Geophysical Data Interchange Assessment, 1978 (1979)

Pages
81

Size
5 x 9

ISBN
0309332613

Committee on Data Interchange and Data Centers; Geophysics Research Board; Assembly of Mathematical and Physical Sciences; National Research Council

Geophysical Data Interchange Assessment 1978

Committee on Data Interchange and Data Centers
Geophysics Research Board
Assembly of Mathematical and Physical Sciences
National Research Council

NATIONAL ACADEMY OF SCIENCES
Washington, D.C. 1979

NOTICE: The project that is the subject of the report was approved by the Governing Board of the National Research Council, whose members are drawn from the Councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the Committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

Available from
Committee on Data Interchange and Data Centers
of the Geophysics Research Board
National Research Council
2101 Constitution Avenue
Washington, D.C. 20418

Committee on Data Interchange and Data Centers

Thomas O. Haig, *Chairman*
Space Science and Engineering Center
University of Wisconsin

Sidney A. Bowhill
Aeronomy Laboratory
Department of Electrical Engineering
University of Illinois

Colin B. Bull
College of Mathematics and Physical Science
Ohio State University

Carl W. Kreitzberg
Department of Physics and Atmospheric Science
Drexel University

Juan G. Roederer
Geophysical Institute
University of Alaska

Bruce A. Taft
Department of Oceanography
University of Washington

M. Nafi Toksöz
Department of Earth and Planetary Sciences
Massachusetts Institute of Technology

Richard Y. Dow, *Executive Secretary*

PREFACE

In discussions with representatives of the National Oceanic and Atmospheric Administration and the National Science Foundation in early 1977, the Chairman of the Committee on Data Interchange and Data Centers (CDIDC) was informed that the previous reports of this Committee were extremely useful and that further reports were needed on the status and effectiveness of the National and World Data Center-A (WDC-A) Geophysical Data Centers, with particular attention to present and future data problems. In March 1977, on the recommendation of an *ad hoc* panel, and with the approval of the Geophysics Research Board, the Committee established *ad hoc* panels to study the WDC-A's and associated National Data Centers (NDC's) in six geophysical disciplines. Each panel was chaired by a CDIDC member representing the appropriate discipline; panel members were recommended by relevant National Research Council (NRC) Committees.

Atmospheric Sciences: Carl Kreitzberg, Drexel University, *Chairman*; Lance Bosart, State University of New York at Albany; Roy Jenne, National Center for Atmospheric Research; Paul Julian, National Center for Atmospheric Research.

Glaciology: Colin Bull, Ohio State University, *Chairman*; William O. Field, American Geographical Society; Wesley Pietkiewicz, U.S. Army Cold Regions Research and Engineering Laboratory; John Hollin, University of Colorado; Paul McClain, National Environmental Satellite Service.

Oceanography: Bruce A. Taft, University of Washington, *Chairman*; Douglas R. McLain, Department of Commerce, NOAA.

Solar-Terrestrial Physics: Sidney A. Bowhill, University of Illinois, *Chairman*; Donald A. Gurnett, University of Iowa; Erwin Schmerling, NASA.

Solid-Earth Geophysics: M. Nafi Toksöz, Massachusetts Institute of Technology, *Chairman*; Al Balch, U.S. Geological Survey; Michael A. Chinnery, Lincoln Laboratory, Massachusetts Institute of Technology; Bruce Doe, U.S. Geological Survey; Adam Dziewonski, Harvard University; E. R. Engdahl, U.S. Geological Survey; J. C. Harrison, University of Colorado; Martin Kane, U.S. Geological Survey; Carl Kisslinger, Cooperative Institute for Research in Environmental Sciences, University of Colorado; William H. K. Lee, U.S. Geological Survey, Menlo Park; Tom Simkin, Smithsonian Institution; James G. Tanner, Department of Energy, Mines and Resources, Ottawa, Canada; Wayne Thatcher, U.S. Geological Survey, Menlo Park; Anthony Watts, Lamont-Doherty Geological Observatory, Columbia University.

Space Science: Juan G. Roederer, University of Alaska, *Chairman*; James W. Head, III, Brown University; George A. Paulikas, Aerospace Corporation.

Each Panel visited the WDC-A's and NDC's in its field. The schedule of these visits follows:

The Panel on Space Science visited WDC-A for Rockets and Satellites, and the National Space Science Data Center on May 13, 1977.

The Panel on Solar-Terrestrial Physics visited WDC-A for Rockets and Satellites and the National Space Science Data Center on May 11, 1977, and the WDC-A for Solar-Terrestrial Physics and the National Geophysical and Solar-Terrestrial Data Center on November 3, 1977.

The Panel on Atmospheric Sciences visited the WDC-A for Meteorology and the National Climatic Center on September 15, 1977.

The Panel on Glaciology visited WDC-A for Glaciology on July 19, 1977.

The Panel on Solid-Earth Geophysics visited the WDC-A for Solid-Earth Geophysics and the National Geophysical and Solar-Terrestrial Data Center on November 3-4, 1977. They met with staff members of the data centers and formed

working subgroups to look into the data-exchange problems of the subdisciplines in solid-earth geophysics.

The Panel on Oceanography, which was not established until February 1978, visited the WDC-A for Oceanography and the National Oceanographic Data Center on February 28 and March 1, 1978.

Reports prepared by these Panels were reviewed by the CDIDC at its meetings on November 9-10, 1977, and March 15, 1978, and are included as Appendixes to this report. Based on the findings of the panels, the Committee developed its report and recommendations.

CONTENTS

1
Findings and Recommendations

1. In the United States, the growth of concern about the environment is producing a corresponding increase in the flow of geophysical data. To accommodate this growth, we recommend continued expansion and automation of data-center facilities.

2. The scientific community will become more heavily dependent on effective data-center services in the future; thus, we recommend that this community lend its strong support to the achievement of an adequate level of funding for those services.

3. In planning data-intensive projects, we recommend that the funds necessary for preparation of data for archiving and for long-term preservation and distribution be included from the outset.

4. We recommend that the application of modern techniques to make the data available to users be accelerated. When necessary, special action should be taken to accelerate automatic-data-processing (ADP) procurement procedures.

5. Some geophysical observations made regularly by federal agencies and scientists are not deposited in National Data Centers (NDC's). Where current or anticipated circumstances could result in the loss of unique and valuable data sets, we recommend that the Data Centers be provided the resources necessary to take custody of such data.

6. We recommend that the National Aeronautics and Space Administration (NASA) re-evaluate the arrangements for processing and archiving space data in the National Space Science Data Center (NSSDC), so that adequate capability is assured to meet NASA's responsibilities to the user community.

7. We recommend that scientific use of the data be promoted by encouraging and financing scientists and students, perhaps through a program of data scholarships, and, in some

cases, by stimulating scientific activity by the Data Center personnel.

8. We recommend expansion of contacts with WDC's in other countries to reach agreement on common formats for related data, to ensure more timely data exchange, to expand the types of data that are exchanged, and to explore the possibility of digital data links.

9. We recommend that the Geophysics Research Board (GRB) consider the need for regional or specialized data centers in such subjects as mesoscale climatology, water resources, and snow and ice data.

10. Because most records at the Data Centers would be difficult or impossible to replace, and most of the present archives are inadequately protected against loss through natural or human causes, we recommend that duplicate copies be made of all physical data and archived in a secure place.

11. In view of the importance of the *Guide to International Data Exchange Through the World Data Centers*, we urge the International Council of Scientific Unions (ICSU) Panel on WDC's to issue a new and revised *Guide* promptly.

12. Believing that there is need for a unified national geophysical data policy, the Committee on Data Interchange and Data Centers proposes to assemble a draft policy during the next several months. Input by interested government agencies, National Research Council (NRC) committees, professional societies, and other organizations and individuals will be solicited as part of this process.

2

INTRODUCTION

The term National Data Center (NDC) is used repeatedly in this report; yet it is difficult to identify all NDC's precisely. Geophysical data are held by a large number of agencies, only a few of which have titles indicating NDC status. National Data Centers are facilities operating within the federal government to serve the needs of users of geophysical data. Those with which this report is concerned are

National Climatic Center (NCC), Asheville, North Carolina
World Data Center-A Glaciology, Boulder, Colorado
National Oceanographic Data Center (NODC), Washington, D.C.
National Space Sciences Data Center (NSSDC), Greenbelt, Maryland
National Geophysical and Solar-Terrestrial Data Center (NGSDC), Boulder, Colorado

The World Data Center (WDC) system was established in 1957 to provide international access to data collected during the International Geophysical Year. There are three sets of WDC's: WDC-A in the United States; WDC-B in the Soviet Union; and WDC-C distributed among several other countries. Fuller information concerning the WDC's is available from the *Guide to International Data Exchange Through the World Data Centers*, and from an earlier report* of the Committee on Data Interchange and Data Centers (CDIDC).

An Assessment of the Impact of World Data Centers on Geophysics (National Academy of Sciences, Washington, D.C., March 1975).

In general, related NDC's and WDC-A's are co-located in the United States; however, there is not a one-to-one correspondence between NDC's and WDC-A's. For example, there is no designated NDC for glaciology, and the WDC-A Glaciology serves both functions. Co-located WDC-A's tend to serve as a means of special access to the NDC holdings, with little or no separate WDC archive.

The NSSDC is operated by the Goddard Space Flight Center of the National Aeronautics and Space Administration (NASA); the other NDC's listed above are operated by the Environmental Data and Information Service (EDIS) of the National Oceanic and Atmospheric Administration (NOAA). Many other data centers of national and international importance exist in other agencies, such as the WDC-A Rotation of the Earth at the U.S. Naval Observatory, Washington, D.C.

The organizational complexity of the data centers reflects the multiplicity of interests and disciplines involved in the collection and use of geophysical data. This variety of attitudes also is demonstrated in the different approaches taken by the panels that have contributed the appendixes to this report.

During the past decade, a significant consolidation of data centers has occurred, especially within NOAA. Considerable progress has been made in coordination of data center holdings and operations with the initial implementation of a computer-to-computer network among the major EDIS facilities. All the data centers are actively engaged in updating their facilities and practices to take advantage of new technology to improve services and reduce costs. Especially noteworthy are the activities of EDIS to procure and install automatic-data-processing (ADP) equipment tailored to fit the needs of each center.

The number of requests for data from the centers has been increasing steadily. For all EDIS centers, the increase has averaged 10.3 percent per year for the last five years. Greatest increases in numbers of users have occurred in "General Public" and "Industry" categories. In 1977, users of the EDIS centers were distributed as follows:

Foreign	4.4%
Academic	10.7%
Industrial	25.7%
Government	20.4%
General public	38.7%

The rapidly increasing use of the centers and the even more rapid increase in the volume of archived data continue

to challenge the agencies responsible for operation of the centers. The objective of the CDIDC and its contributing panels has been to understand better the dynamics of present NDC/WDC-A activities and to recommend ways to improve data services. In many cases our recommendations are intended to reinforce an action already started by one or more of the responsible agencies.

3
More Data and Why

IMPACT OF ENVIRONMENTAL CONCERNS

The Geophysical Data Panel, chaired by Carl Savit (see Appendix A), considered the impact of large-scale geophysical research programs such as BOMEX, GATE, and FGGE on the data centers and found that expansion of the facilities of the centers was necessary to cope with these special data-collection programs. The present level of data-center activities confirms the increases in data quantities predicted by the Savit Panel. However, it now seems clear that the increased quantity of data is a consequence of factors of greater scope than the data-collection programs themselves.

During the past dozen years the attitudes of the people and government of the United States has shown a profound change in the way the environment is regarded. The federal government has established a large number of programs to understand, protect, predict, and improve the geophysical environment. Actions that give evidence of ever-growing concern for the environment include the Clean Air Act, legislation requiring environmental impact studies before the approval of projects that a few years ago would have proceeded without question, the National Climate Program, restrictions on release of aerosols, and the formulation of a National Water Policy. One of the consequences is a rapid increase in the demands for data service placed on the geophysical data centers by government, industry, schools, and scientists. One cannot write an environmental impact report without having researched past and present environmental conditions. The sources of data for such studies are alsmot always the NDC's and WDC's.

As the governmental and public concern for the environment has grown, the activities of scientists and the

operating agencies have increased to meet the demands for greater understanding of geophysical processes and for closer monitoring of environmental factors. The latest technological developments--remote sensing, satellites, and computers--have been adapted to make the needed measurements. Both experimental and operational programs have been initiated that have greatly increased the volume of geophysical data being produced. The rate of data production continues to increase and can be expected to do so for the foreseeable future.

The large-scale data-collection programs have served to focus the attention of scientists and governmental agencies on specific research objectives and have caused collected data to be organized in identifiable sets. However, these research programs probably have not, by themselves, greatly changed the volume of data that the centers must handle. The data volume would have increased at nearly the same rate because of the general growth in environmental concern and the utilization of new data-collection technology by government agencies to meet that concern. Expansion of data-center facilities was recommended by the Savit Panel in consequence of a predicted continued exponential increase in the quantity of data being collected by the geophysical community. Their recommendations were not made in response to a one-time situation created by the large-scale programs. Since those recommendations were made, the increased demands on the centers, in terms of both data input and services to users, have materialized as the result of a national awakening to the importance of geophysical environments. It is essential that those responsible for the management of the NDC's and WDC's realize that continuing expansion and automation of facilities are inevitable and must be accommodated.

NEED FOR SCIENTIFIC COMMUNITY SUPPORT

The task of archiving data has usually been regarded as a necessary chore but not an activity worthy of the attention of top-flight scientists. The glamorous tasks in science are the development and use of systems to collect and analyze data. Traditionally, the work of formatting, cataloging, storing, and retrieving data has been left for assistants or ignored entirely. Increasingly, scientific investigations are being organized into large-scale data-collection programs in which the objectives of the programs cannot be achieved unless the data are processed, assembled, and stored with care. By the nature of the large-scale

programs, it is nearly impossible for participating scientists to collect and analyze their own data independently. The geographic and time scales, the data volume, and the number of people involved are all too large. Observations must be collected, processed, and assembled into data sets before they become useful for scientific analysis and investigation. It is important that the scientific community recognize its increasing dependence on the NDC's and WDC's and lend active support to the continuing expansion and automation of center facilities.

4
Data Acquisition

Most data archived by the NDC's are produced by the federal agencies responsible for monitoring the atmosphere, the oceans, and the solid earth. For example, weather information is regularly deposited at the National Climatic Center by the National Weather Service, the Federal Aviation Administration, the Air Weather Service, the Naval Weather Service, and several other agencies that maintain atmospheric observation facilities. Lesser amounts of data are received from experimental programs, most of which are federally funded. The mix of data differs greatly at each of the NDC's. At the NCC, most of the data are received from government agencies on a regular basis. At the NSSDC, data are received from experimental satellite and space-vehicle programs sponsored by NASA and conducted by scientists in government laboratories, universities, and sometimes private corporations. At the NODC and NGSDC, both types of data are received, with the proportion from regular observing services increasing as satellite observations become available.

RESPONSIBILITY FOR ARCHIVING DATA

The Environmental Data and Information Service (EDIS) of NOAA is responsible for five of the six data centers visited by the Panels. The EDIS has negotiated agreements with several government agencies not under NOAA jurisdiction that provide for the smooth and regular transfer of data to the NDC's. Nevertheless, many observations are made regularly by federal agencies, as well as by individual investigators, that do not find their way into the NDC's. Specific cases are cited in the panel reports that appear as appendixes to this report. In particular, the loss of

precipitation data cited in the report of the Atmospheric Sciences Panel (Appendix B) should be corrected.

The general philosophy of data-center operation places the responsibility for preparing data and accompanying documentation for archival purposes with the agency making the observation, not with the data centers. When the data collection is a routine function, this policy appears to operate well and is probably the most efficient alternative. However, difficulties are experienced when the data in question are collected as part of an experimental program. The problems experienced by the NSSDC, operated by NASA, illustrate the point.

For many years it has been NASA's policy that data being collected in experimental programs should be available to the principal investigator first and, after a reasonable length of time, should be placed in the NSSDC to be available to all interested parties. The responsibility for placing the data in the NSSDC rested with the principal investigators and the NASA Project Office responsible for the experimental programs. However, it frequently happened in the past that the principal investigators had neither the interest nor the funds to process the data for archiving after the original experimental objectives had been reached, and the NASA Project Office had frequently ceased to exist following the data-collection phase. Data often were received by NSSDC with inadequate documentation and in formats that were incompatible with the data-center's storage system. Consequently, the NSSDC staff developed the capability to reprocess the data into archival form. In spring 1977, NASA management re-established the original policy and again placed responsibility for data processing for archival purposes with the originating program and its principal investigators. The staff of NSSDC was reduced to eliminate much of the capability for data processing. Analysis of, and concern for, the impact of this action is reported in detail by the Space Sciences Panel in Appendix G.

PRIVATE DATA SETS

Concern for the problems of obtaining data reliably from experimental programs is not limited to space programs. Each of the panels independently found problems of this sort. Experimental programs, some of which continue for several years, frequently generate data sets that are unique and have great value to other investigators. However, these private data sets seldom find their way into the NDC's.

A mechanism should be created to search for and to identify those private data sets that are of sufficient value to other data users to justify the cost of documenting and formatting the data for inclusion in the NDC's. It has been suggested that the scientific discipline-based boards and committees of the NRC* are best suited to accept responsibility for identifying and certifying the value of the private data sets. Documentation and processing of the data for archiving at the appropriate center probably could be done best by those currently holding the data sets, but if this were not possible, the Center for Experimental Design and Data Analysis (CEDDA) of the EDIS is ideally equipped with personnel and facilities to undertake this task for a limited number of data sets. With some exceptions, the data centers themselves have neither the facilities nor the personnel to process the private data sets for archiving.

*Geophysical Research Board and its Committee on Solar-Terrestrial Research, Ocean Sciences Board, Polar Research Board and its Committee on Glaciology, Space Science Board, Committee on Atmospheric Sciences, Committee on Geodesy, Committee on Seismology, U.S. Geodynamics Committee, U.S. National Committee for Geochemistry, and U.S. National Committee on Rock Mechanics.

5
DATA DISTRIBUTION

Data centers, like libraries, are of little value if people do not or cannot gain access to the information residing therein. Without exception, the panels found that the directors and personnel of the NDC's and the associated WDC-A's were sensitive to the needs of data users and were cooperative and responsive to the requests received from them. Nevertheless, several improvements were suggested.

IMPROVED DATA ACCESS

Data centers may store both the basic raw data and data products obtained from different levels of processing. The greatest number of requests will be for partially processed (Levels 2 or 3)* data. Catalogs of data held by the centers are essential information to all users. These should be available in printed and, in some cases where holdings are large, magnetic-tape forms. Computer-to-computer access to

---

*Levels of data processing are defined differently by many groups. For this report we have used those that were defined for the GARP:
Level 1. Raw Data or Observations: Numbers obtained from direct reading of instruments or from conversion of telemetry signals by calibration or conversion algorithms.
Level 2. Determinations: Meteorological parameters, at the actual time and place of observation, converted to standard data formats where applicable.
Level 3. Initial State Parameters (ISP): Values of the atmospheric state parameters derived from the Level 2 determinations at the preselected set of geographic grid points for use in numerical models.

information on data availability and processing history and to the data sets themselves should be implemented in the near future to serve large-scale data users, such as the National Center for Atmospheric Research (NCAR) and other laboratories in government agencies, and universities having adequate computer resources.

Correlative studies, in which relationships between different parameters are sought, generally require data from various measurements and disciplines. Ability to access and retrieve information from different data sets, and to combine these into uniform formats, will enhance such scientific studies. Communication and direct access between data centers to complementary data sets could be an important capability that would reduce the cost and time required to obtain data service.

Most of the data centers attempt to accommodate visiting scientists and to provide them with at least minimal facilities to access the data stores within the center. This service for visiting scientists should be expanded and improved to provide user facilities for both short-term (a day or week) and longer-term (several months) studies by visiting scientists. Data centers, too, could profit by the interactions with visiting scientists, thus an expanded program of this type should be considered.

DATA SCHOLARSHIPS

Costs incidental to filling a request for data are normally borne by the requester. There are two exceptions to this policy: when the cost of filling the request is very small (less than the cost of billing), the service is usually provided at no cost; and when a request is filled through the WDC-A, in accordance with the provisions for data reciprocity as described in the *Guide to International Data Exchange Through the World Data Centers*, no charge is made. It has not been established that the cost of obtaining data has denied service to a significant number of potential users. Also, it is reasonable to question whether it is the responsibility of the data center to take special action to sponsor data users. Nevertheless, the concept of making data scholarships available to students was suggested by several of the panels. The idea would be to offer to meet the costs of data service in relatively small amounts, perhaps up to $1000 per year, to individuals on certification of their need for data by a responsible scientific authority and on agreement with the validity of the data

request by the appropriate data center. Such funds could be administered through appropriate committees of the NRC or through the data centers themselves. It is estimated that, initially, acceptable requests for such funds would not exceed $10,000 per year, and probably would never become a large cost item.

REGIONAL DATA CENTERS

The value of the NDC's is undisputed, but it has been suggested that there is also a need for regional data services to meet special needs. Examples of such regional facilities are the Environmental Studies Service Centers (ESSC), operated by the National Weather Service in four locations, which provide specialized agricultural meteorological data. The NOAA is currently considering proposals to establish regional coastal information centers to provide information on the availability of environmental data for groups concerned with coastal zones. We suggest that the need for regional centers devoted to mesoscale climatology, regional water resources, seismology, and regional snow and ice data be considered by the appropriate scientific discipline-based committees of the NRC. Such regional centers could relieve the national centers of a significant portion of their data service requests and provide improved user services at little or no additional cost.

6
The World Data Centers

The WDC system, which was inaugurated in 1957 specifically to facilitate the exchange of data collected during the International Geophysical Year (IGY), has continued during the last two decades to be a most important element in fostering international scientific cooperation. The great importance of the system in this role was recognized even during the IGY, so that in 1959 the centers were converted into organizations charged with permanent responsibility for data collection and dissemination. The operations of the centers have been guided by ICSU, through a series of "Guides to World Data Centers," which have been revised and reissued aperiodically to reflect the changing needs and priorities of the scientists using the data archived in the centers, the changing technologies available for the handling and storage of data, and to some extent the changing sources of the data and the data users.

The most recently issued consolidated *Guide to International Data Exchange Through the World Data Centers* is dated December 1973, although a summary of proposed and adopted revisions through November 1977 has been prepared by the Coordination Office of WDC-A. Several of the reviews of the branches of WDC-A, now reported, have disclosed the need for further revisions. These revisions are being considered by the relevant international bodies. These should be collated by the Coordination Office, considered by the WDC-B, and, where appropriate, by WDC-C, and acted upon by the ICSU, so that a new consolidated guide may be prepared and issued.

In particular: (a) The Glaciology Panel has recommended the change of the name of the Center to "WDC--Snow and Ice"; (b) the Solar-Terrestrial Physics Panel notes that commitments exist in the *Guide* for furnishing data to other WDC's that no longer can be honored because of budgetary limitations,

and it recommends that the Center's activities be augmented to allow this, for in view of the major international voluntary effort in sending data to the WDC's, it is of great importance that the WDC's honor *their* commitments; (c) several of the branches of WDC-A are experiencing problems in honoring the full data interchange prescribed in the *Guide* because of significant incompatibilities in the data-handling systems employed in the various WDC's, the cost of exchange of all the available photographs of glaciers, snow cover, and sea ice, and like problems; therefore, the rewritten *Guide* should recognize these difficulties; (d) the *Guide* should reflect more closely the perceived need for up-to-date indexing of all available data within the various geophysical disciplines, both those archived in the center and those known to exist elsewhere.

The Committee has determined that, in general, the branches of WDC-A are working well and are well integrated with the corresponding NDC's and other WDC's.

The Committee agrees with the suggestion that the International Gravimetric Bureau (IGB) be designated as WDC-C for gravimetry.

Recognizing the significant impact of the recent visit to the branches of WDC-A by representatives of WDC-B, the Committee recommends that reciprocal visits be made to WDC-B and to some parts of WDC-C by representatives of WDC-A. As with the WDC-B delegation, the WDC-A group should include working staff of the WDC-A and scientists outside the WDC's who are actively involved with data-collecting policy and problems. If possible, members of CDIDC should be included. The CDIDC should review the plans for these visits. In view of the division of WDC-A into seven subcenters, the visit to WDC-B should be organized in several sections. The plan may include visits not only to WDC-B but also to relevant NDC's, research institutes, and field sites.

7
Need for a National Geophysical Data Policy

Assignment of responsibility for the selection and preparation of data for archiving, the determination of the level of data to be archived, decisions in data centers on which data to retain and which to eliminate, and policies concerning responsibility for costs incurred in data retrieval are all matters currently settled by the data-center managers or by their parent organizations on a case-by-case basis, without the benefit of general guidance from the community of data users. To some extent, the planners of the large data-collection programs (BOMEX, GATE, IFYGL, FGGE, and others) have led the way in formulating data plans, which are frequently cited as useful guides in determining geophysical data archival matters.

The CDIDC believes that it is possible to formulate a national geophysical data policy that can provide comprehensive general guidelines of benefit to data-center operators and users. We propose to draft such a policy document during the next few months. We recognize that, to be effective, the national geophysical data policy must be accepted by the vast majority of NDC users, that it must be consistent with the commitments for WDC operations, and that it must be acceptable to agencies responsible for the operation of the NDC's.

APPENDIX A
RECOMMENDATIONS OF THE GEOPHYSICAL DATA PANEL

Following a recommendation of the Geophysics Research Board (GRB), in May 1974, a special study was conducted of the impact of new demands placed on World Data Centers (WDC's) and associated National Data Centers (NDC's) by large-scale geophysics programs. The Geophysical Data Panel, chaired by Carl Savit, reported the following:

"1. *The facilities of national and world data centers must be expanded and improved if they are to continue to carry out their assigned functions.*

"2. *The planners of large-scale geophysics programs are making increasing use of advanced sensor and computer technology to meet scientific objectives. The data sets so generated are orders of magnitude larger than those of programs using traditional methods of observation. This growth can be expected to continue, creating increasing demands to store data and to provide data products.*

"3. *Dealing with these data sets will create a need for large-scale computers, special-purpose processors, and information storage and retrieval systems at data centers. Such systems require sophisticated knowledge and procedures to permit access to specific data sets. Staff to provide programming services and scientific advice will be necessary both in the storage and access phases.*

"4. *Summarized, generalized, or abstracted data will be required by much of the user community. The costs of deriving such data products from raw data must be defrayed by means that will be compatible with the resources of both the generator of the data and the user community.*

"5. *In the decisions leading to acquisition of large data sets in major geophysical programs, a strategy should be formulated for data handling and archiving. Cost tradeoffs between acquisition and handling and archiving*

considerations should be an integral part of planning. The strategy should include provisions for any thinning or compression of data. The starting point in planning should be consideration of the ultimate beneficiaries of the information. Program proposals should include the strategy for and costs of archiving data and making them available to users.

"6. *General criteria should be formulated for data acquisition, taking into account, among other things, format, data quality, and handling requirements.*

"7. *Decisions to thin or discard geophysical data may be a necessary consequence of limited resources. Such decisions entail an obligation on the part of data centers to*

(a) *give notice in a timely way to the international scientific community of the intent to destroy data so as not to foreclose action to preserve the data;*

(b) *consider the relative economics of retention, thinning, and compression of data;*

(c) *preserve geophysically significant data samples or subsets;*

(d) *give priority to retention of data that are not reproducible.*

"8. *Regarding the operation of the WDC system:*

(a) *The principles and practices of operation as set down in the* Guide to International Data Exchange through the World Data Centres *should be adhered to.*

(b) *World Data Centers should be consulted in the planning of programs that will lead to new demands on the WDC system.*

"9. *Both technologically developed and less developed countries have found the World Data Centers to be a valuable resource. Future planning regarding World Data Centers should be directed toward increasing their responsiveness to the needs of less developed countries.*"

APPENDIX B
REPORT OF THE
AD HOC PANEL ON ATMOSPHERIC SCIENCES

At a meeting of the Committee on Data Interchange and Data Centers (CDIDC) of the NRC Assembly of Mathematical and Physical Sciences, on March 22, 1977, it was agreed that the members should each form an *ad hoc* panel to assess the status and effectiveness of data exchange and identify problems in collecting and making data available to the scientific community for the discipline for which they were responsible. Accordingly, an *ad hoc* Panel for Atmospheric Sciences was set up, comprising:

Carl W. Kreitzberg, Drexel University, *Chairman*
Lance Bosart, State University of New York at Albany
Roy Jenne, National Center for Atmospheric Research
Paul Julian, National Center for Atmospheric Research

The members of the Panel accompanied by Thomas O. Haig, Chairman, and Richard Y. Dow, Secretary, of the CDIDC, visited the National Climatic Center (NCC) on September 15, 1977, and received an effective tour of the facilities of WDC-A and NCC, followed by discussions with all their key personnel. Their assistance was most gracious and was appreciated.

1. *Difficulties in Procurement of Automatic Data Processing (ADP) Equipment*

The revolution in the electronics industry has permitted meteorological instrumentation to advance rapidly, with a resulting rapid increase in the amount of observational data. Therefore, the NCC must use ADP to accomplish its primary functions. Unless the NCC is permitted to keep up to date with the data collectors, a significant amount of new information cannot be achieved. The program to

provide an on-line data base that would permit rapid response to about 80 percent of the data requests that NCC receives is especially worthwhile. The Panel believes that this program has been seriously delayed by the ADP equipment procurement review process. Because the ability of NCC to keep up with its obligations is at stake, the Panel strongly recommends that delays in acquisition of ADP equipment be minimized in the future.

2. *Identification of Unique Data Sets That Should Be Archived by the National Climatic Center*

The majority of NCC data acquisitions come from NOAA organizations with routine observational programs and large international field projects affiliated with GARP. The NCC personnel do not believe that it is their responsibility to seek data collections unless a firm requirement for data has been established that their current holdings cannot satisfy. In general, they do not initiate actions to acquire data. The NCC has established criteria that incoming data must meet to be accepted; it is the responsibility of data suppliers to meet these standards. Several examples of unique and important data collections held by individuals or institutions outside EDIS, to which access by others is limited, were cited by NCC personnel and panel members. It was recognized that transfer of these collections to NCC could be of real benefit to scientists, especially climatologists, and that the collections will probably be lost if no action is taken to do so. Nevertheless, NCC staff does not consider it to be their mission to seek out these collections, nor do they have the resources to do so.

During the discussion of this subject it was pointed out that if NCC were aggressive in acquiring data collections, such activity might be considered self-serving and might increase NCC operating costs significantly. It was suggested that it might be better to place the responsibility for finding and acquiring significant private data collections with the Center for Experimental Design and Data Analysis (CEDDA). The technical personnel in CEDDA and their prior experience in preparing BOMEX, IFYGL, and GATE data for archives qualify this organization uniquely for the task. The Panel recommends that serious consideration be given by EDIS to assigning responsibility to CEDDA for seeking, acquiring, and preparing for archives significant private data collections. The CEDDA could then encourage suggestions and advice from outside scientists

through any of a number of channels, including the Committee on Atmospheric Sciences or a panel thereof.

The acquisition and handling of unique data sets would be significantly enhanced if there were an exchange of personnel between NCC, CEDDA, and other NOAA laboratories on a temporary reassignment basis. After a period of years, every NOAA laboratory would have someone who had spent time at the NCC, and the NCC would have people who were familiar with the activities of all the NOAA laboratories.

The National Climate Plan may result in the necessary attention being given to global data sets, but this cannot be taken for granted. The problem of searching out and recognizing the significance of global data sets is particularly important to WDC's and may require extra effort by both NCC and CEDDA.

There are three other categories of unique data sets that present particularly important and large problems.

(a) *Precipitation Data Collection and Analysis* Precipitation rates and accumulations are particularly valuable but difficult quantities to treat climatologically. The difficulty arises from the large spatial and temporal variation of precipitation and the need for precipitation information on many different spatial and temporal scales. For example, microwave communications are disrupted by very intense, local, short-term precipitation events, whereas flash flooding occurs on longer-time and larger-space scales. Hydroelectric power generation, agriculture, and long-term water requirements all rely on precipitation information on even larger scales.

It is only recently that quantitative precipitation forecasting has become feasible using numerical weather prediction techniques in high-resolution, limited-area models. The development of these techniques over the next few years would be greatly enhanced if the precipitation observations were systematically analyzed, using all available observations to give a clear picture of the scale and intensity of precipitation events against which the forecast models could be verified.

At the present time precipitation observations are made by many different government agencies and a variety of private agencies. The large volume of precipitation information handled by NCC represents only a portion of that available from all sources. The NCC handles precipitation data collected by the National Weather Service and the Cooperative Observer Program of the National Weather Service. The National Water Data Exchange of the U.S. Geological Survey (USGS) has valuable information on precipitation observations along streams taken by several different agencies.

Certainly, the question of systematic collection and analysis of precipitation data should become an integral part of the National Water Policy.

The Panel recommends that the Federal Committee for Meteorological Services and Supporting Research give high priority to a review of available precipitation observations and the future of the network of such observations, including the systematic analysis and archiving of these data. This review should serve as input to the development of the National Water Policy.

(b) *Regional Data Centers within the United States for Mesoscale Climatology* In the past, the federally supported state climatologist became familiar with weather observations in his locality and performed the compiling and archiving of some of the more important data sets. When this program was abolished a few years ago and the climatology responsibility was concentrated in the NCC, it was impossible for the Center to maintain the close contact with non-NOAA groups in areas around the country. During the past year, NCC has encouraged the activities of the American Association of State Climatologists (mostly state funded), which is an excellent example of intergovernmental cooperation.

The historical development of climatology has generally been on the microscale, using data from local field stations, or on the large scale, using national and international data sources. Regional-scale mesoclimatology has only recently become viable with technological breakthroughs in methods of observation and data collection and processing. Provision for mesoclimatological data bases must involve the close coordination of NCC and regional information centers of all types in all agencies. It is not yet clear how the National Climate Plan will deal with mesoclimatology or the associated data-management task. It is clear that agricultural regions and river basins, within which irrigation can compensate for more local droughts, are both on scales between the traditional large-scale climatology and microclimatology. Therefore mesoclimatology is central to agricultural problems.

It must be recognized that in mesometeorology there are many sources of weather information outside of NOAA. It should also be recognized that a great deal of satellite information is available at high resolution that would be appropriate for archiving and for mesoscale climatological studies at regional centers. Furthermore, when the weather service radar program was established, it was planned to have radar climatologies developed at each site, but this plan was cancelled because of a restriction on manpower at

the radar sites. Development of radar climatologies and satellite climatologies would be ideal tasks for regional climatic centers.

There is a recurring need for regional information covering more than a state but with more focused attention than can be provided by national information centers. Examples of such regional data centers are the Environmental Studies Service Centers (ESSC) operated by the National Weather Service in 4 locations serving about 15 states. These centers provide information for agricultural purposes only. In addition, Regional Coastal Information Centers (RCIC) are being proposed by NOAA to provide environmental data for groups concerned with coastal zones.

The Panel recommends that a review be undertaken of the need for and availability of mesoscale information that could be systematically compiled at a regional data center and of the most efficient method of interaction between regional environmental centers and NCC.

(c) *Identification of Raw Data Sets That Deserve to Be Preserved* The formation of private data collections appears to be encouraged by the tendency to archive processed data. The Panel was unable to study this possible problem in sufficient detail and recommends a more thorough study. The combination of a large volume of raw data, special equipment or techniques required to treat raw data, and limited resources tends to force planners to archive processed data sets rather than raw data. However, most of the large-volume data sets come from new sensor systems in which the data-processing techniques are experimental and will probably be improved significantly in the near future. Unless the raw data are preserved, the opportunities to test new processing techniques and to improve the archived data set are lost.

Present FGGE plans call for retention of some raw data by individuals or in various national data repositories, but only for from 3- to 5-year periods. Some raw data will be lost. This stiuation may well lead to accumulations of private data sets by concerned scientists. It is not clear that private collections are to be discouraged in all cases, but the principle of unrestricted secondary access to all such collections is difficult to maintain unless the national and WDC archival agencies exercise some surveillance over the private collections.

3. *Data Scholarships*

The NCC staff believes that the majority of customers with important data requirements can obtain the funds required by NCC for retrieval of the data. On major research projects at universities there are funds provided for data acquisition. However, there are a number of cases in which students are interested in investigating new ideas that have not developed to the point of requiring large volumes of data and have yet to be funded as regular research projects.

It would be helpful if, with a minimum of paper work, a student could apply for a small amount of data resources, that is, funds that could only be spent for data acquisition at the NCC. The request would have to be approved by a professor at the university and a meteorologist at the NCC, who would certify availability of the data at the estimated cost and its suitability for the problem that the student proposed to investigate. Individual data scholarships might be in amounts ranging from $25 to $1000 for one student in a single year. The total amount of scholarship funds could begin at an annual level of $10,000 and increase as the volume and quality of the requests for such scholarships increase.

We recommend that the EDIS and the NSF explore the feasibility of such a data scholarship program.

4. *Data Communication to Users*

The Panel wishes to record its strong endorsement of the NCC program to complete the transfer of its considerable store of station records to microfiche and microfilm. Its microfiche files will include its publications, as well as the original observation records and miscellaneous summaries and tabulations prepared by NCC. These microfiche files are easily accessible to the NCC meteorologists responsible for communicating information from the Center to the data users.

The future plans of NCC, outlined to the Panel during the site visit, demonstrate the determination of NCC further to improve its capability to effectively archive meteorological information and communicate this information rapidly to nonmeteorologists and meteorologists. Small requests are often communicated by telephone or by mail on a few sheets of paper in a short time. The care being exercised by NCC personnel to ensure data quality and accessibility, yet to keep costs low, demonstrates a high level of professional excellence.

Plans by NCC for computer transmission of information are progressing slowly. Over the next five years Automation of Field Observations and Services (AFOS) will replace the former means of communicating information within the Weather Service and between the National Weather Service and the NCC. This change will lead to a significant reduction in the number of paper observation forms that are sent from the Weather Service to NCC, with a corresponding substantial reduction in the amount of transcription from paper to digital form at NCC. It is recognized that in the 1980's there will have to be computer links between NCC and a few other NOAA organizations.

Any data system must provide for each exchange of data on magnetic tape, in most cases copied as a bit stream. In the usual mass storage device, the average size of a data volume should be at least 10^7 bits to allow reasonable throughput. This means that for the foreseeable future major sets of short reports must be stored in both synoptic and chronological order if both modes are to be used. It simply is not economical to file these data in a single order and then respond to requests in the other order.

The product of a data center is utterly dependent on its data sets, and the major ones must be as clean and well structured as possible. System improvements must be balanced with resources so that the quality of the basic data sets is not jeopardized.

As NCC designs new data-processing systems, we urge that it include enough checks in the system to ensure that none of the data have been lost or altered since they were prepared. After data leave the central memory of a computer for storage, they go through many hardware and software systems before getting back into memory for use. To guard against rare data changes in these data paths, it is desirable that major data centers keep checksums with the data. Secure backup copies of most data sets must also be available.

The problem of computer-to-computer information exchange with nongovernment facilities has not been resolved by NCC. Since plans must be made now to prepare for developing the system to operate throughout the 1980's, NCC should explore the acquisition of a computer-to-computer dial-up capability. It should keep abreast of the status of costs and capability (and likely future development) of packet-switched networks (a system of data transfer in units that provides for high-speed communication with reliability checks) to use in the exchange of programs and data between NCC and various organizations.

The cost of commercial communications satellites will continue to decrease, but magnetic tape is likely to remain the most efficient way to transfer large quantities of data that are not needed in real time. Consideration should be given to the use of geosynchronous *meteorological* satellites for computer-to-computer communication of meteorological information between different computer centers around the country, both government and private.

5. *World Data Center Operations at the National Climatic Center*

The meteorology section of the *Guide to International Data Exchange through the World Data Centers* (1973) should have been updated earlier. Some questions of mode of operation of WDC-A are raised in this section; their impact on the *Guide*, if any, should be assessed.

Adherence to the spirit of the WDC data-exchange concept is observed meticulously by NCC, and the system is working well.

WDC-A Meteorology is simply another part of NCC. There is no separate budget or identifiable line item for the WDC activities. Support of WDC-A Meteorology comes from the NCC budget, which might pose a problem during times of budget reduction. Although there is no apparent problem at this time, stability of WDC operations in the future may make it advisable to establish a budget independent of NCC.

Data exchanges with WDC-B are conducted routinely, and good communications exist. Personnel at NCC have worked out solutions to problems of data format differences as they have been identified and have shown considerable initiative in solving exchange problems. For example, during the visit of WDC-B personnel in summer 1977, NCC personnel discussed changes in rocketsonde data format and obtained Russian agreement to change to a format that can be accessed directly. As another example, the Environmental Protection Agency (EPA) stopped publishing nuclear-radiation monitoring data in 1974. Since that date, NCC has obtained the data from EPA and continued the data-exchange program. There have been no initiatives regarding data exchange with the Peoples Republic of China by either WDC-A or NCC.

The BOMEX, IFYGL, and GATE data have been well prepared either by CEDDA or to meet criteria established by CEDDA; this may not be so for FGGE. The involvement of CEDDA, and of the research community to which CEDDA relates, has so far been less evident in FGGE data planning than in GATE. Since most of the FGGE data will reach NCC from the operational

weather services, the relative absence of input from the research community poses no problem to the NCC, but the FGGE data actually archived may be far less useful for research than they might be.

Because prior planning for the large data-collection programs has been done well, and because NCC was given additional resources this year to handle the data, there appear to be no problems caused by the data-intensive programs that are not under control.

The ENDEX catalog at NCC is a good start on the "data inventory" problem, and it does include information on some data in collections outside of EDIS. An aggressive program by CEDDA to include more outside collections in ENDEX would be a significant service. The Panel also recommends two additional steps related to inventory. Summary information on data sets, including the volume of data and the number of stations, would be helpful; about a 100-page listing should be adequate for this summary. Also, a weather-station-history library is necessary that can be accessed by computers doing digital processing. This master tape would be used by other computer centers that process different data sets at different times.

No problems were identified in WDC-A operations that require further CDIDC attention at this time. However, there are problems in the parent NCC operations, which are discussed elsewhere in this report, that could affect WDC-A operations and that are of continuing interest to the CDIDC.

Appendix C
Report of the Ad Hoc Panel on Glaciology

At a meeting of the Committee on Data Interchange and Data Centers (CDIDC) of the NRC Assembly of Mathematical and Physical Sciences, on March 22, 1977, it was agreed that the members should each form an *ad hoc* panel to assess the status and effectiveness of data exchange and identify problems in collecting and making data available to the scientific community for the discipline for which they were responsible. Accordingly, an *ad hoc* Panel for glaciology was set up, comprising:

Colin Bull, Ohio State University, *Chairman*
William O. Field, American Geographical Society (formerly Director of WDC-A Glaciology)
Wesley Pietkiewicz, United States Army Cold Regions Research and Engineering Laboratory
John Hollin, Institute of Arctic and Alpine Research, University of Colorado
Paul McClain, National Environmental Satellite Service

Further input has been gained from others at the WDC-A Glaciology and in the Institute of Arctic and Alpine Research (INSTAAR), especially Mischa Plam and E. R. LaChapelle.

On July 19, 1977, members of the Panel met in Boulder, Colorado, with Roger Barry, Acting Director, and Marilyn Shartran, Assistant Director, of WDC-A Glaciology and Richard Y. Dow, Secretary of CDIDC. Alan H. Shapley, Director, National Geophysical and Solar-Terrestrial Data Center, and James F. Lander, Director, WDC-A Solid-Earth Geophysics, were also present for a part of the meeting. This meeting followed a meeting of representatives of WDC-B1, Moscow (USSR) and WDC-A, Boulder, held on June 13-14, 1977.

The review of July 19 took the form of a reconsideration of the questions that follow, most of which had been answered

briefly by Roger Barry and Marilyn Shartran at the CDIDC meeting of March 22.

Where appropriate, the information gained at the July 19 meeting has been supplemented by that received by Barry and Shartran at their meeting with members of WDC-C Glaciology, August 12-16, 1977, in Cambridge, England.

At the review meeting the questions were considered *seriatim*.

1. *Are the Provisions in the* Guide *(the* Third Consolidated Guide to International Data Exchange through the World Data Centers, *December 1973) appropriate?*

At the meeting of June 13-14, the glaciology representatives of WDC-A and WDC-B1 recommended changes to the Glaciology Section of the *Guide* to demonstrate a desire by the personnel of these two data centers to take a much more active role in the collection and dissemination of published and unpublished glaciological data than they had previously. WDC-C is not in a position to solicit minimally processed data but is willing to archive deposited data within its constraints of space and staff.

The Panel endorses the recommended changes and, in addition, recommends a title change for the Center to WDC-A Snow and Ice or WDC-A Glaciology (Snow and Ice).

2. *Are the Provisions of the* Guide *Being Met?*

The responsibility of institutions and scientists in forwarding data to the Center is not in general being met. The responsibility of investigators to supply interchangeable data to the Center is not appreciated. WDC-A Glaciology is attempting to improve this situation by describing the Center's functions in the widely distributed *Glaciological Data*. Of a total of 700 copies of the first issue, 400 have been distributed outside the United States.

Officers at NSF have indicated that budgets for individual projects could contain an item for data collection and storage. The Panel considered that the largest problems in producing effective data centers were reminding research workers to contribute their data to one of the centers and ensuring that the bibliographies in *Glaciological Data* are complete. An international list of research workers, including their current interests and projects and their contributions to the centers, would be valuable.

3. *Is WDC-A Glaciology Well Integrated with National Data Centers and Other Branches of WDC-A?*

The Panel found that significant improvements have been made in the integration of WDC-A Glaciology with other branches of WDC-A, as well as some improvement in integration with related national data storage facilities (NOAA/NESS, U.S. Navy Oceanographic Office, USGS, and U.S. Forest Service). The chief point of contact so far has been methodology.

There is no NDC for glaciology. The closest approach to such a center is the U.S. Army Cold Regions Research and Engineering Laboratory (CRREL) whose mission in glaciology is small. It stores little glaciological data, except for those generated by scientists working at CRREL.

However, integration with CRREL's bibliographic activities is progressing well. WDC-A Glaciology could help in increasing scientists' awareness of material from the Soviet Union. The CRREL bibliography indexes about 3000 Russian publications, only about 100 of which have been translated. WDC-A Glaciology intends to index all of its own holdings, including the Russian material, although it does not have the staff to undertake the preparation of abstracts of these items. The CRREL bibliography is going on-line, in cooperation with Systems Development Corporation (SDC). WDC-A Glaciology material could be incorporated in this, through indexing by the Library of Congress. This system could be compatible with the Library of Congress Antarctic Bibliography, if that, too, goes on-line. The development and implementation of an information storage and retrieval system would be greatly facilitated by the location of a computer terminal at the Center.

The Panel noted the desirability of copying the CRREL microfiche library for the period 1949 to 1969. At present only one set of this film exists.

The Panel commends WDC-A Glaciology on its continuing development of the data base and particularly on the close relationship it has developed with the CRREL Bibliographic Project.

At present two large holdings of glacier photographs exist that are of direct concern to WDC-A Glaciology: those of W. O. Field and of Austin Post.

The Panel recommends that WDC-A Glaciology pursue transfer of these collections, with adequate photograph identification and geographical indexing, to the Data Center. Microfilming of these and other collections, for archiving in WDC-A Glaciology, is highly desirable. Possible funding

sources for these projects, including some hardware, should be explored with NOAA/EDIS.

Progress should be made in integrating WDC-A Glaciology (and other WDC's) with appropriate parts of the work of the Scientific Committee on Antarctic Research (SCAR) so that the SCAR reports could include information on the material held in WDC-A Glaciology (and the other WDC's) and so that SCAR could take action to ensure that appropriate data are sent to the WDC's. The National Environmental Satellite Service (NESS) is charged with the operation of a system to observe routinely environmental conditions from satellites. The processed satellite data give information on sea ice and snow cover. Probably, WDC-A Glaciology should not attempt to duplicate this material but should have a complete *inventory* of the relevant imagery held by NESS. However, the Panel did point out the need to have some back-up to the NESS holdings. For example, the data on sea ice and snow cover for December 1967 and January 1968 have been lost.

The glaciology community is largely unaware of the holdings of materials by national agencies; therefore, the Panel is pleased that the next issue of *Glaciological Data* will contain descriptions of some of the glaciological holdings of NESS, the Canadian Ice Forecasting Centre, and the U.S. Navy, together with the access procedures.

The Panel sees a need for increased cooperation between WDC-A Glaciology, the Permanent Service on Glacier Fluctuations, and the UNESCO Technical Secretariat on Glacier Inventory and Glacier Mass Balances. WDC-A Glaciology should hold a complete inventory of the Permanent Service's holdings and should not attempt to duplicate these data holdings.

WDC-A Glaciology should generate a worldwide inventory of centers collecting data on ice and snow, along with details of their field studies. These inventories should be updated routinely. Robert Vivian, Visiting Scientist in WDC-A Glaciology, has drafted a questionnaire that could be used to generate such an inventory for glacier research and is adaptable for similar inventories in other branches of glaciology.

The Panel saw the need for increased dissemination of information on holdings of other agnecies in several fields. Sea-ice data, for example, are scattered among several agencies. The importance of variations of sea-ice cover in relation to climatic changes is now being realized. A proposal for storage of data on snow cover and sea-ice extent in WDC-A Glaciology is being reviewed by the

Glaciology Committee of the Polar Research Board and considered by NOAA (EDIS).

4. *Are Data Exchange and Communication with WDC-B and WDC-C Effective?*

The Panel considers that the June 1977 meeting between representatives of WDC-B1 and WDC-A Glaciology and the August 1977 meeting with WDC-C Glaciology to be most important and valuable improvements in communication. Further improvements could follow from a reciprocal visit from WDC-A Glaciology to WDC-B. Interchange of bibliographic data among the centers is routine, but that with WDC-B1 is not complete. WDC-C Glaciology is interested in cooperating with WDC-A Glaciology on its bibliographic data-base project, particularly in view of recent offers by CRREL to be involved cooperatively.

When WDC-A Glaciology expands its functions to include the archiving of photographs and primary glaciological data, the Center should not attempt to supply WDC-B1 and WDC-C with copies of the material but must ensure that the other centers do have complete and up-to-date inventories of the holdings. This intention should be reviewed with the ICSU Panel on Data Centers. Changes in the *Guide* may be necessary to cover the proposed change in operating procedures.

5. *Are There Identifiable Problems in the Functioning of the WDC?*

There is still a need to define closely the scope of the center responsibility for collection and dissemination, as discussed above, and to resolve some problems such as the storage and interchange of photo coverage. For example, WDC-B1 has no microfiche capabilities but can handle 16-mm or 35-mm film.

At present the WDC-A Glaciology has not defined the limits of its responsibilities in sciences that relate to glaciology. The Panel considered that, as examples, the bibliography of bedrock geology in Antarctica definitely should not be included in the Center, and that inclusion of the bibliography on the glacial history of Quebec would be questionable. On the other hand, the Panel considered that the Center should accept the responsibility of answering such questions as: "Who holds the satellite imagery of the Quelccaya Ice Cap, southern Peru?" and perhaps, even, "I intend to travel in Nepal; what glacier photographs should I take that would be of value to the Center?" The Panel

also noted the significant difference in the responsibilities of WDC-A Glaciology from those of WDC-B1 (which does not include permafrost) and those of WDC-C Glaciology (which is completely integrated into the library of the Scott Polar Research Institute).

At this stage in the development of WDC-A Glaciology, neither the Panel nor the staff of the Center can identify *major* problems that need study or advice from CDIDC.

6. *What "Guides" Govern National Data Collections outside the WDC Arrangements?*

No U.S. NDC exists for glaciology; the Panel did not see the need for such a center, distinct from WDC-A Glaciology. An overview of the needs for data collection in snow research was undertaken by a panel of the Polar Research Board (PRB) of the NRC Assembly of Mathematical and Physical Sciences. Other needs will be identified by the Committee on Glaciology.

Recently some funding agencies have written budget lines into projects specifically for the archiving of interchangeable data. The NSF could write stipulations for data collection into project grants. WDC-A Glaciology should become a major referral center and could be a collection point for appropriate data when projects are terminated. For example, WDC-A Glaciology is probably the logical place to store much of the material from the Ross Ice Shelf Project. (This matter is addressed further under question 7.)

7. *Are Arrangements Reasonably Systematically Made to Retrieve Data Collections When the Housing Institution Loses Interest?*

WDC-A Glaciology is now alerted to the necessity to collect data from discontinued projects. With projects, such as AIDJEX, in which many disciplines are involved, some meteorological, some oceanographic, and some glaciological, the Panel saw no advantage in splitting the collections. All should be housed in one accessible place, but perhaps with duplicates of the most valuable glaciological subsets available in WDC-A Glaciology. *Glaciological Data* should carry full descriptions of the material available and the mode of gaining access to it.

With U.S. projects, little difficulty should be experienced in gaining inventory information on collections for WDC-A Glaciology, but for foreign glaciological work, the only effective way of data acquisition is by advertising

the need, initially in *Glaciological Data*, and by promising reciprocity.

8. *Are Data-Management Plans Adequate for Major Data Programs?*

The Panel considered that few programs in glaciology are likely to produce the great quantities of data that are associated with some meteorological and geophysical projects. WDC-A Glaciology is now closely linked with the NOAA data storage and retrieval systems, and adequate computer time appears to be available for the foreseeable future.

The emphasis of the Center should be on gaining and disseminating information on the location and availability of glaciological data and not necessarily on the duplication of those data in WDC-A. WDC-A Glaciology, for example, should have a current file on ice cores from polar areas, with information on the work that has been done on them, by whom, and where these data are stored; it is not necessary to have the actual data. The Panel did see a need to archive in WDC-A Glaciology some unpublished basic data. The scope of such collections should be identified by the Committee on Glaciology.

9. *What Can Be Done to Improve Data Inventory?*

The Panel approved the WDC-A Glaciology plans for the systematic collection, by subdiscipline, of complete bibliographies, to be described and published in *Glaciological Data*. The retrievability of material in WDC-A Glaciology, in CRREL, and in the Antarctic Bibliography will be greatly enhanced when the SDC on-line system is operational.

Russian items in WDC-A Glaciology can be indexed, on behalf of CRREL, by the Library of Congress. Those that have already been indexed, either through the CRREL Bibliography or the Antarctic Bibliography, will not be reindexed, but new material will be placed on microfiche and indexed and the bibliographic information published within about one month. At present CRREL produces five-year indexes of its holdings, which are valuable but not nearly so useful as the forthcoming SDC on-line system.

10. *Are There Initiatives for Data Interchange with the People's Republic of China?*

The Panel considered that the most efficient way to gain access to glaciological information from China will be

through direct contact with Chinese glaciologists. At present no Chinest glaciological research is listed in either WDC-A Glaciology or in CRREL. Information on all branches of WDC-A should be given to the Chinese liaison unit at the United Nations.

11. *How Much Do Users Have to Pay for Data?*

The Panel considers that the present policies of WDC-A Glaciology are reasonable. Individual requests for reprints of material are handled in a way similar to WDC-A Oceanography. WDC-A Glaciology has an "earnings" account through the University of Colorado, but requests have been met without charge when billing costs would exceed the cost of duplication.

12. *Is There a Reasonably Clear Policy to Determine What Data Are Archived?*

A clear policy on the types of data to be archived in WDC-A Glaciology has not yet been established. The staff of the Center, however, are well aware of the need for such a policy and the guidelines are now being developed, in collaboration with WDC-B1, with input from the Committee on Glaciology of the NRC Polar Research Board.

13. *How Well Is the Center Reacting to Indicated Needs for Upgrading Data-Handling Equipment?*

The Panel considers that, other than the need for a reader-printer for microform and a computer terminal in the Center, for better access to the NOAA computer, the WDC-A Glaciology appears capable of handling such data from glaciological programs as we can now envisage. The Center has what appears to be an adequate allocation of time with the NOAA computers for its data storage and bibliographic needs. However, if major additions are made to the tasks of the Center, for example, the storage and indexing of sea-ice data, physical expansion will be necessary.

The visiting scientists program of WDC-A Glaciology has proved to be of great value to the Center. The Panel approved the idea of the Center's organizing workshops on selected aspects of data collection, storage, and retrieval, if possible with representatives of the other WDC's Glaciology being present. WDC-C Glaciology also favors workshops on data-management techniques and might be willing to sponsor one on radio echo-sounding data.

Appendix D
Report of the Ad Hoc Panel on Oceanography

On February 28 and March 1, 1978, the *ad hoc* panel visited WDC-A Oceanography and the National Oceanographic Data Center (NODC). Members of the panel were Bruce A. Taft, University of Washington, *Chairman*; and Douglas McLain, Pacific Environmental Group, NOAA/National Marine Fisheries Service (NMFS). Both are physical oceanographers, and this report is largely focused on the physical data handled at the centers. The NODC also archives chemical and biological data, but we were not able to treat the chemical and biological fields as thoroughly as was desirable.

The report will be organized around several broad areas of concern that were identified by the Panel. Following each element of the report are recommendations for dealing with the problems that we perceived.

1. *Response to New Ocean Data Types*

The NODC has the responsibility for archiving the physical, biological, and chemical oceanographic data sets that have been generated. As oceanographic science has developed, new types of oceanographic data have become available. These new types arise from both the use of new sensors, for example, *in situ* measurement of salinity and temperature, and the increase in the range of variables that are measured, for example, trace metals of importance to environmental assessment. The NODC has excellent techniques for handling many of the traditional oceanographic measurements. To serve the oceanographic community, formats and data-handling procedures will have to be developed to respond to the new demands placed on NODC. This process requires the development of criteria for data compression and storage that are generally acceptable to the scientific community. The decisions may involve loss of some information,

so that value judgments must be made in developing the procedures. Several types of physical data, for example, continuous vertical temperature-salinity profiles and current velocity (current meters and floats), which have been generated in large amounts in the past decade, are still not archived at NODC in standard format. New types of biological and chemical data are now being sent to NODC, and each of these data bases will have to be developed in a scientifically sound manner. It is obvious that decisions on these data sets cannot be postponed or the general usefulness of the NODC ocean data will be degraded.

Recommendation: High priority should be attached to the development of formats for accession of important data bases for which there are no efficient means of handling the data at NODC. In addition, similar time lags should not be allowed to occur for new ocean data bases. Steps that can be taken to deal with these problems are (a) consultation with outside experts on the effects of data compression and the proper standards of quality control; (b) use of software development capability of private industry when in-house personnel cannot be assigned to and kept on the problem; and (c) consultation with other groups (federal and academic) in the United States who have been working on similar problems and may have developed acceptable solutions, for example, Fleet Numerical Weather Central at Monterey, California.

2. *Regional Data Centers*

During the past few years, NOAA has been developing several types of regional ocean information centers. Some examples are the Sea Grant Marine Advisory Offices, the Regional Coastal Information Centers (RCIC), and the National Weather Service Marine Forecast Offices. In addition, the NODC maintains five regional offices to interact with the data center and the local oceanographic community. These regional offices are set up to promote the exchange of data, whereas the RCIC's are designed to inform users of where data may be found.

Recommendation: Believing that the regional data center concept is good and that these centers have the potential for increasing the use of ocean data, we recommend that NOAA take steps to coordinate these groups on a regional level. These regional offices should have rapid, interactive, graphic access to appropriate NOAA data facilities such as NODC. In addition, these centers should work closely with state climatologists in those states that have them.

3. *Need for Better Inventories*

As the ocean data base expands, it is clear that all data cannot be stored at NODC. However, all data should be inventoried so that data users can identify where potentially interesting data might be held. The EDIS now maintains a variety of data inventory systems. Although these systems apparently serve their intended purposes, there is little commonality of design or data flow between them.

Recommendation: The Panel believes that the evolving Data Base Management Systems (DBMS) technology might be effectively applied here. Its application to this problem appears to be appropriate, for DBMS technology allows ease of data modification and updating. Many DBMS support remote inquiry and networking and operate computers of several different manufacturers. The design goal of such a system should be that a user could easily describe his desires by key words and then quickly locate any particular batch of data, whether it be a published reference or a data set in any of several data centers. We were encouraged to find that NODC is aware of the potential of DBMS and is planning for implementation with the NOAA replacement of the 360/365 computer.

4. *Expand Coverage of Biological Oceanographic Data*

The NODC now has adequate systems for handling many types of physical-chemical data but still has limited amounts of biological oceanographic data and relatively primitive programs for validating such data. In view of the increasing public concern about the effects of climate and man's activities on marine biological populations, we anticipate a considerable increase in the volume and demand for this data base. There is a great range in the types of data; therefore, a wide range of expertise is required to deal with this complex data base. In response to this problem, NODC has undertaken to establish a common taxonomic code for all data. At present, fishery catch and effort data are not collected nor inventoried by any central data center. It is generally felt that a substantial amount of biological data is now unrecoverable because there has been no provision for systematic inventories and collection.

Recommendation: An increased effort should be made to inventory and archive biological data, including data on chemical constituents in organisms. At present, there are four staff biologists at NODC. Consultation with outside experts will be required to carry out this program. When

NODC was formed, it was assumed that eventually all oceanographic data could be assembled in NODC's files. This goal proved to be impossible, and other data centers were established, for example, marine geological and geophysical data are handled at a separate facility. In dealing with the problem of biological data, the possibility of a separate facility should be considered. This alternative will become attractive if the biological data set becomes large and requires a major effort that does not fit into the structure of the present NODC. Because of the large differences in biological and, especially, fishery data between regions, perhaps regional biological data centers might be a possibility. We understand that certain of the NOAA-NMFS regional fisheries centers already have some data archival abilities, and they should cooperate with NODC on archiving fishery catch and effort data. Similarly, NODC should cooperate with NMFS and EPA on archiving data on chemical constituents in marine organisms. The NOAA should take the lead in determining what should be done with fishery catch and effort data; NODC should increase its efforts to inventory biological data sets.

5. *Need to Work Closely with Users*

The NODC must work closely with its users to ensure that its services continue to meet needs. Two specific benefits of such interaction are quality control and development of new products. Although NODC can perform many quality-control checks by itself, many errors are discovered only in the detailed analyses and plots made by users. The need for new products and revision of old ones becomes obvious only through interaction with users.

Recommendations: To increase contact with the user community, we believe that the following steps should be considered: (a) NODC users could provide annual reviews of their programs to a representative group of users and benefit from their critical assessments. (b) Selected users could be invited to visit NODC for periods of several months or longer. The facilities and cooperation of NODC would be available to them in such a way that these individuals could carry out their research more effectively by coming to NODC. (c) NODC could assemble small groups of experts for advice on specific problems. To obtain the services of the best people, a consultant's fee could be offered.

6. *Need to Expand World Data Center-A*

The mechanism for data exchange provided by the WDC system is essential for ocean scientists. For many countries active in marine sciences, the exchange of data outside this system would be impossible. Communication between WDC's-A and -B is already good and is improving. However, it appears that all the data from the Declared National Programs (DNP) are not being submitted to the WDC's. There are large batches of certain kinds of data that have never been received. For instance, no U.S.S.R. *in situ* temperature-conductivity measurements have been put through the WDC system.

The WDC-A has operated with the same size staff since 1969, and its data-accession system has not been automated. During this time the volume of data exchanged has increased sharply and continues to increase. We were disturbed by the relative slowness with which foreign cruise reports are entered into the NODC inventory system. To respond rapidly to demands created by new international scientific programs, the data-managing capabilities of WDC-A must be increased. It does not appear to be reasonable to expect NODC, however, to devote more resources to solve the problems of the WDC system.

Recommendation: The capability of WDC-A should be enhanced. Because WDC-A is dependent on NODC ADP equipment, a speedup of data accession will require a commitment of resources by NODC. A data entry terminal located at WDC-A should be considered. If a data entry terminal were available to WDC-A, it could enter foreign reference reports as they are received. An effort should be made to encourage nations to ensure that all DNP data are contributed to the WDC system. In the United States, the NSF makes continued funding of scientists conditional on timely submission of data to EDIS. This practice should be extended to all federal agencies. To explore common problems, some interaction between the WDC-A's would be beneficial. At present they communicate little; we believe that it would be mutually beneficial if the directors were to meet occasionally. The frequency of meeting would be dictated by the usefulness.

7. *Ability of NODC to Adjust to Impact of Large Programs*

During the last decade the NODC has been asked to respond to data-handling requirements generated by large, intensive, oceanographic programs, for example, POLYMODE, NORPAX, and OCSEAP. It has been imperative for NODC to be aware of the

advent of these programs and to anticipate their impact. NODS personnel have participated in planning meetings so that data-archiving aspects are not ignored. Sometimes managers of short-term programs, such as OCSEAP, request NODC data-handling support and are able to offer funds to cover the expenses. However these funds are not accompanied by an authorization to increase personnel ceilings. Thus the additional work must be done by the present staff, and many important data-base development programs are interrupted.

Recommendation: We believe that NODC should use outside contractors to a greater extent that it has in the past. In this way it can reduce some of the negative impact of accepting reimbursable funds to carry out large data-handling programs. Some combination of contracting efforts, judicious use of consultants, and internal reorganization of the NODC staff would permit it to do the jobs for which it is uniquely equipped without delaying continuing work.

APPENDIX E
REPORT OF THE
AD HOC PANEL ON SOLAR-TERRESTRIAL PHYSICS

At a meeting of the Committee on Data Interchange and Data Centers (CDIDC) of the NRC Assembly of Mathematical and Physical Sciences, on 22 March 1977, it was agreed that the members should each form an *ad hoc* panel to assess the status and effectiveness of data exchange and identify problems in collecting and making data available to the scientific community for the discipline for which they were responsible. Accordingly, an *ad hoc* Panel for Solar-Terrestrial Physics was set up, comprising:

Sidney A. Bowhill, University of Illinois (Chairman)
Donald A. Gurnett, University of Iowa
Erwin Schmerling, NASA Headquarters

The Panel visited both the National Space Science Data Center (NSSDC) and its associated WDC-A for Rockets and Satellites at Greenbelt, Maryland, and the National Geophysics and Solar-Terrestrial Data Center (NGSDC) and its associated WDC-A for Solar-Terrestrial Physics (STP) at Boulder, Colorado.

Because it is difficult to separate the activities of the NGSDC from the WDC-A for STP, and the archives in the NSSDC are available from the WDC-A for Rockets and Satellites, when speaking of these data centers in this report, we will refer to the Boulder complex as the STP Data Center and the Greenbelt complex as the Space Science Data Center.

KEY RECOMMENDATIONS

1. The STP Data Center has established a unique facility at Boulder for handling and disseminating data. We *recommend* that its activities be augmented in the following ways:

(a) All Level 3* solar-terrestrial data, whether taken from the ground or from space, should be available at Boulder.

(b) The STP Data Center should advertise its willingness to take custody of data sets now held in individual institutions, where the scientific interests may have changed, resulting in less active use of the data at that institution.

(c) Scientific use of the data at the STP Data Center should be promoted by upgrading the retrieval and processing capability of the Center, by encouraging and financing visiting scientists, and by stimulating scientific activity (including attendance at scientific meetings) by center personnel.

(d) Catalogs of data holdings should be updated to reflect the present status.

(e) Commitments of the *Guide* for furnishing data to other WDC's that are not currently being met because of budget limitations should be honored.

(f) Planning should be initiated for the extensive data requirements of future international programs such as the Middle Atmosphere Program (MAP).

(g) Since most of the records at the STP Data Center would be difficult or impossible to replace in case of loss, and the present archives are inadequately protected against fire or flood, duplicate copies should be made of all data where practical and archived in a secure place.

(h) Studies should be made of an overall data-management system to improve the capabilities of the STP Data Center to handle a much increased volume of data.

2. The WDC-A for Solar-Terrestrial Physics has maintained contact with the other WDC's by visits, dialogues at scientific meetings, and data exchange. We *recommend* that these contacts be expanded, particularly in establishing common formats for digital data, ensuring more timely data exchange, expanding the types of data that are exchanged, and exploring the possibility of digital data links between WDC's within the United States.

3. Substantial additional funding will be necessary to implement the above recommendations, and user charges cannot cover these costs. We therefore *recommend* that the EDIS request a major augmentation of funds for the STP Data

---

*See footnote in section on "Improved Data Access" in Chapter 5, where data levels are defined.

Center for this purpose, and we *further recommend* that other interested agencies (NASA, NSF, DOD) also provide support.

4. We *recommend* that CDIDC continue to offer assistance to the STP Data Center in planning its priorities for the future.

SUMMARY

The collection of data at the two Data Centers has a somewhat different character. The data collection at the Space Science Data Center consists mainly of data resulting from space experiments that have been telemetered to a ground station and therefore exist in digital form on magnetic tape from the outset. Most of the data in the STP Data Center collection were derived from ground-based experiments, and many of them exist as hard copy (pages of numerical data) or as photographically or mechanically recorded graphical material. This difference implies further differences in the way data are handled, stored, retrieved, and disseminated in the two Centers. With the increasing use of automated digitization of ground-based experiments, however, this distinction is beginning to disappear.

There seems to be no clear distinction of solar-terrestrial data that are archived at the STP and Space Science Data Centers. Since the international contacts of the WDC-A at Boulder are far more extensive than those of NSSDC at Greenbelt, and we believe centralization of such data to be desirable, we *recommend* that all level III solar-terrestrial data be available at the STP Data Center, whether or not they were taken from space vehicles. This recommendation should not be taken as a downgrading of the role of the NSSDC at Greenbelt. By far the greatest part of the data contained there is of levels I and II, or data that have been either insufficiently processed to be used by anyone except the original investigator or are still part of an active spacecraft program; it is clearly better that such data should continue to be retained and handled at the NSSDC. However, level III solar-terrestrial data should clearly be held at the STP Data Center to assist in correlative studies.

The use of the STP archives continues to increase. The STP Data Center is also being pressed to take on new files, thus increasing the archiving and the processing and user service costs. More users are taking advantage of the equivalent exchange principle, which costs more money. The

Data Center is being asked to do more central data processing and to do more elaborate data reformatting. Inflationary increases are only partially compensated. All these demands put increasing pressure on the Center's base funding. Those concerned with resource allocation have not been able to take this into account; rather, the pressure has been for reductions, particularly of staff.

Increased use of computers increases efficiency and effectiveness but does not decrease the need for operating funds and people. The NOAA is providing the STP Data Center with a minicomputer that will further increase effectiveness and is providing funds for the IMS period for archiving, checking, and some data products. But the computer needs and computer costs steadily rise, thus putting added pressure on resources available for essential work by human beings.

The STP Data Center has for some time lost marginal employees through retirement, illness, and the like, and in many cases has filled these places with more-junior assistants. The Center is doing about as much as it can through contracts, while trying to maintain its effectiveness.

With all these pressures for user services, increased use of computers, and the like, it has not been able to do a good job with or even keep up with the basic archiving activities. Manpower and funds are insufficient to put into effect a common data-management system. Further studies are needed of the priorities within the STP Data Center among

1. Response to routine user requests
2. Response to nonroutine user requests
3. Regular data compilation (SGD)
4. Nonregular data compilation (UAG)
5. Efforts to get more data and inventories into computer format
6. Clean up and completion of inventory of existing archives
7. Search for and inventoring/archiving of new data files
8. Support for taking of key observations
9. Quality review of incoming data
10. Quality review of data in the archives
11. Organizing and advertising of "odd" data sets
12. Summarizing of data (e.g., geomagnetic indices)
13. Development of new data products
14. Assistance to visiting scientists
15. Cost/benefit studies on data services.

RESPONSES TO QUESTIONS

1. *Are the Provisions in the* Guide *Appropriate?*

The STP section of the *Guide* is among the more detailed and carefully prepared portions and fills the requirements of the community satisfactorily. Unfortunately, the Rockets and Satellites section of the *Guide* is much less detailed and is not really helpful for dissemination of solar-terrestrial data from space. It might be better if such data were described in detail in the STP section of the *Guide*.

2. *Are the Provisions of the* Guide *Being Met?*

Constant vigilance and followup are exercised by WDC-A to maintain continuity of participation by contributing institutions, generally with good success. WDC-A has itself been unable to comply with the provisions of the *Guide* because of budget limitations; we *recommend* that it begin to comply as soon as is practical.

3. *Is the WDC Well Integrated with National Data Centers and Other Branches of WDC-A?*

We *recommend* complete integration of the international activities of WDC-A with the activities of NDC's. Although liaison with other branches of WDC-A seems satisfactory, we *recommend* a redeployment of responsibility between the STP and Space Science Data Centers as described in the Summary.

4. *Are Data Exchange and Communication with WDC-B and WDC-C Effective?*

Contacts between WDC-A and WDC-B are improving but fall short of the complete data compatibility envisaged in the *Guide*. We therefore *recommend* increased contact with WDC-B and WDC-C at both the technical and administrative levels.

5. *Are There Identifiable Problems in the Functioning of the WDC?*

The dominant problem in the functioning of WDC-A for STP is the lack of adequate funding. This situation is described in detail in the Summary. Our recommendation (c) speaks to this point.

We found the practice at WDC-A of making extensive in-house scientific use of the data to be extremely valuable in improving the quality of the data and enthusiasm of the personnel. We therefore *recommend* that this aspect be augmented, including expanding the program of visitors to the Center. We were disturbed to find that funds are not available adequately to protect irreplaceable data at WDC-A from loss by fire or flood, and we *recommend* that duplicate copies should be made as soon as possible and archived in a secure place. The extensive and expanding data holdings of WDC-A require a data-management system to be implemented, for which it clearly has no funds. We *recommend* that this system be implemented. The Summary outlines further matters on which the advice of CDIDC is needed, and we *recommend* that this advice be provided as appropriate.

6. *What "Guides" Govern National Data Collections and Other Major Collections outside the WDC Arrangements?*

Formal agreement exists with the Air Weather Service (AWS) for DMSP and SOON photographs, and there are informal arrangements with NSSDC, USGS, NOAA/SEL, and other agencies for data exchange. However, these arrangements fall short of a national data plan; consideration should be given to whether such a plan is necessary or desirable.

7. *Are Arrangements Reasonably Systematically Made to Retrieve Personal/Institutional Data Collections when the Housing Institution Effectively Loses Interest?*

WDC-A has taken the initiative in acquiring some data sets of this kind; for example, from the Department of Terrestrial Magnetism and from McDonnell-Douglas. However, we *recommend* that its willingness to assume responsibility for such data sets be widely advertised.

8. *Are Data-Management Plans Adequate for those Programs whose Overall Planning is Well Advanced? For Programs at an Earlier Stage, Are the Planners Well Informed about the Savit Panel Recommendations? Are "Independent" Data Sets Being Generated and Archived by Individuals and Groups outside the NDC-WDC?*

The major data plans for IMS involve the ground-based magnetometer network and the DMSP satellite photographs. Both of these are about to be implemented and seem satisfactory.

For the NOAA and GOES satellites, implementation is under way.

For the MAP, the international Steering Committee has just been appointed, and a Data Management Panel will be appointed shortly. This Panel will develop plans for the different kinds of data handling required by MAP; for example, implementation of new data sets such as ozone and other middle-atmosphere composition measurements. We *recommend* that WDC-A devote some additional resources to planning for this activity. Some independent data sets exist, but these Centers are largely for local use, hence not duplicative.

9. *What Can Be Done to Improve Data Inventory?*

Some progress has been made toward issuing detailed catalogs (e.g., UAG-49 for geomagnetic data, UAG-54 for vertical sounding data), but additional catalogs are needed. We *recommend* that WDC-A put further effort into preparing these catalogs and disseminating them widely.

10. *Is There Data Exchange with the People's Republic of China?*

At present there are no data-exchange agreements with the People's Republic of China regarding STP data.

11. *How Much Do Users Have to Pay for Data?*

The charges made for STP data recover some fraction (about 20 percent) of the cost and do not seem excessive. We *recommend* maintenance of the current billing practices.

12. *Is There a Reasonably Clear Policy to Determine What Data Are Archived?*

Policies for archiving data are determined by the *Guide*, but there has been no need for thinning of data so far.

APPENDIX F
SOLID-EARTH GEOPHYSICAL DATA

Solid-earth geophysical data cover a broad range of disciplines and measurements that include seismology, tsunamis, gravimetry, earth tides, recent movements of the earth's crust, magnetic measurement, paleomagnetism and archeomagnetism, volcanology, geothermics, marine geology and geophysics, and isotope age dating. Coverage of such wide areas is important for studies that require synthesis of complementary data sets. Amounts and requirements of data generated by these fields vary greatly. For example, in geothermics less than about 1000 new measurements are added each year to the present data set of 6000. In seismology, however, the present data rate exceeds 10^{14} bits/year. The tsunami data set is modest, yet for proper interpretation it requires the seismology input. Recent crustal movements are strongly tied to seismicity. Proper interpretation of oceanic heat-flow data requires other data sets from marine geology and geophysics. Thus the archiving plans and data centers need to cope with this diversity of disciplines and data sets.

This appears to be a time of major changes in solid-earth geophysical studies. Amounts and types of data produced follow these changes. Geophysical data centers must evolve to meet the new requirements. A few examples may clarify the situation. In marine geology and geophysics, the data base has increased significantly for a number of reasons. Technological innovations have opened the opportunity for continuous shipboard data acquisition; in the past only a limited number of measurements could be made. Satellite altimeter and gravity measurements over the oceans and continents are reaching the resolution of great importance to geophysics. It is anticipated that satellite data will increase even more rapidly in the coming years. Resource-exploration incentives have resulted in multichannel marine

seismic profiling over about 20,000 km of the continental shelf of the United States. Approximately 10,000 reels of digital seismic data tapes have been generated in the past few years. These data, of great scientific and economic importance, are increasing rapidly.

Another major development has been in earthquake seismology. The earthquake prediction program required the installation of a number of dense seismic networks. More than 500 seismic stations are currently operating in California. The number of stations is increasing in other areas of the United States, as well as in other countries subject to earthquake risk. Data acquired from these networks are important for earthquake prediction, as well as for other scientific studies, including detailed investigations of the structures of the earth's crust and mantle, earthquake mechanisms, and plate tectonics. These network data, which are rapidly being converted to digital form, increase by orders of magnitude the data rate in seismology.

It is clear from this discussion and the cited examples that the two major developments in solid-earth geophysical data are (1) the rapid increase of the amount of data and (2) a shift from primarily analog form to digital data. These developments, although providing a great opportunity for scientific studies, are also creating major problems in data management.

PROBLEM AREAS IN SOLID-EARTH DATA

The impact of a rapid increase in data, especially digital data, is being felt in many fields. Solid-earth geophysics, however, faces some unique problems.

1. *Rapid Increase of Data, Especially Digital Data*

The increase of the data has been unprecedented. The amount of level 1 (raw) digital seismic data from local and global seismic networks will be about 10^{15} bits/year. Even in processed and condensed form, these will amount to 10^{13} bits/year or 10,000 reels of high-density magnetic tape. Seismic profiling in outer continental shelves of the United States has already produced 10,000 reels of data, and this amount is increasing at the rate of 40 percent per year. These data are described in greater detail in the addendum to this report.

2. *Diversity of Data Sources*

Data acquisition and research in solid-earth geophysics are being supported under a wide range of programs by different agencies. The USGS, NSF, NOAA, DOD, NASA, and private institutions contribute to the exploration and data-acquisition program. A large number of small and some large projects contribute to the data base. A majority of the programs do not include a long-term, comprehensive, data-management plan. Provisions and schedules for orderly archiving of data are not part of the contracts in many cases.

3. *Separation of Data Center and Data Sources*

The present situation is such that a majority of solid-earth data are generated or supported by agencies other than NOAA. The majority of the data users are also outside the NOAA domain. Theoretically, this situation should not introduce any problems. In practice, however, the NGSDC has only small leverage on the data generators. Without specific agreements with other agencies, NOAA and NGSDC can play only a limited role in data-management plans.

4. *Changing of Agency Roles*

A transition is taking place in the traditional roles of agencies responsible for solid-earth programs. The DOD (or DARPA), which had a major responsibility for the installation and support of worldwide and other seismic networks under the VELA-Uniform program, is taking a less active role. The NOAA, which had principal responsibility in seismology at one time, does not have a major program at present. In these major transitions, responsibilities are blurred. There is a real danger that networks and data collected from networks could be lost.

5. *Limited Resources of NGSDC*

The NGSDC, especially its solid-earth group, is facing the problem of increased data with inadequate resources. Manpower is limited, and distribution of specialties is uneven. It is difficult to cover all diverse fields. Although in marine geology and geophysics staffing appears to be adequate, in other fields it is subcritical. The greatest need appears to be acquisition, management, and distribution of digital data. The magnitude of this task is beyond

the present manpower, equipment, and financial capability of NGSDC. Furthermore, there is no established precedent or policy in solid-earth geophysics for the acquisition and management of digital data. This is a task that needs to be faced by the scientific community, CDIDC, and the EDIS together. Some guidelines and policy need to be developed before implementation plans can be made.

RECOMMENDATIONS

The solid-earth data problems are substantial and dealing with them requires some major actions. General recommendations on this subject include the following:

1. The Committee on Data Interchange and Data Centers should collaborate with other NRC committees that have a direct interest in solid-earth data (e.g., Committee on Seismology). Immediate problems of preserving and archiving existing data sets should be addressed.
2. Dealing with solid-earth problems requires an interdisciplinary approach and interagency coordination. This can best be achieved if there is a National Data Policy.
3. Funding agencies and project planners should require data management and archiving plans at the onset of data-intensive projects. Resources for archiving, preservation, and distribution should be included in the planning.
4. Impacts of present and expected large data sets should be assessed: decisions must be made on what to save and plans developed for storage, management, and distribution.
5. Some complementary data sets are being stored in different data centers. For example, gravity and magnetics satellite data are stored at NSSDC, and land and marine data are at NGSDC. A cross-referencing system for all data holdings should be developed.
6. Remote computer access to the most-used data sets, availability information, and indexes should be developed. This access should be extended to reduced and partly processed (Level 2 and Level 3) data.

ADDENDUM TO THE SOLID-EARTH DATA REPORT

In this addendum we provide more-specific information from the working group reports on seismic data and marine geophysical data. These represent the two largest categories in solid-earth data.

A. Seismic Data

At present there are large quantities of seismic data. Only a small part of these is readily available from recognized data centers. Even larger quantities of data are likely to be generated as more digital seismic stations become operational.

The data are classified in the following levels:

- Level 0: instrument location, response, and configuration
- Level 1: raw seismic data
- Level 2: event-associated data (waveforms observed for each event attached in some way to the desired event parameters)
- Level 3: station phase data (here, instead of the waveform itself, a series of parameters describing the waveform, such as arrival time, amplitude, and period of each phase, are retained)
- Level 4: event lists (containing the derived values of source parameters)

Seismic data were discussed in the following categories, and specific recommendations were made when appropriate.

1. *Global Analog Data*

The current collection of microfilmed copies of seismograms from the WWSSN at EDIS works well and should be continued for the foreseeable future. Similar data from other stations are available for special-interest events.

2. *Global Digital Data*

Within one or two years, the United States will be operating between 30 and 40 digital stations around the world. The digital Level 1 data from these stations will be assembled in the form of a network-day tape (each day of recording will occupy two 2400-foot tape reels).

Recommendation 1: The NGSDC/EDIS should archive these network-day tapes and duplicate them for users at the lowest allowable cost.

Recommendation 2: Analog versions of these digital seismograms should be microfilmed by NGSDC/EDIS and maintained along with the WWSSN library.

The agency currently operating this digital network is DARPA (transfer to USGS is expected in fiscal year 1979), and efforts are under way to develop a Level 2 data format. This should be operational during fiscal year 1980.

Recommendation 3: Event-associated waveform data, perhaps for some subset of special-interest events yet to be defined, should be maintained at NGSDC/EDIS for general distribution.

The introduction of digital data opens many new possibilities for the storage and distribution of data, many of which will depend on facilities available at the Data Center.

Recommendation 4: The Data Center should explore the possibility of producing seismic data according to user-specified requirements. This task might include special data subsets or special formats.

The working group is also concerned about the hardware and software for researchers.

Recommendation 5: The large new digital data base may require the provision of more extensive computer facilities than those that are currently available to many seismologists. We urge that possibilities such as a central computing center, with associated data facility and computer network or telephone access, be explored.

Recommendation 6: We recognize that software development is a major and expensive task and urge that attempts be made to centralize the documentation and distribution of seismic software, particularly for routine data processing.

In addition to the U.S. network of digital stations, there are already a number of foreign digital stations, and this number is likely to increase in the next few years.

Recommendation 7: For certain special-interest events, data from as many digital stations as possible should be accumulated at the Data Center.

The proper designation of special-interest events needs some modification for digital networks, since many digital stations will clip for the WDC special events ($M \geq 7.5$).

Recommendation 8: A small panel should be established at the NDC to identify special-interest events.

3. *Data from Seismic Arrays*

A large amount of data has been, and continues to be, generated by U.S. seismic arrays.

Recommendation 9: The data from large seismic arrays, particularly LASA and NORSAR, form a valuable and irreplaceable data set. We urge that a panel explore ways to preserve and archive these data.

4. *Historical Seismic Data*

There has been a resurgence of interest in the instrumental seismic data collected prior to 1960. As much of this data set as possible should be systematically preserved.

Recommendation 10: We support the IASPEI resolution concerning historical seismic data. We particularly urge funding agencies, such as NSF and USGS, to support this collection task.

5. *Regional Seismic Networks*

The amount of data currently being collected by regional seismic networks for microearthquake and earthquake-prediction studies is enormous. One estimate for the size of this data set is as follows:

Level 1: 10^{14} bits/year
Level 2: 5×10^{12} bits/year
Level 3: 2×10^{9} bits/year
Level 4: 4×10^{6} bits/year

We estimate that at least five times these annual totals already exist, and we are concerned that many of these data are not properly archived.

Recommendation 11: The task of preserving data already collected is large, and we can see no easy solution. We urge that a panel explore this problem.

Recommendation 12: We recommend that steps be taken to ensure that future data of this kind are archived; the cost of archiving is small compared with the cost of collection. Funding agencies should be strongly urged to require that all proposals to gather data be accompanied by plans and funds (a few percent of the total proposed) for data archiving. General guidelines should be formulated for all government funding networks. In general, the archiving should take place at the institution collecting the data.

Recommendation 13: Level 4 (and perhaps Level 3) data should be stored in the NDC.

Recommendation 14: We recommend the general adoption of the following procedure: (a) All data collected with government funds should be made available to the scientific community within a fixed period (perferably one year) after collection, and (b) any paper based on scientific data should contain a statement about the availability of the original data on which the paper is based.

6. *Special Experiments and Events*

We are concerned that data from certain large-scale experiments may not be preserved properly for future reanalysis. Examples include the Early-Rise experiment, data from the LRSM stations, and aftershock data from unique events such as the Alaskan earthquake. The quantity of such data is exceedingly large.

Recommendation 15: We urge that a panel review the current procedures for preserving data from special events and experiments and ensure that as much as possible of this data base is archived.

7. *Macroseismic and Strong-Motion Data*

The NGSDC/EDIS have developed appropriate procedures for the handling of these data; therefore, we offer no recommendations.

8. *Seismic Exploration (Reflection)*

Large amounts of data have been, and continue to be, collected during assessment of offshore oil and gas reserves. Those portions that are acquired by government agencies should be preserved as long as possible. This statement also applies to large-scale reflection experiments on land.

Recommendation 16: That marine and land-reflection data be preserved. We urge that a panel explore suitable procedures for accomplishing this.

9. *Seismic Exploration (Refraction)*

We are concerned that a large amount of raw seismic refraction data is not being systematically retained in accordance with the WDC *Guide*.

Recommendation 17: A panel should explore the problem of the preservation of the raw seismic data collected in land and marine-refraction experiments.

10. *Global Earthquake Catalogs*

Global event lists (Level 4 data) are currently distributed by USGS and the ISC. The ISC also publishes, in its *Bulletin*, Level 3 data, as does USGS in its EDR reports.

Recommendation 18: We recognize some difficulties in the general availability of computer-compatible versions of the Level 3 data (detailed station reports). We urge that methods be explored for the archiving and distribution of these data at NGSDC/EDIS.

Recommendation 19: We notice one type of information that is not currently available, namely, an inventory of available seismic data for a given event. Ideally, these data should be accessible via computer link. We recommend that NGSDC/EDIS explore the routine compilation of this type of data.

Recommendation 20: When large events occur, there is usually a large amount of research into a variety of source parameters. We recommend that these results be accumulated at the Data Center for events designated of special interest. Results should be accompanied by appropriate references to the source of the material.

B. MARINE GEOPHYSICAL, GRAVITY, MAGNETICS, AND CRUSTAL-MOVEMENT DATA

In marine geophysics, gravity, and magnetics, data are gathered by different organizations. In some cases large amounts of data result from a single experiment. In most cases, however, data gathering is a continuing process. The support agencies and data-gathering organizations are diverse. Thus identifying sources, indexing all available data, and archiving and distribution are all important tasks.

General Recommendations

Recommendation 1: Indexing: In addition to indexing currently available data, WDC-A should develop and maintain an up-to-date index of all available global gravity, magnetic, marine-seismic, and crustal-movement data. For this,

the WDC-A should develop a list of agencies that can be contacted for such data. The index should be updated as appropriate, and it should be computer accessible.

Recommendation 2: WDC-A should make every effort to archive all recent, high-quality data, especially those from projects in which large quantities of data are obtained in a uniform manner. The three-component aeromagnetic data from Project Magnet are an important example. All organizations engaged in such projects should be strongly encouraged to make their results available to the WDC-A in a mutually acceptable format suitable for archiving.

Recommendation 3: WDC-A should request that contributors of data provide information on the quality and accuracy of individual measurements, as well as information on types of instrumentation and navigation used. If this information is not provided, WDC-A should advise users of the data. The "Marine Geophysical Data Exchange Format--MGD 77" key to geophysical records documentation no. 10, published by NGSDC, September 1977, provides a suitable format for this information.

Recommendation 4: Principal funding agencies should be encouraged to provide support for the storage of large data collections assembled either by in-house activities or those supported at institutes and universities and, where appropriate, to provide at least the initial support for transfer to WDC-A. In all major experiments, plans and budgets for data archiving and distribution should be included.

Recommendation 5: WDC-A should seek broad technical advice from a panel or an advisory group with members representing the various disciplines concerned.

Specific Recommendations

1. We *recommend* that

(a) Present activities in the management of Project Magnet data be maintained and, further, that the Project Magnet analog data be digitized, wherever possible;

(b) The Center either manage the permanent storage and distribution of the NURE magnetic surveys of DOE or make arrangements with DOE for storage and distribution of these data;

(c) The Center either manage the permanent storage and distribution of the USGS offshore aeromagnetic data or make arrangements with USGS for storage and distribution of these data subject to any contract constraints on them.

2. We *recommend* that WDC-A increase its efforts to serve as an NDC for land-based gravity observations. In particular, the WDC-A should conclude an arrangement with the National Geodetic Survey whereby it can acquire and distribute descriptions and gravity values for stations in the U.S. National Gravity Base Net. This agreement should make provisions for receipt on a regular basis of updates to the net. With respect to regional gravity data, we *recommend* that WDC-A pay particular attention to acquiring information relating to the quality of the observations. A clear indication of the lack of this information should be given.

3. The International Gravimetry Bureau is recognized as the data center for the International Association of Geodesy. Because of its importance, it should have a more formal association with the WDC system. We therefore *recommend* that the IGB be designated as WDC-C for Gravimetry.

4. We *recommend* that WDC-A monitor closely the activities of NASA in satellite radar altimetry. In addition, we *recommend* that WDC-A, in conjunction with the appropriate national committees, monitor the need of the geophysical community for altimeter data and, if such a need is indicated, acquire these data. We believe that, with the high sampling rate and coverage of all oceans, altimeter data could become useful to geophysical research.

5. We *recommend* that the WDC-A be encouraged to maintain a collection of publications and reports and an index of available data sources in recent crustal movements. These data are to include mareograph, tilt and strain meter records, secular gravity, level lines, and other survey networks. Where there is danger that important records may be lost, the Center should be prepared to make arrangements to preserve such data.

6. We recognize that there is a major problem in the acquisition by WDC-A of large amounts of multichannel, seismic-reflection profile data. We *recommend* that a panel consisting of scientists, computer programmers, and representatives of the principal funding agencies be appointed to assess this problem. We *recommend* that WDC-A continue to archive single-channel seismic data and to increase its holdings by acquiring data currently held in IPOD and DSDP data banks.

APPENDIX G
REPORT OF THE
AD HOC PANEL ON SPACE SCIENCE DATA

PURPOSE AND SCOPE OF THE REPORT

Two basic objectives of the Committee on Data Interchange and Data Centers (CDIDC) of the NRC Assembly of Mathematical and Physical Sciences are (a) to assess the status and effectiveness of international data exchange through the WDC system and associated activities of the related NDC's and (b) to look ahead to future opportunities and challenges. To accomplish these objectives, the CDIDC established several *ad hoc* panels, one of which was charged with advising it on matters related to the National Space Science Data Center (NSSDC) and associated WDC-A on Rockets and Satellites. Panel members included Juan G. Roederer, University of Alaska, *Chairman*; James W. Head III, Brown University; and George A. Paulikas, Aerospace Corporation. The Panel addressed the following question: Are there identifiable problems in the functioning of the Data Center? This report presents the Panel's findings and recommendations, based on its May 13, 1977, visit to the NSSDC.

SUMMARY

A valuable and enduring asset of a space mission is the information contained in the acquired data base, which future investigators can analyze and use in ways not yet imagined. The immense volume of space data accumulated is a national treasure that ought to be given as much attention as the missions themselves. The NSSDC and the associated WDC-A for Rockets and Satellites play an extremely significant role in space-data archiving and national and international distribution. The *ad hoc* Panel on Space Science Data, noting that the NSSDC has recently been

plagued by problems of reorganization and reduction that have begun to jeopardize seriously the efficiency of its operations, *recommends* that the CDIDC advise NASA to take the following steps:

1. Initiate immediate action to stabilize the present situation of the NSSDC and to prevent any further erosion by securing broadly based institutional support of baseline activity.
2. Establish an overall data-management plan adjusted to NASA's goals and missions, which, if necessary, would allow for a cost-effective expansion of the NSSDC activities by incorporating the latest techniques to minimize the human interaction and maximize the computer-compatible stages of data handling and transfer.
3. Maintain an experienced and stable NSSDC staff that, in addition to attending to the normally required tasks, is able to implement the needs of the communities of data collectors and data users and is currently informed on future space missions and related data needs.
4. Establish a working group on NSSDC operation that would act as the liaison with user groups and the data-collecting community.
5. Consider the establishment of a facility at the NSSDC to provide visiting scientists with the opportunity for *in situ* work with the data and computerized display systems in those disciplines that require the use of large amounts of data in digital form.

INTRODUCTION

The acquisition of information on some designated region of space, celestial body, system, object, or process is a fundamental goal of many space missions. To be useful to science or technology, the acquired information must pass through different stages of processing and storage, the last of which is synthesis and conversion into logical statements that are intelligible to the community at large, such as the formulation of a scientific theory, the construction of an environmental model, or the description of a planetary landscape. This reduction and conversion process usually incorporates knowledge and information from other sources or disciplines that may not be unique, may change, or may be revised over time, although the original data base acquired during the given mission remains invariant. If this data base becomes unretrievable, is obliterated, or

is discarded, a fundamental goal of that mission will have been defeated. Hundreds of years of scientific research have demonstrated that, especially in the natural sciences, reinterpretation of earlier data is as common and necessary a practice as is the acquisition of new data. Therefore, a most valuable and enduring asset of a space mission is the information contained in the acquired data base--whether these data have in the meantime been partly analyzed and used in conjunction with information from other sources--which can be made available to future investigators who will analyze and utilize the data sets in ways not yet imagined.

The immense volume of space data accumulated from past missions, currently being acquired, and expected in the future is a national treasure that ought to receive as much attention as the missions themselves. After the initial proprietary stage for data acquired by individual experimenters or by teams during a given space mission, it is necessary that these data be deposited in a national data facility where they are available to the scientific community. It is essential that the flow of data into such a facility proceed according to contractual regulations, that data quality be systematically checked, that storage be safe, and that retrieval be adequate, expedient, and affordable. If some editing, condensation, or other derivative process before archiving is advisable for practical reasons, it must be ensured that the associated information loss has a minimum impact on the general value of the data to the user community. Information on content, format, and addresses of stored data must be readily available to make retrieval procedures independent of the availability of the original experimenter. Finally, as higher-generation data storage, retrieval, and processing systems become available, it must be ensured that the space data facility possesses the ability to incorporate such new systems on a short time scale. To fulfill these conditions, adequate financial support is required--a small investment, taking into account the original costs of the missions that provided the data--and it is necessary to have an able and stable staff of scientists and data-systems experts, with a strong link to the data-collecting community and a short response time to satisfy the demands of users.

ROLE OF THE NATIONAL SPACE SCIENCE DATA CENTER

The NSSDC, established by NASA in 1968 at the Goddard Space Flight Center, plays an extremely significant role in

space-data archiving and distribution activities. A major charge to NASA when it was created was to "provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."* Data archiving at the NSSDC ensures (a) that the data currently being obtained are treated in an orderly manner by the principal investigator and are deposited at a national center; (b) that other qualified investigators can gain access to these data (this is particularly important, for the new mode of research requires multiple satellite measurements, and successful missions such as Viking generate much excitement in the educational and scientific community); (c) that future investigators can obtain access to the data sets to analyze them in ways not yet imagined. Data distribution lies at the heart of the basic charge to NASA in its enabling legislation. Data distribution from a national center such as NSSDC avoids costly duplication of materials and manpower at the institution of each potential investigator. The attachment of WDC-A for Rockets and Satellites to the NSSDC represents a U.S. commitment to the international space-science community that allows the sharing of an important part of the sceintific information obtained in our space missions with other scientists of the world.

Because of the tremendous amount of constriction that normally takes place in the data-flow process, it is difficult to demonstrate quantitatively the usefulness of the NSSDC and measure its cost-effectiveness on the basis of simple statistics relating the number of user requests to the data input flow. Rather, the value of the NSSDC should be appraised in the whole context of long-term development and achievements of space research.

The significant publications of the NSSDC, many issued periodically, such as *Data Catalog of Satellite Experiments*, *Data Users' Notes*, *Data Announcement Bulletins*, and *Satellite Situation Center Reports*, are of extreme importance in ensuring knowledge of data availability and dissemination. It must be pointed out that the cost of the publication services of NSSDC is trivial compared with the total operational budget. The NSSDC also carries out a series of collateral activities that are of fundamental value to the space-science community but that, for practical reasons, could not be carried out cost-effectively elsewhere.

*National Aeronautics and Space Act of 1958; Public Law 85-568; Section 203(a)(3), Functions of the Administration.

For example, the NSSDC develops and updates quantitative models of the near-earth space radiation and distributes them to a large group of users in NASA, DOD, NOAA, and many U.S. industrial organizations concerned with spacecraft systems, as well as to foreign governmental and industrial space scientists and engineers. The models of the space radiation environment are used during the design, development, and operational phases of all U.S. spacecraft systems as the definitive, authoritative design criteria. In this role the NSSDC offers a unique service, and the radiation models provided and updated by NSSDC should be viewed as a national resource. Development and updating of these radiation models rest on the availability at NSSDC of the massive data base on radiation-belt phenomenology necessary for the construction of accurate models. Indeed, the construction of environmental models can be viewed as the ultimate use of archived data.

We anticipate that the expansion of national capabilities in space, the development of the space transportation system, and the resumption of frequent manned spaceflight will require additional emphasis on model development and additional effort by NSSDC. At least the maintenance, if not the expansion, of the role of NSSDC as the national center for the development and distribution of quantitative models of the space environment is essential.

Another service that the NSSDC provides to the international community is the operation of the Satellite Situation Center, established for the IMS as a new approach to coordinating extremely complex combinations of multidisciplinary observations using different techniques. The main function of the Satellite Situation Center is to identify and recommend intervals of special interest on the basis of predicted configurations of satellites so that observations from the satellites, rockets, balloons, and ground stations can be carefully timed to maximize scientific gains. As a by-product of this activity, the Satellite Situation Center produces plots of predicted orbits of the major IMS-related satellites. Thus, beginning in 1975, the Satellite Situation Center has published reports providing plots of predicted orbits of several satellites and information on their orbital positions relative to various boundaries and regions of the magnetosphere. The Satellite Situation Center also compiles and updates a summary list of the IMS-related satellite experiments to keep the IMS community informed of the developments in the satellite program. In addition, the computer facility of the Center has been used to create the IMS file, which is a

computer-sensible information source concerning IMS participants and their projects.

The practical usefulness of the Satellite Situation Center was demonstrated recently when the ESA IMS satellite GEOS failed to achieve the planned geostationary orbit. The personnel of the Satellite Situation Center worked around the clock for several days to recommend an orbit that maximized the possibility of achieving part of the original goals of the mission while at the same time minimizing radiation exposure, maximizing tracking coverage, and opening a new possibility of using the second GEOS apogee to coordinate these observations with the extensive IMS ground-based network in Alaska and Canada.

RECOMMENDATIONS

The NSSDC has had to curtail some of its services as a result of a reduction in work force. However, even with this reduction, the NSSDC would be able to perform a valuable service to the scientific community if NASA, on both policy and the program levels, were to re-emphasize the importance of data management. The usefulness of the NSSDC to the user community is documented in replies to a questionnaire distributed by the General Accounting Office in which about 90 percent of the 392 users who replied indicated that the NSSDC serves a useful scientific purpose. The need for the NSSDC is borne out by this survey; if it did not exist, it would have to be created. Finally, it should be noted that the cost of operating the NSSDC is trivial when compared with the cost of the missions that provided the data.

NASA Policy Directive NPD 8030.3 of January 7, 1967, section 6.b states: "Each Program Director within Headquarters Program Offices is responsible for managing the data reduction, prime analysis and delivery of reduced data records to the NSSDC from space science flight experiments for which he has program management responsibility." If the Program Directors implemented this directive fully, the NSSDC might possibly provide the other services for which they are responsible.

To improve the effectiveness of the NSSDC and to prevent any further deterioration in its services, we recommend that NASA

1. Take immediate action to stabilize the present situation of the NSSDC and prevent any further erosion by

securing support from NASA Headquarters Program offices of a baseline activity that includes (a) continued handling of data collection, storage, and distribution; (b) continued work on radiation environment modeling; (c) continued operation of the Satellite Situation Center; (d) continued publication of the present series of reports, catalogues, and bulletins; (e) continued commitment to operate WDC-A for Rockets and Satellites.

2. Establish an overall NASA data-management plan encompassing all programs, which would ensure stability and, if necessary, allow for a cost-effective expansion of the NSSDC activities, adjusted to NASA's goals and mission. An objective of this plan should be to achieve maximum automation of the data-handling and -transfer operations by incorporating the latest techniques that minimize the human interface and maximize the computer-compatible stages of data handling wherever possible and advisable.

3. Maintain an experienced and stable staff that, in addition to attending to the normally required tasks, is able (a) to implement the needs of the communities of data collectors and data users; (b) to keep currently informed on future space missions and related data needs; (c) to carry out all special programming, data quality check, research, and publication operations.

4. Establish an NSSDC operations working group that would (a) act as the required liaison with user groups and the data-collecting community; (b) help establish rules and procedures for quality control; (c) help make decisions on, or set priorities for, data formatting, thinning, abstracting, discarding, and the like; (d) make recommendations regarding policies on the sale and exchange of data.

5. As part of a future development of the NSSDC, establish a facility to provide visiting scientists with the opportunity for *in situ* work with the data and computerized display systems in those disciplines that require the use of large amounts of data in digital form.

Appendix H
Abbreviations and Acronyms

ADP	Automatic Data Processing
AIDJEX	Arctic Ice Dynamics Joint Experiment
AFOS	Automation of Field Observations
AWS	Air Weather Service
BOMEX	Barbados Oceanographic and Meteorological Experiment
CDIDC	Committee on Data Interchange and Data Centers
CEDDA	Center for Experimental Design and Data Analysis
CRREL	Cold Regions Research and Engineering Laboratory, U.S. Army
DARPA	Defense Advance Research Projects Agency
DBMS	Data Base Management Systems
DMSP	Defense Meteorological Satellite Program
DNP	Declared National Programs
DOD	Department of Defense
DOE	Department of Energy
DSDP	Deep Sea Drilling Project
EDIS	Environmental Data and Information Service, NOAA
EDR	Epicenters Data Report
ENDEX	Environmental Data Index
EPA	Environmental Protection Agency
ESA	European Space Agency
ESSC	Environmental Studies Service Centers
FGGE	First GARP Global Experiment
GARP	Global Atmospheric Research Program
GATE	GARP Atlantic Tropical Experiment
GEOS	Geodetic Satellites
GOES	Geostationary Operational Environmental Satellite
GRB	Geophysics Research Board
IASPEI	International Association of Seismology and Physics of the Earth's Interior
ICSU	International Council of Scientific Unions

IFYGL	International Field Year for the Great Lakes
IGB	International Gravimetric Bureau
IGY	International Geophysical Year
IMS	International Magnetospheric Study
INSTAAR	Institute of Arctic and Alpine Research
IPOD	International Phase of Ocean Drilling
ISC	International Seismological Center
ISP	Initial State Parameters
LASA	Large Aperture Seismic Array
LRSM	Long Range Seismic Measurements
MAP	Middle Atmospheric Program
MGD	Marine Geophysical Data
NAS	National Academy of Sciences
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCC	National Climatic Center, Asheville, North Carolina
NDC	National Data Center
NESS	National Environmental Satellite Service
NGSDC	National Geophysical and Solar-Terrestrial Data Center
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center, Washington, D.C.
NORPAX	North Pacific Experiment
NORSAR	Norwegian Seismic Array
NRC	National Research Council
NSF	National Science Foundation
NSSDC	National Space Sciences Data Center, Goddard Space Flight Center, Greenbelt, Maryland
NURE	National Uranium Resource Evaluation
OCSEAP	Outer Continental Shelf Environmental Assessment Program
POLYMODE	Expanded Mid-Ocean Dynamics Experiment
PRB	Polar Research Board
RCIC	Regional Coastal Information Centers
SCAR	Scientific Committee on Anarctic Research
SDC	Systems Development Corporation
SEL	Space Environment Laboratory
SGD	Solar Geophysical Data
SOON	Solar Optical Observatory Network
STP	Solar-Terrestrial Physics
UAG	Upper Atmosphere Geophysics
UNESCO	United Nations Educational, Scientific and Cultural Organization
USGS	U.S. Geological Survey

VELA	A program for the detection and identification of underground nuclear detonations
WDC	World Data Centers
WDC-A	World Data Centers operated by the United States
WDC-B	World Data Centers operated by the Soviet Union
WDC-C	World Data Centers located in other countries
WWSSN	Worldwide Standardized Seismograph Network