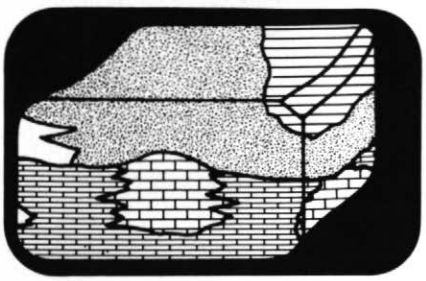

Computer Applications in Strategic and Sedimentary Geology: Notes from an Iconoclast

A. D. Miall
Institute of Sedimentary and Petroleum Geology,
Geological Survey of Canada
3303 - 33rd St. N.W.
Calgary, Alberta T2L 2A7

Introduction

What use a small admonitory finger raised against a mad stampede? Probably very little; but for those who are caught up in the impetuous rush to computerize everything in sight, here are a few doubts and warnings. The discussion is limited to a small part of the total spectrum of earth sciences for this is the part from which my experience derives, but specialists in other fields no doubt will suspect that there is a more general applicability in much that follows.

My main thesis is that computerization in geology is being overdone, that many of the advantages which are supposed to accrue from this process are in fact questionable, and that the immense costs entailed may not be justifiable. These opinions have developed over a period of approximately four years during fairly continuous usage of computers and computer applications both in a major oil company and in the Geological Survey of Canada (they are personal opinions and should not be read as either reflections or contradictions of GSC policy). I am more than willing to be refuted in this argument and am well aware that even if I am correct, very little is likely to change. Computers are certainly not going to go away, and there is no doubt that they are awesome beasts with which a scientist can impress colleagues, supervisors and grant-awarding bodies.

A Critique of the Well-Data File

As an example let us examine the well-data file. Many oil companies and several government agencies have established such computer-based files in recent years. These files, covering areas such as the western plains of Canada, contain millions of items of information including for each well, its name, location, completion date, name of operator, geological formation tops, drill stem test results, and possibly much more. The information is derived from many sources including operator's reports, commercial data services and in-house operations personnel. The file generally is stored on magnetic tape and may be accessed through time-sharing terminals using a special data-management computer language.

Why are such files created? Well, first they are obviously a "good thing". Here are all these dog-eared logs accumulating in a musty file room guarded by an old lady who wants only to knit. Sweep away the file room and replace it with what? A shiny piece of humming technology and a bevy of beautiful people to run it. Instant modernization.

But wait. The information is still the same old stuff. What are we going to use it for? Why do we need to be able to retrieve so much of it, so fast? These are very serious questions and are worth an honest answer, for data files of this type are extremely expensive to establish and maintain. The reasons given for having such a file include the following:

1) ease of housekeeping and updating
2) ease of retrieval of geological data for running through secondary programs to plot structure, isopach, lithofacies maps, fluid pressure maps and so on.

I have very little quarrel with the first of these new-found benefits. If the update routines have been properly thought out, and sufficient computer time and manpower are allocated so that updating can be carried out on a *routine* basis, there is no question that the computerized operation is more efficient. But this in itself is merely maintenance; it is not productive in the sense of enabling the operation to pay back something of the costs involved. The second supposed benefit is where the real advantages of a data file are thought to lie.

In order to discuss the second reason for setting up a well-data file it is necessary to examine what stratigraphic or sedimentary geologists actually do. Stratigraphic correlation forms a basic part of their work, and from it stems everything else, whether they are engaged in academic research or developing a hydrocarbon play. For this purpose a well-data file is of very little use to them, for the simple reason that many items of information will remain suspect until the users have checked the information themselves. All stratigraphers know that even in areas where structure and stratigraphy are simple, disputes arise over dating and correlation, and this problem becomes compounded when the operatives who assemble the data for a data file are poorly trained or inexperienced, as is commonly the case. Use several different sources for your file and the problems mushroom. Who decides whose data takes precedence in the file? The administrative problems, alone, are dismaying. It is necessary, therefore, for stratigraphers to do what they always did, and that is to plod through the raw information themselves and to develop their own ideas and conclusions. Unless they do this they are wasting their time - merely repeating what someone else did before. Admittedly, this in itself may be a useful function in giving geologists their first introduction to an area new to them, but the results will only be of passing interest if the intention is to proceed to serious research.

What about the much vaunted ability of the computer to take masses of information and make maps out of them? In mature areas such as the western plains of Canada all the important structure, isopach and lithofacies maps were drawn years ago by hand (see, for example, the Canadian Society of Petroleum Geologists publication *The Geological History of Western Canada*. Many more detailed maps have also been published from time to time in geological journals) and have long since been milked of all the information they can provide. The ability

to redo all this every day as each new well is inserted into the file will, I maintain, provide virtually nothing new or startling. If there was anything startling to be seen it would not need a computer to show it.

In an immature area such as the Arctic Islands there will be insufficient control points for the computer to be able to draw a useable map without much manual correction. And here is another point: that manual correction will probably be the most valuable thing on the map, for it will result from the geologist's experience, and will reflect his or her *ideas* about the area. In all fields of geological research, and especially a highly competitive field such as commercial hydrocarbon exploration, the generation of ideas which go beyond established fact is the most important contribution a geologist can make. Computerizing such an idea is not possible or necessary, for it hinges on such subtle features as an extra twist in a contour or the emphasis of a certain directional trend, which cannot readily be absorbed by most computerized plotting systems. It could be done using a CRT terminal (cathode-ray tube) and an "interactive graphics" program, which allows the operator to continuously modify the output, but it is difficult to see what is to be gained by such an approach.

The Computer as a Field Tool

Turn now to another rapidly growing area of computer applications. Many geologists are being encouraged to computerize their field data, even as they stand on the outcrop. Is this a good thing? Computerized field-data files vary in scope and purpose, but in general they are intended as data banks to which everyone working in a given area or on a given subject will contribute using standardized techniques for recording of observations. Retrievals may take the form of unedited field notes selected using keywords or pre-determined codes, or plots of such information as structural dip or mineralogical or geochemical data in map or graph form.

It is instructive to examine a recent publication in this field edited by Hutchison (1975), for there are many warnings to be gleaned from the writers who contributed to this volume. Amongst the lengthy technical descriptions explaining input and output and file structure are the following remarks:

> "Field data files may quickly grow and with large files we tend to slip into one of the major pitfalls, namely: the amount of energy required for file management, editing and updating a large file tends to consume so much time that more time is spent managing the data than attaining the ultimate objective" (Hutchison, *op cit.*, p. 5).
>
> "In geological field work, each region presents its special problems and demands an 'individual' approach" (Pipping, *op cit.*, p. 20).
>
> "A geologist's way of thinking in terms of interpolation and extrapolation with his data does not always agree with the logic of the computer, but one could hardly accuse the results of geological field work of being illogical" (Pipping, *op. cit.*, p. 20).
>
> "Historically, Canadian field geologists have been strong individualists . . . the production of a map is thus a highly personal task - with no one more qualified to make geological decisions about the project than the scientist in charge. There is thus no practical way to arbitrarily impose standards for the recording of field data" (Gordon, *op. cit.*, p. 30).

The consensus in Hutchison's (1975) publication seems to be that files which are set up to handle all the data input from a given area for all the geologists working in that area can only be generalized in scope. They can help geologists systemize their observations and their field notes (and this is certainly a beneficial effect) but it is doubtful if they add any new dimension to the results. Good research workers carry the critical items of information in their heads and should be familiar enough with their own field notes to have no need of a computerized retrieval system.

Is all the effort worth it? I suspect that a proper cost-benefit analysis would arrive at a negative conclusion.

Too Much Data

The large scale files I have been discussing contain far more data than an individual can ever use, and in an inexact science such as geology the data are subject to a vast potential for error, misinterpretation or irrelevance. No file is better than the data that go into it, and where the origins of the data are diverse the results are likely to be dubious indeed.

There are many other more general-purpose computer applications about which I am deeply suspicious for the same reasons. Who, for example, ever uses a keyword index to find a useful reference? Most keywords are impossibly generalized, and even if combinations of keywords are permissable in a given retrieval system, the result is likely to be a lengthy bibliographic printout containing many references of little or no use to the customer. A tedious editing and checking procedure then become necessary. Is it not equally efficient to go to a textbook or a recent paper in one's own field of interest and check the list of references at the end of it? The references retrieved in this manner already are sorted selectively, for the writers who chose them did so because the articles were known to be useful and relevant.

Computerized bibliographic services have been set up to keep subscribers informed of current publications, and the same comments apply. They provide far more information than a single individual could ever wish for. Far better to check regularly a dozen journals in one's specialty at the nearest library, making sure to see each new issue as it is published. Most libraries maintain (or have access to) perfectly good, old fashioned bibliographic systems that have provided adequate services for many years without the aid of any computer.

Except insofar as a musty file room can be replaced by a hygenic spool of tape, I have only a very limited use for large scale data banks. The "information blizzard" that the telephone companies love to warn us about in their commercials on television, is just that, a blizzard of meaningless digits. I am reluctant to believe those rosy predictions of inter-continental data networks. They will certainly work eventually, but only after millions of dollars have been spent on what the specialists like to disguise with the impressive sounding title "interfacing". This is what becomes necessary when Joe finds that Fred's data file was compiled with a different lay out, with the items in a different order, using different codes, on a different type of tape-drive. To make Fred's file compatible with Joe's an interface program has to be written, and up soars the computer budget.

The Computer at its Best

Let us not misunderstand one another here. I am no Luddite, a smasher of machines. There are numerous applications such as income tax record, payroll and accounting services, sending rockets to the moon, where computers are undeniably worth their weight in T5 slips and green cheese. I am concerned with the hazy edges of computer applications where it tends to be assumed automatically, and falsely, that to computerize is always the right thing to do.

The computer can do two things supremely well:

1) complex or lengthy numerical calculations
2) repetitive operations.

There are numerous areas of endeavour where the "housekeeping" abilities of the computer can be used to save much tedious work. Maintaining large data files, with continual updating capability, is certainly one such application, (although even here a properly maintained card index system may perform the same functions much more cheaply), but beyond this point somebody, sometime, should ask the question: *what is the file to be used for?* How much of the wonderful new facility for selctive and special-purpose data-retrieval will never be used because it is, in fact, useless?

In my view there are three secrets involved in making use of computers in geology. I believe they have rather general application:

1) keep the computer project small
2) keep it only for a specific, narrowly defined purpose
3) keep it individualized for a specific geologist.

The "specific, narrowly defined purpose" should, of course, be one which makes use of the special abilities of the computer, as noted above.

Files of numerical data which a geologist collects himself, such as grain size data or paleocurrent data, are ideal raw material for a computer project. Routine calculation of statistical parameters, automated plotting of the results in graph form, and the application of research statistical techniques such as regression analysis and factor analysis, are all readily accomplished, so long as the files are kept specialized and the geologist does not attempt to put everything he has into one massive master file. Numerous applications of this type have now been published. See, for example, the compilation by Gordon and Hutchison (1974), which contains short papers on many of the subjects in progress in the Geological Survey of Canada (including one paper by the writer).

Stratigraphic sections are boring things to write out and many attempts have been and are being made to computerize their description and reproduction. They are part of the "hazy edge" of computer applications that I referred to above. The snag is that a program generalized enough to satisfy every stratigrapher will not accomplish enough to be useful. Begin insertion of special items such as detailed sedimentological terms and the program will become so cumbersome nobody will want to use it. The secret is to write a program that is modular, and then each geologist can modify the modules to suit his own interests or the particular type of geology with which he is involved. Aitken and Carswell (1975) have attempted to construct a program of this type, with promising results.

Simulation studies are a very exciting area of computer application. In fact they were virtually impossible until the computer was developed because of the necessity for large amounts of mathematical computation. The purpose of such studies is to imitate reality and, in the process of developing the imitation, to learn more about the reality itself. Harbaugh and Bonham-Carter (1970) provide some excellent examples of the uses of simulation methods in stratigraphic and sedimentary geology.

It helps greatly if the geologist can carry out much of his or her own program development. Programming is often scorned by the researcher as just a technical chore to be done by a tame assistant. But it is, or can be, a highly creative activity giving the geologist great flexibility in the testing of new ideas without the problem of a communication gap between himself and a programmer with little or no geological training. For this reason I strongly recommend that geologists learn FORTRAN while still in their infancy. Like all other foreign languages, it is much more easily absorbed while the student is young and impressionable.

Conclusions

Most original geology is a very *individual* enterprise. All geologists follow, to a greater or lesser extent, their particular interests and talents. Their data collecting will be highly personalized and so will be their results and conclusions. This is how the science progresses, and it is for this reason that their use of the computer should be run in the same personalized way. Many geologists may not agree with all (or any) of my conclusions, and if so I would be interested to hear about it. My concern is only that computers are approached nowadays like the subject of plate tectonics: with boundless enthusiasm and an assumption that their invocation will automatically solve all problems, regardless of the nature of the project in hand.

Another danger of the computer is that it exposes the ignorant to a lot of high powered statistics. I watched once, while a geologist I know took a structurally simple piece of Alberta and subjected it to a fourth degree trend surface analysis. He coloured in all the positive residuals because, as an oil geologist, he had been encouraged to look for anticlines. I could go on, but this is really another story . . .

References

Aitken, J. D., and D. A. Carswell, 1975, Computer-compiled descriptions of stratigraphic sections: Geol. Surv. Can. Paper 75-1, Part B.

Gordon, T., and W. W. Hutchison, eds., 1974, Computer use in projects of the Geological Survey of Canada: Geol. Surv. Can. Paper 74-60.

Harbaugh, J. W., and G. Bonham-Carter, 1970, Computer Simulation in Geology: Wiley Interscience, 575 p.

Hutchison, W. W., ed., 1975, Computer-based systems for geological field data: Geol. Surv. Can. Paper 74-63.

MS received June 23, 1975.

Discussion (1)

H. R. Wynne-Edwards
Department of Geological Sciences
University of British Columbia
Vancouver, B.C. V6T 1W5

Dr. Miall's article is not really as revolutionary (reactionary?) as his introduction and conclusions would suggest. In fact it issues some welcome warning about the costs and dangers of mindless generation of computer-processible files. He rightly points out that one should have an objective in creating a file, and be satisfied beforehand that that objective is realizable. This means that most computer-processible files will be somewhat specialized and somewhat personal things. Most of the successful applications have been of this sort, and probably will continue to be so.

The prospect that appears to frighten Dr. Miall the most is the wholesale generation of computer files to replace hard copy files. One can argue, however, that if the files exist and are worth maintaining, even if only by "an old lady who wants only to knit" (Committee on the Status of Women please note), they would be much better off in the flexible, accessible, and compact storage that a computer system offers. One great misconception is that the costs of computing are horrendous; the costs of the project are, but most of them involve the manpower necessary to generate and maintain the data base. If the files are assembled in a computer-processible form at the outset by the people most qualified to generate them, the cost of computer storage and the cost of file storage are not really that different. If the system is properly organized periodic printouts arranged in various convenient formats will look after the ordinary retrievals, and specific research projects based on the file can gain the necessary data much more quickly by simple programming than by hunting through a file room.

The costs of computerizing existing files and logs, particularly those that are already thoroughly processed and analyzed anyway, are probably not warranted. Nevertheless, computers are here to stay and, as the author states, represent an extraordinary convenience. There comes a point when any new projects involving the analysis of large volumes of data should routinely be compiled in computer-processible format. In my view that point was reached about 1968.

The author overlooks a great number of modelling and analytical studies of a fundamental and worthwhile character and *only possible* with the help of a computer. Great advances in both precision and insight have already come from studies of this sort in many fields. The conceptual analyses of the world situation offered by Reports I and II to the Club of Rome, for example, have an elegance and perceptiveness quite unapproachable by non-computerized means. They are a welcome alternative to hand-wringing over "everything being somehow dependent on everything else." To say that we would be better off without those new-fangled computers is a point of view doubtless shared by flat-earth societies, back-to-the-bush communes, and others seeking to return to the Stone Age, but is hardly one to be endorsed by the scientific community.

MS received August 19, 1975.

Discussion (2)

C. F. Burk, Jr.
Canada Centre for Geoscience Data
580 Booth Street
Ottawa, Ontario K1A 0E4

Dr. Miall's view that the excessive and unwarranted pressure to "computerize everything in sight" has raised a problem of "too much data", belies the fact that only a few per cent of geological data are currently available in computer-processable form on a practical and timely basis. Far from being "overdone", the application of computer technology to the management of geological data is only just beginning in earnest. There are serious problems to be overcome (a few of the lesser ones were raised by Miall) and the process will not be painless and without frustrations, but its potential significance as an important tool for the management of non-renewable earth resources is such that many provincial and national governments are now active in computer-oriented information and data management. Many companies in the private sector have long since developed and benefited from this approach.

In proportion to the volumes generated at the tax-payer's expense, the amount of geological data currently available in computer-processable form (please, not "computerized" data!) is minor. A study of this question by the Canada Centre for Geoscience Data in 1974 revealed that the Department of Energy, Mines and Resources (to which both Miall and I belong) had only 11 fully operational data files; nevertheless, it is clear there is no imminent danger of becoming blinded by a blizzard of digits - meaningless or otherwise. In the present context, it is interesting that not one of the 11 files is the result of, or directly applicable to, stratigraphic or sedimentologic research. Where are all the computer-processable data implied by Miall's analysis?

As one formerly active as a stratigrapher and petroleum geologist, I am tempted to comment in detail on Miall's critique of well-data files and applications thereto. However, this topic should be left to those who are currently active, since the technology has advanced greatly during the past seven years. On a more philosophic bent, his discussion on "what stratigraphic or sedimentary geologists actually do"

supports my view that there are several fields of geoscience wherein some practitioners chose to live in a world without reproducible data - the resulting work thus falls outside the realm of science.

As for computer-based field data systems, readers should consult the complete papers from which Miall's quotations were excerpted (Hutchison, 1975), where the disadvantages and pitfalls he cites are balanced by descriptions of the benefits and utility of these techniques. The papers indicate that, contrary to Miall's implication, field data systems have been and should be, applied on a *project* basis; they should not be used simply to provide fodder for an external "data bank". Each system should be evaluated solely in terms of meeting the project objective; there is of course no basis for criticism just because computers were used!

In reply to the question "Who, for example, ever uses a [computer-based] keyword index to find a useful reference?", I would say: Anyone who wishes to search a definable bloc of the literature of geoscience in a practical manner. For all the difficulties of using computer-based bibliographic systems, as outlined by Miall, it is inconceivably naive to suppose that this task can be performed better by consulting a textbook or paper (even if one could know which to consult), or using a manual "old-fashioned bibliographic system" (even if one could have access to an international, complete library). My best estimate on the number of papers published on Canadian geoscience is in the order of 250,000 (G. Lea, report to Canada Centre for Geoscience Data, 1970); for worldwide geoscience Lea estimates that about 100,000 papers are published *per year*. Would anyone like to use a set of edge-punched cards and some *long* needles?

Reference

Hutchison, W. W., ed., 1975, Computer-based systems for geological field data: Can. Centre Geoscience Data, Geol. Survey Can. Paper 74-63, 100 p.

MS received September 5, 1975.