Status of Other New Data Sets for COADS

Steven J. Worley,
National Center for Atmospheric Research

Abstract

The number of input data sources available for the upcoming Releases la, 1b, and 2 is greater than was used in COADS Release 1 (April, 1985). The reasons for the increased availability include: new and more durable automatic recording devices for the ocean environment, data collection using satellite telemetry, more national and international data exchange programs, and large scale research data collection efforts such as TOGA and WOCE. Each new input data set brings with it new questions concerning data format translation and proper methods for merging with other input data sets. Five of these new data sources will be discussed in order to illustrate their likely contribution to future releases of COADS.

1. Introduction

Modern technologies (satellites and electronic data transfer), more durable automated instrumentation, and data exchange agreements and data collection from research programs like WOCE and TOGA all serve to make more data available for COADS. Therefore, the number of input data sets available for Releases la, 1b, and 2 is greater than what was available for COADS Release 1 (April 1985). Since in COADS the individual reports have uniform format a format translation is required for each input data source. This seemingly simple task is sometimes not as straightforward as one might think. Preexisting errors in the original data and misinterpretation of format rules can all lead to unexpected results. To guard against these pitfalls the computer programs that translate the format are independently doubly written, i.e. two programmers independently write translation programs and strictly compare the output. COADS Release 1 was primarily a collection of shipboard observations. Future COADS Releases will again be dominated by ship observations, but other types of data will be included. Moored and drifting buoy data will provide new opportunities to enhance the COADS, but these types of data need special consideration. Automated buoy measurements are inherently different than the shipboard observations. Buoy instrumentation has built-in averaging unlike synoptic observations taken on ships. The temporal sampling at buoys is also often much more frequent than standard procedures followed on ships. These differences raise questions as to how to properly merge these and other data types.

Several data sets that have not been specifically described in other papers in this volume will be briefly discussed (Woodruff et al. 1992 list all data sets planned for COADS Releases la, 1b, and 2). The usefulness and uniqueness of these data to COADS will be illustrated. In the following text, the archive ending dates should be regarded as flexible since there are cooperative agreements in place that will provide newer data as it becomes available.

2. Inter-American Tropical Tuna Commission Data Set

The Inter-American Tropical Tuna Commission (IATTC) has provided a tropical Pacific Ocean environmental observational data set for 1972-1991 from the tuna fishing fleet. Environmental observations are taken by the captain and crew (called logbook data), and by fisheries management observers (called porpoise observers data). The data set has approximately 0.5 x 10^6 reports of SST, wind components, visibility, and cloud cover. For example, the locations of logbook and porpoise observer data for 1989 are shown in Figure 1. The incorporation of these reports will improve the regional statistics in the eastern tropical Pacific Ocean. Through the continuing data exchange agreement with IATTC, reports for the western tropical Pacific Ocean will also be included in COADS.

3. Pacific Marine Environmental Laboratory Data Set

NOAA's Pacific Marine Environmental Laboratory (PMEL) has made available 30 tropical Pacific Ocean time series. Two groups have provided data for COADS. From the first group (M. McPhaden and P. Freitag) we have obtained the 1980-1989 quality controlled records from eight EPOCS moored buoys and four island stations with very low topographic elevation. These reports are all daily averages of measured wind components, air temperature, and SST (at the buoys). Strictly speaking, daily averaged data are very different than near instantaneous observations taken from ships. Nevertheless, since these data have been carefully quality controlled and calibrated, and the mooring locations are in relatively sparsely sampled regions (Figure 2) they are an important addition to COADS. Although these data cannot be treated as synoptic marine observations they will be included in COADS as individual reports and may be combined with the synoptic data in higher level data products, e.g. the Monthly Summary Trimmed Groups. The second PMEL group (S. Hayes, L. Mangum, and L. Stratton) has contributed the time series data (1984-1989) from the TOGA/TAO Atlas moored buoys. The spatial distribution for these ocean moorings is shown in Figure 2. There are two forms of this data, daily averaged data and basic averaging interval data, where basic averaging interval denotes the averaging period built into the electronic instrument package. The basic format data is more compatible with synoptic ship reports, but still is significantly different. For example, SST basic averaging intervals vary from 4 to 24 hours while wind component records are at every 1 to 6 hours. Although these measurements are generally better than daily averages they still remain different enough that special consideration must be given to them when the data are merged with other synoptic observations.

4. Marine Environmental Data Service Data Set

Drifting buoy data from the Marine Environmental Data Service of Canada (MEDS) will also be a component of COADS Releases la and 2. A cooperative agreement was made with MEDS in 1989. NCAR and NOAA/ERL supplied the NMC moored and drifting buoy data for 1980-1985 to MEDS. MEDS then returned the data after performing quality control that included location track, data range limits, and data time series comparison checks. MEDS has also collected, quality controlled, and provided 1986-1991 drifting buoy data for inclusion in COADS. This archive now has over 4 million reports of sea level pressure, air temperature, and SST. The data are important for COADS because they provide data measurements in some very poorly sampled ocean regions. As an illustration, Figure 3 shows all the buoy locations and tracks for 1989 (note especially the reports coming from the tropical Pacific Ocean and the high southern latitudes). These data require a “good” method for combining frequent buoy reports with sporadic ship reports.

Consider how the many sequential observations from a single drifting buoy in a remote area could easily overwhelm all other measurements taken from passing ships. Furthermore, this circumstance could lead to a monthly mean estimated parameter that indicated statistical significance (many samples and small standard deviation), but nevertheless would be strongly biased by the buoy data and totally erroneous if the buoy instrumentation happened to be faulty.

5. Russian Marine Meteorological Data Set

Data exchange agreements between the U.S. and the former Soviet Union’s World Data Centre B have made available for COADS the Russian historical set of marine meteorological data (MORMET) (Yudin et al. 1992). This is a basic set of observations taken from ocean vessels covering the globe for the 1888-1989 time period. The archive has more than 25 million reports with measurements of wind speed and direction, air temperature, SST, sea level pressure, visibility, cloud cover, and wave observations. This archive has existed for many years. During that time it has been in the custody of several different agencies, gone through format changes, and has been updated using various methods. Because of these circumstances some of the historical background is now unknown (as is the case with most very long term archives). For these reasons and others we are not certain how many MORMET reports are unique and not part of other input data sources already scheduled for COADS. The following description of MORMET is based on the archive that was resident at NCAR in February 1991, i.e. data taken during 1888-1989. Table 1 shows the major contributing countries (as specified by the country code within the data format) to the MORMET. Nearly the entire archive (22.8×10^6 reports) is indicated to have come from Russian ships or ships that have no identifying country. If the number of reports in the MORMET and COADS are compared over time (Figure 4) we note similar fluctuations in report count since 1960. The spatial distribution of the MORMET ship reports also looks much like that of COADS, with largest counts in the North Atlantic and North Pacific Oceans, and smallest in the Southern Ocean surrounding Antarctica. From these general data summaries it cannot be determined with any certainty how many unique ship reports the MORMET will contribute to COADS. Only studies that match the MORMET against COADS on a report by report basis can accurately answer that question. This will be done as the entire MORMET set is processed for inclusion into future COADS Releases.

6. National Oceanographic Data Center Data Sets

Five subsets of data from the National Oceanographic Data Center (NODC) will be used in COADS. Some but not all of these subsets, in a different form, were included in Release 1. All the subsets have in common the facts that, the stations have global distribution, have surface meteorological observations, and have ocean vertical profile data. For COADS we only require the surface meteorological data and a good estimate of the sea surface temperature. The Oceanographic Station Data (SD) archive has ocean temperature measured using reversing thermometers for the 1900-1991 period (815×10^3 reports). The Mechanical Bathythermograph Observations (MBT) collection has values of ocean temperature determined with mechanical thermometers (979×10^3 reports). This method of sampling was not often used after 1970.

Nevertheless, the data extend back to 1900 and provide very valuable information for the first 70 years of the century. The Expendable Bathythermograph (XBT) superseded the MBT instrumentation and uses a thermistor sensor to estimate water temperature. This data subset has 818 x 10^3 reports for the 1986-1991 time period. Another XBT data set, the Selected Level XBT observations (SBT) provide another 240 x 10^3 reports. And finally, the Compressed CTD/STD Oceanographic Stations (Co22) estimate ocean temperature using an electronic sensor package. This data set has 65 x 10^3 stations during 1969-1991. An illustration of the XBT station distribution in the Pacific, Atlantic, and Indian Oceans for the two year period 1985-1986 is shown in Figure 5a-c. In addition to the standard considerations when translating the NODC formatted data into COADS special care will be taken to acquire the best estimate of SST. This is especially true with the XBT archives. The ocean temperature measurements from this instrument can exhibit a thermal spike at the shallowest profile level. The spike is caused by the thermistor time lag as it adjusts from ambient air temperature to sea surface water temperature. Obviously, simply using the shallowest profile level temperature from an XBT as an estimate for SST would introduce unnecessary inaccuracy into COADS. When the NODC data is translated into COADS several near surface estimates from the profiles will be used to establish a “best” SST.

Table 1. The most frequent number of reports by country in the USSR Marine Ship Data Archive (1888-1989), as determined from the country code indicator.

Country	Number of Reports
USSR	11491646
no country code	11360601
West Germany	305929
United Kingdom	302560
United States	237077
Japan	230352
East Germany	137574
Netherlands	110990
France	41205

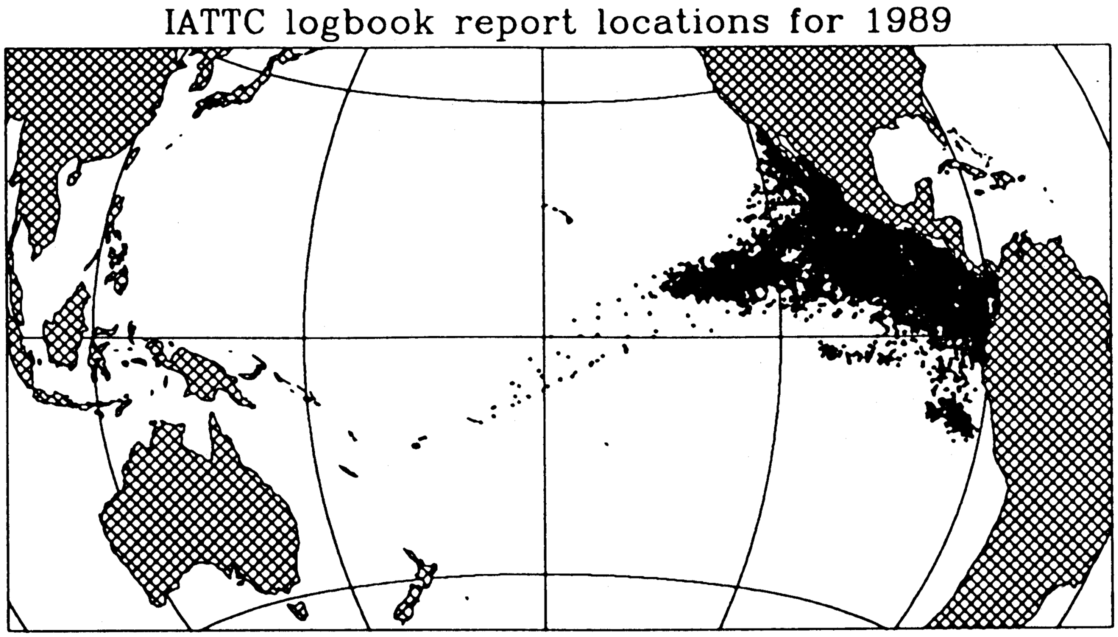
7. Conclusions

The variety and number of data sets available to create future COADS Releases far exceeds that used to develop Release 1. These new data will certainly enhance and improve the quality of COADS and thereby provide a better tool to understand the global ocean- atmosphere boundary layer as well as the coupled oceanic and atmospheric circulations. The variety of these new data types introduces many problems in creating a merged uniform format data set. Nevertheless, the potential usefulness of these combined data sets for many research efforts requires that these problems be properly addressed and all the available data be included in future COADS releases.

References

Woodruff, S.D., S.J. Lubker, and M.Y. Liu, 1992: Updating COADS - Problems and Opportunities. *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*. H.F. Diaz, K. Wolter, and S.D. Woodruff, Eds., NOAA Environmental Research Laboratories, Boulder, Colorado, (this volume).

Yudin, K.B., I.G. Ylyanich, V.N. Popova, and Ye. M. Krakanovskaya, 1992: Russian Marine Meteorological Data Set. *Proceedings of the International COADS Workshop, Boulder, Colorado, 13-15 January 1992*. H.F. Diaz, K. Wolter, and S.D. Woodruff, Eds., NOAA Environmental Research Laboratories, Boulder, Colorado, (this volume).



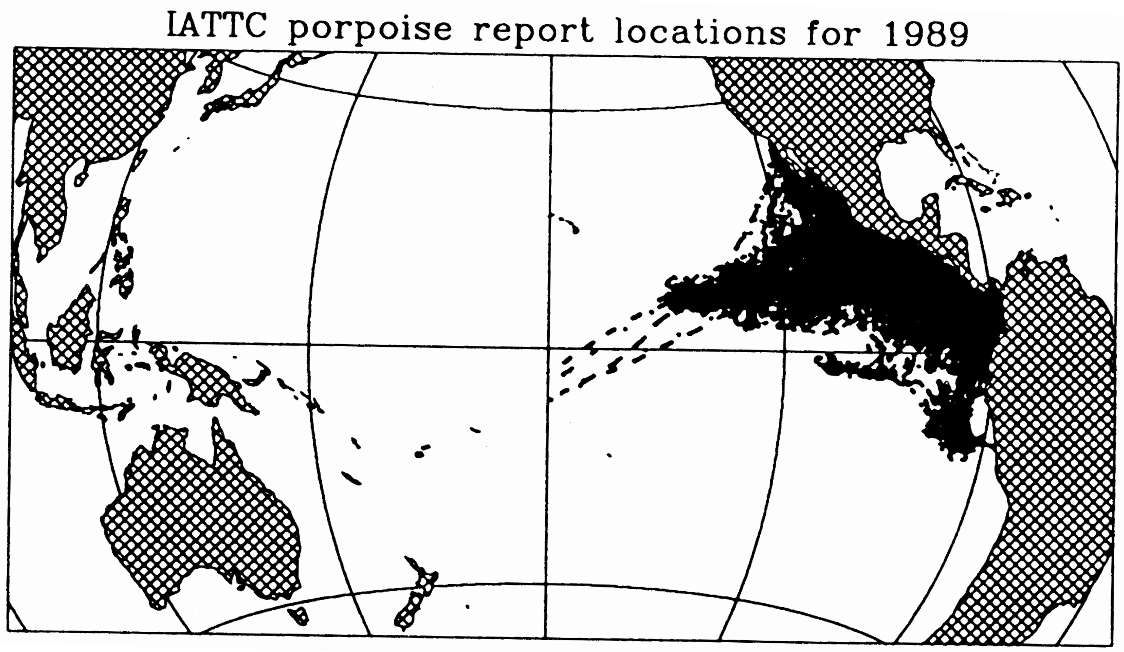


Figure 1. IATTC logbook and porpoise observer report locations for 1989. there are 15493 logbook and 32836 porpoise observer reports during the year.

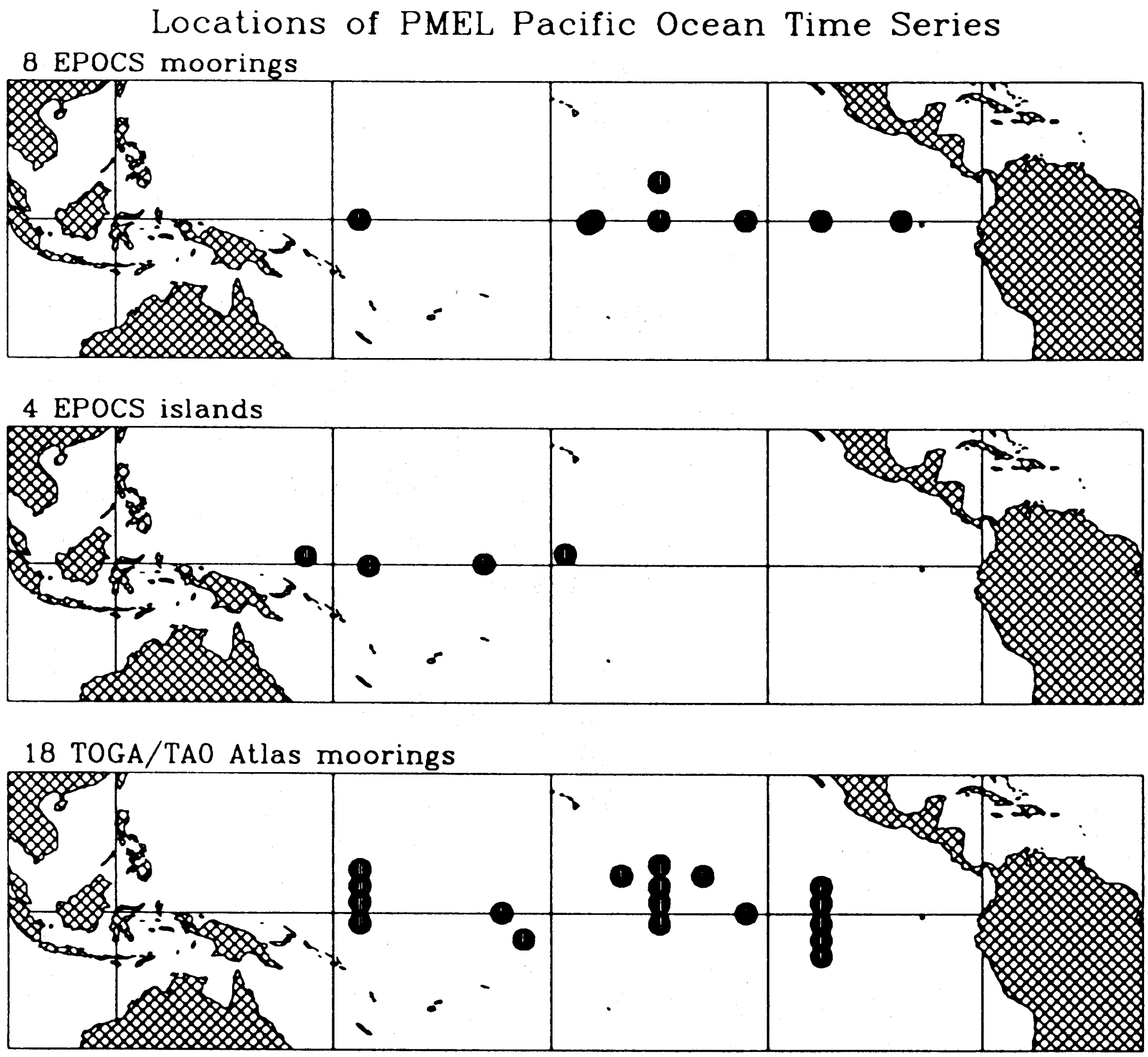


Figure 2. NOAA's PMEL EPOCS moorings and low topography island stations, and TOGA/TAO Atlas mooring locations.

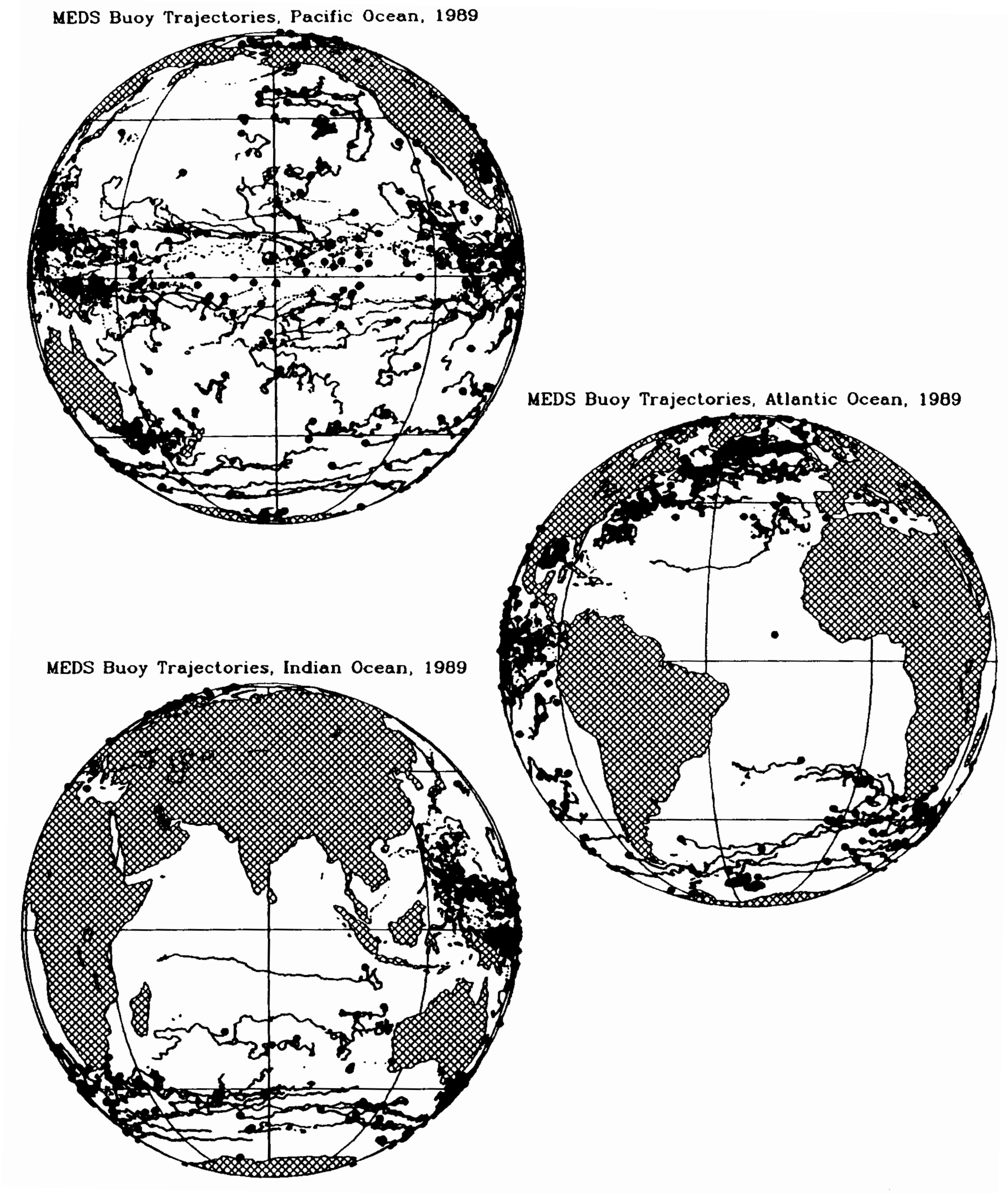


Figure 3. MEDS drifting buoy positions and trajectories during 1989. A large solid dot indicates the buoy starting location and small dots the report positions along the drifter trajectory.

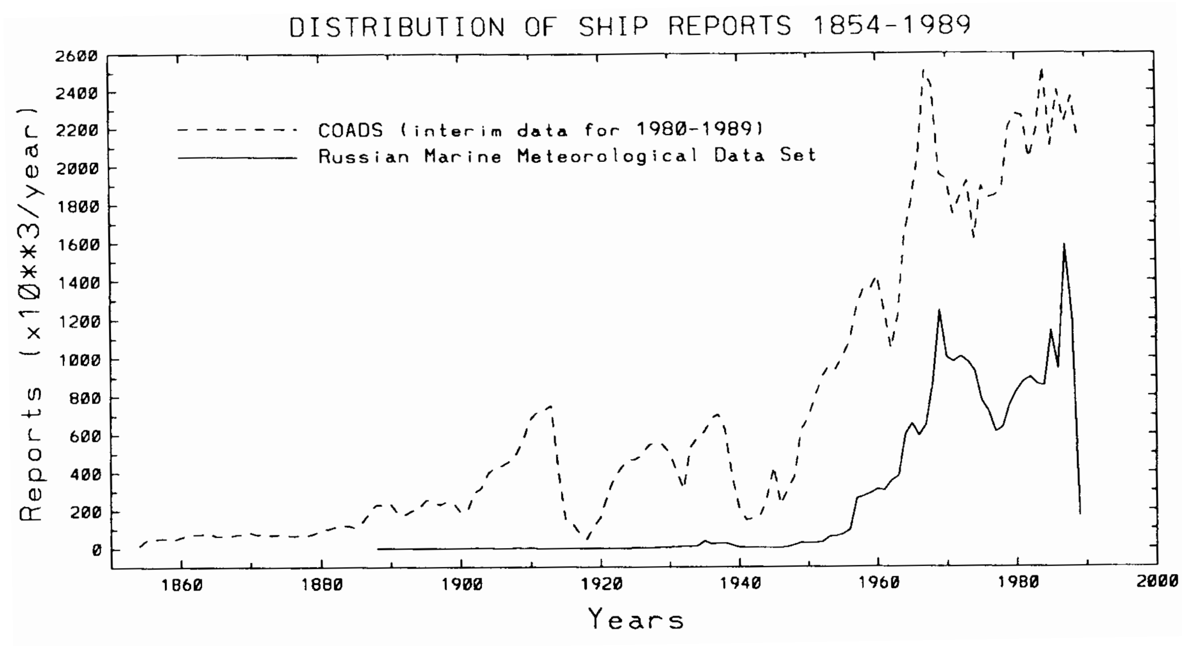


Figure 4. Distribution of ship reports in COADS and the Russian Marine Meteorological Data Set (MORMET). Values plotted are number of reports per year with COADS values for 1980-1989 taken from the interim data processing. Also the low value for 1989 in the Russian data shows that report entry for this year is only partially complete.

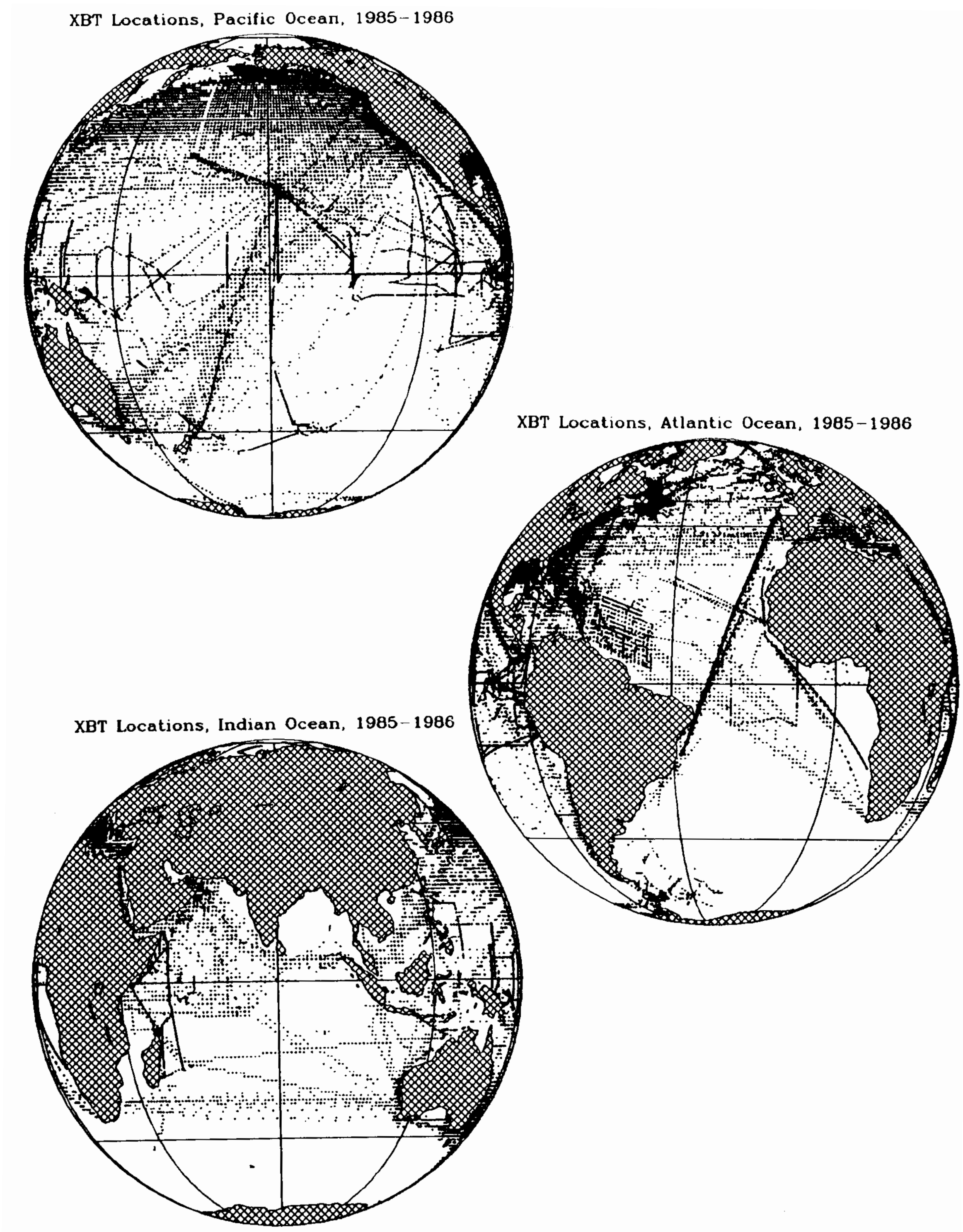


Figure 5. Distribution of NODC XBT stations for 1985 through 1986.