

Weather Rescue at Sea: recovering historical weather observations from 19th century British naval ships

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Teleti, P., Hawkins, E. ORCID: <https://orcid.org/0000-0001-9477-3677> and Wilkinson, C. (2026) Weather Rescue at Sea: recovering historical weather observations from 19th century British naval ships. *Geoscience Data Journal*, 13 (2). e70056. ISSN 2049-6060 doi: 10.1002/gdj3.70056 Available at <https://centaur.reading.ac.uk/128857/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/gdj3.70056>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

DATA ARTICLE 

Weather Rescue at Sea: Recovering Historical Weather Observations From 19th Century British Naval Ships

Praveen Teleti^{1,2} | Ed Hawkins²  | Clive Wilkinson³¹College of Integrative Studies, Singapore Management University, Singapore, Singapore | ²Department of Meteorology, National Centre for Atmospheric Science, University of Reading, Reading, UK | ³Department of Meteorology, University of Reading, Reading, UK**Correspondence:** Praveen Teleti (praveent@smu.edu.sg)**Received:** 4 September 2025 | **Revised:** 7 December 2025 | **Accepted:** 7 January 2026**Keywords:** climate | data rescue | historical | ship logbooks | weather

ABSTRACT

Ship logbooks represent a critical source of historical meteorological data, providing valuable observations of barometric pressure, air temperature, sea surface temperature, wind force and direction, and other variables. Substantial quantities of these records are unavailable to climate science as they have not yet been transcribed. We present ‘Weather Rescue at Sea’, a citizen-science project which transcribed millions of weather observations contained in 19th Century UK Royal Navy ship logbooks. We describe the logbook structure and weather observation-taking instructions and discuss significant challenges with the translation of handwritten text into accurate data due to errors arising from ambiguous handwriting, historical terminology, and inconsistent metadata. We present the dataset and explore its spatio-temporal characteristics. The corrected and quality-assured datasets will enhance climate reanalyses and other historical reconstructions of the pre- and early industrial climate by providing more input meteorological data. Furthermore, we highlight emerging tools, such as AI-driven transcription correction, and outline remaining challenges in fully leveraging these historical records to advance climate science.

1 | Introduction

Throughout the 19th century, the British Royal Navy maintained meticulous logbooks as part of its operational requirements. These logbooks often contained detailed meteorological observations, including air and sea temperatures, wind force (speed) and direction (frequently using the Beaufort wind scale), barometric pressure, and descriptive weather conditions (Allan et al. 2016; Luterbacher et al. 2024). Today, these records provide an invaluable source of climate data for the pre-modern period, spanning vast geographic areas in all ocean basins (Allan et al. 2016). The temporal and spatial coverage makes these datasets particularly valuable for areas and periods where available instrumental weather data is currently sparse or non-existent. There is a scientific interest in understanding the climate of the early industrial era against which our present climate could be measured (Brönnimann et al. 2019).

Recognising the critical importance of these records, numerous data rescue and digitization activities have previously been undertaken. Over the past two decades, significant efforts have been undertaken to preserve, digitise, and make these logbooks accessible for climate research. These include Zooniverse-based citizen science projects such as Old Weather: WWII and Southern Weather Discovery, which leverage the work of thousands of volunteers to transcribe handwritten logbooks into machine-readable datasets (Teleti et al. 2023; Lorrey et al. 2022), as well as other notable data rescue projects such as CLIWOC, RECLAIM, and those associated with the ACRE initiative.

Despite these successes, recovering historical logbook data for scientific use presents considerable challenges. According to the RECLAIM report (Wilkinson et al. 2011) there are at least several hundred thousand ship logbooks yet to be imaged or transcribed. However, this could actually be millions

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Geoscience Data Journal* published by Royal Meteorological Society and John Wiley & Sons Ltd.

of logbooks when combined with yet to be discovered sources and non-English language documents. At the same time, handwriting ambiguities and unfamiliarity with historical notations can introduce errors in transcriptions, while systemic biases in 19th-century observational practices—such as the calibration differences in barometers or thermometers—necessitate careful validation and correction (Luterbacher et al. 2024).

To address these issues, innovative validation and correction methods have been developed. Firstly, emerging technologies, such as machine learning (ML) and optical character recognition (OCR), tailored for historical documents, hold promise for refining transcription and validation workflows (Zhang and Sieber 2025; Muheki et al. in review). However, these technologies are still in nascent stages, and they serve as complements to traditional citizen science workflows, reducing reliance on human labour while enhancing the accuracy and scalability of projects like Old Weather and Old Weather: WW2 (Teleti et al. 2023; Luterbacher et al. 2024).

To correct for transcription errors, statistical approaches, such as consensus scoring (resolving discrepancies between multiple transcriptions) and plausibility checks (e.g., comparing barometric readings against climatological norms), are increasingly used to ensure the accuracy of these datasets (Hawkins, Alexander, and Allan 2023; Hawkins, Brohan, et al. 2023; Teleti et al. 2023). Additionally, citizen scientists have introduced collaborative innovations, such as the Old Weather Forum's transcription tool, which demonstrates the evolving role of volunteer communities in digitization workflows (Old Weather Forum 2025).

Once validated, these datasets can be integrated into global datasets, such as the International Comprehensive Ocean–Atmosphere Data Set (ICOADS; Freeman et al. 2017) and NOAA's 20th Century Reanalysis (Slivinski et al. 2019), where their impact is transformative. These projects can help extend atmospheric reconstructions back into the pre-industrial era and improve the fidelity of early climate baselines. Contemporary studies using these datasets have revealed key insights into historical climate conditions. For example, analyses of Arctic logbooks have illuminated regional climate trends during the early 19th Century, while distant datasets have contributed to understanding phenomena like the North Atlantic Oscillation, El Niño–Southern Oscillation, Antarctic sea ice variations, and extreme storms (Brohan et al. 2010; Cropper et al. 2015; Edinburgh and Day 2016; Barrett et al. 2018; Hawkins, Burt, et al. 2023).

However, large parts of the Pacific and Southern Oceans, and regions away from trade routes, are poorly represented in many global climate datasets, making any ship logbooks that might contain observations from these regions invaluable for global reconstructions (Allan et al. 2016; Luterbacher et al. 2024). England and later Great Britain have a seafaring tradition that goes back to the age of discovery, which goes back to the 16th Century, and has historical trading links with many nations across the globe. Spanning decades, Royal Navy logbooks provide a crucial source addressing the scarcity of high-quality observations in these periods (CORRAL 2011). Royal Navy logbooks archived at

the UKHO (UK Hydrographic Office in Taunton) are well suited to produce data in the targeted time period with global coverage.

This paper explores the development of a 19th-century Royal Navy weather observations dataset from Weather Rescue At Sea, a citizen-science project hosted by [Zooniverse.org](https://www.zooniverse.org). Thousands of volunteers transcribed millions of sub-daily weather observations from ships travelling through the Atlantic, Indian, and Pacific ocean basins in the 1860s.

In the following sections, we provide an overview of logbooks used in the project, outline instructions and meta-data associated with observation taking, set up of citizen-science project and steps taken to standardise the extracted data. In the next section, we explore the resulting dataset and how it compares with existing large datasets. We then assess applications of this dataset in climatology, focusing on their integration into larger datasets, assimilation into reanalysis models and the insights they have provided into 19th-century climate variability. Finally, we explore the legacy of this project, ongoing developments in citizen science, data rescue frameworks, and the potential future role of machine learning and AI-driven approaches.

2 | Data and Methods

2.1 | Description of Logbooks

Royal Navy logbooks are notable for their geographical and temporal coverage. The present collection of 19th-century (1860s–1890s) naval voyages crossed vast, under-sampled regions, including the Pacific, North Atlantic, and Southern Oceans. The remark books, which date back to the 1750s, were meant to complement the ship's logs and were specifically kept to record hydrographic information to update charts and to provide information on ports and coasts where naval vessels might operate (Wilkinson et al. 2011). To follow widely used nomenclature, we refer to Remark books as Logbooks here.

These logbooks were issued by the Admiralty, remained the principal source of information on navigation and day-to-day running of ships, and were duly retained for legal and insurance purposes. They include records of weather and positional information. Hand-written Captains' logbooks are the most common type of logbook in this collection, and they record observations at a 4-hourly frequency (Figure 1).

Usually the logbook page image contains 4 days, with 2 days on each leaf. Observations start from the top-left corner of the page, continuing to the bottom-right. Generally, the name of the vessel, commanding officer or the Captains' name, dates of journey and name of the origin, destination, or last port of call are written at the top of each page.

Other information, such as the ship's position, is recorded in various columns. The logbooks often include observations of atmospheric pressure, wind force (speed) and direction (based on the Beaufort scale), air and sea surface temperatures, qualitative weather descriptions (e.g., visibility, waves, or storm observations) recorded in their respective columns for each day. Sometimes, Ditto or 'are written in place of values'. Note that

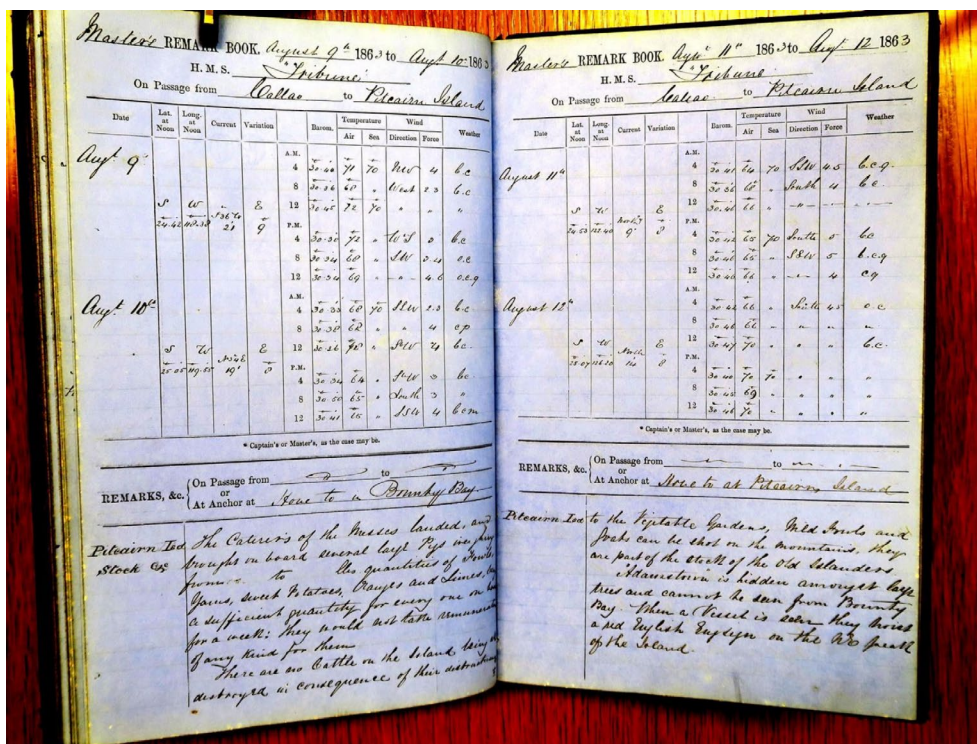


FIGURE 1 | A typical example page of a Master's or Captain's logbook during the 1860s. The image shown here is the logbook page of H.M.S. Tribune dated 9th–12th August 1863. Image courtesy of the UK Hydrographic Office (UKHO) archive.

the daily uninterrupted observations begin at 4 a.m., leading to 8 a.m., 12 a.m., 4 p.m., 8 p.m., and lastly 12 p.m. each day at Local time. This usage of 12 a.m. for noon and 12 p.m. for midnight is opposite to modern day standards.

Importantly, shipboard time was reset at local noon to synchronise the ships' clock to the time zone of the current position of the ship. The actual sub-daily time recorded in the log might not coincide with local time by several minutes (depending on the vessel's course) until the shipboard time is reset again at the subsequent noon. The intervals between the observations would still be 4 h within a slight margin.

Each weather record must contain a valid position, time and day to be useful. The ship's position (latitude and longitude) was observed and recorded once a day at noon as a usual practice, noted in degrees and minutes of four cardinal directions (e.g., North, South, East and West). In some instances a place name is written instead of the ship's position, usually when a ship called at a port or was moored at a known location. At this time, each ship carried a set of chronometers (up to three) to accurately calculate the ship's current longitude. Observations were recorded up to six times a day, and it is assumed that the current collection of logbooks follow local time for recording observations.

2.2 | Instructions

The British Admiralty issued an instruction manual (Herschel 1859; Figure 2) which was a scientifically sound and practical guide for observing various natural phenomena including weather on the high seas. Captains, masters of the Royal Navy and mercantile vessels, along with lieutenants, surgeons,

and midshipmen in the Royal Navy, and mates and inferior officers in the mercantile vessels were expected to study and put it into practice in the course of their duties.

The Manual drew from the latest scientific understanding and principles of the time. Frequent contributors were members of eminent scientific bodies such as the Royal Society, the Royal Geographic Society, and the Royal Observatory at Greenwich, including John Herschel. The Manual outlines crucial methods and procedures for conducting meteorological observations at sea, where such observations were invaluable for navigation and scientific study. The Manual emphasised the need for consistency and precision necessary for observing and recording weather data, including sea-level pressure, air and sea temperature, wind force (speed) and direction, and other weather phenomena.

The Manual strongly recommended maintaining a regular meteorological register, which involves systematically noting instrument readings and weather conditions at fixed hours, as gaps or inconsistencies could diminish the data's scientific value. Though the manual recommends taking observations at 3 a.m., 9 a.m., 3 p.m., and 9 p.m., the Admiralty issued logbooks at 4 a.m., 8 a.m., 12 p.m., 4 p.m., 8 p.m. and 12 a.m. In addition to scheduled observations, the Manual advises recording any unusual weather events as supplementary observations, providing a fuller picture of climatic variations.

The instructions for weather observations include detailed methodologies for measuring and recording atmospheric pressure, temperature (air and sea), wind, and other meteorological elements. These standardised procedures were designed to ensure consistent and reliable data collection across different environments, particularly at sea. However, meteorological observers on board would

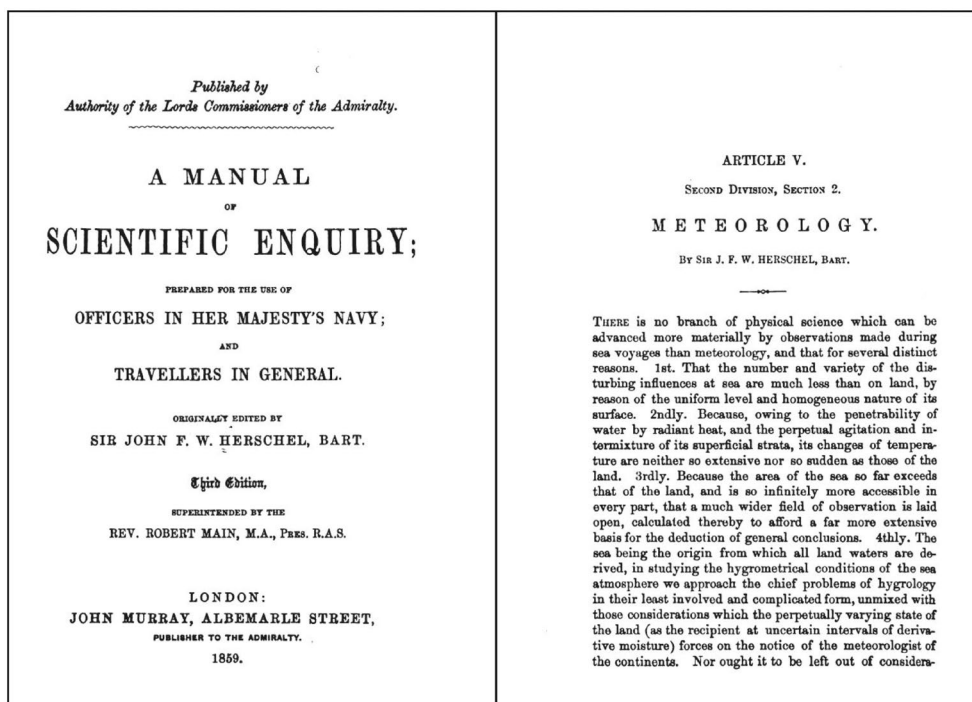


FIGURE 2 | Cover page of British Admiralty instruction manual to be used for observations on board ships (left panel) and in the Meteorology Section (right panel). Published by John Murray, London in 1859.

follow the instructions as much as they could given the weather and instrument conditions. Sometimes adjustments were made without careful computations leading to offsets and biases.

Barometric pressure is measured using mercury barometers, often mounted on gimbals aboard ships to counteract vessel movement. The barometer is placed indoors, shielded from sunlight and sudden temperature changes, usually near midships. Observers read the mercury column at eye level, aligning the vernier scale with the meniscus, and apply corrections for temperature, capacity, capillarity, and altitude. Aneroid barometers, though portable, are less precise and require periodic calibration against mercury barometers.

Air temperature is measured with calibrated mercury thermometers recommended to be placed in sheltered screens (similar to the later standard Stevenson type) and ventilated enclosures. The screens are positioned 1.25–2 m above deck level, away from heat sources or reflective surfaces. Self-registering thermometers are used to record daily maximum and minimum temperatures, reset after each 24-h period. Calibration against a reference thermometer ensures accuracy, and adjustments account for systematic errors and environmental influences.

Sea surface temperature (SST) is measured using bucket thermometers or subsurface thermometers. Bucket thermometers require a water sample taken from about 1 m below the surface, with the reading recorded immediately to prevent air temperature interference. Consistency in sampling depth and regular calibration against standard thermometers ensure reliable SST data. Even though wind force (speed) was recommended to be recorded using Lind's anemometer, due to the unreliability of the instrument, wind force (speed) was visually estimated in

Beaufort's 12-point wind scale. The wind direction was determined by compass-aligned wind vanes.

Observers often note shifts in wind direction and cross-reference cloud movement at various altitudes to understand multi-layer wind patterns. Corrections for ship movement and environmental conditions, such as strong gusts, are applied and corrected values are recorded. Cloud cover, type, movement, and altitude are recorded using Howard's cloud classifications (e.g., cirrus, cumulus). Weather types, such as rain, fog, or squalls, are documented using standardised abbreviations.

Unique phenomena like auroras or thunderstorms are described in detail. These standardised methods allowed comprehensive parameters of weather to be recorded which are now critical for historical climate research. The most important instructions were printed and attached to the logbooks to aid memory and offer quick reference to the observers (Figure 3).

2.3 | Citizen-Science Data Recovery

Weather Rescue at Sea, a citizen-science project, was launched on the Zooniverse platform in October 2021 (Figure 4a). We asked volunteers to transcribe weather observations along with positions and dates recorded in the 19th Century Royal Navy ship logbooks. The transcription process was grouped by different weather variables to ease typing and reduce typographical errors. Five workflows were developed and deployed: Navigation (for position and time), Air-pressure, Air temperature, SST, Wind direction and force (speed). Each workflow contained data-entry spaces for typing observations for six 4-hourly observations (except for Navigation: only one observation at Noon) across 4 days visible on the page.

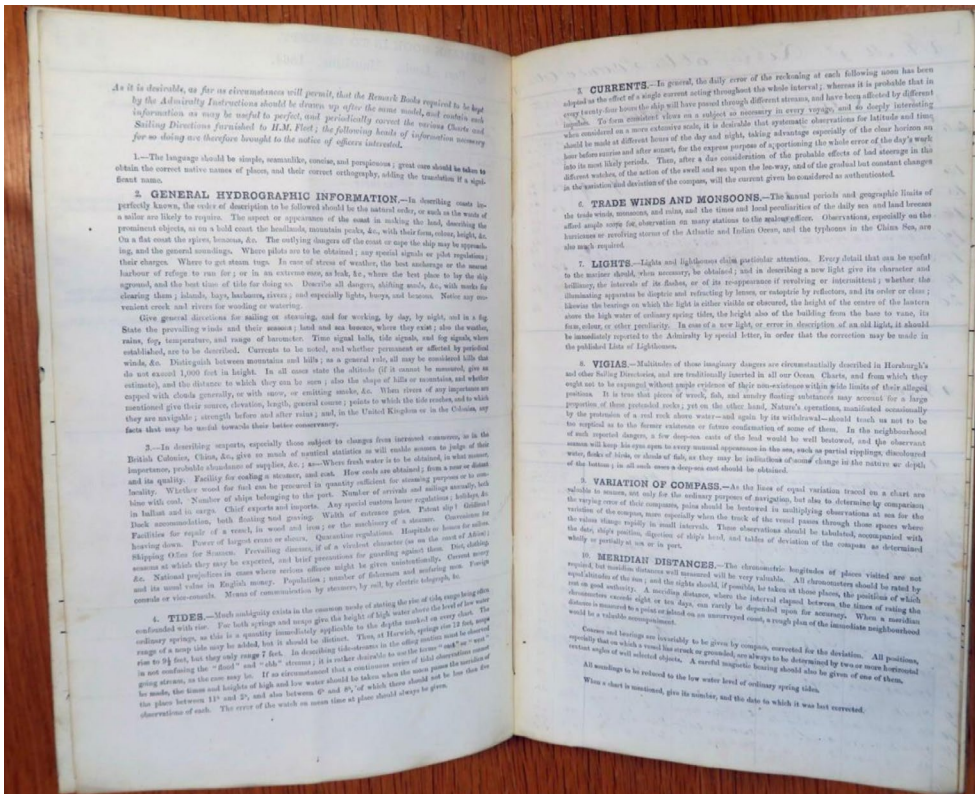


FIGURE 3 | Instruction pages inside of the ship logbooks to provide ready reference to the main objectives and procedures of weather observation taking at sea. Image courtesy of the UK Hydrographic Office (UKHO) archive.

The project was carefully designed to maximise accessibility and participation while maintaining scientific rigour. The project invited volunteers to engage with scanned images of ship logbook pages. Participants transcribed key weather data, guided by intuitive tools, clear instructions in tutorials and a field guide, and an active community forum where questions can be addressed. This setup ensures that individuals with no prior experience in climate science or historical transcription can contribute meaningfully to the project.

A total of 190 logbooks were transcribed by 4500 volunteers over an 18 month period. Central to the project's success is its ability to attract and retain a global community of volunteers. The design emphasizes inclusivity, presenting tasks in manageable increments and highlighting the tangible scientific impact of participants' contributions. Volunteers are not only transcribers but also collaborators in a global effort to recover vital climate data. Regular updates on project progress, milestones, and scientific outcomes further reinforce their engagement and showcase the value of their efforts.

Each image was transcribed by four volunteers for all workflows except wind direction and force (speed) where six were required. Transcription of the same image by different volunteers helps to correct any mistakes made by any individual volunteer. Volunteers were constantly engaged with the project, as Figure 4b shows when volunteers are ordered by number of classifications completed. The first thousand volunteers contributed almost 62% to the total transcription, the next thousand volunteers contributed 25%, while the remaining number of volunteers contributed 13% to the total transcription.

The speed at which images are retired (when enough volunteers have transcribed the page) varies due to a number of factors (Figure 4c). For example, at the launch of the project, when a lot of publicity about the project is circulated, a lot of volunteers begin to transcribe. As the project continues, some volunteers stop transcriptions, or do less of it; hence, we see a slower increase. We have observed that the number of images being retired stagnates, and if there is any news article or social media post made, transcription activities increase.

Several holiday periods such as Summer and Christmas breaks do have an impact on the number of transcriptions per week. In general, autumn and winter months are observed to be high activity periods.

2.4 | Standardisation

Like data from any historical source, data transcribed from historical ship logbooks contain some errors. There are three main sources: (1) incorrectly reading from instruments, or using uncalibrated instruments at the time by the original observer; (2) incorrectly recording the observations, or writing in poor, illegible handwriting could lead to erroneous values when transcribed; (3) volunteers could mistype numbers and text written on the logbook image onto the data entry form.

The first two sources are difficult to identify or correct as the original value shown on the instrument scale is lost forever. The third source can be readily rectified and corrected by choosing a value that has been written most times by the volunteers. As the

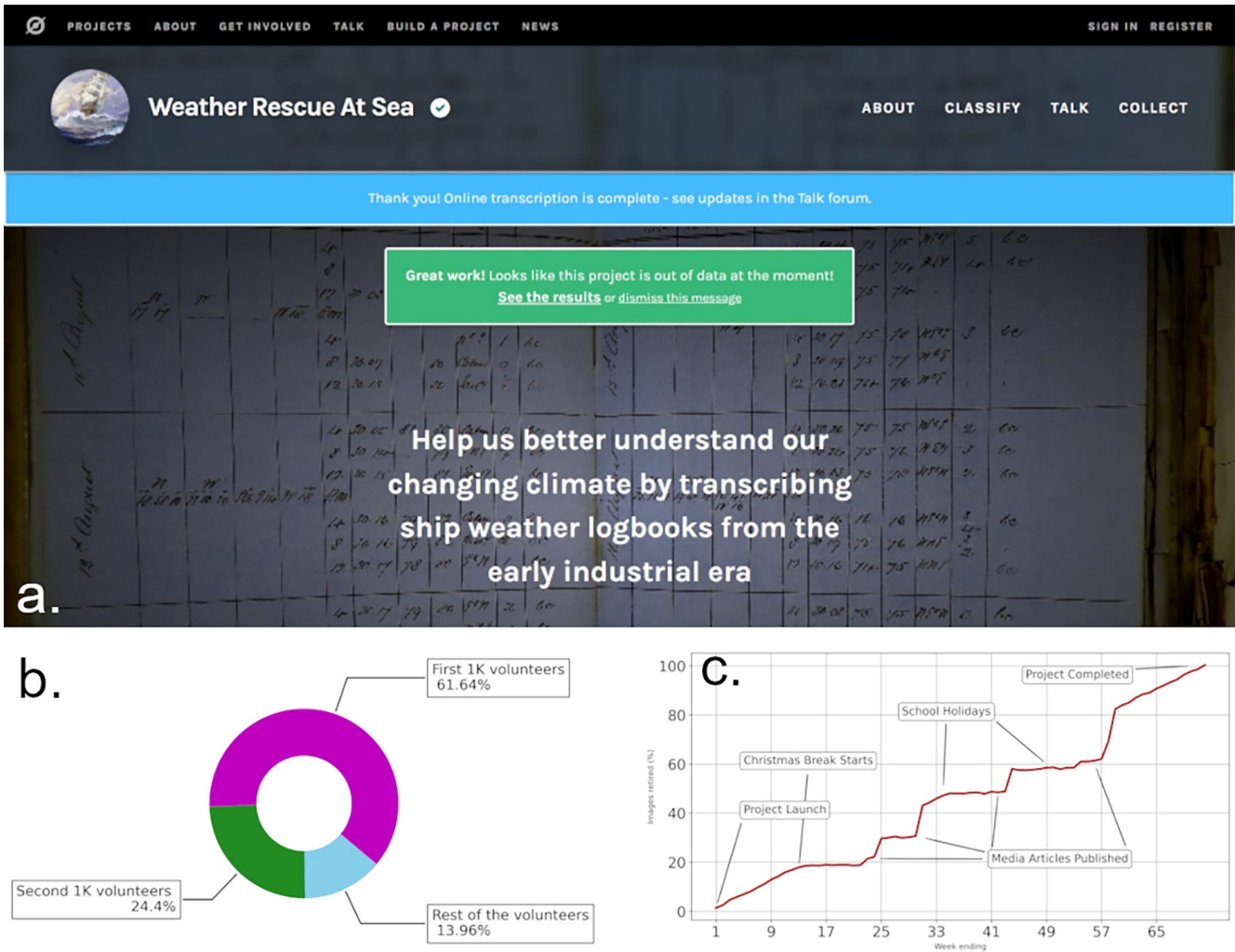


FIGURE 4 | (a) Screenshot of Weather Rescue at Sea project homepage hosted on the Zooniverse platform. (b) Engagement of volunteers with the project in terms of number of classifications ranked by contributions. (c) Percent of images being retired at the end of each week counting from launch date.

Zooniverse platform provides redundancy in its workflows, we allow multiple volunteers to transcribe each logbook entry. After the transcription process, the raw data were combined and compared. If all four (six in case of Wind Direction and Force) entries matched, the value was accepted. If a value reaches majority, then it is accepted. If no clear majority exists then the nearest or first value is accepted. All values including positional and meteorological observations are accompanied by a confidence flag variable which stores the fraction of value that appeared out of all values transcribed. This flag provides an objective method to assess trustability of transcribed observations (discussed in more detail in Section 4). If all values differed, the field was left blank and flagged as inconsistent. Errors happened more often when volunteers entered free text than when they recorded single observations.

Ships are often moored at a place for several weeks and months, or frequently called at ports. In these cases the ship positions are represented by place names instead of latitude and longitude. Some of the names of landmarks and ports are historical and obscure, so learning their accurate position is difficult, especially when these ports/landmarks have modern names and different

places have identical names. To solve this, a geolocation table was built with the help of volunteers, where each unique name found in the logbooks was assigned a latitude and longitude position. All variations in names were listed and assigned a position. For identical names, each name was attached to a country or region to differentiate. We have used simple linear interpolation to infill gaps in the ships' positions within physically plausible distances covered by 19th Century naval ships.

For weather observations, we have not removed or filled in any observations in the records. A major focus of this project has been transcription error correction, validation, and metadata cross-referencing to ensure high-quality outputs for climatological use.

3 | Dataset

A dataset with more than 255,000 records is produced containing position, date and time, along with air pressure, air temperature, SST, wind direction and wind force (speed) observations. Confidence levels for each weather variable are also produced to

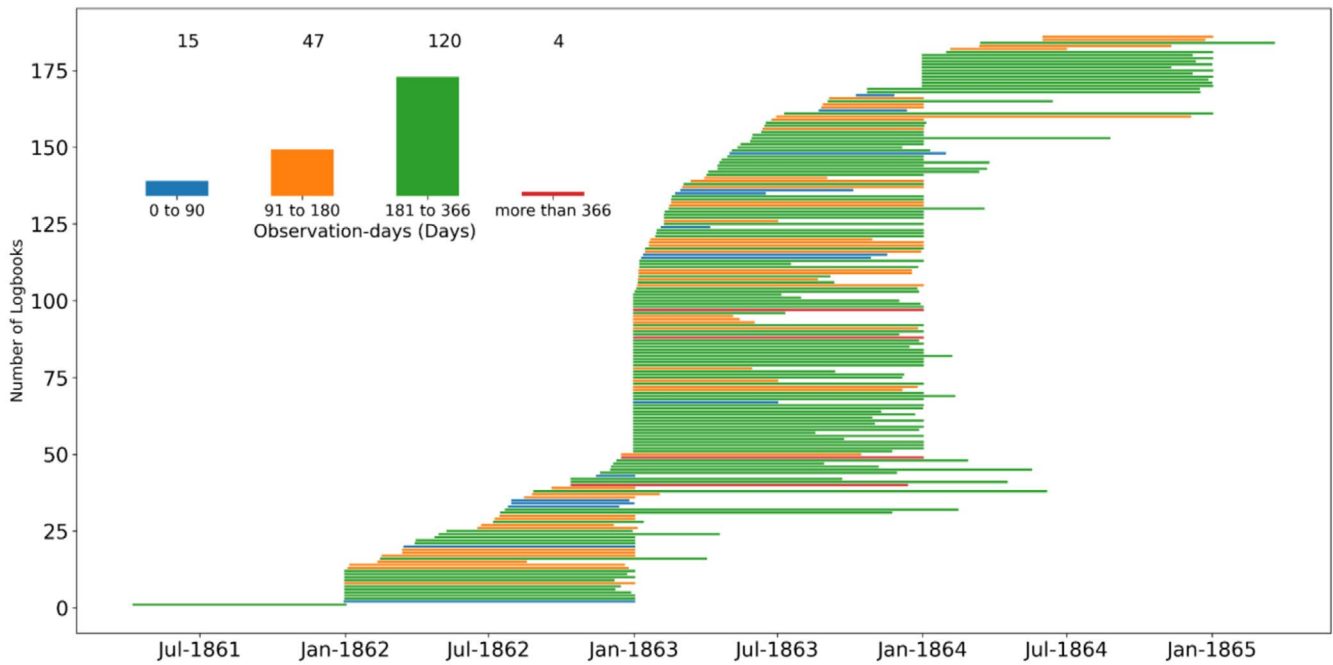


FIGURE 5 | The timings and number of days contained in each logbook transcribed in this project denoted by colour corresponding to the histogram shown in the inset. Each bar spans between the first date and last date in the logbook, while the colour shows the actual number of unique days recorded in the logbook.

expand reliability metrics for down-stream data users. The dataset spans from April 1861 to March 1865 and is spread across almost all ocean basins. The earliest logbook in the present collection is from HMS *Jason* starting in April 1861, and the latest is HMS *Adventure* starting in March 1865. The longest logbook in the collection is from HMS *Himalaya*, which has a duration of 646 days, whereas the shortest logbook is HMS *Gibraltar* having only 45 days in the logbook.

Most logbooks were printed for a calendar year, starting from 1st of January to 31st of December, which is supported by the histogram of the lengths of logbooks. 47 and 120 logbooks have a length between 90 to 180 and 181 to 366 days, respectively. Only a very few of the logbooks are longer than a year. The majority of the observations in the dataset are from 1863, with smaller proportions from 1862 to 1864 (Figure 5).

To understand the spatial distribution of the dataset, all observation counts within a 2×2 degree grid box are summed and divided by duration of the whole dataset in months to construct a monthly density plot (Figure 6). The dataset contains observations from remote regions covering the North Atlantic, South Atlantic, and Indian Ocean basins, and parts of the Pacific Ocean. The dataset covers commercial shipping lanes across the globe and also contains marine observations from less frequented regions. For example, the Pacific Ocean tracks near Hawaii, the west coast of South America, and the Southern Ocean.

Several regions have less than one observation per month (pale green) on average; however, coastal regions, restocking stops (e.g., Bermuda Is, Madeira Is) near shipping lanes to and from the UK have up to 200 monthly observations. The number of observations in high density grid cells is equivalent to six sub-daily observations per day for the whole period. Several noteworthy

regions include the far east; for example, the Sea of Japan and South China Sea, where observations have previously been scarce in existing datasets (e.g., ICOADS), and equatorial West Africa and the east coast of South America.

Although the present dataset is a relatively small addition to existing marine datasets such as ICOADS in terms of total observation numbers, the global coverage and number of observations in this specific period surpasses the ICOADS dataset (Figure 7). The comparison is made between the same weather variables contained in both datasets. For example, the monthly number of marine air temperature (MAT) observations contained in ICOADS in the 1860s decade is around 9000 and covers about 20% of global oceans. The present dataset has around 11,000 observations, but these observations are concentrated in fewer regions and the fraction of ocean grid cells covered is $< 2\%$.

Similarly, for SST, the present dataset has a roughly equal number of observations but ocean grid cells covered are very low as compared to ICOADS. However, it is the air pressure (SLP) observations where we see the greatest relative gain. The number of observations in the present dataset is almost double the number in ICOADS for the same time period. The present dataset is from a small data-rescue effort with limited means and time; however, this demonstrates that targeted data-rescue activities can have a significant impact on overall numbers and coverage of historical marine observations.

4 | Discussion and Future Directions

A new dataset consisting of weather observations transcribed from 19th Century Royal Navy logbooks is presented here. These observations can now be integrated into larger, long-term

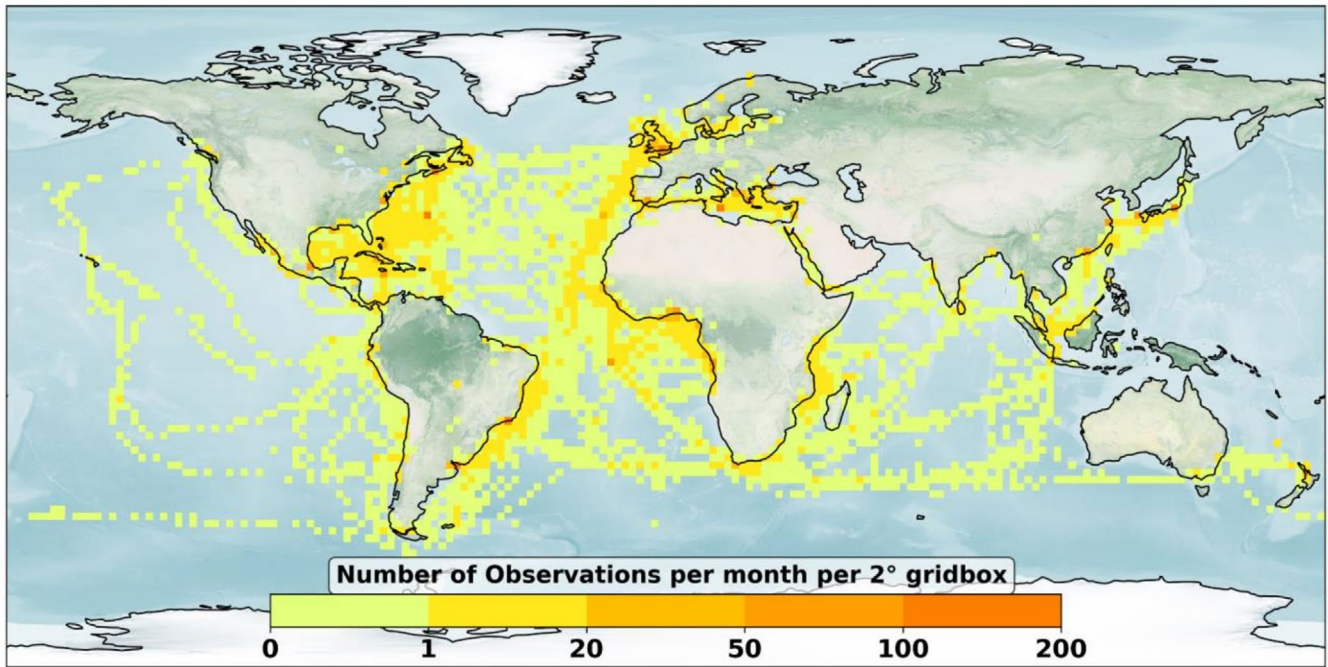


FIGURE 6 | Average monthly observational density for April 1861 to March 1865 in each 2×2 degree grid box for the dataset. Pale colours show where there are fewer than one observation per month on average, with darker colours denoting observation-rich areas with up to 200 observations per month on average.

observational marine datasets such as ICOADS, helping to fill gaps in both space and time, particularly in under-sampled regions such as the Arctic and Southern Oceans. Richer and denser datasets can significantly contribute to constructing long-term baselines for historical and modern climate variations (e.g., Morice et al. 2024). Such datasets also provide critical inputs for long-term reanalyses such as NOAA's 20th Century Reanalysis, which integrates air pressure observations to reconstruct large-scale atmospheric circulation patterns and past climate going back to the pre-industrial and early-industrial periods (Slivinski et al. 2019).

Over the years there have been many attempts to measure the trustability of historical observations in turn to estimate their usability in understanding long-term climate (e.g., Auer et al. 2007; Chandler et al. 2012, Sieber et al. 2022). To objectively assess different kinds of historical data requires a robust framework; Sieber et al. (2022) provide such structure. All historical observations are marked on three attributes—ability (skill), benevolence (motivation to do good), and integrity/consistency—and apply these to four stages: original observers, their observations, modern transcribers, and the transcription process itself.

We can apply their 19-point checklist to our data rescue project as follows: the trust framework treats officers and crew as the historical 'observers' and the logbooks volumes as the 'observations'. Their mandated, legally significant record-keeping, standardised formats, and routine subdaily entries support trust in observer skill and consistency, while navigation and safety provide a strong, if indirect, motivation to make accurate weather observations. We show that 19th Century Royal Navy observers were instructed to take weather observations in accordance with the Instruction Manual issued by the British Admiralty. The

methodologies for measuring weather variables in the Manual indicate consistent protocols, carefully positioned instruments, and precise corrections.

For example, methods like using sheltered screens for air temperature and standardised sampling depths for SST, the Manual established detailed practices to ensure data accuracy and consistency. There are instructions for zero-point calibration and environmental adjustments to create reliable and accurate observations. These standardised techniques allowed seafarers and weather observers to gather high-quality data, contributing to their knowledge of ambient weather at that time and laying the foundation of contemporary understanding of knowledge of climatology and marine weather. Although we cannot know whether all observations were made following these protocols. But, we can use the document to know who observed (rank, training, ship role), how often and in what format they recorded, what instruments and scales were used, and whether sample pages show plausible, internally consistent weather for season and location.

On the modern side, citizen scientists and the online platform become the 'transcribers' and 'transcription process'. The project (Weather Rescue at Sea) demonstrates clear interfaces (e.g., Zooniverse platform), tutorials, forums, and multiple independent transcriptions per each observation. Volunteers have communicated that motivations include contributing to science, interest in history and climate, and enjoyment of the task, which map cleanly to the framework's benevolence dimension. Multiple volunteers typically transcribe the same page, with consensus checks in our project before data are accepted into the dataset (as described in Section 2.4), which is exactly the kind of pre-validation step the checklist recommends. Progress updates, media stories, and listening to researchers working on the project all strengthen perceived integrity and help sustain careful work.

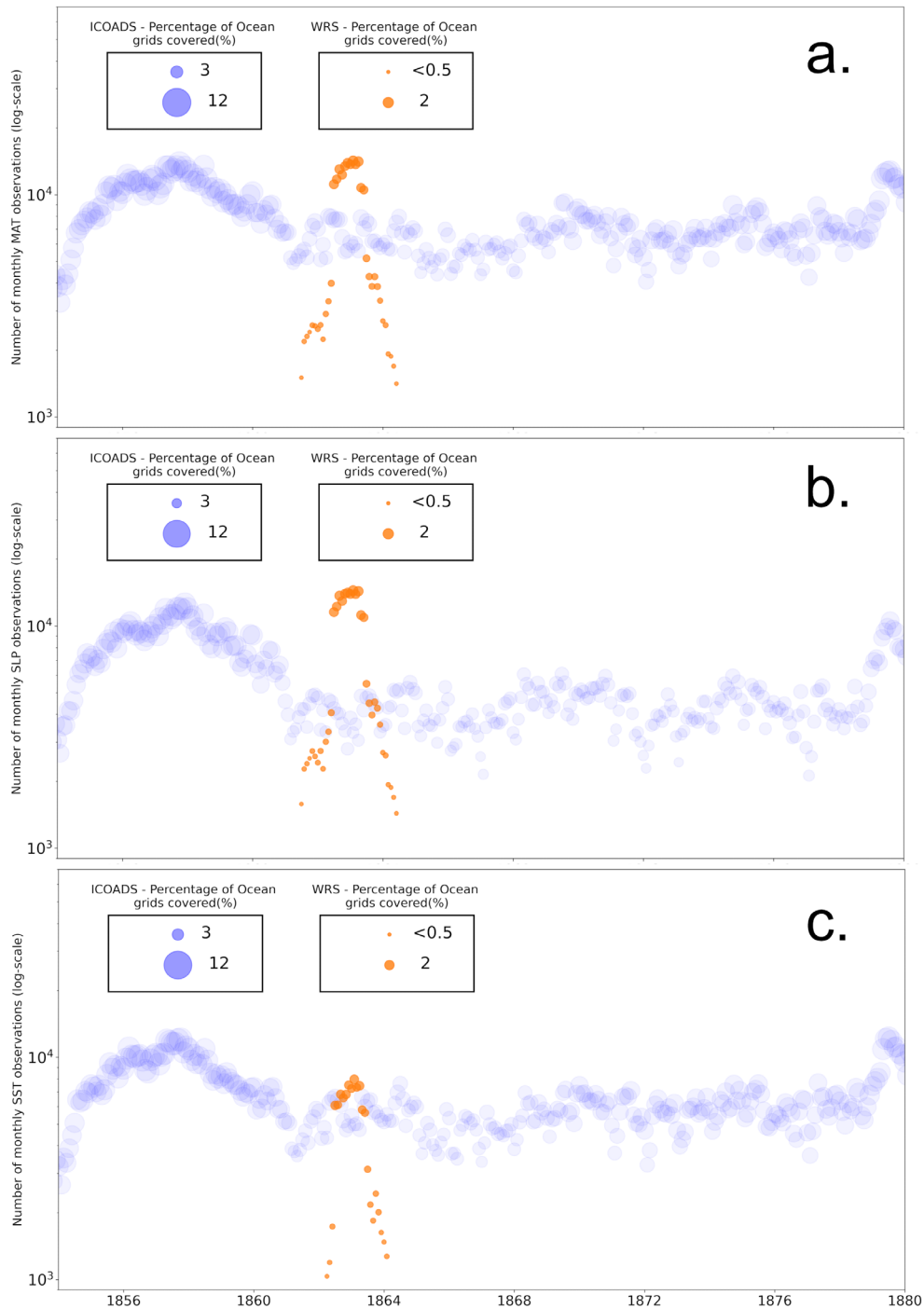


FIGURE 7 | Total number of monthly observations for ICOADS (blue) and Weather Rescue at Sea (WRS; orange). The size of the marker is proportional to the percentage of ocean grids covered by the observations in each month. (a) MAT observations, (b) SLP observations, (c) SST observations.

As all elements of the project can be mapped explicitly onto the trust framework, it contains all attributes of ability, benevolence, and consistency in transcription. Hence, our project satisfies requirements for trustability. Perhaps a future data rescue project could publish a short ‘trust profile’ with the rescued dataset so downstream climate users can gauge how much qualitative confidence to place in each variable.

Despite this work and other similar works the number of logbooks yet to be transcribed is at least hundreds of thousands

(Allan et al. 2016). To expedite the transcription process, an automated system would be necessary to handle such a volume of documents. Fortunately, several attempts using Machine Learning (ML) and Optical Character Recognition (OCR) are potentially emerging, tailored to historical documents (Muheki et al. 2025; Zhang and Sieber 2025). As these methods are in their nascent stage, the combination of ML, OCR, and human review is likely to make quick progress in efficiently processing a massive number of logbooks. These approaches will augment traditional volunteer-based workflows and improve efficiency

(Lorrey et al. 2022). However, software that can reliably transcribe historical hand-written, tabular logbooks is not yet available.

Filling in the gaps in our knowledge will remove ambiguity in how the climate varied historically in many regions where observation availability is currently poor or non-existent. The data generated through this project will help fill many crucial gaps in the large climate datasets (e.g., ICOADS) which will be used to generate new estimates of the industrial and pre-industrial era baseline climate. But more generally, this data and data from other historical sources are used to improve the models and reanalysis systems used for climate and weather research.

We produced and published raw dataset without any additional corrections and adjustments other than standardisation of measurement units. No positional or weather variables have been interpolated to retain original data as much as possible. We are aware that observations at this stage of processing are not comparable to existing climate datasets such as ICOADS. Our aim was to make raw observational dataset with meta information available to the climate community, free to apply required adjustments for specific climate applications.

The transcribed data serves a dual purpose: supporting climate research and preserving maritime history. In addition to feeding into climate reanalyses and models, the rescued records also provide valuable insights for historians studying the intersection of human activity and the environment. The collaboration between citizen scientists and researchers fosters a unique interdisciplinary dialogue, where public participation enhances scientific discovery while connecting participants to the narratives of history. We have also highlighted the importance of citizen science in overcoming barriers to data accessibility and discuss how initiatives like this not only advance climate science but also strengthen public understanding and appreciation of its relevance.

Acknowledgements

We thank the 4300 volunteers who contributed so much of their time for the transcription of the observations, and the many volunteers who continued to provide invaluable input during the construction of the final dataset. This is the data rescue component of GloSAT (grant number NE/S015574/1), a project funded by the UK Natural Environment Research Council to develop and analyse an extended and consistent global surface temperature climate record back to the 1780s, based on air temperature observations recorded across land, ocean and ice. E.H. was supported by the UK National Centre for Atmospheric Science.

Data Availability Statement

The dataset contains 255,000 records containing position, date and time, along with air pressure, air temperature, SST, wind direction and wind force (speed) observations. The dataset spans from April 1861 to March 1865 and is spread across almost all ocean basins. The dataset is accessible at Zenodo (<https://doi.org/10.5281/zenodo.17051534>).

References

Allan, R., G. Endfield, V. Damodaran, et al. 2016. "Toward Integrated Historical Climate Research: The Example of Atmospheric Circulation

Reconstructions Over the Earth." *Wiley Interdisciplinary Reviews: Climate Change* 7: 161–313.

Auer, I., R. Böhm, A. Jurkovic, et al., 2007. "HISTALP—Historical Instrumental Climatological Surface Time Series of the Greater Alpine Region." *International Journal of Climatology* 27: 17–46. <https://doi.org/10.1002/joc.1377>.

Barrett, H. G., J. M. Jones, and G. R. Bigg. 2018. "Reconstructing El Niño Southern Oscillation Using Data From Ships' Logbooks, 1815–1854. Part I: Methodology and Evaluation." *Climate Dynamics* 50, no. 3: 845–862. <https://doi.org/10.1007/s00382-017-3644-7>.

Brohan, P., C. Ward, G. Willetts, C. Wilkinson, R. Allan, and D. Wheeler. 2010. "Arctic Marine Climate of the Early Nineteenth Century." *Climate of the Past* 6: 315–324.

Brönnimann, S., R. Allan, L. Ashcroft, et al. 2019. "Unlocking Pre-1850 Instrumental Meteorological Records: A Global Inventory." *Bulletin of the American Meteorological Society* 100, no. 12: ES389–ES413. <https://doi.org/10.1175/BAMS-D-19-0040.1>.

Chandler, R. E., P. Thorne, J. Lawrimore, and K. Willett. 2012. "Building Trust in Climate Science: Data Products for the 21st Century." *Environmetrics* 23: 373–381. <https://doi.org/10.1002/env.2141>.

CORRAL. 2011. "CORRAL: Historical Meteorological Recordings From the UK Colonial Registers and Royal Navy Logbook Images." NCAS British Atmospheric Data Centre, 2025. UK Colonial Registers and Royal Navy Logbooks Project; Met Office, University of Sunderland. <http://catalogue.ceda.ac.uk/uuid/c0001a8c47724dd514e64daf358fc60>.

Cropper, T., E. Hanna, M. A. Valente, and T. Jónsson. 2015. "A Daily Azores–Iceland North Atlantic Oscillation Index Back to 1850." *Geoscience Data Journal* 2, no. 1: 12–24. <https://doi.org/10.1002/gdj3.23>.

Edinburgh, T., and J. J. Day. 2016. "Estimating the Extent of Antarctic Summer Sea Ice During the Heroic Age of Antarctic Exploration." *Cryosphere* 10: 2721–2730. <https://doi.org/10.5194/tc-10-2721-2016>.

Freeman, E., S. D. Woodruff, S. J. Worley, et al. 2017. "ICOADS Release 3.0: A Major Update to the Historical Marine Climate Record." *International Journal of Climatology* 37, no. 5: 2211–2232. <https://doi.org/10.1002/joc.4775>.

Hawkins, E., L. V. Alexander, and R. J. Allan. 2023. "Millions of Digitized Historical Sea-Level Pressure Observations Rediscovered." *Geoscience Data Journal* 10, no. 3: 385–395. <https://doi.org/10.1002/gdj3.163>.

Hawkins, E., P. Brohan, S. N. Burgess, et al. 2023. "Rescuing Historical Weather Observations Improves Quantification of Severe Windstorm Risks." *Natural Hazards and Earth System Sciences* 23, no. 4: 1465–1482. <https://doi.org/10.5194/nhess-23-1465-2023>.

Hawkins, E., S. Burt, M. McCarthy, et al. 2023. "Millions of Historical Monthly Rainfall Observations Taken in the UK and Ireland Rescued by Citizen Scientists." *Geoscience Data Journal* 10, no. 2: 246–261. <https://doi.org/10.1002/gdj3.157>.

Herschel, J. F. W., ed. 1859. *A Manual of Scientific Enquiry; Prepared for the Use of Officers in Her Majesty's Navy; and Travellers in General*. Admiralty. John Murray.

Lorrey, A. M., P. R. Pearce, R. Allan, et al. 2022. "Meteorological Data Rescue: Citizen Science Lessons Learned From Southern Weather Discovery." *Patterns* 3, no. 6: 100495. <https://doi.org/10.1016/j.patter.2022.100495>.

Luterbacher, J., R. Allan, C. Wilkinson, et al. 2024. "The Importance and Scientific Value of Long Weather and Climate Records; Examples of Historical Marine Data Efforts Across the Globe." *Climate* 12, no. 3: 39.

- Morice, C. P., D. I. Berry, R. C. Cornes, et al. 2024. "An Observational Record of Global Gridded Near Surface Air Temperature Change Over Land and Ocean From 1781." *Earth System Science Data Discussions* 17: 1–32. <https://doi.org/10.5194/essd-2024-500>.
- Muheki, D., B. Vercautse, K. K. T. Chandrasekar, et al. 2025. "MeteoSaver v1.0: A Machine-Learning Based Software for the Transcription of Historical Weather Data." *EGUsphere* [Preprint]. <https://doi.org/10.5194/egusphere-2024-3779>.
- Old Weather Forum. 2025. "The Origin and Evolution of Old Weather - Old Weather Forum." Accessed on June 28, 2025. <https://www.oldweatherforum.org/viewtopic.php?p=20883&hilit=transcription+excel#p20883>.
- Sieber, R., V. Slonosky, L. Ashcroft, and C. Pudmenzky. 2022. "Formalizing Trust in Historical Weather Data." *Weather, Climate, and Society* 14, no. 3: 993–1007. <https://doi.org/10.1175/WCAS-D-21-0077.1>.
- Slivinski, L. C., G. P. Compo, J. S. Whitaker, et al. 2019. "Towards a More Reliable Historical Reanalysis: Improvements for Version 3 of the Twentieth Century Reanalysis System." *Quarterly Journal of the Royal Meteorological Society* 145, no. 724: 2876–2908. <https://doi.org/10.1002/qj.3598>.
- Teleti, P., E. Hawkins, and K. R. Wood. 2023. "Digitizing Weather Observations From World War II US Naval Ship Logbooks." *Geoscience Data Journal* 11, no. 3: 314–329. <https://doi.org/10.1002/gdj3.222>.
- Wilkinson, C., S. D. Woodruff, P. Brohan, et al. 2011. "Recovery of Logbooks and International Marine Data: The RECLAIM Project." *International Journal of Climatology* 31, no. 7: 968–979. <https://doi.org/10.1002/joc.2102>.
- Zhang, Y., and R. E. Sieber. 2025. "Automation of Historical Weather Data Rescue." *Geoscience Data Journal* 12, no. 1: e261. <https://doi.org/10.1002/gdj3.261>.