Newly Digitized Historical Marine Data and Metadata Becoming Available for COADS and the UK Meteorological Office Marine Data Bank

J.D. Elms (1), S.D. Woodruff (2), D.E. Parker (3), and S.J. Worley (4)

(1) NOAA/NCDC, Asheville, NC, USA.
(2) NOAA/ERL Climate Diagnostics Center, Boulder, CO, USA.
(3) Hadley Centre, Met. Office, Bracknell, UK
(4) National Center for Atmospheric Research, Boulder, CO, USA.

Abstract. Accounts are given of progress made, obstacles overcome, and efforts still required, to make recently digitized historical marine data and metadata from a number of maritime nations available to climate researchers and other data users worldwide. Recently keyed data are being uniformly reformatted and integrated into the Comprehensive Ocean-Atmosphere Data Set (COADS) and its blend with the UK Meteorological Office Marine Data Bank (MDB). Then the observational data, metadata, and associated summary products can be made readily and widely available. Recently digitized collections from the US, Norway, and the UK are highlighted. Data or digitization contributions by countries such as China, Japan, Russia, and the Ukraine also are discussed. In addition, prospects are assessed for the future availability of still undigitized ship logbooks housed in national archives or data centers around the world. An indication will be given of the possible benefits to future analyses of climatic variations.

1. Introduction

Heightened concern over climate variability and global change, and the role of the Oceans, has lead to increasing interest in historical marine surface data. Data inventories completed from Release 1 of the Comprehensive Ocean-Atmosphere Data Set (COADS; Slutz et al., 1985) revealed major data gaps during World War I and II, as well as significant spatial variations in data coverage (see Fig. 5 in Woodruff et al., 1987). Even after more recent COADS updates (Woodruff et al., 1998), Figure 1 illustrates the extent of continuing gaps and variations in the historical record.

Decreases and changes in historical ship data coverage are to some extent intrinsic (e.g., changes in ship routes to reflect global trading patterns, advances in ship design,[1] and historical events). Nonetheless, over the last decade much progress has been made in identifying undigitized historical marine collections that could significantly improve our knowledge of the global ocean climate. An important corollary concern relates to the archival of original paper records; these deteriorate through time, and in some cases digitization is required to ensure preservation of information.[2] The "International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata" held in Toledo, Spain in September 1997 (Diaz and Woodruff, 1999) proved valuable in bringing together representatives from many maritime nations to discuss the status and availability of additional data and metadata for the instrumental and pre-instrumental periods back to about 1500. Untapped ship data resources and other contributions towards making these important data available to the scientific community, were discussed at the workshop by researchers or representatives from Canada, China, France, Germany, Japan, Netherlands, Norway, Russia, Spain, Ukraine, UK, US, and WMO. The meeting highlighted work in progress, and has encouraged new efforts, by organizations to locate the significant resources needed to digitize logbook records, and to quality control and convert them into a common format. It also helped

[1] Transition from wind to steam power affected the shipping routes and observing practices (e.g. wind speed and wave height estimates). With the introduction of larger ships (e.g. supertankers) fewer voyages are required resulting in fewer observations and some of the larger ships are unable to use the shipping canals thus altering routing.

[2] Microfilm, or ideally digital images of deteriorating original records are also highly desirable since digitization for climate research may be a selective process not well tailored to other uses; e.g. remarks; in ship logbooks typically have not been digitized or fully digitized).

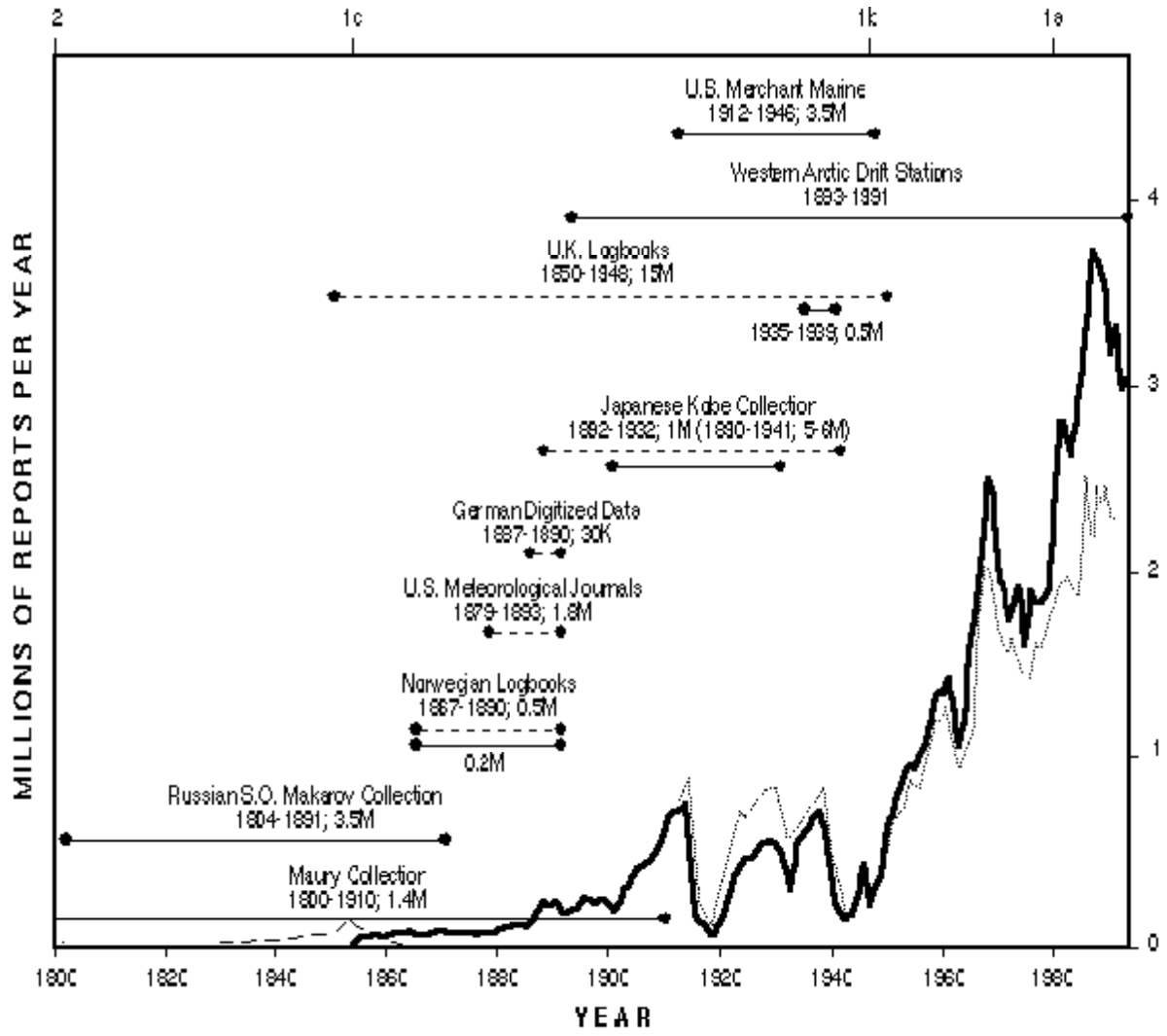

Figure 1. Annual global marine reports (solid curve) for COADS Release 1 through 1949, continued by Release 1b for 1950-79, and by Release 1a for 1980-95 (after removal of certain and uncertain duplicates, and "landlocked" reports). For comparison, the dash/dot curve shows 1950-79 Release 1 and 1980-91 "interim" data. Horizontal lines span the time periods for data that have been collected and digitized (solid), or are partially digitized (dashed line above a solid line) or proposed for future work (dashed), with the approximate numbers of reports shown in millions (M) or thousands (K) (see Elms et al., 1993). Data additions available from the Maury and U.S. Merchant Marine 1912-46 Collections are shown by the dashed curve. Currently, only 1.1 M merchant marine reports from the Kobe Collection for 1901-1932 have been digitized, representing a subset from the 5-6M unkeyed reports (in parentheses). Labeled ticks along the upper horizontal axis mark the starting years for Releases 1a (1980), and 1b (1950) and planned for Releases 1c (1854) and Release 2 (~1820, or earlier).

focus attention on areas and time periods where little digital information was previously available, e.g., the two World War periods.

Finding merchant marine observations to enrich periods around the two World War has proven especially difficult, because it appears few survived. Fortunately, Japan's Kobe Collection contains good coverage in the Pacific and to a lesser extent in the North Atlantic around WW I. Data for this period were recently digitized by Japan and issued on CD-ROM (the subject of a separate conference paper). For WW II, and possibly WW I, limited data enhancements might be available via extraction of observations from daily weather charts, e.g., the US Weather Bureau Northern Hemisphere Synoptic Charts[1] (US EDIS, 1899-1971).

[1] However, when observations were plotted, the exact position and time, were frequently lost, wind speeds plotted to the nearest five knots, etc. Before any full-scale project, a pilot test and comparison against existing COADS coverage would be advisable.

We anticipate that future research into climate variations will significantly benefit from availability of these data to flesh out past climate signals from the available historical record. Individual observational records can be used to create analyzed products, and which in turn provide more easily used inputs to climate model simulations and a wide variety of other research. In preparation for the Intergovernmental Panel on Climate Change (IPCC) Scientific Assessment planned for around the year 2001, the UK Meteorological Marine Data Bank (MDB) is being blended with COADS (as described in a separate conference presentation), to also include new historical sources that can be converted in time. Given limited resources, conversion and analysis efforts need to be prioritized to enhance coverage in data sparse times and regions, and on critical variables, e.g., sea surface and air temperatures (SST and AT), wind, and pressure. Statistical analyses must consider a variety of factors, including uneven distributions of the data spatially, temporally, and diurnally, and inhomogeneities introduced by observational practices, instrumentation, reporting and platform type changes through time. Some of these factors are considered in separate conference papers.

Development and usage of robust modernized keying formats and procedures to maximize the capture of data and metadata from older historical collections has also raised the scientifically important question of how some of the oldest existing digital collections already in COADS were created. Some early collections were prepared and keyed as early as the first half of the 20th century, and typically little or no documentation exists about how the data were obtained or handled. To address the question of whether spurious climate signals might be embedded in the oldest collections as an artifact of divergent procedures, it may be possible to perform intercomparisons of overlapping data sources.

Section 2 of this paper provides background on some of the logistical considerations surrounding digitization of data and metadata, and preparing them for availability to the scientific community. Sections 3-5 describe recently digitized (or ongoing digitization of) collections of US, Norwegian, and UK data, respectively. Section 6 briefly discusses collections planned for, or in the process of digitization, by other countries. Section 7 describes future plans.

2. Background on digitization and conversion to common formats

Effective procedures for digitization, and for conversion of historical marine data and metadata to a common format, pose difficult challenges and require substantial resources. However, it is critical that these foundation steps be implemented with sufficient reliability to maximize the probability that legitimate scientific conclusions can be drawn from the historical record.

A generalized approach to keying historical marine reports was designed for the US Merchant Marine 1912-46 Collection (Elms et al., 1993), and has since been refined and used more widely. Ship logbooks or log forms typically are designed with an initial set of metadata in the front of the logbook or at the top of the log form, followed by multiple observations. The keying approach mimics this design by creation of a "header" record and multiple "observational" records, which can be linked through a unique "voyage" number.

Furthermore, we have adopted a general strategy to key data in their original units or form to the maximum extent practical, with conversion to modern units, if appropriate, delayed until later conversion to a common format. It is advantageous to delay these decisions until conversion, since the data do not need to be re-keyed if a conversion is re-thought, but it should be noted that the conversions are therefore complex and expensive to implement. These approaches minimize keying, and, provided the header and observational records are permanently archived, conversions to a common format can be implemented iteratively when improvements are made in conversions into modern units, instrumental adjustments, etc.

The COADS Long Marine Report (LMR) format (variable-length, mixed binary and characters) currently is used for production COADS processing (updating the database). This is more complex than the fixed-length binary available (and simple ascii formats under development) for typical user access, but it possesses several advantages for production work. For example, source identification is tracked at several levels, so that data can be screened as carefully as possible by platform type and other characteristics. For each collection, we have found it extremely helpful as part of the metadata to document in a standard text format specifications for the rules used to convert each original data field into the corresponding regular LMR field.

Regular fields within the LMR format are based largely on WMO's contemporary Ship and Buoy Codes, which are considerably different from observational practices of a century or more ago. The regular fields may be followed by an arbitrary number of "attachments," currently defined for supplemental or erroneous data, and for quality control information. One of the major strengths of the LMR format is the supplemental attachment, which allows information that cannot be converted into the regular fields to be stored in original form to be tapped for later use. This approach also allows some prioritization of variables since conversion can be tackled via an iterative approach (e.g.,

initial conversions may be scaled down to key variables, while carrying along the necessary information to expand the scope of data availability at a later date).

3. US collections

Presently, US merchant ship data generally only extend back to 1949 and some US Navy data back to 1920 in COADS. Sections 3.1 through 3.4 provide details on the digitization and progress in preparation of four important new US collections planned for inclusion in COADS in the near future, or as part of later updates.

3.1 US Merchant Marine Collection (1912-46)

In recognition of the major gaps surrounding the two World Wars in the existing digital record, efforts initiated in in the late 1980s to locate and digitize US merchant logbooks encompassing these periods. Data for 1912-46 were requested from NCDC archives located off-site in East Point, Georgia (a few records also were keyed for 1910-11 which were inadvertently misfiled with the requested years). For the WWI period, there was a steep reduction in the number of available records (as illustrated in Figures 1), which may reflect reduced merchant shipping or a loss of original records. In contrast, for the WWII period, few records appear because the merchant weather reports were under custodianship of the Maritime Administration, which unfortunately destroyed the logs in 1974. The certificate of destruction is shown in Figure 2a; the accompanying appraisal report, Figure 2b stated that " the item for which disposal authority is requested is disposable because it does not have sufficient value for purposes of historical or other research".

Approximately 3.5 million records from the collection were keyed at NCDC. Since this was the first historical ship data keying effort initiated in recent years at NCDC, many lessons were learned in development of keying formats and procedures, and in the application of quality controls. For example, 25 different format variants were established to store data in different units (e.g., temperatures in Fahrenheit vs. Celsius) and to accommodate different original form types (e.g., Figure 3). To add to the complexity, two different database systems were used for storage of the keyed data. Budgetary constraints furthermore made double-keying impractical, which would have reduced the number of keying errors and improved the overall quality.

A significant intrinsic deficiency is evident for the period 1912 though the mid-1930s in this Collection because observations were requested only at Greenwich noon by the US Weather Bureau. Spurious trends might be introduced into the climate record by false diurnal signals depending on where the observations were reported around the globe (e.g., varying between times of maximum and minimum heating, depending on the ship location), thus these data will have to be utilized with caution when combined with ship reports taken at different schedules. However, it should be noted that the

RG: 178 +
RG: 357

REQUEST FOR AUTHORITY
TO DISPOSE OF RECORDS

LEAVE BLANK
DATE RECEIVED: 26 APR 1974
JOB NO.: NC 174-221
NOTIFICATION TO AGENCY

TO: GENERAL SERVICES ADMINISTRATION
NATIONAL ARCHIVES AND RECORDS SERVICE, WASHINGTON, DC 20408

1. FROM (AGENCY OR ESTABLISHMENT): Dept. of Commerce - Maritime Administration
2. MAJOR SUBDIVISION:
3. MINOR SUBDIVISION:
4. NAME OF PERSON WITH WHOM TO CONFER: William J. Fitzpatrick
5. TEL. EXT.: 189-2477
6. CERTIFICATE OF AGENCY REPRESENTATIVE:

In accordance with the provisions of 44 U.S.C. 3303a the disposal request, including amendments, is approved except for items that may be stamped "disposal not approved" or "withdrawn" in column 10.

5-17-74 James E. O'Neill (Date) Archivist of the United States

I hereby certify that I am authorized to act for this agency in matters pertaining to the disposal of the agency's records; that the records proposed for disposal in this Request of 1 page(s) are not now needed for the business of this agency or will not be needed after the retention periods specified.

5/1/74 (Date) William J. Fitzpatrick (Signature of Agency Representative) Records Management Officer (Title)

7. ITEM NO.	8. DESCRIPTION OF ITEM (With Inclusive Dates or Retention Periods)	9. SAMPLE OR JOB NO.	10. ACTION TAKEN
	Records of the U.S. Maritime Commission and the Maritime Administration.		
1.	Deck Department Log Books dated from ca. 1940 through December 31, 1947.	II-NNA-177A and NN-162-92	
	Disposition: Destroy immediately.		
	These log books consist of about 7000 cubic feet of records that are presently in the New York and San Francisco Federal Archives and Records Centers.		

Figure 2a. Destructive Certificate signed in 1974 for the disposal of the US Merchant Marine Logbooks (circa 1940-1947).

APPRAISAL REPORT
Disposal Job No. NC-174-221

Approved for Disposal:

The Item for which disposal authority is requested is disposable because it does not have sufficient value for purposes of historical or other research, functional documentation, or the protection of individual rights to warrant permanent retention by the Federal Government.

Appraiser: Ronald L. Heise — RONALD L. HEISE — 10 May 74 Date

Approval Recommended: Thomas W. Wadlow — THOMAS W. WADLOW, Acting Director, Records Disposition Division — 5/10/74 Date

Approval Recommended: Meyer H. Fishbein — MEYER H. FISHBEIN, Director, Records Appraisal Staff — 5/14/74 Date

Figure 2b. Appraisal Report for the disposal of the US Merchant Marine Logbooks (circa 1940-1947).

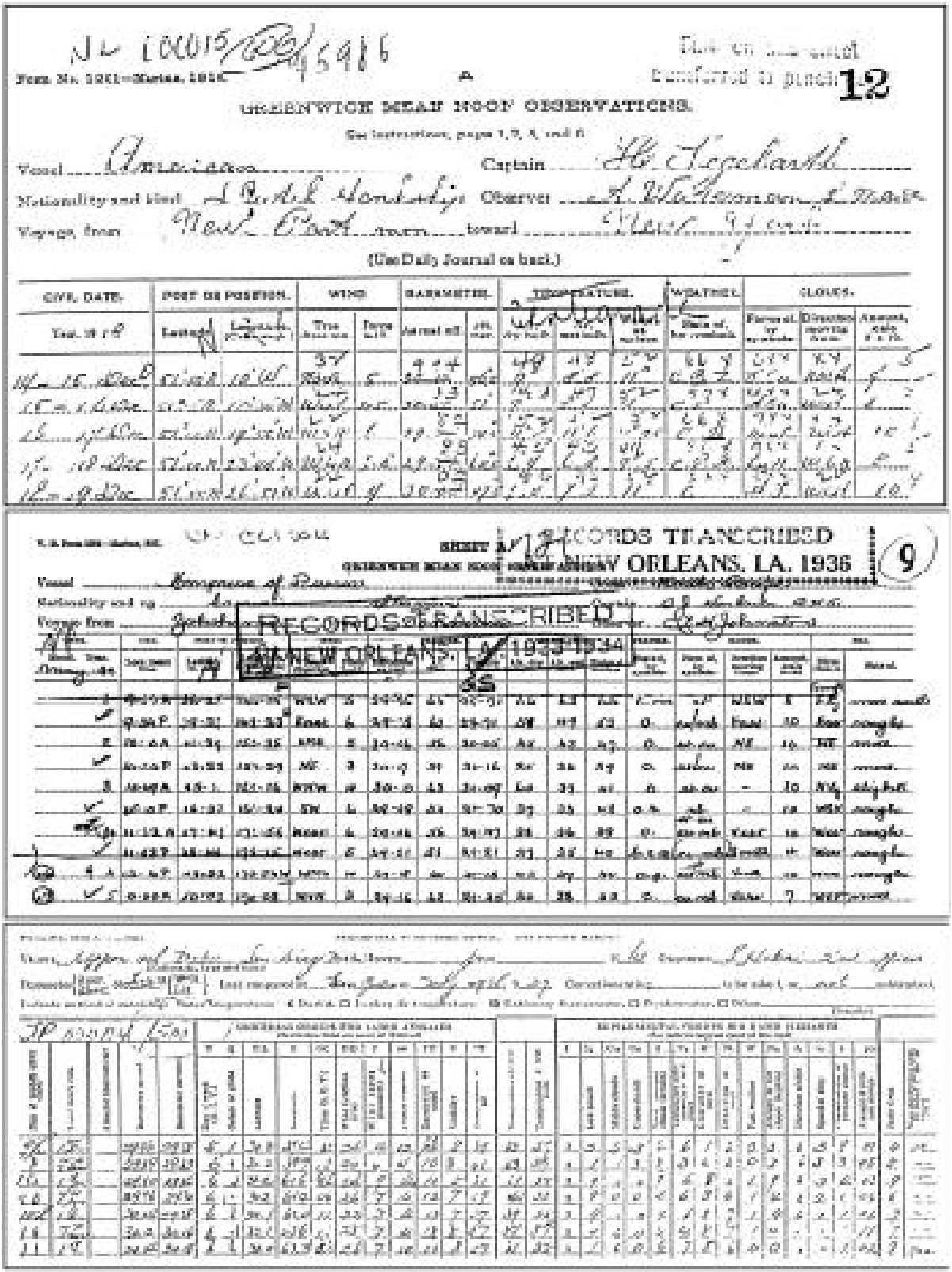

Figure 3. Three observational form type samples from the US Merchant Marine Collection (1912-1946).

available digital record already contains some significant diurnal variations in the reporting procedure (see Figure 4 in Woodruff et al., 1998).

Extensive manual and some automated quality control processing has been applied to a large portion of the 1912-46 data to help ensure correctness, but more work is needed. Nevertheless, much progress has been made on development of the complex conversion program, and we expect that NCDC will complete an initial conversion of the available data (QC'd, plus some partially QC'd data) near the end of 1999.

3.2 Maury Collection (1792-1910)

US Navy Lieutenant Matthew Fontaine Maury encouraged the systematic collection of marine data for the construction of pilot charts (Lewis, 1996), and later was a key participant in the development of a uniform international system for meteorological observations (Maury, 1853). The Maury Collection includes many of the earliest ship logbooks in the US National Archives and Records Administration (NARA) collections. The majority of the logs were collected by Maury and his staff, ranging from the early 1830s until the beginning of the Civil War in 1861 when he left his post as Superintendent of the Naval Observatory (see Figure 4), but the collection also contains some data as late as 1910 and as early as 1792.

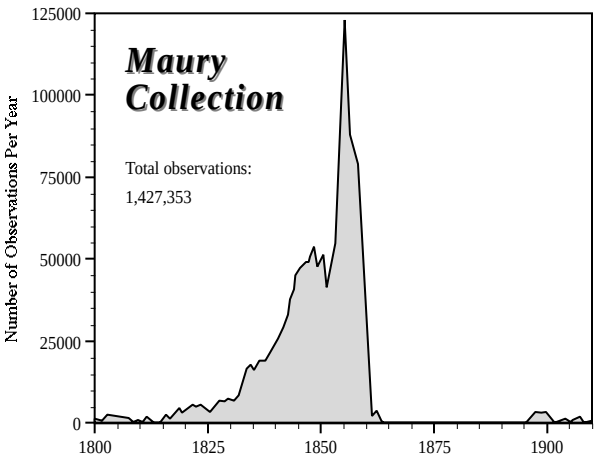

Figure 4. Number of digitized observations per year in the Maury Collection

Under a cooperative agreement, the Maury Collection was keyed by China during 1994-1996 and is available in the original keying format on CD-ROM from NCDC. The keying format represents an advance over the very complex formats used for the 1912-46 data, with indicators to flag different data units within one generalized format, and other simplifications. However, due to differences in early observational and reporting procedures, and gaps in the documentation, the collection poses difficult challenges in conversion to LMR format. But since these records are among the earliest digitized, it is critical at least that high-priority data elements including SST be made available within COADS as soon as practical.

Contemporary ship observational practices dictate that data are to be reported within ten minutes of the observational time for most data elements (although some elements such as winds and waves estimated from the ocean surface patterns may reflect longer time scales). In addition, voluntary observing ship data now are typically reported at 6-hourly synoptic times. In contrast, early reports in the Maury Collection (e.g., Figure 5.) usually were made just once a day (position reported at local noon), but included data for the previous 24 hours: temperatures and possibly pressure at 9AM, and wind direction (assumed magnetic) reflecting prevailing direction over three 8-hour parts of the day. Wind force and present weather conditions were generally entered as remarks. Questions also exist about whether pressure values were corrected or "as read."

Elms (1997) outlines the proposed approach for initial conversion of the Maury data into LMR format, involving creation of a (nominal) local noon report containing data elements whose conversion seems practical with present knowledge, plus accompanying supplemental data. After conversion, comparisons will be desirable against available climatological information to verify that the converted values are in an acceptable range for a given geographical region. In contrast to the Greenwich noon observations in the 1912-46 collection, the local noon observations in the Maury Collection provide somewhat of an improvement in data continuity, since they were all taken near maximum heating, and would generally reflect the temperature changes produced by either cold or warm air advection.

3.3 Western Arctic Drift Station Data (1922-1976)

The Arctic is an area where data coverage is limited, but data are critical to address climate variability and global change questions. NCDC, in cooperation with the National Snow and Ice Data Center (NSIDC) and NOAA's National Geophysical Data Center (NGDC), located or keyed data for the period 1922 - 1976 for a number of manned stations on ice floes, or ships overwintering in the ice. These data will be converted to LMR format as resources permit.

3.4 US Marine Meteorological Journals (1879-1893)

This large NARA collection will be available to help enrich coverage in the late 19th century, once

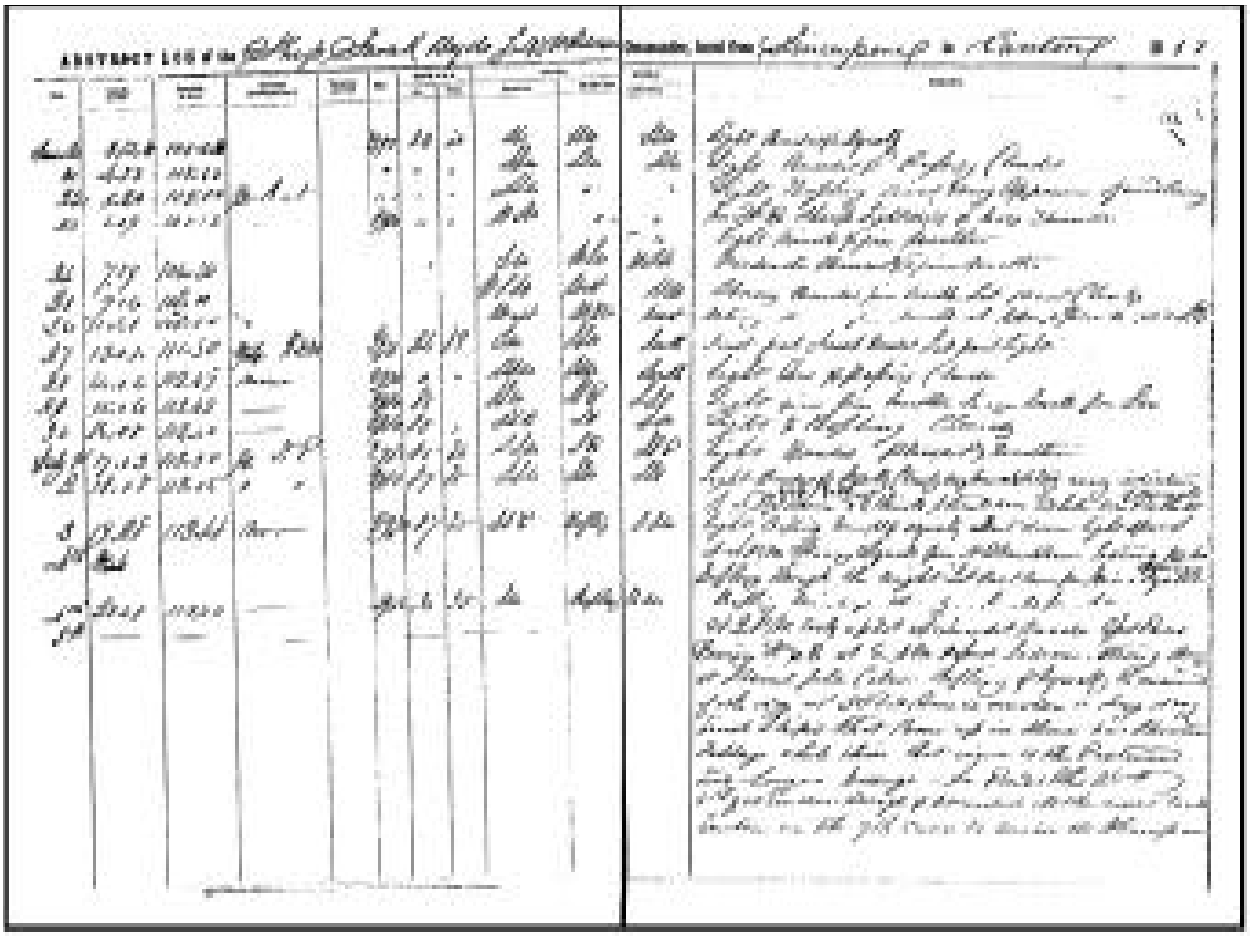

Figure 5. Example of a typical abstract log sheet from the Maury Collection.

digitization is complete (planned for 2000). Agreement was reached May 1998 for the Chinese National Marine Data and Information Service (NMDIS) to digitize and quality control the records, and in exchange NCAR provided NMDIS with a large land station telecommunications archive. Resources were obtained for NARA to microfilm all the approximately 1,955 logbooks (estimated 378 rolls). As the Journals are filmed by NARAthey are periodically shipped to NCDC where paper copies are generated and shipped to China. The digital records will then be shipped back to NCDC, and later converted into LMR for future merger (estimated 1.8 million observational records).

The content of the journals in this series is quite consistent, and generally comprised of 2-hourly observations (12 per day). Figure 6, is an example of a typical journal page. The basic elements reported are wind direction and Beaufort force, air temperature, water temperature, pressure, wet bulb temperature, present weather, proportion of clear sky in tenths, cloud types and state of the sea. Positions were typically reported only at local noon (due to navigational limitations during the 19th century), so conversion will require interpolation of positions (with a flag set in LMR format to differentiate reported from interpolated values). The beginning of each journal usually contained an extensive set of instructions to the observer, which were crucial in developing the keying format, and also form important metadata, which we hope to be able to transcribe into an electronic format for Web access at a future date. The keying format developed in cooperation with China represents an extension of that used for the Maury Collection, with additional features to adequately and efficiently capture the unique structure of this collection, and all important data elements.

Although the collection is more uniform than the Maury collection, a variety of data problems, inhomogeneities, and questions already have been encountered. For example, the observer's instructions documents the Beaufort force scale including the typical descriptive phrases, but in addition contains an equivalent scale in pounds per square foot and in miles per hour which does not convert closely to the equivalent scale in knots adopted by the International Meteorological Organization, Washington, 1947; the extent to which this difference may have affected the

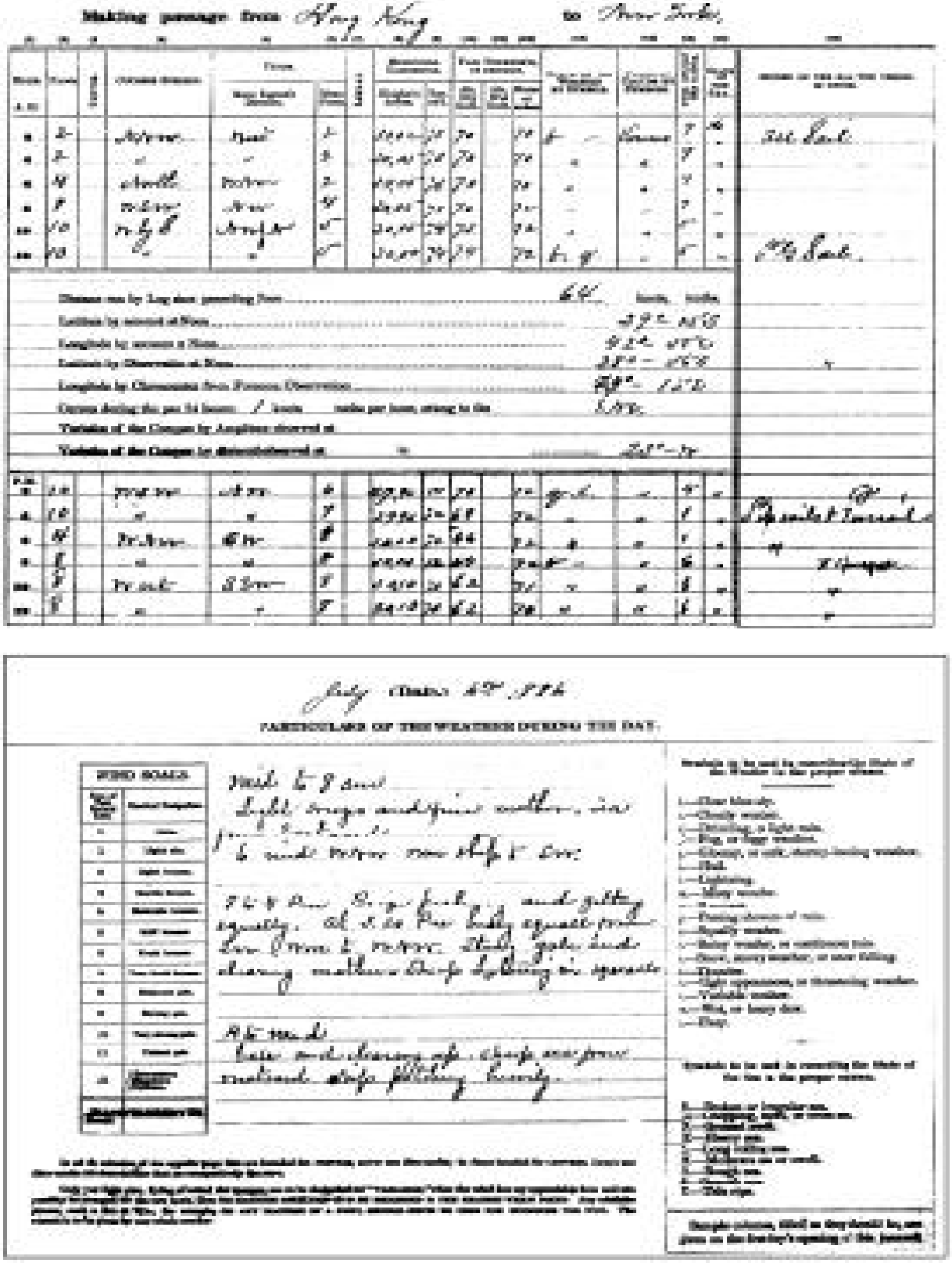

Figure 6. Example of a typical log sheet from the US Marine Meteorological Journals (1879-1893).

observers selection is not known. We have also found that observers frequently did not follow the instructions precisely, leading to data that may be unusable or difficult to interpret or process (e.g., fractional wind directions, or amounts of cloud, rather than the expected proportion of clear sky, in tenths). Figures 7a and 7b illustrates a problem in the cloud fields where the observers were confused about the symbols that should be entered, rendering the data unusable. It appears that wind , temperature, and pressure data will be the most comparable to current observational practices and can be blended into the historical record with the most confidence.

covering 1867-1890 and containing approximately 600,000 individual observations. Of these, 200,000 have been keyed to date by the AmS and provided for eventual inclusion in COADS (Woodruff et al., 1999). Analysed historical pressure fields were provided by NCAR to AmS for research purposes to help initiate the project. The format was designed cooperatively with the US along the lines of that used for the Maury Collection (including header and observational records), with additional features to accommodate keying requirements at AmS (e.g., a comma-delimited format) and the unique requirements of this collection. The available 200,000 reports await conversion to the LMR format before they

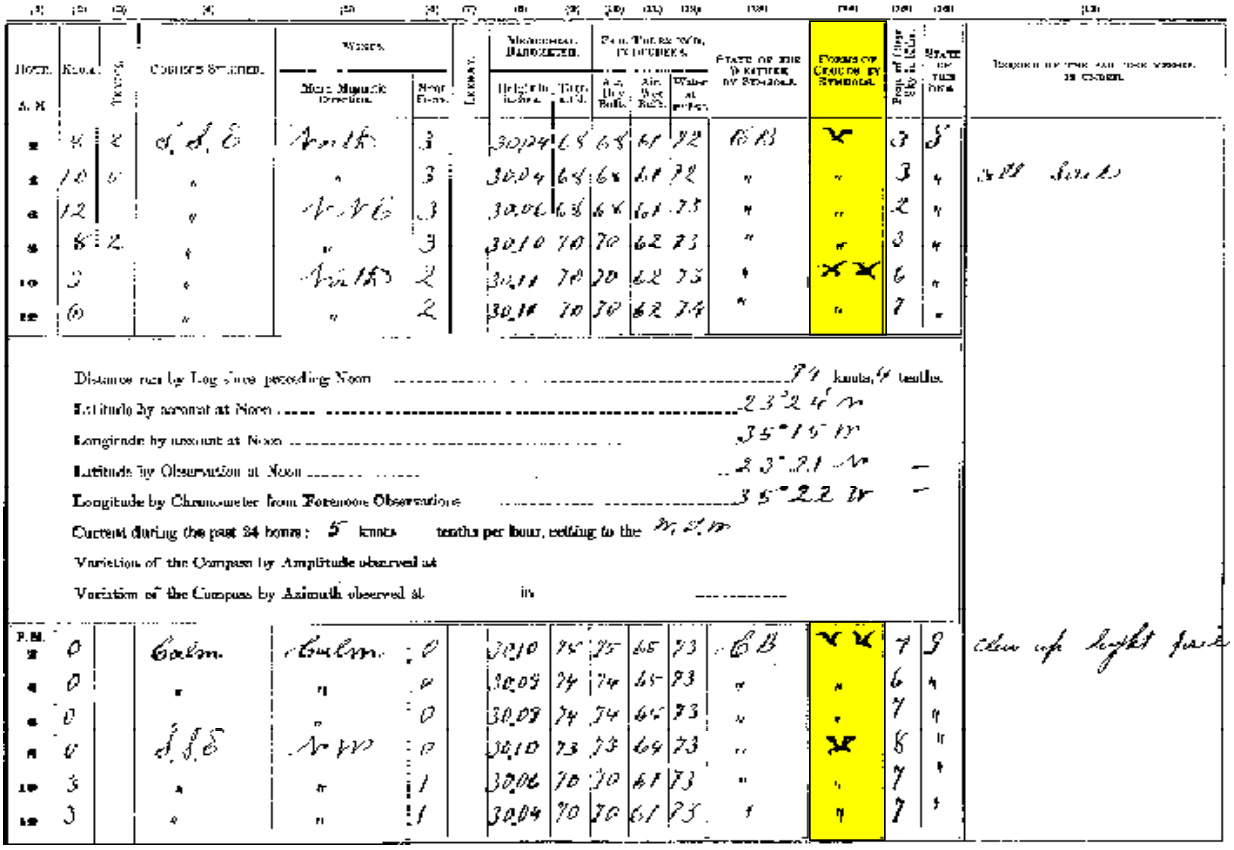

Figure 7a. A sample logbook page from a voyage in the US Marine Meteorological Journals(1879-1893) where the observer tediously drew in bird symbols to represent cloud types. Unfortunately, the same symbol could represent two different cloud types because the cross reference (number of birds; one to four) had one meaning on the primary cloud chart and a second meaning on the secondary cloud chart. Both these reference charts were published in the instructions to the observer.

4. Norwegian Logbook Collection (1867-1890)

In 1992, we became aware through Erik Wishman at the Archeological Museum in Stavanger (AmS) of the existence of approximately 600 Norwegian logbooks from the Norwegian Meteorological Institute (DNMI) can be merged in with the other historical data.

The logbooks in this collection are limited to two different editions, each related to a set of published instructions to the observer. The instructions were translated into English and provided in digital form by Dr. Wishman, and we plan to make these important metadata available for Web access at a future date. In

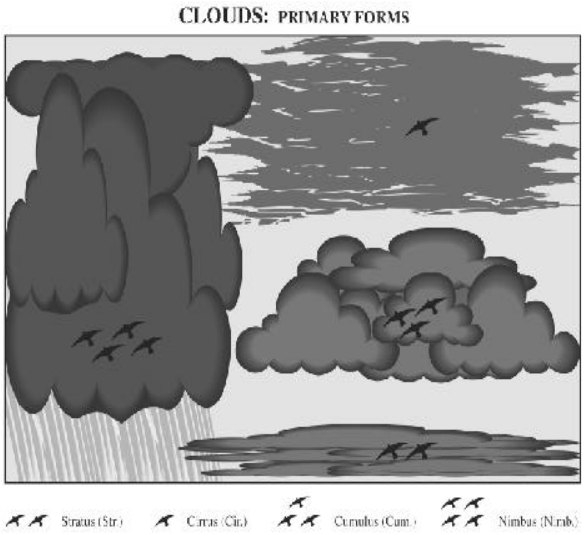

Figure 7b. In the instruction to the observer a figure similar to the one above (recreated as an example since the original images were to weak to reproduce) illustrating primary cloud types (a second illustration was include for the secondary cloud types) was published to help the observer identify the various cloud forms. The bird sketches, ranging from one to four birds, were used as a cross reference between the illustration and the cloud form name listed at the bottom of the chart along with the abbreviation to be entered on the log sheet to report the type of clouds observed. . The bird symbols regrettably cannot be converted to the proper cloud type without knowing if they were primary and secondary clouds.

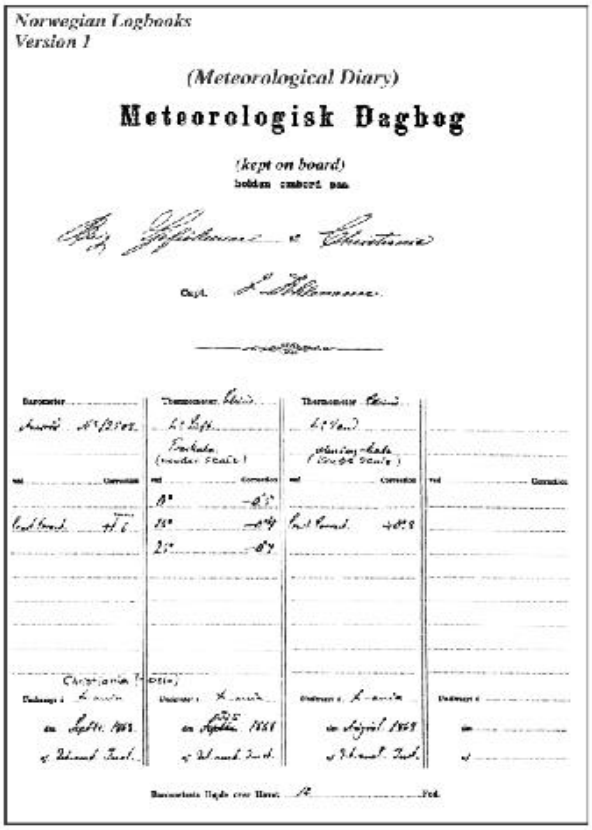

Figure 8a. An example of the type entries (metadata) generally available in the Norwegian logbooks regarding the voyage and observational instrumentation (1st edition).

general, the instructions were strongly based on the 1853 Maritime Conference (Maury, 1854) at which Norway was one of the 10 participating nations. Exceptions exist, however, including the use of a 0-6 (half) Beaufort scale, which may complicate merger of the wind speed data with existing digital records.

Figures 8a and 8b illustrate the header information and observational data available from the first edition of this collection. As in the US Marine Meteorological Journal Collection (1879-1893), positions are generally reported only at local noon, and interpolation will be required to make use of the intervening observations (typically 4-hourly according to local time). One substantial difficulty with this collection is that observers were instructed to enter uncorrected instrument values in the logbook pages, and any necessary correction metadata (e.g., calibration of a thermometer in melting snow) were to be entered once in the logbook header pages. Sometimes the calibration factors were left missing (e.g., if a logbook represents the continuation of a voyage, the observers may not have bothered to re-enter the metadata), or it may be unclear whether they should be added or subtracted.

5. UK Logbooks (1935-1939)

The UK Met. Office digitized nearly 460,000 observations from merchant ships logbook for the period 1935 to 1939 that were located in their archives to help fill the data gaps around World War II. For wartime reasons, data cease in 1939 (Parker and Jackson, 1999). These data were keyed in a structure similar to that used for the Maury and other collections, including a header record containing information such as record type, folio number, ship type and name, instrumentation information, etc. and data records. There were two formats established for the data records depending on the information available and this information is tied to the header record by the folio number.

In UK archives at the Public Record Office (see Millard, 1999 and Chenoweth, 1999) there remain approximately 15 million undigitized marine records (7 million for the period 1911-1920 and 8 million for the 1850-1900 period), plus earlier largely pre-instrumental records from the British East India Company. The cost

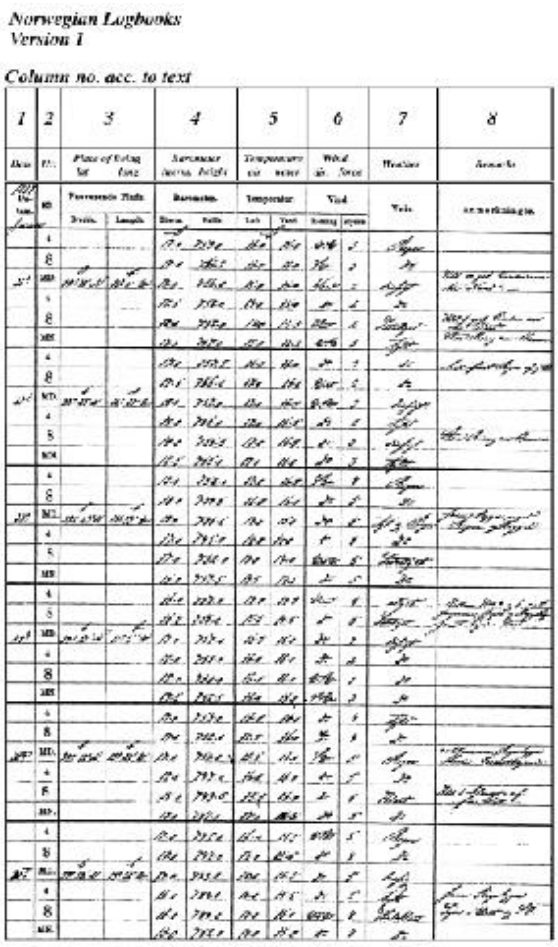

Figure 8b. A sample observational page from the Norwegian Collection (1st edition)

to digitize these records would be substantial, but the value to climate research at least for the instrumental data would be potentially very significant.

6. Other Ongoing Data Digitization and Recovery Activities

As noted in the section 1, Japan recently issued a CD-ROM (edition 1998) containing recently keyed Kobe Collection data for the data-sparse period around World War I. The project will be discussed in detail by Ms. Teruko Manabe of Japan Meteorological Agency in a separate conference paper.

In addition to a large digital archive of Russian and international ship data called MORMET covering approximately 1890-1992, which has already been partially merged into COADS, Russia has provided a small collection of very early ship data in the S.O. Makarov Collection (1804-1891; 3,500 reports). Other nondigital collections or digital archives on deteriorating magnetic media have been identified in Russia (Gemish et al., 1999). About six million reports taken aboard Russian oceanographic research vessels during 1936-1996, with most during the 1970s, form the most important collection. These observations have global coverage, are taken by experienced meteorological personnel, and include about 600,000 surface radiation measurements. Also about three million reports (1837-1969) reside on aging magnetic tapes. Joint US and Russian proposals with objective to digitize and rescue these data have been written. At this time approval has not been received to support activities to save these data.

The Ukraine (Polonsky, 1999) has at its Marine Hydrophysical Institute (MHI) in Sevastopol worldwide marine observations for 1960-1985. The Ukraine also possesses data from approximately 60,000 hydrographic stations in the Black Sea of which most have not been digitized. The digitization format for the surface marine observations has been defined and approximately 750 reports have been forwarded to NCAR. More reports are expected as resource become available at MHI.

7. Future plans

As part of the blend of COADS with the UK MDB in preparation for the upcoming IPCC Scientific Assessment (described in a separate conference presentation), we are attempting to include as many as possible of the recently digitized data from the US, Norwegian, and UK collections described in previous sections, as well as the Kobe Collection data for years around World War I. However, some of these data may have to be omitted from the first version of the COADS-MDB blend due to resource constraints, although we plan to employ scaled-down conversions of priority parameters (firstly SST) to the extent needed to maximize critical data coverage. Follow-up COADS-MDB blend plans will include more complete conversions of these data and other newly available data.

Additional undigitized collections of historical marine data and metadata dating as far back at 1500 are described in Diaz and Woodruff (1999). These include millions of still undigitized records from the Kobe Collection, and large collections in Russian, Dutch, German, and Spanish archives for example. Many of the observations prior the mid or early 19th century are probably pre-instrumental and would not be the first choice for digitization for climate research. However, many collections throughout the archives of the maritime nations would significantly benefit the global database and future research if the resources could be found to digitize and prepare the data and metadata so that it can be made readily and widely available for climate variability and other research.

8. Acknowledgments

COADS is the result of a continuing cooperative project between the National Oceanic and Atmospheric Administration (NOAA)--specifically its Environmental Research Laboratories (ERL), National Climatic Data Center (NCDC), and Cooperative Institute for Research in Environmental Sciences (CIRES, conducted jointly with the University of Colorado)--and the National Science Foundation's National Center for Atmospheric Research (NCAR). The NOAA portion of COADS is currently supported by the NOAA Climate and Global Change (C&GC) Program and the NOAA Environmental Services Data and Information Management (ESDIM) Program.

9. References

Chenoweth, M., 1999: Historical marine data in American and UK archives, 1775-1900. Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROA Report No. 43 (in press).

Diaz, H.F., and S.D. Woodruff (Eds.), 1999: Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. WMO/TD-No.957, MMROA Report No. 43 (in press).

Elms, J.D., 1997: Digitizing US historical marine collections (1792-1976). Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROAReport No. 43 (in press).

Elms, J.D., S.D. Woodruff, S.J. Worley, and C. Hanson, 1993: Digitizing historical records for the Comprehensive Ocean-Atmosphere Data Set (COADS). Earth System Monitor, 4, No. 2, 4-10.

Millard, C., 1999: Surface marine data holdings of the Public Record Office, UK. Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROA Report No. 43 (in press).

Gemish, Y.V., V.N. Popova, E.M. Krakanovskaya, S.M. Somova, G.M. Nazarova, and E.D. Vyazilov, 1999: Digitization and preparation of historical surface marine data from Russian research vessels. Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROA Report No. 43 (in press).

Lewis, J.M., 1996: Winds over the world sea: Maury and Koppen. Bull. Amer. Meteor. Soc., 77, 935-952.

Maury, M.F., 1854: Maritime Conference held at Brussels for devising a uniform system of meteorological observations at sea, August and September, 1853, Explanations and Sailing Directions to Accompany the Wind and Current Charts, 6th Ed., Washington, D.C., pp. 54-88.

Parker D. and M. Jackson, 1999: Digitization of UK merchant ships' data for 1935-39. Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROA Report No. 43 (in press).

Polonsky, A., 1999: Ukranian historical surface marine data, quality control, and interpolation. Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROA Report No. 43 (in press).

Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer, and J.D. Elms, 1985: Comprehensive Ocean-Atmosphere Data Set; Release 1. NOAAEnvironmental Research Laboratories, Climate Research Program, Boulder, CO, 268 pp.

.

US Environmental Data and Information Service (EDIS), 1899-1971: Northern Hemisphere sea- level and 500-millibar charts, Jan. 1899-June 1971 [monthly vols.]. Daily series, synoptic weather maps, part 1. (Title and responsibility vary.). N.p., n.d. ; Asheville, N.C. : National Climatic Center, 1972-1981 (oclc 3823539).

Woodruff, S.D., R.J. Slutz, R.L. Jenne, and P.M. Steurer, 1987: A comprehensive ocean-atmosphere data set. Bull. Amer. Meteor. Soc., 68, 1239-1250.

Woodruff S., E. Wishman, M. Nitter, S. Lubker, and F. Silva, 1999: Plans for incorporation of the Norwegian Ship Logbook Collection into COADS. Proceedings of International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata, 15-17 September 1997, Toledo, Spain. H.F. Diaz and S.D. Woodruff, Eds., WMO/TD-No.957, MMROA Report No. 43 (in press).

Woodruff, S.D., H.F. Diaz, J.D. Elms, and S.J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. Phys. Chem. Earth, 23, 517-527.