2011

Recovery of logbooks and international marine data: the RECLAIM project

Clive Wilkinson
University of East Anglia, c.w.wilkinson@uea.ac.uk

Scott D. Woodruff
NOAA Earth System Research Laboratory, Scott.D.Woodruff@noaa.gov

Philip Brohan
Met Office Hadley Centre

Stefan Claesson
University of New Hampshire - Main Campus

Eric Freeman
Sourcecorp

See next page for additional authors

INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **31**: 968–979 (2011)
Published online 2 March 2010 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2102

RMetS
Royal Meteorological Society

Recovery of logbooks and international marine data: the RECLAIM project

Clive Wilkinson,[a,b]* Scott D. Woodruff,[c†] Philip Brohan,[d‡] Stefan Claesson,[e] Eric Freeman,[a,f†] Frits Koek,[g] Sandra J. Lubker,[c†] Catherine Marzin[h†] and Dennis Wheeler[i]

[a] *NOAA Climate Database Modernization Program (CDMP), Asheville, NC 28801, USA*
[b] *Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK*
[c] *NOAA Earth System Research Laboratory, Boulder CO 80305, USA*
[d] *Met Office Hadley Centre, Exeter, Devon EX1 3PB, UK*
[e] *University of New Hampshire, Ocean Process Analysis Laboratory, Durham NH 03824, USA*
[f] *Sourcecorp, Dallas, TX 75204, USA*
[g] *Royal Netherlands Meteorological Institute (KNMI), de Bilt, Netherlands*
[h] *NOAA Office of the National Marine Sanctuaries, Silver Spring, MD 20910-3278, USA*
[i] *University of Sunderland, Chester Road, Sunderland SR1 3SD, UK*

ABSTRACT: The RECovery of Logbooks and International Marine data (RECLAIM) project is a concerted, international effort to facilitate and encourage the recovery – through imaging and digitisation – of archived marine weather observations, platform and instrumental metadata and historical documentation, from many different countries. Non-instrumental observations of wind and weather have been recorded in ships' logbooks for hundreds of years, augmented by systematic instrumental observations of sea surface and air temperatures, barometric pressure and other meteorological (and oceanographic) elements largely since the mid-nineteenth century. Once digitised, these data are widely useful for climate studies and other avenues of scientific research – including oceanography, fisheries, maritime history and ecology – thus improvements seeking to address gaps and weaknesses in the currently available data record can be of major importance. In addition to documenting and prioritising archived ship logbook records for imaging and digitisation in support of that goal, RECLAIM provides expert assistance in the interpretation of historical marine records in relation to navigation, observational practices and recording. Since the project inception in 2005, RECLAIM has facilitated, for example, recovery of logbooks of the Dutch and English East India Companies in the nineteenth century, and of the British Royal Navy of the twentieth century. Currently the project website includes reference lists, rescued UK and US marine documents, extensive UK archive reports and detailed inventories of ship movements. A variety of other developing international linkages are discussed, including to Chilean, Dutch, French, German, Finnish, Norwegian, Swedish and US historical records. Copyright © 2010 Royal Meteorological Society and Crown Copyright.

KEY WORDS marine meteorological data; ship's logbooks; archives; data rescue; digitisation; metadata

Received 2 July 2009; Accepted 10 January 2010

> The ancients long since discovered that papyrus and parchment were destructible and that burned clay tablets or bricks were the most enduring of all. We have learned that if the light of the sun and the oxygen of the air are shut out we may preserve the very best of our tender paper records, in properly built libraries, for only a few centuries. They are worthy all the care that we can bestow upon them; it is a sad sight to behold unique, invaluable records, inscribed on the poorest paper, crumbling to dust under the influence of mildew, sunlight, and the noxious vapors given off by our gas lights. They should be studied from many points of view, indexed, and summarized before they disappear. – C. A. (Anon. 1906)

1. Background

The coordinated and systematic collection of instrumental shipboard marine meteorological observations began with the 1853 Brussels Maritime Conference (Maury, 1854; JCOMM, 2004). That landmark conference provided the initial international framework for reporting ships' observations, which has evolved into the contemporary voluntary observing ship (VOS) program. Historically, meteorological observations were extracted from ships' logbooks for climatological applications, such as wind and current charts (e.g. Maury, 1854; see also Lewis, 1996) and meteorological atlases (e.g. Bartholomew and Herbertson, 1899). These applications eventually led to

* Correspondence to: Clive Wilkinson, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.
E-mail: c.w.wilkinson@uea.ac.uk
† The contributions of scott D. Woodruff, Eric Freeman, SJ Lubker and C Marzin to this article were prepared as part of their official duties as United States Federal Government employees.
‡ The contributions of P Brohan was written in the course of his employment at the Met Office, UK and is published with the permission of the Controller of HMSO and the Queen's Printer for Scottland.

national and international collections of digitised weather observations, with the coverage and quantity of observations generally increasing from the mid-nineteenth century (Woodruff *et al*., 2005). These collections have proved invaluable for the assessment and modelling of climate change.

There are, however, many spatial and temporal gaps in the observational collections, especially in infrequently travelled regions such as the Arctic and Southern Oceans, and around the two world wars, and in the nineteenth century and earlier. Indeed, very few observations are available from before 1850. These gaps impose a severe restriction on the use of the records. A secondary problem is that, even when observations of high intrinsic quality have been observed, they are sometimes of poor quality in their existing digital form because of early technological limitations such as punched cards, and collections digitised decades ago can be imperfect or incomplete (e.g. lacking numerical resolution and originally reported metadata) (see Woodruff *et al*., 2005).

The Netherlands, United Kingdom, United States and other national archives hold many thousands of untapped ships' logbooks, which record voyages over most of the world's oceans and cover at least the last 300 years. Many of these logbooks contain meteorological and oceanographic observations – though the earliest records are mostly non-instrumental estimates of the weather and wind conditions but could, nevertheless, dramatically improve the observational coverage. However, to make use of these records, the many thousands of logbooks containing weather observations must be identified, catalogued and prioritised. Next, the selected observations need to be digitised, homogenised into modern scientific units in uniform formats and ultimately blended into widely used climate databases including the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) (Worley *et al*., 2005; Woodruff *et al*., 2010). These are very expensive activities.

Paper historical documents such as logbooks (and obsolete computer media such as punched card decks) can consume expensive physical space in archives, and therefore occasionally become subject to ill-advised destruction due to cost considerations. Examples include Manila Galleon logbooks starting in the sixteenth century (Burt, 1990), '...the huge mass of log books collected by Matthew F. Maury between 1840 and l860, which were destroyed as waste paper about 1869 or 1870' (Anon., 1906), and US merchant World War II logbooks apparently destroyed in the 1970s (Elms *et al*., 1993). Under different circumstances, many historical (1854–1938) Dutch logbooks (*ca* 20 K) were apparently moved to Germany at the end of World War II and are now presumed lost (Wallbrink *et al*., 2003).

However, the irreplaceable nature and great value of these historical data for climate studies, as well as for many other avenues of scientific research, is increasingly being recognised (e.g. Kent *et al*., 2007a; Charpentier *et al*., 2008), and several projects have been undertaken in support of rescuing marine meteorological data (e.g. Elms *et al*., 1993, Mierzejewska *et al*., 1997; Diaz *et al*., 1999; Manabe, 1999; Woodruff *et al*., 2004; WMO, 2008). But much remains to be done, and the RECLAIM project is a concerted, international effort to facilitate and encourage the recovery of archived marine weather observations, platform and instrumental metadata and historical documentation wherever they are to be found.

2. The RECLAIM project

A pioneering effort in the recovery of historical marine weather observations was made by the European Union funded 'CLIWOC' project (García-Herrera *et al*., 2005a), which succeeded in digitising and interpreting approximately 300 K wind and other observations from the period 1750–1850, from British, Dutch (e.g. Figure 1), French and Spanish logbooks, and in doing so demonstrated their value for climate and historical research. To build on the success of CLIWOC, the RECovery of Logbooks and International Marine (RECLAIM) data project was established in 2005 (http://icoads.noaa.gov/reclaim/).

Following completion of the CLIWOC project (2001–2003), a meeting was hosted in 2004 by the NOAA National Climatic Data Center (NCDC) to develop new plans for digitising additional logbooks from Europe, the United States and other countries, and to build on the results and knowledge gained during CLIWOC. This meeting – which led to the establishment of the RECLAIM project – concluded that the largest known undigitised logbook holdings reside in the UK archives, with smaller, but still significant, amounts in Dutch, French, German and other European archives. Considerable amounts of undigitised data were also identified in US archives, including 'Simultaneous' [Greenwich Mean Noon; (GMN)] marine observations starting in the late nineteenth century and US Navy logbooks back to 1801.

However, the marine archives of other maritime nations both large and small are also of interest: in particular, those merchant or military navies that operated in regions presently under-represented in the historical climate record. Most recently, representatives of RECLAIM have investigated the marine archives of the Chilean Navy in Valparaiso through collaboration with the Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative (http://www.met-acre.org/).

One of the main objectives of RECLAIM is to assist in the organisation of new international projects for imaging and digitising data, and collating any available platform and instrumental metadata, from historical ships' logbooks and other marine sources. Three major steps are typically involved in preparing data and metadata for research applications, and ultimately for blending into ICOADS and other climate archives:

- *Imaging*: The photographing of an original hard-copy document, with the results output into recognised archival digital image formats (e.g. PDF and TIFF). This step not only helps in preserving the originals,

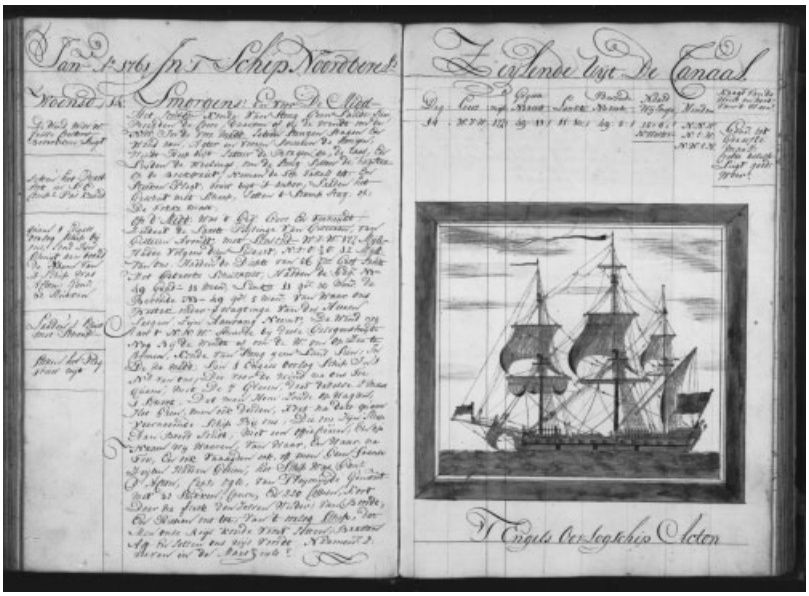

Figure 1. Typical page from an eighteenth century Dutch logbook. This example differs only in the inclusion of a pen-and-ink sketch of an encounter with a British vessel. *By courtesy of National Archief, The Netherlands*.

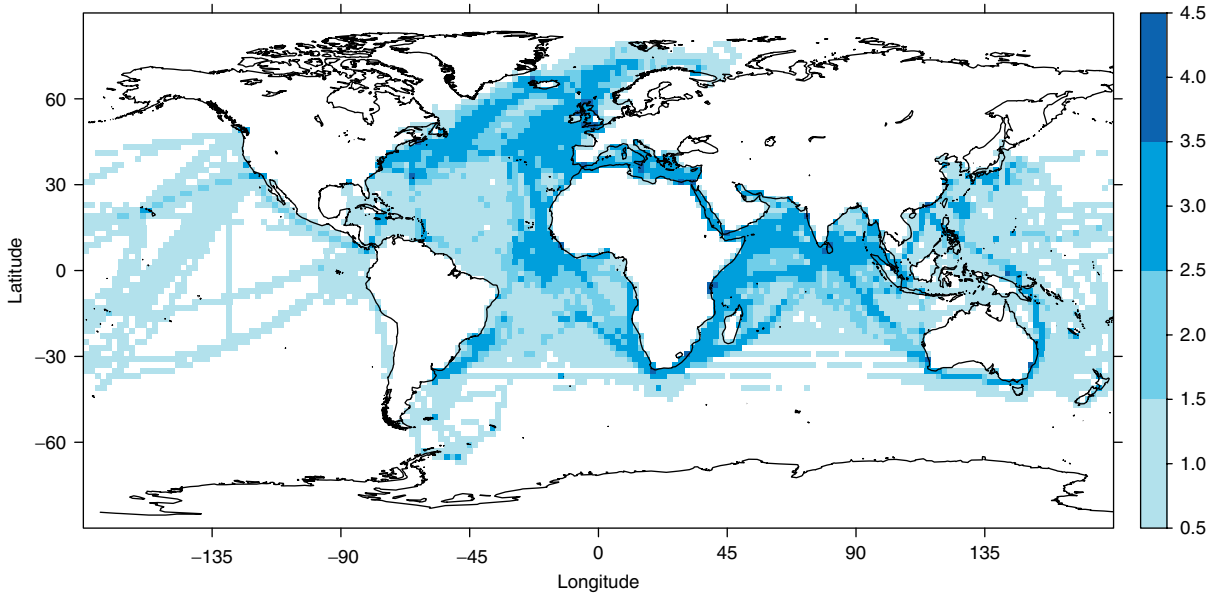

Figure 2. Distribution of the observations in the World War II UK Royal Navy ship's logs, the map shows the mean number of observations per month classified into five categories (darker areas are more frequently observed). The distribution of observations is determined by the distribution of RN activity in the period, except that the majority of the ships operating in the North Atlantic, Mediterranean and UK home waters were not digitised.

assuring that deteriorating paper or other media such as microfilm need be manipulated only once, but also makes it possible to view and use the records remotely from the archive where the originals are kept.

- *Digitisation*: This means the extraction of the data deemed most relevant from the originals or images, and their entry in a computer-readable format (usually ASCII text). This requires manual key entry (and/or sometimes automated optical character recognition). Difficult cost-benefit decisions must often be made on the scope of information to be digitised. For example, many older ships' logbooks contain 'remarks' (e.g. on employment of the crew and detailed navigation information) not directly connected to the coincident meteorological or oceanographic observations, but nevertheless of potential interest to historians and other non-climatic research applications – but in many cases digitisation projects for climate research have omitted these for cost reasons (in this case, however, having the above images readily available can partly satisfy requirements from other disciplines).
- *Translation*: The digitised computer-readable records must be converted into standardised measurement units (e.g. by conversion of barometer measurements from

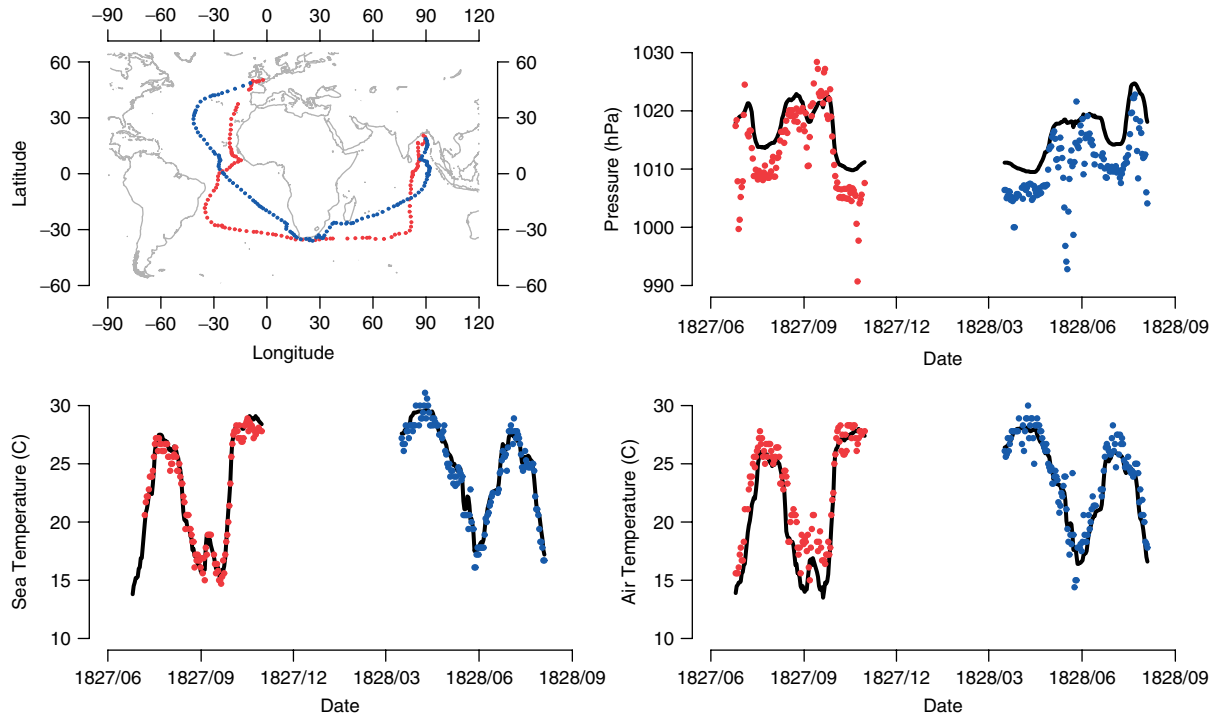

Figure 3. Observations from the outward (red) and return (blue) voyage of the East-Indiaman *Thomas Grenville* in 1827–1828. Coloured points are the daily (noon) positions and weather observations recorded in the logbook, black lines show the expected values from 1961 to 1990 climatologies for a ship following the same route at the same time of year.

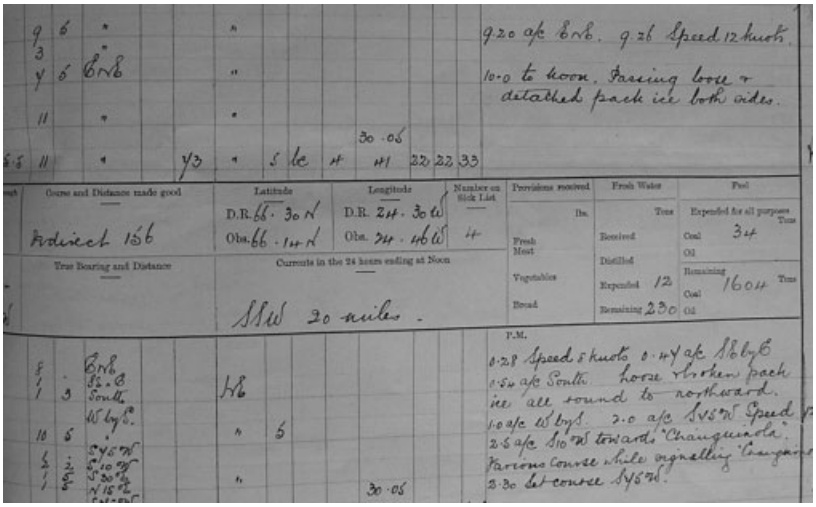

Figure 4. Section of the logbook of HMS *Ebro*, armed merchant cruiser, March 1916: '10.0 (10 a.m.) to noon. passing loose and detached pack ice both sides ... 0.54 (12 : 54 p.m.) ... loose and broken pack ice all round to northward'. Observed position at noon 66°.14′N 24°.46′W (Denmark Strait). *Image (ADM 53/40533) reproduced by permission of the UK National Archives*.

inches of mercury into hectopascals, and application of temperature and gravity corrections; or by conversion of wind speeds from Beaufort force to metres per second) and output in a standard archival format. For ICOADS, this is the International Maritime Meteorological Archive (IMMA; Woodruff, 2007) format. A further problem is posed by non-instrumental records, particularly in their frequent use of wind force descriptions that do not conform to later Beaufort standards, much less modern estimated wind conventions. CLIWOC resolved many such issues of translation from archaic to modern terms (García-Herrera *et al*., 2005b).

The assistance provided by RECLAIM consists of documenting archive holdings, prioritising those holdings and (if required) pre-selecting the best and most suitable material for imaging and digitisation. This pre-selection is aimed primarily at filling temporal and geographic data gaps in the historical climate record.

RECLAIM also provides expert assistance in the interpretation of historical marine records, in relation to

navigation, observational practices and recording. The RECLAIM website is extensively populated with reference lists, rescued marine publications and documents (including US and UK original card deck reference manuals), inventories of ship movements, associated platform and instrumental metadata and manuscript references. In addition, there is an extensive compilation of (mostly) nineteenth-century literature on marine meteorology and oceanography primarily associated with printed journals and accounts of voyages, hydrographic surveys and exploration. The RECLAIM website also includes reports of visits to archives, where the most recent finds are documented and commented upon. Eventually, there are also documentation and ancillary information associated with the projects that RECLAIM has directly assisted.

3. The NOAA Climate Database Modernization program

The Climate Database Modernization Program (CDMP), a partnership between NCDC and private industry (Dupigny-Giroux *et al*., 2007), has since its inception in 2000 provided significant funding and infrastructural resources for the general tasks of recovering, imaging and/or digitising (and translating) climate data records – and has provided substantial support for the RECLAIM project. Tasks to date have included many important recoveries of marine data and metadata, some of which could be subject to deterioration or loss in their original paper form. Recently, the program also began expanding its international connections, and, subject to the constraints of the NOAA budget, it is hoped that a reasonable level of activity in the marine area can continue to be funded. CDMP's contractor-developed image storage software – the Environmental Document Access and Display System (EDADS) (contact Cynthia.B.Karl@noaa.gov for access information) – provides web-based access to a large digital image collection (encompassing marine, terrestrial and other data) for US government employees and their contractors, educational institutions engaged in environmental research and other researchers associated with NOAA projects. In the marine area, the images include US and international logbooks and forms, observing handbooks and legacy documentation.

4. International data recovery projects currently supported by RECLAIM

4.1. Netherlands data rescue activities

In 2006, the Royal Netherlands Meteorological Institute (KNMI) produced approximately 15 K digital images of a collection of Dutch logbooks from the nineteenth century that resides in the archives of the institute. These images are planned for digitisation by CDMP, as resources permit. Support in the form of interpretation of the images is given by KNMI.

Under the umbrella of the KNMI project HIStorical CLIMate (HISKLIM), more inventories of national archives and museums will be completed and more historic marine data will be made available (http://www.knmi.nl/onderzk/hisklim/index-en.html). Surveys will be conducted to find original logbooks or observations that were used to produce several Dutch atlases and publications in the nineteenth century (e.g. Van der Stok, 1897). These surveys are in addition to the inventory of the Dutch part of CLIWOC that only concentrated on the period 1750–1850. As resources permit, an investigation into the availability of early Dutch East India Company logbooks (pre-1750) will be carried out.

4.2. UK Royal Navy ship's logs

In 2007, about 10 K logbooks from some 300 Royal Navy (RN) warships were imaged by CDMP and the Met Office Hadley Centre. These 'Ship's Logs' are from the collection of the UK National Archives, and cover the otherwise data sparse World War II period 1938–1947. They include more than 1.5 million weather observations from both surface vessels and submarines (Figure 2). In this cooperative UK–US project, selected logbooks were imaged by the UK partners and digitised by CDMP. The Hadley Centre implemented a translation into the IMMA format (Woodruff, 2007), and the data were blended into ICOADS Release 2.5 (Woodruff *et al*., 2010). This improved the records of pressure, air temperature and sea surface temperature (SST) for the period, especially in the Indian Ocean and at high northern latitudes in the Atlantic basin (Brohan *et al*., 2009). Advice from RECLAIM was used to select the 10 K logbooks to be digitised from the more than 30 K available. As the North Atlantic, Mediterranean and UK home waters were already well represented for this period, it was important that the logbooks from ships travelling to more remote parts of the world were selected preferentially.

4.3. English East India Company instrumental observations

In 2008, over 900 logbooks of the English East India Company (EIC) for the period 1790–1834 were imaged by the Met Office Hadley Centre. These were a selection from the British Library's collection of about 2 K EIC logbooks of Company ships sailing from Europe to India and China in this period. The main goal in digitising the EIC logs is to capture all available noon observations: location; instrumental observations of pressure and air temperature; and visual estimates of winds, state of the sea and state of the weather. The project aims to extract approximately 285 K daily observations, which should be available by early 2010 (for cost reasons, many sub-daily observations of the visual data and other detailed remarks could not be captured for this project). The data are being prepared and translated for blending into ICOADS by CDMP and the Hadley Centre.

The weather observations in these logbooks are particularly valuable as they are amongst the earliest large,

systematic collections of marine instrumental weather observations known, and cover the climatically interesting period of the Tambora (1815) eruption and the Dalton Minimum of solar activity between 1810 and 1820. The routes of the ships passed through the Atlantic and Indian Oceans around the Cape of Good Hope before sailing north along the meridians of Madras, or Calcutta or towards Indonesia, allowing them to provide information on the climate of a wide area of tropical and sub-tropical ocean and their geographic coverage is therefore extensive and unique for the age (Figure 3).

Along with the standard weather elements, a few EIC logbooks contain rare observations such as sympiesometric pressure and sea temperature measurements that are also being captured. Recordings of the sympiesometer, an early marine barometer made with coloured almond oil and hydrogen gas, were often taken in tandem with the measurements from a mercury barometer, so that comparative studies may be possible. Also, at least one ship, the *Thomas Coutts*, carried three different barometers and recorded simultaneous observations from them all. As they also recorded the make of each instrument, it should be possible to estimate the observational errors and biases in these pioneering measurements. There are also unique observer notations, such as temperatures recorded in degrees and minutes rather than the modern decimal degrees format [e.g. 71°30′ (EIC) = 71.5°(contemporary)].

With the EIC images now permanently preserved and made available to researchers, advantage can be taken of the additional information beneficial to disciplines outside of meteorology including, but not limited to, oceanographic data (i.e. currents and ocean depth/composition) and biological data (i.e. animal sightings).

4.4. UK Royal Navy World War I ship's logs

In 2008–2009, RECLAIM identified and selected 4.7 K RN ship's logbooks for the 1914–1923 period in the UK National Archives and these have been imaged by the Hadley Centre. As for the World War II digitisation, geographic coverage is global though less satisfactory for the Pacific. There is much valuable ice data from the activities of the Northern Patrol during the war period, which every spring would report on ice conditions in the Denmark Strait and the sea passages between Iceland, the Faroes and the Shetland Islands. This was to determine when the passages were ice free and in need of regular patrolling (Figure 4).

As with the World War II ship's logbooks, these records, if digitised, would help fill a data sparse period in the climate record. Moreover, as they follow a very similar format, digitisation by CDMP is relatively efficient. Therefore, we anticipate that digitisation could be proposed as an important priority starting in 2010, should sufficient CDMP resources become available.

4.5. UK colonial registers and Royal Navy logbooks

RECLAIM is also working with the colonial registers and Royal Navy logbooks (CORRAL) project (http://www.corral.org.uk) to image logbooks and other records, from RN survey and exploration voyages, mostly in the nineteenth century, again in the collection of the UK National Archives. Many of these images are now online at the British Atmospheric Data Centre (http://badc.nerc.ac.uk/data/corral/), and digitisation of selected weather observations is in progress. Observations from these vessels are of particular importance due to the work they were engaged in. Officers commanding these types of expeditions often had an interest in marine science, in its earliest manifestations, and were meticulous observers and recorders. This set includes images from the logbooks and journals kept by Cook, Bligh, Ross, Parry, Franklin and Flinders as well as a host of other commanders less well known. The ships were also travelling waters less frequented, often in the Pacific or at high latitudes, where such early meteorological and oceanographic observations are of particular value.

4.6. UK Royal Navy Met logs

An example of records already digitised, but of limited quality, are those of the RN meteorological logbooks ('Met Logs') in the UK Met Office's National Meteorological Archive. Many of the Met Logs are from the same RN ships whose logbooks are in the UK National Archives, but the Met Logs provide much more detailed and comprehensive weather records often including barograph charts and records of pilot balloon ascents or temperatures at altitude recorded from aircraft. They were also apparently more rigorously observed and thoroughly documented. Checks against ICOADS have shown that some of these observations have already been digitised, but that much earlier digitisation process only included a subset of the observations in the paper logs, and did not capture the metadata (e.g. ship names, instrument positions and calibration details) that were also recorded. Consequently re-digitising information from these logs would add substantially to the quantity and quality of the observational record. Unfortunately, because imaging and digitisation are costly, it can be difficult to obtain resources to rehabilitate already-digital, as opposed to entirely new, collections. Nevertheless, rectifying omissions and biases in existing digital collections can be as important scientifically, and better methods of assessing the value of digitising data and metadata are required.

5. Finding and cataloguing additional sources of undigitised logbooks

RECLAIM has also assisted in the identification and documentation of further data-rich sets of logbooks and other manuscripts for future imaging and digitisation projects. One of these exercises was the documentation (funded by the Met Office Hadley Centre) of *ca* 6 K 'remark books' (1820–1909) held by the UK Hydrographic Office (UKHO). The bulk of remark books were kept on vessels engaged in exploration and hydrographic survey and they are most often associated with the development of

marine charting and coastal mapping, but their scope goes well beyond this. Almost all of the remark books after 1850, and a modest number before that date, contain frequent, usually sub-daily, observations of air pressure and temperature, and SST. Along with these are observations of wind direction, wind force and notes on ocean currents. Humidity and the specific gravity (density) of seawater were less frequently recorded. These remark books therefore constitute an important source of early meteorological and oceanographic data. Moreover, a set of 'daily remark books' (1866–1875) is catalogued among the records of the US Hydrographic Office at the National Archives and Records Administration (NARA), and whether similar books might exist within hydrographic archives of other nations remains to be explored.

Once completed, a remark book was sent to the Admiralty along with the various journals of a voyage kept by the captain and master (or navigating officer). The officers' and ships' journals along with the remark books constitute the complete record of a voyage, and as such would have been archived and consulted together. However, the remark books became part of the UKHO Archive, and the journals or logbooks eventually left the Admiralty to be housed at the UK National Archives. The separation of these collections was logical, given the importance of the remark books in the compilation of charts and sailing directions. However, this division has resulted in the relative neglect of the remark books by the majority of maritime and naval historians, except for those with an interest in hydrographic work and exploration. Consequently, their 'rediscovery' as an important source of scientific data is especially fortuitous (Wilkinson *et al*., 2008).

A further potential source of historic marine climate data has been identified through a joint ACRE/RECLAIM initiative. This involved a preliminary investigation of the Chilean Naval Museum archive in Valparaiso and the National Archive in Santiago in May 2009. Information on this collection awaits detailed reports and additional investigation. However, 6 K logbooks are held at the Naval Museum, the earliest of which date from the 1860s. The significance of this collection rests largely on the relative scarcity of historic data for the southeast Pacific, the proximity of western Antarctica and the potential for detecting more precise signals of past El Niño events. In anticipation of a data recovery project based on the Chilean maritime collections, RECLAIM has also compiled corresponding documentation on the ship movements and logbooks of all RN vessels in the eastern Pacific in the nineteenth century.

Future UK projects will include logbooks from both the RN and whaling ships operating in high latitudes. A number of UK archives hold small collections of whaling logbooks (the largest as described in Section 7), the activities of these vessels were mostly concentrated on the Davis Strait and the seas off Greenland in the nineteenth century. In addition to these, there are the logbooks of the Hudson's Bay Company, held on microfilm by the UK National Archives. The earliest of these logbooks is 1751 and the last 1870. They provide valuable data on ice conditions in the Hudson Strait and in Hudson's Bay, along with the recording of wind vectors and weather conditions.

Additional projects will also attend to the extensive collection of nineteenth and twentieth century Met Logs held in the UK Met Office's National Meteorological Archive. For the nineteenth century, there are nearly 7 K Met Logs, of which 600 are from Royal Navy vessels and the remainder from merchant ships (Rhodes, 1995). The registers consist of detailed tabulated data, recorded on pre-printed pages at 2-h intervals. Importantly, the registers contained much metadata concerning the instruments, usually type and number, position, height above sea level, etc. The necessary correction to be applied to the observations was also recorded.

6. Observation metadata and non-logbook records

Other tasks being undertaken by RECLAIM include the documentation of early ship instrument and platform metadata and compendiums of UK and US national ship reporting practices and publications. WMO only began publication of VOS platform and instrumental metadata in 1955 (see Kent *et al*., 2007b), and previously much of this information was not digitised due to restrictions, such as space, imposed by now obsolete technologies (e.g. punched cards). RECLAIM is capturing this information for some existing UK records and documenting the existence of metadata for other records yet to be digitised.

For example, RN Met Logs and many of the late nineteenth century Ship's Logs contain detailed instrumental metadata, recording the maker and type of instrument, its height above sea level and whether screened. Similar metadata may be found in logbooks from many other nations (e.g. Chilean Navy), and there are a number of modern publications containing the details of warships of most Navies. Lloyd's Register (1997) can provide metadata for some international merchant vessels back to 1764. However, such 'ship particulars' (e.g. ship numbers, tonnages, dimensions and superstructures) from commercial organisations may need to be kept confidential, and how to make effective and appropriate use of them is an ongoing subject of discussion with JCOMM and the International Maritime Organization (IMO). Moreover, to facilitate utilisation for scientific research, much of this important information needs to be matched and blended with the corresponding marine data observational records in databases such as ICOADS.

A further rich and surprisingly accessible, though underused, source of historic marine data is to be found in printed accounts of journals and voyages. These volumes often have appendices containing abstract logs or meteorological registers, usually in a tabulated format. This lends itself to capture by optical character recognition (OCR) software. Furthermore, some of these printed accounts contain valuable additional information that is not found in the original manuscript logbooks and

journals. For instance, the description of the fitting out of an expedition might mention the types of scientific instruments carried. Sometimes there are detailed and useful descriptions of observing practices and methods of recording. Much of this material has the added advantage of being available in electronic (usually PDF) form from online sources such as Google Books or the Internet Archive. RECLAIM is documenting the data content of many of these publications, providing details of at least one archive or library where they can be found and the availability of electronic copies. Additionally, RECLAIM is working through the ACRE initiative in documenting which of these publications have had their data captured and incorporated into existing climate data sets.

7. Additional future priorities for RECLAIM and ICOADS

Figure 5 illustrates a number of important new marine data collections linked to RECLAIM and targeted for recovery and future blending into ICOADS as resources permit. Some of these datasets were discussed above, with the others reviewed briefly as follows:

7.1. UK Hull Whaling logbook collection (1790–1850)

The city of Kingston-upon-Hull holds nearly 200 logbooks of whaling vessels providing important (non-instrumental) observations for the far northern Atlantic latitudes. The Universities of Sunderland and of Hull are currently seeking means of abstracting and digitising these data. This appears to be the largest single UK whaling collection.

7.2. German Maury collection (1845–1867)

This collection was loaned to NCDC by the Deutscher Wetterdienst (DWD) for imaging (Braun, 2000), and digitised by CDMP. Translation into IMMA format will likely require attention to issues of data homogeneity with the US Maury Collection (Woodruff *et al*., 2005), and characteristics of the data (e.g. biased sea level pressure observations; Ansell *et al*., 2006) also being studied by KNMI (Wallbrink, 2009).

7.3. DWD historical archive (1850–1939)

Historical marine digitisation activities at DWD began in the 1940s and are continuing in the framework of DWD's 'HISTOR' project (Zöllner R, personal correspondence). Starting in 1999, several high-priority selections from the DWD historical marine archive have been made available so as to help enrich ICOADS and thereby advance historical reanalyses and other urgent climate research applications, and we are hopeful that additional archive data can be made available in due course.

7.4. Swedish (1860–1922) and Finnish (1900–1916, 1919–1957) lightvessel data

Logbooks were imaged from the US National Oceanographic Data Center (NODC), Swedish Meteorological and Hydrographic Institute (SMHI), and Finland Institute of Marine Research (FIMR). A combination of oceanographic (e.g. temperature, salinity, and current speed and direction) and meteorological (e.g. air temperature, wind and barometric pressure) data was recently digitised from the collections. The Swedish data were collected in the Skagerrak, Kattegat, Baltic Sea and Gulf of Bothnia by 14 Swedish Lightvessels, and the Finnish data in the Baltic Sea, Gulf of Finland and Gulf of Bothnia by 32 lightvessels.

7.5. Arctic Norwegian logbook data (1867–1912)

This set extends past the period (1867–1899) of the 600 logbooks from which the Norwegian Logbook Collection included in ICOADS Release 2.0 was extracted

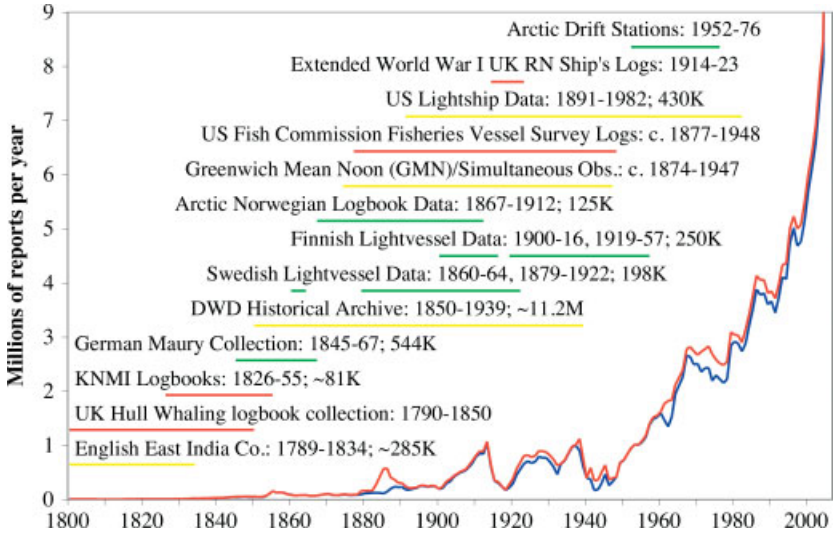

Figure 5. The time periods of candidate historical data sources connected with RECLAIM, and to be blended into ICOADS, are spanned by horizontal coloured lines: green candidates are fully digitised but require format translation, yellow are partially digitised, and red are in the planning stages for digitisation. Each dataset name is appended with the date range and approximate number of reports (if known). The solid blue curve is the number of reports in ICOADS Release 2.4; the solid red curve is the number in Release 2.5.

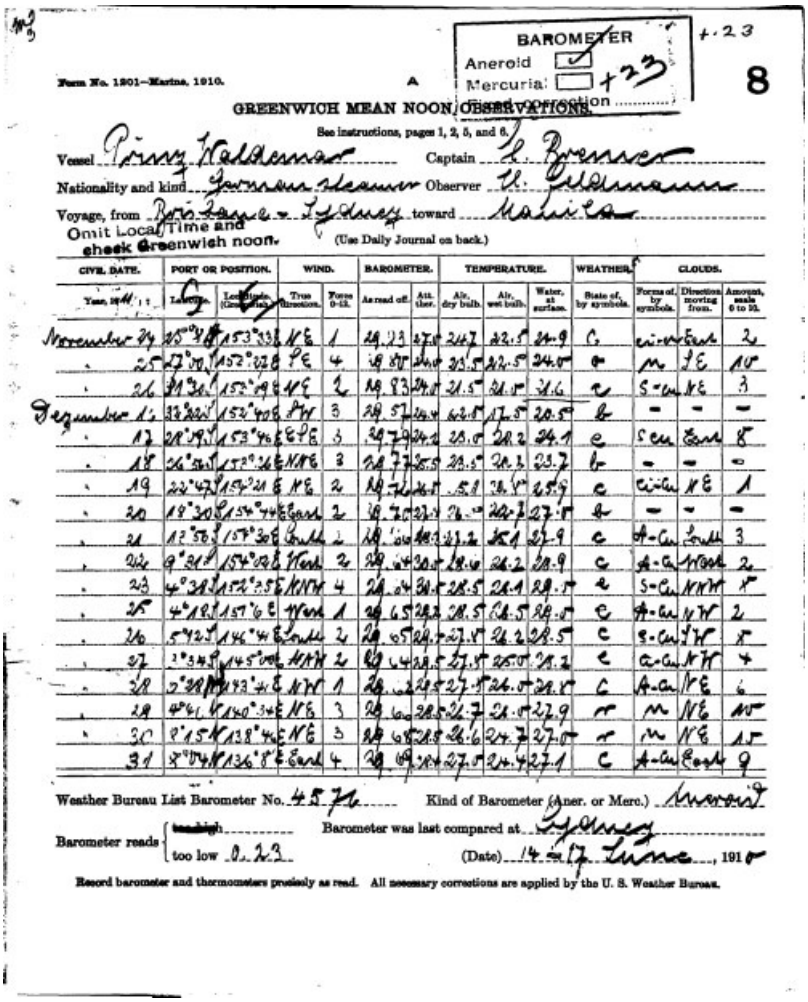

Figure 6. Daily Greenwich Mean Noon (GMN) Observations (US Weather Bureau Form 1201 – Marine, 1910) from the German Steamer *Prinz Waldmar*, 1911. This example form includes stamps posted by Weather Bureau after receipt for barometer type and corrections, and a note to check GMN time. The observations were to be taken at the local time corresponding to GMN. As a time was not entered on this form, it is assumed that the time corresponds accurately to GMN. Stamps on other forms include 'Data on this sheet transferred to punch card' or 'Not Used for Punch Cards'. In the case that the data were transferred to punch cards, it is not clear if these punch cards exist and the possibility remains that they could have been destroyed as with other original records mentioned in Section 1.

(Woodruff *et al*., 2005). However, only SST and air temperature data were keyed. Moreover, that earlier project only succeeded in keying about one-third (200 logbooks) of that collection, so much more potential to rescue additional historical data still appears possible.

7.6. Greenwich Mean Noon (GMN)/simultaneous obs. (*ca* 1874–1947)

Around 1888, the United States started switching from systematic observations made throughout the day in Marine Meteorological Journals to 'simultaneous' observations taken once daily worldwide at GMN (i.e. 12 : 00 UTC) (see Woodruff *et al*., 2005). CDMP work is underway to image and digitise different components of this very complex overall collection, beginning with monthly and daily marine forms sent to the US Weather Bureau covering 1910–1947 (e.g. Figure 6). The forms have been imaged and added to CDMP's EDADS database and digitisation of the data is underway. These data should be available in 2010 and will be translated into IMMA format for inclusion in a future ICOADS update. Another source of original paper forms exists for the period 1886–1902 and is under investigation for potential digitisation.

A variety of published (or tabulated) observations in what was originally titled the *Bulletin of International Meteorological Observations* (e.g. US Signal Service 1883, Figure 7) also form part of the collection (see Anon., 1914, also noting a 'valued supplement in the *Tägliche synoptische Wetterkarte des Nordatlantischen Ozeans*, issued jointly by the Deutsche Seewarte and the Danske Meteorologiske Institut beginning with 1884'). The 1875–1889 Signal Service volumes include both terrestrial and marine observations. Daily observations were included until July 1884. Monthly summaries were produced from July 1880 to December 1887, and semi-annual publications continued thereafter through 1889 (NOAA, 1978). The semi-annual publications have yet to be recovered, but CDMP has imaged the entire 1875–1887 daily and monthly collections and is currently

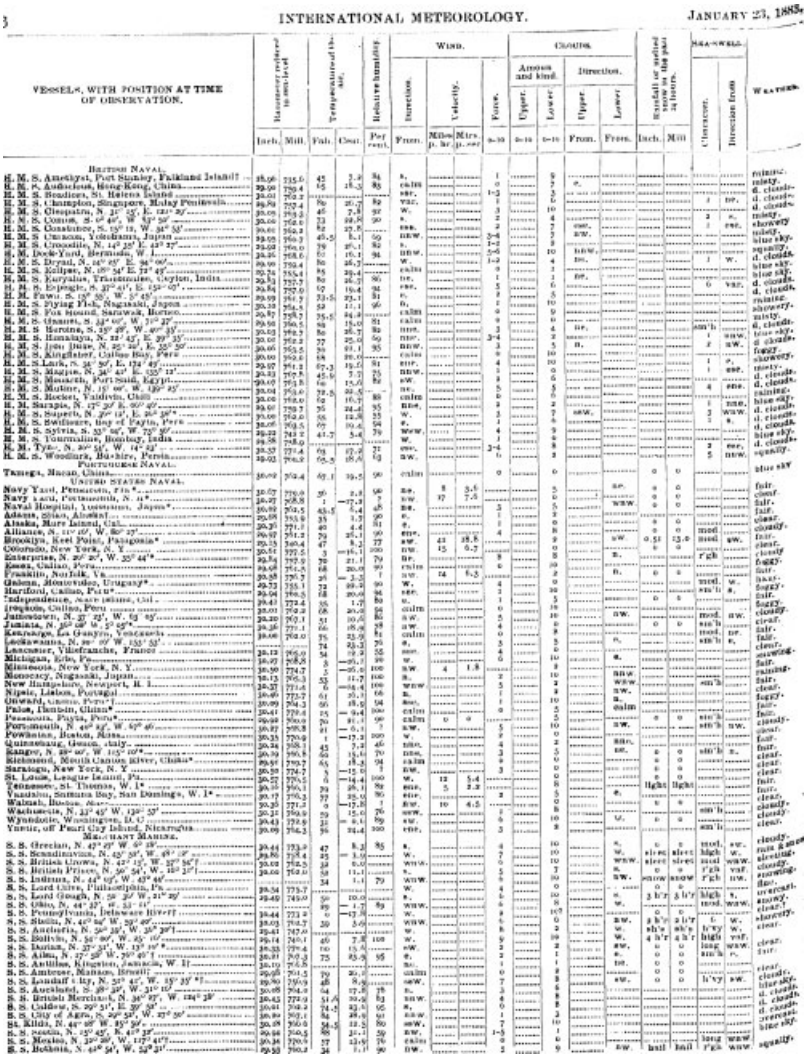

Figure 7. Greenwich Mean Noon Observations printed in the *Bulletin of International Meteorological Observations* (US Signal Service, 1883). The observations are global and, depending on the publication, sorted by contributing country, ship owner or observed section of the globe (North Atlantic, South Pacific, etc.). Note that pressures are included in both inches and millimetres, temperatures in Fahrenheit and Celsius, and winds in miles per hour and metres per second. Wherever possible, elements (i.e. barometer, wind speed) were corrected by the US Weather Bureau and were noted at the foot of every page. There are also many observations while at port with no latitude and longitude recorded (however, future translation work will be able to make use of many such positions as already determined by Brohan *et al*., 2009).

initiating digitisation only of the daily marine observations. An additional source of tabular observations in oversize volumes exists for the period 1874–1892. These are noted to be in poor shape and will need further investigation to determine if the records can be salvaged.

7.7. US Commission of Fish and Fisheries vessel survey logbooks (*ca* 1877–1948)

Original logbooks of the US Commission of Fish and Fisheries and US Bureau of Fisheries survey vessels are archived at the US Smithsonian Institution and at NARA. The environmental variables recorded include water depth, character of the sea floor, weather, air and ocean temperatures, barometric pressure, winds, salinity, currents and species brought back from trawls or dredges, for specific locations. The original forms are being imaged by CDMP, thus laying groundwork for follow-up studies in which the geo-referenced biological, oceanographic and marine meteorological data from these logs will be extracted to blend into ICOADS and into NODC's World Ocean Database Project (http://www.nodc.noaa.gov/General/NODC-dataexch/NODC-godar.html).

7.8. US Lightship data (1891–1982)

This source provides observations from ships anchored around the US coastline, most recently operated by the Coast Guard, as an aid to navigation (Woodruff *et al*., 2005).

7.9. Arctic Drift stations: (1952–1976)

These data were collected and organised by NCDC and the US National Snow and Ice Data Center (NSIDC) (Arctic Climatology Project, 2000). Only some of the data, including from the ships *Maud* and *Fram*, have thus far been blended into ICOADS.

8. Summary and conclusions

This paper has provided an overview of RECLAIM, a cooperative international project to image historical ship logbooks and related marine data and metadata and digitise the meteorological and oceanographic observations for merging into ICOADS and for utilisation for climate research. RECLAIM seeks to identify archives of historical records containing undigitised marine weather observations to catalogue and document those archives, to select high-priority subsets of the archives to assist imaging and digitisation projects and to provide specialist knowledge and advice on the interpretation of the records (platform metadata and observing techniques).

It has already facilitated several large-scale digitisation projects, and characterised additional archives, which are high priorities for digitisation when resources can be found. The metadata records and resources identified by RECLAIM are being made available on the RECLAIM website, which, linked to ICOADS for access to the actual meteorological and oceanographic observations, provides a comprehensive 'one stop' point of reference for global historical marine climate data.

So far, much of the RECLAIM work has been focused on assembling and assessing information about United Kingdom (and to a lesser extent United States) data, metadata and archives. However, one of the original goals from the RECLAIM planning meeting in 2004 was compilation of a detailed documentation of our current knowledge of national logbook holdings, extending eventually beyond the current primary UK and US foci to Danish, Dutch, French, German, Italian, Portuguese, Spanish and other historical archives, and much more work is still needed in this area. For example, French National Archives have a marine hydrographic section (4JJ and 5JJ) that includes logbooks ('journal du bord') from the French voyages of discovery and from scientific expeditions. The archives also contain logbooks of the French East India Company, which ceased trading in 1769.

Despite ongoing digitisation and data rescue activities such as RECLAIM, there remain many undigitised historical ship data, some not widely known outside archives and libraries (including published data). Work continues to actively catalogue, image, digitise and ultimately convert digitised data into the IMMA format. However, these are all expensive tasks, and better methods are needed for prioritising and evaluating the potential of specific collections. This allows for a more strategic approach for different climate applications, as well as for related research disciplines, including oceanography, fisheries, maritime history and ecology.

Acknowledgements

We are grateful to Dr Andrew Cook, India Office, British Library, for information regarding remark books and hydrographic surveys; to Caroline Kimbell, UK National Archives, Kew, for assistance with imaging projects; to Catharine Ward and Fiona Williamson for assistance in the documentation of the marine collections of the British Library India Office and the UK Hydrographic Office; and to Godfrey Trammell, US NODC, for providing the Swedish and Finnish lightvessel data and further information about those collections. This work, including development of the RECLAIM website (hosted at the NOAA Earth System Research Laboratory under the ICOADS project), has been partly supported by the NOAA Climate Database Modernization Program (CDMP). This work was also supported by the Joint DECC, Defra and MoD Integrated Climate Programme – DECC/Defra (GA01101), MoD (CBC/2B/0417_Annex C5).

References

Ansell TJ, Jones PD, Allan RJ, Lister D, Parker DE, Brunet M, Moberg A, Jacobeit J, Brohan P, Rayner NA, Aguilar E, Alexandersson H, Barriendos M, Brandsma T, Cox NJ, Della-Marta PM, Drebs A, Founda D, Gerstengarbe F, Hickey K, Jónsson T, Luterbacher J, Nordli O, Oesterle H, Petrakis M, Philipp A, Rodwell MJ, Saladie O, Sigro J, Slonosky V, Srnec L, Swail V, García-Suárez AM, Tuomenvirta H, Wang X, Wanner H, Werner P, Wheeler D, Xoplaki E. 2006. Daily mean sea level pressure reconstructions for the European – North Atlantic region for the period 1850–2003. *Journal of Climate* **19**: 2717–2742, DOI: 10.1175/JCLI3775.1.

Anonymous ("C.A."). 1906. The indexing of marine reports. *Monthly Weather Review* **34**: 159. [Note: Initials probably refer to Cleveland Abbe, then Editor].

Anonymous ("C.A."). 1914. International simultaneous observations. *Monthly Weather Review* **42**: 675–676. [Note: Initials probably refer to Cleveland Abbe, then Editor].

Arctic Climatology Project. 2000. *Environmental Working Group Arctic meteorology and climate atlas*, Fetterer F, Radionov V (eds). National Snow and Ice Data Center: Boulder, CO (CD-ROM).

Bartholomew JG, Herbertson AJ. 1899. *Atlas of Meteorology: A Series of Over Four Hundred Maps*. Prepared by and edited by Buchan A. Constable: London.

Braun DS. 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†.

Brohan P, Allan R, Freeman JE, Waple AM, Wheeler D, Wilkinson C, Woodruff S. 2009. Marine observations of old weather. *Bulletin of the American Meteorological Society* **90**: 219–230, DOI: 10.1175/2008BAMS2522.1.

Burt WV. 1990. The search for the Manila Galleon log books. *Bulletin of the American Meteorological Society* **71**: 1630–1636.

Charpentier E, Harrison DE, Keeley JR, Kent E, Mietus M, Rayner N, Rutherford M, Swail V, Woodruff S. 2008. Third JCOMM Workshop on Advances in Marine Climatology (CLIMAR-III). *MeteoWorld*. December 2008, WMO: Geneva.

Diaz HF, Woodruff SD (eds.) 1999. *Proceedings of the International Workshop on Digitization and Preparation of Historical Surface Marine Data and Metadata (Toledo, Spain, 15–17 September 1997)*. WMO/TD-No.957, MMROA Report No. 43. World Meteorological Organization: Geneva; 114 pp.††.

Dupigny-Giroux LA, Ross TF, Elms JD, Truesdell R, Doty SR. 2007. NOAA's climate database modernization program: rescuing, archiving, and digitizing history. *Bulletin of the American Meteorological Society* **88**: 1015–1017.

Elms JD, Woodruff SD, Worley SJ, Hanson C. 1993. Digitizing Historical Records for the Comprehensive Ocean-Atmosphere Data Set (COADS). *Earth System Monitor* **4**(2): 4–10††.

García-Herrera R, Jones P, Wheeler D, Können G, Prieto MR (Guest Eds.) 2005a. CLIWOC: Climatology of the World's Oceans, 1750–1850 (Special Issue). *Climatic Change* **73**: 1–194.

García-Herrera R, Wheeler D, Können G, Koek F, Prieto MR. 2005b. *CLIWOC Multilingual Meteorological Dictionary*. Royal Dutch Meteorological Institute (KNMI): publication no.2005: De Bilt, The Netherlands.

JCOMM. 2004. An International Seminar to Celebrate the Brussels Maritime Conference of 1853 – an historical perspective of operational marine meteorology and oceanography: proceedings.

WMO/TD–No. 1226, JCOMM Technical Report No. 27, JCOMM: (CD-ROM).

Kent EC, Woodruff SD, Berry DI. 2007b. WMO Publication No. 47: Metadata and an assessment of observation heights in ICOADS. *Journal of Atmospheric and Oceanic Technology* **24**: 214–234.

Kent E, Woodruff S, Rayner N, Arbetter T, Folland C, Koek F, Parker D, Reynolds R, Saunders R, Smolyanitsky V, Worley S, Yoshida T. 2007a. Advances in the use of historical marine climate data (Second International Workshop on Advances in the Use of Historical Marine Climate Data). *Bulletin of the American Meteorological Society* **88**: 559–564.

Lewis J. 1996. Winds over the World Sea: Maury and Köppen. *Bulletin of the American Meteorological Society* **77**: 935–952.

Lloyd's Register. 1997. *Register of Ships 1997–1998, A–G. Lloyd's Register of British and Foreign Shipping*, London, 1773 pp.

Manabe T. 1999. The digitized Kobe Collection, phase I: historical surface marine meteorological observations in the archive of the Japan Meteorological Agency. *Bulletin of the American Meteorological Society* **80**: 2703–2715.

Maury MF. 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 edn. E. C. and J. Biddle: Philadelphia, pp 54–96†.

Mierzejewska AW, Wu Z, Newell RE, Miyashita T. 1997. Japanese whaling ships' sea surface temperature 1946-84. *Bulletin of the American Meteorological Society* **78**: 443–447.

NOAA (National Oceanic and Atmospheric Administration). 1978. *Catalog of the Atmospheric Sciences Collection in the Library and Information Services Division, Environmental Science Information Center, Environmental Data and Information Service, National Oceanic and Atmospheric Administration–Acquisitions 1890-August 1971*, vol. 3. NOAA: Boston; 330.

Rhodes M. 1995. *Additional Logbook Data 1854-1899*. (Unpublished UK Met Office report)†.

US Signal Service. 1883. *Bulletin of International Meteorological Observations*. Office of the Chief Signal Officer of the Army: Washington, DC.

Van der Stok JP. 1897. *Wind and Weather, Currents, Tides and Tidal Currents in the East Indian Archipelago*. Government Printing Office: Batavia; 209 pp.

Wallbrink, H. 2009. *The U.S. Maury Collection Metadata 1796-1861*. Available from http://www.knmi.nl/onderzk/hisklim/Publicaties/publicaties.html.

Wallbrink H, Koek F, Können GP, Brandsma T. 2003. Sea-level pressure observations from Dutch ships 1854–1938 incorporated in COADS Release 1c climatology. *International Journal of Climatology* **23**: 471–475.

Wilkinson C, Ward C, Williamson F. 2008. *Report on the Royal Navy Remark Books Held at the Hydrographic Office: Taunton UK*. (Unpublished report)†.

WMO. 2008. *Proceedings of the International Workshop on Rescue and Digitization of Climate Records in the Mediterranean Basin (Tarragona, Spain, 28–30 November 2007)*. WCDMP-67 (WMO-TD-1432), World Meteorological Organization: Geneva.

Woodruff SD. 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. [These pages and related format documentation available from: http://icoads.noaa.gov/e-doc/imma/].

Woodruff SD, Diaz HF, Worley SJ, Reynolds RW, Lubker SJ. 2005. Early ship observational data and ICOADS. *Climatic Change* **73**: 169–194.

Woodruff S, Elms J, Reynolds R, Garcia R, Guo F, Worley S, Yoshida T. 2004. Rescuing marine data. *World Climate News No. 25* (June 2004). WMO: Geneva; 10.

Woodruff SD, Worley SJ, Lubker SJ, Zaihua J, Freeman JE, Berry DI, Brohan P, Kent EC, Reynolds RW, Smith SR, Wilkinson C. 2010. *ICOADS Release 2.5 and Data Characteristics* (this volume).

Worley SJ, Woodruff SD, Reynolds RW, Lubker SJ, Lott N. 2005. ICOADS Release 2.1 data and products. *International Journal of Climatology* **25**: 823–842, DOI: 10.1002/joc.1166.

Note: references followed by † are available online at http://icoads.noaa.gov/reclaim/; references followed by †† are available online at: http://icoads.noaa.gov/publications.html or from the underlying selected research bibliography

