

UNITED STATES DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

**An overview, with retrieval request information,
of the Rock Analysis Storage System (RASS),
a database maintained by the
U.S. Geological Survey**

By
C. M. McDougal

Open-File Report 84-653

This report is preliminary and has not been reviewed for conformity with U.S. Geological Survey editorial standards. Any use of trade names is for descriptive purposes only and does not imply endorsement by the USGS.

1984

The Rock Analysis Storage System (RASS) was developed in 1968 as a database to store geochemical and related data received and generated by the Branch of Exploration Research, which is now the Branch of Exploration Geochemistry (BEG). These data are acquired from geologic information and coordinates provided by geologists, and from analyses performed by spectrographers and chemists.

One of the advantages of RASS over other data storage systems is the capability to permanently store large amounts of geochemical data and to rapidly access the data by computer. This rapid access allows Survey scientists to use the data to evaluate new techniques and new sampling media, and to utilize modeling for ore deposits and/or geochemical processes. Ready retrieval of the RASS data from previously completed studies assists in the avoidance of sample duplication.

Sample media, for which data are in RASS, consist of rocks, unconsolidated sediments, heavy-mineral concentrates, soils, waters, plants, and other organics. Sample media are identified and classified in RASS through the geologic coding provided by the submitter.

The chemical components of the samples are identified through laboratory analysis. This information is included in the RASS database.

The RASS database contains information for 586,912 samples. This information is stored in 273 files. The smallest RASS file contains data from 11 samples, while the largest file contains data from 14,900 geochemical samples. The average file holds information from approximately 2,000 samples. Of the 273 files, 216 files contain latitude and longitude, 51 files are stored with X-Y coordinates, and six files are stored with no coordinates. (BEG policy now requires coordinates prior to RASS entry.) These files encompass data from 29 states as well as topical, miscellaneous, and foreign studies. One hundred and ten of these files, 30 percent of the database, contain data from samples collected in wilderness studies. Twenty-seven percent of the data is from samples collected in the state of Alaska.

Each file may contain several jobs. Jobs numbers are assigned to batches of samples when they are received by the sample control unit. Job numbers are used primarily to monitor the progress of the samples through the analytical procedures. A file of 11 samples could conceivably have 11 job numbers, or it could contain only one job number.

Files are now maintained along program and/or project lines. If a geographical area for a project is quite large and extensive sampling is anticipated, a file is established. If the area is small or sparsely sampled, the jobs may be stored in a larger composite file identified by program and geographical area (e.g., BLM, Southwest Wilderness).

Various methods of chemical analysis have various biases and limits of detection. RASS allows for differentiation of these methods. For example, mercury can be analyzed by use of the mercury detector, atomic absorption, or colorimetry. RASS allows three different variables for the three different methods, rather than combining all three methods into one field identified as mercury. There are 643 variables in the RASS database of which 340 are presently active. Figure 1 demonstrates the types of variables in RASS, and lists the total files and samples in RASS. All elements analyzed by various methods follow the same scheme as mercury. The methods used are noted by the chemist on the chemical report forms for data entry as well as other purposes.

The data entry personnel scan all reports prior to entering data. The report sheets and submittal forms are inspected for legibility; blanks in coding or in the analytical report forms are double checked; and analytical results are examined for agreement with standard reporting format and standard reporting intervals. Then, using the information provided by the analyst, the data-transcriber assigns the variable to be used for computer storage of the analytical data.

The data are entered via microcomputer, processed through the U.S. Geological Survey's computer system, and updated to the RASS database. Historically, there is less than a one-percent error rate on data entry.

The ability to recover information rapidly from previous studies is one of the prime assets of RASS. Retrievals may be done on various parameters or combinations of parameters. There are several factors to balance when requesting a retrieval. Detailed retrieval requests are more likely to achieve desired results. However, as more criteria are used to accomplish a retrieval, the cost increases for computer usage and for human resources used in completing the retrieval.

Retrieval cost is also affected by the number of samples in a file. It will take less computer time to search a small file than to retrieve from a larger file. Figure 2 illustrates criteria used for RASS or STATPAC retrievals.

The information in RASS pre-dates many of the data entry and BEG policies now in effect. In many instances, coding was insufficient or improper on the jobs initially entered in RASS. This has made it difficult to fulfill retrieval requests unless the information is requested by tag number. Similarly, retrievals made on coordinates may only be done on files where latitude and longitude were provided. Retrievals may be made on any geologic variable, if coding was provided and entered. Coding should be sufficiently detailed to allow for retrieval on the criteria needed for total identification of the sample.

Retrievals are accomplished using a computer program entitled "b860." Retrievals may create a STATPAC file (a file used for statistical manipulation) or a RASS file (a copy of the file stored in the RASS data base). The control cards for executing this retrieval program can be created using the prompting program "rass header." These programs require a file which can be obtained from the RASS Management personnel prior to creating the control cards. To obtain this file, send a written request to the RASS personnel stating program name, project name, and tag numbers. You will be contacted, when access for the file has been accomplished.

STATPAC retrievals will not be done by RASS personnel for requestors who have computer expertise. However, if a STATPAC retrieval is requested, a written request should be submitted to the RASS personnel. This request should state the program name, project name, tag numbers, and other pertinent information. Figure 3 is an example of a STATPAC retrieval request.

Reference

VanTrump, George, Jr., and Miesch, A. T., 1977, The U.S. Geological Survey RASS-STATPAC system for management and statistical reduction of geochemical data: Computers and Geosciences, v. 3, p. 475-488.

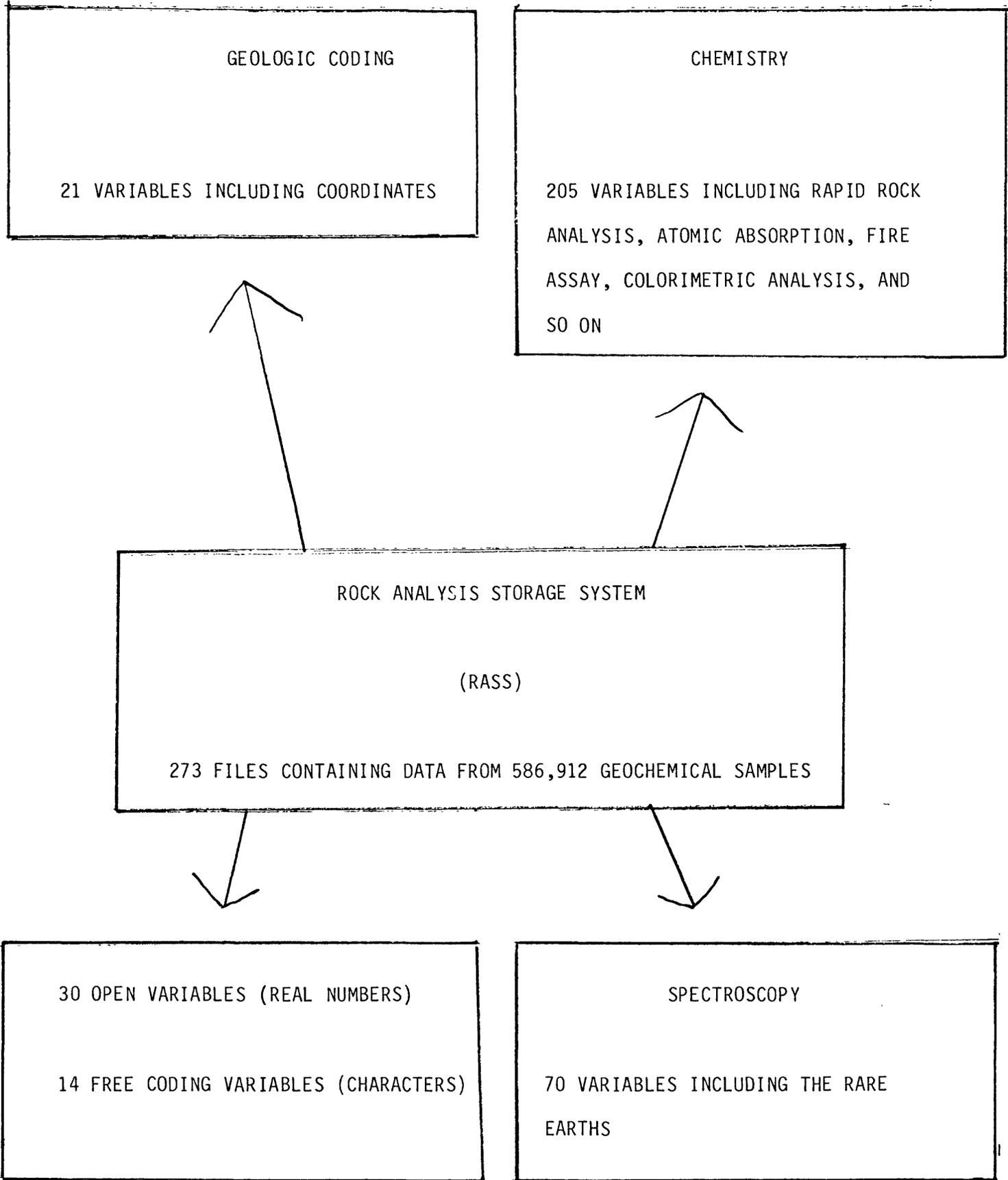
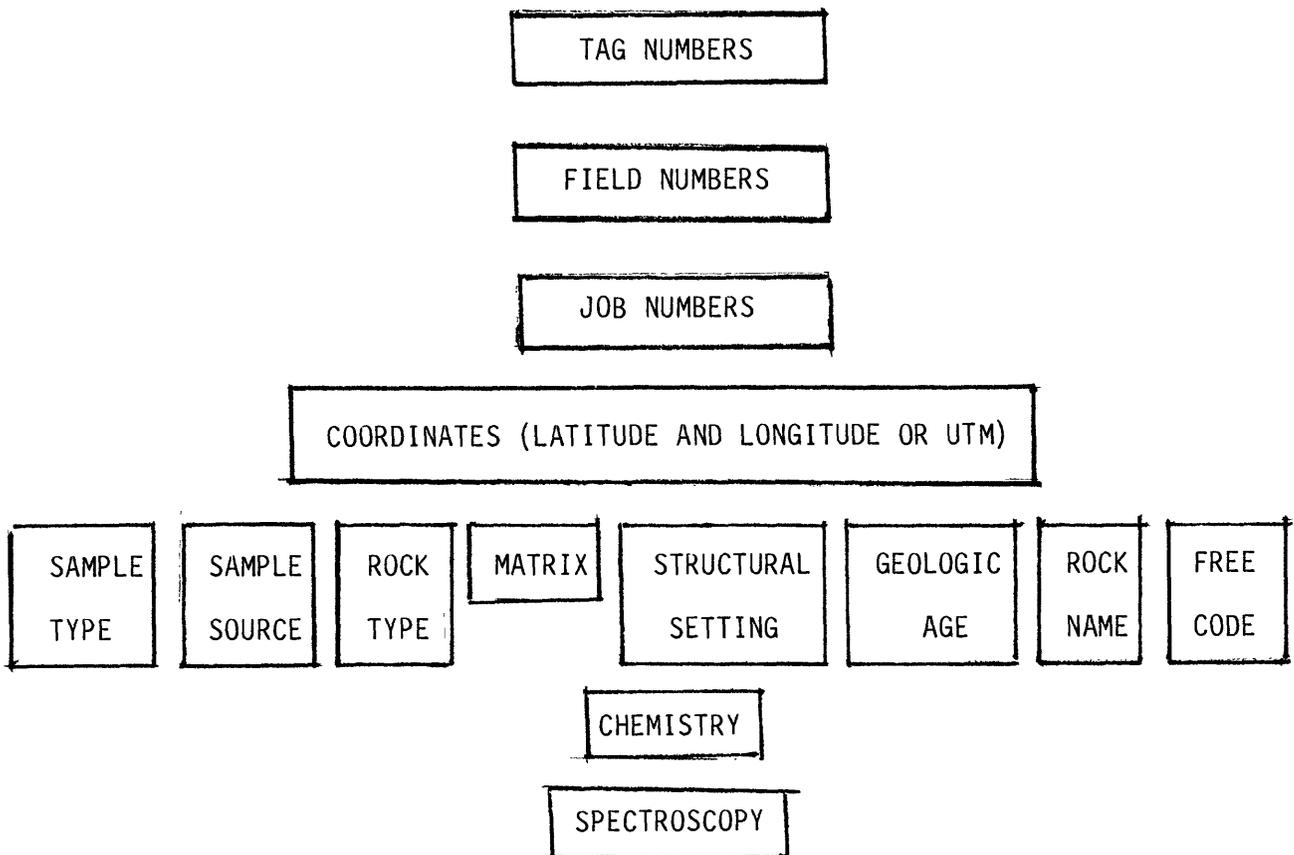


Figure 1.--RASS variables, number of files and number of samples.



ADDITIONAL NOTE

Combinations of various criteria are acceptable retrieval requests (e.g. all rocks within stated coordinates and spec Au values >500).

Figure 2.--Retrieval criteria.

TO : RASS Personnel Feb. 29, 1984

FROM : XYZ SMITH

SUBJECT : STATPAC Retrieval of BLM Lost-in-the-Woods Wilderness, Calif.

Please retrieve a STATPAC file for tag numbers:

AAA001-095

BBB093-100

CCC150-171

There should be three data sets in the final file. Rocks were coded "A", stream sediments were coded "B", and soils were coded "D" under your sample submittal sheet column entitled "Sample Type."

Retrieve only the samples within the coordinates 30°25'N, 119°44'W; 38°44'N, 120°13'W. Do not retrieve atomic absorption results on the soil data set.

Please notify me when access to this information is available.

Figure 3.--Example of STATPAC retrieval request.