURASSIC*MIDDLE
MZJE	MESOZOIC*JURASSIC*EARLY
MZT	MESOZOIC*TRIASSIC
MZTL	MESOZOIC*TRIASSIC*LATE
MZTM	MESOZOIC*TRIASSIC*MIDDLE
MZTE	MESOZOIC*TRIASSIC*EARLY
PZ	PALAEOZOIC*
PZL	PALAEOZOIC*LATE
PZM	PALAEOZOIC*MIDDLE
PZE	PALAEOZOIC*EARLY
PZP	PALAEOZOIC*PERMIAN
PZPL	PALAEOZOIC*PERMIAN*LATE
PZPM	PALAEOZOIC*PERMIAN*MIDDLE
PZPE	PALAEOZOIC*PERMIAN*EARLY
PZC	PALAEOZOIC*CARBONIFEROUS
PZCL	PALAEOZOIC*CARBONIFEROUS*LATE
PZCE	PALAEOZOIC*CARBONIFEROUS*EARLY
PZD	PALAEOZOIC*DEVONIAN
PZDL	PALAEOZOIC*DEVONIAN*LATE
PZDM	PALAEOZOIC*DEVONIAN*MIDDLE
PZDE	PALAEOZOIC*DEVONIAN*EARLY
PZS	PALAEOZOIC*SILURIAN
PZSL	PALAEOZOIC*SILURIAN*LATE
PZSM	PALAEOZOIC*SILURIAN*MIDDLE
PZSE	PALAEOZOIC*SILURIAN*EARLY
PZO	PALAEOZOIC*ORDOVICIAN
PZOL	PALAEOZOIC*ORDOVICIAN*LATE
PZOE	PALAEOZOIC*ORDOVICIAN*EARLY
PZK	PALAEOZOIC*CAMBRIAN
PZKL	PALAEOZOIC*CAMBRIAN*LATE
PZKM	PALAEOZOIC*CAMBRIAN*MIDDLE
PZKE	PALAEOZOIC*CAMBRIAN*EARLY
PC	PRECAMBRIAN*
PCP	PRECAMBRIAN*PROTEROZOIC
PCPL	PRECAMBRIAN*PROTEROZOIC*LATE
PCPM	PRECAMBRIAN*PROTEROZOIC*MIDDLE
PCPE	PRECAMBRIAN*PROTEROZOIC*EARLY
PCA	PRECAMBRIAN*ARCHAEAN

Rock class	Data type	Description
D	LITH	SLST
	COLR	GYRD (10R4/2)
	COMP	VOLCANICLASTIC ?
	GNSZ	SL-VF; OCC. F SAND STREAKS
	BEDG	TN-M
	IBED	LAM, GRADING IN PLACES
	XBED	RXL (WITH RM)
	SEDS	MUDCRACKS, CALCRETE NOD.
	REM	SLIGHTLY HORNFELSED BY DYKE
		AT 5·8 & 7M, SDST AT 3·5-4·75M.
	INTV	0-3·5m; 4·75-8·5m
	LFAC	Fm, Fl, minor P
D	LITH	SDST
	COLR	LT GY (N7)-PK GY (5YR 8/1)
	COMP	Q
	GNSZ	C-VC; PBBLY (UP TO 2CM)
	CLST	MAINLY QTZT, MINOR GRNT
	BEDG	TK-UTK
	IBED	FLAT LAM AT TOP, WITH PL
	XBED	SSTR & MSTR, MINOR MSPL
	SEDS	ERO. BASE, COSETS BOUNDED BY
		CONVEX-UP TOPS IN PLACES,
		MOSTLY EROSIVE. SHARP PLANAR TOP
	INTV	3·5-4·75M
	FOSV	RARE FISH BONE FRAGS ON XBEDS.
	SORT	GOOD
	LFAC	St, Sh, minor Sp
B	LITH	RHLT
	COLR	PK
	GNSZ	F
	TEXT	PORPHYRITIC
	MIN	PY ASSOC. WITH Q VEINS
	VEIN	TN (1CM) Q VEINS, WITH PY
	SMPL	FOR T.S.

Carry over [✓] Sheet [1 of 2]

REGMAP - PROMPT LIST

** PLUTONIC (CLASS A)
** HYPABYSSAL (CLASS B)

LITH	LITHOLOGY
COLR	COLOUR
GNSZ	GRAIN-SIZE
TEXT	TEXTURE
MIN	MINERALOGY
MCOM	MINOR COMPONENTS
MEGX	MEGACRYSTS
PHNX	PHENOCRYSTS
MTX	MATRIX
GMAS	GROUNDMASS
XENL	XENOLITHS
XENX	XENOCRYSTS
IGFL	IGNEOUS FOLIATIONS
IGLN	IGNEOUS LINEATIONS
IGST	IGNEOUS STRUCTURES
LAYR	LAYERING
FOLD	TECTONIC FOLDS
TCFL	TECTONIC FOLIATION
TCLN	TECTONIC LINEATION
VEIN	VEINS
REM	REMARKS

** VOLCANIC (CLASS C)

LITH	LITHOLOGY
COLR	COLOUR
GNSZ	GRAIN-SIZE
COMP	COMPOSITION
TEXT	TEXTURE
PHNX	PHENOCRYSTS
CLST	CLASTS
GMAS	GROUNDMASS
MTX	MATRIX
MEGX	MEGACRYSTS
XENX	XENOCRYSTS
XENL	XENOLITHS
MIN	MINERALOGY
MCOM	MINOR COMPONENTS
IGFL	IGNEOUS FOLIATIONS
IGLN	IGNEOUS LINEATIONS
IGST	IGNEOUS STRUCTURES
BEDG	BEDDING
IBED	INTERNAL STRATIFICATION
XBED	CROSS BEDDING/STRATIFICATION
SEDS	SEDIMENTARY STRUCTURES
FAB	FABRIC
SORT	SORTING
YUNG	YOUNGING
CEMT	CEMENT
CHEM	CHEMICAL
COHR	COHERENCE
ROND	ROUNDING
SPHY	SPHERICITY
FOLD	TECTONIC FOLDS
TCFL	TECTONIC FOLIATION
TCLN	TECTONIC LINEATION
CLEV	CLEAVAGES
FOSI	FOSSIL INVERTEBRATES
FOSP	FOSSIL PLANTS
FOSV	FOSSIL VERTEBRATES
FOST	TRACE FOSSILS
VEIN	VEINS
REM	REMARKS

** SEDIMENTARY (CLASS D)

LITH	LITHOLOGY
COLR	COLOUR
COMP	COMPOSITION
MCOM	MINOR COMPONENTS
GNSZ	GRAIN-SIZE
SORT	SORTING
CLST	CLASTS
GTYP	GRAINTYPE
MTX	MATRIX
ROND	ROUNDING
SPHY	SPHERICITY
FAB	FABRIC
BEDG	BEDDING
IBED	INTERNAL STRATIFICATION
XBED	CROSS BEDDING/STRATIFICATION
SEDS	SEDIMENTARY STRUCTURES
CEMT	CEMENT
YUNG	YOUNGING
CHEM	CHEMICAL
COHR	COHERENCE
PORE	POROSITY
FOSI	FOSSIL INVERTEBRATES
FOSP	FOSSIL PLANTS
FOSV	FOSSIL VERTEBRATES
FOST	TRACE FOSSILS
FOSM	MICROFOSSILS
FOLD	TECTONIC FOLDS
TCFL	TECTONIC FOLIATION
TCLN	TECTONIC LINEATION
CLEV	CLEAVAGES
FACE	STRUCTURAL FACING
VEIN	VEINS
REM	REMARKS

** METAMORPHIC
** CONTACT (CLASS E)
** REGIONAL (CLASS F)
** DYNAMIC (CLASS G)

LITH	LITHOLOGY
COLR	COLOUR
MIN	MINERALOGY
TEXT	TEXTURE
PORP	PORPHYROBLASTS
GNSZ	GRAIN-SIZE
MCOM	MINOR COMPONENTS
LAYR	LAYERING
TCFL	TECTONIC FOLIATION
TCLN	TECTONIC LINEATION
CLEV	CLEAVAGES
FOLD	TECTONIC FOLDS
YUNG	YOUNGING
FACE	STRUCTURAL FACING
VEIN	VEINS
REM	REMARKS

FOR RELICT FEATURES, PROMPT CODES FROM ANY OTHER CLASS CAN BE USED

** OTHER DATA TYPES

SMPL	SAMPLE DETAILS
GSQL	GSQ FOSSIL LOCALITY NO.
GSQD	GSQ RADIOMETRIC DATE NO.
GSQR	GSQ REGISTERED ROCK/SLIDE NO.
PHOT	PHOTOGRAPH DETAILS
BNDY	BOUNDARY DETAILS
SOIL	SOIL DETAILS
TOPO	TOPOGRAPHIC DETAILS
VEG	VEGETATION DETAILS
WETH	WEATHERING DETAILS
TKNS	THICKNESS (OF SECTION)
HT	HEIGHT (OF CLIFF SECTION)
SECT	SECTION
SEQ	SEQUENCE
MAP	MAP DETAILS
JOB	JOB DETAILS OR NO.
GREF	GRID REFERENCE SIX FIG. OR YD

Figure 2. Prompt sheet for REGMAP 'data types'

ADML	ADAMELLITE
AGLM	AGGLOMERATE
ALSK	ALASKITE
ALLV	ALLUVIUM
AMPH	AMPHIBOLITE
ANDS	ANDESITE
APLT	APLITE
ARNT	ARENITE
ARGL	ARGILLITE
ARKS	ARKOSE
BIF	BANDED-IRON-FORMATION
BSLT	BASALT
BDST	BOUNDSTONE
BREC	BRECCIA
BKFM	BROKEN FORMATION
CASI	CALC-SILICATE
CAAR	CALCARENITE
CALU	CALCILUTITE
CLRD	CALCIRUDITE
CACT	CALCRETE
CASI	CALCSILICATE
CACH	CALICHE
CATA	CATACLASITE
CHRT	CHERT
CRMT	CHROMITITE
CLST	CLAYSTONE
CPXT	CLINOPYROXENITE
CLLV	COLLUVIUM
CONC	CONCRETION
CGLM	CONGLOMERATE
COQT	COQUINITE
DACT	DACITE
DIAT	DIATOMITE
DIOR	DIORITE
DOLR	DOLERITE
DOLM	DOLOMITE
DUNT	DUNITE
DURI	DURICRUST
DYKE	DYKE DESCRIPTION
FSPO	FELDSPAR PORPHYRY
FLST	FELSITE
FECT	FERRICRETE
GBBR	GABBRO
GLSS	GLASS
GNSS	GNEISS
GOSS	GOSSAN
GOUG	GOUGE
GSTN	GRAINSTONE
GRNT	GRANITE
GRTD	GRANITOID
GRDI	GRANODIORITE
GRFL	GRANOFELS
GRNP	GRANOPHYRE
GRNL	GRANULITE
GRVL	GRAVEL
GNSC	GREENSCHIST
GNST	GREENSTONE
GREI	GREISEN
GYWK	GREYWACKE
GPSM	GYPSUM
HARZ	HARZBURGITE
HFLS	HORNFELS
IGNM	IGNIMBRITE
INTR	INTRUSIVE
FEST	IRONSTONE
JASP	JASPER
JSPT	JASPILITE
KPHR	KERATOPHYRE
LAMP	LAMPROPHYRE
LATR	LATERITE
LCGR	LEUCOGRANITE
LGRT	LEUCOGRANITOID
LGNT	LIGNITE
LMST	LIMESTONE
MRBL	MARBLE
MLNG	MELANGE
MAND	META-ANDESITE
MARN	META-ARENITE
MBAS	METABASALT
MDAC	METADACITE
MDIO	METADIORITE
MDLR	METADOLERITE
MGBR	METAGABBRO
MGYW	METAGREYWACKE
METS	METAMORPHICS
MPEL	METAPELITE
MPOR	METAPORPHYRY
MRHY	METARHYOLITE
MSED	METASEDIMENT
MSED	METASEDIMENTS
MTUF	METATUFF
MVOL	METAVOLCANIC
MCRT	MICRITE
MCDI	MICRODIORITE
MCGR	MICROGRANITE
MCGT	MICROGRANITOID
MCGD	MICROGRANODIORITE
MCMZ	MICROMONZONITE
MCSY	MICROSYENITE
MCTO	MICROTONALITE
MIGM	MIGMATITE
MODI	MONZODIORITE
MOGB	MONZOGABBRO
MOGR	MONZOGRANITE
MONZ	MONZONITE
MOSY	MONZOSYENITE
MDST	MUDSTONE
MYLN	MYLONITE
NEPH	NEPHELENITE
OBSD	OBSIDIAN
OSHL	OIL SHALE
OPXT	ORTHOPYROXENITE
OQTZ	ORTHOQUARTZITE
PKST	PACKSTONE
PAMP	PARA-AMPHIBOLITE
PEGM	PEGMATITE
PELT	PELITE
PRDT	PERIDOTITE
PERL	PERLITE
PHON	PHONOLITE
PHOS	PHOSPHORITE
PHYL	PHYLLITE
PCST	PITCHSTONE
PORT	PORPHYRITE
PORP	PORPHYRY
PSCG	PSEUDOCONGLOMERATE
PXNT	PYROXENITE
QDIO	QUARTZ-DIORITE
QFPO	QUARTZ-FELDSPAR PORPHYRY
QMDI	QUARTZ-MONZODIORITE
QMON	QUARTZ-MONZONITE
QZPO	QUARTZ-PORPHYRY
QSYN	QUARTZ-SYENITE
QTZT	QUARTZITE
RHDA	RHYODACITE
RHLT	RHYOLITE
SDST	SANDSTONE
SCHT	SCHIST
SERP	SERPENTINITE
SHLE	SHALE
SICT	SILCRETE
SLST	SILTSTONE
SKRN	SKARN
SLAT	SLATE
SPST	SOAPSTONE
SPIL	SPILITE
SGYW	SUBGREYWACKE
SYEN	SYENITE
SYGR	SYENOGRANITE
TILL	TILLITE
TONL	TONALITE
TAND	TRACHYANDESITE
TBAS	TRACHYBASALT
TRAC	TRACHYTE
TRAV	TRAVERTINE
TRON	TRONDHJEMITE
VEIN	VEIN DESCRIPTION
VOLC	VOLCANICS
WKST	WACKESTONE
XENL	XENOLITH

Field no. [1WWV 0056] Date: 19 : 7 : 86

1:250 000 map: SE55 13 1:100 000 map: 7857 GR: ________

Descriptive location: DRY RIVER 5 KM UPSTREAM OF GREGORY RIVER

Geol. age: PCP to ________

Province: LOLWO Subprov: ________ Group: ________ Subgroup: ________

Formation: CAPE Member: ________ Informal unit: ________

Exposure type: OUTC

Photo Name: WAND. Run: 8 No.: 96 Scale: 25 Year: 76 Store: G

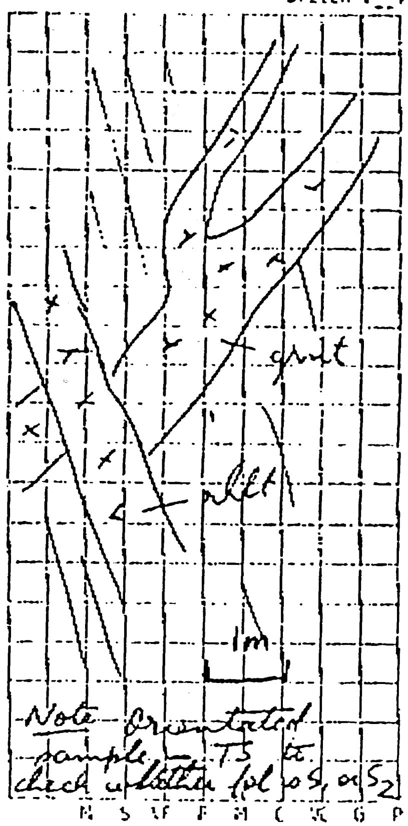
Sketch [✓]

Structural and palaeocurrent data

Pal	Data type	Dip/Azimuth Plunge/Axis	Young	Mode vector
[ ]	S1	56/075		
[ ]	S2	85/027		
[ ]	F21	55/115		
[ ]	L11	56/075		335
[ ]		/		
[ ]		/		
[ ]		/		
[ ]		/		
[ ]		/		
[ ]		/		
[ ]		/		
[ ]		/		

Other structural data: RHLT

DYKES TREND NW



General comments

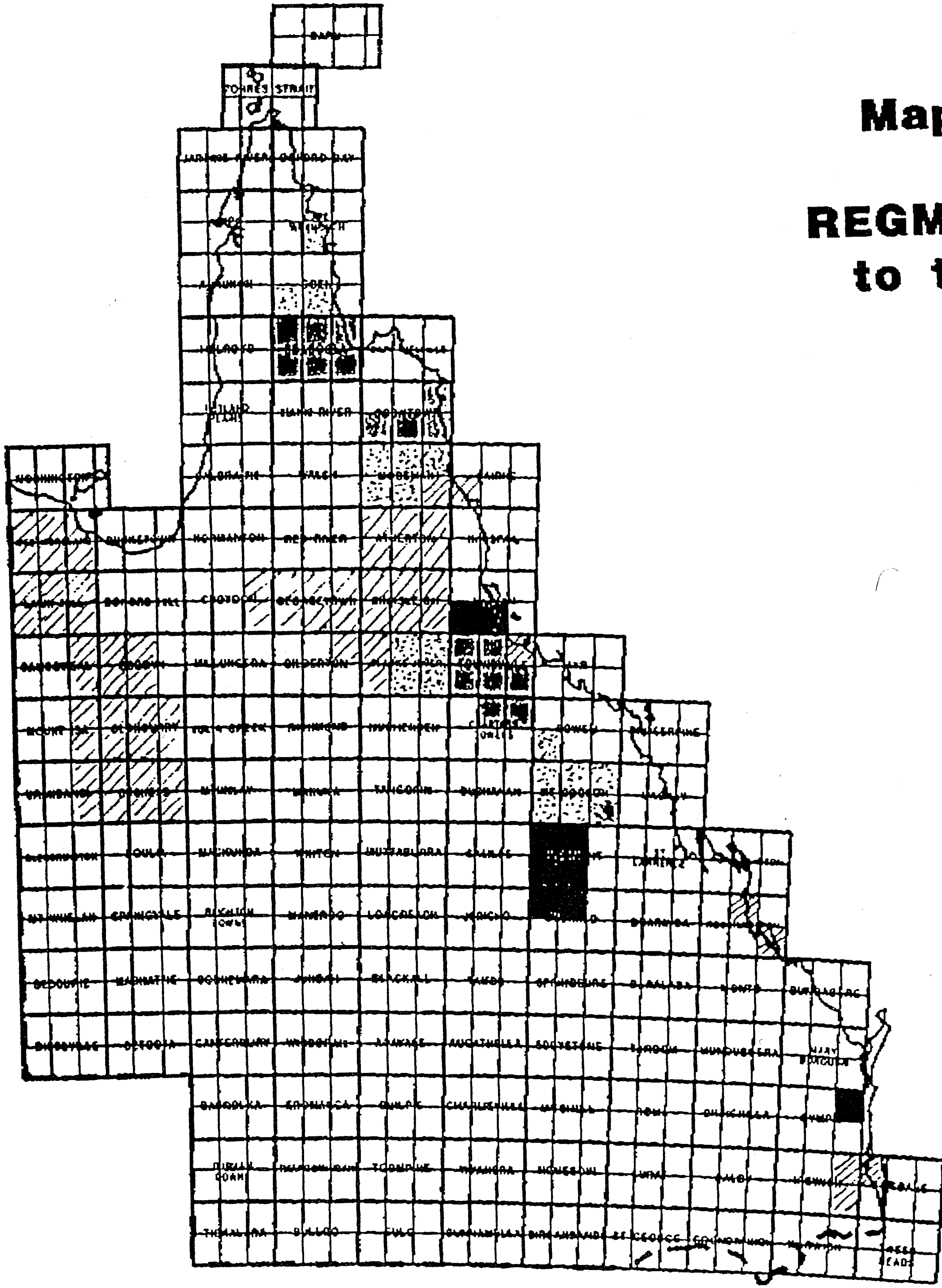
[TOPO]	DYKES FORM PROMINENT
[ ]	RIDGELINE TO WEST OF RIVER
[ABST]	gy bi gnss, mu-bi schst & qtzt
[ ]	w/ grnt veins
[ ]	

Rock class	Data type	Description
(DF)	LITH	GNSS
	COLR	GY
	GNSZ	REL. C.
[ ]	TCFL	STRONG WITH CROSS-CUTTING
	[ ]	CREN RELATED TO TIGHT
	[ ]	F21 FOLDS
[ ]	FOLD	F21 FOLDS ARE ~~[struck out]~~ REL
	[ ]	TIGHT
	REM	SIMILAR TO GNSS FURTHER
[ ]	[ ]	E AROUND BLACK KNOB
	MIN	BIOTITE - MUSCOVITE
	LAYR	THIN BANDS UP TO 5CM
(DF)	LITH	SCHIST
	MIN	MUSCOVITE - BIOTITE
	PORP	MUSCOVITE TO 1CM AFTER
[ ]	[ ]	SILL, CUTTING ACROSS FOL
	TCFL	STRONG WITH DIFFERENTIATED
	[ ]	CREN CLEAVAGE AXIAL PLANE
[ ]	[ ]	TO F21 FOLDS
	COLR	BUFF
	SMPL	ORIENTATED 45/060
(DF)	LITH	QUARTZITE
	MIN	MICACEOUS
	MCOM	SPOTS OF CALC-SILS
[ ]	LAYR	1-2M THICK
	REM	INTERLAYERED WITH GNSS
	[ ]	AND SCHIST
[ ]	[ ]	
	[ ]	
	[ ]	
(A)	LITH	GRNT
	COLR	WH - PK
	TEXT	EQ - PQ

Carry over [✓] Sheet [1 of 2]

Figure 4. Example field sheet with carry over

Map of Queensland showing REGMAP data coverage to the end of 1991.



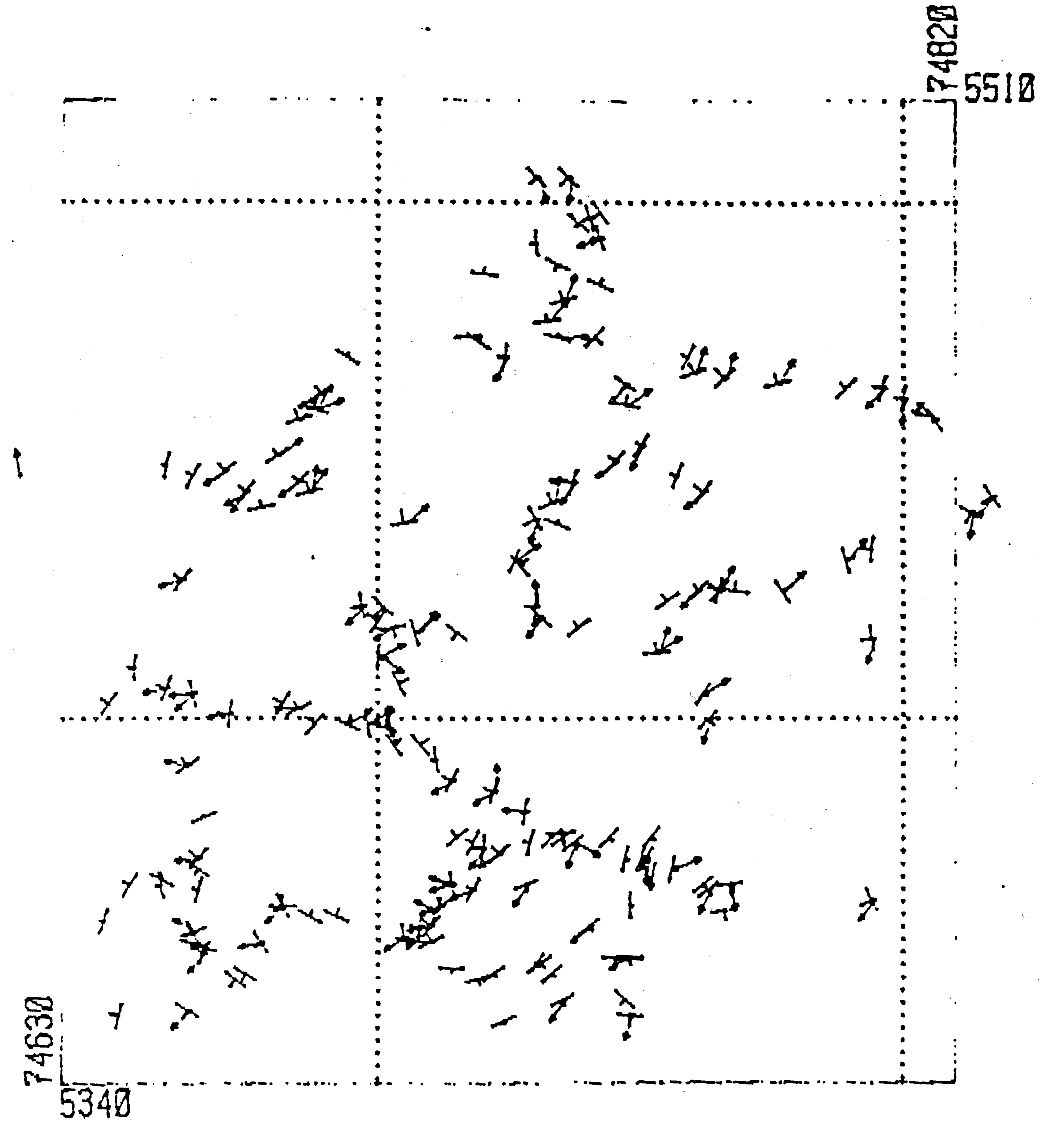
- FIELD WORK COMPLETED
- FIELD WORK COMPLETED (partial REGMAP coverage only)
- FIELD WORK IN PROGRESS
- Second pass mapping (pre REGMAP)

Example 2.

Horizontal format with first 180 characters of data (Option 2 from Submenu within program) using same data_types and filters as in Example 1. Note that semicolons are omitted; causing a wraparound to the next line instead.

ARENITES AND CONGLOMERATES - EWAN BEDS

FIELDNO	LITHOLOGY	COMPOSITION	CLAST TYPE	GRAIN-SIZE	BEDDING	INTERNAL BEDDING
IWT134	CGLM	POLYMICTIC	MAINLY QTZT, SOME Q	PB - CB	VTK	MASS
	ARNT	L	Q, QTZT, BSLT?	C - VC, PBLY	TN - VTK	MASS
IWT136A	CGLM	POLYMICTIC, V. MICACEOUS	Q, MICA SCHIST, QTZT	GL-PB	?, MAINLY FLOATERS ON S SIDE OF GULLY - SOME O/C NEAR HEAD	
IWT145	ARNT	L	Q, SCHIST, CHERT, QTZT	VC, PB	TK	MASSIVE
IWT151	ARNT	L		M-VC, PBLY	TK	TK LAM GRADED LOCALLY
	CGLM	POLYMICTIC, LITHIC	Q ARNT, MINOR Q	PB	TK-M	MASSIVE, ALIGNED PB
IWT163	CGLM	POLYMICTIC	MAINLY Q ARNT, SOME Q & SHALE Q ARNT IS VEINED	PB-CB	TK-VTK	MASSIVE



Anakie Metamorphics - Oaky Creek area

↑ L21 (1Ø4)
⊥ S2 (159)

Spatial plot of selected structural data imported from REGMAP into GENERIC CADD

1990 Field Season Samples

Field No	1:100 000	Grid Ref	Formation	Informal Unit	Lithology	Description
1WCL001	8353	177535	ANAKI		ARNT	Q ARNT (A)
					MDST/PHYLL	CLEAVED PHYLL MOST (B)
1WCL002	8353	289329	ANAKI		SCHT	SPOTTED SCHIST
1WCL005	8452	572875	ANAKI		ARNT	ARNT W/ DISSOLUTION CLEV (A)
1WCL006	8452	575870	ANAKI		ARNT	ARNT W/ LAYR (A)
					ARNT	ARNT W/ FOLDED LAYR (B)
1WCL009	8452	591802	ANAKI		META-ARNT	FOLD HINGE
1WCL010	8352	462895	ANAKI		ARNT	Q ARNT
1WCL012	8352	407899	ANAKI		PHYL	PHYLL ORIENT 45 -> 260
1WCL014	8352	399878	ANAKI		PHYL	PHYLLITE
						WITH ISOCLINALLY FOLDED Q VEINLETS (FLOAT - NOT ORIENTATED)
1WCL015	8352	369872	ANAKI		QTZT	QTZT
1WCL017	8452	513740	ANAKI		PHYL	PHYLL - ORIENTED 25 -> 210
1WCL018	8352	494733	ANAKI		GNST	GREENSTONE (GEOCHEM) (A)
1WCL019	8352	489731	ANAKI		GNST	MASSIVE GNST (A), FOL GNST (B) - BOTH GEOCHEM?
1WCL020	8352	465725	ANAKI		SCHT	SCHT ORIENT 32 -> 310
1WCL021	8352	430722	ANAKI		SCHT	SCHIST (NOT ORIENTATED) - SLAB
1WCL023	8352	415717	ANAKI		GNST	GREENSTONE (GEOCHEM?)
1WCL025	8352	388699	ANAKI		MRBL	MARBLE
1WCL026	8352	341717	ANAKI		ARNT	Q ARNT W/ DISSOLUTION SEAMS
					PHYLL/SCHT	F SCHIST W/ FINE DIFFD LAM
1WCL027	8452	634744	ANAKI		PHYLL/SCHT	PHYLL/SCHT ORIENT 45 -> 170
1WCL028	8452	614714	ANAKI		FEST	FEST (HEMATITE) (A) - FEST (MAGNETITE W/ SULPHIDES (B)
					SCHT	CHL SCHT (C)
					BIF/QTZT	BIF (D)
1WCL029	8452	607696	ANAKI		QTZT	FOL QTZT (NOT ORIENTATED)
1WCL031	8452	0 0	THERE		ANDS	ANDS
1WCL032	8452	575586	RETRE		MONZ?	M EQ MONZ (GEOCHEM) (A)
					MCMZ	PORPH MICROMONZ
1WCL033	8452	940057	ANAKI		GNSS	SEVERAL ORIENTATED SAMPLES FOR SLABBING
1WCL034	8452	947056	ANAKI		CASI?	CALC-SILICATE (A)
					GBBR?	GABBRO?
					GNSS	GNSS

Listing of sample information

Rapid Field Data Capture Techniques

By: Graeme Mitchell
CRA Exploration Pty Ltd
Western / Central Region
20th March, 1992

Abstract

With the demands for digital data increasing at an exponential rate, methods need to be implemented to capture data digitally at the source. This provides the data where the users most require it - "on the spot" in the field. Decisions can then be made on further follow-up work or other courses of action to facilitate a more efficient operation. The data types include drilling, stream sediment / soil sampling, real-time "digitising" of lineal features - roads streams etc via field data recorders, GPS and GIS, and remote sensing techniques.

Introduction

There are increasing demands for more cost effective data capture techniques to improve the efficiency of field operations in the highly competitive 90's. New methods are needed for faster, more cost effective data acquisition at higher accuracy's, providing data in a format suitable for efficient in-field processing. The issues include improvement in field observation recording methods, 3-D location requirements, data verification and in-field processing.

Problem Definition

The trend in recent years is for more data capture and processing to be done in the field. This is driven by the need for more efficient utilisation of available funds and the technological advancements that allow in-field data processing. Digital data capture is inherently expensive due to both the equipment / technology costs and the man-power intensive nature of the operation. Land access is becoming more of a problem, and hence non-intrusive data capture techniques must be investigated.

Data Capture Requirements

Locational Accuracy

The locational accuracy of field observations is generally a function of the observation type, the capture cost and the map scale at which the data will be used. Whilst it would be preferable to have all locations to millimetre accuracy's, the cost of achieving that would be prohibitive. In reality, a regional sampling survey may only need horizontal locations accurate to SO - 100 metres, but greater accuracy would be required for a RAB drilling program on a 100 x 50 metre grid. Gravity readings may only require 10-20 metre accuracy's horizontally, but require 0.5-1.0 metre accuracy's vertically. The technique used therefore is dependent entirely on the accuracy requirements.

Flexible Capture Mechanisms

In a geotechnical environment, there is a wide range of field applications that data capture / logging software needs to address. The software must be sufficiently flexible to satisfy the numerous data types encountered, without the need to write another program for each new field operation. One effective method is to have a program "shell" that can be used to tailor a data capture exercise exactly to the geologist's requirements. The program "shell" would require the definition of the data fields to be recorded, the screen prompts the program would present to the user, and any on-line help that will be available during the operation of the program. Such a system gives the geologist the flexibility needed, without having to compromise his data recording requirements. A data capture system needs the ability to easily revi ew and edit the data entered, as data errors are inevitable.

Verification
It is essential that data verification is done in the field, where any errors can be quickly corrected. Verification can take the form of range and validation checks at capture time, repon generation at the completion of a capture program, or plotting of the data collected. Other verification procedures can include the routine generation of statistics on the data collected to see if any outliers exist.

Storage Requirements
Rapid field data capture normally results in high volumes of data collected, particularly in GPS data collection systems. A lot of satellite ephemeris information is recorded for post processing of differential GPS position calculations. This requires larger capacity storage facilities together with data management tools.

Processing Requirements

Efficient Processing

Efficient processing tools need to be made available to the field operator to manage the volumes of data collected. Improvements are required in the area of differential GPS position calculations to enable non-surveying staff to easily calculate the accurate locations.

Standardised Formats
To improve data management capabilities, it is important that some standards are defined for data storage on field computers. Databases such as Oracle, Dbase or Foxbase provide this management / storage functionality. Another alternative is to define a fileformat that describes the contents of the file in the header. Such a system provides all the documentation necessary to identify the contents and source of the data contained within the file. This, together with a well organised directory structure is necessary to properly manage the large number of files that can be generated.

Data Integration
Format standards are also important in the area of data integration. A range of different programs may need access to the data collected, and hence must be able to recognise the format of the file. The uploading of this field data into central databases can also take advantage of this format standardisation.

Data Types

Rock Samples
The data capture requirements for rock samples is quite varied, depending on the geological environment, the type of sample being collected and the geologist involved. The type of information collected includes the sample identifier, location, type and geological and geophysical attributes. A system needs to be sufficiently flexible to cater for the entry of all of these fields.

Field Mapping
This is an area where little advancement has been made over the original methods used. The old techniques of plotting field observations on aerial photographs or maps is still widely used, but satellite TM imagery has been found useful for delineation of some surface features. There are some projects under development that may further automate some of these time consuming tasks.

Geophysics
A wide range of both ground based and airborne geophysical data types are collected . These include magnetics, radiometrics, spectrometer, gravity and electromagnetic data.

Topographical Data

Topographical information is an important basic ingredient in the building of a digital map. Methods other than manual digitising of the data needs to be investigated to improve the efficiency and accuracy of the data capture.

Data Capture Techniques

Data Loggers

The definition of a Data Logger is becoming a little blurred with the continuing reductions in the physical size of PC's. The HP and Sharp "palm top" computers are an example of this trend with their full MS-Dos functionality and keyboard in a size less that half of the normal "notebook" computer. There is a wide range of data loggers available on the market, with similar features, the Husky being the most rugged and widely used unit. It was tested in British army conditions, is water resistant and reasonably shock proof.

Bar Code Entry

Most data loggers have a full QWERTY type keyboard, with the option of a bar-code readers for entry. The use of Bar codes is very effective in areas where the geology or expected data values are very well defined. This allows "pigeon holing" of recordings into the categories provided by the bar codes. If there is a wide variability of the expected data values, or a lot of keyboard entry is required, then bar codes may not be the answer.

Touch Screen

A different approach has been adopted on the Epson data loggers. These units have a touch sensitive matrix of cells on the screen which is programmable and can display all the options available to a operator when running a program. The user presses the touch sensitive section on the screen displaying the required option, which will bring the next set of options to the screen. Whilst this logger was very easy to use, it was constrained by lack of programmable memory, making it difficult to write an application of any size. They also lacked the durability of the Husky's, but are worth consideration.

G PS

The development in GPS technology is moving at a fast pace with the launching of new satellites annually. Unfortunately, the US Government has decided for security reasons, that "selective availability" will be switched on for the foreseeable future. This in effect degrades the quality of the satellite signal to all non-military GPS receivers, resulting in locational accuracy's of generally 50-80 metres. To be guaranteed of better accuracy's, the "differential GPS" approach needs to be applied. This technique uses one GPS receiver stationed on a known location, with a second "roving" receiver being used to locate the points of interest. Both receivers need to view the same satellites, with the accuracy's decreasing with the GPS receiver separation. This technique effectively eliminates the signal degradation, as the GPS on the known point can measure the location variations. These variations can then be applied to the "roving" GPS to get 2-5 metre accuracy's. Most GPS systems now provide output in AMG coordinates, with some providing the ability to add limited attributes to the recorded location.

Integrated GPS / Data Recorder.

There are several new systems available in the US which integrate the GPS into a PC to create a data logger with locational measuring capabilities. These systems are fully MS-Dos compatible, and will be a very useful tool in the future.

Voice Activated Recorders

Systems have been developed which combine locational mechanisms with voice data recording, allowing almost hands free operation of the recorder which is important in many situations. A system for recording road defects was developed in Western Australia, which was linked to the vehicle odometer. It recorded a simple voice command from the user describing the road defect - "pothole" - together with the "chainage" measured from the odometer. These two values were later down loaded into a larger system in the head office which automatically plotted the location of the road defect and its type. This approach has possible applications in recording drill hole logs and sample descriptions.

Geophysical Data Recorders

Ground

The on-board storage capabilities of the memory magnetometers have been widely used for a number of years, but they still require manual entry of locations. The other ground geophysical recorders generally do not have any data storage capabilities, and have to be used in conjunction with a data logger.

Airborne

Developments are progressing in the area of real-time data capture using helicopter-borne magnetometers in conjunction with a GPS and data recorder. This system records the magnetic response as the helicopter flies along a traverse line at the same time plotting the magnetic profiles on a data logger screen in the helicopter using GPS readings. This clearly highlights any magnetic anomaly, allowing immediate follow-up with detailed traverses over that area. This would result in significant improvements in helicopter utilisation, which at a rate of $5-600 / hour represents a considerable cost saving. In addition, the ability to immediately follow up anomalies is important, because current procedures would require a re-visiting of the site which may not be justified.

Conclusions

There are many opportunities for improvement in the data capture techniques currently employed, the major constraints being the technological tools available. There are considerable cost and operational benefits to be gained from improvements in this area, but the danger of "information overload" is a very real possibility.

DIGITAL MAP COMPILATION

Christopher Carter
Cartographic Services Unit
THE BUREAU Of MINERAL RESOURCES, ENERGY & GEOPHYSICS
CANBERRA.

Introduction

The 1990 North Queensland field season saw the trialling of what we believe to be a first in Australia, and that is the compilation of a geological map digitally, from field interpreted air photo overlays.

It would be correct to say, that prior to trialling this method of data collection, opinion was divided and debate was heated as to the viability or otherwise of this system. In fact, it was quite clear, that the majority of field experienced geological cartographers were committed to the view that this system (in theory) would not work in the general sense, that is, to automate the compilation of geological maps per say.

In qualifying this view, they suggested the proposed system may work in areas presenting well defined control or monument points, but will fail in areas where control points were few or poorly defined.

. there is of course a predominance of this terrain in Australia.

Preparation

When first looking at the procedures (used by most geological surveys throughout Australia) for compiling geological maps, our first impression was, that software packages currently available, would provide all the tools needed to perform this function. Thus the following equipment was resourced and assembled:

- 286 sx personal computer:
  - RAM extended to 2 megabytes.
- 14 inch colour VGA screen.
- Software:
  - Bentley Systems Microstation PC, Cartomap and Cartofit.
- A3 Summagraphics digitising tablet.
- Hewlett Packard 7475A A3 Pen plotter
- Power conditioner

Evaluation

Prior to the field trip, the system was bench tested.
A test area was selected and Cartomap was employed to create a grid and graticule .

Whilst the graticule was generated within the package, the grid needed to be manually constructed through Microstation.
This was an untidy process and so we chose to generate the grid and graticule in the Intergraph environment(World Mapping System) and then downloaded the result to the PC.

. at this point Cartomap was abandoned as a tool in our system.

Using Microstation, the topographic base of the test area was manually digitised.

. in the ideal circumstance the topographic base would be obtained in digital form from AUSLIG.

Air photos of the test area, with the geological overlays were sorted and monument points were selected on the photos, coincident with clearly defined and respective features on the topographic base.

The photo and overlay was then placed on the digitising tablet and the geological overlay was digitised into the design plane.

Cartofit was then run, in an attempt to fit the geological detail file to the topographic base file.The result was unsatisfactory, principally because, the affine transformation imposed upon the geological data had dispersed the distortion equally around the monument points. This resulted in an unacceptable miss register of detail.
To rectify these resultant errors proved to be labour intensive and inefficient and so Cartofit was therefore abandoned.

Failure and Re-think

An entire re-think was required, as the system we had devised had failed.

To find the answer, we went back and analysed the manual processes ie:

1. Assemble the photos and geological overlays.

2. Scale the photos and calculate a mean scale for the coverage.

3. Produce a photo scale topographic base on a transparent medium.

4. Individually orientate photos to respective detail on the photoscale topographic base.

5. Pencil plot boundaries and other detail.

Compilation revisited

The answer to our problem was sitting there in front of us, for we simply needed to duplicate the manual process,incorporating appropriate modifications to suit the new technology employed ie:

1. A photo scale base was produced on a stable base and oriented on the digitising tablet, to it's relative position within the Microstation topographic design file.

2.The geological overlays, including readily identifyable topographic monuments, are positioned and registered to their corresponding topographic features on the photo scale base.

3.The geological detail was then digitised into the design plane photo by photo using established BMR specifications.

4.Overlay joins are effected interactively.

5.Check plots are provided to the scientist for editing.

A Step in Faith

This system was taken into North Queensland and proved to be an great success, not withstanding the significant limitations experienced from the slowness of the 286 PC,and the A3 formatof both th digitiser and pen plotter.

Irrespective of these limitations, the experimental map The Coen Inlier was produced.
This, with the considerable support provided by the North Queensland field geologists, provided the impetus to get our cartographer and the system into the field again in 1991.

The Acid Test

This time with a 386 PC laptop and a A1 digitiser the 1991 field season proved to everybody that the efficiency benefits attributed to our process of digital map compilation were nothing short of outstanding. This time the only niggly inconvenience, was the limitation of the A3 pen plotter.

The Pudding Is Proven

Two months after their return from the field the North Queensland Project team, at a semi formal product release function, were released the completed compilations of six, 1:100 000 Regolith Geology Maps, two,1:100 000 and eight 1:50 000 Basement Geology maps and sundry other Geochemical and Magnetic map products.

A salable product range available this soon after a field season was unheard of in BMR until now and thus has established our digital compilation system as an integral and necessary part of our map production routine.

And Now

Should groups or individuals wish to obtain more information on our system, or wish to challenge any of our claims, please do not hesitate to call us on (06)245 1252, or drop us a FAX(06)247 2728. We are alway pleased to assist.

PRODUCT RELEASE 20 DECEMBER 1991

BASEMENT GEOLOGY

1:100 000 Ebagoola.
1:100 000 Kalkah.

and

1:50 000 coverage of both above.

REGOLITH

1:100 000 Strathburn
1:100 000 Ebagoola
1:100 000 Marina Plains
1:100 000 Princess Charlotte Bay
1:100 000 Strathmay
1:100 000 Kalkah

1990 HARDWARE INVENTORY

- 286 PC Laptop
- 14" Colour VGA Screen
- Summagraphics A3 digitising tablet
- Hewlett Packard A3 Pen Plotter
- Bently Systems Microstation
- Power conditioner.

GRAVITY DATA - DIGITAL GEOPHYSICAL SPATIAL DATA

Michael P Morse

Geophysical Mapping Section
Bureau of Mineral Resources, Geology and Geophysics
Canberra

Abstract

The aim of the gravity method is to generate displays of the gravity field reduced to a common datum so that lateral crustal density inhomogeneities are identified, and then to interpret the information to formulate geological models. Gravity data are also used for geiod studies, GPS navigation and surveying, and in fundamental physics. Gravity data acquisition involves the measurement of position, elevation and relative gravity. In the field, the BMR Gravimetry Group logs most of these data in digital format, an exception being positions which are digitised off maps when GPS data are not available. The data are then processed and point located gravity data are stored in the Australian National Gravity Database. This database is not relational. However, it is in the process of being migrated to a RDBMS. The data at this stage are suitable for inclusion into GIS packages as point attribute data. The use of gravity data in a GIS environment is limited by their spatial density. The power of gravity and magnetic data lies mainly in their ability to provide quantitative information on bodies beneath the surface. The integration of this information into a GIS to give a 3-d view needs further development.

Keywords Gravity/Australian National Gravity Database/GIS

Introduction

The attraction exerted by the earth on a mass above its surface causes the mass to accelerate towards the earth's centre of gravity. The resulting acceleration is denoted by *g*, which is approximately equal to 9.8 ms^{-2}. Small variations in *g* result from inhomogeneities in the earth's internal density distribution. It is the aim of the geophysicist to measure, isolate and interpret these variations in *g*. (Garland 1979) Gravity measurements can be made using two different classes of instrument:

i. absolute gravity meters which determine the earth's true gravitational attraction at a point.
ii. relative gravity meters which measure the difference in the earth's gravitational attraction from one point to another.

Relative gravity meters are used for geophysical mapping as they portable and efficient. These meters are very precise instruments capable of measuring the difference in *g* between two points with a precision of one part in 10^8. The observed gravity data need to be corrected so that the resulting gravity anomalies depend on variations in the density of earth materials in the vicinity of the measurement point, and not on other phenomena such as latitude and elevation which cause changes in gravity with position. The corrections are quite complicated.

Relative gravity surveys have been carried out in Australia since the early 1940's. The Australian National Gravity Database currently holds approximately 600 000 gravity station records and continental and marine data are being added continually.

Data Acquisition

The accuracy of a gravity value determined at a point is largely dependent on the accuracy with which the elevation of the point can measured and to a lesser extent the position. A 1 m error in the elevation of a station causes a 3 μms^{-2} error in the gravity value. This is a significant error as the anomalies caused by many orebodies are of the order of 10 μms^{-2} or less. Elevations derived from 3rd order levelling are usually accurate to 0.1 m. Thus, much effort is put into obtaining quality elevation data using GPS and precision electronic barometers. The data are logged in the field into hand held ruggedized PC's, the base barometer is also logged by a PC. The data are currently copied to disk and sent to Canberra for reduction.

Data Reduction

To reduce the gravity meter readings to gravity anomalies which can be related to density, many corrections are applied. The major corrections are earth tides, meter drift, normal gravity (latitude correction), free-air correction which accounts for changes in station elevation, Bouguer correction which accounts for the attraction of material between the station and the datum plane ignored by the free-air correction, and terrain effects. (Telford *et.al.* 1976) When all the corrections have been applied to the observed gravity reading we obtain the value of *Bouguer gravity* at a station.

Australian National Gravity Database (ANGD)

Table 1 gives the primary and derived attributes for a gravity station in the ANGD. Only the primary attributes are stored in the database. The database in its current form is implemented under custom FORTRAN software on the DG MV20000, this system is being moved to a UNIX environment. The database will be transformed into a relational model implemented in ORACLE when the new BMR database server is operational. A view of the database including the derived parameters (Bouguer and free-air anomalies) would be accessible by the corporate GIS. Table 2 gives a breakdown of the database on a stations per 1:1 000 000 sheet area basis for the Australian continental sheets.

Table 1. Attributes for a gravity station record.

Field	*Name*	*Format*	*Units*	*Comments*
Primary Attributes				
1	Station Number	A10		Unique identifier
2	Latitude	F10.5	degrees	Position
3	Longitude	F10.5	degrees	Position
4	Meter Height	F9.2	metres	Australian Height Datum
5	Observed Gravity	F10.2	μms^{-2}	Reduced gravity value
6	Ground Height	F9.2	metres	Australian Height Datum
7	Terrain Correction	F8.2	μms^{-2}	Due to effect of local terrain
8	Informal Name	A30		BM's, Control Stations, etc
9	Datum	A10		Contains units, format, scale, datum, security, reliability
10	Date of last update	A10		Last modification of gravity station.
Derived Attributes				
11	Free Air Anomaly	F8.2	μms^{-2}	Correction to mean sea level
12	Bouguer Anomaly	F8.2	μms^{-2}	Correction to mean sea level assuming a density for the material between the station and MSL.

Gravity Data in a GIS

The incorporation of gravity data into GIS packages is a simple process. However, the effective use of gravity data in a GIS is not as simple.

Potential field data in a GIS differs from other data types such as radiometrics in that the value at given point may be important to studies of fundamental physics, but for most geological applications is of little use to the interpreter. It is the field patterns when contoured or gridded which provide important 2-d information for interpretation. The gravity value of an individual station is of little use to a geologist.

Apart from the use of the located gravity data in a GIS, an effective way to integrate gravity data into a GIS environment is to use raster images and contours of the Bouguer gravity anomalies. This would only be necessary where gridding and contouring facilities are not present in, or linked to the GIS. The full value of gravity data is obtained by the use of three dimensional modelling parameters derived from the gravity data. The incorporation of this data into a GIS is a problem to be solved.

The simplest use of gravity data in a GIS is possibly the transfer of the Bouguer anomaly contour map from paper to the screen. Limits on the usefulness of this procedure are imposed by the spatial density of the gravity stations.

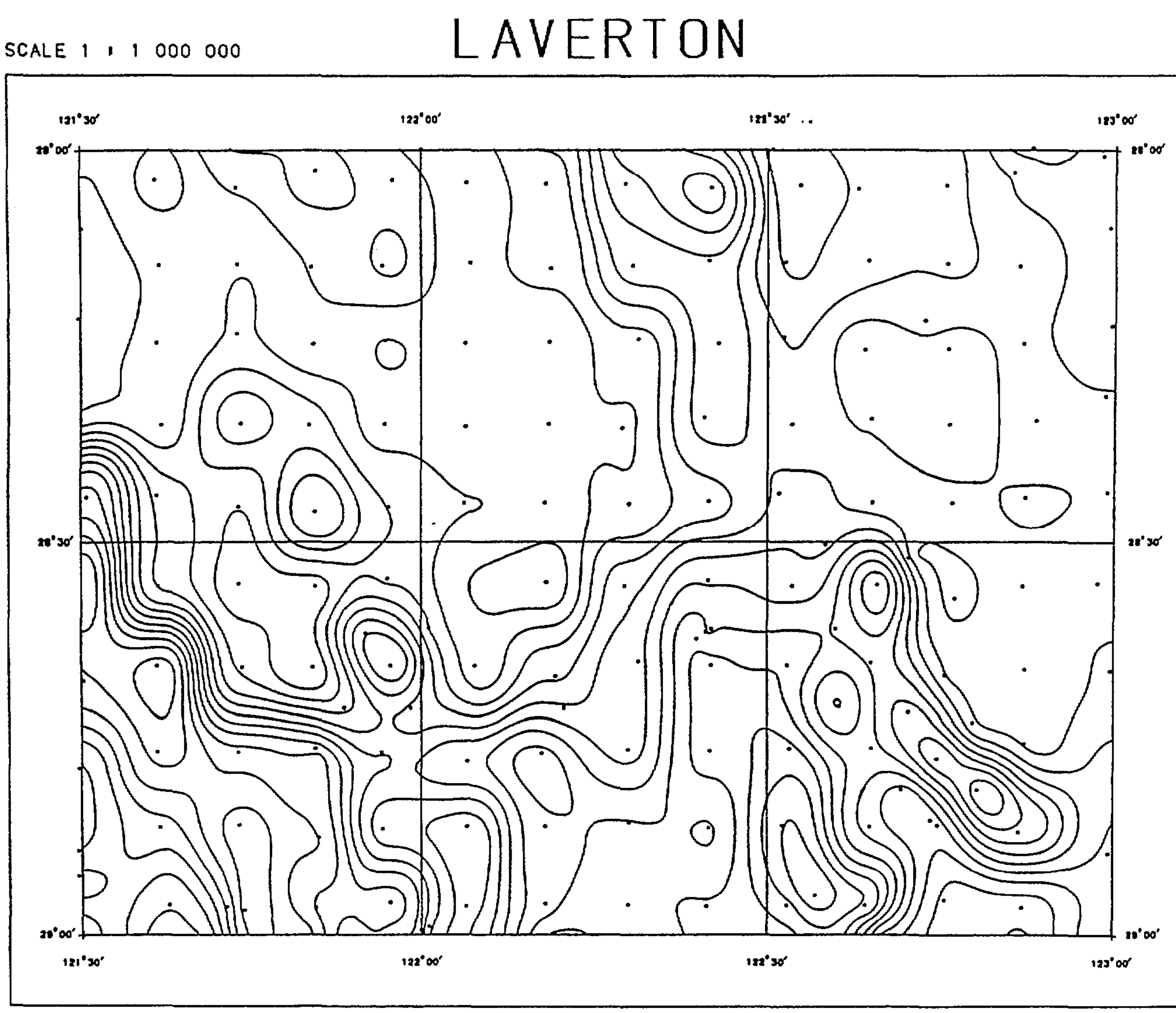
Table 2. Australian continental 1:1 000 000 map sheets with approximate station numbers. From 1991 ANGD data release.

Map Sheet	Sheet Numbers	Stations
Torres Strait	C54	2900
Brunswick Bay	D51	6000
Darwin	D52, part of C52	10000
Roper River	D53, part of C53	3000
Mitchell River	D54	2000
Cooktown	D55	3300
Broome	E51	37000
Halls Creek	E52	4200
Newcastle Waters	E53	6800
Normanton	E54	5500
Townsville	E55	3100
Hamersly Range	F50, part of F49	12000
Oakover River	F51	8300
Lake Mackay	F52	15000
Alice Springs	F53	12700
Cloncurry	F54	10600
Clermont	F55	13100
Rockhampton	F56	2500
Meekatharra	G50, part of G49	12000
Wiluna	G51	20700
Petermann Ranges	G52	14800
Oodnadatta	G53	19500
Copper Creek	G54	24400
Charleville	G55	25700
Brisbane	G56	7200
Perth	H50	9700
Kalgoorlie	H51	8500
Nullabor Plain	H52	6500
Tarcoola	H53	16000
Broken Hill	H54	38400
Bourke	H55	13700
Armidale	H56	6300
Albany	I50	10300
Esperance	I51	3000
Port Augusta	I53	13000
Adelaide	I54	25000
Canberra	I55	15000
Sydney	I56	3600
Hamilton	J54	22500
Melbourne	J55, part of J56	14000
Tasmania	K55, part of J54, J55, K54	33000

Figure 1 shows a 1:250 000 sheet with the standard 11 km station spacing with approximately 150 stations per sheet or 25 stations per 1:100 000 sheet. This figure shows the level of detail that can be expected from this density of data. The shortest wavelengths that can be identified are 22 km - i.e. resolution is low and only gross geological features can be identified.

Figure 2 shows a 1:250 000 sheet which has been infilled to approximately 4 km station spacing, the total number of stations being approximately 1000. This maps shows improved detail, nearly three times the resolution of the existing 11 km data, and gives more confidence in the contours. This level of station density allows for good interpretation at 1:250 000 scale. Figure 3, a 1:100 000 scale map from Figure 2 shows that this station density is minimal for effective interpretation at this scale. As a comparison, the National Geoscience Mapping Accord (NGMA) airborne magnetic data can resolve wavelengths of 12 m.

Figure 1. Laverton 1:250 000 sheet from the Eastern Gold Fields of Western Australia. Grid cell size 1.0 x 1.0 minutes. Contour interval 40 μms^{-2}.

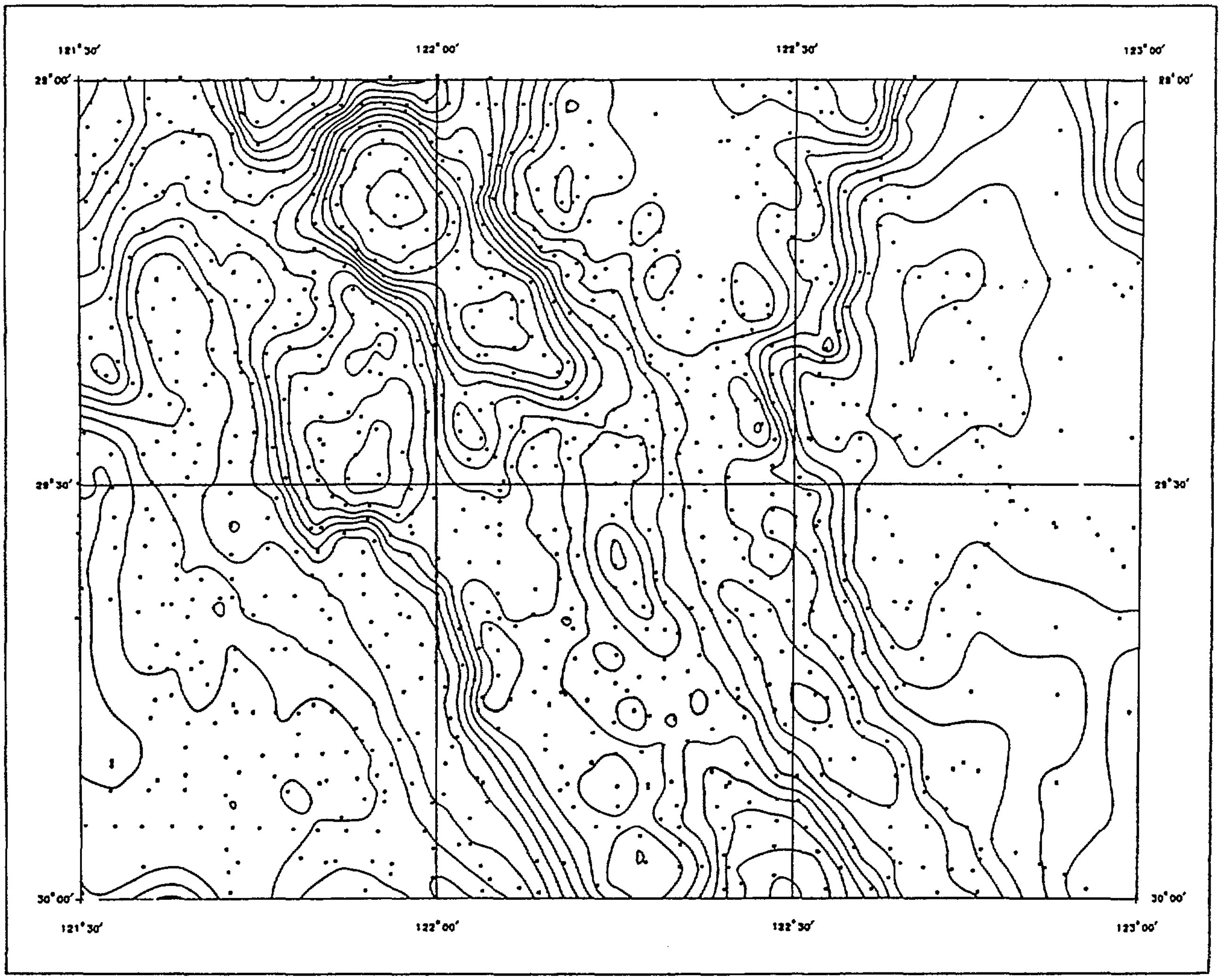


The use of grids as raster overlays and GIS database attributes is also useful. Their use again is limited by the spatial density of the gravity data. The standard grid of Australia has a grid spacing of 3 minutes (approx. 5 km), the data displayed in figure 2 can be gridded to 1 km. These grids have large cell sizes compared with aeromagnetic and radiometric grids and this can cause difficulties in comparing gravity with magnetic or radiometric data (also geological data). Also it is necessary to use an enhanced version of the Australian bouguer/free air grid, to obtain the full information potential from this data presentation.

Figure 2. Edjudina 1:250 000 sheet from the Eastern Gold Fields of Western Australia. Grid cell size 1.0 x 1.0 minutes. Contour interval 40 μms-2.

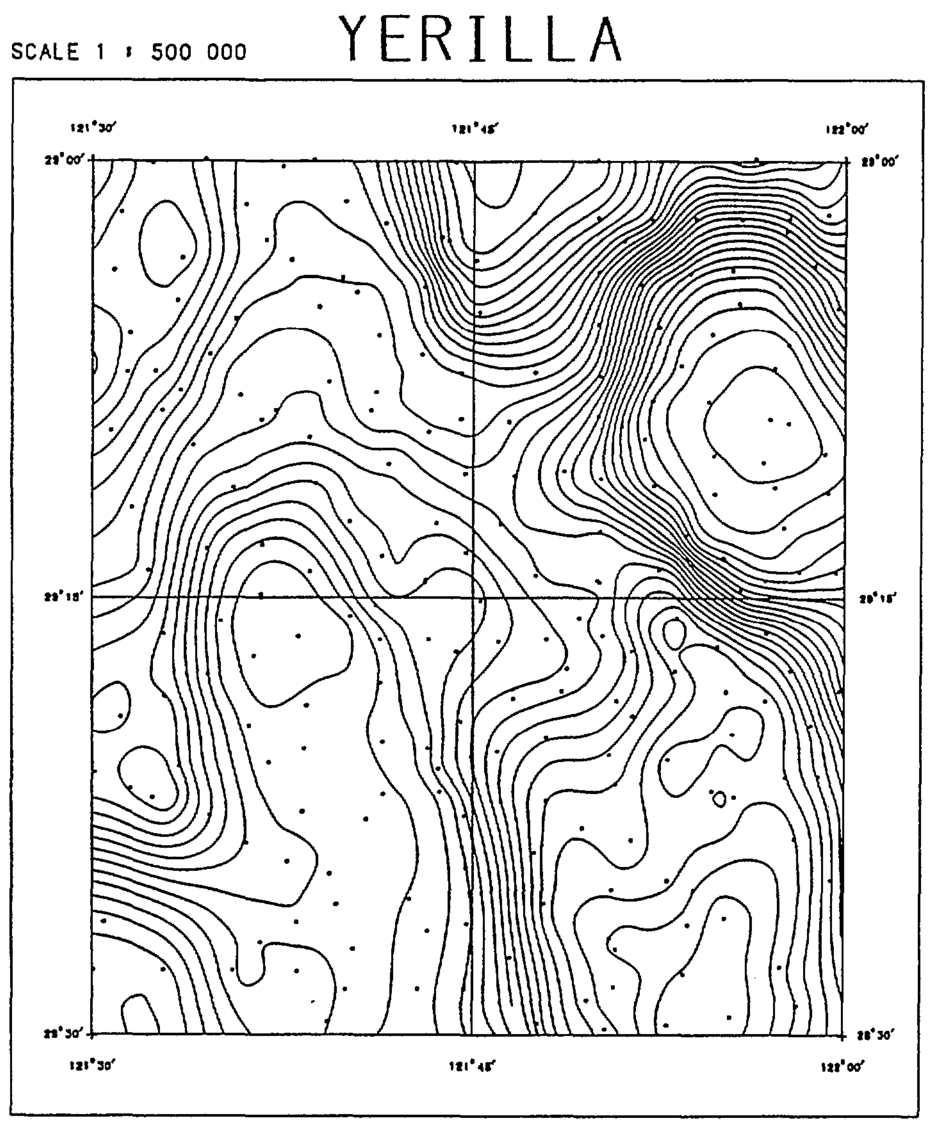
SCALE 1 : 1 000 000

EDJUDINA



In summary the comparison of gravity values in a GIS with other attributes (eg. geological, geochemical, geophysical) must take account of the possible aliasing (ie. undersampling) of the gravity data, and the difference in sampling density between the gravity data and that of the other attributes.

Figure 3. Yerilla 1:100 000 sheet from Edjudina. Grid cell size 1.0 x 1.0 minutes. Contour interval 20 μms^{-2}.



Derived Parameters

Limiting the use of gravity data in the GIS to the approaches given above limits the information available to the interpreter. Indeed potential field data is collected to not only give the field pattern but to determine quantitative models for subsurface structures. The ability to include these derived parameters into the GIS would greatly enhance the usefulness of potential field data to the geologist. If we take the example of gravity traverses, of which there are many in the ANGD, these data add little to the overall 2-d field patterns in contours or grids. Most of the close spaced data in the traverse can be removed without causing any degradation in the contours or grid. How do we effectively use this detailed gravity data in a GIS? These data are important in that they can be interpreted by forward modelling or inversion to give a picture of density inhomogeneities within the earth's crust. The use of the model parameters derived from these data provide an effective vehicle to transfer the information content of this data to the GIS. The modelling of traverse data gives another 2-d data set orthogonal to the standard 2-d plane used in a GIS. The extension of this concept to three dimensions is more difficult, it involves the GIS being able to manipulate and display three dimensional data not just any two dimensions at one time.

Conclusion

Gravity data can be easily incorporated into GIS packages as attribute data, contour maps and raster overlays. However, the ability to include and/or derive model parameters in a GIS will allow more information to be obtained from gravity data.

References

Garland G.D., 1979--Introduction to Geophysics second edition. *W.B. Saunders Company*

Telford W.M., Geldart L.P., Sheriff R.E., and Keys D.A., 1976--Applied Geophysics. *Cambridge University Press.*

AIRBORNE GEOPHYSICAL MAPPING - A DIGITAL DISCIPLINE LOOKING FOR AN EFFECTIVE DIGITAL USER-INTERFACE

Colin V. Reeves

Geophysical Mapping Section,
Bureau of Mineral Resources, Geology and Geophysics,
Canberra.

Abstract - Digital data capture was introduced into airborne geophysical mapping almost 30 years ago. Present day instrumentation and navigation/position fixing techniques generate large quantities of digital data which, while not matching the quantities generated by seismic surveying, present real problems with data volumes. When 'reduced' to rasters, the geophysical data are much in demand as basic data-planes for GIS users but the resources of hardware, software and humanware needed to handle the geophysical information within a GIS environment without seriously compromising access to the wealth of original profile data tend to make current GIS ambitions look rather modest. Meanwhile geophysical mappers themselves have need of better access to other GIS layers such as planimetry, digital geology and mineral occurrences.

Keywords - Aeromagnetics / Gamma-ray spectrometry / databases / GIS.

Within the earth sciences, geophysics has often set the pace in the application of computers. Our brothers in seismic exploration for oil and gas have claimed to be second only to the world's tax departments in terms of volume of computer usage. Within BMR, and within airborne geophysical data acquisition, records show that we already had what was probably the world's first airborne paper-tape punch in regular action in the early sixties (Tucker et al 1988).

Digital data acquisition became fairly general among the world's geophysical survey aircraft in the seventies but it is only in the relatively recent past that this process has reached its logical conclusion and *all* data associated with airborne surveys have been acquired digitally.

The last - and arguably the most important - phase of 'going digital' was in navigation and flight-path recovery. The traditional and laborious process of recording the flight-path using a downward-looking 35 mm frame camera and using this to recover the aircraft's track against topographic maps, conventional aerial photography or satellite imagery has only recently been fully superseded by electronic navigation systems such as Syledis and, more recently, by GPS. Now positions sufficiently accurate for almost all airborne survey criteria can be captured directly in *x,y* coordinates; survey results can be plotted directly onto a sheet of paper without ever referring to a map or photograph. More importantly, digital navigation systems allow direct feedback to the pilot so that survey lines can be flown closer to their desired position, with the result that survey flight patterns become more regular - a bonus in data quality when uniform coverage is a desired objective. In older surveys, even if flight-path *recovery* is accurate, the flight path itself may deviate widely from that desired.

Concurrent with the general advance of digital acquisition has been an increase in the *volume* of data to be captured. What started with the capture of a five-figure number from a proton magnetometer once every second has evolved into a significant exercise of digital data acquisition. Airborne magnetometers can now measure to an accuracy of seven significant figures and sample every tenth of a second. Recording these results, plus ancillary numbers from the navigation system(s), altimeter(s), gamma-ray spectrometer etc leads to an incoming data-flow of about 3 kilobytes per kilometre flown while the aircraft is on line. This can still be stored on a small number of diskettes for the average survey flight of 1000 to 1500 line-km (Figure 1).

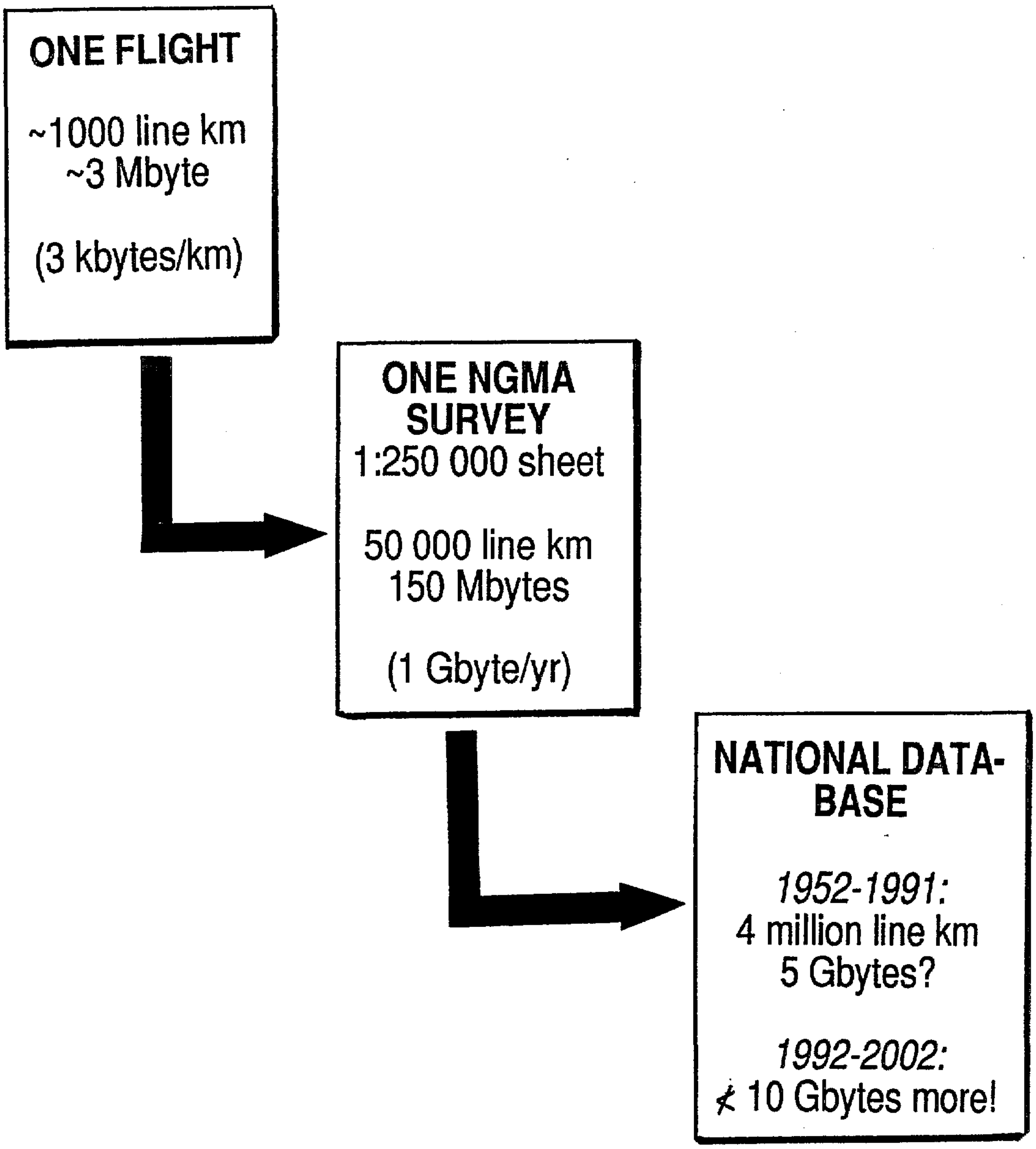


Figure 1: Estimated accumulation rates for airborne survey data.

While we are still working on survey coverage of the remaining few per cent of onshore Australia within the framework of the reconnaissance aeromagnetic survey programme, the thrust has now changed towards the National Geoscience Mapping Accord and its requirement for closer line spacing (400 m usually, 200 m where support from funding outside BMR can be obtained). A consequence is a high density of data in selected 1:250 000 sheet areas.

One 1:250 000 sheet at 400 m spacing represents about 50 000 km of flying (in round figures) and consequently about 150 megabytes of profile data which - after they have been reduced and levelled - is the quantity of data bought by our customers when they acquire digital located profile data from the BMR for one NGMA map sheet. This is certainly more than we can fit on one magnetic tape which is still the accepted medium for despatch of digital data to the end user.

Each survey in turn represents a unit contribution to the national airborne survey database (Figure 1). In 1991 we passed the milestone of 4 million line-kilometres of survey flown. Just how big this data-base is is somewhat uncertain since it spans almost 40 years of improving acquisition techniques and quite a lot of retrospective digitization, but much of it will be far less data-intensive than present acquisition. My guess for the total is 5 gigabytes, ± 3 gigabytes. But note that now, with the acquisition of about five new mapsheets (or their equivalent) per year, the national data-base of profile data is expanding by about 1 gigabyte per year.

While this may now seem to be a lot of data, what happens if we start to record all 256 channels of the gamma-ray spectrometer? Or if we add other geophysical sensors? Or if airborne gravity suddenly becomes feasible in a small aircraft? Or if we get funding to fly the full Mapping Accord target of ten sheets a year for the next twenty years? We could face a data deluge in an area where we already suffer hardware constraints.

My pre-occupation with located profile data may need a little explanation since many people will consider a grid to be the ultimate product of a geophysical data-set.

One reason for this is that, from the moment the diskette comes out of the aircraft, we have a geo-referenced digital data set. Many, if not all, the values recorded will need some adjustment during the routine sequence of data-reduction. User-friendly and systematic access to these data during the data reduction process could well be seen as a GIS application which, if it could be fulfilled suitably, might be an improvement on present data reduction systems.

A second reason is that the serious interpreter of geophysical data cannot function satisfactorily without access to the original profile data. While a gridded data set with a suitable enhancement may provide an excellent representation of much of the geo-information recorded, the interpreter will often need to look more closely at a magnetic profile, or wonder if the ground clearance of the aircraft was in specification, and so on. He is handicapped quite severely if he cannot have access to these data conveniently. This, again, is an opportunity for GIS application. In a presentation to the BMR a few months ago, Nick Sheard of Mt Isa Mines expressed concern that no satisfactory method presently exists to 'browse through' what is arguably the largest and most intensive airborne data set in corporate hands, namely the 600 000 plus line-km of data acquired for MIM over the Mt Isa Block in the last few years. Accessing BMR's national airborne survey database of Australia may be seen as a similar existing problem for which GIS application could offer a solution.

A third reason is that access to profile data within the context of the national data base means that data from surveys flown at different times can be brought together to provide coverage of any area covered by more than one survey. To adjust surveys for secular variation etc by making adjustments at the intersections of lines flown in different surveys is a solid approach to linking together old surveys and should be possible within a GIS framework on the national data-base if it can be implemented at the level of a single survey.

Perhaps the main point I want to make in this brief presentation is that these volumes of real data appear large in the present day vocabulary of GIS. I hope I can be proved wrong in the near future because I am encouraged by what I have seen GIS do in other areas of airborne surveys. The simple indexing of airborne surveys and their outlines on a PC-based GIS was a tremendous asset in the compilation of all available airborne magnetic surveys for Africa which is presently reaching its culmination with my colleagues in Europe, North America and Africa itself.

Even if we are sufficiently modest to content ourselves with grids of airborne survey data, the problem of data volumes within GIS applications does not go away. We are presently well advanced in compiling a digital magnetic anomaly grid for all onshore Australia on a 15-second (about 400 m) grid. This, I am sure, is a fundamental geo-data set for Australia that is long overdue and will be required by many users for a multiplicity of applications. To store this grid as real numbers - i.e. without compromising the dynamic range of the data by reducing them to 256 or even 1024 grey levels - requires some 360 megabytes of storage space.

By the same token, a single 1:250 000 scale sheet of Mapping Accord data can be reduced to six grids (one magnetic anomaly, four spectrometer channels and a digital elevation model) of grid dimension about 90 metres, each of which requires 4 megabytes of storage space.

As a geophysical mapper, it is exciting to experience the demand for these grids as basic data layers from many users; it is clearly our mandate to make these data as easily available as possible to users within BMR, within state geological surveys and the outside community at large. However good we can make the grids, I still worry when access to the original profile data is denied because, as an interpreter, I know the value of that profile data at many stages in the interpretation process.

I also worry that, once imported as grids, geophysical data are often immediately subjected to enhancement processes such as contrast stretching which are inherent - or endemic! - to image processing systems. Relatively little emphasis seems to be placed in the broader user community on data processes that can be applied to Laplacian (potential) fields to enhance aspects of the data while retaining physical meaningfulness. Amongst these are the calculation of vertical derivatives, regional-residual separations, migration to the poles or equator and the derivation of pseudo-rock property maps. These processes have been the stock-in-trade of geophysical interpreters for many years and do not in any way preclude the application of more general image-processing once these processes have been carried out. I would be happy to see their inclusion in the toolbox of any GIS boasting image processing capacity likely to be used for examining geophysical data.

Finally, while I applaud the growing recognition of geophysical data as essential data-planes for GIS users and am happy to be part of the process of provision of these data, the exchange of basic data must ultimately be a two-way traffic within a corporate structure. My hope is that the near future will bring the availability to workers in geophysical mapping of data planes from other discipline areas. Some of these, such as planimetric data, are basic to planning, execution, compilation and presentation of airborne surveys, while others such as digital geological maps and mineral occurrence maps are important elements in the thorough interpretation of geophysical data sets for exploration purposes.

Reference:

Tucker, D.H., Hone, I.G., Downie, D., Luyendijk, A., Horsfall, K., and Anfiloff, V., 1988 -- 'Aeromagnetic regional survey of onshore Australia', Geophysics, vol 53, pp 254-265.

MARINE GEOSCIENCE RESEARCH - ONE MARINER'S PERSPECTIVE OF THE APPLICABILITY OF GIS

Howard M.J. Stagg

Marine Geoscience & Petroleum Geology Program, BMR

Abstract

Petroleum-oriented marine geoscience research has traditionally expended intensive effort and expense in the acquisition, processing, and interpretation of profile-based seismic data, while the common end product, maps, have generally been interpreted in a comparatively simplistic fashion, in part due to the scarcity of suitable tools. Geographic information systems (GIS) have the potential to raise the quality of map interpretation to a level similar to that which is expected in the interpretation of seismic profile data. While the marine research community has lagged the onshore (particularly hard rock) geoscience community in adopting GIS as an interpretive tool, the fact that the majority of marine data (including maps) is available in a variety of standard digital formats should facilitate the importing of such data to a GIS in the future.

Background

While the use of geographic information systems as an aid to oil exploration appears to be on the increase in mature exploration environments (for example, the USA and the North Sea), within Australia the level of useage is still low, both in government and in industry.

Typically, acquisition costs for the most fundamental petroleum exploration and marine research tool, multichannel reflection seismic (MCS) data, run at about \$500-600 /km, while processing costs currently average \$200 /km, though they have been higher in the immediate past. The costs of interpretation vary considerably with the organisation and the level of its available technology, but are probably in the range of \$10-100 /km. The outcome of this expense, up to \$1000 /km is the drilling of an exploration well which, will probably cost from \$10-25 million in the offshore environment.

It is evident from these 'ball-park' figures that the interpretation phase is the poor relation in the exploration family, despite the fact that it is the most critical stage in discovering petroleum. Given that seismic and drilling costs are unlikely to decrease in real terms and that these costs cannot be avoided, then it is evident that efficiency and quality of the interpretation phase has to be maximised.

Do geographic information systems provide an opportunity for improved standards and efficiencies in interpretation? Before attempting to answer this question, it is necessary to outline briefly the stages in the interpretation of an MCS data set, from the completion of data processing.

Marine Geoscience - Processing to Project Completion

In petroleum-oriented marine geoscience research, the sequence of steps from the end of data processing to the release of a final interpretive report is fairly clearly defined. For a 'typical' project these steps may encompass the following:

1) Preparation of base maps and seismic sections at the working scale.

2) Interpretation of the seismic, potential field, and geological data; this step usually requires several iterations and accounts for the majority of the intellectual input.

3) Input of the interpretation to a database. In industry surveys with closely spaced lines, this is frequently facilitated by the use of interpretation workstations in which the horizon data are automatically entered in the database as the interpreters make their horizon picks. Where this technology is not available, or where the processed digital seismic data are not readily accessible, the horizon digitising must be done via a digitising table.

4) Production of structure and isopach maps of the selected horizons and intervals. The data may be either machine- or hand-contoured, depending on the availability of software and the geological complexity.

5) Final reports are then based on interpretation of the structure and isopach maps and of the original seismic data.

Steps 1, 3, and 4 are normally undertaken with a commercial basin analysis software package. Such software typically includes facilities for digitising maps and seismic sections, time-to-depth conversion, map production, gridding and contouring, and ancillary displays such as 3-D fence and mesh diagrams. In most marine geoscience organisations, whether exploration or research, the basin analysis software is central to all office-based functions. In BMR Petroleum Group's case, basin analysis is tackled with the Australian PETROSEIS package.

Unfortunately, in many cases, the level of interpretation of the structure and isopach contour maps is not commensurate with the intellectual effort that went into interpretation of the original seismic reflection data. While commercial basin analysis software packages are very good at what they do, the level of map analysis that can be carried out is at a rather simplistic level, compared to what can be achieved with a full GIS. It is at this 'down-stream' end of the interpretation scheme that GIS could find a niche in petroleum exploration.

Potential Role for GIS

Before defining tasks that could be undertaken in petroleum exploration research with a GIS, it is necessary to list those themes or coverages that would be fairly readily available to a petroleum GIS, at least within BMR.

Basin Outlines: These data are available as part of the Petroleum Exploration and Development Titles map on Intergraph and could be transferred to GIS with attributes added after transfer.

Tectonic Elements: Such elements include faults (normal, reverse, strike-slip), depocentres, high trends, hinges. Much of these data are currently being digitised in PETROSEIS and could be rapidly made available to a GIS.

Petroleum Exploration Tenements: As with basin outlines, these data are available on Intergraph. Attributes, such as tenement holders, operators, and relinquishment dates would have to be added after transfer.

Petroleum Exploration Wells: All basic information relating to petroleum exploration, appraissal, and development wells, onshore and offshore, is directly available through the PEDIN Oracle database.

Oil & Gas Fields: At a national level, these data are available on Intergraph; however, at regional or detailed level they would probably require re-digitising from maps at a suitable scale.

Marine Parks: This coverage would need to be digitised or manually entered.

Seabed Boundaries: national and international; available through Petroleum Tenements and Law of the Sea maps.

Palaeogeography: Some palaeogeographic maps are available on Intergraph and could be fairly readily transferred. However, others would need to be digitised as needed.

Structure & Isopach Maps: While conventional structure and isopach contour maps are probably of limited value for importing to a GIS, derived maps may well be valuable. Such derived maps could include: isopachs of a source interval buried beneath a certain minimum thickness considered necessary to generate petroleum; structure of a potential reservoir horizon edited to show only structural highs. These maps could be fairly readily transferred from the existing basin analysis system.

It can be seen that much of the above data are already available in a digital form that could be imported to a GIS, albeit with some reformatting and with attributes still to be added. There are probably many ways in which the coverages above could be combined and analysed to yield new information; two that are immediately obvious are:

* To carry out straightforward combination of coverages - for example, to produce a map of a particular basin showing the location of mature source rocks that are outside marine parks and within recently-relinquished exploration acreage.

* To test a geological model - for example, suppose it is postulated that within a particular basin transfer faults reactivated in the Paleogene are a controlling influence on the presence of gas fields. This postulate could be tested by combining coverages of the hydrocarbon fields (selecting gas fields) with tectonic elements (selecting transfer faults reactivated in the Paleogene), and carrying out a proximity analysis (buffering). In the event of a successful correlation, the model could be applied in conjunction with other coverages (eg presence of mature source rock) to predict the most likely localities for further gas fields.

To some extent, the above tasks could be handled with the existing basin analysis software. However, the spatial analysis capabilities of such software tend to be 'add-ons' of somewhat limited scope, which means any reasonably sophisticated analysis is very cumbersome.

Pilot Studies

While the BMR Petroleum Group has lagged other areas of BMR in adopting GIS as a tool, three pilot studies are currently being investigated. Should these studies receive approval to proceed, then they should be underway during 1993.

The first of these pilots is at national scale and is designed to provide coverages of extent of sedimentary basins and of petroleum exploration tenements, both of which are currently available on Intergraph where they are used in the updating and production of the Petroleum Exploration and Development Titles map.

The remaining two pilots are at a regional level, one onshore and one offshore. The proposed areas are the Taroom-Mundubbera map sheet in Queensland, and either the Vulcan Graben or the northern Carnarvon Basin on the North West Shelf. Both regional study areas contain comprehensive digital data sets, and both the regional pilots would be capable of integration with the national pilot.

Conclusions

While petroleum geoscience researh in Australia has generally been tardy in adopting GIS as an interpretation tool, there are signs that this is changing, both in industry and in government. The fact that the acquisition and processing, and to a lesser extent the interpretation of seismic data has been largely digitally-based for many years should reduce the pain of adopting GIS technology since the data digitising and verification phase has to some extent already been carried out.

There is little doubt that GIS has a role to play in petroleum exploration and research, both onshore and offshore, as it should provide for more sophisticated interpretation and analysis of maps than is currently the case. However, the particular concern that we are facing in BMR Petroleum Group at the moment is what relative importance should be given to the adoption of GIS in relation to other work priorities.

SDTS - THE SPATIAL DATA TRANSFER STANDARD AND GEOSCIENCE DATA

G. Wood[1] and B.J. Foley[2]

[1]Bureau of Mineral Resources, Canberra
[2]Australasian Spatial Data Exchange Centre, Melbourne

Abstract

The need for standards for file interchange of geoscience data is becoming more pressing as the amount and complexity of digital data increases. The lack of adoption of standards causes inefficiencies and on-going problems as users revert to developing one-off conversion routines or use proprietary one-on-one translators.

Fourteen leading Australian organisations associated with the minerals industry have sponsored AMDEX, a project to define standards for file interchange and archiving for geological and mining data sets. The proposed new Australian Spatial Data Transfer Standard, SDTS, has been found to be suitable for all data sets so far examined. It is a clone of SDTS, developed by the U.S. Geological Survey.

This paper examines progress on the AMDEX project to date, highlights areas for further examination and raises a number of issues related to SDTS, including its impact on data quality, the need for data dictionaries, its role in archiving and its future in the geosciences.

In late 1991, the Australasian Spatial Data Exchange Centre (AUSDEC) was established to foster the development and use of SDTS throughout the spatial data user community. The role of AUSDEC in liaising with the USGS and promoting SDTS to data users, data providers and system vendors is detailed.

Keywords

Spatial data transfer, SDTS, AMDEX, transfer standards, AUSDEC

Introduction

The efficient use of geographic information system (GIS) technology to produce a new generation of geoscience datasets and exploration models under the National Geoscience Mapping Accord (NGMA) will require effective spatial data sharing among the accord contributors. That means moving greater volumes of data more rapidly.

Most practitioners in the geoscience computing field are familiar with the process of 'massaging' data to alter an incoming data set to suit the peculiarities of the software used by their own organisation. This is usually done by using an editor (assuming the incoming data is in ASCII) or by writing a utility (often, a one-off program).

The problems are endemic throughout the exploration and mining industry and waste time, money and man-power. In recognition of this, some software package vendors provide translators for certain data sets for certain other packages. However, given there are more than one hundred mining and exploration packages available in Australia, and over fifty common data sets, this is, at best, a piecemeal approach.

In early 1990, the Australian Minerals Industry Research Association (AMIRA) received a proposal from Mining Research Associates to develop a set of standards for interchange of mining (including exploration) data. The proposal was accepted and subsequently funded by fourteen leading organisations associated with the minerals industry. The project, known as AMDEX (Australian Mining Data Exchange), began in 1991.

Data interchange - the problems

The AMDEX prospectus detailed the following common problems:

- Organisations receiving data from different sources and having to handle each parcel separately.
- Sending data to another organisation.
- Joint ventures having different software but needing access to a common database.
- Different departments or sites needing to share data.
- Need to access old computerised data.
- Submitting (or receiving) exploration data to/from Government authorities.
- Archiving data for extended periods.

A poll of fifty users, consultants and vendors showed widespread problems and overwhelming support for the concept of a standard for file interchange.

Amdex - design criteria

The initial meeting of sponsors in April 1991 laid out the following guidelines for the development of an Australian mining data exchange format:

- The standard must be simple
- It must be unambiguous
- It must cover all the data types in common use in the minerals industry
- The standard should be sufficiently well developed to qualify as an Australian standard
- It must be independent of transfer media considerations
- All other relevant standards should be examined to determine their relationship to AMDEX.

This last qualification led to the connection with the spatial data transfer standards.

SDTS

Australia currently uses AS2482, a standard for interchange of feature-coded digital mapping data. However, for the spatial data user, as Clarke (1990) points out, AS2482 is now seen to have some serious conceptual shortcomings, including:

- Not supporting polygon, grid or raster data types
- Not supporting topologically structured data
- Minimal provision for data quality data
- Minimal provision for attribute data.

Standards Australia committee IT/4 (Geographical Information Systems) has accepted the recommendation of its sub-committee IT/4/2 (Geographic Data Exchange Formats) to 'clone' the U.S. Spatial Data Transfer Standard (SDTS), developed over the last ten years by the U.S. Geological Survey. SDTS handles spatial objects commonly used in mining and exploration data sets - points, strings, polygons, grids, triangulated surfaces etc. SDTS is currently undergoing certification by the U.S. National Institute of Science and Technology and, when completed, will become an FIPS Standard, making it mandatory for all data importing/exporting to and from U.S. Government agencies. It is proposed then to submit it for ANSI accreditation (although this is not certain as a competing standard, DIGEST, may go forward instead).

As part of its examination of relevant standards, the AMDEX project team concentrated its efforts on mapping known data sets to SDTS format to see if it was universal. So far, results are extremely encouraging.

SDTS vs. AMDEX design criteria

Does SDTS meet all the AMDEX design criteria ? No! It patently fails the first test i.e. must be simple. However, at sponsor's meetings in November, 1991, the AMDEX project team (Hume, 1991) successfully argued that:

i) Simplicity and functionality are mutually exclusive

ii) The apparent complexity of SDTS will not, in general, be a problem because the input/output routines will be embedded in commercial software packages and thus be transparent

iii) Initial unfamiliarity with SDTS can be overcome by education and training

iv) For software developers, a range of SDTS utilities can be made available e.g. editors, input/output routines

v) The benefits of building on ten years detailed work by the USGS easily outweigh the apparent complexity

vi) The use of SDTS provides a bridge between mining and exploration systems and other spatial data systems (e.g. Geographic Information Systems)

vii) Because SDTS will become an Australian Standard, the AMDEX project can now concentrate on developing geoscience implementations of the standard, rather than the standard itself (e.g. profiles, data dictionaries).

viii) AMDEX will have a much better acceptance by vendors and users because it is based on an international standard with a guaranteed long life.

ix) The body of knowledge built up in support of SDTS can be utilised by the Geoscience community.

With some reservations (related to complexity) the sponsors of AMDEX unanimously decided to back SDTS, and work is proceeding on this basis.

SDTS and ISO8211

SDTS defines the way spatial data is modelled and how it is structured into records within modules. The actual encoding into logical records is done via the international standard ISO8211 (AS3654 - 1989). ISO8211 serves as the basis for a number of GIS packages, and should be considered, in any event, as a basic data interchange format (rather than flat ASCII transfer files). Again, utility software is available to assist developers.

In general, ISO8211 describes self defining files. Many packages have invented their own version of these, but some order would result from a general commitment to ISO8211.

It is important to make the point that the proposed usage of both SDTS and ISO8211 is solely for external ASCII transfer files exported by a computer system. No assumptions are made on how each system handles or stores the data internally.

Data Quality

Clarke (1991) examined a number of implications for data users and data producers when SDTS is implemented. The standard encourages (requires) users to supply, with each data file, a data quality report with information on data lineage, positional accuracy, attribute accuracy, logical consistency and completeness. There are no guidelines on what should go in these reports. They are freeform text areas. Over the next few years, it is likely that users in many areas (including geoscience) will find benefit in prescribing data quality statements. This will be investigated as a possible extension of the AMDEX project.

Archiving Data

As noted above, one recurring difficulty with geoscience data is that of archiving and retrieval. Several of the AMDEX sponsors have joined the project solely in the hope that some solution will be offered to the problem of gathering costly data, storing it for say ten years and then, on retrieval finding it can't be read by existing software (without massaging) or, worse, it was stored with only local grid co-ordinates, and the tie in with the ground, or AMG, has been lost.

SDTS certainly offers sufficient facilities to meet the requirements of secure archiving. Work is continuing to explore the trade-offs between security and volume; SDTS sets of files are bulky, especially for high density raster data. Nevertheless, current indications are that it will be suitable.

Multi-level Conformance

The AMDEX sponsors gave their blessing to SDTS with the proviso that the issue of multi-level conformance be thoroughly examined. The general concern was, that to do a simple transfer of, say, 10Kb of string data from package A to package B on the same computer, on the same disc, as a temporary, one-off job, then the overheads in using SDTS would be unwarranted.

As an example, if users are not interested in data quality or external grid references but merely want to transfer some sectional ore outlines, how can they do a simple job without having to answer a number of irrelevant questions and producing up to seventeen data files?

The main problem with allowing any kind of "cut down" transfer is that, invariably, it will be used for the wrong purpose i.e. archiving, thus negating the whole idea. On the other hand, rigid adherence will dissuade users and SDTS will be overlooked in favour of one-on-one conversion routines.

One answer is that the output routine on the source system could intelligently check on what data is available to it. If it already has data quality and external reference parameters, it assumes it is a high level (of conformance) transfer and proceeds accordingly. If it does not have this data, it asks 'Is this a permanent file'. If 'No', then it will output blank fields; otherwise it will ask for the input.

There is an obvious need for an SDTS editor, so that a 'temporary' or low conformance level file can be upgraded to a 'permanent' or high level conformance file by addition of the necessary data. An SDTS editor should be available within twelve months. Note that the AMDEX project expressly excludes any software development in its current phase.

Data Dictionaries

SDTS encourages the use of standard data dictionaries, so that cooperating systems can recognise the 'meaning' of data transferred. As part of AMDEX, it is proposed to build a mining data dictionary.

A data dictionary is a collection of terms, usually within a theme (e.g. forestry, mapping, geology), each of which has a standard label, source definition and authority. It allows for synonyms and 'included' terms. It may be hierarchial, but not necessarily so.

The mining data dictionary will include terms such as "Toe", "Crest". "Batter", "Dip", "Strike", "Northing" etc. For many of these, not only will there be a definition, but also a standard convention will have to be nominated e.g. "Strike is in Degrees and Decimal Degrees, Clockwise fron North = 0. These conventions will form part of the AMDEX mining data dictionary, ensuring two systems communicate with the same understanding of terminology and sense.

Developing a data dictionary is a lengthy procedure, based on issuing drafts,seeking comments, issuing revisions etc. A formative mining data dictionary will be available in May, and finalised by December, 1992.

There is a corresponding need to develop other geoscience data dictionaries e.g. geological, geophysical, geochemical etc. It is hoped that AMDEX, together with SDTS will spur on activity in the area.

Profiles

SDTS allows for subsets (termed 'profiles') to be used. A profile is a convention to use only certain features of the entire standard for a particular application. Debate continues about the desirable number of profiles which should exist, and their certification (e.g. should they be registered with Standards Australia). If a proliferation of profiles occurs, then the benefits of the standard may well be diluted.

As an example of a perceived problem of SDTS for geoscience applications, the mapping of commonly used data structures to the standard reveals that the SDTS representation of a string (Spatial object 'LS') is unsatisfactory for normal geoscience applications. Frequently, it is necessary to define a series of ordered X,Y,Z points with attributes (and different values) at each mode. Typical examples would be down hole assays or face mapping traverses.

The SDTS 'LS' spatial object does not cater for this case, so such a string must be represented by a composite of points. Therefore, a mining profile well may exclude the 'LS' object. If it did, then there are problems if a user with a package based on the mining profile imported data from e.g. a GIS system which supported the 'LS' object - what would the target system do with the 'unknown' object?

It should be noted that the use of SDTS will not make up for major differences between systems. For instance, the package 'Surpac' has a data type called an 'exclusion' i.e. a closed string (polygon) digitised in an anticlockwise manner means it is to be subtracted from the total area (it is bounded by a clockwise polygon).

The whole object could be described in SDTS by a 'PG' object (a G-polygon). However, there is little point in encoding such a structure if the target system has no concept of 'exclusions'. So, to cater for Surpac, we must include 'G - Polygons' in the mining profile, yet some systems will not be able to handle them anyway.

Handling of 'Time' in SDTS

SDTS has no concept of 'time' as such. Preliminary examination in the AMDEX project shows that time can be adequately handled as an attribute. This is extremely important in any situation where time dependent readings are to be processed. The whole topic of n-dimensional space for representing point readings is an area for ongoing work. The Julius Kruttschnitt Mining Research Centre report that they require nine dimensions (based on taking stress readings at the same x,y,z at different times in different directions) .

3-D and Surfaces

Because points and nodes have X, Y, Z co-ordinates, SDTS can represent strings and polygons in any orientation e.g. northing sections. Therefore, such features as structure contours can be represented.

SDTS allows for the definition of 'composite objects', so the triangles of a triangulated irregular network (TIN) can be described by composite points, with each mode having attributes and the triangular facet itself having attributes. Not only can surfaces (topography, water table, faults) be represented in this way, but so can triangulated wireframe solids.

AUSDEC

In late 1991, the Australia New Zealand Land Information Council (ANZLIC) let a contract to Barrett, Fuller and Partners, a trans-Tasman resource engineering group, to establish an SDTS support centre. The Australian technical computer user's group, ACADS, was invited to participate in setting up the centre, bringing to bear its extensive experience and contacts in developing and implementing CAD standards (IGES in particular) and its capabilities as a technical conference organiser.

The resultant organisation, known as AUSDEC (the Australasian Spatial Data Exchange Centre) is in the process of establishing a range of educational, training, consulting and research facilities, with the overall aim of helping users understand SDTS and facilitate its penetration into the spatial data community.

AUSDEC will be represented on the relevant Standards Australia committees and play an active role in their activities. An implementation plan for SDTS in Australia and New Zealand is currently being developed and when ratified by ANZLIC will provide the framework for SDTS activities over the next three years.

AUSDEC is charged with developing fee-for-service activities within three years to achieve self-funding. It is intended to publish an SDTS newsletter, run a hot-line, present regular seminars around Australia and New Zealand, present papers on SDTS developments at all relevant seminars and publish technical reports regularly. Whilst it is envisaged that a broad-based low fee subscription scheme will fund basic education and promotional activities, high cost endeavours such as software development will require project funding, either from the user community or from research grants.

AUSDEC will work closely with the U.S. Geological Survey as it pursues its implementation plan for SDTS. The USGS has nominated key short, intermediate and long term goals and AUSDEC will provide the liaison between the Australian spatial data community and their programs. Discussions are underway to determine in which areas Australia can make the most contribution. One possibility is that of developing conformance testing software i.e. routines to check that transfer files produced by a system that purports to conform to SDTS actually do meet the SDTS specifications.

Summary

SDTS is about to make a substantial impact on the Australian and New Zealand spatial data user community. Whilst the topographic mapping and surveying disciplines, led by AUSLIG, are showing the way, the geoscience community is also well placed to gain immediate benefits. The need exists, and the AMDEX project is providing an encouraging pilot study in a limited area.

There are compelling arguments for all geoscience users to examine SDTS as the vehicle for development of standard interchange formats. Perhaps it may also prove to be suitable as an alternative to other, less formal, standards (e.g. laboratory interfaces, ASEG-GDF and various in-house formats).

With the benefit of a full-time support group, close involvement of Standards Australia, liaison with the USGS and the impetus of the GIS industry, geoscience users in Australia and New Zealand can look forward to more open systems over the next five to ten years.

References

Clarke A.L. 1990 - A New Australian Standard for Spatial Data Transfer, *Standards Australia Committee IT/4, Sydney.*

Clarke A.L. 1991 - Data Quality Reporting in the Proposed Australian Data Transfer Standard, *Proceedings of a Symposium on Spatial Database Accuracy, Department of Surveying and Land Information, University of Melbourne* ed. G. Hunter 252-259.

Hume R.G. 1991 - Review of the Spatial Data Transfer Standard for the Adoption or Adaptation to the Australian Mining Industry Data Exchange (AMDEX) Standard, *AMIRA Project P347, AMIRA, Melbourne*

Standards Australia, 1989 - AS2482 Geographic Information Systems - Geographic Information Systems - Geographic Data - Interchange of Feature Coded Mapping Data, *Standards Australia, Sydney*

Standards Australia, 1989 - AS3654 Information Processing - Specification for a Data Descriptive File for Information Interchange, *Standards Australia, Sydney*

ACCESS TO DATA THROUGH FINDAR

E P Shelley
National Resource Information Centre
GPO Box 858
Canberra ACT 2601

Abstract -- The aim of NRIC's FINDAR directory system is to provide up-to-date information on what resource data is available, who is responsible for it, where it is located, and how it may be accessed. FINDAR provides a comprehensive description of a dataset and permits retrieval on any or all of keywords, dataset attributes, and the dataset's spatial coverage. The two main directory nodes at NRIC currently contain descriptions of 300 datasets. External directory nodes are operating at ERIN and at Western Australia's Integrated Land Information Program. Access to FINDAR will be through the BMR/DPIE network for internal users and through X.25 or modem links for registered external users.

Keywords - Information management/ Natural resources data/ Databases/ Natural resources management/ Metadata/ Directories

Introduction

Increasing public and political concern for proper environmental management is resulting in a demand for high quality data on which land-use decisions and policy making can be based. As natural-resource datasets tend to be constantly evolving, the cost of keeping a centralised database up-to-date would be prohibitive. The National Resource Information Centre (NRIC) has addressed this issue by setting up a national directory of datasets relating to natural resources. It is linked to similar directories in other Commonwealth and state agencies and thus will provide up-to-date information on what resource data is available, who is responsible for it, where it is located, and how it may be accessed.

Special software -- FINDAR -- was developed for the directory. The software consists of three components: tables which describe the attributes of the datasets in considerable detail; a structured thesaurus of keywords with which to describe datasets; and a gazetteer of spatial definitions. Searching of the directory can be carried out on the basis of dataset attributes, keywords and location. Further information on the FINDAR software and on the directory concept can be found in Johnson, Taylor & Lovett (1990), Johnson & others (1990), Taylor & Hodgson (1990), Johnson & Robey (1991), and Johnson & others (1991).

A key feature of FINDAR which sets it apart from other directory software packages is its sophisticated spatial searching technique. This uses the SIRO-DBMS software (developed by CSIRO's Division of Information Technology) and a gazetteer of spatial definitions. This enables the spatial coverage of datasets as points, lines, map sheets or polygons to be precisely defined in the directory. A dataset description is retrieved whenever the search window intersects its defined coverage.

Scope of the directory

The objective of the directory is to provide up-to-date descriptions of datasets relating to Australia's natural resources. Natural resources include: geoscience, soils, forestry, vegetation, agriculture, fisheries, ecology, wildlife, environment, land cover, land use, climate, and hydrology. The directory will contain descriptions of 'framework' datasets (e.g. topographic), and also socio-economic data, primarily datasets which relate to geographic location.

In addition to descriptions of the large number of datasets under the custodianship of federal and state government agencies, it is intended that the directory will contain references to datasets held by universities and, where the datasets are publicly available, private companies.

Directory content

Many previous database directories were made available solely in hard copy form (e.g USGS, 1983; Shelley, 1985). There were thus practical limitations on the amount of descriptive information that could be provided for each dataset.

As the NRIC directory is primarily directed towards online interrogation, it provides for a comprehensive description of each dataset. The dataset attributes comprise:

Identification

Name, acronym, abstract, owner, other organisations associated with the dataset.

Data items

Name, description and, where applicable, spatial resolution of each item in the dataset. Items can be grouped where appropriate.

Spatial identification

Type of spatial referencing, projection, coordinate units, feature types.

Spatial coverage

General and detailed information on the area covered by the dataset.

Dataset information

Working form, working medium, size, applicable hardware and software, interchange format, supporting documentation.

Data currency

Custodian details, data collection start/end/update, future proposals, archive details.

Data lineage and quality

Data collection method, source material, data processing details, positional and attribute accuracy, consistency, completeness.

Ordering information

Access restrictions, output products and charges, supplier information, order procedure.

Keywords

Keywords describing the dataset suggested by person providing the entry. (Controlled keywords are allocated by the group or node administrator.)

Organisation/position information

Additional information about the dataset custodian or supplier organisation/position.

Record creation and update dates and other directory management information are also included. A user can search on any of the above attributes.

Searching the directory

The directory is searched on the basis of one or more of indexed keywords, dataset attribute fields, and spatial extents. An example of the query screen is in Figure 1.

Boolean AND/OR logic can be applied to the keyword and attribute field searches. In addition, a 'PICK' facility is available for most fields on the query screen; this presents the user with a selection of valid values and the selected value is transferred to the query screen with a single keystroke. Context-sensitive help screens are also available with a single keystroke.

Each of the three types of query is executed independently and the results combined to give a final 'hit' list. In addition, the results of different searches can be combined (union, intersection, or minus) to further refine the 'hit' list. These procedures produce a powerful search facility.

The results of a search are presented as a list of dataset titles (Fig. 2). Entries can be deleted to produce the final list from which full dataset descriptions can be printed or written to a file. An example of a full description is in Appendix 1.

Search routines can be saved (Fig. 3); this is a very useful feature when searches are complex and are required to be run regularly.

```
FINDAR F3121.39          DIRECTORY  QUERY  MAINTENANCE       16:01 11-MAR-92
DPIE Production
*---------------------------*-----------------------------------------------------*
|GEOCHEM                    |Kw =/<> Keyword Value    (Total Entries: 117   ) &/or|
| Hits Keyword         4    *--*----*--------------------------------------*--*---*
|      Field          55    |Kw| =_ |Geochemistry__________________________|Y_|___|
|      Spatial        65    |Kw| __ |______________________________________|__|___|
|      Combine         3    |Kw| __ |______________________________________|__|___|
|=================================================================================|
|      Field name            Comparer    Field value (Total Entries: 172   ) &/or|
*-----------------------------*--*----------------------------------------*------*
|Owner_Org____________________|=_|FED/BMR_________________________________| ___  |
|_____________________________|__|________________________________________| ___  |
|_____________________________|__|________________________________________| ___  |
|=================================================================================|
|Spatial Type     Id        Name     In State: QLD      (Total Entries: 122   )   |
*-----------*-----------*---------------------------------------------------------*
|MAP-250___ |SE55-14___ |TOWNSVILLE_______________________________________________|
|__________ |__________ |_________________________________________________________|
|__________ |__________ |_________________________________________________________|
*-----------*-----------*---------------------------------------------------------*
 Press [Exe_Query] to perform Keyword search, [Enter_Query] to combine/save hits
Refine & rerun search or press [Exit] to examine results
         Char Mode: Replace  Page 1                       Count: *1
```

Fig. 1. Example of directory query screen.

```
FINDAR P5116.09          SUMMARY REPORT OF SEARCH RESULTS         16:05 11-MAR-1992
DPIE Production          --------------------------------                  Page   1
For query : GEOCHEM

Entry Id  Directory Entry Name                                   Organisation
--------- ------------------------------------------------------ ------------
100A01    International Database for Igneous Petrology           FED/BMR
210A01    Petroleum Source Rock Database                         FED/BMR
225A01    BMR whole-rock geochemistry                            FED/BMR


                          --- END OF REPORT ---
```

Fig 2. Example of search results

The node concept

The national directory will comprise a national node at NRIC and subordinate nodes both at NRIC and at other federal and state/territory government agencies. Subordinate nodes that have been installed to date are:

- Federal Department of Primary Industries & Energy (DPIE);
- Environmental Resources Information Network (ERIN) (Department of the Arts, Sport, the Environment & Territories);
- Western Australia Integrated Land Information Program; and
- 'Others' -- a node installed at NRIC to hold descriptions of datasets under the custodianship of organisations that do not, as yet, have a home node.

FINDAR F3108.10 MAINTAIN/EXECUTE QUERY 16:00 11-MAR-92
DPIE Production

ACT	Query Name	Description	Hits	Date
	CHECKBMR	bmr_entries	55	10-FEB_17:04
	CHECKBRR	brr_entries	18	10-FEB_17:05
	CHECKMAPS	mapcat_entries	79	10-FEB_16:58
	CHECKNRIC	nric_entries	7	10-FEB_17:04
	GETALL	report_all_entries_in_dpie_prod.	174	10-FEB_16:58
	TEST1	marine/bmr	6	13-JAN_15:24
	GEOCHEM	geochem/townsville	3	11-MAR_16:04

Press [Help] to list available actions. (D, E, C, R, X, SO)
Char Mode: Replace Page 1 Count: *7

Fig. 3. Example of saved queries

The FINDAR software has also been installed at the South Australia Department of Mines & Energy for test purposes.

Dataset descriptions on subordinate nodes which have been flagged as suitable for transfer to the national node will be regularly uploaded to that node together with records that have been updated. Where agencies or states develop directories using other software, transfer routines will be developed to allow transfer of dataset descriptions to the national node. A record format and software for direct loading of dataset descriptions from other directories has been developed and tested.

We propose to develop import and export routines that use the international directory interchange format (DIF) currently being developed (NASA, 1990). This will facilitate links and exchange of data with directories operated by other agencies, e.g NASA's Master Directory and UNEP/GRID.

The national node will be the repository of the master keyword thesaurus and gazetteer of spatial objects. Amendments and additions to these tables will be regularly downloaded to the subordinate nodes to preserve uniformity. Keywords and spatial objects can be added at the subordinate node level; however, they only apply at that node and are not uploaded to the national node.

All nodes may be used for public enquiries although it is envisaged that the national node will be the primary contact point.

Directory status

The directory content at present is as follows:

- BMR datasets - 55
- BRR datasets - 18
- NRIC datasets - 7
- State/territory geoscience datasets - 220 (only about 10% of these are up-to-date)

Descriptions of some 1500 land-resource surveys are to be added shortly. In addition, major data capture programs are about to commence for soils and forestry data, datasets relating to the Cape York Peninsula Land Use Strategy, and datasets relating to the Resource Assessment Commission coastal zone enquiry.

The ERIN node contains 130 records and the ILIP node contains about 400. These will be uploaded to the national node in the near future.

Future information capture at the DPIE and 'Others' nodes will be driven by national priorities as reflected in NRIC and other departmental projects. A broad-based capture of descriptions of datasets covering all geographic areas and natural-resource disciplines is not practical.

The directory is now available for searching by NRIC staff on behalf of enquirers. Online access will be available for both *ad hoc* and regular users in the near future when updating of existing records has been carried out. The most efficient method of communication will be an X.25 link via Telecom's AUSTPAC service. It is possible to use a modem and ordinary voice line connection or a combination of voice line and AUSTPAC, but the response will be slower and its quality and reliability will be lower. The choice of communication link will depend on how often a user will want to search the directory. A user will also need a VT100 or VT220 terminal or a PC with VT emulation software.

References

Johnson, B D, Malmberg, H, Taylor, M, Hubble, C, & Lovett, B, 1990 -- The meta-database: a generic approach to information storage and retrieval. Paper presented at *Oracle 1990 Asia-Pacific Users Conference*, Adelaide, 11-14 November 1990.*

Johnson, B D & Robey A, 1991 -- Database management: an approach to information storage and retrieval - the FINDAR directory system. Paper presented at the *Government Technology Event*, Canberra, 26-28 February 1991.*

Johnson, B D, Shelley, E P, Taylor, M M, & Callahan, S, 1991 -- The FINDAR directory system: a meta-model for metadata. In Medyckyj-Scott, D, Newman, I, Ruggles, C, & Walker, D, 1991 (Eds.) -- Metadata in the Geosciences. Group D Publications Ltd, Loughborough, UK.

Johnson, B D, Taylor, M, & Lovett, B, 1990 -- Independent cooperating databases: requirements and responsibilities. Paper presented at *Oracle 1990 Asia-Pacific Users Conference*, Adelaide, 11-14 November 1990.*

NASA, 1990 -- Directory interchange format manual. *National Aeronautics and Space Administration Publication* NSSDC/WDC-A-R&S 90-27.

Shelley, E P, 1985 -- Directory of government geoscience databases in Australia. *Bureau of Mineral Resources, Geology & Geophysics, Australia, Report* 269.

Shelley, E P (Editor), 1986 -- Proceeding of the 3rd International Conference on Geoscience Information, Adelaide, South Australia, 1-6 June 1986. Vol.2: Keynote and theme papers, and general proceedings. *Adelaide: Australian Mineral Foundation.*

Taylor, M & Hodgson, P, 1990 -- An intelligent thesaurus: power through structure. Paper presented at *Oracle 1990 Asia-Pacific Users Conference*, Adelaide, 11-14 November 1990.*

USGS, 1983 -- Scientific and technical, spatial, and bibliographic databases and systems of the US Geological Survey, 1983. *Geological Survey Circular* 817.

* These papers are available from NRIC.

APPENDIX 1: EXAMPLE OF FULL DIRECTORY LISTING

```
NRiC FINDAR R5115.09          DIRECTORY LISTING          15:21 03-MAY-1991
                                                                    Page 1

--------------------------------------------------------------------------
 Id.: 1003A01       Name: Australian Geochronology
 Ref: OZCHRON      Owner: BMR                 Type: DAT     Status: Entered
--------------------------------------------------------------------------
 >> IDENTIFICATION <<

          Full Name: National Database of Australian Geochronology
           Abstract: Main tables are SAMPLES, K_AR, AR40_39, RB_SR, SM_ND,
                     ZIRCON, SHRIMP, REFERENCES and STOREBOXES. They are
                     briefly described under DATA ITEMS.
              Owner: FED/BMR (Bureau of Mineral Resources)
         Other Orgs:
    Other Orgs Desc:
  Authors/Collators: R W Page
     Other Dir Refs:
Master Dir Entry Id:
    Is this a Copy?:

 >> DATA ITEMS <<

          Item Name: Table SAMPLES
   Item Description: includes attributes for sample no., stratigraphic unit,
                     rock type, country, state, region, AMG reference and
                     decimal latitude and longitude
 Spatial Resolution: nominally 1:100 000

          Item Name: Table K_AR
   Item Description: laboratory results that go to make up K-Ar dating results
 Spatial Resolution: 1:100 000

          Item Name: Table AR40_39
   Item Description: laboratory results that make up Ar-Ar dating results
 Spatial Resolution: 1:100 000

          Item Name: Table RB_SR
   Item Description: laboratory results that make up Rb-Sr dating results
 Spatial Resolution: 1:100 000

          Item Name: Table SM_ND
   Item Description: laboratory results that make up Sm-Nd dating results
 Spatial Resolution: 1:100 000

          Item Name: Table ZIRCON
   Item Description: laboratory results that make up Pb-Pb zircon ages
 Spatial Resolution: 1:100 000

          Item Name: Table SHRIMP
   Item Description: laboratory results that go to make up Pb-Pb dates
                     obtained by ion microprobe
 Spatial Resolution: 1:100 000

          Item Name: Table REFERENCES
   Item Description: bibliographic references relevant to samples &/or results
 Spatial Resolution:

          Item Name: Table STOREBOXES
   Item Description: location of samples in BMR rock store
 Spatial Resolution:
```

>> SPATIAL IDENT <<

```
 Direct Spatial Ref: Y
    Projection Type: AMG (Australian Mapping Grid)
   Coordinate Units: DED (Lat/Long (Decimal Degrees))
                     PSW (Positive South/West)
                     *1:100 000 map no. and 6-digit AMG reference
    Spatial Feature: PNT (Point)
Topological Struct?:
   Topological Desc:
Indirect Spatial Id: AUS (Australia)
                     STA (State)
                     MAP (Map Sheet)
```

>>SPATIAL COVERAGE<<

```
       General Area: AUS (Whole of Australia)
   Onshore/Offshore: BTH (Both onshore and offshore areas)
Spatial Description: Australia-wide
  Future Spat Cover: possibly Antarctica and PNG
```

>> DATASET INFO <<

```
          Data Form: RDB (Relational Database)
      Working Media: MDI (Magnetic Disk)
     Supporting Doc: Ryburn, R J; Page, R W; Laynne, V; Shelley, E P
                     OZCHRON - A national database of Australian Geochronology

      Digital Data?: COM (Completely Digital)
           Software: ORACLE
           Hardware: DG MV20000
 Interchange Format:
               Size: 10 Mbyte
```

>> DATA CURRENCY <<

```
      Custodian Org: FED/BMR (Bureau of Mineral Resources)
     Custodian Posn:
 Data/System Status: INP (In Progress)
 Started Collection: 01-JUL-89
   Ended Collection:
   Update Frequency: continuous
   Future Proposals: possible direct updates from some universities
Archive Information:
```

>>LINEAGE/QUALITY<<

```
      Primary Data?: PAR (Partially Primary Data)
  Collection Method: FNO (Field Notes)
                     FSP (Field Specimens)
                     LME (Laboratory Measurements)
  Data Collect Desc:
     Source Dataset:
Source Dataset Desc: various scientific papers & publications referenced in DB

 Processing Details:
Positional Accuracy: 5 (10 Metres to 100 Metres)
                     6 (100 Metres to 1 Kilometre)
                     7 (1 Kilometre to 10 Kilometres)
Posn Accuracy Deter:
 Attribute Accuracy:
Logical Consistency:
       Completeness:
```

>> ORDERING <<

```
Access Restrictions: PUB (No Restrictions)
Access Restric Desc:
     Output Product: not yet determined
            Charges:
       Supplier Org: FED/BMR (Bureau of Mineral Resources)
      Supplier Posn:
    Order Procedure: enquiries to Dr R W Page, BMR
    Order Lead Time: 1 year
```

>> KEYWORDS <<

```
 Suggested Keywords: Geochronology
                     Rock samples
                     Isotopes
                     Radiometric dating
```

>> CONTROL <<

```
Visible to Public ?: Y
Allow Transfer Up ?: Y
Entry in Progress ?: N
      Dir Maint Org: FED/BMR (Bureau of Mineral Resources)
     Dir Maint Posn: UPDATER (Updater)
 Entry Completeness:
        Text Source: N (Data Entry by Node Administrator)
     Spatial Source:
      Last Verified: 09-JAN-90
      Verify Period:
     Verify Request:
     Record Created: 09-JAN-90

       Last Updated: 30-APR-91
         Updated by: FINDAR
        Update Node: A01
  General Checked ?: N
 Keywords Checked ?: N
  Spatial Checked ?: N
     Admin Comments:
```

>> SPATIAL DEFINITION <<

Type	Id	Name
AUSTRALIA	000020A90	AUSTRALIA inc TASMANIA

--- END OF REPORT ---

RESOURCE AND ENVIRONMENTAL INFORMATION: THE INTERNATIONAL PERSPECTIVE

Roger Bradbury

National Resource Information Centre
Canberra, ACT

Abstract: The international dimension of resource and environmental information is growing rapidly. Many of the issues in Australia, such as standards, ownership, directories, integration and so on, find counterparts overseas. This paper explores the consequences of the growing internationalisation of information for Australia.

INTRODUCTION

What I plan to do in this paper is to give an international perspective on the problem of resource and environmental information. It is easy to imagine that we are alone in facing the difficulties and challenges in this area. However, the problems are not confined to the geosciences or even the natural sciences: all disciplines that seek to know about man's interaction with his environment are involved. Nor are the problems confined to Australia: all the OECD countries are currently grappling with them, as indeed are some developing countries. This internationalisation of the problem itself feeds back into the problem, making it at once more interesting and more challenging.

There are three interacting areas that are impacting on the problem of resource and environmental information. In each of these areas, there is an international dimension of growing importance. I will discuss these areas in turn, and then examine how they interact and are likely to affect how we manage environmental information through to the turn of the century.

The three things I see impacting on our problem are these: first, there is a political sea change under way; second, there is an impressive technology push under way; and third, there is an emerging new scientific discipline.

THE POLITICAL SEA CHANGE

The political environment in which we are operating is changing. The idea of sustainable development is gaining acceptance rapidly as the world community gears itself up for the United Nations Conference on Environment and Development (UNCED) later this year. This has resulted in a whole systems approach to issues rather than disciplinary approaches. In turn the whole systems approach is a voracious consumer of information. Both of these consequences impact strongly on the way in which we handle environmental and resource information.

This sea change is being given concrete expression in various ways throughout the OECD countries. In Australia, it is seen in the creation of new organisations such as the National Resource Information Centre (NRIC), the Environmental Resources Information Network (ERIN), and the Resource Assessment Commission (RAC). It also comes through very strongly in the Australian Science and Technology Council (ASTEC) reports on environmental research in Australia (ASTEC, 1990a, b, c) and in the reports of the Ecologically Sustainable Development Working Groups (e.g. ESD (1991)).

In Montreal, in May 1991, the Canadian government sponsored a major international forum on environmental information (Environmental Information Forum Secretariat, 1991). Canada proposed this forum to the 1990 Group of Seven (G7) summit as a way of improving environmental information on a world-wide basis. This meeting developed a comprehensive framework for achieving a 'second generation' of environmental information, and prepared a communiqué in the form of an Environmental Information Statement.

Several perceptions crystallised at the forum, each of which has profound effects on the way we go about using resource and environmental information in our work:

- *Environmental information is a public policy issue in its own right.*
  It should sit on the policy table and be considered in and of itself without being filtered through other perspectives, such as, say, the economic or social aspects. That is to say, it should provide input to public policy deliberations coequally with other considerations.

- *The citizen has a right to know about his environment.*
  The citizen has an inalienable right to know, to understand and to access environmental information in an effective way.

- *Environmental information should be free.*
  The citizen has, in large measure, this belief goes, already paid for this information through her taxes. She should not be charged again merely to access it. The cost of accessing information, therefore, should be kept to a minimum and should reflect only the cost of the media and not the punitive cost of the information itself.

THE TECHNOLOGY PUSH

The second area that is impacting on managing environmental information is the technology push and it, too, is of fundamental importance in the way the whole process is managed and moves forwards. This technology push is apparent on all fronts:

- *Data acquisition*
  New sensors are acquiring information at rates unheard of a few years ago and there is no levelling out of this process. The rates at which we acquire environmental information are increasing exponentially as far forward into the future as we can see. We are moving from megabits of information to terabits without even stopping for breath at the gigabits.

- *Computer hardware*
  New computer platforms are emerging with processing rates of hundreds of megaflops with new architectures. (There are about 10 mips to a megaflop, most workstations are in the 10s of mips or single megaflops). Some of the massively parallel computers on the horizon now are about 3 orders of magnitude more powerful than the things available on scientists' desks now. Supercomputers are becoming available, transputers are becoming available. This hardware push is coinciding with the data push described above.

- *Computer software*
  New software is becoming available. In the domain of spatial information, it is blurring the distinction between raster and vector data models. It is blurring the distinction between database and GIS. Importantly, it is blurring the distinctions between data, database and GIS. Objects called metadata are now objects of intense interest. The idea of topology is being extended to aspatial domains. There is no necessary reason that we should consider our data to exist in a three dimensional euclidean space. Euclidean spaces are trivial. Problems are now being conceptualised in higher, more complex spaces. New geometries, such as Wagner, Finsler and Riemann, will coexist with euclidean ones. New ways of visualising these data are evolving rapidly, dependent utterly on the power of supercomputers.

- *New computing paradigms*
  New paradigms, particularly object orientated programming systems, are encroaching rapidly into areas of database, areas of GIS, into modelling and into analysis. These approaches are alien to those of us brought up on the certainties of Fortran, Basic and Pascal. They have crept in from the quiet corners of computing science and will change the way we compute.

These largely quantitative changes in technology will lead to qualitative changes in our science. They will allow us to do, not more science, but *different* science. Indeed some of the supercomputer users speak of their work as a new way of doing science.

THE EMERGING SCIENCE OF ENVIRONMENTAL INFORMATICS

In the beginning there was a vacant piece of intellectual turf. It got invaded on the one side by the surveying profession, hot from its triumphs of digital map making. It was also invaded on the other side by field scientists, mainly geologists and biologists, who had field notebooks bulging with raw data. Now they each had very different mindsets about what this turf was. For the one, the game is the end of a process: producing a digital dataset is the result. For the other, the production of a digital dataset marks the beginning of the real work of analysis and interpretation. For one, talk of updating digital datasets is talk of closer and closer approximations to reality, to truth. For the other, the talk is more of data existing as a time series: indeed the talk is of capturing change not eliminating it. One talks of owning the data, indeed places great emphasis on 'signing off' the survey - professional pride is embedded in that process. The other talks of sharing the data, indeed the more it is used and dissected by others, the more professional kudos accrue to the originator.

Both disciplines have converged on the same set of technologies, and now both coexist uneasily, with their very different world views, working with similar datasets as a new science emerges around them.

We might call this emerging science environmental informatics (as the europeans do) or earth system science (as the americans do) but whatever we call it, it offers new possibilities of treating the planet as a system. It too generates pressures, the most compelling of which is a pressure to go from pattern to process - to go from statics to dynamics, to put time into the equation. The demands of sustainable development are especially crucial here, since it forces us to try to hindcast and to forecast, that is to model, the world system in ways and on a scale unimaginable only a few years ago. This will increase our computing demands by three to six orders of magnitude in the next few years. It will also increase our data and software demands commensurately.

THE INTERNATIONAL DIMENSION

Together these three areas of change confront us with a series of problems in how we handle environmental and resource information. The solutions being developed all have strong international flavours.

First problem: environmental information is arcane

The data sources are hard to find. Terabit information flows overwhelm even the most experienced users. There are no easily understood standards for sharing and exchanging data. Moreover, environmental processes operate simultaneously over many scales from the most local to global and we need to comprehend them over those scales. The analysis, modelling and interpretation of environmental information is still an art, not yet a science. The visualisation of environmental information still needs the intercession of a priesthood, of a trained cadre of specialists who are the only ones able to handle these complications.

Possible solutions

Firstly, unchain the bibles and secondly, get rid of the priesthood. To unchain the bibles we need to build the things that will help find the information such as directories, catalogues and associated networks which can operate at all scales from local to global. NRIC, in collaboration with the states, is making an effort in this area with its national directory of natural resource data. NASA, together with Japan's NASDA and the European Space Agency, is making an effort internationally with its Master Directory. NASA's Master Directory and NRIC's directory will soon be linked. We will soon be able to find Australian information that exists overseas.

We also need to build things to handle the data flow. We need new graphical user interfaces and new DBMSs - an extremely active area of research in the computer science departments of the world. We need to get the bible translated into the patois by developing data sharing and interchange standards that reflect the needs of users. The emerging standards are all being generated overseas: the Australian spatial data transfer standard currently under development is a knock-off of the US one, and the VPF (Vector Product Format), the new data format for the new Digital Chart of the World (DCW) has been jointly developed by the US, UK and Australian armies. It will have an enormous impact on the use of spatial data because of its simplicity and cheapness, but will be an imposed standard in every sense.

This imposition highlights a particular tension. While some of these international efforts have a colonial aspect - by their sheer size and momentum, they impose themselves - others have an empowering aspect - they liberate scientists to work more fruitfully. These efforts provide us with tools, and we need to find ways to encourage them even though they originate overseas. For every tool that passes into the hands of the users, we have a need for one less priest. We need to encourage education and training of users at home even as we encourage the development of user-friendly software and hardware globally. For example, we need to encourage the development of the easy-to-use analysis, modelling and visualisation technologies being developed by research centres in the United States and Europe. The

National Center for Supercomputing Applications in Illinois, the Ohio Supercomputer Center and the International Institute for Applied Systems Analysis in Austria are good examples.

Second problem: Ownership of environmental information

There is a fundamental inconsistency between the idea of ownership of information and the idea of a citizen's right to know. There is also a fundamental inconsistency between the policy of cost recovery and the idea of freeing up the sharing and exchange of natural resource information. The current stop-gap in the government area (where most of these data are held) is the idea of custodianship. The problem is seen in its most pathological form when a government cannot afford to buy its own data. Moreover, beyond this ludicrous situation lies a serious micro-economic reform issue: the working of efficient markets requires perfect information, and to the extent that information flow is restricted by whatever process, then those markets themselves become inefficient.

Possible solutions

In a sense this is not the problem it seems because individuals will not put up with it. This situation will change and it will change sooner rather than later. The present confused situation is mostly the result of society's legal and social structures not being able to keep up with scientific change. This whole creaky edifice of information ownership must soon collapse. This is because the instruments, such as copyright, that were invented for Caxton's revolution are no longer relevant in an era where I can have access to more information in one day than was ever assembled in all the world in all its yesterdays.

If scientists cannot afford data from official national data centres, they will ignore them and acquire data overseas. The range of data becoming freely available from international sources will soon rival the offerings of traditional national suppliers. For example, the Global Resource Information Database (GRID) established in Nairobi by the United Nations Environment Program has a significant holding, while the World Conservation Monitoring Centre in Cambridge has more specialised holdings. NRIC has close links with both these organisations.

But these are interim solutions. What should we be seeking beyond these bandaids? Here are some elements that will be found in any real solution.

- A recognition that information is evanescent, ephemeral and will pass through our hands at an ever increasing rate

- An insistence that the metadata (that is, the descriptions of the data) are captured and distributed routinely. Murray Darling Basin Commission and the ASTEC report are both arguing for this as a condition for granting of research funds.

- Encouragement of the sharing and exchanging of environmental information between levels of government and ultimately beyond government. NRIC has some role here in that it has the carriage of an MOU between the Commonwealth and States for the sharing of this information.

- A norm of adding value to information as it passes through our hands and then forgetting about it. For example, NRIC has released a national digital geology data set with the BMR and a national digital soils set with CSIRO. Copyright of environmental information is about as sensible as copy protection of software. It is a 16th century solution to a 21st century problem. The most successful software houses learnt this years ago and they have ignored copy protection and concentrated instead on constantly improving their product so that they have repeat sales.

Third problem: Policy for environmental information is dominated by a producer cartel

This is an understandable thing historically since the producers were the first to see the possibilities of digital data. Because there is a producer cartel, there is consequently no real market in information and there is no real way to assess its value. A market is now emerging as agencies, both private and public, both at home and abroad, begin to generate digital data. Think of SPOT Image from France, EOSAT from the US, and the several international efforts like GRID and the DCW I mentioned above.

As markets emerge, I would expect "prices" to tumble by a couple of orders of magnitude at least. More fundamentally, their may be no market in the base information at all since the surveyors may give if away. EOSAT, the US agency, has just gone broke trying to sell satellite imagery and I understand will once more give it away for the price of the medium.

Possible solutions

Two major reforms are needed to accelerate the breakdown of this cartel:

- We need to rebuild the organisational structures, the gatekeepers of technology and science, so that users as well as producers are represented on, and have an effective say in, these gatekeepers and the ways in which environmental information flows throughout our society.

- Agents of change in this game, such as NRIC and ERIN need to be encouraged as this transitional landscape evolves, while the traditional producers of environmental information need to be reassured that the changes will ultimately benefit them too.

CONCLUSIONS

First, I think it has now become possible to make decisions about complex natural resource issues based on environmental information. That sounds trite to an audience like this, but to have complex decisions based on data is, I would suggest, a huge change in the way we build our society.

Second, do not underestimate this change.

Third, this whole process affects policy makers as well as policy. These in turn feedback strongly and positively (I stress the positively) to the technology and science. So the technology, science and politics of the problem are linked **but** they are linked in a positive feedback loop. Those of you who remember enough of your elementary general systems theory will remember that positive feedback loops are intrinsically unstable.

We are in for a rocky road over the next few years.

REFERENCES

ASTEC (1990a) Environmental science in Australia: A compendium (Australian Government Publishing Service, Canberra)

ASTEC (1990b) Environmental science in Australia: A review (Australian Government Publishing Service, Canberra)

ASTEC (1990c) Environmental science in Australia: The issues (Australian Government Publishing Service, Canberra)

Environmental Information Forum Secretariat (1991) Environmental Information Statement (Environment Canada, Ottawa)

ESD (1991) Ecologically sustainable development working groups: Draft report - executive summaries (Australian Government Publishing Service, Canberra)

NATIONAL TOPOGRAPHIC DATA - COLLECTION AND AVAILABILITY

Andrew L. Clarke

Australian Surveying and Land Information Group

Box 2, Belconnen ACT 2616

ABSTRACT

The Australian Surveying and Land Information Group is a major producer of national spatial data products which support the needs of geographic information system, cartographic and geoscience users. These products include topographic, administrative boundary, resource and environmental, and satellite remote sensing data. This paper briefly describes the range of products; details the Australian Geographic Database program; and concludes with some comments and how they support national geoscience activities.

INTRODUCTION

The Australian Surveying and Land Information Group (AUSLIG) is a major producer of national spatial data products. These products include topographic, administrative boundary and resource/environmental data (marketed as AUSMAP Data products), aerial photography, and satellite remote sensing data (marketed as ACRES products). Table 1 details present AUSMAP Data products and table 2 details present and future ACRES products (from Bell and Puniard, 1991).

AUSTRALIAN GEOGRAPHIC DATABASE PROGRAM (AGDB)

Introduction

The arrival of the AGDB marks a significant change of direction for AUSLIG. While AUSLIG has been producing digital data products for many years, the AGDB is different in a number of fundamental ways:

all AGDB data will conform to high levels of quality and documentation so they can be easily used in a wide variety of applications;

AGDB products will be available in a relatively short timeframe;

the AGDB program will be client driven and feedback from a wide range of potential customers will be the major determining factor for product specifications and priorities.

Table 1. Present AUSMAP Data products

The following is a summary of existing (and proposed) AUSMAP data products.

Description	Capture Scale	Transfer Format
Topographic		
1. Baseline and Bay Closure Lines	1:100k	AS2482
2. Coastline	1:100k	AS2482
3. DEM - Critical Aeronautical Heights	NA	Formatted ASCII
4. DEM - Spot Heights	NA	AS2482
5. DEM - 18" Grid	NA	AUSLIG DEM format
6. Master Names File	1:100k, 1:250k	Formatted ASCII
7. Topographic Base	1:250k	AS2482, GINA
9. Topographic Base	1:1m	AS2482, GINA, ARC/INFO
10. Topographic Base	1:2.5m	AS2482, GINA
11. Topographic Base	1:10m	AS2482, GINA
Administration		
12. AWRC Representative Drainage Basins	1:250k	GINA, ARC/INFO
13. Census Boundaries 1986	Various	AUSLIG Census Format
14. Commonwealth Electoral Boundaries	Various	GINA
15. Legal Local Government Areas	Various	GINA
16. National Estate Areas	Various	GINA
17. Public Lands (Land Tenure)	1:5m	GINA
Resources and Environmental		
19. Dams and Storages	NA	GINA
20. Minerals	1:2.5m	GINA
21. Vegetation - Natural	1:5m	GINA, ARC/INFO
22. Vegetation - Present	1:5m	GINA, ARC/INFO

Table 2. Present and future ACRES products

Present products

Satellite/sensor	Sensor Bands	Ground Resolutions (m)	Swath Width (km)	Repeat Cycle
NOAA/AVHRR (USA)	5	1100	240	4 per day
Landsat 5 (USA)				
MSS	4	80	185	16 days
TM	7	30	185	16 days
SPOT 2 (France)				
XS	3	20	60 x 2	26 days off nadir on request
PA	1	10	60 x 2	as above
MOS-1 and 1b (Japan)				
MOS-1/MESSR	4	50	100 x 2	17 days
ERS-1 (European)	C band	30	100	3/35/176 days

Future products

Satellite/sensor	Sensor Bands	Ground Resolutions (m)	Swath Width (km)	Repeat Cycle
Synthetic Aperture Radar (July 91)	57 mm			
Landsat 6 (USA) (July 92)				
TM	7	30	185	16 days
PA	1	15	185	16 days
ERS 1 (Japanese)				
OPS	8	20	75	44 days
SAR	L Band (300mm)	20	75	44 days

(Note: OPS have in track stereo capability)

SPOT 3 (January 93). Details as per SPOT 2

Product family

The AGDB is a family of spatial data products derived from AUSLIG's series of topographic and thematic maps at scales ranging from 1:100,000 to 1:10,000,000. The product family is described conceptually in table 3.

Table 3: AGDB product family concept

Theme	1:100k	1:250k	1:1M	1:2.5M	1:5M	1:10M
Hydrography	✓	✓	✓	✓	✓	✓
Infrastructure	✓	✓	✓	✓	✓	✓
Relief	✓	✓	✓	✓	✓	✓
Administrative Boundaries		✓			✓	
Resource & Environment			✓		✓	✓

Product quality

Structure. AGDB data will be suitable for use within GIS applications. The data model for AGDB vector products is feature-based and includes strict adherence to a topological data structure. Each feature consists of spatial and attribute components. For distribution purposes, features will be grouped into themes. Each theme will contain logically-related geographic information.

AGDB data will comply with the following rules for spatial data integrity:

data will be "clean" (ie. no overshoots, undershoots, broken lines, pseudo nodes or artefacts);

data will have a node/chain structure (ie. all linear features within the same network layer will be broken by a node at intersections or at the point where an attribute of the feature changes);

polygons bounding area features will be closed;

every polygon will have a centroid (polygon label point);

multiple digitising of lines within the same network layer will not occur.

Data point reduction filters will be applied so that locational information is conveyed by the minimum number of points.

Features that cross the edges of tiles (ie. map areas) will be positionally and attribute matched to adjacent tiles.

Accuracy and reference system. The positional accuracy of most products will be the sum of errors from two sources:

accuracy of the source material;

accuracy of the data acquisition process.

This means, for example, that no more than 10% of well defined features in the 1:250,000 scale hydrography and infrastructure themes will be more than 132 metres from their true position. An absolute accuracy check will be undertaken on a sample of features to ensure that the accuracy statement is correct.

Products will be available in either geographic or projection co-ordinates. The 1:250,000 scale product, for example, will use geographic co-ordinates based on the Australian Geodetic Datum (1966) and Australian Map Grid projection co-ordinates. Parameters for translation of co-ordinates to the geocentric World Geodetic System (1984) will be provided.

Documentation. All products will be supplied with data quality information which will define the lineage, positional accuracy, attribute accuracy, logical consistency and completeness of the data. Data quality information will be provided in three ways:

as part of the user documentation;

in a data quality statement file that will accompany each spatial data file;

as a data quality attribute that will be held against each feature.

Each product will be provided with comprehensive user documentation.

Operations. Quality checks will be carried out at critical stages during production. A quality assurance team will independently validate data from production areas against product specifications before the data is archived. Data will be validated before it is supplied to users to ensure that it is correctly formatted and that the transfer media is sound. All products will undergo beta testing in selected user sites before being released. These procedures are designed to meet the requirements of the Quality Management and Quality Assurance Standards as issued by Standards Australia, in particular AS3901.

1:250,000 scale product

Availability. This will be the first major AGDB product to become available. The product will be available from 1 August 1992 and national coverage will be completed by 30 June 1994. Figure 1 shows the planned timing of availability of hydrographic and infrastructure themes. The relief theme is planned to be available nationally by 1 September 1992.

Figure 1. Planned timing of availability of 1:250,000 scale hydrography and infrastructure themes

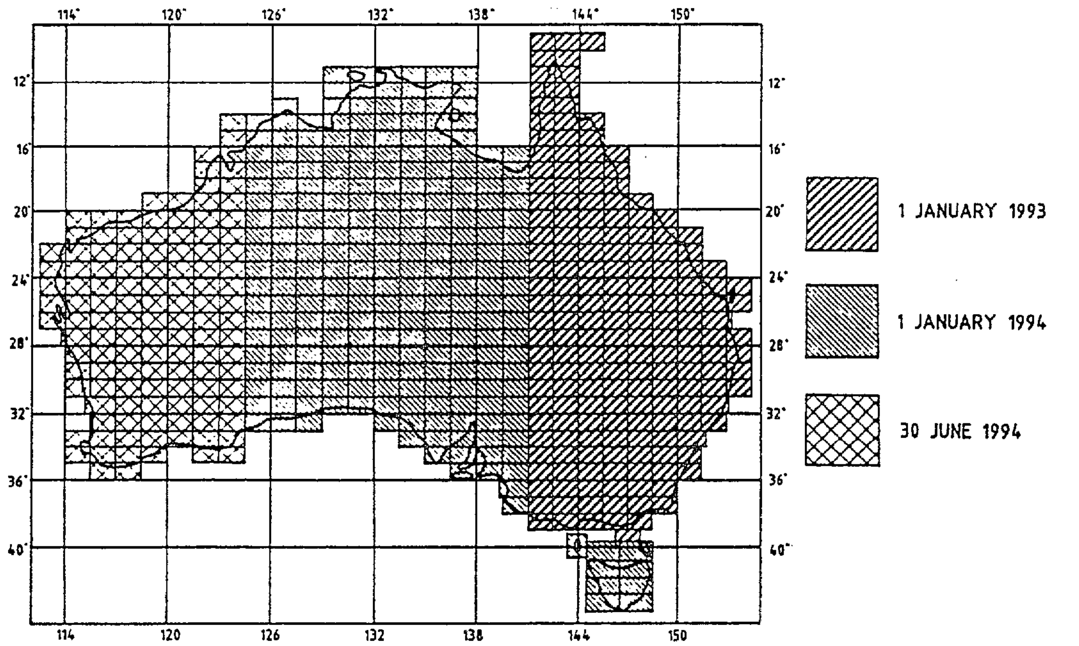


Table 4. Features and attributes envisaged for 1:250,000 scale hydrography and infrastructure themes

Hydrographic Feature and Theme Layer Structure

Layer	Layer Type	Features	Attributes	Spatial Objects
Drainage	Linear Network	Watercourse		Chain
		Canal		Chain
		Connector		Chain
Waterbody	Polygon	Lake	Hydrographic category, Name, Water Quality	Polygon
		Watercourse	Hydrographic category, Name	Polygon
		Foreshore Flat		Polygon
		Swamp	Hydrographic category, Name	Polygon
		Saline Coastal Flat		Polygon
		Land Subject to Inundation		Polygon
		Reservoir	Name	Polygon
		Mangrove		Polygon
		Waterline		Chain
		Junction		Chain
Framework	Polygon	Coastline	Coast Type	Chain
		Island	State/Territory, Name	Polygon
		Mainland	Name, State/Territory	Polygon
		Sea	Sea/Ocean Name	Polygon
		State Border		Chain
		Tile Edge		Chain

Infrastructure Feature and Theme Layer Structure

Layer	Layer Type	Features	Attributes	Spatial Objects
Road	Linear Network	Road	Number of lanes, Name, Road usage, Surface, National highway number	Chain
		Bridge 1	Traffic, Name	Node or Chain
		Tunnel 2	Traffic, Name	Node or Chain
		Ferry Route	Name	Chain
Railway	Linear Network	Railway	Tracks, Name, Status, Gauge	Chain
		Railway Station	Name	Node
		Bridge 3	Traffic, Name	Node or Chain
		Tunnel 4	Traffic, Name	Node or Chain
Aeronautical	Point	Aircraft Facility	Aircraft Facility Type, Name	Point
Localities	Polygon	Locality	Name	Point
		Built-up-area	Name	Polygon
Framework	Polygon	Coastline	Coast Type	Chain
		Island	Name, State/Territory	Polygon
		Sea	Sea/Ocean Name	Polygon
		Mainland	Name, State/Territory	Polygon
		State Border		Chain
		Tile edge		Chain

1 With attributes of road or road/rail

2 With attribute of road

3 With attributes of rail or road/rail

4 With attribute of rail only

Content. Table 4 contains details of the features and attributes envisaged for the hydrographic and infrastructure themes. The relief theme contains spot heights and points at changes of grade.

Depending on the level of interest, additional themes may be added to the product. Themes currently being considered are:

> monochrome raster underlay: this theme could consist of a black and white raster image of the key features on the printed map sheet;
>
> colour raster underlay: this theme could consist of a full colour raster image of all features on the printed map sheet;
>
> gridded DEM: the relief theme could be enhanced by the computation of a regular, gridded digital elevation model; a grid of 250 metres could be produced from the relief theme.

Due to resource and timeframe constraints not all features shown on the printed map are included in the product. At this stage the main features that are being considered for omission are:

cartographic names: names of features and places are included as feature attributes only; no information is given about the cartographic position or representation of names;

vegetation;

50 metre contours;

buildings: only buildings important enough to be shown on 1:1,000,000 scale maps (usually homesteads) are included.

Currency. Initially the currency of information in the product will be the age of the associated printed map. Some features and attributes, however, will be brought up to date as the product is being acquired (eg. changes to major roads, railways, boundaries of major cities, and new populated places).

AUSLIG currently plans to provide an annual update service for this and other AGDB products. The update service would include:

revision of positional and attribute information, with special emphasis on key features such as roads;

addition of attribute information to certain features where initially the attribute may have been classed as "unknown";

correction of any errors that may have been detected.

In addition to the traditional update method of full resupply of data, a "change only" update service will be offered which will utilise the unique feature identifier attribute held against each feature.

Other products

1:100,000 scale product. At this stage most other products are still in the conceptual stage. There are indications, however, of strong demand for certain themes of a 1:100,000 scale product over certain areas. This is likely to result in release of a 1:100,000 scale product as soon as national coverage of the 1:250,000 scale product has been achieved, or earlier if resources are available. The main differences between these two products are likely to be:

the positional accuracy of the 1:100,000 scale product will be higher;

the 1:100,000 scale product would eventually contain almost all of the features on the printed map.

Sale and distribution

Price. Under current government policy, AUSLIG is obliged to recover a proportion of the costs associated with generating, distributing and maintaining AGDB products. The current policy does not discriminate between government and non-government users. Within these constraints, however, AUSLIG intends to encourage as wide as possible use of AGDB data. To this end, users are presently being asked for their views on several approaches to pricing the 1:250,000 scale product.

A number of basic concepts will apply to use of the product:

- copyright is retained by AUSLIG;
- a one-off payment provides a user with a licence to use the product on an ongoing basis within the user's organisation; external uses such as generating publications, repackaging data and on-selling require separate negotiation;
- the licence fee varies depending on the extent of use - single-user (base fee), site (two times base fee), corporate (three times base fee);
- bulk discounts apply to large orders;
- genuine researchers can obtain small amounts of data for a nominal cost.

In addition to the one-off payment, consideration is being given to an annual licence fee which would be available at a cheaper price.

AUSLIG's current pricing policy for data produced under Community Service Obligation programs will be reviewed during 1992. The primary alternative to current policy is one in which data are exchanged by all budget-funded agencies, including between State and Federal agencies, at the marginal cost of transfer for use in non-commercial government activities. Commercial use by government agencies, or sale to the off-budget government and private sectors, would be at a more commercial price. The new Commonwealth Spatial Data Committee and the Intergovernmental Agreement on the Environment will provide the political framework for this review.

Transfer format and media. Products will be supplied in a variety of national and vendor standard formats, including:

- AS2482 (Standards Australia, 1989);
- ESRI ARC/INFO;
- GeoVision GINA.

The proposed new national standard, the Spatial Data Transfer Standard, will be supported as soon as it is released.

Other vendor specific formats are being considered and will be implemented given sufficient user demand.

Products will be supplied on a variety of media, including magnetic tape, floppy disk, and unix cartridge. Other media will be utilised given sufficient user demand.

CONCLUSION

The national spatial data products from AUSLIG may be used in a variety of geoscience applications, such as:

- geological analysis and mapping;
- oil, gas, and mineral exploration;
- mining and infrastructure development;
- natural resource management;
- environmental impact studies;
- pollution monitoring and toxic waste management;
- resource and environment policy development.

Programs such as the National Geoscience Mapping Accord (NGMA), and the Cape York Peninsula Land Use Strategy (CYPLUS), require base data sets in the form proposed for the AGDB, and those currently available as AUSMAP Data and ACRES products. AUSLIG is actively consulting with NGMA and CYPLUS participants to ensure that product availability best meets user needs.

REFERENCES

Bell, K.C. and Puniard, D.J. 1991. Geographic information and remote sensing in support of the geosciences. Proceedings of the National Conference on the Management of Geoscience Information and Data, Adelaide.

Data Models, Attributes and Standards: Workshop Overview

David M L Berman

Head, Information Systems Branch
Bureau of Mineral Resources
GPO Box 378, Canberra City 2601

This workshop has brought together several major themes and issues in the presentations and discussion over the three days.

Standards Initiatives

The workshop event has been used to hold, in parallel, the first meeting of a Working Group on Geoscience Data Standards. This group was convened by the Government Geologists Database Policy Advisory Committee (GGDPAC), with the task of collating and promoting, to Australian Standard, core geoscience data classes and attributes (Schmidt). Interaction has taken place between the working group and workshop attendees. In particular, advantage was taken of the expertise present to survey opinion on proposed core data classes and attributes. Outcomes from this interaction were discussed in the workshop, and will be reported on through the GGDPAC newsletter (Geoscience Database News). Further consultation will take place between government and industry under the Standards Australia umbrella.

GIS & RDBMS

Data management software has been discussed at length. GIS systems are seen as "the way to go" for managing, relating and analysing spatial attributes and layers (Williams, Gallagher, Parker, Rattenbury& Williams, Reeves, Morse, Stagg) - for traditional geology, geophysics and marine geoscience. Relational database systems are modern tools for managing traditional attribute data collections. Their role in maintaining and helping define de facto standards in setting up authority tables was also emphasised (Ryburn, Blewett, Belperio, Lenz & Pain, Withnall). Parker reminded us in discussion not to overlook existing standards (eg cartographic codes and colour attributes) when adopting new technology.

Data Models

Papers presented brought out the overall importance of data models - conceptual views - that allow integration of data to become a reality. The distinctions between data themes in a GIS (Arc/Info - Belperio), layers in a cartographic package (Intergraph - Hillier), types in datasets (vector/raster - Creasey & Chopra), and entities in a RDBMS (Oracle - Blewett & Ryburn) were brought out. The technology to enable integration of data has now arrived (Musto) - but robust data models must first be designed. Examples of this process were discussed by Gallagher (Arc/Info data layers, data dictionary and descriptor file design; Intergraph and attribute exchange problems), and Ryburn (ensuring spatial keys are used to link attribute tables, for GIS access). Musto and Ryburn also questioned whether current GIS software was adequate for 3-D geoscience models and process representation.

Data Quality

Several speakers raised the issues of data quality and lineage as being critical. Kelly described the problems of data matching when working with poorly checked cadastral data. Nanninga, and Brodie & Tucker, gave good examples of projects where it was impossible to compare data in regional studies from sources where different accuracies and classification techniques had been employed. A plea for common digitisation standards, from James, also alerted the workshop to another source of error. Examples of the effect of scale and map projection were also described (Gallagher, Lemon). Musto raised the practical problems of agencies such as NRIC being forced to digitise poor source data (geology, soils) which then tended to be taken as authoritative, being digital, rather than questioned and updated.

Data Exchange

Forthcoming standards to enable effective data exchange and archiving were detailed (Wood & Foley) in the context of models to describe data and metadata, based on the Spatial Data Transfer Standard (SDTS). The workshop followed on from an AUSLIG/ AURISA workshop on the SDTS, which gave the presenter, Jim McDermott (USGS), an opportunity to attend and contribute to the discussion.

The AMIRA Australian Mining Data Exchange project (AMDEX) was also outlined by Foley, as an example of an immediate practical application of the SDTS model. In the context of the NGMA data exchange time frames, however, it was observed by the meeting that the SDTS is not likely to be fully implemented in widely used commercial software, quickly enough. The importance of intermediate format translations (eg DXF for data exchange), and compatible data dictionaries where organisations use the same product (ie Oracle and Arc/Info most commonly) was stressed as the practical interim alternative.

Data Providers

Early in the workshop, industry's interest in the pricing policies of government agencies (State surveys, BMR, AUSLIG) was raised as an issue (Rutland, Mitchell). Recent W.A.,Tasmanian (James) and forthcoming NSW policy was referred to (Kelly), and the contrasting lack of a Commonwealth policy was noted. However, developments late last year towards setting up a Commonwealth Spatial Data Committee to deal with these issues were mentioned. In his paper, Clarke outlined the role and products of AUSLIG as a data provider . At a higher level of abstraction, the role of NRIC in providing metadata about such datasets was described by Shelley. The need to coordinate national data collection programs between agencies working on different layers came out in discussion of AUSLIG's Australian Geographic Data Base. Speculation about the issue of future information ownership and pricing was quite invigorating (Bradbury), and well fielded by Clarke, who mentioned in discussion new initiatives at the policy level which may see public good data become more freely available.

In Summary

- formal structures are required to make things happen in a coordinated manner at the national level (GGDPAC/ AUSDEC/ AMDEX/ CSDC).

- we need directories to tell us what data exist (FINDAR), and details about it (data dictionaries, SDTS).

- we need to know who our clients are, their data needs, and their technological capabilities. This will extend to interaction with other disciplines, and industry/ government cooperation.

- GIS (spatial), RDBMS (attribute) and CAD (cartographic) software and data models are linkable technically, as long as compatible data models are used - so we must exchange models!

- data quality, capture techniques and issues of scale must be considered at the start of all projects.

- we need to build standard envelopes to package our data (SDTS, and profiles such as AMDEX) for exchange, so that models can accompany the data.

In Conclusion

The workshop successfully raised a number of cross-disciplinary issues to a group of people who have a common interest in spatially related data, but who rarely have cause to meet. Whilst the workshop focused on National Geoscience Mapping Accord needs, the problems, principles and solutions will apply to general geoscience and related natural resources data.

Closing Address to the Workshop

Neil Williams

Associate Director
Minerals and Environment Group
Bureau of Mineral Resources

A good measure of the success of a conference is the number of people still remaining towards the end - by this measure alone we have done well !

This workshop has brought together professionals in the fields of geoscience, computing and cartography from government (Commonwealth and State), industry and university to address the issues raised by the increasing use of geographic information systems. The framework for this has been the National Geoscience Mapping accord: a joint Commonwealth State initiate to provide a new generation of geoscience maps for Australia.

It is clear from the level of support for this workshop that key issues have been raised and that there is a growing awareness of the potential and problems associated with the use of geographical information systems.

Important issues that emerged include:

- the growing need for digital data, particularly the demand from Government and semi-government environmental and planning agencies;

- the need for data standards and data compatibility in the geosciences, promoted to Australian standard in conjunction with industry;

- the increasing responsibility of organisations to use structured methods of providing data to ensure full documentation of data;

. the recognition that the best custodians of data are the primary data gatherers/users and the increasing demands for data custodians to ensure data integrity, data availability and maintenance;

. the issue of availability of data - the need for better national agreement on data sharing, copyright, and costs;

. the need for directories of data availability eg FINDAR;

. the explosion of data quantity - especially space imagery. The rate of production is accelerating, and we must address these problems now.

This workshop has identified a number of issues and made some recommendations touching on standards for digitised compilation of maps, exchange of information about data base structures, and the need for ongoing communication. One obvious follow on for this workshop is the need for followup workshops/meetings of specialist groups to address specific issues. The important thing is not lose the momentum and to build on the obvious goodwill evident at the workshop.

Finally, I would like to thank the organisers, all the people who have helped behind the scenes in the smooth administration of the workshop, as well as David Berman and Lynton Jaques who have been more visible up front.

Thank you for attending and taking part.

I hope the contacts initiated here will be developed and strengthened.

A safe trip home.

Participant	Position	Organisation	Telephone	Fax
Mr Garry Andrews	Cartographer	Dept Mines & Energy GPO Box 3783 Darwin NT 0801	089 895 339	089 814 806
Ms Heika Apps	Technical Officer 3	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1370	
Mrs Vera Ashby	Technical Officer 3	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1204	06 247 2728
Mr John Bain	Project Manager	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9282	06 257 6465
Mr Geoff Baisden	Graphics Project Manager	Dept of Resource Industries GPO Box 194 Brisbane QLD 4001	07 237 1396	07 221 9517
Mr Stephen Bandy	Project Leader	Dept of Mines (W.A.) Mineral House 100 Plain St East Perth	09 222 3266	09 222 3342
Mr Rob Barnes	Geologist	NSW Geological Survey Mineral & Energy House PO Box 536 St Leonards NSW 2065		02 901 8468
Mr Keith Bell	Manager AUSMAP DATA	AUSLIG PO BOX 2 BELCONNEN ACT 2616	06 201 4330	06 201 4366
Dr Tony Belperio		Dept of Mines & Energy 191 Greenhill Road Parkside S.A. 5063	08 274 7535	08 272 7597
Mr David Berman	Head Information Systems	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9602	06 257 6376
Mr Phil Bierwirth		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9231	
Dr Richard Blewett	Research Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9713	
Mr Evert Bleys		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9505	06 257 3667
Mr Bob Blight	Supervising Cartographer	Dept Resource Industries GPO Box 194 Brisbane QLD 4001	07 237 1669	07 221 9517
Mr Lyn Bond	Principal Geologist	RGC Exploration Pty Ltd PO Box 590 Fyshwick ACT 2609	06 280 6000	06 280 5915
Dr Roger Bradbury	Director	NRIC GPO Box 858 Canberra 2601	06 272 4688	06 272 4687
Mr Chris Brittan	Computing Specialist	Dept Mines & Energy GPO Box 3783 Darwin NT 0801	089 895 339	089 814 806
Mr Ross Brodie	PO 2	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9396	
Mr Lyle Burgess	Manager Remote Sensing	Western Mining Corp 168 Greenhill Rd Parkside SA 5063	08 372 7200	08 372 7247
Ms Karen Butler	Hydrogeologist POG1	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9683	06 248 8178

Participant	Position	Organisation	Telephone	Fax
Mr Charles Butterworth	Database Support	Schlumberger SEACO Level 3, 312 St Kilda Road Melbourne Vic 3004	03 696 6266	03 690 0309
Mr Max Bye	Princ. Geol.Inf. Systems	BHP Minerals PO Box 619 Hawthorn VIC 3122	03 810 7700	03 810 7722
Mr Graeme Carey	Analyst Programmer	CRA Exploration Pty Ltd 139 Canberra Avenue Fyshwick ACT 2609	06 239 6316	06 295 2655
Mr Chris Carter		Bureau of Mineral Resources PO Box 378 Canberra City 2601		
Mr Greg Cayzer	Chief Draughtsperson	Aberfoyle Res Ltd 123 Camberwell Rd East Hawthorn VICTORIA	03 882 2226	03 813 1086
Mr Bill Chestnut	Principal Geologist	NSW Geological Survey Mineral & Energy House PO Box 536 St Leonards NSW 2065	02 901 8343	02 901 8777
Dr Prame Chopra	Senior Reseach Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9540	06 257 6376
Mr Drew Clarke	A/g General Manager	AUSLIG PO BOX 2 BELCONNEN ACT 2616	06 201 4262	06 2014368
Mr John Clarke	Deputy Director	Geological Survey of W.A. 100 Plain Street East Perth 6004	09 222 3171	09 222 3633
Mr Richard Cooper		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9765	06 257 6120
Mr Denis Copley		Dept of Mines & Energy 191 Greenhill Road Parkside S.A. 5063	08 274 7535	08 272 7597
Ms Ann Cowling		ABARE Macarthur House Lyneham	06 246 9306	06 246 9699
Dr Simon Cox		Dept of Earth Sciences Monash University Clayton Vic 3168	03 565 5762	03 565 5062
Mr Terry Crabb		Dept of Mines & Energy 191 Greenhill Road Parkside S.A. 5063	08 274 7535	08 272 7597
Mr Mike Craig		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9453	
Mr John Creasey	Manager Image Processing	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9395	
Mr Eamonn Doyle	Sales Manager - SPANS	Informaster Australia 2/15-17 Forest Road Hurstville NSW 2220	02 587 1311	02 587 1449
Dr Morrie Duggan	Research Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9284	06 257 6465
Mr John Dunnell	Regional Manager	Dept of Mineral Resources 29-57 Christie Street St Leonards NSW 2065	02 910 8244	02 901 8246
Mr Andrew Eaton	Geoscience Analyst	Dept Manufacturing & Indust Development PO Box 173 East Melbourne Vic 3002	03 412 7000	03 412 7988

Participant	Position	Organisation	Telephone	Fax
Mr Neal Evans	Information Geologist	CRA Exploration Pty Ltd 139 Canberra Avenue Fyshwick ACT 2609	06 239 6316	06 295 2655
Mr Robin Fegeas	Chief Branch of Applications	USGS c/o AUSLIG PO Box 2 Belconnen ACT 2616		
Mr Jack Foley		AUSDEC 407 Riversdale Road East Hawthorn VIC 3123	03 882 3335	03 882 4581
Mr John Fryer	Chief Cartographer	Dept. Mineral Resources 29-57 Christie St St Leonards NSW 2065	02 901 888	02 091 8777
Ms Robyn Gallagher	Consultant	GISolutions 3 Hull Place Hackett ACT 2602	06 247 6994	06 247 6994
Dr George Gibbons	Assistant Director	NSW Geological Survey Mineral & Energy House PO Box 536 St Leonards NSW 2065		02 901 8468
Mr Ian Grierson	Managing Director	ENCOM Technology Pty Ltd PO Box 422 Milsons Point NSW 2061	02 957 4117	02 922 6141
Mr Bob Harris	Managing Consultant	EASAMS Unit 3, 2 Giffnock Avenue North Ryde NSW 2113	02 805 1066	02 887 3189
Mr Ingo Hartig	Technical Officer 4	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1204	
Mr Bob Harvey	Geophysicist	CRA Exploration Pty Ltd 139 Canberra Avenue Fyshwick ACT 2609	06 239 6316	06 295 2655
Major Mark Heinrich	SO2 Plans	Army Directorate of Survey-Army Campbell Park Offices CP2-4-24 Canberra 2600	06 265 3897	06 265 3601
Mr Steve Henshaw	Sales Rep	Intergraph Corp Unit C, 61 Talavera Road North Ryde NSW 2113	02 888 9900	02 887 2039
Mr Ross Hill	Technical Officer 4	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1226	06 247 2728
Mr John Hillier	Chief Cartographer	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1295	06 247 2728
Mr Graeme Holdaway	Chief Draftsman	Ampolex Limited Level 15, 580 George Street Sydney NSW 2000	02 364 4999	02 364 4930
Mr Gary Isbel	Senior Cartographer	Dept of Lands Qld PO Box 40 Woolloongabba QLD 4102	07 896 3250	07 896 3275
Mr Peter James	Manager Computer Services	Dept of Mines PO Box 56 Rosny Park Tasmania 7018	002 338 301	002 442 117
Dr Lynton Jaques	Chief of Program	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9745	06 257 6465
Mr Colin Johnson	Technical Officer 3	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9237	
Dr Dave Johnson		Etheridge & Henley Suite 5, 50 Geils Crt Deakin ACT 2600	06 285 2402	06 285 2586

Participant	Position	Organisation	Telephone	Fax
Mr Bob Kennard		AUSLIG PO BOX 2 BELCONNEN ACT 2616	06 201 4330	06 201 4366
Mr Dan Kennedy	ITO2	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1278	06 247 2728
Mr Paul Kelly	Deputy Surveyor General	Land Info Centre Dept Conservation & Management PO Box 143 Bathurst NSW 2795	063 328 414	063 322 320
Mr Jim Kinder	Manager Draughting Service	Dept of Manufacturing & Indust Development PO Box 173 East Melbourne VIC 3002	03 412 7855	03 412 7803
Ms Kathy Kirk	System Manager	CRA Exploration Pty Ltd 139 Canberra Avenue Fyshwick ACT 2609	06 239 6316	06 295 2655
Mrs Karen Lawson	A/Senicr Cartographer	Dept. Mineral Resources 29-57 Christie St St Leonards NSW 2065	02 901 888	02 901 8777
Colonel Simon Lemon	Director of Survey - Army	Dept of Defence (Army) Campbell Park Offices Canberra ACT 2600	06 266 3359	
Ms Sonja Lenz	Database Group	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9676	
Mr Ian McNaught	Project Scientist	NRIC GPO Box 858 Canberra 2601	06 272 3019	06 272 4687
Mr Steve Mann	Software Development Manage	ENCOM Technology Pty Ltd PO Box 422 Milsons Point NSW 2061	02 957 4117	02 922 6141
Mr Andy Marosszeky	Systems Analyst	GEOPLAN PO Box 166 St Ives NSW 2075	02 988 4669	02 449 6394
Mr Steve Massey	Geophysicist	Placer Exploration PO Box 764 Cloverdale Perth WA 6105	09 478 1199	09 478 3279
Mr Yanis Miezitis	SPOA	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9809	06 248 5619
Mr Joe Mifsud	Technical Officer 3	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1257	
Mr Bill Mitchell		Dept of Mines & Energy 191 Greenhill Road Parkside S.A. 5063	08 274 7535	08 272 7597
Mr Graeme Mitchell	Manager Information Sys	CRA Exploration Pty Ltd PO Box 175 Belmont WA 6104	09 277 8833	09 277 6272
Mr Kirk Mitchell	Cartographer	Stockdale Prospecting 60 Wilson St South Yarra Vic 3141	03 827 7522	03 826 0974
Mr Michael Morse	Project Manager Gravity	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9251	06 257 6041
Mr Ian Musto	GIS Manager	NRIC GPO Box 858 Canberra 2601	06 272 4688	06 272 4687
Mr Paul Nanninga	GIS Scientist	Murray-Darling Basin Comm GPO Box 409 Canberra ACT 2601	06 279 0124	06 248 8053

Participant	Position	Organisation	Telephone	Fax
Mr Ian O'Donnell	Resource Manager	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1363	06 247 2728
Dr Kerry O'Sullivan	Principal Geologist	CRA Exploration Pty Ltd 139 Canberra Avenue Fyshwick ACT 2609	06 239 6316	06 295 2655
Mr Brian Oversby	Senior Reseach Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9387	
Dr Colin Pain		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9469	
Ms Eva Papp		Dept of Earth Sciences Monash University Clayton Vic 3168	03 565 5762	03 565 5062
Dr John Parker		Dept of Mines & Energy 191 Greenhill Road Parkside S.A. 5063	08 274 7535	08 272 7597
Mr Chris Parvey	Project Leader Computing	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9775	06 257 6465
Mr Hamish Paterson	Computing Manager	Western Mining Corp PO Box 114 Daw Park SA 5041	08 372 0100	08 372 0130
Mr Rod Paterson	Chief Geologist(Computer Sc)	Aberfoyle Res Ltd 123 Camberwell Rd East Hawthorn VICTORIA	03 882 2226	03 813 1086
Dr Dave Patterson	Senior Research Petrologist	M.I.M. Holdings Ltd Mount Isa Mines Ltd Mount Isa QLD 4825	077 44 2931	077 43 9633
Dr Jim Peterson		Dept of Earth Sciences Monash University Clayton Vic 3168	03 565 5762	03 565 5062
Mr Eugene Petrie	Executive Officer	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9217	
Mr Peter Petkovic		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9278	06 257 6041
Mr Brian Pettigrove	Mapping Analyst	Dept Manufacturing & Indust Development PO Box 173 East Melbourne Vic 3002	03 412 7000	03 412 7988
Mr Peter Pieters		Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9310	06 257 6465
Mr Dennis Pogson	Manager Regional Mapping	NSW Geological Survey Mineral & Energy House PO Box 536 St Leonards NSW 2065	02 901 8341	02 901 8777
Mr Keith Porritt	Senior IT Officer (C)	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9554	06 257 6465
Ms Sandy Radke		Bureau of Mineral Resources PO Box 378 Canberra City 2601		
Dr Mark Rattenbury	Research Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9215	06 257 6465
Prof Colin Reeves	Head Geophysical Mapping	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9226	06 257 6041
Mr Dave Richards	Principal Geophysicist	CRA Exploration 139 Canberra Avenue Fyshwick ACT	06 239 6316	06 295 2655

Participant	Position	Organisation	Telephone	Fax
Mr John Richardson	Senior Geologist	M.I.M. Holdings Ltd Mount Isa Mines Ltd Mount Isa QLD 4825	077 44 2931	077 43 9633
Mr Greg Rogers	Geologist	CRA Exploration Pty Ltd 139 Canberra Avenue Fyshwick ACT 2609	06 239 6316	06 295 2655
Dr Rod Ryburn	Database Liaison	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9605	
Ms Cherryl Schmidt	Data Administrator	Dept of Resource Industries 61 Mary Street GPO Box 194 Brisbane QLD 4001	07 237 1437	07 221 9517
Mr Mike Sexton		Bureau of Mineral Resources PO Box 378 Canberra City 2601		
Mr Paul Shelley	Manager National Directory	NRIC GPO Box 858 Canberra 2601	06 272 4688	06 272 4687
Mr Col Simpson		Bureau of Mineral Resources PO Box 378 Canberra City 2601		
Mr Val Sobol	Survey Frafting Officer	Dept of Mineral Resources 29-57 Christie Street St Leonards NSW 2065	02 901 8245	02 901 8246
Mr Phil Soole	Experimental Scientist	CSIRO Geomechanics PO Box 63 St Lucia QLD 4067	07 377 7817	07 371 7435
Mr Howard Stagg	Principal Research Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9343	06 257 6041
Mr Barry Steele	Consultant Land Geog Info Sys	BHP Engineering PO Box 1794 Wollongong NSW 2500	042 280 411	042 280 893
Mr Garth Stewart	Geologist	RGC Exploration Pty Ltd PO Box 590 Fyshwick ACT 2609	06 280 6000	06 280 5915
Mr Greg Stewart	Programmer	Dept of Mineral Resources 29-57 Christie Street St Leonards NSW 2065	02 901 8681	02 901 8777
Mr Jon Stirzaker	TO 4 Cartography Group	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1265	06 247 2728
Dr Peter Stuart-Smith	Senior Reseach Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9293	
Ms Annette Sugden	Map Librarian	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 245 1258	06 247 2728
Mr Dave Suppel	Manager Metallogenic Mapping	NSW Geological Survey Mineral & Energy House PO Box 536 St Leonards NSW 2065	02 901 8345	02 901 8468
Dr Chris Tarlowski	Senior Reseach Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9265	06 257 6041
Mr Phil Tickle	Project Scientist	NRIC GPO Box 858 Canberra 2601	06 272 4688	06 272 4687
Mr Paul Treszise	Director Marketing	AUSLIG PO BOX 2 BELCONNEN ACT 2616	06 201 4330	06 201 4366

Participant	Position	Organisation	Telephone	Fax
Dr Liz Truswell	Senior Principal Res. Scientis	BMR GPO Box 378 Canberra City 2601	06 249 9427	
Mr Andrew Tucker	Software Consultant	BMR GPO Box 378 Canberra City 2601	06 249 9248	06 257 6120
Mr Vivek Tulpule'	.	ABARE Macarthur House Lyneham ACT 2602	06 246 9496	06 246 9699
Mr Andrew Valja	A/Assist Senior Cartographer	NSW Dept. Mineral Resources 29-57 Christie St St Leonards NSW 2065	02 901 888	02 901 8777
Mr David Walton	Senior IT Officer (C)	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9348	06 257 6376
Mr Jim West	A/Chief Cartographer	NSW Dept. Mineral Resources 29-57 Christie St St Leonards NSW 2065	02 901 888	02 901 8777
Mr Alan Whitaker	Geologist/Geophysicist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9702	
Mr John Wilford	Science II	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9455	
Dr Neil Williams	Associate Director	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9471	06 257 4614
Dr Peter Williams	Senior Reseach Scientist	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9389	06 257 6465
Mr Peter Williams	Senior Research Geophysicist	Western Mining 68 Daly St Belmont Western Australia	09 479 8400	09 479 8444
Mr Ian Withnall	Senior Geologist	Geological Survey of Qld GPO Box 194 Brisbane QLD 4001	07 237 1516	07 221 9517
Mr Geoff Wood	Data Manager	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 249 9398	06 257 6376
Mr Ray Wood	Marine Geophysics	DSIR Geo PO Box 1320 Wellington New Zealand	0011 64 473808	0011 64 4710977
Miss Gayle Young	Technical Officer 4	Bureau of Mineral Resources PO Box 378 Canberra City 2601	06 272 5007	06 247 2728
Mr Graham Young		Dept of Mines & Energy 191 Greenhill Road Parkside S.A. 5063	08 274 7535	08 272 7597