

**NOAA'S CLIMATE DATABASE MODERNIZATION PROGRAM
MARINE DATA RESCUE PROJECTS**

J. Eric Freeman *
Sourcecorp, Asheville, North Carolina

Heather Anderson
STG, Inc., Asheville, North Carolina

Mark Seiderman
NOAA National Climatic Data Center, Asheville, North Carolina

1. Introduction

The Climate Database Modernization Program (CDMP) supports the National Oceanic and Atmospheric Administration's (NOAA) mission to collect, integrate, assimilate and effectively manage Earth observations on a global scale, ranging from atmospheric, weather, and climate observations to oceanic, coastal, and marine life observations. Many of these holdings, which are part of the U.S. National Archives, were originally recorded on paper, film, and other fragile media, and stored at various NOAA Centers. Prior to CDMP, not only were these valuable data sources mostly unavailable to the scientific community, but storage technology for the archives was obsolete. Without proper preservation of the media, the information they contained was in danger of being lost forever. Today, CDMP has greatly improved the preservation and access to NOAA's holdings by migrating many of these resources to new digital media. CDMP partners with private sector contractors to image and key these records (Dupigny-Giroux et al. 2007) and have placed over 54 million weather and environmental images online for viewing through contractor-developed software: Environmental Document Access and Display System (EDADS; account and access information can be located at <http://www.ncdc.noaa.gov/oa/climate/cdmp/edads.html>), supporting researchers around the world via the Internet. In 2009, CDMP supported 86 NOAA data rescue/modernization projects. While CDMP strives to preserve all types of environmental data, this paper strictly highlights marine data rescue activities. During CDMP's ten year existence, more than 6.7M individual surface marine observations have been digitized from major archival collections such as the U.S. Merchant Marine records (3.5M) held at the National Archives and Records Administration (NARA) in the United States, as well as the German Maury Collection (544K) loaned from the Deutscher Wetterdienst (DWD) in Germany (Braun 2000).

Corresponding author address: J. Eric Freeman, NOAA/NCDC, 151 Patton Avenue, Asheville NC 28801; Email: Eric.Freeman@noaa.gov.

During this time, special attention has also been given to sources of marine metadata and weather observing practices of the present and historical ocean observer. These include instructions to observers, rare journals and publications, and ship and other marine platform metadata sources.

By preserving all types of historical data and metadata, CDMP is devoted to saving the past and making it available for future generations. By initiating and assisting in new projects, national and international collaboration has never been greater. The marine community is working together for the greater purpose of exposing the earth's past climate, so that the future global climate can be understood, correctly predicted, and adapted to by future inhabitants.

2. Current Marine Projects

CDMP is currently involved in numerous projects to extract surface marine observations from historic logbooks and tabular publications. With every digitization project, a plan must be devised specifically for each source and tailored to meet the projects' goals. The main goals typically include extracting all defined environmental variables as well as all associated metadata contained therein. Depending on the scope of the project, some may take a few months to complete while others may take multiple years. Often times, the duration of the project is dependent on resources and the annual budget appropriated for that task.

2.1 English East India Company (EIC) Instrumental Observations 1789-1834

In 2008, the Met Office Hadley Center imaged over 1100 of the original 2000 logbooks of the English East India Company (EIC) held at the British Library. The logs were selected based on their holdings of weather observations, visual and instrumental, as well as the significant spatial coverage of voyages from England to India and China through the Atlantic, Indian and Southern Oceans during the years 1789-1834. Commencing well before the landmark 1853 Brussels Maritime Conference (Maury 1854) devoted to coordinating an international effort for global systematic collections of marine instrumental and

visual observations, the EIC collection is likely the largest and earliest collection of such systematic instrumental observations.

The goal is to capture all noon observations containing: location; instrumental observations of pressure and air temperature (and occasionally sea surface temperature); and visual estimates of winds, state of weather and state of sea. Less than 900 of the 1100 already-imaged logbooks contain instrumental data and are being digitized. From the digitized logbooks, approximately 285K observations are expected, significantly increasing early instrumental coverage, spatially and temporally.

2.2 Greenwich Mean Noon (GMN)/ Simultaneous International Observations 1874-1947

During the late 19th century, the US began to coordinate global simultaneous ship observations at Greenwich Mean Noon (GMN). The goal was to collect as many simultaneous observations possible and standardize those observations in order to study the dynamics of the atmosphere (MWR 1914). This collection is in various formats including original logbooks and tabular publications, all together covering the full period of record. Currently two of the four parts are being digitized by CDMP and information on those follows in sections 2.2.1 and 2.2.2.

2.2.1 Bulletin of International Simultaneous Observations 1875-1889

In 1873 the International Meteorological Congress established the framework for the International Simultaneous Observations to all be taken at Greenwich Mean Noon (GMN). The U.S. Weather Bureau standardized and published these daily global observations in printed tabular publications from 1875 until 1889 (MWR 1914). Over 5M daily land and marine surface observations were submitted by thirty cooperating countries.

While all of the data contained within the publications will be captured, the initial goal is to capture the marine data from 1877-1884 submitted by over 600 vessels, providing over 200K daily marine observations around the world at the same time (GMT). Primary elements being captured include: locations; pressures; temperatures; winds; clouds; precipitation; sea and weather conditions. The land observations will be captured later in a similar process.

2.2.2 GMN Monthly and Daily Marine Forms 1910-1947

Approximately 500K paper records from the NCDC archives containing global marine records are being imaged with 190K of these records being digitized. The marine elements captured during this process

include: winds, barometric pressure, air (dry and wet bulb) temperatures, sea surface temperature, present and past weather, visibility, lower/middle/upper cloud types and amount, along with sea and swell data. Once this project is complete an additional 2.6M observations will be available, assisting in the understanding of the global marine meteorological environment.

2.3 U.S. Fish Commission Scientific Expedition Logbooks 1877-1948

In 2009 CDMP imaged nearly 400 original logbooks and scientific records from the US Fish Commission vessels archived at the Smithsonian Institution, including those of the *Albatross* and the *Fish Hawk* during the years 1877-1948. The records of the Fish Commission are significant in their coverage of the Pacific and Atlantic Oceans (including extraordinary 19th century voyages in the Bering Sea and waters in the far Northern Pacific) as well as a wealth of environmental data: meteorological, oceanographical, and biological. The environmental variables contained in these records include: air and ocean temperatures (multiple depths); state of weather; state of sea; barometric pressure; winds; cloud amounts; salinity; alkalinity; density and specific gravity of water; currents; tides; water depth; character of bottom; and species from trawls or dredges. By combining all of these variables into one dataset, CDMP assists researchers interested in the relationships between the climate and its effects on the environment (i.e. fish catch populations related to changes in sea surface temperature).

Also of high value is the remarkable and often recorded information on instrument makes and models used to record these elements during the voyages. Rich metadata on thermometers and trawl/dredge equipment are contained throughout the records. Digitization of these records will begin in 2010. An additional 100 deck logs from fisheries survey vessels were identified at NARA in 2009 and are recommended for future imaging and digitization.

2.4 U.S. Lightships 1916-1982

Throughout much of the 19th and 20th centuries lightships were anchored around the U.S. coastline to protect hazardous shipping routes and to keep merchant and navy ships from running aground on voyages along the U.S. east coast. The lightships were in operation from the 1820's to the 1980's, with a total of ~200 commissioned lightship vessels. As buoys and towers began to be used more and more in the latter 20th century, the lightships were ultimately phased out since the same job could be done permanently without the need for a manned crew. In the early 1980's the last lightship retired ending a 150 year lightship tradition in the US. The observations recorded on board these light vessels contain valuable meteorological data that is currently being

extracted from the original logbooks. Observations from 14 lightships are being captured to produce 430K daily observations. Elements being digitized include: wind direction; wind velocity; sea-level and station pressure; dry bulb, wet bulb and dew point temperatures; relative humidity; sea surface temperature; wave/swell; state of sea; visibility; present/past weather; clouds; and ceiling.

3. Partnerships

With a growing need for historical climate data over the global oceans many groups are now interested in locating, documenting and digitizing all sources of marine observations and metadata that can be located, past or present. Often times documents lie tucked away in various archives, libraries, or personal collections with the potential to be lost forever due to the fragile media on which they were originally recorded. By working with various national and international groups, coordinating efforts and information sharing are opening new doors in the world of data preservation and new sources are being located on a frequent basis. Working with multiple organizations also aids in eliminating duplication of effort. CDMP works diligently with the U.S. Voluntary Observing Ships (VOS) fleet, the International Comprehensive Ocean-Atmosphere Dataset (ICOADS), RECOVERY of Logbooks And International Marine data (RECLAIM), and the Atmospheric Circulation Reconstructions over the Earth initiative (ACRE) to facilitate and coordinate marine data rescue efforts around the globe.

3.1 U.S. VOS

Working with the US VOS, delayed mode ship observations recorded on paper logbooks are being digitized monthly by CDMP and included in NCDC databases as well as in international data exchanges with the Global Collecting Centers (GCC) in the United Kingdom and Germany. The delayed mode observations are used to verify observations that were transmitted over the Global Telecommunications Satellite (GTS) in real time, and many times the delayed mode observations include important metadata for Voluntary Observing Ship Climate (VOSCLIM) vessels not transmitted in the original electronic record. CDMP also imaged the complete *Mariners Weather Log* collection for the VOS in order to preserve that historical publication (Freeman et al. 2009).

3.2 RECLAIM

The RECLAIM project started in 2005 as an international effort to locate and document historical sources of marine data and metadata around the globe. Upon discovering new sources of information, RECLAIM has been very successful in documenting and prioritizing marine data sources for imaging and digitization (Woodruff, Freeman, Wilkinson et al.

2009), and has also given valuable expert assistance in interpreting historical marine records in relation to navigation, observational practices and recording (Wilkinson et al. 2009). CDMP works with RECLAIM to assist in these valuable duties. RECLAIM gave expert advice in a successful 2007 joint project between CDMP and the UK Met Office Hadley Center by selecting 10K scientifically valuable logbooks from the UK Royal Navy's logbook archive at the British National Archives where over 1.5 million global marine observations were rescued from the World War II (WWII) period - a time when global observations over the oceans are very sparse (Brohan et al. 2009). RECLAIM has also been a very valuable resource for the EIC digitization project by selecting all available logs with instrumental data and thoroughly documenting their contents.

3.3 ACRE

The ACRE initiative specializes in bringing together various groups around the world interested in data preservation and recovery in order to reconstruct and visualize (in 4D) the global climate over the past 250 years. Through major international collaborations, ACRE ensures that data is being discovered, recovered, digitized and used in reanalysis products that can be tailored and downscaled to seamlessly flow into various climate applications and production models (Allan 2009). As with RECLAIM, CDMP assists ACRE in locating and documenting new sources of marine data and metadata.

3.4 ICOADS

Another group interested in the preservations of historical marine data is the International Comprehensive Ocean-Atmosphere Dataset (ICOADS) (Worley et al. 2005, Woodruff et al. 2009). CDMP, as well as RECLAIM and ACRE, work together with ICOADS to assure that new sources of historical marine data are discovered, catalogued, digitized, and eventually translated into the International Maritime Meteorological Archive (IMMA) format (Woodruff 2007) so that the new data sources can be blended into the largest collection of surface marine meteorological observations, currently ICOADS Release 2.5 (Woodruff et al. 2009). With all groups working in tandem, assurance is given that the data is of the highest quality, has been thoroughly quality controlled, and is suitable for global climate studies.

4. Conclusion

As the world becomes more knowledgeable yet more complex, every piece of information that clarifies our knowledge of the earth's physical processes becomes more and more valuable. To better understand climates of the future, climates of the past must be appropriately represented and reconstructed. CDMP

is devoted to ensuring that any data that will help better understand the climates of the past and future is uncovered, preserved and made accessible online for researchers. While there are many sources of marine data available for future digitization with more being discovered regularly, the need to recover and preserve historic marine data is at an all time high. In order to do this, the experience of CDMPs ongoing projects combined with the collaboration of its' many partners assure that the records and their contents are never lost due to media degradation, and so that they can be included in the earth's historical climate record. Once converted into electronic formats, the data are more portable, can quickly and easily be shared, and can be analyzed to contribute to global climate studies. As more sources of data are discovered, the task of permanent preservation and data digitization will continue to be of high importance to CDMP and a fundamental necessity to the climate research community.

5. References

- Allan, R., 2009: Atmospheric Circulation Reconstructions over the Earth. 11/13/2009 <<http://www.met-acre.org/>>.
- Braun, D.S., 2000: Scientific vision, a passion for observation, and the impetus for change: Germany loans Maury logs to the National Climatic Data Center. *Earth System Monitor*, 11(1), 4-7.
- Dupigny-Giroux L.A., T.F. Ross, J.D. Elms, R. Truesdell, and S.R. Doty, 2007: NOAA's Climate Database Modernization Program: Rescuing, Archiving, and Digitizing History. *Bull. Amer. Meteor. Soc.* 88: 1015–1017.
- Freeman, E., T.F. Ross, 2009: 'NOAA's Climate Database Modernization Program (CDMP): Sailing to the Future with Historical Marine Observations'. *Mariners Weather Log*, Vol. 53 No.1, 04/2009.
- MWR (authors not available), 1914: International Simultaneous Observations. *Monthly Weather Review*, 42, 675-676.
- 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., E. C. and J. Biddle: Philadelphia, pp. 54-96.
- Wilkinson, C., S.D. Woodruff, P. Brohan, S. Claesson, E. Freeman, F. Koek, S.J. Lubker, C. Marzin, D. Wheeler, 2009: RECOVERY of Logbooks and International Marine Data: The RECLAIM Project. *Int. J. Climatol.* (submitted).
- Woodruff S.D., 2007: Archival of data other than in IMMT format: The International Maritime Meteorological Archive (IMMA) Format. *Second Session of the JCOMM Expert Team on Marine Climatology (ETMC), Geneva, Switzerland, 26-27 March 2007, JCOMM Meeting Report No. 50, 68-101.*
- Worley S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, N. Lott., 2005: ICOADS Release 2.1 data and products. *Int. J. Climatol.*25: 823-842. DOI: 10.1002/joc.1166.
- Woodruff, S., S. Worley, S. Lubker, Z. Ji, E. Freeman, D. Berry, P. Brohan, E. Kent, D. Reynolds, S. Smith, C. Wilkinson, 2009: ICOADS Release 2.5: Extensions and Enhancements to the Surface Marine Meteorological Archive. *Int. J. Climatol.* (submitted).
- Woodruff, S., E. Freeman, C. Wilkinson, et al., 2009: ICOADS Marine Data Rescue: Status and Future CDMP Priorities. Online document at <http://icoads.noaa.gov/reclaim/pdf/marine-data-rescue_v15.pdf>.