The COADS Ship Data Project and Publications, 1980 – 2003

- Some history of COADS, about 1980 – on.
- Brief description of the datasets of observations and statistics.
- Publications from COADS conferences.
  - A few pages from each.
- Added observations that will go into COADS.
- Plots of observations per year, and data coverage per decade.
- Ready Jan 8, 2004 (76 pages), document RJ0326

Document RJ0326

Roy Jenne
Jan 8, 2004
The COADS Ship Data Project and Publications, 1980 – 2003

Roy Jenne
Jun 0, 2003
Rev Jan 5, 2004

This document has introductory information about several documents that describe the world ship and buoy data (COADS) for over a century. A number of COADS documents are online at NOAA/CDC in Boulder, Colorado. A list of those by Scott Woodruff is included here.

1. COADS release 1 doc (1985, 2 pages here)

2. Ship data files (COADS) (R. Jenne, 1988, 9 p)
   • Describe the datasets for users.

3. A brief description of COADS documents (online at NOAA, Boulder). The new name of COADS is I-COADS (I = more international help), (Scott Woodruff, Jun 2003, 2 p).

4. Ship data history (R. Jenne, Jan 1989, 3 p)
   • Inputs from Rob Quayle, NCDC, Asheville

5. Programs of the National Data Buoy Center (Bull. AMS, July 1992, 2 p)
   • And map of fixed buoys in 1980

   • Has early history of COADS (2 p)
   • The importance of COADS for global reanalysis


10. Ship data (R. Jenne, S. Woodruff, Jul 1989, 4 p)
    • Data counts from sources.

11. COADS update notes (S. Woodruff, Feb 1989, 3 p)

12. Two pages from NCDC, Asheville
    • Counts of ship obs from Countries, 1968 – 1984
    • Estimate of US reports on ship forms for 1913 – 1948, done 1984

   - This has data coverage plots for decades for 1810 – 1949

15. Doc RJ0159: Selected Buoy and Sea Ice Information, 91 p, Jan 2002
   - Includes information about arctic buoys
   - Shows other arrays of buoys
   - Online at Data Support, NCAR ([http://dss.ucar.edu/docs/papers-scanned/papers.html](http://dss.ucar.edu/docs/papers-scanned/papers.html))

   See this online document about buoys.

Also remember: There were good surface and upper air observations from drifting Russian Ice Islands in the Arctic, mainly from 1954 – 1991.

16. World coverage of COADS reports, plots by decade (from Scott Woodruff, July 2003)
    Elements: SST, wind speed, SLP, relative humidity plots available.
    - Only SST and SLP given here.
    - Look at the NOAA online Web for more plots and plots in color.
Comprehensive Ocean-Atmosphere Data Set
Release 1

This doc is 1.5 cm thick

CIRES  University of Colorado/NOAA
        Cooperative Institute for Research in Environmental Sciences
        Ralph J. Slutz
        Sandra J. Lubker
        Jane D. Hiscox

ERL    U.S. Department of Commerce
        National Oceanic and Atmospheric Administration (NOAA)
        Environmental Research Laboratories
        Scott D. Woodruff

NCAR  National Science Foundation sponsored
      National Center for Atmospheric Research
      Roy L. Jenne
      Dennis H. Joseph

NCDC  U.S. Department of Commerce
      National Environmental Satellite, Data, and Information Service
      National Climatic Data Center
      Peter M. Steurer
      Joe D. Elms

Boulder, Colorado
April 1985
Foreword

To understand climate variability we must first delineate what kind of behavior must be understood. Do changes in the more energetic parts of the global climate machine occur gradually or suddenly? If there are clear "climate signals," where in the global domain do they appear first? How do they evolve in time? Do the signals reflected in various geophysical fields relate to one another in physically consistent ways? Do the forcing fields exhibit time variability that is consistent with the response fields? What does the behavior tell us about possible causes of climatic variability?

The opportunity to explore such questions has been severely limited by the availability of observations reflecting past behavior. Only since the advent of satellites have we been able to observe some few parameters on a global basis. Only since World War II have there been enough upper air observations to explore the vertical dimension and they are sparsely distributed. Only with surface observations can we extend the record of past behavior back into the last century.

In doing so, we find that the land stations having long records are too few to delineate spatial variability over the planet. Over the ocean areas, however, ship observations provide a richer record. They are good enough to delineate the time variability of the major wind systems and related fields of surface pressure and temperature.

The incentive for developing the Comprehensive Ocean-Atmosphere Data Set (COADS) was to make this record available to the individual investigator in a form that is reliable and easy to use. The global marine surface data set contains the most detailed record we will ever have of the dynamics of the global climate system over the last century and more. It should trigger rapid progress in understanding by making it possible to delineate the spatial and temporal characteristics of the several sharp adjustments of the global circulation that have occurred, and to glean from them clues to the nature and causes of global climate variability. COADS provides the material for diagnostic research to identify and explore the key questions. It also provides the needed boundary conditions for model simulation of the climate system variability.

It has taken four years and much effort by many individuals and several institutions to obtain and process the hundreds of tapes containing the basic data input. All of this effort was provided from ongoing activities; there was no appropriation identified for the task. It is a tribute to the spirit of cooperation among the participating organizations that the task has been successfully completed.

Throughout the effort, the support and encouragement of Dr. Wilmot N. Hess was crucial, as Director of ERL during the early stages and as Director of NCAR during the later stages.

Joseph O. Fletcher

Apr 1985
SHIP DATA FILES (COADS)

From 1982, NCAR (Boulder, CO), ERL/CIRES (Boulder, CO), and NCDC (Asheville) have worked together on a cooperative project to clean up several existing large files of world ship data and to merge files into a consolidated dataset with duplicates eliminated. The major component files were the "Atlas" files from Asheville that were used in the Marine Atlases and the Historical data files for 1861–1960 that were prepared under a major IDOE project involving several countries (especially Germany, UK, Netherlands, and US). In addition, all available exchanged ship reports (WMO Resolution 35 data) from 1961–79 were included. About 9 other smaller files were also merged in. This included surface SST from XBT reports. About 100 million ship reports were processed, resulting in 72 million after duplicate elimination. There were 53,185,975 reports output for 1854–1969 and 18,682,484 for 1970–79.

The processing steps to cleanup the data, run sort/merges, and calculate statistics took many hours on the CRAY computers at NCAR. Additional time was spent on other NOAA computers. The major sets of data and statistics through 1979 became available in mid-1984.

The volume of data is often rather high for the whole world ocean, but much of the observed data is organized by 10° lat-lon boxes, so that part of an ocean basin can be studied without many problems due to volume.

The data are divided into several separate sets of observations and statistics. The major sets are identified by product numbers in [3]. Where applicable these product numbers are given in the dataset descriptions which follow.

1. Basic Ship Synoptic Observations (including SST from surface level XBTs, buoys, etc).
   a. All 72 million observations in character form, 148 characters each report. This dataset is located at NCDC, Asheville. The volume is about $8.46 \times 10^9$ bits for the world, 1854–1979, (product #19).
   b. Basic binary dataset of observations at NCAR (called LMR, long marine reports, product #1).
      This set has all of the basic data, and data check information. It retains original data that was changed due to checking. For example, a temperature

---

1. Also Dennis Joseph (NCAR), Scott Woodruff (NOAA-ERL), Sandy Lubker (CIRES), and Dick Cram (NCDC, Asheville, N.C.)
2. Comprehensive Ocean-Atmosphere Data Set (Dataset #s at NCAR are DS 540.0 and DS 540.1)
with alphabetic characters would have been set to missing, but the original information was kept in an attachment. Only near-identical duplicate reports were eliminated. It retains some near duplicates not kept in set #a above. It also has a little more report status information.

The data for 1854-1969 and for 1970-79 are in two separate data streams. All data for each 10° ocean box is together within each data stream. Thus, it is easy to select data for one region of the ocean.

The total volume is about 39.5 x 10⁹ bits.

c. NCAR may prepare a 24 x 10⁹ bit version of the basic observations (in b) that has ship ID, and all basic data, but not all flags and comments. This may be sorted into synoptic order. We will do this large task only if several people need it.

d. Abbreviated binary dataset of observations (CMR, product #10)

This has nearly all of the basic data in observations, but little of the status or QC information. The information dropped was waves, swell, ship call sign, source ID and QC flags. It does have trimming flags from the statistics program — whether the data was outside of 2.8 or 3.5 standard deviations. Each report is 192 bits long. Total volume is 13.7 x 10⁹ bits.

e. Interim Product CMR data for 1980-1986

The sort is such that data in all 2° boxes over the world for one year-month are together (Sort: yr; mo; 2° box, day, hour, space). Exact duplicates eliminated. Some problems with selected data will be taken care of in the final version. About 15.8 million reports for the 7 years. Volume 3.03 x 10⁹ bits.

There is a version of the 1970-79 CMR data that is in this same sort.

2. Monthly Summaries trimmed, (two sort orders: product #15, sort is yr, mo, 2° box; product #16, 10° box, 2° box, yr, mo). (Period: 1854-1979)

Statistics were calculated for 19 variables, 14 statistics each. Any year-month-2° lat-lon box having one or more ship reports caused the batch of statistics to be produced that is somewhat larger than the set listed on Table B1-1a, attached.

In the whole 1854-1979 period there were 4,470,346 2° box-months with data, 3,712 bits each. Thus, the volume is 16.6 x 10⁹ bits.

One data order is a series of statistics within 2° boxes for the whole period of time. Data from 10° boxes can be selected so that a user can obtain data, for one area of the world, without having to consider the whole world. A second sort order has all global data together for each year-month.

Caution: All basic statistics are derived from all observations within a 2° box-month. Extremely wild observations are screened out, but statistics for adjacent boxes for the same month may still be quite different due to different numbers of samples, times of samples, and some data errors. Thus, the raw statistics will need to be further analyzed and smoothed in space and time for many uses.
3. Monthly Summaries in Group Files (product #18, trimmed)

These year-month statistics are in synoptic sort—all the world at one time is together. Extracted from item 2 above.

Group files: to make related elements easier to access, we have separated the data into five datasets, each with four variables and a reduced set of statistics for each (as given in Table B1-1a attached). This is the most widely used set of COADS statistics. If you only need SST for 1946-79, it is all on one tape.

The volume of each group of four variables is $1.72 \times 10^9$ bits, on two tapes, 6250 BPI. Statistics for a given lat-lon region can be selected from the whole set. The attached listing shows how the variables are grouped together.

The first tape of each group (trimmed), has statistics through 1945. The second has 1946-1979.

The format of the untrimmed statistics (product #17), is the same as for the trimmed except that the "fraction of observations in daylight" in the trimmed set was the "mean hour of observations" in the untrimmed.

4. Interim Product Statistics in Group Files for 1980-86

The data are put into the same five group files, but all statistics haven't been calculated. It has the number of observations and the mean. The data for all five groups are on one tape (6250 BPI). The format differs slightly from the pre-1980 data, but the change affects only the group number and checksum.

5. Monthly Summaries, untrimmed (product #13 and #14)
(Period: 1854-1979)

These data do not have as many variables as the trimmed set. Very wild values were excluded from use, but questionable data were included that were excluded from the statistics in items 2 and 3 above. Most users should use the trimmed data, not this.

The "trimmed" statistics (described above), include such variables as surface stress components, which are not included in the untrimmed set.

6. Decadal Summaries, trimmed (product #9)

These means for each decade were made by taking all available ship reports for the whole decade. Thus, if one year had many more reports than the others, it would have too much weight in the mean. It is better to prepare a mean directly from the yr-month data, perhaps with partial weighting of the ship counts.

7. Long-Period Means

Users have prepared some long-period mean products, based on the COADS statistics. There are three or four such products for SST.

For 1950-79, NOAA and NCAR can make available long-period means for 19 variables (each month, each 2° square). In this product, there is no smoothing across boxes and boxes with no data are not filled in.
8. User Software

Software for reading most of the COADS formats is available in listings (see ref [3]) and in computer compatible form on tape or floppy disk.

DATA VOLUME

Most of the data are stored on 6250 BPI tapes which can hold about $1.0 \times 10^9$ bits each. A 1600 BPI tape holds about $0.3 \times 10^9$ bits and an 800 BPI tape about half that. Details on data volume are available in [3].

UPDATE PLANS


DATA NOT YET INCLUDED

NCDC has identified many ship log forms for the World War I and II periods; about 17,000,000 observations. Other data also are not digitized:

- Germany: about 6 million ship obs., mostly for 1860–1890.

(Info re Germany and UK is as of 1980 (see WCP–19).

NOTE: $10^9$ bits of data are called one gigabit (one Gbit). It is also 125 megabytes. It will fit onto one tape, 6250 BPI.

References


[3] Slutz, et al, 1985: Comprehensive Ocean-Atmosphere Data Set; Release 1, NOAA ERL, Boulder, CO, 268 pp (NTIS PB86 - 105723). This has details about processing, contents, formats, etc.
SHIP STATISTICS; INFORMATION ABOUT GROUP FILES

0. Introduction

The seven group files are relatively compact alternatives to the full Monthly Summary Trimmed and Untrimmed (MST or MSU) formats, intended for studies using only a few variables or statistics. Eight important statistics for each of four related variables are grouped together in each file using a packed binary format. Thus five files are needed to represent all 19 MST variables, and two files are needed to represent all eight MSU variables. The statistics were chosen to bring together information that can be used to analyze the variability of the data and inhomogeneities of their distribution in time and space.

Cross reference is made to supp. A for standardized unpacking information, and the same notation for variables and statistics is followed or extended, also using the same type of two-dimensional table presentation.

1. Monthly Summary Trimmed Groups (MSTG)

The five trimmed groups were derived from MST (described in supp. A). Each MST was split into five MSTG records; these were written out onto the five separate group files even if every \( \alpha\beta \) (the value of the statistic \( \alpha \) for the variable \( \beta \)) was missing. Thus the record structure is identical for all the groups and their parent MST file. The five trimmed groups are numbered 3-7 to distinguish them from the untrimmed groups (numbered 1-2). Groups 3-7 contain four variables each: 3 = (S, A, Q, R), 4 = (W, U, V, P), 5 = (C, R, X, Y), 6 = (D, E, F, G), and 7 = (I, J, K, L). Table B1-1a shows the bit layout in common to any MSTG, and Tables B1-1b through B1-1f show the bit layout of each of the 64-bit or 16-bit sections of groups 3 through 7, respectively, in sequential bit-order reading from top to bottom. An example showing the bit-order is given following Table B1-1f.

The table numbers refer to the large COADS text.

2. Monthly Summary Untrimmed Groups (MSUG)

The two untrimmed groups were derived from MSU (described in supp. A). Each MSU was split into two MSUG records; these were written out onto the two separate group files even if every \( \alpha\beta \) (the value of the statistic \( \alpha \) for the variable \( \beta \)) was missing. Thus the record structure is identical for all the groups and their parent MSU file. Groups 1-2 contain four variables each: 1 = (S, A, P, Q), 2 = (W, U, V, C). Table B2-1a shows the bit layout in common to any MSUG.

The format of the untrimmed data is the same as the trimmed except that there is "mean hour of observations" (untrimmed) instead of "fraction of observations in daylight" (trimmed).

3. See the following page for the contents of the groups

- Data Groups 1 and 2 have untrimmed data
- Data Groups 3 thru 7 have trimmed data
### Table B1-1a
MSTG.1

<table>
<thead>
<tr>
<th>#</th>
<th>Statistic</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rptin</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>year</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>month</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>2 * box</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>10 * box</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>identification checksum</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>3/6 sextile (the median)</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>mean</td>
<td>64</td>
</tr>
<tr>
<td>9</td>
<td>number of observations</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>standard deviation estimate</td>
<td>64</td>
</tr>
<tr>
<td>11</td>
<td>mean day-of-month of observations</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>fraction of observations in daylight</td>
<td>16</td>
</tr>
<tr>
<td>13</td>
<td>mean longitude of observations</td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td>mean latitude of observations</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>total</td>
<td>384</td>
</tr>
</tbody>
</table>

*For the Interim product (80–86) files these 12 bits are divided into 4 bits for group number followed by 8 bits for a modulo 255 checksum.

### Table B1-1b
Group 3
64-bit or 16-bit Sections

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Bits</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S  sea surface temperature</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>A  air temperature</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Q  specific humidity</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>R  relative humidity</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>total</td>
<td>64</td>
<td>16</td>
</tr>
</tbody>
</table>

### Table B1-1c
Group 4
64-bit or 16-bit Sections

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Bits</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>W  scalar wind</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>U  vector wind eastward component</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>V  vector wind northward component</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>P  sea level pressure</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>total</td>
<td>64</td>
<td>16</td>
</tr>
</tbody>
</table>
Table B1-1d  
Group 5  
64-bit or 16-bit Sections

<table>
<thead>
<tr>
<th>#</th>
<th>β</th>
<th>Variable</th>
<th>Bits</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>C</td>
<td>total cloudiness</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>R</td>
<td>relative humidity</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>X</td>
<td>WU</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>Y</td>
<td>WV (14-15 are wind stress parameters)</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>total</td>
<td>64</td>
<td>16</td>
</tr>
</tbody>
</table>

Table B1-1e  
Group 6  
64-bit or 16-bit Sections

<table>
<thead>
<tr>
<th>#</th>
<th>β</th>
<th>Variable</th>
<th>Bits</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>D</td>
<td>$S - A = \text{sea-air temperature difference}$</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>E</td>
<td>$(S - A)W = \text{sea-air temperature difference \times wind magnitude}$</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>$Q_s - Q = (\text{saturation } Q \text{ at } S) - Q$</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>G</td>
<td>$FW = (Q_s - Q)W$ (evaporation parameter)</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>total</td>
<td>64</td>
<td>16</td>
</tr>
</tbody>
</table>

Table B1-1f  
Group 7  
64-bit or 16-bit Sections

<table>
<thead>
<tr>
<th>#</th>
<th>β</th>
<th>Variable</th>
<th>Bits</th>
<th>Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>I</td>
<td>$UA$</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>J</td>
<td>$VA$</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>K</td>
<td>$UQ$</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>L</td>
<td>$VQ$ (16-19 are sensible and latent heat transport parameters)</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>total</td>
<td>64</td>
<td>16</td>
</tr>
</tbody>
</table>

For example, group 3 contains, in order: rptin, year, month, 2 * box, 10 * box, and identification checksum, followed by

$$((\alpha \beta, \beta = S,A,Q,R), \alpha = s,m,n,e,d,h,x,y)$$

which defines the following matrix, with 4 rows and 8 columns:

etc., See [3]
Figure 2. Indian Ocean region log base 10 of reports (LMR) in 10 degree boxes for selected decades.
Figure 3-1. Global reports after duplicate elimination.
A brief description of COADS publications

Documentation On-line at the I-COADS Website (hosted at NOAA/CDC)
6 June 2003, S. Woodruff

The following published documents are available. In addition, electronic documentation for products (~3MB; mostly simple ASCII), Web information pages (HTML pages describing products or technical matters), ship metadata (copies of WMO Pub. 47 ship characteristics metadata currently covering 1973-99), and user software also reside on the I-COADS website. The possibility of mirroring or otherwise backing up all these holdings has been discussed, but no action taken yet.

Publications
Many I-COADS project publications are available on-line in PDF or HTML format via links from this webpage:

http://www.cdc.noaa.gov/coads/publications.html

Links (on the webpage) to on-line material are shown below in blue.

[Notes: Updated HTML edition including corrections (in red) with respect to the original printed volume. Includes links to PDF versions of the original printed document and of Jenne and Joseph (1974; Tech. for the Proc., Storage and Exch. of Data), both scanned by NCAR.]

[Notes: Not yet available in digital form.]

[Notes: Available from AMS website.]

[Notes: PDF format (each paper is a separate document).]

[Notes: HTML version, with some minor revisions.]

[Notes: HTML version, with some minor revisions.]

[Notes: PDF format (each paper is a separate document).]

[Notes: HTML version.]
[Notes: Supposed to become available in digital form on WMO website, someday.]

[Notes: HTML versions.]

[Notes: PDF obtained courtesy of WMO.]

**Selected Research Bibliography**

Three publications closely related to I-COADS are available in PDF format within the “Historical” category of this webpage:  
http://www.cdc.noaa.gov/coads/selected.html

[Notes: Obtained from NOAA/NODC website (containing Earth System Monitor articles currently back to 1995).]

[Notes: Imaged by NOAA/CDMP, with permission of the UK Met. Office.]

[Notes: Imaged by NOAA/CDMP. There are also ongoing CDMP plans to image the original marine card deck reference manuals (decks input to TDF-11), and other (non-marine) TD manuals at NCDC.]
SHIP DATA HISTORY
(Based partly on talk with Rob Quayle, NCDC)

1. U.S. Marine deck 116 started with data for 1948. It had a ship # or call sign. Deck 116 was later translated to deck 128 format. The ship name was probably carried along.

2. U.S. Marine prior to 1948. Very little of this has ever been key-entered. A little Navy data for the 1920s was punched. Put into deck 281 or something.


   This was defined about 1961. It is similar to the international IMM format which started about 1962. Deck 128 had 4 characters for ship name. The U.S. used them for U.S. ship names.

4. IMM International Data

   This started about 1962. It had 4 columns for ship name. He isn’t sure if other countries entered a name. If they did, he isn’t sure if Asheville saved the name in the conversion to TD-11. The names for U.S. ships were in the IMM code from at least 1973-on.

5. About 1971 Quayle and I talked about getting a ship ID routinely put into IMM data. Other countries balked for a long time. There were 4 characters they could use. Some names were longer than 4 characters. They did not want to give up the 80 column punch card format.

   The new practice, with ship name did not start until 1982.

   Questions: How many countries did put ship names into the IMM format prior to this time?

   If the names were not put into the data, are they still in national formats of the data?

6. HSST Data (1860–1960)

   Several countries prepared old data into this format which did not include a ship name. Was the data largely reformatted from other formats that included a ship name? How much new key entry was done for the project, where the ship name was left out?

7. Fixed Weather Ships

   Some of the European ships are still running, I think. Can we make sure that we have good time series of observations from these ships. Is it easy to get a tape?
8. Fixed Buoy Data

The U.S. has special datasets from off-shore fixed buoys. Do such clean datasets exist for some buoys near Canada, near the UK, and in the North sea? Does Japan have some?

9. Ship Identification

In pre-1970's data, the NCDC TD-11 deck only has ship #s for U.S. ships. There were books that gave the correspondence between ship names and numbers. The guess at NCDC is that these were thrown away in a riff in 1973. In 1976, when the program restarted, people could not find the books.

But didn't the old card decks have the ship name? If these card decks are around we could match up reports and derive the correspondence rules, if necessary.

Why identify ships?

• Can make a good track check to verify (and often correct, if necessary) the ship location

• Some ships have a bias in the pressure. The name allows one to detect the bias and correct for it.

• If a ship track is available, a very tight check on the consistency of pressure, temperature, SST can be made for use in an analysis. A correct value can even be derived.

• Given a ship name (or number), ships can be used to calibrate each other.

Sette (marine fisheries) made an Atlas of SST over the Pacific Ocean, each month, for about 1948–62. He compared each ship to every other ship to determine ship biases, etc. To do this, a person has to have the ship identification. Where did he obtain his data?

10. Ship Size, Speed, Etc.

These numbers aren't in the WMO dataset, but are in some other ship datasets. The Coast Guard had a dataset (AMVER?) that gave ship size, speed, and whether a doctor was on-board. For international search and rescue. Since 1982 the data has been in private hands and hard to get.

11. Joe Elms goes to a meeting in Paris, leaves 29 Jan. Joe, could you ask whether the main countries have:

• Data on ship size (tonnage and length). Do we care about length?

• Is it an intake SST temperature

• When ship was launched

• Name of ship as given in the ship reports

• Quayle says that the call signs may change when the radio changes. How is this handled?
- 3 -

- I understand that ship # and name may be reassigned. The launch data should help to determine whether there is a ship change. As long as a name is unique at a given period of time, the logic for track checks will be OK.

- Data from a port check. If a port officer changes an anemometer, by 3.6 mb, we should get a record of this fact.

12. Unpunched U.S. Ships, Pre-1948

The attached sheet shows an estimate of the number of unpunched U.S. ship obs prior to 1948. The info was prepared by Bruce Blankenship, NCDC. Bruce was mostly trying to tell us about U.S. Marine data to fill the war gaps.

- Is there also unpunched U.S. data prior to 1913?
- Was U.S. data for 1920-1936 ever punched?
- Note that Bruce listed obs for 1948. Wasn’t all of 1948 punched?

13. Unpunched Foreign Ship Data

What can other countries tell us about how much of their own old data has not been key-entered?

cc: Bob Quayle
Joe Elms
Dennis Joseph
Scott Woodruff
Steve Worley
Programs of the National Data Buoy Center

Abstract

Platforms of the National Data Buoy Center provide vital meteorological and oceanographic observations from data-sparse marine areas worldwide. The data are essential for real-time weather forecasting and research programs. This paper provides information on the data-acquisition systems, networks, monitoring capabilities, data processing and dissemination, data quality and availability, and related technology development for these platforms.

1. Introduction

The National Data Buoy Center (NDBC), a part of the National Weather Service (NWS), which is a component of the National Oceanic and Atmospheric Administration (NOAA), operates over 110 moored buoy and Coastal-Marine Automated Network (CMAN) observation stations for the primary purpose of acquiring environmental data for operational weather analyses and forecasts and other uses. Often, the first indication that forecasters have of rapid intensification or change in movement of storms come from these stations. In U.S. offshore and coastal waters, about 60% of all marine warning actions are instigated by these reports. Data from these stations are a vital part of the hurricane warning system (Sheets 1990).

The high quality of moored buoy measurements compared to other sources is discussed in Willkerson and Earle (1990) and Pierson (1990). As a result, buoys are used to provide ground truth for surface measurements from satellites. Space-based sensors are vulnerable to systematic biases that can only be compensated by reference to such ground-based buoys. Buoys are being used to develop algorithms for satellite retrieval of winds and other parameters.

Data requirements of NDBC systems are established through two general means. First, the NWS Office of Meteorology, Branch of Marine and Applied Services, identifies the location and measurement requirements for the basic NDBC station network. The NWS provides funding for ongoing operation and maintenance to support the NWS mission to provide routine forecasts and warnings. Additional requirements may be established by groups and agencies other than the NWS, who require special environmental monitoring. Data from all stations are released in real time and submitted to the archive centers regardless of the source of funding.

In addition, a network of approximately 35 drifting buoys is maintained in the oceans of the Southern Hemisphere for the Tropical Ocean and Global Atmosphere (TOGA) research program. National meteorological centers use NDBC data for hemispheric and global numerical analyses. NDBC data are also an important source of observations for research studies. Research programs on the marine boundary layer, wave generation and propagation, coastal ocean circulation analysis and modeling, ocean mixed-layer dynamics and thermodynamics, climate, pollution, etc., frequently use NDBC data.

Descriptions of previously existing programs were provided by Hamilton (1980, 1986). Since the 1986 article, new programs have evolved, and the number of observations from NDBC stations has more than doubled, to well over 1.2 million messages per year. These earlier papers have been referenced in the scientific literature for descriptions of NDBC data, but are now out of date. This article provides updated information about NDBC programs.

2. Moored-buoy network

Figure 1 shows the location of NDBC moored buoys. NDBC has standardized on the 6-m, boat-shaped NOMAD (Navy Oceanographic and Meteorological Automatic Device) buoy (Fig. 2) for deep-ocean deployments. It survives in the most severe environment, can be carried on the deck of U.S. Coast Guard (USCG) vessels for deployment, can be shipped on flatbed trucks to NDBC for refurbishing purposes, and is less costly to maintain and operate than large discus buoys. The 12-m-diameter discus buoys are used in areas of harsh environmental conditions, such as the Bering Sea and, in recent years, on two Great Lakes stations that were evaluated in a test program for year-round performance (most Great Lakes buoys

1992 American Meteorological Society

Bulletin American Meteorological Society

July 1992
are removed in the fall before the onset of the ice season. Seven USCG 12-m-discus, large navigational buoys (LNBS) have been instrumented by NDBC as part of C-MAN to provide observations to NWS. Ten-meter discus buoys (Fig. 3) have been found to capsize in the northeast Pacific Ocean and are now used in less-severe environments, such as coastal areas and the Gulf of Mexico. The 3-m discus buoy (Fig. 4), described in Hamilton (1988), is the hull that is normally used for coastal locations, Great Lakes stations, and directional wave measurements.

In the buoy network, there are normally 33 buoys (including the USCG LNBS) providing data to directly support NWS. There are presently 21 buoys funded by reimbursable users for their own purposes; however, the data are also available to NWS. Reimbursable users include the Minerals Management Service (MMS), the National Ocean Service (NOS), the U.S. Army Corps of Engineers, the National Aeronautics and Space Administration, and the Office of Naval Research. The location of all reimbursable buoy stations are subject to change with time. There is also one NDBC developmental buoy transmitting data.

3. Coastal-Marine Automated Network

In 1981, in response to the need for more coastal and offshore observations, C-MAN was established as a result of an NWS budget initiative. The core network of 40 stations was completed in 1985. Since then, additional C-MAN sites have been installed to meet a variety of requirements. In addition to NWS, other organizations sponsor 11 C-MAN installations to meet their requirements for NDBC high-quality marine/coastal observations. Among these are five major oil companies (through a joint government-industry program) and the Florida Institute of Oceanography. The total existing C-MAN network of 59 stations consists of 7 USCG LNBS and 52 fixed stations located on USCG lighthouses, beach areas, exposed fishing piers, and offshore oil platforms. A C-MAN location map is shown in Fig. 5, and a typical installation is shown in Fig. 6.

Baseline C-MAN measurements include wind speed and direction, peak wind, sea level pressure, and air temperature. At various sites in the network, this capability has been augmented to satisfy other spe-
US fixed buoys, 1980

Fig. 1. NDBO buoy locations.

See Bul AMS Sept 1980

Editors:
Henry F. Diaz  
NOAA/ERL Climate Monitoring and Diagnostics Laboratory  
Climate Research Division
Klaus Wolter  
University of Colorado/NOAA  
Cooperative Institute for Research in Environmental Sciences
Scott D. Woodruff  
NOAA/ERL Climate Monitoring and Diagnostics Laboratory  
Climate Research Division

August 1992

U.S. Department of Commerce
National Oceanic and Atmospheric Administration
Environmental Research Laboratories  
Climate Monitoring and Diagnostics Laboratory  
Climate Research Division  
Boulder, Colorado
Contents

(† Only abstract is included; ‡ contributed paper for this volume; presenter in italics)

Acknowledgments ........................................................................................................ iii
Abstract ....................................................................................................................... 1
Background .................................................................................................................. 1

Part 1. Overview

1.1 The Early History of COADS (J.O. Fletcher) ....................................................... 7 √
1.2 The Importance of COADS for Global Reanalysis (R.L. Jenne) ....................... 9 √

Part 2. Improvement of the COADS Database

2.1 Updating COADS—Problems and Opportunities
(S.D. Woodruff, S.J. Lubker, and M.Y. Liu) ............................................................... 19 √
2.2 Status of NCDC Keying of Historical Marine Data (J.D. Elms) ......................... 37
2.3 The Collection of Historical Ships' Data in Kobe Marine Observatory
(T. Uwai and K. Komura) ......................................................................................... 47
2.4 Blending of COADS and UK Meteorological Office Marine Data Sets
(D.E. Parker) ............................................................................................................. 61
2.5 Russian Marine Meteorological Data Set (K.B. Yudin, I.G. Ylyanich,
V.N. Popova, and Ye.M. Krakanovskaya; presented by V.N. Razuvaev) .............. 73
2.6 Status of Other New Data Sets for COADS (S.J. Worley) ............................... 81
2.7 Sifting Out Erroneous Observations in COADS—The Trimming Problem
(K. Wolter) .............................................................................................................. 91
2.8 An Analysis of Differences Between Ship and Buoy Observations
(U. Radok, S.J. Lubker, and T.J. Brown) ................................................................. 103

Part 3. Evaluation of COADS Wind Fields

3.1 Thirty Years of Tropical Pacific Winds from COADS: The FSU Winds†
(J.J. O'Brien) .......................................................................................................... 117
3.2 The Quality of Ship Observations in the Equatorial Western Pacific
(M.L. Morrissey) ...................................................................................................... 119
3.3 A Comparison of Surface Geostrophic Winds with COADS Ship Wind
Observations (H.F. Diaz, Fu C., and Quan X.) ...................................................... 131
3.4 Comparison of Estimated and Measured Marine Surface Wind Speed
(H.-J. Isemer) .......................................................................................................... 143
3.5 The Scientific Beaufort Equivalent Scale: Effects on Wind Statistics and
Climatological Air-Sea Flux Estimates in the North Atlantic Ocean
(H.-J. Isemer and L. Hasse) .................................................................................... 159
3.6 Estimates of Wind Stress and Heat Fluxes with Bias Corrections
(A.M. da Silva, C.C. Young, and S. Levitus) .......................................................... 179
3.7 The Wind Problem in COADS and Its Influence on the Water Balance
(Z. Wu and R.E. Newell) ....................................................................................... 189
3.8 Towards a Dynamically Constrained Analysis of Sea Level Pressure and Winds‡
(Y. Kushnir, S.E. Zebiak, M.A. Cane, and R.L. Tagett) ......................................... 201
3.9 Statistical Analyses of COADS Wind Data in Coastal Regions of the United
States‡ (F.A. Godshall and H.A. Walker) ............................................................ 211
Part 4. Evaluation of COADS Sea Surface Temperature Fields

4.1 Corrections to Pre-1941 SST Measurements for Studies of Long-Term Changes in SSTs (P.D. Jones and T.M.L. Wigley) .......................................................... 227
4.2 Transformation of 2° into 5° Grids for Global SST Trend Analyses (A. Spekat) ................................................................. 239
4.3 Improvements to an SST Climatology Using COADS (D.J. Shea, K.E. Trenberth, and R.W. Reynolds) .................. 245
4.4 Effect of Recent Volcanic Eruptions on Satellite-Derived Sea Surface Temperatures (R.W. Reynolds) .................................................. 253
4.5 Preliminary Comparison of North Atlantic SST Anomalies Between COADS and AVHRR-Derived Fields (G.R. Halliwell) ........................................... 261

Part 5. Evaluation of COADS Hydrological Variables

5.1 The Use of COADS Ship Observations in Cloud Climatologies (C.J. Hahn, S.G. Warren, and J. London) ...................... 271
5.2 A Comparison of COADS and Nimbus-7 Cloud Statistics (B.C. Weare) .......................................................... 281
5.3 Use of COADS in Global Water Budget Studies† (R.E. Newell and Z. Wu) ..................................................... 291

Part 6. COADS as an Analysis Tool I

6.2 The Value and Use of Averaged Homogenized Data Derived from COADS (P.B. Wright) ................................................... 303
6.3 Intercomparison of Global Marine Climatologies (S.L. Groth) .............................................................. 307
6.4 COADS as a Diagnostic Modeling Tool: Air-Sea Interaction and the Role of the Deep Ocean in Climate Change (H.P. Hanson) ............................................. 315
6.5 Monitoring Global Monthly Mean Surface Temperature (K.E. Trenberth and J.W. Hurrell) ......................................... 323

Part 7. COADS as an Analysis Tool II

7.1 Combined Southern Oscillation Index and Sea Surface Temperatures as Predictors of Seasonal Rainfall (J.S. Russell, I. Mcleod, and M.B. Dale) ............ 333
7.2 Effects of Arabian Sea Parameters on Indian Monsoon (S.S. Aralikatti) ...................................................... 341
7.3 Scientific Work at MPI in Hamburg in Context with COADS† (I. Jessel) .............................................................. 351
7.4 Relationships of Global SST Patterns to Australian Rainfall Variations† (W. Drosdowsky) ...................................................... 353

Part 8. Computer Applications

8.1 Climate Research Data Tools (CRDtools) and COADS (D.R. Mock and A.R. Messenger) .................................................... 357
8.3 Interactive Data Access System for Climate Researchers: The NASA Climate Data System Experience† (J. R. Vanderpool) ........................................ 375

Workshop Summary and Recommendations ........................................... 377

Address List of Participants and Contributors .................................................. 385
The Early History of COADS

J. O. Fletcher
NOAA Environmental Research Laboratories

We all learn at an early age that the first and most essential step in answering a question is to understand what the question is—and then, all too often we see that obvious principle violated.

In laboratory science we can often formulate the question ourselves and design experiments to test possible answers. For geophysical systems nature defines the behavior we seek to understand and we learn what that behavior is by observing nature. We must, also, test our hypotheses against this observed behavior. This is not as neat as a laboratory experiment but is unavoidable.

The problem of understanding climate change is further exacerbated by the long time scale. We have no choice but to look backward in time because we cannot afford to wait for the future to unfold.

When we look backward in time we find the observational record frustratingly incomplete. We have satellite records for only a couple of decades and very incomplete radiosonde coverage for about four decades. To document longer term behavior we have only surface observations and for reasonable spatial coverage of this water planet they must include data from the ocean domain.

I ran squarely into this limitation when I began to study climate change in the mid-1960s. My general approach was to look at the global system as a whole, considering first its most energetic parts and asking:

What factors force the motion?
Do they exhibit variation in time?
What physical processes are related to these variations?
Can they be described quantitatively?
What observational tests can be designed to evaluate concepts of how the machine works?

One of the strongest thermal forcing feature of the global system is Antarctica and the most steady energetic dynamical feature is the southern hemisphere westerlies, so an early effort was to look at the time variability of the Antarctic heat sink and the wind field. The record for the ocean area seemed to be a blank wall and in the effort to find such data I visited the U.S. Navy’s then Fleet Numerical Weather Center (FNWC) at Monterey, California.

FNWC was already struggling with the problem of compiling historical marine observations and in working with NOAA, NSF and WMO to organize international efforts. Taivo Levastu was the project officer and Paul Wolff, the Center Director, was pushing hard for progress. I think they deserve recognition for helping to generate the international effort to digitize the surface ocean record.

When I came to NOAA in 1974 I again tried to access this record. I found that an enormous effort had been invested by the international community in recovering and digitizing the marine data but it was not in a useable form for global studies.
The data collection effort really had started with Matthew Maury in the 1840s. In the 1930s Willard McDonald of the U.S. Weather Bureau initiated the first large-scale comprehensive marine climatic data base, resulting in publication of an atlas (McDonald, 1938). This was followed by Navy-supported work in Asheville, North Carolina led by Harold Crutcher, Norman Canfield, Dick Davis, Robert Quayle, and Joe Elms spanning the years from the 1950s to the early 1980s. The resulting “Atlas” data set of about 50 million observations on about 300 magnetic tapes formed the foundation for COADS.

Another major source of data eventually used in COADS consisted of observations from the Historical Sea Surface Temperature (HSST) Data Project, undertaken in the early 1970s and coordinated in the U.S. by Rob Quayle. In an effort to assure better data quality, only those marine reports containing SST observations taken by a bucket were to be included in the project. Unfortunately, the HSST data contained some gaps, as exchanges with all countries were not completed for some years and even then not all weather elements were included in the HSST data format.

Then in 1973 funding in Asheville for routine data entry and quality control of surface marine observations was cut by the Navy (about half of the data were already being received in digital form via telecommunications).

From ERL I tried to encourage NCDC and their parent EDIS (Environmental Data and Information Service) to push this important work but year after year in the 1970s their budget requests were given too low priority to be funded. Finally, in 1980, I went to Tom Potter, who was then heading EDIS, and proposed that we undertake the effort in concert, without funding, and without commitment to deadlines. Tom agreed and NCDC was given a green light.

Ralph Slutz, Scott Woodruff, and Sandy Lubker were the primary ERL and CIRES participants and it soon became clear that a huge amount of computer time would be needed, more than could be extracted from ERL’s CDC Cyber. The Director of NCAR, Bill Hess, agreed to provide this computer resource during the test phase of their new Cray supercomputer, before it was fully allocated to other users. Roy Jenne became the NCAR project supervisor, and Dennis Joseph and other staff members of NCAR’s Data Support Section were instrumental in getting the job done. Pete Steurer and Joe Elms became vigorously involved from NCDC.

We did not know at the time what an enormous task it would be to find and correct all the errors from various sources present in the many data sets that were collected nor did we fully anticipate all the other pitfalls. The work is still in progress a decade later and there are still many improvements to be made. I believe that this work deserves high priority for the reasons outlined above: only nature can define the behavior we seek to understand; we must look backward in time to define this behavior. Only the surface marine record provides the spatial and temporal extent that can help understand climate behavior; this record is the Rosetta Stone for interpreting the longer proxy records that can be extracted from countless other sources.

Reference

The Importance of COADS for Global Reanalysis

Roy L. Jenne
National Center for Atmospheric Research

Abstract

The existing daily atmospheric analyses for the period 1950-present are not of sufficient quality to permit the type of climate studies we need. There are several reasons for this: [1] there have been many changes of analysis procedures, [2] much of the observed data did not reach the operational centers in time for inclusion in the analyzed fields, and [3] in some cases the methods used for the analyses were not adequate. For example, it is only since about 1987 that the ocean surface-wind analyses have been good enough to drive ocean models, after some adjustment.

The COADS dataset that describes the surface marine conditions is of fundamental importance for reanalysis. It includes the world’s collection of ship observations, drifting buoys, moored buoys, ice buoys, and the surface level from ocean station observations, e.g., XBTs. The COADS dataset has much more data than can be obtained in real-time.

The archives from reanalyses will contain data each six hours. There will be the normal fields such as temperature, humidity and winds. Boundary layer analyses will be archived. Many diagnostic fields such as precipitation, evaporation and surface radiation fluxes will also be available.

Introduction

As early as 1985 there were discussions about the possibility of making a reanalysis that would help achieve the goals of the TOGA Experiment. J. Shukla (University of Maryland) and others have helped to push this task forward. Some efforts arguing for a reanalysis project can be traced to as early as about 1982. From 1987-1991 there has been an increasing interest in making daily reanalyses of the global atmosphere. The data assimilation methods used to analyze the state of the atmosphere have shown major advances during the 1985-1991 period. A forecast is a part of these methods; therefore, the analyses improve when the forecast model is improved and when the methods improve. There have been large advances in the capability of forecast models and analysis methods. These improvements mean that the benefits of making a reanalysis have become much greater than before.

Bengtsson and Shukla (1988) published a paper that helped to start the movement toward planning for reanalyses. In early 1989 a small workshop was held to consider the initiative of making Tropical Ocean, Global Atmosphere (TOGA) reanalyses (Kinter and Shukla, 1989). There was a clear interest in analyzing the period 1979-1990 starting with two years in the 1980s (1982-1983) as a pilot project. Interest in pursuing this effort has grown, and now there is a NMC/NCAR project to reanalyze the entire period from 1958 to the present. An associated reanalysis workshop was held in April 1991 (Kalnay and Jenne, 1991). Representatives from ECMWF and the University of Maryland were present.
1938-1948, including the fixed weather ship observations for 1940-1945, were maintained, at one time, on punched cards designated as Card Deck 115 (National Weather Records Center, 1953). These totaled approximately 946 thousand observations. In December 1959 it was decided by the "Card Deck Disposal Committee" that the number of errors existing in this particular deck warranted the cards disposal. In addition, the cards were not even filmed since the committee believed that many of the errors could not be corrected and it was likely that many of the observations were duplicated in other decks, particularly those acquired from foreign sources. In November 1960, the cards were destroyed and because of this action it is now impossible to determine if any unique observations were lost. It is likely that a) a significant number of the observations were from the fixed ships which have been retained in other digital files, b) the foreign observations have been received through subsequent data exchanges and c) most of the U.S. observations are now being re-keyed under the current scheme; therefore, few observations should have been lost because of this decision.

![Graph](image)

**Figure 1. Number of global observations per year**

In 1989, a pilot project began at NCDC to key the U.S. observations located in the National Archives that did not already reside in the digital data base for the period 1912-1946. It was the consensus at that time that the production keying to follow would likely be done outside the U.S.; therefore, CLICOM was selected as the system to be used for keying and quality controlling the data. A software development program was initiated (Bissing and Manns, 1991) to adapt the CLICOM system originally designed for fixed land stations for keying mobile ship observations. In an effort to replicate as much detail as possible on the original ship forms, such as the units (e.g. Celsius, Kelvin, or Fahrenheit), over 20 different form types were created for the CLICOM keying effort.
Figure 1. Annual numbers of observations in the area 10° - 20°N, 60° - 70°E. Hollow bars, COADS + MDB duplicates; light shading, unique to COADS; heavy shading, unique to MDB; heavy line, maximum of COADS and MDB if not blended. From Woodruff (1990).
Wind: → = 4 m/s.  SST: black < 15°C ≤ blue < 20°C ≤ green < 25°C ≤ red
A Comprehensive Ocean-Atmosphere Data Set

Abstract

Development is described of a Comprehensive Ocean-Atmosphere Data Set (COADS)—the result of a cooperative project to collect global weather observations taken near the ocean’s surface since 1854, primarily from merchant ships, into a compact and easily used data set. As background, a historical overview is given of how archiving of these marine data has evolved from 1854, when systematic recording of shipboard meteorological and oceanographic observations was first established as an international activity. Input data sets used for COADS are described, as well as the processing steps used to pack input data into compact binary formats and to apply quality controls for identification of suspect weather elements and duplicate marine reports. Seventy-million unique marine reports for 1854–1979 were output from initial processing. Further processing is described, which created statistical summaries for each month of each year of the period, using 2° latitude × 2° longitude boxes. Monthly summary products are available giving 14 statistics (such as the median and the mean) for each of eight observed variables (air and sea-surface temperatures, scalar and vector wind, pressure, humidity, and cloudiness), plus 11 derived variables. Examples of known temporal, spatial, and methodological inhomogeneities in marine data, and plans for periodic updates to COADS, including an update through 1986 scheduled for completion by early 1988, are presented.

1. Introduction

The world ocean covers over 70 percent of the earth’s surface. The history and future of global climate therefore cannot be understood without ocean weather observations. Sailors were among the first to systematically record the weather because the states of ocean and atmosphere controlled their progress and survival (Quayle, 1977). These observations were eventually collected in different seafaring countries and published as climatological summaries and atlases. Later the observations were used in constructing synoptic charts to support weather forecasting. Such static descriptions of ocean climate and weather undoubtedly retain great value. However, growing concerns with the effects of climatic anomalies have brought the realization that climate itself is dynamic and that observations from the oceans require the same detailed study that has been devoted to observations from stations on land for over a century.

These considerations formed the starting point for the construction of a Comprehensive Ocean-Atmosphere Data Set (COADS). Development of COADS is a continuing cooperative effort between the National Oceanic and Atmospheric Administration (NOAA)—its Environmental Research Laboratories, National Climatic Data Center (NCDC), and Cooperative Institute for Research in Environmental Sciences (CIRES)—and the National Science Foundation’s National Center for Atmospheric Research (NCAR). Initiated in 1981 (Fletcher et al., 1983; Woodruff, 1985), the COADS project has used modern formats to minimize storage volume and other innovations in data organization to provide for the first-time convenient and efficient access to a unique record of ocean-atmosphere behavior, beginning in 1854 and continuing into the future.

In this paper we review the history of ocean-weather observations, the data sources of COADS, and the methods used to create a continuing and readily accessible archive of global-climate information. Individual "marine reports" form the basis for this archive, each combining quality-controlled observations with the date, position, and other elements giving information about the report’s source. Acceptable observations of important climate variables are used to construct monthly statistical summaries for each year, using 2° latitude × 2° longitude boxes. Detailed and extensive documentation (Slutz et al., 1985; hereafter Release 1) describes data and software products that can be obtained from NCAR and NCDC in a first release of COADS, covering the period 1854–1979. A second data release, updating selected products through 1986, should become available by early 1988.

2. Historical background

For hundreds of years, mariners have narrated and later codified observations and measurements of weather at sea into ships’ logbooks (an example is shown in Fig. 1). Table 1 summarizes...
COADS Workshop
Jan 1986 meeting

NOAA Technical Memorandum ERL ESG-23

PROCEEDINGS OF A COADS WORKSHOP
BOULDER, COLORADO, JANUARY 22-24, 1986

Scott D. Woodruff, Editor

218 pages

Climate Research Program
Environmental Sciences Group
Boulder, Colorado
July 1986

UNITED STATES
DEPARTMENT OF COMMERCE
Malcolm Baldrige,
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION
Anthony J. Calio,
Administrator

Environmental Research
Laboratories
Vernon E. Derr,
Director
Contents

Acknowledgements ........................................... iii
Abstract .................................................. 1
Introduction ............................................... 1
Workshop Summary (Diaz) .................................. 3

Part 1. COADS Overview
Available COADS data (Jenne/Woodruff) .............. 7
Distribution of COADS 1854-1979 (Woodruff/Joseph/Slutz) ... 15

Part 2. Marine Data Processing Overview
Marine data processing procedures at the National
Climatic Data Center (Cram) ............................ 25
COADS 1980-85 update (Woodruff/Lubker) ............. 36
Observations from fishing fleets of the Inter-American
Tropical Tuna Commission (Sadler) ..................... 49

Part 3. Scientific Papers
Data evaluation
Some gross errors and biases in COADS and their
probable causes (Sadler/Lander/Hori) .................. 55
A data gridded set for the global oceans,
1854-1979 (Wright/Jones) ............................... 62
Problems in using the COADS data for investigating
secular climate change (Jones) .......................... 66
A comparative analysis of marine weather records from
Ocean Weather Stations and merchant-ships (Ramage) 72
How reliable is the COADS sea level pressure data set?
(Jones) .................................................. 78
Data comparison of pressure fields from COADS and Fleet
Numerical Weather Central (Mitchell/Deser/Wallace) . 84
Observed and geostrophic ship winds (Rado/Steiner) ... 87

Analysis and applications
Analysis strategies and selected real time data products
(Jenne) .................................................. 100
High-latitude coverage and applications of COADS
(Walsh) .................................................. 110
Analysis strategy for the COADS and some comparison of
products (Sadler) ....................................... 116
Surface wind stress from monthly mean winds
(Sadler/Lander/Maliekal/Hori) .......................... 121
COADS: a climatological atlas and other uses
at NCAR (Shea) ......................................... 127
Climatology of Ocean Weather Stations
(Diaz/Ramage/Woodruff) ................................ 132
Fourier analysis of SST climatologies (Levitus) ....... 146
Interannual variability of surface marine fields
(Oort/Pan) ............................................. 149
Recent large scale variability in the North Atlantic
SST field (Cayan) ....................................... 156
Regional scale circulations produced by the orography of Central America (Sadler/Lander) . . . . . .169
An "anomalous" long-term mean surface wind pressure relationship in the subtropical ridge and its probable cause (Lander) . . . . . . . . . . . . . . . . . .172
On the contribution of the sea surface temperature pattern to regional differences in the marine boundary layer shear (Lander) . . . . . . . . . . . . . . . . . .176

Part 4. Working Group Reports
Scientific (Ramage) . . . . . . . . . . . . . . . . . . . . . . . .181
Technical (Jenne) . . . . . . . . . . . . . . . . . . . . . . . .183

Appendices
A. Participants . . . . . . . . . . . . . . . . . . . . . . . . .193
B. Navy analyses (Jenne) . . . . . . . . . . . . . . . . . . . .195
C. Marine data inventories . . . . . . . . . . . . . . . . . . . .202
Proceedings of a COADS Workshop
Boulder, Colorado
January 22-24, 1986

Abstract

The proceedings of a workshop held January 22-24, 1986, in Boulder, Colorado, are presented. The workshop was organized to discuss the Comprehensive Ocean-Atmosphere Data Set (COADS), the most complete set of global surface-marine data available for the period 1854-1979. A compilation of scientific papers and technical material presented or contributed to the workshop is provided, together with reports from its Scientific and Technical Working Groups, and a list of participants. Scientific papers are grouped according to Data evaluation, and Analysis and applications.

Introduction

During January 22-24, 1986, the Climate Research Program of NOAA/ERL (National Oceanic and Atmospheric Administration/Environmental Research Laboratories) hosted a workshop in Boulder, Colorado. The workshop was held at the suggestion of the Equatorial Pacific Ocean Climate Studies (EPOCS) Advisory Committee, to stimulate use of the recently completed Comprehensive Ocean-Atmosphere Data Set (COADS) and to discuss needs for future updates and modifications of these data. Other goals of the workshop were to: 1) acquaint users with data processing procedures; 2) discuss work done to date on COADS data evaluation; and 3) present the results of scientific research based on COADS in order to get a feel for the number of applications already in progress, as well as to share the results of such work.

COADS is the most complete set of global surface-marine data now available for the period 1854-1979. Individual ship observations are available from the National Center for Atmospheric Research (NCAR) or NOAA's National Climatic Data Center (NCDC), and monthly summaries for $2^\circ$ latitude x $2^\circ$ longitude boxes and other products are available from NCAR. Parts 1 and 2 (COADS Overview and Marine Data Processing Overview) give further background on COADS, including an update to extend the period of record through 1985 planned for completion by 1987.
Available COADS Data

Roy L. Jenne
National Center for Atmospheric Research
Boulder, CO 80307

Scott D. Woodruff
NOAA/Environmental Research Laboratories
Boulder, CO 80303

1. Introduction

For several years starting in 1982, NCAR, ERL, CIRES (Cooperative Institute for Research in Environmental Sciences), and NOAA's National Climatic Data Center (NCDC) have worked together on a cooperative project to clean up several existing large files of world ship data and to merge files into a consolidated data set with duplicates eliminated. Release 1 of this Comprehensive Ocean-Atmosphere Data Set (COADS) covers 1854-1979 (Fletcher et al., 1983; Woodruff, 1985; Slutz et al., 1985). Release 2, updating COADS through 1985, is planned for availability by 1987 (Woodruff and Lubker, 1986).

The major component files were the "Atlas" data set from NCDC that was used in the construction of marine atlases (e.g., U.S. Navy, 1977) and the Historical Sea Surface Temperature (HSST) Data Project files for about 1861-1960 that were prepared under the auspices of the International Decade of Ocean Exploration (IDOE) and involving several countries (especially Federal Republic of Germany, Netherlands, and the United States). In addition, all available ship reports from 1961-79 exchanged under WMO Resolution 35 were included. About 9 other smaller files were also merged in. This included buoys and sea surface temperature from XBT reports.

About 100 million ship reports were processed, resulting in 72 million after duplicate elimination. There were 53.19 million reports output for 1854-1969 and 18.68 million for 1970-79. Monthly summaries of acceptable observations within each 2° latitude x 2° longitude box give 14 statistics for each of 19 observed and derived variables. The processing steps to clean-up the data, run sort/merges, and calculate statistics took many hours on the CRAY computers at NCAR. Additional time was spent on NOAA computers.

The volume of data is often rather high for the whole world ocean, but many products are organized by 10° boxes so that part of an ocean basin can be studied without volume problems. Figures 1 and 2 are illustrations of data coverage in time and space. There are additional products such as decade-month summaries and report inventories; for more information, the reader should refer to the overall COADS text (Slutz et al., 1985). Fortran 77 software is available to read all of the packed binary products (Woodruff et al., 1986).
Figure 1. Global percentage of possible (sea or coastal) year-month-2 degree boxes per year containing at least the indicated number of observations of sea surface temperature, subject to trimming (solid curves). The dashed curve is for the equivalent area covered by 2 degree boxes containing at least 1 observation. Horizontal lines represent an approximate upper-limit on the number of boxes (solid) or area (dashed); they are positioned at 100% minus the percentage of boxes or area poleward of 60 degrees N or S. Letters are ordinates of the 126-year-average percentage of 2 degree boxes containing at least 1 observation of variables: S = sea surface temperature, A = air temperature, W = wind, P = sea level pressure, C = total cloudiness, Q = specific humidity.
The following is a summary of the major topics discussed by the technical working group. Some additional notes that are referred to follow sec. 11.

1. Problems in COADS Release 1 (1854-1979)

   The data set is much cleaner, much more comprehensive, and easier to use than any previous set of ship data. It still has some problems that have not been explained (see also Notes A and B). Examples were shown (Sadler et al., 1986) of a few year-month-20° boxes with bad values in the untrimmed monthly summaries. The original ship data would need to be examined to try and see what the problem is, and the trimmed summaries should be checked to make sure that these bad values were trimmed.

   Large amounts of ships mislocated over land were also noted during 1975. A test tape of keyed merchant data (deck 927) containing erroneous longitudes for February-July 1975 was inadvertently included in COADS (Cram, 1986). The problem has been corrected by NCDC in their copy of the 1970-79 TD-1129. NCDC is supplying data so that similar corrections can be made in Boulder when time permits.

   Monterey Telecommunication data (deck 555) were the only available real-time source for about 1966-73, containing roughly 4 million reports. Comparisons that NCDC will provide show temperatures (except SST) frequently as much as 1.5°C too high. These errors may have biased the trimmed summaries (deck 555 was omitted from the untrimmed summaries and thus could not contaminate the limits used for trimming). More evidence must be collected before any assessment can be made of the severity of this problem.

2. Errors in the 1980-85 Update Data

   Some errors have been discovered by NCDC, primarily in NOAA/National Meteorological Center (NMC) data. Examples are a double conversion of wind speed, improperly decoded data that ended up in swell fields, and switched dew point and wet bulb. All of these errors can be fixed readily, or are confined to limited time periods and will be documented (Note C).

   One significant error that cannot be fixed readily is the lack of the weather indicator, iX, during 1982-83 in data from NMC. This indicator, a part of the 1982 WMO code change, is used to flag the omission of present and past weather fields when no significant (i.e., bad) weather was reported. Without iX, there is no way to distinguish between missing and good weather (e.g., present weather codes 00-03) after 1982. Unfortunately, there are problems with iX in other data sets between 1982-84, and the
potential exists throughout that time period for serious biases toward good or bad weather (depending on whether missing weather is interpreted as good or bad) in statistical summaries. More study is required to determine how frequently this condition exists and what can be done about it.

Foreign fixed buoys (probably only one or two) are missing prior to July 1985 (drifting buoys are included). This omission could be recovered from the basic NMC tapes when time permits, as with 1x for 1982-83 (see Cram, 1986).

3. Hourly Data From Fixed Buoys

Hourly data from about 33 buoys have more observations than the U.S. merchant marine (currently, 20 moored U.S. buoys are "offshore," plus 29 within 150 km of shore). It was suggested that only 3-hourly data from buoys should be retained in COADS, with hourly data in a side set. NCDC prefers to keep all observations in COADS (except sub-hourly resolution). In any event, future computation of monthly summaries should exclude off-3-hourly data to increase the influence of passing ship observations on statistics for 2° boxes containing fixed buoys.

Monthly counts of global data received at NCDC for 1980-84 average approximately as follows (see also Appendix C):

<table>
<thead>
<tr>
<th>source</th>
<th>reports/month</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. merchant marine (ready within 3 months)</td>
<td>23,590</td>
</tr>
<tr>
<td>International exchange</td>
<td>60,000</td>
</tr>
<tr>
<td>NMC (telecommunications data without U.S. fixed buoys)</td>
<td>90,306</td>
</tr>
<tr>
<td>U.S. fixed buoy data (hourly)</td>
<td>24,797</td>
</tr>
<tr>
<td>NODC surface-level XBTs (not incorporated after 1977)</td>
<td>n/a</td>
</tr>
</tbody>
</table>

(For U.S. observations only, it is estimated that 59% of the NMC data are not duplicated in delayed manuscript sources; conversely, about 72% manuscripts are unique.)

4. Boulder/NCDC Agreement on COADS Update

A batch of data for about 1980-85 will be prepared by continued cooperation between the groups. NCDC will always prepare annual updates. A major cooperative update will be done about each five years (see Note D and Woodruff and Lubker, 1986).

5. Deficiencies in Global International Exchange

WMO's data flow plan is intended to gather ship reports into eight regional centers operated by the following responsible countries: Federal Republic of Germany, Hong Kong, India, Japan, Netherlands, U.K., U.S., and USSR. NCDC tries to obtain a global set by negotiating with countries and regional centers, but where
Appendix B
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Roy Jenne
9 Oct. 1985
Rev. 9 Jan 86

NAVY ANALYSES

A. General Comments

— Navy UA (upper air) analyses on hemispheric grids never included wind analyses before 1983. The new UA 2.5° global archives starting Jan 1983 include winds.

— Navy surface archives on hemispheric grids include winds, but these were derived from the pressure field and stability prior to 1974. Real winds were not used before that time. The derived winds appeared to be reasonably good. They were used to make wave forecasts and for other purposes.

— The global band surface and UA analyses (40S-60N on 2.5° grids) use reported wind data to make wind analyses. The global forecast model (1983-on) does not use any global band data.

— The new 2.5° global archives starting Jan 1983 include wind analyses, which also directly use observed wind data in the upper air.

— The initial state for the forecast model does not use wind grids at the surface that include the direct use of real winds.

— A user could prepare global Navy analyses of winds from 1974-on by using global band winds (40S-60N), and blending these with geostrophic winds for the polar areas.

B. Summary of available Navy surface and UA analyses

1. N. Hemisphere surface analyses (63x63 grid points) 1961 to 1985, and later

   Most grids start about 1961, some before. SLP (from Nov 45), SST (Nov 61), T air (May 65), E air (May 65), N clouds (Jan 68), winds (calc from pressure fields from 1945 until Aug. 1974, then analyzed).

2. N. Hemisphere UA analyses (63x63) 1961 to Jan 1983

   1000 to 100 mb, mostly start about 1961-63. See data lists in NCAR TN/IA-111. (Data Sets for Meteorological Research, 1975, by R. Jenne).

3. S. Hemisphere surface analyses (63x63 grids) July 1973 to 1985 and on

   SST starts 1 Jul 73. Also SLP. By 1983 the data included SST, SLP, Air T, surface vapor pressure. Winds start Dec 1978.

4. S. Hemisphere UA (63x63) Aug 1974 to Jan 1983

   1000, 925, 850...-100 mb, start Aug 1974.
Appendix C

Marine Data Inventories

Table 1 was compiled from an inventory that NCDC completed by deck, month and year for the corrected '70s COADS and '80s period-of-record tapes (as of 13 February 1986).

Tables 2-5 were compiled from inventories (product 2) described in Release 1 (Slutz et al., 1985). Table 2 summarizes duplicate elimination input (I), output (O), and uncertain duplicates (D) for three periods (1854-1969, 1970-79, 1854-1979). The number of Long Marine Reports (LMR) is given for a few general categories and separately for each card deck*. Table 3 gives related percentages with respect to the complete, global LMR file of a particular type (I, O, D) and time period. Table 4 is a similar breakdown for ((I - O) + D)/I and D/O. ((I - O) + D)/I is the fraction of input that was identified as duplicate (certain or uncertain). D/O is the fraction of output that was identified as uncertain. Each fraction is expressed as a percentage with respect to the input or output for a specific category (GTS, etc.) and time period, only. Thus ((I - O) + D)/I in each case is really ((I - O) + D)/I x 100. Table 5 lists thousands of LMR per source ID (see Release 1's Table F1-2, reprinted within Table 2) separately for each year 1854-1979.

* Global telecommunication system (GTS) data were identified by card deck (see Release 1's Table F1-1, reprinted within Table 2): 555, 666, 849, 850, 888, 889, 999. Non-GTS data comprise all other card decks, as well as identifiable data from the remaining categories: buoy decks 143, 876-882; NODC 891; IMM 926; U.S. IMM 128, 927-928; foreign 118-119, 184-185, 187-189, 192-194, 196-197, 898-900, 902; U.S. 110, 116-117, 195, 281; Historical Sea Surface Temperature (HSST) Data Project 150-156; and other 186, 897, 901. It should be noted that U.S. International Maritime Meteorological (IMM) exchange data (deck 927) include some foreign IMM prior to 1980. Therefore, counts for IMM and U.S. IMM are inaccurate.
Proceedings of the International COADS Winds Workshop, Kiel, Germany, 31 May - 2 June 1994

Editors:
Henry F. Diaz,
Hans-Jörg Isemer
May 1995
SHIP WIND SPEEDS
(knots)

Based on selected 10 degree regions:

Marsden Squares 79, 80, 122, 123, 141, 142, 184, 185, 199, 200, 217 & 252

U.S. conversion error

USSR

Excluding USSR

Figure 1: GTS ship wind speeds averaged for selected 10 Marsden Squares in the North Atlantic North Pacific, and Mediterranean. 79, 80, 122, 123, 141, 142, 184, 185, 199, 200, 217, and 252. Curves shown for USSR and all other data are displaced possibly due to biases from reporting wind in meters per second versus knots (see text). The effect of a U.S. conversion software error is also strongly evident during February-June 1984.
Figure 2: Annual global marine reports after duplicate elimination (curve) for COADS Release 1 through 1979, continued by Release 1a through 1992. Horizontal lines span the time periods for data now being collected and digitized, or proposed for future digitization (*), with the approximate numbers of reports shown in millions (M) or thousands (K) (Elms et al., 1993). Also listed are major existing digital data inputs proposed for inclusion in Release 2 or following Release 2. Labeled ticks along the upper horizontal axis mark the starting years for Release 1a, and those planned for Release 1b (1947) and Release 2 (1854, or earlier).
Figure 5: Air temperature at New Haven, CT with longitudinal average air temperature anomaly from the latitude band of the mic-Atlantic region (from Ingham, 1982, and Hansen and Lebedeff, 1987).
Ship Data

In the following, we give ship counts in NMC data, and in COADS. We also analyze how the data counts for a given year improve as the delayed data is received within 1, 2 or 5 years.

We decided not to make counts of the Navy ship data because the high duplication rate of about 20 percent found by NOAA/ERL (Woodruff) would make it impossible to use the counts. The essential facts about ship data are as follows. A table is available in a 24 Feb. 89 memo by Scott Woodruff. It has many of the following counts. The counts are based on COADS experience for data observed in the 1980s.

- Facts about marine ship synop data, not buoys, etc.:

NMC Ships

-- 1200K obs/year, available in real time
-- After five years, 75% of these reports are still unique, and 75% are retained
-- 900K obs/year remain after five years
-- In COADS a delayed logbook ship report is usually saved instead of a GTS report, if they are duplicates.

NAVY DECASSIFIED SHIPS

-- 60K per year
-- These have been declassified within one year delay
-- 95% of these are unique
-- nearly 60K obs/year remain after five years

US DELAYED MARINE LOGS

-- 320K obs/year, all available within 1 year
-- After five years, 90% of these are retained in COADS
-- After five years, 70% of these are new data, added to GTS
-- 290K obs/year remain after five years

IMM DELAYED FOREIGN MARINE

-- 750K per year, available within one year
-- 960K per year, available within two years
-- 1100K per year, available within five years
-- After five years, 90% of these are retained
-- After five years, 50% of these are unique data, added to GTS

SUMMARY FOR COADS (How the archive improves with time)
-- These summary counts are only for ships. (Buoys are also in COADS)
-- Total unique ship reports in real time is about 1200K/year
-- Total received by one year delay is 1130K additional reports.
   About 656K of these are unique. So, total unique ship reports
   is about 1856K/year.
-- After two years delay, total unique ship reports is about 2100K/year
-- After five years delay, total unique ship reports is about 2220K/year

LONG DELAY SHIP DATA

In Jan - March 1989, NCAR received about 5,000,000 USSR ship obs for the
N. Atlantic, mostly for 1957-87 (about 165,000 per year). We hope to obtain
USSR data for other regions also. These delayed USSR data are not reflected in the
above ship counts.

SHIP DATA FOR WWI AND WWII

The COADS files have very few observations during the two world
wars. Data on US ship logs would fill in this gap. These could be key-
entered for about $1M. There would be about 15 million observations. For
comparison, COADS for 1854-1979 now has 72 million observations.
Counts of Navy (FNOC) and NMC Ship and Buoy Reports

NOAA/CIRES (Boulder, CO), has compared Navy and NMC surface marine data for two months. A report will be available. Info is from Scott Woodruff. A summary follows:

Table 1: Surface Marine reports from Navy and NMC, for two months.

The dates are December 1986 and June 1988. Duplicate elimination counts (interim dupelim; all hours): "Input" is the reports per month on the tape. "Output" is the same after internal duplicate elimination.

The column "unique" shows unique reports for each center. The other reports are common to both centers. The column "Input 3" here is the same as Output, except for off-hour reports (ships and moored buoys reduced to 3-hourly). The input (Input 3) to a cross duplicate check and the number unique to each dataset are shown below:

<table>
<thead>
<tr>
<th></th>
<th>December 1986</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FNOC</td>
<td>NMC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input</td>
<td>Output</td>
<td>Input 3</td>
<td>Unique</td>
</tr>
<tr>
<td>Ships</td>
<td>120,579</td>
<td>99,081</td>
<td>97,108</td>
<td>10,049</td>
</tr>
<tr>
<td>Moored buoys</td>
<td>31,404</td>
<td>30,250</td>
<td>11,267</td>
<td>638</td>
</tr>
<tr>
<td>Drifting buoys</td>
<td>57,502</td>
<td>53,188</td>
<td>53,188</td>
<td>19,010*</td>
</tr>
<tr>
<td>NMC</td>
<td>Input</td>
<td>Output</td>
<td>Input 3</td>
<td>Unique</td>
</tr>
<tr>
<td>Ships</td>
<td>96,532</td>
<td>92,923</td>
<td>92,006</td>
<td>4,947</td>
</tr>
<tr>
<td>Moored buoys</td>
<td>12,915</td>
<td>12,373</td>
<td>11,680</td>
<td>1,051</td>
</tr>
<tr>
<td>Drifting buoys</td>
<td>56,631</td>
<td>48,801</td>
<td>48,801</td>
<td>14,623*</td>
</tr>
<tr>
<td></td>
<td>Input</td>
<td>Output</td>
<td>Input 3</td>
<td>Unique</td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>Ships</strong></td>
<td>159,801</td>
<td>122,869</td>
<td>120,875</td>
<td>13,857</td>
</tr>
<tr>
<td><strong>Moored buoys</strong></td>
<td>45,722</td>
<td>41,865</td>
<td>17,328</td>
<td>1,342</td>
</tr>
<tr>
<td><strong>Drifting buoys</strong></td>
<td>26,020</td>
<td>24,750</td>
<td>24,750</td>
<td>11,573*</td>
</tr>
<tr>
<td><strong>NMC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ships</strong></td>
<td>114,545</td>
<td>110,117</td>
<td>109,281</td>
<td>2,263</td>
</tr>
<tr>
<td><strong>Moored buoys</strong></td>
<td>19,833</td>
<td>17,699</td>
<td>16,954</td>
<td>968</td>
</tr>
<tr>
<td><strong>Drifting buoys</strong></td>
<td>26,430</td>
<td>23,182</td>
<td>23,182</td>
<td>10,005*</td>
</tr>
</tbody>
</table>

*Even the estimates of uniqueness for ships and moored buoys may still be slightly distorted by errors in identifying duplicates. Spot checks have shown the numbers of unique drifting buoy reports to be significantly inflated due to the presence of apparent duplicates with small variations in position. Ships appear to be 3% high, moored buoys 1% high, and drifting buoys 17% high.

**Table 2. All unique reports from FNOC plus NMC**

After internal duplicate elimination and reduction of the ship and moored buoy data to 3-hourly (i.e., 00, 03, ..., 21 UTC), following are the total unique reports obtained by combining the two datasets using the same cross-duplicate check:

<table>
<thead>
<tr>
<th></th>
<th>December 1986</th>
<th>June 1988</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ships (3-h)</strong></td>
<td>102,055</td>
<td>123,138</td>
</tr>
<tr>
<td><strong>Moored buoys (3-h)</strong></td>
<td>12,318</td>
<td>18,296</td>
</tr>
<tr>
<td><strong>Drifting buoys (all-h)</strong></td>
<td>67,811*</td>
<td>34,755*</td>
</tr>
</tbody>
</table>

The above total unique reports is the sum of the number of reports shared between the two datasets (not shown), plus the number unique to each dataset.

*See the * in Table 1.
COADS Update Notes

8 February 1989
S. Woodruff

1. Drifting buoy QC
MEDS (Marine Environmental Data Service, Dept. Fisheries and Oceans, Canada) is amenable to QC the global drifting buoy archive. We will provide the data for 1980-85 (1986-date at MEDS); format (undecided) may impact timing.

2. Extension of trimming limits
Drifting buoys traverse regions with insufficient data to define 1950-79 trimming limits. Limits need to be interpolated over these regions, or the data will be automatically rejected from statistics.

3. Tuna boats
Inter-American Tropical Tuna Commission may agree to release individual observations (SST, wind) from fishing boats, under conditions including anonymous integration into monthly summaries.

4. Duplicate elimination
Modifications are needed to adapt the method used for 1970-79 data for 1980s data to avoid erroneous elimination of a small number of reports. The method also failed to eliminate some (including wildly mislocated) duplicate reports; not a tractable problem at this time.

5. Format conversions and error corrections
NCAR is progressing steadily, but there are many formats, and the error corrections are not fully defined. Testing reveals too many rejections—we will work with NCDC to solve problems.

6. Platform blending
FNOC/NMC data comparisons (ERL report planned) show large differences between selected platform types (Table 1).

7. Incorporation of UK corrections
We may wish to consider statistics for nighttime/daytime air temperature. Adjustment of existing COADS 2° SST (1854-1941) could be made by obtaining adjustment climatologies and method from UK Met. Office in digital form (regridding needed), or by calculating new adjustment data.

8. Schedule considerations
Interim 1980-87 products are available at NCAR, but some users (e.g., Oort) need statistics not included. NCDC is hoping for its TD-1129 product in 1989. However, NCAR computer replacement will slow at least final sorting and statistics production until June-July. Considering the NCDC data preparation cycle, we plan to include 1988 data (available June-July): NMC data (maybe part of 1989), US-punched data, and any international exchange data that can be gathered. Unforeseen problems, for example in data from drifting buoys or buoys-on-ice, could impact schedule.

9. Interest in new products
- Klaus Weickmann and others (CRD): 2° daily/pentad summaries (Monterey).
- WMO ship history tapes provide instrumentation, routes, etc. (since 1956).

cc:
Maurice Blackmon
Roy Jenne

Table 1. Preliminary comparisons of NMC platforms, December 1986.
Means were computed for 2° boxes (COADS 1854-1979 method; sample plot attached), separately for each platform type: ships, moored buoys, drifting buoys, and the Coastal Marine-Automated Network (C-MAN). Differences were calculated between corresponding 2° boxes (each with at least 1 observation) for each comparison. The following "global" differences are the mean of all such 2° box differences.

<table>
<thead>
<tr>
<th>Global difference</th>
<th>SST</th>
<th>AT</th>
<th>SLP</th>
<th>scalar wind (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ship minus moored buoy</td>
<td>.07</td>
<td>.26</td>
<td>-.06</td>
<td>1.13</td>
</tr>
<tr>
<td>ship minus drifting buoy</td>
<td>.05</td>
<td>.28</td>
<td>-.18</td>
<td>3.92*</td>
</tr>
<tr>
<td>ship minus C-MAN</td>
<td>2.01</td>
<td>2.33</td>
<td>-.47</td>
<td>1.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentage of boxes** in global difference</th>
<th>SST</th>
<th>AT</th>
<th>SLP</th>
<th>scalar wind</th>
</tr>
</thead>
<tbody>
<tr>
<td>ship minus moored buoy</td>
<td>0.5%</td>
<td>0.6%</td>
<td>0.5%</td>
<td>0.5%</td>
</tr>
<tr>
<td>ship minus drifting buoy</td>
<td>2.7%</td>
<td>2.8%</td>
<td>3.4%</td>
<td>0.3%</td>
</tr>
<tr>
<td>ship minus C-MAN</td>
<td>0.1%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

* Although some drifting buoys (e.g., French) may report wind speed and direction, data from other platform types may be inadvertently included.
** Percentages with respect to the total number (11,827) of ocean/coastal 2° boxes, as defined by an approximate file (LlN2F1) with 10,946 ocean, 4,375 land, and 881 coastal boxes.

Maurice:

I talked Table 1 over with Joe Fletcher. Since you will be meeting with concerned individuals (e.g., Don Hansen, Stan Hayes, and Ed Harrison), he recommended that Table 1 be transmitted to you for possible discussion. (Caution: these are very preliminary results, for one month.) For COADS, one question has been at what frequency (e.g., 3-hourly) to blend data from different platform types (including "surface" XBT data; Reverdin et al., November 1988 TO-AN). Table 1, if representative, also poses the question of whether blending data from a diversity of platforms (without adjustment for anemometer height, type, and drift, averaging period, etc.) will introduce unacceptable biases, although exclusion of drifting buoy data, for example, would reduce data coverage.
December 1986
Zonal means: NMC ships minus NMC drifting buoys

- mean difference
- = s.d. of differences
= no. of boxes used
Ship Data Counts, Next Two Pages

1st page: Ship reports from countries, 1968 – 1985

I think that our COADS ship format remembers the country of origin. We should generate a list like this by country, from COADS.

2nd page: # of unpunched ship observations, NCDC, 1984

Bruce Blankenship (NCDC, Asheville) prepared this list of ship observations not yet key-entered. He estimated 17.7 m observations. Bruce was good to talk with. It seemed to me that he made this estimate in a sensible way.

Things went bad somehow; Bruce took his own life. Later, NCDC could only locate about 2.7 m observations. The question is whether some more observations either did exist or even still exist.

I talked with Hal Crutcher who had been a scientist and leader at NCDC for many years. He said that Bruce was a reliable and careful person. It seems to me that Bruce really did count the number of boxes with ship forms. Beyond that, we do not know what happened.

Roy Jenne
Sep 2003
<table>
<thead>
<tr>
<th>RESPONSIBLE MEMBER</th>
<th>68</th>
<th>69</th>
<th>70</th>
<th>71</th>
<th>72</th>
<th>73</th>
<th>74</th>
<th>75</th>
<th>76</th>
<th>77</th>
<th>78</th>
<th>79</th>
<th>80</th>
<th>81</th>
<th>82</th>
<th>83</th>
<th>84</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>551</td>
<td>1510</td>
<td>1841</td>
<td>2156</td>
<td>2453</td>
<td>2717</td>
<td>3000</td>
<td>3269</td>
<td>3532</td>
<td>3797</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>1792</td>
<td>1510</td>
<td>1841</td>
<td>2156</td>
<td>2453</td>
<td>2717</td>
<td>3000</td>
<td>3269</td>
<td>3532</td>
<td>3797</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td>3908</td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>788</td>
<td>3720</td>
<td>7084</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td>9185</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>100</td>
<td>3097</td>
<td>19366</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td>24032</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>100</td>
<td>3097</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td>1201</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>581</td>
<td>3252</td>
<td>20938</td>
<td>33283</td>
<td>37714</td>
<td>45295</td>
<td>5032</td>
<td>5510</td>
<td>60268</td>
<td>66428</td>
<td>69166</td>
<td>72513</td>
<td>7622</td>
<td>7944</td>
<td>8212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany, Federal</td>
<td>12164</td>
<td>53168</td>
<td>10345</td>
<td>85635</td>
<td>55287</td>
<td>10</td>
<td>44</td>
<td>137</td>
<td>2656</td>
<td>14194</td>
<td>58311</td>
<td>10539</td>
<td>13006</td>
<td>14194</td>
<td>14194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Republic</td>
<td>2793</td>
<td>1537</td>
<td>1761</td>
<td>8981</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>12164</td>
<td>53168</td>
<td>10345</td>
<td>85635</td>
<td>55287</td>
<td>10</td>
<td>44</td>
<td>137</td>
<td>2656</td>
<td>14194</td>
<td>58311</td>
<td>10539</td>
<td>13006</td>
<td>14194</td>
<td>14194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>851</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Iceland</td>
<td>6337</td>
<td>13390</td>
<td>16153</td>
<td>14242</td>
<td>187</td>
<td>25857</td>
<td>14153</td>
<td>33246</td>
<td>68520</td>
<td>18816</td>
<td>14863</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Israel</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jamaica</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jordan</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kenya</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanzania, United</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Republic of</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Kingdom of</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Great Britain and</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern Ireland</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States of</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>America</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yugoslavia</td>
<td>4214</td>
<td>714</td>
<td>1085</td>
<td>1085</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The US kept about 30,000 reports per month in 1980-82 etc.

1981 about 35,000 in year
Unpunched US Ship OBS

<table>
<thead>
<tr>
<th>Years</th>
<th>Merchant</th>
<th>Other</th>
<th>Ships Obs per Form</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1913-1919</td>
<td>237K</td>
<td></td>
<td>v19</td>
<td>4,503K</td>
</tr>
<tr>
<td>1937-39</td>
<td></td>
<td>101K</td>
<td></td>
<td>1919</td>
</tr>
<tr>
<td>1940</td>
<td></td>
<td>36</td>
<td></td>
<td>684</td>
</tr>
<tr>
<td>41</td>
<td></td>
<td>33</td>
<td></td>
<td>627</td>
</tr>
<tr>
<td>42</td>
<td></td>
<td>6</td>
<td></td>
<td>114</td>
</tr>
<tr>
<td>43-45</td>
<td></td>
<td>3</td>
<td></td>
<td>57</td>
</tr>
<tr>
<td>46</td>
<td></td>
<td>6</td>
<td></td>
<td>114</td>
</tr>
<tr>
<td>47</td>
<td></td>
<td>27</td>
<td></td>
<td>513</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td>33</td>
<td></td>
<td>627</td>
</tr>
<tr>
<td>Mid War II</td>
<td></td>
<td>450K</td>
<td></td>
<td>8550K</td>
</tr>
<tr>
<td></td>
<td>482K</td>
<td>450K</td>
<td></td>
<td>9158K</td>
</tr>
<tr>
<td></td>
<td>932K</td>
<td></td>
<td></td>
<td>17,708,000</td>
</tr>
</tbody>
</table>

Costs to microfilm the ship forms (Count of Forms)

<table>
<thead>
<tr>
<th>Years</th>
<th>Ship Forms</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1913-1919</td>
<td>237K</td>
<td>$35,550</td>
</tr>
<tr>
<td>1937-1948</td>
<td>625K</td>
<td>$104,250</td>
</tr>
<tr>
<td></td>
<td>932K</td>
<td>$139,800</td>
</tr>
</tbody>
</table>

932K forms.
- $1.15 = $139,800.
- $1.4 = 138,400.
- $1.13 = 121,160.
- $1.12 = 111,840.
- $1.11 = 102,526.
- $1.10 = 93,200.

Total obs about 9,158,000 ± 8,900,000 ± Total x 2 = 17,708,000 ships not punched.
More ship data for I-COADS
by Scott Woodruff, Sept 2003

- Briefly describes a number of ship datasets that are being added (when available) to COADS

- A graph shows ship counts from about 1820-1995

- A chart shows the status of US logbooks of ship observations, 1784-recent, for US merchant ships and US Navy ships.

(6 pages follow)

Roy Jenne
Jan. 6, 2004
Hi Roy,

On Tue, 16 Sep 2003, Scott D Woodruff wrote:

> Hi Roy,
> Sorry I missed your phone call last week, but if there are questions
> on the plots feel free to let me know.
> There is a fairly significant historical source I should have
> mentioned that does not appear on Plot 1: Arctic Norwegian Logbook data
> (1867-1912; 125K reports). A limited set of fields was keyed by Russia.
> Steve was sorting out some questions with Syd Levitus on these data,
> which may be the same as a similar or identical data set that NODC heard
> about. The data are already archived at NCAR in their native format:
> http://dss.ucar.edu/datasets/ds539.1/
> Regards, -Scott

To follow-up on our discussion of the above dataset: Below is a more stand-alone text that incorporates what is known about this new dataset. I made a few other changes to clarify the text.

Let me know there are questions or additional changes would be helpful. Cheers, -Scott

Discussion to Accompany Plots of International Comprehensive Ocean-Atmosphere Data Set (I-COADS) Blend Candidates

17 September 2003, Scott Woodruff (NOAA/CDC)

Plot 1

The first attached plot shows the time spans (and data volume, if known) of some major data additions available or becoming available (colored horizontal lines), but not yet converted into the common format or blended (both

9/17/03
of which can be major tasks requiring adequate resources).

These lines are placed above the curve for the current 1-COADS Release (2.0) versus the original COADS Release 1 (shown continued by "interim" data--dashed curve).

Note: The emphasis of this plot was on historical blend candidates. There are a number of additions available for recent years not shown (e.g., Canadian and other buoy data).

Although this plot includes most of the major historical candidates, more work will be required to develop a more complete plot. The following are examples of two potential historical additions not yet included on the plot:

- Arctic Norwegian Logbook Data (1867-1912; 125K reports). This consists of a limited set of fields (air temperature and SST were the only meteorological elements) keyed by Russia and received from the Norwegian Polar Institute in 1999. NODC may have obtained an identical or similar version of these data (not yet resolved).
- US Lightship data back to about 1891. These Coast Guard data are being imaged, and digitized for the North Atlantic, by NOAA/CDMP.

Here is some brief background (from the top down) on the blend candidates that are currently plotted:

a) Russian R/V Marine: This line should be changed to yellow, since the data are being digitized (NCAR is receiving data periodically). It now appears that 6M reports was an overestimate by Russia.

b) Polish Baltic: A fairly small addition, but illustrating the international contributions being made to 1-COADS.

c) Ukrainian Hydrometeorological: Ditto.

d) Japanese Whaling Data: This is considered important e.g. by Chris Folland since it would add pressure data in the sparse Southern Ocean (distinct from Kobe Collection).

e) UK Marine Data Bank (MDB): We blended the MDB into the time periods before and after, but couldn’t fit this in to the last update.

f) Arctic Drift Stations: Some fairly scattered records and stations keyed by NCDC several years ago. Lots of formats (complicated).

g) US Navy Hourlies: Only a very small amount of data from this deck was included in COADS Release 1—presumably it would be appropriate to include the entirety (maybe needing some checks on data quality/compatibility first).

h) Japanese Kobe: A 2003 Edition recently became available (1889-1940; 3.2M), which apparently exhausts the usable data (e.g., they decided some Navy data were unusable?),
so probably this line should be colored green to indicate that the source is fully digitized. Only the first Kobe edition (prior to 2001 and 2003 editions) is now blended.

i) US Marine Meteorological Journals: Keyed by China. A major pending addition. Very recently about 30 additional undigitized journals (about 28K reports) were discovered mis-filed at NARA. We hope to be able to digitize these under an upcoming US-China bilateral agreement.

j) German Data: A small set that was supposed to have been keyed by Germany several years ago. (However, Germany was actually keying a lot more data—the line could be longer. Also, there were informal discussions about the possibility of re-blending their complete archive to overcome some of the problems in the HSST format, etc.).

k) German Maury: The largest new source to be keyed by China under the new bilateral agreement mentioned above. Estimated reports: 500-750K. We are also bundling into this agreement the re-keying of two microfilm reels from the original US Maury Collection (30-50K reports), which were not keyed or not usable.

Plot 2

The second plot shows the status of US Logbooks (Merchant and Navy). The colors have a somewhat different meaning on this plot:

- Green: Fully digitized
- Yellow: Fully digitized, but unblended
- Red: Undigitized

The plot needs some editing and clarifications, and indicates some unresolved questions. E.g., for the Navy data, what is the relationship between the holdings at NARA (National Archives) through 1947, and early Navy decks already blended (281, 195, and part of 110)?

The green line showing that all Navy logbooks since 1963 have been digitized is partly incorrect. Keying of Navy data at NCDC was halted around the early 1980s, but recently restarted (under NOAA/CDMP?). So the line should include a red segment (years are uncertain).

The plot needs to be clarified for 1947 Merchant data (no status shown). Probably 1947 data were thrown out in 1974 by the Maritime Commission with the rest of the WW II data, but they were not part of the NCDC project to key 1912-46 so this isn't certain. The location of the 1948 Merchant logbooks, assuming they still exist, is currently unknown (NCDC's East Point archive?).

The green line showing that all Merchant logbooks since

9/17/03
1963 have been digitized is also partly incorrect. Due to the MOPS project at NCDC, keying was halted about 1994. This was also restarted fairly recently (the dates when keying was ceased and restarted at NCDC would be useful to document, for both Navy and Merchant data, if possible).

A significant unkeyed data holding shown by this chart is the Simultaneous (Greenwich Mean Noon) Obs for 1886-1911, divided between NARA and NCDC. Similarly to the Navy data, the significance of the overlap between the early GMN obs and the US Marine Met Journals Collection for 1886-94 is unknown. If the GMN obs were derived from the much more complete Journals for 1886-94, then the GMN obs for the 1895-1911 period should probably receive priority for keying (e.g., by NOAA/CDMP).

Another item from Scott Woodruff's message dated Sep 12, 2003:

Attached is a powerpoint file containing the two plots we recently discussed. Both plots are in need of some updating, but I decided to ship them up to you with the following explanatory notes, since it will be difficult to quickly iron out some questions and make new plots.

Joe Elms at NCDC
P.S.: Joe: Maybe we should think about the 1886-1911 (or 1895-1911) GMN forms discussed under Plot 2 as a potential FY2004 CDMP project, as Gil Compo has suggested?
Release 2.0 (dark blue) vs. Release 1 (light blue) reports/year
Future blend candidates: red (undigitized), yellow (partly), green (fully)
# U.S. Logbook Status

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1784-1819</td>
<td>Sparse data</td>
</tr>
<tr>
<td>1863-1894</td>
<td>US Marine Met. Journals (1.8M)</td>
</tr>
<tr>
<td>1886-1902</td>
<td>Simultaneous (GMN) Obs (NARA)</td>
</tr>
<tr>
<td>1903-1911</td>
<td>Simultaneous (GMN) Obs (NCDC)</td>
</tr>
<tr>
<td>1903-1946</td>
<td>US Merchant Marine (3.5M)</td>
</tr>
<tr>
<td>(Unknown status)</td>
<td>1948-1948</td>
</tr>
<tr>
<td>Deck 116 (7.15M)</td>
<td>1949-1963</td>
</tr>
<tr>
<td>Within deck 128/927 (WMO Res. 35)</td>
<td>1963</td>
</tr>
</tbody>
</table>

## NAVY

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1801-1947</td>
<td>US Navy Ships, Etc. (NARA; ~73K logbooks)</td>
</tr>
<tr>
<td>1920-1945</td>
<td>Deck 281 (200K)</td>
</tr>
<tr>
<td>1942-1945</td>
<td>Deck 195 (650K)</td>
</tr>
<tr>
<td>1945-1951</td>
<td>Deck 110 (750K)</td>
</tr>
<tr>
<td>1951-1963</td>
<td>Deck 117 (3M; 16K COADS)</td>
</tr>
<tr>
<td>Within deck 128/927 (WMO Res. 35)</td>
<td>1963</td>
</tr>
</tbody>
</table>
World Meteorological Organization (WMO)

Oct 2002

Workshop on Advances in the Use of Historical Marine Climate Data

By Henry Diaz, Chris Folland, Teruko Manabe, David Parker, Richard Reynolds and Scott Woodruff

Article re-printed from WMO Bulletin 51 (4), pages 377-380
October 2002
Workshop on Advances in the Use of Historical Marine Climate Data

By Henry Diaz1, Chris Pollard2, Teruko Manabe3, David Parker2, Richard Reynolds1 and Scott Woodruff1

A Workshop on Advances in the Use of Historical Marine Climate Data was held at the NOAA Climate Diagnostics Center, Boulder, Colorado, USA, 29 January-1 February 2002. It was organized by NOAA, the UK Met Office and the Japan Meteorological Agency and was sponsored by the Global Climate Observing System (GCOS) and WMO.

Scope

The overall intention was to build on the recent blend of the US Comprehensive Ocean-Atmosphere Data Set (COADS) with the Met Office Marine Data Bank and several million newly digitized data. This blend provides the climate research community with an unprecedented assembly of in situ marine data. There have been major improvements in data availability up to the mid-20th century. The new observational archive has been named the International Comprehensive Ocean-Atmosphere Data Set (ICOADS).

Proceedings

The Workshop began with presentations on historical marine datasets, sea-surface temperature (SST) and sea ice; marine air temperature, mean sea-level pressure (MSLP) and wind; and recommendations from the second CLIVAR Climate of the Twentieth Century (C20C) Workshop. Three breakout groups covered SST, air temperature and sea-ice; MSLP and wind; and technical requirements. These groups made recommendations, summarized below. General background to the recommendations includes the need to reduce the remaining biases in the data; to increase, where possible, coverage and temporal resolution; to specify uncertainty in analyses; to distinguish versions of datasets; and to promote easy access to all data. A staged timetable for implementation was agreed: firstly, a two-year period would lead to the third C20C Workshop around April 2004; and, secondly, a period of about five years would lead to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC).

Recommendations

SST, air temperature and sea-ice

Within two years:

• Re-examine the historic bias corrections to SST, especially for the late 1930s until the end of the 1940s;
• All the metadata in the issues of publication WMO No. 47 (International List of Selected, Supplementary and Auxiliary Ships) should be digitized; biases in recent night marine air temperature (NMAT) data should be evaluated, and NMAT interpolation techniques should be reassessed;
• Use geostationary satellite and moored-buoy data to analyse the diurnal cycle of SST, particularly in the tropical west Pacific warm pool. It was recommended that the Voluntary Observing Ship Climate (VOSClim) Project be extended or a parallel project be initiated, to include buoys;
• Commence regular comparisons of the quality control (QC) procedures for SST. For these, common in situ input data should be used;
• Collate NOAA Pathfinder satellite SSTs for inland seas and large lakes;
• Develop sub-monthly analyses of SST since 1950;
• The Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) Expert Team on Sea Ice (ETSI) should provide recommendations on the blending of sea-ice data and the interpretation of microwave observations of sea ice. This would provide much needed information on variations in sea-ice thickness;
• It is desirable that the ETSI should provide an inventory of historical sea-ice data for the Southern Ocean;
• The use of satellite SSTs in relationships between SST and sea-ice concentration should

1 NOAA
2 United Kingdom Met Office
3 WMO
be re-assessed, owing to possible contamination of the SSTs by the sea ice. Improved sea-ice data and relationships should be incorporated into SST analyses.

Within five years
- Cloud-cleared techniques for satellite-based infrared SSTs should be compared;
- Regular comparisons of SST analyses should commence;
- All SST analyses need to include gridded fields of analysis error, including bias correction error. Error covariances are also needed;
- Create monthly and sub-monthly blended SST/sea-ice products. Estimates of errors and indications of sources of data should be included in the product.

Mean sea-level pressure and wind
Well within the two-year timeframe and ideally by early 2003:
- The Hadley Centre global monthly MSLP dataset (HadSLP) should be updated;
- The terms of reference of the COOS MSLP Working Group should be expanded to include surface winds;
- A catalogue of available wind and pressure products should be developed.

Within two years
- Florida State University would have a non-global (Pacific & Indian Oceans) dataset of surface wind and MSLP, fluxes, and related variables from 1950 onwards;
- Appropriate techniques for the adjustment of both estimated and measured wind speed observations should be investigated and applied;
- Monthly wind statistics for 1884 to date should be computed using the adjusted estimated and measured winds;
- The Meteorological Service of Canada has created a high-resolution analysis of winds over the North Atlantic for 1958-1997. The use of historical daily MSLP fields to backdate this analysis should be investigated;
- Biases from the US Maury Collection pressure dataset should be investigated;
- More observations on pressure are needed to improve historical MSLP analyses;
- The new JCOMM buoy metadata base should be populated with current and historical data.
duplication and lack of QC in the interim products;
• Use a new, fully documented format for interim and newly digitized data;
• Continue development and application of new QC techniques and utilization of metadata;
• Continue wide distribution of all data in appropriate formats, and share software to access and analyse the data. Data should be available freely, e.g. over the Internet, or at a minimum cost for media.

Within 2 years
• Real-time data collection centres should keep original copies of the GTS data stream. A comparison of GTS receipts at these collection centres should be made;
• Modern high-quality data, at a higher observational frequency than standard synoptic periods, should be incorporated in I-COADS;
• There should be a mirror data site for the new I-COADS database.

Conclusion
The Workshop achieved its goals by:
• Creating a timetable for further enhancement of in situ marine datasets;
• Developing a strategy for creating and comparing alternative SST, sea-ice concentration and marine air temperature analyses, to provide estimates of uncertainty in analyses and key diagnostics of climate variability and change, and to allow assessment of the effects on atmospheric general circulation models of legitimate uncertainties in the analyses;
• Taking account of recommendations made by the second Workshop of the CLIVAR C2OC Pro-

The Workshop thanked Dr Joseph O. Fletcher (left), who inspired the original COADS project in the 1980s. Several speakers acknowledged his major contribution, and participants signed a certificate in his honour.
ject. These included acquiring current and historical SST data for inland seas; archiving quality-controlled SSTs and their uncertainties for assimilation into coupled GCMs; assembly of tropical skin SSTs to test model sensitivity to their use and to the diurnal cycle; provision of analyses with estimates of error associated with each grid-box; testing the sensitivity to use of alternative SSTs; creation of sub-monthly SST analyses from 1950; acquisition of sea-ice thickness information to improve heat fluxes; incorporation of historical Russian sea-ice data;

- Proposing the further development of analyses of marine surface pressure and winds, with support from the new CCOS MSLP Working Group.

In its final plenary session, the Workshop voted in support of the name International Comprehensive Ocean-Atmosphere Data Set (ICOADS) for the new blended observational database. This name recognizes the multinational input to the database while maintaining continuity of identity with COADS, which has been widely used and referenced. The Workshop was an appropriate lead-in to the conferences planned by JCOMM for September 2003 in Brussels, to commemorate the 150th anniversary of the conference convened in Brussels in 1853 by US Navy Lt. Matthew Fontaine Maury to establish inter alia, the standardization of meteorological and oceanographic observations from ships at sea. Maury’s work (see Lewis, 1996) remains the foundation of much operational and research maritime meteorology and oceanography.

Reference

Coverage of COADS data with sea level pressure goals

Older version of COADS

1810-1819 SLP
1820-1829
1830-1839
1840-1849
1850-1859 SLP
1850-1859
1860-1869
1860-1869
1870-1879
1870-1879
Older version of COADS

Recent version COADS

1880–1889 SLP

1890–1899

1900–1909

1910–1919

1920–1929

1930–1939

1940–1949
COADS data coverage plots online
(with color for decades 1800 - 1997)

- six pages follow (black and white)
  - please see the online color version
  at NOAA Boulder
  - thanks to Scott Woodruff
  - Roy Hanna
  Jan 2004

Hi Roy,

CoADS ship (and buoy) data coverage

Follow-up on our discussion yesterday of the decadal coverage maps:

The maps we currently have on-line compare the original Release 1 (left-hand column) coverage with I-COADS Release 2.0 (right-hand column).

Enclosed are printouts from two of our webpages:
1) Release 2.0 (1784-1997) Overview: Figs. 1a through 4d compare the coverage for four different variables.
2) Publications: PDFs of the WMO Bulletin article (figures are poor quality), and from that article of the decadal coverage maps for SLP (see text for figure caption).

Comments:
-- Wind data extend the farthest back (US Maury Collection).
-- SLP data coverage during the 1850s through 1890s decades was dramatically improved (compare SST for the same decades). The main factor was recovery and gravity correction of Dutch (deck 193) supplementary SLP data. Deck 193 was a crucial 19th century deck for COADS Release 1, but the supplementary SLP data were not extracted.

If at some point a certain selection of coverage maps in b/w or color would be useful, I think we could assemble them without too much difficulty. Let me know if there are additional questions.

Regards,

From: Scott Woodruff, 22 July 2003

Scott Woodruff
NOAA/OAR/CDC (R/CDC1)
325 Broadway
Boulder, CO 80305 USA

phone: +1 303-497-6747
fax: +1 303-497-7013 or -6449
e-mail: Scott.D.Woodruff@noaa.gov
web: www.cdc.noaa.gov/coads/
Release 2.0 (1784-1997) Overview

Background

The observational data, summary products, and new capabilities that form I-COADS Release 2.0 were completed around August 2001. The underlying data for Release 2.0 are from three intermediate releases, described on the following webpages:

- Release 1a (1980-97)
- Release 1b (1950-79)
- Release 1c (1784-1949)

The full observational archive (1784-1997) and products (1800-1997) in binary formats are now freely available on-line. We also offer processing to create subsets in either binary or ascii formats. These and other data distribution options are available at NCAR for marine surface observations and for monthly summary statistics. Or, NOAA/CDC offers netCDF format for the monthly summary statistics.

Figures 1a through 4d below illustrate the enhancements in coverage gained in Release 2.0, in comparison to the original COADS Release 1 data. These show total numbers of observations per decade for 2° latitude x 2° longitude boxes for the indicated variables:

- Fig. 1a: sea surface temperature (SST)
- Fig. 1b: wind speed (WSPD)
- Fig. 1c: sea level pressure (SLP)
- Fig. 1d: relative humidity (RHUM)

Figure letters indicate the decadal coverage:

- 1800-09,...,1840-49 (a figures)
- 1850-59,...,1890-99 (b figures)
- 1900-09,...,1940-49 (c figures)
- 1950-59,...,1990-97 (d figures)

The Release 1 data (left-hand column on each figure) are from the "trimmed" version, and the Release 2.0 data (right-hand column) are from the "enhanced" version. Decades prior to 1854 and after 1979 are blank for Release 1 because it only covered 1854-1979, whereas statistics for Release 2.0 cover 1800-1997.

A wide range of colors was used to show different levels of detail during earlier and later time periods. Colors were assigned to ranges of total observations per decade, each equivalent to an average maximum number of observations per year or month (note that all observations within a given decade could fall within one month):

- Decadal average
- Range maximum

http://www.cdc.noaa.gov/coads/v2.html
Release 1

Release 1 had data from 1854-on.

Release 2.0

1800–1809 SST

1810–1819 SST

1820–1829 SST

1830–1839 SST

1840–1849 SST

COADS 2001–12–12