NOTE

In common with general usage, I have employed the acronym GIS to refer to either a single geographic information system, or to several systems, or to the field of geographic information systems as a whole. For the plural case, I have not added an "s" or "es" to the basic three letters. For example, the acronym is used in the following ways:

* "This GIS has a powerful analytical module..." (used as a singular noun), or

* "These GIS operate on personal computers..." (used as a noun in the plural), or

* "GIS is a rapidly growing field..." (used as a singular noun referring to the whole technology), or

* "The GIS approach to the solution..." (used as an adjective referring to the whole technology).

CHAPTER 1

Introduction to GIS

A minority of geologists and other geoscientists have been using computers for manipulation of spatial data since the 1960s. During the 1980s, advances in computer hardware, particularly processing speed and data storage, catalyzed the development of software for handling spatial data. The emerging capabilities for graphical display played an important role in this development. One of the significant products of this period of rapid technological change was GIS. The impact of GIS has been widely felt in all fields that use geographic information, in resource management, land-use planning, transportation, marketing, and in many applications in the geosciences and elsewhere. The majority of large geological organizations are now using GIS, and the handling of spatial data of all types by computer is widespread.This chapter introduces GIS by describing what GIS means, the purposes and functions of GIS, how GIS relates to other kinds of software for spatial data handling, and presents a typical geological application.

WHAT IS GIS?

A geographic information system, or simply GIS, is a computer system for managing spatial data. The word geographic implies that locations of the data items are known, or can be calculated, in terms of geographic coordinates (latitude, longitude). Most GIS are restricted to data in two spatial dimensions, although some systems of particular interest to geologists have true three-dimensional capabilities and can represent objects such as recumbent folds. The word **information** implies that the data in a GIS are organized to yield useful knowledge, often as coloured maps and images, but also as statistical graphics, tables, and various on-screen responses to interactive queries. The word **system** implies that a GIS is made up from several interrelated and linked components with different functions. Thus, GIS have functional capabilities for data capture, input, manipulation, transformation, visualization, combination, query, analysis, modelling and output. A GIS consists of a package of computer programs with a user interface that provides access to particular functions. The user may control GIS operations with a graphical user interface, commonly called a **GUI**, or by means of a **command language**, consisting of program statements that dictate the sequence and type of operations.

GIS are computer tools for manipulating maps, digital images and tables of **geocoded** (geographically located) data items, such as the results of a geochemical survey. GIS are designed to bring together spatial data from diverse sources into a unified database, often employing a variety of digital data structures, and representing spatially varying phenomena as a series of data layers (such as bedrock geology, depth to water table, Bouguer gravity anomaly, and so on), all of which are in spatial **register,** meaning that they overlap correctly at all locations.

GIS is filling a very real need in the face of the rapid growth of digital spatial data in the geosciences. Many spatial datasets are now being generated by government agencies, private companies and university researchers, and they would be ineffectively used and result in wasted resources without good systems of data management. Satellite images are a prime example of this data explosion. Without digital systems for the processing and display of images, the enormous volumes of remote sensing data collected daily would simply remain on computer storage devices, and the wealth of information they contain would remain unrevealed and indigestible. Data collected from geophysical instruments, airborne, ship-borne, ground-based, down-hole and others, also make up enormous volumes of numbers that reveal little until they are properly organized and displayed. Similarly, geochemical surveys of rocks, soils, water, sediments and plants, often with thirty or more elements determined from each sample, yield huge amounts of spatial data whose information content cannot be assessed without efficient spatial data systems.

GIS has made a tremendous impact in many fields of application, because it allows the manipulation and analysis of individual "layers" of spatial data, and it provides tools for analyzing and modelling the interrelationships between layers. Geoscientists need to understand the spatial relationship between all the various kinds of spatial data that they collect. For example, mineral exploration requires the simultaneous consideration of many kinds of spatial evidence for mineral deposits, such as the geology, structure, geochemical and geophysical characteristics of a region, as well as the locations and type of past mineral discoveries. In environmental problems, the interaction of many processes must be considered, and the simultaneous analysis of multiple datasets is imperative.

Before widely-available commercial GIS appeared on the scene in the late 1980s, most geoscientists working with multiple spatial datasets were doing their work on light tables. A relatively small proportion of the geoscience community was computer-orientated, working mostly with mainframe computers and locally-developed software. GIS technology is still in its infancy (1993), with the majority of applications being carried out by specialists. However, GIS has the potential to change the geological workplace drastically. In a few short years, the personal computer has virtually eliminated the typewriter and the hand calculator from the office, and GIS is likely to replace the light table and the map cabinet. Like many fields where computers are employed, GIS has the potential to free the user from the technical, slow, laborious problems of data handling, and to enhance the capability for creative data analysis and interpretation.

The acronym "GIS" has come to mean much more than a type of computer program. GIS implies the science of geographic information management and analysis. In many countries there are now academic institutes devoted to all aspects of GIS. In the United States, the National Center for Geographic Information and Analysis (NCGIA) is headquartered at the University of California, Santa Barbara. In the United Kingdom, a group of regional laboratories has been established for GIS research. GIS courses are now part of the academic curriculum in many universities and technical colleges. There are a large number of GIS periodicals, many of them commercially-orientated, but some entirely for the publication of scientific papers, such as the International Journal of Geographical Information Systems. Numerous conferences are devoted to GIS, catering to the needs of those working in GIS in academia, government and business. GIS is enjoying a period of popularity because it is new and topical. Over time, GIS may simply be

absorbed into the various disciplines that it affects, like some other computer applications. It seems more likely, however, that the discipline of GIS will become permanently established, with its own agenda of research, to meet the needs of the many fields that deal with spatial data.

PURPOSE OF GIS

The ultimate purpose of GIS is to provide support for making decisions based on spatial data, as illustrated by a few geological examples. The exploration manager may use GIS to assemble data in the form of a mineral potential map to decide priorities for future exploration; the mining geologist may evaluate the effects of acid mine drainage with GIS to decide the kinds of remediation that would be cost-effective; the engineering geologist may evaluate slope stability conditions with GIS to decide the best route for a new road. Sometimes the purpose of using a GIS is to support general research. For example, a geochemist may use GIS to investigate the spatial association between the distribution of selenium in surface water, and the distributions of local rock type, pH of the water and the local vegetation; or the geophysicist may employ GIS to study the spatial factors related to earthquakes. Of course, GIS is invaluable for collecting, maintaining and using spatial data in a database management role, as well as for producing both standardized and customized cartographic products. Application of GIS achieves these major goals through one or more of the following activities with spatial data: organization, visualization, query, combination, analysis and prediction.

Organization

Anyone who has collected a large mass of data for a particular purpose knows the importance of data organization. Data can be arranged in many different ways, and unless the organization scheme is suitable for the application at hand, useful information cannot be easily extracted. Schemes for organizing data are sometimes called **data models**, as discussed in Chapter 2. The principal characteristic for organizing GIS data is spatial location. A table of geochemical data may be interesting for analyzing the relationships between geochemical elements on scatterplots, but without knowing the *locations* of samples, the interpretation of spatial patterns and relationships with other spatial data, such as rock type, cannot be established. GIS data are also organized according to non-spatial characteristics. For example, the interpretation of geochemical data may depend on recognizing spatial patterns of element ratios, or of groups of observations based on the method of analysis or the year of collection. Data models must, therefore, organize observations both by spatial and non-spatial attributes. The efficiency and type of data organization effects all the other five activities and is therefore of fundamental importance.

Visualization

The graphical capabilities of computers are exploited by GIS for visualization. Visual display is normally carried out using the video monitor, but other output devices such as colour printers are used for hardcopy displays. Humans have an extraordinary ability to understand complex spatial relationships visually, whereas the same information may be quite unintelligible when presented as a table of numbers. For example, given a table of geochemical analyses, a geologist is normally unable to recognize the spatial distribution of highs and lows in the data, but when the same table is converted to an effective map display, spatial patterns are immediately revealed. Visualization is achieved in GIS with colour and symbols, and by specialized methods using perspective, shadowing and other means.

Spatial Query

Visualization reveals spatial pattern amongst collections of organized data items. However, visualization is not so helpful for answering questions about special instances in the data, such as the value of particular data items. Spatial query is a complementary activity to data visualization. For example, a particular display of a combination of mineral deposit points and a geochemical map might suggest the existence of a spatial relationship in some areas of the display, but not in other areas. Spatial query permits the user to find out the special circumstances of each case, by finding out the name and other details of individual mineral occurrences, and the characteristics of individual geochemical samples in the selected neighbourhoods of interest. This often helps to find the reason behind the spatial pattern. GIS provides tools for two types of interactive query. The first type is the question "What are the characteristics of this location?" The second type is "Whereabouts do these characteristics occur?"

For example, suppose an aeromagnetic map is displayed on the video monitor, one might wish to know in detail the rock formation, the distance to the nearest road, the topographic elevation, the value of the Bouguer anomaly and the location and analysis of the nearest geochemical sample for any specified location. Many GIS allow the user to generate a summary table of selected characteristics (appearing in a display window) pertaining to a specific location. The location is often identified interactively with the cursor, and the table is instantly updated as the cursor is moved to each new location. Clearly this requires that data is efficiently organized by spatial location to permit fast retrieval.

As an example of the second type of question, one might wish to know all the locations on the map where a particular set of conditions is satisfied. In a simple case, it might be useful to know all the locations where arsenic in soil is greater than 250 parts per million (ppm). Many GIS use a query language for constructing specific questions. The user types in one or more statements, causing a search, from which the result may be a table or a map with the identified locations. The query might be more complex, asking for all locations within 250 m of a lake, at elevations greater than 1000 m above sea level, underlain by Carboniferous limestone. Efficient searches of this type require organization of data both by spatial and non-spatial attributes.

Other kinds of questions may be related to distance, orientation, and conditions of adjacency or containment like "Find all instances of granite in contact with a limestone?" Such questions require not only efficient search of data items, but also the capability for deriving their geometric and topological attributes. The term topology refers those

characteristics of a data object, like adjacency and containment, that are not affected by spatial transformations. Thus Africa is *adjacent to* the Mediterranean Sea, no matter what coordinate projection is used to represent the boundaries of these spatial objects.

Combination

The ability to merge spatial datasets from quite different sources and display and manipulate combinations can often lead to an understanding and interpretation of spatial phenomena that are simply not apparent when individual spatial data types are considered in isolation. For example, by superimposing a digitized geological map on to a satellite image, it may become clear that a particular lithology has a distinctive texture or spectral response on the image. The process of combining layers of spatial data is sometimes called data integration, and can be carried out either by visualizing composite displays of various kinds, or with integration models that effectively create a new map from two or more existing maps. Integration models are symbolic mathematical models, using arithmetic and logical operations to combine layers of data together. A simple example is the combination of a map of lakes, derived by digitizing a topographic base map showing drainage features, with a digital map of radiometric data, derived from an airborne survey. The radiometric map contains, say, 100 colours, each colour representing the intensity of the radioelement measurement, whereas the lake map contains only 2 colours, lake present and absent. The algebraic statement that combines these two maps might indicate that the new map be "equal to the class (in this case colour) of the radiometric map except where lakes are present, class 0 otherwise". Such statements are usually written in a programming language, sometimes known as "**map algebra**", specific to the GIS. The combined map can now be treated as a single map, revealing the spatial relationship of water bodies to radiometric patterns.

One of the really powerful features of GIS is the ability to link several map algebra statements together to form more complex algorithms. Several maps and tables of attribute data can be combined in a single processing step. The process of combining maps together is often called **map** or **cartographic modelling.**

Analysis

Analysis is the process of inferring meaning from data. Analysis is often carried out visually in a GIS, as already indicated. Analysis in a GIS can also be carried out by measurements, statistical computations, fitting models to data values and other operations. For example, an analysis of areas on a map may lead to a table (or histogram) summarizing the proportions of a region underlain by classes of surficial materials, or an area cross-tabulation summarizing the overlap relationships of bedrock classes and surficial classes. A statistical summary might be used to compare the mean and standard deviation of airborne uranium measurements by rock type. Or a regression model might be fitted to bivariate geochemical data and the resulting predicted values displayed as a map to see if deviations from the model were related to other known environmental factors. Sometimes a classification analysis based on clustering together locations with the same multivariate characteristics can lead to useful interpretation. Often the tools for specific analyses are not found as part of a GIS, but require that data be exported to other computer programs. Analysis is carried out either on data organized as maps, or on data organized as tables.

Spatial analysis in a GIS sense means simply the analysis of spatial data. For example, the area cross-tabulation of two maps may lead to useful conclusions about the relationship between the two maps, although spatial coordinates play no direct role in the statistical summary. In the statistical literature, however, spatial analysis often implies analysis specifically involving spatial location. For example, trend surface analysis is a method of fitting a mathematical surface to observed data values, and explicitly uses the spatial coordinates in the calculations. The book by Unwin (1981) is a good introduction to this more specific type of spatial analysis; see also the advanced book by Cressie (1991).

Prediction

The purpose of a GIS study is often for prediction. For example a number of data layers indicative of gold deposits might be combined together to predict the favourability for gold as a new map. Such a map may then be used as a basis for making exploration decisions, or land-use decisions such as "Is this region suitable as a national park?" Prediction in a GIS involves the use of map algebra for defining symbolic models that embody the rules for combining data layers together. Prediction is sometimes a research exercise to explore the outcome of making a particular set of assumptions, often with the purpose of examining the performance of a model. For example, one might be interested in the number and area of sites predicted as unstable with a slope stability model, as a function of changes in slope and soil saturation, simply to evaluate model performance and sensitivity. Alternatively, the purpose might be to use the results for choosing building sites or planning road construction. The modelling tools of GIS provide the means to apply spatial data in problem solving, and take spatial data beyond simply the retrieval and display of information.

GIS AND RELATED COMPUTER SOFTWARE

There are a number of close relatives in the family of computer software products developed for handling spatial data, as illustrated in Figure 1-1. Many of these are computer programs that are similar to GIS in some, but not all, respects. Even amongst the products that qualify as full-fledged GIS, there is a great range in functional capabilities, some excelling at making cartographic products, some being good for map modelling, others offering superior database management, and so on. The following paragraphs provide a cursory survey of some of the software categories. This helps to clarify what is, and what is not, a GIS, and also illustrates some of the factors that have influenced the development of GIS.

Because of the focus of this book, GIS has been placed at the centre of the diagram in Figure 1-1, but in fact any one of the boxes could be at the hub. If the book were about database management, for example, the DBMS box would be the central focus. An alternative representation is to show a series of overlapping boxes, because in reality many of the programs (and systems of programs) overlap in functionality. Furthermore, the situation is

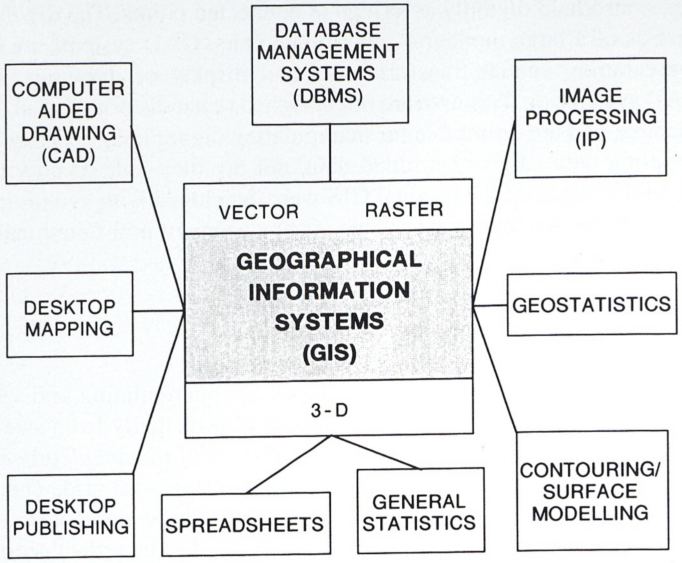


FIG. 1-1. GIS and other related software systems, many of them dealing with spatial data and including some GIS functionality.

always changing, because the commercial systems release new versions with expanded capabilities on a regular basis. This leads to a convergence of functionality, with for example, desktop mapping systems offering advanced functions that make them quite close to GIS, or computer-drawing packages that include the GIS projection transformation functions. This fuzziness between different kinds of systems, and the steady evolution of software systems, also make it difficult to choose a program suitable for a particular application. There are no specific recommendations here, just a note of caution that the spectrum is broad, the categories not very well-defined, and the boundaries between categories move with time. It is rather like buying hardware; you must base your decision on data available at time x, although you know that at time $x+t$ the data and your decision will change.

Two important close relatives of modern GIS are computer aided drawing (CAD) systems and image processing (IP) systems. Both kinds of system deal with spatial data, but they are based on quite different data models and functions. However both CAD and IP have played an important role in the development of GIS.

Computer Aided Drawing

CAD systems were originally developed for engineering drawings. They employ a **vector** data structure for representing points, lines and graphical symbols. The use of *vector* means that a point in a drawing is defined by a pair of spatial coordinates (rather than a single number or scalar), and that lines are built up by a series of ordered points. Areas are represented by boundary lines, also held digitally as strings of connected points. The digital data defining a drawing consists of a large number of coordinate pairs. CAD systems are excellent for the digital representation, spatial transformation and display of drawings, including maps. However CAD systems on their own are not designed to handle non-spatial attributes, except in a basic manner, and are unsuitable for manipulating digital images. They are therefore not good for handling data tables or gridded data, nor are they able to provide the analytical functions of a GIS. Many of the original GIS were developed with vector data structures for handling spatial data, but later added database management and functional capabilities for analysis and modelling.

Image Processing Systems

On the other hand IP systems were developed for manipulating and visualizing digital images in **raster** format, originally images generated principally from satellite sensors and also medical imaging sensors. A *raster* is simply a lattice of **pixels** (*pic*ture *el*ements) similar to those on a TV screen. Each pixel is like a cell in a rectangular grid. The terms raster and grid are used interchangeably in the GIS literature. Spatial locations of pixels are not stored explicitly with each pixel value, but are stored implicitly by the sequence in which the pixel data is held digitally. IP systems are very strong for the display and analysis of digital images, and the gridded data structure makes the overlap and combination of spatial datasets simpler to compute than in vector systems. However pure IP systems are often weak at handling vector data and they lack the linkages to tables containing non-spatial attribute data. They are not generally suitable for producing high quality maps, because the boundaries of map units are jagged, although some high-resolution digital images are of map quality. Some of the original GIS were based on a raster data structure and treated all spatial data as a series of gridded layers, like an IP system. When the data being analyzed is spatially distributed over earth's surface (as opposed to images of thin section of rocks, or images of human insides) an IP system is the same as a GIS. Many of the more advanced IP systems have improved their vector handling, can add cartographic annotation to the hardcopy of images, and possess analytical and modelling capabilities.

3-D GIS

Most commercial GIS have been developed for applications in a variety of fields, and are not specifically orientated towards the geosciences. GIS designed specifically for geological work, particularly for mining and oil exploration, need to be fully three-dimensional, so that each data object is characterized by its location in space with three spatial coordinates (x,y for horizontal and z for vertical position). This allows two or more objects that occur at the same (x,y) location to be distinguished by their z value. For geoscientific data where objects at one (x,y) location only have a single z value, two-dimensional GIS is adequate. Most 2-D GIS have facilities for perspective display of surfaces that are single-valued (not more than one z value per location). This is often referred to as a two-and-a-half dimensional capability. The "height" of the surface can be any attribute, not simply elevation. For projects that involve

data such as multiple points down boreholes, seismic sections, recumbent or complex faulted structures, or three-dimensional geotechnical and geophysical data, three-dimensions are needed. This book does not devote much space to 3-D topics, and those interested in the field are referred to the books edited by Raper (1989), Turner (1992) and Pflug and Harbaugh (1992) as an introduction. There are a number of 3-D GIS available commercially, and there is clearly some overlap of functionality with 2-D GIS. However, the specialized data structures, visualization and analysis needs associated with 3-D problems make 3-D systems a rather different breed from regular 2-D GIS, at least at the time of writing in 1993.

Database Management Systems

Another important relative of GIS is DBMS, which stands for database management system. DBMS are computer systems for handling any kind of digital data. Some form of database management system lies at the heart of any GIS, and many commercial GIS are explicitly linked to a particular DBMS. Many of the data collected by geologists are stored as tables of numbers and text. For example, geochemical data are often recorded in tables with the rows being sample sites, and the columns being the chemical elements. When the location of each site is recorded by a pair of spatial coordinates (i.e. a vector), such tables comprise one of the most important inputs to a GIS. Where a CAD system is used in conjunction with a DBMS, many of the data handling and vector functions of a GIS can be implemented, although the data structures are usually not sufficiently complex for the more advanced analysis and modelling operations.

Desktop Mapping Systems

These are systems mainly for display and query of spatial data. Often they use databases that have already been assembled in a GIS or DBMS. They are less expensive and are easier to learn than a full GIS. They are suitable for users of spatial data that need to visualize and explore an existing database, without getting into database creation or advanced analysis and modelling.

Contouring and Surface Mapping Packages

Where data are organized as a table, with each row being the record of data sampled at a geographic location, the immediate goal of data analysis is often to make contour maps of one or more of the variables, stored as columns of the table. For example, the depth to formation tops held as columns in a file of well data are often contoured to show the shape of a stratigraphic surface in plan view. Surface mapping is a more general term for the estimation of surface characteristics from irregularly-spaced point data, not simply expressing the result as contour lines, but also as gridded or triangulated data structures. Some surface mapping systems offer sophisticated analysis, such as the handling of faults in surfaces, permit operations between multiple mapped surfaces, provide visualization and modelling tools, and approach full GIS functionality. Surface mapping packages have been widely used by geologists, particularly in the oil industry. Several chapters of Davis (1986), one of the standard textbooks for statistical methods in geology, are devoted to contouring and statistical analysis of geological surfaces.

Geostatistics Programs

Geostatistics is the branch of statistics dealing with **regionalized variables**. A regionalized variable is a quantity whose value changes with spatial location, and whose behaviour is somewhere between a truly random variable and one that is deterministic. The behaviour of regionalized variables is studied with variograms, that depict the average squared difference between data values as a function of distance and orientation between data locations. Variograms are used in kriging, a method of estimating the value of a regionalized variable from scattered data points. Kriging is a popular method of spatial interpolation for contouring and surface mapping. Geostatistics program packages are normally designed to handle point location data, similar to contouring and surface mapping programs, and can generate co-variograms and co-kriging estimates based on the spatial co-variation of pairs of variables. A good introduction to geostatistics is the book by Isaaks and Srivastava (1989).

Mathematical Morphology Programs

Mathematical morphology refers to a branch of spatial analysis that deals with the shapes and sizes of geometrical objects in images, Serra (1982). It is an important field for the extraction and analysis of features in images, and has been applied geologically mainly to images of rocks in thin or polished sections, see for example Fabbri (1984) . The extraction of features on satellite images and the analysis of features on maps is a methodology of importance to GIS, although at present few IP or GIS incorporate mathematical morphology routines.

Other software

Other programs often used in association with GIS for specialized tasks are **spreadsheets**, **statistical analysis** programs, particularly for multivariate analysis, and **expert system** shells.

CUSTODIAL VERSUS PROJECT-RELATED GIS

A major proportion of GIS resources, particularly in large organizations, is devoted to the development and maintenance of large **custodial** databases that serve as data sources for a large group of users over an extended period of time. At the opposite extreme is the category of GIS activity that is **project-related**, involving datasets that are assembled for a particular purpose for a small group of users, and are not maintained after the duration of the project. There are some important differences between the typical activities and concerns of those working with custodial databases and those engaged in project-related work.

The size of a custodial database is likely to be much larger than the data gathered for a project-related study. A custodial database must employ data standards that are widely accepted and stable with time, whereas a project-related set of data need maintain standards appropriate only for the purpose at hand. A typical custodial database might be a national geochemical database, or regional geological map database for a province or state. On the other hand, a project-related GIS study might be to map the landslide potential for a small region for the purpose of planning the location of a new housing subdivision, road or pipeline, using data generated just for that application.

In order to create a custodial database, data standards must be established for data entry and data maintenance that can be applied by different groups of people over time. Thus both the geochemical and the geological map database require detailed data definitions, and data models that establish exactly how data items are to be structured and organized. The GIS chosen for custodial use must excel at providing the functions needed for updating and editing of spatial data. The project-related GIS need not be as rigorous on the data standards issue (although the data model must still be unambiguously defined), nor so strong on the data maintenance side, but must be suitable for data analysis and modelling, providing a flexible computing platform for research rather than production work.

Note that we have used the word *database* for the custodial GIS, whereas for the project-related data the words *collection of spatial datasets* are more appropriate. The very notion of a database usually implies long-term custody, maintenance and access by multiple users. (In practice, however, the term spatial database is applied by those using GIS for both custodial and project-related types). In project-related studies there are seldom the resources for long-term maintenance, and once the immediate objectives of the project are complete, the datasets are backed up on to mass-storage devices, such as tapes or disks, for long-term "mothballing". In practice, the data collected together for geological GIS projects are generally obtained in part from custodial databases (often the topographic data, and data from regional geochemical and geophysical surveys) and in part from less formally organized sources and from custom-digitized maps.

GEOLOGICAL APPLICATION OF GIS

Mineral potential mapping has been selected to illustrate a typical geological application of GIS. However the operations that are carried out for mineral potential mapping are in many respects similar to those employed for a variety of other GIS applications, and the aim here is to illustrate the general approach rather than to dwell on the particular details of mineral deposits. Search for other kinds of geological resources, evaluation of hazards, environmental impact and site selection studies also require the simultaneous appraisal of spatial data from several sources. Some other earth science applications of GIS that could be used to illustrate GIS methodology are:

1. Hazard mapping related to slope stability and landslides, earthquake damage zonation, volcanic eruption impacts, flood damage from rivers and tsunamis, coastal erosion, impacts of pollution as a result of mining or industrial activity, global warming.

2. Site selection for engineering projects, such as waste disposal (municipal landfills, nuclear waste in disposal wells), pipeline, road and railway routing, dams, building developments.

3. Resource evaluation for a variety of geological commodities besides metallic minerals, such as water, sand and gravel, building stone, petroleum, natural gas, coal, geothermal energy.

4. Investigation of possible cause and effect linkages of environmental interest between different spatial datasets, such as the incidence of disease (in plants, animals or man) in relation to geochemical patterns in rocks, or soils, or water. Diseases may also be related to complex combinations of spatial environmental factors.

5. Exploratory investigations of spatial inter-relationships between datasets during the course of geological research, such as understanding the regional geochemical and geophysical signatures of I- and S-type granites, or the evaluation of spectral signatures from satellite images in relation to lithology and vegetation.

Mineral Potential Mapping

Mineral exploration is a multi-stage activity that begins at a small scale and progresses to a large scale, ultimately leading to the selection of sites as targets for drilling for buried deposits. At a small scale, exploration companies must delineate general zones that may be of potential interest for mineral deposits of a selected type, usually based on broad geological considerations. At a medium scale, parts of these general zones are selected for more detailed follow-up exploration, based on evidence from geological mapping, regional geochemical and geophysical surveys and the locations of known mineral occurrences. Having identified those more specific zones of favourability, targets may be selected directly, or a further stage of detailed survey work undertaken. Ultimately, this process leads to a large-scale map showing the locations and ranking of potential sites. Of course, actual drilling decisions are based also on other considerations such as physical access, and economic factors. The decision making process must consider many types of spatial data together.

Exploration is similar to mineral resource assessment, an activity often undertaken by governments in order to compare the mineral value of a tract of land with its value for other activities such as forestry, agriculture, or recreation. Like mineral exploration, mineral resource assessment involves the subdivision of land into zones according to mineral favourability. However, in order to assign a monetary value to the mineral resources (needed for making comparisons with competing land uses), the number of estimated deposits, plus their size and grade, must be attached to the favourable tracts. Such information can then be used to build an inventory of mineral resources, and to facilitate land use decisions like the siting of parks, or the settling of land claims.

Both mineral resource estimation and mineral exploration utilize spatial data from a variety of sources. In the past, the selection, evaluation and combination of evidence for mineral deposits was undertaken with the aid of a light table, the various maps being physically

superimposed on one another to determine the overlap relationships between anomalies. GIS has greatly improved the efficiency with which this process can be carried out, and expanded the possibilities for specialized data processing and spatial data analysis.

The application described in this section deals with evaluation of mineral potential of a small area in Manitoba, Canada. This work was undertaken as part of a research project at the Geological Survey of Canada dealing with new technology applied to exploration for base metals, focused on volcanogenic massive sulphide (VMS) deposits in the Snow Lake greenstone belt. An early phase of the work is described in Reddy et al. (1992), and the illustrations shown here come from a second phase of study due to be published by the Geological Survey of Canada in 1994.

Conceptual model

Mineral deposits can be grouped or classified into different types, depending on their characteristics. No two deposits of a single type are completely identical, and sometimes one class may include a range of variation. Each class can be represented by an idealized mineral deposit, known as a **mineral deposit model**, one that has all the typical characteristics of the group. Mineral deposit models are conceptual models, usually described in words and diagrams. They describe the typical characteristics of a group of deposits, are accompanied by an interpretation of the processes of deposit formation, and are useful for providing criteria for mineral exploration (Hodgson, 1990). Mineral deposit models are important for providing the theoretical framework to guide GIS studies of mineral potential. They help in data selection and data modelling, in deciding which features to enhance and extract as evidence, and for deciding how to weigh the relative importance of evidence in estimating mineral potential. Conceptual models of different types are important for all kinds of GIS applications.

Volcanogenic massive sulphide deposits are formed at volcanic vents on the ocean floor. For a description of the characteristics and genesis of VMS deposits in general, see Lydon (1984, 1988). The Chisel Lake VMS deposit occurring in the Lower Proterozoic Snow Lake greenstone belt is described by Bailes and Galley (1989), who summarize the geology and mineral deposits of the Anderson-Chisel-Morgan Lake area. In this region (the area of study), the principal characteristics of the deposits useful for potential mapping are as follows.

1. The deposits occur within thick sequences of subaqueous volcanic rocks. Deposits are found mainly in felsic lava flows, but also in other volcanic rock types. Proximity to the contact between felsic flows and volcanoclastic units may be important.

2. The volcanic rocks are associated with large felsic intrusions, believed to act as a subvolcanic heat source driving a hydrothermal circulation. Seawater is thereby pumped through the volcanic rocks, from which metals are dissolved and re-deposited at or near vents to the seafloor.

3. The vent areas are associated with dykes and vertical fractures that cut through the volcanic rocks, localizing the upward flow of hydrothermal fluids. Heat from the intrusive dykes also helps to drive the circulation.

4. The hydrothermal fluids react with the volcanic rocks to produce alteration minerals. Silicification and the presence of amphibole minerals occurs in large zones, semiconformable with the volcanic stratigraphy. Ankerite occurs in more restricted zones.

5. Weathering and erosion of deposits produces a dispersion halo of metallic elements in drainage sediments and till.

6. Regional magnetic surveys show variations in magnetic susceptibility of the rocks, mainly due to differences in rock type. Sharp gradients in the magnetic field measurements usually reflect changes in rock type at geological contacts.

7. Geophysical data can be particularly important evidence for VMS deposits, which have distinctive electrical characteristics. Airborne and ground EM data are widely used to explore for these deposits, particularly in association with magnetic data.

These criteria constitute part of the deposit model used in the GIS study, providing a framework for data selection and analysis.

Brief Description of GIS study

Most GIS projects can be boiled down to three major steps or stages, as illustrated in Figure 1-2 for mineral potential mapping. The first step is to bring all the appropriate data together into a GIS database. The second step is to manipulate the data to extract and derive those spatial patterns relevant to the aims of the project, which in this case are the patterns critical as evidence for VMS deposits. The third step is to combine the derived evidence to predict mineral potential. Consider these steps in more detail for the VMS example, with reference to Figure 1-3.

Step 1

The initial database-building step was the most time-consuming phase of the project. This is typical of GIS studies in general. It involved establishing the spatial extents of the study area, deciding an appropriate working projection, and assembling the various spatial data to be used in the study in digital form, properly registered so that the spatial components overlap correctly.

The most important source of data was a recent geological map published by Bailes and Galley (1992). The paper map was digitized, transformed from table coordinates into a working projection, converted from vector into raster mode, and simplified by grouping some

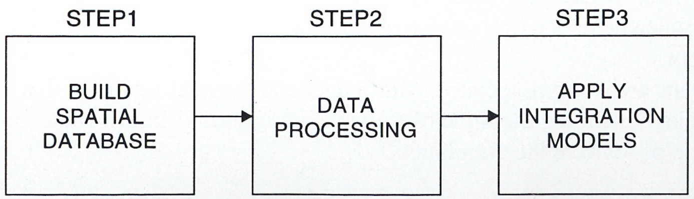


FIG. 1-2. Mineral potential mapping with a GIS as a 3-step process.

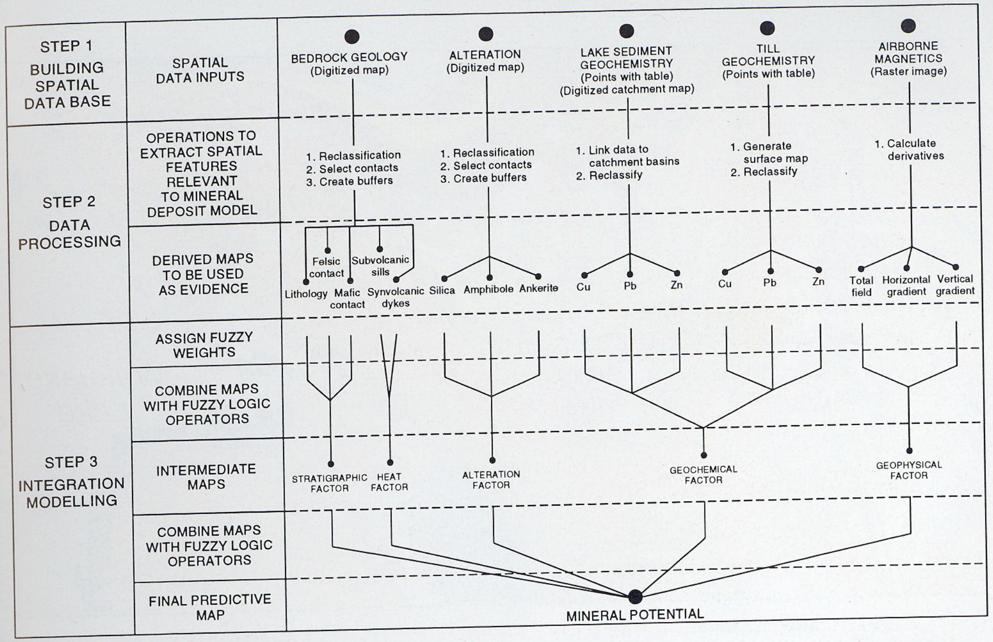
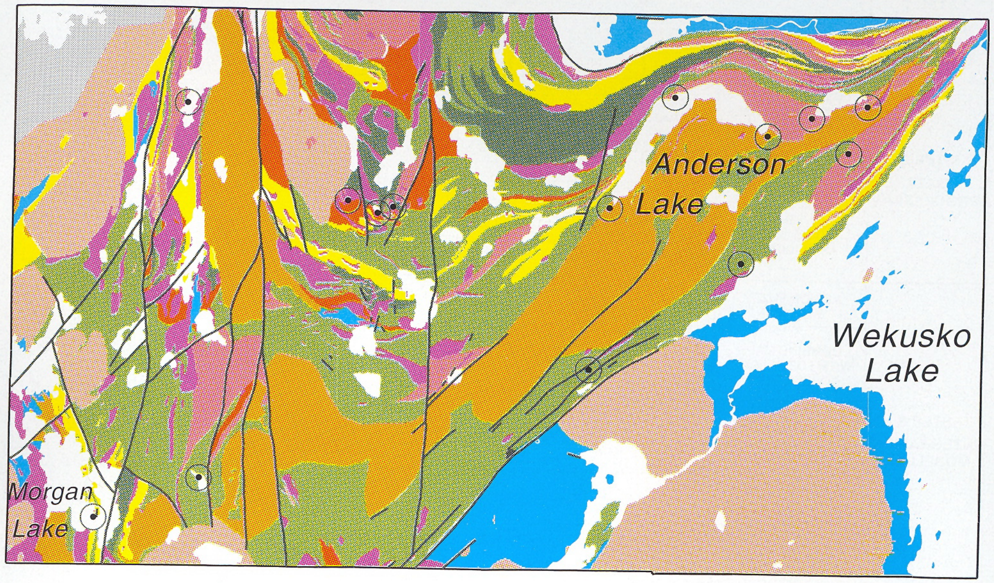
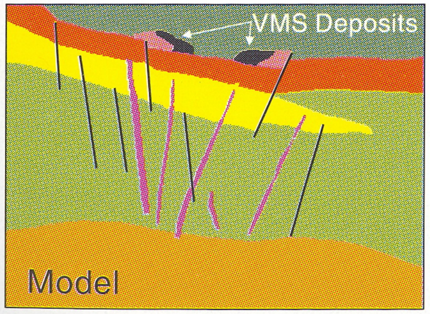


FIG. 1-3. Flowchart for the 3-step mineral potential study of part ot the Snow Lake greenstone belt, Manitoba. The three steps in the study are: 1) A database phase of assembling the available spatial data in a GIS; 2) Extracting and enhancing those features of the primary datasets important for predicting volcanogenic massive sulphide deposits; and 3) Integrating the spatial evidence together using modelling to predict mineral potential. The final result is a map showing regions ranked according to favourability.

units together, see Figure 1-4. The alteration map was also digitized from the same source, and subjected to a similar treatment, Figure 1-5. Geochemical data were available from a regional survey of lake sediments. Samples taken from the many lakes in the area were analyzed for a variety of metallic elements. The data were obtained in digital form as a table, one row per sample, with columns containing spatial and non-spatial attributes. The spatial locations were recorded as (latitude, longitude) coordinates, and the chemical attributes as numerical fields. Each sample was treated as being representative of the local drainage catchment basin in which it occurs, so the data table was applied to basins instead of to the sample points, see Figure 1-6. The catchment basins were digitized from a topographic basemap, as were the drainage features of the area. Aeromagnetic data were brought into the database in a gridded raster format, and converted to the working projection, Figure 1-7. These geophysical data had already undergone several processing steps before being obtained as a digital file for the project. The original flight-line data had been transformed from a series of point measurements along lines and interpolated on to a regular grid. Likewise, the mineral deposit and geochemical sample locations were already in digital form when they were obtained.



-1 0 5 km



LEGEND

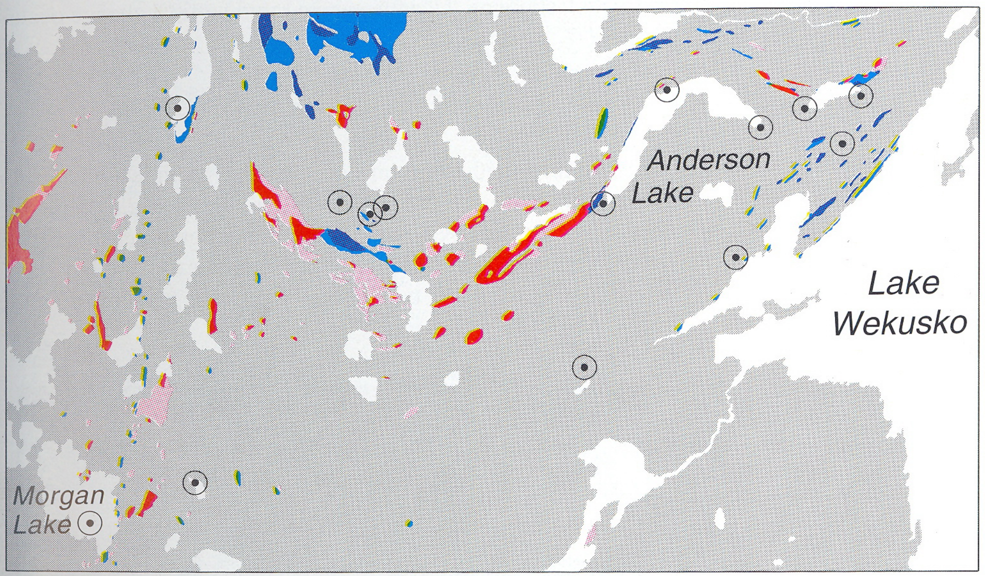
- Post Amisk map units
- Felsic Flows
- Mafic Flows
- Felsic Volcaniclastic Rocks
- Mafic Volcaniclastic Rocks
- Mixed Volcaniclastic Rocks
- Other sedimentary rocks
- Subvolcanic tonalite sills
- Synvolcanic intrusive rocks
- Unmapped

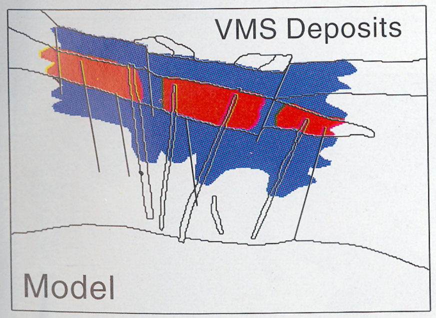
Mine/Occurrence . . . . ⊙
Faults(Unclassified) . . . ———

FIG. 1-4. Simplified geology of the Morgan-Chisel-Anderson lakes area, digitized and generalized from a 1: 15 000 scale map by Bailes and Galley (1992). A thick sequence of submarine volcanic rocks (green and red units) are underlain by large tonalite sills (orange), believed to have acted as a heat source for a hydrothermal system. Fracturing and intrusion of dykes (purple) helped to focus the transport of metal-rich fluids during various phases of submarine volcanism. The small model diagram shows a typical cross-section, with sulphide deposits often associated with felsic flows, being deposited at or close to the paleoseafloor. The points show the known VMS deposits, some of which are mines, from Fedikow et al. (1989).

Step 2

The second step was to process the input layers to extract the evidence critical to the prediction of VMS deposits. The geological map was generalized into a smaller number of map units or classes, preserving the information believed to be significant for mineral prediction. The contact between felsic flows (rhyolites) and volcanoclastic units was selected from the geological map and dilated, or buffered, to produce a proximity map, because some





LEGEND

- Si (Strong)
- Si (Weak)
- Amphibole (Strong)
- Amphibole (Weak)
- Ankerite/Epidote

Mine/Occurrence . . . . . ⊙

-1 0 5 km

FIG. 1-5. Map showing the extent of alteration zones, also digitized from the Bailes and Galley (1992) map. Large zones of semi-conformable alteration, mainly silicification and amphibole alteration. Alteration minerals indicate the occurrence of past chemical reactions between volcanic rocks and hydrothermal fluids.

VMS deposits show a spatial association with this type of contact. Likewise, proximity maps were derived from the geological data showing distance to dykes and sills, used as evidence of proximity to a heat source, as shown in Figure 1-8. The different zones of alteration were separated and buffered, allowing the presence, and proximity to, alteration of various types to be modelled. The geochemical data from elements believed to be important indicators of base-metal deposits were classified on a scale designed to accentuate and enhance those areas with anomalously high values, and converted from the data table to catchment basin maps, Figure 1-6. The magnetic data was enhanced by choosing a suitable classification scheme, and a derivative map was constructed, showing the areas where rapid vertical change occurs in the local magnetic field, Figure 1-7.

Step 3

The third step consisted of combining together the various maps that provide evidence for VMS deposits. This was carried out in stages, with the ultimate product being a predictive map showing the relative favourability, or VMS potential. The combination process involves the weighting and fusion of evidence and can be carried out in a number of different ways, as discussed at some length in Chapter 9. In this case, fuzzy logic was employed as the

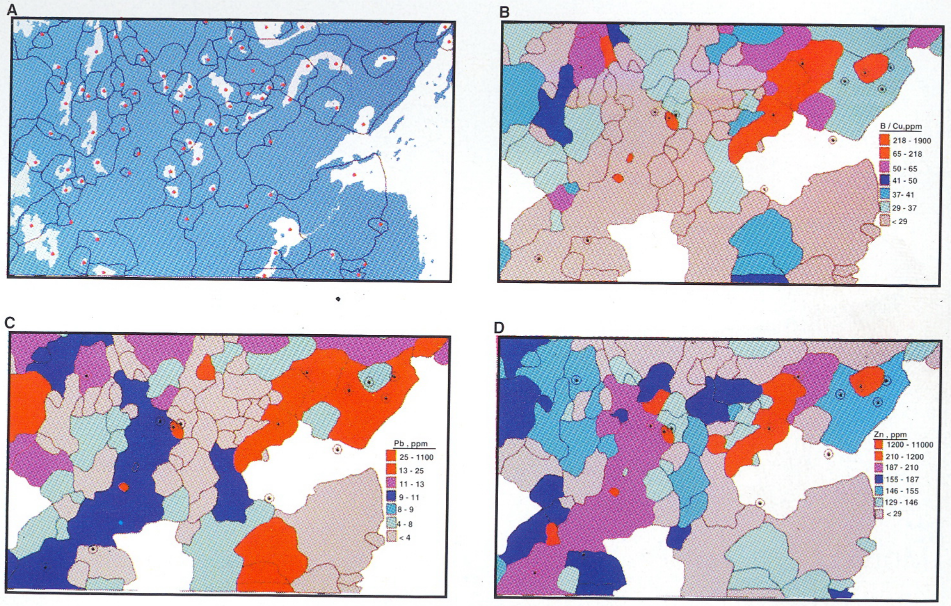


FIG. 1-6. Geochemical maps based on element concentrations in samples of lake sediment. **A.** Sample locations, and boundaries of digitized catchments used to make element maps. **B.** Basins classified according to the copper (Cu) content. **C.** Lead (Pb) content. **D.** Zinc (Zn) content. The dots enclosed by circles are the deposit locations.

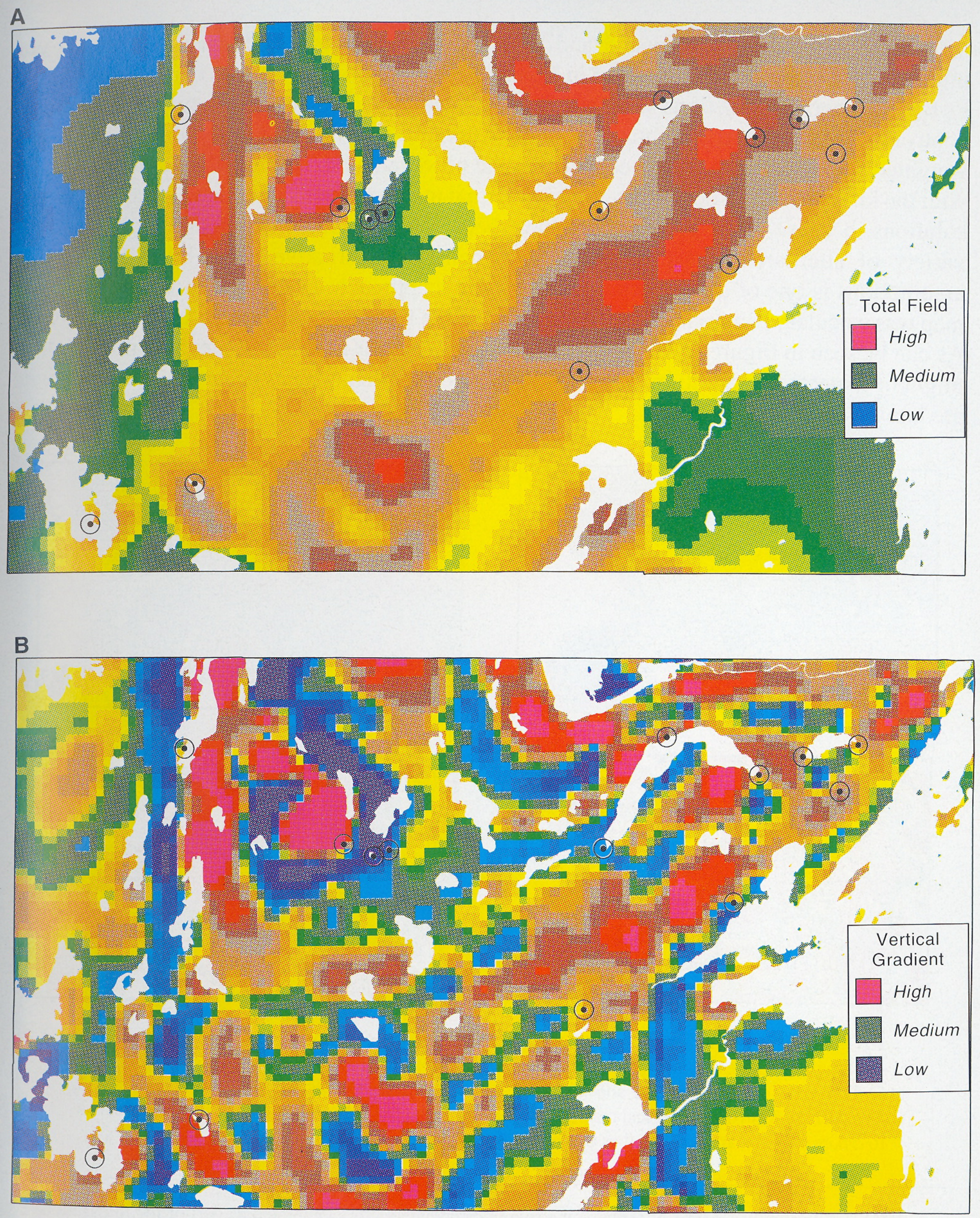


FIG. 1-7. Geophysical maps, from an airborne magnetic survey. The original data were measured at points along flight lines, but now have been interpolated on to a grid and coloured by intensity. **A.** Total field data. **B.** Vertical gradient data. The dots are locations of deposits.

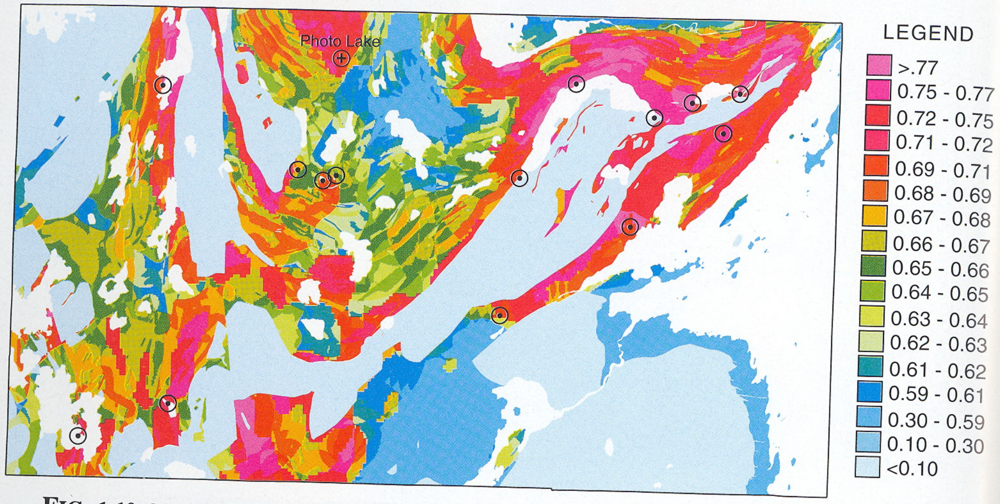


FIG. 1-10. Map showing the favourability for volcanogenic massive sulphide deposits, derived by combining evidence from the five intermediate factors. This is the mineral potential map that can be used for exploration decision-making, for example to decide priority areas for additional detailed surveys. The known deposits (black dots) occur in highly favourable zones, as expected. A new previously unknown deposit (ca. 1 m tonnes) was discovered in 1994 within the large favourable zone that intersects the top of the map.

Not shown here is the ability to make interactive spatial queries of the data and results at specific locations as determined by the cursor position. This greatly helps to understand the reason why particular locations have high VMS potential. Also not shown here is the kind of specialized output, such as perspective displays, that help to visualize the results.

It is emphasised that the conceptual model was essential for guiding the study. The choice of data, the kinds of information extracted from it, and the assignment of weights to evidence were all dependent on the deposit model. For example, vegetation maps, airborne radiometric maps and Landsat imagery were not employed, because they were not directly usable as VMS evidence. Unfortunately, some evidence, defined in the model as being valuable, such as airborne EM data, was also not available for the study. The extraction of critical evidence, employing transformation methods to select contacts, buffer lines, calculate derivatives, interpolate from points to surfaces, generalize maps by reclassification, was greatly influenced by the conceptual model. Most importantly, the assignment of fuzzy weights and the choice of fuzzy logic operations for combining evidence were due to the deposit model.

In short, analysis and modelling of spatial data in a GIS is not simply a matter of throwing the data layers into a "black box" computer program. A conceptual model, preferably formulated in the early stages of a study, is used to guide the various stages of GIS processing.

REFERENCES

Bailes, A.H. and Galley, A.G., 1989, Geological setting of and hydrothermal alteration below the Chisel Lake massive Zn-Cu sulphide deposit: *Manitoba Energy and Mines, Minerals Division, Report of Field Activities*, p. 31-37.

Bailes, A.H. and Galley, A., 1992, Chisel-Anderson-Morgan Lakes Preliminary Map 1992 S-1: *Manitoba Energy and Mines*.

Cressie, N.A.C., 1991, *Statistics for Spatial Data*: John Wiley & Sons, Inc., New York, 900 p.

Davis, J.C., 1986, *Statistics and Data Analysis in Geology*: Second Edition, John Wiley and Sons, New York, 646 p.

Fabbri, A.G., 1984, *Image Processing of Geological Data*: Van Nostrand Reinhold, New York, 244 p.

Fedikow, M.A., Ostry, G., Ferreira, K.J. and Galley, A.G., 1989, Mineral deposits and occurrences in the File Lake area, NTS 63K/16: *Manitoba Energy and Mines, Mineral Deposit Series*, 277 p.

Hodgson, C.J., 1990, Uses (and abuses) of ore deposit models in mineral exploration: *Geoscience Canada*, v. 17 (2), p. 79-89.

Isaaks, E.H. and Srivastava, R.M., 1989, *Applied Geostatistics*: Oxford University Press, New York-Oxford, 561 p.

Lydon, J.W., 1984, Ore deposit models - 8: Volcanogenic massive sulphide deposits, Part 1: A descriptive model: *Geoscience Canada*, v. 11, p. 195-202.

Lydon, J.W., 1988, Ore deposit models - 14. Volcanogenic massive sulphide deposits, Part 2: genetic models: *Geoscience Canada*: v. 15, p. 43-65.

Pflug, R. and Harbaugh, J.W. (eds.), 1992, *Computer Graphics in Geology: Three-Dimensional Computer Graphics in Modeling Geologic Structures and Simulating Geologic Processes*: Lecture Notes in Earth Sciences, Springer-Verlag, Berlin, v. 12, 298 p.

Raper, J., (editor), 1989, *Three Dimensional Applications in Geographical Information Systems*: Taylor and Francis, London-New York, 190 p.

Reddy, R.K., Bonham-Carter, G.F. and Galley, A.G., 1992, Developing a geographic expert system for regional mapping of volcanogenic massive sulphide (VMS) deposit potential: *Nonrenewable Resources*, v.1(2), p. 112-124.

Serra, J., 1982, *Image Analysis and Mathematical Morphology*: Academic Press, London-New York, 610 p.

Turner, A.K., (editor), 1992, *Three-Dimensional Modeling with Geographic Information Systems*: Kluwer Academic Publishers, Dordrecht, 443 p.

Unwin, D., 1981, *Introductory Spatial Analysis*: Methuen, London-New York, 212 p.