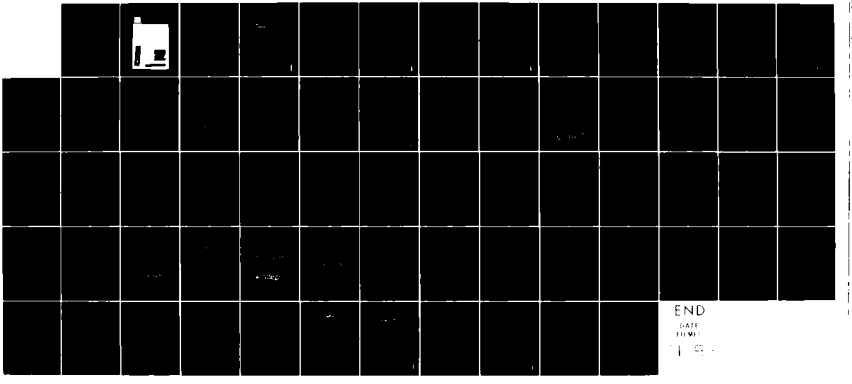
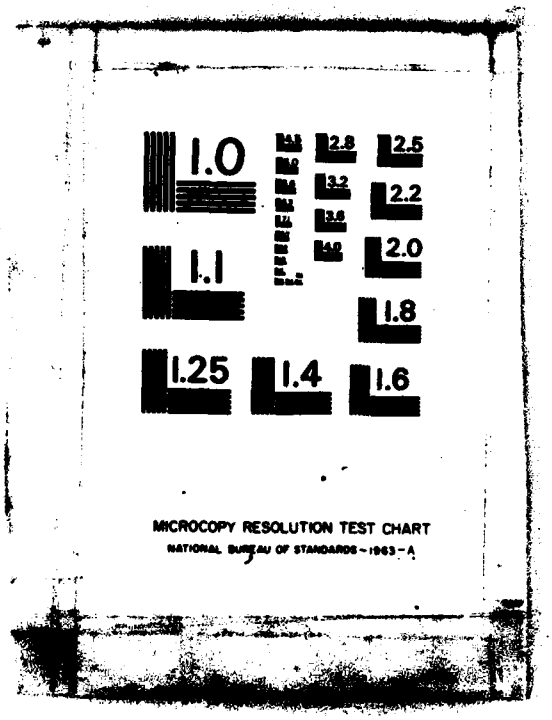
AD A134 077 EFFECTIVE USE OF EARTHQUAKE DATA(U) NATIONAL RESEARCH 1/1
COUNCIL WASHINGTON D C COMMITTEE ON SEISMOLOGY
T V MCEVILLY ET AL. 1983 N00014-82-G-0016

UNCLASSIFIED F/G 8/11 NL



END


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

Effective Use of Earthquake Data

NATIONAL RESEARCH COUNCIL

COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND RESOURCES

2101 Constitution Avenue Washington, D.C. 20418

BOARD ON EARTH SCIENCES
COMMITTEE ON SEISMOLOGY

(202) 334-2160

October 24, 1983

Defense Technical Information
Center
Office of Naval Research
Building 5, Cameron Station
Alexandria, VA 22314

Dear Sir:

The report Effective Uae of Earthquake Data, two copies of which are enclosed, was prepared under Grant No. N00014-83-G-0016 in partial support of the Committee on Seismology.

Sincerely,

Joseph W. Berg, Jr.
Executive Secretary

Enclosures

Effective Use of Earthquake Data

Panel on Data Problems in Seismology
Committee on Seismology
Board on Earth Sciences
Commission on Physical Sciences, Mathematics, and Resources
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1983

NOTICE: The project that is the subject of this 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.

The National Research Council was established by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and of advising the federal government. The Council operates in accordance with general policies determined by the Academy under the authority of its congressional charter of 1863, which establishes the Academy as a private, nonprofit, self-governing membership corporation. The Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in the conduct of their services to the government, the public, and the scientific and engineering communities. It is administered jointly by both Academies and the Institute of Medicine. The National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively, under the charter of the National Academy of Sciences.

Available from
Board on Earth Sciences
National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

COMMITTEE ON SEISMOLOGY

THOMAS V. McEVILLY, University of California, Berkeley, Chairman
C. ALLIN CORNELL, Stanford University
RICARDO DOBRY, Rensselaer Polytechnic Institute
ROBERT B. HERRMANN, St. Louis University
HIROO KANAMORI, California Institute of Technology
FRANKLYN K. LEVIN, Exxon Production Research Company
PAUL W. POMEROY, Rondout Associates
PAUL G. RICHARDS, Lamont-Doherty Geological Observatory, Palisades
DAVID W. SIMPSON, Lamont-Doherty Geological Observatory, Palisades
ROBERT B. SMITH, University of Utah
ROBERT E. WALLACE, U.S. Geological Survey, Menlo Park

Liaison Members

LEON L. BERATAN, U.S. Nuclear Regulatory Commission
WILLIAM J. BEST, Air Force Office of Scientific Research
MICHAEL A. CHINNERY, National Oceanic and Atmospheric Administration
JOHN R. FILSON, U.S. Geological Survey
EDWARD A. FLINN, National Aeronautics and Space Administration
JOHN G. HERCOCK, Office of Naval Research
LEONARD E. JOHNSON, National Science Foundation
GEORGE A. KOLSTAD, U.S. Department of Energy
PAUL F. KRUMPE, Agency for International Development
JAMES F. LANDER, National Oceanic and Atmospheric Administration
JAMES R. MACDONALD, Office of Naval Research
UGO MORELLI, Federal Emergency Management Agency

CARL F. ROMNEY, Defense Advanced Research Projects Agency
EDWARD SCHREIBER, U.S. Department of Energy
JOSEPH W. SIRY, National Aeronautics and Space Administration
K. THIRUMALAI, National Science Foundation

Staff

JOSEPH W. BERG, JR.
ROY E. HANSON

PANEL ON DATA PROBLEMS IN SEISMOLOGY

FREEMAN GILBERT, University of California, San Diego, Chairman (until February 1983)
SHELTON S. ALEXANDER, Pennsylvania State University, Chairman (after February 1983)
MILO M. BACKUS, University of Texas, Austin
RAYMOND P. BULAND, U.S. Geological Survey, Denver
RICHARD T. LACOSS, Massachusetts Institute of Technology
KENNETH L. LARNER, Western Geophysical Company of America
WILLIAM H. K. LEE, U.S. Geological Survey, Menlo Park
KAREN C. McNALLY, University of California, Santa Cruz
DAVID W. SIMPSON, Lamont-Doherty Geological Observatory, Palisades
M. NAFI TOKSOZ, Massachusetts Institute of Technology

Liaison Members

LEON L. BERATAN, U.S. Nuclear Regulatory Commission
WILLIAM J. BEST, Air Force Office of Scientific Research
FRANKLIN W. BURCH, National Archives and Record Service
E. R. HEDDAHL, U.S. Geological Survey
JOHN R. FILSON, U.S. Geological Survey
PAUL R. FISHER, U.S. Army Corps of Engineers
EDWARD A. FLINN, National Aeronautics and Space Administration
LEONARD E. JOHNSON, National Science Foundation
ANN U. KERR, Defense Advanced Research Projects Agency
JAMES F. LANDER, National Oceanic and Atmospheric Administration
RICHARD A. MARTIN, U.S. Bureau of Reclamation
JAMES M. McDONALD, Office of Naval Research

WILLIAM E. STRANGE, National Oceanic and Atmospheric Administration
K. THIRUMALAI, National Science Foundation

Staff

JOSEPH W. BERG, JR.
ROY E. HANSON

BOARD ON EARTH SCIENCES

WILLIAM R. DICKINSON, University of Arizona, Chairman
SAMUEL S. ADAMS, Adams and Associates, Boulder
LLOYD S. CLUFF, Woodward-Clyde Consultants, San Francisco
WALTER R. ECKELMANN, Sohio Petroleum Corporation, Houston
MICHEL T. HALBOUTY, The Halbouty Center, Houston
WILLIAM W. HAY, University of Colorado
MELVIN J. HILL, Gulf Oil Corporation, Houston
CARROLL ANN HODGES, U.S. Geological Survey, Menlo Park
WILLIAM C. LUTH, Sandia National Laboratories, ALbuquerque
CHARLES J. MANKIN, Oklahoma Geological Survey, Norman
V. RAMA MURTHY, University of Minnesota
JACK E. OLIVER, Cornell University
STEPHEN C. PORTER, University of Washington
J. WILLIAM SCHOPF, University of California, Los Angeles
E-AN ZEN, U.S. Geological Survey, Reston

Liaison Members

LEON L. BERATAN, U.S. Nuclear Regulatory Commission
ROBIN BRETT, National Science Foundation (until September 30, 1982)
PHILIP COHEN, U.S. Geological Survey
KENNETH DAUGHERTY, Defense Mapping Agency
PAUL R. FISHER, U.S. Army Corps of Engineers
BRUCE B. HANSHAW, U.S. Geological Survey
JAMES F. HAYS, National Science Foundation (from October 1, 1982)
JOHN G. HEACOCK, Office of Naval Research
LYNN HOOVER, U.S. Geological Survey
GEORGE A. ROLSTAD, U.S. Department of Energy
JOHN F. LANCE, National Science Foundation
DALLAS L. PECK, U.S. Geological Survey

COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND RESOURCES

PREFACE

Seismic data from earthquakes, especially those recorded digitally, have been accumulating rapidly in recent years. At present, diverse types of high-quality data are being generated under a wide range of programs supported by agencies such as the U.S. Geological Survey (USGS), the National Science Foundation (NSF), the Departments of Defense and Energy (DOD, DOE), the U.S. Nuclear Regulatory Commission (NRC), and other state and federal government and private institutions. The recent report of the National Research Council's Committee on Seismology, Seismographic Networks: Problems and Outlook for the 1980s (National Academy Press, Washington, D.C., 1983) provides insight into current data-acquisition aspects of earthquake seismology and difficulties confronting those operations and gives recommendations on ways to help eliminate problems in data collection.

It is clear that with the development of high dynamic range, broadband digital seismic systems, the distinction between strong-motion recording and sensitive high-gain seismic recording is disappearing. This means that earthquake engineers and seismologists soon can share a common seismic data base for their respective applications. Therefore, the discussion in this report should be taken to apply to all types of earthquake ground-motion data.

Digital data are opening exciting new areas of research and applications that until now have not been possible even with the best analog data. To realize these potential scientific breakthroughs in seismology fully, these data must be effectively disseminated to a wide user community concerned with both basic research and applications of seismic data. This report is addressed primarily to the problems of data management in

seismology, because shortcomings in the present data-handling and -distribution practices constitute the greatest barrier to effective use of earthquake data of all types.

The Panel on Data Problems in Seismology was established by the Committee on Seismology to review seismic data problems and make recommendations for the organization, selection, storage, and distribution of the data. It was specified that the study should include both existing and anticipated analog and digital data from global, national, regional, and local seismic networks and strong-motion data. Further, the panel was asked to consider procedures and policies to ensure the availability, timely distribution, and analysis of these data. (The charge to the panel is given in Appendix A.)

The results of the panel's findings presented in this report focus on seismological data from earthquakes. The general discussion and conclusions should be relevant for several years, but details of the report reflect the current situation in a rapidly changing era of data collection and distribution; of advancing capabilities and availability of computers; of varying data storage capacities of present systems; of increasing need for computers for solving complex analytical problems; of increasing availability and use of digital as compared with analog data; and of perpetual uncertainties in other factors, such as the yearly funding levels of federal agencies.

The data problems in seismology are of such key importance for achieving potential scientific advances and so changeable with time that continued vigilance will be needed to ensure that new developments in technology are implemented in a timely manner, enabling United States scientists and engineers to stay at the forefront of modern seismology.

The support of the Committee on Seismology by the following federal agencies is appreciated: National Oceanic and Atmospheric Administration; National Science Foundation, Division of Earth Sciences (AABO); U.S. Geological Survey; Office of Naval Research; Air Force Office of Scientific Research; National Science Foundation, Division of Civil and Environmental Engineering (ENG); U.S. Nuclear Regulatory Commission; Defense Advanced Research Projects Agency; U.S. Department of Energy; and Federal Emergency Management Agency.

Shelton S. Alexander, Chairman
Panel on Data Problems in Seismology

CONTENTS:

EXECUTIVE SUMMARY

After considering a wide variety of data problems in seismology, the Panel has identified as the primary challenge in the immediate future the development of a coordinated national effort in the collection, storage, and dissemination of digital earthquake data to assure that our most advanced technology is used effectively in seismological research and engineering applications. Indeed many of the most important and challenging seismological studies of today require both digital data from state-of-the-art instruments and computer facilities capable of analyzing large data sets and modeling the processes that explain the observations. Providing seismologists* with easy (remote) access to desired event information and waveform data (including real-time recordings from global or national networks) at a National Center for Seismological Studies should increase scientific productivity significantly without appreciably increasing overall data management and dissemination costs.

We foresee major advances in seismology resulting from the increasing use of digital data. Many of these advances will simply follow from refinements in doing better what has already been done, e.g., improved velocity models, more accurate hypocenter locations, more complete earthquake catalogs, and more efficient searching and sorting of data bases. There are many areas of research, however, where digital data are opening

*In this report "seismologist" refers interchangeably to both scientists and engineers and "seismology" includes the full range of science and engineering studies, including strong motion.

possibilities that have until now been beyond the limitations of the best analog data; among these are detailed structure and heterogeneity of the Earth's interior from high-resolution eigenfrequency spectra of free oscillations; frequency-dependent attenuation (Q) structure of the Earth; detailed dynamic models of the earthquake source from broadband waveform studies; real-time or near-real-time analysis of unclipped, broadband records from very large earthquakes recorded both teleseismically and in the near field; and the routine determination of moment-tensor representations of earthquake sources.

If no steps are taken to address the problems identified in Chapter 3, the situation is likely to develop that only the operators of seismograph stations will have ready access to the data they record. Effective data dissemination to the seismological user community currently exists only for the World Wide Standardized Seismograph Network (WWSSN) analog (film) data, and that service is threatened by escalating costs. There is no comprehensive directory of information that describes what data sets exist, who has them, and how data can be obtained. There is no comprehensive national earthquake catalog, and additional phase data are not generally available at all. Information is fragmented and must be obtained from many different organizations. Distribution of digital data is limited because of current policies and procedures for dissemination, combined with significant direct or indirect costs to users.

These problems, coupled with the Defense Advanced Research Projects Agency's (DARPA) recent implementation of a modern computational capability for seismological data storage, retrieval, and dissemination to their research contractors, have led the Panel to conclude that <u>establishing a national seismological data base at a National Center for Seismological Studies is both desirable and feasible at this time.</u> A National Center that meets the functional requirements elaborated in Appendix F would overcome key data-management and -dissemination problems and enhance significantly the availability of high-quality data sets and their effective use by the entire seismological community.

Effective use of this national data base will require the following:

(a) Upgrading the present data-management systems to provide adequately for user needs.

(b) Development of a comprehensive directory of available data, data sources, quality, and formats, together with a computerized search capability implemented at a National Center and regional centers.

(c) Development of a comprehensive catalog of national and global hypocenters, including associated phase data for at least the larger events, together with a computerized search capability implemented at a National Center.

(d) Standardization of data formats for the purpose of data exchange.

(e) Establishment of a policy for data retention and archiving.

These subjects are discussed in the text and appropriate recommendations are given.

The principal recommendations of the Panel are as follows:

1. A National Center for Seismological Studies (NCSS) should be established to ensure the effective use of global, national, selected regional, and strong-motion earthquake data. New capabilities afforded by this center would include real-time or near-real-time access to global and national network recordings and remote terminal access to a well-organized seismological data base. Suggestions for the funding and management of an NCSS are given in this report.

2. In calendar year 1983 the NRC Committee on Seismology should initiate steps to convene a meeting of representatives of funding agencies, user groups, and the university research community to discuss implementation and funding of a National Center for Seismological Studies using DARPA's Center for Seismic Studies (CSS) as its prototype.

Other important recommendations are as follows:

3. The present NOAA/National Geophysical Data Center's analog (film) facility for seismic data should be maintained, and the costs of operating should be shared as described in the text; alternative techniques for disseminating analog data should be considered by NOAA to keep user costs at a minimum.

4. Alternative types of Global Digital Seismograph Network (GDSN) digital data subsets should be made available routinely to users in addition to the standard 24-hour network-day tapes.

5. The Regional Data Centers should continue to provide Global Digital Seismic Network data to visiting researchers.

6. Data from regional and local networks should continue to be collected, analyzed, and archived by individual network operators and made available in a standard format to other users directly or via the National Center or other appropriate regional data centers.

7. Strong-motion digital data should be archived and distributed by the National Center.

8. At least a subset of the Air Force Technical Application Center's Global Surface System (GSS) data for the larger earthquakes should be made available for general use.

*9. Selected subsets of special data collections currently at Teledyne-Geotech's Seismic Data Analysis Center (SDAC) in Alexandria, Virginia, should be kept indefinitely as part of an active, accessible national seismological data base; these subsets should include digital event tapes of all m_b 5.3 and greater, teleseismic event recordings and all m_b 3.5 and above, regional and local event recordings.

10. A centralized data directory should be developed and maintained as part of the national seismological data base accessible to users via computer terminals, and global hypocenter catalogs should be augmented with a comprehensive national catalog derived from local and regional networks and associated phase data be provided for the larger events. These catalogs should be available from a national facility that provides computer search capability to users.

11. Research granting and contracting agencies should plan allocation of approximately 10 percent of awarded monies for management and dissemination of data for studies involving the acquisition of new data or analysis of existing data.

12. The National Research Council's Committee on Seismology should organize a workshop for the purpose of establishing standard data exchange formats and standardization of event catalogs. For the present, the International Association of Seismology and Physics of the Earth's Interior (IASPEI) standards for international exchange of digital data should be used.

*This item needs immediate action.

13. The National Research Council's Committee on Seismology should assume the lead role in establishing overall policy for the long-term retention of seismological data currently being collected; it is essential that funding agencies, user organizations, and the university research community actively participate in establishing overall policy to ensure that useful data are not discarded.

14. Because of the rapid technological changes in computational and data-handling capabilities, the National Research Council's Committee on Seismology should review the status of data problems in seismology on a year-to-year basis to ensure effective use of earthquake data.

INTRODUCTION

Seismology, in common with most of the physical sciences, has entered the 1980s with increasing challenges and opportunities presented by advances in the technology for gathering and analyzing data. These developments have resulted in a rapid increase of the amount of data, a shift in data collection from primarily analog to digital recording, and the growing application of advanced computer technology, and they provide opportunities to increase significantly the scientific returns from seismic data of all types.

These developments pervade all aspects of seismology and are apparent in strong-motion studies; exploration activities; continental and marine reflection profiling; global, regional, and local network and array studies; and large-scale global studies of structure, attenuation, and earthquake source mechanisms. Each of these areas has its own set of specific data problems. This report is focused primarily on data problems of the global networks, national and regional networks, and strong-motion observations. Problems associated with seismic profiling or more specialized data-gathering activities in the other areas of seismology mentioned above have been deferred until the major problems of earthquake data management have been addressed.

The impact of the rapid increase in the amount of data, especially digital data, is being felt in many fields, seismology in particular. The increase in the amount of earthquake data is a consequence of the rapid expansion during the past decade of the number of seismographic observatories throughout the world and the growing use of digital data-acquisition methods. Table B.1 in Appendix B gives an indication of the amount of data produced by typical global and regional networks and

the amount of storage capacity required to save these data.

On a global basis the number of countries and organizations currently engaged in digital seismic monitoring is already impressive, and others are moving rapidly toward digital recording. A joint NOAA/USGS publication, Directory of World Digital Seismic Stations (1982), describes many of the existing operations. Several additional countries are installing digital systems, and in others they are planned or initiated.

One of the major advances in observational seismology in the predigital era was the development and installation of the World-Wide Standardized Seismograph Network (WWSSN) in the early 1960s. The network was installed with DARPA support under the VELA UNIFORM program for fundamental studies, including those related to the detection and discrimination of nuclear explosion. This network succeeded in its primary goal of upgrading observational seismology. For example, modern theories of plate tectonics are rooted in global patterns in seismicity and earthquake focal mechanisms; the observation of these systematic patterns can be traced directly to WWSSN data. Global earthquake catalogs, used in a variety of applications from earthquake prediction to tectonic studies and seismic risk analysis, have relied heavily on WWSSN data. Our knowledge of the internal structure of the Earth has also been advanced significantly by studies using WWSSN data. Because of the widespread use of these data and the operation of stations by many university groups, the WWSSN has also played an important role in the education and training of seismologists, both in the United States and internationally.

The primary reasons for the outstanding success of the WWSSN are inherent in its title: it is worldwide, providing global coverage; it uses a standardized set of instruments, providing uniform responses at all stations; and it is a coordinated network, with a successful mechanism for data collection and distribution through the National Geophysical Data Center. Our challenge as we enter the era of digital seismology is to repeat these successes of the WWSSN, while taking advantage of the enhanced quality and resolution of digital data.

The sensors used in the WWSSN incorporated state-of-the-art seismometry of that time (circa 1960). The instrumentation selected followed a careful consideration of the available technology and anticipated uses of the data. The wisdom of the choices made is reflected in the

continued viability of these stations today. However, because of the limitations imposed by the photographic recording medium, the WWSSN observations are limited in both frequency and dynamic range. The WWSSN data have been most useful in those studies that require measurements of time and amplitude of particular phases, e.g., hypocenter location, magnitude determinations, fault-plane solutions, and velocity structure. They have been used considerably less in detailed waveform and spectral analysis because of the necessity that they first be converted to digital form. Some of the long-period records have been successfully digitized by hand (a tedious and time-consuming process, with limited spectral resolution) for more-detailed waveform and spectral analyses; the short-period data have seen more limited use in waveform and spectral studies, because of the difficulties in digitization from the compressed time scale of the photographic records.

Recent developments in sensor technology (especially feedback seismometers), coupled with stable, broadband amplifiers and digital recording and playback methods, now provide the technological capability to record and analyze ground motion with high fidelity over a broad frequency and amplitude range. In a frequency band spanning many orders of magnitude, from periods of thousands of seconds to frequencies of hundreds of hertz, there are no longer any technical limitations in detecting and recording ground motion ranging in amplitude from background noise at the quietest sites to the strong motions produced in the near field by large earthquakes. As discussed elsewhere in this report, the major problems now lie in determining what data will be collected, how much will be archived, and how data will be distributed to users (i.e., data-management policies and procedures) and funding of data-management and -distribution activities. While full advantage of these new technologies has not yet been taken in many areas of seismology, the emerging implementation of digital recording, especially in teleseismic and near-field studies, already points to the scientific advantages gained from the use of high-resolution data [see Global Earthquake Monitoring: Its Uses, Potentials, and Support Requirements (1977); Strong-Motion Earthquake Instrument Arrays (1978); U.S. Earthquake Observatories: Recommendations for a New National Network (1980); U.S. Strong-Motion Earthquake Instrumentation (1981); and Seismographic Networks: Problems and Outlook for the 1980s (1983)].

In this report we deal only peripherally with questions of instrumentation for primary data collection (seismograph systems). These instrumentation issues are discussed at length in the network reports listed above and (except for regional networks, which are discussed briefly below) excellent prototype seismographic systems with broadband digital recording are already available. We also have not emphasized the computer hardware to be used in accessing the data base, nor have we considered in detail the computer facilities required for extensive computational studies using the data. In these areas, technology is developing rapidly in response to demands from a wide variety of scientific and nonscientific users, and the data volumes anticipated in seismology do not appear to present any significant technological problems that will not be satisfied by the advancing state of the art. Here, as with seismographic instrumentation, our problems lie not in encouraging the development of new technologies but in ensuring that modern facilities are made available to the seismological community and used wisely. This will require adequate funding and good data-management policies and procedures.

To realize fully the potential scientific returns in seismology requires that high-quality data not only be gathered but also effectively disseminated to a wide user community concerned with both basic research and applications of seismic data. This report is addressed primarily to the problems of data management in seismology because of shortcomings in the present data-handling and -distribution practices.

3

CURRENT STATUS AND PROBLEMS

INTRODUCTION

It is convenient to characterize earthquake data sources by function into near-field or strong-motion recordings and far-field or network recordings. Network data may be further categorized by areal extent into local, regional, national, and global scales. Ideally, each data source is described by a station catalog listing the position and operating characteristics of each receiver as a function of time. Each data source produces raw data recorded in analog and/or digital format. Routinely, phase arrival times and amplitudes or event durations are derived from raw network data, and event hypocentral locations, origin times, and magnitudes are derived from these phase data.

In general, the success of efforts to provide station catalogs, event catalogs, event-associated phase data, and raw data to other users reflects both the maturity of a particular data source and the national perception over the past several decades of its social relevance. Thus, the distribution of strong-motion data is relatively well developed because of its relevance to lifesaving earthquake-resistant building design and because the data volume is small. Similarly, global network data distribution is better developed than that for regional or local networks because of its applicability to a wide range of seismological and tectonic scientific problems and to practical issues such as monitoring nuclear explosions and assessments of seismic risk over large areas. The smaller networks tend to be of more specialized interest and limited in lifetime, resulting in little or no distribution of data to secondary users.

The distribution of analog recordings is better developed than that for digital data for several reasons. The latter is based on new technology and have been less readily available for perusal; many data users have limited familiarity with the use of digital data and commonly lack facilities for exploiting digital analysis techniques in their research. The following discussion focuses primarily on the status and problems with global and national network data and strong-motion data.

DATA SOURCES

The Global Seismic Network (GSN) was developed by the Defense Advanced Research Projects Agency (DARPA) in a series of successive projects, and it is currently operated by the USGS. The oldest component, the WWSSN, had an intended size of 125 stations. About 105 stations were installed in the early 1960s, and 96 are still operating. These stations feature analog recording of three-component long-period (~15-sec peak period) and three-component short-period (~1-sec peak period) instruments. In the late 1960s and early 1970s, the GSN was augmented with 10 digitally recorded high-gain long-period (HGLP) instruments featuring three-component long-period data sampled at a rate of one sample per second (sps). During the mid-1970s 12 Seismic Research Observatories (SROs) were added. Borehole, force-balance instruments in the SROs feature digital recording of three-component long-period data sampled at 1 sps and vertical component short-period data at 20 sps. In the late 1970s, a short-period vertical component sensor was added to five of the HGLPs, which were upgraded to an SRO-compatible recording system and dubbed Abbreviated Seismic Research Observatories (ASROs). The other HGLP stations were closed. In the early 1980s, 14 of the WWSSN stations were upgraded to digital recording and renamed Digital World Wide Standardized Seismograph Network (DWWSSN) stations. The DWWSSNs feature three-component long- and intermediate- period channels sampled at 1 sps and 10 sps, respectively, and a vertical component short-period channel sampled at 20 sps. The SROs, ASROs, and DWWSSNs are collectively referred to as the Global Digital Seismograph Network (GDSN). All GDSN stations feature continuous recording of long-period channels and field event-triggering to record signals from short- and intermediate-period

channels. The GDSN is complemented by the International Deployment of Accelerometers (IDA) operated by the University of California at San Diego. The 17 force-balance gravimeters in this global network supply digitally recorded, vertical-component information at periods longer than 60 sec, sampled every 10 sec.

Several new networks are being added to the GDSN. The instruments feature borehole force-balance operation, real-time satellite telemetry, and three components in each of three period bands all continuously digitally recorded at rates of 1, 4, and 40 sps for long-, intermediate-, and short-period bands, respectively. A five-station North American network of such stations, known as the Regional Seismic Test Network (RSTN) operated by Sandia National Laboratories is already in operation. A subset of these continuous data (with laboratory event-triggered short- and intermediate-period channels) is currently being included on the network-day tapes (NDTs). Another similar network, known as the Global Telemetered Seismic Network (GTSN), is now being planned by the USGS. It will consist of 17 stations in the southern portion of the western hemisphere. These 17 new stations will represent a significant increase in the GDSN data volume that will require substantial increases in the level of effort required for effective archiving and distribution. In turn, costs to users will be significantly increased if the smallest unit of data that can be obtained continues to be a network day (26 hours of data for all stations). Appendix E describes new global data collection activities of the Air Force Technical Applications Center (AFTAC) utilising similar instrumentation.

There is currently no national seismic network. However, by telemetering about 60 short-period vertical signals from existing observatories in the continental United States and Alaska to Golden, Colorado, the USGS has created what is de facto a rudimentary national network. The data produced are digitally recorded at 20 sps for use by the National Earthquake Information Service (NEIS) of the USGS. Because these are narrow-band and low-dynamic-range data, no effort has been made to distribute them to secondary users. However, the Panel believes there is a possibility of establishing selected high-dynamic-range, broadband regional network stations that together with existing GSN stations (including RSTN) would form a national network of high quality.

There are numerous regional and local networks throughout the United States. They are typically operated by university personnel and funded by either the USGS or the Nuclear Regulatory Commission (NRC). In aggregate, they comprise about 1,600 stations, of which about half are digitally recorded (event windows) at rates of 50-250 sps. For the most part, the instrumentation is vertical-component only, narrow band, and low dynamic range, although there are an increasing number of broadband, high-dynamic-range, three-component stations. These networks are used mainly to measure arrival times, signal durations, and amplitudes of microearthquakes for earthquake-hazard-reduction research. The chief products of these networks usually consist of an event catalog and associated seismicity maps.

Although some of the larger networks share common hardware and software, there are many data formats extant. There is little data exchange at present and then only on an informal basis. There is no comprehensive catalog of stations and events for these networks. Thus potential users cannot easily determine whether appropriate data even exist for their problems.

Irreplaceable data from many limited-duration station deployments have been gathered during the past 35 years and now are in danger of being lost forever unless immediate action is taken. Millions of dollars have been spent on collecting these special research data sets from networks such as the national network of temporary stations, which comprises the Long Range Seismic Monitoring (LRSM) network and permanent observatories for nuclear test ban monitoring that were in place in the late 1960s and the early 1970s; the HGLP network; the Seismic Data Collection System (SDCS); and from special-purpose research arrays such as the Large Aperature Seismic Array (LASA) and the Alaskan Long Period Array (ALPA). To the present, Teledyne-Geotech's Seismic Data Analysis Center in Alexandria, Virginia, has archived these analog and digital data, but the operation of that facility is scheduled to terminate on October 1, 1983. Large subsets from these data should be archived indefinitely because of their value for current and future seismologic studies. To do this effectively will require a policy for data retention for all earthquake seismic data, since none exists at present, and the assignment of organizational responsibility for the preparation the data subsets selected for archiving.

The National Strong-Motion program has two components: (a) the National Science Foundation is responsible for the research program, and (b) the USGS operates and manages the strong-motion networks, including data handling. An extensive body of strong-motion data has been collected under this program and organized into a data base that includes both analog and digital waveforms. A strong-motion instrumentation program is operated by the state of California, and selected data from this program should be included in a national data base. Some arrangement should be made for getting strong-motion data routinely on a global basis. The details of these strong-motion data collection systems are summarized in the reports entitled Strong-Motion Earthquake Instrument Arrays (1978) and U.S. Strong-Motion Earthquake Instrumentation (1981).

DATA ARCHIVING AND DISTRIBUTION

Since its beginning in 1961, more than 5 million original WWSSN analog records have been copied and 60 million copies supplied to users. Currently, there are several hundred requests per year for seismogram microfiche. Originally, the seismograms were filmed on specially designed 70-mm panoramic cameras at 8x reduction. In 1978 filming was changed to put 24 images (4 days of normal operation) on a single 105-mm microfiche at 32x reduction. Standing orders of the whole network have been purchased by eight institutions [Lamont-Doherty Geological Observatory; Institute of Geological Sciences, Edinburgh, U.K.; University of Tokyo; California Institute of Technology; Massachusetts Institute of Technology; USGS/Menlo Park; USGS/Golden; National Geophysical Data Center (NGDC)], and substantial parts of the network have been supplied to five institutions [University of Texas/Galveston; Cornell University; University of Otago (New Zealand); Los Alamos National Laboratory; USGS/Albuquerque].

The WWSSN network data are augmented by copies of the visible (analog) records from the ASRO (5) and SRO (12) networks, from the Canadian network on 35-mm film since 1966, and from the People's Republic of China since 1979. Large-magnitude or seismologically important earthquakes from several hundred additional stations, including those of the Union of Soviet Socialist Republics, are provided under the International Data

Exchange (IDE) agreement. Also, selected historical analog seismograms are being filmed under a joint USGS/NOAA project.

This analog data distribution system is operating primarily with contract labor and with about 8 weeks being required for the cycle from receipt of original records to supplying of copies to users. Fifty percent of the network data is generally available for distribution within 8 months after the recording interval. The archival film copy is made at NOAA expense, with the cost of each additional copy being borne by the user, which is a direct reflection of the policy of the U.S. Department of Commerce. Present costs to users have increased significantly for some services, which poses a barrier to continued wide use of these analog data.

NOAA has compiled station catalogs of GSN stations and historical stations (both U.S. and foreign) on hardcopy. The USGS maintains, in a computer data base, a current station catalog of U.S. and foreign stations that contribute derived phase data to the NEIS, but it is not currently available to secondary data users. The NEIS collects phase arrivals and amplitudes from some 2,000 contributing stations by telex and airmail letter. Some 500,000 of these data per year are culled, associated, and used to locate 5,000 to 6,000 earthquakes. Earthquake bulletins (hardcopy) are prepared and distributed by the USGS. Event lists are also merged into a NOAA catalog for distribution on both hardcopy and computer-compatible media. In addition, a catalog of historical U.S. earthquakes is being compiled by the USGS. Associated phase data are distributed only on hardcopy listings and even that distribution is limited. Both of these limitations constitute a major shortcoming of the present data-management procedures.

Since 1976, GDSN data have been collated into NDTs first in a raw binary format by DARPA and since 1980 by the USGS in a binary format with American National Standard Code for Information Interchange (ASCII) headers (including considerable information about station operating characteristics). Raw station tapes containing about 14 days of data and operator logs covering the same time period are collected from each station in the GDSN by the Albuquerque Seismological Laboratory (ASL) of the USGS. Station-channel headers incorporating the log information are generated for each day. Finally, all headers and station tape data for each day are included in NDTs, each station-channel set appearing in sequence

as a file on the tape. The NDTs are made available to users approximately 12 weeks after the data are initially recorded. This operation is also plagued by rapid cost increases due to increases in the cost of personnel and materials and because of the rapid growth in the size of the GDSN. Further, the GDSN is already in imminent danger of saturating available computer resources, necessitating a considerable investment in new computer hardware. To date, all costs of generating the NDTs have been borne by the USGS as a service to the user community. The NDTs are currently archived by the USGS and along with IDA data tapes are archived by NGDC for distribution of copies to users.

There are a number of difficulties with the GSN data collection, archiving, and distribution. Although the analog data handling and distribution is smooth and effective, this operation is necessarily labor-intensive, resulting in high costs that have chronically threatened its continued existence. Of the catalog data, only the event catalog may be computer searched at present, and each user must supply his own search software or request a search by NGDC, which necessarily involves some delay in receiving the desired event information. The digital data distribution problem is the most severe of all. These data are available only in increments of one day (network day tapes) at $100 per tape. Coupled with alternative means of acquiring NDTs this has represented a significant barrier for users (especially university researchers), resulting in low demand at NGDC.

To illustrate the nature of this barrier, a typical university study of the nature of earthquake source mechanisms that characterize a seismic area of interest reasonably would involve at least 20 earthquakes, resulting in a cost of $2,000 for 20 network-day tapes. The researcher must then preprocess these 20 full tapes to extract the event windows of interest; thus substantial additional costs are likely to be incurred before any data analysis is done. In practice, users have acquired GDSN data informally, essentially at no cost, from DARPA and the USGS. In an effort to alleviate this situation and encourage the use of GDSN data, the USGS recently established three regional centers that have acquired complete sets of NDTs (at media cost). It is understood that visitors may use regional center facilities to copy NDTs on the same media cost basis. However, this has only partly alleviated the cost problem, because travel costs are incurred and the number of tapes that must be

obtained and preprocessed by the user remains the same. These cost-induced barriers will continue to ensure excessively high per-user cost of acquiring and utilizing GSN data, unless alternative means of distribution are implemented.

These distribution problems raise serious questions of how effective distribution at affordable costs can be achieved and which organization should ultimately be responsible for distributing GDSN data. This whole situation will be further exacerbated by the impending addition to the GDSN of several new networks representing a new generation of technology, as discussed earlier. Very soon a network day will not fit on a single 1600 bpi 9-track magnetic tape.

The national strong-motion program is funded and managed by the USGS from Menlo Park, California. Event-recording catalogs have been compiled and made available to users via on-site and remote interactive access to a relational data-base manager. Original analog film recordings from the western hemisphere are archived in Menlo Park, California, and derived digitized waveforms are available from NGDC. However, because of a constant state of flux in the USGS computing environment and a shortage of data-handling resources, it has been difficult to maintain accessibility to the strong-motion data base to meet the needs of the engineering community or even to keep it current. To ensure its viability in the future, this invaluable and already highly organized data resource needs to be housed and managed in some more stable, central, national earthquake data archiving and distribution facility where it can be maintained in a readily accessible form for users.

A PROTOTYPE NATIONAL CENTER FOR SEISMOLOGICAL STUDIES

An item of immediate importance, with long-term implications for the future of seismology, and specifically for overcoming many of the data problems discussed above, is the current development by DARPA of the Center for Seismic Studies (CSS), recently established in Roselyn, Virginia. Appendix D describes this facility more fully. When completed in 1984, the Center will include state-of-the-art computational and data-management capabilities designed for seismological applications and research, a small resident research staff, and provision for visiting scientists working on DARPA research programs in test ban treaty verification.

To support DARPA objectives, data are being collected from a variety of sources through several types of data links. These range from RSTN data received via a satellite Earth station located at the Center to alphanumeric data received via the World Meteorological Organization's (WMO) Global Telecommunications System. GDSN and other data are received on tape. At various times special experiments will be conducted as part of the United Nations Committee on Disarmament Group of Scientific Experts activities and will include national and international data exchange.

As currently conceived, this facility is developing many of the capabilities desired for general seismological data management and distribution. The seismic data base will consist of basic event information (such as station phase and amplitude readings, hypocentral locations, magnitude, and other source characteristics) and recorded signals, consisting of event waveforms, some available as the events occur, but others, such as GDSN day tapes, delayed by weeks to months. A variety of analysis software is also being developed and implemented to provide routine user computational services.

In addition, remote user access to the Center is being developed by DARPA in the form of relatively inexpensive ($30,000 to $50,000 each) Remote Seismic Terminal (RST) work stations. These terminals will be capable of storing and displaying a significant amount of event information and waveform data; powerful local processing functions can also be carried out on the RSTs. Designing interfaces with many other remote user systems should be reasonably straightforward as well.

The Center will be accessible via the ARPANET, Tymnet, dial-up, the RSTs, and, in a few cases, dedicated lines. Through these means, DARPA researchers will have access to both the data and computing resources. These resources will include two Digital Equipment Corporation (DEC) VAX 11/780 computers and three DEC PDP 11/44s linked by a high-speed (10-MHz) local network. Analysis may be carried out on various interactive graphics terminals. The Center has been designed to be as flexible and accessible to researchers as possible.

A small number of researchers have already begun using the CSS computing and data-handling resources on a trial basis to help to evaluate it and to recommend improvements to the system as development continues. Beginning in fiscal year 1984, the Center expects to support more extensive use of the facilities by DARPA-sponsored

researchers. Within at most a few years the operational responsibility for the Center is planned to be passed to some other organization. The Center would then be jointly funded by DARPA, the operating organization, and other user agencies.

The panel views the development of the CSS as a significant opportunity to benefit immediately from the rapid advances in computer technology and data-base management that have profoundly increased digital data storage, retrieval, and analysis capabilities during the past decade. Other DARPA-sponsored efforts in the past have led to major seismological advances; noteworthy examples are the WWSSN and film chip data distribution in the 1960s and the new GDSN global digital network with standard network-day tape data distribution. Experience with this advanced system at the Center will be exceedingly important to the development of intermediate- and long-term capabilities needed by the seismological community for research and applications in the digital era.

SUMMARY OF PROBLEMS

The above discussion of the current status of seismographic network data handling includes problems of operations and data management. Specifically the problems that now seriously limit or threaten to limit the effective use of earthquake seismic data are the following:

1. A comprehensive directory of information on earthquake data sources does not exist in hard-copy or computer-accessible form. Thus, for potential users it commonly requires a time-consuming, laborious effort just to determine what relevant data might exist, where they are archived, and how they can be acquired and in what forms or formats. This situation is exacerbated by rapid increases in data volume, changing station distribution, and a steady transition to digital recording.
2. There does not exist a comprehensive catalog of global and national hypocenters with associated phase data. Even the limited data that exist are mainly distributed in hard-copy form. Direct computer access to such data by users is severely limited or impossible at present.

3. The handling and distribution of Global Seismic Network analog data is well run by the NGDC personnel, but current Department of Commerce policy requires that copying and distribution of seismograms be charged to the users of the data, and these costs have already increased so greatly that the continuation of the present data services at affordable costs to users is seriously threatened.

4. Except for the GDSN network-day tapes, there are no standard formats for digital data exchange. Because of the variety of formats used for recording and archiving data among the past and current seismic data-collection operations, direct exchange of data is exceedingly cumbersome and difficult. There needs to be a standardization of national and international data-exchange formats for global, regional, and local networks.

5. Because the operation of smaller regional and local networks is focused on seismic problems of specialized interest, there is little or no distribution of data to secondary users and there is no standard exchange format. Because of limited funding, these network data commonly are not archived in a readily accessible form for distribution. There is also no readily available central directory of information on the events recorded or even on the station locations and instrumentation that make up these networks.

6. The extensive data base of high-quality strong-motion data is not adequately available to users as it would be if the strong-motion data were part of a national seismological data base housed in a stable, national earthquake data archiving and distribution facility.

7. Because of Department of Commerce cost-recovery policies and a low initial demand, network-day tapes of the GDSN have been priced by the NGDC at $100 per tape copy. Because the network-day tape is the smallest increment of data currently available, the user typically must obtain (and preprocess) many network-day tapes just to obtain the event data of interest; these event data usually occupy only a small fraction of the tapes that must be acquired. The high costs have resulted in network-day tapes being obtained by users, with less convenience but essentially at no cost, on user-supplied tapes from DARPA, the USGS, and at four regional centers. Organizational responsibility for distributing the GDSN data is fragmented and unsettled, and it needs to be determined now, under the condition that standard

types of digital data subsets must be readily available to users at low cost.

8. New digital data from stations currently being established will soon result in a significant increase in the volume of digital data to be handled. This will require a substantial increase in the level of effort for data handling and distribution. Facilities need to be augmented to manage these data so that they can be used effectively by government agencies and the scientific community in general. Planning is needed now to establish an effective data-handling system that can accommodate the growth in the volume of data recorded and provide desired data to users in unit amounts other than the present 26-hour network-day tape.

9. There is a need for a policy for data retention for all earthquake seismic data, with an organized program to prepare the selected data sets for long-term archiving. This problem is immediate and acute for the high-quality special data bases acquired by the LRSM, LASA, ALPA, NGLP, SDCS, and other network or array operations of the past 35 years that are currently archived at Teledyne-Geotech's Seismic Data Analysis Center in Alexandria, Virginia. This facility is scheduled to be closed at the beginning of fiscal year 1984. Immediate action is needed to assure that this invaluable data source is preserved.

10. The development by DARPA of the Center for Seismic Studies, a state-of-the-art data analysis and seismic data-base management facility, represents simultaneously a problem and an opportunity. The problem is that once fully developed and evaluated, the facility is to be turned over to some other, as yet undetermined, organization that will have to assume the management responsibility for its continued operation. Thus at present its fate is unknown. The opportunity is that this advanced facility, properly configured for a wider community of users, could serve as the prototype of a first-rate national center for seismological studies. Such a facility, properly managed and adequately funded, could overcome most of the current problems that the panel has identified and greatly enhance the use of available earthquake seismic data by both U.S. and international scientists.

4

RECOMMENDATIONS

If no steps are taken to address the problems identified in Chapter 3 the situation is likely to develop that only the operators of seismograph stations will have ready access to the data they record. Effective data dissemination to the seismological user community currently exists only for the Worldwide Standardized Seismograph Network (WWSSN) analog (film) data, and that service is threatened by escalating costs. There is no comprehensive directory of information that describes what data sets exist, who has them, and how data can be obtained. Existing global earthquake catalogs are difficult to computer search, and there is no comprehensive national earthquake catalog. Also, additional phase data are not generally available at all. What information exists is fragmented and must be obtained from many different organizations. Distribution of digital data is limited because of fragmented responsibility, current policies and procedures, and direct and indirect costs to users.

These problems, coupled with the Defense Advanced Research Projects Agency's (DARPA's) recent implementation of a modern computational capability for seismological data storage, retrieval, and dissemination to users (the DARPA Center for Seismic Studies) and for data analysis have led the panel to conclude that establishing a national seismological digital data base and a national seismic data facility is both desirable and feasible at this time. Therefore, the most important recommendation of this panel is that a National Center for Seismological Studies be established that will overcome the key data-management problems that we have identified and enhance the availability and effective use of high-quality data sets by the entire seismological community.

Effective use of this national seismological facility will require the following:

(a) Upgrading of the present data-management systems to provide adequately for user needs.
(b) Development of a comprehensive directory of available data, data sources, data quality, and format, together with a search capability implemented at the National Center and at regional facilities, accessible also by remote terminals.
(c) Development of a comprehensive catalog of national and global hypocenters including associated phase data for at least the larger events, together with a computer search capability implemented at the National Center.
(d) Standardization of data formats for the purpose of national and international data exchange.
(e) Establishment of a policy for long-term data retention and archiving.

To meet these goals, the panel's conclusions and recommendations are as follows:

1. DARPA's Center for Seismic Studies (CSS) has several objectives of direct relevance to this panel's objectives. The first is to develop the capability to use digital data from global stations and networks effectively. The second is to support U.S. initiatives in the United Nations Committee on Disarmament (UNCD), Group of Scientific Experts (GSE), to develop improved data-exchange provisions for future test ban treaties. The third objective is to provide a test facility for developing new tools for seismic data analysis and to provide access to organized data bases to support DARPA research.

The panel views the development of this Center as a significant opportunity to benefit immediately from the rapid advances in computer technology that have profoundly increased digital data storage, retrieval, and analysis capabilities during the past decade. Experience with this system will be exceedingly important to the development of intermediate- and long-term capabilities needed by the seismological community. As discussed in Chapter 3, other DARPA-sponsored efforts in the past have led to major seismological advances, and this one clearly will also.

<u>DARPA's Center for Seismic Studies should be considered a prototype for the development of a national</u>

Center for Seismological Studies. Its operation should be structured to provide data and services needed by the seismological research community in general and by several interested federal agencies. Operational costs of a national center are estimated to be between $2 million and $3 million per year. The required level of funding should be provided by the participating user organizations; federal agencies that can be shown to have a need are the National Science Foundation (NSF), DARPA, the U.S. Geological Survey (USGS), the Department of Energy (DOE), the Air Force Technical Application Center (AFTAC), the Agency for International Development (AID), the National Oceanic and Atmospheric Administration (NOAA), the Federal Emergency Management Agency (FEMA), and the National Aeronautics and Space Administration (NASA).

Therefore, the highest priority recommendation of the panel is that a National Center for Seismological Studies be established to assure the effective use of global, national, and selected regional and strong-motion earthquake data. The National Center should meet at least all the functional requirements given in Appendix F of this report.

The main goal of the National Center for Seismological Studies should be to encourage the continuing application of state-of-the-art equipment and techniques to current problems in seismology. A component of in-house research and facilities for visitors will be essential to ensure the continued effectiveness of the Center; one of its major functions will be to ensure the availability and easy distribution of data to external users. Establishing formats and a data-base structure for the archived data and developing standardized software for the routine manipulation of both waveform and parameter data within the Center will de facto solve many of the problems of standardization that currently inhibit the easy exchange of data. Many of the current activities of federal agencies responsible for earthquake monitoring and research can be coordinated through the Center with significant improvement in quality, efficiency, and response time for major seismic events. Some or all of the functions of the National Earthquake Information Service (NEIS, USGS) can be incorporated within the Center. The close coordination of the activities of the Center with the ongoing work of DARPA in testing and implementing new technologies will help to ensure that modern facilities are maintained. The Center will provide a natural focus for the analysis and distribution

of data, either via direct access or through a distribution facility such as that now operated by the NGDS.

2. The operation of the National Center for Seismological Studies should be the designated responsibility of a lead organization, such as a private nonprofit corporation like that of the National Center for Atmospheric Research, the National Science Foundation, or another federal agency. However, the policies and procedures for the National Center for Seismological Studies should be established by the participating user organizations through an advisory board. The funding for operation of this facility should be planned and formally committed on a continuing basis by a consortium of funding agencies to minimize future support problems. Figure 1 shows how the proposed Center could be organized to meet the desired objectives discussed in this report.

Using the Center for Seismic Studies (CSS) as a prototype or a nucleus to develop a national data facility will require immediate action. The CSS will soon be operational (at the end of fiscal year 1984) as a model facility designed for use in comprehensive monitoring of a test ban on underground nuclear explosions. As such, it has a limited operational life for these requirements, and then it will be transferred to some other (currently undetermined) organization or shut down. The CSS represents a rare opportunity to advance our seismological research capabilities taking full advantage of the state-of-the-art computing and data base management technology that has been implemented. The panel recommends that in calendar 1983 the National Research Council's Committee on Seismology initiate steps to convene a meeting of representatives from funding agencies, other user organizations, and the university community to discuss the possible implementation and funding of a National Center for Seismological Studies using the CSS as its prototype.

3. The present distribution system for analog data should be maintained and supported. The technology for acquiring and analyzing data is changing rapidly, but many productive and innovative seismologists in the United States and in other countries will continue to use analog data for at least another decade because new technology will not be available to them and, perhaps, because of lack of training in the use of digital technology. Too, some seismographic stations recording analog data fill important gaps in the global network coverage. Even though the trend is to replace analog by

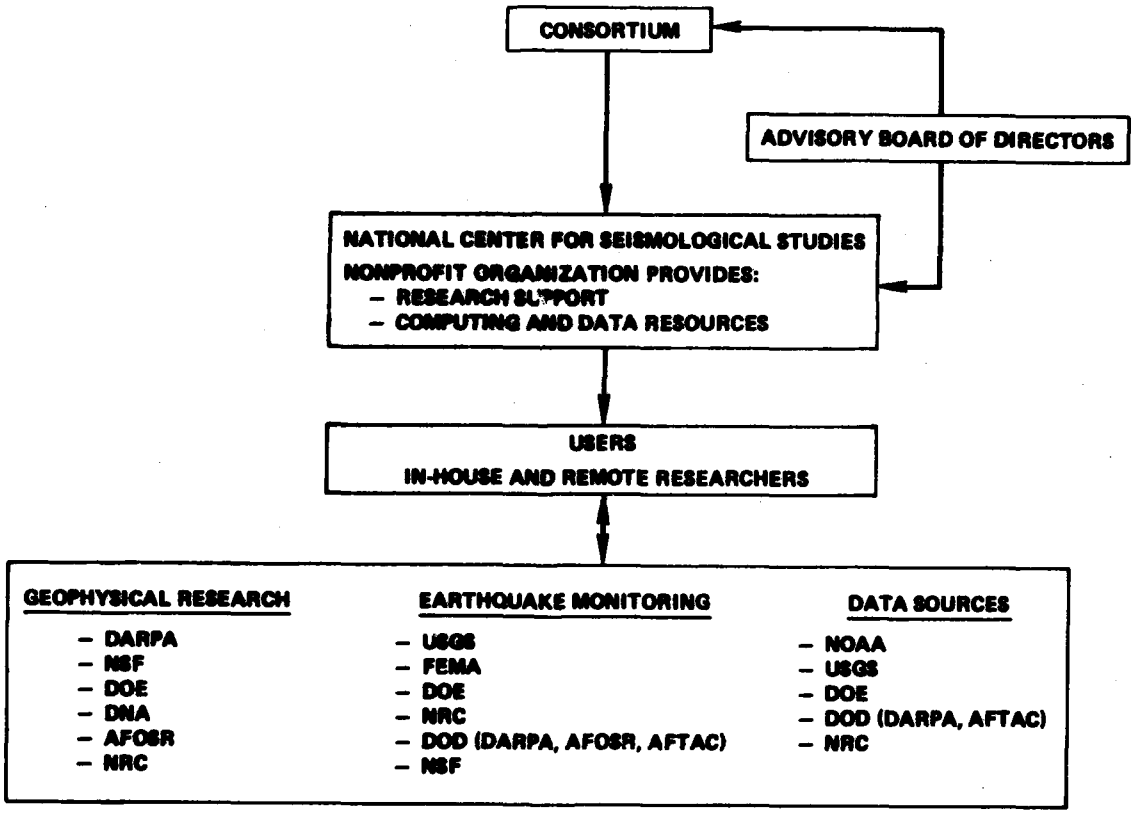


FIGURE 1 Functional organization for a National Center for Seismological Studies.

digital recording in the long term, provisions must be made for handling both types of data.

The National Geophysical Data Center (NGDC), of NOAA, has had the primary responsibility of handling analog data from the Worldwide Standardized Seismograph Network (WWSSN) from its inception; responsibilities include the microfilming of original records and the provision of high-quality film copies to users. These services have been and will continue to be valuable to the seismological community for the foreseeable future. The policy of the Department of Commerce is that the costs of reproduction and distribution are recovered from the users. Rather than considering curtailment of this activity, NOAA should recognize the importance of this national data resource to the seismological community and continue to provide funding for manpower, equipment, and storage facilities adequate for this analog data preparation and distribution. Funding organizations should provide sufficient monies through their grants or contracts to researchers for copying and distribution of the analog records. The panel recommends that the NGDC analog data archiving and distribution facility be maintained and that the costs of operation be shared as decribed above. Alternative techniques for archiving and disseminating analog data should be considered by NGDC to keep costs to users at a minimum.

4. The Albuquerque Seismological Laboratory (ASL) of the USGS has the primary responsibility for assembling the digital data from the evolving GDSN. This is being done as described in Chapter 3. The ASL monitors global and national digital data collection, collates the data and generates network-day tapes (NDTs). The data are available through the NGDC in increments of one day (26 hours) for $100 per tape and NGDC/NOAA should make efforts toward reducing this price. To avoid this high cost, users have acquired data informally at essentially no cost (except tapes) from DARPA and USGS. The USGS has established three regional centers--at the California Institute of Technology, Saint Louis University, and Harvard University--all of which have acquired complete sets of recent NDTs. Visitors can use regional center facilities to copy NDTs.

All of these current modes of distribution of digital data are either too costly or overly cumbersome to users, because the minimum increment of data available is one NDT consisting of 26 hours of data from all the stations in the GDSN. Thus a large number of tapes may be

involved in a single research project and the user commonly must engage in a significant preprocessing computational effort just to get the desired event time windows for analysis. This situation will worsen considerably in the near future as all of a network day will not fit on a single 1600-bpi 9-track tape. It is the panel's understanding that at present almost a network-month of GDSN data resides in active memory or mass storage on the system used to generate the NDTs at the ASL. Thus network tapes of any time increment could be generated with minimal perturbation to the present procedures and no change in format. With modest upgrading of the aging computer systems now used, several network months could reside in the active archive and thus be available for supplying data increments other than 26-hour network days. Other simple types of subsets could easily be generated, such as only selected stations or long, continuous intervals of long-period data needed for free oscillation studies. The panel recommends that alternative types of GDSN digital data increments be made routinely available to users in addition to the standard network-day tape.

5. The problem of which organization is responsible for distributing the GDSN data should be settled without delay. A new generation of technology is coming on line, and adequate planning is required to assure effective use of the resulting data. The current and future earthquake digital data are and will be computer-intensive. User organizations should help with the planning and should share in the data-handling costs. As with the analog data, adequate support of data distribution to users should be provided in the contracts and grants made to researchers.

The panel recommends that GDSN data be made available at a National Center for Seismological Studies and that the Regional Data Centers continue to provide GDSN data to visiting researchers. It is recognized that providing NDTs at the regional centers is only an immediate, partial answer to present distribution problems, but it does provide relatively easy access to the data for some researchers and has the advantage that new users can gain initial familiarization with the data from regional center personnel. Eventually these centers should also have available other standard types of GDSN sets that might be generated, such as event tapes.

6. Data from regional as well as global and national networks must be considered in developing a national

seismological data base. Regional networks have almost exclusively used large numbers of inexpensive, narrow-band, low-dynamic-range stations for gathering arrival time, duration, and amplitude data for accurate location of hypocenters and local magnitude determinations. The earthquake catalogs produced by regional and local networks should be included in the national seismological data base. Procedures should be established to monitor the quality of these catalogs and provide at least minimum uniformity in location and magnitude determination procedures.

In order to expand the use of regional networks beyond the simple cataloging of earthquakes, it is essential that waveform data of high dynamic range and broad frequency content be gathered from a subset of stations within each network (as discussed in Seismographic Networks: Problems and Outlook for the 1980s, 1983). Such data are necessary to extend the understanding of regional propagation and seismic hazard. The mission-oriented agencies concerned with these problems should provide the funding support necessary to develop and install improved instrumentation. Waveform data from such high-quality broadband stations should be incorporated into a national seismological data base.

The panel recommends that data from regional and local networks should continue to be collected, analyzed, and archived by individual network operators but made available in a standard data-exchange format to the National Center as part of the national seismological data base.

7. As discussed in Chapter 3, a high-quality digital and analog strong-motion data base has been established, but its continued effective maintenance and distribution to users by the USGS and NGDC is problematical. Therefore, the panel recommends that strong-motion data be included as part of the national seismological data base located at the National Center for Seismological Studies.

8. High-quality data sets will be generated by AFTAC's new GSE network as described in Appendix E, but data from it will not be made routinely available. However, the panel recommends that AFTAC provide at least event data for large earthquakes for general distribution to other users.

9. The value of most seismological data does not depreciate with time. Large amounts of money have been expended (millions of dollars) by DARPA, DOD, USGS, DOE, NRC, and possibly other agencies, in obtaining seismic

data for special purposes. Examples of such data are those from the LRSM network, LASA, and the WGLP network of stations. The data from such operations can be considered analogous to other valuable, unique works; the stewards of such unique data should have a moral obligation to preserve it for future use of science and technology. A large collection of these unique sets of data are contained in Teledyne-Geotech's Seismic Data Analysis Center (SDAC) in Alexandria, Virginia; but it is scheduled to terminate operations on October 1, 1983. The panel recommends that selected subsets of these special data collections be kept indefinitely as part of an active, accessible national seismological data base. Specifically, with regard to the analog and digital magnetic tape data at SDAC, we recommend that digital data subsets for all teleseismic events of m_b 5.3 and greater and all regional or local events of M_L 3.5 or greater be generated and saved. These data should be culled and digital events tapes made from the usable analog and digital magnetic tapes. The original recordings could then be stored in data warehouses, given away to any interested organization, or discarded. The digital event tapes should be put into a national seismological data base. The panel recommends that the lead organization for the National Center should have the responsibility to archive these data subsets as part of a national seismological data base. The costs should be distributed among the consortium of funding agencies listed earlier in this chapter, apportioned according to the current and projected future use of the data.

10. Who has what data? What are the storage media and formats? Are services, facilities, or special software available to deal with the data? The need for a centralized, computer-accessible directory of such information is real and immediate because none exists at present. Objections to a centralized service providing this kind of information are that the information will not be complete or accurate and that the organization and documentation burden on those who must ultimately provide the information to the centralized facility will be unacceptably large. Our view is that these problems can only grow worse with time, and only by starting now will we have a chance of providing this directory information in later years when the volume of data will have increased significantly. Eventually global, national, regional, and local catalogs should be included.

A first step should be to select an agency or organization that is appropriate for developing and providing such services. One logical immediate choice appears to be NOAA/NGDC. They already provide many related data services and have recent experience with prototype directory development. Alternatively, this could be a designated activity at the National Center. In initiating such a service, major potential contributors should be approached to determine their willingness to participate and to indicate just what information about their data they would be willing to furnish routinely. Information should not just indicate what data exist but should also indicate intended retention time. If data are discarded or become unusable, this information should also be made available to the data information center. As an incentive the contributing organizations should be provided with computer-accessible copies of the directory. The panel recommends that a centralized data directory be developed and maintained as an integral part of the data base that is accessible to users via remote computer terminals. This activity should be considered as part of the mission of NOAA/NGDC or the National Center in developing and maintaining a national seismological data base.

11. Good, well-planned data management is an essential element in effective utilization of seismic data, and it is a keystone to future advances in earthquake seismology. The problems of data handling result from poor planning; fragmented, uncoordinated responsibility for different operations; and poor financial support. These problems must be rectified to derive desired societal benefits from the advancing technological capabilities of the nation and to ensure the health of seismology as a science. The recommendations given above address all the major issues except the sustained funding implicit in the recommendations. The panel recommends that research granting and contracting agencies plan allocations of approximately 10 percent of awarded monies for data management and distribution in studies that involve acquisition of new data and analysis of existing data. In return the users should expect (and demand) effective access at low cost to data subsets that meet their particular needs. Collectively this will be a benefit because the users can devote more effort to the analysis and interpretation, and new users will not be discouraged by the effort required to obtain the data they want.

12. There is a need for standard digital exchange formats. Internationally, IASPEI has formulated standards

for data exchange. The panel recommends the adoption of the IASPEI standards for international exchange of digital data (see Appendix C). This currently incorporates the network-day tape format as a standard for exchange. A few different standard formats may be acceptable, but no more than a few should be adopted.

The data derived from regional networks have been previously described in Chapter 3. The problems associated with network operations are discussed in Seismographic Networks: Problems and Outlook for the 1980s (1983). There are about 50 regional networks operating about 1,600 seismographic stations in the United States. The purpose of the networks is to provide data fundamental to research on seismotectonic processes and earthquake occurrence in the region. There is need for coordination of regional activities, because the seismic data have use to other network operators as well as to private users such as utility companies. Standardization of data formats or at least data-exchange formats will enhance the usefulness of the data. The panel recommends that the Committee on Seismology take steps to convene a workshop for the purposes of establishing standard data and data-exchange formats and standardization of data catalogs.

13. The panel recommends that the National Research Council's Committee on Seismology should assume the lead role in establishing overall policy for the long-term retention of data currently being collected. However, the consortium of supporting organizations should participate in the development of the policy that is adopted. This policy should include saving indefinitely the NDTs, network event tapes, triggered waveform data from a national network, near-field (strong-motion) data, historical and current special data sets, and special data sets from regional networks.

With the advances in data storage capacity and decreasing costs per unit volume of storage, it likely is technically feasible to keep indefinitely a large fraction of the data collected (see Appendix B for current and anticipated data volumes). However, data-base-management costs could be unacceptably large if all the data are kept in active (immediately accessible) mass storage. The guidelines that the panel suggests for waveform data retention are as follows: (a) Keep indefinitely as much of the recorded event data as possible, but use a storage-retrieval archiving architecture that moves less-frequently used data into "deeper,"

less-costly storage where it is still accessible but with some delay. (b) Apply a uniqueness criterion to the recorded events that are considered for deletion, that is, how common is the event, and how many opportunities exist to obtain equivalent data if needed. Example candidates for deletion on that basis include microearthquakes in areas of high seismicity or repeated quarry blasts at the same location. (c) Update the older archived data sets to conform to current storage and retrieval formats, that is, older data sets should be as readily accessible as the newer data.

14. Finally, because of the rapid technological changes in computational and data-handling capabilities, the panel recommends that the Committee on Seismology review the status of data problems in seismology on a year-to-year basis to ensure effective utilization of earthquake seismic data.

REFERENCES AND BIBLIOGRAPHY

Directory of World Digital Seismic Stations. 1982. World Data Center A for Solid Earth Geophysics, Report SE-32. R. Ganse and C. R. Hutt. Boulder, Colorado. 439 pp.

Global Earthquake Monitoring: Its Uses, Potentials, and Support Requirements. 1977. Panel on Seismograph Networks, Committee on Seismology, Assembly of Mathematical and Physical Sciences, National Research Council. National Academy of Sciences, Washington, D.C. 75 pp.

Seismographic Networks: Problems and Outlook for the 1980s. 1983. Report of the Workshop on Seismographic Networks, Committee on Seismology, Commission on Physical Sciences, Mathematics, and Resources, National Research Council. National Academy Press, Washington, D.C. 62 pp.

Strong-Motion Earthquake Instrument Arrays. 1978. Proceedings of the International Workshop on Strong-Motion Earthquake Instrument Arrays, May 2-5, 1978, Honolulu, Hawaii. Sponsored by the National Science Foundation and United Nations Educational, Scientific and Cultural Organization. W. D. Iwan, ed. California Institute of Technology, Pasadena, Calif. 103 pp.

U.S. Earthquake Observatories: Recommendations for a New National Network. 1980. Panel on National, Regional, and Local Seismograph Networks, Committee on Seismology, Assembly of Mathematical and Physical Sciences, National Research Council. National Academy Press, Washington, D.C. 122 pp.

U.S. Strong-Motion Earthquake Instrumentation. 1981. Proceedings of the U.S. National Workshop on Strong-Motion Earthquake Instrumentation, April 12-14, 1981, Santa Barbara, California, Earthquake Engineering Research Institute Universities Council for Earthquake Engineering Research, and at the California Institute of Technology.

APPENDIX A:
CHARGE TO THE
PANEL ON DATA PROBLEMS IN SEISMOLOGY

Seismic data, and especially digital seismic data, have been increasing rapidly in recent years. Furthermore, data sources are diverse, and data are being generated under a wide range of programs supported by USGS, NSF, DOD, NOAA, NRC, and other government and private institutions. In order to obtain maximum scientific benefit from seismic data, their collection, storage, distribution, and analysis require careful planning. The Panel on Data Problems in Seismoloyg, of the Committee on Seismology, should review the seismic data problems and make recommendations for organization, selection, storage, and distribution of the data. Study should include both existing and anticipated analog and digital data from global, regional, and local seismic networks and strong-motion instruments. Seismic profiles, ground deformation, and other complementary data sets should also be considered. The panel should recommend procedures and policies to ensure the availability, timely distribution, and analysis of these data.

APPENDIX B:
ESTIMATES OF PRESENT AND FUTURE DIGITAL DATA QUANTITIES

The following characterizes the magnitude of the present and future digital seismic data waveform problem for global and regional seismic networks.

The Regional Seismic Test Network (RSTN) stations are representative of the most advanced and digital-data intensive stations in global seismic networks. Each station produces three short-period channels sampled at 40 sps, three medium-band channels sampled at 4 sps, and three long-period channels sampled once per second. All samples are 16 bit, encoded in gain-ranged (pseudo-floating point) format to handle very wide dynamic ranges. Other stations, such as ASRO stations, have only short- and long-period bands, and the short-period sampling rate is reduced to 20 sps. It is also possible to have broadband instruments, sampled at a rate appropriate for short-period data, from which all bands can be computationally extracted. But the RSTN represents a good basis on which to characterize the data problem for global networks.

Regional or local networks are even more oriented toward short periods and high-frequency data. Often only a single vertical component is of interest. Sampling rates for digital data are typically in the 50-100 sps range. In what follows we assume that a single station in a regional or local network will produce 100 samples of 16-bit data each second.

The number of advanced digital stations available today for inclusion in a global network is on the order of 30. It is not unreasonable to assume that may grow to be as much as 100. The basis for this assumption is a review of WWSSN data. A subset of some 60-100 WWSSN

stations has produced the vast majority of all WWSSN arrival times and film data. It seems reasonable to assume that the digital network may grow to that size but probably no larger.

The number of seismometers in regional and local networks is very large already, although most are not yet digital. In the United States alone there are some 1,600 stations in such networks (Seismographic Networks: Problems and Outlook for the 1980s, 1983), with perhaps half of those collected as analog data but processed and handled in digital form. This processing is done by various universities and government agencies, which act as regional network analysis and processing centers. Based on this we will assume that there might be as many as 2,000 digital stations at some future time.

Table B.1 summarizes the magnitude of the digital seismic waveform data problem for global and regional networks. First, it is clear that in terms of bulk of data the regional networks represent the largest potential problem. Even after selection of intervals for retention, the amount of data is probably beyond what is practical to retain for long periods of time. Also, the nature of such networks and their data may not justify saving large amounts. In the case of global networks, the total amount of data may be impractical to save using current technology and reasonable funding expectations. But by selecting short-period event windows to save permanently, the quantity of waveform data is reduced to a feasible level for archiving.

Figure B.1 presents additional information concerning seismic data accumulation and storage media. The figure can be used to gain an appreciation of the size of digital seismic data bases. It shows graphically how the size of the data base increases with the average sample rate, the number of stations, and the time of operation. For example, consider the global network station parameters used in the Table B.1 that resulted in 10 selected tapes per year. From Figure B.1 we see that this corresponds to an average sampling rate of about 7 three-component sps, which is correct.

The exact numbers of stations, station characteristics, tape storage capacity, data selection policies, and other related factors are not important for this discussion provided they are reasonable. There are two main points to be noted: first, that digital global networks will produce data quantities that are probably

TABLE B.1 Examples of Data Volume from Global and Regional Networks

Global Network Data Production[a]

Typical Station	Per Station			Per 100-Station Network		
	Raw Bits/Yr	Raw Tapes/Yr[b]	Tapes/Yr[c]	Raw Bits/Yr	Raw Tapes/Yr[b]	Tapes/Yr[c]
3-component SP @ 40 sps	6×10^{10}	60	2.5	6×10^{12}	6000	250
3-component MP @ 4 sps	0.6×10^{10}	6	6	0.6×10^{12}	600	600
3-component LP @ 1 sps	0.15×10^{10}	1.5	1.5	0.15×10^{12}	150	150
Total	6.75×10^{10}	67.5	10	6.75×10^{12}	6750	1000

Regional Network Data Production

Typical Station	Per Station			Per 2000-Station Network		
	Raw Bits/Yr	Raw Tapes/Yr[b]	Tapes/Yr[d]	Raw Bits/Yr	Raw Tapes/Yr[b]	Tapes/Yr[d]
1 component @ 100 sps	5×10^{10}	51	0.36	1×10^{14}	100,500	715
3 component @ 100 sps	15×10^{10}	153	1.1	3×10^{14}	301,500	2145

[a] While the reader may wish to change the parameters, the main point of the table is to show that
(a) Event recording from a global network is no problem in terms of data quantity.
(b) Continuous recording of a global network is possible but difficult to store.
(c) Continuous recording and saving of all regional network data is highly impractical.

[b] Assumes 16-bit data samples and 6250-bpi, 9-track, 2400-ft tapes with 800-byte block size, which provides for 10^9 bits/reel or 6.3×10^7 samples/reel.

[c] Assumes all LP, MP data saved and that on average some 2 minutes of SP data are saved 30 times per day at each station (SP data reduced from 100% to 4%).

[d] Assumes that one minute of regional event data are saved 10 times per day for a net reduction from 100% to 0.7% of all data.

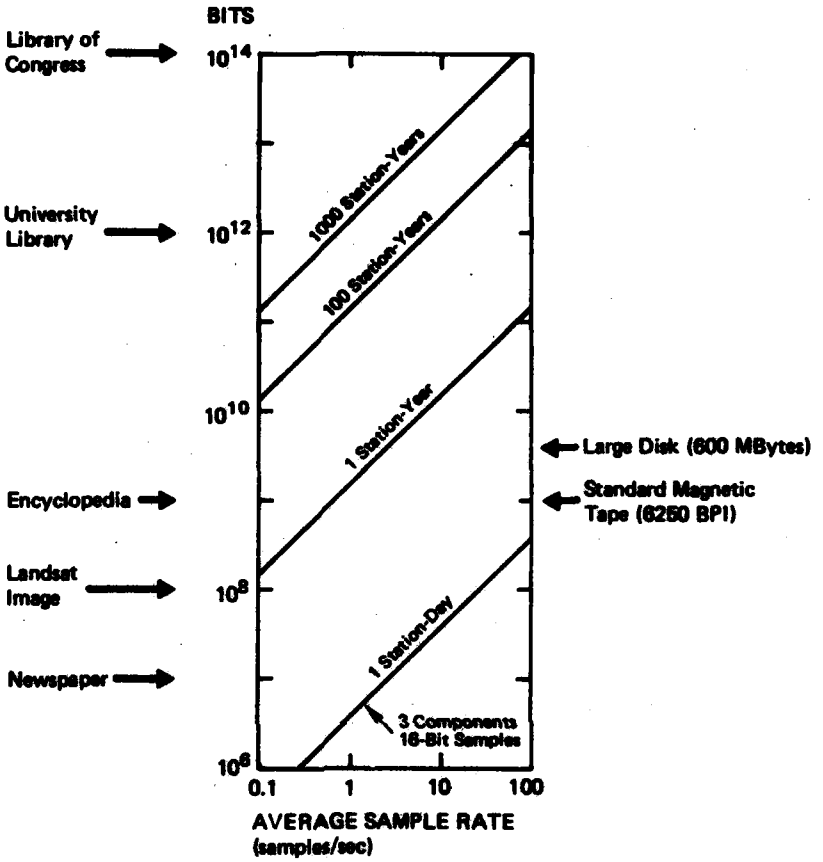


FIGURE B.1 Comparative examples of data volumes.

too large for total retention by currently available means, but judicious selection should reduce the data to manageable quantities while saving almost all data likely to be of significant scientific value; second, the potential quantity of digital data from regional networks is at least an order of magnitude larger.

APPENDIX C

11. NEAR FIELD DATA

IASPEI,

Considering the lack of data obtained from the near field of large earthquakes,

Urges that every attempt be made to obtain near field data, for example by installing strong motion instruments in recognized seismic gaps or in regions of predicted earthquakes, and by moving portable arrays of strong motion stations into the epicentral region as soon as possible following a major earthquake.

12. INTERNATIONAL INSTITUTE OF SEISMOLOGY AND EARTHQUAKE ENGINEERING

IASPEI,

Noting that the UNDP-Japan joint project, established in 1963 by the *International Institute of Seismology and Earthquake Engineering* (IISEE), Tokyo, and operated from 1972 by the Japanese Government, contributed much to the training of seismologists and earthquake engineers in the developing countries, especially with the help of Unesco experts before 1972.

Recommends that the IISEE endeavour to resume its former practice of inviting professors from abroad, seeking national or international funds to achieve this.

13. STORAGE OF HISTORICAL DATA

IASPEI,

Recognizing the importance of retaining old seismographs, bulletins, unpublished readings, clock correction and calibration records and other seismological information,

Noting the danger that these data may be lost due to lack of space or facilities, or for other reasons,

Encourages all stations and institutes to take appropriate measures to improve storage conditions in order to preserve these invaluable data, if necessary seeking financial and technical help from national or international sources.

14. REGIONAL SEMINARS

IASPEI,

Recognizing the usefulness of regional seminars which have been devoted to specific seismological topics, such as that held by CERESIS/OAS on microzonation in Lima in November, 1978, and that held by the IISEE on engineering seismology in Japan in April 1980,

Noting that these seminars have particular value for scientists from developing countries,

Resolves to encourage national and international bodies to organize similar regional seminars and symposia in and/or for the developing countries.

*From I.U.G.G Chronicle, No. 152, November 1981.

APPENDIX C

15. DIRECTORIES OF DIGITAL STATIONS

IASPEI,

Noting the continuing development of new digital seismograph stations by many countries,

Urges that directories of digital recording stations be updated at least annually, and that these directories be made available in computer accessible format.

16. DIGITAL RECORDING FORMAT

IASPEI,

Noting the variety of data formats currently in use for digital data,

Recommends that the global digital seismograph network day tape format be adopted as the initial standard for international data exchange, and that data sets in this format be made available for arbitrary (user defined) event time windows.

Further, recommends that one or more demonstration data tapes be developed to help users.

17. DIGITAL ANALYSIS

IASPEI,

Recognizing that digital waveform analysis is a detailed procedure,

Urges that software for simple types of analysis be made available to seismologists who are relatively inexperienced in digital seismometry.

18. DIGITAL DATA EXCHANGE

IASPEI,

Recommends that digital seismograms in an internationally accepted format be includes in the International Data Exchange data sets.

19. RESOLUTION OF THANKS

IASPEI,

Considering the success of the 1981 Assembly,

Recognizing that much work and time were involved in preparation,

Expresses its thanks to the University of Western Ontario, the Department of Geophysics and the local organizing committee for the fine facilities which were made available and for all the preparation which contributed to an excellent Assembly.

APPENDIX D:
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY'S CENTER FOR SEISMIC STUDIES

INTRODUCTION

DARPA is supporting development of a new facility called the Center for Seismic Studies for which there are two major objectives. The first is to enhance the effectiveness of DARPA-supported research to improve the U.S. capability to monitor nuclear test ban treaties. The second is to develop the capability to meet the international data exchange obligations likely to be needed for a Comprehensive Test Ban Treaty. Satisfaction of the objectives requires development of effective means for organizing and managing large volumes of digital data and providing convenient access to explosion and earthquake data. This effort provides a model for the kind of capability needed to fully exploit modern digital seismic data for broader geophysical research objectives. An advanced data center prototype based on a network of minicomputers has been designed and is being developed at the Center. Seismic data are being collected from sources ranging from digital data transmitted via satellite (Regional Seismic Test Network) to parameter data transmitted via the WWO/GTS telex and the ARPANET. Data from the Global Digital Seismic Network and other data sources are collected on tape. The data-base-management system will organize data for access on such user-selectable criteria as event, station, and geographic region.

The Center for Seismic Studies facility was recently established to house the prototype, a small resident research staff, and visiting scientists. It supports data-exchange experiments being conducted by the United Nations Committee on Disarmament to develop the concept and functions of an International Data Center for treaty

monitoring. A Remote Seismic Terminal is also being developed for ease of international data exchange and remote access to the resources of the Center. Completion of the operational prototype to support the research and international data center functions is scheduled for late 1984.

COMPUTING AND DATA REQUIREMENTS

To support these objectives, capability is required to collect, process, and organize high-quality digital data and make them conveniently available. Seismic-data-analysis capability is required to support easy display and manipulation of seismic waveform and nonwaveform (maps, focal solution) data. The nature of seismic data makes the problem complex in that it is recorded on various media including paper records, film, and magnetic tape in analog and digital format. Receipt of data varies from close to real time to several months and sometimes years. Parameter data must be integrated with waveform data, and historical data must be combined with current data to construct organized data bases for research. A prototype seismic data center was designed and developed to provide these functions and to meet current and evolving research and data-exchange needs.

SYSTEM ARCHITECTURE

The required functions and the need for expansion have been satisfied through the development of a distributed computer system architecture. A number of minicomputers, connected by a local computer network, form the major subsystems of the prototype: Communications, Database, Seismic Analysis, and Remote Access. The computer resources include six Digital Equipment Corporation computers, three PDP VAX 11/780 computers, three PDP 11/44 computers, a Megatek waveform graphics terminal, and a Megatek nonwaveform graphics terminal. The operating system controlling the computers is UNIX, a Bell Laboratories operating system. Each of the subsystems performs a set of functions.

The Communications Interface Subsystem receives all incoming data including that from the NSN Regional Seismic Test Network, which is received via a satellite terminal at the Center. It also provides a variety of

ways to communicate with the Center for convenient access and data exchange. These include the ARPANET, TYMNET, dedicated lines, dial-up capability, mail, and the Remote Seismic Terminal. Through the Communications Interface researchers access the Center's data and computing resources.

The Database Management Subsystem receives, organizes, and archives all the data at the Center. It is based on Ingres, a relational data base supported by UNIX. Through the data base, requests for data can be formulated in seismological terms and efficiently stored and retrieved. Data requests can take the form "get all the data from a particular seismic region" and the system will provide the parameter and waveform data in an integrated form independent of its storage location on disk and tape. The digital data at the Center will include all historical and future Global Digital Seismic Network data condensed into archive format and referenced by event rather than time. The International Data Collection Experiment Database prepared by Sweden for the Group of Scientific Experts and other selected research data bases will also be in the Center data library. The parameter data base includes the National Earthquake Information Service and International Seismological Centre (ISC) catalogs, parameters received from Canadian and UK arrays, and data received over the World Meteorological Organization telex system. The entire World Wide Seismic Station Network film library is also at the Center.

The heart of the system is the Seismic Analysis Subsystem, which has been designed to provide a suite of algorithms for testing large data bases as well as the capability for easy integration of new ideas in seismic signal processing. Interactive signal analysis is possible through the use of the graphics terminals. One terminal displays waveform, while the other can display maps, focal solutions, ray-tracing solutions, and other analysis tools.

The capabilities and tools developed for the Center for Seismic Studies are representative of the capability desired for the National Center discussed in this report.

APPENDIX B:
AIR FORCE TECHNICAL APPLICATION CENTER'S SEISMOLOGICAL ACTIVITIES

The United States Air Force is developing a major improvement in the seismic segment of the Atomic Energy Detection System (AEDS). This improvement known as the Global Surface System (GSS) will provide the AEDS with a fully integrated digital-technology-based system capable of detecting and locating seismic events in the Union of Soviet Socialist Republics (USSR) and the Peoples Republic of China (PRC) shortly after their occurrence. The system will be designed for expansion to accept and process data from additional sources that will eventually lead to worldwide coverage and, when combined with additional system improvement, will allow prompt explosion identification.

Selected AEDS stations are being upgraded to provide an unmanned digital data collection and transmission capability (facility) to forward data from an expanded set of sensor arrays. Seismic waveform data collected at these remote stations will be transmitted through high-quality digital transmission circuits using error-detection and error-correction techniques. Transmission paths are predominately comprised of satellite links, with the delivery of a data circuit to an Earth station at the central headquarters facility.

A central headquarters Technical Operation Subsystem (TOS) is being designed to provide a high degree of automation in the processing of the seismic data. State-of-the-art signal detection and signal association algorithms will perform the primary seismic data screening such that experienced seismologists can devote time to evaluating events of interest using advanced techniques. Automation will provide a much broader event detection capability with less manpower than previous systems. A data archive will provide historical files for subsequent research or development efforts.

APPENDIX F:
PROPOSED FUNCTIONAL REQUIREMENTS FOR THE NATIONAL CENTER FOR SEISMOLOGICAL STUDIES

The National Center should have at least the following functional requirements.

A. Maintaining an informational data base containing:

o A comprehensive directory of national and international seismic data sources.

o Bulletins of hypocenter and associated phase data [International Seismological Centre (ISC), NEIS, and regional and local reports)].

o Special descriptions of earthquakes (e.g., Smithsonian bulletin of short-lived phenomena, damage and geologic observations, and intensity data, i.e., MMI).

B. Maintaining a digital and analog waveform data base containing:

o GDSN data (network-day tapes) and latest segment of continuous data (e.g., RSTN, GTSN).

o International Deployment of Accelerometers (IDA) network data.

o Selected high-quality data from foreign stations.

o National network data (consisting of selected stations in regional networks that feature digital recording, high dynamic range, and broad bandwidth).

o Special event data sets (e.g., digital data and/or data from digitized analog records from current systems, LASA-type arrays, LRSM, and analog magnetic tape, paper, or film recording from permanent observatories.

o WWSSN film chips organized both by station and network-day sequences.

o Analog film recordings for other historic data sets (e.g., LRSM and permanent observatories).

o Near-field (strong-motion) analog and digital data.

C. Developing and maintaining an effective data-base-management system to provide users with desired seismological information and data in a timely manner. The system should be versatile to meet the types and combinations of features as specified by the data users.

D. Providing user services at the NCSS as follows:

o Quick look and browse for waveform data.

o Preprocessing of data (e.g., rotation of components, filtering, spectra, record sections, reduction to a common instrument response).

o A variety of graphical display choices.

o Limited access by users to computers for intensive data manipulations and preliminary analyses of data.

E. Implementing new technology and software to improve the NCSS's capabilities to provide data services and products.

F. Providing periodic training sessions to educate new users.

G. Maintaining software necessary for the following:

o Routine event location and bulletin generation.

o Preprocessing of data.

o Selected analytical techniques.

o Real-time data manipulation for event detection and location using a limited number of stations.

H. Distributing data to users under several options:

o Network-day tape (in standard format)

o Network event tape (in standard format)

o Remote access via terminals

o Analog waveforms

o Access to relational data bases

o Special event tapes (in standard format)

o International data exchange

I. Continuing involvement of research seismologists in the use of the NCSS, possibly through a visiting scientist program.

J. Disseminating selected event data sets generated at the NCSS. These data sets should be constructed to minimize or eliminate (e.g., by means of instrument calibration or response matching) the need for preprocessing so that seismologists with limited experience with digital processing can experiment with the data.

K. Creating a comprehensive event data base consisting of all global network digital data for events of m_b 5.5, plus selected special data sets (e.g., mainshock and aftershock sequences) that also include strong-motion and regional network observations, and other ancillary information characterising the source region.

L. Experiments designed to provide near-real-time capability to determine source characteristics of large earthquakes needed for purposes such as warnings (e.g., tsunami), damage assessment, and deployment of special instrumentation in the epicentral area.

M. Flexibility to incorporate new types of event and waveform data in the future and to accept special digital event data, such as digitized analog seismograms.

N. Archiving of all data, with retrieval architecture structure to accommodate specific user needs and frequency of use.

O. Developing and distributing of experimental "seismic analysis software packages" nominally using higher-order languages such as FORTRAN, that are easy to operate on the most common computer systems extant in the user community.

P. Wide dissemination to the seismological community of information about the NCSS capabilities and the available data bases.

Q. Periodic review sessions, or special symposia at national/international meetings, to discuss research results and capabilities and to identify additional user needs.

APPENDIX G:
ABBREVIATIONS AND TERMINOLOGY

AEDS	Atomic Energy Detection System
AFTAC	Air Force Technical Application Center
AID	Agency for International Development
ALPA	Alaskan Long Period Array
ASL	Albuquerque Seismological Laboratory
ASCII	American National Standard Code for Information Interchange
ASRO	Abbreviated Seismic Research Observatory
CSS	Center for Seismic Studies (DARPA)
DEC	Digital Equipment Corporation
DARPA	Defense Advanced Research Projects Agency
DOD	Department of Defense
DOE	Department of Energy
DWWSSN	Digital World Wide Standardized Seismograph Network
FEMA	Federal Emergency Management Agency
GDSN	Global Digital Seismograph Network
GSE	Group of Scientific Experts
GSN	Global Seismograph Network
GSS	Global Surface System
GTSN	Global Telemetered Seismograph Network
HGLP	high gain long period
IASPEI	International Association of Seismology and Physics of the Earth's Interior
IDA	International Deployment of Accelerometers
IDE	International Data Exchange
ISC	International Seismological Centre
LASA	Large Aperture Seismic Array
LRSM	Long Range Seismic Monitoring
MMI	Modified Mercalli Intensity
NASA	National Aeronautics and Space Administration
NCSS	National Center for Seismological Studies
NDT	network-day tape

NEIS	National Earthquake Information Service
NGDC	National Geophysical Data Center (NOAA)
NOAA	National Oceanic and Atmospheric Administration
NRC	National Regulatory Commission
NSF	National Science Foundation
RST	Remote Seismic Terminal
RSTN	Regional Seismic Test Network
SDAC	Seismic Data Analysis Center
SDCS	Seismic Data Collection System
sps	samples per second
SRO	Seismic Research Observatory
TOS	Technical Operations Subsystem
UNCD	United Nations Committee on Disarmament
USGS	U.S. Geological Survey
WDC	World Data Center
WMO	World Meteorological Organization
WWSSN	Worldwide Standardized Seismograph Network

Event Data: Information such as station phase and amplitude readings, hypocentral locations, magnitude, and other source characteristics.

Waveform Data: Consisting of event signatures, some available on-line in real time but others, such as GDSN network-day tapes, delayed by days to weeks.