Our understanding of climate change and climate variability depends critically on historical weather observations. Studies of recent changes use instrumental observations directly (Trenberth et al. 2007), while studies of longer-term changes make use of climate proxies, which are calibrated to observations (Jansen et al. 2007). The currently available instrumental observations allow some surface climate fields to be reconstructed for the last 150 yr (Fig. 1).

Weather observations, including measurements of temperature and pressure, have been made on ships, in large numbers, back to the late eighteenth century, and are recorded in the logbooks of those vessels. Subjective but reliable estimates of wind force and direction are available even further back (Wheeler 2005). National programs to digitize marine observations from some logbooks (and from later specialized meteorological forms) appear to have started around the turn of the twentieth century. By the middle of the century there were international exchanges of large national collections (Woodruff et al. 1987, 2005), and these developed into today’s international scheme for routine digitization and exchange of delayed-mode voluntary observing ship (VOS) data. In spite of these early efforts, many as yet untapped logbooks still exist in various archives around the world.

Unfortunately, owing to early technological limitations, some of the early observation keying efforts resulted in incomplete and possibly inaccurate data and metadata (Woodruff et al. 2005). This makes new digitization programs even more valuable: if missing or deficient observations were digitized and made available to climate researchers, they might extend our knowledge and understanding of climate change.

Recent work to digitize observations in the U.K. archives has been confined to reconstructing the weather of naval battles (Wheeler 2001), specific
areas (Farrington et al. 1998), large volcanic eruptions (Chenoweth 2001), or specific phenomena, such as the Southern Oscillation (Garcia et al. 2001). The exemplary work of Catchpole (1992) uses ship logbooks of the Hudson’s Bay Company to study ice cover in the nineteenth century across the seas of northern Canada. More comprehensive has been the Climatological Database for the Worlds Oceans, 1750–1850 (CLIWOC) program, which tested the utility of noninstrumental logbook data (principally wind force and direction) in a modern-day scientific context (Wheeler et al. 2006). A database of some 300,000 quality-controlled and processed data are available (online at www.ucm.es/info/cliwoc).

Subsequent to the CLIWOC program, an international project for the Recovery of Logbooks and International Marine Data (RECLAIM; http://icoads.noaa.gov/reclaim) has recently begun to recover some of the vast quantity of logbook data that remains untouched in the archives. This project has begun to identify data worth extracting, and has completed a project to recover observations from Royal Navy logbooks for the period of 1938–47.

**U.K. Observations Available for Digitization.** Historical marine observations are archived in many parts of the world, but particularly in the United Kingdom, where over 300,000 undigitized logbooks, mostly from Royal Navy vessels, are available for the period from 1670 to the late twentieth century (Fig. 2). The near-global extent of the British Empire meant that British ships sailed extensively across the world’s oceans, each making its daily logbook record as it went. While some U.K. marine data have been available in digital form for decades (Woodruff et al. 2005), there has been no systematic program for digitizing more of these observations, and most remained unused.

The majority of the undigitized U.K. logbooks are in the National Archives (Kew), the British Library (central London), and the National Maritime Museum (Greenwich). Until the early nineteenth century each senior ship’s officer of the Royal Navy kept his own record, giving rise to captains’, masters’, and lieutenants’ logbooks; there is a lot of redundancy in these records, because a single ship may produce one captain’s, one master’s, and several lieutenants’ logbooks from each voyage. The general ship logbook became commonplace only as the century advanced. In addition to Royal Navy records, there is also a collection of logbooks of ships in the service of the Honorable East India Company (EIC). These, though fewer in number and dating only from the 1600s to the 1840s (Farrington 1999), are important in their coverage, often instrumental (after about 1780), of the southern and Indian Oceans as they made passage to and from India and China via the Cape of Good Hope. Figure 2 summarizes the distribution of these logbooks by category and decade.

Another significant source of historical observations is the Met Office archives, which contain a large number of meteorological logbooks (“Met logs”), in some cases from the same Royal Navy ships whose logbooks are in the National Archives, but the Met logs are apparently more rigorously observed and thoroughly documented. To a limited extent these Met logs have already been digitized and blended into the International Comprehensive Ocean–Atmosphere Data Set (ICOADS), but that much earlier digitization process only included a subset of the observations in the paper logs, and did not capture the metadata (e.g., instrument positions and calibration details) recorded in them.

Elsewhere in the United Kingdom only small numbers of logbooks are known to exist, and few historical U.K. merchant marine logbooks have survived. Most important among these are the logbooks of whaling vessels from the late eighteenth and early nineteenth centuries. Whaling activities were concentrated along the ice limits around Greenland. As such, while their numbers are small (only about 250), they contain climatic information for the high northern latitudes at a time for which no other sources are available. Most of these documents are to be found in the Kingston-upon-Hull City archives (Wheeler 2007). Attention has been given to this collection of logbooks by the British Arctic Whaling group (infor-
mation online at www.hull.ac.uk/baw/). Similar collections are in the archives of New England whaling ports (Sherman et al. 1986).

In the age of sail, before the mid-nineteenth century, most logbooks contained subjective estimates of wind force and direction, and an example of such a logbook is given in appendix A. These observations, derived by procedures similar to those employed today by the VOS network, can be retrieved and used, as shown by the CLIWOC project (Wheeler et al. 2006). More directly valuable, however, are the later logbooks containing thermometer and barometer observations. Instrumental data became widespread only following the Brussels Maritime Conference of 1853 (Maury 1854). Before that date the East India Company logbooks, which number about 4,000 volumes, provide the only consistent source of marine instrumental data, many of them having daily barometer and thermometer observations going back to before 1800. Appendix B gives examples from an East India Company logbook.

The digitization of all the U.K. data would be a monumental undertaking. The limited resources currently available provide only for a pilot project using logbooks from the period of the Second World War (WWII). This has been completed and demonstrates the potential of the source.

A PILOT PROGRAM TO RETRIEVE WW2 DATA. The period around the Second World War (taken here as 1938–47) is a good candidate for digitization for several reasons: first, there is a shortage of digitized observations; second, during this period there were big changes in the way marine observations were made; and finally, there is evidence of large and sudden changes in the observed temperature record that require substantiation. From a practical point of view, logbooks for this period are easily accessible in the National Archives and are in good condition.

Over the past century, periods of major conflict are marked by deficiencies in the global climate record. Figure 3 shows the drop in global coverage of the existing ICOADS (Worley et al. 2005; Woodruff et al. 2003, 2005, 1987) of marine observations, which falls from 65% in 1937 to 35% in 1941, effectively reducing surface observation coverage to that available from the 1880s. Such shortcomings in the record substantially reduce the precision and reliability of our understanding of regional and global climatic variability and change over this period (Rayner et al. 2006).

The period between 1938 and 1942 saw significant changes in marine observational practices, and the SST datasets used for climate research contain adjustments to remove the resulting temperature biases (see the “SST biases” section). However,
uncertainties remain in the size of the biases, and how they change over time. Figure 1 shows that the current estimates of SST (after adjusting for these instrumental biases) suggest a noticeable increase in the WWII period. This important feature may be a result of natural climate variability, but concerns that it might be at least partly an artifact of uncorrected biases in the SST measurements require that as much additional data as possible for the period be gathered in order to resolve the issue.

Choosing logbooks for digitization. Investigation in the U.K. National Archives revealed around 30,000 Royal Navy ships’ and submarines’ logbooks for the period in question. Each logbook covers a month, and consists of about 30 pages, each detailing 1 day. A typical page contains one or more recordings of the ship’s position and six sets of meteorological observations, including air and sea temperatures and air pressure, taken at different times through the day. Figure 4 shows a typical example. Approximately 8,000 logbooks were photographed (about 250,000 pages). Given that existing data for the period are concentrated on European waters and the North Atlantic, logbooks were selected predominantly from ships undertaking long voyages into the Southern Hemisphere. Moreover, the existing data shortage is most acute for the years from 1941 to 1946, and attention was concentrated on that period, but not to the exclusion of some sampling from 1938, 1939, and 1947. Selection was done by deployment, rather than by individual logbook, so if, for example, a ship was posted from Plymouth into the Indian Ocean for a year, all of her logbooks for that year were selected. Also, all observations from each selected logbook were digitized, not just those from data-sparse areas. This gives a comprehensive set of observations, which is helpful for quality control and allows for the widest possible use of the new data, but does mean that not all observations are in places ideally suited for studies of marine climate; more than half of the new observations are from ships in port.

Imaging and digitizing the logbooks. The imaging and subsequent digitization (the latter is here taken as the manual transcription of data into electronic form, because automated optical reading is not practical) required cooperation between the U.K. and U.S. partners. The U.K. National Archives were contracted by the Met Office Hadley Centre to produce the images. Each logbook was photographed, and the images were stored on DVD and shipped to the National Oceanic and Atmospheric Administration’s (NOAA’s) National Climatic Data Center (NCDC), which administers the Climate Database Modernization Program (CDMP; Dupigny-Giroux et al. 2007).
The 267,874 images are stored at NCDC, and are available over the Internet to researchers through their Web Search, Store, Retrieve, Display visual archive system (WSSRD; www.ncdc.noaa.gov/oa/climate/cdmp/wssrd.html).

A keying process was designed and implemented using CDMP’s Pre-keying Inventory Comments and Summary (PICS) process. Keying of the data was then completed with a minimum 99% accuracy rate, verified by blind-double keying. The keyed data were randomly sampled to verify consistency with the output format. The keyed information includes the several sets of weather observations that appear on each page (see Fig. 4). The data and time of these observations is usually evident, but corresponding information on the ship’s location is not consistently recorded. Such reservations notwithstanding, by these means a total of about 1,500,000 observations were made available for addition to the datasets.

A sample logbook page and an illustration of the digitized observations for one Royal Navy ship is given in appendix C.

Observation processing and quality control. For scientific use, and inclusion in ICOADS, the observations had to be converted into International Maritime Meteorological Archive (IMMA) format (Woodruff 2007). [ICOADS data and products can all be accessed via the ICOADS web portal (http://icoads.noaa.gov/). This includes links to the National Center for Atmospheric Research (NCAR), from where the WWII data can be accessed as an ICOADS auxiliary dataset (http://dss.ucar.edu/datasets/ds530.0/); and to NCDC’s Climate Data Online (CDO; http://cdo.ncdc.noaa.gov/CDO/CDOMarineSelect.jsp), from where ICOADS data are also available. At this writing the WWII data are being blended into ICOADS, and will be included in a new release later in 2008.] As well as reformatting the data, the opportunity was taken to perform some basic quality control on the observations. Location proved to be a particular problem, and, for instance, many longitudes were significantly in error, or assigned to the wrong hemisphere.

When at sea, a typical logbook page contains many more weather observations than position records (Fig. 4): many ships recorded positions less often than 3 times a day. Of the digitized observations where the ship was at sea, 177,000 (29%) had positions given in the logbooks, the locations for other observations could often be estimated by interpolation between known positions, and this raised the number of usable observations to 508,000 (83%) of the 612,000 observations made at sea.

When in port, the logbooks only record the port name, not its latitude and longitude; thus, assigning a position requires a conversion table between port names and positions. The digitized observations included more than 1,700 port names, and not all of these could be unambiguously identified, but locations were found for 924,000 (97%) of the 952,000 observations made in port.

Coverage changes. The new observations have provided a significant increase in both the number of observations and the global coverage of observations (Figs. 5 and 6). Much of the improvement in coverage is due to the judicious selection of logbooks from ships crossing oceanic areas that have hitherto been poorly represented.

Using the new observations in climate analysis. A principal use of historical marine observations is to

![Graphs showing coverage and number of observations over time.](http://example.com/graph.png)

Fig. 5. Effect of the new observations on SST global coverage and number of observations.
Fig. 6. Spatial distribution of observations coverage for 1941–46. The map shows the mean number of observations per month in each 2° square classified into four categories. Areas where the new observations improve the coverage by at least one category are shown in red, otherwise the preexisting coverage from ICOADS is shown in blue.

The new observations contribute to gridded datasets showing the changes in SST through time. One such dataset is HadSST2, produced by the Met Office Hadley Centre (Rayner et al. 2006). The new observations have been used to make two new and improved versions of this dataset—one using only the new observations, and one made from a blend of the new observations with HadSST2. In both cases the observations have been subjected to the quality control, gridding, and bias adjustment methods used in HadSST2 and described in Rayner et al. (2006).

The improvement is particularly noticeable for the Indian Ocean. The upper half of Fig. 7 is a copy of Fig. 3.5 from Trenberth et al. (2007), and illustrates the problems caused by missing observations in the 1940s, with the extent and timing of the temperature maximum being almost impossible to estimate. The lower half of the figure is produced in the same way, but with the new observations added. The improved coverage for the 1940s needs little elaboration and is now sufficient to characterize the mid-century temperature maximum with greater precision.

Fig. 7. Latitude–time sections of zonal-mean Indian Ocean SST anomalies (°C) for 1900–2005, relative to the 1961–90 mean (after Fig. 3.5 of Trenberth et al. 2007)). (top) HadSST2 (Rayner et al. 2006) and (bottom) HadSST2 with the newly digitized observations.
SST Biases. Observed SSTs revealed two large changes over the WWII period and these may not be wholly climatic in character. First, observational procedures changed from measuring SST by taking samples in canvas buckets to using measurements of engine-room intake temperatures. This is known to have produced a systematic change in the temperature observations (Folland and Parker 1995) leading to a significant increase in measured temperatures between 1939 and 1941 [about 0.5°C in the global average; see Rayner et al. (2006); Smith and Reynolds (2004); Folland and Parker (1995); Barnett (1984)]. Even after correcting for these biases, measured SST shows a distinctive if attenuated peak in the mid-century, followed by an abrupt fall in 1946 (Fig. 1). Largely because of the war, there are also big changes in the list of ships contributing observations over this period. Different ships, particularly if from different nations and fleets, had different observing practices, and it is possible that there are some additional, as yet unknown, biases in the record. The new Royal Navy observations provide a useful check on this and, because all the ships presumably operated to the same set of meteorological observing practices, they should provide a time series prepared to be consistent observational standards. As a result, comparison between the Royal Navy-only data and the existing observations from a mixture of sources should draw attention to any effects caused by non-climatic effects in the latter.

Comparison of SST time series from the old (multisource) observations and from the new Royal Navy observations shows good agreement in the overall trend over the period (Fig. 8). There are fewer new than old observations, a feature that may account for the higher variability of this series. Data shortage is a yet more probable explanation for the departures in the record for 1939 and 1947. Such differences notwithstanding, the two series agree well over the 1938–45 period, and both then show a fall in 1946 and 1947, though the shape of the curve over these 2 yr is quite different in the two datasets. This suggests that the abrupt fall in SST seen in the older observations in 1946 may be an artifact of changes in measurement methods, but further analysis will be required before this hypothesis can be offered with complete confidence.

Assessing Uncertainty and Bias in Observations of SLP. One virtue of having many observations from ships in port is that several ships are likely to use the same harbor at once, and thus sets of observations for cross-validation of the systematic and random errors in each ship’s measurements are provided. During August and September 1938, 11 of the ships whose logbooks were digitized spent some time in Portsmouth Harbor. Figure 9 shows the SLP data recorded by each ship, compared to observations from...
Future uses. These logbook data have a wide application. More extensive work on the SST measurements from the logbook observations is expected to contribute to new versions of the HadSST dataset (Rayner et al. 2006). The observations also include dry- and wet-bulb air temperature measurements, which will contribute to any future versions of the nighttime marine air temperature dataset HadMAT (Ryaner et al. 2003) and the humidity dataset HadCRUH (Willett 2007). The pressure observations will be used not only in future regional and global gridded pressure datasets like EMSLP (Ansell et al. 2006) and HadSLP (Allan and Ansell 2006), but also will be assimilated in general circulation model reanalyses, notably the Twentieth Century Reanalysis Project (Compo et al. 2006).

Only a small proportion of available logbooks have been included in this exercise, and many data remain untapped. A succession of reanalyses using only surface observations, including and building from the Twentieth Century Reanalysis Project, are being planned under the Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative (R. Allan 2007, personal communication). These analyses will require the recovery and digitization of the bulk of the outstanding ship logbook observations held as hard-copy volumes in U.K. repositories.

Conclusions. Our understanding of large-scale climate change and variability is based largely on a database of digitized historical weather observations. There are few such observations available from before about 1850, and there...
are also significant shortages of observations in the late nineteenth century, and in the twentieth century during the two World Wars (Fig. 3). Observations exist that could be used to improve the database, and to extend the data back to 1800 or before, but these observations are only available on paper—in logbooks and reports in libraries and archives—and are currently unavailable to the climate research community. The Joint World Meteorological Organization (WMO)—Intergovernmental Oceanographic Commission (IOC) Technical Commission for Oceanography and Marine Meteorology (JCOMM) Workshop on Advances in Marine Climatology (CLIMAR; Parker et al. 2004) series, and the International Workshop on Advances in the Use of Historical Marine Climate Data (MARCDAT-II; Kent et al. 2007) series, have both emphasized the importance of digitizing such observations and making them available to researchers.

This campaign to digitize meteorological observations in Royal Navy logbooks for the period between 1938 and 1947 has substantially improved the observational coverage for the Second World War. It has also confirmed the benefits of utilizing this data source. Similar undertakings could provide similar improvements for the First World War and the late nineteenth century, and could eventually enable instrumental estimates of large-scale climate variability and change for the whole of the last 200 yr.

ACKNOWLEDGMENTS. Funding for this project was provided by the Climate Database Modernization Program of the US National Climatic Data Center, and by the Joint Defra and MoD Programme. (Defra) GA01101 (MoD) CBC/2B/0417 Annex C5. The logbooks used are in the collection of the U.K. National Archives, who also made the digital images.

APPENDIX A: LOGBOOK OF HMS RATTLESNAKE, 1796.
Royal Navy officers began to maintain logbooks in their currently recognizable form in the 1670s (Wheeler 2004). From that point until the mid-nineteenth century the presentational style of these documents was broadly consistent and fell into two categories, the more commonly used of which is exemplified in Figs. A1 and A2. In this style, logbooks consisted of facing pages. The left-hand page (Fig. A1) contained essential navigational information (latitude, longitude, and bearings to landmarks), the date, and notes on wind direction, with the latter commonly on a 32-point compass. The facing page (Fig. A2) was narrative, but each day’s entry opened with a note of the wind force and the weather at noon of that nautical day. In the example shown, from November 1796, HMS Rattlesnake lies some 52 leagues (approximately 150 miles) west of Table Bay (South Africa), and the winds on 3 November are south by east veering to south later, and are described as “fresh gales and squally” [Beaufort force eight according to CLIWOC (2003)]. The alternative style popular in the EIC is illustrated in appendix B. Both contain the same types of information although the EIC logbooks in latter years often also contained daily instrumental observations.

APPENDIX B: LOGBOOK AND OBSERVATIONS FROM EAST INDIA COMPANY SHIPS CASTLE HUNTLY AND HINDOSTAN, 1780 AND 1820. From about 1780, Alexander Dalrymple, chief hydrographer to the British East-
India Company, set up a system for officers to make daily records of air pressure and temperature on their voyages to and from India and China. Observations were made at midday and recorded in forms in the logbooks (Fig. B1). Most of these logbooks are archived in the British Library.

Some insight into the methods and procedures that Dalrymple expected of his crew is provided when he describes one such journey in detail and includes also the data to which it gave rise (Dalrymple 1778). But this voyage was by no means unique: Fig. B2 shows temperature and pressure records from a further voyage: that of the Hindostan, which left Canton in April 1780 and returned to London via the Sunda Strait, the Indian Ocean, the Cape of Good Hope and the Atlantic, arriving in England in September. The graphs summarize the quantity of data that are available in addition to the more standard wind force and direction described above. Thus far no comprehensive attempt has been made to recover these data, which are of particular value as they include passage through the southern oceans where even today data are scarce. It is estimated that as many as 800 logbooks contain data of this character, covering the period from the 1780s to the 1840s.

**APPENDIX C: OBSERVATIONS FROM HMS WARSPIE, 1941–43.**

The battleship HMS Warspite (Fig. C1) sailed between 1913 and 1947 and was one of the most famous ships of the Royal Navy. Her logbooks are in the National Archives, and those for the period 1941–43, when she was serving outside U.K. home waters, have been digitized as part of the project. Figure C2 shows the route of the Warspite during this period, as obtained from newly digitized observations from her logbooks.

Operational requirements took the Warspite over a wide area of the globe. On
FIG. C2. The route of HMS Warspite between 1941 and 1943. Warspite served as flagship of the Mediterranean Fleet from May 1940. She was damaged by air attack in May 1941 during the battle of Crete, and, after temporary repairs at Alexandria, steamed via Singapore to Bremerton, in the U.S. Pacific Northwest, for repairs. She left Bremerton in January 1942, by which time Japan had entered the war, so she steamed to her new station in the Indian Ocean (as part of the British Eastern Fleet) via the South Pacific and Australia. She served in the Indian Ocean until June 1943, when she was recalled to the Mediterranean; and in September 1943 she was once again damaged in an air attack, this time by armor-piercing glider bombs (during the Salerno landings). Returning to Rosyth in March 1944, Warspite was never fully repaired, but took part in support of the Normandy landings in Jun 1944 (Coward 1986).

all these journeys weather observations were made, and recorded in the logbook, several times a day. An example of such a page from the ship is shown in Fig. 4. As with the earlier logbooks, this one contains precise navigational data and a column of general remarks. But now there are instrumental weather observations included: pressure as well as dry-bulb, wet-bulb, and sea temperatures.

REFERENCES


——, S. Lubker, U. Radok, and S. Woodruff, 1998: South Atlantic winds and weather during and following the little ice age—A pilot study of English
——, 2007: Understanding arctic climates using old logs. Meteorologisch Informatiebulletin Maritim, KNMI.