Report
of the
OCEANOGRAPHIC DATA ARCHAEOLOGY WORKSHOP

September 13 - 14, 1990

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data, and Information Service
National Oceanographic Data Center

Report
of the
OCEANOGRAPHIC DATA ARCHAEOLOGY WORKSHOP

September 13 - 14, 1990

BRUCE B. PARKER, Chairman
National Oceanographic Data Center
Washington, D.C.

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary

National Oceanic and Atmospheric Administration
John A. Knauss, Under Secretary and Administrator

National Environmental Satellite, Data, and Information Service
Thomas N. Pyke, Jr., Assistant Administrator

National Oceanographic Data Center
Gregory W. Withee, Director

EXECUTIVE SUMMARY

A critical requirement for climate and global change research is the availability of global oceanographic data covering long time periods. Substantial resources are being allocated for global ocean programs such as TOGA, WOCE, and JGOFS, and for the establishment of a true global ocean observing system. However, researchers studying long-term changes in the ocean will have to wait many years, even decades, for long enough data sets to accumulate from these projects. Our only recourse, therefore, is to take full advantage of data that have been collected over past decades. Unfortunately, it has been estimated that perhaps half of all global historical oceanographic data may not presently reside in any of the maintained data archives around the world, and thus are not available to climate and global change researchers. Many of these data may be at risk of being lost forever. It is time to put forth an ambitious integrated effort to recover these data before it is too late.

The Oceanographic Data Archaeology Workshop was sponsored and hosted by the U.S. National Oceanographic Data Center (U.S. NODC) in Washington, D.C. on September 13 and 14, 1990. It was funded by the NOAA Climate and Global Change Program as a result of a reviewed proposal for FY 1990 funds. The participants included scientists and data managers from seven countries (USSR, Japan, Korea, Australia, Denmark, Chile, and the USA) who had an understanding of the needs of the research community and/or a broad knowledge of the availability of historical data and possible methods for finding and acquiring those data.

The main objective of the Workshop was to obtain enough information from the participants to design and implement an integrated data archaeology approach for the efficient and extensive acquisition of historical global oceanographic data for use by the climate and global change research community. A second objective was to gain an initial sense of the priorities -- specifically, what are the areas where data are most needed, and what aspects of data archaeology are most important to pursue.

An integrated approach to oceanographic data archaeology will include the following activities:

- setting priorities based on geographic and temporal needs of the research community and the need to "rescue" data at risk,
- summarizing existing archived data sets in the world's data centers and data gathering institutions,
- summarizing known manuscript and analog data that should be digitized,
- discovering the existence of unarchived historical data,
- convincing institutions to provide the newly discovered historical data (preferably in digitized form) and the necessary documentation,
- digitization (optical or manual) of manuscript and analog data,
- performing quality control activities,
- making the data accessible to the scientists who need them, and
- increasing communication among international scientists and data managers about oceanographic data sets and data archaeology efforts.

It was agreed that initially data archaeology efforts should concentrate on hydrographic profile data, including nutrients. Participants at the workshop have worked with historical global data sets in producing climatologies or carrying out dynamic analyses, and they provided the Workshop with information on some specific data needs and deficiencies.

Summarizing the data held in the world's maintained archives will allow researchers and data managers to see what data are available, where the gaps in data coverage are, and what data they have or know about that are not in the archives. Data distribution maps for various time periods will provide an economical means to provide this information. Descriptions from each data center about their holdings, including data in manuscript and analog form, will also be useful. (In preparing such a report for the Workshop the ICES Hydrographer discovered 200,000 cards of hydrographic data in the cellar of ICES, which included deep-water data from the eastern North Atlantic, an area where there have been little historical data.)

It will be important to inventory known manuscript and analog data sets (e.g., 158,000 MBTs at Scripps, the newly discovered ICES data, manuscript data at the World Data Center A (WDC-A), Oceanography, etc). This inventory can be used to determine priorities for expending resources to digitize data, and to discover if the data may have already been digitized. This inventory will be continually updated as new data sets are discovered.

Mechanisms for discovering the existence of data never sent to a maintained archive include: the use of data distribution maps sent to data centers and other data collecting institutions around the world, where each can see if their own data sets are included; ROSCOP forms; historical yearly lists of the National Oceanographic Programs and Declared National Programs; and various WDC publications, which also point to potential sources of data. Participants at the Workshop also felt that letters could be sent out to the Directors of all the world's oceanographic data centers and data collecting institutions, emphasizing the importance of climate and global change research, and asking for their help in finding historical global oceanographic data. Such a letter might come from the Administrator of NOAA. It might accompany the data distribution maps to be sent out by NODC/WDC-A.

Once the existence of a particular data set is known, the next and often most difficult step is to convince the institution or scientist to provide the data (and metadata/ documentation), hopefully in digitized form. Methods for obtaining these data include: monetary or ADP support for the institution; enhanced visiting scientist programs; joint research projects; trips to institutions holding valuable data (including the provision of training in quality control techniques and format conversion); and the development of a PC software program to make data entry, format conversion, and quality control easier and more reliable.

It will be helpful to the entire quality control process to produce a compilation of all instrumentation and measurement techniques used over the years for a particular data type, including known data problems; this is especially important for nutrient data.

It was agreed that a frequently published Ocean Data Archaeology Newsletter would serve many useful purposes. It would help to make and keep the scientific and data management communities aware of data archaeology efforts and could provide the kinds of information that will stimulate the discovery of unarchived historical data sets. A telemail mailbox (or possibly a bulletin board) would allow for rapid communication on archaeology issues and potential data sets.

CONTENTS

1. BACKGROUND

A critical requirement for climate and global change research is the availability of global oceanographic data covering long time periods. Substantial resources are being allocated for global ocean programs such as TOGA, WOCE, and JGOFS, and for the establishment of a true global ocean observing system. However, researchers dealing with long-term changes in the ocean will have to wait many years, even decades, for long enough data sets to accumulate from these projects. Our only recourse, therefore, is to take full advantage of data that have been collected over past decades. Unfortunately, it has been estimated that perhaps half of all global historical oceanographic data may not presently reside in any of the maintained data archives around the world, and thus are not available to the climate and global change researchers. Many of these data may be at risk of being lost forever. It is time to put forth an ambitious integrated effort to recover these data before it is too late. This critical juncture, when interest in climate and earth history is high and funds are available, may be our last chance to save many of these valuable old data.

In a report to the NOAA Panel on Climate and Global Change by the Working Group on Data Management (chaired by Francis Bretherton, University of Wisconsin) the importance of data archaeology is emphasized. In that report "data archaeology" is defined as "the reconstruction of past climate and other aspects of global change from existing data. It involves a mix of seeking out, restoring, evaluating, correcting and interpreting past data sets."

The large amounts of valuable additional data that do exist are scattered among many institutions around the world and are frequently known only to the particular institutions housing them. To obtain such data a data center must discover their existence, convince the institution to provide them (preferably in digitized form) and the necessary documentation, assess their quality, digitize them (if they are in manuscript form), carry out quality control checks, archive them, interpret them, and make the data accessible to the scientists who need them. An important part of this process is the setting of priorities based on input from the scientific research community on their requirements, i.e., identifying important gaps in geographic and temporal coverage for a specific data type.

Data archaeology activities have heretofore been performed on a random basis. Data centers have obtained historical data sets as they found out about them, but they have been prevented by limited resources from carrying out major data archaeological efforts. A few scientists have obtained as much data as they could from data centers or fellow scientists. Some have sought data out from contacts in other countries, or requested that the data centers search them out.

The World Data Centers for Oceanography located at the NODCs in Washington, USA, Obninsk, USSR, and Tianjin, PRC, are important sources of historical global oceanographic data, because they receive data on a regular basis from data centers and data collecting institutions around the world. These Centers serve as the final archive for data sets from most global ocean programs; they also seek out historical data sets when they become aware of their existence (as their resources allow). For example, WDC-A's Catalogue of Data and its annual Change Notices constitute a comprehensive global inventory of data sets. Automation of this Catalogue's Change Notices has begun, as well

as efforts to produce a common catalogue for World Data Centers A, B, and D. WDC-A also publishes an annual summary of its oceanographic data exchange activities and maintains an archive of grey literature, which are summarized biannually in its Catalogue of Accessioned Publications. This archive contains references to data from around the world. WDC-A, Oceanography, and other data centers, also archive ROSCOP marine data inventory forms and other records of data gathering activities, which point to sources of data. In addition, the U.S. NODC and other NODCs have bilateral data exchange agreements which bring in valuable data.

The World Data Centers and most NODCs now have either ongoing research activities at their institutions or cooperative arrangements with universities. Interaction with research scientists is critical both for setting priorities on data acquisition and for quality control activities.

Data centers and scientists working together, taking advantage of modern PC-based ADP capabilities and telecommunications, should be able to locate and obtain important historical data sets. What is needed is a well thought out, totally integrated approach, along with support from such programs as the NOAA Climate and Global Change Program.

2. OBJECTIVE OF THE WORKSHOP

The main objective of the Workshop was to obtain enough information from the participants to design and implement an integrated data archaeology approach for the efficient and extensive acquisition of historical global oceanographic data for use by the climate and global change research community. A second objective was to gain an initial sense of the priorities -- specifically, what are the areas where data are most needed, and what aspects of data archaeology are most important to pursue, so that an archaeology program can begin right away and take advantage of whatever funding can be obtained.

3. WORKSHOP PARTICIPANTS

The Workshop was sponsored and hosted by the U.S. National Oceanographic Data Center (U.S. NODC) in Washington, D.C. on September 13 and 14, 1990. It was funded by the NOAA Climate and Global Change Program as a result of a reviewed proposal for FY 1990 funds. The participants included scientists and data managers from seven countries who had an understanding of the needs of the research community and/or a broad knowledge of the availability of historical data and possible methods for finding and acquiring that data. The participants also collectively provided good coverage of all regions of the globe.

The Workshop was chaired by Bruce Parker, Director of World Data Center A (WDC-A), Oceanography and Chief of the Information Services Division at the U.S. NODC. Other U.S. participants included: Ronald Moffat, Associate Director of WDC-A, Oceanography; Sydney Levitus, Chief of the Product Development Branch at the U.S. NODC (the applied research group at the U.S. NODC); Joseph Reid and Warren White of the Scripps Institution of Oceanography; Pembroke Hart, WDC-A Coordination Office, National Academy of Sciences; and several other members of the U.S. NODC (see Appendix I for complete list with addresses and phone numbers). Participants from other countries included: Harry Dooley, the Hydrographer for the International Council for the Exploration of the Sea (ICES), headquartered in Denmark; Nikolai Mikhailov, Chief of Laboratory, World Data Center B, Obninsk, USSR; Arkady Alekseev, Chief Deputy Director, Pacific Oceanological Institute, USSR Academy of Sciences, Vladivostok, USSR; Shin Tani, Chief, International Programs Group, Japan Oceanographic Data Center, Tokyo, Japan; Ben Searle, Head, Science and Oceanography, Australian Oceanographic Data Centre; Gary Hopwood, Australian Oceanographic Data Centre; Kee-Soo Nam, Head, Dept. of Data Management and Computer Services, Korea Ocean Research and Development Institute, Seoul, Korea; and Ricardo Rojas, Instituto Hidrografico de la Armada, Chile.

A month prior to the Workshop a Working Draft Plan for an Oceanographic Data Archaeology Project was sent to each Workshop participant (as well as other interested individuals who were unable to attend the Workshop), in order to provide a framework for discussions at the workshop and to help stimulate ideas. Responses to and ideas stimulated by the draft plan were incorporated into the Workshop. The entire Workshop was (audio) taped.

4. GENERAL OVERVIEW OF THE ISSUES DISCUSSED AT THE WORKSHOP

As was suggested in the Working Draft Plan sent to each participant, an integrated approach to oceanographic data archaeology would likely address each of the following activities:

- o setting priorities based on geographic and temporal needs of the research community and the need to "rescue" data at risk,
- o summarizing existing archived data sets in the world's data centers and data gathering institutions,
- o summarizing known manuscript and analog data that should be digitized
- o discovering the existence of unarchived historical data,
- o convincing institutions to provide the newly discovered historical data (preferably in digitized form) and the necessary documentation,
- o digitization (optical or manual) of manuscript and analog data,
- o performing quality control activities,
- o making the data accessible to the scientists who need them, and
- o increasing communication among international scientists and data managers about oceanographic data sets and data archaeology efforts.

The Agenda for the Workshop (Appendix II) highlighted each of these activities. This report will describe the discussions and conclusions reached by the Workshop participants, in approximately the chronological order that they occurred, and thus topics will appear in approximately the same order as they appear in the Agenda.

Initial discussions at the Workshop confirmed the general belief that there was a critical need for historical global data and this was the time for an ambitious integrated approach to oceanographic data archaeology. Participants cited various best (but rough) estimates of the amounts of data around the world that have not been sent to the world's maintained data archives. For example, it was estimated that approximately 50% of all the temperature profile data ever taken is not in the world's data centers. Such estimates are, of course, difficult or impossible to prove, but they are an indication of the general feeling of scientists and data managers who work with these data, and an indication of "treasures" that may be acquired from a concerted data archaeological effort.

5. THE OCEANOGRAPHIC DATA NEEDS OF CLIMATE AND GLOBAL CHANGE RESEARCHERS

It is the requirements of climate and global change (C&GC) researchers for historical global data that drive this data archaeology effort. Thus, it is important to obtain their input in determining priorities. Several participants at the workshop have worked with historical global data sets in producing climatologies or carrying out dynamic analyses, and they provided the Workshop with their priorities. In addition, it was felt that there should be mechanisms for determining the needs of other C&GC researchers, specifically what types of historical data they need most, from what areas of the globe, and for what time periods.

A second view was also presented at the Workshop. This viewpoint did not deny that as many C&GC researchers as possible should be able to provide input and be a part of the process, but this view held that, until we begin to discover many of these unarchived data sets, there is no need to set priorities. We need to go after all the data sets that we are presently aware of -- there will be uses for all of these data. The only priority should be determined according to whether particular data sets may be at risk, that is, in real jeopardy of being lost due to: deteriorating storage medium, the possibility of being inadvertently thrown out, or retirement of key individuals who understand the methods and formats used. If, however, the archaeology effort turns up many data sets, then the limited resources for this effort demands some type of prioritization, which will probably strike a balance between consideration of researchers' needs (as best they can be determined) and whether data are at risk.

There was agreement that initially the ocean data archaeology effort should at least concentrate on one type of data, or a group of related data parameters. It had been suggested in the Draft Working Plan that salinity and temperature profile data might be selected as the first data type on which to concentrate, for the following reasons: (1) these data are important to climate and global change researchers; (2) they represent a large data set that has been analyzed by a number of scientists, but which still has important gaps relative to geographic and temporal coverage; and (3) they are a type of data apparently residing at many institutions around the world. However, there is also a critical need for nutrient and chlorophyll data, which are often measured simultaneously with temperature and salinity profile data, and thus should also be obtained when discovered. There appeared to be general agreement at the Workshop to concentrate on hydrographic profile data, including nutrient data.

(When asked how far back temperature and salinity data can be considered reliable, many participants agreed that temperature data were reliable at least since the invention of the reversing thermometer, and that salinity data were reliable since the establishment of standard sea water. However, some participants felt that this should not be used to limit the data we obtain; we should get all the historical data we can and flag the poor data after quality control procedures have been carried out.)

It was felt that the data archaeology approach developed with these initial data types should be applicable to other data types. It was also felt that there should be mechanisms to keep track of data of other types discovered while carrying out this initial effort.
Suggested data types that might warrant a data archaeology effort included: ship drift,

currents, float data, and water levels.

Regarding input from other C&GC researchers, it was suggested that it would be useful to compile a list of all scientists working with the data types being sought initially. The Workshop participants will provide the names of others working with these data, who in turn can point to others. Other suggested sources of names of scientists working with global data included: the AGU Membership Directory (which lists members by areas of interest), IOC, ASFA, various data centers and hydrographic offices, and NASA. A comprehensive bibliographic search using on-line bibliographic services (which the NOAA library does on a regular basis) and the WDC-A Catalogue of Accessioned Publications can add to this list. [This activity may also point to particular data sets not presently in a major archive.] The resulting names and institutions could form a database of international scientists working with these types of data. Such a database could reside on a PC and would include the areas of research in which the scientists have used these data, their publications (refereed and grey), the data sets they have used or are using (with geographic areas and time periods), and their needs. Individuals and institutions in such a database would also be included in the initial mailing list for the data archaeology newsletter that will be produced (See Section 11).

Some specific needs and deficiencies were noted by participants at the Workshop. The Indian Ocean was cited by Warren White as an area where more data are needed and where there is a feeling that there are many data sets that are not in the system; he understood that the USSR, Japan, and India have made many cruises in the Indian Ocean. Shin Tani has looked at the data density of the Indian Ocean at JODC and has found a data density approximately one half that of the global average data density. He also stated that IGOSS data obtained from the Indian Ocean is about one tenth the amount obtained from the Pacific Ocean. Ron Moffatt mentioned that India has never sent in ROSCOPs, but the National Institute of Oceanography of India publishes a newsletter which includes information (e.g. ships and cruises) found on ROSCOPS. It has been difficult to obtain data from India in the past, but this may be primarily because a shortage of personnel at the Indian National Oceanographic Data Center. India made some very positive statements at IODE-XIII last January at the United Nations about providing their data to other data centers. Nikolai Mikhailov stated that the Soviet data center recently received a letter from the Indian National Oceanographic Data Center requesting data for the Indian Ocean; in looking through their archive they found some Indian Ocean data from Soviet research ships and a small volume from the international system through WDC. They are planning to organize an Indian Ocean water bottle data set based on bilateral work between the Indian and Soviet data centers. There is a need to define the status of data in this area. He also spoke of obstacles to obtaining the data, the need to set up bilaterals (e.g., between the USSR and USA) to obtain the data taken by Soviet ships, and the use of the IODE system. He believes only 30-50% of Soviet Indian Ocean data have ROSCOPs.

Joe Reid stated that there are not enough deep-water data in the Bay of Biscay area in the eastern part of the North Atlantic Ocean. There is a need for oxygen and nutrient data everywhere; until the last ten years or so the North Atlantic has had lower quality oxygen and nutrient data. There is also a great scarcity of really deep hydrographic stations or CTDs in the Equatorial Zone. The density at 3000m is critically important in all oceans between 10° N and 10° S.

Syd Levitus stated that we need to acquire data from all areas of the globe and that we should set a goal that perhaps in five years we should have acquired and digitized all known global hydrographic profile data.

Harry Dooley agreed with Joe Reid that there is little data from the eastern slope area of the North Atlantic (but he has found some, see Section 5 of this report); Syd Levitus also agreed that this was an important area.

Steve Patterson felt there is a need for more data around Antarctica and suspects there are more data out there.

6. SUMMARIZING KNOWN DATA SETS

It is important to be able to provide summaries of the digital data that are presently in the world's maintained archives in order to show where the gaps are and to make it easy for various scientists and data managers to see whether their data are in these archives. It is also important to be able to inventory all known manuscript and analog data (and add to this inventory other data as they are discovered) in order to prioritize what data should be digitized first.

Data Distribution Maps

Summarizing the data held in the maintained archives would allow researchers around the world to see what data are available, where the gaps in data coverage are, and what data they have or know about that are not in the archives. Although a geographically-based PC inventory program would be a convenient way to provide such information, its development and population would require a large effort It was felt that simple data distribution maps for various time periods would provide a much more economical means to provide this information and such an effort could be started immediately.

Syd Levitus has disseminated data distribution maps and has received a tremendous amount of feedback about data that were not shown on those maps. Such maps show up cruise tracks that can easily be identified by those familiar with particular data collection cruises. Ben Searle felt that many large directories or inventories are not successful, because it takes too much effort for many data centers (who may only have a few people) to send their inventories in. He felt that the data distribution map approach was much preferred, because it would be much easier for these data centers. He stated that there are various inventories which we could collect and which may turn up useful information, but we could better put our efforts elsewhere than to produce a large inventory of the world's holdings of hydrographic data. All the participants felt that we continually need to prioritize our activities to bring in the most data with our limited resources and that we must carry out activities that we can begin quickly.

It was felt that the approach that would bring the most immediate results, using a minimum of resources, would be for the U.S. NODC/WDC-A to produce data distribution maps of its hydrographic data holdings. These maps should be done on a country-by-country basis, and be broken down into decadal or shorter periods (to assure that individual cruise tracklines will be visible). These maps would be sent to the appropriate data center and data gathering institutions in each country. Those institutions would be able to tell from these maps whether all their data are in the U.S. NODC/WDC-A archive (and what data may be in that archive that they don't have). (Samples of some of the data distribution maps that the U.S. NODC has produced since the Workshop are shown in Appendix IV.)

These countries could then provide information about data sets that do not show up on the U.S. NODC/WDC-A maps. They could also produce data distribution maps of their own hydrographic data holdings (or provide latitudes and longitudes and appropriate information to the U.S. NODC for the maps to be plotted there).

Descriptions of Data Holdings

In addition to the data distribution maps, it was agreed that it will be very useful to receive descriptions from each data center about their holdings, including data in manuscript and analog form. Each of the participants agreed to provide such information for his data center.

In preparation for this workshop Harry Dooley, the ICES Hydrographer, prepared a report entitled "An overview of the ICES Oceanographic Data Bank" (attached to this report as Appendix III). This report provides very interesting and important information about a century of hydrographic data stored at ICES. It covers, country by country, primarily the data prior to 1957, but Table 1 includes data up to the present. The purpose was to look at the years for which ICES doesn't have data for certain countries and try to find out why. This approach was successful because it has identified many important gaps and has already turned up substantial missing data for an area where data are badly needed.

For example, Dr. Dooley discovered 200,000 cards of hydrographic data in the cellar of ICES. These cards were all originally thought to be included in the ICES computer database, but that database was found only to include data from Bulletin Hydrographique (which ICES published from 1908 to 1956), and these cards were found to contain considerably more data, including data from Rapport Atlantique, which the ICES Atlantic Slope Committee published from 1921 to 1935. Table 2 of his report lists data in Rapport Atlantique not in the ICES computer database, but all found on these cards. At least 5000 of these cards contained deep-water data, down to 3000 meters, from the eastern Atlantic. Many of the French data on these cards are from the cruises of the "Pourquois Pas". The data that ICES has from these cruises are only for the eastern Atlantic, but those cruises covered a much greater area; all the data from those cruises (at least ten times the amount at ICES) should be found in the French publication Annales Hydrographiques. So another significant result of this exercise has been a pointing to at least one other source of important historical data.

Dr. Dooley also discovered another 20,000 cards in the basement of Charlottenlund Castle (the form ICES headquarters), which are geosorted Danish data that were also found not to be on the computer. These cards included data from the 1896 Ingulf expedition in the Norwegian Sea.

Dr. Dooley stated that he is hoping some of the data listed in Table 2 of his report will exist in digitized form at other data centers. The German data center has digitized the Meteor data from 1925-27, which the U.S. NODC has. The U.S. NODC also has the Discovery data set. The intercomparison of the data distribution maps should help sort some of this out.

Robert Gelfeld, who is helping IOC develop its Marine Environmental Data Inventory (MEDI), was asked if that inventory was at too high a level to be helpful in determining what data the various data centers have. While the inventory consisted of general descriptions of entire data sets and doesn't get down to individual observations, it does give geographical areas, time periods, measured parameters, institution and contact person, so that this inventory may point to some data that we were not aware of. There are presently

approximately 250 file descriptions from 23 countries, but mailings are going out to all 149 IOC member states.

Manuscript and analog data that should be obtained and/or digitized

It will be very important to inventory manuscript and analog data sets that the Workshop participants are aware of now. With this information priorities for expending resources to digitize data can be decided upon. By distributing this list to various data centers we can also find out if any of these data may have been already digitized (especially those in WDC-A; see below). This inventory should be continually updated as new data sets are discovered.

Syd Levitus mentioned 158,000 MBT profiles at Scripps, from the Pacific and Indian Oceans, extending as far back as 1942; these need to be digitized soon before it is too late. He also felt the eastern Atlantic data that Harry Dooley has found at ICES are very important because these data are for an area where there has been little historical data (the NOAA Atlantic Climate Change Program will need these data). Irv Perlroth mentioned the backlog of XBT analog strip charts at FNOC that need to be digitized (on the order of 180,000, of which about half are open ocean); they are world-wide data.

Ben Searle mentioned a couple of thousand analog XBT profiles that AODC has that have not been digitized yet. He mentioned that the New Zealand navy has analog XBT data, as does the New Zealand Oceanographic Institute. AODC has six years of data from the naval research ship HMAS Cook, the majority in digital form, but not in an exchange format. Their Antarctic Division of the Dept. of Science has made a number of cruises with XBT lines. He also gave the Workshop an unofficial copy of a report from the IODE mission to southeast Asia, which doesn't talk about data set specifically, but points to some potential data sets. [See also Bob Gelfeld comments below on Indonesia.] He believes that in the South Pacific there are many small data sets to be found.

Shin Tani said that many of JODC's data are in standard level form, but they still have the analog records, so that they could still be redigitized into inflection-point records. Syd Levitus agreed that it would be nice to have the data in inflection point form, so that we can do studies of mixed layer depth with a little more confidence. Many of the Japanese fisheries data that are at the U.S. NODC is in standard level form, and it would be useful to have these in inflection-point form. He also mentioned that there is such a demand for chlorophyll data that the manuscript holdings at WDC-A should be strong candidates for digitization.

At WDC-A there are thousands of profiles in manuscript form from various data collecting institutions and data centers around the world. These must be digitized, but before that large effort is put in, we must contact these institutions to see if any of these data may have already digitized. Syd Levitus suggested that we can start with the largest data sets and largest data centers first. Shin Tani stated that JODC has spent approximately $100K digitizing manuscript data and that all of JODC's manuscript data have been digitized, except for their Fisheries data. Ron Moffatt said that NODC/WDC-A has approximately 200,000 unprocessed Japanese profiles from the Japanese Fisheries Agency (JFA) Publication (1963-

1985) that are in manuscript form. He showed chartlets depicting JFA tracklines seasonally for selected years.

Ricardo Rojas mentioned that there are data from Chilean Fisheries cruises, but these have not been available, because they have been concerned about how the data will be used; they have been afraid that the data centers will make the data available to everyone. Those who do give them data do so with the stipulation that it not be sold to any private companies without authorization.

The chairman mentioned that while at IODE-XIII last January he was told by Commander Persio Souto, the Diretoria de Hidrografia e Navegacao of Brazil that they had data from repeated lines off the Brazilian coast. Lt. Felix Espinoza of the Chilean Oceanographic Data Center told him that he felt there were probably many other cases like that; Lt. Espinoza had suggested that the U.S. NODC/WDC-A should send some one to all the data centers in South America to find out about such data. Syd Levitus mentioned that Don Olson of the University of Miami is in the process of acquiring BT and hydrocast data from Brazil.

Ron Moffatt mentioned that the U.S. NODC has a bilateral agreement with the Chinese NODC and has received quite a bit of data recently, most of it attributed to their TOGA efforts; but there may also be some historical data to be obtained through that means. In addition to that, the Chinese Academy of Sciences apparently has a different set of data which may not be at CNODC; he was told by some Academy of Science members who visited the U.S. NODC recently that we could ask for their data through the WDC system.

Kee-Soo Nam stated that the Korean Oceanographic Data Center is located in FRDA and that they produce an annual report of manuscript data. A couple of years ago they were trying to key in their MBT data, and he thinks they finished it. But for some reason KORDI couldn't get the data from them. The Department of Defense was digitizing the data for FRDA in return for being able to use them. Ron Moffatt mentioned that WDC-A has data for approximately 37,000 Korean profiles in manuscript form covering 20 years; approximately 7000 of those were digitized by JODC serving as the Kuroshio Regional Data Center during the CSK program. Kee-Soo Nam mentioned that KORDI keyed in 18 years of Yellow Sea data, which is archived at KORDI now.

Robert Gelfeld stated that we are aware of other national data centers that have data, but we just don't know their exact holdings. We need to contact each and find out what they have. He and the previous U.S. NODC Director made a trip to the Indonesian Oceanographic Data Center where they identified their data holdings and upon returning to the US put together a plan to help them automate it; that effort apparently died because of the bureaucracy over there. The Indonesian data center had substantial holdings of hydrographic data that were not digitized. Pete Topoly mentioned that this was a good example of where many of the data were not taken by researchers, but were collected by some one for other purposes, and thus are not easy to obtain. In Indonesia the Fisheries Department did a lot of hydrographic work, but they haven't wanted to release the data because they are economically sensitive.

A key question may be: are there ways of overcoming some of these economically

induced restrictions that countries (especially in the Third World) are putting on these data? Ron Moffatt stated that in going through the Netherlands data center (which has had ties with Indonesia over the years) it has become obvious that Indonesia has not let its data out. Ben Searle stated that there was an IODE mission to southeast Asia in 1987 (Indonesia, Thailand, Philippines, and Malaysia) and a large number of organizations were visited. They found that the problem tended to be resources; they have data in manuscript form and many of the organizations had only one PC as the data system. He provided a draft report to the workshop.

Arkady Alekseev stated that the Pacific Oceanology Institute sends data to and gets data from the Soviet data center in Obninsk, but other Soviet institutions around the Northwest Pacific may also have data available.

The participants agreed that we should try to obtain the data at the Indian Oceanographic Data Center (as discussed earlier in the Workshop). Ben Searle's supervisor has just come back from a visit to that data center. Nikolai Mikhailov mentioned earlier that the Soviet and Indian data centers will be working on a joint Indian Ocean data project.

Ricardo Rojas mentioned that the Peruvians have a regular XBT monitoring program. Irv Perlroth noted that Don Hansen at NOAA's Atlantic Oceanographic and Meteorological Laboratory (AOML) has worked with Peruvian and Ecuadoran data. These data have been proved to the U.S. NODC through AOML; they have been mostly contemporary, not historical.

Syd Levitus mentioned that Susumu Tabata (Institute of Ocean Sciences, BC, Canada) has Canadian data from the North Pacific dating from the 1930's.

Ron Moffatt mentioned that Swedish coastal section profile data are available in manuscript form in WDC-A, Oceanography. Harry Dooley noted that some of these data have been digitized. A planned updating of the ICES and U.S. NODC/WDC-A data exchange will provide additional information about the availability of Swedish data in automated form after which the Swedish data center will be contacted for missing data.

Harry Dooley provided the Workshop with a copy of the Directory of Source Materials For the History of Oceanography, UNESCO Technical Paper In Marine Sciences, No. 58, which may be useful in the archaeological efforts.

Ron Moffatt asked the advice of the Workshop regarding the value to researchers of Lightvessel data available in manuscript form at WDC-A and elsewhere. Were such data too close to shore or in too shallow water to be useful? Syd Levitus noted that there is a need for these data. As an example, Moffatt distributed a chart showing the locations of German Lightvessels for which profile data are available in WDC-A, noting that the number of stations in manuscript form exceeds 8,000. Harry Dooley noted that ICES has German Bight data since 1930 available in automated form. Steve Patterson felt that identifying where long records exist would be useful, although he felt that they would be primarily from coastal stations.

Harry Dooley noted that most research ships take thermosalinographs (which are

good to .01), which could be useful. Warren White and Syd Levitus agreed that knowing surface temperature and salinity is important. Ron Moffatt noted that WDC-A has substantial amounts of surface temperature and salinity data in manuscript form.

The chairman raised the question about possible useful data from oil companies? Ben Searle noted that temperature and wave data are collected by oil companies and do become available after a period of time for "declassification". Margaretta Conkright mentioned that some oil platforms are in deeper water and have been operating longer than five years. Shin Tani stated that JODC gets oil rig data, including surface temperatures, but has less than ten years.

The Workshop participants agreed that it will be important to keep track of all data discovered as a result of the archaeological effort, as well as all leads pointing toward potential sources of additional historical data. Here a PC database inventory should eventually be useful. Initially such a database would include descriptions of data that we are presently aware of that are available only in manuscript or analog form and which need to be digitized. It might also include important data sets not yet in the archives but identified by ROSCOP forms or other sources, so that a researcher aware of these data could attempt to locate them. (A fully automated ROSCOP tracking system, as a mechanism for finding additional data, is talked about in the next section.) Data perceived to be at risk and in need of rescue and/or digitization could also be flagged.

7. DISCOVERING THE EXISTENCE OF HISTORICAL DATA

An important part of the ocean data archaeology program will be the development of mechanisms for discovering the existence of data not sent to a maintained archive.

The data distribution maps discussed in the previous section should provide a good mechanism for discovering data sets. They will be sent to data centers and other data collecting institutions around the world, where each can see if their own data sets are included. These institutions may also make data distribution maps of their own holdings, or provide the necessary information to the U.S. NODC/WDC-A (or another data center) for such maps to be produced there.

The inventory of ROSCOP forms available at the World Data Centers (almost 20,000 cruises), ICES (about 11,000 forms), and other institutions provides an important means of determining what data have been collected. A geographically-based PC ROSCOP inventory program would be needed to carry out efficient geographical and temporal searches. Comparison of this inventory against an inventory of data in the archives would identify data that should exist somewhere and may hopefully be obtained in some form. The present ROSCOP log system at WDC-A is well organized but not automated. ICES has developed a PC inventory program for ROSCOP forms, which perhaps could be used or modified to develop a comprehensive WDC-A ROSCOP database. ICES also has a convenient ROSCOP data entry program. (Both these programs were demonstrated for the Workshop by Harry Dooley.) Most of the effort will be in entering the information from the ROSCOP logs. Until such a program is developed, the present WDC-A, ICES, and other ROSCOP systems can be used to obtain initial insights as to where missing data might be found.

Not all cruises result in ROSCOPs being filled out and sent in. Harry Dooley found that half of some data sets were not on ROSCOP forms. Nikolai Mikhailov mentioned earlier that he believes only 30-50% of Soviet Indian Ocean data have ROSCOPs. We must also search out cruise reports and other reports that describe data collecting cruises. Ron Moffatt mentioned earlier that India has never sent in ROSCOPs, but the National Institute of Oceanography of India publishes a newsletter which includes information (e.g., ships and cruises) found on ROSCOPS. Joe Reid searched and found French cruise reports for the Equatorial Pacific.

Warren White asked whether there were other ways besides ROSCOPs to identify data taken on cruises. Ron Moffatt noted that the historical yearly lists of the National Oceanographic Programs and Declared National Programs, published by IOC for cruises since 1960, represented an excellent source. These lists, by country, are maintained by WDC-A and are compared yearly against WDC-A's data accessions.

Statistics in the World Data Center A, Oceanography data exchange report should also provide some insight, as would similar reports from other data centers. The World Data Center Catalogue of Data provides another excellent source of information on known data sets. (Complete automation on a yearly basis of a common catalog for the three WDC's would be of great benefit to this effort, but may not be ready for some time.)

A comprehensive bibliographic search using on-line bibliographic services (which the

NOAA library does on a regular basis) should turn up refereed and grey literature papers that make use of data sets that we may not be aware of. WDC-A also maintains an archive of refereed and grey literature, which are summarized biannually in its Catalogue of Accessioned Publications.

It was also felt that letters could be sent out to the Directors of all the world's oceanographic data centers and data collecting institutions, emphasizing the importance of climate and global change research, and asking for their help in finding historical global oceanographic data. Such a letter might come from the Administrator of NOAA. It might accompany the data distribution maps to be sent out by NODC, or be sent out before those maps are completed. If the NOAA Climate and Global Change Program funds the Oceanographic Data Archaeology Proposal and these letters go out after that, they could mention the possibility of funding support for locating and digitizing such historical data.

As will be discussed in Section 11 (dealing with communication with researchers and data managers), a frequently published Ocean Data Archaeology Newsletter will help to maintain an awareness by the scientific and data management communities of the data archaeology effort and could provide the kinds of information that will stimulate the discovery of unarchived historical data sets (e.g., data distribution maps, lists of known manuscript and analog data, rumors and leads that need substantiation, etc.).

Other techniques for finding out about data include: supporting archaeological activities at key data centers and institutions (at least one for each key region of the globe), supporting expanded visiting scientist programs and joint international research projects, attending key international data management and scientific meetings, using electronic bulletin boards, writing articles for appropriate trade magazines and newsletters, etc.

An important part of this integrated approach will be to keep records of all search activities, successes, failures, leads, etc.

8. ACQUIRING HISTORICAL DATA AND DOCUMENTATION

Once the existence of a particular data set is known the next and often most difficult step is to convince the institution or scientist to provide the data (and metadata/ documentation), hopefully in digitized form. A number of methods will utilized.

It may take some type of monetary support to convince an institution to put in the effort required take to "dig out" historical data and all accompanying documentation, and to digitize and quality control them. Such support might be in the form of ADP equipment.

Another method would be an enhanced visiting scientist program, where the scientists bring their data with them to work on, or some type of a joint research project, even if the scientists stayed at their home institutions. The scientific quality control of the data resulting from their use in a research project is then a fundamental part of the process, in addition to the standardized quality control algorithms to be used in the routine data processing procedures.

Another method would be to make trips to institutions holding valuable data. Such trips could include the provision of training in quality control techniques and format conversion. Joint research projects could be developed with these institutions.

Another activity which might lead to acquiring more data in digitized form would be the development of a PC software program to make data entry, format conversion, and quality control easier and more reliable. Such a program could carry out the following functions: (1) easy, double-entry, quality-controlled data entry, (2) easy metadata/ documentation entry (answering the questions asked by the program), (3) easy format conversion (to and from: GF-3, BUFR, formats of NODCs and regional data centers, and formats specified by the user), (4) scientifically-based quality control algorithms and flagging of data, and (5) graphics and statistical analyses. The idea would be to have a program useful enough to convince the scientist or data manager to use it. This is the same concept that led to the CLICOM program run by the World Climate Data Office of WMO. That system has become quite large and relies on commercial packages such as DATAEASE, which are revised on occasion, causing potential compatibility problems. It is suggested here that the key functions (listed above) can be carried out with a single modest PC program with well structured, documented, maintainable code. IOC has recently begun looking into the possibility of producing a similar type of program (which they call "OCEAN PC").

The best method of actually acquiring newly discovered historical data will vary with each situation. The strengths and needs of each potential cooperating institution will be assessed (e.g, manpower and funding resources or lack thereof, ADP equipment, etc.) to decide on the best approach. Cooperative data archaeology projects could be developed as a result.

9. DATA DIGITIZATION

There are numerous manuscript data sets residing in the World Data Centers and other archives. These will probably be added to by newly acquired data from institutions willing to provide data but not digitize them. Digitization of these data will be an important part of the data archaeology effort, and may rely on optical scanners.

As mentioned earlier, we will need to inventory the data available only in manuscript form in the world's archives, making use of the World Data Center Catalogue of Data and other sources. The next step will be to identify those manuscript data sets that have been digitized since they were originally sent in, find out whether there is additional documentation, and find out whether the institution would be willing to digitize the data if provided some type of support (monetary or ADP).

Digitization will require support for the use of optical scanners to digitize the data, the use of manual data entry (perhaps as part of the computer program described above), the use of outside contractors, and some type of support (financial or ADP) for institutions with data requiring digitization.

The question was raised whether the data should be digitized at the institution where the data were found, at a regional center, or at the ultimate archives. It would be a very large job to have all the digitization done at the ultimate archives, which already have formidable backlogs of manuscript data to be digitized. On the other hand, it was suggested that the most important thing is to obtain the data, before they are lost forever. Waiting for data to be digitized at the home institution might in some cases jeopardize acquisition of the data or at least delay it substantially. It was suggested that digitization at the institution where the data were found, by those who know the data best, would be preferable. However, one worry expressed was that corrections might be made to the data that were not documented; and, in fact, there may no longer be anyone at an institution who was involved in collecting the older data. This worry might be lessened if we supplied a program (like that mentioned above) for each institution to use, with accepted QC algorithms and flagging procedures. Wherever the data is digitized, some immediate QC should be done to catch digitization errors (and manual digitization should include double entry). The digitized data would be QC'd again at the regional data center or the ultimate archive. It may come down to needing to offer funding support for digitization in order to get the data at all. It was agreed by all that it is critical to obtain all metadata that relate to the data.

Another important question will be: which data sets merit digitization, especially if our resources are limited? We will have some idea from the geographic area and time period of the data whether the data are important for climate and global change purposes, but we may have little idea about the quality of the data. This supports the idea of getting the data in manuscript form first, examining the data, and then giving the go ahead to whomever is chosen to digitize it; special instructions could also be given to the digitizers based on this examination of the data.

The overriding issue may be the magnitude of the digitization task, which would suggest spreading it around as much as possible, or it will not get done. We also may not get some of these data without financial or other support being provided to the institution

that has the data. This support could be provided to do the digitization at institutions which have the data, or to regional centers which might collect the data for a particular region.

The question of priorities as determined by the needs of the researchers was discussed earlier in the workshop, but this subject was raised again in relation to determining some guidelines for making decisions on which data were most important to digitize quickly and which data could be given a lower priority. A few participants felt that there would be certain geographical areas and time periods that might be considered more important. Others felt that higher priority should be given to large data sets or data sets at risk. Others felt that quality of the data was the most important consideration; there may be some discovered data sets that clearly will not be worth digitizing because of poor quality. All agreed that the data cards found by Harry Dooley in the ICES cellar should be given high priority because they come from an important area that had little data. It was also agreed that the MBTs at Scripps were important and should be given high priority.

But it was also agreed that priorities will be constantly changing as new data sets are discovered. Even within large data sets there may be some data that are more important than others. Resources will always be limited. There should be some mechanism for reviewing and prioritizing the manuscript and analog data sets that are candidates for digitization, perhaps an informal committee. It was agreed that the first step is to itemize known data sets that are candidates for digitization (i.e., the inventory noted earlier; see Section 6). The next step is to provide this inventory to pertinent institutions and data centers to see if the data have been digitized. This inventory should also be provided to those researchers who might be interested in this data type, to see which data sets they might need. Communication (discussed later in the workshop; see Section 11) is an important part of this, and an Ocean Data Archaeology Newsletter could be used to keep everyone updated on the status of these data sets.

There were discussions about the workshop participant's experience with optical scanning. This seems to be a very promising way to digitize manuscript data, but more work needs to be done. Software is the key, both the software supplied by manufacturers that allows the scanner to "learn" to recognize particular characters, and the software that converts the resulting screen of data to the appropriate format. Scanning would appear to be most efficient when dealing with many manuscript sheets having the same format. Hand entry (with software allowing double-entry QC) may still prove to be most efficient for many cases. It was felt that QC should be done as the data are digitized, since some of the problems could result from the digitization.

10. QUALITY CONTROL

Quality control (QC) is a basic issue whether one is dealing with data archaeology or simply the processing of recent data. It becomes more difficult, however, dealing with old data, which may have incomplete documentation or may have been produced using old instrumentation with unknown calibration. When looking for changes in the ocean, as supposedly indicated by differences between historical data and more recent data, the added problems of different measurement techniques, different procedures, calibration differences, different digitizing techniques, etc. can cause doubts in the results.

It will be helpful to the entire quality control process to produce a compilation of all instrumentation and measurement techniques used over the years, including known data problems.

It was also suggested that it would be helpful to have a compilation of the various QC procedures being used today. Data centers and other data collecting institutions could exchange QC techniques. This project could serve as a reason for an assessment of the various QC algorithms in use today (for a particular data type) and agreement on some standardized QC algorithms that could be provided to those institutions digitizing manuscript data (and could be included in the PC program mentioned above). MEDS in Canada has recently produced a QC Manual for the IGOSS-IODE supported GTSPP project. Additionally, the three WOCE hydrographic data assembly centers (at Scripps, AOML, and CSIRO) have also prepared a QC manual. These efforts will likely serve this purpose, and it would probably not be wise to use our limited resources to repeat their efforts.

Although the methods have not been standardized, the QC procedures for temperature and salinity profiles are probably reasonably good. The real need for QC techniques is with nutrient data. It is well known that there have been systematic errors in some nutrient measurements in the past. There may also be more variability than in temperature and salinity data. Syd Levitus mentioned that there can be large vertical gradients because of blooms and other phenomena.

It was felt that there needed to be a compilation of descriptions of methods used with nutrient data. Climatologies also need to be produced. (Syd Levitus has produced some initial world ocean nutrient climatologies using Joe Reid's version of the NODC nutrient data set, which show reasonably smooth fields.)

The quality control of the nutrient data is a large undertaking and will take time. There needs to be an exchange of information on QC techniques used with nutrient data and a compilation of those techniques. Harry Dooley stated that ICES does a mapping of nutrient distributions onto temperature and salinity diagrams (based on the assumption that in the deep ocean nutrients are conservative) and that seems to work quite well. (Joe Reid mentioned that one only has to assume weak lateral variability.) In fact, ICES can separate out anomalous or noisy data sets. They have found, for example, that the Bedford Institute of Oceanography (in Canada) provided the best nutrient data in the Atlantic, whereas data from some of the European labs had more variability. He also cited one example to show the problems that can occur with taking nutrient data. ICES received nutrient data from several institutes taken during a 12-ship survey in the Baltic Sea; it turned out that the total

variability from one institute to the other in the same water mass, using the same techniques, was 52%.

Scientific use and assessment of data is obviously the best means of quality control. Thus, visiting scientist programs and joint archaeology projects become important parts of quality control, as does a mechanism to receive feedback from all scientists working with the data.

11. COMMUNICATION WITH RESEARCHERS AND DATA MANAGERS

It is obviously important to make newly acquired historical data available to the scientists who need them as soon as reasonably possible. Because the project should be driven by the needs of the climate and global change researchers, communication with them regarding recent data acquisitions (or discovered data sets not yet obtained) and obtaining their feedback on quality issues is important and should occur on a regular basis.

It was agreed that a frequently published Ocean Data Archaeology Newsletter would serve many purposes and would be important to produce. It would help to make and keep the scientific and data management communities aware of data archaeology efforts. It could provide the kinds of information that will stimulate the discovery of unarchived historical data sets. For example, it could include: data distribution maps, lists of known manuscript and analog data, updates on newly discovered data, rumors and leads that need substantiation, articles about archaeology and research projects at various institutions, and even interesting historical articles relating to cruises and data.). A telemail mailbox (or possibly a bulletin board) would allow for rapid communication on archaeology issues and potential data sets.

Further into this project it would be useful to hold a conference/workshop devoted to a particular data type or types. Such a conference would include presentation of scientific papers making use of these data, but would also include working group discussions on quality control issues, potential sources of other historical data, and priorities. It would also lead to increased communication among the scientists using the data.

APPENDIX I

List of Participants

PARTICIPANTS

Oceanographic Data Archaeology Workshop

September 13-14, 1990

U.S. National Oceanographic Data Center
Washington, D.C. USA

Arkady V. Alekseev
Chief Deputy Director
Pacific Oceanoological Institute
Far Eastern Branch
USSR Academy of Sciences
7, Radio Street
Vladivostok 690032 USSR

phone: 96-6-13
telex: 871 213121 SVTSU

Margarita Conkright
U.S. National Oceanographic Data Center
E/OC13, NESDIS, NOAA
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5643
fax: (202) 673-5686
telemail: NODC.WDCA (Omnet)

Rene Cuzon du Rest
World Data Center A, Oceanography
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5583
fax: (202) 673-5586
telemail: NODC.WDCA (Omnet)

Harry Dooley
ICES Hydrographer
International Council for the
Exploration of the Sea
Palaegade 2-4
DK-1261 Copenhagen K
Denmark

phone: 33 93 42 14
telex: 22498 ICES DK
telemail: ICES.DK (Omnet)

Robert Gelfeld
World Data Center A, Oceanography
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5583
fax: (202) 673-5586
telemail: NODC.WDCA (Omnet)

Doug Hamilton
U.S. National Oceanographic Data Center
E/OC13, NESDIS, NOAA
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5643
fax: (202) 673-5686
telemail: NODC.WDCA (Omnet)

Pembroke J. Hart
WDC-A Coordination Office
National Academy of Sciences, HA-40
2101 Constitution Avenue
Washington, D.C. 20418 USA

phone: (202) 334-3368
telex: 248664 NASW UR
telemail: NRD.GEOPHYSICS/OMNET

Gary Hopwood
Australian Oceanographic Data Centre
P.O. Box 1332
North Sydney NSW 2090
Australia

phone: (02) 925 48 70
fax: (02) 925 48 35
telemail: B.SEARLE

Sydney Levitus
U.S. National Oceanographic Data Center
E/OC23, NESDIS, NOAA
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5411
fax: (202) 673-5586
telemail: S.LEVITUS (Omnet)

Nikolai Mikhailov
Chief of Laboratory
All-Union Research Institute of
Hydrometeorological Information
- World Data Centre
USSR State Committee for Hydrometeorology
6, Korolyov Street
Obninsk, Kaluga Region
249020 USSR

phone: 2-56-78
telex: 412633 INFOR SU

Ronald Moffatt
Associate Director
World Data Center A, Oceanography
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5583
fax: (202) 673-5586
telemail: NODC.WDCA

Kee-Soo Nam
Head, Dept. of Data Management
& Computer Services
Korea Ocean Research & Development Institute
Korea Institute of Science
& Technology
Ansan P.O. Box 29
Seoul, 425-600 Korea

phone: (02) 863-4770 (ext.501)
fax: (0345) 82-6698
telex: KORDI K27675

Bruce B. Parker
Director, WDC-A, Oceanography
and Chief, Information Services Div.
U.S. National Oceanographic Data Center
E/OC2, NESDIS, NOAA
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5411
fax: (202) 673-5586
telex: 7401815 NODC UC
telemail: B.PARKER/OMNET

Steve Patterson
U.S. National Oceanographic Data Center
E/OC23, NESDIS, NOAA
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5411
fax: (202) 673-5586
telemail: NODC.WDCA (Omnet)

Joseph Reid
Scripps Institution of Oceanography
MLR Group, A-030
La Jolla, CA 92093

phone: (619) 534-2055
telemail: J.REID (Omnet)

Ricardo L. Rojas
Departamento de Oceanografia
Instituto Hidrografico de la Armada
Errazuriz 232 - Casilla 324
Valparaiso, Chile

phone: 282704 - 282697
fax: 56-32-283537
telex: 230362 HIDRO CL
telemail: TOGA.CHILE

Ben Searle
Head, Science & Oceanography
Australian Oceanographic Data Centre
P.O. Box 1332
North Sydney NSW 2090
Australia

phone: (02) 925 48 70
fax: (02) 925 48 35
telemail: B.SEARLE

Shin Tani
Chief, International Programmes Group
Japan Oceanographic Data Center (JODC)
5-3-1, Tsukiji 5-chrome,
Chuo-Ku, Tokyo 104
Japan

phone: +81-3-545-3811
Ext. 732
fax: +81-3-545-2885
telex: 252 2452 HD JODC J
telemail: T.MORI (Omnet)

Peter Topoly
U.S. National Oceanographic Data Center
E/OCx4, NESDIS, NOAA
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5591
fax: (202) 673-5586
telemail: NODC.WDCA (Omnet)

Godfrey Trammell
World Data Center A, Oceanography
1825 Connecticut Avenue, NW
Washington, D.C. 20235 USA

phone: (202) 673-5571
fax: (202) 673-5586
telemail: NODC.WDCA (Omnet)

Warren White
Scripps Institution of Oceanography
A-030, UCSD
La Jolla, CA 92093

phone: (619) 534-4826
telemail: W.WHITE (Omnet)

APPENDIX II

Agenda of Workshop

AGENDA

Oceanographic Data Archeology Workshop

U.S. National Oceanographic Data Center
Washington, D.C.

September 13-14, 1990

Welcome and Introductions

Overview of Workshop

Agenda and Time Table
Final Report and Plan
Dinner

Background and Overview of Oceanographic Data Archaeology Project

Definition of "Archaeology"

The Ocean Data Needs of Climate and Global Change (C&GC) Scientists

Types of data needed most
Selecting initial data types for archaeology project
Critical gaps in data (geographical and temporal)
Obtaining the priorities of additional researchers
    Lists of key scientists and data managers
    Mechanisms for obtaining priorities

Summarizing Known Data Sets

Should we develop a comprehensive inventory of **archived digital data sets** around the world?
    (A big job, but might find out about a digitized data set we didn't know about)
Or, should we concentrate initially on inventorying **data not yet in an archive in digital form** that we know about (e.g. data in manuscript or analog form)?
How important is this activity (relative to other activities in plan)?
Best mechanism(s) for producing and providing inventory; benefits of being easily distributed in computer compatible form or hardcopy maps?
Sources of information.

Discussion on data sets known to the workshop participants that should be obtained; produce summary report of initial findings.

Discovering Data Not In Any Archives (Mechanisms)

ROSCOPs -- automated inventory
World Data Centers catalogues
Grey Literature and Journal Bibliographic Database
Supporting archaeology activities at various data centers and institutions
Enquiries to institutions around the world; basic detective work -- following up leads
Visiting scientist programs
Attending key scientific and data management meetings
Newsletter and electronic bulletin boards
Systematic record (and/or database) of all findings

Acquiring the Data and Documentation

Inducements to provide data and documentation
    Monetary or ADP Hardware support
    Useful PC-program (data entry, format conversion, Quality Control, etc.)
    Visiting scientist programs and cooperative projects
    Visits to institutions with data (including training)
Supporting regional archaeology activities at various data centers and institutions
Who should digitize the data and documentation?

Digitization

The role of optical scanners versus quality controlled hand entry
Digitizing the manuscript and analog data we already know about (its importance relative to other activities in the plan)
Who should do it? Who can quality control it the best?
Supporting regional archaeology activities at various data centers and institutions

Quality Control Issues

Added problems with historical data
Compilation of instrumentation and measurement techniques (and typical problems)
Standardized QC algorithms for particular data types (Would be included in above PC program, if that program considered useful. Use White-Meyer-Molinari document, MEDS-produced GTSPP report, Levitus techniques?)
Should this be a part of this archaeology effort?
QC by scientific use of data

Communication Among Researchers and Data Managers

Newsletter
Electronic Bulletin Board
Specialized Conferences/Workshops

Consensus on Most Important Archaeology Activities and Overall Approach

Consensus on Initial Data Types to Obtain and Initial Geographical and Temporal Priorities

Consensus on Most Efficient Use of Resources

Follow-ups to Workshop

Sent to Workshop Participants:

Report on Workshop
Final Draft of Plan for Integrated Approach to Oceanographic Data Archaeology
Summary Report of known data sets that should be obtained and/or digitized

Received from Workshop Participants:

Final comments on Report and Plan
Additional information from participants about data holdings, and data that should be obtained and/or digitized

Adjourn Workshop

APPENDIX III

ICES Report Submitted to the Workshop

INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA

An overview of the ICES Oceanographic Data Bank

For consideration by the
Ocean Data Archaeology Workshop

September 13-14, 1990

U.S. National Oceanographic Data Center
Washington, DC

OCEANOGRAPHIC DATA ARCHAEOLOGY PROJECT (ODAP)

The ICES (pre-1957) Oceanographic Data Bank

Historical Perspective

In 1902, ICES commenced publication of *Bulletin Hydrographique*. This publication, which appeared annually until 1956 (apart from the war years), lists data from hydrographic stations in the North and Baltic Seas, and the northeast Atlantic, including waters around Greenland. From the beginning, these data were accompanied by details of methods and analysis, as well as track charts and distribution maps. Although the primary concern was temperature and salinity, *Bulletin Hydrographique* also contained listings of chemical and biological (plankton) data. In the beginning all these data were listed by country and date, but in 1936 the data were geosorted by date of measurement.

From 1921 to 1935 ICES also published data in the reports of its Comité du Plateau Continental Atlantique (which was translated to Atlantic Slope Committee) in publications entitled *Rapport Atlantique* of the *Rapports et Procès Verbaux* series. These data were supplementary to those published in *Bulletin Hydrographique* (from 1932-34 they were published in both volumes) and covered the continental shelf and deep water areas to the south and west of UK, Ireland, France and Iberia. Although the interests of the Atlantic Slope Committee were primarily related to fisheries questions, this series of reports additionally provided detailed scientific interpretation of the hydrographic data published therein.

In the mid-1930s ICES decided to "automate" its data collections and extend their range to the entire North Atlantic with the creation of the Hydrographic Card Index which is still preserved in the archives of the ICES Secretariat. In 1956 when this particular activity stopped along with the publication of ***Bulletin Hydrographique*** there were 200,000 such cards, also geosorted by month. This Card Index served as an aid to arranging the entries in "***Bulletin Hydrographique***", but the former contained considerably more data than that published in the Bulletin. During the establishment of the Card Index, all publications were scoured for data lists and their origin noted on the back of each card entry, if different from the Bulletin. Thus the card index additionally contained the data that was published in ***Rapport Atlantique*** as well as a number of non-ICES publications. The publications used in the transfer of data to the Card Index, along with handwritten notes cross referring to the Card Index are still present in the ICES library. Further details of these non-ICES publications are given below in the discussion of the data bank contents.

From 1957, data sets submitted to ICES were prepared on punch cards, formatted using the ICES punch card format which is still in use today. These data, which covered a similar area to that published in "***Bulletin Hydrographique***" were printed in a publication called "***ICES Oceanographic Data Lists" (IODL)*** but this publication was ceased in 1962, apparently because data volumes were becoming unmanageable. In its place an inventory of cruises was published, a precursor to the present

ROSCOP form. Following the formation of a number of National Data Centres, member countries were no longer obligated to submitting data to ICES directly.

ICES data holdings on Computer

Following the introduction of punch cards in the mid 1950s, ICES data holdings, including those acquired during the first half of the century, were transferred to computer. This transfer was done in cooperation with NODC Washington, who supplied ICES with geosorted tapes created from the punch card holdings. For various reasons the job of identifying these data, and resorting them into ICES's data management structure of data files by country, ship and year is still under way. The present document is a first attempt to review these data holdings for the years prior to 1956, ie during the period when the Hydrographic Card Index provided the backbone to the data bank.

A summary overview of the Data Bank, by period and country is presented in Table 1. Since the objective of this table is to identify any major gaps in data, information is limited to the years for which data have been identified in the data bank, for each country.

In fact large gaps in the data bank have been uncomfortably easy to find, which points to the need for caution in interpreting the contents of this paper. In may well be that the review of the data holdings, which are still under way, has not yet uncovered large sections of it that may have been prepared in the 50s, or the re-transfer of data from WDCA was not completed. Thus the gaps identified here may be, hopefully, of data already in WDCA.

Table 1 shows the distribution of data in three periods, viz the period of publication of *Bulletin Hydrographique* (also *Rapport Atlantique* and the Hydrographic Card Index), the publication of *ICES Oceanographic Data Lists* and the period since 1963 when no data has been published. The remaining discussion in this paper concerns only the first period, i.e., 1902-1956.

Gaps

Using Table 1 several gaps in the Data Bank were readily identified. In particular the pre-1956 data bank contains:

- all data published in the *Bulletin Hydrographique*
- No non *Bulletin Hydrographique* data published in *Rapport Atlantique* and included in the Hydrographic Card Index can be identified in the Data Bank.

This lack of data from the *Rapport Atlantique* means that the ICES data bank lacks significant data banks from a number of countries who are currently not well represented in the Data Bank, viz Ireland, France, Spain and Portugal. Furthermore the lack of data means that ICES has a serious shortfall in coverage in an important area of the North

Atlantic, i.e., its eastern margin.

Table 2 summarizes the data reported in *Rapports Atlantique*. In the period 1921-1934, 5900 stations were reported. All of these are also reported in the Hydrographic Card Index and only those collected in the period 1932-34 are in the data bank when they were also published in *Bulletin Hydrographique*.

There are a number of other large data sets not presently identified in the ICES data bank but which were included in the Hydrographic Card Index. These, and some others, will be described in the following Country-wise review of the ICES Data Bank. In compiling this review, further help was acquired from publication (1), which included a compilation of mean values by 1° Marsden Square in the North Atlantic in the period 1868 to 1953. This publication used *Bulletin Hydrographique* as a basis, but additionally data were extensively sought and used in the calculation of the means. The sources of these additional data were published in a number of Tables in that paper, one of which is reproduced here as Annex 1.

Country-wise review

Belgium

No data exist for long periods, notably 1914-1924, and 1939-1957. It is not known whether data do exist in these periods. A small amount of Belgian profile data (Southern Ocean) has been located in (2).

Canada

There were no contributions to ICES data publications.

Denmark

As National Data Centre, ICES has a responsibility to maintain full data sets. Extensive data sets are maintained in the publications (3 & 4) as well as the data from the Ingolf Expedition in 1896. Although these particular data sets are not included in the Hydrographic Card Index, it may be that they are included in a supplementary card index stored in the cellar at Charlottenlund Slot, the former home of the Service Hydrographique.

Germany

This data set, has been re-compiled recently in cooperation with DOD. All data in ICES for the pre-57 period is from north of 50°N. Thus the results of the Altair/-Armauer Hansen cruise published in (5) are absent. Whilst the results of the latter ship are in the ICES data bank, there is no trace of the former (there is not even an ICES ship code for the Altair! The 900 or so surface t,S data from the stations worked during the 1925-1927 German Atlantic Expedition have been published in Band IV of (1) -the chemical data are in Band VI. There is no evidence of any of these data in any ICES

publication, nor are they in the Hydrographic Card Index. These two examples illustrate that the ICES data holdings may fall far short of the actual amount collected.

Finland

It is believed that this data set, comprising some 6000 stations, is essentially complete. Table 1 does indicate some gaps, e.g. 1937-1953 and 1955-1956. Substantial data sets for most of these "missing" years do exist in printed form in publication series (6). Data are, however, primarily from the Baltic Sea, and some from the Barents Sea.

France

This is one of the most incomplete of data sets in the ICES data bank. However there are substantial data sets in printed form, many of which are shown in Table 2 and Annex 1. Many of these data are included in the Hydrographic Card Index, and are used in the work of Krauss (1). The most substantial of these data are those of the cruises of "Pourquoi Pas?" which have been published in part in *Rapports Atlantique* and in there entirety in the serial publication (7).

Iceland

There are apparently no gaps in the Icelandic data sets. The period 1950-56 has not been located in the ICES data bank in spite of the fact that *Bulletin Hydrographique* includes data for this period. Note also that Annex 1 item 114 refers to data from the Maria Julia in the early 1950s. These data are reported in *Bulletin Hydrographique*

Ireland

Since few data were reported in *Bulletin Hydrographique*, not many data sets have found there way to the data bank, but most are in the Card Index as a result of their publication in *Rapport Atlantique*. There is no trace of data collected subsequent to the last war.

Netherlands

There are a few gaps surrounding the war years. However the gap 1914-1930 points to the absence of data from the Snellius Expedition 1929-1930. Data have however been transferred from Vol 1 of the Expedition Proceedings (8) to the Hydrographic Card Index. Other gaps include time series from a number of Dutch light vessels but these are believed to be in the safe keeping of a Dutch Institute.

Norway

There are 28199 hydrographic stations in the ICES data bank for the period 1900-1967. Table 1 does suggest that holdings are more or less complete for the North Atlantic, probably because the Norwegian data centre sought out non - *Bulletin Hydrographique* data some years ago. However it has to be pointed out that the extensive quality control undertaken over the years by ICES has not been transferred to either punch card or computer, either by ICES or the Norwegian data centre. Whether this omission is applicable to other data sets is currently not known. The data from "Morild" (see comments under Sweden) have been identified in the Data bank except those for 1926. In this year all station positions were given as bearings, which may have discouraged data entry (the same applies to a number of other stations in different years). One other data set known not to be included is listed in publication (9).

Poland

The one non-war gap, i.e 1949-1951 (Table 1) points to incomplete holdings. Indeed there is a publication (in Polish) by Wanda Piacek in the 1959 issue of the Sea Fisheries Institute journal, which makes use of data from Polish Hydrographic Stations (9 of them worked several times) during this missing period

Portugal

Publication (1) supports the view that the main source of data prior to the late 1930s was the "Albacora" whose data were published in *Rapport Atlantique*. However it appears that not all "Albacora" data were published here, for example its 500 or so stations worked in 1937 were published in (10). Only occasional copies of this publication series are available in ICES, and it may be that many more data are published than the impression given here. (In 1971 the Lisbon Institute of Hydrography started publication of data in its technical report series).

Spain

Similar comments to above, except main ships concerned included Admirante Lobo, Xaeun and Eduardo Duado amongst others.

Sweden

There are no known gaps. Publication (1) refers to data from the "Skagerak". This appears to be an error, the hydrographic observations were in fact undertaken by the Norwegian ship "Morild" - see comments under Norway.

UK

Table 1 indicates a fairly complete data set. This will certainly be true for the UK fisheries institutes, but the completeness of from other institutes cannot be verified easily. This is because UK data sets were not routinely published (except for the Discovery Reports which refer to the Southern Ocean, and the Challenger Expedition). Table 2 shows the existence of extensive listing from "Salpa", a ship that worked out of an institute in Plymouth. Apart from these data, ICES has not received any other data from Plymouth in spite of extensive activities by the relevant institute over a long period of years.

USA

There are many entries in the Hydrographic Card Index, but no data resides in the data bank.

USSR

There are a number of substantial gaps. Present holdings for this period refer mainly to Baltic investigations. There are some known gaps from Latvia referred to in the 1938 *Bulletin Hydrographique*

1. This is a 15 volume series based on the 1925-27 "Meteor" Expedition to the Atlantic (D.A.E.). All hydrochemistry data for the expedition are published here as well as several detailed inventories of other cruises (see e.g., Annex 1). Related studies included calculations of mean temperature and salinity of the North Atlantic during the 86 year period 1868-1953, including a discussion on the relation between coastal station data and open sea data. *Bulletin Hydrographique* is mentioned as a source of data.

Deutschen Alantischen Expedition Auf dem ForschungsundVermessungsschiff "Meteor" 1925-1927, 1936, Berlin

2. A few stations off Antarctica given temperature profiles in an expedition held in 1897-99.

Expedition Antarctique Belge. Resultats du SY Belgica en 1897-1898-1899, Anvers 1908.

3. Describes a series of cruises by "Thor" to the Mediterranean in 1908-1910, and also to the Atlantic south of Iceland in 1905 and 1906. Supplementary stations by various vessels to the Mediterranean and the Atlantic are also given. Contains a reference number to the Card Index (C.I.D 34)

Report on the Danish Oceanographical Expeditions 1908-1910, Copenhagen 1912.

4. Describes the Dana Expeditions 1920-1922 to the North Atlantic and Gulf of Panama. There is a note saying it was card-indexed on 8 Dec 1936. Lists Station numbers from 1100 to 1387.

The Danish "Dana"-Expeditions 1920-22, Oceanographical reports edited by the Dana Committee, No 1., Copenhagen 1929.

5. Description of the results of the Altair/Armauer Hansen Cruise to the northeast Atlantic in May-June 1938. There are 102 stations.

Annalen de Hydrographie, 1941. The stations of the Altair by Wust.

6. Series of publications from the Finnish Institute for Marine Research. As well as publishing data from cruises, data from coastal stations are regularly listed.

Merentutkimsdlaitoksen Julkaisu Havsforskningsinstitutets Skrift.

7. Series of publications containing data ca 1912-1938. Features mainly the cruises of the "Pourquoi Pas?" some of which are duplicated in *Rapports Atlantique*.

Annales Hydrographiques, Service Hydrographique da la Marine, Impremerie Nationale, Paris

8. This is in several volumes, volume 1 of which contains listings of t,S data. It describes the results of a voyage of the "Snellius" from Holland to the Dutch East Indies from 1929 to 1930.

The Snellius Expedition in the eastern part of the Netherlands east Indies, 1929-1930. Leiden, 1937.

9. This "treatise" gives the general results on the Hydrography and pelagic life of the Northern Ocean based on observations collected at the end of the 19th century. It list data from a number of sections, and presents beautiful coloured pictures of the sections.

Report on the Norwegian Marine Investigations, 1895-1897 by Johan Hjort et al. Bergens Museum, Bergen 1899.

10 One of a number of reports that may contain full data listings from "Albacora"

Travaux de la Station de Biologie Maritime de Lisbonne, No 43, 1938 Lisbon.

Table 1 - The ICES Oceanographic Data Centre - Location of Data holdings by Country - 1902 to present				
Country	1902-1956 (BH)	1957-1962 (IODL)	1963-	Remarks (post 1963 data only)
Belgium	03-13,25,30-38	57-62	63-65, 69-85	Prior to mid-70s, mostly light vessel, latterly surface nutrient surveys adjacent to coast (77-).
Canada	-	-	63,73,76	Includes Overflow '73 data.
Denmark (incl Greenland excl Faroes)	02-13,22-25,27-56 (26 in BH)	57-62	63-90	Nearly complete and up-to-data. Mainly from North Sea/Baltic. Lack data from University (Inst for physical Oceanography).
German D.R		57	64,69,70,79-90	So called "National" Stations only ca 10% of that collected - Baltic only. Other data now being made available.
F.R Germany	02-14, 27-56	57-62	63-90	Mainly data from Fisheries and Hydrographic Institutes (North Sea/Baltic). Large quantity other data (mainly Atlantic) not available.
Finland	02-14, 27-36, 54	57-62	63-87	Only Baltic (from IMR Helsinki) (ca 6000 stations since 1902
France	30-38,47-55	57-58, 60-62	63-70, 75-81	Mostly small data sets (Atlantic) from individual scientist.
Iceland	34-39, 47-49 (50-56 in BH)	57-62	63-88	Complete data set from IMR, Reykjavik
Ireland	05-14, 32-39	-	-	Data do exist for more recent years - an unreadable tape submitted in mid 1970s.

Table 1 - The ICES Oceanographic Data Centre - Location of Data holdings by Country - 1902 to present				
Country	1902-1956 (BH)	1957-1962 (IODL)	1963-	Remarks (post 1963 data only)
Netherlands	02-14, 30-37, 48-50, 52-56	57-62	63-83, 86-88	Mostly light vessel. Also nutrient surveys adjacent to coast and 3 Atlantic/Med data sets 83, 86 (Saclant) and 87.
Norway	00-06, 10-14, 18-20, 22-56 (NODS holdings (some extra to BH)	57-62	63-85	North Sea only from 1968. Huge gaps.
Poland	28-30, 36-38, 46-48, 52-56	57-62	63-77, 79-85	Mainly Baltic/North Sea.
Portugal	34-37	-	-	Potentially significant data source
Spain	32-34	-	71-74, 76	In connection with CINECA only.
Sweden	02-14, 25, 27-56	57-62	63-87	Mostly Baltic and Kattegat. A number of source gaps (eg universities).
United Kingdom	02-14, 20-39, 46-49, (50-56)*	(57-62)*	63-89	Mostly fisheries labs data. Very large gaps related to University and Research Councils. * Information uncertain - data yet to be catalogued.
United States of America	-	Some punch cards	-	
USSR (incl Lithuania, Estonia and Latvia)	02-08, 30-39, 56	57-62	63-82	Mainly Atlantic. All submitted via WDCA

Table 2. Data listed in Rapport Atlantique (R.A.) of Rapports et Proces Verbaux (RPV)

R.A.	RPV	Country	Ship	Station Nos	Comments
1921	XXIX	France	Tanche	60	
			Pourquoi Pas	60	
			Pétrel	13	
			Canonière(s)	75	
		UK	Salpa	80	
		Ireland	Helga	27	for 1920
			Helga	70	
1922	XXXI	France	Tanche	97	
			Pourquoi Pas	23	
			Canonière(s)	86	
			Sentinelle	3	
		GB	Salpa	167	
		Ireland	Helga	33	
1923	XXXV	France	Tanche	84	
			Pourquoi Pas	11	
			Canonière(s)	24	
		GB	Salpa	125	
		Portugal	Cinco de Oetubro	78	
1924	XXXVII	France	Tanche	28	
			Pourquoi Pas	34	
		GB	Salpa	147	
		Ireland	Muirchu	58	
		Spain	Admirante Lobo	540	for 1923
1925	XL	France	Tadorne	29	
			Tanche	23	
			Pourquoi Pas	17	
			Canonières	36	
		GB	Salpa	118	
		Ireland	Muirchu	65	
		Spain	Marques de la Victoria	48	
		Portugal	Albacora	75	
1926	XLIV	France	Canonières	55	
			Estafette	13	

Table 2 (ctd)

R.A.	RPV	Country	Ship	Station Nos	Comments
			Pourquoi Pas	11	
			Tanche	103	
			Hébe	24	
		GB	Salpa	100	
		Ireland	Muirchu	34	
		Spain	Proserpina	32	
		Portugal	Albacora	41	
1927	LV	France	Canonière(s)	41	
			Tanche	92	
		Ireland	Muirchu	66	
		Spain	Eduardo Dato	64	
		Portugal	Albacora	71	
1928	LXII	France	Arras	12	
			Estafette/Sentinelle	45	
			Tanche	72	
			Pétrel	8	
		UK	Salpa	145	
		Ireland	Muirchu	58	
		Spain	Eduardo Dato	61	
		Portugal	Albacora	37	
1929	LXX	France	Sentinelle	33	
			Pourquoi Pas	23	
			Pétrel	18	
		UK	Salpa	95	
		Ireland	Muirchu	73	
		Spain	Xauen	40	
		Portugal	Albacora	62	
1930	LXXVII	France	Sentinelle	6	3 other ships mentioned
		UK	Salpa	119	
		Ireland	Muirchu	66	
		Spain	Xauen	88	
		Portugal	Albocora	0	Data published elsewhere?

Table 2 (ctd)

R.A.	RPV	Country	Ship	Station Nos	Comments
1931	LXXXIV	France	Estafette	95	Pourquoi Pas and Ville d'Ys mentioned
		UK	Salpa	60	
		Ireland	Muirchu	51	
1932	LXXXVII	France	Estafette	34	
			Notre Dame	21	
			Alexandrine	13	
			Capitaine Armand	21	
			Sentinelle	20	
			Excellent	9	
		UK	Salpa	27	In BH also
		Ireland	Muirchu	92	In BH also
		Spain	Xauen	86	In BH also
1933	XCIII	France	P.T Tissier	46	
			L'Ardent	50	Pourquoi Pas & Ville d'Ys also
		UK	Salpa	40	In BH also
		Ireland	Muirchu	76	In BH also
		Spain	Xauen	20	In BH also
1934	CIV	Spain	Xauen	12	8 in BH
1935	CIV	Table of cruises of Salpa, Muirchu, Quentin Roosevelt, Tissier, Pourquois Pas, Ailette, Ville d'Ys only			
1936	CIV	Only very general information given, including a report ona reunion of experts held in Plymouth in 1937.			
1937	CXI				
1938	CXI				

Annex 1 (from Ref(1))

Tabelle 39

Anzahl der Stationen im offenen Atlantischen Ozean mit Reihenmessungen des Salzgehaltes

(Stand Anfang 1934)

Abkürzung	Schiff	Jahr	Anzahl der Stationen, die für ein bzw. mehrere Niveaus der Horizontalkarten zwischen — und — m Werte des Salzgehaltes geliefert haben				
			200—1000 m	1250—2000 m	2500—4000 m	4500—5000 m	Bem.
A	»Acadia«	1915	28	—	—	—	
Ab	»Albatross«	1884—85	11	—	—	—	[1]
		1919—20	7	1	—	—	
Ah	»Armauer Hansen«	1913—14	76	68	11	—	
		1922	29	12	—	—	
Al	»Albacora«	1925—26	4	1	—	—	
		1927—29	47	12	—	—	
An	»Antarctic«	1901—02	10	—	—	—	
At	»Atlantis«	1931—32	205	171	100	3	[2]
Ba	»Bache«	1914	28	13	—	—	[3]
Bl	»Berlin«	1924	9	9	9	—	[4]
Cdo	»Cinco de Outubro«	1923	91	—	—	—	
Cg	»Carnegie«	1928	33	25	19	2	[5]
Ch	»Challenger«	1873—76	28	3	—	—	[6]
Cha	»Challenger«	1932	24	21	8	—	
Cn	»Chance«	1926	5	—	—	—	
D	»Deutschland«	1911—12	76	44	12	—	[7]
Da	»Dana«	1921—22	92	57	35	11	
		1925	6	4	—	—	
		1931	18	16	—	—	
		1932	17	4	—	—	
Di	»Discovery«	1926—27	81	31	19	3	[8]
Di*	»Discovery II«	1930—31	174	132	97	16	
Eo	»Eduardo Dato«	1927—28	61	—	—	—	
Ex	»Explorer«	1929—31	12	—	—	—	
F	»Frithjof«	1910	21	5	—	—	
Fk	»Falken«	1912	21	—	—	—	
Ga	»Gauß«	1901—03	17	12	1	—	[9]
G.Gr.	»General Greene«	1931—32	62	46	—	—	[10]
Gh	»Godthaab«	1928	26	19	16	—	
Gr	»Grampus«	1914—16	6	—	—	—	
Gs	»Goldseeker«	1906—07	91	—	—	—	
		1909	6	2	—	—	
		1910—11	6	2	—	—	
Hl	»Helga«	1905—13	81	—	—	—	
Hx	»Huxley«	1909	4	—	—	—	
I	»Ingolf«	1895—96	47	17	—	—	[11]
Ma	»Margrethe«	1913—14	81	4	1	—	[12]
Mo	»Modoc«	1925—26	—	—	—	—	[13]
Mö	»Möwe«	1911—12	41	18	—	—	[14]

[1]) Manche Werte zweifelhaft. [2]) Umkipptiefe durch thermometr. Tiefenmessung bestimmt. [3]) Einige Werte zweifelhaft. [4]) Einige Werte herausfallend (Serie vorzeitig gekippt?). [5]) Mit salinity-tester ermittelt, in größeren Tiefen systematische Abweichung. [6]) Aräometrisch bestimmt, neu reduziert, manche Werte herausfallend. [7]) Meist auf wahre Tiefe neu reduziert. [8]) Manche Werte zweifelhaft. [9]) Bordtitrierungen verwendet, einige herausfallend. [10]) Mit salinity-tester ermittelt. [11]) Salzgehalt mit Hydrographischen Tabellen neu ermittelt. Manche Werte zweifelhaft. [12]) Manche Werte herausfallend. [13]) Vgl. am Schluß unter Ice Patrol Service. [14]) Einige Titrierungen zweifelhaft.

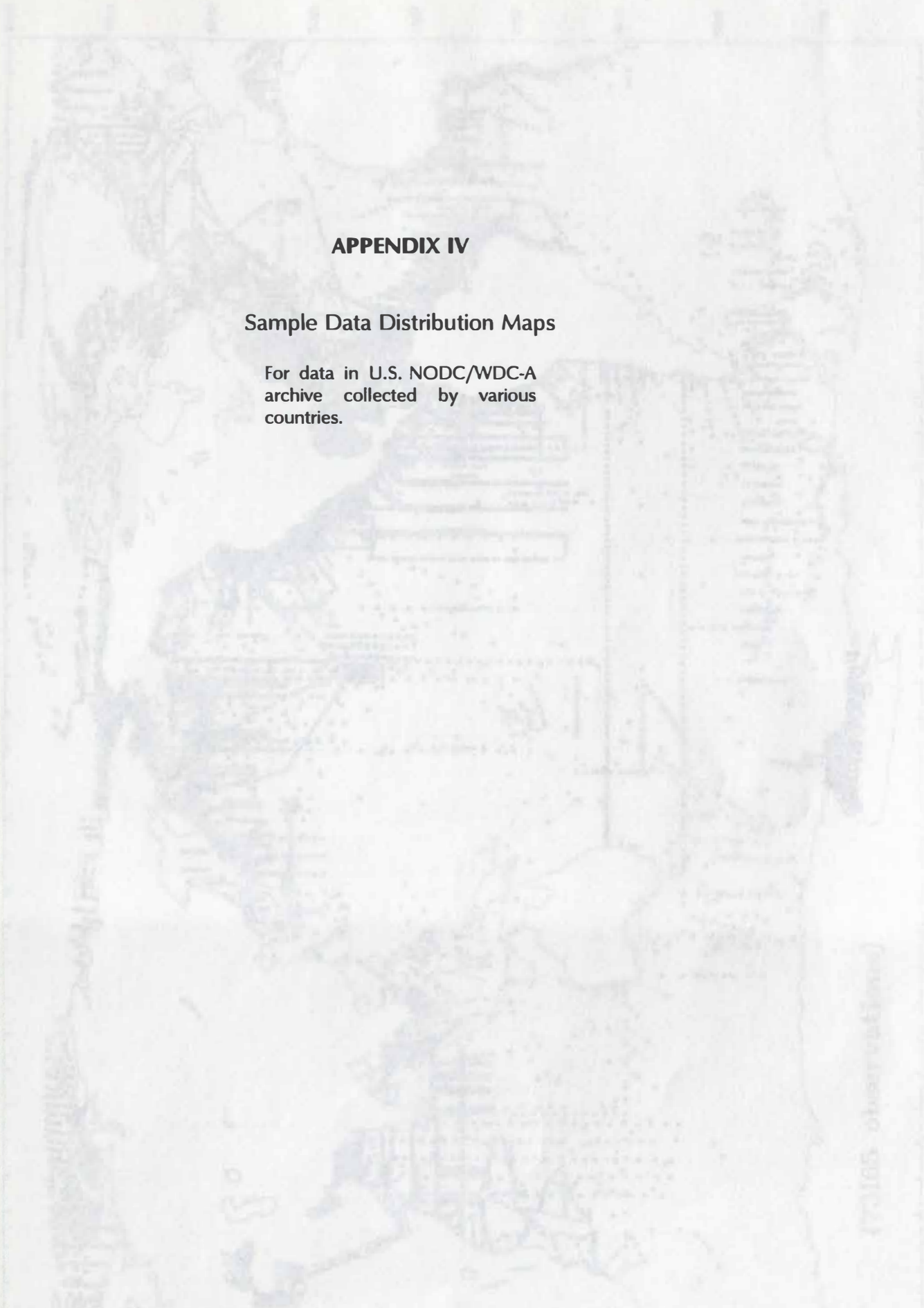
Abkürzung	Schiff	Jahr	Anzahl der Stationen, die für ein bzw. mehrere Niveaus der Horizontalkarten zwischen — und — m Werte des Salzgehaltes geliefert haben.				
			200—1000 m	1250—2000 m	2500—4000 m	4500—5000 m	Bem.
Mr	»Muirchu«	1925—30	8	—	—	—	
		1932	4	1	—	—	
Ms	»Michael Sars«	1900—02	1	—	—	—	
		1902	11	—	—	—	
		1910	51	28	4	3	
		1924	5	2	—	—	
Mt	»Meteor« D. A. E.	1925—27	284	246	226	72	1)
	»Meteor« Grönlandfahrten	1929—30	47	29	21	—	
Mv	»Marques de la Victoria«	1925	38	—	—	—	
Nr	»Norvegia«	1927—31	38	13	—	—	
Pb	Portugiesische Bewachungsschiffe	1910—11	1	—	—	—	
Pl	»Planet«	1906	21	9	3	—	
Pn	»Proserpina«	1926	20	—	—	—	
Pol	»Polaris«	1932	5	—	—	—	
Pp	»Pourquoi-pas?«	1909	1	—	—	—	2)
		1912—13	11	1	—	—	2)
		1921—22	4	—	—	—	2)
Pr	»Princesse Alice«	1902—03	6	4	2	—	3)
		1904	4	4	1	—	3)
Ry	»Rosemary«	1929—30	40	12	3	—	4)
Sc	»Scotia«	1903—04	10	7	6	—	5)
		1913	9	—	—	—	5)
St	»Stephan«	1911	22	—	—	—	
T	»Thor«	1903	26	4	—	—	
		1904—05	43	10	—	—	
		1906—10	38	17	1	—	
Ta	»Tanche«	1923	1	—	—	—	
Td	»Tadorne«	1925	1	—	—	—	
Tp	»Tampa«	1925—29	—	—	—	—	6)
Va	»Valdivia«	1898	2	1	—	—	7)
Vi	»Viking«	1913	5	—	—	—	
Vk	»Vikingen«	1929—30	14	—	—	—	
WS	»William Scoresby«	1926—27	31	9	—	—	
		1928—29 / 1929—31	195	79	36	—	
X	»Xauen«	1930	8	—	—	—	
		1932	27	—	—	—	
o. N. Tp, Mo u. a.	ohne Namensangabe Beobachtungsschiffe des Ice Patrol Service (»Tampa«, »Modoc« u. a.)	1927	5	—	—	—	
		1922—30	594	1	—	—	8)
		Summe	3047	1226	622	110	

1) Umkipptiefe durch thermometr. Tiefenmessung bestimmt. 2) Meist aräometrisch bestimmt, manche Werte zweifelhaft. 3) Aräometrisch bestimmt, manche Werte zweifelhaft. 4) Manche Werte zweifelhaft. 5) Aräometrisch bestimmt, manche Werte zweifelhaft. 6) Vgl. am Schluß unter Ice Patrol Service. 7) Aräometrisch bestimmt. 8) Meist mit salinity-tester bestimmt, manche Werte herausfallend.

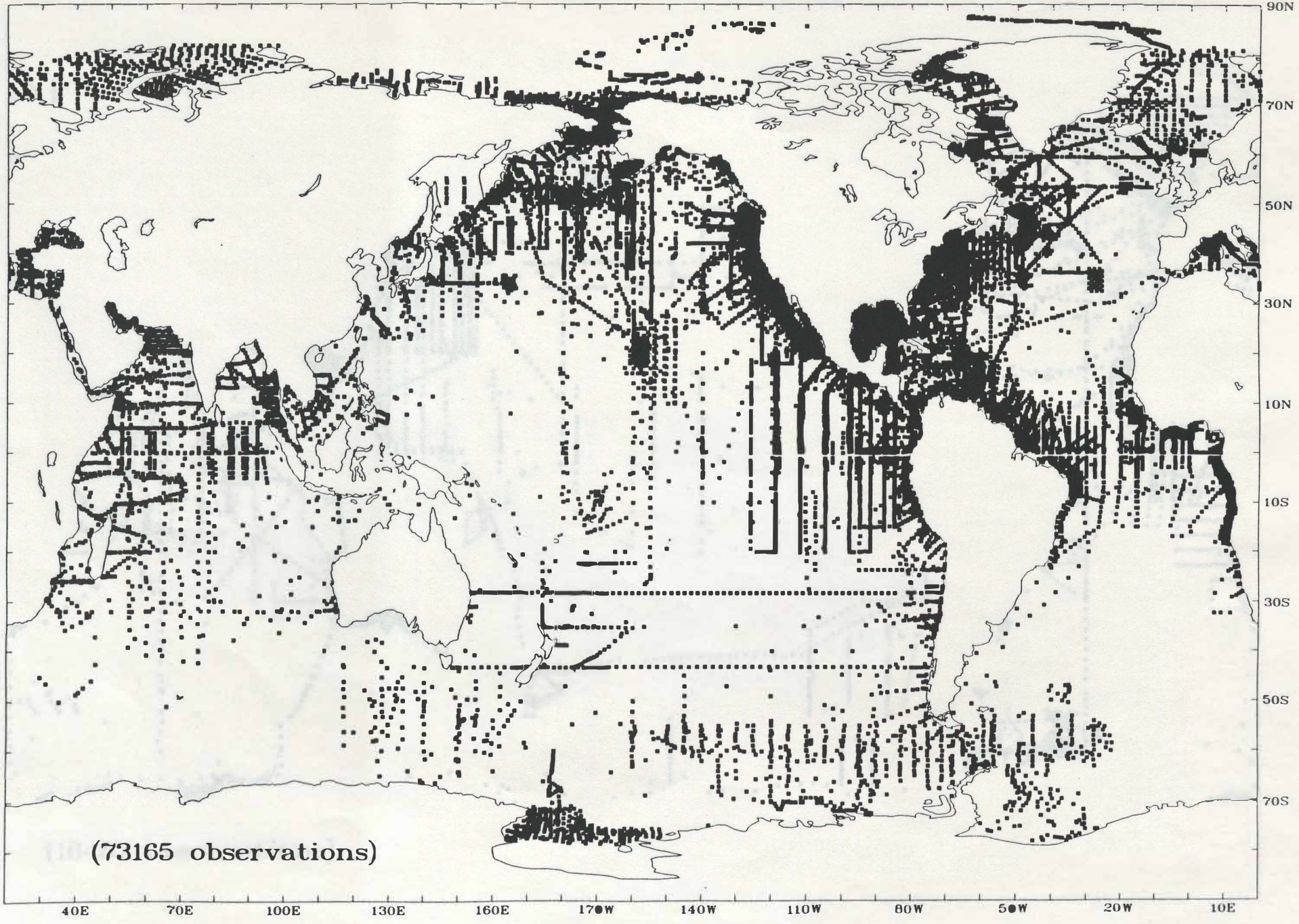
APPENDIX IV

Sample Data Distribution Maps

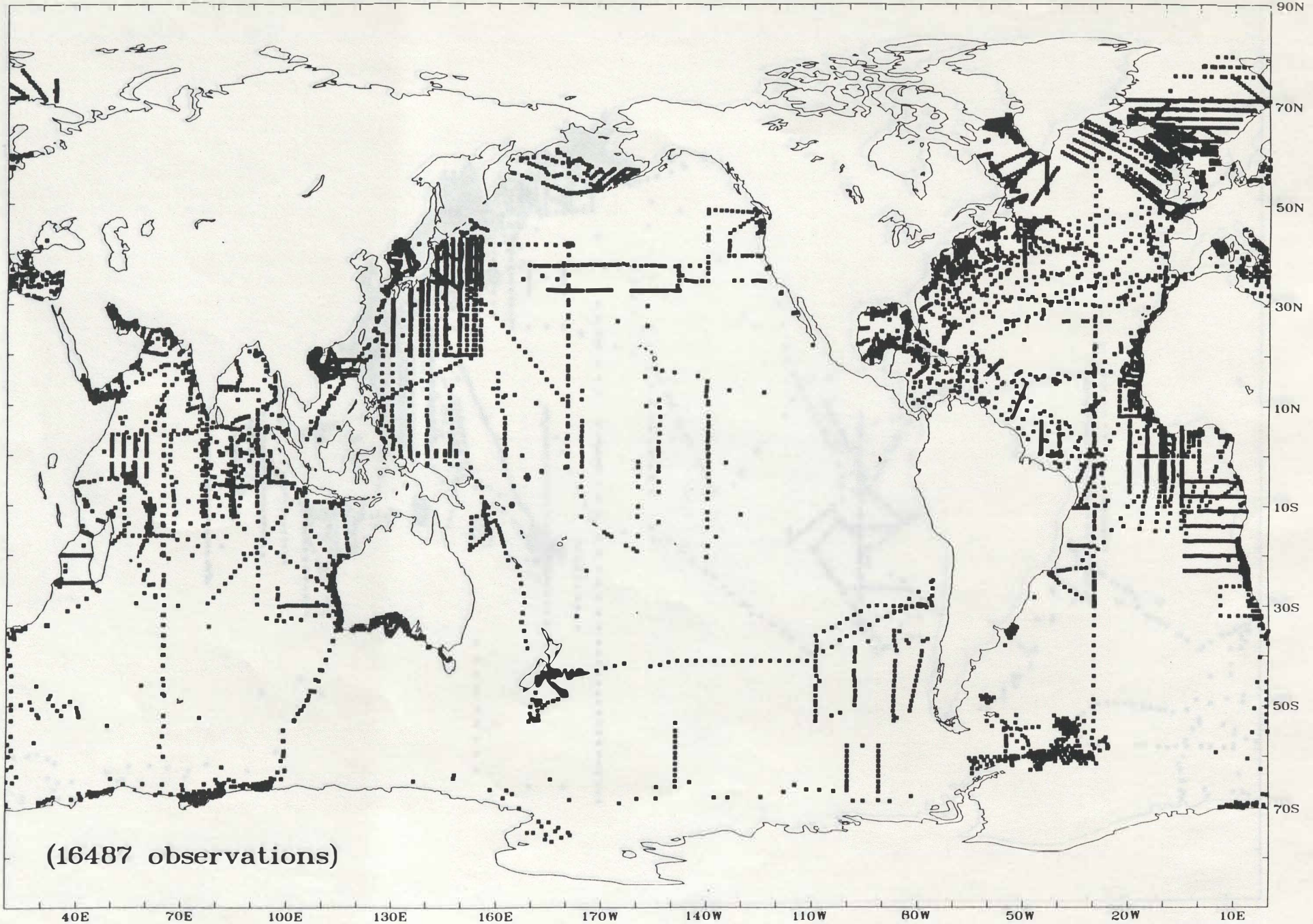
For data in U.S. NODC/WDC-A archive collected by various countries.



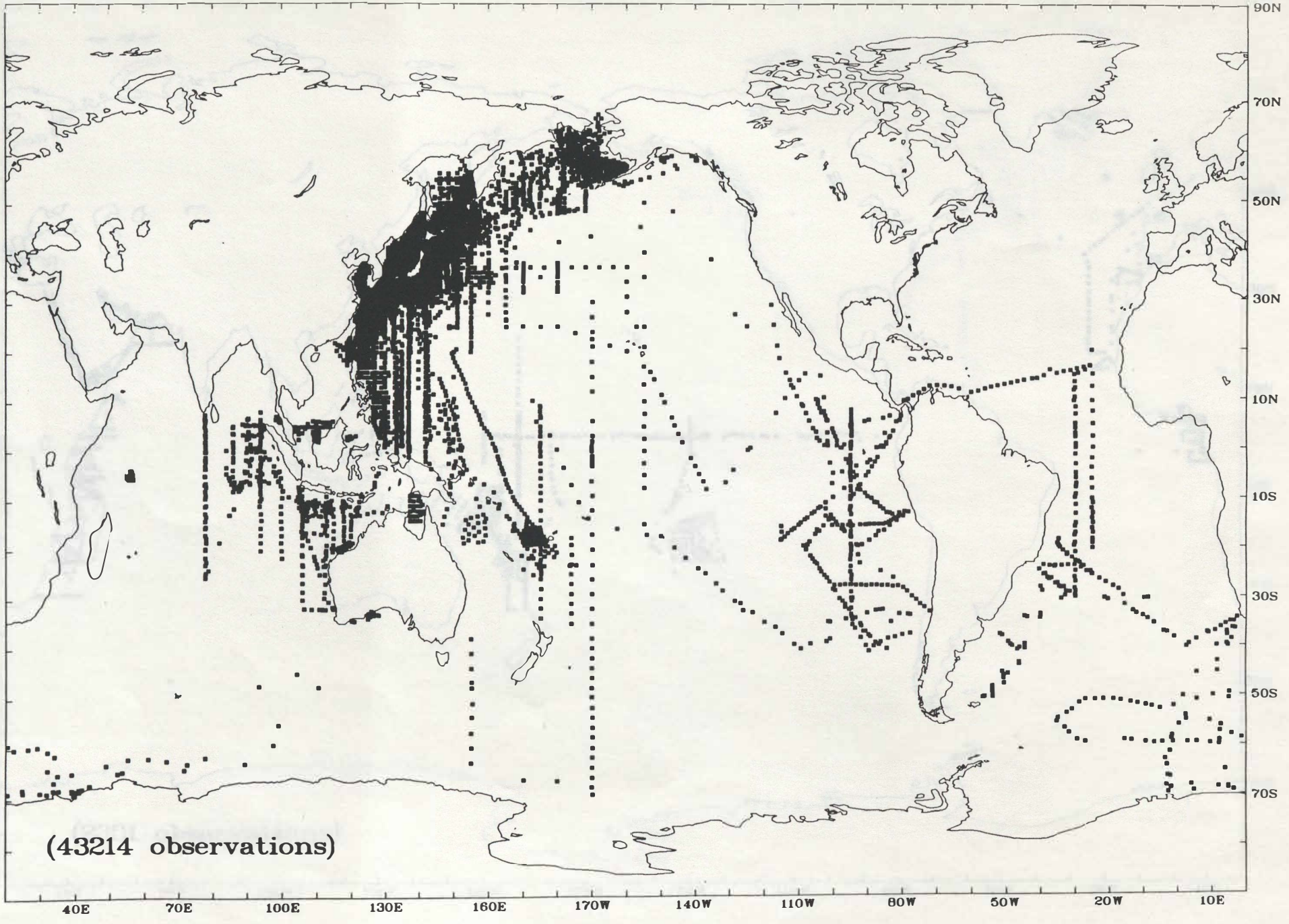
Nansen Cast Data for UNITED STATES (1960–1969)



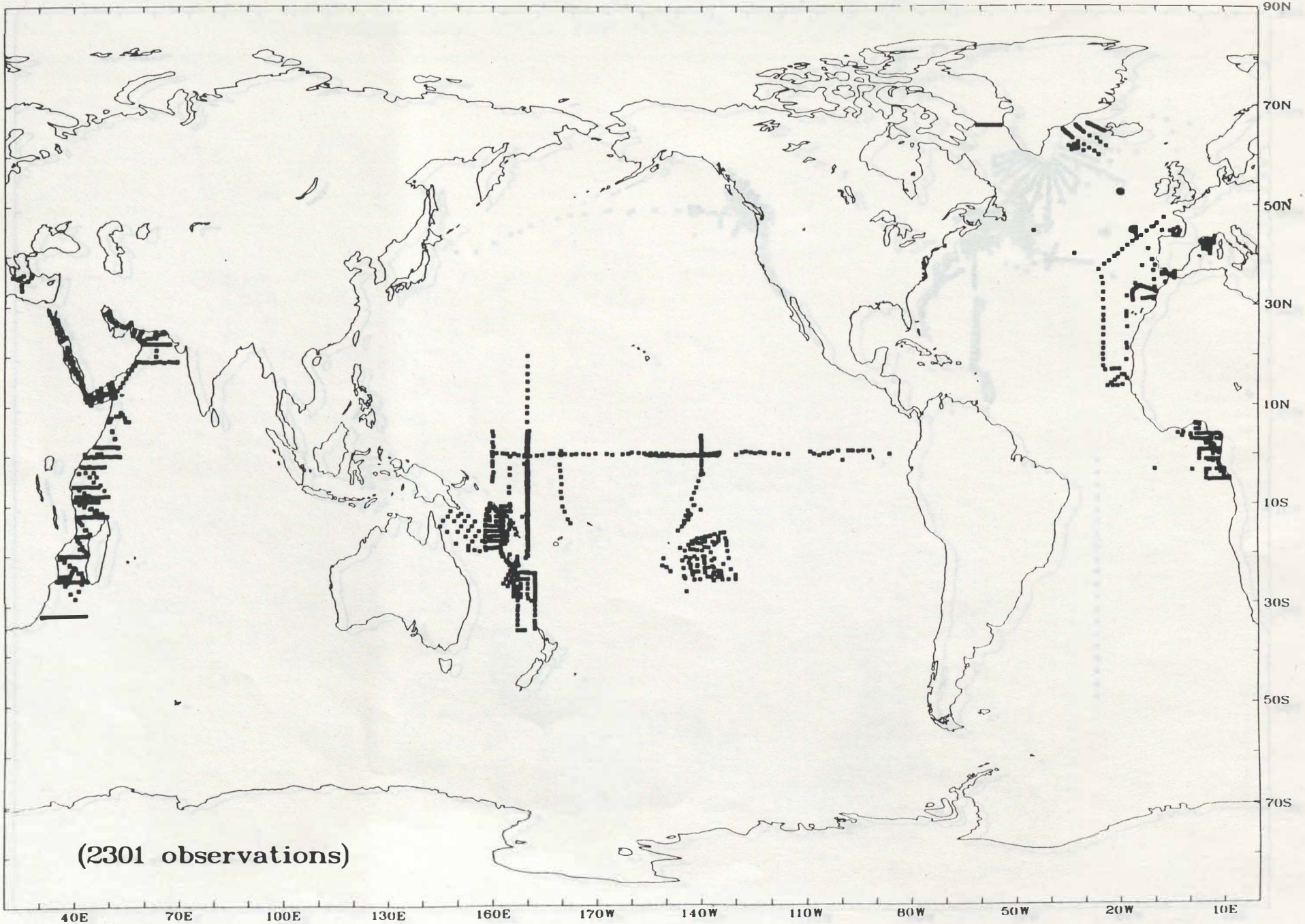
Nansen Cast Data for USSR (1960–1969)



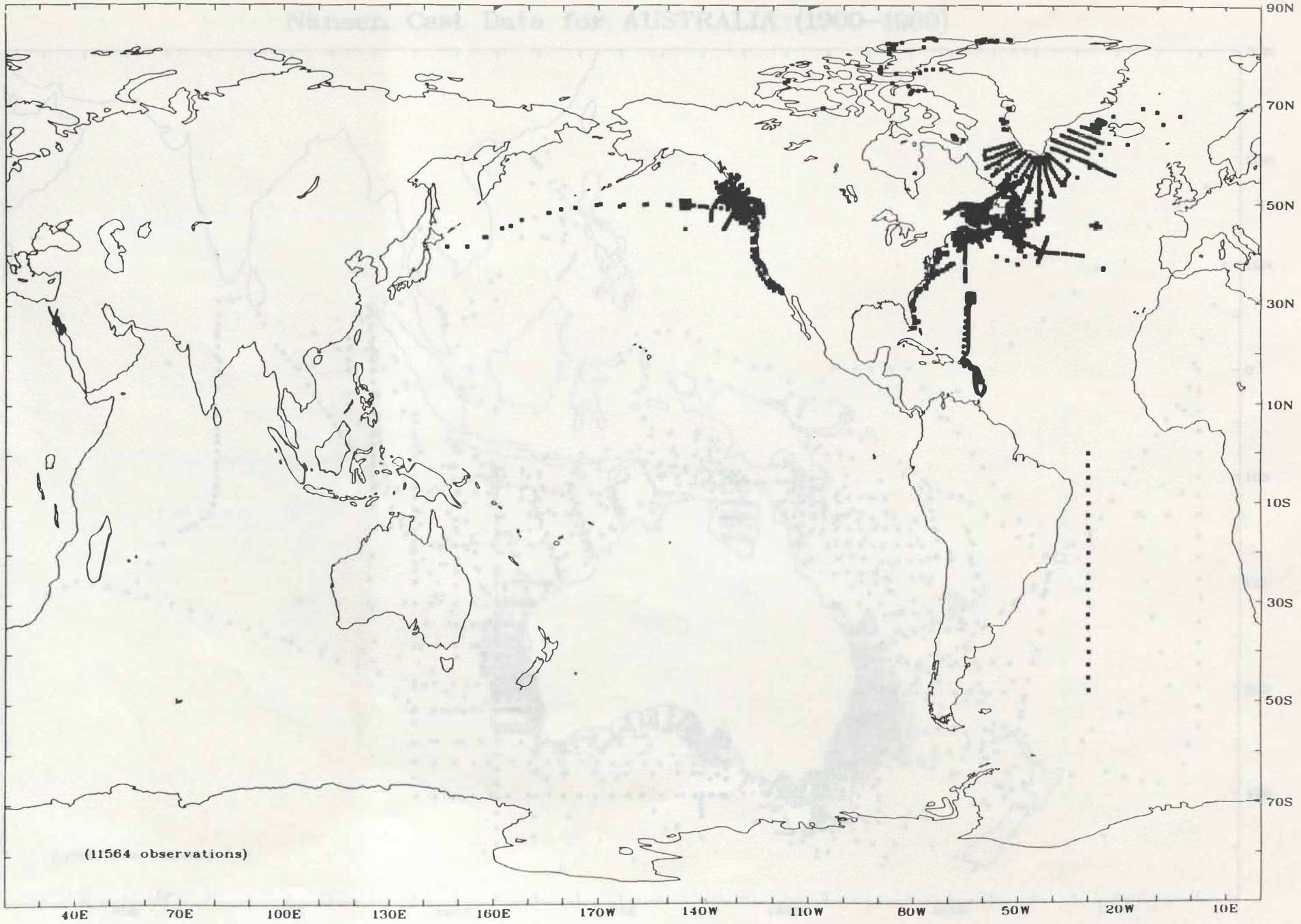
Nansen Cast Data for JAPAN (1960–1969)



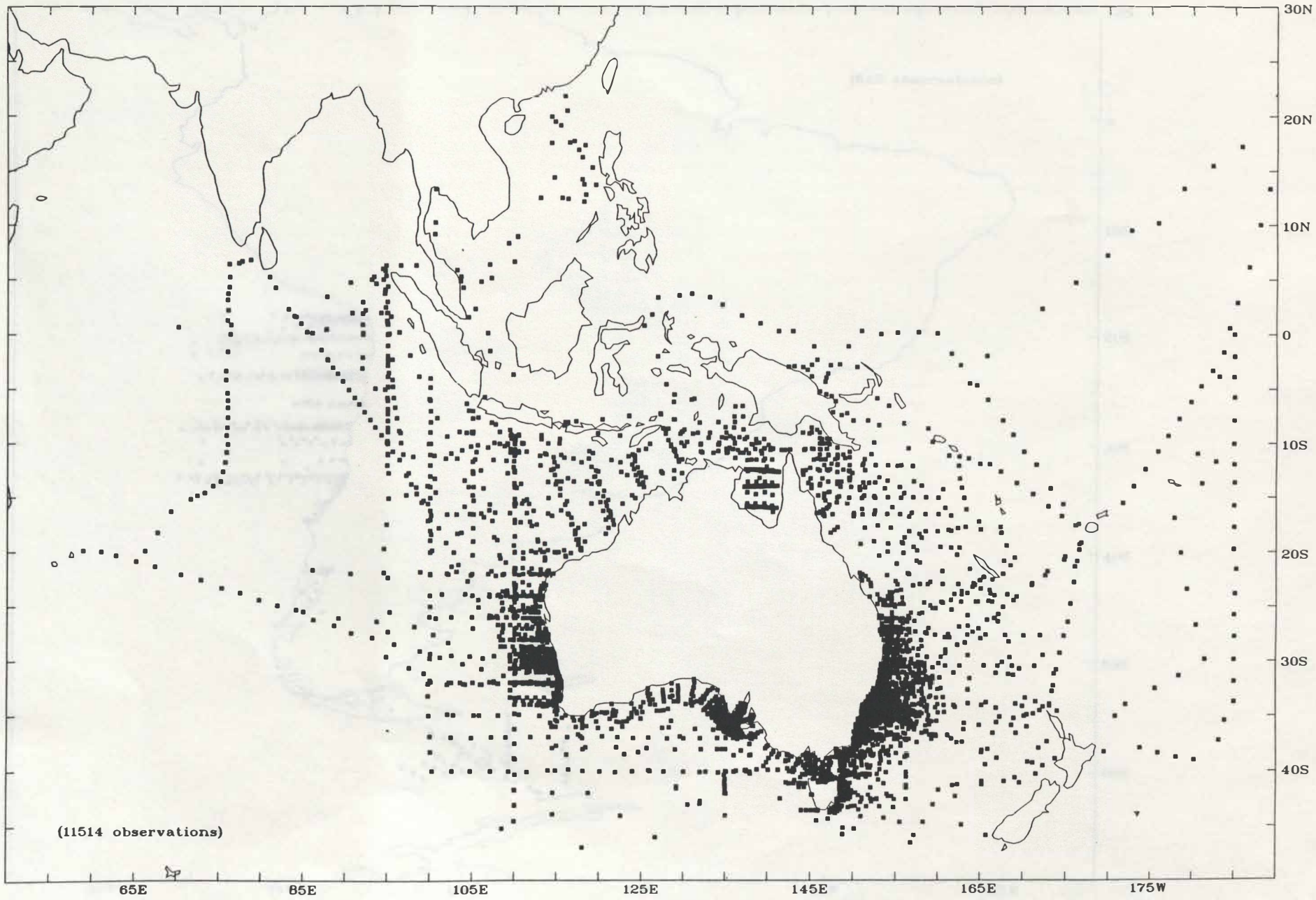
Nansen Cast Data for FRANCE (1960–1969)



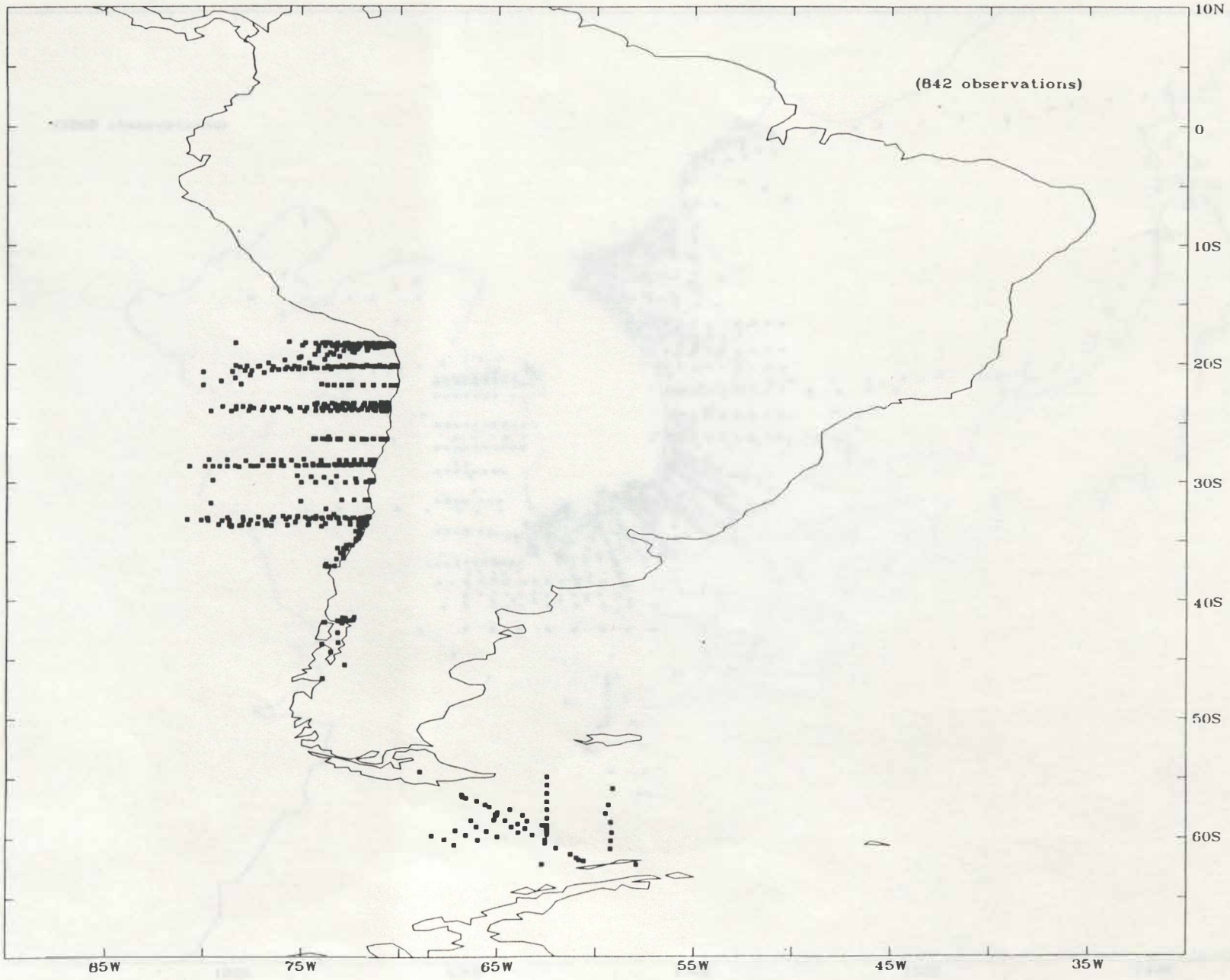
Nansen Cast Data for CANADA (1966–1969)



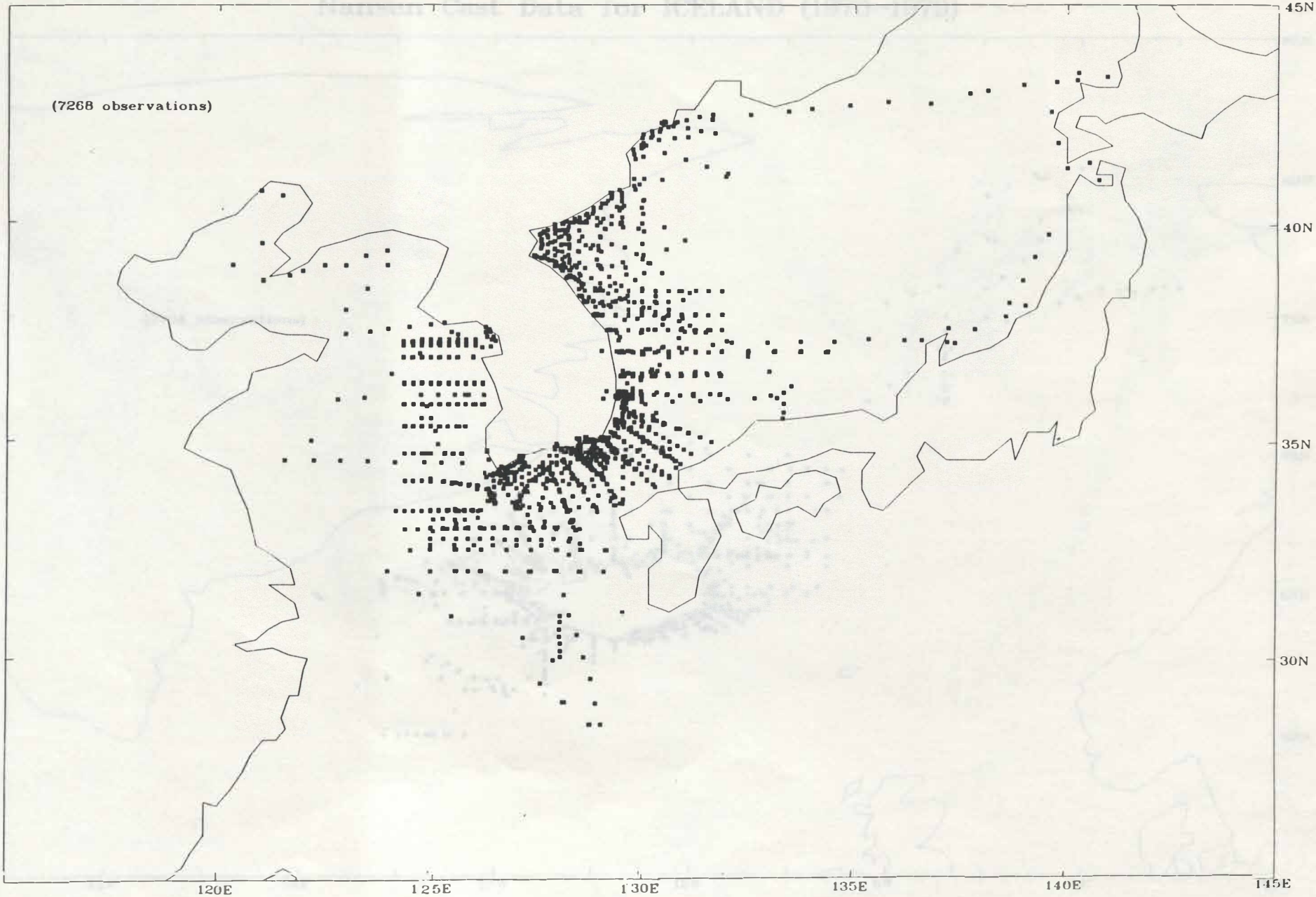
Nansen Cast Data for AUSTRALIA (1900–1989)



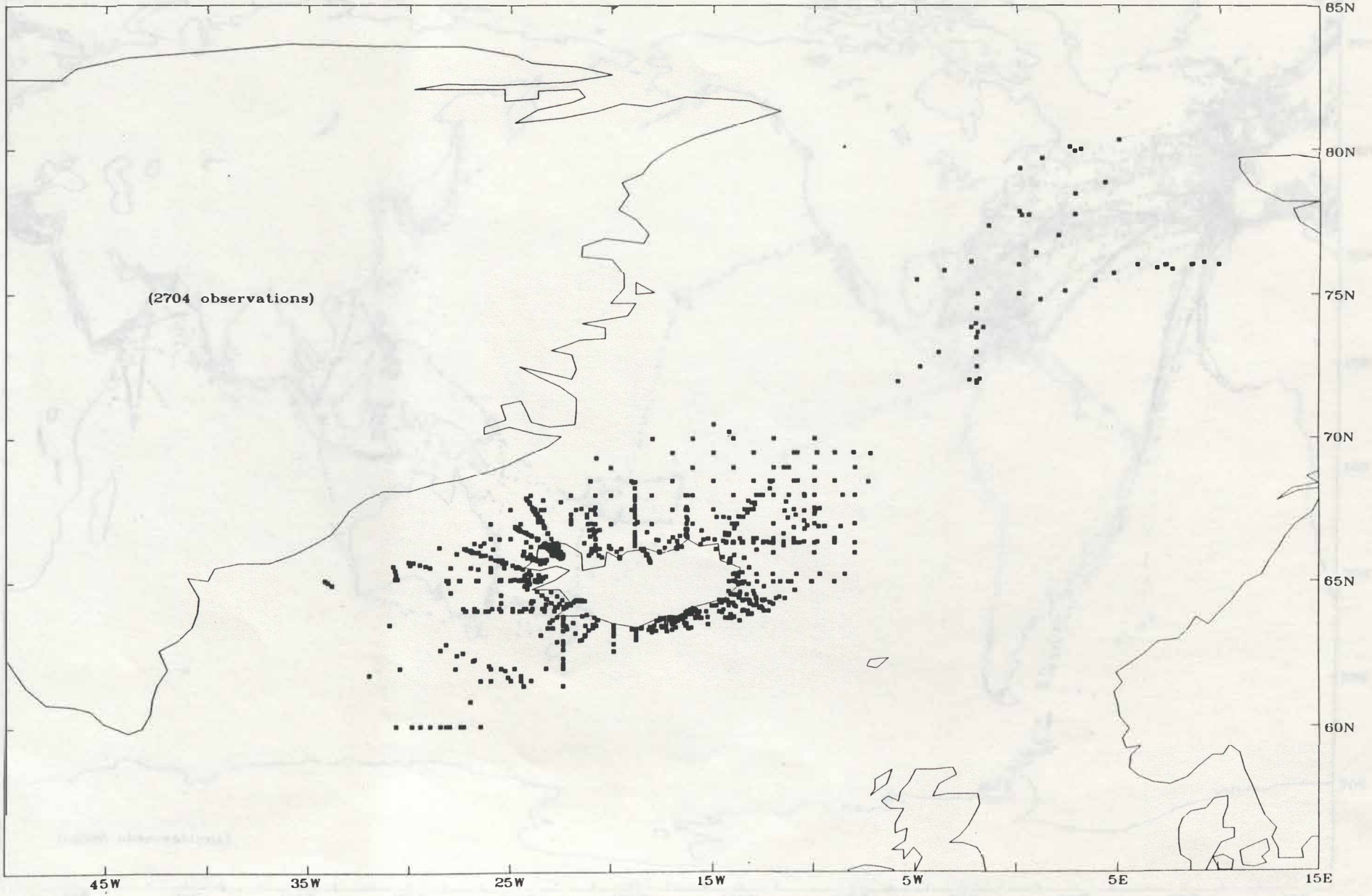
Nansen Cast Data for CHILE (1900–1989)



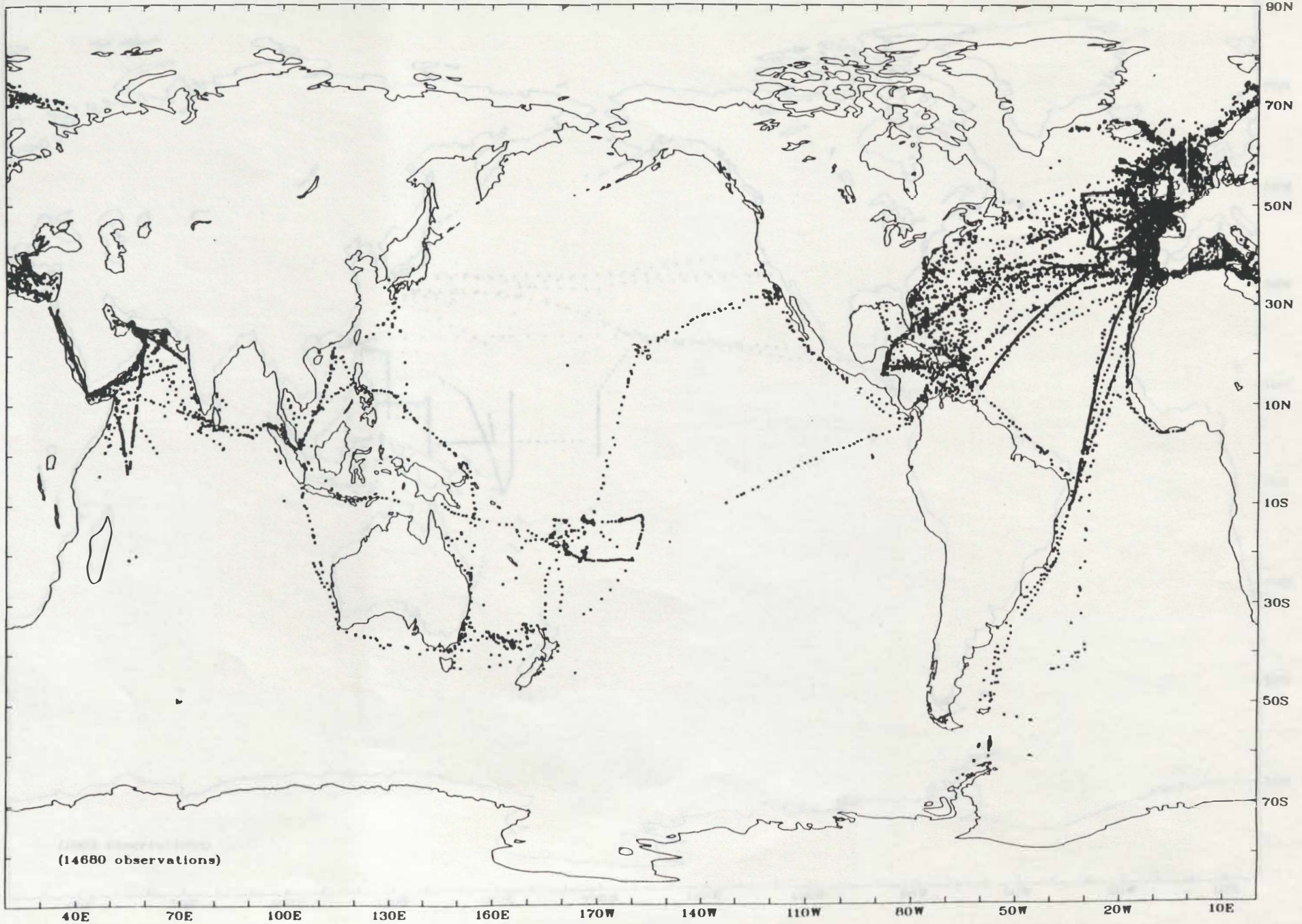
Nansen Cast Data for REPUBLIC OF KOREA (1900–1989)



Nansen Cast Data for ICELAND (1970–1979)



Expendable BT Profiles for UNITED KINGDOM (1975-1979)



Expendable BT Data for CHINA (1982–1989)