74-60

GEOLOGICAL SURVEY OF CANADA

DEPARTMENT OF ENERGY, MINES AND RESOURCES, OTTAWA

PAPER 74-60

COMPUTER USE IN PROJECTS OF THE GEOLOGICAL SURVEY OF CANADA

TERRY GORDON
W. W. HUTCHISON

1974

Energy, Mines and Resources Canada — Énergie, Mines et Ressources Canada

GEOLOGICAL SURVEY
PAPER 74-60

COMPUTER USE IN PROJECTS OF THE GEOLOGICAL SURVEY OF CANADA

TERRY GORDON
W. W. HUTCHISON

1974

Available by mail from *Information Canada*, Ottawa

from the Geological Survey of Canada
601 Booth St., Ottawa

and

Information Canada bookshops in

HALIFAX — 1683 Barrington Street
MONTREAL — 640 St. Catherine Street W.
OTTAWA — 171 Slater Street
TORONTO — 221 Yonge Street
WINNIPEG — 393 Portage Avenue
VANCOUVER — 800 Granville Street

or through your bookseller

A deposit copy of this publication is also available
for reference in public libraries across Canada

Price: $3.00 Catalogue No. M44-74-60

Price subject to change without notice

Information Canada
Ottawa
1974

CONTENTS

COMPUTER USE IN PROJECTS OF THE GEOLOGICAL SURVEY OF CANADA

INTRODUCTION

This publication brings together a sample of current activities of the Geological Survey of Canada in which computer use plays an important role. It is somewhat distorted presentation of Survey projects in that the focus in computers - essentially merely tools - rather than project objectives. It is also incomplete in that some important computer applications are discussed only briefly or not at all. Nevertheless this publication should serve the important dual purposes of (1) providing in one place a reference to computer users outside the Geological Survey who wish to maintain contact with applications and users in our organization, and (2) providing a basis for Branch planning for future development and acquisition of computer hardware and software.

Computer use at the Geological Survey is receiving particular attention at this time for several reasons. (a) Third generation computers and their associated software now provide powerful data storage and retrieval, text manipulation, and time-sharing graphics capabilities - essential to any scientific program requiring interpretation of, and instantaneous interaction with, large volumes of data. (b) There is a growing demand for release of primary machine-processable data to other government organizations and the public. Such requests are often based on the naive assumption that if data can be "read" by a computer they are immediately usable by people. In fact, although providing the computer technology required for the transmission of the data is itself a difficult task, far more complex and poorly understood is the problem of transmitting the meaning, accuracy, precision and logical links between the various data items in a fine. Nevertheless, the problem must be faced. If solutions are to be found they must be based on an intimate knowledge of current practices of data capture, manipulation and display. (c) Computers, large and small, can now provide excellent process control for operations varying from chemical analytical procedures to production of coloured geological maps. The tremendous savings in time and dollars provided by this technology must be carefully weighed against the high capital costs and problems of software development associated with computer use.

The papers contained in this publication will provide the reader with an insight into some of these problems and practices. Like all publications in the rapidly changing field of computer technology, its useful lifetime will be short. We hope, however, that it will serve as a point of reference for the future development of computer use in the Geological Survey.

Cette publication a pour but de donner une idée des activités courantes de la Commission géologique du Canada où l'ordinateur joue un rôle important. C'est une présentation un peu modifiée des projets de la Commission puisque l'on insiste sur l'utilisation des ordinateurs, qui sont essentiellement des outils, plutôt que sur les objectifs des projets. Elle est également incomplète puisque certaines applications importantes de l'ordinateur sont décrites très brièvement et parfois même non mentionnées. Néanmoins, la présente publication vise deux buts importants: premièrement, servir de référence aux usagers d'ordinateur qui ne font pas partie de la Commission géologique et qui désirent garder contact avec les applications et les usagers de notre organisation; deuxièmement, fournir à la Direction une base pour planifier les améliorations et acquisitions futures de matériel et de programmes.

On porte en ce moment une attention particulière à l'utilisation de l'ordinateur, à la Commission géologique, pour plusieurs raisons. (a) Les ordinateurs de troisième génération et la programmation qui s'y rattache offrent maintenant des services très efficaces d'enregistrement et de localisation de données, de traitement de textes et de graphiques en temps partagé, éléments essentiels à tout programme scientifique nécessitant l'interprétation et l'interaction instantanée de grandes quantités de données. (b) Les autres organisations gouvernementales et le grand public demandent de plusen plus de données de base qui peuvent être traitées. De telles demandes s'appuient souvent sur la supposition naive que puisque des données peuvent être "lues" par un ordinateur, elles peuvent être utilisées immédiatement par le public. En fait, bien qu'il soit déjà difficile de fournir à l'ordinateur la technologie nécessaire à la transmission des données, il est encore beaucoup plus complexe, et ce fait est mal compris, de transmettre la précision, la signification et l'exactitude des données d'un dossier ainsi que leur interrelation. Néanmoins, il faut envisager ce problème. Si des solutions sont retenues, elles devront s'appuyer sur la connaissance exacte des méthodes courantes de rassemblement, de traitement et d'affichage des données. (c) Les ordinateurs, petits et gros, peuvent maintenant effectuer de façon très précise des opérations qui vont de l'analyse chimique au tracé de cartes géologiques en couleur. Il faut bien soupeser, d'une part les très grandes économies de temps et d'argent que permet cette technologie et, d'autre part, les investissements importants et les problèmes techniques qu'entraîne l'utilisation de l'ordinateur.

Le texte de la présente publication donnera au lecteur une idée de certains de ces problèmes et procédés. Comme toutes les publications qui touchent le domaine de l'informatique où tout change rapidement, celle-ci sera vite périmée. Toutefois, nous espérons qu'elle servira de point de référence pour des modifications futures dans l'utilisation de l'informatique à la Commission géologique.

TMG and WWH
August/Août, 1974

1. TOWARDS COMPUTER-BASED 'SYSTEMS' FOR RECORDING, STORAGE, RETRIEVAL, ANALYSIS AND PRESENTATION OF GEOLOGICAL FIELD DATA IN THE GEOLOGICAL SURVEY OF CANADA

W. W. Hutchison,
Computer Systems Group

Introduction

In May 1971 I was asked by the then-Director of the Geological Survey of Canada, Dr. Y.O. Fortier, to initiate an investigation of how computer-based systems* might be applied to the collection, storage, retrieval, manipulation of geological field data. At that time, Terry Gordon and I decided that instead of prejudging the situation and putting all the emphasis on the development of one, unique system, we should instead initiate a series of pilot projects, each of which would lead to an investigation of the collection of data from different geological terrains and whose projects also dealt with different geological problems.

Evaluation of Initial Results of Pilot Projects

It became apparent that certain projects were much more amenable to computer-based systems for recording the geological field data than were others. The criteria for and against are listed below. Types of geological settings in which computer-based systems could be used were:

1. Those in which there were large field parties with a number of geologists.
2. Where the terrain did not have a well-defined, sufficiently continuous stratigraphy.
3. Where many measurements were being taken repetitively and where such measurements would have to be plotted in various forms at a later time.
4. When a project was going to last at least two years.
5. When there was some knowledge of the nature of the terrain to be mapped.

The major problems encountered by those who did not proceed further with the techniques were as follows:

A. Bottle necks develop in key punching at the end of the field season.
B. Much time was spent in pre- and post-keypunch editing of data.
C. Field data did not appear to lend themselves easily to computerization.

As a result of these initial investigations, we become very quickly convinced of the following:

(a) That the geologist himself decides whether or not a computer-based system allows him to do his job more effectively or fits more naturally into his work style.

*The term "system" is here used loosely to include the combination of file design and built-in processing programs.

(b) That rather than establishing one system we were probably faced with establishing a number of systems.

(c) That even if we did develop essential computer-supported facilities for handling geological field data, there still did not seem to be the expertise to handle the cartographic analysis of this data and information. Furthermore, we were not tooled-up to handle these data and information in terms of map production. Because of this, it was decided that we should concentrate on the flow of information from the collection of data in the field to the production of final product.

In Figure 1, we demonstrate what is meant by the above statement. If we look at the field collection stage on the left-side of the figure, it can be seen that we are attempting to devise systems which would assist each individual project leader obtain his assigned objective as expediently and effectively as possible. Accordingly A^o is the starting point and A' is the obtained, desired objective.

In Figure 2, there are arrows pointing upward toward the "Data Base". This is to emphasize that whereas we were willing to investigate and address ourselves to the task of producing computer-assisted production line systems, we were <u>not</u> concerning ourselves, at that time, with the development of archival files to be used as part of the Branch Data Base Management system.

At this point we decided that we had to obtain support in two areas: (1) office compilation and cartographic analysis, and (2) final stage production. Accordingly a plan was developed which was contingent upon the establishment by the Geological Survey of an autocartographic system to be 'plugged in' to the already established facility at Surveys and Mapping Branch. The plan that we set out is sketched out in Figure 3.

The flow of data (Fig. 3) in the autocartographic system goes from the digitizing table to a small mini-computer called the PDP 11/45, which would then be hooked to the PDP 11/20 at Surveys and Mapping Branch, which is in turn linked to the PDP 10 main computer. It is this final computer which is the process control unit that drives the plotting table serving Surveys and Mapping Branch. On this basis we also wanted the facility to do cartographic analysis, using a flat-bed plotter (so that we could plot information and data directly onto geological or topographic maps), and in addition a cathode-ray display tube for interactive graphic analysis.

At the present time we have a Bendix digitizer and the PDP 11/45, linked to the Surveys and Mapping unit so that we can operate on-line with their process control unit (*see* paper no. 32 by Debain in this publication). Our EA1 430 plotter is operational. We

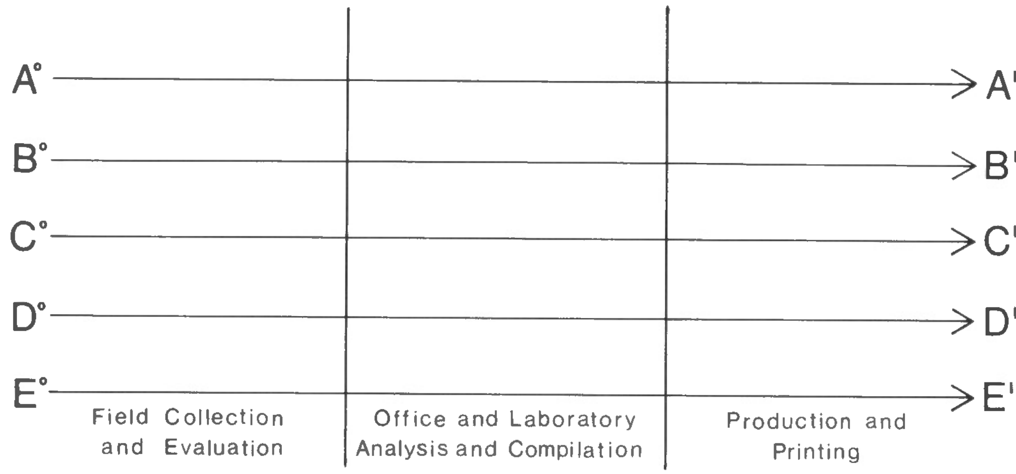


Figure 1.

This figure shows a series of projects with the various stages each goes through from initial inception (e. g. A^o), via data collection in the field, office compilation, and finally production and attainment of final objectives. During 1971-1974, the major aim was (1) to apply computer techniques where appropriate in each of the stages to expedite attainment of assigned objectives and (2) to increase efficiency of techniques utilized.

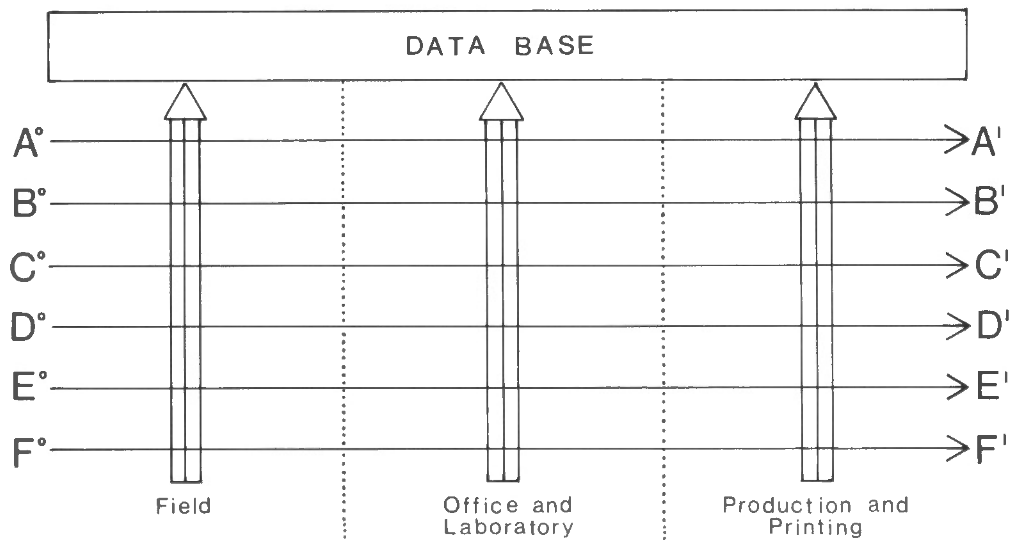


Figure 2.

This figure emphasizes that planning priority was given to supporting projects (X-direction) and increasing efficiency of techniques. Consideration of moving in Y-direction (to form an integrated data base from projects) was low because: -

(a) it would have handicapped instead of assisted attainment of project objectives, through effort required to establish 'standards' and/or appropriate linkages.

(b) operational efficiency at the project level must first be established to guide development.

have not yet developed the system whereby we can capture data directly from the digitizer and return this for plotting directly on the EA1 430 plotter. The interactive display unit is a small CRT display which is operated through computers in Toronto. As a result, it has been used only for pilot projects (*see* papers by Crain, Picklyk and Ridler in this publication) and we would not anticipate using telephone line linkage as part of a system in which we have to transmit large volumes of data.

In summary then, the emphasis was placed on ensuring that as each stage of work is completed the next stage 'downstream' would have the capability and facility to handle information that had been passed on. In a sense then, one can make an analogy with a river. Our planning has been from the mouth of the river back upstream.

If we start on the right-hand side of Figure 4, there is a wide, broad channel (near the mouth of a river), through which nearly all information flows. This is the production stage. In the office-data-reduction/cartographic-analysis-compilation stage, there are substantial tributaries feeding into the mainstream. These could be concerned with such features as paleontology, age-date information, chemical analyses, trace element geochemistry, paleomagnetic results, all of which could conceivably be computerized or handled in partly computerized form. In the third sector upstream, namely that of the field work, there are even fewer tributaries and the mainstream itself is much smaller. This simply reflects the degree to which computers can be applied to support the various activities in this particular area. The hope is that this stream pattern, once established, will be comparable to that of a young recently uplifted terrain that will gradually cut back upstream with time.

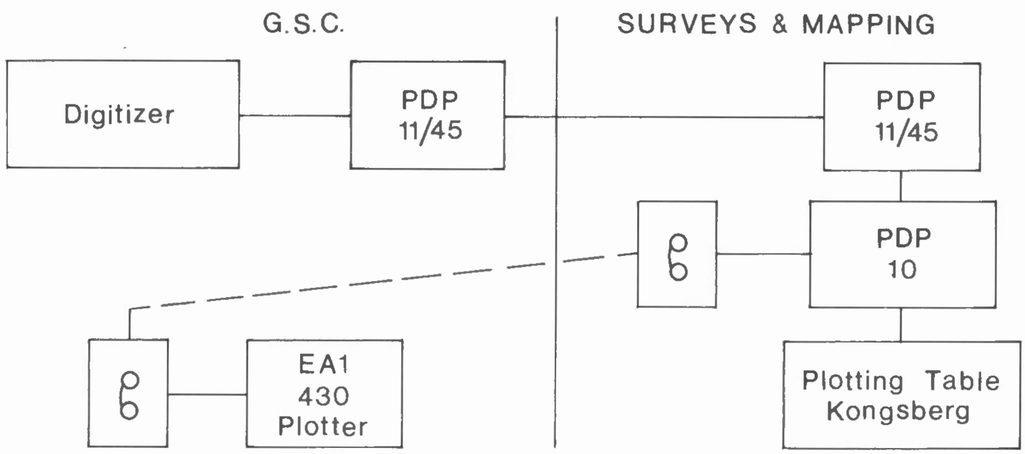


Figure 3. Shows disposition of hardware for auto-cartography at the Geological Survey of Canada. A digitizer (digitizing table) is linked to mini-computer (PDP-11/45) which in turn is linked to its counterpart in the Surveys and Mapping Branch. This mini-computer in turn is linked to the large PDP 10 computer which performs the major processing and drives the large Kongsberg plotting Table. Also indicated is the EAI 403 plotter used for precise, working plots on a cartographic base. The dashed line indicates intertion to transfer digitized data by tape from the PDP 10 to the EAI 430 flat bed plotter.

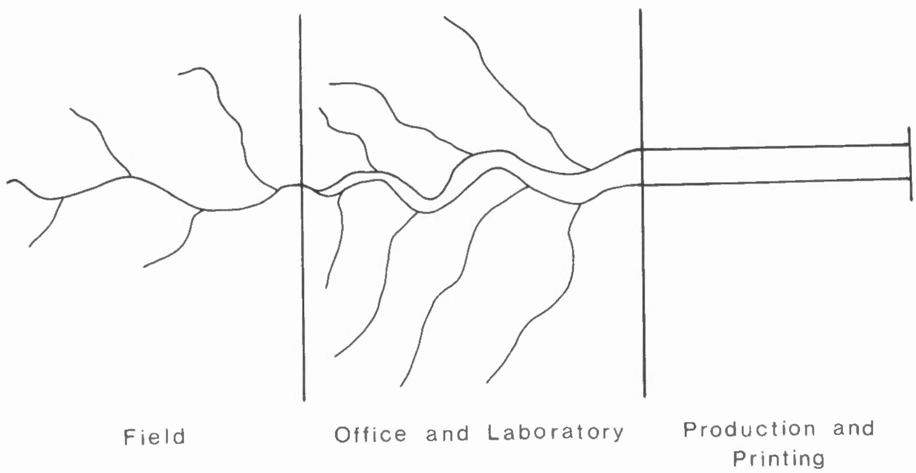


Figure 4. Illustrates that plan for development of support took form of a stream pattern.

Computer support within the Geological Survey of Canada

The Geological Survey of Canada is a branch of the Department of Energy, Mines and Resources. In Ottawa, the Department has a CDC 6400 (currently being upgraded to CYBER 74) which operates under the SCOPE 3.4 system. This facility can be used either in batch or by use of remote batch terminals, interactive terminals with 110 and 300 BAUD ports. Survey personnel may use commercial computers, but the tendency during the last two years has been to have more and more work done through the in-house facilities.

There are a number of major problems here however; namely the CDC 6400 has only a 64-character set, and because its number of bits per word differs from those of other computers, there are problems attaching peripheral hardware. Attempts to have all geological work done through the CDC 6400 has led to certain groups being handicapped.

The biggest factor in use of a computer by geologists is to break down or minimize the initial energy barrier to be surmounted, so that the geologist can interface with the computer as easily as possible. There are three aspects to this problem: (a) Educa- (a) Education, (b) Personnel, (c) Hardware. It is in the sphere of education that we have had the strongest support from the Computer Science Centre. Through their efforts, particularly those of Gwynneth Martin, a course has been developed based on geological examples which displays both the facilities and the manner in which geological data can be recorded, stored, retrieved and analyzed. In terms of programming support, we have been fortunate in having some substantial assistance from the Computer Science Centre. In addition we have appointed a data manager for the Regional and Economic Geology Division of the Geological Survey, part of whose role will be to educate and assist newcomers to the computer scene. The last problem is the one of hardware, and what I refer to here are the new techniques of data entry, in particular, either off-line terminals (with cassettes or paper tape), or even "intelligent" terminals whereby data can be recorded on small cassettes, and partially edited before being submitted to the main computer. Current estimates are (Askevold, pers. comm.) that such data entry systems are even cheaper than typing. Such systems then have potential to greatly reduce the problems of bottle-necks in data entry and editing of the data.

Summary of Current Practices

At the present time any detailed discussion under this heading would, instead of showing some sort of coherence of effort within the Geological Survey, point to the varieties of individual systems and techniques that have been developed largely independently. Rather than be dismayed by this, one should remember this has resulted from a demand at the 'grass roots' level.

At the stage of data reduction and analysis, various means are at the disposal of the geologists. There are one or two terminals available which allow the geologist to work interactively with this data. The geologists using this technique are using INTERCOM (the interactive terminal system on the CDC 6400), and frequently use the MARS VI file management system. Most geologists using the computer, however, utilize the remote batch terminal in the basement of the Geological Survey, building , and have their results returned to the printer at that same terminal. Many geologists who have had the initial education course are able to describe the manner in which they would wish the data to be treated, and within the Regional and Economic Geology Division, such information is handled by the data manager, who, in this particular role, is really acting as a dispatcher and expediter. He then directs the work to where it can be key punched, so that the data can be processed by an applications program already written, or request development of a particular program, or even alternatively, prepare the data for plotting on the EAI flat bed plotter. Many field geologists are using this particular plotter so that they can get clean, true-to-scale plots, which may be done on their geological maps, or on overlays.

For final production we hope soon to be using, as a matter of course, the autocartographic system described by P. Debain in this publication.

Future requirements of Project Geologists

The main aim during the last three years has been to provide support to the individual project geologist. There are two major areas in which we are endeavouring to elicit greater support. One is in the area of easier data capture, or data entry into the computer. This has already been commented upon and is probably contingent upon acquisition of semi-intelligent terminals which can operate off-line. The other area is the ease with which the geologist can work with his data. On this basis, I believe that there are probably at least three levels of support required: (a) interactive (conversational) mode of operation, using a high level command language; (b) the use of professional programming support to develop specific application programs; and (c) programming by the geologists themselves.

For the average geologist to use the computer, the system has to be such that, after many months away from the computer, he can sign-on, utilize it, without having to relearn the rudiments of the system, and secondly, without having to expend substantial amounts of valuable time on programming and debugging. It is on this basis that the more expensive (from the point of view of efficient computer utilization) conversational mode with simple but powerful command language and text editor is probably substantially cheaper but more powerful and effective in the long run than the batch processing approach. This becomes particularly true when the geologist wishes to work with large volumes of free text which are an essential element of most field geologist's notes.

In terms of hardware, we are completing a project so that the EAI flat bed plotter becomes operational in free-line mode, and secondly, we have in conjunction with A. Porter and J. Linders of Surveys and Mapping Branch planned for the Extended Cartographic Monitor (XCM) to handle geological field data and symbols, at the digitizing stage.

Data base management systems for Geological Survey data

When talking about data base management systems, we should really state whether we are talking about systems to support data for individual projects (really file management systems) or whether these are data that might be accessed by large numbers of users.

In the first case, the geologist probably is only concerned with keeping those data alive until his project is completed and to all intents and purposes (so far as he is concerned), those data can then be discarded. In the second case, we are focussing on the institution's approach to handling data and information. Within the Geological Survey of Canada, the results of synthesis of work are usually partly reduced, and in all cases, generally compiled and presented in either map form and/or reports. Rarely are the raw data themselves published or made available so that they can be easily accessed by the public.

The first and most important stage here is to establish what the need or market will be for 'raw' or partly reduced data and information related to field and laboratory studies of officers of the Geological Survey of Canada. The current rate of acquisition of laboratory data and information alone is such that a co-ordinated system for handling this must be developed. Again there are three levels: one is the index to data and reports on information (currently partly handled by the Canadian Index). Two, the partly reduced, or partly synthesized, data as may appear on certain inventory maps; and third, would be the raw data themselves. Early in 1974, it was decided that some substantial effort should be made, both to co-ordinate activities within the Geological Survey of Canada in terms of the support of projects, and at the same time to look at long term data management needs within the Geological Survey.

Before looking in any greater detail at data base management, a number of criteria must be kept in mind. Firstly, it ought to be clear as to what the short and long term objectives of the Geological Survey of Canada will be. Secondly, identify the scientific support required to attain these, and thirdly the potential computer component of this support. We can then establish what data or information must be instantly accessible, what data and information can wait for an hour or a day for retrieval, and furthermore, what, if any, of that information need necessarily be in computer-processible form. Decisions on implementation will then take place as a result of pilot studies, cost analyses and demands and priorities.

Philosophy

Why bother to use computer? At the present we are chiefly forced into using it because of the need to store and/or process large volumes of data. Only a limited number of field projects stand to benefit from use of such equipment and associated techniques.

The computer, as used in the past, is only a tool of limited usefulness. However, few geologists have bothered to modify or shape the tool to serve them better. Field geologists must have their data easily captured, do selective retrievals of data for map units (or commodities), store and tretieve large volumes of free text (allowing for character string searches) and to work interactively with these data and the map (graphics) so that they can select the best portrayal of the distribution of a particular set of data. When we can succeed in doing this so that training of a geologist in computer use takes a minimum of time, then we will have adapted the computer so that it does the 'leg work'*.

In addition, there is the prospect, more a hope than a reality at present, that the computer will become more than a tool - perhaps also a medium. In my opinion the computer's potential has been scarcely exploited. To do so properly requires money, manpower and the devotion of scientists' time for system development and evaluation. As a result we are taking a positive yet cautious approach through pilot projects, some of which are referred to in this publication.

*Using the SARS System I discovered that descriptions of quartz diorites (from Bute Inlet) with sulphides containing the word, 'green' or 'greenish' occurs ten times more frequently than descriptions of quartz diorites without sulphides. Because there was no coding for colour in fixed format this simple yet striking relationship could only be detected by intimate scanning of the notes - in this instance done by computer using a character string search.

Summary and Conclusions

Projects within the Geological Survey of Canada have many diverse objectives. They span the spectrum of disciplines in geology, and furthermore, range from the realm of true data, through a continuum to the realm of soft qualitative information. Many project data bases are a mixture of both. Superimposed on this is the fact that even in one project, people may have different work styles. Accordingly, a combination of all of the above factors lead to the conclusion that no single system can probably be established for handling project data.

Rather than concentrate on such concepts as standards and data exchange, it is felt that greater emphasis should be placed on increasing the efficiency whereby geologic data can be collected, interpreted, analyzed and compiled through computer-based techniques. This can probably be done by a series of continuing pilot projects, whereby there is iteration as follows: the efficiency with which the data are collected, stored, retrieved, and analyzed is evaluated so that modifications may be incorporated in the next project. I feel that with current trends in technology, we are close to reaching a plateau whereby we can provide facilities that geologists can 'capture' data as easily as have a manuscript typewritten. An impressive example of this is the current work of Hudson et al. (1974).

At this time it would be unfair and unreasonable to comment on the approach we will take towards institutional data management. One thing, however, is certain - we must spend substantial time with the working geologists to find out what data they will require in the future and to try to plan accordingly. Presumably, we will attempt to establish pilot projects and the usefulness of these projects must be evaluated by the working geologists themselves.

This publication has been prepared to present to the reader some idea of the spectrum of computer-based activities currently under way within the Geological Survey of Canada. In view of the diversity, I am relieved that the Geological Survey of Canada did not, early on, take major action towards an institutional data base management system which could have been premature in implementation, and could have had a long-term stultifying effect. In fact, these same comments must be kept continuously in mind in planning for the future, and in ensuring that we do not become 'locked in' to some system that ultimately handicaps our work. Above all, the main theme of this work must be the attainment of the objectives set by the Geological Survey of Canada.

Reference

Hudson, T., Askevold, G., and Plafker, G.
1974: A new approach to computer-assisted methods in geology; Geol. Soc. Amer., abstracts with programs, v. 6, no. 3, p. 194-195.

2.

COMPUTER-BASED DATA MANAGEMENT IN THE GEOLOGICAL SURVEY OF CANADA - A PRELIMINARY APPRAISAL

Terry Gordon,
Regional and Economic Geology

Gwynneth Martin,
Computer Science Centre, Department of Energy, Mines and Resources

Introduction

The management of scientific data in the Geological Survey of Canada (G. S. C.) is the subject of an intensive study currently in progress (*see* Hutchison, this publication). Our paper is a preliminary appraisal of the use of computer-based data management systems for information processing in the context of Geological Survey applications. It is based on several years of practical experience in computer use in geological projects (e. g. Gordon, 1972), system design, training, consultation and system evaluation; as well as informal discussions with staff scientists.

It is important at the outset to distinguish file management from data base management (*see* our other paper in this publication). For purpose of introduction we make the following distinction. File management implies management of simple related data for a single project with well defined goals. A small number of people are responsible for collection, maintenance and use of the data. This has been the prime information management need in the Geological Survey and it will remain so for the immediate future. Data base management implies management of complexly structured data for a wide number of users with a variety of applications. The collectors and users of data are often separate groups. Although there is a growing requirement for data base management in the Geological Survey, this is a new function, divergent from the Survey's historical role, and with far-reaching implications for both administration and the direction of individual projects. Some of the problems of the data base management are discussed later in this paper.

This report will thus deal primarily with file management. A classification of existing Geological Survey files, their requirements for computer assistance, and our preliminary evaluation of several systems to meet these requirements will be presented. We also offer a caveat on data base management - separating it from project-oriented file management.

Geological Survey Files

To simplify our evaluation of file management systems, we have distinguished four general types of Geological Survey files currently being computer processed. In addition to physical parameters such as file size and lifetime we have considered:

1) File contents - the proportion of numeric to text data.
2) Project output - the degree of synthesis and interpretation of raw data required to meet project goals.
3) User sophistication - the amount of computer science expertise required by the person or persons using or maintaining the file.
4) Predictability of file output - the proportion of "standard" or "production" queries concerning the file - as opposed to ad-hoc retrieval.
5) Response time - both the length of time taken to produce information in response to a query or the time taken to update the file.

TYPE I Numerical project files

The files contain almost exclusively quantitative data - geophysical or geochemical, often initially captured automatically (for examples *see* Grasty, Holroyd, Shih, in this publication). The ultimate project output is usually a map or series of maps with very little interpretive content that summarize or otherwise depict the data approximately as it was initially recorded. The people responsible for collecting and processing the data, and for producing the final output comprise a group with a high level of expertise in the use of computers for mathematical operations. Retrievals from the file tend to be simple extractions of large amounts of numerical data and used for production purposes - either statistical processing or map production. Users normally wish frequent and rapid access when the data are being edited at the initial stages of acquisition. In the production stages of the project, overnight turn-around is adequate. Because these data are objective, well defined, and in machine processable form from the time of capture to production of the final map, they readily lend themselves to archival storage. Hence, although the file size for a single mapping project may be in the range of ten million to twenty million characters, there is an accumulation of such files resulting in archival "superfiles" of relatively restricted content.

TYPE II Text Project files

These files contain a large percentage of non-quantitative information - often geological field data of various types (For example Lambert and Reesor, Hutchison and Roddick, this publication). Project output usually contains a high degree of subjective interpretation of the raw data in the file. Usually only one or two people are responsible for data collection, processing, and interpretation. With rare exceptions they have very limited training and exposure to computer science. Although there are a few standard 'production', uses of the data (e.g. posting maps), most retrievals are complex boolean logical requests that select limited parts of the file. In fact, the interpretation process usually requires study of the output of one retrieval request in order to formulate the next (*see* Picklyk and Ridler, this publication).

This personal interaction with the data makes rapid turn-around essential for both updating and retrieval of information in the file. These files are small, less than one million characters, and have lifetimes of three to four years.

Non-quantitative data, selected in the field to aid the interpretation process, do not readily lend themselves to archival storage. Definitions of terms, accuracies and spacing of observations and, most importantly, relationships between sets of observations all vary drastically from file to file. As the main product of the investigation is synthesis and interpretation, it is this secondary information rather than raw data that is of greatest interest to other users.

TYPE III Internal archival files

These are files containing both quantitative and non-quantitative data used for on-going programs rather than short term projects - for example mineral deposits, oil and gas pools or bore-hole data (*see* Belanger, Shih et al., Whitmore in this publication). These files provide input data to various other projects and users, hence the only easily defined output are the data themselves. Often the separate responsibilities of data capture, data validation, computer operation, and provision of retrievals have been assigned to only one or two people. The computer science expertise of such people varies widely. Retrievals from such files can be as unpredictable and complex as for project files, but more often large amounts of data are extracted to produce either a summary map or a printed report on a 'production' basis.

These files are large, have a steady growth rate and expected lifetimes of tens of years. Retrieval turnaround should be within one hour, update transactions on an overnight basis.

Although the "multi-user" nature of these files has echoes of our definition of "data-base", they nevertheless have a simple record structure. File maintainers usually work closely with the groups using the data so the hiatus between collection and use of information is minimized.

TYPE IV Public archival files

The data in these files varies from purely reference information to archived numerical data. This type of file is similar to type III files in terms of project output (the data themselves), retrieval queries (varied) and physical file characteristics (response time, size etc.) However, this information is available to the public and uses are not well known.

The Branch has only a few of this type of file. They represent prototype "data base" systems although they were designed and built before the computer industry provided and supported data base management packages to any extent.

There are other miscellaneous file types within the Geological Survey and files which could be placed in either of two categories of our admittedly simplified classification. Nevertheless, this spectrum of the file types with their associated processing requirements does provide a basis for evaluating general file management systems.

Requirements of file management systems

To begin with, files of type I - numerical project files - do not lend themselves to processing by currently available general file management systems. The large quantities of numerical data, high computational requirements and restricted types of output make them ideal for the custom-designed systems described elsewhere in this publication. Accordingly, features of file management systems will be discussed in the context of file types II, III and IV.

There are ten criteria that may be used for preliminary analysis of the suitability of file management systems currently installed in Ottawa for Geological Survey applications. We have not considered cost for two reasons - one, there is insufficient data to make accurate comparisons; two - costs are probably similar between systems and thus a less important factor for comparative evaluation than the efficiency of the system in attaining project objectives. A complete examination of all commercially available systems would, of course, be much more stringent (e.g. Kuss and Madill, 1972), and such an evaluation may well be undertaken on completion of Hutchison's study (Hutchison, this publication). The criteria are:

1) System Maintenance and support.

File management systems are complex pieces of software inevitably subject to problems both at time of delivery and as the computer installation changes operating systems, compilers, and hardware. Unless the Department is willing to commit its own resources to the highly specialized work of maintaining the file management system, there must be a source of readily available system expertise. Most commercial systems come with this support as part of the purchase of lease price, but the quality of service varies. The best test is other users' experience with any particular vendor.

2) Training and experience required.

It is particularly essential for type II and useful for type III files that the users have personal interaction with their data. That is, a scientist with very little computer science training must be able to retrieve and update from his file with a minimum of help from a computer specialist. On the other hand, there will always be a requirement for "system managers" (either with Geological Survey or the departmental Computer Science Centre) to assist users, and prepare and run complex or non-standard jobs. "System managers" thus require a good background in computer science. There are then, two types of training required; a course for non-programmers who need to have a general view of the system and to know the essentials of retrieval and possibly update operations; and a more extensive course for analyst-programmers who need to have a more detailed understanding of all the features of the system.

3) Documentation.

The manuals provided with a system should give an analyst or system manager sufficient information to enable him to use the basic system functions with a minimum of training and without recourse to consultation with the supplier. Documentation should also be made available for non-programming users, to explain system procedures and at least the retrieval and update commands.

4, 5, 6, 7) File definition, data validation, update and retrieval flexibility.

These operations are briefly described with reference to field data (type II) files in our other paper in this publication. Type III and type IV files require essentially the same features as field data files.

Some systems provide the additional capability of cross-linking files - that is, allowing data in a record in one file to be linked with data in other, different, files. This is extremely useful for type III and IV files which may ultimately serve as indexes to type II files. For example, laboratory data may be linked to field observations or a mineral deposits archival file may reference a more detailed project file.

8) Response time.

As a general rule, the more interpretation and synthesis performed with the data, the more rapid access is required. Hence, users of type II files need on-line updating and retrieval capabilities for at least part of the file life-time. On line access is less essential for type III and type IV archival files because the data are relatively static and file output is normally studied extensively before a new retrieval request submitted.

9) Output options.

File management systems will usually produce a simply formatted printout of retrieved data. This is often all that is required in the interpretive type of retrieval common with type II files. Archival files (types III and IV) are often the source of more elaborately formatted reports and summaries. Thus a report generator is an extremely useful system capability for these applications.

Type II and type III files often contain information best presented graphically or processed statistically to be of use. Since these types of output are application dependent, they are not provided by the file management system but require custom-designed programs. Thus an essential retrieval capability is to produce a data file for subsequent computer processing. This is commonly done by producing a file of fixed length formatted records.

10) Escape routes.

Computer technology in general is subject to rapid change and this includes software packages such as file management systems. This, together with the fact that the Geological Survey has relatively little experience in the use of such systems, makes it inevitable that files of all types will have to be transferred from one system to another. There are two ways of accomplishing such a transfer; one, by starting from the original input data; the other, by simply reformatting the existing file for processing by the new system. This latter method is preferable as the original data will have been updated and modified.

If reformatting is not possible, then the original data may have to be used. This may present some problems if the original system required special delimiters to identify data items. These would have to be removed or modified for processing by another system.

Table I summarizes our evaluation of several file management systems available to the Geological Survey in Ottawa. As mentioned above these systems, as well as Geological Survey requirements, are subject to continuing change. Thus this table represents only our current assessment and will be subject to periodic revision.

Problems in data base implementation

With the increased use of computers for processing geological data of all kinds, there is the possibility of providing rapid access to G. S. C. information by establishing "data base" on a computer. These would be created, maintained, and used through a data-base management system or systems. Our experience, as well as that of others (Patterson, 1972), has made us aware of some of the major problems involved in the implementation of such a scheme. These problems can be divided roughly into three categories:

1) computer systems - the technical aspects of providing the hardware and software to deal with all aspects of data handling.

2) information systems - the tasks of ensuring that the needs of the data users are understood and that the data, their accuracy, precision and validity satisfy these needs - i. e. ensuring that data transmission is true communication.

3) administration of systems - the problems of maintaining control over costs, data security and validity whilst ensuring that the system meets users needs - all within the framework of Geological Survey budgets and priorities.

Computer systems

Data base management systems are expensive. Aquisition and maintenance of hardware and software are themselves costly, but even more important are the day-to-day costs associated with using the data base. There are many ways of organizing data and

TABLE 1

SYSTEM AND INSTALLATION	MAINTENANCE	EXPERIENCE AND TRAINING REQUIRED	DOCUMENTATION
MARK IV Systems Dimensions Ltd.	Informatics, Inc. Canoga Park, California Good	System manager - 1 week System designed for use by experienced programmers.	Good for system manager. Good for users.
MARS VI PARTIAL INVERSION Departmental Computer Science Centre	Control Data Corporation Minneapolis, Minnesota. Guaranteed to June 1975 Fair	System manager - 1 week Users - 3 days	Fair to good for system manager. Fair to good for users.
SAFRAS Departmental Computer Science Center	University of Western Ontario and users. Will not be supported by GSC after Sept. 1/74.	System manager must be experienced COBOL analyst - 2 weeks Users - 1 week	None for system manager. Fair for users.
GDMS[1] Departmental Computer Science Centre.	Will be supplied at users' risk. System under development. Support function not assigned.	As SAFRAS	None for system manager. Good but incomplete at 01/08/74.
G-EXEC[2] Single file system at Departmental Computer Science Centre	Supplied at users' risk by Institute of Geological Science, London. Good informal support from originators.	System manager must be experienced FORTRAN analyst - 1-2 weeks Users deal through system manager.	Fair to good for system manager. Fair for users.
SYSTEM 2000[3] Departmental Computer Science Centre	MRI Systems Corp. Austin, Texas Good	System manager - 2 weeks System designed for use by experienced programmers.	Good for system manager. None for users.

[1] Reference Williams et al., 1972.

[2] Reference Martin 1974. The full G-EXEC system provides much more than file management. It contains many data analysis and display programs. This table refers only to the portable file management subsystem.

[3] System 2000 is a full data-base management system, not a file management system and is included in this table for comparative purposes.

Table 1 (cont'd.)

SYSTEM AND INSTALLATION	FILE DEFINITION	VALIDATION	UPDATE	RETRIEVAL FLEXIBILITY	RESPONSE TIME
MARK IV Systems Dimensions Ltd.	Very flexible Complex hierarchies Cross-file linking	Good	Good	Good but form-oriented	Designed for batch processing. No time sharing at SDL.
MARS VI PARTIAL INVERSION Departmental Computer Science Centre	Fixed length records Repeating groups 1 level of hierarchy Cross-file linking	Good	Fair	Good	Time sharing and batch.
SAFRAS Departmental Computer Science Centre	Variable length records Repeating groups 1 level of hierarchy No cross-file linking	None	Poor	Fair to good	Batch processing ½ hour turnaround.
GDMS Departmental Computer Science Centre.	Variable length records Repeating groups 2 levels of hierarchy No cross-file linking	Good	Incomplete at 01/08/74	Incomplete at 01/08/74	Batch processing 1-2 hour turnaround.
G-EXEC Single file system at Departmental Computer Science Centre	Fixed length records No repeating groups Hierarchy requires cross-file linking in full system	Fair	Not provided by system but easily managed with utilities	Poor to fair	Batch processing ½ hour turnaround.
SYSTEM 2000 Departmental Computer Science Centre	Very flexible Complex hierarchies Cross-file linking not required	None	Good	Good - but complex data structures can cause complicated and expensive retrievals.	Time sharing and batch.

Table 1 (cont'd.)

SYSTEM AND INSTALLATION	OUTPUT OPTIONS	ESCAPE ROUTES	COMMENTS
MARK IV Systems Dimensions Ltd.	Formatted records Report generator	Good Input data on standard IBM files. All files except hierarchic can be COBOL processed.	Requires programmer interface between user and system. Good for type IV files, fair for type III files, poor for type II files.
MARS VI PARTIAL INVERSION Departmental Computer Science Centre	Formatted records Report generator	Fair to good Input data are fixed length records. Output of whole records requires sort.	Requires programmer for some operations. Good for type II and type III files. Fair for type IV files. Retrieval and update easy to perform and fairly powerful.
SAFRAS Departmental Computer Science Centre	Formatted records No report generator	Poor to fair Input data has delimiters. Output contains zeros for missing data.	Problems with computer and operating system changes. Poor to fair for type III and type IV files, poor for type II files.
GDMS Departmental Computer Science Centre	Formatted records Report generator incomplete at 01/08/74	Fair Input data has delimiters. Output "work files" are fixed length coded.	System still under development. Will have problems with compiler and operating system changes.
G-EXEC Single file system at Departmental Computer Science Centre	Formatted records No report generator but special purpose reports	Fair to good Input and stored data on fixed length records with twelve header records	Problems with compiler and operating system changes. Requires programmer for some operations. Fair for type III files. Fair to good for type IV files; too limited for type II files. Full system (not on Departmental computer) provides more features.
SYSTEM 2000 Departmental Computer Science Centre	Formatted records No report generator	Fair to good Input data has delimiters. Output of whole records requires small FORTRAN job.	Requires programmer interface between user and system. Powerful retrieval capabilities and slow expensive updates. Sequential file processor and report generator promised. Under test in GSC applications.

designing retrieval and update methods. The permutations of these logical operations are compounded by the many different ways of physically placing data on computer storage devices (Lefkovitz, 1969). Each method provides economies in the use of some expensive resources (e.g. core usage, CP time) for certain data base operations - but with resulting increased costs for others.

Thus, although it is claimed that data base management systems make data "independent" of their applications or uses, in fact the implementation and efficient use of such systems require considerable knowledge of the applications - i.e. how often, and which data will be acquired, modified and retrieved.

In business applications, how and when the data will be accessed can be predicted with considerable precision, for example, to produce payrolls, inventories or billings. This is in stark contrast with scientific data systems, a prime characteristic of which is their unpredictability (Crain, 1974). The task of the computer specialist in determining system requirements and designing a data base for scientific needs is thus considerably more difficult than that of his business-world counter-part.

It can be safely stated that implementation of a scientific data base system will be considerably more expensive both in time and money than a commercial system of comparable size.

Information systems

Although the computer science problems are formidable, the establishment of a data base requires answers to the even more difficult questions of which users the system is to serve, what information they need, and how it is to be provided to them. Wildavsky (1973) gives an excellent discussion of the pitfalls of over-enthusiastic development of technical information systems in government. One of his ten rules of thumb is this: - "Unless you understand precisely who will use each data bit, how often, at what cost, relevant to which decisions they are empowered to make, don't proceed". Although this represents an extreme attitude, a careful analysis of the data use must obviously accompany any data base implementation in the Geological Survey.

At present it is not entirely clear what data services the users of G.S.C. information would like to have. It is even less clear whether or not all their potential demands can be met at reasonable cost. Hutchison's project will attempt to answer these questions.

It is evident that the services to be provided will have great impact on computer systems design. At the same time, realization of both the costs and benefits of computer technology will force changes in users' requirements. Thus, although we have discussed these two aspects separately, they are inextricably linked. Both computer system design and user demands will necessarily undergo considerable modification before optimum performance is achieved.

Administration of systems

Looming over the technological problems of data base operation are the administrative problems of managing such a system. If data are to be released to users, they must be correct. Internal mechanisms will have to be created to ensure that information in any data base is valid, up to date, and that confidentiality requirements are met. We would estimate the administrative costs of data base management at ten times the computer systems costs.

Careful administration is required to ensure that resources are not diverted into collecting and maintaining data that is never used; and that conversely, important decisions are not based on obsolete, incorrect or incomplete information merely because responsibility for that particular data has not been assigned.

CONCLUSIONS

General file management systems can be used to advantage in the processing of Geological Survey data. Although the information in Table 1 shows that no single system will provide <u>all</u> of the features that might be considered desirable, some system can be found to meet <u>most</u> requirements. Careful attention to specific needs, evaluation of the ten criteria we have attained, and consultation with experienced users will provide a good basis for chosing a system for any particular application.

Our comments on file management systems and their uses may seem a small contribution to the obviously more difficult and costly task of data base management in the Geological Survey. We are confident, however, that applying simple systems to data management on a small scale will provide necessary insight into solutions to the larger problems we have outlined. Our point is that successful data base implementation depends on careful and extensive study, combined with a realistic appraisal of the large amounts of time and money that will be required.

REFERENCES

Crain, Ian K.,
1974: The analysis of scientific computer systems; <u>in press</u> INFOR.

Gordon, T.M.
1972: Report to the working group on the storage and retrieval of geological field data, February 1972. Geological Survey of Canada internal report.

Kuss, J.K., and Madill, J.P.
1972: A technique for evaluation of generalized data management systems. INFOR, vol. 10, p. 311-319.

Lefkovitz, D.
1969: File structures for on-line systems. Spartan Books, N.Y., 215 p.

Martin, Gwynneth
1974: An evaluation of the G-EXEC system for use in the Geological Survey of Canada. August 1974. Geological Survey of Canada internal report.

Patterson, Albert C.
1972: Data base hazards. Datamation July 1972, vol. 18, p. 48-50.

Wildavsky, Aaron
1973: Consumer report (review of "Politicians, Bureaucrats, and the Consultant" by Gary D. Brewer) Science, vol. 182, p. 1335-1338.

Williams, G.D., Dickie, G.J., Gordon, T.M., Groen, H.A., and Kelly, A.M.
1972: Future development of the SAFRAS generalized data base management system. July 1972. Geological Survey of Canada internal report.

3.

COMPUTER-BASED STATISTICAL ANALYSIS OF GEOLOGICAL DATA: PAST AND CURRENT ACTIVITIES

Frederik P. Agterberg,
Regional and Economic Geology Division

Computer Usage

Medium to large electronic computers are particularly suitable for the digital manipulation of large arrays of numbers (matrices). Three basic operations that can be performed on the computer are: (1) matrix inversion; (2) calculation of Fourier transforms; and (3) extraction of eigenvectors from a square matrix. One or more of these techniques is used in nearly all statistical methods for interpolation, prediction and multivariate analysis developed and applied by us since 1962.

From 1968 we have been concerned with the development of computer-based methods for correlation of the occurrence of various types of mineral deposits with combinations of parameters systematically quantified for cells superimposed on the geological framework of a region. The objective of this project is to estimate the probability that a cell is a hit, viz. of the same type as a number of control cells within the same region or located elsewhere which are known to contain one or more deposits of the type being studied.

Such probabilities can later be added for overlapping unit areas and their sums can be automatically contoured. The contours then represent the number of hits expected per unit of area. The current work on statistical methods is best illustrated by practical examples.

Example 1: Kidd Creek deposit. Figure 1 shows a set of cells measuring 10 km on a side. The cell marked C is a single control cell containing the large (Cu, Zn, Pb, Ag) orebodies of the Kidd Creek Mine in the Abitibi Volcanic Belt of the Canadian Shield. Cells with a minus sign were not used for statistical analysis because of lack of data. Percentage values for Archean rock types in combination with Bouguer anomaly and aeromagnetic data were used for 644 cells. The data for this example were the same as those used by Agterberg et al. (1972).

The input matrix X for this test run consisted of 644 rows (for cells) and 52 columns (for parameters). A single column Y was defined consisting of 644 elements equal to zero for every cell except for the control cell to which a one was assigned. By means of the computer it is feasible to compute a new column $Y_c = X(X^TX)^{-1}X^TY$ which also consists of 644 values. (The superscript T denotes matrix transposition and -1 represents inversion.) If a number of conditions are satisfied, every number of Y_c expresses the similarity of a cell with the control cell. The resulting 644 numbers were arbitrarily grouped into three classes as shown in Figure 1. Empty cells correspond to negative values in Y_c; cells with a heavy frame and black cells represent positive values. The 12 black cells have the largest positive values and would have a geological setting similar to that of C; the 123 other cells with positive values have a similarity index which is greater than average. The potential usefulness of this approach is that we could speculate that similar cells are more likely to contain other large sulphide deposits. For comparison, the other cells known to contain at least 1,000 short tons of copper or zinc (from published data for past-producers, producers and prospects) are represented by symbols in Figure 1. For statistical reasons, better results are achieved when the number of control cells is much greater than one.

Example 2: Nickel deposits in Archean rocks of UTM-Zone 17. A trend in our approach during the past few years has been to divide large regions into many relatively small cells using the UTM-grid system, and to restrict the statistical analyses to specific types of deposits. A major difficulty for statistical work in prediction of occurrence of mineral deposits is that we are dealing with rare events. Another problem is that our geological variables for small cells are dichotomous, viz. a variable assumes a value only if it is present in a cell, otherwise its value equals zero.

In the data block of Figure 2, there are 2,542 cells but only 16 of these contain one or more sulphide deposits with nickel in concentrations sufficient for mining at present or possibly in the future. For statistical analysis, this set of 16 control cells was enlarged to 25 by including other Archean deposits of this type located within the Superior Province but outside of UTM-Zone 17.

Nevertheless, the linear statistical models which previously had given good results for polymetallic Archean sulphide deposits failed to produce realistic results for the nickel deposits. In 1972, Dr. J.W. Tukey of Princeton University had commented on our approach suggesting that nonlinear models including the so-called logistic model might provide us with better results. This model indeed yielded greatly improved results for the Archean nickel deposits with the result shown in Figure 3. Plots of this type are obtained automatically on the CALCOMP-plotter at any scale and, if required, in the Lambert conformal projection to superimpose it on topographic maps at scale 1:1,000,000.

Crosses in Figure 3 represent individual nickel deposits whereas the contours represent number of hits expected per unit area measuring 40 km on a side. We recall that a hit was defined as a 10 by 10 km cell with one or more deposits so that deposits which are close together may be counted as a single event for statistical analysis. Although the linear model did not produce a good contour map in this example, it identified several factors known to be related to occur-

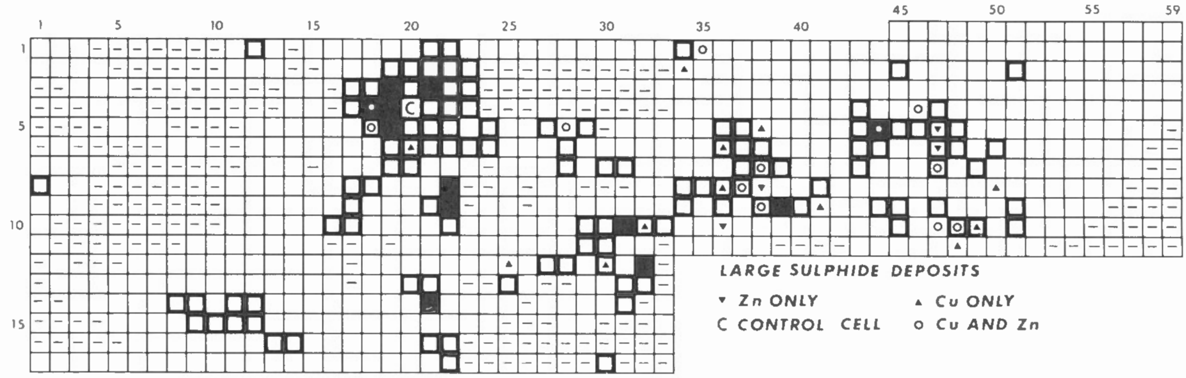


Figure 1. Similarity of (10 km x 10 km) cells in comparison with individual control cell. For explanation, see text.

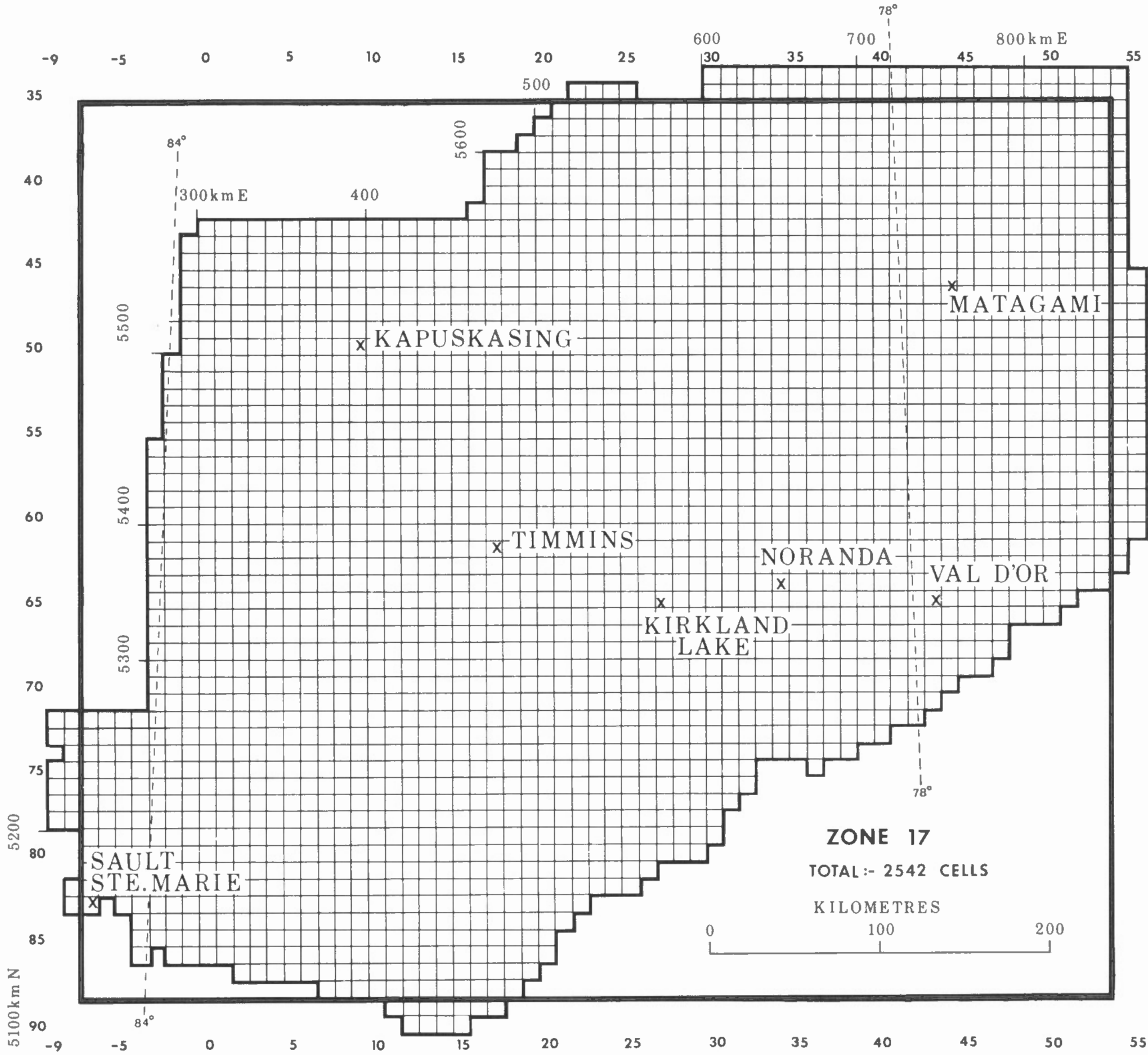


Figure 2. Block of cells in UTM. Zone 17 for which geological data were coded. Heavy frame is outline of automatic contour map shown in Fig. 3. Cell numbers along margins correspond to those in Fig. 4.

Figure 3. Automatic plot showing nickel deposits (crosses) and contours for number of "hits" expected per unit area of 40 km x 40 km. For location, see Fig. 2.

```
  -9   -4    1    6   11   16   21   26   31   36   41   46   51   56
34 
35 
36                                   00
37                                  010
38 
39                            0
40                            0
41                            00
42                           0 0
43                           0
44 
45                          1
46                         13
47                         00
48 
49 
50 
51                  0
52 
53                                                    00
54                                                    0
55                         UU     0     0
56                         00010  01    00
57                         000010 0000     000
58                         0 0 011 00000     00
59                         000010U    0           00
60                         00 0000 00000           0 00 00`
61                           00000  000000    1       0 00
62                           00000   00        0    010
63                         000110    0            00120
64                          UOJUU                    02002
65                      00   01  0            0   0  010 000
66                            00        0                 0
67                            00  00   0000
68                           000
69                            0
70 
71 
72 
73 
74                                          00
75                                          0
76 
77                                      00 0
78 
79 
80 
81 
82 
83 
84 
85 
86 
87 
88 
89 
```

Figure 4. Print-out of data on ultramafics (variable no. 6). Numbers correspond to cells shown in Fig. 2;

0 = 0–10% 2 = 20–30%
1 = 10–20% 3 = 30–40%

of cell area underlain by ultramafics.

rence of nickel deposits. One of these is the occurrence of ultramafic rocks whose input pattern is shown in Figure 4. This final diagram illustrates that the printer of the computer can be used to produce inexpensively selected patterns for individual variables stored in the data base.

Bibliography

Agterberg, F.P.
1967: Computer techniques in geology; Earth Sci. Rev., v. 3, p. 47-77.

1974: Geomathematics. Mathematical Background and Geo-science Applications; Elsevier, Amsterdam, appr. 600 p.

Automatic contouring of geological maps to detect target areas for mineral exploration; J. Int. Assoc. Math. Geol. (in press)

Agterberg, F.P., and Cabilio, P.
1969: Two-stage least-squares model for the relationship between mappable geological variables; J. Int. Assoc. Math. Geol., v. 1, p. 137-153.

Agterberg, F.P., and Chung, C.F.
1973: Geomathematical prediction of sulphur in coal, new Lingan Mine area, Sydney coalfield; Can. Mining Met. Bull., October 1973, p. 85-96.

Agterberg, F.P., Chung, C.F., Fabbri, A.G., Kelly, A.M., and Springer, J.S.
1972: Geomathematical evaluation of copper and zinc potential in the Abitibi area of the Canadian Shield; Geol. Surv. Can., Paper 71-41.

Agterberg, F.P., and Fabbri, A.G.
1973: Harmonic analysis of copper and gold occurrences in the Abitibi area of the Canadian Shield; Proc. 10th Symp. Application Computers in Mineral Industry, Johannesburg, S.A., April 1972, p. 193-201.

Agterberg, F.P., and Kelly, A.M.
1971: Geomathematical methods for use in prospecting; Can. Mining J., v. 92 (no. 5), p. 61-72.

Agterberg, F.P., and Robinson, S.C.
1972: Mathematical problems of geology; Bull. Int. Statist. Inst., v. 38, p. 567-596.

Fabbri, A.G.
Design and structure of geological data banks for evaluation of mineral resource potential; for publication in Can. Mining Met. Bull. (in preparation)

BARIUM, FLUORINE, STRONTIUM FILE

K.R. Dawson,
Regional and Economic Geology Division

The Geological Survey of Canada project 720095, a study of the Canadian deposits and occurrences of barium, strontium, and fluorine, is supported by a computerized file. The file is operated in batch mode using the MARS VI (Multiple Access and Retrieval System) software available on the Department of Energy, Mines and Resources CDC6400 computer. The contents of the file have been compiled from a thorough literature search supplemented by information from the Canadian Index to Geoscience Data, Mineral Resource Section files, Mineral Resources Branch files, and CAN/SDI service. The unit of record in the file is a single deposit or occurrence. There are 746 records in the file each having a maximum of 175 subitems (items of data) and a maximum total length of 3,607 characters. The length of the current file exceeds 2.7 million characters. Individual records consist of as many as nine 59-character subrecords and some subitems have value tables having a maximum of 44 values. The file is resident on disc when in operation and on magnetic tapes for backup purposes. Retrievals and other processing are done as a single batch mode computer job. The subject matter is descriptive of the individual deposits and the subitems stored include many of those recommended by National Committees that studied mineral deposits files.

Operating Procedure

The initial step in the creation of the file was the design of an appropriate file definition and the drafting of suitable input documents for the collection and organization of the data. Subsequent entry of fragmentary records is done on standard 80-column forms and the contents of both types of form are punched onto IBM cards. The special forms generate cards for either the preliminary load or subsequent "add-on" loads whereas the standard 80-column forms generate input for partial records for update runs.

The loader module which creates a file that will be processable by the MARS software requires three steps once the cards have been punched and verified: (1) the cards are sorted and read into permanent file, (2) the contents of the permanent file are passed through a COBOL program that formats the data according to the file definition and places the output on permanent file, (3) the contents of the second permanent file are processed by the loader module to generate the record file and the new related data base file both of which are also placed on permanent file or loaded on magnetic tapes. The add-on-load routine is a variant of this procedure that inserts new records into the existing record file and generates a modified record file and data base.

The update module inserts all subitems into existing records but is not used to add whole records or to change key values. The corrections and additions are collected and coded on the standard 80-column form, punched into cards having the accession number on the left and the subitems on the right. A dummy update run is made to identify errors which are then corrected and the final run is made. The updated data base is placed on permanent file and it replaces the existing data base for the operation of the file.

The retrieval module is very efficient permitting a geologist to search for and retrieve at all subject levels from single subitem records to complete records (*see* Fig. 4). It has an additional function that it serves as the entrance into several collection specifications including SUBSET, QUALIFY, PRINT, REPORT, COUNT, RECAP etc. The subscript "ANY" permits retrieval from a repeating set of subitems and the subscript "RANGE" tests ranges of values for a given subitem. "ANY" makes it possible to search within a repeating subitem for one code of many, for example to discover whether there are any records having lead listed in COMM1. "RANGE" is useful in searching for records that occur within a specified range of values. "SUBSET" has been found useful for filing retrieved sets of data that will be input into FORTRAN or other programs for computation purposes.

REPORT (*see* Fig. 1) is frequently used to generate rudimentary tables of subitems having captions positioned at the heads of columns. There is no facility to format these other than the default spacings but the resulting output can be obtained readily and has proved very valuable for editing purposes. Such errors as missing subitems, mispositioned subitems, and alphamerics in numeric fields and vice versa are readily identified.

```
MARS VI V2.1   05/01/72
LOAD DATA BASE FROM DTBS429:
SAVE/LOAD COMPLETED
RETRIEVAL:
REPORT ACNO,NAME,NTSNO BY HIGH ACNO,
WHERE C1 EXISTS:

ACNO      NAME                                NTSNO


 8202     CHRYSLER LEAD PPTY                     31
 8203     ROCKY POND                              1
 8204     SHOAL COVE POND                         1
 8205     DEVIL'S KITCHEN                         1
 8206     LUNCH POND CLAIM                        1
 8207     ANCHOR DROGUE                           1
 8208     CRANBERRY BOG                           1
 8209     SPICES COVE                             1
 8210     HAYPOOK POND VEIN                       1
 8211     MOUNT MARGARET SPAR FLOAT               1
 8212     CHAPEAU ROUGE VEIN                      1
 8213     GRAND BANK SULPHIDE                     1


--- END OF RETRIEVAL ---
EXIT:
```

FIGURE 1

REPORTER (*see* Figs. 2 and 3) is a much more sophisticated module that permits the formatting of tables and the placing of titles, headings, and data lines, and footings. Arithmetic operations on data items can be performed as well. The spacings between lines and between columns can be controlled and blank lines can be inserted in tables making it convenient to build large tables across more than one sheet of listing. This is especially valuable in preparing tables for final reports.

```
MARS VI V2.1  05/01/72
LOAD DATA BASE FROM DTBS429:
SAVE/LOAD COMPLETED
RETRIEVAL:
QUALIFY ACNO.CHEMEL(1).PROD(1).CHEMEL(2).PROD(2).CHEMEL(3).PROD(3).
CHEMEL(4).PROD(4) BY HIGH ACNO:
--- END OF RETRTEVAL ---
REPORTER:
FORMAT;
LINES PER PAGE = 60:
H IS ACNO.CHEMEL.PROD.CHEMEL.PROD.CHEMEL.PROD.CHEMEL.PROD:
D IS ACNO.CHEMEL.PROD.CHEMEL.PROD.CHEMEL,PROD.CHEMEL.PROD:
GENERATE:
```

ACNO	CHEMEL	PROD	CHEMEL	PROD	CHEMEL	PROD	CHEMEL	PROD
6408	NB	0		0		0		0
7443	AG	0		0		0		0
7468		0		0		0		0
7469		0		0		0		0
7470	BA	0.6615		0		0		0
7471		0		0		0		0
7472	BA	1.2		0		0		0
7473	BA	0		0		0		0
7474	BA	4		0		0		0
7475		0	F	0		0		0
7476		0		0		0		0
7477		0		0		0		0
7478		0		0		0		0
7479		0		0		0		0
7480		0		0		0		0
7481	BA	5		0		0		0
7482		0		0		0		0
7483	BA	003817	CU	002884	PB	019532	ZN	009144
7484		0		0		0		0
7485	BA	0		0		0		0
7486		0		0		0		0
7487		0		0		0		0
7488		0		0		0		0
7489		0		0		0		0
7490		0		0		0		0
7491		0		0		0		0
7492	BA	000376		0		0		0
7493		0		0		0		0
7494		0		0		0		0
7495		0		0		0		0
7496		0		0		0		0
7497		0		0		0		0
7498	AG	0		0		0		0
7499	AG	0		0		0		0
7500		0		0		0		0
7501		0		0		0		0
7502	F	000025		0		0		0
7503	PB	00.026		0		0		0
7504	BA	.00002		0		0		0
7505		0		0		0		0
7506		0		0		0		0

FIGURE 2

```
REPORT COMPLETE
RETRIEVAL:
QUALIFY NAME,PROVCOD,NTSNO,SUBNO,ACNO BY HIGH NAME:
--- END OF RETRIEVAL ---
REPORTER;
FORMAT;
LINES PER PAGE = 60;
TITLE IS INDEX TO CANADIAN OCCURRENCES, "D:
H PROPERTY NAME,PROVINCE,NTSNO,SUBNO,ACNO;
SPACE TO 31,41,48,57 IN H1 BEFORE I2,I3,I4,I5:
D NAME,PROVCOD,NTSNO,SUBNO,ACNO:
SPACE TO 31,41,48,57 IN D1 BEFORE I2,I3,I4,I5:
GENERATE;
```

INDEX TO CANADIAN OCCURRENCES   27 MAR 74

PROPERTY NAME	PROVINCE	NTSNO	SUBNO	ACNO
	QUE	31	F15E	7687
	BC	82	K 2E	7734
	QUE	31	G12E	7942
	QUE	31	F 9E	7946
	QUE	31	F 9E	7947
	QUE	31	F 9E	7948
	MACKENZIE	75	L 8W	8015
	ONT	52	A10	07913
	ONT	52	A15E	7917
	QUE	31	J 6	8108
A J MACLEAN VEIN	NS	11	K 3E	7477
ACHOOK IS	MACKENZIE	85	K 5W	7578
AILPORT PROSPECT	BC	93	L 8E	7729
ALASKA HWY MI 397	BC	94	K10W	7539
ALBERT CANYON-SILVER CREEK	BC	82	M 4W	7858
ALLIED MINES LTD	BC	92	I16	8085
ALMA TP MATACHEWAN	ONT	42	A 2	7517
ALMA(MYRTLE) GRP	BC	82	F	7854
AMEDEE COMUS	NS	11	K10	7500
AMHERSTBURG	ONT	40	J 3	7783
ANCASTER	ONT	30	M 4	7510
ANCHOR DROGUE	NFLD	1	M 3	9007
ANDERSON VEIN	ONT	52	A10E	7726
ANDERSON-BINGHAM PPTY	ONT	52	A16	7836
ANNA GRP	BC	82	F	7853
ANNIE LAURIE CL	BC	103	I 5	7731
ARISAIG AGLT QUARRY	NS	11	F16E	8184
ARIZONA CR	YUKON	116	A 2	7584
ARSENO GRP	MACKENZIE	85	J W	7715
ASHNOLA-PETERBOROUGH GRP	BC	82	E15	7859
ATAN-AUGUST CL GRPS	BC	104	P 3E	8068
AUGUSTA MINE	ONT	52	A 5	7790
AUMOND	QUE	31	J12	7938
BADGER MINE	ONT	52	A 5	7802
BAIE ST PAUL MOULIN R	QUE	21	M 7E	7909
BAILEY MINE	ONT	31	C11	7522
BALMORAL MILLS I	NS	11	E 6	8141
BALMORAL MILLS 11	NS	11	E 6	8142
BARITE FAULT	NFLD	1	L14W	8156

FIGURE 3

```
MARS VI V2.1  05/01/72
LOAD DATA BASE FROM DTBS426;
SAVE/LOAD COMPLETED
RETRIEVAL;
PRINT RECORD WHERE C1 EQ 7616;
RECORD= 7616 2F  BA    20B          72  950274KRD    CORAL PROSPECT             5
                                    RABBIT RIVER SHEET ALASKA HWY LIARD SP75930001260600 93M 8E              7M N LIARD R BR

                                F  BA               FLRT  WTRT  BRIT  BRCL

                                    F     1

                          2 2   70600     3

                        SHLE2  2                                                      LMSN3  2

                      BESA RIVER FM        2                                        DUNEDIN FM          3

                   S 250000 62    1

              72  LIARD FLUORITE MINES LTD

                                                                       1972WOODCOCK

                                                                  LMSN REPLACED ALONG FAULTS                                 FOLL
  OWS LMSN-SHALE CONTACT DOWN LO DIP S UNCER SHALE        12DDH 200FT GRID 30FT 35PC CAF2 70FT 36PC CAF2                  87FT
  39PC CAF2 COMPLEX N-S ANTICLINAL STRUCTURE             REPLACEMENT LMSN BRCC FILLS SHLE BRCC COLLAPSE STR

  --- END OF RETRIEVAL ---
  PRINT RECORD WHERE C1 EQ 7650;
  RECORD= 7650 2AG BA    30B          72  951073KRD    RUBY GRP TORIC MINE         1
                                    HASTINGS ARM SHEET UPR KITSAULT VALLEY75541301293030103P12               17MI FROM ALICE AR
  M
                                BA PB AG             BRIT  PYRT  GLEN  SLVR

                              JSPR HMTT ANKR QRT* PYRT
                            30.5333AG2829                                          29.1072PB2829

                            33   28 80120             53

                         FLST 3 3  BRCC 3 3

                       HAZLETON GRP          3      12 2

                   S  50000 72                    1

              2427CONSOLIDATED HOMESTAKE M DEV                              2729TORIC MINES LTD
             29  TORBRIT MNG CO
```
FIGURE 4

The PRINT (*see* Fig. 4) instruction also produces listings but the organization of the output differs from REPORT in that the subitem names appear one above the other along the left margin of the listing with the retrieved values positioned to the right of the names. It is less convenient for comparing the contents of subitems and has not been as useful for editing purposes.

The SUBSET instruction is useful to reformat the retrieved subitems and place these on permanent file in preparation for use as input to computational programs. This is proving useful in preparing input data for plotting index maps of the data in the file.

Research Benefits

Each deposit is described in a precis or summary that is available in the computer based file. This facilitates the filing of standard sets of data for each deposit resulting in a greater efficiency in research and manuscript preparation.

Parts of whole records can be retrieved with greater precision and output in printed form in a fraction of the time required to make a comparable retrieval from a conventional file. The saving in manhours and the manual transcription of the data from file to output form provide significant advantages.

Data sets are readily maintained as a result of the discipline required to complete the collection forms and by the record (data set) printout capability. The former makes it more difficult to leave out subitems, the latter clearly identifies the missing subitems for the next update run.

As a result of retrieval software the computer based file has multiple cross references in contrast with the single reference system in the conventional manual file. Data can be retrieved as printed output sequenced by any of the available subitems or by a hierarchical sort based on more than one subitem. This is a tremendous aid to research procedures.

Numerical data can be retrieved and then processed by means of arithmetic or statistical procedures without any need to copy the data manually or to personally do the computations. This is done without manually transcribing data with obvious advantages.

FILE MANAGEMENT SYSTEMS AND GEOLOGICAL FIELD DATA

Terry Gordon,
Regional and Economic Geology Division

Gwynneth Martin,
Computer Science Centre, Department of Energy, Mines and Resources

Introduction

This paper is intended as a general introduction to the concepts of data management for field geologists with little or no experience with computers. It does not deal with the myriad of scientific applications of computer technology to geological problems - some of which are described elsewhere in this publication. Hopefully, it will provide some insight into both the institutional commitments and individual effort required to apply current computer techniques to storage and retrieval of field data.

Classical Geological Mapping

A simplified view of geological mapping suggests six phases of activity. These are:

1. Determining the types of data to be recorded in the field.
2. Establishing some standards for note-taking to keep the observations of various geologists as consistent as possible.
3. Mapping and recording field observations and modifying 1 and 2 as required.
4. Checking the data for blunders, then storing and retrieving various pieces of information in order to compile the map and report.
5. Using the retrieved data to produce summaries, perform computations, and generate plots in order to define map-units and their relationships.
6. Preparing final map and report.

Particularly in projects with large quantities of data and/or map-units separable only after considerable office work, steps 4 and 5 can become tedious and time consuming. It is at these stages of a mapping program that computer technology can provide considerable aid to a field geologist.

Field Data Systems

Details of many field data "systems" have been described, for example Wynne-Edwards et al., 1970, Roddick and Hutchison, 1972; Reinhardt and Jackson, 1973; Platou, 1971; Berner et al., 1971a, 1971b; Gaál and Suokonautio, 1973. Most published systems deal with all of the six steps outlined above. The authors describe the data items recorded, codes, input forms and various computer programs to store, retrieve, manipulate and plot the collected data. Such systems are generally efficient when applied to the mapping projects and geological problems for which they were designed. Many, however, suffer from being data dependent, that is, the computer programs written for storage, retrieval, manipulation and plotting will work only for field data collected with the forms and codes specified for that particular system. The disadvantages of such data dependency are obvious. A geological decision to modify the format or kinds of field data recorded may require extensive and costly re-programming. If areas of strongly contrasting geology are being studied, it is difficult, if not impossible, to design a simple input document that will handle all possible mapping situations.

Data Management

One solution to this problem is to explicitly separate the geological data recording tasks (steps 1 to 3) from the data management problem (step 4) and separate that in turn from data manipulation and display (steps 5 and 6).

In other words, it would be desirable to be able to design a field data checklist or input form and collect field observations without concern for how the data are to be stored. It would also be desirable to select, sort, print, manipulate, and plot various pieces of information independently of the way in which they were originally recorded. This is not to imply complete independence of field data collection and office work. Field observations must be made with foreknowledge of the types of office study to be undertaken. And no amount of sorting, searching, and plotting will recover data that were not originally recorded. The connections between the notes as collected and office manipulations are, however, logical ones. What it is necessary to do is to separate the details of data recording (e.g. what codes are used and in what format) from the computer programs that do the general tasks of data storage, retrieval, and updating - i.e. data management. Similarly, post-retrieval data manipulation and display should be considered as separate operations from data management per se.

File Management Systems

There are now many commercially available file management systems - packages of programs that simplify the problems of data management and, more importantly, permit the various operations of data storage and retrieval to be easily performed on widely varying sets of data. The remainder of this paper will discuss the various operations performed by most file management systems, their specific application to geological field data, and the general implications of using such an approach in a mapping organization.

The basic steps in building a file and using any file management systems are:

1. File definition - organizing the form of the data and describing it in computer terms.
2. Data loading - transferring the data from its initial form (usually punched cards) to the storage medium from which it will be used (usually tape or disc).
3. Data validation - checking the data for gross errors or blunders - incorrect codes, numeric values out of range, misspellings, or keypunch errors.
4. Updating - adding, deleting, or changing data in the file.
5. Retrieval - extracting from the file desired data items.

The first process, file definition, requires that field notes be organized into records. For purposes of most field work, the notes taken at a single station form a single record, although such notes may include inter-station observations. Each record in turn will consist of a set of data items - usually corresponding to a single identifiable observation or a comment on some particular aspect of the geology. Station number, dominant lithology, strike of axial plane are all data items. Most file management systems permit repeating groups which are sets of related data items that may vary in number of occurrences from record to record. For example, a repeating group called "planar features" might consist of the data items "type of surface", "strike", and "dip". At a single station the particular group of data items might be recorded anywhere from zero to one or two hundred times, depending on geology and geologist. Most systems require, however, that the maximum number of occurrences of any particular repeating group be specified in advance. In fact there may well be an upper limit to the total number of characters permitted in a record. For most applications a record length of up to 4,000 characters should be sufficient.

A few systems permit several levels of hierarchy in repeating groups so that, for example, a variable number of mineral names might be recorded for each of a variable number of rock names at a single station. Although such a feature has immediate appeal, in practice it becomes difficult to use such a "multi-level hierarchic file" - primarily because the hierarchical relationships add a great deal of complexity to retrieval requests.

Depending on the system, the geologist will have to specify:

1. A name for each data item.
2. The length (number of characters or digits) of each data item.
3. The type (alphabetic, numeric, or both) of each data item.
4. The maximum number of occurrences of each repeating group.

File definition is an iterative process. The first field note recording form designed for a project will have inadequacies. The best procedure is to create a small file, even with artificial data, and attempt to work with it, then revise the file definition as needed. Once a satisfactory record design has been obtained, a computer job is run which loads the data (see below) using the definition thus permitting other file management programs to access the file.

Data loading (sometimes called file generation) is a computer job, or more commonly a series of computer jobs, in which the data, usually on punched cards, are read into the machine, checked against the data definition and stored either on magnetic tape or magnetic disc. This makes the data available for subsequent updating and retrievals. It may be necessary to preprocess the data punched directly from field data sheets to put it into a form recognized by the file management system. This usually consists of running a small computer program written specifically for the input forms and file management system used.

At the loading stage it is important that data validation be performed. This consists of checking each data item against either a permitted set of values (for example a list of lithology codes) or a permitted range of values (e.g. 0-360 for "strike"). Blunders and inconsistent note-taking in the field, illegible writing, and keypunching can all contribute to errors that may render notes unusable by the computer. In fact it is advisable that manual checking of both field sheets and the data as keypunched be performed even before it is processed by the file management system. This usually consists of quickly scanning a print-out of the data cards. A small amount of time invested at this stage can minimize frustrating delays caused by trying to enter "dirty" data into an unforgiving system.

Fast retrieval of specific data items is the prime reason for using computer technology in compiling a map. A retrieval request usually consists of two parts - a qualifying clause which specifies the records (stations) of interest - and a selection clause which specifies that particular data items recorded at that station be retrieved. The qualifying clause IF LITHØLØGY EQ GRANITE would recover all stations where the characters "GRANITE" had been recorded in data item "LITHØLØGY". This could be combined with the selection statement PRINT NØRTHING, EASTING so that the locations of such stations would be retrieved. An even simpler retrieval request has no qualifying clause. In this case the selected data items for every record on the file are retrieved. In addition to printed output, most file management systems permit the creation of a computer processable output file. These data can then be passed directly to other computer programs - for example to automatically plot a map.

In specifying retrieval requests, decisions made while defining the file have a great deal of influence. Two important concepts should be remembered. First, a data item is the smallest accessible unit of information. It is not possible to test for or retrieve only part of a data item - i.e. a single word within a data item "comments" cannot be used in a qualifying clause. Secondly, field data can be divided into two classes. Those data items that will appear in qualifying clauses must have limited and defined values. Without this restriction it is impossible to write meaningful search criteria - simply because the data item being tested has undefined values. Conversely, there are data items that will never appear in qualifying clauses, but will be required as retrieved data. Data item "comments" is an example, and "comments" may contain any legitimate characters or numbers, even gibberish. It is possible, however, to use a limited set of codes in a separate data item to identify the subject matter of "comments" (*see* Hutchison and Roddick, this publication) and thus retrieve an otherwise undefined data item[1].

The qualifying clause is the link between the scientific questions posed by the geologist and the field data itself. Obviously, the more powerful the retrieval capability of the file management system, the more useful it is. Most systems permit the testing of individual data items for: presence or absence of data, equality, inequality with a specified value, and relative magnitude (greater or less than) to a specified value. Such tests can usually be linked by AND, and ØR to permit multiple conditions to be tested simultaneously.

Updating is the process of changing, deleting or adding values of either individual data items or complete records in the file. Most systems permit both complete updates - in which the same data item in each record is modified - and conditional updates in which a condition similar to that of a retrieval request must be satisfied before a record is modified.

A conditional update might thus take the form "IF MAP-UNIT EQ 12 UPDATE AGE = TRIASSIC". This command would cause data item "AGE" to contain the value "TRIASSIC" in all records where data item "MAP-UNIT" contained the value 12. In field data files, updates could include both correction of errors detected during validation and the addition of new information to the file.

The basic file management operations - definition, validation, loading, update, and retrieval - must be performed on any file, whether a general file management system or custom designed programs are used. Use of a general system eliminates redundant programming as well as permitting easy modification of file structures as requirements change. Our suggestion is that most mapping organizations have several different applications of file management techniques, both in field data and other areas, hence would benefit by using a general system.

Institutional Requirements

Computer access

What then are the minimum requirements for a geological mapping organization to implement computer-based field data processing? First, in terms of hardware - actual equipment - there obviously must be access to a computer, i.e. the facility to submit computer jobs and receive output.

It is usually possible to submit jobs to the machine in two distinct modes: -

1. Batch

Traditionally this would involve making up a deck of cards which would define the job to be done. The deck would be read into the computer and at some later time (dependent upon what the job involved and upon the operational parameters of that particular installation), a listing will be produced. The job is wholly defined by the card deck and there is no access to the computer possible during the execution of that job.

2. Time-sharing

In this case, access to the computer is through a terminal, usually with keyboard, such as a teletype. An instruction may be given to the machine by typing a command, the computer will obey (often a response will be typed back) and another instruction can then follow. Obviously the response to one instruction may influence the next instructions i.e. the 'job' need not be wholly defined before typing begins. Clearly this mode of operation allows interaction between the user and the computer.

Due to physical limitations, an installation may only have one access option. However, in an ideal situation, both modes of access would be possible and, for any job, the user may choose between them. In general, it will cost more to use time-sharing; the actual computing charges are higher and the man-hours spent running the job greater. If interaction is beneficial then extra cost can be justified.

Consider the data management functions which have been outlined - the loading process involves putting the data onto computer storage. There may be large amounts of data. Obviously, typing it on a terminal connected to a large expensive computer is not economic, so the loading process is more sensibly done as a batch job including the data on cards. Time sharing may, however, be particularly useful for retrieval. For example, the response to a particular retrieval request may suggest the form of subsequent retrievals that can be immediately requested.

This "browsing" through the file can be extremely productive. The same process using batch procedures would involve several job submissions with resulting delays waiting for the batch jobs to be returned. The process of validation may also lend itself to interactive use - using teletype commands to inspect parts of the data and modify it, if necessary, using update commands.

Graphics

Graphical display of geological data will always be the primary technique for both analysis of observations and communication of data and interpretations. Such graphics might include interactive packages described by Crain (this publication) and Picklyk and Ridler (this publication). Less expensive and larger displays are provided by computer-driven plotters.

There are two types of plotter in general use. A drum plotter, as the name implies, produces graphical

[1] Use of a special computer program called a text-editor will permit 'character-string searches' i.e. searching for a word within a data item. This, however, involves at least an extra step in the retrieval process.

output on a continuous roll of paper or other base material. A flat-bed plotter produces output on a single flat sheet. Both types are useful, although use of a flat-bed plotter has the advantage of plotting data on previously prepared base material - for example a topographic map. Whatever graphics facility or service is chosen will require that the computer installation maintain a set of programs to support such hardware.

Digitizing

Use of automatic plotting techniques depends, of course, on the pre-recording of topographic location information for each record in the field data file. In Canada, the Universal Transverse Mercator (UTM) grid, printed on most topographic maps, provides a convenient reference system. This information normally not recorded in a field data file, can be obtained by manual scaling in the field, or by using a digitizer, a device for automatically recording X and Y co-ordinates of lines and points as selected from a base map by an operator.

Access to a digitizer can provide a mapping organization with a rapid and relatively error-free method of obtaining digitized locations of previously plotted stations. There are the additional expenses of paying for a digitizer operator as well as running computer programs to convert digitizer output to actual topographic co-ordinates and merging this data with the field data file.

File Management Systems

Access to a computer implies access to a certain number of 'package' programs, which may include a general file management system. To be of use in managing field data files, the system should have at least the features mentioned in our description of general data management operations. If such a system is not provided by the computer installation, then it must be purchased or leased from another source. Several commercially available systems provide most of the required functions; some provide additional capabilities. Almost inevitably any one system will be lacking some feature which may seem very attractive, so there may be a temptation to wait until the 'perfect' system becomes available. It is our opinion that it never will. Once a system is put into operation, features which were not obviously useful will be discovered; attractive features will be found to be deficient in some way; and, possibly most important, the user will discover those features which are most important in her/his application. The important thing is to get a system that works, that satisfies most of the organization's requirements and then use it.

Data base management systems

As said above, there are several file management systems currently available. Also available, and aimed at the same general type of problem (information handling) are data base management systems and although the two terms are often used synonymously, the two types of systems should not be confused. A data base management system is much more complex in the following particular respects: -

1. It aims at efficiently handling many files which have items of information in common, by using the common items as a base. In other words, although the files may be considered as consisting of separate records, the data items they have in common are stored only once, the system providing the connections between different items for different files.
2. It is designed to serve many users and uses. A field data file is primarily for the use of one or two project geologists. A data base, for example of mineral deposits data, could serve regional geologists, exploration geologists, economists, and planners, each with their own requirements for specific pieces of data, but with some data items common to all their applications.
3. The data handled are described in 'structural' terms. What has been called a data item would be defined in terms of its relationship to other items in the data structure, i.e., the information is considered in a hierarchical fashion.

Although these systems may seem attractive as a long-term goal, their complexity necessitates a greater degree of knowledge of the system itself for efficient and effective usage. Such systems are expensive to acquire and operate. Field data file handling as described in this paper does not warrant that level of sophistication.

Personnel

The software and hardware requirements outlined imply another requirement - people. These do not have to be employees of the mapping organization, but could work for the computer installation. Obviously there must be personnel to actually operate the equipment. However, when considering the use of a package such as a file management system, there are at least two other computer-related functions which must be provided: -

1. Question-answering Service.

   Usage of a system will give rise to problems which will not be due to the nature of the data file, but instead due to the file management system itself. For example, the supplied documentation may not adequately explain the use of particular features. If a problem arises of this type then there must be access to expertise on the general use of the system independent of applications. Provision of such expertise is normally the function of the computer installation, not the mapping agency.

2. Maintenance.

   A package of this nature is subject to revision by its manufacturer; the computer on which it runs is also subject to changes (of both hardware and software). Again these are tasks in the realm of the computer installation rather than the mapping group.

In addition to personnel for computer systems support, there is a requirement for 'middle-men'; people who have some knowledge and interest in the computer systems, but even more importantly, have a good understanding of the application of these systems to the specific problems of geological mapping. That is, there is a definite requirement for a "data-manager" or "geologist-programmer" to provide assistance to individual geologists. Such a person would:

1) provide instruction as to the computer and file management facilities available and their use;
2) assist geologists in the design of their own file;
3) perform the data loading and possibly updating tasks of file management;
4) assist in retrievals of a complex nature;
5) maintain a library of computer programs to perform standard operations such as map plotting, norm calculations, or stereo net contouring.

Such a person should be employed by the mapping organization itself.

Individual Effort

The effort required of a single field geologist to "computerize" depends on the level of support he receives from the data manager and computer installation. At one extreme he may merely hand his problems to a data processing group and wait for results, at the other he may be required to do extensive programming himself.

Education

At a minimum the geologist must learn how to clearly describe his data, formulate retrieval requests, and specify the types of computations and plots he requires. One route to this goal is to learn simple programming in a scientifically oriented language (for example BASIC or FØRTRAN), followed by training in the retrieval language of the file management system in use. Eight days of instruction spread over three weeks can provide sufficient background, actual use of the system providing the remainder of knowledge required.

Changes in work style

The organization of field notes into station records does not differ drastically from most classical notebook methods. There are, however, at least four areas in which computer-based systems require departures from normal procedures.

1. Station locations are traditionally recorded in analogue form as dots on photographs and base maps. Although this can still be done with a computer-based system, the addition to the file of location information in digital form permits the exploitation of many automated plotting devices. This benefit is gained at the expense of additional effort in recording and editing the extra data required.

2. Traverse summaries, generalizations, and observations of trends cannot be easily referred to a single station location.

3. Sketches, photographs, and lengthy or unusually detailed descriptions have to be stored separately from the bulk of the field data. This problem, and 2 above, can be overcome by creating sub-files that are linked to the main file by common data items - such as station number.

4. Computer use requires that geologists and their assistants be consistent and follow specific rules, albeit of their own making, for note-taking in the field. Because the notes are to be "read" by machine, there can be no ambiguity in the spelling of data items that are to be used in retrievals. For example, a request for all stations at which ultramafic rocks were observed will not provide a correct answer if there are misspellings of the code for this rock-type. This does not imply that a single code must be defined as the correct one, but that whatever codes are chosen are used consistently within a particular file.

It is unlikely that the first year's attempt at using computer technology in field mapping will be economic either in dollars or geologist's time. A beginner must expect to invest extra effort in his project to "break even" during the initial field season. Careful planning can minimize this effort and, with the experience gained, produce real benefits in the second and following years of a mapping project.

CONCLUSION

This paper has been necessarily general, introducing terms and concepts rather than outlining a definite course of action for applying computer techniques to field mapping. The details of the approach for any particular mapping organization will depend on the resources and requirements of that organization. By first carefully defining the organization's goals and secondly exploiting well established computer techniques an effective merger of field geology and modern computer technology can be attained.

References

Berner, H., Ekström, T., Lilljequist, R., Stephansson, O., and Wikström, A.
1971a: Data storage and processing in geological mapping. I Field data sheet: Geologiska Föreningen i Stockholm Förhandlingar, v. 93, p. 86-101.

1971b: Data storage and processing in geological mapping. II Data file: Geologiska Föreningen i Stockholm Förhandlinger, v. 93, p. 693-705.

Gaál, G., and Suokonautio, V.
1973: An automatic data processing system for exploration mapping in Precambrian terrain: GEØKU, Finl. Comm. Geol. Bull. 266.

Platou, S.W.
1971: An electronic data processing system for geological field and laboratory data. The E.D.P. system Agto: Grønlands Geol. Unders. Rept. 39.

Reinhardt, E.W., and Jackson, G.D.
1973: Use of a geological field data collecting form on Operation Bylot, North-central Baffin Island, 1968: Geol. Surv. Can., Paper 73-35.

Roddick, J.A., and Hutchison, W.W.
1972: A computer-based system for geological field data on the Coast Mountains Project, British Columbia, Canada. 24th Int. Geol. Cong., Session 16, p. 36-46.

Wynne-Edwards, H.R., Laurin, A.F., Sharma, K.N.M., Nandi, A., Kehlenbeck, M.M., and Franconi, A.
1970: Computerized geological mapping in the Grenville Province, Quebec: Can. J. Earth Sci., v. 7, p. 1357-1373.

6. COMPUTER STORAGE AND RETRIEVAL OF ABSOLUTE AGE DATA

R.A. Frith,
Regional and Economic Geology Division

At the present time several active files on Canadian geochronological data are being kept (University of Toronto and Geological Survey of Canada). Neither of these files are readily available to users. This prompted the writer in 1972 to begin a separate file on absolute ages of the Canadian Shield and nearby localities that could be published as a compilation every few years. The present file is being kept up to date for such a publication (e.g. Weatherill et al., 1965).

In the past, attempts to co-ordinate absolute age data on a national scale were not successful. A Canadian Isotopic Age Data Centre, developed at the Geological Survey of Canada, in response to national and international attempts to centralize data (International Council of Scientific Unions) relied heavily on voluntary data input that was not forthcoming (Wanless, 1968).

In order to overcome these problems, a simple system of storage and retrieval was adopted with data gleaned from existing files and updated from abstracting and CAN/SDI services. The system has the following features:

(1) Age data input is restricted to a one-card format which forms one file.
(2) Each piece of age data is indexed to a second reference file.
(3) The data are stored in increasing chronological order and may be retrieved on any one or combination of the following parameters:

a) Absolute ages ranges;
b) Dating method;
c) Locations between longitudes and latitudes;
d) Rock type;
e) Geologic Province;
f) NTS grid number(s).

(4) The data are amenable to computer plotting.
(5) The output lists original references, complete with page number.

References

Wanless, R.K.
1968: Isotopic age data file reporting format and codes; Geol. Surv. Can., Internal Circular.

Wetherill, G.W., Bickford, M.E., Silver, L.T., and Tilton, G.R.
1965: Geochronology of North America; Nat. Acad. Sci., National Research Council, Washington, D.C. Publication No. 1276.

7. SUB-AREA RETRIEVAL SYSTEM USED ON THE COAST MOUNTAIN PROJECT OF THE GEOLOGICAL SURVEY OF CANADA

W. W. Hutchison and J. A. Roddick,
Regional and Economic Geology Division

In late 1971 the Geological Survey of Canada initiated a program of pilot projects to evaluate the usefulness of computer-based techniques for collection, storage and manipulation of geologic field data and information. The following is a description of the most highly evolved of these projects.

The system used in the Coast Mountains Project had been in existence since 1965 (Hutchison and Roddick, 1968). Major evolutionary stages in this system have been:

1965 - Computerization of limited amount of field data.
1969 - Plotting of strikes and dips on stable base to go directly to the printer.
1970 - Use of free text to complement data (all data and handwritten notes were keypunched at a cost of 2% of field budget).
1973 - Compilation (using interactive terminal) by instantly retrieving the complete data base for a map-unit sub-area at one command and by interrogation, data reduction and synthesis of the generated sub-file.

For the field geologist the last two stages of evolution were the most beneficial. Because these are potentially important and fundamental and because they are commonly lacking in systems developed to date, this report illustrates these aspects as an example of one pilot project of the Geological Survey of Canada.

Coast Mountains Project

The Coast Mountains Project of the Geological Survey of Canada was set up in 1962 to conduct geological reconnaissance mapping of the Coast Mountains in British Columbia from latitude 56°N to Vancouver (near 49°N). The area comprises about 43,000 square miles (110,000 km^2). So far data and information have been collected for 25,000 stations and 25,000 specimens. This information pertains to possibly the largest post-Precambrian plutonic complex in the world so it was decided some effort should be made to systematically synthesize the vast quantities of data and information and maintain these for further study.

Evolution of System

The development of the system is documented in Roddick and Hutchison (1972).

Current System

Input

The present system using pre-printed check sheets using standard 80-column cards is still in use (*see* Roddick and Hutchison (1972) for codes). An example of the current field station input document (Fig. 1) is shown to illustrate the much greater allowance now made for recording free text. Station locations are measured at base camp using Romer measuring grids on a map printed with a Universal Transverse Mercator grid. A separate input document is used for petrographic information such as rock specimen data, specific gravity, mineral abundances and revised rock names (*see* Roddick and Hutchison, 1972, Fig. 2).

Editing

Editing is the most time consuming part of this endeavour. Prime information can usually only be checked by the field geologist and this results in his being severely handicapped. At present we do not have a wholly systematic approach to editing.

The three aspects of editing are as follows - editing of station locations, editing for invalid codes or numbers and scanning of some prime data and comments.

Editing can also be done during retrieval and this is described below.

Retrieval

For economical access to our now large master file, a system of on-line retrieval, called SARS (Sub-area Retrieval System) was developed to run on the University of British Columbia's IBM 360-67 under the MTS (Michigan Terminal System) operating system. SARS, however, could be adapted to other comparable operating systems, as the techniques developed in its implementation are applicable to more general systems.

SARS consists of a master data file, several ancillary files, the on-line retrieval program, and file maintenance programs. It includes various features of the MTS operating system, particularly the file-handling capabilities and the MTS editor.

The master file represents a merge of both the station file and the petrographic file. It is a sequential file with line numbers related to station numbers. Previously developed programs permit us to plot most of the hard data, such as attitudes, dykes, mineral occurrences, etc., on either a drum or flat-bed plotter (the latter being used for direct plots onto a topographical base map).

To manipulate the information in the master file and to retrieve selected parts of it, a program called SUBMON was developed. This program performs operations on the master file consulting where required a file called 'Subspecs' which contains a list of stations pertaining to a lithologic unit. This file must be

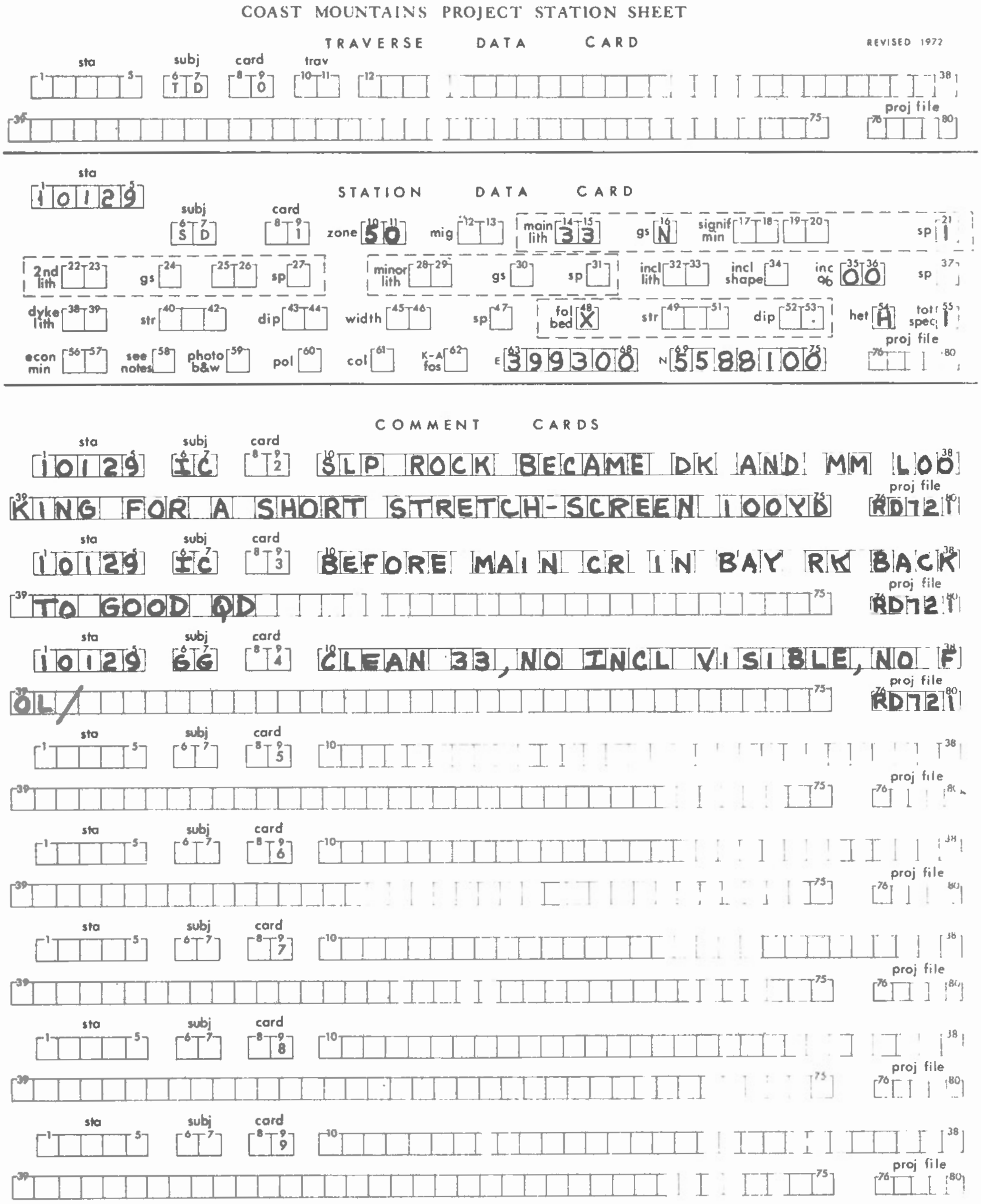


Figure 1.

Typical input document for geologic field data. This sheet corresponds to lines with asterices in Fig. 2 Note however this document does not contain petrographic data which was merged and added into field data file as shown in Fig. 2.

previously built by the user who selects and inputs the station numbers. This onerous task is somewhat simplified by a provision which makes ranges of consecutive stations acceptable, and could be done entirely by the computer in the type of geological work where the lithologic unit is known at the time the stations are made.

The subspecs file permits listing of all information concerning a rock unit or specific types of information that can be retrieved by simple character-string searches. Conditional searches require short fortran programs, as do calculations on certain specified data and the production of special tables.

Considerable progress has been made in developing editing routines, but a general, verification-edit routine is not presently operational. The Submon program was developed with file maintenance in mind. Editing can be done on any sub-file retrieved. The corrected data are then consigned, by one command, to an update file. Before searching the master file, the program goes first to an update file and if it finds the station in question, then it bypasses the master file containing the uncorrected data. When the update file is large enough, it is incorporated into the master file (the program to do this, however, remains to be written).

Use of SARS

The basic concept of SARS is that it allows the field geologist to mimic the procedure he would normally use in compilation. Through SARS he may retrieve the complete data base for any map-unit or sub-area as part of that map-unit. He can then interrogate the sub-file for all data and comments concerning such topics as

```
*  74    39-58   TOBA QD-GD
  10022SD0150   04CSP   1                                 F33090 1        3987005585700RD701P12690804078103
  10022GG 2 ALMOST AUGEN STREAKINESS IN PLACES. ONCE EUHED BIOT HAS GONE TO F
  10022GG 3 INE GRAINED MAFIC PSEU DOMORPHICALLY
  10023SD0150   05F     1            00                   X       S1      3976005582000RD701P12630213147104
  10023GG 2 VERY WHITISH LEUCOCRATIC ROCK/SUGARY TEXTURE
  10275SD0150   05F     1                                 M33590S1        3977005581000R0701P126302151C7304
  10275GG 2 QM BUT MAYBE INSERTED BETWEEN QD AND GD/QUITE A FEW SCHLIEREN
  10274SD0150   33N     1              E02                G31060M1        3953805580450RD701P12771703057543
  10274GG 2 DEFINITELY NOT THE GD
  30292SD0156   33C     1            AME05                G33090 1        3856605588240RD701P12711201226543
  30292GG 2 STRONGLY FOL 33 -BUT NOT ALTERED
  30293SD0156   33M     1FP    2             CV03090   3G29090S3    2     3879205590450RD701P12690912166344
  30293SD0156   33M     1FP    2             CV03090   3G29C90S3    2     3879205590450R0701D3291        BA
  30293SD0156   33M     1FP    2             CV03090   3G29090S3    2     3879205590450R0701P22660905C97724
  30293GG 2 WELL FOL 33/ MINOR SCHLIEREN/
  30293DY 3 SWARM OF DARK GREEN DY- SYNPLUTONIC W CHILLED MARGINS-SEE
  30293DY 4 PHOTOS/1/3 OC IS FP POSS EQUIV OF PO GRANOD TO W THIS CUTS
  30293DY 5 THE 33 AND SP DY/
  30293PH 6 3-2-25A,26A DIKE W CHILLED MARGINS AND SYNPLUTONIC ÔFFSETS
  30395SD0156   24M     1         *     00 N 29090   2X      H2           3898005591310RD701P12680912097044
  30395SD0156   24M     1               00 N 29090   2X      H2           3898005591310R0701D2289        AN
  30395GG 2 CLEA HOMO 24- V MINOR AP VEINS/ALSO V MINOR EP STRINGERS/
  30395GG 3 DY SWARM WELL EX ON CIRQUE WALL- SOME AS WIDE AS 25FT-ALSO AL
  30394SD0156   24M     1            AMR01                I29090H1        3926805590540RD701P12640615186134
  30394GG 2 CLEAN 24-BUT TX DOES NOT LOOK NORMAL-MAFICS- ANHEDRAL
* 210129SD0150   33N     1            00                   X       H1      3993005588100RD721P12741503156733
* 210129IC02SLP ROCK BECAME DK AND MM LOOKING FOR A SHORT STRETCH SCREEN 100YD
* 210129IC03BEFORE MAIN CR IN BAY RK BACK TO GOOD QD
* 210129GG04CLEAN 33,NO INCL VISIBLE, NO FOL/
 210130SD0150   33N     1                                  E34590 1        3973505587650RD721P12691512155844
 210130IC02PASSED SUDDENLY FROM MASSIVE QD INTO EXTREMELY ELONG AGMATITE SCHL
 210130IC03IEREN GNEISS AND SCREEN GNEI SS/ E4 HAS INCL UP TO 10FT OR MORE IN
 210130IC04LENGTH
 210130GGC5ATP STRONGLY FOL TO GNEISSIC QD/ COARSE PATCHES OF FG MAFICS
 210130SU06THE TRAV SEEMED TO PASS DOWN SECTION THRU A THICK SEQUENCE OF ISO
 210130SU07CLINALLY FOLDED BI HO QTZ SC AND MINOR QTZITE WITH A COUPLE OF ZON
 210130SUC8ES OF GRANITOID RX/ SOME OF THE METASEDS ARE LIMY,SOME LAYERS DIOP
 210130SU09THAT SEQUENCE IS UNDERLAIN BY ABOUT 4000FT OF CONGL THAT CONSISTS
 210130SU12LARGELY OF QD OR GD CLASTS UP TO 3FT ACROSS IN A LT GREY SCHISTOSE
 210130SU13ARKOSIC MATRIX IN WHICH BIOT AND SMALL GARNETS HAVE GROWN/ THE CLA
 210130SU14STS ARE SOMEWHAT STRETCHED BUT HAVE NOT DEV ANY INTERNAL FOL/STRON
 210130SU15G FOL IN MATRIX BENDS AROUND CLASTS/ WE PASSED NEXT INTO QD THAT
 210130SU16MATCHES THE CLASTS,SUGGESTING AN UNCONFORMITY/ BETW 129 AND 130 IS
 210130SU17A 4 5C00FT WIDE ZCNE OF SPECTACULAR EL AGM AND SCHLIEREN GNEISS
 210131TD0016JUNE 9/ TOBA INLET UNCONFORMITY HUNT " CONT OF TR 15/ R.T./ CAVU
 210131SD0150   33N     1            AMR01                 F0C090S1        3997505588200RD721P12782002C96933
 210131AP0292K 40 81
 210131IC03FROM GOOD CONGL NE OF STA 128 TO HERE WE HAVE EXAMINED THE OTCP
 210131IC04CLOSELY,WALKING PART OF IT/ CONG SEEMS TO EXTEND TO THIS SIDE OF
 210131IC05128,BUT IT BECCMES VERY DEFFICULT TO RECOGNIZE BECAUSE MATRIX AND
 210131IC06CLASTS ARE OF SAME MATERIAL, AND ONLY ON CERTAIN W SURFACES CAN
 210131IC07CLASTS BE DISCERNED/ WOULD TEND TO PUT CONTACT IN NO OTCP AREA BET
 210131IC08W TP AND CREEK MOUTH TO NE
 210131GG09ATP CLEAN QD,HIGH QTZ AS IN CLASTS/ V FAINT FOL/ A FEW NEB INCLU/
 210132SDC150   43N     1            AMEC3                 E3409CH1        3967005587150RD721P12703C15134244
 210132AP0292K 39 165
 210132IC03FROM STA 130 TO HERE, SOS WITH SLIVER AM INCLUS/
 210132GG04ATP SOS,BUT BIOT DEF DOMINANT V NARROW SLIVER INCLUSIONS/ QD IS
```

Figure 2. To illustrate retrieval of complete data base for one sub-area, based on one command, this is start of listing for sub-area 74 -- TOBA QUARTZ DIORITE - GRANODIORITE. Lines with asterices appear in Fig. 1.

```
:mts r nea
LINE DELETED
_mts
#r mean 5=-d 6=-74
#EXECUTION BEGINS
#EXECUTION TERMINATED
#1 -74
>    1
>    2
>    3
>    4
>    5
>    6
>    7
>    8
>    9
>   10
>   11
>   12
>   13
>   14
>   15
>   16
>   17
>   18
>   19
>   20
>   21
>   22
>   23
>   24
>   25
>   26
>   27
>   28
>   29
>   30
#END OF FILE
#(
LINE DELETED
```

STN NO	INCL	FOL	HET	SPGR	MAF	KFP	QTZ	ROCK
10022	0	F		2.69	8.	4.	7.	03
10023	0	X	S	2.63	2.	13.	14.	04
10274	2	G	M	2.77	17.	3.	5.	43
10275	0	M	S	2.63	2.	15.	10.	04
30292	5	G		2.71	12.	1.	22.	43
30293	0	G	S	2.69	9.	12.	16.	44
30293	0	G	S	2.66	9.	5.	9.	24
30394	1	I	H	2.64	6.	15.	18.	34
30395	0	X	H	2.68	9.	12.	9.	44
210129	0	X	H	2.74	15.	3.	15.	33
210130	0	E		2.69	15.	12.	15.	44
210131	1	F	S	2.78	20.	2.	9.	33
210132	3	E	H	2.70	30.	15.	13.	44
210133	0	M	S	2.70	18.	5.	10.	33
230206	2	E	S	2.66	12.	14.	21.	44
230207	3	E	S	2.71	13.	3.	6.	43
230208	0	F	H	2.74	15.	12.	19.	24
230209	1	G	S	2.68	21.	17.	12.	24
230210	0	E	S	2.65	7.	13.	12.	44
230212	0	M	S	2.69	9.	3.	17.	43
230213	5	F	M	2.81	21.	0.	2.	2
230214	1	X		2.84	22.	0.	2.	32
250133	0	M	M	2.66	10.	10.	12.	54
250134	20		M	2.68	15.	15.	12.	44
250135	8	F	M	2.64	10.	20.	12.	54
250136	5	M	M	2.68	15.	0.	10.	53
MEAN OF 26 STATIONS				2.70	13.	9.	12.	

Abbreviations

STN No = station number,
INCL = inclusions in per cent of outcrop,
HET = outcrop heterogenite,
H = homogeneous,
S = slightly heterogeneous,
M = moderately heterogeneous,
SPGR = specific gravity,
MAF = total mafic minerals,
KFP = potash feldspar,
QTZ = quartz,
ROCK = rock name - first digit is hornblende/biotite ratio, second digit is 3 = quartz diorite, 4 = granodiorite.

Figure 3.

This is a tabular summary of main features of this sub-area.

```
LINE DELETED
zed -s74DY
Tcol 8
Tm@a@nl /file 'dy'
:  30293DY 3 SWARM OF DARK GREEN DY- SYNPLUTONIC W CHILLED MARGINS-SEE
:  30293DY 4 PHOTOS/1/3 OC IS FP POSS EQUIV OF PO GRANOD'TO W THIS CUTS
:  30293DY 5 THE 33 AND SP DY/
: 230207DY05DARK GREEN SC DYKES FORM 3( SCREENS/ LOCALLY THEY ARE CROSS
: 230207DY06CUTTING
: 230209DY04NOTE DYKES ARE NOT SYN PLUTONIC BUT ARE PREMETA.
: 230209DY06AMPHIBOLITE DYKE SWARMS ARE COMMON THEY GENERALLY HAVE A      } ONE
: 230209DY07CONSISTENT TREND BUT COMMONLY CROSS CUT THE QD LOCALLY THEY   }
: 230209DY08FORM A STOCKWORK BUT NOTE THE DYKES AND THE PLUTONIC ROCK HAVE } STATION
: 230209DY09BOTH BEEN META. THIS PLUTON IS PROB PRECONGL
: 230210DY08DARK GREEN AM DY ARE COMMON THEY GENERALLY FORM SWARMS AND
: 230210DY09MAKE UP 10( OF OTCP
: 230212DY07TWO AGES OF DY PARTIALLY GRANITIZED DY THEN YOUNGER BUT META
: 230212DY08AN DYKE SWARMS SUCH AS AMPHIBOLITE IN SPECD2/ THESE DYKES DO
: 230212DY09NOT HAVE CHILLED MARGINS INSTEAD FINE GRAINED SCHISTOSE MARGNS
: 250133DY05ANDESITE THAT NOW HAS A'FOLIATION RESEMBLING AN AMPHBOLITE
: 250134DY05S FRACTURES ALONG SIDE OF DYS/ THEY IN THICKNESS FROM 6 INCHES TO
: 250134DY065 FEET/ THEY WEATHER TO A DARK GRAY%SOOTY COLOR<CHILL MARGINS
: 250135DY06ANDESITE/ SLIGHTLY PO WITH MUCH EPIDOTE " CHL/ SMALL PEG VEINS WI
: 250135DY07TH FRAGMENTS OF DIKE ROCK ARE A LATER CROSSCUTTING FEATURE
: 250136DY04BLACKISH GREEN WEATHERING/ MUCH EPIDOTE " CL/ MAKES UP 40( OF OT
: 250136DY05CP HERE/ TRENDS WERE SAME AS PREVIOUS STA/
:)
LINE DELETED
```

Figure 4. Retrieval of comments on dykes through searching two letter subject code DY in columns 8 + 9.

```
LINE DELETED
zm@a@nl /file 'su'
: 210130SU06THE TRAV SEEMED TO PASS DOWN SECTION THRU A THICK SEQUENCE OF ISO
: 210130SU07CLINALLY FOLDED BIHOQTZ SC AND MINOR QTZITE WITH A COUPLE OF ZON
: 210130SU08ES OF GRANITOID RX/ SOME OF THE METASEDS ARE LIMY,SOME LAYERS DIOP
: 210130SU09THAT SEQUENCE IS UNDERLAIN BY ABOUT 4000FT OF CONGL THAT CONSISTS
: 210130SU12LARGELY OF QD OR GD CLASTS UP TO 3FT ACROSS IN A LT GREY SCHISTOSE
: 210130SU13ARKOSIC MATRIX IN WHICH BIOT AND SMALL GARNETS HAVE GROWN/ THE CLA
: 210130SU14STS ARE SOMEWHAT STRETCHED BUT HAVE NOT DEV ANY INTERNAL FOL/STRON
: 210130SU15G FOL IN MATRIX BENDS AROUND CLASTS/ WE PASSED NEXT INTO QD THAT
: 210130SU16MATCHES THE CLASTS,SUGGESTING AN UNCONFORMITY/ BETW 129 AND 130 IS
: 210130SU17A 45000FT WIDE ZONE OF SPECTACULAR EL AGM AND SCHLIEREN GNEISS
: 230208SU12AGMATITE " DIORITE 123 ON MIKES TRAV/ 124126 PALE PINK
: 230208SU13APLITIC GD/ 127132 DARK MAFIC PHASES DIOR/QD VARIED AGMAT
: 230208SU14THEN SHEARED V STRONGLY SHEAR FOL GRANOD THEN V SMEARED
: 230208SU15ALMOST LIKE ARKOSIC MATRIX
: 230208SU16NOTE  JIMS GRANOD 128 IS EDNETICAL TO CLASTS %EG. 30200EI<
: 230208SU17FOUND IN CONGL./ NOTE ALSO JIMS PRECONGL GD IS V SIM TO 30207
: 230208SU18%IE HIS 3 SAMPLES ARE SIM TO THIS SECTION<
: 250136SU02LL FOLIATED IS FOUND/ THEN TO THE END WAS A SHEARED PINK AUGEN OR
: 250136SU03THOGNEISS %GRANOD</ ALSO SHEARING OF ANDESITE DIKES IS FOUND/ NOTEE
: 250136SU04D FEATURES WERE THE ELONGATE AGMATITE ZONES AT STA 121 " 132 " TH
: 250136SU05E DIKE SWARM AT STA 134135
: 250136SU07THE TRAVERSE BEGAN AT SNOUT PT WITH A SLIGHTLY FOLIATED QTZ DIORI
: 250136SU08TE/ THIS CONTINUED FOR 3 STATIONS/ NEXT STATIONS WAS A FINE GR PI
: 250136SU09NK APLITIC FOLIATED GRANOD/ UNTIL STA 133 QTZ DIORITE DIORITE WEL
:
```

Figure 5. Retrieval of traverse summaries through searching two letter subject code SU in columns 8 + 9 Note that first 10 lines could be almost read by someone not familiar with the system.

```
 mts
LINE DELETED
zmts
#ed -s74
Ts@a@nl /file 'congl'
: 210130SU09THAT SEQUENCE IS UNDERLAIN BY ABOUT 4000FT OF CONGL THAT CONSISTS
: 210131IC03FROM GOOD CONGL NE OF STA 128 TO HERE WE HAVE EXAMINED THE OTCP
: 210133GG06DIST LARGE PATCHES/ RESEMBLES CLASTS IN CONGL/
: 230208SU17FOUND IN CONGL./ NOTE ALSO JIMS PRECONGL GD IS V SIM TO 30207
: 230209GG04TO CLASTS IN CONGLOMERATE FOUND YESTERDAY
: 230209DY09BOTH BEEN META. THIS PLUTON IS PROB PRECONGL
Tmts
```

Figure 6. Retrieval of all comments on conglomerate through searching everywhere in each line for the character string 'CONGL'.

lithology, mineralogy, structure, grade of metamorphism, migmatites, dykes, relationships to other units, etc. From this point he can summarize the data, edit the data or map boundaries and then re-compile rapidly. The technique is basically very powerful, fast and flexible. Furthermore it contains the logic of a cartographic analysis system.

Examples of retrieval are attached. Figure 2 illustrates the start of the retrieval of the complete data base for a map-unit - namely the TOBA QUARTZ DIORITE - GRANODIORITE and Figure 3 is an example of the tabular summary of the main features of this sub-area. Comments on dykes selected by retrieving on DY in column 8 are listed in Figure 4. Interstation comments appear in Figure 5 and illustrates the value of such textural information which would be lost using codes alone. Character string searches anywhere within free text are also possible (Fig. 6).

Conclusions

The following conclusions continue to be valid. The system requires (a) less time on outcrop, (b) less information is lost through illegible handwriting (c) consistency of recording information is greatly improved because the pre-printed sheets serve as check-lists and (d) manipulation and display of geologic data can be accomplished in many more ways, both for analysis and publication, than now hitherto been thought practical.

Efficiency is greatly reduced by time consumed on editing of input data, by programming efforts on part of geologists and a deceptive amount of effort required for file and data base management. Hopefully the pitfalls and progress of this project will increase the efficiency of successor projects of this nature.

References

Hutchison, W. W., and Roddick, J. A.
1968: Recording geologic field data for machine retrieval and processing; Western Miner, v. 41, p. 39-43.

Roddick, J. A., and Hutchison, W. W.
1972: A computer-based system for geological field data on the Coast Mountains Project, British Columbia, Canada; 24th Int. Geol. Cong., Session 16, p. 36-46.

8. A FORMAT FOR FIELD AND LABORATORY DATA COLLECTION

M.B. Lambert and J.E. Reesor,
Regional and Economic Geology Division

A system for computer processing of field and laboratory data, devised by J.E. Reesor, enjoys considerable success because of the simplicity and flexibility of the input data files. The free format for recording data appeals to those who are reluctant to give up the creative element that manifests geological descriptions in the "hand written" method of recording field data. The system applies to both detailed and reconnaissance types of mapping in all types of geological terrains.

Data are recorded on three formats (5 by 9 inch loose-leaf paper) for use in the field, and one (8½ by 14 inch) for use in the laboratory. These formats (Figs. 1 and 2) include: File 1, Lithology and Structure; File 2, Mesoscopic Fabric Data; File 3, Petrographic Data; and File 4, Traverse Summary. File 1 comprises six cards each having space for 80 characters. The first card of this file contains complete station identification and location and a summary of the units described, their lithologies and mineral assemblages. Each subsequent card has spaces for station identification, category numbers and notes. Category numbers indicate the subject of the notes on each card. A new card is used for each category of data discussed, and distinctive category numbers indicate notes taken between stations (Fig. 1d). Notes are recorded in free format using acceptable abbreviations for rock and mineral names and for structural features. All geologists on a field party use the same rock and mineral abbreviations.

File 2 is a fixed format for recording attitudes of planar and linear structures and fabric elements. Space is provided for a brief description and classification of each element.

File 4 provides for notes summarizing general impressions, interpretations, conclusions or special problems encountered in one or more traverses. Each type of data is recoverable through the appropriate usage of category numbers as in File 1. General traverse location, weather conditions, maps and air photos used are noted at the top of this form, but are not retrievable data.

File 3 is a format for recording petrographic data in the laboratory. The first card features a complete summary of mineral assemblage and their estimated amounts along with station identification data. Subsequent cards are used for describing hand specimens information and features of individual minerals and textural relations. This information is retrievable by reference to abbreviations for minerals entered in spaces 15 and 16, or a Summary Table (Fig. 3) of mineral assemblages can be retrieved automatically.

All data processing is accomplished using simple FORTRAN programs. Data output aids in sorting and collation of vast quantities of field and laboratory data, compilation of maps and preparation of written reports. A complete listing of all field data from File 1 can be easily read and understood without reference to the format of the original input file. A copy can be taken to the field in subsequent years without risking the original notes. The data files are sorted by computer and tables of frequently used data are automatically compiled. Such tabulations may include complete lists of specimens, specimens taken for specific purposes (thin sectioning, chemical analysis, or other laboratory procedures), mineral assemblages, photographs, structural observations and specific mineral occurrences. In addition, all data pertaining to specific quadrants of a map sheet can be collated.

A flat bed plotter is used to post station locations, accompanied by a variety of data such as structural measurements and mineral occurrences, on a base map or overlay of any specified scale. Figure 4 shows a sample of a station posting and structural plot showing planar and linear elements.

Preparation of the written report is greatly facilitated by computer sorting and collation of all data pertaining to each formation and of structural relations, features of economic interest, conclusions, interpretations and problems noted in the field.

The open format of the input data files permits the geologist to make unlimited amounts of verbal description. In contrast to many other systems in use, this format is not a check list of information to be gathered: only structural measurements and station identification are recorded on a fixed format. A separate check list of information to be observed may be appended to the field note book, to aid in the completeness of data collected by each geologist. A valuable part of this system is the provision for recording and retrieving subjective and interpretative data, data gathered between stations and impressions arising from data gathered over a range of stations. The input format appeals to the field geologist because it is relatively uncomplicated, easily understood and used, yet has complete flexibility with respect to the type and subjectivity of the information recorded.

Figure 1a

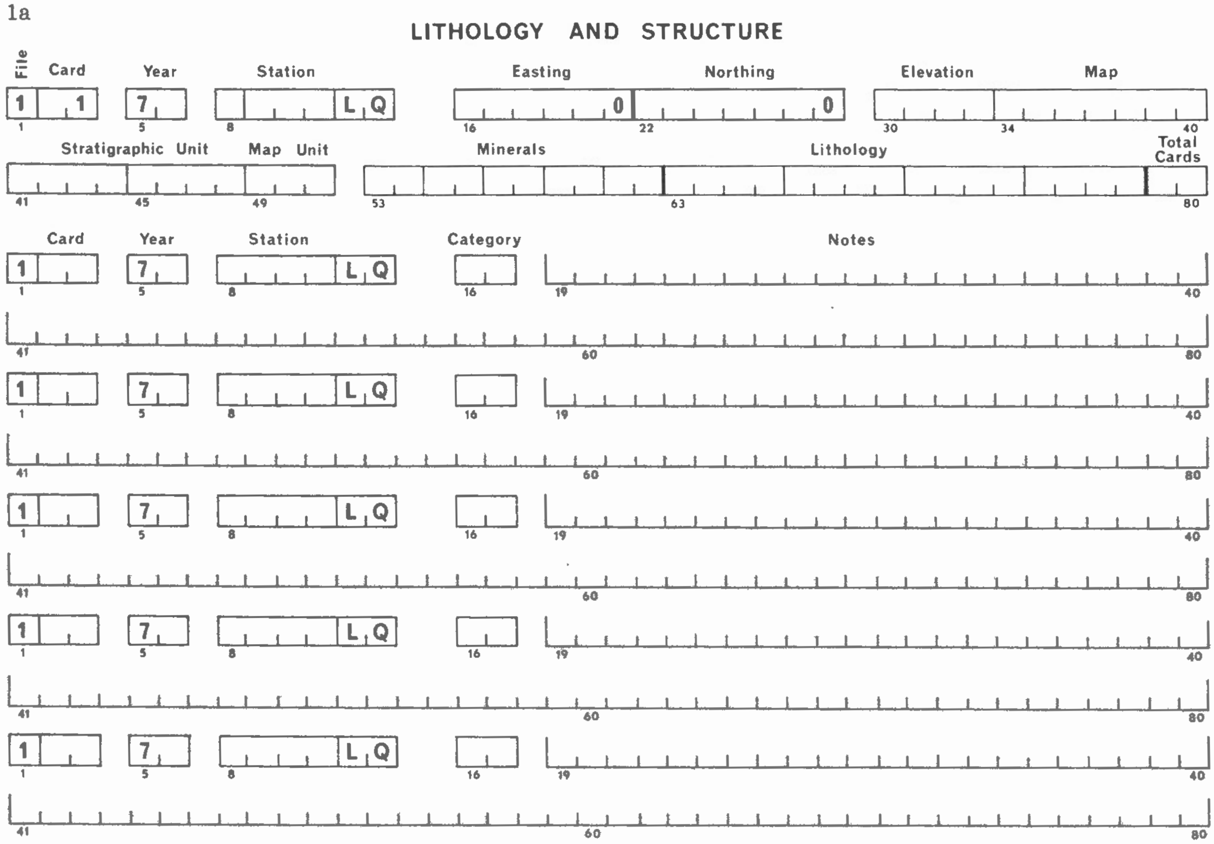


Figure 1b

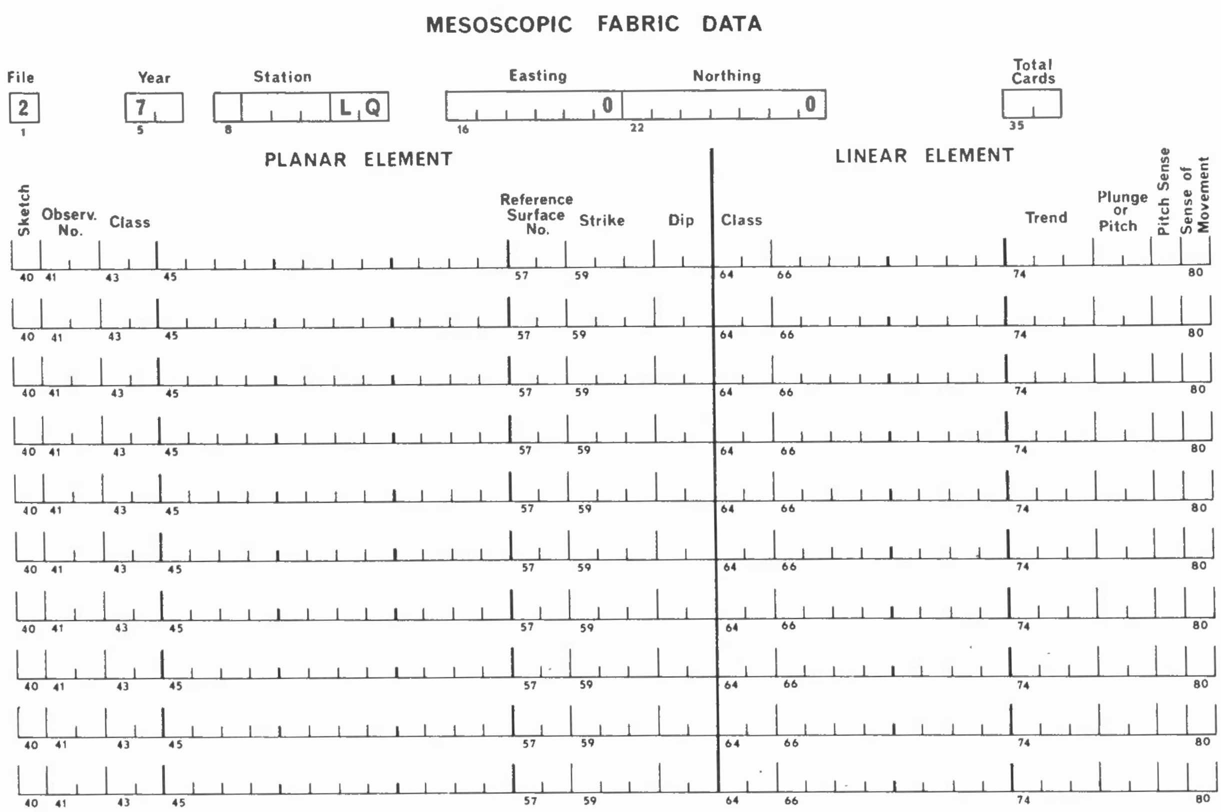


Figure 1c

Traverse: ______________________ **TRAVERSE SUMMARY**

Traverse Location: ______________________ Date: ______________________

Camp Location: ______________________ Temperature: ______________________

Photo Nos.: ______________________ Weather: ______________________

Manuscript Map No.: ______________________ B in C: A.M. __________ P.M. __________

File	Sum No.	Card	Year	Station		Station	Map	Total Cards
4		1	7		TO			
1		5	8	11		21	28	36

File	Sum No.	Card	Category	Notes
4				
1		5	8	11 ... 40
				41 ... 80
4				
1		5	8	11 ... 40
				41 ... 80
4				
1		5	8	11 ... 40
				41 ... 80
4				
1		5	8	11 ... 40
				41 ... 80
4				
1		5	8	11 ... 40
				41 ... 80

Figure 1d. Instructions for Notes

Category

1. Metasediments
2. Gneiss
3. Massive granitoid rocks
4. Structure
5. Dykes
6.
7.
8.
9.
10. Rocks Specimens - 3/card: 1 - spaces 19-40, 2 - spaces 41-60, 3 - spaces 61-80. [An oriented specimen can be indicated by ØRTD].
11. B & W Photos: one description per card.
12. 35 mm colour: one description per card, though more than one frame can be taken.
13.
14.
15. Glacial features

NOTE - Continuity between stations very important - should give concise account of what is seen, proportions of various rock types, relation to previous station, etc. A new station is necessary at any abrupt changes in lithology or structure. Account can be brief.

50. Interstation continuity - unconsolidated cover no outcrop, gravel, sand, soil, swamp.

51, 52, 54, etc., indicate interstation continuity in metasediments, gneisses, structure, etc.

NOTES: A. Measure structure so dip always to right. Overturned beds Ø1 instead of S1

B. Can use, 1 - / * () + but not ? Letters Ø, I, Ƶ should always be marked as shown.

C. Use metric system of measurement, in so far as possible.

Figure 1. Field input data files. (a) File 1, Lithology and Structure, (b) File 2, Mesoscopic Fabric Data, (c) File 4, Traverse Summary, and (d) Instructions for notes.

Figure 2

PETROGRAPHIC DATA

Column headings (cols. 37–80, mineral codes): AC Actinolite, AD Andalusite, AN Antigorite, AP Apatite, BI Biotite, CC Calcite, CH Chlorite, CO Cordierite, DI Diopside, DO Dolomite, EN Enstatite, EP Epidote, FL Fluorite, FO Forsterite, GA Garnet, GR Graphite, HT Hematite, HB Hornblende, HY Hypersthene, KF K-feldspar, KY Kyanite, MT Magnetite, MU Muscovite, PH Phlogopite, PC Plagioclase, PY Pyrite, PX Pyroxene(clino), QZ Quartz, SC Scapolite, SE Sericite, SP Serpentine, SI Sillimanite, SP Sphene, ST Staurolite, TO Tourmaline, TR Tremolite, ZR Zircon, ZO Zoisite

File (1)	Card (2–3)	Year (4–5)	Station (7–11)	Spec. No. (12–13)	Columns 14–80 (X, Y, NTS, Total Cards, minerals / descriptions)
3	1	71	19RA	-2	L07600 7440900 46N-1 · XZ 1X 2 X 3 4 2X X
3	2	71	19RA	-2	Hand Spec. Description: HS-MED GRN BI GA SI FS QZ SCST W FS-QZ LYRS
3			RA		HS-
3			RA		HS-
3	3	71	19RA	-2	Thin Section Description: QZ E35-SL STRAINED, ANHED, INCL IN GA W. SI
3	4	↓	↓ RA	↓	BI 15-RED BRN EXCEPT CONTACT GA WHERE GREEN
3	5		RA		GA 10-ELONGATE GRNS W.QZ INCL
3	6		RA		SI 10-CSE NEEDLE AND FINER WATS ASSOC GA AND CO
3	7		RA		CO <1-INTIMATE WITH SI AND QZ IN RIM OF GA
3	8		RA		MT <1-OPAQUES
3	9		RA		PC 25-TWINNED FRESH AN 38
3	10		RA		CH 5-INTERGROWTH WITH BI AND QZ NEXT GA
3	11		RA		AS <1- AP ZR ABUN IN BI SP
3	12		RA		TX -STRONGLY FOL, DISTORTED -AV GRN SIZE OF FS-QZ IS 1.5
3	13	71	19RA	-2	-MM BI FINER SI NEEDLES UP TO 0.6 MM -V. FRESH
3			RA		-
3			RA		-
3			RA		-

Note: HS - hand specimen; AS - accessories; TX - texture; SR - structure.

METAMORPHIC MINERAL ASSEMBLEGES - MELVILLE 1973

LEGEND

AC-ACTINOLITE
AD-ANDALUSITE
AP-APATITE
AT-ANTHOPHYLLITE
BI-BIOTITE
CC-CALCITE
CH-CHLORITE
CM-CUMMINGTONITE
CO-CORDIERITE
DI-DIOPSIDE
DO-DOLOMITE
ED-EDENITE
EN-ENSTATITE
EP-EPIDOTE
FL-FLUORITE
FO-FORSTERITE
GA-GARNET
GE-GEDRITE
GR-GRAPHITE
GU-GRUNERITE
HB-HORNBLENDE
HT-HEMATITE
HU-CLINOHUMITE
HY-HYPERSTHENE
KF-K-FELDSPAR
KY-KYANITE
MT-MAGNETITE
MU-MUSCOVITE
PC-PLAGIOCLASE
PH-PHLOGOPITE
PR-PREHNITE
PU-PUMPELLYITE
PX-PYROXENE(CL)
QZ-QUARTZ
SC-SCAPOLITE
SE-SERICITE
SI-SILLIMANITE
SL-SULFIDE
SP-SPHENE
SR-SERPENTINE
ST-STAUROLITE
TA-TALC
TO-TOURMALINE
TR-TREMOLITE
ZO-ZOISITE
ZR-ZIRCON

X - PRESENT LESS THAN 1 %
1 - BETWEEN 1 AND 10 %
2 - BETWEEN 11 AND 20 %
3 - BETWEEN 21 AND 30 %
4 - BETWEEN 31 AND 40 %
5 - BETWEEN 41 AND 50 %
6 - BETWEEN 51 AND 60 %
7 - BETWEEN 61 AND 70 %
8 - BETWEEN 71 AND 80 %
9 - BETWEEN 81 AND 100 %

STATION	AC	AD	AN	AP	BI	CC	CH	CO	DI	DO	EN	EP	FL	FO	GA	GR	HT	HB	HY	KF	KY	MT	MU	PH	PC	PY	PX	QZ	SC	SE	SR	SI	SP	ST	TO	TR	ZR	ZO
1RA 1				X	2	1						1				2		1		3					3			4		1			X				X	
1RA 3				X	1	X	1									1		1		3					2			4		2			X				X	
1RA 8				X	1	X	1											1		1		1			3			6		1			X					
1RA13				X		5	2					4								3		X			X		2	X		2			X					
2RA 3				X		X												1		8		X	1		1		1	2	1				X					
2RA 4				X		2	1													2		1			3		2	3		2			X					
2RA 5						1														3		X					5	1	2	1			X					
2RA5A									X											2		X	X					3	1				X			5		
6RA 3				X	1													2		1		1			5			4		1			X				X	
8RA 1				X	1							1					X			4		X			3			3		1			X					
11RA 1				X	1	X	X											3		2		X			2		3	X	2	3							X	1
19RA 2				X	2		1	X							2					1		1			3			4				2	X				X	
22RA 1							6					1								1		1	1					4		1			1					
22RA 2				X	2							1								3		1	X		2			5				1	X				X	
22RA 3				X	2	X	1													2		1	X		3			4	X	1		1	X				X	
22RA 4				X	2													1		1		X	1		3			4		2	2		X				X	
23RA 1				X	2													1		3		1			4			3					1				X	
26RA 1		1		X	2			1												1		X	X		1			6		2		2	X				X	
26RA 2		1		X	2															1		1			3			5		1		1					X	
28RA 1				X	3							1								1		X	1		1			6		1		1					X	
31RA 2				X	2							X						3		4		X			2		1		1	1			X				X	
31RA 3		X			2								X							2		1	1		3			5		1		1	X				X	
31RA 4				X	2			1							X		1			3			1		2			4		1		1	X					X
32RA 2						1	5													1		X	1				5	1	X				X					
35RA 1				X	2										X	1		1		2					3			5					X					
35RA 2				X	2	X	X									1		1		2					3		X	4		X			X				X	
35RA 3				X		1										1	X			2		X			X			8	1				X					2

Figure 3. Legend and mineral assemblage table retrieved from petrographic data (Fig. 2).

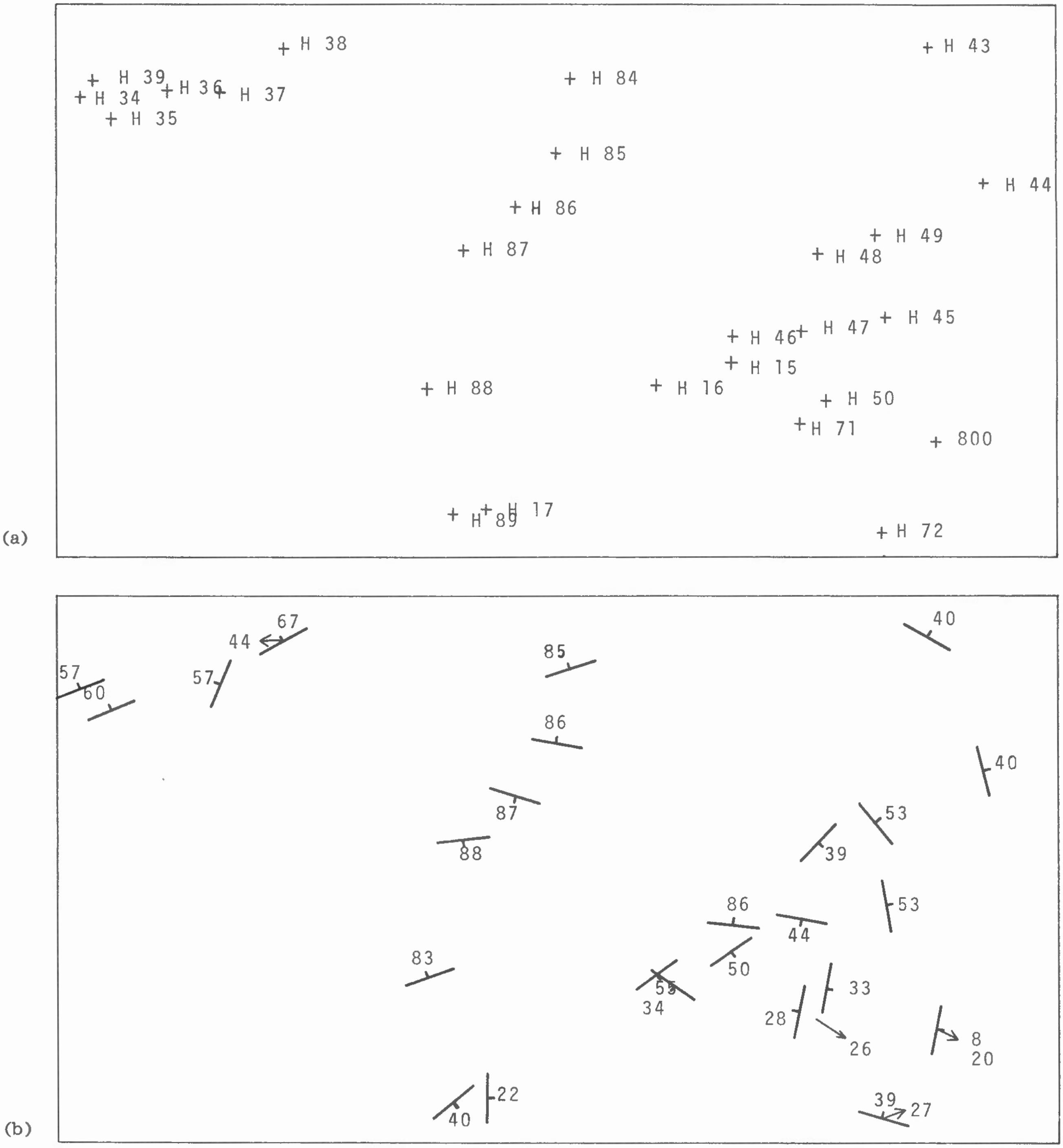


Figure 4. Station posting (a) and plot of planar and linear elements (b).

9. AN INTERACTIVE COMPUTER GRAPHICS APPROACH TO THE VOLCANIC GEOCHEMISTRY AND STRATIGRAPHY OF THE KIRKLAND LAKE AREA

D. D. Picklyk and R. H. Ridler,
Regional and Economic Geology Division

In order to test the feasibility of the use of interactive computer graphics as a research tool by geologists, particularly at the Geological Survey of Canada, a pilot project was initiated by the first author, under a N. R. C. Postdoctoral fellowship, in the spring of 1973 to design a system based on a relatively small geographic area (700 sq. miles) and data set (325 samples analyzed for 13 oxides). The system was designed to expedite the research of R. H. Ridler concerning the Archean volcanic stratigraphy of the Kirkland Lake area, Ontario but could readily be adapted to the analysis of a wide variety of analogous geological problems and much larger data sets.

The function of the system is to provide rapid access to, and selection of, rock geochemical data for display on a cathode ray tube. There are three basic types of diagrams on which the data may be displayed. These are: (1) a ternary diagram of any three of the major oxides, (2) a variation diagram of any two of the major oxides, or (3) a simplified geological map with the sample locations of the selected data indicated. Data may be selected by employing a variety of geologically logical parameters. The strictly geological criteria that may be used are the volcanic composition, stratigraphic age or gross rock fabric (principally primary). Data may also be selected by applying limiting criteria

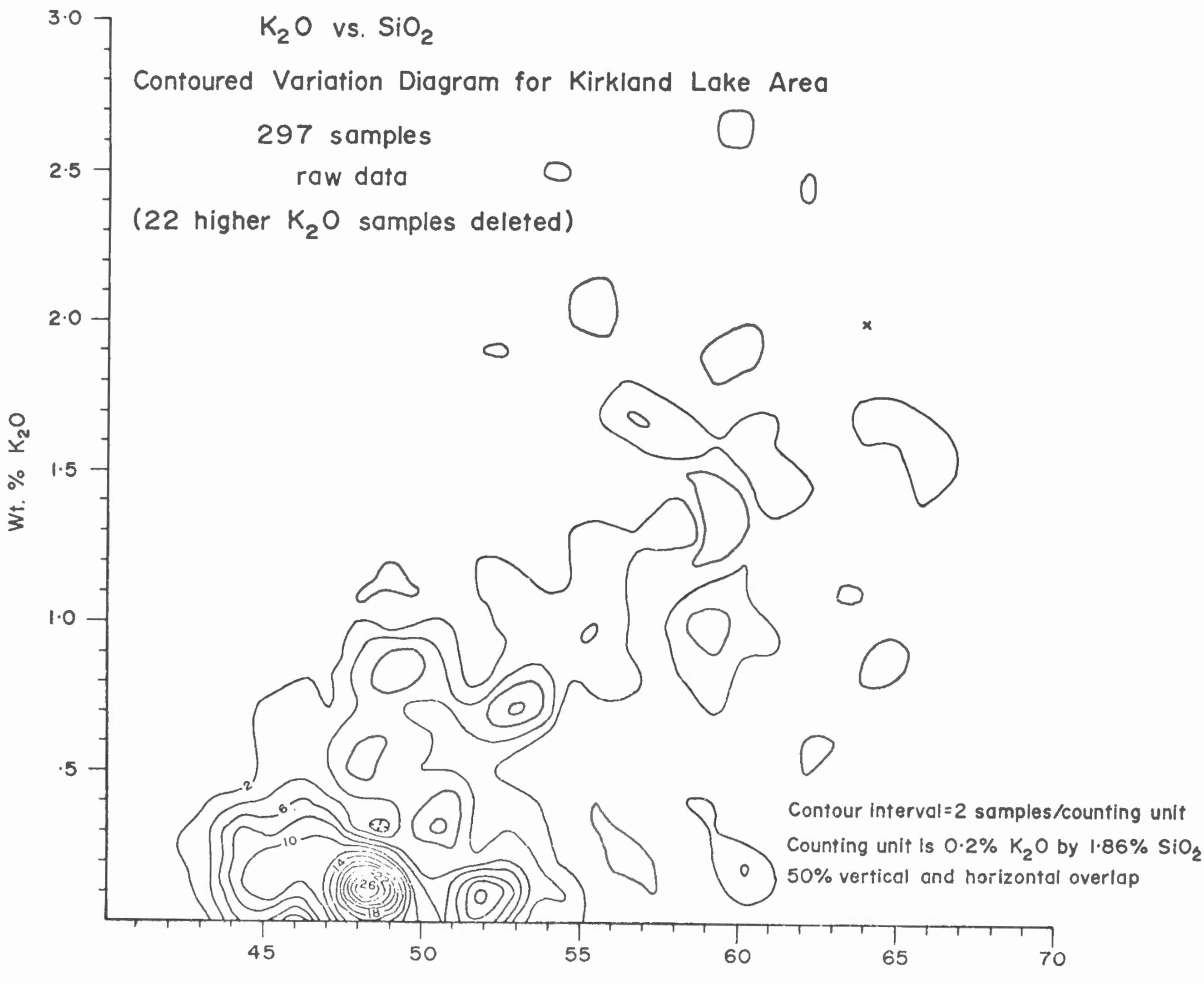


Figure 1

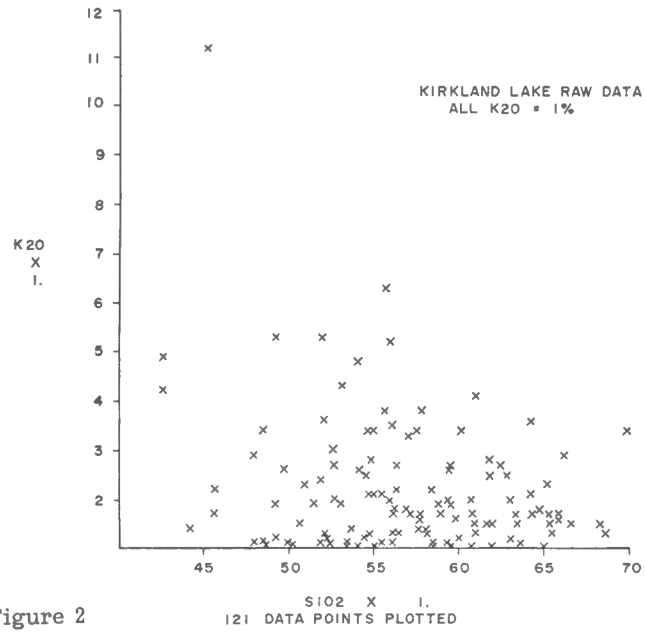


Figure 2

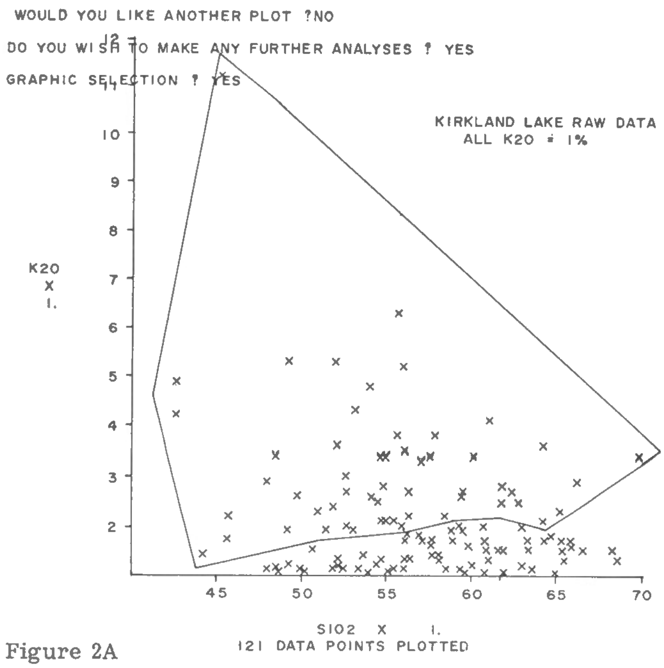


Figure 2A

to the geochemical data. The options include minimum, maximum and bracketing values specified for any three of the major oxides. A very "fine" selection of data is possible when three oxides are subject to limiting criteria. Less strictly geological criteria available are inclusive sample number sets and geographic area defined by latitude and longitude. The procedure of perhaps the greatest importance involves the selection of clusters (Ridler, 1973) of points on any of the diagrams by enclosing them in a closed polygon using the joystick control. Various criteria may be applied one after the other, listing the selections, if desired, at each stage, until the desired selection of data is obtained.

One of the main aims in developing the system was that it be as user (i.e. geologist) oriented as possible. This requirement was met by keeping the operator responsibilities very simple. Not even a moderate degree of data processing experience is required of the operator. For the most part the system is directed by "yes" or "no" responses to questions displayed on the screen. In cases where names or numbers are required they are always entered singly. In the case of numeric data the input format to be used is also supplied. One of the most useful features is that the operator can obtain assistance anytime input is required by entering HELP on the keyboard. The system will respond with not only a short message outlining what is required at the moment but also a listing of all possible valid responses.

The system is designed around a Tektronix Model 4002A graphics terminal with a 4603 Hardcopy unit and joystick as accessories. The terminal consists of an 11 by 11-inch storage cathode ray tube and keyboard through which all communication with the main computer is carried out. The Hardcopy unit provides the facility to make a permanent record of any output appearing on the screen. The Joystick control is used to provide a means of graphic input to the computer by providing positional control of a crosshair which appears on the screen. The terminal operates with a Digital Equipment Corp., PDP10/50 owned by Dataline Systems Ltd., of Toronto. All communication with the central computer is via telephone at 300 baud in full duplex. Although the following example is of considerable geologic significance (Ridler, 1970), it is presented here mainly to illustrate the facilities available on the interactive graphic system.

The Kirkland Lake area northeastern Ontario is a unique Archean potash-rich province with, in particular, chemically distinct volcanics. A contoured variation diagram for the entire K_2O vs SiO_2 population displays a distinct distribution of clusters (Fig. 1)*. By way of example we select a plot of all those samples containing more than 1% K_2O against the standard SiO_2 range (Fig. 2). By reference to the standard contoured variation diagram (Fig. 1) we select a group of clusters having relatively high K_2O contents by constructing a closed polygon with the Joystick control (Fig. 2a). This selection may be listed before continuing with the analysis and a partial listing of the output is shown in Figure 3 (the first 35 data records). Figure 4 is a ternary plot of the samples within the polygon and is presented here to illustrate another facet of the system. Figure 5 is a sketch map showing the various Archean formations, and the locations of the previously selected samples. By plotting the selected points before plotting the map, a convenient opportunity is provided for a separate copy showing only the points, if this is needed for clarification. A majority of the samples fall within the

---

*Editors' Note: Although the Hardcopy unit produces adequate prints for working use, these figures have been re-drawn for purposes of publication.

FIGURE 3

PAGE 1

NUM	FAB	TYP	AGE	SI	AL	K	MN	TI	CA	MG	FE2	FE3	NA	P	C02	H20
12.	Ø1	1	5	45.6	14.4	1.7Ø	.13	1.12	7.8	6.4	8.Ø	1.1	2.4	****	1.Ø	-1.Ø
13.	Ø4	2	5	56.3	15.3	2.7Ø	.14	.82	3.1	4.6	8.1	2.9	1.1	.16	Ø.3	4.3
16.	Ø1	2	3	52.5	17.6	3.ØØ	.Ø8	.53	5.1	3.Ø	3.3	Ø.1	3.3	****	-1.Ø	-1.Ø
21.	Ø3	2	3	55.4	15.4	2.1Ø	.14	.69	3.6	4.9	3.1	3.7	5.2	****	-1.Ø	-1.Ø
22.	Ø3	2	3	55.9	18.1	2.ØØ	.14	.85	3.5	2.5	4.5	1.5	4.1	.23	2.5	3.4
24.	Ø4	2	5	52.6	18.3	2.7Ø	.34	.63	6.3	2.7	8.Ø	1.9	3.5	****	-1.Ø	-1.Ø
32.	Ø4	6	5	57.Ø	14.1	3.3Ø	.12	.66	5.5	4.8	5.9	1.8	4.Ø	****	-1.Ø	-1.Ø
41.	Ø2	1	4	44.2	17.3	1.4Ø	.35	1.18	9.2	6.4	9.4	3.8	2.8	****	-1.Ø	-1.Ø
57.	Ø2	1	4	47.9	13.4	2.9Ø	.22	.92	7.7	6.3	6.5	7.2	2.3	.Ø5	Ø.1	2.9
62.	Ø2	1	4	49.2	13.2	1.9Ø	.13	.72	7.8	4.9	5.4	1.3	2.9	****	-1.Ø	-1.Ø
67.	Ø3	2	3	54.8	13.7	2.8Ø	.37	.56	6.3	4.4	5.3	Ø.5	3.2	****	-1.Ø	-1.Ø
73.	Ø4	3	4	59.5	15.2	2.7Ø	.Ø5	.71	4.4	3.7	6.5	1.Ø	2.3	****	-1.Ø	-1.Ø
75.	Ø3	6	3	55.9	14.5	5.2Ø	1.24	.45	5.Ø	3.5	3.9	1.4	Ø.7	****	-1.Ø	-1.Ø
68.	Ø1	1	5	54.9	14.5	2.1Ø	.18	.38	1Ø.8	8.8	7.1	2.3	1.6	****	-1.Ø	-1.Ø
87.	Ø2	6	5	54.Ø	16.4	4.8Ø	.28	.52	7.7	3.2	5.5	2.6	1.1	.Ø5	-1.Ø	2.2
1Ø2.	Ø3	3	3	61.8	15.5	2.8Ø	.Ø9	.48	4.6	2.6	3.7	Ø.1	3.2	****	-1.Ø	-1.Ø
1Ø5.	Ø4	4	5	64.2	15.8	3.6Ø	.Ø5	.33	1.8	1.4	3.Ø	Ø.2	4.Ø	****	-1.Ø	-1.Ø
116.	Ø3	3	3	61.Ø	15.3	4.1Ø	.37	.47	7.9	3.8	3.Ø	2.5	Ø.1	.18	Ø.2	2.3
117.	Ø4	2	5	50.9	15.4	2.3Ø	.15	.73	7.6	7.9	8.4	1.2	2.8	.16	Ø.1	2.Ø
128.	Ø2	1	5	45.6	14.8	2.2Ø	.18	.84	7.8	7.6	9.Ø	2.4	2.6	.Ø5	2.8	2.4
149.	Ø3	6	5	57.5	15.1	3.4Ø	.11	.59	4.8	3.6	1.8	4.2	3.1	.35	3.6	2.5
179.	Ø1	1	5	49.7	11.3	2.6Ø	.18	.69	8.5	9.8	5.7	2.Ø	2.7	****	-1.Ø	3.4
182.	Ø1	2	5	56.3	14.7	2.2Ø	.11	.65	5.5	5.4	5.2	1.7	3.4	.26	3.4	3.4
195.	Ø1	3	3	61.8	15.9	2.5Ø	.Ø7	.43	4.4	3.Ø	3.6	Ø.8	1.7	****	3.3	3.3
197.	Ø5	3	3	58.4	15.1	2.2Ø	.1Ø	.52	4.4	5.Ø	3.8	1.3	5.4	.22	Ø.1	1.6
203.	Ø5	4	3	64.2	16.3	2.1Ø	.Ø9	.44	5.4	3.2	1.8	2.7	2.6	.17	Ø.4	2.3
209.	Ø4	4	4	66.2	19.3	2.9Ø	.Ø3	.42	1.5	1.9	Ø.9	Ø.3	5.4	.14	Ø.2	1.4
231.	Ø3	2	3	54.5	18.5	2.5Ø	.13	.59	6.5	4.9	4.4	2.Ø	2.4	.23	Ø.8	3.9
25Ø.	Ø3	2	5	52.6	13.Ø	2.ØØ	.15	.61	7.5	6.4	5.5	2.Ø	4.2	.49	3.8	2.7
251.	Ø4	2	5	51.8	18.5	2.4Ø	.19	1.Ø9	3.4	3.9	5.9	2.Ø	5.6	.41	-1.Ø	2.8
252.	Ø4	6	5	55.6	15.4	3.8Ø	.18	.5Ø	5.Ø	4.8	5.6	Ø.6	4.3	.28	Ø.3	1.3
254.	Ø4	2	5	53.8	14.3	1.9Ø	.11	.87	4.9	7.5	5.Ø	2.3	4.1	.37	1.3	3.7
255.	Ø4	5	5	42.5	15.5	4.2Ø	.23	.95	1Ø.4	4.Ø	5.4	3.1	3.3	.7Ø	5.1	2.5
266.	Ø3	2	5	54.1	15.9	2.6Ø	.13	.53	3.8	2.5	2.8	2.9	5.9	.23	5.1	2.5
273.	Ø4	3	5	6Ø.1	17.Ø	3.4Ø	.11	.69	1.8	3.6	5.8	Ø.6	4.7	.19	Ø.5	2.6

TO CONTINUE DEPRESS SPACE BAR TWICE

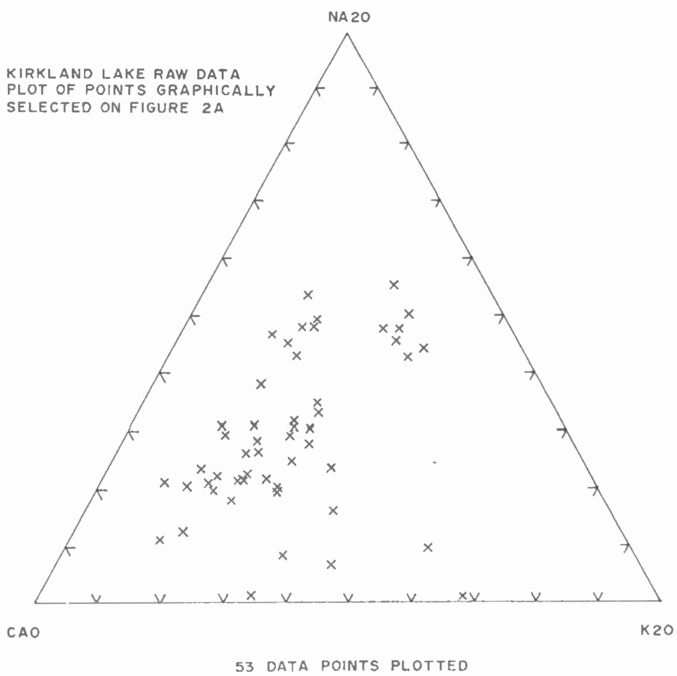


Figure 4

Timiskaming Group while most of the rest occur in the Skead pyroclastics. Without going into details, the greatest geological ramification of this finding is that it is consistent with the structural/stratigraphic interpretation of the map and inconsistent with the traditional interpretation which disallowed correlation of formations across an alleged major east-west fault supposedly bisecting the area. It is worth emphasizing that the above illustration is extremely simplistic. The system's capability to assist in more complex cluster analysis as well as many other ways of analyzing the data has not yet been fully utilized. Hopefully, the system offers the opportunity to plug the thousands of chemical analyses buried in tables or on uncontoured variation diagrams back into their geologic setting and get something really useful out of them ... in particular, chemical "finger prints" or "index fossils".

Some of the features of the command language used to guide the system can be noted by comparing Figures 2 and 2a. Figure 2 is a plot of the selected data. In Figure 2a it will be noted that the upper left part of the figure has been overprinted. These are the system messages and responses necessary to make the graphic selection by enclosing the points within the polygon.

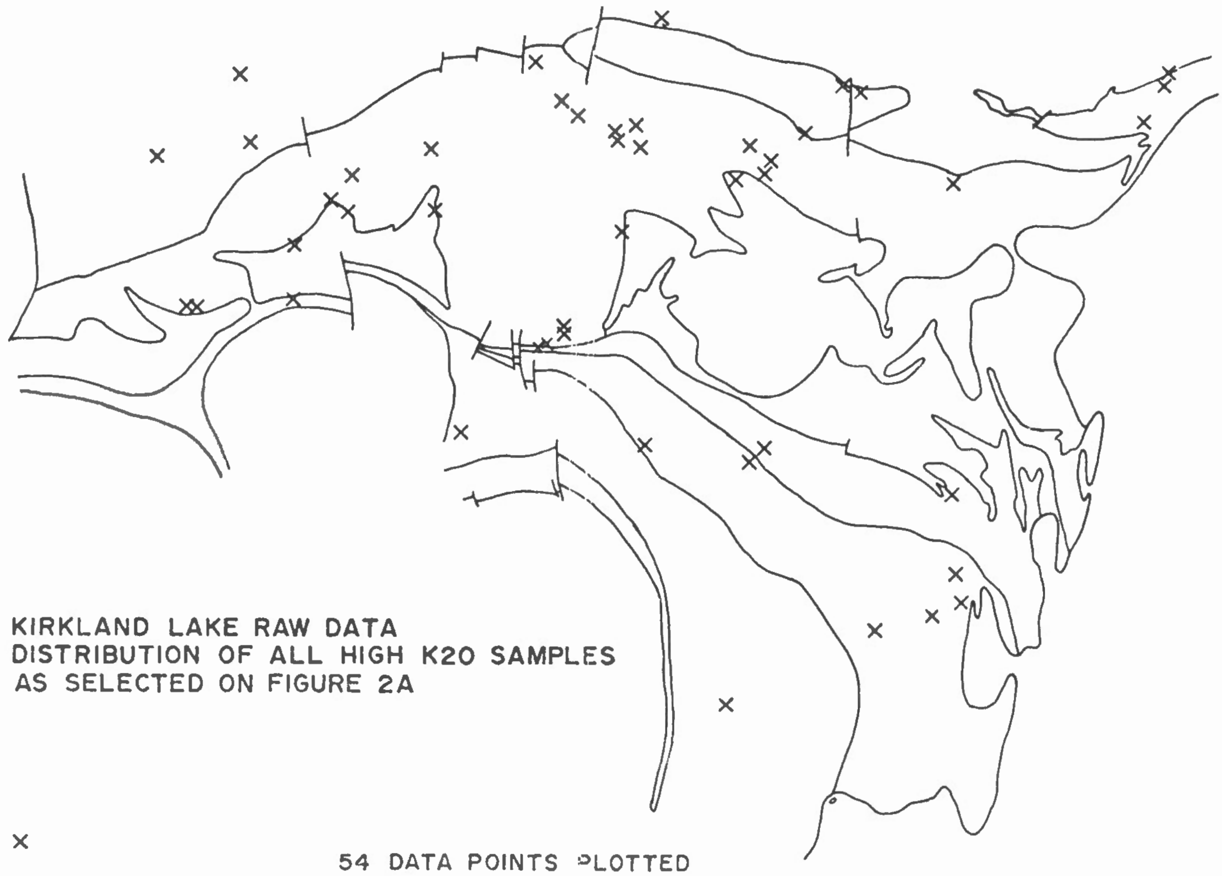


FIGURE 5

It is possible to annotate any of the figures by placing characters anywhere on the screen.

At the present all programming and most of the testing, including a trial run by Ridler, have been completed. A User Manual and System Reference Guide will be prepared as soon as all testing is satisfactory. A complete description with examples will be published in the Bulletin of Canadian Institute of Mining and Metallurgy (Picklyk and Ridler, in press). Queries regarding the system may also be directed to either R.H. Ridler or D.D. Picklyk at the Geological Survey.

References

Picklyk, D.D., and Ridler, R.H.
1974: Computer Graphics - An Interactive Approach to Volcanic Geochemistry and Stratigraphy: Illustrated by data from the Kirkland Lake Area, Ontario; Can. Mining Met. Bull., v. 67, no. 743, p. 71 (Abs.).

Ridler, R.H.
1970: Relationship of Mineralization to Volcanic Stratigraphy in the Kirkland-Larder Lakes area, Ontario; Geol. Assoc. Can. Proc., v. 21, p. 33-42.

1973: Automated contoured variation diagrams, *in* Volcanism and Volcanic Rocks; Geol. Surv. Can., Open File 164, p. 65-71.

N. Prasad, R.V. Kirkham and F.M. Vokes,
Regional and Economic Geology Division

MOLYFILE, a simple computer-based file on molybdenum deposits in Canada, was designed to store and process a limited amount of mineral deposit data. Increase in amount of data being handled and constantly changing demands for information make it desirable to use a computer-based system for storage, retrieval and accurate, efficient manipulation of data. Some of its main functions are as follows: comprehensive but shallow coverage of molybdenum deposits and occurrences on a national scale; organization of data for rapid scanning, sorting and indexing in a variety of ways; immediate availability of data for plotting maps of variable scale and generating tables of variable form and content; efficient update capacity directly into a computer-processable form; rapid tabulation and calculation of production and reserve data necessary for resource evaluation studies; rapid interactive data manipulation, extraction and editing; table and note production to aid in report writing.

The flow chart in Figure 1 illustrates various stages of development, operation and use of MOLYFILE.

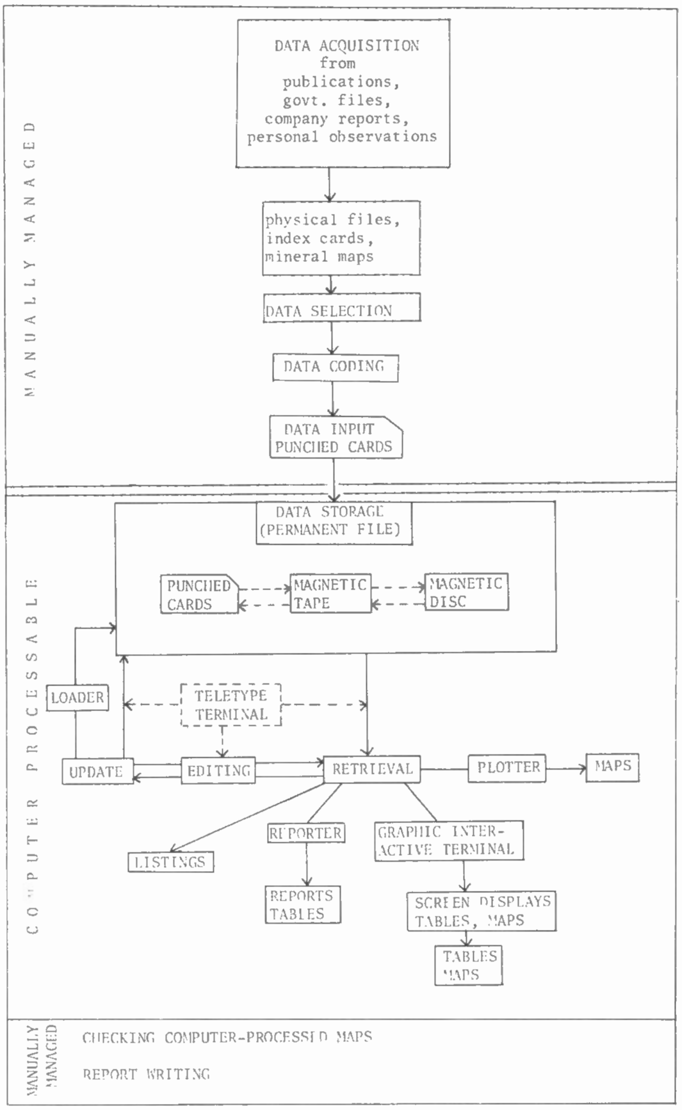
The data items are coded and punched on 80-column cards (Fig. 2) according to a fixed format. The coding scheme and the format used were devised specifically for this file. In order to keep the file simple, data on one deposit was normally coded on one 80-column card. However provision to supplement reserve, production and other information permits the use of up to four cards. A classification of molybdenum deposits designed by F.M. Vokes was integrated into the file.

MOLYFILE utilizes the MARS VI system (Multi Access Retrieval System), which is controlled by the SCOPE 3.4 operating system on a CDC 6400 computer. FORTRAN is used for reading and repacking punched data cards for input into MARS on magnetic tape or disc. Definition of codes and field are done in MARS, thus creating a permanent record file and a permanent data-base file. The file consists of 1,537 records in the "record file" which together with the "data-base file" occupy 185 blocks of the computer memory. Five and one half man-months by a trained individual were necessary to complete the data coding and editing and one man-month of computer programming* was necessary in construction of the file. About $500 of computer time were spent in erection and testing of the file. Aspects of the file are still being developed and tested.

The file is accessible by an interactive teletype-terminal which can be used for retrievals and updates. Future use could also be made of graphic terminals (*see* Picklyk and Ridler elsewhere in this volume). Some examples of actual update and retrieval runs are given in Figure 3. Figure 4 shows a computer plotted map (redrafted) of the distribution of moly occurrences in NTS block 21.

*The computer programs were kindly written and tested by W.N. Houston.

FIGURE 1. FLOW CHART



MOLYFILE, in its present form, has been designed for personal project activities. Because of manpower and other limitations and constantly changing demands for information the file was purposely not made to be exhaustive or necessarily to have a long-life span. Nevertheless, because of its simple fixed format many useful data could be easily transferred to other files. Yet for current uses most of the important data have been included and it is hoped that it will greatly expedite report and map generation and other project activities. Limitation on use of the file are mainly due to its restricted data content, personalized nature and, of course, to the normal limitation of computer-based mineral deposits files, such as recurrent periods of computer breakdown and manpower shortages.

Figure 2

MO DEPOSIT DATA INPUT FORM

Accession No. (1–6) | Name (7–23) | Total Cards (24) | Lat. (25–30) | Long. (31–37) | Accuracy (38)

Accuracy:
1. Reliable Publication
2. From 1:50,000
3. From 1:250,000 map
4. From 1 million map
5. From MRBNMI
6. From Geoscience Index

N.T.S. No. (39–45)

Prov./Terr. (46)

C = B.C.
Y - Ykn.
A = Alta.
S = Sask.
M = Man.
W = NWT
O = Ont.
K = Que.
B = N.B.
N = N.S.
F - Nfld.
L = Labr.

@Status (47)

P = Producer
F = Former producer
R = Deposit with reserves
I = Investigated Deposit
O = Mineral Occurrence

Deposit Type (48)

According to F.M. Vokes classif.

Geol. Prov. (51)

C = Cordillera
B = Bear
L = Slave
H = Churchill
S = Superior
G = Grenville
N = Nain
O = Southern
A = Appalachian

Stratig. age (52)

Q = Quaternary
T = Tertiary
K = Cretaceous
J = Jurassic
R = Triassic
P = Permian
V = Pennsylvanian
M = Mississippian
D = Devonian
S = Silurian
O = Ordovician
C = Cambrian
H = Hadrynian
N = Neohelikian
L = Paleohelikian
A = Aphebian
X = Archean

Elements — Major (54) — Minor (59)

M = Mo
C = Cu
W = W
B = Bi
U = U
H = Th
T = Sn
5 = Te
Z = Zn
P = Pb
N = Ni
L = Li
J = Co
R = RE
X = Sr
6 = Pt
Y = Be
G = Au
S = Ag
Q = Cb
K = As
A = Ta
I = Fe
7 = Py
F = Fluor
E = Apat
D = Fels
O = Sb
1 = Graph
2 = Mica
3 - Zircon
4 = Corund
8 = S

@Res. &/or Prodn. in '000 tons (60–65)

Grade Mo (66) — (.) before column 66

Metal X (69)

* Grade (70) — (.) between columns 70 & 71

Metal y (73)

* Grade (74–76) — (.) between columns 74 and 75

@Date of res./end prodn./invest. (77)

Ref. (79)

1 = Mo Card index
2 = Mo Card + file
3 = Mo Card index + MRBNMI sheet
4 = Mo Card index + file + MRBNMI sheet

No. of card (80)

*when Metal x/y is Au or Ag grades are in ppm.(.) between columns 71 and 72 or 75 and 76.

@ Reserve & production values and date are in conjunction with the status. If both reserve and production data are available more than one card is used.

```
MARS VI V2.1   05/01/72
LOAD DATA BASE FROM MYDB7:
SAVE/LOAD COMPLETED
RETRIEVAL;
IF DEPTYPE(1) EQ 482 AND GEOLPROV(1) EQ C AND RESPRD(1) GE 100000 REPORT
NAME(1),NTS(1),METALX(1),GRADEX(1),C1,BY HIGH C1:

NAME(1)                NTS(1)     METALX(1)    GRADEX(1)     C1


HIGHMONT               092I07W    CU                 028     092025
MAGGIE                 092I14W    CU                 040     092032
ISLAND COPPER(BAY      092L11W    CU                 052     092044
VALLEY COPPER          092I06E    CU                 050     092053
CATFACE                092F05W    CU                   0     092057
BETHLEHEM-JA ZONE      092I07W    CU                 043     092091
ENDAKO (STELLA)        093K03E                         0     093018
AJAX                   103P11W                         0     103022
LIARD COPPER.SND       104G07W    CU                  04     104001
ADERA                  104N11W                         0     104005
CASINO-PATTONHILL      115J10W    CU                 037     115009

MASS UPDATE BY C1;
C1=092036,DEGLAT(1)=51,MINLAT(1)=00,DEGLONG(1)=120,MINLONG(1)=26,GEOLPROV(1)=C;
READY
C1=092096,NAME(1)=KR   K,LAT(1)=502918,LONG(1)=12031,ACCUR(1)=1,NTS(1)=092I07E,
READY
PROV(1)=C,STATUS(1)=I,DEPTYPE(1)=0Y0,GEOLPROV(1)=C,MAJELS(1)=CM,REF(1)=1,
READY
NTS(2)=*,CARDTOT(1)=1;
```

Figure 3.1

```
                                     REPORT ON
                                MOLY DEPOSITS IN CANADA


                                 TYPE 4A1-VEINS RELATED
                                 TO HOLOCRYSTALLINE INTRUSION


ACC NO     001001     NAME     LAWN                    REF     MO CARD INDEX + FILE + MRBNMI SHEET
LAT     465706     LONG     0553148     PROV/TERR     NFLD.     NTS     001L13E
ACCURACY     FROM MRBNMI               STATUS     INVESTIGATED DEP     DATE     0
GEOLPROV     APPALACHIAN     AGES     DEV
MAJOR ELEM     MO                          MINOR ELEM     FLUOR
RES/PROD          0          0     GRADE OF MO          0


ACC NO     001006     NAME     SALMONIER POND CV     REF     MO     CARD INDEX + MRBNMI SHEET
LAT     4734     LONG     0553036     PROV/TERR     NFLD.     NTS     001M12E
ACCURACY     FROM MRBNMI               STATUS     OCCURRENCE               DATE     0
GEOLPROV     APPALACHIAN     AGES     DEV
MAJOR ELEM     MO                          MINOR ELEM
RES/PROD          0          0     GRADE OF MO          0

ACC NO     001015     NAME     HUNGRY GROVE POND     REF     MO     CARD INDEX + MRBNMI SHEET
LAT     475440     LONG     05507     PROV/TERR     NFLD.     NTS     001M14E
ACCURACY     FROM MRBNMI               STATUS     OCCURRENCE               DATE     0
GEOLPROV     APPALACHIAN     AGES     DEV
MAJOR ELEM     MO                          MINOR ELEM     BI
RES/PROD          0          0     GRADE OF MO          0


ACC NO     011001     NAME     L.CHARLOTTE             REF     MO     CARD INDEX + MRBNMI SHEET
LAT     4452     LONG     0625830     PROV/TERR     N.S.     NTS     011D15W
ACCURACY     FROM MRBNMI               STATUS     INVESTIGATED DEP     DATE     39
GEOLPROV     APPALACHIAN     AGES
MAJOR ELEM     W                           MINOR ELEM     MO          AU          AG
RES/PROD          0          0     GRADE OF MO          0


ACC NO     021001     NAME     WILCOX(KEDDY)PROS     REF     MO CARD INDEX + FILE + MRBNMI SHEET
LAT     4443     LONG     0642945     PROV/TERR     N.S.     NTS     021A09W
ACCURACY     FROM MRBNMI               STATUS     INVESTIGATED DEP     DATE     63
GEOLPROV     APPALACHIAN     AGES     DEV
MAJOR ELEM     MO                          MINOR ELEM
RES/PROD          0          0     GRADE OF MO          0

ACC NO     021002     NAME     L.RAMSAY                REF     MO CARD INDEX + FILE + MRBNMI SHEET
LAT     444430     LONG     0643040     PROV/TERR     N.S.     NTS     021A10E
ACCURACY     FROM MRBNMI               STATUS     INVESTIGATED DEP     DATE     63
GEOLPROV     APPALACHIAN     AGES     DEV
MAJOR ELEM     ZN                          MINOR ELEM     FLUOR     MO          CU
RES/PROD          0          0     GRADE OF MO          0
```

Figure 3.2 - PAGE 1 OF A GENERATED REPORT

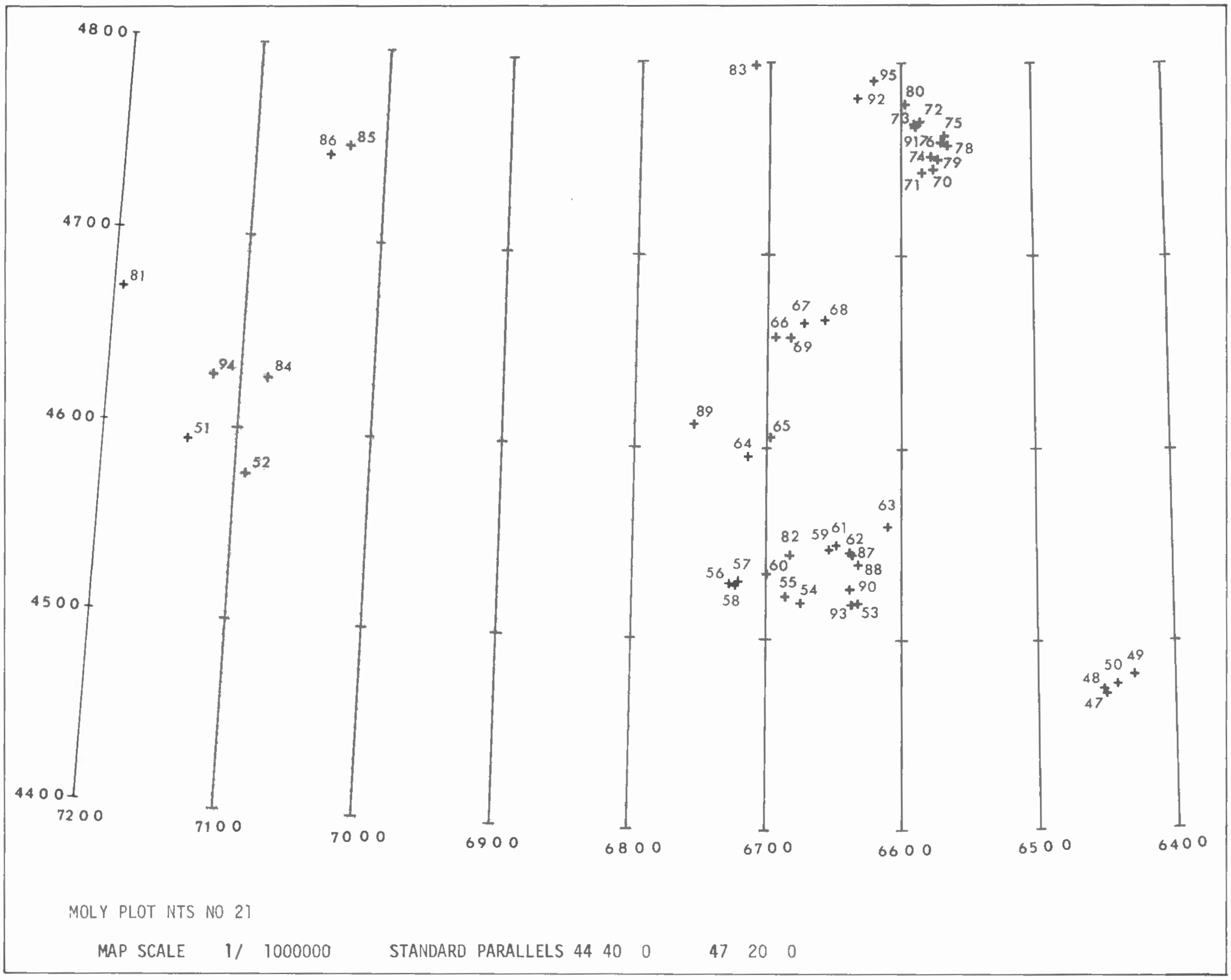

4800
4700
4600
4500
4400
7200
7100
7000
6900
6800
6700
6600
6500
6400
MOLY PLOT NTS NO 21
MAP SCALE 1/ 1000000 STANDARD PARALLELS 44 40 0 47 20 0

Figure 4

11. GNEISS DISTINCTIONS IN THE HAYES RIVER REGION, DISTRICT OF KEEWATIN

Mikkel Schau,
Regional and Economic Geology Division

S.W. Campbell,
University of British Columbia

Gneisses in the Precambrian Shield are difficult to subdivide. In the Hayes River region located south of Boothia Peninsula, it is known that gneisses of different ages occur in the region (Heywood, 1961; Schau, 1973, 1974). Since outcrops rarely show the significant features which allow classification of gneisses into those of probable Aphebian age and those of probable Archean age, it was thought that the interior fabrics of the gneisses might yield a clue as to which group they should be assigned.

In a small pilot project some gneisses are examined in the field according to a questionnaire (Fig. 1). The answers yield a vector composed of Yes, No type data (Fig. 2). This type of data is divided into groups of related data (Fig. 3) using a non-parametric clustering program (Bonham-Carter, 1965). This program produces a listing of the samples in the order that they are grouped on the resultant dendrogram so that the arbitrary assignment of samples to subgroups can easily be accomplished by simple commands in EDITOR on the SCOPE 3.4 INTERCOM system. Once the gneiss is assigned to a subgroup it is then put into a data base management system such as MARS VI along with other data collected in field and laboratory. The MARS retrieval capabilities are such that the files are linked and the relevant data retrieved. (Repeated retrievals can be accessed using macro calls.) Some of these data may be best displayed on a map; in this case it is put onto a file and the required maps are produced using appropriate programs (Crain and Morris, 1972).

Example

In the Laughland Lake (56K) and Walker Lake (56J) sheets of the Hayes River area, two, possibly three, periods of gneissification occurred (Schau, 1973). One was prior to the deposition of the "Archean" Prince Albert Group (Reesor, pers. comm.) and it may include gneisses whose precursors may include layered and chemically heterogeneous units as well as certain plutonic "tonalitic" complexes (Schau, 1974). A second gneissification event occurred after the emplacement of ultrabasic sills and dykes in the Archean (?) Prince Albert Group but prior to the emplacement of a white muscovite garnet granite (Schau, 1973). This affected mainly layered rocks of the Prince Albert Group and was most effective east of 92°W or so. A third event, perhaps of very local extent, is associated with the emplacement of the above-mentioned granite (Schau, 1973).

Where requisite geological relations are not available lithologic correlation must be resorted to in the study of Archean rocks. To illustrate the method previously described, seven gneisses, of which geological relations are known for three, are described below:

GNEISS No. 4 (5024)
: Faintly layered, fine-grained, gneiss whose 5 to 10-mm-thick layers are delineated by abundant biotite clots and grains set in a quartz-rich granitoid matrix.

GNEISS No. 8 (0136)
: A distinctly layered gneiss in which lenticles of schistose chloritized biotite and lineated hornblendes about 10-20 mm apart are set in an augen-bearing, quartz-rich granitoid matrix. The gneiss is interbedded with chemically heterogeneous rocks and clearly belongs to the Prince Albert Group; it is cut by quartz veins and tension gashes.

GNEISS No. 11 (5091)
: Faintly layered, medium-grained, gneiss whose 5 to 10-mm-thick layers are delineated by abundant (20 + %) biotite clots set in a quartz-eye rich granitoid matrix.

GNEISS No. 26 (0383)
: A distinctly layered gneiss with warped layers 1-10 mm apart of isoclinal fold noses, lenticles and boudins consisting of biotite and abundant lineated hornblende, set in quartz-rich equigranular matrix in which K-spar is rare. This too is part of the Prince Albert Group and is cut by nearby "white granite".

GNEISS No. 31 (8004)
: Distinctly layered porphyroblastic gneiss with abundant lenticular layers 5-10 mm apart of biotite set in inequigranular plagioclase, and quartz matrix. A minor amount of pyrite and local epidotization provide variety on the outcrop. To the east, rocks of the Prince Albert Group are folded about a core of similar looking gneiss. There, near the "supposed base" of the Prince Albert Group, are found units with many small lenticular pods which may be remains of a basal conglomerate. Therefore, these rocks are considered to be part of the pre-Prince Albert Group Gneisses.

GNEISS No. 37 (5116)
: Fine-grained, finely but faintly layered gneiss with schistose biotite clots set in granitoid matrix. Cut by coarse pegmatite. Grades in outcrop to more granitic members.

GNEISS No. 44 (5140)
: Light pink to grey weathering fine-grained, thin layered mafic gneiss consisting of equally abund-

Figure 1. Field Classification of Gneiss

If "yes" put 2 in column, if "no" put 1, if "not applicable" put 0.

Column Number	Nature of attribute
1	Distinctly layered.
2	Faintly layered.
3	Continuous (planar) layering.
4	Discontinuous or lenticular layering.
5	Mafics aggregated in layers or clots.
6	Evenly dispersed mafic minerals.
7	Layering due to grain size change?
8	Layering due to compositional change?
9	Spacing between layers regular?
10	Is spacing rhythmic?
11	Are layers due to very scattered mafic blebs?
12	Spacing of layers is 1-5 mm apart?
13	Spacing of layers is 5-10 mm apart?
14	Spacing of layers is 10-20 mm apart?
15	Felsic layers isoclinally folded?
16	Felsic layers warped?
17	Mafic layers isoclinally folded?
18	Mafic layers warped?
19	Are felsic layers boudinaged?
20	Are mafic layers boudinaged?
21	Are mafic portions discontinuous within felsic material?
22	Are felsic layers more than 50% rock?
23	Are mafic layers more than 20% rock?
24	Equigranular felsic layer?
25	Inequigranular felsic layer?
26	Equigranular mafic layer?
27	Inequigranular mafic layer?
28	Coarse or finer grained?
29	Medium or finer grained?
30	Granitoid fabric in felsic part?
31	Granitoid fabric in mafic part?
32	Granoblastic fabric in felsic part?
33	Granoblastic fabric in mafic part?
34	Schistose fabric in mafic part?
35	Lineated fabric in mafic part?
36	Augen in felsic layer?
37	Porphyroblasts in felsic layer?
38*	Porphyroblasts in mafic layer?
39	Has it been sheared?
40	Evidence for 2 or more generations of mineral?
41	Quartz rich?
42	Has quartz?
43	Has K-Spar?
44*	Has plagioclase?
45	Has garnets?
46	5 mm across or larger?
47*	Has sillimanite?
48	Has fibrolite?
49	Has epidote?
50*	Has staurolite?
51*	Has andalusite?
52*	Has kyanite?
53*	Has pyroxene?
54	Has green plag?
55	Has amphibole in abundance?
56	Is amphibole present?
57	Has biotite in abundance?
58	Is biotite present?
59*	Has muscovite in abundance?
60	Is muscovite present?
61	Is hornblende aligned?
62*	Is hornblende in rosettes?
63	Random hornblende?
64	Is chlorite present?
65	Does it clear post date other mafics?
66*	Is tremolite present?
67*	Is talc present?
68*	Is calcite present?
74-80	Specimen designation

*These questions are redundant. All are one with exception of 44 which is two.

GNEISS No. 44 (5140) (cont'd.)

ant granitoid felsic material and locally epidotized schistose biotite rich mafic layers.

Outcrops of these gneisses yield Yes, No vectors as seen in Figure 2. After the number of similar and dissimilar characters are counted, a matrix of similarity coefficients are generated as shown in Table 1.

This matrix shows that 4 and 11 are most like each other and that 4 and 8 are least like each other. If the Yes or No were randomly chosen, the similarity coefficient used here would be around .562. The procedure of clustering then considers the pair with the highest similarity coefficients (i.e. 4-11) as though they were one; and the similarity coefficients are re-calculated between this pooled pair and all other vectors. This procedure is continued until all vectors (i.e. samples) have been related to each other. This will yield a dendrogram whose horizontal axis is the similarity coefficient discussed above and whose vertical axis is composed of samples (Fig. 3).

Four groups have been arbitrarily distinguished using the dendrogram as a guide:

DIVISION 1:

5136, 5141, 5140, 5116, 5133, 5134, 5111, 5109, 8003, 8007, 8004, 8002, 8010, 5110, 5112, 5113, 0377, 5083, 5123, 8005.

GNEISSES OF PRINCE ALBERT BELT . U.MO

PRINTOUT OF DATA MATRIX

SAMPLES NAME	NO.	CHARACTERS
5004SM	1	211221121111121111112222121111111121211112222111111111111212111111111
5010SM	2	121221121121211111111212212111111121211111222111111111111211211111111
5019SM	3	211221121121211212112212121111111121111112222111111111111212121111111
5024SM	4	121221121121211111112212121111221121111112222111111111111221111111111
5028SM	5	211221121122111212112122212111211121111211222111111111111221111111111
5040SM	6	211221121112111112112222121112211121111112222111111111111221111111111
0135SM	7	211212221111212211111211211121111111111212222111111111111221111101111
0136SM	8	212121221121121111111221121212211121211112122111111111111212112112211
0138SM	9	211221222121122222222221122112211121121211222111211111111211111111111
5083SM	10	121221121122111111111221121212221121211112222111111111111211111111111
5091SM	11	121221121121211111111222212112221121111112222111111111111221111111111
5092SM	12	121222121122111111111212212111221121111111222111111111111211111111111
5093SM	13	121221121122111212112122212112221121111112222111111111111221111111111
5096SM	14	121212121111211111112122212112211121111112222111111111111221111111111
5097SM	15	121221121121211111112122212112211121111112222111111111111211111111111
5098SM	16	12122112112211121211212211212211121111112222111111111111221111111111
5101SM	17	121222121122111111112122212111221121111112222111111111111221111111111
5102SM	18	121222121121211111112122212112221121111111222111111111111221111111111
5109SM	19	12121212112211111111121212211111112112111122211111111111*221111111111
5110SM	20	121221121121211212112211221121111121111112222111111111111211111111111
5111SM	21	121221121122111212112121221111111121111111222111211111111221111111111
5112SM	22	121221121121211111111221122112211211111112222111111111111211111111111
5113SM	23	121221121121211111111221122121221121121112222111111111111211111111111
0375SM	24	12122122221121111122222122112211121211221222111211111111211111111111
0377SM	25	12122112112121111111122112211221121111211222111111111111121111122111
0383SM	26	211221222111221122212222211212211111111121211111111111112212112111111
0388SM	27	121212211122211111112211221122111111121122211111111111111212111111111
0410SM	28	121212221112111111111211212121111121121212211111111111111211111111111
8002SM	29	121221121121211111111212121211111111111112121111211111111221111111111
8003SM	30	121221111121211111111212121211112112111121211112111111111221111111111
8004SM	31	121221111121211111111212121211111121121112121211111111111221111111111
8005SM	32	121221111121211111111221121111111121121112122112111111111111111111111
8007SM	33	121221111121211111111212121211111121121112121111211111111221111111111
8009SM	34	112212111122111111111221212111112211111112212111121111111121111111111
8010SM	35	121221111122111111111212121211112111211112121111211111111221111111111
0467SM	36	121221211111211111111222121222211122211211222111211111112221111211111
5116SM	37	121221121122111111111122121211121112111111222111111111111221111111111
5123SM	38	121221121121211111111222121212211121121121222111111111111221111111111
5133SM	39	121221121122111111111212122211121112111111222111111111111221111111111
5134SM	40	121221121122111111111222121211211121111111222111111111111221111111111
5135SM	41	211222122112111111112122112121111121111111222111111111221211111111111
5136SM	42	121221121122111111111212221111111121111111222111111111111221111111111
5139SM	43	121221121121211111111122121211221112111111222111111111211121111111111
5140SM	44	121221121122111111111212212111111121111111222111211111111221111111111
5141SM	45	121221121122111111111212212111111121111111222111111111111221111111111
5146SM	46	121221122121211111111221212112211121111112222111111111111211111111111
5147SM	47	121222121122111111111212211212111121111111222111111111122221111211111
5159SM	48	211221111121211111111212121211111211111112221111111111112211112111111
0527SM	49	121222211111211111122221212221112111121121211112111121111211111111111
0512SM	50	121222121121121111111211221222111211111112212111121111211221111111111
0531SM	51	121111211121211111112212112222111211112112221111211112111121111111111

DIVISION 2:

5092, 5101, 5102, 5096, 5139, 5146, 5097, 5024, 5091, 5010, 5028, 5040, 5093, 5098, 5004, 5019.

DIVISION 3:

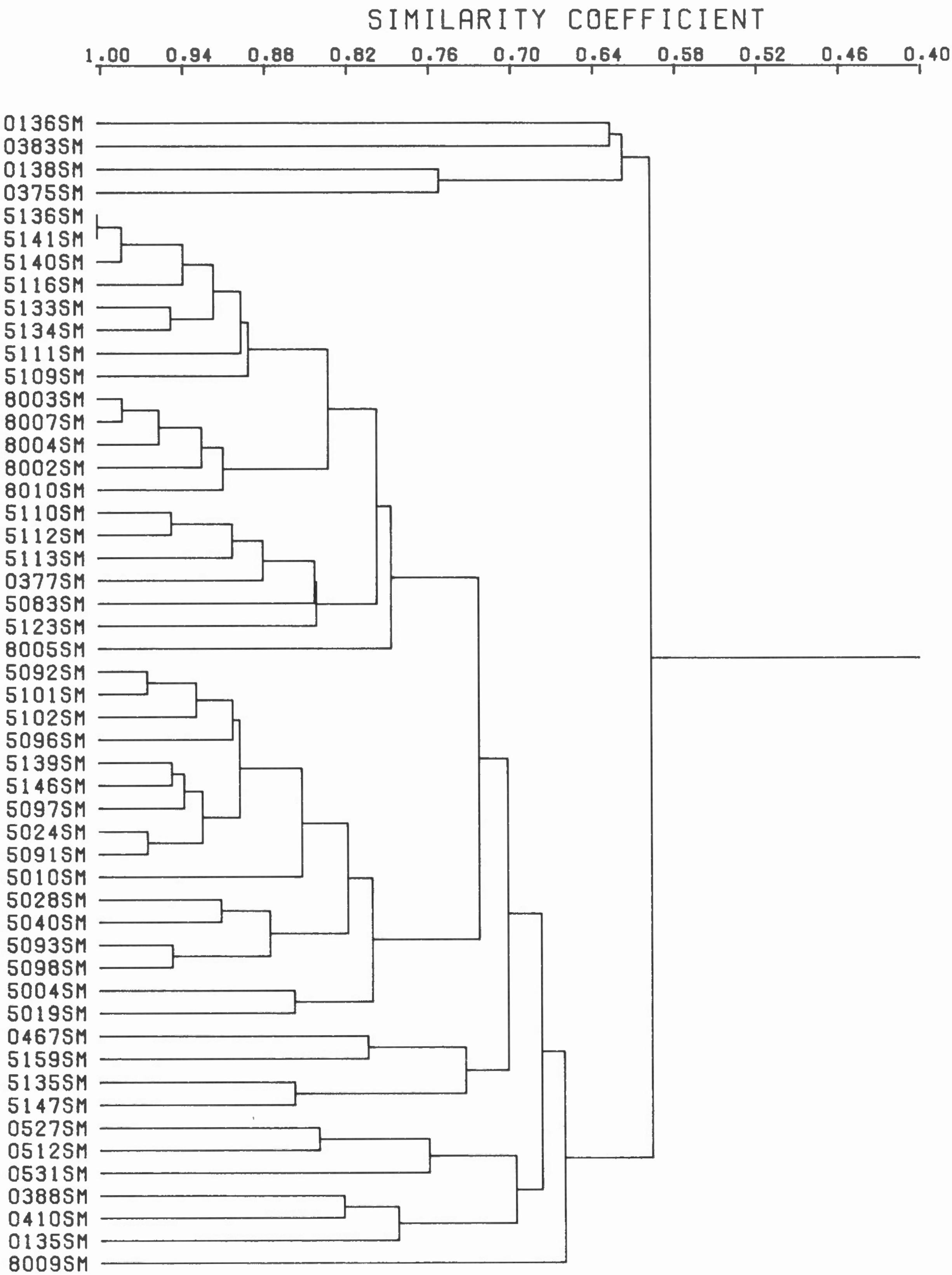
0467, 5159, 5135, 5147, 0527, 0512, 0531, 0388, 0410, 0135, 8009.

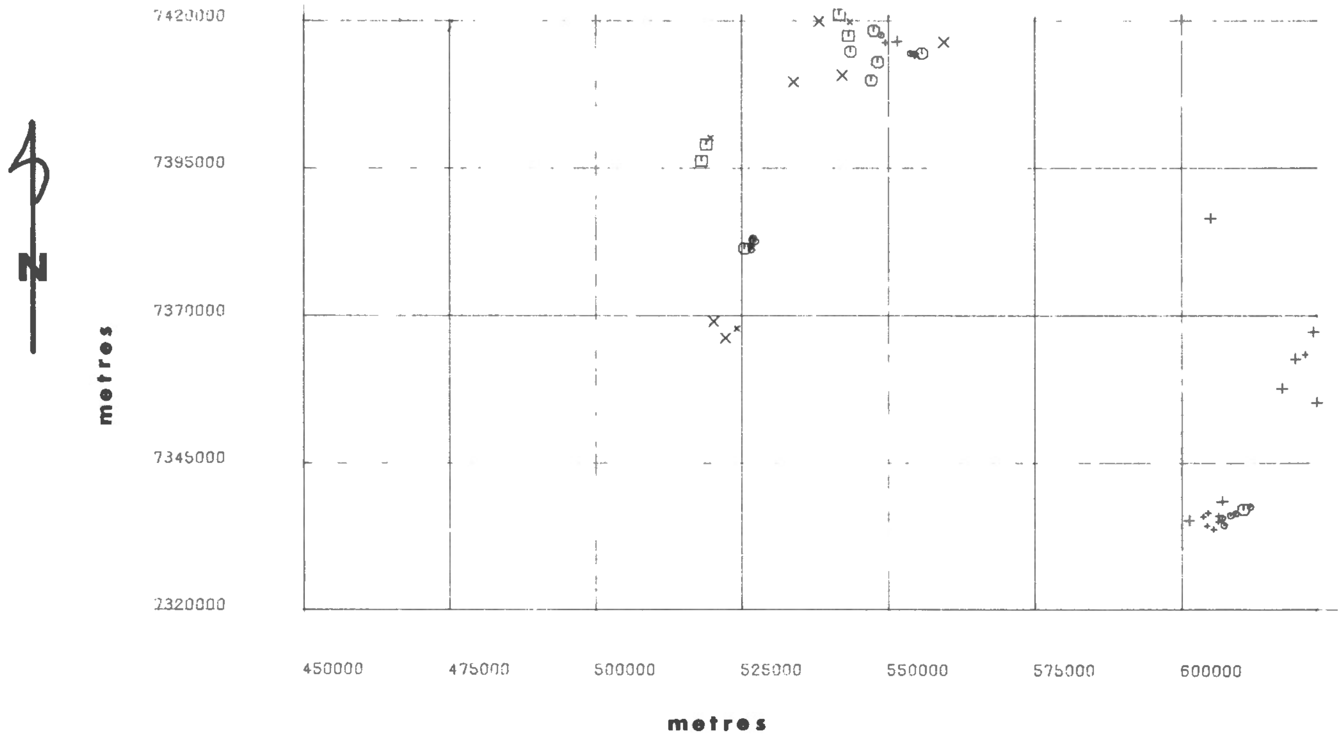
DIVISION 4:

0136, 0383, 0138, 0375.

The seven samples have been assigned to three groups. Division 1 contains the gneiss thought to be from pre-Prince Albert Group complexes as well as two other specimens whose geological position is less well known. Division 2, which includes granite gneisses

FIGURE 3


SIMILARITY COEFFICIENT
1.00
0.94
0.88
0.82
0.76
0.70
0.64
0.58
0.52
0.46
0.40
0136SM
0383SM
0138SM
0375SM
5136SM
5141SM
5140SM
5116SM
5133SM
5134SM
5111SM
5109SM
8003SM
8007SM
8004SM
8002SM
8010SM
5110SM
5112SM
5113SM
0377SM
5083SM
5123SM
8005SM
5092SM
5101SM
5102SM
5096SM
5139SM
5146SM
5097SM
5024SM
5091SM
5010SM
5028SM
5040SM
5093SM
5098SM
5004SM
5019SM
0467SM
5159SM
5135SM
5147SM
0527SM
0512SM
0531SM
0388SM
0410SM
0135SM
8009SM



'OLD' GNEISS — GROUP 2 — GROUP 3 — GROUP 4

MERCATOR PROJECTION

SCALE 1 : 1,000,000

FIGURE 4. 'OLD' GNEISS IN HAYES RIVER REGION.

Table 1.

Similarity coefficient matrix for seven gneisses. Numbers at top and side of matrix are gneiss numbers as described in text.

	4	8	11	26	31	37	44
4	1						
8	.643	1					
11	.964	.643	1				
26	.625	.625	.661	1			
31	.768	.589	.768	.607	1		
37	.929	.643	.893	.661	.768	1	
44	.857	.571	.857	.625	.804	.929	1

Table 2.

Additional data for gneiss groups one to four

Specific Gravity	1 2.66(16)†	2 2.66(12)	3 2.74(10)	4 2.80(3)	Specific Gravity File
Scintrex Readings in μR/hr	15(19)	13(15)	12(9)	12(4)	Geophysical File
Relative Reading Magnetometer in γ (base camp 4000 γ)	3517(12)	4160(15)	3550(2)	4400(3)	Geophysical File
Grain Size	∿2mm(17)*	∿2mm(15)	∿3mm(11)*	1½mm(4)	Field Note File
Layer Thickness	∿1cm(12)	∿1cm(13)	∿5cm(8)	∿2cm(4)	Field Note File

* Grain size varies 4 orders of magnitude for these groups.
† Number of readings used to establish arithmetic average.

already separated by Heywood (1961) contains two specimens. Two specimens are assigned to Division 4; these are the Prince Albert Group gneisses whose protolith was the Prince Albert Group.

In an effort to define the "old gneisses" more accurately other data were retrieved from a multiple file retrieval system (MARS). The only reason for utilizing this system is that retrievals are rapid and orderly. These data are shown in Table 2.

The specific gravity, scintrex readings, and magnetometer values are similar to each other, showing these measurements are not particularly useful as discriminators by themselves. Perhaps one can say, but with the utmost of caution, that the old complex may be slightly more radioactive than the other groups. This is a hypothesis which would have to be evaluated through large expensive sampling schemes and is probably not justified at this time.

If the assignment of group 1 gneisses to "old gneiss" is correct then some possible localities of "old gneiss" in 56J, K are plotted on 1:1,000,000 scale (Fig. 4).

Conclusion

That "old gneiss of heterogeneous origin underlying the Prince Albert Group may be widespread" is a plausible hypothesis which should be followed up with more field work.

The procedure outlined above is, of course, no different from the usual hypothesis-generating procedures of geologists. The main advantage is that numeric data collected at different times is co-ordinated in less time than it would take by hand. This method also ensures that uniform data are collected so that the "inexplicable" outcrops of the early summer can be related to the "well-understood" outcrops of the late summer. Because the clustering requires the generation of Yes, No data in the field, it has been suggested that the method is too time-consuming. This criticism is valid; on the other hand, only after many more man-years have been spent carefully observing gneisses will we know the right questions to ask. At such a time, results more definite than a "plausible hypothesis" may materialize.

Acknowledgments

T.M. Gordon, W.N. Houston and P. Van Western helped with numerous software problems.

References

Bonham-Carter, G.F.
1965: A numerical method of classification using qualitative and semi-qualitative data as applied to Facies Analysis; Bull. Can. Pet. Geol., p. 482-502.

Crain, I.K., and Morris, J.A.
1972: Plotting geological field data, a users guide; Computer Science Centre, Dept. Energy, Mines and Resources.

Heywood, W.W.
1961: Geological notes, Northern District of Keewatin; Geol. Surv. Can., Paper 61-18.

Schau, Mikkel
1973: Volcanic rocks of the Prince Albert Group; *in* Report of Activities, April to October 1972, Geol. Surv. Can., Paper 73-1A, p. 175-177.

1974: Volcanic rocks of the Prince Albert Group; *in* Report of Activities, April to October 1973, Geol. Surv. Can., Paper 74-1A, p. 187-188.

12. COMPUTER USE IN THE GEOCHRONOLOGY SECTION

R.K. Wanless,
Regional and Economic Geology Division

A mini-computer and associate peripherals has been obtained and is being interfaced with six mass spectrometers in the Geochronology Laboratory. The system will be programmed to permit the simultaneous input of 'raw' isotopic data from five of the mass spectrometers. Preliminary isotope abundance calculations will be available, at teletype terminals, immediately following the input stage. From this information the mass spectrometer operator will decide whether to terminate the analytical procedure or to continue to input more data. Programs are being prepared to handle the raw data and to carry out all isotope ratio, age, and statistical calculations. A final phase of development will comprise the computer control of instrumental parameters.

The hardware now installed comprises a Nova 1200 computer with 24k of core, two teletype terminals, a fast paper tape punch and reader and a Nova fixed-head disc with a capacity of 256k words. Cabling between the mass spectrometers and the computer has been installed and some interfacing logic has been assembled. Work will continue on the latter during the coming summer and hopefully all aspects, except the final computer control phase, will be operating by late 1974.

13. STATUS OF COMPUTERIZED FILES IN THE MINERAL DATA BANK, GEOLOGICAL SURVEY OF CANADA

D. R. E. Whitmore,
Regional and Economic Geology Division

Beginning with an experimental prototype (M-file) initiated in 1968, several computer-oriented files have been set up within the Geological Survey Mineral Data Bank. The M-file, a semi-fixed format file was expanded to some 7,000 entries before it became inactive early in 1970. Had the MARS VI file management system been available at the time, the file, though experimental, would probably have been further developed. Valuable retrievals have been made from it at intervals since then, notably in 1972 for listing of gold deposits and their production. Although some of the data in the M-file are faulty and the file as a whole requires editing, it contains much valuable data, particularly on nickel deposits. When priorities permit, it will be converted to allow it to be managed by MARS VI.

A somewhat similarly designed file, the Nb-Ta deposit file was developed in 1970 by K. R. Dawson and used by him for a limited range of retrievals. It contains 600 deposits and is now inactive.

Subsequently Dawson, profiting from the experience of the two earlier files, created a more elaborate Ba-Sr-F deposit file designed for management by MARS VI. It is described in detail by him elsewhere in this report.

When the M-file was rendered dormant, the emphasis in mineral deposit files was then directed, at the urging of the Canadian Centre for Geoscience Data to free-format files (e.g. SAFRAS) and a file (Mineral Deposits File) was built up by A. M. Kelly and assistants. This file also is at the present time inactive. It contains over 4,000 entries, concentrated largely on the location, ownership, development status, and production of deposits, with relatively little geological data.

Difficulties in the use of SAFRAS have led Kelly, together with G. Williams of the University of Alberta and others, to design a more flexible free-format file management system (GDMS). The currently active Mineral Deposits File (OPSEP file) is designed to be managed by this system, which became operational on March 20, 1974, so that it was possible to generate a test file of 102 deposits selected from those included in a recent mineral potential study (Operation September) and make some preliminary retrievals. If successful, the file will be extended to include all 700 Operation September deposits. The application of GDMS is not certain at the moment. Attention is being given to possible conversion of the file so that it, or parts of it, could be managed by MARS.

In existence also within the Data Bank are three other files: (1) Molybdenum file, a restricted version of Dr. Dawson's barite fluorite file compiled by N. Prasad and currently operational with MARS (*see* Prasad et al., this publication). It contains data on some 1,500 molybdenum deposits. (2) Update file used to store development and geological data on deposits currently being explored and reported in the mining press. Using UASAFRAS it produces a periodic publication "UPDATE", and (3) Minplot file, a simple fixed-format file containing development status, name, and location of mineral deposits from which annoted plots of mineral deposits are currently produced on UTM or Lambert conformal projections at various scales.

Two other mineral deposit files not in the Geological Survey should be mentioned. They are the MEPI (Mineral Establishment and Property Index) and the Reserves file, both compiled in the Mineral Development Sector. Both are currently operational using MARS, and are linked to the OPSEP file by the use of a common link number. Their existence will make it possible to concentrate the resources of the Geological Survey on the computerization of geological data in the strict sense. Interaction between the three files is planned.

Proposed use of the OPSEP file is to:

(1) Locate and plot, using various symbols on any required projection and scale, mineral deposits classified as to geological type, mineralogy, development status, commodities, enclosing rocks (specifically or by general, compositional, or genetic class), geological formation, geological province and related metallotect.

(2) Make statistical studies of the occurrence of deposits in terms of the above parameters.

(3) Compile, total and analyze production and reserve figures for deposits as characterized by the same parameters.

In geological terms the output of many mineral deposit files has so far been perceived by geologists as disappointing or trivial. Largely this is because the files contain relatively little geological data in the first place, but also because rapid retrieval has not been obtained before the interest of the person asking the question has vanished. With the OPSEP file the intention is to concentrate on retrieval software from the outset and thus to be in a position to work with the users in the modification of the file and the generation of the data it will require to be geologically useful.

To be fully useful a file must contain "all" of the deposits it purports to contain. In a country with the geological diversity and geographic extent of Canada this is a substantial requirement. For example, there are about 450 significant present and past producers indicated on the Mineral Map of Canada. A study in 1950 by the Mining Association of Canada indicated that only 1 in 1,000 occurrences staked and drilled developed to the stage where they became taxpaying entities (i.e. mines). The number of investigated occurrences on this basis would approach half a million. For a file of any depth a smaller population must be selected and for this reason the OPSEP file will contain only deposits with known or inferred reserves.

Figure 1a

Figure 1. UPDATE sample page.

```
          1         2         3         4         5         6         7
1234567890123456789012345678901234567890123456789012345678901234567890123456789 0
*0101    IDENTIFICATION                         IDENT                 P         01
    REFERENCE-NUMBER                    REF-NUM     I    8            P         02
    RECORD-NAME                         RECNAM      X   32            P         03
    ENTITY-CODED                        ENTCOD      A   10            P    T    04
    LINK-NUMBER                         LINK        I    6                      05
 SS06    OTHER-IDENTIFIERS                      OTHIDENT                        06
    NAME-OR-IDENTIFIER                  NAMOID      X   32                      07
    YEAR-OF-NAME                        YON         I    4                      08

*0201    LOCATION                               LOC                   P         09
    OBJECT-LOCATED                      OBJ         X   16                      10
    LATITUDE                            LAT         F    7     4      P         11
    LONGITUDE                           LONG        F    8     4      P         12
    UTM-ZONE                            ZONE        I    2                      13
    NORTHING                            NORTH       I    7                      14
    EASTING                             EAST        I    6                      15
    UNCERTAINTY-METRES                  UNCRT       I    4                      16
    LATITUDE-DECIMAL-MINUTES            LATDM       X    7                      17
    LONGITUDE-DECIMAL-MINUTES           LONGDM      X    8                      18
 SS04    NTS                                    NTS                             19
    NTS-PRIMARY-QUADRANGLE              NTSP        I    3                      20
    NTS-LETTERED-QUADRANGLE             NTSL        A    1                      21
    NTS-TERTIARY-QUADRANGLE             NTST        I    2                      22
    PROVINCE-TERRITORY                  PROV-T      X   16                 T    23

*0303    PRODUCTION-STATUS                      PRODSTAT                        24
    STATUS                              STAT        A    8                 T    25
    DATE-OF-STATUS                      DOS         I    4                      26
 SS06    COMMODITIES                            COMM                            27
    COMMODITY-NAME                      COMNAM      X   16                      28

*0408    DEPOSIT-CLASSIFICATION                 DEPCLAS                         29
    CLASSIFICATION-SCHEME               CLAS-S      X   20                 T    30
    DEPOSIT-TYPE                        D-TYP       X   16                 T    31
    CLASSIFIED-BY                       CLAS-BY     X   16                      32

*0507    ASSOCIATED-ROCKS                       ASSRCKS                         33
    ROCK-NAME                           ROCK-N      X   32                      34
    MAJOR-ROCK-CLASS                    M-CLAS      X   14                 T    35
    COMPOSITION-CLASS                   C-CLAS      X   14                 T    36
    GENETIC-CLASS                       G-CLAS      X   14                 T    37
    METAMORPHISM                        METAM       A   16                      38
    OTHER-MODIFIER                      OMOD        A   12                      39
    RELATIONSHIP-TO-DEPOSIT             REPDEP      X   16                 T    40
    RADIOMETRIC-AGE-MILL-YR             RAGE        I    4                      41
    RADIOMETRIC-AGE-REF                 RAGE-RF     X   16                      42

*0605    GEOLOGICAL-UNIT                        GEOL-U                          43
    UNIT-NAME                           UNAM        X   20                      44
    UNIT-CATEGORY                       UCAT        X   12                      45
    MINIMUM-AGE-CODE                    MINAGE      I    3                      46
    MAXIMUM-AGE-CODE                    MAXAGE      I    3                      47
    RELATION-TO-DEPOSIT                 RELDEP      X   16                 T    48

*0702    GEOLOGICAL-PROVINCE                    G-PROV                          49
    PROVINCE-NAME                       PROV-N      A   20                 T    50
    TECTONIC-UNIT                       TECT-U      A   32                 T    51

*0801    MINERALS-OBSERVED                      MINOBS                          52
 SS20    MINERAL-LIST                           MINLIST                         53
    MINERAL-NAME                        MINPL       A    4                      54

*0901    DEVELOPMENT                            DEV                             55
    DEVELOPMENT-STAGE                   DEV-S       A    8                 T    56
 SS05    DEVELOPMENT-DATA                       DEV-D                           57
    TYPE-OF-DEVELOPMENT                 DEVT        X   20                      58
    DATE-AT-WHICH-COMPLETED             D-COMPL     I    4                      59
    TOTAL-LENGTH                        T-LENGTH    F    6     1                60
    LENGTH-UNITS                        LENGTH-U    X    8                 T    61
    NO-OF-DEVELOPMENT-UNITS             NODEVU      I    4                      62
```

Figure 1b

```
*1003     PRODUCTION                             PROD                                        63
     ORE-OR-CONCENTRATE                     ORCON       A     4               T        64
     ORE-CONCENTRATE-AMOUNT                 ORC-A       E    12    6                   65
     ORE-CONCENTRATE-UNITS                  ORC-U       X     8                        66
     DATE-FROM-P                            DFP         I     4                        67
     DATE-TO-P                              DTP         I     4                        68
 SS05     PRODUCTION-DATA                        PRDAT                                       69
     CONTAINED-COMMODITY-NAME               COMNAM      X    16                        70
     CONTAINED-COMMODITY-GRADE              COMGRD      F     6    2                   71
     GRADE-UNITS                            GRD-U       X     8               T        72
     COMMODITY-AMOUNT                       COMAMT      E    12    6                   73
     UNITS-OF-AMOUNT                        U-AMT       X     8               T        74

*1102     METALLOGENY                            METALL                                      75
     METALLOTECT-NAME                       M-NAM       A    20                        76
     METALLOTECT-TYPE                       M-TYP       A    16                        77

*1205     RESERVES                               RSRVS                                       78
     ORE-ZONE-IDENTIFIER                    OZID        X    32                        79
     TOTAL-PARTIAL-ASCRIBED                 TOPAS       A     8         P     T        80
     ASCRIBED-TO                            ASTO        X    16                        81
     ASCRIBED-FROM                          ASFROM      X    32                        82
     QUANTITY                               QTY         F    12    6                   83
     QUANTITY-UNITS                         QTYU        X     8               T        84
     ORE-CATEGORY                           ORCAT       X    12               T        85
     RESERVE-SOURCE                         RESS        X    32                        86
     DATE-AT                                DAT         I     4                        87
     ALTERNATIVE-ESTIMATE                   ALTEST      A     3                        88
 SS05     RESERVES-DATA                          RES-D                                       89
     COMMODITY-NAME-R                       COMMNR      X    16                        90
     COMMODITY-GRADE-R                      COMMGR      F     6    2                   91
     GRADE-UNITS-R                          GRUR        X     4               T        92

*1303     OWNERSHIP                              OWNR                                        93
     OWNER-OPERATOR                         OWNOP       X    32                        94
     WHETHER-OWNER-OPERATOR                 OWN-OP      A     8               T        95
     DATE-AT-OWNS                           DATW        I     4                        96

*1403     COMMENTS                               COMTS                                       97
     APPLICABILITY                          APPL        X    20               T        98
 SS20     REMARKS                                RMRKS                                       99
     COMMENT                                COMT        X    32                       100

*1508     REFERENCES                             REFS                                       101
     AUTHOR-NAME                            AUTH        X    32                       102
     TYPE-OF-DOCUMENT                       TOD         X    12                       103
     YEAR                                   YRD         I     4                       104
     TITLE                                  TIT         X    32                       105
     SERIES-NAME-NUMBER                     SNAN        X    60                       106
     PAGE-REFERENCE                         PAGE        X    16                       107
     DATA-ITEM-STRING-XREF                  DISX        X    30                       108
 SS03     MAPS                                   MAPS                                       109
     MAP-IDENTIFIER                         MAPID       X    12                       110
     MAP-TYPE                               MAP-T       X    12                       111
     MAP-SCALE                              MAP-S       X    12                       112

*1601     COMPILATION                            COMPIL                                     113
     COMPILED-BY                            COM-BY      X    16         P             114
     DATE-COMPILED                          DATEC       I     6         P             115

*1701     REVISIONS                              REVS                                       116
     REVISED-BY                             REVBY       X    16         P             117
     DATE-REVISED                           DREV        I     6         P             118
 SS10     REVISION DETAILS                       REV-D                                      119
     ITEMS-REVISED                          ITSREV      X    30         P             120
1234567890123456789012345678901234567890123456789012345678901234567890
         1         2         3         4         5         6         7
```

Figure 2. SPECIFICATIONS OPSEP FILE.

VOL. 2, NO. 5 P. 11

ITEM	COMMODITIES	PROPERTY , LOCATION & COMPANY	PROV	N.T.S.	ACTIVITY & DATA	REFS
19	PB ZN	- CUVIER PROPERTY INCLUDES ORIG. GAYS RIVER PROP - GAYS RIVER AREA APPROX. 30 MI NE OF HALIFAX - CUVIER MINES LTD (OPT. TO IMPERIAL OIL LTD)	NS	11 E 3 45.0319 063.3583	DRILLING HAS INTERSECTED MINERALIZATION BOTH EAST + WEST OF AN AREA PARTIALLY DRILLED IN 1950'S. 15 OF 36 HOLES HAVE RETURNED 'SIGNIFICANT' PB-ZN VALUES. OFFICIALS FEEL THERE COULD BE A SERIES OF DEPOSITS IN THE MISSISSIPPIAN LIMESTONE. ASSAY RESULTS : HOLE NO. 99 (1.25 MI E OF AREA DRILLED IN 1950S) - 24.5 FT (417.5-442 FT) - 1.88 % ZN, 6.10 % PB. (LAST 10 FT - 3.32 % ZN, 11.90 % PB.)	NORTHERN MINER MAY 10, 1973 (73 05 10) P. 23 UPDATE VOL 1 NO 3 (72 11 00) P. 9 ALSO... MIN. OCCUR INDEX 11 E 3 PB I

VOL. 2, NO. 5 P.

ITEM	COMMODITIES	PROPERTY , LOCATION & COMPANY	PROV	N.T.S.	ACTIVITY & DATA	REFS
20	AG PB ZN AU CU	- HORN SILVER MINE - WEST SLOPE, RICHTER MTN. 12 MI SE OF KEREMEOS - DANKOE MINES LTD (FORMERLY UTICA MINES LTD)	BC	82 E 4 49.0583 119.6889	COMPANY CONSIDERING RESUMING PRODUCTION. A CONSULTANTS' REPORT RECOMMENDS EXPLORATION AND DEVELOPMENT ON THE NORTHWARD PROJECTION OF THE 'N' VEIN AND ON THE DOWN DIP PROJECTION OF THE 'B' VEIN. ORE RESERVES (MARCH,1970) : PROBABLE ORE- 71,250 TONS @ 11.7 OZ/TON AG. POSSIBLE ORE- 73,700 TONS @ 6.8 OZ AG. PREVIOUS PRODUCTION (AUG 1967- MAR 1970): 258,009 TONS @ 9.3 OZ/TON AG.	NORTHERN MINER MAY31 ,1973 (73 05 31) P. 17 ALSO... MIN OCCUR INDEX 82 E 4 AG 1

THE GEOCHEMICAL INTERACTIVE GRAPHICS SYSTEM

Ian K. Crain,
Chief, Planning and Evaluation - EDP, Health Protection Branch,
Health and Welfare, Canada

Introduction

The Geochemical Interactive Graphics System (GIGS) was designed for the Regional and Economic Geology Division of the Geological Survey. The system was designed and implemented by the EMR Computer Science Centre during the period April to February 1973. This was the first operational graphics system in the Department of Energy, Mines and Resources.

The basic objective of GIGS was to provide a facility for displaying and manipulating rock geochemical data through machine - scientist interaction, thus enabling rapid data-editing and verification and on-line interactive investigation of geochemical relationships using a cathode ray tube display. It was intended that the system would be an initial phase of a more elaborate interactive graphics facility which would encompass more sophisticated numerical data analysis techniques, and provide service to some other geological disciplines. At the same time, it was felt that the experience gained from the system development would assist in weighing the capabilities and utility of interactive graphics against their cost.

Design Criteria

The global design criterion was simply that the system facilitate the graphic presentation of rock-geochemical data, allowing the scientist rapid access to, and interaction with his data. More specific criteria follow the guidelines of Crain (1974).

More specific criteria were:

(1) Simplicity of operation

The system was to be operated by scientists, not computer-scientists. For this reason, it was essential that the system be accessible by a small number of simple commands and that the programs themselves be as self-instructing as possible.

(2) Interactions

It was obviously desirable that the scientist be able to control the operation of the programs and input his own specifications for the particular displays. On the other hand, it was also desirable to have reasonable and useful "defaults" to produce displays with minimum human intervention for standard of preliminary displays.

(3) Convertibility

The system was implemented on a PDP-10 computer time-sharing system (Dataline Systems Ltd.). It was expected that interactive graphics support would be available from the EMR Computer Science Centre in the future, so that it was highly desirable to have programs which could be converted easily to other systems.

To accomplish this flexibility, the graphics programs were coded wherever possible in a universal subset of FORTRAN IV. Some portions of the system are inevitably dependent on the Dataline Systems operating system FORTRAN requirements and on the selective retrieval programs (called MPLDR) used by the graphics programs. This dependent coding was kept in separate modules to ease conversion efforts.

The generation of displays through FORTRAN subroutine calls makes use of programs written by the terminal manufacturer (the Tekplot package). The manufacturer has threatened major upgrading (hence changes) of these programs, and various models of their hardware require various software drivers. In expectation of problems in this area, all display generation in the main applications programs were coded with 'Calcomp' style subroutine calls. This serves two purposes: (1) allows a standardization of coding, using a series of calls which are familiar to many and fairly common in the computer industry, and (2) allows any programs currently using Calcomp pen-plotter to be instantly usable on the CRT display.

This feature was achieved through a series of subroutines which emulate the functions of the Calcomp routines by calling equivalent routines in the Tekplot package. This emulation is strictly dependent on the 'driver' or software for the particular terminal.

(4) Low-cost operation

Criteria 1 through 3 are partially in opposition to this criterion. However, within the framework of the system, coding was written with an eye to economy of operation.

System Configuration

General

Graphic output is accomplished using a Tektronix 4002A Graphic Computer Terminal, connected to a 4601 Hardcopy Unit. For the GIGS system the terminal is connected by telephone lines at 300 baud to Dataline Systems Ltd. in Ottawa and then by higher speed lines to the Dataline PDP-10 computer in Toronto.

The basic graphics software is the 'Tekplot' package provided by Tektronix. Supervisory programs were written by the EMR Computer Science Centre and are described in detail in system documentation at the Geological Survey.

Selective retrieval of items from the main data file is accomplished by a Dataline System program called MPLDR. All this software is disc-resident in the PDP-10. Figure 1 shows the system configuration.

Graphic Hardware

The Tektronix 4002A Graphic Terminal features

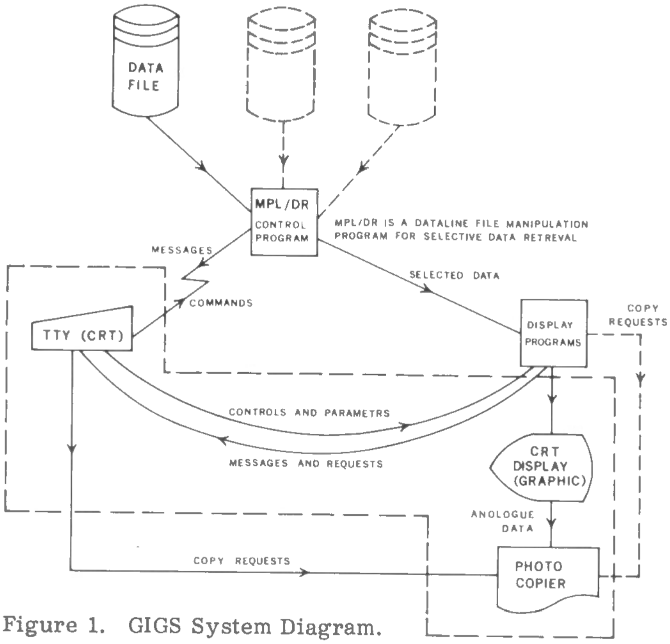


Figure 1. GIGS System Diagram.

a display screen approximately 10 inches wide by 8 inches high with a resolution of about 1/100 inch. Data may be entered through a teletype compatible keyboard, which contains the usual character set and control keys, with some additional control buttons for erasing the screen and repositioning the graphics cursor. The device can operate at information transfer rates as high as 4800 baud and is capable of accepting "graphic" input from a joystick operated cross-hair. The joystick feature is not used by GIGS and the transfer rate is restricted to 300 baud by the computer installation.

The CRT screen is illuminated by a long persistence phosphor which produces a continuous unflickering display for about sixty seconds. The display may be refreshed electronically as many times as desired by user control.

The 4601 Hardcopy Unit is capable of making paper copies of the display screen both through manual or program control. The display is automatically disabled for about 5 seconds while the print is being made. Prints therefore cannot be made while a plot is still in progress.

Software

Subroutines for drawing line segments and annotating plots are contained in the 'Tekplot' package purchased from Tektronix Inc. For selective storage and retrieval of data items, the system employs a Dataline Systems Ltd. program called MPLDR. Supervisory programs were written by the Centre for performing the specific graphics tasks.

Typical Output

The system is capable of producing four types of displays, as follows:

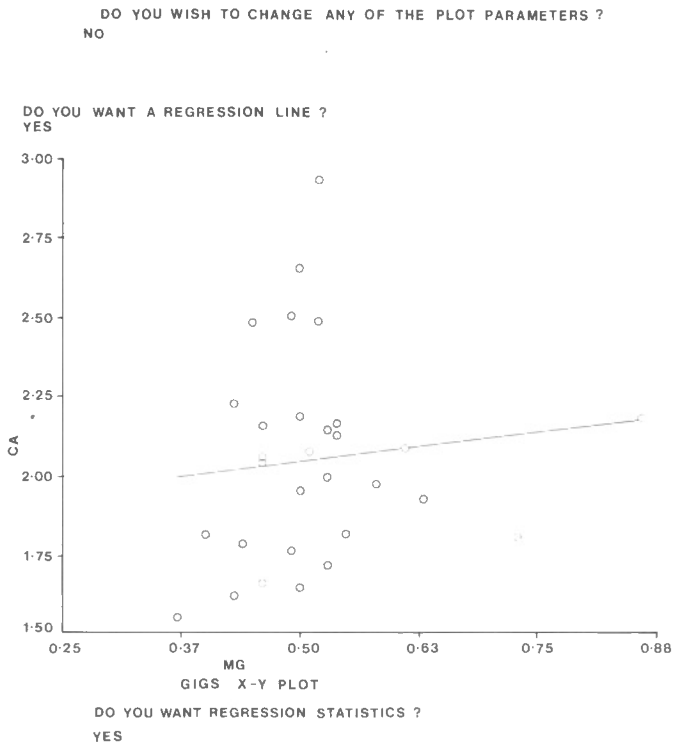


Figure 2. Example of X-Y plot output.

2BUTESØ1A	8.34ØØ	Ø.73ØØ	1.81ØØ	1.85ØØ
2BUTESØ1B	8.45ØØ	Ø.96ØØ	2.19ØØ	2.Ø2ØØ
2BUTESØ2A	8.59ØØ	Ø.61ØØ	2.Ø9ØØ	2.35ØØ
2BUTESØ2B	8.23ØØ	Ø.44ØØ	1.79ØØ	2.Ø4ØØ
2BUTESØ3A	7.7ØØØ	Ø.46ØØ	2.16ØØ	1.77ØØ
2BUTESØ3B	1Ø.ØØØØ	Ø.52ØØ	2.94ØØ	2.53ØØ
2BUTESØ4A	8.17ØØ	Ø.45ØØ	2.49ØØ	1.75ØØ
2BUTESØ4B	7.7ØØØ	Ø.5ØØØ	2.66ØØ	1.63ØØ
2BUTESØ5A	7.39Ø9	Ø.37ØØ	1.55ØØ	1.77ØØ
2BUTESØ5B	7.87ØØ	Ø.54ØØ	2.13ØØ	2.Ø9ØØ
2BUTESØ6A	7.98ØØ	Ø.46ØØ	2.Ø6ØØ	2.Ø8ØØ
2BUTESØ6B	8.28ØØ	Ø.53ØØ	2.15ØØ	2.2ØØØ
2BUTESØ7A	7.97ØØ	Ø.52ØØ	2.49ØØ	1.67ØØ
2BUTESØ7B	6.73ØØ	Ø.43ØØ	1.62ØØ	1.45ØØ
2BUTESØ8A	8.2ØØØ	Ø.55ØØ	1.82ØØ	2.13ØØ
2BUTESØ8B	8.44Ø9	Ø.63Ø9	1.93ØØ	2.12ØØ
2BUTESØ9A	8.75ØØ	Ø.58ØØ	1.98ØØ	2.Ø7ØØ
2BUTESØ9B	8.27ØØ	Ø.53ØØ	1.72ØØ	1.97ØØ
2BUTES1ØA	8.13ØØ	Ø.5ØØØ	1.65ØØ	2.Ø6ØØ
2BUTES1ØB	7.81ØØ	Ø.46ØØ	2.Ø4ØØ	1.93ØØ
2BUTES11A	7.71ØØ	Ø.46ØØ	1.66ØØ	1.94ØØ
2BUTES11B	8.6ØØØ	Ø.49ØØ	2.51ØØ	2.Ø6ØØ
2BUTES12A	8.19ØØ	Ø.51ØØ	2.ØØØØ	1.96ØØ
2BUTES12B	8.15ØØ	Ø.54ØØ	2.17ØØ	1.83ØØ
2BUTES13A	7.97ØØ	Ø.4ØØØ	1.82ØØ	2.Ø6ØØ
2BUTES13B	9.Ø4ØØ	Ø.5ØØØ	1.96ØØ	2.ØØØØ
2BUTES14A	8.42ØØ	Ø.53ØØ	2.ØØØØ	2.Ø7ØØ
2BUTES14B	8.51ØØ	Ø.49ØØ	1.77ØØ	2.ØØØØ
2BUTES15A	7.Ø6ØØ	Ø.43ØØ	2.23ØØ	1.68ØØ
2BUTES15B	8.ØØØØ	Ø.5ØØØ	2.19ØØ	1.83ØØ

END OF PAGELIST

Figure 3. Example of data tabulation output.

1) X-Y co-ordinate plots with suitable scaling (Fig. 2)*.
2) Table listings of selected data items from the file (Fig. 3).
3) Ternary diagrams (Fig. 4).
4) Frequency distributions (Fig. 5).

The user commands which request these plots have a simple English-like structure and are self-prompting.

---

*Editors' Note: Although the Hardcopy Unit produces adequate prints for working use, these figures have been re-drawn for purposes of publication.

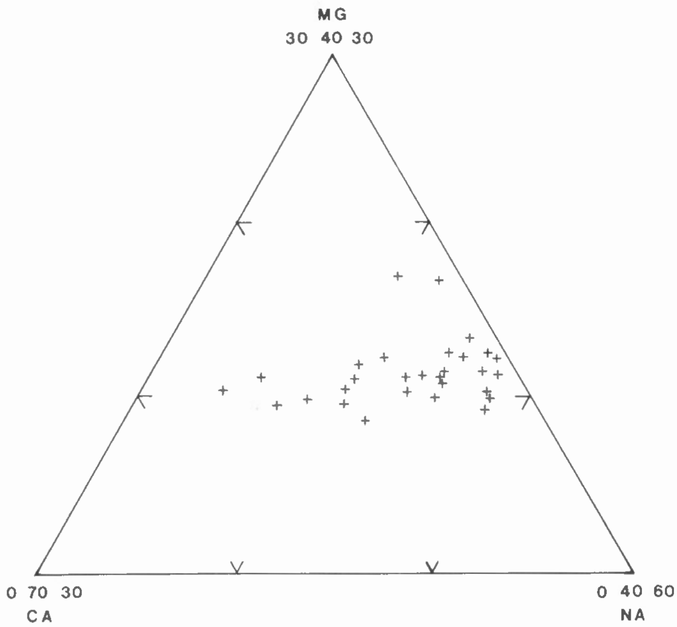


Figure 4. Example of ternary diagrams output.

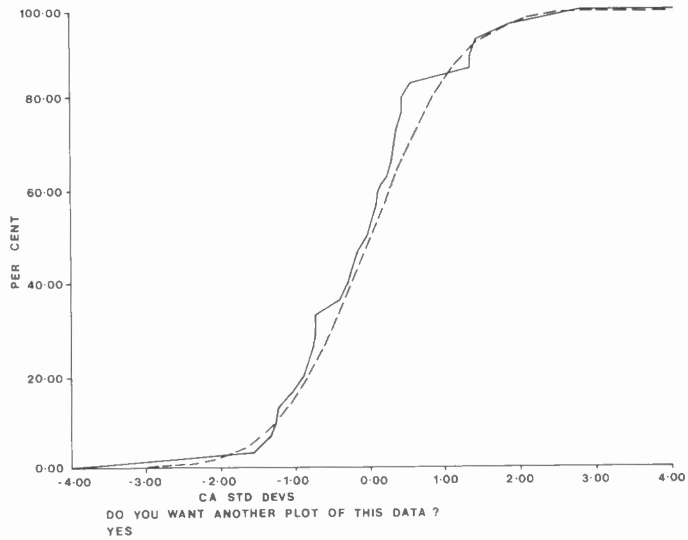


Figure 5. Example of frequency distribution output.

Preliminary System Evaluation

The operational benefits of the system are still being evaluated, with the initial impression being highly favourable. The cost for developing the system has been studied in some detail and can be summarized as follows:

Computer costs of developing software	$ 8,000.00
Disc storage of data files during development	5,000.00
Tektronix terminal and Hardcopy Unit purchase	8,000.00
TOTAL -	$21,000.00

Manpower: Senior Systems Analyst (6 man-months)
Programmers (4 man-months)

Cost of systems use:

A typical modest production use of the system might be to selectively retrieve the data for an X-Y plot (about 200 points) and obtain two plots at differing scales. The cost would be approximately:

CPU time	$10.00
Connect time	$ 5.00
Disc storage (prorated))	$ 5.00
TOTAL	$20.00

These costs demonstrate that the development and use of interactive graphics systems is fairly expensive. This cost must be weighed against the well-known benefits of man-machine dialogue (Martin, 1973). The saving of a scientist's time with this system is immense, thus on the assumption that his time is valuable, the system is probably cost effective.

Acknowledgments

Programmers Mr. J. Cameron and Ms. J. Dawes assisted in the implementation of the system. Dr. D.K. Dawson provided overall supervision of the computer aspects and Mr. M. Singer of Dataline Systems Ltd. was extremely helpful in areas concerned with their software. Dr. R.G. Garrett provided the initial impetus and overall direction to this project.

References

Crain, I.K.
1974: The analysis of scientific computer systems; INFOR, in press.

Martin, J.
1973: The design of man-computer dialogues; Prentice Hall, Englewood Cliffs, N.J., p. 557.

15. COMPUTER USE IN THE ELECTRICAL METHODS SECTION

L. S. Collett,
Resource Geophysics and Geochemistry Division

A mini-computer, Datagen Nova 1230, Type 8164, is presently being interfaced and programmed for controlling electrical measurements on rock and soil samples in the Electrical Rock Property Laboratory. Some use of this mini-computer will be applied to calculations of the measurements for determining the conductivity, dielectric constant and dissipation factor characteristics.

Geophysical field data are now recorded on cards. The information on these cards is punched on computer cards. Programs have been written to reduce the field measurements to geophysical parameters commonly used and to plot profiles of the compiled data.

The Calcomp Plotter is being used to plot charts and graphs of geophysical data. This plotting was formerly hand drawn by a draftsman. This procedure has greatly speeded up the preparation of geophysical papers for publication of papers and reports.

Robert G. Garrett,
Resource Geophysics and Geochemistry Division

A review of EDP usage in the Geochemistry Section of the Resource Geophysics and Geochemistry Division

Electronic data processing is used in all phases of the work of the Geochemistry Section; however, usage can be split into three major groups. Firstly, process control applications in the laboratories related to instrument automation. Secondly, the merging of machine-acquired data with manually-acquired field and laboratory data, and the subsequent archival storage of project data. Third, and lastly, the mathematical and statistical manipulation of the total data to aid the geochemist in data quality control, interpretation and presentation. These three general themes will be discussed in greater detail and some examples cited of specific applications.

Other than programmable calculators, the Section operates two minicomputors, a Hewlett Packard 2114A (8k) and a Texas Instruments 960A (16k). The former is interfaced to an ARL 40-channel optical spectrometer. This spectrometer is used in a variety of modes with the appropriate process control and data reduction program being entered from paper tape. The programs include report generators and results are both neatly tabulated and punched onto paper tape for subsequent processing. For major and minor element determinations the sample is fed, after a $Li_2B_4O_7$ fusion and grinding, through an AC spark on a continuous tape feed (Cameron and Horton, 1967; Danielson, 1968; and Jones and Horton, 1971). For rapid scan work where a large number of elements are to be determined in samples with a relatively constant matrix composition a DC arc method is used (Timperley et al., 1973). The Texas Instruments 960A is used for process control and data reduction with a Perkin Elmer 303 atomic absorption spectrophotometer (Bristow, 1974). The automated A.A. system has been used for both major and trace element determinations. Currently, the results are punched onto paper tape for transmission to the Departmental Computer Science Centre. In the future it is proposed to store the results on cassette magnetic tapes for entry to the centralized computing facility via normal telephone lines. Additionally, it is intended to service a second A.A. with the T.I. 960A; development work to attain this objective has been commenced.

Field data is collected on pre-printed notebook forms which act as keypunch input documents. These field cards have been developed over the last 10 years specifically for exploration geochemical surveys (Garrett, 1974). The results of analyses made in the Section laboratories are either entered onto pre-printed report sheets designed as keypunch input documents, or are recorded on punched paper tape. With small sets of data the resulting punched cards are often merged by hand and then the total field and analytical data file stored on magnetic tape or disc for processing. Large data sets are merged by machine using the SORT/MERGE package or small custom programs that simultaneously make a variety of verification checks. Machine merging has become mandatory where analytical results are recorded on punched paper tape as cards do not exist to be physically merged. In 1974 M.T. Holroyd developed an extremely efficient merge package for the Geochemistry Section field and analytical data which took full advantage of the information content of the sample numbering system which includes the N.T.S. map-sheet number, year of collection and field party. This package will accept data from cards, paper tape, or directly off magnetic tape or disc storage and will see increasing use in the future.

The Geochemistry Section project data files are not currently stored within the GEODAT System. GEODAT is used for relatively small sets of data from samples often analyzed for different subsets of elements or compounds. Also it does not contain the field data that are considered essential for an adequate interpretation in exploration or applied geochemical studies. In contrast the geochemical project files are integrated field and analytical files which are internally consistent and often large, e.g., the Bear-Slave lake sediment file of 90 variables for 4,000 samples or the Yukon granitoid study with 40 variables for some 3,000 samples. Currently, all machine processible files going back as far as 1964 are being checked and prepared for archival storage using GAS (Geochemical Analysis System), a simple file management system, with some selective retrieval and post-retrieval processing capability. GAS was developed at the Department of Geological Sciences, Queen's University, under the general direction of I. Nichol. The project resulted from suggestions made by the Geochemistry Section of the Geological Survey and was funded through an E.M.R. Research Agreement. A GAS file consists of two parts. Firstly, an extended text section in which all relevant historical and other details of the project, field and analytical methods, and codes, etc. used in the data file are described. Secondly, the project data follows as an internally consistent block of integrated field and analytical data. On the completion of data preparation and magnetic tape file generation the files relevant to completed projects, or Open Filed projects, will be made available to the public in the GAS file format.

The merged field and analytical data files are routinely processed using a variety of mathematical and statistical techniques. The extent to which machine assisted interpretation is taken is largely a function of the interests of the project leader. At one end the processing may consist of studying overall data statistics, and at the other, the application of sophisticated multivariate methods. Prior to data analysis the quality of the data is assessed in terms of the field sampling and analytical errors using a variety of

TABLE 1

URANIUM IN LAKE SEDIMENTS

DATA (PPM) GROUPED BY LAKE CATCHMENT AREA MAJOR LITHOLOGY

Structural Province	Division Sub-Province	Rock-type	No. of Samples	Range	Arithmetic Mean	Std. Dev.	C.V.%	Geometric Mean	($\log_{10}$) Std. Dev.
Bear			1078	0.1 - 300	5.6	13.7	245	2.7	.4770
	Bear Batholith		459	0.2 - 160	7.8	12.1	155	4.7	.4089
		Porphyries	54	0.4 - 35	4.4	5.1	117	3.0	.3626
		Granodiorite	120	0.8 - 70	8.3	9.8	118	5.6	.3653
		Granite	244	0.2 - 90	8.3	11.0	132	5.1	.4112
	Hepburn Batholith		469	0.1 - 300	4.8	16.6	346	2.2	.4337
		Granodiorite	49	1.1 - 80	7.0	12.0	171	4.0	.4039
		Granite Gneiss	172	0.1 - 65	2.8	6.2	223	1.6	.3793
		Granite	58	1.0 - 300	14.7	42.9	292	5.0	.4965
		Migmatite	113	0.3 - 43	3.7	5.0	136	2.3	.3944
		Quartzite	30	0.8 - 10	2.4	2.1	89	1.9	.2622
	Epworth Group		150	0.1 - 14	1.3	1.5	116	0.9	.3563
		Dolomite	45	0.2 - 14	1.7	2.0	122	1.2	.3138
		Shale	55	0.1 - 8	1.0	1.1	112	0.7	.3540
		Sandstone	27	0.3 - 8	1.7	1.6	92	1.3	.3267
Slave			2692	0.1 - 70	1.7	2.9	174	1.0	.3992
		Cordierite Schist	707	0.1 - 7	0.9	0.8	81	0.7	.3401
		Granodiorite	911	0.1 - 70	1.8	4.0	220	1.1	.4143
		Granite Gneiss	318	0.1 - 15	1.8	1.9	101	1.3	.3814
		Granite	489	0.1 - 45	2.4	2.9	121	1.8	.3287
Churchill			114	0.1 - 3	0.9	0.5	62	0.7	.3212
		Granite Gneiss	107	0.1 - 3	0.9	0.5	63	0.7	.3277

analysis of variance methods (e.g. Garrett, 1972, 1974c). If the quality of the data is not sufficient to warrant rigorous data analysis only simple data studies are undertaken. A major portion of the effort expended in the Geochemistry Section in the EDP sphere has gone into the testing and evaluation of known data analytical methods and to the development of new methods. It is fully realized that the computer cannot replace the geochemist in making an adequate data interpretation. Nevertheless, the computer can be used to manipulate and present the data in such a way as to aid interpretation and data presentation.

Historically, most of the Section's EDP work has been carried out in the batch mode. However, interactive graphics (GIGS, *see* Crain, in this publication) and other time sharing services are now being used. Interactive graphics provides the tool by which the geochemist can efficiently search for the best methods of data manipulation and presentation so as to convince his audience of the substance and validity of his interpretation. Normal time sharing is used in the sorting and merging of files and to obtain solutions to simple problems on large sets of data.

One of the most useful procedures used in the Section to aid interpretation is the splitting of the data into a variety of groups, based on geology or other field observations, and subsequent computation of simple statistics to define the separate geochemical populations. An example is given in Table 1 where data for uranium in lake sediments from Operation Bear-Slave (Allan et al., 1973) has been split into groups on the basis of geological province, subprovince, and rock type. A variety of more sophisticated methods have been used over the years. Discriminant analysis has been frequently employed in both published and unpublished studies (Cameron, 1969; Cameron et al., 1971), the particular strength of this method comes in using the function generated from the training data set to classify unknown samples into one of two or more groups, one of which may be directly, or indirectly, related to mineralization of economic interest. Principle component and factor analysis have been used in a variety of studies to transform the response oriented chemical data to a new co-ordinate system which is process oriented and thus easier to interpret. These methods have found particular, but not exclusive, application in bedrock geochemical studies, e.g., Cameron (1968), and Garrett (1969, 1972). A variety of forms of multivariate and univeriate cluster analysis have been investigated. A nonlinear mapping algorithm has been used to group rock geochemical data into clusters of similar plutonic intrusives on the basis of major element composition, Garrett (1972). Linear clustering has been used to elucidate individual element frequency distributions in addition to histograms and cumulative frequency plots (Garrett, 1974). Multilinear regression has been employed with considerable success on data from both the primary and secondary geochemical environments. The method was used to compensate for variations in the composition of granitoid rocks and provide a dynamic threshold for base metal levels, whose content in the rocks under study was known to vary partly as a function of the mineralogical, and thus major element chemical, composition of the rock (Garrett, 1972). In lake sediments the effects of iron, manganese and organic content on the zinc content of the sediments was studied with the aid of multilinear regression (Davenport et al., 1974). An inspection of the zinc residuals after regression and the raw zinc data, both plotted as maps, shows that the residuals correlate better with geological features known to control the zinc distribution than does the raw zinc data.

The reader will have possibly noticed that there has been a lack of references to contouring methods. Both the author, before he joined the Geological Survey, and Cameron (1968) have experimented with polynomial surface fitting. From this work we have concluded that the method has a limited application to either very broad regional surveys, or very detailed local surveys, and in either case the order of the surfaces should not exceed quartic at the maximum. Bearing in mind the heterogeneity of individual rock-units and whole survey areas we are not convinced that a polynomial model is appropriate as we do not have, in the vast majority of cases, a sufficient knowledge of the true pattern of areal variation and the included discontinuities. Instead, we have studied the individual samples within a univariate or multivariate system to see where the samples' chemistry placed it relative to known anomalous populations. The multivariate studies include those referred to already and univariate studies have been made relative to such elements as sulphur, mercury and tungsten (Cameron, 1974; Garrett, 1971, 1974b, c). However, we are still faced with the problem of areal data presentation. Two approaches have been used to date, contoured maps, e.g., Allan et al., 1974, and symbol maps where symbols are used to indicate which histogram group of the data any particular sample falls into, e.g., Garrett, 1972. The contoured maps were prepared using the Calcomp GPCP package and its interpolation function is not considered entirely suitable for geochemical data. Thus, J.D. Hobbs has been preparing a contouring package which it is felt will produce a more realistic geochemical contoured map. This program is being written to produce EAI plot tapes so maps can be drawn directly onto prepared bases for Open-File quality publication.

In conclusion, our experience with EDP continues to grow and our objective is twofold. Firstly, to relegate those routine tasks which can be carried out by computer to the machine so that trained technical staff can be released for more productive work. Secondly, to gain sufficient experience of the various mathematical and statistical aids to interpretation so that the correct and most applicable one can be chosen to solve any particular interpretational problem.

References

Allan, R.J., Cameron, E.M., and Durham, C.C.
1973: Reconnaissance geochemistry using lake sediments of a 36,000-square-mile area of the northwestern Canadian Shield; Geol. Surv. Can., Paper 72-50, 70 p.

Allan, R.J., Cameron, E.M., and Jonasson, I.R.
1974: Mercury and Arsenic levels in lake sediments from the Canadian Shield. Paper presented at the First International Mercury Congress, Barcelona, Spain.

Bristow, Q.
1974: Geochemical instrumentation; *in* Report of Activities, April to October 1973, Geol. Surv. Can., Paper 74-1A, p. 51.

Cameron, E.M., and Horton, R.E.
1967: Analysis of rocks using a multichannel emission spectrometer; Chem. Geol., v. 2, p. 135-145.

Cameron, E.M.
1968: A geochemical profile of the Swan Hills Reef; Can. J. Earth Sci., v. 5, p. 287-309.

1969: Regional geochemical study of the Slave Point carbonates, Western Canada; Can. J. Earth Sci., v. 6, p. 247-268.

1974: Sulphur in Archean volcanic rocks of the Canadian Shield; Geol. Surv. Can., Paper 74-18, 9 p.

Cameron, E.M., Siddeley, G., and Durham, C.C.
1971: Distribution of ore elements in rocks for evaluating ore potential; nickel, copper, cobalt and sulphur in ultramafic rocks of the Canadian Shield; *in* Geochemical Exploration, Can. Inst. Mining Met., Spec. Volume 11, p. 298-313.

Danielsson, A.
1968: Spectrochemical analysis of geochemical purposes; *in* XIII Colloquium Spectroscopium Internationale, Ottawa. Hilger, London, p. 311-323.

Davenport, P.H., Hornbrook, E.H.W., and Butler, J.A.
1974: Regional lake sediment geochemical survey for zinc mineralization in western Newfoundland; Paper presented at the 5th International Geochemical Exploration Symposium, Vancouver, in press.

Garrett, R.G.
1968: Factor analysis as an aid in the interpretation of regional geochemical stream sediment data; Quart. Col. Sch. Mines, v. 64, p. 245-264.

1971: Molybdenum, tungsten and uranium in acid plutonic rocks as a guide to regional exploration, S.E. Yukon; Can. Mining J., v. 92, p. 37-40.

1972: Regional geochemical study of Cretaceous acidic rocks in the northern Canadian Cordillera as a tool for broad mineral exploration; *in* Geochemical Exploration 1972, Inst. Mining Met., London, p. 203-219.

1974a: Field data acquisition methods for applied geochemical surveys at the Geological Survey of Canada; Geol. Surv. Can., Paper 74-52, 36 p.

1974b: Copper and zinc in Proterozoic acid volcanics as a guide to exploration in the Bear Province; Paper presented at the 5th International Geochemical Exploration Symposium, Vancouver, in press.

1974c: Mercury in some granitoid rocks of the Yukon and its relation to gold-tungsten mineralization; J. Geochem. Expl., v. 3, p. 277-289.

Jones, F.W., and Horton, R.E.
1971: A data acquisition system with computer control for an optical emission spectrometer; Can. Spectroscopy, v. 16, p. 1-4.

Timperley, M.H., Horton, R.E., and Lynch, J.J.
1973: The analysis of metals in geological materials by D.C. Arc direct reading emission spectrometry; *in* Report of Activities, April to October 1972, Geol. Surv. Can., Paper 73-1A, p. 63-64.

COMPUTER PROCESSING OF AIRBORNE GAMMA-RAY SPECTROMETRY DATA

R. L. Grasty,
Resource Geophysics and Geochemistry Division

Since 1967, the Geological Survey of Canada has been developing a high-sensitivity airborne gamma-ray spectrometry system for the purpose of mapping surface concentrations of potassium, uranium and thorium. This report summarizes the data processing procedures carried out on the airborne data from the original magnetic tapes to the final corrected profiles and machine drawn contoured maps.

The original magnetic tapes are created by a Digi-Data tape recorder at 200 BPI and have 68 character records, each followed by an inter-record gap. These 68 characters are separated into 17 parameters of information, each of four characters. The data are recorded as shown in Figure 1. The integral count or total radioactivity is recorded every 0.5 second and all other parameters every 2.5 seconds. The integral count, average altitude of the aircraft over the 2.5 second counting period, and thorium, uranium and potassium information are recorded in binary coded decimal (BCD), in the form of a three digit number and an exponent which has a power of ten. For example, a four digit number written on the tape as 3,451 is interpreted as $345 \times 10^1 = 3,450$. The remaining eight parameters are conventional four digit numbers. The date, line number and operating code are manually recorded. The last three digits of the line parameter are normally used for the line number, the reflight of a line due to a false start being indicated by a 1 in the first digit. In the case of the operating code, the last two digits are used to indicate to the program whether the data are for a survey line (code 10) or a background over a lake (code of 30). Fiducials, recorded in the first two digits of the operating code, are manually inserted by the operator when the aircraft passes over an obvious feature, normally every 5 miles or so, for track recovery. The fiducials normally run consecutively from 1 upwards along the line and serve as a check on the direction the line was flown. Long and cross-track information from a Doppler navigational system is also recorded.

Seven Fortran programs are used in the computational procedure and are supplied on disc file or can be used as a source deck. The flow charts relating the various programs are shown in Figure 2. Complete details of all the programs may be found in Grasty (1972).

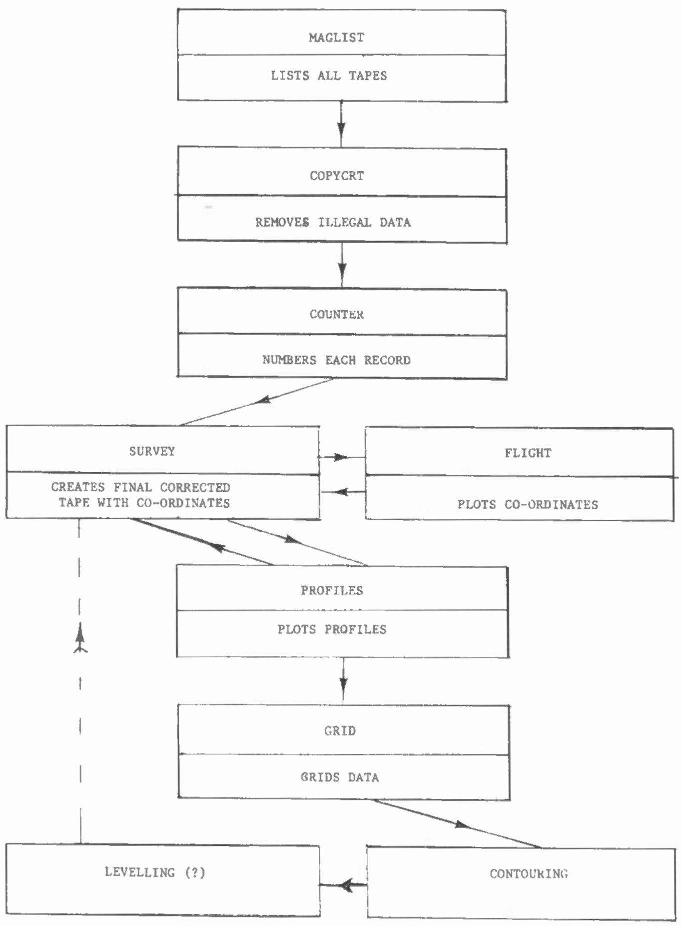


Figure 1. Data Flow Chart.

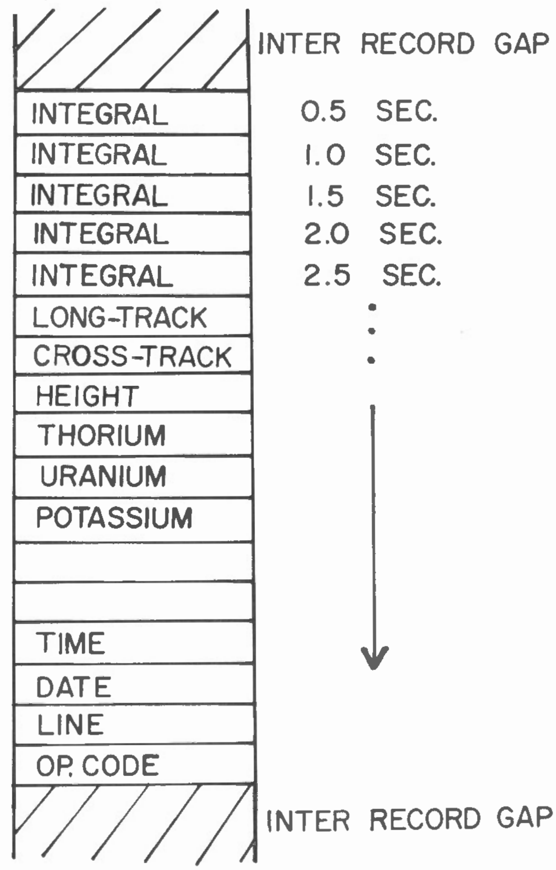


Figure 2. Layout of Magnetic Tape.

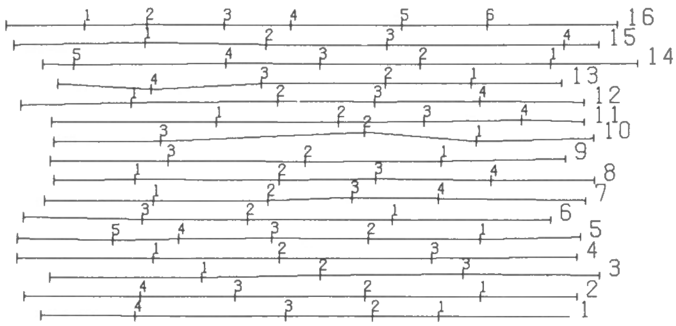


Figure 3. Sample Flight Line Map.

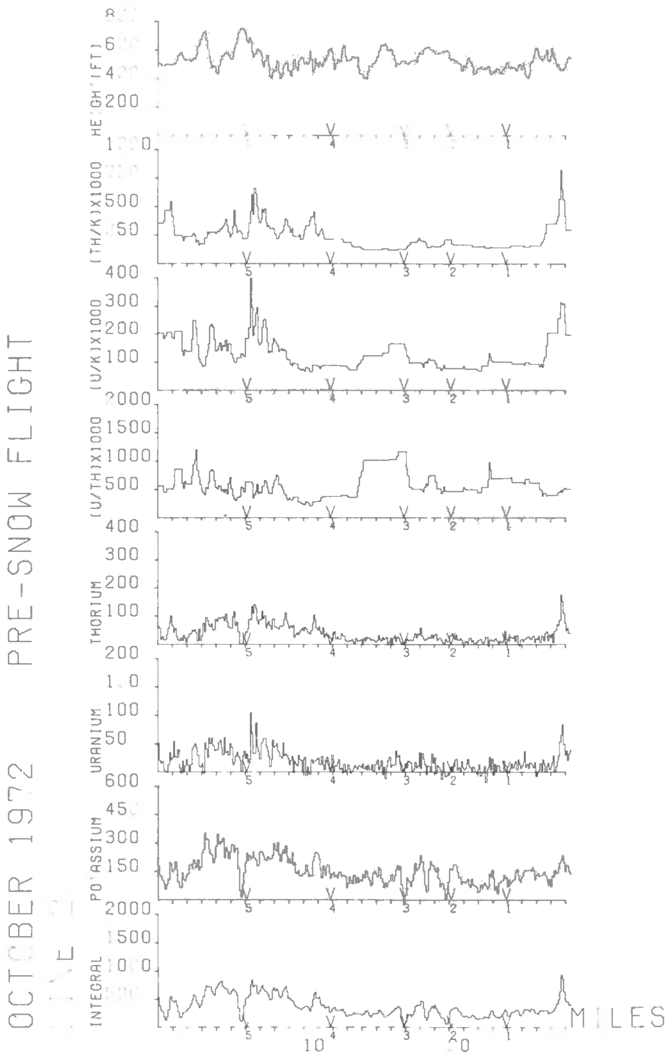


Figure 4. Sample Profile.

On completion of a survey, all raw data tapes are listed using the program MAGLIST to establish that the data were recorded satisfactorily. The occasional bad tape is copied using the copying routine available on the Systems Dimensions Limited IBM machine. All raw data tapes from any one survey are copied onto one Master tape as one job, using the program COPYCRT. Any non-integer numbers are converted to zeros and only codes of 10 and 30, i.e. background and 'on-line' information is copied. If any parity errors occur, the previous good record is copied. As a safety precaution, the original Master tape is duplicated using the program COUNTER. In this program the seventh parameter on the tape, the cross-track Doppler position, which has proved unsatisfactory, is replaced by a counter value for easy identification of each record. The counter value runs consecutively from 1 upwards and is reset to 1 after reaching the maximum four digit number of 9999, in order that the tape will be compatible with the previous one.

Program SURVEY produces a final corrected data tape with X and Y co-ordinates. The program consists of a main program, nine subroutines and one function subprogram. Co-ordinate information is read in from cards from the digitized fiducial positions. The complete data correction procedures may be found in Grasty (1972). The format on this final tape is compatible with all previous tapes. In order to check the fiducial co-ordinates on the data cards used in the SURVEY program, a computer plotted map is drawn to the same scale as the original flight line map so that the two can be checked accurately against each other. A typical flight line map is shown in Figure 3.

Figure 5. Sample Contour Map.

Computer plotted profiles of total radioactivity, potassium, uranium, thorium, the three element ratios and the aircraft altitude are shown in Figure 4. These profiles are plotted using the program PROFILE which has several input parameters for varying the smoothing of the data and the scale of the profiles.

For any automatic contouring procedure it is necessary to interpolate the data values to a regular network of grid points. In the program GRID, the grid intersections are located along the original flight lines and hence a minimum of interpolation is necessary and the map will accurately represent the profile data. A linear interpolation procedure is used. There is a facility in the program for smoothing the data over any predetermined number of data points by a running average technique. The calculated grid values are then written on an output tape in the format required by the Calcomp General Purpose Contouring Package (GPCP). A sample contour map of the final output using GPCP is shown in Figure 5.

Reference

Grasty, R. L.
1972: Airborne gamma-ray spectrometry data processing manual; Geol. Surv. Can., Open File 109.

THE AEROMAGNETIC DATA AUTOMATIC MAPPING (ADAM) SYSTEM

M. T. Holroyd,
Resource Geophysics and Geochemistry Division

Aeromagnetic Survey and Compilation Practice

The normal practice for aeromagnetic surveys is to cover the designated area with a set of equispaced parallel "traverse lines" and a further set of "control lines" flown perpendicular to the traverses. The control line spacing is usually about 5 to 10 times greater than the traverse spacing. Usually the true flight path of the aircraft is recovered, after flying, upon a photo mosaic base map of the survey region using aerial photographs taken continuously during survey flights.

The magnetic field values on a traverse and control line at their apparent intersection point are not usually identical. The discrepancy is due to flight path position errors, difference in altitude between the lines and the diurnal variation of the Earth's magnetic field. The intersection discrepancies over the whole survey are resolved by a process known as "levelling" and a preliminary contour map of the traverse data prepared. Imperfectly resolved level differences and other types of error cause "Herringbone" in the contours. Any such residual errors are traced back and resolved and the final contour maps are produced.

The manual compilation process of aeromagnetic map making mainly entails the re-arrangement and combination of large analogue data sets. It is a well established process but includes many of those human capabilities difficult to simulate efficiently with the computer, e.g. hand-eye co-ordination and immediate recognition of a deviant region within an otherwise orderly pattern. It is, however, basically a sequence of sort-merge operations interspersed with numerically definable processes and is therefore also related to well established computer practices.

The hoped for advantages of computer compilation systems are increased objectivity and accuracy and decreased time and cost. These advantages can be achieved even if the data are recorded in an analogue mode and are subsequently digitized, but are easier to achieve if the airborne data are digitally recorded.

System History

About six years ago, the Geological Survey began to acquire digitally recording high sensitivity aeromagnetic data; at first from contracted surveys, then later from its own airborne equipment.

The high sensitivity magnetometer is an instrument capable of absolute measurements to a precision of 10^{-3} gamma (10^{-8} gauss) over a range exceeding 10,000 gammas. Full exploitation of such a capability places great demands upon the recording and processing systems to be employed and it was obvious that conventional analogue recording and manual compilation techniques were inadequate. Consequently inflight digital recording was used and work began on computer software systems to automate the data compilation.

After testing alternative approaches, the nucleus of a system emerged and by early 1972 the first maps were produced by what would later be called the ADAM System.

Source Data Characteristics and Specifications

In addition to the magnetic field measurement, such data as time, radar and/or barometric altitude and along- and across-track Doppler are usually recorded in flight. Any chronologically continuous sequence of records (termed a 'line') must possess a unique 'line name'. Very long lines must be broken into 'line segments' each of less than 10,000 data point records and each labelled with the original line name. The lines or line segments are not required to be in any particular order.

Due to various reasons including cost, reliability, accuracy and load capacity of the aircraft, a navigation system capable of providing digital geographical co-ordinates in flight is not employed in the Geological Survey's system. Instead, the pilot attempts to fly straight lines as marked in a photo mosaic "Navigation strip" and a 35 mm camera takes a continuous sequence of vertically-downward photographs. The flight path is later recovered by plotting the position of these in-flight photographs on a photo mosaic map, and converted to digital form on a digitizer.

The size, shape and scale of the mosaic maps on which the F/P* was recovered, the orientation and scale of the digitization axes with respect to the geographical axis of the map, and the projection of the map (provided its projection equations are known), are all irrelevant to the system. (N.B. other considerations beyond the system requirements place practical limits on these factors.)

The mandatory conditions are that there must be some slight overlap between adjacent maps, or during the plotting of F/P points, some points must be plotted beyond the mosaic boundary in the position they would occupy on an overlap.

At least four control points of known latitude and longitude must be digitized from the map boundary and the same points must appear on and be digitized from all maps contiguous in the region of the point. The end points of a F/P line segment on a map must extend into the overlap region unless such an end point is a termination within the map or on a survey boundary.

Within a map the F/P points may be digitized in any order provided that any sequence of points possesses the same line name as its corresponding I/F* data.

* (I/F = in-flight, F/P = flight path.)

For the logical process of co-ordinate transformation, the only strict requirements are that at least two latitude-longitude points appear somewhere within a map of known projection; overlaps between adjacent maps are not required. The system requirements as specified in excess of this basic minimum are included to enable the system to detect and rectify errors of displacement or distortion present in the mosaic. If such errors are of greater magnitude than the limits of applicability of the correction algorithm, the system indicates the fact so that alternative means can be applied.

System Structure

The system is subdivided into separate program modules so as to provide greater flexibility of use but also to minimize core requirements without significantly increasing running time. Because of the massive data sets involved (the system is designed to deal with up to 500,000 line miles of data as a single throughput = (approx.) 50 million words of data) and the nature of the processes applied, data are passed through a process to temporary disc residence then re-passed through the next process etc. The natural hiatus between processes allows the use of interchangeable program modules to vary the nature of the processes applied according to requirements.

Program modules within the structure are labelled by a 4-character purpose identifier, a one digit "stack sequence" number and a one character "version code". Many programs are used in concatenate groups (theoretically, for error free data, the whole system of approximately 10,000 FORTRAN statements could be concatenated into a single deck and run as one job!!). The "stack sequence" number indicates the relative position of a program in a stacked group. All programs with the same purpose identifier and stack sequence may be interchanged without modification. For example, the program group SGRD interpolates a square grid from the final levelled magnetic data. Its basic configuration contains 3 program modules: - SGRD1A, SGRD2A, SGRD3A. If SGRD1A is replaced by SGRD1B the input magnetic data will undergo a filtering process before gridding (i.e. to produce a vertical derivative map rather than a total field map). If SGRD2A is replaced by SGRD2B the data along the grid lines can also be filtered. If SGRD3A is left out the final output will be a total survey grid; if included the grid will be subdivided.

Some general purpose utility programs (such as MROT1A for matrix rotation) occur in several program groups, e.g. the actual program stack for filtering and gridding aeromagnetic line data with no filtering along the grid lines and with subdivided grid output would read:

SGRD1B
MROT1A
SGRD2A
SGRD3A

The user's task simply involves attaching the permanent files containing the required programs, specifying in their execution in the correct order and supplying the control parameter cards required by the programs used.

Input/Output Options

The initial data sets and all subsequent intermediate data are output in binary mode by unformatted FORTRAN write statements. The choice of data code for the final output is a user option.

As all key stages in the system involve intermediate disc storage of the data, the user may:

i) allow the disc resident data to be a transient product which will be overwritten by a more advanced data set by assigning the same file name to the advanced data set.

ii) between the execution of consecutive program modules, insert commands to dump the disc data to tape.

iii) define a file so that the data is written directly to tape rather than disc.

iv) with fiendishly cunning control language commands, execute a job containing a stack of program modules in such a way that if any program aborts the immediately preceding data sets are dumped to tape thus allowing acceptable stages to be by-passed when restarting the job after the trouble has been traced.

v) employ any combination of the above to achieve user requirements.

System Operations

These are illustrated by Figure 1.

Pre-processing F/P and I/F Data

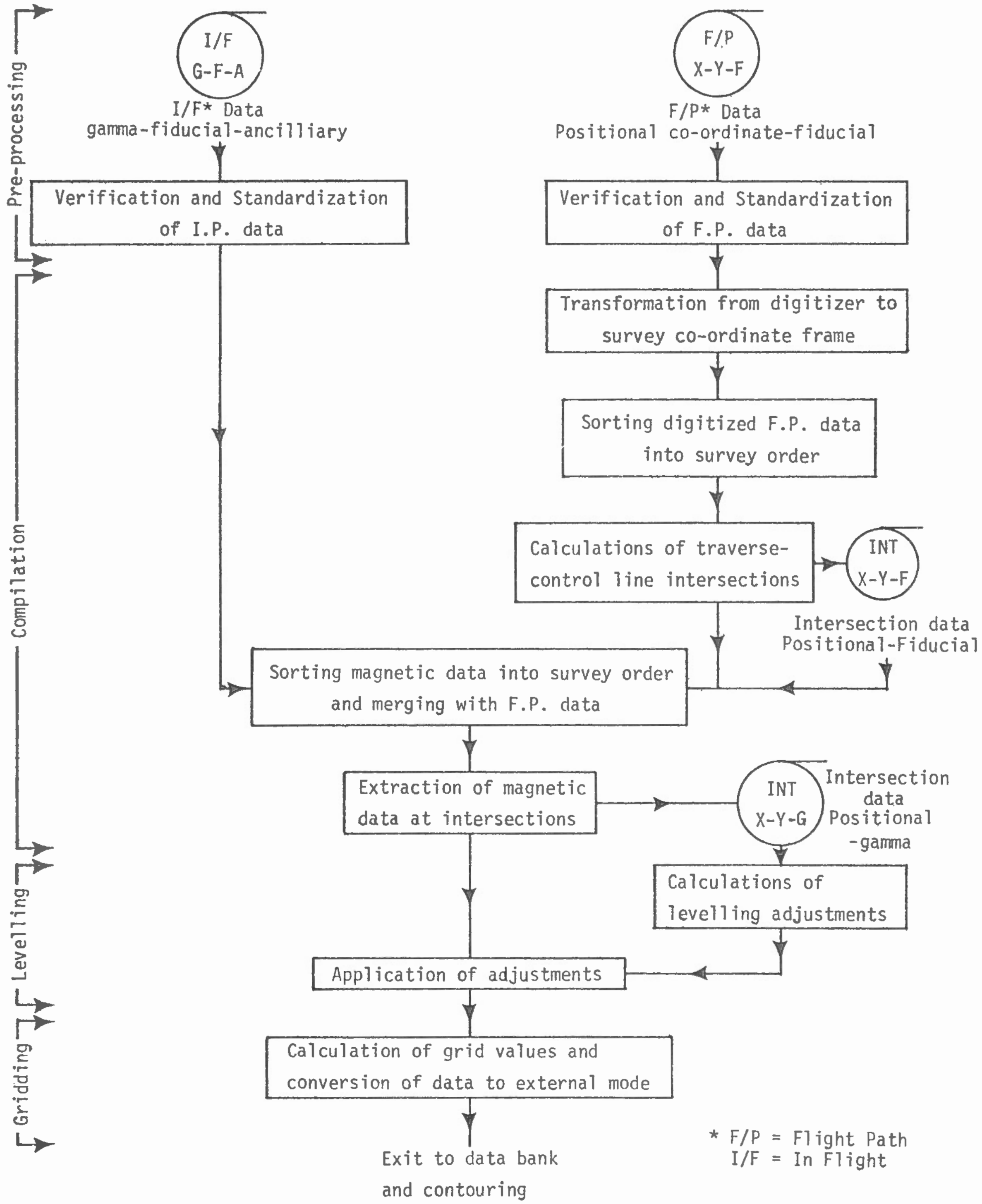
The coding and order of the source data sets are both tailored to the requirements of the acquisition environment. The first task of the system is to convert the data to a form more suitable to the processing environment, and to detect and remove data errors. This phase is named pre-processing and is regarded as external to the system proper as the program content is largely dependent on the in-flight recording and digitizer hardware and thus not generally applicable. The output from this phase is in the form of large buffers of binary data but with the general order of the data unchanged from the order of acquisition.

Compilation

Flight path data are digitized from a photo mosaic base map or planimetric map. Many such maps are usually required to cover a survey area and as the survey is not flown by map-sheet the F/P data are therefore acquired in a radically different order from the complementary I/F data. The system allows that the order in which the map-sheets are digitized, the orientation and positioning of any map with respect to the digitizer axes, the digitizer units and the order of digitization of the flight line segments within a map-sheet, may all be arbitrary. The user defines a survey origin by its latitude and longitude, and survey axes by defining the lat.-long. of a second point through which the Y axis will pass. The X axis is taken to be perpendicular to the thus defined Y axis at the origin point.

The first stage of compilation is to transform the co-ordinates on each map in turn from their existing arbitrary units and reference frame to standard units and defined survey frame. During this process the digitized points of known latitude and longitude are used to detect the presence of distortions in the mosaic.

Figure 1. General Flow Chart of the ADAM System.



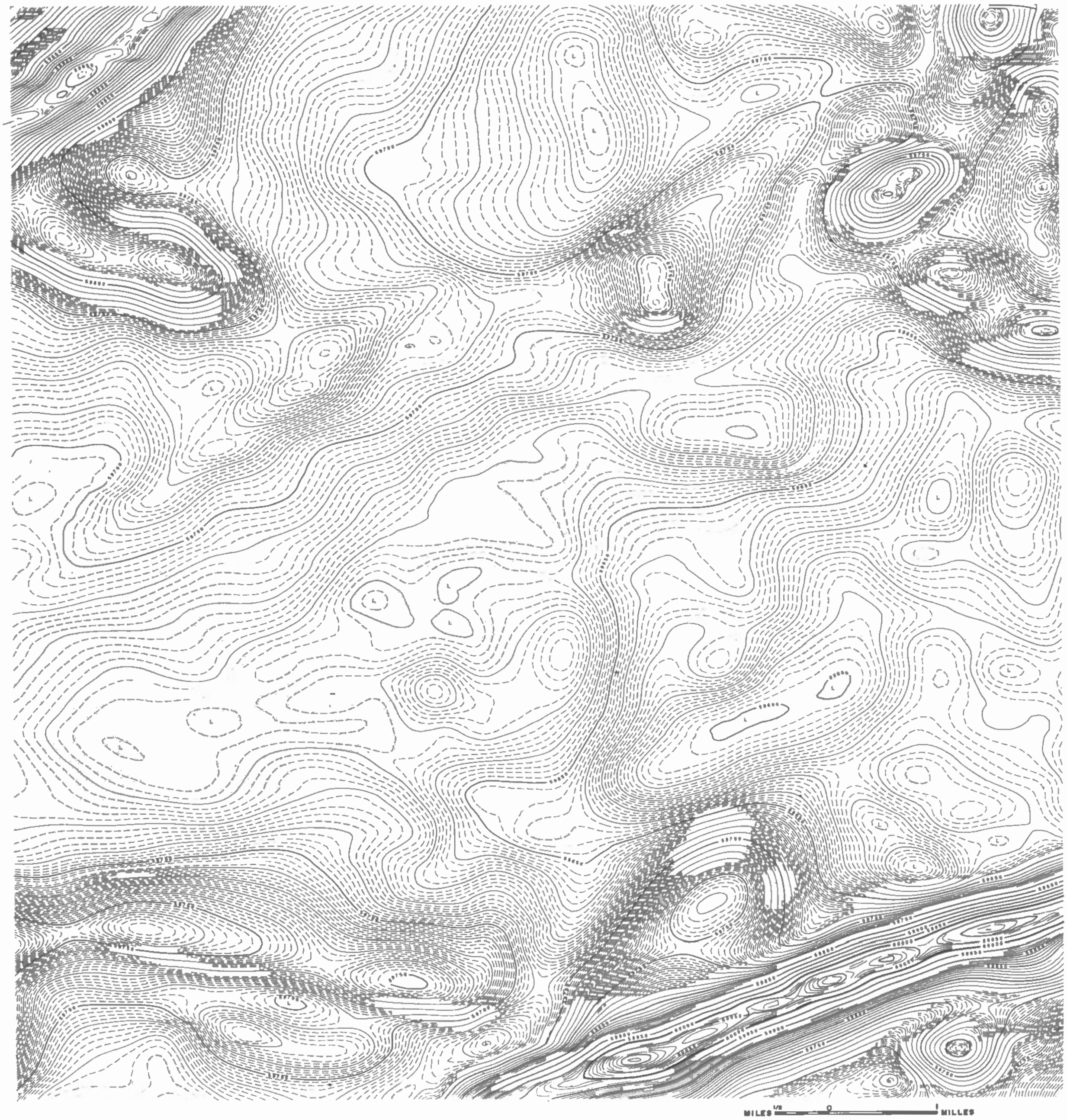


Figure 2. High resolution aeromagnetic map compiled by the ADAM system.

Differences in scale between the X and Y axial directions and a change in X and/or Y scale across the map are assessed and magnitude of such distortions listed. X and Y scale factors are then calculated so as to eliminate or reduce distortion.

The second stage is to sort the line segments from map-order to total survey order and merge the separate segments from adjacent maps into continuous lines. At this time the overlap of segments into adjacent maps is used to detect erroneous shifts between sheets due to the mispositioning of lat.-long. points and if possible to resolve the shifts.

The data are now in good order and all positioned co-ordinates refer to the same reference frame.

The third stage is to calculate all possible intersections between traverses and control lines and to establish a separate data set containing the positional co-ordinates and fiducials of the lines at the intersection points.

The fourth stage sees the re-entrance of the I/F data which is sorted into survey order and merged with the F/P data. During this stage the intersection data set is also briefly re-introduced and the magnetic data at the intersection added to it.

Compilation is now complete and all data exists in the desired order and form for Phase 3 of the system.

Levelling

The levelling process is similar to those required in gravity surveys and topographic mapping. Differences along measurement "legs" forming a closed loop have a non-zero sum. The problem is to adjust each difference so as to distribute the error evenly and bring about a zero sum. Many solution methods exist. The one chosen for the ADAM system depends on flying double, overlapping control lines but all require an intersection data set of some form and alternative program modules with different solution methods could be substituted at this point. Let it suffice to say that from the levelling program emerges a set of adjustments to be applied to the data at control line traverse intersection points.

The adjustment to be applied to a data value not falling on an intersection must be interpolated from the two intersections which bracket it. No extrapolation is permitted. A choice of interpolation functions is allowed, once again by the substitution of program modules.

Gridding

Many approaches exist to the interpolation of a regular grid over irregularly spaced data. For the ADAM system one well suited to the nature of aeromegnetic data was chosen. A set of closely spaced parallel lines are run approximately perpendicularly to the traverse direction. The values at the points where these TDS (traverse data sections) intersect the traverses are extracted and values interpolated along the TDS by a cubic spline interpolation algorithm originally designed by B.K. Bhattacharyya. The grid lattice major axes are parallel to the survey axes, but as the traverse direction may not be parallel to one or other of the survey axes, and the TDS should be close to perpendicular to the traverses the TDS need not be parallel to one of the major axes of the grid. Eight different directions are permitted for the TDS with respect to the survey X axis. They are (for those familiar with crystalographic notation), 1: 0, 2: 1, 1: 1, 1: 2, 0: 1, -1: 2, -1: 1, -2: 1. Thus a gridding direction can be chosen to lie within approximately 12 degrees from the perpendicular to the traverse. The grid size was chosen so that contours produced by short linear interpolations across the cells would be satisfactorily smooth but still economic to produce. A grid cell size of approximately 0.1 inch is a fair compromise between these two opposing objectives. An exact figure of 10.16 grid cells/inch (4 grid cells/cm) was chosen to fit in with metric publication scales. At the usual scale of 1: 25,000 this results in a ground distance of 62.5 metres/cell. Hence super grids can be extracted at 125, 250, 500M, 1 km, etc. and contoured still at 4/cm to produce maps at scales of 1: 50,000, etc.

Conversion to External Mode

Due to the redundancy of the 60 bit CDC (cm = central memory) word with regard to the precision required in many processing phases, "packing" routines were designed into the system. In various stages of the work 2, 3 and 4 binary data values may be packed into one cm word. This is a user option, obtainable by program module substitution. The advantage of word packing is a significant reduction in central core memory requirement at the expense of some increase in central processor time. Such packed words mean little, however, if on transfer from the system the recipient software is unaware of the coding. Hence program modules exist to decode such data into more common form.

Contouring

The existence of many excellent contouring programs outside the Departmental Computing Science Centre and (at the time of development of the system) the absence of a flat bed plotter, led to the decision to "go outside" for final contouring of maps for publication. The attached example (produced by Survair Ltd. of Ottawa) is a special map originally drawn at a scale of 1: 25,000. The maps so far published (all contoured by Dataplotting Services of Toronto) were contoured at 5 gamma intervals to prevent overcrowding in the many high gradient areas. To show the fine detail in a low gradient area of the Abitibi survey, the area was contoured at both 1 and 2 gamma intervals. Figure 2 is the 2 gamma version.

Data Bank

In our opinion, the high resolution and precision of the original data and the objectivity and accuracy of the processing system have relegated the contour map to the status of a by-product and index rather than the end product of digitally recorded and processed high resolution aeromagnetic survey. Far more information than can be displayed on a contour map, exists in the digital data bank. Furthermore, digital data allows re-processing into any of the various "derived" maps that are of specific use to the interpreter of aeromagnetic data. A system is being developed whereby a potential user may acquire the digital data of any surveyed region in a form suited to his processing software, or, if so required, a user can obtain "one-of" specially processed maps of a region of interest. Such special processes include band, high and low pass filtering to separate the magnetic effects of different depth horizons, downward continuation or vertical derivative maps to emphasize low amplitude features, and special contour effects or grey-shade mapping to delineate texture.

Conclusion

The results of the application of the ADAM system have demonstrated that the potential advantages offered by such an automatic compilation system are attainable. Distinct increases in objectivity and accuracy have emerged along with definite decreases in time and cost. Caution is, however, called for. Once manually prepared data are exposed to the pedantic attention of the computer, hitherto undetected or disregarded flaws become critically apparent. The standards of quality of base mosaics for flight path recovery have regularly been found inadequate. The high potential of high resolution aeromagnetic data should not be degraded by poor quality ancilliary data. The most fitting conclusion is to hope for the advent of an in-flight, digitally recording navigation system whose cost, reliability, weight and precision are concomitant with the same factors within the aeromagnetic measuring-recording system.

19. COMPUTER USE IN THE SEISMIC SECTION

A. Overton,
Resource Geophysics and Geochemistry Division

The Seismic Section of the Resource Geophysics and Geochemistry Division uses matrix inversion routines for high order linear simultaneous equation solutions, standard statistical evaluations (standard error, significance tests, least squares curve fitting, and correlation tests), standard reduction of seismic data using refraction, reflection and diffraction formulae, and geodetic computations. Some use has been made of a portable remote terminal in the field office in a time-sharing system. Before the advent of the electronic computer these problems were solved using mechanical devices, slide rules, desk calculators, nomograms, mathematical tables, and graphical methods. We now use the electronic computer when applicable.

20. COMPUTER-BASED ACTIVITIES IN PALYNOLOGY

W. W. Brideaux,
Institute of Sedimentary and Petroleum Geology, Calgary

The application of computer-based techniques for palynologic analyses in the Geological Survey is still in a formative stage, although such techniques are used widely in the petroleum industry. Computer files and allied retrieval programs are being developed along two main lines: those using data primarily generated from outside the Survey, and those using data generated primarily from Survey files and activities. These of necessity overlap to some degree.

In the first case, the Geological Survey will acquire palynologic data load forms for some 5,300 published papers by participation in the Kremp Palynologic Computer Retrieval Project, initiated by Dr. Gerhard Kremp of the University of Arizona, and supported by seven major petroleum firms. Further updating of this data base will include approximately 1,500 more papers. The file will then contain abstracted pre-Quaternary spore-pollen-dinoflagellate data from about 90 per cent of the western world literature, and 15 per cent of the Sino-Russian literature. Retrieval programs must now be developed to enable users to search the file for taxonomic, biostratigraphic, paleogeographic, and bibliographic purposes.

In the second case, Geological Survey of Canada palynologists have been actively developing the framework for the recording and retrieval of palynologic data, both published and unpublished, generated by Survey personnel. Standard data forms, standard methods of recording taxa, and final aims for the use of the file have been worked out at a series of annual meetings initiated at Calgary in October, 1971. The system envisaged is flexible, allowing for both published and unpublished information to be entered and, if desired, stored permanently. The retrieval programs (to be developed) are intended to provide a means of ordering data in range charts, ratio plots, percentage plots, printouts of assemblages, and in other useful statistical arrays (recurrent groups, cluster analysis).

Of much value, but less general in application is the recently devised DINOFILE, which lists all dinoflagellate-acritarch taxa described in publications on assemblages from western, northern and offshore eastern Canada, and includes authors of taxa, date of publication, locality and geologic age ranges for each taxon. Similar future files might include megaspore taxa files and spore-pollen files. DINOFILE will be expanded in future to include all North American publications and, in addition, verified but unpublished data from Geological Survey files.

Paul R. Gunther,
Institute of Sedimentary and Petroleum Geology, Calgary

Figure 1

```
READY
PTAPE

READY
8Ø DATA 65,63,72,45,67,63,57,54,66,44,64,41,51,54,63,58,55,63,63
81 DATA 48,63,7Ø,63,53,54
17Ø DATA 66,54,63,49,65,7Ø,69,68,57,66,62,51,57,67,53,55,49,69
171 DATA 49,55,67,68,73,61,63
RUN
THE SUM OF COLUMN ONE IS       14.59
THE SUM OF COLUMN TWO IS       15.26
THE AVERAGE OF COLUMN ONE IS   .5836
THE AVERAGE OF COLUMN TWO IS   .61Ø4
THE GRAND TOTAL IS             29.85
A COMPAR. OF SUBAVE. IS       -2.68       /1ØØ
THE AVE. MAX. REFLECTANCE IS   .597
THE STANDARD DEVIATION IS      7.9Ø823    /1ØØ
THE VOLATILE MATTER IS 42   %

THE V-TYPES ARE
            V 4    = 7
            V 5    = 15
            V 6    = 24
            V 7    = 4
THE HISTOGRAM FOLLOWS:
                         Ø123456789Ø123456789Ø123456789Ø
. 4Ø   -. 44             1 x
. 45   -. 49             1    x
. 5Ø   -. 54             1       x
. 55   -. 59             1      x
. 6Ø   -. 64             1           x
. 65   -. 69             1           x
. 7Ø   -. 74             1   x
```

In the first application (Fig. 1) fifty observations are statistically analyzed to give an average reflectance of a coal sample. Standard deviation, volatile matter, V-types and a histogram are provided to numerically and visually portray the data. In the second application (Fig. 2) coal constituents and vitrinite rank variations are combined by a percentage loading technique to predict the resultant coke stability.

Two facets of coal petrology research are greatly facilitated by computer applications. These are reflectance studies and coke stability calculations. Reflectance of vitrinite (the main constituent of coal) is used to determine the rank of a coal which, in turn, can be used to evaluate the degree of organic metamorphism and the burial history of a sedimentary basin. Coke stability calculations are used to predict whether or not a coal can be used to produce a satisfactory metallurgical coke.

Figure 2

```
READY
PTAPE

READY
RUN
THE V-TYPES ARE
                    V 9    = 1          V 1Ø   = 6
                    V 11   = 27         V 12   = 5Ø
                    V 13   = 15         V 14   = 1
                    V 15   = Ø          V 16   = Ø
TOTAL INERTS % 44.7Ø2
LOOK UP TABLES FOR INPUT
?3.55,3.7,4.55,5,4.31,2.6,Ø,Ø
THE TOTAL REACTIVES % 55.2872
THE CORRECTION ERRORS ARE 1ØØ   99.9893
THE OPTIMUM INERT INDEX IS 17.9418
THE BALANCE INDEX IS 2.49151
THE OPTIMUM STRENGTH IS 257.556
THE STRENGTH INDEX IS 4.6585

READY
```

COMPUTER APPLICATIONS IN THE STUDY OF A SEDIMENTARY BASIN

Andrew D. Miall,
Institute of Sedimentary and Petroleum Geology, Calgary

The Institute of Sedimentary and Petroleum Geology currently maintains a CRT terminal linked to a Univac 1106 computer operated by Digitech Ltd. The author has developed a number of statistical and data storage/retrieval programs for use on this system, to aid in sedimentological studies of a suite of Cretaceous and Tertiary clastic rocks on Banks Island, Northwest Territories. All the programs are designed for use by untrained operators in that instructions for data input and requests for program options are printed at the terminal during program execution.

The advantage of a sophisticated terminal system is its great flexibility. Programs may be designed, entered into the system and implemented with great rapidity to suit any immediate requirement. The main computer applications in use at the present time are listed below, with a brief accompanying discussion.

1. Numerous measurements are being made in the field on sedimentary structures. Parameters studied include structure type and size (thickness), indicated current direction and, where appropriate, the dip of foreset beds within the structure. These data are stored on disc together with numerical codes to identify location, stratigraphic age and level within a measured stratigraphic section. This data set is amenable to processing in several different ways, as discussed below in paragraphs 2, 3 and 6.

2. A program has been designed to calculate mean current azimuth using the vector method of Curray (1956). Output also includes vector magnitude, Rayleigh significance test and a listing of the readings falling in 20 degree azimuth classes, by number and by per cent of total. An option which is planned for this program but is yet to be developed, is a plotting subroutine to draw current rose diagrams. Another possible option (not developed owing to a lack of present need) would be to build in a routine to calculate the necessary corrections for structural dip.

This program is designed so that any set of data may be retrieved for processing from the paleocurrent data file by using retrieval options for stratigraphic age, location and sedimentary structure type. Detailed studies of local and regional paleocurrent variation are thus possible.

3. Several simple programs are being used to study vertical sequences of sedimentary structures in stratigraphic sections of some fluvial deposits. Calculations of moving average azimuth, crossbed set thickness and directional variance serve the purpose of clarifying vertical trends, leading to interpretations of varying stream sinuosity, stream rejuvenations, and migrating channel patterns.

4. Analysis for memory effects in stratigraphic successions is being undertaken using Markov chain methods. Successions with strong memory effects are said to be cyclic (e.g. coal-measure cyclothems), but the cyclicity may be far from obvious, and a statistical approach is often useful in pointing out subtle facies inter-relationships that would otherwise not be noticed.

The technique is well described in the literature, for example by Harbaugh and Bonham Carter (1970, chap. 4) and has been successfully applied by the author to stratigraphic successions formed under several different alluvial environments (Miall, 1973). The author's Markov program is at present designed to test only single-dependency relationships, which are the easiest to interpret. The data consist of coded stratigraphic successions in which lithologies are classified into a limited number of categories (four to six is ideal) and each category is assigned an arbitrary code number. Strings of these numbers with accompanying values for bed thickness are then used as input to the program. Some preliminary geological results from the present studies, without statistical details, are discussed by Miall (1974). They have allowed the definition of a typical coarsening-upward clastic cycle that is widely developed in a succession of Tertiary deltaic deposits in the Banks Island project area.

5. An option in the Markov program enables similar processing to be carried out on interval and ratio type data such as digitized mechanical logs. The data are transformed into a limited number of states and analyzed in the same manner as the surface stratigraphic sections. This routine (plus a second program testing for double dependency relationships) has been applied to logs of wells penetrating the deltaic deposits, in an attempt to trace the same type of clastic cycle into the subsurface.

6. Another application of the Markov technique is in a study of the vertical changes in paleocurrent direction from structure to structure. The question is posed: how random are the changes in orientation between successive, superimposed current structures? A perfect Markovian memory would indicate a complete sedimentary record of a slowly evolving current environment, such as a migrating river meander. Statistical randomness would indicate imperfections in the sedimentary record, such as erosional breaks. The results of such experiments will, it is hoped, provide information on the preservation potential of sedimentary structures and on the way in which fluvial deposits accumulate.

7. Fourier analysis is a powerful tool in the detection of cyclic character in strings of numerical data. Geological applications are discussed by Preston and Henderson (1964) and Young (1969). In the present project a program developed by Digitech Ltd. has been used to study digitized logs through the Tertiary deltaic successions. Amplitude spectra reveal a persistent cycle similar in thickness to the clastic cycles that are described in the surface sections. In detail, however, the thickness varies across the project area.

The variation closely parallels the structural configuration of the sedimentary basin, such that the cycles are thicker in deeper portions of the basin. This demonstrates that the basin is a true depositional basin, and has not been appreciably modified by subsequent tectonic activity.

These are the main computer-aided projects under way at the present time. A number of other basic utility programs are in use, such as cluster analysis, linear regression, calculations of arithmetic and moment measures of the normal distribution, calculations of lithologic percentages and ratios in sedimentary successions. None of these programs are novel; their importance lies in their ready availability for the rapid testing of ideas during the normal course of stratigraphic and sedimentological studies.

References

Curray, J.R.
1956: The analysis of two-dimensional orientation data; J. Geol., v. 64, p. 117-131.

Harbaugh, J.W., and Bonham, Carter G.
1970: Computer simulation in geology; Wiley Interscience, 575 p.

Miall, A.D.
1973: Markov chain analysis applied to an ancient alluvial plain succession; Sedimentology, v. 20, p. 347-364.

1974: Bedrock geology of Banks Island, District of Franklin; *in* Report of Activities, April to October 1973, Geol. Surv. Can., Paper 74-1, Part A, p. 336-342.

Preston, F.W., and Henderson, J.H.
1964: Fourier series characterization of cyclic sediments for stratigraphic correlation; State Geol. Surv. Kansas, Bull. 169, p. 415-425.

Young, F.G.
1969: Sedimentary cycles and facies in the correlation and interpretation of Lower Cambrian rocks, east-central British Columbia; Unpubl. Ph.D. thesis, McGill Univ.

23. CURATORIAL RETRIEVAL SYSTEM, INSTITUTE OF SEDIMENTARY AND PETROLEUM GEOLOGY

L.L. Price,
Institute of Sedimentary and Petroleum Geology, Calgary

The purpose of a computer record in this curatorial system is first, to maintain a reliable compact storage system for geological field specimens, corresponding processed material and laboratory reports, so that this physical material can be retrieved efficiently; secondly, to coordinate effort in laboratory procedures and prevent duplication, and, thirdly, to access selectively all of the material pertinent to a specific study.

The system will perform the following functions:

Identify field specimens initially by

1. Collector and year.
2. Geographic location.
3. Purpose of collection (e.g. paleontology).
4. Source (outcrop, talus).
5. Probable age.
6. Formation.

Summarize the types of laboratory analyses and resulting material, residues, slides, photos, etc.

Record permanent and temporary use locations of specimens or laboratory material.

Retrieval may be based on:

- Geographic location expressed in Lat.-Long., UTM Grid, National Topographic System, or Township and Range.
- Geological age.
- Formation.
- Purpose of collection.
- Type of laboratory analysis.
- Collector and year.
- GSC specimen number.

24. QUANTITATIVE ASSESSMENT OF RESOURCE POTENTIAL USING MONTE CARLO SIMULATION

K.J. Roy,
Institute of Sedimentary and Petroleum Geology, Calgary

Since 1972, the Geological Survey has been involved in quantitative assessment of Canada's hydrocarbon potential. During this period, applications of Monte Carlo simulation techniques have been developed and are being used in the assessment. Because of the inherent variability in the data and because of the uncertainty associated with many of the parameters, the ability to define variables in a probabilistic sense is essential. The output that is required is the frequency distribution of estimates of potential.

The basis of the method is the derivation of a functional relationship between a series of variables and constants and the resource potential.

Prospect Potential = Pore Volume *% fill *% oil * etc.

The parameters are described by cumulative frequency distributions derived from "hard" data or from estimates by informed sources or both. These curves are digitized and entered into the computer. Programs have been written to manipulate the input information in a Monte Carlo simulation which, in effect, cycles through the "potential" equation with random collections of values for the input variables. This procedure generates a frequency distribution of "potential" equation solutions which describes, in probabilistic terms, the variability of the outcome resulting from variability of input data.

The definition of distributions associated with input parameters has been approached by consideration of joint, conditional and marginal probability. (The probability that the parameter is non-zero and has a value 'Y', the probability that the parameter has a value 'Y' given that the value is non-zero, and the probability that the parameter is non-zero.)

$$P(X > O \text{ and } = Y) = P(X = Y \mid X > O) * P(X > O)$$

The P(A) term is considered the risk term. The risks (the probability of occurrence of non-zero values) associated with the various factors controlling hydrocarbon occurrence at present are assigned subjectively (on the basis of informed judgment) as single numbers. These probability values are then multiplied to arrive at the total risk. This risk is applied in the system as a single number reduction of the probability of occurrence of any given value of potential estimate.

The manipulation procedure is divisable into two parts: (1) the "system" - the programs; and (2) the "logic" - the equations of potential. The system is relatively standardized but the "logic" can be of any nature. Thus, depending on the data available, the geometric configuration considered, and the nature of the approach, a large variety of logic statements has been used. This capacity for variety in approach is also useful to check results.

The system is being used via terminal facilities at the I.S.P.G. in Calgary and has been used from Halifax during meetings there. On the average, five curves are drawn for each problem. In general, 10,000 solutions are made for each problem and, in the assessment, over 200 problems were solved. A major time factor is the input of digitized data from the curves. Automatic digitizing, however, has reduced the time considerably.

The programming and the philosophy are still in the prototype stage. Considerable change, improvements and refinements are taking place. The "system" can be applied to any problem in which empirical solutions to equations are required and the variables can be described by frequency distributions.

Selected Bibliography

Grayson, C.J., Jr.
1960: Decisions under uncertainty drilling decisions by oil and gas operators; Harvard Business School, Division of Research, 402 p.

Hertz, D.B.
1964: Risk analysis in capital investment; Harvard Business Review, XLII (Jan.-Feb. 1964) p. 145-156.

Mihram, G.A.
1972: Simulation statistical foundations and methodology; Academic Press, New York, 526 p.

Mize, J.H., and Cox, J.G.
1968: Essentials of simulation; Prentice Hall Inc., Englewood Cliff, New Jersey, 234 p.

Smith, M.B.
1968: Estimate reserves by using computer simulation method; Oil and Gas J., Mar. 11, 1968, p. 81-84.

25. NATURAL RESOURCE MAP PREPARATION AND POTENTIAL FIELD DATA INTERPRETATION

R. T. Haworth,
Atlantic Geoscience Centre, Dartmouth

The Canadian Hydrographic Service has been collecting bathymetry, gravity and magnetic field data on routine surveys of the Atlantic Canada continental margin for the past ten years. During that time approximately 180,000 km of track have been surveyed at an average track spacing of 5 km. Bathymetry data is used by the Hydrographic Service in the preparation of bathymetric and geomorphic maps. The geophysical data are used by the Atlantic Geoscience Centre in its regional geology program. The data on earlier cruises were prepared by hand for publication at a scale of 1: 250,000 in the Natural Resource Map series. The quantity of data collected and the desire for consistency in its presentation led to a request being made to industry for examples of what compilation by computer might accomplish. As a result, all potential field data now undergo only preliminary editing before being submitted to a contractor. The data are then compiled, edited and corrected on the basis of comparisons between data at line intersections, and then computer contoured to produce draft maps. Only minor changes (such as the insertion or deletion of contour values) are made before the map is photographically superimposed on a bathymetry background for publication.

The gravity and magnetic field data collected as part of the program described above are interpreted by means of a set of standard computer programs. In certain situations, optimization techniques may be used to predict the geometry of bodies having specified densities or magnetic susceptibilities which will produce the observed potential field anomalies. Where this is the case, the amount of computer-geophysicist interaction is relatively small. However, composite features may only be interpreted on the basis of postulating the properties of the causative bodies and calculating the resultant anomaly. Any difference between the observed and the calculated anomalies must be resolved by changing the geometry or the properties of the bodies. This interactive process can be time consuming on a batch processing computer - one job, one answer. With an interactive graphics terminal on line to the computer, the results may be displayed quickly and changes to the model may be deduced and introduced quickly also via the terminal. Experiments with a mini-computer having a scope display have demonstrated the usefulness of such a system. Investigations are now underway towards developing a system on line to a larger time-sharing machine so that advantage may be taken of the larger memory and more rapid access to the input data.

26. REAL TIME SEISMIC SIGNAL PROCESSING

D. Heffler,
Atlantic Geoscience Centre, Dartmouth

A small computer has been added to the seismic reflection system in use at the Atlantic Geoscience Centre. Development of programs to perform signal processing in real time has been taking place for two years and the system should be operational in the summer of 1974.

The computer used is a 16K, PDP-11/20 with Dec Tape and extensive analogue to digital and digital to analogue converters. A clock, hardware arithmetic unit and CRT display are also used. All programming has been done in assembly language for maximum speed and efficient core usage.

The processing program, called The Core Resident Monitor (CRM), controls the execution of an extensive set of processing sub-routines. It also provides operator interaction with all the important processing parameters, such as sample rate, output gains or signal length. An operator or watchkeeper with no programming knowledge and a brief introduction to the CRM can create a sequence of the processing sub-routines and control its execution.

The program is designed to process deep seismic reflection data (over 400 m water depth) from a single channel hydrophone array. The most important processing is dereverberation or elimination of multiple reverberations in the water column. Enhancement of the reflections from deep layers is also possible through vertical stacking, or summing successive shots.

The dereverberation scheme adopted is a two-point Backus filter, a very simple method which has not proved effective in the past. However, a new correction for the sound source receiver horizontal separation has dramatically improved the multiple reduction. Also, by allowing the reflection coefficient to change along a single shot, further improvement has been achieved. Tests on several pre-recorded sections show dereverberation as adequate as many more computationally involved schemes.

Effective display of data is considered to be an important aspect of signal processing. Several subroutines are included to normalize, filter, and adjust the gain of the signal. The signals at any stage of the processing may be displayed on a storage CRT display to monitor progress of the processing. The CRT may be used to simulate a variable area wiggly trace seismic recorder. Also, signals may be output to a seismic recorder or stored on analogue or digital tape.

The system should be completed, documented and ready for extensive field testing on the cruise of C.S.S. HUDSON to Baffin Bay in July of 1974.

27. ATLANTIC GEOSCIENCE CENTRE MARINE GEOSCIENCE DATA STORAGE AND RETRIEVAL SYSTEM

K. G. Shih,
Atlantic Geoscience Centre, Dartmouth

Introduction

The Atlantic Geoscience Centre data acquisition system was developed for processing the marine geoscience data on the shipboard computer at sea and on the shorebase computer at the Institute. The computer-based data storage and retrieval system was also developed for rapidly handling the considerable amount of data collected on the Bedford Institute of Oceanography ships (Fig. 1).

Shipboard Computer-Based Data Processing System

Bathymetric, gravity and magnetic data collected on the BIO ships have been punched since 1965 on paper tape using the automatic data logging system (BIODAL). Magnetic tape is now being used to replace the paper tape on the survey chartered ship, Minna. In order to check the quality of the paper tape data and to analyze the data at sea, a shipboard computer is required to investigate the errors arising from navigation or instrumental faults and process the data. The processed data not only provide rapid quality control of the data but also provide necessary information for controlling survey parameters, such as line spacing or area extent.

To increase the efficiency of data processing, 16K HP-2100 computers with magnetic tape transport units have replaced the 4K or 8K PDP-8 computers on the BIO ships since 1973. The 16K HP-2100 computers are equipped with two magnetic transport units, a line printer, a high speed paper tape reader and punch, a teletype, a fixed head disc platter, and a removable disc platter and operated with a moving head disc operation system (DOSM). The use of the HP-2100 computer with magnetic tape facility for geophysical data processing has demonstrated that the data collected at sea can be quickly and completely processed onboard ship.

Shipboard computer programs have been developed for quality control, data compilation, data reduction and graphic presentation. The data reduction methods are documented in BIO report (Shih, 1973) and the instruction manual for operating the system is also available in the AGC marine geoscience data section. A Calcomp plotter interfaced with the shipboard computer can provide a quick visual display of the data in various forms, such as profiles, track plot, a posting of the numeric data, etc. Calcomp output in various scales is convenient for matching to available charts, maps or records.

This year the shipboard computing facilities will be used to process the geological data collected on boardship so that it can be immediately transferred to the shore-based data storage system when the ship completes the cruise. This will include the data coding, data digitizing and data compilation on the HP-2100 computer.

Marine Geophysical Data Storage and Retrieval System

To provide facilities for storage and retrieval of the considerable amount of the marine geophysical data which has been and is being collected at sea required a computer-oriented data file. The marine geophysical data storage and retrieval system has been developed and used since 1968 around the CDC-3150 computer (Ross et al., 1973). The 32K CDC-3150 computer is equipped with three magnetic transport units and six disc drives and operated with a mass storage operating system. In the data system, magnetic tape has been used as the basic storage medium.

Since the shipboard data processing and data compilation can provide excellent control of the reliability of the navigation data and the raw data collected at sea, the data files for a cruise can be completed on the CDC-3150 computer within a week after ship returns to the Institute. The data files of a particular cruise usually contain the navigation data file, raw data file and the processed data file.

The raw data file contains all the geophysical data and the navigation data concurrently recorded by the BIODAL system on a cruise. Access to the raw data file is by cruise, day and time (GMT). This file is used for obtaining high quality processed data in combination with the navigation data file. The navigation data file contains the day, time, latitude, longitude of the start fix and end fix of the tracks as well as the navigation method used. The navigation data file is also used for plotting on the Calcomp plotter to show the data coverage for publication or for producing slides.

The processed data give the day, time, latitude, longitude, bathymetry, magnetic anomalies, total magnetic field, free air gravity, total gravity field and simple Bouguer anomalies along the ship's track. This file is used for the majority of geophysical interpretation, analysis and distribution of the data. Copies of this file can be readily exchanged with other scientific institutions.

During the past decade, BIO has obtained about 600,000 kilometres of gravity, bathymetric and magnetic data in the northwestern Atlantic Ocean and Baffin Bay area along the eastern Canadian coast and some other selected areas such as Mid-Atlantic Ridge and the Hudson '70 cruise circumnavigating the Americas. Figure 2 shows the ships' tracks of the geophysical cruises during the period from 1963 to 1971. Table 1 gives the number of geophysical data stations available in the AGC data system.

The processed data files are also maintained

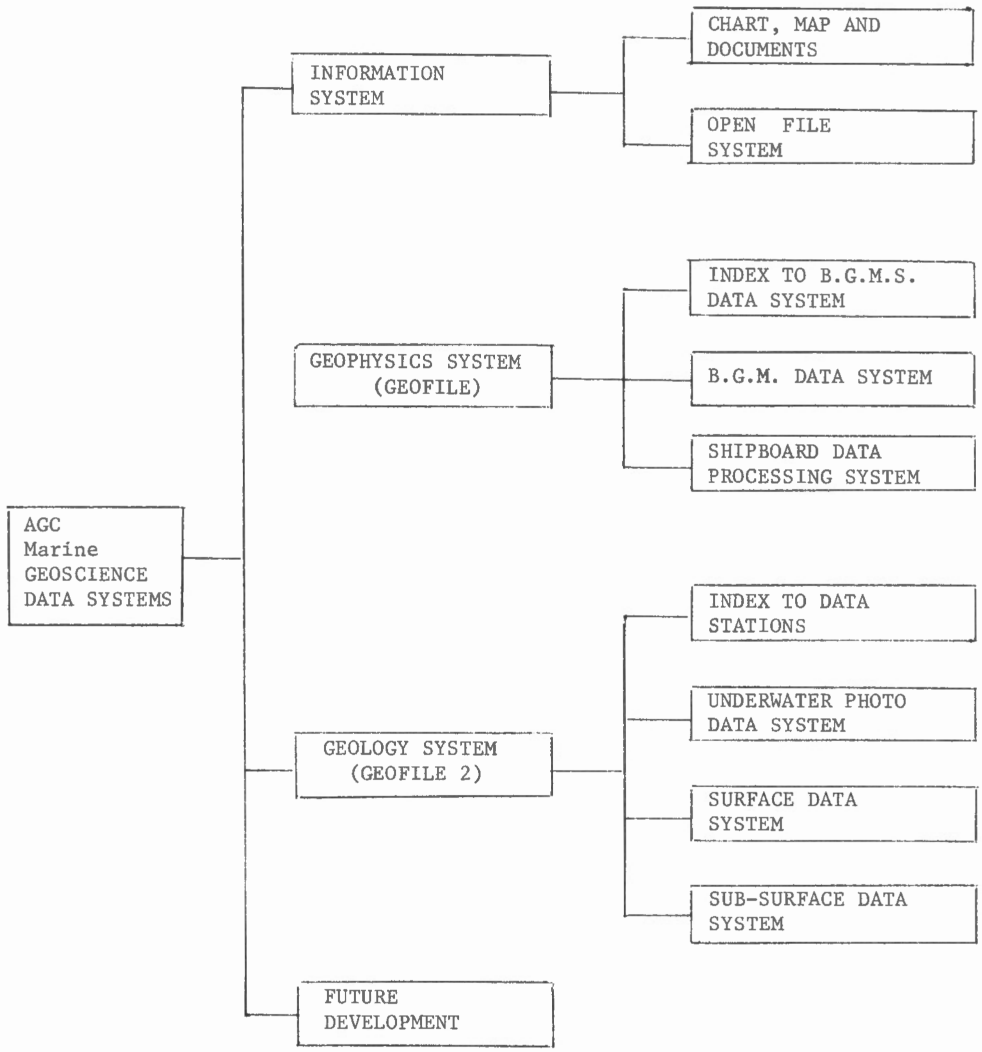


Figure 1. AGC computer-based data system for processing, storing, retrieving and displaying marine geoscience data on the shorebase CDC-3150 computer and on the shipboard HP-2100 computer.

according to area, such as the GEODATABASE data file (Shih and Heffler, 1972) and the binary processed data files. These files are used for retrieving the data in a defined area. A seismic index data file has also been prepared. This file contains the same information as the navigation data file but additional information is available, such as line number, area code, survey methods, etc. This file can be used for plotting the location of the seismic data on the Calcomp plotter with various defined output options.

Marine Geological Data Storage and Retrieval System

Since 1973, methods and techniques for constructing marine geological data files have been investigated and developed. This file gives the station identification, instrumentation, field analysis and field observation of the geological, geochemical or biological stations carried out on the AGC cruises, such as grab, dredge, sediment core, underwater camera, plankton tow, water sample, etc. The data coding sheets are available in the AGC data section. Various output options have been developed for retrieving this file according to area, cruise, year, etc. Figure 3 is an example of one use of this file to show where the grab stations are located. Only about 20 per cent of the geological data collected on AGC cruises has been entered in this file at this time. The size of this file can be found in Table 1.

A well data system is presently under investigation and development. An index file and a geological contact data file are now in preparation. Retrieval routines and data displaying methods are being developed around the CDC-3150 computer output facilities, such as line printer, Calcomp plotter and the EAI flatbed plotter.

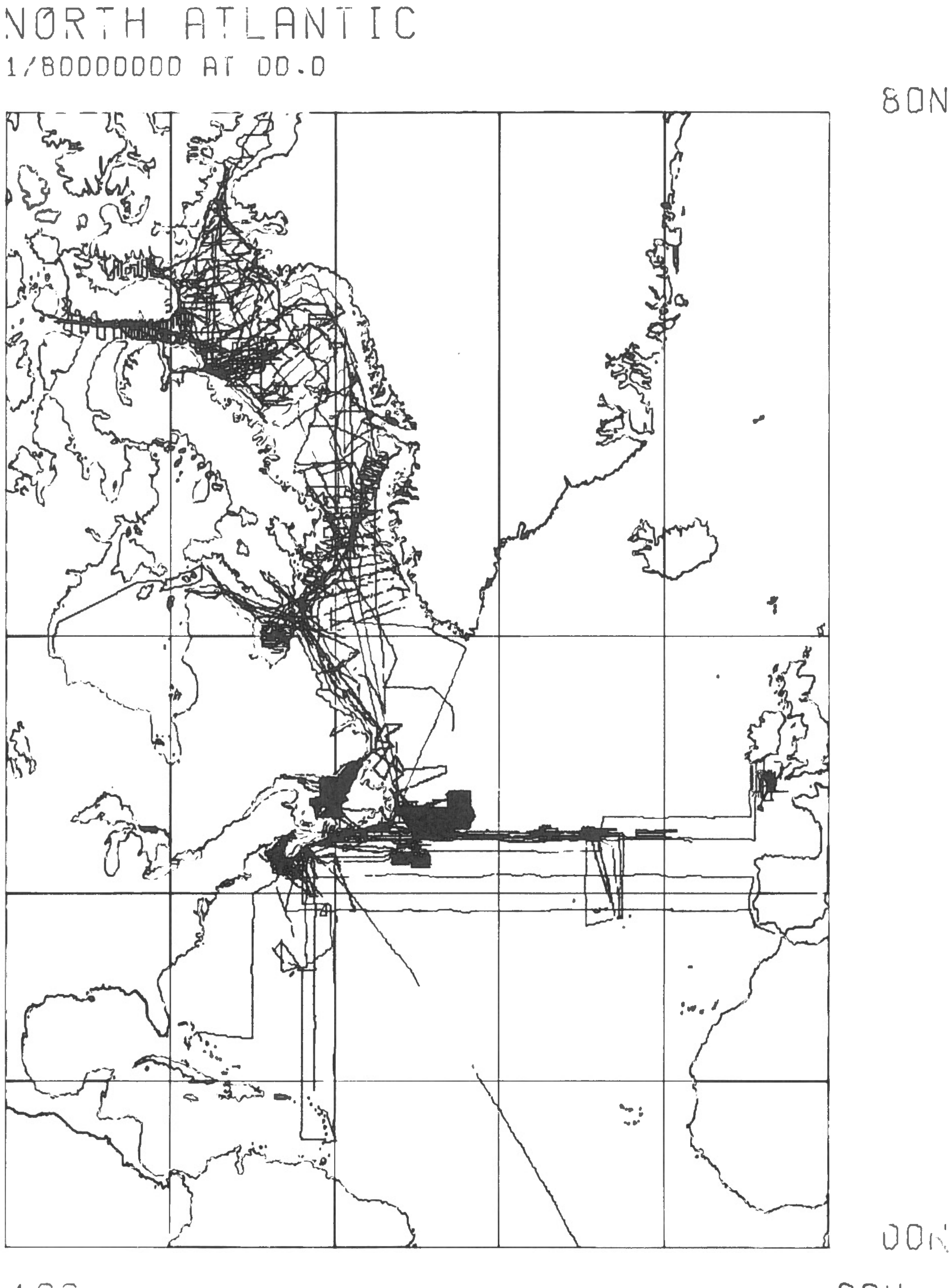


Figure 2. Ship's track of AGC Geophysical Cruises to the end of 1971.

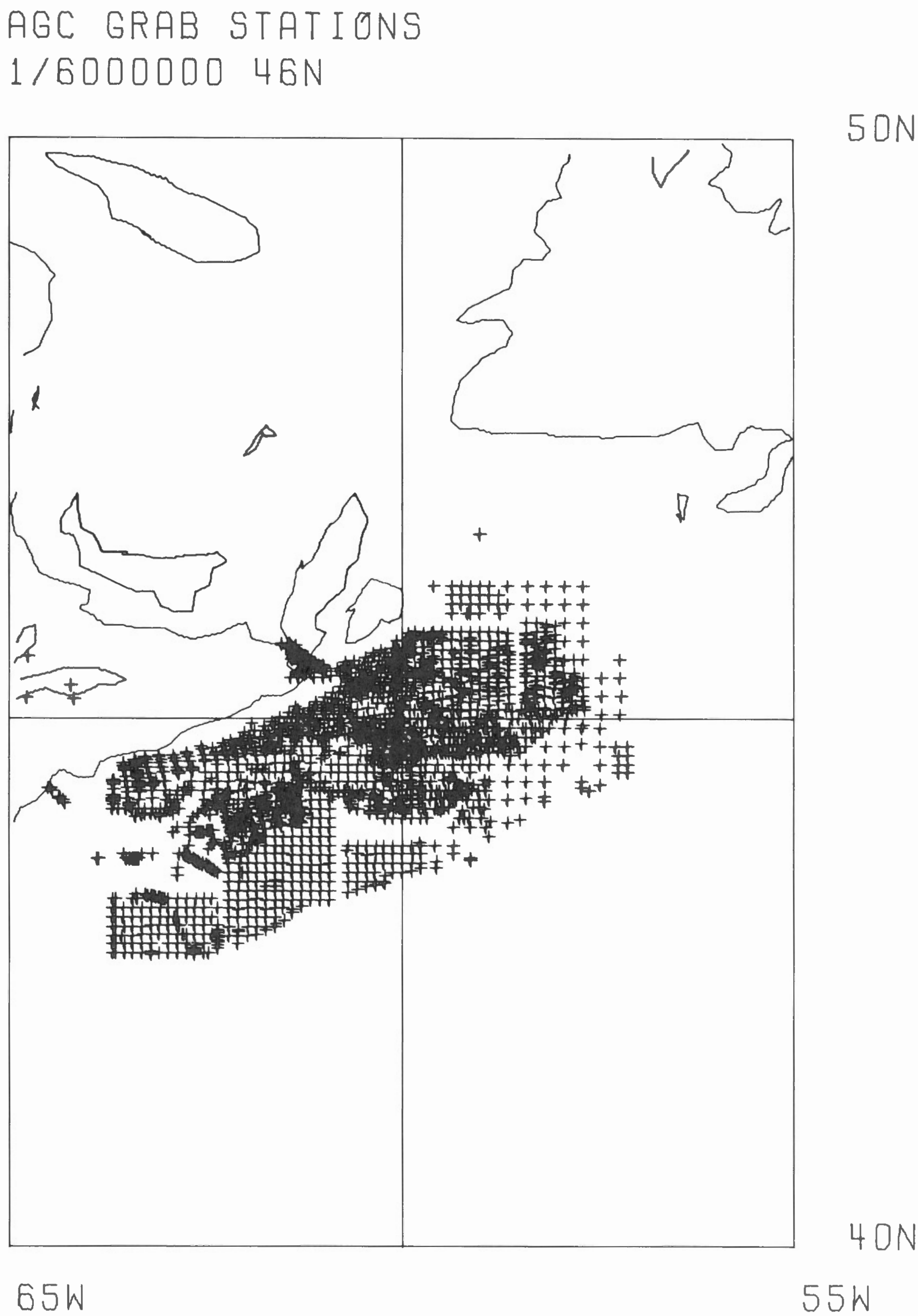


Figure 3. Digitized AGC grab station information to the end of March, 1974.

TABLE 1

Computer-processible Data in AGC Data Systems

(March 1974)

Programs	Ship's Tracks (km)		Data Stations	Analyzed Data Records	File Start Year
	AGC Cruises	Other			
Bathymetry	600,000	10,000	1,500,000	-	1968
Gravity	400,000		800,000	-	1968
Magnetics	600,000	10,000	1,500,000	-	1968
Seismic (Reflection & Refraction)	40,000	120,000			1973
Grab	-	-	2,568		1973
Underwater Camera	-	-	70	7,000 photos	1972
Coring	-	-	254		1973
Dredging	-	-	26	550 specimens	1973
Plankton Tow	-	-	18		1973
Well	-	-	30		1973
In Situ	-	-	509		1973

Pilot projects for storing and retrieving the analyzed results of the geological information obtained from geological samples have also been developed and tested. These include an underwater photograph data file containing the geological interpretation of 7,000 photographs and a geological and geophysical data file containing the analyzed results of 550 specimens obtained from the dredge samples. Computer programs for making a statistical study of these files have been developed and tested. The coding methods and the techniques of system development will be documented in the future.

References

Ross, D.I., Shih, K.G., Johnston, B.L., and Porteous, D.M.
1973: Geofile: a revised manual on the storage and retrieval of geophysical data; BIO computee note, BI-C-73-03, 171 p.

Shih, K.G.
1973: Shipboard computer systems for processing and displaying bathymetric, gravity and magnetic data; BIO report, BI-R-73-13, 37 p.

Shih, K.G., and Heffler, D.E.
1972: Bedford Institute geographically ordered marine geophysical data storage and retrieval system; 24th International Geological Congress, Montreal, Canada. Section 16, p. 76-81.

SUBSURFACE DATA SYSTEM AT THE ATLANTIC GEOSCIENCE CENTRE

K. G. Shih, I. A. Hardy and A. G. Sherin,
Atlantic Geoscience Centre, Dartmouth

A project for setting up a subsurface data system at the Atlantic Geoscience Centre was initiated by Hardy in the summer of 1972 and formalized as project 740007 in early 1974. This system will be utilized to store and retrieve geological information obtained from well data. Two sets of files are currently being constructed: (i) an index data file containing information on drilling operations and performance of wells; and (ii) the geological data file containing the name, age and lithology of the geological formations encountered in each well. Other subject data files will be initiated to include data originating from well studies in geochemistry, micropaleontology and palynology, etc.

In the development of a data storage and retrieval system, the most important part is the definition of the data file which includes the contents, data format and data codes. The data file should be useful to the geologist, simple in file construction, and suitable for data exchange purposes. To simplify the file construction procedures and to consider the data exchange problem, various subject data files and the full geological names, are used in the well data storage and retrieval system. The use of subject data files will keep the files to a reasonable size for ease of updating and decrease the file search time. If a full name is used a coding dictionary will not be required when the data are trans-

Table 1

Contents of the Index Data File

Parameter Code	Record Type	Contents of Data Information
0101	Label	Well number, confidential code, country code, region code, latitude, longitude, NTS, R.T./K.B., length unit code, elevation, ground, total depth, PBTD, sample tray number and pool/field.
0102	Region	Region name (20 characters) and country name (16 characters).
0103	Area	Area name (20 characters).
0104	Well name	Well name (48 characters).
0106	Location	Well location name (48 characters).
0107	Survey	Survey system description (48 characters).
0108		Non-operating description (48 characters).
0109	Scientist	Geology by whom and date.
0110	Operator	Operator's name (48 characters).
0302	Performance	Spud date, finished date, completed date, suspended date, abandoned date, rig released date, and date released date.
0303	Contractor	Name of contractor (40 characters).
0304	Class	Classification (40 characters).
0305	Reason	Drilling reason (40 characters).
0306	Status	Status (48 characters).
0402	Initial gas	Interval, natural flow and S. I. Pressure.
0502	Initial oil	Interval, estimated flow, S. I. Pressure.
0602	Water	Interval and flow type.
0701	Logging	
0801	Casing	
0901	Coring	
1001	Results	Final results information (code by yes or no).
1101	Comments	
9999	END	

Table 2

Contents of the Geological Data File

Parameter Code	Record Type	Contents of Data Information
1201	Label	Well number, confidential code, country code, region code, latitude, longitude, NTS, R.T./K.B., elevation, ground, sea floor, total depth, length unit code and area code.
1202	Section	Well number, length unit code, top, elevation and thickness of the formations, age (106 = 10 million years), geological periods (in numerical code), rock types (15 characters) and geological formations (16 characters).
1203	Comments	
9999	END	

```
                              WELL DATA STATION

 IDENTIFICATION

          WELL NUMBER-BOL 661212
          REGION------HUDSON PLATFORM      COUNTRY-----CANADA
          AREA--------HUDSON BAY           INDEX NO.---
          WELL NAME---SOGEPET AQUITAINE KASKATTAMA PROV. NO. 1
          NEW NAME----
          LOCATION----DELTA * MOUTH OF KASKATTAMA RIVER, MANITOBA
          SURVEY SYS--
          NON-OPERAT.-
          GEOLOGY BY--                     GEOL. DATE--
          OPERATOR----BANFF OIL LTD (BOL)

OPERATION

          RIG TYPE     DIAMOND D
          SPUD DATE            16  9 1966
          FINISHED DATE         0  0    0
          COMPLETED DATE       12 12 1966
          SUSPENDED DATE       17 12 1966
          ABANDONED DATE        0  0 1967
          RIG RELEASED         17 12 1966
          DATARELEASE          17  7 1969
          CONTRACTOR------BIG INDIAN DRILLING CO. LTD.
          CLASSIFICATION--
          REASON(DRILL)---
          STATUS--ABAN/DRY LICENCE NUMBER 2168
          REMARK--OPERATION RESUMED 01/07/67 AFTER SUSPENDED 17/12/66

COMPLETION DATA AND FINAL RESULTS

          PERFORMED          NO
          OPEN HOLE          NO
          PROD. FORMATION    NO
          OIL                NO
          GAS                NO
          PRESSURE           NO
          PROD. INTERVAL     NO
```

Figure 1. Index data file: Output sample.

```
WELL NO.    LATITUDE  LONGTITUDE  N.T.S.      ELEV.   GROUND  SEA FLOOR   T.D.          AREA
JBB0300000    50 47 50    81 20 15     0  0             0       100            0        610  JAMES BAY LOWLAN
  TOP  ELEV. THICK                          GEOLOGICAL PERIOD                    GEOLOGICAL FM.       ROCK TYPE
   37       0    119               DEVONIAN                                    WILLIAMS IS          CLAY+LS+SH
  156       0     20               DEVONIAN                                    ABITIBI R            LS                  UPPER
  176       0    291               DEVONIAN                                    ABITIBI R            LS+GYPSUM+AH        MID(MOOSER
  476       0    131               DEVONIAN                                    ABITIBI R            LS+CONG+BRECCIA     LOWER


WELL NO.    LATITUDE  LONGTITUDE  N.T.S.      ELEV.   GROUND  SEA FLOOR   T.D.          AREA
OUP 301213    50 35  1    81 31  5    42114             0       187            0       1060  JAMES BAY LOWLAN
  TOP  ELEV. THICK                          GEOLOGICAL PERIOD                    GEOLOGICAL FM.       ROCK TYPE
  250       0    285  UPPER    DEVONIAN                                        LONG RAPIDS          SH+CLAY             PORTAGE
  535       0    301  UPPER    DEVONIAN          MIDDLE   DEVONIAN             WILLIAMS IS          LS+SH               TULLY-HAMI
  836       0     37  MIDDLE   DEVONIAN                                        ABITIBI R            LS+CLAY             ONONDAGA
  873       0    154  MIDDLE   DEVONIAN          LOWER    DEVONIAN             MOOSE R              LS+SS+GYPSUM
```

Figure 2. Geological data file: Output sample of geological formations of Devonian age.

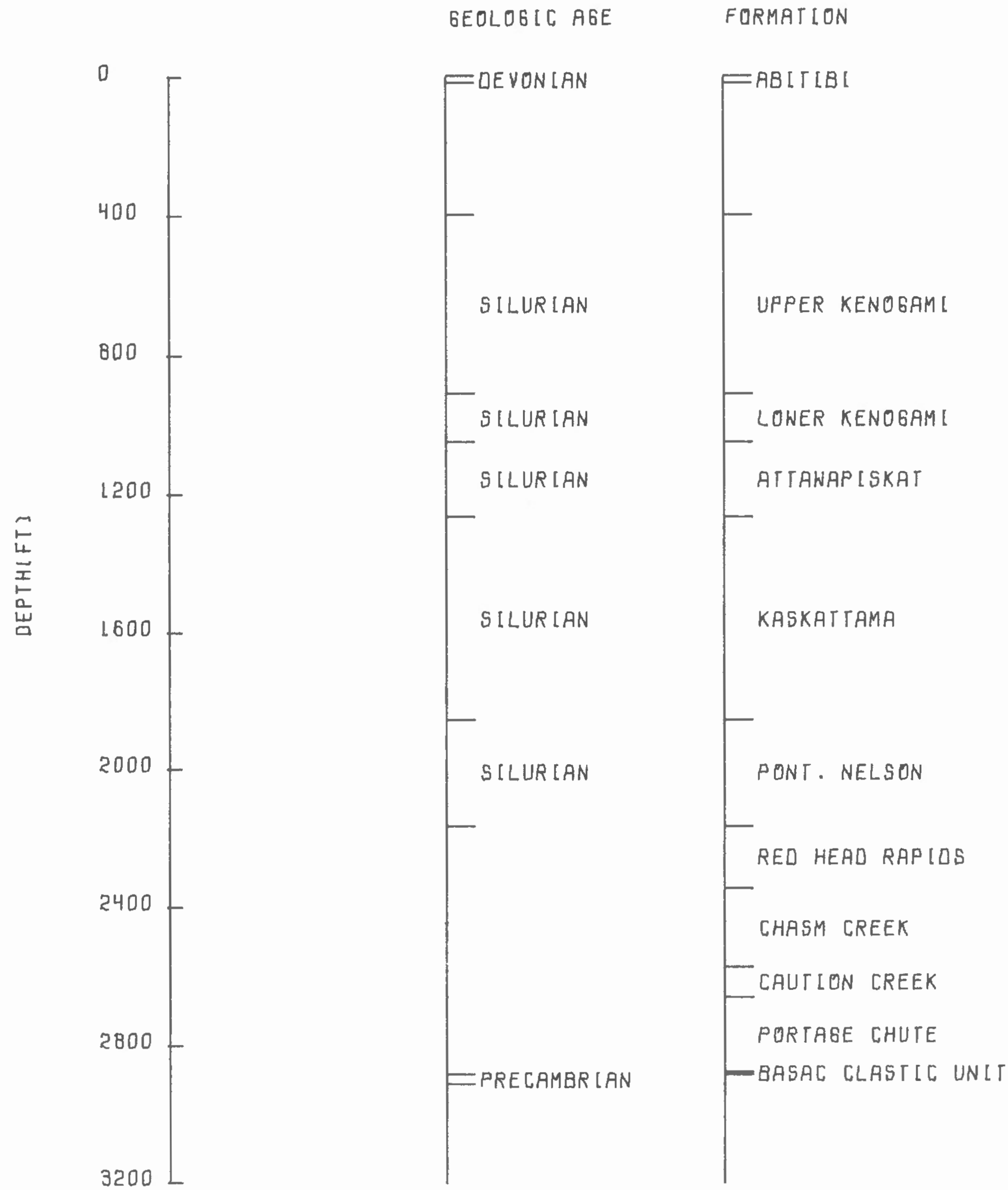


Figure 3. Geological data file: geological age and geological formation plot.

ferred to the data coding sheet, retrieved from the data file, or reformatted for data exchange purposes.

The contents of the index data file are given in Table 1. The parameter code is used for defining the data format and specifying the required parameter when the data is retrieved. Table 2 gives the contents of the geological data file.

The input portion of the system consists of coding sheets and software for loading the data onto magnetic tape. The data on magnetic tape are card images and are blocked in the BCD mode. The retrieval portion of the system consists of the software to retrieve the information according to the given output option and provide the required parameters such as latitude, longitude, depth, name of geological terms, etc. The display portion of the system presents the data in various forms useful to the geologist. The displaying techniques are developed around the output facilities such as the line printer, Calcomp plotter and EAI flatbed plotter. Figures 1, 2 and 3 show some of the available display formats in the well data system to date. The update portion of the system will be designed to correct, delete and add data as required to the data files. The merge portion will make it possible for retrieving data from two or three subject data files, permitting correlation studies to be made by the geologist.

The prime concern in the development of this system is its usefulness to the geologist. The output of this system, especially in the area of graphic presentation, is designed to present the data in a format often used by the geologist. This system has the potential of becoming a powerful tool to the subsurface geologist, who may now solve many geological problems without the task of editing large volumes of data.

URBAN GEOLOGY AUTOMATED INFORMATION SYSTEM (U.G.A.I.S.)

J.R. Belanger,
Terrain Sciences Division

In 1970, the Terrain Sciences Division initiated a pilot project called "Environmental Geology Prototype Study, Ottawa-Hull Region", to investigate the requirements for geological and related earth sciences data for efficient land-use planning in urban areas. One of the goals of the project was to develop a computer-oriented system to compile, evaluate and present geological information to meet the needs of people concerned with regional planning.

In the fall of 1971, the first phase of the project was completed and the first maps of the surficial geology of the National Capital region were produced with an adapted version of SYMAP, a standard printer-produced map plotting package. This success encouraged the Geological Survey to undertake a second phase of the project: to adapt the system to the major urban centres of Canada and to develop a retrieval and cartographic output adapted to the needs of the users of urban geology.

Input

Although the system can accept almost any type of geological information, the main source of information is the analysis of engineering boreholes. The lithologic logs on the original records (1)* are coded and transferred (2) on standard forms called Data Record Sheets. To permit the manipulation of the records with standard filing systems (manual data bank) (4) as well as with electronic data processing techniques, the mnemonic codes are of standard length but as close as possible to the English word.

Processing

The processing of the basic information (5) is done in two steps to faciliate the correction of errors and reduce the overall cost. Each city is processed separately since the format and mnemonic codes can vary from one region to another.

Preparation of Data

Since the Data Record Sheet consists of two to eleven punched cards depending on the amount of information available, the first step is to transfer each borehole from several card images to one record image. While transforming the format of the records the program (9) verifies for possible sequence or key-punching errors and thus creates a "good" (error-free) file (11) and an error file (12). The error file is then corrected (14) and verified again until all the information is corrected. During the first step of the processing, the basic information is also standardized since the Date Record Sheet permits the inclusions of non-standard parameters in order to be more universal.

Correction of Geological Parameters

The second control of the information is to verify the validity of the mnemonic codes and geological information (16), and to standardize (21) the final format of the data bank so the information can be accessed by only one retrieval program. All mnemonic codes in the form of characters are changed to numeric format to facilitate further analysis of the information.

Output

The retrieval and processing (25) is oriented toward two types of output: a reproduction of the information stored in the data bank in a format similar to an engineering report (28) or a graphic display of the interpreted data (27).

Until now the graphic display of the information consisted of location maps and contour maps. The location maps are used to indicate the position and nature of the information stored in the data bank. The scale, limits and type of phenomenon are specified in a driver program (26) which activates the retrieval system.

The contour maps are produced in three steps: a search of the data bank to retrieve the control points, the generation of a systematic grid by an interpolation algorithm and the contouring of the gridded data. The interpolation algorithm was developed for the Urban Geology Project and therefore tries to interpolate values in the same way a geologist would when dealing with continuous phenomena.

The contouring program is developed for the EAI flat-bed plotter using the momentum of the plotter "head" to produce a smooth line joining the interpolated points. Figure 2, part of the urban geology of Ottawa series, shows the current stage of the project for the production of automated contour maps. The entire map (including the contour, legend, scale, etc.) is produced by the plotter. The dots show the location of the original data points. The map was reduced photographically from the scale of 1:25,000.

Note that the contour lines are produced by a series of straight segments joining the gridded or sub-gridded points, since the plotter software for the "Freerun mode" is still in development.

Future Development

A series of papers and reports will be issued in the near future on the Urban Geology Project to describe the development and application of the U.G.A.I.S. They will include: a user's reference manual for data

*The numbers refer to the U.G.A.I.S. flow chart - Figure 1.

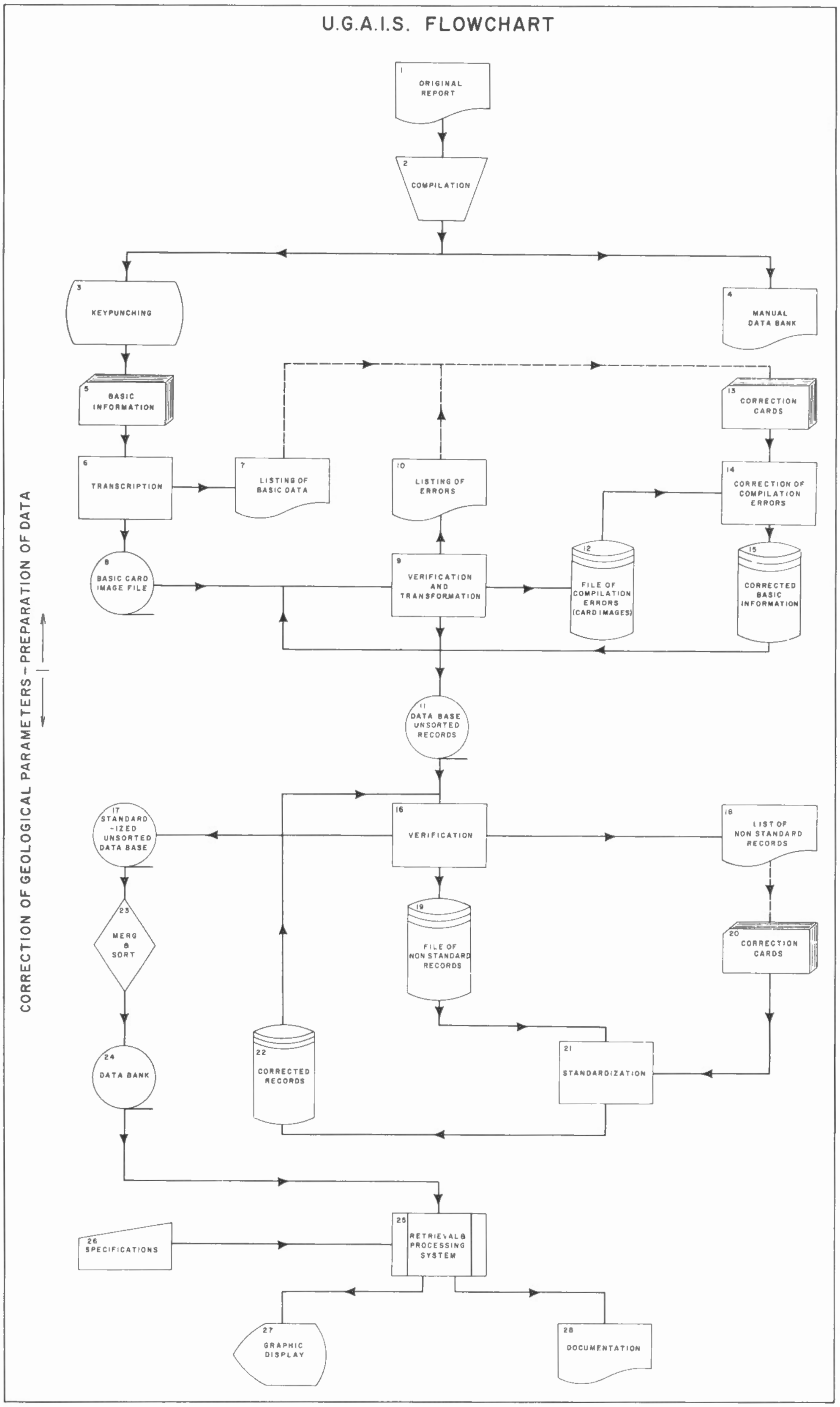


Figure 1. Flow chart of the Urban Geology Automated Information System.

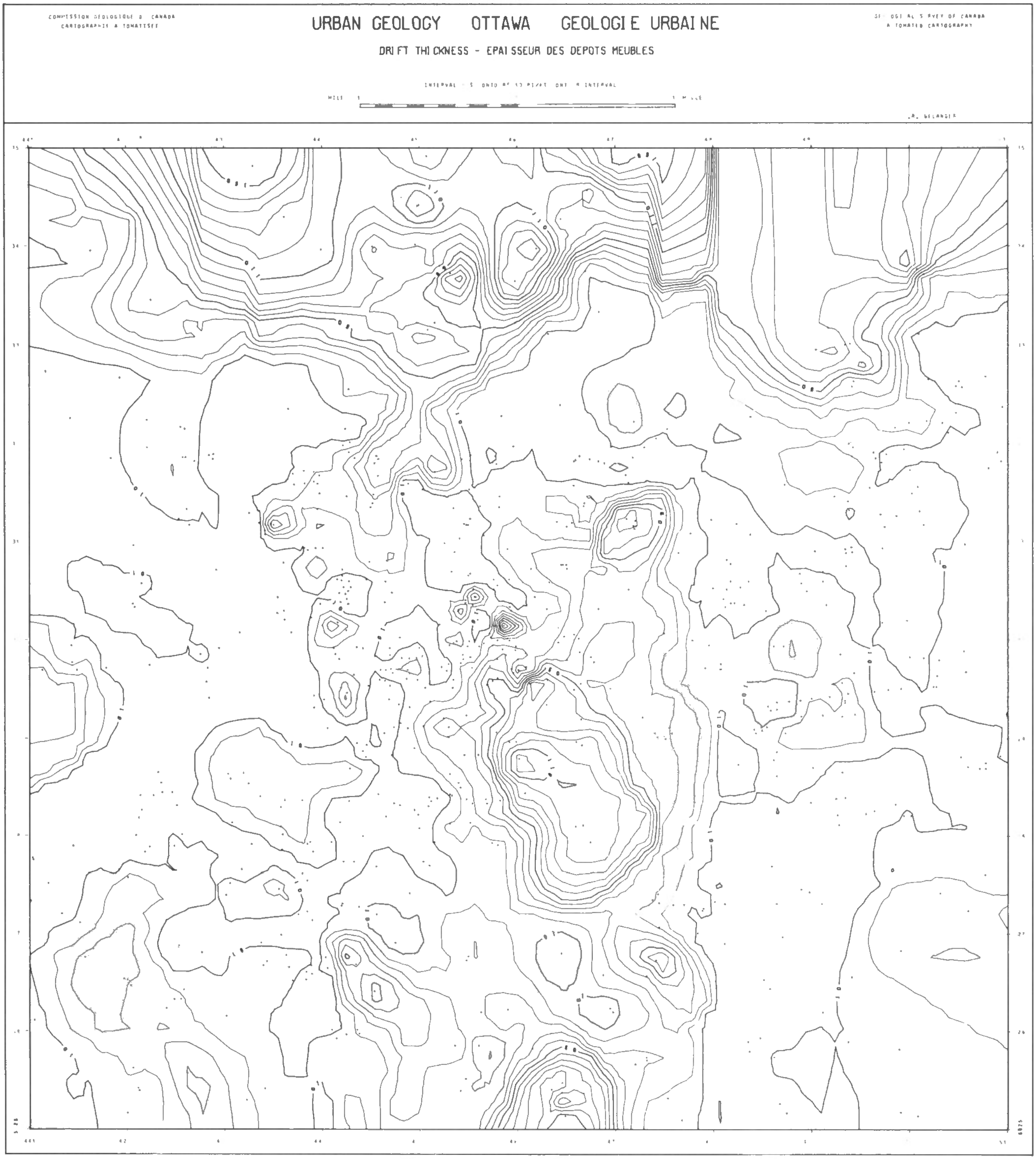


Figure 2. Typical contour map output of the U. G. A. I. S.

collection and processing of the information, technical reports on the methods used for the automated cartography, computer programs and an atlas of the Ottawa-Hull region showing several applications of the U. G. A. I. S.

The next stages of the project will be the optimizing of the processing and retrieval systems, a detailed study of the applications of the current system, the improvement of the graphic output and the development of new programs such as "interpolated stratigraphic sections".

References

Bélanger, J. R.
1974: L'Axe géologique de l'environnement; GEOS, E. M. R., hiver 1974, p. 16-17.

Bélanger, J. R., and Morin, J. M.
1971: Synagraphic computer mapping; Manual of detailed geomorphological mapping, Brno, Czechoslovakia, Chap. 8.

Bélanger, J. R., and St-Onge, D. A.
1973: Environmental geology prototype study - Ottawa-Hull region; *in* Report of Activities, Part A, April to October 1973, Geol. Surv. Can., Paper 74-1, pt. A, p. 215.

Scott, J. S.
1970: Environmental geology prototype study - Ottawa-Hull region; *in* Report of Activities, Part A, April to October 1970, Geol. Surv. Can., Paper 71-1, pt. A, p. 156.

1971: Environmental geology prototype study, Ottawa-Hull region; *in* Report of Activities, Part A, April to October 1971, Geol. Surv. Can., Paper 72-1, pt. A, p. 147-148.

1972: Environmental geology prototype study, Ottawa-Hull region; *in* Report of Activities, Part A, April to October 1972, p. 186-187.

St-Onge, D. A., and Bélanger, J. R.
1973: Urban geology liaison; *in* Report of Activities, Part A, April to October 1973, Geol. Surv. Can., Paper 73-1, pt. A, p. 227.

30. MATHEMATICAL SIMULATION OF AGGRADATION AND DEGRADATION OF PERMAFROST

R.M. Isaacs,
Terrain Sciences Division

The aggradation and degradation of permafrost due to climatic changes and engineering activities are being simulated mathematically with a computer program (TRAMPS - Temperature Analysis of Multi-phase Systems) developed by P. Hibbert of the Computer Science Centre, Department of Energy, Mines and Resources. An important assumption inherent in the program is that heat transfer can occur by conduction only.

The analysis of the heat flow problem is being accomplished using the finite element method. There are several examples in the literature on the use of this approach with respect to linear cases; because of the large computational times involved, however, it has not been applied extensively to highly non-linear problems encountered in phase change phenomena.

Some of the assumptions inherent in the method are as follows:

a) Every cross-section is the same as every other cross-section. This implies that there is no heat flow parallel to the continuous heat source.

b) The cross-section being investigated is composed of a large number of triangular elements which are interconnected at the nodes or joints.

c) Each element may be composed of a different material but the material is uniform throughout the area of each element.

d) The temperature variation across an element is constrained to being linear.

e) Specified nodes may be designated as heat sources or sinks by specifying their temperatures, which may be time dependent.

f) The latent heat effect does not occur instantaneously at a particular temperature but is considered to be part of the specific heat and is spread over a small but finite temperature range.

g) The variation of thermal conductivity with temperature also is taken into account.

When the network of elements is being constructed, care must be taken to ensure that small elements occur in regions of steep changes in the thermal gradient so that the temperature profile may be modelled adequately. The network also must be large enough in extent so that the thermal changes will be negligible at the boundaries at the latest times investigated.

The solution of this non-linear problem requires that a large number of linear solutions for small time-steps be carried out and accumulated. Thus in the limit, the model approaches reality as the time increment approaches zero. It is important that the time interval is small enough so that each element, as it passes through a phase change, exists in the temperature range of the change for a number of time-steps so that the latent heat may be incorporated fully.

Studies of the thermal degradation of permafrost beneath a road built about 30 years ago in the Norman Wells, N.W.T. region have shown that the computer program seriously underestimates the amount of thaw that may be caused by the road building. As a result, the hypothesis has been advanced that convection of water into the permafrost may be occurring; thus, it is intended to modify the existing program to take this into account using theoretical data of moisture migration under thermal gradients supplied by R.L. Harlan, formerly of the Department of the Environment.

Once modified, this program may be used to study the build-up of ice lenses around a chilled gas pipeline, taking into account the existing thermal gradients, thermal properties, and varying temperatures in the pipeline.

31. MACKENZIE VALLEY GEOTECHNICAL DATA BANK

D.E. Lawrence,
Terrain Sciences Division

Environmental assessment and geotechnical investigation have been widespread in the Mackenzie Valley in the past few years especially with regard to investigation for possible routes for the Mackenzie Highway and for gas and oil pipelines through the valley. Since construction problems and control of environmental and engineering concerns are most difficult in permafrost terrain, a project was undertaken to assess the presently available data in order to establish the engineering characteristics of materials and to correlate the controlling geological, engineering, and hydrological factors which influence the occurrence of ground ice and permafrost.

The principal source of geotechnical data was the reports of engineering investigations for the proposed highway and pipelines and are in the form of borehole logs. Geological data was available in the form of 1: 125,000 scale surficial geology maps produced by the Geological Survey of Canada.

Because of the large amount of available information (approximately 7,500 boreholes with up to 25 different geological, hydrological, morphological, climatic and engineering test parameters recorded for each) it was decided at the inception of the project to establish a data bank to permanently store all the information so that it would be available for future analysis.

The primary purpose of the project was to establish the properties and their ranges for various natural materials under various conditions within the Mackenzie Valley and Delta, and to establish the permafrost condition and ground ice content so that engineering design and construction could be carried out with minimum environmental impact. Secondly to assess the various methods of drilling and recovery of samples in permafrost terrain and to assess the relative merits of the various field and laboratory tests carried out on the samples. Should current practice not prove fruitful alternate procedures would be suggested. Thirdly, in areas of little information the knowledge gained from the above would be used to predict the occurrence of ground ice and thus anticipate some of the engineering-environmental problems before they are encountered and remedial measures are required.

The data bank is based on drillhole records contained in the reports of various geotechnical consultants. Information is coded along with data on various site characteristics. Each record contains a five-digit identifying number, a header line of 74 characters which includes 27 variables specifying location, topography and technical aspects of the drilling site; 18 lines of 74 characters in which up to 30 geotechnical and geological characteristics of each soil horizon can be recorded, the last recorded horizon being the bottom of the hole, and a 75 character line in which other pertinent or explanatory data can be recorded. Data records were verified and amended prior to being incorporated into the data bank.

The completeness and accuracy of the data stored for any drillhole is dependent on the quality of the original record. In general, the capacity of the data bank is far in excess of the amount of information currently available. Actual values are used for most variables, with only limited use being made of coding or grouping so that the accuracy of the information in the data bank depends almost entirely on the quality of the input data.

The information is stored as records on seven-track tape, each record being 1,486 characters in length. The size of each record and the number of records on file lend themselves readily to manipulation using COBOL. The most efficient procedure is to extract the records which are to be analyzed and copy them onto a new tape which is then used as the input file for further programming.

Reference

Lawrence, D.E.
1973: Review of geological and geotechnical data Mackenzie Valley unpublished report, prepared for Dept. of Indian and Northern Affairs, Environmental Working Group, Mackenzie Highway, 64 p.

AUTO-CARTOGRAPHY IN THE GEOLOGICAL SURVEY OF CANADA – DEVELOPMENT IN AUTOMATION

P. Debain
Geological Information Processing Division

The Geological Survey's Auto-cartography Unit was formed in June 1973. Since then one experimental map, Edehon Lake has been produced. The adjacent sheet Nueltin Lake is in the coding stage. The revised base of the 1/5,000,000 Map of Canada is ready for coding and three 1/1,000,000 series maps are in progress.

The relatively easy success in producing the first geological map by means of automation is due to:

1) Surveys and Mapping Branch development and implementation of an effective system.
2) The training GSC Auto-cartography staff received from Surveys and Mapping Branch.
3) The enthusiasm of the cartographers of the Survey toward this new approach in cartography.

Automated Cartography at Surveys and Mapping Branch has been described in different publications (Montagano, 1971, Harris, 1972 and Linders, 1973). In this publication, the applications of their methods to our specific needs for geological maps will be described.

Auto-cartography is controlled by the same principles as in Surveys and Mapping. Only the procedures are different as shown in the following table and in Figure 1.

EQUIPMENT	SURVEYS AND MAPPING AUTO-CARTOGRAPHY	GSC AUTO-CARTOGRAPHY
Digitizing system	Five digitizing tables	One digitizing table "Digi-Grid" Table size work area: 36" x 48" Resolution 0.001" Repeatability ± 0.001" Absolute accuracy ± 0.005" One Alphanumeric CRT Terminal and seperate keyboard
Controller	One PDP 10 One PDP 11/20	One PDP 11/45 (hooked-up to Surveys and Mapping PDP 11/20)
Atomatic Drafting Machine	One table Kongsberg Kingmatic Table size 120 cm by 150 cm Maximum speed 10 metres per minute Resolution (step size) ± 0.0062 mm Positional accuracy ± 0.05 mm Repeatability ± 0.015 mm	No table The GSC plotting in done on Surveys and Mapping table
OPERATIONS		
Source Material	Air survey compilation transferred on digital scribecoat	(1) Topographical base-map (2) Geological manuscript The two are combined to produce an image on a scribecoat. The geological linework is scribed, transferred on a reverse positive and then onto the digital scribecoat.
Codes	A permanent list of codes is established. Each code pertains to a specific feature. Three types of codes: Point feature, open feature and closed feature.	A general list of codes is established for the GSC. For each geological map a distinct dictionary of codes is prepared.
Digitizing	Is accomplished on an etched image with a point cursor for time mode, ordinary cursor for point mode. Choice of scan rate. UTM grid coordinates control the map features.	Same procedure
Plots	Made from standard codes	Prepared for each geological map according to specifications written by GSC Auto-cartography.

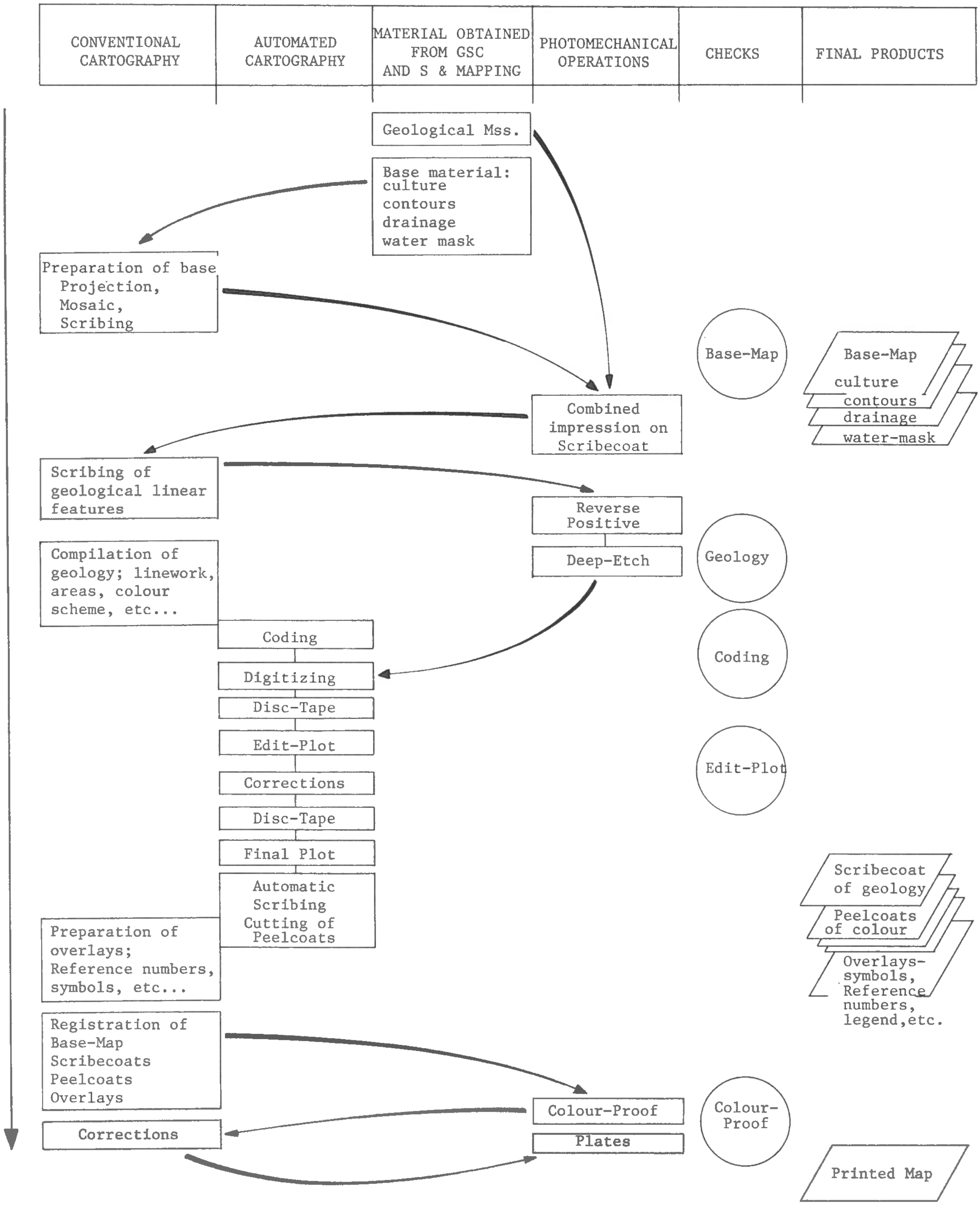

CONVENTIONAL CARTOGRAPHY
AUTOMATED CARTOGRAPHY
MATERIAL OBTAINED FROM GSC AND S & MAPPING
PHOTOMECHANICAL OPERATIONS
CHECKS
FINAL PRODUCTS
Geological Mss.
Base material: culture contours drainage water mask
Preparation of base Projection, Mosaic, Scribing
Base-Map
Base-Map culture contours drainage water-mask
Combined impression on Scribecoat
Scribing of geological linear features
Reverse Positive
Deep-Etch
Geology
Compilation of geology; linework, areas, colour scheme, etc...
Coding
Coding
Digitizing
Disc-Tape
Edit-Plot
Edit-Plot
Corrections
Disc-Tape
Final Plot
Automatic Scribing Cutting of Peelcoats
Scribecoat of geology
Peelcoats of colour
Overlays-symbols, Reference numbers, legend, etc.
Preparation of overlays; Reference numbers, symbols, etc...
Registration of Base-Map Scribecoats Peelcoats Overlays
Colour-Proof
Colour-Proof
Plates
Corrections
Printed Map

Figure 1

The three areas requiring improvements are: Source Material, Codes and Plots.

Source Material: As geological maps are not necessarily prepared within the UTM or the NTS system the Geological Survey's cartography section assembles topographical maps or parts of them together to produce a base-map. Very often projections and UTM grids are calculated.

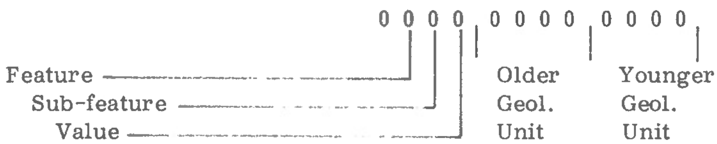
When a complete coverage of Canada will be retrievable at Surveys and Mapping Branch this situation will be greatly improved.

Codes: Coding geological features is the major problem we have encountered up to now. The digitizing, the correction and the plotting have been solved to a reasonable degree by the Auto-carto of the Surveys and Mapping Branch. We have used their system but had to code our maps by a method different from theirs. At the beginning we adapted topographical codes to geological features; now we have our own geological codes.

CODES FOR GSC MAPS

0 000 0000 0000

Geological boundary or associated features: Fault, Dyke, Isograd (Feature qualifiers 0, 1, 2, 3, 4)



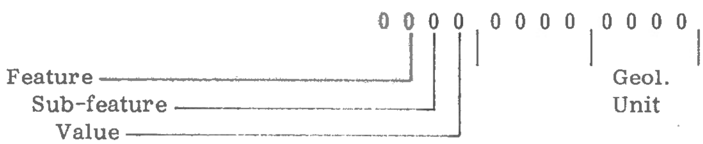
Example: Normal, Fault, approximate, is the geological boundary between Unit 16a and Unit 123

0222 0161 1230

The subdivisions of a Unit such as a, b, c, etc........ will be coded respectively as 1, 2, 3 etc....

Fault, Dyke, Isograd not being a boundary (feature qualifiers 2, 3, 4)

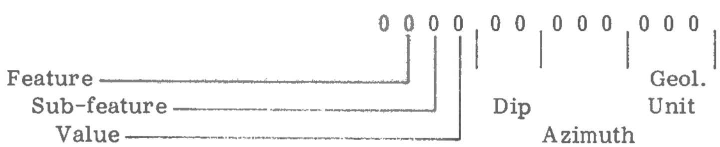
Folds, structural trend, lineaments (Feature and sub-feature qualifiers 50, 51, 52, 53, 54, 55, 56, 80, 81, 82)



Example: Syncline, defined, is located in Unit 16c

0521 0000 0163

Bedding, schistosity, gneissosity, cleavage, foliation, axial plane, lineation (Feature and sub-feature qualifiers 70, 71, 72, 73, 74, 75, 76, 77, 78, 79)



Where these symbols are in use, legend blocks will not exceed the number 90.

Example: Bedding, top known, inclined, dip 23, azimuth 212, is located in Unit 5

0712 23 212 050

With these codes retrieval of information is possible automatically, The Feature qualifier should always be on call.

		Feature Qualifier		Sub/Feature Qualifier		Value Qual.	Geol. Units Codes
0		0		0		0	
0	Not labelled	0	Not labelled	0	Not labelled	0	
	Geological Boundary	1		0	Defined	1	
		1	Limit of Geological Mapping	1	Approximate	2	
		1	Border	2	Assumed	3	
		1		3	Gradational	4	
		1		4	Inferred	5	
		1		5	Under water	6	
		1		6	Colour	7	
		1		7	Overturned	8	
		1		8		9	
		1		9			
	Faults	2		0			
		2	Transcurrent Fault	1			
		2	Normal Fault	2			
		2	Vein Fault	3			
		2	Thrust Fault	4			
		2		5			
	Dykes	3		0			
	# 1	3		1			
	# 2	3		2			
	etc.	3		3			
		3		4			
		3		5			
	Isograds	4		0			
	# 1	4		1			
	# 2	4		2			
	etc.	4		3			
		4		4			
		4		5			
		4		6			
		4		7			
		4		8			
		4		9			
	Fold	5		0			
		5	Anticline	1			
		5	Syncline	2			
		5	Anticline or syncline	3			
		5	Antiform	4			
		5	Synform	5			
		5	Antiform or synform	6			
		5		7			
		5		8			
		5		9			
	Bedding	7		0	Not labelled	0	
		7	Tops known	1	Horizontal	1	
		7	Tops unknown	2	Inclined	2	
	Schistosity	7		3	Vertical	3	
	Gneissosity	7		4	Dip unknown	4	
	Cleavage	7		5			
	Foliation	7		6			
	Axial plane	7		7			
	Lineation	7		8			
		7		9			
	Structural trend	8		0			
		8	From air photo	1			
	Lineaments	8	From air photo	2			

Alphanumerics

Due to the complexity of geological maps the Auto-Cartographic Unit had to develop concurrently with the 12 digit code just described a sequence of alphanumerics to code the geological information on the deep-etch.

These pairs of alphanumerics are recognized by the computer as valid codes at the digitizing and at the plotting stages. A dictionary of the equivalence of the 12 digits and the 2 alphanumeric codes is prepared for each geological map.

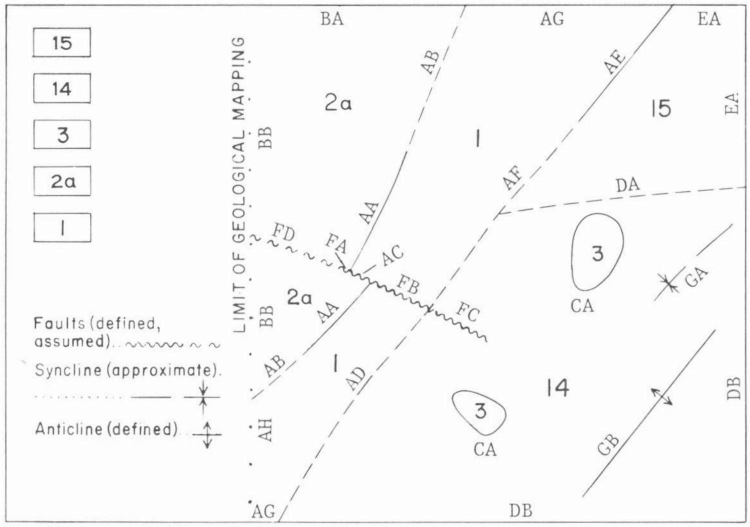


Figure 2.

EQUIVALENT ALPHANUMERICS
FOR CODING ON DEEP-ETCH

This part reserved for corrections.

AA, AB, AC, AD	AZ	A0, A1 .... A9
BA	BZ	B0, B1
CA	CZ	C0, C1
DA	DZ	D0,
EA		
FA		
GA		
HA		
IA		
JA		
KA		
LA		
MA		
NA		
OA		
PA		
QA		
RA		
SA		
TA		
UA		
VA		
WA		
XA		
YA		
ZA		
0A		
1A		
2A		
3A		
4A		
5A		
6A		
7A		
8A		
9A, 9B 9Y, 9Z		90, 91 99

Plots

After digitizing the linework, we obtain edit-plots of each list of codes and make the corrections on the digitizing table to obtain the final plots, cutting of peelcoat and scribing of linework.

Beside the specifications needed for each map, we need a flexible program to establish the origin of each map on the Automatic Drafting Machine (ADM), that is to register each negative, scribecoat or peelcoat according to the press regulations. This program is in preparation at the present time.

A portion of the first geological map produced partly by automation, Edehon Lake Experimental Map 1, is reproduced below (Figure 3). The geological boundaries have been scribed on the ADM, and the colour separation was made by retrieval of geological information. For this publication, colours have been replaced by black screen, however the basic separation process remains the same - (Fig. 3).

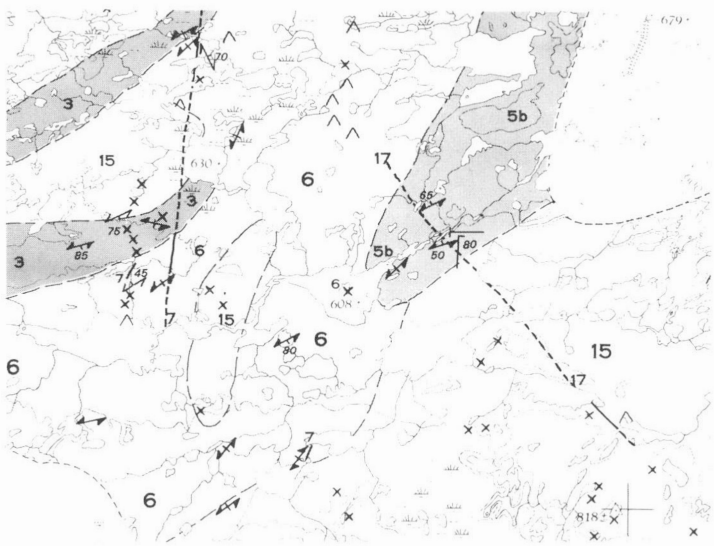


Figure 3.

As an example follow the preparation of codes, to digitize Figure 2, to scribe the linework and to separate the colours by the peel-coat method.

Map-units	12 digits codes	Alphanumerics
1/2a	0101 0010 0021, 0102 0010 0021, 0201 0010 0021,	AA, AB, AC
1/14	0102 0010 0140,	AD
1/15	0101 0010 0150, 0102 0010 0150	AE, AF
1/Border	0127 0010 0000	AG
1/Limit of Geol. Mapping	0117 0010	AH
2a/Border	0127 0021 0000	BA
2a/Limit of Geol. Mapping	0117 0021 0000	BB
3/14	0101 0030 0140	CA
14/15	0102 0140 0150	DA
14/Border	0127 0140 0000	DB
15/Border	0127 0150 0000	EA
Features		
Fault (defined)	0201 0000 0021, 0201 0000 0010, 0201 0000 0140	FA, FB, FC
(assumed)	0203 0000 0021	FD
Syncline (approx.)	0522 0000 0140	GA
Anticline (def.)	0511 0000 0140	GB

Assuming 1 to be printed in blue
2a in green (blue & yellow)
3 in orange (red & yellow)
14 in red
15 in yellow
we will have three colour-plates: blue, yellow and red
Colour scheme: blue plate .....1, 2a - yellow plate 2a, 3, 15 - red plate 3, 14

List of codes to call for cutting the peel coats on the Automatic Drafting Machine (ADM)

blue	.1.. ..1. ...., .2.. ..1. ....
	.1.. ..2. ....
yellow	.1.. ..1. ..21, .1.. ..21 ....
	.1.. ..3. ....
	.1.. ..1. .15.
	.1.. .14. .15.
	.1.. .15. ....
red	.1.. ..3. ....
	.1.. ..1. .14.
	.1.. .14. ....

List of codes to call for scribing the linework on the ADM

boundaries:	defined	.1.1 ....
	approximate	.1.2 ....
faults	defined	.2.1 ....
	assumed	.2.3 ....
syncline	approximate	.522 ....
anticline	defined	.511 ....

With the present equipment and the software facilities at hand we can start to produce geological maps by automation. The 1/1, 000, 000 series will be on tape and thus available for easy revision. We hope that in the future many other applications will be used by geologists such as generalization, retrieval of geological units over large areas and many others the Cartographers are not aware of in specialized fields of geology.

References

Montagano, G.A.
1971: Automated Cartography Development at Surveys and Mapping Branch, Technical Report No. 70-11.

Harris, Lewis J.
1972: Automated Cartography in Federal Mapping in Canada, The Canadian Cartographer, v. 9, no. 1.

Linders, J.G.
1973: Computer Technology in Cartography, International Yearbook of Cartography XIII.

COMPUTER USE IN THE GEOLOGICAL SURVEY LIBRARY

Doreen Sutherland,
Geological Information Processing Division

The library of the Geological Survey is using two computer files to control its record keeping. Briefly they are:

PERFIND. A list for periodicals and serials showing title, call number and stack number, with a few added codes to enable the library to identify such records as method of ordering, frequency, etc. It is used by the library staff and users, to find the shelf location of the periodicals. The advantage over manual methods is greater frequency in updating information, ease of producing copies and greater detail possible in the records themselves. Updated quarterly. On the CDC 6400.

FAMULUS. This is a system for small files put on the CDC 6400 by the Earth Physics Branch. The library is using it as an "on order" file, which lists under author, title and organizations all monographs on order. Updates add new orders and delete those received. Advantages are the possibility of more cross-referencing, and availability to users. Updated monthly. Written in Fortran.

GEODAT is an acronym for a computerized, multiple cross-reference geoscience data bank for the numerical results produced by the laboratories in the Geological Survey of Canada ... the file consists mainly of chemical and spectrographic analyses reported in oxide per cent or parts per million; and published C^{14} and D/Ar age determinations.

Originally accessed by COBOL, now by using the MARS VI retrieval system on the CDC 6400, and with continuous additions, the file now contains some 80,000 - 100,000 analyses reported from 1956 to the present. The retrieval programs are still under development, but the results of the analyses up to within three years of the date of the request are available to users in the private sector. Prices vary, but would probably not exceed $100.00 for results of any specific analysis.

References

Dawson, K. R.
1970: Description of the Geological Survey analysis requisition forms; Geol. Surv. Can., January, 1970, 20 p. Unpubl. report.

1971: Description of the GEODAT geoscience data-bank codes, retrieval procedure, output formats; Geol. Surv. Can., Open File no. 43, Ottawa, 1971, 35 p.

34. COMPUTER-BASED ACTIVITIES IN THE CENTRAL LABORATORIES AND TECHNICAL SERVICES DIVISION, GEOLOGICAL SURVEY OF CANADA

G. R. Lachance,
Central Laboratories and Technical Services Division

A. Departmental Computer Science Centre

1. Electron microprobe analysis of geological materials involves the correction of measured intensities for matrix effects and the calculation of weight fraction by comparison to appropriate standards. The program EMPADR VII (Rucklidge and Gasparrini, 1969) is used to correct for absorption, enhancement and atomic number effects for up to 20 constituents.

2. Regression analysis programs are used to calculate inter-element correction coefficients in order to correct for matrix effects.

3. When a beam of X-rays is directed at a mass of finely divided crystals, the resultant pattern may be recorded on photographic film. With the aid of a Strip Scanner and Hewlett-Packard Mini-Computer, these patterns are digitized and stored on a nine track magnetic tape. In creating this tape, noise is introduced and distorts the true nature of the patterns. Program DSPACE analyzes this tape, separating the true patterns from the noise and calculates from measurements on these patterns the lattice parameters from Bragg's Equation.

B. Mini-Computers

1. Hewlett Packard model No. 2114A (8K) with high speed reader and magnetic tape, on-line with Strip Scanner. It is primarily used to read plates produced by the laser probe in order to obtain qualitative and quantitative data. The computer positions the light spot on the plate, finds peaks as a function of wavelength, calculates the area under peaks and subtracts the background. The Stripscan can also be used to read diffraction films for mineral identification and calculate 2d spacings. This computer is used off-line, when time permits, for data reduction in X-ray fluorescence.

2. Hewlett Packard model No. 2100 (16K) with high speed reader, on-line with MAC electron microprobe. It is primarily used for data reduction of an energy dispersive spectrometer. Intensity spectra from a 1024 channel analyzer are read into the computer, the data are convoluted and peak intensities are calculated as a function of energy. The peak intensities are then corrected for background and non-resolution, compared to appropriate standards to obtain weight fractions which are then corrected for matrix effects. The results are then expressed in wt. % and/or mol. % if desired. The computer is also used on-line to process data from three wavelength spectrometers, i.e., measured intensities are converted to wt. % after background corrections. When time permits, the computer is used off-line for regression analyses, for data reduction in X-ray fluorescence and for data reduction of measured intensities from a second electron microprobe.

3. Nova model No. 1200 with high speed reader. Used to develop computer control systems. It is also available, part-time, for instruction and training.

4. Nova model No. 1220 on-line with direct reader optical spectrometer for the quantitative analysis of 22 elements using multiple regression methods. Used off-line when time permits for data reduction of analytical measurements.

Reference

Rucklidge, J., and Gasparrini, E. L.
1969: Specifications of a computer program for processing electron microprobe analytical data; EMPADR VII. Dept. Geology, Univ. Toronto, Toronto, Ontario.