NATIONAL ● OCEANOGRAPHIC ● DATA ● CENTER

Final Report of the NODC/ERL Workshop on Ocean Data Files

13-14 June 1988
National Oceanographic Data Center
Washington, D.C.

September 1988

Sponsored by:
National Oceanographic Data Center
Washington, D.C.

and

Environmental Research Laboratory
Boulder, Colorado

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

Final Report of the NODC/ERL Workshop on Ocean Data Files

13-14 June 1988
National Oceanographic Data Center
Washington, D.C.

Sponsored by:

National Oceanographic Data Center
Washington, D.C.

and

Environmental Research Laboratory
Boulder, Colorado

1 September 1988

Table of Contents

Final Report of the
NODC/ERL Workshop on Ocean Data Files
13-14 June 1988

I. INTRODUCTION

The National Oceanographic Data Center (NODC) and the Environmental Research Laboratory (ERL) hosted a workshop on Global Ocean Data Files on 13-14 June 1988 in Washington, D.C. The workshop, which was chaired by Irving Perlroth (NODC), brought together 26 representatives from operating activities of the National Oceanic and Atmospheric Administration (NOAA) including the Atlantic Oceanographic and Meteorological Laboratory (AOML), ERL, the Geophysical Fluid Dynamics Laboratory (GFDL), the National Ocean Service (NOS), the National Weather Service (NWS), the Office of Oceanic and Atmospheric Research, and NODC; from activities of the U.S. Navy including the Institute for Naval Oceanography (INO), the Naval Ocean Research and Development Activity (NORDA), the Naval Oceanographic Office (NAVOCEANO), and the Office of Naval Research (ONR); and from other interested groups at Compass Systems Inc., the Marine Environmental Data Service of Canada (MEDS), the National Center for Atmospheric Research (NCAR), and Planning System Inc. (PSI).

In his opening remarks, Greg Withee (NODC) noted that the proposed objectives of this workshop were to discuss the general subject of ocean data archiving and distribution, and specifically to focus on two areas:

1) the development and maintenance of a continuously-updated, scientifically quality-controlled, global temperature/salinity (T-S) data base. This data base would rely on both real-time and delayed-mode data sources and might be available on-line.

2) the development of a comprehensive and scientifically quality-controlled data base of historical hydrographic station data which might be distributed on a compact disk with read-only memory (CD-ROM).

Withee noted that there are several advisory groups that have recently been formed to review and provide guidance to data mananagement development efforts. The Committee on Geophysical Data of the National Academy of Sciences (NAS) is serving as an advisory group to the Interagency Working Group on Data Management for Global Change. This NAS committee, which is chaired by Dr. Ferris Webster (University of Delaware), has also agreed to advise NODC. A scientific subpanel for this purpose has not yet been formed. A NOAA Data Management Council, chaired by Greg Withee, has also recently been formed.

As a result of increased interaction with the academic research community, NODC has become involved in several new data management development efforts including the Joint Environmental Data Analysis (JEDA) Center with Scripps Institution of Oceanography (SIO), the NODC Ocean Science Information Exchange (NOSIE), collaboration with Barney Walsh of University of Miami in compiling a data base of historical ocean pigment observations, with Glenn Flierl of Massachusetts Institute of Technology (MIT) in planning for Global Ocean Flux Study (GOFS) data management, and Ferris Webster of University of Delaware in planning for World Ocean Circulation Experiment (WOCE) data management and the Joint Center for Research on Management of Ocean Data. It will be proposed that WOCE adopt the JEDA model in development of an Atlantic thermal data analysis center under the direction of Bob Molinari (AOML).

NODC is currently linked to several ocean data management and research activities via networks and more network connections are planned. Active data acquisition efforts have resulted in more data arriving at NODC in a more timely fashion than in the past. Half of the data arriving at NODC are now in a master file within 45 days. The quantity of data added to the archives in the last decade justify an update of the Levitus atlas. In 1985-86 alone, the archives increased in volume by 30 percent.

II. ONGOING TIMELY DATA PROCESSING

A. Real-time Update/Quality Control - Paul Friday/Roger Bauer (Compass Systems, Inc.)

The NOS Office of Ocean Services, in cooperation with the Ocean Products Center (OPC), is focusing on the real-time data stream and as part of its ocean monitoring mission, is taking the lead in providing end-to-end real-time data management. Problems that are currently encountered in this area are:

1) not all useful data are entered into the real-time data stream,

2) some data in the stream are not useful because of deficiencies in quality,

3) from an operational perspective, data perish if they arrive too late.

Issues:

1) There is a need to maximize utility of existing data (especially since there are minimal funds for new observational systems). To this end an inventory of all federally funded data sources (mostly fixed platforms) is currently being compiled. Sources are being judged on the basis of data quality. Eventually the inventory will include international, state, and local data sources as well.

2) Non-standard formats are a problem for some data handled by Argos. There is a need to expedite reformatting so that these data can be entered into the Global Telecommunication System (GTS).

3) Quality of real-time data is being improved by the implementation of the Quality Improvement Profile System (QUIPS) 2 at OPC. The Fleet Numerical Oceanography Center (FNOC) is also performing quality control (QC). QC should be performed as close to the source as practical. This should be interpreted as distribution of effort, not duplication of effort.

OPC Objectives:

1) Provide for operational QC of all marine observations which impact NOAA's ability to generate forecast guidance.

2) Ensure dissemination of marine data and products on the GTS.

3) Monitor all marine forecast guidance dissemination from the OPC/National Meteorological Center (NMC) complex.

4) Interactively generate global and regional marine forecast guidance products and analyses.

The QC operations at OPC are being phased in over three or more years (1988-90). There are both automated and interactive parts to the QC process.

Discussion:

Roy Jenne (NCAR) cautioned that any modifications or deletions to the real-time data stream should be reversible. He recommended retaining a copy of the original unaltered data.

Paul Friday pointed out that NOS needs to expand into biological and chemical data for coastal applications and GOFS.

Warren White (SIO) wanted to know if problems detected by the QC process were investigated in an effort to correct the problems at the source. For example, in the case of expendable bathythermographs (XBTs), research vessels experience significantly fewer probe failures than Voluntary Observing Ships (VOS). This is probably due to the improper handling and storage of probes before launch. The age of the probes is another factor in failure rate. Pressure measurements from drifting buoys also exhibit suspiciously high failure rates. Roger Bauer indicated that such remedial efforts are being undertaken. Sydney Levitus (GFDL) noted that some data problems could be detected almost immediately aboard ship by comparing new observations with historical climatology.

Roger Bauer reported that Master Oceanographic Observations Data Set (MOODS) is no longer being regularly updated. NAVOCEANO has assumed responsibility for MOODS, but they plan to update it only when their needs dictate. Furthermore, only about 40% of the newly acquired data are unclassified; more should be.

B. Monthly Models - Ants Leetmaa (NWS)

Global sea surface temperature (SST) analyses are constructed using data from many sources. For many parameters, observational data are too sparse. In some cases (e.g., wind fields) we must rely on models to fill in gaps in observational data.

An ocean model assimilates SST observations and is forced by model winds; the ocean response provides fields of sea level, temperature, and velocity. Future plans are to provide weekly products (at present they are provided monthly) to expand the area of coverage to the entire Pacific basin, and to begin assimilating sea level and current data.

The Ocean Analysis System's ocean global circulation model extends from 30 S to 50 N and has a one-third degree N-S by one degree E-W grid with 27 depth levels. Real-time SST input comes from XBTs, Advanced Very High Resolution Radiometer, drifting buoys, etc. XBT observations appear to be fewer for 1988 than for previous years. Delayed-mode data come from moored current meters and the Navy's geodetic satellite (GEOSAT) altimetry.

A comparison between model output and observed currents are quite good (especially when assimilation is used to correct the model). Model sea level compares favorably with Klaus Wyrtki's (University of Hawaii) tide gauges and GEOSAT. Interannual surface dynamic height variability as revealed by the model between 5 N and 5 S is consistent with the hypothesis of White and Pazan (SIO) that off equatorial Rossby waves are related to El Nino and can be used as a forecasting aid.

C. Temperature/Salinity Project Overview - Ron Wilson (MEDS)

Ron Wilson observed that we all have the same objective of wanting to understand what is going on in the ocean. Observations and modelling are both needed. Since we are never going to have "enough" data, we should save all that we can. However, he cautioned that output from numerical models will rival satellite observations in terms of data volume.

T-S project goals:

1) To create, from real-time and delayed-mode sources, a timely and complete data base of ocean temperature and salinity data of known quality in support of the World Climate Research Program and of national requirements (e.g., fisheries) for these data.

2) To improve the performance of the Intergovernmental Oceanographic Commission (IOC)/International Oceanographic Data Exchange (IODE) system by exercising the data inventory, data management, and data exchange mechanisms as they are intended to work and recommending changes where necessary to meet national and international requirements.

3) To demonstrate through a widely distributed monitoring report produced on a regular basis (e.g., quarterly), the performance of the IODE system of world data centers (WDCs), responsible national oceanographic data centers, and national oceanographic data centers. It may turn out that the system does not work and that it is impossible to make it work. We need to identify

and demonstrate the problems.

4) To improve the state of historical data bases of oceanographic data by developing and applying improved QC systems to these data bases.

5) To distribute copies of the data base or portions of it to interested users and researchers in forms compatible with modern computer technologies.

Discussion:

The attendees agreed that data that have been converted to a common format are much more likely to be used. The initial goal would be to get all delayed-mode data into the data base within six months of observation; then they would be available for modellers and other users.

With regard to QC, Ron Wilson noted the need to develop improved climatologies. Bob Molinari expressed skepticism about the quality of salinities made available in "real-time"; raw salinity profiles must be calibrated before their accuracy can be assured. Warren White cautioned that even data that have been subjected to QC are not to be trusted.

D. Future JEDA Projects - Steve Pazan

The JEDA Center is based on the belief that distributed data management, which involves those actively doing research with the data, leads to better quality data.

At present, objective analysis techniques are used to construct the following monthly products for the tropical Pacific:

1) Data distribution maps
2) Normalized error maps
3) Sea surface temperature
4) Average temperature (0-400 m)
5) Depth of 14 C isotherm
6) Dynamic height (0-400 db)

These products are being compared to model output and independent observations (e.g., GEOSAT).

JEDA plans to expand analyses to the entire Pacific for WOCE. A monthly climatology will be developed; El Nino periods will be segregated first. The procedure will employ robust regression techniques and median filtering; the median of a population is much more resistant to outliers than the mean.

The availability of observed salinities will allow significant improvement in the computation of dynamic height.

A paradigm for future JEDA Center activities was presented. In situ data will be subjected to QC and analyzed to determine temperature, salinity, dynamic height, and currents. These fields will be used in verification of analogous fields derived from remotely-sensed observations and those derived from ocean models. The intercomparison between observations and models will allow the production of research-quality mesoscale-resolution products.

Discussion:

With regard to QC, it was suggested that artificial intelligence or expert systems might help with the more routine QC procedures, especially for voluminous data sets. Roger Bauer noted that institutional QC procedures (as well as those employed by an intelligent system) become obsolete over time. He cautioned that artificial intelligence is no panacea. Ants Leetmaa questioned the need for data centers to perform QC. He observed that quality involves very subjective judgement; most researchers will want to do their own QC in any case. Even when QC is performed, it is important to save the original data. In this regard Steve Pazan reminded the group of the need to transfer data to modern media before the original medium becomes obsolete.

Roger Bauer suggested that climatologies need to be made dynamically consistent; they should not be merely a statistical representation of the observations. However, Sydney Levitus countered that the dynamics are still not completely known, which accounts for part of the problem with present model predictions.

E. Continuously Updated Data Base - Chris Noe (NODC)

The envisioned continuously updated data base (CUDB) is characterized by being "on-line," accessible via networks, and interfaced to QC. The data base is to retain only the "best" version of the data; a suggested data replacement hierarchy was presented. An integrated inventory will provide indexed access to the data.

Why is a CUDB needed? Users require access to large datasets. The existing system at NODC is not well integrated; multiple archives confuse the users and invite duplication of data and inventory entries. The system is not adequately responsive to the user and has poor editing access. Noe felt that the current practice of grouping data by cruise is outmoded. The CUDB schema is based on the station model (i.e., vertical profile is the basic unit).

Discussion:

Someone cautioned that there is still a need to be able to identify and retrieve data from an individual cruise. Roy Jenne urged that only exact duplicates should be purged, otherwise the objective choice of which entry to purge might be incorrect.

III. DELAYED-MODE DATA PROJECTS

A. MOODS Update - Bill McQueary (NAVOCEANO)

MOODS will consist of all T-S profile data from NODC as well as classified data not at available from NODC. Perhaps products generated using the complete (i.e., including classified data) MOODS archive could be unclassified.

MOODS Update (continued) - Eugene Molinelli (PSI)

Objectives of MOODS project include:

1) accommodate all data available in MOODS plus data from other sources including high resolution data,

2) allow generation of existing products (e.g., General Digital Environmental Model) plus new products,

3) permit updates and editing,

4) permit flexible and efficient data retrieval,

5) achieve computer independence.

The data base consists of two files: a header file and a data file. The system supports two formats, FEB or MOODS format, but not all fields in MOODS.

Discussion:

The attendees were pleased that NODC would be the outlet for unclassified MOODS and urged that the efforts for declassification continue.

B. Historical Data QC - Sydney Levitus

In all cases QC procedures should be documented (preferably published). Data errors can be grouped into two classes: 1) impossible values and 2) suspicious values. Vertical stability provides a useful QC tool for quality checking hydrostation data. In checking the NODC archive, data from certain countries (e.g., USSR) seemed more error-prone than others. The fact that NODC retains source information such as country and ship with the data in their master files makes such intercomparisons possible.

Analysis of the "cleaned-up" Levitus data base indicates that the ocean T-S characteristics are changing over time. The Panularius time-series (near Bermuda) shows two abrupt changes in 1964 and 1969. Cooling and freshening has occurred at 50 W between 30 N and 40 N.

Although there has been a significant increase in the data holdings at NODC in the last decade, there are still many old data missing. These need to be obtained. Even though the old data may be of dubious quality, with enough of them, the major signals in the historical record may be resolved.

Discussion:

The importance of such "cleaned-up" data bases, as that assembled by Levitus, to identifying and correcting problems in the observational system, to numerical modelling efforts, and to detecting global climate trends is fully appreciated and enthusiastically endorsed by the attendees.

C. CD-ROM Project and Possibilities - Pete Topoly (NODC)

It is proposed that a CD-ROM may be prepared which contains subsurface thermal data. Beyond that, there is still much to decide:

1) type of data: digital/graphical/both
2) distribution: global/single ocean/time-series
3) sensors: Nansen/XBT/MBT/STD
4) access: data only/indices/both
5) organization: spatial/temporal/sensors
6) software: retrieval/display/commercial
7) format

The tentative plan is to include all subsurface thermal profiles except mechanical bathythermographs (MBTs) from 1940 to the present. The CD-ROM will contain indices to expedite search and retrieval. Some applications software will be provided separately to extract data and display index counts and summaries of the data.

Discussion:

There were some reservations expressed about the choice of CD-ROM as a storage medium. Admittedly, searching and retrieving data from a CD-ROM can be very slow, but this standard medium offers the advantages of relatively low cost, convenience, and high-capacity. It was suggested that if NODC puts comprehensive data files on a CD-ROM, the users will develop their own access and display software. This software might then be shared (with distribution via SPAN).

Demonstrations:

There were demonstrations of a U.S. Geological Survey CD-ROM showing Geological Long Range Inclined Asdic (GLORIA) data (revealing a river-like feature on bottom of Gulf of Mexico), and a National Aeronautics and Space Administration (NASA) CD-ROM showing Coastal Zone Color Scanner image data.

D. Data Access - NOSIE/SONIC

SONIC/OCEANIC - Jim Churgin (NODC)

The SPAN (Space Physics Analysis Network) Ocean Network Information Center (SONIC) is the first version of an on-line ocean information system developed at the University of Delaware. (The next and improved version will be called OCEANIC.) Access is via SPAN, OMNET, or dial-up modem. In the demonstration of SONIC, Jim Churgin showed graphical presentations of some North Pacific time-series (repeated hydrographic sections near Japan), WOCE proposed track lines, and a sample ocean product (O'Brien wind stress vectors). Text string search capability allows one to selectively retrieve entries.

NOSIE - Doug Hamilton (NODC)/Jan Ward (NODC)/Jim Churgin

The demonstration of the NODC Ocean Science Information Exchange (NOSIE) system by Jan Ward included access to a user-queriable archive summary.

E. NODC/SIO Hydrosearch - Pete Topoly/Mike Simmons (NODC)

Mike Simmons demonstrated the SIO-developed Hydrosearch system. In the future it may be possible to access this system from NOSIE. It may also be possible to by-pass NOSIE and access Hydrosearch directly via SPAN or OMNET. Hydrosearch is not currently self-explanatory (i.e., user-friendly), but it has a user's guide.

F. INO Introduction - Greg Han (INO)

The Institute for Naval Oceanography (INO) is a University Corporation for Atmospheric Research (UCAR) institution at the National Space Technology Laboratory funded by ONR. The mission of INO is to develop and implement global eddy-resolving models for Navy applications. Greg Han is working on the problem of providing access to needed ocean data from existing national sources. INO is likely to become a major user of these data sources. INO needs a state-of-the-art data handling system. There is also a need for much cross-fertilization with the academic community.

Han is in the process of visiting most national facilities involved in data handling activities: Compass Systems, Inc., ERL, FNOC, Georgia Tech Research Institute, Goddard Space Flight Center (GSFC), Jet Propulsion Laboratory, NCAR, National Climatic Data Center, NMC, NOAA/Ocean Applications Group, NODC, NAVOCEANO, Navy Post-Graduate School, SIO, UCAR/UNIDATA, University of Delaware, University of Miami, University of Washington, and University of Wisconsin.

To date Han is most impressed with the NASA Climate Data System (NCDS) at GSFC. NCDS deals mostly with satellite data. The catalog is based on the Oracle data base management system. Each data set is accompanied by a format translator which converts to Common Data Format.

Han has observed that are many groups working on the data management coordination problem; several are in the Navy. The Earth Science and Applications Data Systems initiative is attempting coordination within NASA.

IV. DISCUSSIONS

A. Temperature/Salinity Project - Ron Wilson

The following issues were identified for the proposed T-S project during the previous presentations:

1) Parameters to be included in the data base:

Should we include salinity? Someone noted that we do not even have thermal data adequately compiled and organized yet. However, salinity is important in the upper ocean where T-S properties are highly variable. Physical oceanographers would not be the only beneficiaries of the inclusion of salinity. Fisheries can utilize salinity data. Modellers would benefit from good salinity data for boundary conditions (infer precipitation, thermohaline forcing, etc.).

Roy Jenne suggested that we should call for more in situ rainfall observations. Ants Leetmaa noted that atmospheric models predict rainfall.

Instruments, such as the expendable conductivity-temperature-depth (XCTD) profiler, need to be checked out for accuracy. Bob Molinari stated that salinity sensors should not only include XCTD, but also surface thermosalinograph data.

Even with XBTs, we need to study and develop the optimal sampling strategy. Bob Molinari wondered how useful the existing models might be in helping in the design of sampling strategies. The tropical models are inappropriate in mid-latitudes.

Should we include current velocity? Ants Leetmaa noted that the Acoustic Doppler Current Profiler (ADCP) is now an operational sensor. However, Jim Churgin reminded the group that there was much disagreement on the utility of these data. Furthermore, there is trouble with assimilating these observations into models, since they are contaminated by non-baroclinic motions. Bob Molinari noted that in comparisons with PEGASUS profiles, AOML concluded that the ADCP is good to within 10 cm/s. Ants Leetmaa observed that once again data usefulness depends on what one intends to do with the data.

2) Time frame to start project:

Greg Withee asked what time constraints are being imposed by WOCE on the JEDA effort? Warren White and Steve Pazan indicated that NODC should be prepared to send thermal data for the entire Pacific to SIO in October. Chris Noe expressed some concern over being able to meet this schedule. Subsequent discussions have delayed the start for this expanded effort to December.

3) Time frame of data availability:

The real-time data will be merged with available delayed-mode data and the CUDB will initially be updated every two months. This update frequency should eventually be increased to weekly.

We should initially strive to have most delayed-mode data into the data base within six months. Over time this lag should be reduced to one month.

4) Data flows:

Sydney Levitus emphasized that we need to get all the data in. Station data alone are too sparse. Roger Bauer noted that a major potential benefit of archiving real-time data will be realized if the data are monitored to correct problems in the observational system. He expressed the sentiment that NODC should be located with an operational analysis center.

There are many problems with the GTS data stream. There appears to be an inconsistency between the amount of data entered and that which is subsequently captured downstream. Roy Jenne noted that approximately 10-15% of the data are apparently lost. Some data are entered many times. There is a need to do a better job of monitoring the system to determine the causes. Scott Woodruff (ERL) indicated that ERL saves all rejected duplicates to study this problem. Roy Jenne reiterated that only exact duplicates should be rejected. Ron Wilson indicated that at MEDS they save everything and only flag duplicates. Roy Jenne noted that many of the erroneous data in GTS can be attributed to erroneous positions. Perhaps we should think about modifying the GTS codes if they contain features that continue to create confusion.

5) Quality:

Should we include all sources of thermal data, such as drifting buoys with thermistor chains, etc.? There are important differences in data quality, principles of sensor operation, etc. Ants Leetmaa contended that for any particular parameter, we need to use all available sources in spite of varying accuracies. Roger Bauer cautioned that we should be careful here because the availability of any model output (regardless of quality of input observations) can affect the policy makers and have far-reaching consequences. Also, the in situ data are being used as ground truth for satellite observations. Steve Patterson (NODC) asked whether it will be possible to identify sensor type in the T-S data base? Roger Bauer answered that this wouuld not be possible for the data captured from GTS because sensor type is not indicated in the transmitted message. There needs to be additional supporting documentation with the data.

Flags that are added by QC efforts at FNOC and/or NMC are not presently accommodated at NODC and are consequently lost. Since subsequent analyses might benefit from these flags, an effort should be made to preserve them.

All agreed that QC should be applied as early as possible within the data stream. However, there is a need to build up climatologies before we will be able to do this very well. There are regional variations in climatology that need to be considered in QC. There may be a role for expert systems. Sydney Levitus felt that we will be able to address this issue better after we have computed more statistics on larger data bases. We should retain data judged as suspect because not all suspicious values are actually bad. On the other hand, Steve Patterson pointed out that many oceanographic cruises purposefully seek out "anomalies", such as fronts and eddies, to sample in detail. Therefore, climatologies may be based on historical data bases that contain a disproportionate amount of these types of observations. Levitus agreed and cited the deep high-salinity "Meddies" as an example. Roy Jenne contended that in the case of bad position or date, the value should be corrected, not merely flagged. Roger Bauer suggested that a CD-ROM could contain bathymetry and an atlas of climatology. Bob Molinari agreed and noted that he still takes a paper atlas to sea as a first check on data quality.

6) Formats:

Roger Bauer expressed concern that the rules governing the use of the Binary Universal Format for Records (BUFR) have not been defined and officially approved by the World Meteorological Organization. This format will eventually have important impact on the T-S project. If we are to start this year, we have to take the initiative; we can't wait.

7) Products:

It was noted that a prototype bulletin of graphical products called Ocean Climate Environmental Analysis News (OCEAN), distributed by NODC, has been well-received by the ocean community. One notable reaction to the bulletin was that people wanted access to the gridded data used to generate the products. (OCEAN has subsequently evolved into the Climate Diagnostics Bulletin: Delayed Mode Ocean/Atmosphere Analyses, distributed by the Climate Analysis Center of NOAA.)

Any proposed salinity products will likely be gappy because of sparse data coverage.

Greg Han noted that INO modelling applications will need data products rather than original data. One objective is to upgrade the quality of FNOC analysis products. Roy Jenne expressed concern that shared processing gives SST responsibility to the Navy. A lot of NESDIS (National Environmental Satellite, Data, and Information Service) expertise in SST products is being abandoned.

How do we decide which model output to archive? Warren White expressed concern that operational models tend to be kluged together to try to deliver a product as soon as possible. Hence, it is not clear that they merit archiving. Bob Molinari observed that NODC is not adequately handling in situ observational data. These data should be the first priority. Does it make sense to even consider archiving model output?

Benefits:

The attendess agreed that benefits to be derived from the proposed T-S data base iclude the following:

1) Monitoring the data system will potentially improve its effectiveness of observational system.

2) More process oriented studies will be possible with more timely and comprehensive information on data already collected.

Implementation:

Bob Molinari asked how the groups participating in the T-S project (e.g., MEDS, NODC, SIO, AOML, etc.) will link together and coordinate their activities?

Roger Bauer explained that currently the AUTODIN network, which is a secure communications link, brings data into FNOC. FNOC subjects the data to QC using QUIPS 1. The unclassified portion of these data are sent via AWN to NMC where they are subjected to further QC before they are injected into GTS. NMC also receives data from other sources (e.g., coastal stations, Geostationary Observational Earth Satellite, international GTS, etc.) All U.S. data are subjected to QC before entry into GTS. It is planned that NMC should implement QUIPS 2, which does "context" checking, by the end of 1988. NODC is currently receiving data directly from FNOC via the Naval Oceanographic Data Distribution System and from NMC, and in delayed-mode from many sources.

The modifications to the existing flow diagram would send data directly to MEDS from FNOC and NMC. MEDS will perform QC and send the result (GTS-prime) to NODC over SPAN. NODC will reformat and merge GTS-prime with delayed-mode data and forward appropriate subsets monthly to SIO and AOML where additional QC occurs as part of research efforts. The scientifically quality-controlled data will be returned to NODC on a bimonthly schedule to update the CUDB.

The attendees advocated acceptance of the MEDS offer to QC data because 1) it provides an automatic archive backup, and 2) the additional QC might detect data problems that others miss.

Since Ants Leetmaa and others have a need for data on a weekly basis, updates should eventually accommodate these needs. Bob Molinari asked why update every two months, since the continual addition of delayed-mode data means the data base will not be stable for at least three years. This is to make data available to outside users and to obtain timely feedback on effectiveness of data sources. Of course, update frequency will depend to some extent on resources available.

Scott Woodruff reported that the Comprehensive Ocean-Atmosphere Data Set (COADS) is to be updated on three time cycles. Real-time data will be summarized in monthly reports; delayed-mode data will be merged and used to produce reports and summaries on both a one-year and a five-year cycle.

B. Historical Data Validation Project - Sydney Levitus

1) Availability:

The ocean's T-S characteristics appear to be changing. Are the changes that we see real? To answer this question, we need to locate and preserve as much historical data as possible. For example, some hydrographic data from the 1920's and some post-war MBTs are not in the archive (See Appendix D). Once the data are available, they can be analyzed. Scott Woodruff noted that there are gaps in COADS during the world wars that might be filled with data that have not yet been archived.

Regarding old data that are not digitized, how good are they? Is going after these old data a worthwhile expenditure of resources? Levitus reiterated that climate studies are limited by present data availability. Even marginal data may be valuable. We should initiate a pilot data recovery project to assess the magnitude of effort required. Bob Molinari contended that this should be a big effort, not a pilot project. He argued that retrieving existing data must be less expensive than obtaining new observations. Irv Perlroth reminded the group that the Tropical Ocean and Global Atmosphere (TOGA) XBT program in the tropical Pacific has demonstrated that such an effort can work.

Jim Churgin asked how NODC should prioritize data to seek out? ...by area?...by type? How should we deal with duplications? Ron Wilson pointed out that some "duplications" may have been purposefully obtained as "replicates."

Greg Withee suggested that we might select a group of people who are interested in the acquistion of old data and who travel a lot. They might be successful in identifying and retrieving these data. Also, NODC could produce plots showing what is currently in the archive and distribute them for review as to completeness. The Southern Ocean Center (SOC) in Argentina might be able to supply some help in digitizing some of the old data.

2) Quality:

The attendees agreed that QC is not a single process, but a series of steps. Some things are best done during initial data acquisition, others by data centers, and others by the researcher.

Sydney Levitus reiterated his call for putting climatology based on 5-degree square statistics on board ships to be used as a preliminary quality check. Roger Bauer added that a bathymetry data base should also be available. Both of these might be on CD-ROM.

Greg Withee questioned whether, with all the long- and short-term changes that have been described, it makes sense to develop a climatology? What would be the minimum adequate averaging interval? Another asked about potential biases introduced by changes in technology (e.g., MBT to XBT)? Levitus responded that we won't know the magnitude of such effects until we conduct a careful study with as much data as possible.

With regard to QC, Ants Leetmaa reiterated that data centers should just do the gross checking that is absolutely necessary. Then distribute the data with a disclaimer. QC is too subjective. Let each user apply his own criteria. Roy Jenne also reiterated his contention that QC at the data center should ensure that positions, dates, and mechanical things (e.g., format conformity) are right. These should be corrected, not flagged. Ron Wilson noted that in the case of drifting buoy data bad positions cannot really be corrected, they can only be flagged.

Ants Leetmaa noted that although models might be helpful in QC, it is inappropriate to extend the tropical ocean model into the mid-latitudes.

3) Other data types:

Bob Molinari complained that velocity profile data (such as PEGASUS data) and current meter data have not been assembled by NODC. Jim Churgin couuntered that NODC does have a file of current meter data, but since it is from diverse sources and sensor types, he acknowledged that it is of widely varying quality and not "cleaned up" like the Levitus data set. Greg Han observed that the data collectors's archiving requirement is technically met when the raw data have been submitted to NODC; subsequent improved versions of the data do not have to be archived and therefore frequently are not. Bob Molinari suggested that more people might utilize the NODC current meter data if they become aware that they are available. Warren White felt that there is a need to archive ADCP data. He stated that Dean Roemmich (SIO) will equip 10 VOS with ADCPs. These VOS will also be equipped with automatic XBT launchers that have a 12-probe capacity and this will allow the lone technician responsible for launching the probes to get some sleep. White encouraged NODC to produce plots showing the geographic distribution of data holdings for various data types.

4) Products:

Sydney Levitus suggested that helpful products would be parameter statistics and inventories of master files. These could show annual and seasonal distributions, help determine important time and space scales, and serve as a guide in QC efforts.

Anomalies defined relative to climatology obviously depend on what one uses for a climatology. We probably can not all use the same climatology. Roy Jenne argued that we want climatologies that have good spatial resolution, but as we increase the resolution the significance of the computed mean values decreases. Steve Pazan felt that this dilemma illustrates an advantage of the robust regression technique. Greg Han argued that there is a need to compute many statistics of the distributions, not merely means and standard deviations.

Sydney Levitus suggested that a CD-ROM product might include standard level data, T-S relationships, and maps showing data distribution. Greg Han was interested in a sampler of various data types. Someone reinterated that reading a CD-ROM is slow and urged that efforts be made to optimize searching the disk.

C. Ocean Climate Data Management Workshop - Jim Churgin

Jim Churgin announced that an IOC workshop is to be held about a year from now (i.e., spring or summer of 1989) somewhere in the U.S. It will be concerned with improving timeliness and quantity of data exchanged internationally. It is anticipated that there will be 20-25 speakers. The tentative plan is to start the meetings with formal presentations to stimulate follow-up discussions and recommendations. Greg Han noted that INO is also sponsoring a workshop in the spring of 1989 on real-time data flows. Han suggested that the IOC meeting might be held in conjunction with the INO workshop. The potential for such a joint meeting will be investigated.

Bob Molinari felt that there are already too many meetings and questioned whether much could be accomplished with so many people in attendance? It was noted that this workshop might provide a good opportunity to distribute plots showing NODC data holdings.

They would probably provoke some reaction. Roger Bauer also had some reservations and asked what could be accomplished at this proposed workshop that could not have been accomplished at the February 1988 Integrated Global Ocean Services System (IGOSS)/IODE meeting in Monterey or at one of the past project-specific (e.g., WOCE) data management meetings? Jim Churgin responded that IGOSS/IODE meetings consist mostly of data managers. This proposed workshop is intended to provide an opportunity for international dialog between data managers and researchers on important data exchange issues. Furthermore, the subject of this proposed meeting is broader than any single project. Roy Jenne asserted that the proposed meeting needs to be attended by those people who can actually implement efforts and follow through to results. Someone also pointed out that new proposals continue to emerge. For example, the T-S project discussed at the present workshop was not even conceived until after the February Monterey meeting.

Roger Bauer observed that past meetings have not been solving problems. The issues before us now are the same issues ones that have been before us for years. The recommendations coming out of the April 1987 INO workshop were never implemented. The ones coming out of this meeting may not be either. The structure of such meetings is so ad hoc that there is no real responsibility to follow through. Ron Wilson responded that he (MEDS) and Withee (NODC) certainly intend to follow through. The T-S project will be carried out; historical data will be actively sought; IGOSS and IODE data flows will be monitored. Community feedback, such as being provided here, helps in tuning the concepts and setting priorities.

V. RECOMMENDATIONS

A. Temperature/Salinity Project

1) Give responsibility for real-time QC of GTS BATHY and TESAC messages to MEDS. They need a SPAN link. They should monitor and evaluate the performance of the data links (e.g., monitor counts). Those who enter data into GTS should communicate via a separate channel what they entered.

2) Use the results of the QC efforts to identify and correct problems in the data collection system.

3) Include salinity in the data base; carry out a parallel effort in sensor (XBT, XCDT, thermosalinograph, etc.) sampling network design.

4) Initially update the CUDB at NODC every two months. Since Ants Leetmaa and others need data on a weekly basis, update frequency should eventually accommodate these needs. It is acknowledged that update frequency will depend to some extent on resources available.

5) Strive to acquire delayed-mode data within six months.

6) INO may archive global gridded fields.

B. Historical Data Validation Project

1) Recover as much historical data as possible. A pilot effort should be started.

2) Exploit IOC/WDC system to identify and obtain data from international sources.

3) Some old data will have to be digitized. Solicit assistance from SOC in Argentina.

4) NODC should put their data on a CD-ROM. The first NODC "experimental" CD-ROM should be given a high priority and hopefully completed this year Suggestions on subsequent CD-ROMs include: bathymetry, continental outlines, summary data (atlas), T-S profiles, previews of data distributions, and a sampler of different data types at NODC. Once NODC personnel have their "feet wet" on how to produce a CD-ROM, all future CD production efforts should closely involve the scientific community.

C. General

1) NODC should retain the original data as well as the quality-controlled set.

2) NODC should document (publish) QC procedures.

3) Climatologies need to be developed.

4) NODC should investigate the use of expert systems in QC efforts.

5) NODC in conjunction with the scientific community should develop a minimum set of standard products that will be routinely produced (not to preclude others).

6) NODC should distribute plots of its data holdings to community for review as to completeness.

7) NODC should keep velocity data in master files (current meters, profilers, ADCP, etc.)

D. Unresolved Issues

1) Non-standard formats are a problem for some data handled by Argos.

2) Only about 40% of data flowing into FNOC are unclassified.

3) IGOSS messages do not indicate sensor type.

4) NODC should be located with an operational center.

5) MOODS is no longer being regularly updated.

6) Should numerical model output be archived?

7) Archives need to utilize modern exchange mechanisms and modern storage media.

8) Rules governing the use of BUFR have not been defined and officially approved.

NODC/ERL
WORKSHOP ON OCEAN FILES

Room 406 - Universal Building
June 13-14, 1988

Monday - June 13 - 9:30 a.m.

I. Welcome and Opening Remarks - Withee

II. Ongoing Timely Data Processing

9:45am - Real-Time Update/Quality Control - Friday/Bauer

10:05am - Monthly Models - Leetmaa

10:25am - Temperature/Salinity Project Overview - Wilson

10:45am - Future JEDA Projects - Pazan

11:55am - Continuously Updated Data Base - Noe

LUNCH BREAK - 12:30pm - 1:30pm

III. Delayed Mode Data Projects

1:30pm - Moods Update - Teague/McQueary/Molinelli

1:45pm - Historic Data Q.C. - Levitus

2:00pm - CD-ROM Project & Possibilities - Topoly

2:45pm - Data Access - NOSIE/SONIC - Ward/Churgin

3:15pm - NODC/SIO Hydrosearch - Simmons

5:00pm - ADJOURN

Tuesday - June 14 - 8:30 a.m.

IV. Discussions 8:30am - 12:00pm

8:30am	- INO Introduction	- Han
	- Temperature/Salinity Project	- Wilson/Perlroth
	- Historical Data Validation Project	- Levitus
	- Ocean Data Management Workshop	- Churgin

V. Recommendations - 1:00pm

APPENDIX B

NAME	ORGANIZATION	TELEPHONE	E.MAIL
Ants Leetmaa	NWS/NMC/CAC	301/763-8227	A.LEETMAA
Bob Fish	NOAA/NESDIS/ NODC	202/673-5513	NODC.WDCA FISH::NODC*
Bob Molinari	NOAA/AOML	305/361-4344	AOML.MIAMI
Chris Noe	NOAA/NESDIS/ NODC	202/673-5636	NODC.WDCA NOE::NODC*
Darrell Knoll	NOAA/NESDIS/ NODC	202/673-5518	NODC.WDCA KNOLL::NODC*
Doug Hamilton	NOAA/NESDIS/ NODC	202/673-5636	NODC.WDCA DOUGLAS::NODC*
Eugene Molinelli	Planning Systems, Inc.	703/734-3439	E.MOLINELLI
Gregory Withee	NOAA/NESDIS/ NODC	202/673-5594	NODC.WDCA WITHEE:NODC*
Gregory Han	INO	601/688-2567	IND::HAN
Irving Perlroth	NOAA/NESDIS/ NODC	202/673-5598	NODC.WDCA PERLROTH::NODC*
J.R. Wilson	MEDS	613/990-0264	R.WILSON.MEDS
James Churgin	NOAA/NESDIS/ NODC	202/673-5546	NODC.WDCA J.CHURGIN DELCON::CHURGIN*
Paul Friday	NOAA/NOS	202/377-5804	P.FRIDAY
Peter Topoly	NOAA/NESDIS/ NODC	202/673-5591	NODC.WDCA TOPOLY::NODC*
Phil Hadsell	NOAA/NESDIS/ NODC	202/673-5600	NODC.WDCA HADSELL::NODC*
Roger W. Reeves	NOAA/OAR/CAR	301/443-5381	M.HALL
Roger Bauer	Compass Systems, Inc.	619/270-5230	
Roy Jenne	NCAR	303/497-1215	R.JENNE
Scott Woodruff	NOAA/NESDIS/ ERL	303/320-6747	NOAA.CRD
Stephen Pazan	SIO	619/534-3988	S.PAZAN
Steven Patterson	NOAA/NESDIS/ NODC	202/674-5601	NODC.WDCA PATTERSON::NODC*
Steven Worley	NCAR	303/497-1248	R.JENNE
Syd Levitus	NOAA/GFDL	609/452-1215	S.LEVITUS
Thomas Kinder	ONR	202/696-4441	T.KINDER
Warren White	SIO	619/534-4826	W.WHITE
William McQueary	NAVOCEANO	601/688-5675	
William Teague	NORDA	601/688-4734	B.TEAGUE

* SPAN

Appendix C

List of Acronyms

ADCP	Acoustic Doppler Current Profiler
AOML	Atlantic Oceanographic and Meterological Laboratory
BUFR	Binary Universal Format for Records
CD-ROM	Compact disk with read-only memory
COADS	Comprehensive Ocean-Atmosphere Data Set
CUDB	Continuously Updated Data Base
ERL	Environmental Research Laboratory
FNOC	Fleet Numerical Oceanography Center
GEOSAT	Geodetic Satellite
GFDL	Geophysical Fluid Dynamics Laboratory
GLORIA	Geological Long Range Inclined Asdic
GOFS	Global Ocean Flux Study
GSFC	Goddard Space Flight Center
GTS	Global Telecommunications System
IGOSS	Integrated Global Ocean Services System
INO	Institute for Naval Oceanography
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data Exchange
JEDA	Joint Environmental Data Analysis
MBT	Mechanical Bathythermograph
MEDS	Marine Environmental Data Service of Canada
MIT	Massachusetts Institute of Technology
MOODS	Master Oceanographic Observations Data Set
NAS	National Academy of Sciences
NASA	National Aeronautics and Space Administration
NAVOCEANO	Naval Oceanographic Office
NCAR	National Center for Atmospheric Research
NCDS	NASA Climate Data System
NESDIS	National Environmental Satellite, Data, and Information Service
NMC	National Meterological Center
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center
NOS	National Ocean Service
NOSIE	NODC Ocean Science Information Exchange
NWS	National Weather Service
OCEAN	Ocean Climate Environmental Analysis News
ONR	Office of Naval Research
OPC	Ocean Products Center
PSI	Planning Systems Incorporated
QC	Quality control
QUIPS	Quality Improvement Profile System
SIO	Scripps Institution of Oceanography
SOC	Southern Ocean Center
SONIC	SPAN Ocean Network Information Center
SPAN	Space Physics Analysis Network
SST	Sea surface temperature
STD	Salinity-temperature-depth profiler

TOGA	Tropical Ocean and Global Atmosphere
T-S	Temperature - salinity
UCAR	University Corporation for Atmospheric Research
VOS	Voluntary Observing Ship
WDC	World Data Center
WOCE	World Ocean Circulation Experiment
XBT	Expendable Bathythermograph
XCTD	Expendable conductivity-temperature-depth profiler

APPENDIX D

PRESENTATION OF VIEWGRAPHS

BY

GREGORY W. WITHEE

National Oceanographic Data Center

- **MISSION:** TO ACQUIRE, PROCESS, STORE, DISSEMINATE, AND EXCHANGE GLOBAL OCEANOGRAPHIC DATA AND PROVIDE DATA MANAGEMENT SERVICES

- **1960** ESTABLISHED AS INTERAGENCY FACILITY UNDER ADMINISTRATION OF U.S. NAVY

- **1970** TRANSFERRED TO NOAA'S ENVIRONMENTAL DATA AND INFORMATION SERVICE

- **1982** INCORPORATED INTO NOAA'S NATIONAL ENVIRONMENTAL SATELLITE, DATA, AND INFORMATION SERVICE

National Oceanographic Data Center

Characteristics

- ✓ **INTERDISCIPLINARY** - PHYSICS, CHEMISTRY, BIOLOGY
- ✓ **INTERAGENCY** - NASA, NAVY, NSF, DOI, EPA, DOT, DOS
- ✓ **INTERNATIONAL** - IOC, WMO, ICES, ICSU

OCEAN DATA MANAGEMENT

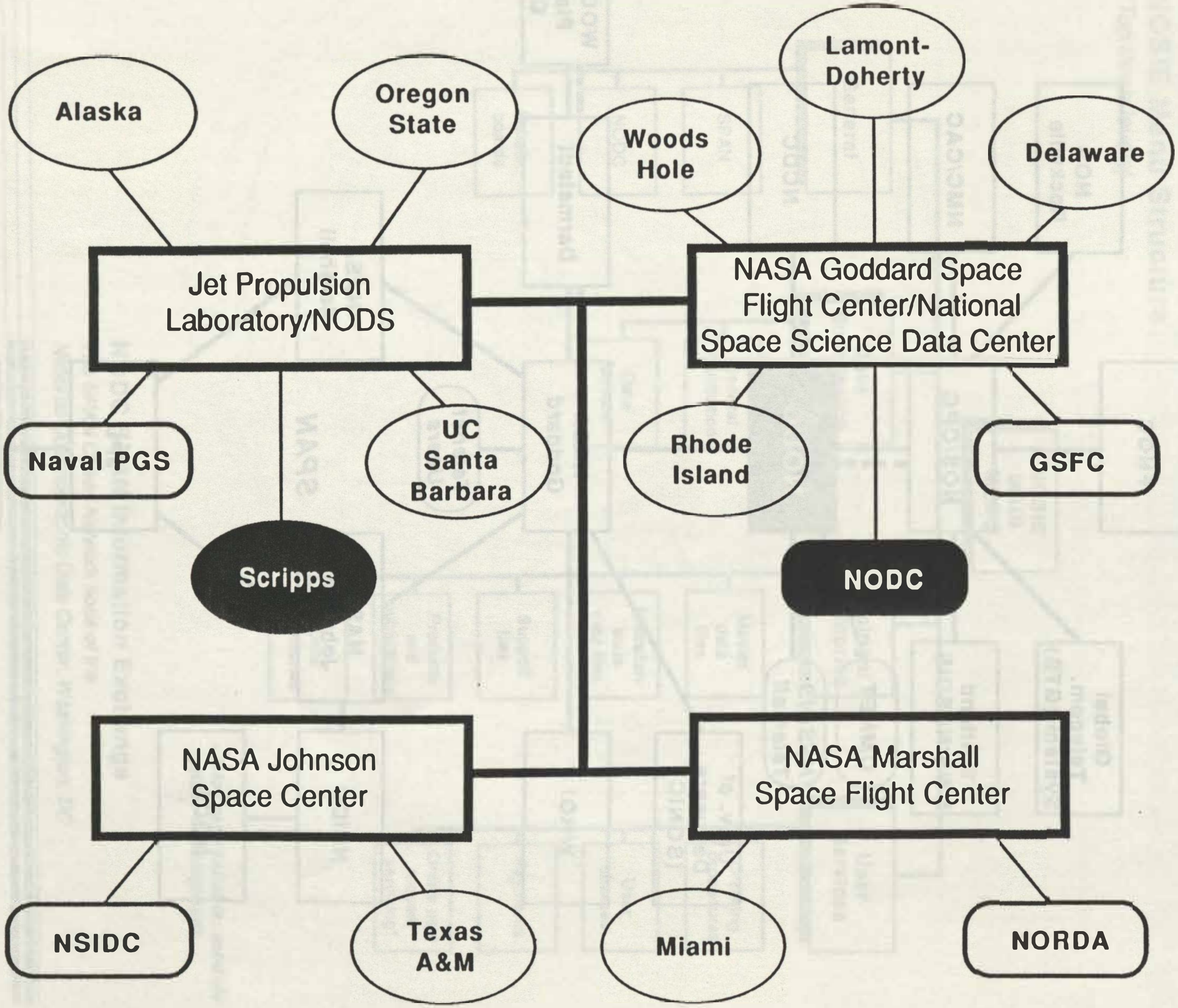
The ***strategic*** organization and operation
of
data handling functions

CONTROL	DATA HANDLING FUNCTIONS
• ADVISORY GROUP(S)	OBSERVE COLLECT TRACK
• REQUIREMENTS	ACQUIRE QUALITY CONTROL VALIDATE
• PRIORITIES	ARCHIVE ACCESS
• STANDARDS	DISTRIBUTE NETWORK

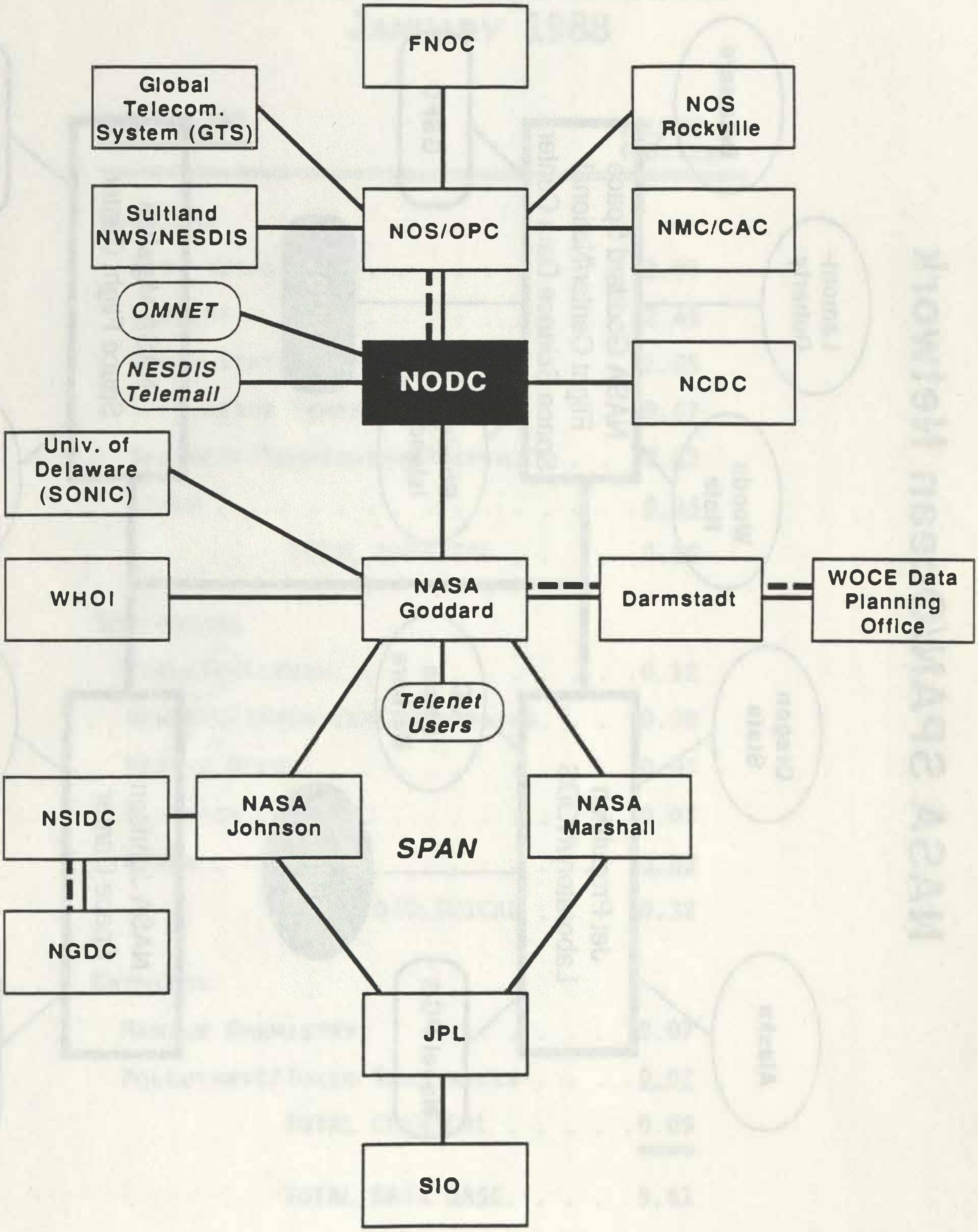
MAJOR NODC DATA HOLDINGS
JANUARY 1988

DISCIPLINE	VOLUME (GBYTES)
PHYSICAL	
BUOY (WIND/WAVE)	3.95
CURRENT	2.45
OCEAN STATION	1.25
SUBSURFACE TEMPERATURE	0.67
SALINITY/TEMPERATURE/DEPTH	0.83
OTHER	0.06
TOTAL PHYSICAL	9.20
BIOLOGICAL	
FISH/SHELLFISH	0.12
BENTHIC/INTERTIDAL ORGANISMS	0.09
MARINE BIRDS	0.05
PLANKTON	0.03
OTHER	0.03
TOTAL BIOLOGICAL	0.32
CHEMICAL	
MARINE CHEMISTRY	0.07
POLLUTANTS/TOXIC SUBSTANCES	0.02
TOTAL CHEMICAL	0.09
TOTAL DATA BASE	9.61

NASA SPAN/Ocean Network

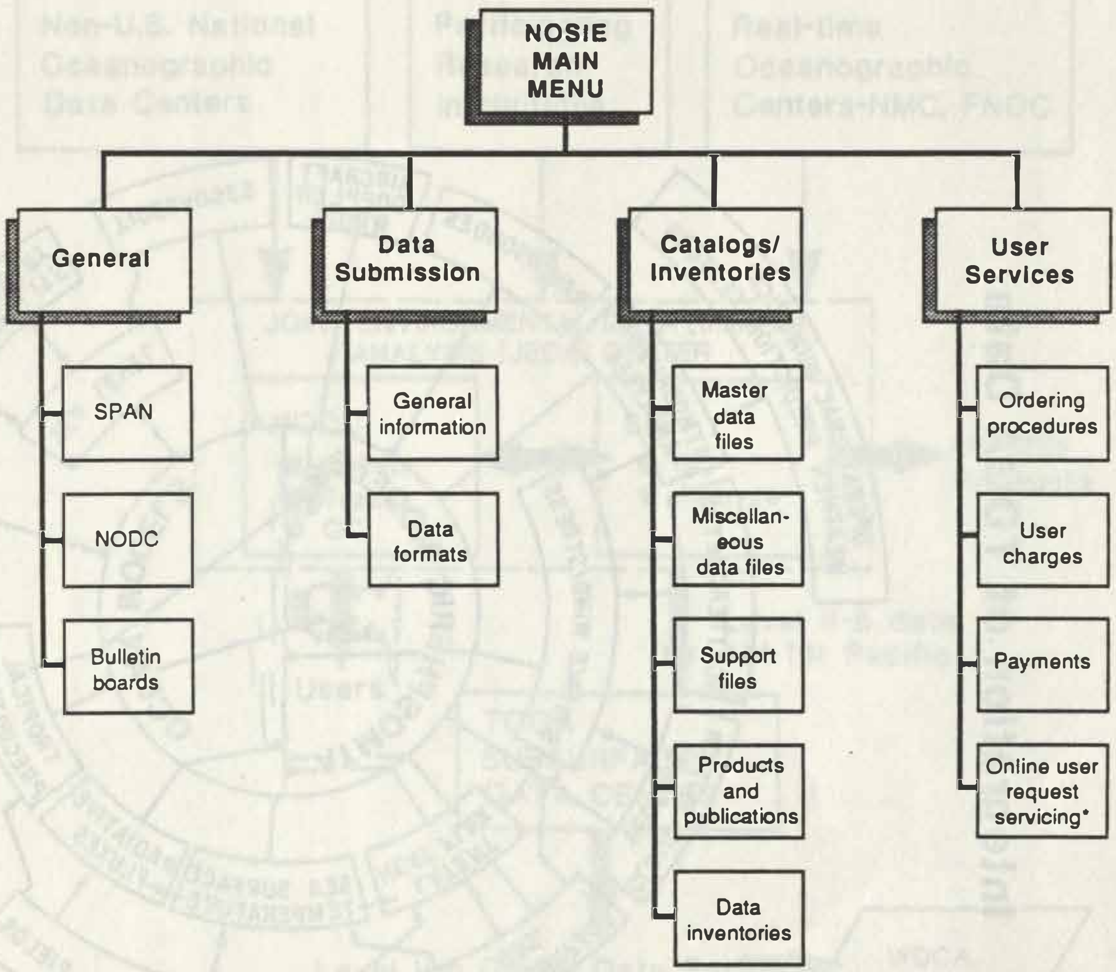


NODC CONNECTIVITY



NOSIE Menu Structure

(Top three levels)

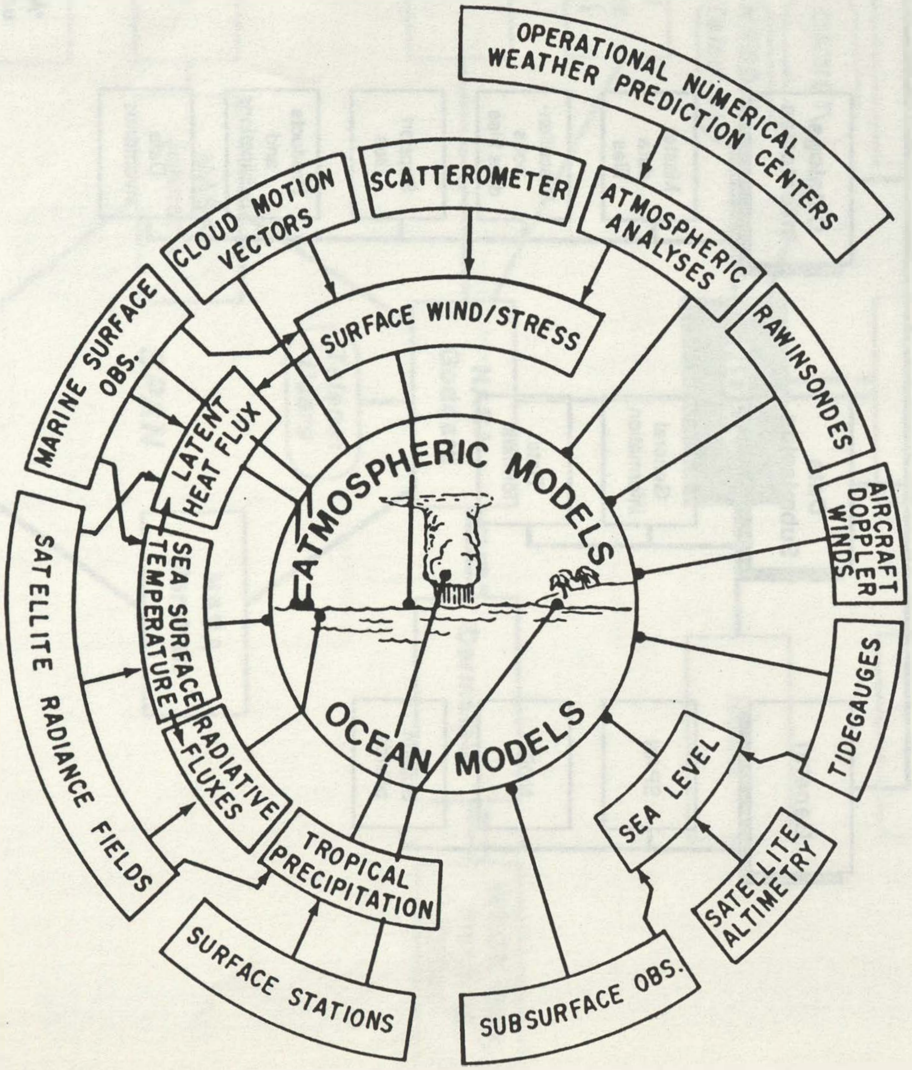


Not yet available; module under development.

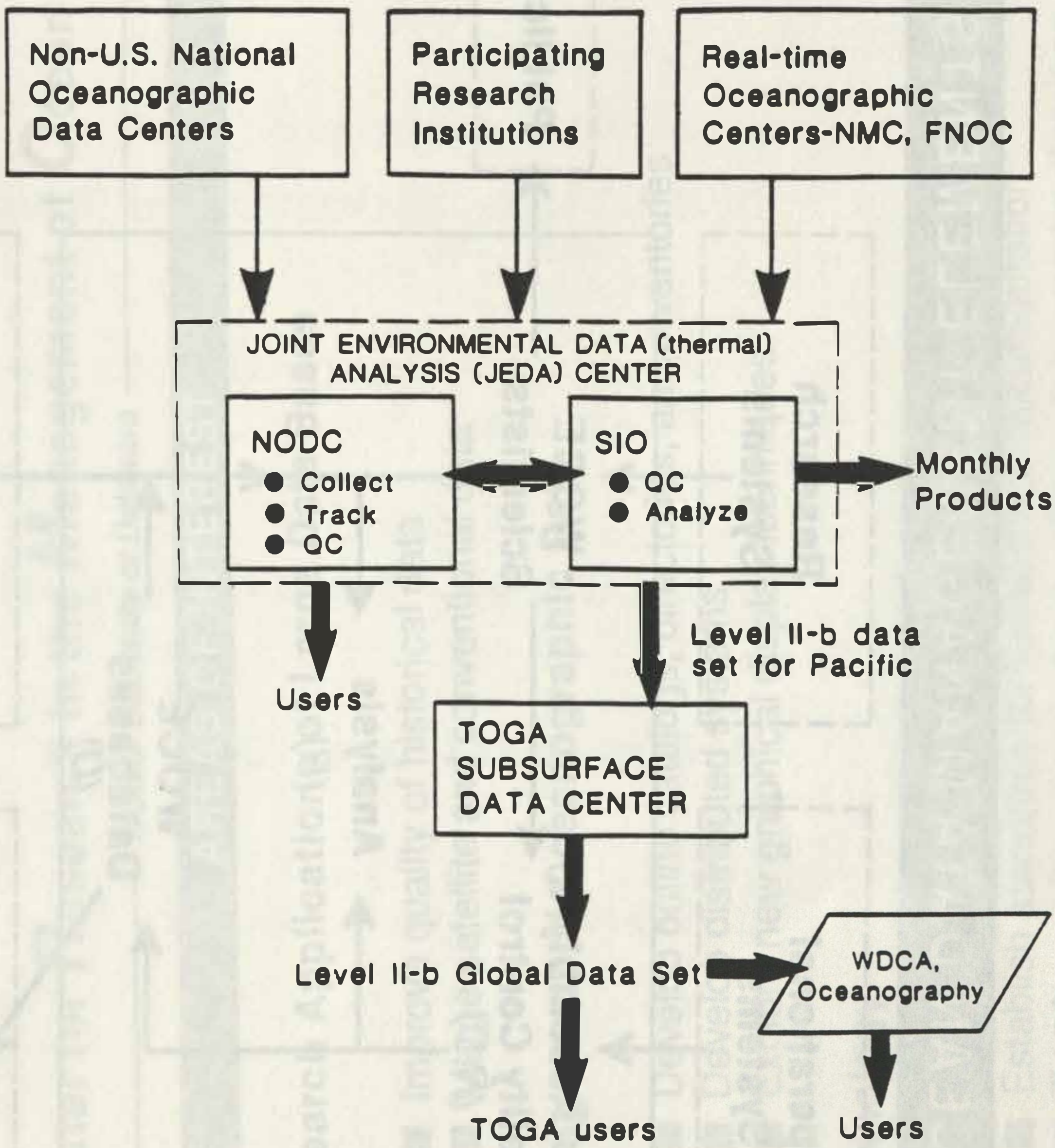
NODC SPAN Information Exchange

The SPAN Ocean Network node of the National Oceanographic Data Center, Washington, DC

Integration of TOGA Data

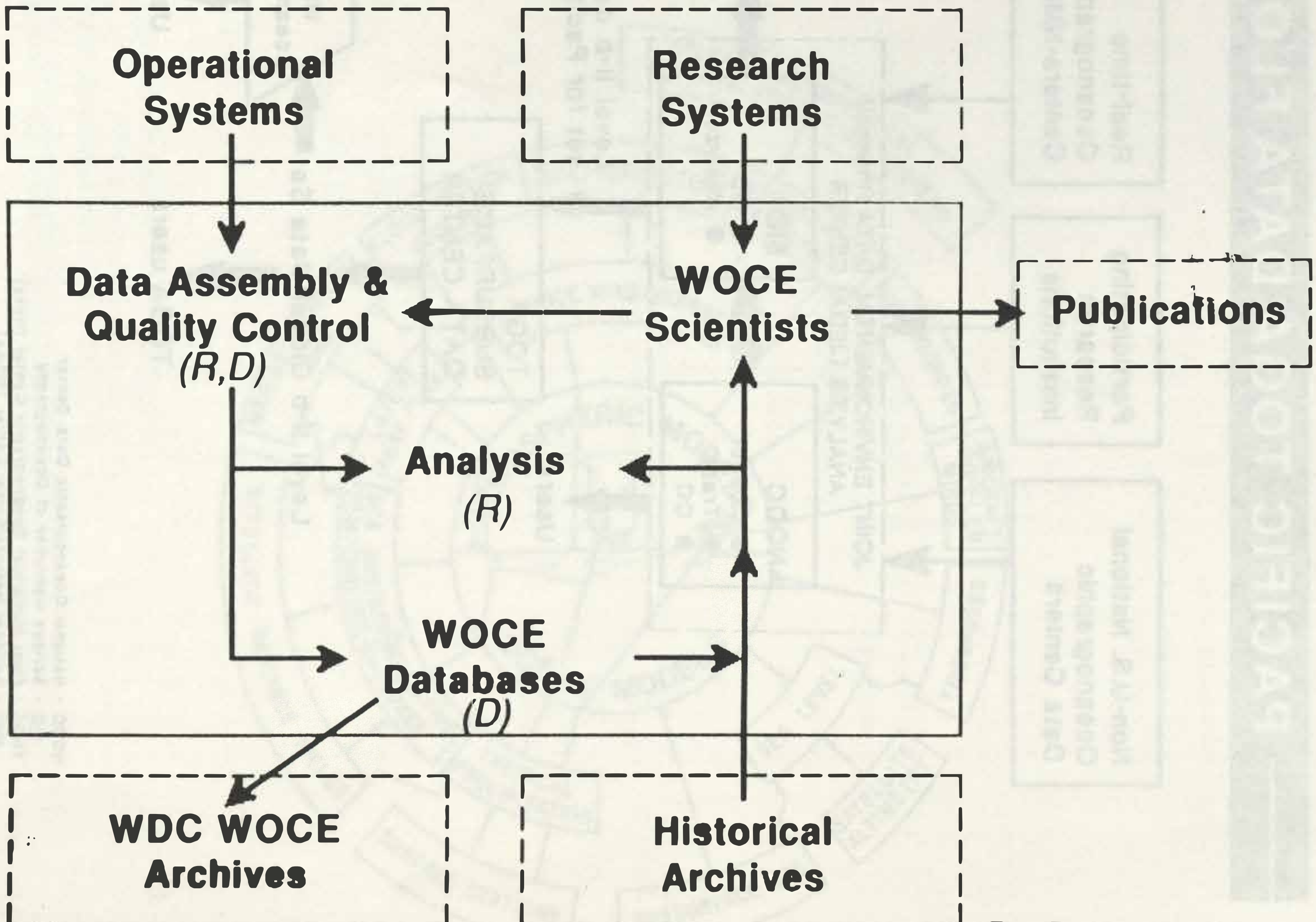


PACIFIC TOGA DATA FLOW



NODC - National Oceanographic Data Center
SIO - Scripps Institution of Oceanography
FNOC - Fleet Numerical Oceanography Center (Navy)
NMC - National Meteorological Center (NOAA)
WDCA - World Data Center A, Oceanography

OCEAN DATA MANAGEMENT ELEMENTS



R = Research Organization
D = National Data Center

JCRMOD

Joint Center for Research in the Management of Ocean Data

NODC/ University of Delaware

AREAS OF INTEREST

Research Applications of Large Data Bases

- Improve quality of historical data
- Merge satellite and conventional data

Management of Oceanographic Data

- Develop online catalogs, directories, and inventories
- Develop distributed systems
- Explore new graphical display techniques

Technology

- Experiment with effective use of computer networks
- Establish standards for formats and documentation
- Test feasability of "standard" software for common algoritms

JOINT ARCHIVE FOR SEA LEVEL DATA (JASL)

- O NODC with University of Hawaii
- O Maintain Expert Analysis and Processing
- O Establish a Permanent Archive
- O Expand Sea Level Data Base
- O Support Global Change

PRESENTATION OF VIEWGRAPHS

BY

PAUL FRIDAY

QUALITY CONTROL OF MARINE DATA CENTER

AT THE

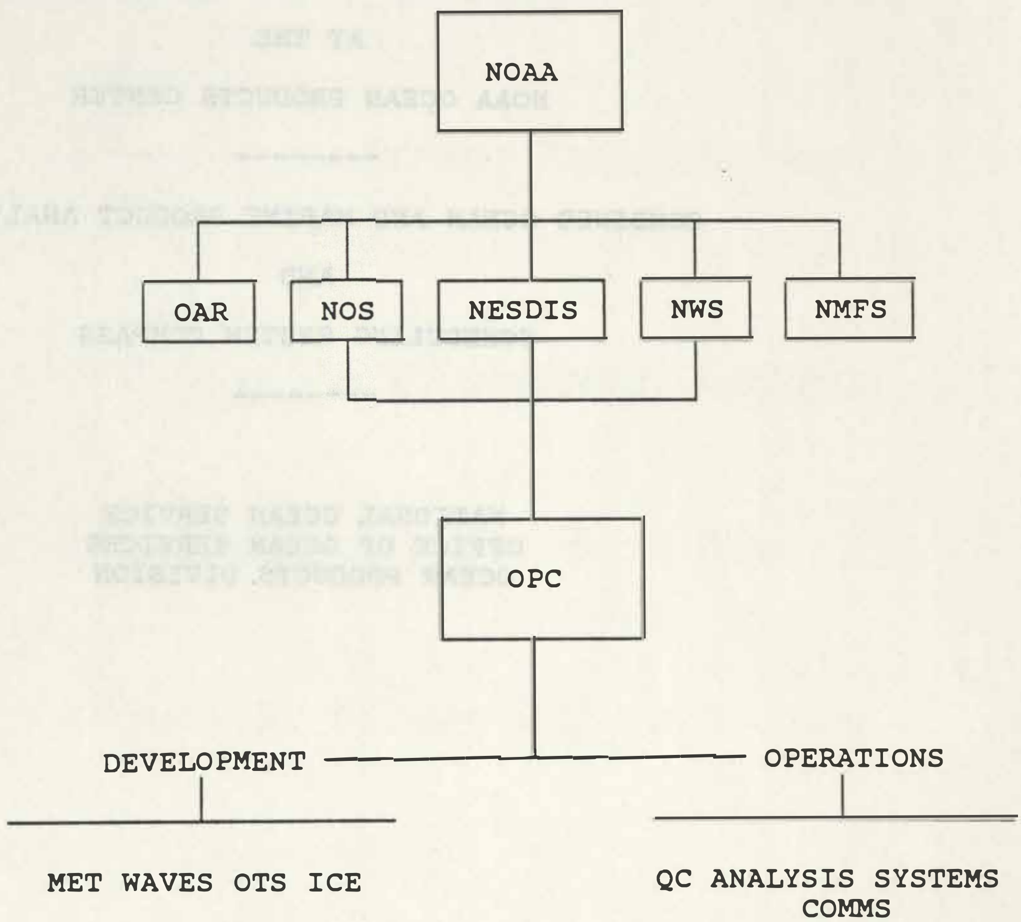
NOAA OCEAN PRODUCTS CENTER

COMBINED OCEAN AND MARINE PRODUCT ANALYSIS

AND

SCHEDULING SYSTEM COMPASS

NATIONAL OCEAN SERVICE
OFFICE OF OCEAN SERVICES
OCEAN PRODUCTS DIVISION


NOAA
OAR
NOS
NESDIS
NWS
NMFS
OPC
DEVELOPMENT
OPERATIONS
MET WAVES OTS ICE
QC ANALYSIS SYSTEMS
COMMS

OPC OPERATIONS

OBJECTIVES:

1. Provide for operational quality control of all marine observations which impact NOAA's ability to generate forecast guidance.

2. Ensure dissemination of marine data and products on the Global Telecommunications System (GTS), the Family of Services, and NOAAPORT.

3. Monitor all marine forecast guidance disseminated from the OPC/NMC complex.

4. Interactively generate global and regional marine forecast guidance products and analyses.

QUALITY CONTROL OPERATIONS

Real-time QC of ship reports, drifting buoys, and XBT/TESAC data, prior to model assimilation and dissemination, using interactive processing techniques.

OPC microVAX quality control system interfaced to NMC/OSO and OAG/FNOC mainframes allows transfer of data sets and model output.

QC unit collocated and integrated with the NMC Meteorological Operations Division.

QC unit has been operational since 1 October 1987 using NMC QC system. MicroVAX system becomes operational January 1988.

QUALITY CONTROL - DATA SETS

Phase I (FY88)

Ship Meteorological Reports	- pressure/winds
Drifting Buoy Data	- pressure/winds
XBT/TESAC Reports	- temperature/salinity

Phase II (FY89)

Ship Meteorological Reports	- all parameters
Drifting Buoy Data	- all parameters
Moored Buoy Reports	- all parameters
C-MAN Reports	- all parameters
XBT/TESAC Reports	- temperature/salinity
Real-Time Water Level Data	- all parameters
Acoustic Doppler Current Profiles	- currents

Phase III (FY90 and beyond)

All of Phase II, plus:

Ocean Satellite Data (scatterometer...microwave...)

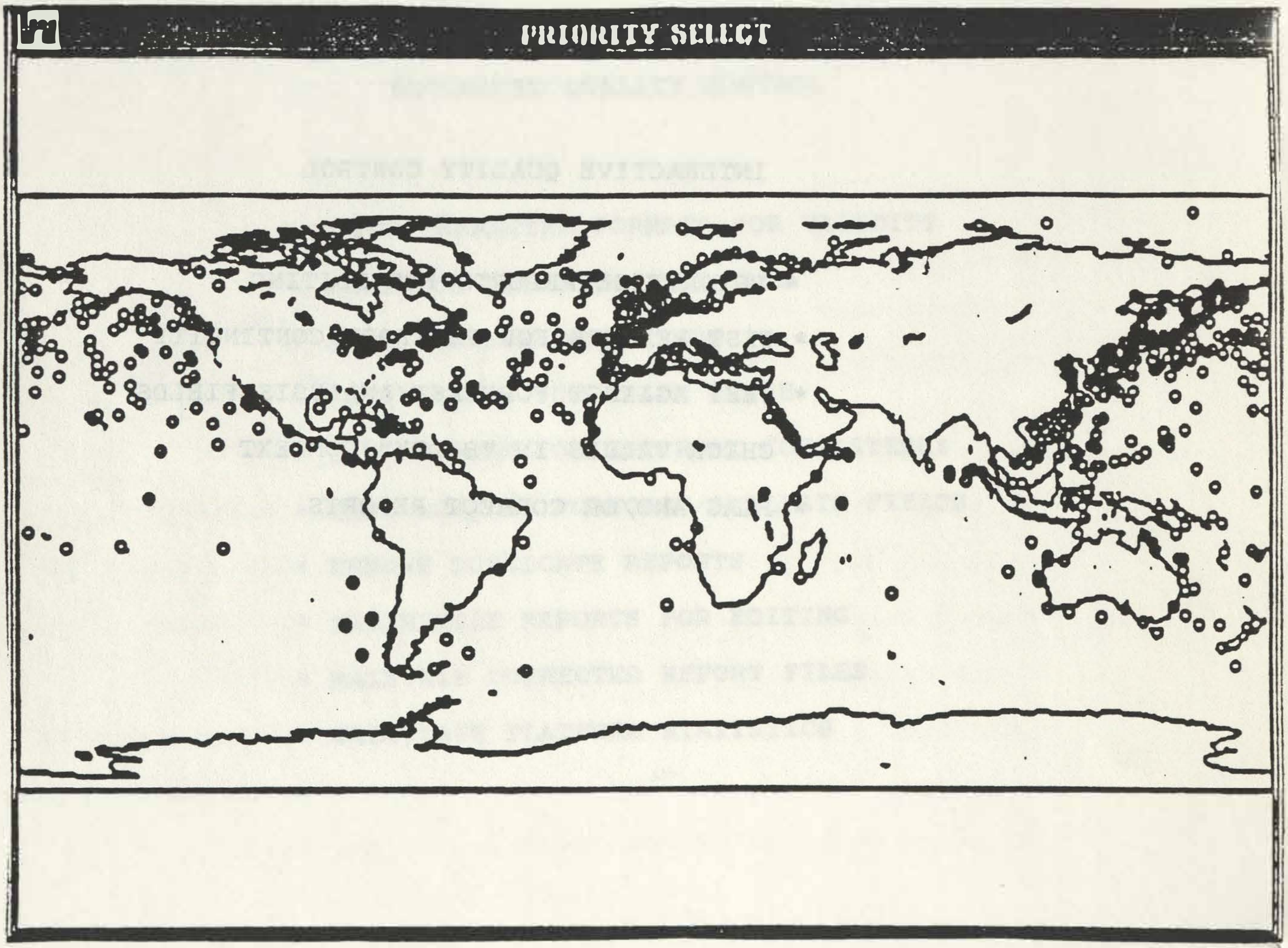
Meteorological Satellite Data (marine soundings)

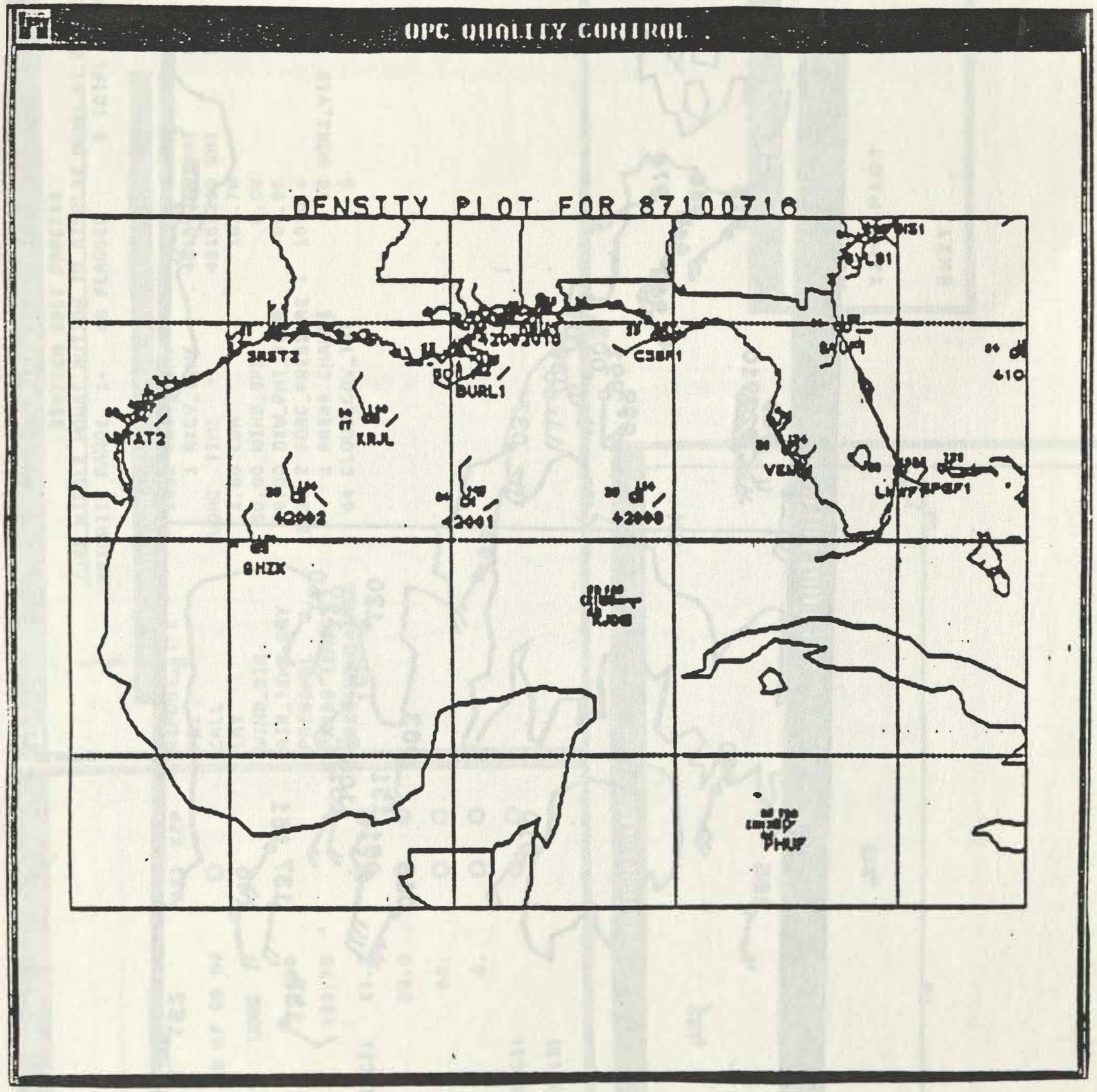
AUTOMATED QUALITY CONTROL

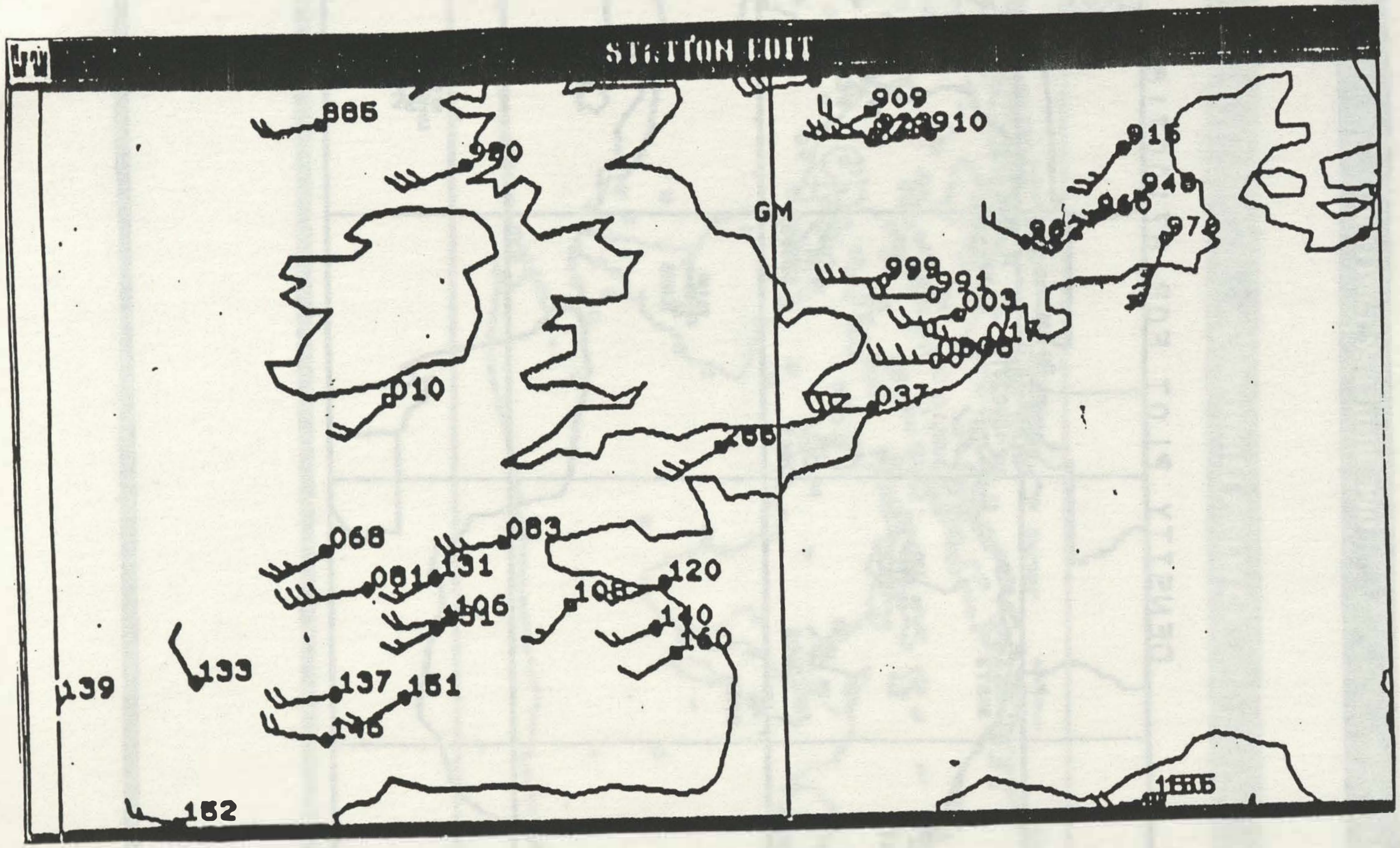
* CHECK CHARACTER FORMATS FOR VALIDITY
* VALIDATE PLATFORM CALL SIGN
* APPLY SENSOR CORRECTION
* TEST VALUES FOR VALID RANGE
* TEST VALUES FOR INTERNAL CONSISTENCY
* TEST AGAINST FORECAST/ANALYSIS FIELDS
* REMOVE DUPLICATE REPORTS
* PRIORITIZE REPORTS FOR EDITING
* MAINTAIN CORRECTED REPORT FILES
* CALCULATE PLATFORM STATISTICS

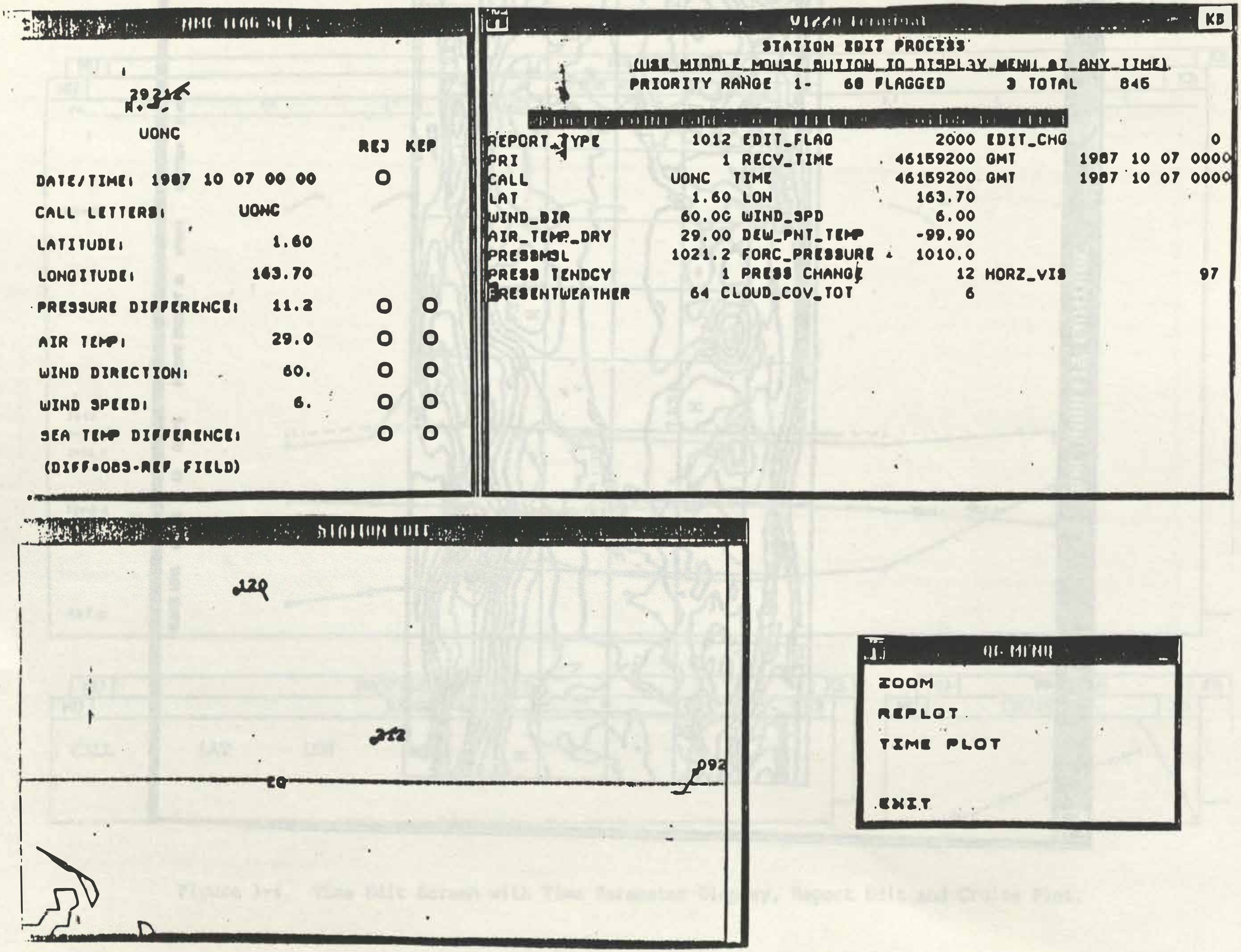
INTERACTIVE QUALITY CONTROL

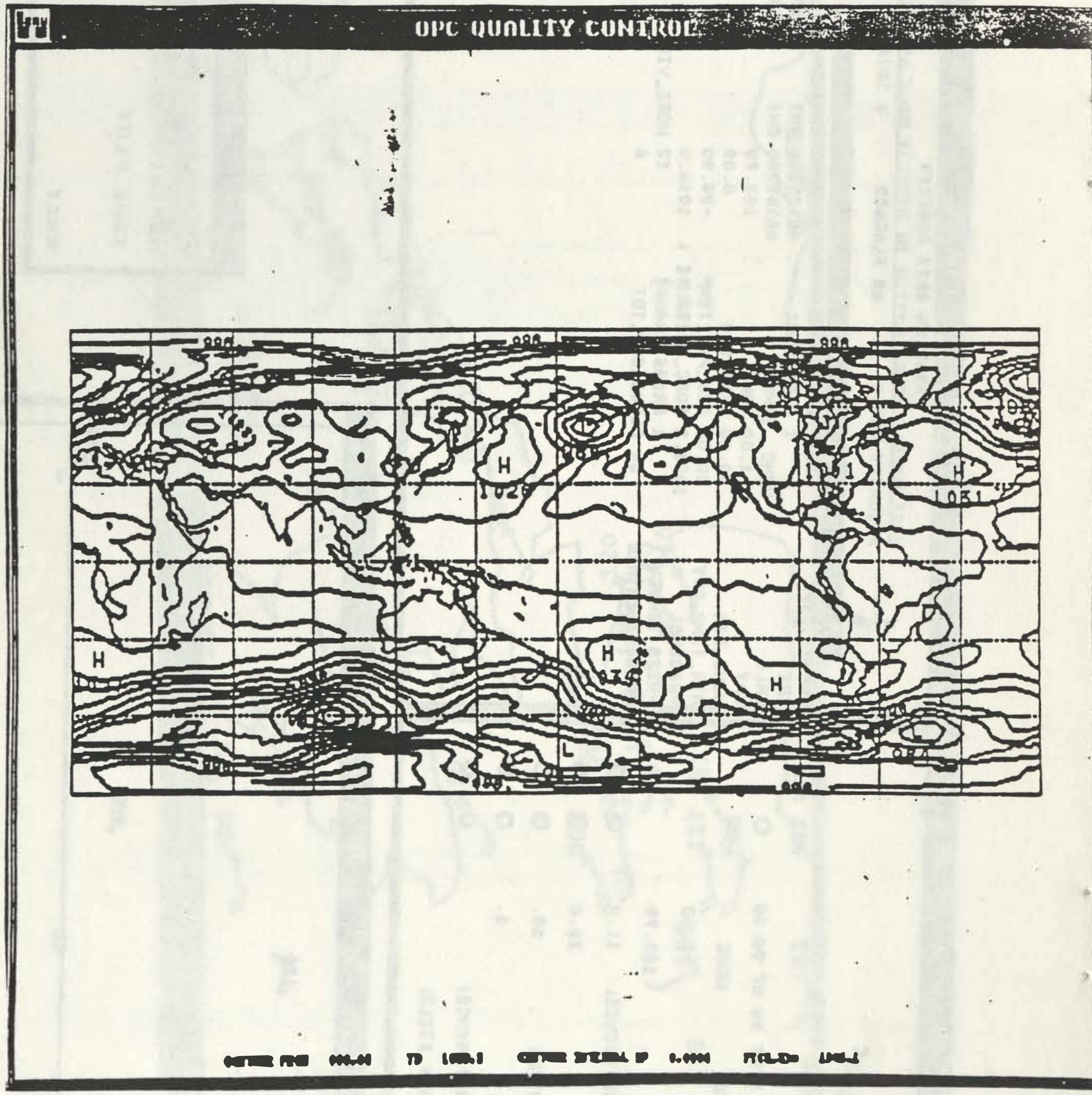
* PRIORITIZE REPORTS FOR EDITING
* TEST REPORTS FOR PLATFORM CONTINUITY
* TEST AGAINST FORECAST/ANALYSIS FIELDS
* CHECK VALUES IN "BUDDY" CONTEXT
* FLAG AND/OR CORRECT REPORTS


PRIORITY SELECT


OPC QUALITY CONTROL
DENSITY PLOT FOR 87100716
BURL1
42002
42001


STATION EDIT
885
909
910
916
948
GM
999
991
003
037
010
068
083
131
120
106
110
133
137
151
139
152


KB
STATION EDIT PROCESS
(USE MIDDLE MOUSE BUTTON TO DISPLAY MENU AT ANY TIME)
PRIORITY RANGE 1- 60 FLAGGED 3 TOTAL 846
REPORT_TYPE 1012 EDIT_FLAG 2000 EDIT_CHG 0
PRI 1 RECV_TIME 46159200 GMT 1987 10 07 0000
CALL UONC TIME 46159200 GMT 1987 10 07 0000
LAT 1.60 LON 163.70
WIND_DIR 60.00 WIND_SPD 6.00
AIR_TEMP_DRY 29.00 DEW_PNT_TEMP -99.90
PRESSMSL 1021.2 FORC_PRESSURE 1010.0
PRESS TENDCY 1 PRESS CHANGE 12 HORZ_VIS 97
PRESENTWEATHER 64 CLOUD_COV_TOT 6
UONC
REJ KEP
DATE/TIME: 1987 10 07 00 00
CALL LETTERS: UONC
LATITUDE: 1.60
LONGITUDE: 163.70
PRESSURE DIFFERENCE: 11.2
AIR TEMP: 29.0
WIND DIRECTION: 60.
WIND SPEED: 6.
SEA TEMP DIFFERENCE:
(DIFF=OBS-REF FIELD)
120
212
092
ZOOM
REPLOT
TIME PLOT
EXIT


OPC QUALITY CONTROL

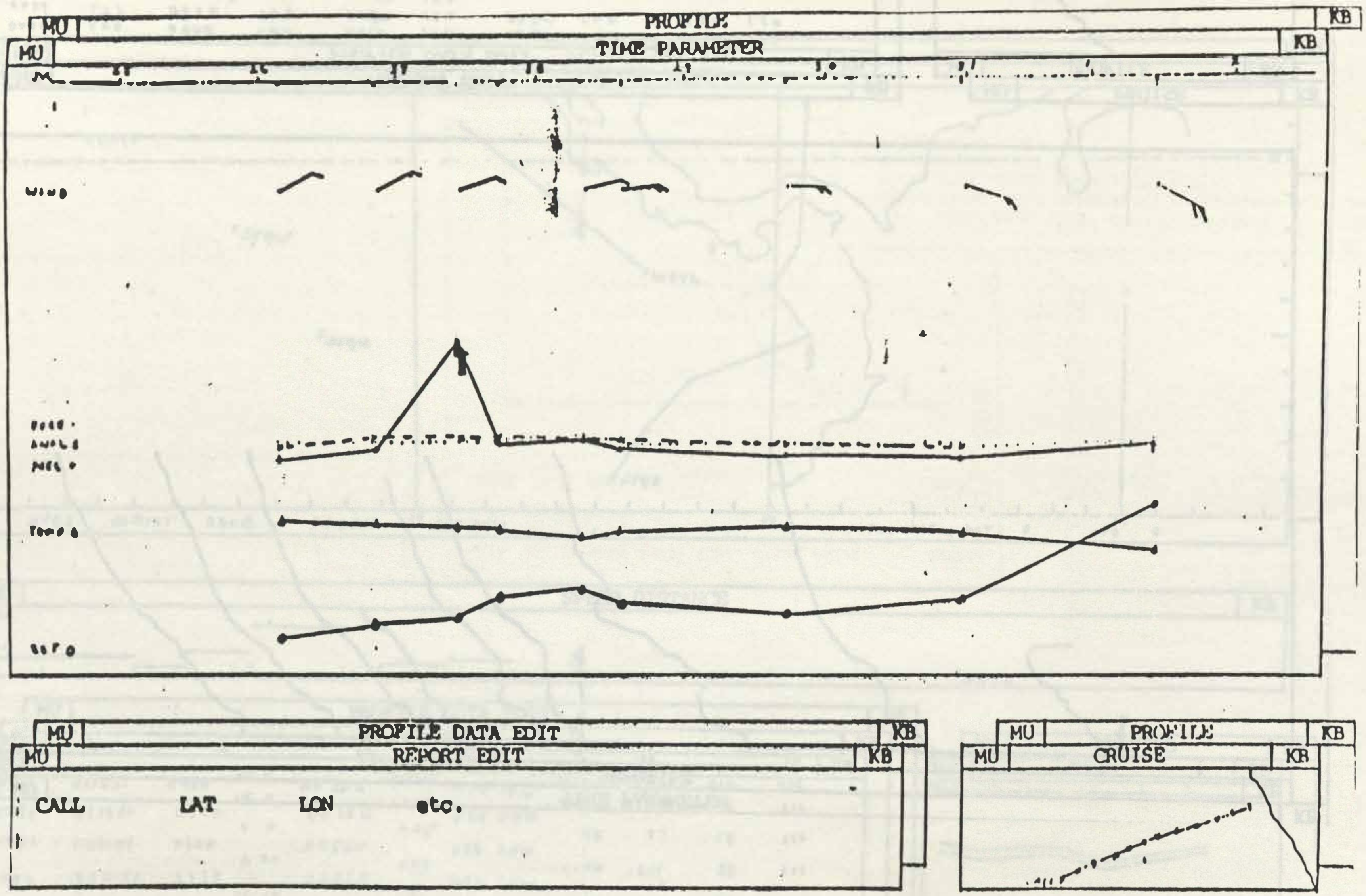


Figure 3-4. Time Edit Screen with Time Parameter Display, Report Edit and Cruise Plot.

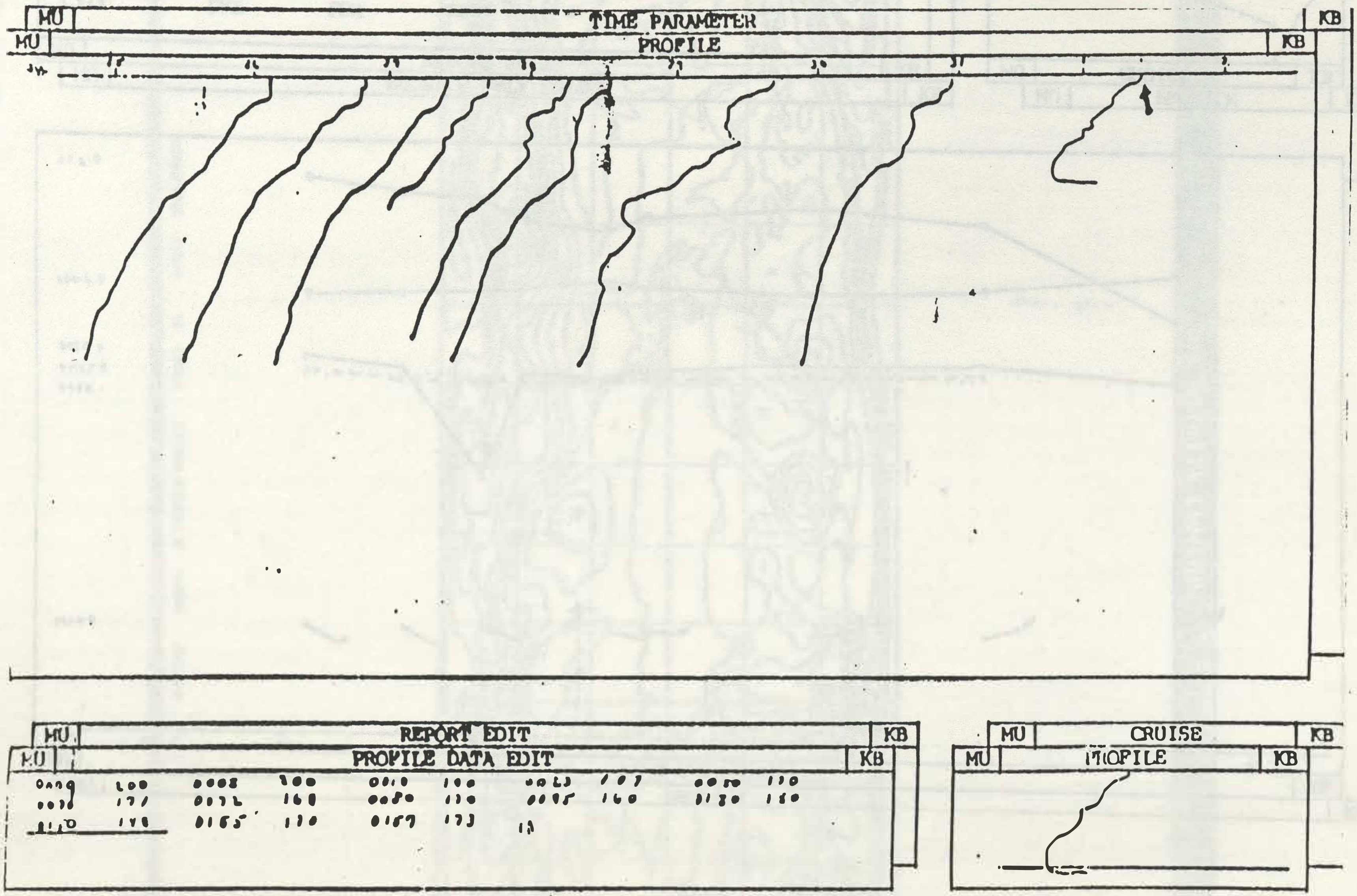


Figure 3-5. Time Edit Screen with Profile and Profile Edit Display.

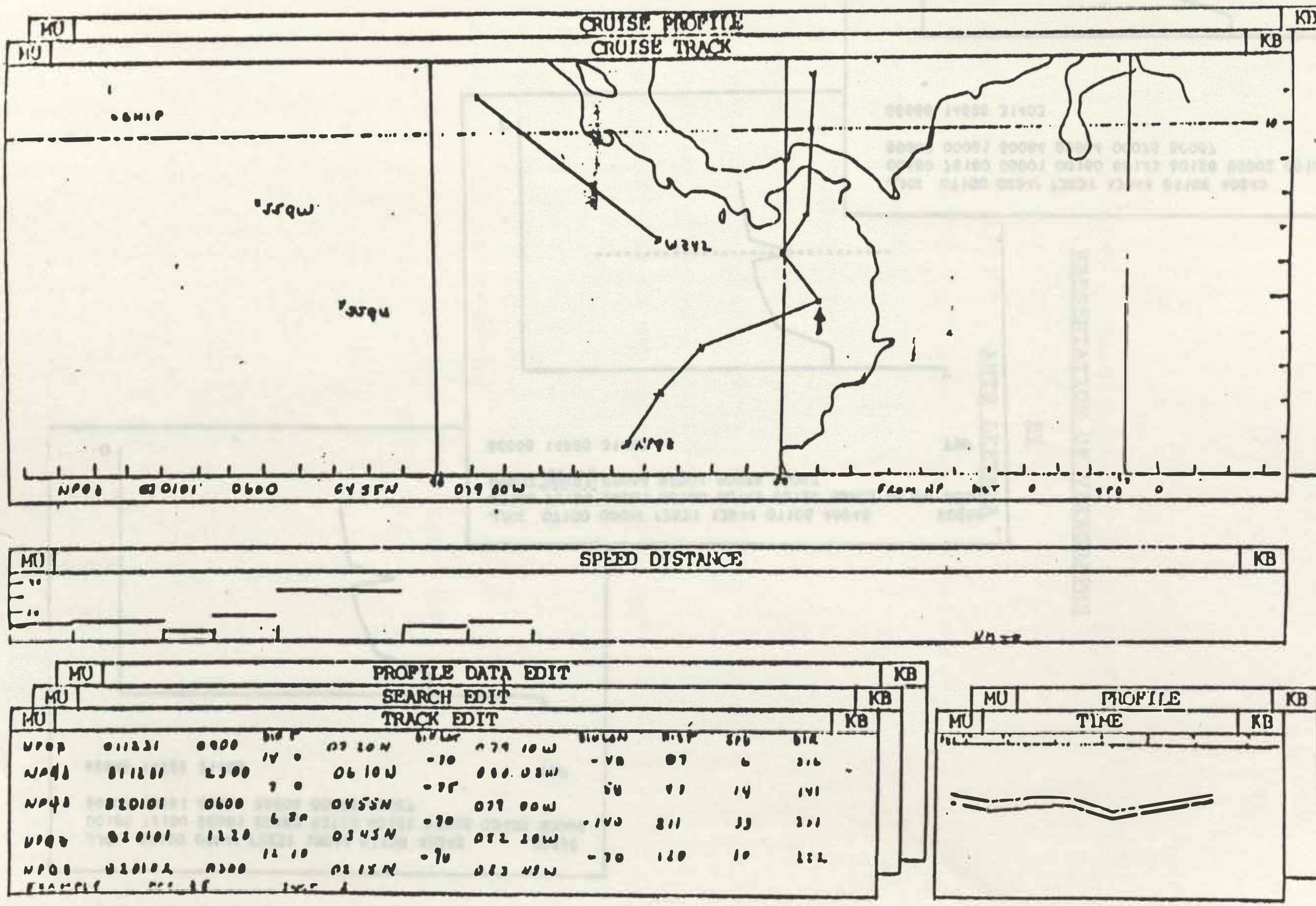
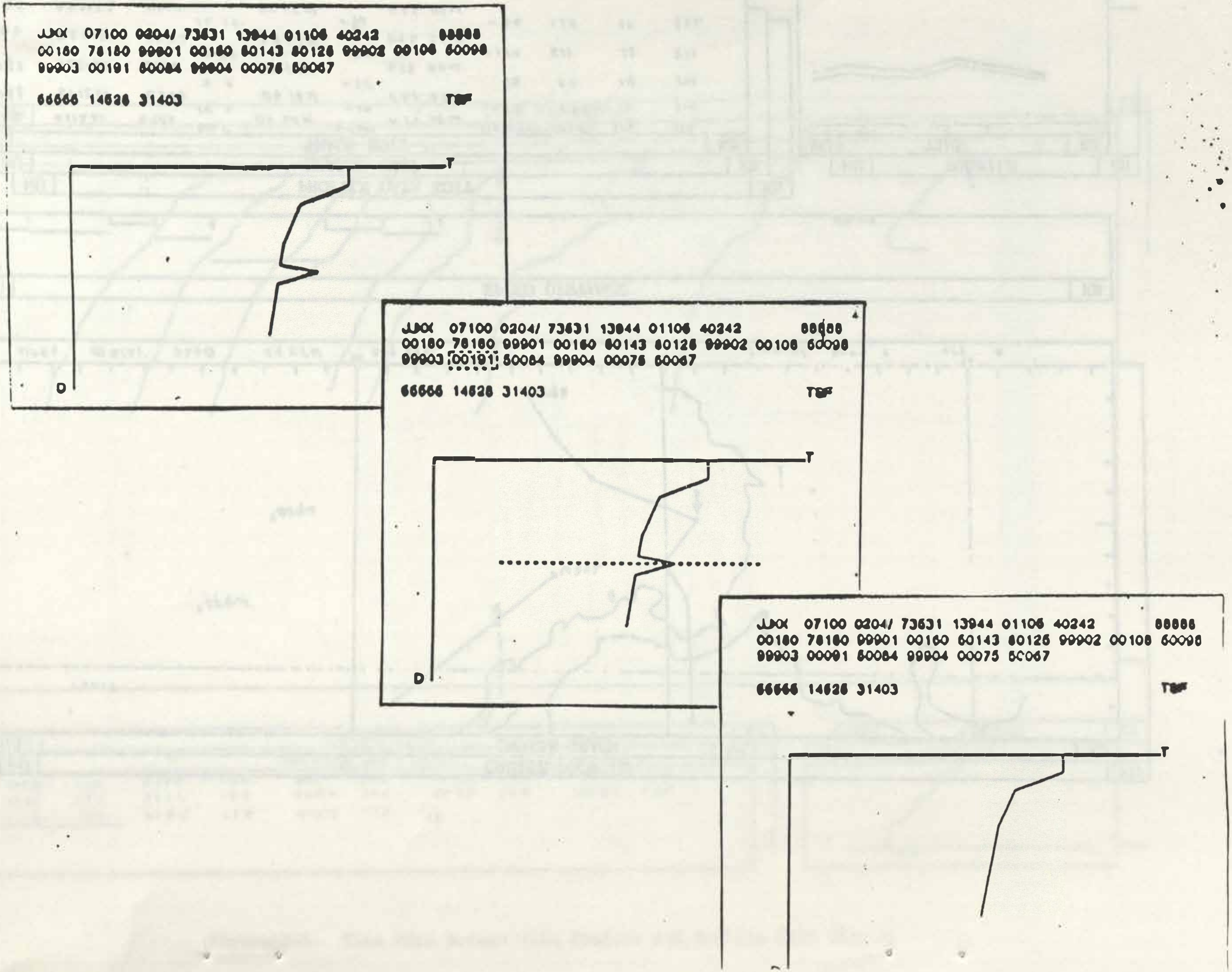


Figure 3-6. Cruise Edit Screen with Cruise Track, Track Edit and Time Windows.


JJXX 07100 0204/ 73531 13944 01106 40242 88888
00160 78160 99901 00160 60143 60126 99902 00106 60098
99903 00191 60084 99904 00075 60067
66666 14626 31403 TSF
T
D
JJXX 07100 0204/ 73531 13944 01106 40242 88888
00160 78160 99901 00160 60143 60126 99902 00106 60098
99903 00191 60084 99904 00075 60067
66666 14626 31403 TSF
T
D
JJXX 07100 0204/ 73531 13944 01106 40242 88888
00160 78160 99901 00160 60143 60126 99902 00106 60098
99903 00091 60084 99904 00075 60067
66666 14626 31403 TSF
T

PRESENTATION OF VIEWGRAPHS

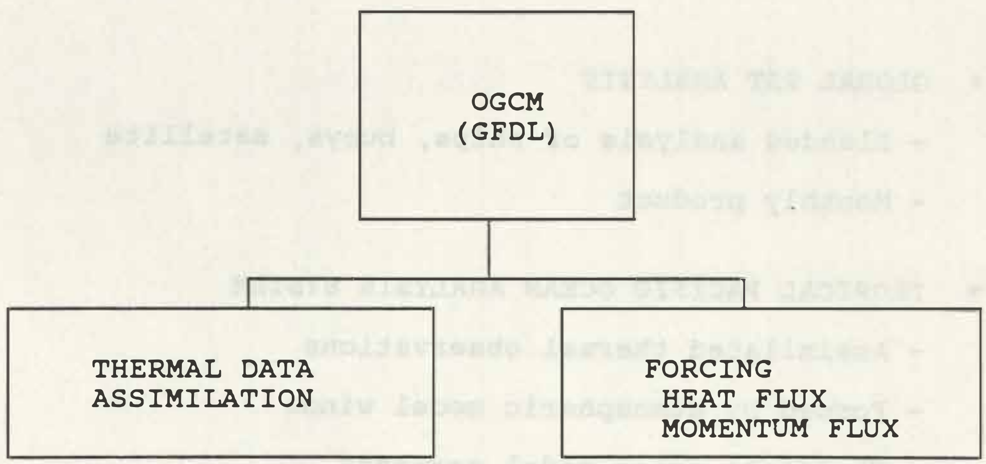
BY

ANTS LEETMAA

NWS OCEAN CLIMATE DIAGNOSTICS AND MODELING

* GLOBAL SST ANALYSIS
    - Blended analysis of ships, buoys, satellite
    - Monthly product
* TROPICAL PACIFIC OCEAN ANALYSIS SYSTEM
    - Assimilated thermal observations
    - Forced by atmospheric model winds
    - Numerical ocean model response
    - Provides fields of sea level, temperature, velocity (figure)
* PLANS
    - Weekly analysis
    - Expand to Pacific Basin
    - Assimilation of sea level and current data

OCEAN ANALYSIS SYSTEM



CAC/NMC

GFDL OCEAN MODEL

PACIFIC BASIN 30 S TO 50 N

HORIZONTAL RESOLUTION

1/3 DEGREE MERIDIONAL 10 N - 10 S

1 DEGREE ZONAL

27 VERTICAL LEVELS

FLAT BOTTOM

1 HOUR TIMESTEP

20 MIN. CY205 PER MONTH

ASSIMILATION PROCEDURES

MONTHLY

UNIVARIATE - TEMPERATURE

MODEL BASIS FOR FIRST GUESS FIELD

VARIOUS QUALITY CONTROL STEPS

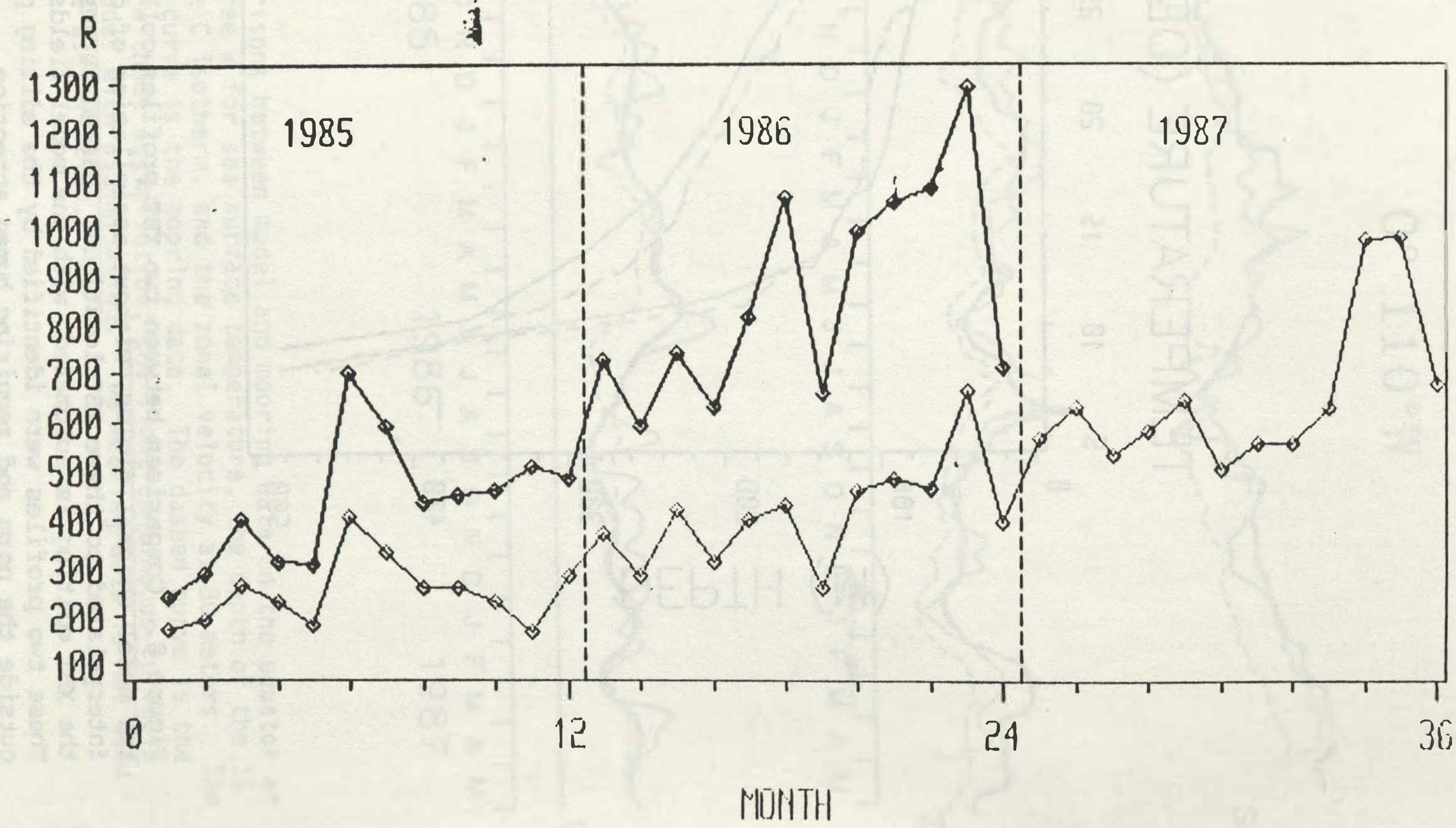
- pre-editing, buddy check
- regional 3 sigma check

INDEPENDENT OBJECTIVE MAPPING EACH LEVEL

SST, MIXED LAYER DEPTHS DONE DIFFERENTLY

XBTs AVAILABLE PER MONTH

BET WEEN 20S – 20N



BLACK – FROM GTS RED – DELAYED PLUS GTS

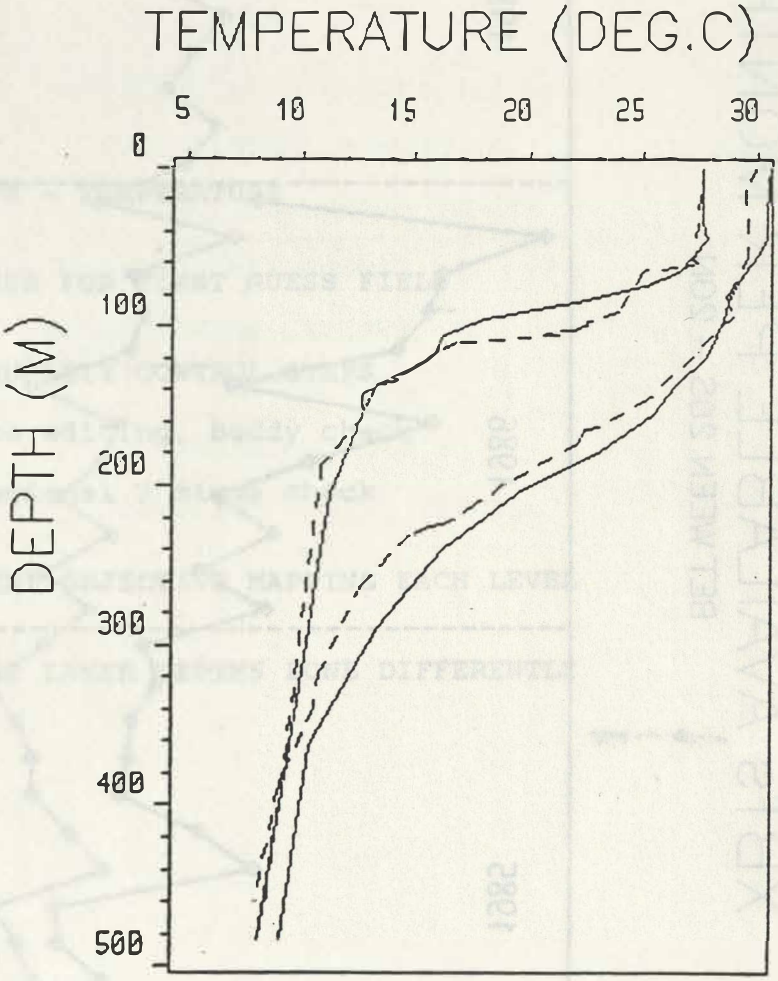


Figure 8. Comparison between two XBT profiles collected in January 1988 and the model thermal fields, without assimilation of data for this month, interpolated to these XBT locations. Dashed curves are the XBTs. Much of the XBT editing is automated, with the model fields being used for reference. These two profiles were identified by the editing procedures as being outside the norm and requiring human attention.

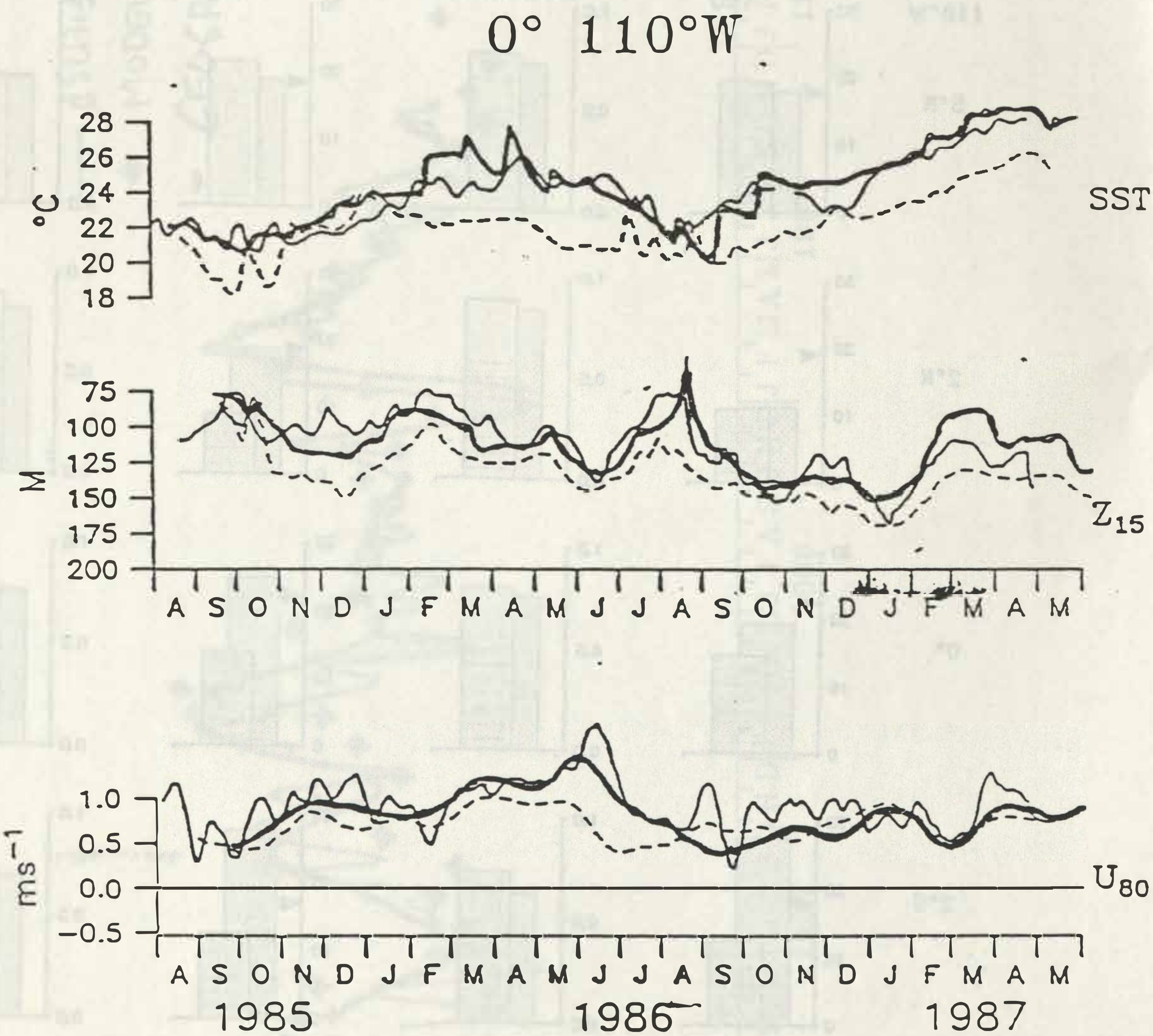
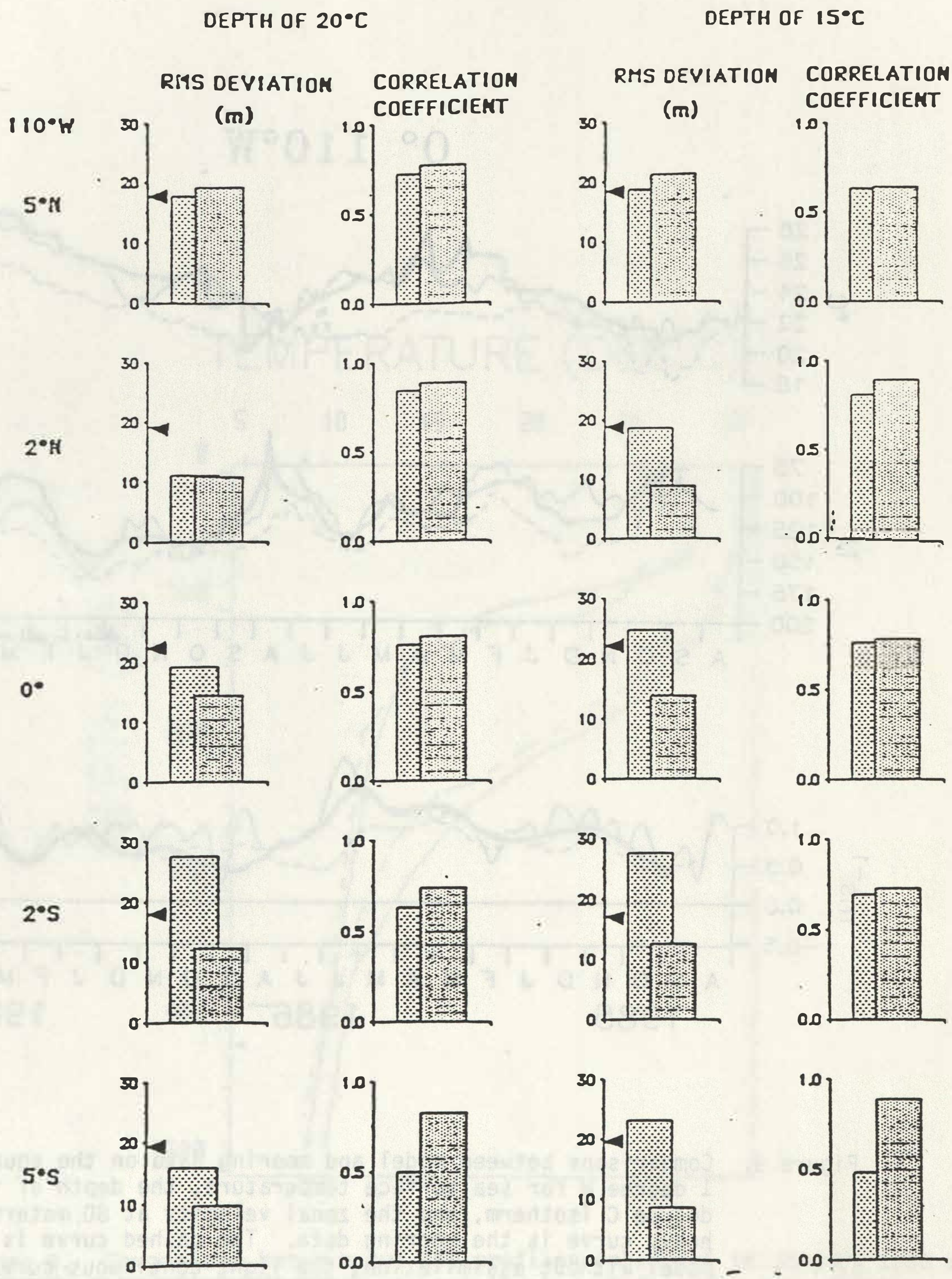
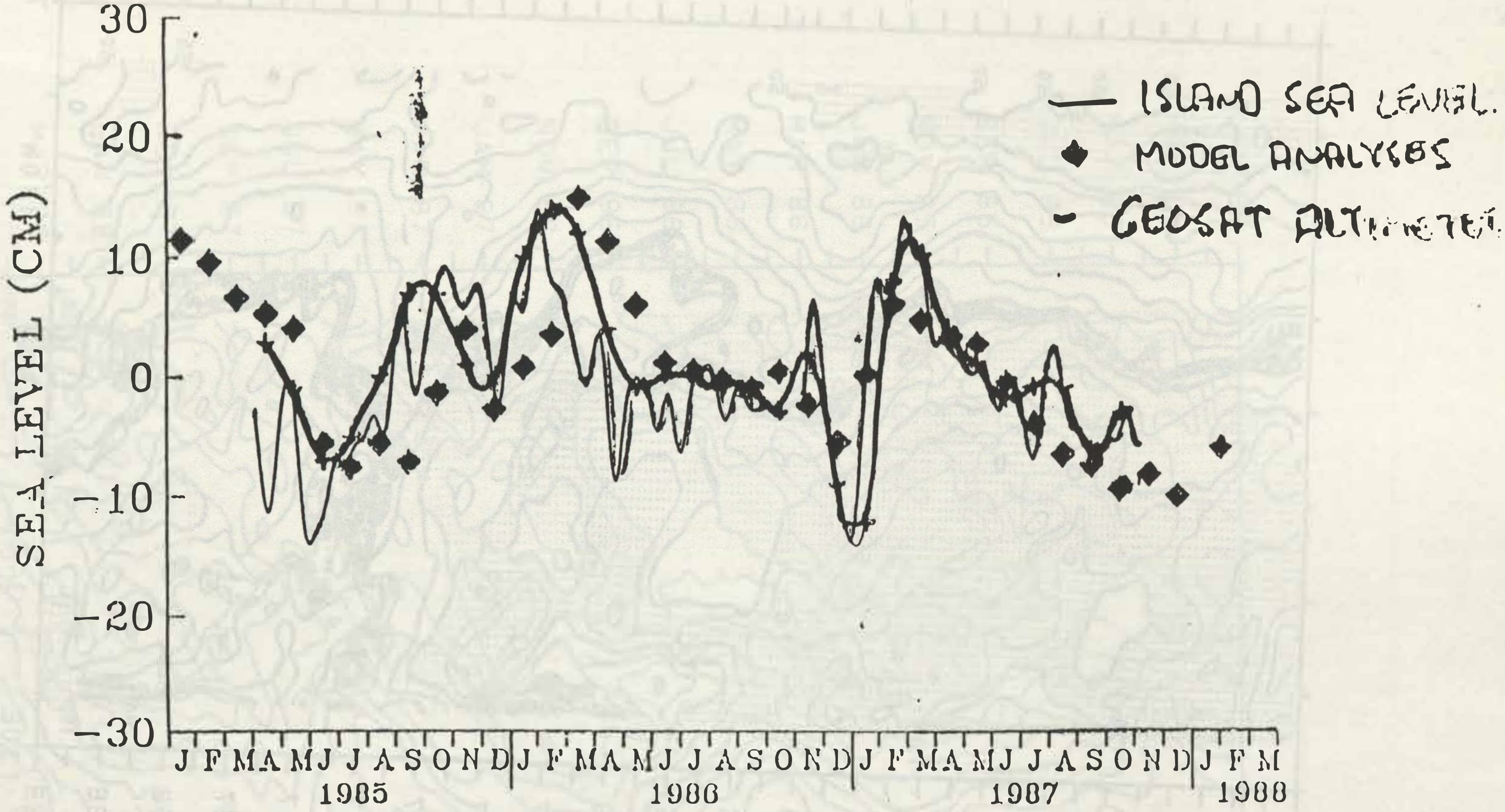
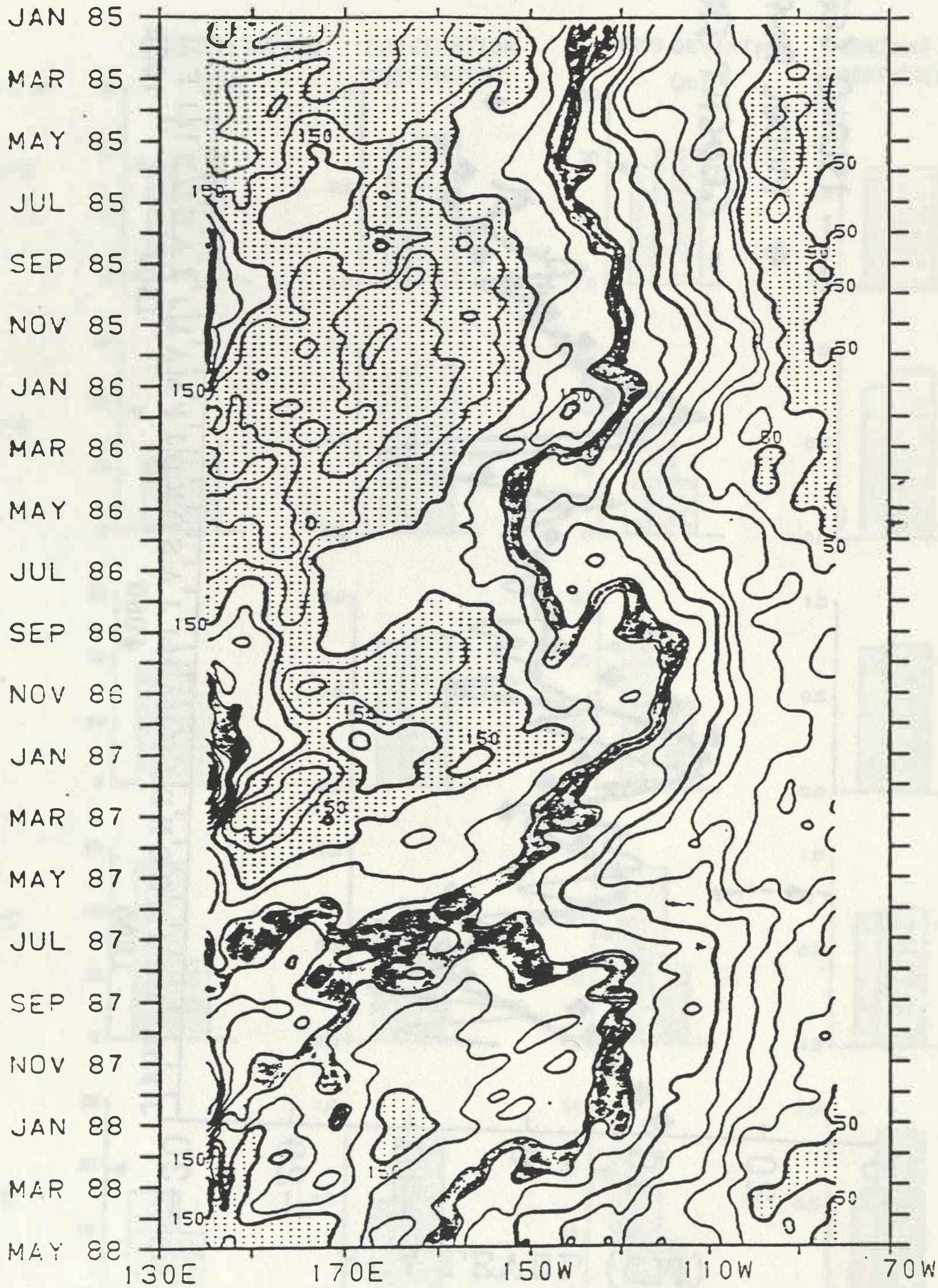


Figure 9. Comparisons between model and mooring data on the equator at 1 degree W for sea surface temperature, the depth of the 15 degree C isotherm, and the zonal velocity at 80 meters. The heavy curve is the mooring data. The dashed curve is the model without assimilation; the light continuous curve is the model with assimilation. The mooring data is not used in the assimilation.

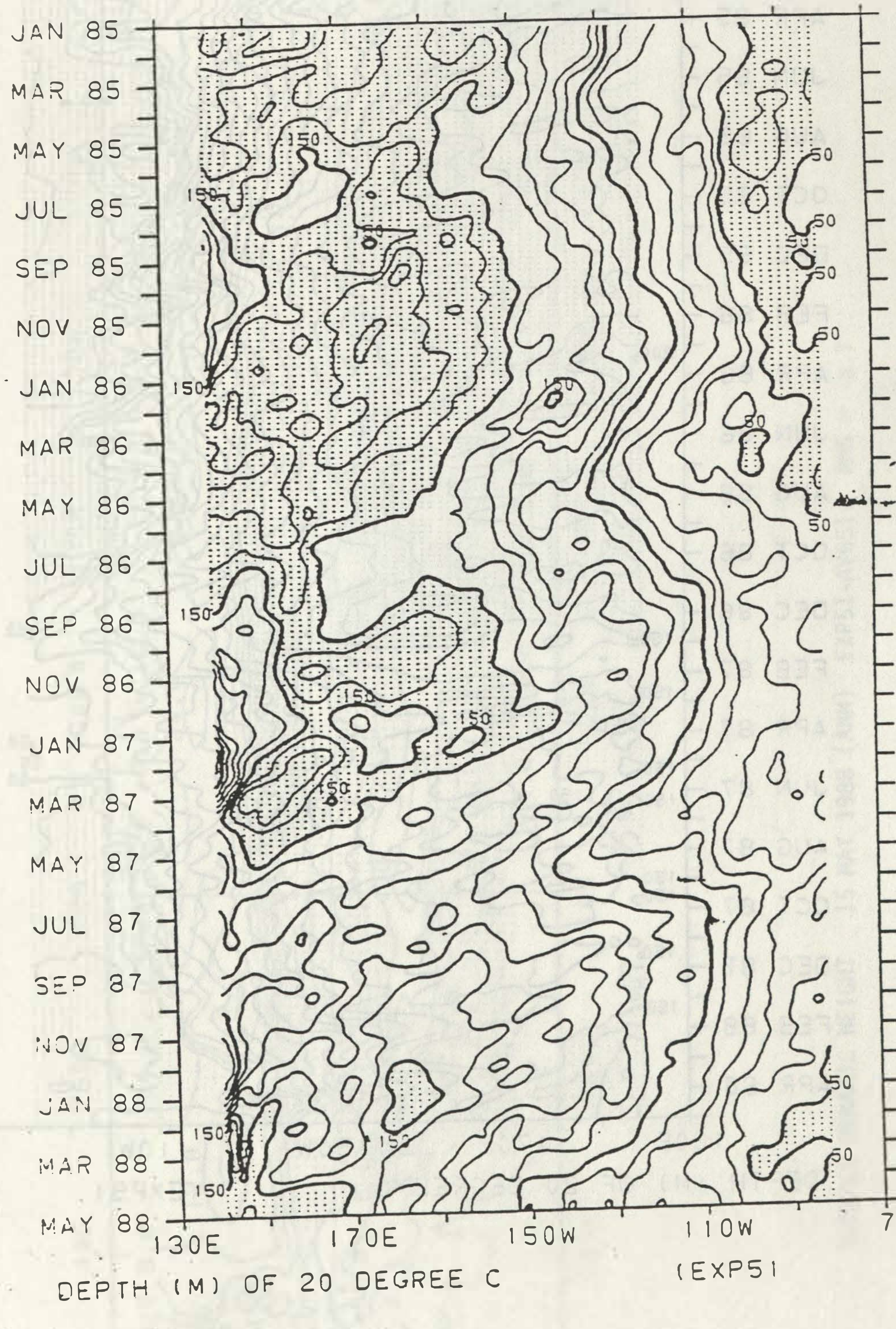

DEPTH OF 20°C
DEPTH OF 15°C
RMS DEVIATION (m)
CORRELATION COEFFICIENT
RMS DEVIATION (m)
CORRELATION COEFFICIENT
110°W
5°N
2°N
0°
2°S
5°S
0
10
20
30
0.0
0.5
1.0



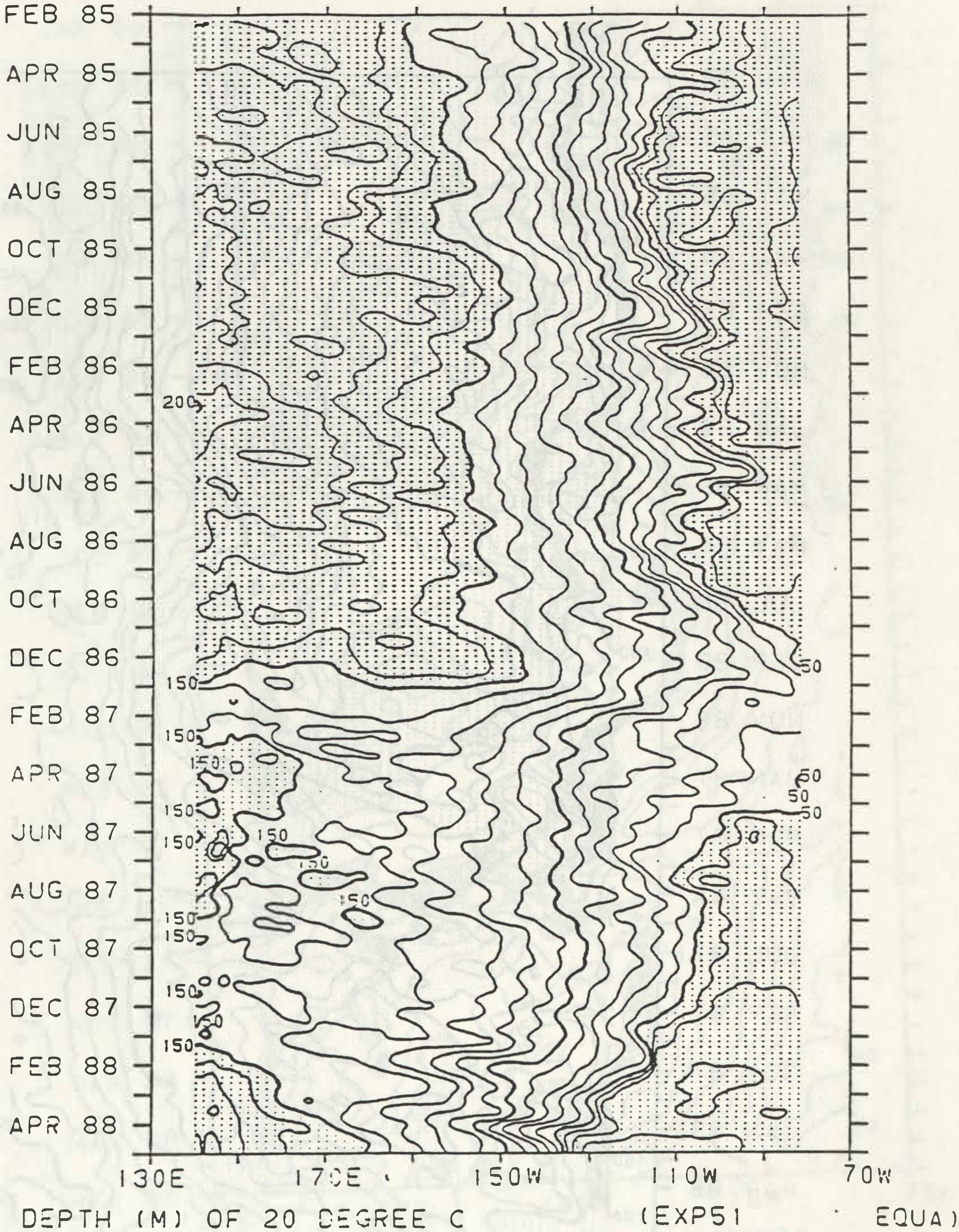
ISLAND SEA LEVEL
MODEL ANALYSES
GEOSAT ALTIMETER
SEA LEVEL (CM)
30
20
10
0
-10
-20
-30
J F M A M J J A S O N D J F M A M J J A S O N D J F M A M J J A S O N D J F M
1985
1986
1987
1988

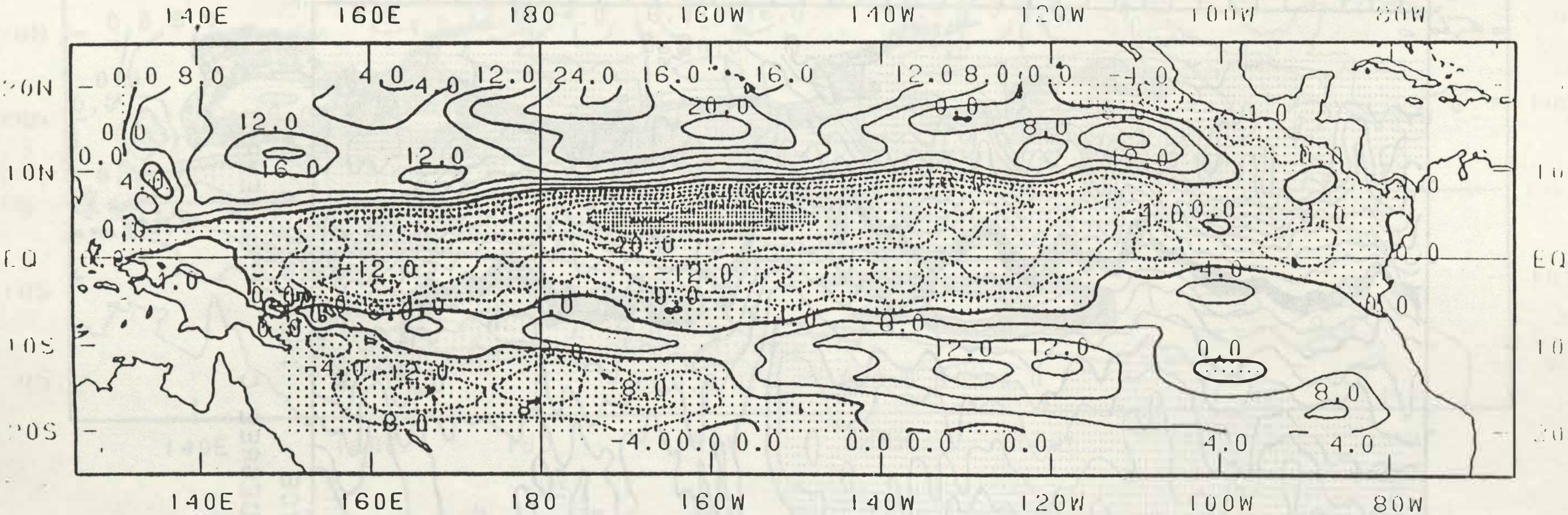


DEPTH (M) OF 20 DEGREE C (EXP51 5N)

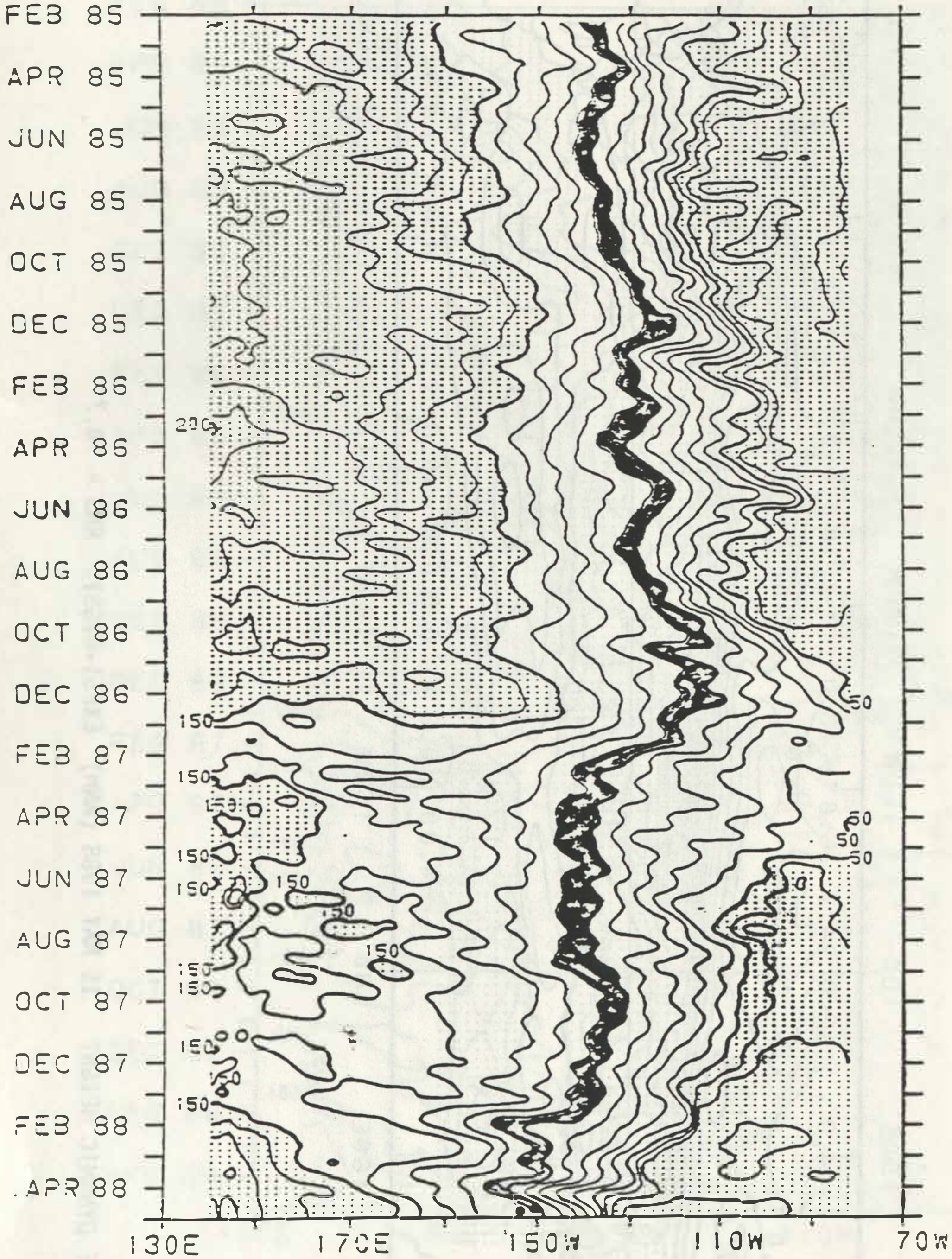


DEPTH (M) OF 20 DEGREE C (EXP51 5N)

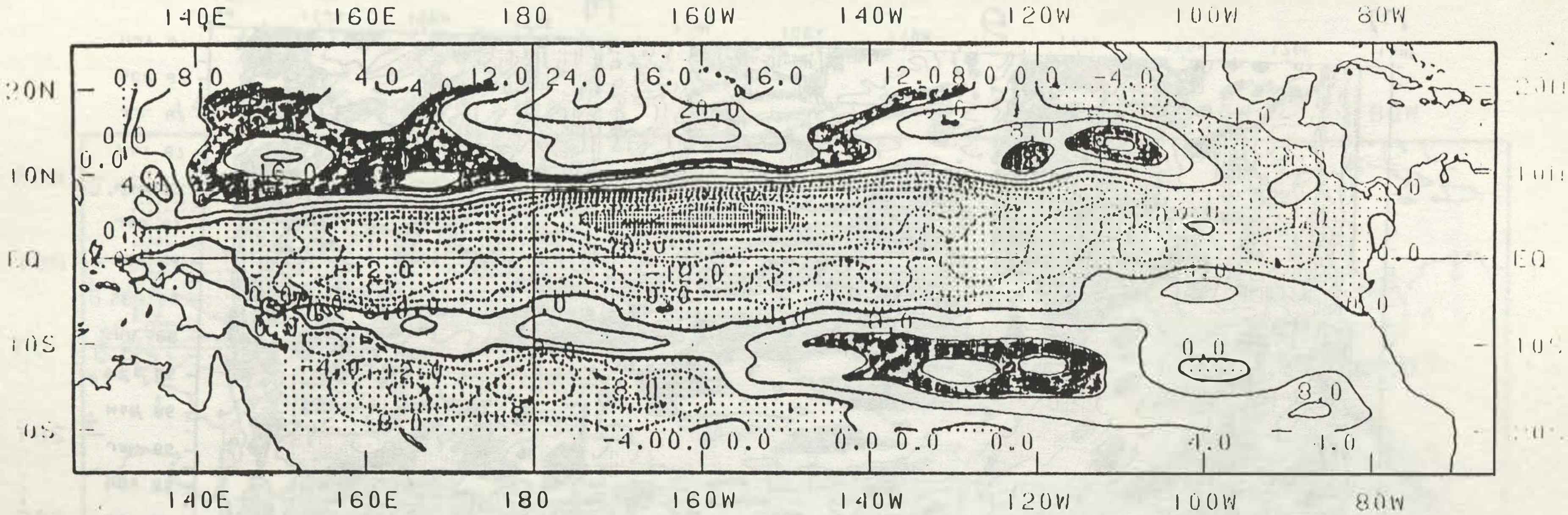

FEB 85
APR 85
JUN 85
AUG 85
OCT 85
DEC 85
FEB 86
APR 86
JUN 86
AUG 86
OCT 86
DEC 86
FEB 87
APR 87
JUN 87
AUG 87
OCT 87
DEC 87
FEB 88
APR 88
200
150
50
130E
170E
150W
110W
70W
DEPTH (M) OF 20 DEGREE C
(EXP51 EQUA)



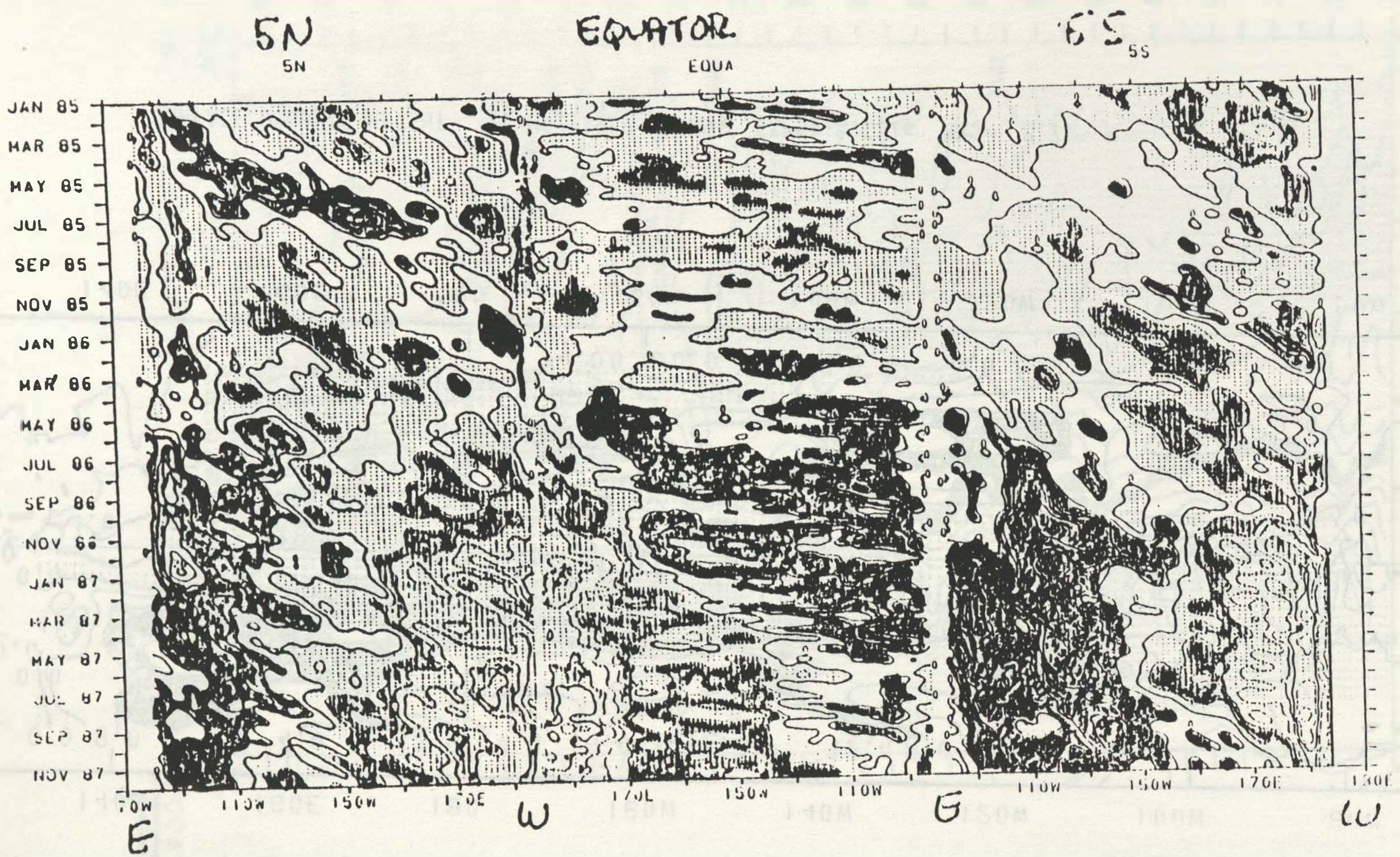
SURFACE DYNAMIC HEIGHT .15 MAY 1988 (ANM) EXP51-AVG51; RMS = 8.7



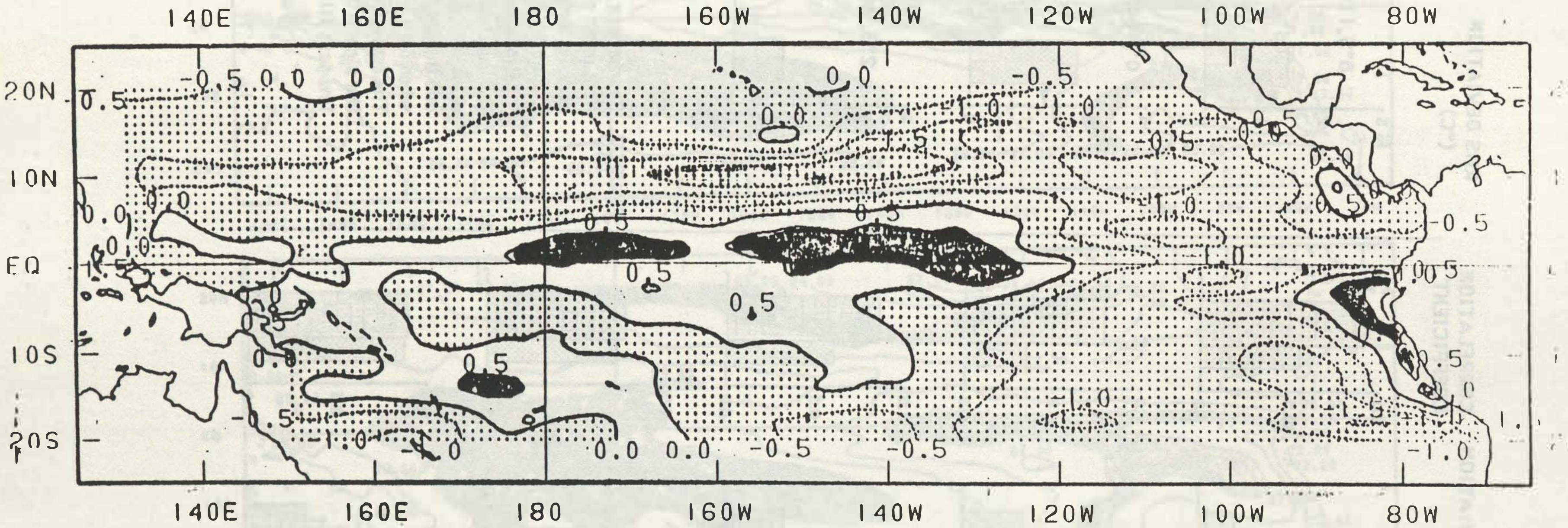
DEPTH (M) OF 20 DEGREE C (EXP51) EQUA)



SURFACE DYNAMIC HEIGHT 15 MAY 1988 (ANM) EXP51-AVG51; RMS= 8.7



INTERNATIONAL SURFACE DYNAMIC HGT. VARIABILITY (C.I. = 4cm)



QUARTERLY AVERAGED TEMP. ANOMALY (5M. DEC-FEB) EXP51 - EXP79

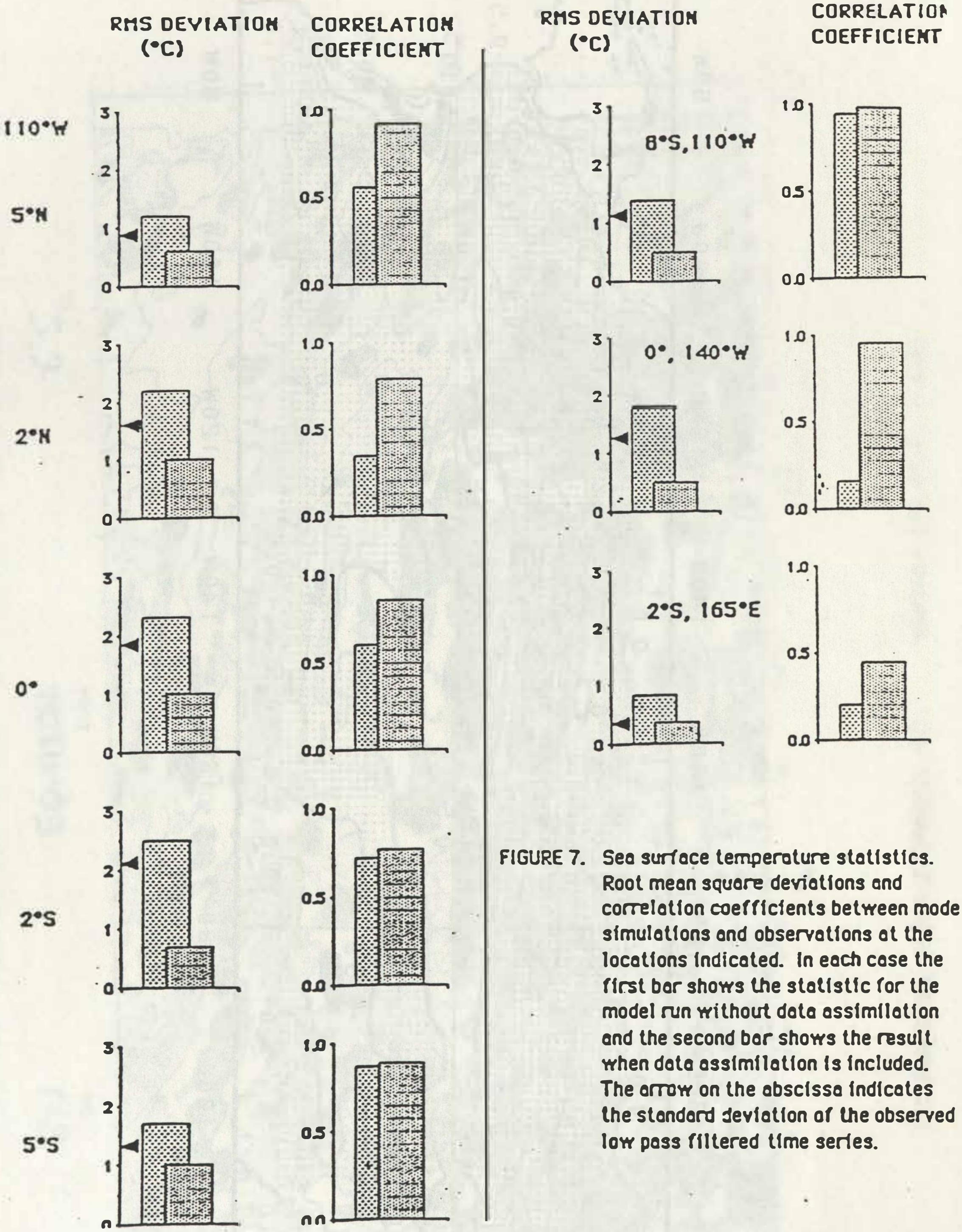


FIGURE 7. Sea surface temperature statistics. Root mean square deviations and correlation coefficients between mode simulations and observations at the locations indicated. In each case the first bar shows the statistic for the model run without data assimilation and the second bar shows the result when data assimilation is included. The arrow on the abscissa indicates the standard deviation of the observed low pass filtered time series.

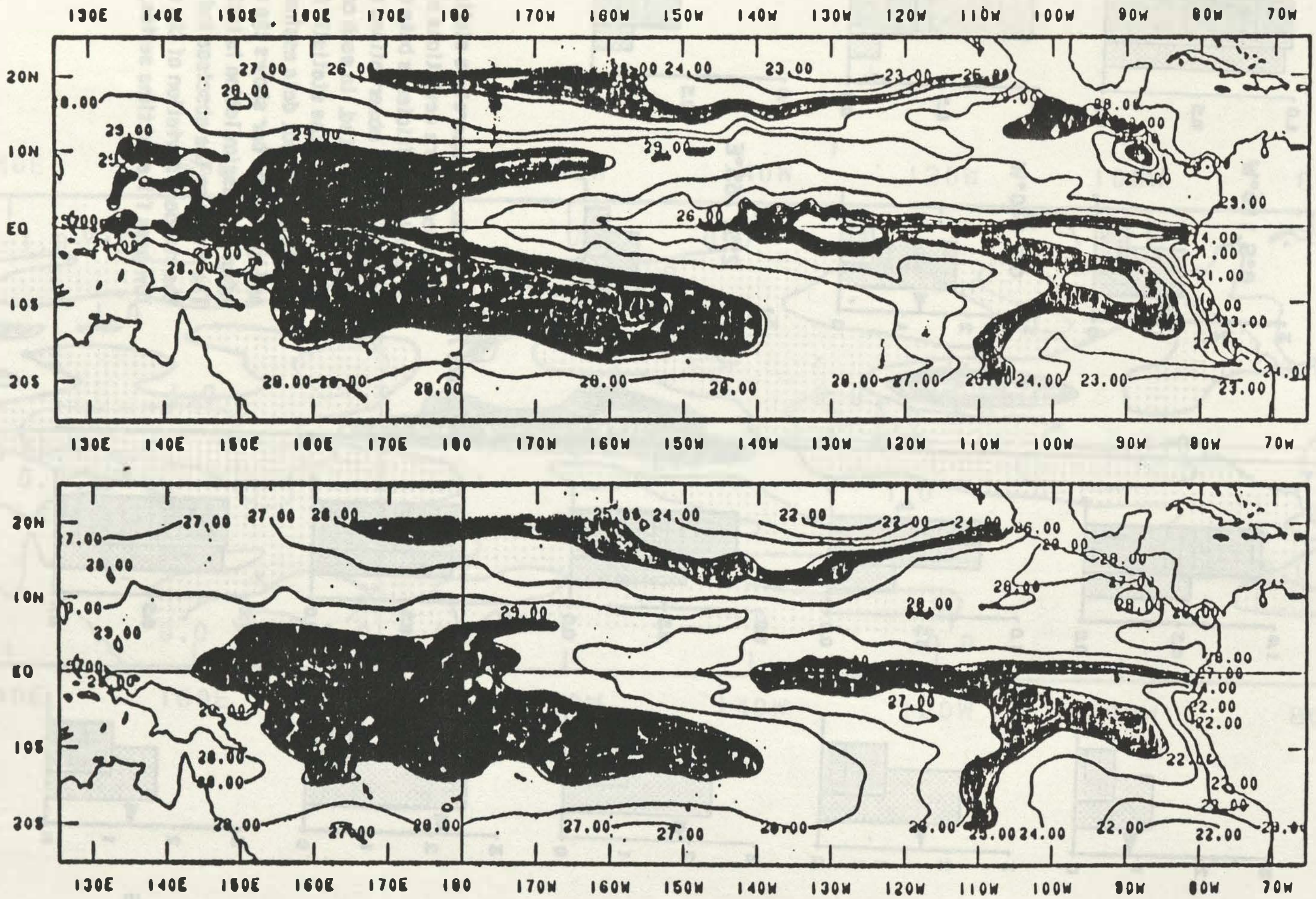
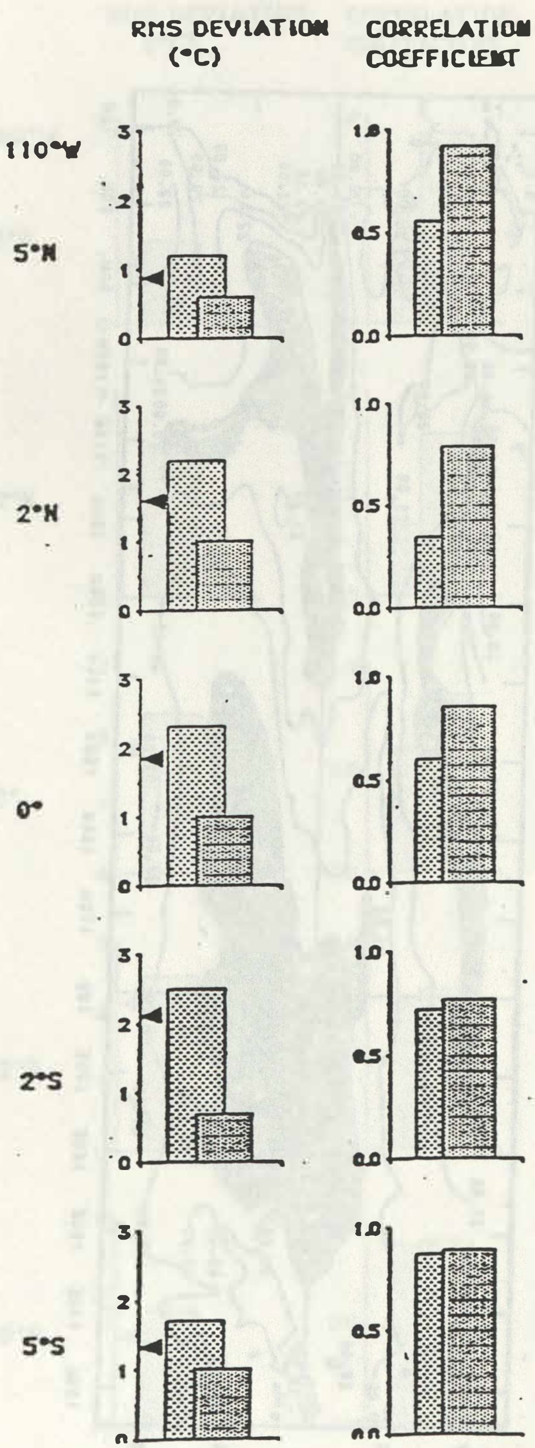


Figure 1. Upper panel: Model estimate for sea surface temperature for January 1988 after a 30 day integration with no corrections with in situ data. Lower panel: The field in the upper panel modified through the use of BSST and XBT information.



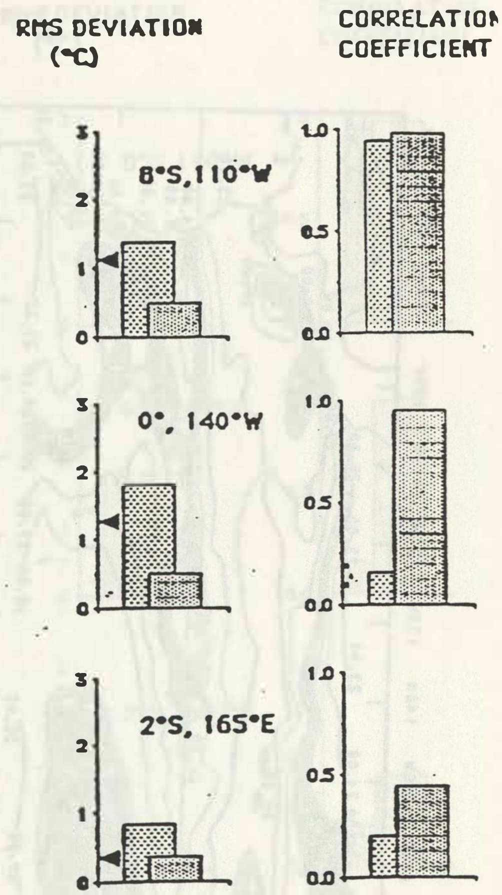
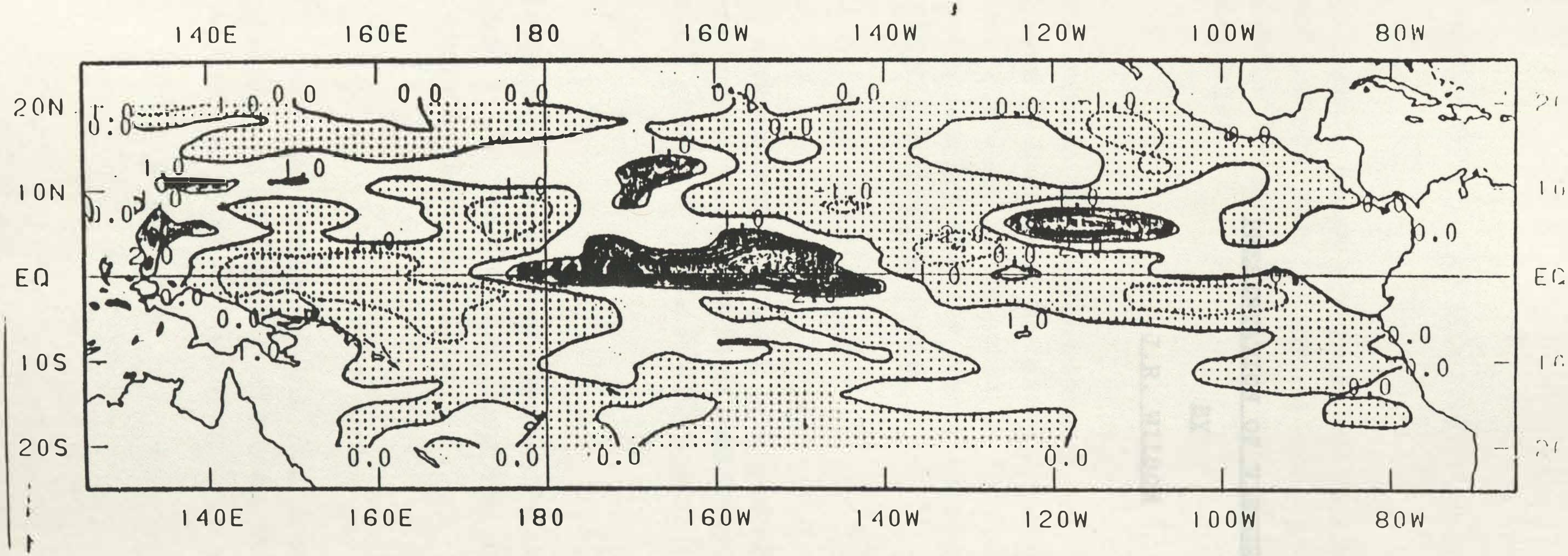


FIGURE 7. Sea surface temperature statistics. Root mean square deviations and correlation coefficients between mode simulations and observations at the locations indicated. In each case the first bar shows the statistic for the model run without data assimilation and the second bar shows the result when data assimilation is included. The arrow on the abscissa indicates the standard deviation of the observed low pass filtered time series.



(CORRECTION)

QUARTERLY AVERAGED TEMP. ANOMALY (95M. DEC-FEB) EXP51 - EXP79

PRESENTATION OF VIEWGRAPHS

BY

J.R. WILSON

GOALS OF US-CANADA THERMAL-SALINITY PROJECT

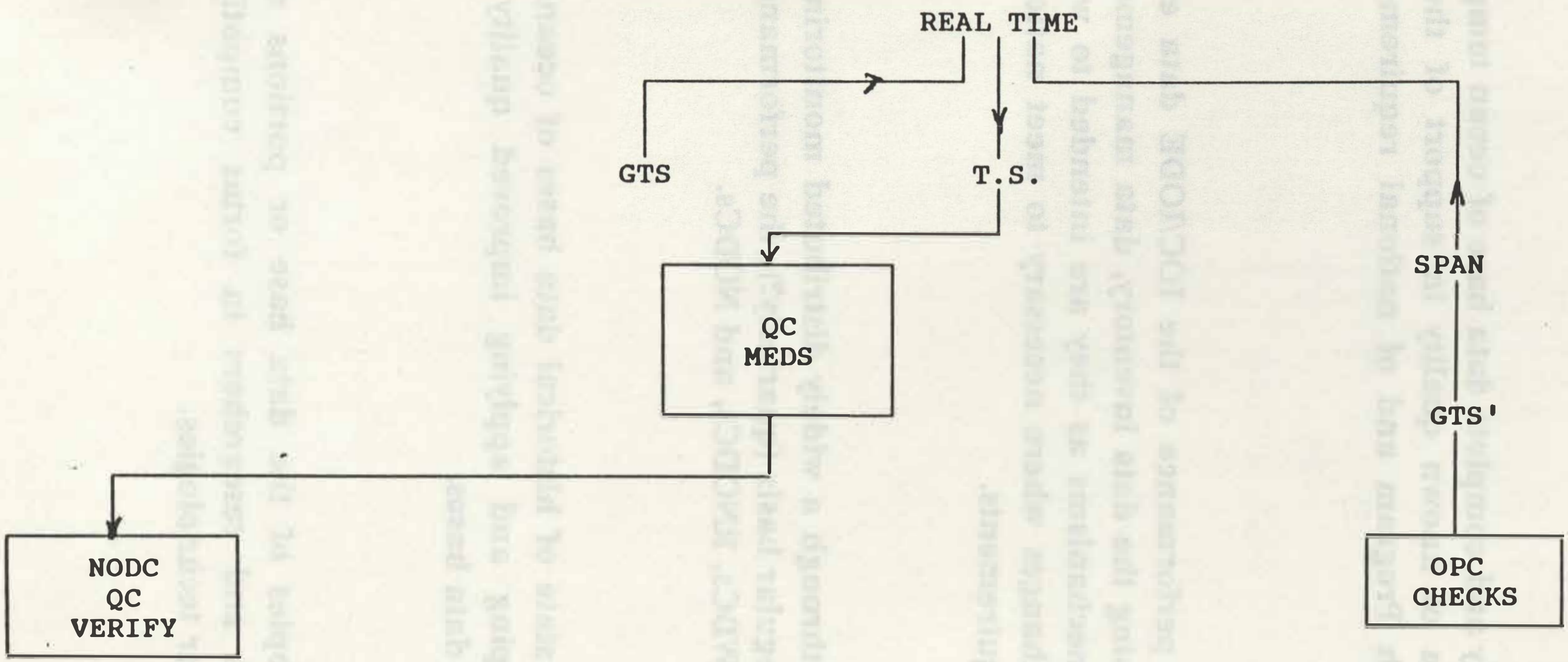
1. To create a timely and complete data base of ocean temperature and salinity data of known quality in support of the World Climate Research Program and of national requirements for these data.

2. To improve the performance of the IOC/IODE data exchange system by exercising the data inventory, data management, and data exchange mechanisms as they are intended to work and recommending changes where necessary to meet national and international requirements.

3. To demonstrate through a widely distributed monitoring report produced on a regular basis (quarterly?), the performance of the IODE system of WDCs, RNODCs, and NODCs.

4. To improve the state of historical data bases of oceanographic data by developing and applying improved quality control systems to these data bases.

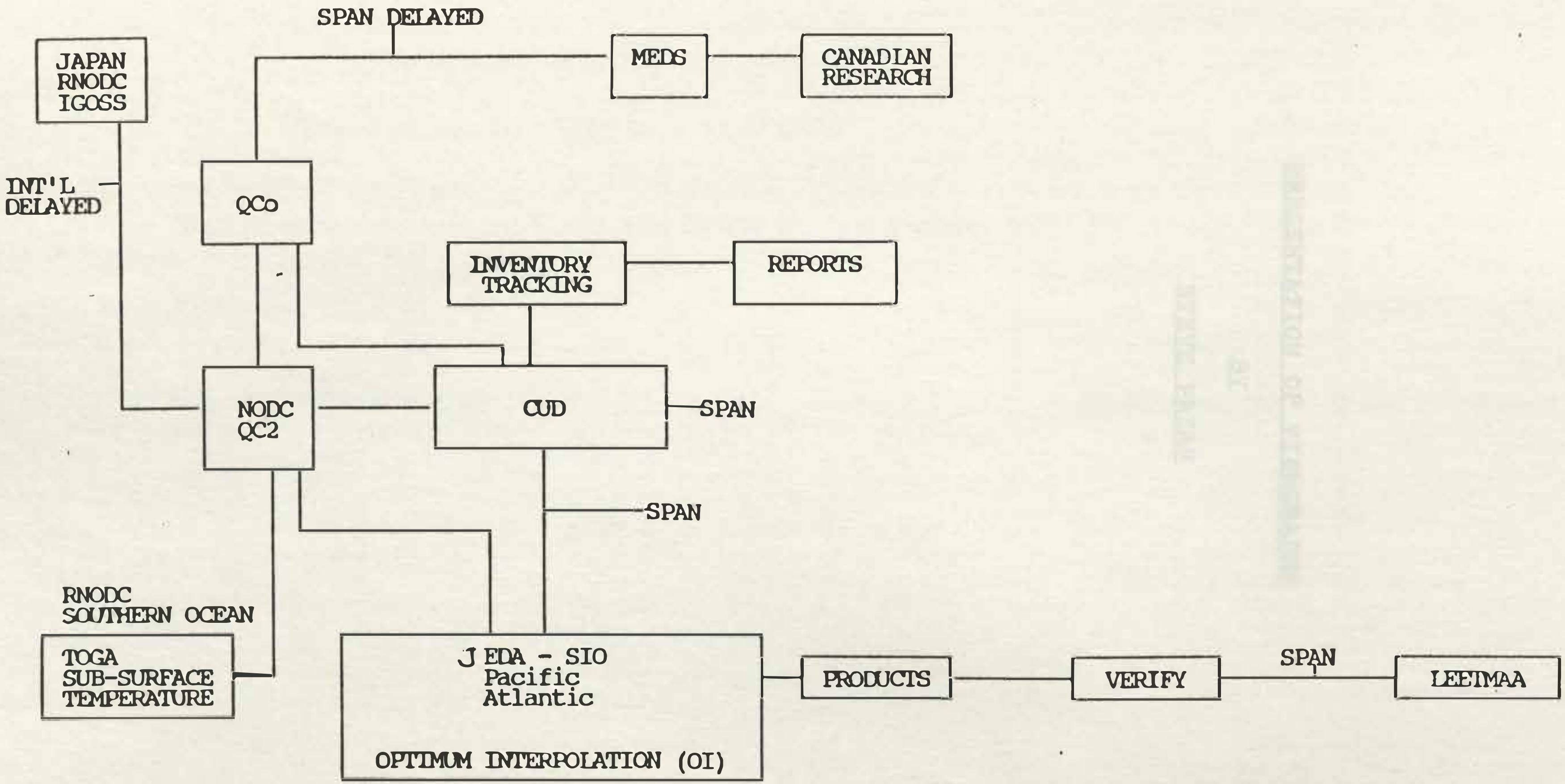
5. To distribute copies of the data base or portions of it to interested users and researchers in forms compatible with modern computer technologies.



MEDS GETS GTS & GTS' FROM NODC
WE GET GTS & SEND TO MEDS
SENT GTS' TO MEDS VIA OPC

MEDS QC INCLUDES:
HEADER
PROFILE
SHIP TRACK
TIME DISTANCE
ETC.

INTERNATIONAL DATA CENTERS



OCo - On line data base

PRESENTATION OF VIEWGRAPHS

BY

STEVE PAZAN

PARADIGM FOR FUTURE JEDA CENTER ACTIVITIES

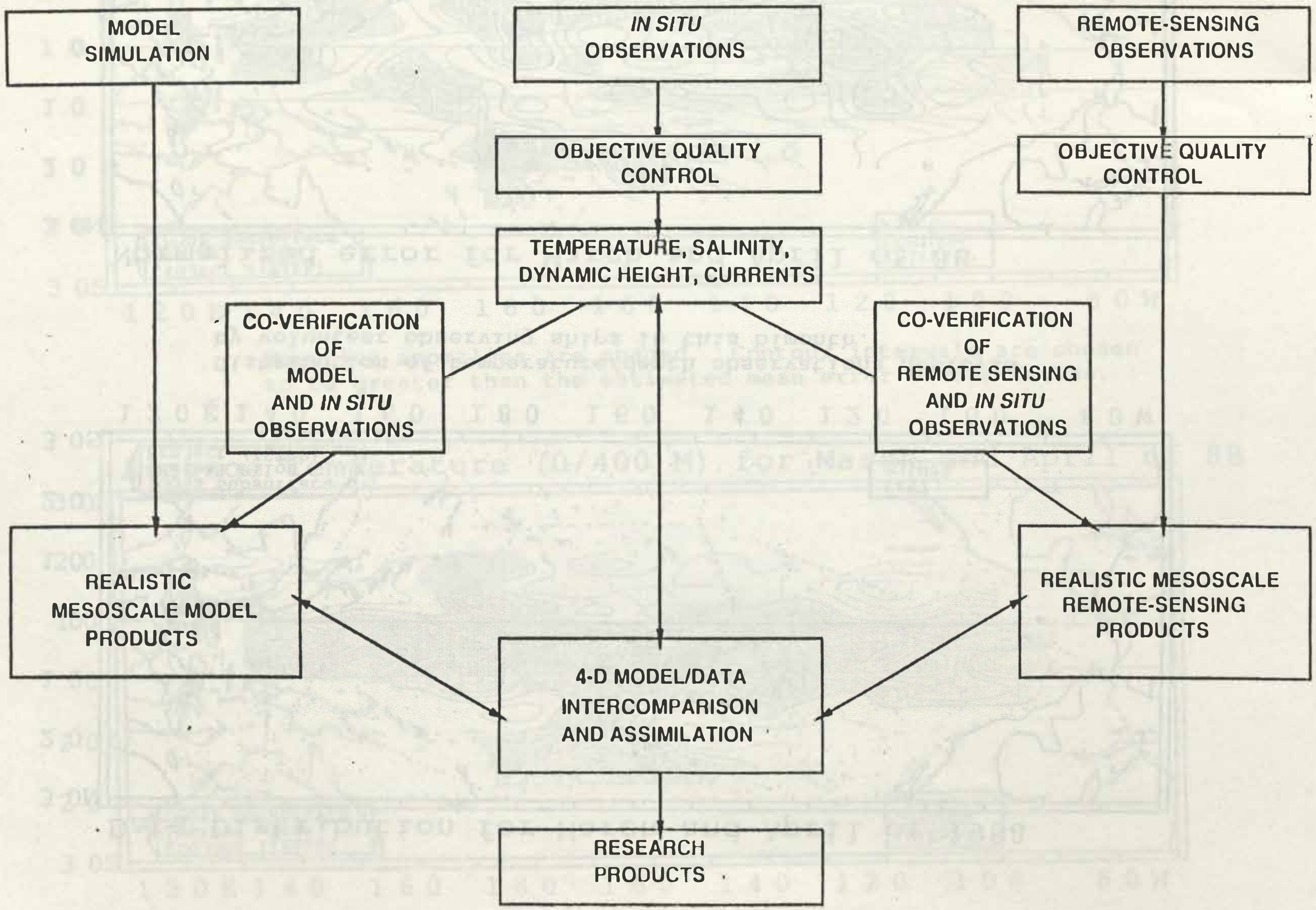
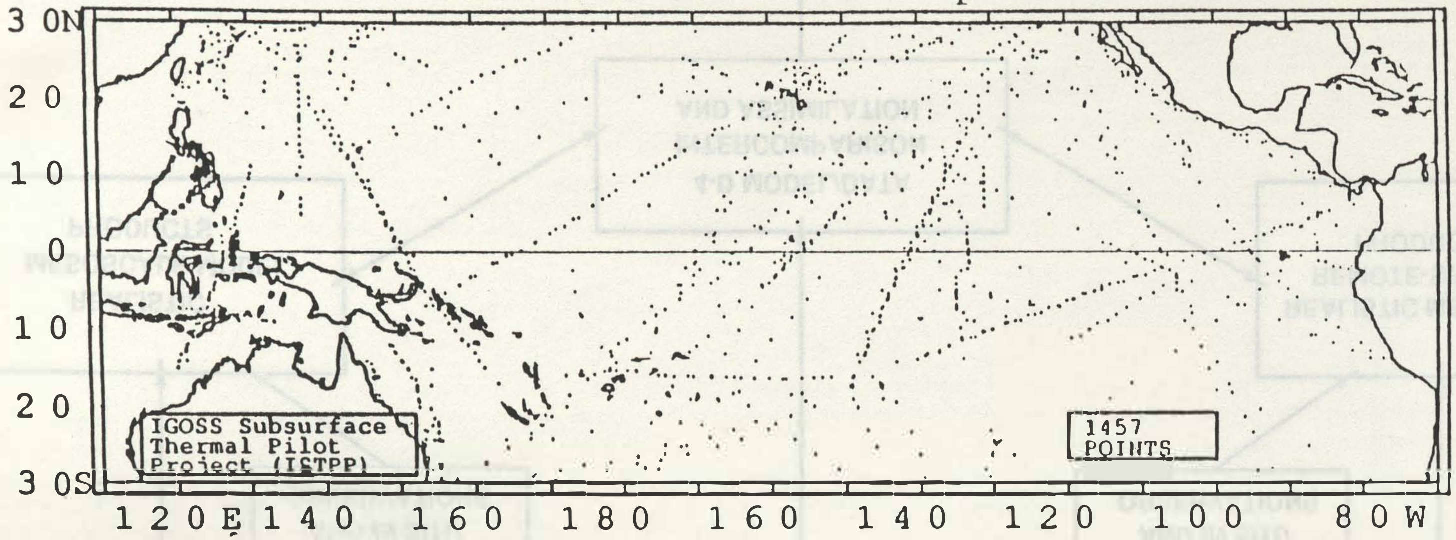


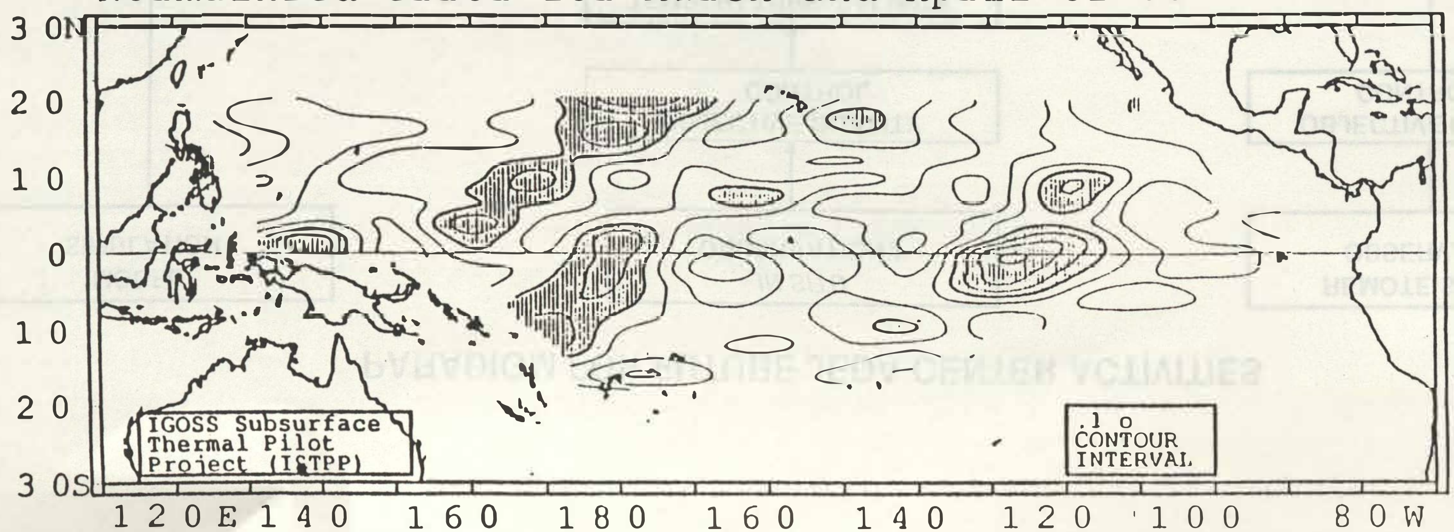
Figure 1

Data Distribution for March and April of 1988



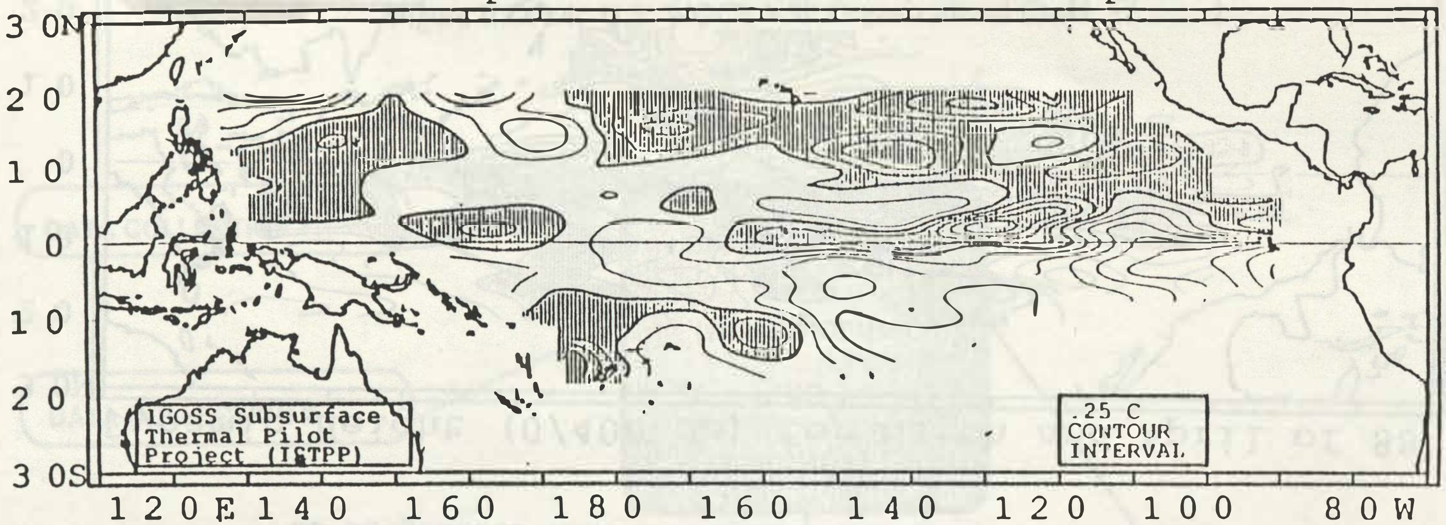
Distribution of temperature/depth observations deployed by volunteer observing ships in this bimonth.

Normalized error for March and April of 88



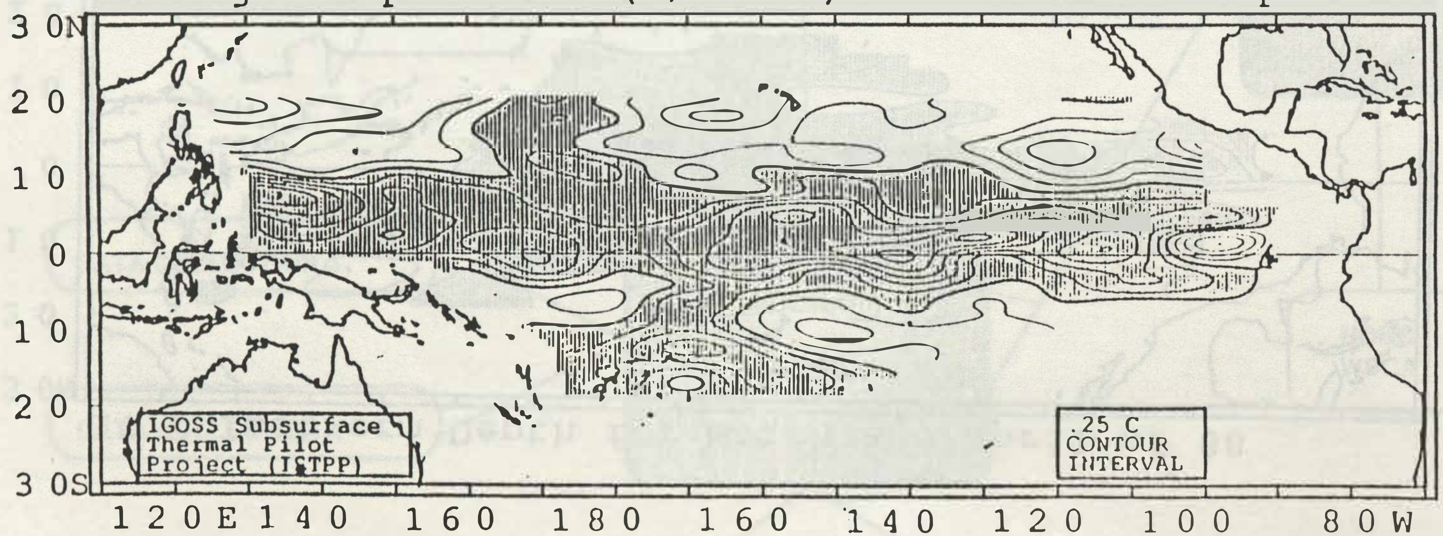
Errors greater than .7 standard deviations are shaded to indicate regions of large error.

Sea Surface Temperature for March and April of 88



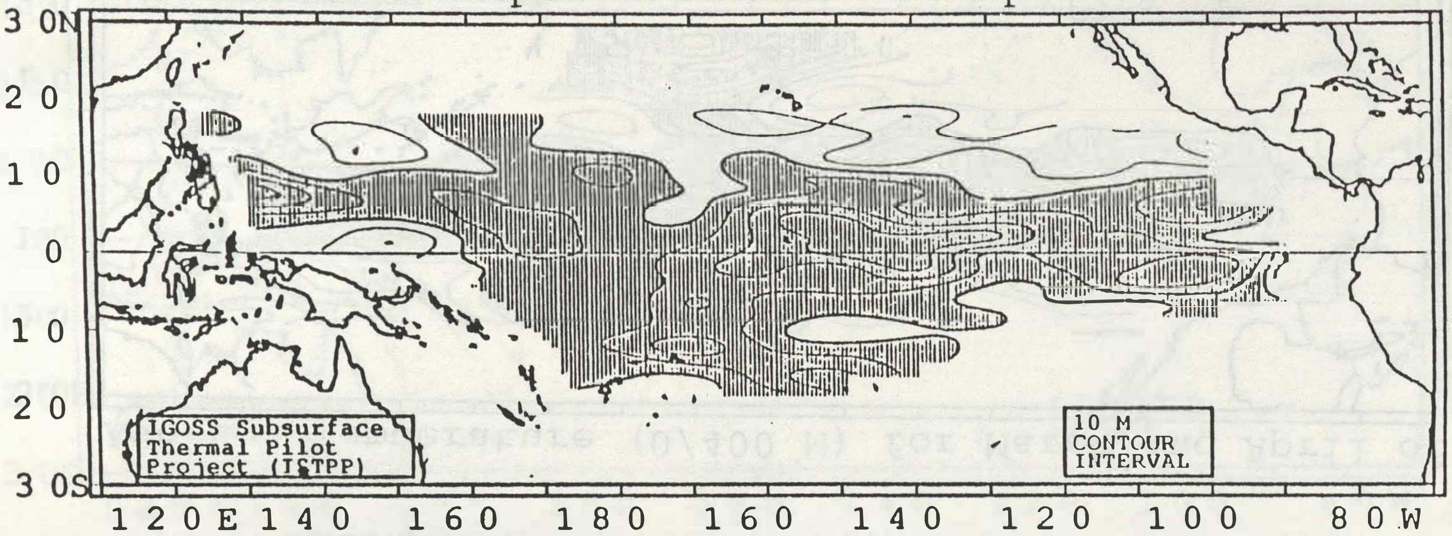
Negative anomalies are shaded. Contour intervals are chosen to be greater than the estimated mean error of the fields.

Average Temperature (0/400 M) for March and April of 88



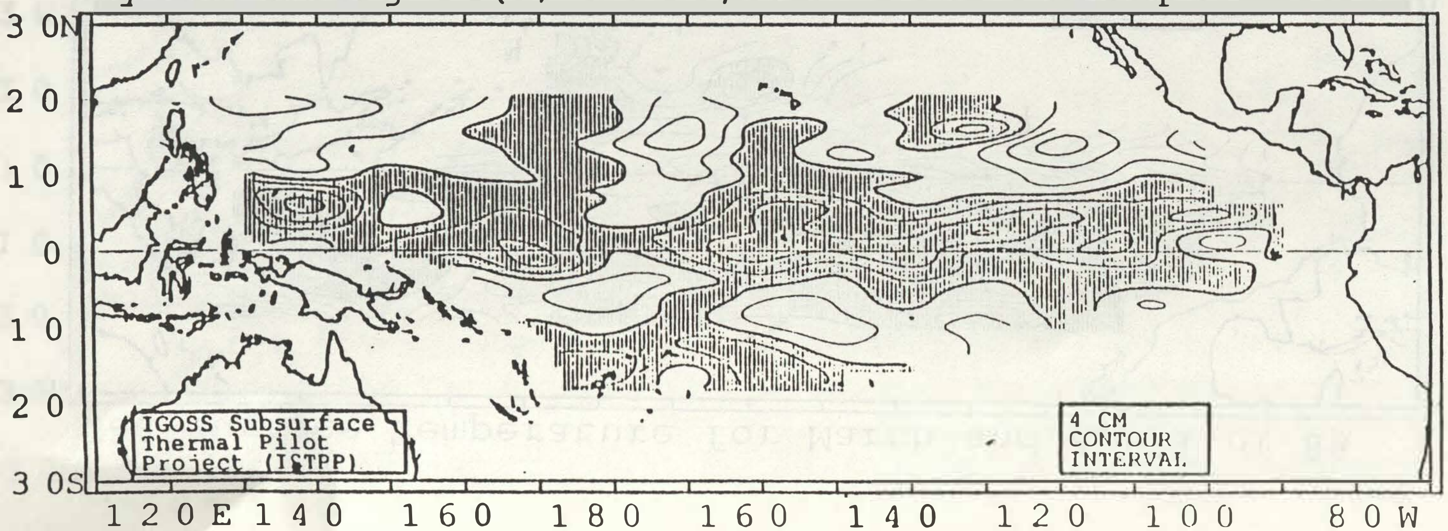
Negative anomalies are shaded. Contour intervals are chosen to be greater than the estimated mean error of the fields.

14 C Isotherm Depth for March and April of 88



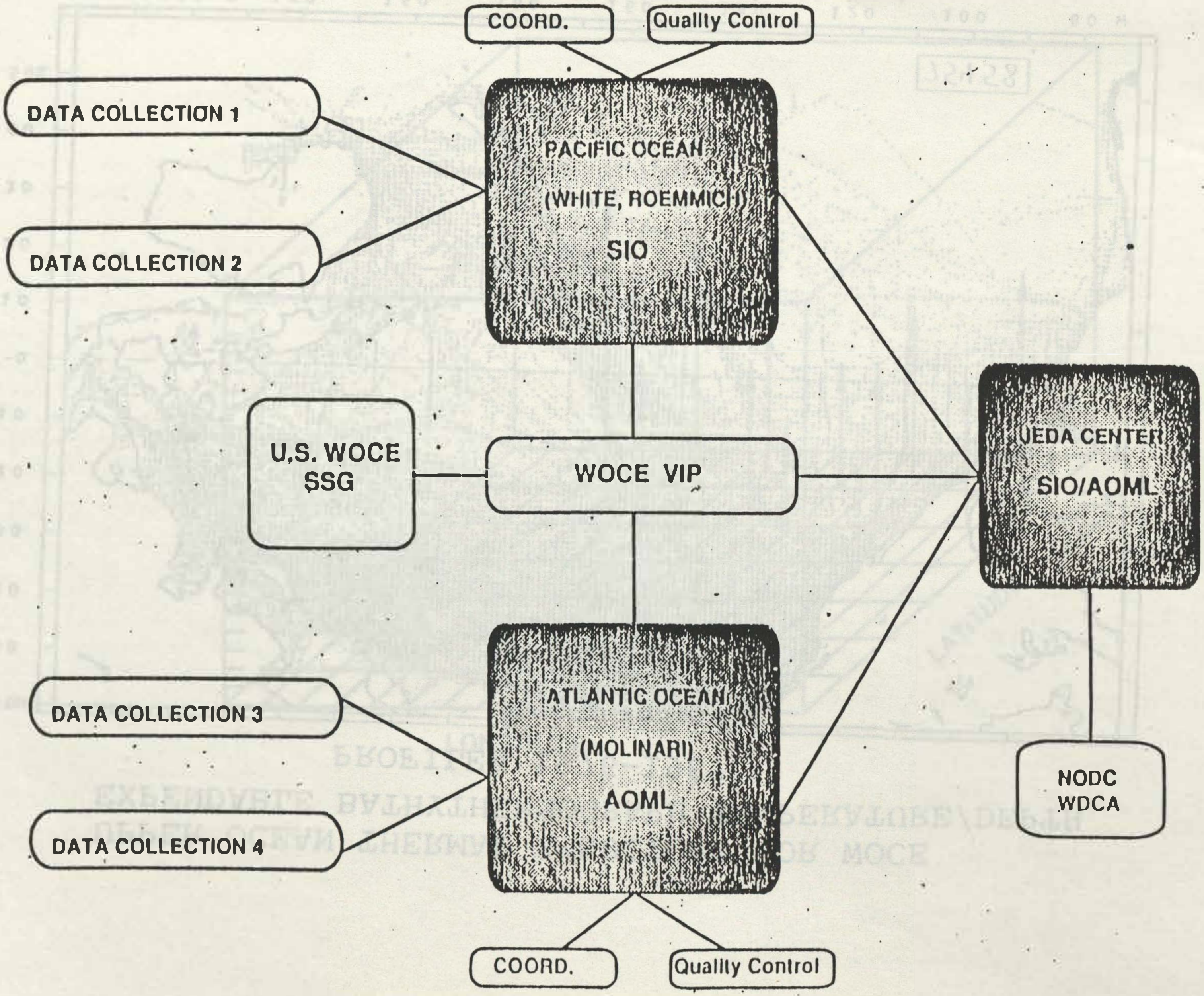
Negative anomalies are shaded. Contour intervals are chosen to be greater than the estimated mean error of the fields.

Dynamic Height (0/400 db) for March and April of 88



Negative anomalies are shaded. Contour intervals are chosen to be greater than the estimated mean error of the fields.

Figure II.D.4.4. U.S. WOCE VSP Organizational Chart



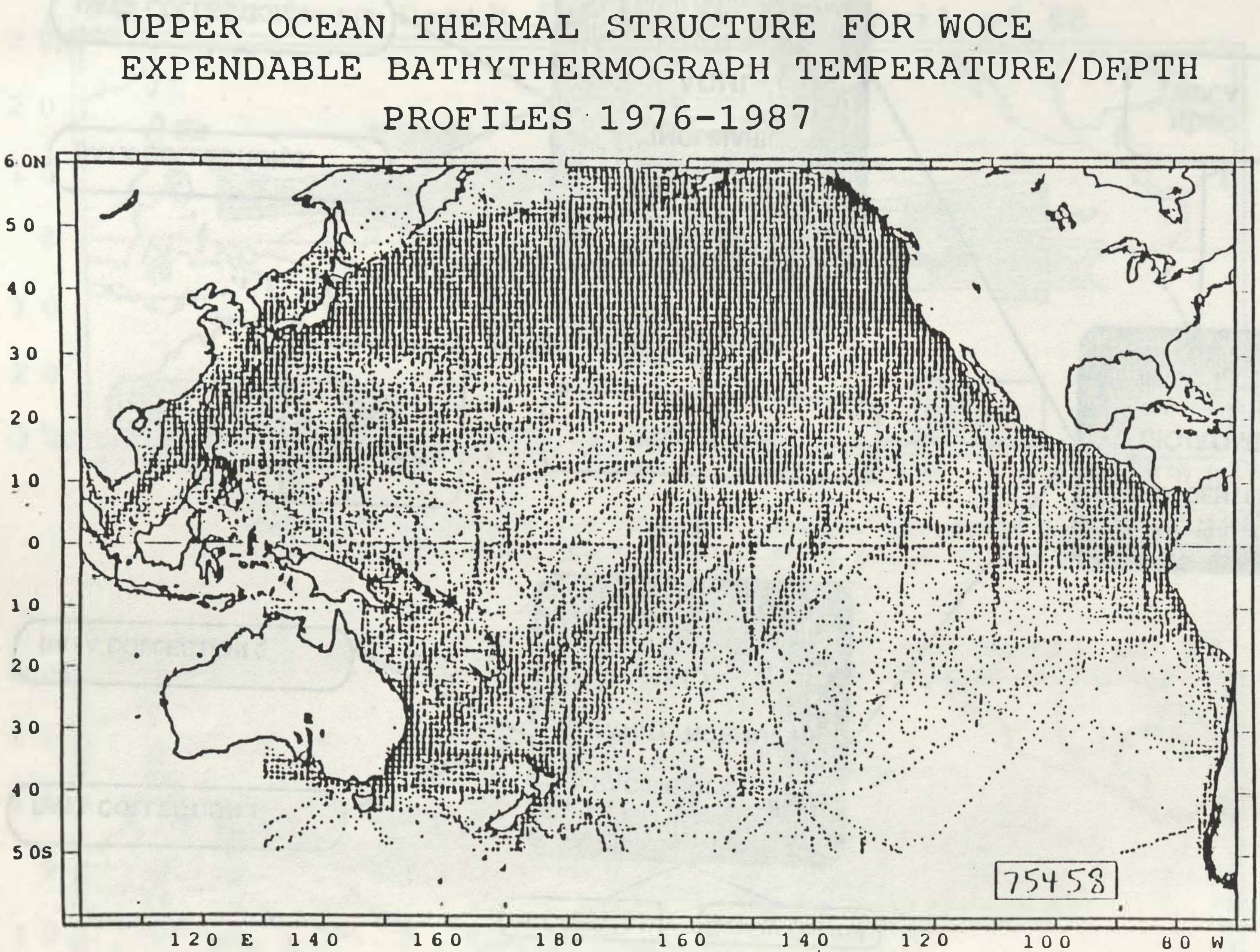
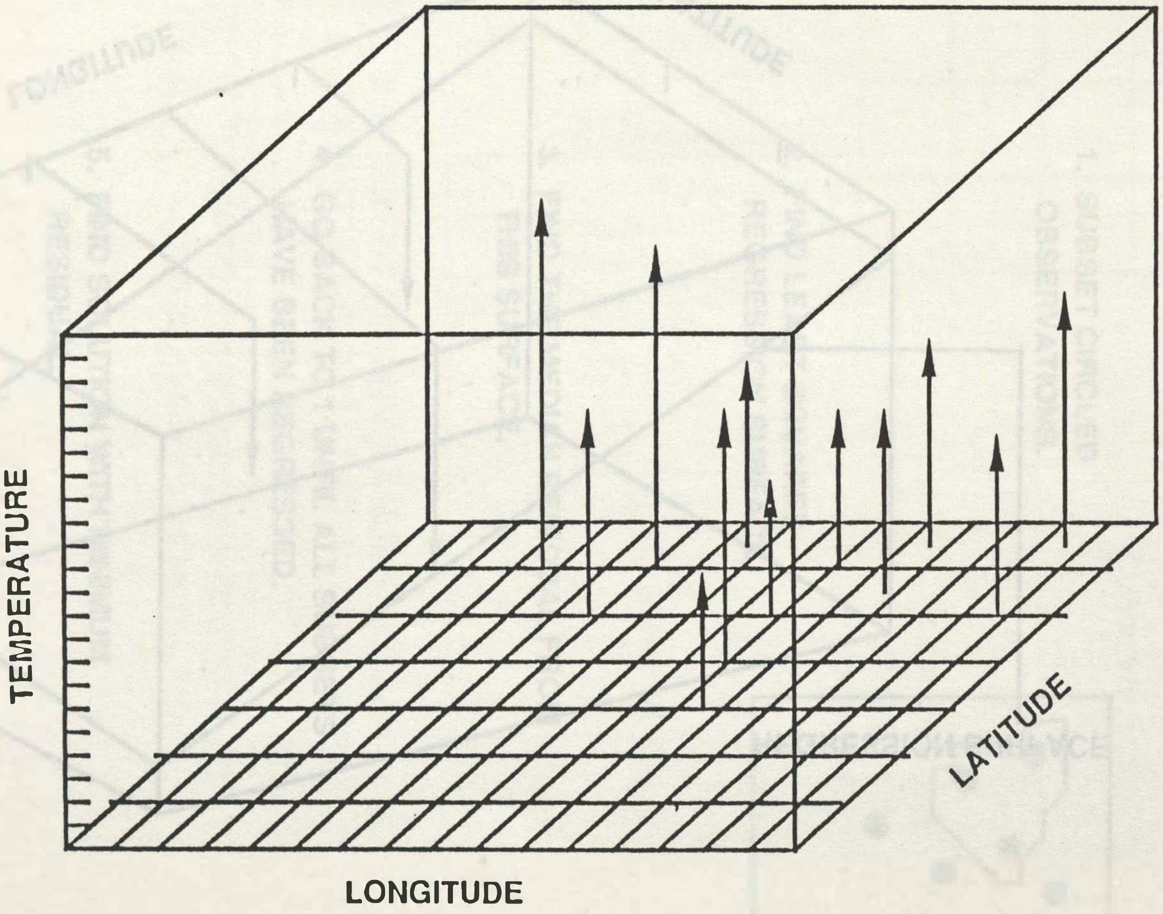
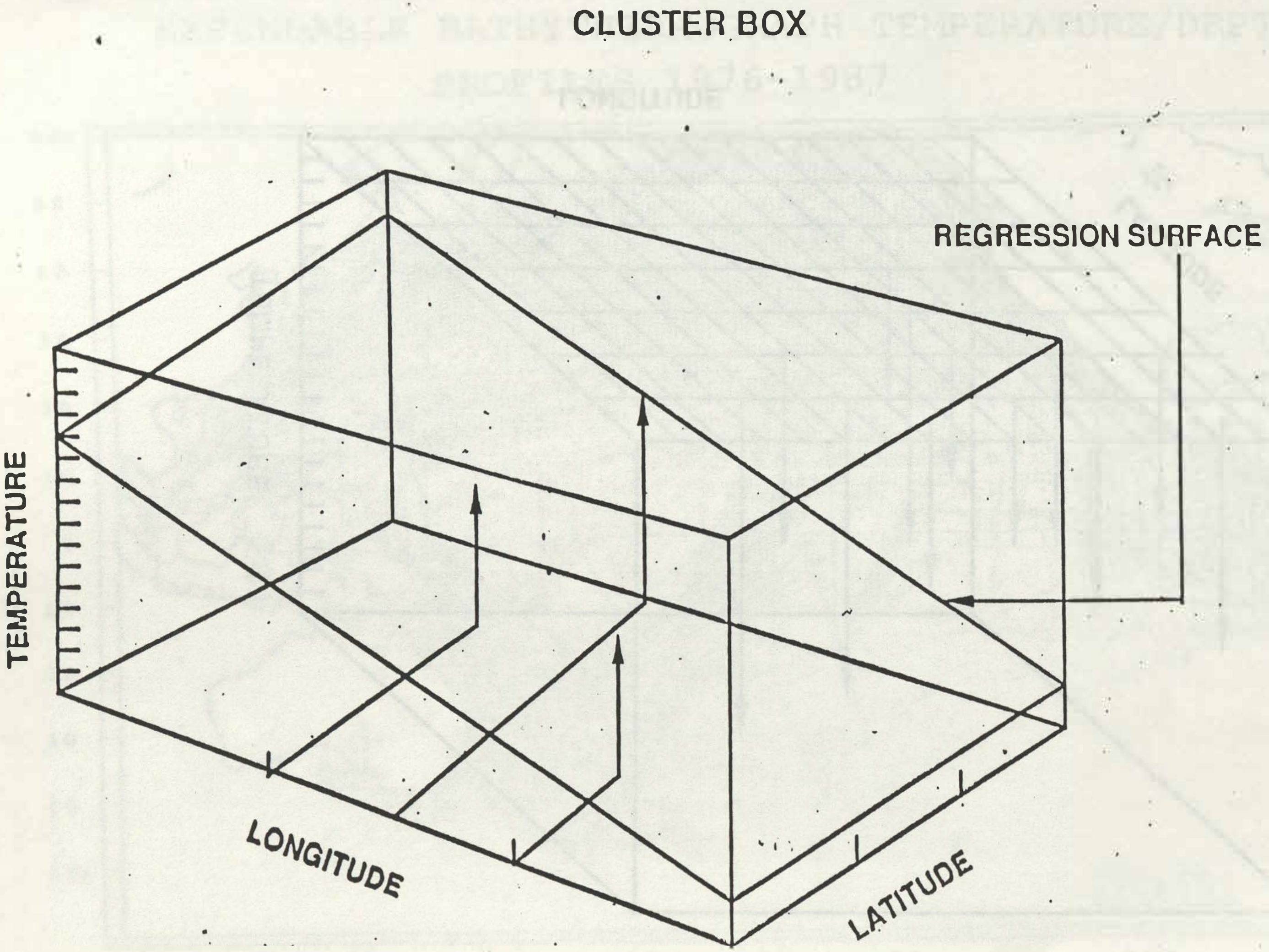


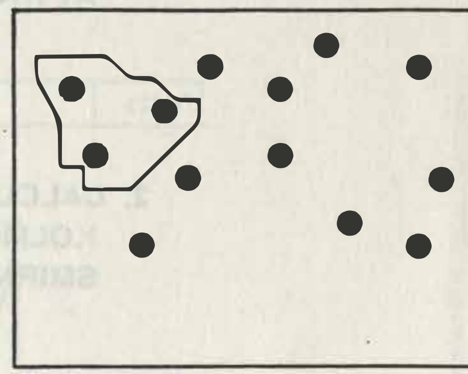
Figure 1.1 Distribution of temperature/depth observations deployed by volunteer observing ships

CLUSTER BOX




CLUSTER BOX
REGRESSION SURFACE
TEMPERATURE
LONGITUDE
LATITUDE

1. SUBSET CIRCLED OBSERVATIONS.



2. FIND LEAST SQUARES REGRESSION SURFACE.

3. FIND THE MEDIAN RESIDUAL FROM THIS SURFACE.

4. GO BACK TO 1 UNTIL ALL SUBSETS HAVE BEEN REGRESSED.

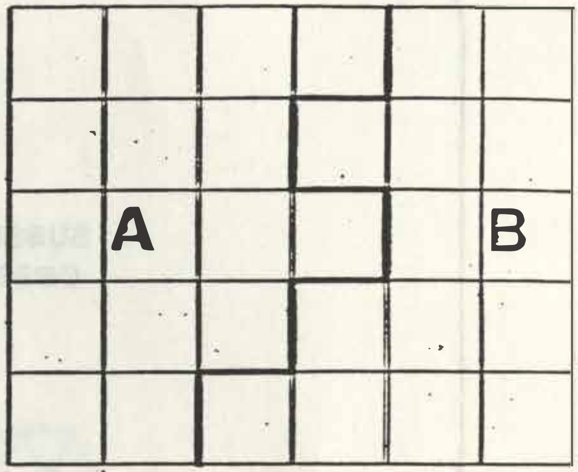
5. FIND SOLUTION WITH MINIMUM RESIDUAL.

THIS SOLUTION IS ROBUST AGAINST 50% OF OUTLIERS IN LATITUDE, LONGITUDE, AND TEMPERATURE!

1. DIVIDE CLUSTER BOX

   ROUGHLY IN TWO

   AT RANDOM.

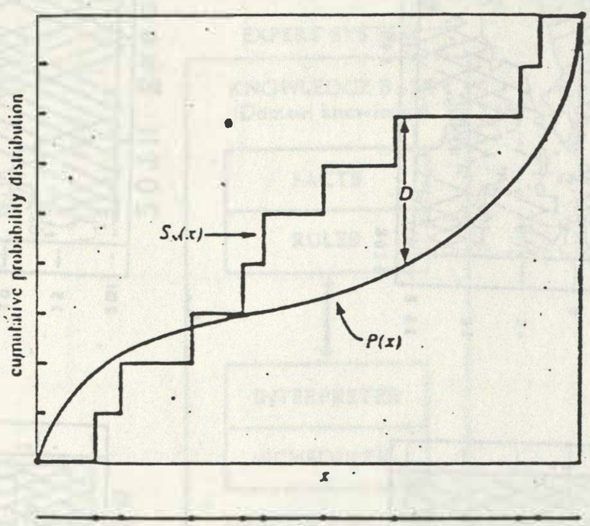


2. CALCULATE KOLMOGOROV-SMIRNOV TEST.

3. GO TO STEP 1.

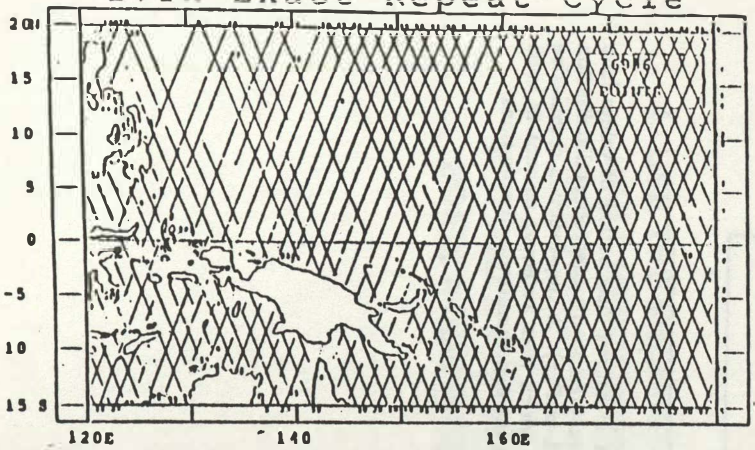
4. USE THE MINIMUM K-S NUMBER TO TEST THE NULL HYPOTHESIS THAT POPULATIONS ARE IDENTICAL.

   IF IT FAILS AT 1% CONFIDENCE, DIVIDE POPULATION INTO TWO CLUSTERS.

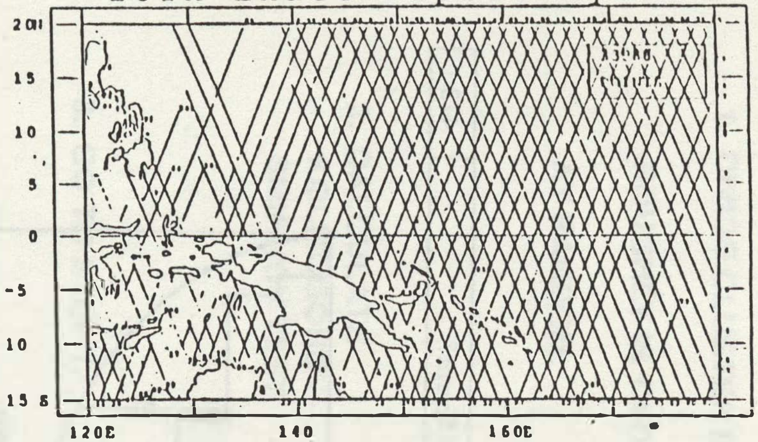


13.5.1. Kolmogorov-Smirnov statistic D. A measured distribution of values in x as N dots on the lower abscissa) is to be compared with a theoretical distribution cumulative probability distribution is plotted as $P(x)$. A step-function cumulative ...ility distribution $S_N(x)$ is constructed, one which rises an equal amount at each ...d point. D is the greatest distance between the two cumulative distributions.

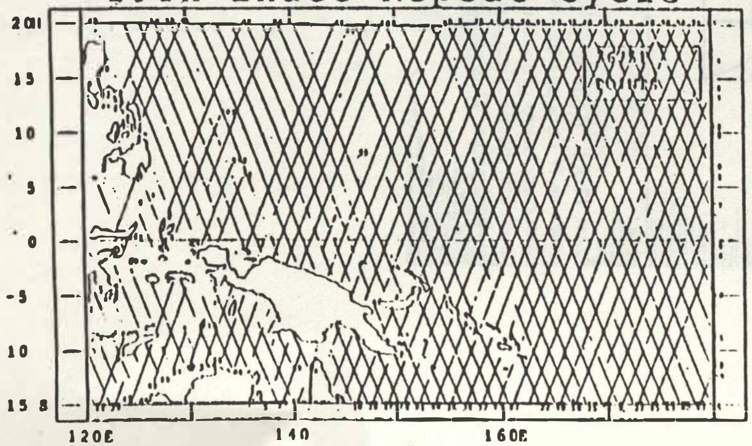
17TH Exact Repeat Cycle



18TH Exact Repeat Cycle



19TH Exact Repeat Cycle



20TH Exact Repeat Cycle

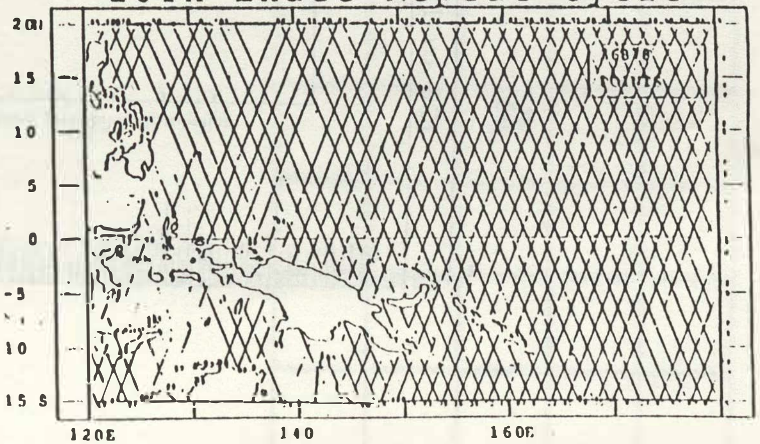


TABLE 2.1 Comparing human and artificial expertise: the good news.

The Good News	
Human Expertise	*Artificial Expertise*
Perishable	Permanent
Difficult to transfer	Easy to transfer
Difficult to document	Easy to document
Unpredictable	Consistent
Expensive	Affordable

TABLE 2.2 Comparing human and artificial expertise: the bad news.

The Bad News	
Human Expertise	*Artificial Expertise*
Creative	Uninspired
Adaptive	Needs to be told
Sensory experience	Symbolic input
Broad focus	Narrow focus
Commonsense knowledge	Technical knowledge

Organizing Knowledge

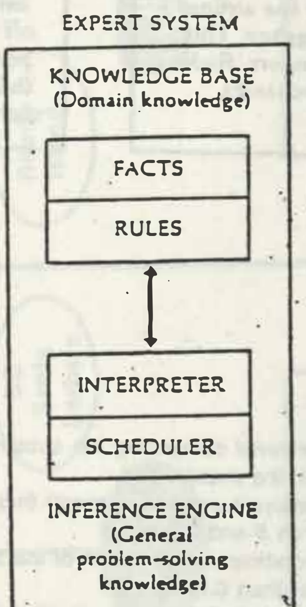


FIGURE 3.3 The structure of an expert system.

Representing Knowledge

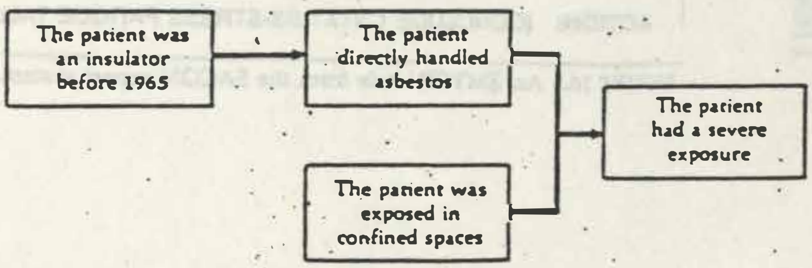
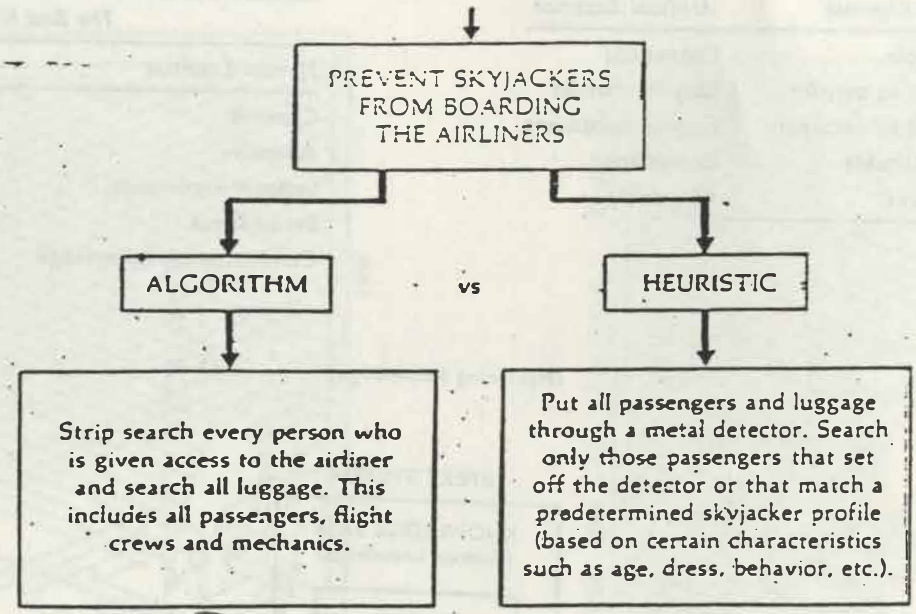


FIGURE 3.4 Inference chain for inferring the severity of asbestos exposure.



IF: 1) The material composing the substructure is one of: the metals, and
2) The analysis error (in percent) that is tolerable is between 5 and 30, and
3) The nondimensional stress of the substructure is greater than 0.9, and
4) The number of cycles the loading is to be applied is between 1000 and 10000

THEN: It is definite (1.0) that fatigue is one of the stress behavior phenomena in the substructure.

English translation of the EMYCIN rule shown below.

```
PREMISE: (SAND (SAME CNTXT MATERIAL (LISTOF METALS))
               (BETWEEN* CNTXT ERROR 5 30)
               (GREATERP* CNTXT NO-STRESS 0.9)
               (BETWEEN* CNTXT CYCLE 1000 10000))
ACTION:  (CONCLUDE CNTXT SS-STRESS FATIGUE TALLY 1.0)
```

Actual EMYCIN rule.

FIGURE 10.3 An EMYCIN rule from the SACON expert system.

NASA SPAN/OCEAN NETWORK

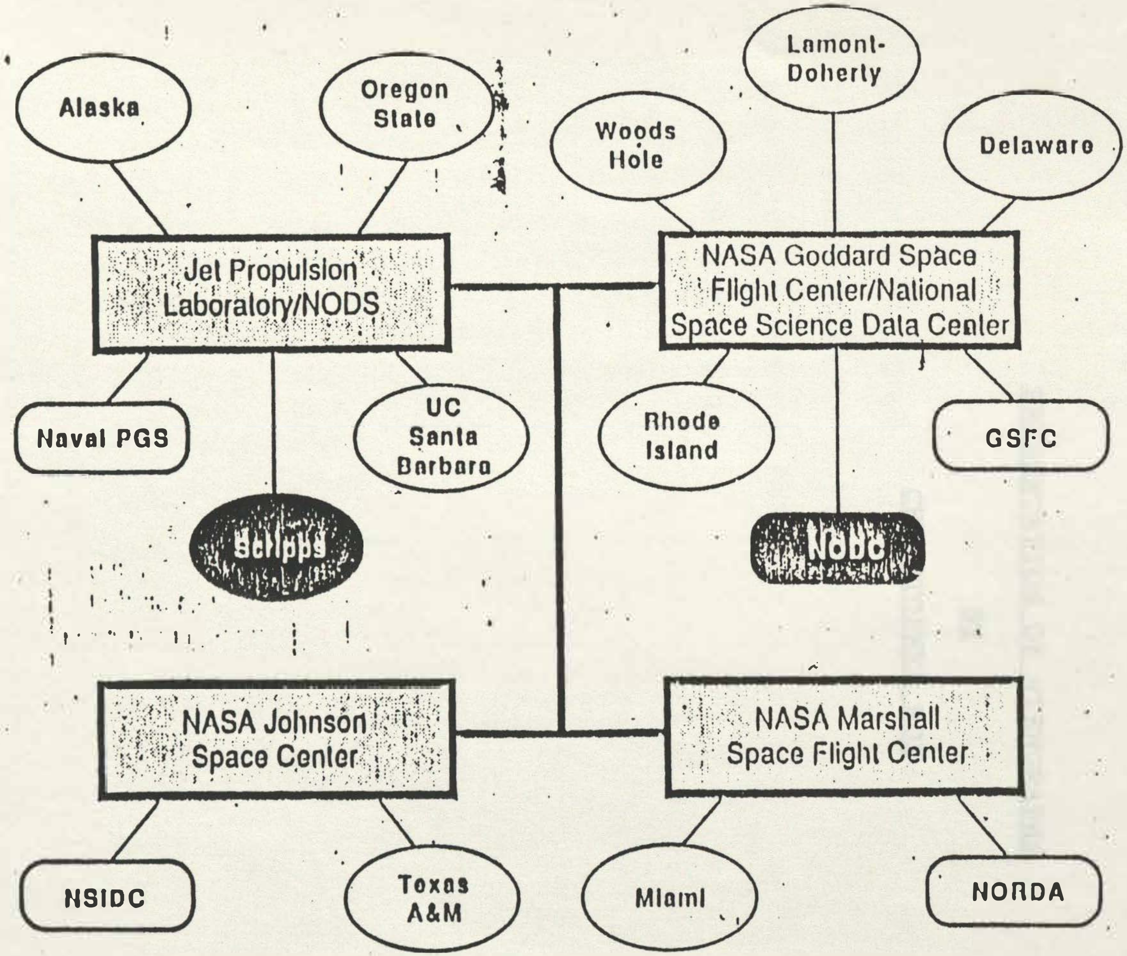


FIGURE 2. DIAGRAM SHOWING THE TELECOMMUNICATIONS LINK BETWEEN NODC AND JEDA CENTER AT THE SCRIPPS INSTITUTION OF OCEANOGRAPHY VIA THE SPACE PHYSICS ANALYSIS NETWORK (SPAN).

PRESENTATION OF VIEWGRAPHS

BY

CHRISTOPHER NOE

CONTINUOUSLY UPDATED
DATABASE

* "ON LINE"
* ACCESSIBLE VIA SPAN
* INTERFACED TO QUALITY SYSTEM
* SAVES "BEST" THERMAL DATA
* INVENTORY INCLUDED
* INDEXED ACCESS

CONTINUOUSLY UPDATED

DATABASE

WHY?

* USERS REQUIRE LARGE DATASETS
* "BATCH" ARCHIVES:
    - NOT USER RESPONSIVE
    - POOR DATA EDITING ACCESS
* "CRUISE" CONCEPT OUTMODED
* MULTIPLE ARCHIVES:
    - CONFUSE USERS
    - DUPLICATED DATA
    - SEGREGATED INVENTORIES

CONTINUOUSLY UPDATED

DATABASE

REPLACEMENT HIERARCHY

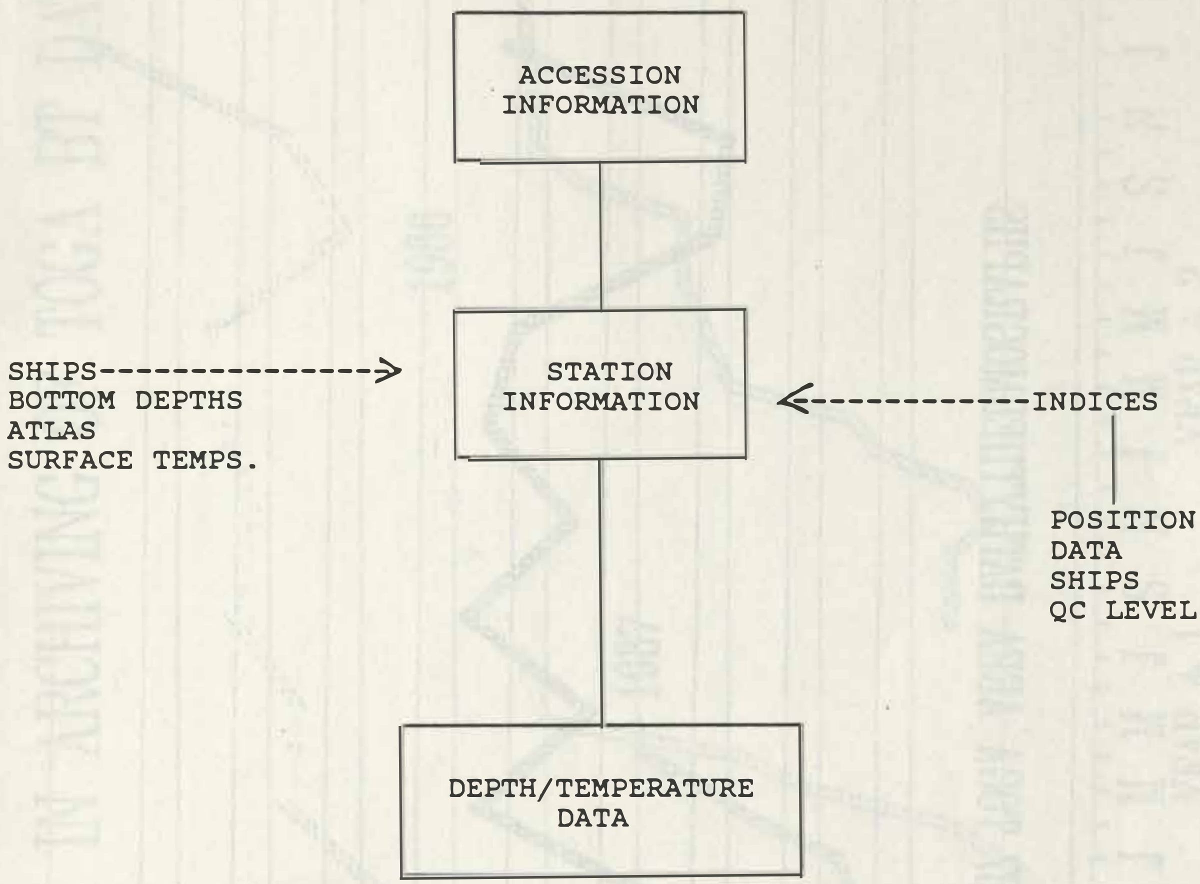
* RADIO MESSAGE
* LEVEL IIB RADIO MESSAGE
* DELAYED DATA (SELECTED LEVELS)
* DELAYED DATA (INFLECTION POINTS)
* LEVEL IIB DELAYED

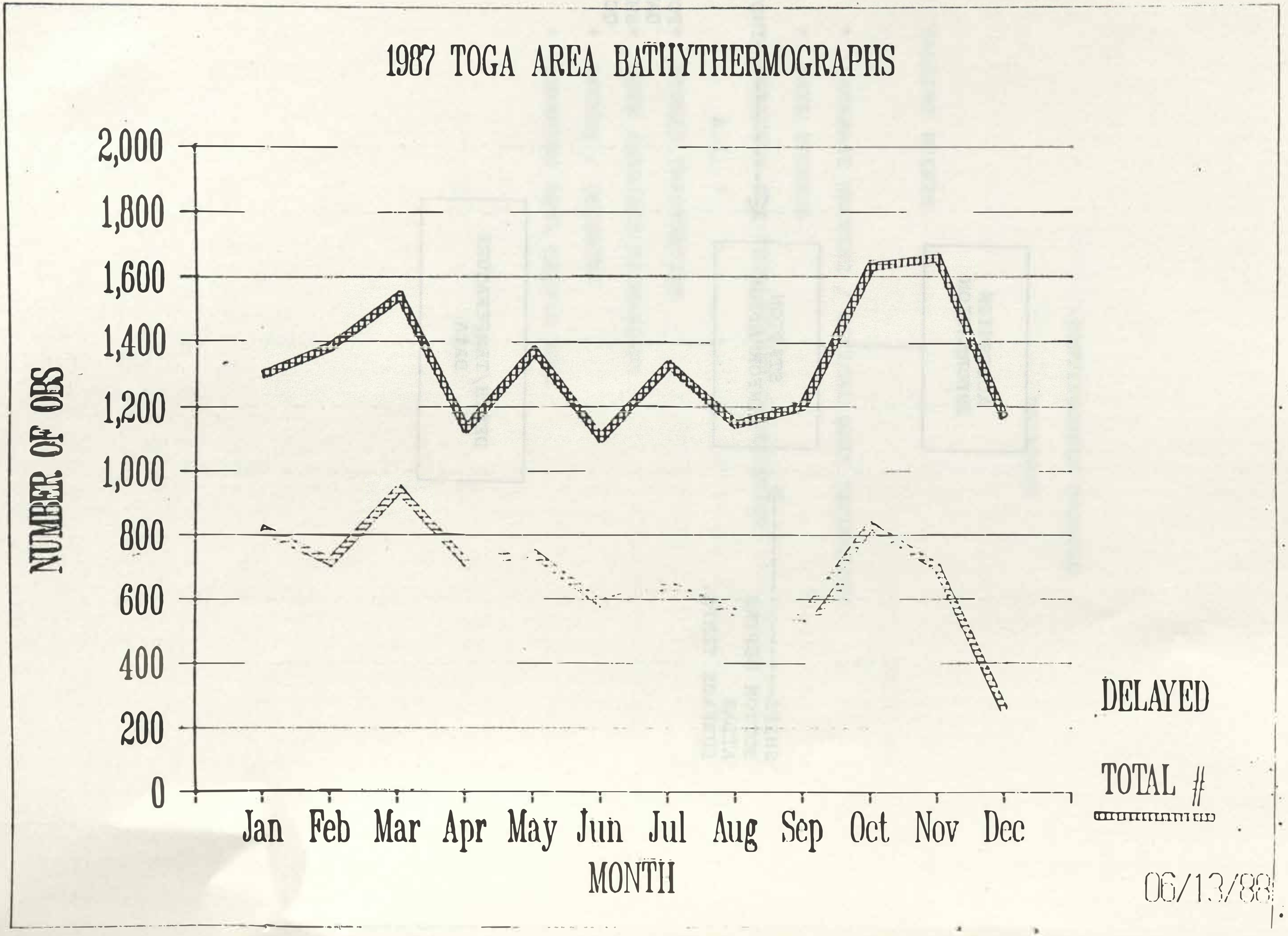
CONTINUOUSLY UPDATED

DATABASE

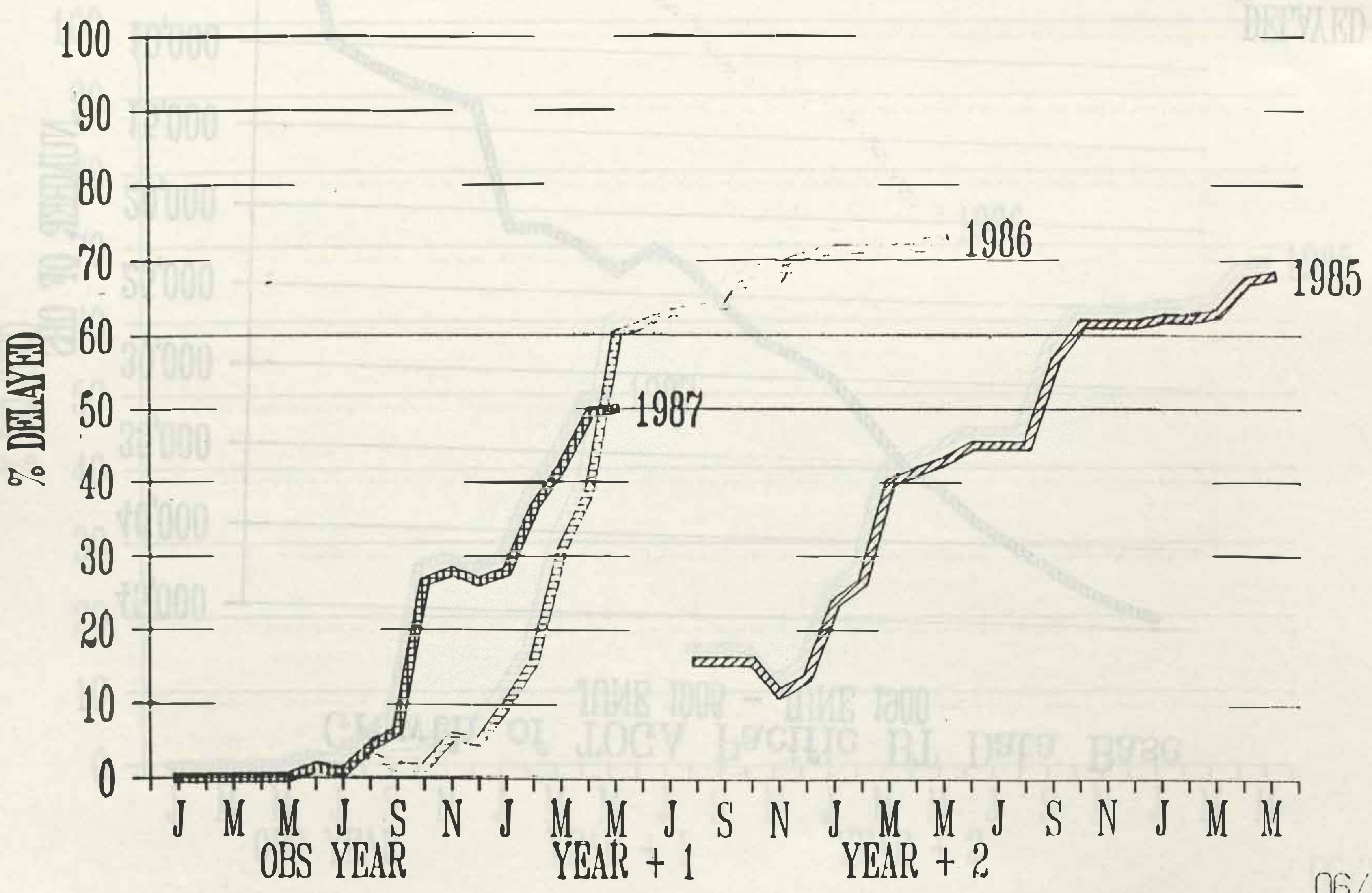
TALKING POINTS

* DATABASE MACHINE / OPTICAL DISK TECHNOLOGY
* DATA STREAMS
* STRATEGY FOR "DUPLICATE" ELIMINATION
    A.I. ?
* STORED INFORMATION
* USER FEEDBACK MECHANISMS
* ACCESS / SECURITY
* CHANGING DATA CENTER ROLE


ACCESSION INFORMATION
STATION INFORMATION
DEPTH/TEMPERATURE DATA
SHIPS
BOTTOM DEPTHS
ATLAS
SURFACE TEMPS.
INDICES
POSITION
DATA
SHIPS
QC LEVEL


1987 TOGA AREA BATHYTHERMOGRAPHS
NUMBER OF OBS
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
MONTH
DELAYED
TOTAL #
06/13/88

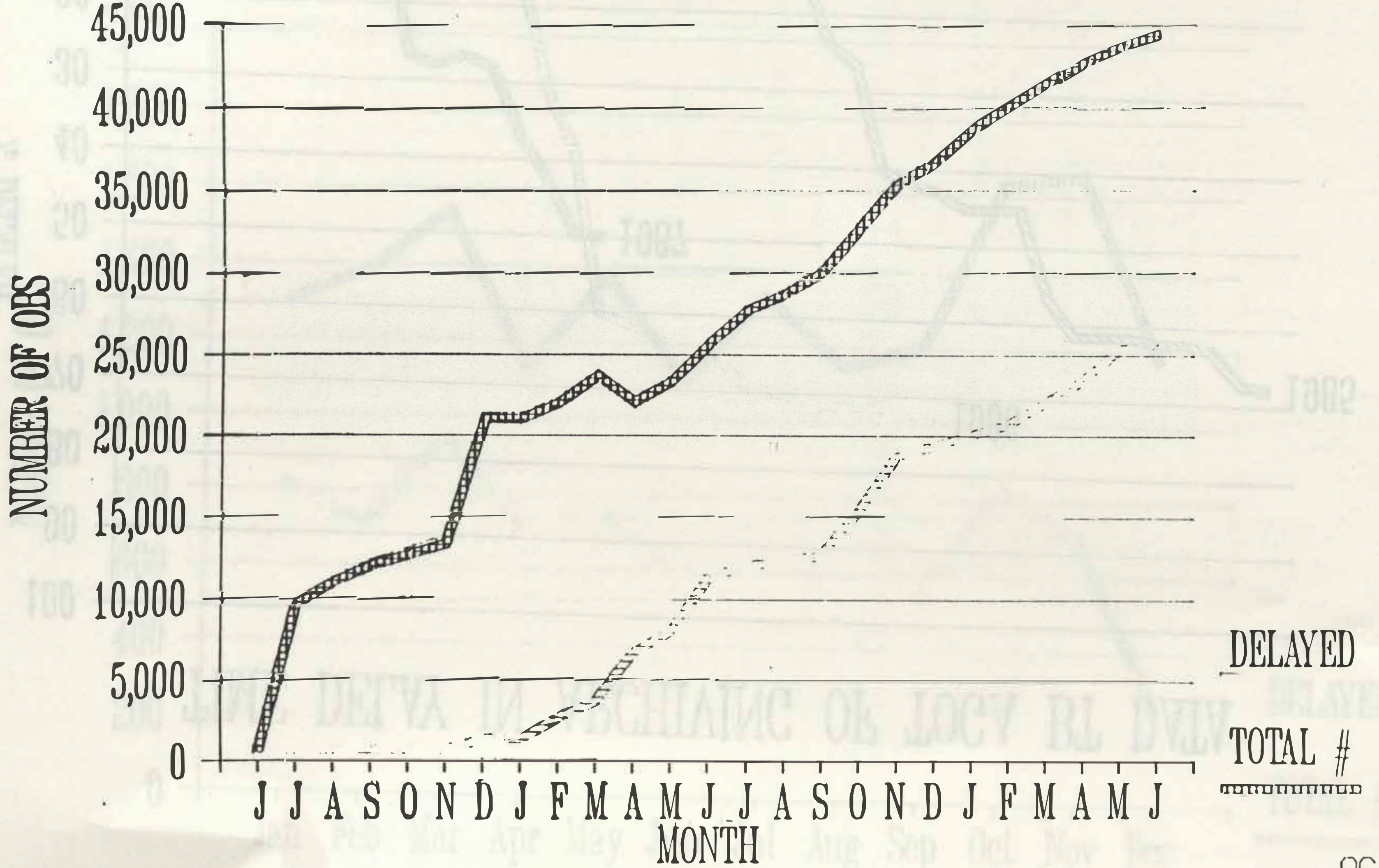
TIME DELAY IN ARCHIVING OF TOGA BT DATA



06/13/88

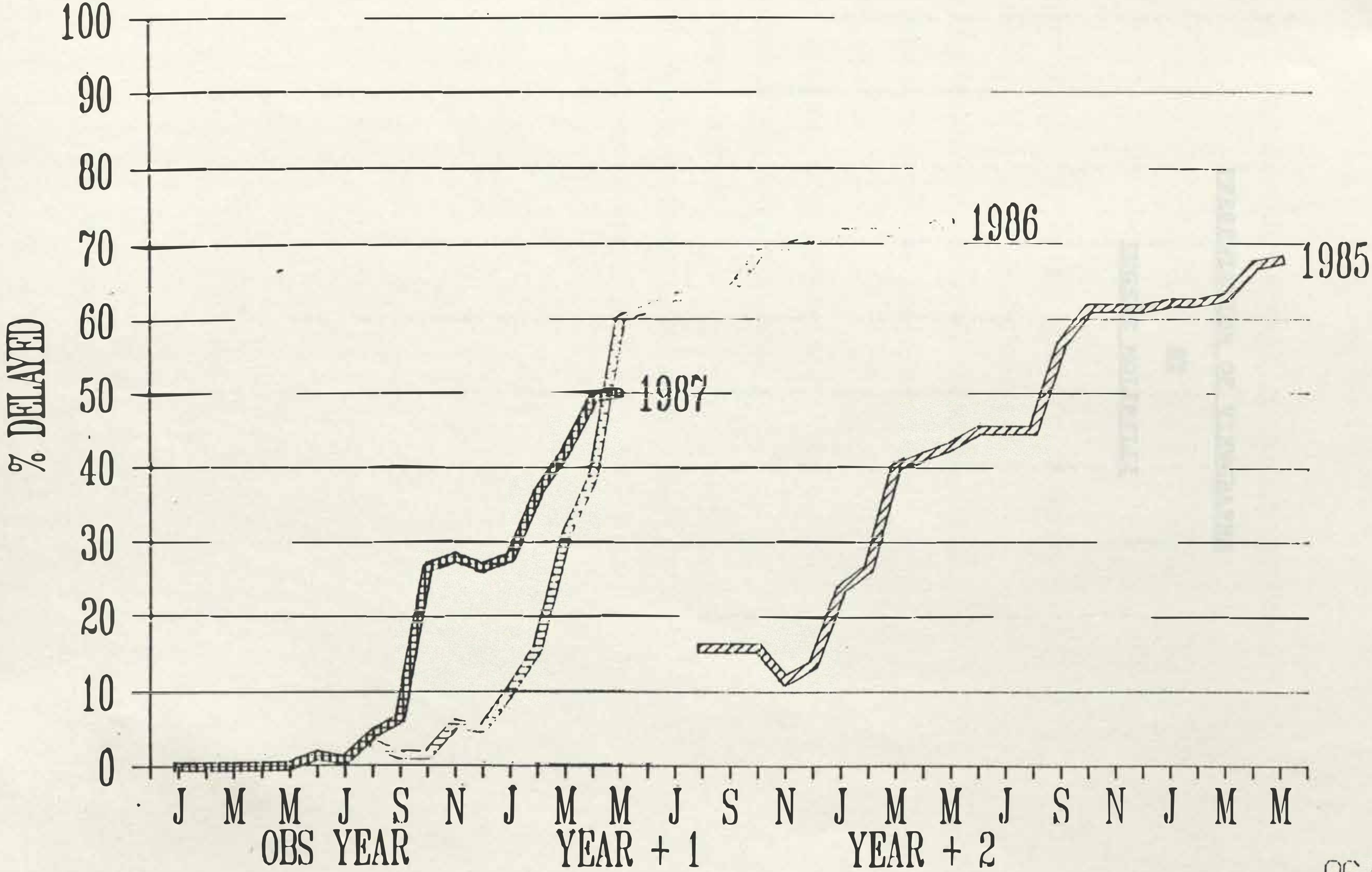
Growth of TOGA Pacific BT Data Base

JUNE 1986 - JUNE 1988



06/13/88

TIME DELAY IN ARCHIVING OF TOGA BT DATA



06/13/88

PRESENTATION OF VIEWGRAPHS

BY

EUGENE MOLINELLI

NEW MOODS MANAGEMENT SYSTEM AT NAVOCEANO

OBJECTIVES

* ACCOMMODATE ALL THE DATA AVAILABLE IN MOODS
* ACCOMMODATE DATA FROM OTHER SOURCES
* ALLOW GENERATION OF EXISTING PRODUCTS (GDEM)
* ALLOW GENERATION OF NEW PRODUCTS
* PERMIT UPDATES
* PERMIT FLEXIBLE DATA RETRIEVAL
* PERMIT EFFICIENT DATA RETRIEVAL
* ACCOMMODATE EDITING
* ACCOMMODATE HIGH-VERTICAL-RESOLUTION DATA
* ACHIEVE COMPUTER INDEPENDENCE

PRESENTATION OF VIEWGRAPHS

BY

SYD LEVITUS

TOTAL NUMBER OF HYDROGRAPHIC CASTS 584,013

FREQUENCY DISTRIBUTION OF THE NUMBER OF INSTABILITIES PER HYDROCAST

NUMBER OF INSTABILITIES →→→

	1	2	3	4	5	6	7	8	9	10
NUMBER OF CASTS	36649	7505	2154	749	308	91	29	11	3	1

INSTABILITY CRITERIA AS DEFINED ON PAGE 4 OF CLIMATOLOGIAL ATLAS OF THE WORLD OCEAN

0-30m $\delta\rho = .00003$

50-400m $\delta\rho = .00002$

500-5500m $\delta\rho = .00$

$$E = \frac{1}{\rho}\frac{\delta\rho}{dz}$$

K

K+1

Adiabetic displacements of water mass from level K+1 to level K.

NOTE: ~~[illegible]~~

σ_θ shows inversions on order of .02 mg/cm³ (N. Atlantic) but the data are O.K.

Entire cruises show up as containing many instabilities

Table 3

Depth Range (m)	Number of Hydrographic Observations	magnitude of instability (10^{-5} g cm^{-3}) 0 - 2	2 - 4	4 - 6	> 6
0-10	60368	7643	2253	981	1682
10-20	58151	5718	1120	393	425
20-30	56274	3807	545	206	197
30-50	54330	1873	409	135	172
50-75	50840	762	169	68	86
75-100	47248	358	88	37	72
100-125	45294	173	55	28	23
125-150	43163	119	44	19	19
150-200	42611	53	34	17	33
200-250	40439	72	47	21	24
250-300	38776	75	26	13	18
300-400	38003	47	26	10	31

Table 3. Frequency distribution of density instabilities as a function of depth and magnitude of instability ($\delta\rho$) for all data in Marsden Squares 1-90 as of 1978.