UC San Diego
SIO Reference

Title
Sediment Data Bank Users Handbook

Permalink
https://escholarship.org/uc/item/0mg5s34f

Journal
SIO Reference, 78(10)

Authors
Frazer, J. Z.
Fisk, M. B.

Publication Date
1978-05-01

2 out of 2

SIO REFERENCE SERIES

SEDIMENT DATA BANK USERS HANDBOOK

J.Z. Frazer

M. B. Fisk

SEDIMENT DATA BANK USERS HANDBOOK

by

J. Z. Frazer and M. B. Fisk

Scripps Institution of Oceanography
University of California, San Diego
La Jolla, California

PREFACE

This report contains a complete description of the SIO Sediment Data Bank and instructions for retrieving selected data. Sections A and B provide a brief history of the data bank and an overview of the system. Later sections contain more detailed information for the data bank user.

Section C describes data collection. Section D includes instructions for retrieving information with the interactive SEARCH program. This program is available to users with remote terminals and enables them to search the data files at will according to any specified criteria. The information in this section would also be useful to other users, since it describes precisely what information is stored and lists the available search options. Sections E and F explain the use of additional computer programs for displaying and manipulating the selected data. The figures in these last sections provide examples of different types of output. The remainder of the material will be useful primarily to those who access the data bank directly.

SIO Reference No. 78-9, Sediment Data Bank Coding Instructions, is a companion volume to this report. Besides complete information on how data are coded for storage, it contains an explanation of how our system of classifying sediment lithologies relates to other systems used now and in the past.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

A. INTRODUCTION

In 1912, John Murray could state, "Nearly all the samples of deposits procured from deep water over the ocean's floor have passed through our hands."[1] There is probably no living scientist who has examined even every sample in his own institution's archives. With seafloor sampling being conducted at a continually increasing scale by more and more institutions throughout the world, the marine scientist working today need not be limited to studies of material he has collected and analyzed himself. Given the large costs of retrieving seafloor samples and the additional effort that has gone into describing, analyzing and storing the material, it is desirable that samples and information about them be made widely available. Such sharing of information cannot take place easily if each scientist seeking information must communicate directly with each possible information source.

To alleviate this difficulty, a number of data banks have been established to collect certain types of information in a central location. Although information can be collected, stored and distributed without the aid of computers, as the collection grows it becomes increasingly time-consuming for humans to search it for specific information. Computers certainly speed up the data handling and lower the expenses. And because computers can provide data cheaply in a variety of display formats (maps or graphs as well as printed listings), information can often be obtained in a more useful form than it was in the original source.

The SIO Sediment Data Bank was set up in 1967 to provide computer access to information about (1) the location of stations in the deep ocean where sediment samples have been collected, (2) the general character of the sediment, and (3) bibliographic references to sources of more complete information about the samples. At an early stage a file of manganese nodule elemental analyses was added.

In 1973, a new disk-based system for storage and retrieval of sediment data was designed to handle more complete information about sediments and manganese nodules and to make the data immediately accessible from remote terminals linked by telephone to the computer. With our interactive mode computer programs the data files can be searched according to any parameter or combination of parameters stored, and selected information can be listed, plotted on maps (in any scale and various projections), or saved as disk or magnetic tape files for further use. The files in the new system presently contain records for more than 57,000 stations, 4,000 nodule analyses, and sediment descriptions for more than 16,000 stations. All data are in the public domain.

The data have been obtained from both the published literature and the unpublished reports of various government agencies, oceanographic research institutions and private companies, and from personal communications.

---

[1]John Murray and Johan Hjort, The Depths of the Ocean (London: MacMillan and Co., Ltd., 1912), p. 144.

Up to about 1955, when the oceans were relatively unknown, the purpose of most cruises was general exploration and sampling. Lamont-Doherty Geological Observatory is now the only U.S. institution which pursues a program of evenly sampling the seafloor. Most samples have been collected for some particular purpose: one cruise may be designed to retrieve samples which would shed light on the carbonate compensation depth, another cruise may be aimed at determining sediment age, while a third cruise may prospect for economic deposits of manganese nodules. Even if these three hypothetical cruises covered the same areas, it is unlikely that the types of observations and measurements made on the recovered samples would be identical.

Since different classification schemes or models of geologic processes are accepted in different times and places, it cannot even be taken for granted that scientists studying the same problem have analyzed and described their samples in the same manner. Further, similar measurements are not always made with the same degree of accuracy or precision.

For these reasons, it can be argued that a collection of information from a wide variety of sources, such as the Sediment Data Bank, is virtually useless. It is clear, however, that such a data collection can serve many purposes. It is wasteful of both time and money to collect new samples when the type of material sought is already stored in the core locker of some institution. Obtaining information from individual institutions regarding samples available from a particular area or of a specified type can take weeks or months, whereas accessing a central data bank from a remote terminal takes minutes.

When a cruise is to be undertaken, it is useful to have some prior knowledge about sediment type in the area to be surveyed. Even the briefest qualitative sediment descriptions can often provide the information that is needed.

Most oceanographic institutions routinely describe the samples they archive to enable their own researchers and others to select samples for study. That each institution uses its own sediment classification scheme and presents information in varying detail makes such descriptions somewhat less useful than they might be, but it does not make them useless. If properly done, collecting these heterogeneous data in one central data bank can preserve a substantial fraction of their original value.

Although the Sediment Data Bank as a whole is not a consistent body of information, one can select subsets of data which do form such consistent bodies: for example, all analyses performed by a single laboratory or all core descriptions from a single institution. Consistency is not always a virtue, however, since an investigator can be wrong. When data come from a number of sources, errors tend to be random and cancel each other.

On the other hand, processing information through a computer and storing it on magnetic tape cannot make it more accurate or more complete than it was before. The most sophisticated data coding scheme or search program cannot produce information that is not present in the original source.

The Sediment Data Bank is analogous to a library, not to a finished piece of research. Not all items of published literature are equally reliable, and unpublished data are likely to be less reliable than data from published sources. The data bank user is advised to note the reference number by which each data record is identified. For a particular project one may wish to select only data from certain authors or analyses performed by certain methods. The SEARCH program used for data retrieval (see Section D) allows for such selectivity.

Present users include scientists from industry, academic institutions and government agencies. The following are examples of some ways in which the data bank has been used:

The United Nations Seabed Assessment Committee, preparing for a Law of the Sea Conference, needed maps of worldwide manganese nodule distribution.

A chief scientist at SIO planning a cruise wanted to know the sites within a certain area that had already been sampled by coring.

A scientist studying copper distribution in the Earth's crust requested a map showing copper content in marine manganese nodules.

A geologist doing research on deposition of calcium carbonate in the Pacific Ocean wanted complete information on all samples for which carbonate determinations had been made.

A graduate student embarking on the study of pteropods requested a list of all cores which were known to contain these organisms and which were archived at U. S. institutions.

The coordinator of the NSF IDOE program on manganese nodules, a project involving investigators from a number of geographically scattered institutions, deposited cruise reports and analytical results in the data bank in order to facilitate communication among project participants as well as to fulfill his obligation to make results available to the general public.

The U. S. Geological Survey, preparing maps for the Circum-Pacific Map Project, requested base maps showing sediment types classified according to their specifications.

The data bank is an active research tool. It functions as both a storage and retrieval facility. The information which users deposit into the bank will be available for their future use as well as that of the scientific community and the general public.

3

B. STRUCTURE AND CONTENTS OF THE DATA BANK

The Sediment Data Bank consists of a series of linked files, each containing separate categories of information. With this structure additional types of information can be stored as the need arises without any major changes in the system.

Each data record is associated with a reference number which indicates the original source of the data. Separate bibliography files contain entries for each published or unpublished source from which data have been extracted.

All files are condensed as much as possible. Fixed-field formats are utilized, with each parameter stored in a partial computer word. For this reason direct copies of the files are useless without the accompanying software and are therefore not available to the user. All printed output in provided in plain English or easily understood codes (see examples in Section E). Magnetic tape or disk files of selected data can be prepared in formats suitable for later processing (see Section F).

Files are stored permanently on magnetic tape and are copied to disk for searching to allow for random access. A file on disk may be searched simultaneously by a number of users.

The remainder of this section outlines the contents of each file. More detailed explanations of data items are given in Tables 1-4 and in SIO Reference 78-9. Every item of information stored adds to the cost of the data bank by requiring additional effort to code, more disk storage space, and increased computer time for searches. We have tried to balance requirements for detailed information with the need to make the data bank economical. Specific items stored were selected according to our own research interests and the requirements of our contributors and users. It is always possible to add new items or revise classifications, and this is done from time to time as the need arises.

1. The STATIONS File

The STATIONS file is the basic file of the data bank and contains one record for each point on the seafloor where a geological sample is known to have been collected. Where a number of different samples have been collected from within a small area, sometimes referred to as a single "station" in cruise reports, each sample is considered a separate "station" for the purposes of this system.

Information contained in the STATIONS record includes latitude, longitude, water depth, ship or cruise name, station or sample number, type of sampling device, length of sample (if a core), institution collecting the sample (only for some major institutions), and a Reference Number indicating data source. Also included, if the information is available,

5

are surface sediment lithology, an indication of the presence or absence of manganese nodules, seafloor nodule coverage or concentration (kg/m^2), and nodule location in a core (above or below 10 cm from the sediment surface). Finally, this record contains the addresses in other files where analyses or descriptions of samples from the station are stored. The SEARCH program uses these links to locate additional information about the station.

Each record in this file is assigned a unique seven-digit Sequence Number. The first three digits of the Sequence Number refer to the 10^0 square on the surface of the Earth from which the sample was collected. The square numbers used in this system are not the same as Marsden Squares, although the principle is similar. Each square is bound by 10^0 parallels and meridians. Numbers run from 1 to 648, beginning at the South Pole at 0^0 longitude. Numbers proceed westward in that latitude band through 360^0 (36 squares). Square 37 then begins at 80^{0}S, 0^{0}W, and so on (Fig. 1).

This numbering system makes it quite simple to calculate the square number for any location. First, express southern latitudes as negative and northern as positive, western longitudes as positive and eastern as 360 minus longitude. The formula is then as follows:

Square Number = 36 x   INTEGER(latitude/10) + 9 +

INTEGER(longitude/10) + 1

where the INTEGER function is defined as the next lower integer. For example, let us calculate in which square the location 32.5^{0}S, 137^{0}W would be found. INTEGER(-32.5/10) = -4, and INTEGER(137/10) = 13. Thus the square number = 36 x 5 + 13 + 1 = 194.

The final four digits of the Sequence Number are assigned in order of record acquisition within each square. Besides being included in the STATIONS record, the entire Sequence Number is repeated in every analytical or descriptive record pertaining to that station in any other file.

Station records are stored in order of Sequence Number, and the file is resorted after each addition of new data records. The first part of the file contains an index in which is stored the starting location of data for each 10^0 square. This increases the efficiency of the SEARCH program; when stations are to be selected for a specific area, only records from squares within that area need to be examined, not all records in the file. Files produced by SEARCH which contain a subset of the STATIONS file have no index.

2. The Manganese Nodule Analyses (NAR and XTRA) Files

The NAR file contains one record for each available analysis of a Mn nodule or crust or part thereof. There may be more than one analysis for a given station. Each record includes the Sequence Number assigned to the station at which the nodule was collected as well as an Analysis Number assigned to each separate analysis and a Reference Number indicating the source of the analysis. Besides the weight percent of each element determined, there is descriptive information about the sample including sample type, dimensions, morphology, section analyzed, analytical method, nucleus, and depth below seawater-sediment interface at which the sample was found.
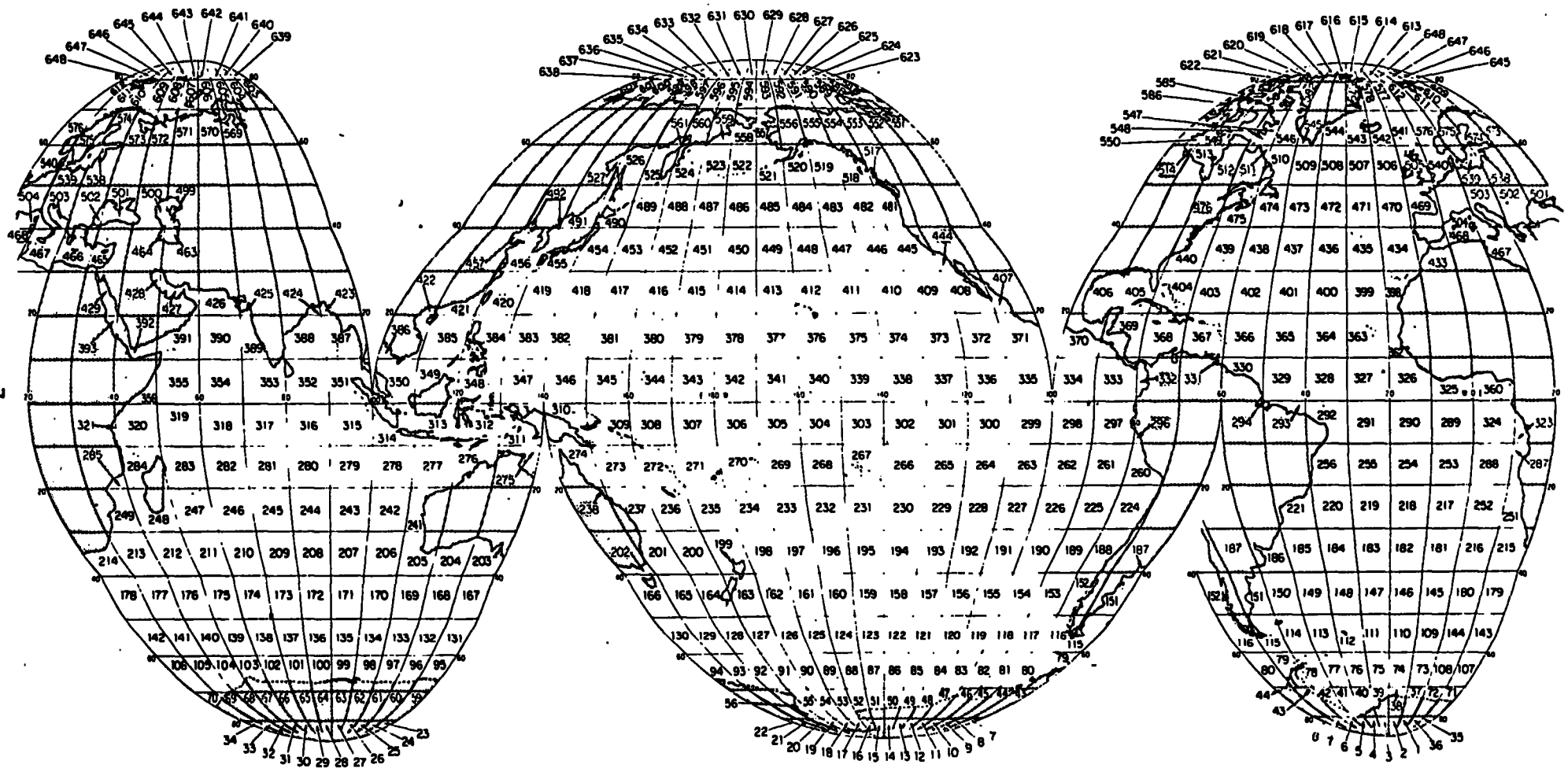
Fig. 1. Map illustrating Sediment Data Bank system for numbering ten-degree squares on the Earth's surface.

In order to save disk storage space and allow for efficient data retrieval, the NAR records contain only the elements and sample characteristics most commonly measured. Additional data, if any, are stored in a separate file titled XTRA. Information in the XTRA file, which can include the weight percent of any element plus some of the descriptive information mentioned above, is linked to the primary analysis record in the NAR file through an address link stored in the NAR record. For each NAR record there may be no XTRA records or as many as are necessary to contain all available data.

Records in the NAR and XTRA files are stored in random order. The first analysis for each station is located through the address link in the STATIONS file, and each subsequent analysis or XTRA record is linked through addresses stored in the NAR file.

3. The Sediment DESCRIPTIONS File

The DESCRIPTIONS file contains descriptive information about the sediment collected at each station, except where no descriptions are available or available data have not been coded. There is one record for each layer of sediment described; thus for a grab sample there is generally only one record, while for long cores there may be more than 100. Major and minor sediment components which can be stored for each layer include the following:

Depth below the sediment-water interface at which the described layer begins and its length in centimeters.

Dominant lithology: rock or gravel, manganese nodules or pavement, sand or silt, mud, siliceous ooze, calcareous ooze, siliceous-calcareous ooze, clay, volcanic ash, or zeolitite.

Color: predominant and secondary color of the sediment expressed in an abbreviated version of the Munsell Color Code.

Grain Size: amounts of rock, gravel, sand, silt and clay present, expressed as a percentage rounded to the nearest 10% or noted as dominant, secondary or minor constituents. Rock types are specified where known, and the sand-silt and clay fractions may be described by a limited number of adjectives.

Calcium carbonate and organic carbon content.

Consolidation.

Microfossils present: amounts indicated where known for foraminifera, pteropods, radiolaria, calcareous nannofossils and diatoms.

Minerals present: indication of presence (but not amount) of quartz, feldspar, pyroxene, chlorite, mica, glass, palagonite, glauconite, barite, phillipsite, other zeolites, pyrite, other sulfides, other dark minerals, other opaque minerals.

he presence of turbidites, grading, bedding, worm
,ng, volcanic ash layers, dispersed ash, H_2S, man-
t or crusts, manganese nodules or micronodules,
flow-in.

lat`can be stored is limited by the quality and quantity
n the source reference. Where only a brief shipboard
.ilable, the record may contain only lithological·type;
:te description is available, most of the fields will

contains the Sequence Number assigned to the station.
:d in random order and are located by means of links in
e. The DESCRIPTIONS file contains no Reference Numbers,
'e information is generally from the same source as data in
/e and would have the same Reference Number. Where addi-
arding samples have been obtained from other sources, there
rence in the BIBLIO/TEXT file.

nt Analyses Files

itaining chemical analysis of sediments will contain similar
nd will be structured in approximately the same way as the
es files. Although collection of analytical data for sedi-
jun and the format of the files developed, this part of the
; not yet been implemented.

floor Photograph (FOTO) File

gh manganese nodule coverage as determined from seafloor photos
in the STATIONS file, it has been suggested that more photo
should be stored in the data bank. Whereas the STATIONS
ains data summarized from a number of photographs at each sta-
uld be useful to include descriptions of individual photographs
studies of nodule distribution. Other information that can be
om photographs, such as fauna, indications of currents, etc.,
be useful for a variety of investigations. A FOTO file has
on an experimental basis, but it contains few data and is not
or general use at this time.

rce Reference (BIBLIOgraphy) Files

able users to identify the original sources of information in
t Data Bank, reference files are maintained with entries for
hed or unpublished source from which data were taken. More
ferences are now included in the BIBLIO files. Most are primary
taining station lists, sample descriptions, or analyses that
oded into the data bank files. Some, however, are secondary
taining additional information that has been used to augment the
In addition to standard reference information, entries which
specific oceanographic cruises include the cruise date and name

9

Indication of the presence of turbidites, grading, bedding, worm
burrows, mottling, volcanic ash layers, dispersed ash, H_2S, man-
ganese pavement or crusts, manganese nodules or micronodules,
disturbance or flow-in.

Information that can be stored is limited by the quality and quantity
of data provided in the source reference. Where only a brief shipboard
description is available, the record may contain only lithological type;
but where a complete description is available, most of the fields will
contain data.

Each record contains the Sequence Number assigned to the station.
Records are stored in random order and are located by means of links in
the STATIONS file. The DESCRIPTIONS file contains no Reference Numbers,
since descriptive information is generally from the same source as data in
the STATIONS file and would have the same Reference Number. Where addi-
tional data regarding samples have been obtained from other sources, there
is a cross reference in the BIBLIO/TEXT file.

4. The Sediment Analyses Files

Files containing chemical analysis of sediments will contain similar
information and will be structured in approximately the same way as the
nodule analyses files. Although collection of analytical data for sedi-
ments has begun and the format of the files developed, this part of the
data bank has not yet been implemented.

5. The Seafloor Photograph (FOTO) File

Although manganese nodule coverage as determined from seafloor photos
is included in the STATIONS file, it has been suggested that more photo
information should be stored in the data bank. Whereas the STATIONS
record contains data summarized from a number of photographs at each sta-
tion, it would be useful to include descriptions of individual photographs
for use in studies of nodule distribution. Other information that can be
obtained from photographs, such as fauna, indications of currents, etc.,
could also be useful for a variety of investigations. A FOTO file has
been set up on an experimental basis, but it contains few data and is not
available for general use at this time.

6. The Source Reference (BIBLIOgraphy) Files

To enable users to identify the original sources of information in
the Sediment Data Bank, reference files are maintained with entries for
each published or unpublished source from which data were taken. More
than 750 references are now included in the BIBLIO files. Most are primary
sources containing station lists, sample descriptions, or analyses that
have been coded into the data bank files. Some, however, are secondary
sources containing additional information that has been used to augment the
coded data. In addition to standard reference information, entries which
pertain to specific oceanographic cruises include the cruise date and name

of the ship or ships, if known. Some sources of analytical information are annotated to include information about analytical method.

Each entry has been assigned a reference number consisting of three digits preceded by a letter. Typical bibliographic items are shown in Fig. 3, page 39.

The complete bibliographic information is stored as EBCDIC records in the BIBLIO/TEXT file. A supplementary file, BIBLIO/INDEX, contains the date, author, institution (for some items), which data bank files if any contain information coded from that source, and from what oceans or areas the samples discussed were collected. This index file can be searched according to any of these parameters by Program BIBSEARCH (see Section D-4); output is either a list of the Reference Numbers for sources meeting the specified criteria or the complete bibliographic information of such references.

After a subset of information has been selected from one of the data bank files, the user may wish to obtain a reference list of all sources for the selected data. Program BIBLIOLIST (see Section E-5) produces either a Reference Number list or a complete bibliography for any stations or nodule analyses file produced by Program SEARCH.

C. DATA COLLECTION

Input to the files is coded from original sources entirely by the data bank staff. Since the format of the input records differs completely from the formats of the stored records or the output records, the actual format used for data coding is of little concern to the user. However, detailed coding instructions and an explanation of how we interpret sediment lithological classes assigned by various authors are contained in a companion volume to this report, SIO Reference 78-9.

Files are updated about every two months. As new information becomes available for stations already in the files, it is added to existing records; or existing records may be deleted or entirely replaced. In order to give the user some idea of the completeness and reliability of the data files, the remainder of this section discusses our general priorities in data coding and some specific problems that occur when one compiles data from a variety of sources.

1. The STATIONS File

Most data items included in the STATIONS file are quite straight-forward and require almost no subjective interpretation on the part of the coder. Therefore, the most likely errors in this file are typographical, either in the original data source or introduced by our coders. Sometimes it is possible to correct source errors; for example, it is obvious that a ship cannot travel from 30°N to 30°S and back again in a few hours, so in such a case we can presume that a latitude is given incorrectly.

The location given for each station is that reported in the most recently available reference. Sometimes cruise reports give ship positions as estimated from shipboard logs; later reports show corrected positions based on final processing of the navigation data. In such cases, the station record is corrected to show the latter position. For a few stations for which the coordinates are not given, it is necessary to interpret station locations from maps. The precision of the coordinates for a given station may be checked by consulting the original source.

Hundreds of seafloor samples are collected every year from the deep ocean. Sometimes station locations and information about the samples are included in the published literature, which we continually survey. More frequently, academic institutions summarize their cruise data in unpublished reports or computer listings which we can obtain on request. Much information, however, is buried in archives throughout the world. Some of it can be dug out if we know what to ask for and whom to ask, but a large number of data are proprietary.

Our staff is too small to code even all the data that are easily available, so we add information to the files according to priorities based on our own research interests and those of our major users. Although we

have coded some data for continental shelf samples, most of our efforts are concentrated on samples from water depths greater than 200 meters. Within this limit, our priorities are as follows:

1) Samples collected by Scripps Institution of Oceanography (SIO station data is currently complete).

2) Samples collected from the Northeast Equatorial Pacific ($0-25^{\circ}N$, $100-180^{\circ}W$) where manganese nodule resources are most concentrated (currently complete for all information available to us).

3) The remainder of the Pacific Ocean (reasonably complete for data from major institutions).

4) The Southeast Asia Area ($15^{\circ}S - 45^{\circ}N$, $90-160^{\circ}E$), which was brought up to date in 1975 in conjunction with the National Science Foundation IDOE/CCOP/IOC SEATAR Program.

5) The rest of the world ocean (we estimate only half of the available station data for this area is coded).

Most stations were added before information about manganese nodules was included in the STATIONS file. Nodule occurrences have now been added wherever data are available. Different investigators are inconsistent in reporting nodules, however, and the distinction between "nodules absent" (reported as absent or would have been reported had they occurred) and "no information" (for stations where the investigator tends to ignore nodules even if they do occur) has not yet been included for many stations. This could lead to significant errors when the data bank is used to estimate worldwide nodule distribution. Reported nodule populations from photographs or concentrations as determined from grab samples have been added for much of the Pacific Ocean, but are incomplete.

Although each oceanographic institution or private company has much more extensive information about samples it has collected, the Sediment Data Bank has probably the largest public file of station information for samples collected by all institutions.

2. The Manganese Nodule Analyses Files

The nodule analyses files contain marine manganese nodule elemental chemical analyses and descriptive information coded from published and unpublished sources from throughout the world. Data are coded as they appear in the source; we do not interpret or evaluate the data.

Occasionally an analytical reference will contain incomplete or ambiguous station information. That is, it may not be clear from which station the sample was actually taken, particularly when multiple sampling devices were used to retrieve samples at the same location. In most cases we choose the sampler type which is most likely to pick up nodules and we assign them to that station; for example, if the nodule could have come from either a grab sample or a piston core, we would assume it came from the grab. A second problem arises when samples from two or more closely spaced stations are analyzed together. In this case, we choose the most

central location as the source of the sample. In both cases, the user
should refer to the original source if questions arise.

Descriptive terms pose the greatest difficulties in coding analyses.
Terminology varies among authors; and as yet there are no standard definitions
for nodule, crust, encrustation or concretion. In general, we code as nodules
samples described as either nodules or concretions. Crust or pavement is
coded as crust, meaning a sample which does not show concentric banding.
Encrustation is coded as manganese-coated rock. Often an author will describe
the samples in the text of the paper. When this is done, we take the
description as the identification. Many times there will be no persuasive
identification, and in these cases the sample type will remain blank. Data
bank studies have shown, however, that when one calculates average composition
for a region there is no significant change in the result if one includes
the unidentified samples with the clearly identified bulk samples.

Since we cannot code every variation in terminology, fine distinctions
may be lost when a description is coded. For instance, outer layer or
outer crust may refer to anything from a 1 mm layer to a 10 mm layer, or to
the coating on a basalt. Middle layer may refer to one of any number of
layers occurring in the middle of a sample. Every effort is made to iden-
tify the true nature of the sample and the section analyzed; however, the
user should refer to the original data source for more precise information
and regard the stored information as a general guide.

If the user wishes to eliminate blocks or types of data, this can be
done at the time of the search using any of the parameters listed in Table
3 (see Section D). When descriptive data items are absent in the original
source, they are coded as zeros.   Elements which are undetermined are
listed as 0.0% concentration.

Each analysis is stored as a single record, and there may be up to
128 separate analyses for a single station. Since the data come from a
number of different sources and analyses are performed by a variety of
techniques, the assays are not consistent. We make no claims about the
precision or accuracy of data from the literature or other sources. We
have checked many apparently anomolous data in our files and find that the
anomolies are present in the original source. However, we are convinced
that analytical errors as well as coding errors are random and tend to
cancel each other.

Although private companies have much more extensive data collections,
the Sediment Data Bank contains the most complete and comprehensive com-
pilation of public data on marine manganese nodules available today.

3. The Sediment DESCRIPTIONS File

Due to space limitations imposed by computer storage and processing
it is necessary to condense an original description of the physical and
chemical properties of sediment into a brief coded abstract. In preparing
the DESCRIPTIONS file it has been our basic policy to be as internally con-
sistent as possible. There are problems, however, in attempting to make
consistent a diverse set of semi-subjective sediment descriptions of

13

varying quality which have been prepared by a large group of people with assorted interests and different levels of experience using a variety of classification schemes. Some descriptions include textural data only, while others list only mineralogical or paleontological constituents. Some published reports may be based on an extensive period of systematic research, while on the other hand some shipboard identifications are made hastily without laboratory equipment.

We feel that it is worthwhile to include all available data in the hope that even preliminary information can be of some use to a researcher. The user is advised to obtain the list of references which can be made for the product of all data searches. The user should be warned that an appreciable number of shipboard descriptions will be proved in error by later analysis. However, preliminary shipboard descriptions are all that are available for a very long time for many samples, particularly from SIO cruises. As soon as a sample has been restudied and the new description becomes available, the old description is replaced.

Ambiguities in lithological classifications caused by the use of different classification schemes are discussed in SIO Reference 78-9. The data bank user can, however, search the data according to any parameters stored to assign new lithologies to any samples. An example of such reclassification is described in Section D-2.

Lacking adequate funding for such a project, we are presently making no attempt to code detailed sediment descriptions for all stations in the data bank. A general indication of surface lithology is included in the STATIONS file, but complete description records are coded for only two types of samples:

1) cores, dredge hauls and grab samples collected by Scripps Institution of Oceanography, and

2) samples from areas of interest to specific funded research projects.

The DESCRIPTIONS file is complete for described samples from SIO expeditions since 1972. Sediment descriptions of samples taken on earlier cruises are in the old data bank and are gradually being added to the new system. In 1975, as part of the SEATAR Program of IDOE/CCOP/IOC, coding was completed for the area 15°S - 45°N, 90-160°E. The area 50-80°N, 120-170°W, including the Gulf of Alaska, was updated extensively in 1976. Many descriptions have also been coded for the Northeast Equatorial Pacific area, 0-20°N, 80-180°W. This is currently the only area for which sediment descriptions other than those from SIO are being added to the file.

14

D. INFORMATION RETRIEVAL

Information stored in the Sediment Data Bank is in the public domain and is available to anyone. However, data users must generally pay for their information retrieval costs. Anyone may request information by letter or telephone, and requests are normally filled within one to seven days unless special programming is required.

Frequent data users may wish to access the data bank directly from remote terminals located in their own offices. Remote interactive access is available on request to any group at SIO or to members of Scripps Industrial Associates; the only requirement is that such users have a valid account at the UCSD Computer Center and pay the actual computer charges. Other users may gain direct access to the data bank by means of a purchase order or contract to the University of California.

Selection of information to be retrieved is accomplished with the interactive SEARCH program described below. Other programs available for displaying or manipulating the data are described in Sections E and F. Additional programs can be written to meet specific needs, but the user must usually pay the cost of such special programming.

All programs are written in Burroughs-6700 Extended Algol, but the user need not be familiar with this language. All our programs are designed for use by persons who are not computer experts. In order to use the SEARCH program one need do only the following:

1) Have a valid computer account and password at the UCSD Computer Center (see Computer Center Documents #037, "The Campus Computer Center, and #018, "Rates, Schedules and Limits"). Let us know the account number so that we can authorize access to the data storage tapes.

2) Copy the necessary program and data files from tape to disk (see Computer Center Document #059, "Opinfo").

3) Follow the instructions in Section D-1, keeping in mind the advice in Section D-2.

To run other programs described in this handbook the user must also be familiar with the Burroughs CANDE (Command-and-Edit) language, Computer Center Documents #251, "CANDE Handbook," and #127, "Primer for Interactive Usage." Individual instruction can be provided by the data bank staff.

Users who plan to request data instead of doing their own searches will still find it useful to read through Section D-1 in order to see what types of searches can be performed and what parameters can be used as search criteria.

1. Program SEARCH

    a. <u>Purpose</u>: SEARCH is an interactive program that selects information from any or all of the files in the Sediment Data Bank or from any disk files made by a previous execution of the program. SEARCH accepts a series of user input statements which specify the files to be searched and the criteria to be used in the search. Files are searched one at a time. Retrieved data are stored in a newly created special disk file. After each search the user may obtain a list of selected stations or analyses. A disk file produced by the program may be removed immediately or saved as the basis of a future search, for printing (programs LISTSTA, LISTDESCRIP, and LISTNAR), plotting (Program MAP), numerical calculations (Program NARCALC, for NAR files only) or copying to tape.

    b. <u>Use</u>: The files to be searched must be present on disk before this program is run. To execute SEARCH, sign on from a remote terminal and enter RUN SEARCH. The program will then type out a series of prompts to which the user must enter a response. Fig. 2 is an example of a complete search session involving searches of all three file types. The prompts and questions printed on the terminal by the program are indicated below by a preceding # sign; appropriate responses and explanations follow:

#NAME OF FILE TO SEARCH:

    Enter STATIONS, NAR, DESCRIPTIONS or the name of any disk file that contains a subset of one of these files. Generally such files have been created by previous searches of the data bank files. If no more files are to be searched, enter the word STOP. If the file specified is not present on disk, the user will receive an error message and a chance to enter a new file name.

#TYPE OF SEARCH FILE:

    Enter STATIONS, NAR, or DESCRIPTIONS. This will be asked only if the search file is not one of the standard (i.e., STATIONS, NAR, or DESCRIPTIONS) data bank files. The answer is the name of the file in the data bank with a format corresponding to that of the file to be searched. For example, all stations selected by a search on the STATIONS file will be stored in a file of type STATIONS.

#STATIONS TO BE SELECTED FROM THOSE IN WHAT FILE?

    This will be asked only for a search on the NAR or DESCRIPTIONS files. Enter the name of a file which is a subset of the STATIONS file, or ALL if the entire file is to be searched.

    Let us consider a hypothetical search for a list of all stations where water depth is greater than 2500 meters and $CaCO_3$ content greater than 20%. This cannot be done with a single search because bathymetry information is stored in the STATIONS file while $CaCO_3$ content is stored in the DESCRIPTIONS file. The user would first search STATIONS for BATHYMETRY > 2500 and store the selected stations in a file called SOMENAME. Then he would search DESCRIPTIONS for CACO3 > 20, stations to be selected from those in file SOMENAME. Not all records in the DESCRIPTIONS file would be searched in this case, but only those records pertaining to the stations already

16

selected and stored in SOMENAME. To search the entire DESCRIPTIONS file, the user would enter the word ALL in answer to this query.

#WRITE SELECTED STATIONS IN FILE TITLED:

Enter the name (up to 30 characters) you want assigned to the disk file in which selected stations will be stored. This must be a new file; records cannot be added to a file already present. If the name entered is the title of a file already on disk, the user will receive an error message and a chance to enter a new name.

In the case of a search on the NAR or DESCRIPTIONS files with stations selected from some file (answer to the preceding prompt was not ALL), the selected stations will be those for which at least one description or analysis meets the search criteria.

#WRITE SELECTED $\left\{ \begin{array}{l} \text{DESCRIPTIONS} \\ \text{NODULE ANALYSES} \end{array} \right\}$ IN FILE TITLED:

This will appear only in a search on an analyses or descriptions file. Enter a new file name.

#ENTER PARAMETERS FOR SEARCH:

Enter parameters according to instructions in Section D-2c below. A file may be searched according to any of the parameters that are stored in that file. See Tables 1, 3 and 4 for parameters stored in the STATIONS, NAR and DESCRIPTIONS files, respectively.

#NNN STATIONS SATISFY SEARCH PARAMETERS.
#DO YOU WANT A LIST OF THESE STATIONS?

Enter YES or NO. This will appear at the end of a search on a stations file. If the user wants a list of selected stations to be printed on the terminal, the answer would be YES. By hitting the break key the user can terminate a listing without terminating the program or losing the file.

Note: Because of the limited line width of most interactive terminals, the station listings that can be obtained at this point are incomplete; they lack the surface sediment type and information about manganese nodules. For a complete listing on a line printer, save the file and then run Program LISTSTA (see Section E-1). For a long listing this procedure is also more economical.

#NNN NODULE ANALYSES SATISFY SEARCH PARAMETERS.
#WOULD YOU LIKE A LIST OF THESE ANALYSES?

Enter YES or NO. This will appear at the end of a search of a nodule analyses file.

Note: terminal listings of analyses files are also incomplete, including the weight percentages of only seven elements (Fe, Mn, Co, Ni, Cu, Zn, Mo) and no descriptive information. Program LISTNAR (see Section E-2) must be used to obtain a complete printout.

17

```
TYPE <CR>
::UCSD B6700 CANDE 29.170: YOU APE TTX1(86)
::ENTER USERCODE PLEASE
GR15
::ENTER PASSWORD PLEASE.
███████████████████
:: DEFAULT PRINT DESTINATION=SAVE
::PRIME RATE SESSION 2258  10:42:03  04/05/78
RUN SEARCH
♦RUNNING 0337
::?
 NAME OF FILE TO SEARCH:
NEQSTATIONS
 TYPE OF SEARCH FILE:
STATIONS
 WRITE SELECTED STATIONS IN FILE TITLED:
DEMO/STA
 ENTER PARAMETERS FOR SEARCH:

LAT = 0 N TO 10 N, LON = 140 W TO 136.3 W;
 SEARCHING. . .PLEASE WAIT


    41 STATIONS SATISFY SEARCH PARAMETERS.
 WOULD YOU LIKE A LIST OF THESE STATIONS?
YES
```

SEQ. NUMBER & STATION NO	SHIP/CRUISE	LATI- TUDE	LONGI- TUDE	SAMPLING DEVICE	WATER DEPTH (M)	CORE LENGTH (CM)		REF NO.
3380040	RC12-58	9.833N	136.383W	PISTON CORE	4823	7	LDGO	A114
3380127	VEMA28-177TW	4.817N	136.450W	TRIP CORE	4709	1	LDGO	A266
3380126	VEMA28-177	4.817N	136.450W	PISTON CORE	4709	15	LDGO	A266
3380041	ALBATRO.15	4.583N	136.900W	DREDG	4727	0	SIO	A009
3380042	ALBATRO.14	6.683N	137.000W	DREDG	5080	0	SIO	A009
3380043	ALBATRO.16	2.633N	137.367W					

```
 LISTING INTERRUPTED BY USER--MAY BE INCOMPLETE

 NOTE:  A VALUE OF ZERO MEANS NOT KNOWN.
 KEEP FILE (GR15)DEMO/STA ON PACK?
YES
 KEEP FILE (GR15)NEQSTATIONS ON PACK?
YES
 NAME OF FILE TO SEARCH:
NAR
 STATIONS TO BE SELECTED FROM THOSE IN WHAT FILE:
DEMO/STA
 WRITE SELECTED STATIONS IN FILE TITLED:
DEMO/NARSTA
 WRITE SELECTED NODULE ANALYSES IN FILE TITLED:
DEMO/NAR
 ENTER PARAMETERS FOR SEARCH:
MN GEQ 0;
 SEARCHING. . .PLEASE WAIT


    43 NODULE ANALYSES SATISFY SEARCH PARAMETERS.
 WOULD YOU LIKE A LIST OF THESE ANALYSES?
YES
```

Fig. 2. Example of a run of Program SEARCH showing computer prompts and
appropriate responses and on-line output of station and analysis records.
The first file searched (NEQSTA) is a subset of File STATIONS made by a
previous search.

SEQUENCE NUMBER	REF NO.	MN	FE	CO	NI	CU	ZN	MO
3380040	A179	21.80	5.80	0.2700	1.2000	0.9400	0.0000	0.0000
3380044	A179	2.50	8.20	0.0500	0.1100	0.1100	0.0000	0.0000
3380045	A197	29.80	4.80	0.2000	1.3600	1.2000	0.1200	0.0540
	A166	48.50	3.38	0.1500	2.4600	1.5400	0.0000	0.0000
	A261	32.00	5.30	0.1800	2.0000	1.4000	0.0400	0.0700
	A201	0.00	6.20	0.2600	1.6000	1.7000	0.0000	0.1100

LISTING INTERRUPTED BY USER--MAY BE INCOMPLETE

NOTE: A VALUE OF ZERO MEANS NOT KNOWN.
KEEP FILE (GR15)DEMO/NAR ON PACK?
YES

 DO YOU WANT A LIST OF STATIONS FOR THESE ANALYSES?
NO
 KEEP FILE (GR15)DEMO/NARSTA ON PACK?
NO

 KEEP FILE (GR15)NAR ON PACK?
YES
 KEEP FILE (GR15)DEMO/STA ON PACK?
YES
 NAME OF FILE TO SEARCH:
DESCRIPTIONS
 STATIONS TO BE SELECTED FROM THOSE IN WHAT FILE?
DEMO/STA
 WRITE SELECTED STATIONS IN FILE TITLED:
DEMO/DESCRIPSTA
 WRITE SELECTED DESCPIPTIONS IN FILE TITLED:
DEMO/DESCRIP
 ENTER PARAMETERS FOR SEARCH:
LIT GEQ 0;
 SEARCHING. . .PLEASE WAIT

     178 DESCRIPTIONS SATISFY SEARCH PARAMETERS.
 WOULD YOU LIKE A LIST OF STATIONS FOR THESE DESCRIPTIONS?
NO
 KEEP FILE (GR15)DEMO/DESCRIPSTA ON PACK?
YES
 KEEP FILE (GR15)DEMO/DESCRIP ON PACK?
YES
 KEEP FILE (GR15)DESCRIPTIONS ON PACK?
NO
KEEP FILE (GR15)DEMO/STA ON PACK?
YES
 NAME OF FILE TO SEARCH:
STOP
*ET=7:49.6 PT=6.2 IO=11.2

REM JOB/=
* 1 FILE IN (GR15).JOB/= REMOVED ON PACK
BYE
*END PRIME RATE SESSION 2258 ET=1:40.2 PT=6.2 IO=11.2
*USER = GR15 10:44:03 04/05/78

Fig. 2 (continued)

#NNN DESCRIPTIONS SATISFY SEARCH PARAMETERS.

This message will appear at the end of a search of a descriptions file. It is for the user's information only, as descriptions cannot be listed on the terminal. Complete printouts of descriptions can be obtained with Program LISTDESCRIP (see Section E-3).

#DO YOU WANT A LIST OF STATIONS FOR THESE $\left\{ \begin{matrix} \text{DESCRIPTIONS} \\ \text{ANALYSES} \end{matrix} \right\}$?

Enter YES or NO. This question will be asked only after a search of the NAR or DESCRIPTIONS file with stations selected from some stations file.

#KEEP FILE XXX ON PACK?

Enter YES or NO. If you want this file as the basis for a future search or if you want a line printer listing, answer YES; files not saved will be removed immediately. Any incorrect entry here will be interpreted as YES, and the file will be removed only if NO is correctly entered. This question will be repeated with different file names until the status of all files utilized in or produced by the search is determined.

#NAME OF FILE TO SEARCH:

At the end of each search, the program goes back to the first query and asks for a new file to search. Enter STOP to stop the program. To terminate the program at any other point during execution, enter ?DS .

c. Search Parameters: Search criteria are specified through a combination of parameters, modifiers, values, and logical operators. Parameters must be selected from Table 1, 3 or 4 depending on the type of file to be searched. Input parameters and modifiers will be truncated to the first three letters. The user may enter the full word, but the first three letters must match the abbreviation code in the tables. This code is usually but not always the first three letters of the parameter name.

Each parameter, operator and value must be delimited by one or more spaces. Parameters or groups of parameters must be separated by commas. A semicolon and then a carriage return must follow the last input. Up to 20 parameters are allowed for each search. Data are entered free form and may take as many lines as needed; but only 72 characters may be entered before a carriage return, and all data lines should start with a parameter name or operator.

If parameter codes are entered incorrectly or required operators are omitted, the user will receive an error message and a chance to enter new data.

For some parameters the only available information is yes or no--the presence or absence of that constituent. For these parameters, no values or modifiers are permitted. For example:

ZEOLITE, QUARTZ;

20

Values may be specified for some parameters. Values are usually numerical but may be alphabetical codes; see Tables 1-4 for possible values. Any of the following relational operators may be used with the value: >, GEQ (greater than or equal to),<, LEQ (less than or equal to), =, NEQ (not equal). At least one blank space must precede and follow each operator. For example:

    LATITUDE GEQ 20 N, INSTITUTION = SIO, SAMPLER = CORE;

If no values are given, the program will check merely for the presence or absence of the constituents. For example, CLAY, CACO3 > 10; selects samples with clay present, >10% $CaCO_3$.

Certain parameters in the DESCRIPTIONS file may also be modified by a term from Table 5 or 6. Modifiers must appear after the noun. Example:

    SAND FINE, ROCK IGNEOUS;

If a parameter or modifier is preceded by the word NOT, samples selected will be those with an absence of such constituents. Example:

    MNNODULES, NOT MNMICRONOD, ROCK NOT BASALT;

If values are connected by the word TO, records selected will be those where the value of the parameter is greater than or equal to the first value and less than or equal to the second. For example, the following two entries are equivalent:

    LAT GEQ 20 S, LAT LEQ 20 N, LON GEQ 160 W, LON LEQ 100 W;

    LAT = 20 S TO 20 N, LON = 160 W TO 100 W;

For longitude, ">" means "to the east of" and "<" means "to the west of." For latitude, "<" "means south of" and ">" means "north of." (See notes under latitude and longitude in Table 1.)

Parameters, values, or modifiers may be grouped by use of the logical operators OR and &. The & sign must be used here instead of the AND used in Algol or Fortran, because AND is the abbreviation for Andesite in this context. As in Algol, & takes precedence over OR. Thus A OR B & C OR D would be evaluated as "A OR (B & C) OR D." Parentheses may not be used. To obtain an expression that would be evaluated as "(A OR B) & (C OR D)" enter A OR B, C OR D.

    ROCK BAS OR IGN, CLAY SILICEOUS & PELAGIC, PHI OR ZEO;

    CU > 1 OR NI > 1.5 OR TYPE = CRUST & CU > 0.6;

    INST = HIG OR LDGO OR SIO;

Some values to be used in searches are alphabetical; but since internal values are all numerical, one can still use operators such as < and >. See notes in Table 4 for specific instructions.

21

d. Symbols for Plotting: If the selected stations are to be plotted,
the user may wish to use different symbols for different points or dif-
ferent categories of data. A symbol may be assigned to each station
selected at the time of the search and will be stored as part of the
station record in the new file for use by Program MAP.

Symbol assignment is accomplished by entering the word SYMBOL after
the criterion which is to determine the symbol, followed by a number from
the Symbol Table (Fig. 8) which specifies the symbol to be drawn. Instead
of a symbol number, one may enter the word VALUE; this specifies that the
symbol is to be the value stored for that parameter. Thus to plot a map
of copper content in manganese nodules with the percent copper shown for
each point, one would search the NAR file and specify

    CU SYMBOL VALUE;

To make a map showing which of several institutions collected each sample
one could specify

    INSTITUTION = SIO SYMBOL 13 OR LDGO SYMBOL 10 OR HIG SYMBOL 14;

This would produce a map with SIO stations indicated by + (symbol No. 13),
LDGO stations by □ , and HIG stations by X . All internal values are
numerical so one could also specify

    INSTITUTION SYMBOL VALUE;

This would produce a map showing all stations for which the collecting
institution is known. (See Tables 1-4 to see what number corresponds to
a particular character string.) Or for a more complete map one could
enter

    INS SYMBOL VALUE OR NOT INS SYMBOL 13;

This would map all stations in the file. Those plotted in the immediately
preceding example would be designated by symbols for the integers 1-7,
and those stations for which the collecting institution is not stored would
be designated by the symbol + .

e. Symbols as Search Parameters:

Once a symbol has been assigned to a station either by a previous
search, Program NARCALC (see Section F-2) or the present search, the
value of the symbol may itself be used as a search parameter. For example,
suppose one wished to select nodule analyses showing copper content
greater than nickel content. One could achieve this by entering the fol-
lowing search parameters:

    NI SYMBOL VALUE, CU > SYMBOL;

The symbol can be replaced any number of times during a search. One could,
for example, enter

    NI SYMBOL VALUE, CU GEQ SYMBOL SYMBOL VALUE OR CU < SYMBOL SYMBOL 10;

22

A plot of the resulting file would show the copper content for stations where Cu > Ni, and a symbol □ for stations where Cu ≤ Ni.

   f. <u>Output</u>: Output may consist of disk files saved for future use and/or listings of selected stations or analyses. Saved files, including files of descriptions records, may also be listed later on a line printer with one of the programs described in Section E. Fig. 2 shows examples of terminal output.

   g. <u>Caution</u>: Saved disk files will remain on disk and incur storage charges until removed by the user. To see what files are stored on disk under his account number, the user need only enter the word FILES followed by a carriage return. It is recommended that this be done before signing off after an interactive session. File removal can be accomplished by entering the following statement:

   REMOVE <FILENAME>

This statement, as all other terminal entries, must be followed by a carriage return.

TABLE 1. PARAMETERS IN THE STATIONS FILE

Parameter	Code*	Possible Values**	Notes
Square number	SQU	1-648	Refers to the ten-degree square on the surface of the Earth, numbered according to the system described in Section B. Only one square number may be specified for each search if the search is on the complete STATIONS file.
Sequence number	SEQ	any integer	Refers to the sequence of records within each square.
Sampling device Sampler type	SAM TYP	One or two letters; see Table 2.	A search on a particular sampling device may be limited to a specific type of that device if a two-letter code is entered. For example, SAM = C or SAM = CORE selects all cores; SAM = C, TYP = P or SAM = CP selects piston cores only.
Latitude	LAT	90 S to 90 N, in degrees and decimal parts of degrees	The letter indicating north or south is required and must be separated from the number by at least one space. In searches "<" means "south of" and ">" means "north of"; for example, LAT > 20 N selects stations north of 20°N.
Longitude	LON	180 W to 180 E, in degrees and decimal parts of degrees	Spacing same as for latitude. In searches "<" means "west of" and ">" means "east of." If the area to be selected includes the zero meridian, the parameter must be preceded by CROSS. Care must be used in specifying longitude ranges. Picture the area to be selected as if on a map with north in the positive vertical direction; the leftmost longitude is then the minimum. For example, LON = 100 E TO 80 W selects all Pacific Ocean stations, whereas CROSS LON = 80 W to 100 E selects everything else, Atlantic and Indian Ocean stations.
Bathymetry (water depth)	BAT	any integer; 0 = not known	Since zero means not known, a search for BAT < 1000 will select all stations with depths reported less than 1000 m plus those with unreported depths. To exclude stations with unreported depths, search for BAT > 0 & < 1000 or BAT = 1 TO 999.
Core length	LEN	an integer or real number; 0 = not known	For cores >20 m long information is stored to the nearest cm. For longer cores, information is stored to the nearest m. See note under BAT for omitting cores of unknown lengths.
Institution	INS	FSU (1), HIG (2), LDGO (3), OSU (4), USSR (5), SIO (6), WHOI (7)	The institution name is stored for these seven institutions only, although data collected by many others are also stored.
Reference Number	REF	letter + 3 digits	Reference number in the bibliography files; see Section B-6.

*Codes are usually but not always the first three letters of the parameter name or value. The names may be entered in full and will be truncated to the first three characters entered.

**Possible values are numerical or alphabetical as indicated. If values are alphabetical, they are shown in all caps below. For some parameters the internal stored value (which would be assigned as the symbol after a search specifying SYMBOL VALUE) are different from the value codes to be entered for searching. In these cases, the internal values are indicated in parentheses.

TABLE 1 (continued)

Parameter	Code	Possible Values	Notes
Station name	STA	The first 1-6 characters of the station name	This parameter may be used to select stations from a particular cruise or ship. The search will be for up to six characters exactly matching the value given, which may contain no blanks. Warning: not all ship names have been coded consistently; for example, the Robert Conrad stations have been coded as both RC and CONRAD. Vityaz stations have been coded as VITYAZ or VIT, but a search for STA = VIT would also select stations for the USS Vital and maybe others.
Surface lithology	LIT	VA = volcanic ash (1) RO = rock (2) MP = Mn pavement (3) PC = pelagic clay (4) SC = siliceous clay (5) CC = calcareous clay (6) SO = siliceous ooze (7) CO = calcareous ooze (8) CS = calcareous-siliceous ooze (9) TM = terrigenous material (10) ST = siliceous-terrigenous (11) CT = calcareous-terrigenous (12)	Note that the lithology classifications stored in the STATIONS file are not exactly the same as those stored in the DESCRIPTIONS file and that the search codes for them are entirely different. A search for LIT with no operator or value selects all stations which contain surface lithology information
Description record Nodule Analysis	DES NAR	none or integer	A search on one of these items with no value specified selects stations for which records are stored in the specified file. Values refer to addresses at which the first record pertaining to the station is stored in the file. Addresses refer to the physical location of records in the file, not to information stored within the record.
Mn nodule occurrence	NOD	none, or PRE = present (2) ABS = absent (1)	If no value is given the search selects all stations with any information, present or absent. If the value is ABS, search selects stations with nodules reported or deduced to be absent but not stations with no nodule information. If a nodule analysis is stored for the station, the internal value of NOD is the same as NAR and the station will be selected in a search for NOD = PRE.
Nodule coverage	COV	none, 0, 20 or 50 (percent) (internal values: 1 = 0%, 2 = <20%, 3 = 20-50%, 4 = >50%)	If no value is given, search selects stations for which coverage data are available, including stations where coverage is zero.
Sampler by which coverage and concentration estimates made	BY	1 = box core or grab 2 = photographs 3 = miscellaneous, other 4 = sonar 5 = television 6 = grab sampler 7 = coverage by photo, concentration from bottom sampler	
Concentration	CON	real number, in kg/m^2	Refers to concentration of Mn nodules
Surface nodules	SUR	none	Nodule present within top 10 cm of the sediment
Buried nodules	BUR	none	Nodule present below the top 10 cm of the sediment
Symbol	SYM	any real number or integer	SYM is a value assigned to the station record by the present search or a previous search or calculated by Program NARCALC (see Section F-1). At any time a new symbol value may replace the value on which the search is being made.

TABLE 2. SAMPLER TYPES

Sampler Code*	Type Code*	
C (1)		Core
	A (1)	Heat Probe (pre-1971)
	B (2)	Box
	C (3)	Camera
	D (4)	Dart
	E (5)	Triple (pre-1971)
	F (6)	Free-fall
	G (7)	Gravity
	H (8)	Gravity with heat flow thermistors
	I (9)	5-inch Gravity
	J (20)	Multiple core, up to 5 gravity cores
	L (11)	Phleger
	M (12)	Miscellaneous
	O (13)	Gravity oriented
	P (14)	Piston
	R (15)	Piston oriented
	T (17)	Trip (with piston core)
	V (18)	Von Herzen
	W (19)	Biological box
D (2)		Dredge or drag haul
	A (1)	Box
	B (2)	Bucket
	C (3)	Chainbag
	M (12)	Miscellaneous
	P (14)	Pipe
	R (15)	Rock
	T (17)	Trawl
F (3)		Photograph (camera station)
	D (4)	Deeptow
	M (12)	Multiple (more than one frame at same spot)
	O (13)	One photo
	S (16)	Survey (more than one frame, different spots)
	T (17)	Television
G (4)		Grab sampler
	A (1)	Sounding
	B (2)	Box
	C (3)	Catcher (on camera core)
	F (6)	Free-fall
	K (10)	Shipek
	M (12)	Miscellaneous
	O (13)	Orange peel
	P (14)	Petterson
	S (16)	Snapper
	W (19)	Biological
H (5)		Deep Sea Drilling Project hole
	D (4)	Drill core
	S (16)	Side wall core

*Alphabetical codes are those to be entered as values for searches. The numbers in parentheses are the internal values which would be assigned as the symbols if SYMBOL VALUE were entered next.

TABLE 3. PARAMETERS IN THE NAR AND XTRA FILES

Parameter	Code	Possible Values	Notes
Square number	SQU	1-648	Refers to the ten-degree square on the surface of the Earth, numbered according to the system described in Section B.
Sequence number	SEQ	any integer	Refers to the sequence number of records within each square.
Analysis number	ANA	0-127	Analyses for each station are numbered in order of acquisition, starting at zero.
Latitude Longitude	LAT LON	90 S to 90 N 180 W to 180 E	Latitude and longitude are not stored in the file; however, the program uses the Square Numbers to select analyses within specified latitude and longitude ranges that are multiples of ten degrees. Values entered must be multiples of ten.
Reference	REF	Letter + 3 digits	Refers to the item in the bibliography files which is the source of the analysis.
Type of sample	TYP	0 = unknown 1 = nodule or nodules 2 = nodule in sediment 3 = micronodules 4 = micronodules in sediment 5 = nodules and micronodules 6 = Mn-encrusted sediment 7 = Mn-coated rock 8 = crust or pavement 9 = Mn-coated organic material 10 = several crusts	Describes the sample analyzed, not all the material collected at the site.
Size of sample	SIZ	integer; size is coded in millimeters	This parameter represents the diameter or the largest dimension of the sample analyzed. Other dimensions are stored but cannot be searched on.
Section analyzed	SEC	0 = unknown 1 = whole 2 = half 3 = quarter 4 = fragments or pieces 5 = core or nucleus 6 = all but core 7 = cross section 8 = layer adjacent to core 9 = middle layer 10 = outer layer or outer crust 11 = all but outer layer 12 = top side 13 = underside 14 = bulk analysis, representative of total nodule material on seafloor at this location 15 = average of several analyses 16 = composite sample, 2-5 nodules or micronodules analyzed together 17 = composite sample, 6-12 18 = composite sample, >12	When section analyzed is a layer, it is necessary to refer to the source for layer thickness.

TABLE 3 (continued)

Parameter	Code	Possible Values	Notes
Core or nucleus	COR	0 = no information 1 = present but undetermined 2 = none apparent 3 = tooth 4 = ear bone 5 = pumice 6 = chert 7 = palagonite 8 = clay 9 = altered basalt 10 = volcanic 11 = nodule fragment 12 = metallic object 13 = sediment, unspecified 14 = rock, unspecified	
Analytical method	MET	0 = unknown 1 = wet chemical 2 = X-ray fluorescence 3 = X-ray energy spectrometry 4 = electron microprobe 5 = atomic absorption 6 = neutron activation 7 = emission spectrometry 8 = other or combination of methods	If a single method was used for all major elements, the method coded is that method, not 8.
Element concentrations (weight percent)	MN, CO, etc.	integer or real number; zero means not determined	Parameter codes for elements are the standard abbreviations for the element names. Concentrations of the elements Mn, Fe, Co, Ni, Cu, Zn, Si, Ti, Ca, Al, Mo and Pb are stored in the NAR file. Additional element concentrations are stored in the XTRA file. Both files can be searched by a command to search NAR, but XTRA must be on disk when the other elements are searched for.
Water content	H2O	real number or integer; zero means not determined	Water content is as reported in the source. Sometimes this is water loss on heating to $110^{\circ}C$, whereas other times it is loss on ignition. May represent water content of saturated or air-dried nodules. Information on how water content was determined may or may not be included in the data source.
Record	REC	integer $\leq$ the number of records in the file	This parameter refers to the physical position of the record in the file being searched, not to any value stored as part of the record itself.
Depth in core	DEP*	integer	Represents the depth below the surface, in centimeters, at which the sample analyzed was found.
Morphology	MOR*	0 = unknown 1 = spherical 2 = ellipsoidal 3 = discoidal 4 = poly 5 = biological 6 = tabular 7 = faceted	Morphology refers to the sample analyzed, not to all samples collected at the site.
Texture	TEX*	0 = unknown	Texture refers to the sample analyzed, not to all samples collected at the station.

*These parameters, as well as concentrations of all elements except Mn, Fe, Co, Ni, Cu, Zn, Si, Ti, Ca, Al, Mo and Pb are stored in File XTRA, which must be on disk for a search on these parameters. Both files are searched by a command to search NAR; there is presently no way to search XTRA or a subset of its records alone with no NAR file present. Although parameters may be entered in any order, the search will be more efficient if parameters stored in NAR are entered before those stored in XTRA.

TABLE 4. PARAMETERS IN THE DESCRIPTIONS FILE

Parameter	Code*	Possible Values**	Notes
Surface layer	SUR	none	Entering SUR as a parameter will limit the search to the topmost layer described for each station.
Depth in core	DEP	numerical, in centimeters	A search on DEP will retrieve all layers of which any part satisfies the DEP criterion.
Layer length	LAY	numerical, in centimeters	
Color	COL	Red (1), Brown (2), Yellow (3), Green (4), Blue (5), Gray (6), Tan (7)	The only operators that may be used with this parameter are = and NEQ. Value may be modified by LIGHT or DARK following the value.
Lithology	LIT	0 = undetermined 1 = rock, gravel 2 = Mn nodules, pavement 3 = sand, silt 4 = mud 5 = calcareous ooze 6 = siliceous ooze 7 = clay 8 = volcanic ash 9 = siliceous-calcareous ooze 10 = zeolitite	See SIO Reference No. 78-9, Sediment Data Bank Coding Instructions, for definitions of these lithological classifications.
Consolidation	CON	0 = soft or not specified 1 = semi-liquid 2 = stiff, partially indurated 3 = hard, indurated	
Foraminifera Pteropods Coccoliths, other nannofossils Diatoms Radiolaria	FOR PTE COC DIA RAD	ABS = absent (1) PRE = present, amount unknown (2) RAR = rare, probably 1-10% (3) LOW = low, probably 10-30% (4) MOD = moderate, probably 30-60% (5) MAJ = major, probably >30% and possibly >60% (6) DOM = dominant, probably >50% (7)	Any operator may be used. A search for > MOD would select major and dominant; a search for < LOW would select absent, present and rare. NOT DIA would select records with no information about diatoms and would not be the same as DIA = ABS.
Sponge spicules, echinoid spines Macrofauna shells Animal debris Plant debris	SPO SHE ANI PLA	none; search is for presence or absence only	
Percent CaCO₃	CAC	numerical values, 0 to 100, or L = LOW (101) M = moderate (102) H = high (103)	NOT CAC would select records with CaCO₃ unknown; it is not the same as CAC = 0. The letter values can be used with = and NEQ operators only and will select only those records for which CaCO₃ was coded with a qualitative indicator. Numerical values will select only those records for which numerical values were coded. Values may be combined; for example, CAC > 30 OR = M OR H.

---

*Codes are usually but not always the first three letters of the parameter name.

**Possible values are numerical or alphabetical as indicated. If values are alphabetical, they are shown in all caps below and must be entered exactly as listed. For some parameters the internal stored value (which would be assigned as the symbol after a search specifying SYMBOL VALUE) are different from the value codes to be entered for searching. In these cases the internal stored values are indicated in parentheses.

TABLE 4 (continued)

Parameter	Code	Possible Values	Notes
Mud	MUD ⎫	numerical values, 0 to 100(%).	Values stored are rounded to the nearest 10%. Note that some amounts have been coded only as dominant (>50%), secondary (10-50%) or minor (<10%). A search for CLA > 10 would select samples with dominant and secondary amounts as well as those with numerical amounts >10%. May be modified by terms from Table 5.
Clay	CLA ⎬		
Sand	SAN ⎪		
Silt	SIL ⎭		
Rock	ROC ⎫	D = dominant (1)	May be modified by terms from Table 6.
Pebbles, gravel	PEB ⎭	S = secondary (2)	
		M = minor (3)	
Organic carbon	COR	0 = not specified	
		1 = <0.5%	
		2 = 0.5=1.5%	
		3 = >1.5%	
Record	REC	Integer less than or equal to the number of records in the file being searched	The search is on the position of the record in the file being searched, not on any value stored as part of the record itself.
Sample disturbed or flow-in	DIS ⎫		
Turbidite	TUR		
Bedded	BED		
Graded	GRA		
Burrows	BUR		
Mottled	MOT		
Volcanic ash layers	VAL		
Volcanic ash, dispersed	VAD		
H₂S	H2S		
Mn crust or pavement	MNC ⎬	none; search is for presence or absence only	
Mn nodules	MNN		
Mn micronodules	MNM		
Quartz	QUA		
Feldspar	FEL		
Pyroxene	PYX		
Chlorite	CHL		
Mica	MIC		
Glass	GLS		
Palagonite	PAL		
Glauconite	GLA		
Barite	BAR		
Phillipsite	PHI		
Other zeolites	ZEO		
Pyrites	PYR		
Sulfides	SUL		
Other dark min.	DAR		
Other opaque min.	OPA ⎭		

TABLE 5. MODIFIERS FOR CLAY, SAND AND SILT

Code*	Modifier
FIN (1)	Fine
COA (2)	Coarse
PEL (3)	Pelagic
TER (4)	Terrigenous
CAL (5)	Calcareous
SIL (6)	Siliceous
SHE (7)	Shell
COR (8)	Coral
VOL (9)	Volcanic
FER (10)	Ferruginous
ZEO (11)	Zeolitic

*Alphabetical codes are those to be entered
as modifiers of search parameters. The numbers
in parentheses are the internal values which
would be assigned as the symbols if SYMBOL
VALUE were entered after the modifier.

TABLE 6. MODIFIERS FOR ROCK AND PEBBLES

Code*	Modifier
IGN (1)	Igneous, unspecified
BAS (2)	Basalt, pillow basalt, basalt blass
GAB (3)	Gabbro, diabase
PUM (4)	Pumice
BRE (5)	Breccia, volcanic
VOL (6)	Volcanic, unspecified
TUF (7)	Tuff
AND (8)	Andesite
GRA (9)	Granite
PER (10)	Peridotite
MET (11)	Schist, or unspecified metamorphic
GRE (12)	Greenstone
SER (13)	Serpentine
SED (14)	Sedimentary, unspecified
SAN (15)	Sandstone, graywacke
MUD (16)	Mudstone, siltstone, claystone, shale
CON (17)	Conglomerate
LIM (18)	Limestone
CHA (19)	Chalk
CHE (20)	Chert, porcelanite
EVA (21)	Evaporite--gypsum, salt
MNN (22)	Manganese nodules
MNP (23)	Manganese pavement or crust
MNC (24)	Manganese-coated
COR (25)	Coral
PHO (26)	Phosphorite nodules
ZEO (27)	Zeolite nodules
GLA (28)	Glacial transport material
SHE (29)	Macrofauna shells
ALT (30)	Altered
FER (31)	Ferruginous
SIL (32)	Siliceous

*Alphabetical codes are those to be entered as modifiers
of search parameters. The numbers in parentheses are
the internal values which would be assigned as the symbols
if SYMBOL VALUE were entered after the modifier.

2. Search Strategies

The previous section provides basic instructions on how to use the SEARCH program. This section contains some tips on how to maximize search efficiency, dangers to avoid, more on symbols and plot files, and some miscellaneous examples of complicated searches. Although by no means complete, it includes examples of the search errors most frequently made by ourselves and our users.

a. Efficiency: The most expensive possible search that can be made involves reading every record in the STATIONS file, checking each against 20 parameters, and writing every record to the new file of selected stations. Such a search was accidentally performed by the authors and cost $36 in computer time. Typical searches usually cost about $1. This cost depends on the number of records that must be read, the number of search criteria to be checked, and the number of records selected, and can be minimized by careful planning of the search strategy.

If a user plans several separate searches of a particular category of stations, for example all stations within a specific area or all stations for which manganese nodule analyses are available, computer time and disk storage charges can be minimized by first searching the STATIONS file for the general data category and then making searches on this smaller file. For example, suppose one wished to prepare a separate list of stations for each Lamont-Doherty Geological Observatory cruise. In this case it would be less costly to first search the STATIONS file for INSTITUTION = LDGO and then search this new file for each individual reference number than to search the entire STATIONS file separately for each reference number.

The STATIONS records are stored in order of Sequence Number, and the file contains an index with the beginning and ending record numbers for each ten-degree square. When latitude or longitude is a search parameter, the program calculates which squares could contain the desired data and searches only the records in those squares. If the specified latitude or longitude boundaries are also the boundaries of ten-degree squares, the program will select stations from the appropriate squares (providing they meet any other specified search criteria) without specifically checking the latitude or longitude for each record. Thus a search for LAT = 10 N TO 20 N, LON = 140 W TO 100 W will take less computer time than a search for LAT = 11 N TO 19 N, LON = 139.9 W TO 100.2 W, even if the same number of stations are selected.

Files made from the STATIONS file contain no index. Therefore, if one wishes to search for several different areas, it is cheaper to search the entire STATIONS file each time than to search a large file made from the STATIONS file.

Search parameters are evaluated in the order they are entered, and only enough are checked to determine if the record meets or fails to meet the criteria. Thus computer time can be minimized by entering parameters in the optimum order. For example, many more sediment samples contain clay than contain chert. If samples containing both are to be selected, a search for ROCK CHERT, CLAY would be quicker than for CLAY, ROCK CHERT because most descriptions could be eliminated after checking only the first

parameter. On the other hand, the latter entry would be more efficient if samples containing either but not necessarily both are sought.

b. Dangers to Avoid: When performing searches involving data from more than one file, one. must keep in.mind the linked file structure. The STATIONS file contains pointers (links) to records in the NAR and DESCRIPTIONS files, and NAR contains pointers to records in the XTRA file as well as to other records in the NAR file itself. Thus after selecting a subset of stations records one can search NAR or DESCRIPTIONS for records pertaining to those stations. To do this, however, one must search the entire NAR or DESCRIPTIONS file because the pointers refer to the actual location of the appropriate record in the original file. Record numbers or addresses are not part of the records themselves. If one attempts to search a subset of the NAR file with stations to be selected from those in some stations file, he will receive the error message LINK TO NAR IS WRONG FOR <SEQUENCE NUMBER>. The user should then terminate the search by entering ?DS .

Although stations records are stored in order of acquisition within each ten-degree square, records selected by SEARCH are sorted before being listed or saved so as to be ordered from east to west within each square. The other files are stored in random order, so it is sometimes necessary to search the STATIONS file first even though none of the information from that file is of interest. For example, if a user wished to select analyses from the entire NAR file but list them in some logical order (east-to-west within each ten-degree latitude band), it would be necessary first to search the STATIONS file for all records with associated analyses (search parameter NAR) and then to search the NAR file specifying stations to be selected from the file just made.

The files contain no reverse pointers. Thus if one searches the NAR or DESCRIPTIONS file or a subset of one of them without reference to any stations file (answer ALL to the question STATIONS TO BE SELECTED FROM THOSE IN WHAT FILE?) it is impossible to go backwards and construct a file of stations to go with the selected records.

It is important to think through the logic of the operators before attempting a search. In particular, it must be remembered that a search for A & B does not select records for which A is true and records for which B is true, but only those records for which both A and B are true. For example, suppose one wished to select all stations with Reference Number A336 as well as all stations with Reference Number A288. The proper search would be REF = A336 OR A288. Since there is only one single Reference Number for each station, a search for REF = A336 & A288 would select no stations.

Up to 20 separate criteria may be specified for each search, but each value specified counts as a new criterion. For example, REF = N126 OR A023 OR B521 would count as three criteria.

c. More on Symbols and Plot Files: If a symbol is to be assigned to a station record, this is done as soon as the parameter that determines the symbol is evaluated. As the search continues, a symbol may be replaced by another symbol. If the symbol is determined according to a parameter stored in the NAR or DESCRIPTIONS files, in which there may be more than one record for each station, the final symbol assigned to the station record will be that for the last analysis or description that meets the search criteria.

To assign a symbol from the DESCRIPTIONS file that represents only the top
layer of a sample, one must include SURFACE as a search parameter. To as-
sign a symbol that represents an average value for all selected nodule
analyses for each station, one must use Program NARCALC (see Section F-2).

It is often desirable to plot maps with each different symbol repre-
senting a range of values. To create a file for such a plot a typical
search might be

        BATHYMETRY GEQ 3000 SYMBOL 4 OR GEQ 2000 SYMBOL 3 OR GEQ
        1000 SYMBOL 2 OR > 0 SYMBOL 1;

Note the sequence of the data entry. If bathymetry (water depth) is
greater than or equal to 3000 the record will be selected and symbol as-
signed after checking only the first value. Therefore, the subsequent
check for bathymetry greater than or equal to 2000 will actually select
stations from depths in the range 2000 to 2999. The values cannot be en-
tered in the reverse order, i.e. BAT < 1000 SYMBOL 1 OR < 2000 SYMBOL 2...
unless one first searches for BAT > 0. Otherwise, stations from unknown
depths will be included in the first group.

Symbols are associated with all criteria between the occurrence of
the word SYMBOL and the last previous comma or symbol assignment. Thus
REF = A101 OR A102 SYMBOL 1 OR A262 OR N341 SYMBOL 2 would have the same
effect as REF = A101 SYMBOL 1 OR A102 SYMBOL 1 OR A262 SYMBOL 2 OR N341
SYMBOL 2. To assign symbols to only two of the references and select the
others but leave the symbol blank or as it was previously, it would be
necessary to enter REF = A102 SYMBOL 1 OR N341 SYMBOL 2 OR A101 OR A262.

d. Examples of More Complicated Searches: users frequently want to
select all stations at which manganese nodules have been observed. Since
nodules can be observed in photographs but not retrieved in a bottom sam-
pler at the same site, the following search parameters would be used:

        NODULES = PRESENT OR COV > 0;

The NAR file contains analyses of crusts, manganese-coated rock,
micronodules, etc., as well as true nodules. Some analyses are of uniden-
tified samples. Some nodule sections analyzed, such as half or quarter
nodules or cross sections, are generally representative of the bulk nodule.
Analyses of a nodule core or outer layer, however, are probably not rep-
resentative of the whole nodule. To select only representative bulk
analyses of true nodules, one could search the NAR file for

        TYP = 1, SEC LEQ 3 OR > 14 OR = 7;

(See Table 3 for the interpretation of these value codes.) If one wanted
to assume that unidentified samples are probably true nodules, at least if
their manganese content exceeded 5%, one could instead search for

        TYP = 1 OR NOT TYP & MN > 5, SEC LEQ 3 OR. . . ;

Here NOT TYP means "type unknown"; manganese content is evaluated for only
samples of unknown type, since the & operator takes precedence over OR.

Some users wish to reclassify sediment samples according to lithological categories different from those used by the data bank. This is possible if the sediment description contains enough information, and it is usually accomplished by evaluating several parameters. For example, one user requested maps showing surface sediment lithologies according to a scheme which differed from ours in the following respects: (1) calcareous ooze was to be divided into two separate categories, 30-60% $CaCO_3$ and >60% $CaCO_3$; (2) clay was to include sediments we normally classify as clay plus mud with at least 50% clay-sized fraction; and (3) mud was to include samples we classified as mud unless the clay fraction was more than 50%, and in addition mud was to include any sample with turbidites reported. The search that accomplished this task was done with the following:

```
SURFACE, TURBIDITES SYMBOL 13 OR LITH = 7 SYMBOL 23 OR
LITHOLOGY < 3 SYMBOL 11 OR LIT = 5 & CACO3 GEQ 60 OR LIT = 5
& CACO3 = H OR LIT = 5 & FORAMS > 6 SYMBOL 15 OR LIT = 5 SYMBOL 10
OR LIT = 6 SYMBOL 12 OR LIT = 4 & CLAY > 50 SYMBOL 23 OR
LIT = 4 SYMBOL 13 OR LIT = 3 SYMBOL 24;
```

Note that parameter names may be spelled in full or abbreviated in any way as long as the first three characters are the correct code.

3. Searching the Bibliography: Program BIBSEARCH

a. **Purpose:** BIBSEARCH enables the user to search the bibliography files according to any parameter or combination of parameters that are stored in the BIBLIO/INDEX file. These include author's name, institution, year of publication (or year of cruise for shipboard logs), area from which samples were collected, and files in the data bank which contain data from the source. The program can also be used to list complete bibliographic information for any specified Reference Numbers.

The list of references obtained may be used as input to the SEARCH program in order to select from the data files information that was extracted from sources meeting the specified criteria.

b. **Use:** File BIBLIO/INDEX must be on disk before the program is run. If output is to include the complete bibliographic information, not just a list of Reference Numbers, File BIBLIO/TEXT must also be on disk.

The program may be run interactively, scheduled as a batch job from a remote terminal, or submitted through a card reader. Interactive jobs are run in the fastest and most expensive queue, and output is listed on the terminal almost immediately. Batch jobs are executed within about 10 minutes for Queue 2 to more than eight hours for Queue 5; output generally goes to a line printer but may go to disk storage (destination = SAVE) for subsequent listing on the terminal.

Workfile setup for scheduling BIBSEARCH from a remote terminal is as follows:

```
?JOB <JOBNAME>; CLASS = <QUEUE>

?TO = <DESTINATION>

?BEGIN

?RUN OBJECT/BIBSEARCH

?DATA

      <DATA RECORDS>

?END JOB
```

To run interactively, enter the following and wait for output:

```
E BIBSEARCH; FILE PRINTER(KIND=REMOTE); FILE READER(KIND=REMOTE)


      <DATA RECORDS>
```

c. **Data Input:** Search criteria are specified through a combination of parameters, values, and logical operators. Parameters or groups of parameters to be searched for must be separated by commas, and a semicolon and a carriage return must follow the last input. Up to 24 parameters are allowed for each search. Data records are free field and may take as many lines as needed. Input parameter names (but not values) will be truncated to the first three letters, although the user may enter the full word. The search parameters and their possible values are as follows:

Parameter	Values	Notes
AUThor	last name	References are indexed according to the name of the first author only. To make a complete search for a particular author use the Command-and Edit Language (see UCSD Computer Center Document #251, "CANDE Handbook"). For anonymous institution reports the name of the institution is given as author.
YEAr	a 4-digit number	
FILe	STAtions NAR (nodule analyses) DEScriptions	Program BIBSEARCH can provide an inexpensive listing of all references from which data have been extracted for a particular complete data bank file. To obtain a list of references for data that have been selected by SEARCH, use Program BIBLIOLIST (see Section E-5).
REFerence	a letter (A, M, B or N) followed by 3 digits	
AREa	NPA (North Pacific) SPA (South Pacific) NAT (North Atlantic) SAT (South Atlantic) ARCtic (everything north of 60°N) INDian MEDiterranean SEA (Southeast Asia area, 15°S - 45°N, 90-160°E)	
INStitution	2-4 characters	Selects reports from an institution, government agency or project. Most published papers are not identified with a particular institution.

Any of the following relational operators may precede a value: >, GEQ, <, LEQ, =, NEQ. If the operator is omitted, it is assumed to be =. At least one blank space must precede and follow each operator. For example, to select all references later than 1970 from which nodule analyses have been extracted, search criteria would be

    YEAR > 1970, FILE = NAR;

Parameters or values may be grouped by the use of the logical operators & and OR to select references that meet a combination of criteria. See page 21 for an explanation of the use of these operators. Examples:

    YEAR > 1965 & LEQ 1972;

    REF A132 OR A329, AREA NPA OR SPA;

38

Two additional input records are required after the record or records containing the search criteria. These are:

1) A record with the word TEXT starting in Column 1 if complete information for selected references is desired, or LIST if only a list of Reference Numbers is required.

2) A record with the word (or first three letters) ALPHABETICAL or · NUMERICAL starting in Column 1 to indicate the order in which the output is to be arranged.

d. Output: Depending upon the input, output is either a list of Reference Numbers or complete bibliographic information for each selected item as illustrated in Fig. 3.

e. Caution: All files on disk will remain there and incur disk storage charges until removed by the user.

A200 AHRENS, L. H., J. P. WILLIS AND C. O. OOSTHUIZEN, FURTHER OBSER-
VATIONS ON THE COMPCSITION OF MANGANESE NODULES WITH PARTICULAR
REFERENCE TO SOME OF THE RARER ELEMENTS. GEOCHIMICA ET COSMO-
CHIMICA ACTA, VOL 31, PP 2169-2180, 1967.

A215 BACON, JOAN IRENE, A GEOCHEMICAL STUDY OF SOME MANGANESE NODULES,
MA THESIS, DEPT OF CHEMISTRY, UNIV OF TULSA, OKLA, 1967.

A166 BARNES, S. S., THE FORMATION OF OCEANIC FERROMANGANESE NODULES.
UNPUBLISHED PHD DISSERTATION, UNIV OF CALIF, SAN DIEGO, 1967.

A365 BROOKE, J. N. AND A. P. PROSSER, MANGANESE NODULES AS A SOURCE OF
COPPER AND NICKEL MINERALOGICAL ASSESSMENT AND EXTRACTION. INST
OF MINING AND METALLURGY, SEC C, VOL 78, C63-C72, 1969.

A322 CALVERT, S. E. AND N. B. PRICE, GEOCHEMICAL VARIATION IN FERROMAN-
GANESE NODULES AND ASSOCIATED SEDIMENTS FROM THE PACIFIC OCEAN.
MARINE CHEMISTRY, VCL 5, NO 1, FEB 1977.

A189 CRONAN, D. S., AND J. S. TOOMS, THE GEOCHEMISTRY OF MANGANESE
NODULES AND ASSOCIATED PELAGIC DEPOSITS FROM THE PACIFIC AND INDIAN
OCEANS. DEEP-SEA RESEARCH, VOL 16, PP 335-359, 1969.

A381 FUERSTENAU, D. W., METAL RECOVERY FROM MANGANESE NODULES. OFFICE
OF RESEARCH SERVICES, COLLEGE OF ENGINEERING, UNIV OF CALIF,
BERKELEY, JUNE 1966.

Fig. 3. Example of output from Program BIBSEARCH or BIBLIOLIST.

E. DATA DISPLAY

This section describes the use of programs for listing and mapping information selected from the data bank. Additional display formats such as point plots or histograms are discussed in Section F.

Instructions are given for scheduling these programs from a remote terminal. Depending on the destination specification, this causes the output either to be printed on the specified line printer or saved on disk for later listing on a line printer or remote terminal. The programs may also be run interactively, with output appearing immediately on the terminal. Unless the terminal has a record length of 132 characters, output will be truncated. Instructions for running the program interactively or listing the output of scheduled jobs on a terminal can be obtained from the data bank staff.

· 1. Program LISTSTA

a. **Purpose:** LISTSTA allows the user to obtain a line printer listing of the STATIONS file or another file with the same format which has been produced by a search of the STATIONS file.

b. **Use:** The program may be submitted through a card reader, scheduled from a remote terminal, or run interactively. If the file to be listed is not the STATIONS file itself, the file title must be specified in a label equation record. Workfile setup for scheduling LISTSTA from a remote terminal:

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?PRINTLIMIT = <NUMBER OF LINES>        Use only if number of lines to be
                                       printed exceeds the default value
                                       for the queue (see Caution below).
?BEGIN
?RUN OBJECT/LISTSTA
?FILE STATIONS(TITLE = <FILENAME>)     Omit if STATIONS is the file to
?DATA                                  be listed
     <OPTIONAL DATA RECORD>
?END JOB
```

c. **Data Input:** The ?DATA record is necessary. After this record there are three options:

1) If there is no data record or if a zero is entered in the data record (in any column), the output will list only the sequence number, ship or cruise and station number, latitude, longitude, sampling device, water depth, core length, institution, reference number, and surface lithology.

41

2) If the digit "1" is entered in any column, information about manganese nodule occurrence, whether nodules occurred within 10 cm of the sediment surface (S) or below (B), coverage and concentration estimates for manganese nodules and instrument from which estimates of coverage and concentration were made will be printed after the station information listed above.

3) If a "2" is entered in the data record, the above information plus the station's links to other files in the data bank will be listed. In addition, if the stations file being listed is the result of a search in which a SYMBOL was assigned to the station, the value of the assigned SYMBOL will also be listed.

d. Output: Fig. 4 is an example of output for a job in which the data record contained the digit "1." A water depth or core length of zero means "unknown." Most output items are in English, but the surface lithology is indicated by the following codes:

> VA = volcanic ash (includes volcanic mud, volcanic sand)
> RO = rock
> MP = manganese pavement
> PC = pelagic clay
> SC = siliceous clay
> CC = calcareous clay
> SO = siliceous ooze
> CO = calcareous ooze
> CS = calcareous-siliceous ooze
> TM = terrigenous material (sand, silt, mud, gravel, clay)
> CT = calcareous terrigenous material
> ST = siliceous terrigenous material (siliceous mud)

e. Caution: Files will remain on disk and incur disk storage charges until removed by the user.

To determine print limit, add an additional seven lines of heading and spaces for each 100 lines of data.

SEQ. NUMBER	SHIP/CRUISE + STATION NO	LATI- TUDE	LONGI- TUDE	SAMPLING DEVICE	WATER DEPTH (M)	CORE LEV. (CM)	INST	REF. NO.	SUF LITH	MN NODULES % OCC	COVG	FROM	KG/M2
1180040	PC12-58	9.833N	136.383W	PISTON CORE	4823	721	LDGO	A114	SC	NAP SB	20-50	FOTO	
1180127	VEMA28-177TW	4.817N	136.450W	TRIP CORE	4709	58	LDGO	A266					
1180126	VEMA28-177	4.817N	136.450W	PISTON CORE	4709	1538	LDGO	A266	CO	NO			
3380041	ALBATRO.15	4.583N	136.900W	DREDG	4727		SIO	A009	CO	NO			
3380042	ALBATRO.14	6.683N	137.000W	DREDG	5080		SIO	A009	CO	NO			
3380043	ALBATRO.16	2.633N	137.167W	DREDG	4465		SIO	A009	CO	NO			
1380129	VEMA28-178TW	4.750N	137.367W	TRIP CORE	4409	28	LDGO	A266					
1380128	VEMA28-178	4.750N	137.367W	PISTON CORE	4409	1569	LDGO	A266	CO	NO			
1380044	FC12-59	8.983N	137.683W	PISTON CORE	4660	86	LDGO	A114	PC	NAP SB			
3380045	ALBATRO.13	9.950N	137.783W	DREDG	4923		SIO	A009	CO	NAP			
3380046	MERO-2P-52	9.943N	137.802W	DREDG	4930		SIO	A189		NAP			
1380047	ALBATRO.17	0.833N	137.900W	DREDG	4507		SIO	A009	CO	NO			
3380153	V405-143GB	6.750N	138.180W	FREEPA GRAB				A332		YES	<20	FOTO	
3380048	AMPH-136GV	0.400N	138.183W	GRAVTY CORE	4330	40	SIO	M005	CS	NO			
3380195	PLDS-107BX	6.157N	138.277W	BOX CORE	4849		SIO	A339					
3380049	VEMA 24-57	2.267N	138.600W	PISTON CORE	4252	1060	LDGO	A105	CO	NO	0	FOTO	
1380050	PC12-60	8.350N	138.783W	PISTON CORE	4956	1114	LDGO	A114	PC	NAP	0	FOTO	
1380051	JOIDES 72	0.433N	138.900W	DSDP	4336	34500		A008	CO	NO			
1380052	JYNV-43G	9.883N	138.933W	GRAVTY CORE	4803	82	SIO	M015	PC	NO			
1380131	VEMA28-179TW	4.617N	139.600W	TRIP CORE	4502	55	LDGO	A266					
1380130	VEMA28-179	4.617N	139.600W	PISTON CORE	4502	2100	LDGO	A266	CO	NO			
3380152	AMIGO-50	7.617N	139.667W	BOX CORE				A213		NAP	<20	BXCOR	
3380112	MN7402-15G24	9.217N	139.700W	FREEPA GRAB	5105	10	HIG	A242	SO	NAP		GRAB	2.12
3380111	MN7402-15G23	9.100N	139.750W	FREEPA GRAB	5040	10	HIG	A242	SO	NAP		GRAB	6.25
3380053	VITYAZ 5078	9.052N	139.755W	GRAVTY CORE	5064		USSR	A148	SC	NO			
1380150	MN7402-15B05	9.000N	139.800W	BOX CORE	5010		HIG	A242	SO	NO			
3380110	MN7402-15G22	8.933N	139.813W	FREEPA GRAB	5175	10	HIG	A242		YES		GRAB	0.10
3380109	MN7402-15G21	8.850N	139.813W	FREEPA GRAB	5085	10	HIG	A242	SO	NAP		GRAB	14.50
3380114	MN7402-15G26	9.000N	139.850W	FREEPA GRAB	5031	10	HIG	A242	CO	NAP		GRAB	15.25
1380054	JYNV-42P	9.000N	139.850W	PISTON CORE	5000	496	SIO	M015	PC	NO			
3380055	VITYAZ 5080	8.108N	139.880W	GRAB	5104		USSR	A148	SC	NO			
3380056	PC11-D21	8.783N	139.883W	DREDG	5086		LDGO	A130		NAP			
3380057	PC11-206	8.783N	139.883W	PISTON CORE	5086	831	LDGO	A116	SC	NAP S			
3380113	MN7402-15G25	8.950N	139.883W	FREEPA GRAB	4975	10	HIG	A242	SO	NAP		GRAB	3.75
1380149	MN7402-15B04	8.933N	139.900W	BOX CORE	4940		HIG	A242	SO	NO			
1380058	PC11-207	7.400N	139.933W	PISTON CORE	5091	1440	LDGO	A116	PC	NO			
1380060	MIDPAC10-C1	4.588N	139.962W	GRAVTY CORE	4330	90	SIO	M038	CO	NO			
3380061	PC11-208	5.350N	139.967W	PISTON CORE	4720	1636	LDGO	A116	CO	NO	0	FOTO	
1380062	VITYAZ 5082	5.952N	139.972W	GRAB	4842		USSR	A148	CS	NO			
1380063	VITYAZ 5082	5.952N	139.972W	GRAVTY CORE	4842		USSR	A148	CS	NO			
1180064	REHOBBS17-18	2.000N	139.983W	GRAB				B736	CO				

Fig. 4. Line printer listing of station information produced by Program LISTSTA. The search which selected data for this listing is illustrated in Fig. 2.

2. Program LISTNAR

a. **Purpose:** LISTNAR allows the user to obtain a line-printer listing of complete or partial nodule analyses from the NAR file or another file with the same format produced by a search of the NAR file.

b. **Use:** The program may be submitted through a card reader, scheduled from a remote terminal, or run interactively. If the file to be listed is not the NAR file itself, the file title must be specified in a label equation record (see below).

Only the most frequently reported information for each nodule analysis is stored in the NAR file or subsets of this file produced by SEARCH. The rest of the information is stored in a file titled XTRA. If a listing of complete nodule analyses is desired, File XTRA must be on disk before execution of LISTNAR. Otherwise, only the information shown in Fig. 5 will be listed.

Workfile setup for scheduling LISTNAR from a remote terminal:

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?PRINTLIMIT = <NUMBER OF LINES>      Required only if number of lines to
                                     be printed exceeds the default value
                                     for the queue. See Caution below.
?BEGIN
?RUN OBJECT/LISTNAR
?FILE PRINT2(KIND = DISK)            Omit if complete listing is desired
?FILE NAR(TITLE = <FILENAME>)        Omit if NAR is the file to be
                                     listed
?DATA
      <OPTIONAL DATA RECORD>
?END JOB
```

c. **Data Input:** The ?DATA record is necessary. After this record there are three options:

1) If listing is to include only information from the NAR file and not from XTRA, as shown in Fig. 5 (except for morphology), no data record is allowed.

2) For a listing of information in the NAR file plus morphology but no additional elements, a data record with the numeral "1" in any column is required. The label equation record shown above for File PRINT2 is also required, and XTRA must be on disk.

3) For a complete listing of all elements stored, which will be in the form of two separate lists, a data record with the numeral "1" is required and the label equation record for File PRINT2 must be omitted. File XTRA must be on disk.

d. **Output:** Output will include one or two lists. One list (from NAR) will contain sample size and type, analytical method, description of the nucleus, and the elemental concentrations in weight percent of the elements Mn, Fe, Co, Ni, Cu, Zn, Pb, Al, Si, Ca and H_2O (Fig. 5). A value

45

of zero for any element indicates that the concentration of that element was not reported. Blanks in the columns for descriptive information indicate unreported information.

The other list (from File XTRA), which is produced only under data option 3 above, will contain the depth in the core from which the sample was obtained plus additional element concentrations which are identified by their atomic number (Fig. 6).

The following tables will aid in decoding the listings:

Analytical Method

WC = Wet Chemical
XF = X-ray Fluorescence Wavelength Spectrometry
SE = X-ray Fluorescence Energy Spectrometry
EM = Electron Microprobe
AA = Atomic Absorption
NA = Neutron Activation
ES = Emission Spectrometry
RF = see reference--other methods and combination of methods

Primary Morphology

S = Spherical
E = Ellipsoidal
D = Discoidal
P = Poly
B = Biological
T = Tabular
F = Faceted

Surface Texture

S = Smooth
R = Rough (granular or microbotryoidal)
B = botryoidal

When two letters are listed on the same side of the slash under morphology, it indicates a grading from the most apparent to the next most apparent type. For example, DE/RS would be read as discoidal to ellipsoidal with a rough to smooth surface texture.

Each line of output is identified by the Sequence Number of the station at which the nodule sample was collected. To obtain the latitude, longitude and other station information for each analysis listed, run Program LISTSTA for the stations file produced along with the analysis file and match the Sequence Numbers.

e. Caution: File XTRA must be on disk when LISTNAR is run if the data record contains the value "1". Allow one extra line of heading and spacing for each three records to be listed. Double the print limit if complete analyses are to be printed.

Files will remain on disk and incur disk storage charges until removed by the user.

SEQ NO.	SAMP. TYPE & SECT SAMPLED	DIMENSIONS CORE	MORPH	PEF	MN	FE	CO	NI	CU	ZN	PB	AL	SI	CA	H2O
3380040 0	NODULE			A179 WC	21.80	5.90	0.27	1.20	0.94	0.000	0.00	0.00	0.00	0.00	0.00
3380044 0	NODULE			A179 WC	2.50	8.20	0.05	0.11	0.11	0.000	0.00	0.00	0.00	0.00	0.00
3380045 1	NODULE X-SECTION	160X100X100		A197 XF	29.80	4.80	0.20	1.36	1.20	0.120	0.06	3.02	6.08	1.47	16.20
3380045 2	NODULE X-SECTION			A165 XF	48.50	3.38	0.15	2.46	1.54	0.000	0.12	0.00	0.00	0.00	0.00
3380045 3	FRAGMENTS	25X25X25		A261	32.00	5.30	0.18	2.00	1.40	0.040	0.05	2.70	7.50	1.80	0.00
3380045 4	NODULE X-SECTION			A166 XF	31.50	0.00	0.12	1.53	0.92	0.000	0.07	0.00	0.00	0.00	0.00
3380045 5	NODULE, INNERMOST LAYER			A201 FS	0.00	6.20	0.26	1.60	1.70	0.000	0.00	0.00	6.20	0.00	0.00
3380045 6	FRAGMENTS	30X30X20		A200 XF	38.10	4.43	0.25	1.70	1.21	0.141	0.05	0.00	0.00	1.66	0.00
3380045 7	NODULE, INNERMOST LAYER	100		A197 XF	30.70	2.30	0.11	1.37	1.34	0.000	0.00	0.00	6.90	1.23	12.70
3380045 8	NODULE, MIDDLE LAYER	100		A197 XF	29.10	3.90	0.16	1.26	1.00	0.000	0.00	0.00	6.90	1.46	15.90
3380045 9	NODULE, MIDDLE LAYER	100		A197 XF	27.70	5.30	0.22	1.33	1.00	0.000	0.00	0.00	5.70	1.56	14.90
338004510	NODULE, OUTER LAYER	100		A197 XF	32.10	3.20	0.20	1.05	0.77	0.000	0.00	0.00	5.80	1.40	11.50
338004511				A215	30.00	1.50	0.08	1.10	0.25	0.163	0.03	1.29	20.34	0.52	0.00
338004512	NODULE			A262 ES	0.00	7.00	0.30	1.50	1.50	0.070	0.07	3.00	7.00	3.00	0.00
338004513	NODULE FRAGMENTS	30X20X10		A388 XF	34.31	5.30	0.26	1.62	1.15	0.175	0.03	0.00	4.49	1.46	0.00
338004514		30X30X20		A388 XF	32.66	4.46	0.22	1.64	1.13	0.167	0.03	0.00	4.67	1.33	0.00
				A185	23.46	5.41	0.14	1.93	1.59	0.000	0.04	0.00	0.00	0.00	0.00
3380046 0	NODULE														
3380046 1	NODULE		T/	A355 AA	22.00	4.60	0.14	1.30	1.04	0.000	0.00	0.00	0.00	0.00	0.00
3380046 2	NODULE			A322 XF	19.60	5.80	0.17	0.97	0.92	0.090	0.04	2.90	8.90	1.70	0.00
3380046 3	NODULE W/O NUCLEUS			A346 RF	27.80	4.10	0.17	1.44	0.95	0.000	0.00	0.00	0.00	0.00	0.00
3380046 4	NODULE COMPOS. >12			A370 AA	25.60	5.00	0.18	1.50	1.20	0.000	0.00	0.00	0.00	0.00	0.00
3380046 5	NODULE			A381 EM	25.70	5.10	0.15	1.90	0.95	0.150	0.00	0.00	0.00	1.25	0.00
3380046 6	NODULE			A381 EM	25.50	4.80	0.20	1.75	1.15	0.150	0.00	0.00	0.00	1.30	0.00
3380152 0	NODULE	NONE	/ P	A326 FF	19.10	5.08	0.19	1.40	1.09	0.125	0.01	0.00	10.84	0.00	0.00
3380152 1	NODULE	NONE	/ P	A326 FF	25.40	5.13	0.21	1.71	1.34	0.152	0.02	0.00	9.16	0.00	0.00
3380112 0	NODULE			A242 XF	24.68	7.67	0.00	1.70	0.87	0.000	0.00	0.00	0.00	0.00	0.00
3380111 0	NODULE			A242 XF	25.49	7.66	0.00	1.55	0.98	0.000	0.00	0.00	0.00	0.00	0.00
3380111 1	NODULE W/O NUCLEUS			A242 XF	26.54	8.54	0.00	1.80	1.03	0.000	0.00	0.00	0.00	0.00	0.00
3380111 2	NODULE W/O NUCLEUS			A242 XF	26.81	7.53	0.00	1.82	1.00	0.000	0.00	0.00	0.00	0.00	0.00
3380111 3	NODULE W/O NUCLEUS			A242 XF	28.72	8.35	0.00	1.87	1.07	0.000	0.00	0.00	0.00	0.00	0.00
3380109 0	NODULE, HALF			A244 WC	30.70	4.65	0.00	1.67	1.52	0.000	0.00	0.00	0.00	1.03	30.15
3380109 1	NODULE, HALF			A244 XF	31.70	4.54	0.00	1.55	1.39	0.000	0.00	0.00	0.00	1.12	29.91
3380109 2	NODULE W/O NUCLEUS			A242 XF	30.06	6.65	0.00	1.72	1.28	0.000	0.00	0.00	0.00	0.00	0.00
3380109 3	NODULE W/O NUCLEUS			A247 XF	31.94	6.37	0.00	1.64	1.19	0.000	0.00	0.00	0.00	0.00	0.00
3380109 4	NODULE W/O NUCLEUS			A242 XF	27.47	6.51	0.00	1.59	1.14	0.000	0.00	0.00	0.00	0.00	0.00
3380109 5	NODULE W/O NUCLEUS			A247 XF	32.02	7.26	0.00	1.73	1.17	0.000	0.00	0.00	0.00	0.00	0.00
3380114 0	NODULE			A242 XF	32.47	5.94	0.00	1.85	1.20	0.000	0.00	0.00	0.00	0.00	0.00
3380056 0	NODULE			A179 WC	6.00	11.60	0.20	1.28	1.32	0.000	0.00	0.00	0.00	1.40	0.00
3380113 0	NODULE, HALF			A244 XF	26.20	6.36	0.00	1.45	1.61	0.000	0.00	0.00	0.00	0.91	33.94
3380113 1	NODULE, QUARTER			A244 XE	31.20	4.00	0.00	1.60	1.52	0.000	0.00	0.00	0.00	1.06	30.70
3380113 2	NODULE			A242 XF	22.56	9.62	0.00	1.27	0.91	0.000	0.00	0.00	0.00	0.00	0.00
3380113 3	NODULE, WHOLE			A334 AA	32.90	4.80	0.23	1.44	1.41	0.182	0.01	2.70	0.00	0.00	0.00
3380057 0	NODULE			A174 WC	16.80	6.70	0.26	0.94	0.98	0.000	0.00	0.00	0.00	0.00	0.00

NOTE: ALL DIMENSIONS ARE IN MILLIMETERS EXCEPT THOSE MARKED BY AN ASTERISK (*) WHICH ARE IN MICRONS.

Fig. 5. Line printer listing of nodule analyses produced by Program LISTNAR. The search which selected data for this listing is illustrated in Fig. 2.

47

```
SEQ NO.  Z  WT%   Z  WT%   Z  WT%   Z  WT%   ?  WT%   Z  WT%   Z  WT%   ?  WT%   Z  WT%   Z  WT%   Z  WT%   Z  WT%

3380045 1 22 0.4400 42 0.0540 15 0.0520 19 0.7900 38 0.0700 56 0.0100
3380045 3 22 0.4300 42 0.0700  5 0.0140 11 3.1000 12 1.7000 19 1.1000 21 0.0008 23 0.4200 24 0.0006 31 0.0007 38 0.0590 39 0.0110
          40 0.0350 56 0.2800 70 0.0015
3380045 5 22 1.4000 42 0.1100 12 2.7000 23 0.0520 81 0.0085
3380045 6 48 0.0021 81 0.0115 83 0.0006

3380045 7 22 0.3300 38 0.0510 56 0.5600
3380045 8 22 0.5200 38 0.0510 56 0.4600
3380045 9 22 0.6100 38 0.0560 56 0.4500
338004510 22 0.3400 38 0.0510 56 0.3300
338004511 22 0.2600 42 0.0950 11 6.5535 12 2.0000 19 1.0800 23 0.0200 38 0.0030

338004512 22 0.7000 42 0.0700  4 0.0002  5 0.0070 11 3.0000 12 1.5000 19 1.5000 21 0.0015 23 0.0300 24 0.0015 38 0.1500 39 0.0150
          40 0.0300 41 0.0030 56 0.1500 57 0.0150 58 0.0700 60 0.0300 70 0.0015 81 0.0150
338004513 42 0.0857 15 0.1004 16 1.0000 31 0.0020 33 0.0090 38 0.0551 39 0.0094 40 0.0270 41 0.0021 81 0.0319
338004514 42 0.0862 15 0.0480 16 0.0000 31 0.0028 33 0.0075 38 0.0532 39 0.0097 40 0.0270 41 0.0023 81 0.0115
3380046 0 22 0.3200 42 0.0430 23 0.0430 24 0.0015 56 0.4630

3380046 2 22 0.4100 42 0.0580  6 0.1700 12 1.1000 15 0.1300 19 1.0400 19 1.0400
3380046 3 22 0.2700 42 0.0940 21 0.0000
3380046 5 22 0.4500 19 0.6200
3380046 6 22 0.6000 19 0.6500
3380112 0 22 0.3600

3380111 0 22 0.2800
3380111 1 22 0.3100
3380111 2 22 0.3300
3380111 3 22 0.3800
3380109 0 22 0.3200

3380109 1 22 0.3200
3380109 2 22 0.2300
3380109 3 22 0.2200
3380109 4 22 0.2800
3380109 5 22 0.2400

3380114 0 22 0.2700
3380056 0 93 0.0000
3380113 0 22 0.4100
3380113 1 22 0.3000
3380113 2 22 0.3200

3380113 3 42 0.0720 24 0.0063
```

Fig. 6. Line printer listing of additional analytical information from File XTRA produced by Program LISTNAR at the same time as the listing shown in Fig. 5.

48

3. Program LISTDESCRIP

   a. <u>Purpose</u>: LISTDESCRIP allows the user to obtain a line-printer listing of the DESCRIPTIONS file or another file with the same format produced by a search of the DESCRIPTIONS file.

   b. <u>Use</u>: The program may be submitted through a card reader, scheduled from a remote terminal, or run interactively. If the file to be listed is not the DESCRIPTIONS file itself, the file title must be specified in a label equation record as shown below. Workfile setup for scheduling LISTDESCRIP from a remote terminal is as follows:

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?PRINTLIMIT = <NUMBER OF LINES>        Required only if number of lines
                                       to be printed exceeds the de-
                                       fault value for the queue (see
                                       Caution below).
?BEGIN
?RUN OBJECT/LISTDESCRIP
?FILE DESCRIPTIONS(TITLE = <FILENAME>)  Omit if DESCRIPTIONS is the
?END JOB                                file to be listed
```

   c. <u>Output</u>: Fig. 7 shows a sample output from this program. Each description is identified by the Sequence Number of the station at which the sediment was collected. To obtain the latitude, longitude and other station information for each description listed, run Program LISTSTA for the stations file produced along with the analysis file and match the Sequence Numbers.

   d. <u>Caution</u>: Because output is in complete words, each record may require from one to ten lines depending on the amount of information stored. In setting the print limit, figure an average of four or five lines per description record.

   Files will remain on disk and incur disk storage charges until removed by the user.

SEQ. NUMBER	INTERVAL IN CORE	LITHOLOGY	PERCNT CACO3	CONSTITUENTS BY GRAIN SIZE	ORGANIC MATERIAL OR FOSSILS	OTHER FEATURES	MINERALS
3380040	0-70	CLAY	000	<10% ROCK MN NODULES 20% SAND SILICEOUS >50% CLAY SILICEOUS PELAGIC	DIATOMS 10-30% RADIOLARIA 10-30%	COLOR: PALE BROWN BROWN BURROWS, MN NODULES	
3380040	70-239	CLAY	000	20% SAND SILICEOUS >50% CLAY SILICEOUS PELAGIC	FORAMS PRESNT RADIOLARIA 10-30% ANIMAL DEBRIS	COLOR: BROWN GRAYISH BROWN TEXTURE: STIFF BURROWS, MN MICRONODULES	
3380040	239-590	CALC. OOZE		30% SAND SILICEOUS CALCAREOUS 10-50% CLAY CALCAREOUS SILICEOUS	FORAMS 10-30% RADIOLARIA 10-30% SPINES, SPICULES ANIMAL DEBRIS	COLOR: BROWN LIGHT BROWN TEXTURE: STIFF BURROWS	
3380040	590-721	CALC. OOZE		<10% ROCK MN NODULES 40% SAND SILICEOUS CALCAREOUS 10-50% CLAY CALCAREOUS SILICEOUS	FORAMS 10-30% RADIOLARIA 10-30%	COLOR: BROWN TEXTURE: STIFF MN NODULES	
3380041	?-?	CALC. OOZE	065	20% CLAY SILICEOUS PELAGIC	FORAMS 30-50% COCCOLITHS PRESNT DIATOMS PRESNT RADIOLARIA 1-10% SPINES, SPICULES MACROFAUNA SHELLS	COLOR: PALE BROWN	FELDSPAR GLASS
3380042	?-?	CALC. OOZE	044	40% CLAY PELAGIC SILICEOUS	FORAMS 30-50% COCCOLITHS PRESNT DIATOMS 1-10% RADIOLARIA 1-10% SPINES, SPICULES	COLOR: PALE BROWN	FELDSPAR GLASS
3380043	?-?	CALC. OOZE	071	20% CLAY PELAGIC SILICEOUS	FORAMS 30-50% COCCOLITHS PRESNT DIATOMS 1-10% RADIOLARIA 1-10% SPINES, SPICULES	COLOR: BROWNISH GRAY	FELDSPAR GLASS
3380128	0-10	CALC. OOZE	HI		FORAMS 30-50% DIATOMS 1-10% RADIOLARIA 1-10%	COLOR: PALE BROWN TEXTURE: STIFF	
3380128	10-80	CALC. OOZE	MOD		FORAMS 30-50% DIATOMS 1-10% RADIOLARIA 1-10%	COLOR: PALE BROWN TEXTURE: STIFF BURROWS	

Fig. 7. Line printer listing of sediment descriptions produced by Program LISTDESCRIP. The search which selected data for this listing is illustrated in Fig. 2.

50

4. Program MAP

a. <u>Purpose</u>: MAP is a program for plotting maps of data previously selected from the data bank files. It is based on UCSD SUPERMAP by R. L. Parker.

b. <u>Use</u>: This program may be submitted through a card reader, scheduled from a remote terminal, or run interactively.

Before running MAP, the user must use the SEARCH program to prepare a disk file containing stations for which symbols are to be plotted. If different symbols are to be used to represent different stations in the same file, these symbols must have been previously assigned by SEARCH, NARCALC, or some other program. (Sections D-2 and F-1 contain instructions for producing a file of data to be mapped.)

c. <u>Input Data</u>: Data records are free form and consist of a series of keywords. These keywords either cause a specific action to occur and stand alone, or else they determine a value and must be followed by an appropriate value. Entries must be separated by at least one space; commas and equal signs may be used if desired, although they will be ignored by the program. Data may be continued on as many cards or terminal records as necessary, but each value must be on the same record as the keyword it follows. Data may be entered in any order, as limited by common sense (e.g., specify data needed to determine size and scale of the map before starting map, end one map before starting another, etc.). Only the first 71 characters of a record may be used for data.

The following keywords must be followed by numerical values. Entering them as data has the effect of assigning a value to the variable named. Note: map dimensions and symbol heights are expressed in inches. Latitudes, longitudes and rotation are in degrees.

IGRID
: Number of degrees between grid lines. First grid lines are at MINLAT and MINLON, with succeeding grid lines at intervals of IGRID. If it is important to have grid lines at latitudes and longitudes which are multiples of some specific number, such as 10, be sure MINLAT and MINLON are multiples of that number.

SYMHEIGHT
: Height in inches for symbols to be plotted on map. (This will be the actual height of the symbols only if the symbol number is 0 through 9 or if numerical values are to be plotted. Otherwise, the symbol will be half the height of SYMHEIGHT.) Symbols will have a more uniform appearance if SYMHEIGHT is a multiple of 0.08. The smallest legible symbols that can be plotted are 0.04.

IPROJ
: An integer which indicates the projection to be used: 1 = stereographic, 2 = orthographic, 3 = Mercator, 4 = Lambert equal area, 5 = gnomonic, 6 = azimuthal equi-distant, 7 = dummy, 8 = cylindrical equi-distant, 9 = transverse Mercator (Mercator with arbitrary pole), 10 = Molleweide, 11 = Mercator with correction for the ellipticity of the Earth, and 12 = Urmaev. The user is expected to know the properties of the projection; however, Table 7 summarizes the scaling information required.

51

TABLE 7. SCALING INFORMATION FOR MAP PROJECTIONS

IPROJ	NAME	SCALING INFORMATION	REMARKS
1	Stereographic	Rmax=1 when θ=90°	map origin opposite POLAT, POLONG
2	Orthographic	Rmax=1 when θ=90°	the near hemisphere plotted
3	Mercator	$\lvert U \rvert \leq \pi$, VMAX=∞, V=ln tan(45°+lat/2)	map origin always at equator; POLAT is ignored; POLONG must be within 180° of MINLON and MAXLON
4	Lambert	Rmax=2 when θ=180°	equal area projection
5	Gnomonic	Rmax=∞ when θ=90°, R=1 when θ=45°	the near hemisphere plotted
6	Azimuthal Equi-distant	Rmax=∞ when θ=180°, R=π/2 when θ= 90°	
7	Dummy		
8	Cylindrical Equi-distant	$\lvert U \rvert \leq$ 180°, $\lvert V \rvert \leq$90°, U=long., V=lat.	simple plot of data in degrees
9	Mercator with arbitrary pole	same as 3	POLAT, POLONG define new pole; origin is always on new equator and ROT degrees to the left of the true North Pole
10	Molleweide	$\lvert U \rvert \leq$2, $\lvert V \rvert$<1	equal area elliptic; POLAT is ignored; origin always on the equator
11	Mercator corrected for ellipticity of the Earth		map origin always at equator; POLAT is ignored. POLONG must be within 180° of MINLON and MAXLON
12	Urmaev Pseudocylindric		map centered on POLONG

— o —

R denotes the distance $(U^2 + V^2)1/2$ in the map projection (U,V) plane, and θ is the angle from the origin to the point on the Earth's surface. The map origin is at POLAT, POLONG except where indicated to the contrary. No corrections for ellipticity have been included except for IPROJ=11. For more information about the properties of the projections see the Encyclopedia Britannica World Atlas.

| MINLAT | Minimum latitude to be plotted on map | } | northern latitudes are positive; south- |
| MAXLAT | Maximum latitude to be plotted on map | } | ern, negative |

| MINLON | Leftmost longitude on map | } | Western longitudes are positive, from 0 to 180. Eastern longi- |
| MAXLON | Rightmost longitude on map | } | tudes may be expressed either as negative numbers (0 to -180 go- ing east) or as positive numbers (180 to 360 going west). |

POLAT)
POLONG) — Latitude and longitude of the map pole. See documentation of SUPERMAP (obtainable from the UCSD Computer Center) and Table 7.

ROT — Rotation; the angle between the V-axis and north at the map origin.

MAPHEIGHT, MAPWIDTH — In any given projection there is a fixed ratio between the height and width of a map of any part of the globe. Therefore, specify either MAPHEIGHT or MAPWIDTH, but not both.

UMIN

UMAX

VMIN

VMAX — Each map is conceived in a Cartesian coordinate system, the U,V plane, where U is along the paper and V across it, with the usual sign convention. The map origin (which may be at any location on or off the map paper) is referenced to some point on the Earth's surface and an orientation is also de- fined. The actual map that is plotted consists of all the material that appears in a "window" UMIN < U < UMAX and VMIN < V < VMAX, which need not include the origin. For many maps UMIN, UMAX, VMIN and VMAX can be calculated directly from MINLAT, MAXLAT, MINLON and MAXLON, and Program MAP will perform the calculation. In such cases it is not necessary to specify values for UMIN, etc. In other cases, including (1) all maps which have been rotated so that north does not correspond to the positive V direction and (2) all maps including the North or South Pole within their limits, these values must be cal- culated and provided as data. In these cases it is necessary to know the range in the units of the projection plane to specify the area required (see Table 7).

ISYM — Determines the symbols to be plotted on the map. If ISYM is a positive number, it represents the number of the symbol from the Symbol Table, Fig. 8, which will be plotted for all data points. Example: if ISYM is assigned a value of 13, all data points will be designated by the symbol "+" until a new value is assigned to ISYM.

If ISYM = -1, each point will be represented by a "+" followed by the value previously assigned as a symbol value to that data record expressed as a real number. Example: ISYM = -1, and the map is to plot weight percentages of Ni + Cu. If the value for a data point is 1.48%, then it will be plotted as +1.5 or +1.48 or +1.480 depending on the value of PLACES.

If ISYM = 0, each point will be represented by the symbol from the Symbol Table (Fig. 8) assigned to it by SEARCH. This means that the symbols plotted will vary according to the data stored in the stations file.

Fig. 8. Symbol Table.

PLACES   Required only if ISYM = -1. Specifies the number of place values after the decimal that will be plotted when real number values are plotted.

Values for certain variables are stored in the program. MINLAT, MAXLAT, MINLON and MAXLON are set so as to include the entire world unless new values are entered. Other default values are:

MAPHEIGHT = 10 (inches)
IGRID = 10 (degrees)
POLAT = 0 (degrees)
ROT = 0 (degrees)
PLACES = 2
SYMHEIGHT = 0.08 (inches)
ISYM = 13 (+)

54

The following keywords, some of which must be followed by values, cause specific action to be taken:

STARTMAP Starts map; this is required for plotting each map, and must be entered after all scaling information is entered.

FRAME Draws a rectangular frame around the map area.

GRID Draws grid lines on map at intervals of IGRID degrees. If IGRID has not been specified previously, grid lines will be at 10^0 intervals.

COAST Draws coastlines on the map from the UCSD WORLDMAP coastline data.

TITLE Puts a title at the top of the map and writes "Prepared by S.I.O. Sediment Data Bank" at the bottom. TITLE must be followed by a real number specifying the symbol height for the title, credit line, and label (if any). Then, on a separate record, enter the words to be printed as the map title, up to 72 characters.

LABEL Puts a map key (legend) at the lower right corner of the map. Must be followed by an integer indicating how many lines are to be in the label. Then, for each line of the label, a separate record with the symbol number (from Fig. 8) in the first two columns (right justified) and the explanation of the symbol in Columns 3-71.

POINTS Plots a point on the map for each station in the data file. Must be followed by the name of the file containing the points to be plotted. The file must be on disk and must have the same format as the STATIONS file. POINTS may appear more than once with different file names to plot information from more than one file on the same map (see Example 1).

ENDMAP Ends map; must appear for each map to be plotted.

RESCALE Indicates that the next map to be plotted has a different scale from the previous map and that the program needs to recalculate UMIN, UMAX, VMIN and VMAX from MINLON, MAXLON, MINLAT and MAXLAT. A new value for IPROJ automatically sets RESCALE. RESCALE must appear in the data before any data pertaining to the new map. Do not set RESCALE after entering values for UMIN, etc., as this will cause those values to be recalculated from MINLON, etc.

    RESCALE is used when the job includes plots of different sizes, different areas, or different scales. Enter RESCALE after ENDMAP. After entering RESCALE, the only values which are required to be entered are MAPHEIGHT or MAPWIDTH. Any other parameters may also be entered according to the specifications required for the next plot. See Example 2 for use of RESCALE.

STOP Signals the end of data; must follow the last ENDMAP.

55

d. <u>Output</u>: This program will produce a line printer summary of the input data along with the calculated scaling information and a disk file titled DISKPLOTS/<account number>/<plot number> for each map. Any number of map plot files may be produced in a single run of Program MAP.

No maps will be plotted by this program. To actually plot the files on the drum plotter, run Program PLOTREQUEST (UCSD Computer Center Document #063). To display plots on a CRT, run Program TEKPLOT (Computer Center Document #065).

<u>Example 1</u>. Below is the workfile setup for producing the map shown in Fig. 9. The map shows nickel content for stations with analyzed manganese nodules as a real number (set by ISYM -1) with two decimal places (using default value of PLACES). Data for these points is stored in File MAPNISTA. Other stations are designated by + if nodules were collected and by □ if no nodules were reported (set by ISYM 0). These stations are stored in File MAPNODSTA.

```
?JOB NIMAP;CLASS=5
?TO=SITE
?BEGIN
?RUN OBJECT/MAP
?DATA
IPROJ 3 SYMHEIGHT .12 MINLAT 10 MAXLAT 15 MINLON 135 MAXLON 130
POLONG 130 MAPWIDTH 6 STARTMAP FRAME TITLE .16
WT. % NICKEL IN MN NODULES:   10-15N, 130-135W
ISYM 0 POINTS MAPNODSTA ISYM -1 POINTS MAPNISTA ENDMAP STOP
?END JOB
```

WT. % NICKEL IN MN NODULES: 10-15N, 130-135W

```
                    +1.08
              +1.08
         +.44
      +1.43
           +            □
                          +1.63

                          .

                 □
            □
                     +1.72
         +1.34
    □ +
 □ +

                     +2.26
```
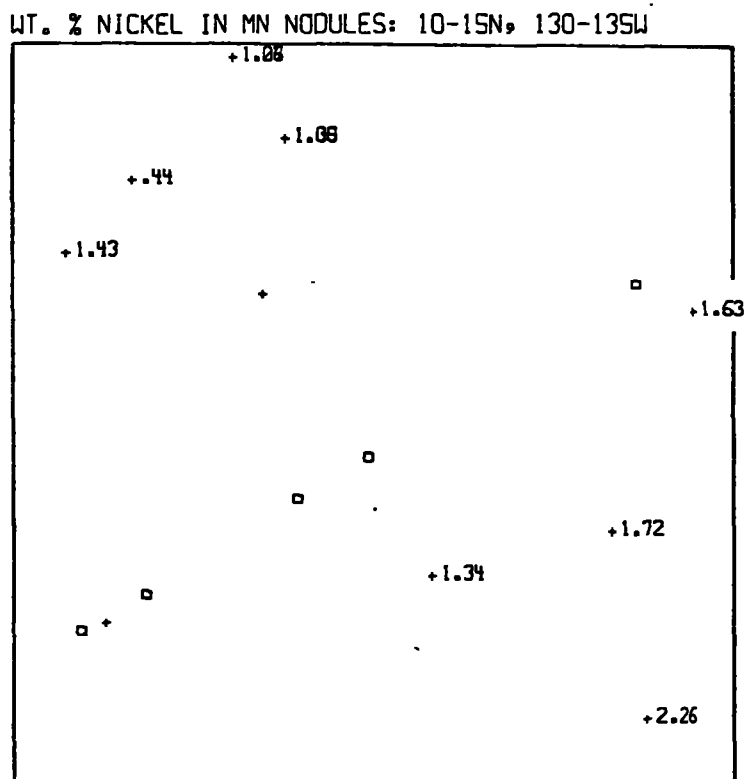
MAP PREPARED BY S.I.O. SEDIMENT DATA BANK

Fig. 9. Plot produced by Program MAP using data from
         Example 1. Mercator projection.

Example 2. The following is a workfile setup which plots three separate maps in one job, all utilizing the same stations subfile GULF. Note that IGRID, IPROJ and ISYM remain the same for each plot but that information for scaling has been changed because each map will be of a different area.

```
?JOB GULFMAP;CLASS = 5
?TO = RJS
?BEGIN
?RUN OBJECT/MAP
?DATA
IGRID 1 IPROJ 3 MINLAT 30 MAXLAT 30 MINLON 116 MAXLON 194
POLONG 116 MAPHEIGHT 29 ISYM 11 STARTMAP GRID COAST POINTS GULF
ENDMAP RESCALE MINLAT 20 MAXLAT 27 MINLON 114 MAXLON 105
MAPWIDTH 37.17 STARTMAP GRID COAST POINTS GULF ENDMAP RESCALE
MINLAT 27 MAXLAT 30 MINLON 116 MAXLON 109 MAPWIDTH 28.91
STARTMAP GRID COAST POINTS GULF ENDMAP STOP
?END JOB
```

Example 3. This workfile produces the map shown in Fig. 10, which uses the Lambert equal area projection.

```
?JOB CONMAP;CLASS = 3
?TO = RJS
?BEGIN
?RUN OBJECT/MAP
?DATA
IPROJ 4 SYMHEIGHT .08 MINLAT -90 MAXLAT 90 MINLON -110
MAXLON 70 POLONG 160 ISYM 14 MAPWIDTH 15 IGRID 10 STARTMAP
COAST GRID TITLE .16
STATIONS WITH MANGANESE NODULE CONCENTRATION MEASURED
POINTS CONDATA ENDMAP STOP
?END JOB
```

STATIONS WITH MANGANESE NODULE CONCENTRATION MEASURED
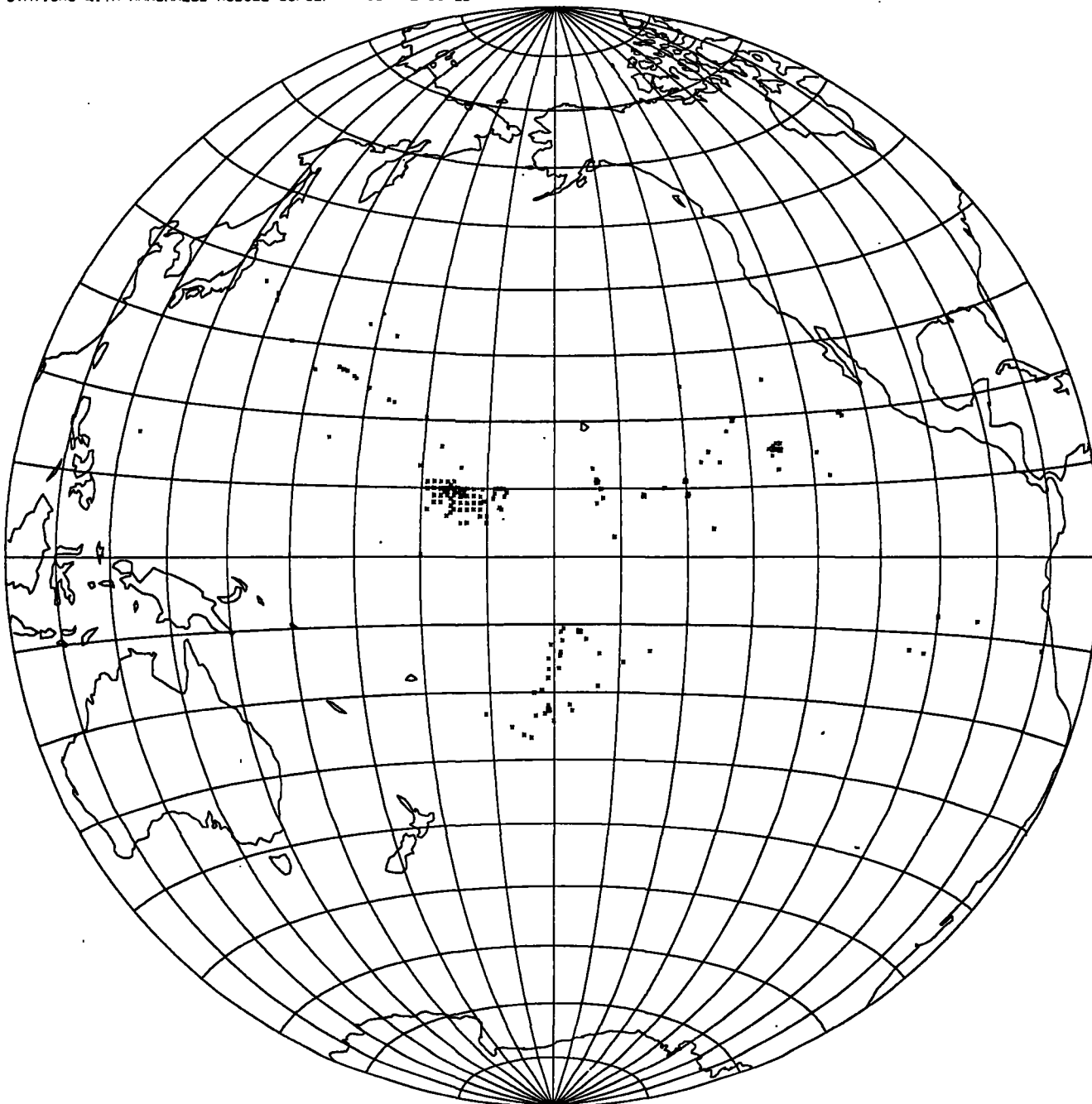
MAP PREPARED BY S.I.O. SEDIMENT DATA BANK

Fig. 10. Plot produced by Program MAP using data from
Example 3, Lambert equal area projection.

58

5. Program BIBLIOLIST

a. __Purpose:__ BIBLIOLIST provides the user with a listing of the bibliographic references for the entire STATIONS or NAR files or for other disk files which are subsets of one of.these files. Either a list of Reference Numbers or the complete bibliographic information may be obtained.

b. __Use:__ The program may be submitted through a card reader, scheduled from a remote terminal for listing to a line printer, or run interactively. File BIBLIO/INDEX must be on disk before the program is run. If complete bibliographic information is desired as output, File BIBLIO/TEXT must also be on disk.

The workfile setup for scheduling BIBLIOLIST from a remote terminal is:

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?BEGIN
?RUN OBJECT/BIBLIOLIST
?DATA
     <DATA RECORD>
?END JOB
```

To run interactively, enter the following:

```
E BIBLIOLIST;FILE PRINTER(KIND=REMOTE);FILE READER(KIND=REMOTE)
     <DATA RECORD>
```

c. __Input Data:__ Data to be entered consists of four words in the order specified, separated by one or more spaces, as follows:

<FILENAME>   name of file for which references are to be listed

<FILETYPE>   type of file, either STATIONS or NAR

NUM   if references are to be listed in numerical order, __or__
ALPHA   if references are to be listed alphabetically by author

LIST   if only a list of Reference Numbers is desired, __or__
TEXT   if complete bibliographic information is desired

For example, a typical data card (which was used to produce the list in Fig. 3) might be as follows:

DEMO/NAR NAR NUM TEXT

d. __Output:__ if the LIST option is selected, output is simply a list of Reference Numbers for which data are stored in the file for which the list was made. If the TEXT option is selected, output will be the complete bibliographic reference; Fig. 3 is an example of such output.

Some bibliography items contain cross references to additional sources of data for the same samples. To obtain these additional references, use Program BIBSEARCH (Section D-3).

e. __Caution:__ Files will remain on disk and incur storage charges until removed by the user.

F. FURTHER MANIPULATION OF SELECTED DATA

The Sediment Data Bank system allows investigators to perform statistical studies on selected data. Possible studies include frequency distributions, factor analysis, correlation coefficients, regression lines, etc. A number of software packages exist for performing statistical calculations, and one of these (SPSS) is described in Section 2 below.

In order that statistical studies or other data manipulation may be performed without need for the user to punch large quantities of data or enter them by hand, programs have been developed to convert information from the data bank into formats suitable for input to other programs. Program NARCALC, which also performs a variety of other calculations, reformats the most frequently used data from the nodule analyses (NAR) file or a subset of analyses. NARCALC is described in Section 1 below and is available to users from remote terminals.

No general program yet exists to reformat data from other files, but such a program is expected to be developed in 1978. In the meantime, programs designed specifically to allow statistical operations on other parameters such as water depth, nodule size, or sediment lithology are written by the Sediment Data Bank staff as the need arises. Users should consult the staff about their particular program needs. If a suitable program exists it can be made available. If special programming is required, the user must cover the cost which will depend on the complexity of the request.

1. Program NARCALC for Nodule Analyses

a. **Purpose:** NARCALC operates on a file of data in the nodule analyses (NAR) format which was previously produced by Program SEARCH. NARCALC can perform any or all of the following tasks:

1) Element concentrations can be averaged for each station where there is more than one manganese nodule analysis. Only the following elements can be averaged or can be used in the calculations described below: manganese, iron, cobalt, nickel, copper, zinc, molybdenum, lead, aluminum, silicon, calcium.

2) Sums or ratios of various elements or multiples of them, or their most common oxides, can be calculated. For example, values such as the following can be calculated and printed for each individual analysis or for the average of analyses at each station:

Cu + Ni + Co

Mn/Fe

$(Cu + Ni)/(MnO_2 + Fe_2O_3 + CoO + NiO + CuO)$ "normalizing" the analysis to remove effects of varying detrital content

Ni + .33Cu   calculating a "nickel equivalent" value

3) If a file of station records corresponding to the analyses is on disk, NARCALC can insert in the SYMBOL location of each station record the result of the calculation or the concentration of a specified element for the appropriate analysis or average. With the resulting file a map can be plotted (see Section E-4) to show the specified or calculated value for each location.

4) NARCALC can create a disk or tape file containing the elements listed above for each of .the analyses, or a file containing element averages at each station in a format suitable for further use with statistical packages such as SPSS (see Section F-2) or other programs such as HISTOGRAM (see Section F-3).

b. Use: A file of nodule analyses selected by SEARCH must be on disk. If Task 4 above is to be performed, the stations file corresponding to the analyses file must also be on disk. NARCALC may be scheduled from a remote terminal, run interactively, or submitted through a card reader. File setup for scheduling from a remote terminal is as follows:

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?BEGIN
?RUN OBJECT/NARCALC
?FILE ANA(TITLE = <FILENAME>)          Filename is the name of the analyses
                                        file on disk if it is not ANA
?FILE STA(TITLE = <FILENAME>)          Filename is the name of the station
                                        file on disk if it is not STA
?FILE NEWANA(TITLE = <FILENAME>)       File to be created by the program if
                                        Option 3 = 1.  If not named, the
                                        resulting disk file will be titled
                                        NEWANA by default.  Omit this record
                                        if Option 3 = 0.
?DATA
      <DATA RECORDS>
?END JOB
```

c. Input: The first data record, which is always required, specifies the output options. The record is free field; entries must be separated by commas, and blanks are optional. A zero or one is entered for each of the four options in order as follows:

Option 1: 1 = print each analysis plus the average and the calculated
               value if they are computed.
           0 = omit printed listing.

Option 2: 1 = if more than one analysis for a station, average the
               elements for that station.
           0 = do not average.

Option 3: 1 = store analyses (only one for each station if Option 2
               = 1; more than one for some stations if Option 2 = 0)
               in disk file NEWANA or file named by user.
           0 = do not store analyses or averages.

Option 4: 1 = store calculated or specified value in the SYMBOL
location of the stations file (STA or its equivalent).
If Option 4 = 1 then Option 2 must also = 1 so that
there is only one analysis for each station. A station
file must be on disk. This file must contain one record
for each station for which analyses are included in the
ANA file and must be in the same order.

0 = do not store value in a stations file.

The second data record, which is optional, contains either the name
of up to five elements to be added together for the numerator of the cal-
culation or the name of the single element to be inserted as the SYMBOL in
the stations file. If Option 4 = 1, this data record is required. Format
of this record is five six-character fields, with each field entered as
follows:

Col. 1   If element concentration is to be expressed in percent,
           enter "%"; if it is to be expressed as a fraction (25% =
           .25), leave blank.

Cols. 2-3 The standard abbreviation for the name of the element.
           These are: Mn, Fe, Co, Ni, Cu, Zn, Mo, Pb, Al, Si or Ca.

Cols. 4-6 If element concentration is to be expressed as an oxide,
           enter the letter "O" in Col. 4. Oxide factors included in
           the program are for MnO_2, Fe_2O_3, CoO, NiO, CuO, ZnO, MoO_3,
           PbO, Al_2O_3, SiO_2 and CaO.

           If the element is to be multiplied by a constant factor,
           enter the factor as an integer in Cols. 4-6, with leading
           zeroes if integer is less than three digits. To multiply
           by a non-integer value, the user can take advantage of this
           fact: omitting the "%" in Col. 1 has the effect of multi-
           plying the concentration, which normally is expressed in
           weight percent, by 0.01. Thus omitting the "%" and entering
           a constant factor of 225 would have the effect of multiply-
           ing the concentration by 2.25.

           If no constant factor or oxide conversion is required, Cols.
           4-6 are left blank.

The third data record, which is used only if a ratio is to be calcu-
lated, contains the names of one to five elements or oxides which are to
be added together to form the denominator for the calculation. Format is
the same as for the previous record.

    d. Input Examples: To produce a disk file for input to SPSS contain-
ing a record for each analysis in the specified file but eliminating all
printed output, enter the following data record:

0,0,1,0   (second and third data records omitted)

To average the analyses, calculate (Ni + Cu + Co)/Mn, obtain a printed
listing, and store the resulting value in the SYMBOL location of a stations
file, enter the following data records:

```
1,1,0,1
%NI    %CU  .%CO
%MN
```

e. <u>Output</u>: If Option 1 = 1, NARCALC will produce a line printer list-
ing (Fig. 11). The element average will be included on the printout if
Option 2 = 1. The VALUE shown on the output is the result of the calculation
or the element value assigned to SYMBOL. There is no printed output if
Option 1 = 0.

If Option 3 = 1, a disk file will be created containing either one
record for each analysis (when Option 2 = 0) or one individual or averaged
analysis for each station (when Option 2 = 1). This file will have the
title NEWANA unless another name was specified in the label equation
record. File NEWANA is written in binary into 13-word records as follows:

WORD[0] for our use only   WORD[9] = Al
WORD[1] = Mn   WORD[10] = Si
WORD[2] = Fe   WORD[11] = Ca
WORD[3] = Co   WORD[12] = result of calculation or the
WORD[4] = Ni   element assigned to SYMBOL;
WORD[5] = Cu   will be zero if no second
WORD[6] = Zn   input data record or if ele-
WORD[7] = Mo   ment values needed for
WORD[8] = Pb   calculation are missing.

All values are stored in weight percent.

If Option 4 = 1, the STA file will have had the calculated value or
specified element for each station inserted in the SYMBOL location. This
file can then be used as input to Program MAP, and the SYMBOL value for
each station can be plotted.

f. <u>Caution</u>: The NAR file itself cannot be averaged because it is in
random order and all analyses for a given station are not stored together.
To create an ordered file of all nodule analyses, first search STATIONS
for NAR. Then using the newly created stations file (those stations with
nodule analyses) search for any element greater than or equal to zero.
Now the available analyses are ordered and can be averaged.

An error message will be printed and the program will terminate if an
element name is not one of those listed above.

If an incorrect character is entered instead of "%", it will be read
as a blank. If an incorrect character is entered in place of the letter
"O" (for oxide), it will be read as an integer factor. No error message
will be printed; the program will use the integer factor in its calcula-
tions, producing incorrect results.

All files used or created will remain on disk and incur storage
charges until removed by the user.

SEQ NO.	SAMP. TYPE & SECT SAMPLED	DIMENSIONS CORE	KEP	MN	FE	CO	NI	CU	ZN	MO	PB	AL	SI	CA	H2O
3380040 0 NODULE			A179 WC	21.80	5.80	0.27	1.20	0.94	0.000	0.000	0.00	0.00	0.00	0.00	0.00
VALUE= 2.1400															
3380044 0 NODULE			A179 WC	2.50	8.20	0.05	0.11	0.11	0.000	0.000	0.00	0.00	0.00	0.00	0.00
VALUE= 0.2200															
3380045 1 NODULE X-SECTION	160X100X100		A197 XF	29.80	4.80	0.20	1.36	1.20	0.120	0.054	0.06	3.02	6.08	1.47	16.20
3380045 2 NODULE X-SECTION			A166 XF	48.50	3.38	0.15	2.46	1.54	0.000	0.000	0.12	0.00	0.00	0.00	0.00
3380045 3 FRAGMENTS	25X25X25		A261 EM	32.00	5.30	0.18	2.00	1.40	0.040	0.070	0.05	2.70	7.50	1.80	0.00
3380045 4 NODULE X-SECTION			A166 XF	31.50	0.09	0.12	1.53	0.92	0.000	0.000	0.07	0.00	0.00	0.00	0.00
3380045 5 NODULE, INNERMOST LAYER			A201 EM	0.00	6.20	0.26	1.60	1.70	0.000	0.110	0.00	0.00	6.20	0.00	0.00
3380045 6 FRAGMENTS	30X30X20		A200 XF	38.10	4.40	0.25	1.70	1.21	0.141	0.000	0.05	0.00	0.00	1.66	0.00
3380045 7 NODULE, INNERMOST LAYER	100		A197 XF	30.70	2.30	0.11	1.37	1.34	0.000	0.000	0.00	0.00	6.90	1.23	12.70
3380045 8 NODULE, MIDDLE LAYER	100		A197 XF	29.10	3.90	0.16	1.26	1.00	0.000	0.000	0.00	0.00	6.90	1.46	15.90
3380045 9 NODULE, MIDDLE LAYER	100		A197 XF	27.70	5.30	0.22	1.33	1.00	0.000	0.000	0.00	0.00	5.70	1.56	14.90
338004510 NODULE, OUTER LAYER	100		A197 XF	32.10	3.20	0.20	1.05	0.77	0.000	0.000	0.00	0.00	5.80	1.40	11.50
338004511			A215	30.00	1.50	0.08	1.10	0.25	0.163	0.095	0.03	1.29	20.34	0.52	0.00
338004512 NODULE			A262 EM	0.00	7.00	0.30	1.50	1.50	0.070	0.070	0.07	0.00	7.00	0.00	0.00
338004513 NODULE FRAGMENTS	30X20X10		A388 XF	34.31	5.30	0.26	1.62	1.15	0.175	0.086	0.03	0.00	4.49	1.46	0.00
338004514	30X30X20		A388 XF	32.68	4.46	0.22	1.64	1.19	0.167	0.086	0.03	0.00	4.67	1.33	0.00
2.6916		AVERAGE		33.04	4.08	0.19	1.54	1.15	0.125	0.082	0.06	2.50	7.42	1.54	
3380046 0 NODULE			A189	23.46	5.41	0.14	1.93	1.59	0.000	0.043	0.04	0.00	0.00	0.00	0.00
3380046 1 NODULE	21		A365 AA	22.00	4.60	0.14	1.30	1.04	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380046 2 NODULE			A322 XF	19.60	5.80	0.17	0.97	0.92	0.090	0.058	0.04	2.90	8.90	1.70	0.00
3380046 3 NODULE W/O NUCLEUS			A346 AA	27.80	4.10	0.17	1.44	0.95	0.000	0.094	0.00	0.00	0.00	0.00	0.00
3380046 4 NODULE COMPOS. >12			A379 AA	25.60	5.00	0.18	1.50	1.20	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380046 5 NODULE			A381 MP	25.70	5.10	0.15	1.90	0.95	0.150	0.000	0.00	0.00	0.00	1.25	0.00
3380046 6 NODULE			A381 MP	25.50	4.80	0.20	1.75	1.15	0.150	0.000	0.00	0.00	0.00	1.10	0.00
3380046 7 NODULE			A381 WC	0.00	0.00	0.20	1.60	1.00	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380046 8 NODULE			A381 WC	0.00	0.00	0.30	1.50	1.10	0.000	0.000	0.00	0.00	0.00	0.00	0.00
2.6433		AVERAGE		24.24	4.97	0.18	1.54	1.10	0.130	0.065	0.04	2.90	8.90	1.42	
3380050 0			A395	0.00	0.00	0.14	0.62	0.47	0.000	0.000	0.00	0.00	0.00	0.00	0.00
VALUE= 1.0900															
3380152 0 NODULE		NONE	A326 RF	19.10	5.08	0.19	1.40	1.09	0.125	0.000	0.01	0.00	10.84	0.00	0.00
3380152 1 NODULE		NONE	A326 RF	25.40	5.13	0.21	1.71	1.34	0.152	0.000	0.02	0.00	9.16	0.00	0.00
2.7700		AVERAGE		22.25	5.10	0.20	1.56	1.22	0.134	0.000	0.02	0.00	10.00	0.00	
3380112 0 NODULE			A242 XE	24.68	7.67	0.00	1.70	0.87	0.000	0.000	0.00	0.00	0.00	0.00	0.00
VALUE= 2.5700															
3380111 0 NODULE			A242 XE	25.49	7.66	0.00	1.55	0.98	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380111 1 NODULE W/O NUCLEUS			A242 XE	26.54	8.54	0.00	1.80	1.03	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380111 2 NODULE W/O NUCLEUS			A242 XE	26.81	7.53	0.00	1.82	1.00	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380111 3 NODULE W/O NUCLEUS			A242 XE	28.72	8.35	0.00	1.87	1.07	0.000	0.000	0.00	0.00	0.00	0.00	0.00
2.7800		AVERAGE		26.89	8.02	0.00	1.76	1.02	0.000	0.000	0.00	0.00	0.00	0.00	
3380109 0 NODULE, HALF			A244 XE	30.70	4.65	0.00	1.67	1.52	0.000	0.000	0.00	0.00	0.00	1.33	30.15
3380109 1 NODULE, HALF			A244 XE	31.70	4.54	0.00	1.55	1.39	0.000	0.000	0.00	0.00	0.00	1.12	29.91
3380109 2 NODULE W/O NUCLEUS			A242 XE	30.06	6.65	0.00	1.72	1.28	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380109 3 NODULE W/O NUCLEUS			A242 XE	31.94	6.37	0.00	1.64	1.19	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380109 4 NODULE W/O NUCLEUS			A242 XE	27.47	8.51	0.00	1.59	1.14	0.000	0.000	0.00	0.00	0.00	0.00	0.00
3380109 5 NODULE W/O NUCLEUS			A242 XE	32.02	7.26	0.00	1.73	1.17	0.000	0.000	0.00	0.00	0.00	0.00	0.00
2.9317		AVERAGE		30.65	6.33	0.00	1.65	1.28	0.000	0.000	0.00	0.00	0.00	1.08	

Fig. 11. Line printer output produced by Program NARCALC. The calculated value is the average Ni + Cu at each station.

65

2. SPSS Statistics Package

Statistical Package for the Social Sciences (SPSS) is one of several systems of computer programs available on the Burroughs B-6700 at UCSD. Although SPSS was designed for students and researchers in the social sciences, it has many useful applications for other fields of research. SPSS can perform a variety of statistical procedures efficiently and can be used by someone with little or no computer experience. The types of procedures available in SPSS as well as a description of how to use the system are well documented in SPSS Statistical Package for the Social Sciences, Second Edition (McGraw-Hill, Inc., 1975). In addition to this manual, the user is urged to obtain UCSD Computer Center Documents #034, SPSS Overview, and #268, SPSS Workflow Control.

a. Purpose: SPSS offers a variety of statistical procedures including descriptive statistics such as minimum, maximum and range of values, mode, median, mean, variance, standard deviation, standard error, skewness and kurtosis; one-way frequency distributions; scatter diagrams with bivariate regression and correlation; multiple regression analysis; analysis of variance and covariance; and factor analysis. It provides a variety of output options in a variety of formats.

b. Use: The user should be familiar with both SPSS and the UCSD Computer Center workflow languages. Before running SPSS programs it is advisable to call the UCSD Computer Center consultant to determine what version of SPSS is currently being used. In the example below, Version 6 was used.

All rules described in the SPSS manual must be adhered to. SPSS can be run in batch or scheduled from a remote terminal, but it is not interactive.

c. Input: The data may be read either from a disk file such as file NEWANA produced by Program NARCALC, or data may be punched on cards. Data bank files such as STATIONS, NAR or DESCRIPTIONS may not be used because they have more than one variable stored in each word of the records. In the examples below, File NEWANA was created by NARCALC. The DATA LIST card will remain the same for all SPSS procedures as long as the file containing the data was created by NARCALC.

Workfile setups for two SPSS procedures that the user may find particularly useful for nodule analysis data follow:

Example 1. Procedure SCATTERGRAM

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?BEGIN
?RUN STATPAK/SPSSV6
?FILE FILE8(KIND = DISK, TITLE = NEWANA, FILETYPE = 7)
?DATA
RUN NAME        FIG12
DATA LIST        BINARY X MN FE CO NI CU ZN MO PB AL SI CA SYM 1-13
INPUT MEDIUM     DISK
N OF CASES       UNKNOWN
MISSING VALUES   MN TO SYM (0)
```

```
SCATTERGRAM    MN WITH NI(0,3)
STATISTICS     ALL
READ INPUT DATA
FINISH
?END JOB
```

This work file will produce a scatter diagram on a line printer for nickel vs. manganese with Ni within the range 0 to 3.0% and Mn within the range calculated from its actual minimum and maximum values. Data for the scattergram is in File NEWANA. The following statistics will also be produced: Pearson's r (correlation coefficient), r^2, significance of r, standard error of the estimate, and the intercept with the vertical axis and slope of the regression line. Zero values will be omitted. See Fig. 12 for the output from this job.

Example 2. Procedure PEARSON CORR

Work file setup for PEARSON CORR can be exactly the same as for SCATTERGRAM except for the procedure record:

```
PEARSON CORR   MN, FE, NI, CU
```

This work file will create and print a correlation matrix for Mn, Fe, Ni and Cu for all values in file NEWANA and will compute and print the standard deviation and mean of each variable. See Fig. 13 for output from this job.

d. Output: Output is to a line printer, disk, tape or cards. See the SPSS Manual for options. Figs. 12 and 13 are examples of printed output.

e. Caution: The workfile you make to schedule an SPSS run must be of type JOB. When you make the workfile using CANDE, you must enter MAKE <FILENAME> JOB. If this is not done, execution will fail.
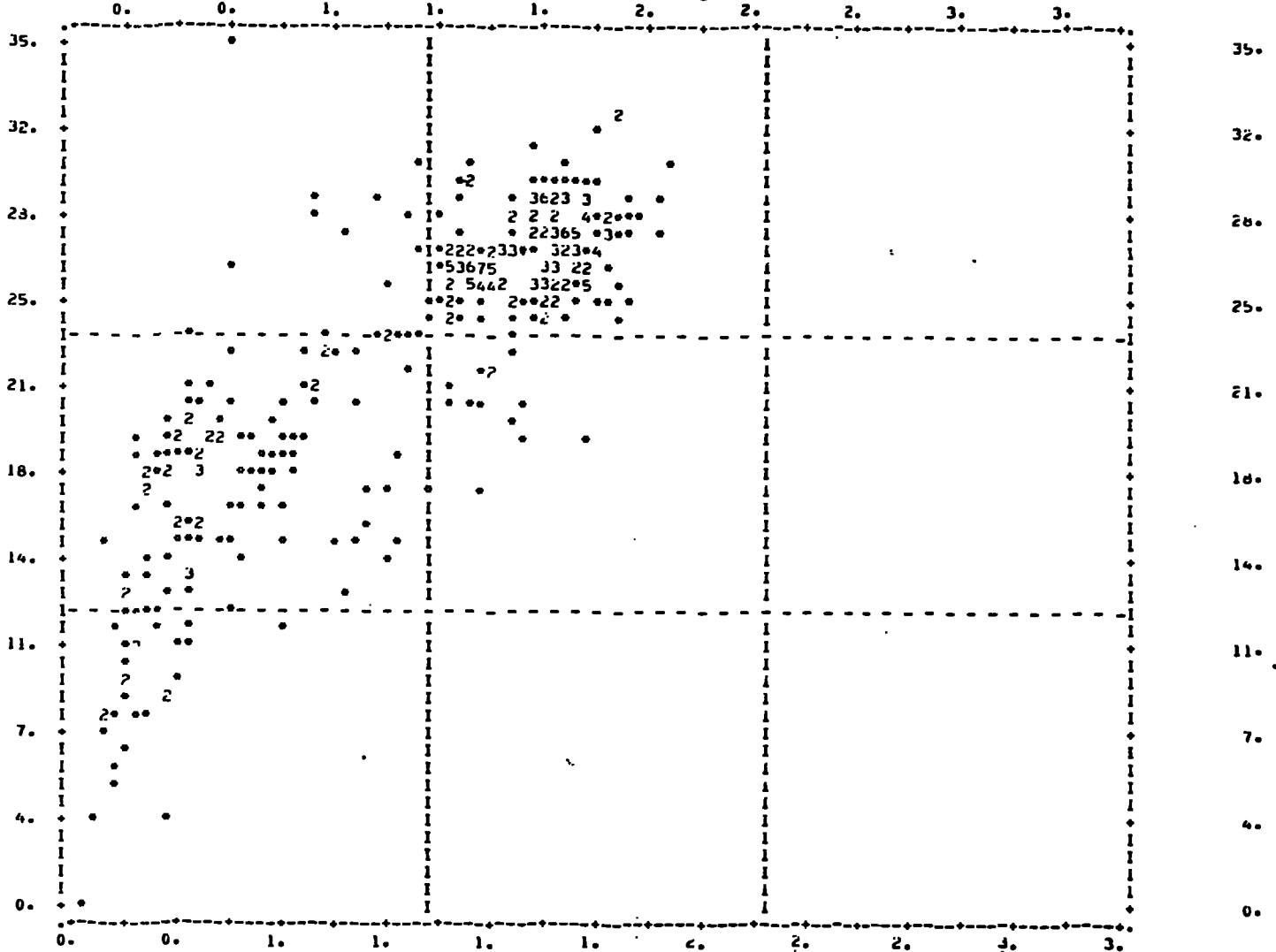
EXAMPLE1   01/29/78   PAGE   2

FILE   NONAME   (CREATION DATE = 01/29/78)
SCATTERGRAM OF   (DOWN) MN   (ACROSS) NI
```
                  0.      0.      1.      1.      1.           2.      2.      2.      3.      3.
         .----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+.
    35.  . *              *                       I                        I                        I    35.
         I                                        I                        I                        I
         I                                        I                        I                        I
         I                                        I                        I                        I
    32.  .                                        I             2          I                        I    32.
         I                                        I          *             I                        I
         I                                     *I  *       *               I                        I
         I                                       I*2   ******              I                        I
    28.  I                              *      *  I*   * 3623 3    *    *   I                        I    28.
         I                           *         *I*   2 2 2  4*2***         I                        I
         I                                    * I*    *  22365 *3** *       I                        I
         I                               *I*222*233** 323*4                 I                        I
         I                               1*53675   33 22  *                 I                        I
    25.  .                           *   I 2 5442  3322*5  *                I                        I    25.
         I                              **2*  *  2**22 * ** *               I                        I
         I                               * 2*  * *  *2 * *                  I                        I
         I- - - - - -*- - - - -  *  - -*24**- - - - - *- - - - - - - - - - - - - - - - -I    25.
         I                    *    *  2* *    * I   *                       I                        I
    21.  .              *  * *      *2     * I  *7                          I                        I    21.
         I            *  2   *              *I  * **    *                   I                        I
         I         *  *2  22 **  ***        I  *        *                   I                        I
         I         * *****2  ****            I                              I                        I
    18.  .         2*2   3   ****  *         I  *                           I                        I    18.
         I         2                    * *  * *                            I                        I
         I        *      ** * *      *        I                             I                        I
         I          2*2          *           I                              I                        I
         I          *** **    * * *          I                              I                        I
    14.  .   *    *  *   3                    I                             I                        I    14.
         I         * *  * *                   I                             I                        I
         I         2   * *          * *       I                             I                        I
         I- - -*.**.- - - -*- - - -  -*- - - -I- - - - - - - - - - - - - - -I                        I
    11.  .      *   * *     **          I                                   I                        I    11.
         I         *2    **                   I                             I                        I
         I         2   *                      I                             I                        I
         I        *  *  2                     I                             I                        I
    7.   .     2* **                          I                             I                        I    7.
         I       *                            I                             I                        I
         I       *                            I                             I                        I
         I                                    I                             I                        I
    4.   .     *         *                    I                             I                        I    4.
         I                                    I                             I                        I
         I                                    I                             I                        I
    0.   . **                                 I                             I                        I    0.
         .*---+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+*.
                  0.      0.      1.      1.      1.      1.      2.      2.      2.      3.      3.
```

STATISTICS..

CORRELATION (R)-	0.84751	R SQUARED -	0.71828	SIGNIFICANCE -	0.00001
STD ERR OF EST -	3.29117	INTERCEPT (A) -	12.04542	SLOPE (B) -	11.10253
PLOTTED VALUES -	382	EXCLUDED VALUES-	0	MISSING VALUES -	13

'*********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

Fig. 12. Line printer output produced by SPSS Procedure SCATTERGRAM using workfile setup in Example 1.

FILE   NONAME   (CREATION DATE = 01/29/78)

VARIABLE	CASES	MEAN	STD DEV
MN	382	22.7320	6.1926
FE	395	11.1147	5.6163
NI	395	0.9418	0.4801
CU	395	0.7103	0.4576

VARIABLES		CASES	CROSS-PROD DEV	VARIANCE-COVAR	VARIABLES		CASES	CROSS-PROD DEV	VARIANCE-COVAR
MN	MN	382	14610.5469	38.3479	MN	FE	382	-9420.7754	-24.7264
MN	NI	382	945.2323	2.4809	MN	CU	382	822.6797	2.1593
FE	MN	382	-9420.7754	-24.7264	FE	FE	395	12427.8090	31.5427
FE	NI	395	-907.4150	-2.3031	FE	CU	395	-805.0400	-2.1955
NI	MN	382	945.2323	2.4809	NI	FE	395	-907.4150	-2.3031
NI	NI	395	90.8070	0.2305	NI	CU	395	80.0188	0.2031
CU	MN	382	822.6797	2.1593	CU	FE	395	-805.0400	-2.1955
CU	NI	395	80.0188	0.2931	CU	CU	395	82.5189	0.2094

- - - - - - - - - - - P E A R S O N   C O R R E L A T I O N   C O E F F I C I E N T S - - - - - - - - - - - - - - -

	MN	FE	NI	CU
MN	1.0000 (382) S=0.001	-0.7042 (382) S=0.001	0.8475 (382) S=0.001	0.7661 (382) S=0.001
FE	-0.7042 (382) S=0.001	1.0000 (395) S=0.001	-0.8542 (395) S=0.001	-0.8542 (395) S=0.001
NI	0.8475 (382) S=0.001	-0.8542 (395) S=0.001	1.0000 (395) S=0.001	0.9244 (395) S=0.001
CU	0.7661 (382) S=0.001	-0.8542 (395) S=0.001	0.9244 (395) S=0.001	1.0000 (395) S=0.001

(COEFFICIENT / (CASES) / SIGNIFICANCE)   (A VALUE OF 99.0000 IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED)

Fig. 13. Line printer output produced by SPSS Procedure PEARSON CORR using workfile setup in Example 2.

70

3. Program HISTOGRAM

a. **Purpose:** To produce frequency distribution plots of any parameter selected from the data bank files.

b. **Use:** Parameters for which frequency distributions can be made must be one of the following: (1) elements or calculated values stored in a disk file (NEWANA) which resulted from a run of Program NARCALC, or (2) values assigned to SYMBOL in a stations file on disk. HISTOGRAM may be scheduled from a remote terminal, submitted through a card reader, or run interactively. File setup for scheduling from a remote terminal is as follows:

```
?JOB <JOBNAME>;CLASS = <QUEUE>
?TO = <DESTINATION>
?RUN OBJECT/HISTOGRAM
?FILE NEWANA(TITLE = <FILENAME>)    Use if filename is not NEWANA
?DATA
      <DATA RECORDS>
?END JOB
```

c. **Input data:** Two data records are required. The first data record sets the range of values to be plotted, the number of intervals (bins) and the location in the file records from which the data are to be read. The record is free field, and entries must be separated by commas. Data must be entered in the following order:

XMIN, XMAX,BINS,WORD

where XMIN is the minimum and XMAX the maximum value to be included. A specified value allows exclusion of outlying data points from the plot; however, all data are used to calculate the mean, standard deviation, etc. If XMIN or XMAX is set at -0, the program will use the minimum or maximum value in the actual data set.

BINS is the number of intervals into which the data are to be divided. If BINS is set to zero, the default of N/10 - 1 is used where N is the number of data points in the file. In this case a frequency distribution table will not be printed.

WORD is the number of the word in the file records that contains the value of the appropriate parameter. See Page 64 for the locations of parameters in files produced by Program NARCALC. If the parameter is stored in a stations file, WORD = 6.

The second data record is the title that will appear on the plot and on the printed output. The title can be up to 60 characters long including blanks and must be followed by a semicolon.

Input example: To produce the plot and listing shown in Figs. 14 and 15 the following data records were used:

```
0,2,20,4
%NI IN DEMO/NAR (SQU=373);
```

d. __Output__: HISTOGRAM produces a line printer listing (Fig. 14) of the total number of cases, mean, standard deviation (sigma), and variance (sigma squared), plus a table showing the number of cases within each interval, the relative frequencies and the cumulative frequencies.

The program also produces a disk file titled DISKPLOTS/<account number> /<plot number>, but it does not actually plot the histogram. To plot the file on the drum plotter, run Program PLOTREQUEST (UCSD Computer Center Document #063). To display the plot on a CRT, run Program TEKPLOT (Computer Center Document #065). A typical example of plotted output from HISTOGRAM is shown in Fig. 15.

f. __Caution__: The DISKPLOTS file will remain on disk until removed explicitly or by running PLOTREQUEST with no SAVE option.

%NI IN DEMO/NAR (SQU=373);
   NO. OF CASES = 314
   MEAN = 1.360
   SIGMA = 0.247
   VARIANCE = 0.061

INT. MIN.	NO.	REL(%) FREQ.	CUM(%) FREQ.
0.00	1	0.32	0.32
0.10	1	0.32	0.64
0.20	0	0.00	0.64
0.30	0	0.00	0.64
0.40	1	0.32	0.96
0.50	0	0.00	0.96
0.60	1	0.32	1.27
0.70	3	0.96	2.23
0.80	6	1.91	4.14
0.90	9	2.87	7.01
1.00	13	4.14	11.15
1.10	21	6.69	17.83
1.20	36	11.46	29.30
1.30	83	26.43	55.73
1.40	68	21.66	77.39
1.50	33	10.51	87.90
1.60	14	4.46	92.36
1.70	18	5.73	98.09
1.80	2	0.64	98.73
1.90	3	0.96	99.68
2.00	1	0.32	100.00

Fig. 14. Line printer output produced by Program HISTOGRAM.
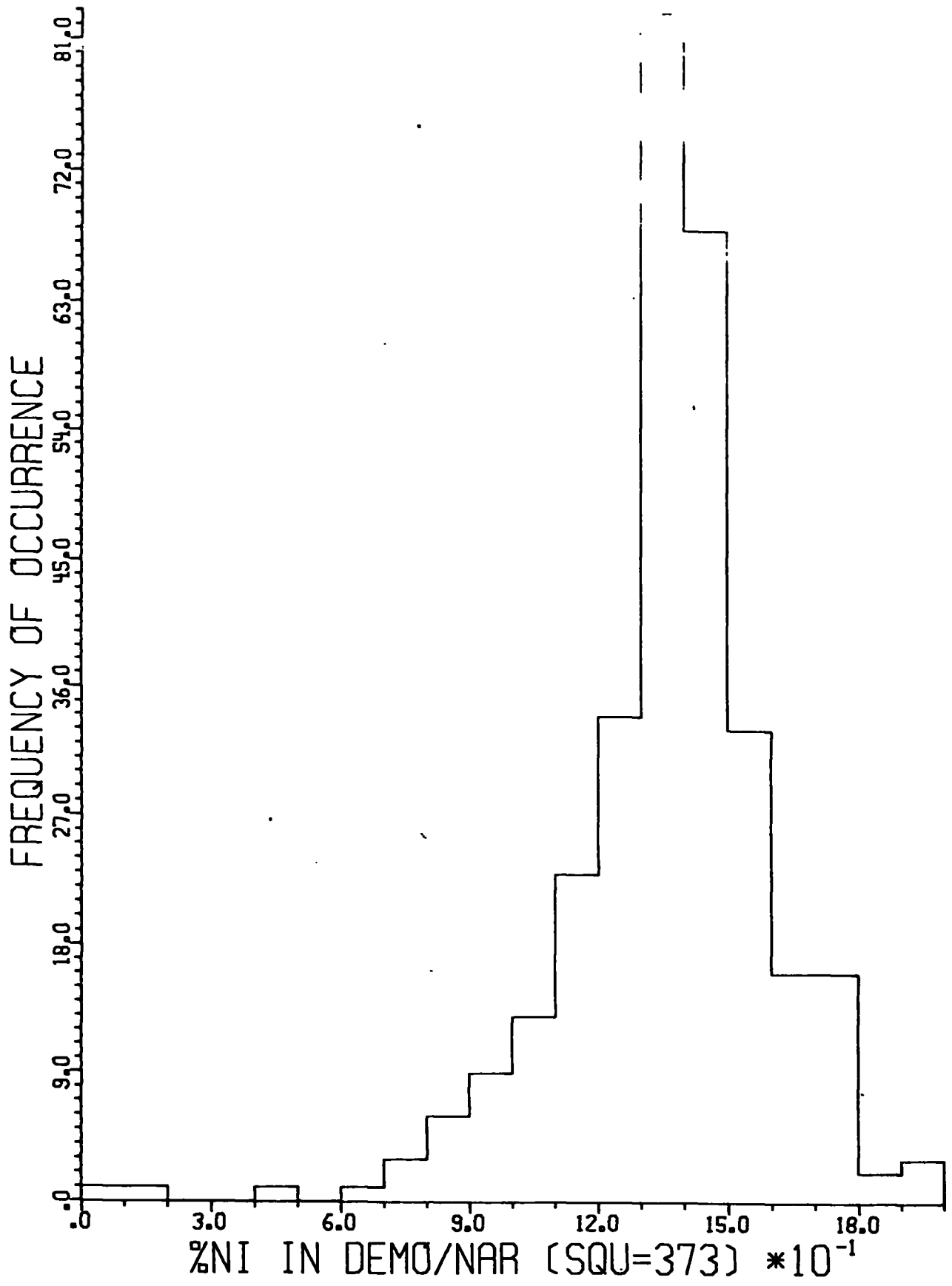
Fig. 15. Frequency distribution plot produced by Program HISTOGRAM.

G. ACKNOWLEDGMENTS