The Use of G-EXEC for Resource Analysis[1]

K. G. Jeffery[2] and E. M. Gill[2]

A brief description is given of the G-EXEC data handling system. Its suitability for resource data handling is noted. Several examples of its use for resource data are described, including sand and gravel, limestone, and minerals. **KEY WORDS**: data systems, data processing, data files.

INTRODUCTION

Resource assessment, and allied fields of resource accounting and management, form a large and complex subject area. Problems tractable by human understanding are of limited size and must be derived from the total subject by reduction. This reduction requires data selection to produce a relevant subset of the total data available. It also requires an analysis of the problem into its component subtasks, and the treatment of the subset data by subtask processes which mimic the equivalent human processing operations.

It is suggested that a computer system which provides these problem reduction facilities is likely to be efficient and useful. G-EXEC has been used in this way in the field of resource data handling.

G-EXEC is a machine-independent data handling system in daily use for many kinds of data including data for resource analysis. It provides database management, and many other facilities including several kinds of statistical treatments, display in several forms, file transformations, sorts and merges, and retrieval from a single file or groups of linked files. All files in the system are of a standard form with a data description at the beginning, which allows any program to act on any data file. User access to the system is by a set of simple commands:

EXEC defines a process to be carried out:

   e.g., EXEC GRETSB calls for single-file retrieval

FIND names a file to be acted on:

e.g., FIND CU401999 COURSE A DATASET

MAKE names a file to be loaded or overwritten

e.g., MAKE CU401998 NEW FILE

Strings of these commands may be linked together to perform complicated processes, as shown in the next section.

The translation of the user's commands is carried out by a program called the Executive Controller, which accesses tables of the files and programs available in order to form a FORTRAN program and Job Control Language corresponding to the processes and files required. This Executive Controller is the interface between G-EXEC and the operating system of the host computer, and allows the maintenance of a constant user interface independent of machine or installation. The Controller is partly machine-dependent and is usually installed by the system engineers. The program library, on the other hand, is written in FORTRAN IV and may be added to or amended by the installation's programmers as required, within certain system standards.

PROBLEM-SOLVING WITH G-EXEC

Problems in resource data handling, as with any other data, are handled in G-EXEC by splitting the overall task into its logical component subtasks. Many subtasks are common to the succession of processing steps for many different kinds of data. Thus, the software for most subtasks is used many times for handling data from different subject disciplines. Furthermore, the structured and modular software design allows a programmer to build software for a new subtask very quickly—for example to perform calculations for a new method of assessing the quality of a particular resource.

To illustrate the general methodology described above, a simple example from the field of geochemistry is worked out.

Problem

It has been discovered that metal A, which was previously only found in a compound mineral which was difficult to smelt, occurs in an easily handled compound with metal B. However, the separation process requires abundant fresh water. Find areas in country Z which are possible sites for mining and separation operations. Data available:

1 file of geological field data for country Z

A file of geochemical analyses for samples equivalent to field sites in the first file

Solution

Merge the two files to produce a composite file with one recorded for each site. Link is on sample number	EXEC GSUBST,,,45,200 001 SAMPNUMB FIND NX001CHE CHEMISTRY OF Z FIND NX001FLD FIELDDATA OF Z MAKE TEMPFILE,,,45,200 .
Select samples with high values of A and B and also with streams of large volume defined by stream code value 7 or above. Also check that data all in country Z	EXEC GRETSB 000 FIND TEMPFILE WHEN A .GE. 50 .LE. 10000 WHEN B .GE. 20 .LE. 5000 WHEN STRMCODE.GE.7 .LE. 10 WORD COUNTRY .EQ. Z HOLD XCOORDTE, YCOORDTE, A ,B END MAKE TEMPFILE,,,4,200
Produce a lineprinter map of sites with a coded magnitude value for A at each	EXEC GRID,,,4,200 000 1 2 3 FIND TEMPFILE

This procedure would establish the position and magnitude of any potential resources of A, and lead the way for a fuller investigation. An exactly similar procedure could be carried out for metals M and N in country Y by using the same methodology (but changing the names where appropriate).

G-EXEC is designed to help the user to approach problems of resource assessment from a systematic viewpoint; the user maintains control of the system but the software performs the routine and tedious tasks required to solve the particular problem.

EXAMPLES OF THE USE OF G-EXEC FOR RESOURCE DATA HANDLING

The Institute of Geological Sciences uses G-EXEC for many of its data handling and analysis needs, among which are several resource analysis projects. These projects are all independent of one another but use the same programs for data banking and analysis, so that the files are compatible with each other. In several cases hierarchies of files have been set up for a project, linked by a locational reference of some sort. If at some future date it becomes

E

necessary to link data from different projects, these locational references will provide a primary key for linking. Nearly all the analysis of each resource is done by the scientists themselves, using the simple G-EXEC commands— the assistance of trained personnel is only required to ensure that the files are set up initially in a logical manner. This approach ensures that users understand what is being done with their data and can take new lines of research as seem appropriate, while the programming staff can concentrate on developing new applications software if the project is very specialized.

A brief description is given in the following pages of some of the current resource-based projects using G-EXEC in one branch of the IGS. Two of these projects, MINTMINES and BRITPITS, are being set up by the Mineral Intelligence Unit and contain information on economic mineral deposits. MINTMINES covers the whole world while BRITPITS is confined to quarries within the United Kingdom. The third data bank, and at present the most advanced, is MINSTATS, a data bank to provide production and trade statistics on an annual basis for mineral commodities throughout the world; this is being developed by the Mineral Statistics and Economics Unit. The Mineral Assessment Unit is currently carrying out surveys into sand and gravel resources and limestone resources in certain areas of UK. The Radioactive and Metalliferous Minerals Unit has a continuing program of geochemical reconnaissance in several areas of Britain.

MINTMINES

MINTMINES (Mineral Intelligence Unit—Mines Data Bank) is a computer-based storage and retrieval system presently under development by the Institute of Geological Sciences. It is designed to operate within the G-EXEC system.

MINTMINES offers a simple method of storing data relating to mining operations throughout the world concerned with the exploitation of metalliferous and industrial minerals. The system could be extended to cover coal mines eventually but would not be suitable for handling data on petroleum production. Although MINTMINES is primarily intended to deal with information concerning active mines, it is possible to incorporate data on specific mineral deposits which, although not being currently mined, are of economic interest.

First-Level Data

Basic data on mines which are to be recorded on MINTMINES are entered on a first-level master file. Each mine is uniquely identified within this file by three alphanumeric fields recording the country in which the mine is located, the commodity produced, and a locality reference. The latter can be

the name of the mine itself, or the place at which it is located, or its latitude and longitude coordinates. Additional alphanumeric fields are provided on this file for basic details of geology and mine ownership. If a mine produces more than one commodity, records are made under each commodity.

The three fields which uniquely describe each record form a keyfield which is also incorporated in each record on second-level data files. Thus by index-linking the two files, second-level data can be accessed via the top-level file as well as directly.

Second-Level Data

In the present initial phase of development of MINTMINES only one second-level file is being built up. This will cover numeric information for the years 1975 to 1979 relating to crude ore treated, production of metal or concentrate, grade of concentrate, tonnage and tenor of ore reserves. In addition, space is provided for a free text note on the reliability of the data entered and any other remarks that are appropriate. It is intended that collection of data for this file should be an on-going exercise. Quinquennial files are best suited to the data processing facilities available and a new file for 1980 to 1984 will be created when necessary. Retrospective files for quinquennia prior to 1975 could be created in response to any specific demand.

Other index-linked second-level files that are envisaged in future developments will also be linked by the same three fields as key. These files might include details of company affiliations, further information on the geology, and references to the literature.

The MINTMINES system is user-orientated in that the input is designed so as to require little formal training and can provide flexibility of storage, retrieval and statistical methods.

BRITPITS

BRITPITS is a data bank using the G-EXEC system designed to provide data on United Kingdom mines and quarries. It will contain information on the type of "mineral" extracted, the name of the quarry or mine, its location, and the geology of the deposit worked.

Location by national grid reference will be the unique identifier for each entry and will form the keyfield for purposes of alterations and linkage to any subsequent files. Further index-linked second-level files will contain such information as the size and thickness and bulk physical–chemical properties of the deposit exploited and the end use application of the minerals produced. Retrieval will be possible by any of the field names or any combination of these. Graphic plotting facilities will enable the data to be printed at varying scales as overlays for geological and regional planning maps.

MINSTATS

The Statistics and Economics Unit of the Mineral Resources Division of IGS is using the G-EXEC system to transfer their current records of world mineral production and trade into a computer data bank called MINSTATS. The statistics to be stored cover annual series of figures for production, exports, and imports of approximately 70 mineral commodities throughout the world. The figures are up-dated every year. The data-base is relational so that each record of a series of annual figures is described by four alphabetic fields: commodity, type of data (i.e., exports, imports, or production), country name, and subcommodity. This unique description serves as the keyfield. The 'commodity/type' fields are sorted alphabetically and numeric codes are attached to the 'country' field in order to sort within 'commodity-type' down to 'subcommodity' and in order to sort within 'country' into a predetermined order. Retrieval is effected using any field within the record to access data for any particular year or range of years.

At present the data-bank comprises about 12,500 records covering the years 1966–1970 and sorting codes have been attached to about 3000 records.

Apart from the general objective of more efficient management of this large number of statistics it is hoped to use MINSTATS as a basis for an annual publication of a five-year series printed direct from magnetic tape and for performing analytical statistics on the data held in store using G-EXEC process programs. Specifically, forecasts based on regression analyses of historical production would complement analyses of reserves data held in MINTMINES.

SAND AND GRAVEL SURVEY SYSTEM

The purpose of the sand and gravel survey system is to estimate the available resources of sand and gravel in UK. The overall system utilizes the G-EXEC data handling system so that the users can easily obtain the required scientific output from their own data.

The detailed purpose of the survey includes the calculation of grade percentages according to several schemes for mineral samples, discrete deposits, and for the total mineral content of the borehole. The sediments are then classified according to the scheme of Folk and also in the Willman Institute of Geological Sciences (IGS) Mineral Assessment Unit classification. Graphical representations of the mineral distribution and type in each borehole are also produced. Future plans include automated volume estimation and computation of means for selected resource blocks.

The input to the system reflects closely the sheets used in the original manual system. There are two separate documents which are used for punch-

ing and the punched cards produced from each document are stored in separate G-STAR files in the G-EXEC system. The first document contains data of reference type and data pertaining to the borehole—these include location, surface level, data, drilling contractor, and drilling information. These data, when punched, are validated and loaded to a file in the system.

Each record in this reference file refers to several records representing units of lithology down a borehole, and records representing each interval form the content of the second file, the data for which are punched from completed second documents. The fields in this file refer to the position of the sample in the borehole and the cumulative-percentage-data of percent of the sample passing each of the sieve sizes.

For each mineral sample (each record in the second file), grade percentages for grouped class intervals are calculated. The percentages of fines, sand, and gravel are computed and the sample is classified by the Folk and Willman/IGS Mineral Assessment Unit schemes.

For each mineral deposit within each borehole mean cumulative grading percentages are calculated, while for each borehole the mean cumulative percentages for the total thickness of mineral deposits in the borehole are computed. A summary log of each borehole is produced, based on mineral deposits rather than on individual samples. Finally, in a separate step, punched cards may be produced for further processing in other systems.

LIMESTONE QUALITY ANALYSIS

A survey is being carried out on the limestone resources of an area in Britain, using G-EXEC for the databanking and analysis. The input data are initially in eight separate files, which are linked by an external index using a composite keyfield of quarter-sheet number, borehole number, and depth in borehole. The files are on two hierarchic levels, the higher level consisting of one file containing site reference data about each borehole. Below this, one file gives for each borehole the depth of top and bottom and average strength. At regular intervals down each borehole various properties have been measured, divided into six files for convenience of input but merged at a later stage. Chemical measurements include the basic limestone parameters (CaO, MgO, SO_3, Na_2O, K_2O) and heavy mineral content (SrO, P_2O_5, Cu, Pb, Zn, MnO, Fe_2O_3, As). Minerals, fossil occurrences, color and etched color, and lithology are also recorded for each depth.

The analysis of this data consists largely of finding overall definitions of each area and facies type, in terms of quality parameters such as absence of impurities and strength. The relative abundance of each type is also investigated using contour plots and simple point plots.

The types and area of limestone are selected by sorting the file or files

on a particular aspect such as area, formation, or facies. Having grouped together all the records pertaining to this unit, these are then treated in one or more ways to illustrate their properties. One approach calculates the percentage of each lithology represented in the unit, or the percentage of each improvement of the insoluble residue. Another technique involves summing simple statistics on each unit (facies, formation, area, or other grouping) and using the results of these as definitions of the unit properties. Point-plot maps of chemical components and contour maps of unit thickness are also useful.

This arrangement is a typical case of user interaction with his data via simple commands, where the delay involved in writing a new program for each application would be unacceptable. This application also illustrates the necessity for files to be processable in many different ways, including some not foreseen at the time of setting up the file.

GEOCHEMICAL RECONNAISSANCE

The Radioactive and Metalliferous Minerals Unit is currently carrying out surveys, for other Government departments, in various parts of the country. The materials of interest are usually heavy metals. Data is collected from stream sediments and soils, and after rigorous checks which include division of samples and randomization of sample order, chemical analyses of several types are done. The output from the analysis equipment is largely automatic, and complete analyses from a given sample location are unlikely to be available until part way through the project. Four or five different chemical data files are built up as the data become available, and are merged as the analyses are finished. Often one or two of each type are missing so that the merged file is a superset of the individual files. Field data and locations are then merged with the chemistry, and several kinds of maps produced, some by systems outside G-EXEC, to which interfaces are provided. The distribution of element populations, and interrelationships between them, are investigated using a large number of statistical techniques all available within G-EXEC.

ACKNOWLEDGMENT