Chapter 8
Geological Mapping by Computer

CHRISTOPHER GOLD

INTRODUCTION

Davis (1973), p. 298 states: ‘Maps are as important to earth scientists as the conventions for scales and notes are to the musician, for they are compact and efficient means of expressing relationships and details. Although maps are a familiar part of every geologist’s training and work, surprisingly little thought has gone into the mechanics and philosophy of geologic map making.’ He then states that most of the work in this field is done by geographers, and gives various examples. One of the difficulties with this approach of merely accepting rather than developing techniques is that many of the wide variety of map types that geologists use are uncommon in other disciplines.

Maps are displays of items of information that have both non-spatial components-without both of these it is not possible to produce a true map. Data that are not in suitable form must be transformed in order to make map production possible. Statistical data, for example, do not usually possess information about the location of each data point in space, but scatter plots or contour plots are frequently generated from data of this type, by arbitrarily considering two variables to represent ‘east’ and ‘north’ and then either plotting data point locations or contouring the value of some other associated variable. A point to note is that, given points within a scatter diagram, it is often desirable to contour their density. This density is a continuous variable expressing the number of points per unit area. Unfortunately, the resulting surface varies greatly with the sampling area chosen. This transformation from point location to density has been studied by various geologists, in particular Ramsden (1975, and in press).

Perhaps surprisingly, one-dimensional maps are commonly used by geologists: the geological section is the basic tool of much field work. The direction is usually vertical and the mapped variable may be either categorical, as in stratigraphic sections showing the rock types and formations present in some locality, or continuous, as with electric log measurements of resistivity, etc., down a drill hole. Much geological work consists of taking these one-dimensional maps and correlating between them to produce two-dimensional cross-sections (Figure 8.1 a) which imply, conceptually if not diagrammatically, three-dimensional models (Figure 8.1 b).
Alternatively, the three-dimensional model may be constructed by generating contour maps of each geological contact in turn.

8.1 Wind River Basin, Wyoming, stratigraphic cross-section and block diagram (Harbrough and Merriam, 1968; reproduced by permission of John Wiley & Sons, Inc.)

If only spatial information is present in a data set it may become necessary to eliminate one dimension by treating it as non-spatial for display purposes. An example is the case of a topographic map where east, north and vertical measurements are obtained for points on the atmosphere-lithosphere contact. In this case the elevation information is treated as non-spatial and displayed in various ways.
on the two-dimensional map sheet—perhaps by plotting spot elevations, colouring elevation zones or some other technique. It is, of course, possible to choose a different pair of map variables and draw a cross-section to the topographic surface, usually with the y-axis of the map representing the elevation.

Geologists basically study the earth, a three-dimensional object, and are frequently concerned with the variation of one or many parameters within three-dimensional space. This differs markedly from political scientists, for example, who rarely consider the vertical variation in voting preferences! (Both groups, however, would frequently like to be able to represent time as another spatial dimension.) This underlying concern with what is happening in the third, undisplayed, spatial dimension is of great importance in some, if not all, branches of geology. Considerable use is made of mapped data containing no intrinsic spatial information; for example, in examining how one measured chemical component in a set of rock samples varies with changes in another component. Most maps make use of data with only two spatial variables—examples of these are maps of the earth’s magnetic field, concentration of some economic mineral or element found in an area, or the type of rock to be found in a particular geographic region. This last is perhaps what most people think of as a geological map. A classical example of this is shown in Figure 8.2, the first true geological map, drawn by William Smith and published in 1815 and showing the major rock units in England and Wales. Nevertheless, a little thought will suggest that all of these examples are two-dimensional only because of the inability to collect samples far from the surface of the earth, not from limitations in the display technique. Indeed, many maps are intended to facilitate the intuitive estimation of variation in the third dimension, based on professional knowledge of the behaviour of the parameters displayed (see the discussion in Gold and Kilby, 1978). This is especially true for maps displaying structure and lithology, where structural boundaries and plotted bedding orientations are designed to enhance the three-dimensional model perceived by the compiling geologist.

Maps that are usually amenable to consideration as two-dimensional include many maps (geomagnetic, gravitational, etc.), geochemical maps (plotting the concentration of some element measured in samples of soil, rock, or water obtained near the earth-air or earth-water contact) and facies maps (partitioning some geological contact into zones with relatively homogeneous properties). Orientation and density maps will usually be two-dimensional as well. Depending on the scale and the objectives of the study, ‘geological’ or, more properly, ‘lithological’ maps may be considered to be two-dimensional, but as previously noted the third dimension may be indicated either by bedding orientation symbols or by the form of the geological contact, especially as it relates to the topographic surface. Topographic maps, structural contour maps, isopach maps, and other derivatives are of course three-dimensional in input, even if the vertical dimension must be treated as a mapped variable for display. These map types, along with related discussions of data capture and data base management, will be described under separate headings later in this chapter.
Perhaps it is worth making a comment here on the role of maps in geology. They are certainly a product—most geologists have produced maps. As well as output they are also input—all geologists have read maps; many kinds of maps. Rather than being attractive products to pass on to the end-user, most maps are an essential, and primary, means of communication between geologists, and most are not expected to
be meaningful even to those geologists not concerned with the particular topic under consideration. Since relatively few maps are intended for immediate comprehension by the general public, a high level of generalization is not usually desirable. Hopefully this explains, and even excuses, the great amount of detail found in many geological maps.

Having mentioned the role of maps in geology, we come to the main purpose of this chapter— to discuss the role of the computer in producing geological maps, and in particular how the advent of the computer has affected geological map production. The main details of these changes are described under each map type. In general, the first change produced by computer availability was the dramatic increase in the quantity of data points generated, with the consequent necessity to display them. This is still the driving force behind map automation. The first mapping techniques developed in response were contouring algorithms, and these are still the leading mapping functions used. It is becoming evident, however, that some of the earlier and more inefficient approaches to contouring are beginning to show signs of wear under the ever-increasing load of generated data.

Apart from contour mapping, the increasing volume of data has mainly generated programs for individual data point display (‘posting’) and statistical and other data transformations. Of course, once it became easy to generate many maps from one data set, many maps were requested. One offshoot of this is that it is becoming more common to standardize the acquisition of field data by using coding forms, thus facilitating data entry for computer generation of base maps that rapidly display the raw data for subsequent interpretation. It is now not unknown for the weekly mail plane to arrive at the bush camp to collect the week’s coding forms and deliver the latest base maps.

The primary effect of computers has thus been to facilitate the production of ever-increasing numbers of simple maps. In this way the geologist may more rapidly reach the stage of data interpretation, where his comprehension of geological processes may help him to make meaningful decisions from the displayed information. Nevertheless much research is in progress in other directions, often with the intention of computerizing or standardizing even more of the routine labour of the practising geologist.

GEOPHYSICAL MAPS

Geophysical mapping involves the display of a wide variety of physical parameters, for example the earth’s gravitational field, the earth’s magnetic field (both strength and orientation) and earthquake data. Also mapped, on a more local scale, are resistivity and self potential data and, even more commonly, seismic data for oil exploration. A wide variety of less common data types are also used.

Most of these measurement techniques have some features in common— large quantities of data are accumulated by automatic means, usually along flight paths or traverse lines. The data processing is frequently sophisticated and beyond the scope
of this study. Nevertheless difficulties exist in the management and display of the large quantities of data of this type that are relevant here. The key point, however, is that the display of large quantities of data necessitates the use of computer-based mapping systems.

While many types of geophysical measurement are possible, examples will be given for geomagnetic surveys (examining the earth’s magnetic field either for regional trends or local anomalies); gravity surveys (as above, but for the gravitational field); seismic ‘profiling’ to identify geological contacts by monitoring the reflection of shock waves; and regional seismic studies to determine earthquake locations.

Dawson and Newitt (1977) produced a regional map of the earth’s magnetic field for Canada, based on 5 1,133 data points (Figure 8.3A) for which the earth’s field had been obtained in three perpendicular directions. These data were obtained from an on-going data collection process from 1900 to the present although, since all three components of the earth’s field had to be measured at each location, the data points actually used were acquired between 1955 and 1974. Observations were obtained either from ground stations or, more recently and for the bulk of the data, from aerial surveys.

Since this was a regional survey, concerned with field fluctuations of the order of 1000 km in wavelength, it was estimated that a sixth-degree polynomial fit in each quarter of the map area would display such features. In addition, since the earth’s field varies with time, a cubic time component was used to adjust for data acquired at different dates. The resulting charts were plotted for estimated 1975 values and a correction applied to correct to sea-level elevation. These charts displayed D. (magnetic declination), H (horizontal intensity), Z (vertical intensity), I (magnetic inclination) and F (total intensity). These parameters (which are merely different ways of displaying the orientation and intensity of the field calculated using the orthogonal x, y, and z vectors), were plotted using the regression coefficients obtained independently for the x, y, and z components. Figure 8.3A shows the map for magnetic declination.

Kearey (1977) described a gravity survey of the Central Labrador Trough, Northern Quebec. Average sample spacing was 5 km, with some detailed traverses having a spacing of 1.5 km or less. Regional and residual gravity fields were estimated. The residual field was then related to the lithologic variation within the Trough.

Howells and Mackay (1977) performed a seismic reflection survey of Miramichi Bay, New Brunswick. A more common type of survey uses the refraction technique, which determines the propagation velocity of seismic energy through rocks and surficial sediments. Contrasting lithologies may be correlated with certain velocities, or velocity ranges, provided borehole control is available. This approach is the one primarily used in oil exploration. In the Miramichi study, however, the water depth and available equipment made a reflection survey more applicable.

The objective of the survey was to determine surficial sediment types and
8.3 Distribution of magnetic survey data for Canada 1955-73 and the resulting estimated magnetic declination for 1975. (Broken lines show annual change in minutes of arc (Dawson and Newitt, 1977).) Reproduced by permission of the National Research Council of Canada from the Canadian Journal of Earth Sciences, 14, 477-487 (1977).
thicknesses in the area, and over 300 line kilometres of seismic profiling was obtained. The equipment used consisted of a launch-mounted echo-sounder receiver with the ‘boomer’ sound generator towed behind on a catamaran. Position fixing for the survey was accomplished using a microwave system with three pairs of land-based slave stations. The overall position-fixing repeatability was about 10 m. The inner part of the bay was less than 7 m deep; the outer part sloping, to 15 m.

In the presence of soft bottom muds a good reflection was received from the harder underlying material. Where the sediments were harder penetration was not achieved, but sea floor roughness was used to distinguish relatively flat sands from the rougher exposed bedrock. This interpretation was aided by several shallow boreholes. On the basis of these measurements various conclusions were made about the geological history of the area.

Basham et al. (1977) examined the last data type we will consider under the heading of geophysics—the incidence of earthquakes. From the catalogue of Canadian earthquakes, they extracted all those with epicentres north of 59 degrees of latitude. Their objectives were to review what is known of the seismicity in northern Canada, and to comment on the spatial relationships between seismicity and other geological and geophysical features. The data consisted of events monitored since 1962, the first year a large enough number of seismographic stations were maintained in the area to provide reasonably comprehensive coverage. They state that the better-defined epicentres are probably located to within 50 km, with the worst having a location error of up to 100 km, depending on the distance to the nearest stations. They then attempted to relate the clustering of these observations to various other features and concluded that they may be explained by known deformational features, by sedimentary loading of the crust and by residual disequilibrium due to crustal unloading at the time of the last major glacial retreat.

In evaluating the impact of the computer on the types of studies mentioned here, it is clear that for the magnetic and gravity surveys the computer is a necessity to handle the many thousands of data points acquired. Indeed, it is probably true to say that the ability to acquire large quantities of data grew from the same technology as, and because of the availability of, the ability to display the resulting information. Thus without computer-based mapping the projects in their present form would not have been performed. In the seismic reflection and earthquake incidence studies sophisticated instrumentation was needed for data acquisition, but the projects could have been completed without the use of computer cartography.

GEOCHEMICAL MAPS

Geochemistry, particularly in mineral exploration, is one of the oldest geological skills. Agricola, for example, in 1546, described six basic flavours of spring water, that were of use in determining mineral deposits. Indicator plants-species that grow near particular types of mineralization-have been known for a millennium. Placer exploration-panning for gold, tin, etc.-is also ancient. Geochemical exploration
today is little different in this respect—elements are searched for in rock, soil, sediment, plant, and water samples. The main differences are in analytic instrumentation, sampling methodology and data processing of the results. It is not commonly realized how many geochemical analyses are, in fact, performed with modern rapid analytic instrumentation. In the 1971-72 year over 800,000 samples were analysed for at least one element in Canada, and over 300,000 in the U.S.A. Soil, rock, and stream sediment samples comprised the vast bulk of these rock samples being more popular in the U.S.A. and soil samples in Canada, presumably due to northern terrain types. Most of the other analyses were performed on water and vegetation samples.

Sample collection has several aspects. It is the job of the geochemist to determine the best position in the stream bed or the best depth in the soil profile to collect a sample. This should be based on his knowledge of the dispersion mechanisms of the element or mineral for which he is searching. Another aspect concerns the spacing of samples—clearly, if samples are too far apart the target anomaly may well be missed; and if samples are too closely spaced considerable money is being wasted. A decision must also be made as to what kind of sampling to perform—it may well be much cheaper to sample vegetation for elemental analysis than to dig a suitable hole for soil sampling. The number of elements to be analysed in each sample is another factor in the economic equation, since several elements may correlate well together, or certain element combinations may indicate either a false anomaly (where the accumulation of the target elements is merely due to the dispersion agents, such as ground water) or strongly suggest a true one.

The manipulation and display of data after sample acquisition and analysis is the final step. As might be expected from the cost of data acquisition, considerable effort is sometimes expended on this stage but, conversely, a surprisingly large number of surveys are satisfied with elementary processing techniques. This is, in part, explainable by the lack of consensus on methodology. In addition, survey objectives may be either regional geochemistry, where the objective is to observe the general behaviour of some component over appreciable distances, or exploration geochemistry, where the objective is to determine anomalously high concentrations of some mineral of economic interest. Since the concentration of the target material frequently varies greatly over even small distances, appreciable smoothing should be performed prior to map production. For regional analysis, trend surface methods (see Davis, 1973) are commonly used for computer-based estimation of smooth regional trends. In exploration geochemistry careful examination of the possible population distributions is necessary before areas possessing appreciable concentrations of the desired material can be satisfactorily distinguished from the ‘background’ range of expected data values (Govett et al., 1975).

The first requirement of data examination is to determine the background and threshold values for the region. These are defined as the average regional value and the upper limit of regional variation exclusive of anomalous zones. This raises two points. First, a region has to be defined within which the background variation is
more or less homogeneous (this may be identifiable from other information, such as rock type). Second, an anomaly is defined as having values outside (usually above) the expected background variation, so the argument is circular—as indeed it is for defining regions or domains as well. This is a major problem in cartographic work in geology—the delineation of zones with similar properties. Determination of regions or anomalies is not usually a major problem when mapping by hand—other aspects of the data can be taken into account by the trained geologist. The advent of computer analysis, however, rapidly demands the introduction of computer-assisted mapping, with a consequent demand to specify the criteria more precisely. In theory this is good, since it forces objectivity. Nevertheless, in practice it has merely deterred the use of the computer on a production basis, since it appears to the untrained individual that much effort and statistical skill is needed merely to simulate the ‘obvious’ manual interpretation.

The above remarks apply to both data posting—where anomalous values may be flagged automatically—or for contour maps. In the second case production of a map honouring all data values is frequently undesirable, as the errors in sampling and analysis may generate surface fluctuations that obscure the subregional behaviour. In statistical terms, the correlation between neighbouring points is not very high. For most purposes, it is necessary to filter or smooth the data.

Data smoothing occurs every time the mathematical surface generated from the data fails to honour, or pass through, all the input points. This is a natural consequence of the gridding stage in most computer programmes, since the requirement that the surface passes through the interpolated values at the nodes of what is usually a relatively coarse intermediate grid precludes the possibility of their passing through the data points, especially if several occur in one cell. It is necessary in most cases to be able to control the extent of smoothing, and this may be achieved using various approaches. For regularly spaced data on a grid, Robinson and Charlesworth (1975) used a Fourier filter to eliminate both excessively high and low frequencies. Their information was concerned with subsurface stratigraphy in western Canada, and for oil exploration purposes they wished to examine features a few kilometres in length that could possibly be buried reefs. Unfortunately controlled smoothing of this type is not readily obtainable for irregularly spaced data.

Contouring is usually performed using an arbitrary grid over the map area and estimating elevations at each node. Clearly enlarging the grid spacing increases the smoothing of the data. This necessitates some interpolation method to estimate grid node values from nearby data points. A typical method is the moving average, where the value at any location is a weighted average of the nearby points, closer points being weighted more heavily than more distant points. Two sources of deviation of the surface from the data values should be noted. First, if the weighting function is such that at any data location any other data points contribute to the surface being generated, then the surface will not pass precisely through the original data. This depends on the weighting function itself. A weighting of \(1/d^2\), where \(d\) is the distance between data location and grid node, will not produce any smoothing. A
weighting of \( \exp(-d) \), on the other hand, will do so. If \( d = 0 \), the first case yields an infinite weighting, while the second gives unity. Secondly, the very existence of an intermediate grid smooths the data uncontrollably, since data points not lying at grid nodes will rarely be honoured.

To meet the need for smoothed maps two competing techniques have been used in geology: trend surfaces and kriging. Trend surface analysis consists of fitting a polynomial of fixed order in \( x \) and \( y \) to the data, using conventional regression techniques. This has the advantage that traditional statistical techniques may be used, although it is assumed in the method that local fluctuations (the ‘error’ component) are uncorrelated. This is reasonable for errors in measurement of a geochemical variable, but is not correct for real small-scale or local variations. In addition, there is no foolproof way to determine how high an order polynomial to use. Figure 8.4A shows a third-order trend surface map of nickel distribution in soil samples in Sierra Leone, and Figure 8.4B shows the same data contoured using moving average techniques.

8.4 Regional distribution of nickel content in soil samples over the basement complex, Sierra Leone: (A) cubic trend surface; (B) moving average surface (Nichol et. al., 1969; reproduced by permission of Economic Geology Publishing Co.)
The usual alternative to trend surface mapping is universal kriging, which is a form of moving average that is, the value at a grid node is a weighted average of some neighbouring points. Several preliminary stages are necessary in order to generate these weights. The first step is intended to handle precisely the point that trend surfaces cannot—that points close to each other in space are likely to correlate strongly with each other. For a given set of data (and, by implication, a single domain) the variation of this correlation is examined with increasing distance between samples. Clearly this correlation is large between close samples, and almost zero for widely separated ones. This effect—expressed on a diagram called a semivariogram—is the primary step in kriging. Obtaining a good semivariogram requires appreciable experience, since it is necessary to fit a suitable mathematical expression to the correlations observed in the data. If this correlation is unity at zero separation distance then no error component is associated with data measurement. If the value is less than unity then some error in measurement occurs that is, two samples at the same location would not be entirely correlated with each other—the ‘nugget effect’) and the surface would be smoothed. An introduction to map kriging is given in Davis (1973). Journel (1975) describes some approaches used in other applications. Govett et al. (1975) provide an excellent description of sampling strategy, relevant frequency distributions and anomaly detection for geochemical data.

In conclusion it could be said that the advent of computers has been a mixed blessing. While technological development has dramatically increased the number of analyses performed, it has not produced a satisfactory ‘hands-off’ display technique for the field geologist who is not comfortable with statistical principles in general and hotly debated ones in particular. Underlying all the statistical approaches is the question of distinguishing between local and regional variation. This is in essence a spatial (and perceptive) problem—a point that must not be forgotten in any statistical analysis.

FACIES MAPS

The object of many geological studies is to classify an area into several facies (zones, domains) each with relatively homogeneous properties. The initial information consists of a set of point samples analysed for various constituents. A facies is defined as a laterally continuous unit, a unit being expressed as any identifiably similar combination of properties. Thus we have two problems: how to classify a set of observations into several facies or categories, and how to display the end product. In many cases, facies are somewhat arbitrary subdivisions of a continuum, although occasionally sharp boundaries between units may exist. An overview of a wide variety of manually produced facies maps may be found in Forgotson (1960).

One may wonder why continuous variables should be broken into distinct categories, rather than being contoured as a continuous surface. The answer is that we may normally contour only one variable per map, and a sea-floor environment,
for example, consists of many interrelated variables. Statistical classification schemes such as cluster analysis may be used to help define several distinct groups of samples such that the within-group similarity is greater than between-group similarity. An alternative scheme is to use factor or principal components analysis in an attempt to generate one or more composite variables (consisting of varying proportions of the analysed variables), that express most of the overall variation. Each composite variable or factor may then be considered as continuous over the map area and contoured by conventional techniques.

These numerical techniques are all relatively recent, whereas facies maps have been constructed for many years. More subjective categories were used before the widespread availability of computers permitted extensive mathematical computations. Indeed, much of geology consists of this classification of continuously varying material into artificial but meaningful categories. It is still useful to describe a rock as a granite even if there are cases where the differences between it and other rock types are artificial or even vague. Thus some traditional ‘geology maps’ (or, more correctly, lithology maps) may have boundaries drawn between continuously gradational rock types. Often, however, the boundaries between the geographical occurrence of various rock types may be defined fairly precisely. The techniques described below are fairly widely used computer methods that have replaced, in some part, the intuitive drawing of facies boundaries from the tabulated or plotted analyses.

In the numerical processing of facies information the data are initially edited for errors. They are then usually standardized so that each variable has a mean of zero and a standard deviation of one. This step is used to prevent one variable from dominating subsequent operations merely because of large measured values.

The next stage, data reduction, is performed to reduce the number of variables used in further processing. In many situations some of the parameters that are measured may well correlate strongly with each other, as they are merely expressions of a few underlying forces. As an example, in a shore environment the aspect of the shore to the prevailing wind and waves may affect particle grain size as well as faunal species and abundance. Thus many of the measured variables may be redundant in that, individually, they add little to the description of the total variability of the environment. The main methods used for grouping variables are R-mode principal components analysis and factor analysis. Both of these methods, which are very similar in application, are used to generate a set of components or factors that are linear combinations of the initial variables. These are frequently obtained so that the first of these composite variables explains as much of the total variation as possible, the second as such of the residual variation as it can, etc. Consequently only a ‘few’ are required to explain ‘most’ of the original variation in the measured variables. Deciding how few, however, is not always easy.

The input to either of these processes is a matrix expressing the similarity of every variable to every other variable. This is often the classical product-moment correlation coefficient, but it need not be. The difference between principal
component and factor analysis lies in the underlying model. In the principal components approach it is assumed that the components being extracted can be described entirely by the measured variables with no need to consider other properties that could have been measured but were not. In true factor analysis, allowance is made for the incomplete description of the factors by the actual variables measured. This depends on a variety of assumptions about the data, and the various methods of attempting this feat will not be described further here.

In many cases the number of properties describing the data variation can be drastically reduced by the techniques described above, and it is frequently advantageous (although not obligatory) to perform an operation of this type to aid in comprehension of the data.

Most techniques for grouping samples (Q-mode analysis), rather than variables, require as input a matrix containing similarity measurements between these samples. Various possibilities exist for this similarity measure those like the product-moment correlation coefficient that increase (from $-1$ to $+1$) with increasing similarity, and those like Euclidean distance that decrease to zero with increasing similarity. Preliminary reduction in the number of variables can provide an appreciable saving at this step, as most grouping methods increase in cost approximately with the square of the number of samples.

Given a similarity matrix, there are commonly three approaches used to group the data. The first is Q-mode principal-components or factor analysis. This is similar to the R-mode methods just discussed, except that the factors or components extracted are expected to be associated with clusters of similar rather than variables. Alternatively, depending on the methodology used for rotating the factors, they may be considered to be theoretical end-members of the suite of samples. While Q-mode component or factor analysis is fairly common, it tends to increase in cost very rapidly with the number of samples.

A second technique is discriminant analysis, although this method is intended for somewhat different initial conditions. A set of samples is defined as falling into two or more groups and a set of ‘discriminant functions’ is obtained to distinguish as well as possible between them. Subsequent samples may then be classified into groups using these derived functions. In many mapping applications, however, initial categories and classified samples are not available, necessitating the generation of sample groups from the measured variables alone.

The third common technique is cluster analysis. This term encompasses a wide range of usually non-statistical grouping techniques. Given the similarity matrix described above the next step is to arrange the samples, usually into a hierarchy, so that samples with the highest similarity are placed together. These groups are then associated with the groups that they most closely resemble and so on, until all of the objects have been placed in the classification.

Just as there are many ways of generating a similarity matrix, so also there are many ways of clustering the groups. Most of the common methods of both are described in Wishart (1975).
Jaquet et al. (1975) collected seventy-six samples from the sediments on the floor of Lake Geneva (Figure 8.5A) and analysed them for twenty-nine chemical components. The data were standardized and a product-moment correlation matrix was obtained for each variable against each other variable. Data reduction was
performed using both principal-components and factor analysis techniques to produce four composite variables. These four variables were then used to generate a sample-similarity matrix using Euclidean distance as the measure. Clustering was performed using the ‘unweighted pair-group’ method, and displayed using the dendrograph approach of McCammon (1968). An example of this is given in Figure 8.5b, which was derived from four standardized principal components. The resulting facies map is given in Figure 8.5c.

Having evaluated various combinations of procedures they suggest that standardizing variables, followed by principal-components analysis for reduction in the number of variables, gave the best results with the fewest statistical assumptions. The effects of varying the clustering approach were not examined. In the end, however, it is the interpretability of the resulting map that is of greatest importance, and most of the techniques used produced plausible maps.

Facies mapping by computer suffers from advantages and drawbacks similar to those of computer-assisted geochemical mapping, with the additional problem that costs increase dramatically with large numbers of samples. Again, spatial closeness is rarely taken into account in the grouping of samples. For a small number of samples with many measured variables automated sample grouping is commonly used, but this is done less frequently when a few parameters are measured on many samples. It is primarily a matter of convenience whether computer mapping techniques are used for data display-partly because little work has been done on techniques for generating boundaries between groups of similar samples.

ORIENTATION DATA AND DENSITY MAPS

Orientation data, a fairly common type of geological information, includes properties such as water runoff direction, flow direction based on current ripple marks, direction of ancient glaciation based on existing abrasion marks, pebble orientations, and rock fractures. The orientation may involve a compass direction only, or it may involve a dip component. It may possess merely an orientation, as with pebble long axes, or it may also possess a direction, as with water runoff. Appreciable difficulties may be encountered in manipulating this kind of data-calculating a conventional average is not trivial when 360 degrees is the same as zero! Much of the research into this kind of data has been performed by geologists (for example Ramsden, 1975; Ramsden and Cruden, in press).

Problems arise in the field work or collection of orientation data because it is typically highly variable, and several local observations of pebble orientation, etc., must be made in order to derive a reasonable average direction. The local results are frequently displayed on the map as rose diagrams, these being circular histograms displaying the number of occurrences of an orientation within each of several sectors. Nevertheless, appreciable difficulties arise if it is desired to display flow directions, for example, as a continuous property over the map area. The
mathematical difficulties are described in Agterbeg (1974), and the cartographic problems are not yet fully resolved.

When the orientation information can be considered to be internal to a body of rock it is referred to as a fabric; fabric being defined as the internal geometric properties of the body. These properties are the cumulation of individual structures having geometric orientation-for example planar joints; bedding or cleavage; and linear features such as the long axes of sedimentary particles or the orientation of ripple marks. Each structure type with a distinct geometric configuration is called a fabric element, and the overall fabric is the sum of all these fabric elements.

One object of fabric analysis is to describe the fabric of the geological body under study, and a major aspect of any such study is the need to subdivide the rock body into a series of spatially distinct portions or domains within which the fabric may be regarded as homogeneous. The selection of fabric domains depends on the size of the smallest portion of the body that may be considered a distinct unit—that is, it depends on the scale of the study. The normal procedure for fabric analysis is to collect data as uniformly as possible over the study area; divide the area into domains visually; plot orientation diagrams for each domain; and then combine any domains that appear similar. It is also possible to contour angular dip measurements, and regions of homogeneous fabric are identified as regions of low contour density.

Ramsden (1975) described a computer-based procedure in which the area is divided into sampling units by the geologist, and the data are treated throughout as three-dimensional, eliminating the need to examine horizontal and vertical components separately. Statistical tests are used to indicate domains with large scatter of orientations, and fabric diagrams are automatically produced. These items are primarily aids in separating the study area into meaningful domains, within which the fabric is basically similar, and between which the fabrics vary. Maps may be produced which indicate the orientation of individual readings by a line whose own orientation indicates the orientation of the sample, and whose length indicates the vertical component (Figure 8.6A). Once domains are defined, the vector mean may be displayed for each (Figure 8.6B). Alternatively, the deviation of each domain mean from the regional mean may be displayed (Figure 8.6C).

In the process of identifying domains it is frequently desirable to display the scatter of all the samples as an aid in interpreting fabrics and in locating relatively homogeneous regions. This is usually handled by considering all the samples within some region to be located at a point, and the measured directions or vectors to be projected onto a hemisphere or sphere. The next step is to contour this surface in units of point density and examine the result for various clusterings of orientations, using statistical techniques if possible. The standard approach is to define some counting area and move this over the map, determining how many observations fall within it at each location. Unfortunately, the resulting map varies drastically with changes in the counting area—from a very smooth map if the areas are large to a delta function (zero where there is no sample, unity where there is) as the counting area gets infinitely small.
8.6 Normals to observed bedding orientation: (A) raw data; (B) grouped data; (C) residuals from overall mean (Ramsden, 1975; reproduced by permission of J. Ramsden)
Ramsden (1975) shows that the common methods of density estimation can be generalized to the form

\[ d = \frac{1}{n} \sum_{i=1}^{n} w(\theta_i) \]

where \( n \) is the number of samples, \( i \) is the distance of the sample point from the counting location, and \( w \) is a weighting function. For the constant-area method (where a point is counted if it falls within a circle of radius \( c \)), \( w(\theta_i) = 1/a \) (the area of the counting circle) if \( \theta_i \leq c \), or zero otherwise.

An alternative weighting scheme is to apply, instead of a constant weight \( 1/a \), a weight decreasing regularly with increasing \( \theta_i \). If the Fisher distribution is used, \( w(\theta) = \frac{k}{2 \sinh k} \exp(k \cos \theta) \), where \( k \) is a suitably-chosen concentration parameter. A third method is to adjust the area of the counting circle until a coefficient of variation reaches a prescribed level, and the constant-area equations are then used.

Ramsden examined these methods in detail with respect to the selection of suitable parameters. He concludes that there is no basis for selecting any set of parameters (that affect the degree of smoothing of the surface) if the model (that is, the expected number of clusters, or density peaks) is not known. Given an expected number of peaks, however, some selection of suitable parameters may be made. This is in accord with work done on conventional contour mapping, where it has not been found possible (on the basis of the data alone) to distinguish between data points that are poorly correlated because they only occur at each peak and pit in the topography (that is, at or close to the Nyquist frequency, or sampling limit); poorly correlated data due to the (accurate) measurement of a surface having features of frequencies higher than the Nyquist frequency; or data having large errors in the x, y, or z measurements.

While much of the work on orientation data has been inspired by the availability of the digital computer, there is no overwhelming need to use the computer for map production. Nevertheless, as a matter of convenience most data display will probably be automated, since engineering geology problems can generate reasonable quantities of data, and domain definition tends to be an iterative (and hence repetitive) process.

**GEOLOGICAL AND LITHOLOGICAL MAPS**

The classical ‘geological’ map displays the rock types present in a region, the level of tectonic activity, or some similar attribute. It is concerned with variation in three dimensions, but with sampling limited to two in most cases. They may be of two types: area coverage maps showing which rock types are to be found in any particular area and ‘posting’ maps indicating the values of particular parameters at observed locations. These values may be the rock type observed at each outcrop; the orientation of any faulting; or any of a wide variety of possible parameters. The
8.7 Geological field data sheet (Berner et al., 1975; reproduced by permission of Geological Survey of Canada from Paper 74-63)
normal procedure is for field forms (such as that of Figure 8.7) to be processed, and postings made of interesting parameters. On this basis the geologist will construct his area coverage map. It should be noted that his boundaries usually indicate more than merely some mean distance between samples of different types—the form of the boundaries frequently indicates the structure, such as faulting or folding interpreted by the geologist to explain the observed information. For this reason it is, and will probably remain, difficult to produce good final geology maps without considerable manual intervention.

Geological rock types may well be sampled in three dimensions if it is of economic interest to do so. Oil company well logs are a prime example. In this case the three-dimensional contacts must be defined, then displayed in two dimensions. Contour maps and cross-sections are frequently used; the one to display the overall behaviour of a single contact, the other to display the behaviour in a limited direction of all of the geological surfaces of interest. The two-dimensional limitations of a map sheet are a great inconvenience in this type of work, and three-dimensional wood or foam models are sometimes used. Many challenging problems remain in displaying this type of data.

Even within the realm of the most traditional of map types—the lithology map—much development work is possible. Two recent examples are the work by Bouille (1976, 1977) and Burns (1975). Both are concerned with the systematic description of the relationships between delineated zones on a map sheet. These relationships may then be manipulated by computer.

The work of Burns concerns the fact that geological events occur in a time sequence, and any historical understanding of the geological processes requires the derivation of this sequence. Unfortunately it is usually only possible to determine the age relations of rock types where pairs of them meet. Figure 8.8A shows a schematic geological map of an imaginary area, possessing lithology types A to L. Arrows indicate contacts between lithology pairs at which relative age determinations may be made on the basis of superimposition, intersection, etc. The arrows point from older to younger rocks. Figure 8.8B indicates the processing of this information. The upper left diagram records the younger member of each lithology pair in the relevant cell. The upper right diagram shows the same matrix reorganized by switching rows and columns until older events are to the left of or above younger events. The lower left diagram shows the result when all other relationships deducible from the previous set are filled in (that is, all columns above any particular entry are filled with the entry value). Finally, the lower right diagram indicates the deduced sequence of events, starting with LDFH and finishing with JCBG. Due to the incompleteness of the final matrix the relations of K and AE to each other are unknown. Reference back to Figure 8A indicates a discontinuity or fault in the top centre. This clearly occurred after deposition of units K, H, F, and D which are dislocated, and before the occurrence of J and A, which are not. This information is sufficient to date K as prior to A, and hence the correct sequence of events has been deduced from the geological map.
8.8 Sequence of geological events: (a) rock types with older/younger relationships across boundaries; (b) derivation of event sequence (Burns, 1975; reproduced by permission of Plenum Publishing Corp.)
This procedure of Burns is a somewhat more elaborate version of what, in computing science, is called a topological sort, for which a simple algorithm is given in Knuth (1968). This approach assumes that any legitimate sequence is satisfactory, as is the situation of the top right of Figure 8b, where sequence K and AE could legitimately be interchanged. The concept of a topological sort is an extremely valuable one in many spheres of computer-assisted cartography. A good example is whenever a series of objects or polygons need to be ordered ‘front to back’ for perspective block diagrams or other applications, and only the relative positions between adjacent objects are known. A development of possible interest is the work of Gold and Maydell (1978) in which a region is defined as a set of triangles, possibly with nodes representing the locations of objects, or else with polygons defined by one or more triangles. A simple algorithm is given to achieve a front-to-back ordering from any direction for any triangulation, merely by starting at the foremost triangle and processing neighbouring triangles in turn. The method also works for a radial expansion out from any specified point.

The work by Bouille on geological maps is also concerned with adjacency relationships, but is based heavily on the availability of SIMULA 67, a very high-level programming language which permits extremely flexible data structures. He has developed an extension to this, HBDS (Hypergraph Based Data Structure), to express the graph-theoretic relationships between geological entities that are frequently displayed as geological maps.

One example of this involves a very similar problem to that of Burns in that a stratigraphic summary is derived from the map, but in this instance the initial problem was to digitize a lithology map so as to facilitate subsequent computer operations. First of all, the nodes (junctions of three or more lines) are digitized and numbered, and then the arcs are digitized between nodes. Associated with the arcs must be the two node numbers and the left and right stratigraphic units. From this a graph (in the mathematical sense) is constructed. A dual of the graph is also constructed. If the graph is thought of as a set of countries with irregular (digitized) borders, the dual of a graph may be visualized as a road network connecting the capitals of each country (located anywhere within its borders) to the capitals of each adjacent country.

This dual graph expresses all the neighbourhood relationships between the different stratigraphic units (or, more correctly, the many possible separate areas of each stratigraphic unit or rock type). From this a stratigraphic summary graph (Figure 8.9A) may be derived, retaining one linkage between each pair of stratigraphic units which were linked on the dual graph. From the summary graph various properties may be observed:

- an arc between a vertex \(i\) and a vertex \(i - 1\) shows a normal stratigraphic situation;
- an arc between \(i\) and \(j\) (\(j \neq i + 1 \) or \(i = 1\)) implies an unconformity;
- on the contrary, a lack of an arc between \(i\) and \(i - 1\) does not imply an
unconformity, but shows that connection between \( i \) and \( i - 1 \) is nowhere observable on the map; and

lack of arc between a vertex \( i \) and any other tops indicates that in the present situation, the layer \( 'i' \) is known only from drilling, (Bouille 1976, p. 380).

Other work by Bouille is more generalized, his HBDS (Hypergraph Based Data Structure) language being based on set theory and the hypergraph.

It deals with four basic abstract data types which are named \textit{CLASS}, \textit{OBJECT}, \textit{ATTRIBUTE} and \textit{LINK}. They respectively represent: set, element, property, and relation. But, these fundamental concepts of the set theory are always available and we then may define and manipulate sets of classes, sets of relations, relations between sets of classes, etc. (Bouille, 1978, p. 2).

This system has been used for a wide variety of problems, primarily of a cartographic nature, since the early work on geological maps described above. Examples include hydrography and hydrographic networks, stream systems, road networks, administrative divisions, and urban development. Two examples only will be given briefly here. Figure 8.9B shows a possible data structure for a simple topographic surface. Various classes have been defined: map, contour level, curve, summit, and reference point. The diagram shows the relations between elements as curved lines. Of special interest are the heavy lines. These form a skeleton to the whole data structure, even in the absence of any data, as they show the relations between classes. With a data structure of this type a wide variety of otherwise difficult questions may clearly be asked.

A final example of the flexibility of the concept is the draughting of a perspective drawing of a map composed of contour lines and a road network:

The method consists of: starting from the highest level and considering the set of its corresponding isolines, drawing them, drawing the parts of the segments which are included in these curves; then we consider a lower level, which is drawn taking the windowing into account; then the parts of the segments intersecting these curves or included between them and their possible upper curves, are also drawn in the same manner; we thus do likewise with the lower level until the last one . . . (Bouille, 1978, p. 21).

Perhaps the most amazing aspect of this example is that a link between two classes need not be a pointer, but may be an algorithm-an intersection test in this case! With this approach the handling of cartographic problems may not always be efficient, but is almost limitless. While the handling of lithology maps by computer is
8.9  (A) Stratigraphic summary graph of relationships between layer types 1–10 (Bouillé, 1976; reproduced by permission of Plenum Publishing Corp.); (B) data structure representing some topographic features (Bouillé, 1977; reproduced by permission of President and Fellows of Harvard University)
still in its infancy, valuable work has already been done in defining some of the ground rules.

TOPOGRAPHIC MAPS

Contour maps are widely used to represent surficial topography. Geologists often need to proceed beyond this and map buried topographic surfaces. These may well have once represented the earth-air contact (or, more commonly, the earth-water contact) but they have been buried and frequently deformed by folding or faulting. Very large sums of money are spent each year by oil companies in collecting and processing this type of data. Consequently computer contouring techniques have been developed by many individuals, and many maps of this type have been and are being produced with varying degrees of success. It should be noted that certain differences exist between subsurface and surface geological mapping because of the differences in the properties of the data used. The properties of a data set for topographic modelling can be described (Gold, 1977) under the heading of sample content, sampling adequacy and sample isotropy. Under sample content would be considered the sample’s x-y location information as well as any elevation or slope information associated with that location, and error estimates associated with each of these. Photogrammetric methods will not generally produce slope information, whereas some borehole techniques will do so. The x, y, and elevation values of a borehole data point will usually also be relatively precise.

The biggest differences between these two data types are noticed when considering sampling adequacy. Photogrammetric methods generally produce from tens to hundreds of thousands of data points per stereo model, whereas, due to the costs, a few tens of sample points must suffice to define a buried surface delineated by drilling. Sample point density must be adequate to resolve features that are of sufficient size to be considered important. Current topographic surfaces often pose no problem, but frequently in subsurface work the sample frequency is less than desired.

The third data property, sample isotropy, has rarely been considered, but is an important technical consideration in many practical mapping problems. It concerns data point distribution over the map sheet. Davis (1973) classifies data point distributions into regular, random, and clustered. He describes various tests used to examine data point distributions, primarily using techniques either of sub-area analysis or nearest-neighbour analysis. The results of some techniques, for example trend surface analysis, are heavily dependent on the uniformity of the data distribution.

Data point isotropy, however, is more concerned with the variation of data point density in differing directions. For most randomly collected data sets the anisotropy is not marked and can be ignored. What is often forgotten by theoreticians is that a very large quantity of data, particularly that acquired by automated means, is collected in strings or traverses, where the sample spacing along the traverse is much...
smaller than the distance between traverses. Many very important data types fall into this category, including seismic profiling, ships' soundings, and airborne surveys of many types (magnetic, gravimetric, radiometric, etc). An interesting data type at the research level consists of the input of contour lines digitized from topographic or other maps. A topographic map may only be considered a general topographic model if its data structure (contour strings) may be converted to some other form (for example, regular grids of elevations) for further study. In addition, many non-automated geological data collection techniques consist of field traverses—along roads, streams, ridges, or valleys.

Three aspects of this discussion are worthy of further note: the concepts of a topographic model; of data collection along relevant features; and any special processing requirements of traverse data.

The Topographic Model

In computer terms at least, a topographic model should be distinguished from a topographic map. As with a balsa-wood or Styrofoam model, it should be viewable in many ways—from any orientation, by slicing it, etc. A topographic contour map is merely one way of displaying the topographic model. The primary requirement for any display of the model—ontour map, cross-section, block diagram, etc.—is that the model may be interrogated to obtain the elevation at any x-y location, and that this value should be obtained in some reasonably efficient manner. Since there will not usually be a data point precisely at each desired location, a topographic model should be defined as a set of data points plus the required algorithms to obtain any requested elevation. Most modelling or contouring algorithms assume that the data point location has no intrinsic meaning. This is often not the case. Where the surface is visible in advance (for example, current topography, but definitely not buried topography), some data points at least would be selected to occur at peaks, pits, ridges, or valleys. In many cases, samples would be selected along a ‘traverse’ of a ridge or valley. Peucker (1972) calls these items ‘surface-specific lines or points’. While manual mapping methods may take breaks in slope into account at these locations, this is difficult with automated techniques.

Because of the widely discrepant distances between data points for traverses (which need not be perpendicular to the x or y axes) some modelling methods may break down. This is typically true of interpolation techniques that perform some form of weighted average on the ‘nearest’ few data points they can find in order to evaluate the elevation at some unknown point. Obviously, having to search the whole data set to find the nearest few points to each location on a grid, for example, may be time-consuming. In addition, the nearest few points will often be obtained from one traverse exclusively, giving no weighting to values from adjacent traverses that would provide information on the behaviour in the second dimension. Many programs acquire ‘neighbouring’ data points until values are obtained in at least six of eight sectors of a circle around the unknown point. This is an attempt to
compensate for the anisotropic distribution of sample locations around the point of interest, but for traverse data this requires much computational effort and frequently little improvement is observed.

The problem arises because a simple metric distance is a poor measure of neighbourliness, especially in anisotropic cases. A non-metric definition is needed. This may be achieved by triangulation techniques (Gold, 1977; Gold et al., 1977; Males, 1977) whereby some ‘best’ set of triangles is defined to cover the map area, with all vertices at data points. Neighbours may thus be found in a fashion independent of the relative data point spacings and without the necessity of performing a search for neighbours at every grid point.

Topographic mapping is used by many disciplines; what is different in the case of geology? Our use of conventional topographic survey maps is much the same as anyone’s, except perhaps for a greater need to extract specific elevation values from the map when determining sample locations. Isopach, or thickness, maps are common; these may be considered as topographic maps of the top of the geological unit, the base of the unit being taken as zero.

A wide variety of contouring programmes have doubtless been used by geologists to construct isopach maps. Most of these programmes require the generation, from the irregularly spaced data, of a regular grid of estimated values. The most common procedure would be to extract the thickness information directly from each sampling or drill hole location by subtracting the elevation of the bottom contact of the geological unit from the elevation of the top contact. The thickness values at each data point are then contoured by estimating the values at each node on a grid, and interpolating contour lines within each grid square.

While this approach produces a reasonable map where the geological unit is continuous, shortcomings become apparent when the stratum is absent in places. In particular, the zero thickness contour line is frequently implausible. This is the result of contouring a computer-generated grid with zero values in regions where the unit is absent, and positive values elsewhere. The interpolated zero contour will therefore follow the grid square edges.

It is clear that the thickness ‘model’ is inadequate and that zero thickness is not necessarily an adequate description of the absence of a geological unit in a particular location, specially if the absence is due to the erosion of pre-existing material. It is therefore more correct to consider an isopach map as a map of the difference in elevation between two complete topographic models, one of the upper contact and the other of the lower contact of the geological unit. With this approach, negative values are legitimate and problems with the zero isopach line disappear.

An interesting example of this type of mapping concerns a coal resource evaluation project in the Foothills of the Canadian Rockies (Gold and Kilby, 1978).

Goal resource evaluation clearly requires information on the mineable tonnages, as well as the-grade, of the coal strata. In open-pit mining the volume of coal economically extracted from a seam is directly related to the amount of overburden that must be removed in order to expose the coal. This may very conveniently be
expressed as an ‘overburden ratio’ map, being a contour map of the ratio of the overburden thickness divided by the coal seam thickness at the same location. While the cutoff value for economic recovery varies due to many factors, it is relatively uncommon for coal with a ratio of greater than 10 to be economically mineable. The overburden ratio map may thus be of value in coal reserve estimation.

The construction of an overburden ratio map requires three components: a topographic model, a model of the top of the coal seam, and a model of its base. In this study the coal seam was fairly consistent in its thickness (of about 30 ft) and therefore no model of the base of the coal seam was used for thickness estimation.

Once the topographic surface has been modelled, the same must be done for the top contact of the coal seam. This poses particular difficulties in that most of the data is obtained from outcrops and drill holes along one line—the intersection of the coal seam with the topographic surface. It is therefore necessary to estimate the geological structure (that is, folding) of the originally flat-lying sedimentary rocks, so as to permit projection of the seam perpendicular to the ‘trace’ of the coal. As in the topographic model, suitable mapping software and computer resources must be available. In addition, geological expertise is necessary to evaluate the geological structures, on the basis of the work of Charlesworth et al. (1976) and Langeberg et al. (1977), who discussed the criteria necessary in order to assume the presence of cylindrical folding within a domain—that is, under certain mathematical conditions, a particular portion of the coal seam may be satisfactorily described by a type of folding similar to a sheet of corrugated iron, linear in one direction and undulating perpendicular to this. If the orientation of the desired geological contact has been observed at several locations in the field, this linear direction, called the fold axis, may be determined by the mathematical techniques described in the previously mentioned references. A cross-section may then be constructed normal to this fold axis and all data points projected on to this ‘profile’. The irregular folding of the coal seam in a portion of the study area is shown in the profile of Figure 8.10.

In the absence of satisfactory orientation data it is possible to estimate a suitable fold axis in a trial-and-error fashion by examining the profiles generated for each estimate.

If a suitable profile can be generated for each domain, it is a suitable description of the geological contact. However, it is described at a different orientation in space from the original map, being perpendicular to the estimated fold axis. Dummy data points may then be generated on the map and rotated into the coordinate system of the profile. Their elevations may then be estimated from the profile, and the resulting points transformed back into the map coordinate system. These values may then be contoured by the methods previously described for the topographic surface (Figure 8.10b).

The final step in the procedure consists of generating the overburden ratio map from the two distinct surface models. The procedure is straightforward: the elevations of the two surfaces must be evaluated at a series of map locations (perhaps on a grid): the elevation of the top of the coal seam subtracted from the topographic
8.10 (a) Profile looking down fold axis; (b) Intersection of geological contact with topographic surface (existing contact indicated by shading); (c) Overburden ratio map (heavy line indicates faulting; shading indicates mineable coal seam (Gold and Kilby, 1978) )
surface; the resulting difference divided by the seam thickness; and a contour map produced of the result (Figure 8.10 c). These results may legitimately be negative where the coal has been eroded out, although normally only positive values would be contoured. On the basis of the local cutoff point for the overburden ratio, tonnages may readily be determined, either manually from the map or by further computer processing.

Two important conclusions may be derived from this example, and in the opinion of this writer they hold true for all map types—whether topographic or thematic. The first is that a model consists of an attempt to generate a space-covering map of a single parameter only. It should conform as closely as possible to the original data and avoid synthetic devices such as grids. The second point is that grids or rasters should be display devices only, to be used as needed to compare and display models. Ideally they should not be preserved or used again to make subsequent comparisons. The large effort expended to create good polygon-overlay procedures and to handle the resulting multitude of small residual polygons may, in some applications, have been better spent on generating good techniques of simultaneously scanning several separate polygon overlays and generating a final raster-type display.

The use of similar techniques for topographic modelling as well as for lithology maps introduces the final example in this section. It concerns a topic very close to the heart of many petroleum geologists—how well can a computer, as compared with a seasoned geologist, estimate the structure of a surface (a buried geological contact)? Dahlberg (1975) constructed an imaginary coral reef with several pinnacles, took random samples of from 20 to 100 elevations from the model and submitted them to two geologists and one computer (using a triangulation-based contouring algorithm). He showed that the computer-based model was always more precise, especially so with the fewest control points; so much for professional judgement! However, when additional geological information was available—such as general regional trends or depositional environment—the geologist’s interpretation of many data sets improved dramatically.

**COMPUTER-BASED DATA HANDLING**

**Data Capture**

The following list gives the collection technique and estimated cost (in U.S.$) of each data point collected from a variety of surveys of interest to geologists:

- Airborne magnetic ($2), ground gravity ($15–$20), marine seismic ($35), land seismic ($250), magneto-telluric ($800–$1000), shallow well ($50,000–$100,000), medium well ($250,000–$500,000), deep well ($1,000,000–$5,000,000). Clearly the quantity of information acquired for each technique varies dramatically. Nevertheless there are still a lot of deep drill holes in existence, since they are clearly of great commercial value. It is perhaps more useful to speak of the density of
observations obtained as a function of the smallest feature of interest. In most petroleum geology work the relative density of seismic results is much greater than that for drill holes. Each drill hole, on the other hand, provides a large number of different items of information. Data storage and retrieval techniques will consequently differ.

The aspect of data collection that has received the greatest published attention is the automation of geological observation in the field. This has traditionally involved the devising of suitable coding forms. The great problem with this is that, to date, computer systems are not nearly as flexible as field geologists whose training is to note any remarkable feature contributing to the overall aim of the study. By restricting him to any standard set of features to be described or identified, one is running directly counter to his training and philosophy—always a bad thing in systems design—as well as drastically reducing his real value in terms of relevant observation. Conversely, a check list is a useful memory aid, especially for subordinate-level field staff, although, like fire, it is a bad master even if a useful servant. The discipline of true observation, the persistence to truly look for clues to the problem in hand for outcrop after outcrop under frequent conditions of discomfort, is a hard thing to acquire with the best of incentives. It is nearly impossible when the view is dominated by a multiple choice questionnaire, the completion of which requires method but little imagination and whose final line irresistibly suggests that all needed information has been acquired. It should also be noted that a field survey is always problem-specific, and since a geologist would view the same piece of terrain very differently depending on the objective, a general-purpose coding form would probably be horrendously elaborate or hopelessly inadequate.

Nevertheless, much work has successfully been done using this approach. The greatest success has probably been achieved where a project has continued for several years with the same, or continually improving, coding forms and objectives (typically merely extending the area of coverage)—and where heavy dependence is put on seasonal or partly trained field staff. Other suitable applications are where the possible field observations are in any case limited—for example in drill-hole work, whether in shallow urban engineering-geology applications or deep, complex oil exploration. In conjunction with the availability of more sophisticated data base management systems, input of free text for each observation station has become more common, but some basic problems still remain in the editing and identification of relevant key words and phrases. Hutchison and Roddick (1975) conclude that such an approach requires less time on outcrop, loses less information due to illegible handwriting, and improves the consistency of recording information. The main disadvantages are the time lost in editing input data and in file and data base management. Garrett (1974) gives a good description of field coding forms for geochemical surveys, as used by the Geological Survey of Canada. These forms vary depending on the type of sample (soil, water, etc.) being collected. Clearly the requirements of such a survey type are simpler than those for general field surveys,
but even here the bulk of the coding form is left available for comments or parameters defined by the project leader.

Another excellent example of the effort required to develop a comprehensive field data system, even for specific objectives, is found in the LEDA system (David and Lebuis, 1976) for Quaternary geology. Six computer-based forms are defined, each convertible to one 80-column punch card (Figure 8.11 A). One additional form is reserved for sketches and photographs. Form 1 describes the sampling location, neighbourhood, and other relevant material. Form 2 describes a single stratigraphic unit at that location, including thickness, structure, texture, etc. Form 3 describes the composition of any stratigraphic unit or other geological entity, and measured parameters that may include fossil types, rock types, mineral types, etc. Form 4 (with the same layout) describes any orientation information, and Form 5 is designed to accept comments. Form 6 is used to record detailed measurements of the thickness of subunits occurring within a composite stratigraphic unit. These forms may be used in various legitimate combinations to describe any geological section, using a carefully selected and open-ended series of codes for each position on the form. Figure 8.11 B shows, without need for further description, how the predefined form types may be used to describe an extremely complex geological section.

Data Base Management

Computer-based storage and retrieval of geological information has a long history (the earliest paper is by Margaret Parker, in 1946, on the use of punched cards in tabulating drilling data). Bibliographies up to 1972 include 447 references (Hruska and Burk, 1971; Burk, 1973). Some of the major geological topics covered include geochemistry (58), geological field mapping (47), geophysics (24), hydrogeology (19), hydrology (15), geophysical logs (19), mineral and fuel deposits (62), mineralogy (13), oceanography (13), palaeontology (47), petrology and lithology (46), stratigraphy (27), and petroleum and gas wells (61). The computer topics discussed include bibliographic files (91), bibliographies (10), data codes and coding (121), computer systems (110), curating (21), data display (117), data files (286), data recording (160), digitizing (151), indexes and indexing (69), information centres (36), national and international systems (53), data standards (48), surveys and questionnaires (14), theoretical topics (95), and thesauri (36). While much other work has been done, clearly a major effort has been made in the handling of computer files of geological data acquired in the field and intended for some form of geographical display.

The phrases ‘data base’ and ‘data base management’ are applicable to the handling, however primitive, of any collection of data, however small. However, in the computer business they imply the relatively sophisticated manipulation of relatively large data sets. Nevertheless, the types of manipulation can be described under the headings of collection, storage, updating, rearrangement, retrieval, display, and analysis. The first of these has already been described; the next three items concern the construction of the data base, and the last three concern its examination.
8.11 LEDA stratigraphic coding system: (A) field data forms; (B) application to a complex
geological section (David and Lebuis, 1976; reproduced by permission of Pergamon Press Ltd)
Display and analysis are clearly subsequent to and dependent on retrieval of the relevant subset of the whole data set (although the form of the retrieval may be affected by the intended use), and some of the analysis or display capabilities may already be part of the data base management system. We are thus left with problems of storage and problems of retrieval. De Heer and Bie (1975) describe some of their requirements for a suitable data system. It should accept input in the form ‘attribute = value’ since the skills required to convert ordinary field information into fixed length codes is rarely available. It should allow hierarchical relationships between data elements. It should allow retrieval by Boolean expressions (for example, if clay is less than 10 per cent). Finally, the data base itself should recognize four main types of data sets (points, lines, areas, and descriptonal data sets independent of geographical coordinates) and any relationships between them. Many industrial systems cannot handle these requirements. Many, geared to textual data, are unable to retrieve all data elements whose values of a particular attribute fall within a numeric range. Even more commonly, the relationships between points, lines, and areas on a map are not readily definable using only the hierarchical relationships available with most commercial systems, since these particular relationships are best expressed as ring structures.

Other required operations are more readily available with many systems. These include the storage, updating, and rearrangement (editing, sorting, merging, subfile creation, etc.) categories mentioned earlier. Although special difficulties may arise in particular applications, these more properly belong in the domain of computing science and will not be described further here. It is this writer’s opinion, however, that until spatial coordinates and relationships are recognized as separate and distinct from other data types, and are stored and treated according to their own rules, truly meaningful and conceptually elegant geological data base management systems will not be achieved. It is not, for example, easily possible to extract all water samples collected within a particular drainage basin unless such a relationship was specifically defined when the water sample data were entered. Nevertheless, such an operation is easily done by humans using a map, and questions of this form are basic to geological thinking. Hutchison (1975, p. 5) states that:

statistical and other analyses . . . can be fed back, as part of an iteration, to allow evaluation of domain boundaries. In this way data [points] may be given geological significance and as domain boundaries are changed, then so must the geological context of all contained data. Many geologists do this every day and it is essential that any computer system for field geology must have the same capacity.

The previously mentioned work on triangulations of a plane is a tentative step in this direction.

Considering the relationship between points, lines, and areas, Hutchison again
comments that geological field data are frequently considered as points, as if they had been acquired by a blind machine. He continues

The geologist does not stop at one point A, look down at his feet and record all significant data within a radius of three metres of his feet, then close his eyes and walk blindly onwards for a hundred metres or one kilometre and then stop, look down, open his eyes at point B and record all data within a radius of three metres. Instead he is more concerned in the first instance in establishing the relationship between data set A and data set B, and recognizing (using his own inboard computer) whether or not there is a difference, and if so, what its nature might be. Each hour and each day in the field is spent working essentially interactively with the rock patterns to build up a picture of the field setting. . . . I would contend . . . that there has been a general failure to recognize that geological contacts are prime data located on geological field maps . . . whose nature results from observations between data set A and data set B. On this basis, ‘spot’ data within units are therefore of lower rank than the data for contacts between units (Hutchison, 1975, p. 3).

So much for the spatial aspect of the data set. Sutterlin and May contend that the slow acceptance of data base management systems by geologists is because:

They are accustomed to dealing with naturally occurring objects and phenomena which are quite complex and only partially understood. As a result, the data about these objects and phenomena are varied and complexly interrelate&that is, complexly structure&and the earlier techniques and concepts designed for more predictable and less highly structured data of the commercial environment were less than completely successful when applied to the management of geological data (Sutterlin and May, 1978, p. 30).

They comment that modern computing advances have largely eliminated the need for data coding, although it is often used in the field to obtain a measure of uniformity and completeness at the data collection stage. While much work has been done on geological data structures, little of it is directly relevant to mapping technology exclusively. Further information may be found in Sutterlin et al. (1974) and Dickie and Williams (1975).

CONCLUSIONS

We have discussed some of the major types of geological map and how various workers are attempting to automate different procedures. We can summarize by saying that contour type map production by computer is alive, well, and hopefully
moving towards second-generation techniques. These will probably involve taking
the topology or neighbourhood relationships of data points into account. Data
transformation (often statistical) prior to display is fairly widely used, but again
spatial relationships are not normally considered. In lithology mapping, topology is
being seriously examined by a few workers, but automation is rare beyond selective
data point posting. This last, along with the necessary field data acquisition and
management, is progressing steadily.

Where contouring is not involved, the use or non-use of computers for map
production seems mainly to depend on whether automated statistical or other data
transformations have already been performed. If so, or if field coding forms were
used, automated posting of data values will probably be used. In geology, computer-
generated point, line, and area symbolism has not been extensively examined,
although some consideration has been given to the problems of the computer-
assisted drawing of boundaries between groups of similar data points. A particularly
bright spot is the recent concern over the topological relationships between
neighbouring geological entities (such as rock units) on a map. Nevertheless, with
this exception (and that of contouring) computer usage in geological map production
is primarily in data transformation rather than display.

REFERENCES

Bouillé, F. (1976). ‘Graph theory and digitization of geological maps’, Math. Geol., 8,
375-393.
Bouillé, F. (1977). ‘Structuring cartographic data and spatial processes with the Hypergraph-
Based Data Structure. In G. Dutton (ed.), First International Advanced Study Symposium
on Topological Data Structures for Geographic Information Systems, vol. 5. Harvard
Laboratory for Computer Graphics and Spatial Analysis.
Burk, C. F., Jr. (1973). ‘Computer-based storage and retrieval of geoscience information:
planes and sections of macroscopic folds using computer-based methods. Can. J. Earth
Sci., 13, 54-65.
files of geological field data’, Computers and Geosciences, 1, 265-278.


Males, R. M. (1977). ‘ADAPT—A spatial data structure for use with planning and design


