GIS Methods for Integrating Exploration Data Sets

Bonham-Carter, G.F.[1]

1. Geological Survey of Canada, Ottawa, Ontario, Canada

ABSTRACT

The decade 1987–1997 has seen the introduction of commercial GIS to the geosciences, and its application to mineral exploration. GIS provides a computing environment for handling images, maps and data tables, with tools for data transformation, visualization, analysis, modelling and spatial decision support. Methods of integrating exploration data sets for mineral potential mapping are facilitated by GIS, and can be either knowledge-driven or expert-system–driven, depending on the level of prior exploration. The experience gained with these methods to date demonstrates that they are invaluable for formalizing exploration models, for providing a basis of communication between individuals with differing backgrounds and perspectives, for quantifying spatial associations of mineral occurrences with data layers, and for identifying prospective areas for follow-up.

INTRODUCTION

The progress in the handling of exploration data sets in the decade since the last of these conferences (Exploration '87–Exploration '97) has been greatly influenced by the development of geographic information systems (GIS). Commercial GIS appeared on the market in the mid-1980s. By the end of the 1980s, systems capable of handling vector, raster and tabular data were available on UNIX workstations and PCs. Although most of the algorithms used in modern GIS were developed and available in the 1960s and 1970s, general-purpose systems for handling a mixture of geocoded images, digitized maps and associated databases were not practical, because of limitations in computer speed, memory and graphical capability. Hardware performance reached a stage in the 1980s that catalysed the development of commercial GIS, although 'proto-GIS' had been around since the late 1960s in government departments and specialized computing laboratories, for example the Canada Lands Data System in Environment Canada (Dupre, 1981), often quoted as being the first GIS (e.g., Tomlinson, 1984). The huge commercial market for digital mapping systems (in natural resources, municipal, military and business applications in particular) ensured the investment in and growth of GIS. The exploration industry, which represents a small but important segment of the GIS market, has benefited from the investment in GIS by other sectors of the economy. The 3-D GIS market is much smaller and more specialised than the 2-D market, resulting in the high cost of commercial 3-D systems, both now and probably in the forseeable future.

The traditional method of handling exploration data sets was the preparation by hand of maps; the map cabinet was equivalent to today's computer storage; the light table was today's computer monitor; and coloured pencils were today's enhancement and visualization tools. Computer applications are certainly not new to the exploration business, and geophysical (and to some extent geochemical) data processing were important mainstays well before 1987. Furthermore, the development of image processing systems for satellite data, and CAD systems for simple map handling as line drawings, occurred in the 1970s. Lacking was the ability to handle different types of spatial data on a unified computing platform, with modular tools for tasks such as changing geographic projection, visualization and analysis. Neither image processing nor CAD systems provided the database linkages so important to modern GIS. GIS today have the tools for creating multi-layer spatial databases, using a variety of data structures, and ensuring the geographic registration between layers, critical for any multi-layer analysis (Bonham-Carter, 1994).

Besides seeing the rise of commercial GIS, this past decade has also seen significant growth of exploration data sets in digital form. This has happened mainly as a consequence of the catalysing effect of computer developments and other aspects of instrument design. Some examples are: satellite sensors able to generate multispectral images with increasingly fine resolution (e.g., Kruse *et al.*, 1993); geochemical methods capable of analysing elements in various chemical phases (e.g., Hall *et al.*, 1996a,b); global positioning systems (GPS) that can provide superior accuracy of spatial positioning (e.g., Leick, 1990); field-portable computers suitable for geological mapping in the field (e.g., Brodaric and Fyon, 1988); and many others. Government agencies now generate large amounts of digital map data (e.g., Broome *et al.*, 1993; USGS, 1996), and the ability to move digital data sets around rapidly has been enormously enhanced by the Internet (e.g., Thoen, 1995). Much of the real progress in the past decade has been in the ability to handle explo-

In "Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration" *edited by* A.G. Gubins, 1997, p. 59–64

ration data sets digitally, particularly in capturing, rectifying and visualizing spatial data, although there have also been developments in data analysis and modelling.

This paper reviews and discusses some current issues in GIS, with some guesses about future developments.

DIGITAL CAPTURE AND REPRESENTATION OF GEOLOGICAL MAPS

The geological map has for over a century been the principal medium for information management in geology. As we convert from a traditional analogue document to a spatial database, the capture, editing, compilation, manipulation and analysis of geological information is profoundly altered. The most immediate effects are in field operations, the uses to which map data are put, and in the production of printed maps. The field notebook is now a computer. Field location is determined by GPS. The database forms a source of information that far exceeds what can be shown on a single map (Brodaric and Harrap, 1997). Data from the database can be updated, revised, searched and visualized in numerous ways: data sets can be analysed in combination with other data sources, they can be archived and disseminated electronically, and they can be converted into printed customized geological maps at a fraction of the cost of traditional map making. John Broome (pers. comm., 1997) estimates that the Geological Survey of Canada, by applying GIS methodology to field operations, database construction and map production, has been able to double map output since 1991, and reduce turnaround from two years to less than three months, in spite of a 20% reduction in cartographic staff. By moving from a static analogue document to a digital database, the possibility of combining the map data with geophysical, geochemical and mineral deposit information by digital methods is created.

However, as an exploration data layer, the digital geological map is probably the most difficult data source to deal with in an integrated study, even when it has been captured in a computer-accessible form. This is because a data model has not been fully developed to permit the storage and manipulation of the information on the map. The data model allows the objects on the maps (map units, faults, mineral occurrences, structural observations, etc.) to be described or structured in the database. Simply replicating the legend as symbols and text greatly restricts the use of the data. At one extreme, a scanned map is a totally unstructured raster image—we may be able to stretch it and change the projection, but we are unable to do more than reproduce the original lines and symbols (although this is still sometimes worthwhile because of the costs of digital structuring). At an intermediate level of structure, the lines and points on the map can be digitized as 'spaghetti' in a vector format, possibly with some tagging of objects with descriptive attributes; this allows some manipulation and search, but still limits analysis. A full structuring of the objects on the map involves the use of both topological (adjacency, containment) attributes, and the linking of objects to a relational database capable of representing lithology, age, structure and other attributes that appear in the traditional legend, field notebook and map report.

At present, geological maps that are held digitally are generally poorly structured, because an acceptable data model has not been developed. The uses to which the digital geology can be put are, therefore, restricted to (1) map production (2) map generalization by reclassification (recoding of units) and (3) search for spatial objects having particular characteristics. Combining two maps from different scales, joining adjacent map sheets, changing the scale (except as zoom in and zoom out) and more sophisticated searches and generalization are simply not possible without manual intervention, except at the simplest level.

There are several groups of geologists currently addressing this problem in different countries (e.g., USGS, 1996). The difficulty is not so much one of computer technology: there is plenty of database expertise that can be applied. The problem is one of getting agreement among geologists about how data can be structured, and applying some standards that will be widely accepted. This blueprint of how map data should be organised must be flexible to allow customization to particular needs, and to allow evolution as initial deficiencies are revealed. When a satisfactory data model is in place, new geological mapping can adhere to the standards, but digitizing existing maps, particularly maps compiled at small scales, will only produce part of the data needed. It will be many years before the effects of a satisfactory data model will be widely felt and appreciated. The impact of the introduction of a well-designed data model for geological maps will be profound for exploration, and for other users of digital geology. Map compilation and generalization will be easier, and extraction of map features pertinent for mineral potential mapping will be possible without necessarily consulting reports or talking to those who have mapped in the area of interest (although this is always desirable).

GEOCHEMICAL DATA

Geochemical data form a significant component in many spatial databases for exploration. Unlike geophysical data, there is little in the way of geochemical theory in mathematical terms to help interpret geochemical data sets. The main problems of dealing with geochemical data in a GIS are (1) whether to convert point observations to surface representations, (2) whether to reduce the dimensionality of the data using multivariate statistical techniques, (3) how to decide on class intervals for symbols, colours or contours (and in the situation of two classes, how to separate anomaly from background).

For the first problem, most GIS provide surface modelling tools for spatial interpolation (although the choice of method is almost always contentious), and the use of catchment basins for displaying and modelling lake and stream samples can be an effective alternative for detailed surveys. The second problem (dimensionality) is sometimes solved using principal components analysis, cluster analysis, or regression techniques, but outside the GIS in a statistical package. The third problem (classification intervals) is usually solved arbitrarily by using percentiles, but may be analysed independently of the GIS with probability plots or other tools. The log-log plots of concentration versus area proposed by Cheng *et al.* (1994), or the contrast method in weights of evidence (e.g., Agterberg *et al.*, 1990; Bonham-Carter, 1994) are new methods for helping to distinguish anomaly from background.

New innovative approaches are needed to help cope with data overload. The application of selective extraction methods compounds the problem of multiple elements and multiple media, and useful information is undoubtedly lost by the inability to cope with too many data layers.

GEOPHYSICAL DATA

Geophysical data processing is not the focus of this paper. However, thematic layers of geophysical data, already processed to produce maps suitable for exploring for particular deposit types, form part of the spatial database used in mineral potential mapping. One of the key developments over the past decade has been the development of forward modelling in three dimensions to aid in the interpretation of regional geophysical data (e.g., Jessell and Valenta, 1996). As discussed in a poster by Jessell, and in a summary in this volume, the approach is to build up a series of geological events (sedimentary layers, folding, faulting, intrusion, etc.) to generate a 3-D model, and having assigned physical properties to the materials in the model, to generate maps showing predicted geophysical response. This approach is not strictly 'GIS', but it clearly has great potential for linking to GIS for integrated geological/geophysical interpretation. If this method can somehow be constrained by actual geological maps from GIS, the interpretation of geophysical images, and the postulation of subsurface geology consistent with observed geophysical patterns, could be greatly enhanced.

MINERAL DEPOSIT DATA

One of the less glamorous but important tasks of data management is the construction and maintenance of first-class mineral deposit databases. Without these data, mineral potential mapping using GIS is constrained to the expert-system approach, discussed later. Although there are now a number of valuable databases, constructed according to relational database principles, produced by government surveys and within individual companies, data quality remains a difficulty in putting the information to use in a GIS. In a recent study of the mineral potential of the Slave Province in Canada, it was determined that the VMS occurrences in the database did not always occur on the volcanic rocks, when they were superimposed digitally on a digital map, although most of the 'delinquent' points were within 1 km of the volcanic contact. Further investigation showed that some of this error was in the spatial locations of the mineral occurrences, and some was in the position of the geological contact. In other instances, occurrences were misclassified according to deposit type, or were missing, or were repeated. The demands of the GIS environment throw data quality into a much more critical light, because of the ability to change scale, and to examine spatial associations between occurrences and any one of a variety of other data layers. Improved mineral deposit databases are essential to the success of GIS mineral potential studies.

MINERAL POTENTIAL MAPPING

Mineral potential mapping with digital data sets was practised in research organisations prior to the arrival of commercial GIS. At the Geological Survey of Canada, for example, a system called SIMSAG, which allowed multivariate statistical analysis of multiple layers of gridded data was in use in the 1970s and early 1980s (Chung, 1983). This system incorporated logistic and Poisson regression methods for predicting mineral potential, based on the use of a well-explored training set of grid cells. Somewhat analogous systems and methods were in use in several other institutions around the world. Developed on contract to the U.S. Geological Survey, the Prospector expert system was applied to regional data sets to predict mineral potential, e.g., Campbell *et al.* (1982). The drawback of these systems was that they could not handle vector data, had only a limited database capability, and were restricted to small grids. When commercial GIS arrived in the mid- to late 1980s, the quantitative approach to mineral potential mapping became easier, because many of the laborious data preparation and handling aspects were simplified.

Table 1: ***Data-driven and knowledge-driven methods of mineral potential mapping.***

Type	Model parameters	Example
Data-driven	Calculated from training data	Logistic regression
		Weights of evidence
		Neural networks
Knowledge-driven	Estimated by an expert	Bayesian updating
		Fuzzy logic
		Dempster Shafer belief

In the year of the last Exploration conference, Bonham-Carter *et al.* (1987) reported on a mineral potential study of gold in the Meguma terrane of Nova Scotia. Since that time there has been a steady growth of GIS applications in all branches of geoscience, and the use of GIS to integrate exploration data sets has become common practice.

Undoubtedly, the most popular approach of examining the spatial relationships between regional data sets is to use visualization techniques. The ability to zoom, pan, query, apply enhancements, superimpose points and lines on colour displays, do hillshading and 'drapes', encourages almost unlimited exploratory spatial data analysis. Beyond this visualization phase, however, there is a need to apply exploration models in an environment that allows statistical analysis, and permits integration scenarios to be developed using expert knowledge about the deposit type under consideration. This is not the preserve of the GIS specialist alone, but requires a close collaboration between GIS staff and the exploration geologist. The ultimate goal should be to put a user-friendly GIS in the hands of the exploration geologist alone, and let the GIS expert provide the spatial database to work on.

Quantitative approaches of mineral potential mapping are essentially (1) selecting and enhancing those data layers that are suitable as predictors of the mineral deposit type being considered, and (2) combining these layers using operations and weights that differ according to the method used. Underlying both these steps is the concept of an exploration model—usually taking the form of a series of statements about the characteristics of a typical deposit that can be determined from spatial data sets. The combination and weighting of selected layers can be subdivided into two types: a 'data-driven' statistical approach, and a 'knowledge-driven' expert approach, as shown in Table 1.

The data-driven approach requires that quite extensive exploration has taken place in the region of interest, so that an adequate sample of mineral occurrences of the right type are known. The spatial association between the known occurrences and the predictive data sets is used to determine weights, and the weights are then applied to predict areas with similar characteristics to the known occurrences. Neural networks are the most recent addition to the stable of data-driven methods.

The knowledge-driven approach can be quite simple, using subjective weights and an additive model, or can employ more sophisticated knowledge-representation tools such as fuzzy membership functions (e.g., An *et al.*, 1991; or Dempster Shafer belief functions (e.g., An *et al.*, 1992; Chung and Fabbri, 1993; Wright and Bonham-Carter, 1996). The thought processes and logic embodied in the exploration model are simulated with an inference network that may allow a variety of combination rules to be applied depending on the situation (e.g., Wright and Bonham-Carter, 1996).

EXAMPLE APPLICATION

As an example of this type of modelling, the results of a study of VMS potential in the Snow Lake area of Manitoba are briefly illustrated. This work was published by Wright and Bonham-Carter (1996). Both data-driven and knowledge-driven methods were applied, although the figures here refer to the use of fuzzy logic to represent the knowledge of the exploration expert. Figure 1 shows the inference network that was applied to define the sequence of operations involved in combining 23 input maps, or data layers. Each map was linked to a table containing fuzzy membership values whose magnitudes reflect the opinion of the

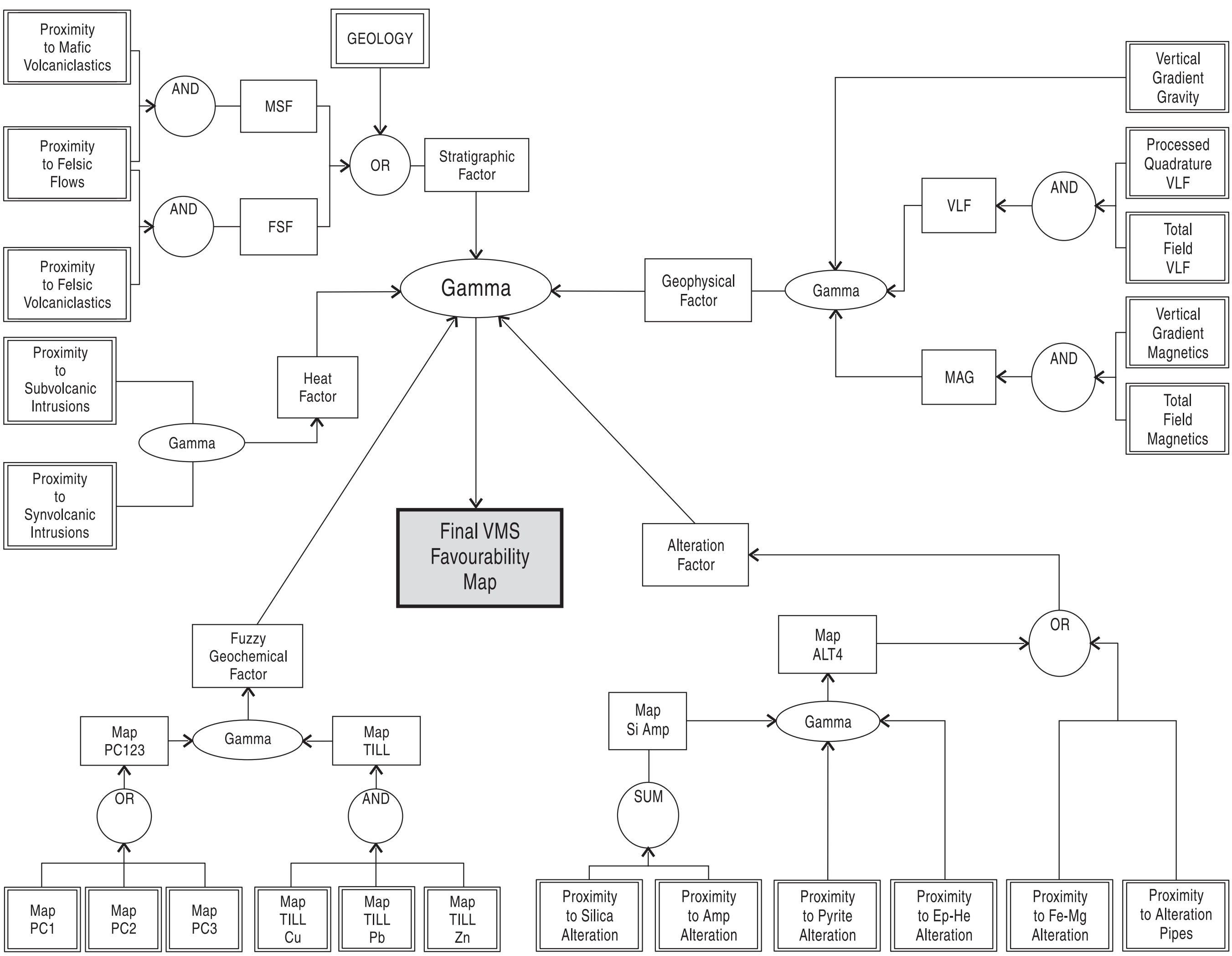

Figure 1: *Inference network for VMS study in Snow Lake greenstone belt (Manitoba). Rectangular boxes represent maps (double line surround implies input map), whereas circular or elliptical shapes represent fuzzy operators. The final product is a favourability map showing relative potential for VMS deposits (adapted from Wright and Bonham-Carter, 1996).*

exploration geologist about each map unit or class. Figure 2 shows the potential map, after the combination steps, and the application of a colour scheme (CD-ROM version) to denote relative favourability. The distribution of the known deposits is also shown. The deposit marked as Photo Lake was discovered by Hudson Bay Mining and Smelting after this study was complete (but before it was published). Clearly there are a number of areas that have a high potential on the map, with characteristics similar to those of the known deposits, and Photo Lake is in one of them. Photo Lake was also within an area predicted by the weights of evidence method.

Some of the advantages of automating the integration process are:

1. It provides an audit trail showing the process of ranking areas on the map. The same mineral potential map can therefore be reproduced, using the same processing steps.
2. It guarantees that an exploration model is formulated in concrete terms, giving the basis for communication of ideas amongst the exploration team.
3. It provides a tool for developing a variety of scenarios, by allowing alternative exploration models to be evaluated with the same data. This experimental aspect allows for a sensitivity analysis to be carried out.
4. Spatial associations previously unknown can be revealed and tested.

There are also some disadvantages:

1. The process could become a black box, so that the exploration geologist fails to understand the modelling process, relying on a GIS person to carry out the analysis instead of getting a hands-on feel for the data.
2. Although some kinds of error analysis are possible, the propagation and effect of errors in the data are not well understood, except in broad terms.
3. Some factors used in exploration models may be difficult or impossible to incorporate in the computer model.
4. Data may be too scattered or patchy to allow complete data layers to be constructed, because there is too much missing information. In under-explored areas, this is probably the greatest problem.

A disadvantage sometimes cited is that the modelling process is too slow and laborious to make this a feasible methodology in the fast-paced world of exploration, where decisions are needed in a short time frame. This may be true if the time taken to create a digital database is included. However, if the database has already been constructed, 95% of the work has already been done. The actual integration modelling can be carried out quickly by an experienced group. On a recent visit to CPRM in Brazil, a series of exploration data sets was studied for an area west of Sao Paulo where a large number of small gold prospects occur, an exploration model was formulated in discussion with the local geologists,

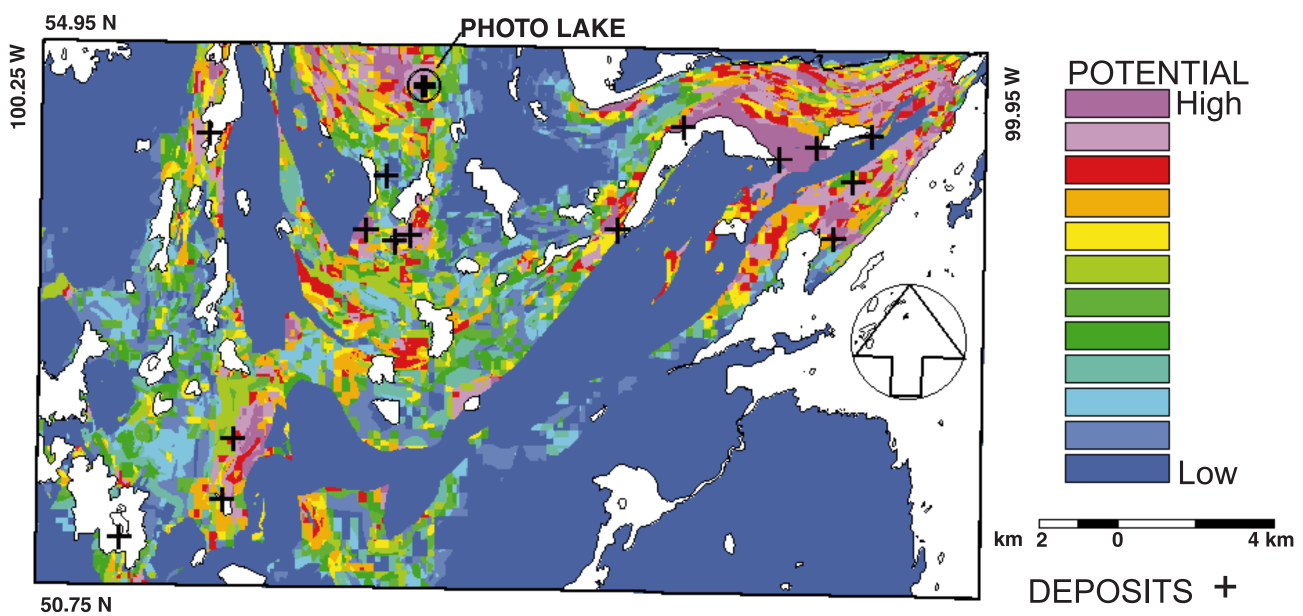

Figure 2: *Favourability map showing potential for VMS deposits (Chisel Lake-Anderson Lake area). Known deposits are shown, as well as the location of a new deposit (Photo Lake) discovered after this study was done.*

spatial associations were evaluated between the known occurrences and various predictive patterns, and a mineral potential map was generated by weights of evidence in less than a day. There are an increasing number of software packages suitable for doing this type of work.

DISSEMINATION OF EXPLORATION DATA SETS

Communication of spatial information has greatly improved over the past decade, and there is no doubt that our ability to move exploration data sets on the World Wide Web around the world, or from one office to another in the same building, is going to improve dramatically over the next few years. Some of the possibilities and problems of this new world of the Internet are discussed by Cox (this volume).

CONCLUSIONS

Information technology in general is making a profound impact on mineral exploration, and GIS in particular have become essential tools to manipulating exploration data sets.

Statistical and expert-system methods can provide useful tools for integrating exploration data sets. Some limited success has been experienced in applying these methods. The methods are valuable for discovering and quantifying new spatial associations between deposits and geological, geochemical and geophysical predictors, and for providing rational support for exploration decisions. The methodology also has value for giving a framework to the organisation and analysis of exploration data, for prioritizing exploration prospects, and providing a mechanism for exchanging ideas about exploration models.

New developments can be expected in several areas: data models for structuring geological map data; 3-D systems capable of linking surface geology to plausible 3-D reconstructions, consistent with geophysical data; improved exploration models, given the accumulated experience gained from practice of the integration methodology; improved methods for dealing with uncertainty and the propagation of errors.

GIS technology is valuable to all stages of the mining cycle: exploration, property evaluation, environmental impact, mine design and operation, and landscape restoration. Spatial data assembled in GIS have application for more than exploration alone. The integration of data sets for exploration decisions need not be a time-consuming and expensive process.

REFERENCES

Agterberg, F.P., Bonham-Carter, G.F., and Wright, D.F., 1990, Statistical pattern integration for mineral exploration: *in* Computer Applications in Resource Estimation Prediction and Assessment for Metals and Petroleum: Gaál, G. and Merriam, D.F., eds., Pergamon, Oxford-New York, p. 1-21.

An, P., Moon, W.M., and Bonham-Carter, G.F., 1992, On a knowledge-based approach of integrating remote sensing, geophysical and geological information: Proc. IGARSS '92, p.34-38.

An, P., Moon, W.M., and Rencz, A., 1991, Application of fuzzy set theory for integration of geological, geophysical and remote sensing data: Canadian Journal Exploration Geophysics, **27**, p. 1-11.

Bonham-Carter, G.F., 1994, Geographic information systems for geoscientists: modelling with GIS: Pergamon, Oxford, 398 p.

Bonham-Carter, G.F., Agterberg, F.P., and Wright, D.F., 1987, Integration of geological datasets for gold exploration in Nova Scotia: Photogrammetry and Remote Sensing, **54**, no. 11, p. 1585-1592.

Brodaric, B., and Fyon, J.A., 1988, OGS Fieldlog: a microcomputer-based methodology to store, process and display map-related data: Ontario Geological Survey, Open File Report 570, 73 p.

Brodaric, B., and Harrap, R., 1997, Fieldlog v3.0: Users Guide and Tutorial: Geological Survey of Canada, Open File, in preparation.

Broome, H.J., Brodaric, B., Viljoen, D., and Baril, D., 1993, The NATMAP digital geoscience data-management system: Computers & Geosciences, **19**, no. 10, p. 1501-1516.

Campbell, A.N., Hollister, V.F., and Duda, R.O., 1982, Recognition of a hidden mineral deposit by an artificial intelligence program: Science, **217**, no. 3, p. 927-929.

Cheng, Q., Agterberg, F.P., and Ballantyne, S.B., 1994. The separation of geochemical anomalies from background by fractal methods. J. Geochem. Explor., **51**, p. 109-130.

Chung, C.F., 1983, SIMSAG; integrated computer system for use in evaluation of mineral and energy resources: Mathematical Geology, **15**, no. 1, p. 47-58.

Chung, C.F., and Fabbri, A.G., 1993, The representation of geoscience information for data integration: Nonrenewable Resources, **2**, no. 2, p. 122-139.

Cox, S., this volume, Delivering exploration information on-line using the WWW: Challenges, and an Australian experience.

Dupre, N., 1981, Introduction to the Canada Lands Data System: Canada Lands Data System, Environment Canada, Report No. R001150, unpaginated.

Hall, G.E.M., Vaive, J.E., Beer, R., and Hoashi, M., 1996a, Selective leaches revisited, with emphasis on the amorphous Fe oxyhydroxide phase extraction: J. Geochem. Explor., **56**, p. 59-78.

Hall, G.E.M., Vaive, J.E., Beer, R., and Hoashi, M., 1996b, Phase selective leaches for use in exploration geochemistry, *in* Bonham-Carter, G.F., Galley, A.G., and Hall, G.E.M., eds., EXTECH I: a multidisciplinary approach to massive sulphide research in the Rusty Lake-Snow Lake greenstone belts, Manitoba: Geological Survey of Canada, Bulletin 426, p. 169-200.

Jessell, M.W., and Valenta, R.K., 1996, Structural geophysics: integrated stuctural and geophysical mapping: *in* Structural Geology and Personal Computers, DePaor, D., ed., Pergamon, Oxford.

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993, The Spectral Image Processing System (SIPS); interactive visualization and analysis of imaging spectrometer data: Remote Sensing of Environment, **44**, no. 2-3, p. 145-163.

Leick, A., 1990, GPS satellite surveying: John Wiley & Sons, New York, 352 p.

Raines, G., 1996,

Thoen, W., 1995, Internet resources for the geosciences, with an emphasis on GIS and mapping: Computers & Geosciences, **21**, no. 6, p. 779-786.

Tomlinson, R.F., 1984, Geographic information systems—a new frontier: Operational Geographer, **5**, p. 31-35.

USGS, 1996, National geologic map database:
URL: http://wwwflag.wr.usgs.gov/ngmdb/ngmdb_home.html

Wright, D.F., and Bonham-Carter, G.F., 1996, VHMS favourability mapping with GIS-based integration models, Chisel Lake-Anderson Lake area, *in* Bonham-Carter, G.F., Galley, A.G., and Hall, G.E.M., eds., EXTECH I: a multidisciplinary approach to massive sulphide research in the Rusty Lake-Snow Lake greenstone belts, Manitoba: Geological Survey of Canada, Bulletin 426, p. 339-376, 387-401.