University of Reading

Digitizing observations from the 1861–1875 Met Office Daily Weather Reports using citizen scientist volunteers

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Craig, P. M. and Hawkins, E. ORCID: https://orcid.org/0000-0001-9477-3677 (2024) Digitizing observations from the 1861–1875 Met Office Daily Weather Reports using citizen scientist volunteers. Geoscience Data Journal, 11 (4). pp. 608-622. ISSN 2049-6060 doi: 10.1002/gdj3.236 Available at https://centaur.reading.ac.uk/114749/

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/gdj3.236

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

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Check for updates

Received: 27 March 2023 | Revised: 5 November 2023 | Accepted: 11 December 2023

DOI: 10.1002/gdj3.236

Geoscience Data Journal | RMetS | WILEY

DATA ARTICLE

Digitizing observations from the 1861–1875 Met Office Daily Weather Reports using citizen scientist volunteers

Philip M. Craig | **Ed Hawkins**

National Centre for Atmospheric Science, University of Reading, Reading, UK

Correspondence
Ed Hawkins, National Centre for Atmospheric Science, University of Reading, Reading RG6 6BB, UK.
Email: ed.hawkins@ncas.ac.uk

Funding information
Natural Environment Research Council; National Centre for Atmospheric Science

Abstract

We describe the transcription and quality control processes for rescuing around 570,000 sub-daily and daily weather observations which were recorded in the UK Met Office Daily Weather Reports during the 1861–1875 period. These data are from the start of coordinated weather observations and were collected with the aim of making the first-ever weather forecasts. The observations were rescued thanks to 3500 volunteers and include sub-daily sea-level pressure, dry and wet bulb temperatures, daily maximum and minimum temperatures, and daily rainfall amounts from 70 different locations across Western Europe, and one in Canada. We highlight how these observations will be used to fill gaps in existing pressure and temperature datasets and use two case studies to show how the pressure observations will likely better constrain the atmospheric circulation during two severe storms. We also compare a sub-sample of the newly rescued observations with data that were previously digitized for a small number of locations for the same dates, finding good agreement in general, although some discrepancies remain.

KEYWORDS

daily weather reports, data rescue, Fitzroy, pressure, quality control, rainfall, temperature

1 | INTRODUCTION

After a severe storm which caused the sinking of the *Royal Charter* ship in October 1859, Vice-Admiral Robert Fitzroy noted that "telegraphic communication can give notice of a storm's approach, to places then some hundred miles distant, and not otherwise forewarned" (Fitzroy, 1860). Fitzroy's vision was to collect

Dataset Identifier: doi: 10.5281/zenodo.5940391

Creator: Philip M. Craig and Ed Hawkins

Title: Weather observations from the 1861–1875 Met Office Daily Weather Reports

URL: https://doi.org/10.5281/zenodo.5940391

Publication year: 2023

Resource type: Data files.

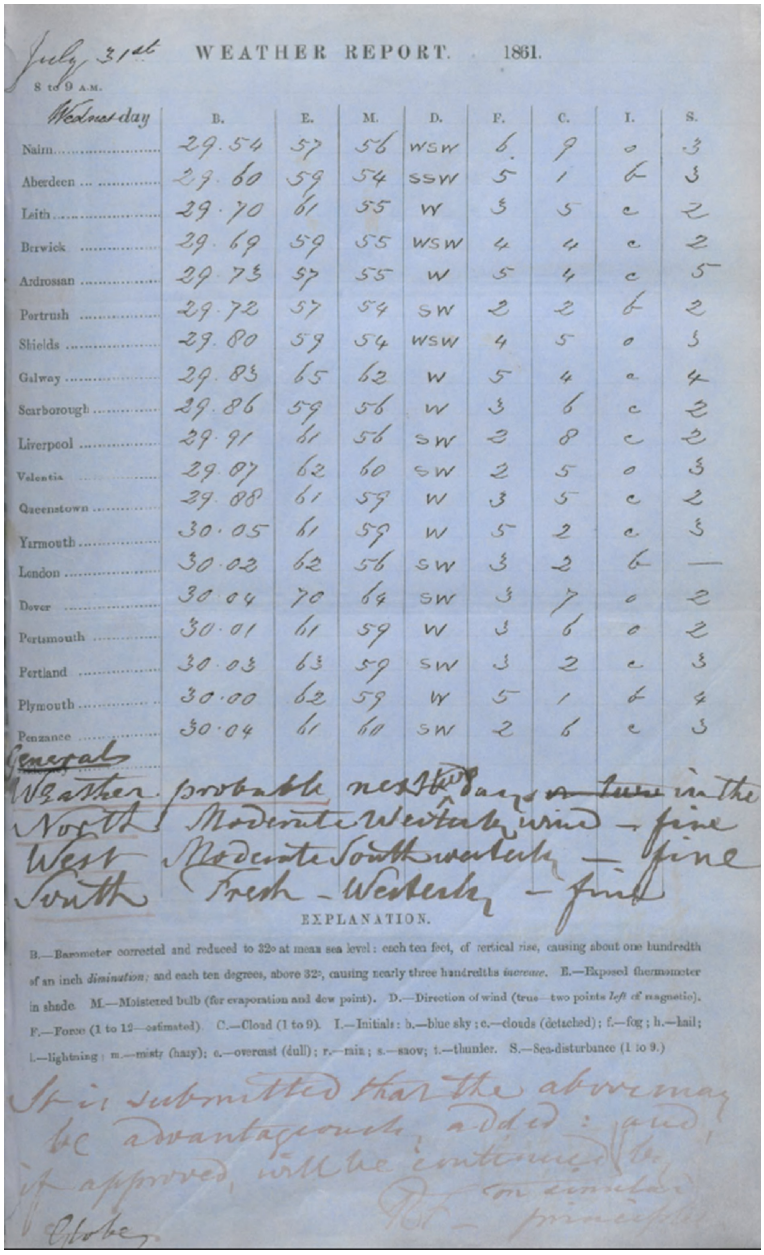
Version: 1.0

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2024 The Authors. *Geoscience Data Journal* published by Royal Meteorological Society and John Wiley & Sons Ltd.

contemporaneous weather observations to understand the current state of the atmosphere and so provide warnings of extreme weather conditions in advance to avoid loss of life and valuable infrastructure such as ships and their cargo.

This endeavour, which ultimately led to modern weather forecasting, began in 1860 with simple observations of atmospheric pressure, temperature, winds and weather type being transmitted using a telegraph cable network from all corners of Great Britain & Ireland, and collated by Fitzroy each day on hand-written sheets of paper in London (see example in Figure 1). Over the subsequent decades, additional observations from more locations across the UK and Western Europe were added to the daily collection (see example in Figure 2). All these sheets – known as the Daily Weather Reports (DWRs) – were archived and have been scanned, but the observations contained on the pages have not generally been transcribed into digital formats to inform modern weather and climate science. There is a high degree of trust (Sieber et al., 2022) in these observations because they were taken by trained meteorologists with high-quality instruments and were used in real-time every day to make forecasts.

This study describes the efforts undertaken to recover some of these important coordinated historical weather observations from the archives and make them available to the global climate community. The approach taken is to use citizen scientist volunteers to transcribe the observations using an online platform, before undertaking manual quality control, in a similar approach to previous studies (e.g. Craig & Hawkins, 2020; Gergis et al., 2022; Hawkins et al., 2019; Lakkis et al., 2022; Sieber &



July 31st WEATHER REPORT. 1861.

8 to 9 A.M.

Wednesday	B.	E.	M.	D.	F.	C.	I.	S.
Nairn	29.54	57	56	WSW	6	9	o	3
Aberdeen	29.60	59	54	SSW	5	1	b	3
Leith	29.70	61	55	W	3	5	c	2
Berwick	29.69	59	55	WSW	4	4	c	2
Ardrossan	29.73	57	55	W	5	4	c	5
Portrush	29.72	57	54	SW	2	2	b	2
Shields	29.80	59	54	WSW	4	5	o	3
Galway	29.83	65	62	W	5	4	c	4
Scarborough	29.86	59	56	W	3	6	c	2
Liverpool	29.91	61	56	SW	2	8	c	2
Valentia	29.87	62	60	SW	2	5	o	3
Queenstown	29.88	61	59	W	3	5	c	2
Yarmouth	30.05	61	59	W	5	2	c	3
London	30.02	62	56	SW	3	2	b	—
Dover	30.04	70	64	SW	3	7	o	2
Portsmouth	30.01	61	59	W	3	6	o	2
Portland	30.03	63	59	SW	3	2	c	3
Plymouth	30.00	62	59	W	5	1	b	4
Penzance	30.04	61	60	SW	2	6	c	3

General Weather probable next two days in the North Moderate Westerly wind – fine West Moderate South westerly – fine South Fresh – Westerly – fine

EXPLANATION.

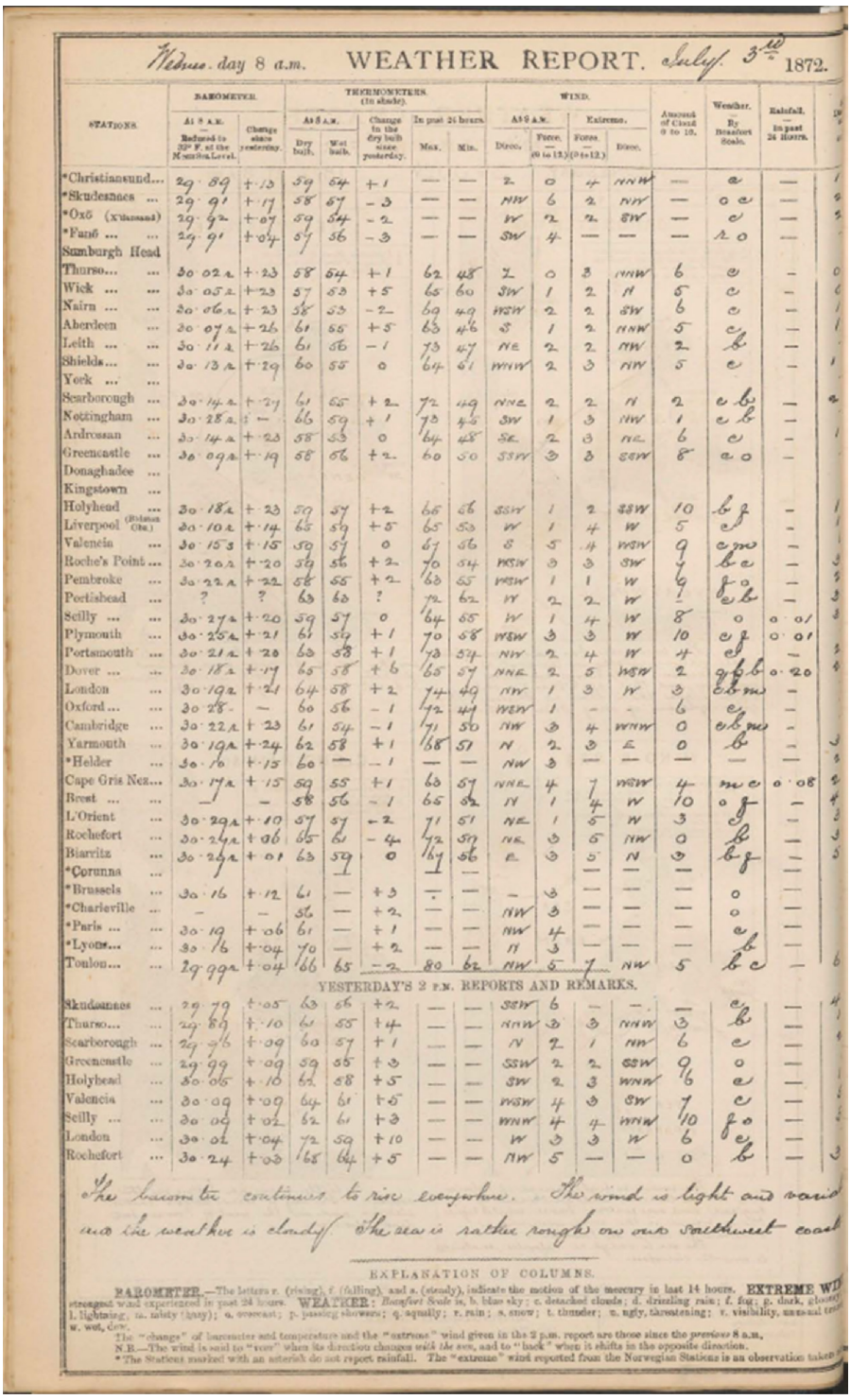
B.—Barometer corrected and reduced to 32° at mean sea level: each ten feet, of vertical rise, causing about one hundredth of an inch *diminution*; and each ten degrees, above 32°, causing nearly three hundredths *increase*. E.—Exposed thermometer in shade. M.—Moistened bulb (for evaporation and dew point). D.—Direction of wind (true—two points *left* of magnetic). F.—Force (1 to 12—estimated). C.—Cloud (1 to 9). I.—Initials: b.—blue sky; c.—clouds (detached); f.—fog; h.—hail; l.—lightning; m.—mistr (hazy); o.—overcast (dull); r.—rain; s.—snow; t.—thunder. S.—Sea-disturbance (1 to 9.)

It is submitted that the above may be advantageously added: and, if approved, will be continued by on similar principles

Globe

FIGURE 1 Example DWR from Wednesday 31st July 1861. Stations are listed on the left-hand side of the page and the observations are presented in columns of pressure (*B*), dry bulb temperature (*E*), wet bulb temperature (*W*), wind direction (*D*), wind force (*F*), cloud cover (*C*), weather type (*I*) and sea disturbance (*S*). This DWR also includes Fitzroy's first public weather forecast that was published in *The Times* newspaper.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Wednesday 8 a.m. WEATHER REPORT. July 3rd 1872.

STATIONS	BAROMETER. At 8 A.M. — Reduced to 32° F. at the Mean Sea Level.	BAROMETER. Change since yesterday.	THERMOMETERS. (In shade). At 8 A.M. Dry bulb.	At 8 A.M. Wet bulb.	Change in the dry bulb since yesterday.	In past 24 hours. Max.	In past 24 hours. Min.	WIND. At 8 A.M. Direc.	WIND. Extreme. Force. — (0 to 12.)	WIND. Extreme. Force. — (0 to 12.)	WIND. Extreme. Direc.	Amount of Cloud 0 to 10.	Weather. By Beaufort Scale.	Rainfall. — In past 24 Hours.	
*Christiansund...	29·89	+·13	59	54	+1	—	—	S.	0	4	NNW	—	a	—	1
*Skudesnæs ...	29·91	+·17	58	57	−3	—	—	NW	6	2	NW	—	a a	—	4
*Oxö (Christiansø)	29·92	+·07	59	54	−2	—	—	W	2	2	SW	—	a	—	3
*Fanö ...	29·91	+·04	57	56	−3	—	—	SW	4	—	—	—	2 0	—	
Sumburgh Head															
Thurso...	30·02 a	+·23	58	54	+1	62	48	S.	0	3	NNW	6	a	—	0
Wick ...	30·05 a	+·23	57	53	+5	65	60	SW	1	2	N	5	a	—	0
Nairn ...	30·06 a	+·23	58	53	−2	69	49	WSW	2	2	SW	6	a	—	1
Aberdeen ...	30·07 a	+·26	61	55	+5	63	46	S.	1	2	NNW	5	a	—	1
Leith ...	30·11 a	+·26	61	56	−1	73	47	NE.	2	2	NW	2	b	—	
Shields...	30·13 a	+·29	60	55	0	64	61	WNW	2	3	NW	5	a	—	1
York ...															
Scarborough	30·14 a	+·27	61	55	+2	72	49	NNE.	2	2	N	2	a b	—	2
Nottingham	30·28 a	—	66	59	+1	73	45	SW	1	3	NW	1	a b	—	
Ardrossan	30·14 a	+·23	58	53	0	64	48	SE.	2	3	NE.	6	a	—	1
Greencastle	30·09 a	+·19	58	56	+2	60	50	SSW	3	3	SSW	8	a o	—	
Donaghadee															
Kingstown															
Holyhead	30·18 a	+·23	59	57	+2	66	56	SSW	1	2	SSW	10	b f	—	1
Liverpool (Bidston Obs.)	30·10 a	+·14	65	59	+5	65	53	W	1	4	W	5	a	—	1
Valencia	30·15 3	+·15	59	57	0	67	56	S.	5	4	WSW	9	a w	—	
Roche's Point ...	30·20 a	+·20	59	56	+2	70	54	WSW	3	3	SW	4	b a	—	3
Pembroke	30·22 a	+·22	58	55	+2	63	55	WSW	1	1	W	9	f o	—	3
Portishead	?	?	63	60	?	72	62	W	2	2	W	9	f b	—	3
Scilly ...	30·27 a	+·20	59	57	0	64	55	W	1	4	W	8	o	o o/	3
Plymouth	30·25 a	+·21	61	59	+1	70	58	WSW	3	3	W	10	a f	o o/	
Portsmouth	30·21 a	+·20	63	58	+1	73	57	NW	2	4	W	4	a	—	3
Dover ...	30·18 a	+·17	65	58	+6	65	57	NNE.	2	5	WSW	2	g b b	o·20	4
London	30·19 a	+·21	64	58	+2	74	49	NW	1	3	W	3	b a w	—	
Oxford...	30·28 —	—	60	56	−1	72	44	WSW	1	—	—	6	a	—	
Cambridge	30·22 a	+·23	61	54	−1	71	50	NW	3	4	WNW	0	a b w	—	
Yarmouth	30·19 a	+·24	62	58	+1	68	51	N	2	3	E.	0	b	—	3
*Helder	30·16	+·15	60	—	−1	—	—	NW	3	—	—	—	—	—	3
Cape Gris Nez...	30·17 a	+·15	59	55	+1	63	57	NNE.	4	7	WSW	4	w a	o·08	2
Brest ...	—	—	58	56	−1	65	52	N	1	4	W	10	o f	—	4
L'Orient	30·29 a	+·10	57	57	−2	71	51	NE.	1	5	W	3	a	—	3
Rochefort	30·24 a	+·06	65	61	−4	72	59	NE.	3	6	NW	0	b	—	3
Biarritz	30·24 a	+·01	63	59	0	69	56	E.	3	5	N	3	b f	—	3
*Corunna															
*Brussels	30·16	+·12	61	—	+3	—	—	—	3	—	—	—	o	—	
*Charleville	—	—	56	—	+2	—	—	NW	3	—	—	—	o	—	
*Paris ...	30·19	+·06	61	—	+1	—	—	NW	4	—	—	—	a	—	
*Lyons...	30·16	+·04	70	—	+2	—	—	N	3	—	—	—	b	—	
Toulon...	29·99 a	+·04	66	65	−2	80	62	NW	5	7	NW	5	b a	—	6

YESTERDAY'S 2 P.M. REPORTS AND REMARKS.

Skudesnæs ...	29·79	+·05	63	56	+2	—	—	SSW	6	—	—	—	a	—	4
Thurso...	29·89	+·10	61	55	+4	—	—	NNW	3	3	NNW	3	b	—	
Scarborough	29·96	+·09	60	57	+1	—	—	N	2	1	NW	6	a	—	
Greencastle	29·99	+·09	59	55	+3	—	—	SSW	2	2	SSW	9	o	—	
Holyhead	30·05	+·10	62	58	+5	—	—	SW	2	3	WNW	6	a	—	
Valencia	30·09	+·09	64	61	+5	—	—	WSW	4	3	SW	7	a	—	
Scilly ...	30·06	+·02	62	61	+3	—	—	WNW	4	4	WNW	10	f o	—	
London	30·01	+·04	72	59	+10	—	—	W	3	3	W	6	a	—	
Rochefort	30·24	+·03	68	64	+5	—	—	NW	5	—	—	0	b	—	3

The barometer continues to rise everywhere. The wind is light and variable and the weather is cloudy. The sea is rather rough on our Southwest coast.

EXPLANATION OF COLUMNS.

BAROMETER.—The letters r. (rising), f. (falling), and s. (steady), indicate the motion of the mercury in last 14 hours. EXTREME WIND. strongest wind experienced in past 24 hours. WEATHER: Beaufort Scale is, b. blue sky; c. detached clouds; d. drizzling rain; f. fog; g. dark, gloomy; l. lightning; m. misty (hazy); o. overcast; p. passing showers; q. squally; r. rain; s. snow; t. thunder; u. ugly, threatening; v. visibility, unusual; w. wet, &c.

The "change" of barometer and temperature and the "extreme" wind given in the 2 p.m. report are those since the previous 8 a.m.

N.B.—The wind is said to "veer" when its direction changes with the sun, and to "back" when it shifts in the opposite direction.

*The Stations marked with an asterisk do not report rainfall. The "extreme" wind reported from the Norwegian Stations is an observation taken

FIGURE 2 Example DWR from Wednesday 3rd July 1872. Stations are listed on the left-hand side of the page and observations are presented from left-to-right in columns of pressure, temperature, wind, cloud, weather, rainfall and sea disturbance. The 2 PM data from the previous day are included in a separate table at the bottom of the page.

Slonosky, 2019). Such studies have demonstrated the value of recovering weather data from a broad range of locations and, by comparing it to long-term reanalyses, provide a stark contrast to a statement made in the 1868 Met Office Annual Report that remote stations are "of comparatively minor value."

We also compare the newly rescued observations with other datasets to evaluate their quality and make the data openly available. These data will enable improved reconstructions of changes in climate by being added to global databases, including the International Surface Pressure Databank version 4 (ISPD v4; Compo et al., 2019) and UK Met Office Integrated Data Archive System (MIDAS) (https://catalogue.ceda.ac.uk/uuid/dbd451271eb04662beade68da43546e1; Met Office, 2023). These observations will then be available for reanalyses such as the 20th Century Reanalysis (Slivinski et al., 2021) and gridded datasets such as HadUK-Grid (Hollis et al., 2019),

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

enabling improved reconstructions of long-term climate trends and also extreme weather events (e.g. Hawkins, Brohan, et al., 2023).

2 | THE DAILY WEATHER REPORTS

The collation of weather observations into the DWRs was started in September 1860 by Vice-Admiral Robert Fitzroy (Walker, 2011). The data on the original documents from September to mid-October 1860 are presented in a clear, tabulated format, but from mid-October 1860 to February 1861 the documents were not clearly written and the station locations are often too faint to be read, or even not written at all. As a result, we have not considered the DWRs before March 1861, but note that many of these earlier observations are available in archives of *The Times* newspaper, which published the observations daily, and could be transcribed from that source.

From 1st March 1861 until 13th January 1872, the DWRs collected observations of mean sea-level pressure (mslp), dry bulb temperature (T_d), wet bulb temperature (T_w), rainfall, sea disturbance and weather type at 8 AM each day in a consistent format. This period includes Fitzroy's first attempt at a weather forecast, handwritten on the DWR on 30th July 1861. His first public forecast was written in the DWR the following day (31st July 1861; Figure 1) and published in *The Times* newspaper the next morning. This period also includes the final DWR published during Fitzroy's life, as he died from suicide on 30th April 1865 (Paul et al., 2013).

No Sunday observations were collected in the early years, apart from three consecutive weekends in June/July 1862, then from October to December 1864, and once in December 1865. Sunday observations restarted in January 1867 and are available permanently from April 1867 onwards. There are also no DWRs for some public holidays, e.g. Christmas Day and Good Friday. The first foreign stations (Copenhagen, Helder, Brest, Bayonne, Lisbon) were included in the DWRs from August 1861 but did not report wet bulb temperatures.

The previous day's 2 PM observations at some stations were published from November 1867 onwards. These were initially inserted as an extra sheet of paper, then included at the bottom of the main page from January 1868 (see Figure 2). Additions & corrections to the observations were included from 7th July 1868, initially daily with each DWR, then as a monthly list from April 1872, as part of a wider update to the DWR format (also see Craig & Hawkins, 2020).

From 14th January 1872 onwards, the daily maximum and minimum temperatures (T_x, T_n) were also published, along with the change in pressure from the previous day. From 1st April 1875 onwards, observations of mslp, T_d, wind and weather from the previous evening were also included. Data from the DWR sheets for 1900–1910 have previously been rescued (Craig & Hawkins, 2020), along with the sub-daily pressure data for 1919–1960 (Hawkins, Alexander, et al., 2022).

In this study, we describe the transcription and quality control of the weather observations contained in the DWRs for 1st March 1861 to 31st March 1875 inclusive. This is the period of easily readable daily morning and 2 PM observations before the format became more complicated with the inclusion of evening observations. We did not transcribe the winds, cloud information, weather type or sea disturbance observations because, given limited resources, we prioritized obtaining longer time series of the higher-priority variables. The images of the DWRs are openly available from the National Meteorological Library & Archive (NMLA; https://digital.nmla.metoffice.gov.uk/).

3 | DATA COLLECTION AND QUALITY CONTROL

3.1 | Method of data collection

The observations were transcribed by over 3500 volunteer citizen scientists using the WeatherRescue.org website between March and August 2019. The project started as part of British Science Week, which encouraged members of the public to get involved in a range of science activities. The website was built using the Zooniverse.org platform.

The set of pages were split into groups by time period with each set having a similar page layout and set of locations. Each group was completed in turn. Volunteers chose a location from a list, were shown a brief tutorial, and were then shown a random page from the group and asked to transcribe the weather observations for the chosen location.

Seven separate volunteers were asked to transcribe each value, which is larger than some other similar projects because it was considered that the handwriting was harder to read. If five or more volunteers agreed on the value it was provisionally accepted, else it was flagged for checking.

The project ended with pages from March 1875 as the page layout changed considerably after that date, with twice-per-day observations included, which would have required additional website development to ensure the pages could be transcribed. As the volunteer engagement had declined over time it was decided not to continue the project.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The choice to split the transcription by location meant that full-page images could be shown to the volunteers, rather than having to segment the images into different smaller components, reducing the time overheads on the project team. This was only possible because of the smaller amounts of data per page than in previous similar projects that have used the DWRs, such as that described in Craig and Hawkins (2020).

3.2 | Quality control

The quality control (QC) procedure for this dataset was broadly similar to the approach taken by Craig and Hawkins (2020) for a different set of the DWRs (1900–1910). A CSV (comma-separated values) spreadsheet was generated from the Zooniverse output as a reduced digital equivalent of each DWR page (Figures 1 and 2). Each spreadsheet was checked for flagged disagreement and the corresponding entry in the DWR was checked manually to determine if it was legible. If the corresponding DWR entry was unclear then the spreadsheet entry was deleted, but if the DWR entry could be read clearly then the spreadsheet entry would be updated accordingly. Flagged values that included a question mark and a numerical value (e.g. pressure value of 29.95) were entered into the spreadsheets as they were written in the DWRs so that in the final dataset files such values could be highlighted in the metadata column as having an uncertain value. This is an improvement on the Craig and Hawkins (2020) approach to delete such values.

The process to find errors in the spreadsheets was partially automated and partially manual. A simple script was used to examine each year's CSV files and print the names of all the files that contained flagged values. The process of comparing the flagged entries with the DWRs and making the appropriate updates was manual and time-consuming. A total of 2948 flagged issues were found in the spreadsheets – around 0.5% of the total number of observations. The combination of volunteers was therefore around 99.5% accurate.

The wet-bulb temperature data before 1872 were not chosen for digitization because not every station recorded these observations. During the QC phase, these data were initially added to the spreadsheets manually. However, this was exceptionally time-consuming so it was decided not to add the wet-bulb temperature data for the period after 1863 as these observations were considered to be of less importance than the pressure, dry-bulb temperature and rainfall measurements. A further complication was that from October 1862 the wet-bulb temperature data were recorded as the difference from dry-bulb temperature so had to be calculated at a later stage. There is therefore a gap in the wet-bulb temperature data for some stations, and it becomes fully available from 1872 onwards. The daily rainfall data were only requested to be digitized by the volunteers from 1864 onwards, and the 1861–1863 data were added manually when available. The 2 PM observations were only transcribed for March 1872 onwards, so for some stations undigitized afternoon observations may exist from 1868 until February 1872.

After assessing the values where there was no volunteer consensus, the data were checked for physical inconsistencies, specifically $T_n > T_x$, $T_w > T_d$ and implausible pressure values such as mslp < 28 inHg and mslp > 31 inHg. There were very few implausible pressure values in the spreadsheets and some were caused by mistakes in the initial QC phase or could be calculated correctly using the change in pressure over the previous 24 h which was published in the DWRs from January 1872. These checks also identified and enabled the correction of some temperature values in the spreadsheets due to transcription errors.

Craig and Hawkins (2020) deleted many flagged values from the spreadsheets where there were physical inconsistencies such as $T_n > T_x$ even if both values were written clearly in the DWR. However, this is not considered best practice (Brunet et al., 2020) so an effort was made in the QC phase for this dataset not to delete suspicious values if they were clearly written. Instead, the dataresq R package v1.0.0 (https://github.com/c3s-data-rescue-service/dataresqc) was used to flag such issues in the resulting output data files. This standard package tests data for various issues such as climatic outliers, duplicate dates, repeated values and WMO (World Meteorological Organization) gross errors (WMO, 1993).

The QC process resulted in a complete set of spreadsheets which were converted to individual Station Exchange Format (SEF; https://datarescue.climate.copernicus.eu/node/80) files, with one file per station per variable. These 371 files are made openly available (see Dataset details). We do not perform any homogenization on the data as this is more appropriate when these data are combined with other sources.

4 | DATASET DESCRIPTION

4.1 | Station locations and metadata

Data from 71 stations from Great Britain and Ireland (GBI), nine other European countries and Canada were recovered for this dataset (Figure 3). The stations included on the DWR changed from year to year and increased substantially as time progressed, from 25 to 50, with 10 stations reporting for the entire 14-years period. Some stations stopped reporting and were replaced by nearby

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

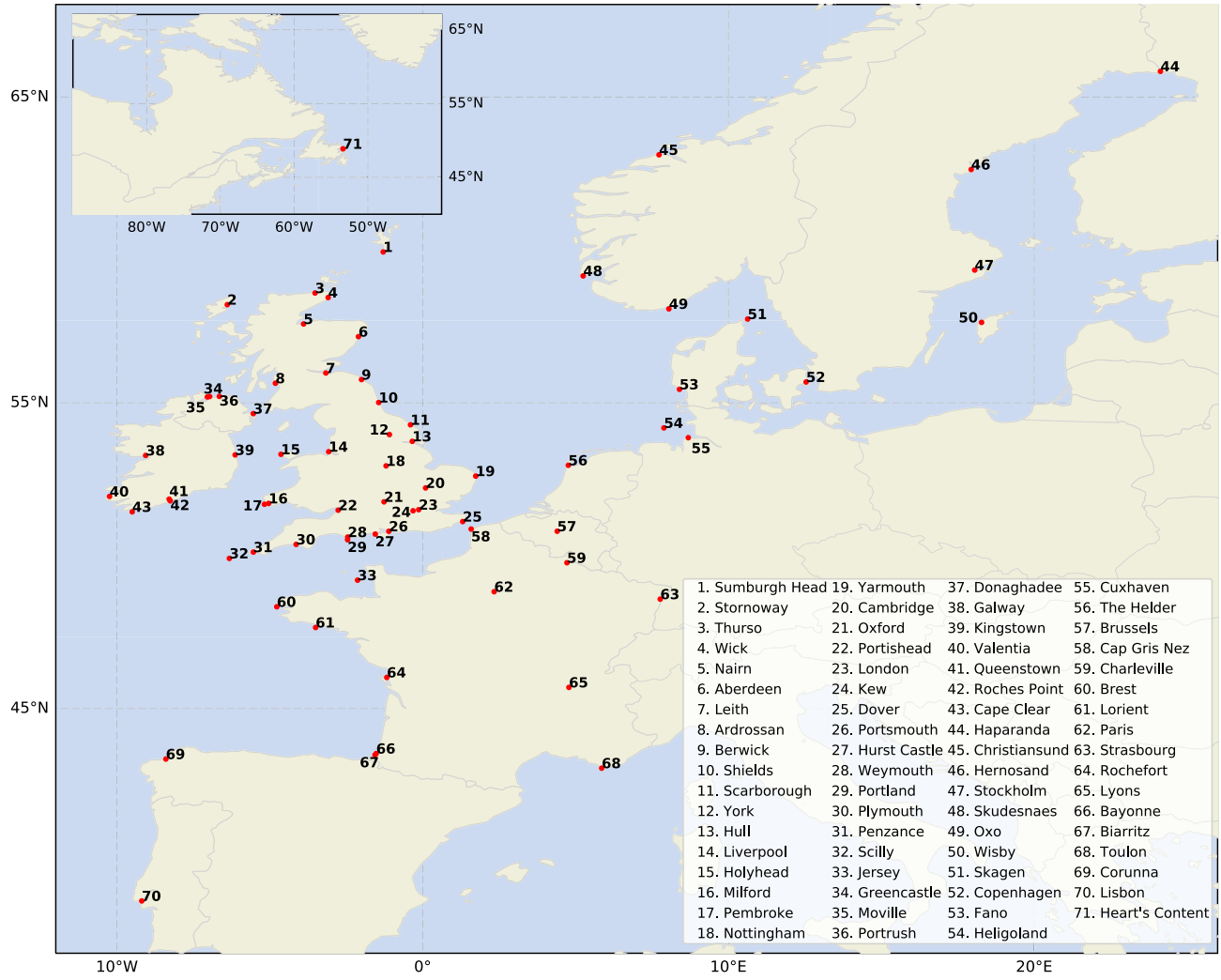


FIGURE 3 Map of all 71 stations included in the dataset from the 1861–1875 DWRs. The country borders shown are the modern-day boundaries.

stations or the equipment was moved to another location. For example, Greencastle was replaced by Moville in 1871, and Queenstown was replaced by Roches Point in 1872. Some stations dropped off the DWRs then reappeared later, such as Dover which reported 1861–1864 and again from 1871 onwards. The addition of Heart's Content in Canada to the DWR was a significant and historic inclusion with a station on the other side of the Atlantic Ocean but it dropped off the DWR after 2 years as the Meteorological Committee of the Royal Society was not prepared to pay the Anglo-American Telegraph Company for the observations.

4.1.1 | British and Irish stations

There were 43 different British and Irish stations included in the DWRs for at least some of the 1861–1875 period. The available metadata for the British and Irish stations is summarized in Table S1. The metadata were identified in a similar manner to Craig and Hawkins (2020) by using various documents made available by the NMLA and some other sources (Table 1). To supplement the available documentary information, various old maps were used to pinpoint a station's location after proving very useful for metadata detection by Hawkins, Burt, et al. (2022). Some local history groups were also contacted for additional information. For stations for which no useful information could be found (e.g. Hull, Berwick), approximate locations were estimated based on the sites of postal/telegraph stations from old maps as the observers were often telegraph clerks. The full list of sources used to locate the British and Irish stations is provided in Table S1.

The longitude and latitude coordinates of the 43 British & Irish stations are presented in decimal form in Table S1 and rounded to 2 decimal places. The Met Office Annual Reports, Climatological Returns and British Rainfall often

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 List of key documents that were used to find the metadata of the British & Irish stations in the 1861–1875 DWRs. Further sources specific to some sites are noted in Table S1.

Document type	Available information	Available years	Available from
Reports of the Meteorological Council to the President and Council of the Royal Society (Annual reports)	Exhaustive list of stations which reported to the Met Office Latitude and longitude coordinates	1857–1905/6	Online at NMLA
British Rainfall	Rain gauge elevation above sea level Information regarding the specific site of the station, for example street or building name	1860–1968	Online at NMLA as yearly documents
Climatological Returns	Station-specific observations Latitude and longitude coordinates Equipment elevations above sea level & ground	1857–1935	Online at NMLA
Old maps	Detailed Ordnance Survey maps of Great Britain and Ireland	1841–1971 1829-present 1832–1986 1847–1913	National Library of Scotland GeoHive Ordnance Survey of Northern Ireland Ordnance Survey of Ireland Historical Maps Boundary Finder v2.0

provided names and/or employment details of the observers which helped to infer the specific locations of the observing equipment. Some of the stations overlap with the 1900–1910 DWRs (Craig & Hawkins, 2020) so the same metadata were used for this earlier period (Table S1) but a considerable amount of research was required to find metadata for other stations.

Several stations had already been described in published meteorological literature: Nottingham (Lowe, 1846; Mellish, 1893), Galway (Hickey, 2003), Oxford (Burt & Burt, 2019) and Kew (Galvin, 2003). Scott (1873) mentioned the Portrush station to criticize the observer's "incorrigible idleness" but did not provide any details regarding the specific site of the equipment. There is also some information in non-meteorological literature, such as Smith (2014) who briefly discussed the Cape Clear telegraph station.

The lack of documentary evidence in the NMLA or published literature required a different approach for some stations. If the name and profession of an observer were available without a specific building or street, we contacted local history groups for further information. Shaw (1877) described the site of the Scarborough weather station but this was difficult to infer from old maps. However, the Scarborough History Society were able to find this specific location (later verified from a map in the 1880 station inspection reports, Table S1) and the nearby residence of the observer. Local history groups were also very helpful in determining the locations of the equipment or observers at Yarmouth and Sumburgh Head.

4.1.2 | International stations

Less metadata are available for the international stations since the information is not available in the NMLA. Despite this, the Annual Reports do contain some useful information about data quality or availability. For example, throughout 1869 the DWRs consistently have an empty row for Paris and other European stations. This may be a consequence of disruption caused by the Franco-Prussian War. Some stations match those from Craig and Hawkins (2020) so are given the same metadata. It is likely that some of the data rescued here for France duplicate data that already exist in the Météo-France digital data archives, but this comparison has not been conducted due to Météo-France's closed data policies.

The precise locations of the two Danish stations (Fanø and Copenhagen) are known from Danish Meteorological Institute (DMI) documents (Brandt, 1994a, 1994b). The 1878/79 Met Office Annual Report states that the Toulon station (68 in Figure 3) was at Cape Sicie and we have therefore inferred that this was the old semaphore station similar to other French stations (Craig & Hawkins, 2020). The Paris station was at Parc-de-Montsouris (inferred from the 1869–72 Annual Reports) where the observatory was located and there is still meteorological equipment. The 1869 Annual Report also mentions that the Oxö/Christiansand station (49 in Figure 3) was on Oxö island south of Christiansand (modern name Kristiansand) and we have assumed that the meteorological station was at the lighthouse.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The NMLA has two documents that discuss a weather station on Heligoland (54 in Figure 3; Kremser, 1891; Thurau & Kaufeld, 1990) but neither provide specific information on the location of the weather station before 1875 (the DWRs cover 1862–67). Kremser (1891) was published in the year after Heligoland was ceded back to Germany by the United Kingdom as part of the Heligoland-Zanzibar treaty. This change of sovereignty may explain why Kresmer did not publish earlier information as it was likely unavailable. The Annual Reports do not provide any information about the Heligoland site.

The Heart's Content Cable Station is a significant historical location as the site where the first ever transatlantic telegraph message was received (Muller, 2016). Observations were published in the DWRs from January 1868 and the Met Office Annual Reports note that the telegraph superintendent was the observer (J Weedon), but Muller (2016) names the superintendent as Ezra Weedon who is mentioned elsewhere (Cilento, 2017; Rowe, 2009) in connection with the cable station. The only other mention of meteorological observations that we could find is in Rowe (2009) who mentions the "small allowance" Weedon received for taking the observations and that in the winter of 1867–68 the thermometers provided for the observations "did not register low enough to track the temperature." Newfoundland was a British colony in the 19th century but the NMLA does not appear to have any relevant documents. We enquired with colleagues in Canada and local experts in Newfoundland but currently nothing further is known about the meteorological site which is surprising given the historical significance of the cable station.

4.1.3 | Observing times

There is very little known about the observing times for these observations. The GBI times are therefore taken as 8 AM, and the international stations follow the same times as Craig and Hawkins (2020). Observing times for Heart's Content were written on the DWR and changed from 6 AM local time to 9 AM local time in December 1868.

4.2 | Weather observations

A total of 569,345 observations of mslp, temperature and rainfall were digitized from the 1861–1875 DWRs using the approach outlined in Section 3. There are 142,259 mslp, 152,355 T_d, 79,445 T_w, 38,007 T_x, 38,188 T_n and 115,091 daily rainfall observations in the dataset.

We examine a small sample of these observations by comparing with existing overlapping datasets where possible, especially for pressure, daily temperature and rainfall.

4.2.1 | Pressure observations

Most of the mslp observations are once daily at 8 AM, but with some 2 PM observations included from 1872. The coverage of pressure observations for GBI has improved substantially compared to the current availability in ISPD across 1861–1875 with improvements in coverage across Europe too, particularly over France and Scandinavia (Figure 4). This provides a basis for better representation of weather events in extended reanalyses such as 20CRv3 (Slivinski et al., 2021) with reduced ensemble spread (Hawkins, Brohan, et al., 2023).

The quality of the newly recovered DWR pressure observations can be assessed by comparing time series at repeated sites across dates where there is overlapping data. For example, the EMULATE project (Ansell et al., 2006) recovered sub-daily pressure series for various locations across Europe and beyond. Not all of this data was used in ISPDv4. For 14 UK sites, the EMULATE pressure observations were transcribed from the DWRs, so provide a direct comparison with our recovered data from the same source. For the 4 longest overlapping series, the agreement is excellent (Figure 5) with differences of >1.6 hPa being present in less than 1% of observations. We choose this 1.6 hPa threshold as it is the assigned uncertainty for individual sea level pressure observations in 20CRv3. Note that for Greencastle, the EMULATE data includes 2 PM observations for 1868–1871 which were not transcribed in this study. This may cause the appearance of more frequent differences in the left column of Figure 5 than when comparing observations at the same time of day (right column). This comparison helps demonstrate the reliability of our volunteer-transcribed data.

In addition, there are three locations for which overlapping pressure data exists in ISPDv4 for the same site, but where the data likely came from different original sources. Specifically, Figure 6 considers these observations for Galway (1861 to 1864) and Stornoway (1873 to 1875). There are also data for Plymouth in ISPDv4 for which the agreement between the two sources is excellent, apart from a small offset (0.4mb), probably due to different station height corrections applied in two different sources (not shown). We also denote in Figure 6 whether the existing ISPDv4 observations were successfully assimilated or rejected by the 20th Century Reanalysis v3 (20CRv3; Slivinski et al., 2019, 2021). Observations that are rejected are those that appear to

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

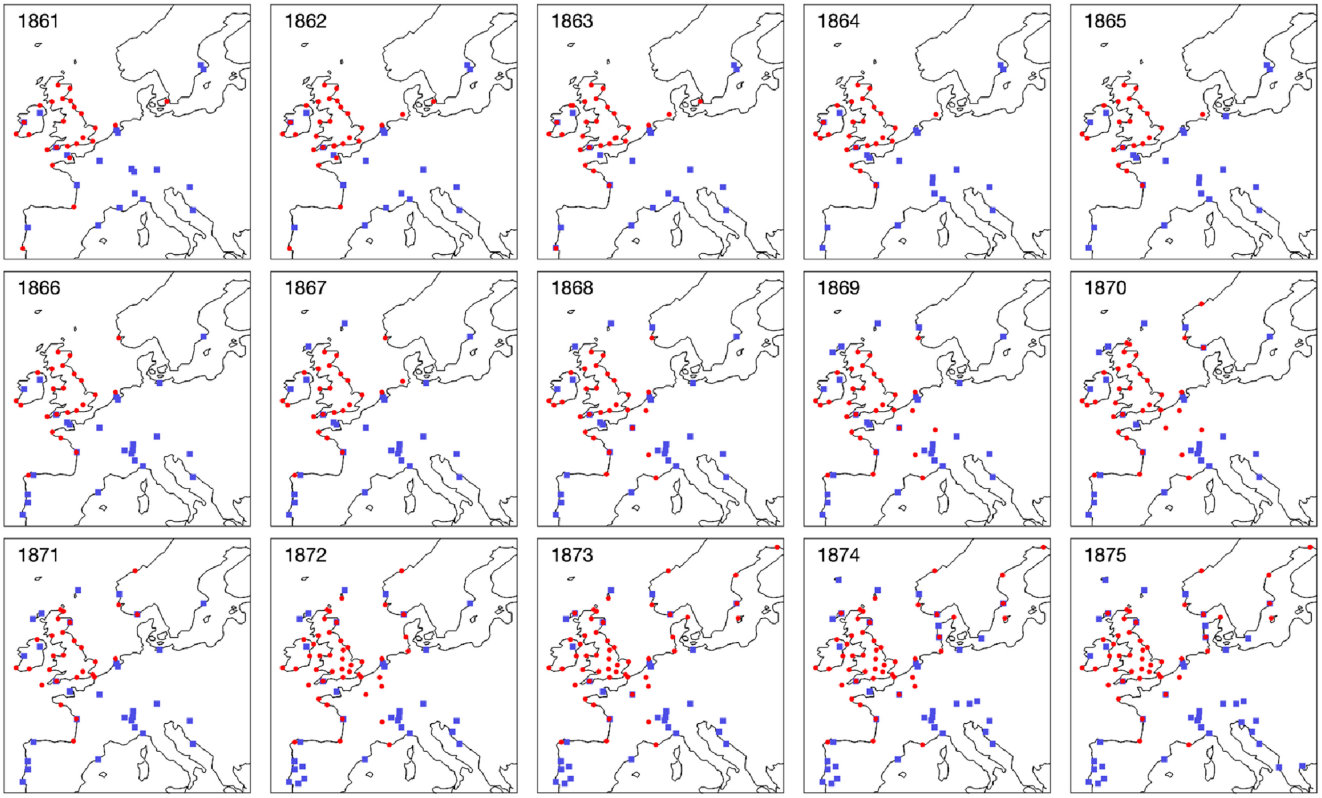


FIGURE 4 Land stations across Europe with pressure observations currently in ISPD (blue squares) and the stations from the DWRs (red circles) for each year in 1861–1875. Some of the DWR stations are the same locations as data already in ISPD but can be at different times of day. Each symbol represents a station with more than 30 observations during the year.

be inconsistent with the dynamical reconstruction constrained by nearby observations.

The mslp pressure observations from the DWRs and surface pressure observations from ISPDv4 for Galway (Figure 6) are rather inconsistent, apart from several months in 1862 where ISPDv4 surface pressure is always about 0.8 hPa greater than the mslp from the DWRs. It is unclear why this consistent offset exists since the mslp observations should be slightly larger for a given time and location. This suggests that there is an issue with the conversion to sea level in one or both sources but the altitude of the presumed Galway site is only a few metres above sea level. Twelve ISPDv4 Galway observations were rejected by 20CRv3 in the time period shown (red dots in right column of Figure 6). Further investigation of the Galway data is ongoing.

For Stornoway (Figure 6), there is better agreement between the data from two different sources. The DWR data is already converted to mslp, whereas the ISPDv4 data is for station pressure. This explains the small seasonal cycle in the differences, as the correction to sea level will be larger in summer than winter due to the warmer temperatures.

Overall, these comparisons highlight two locations where the existing ISPDv4 data appears very reliable (Plymouth, Stornoway) and one location which requires additional assessment (Galway), although it is not clear which dataset may be erroneous. However, the EMULATE project also recovered the same DWR data for Galway and the agreement with this study is excellent (not shown).

4.2.2 | Case studies: 20th January 1863 and 18th January 1872

One use of these new observations will be to better reconstruct the atmospheric circulation in reanalyses such as the 20th Century Reanalysis. We briefly highlight this potential using two case studies of severe storms. The 20CRv3 reanalysis consists of 80 ensemble members to represent the uncertainty in the reconstructed circulation patterns.

On 20th January 1863, a low-pressure system was located between Norway and the Shetland Isles in the 20CRv3 ensemble mean (Figure 7a). Zong and Tooley (2003) noted that on 19th and 20th January there was flooding on the coast of North-East England. On 21st January, *The Times* newspaper published various accounts of the storm, such as a "perfect hurricane" during the night in London and the highest tide for 6 years at Bristol (more than

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

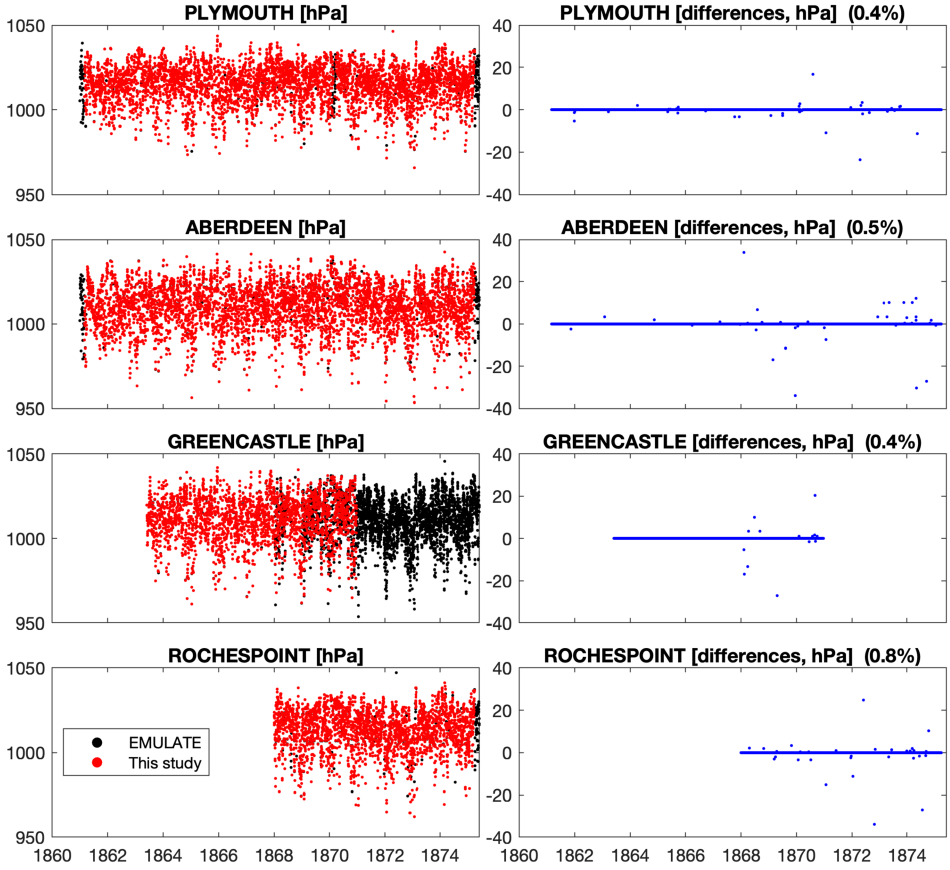


FIGURE 5 Comparing mslp observations from four sites from EMULATE and this study. The data came from the same source but were transcribed in two independent projects. The differences (EMULATE minus DWR) are shown in the right column and the percentages refer to the proportion of data which disagree by more than 1.6 hPa.

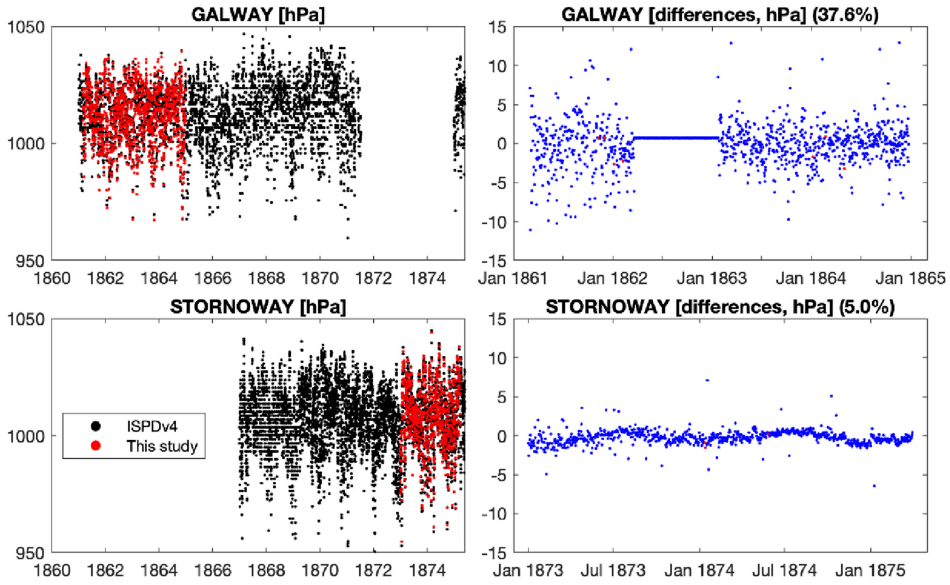


FIGURE 6 (left) Comparing pressure observations from two sites from ISPDv4 (black) and this study (red). The differences (right column; ISPDv4 minus this study) show substantial disagreements for Galway and a small seasonal cycle in Stornoway. The percentages in the right column refer to the proportion of data which disagree by more than 1.6 hPa.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

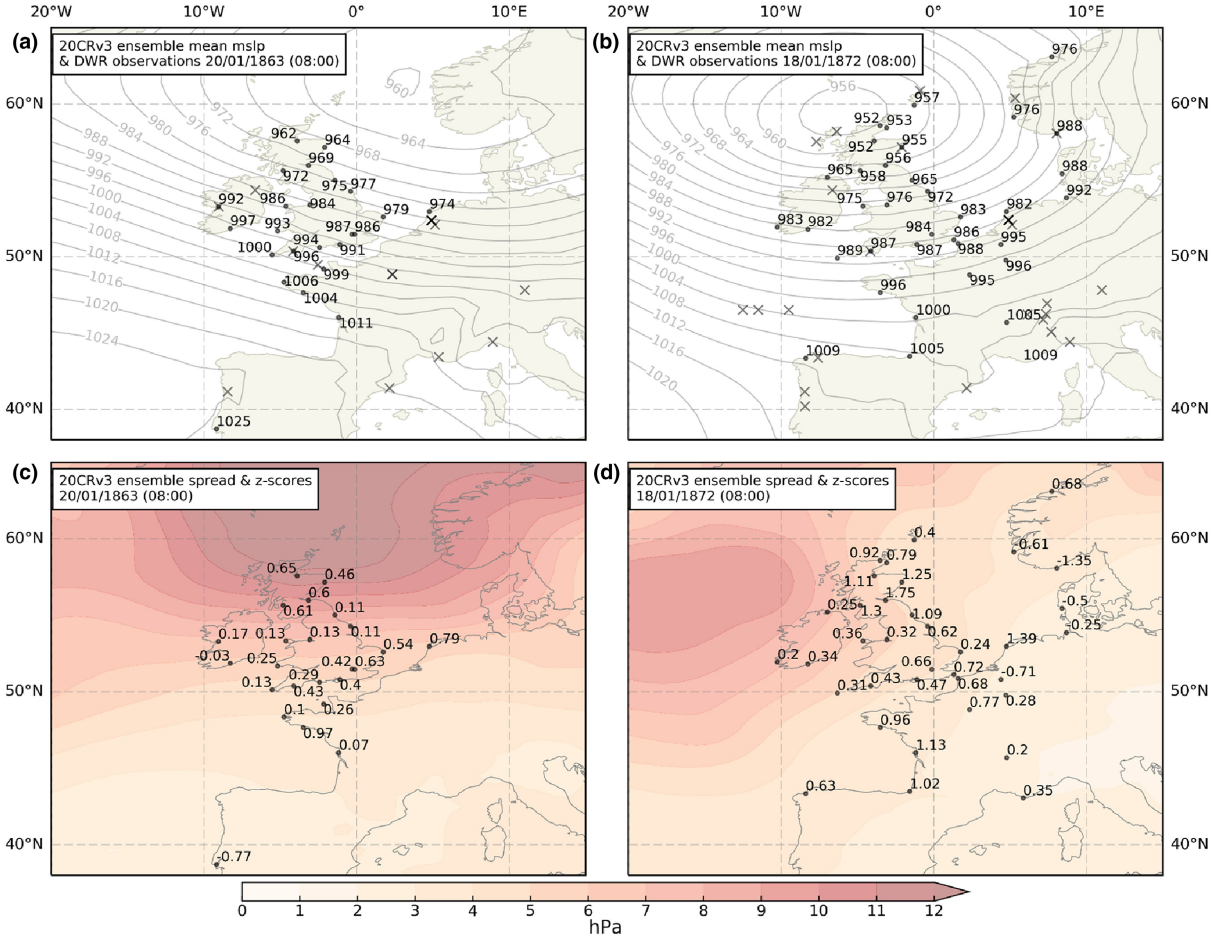


FIGURE 7 Ensemble mean 20CRv3 mslp (contours) and DWR mslp observations (grey circles and black numbers) with approximate gravity corrections applied (a,b) and 20CRv3 ensemble spread (filled contours) and *z*-scores (grey circles and black numbers) for 8 AM 20th January 1863 (a,c) and 8 AM 18th January 1872 (b,d). The time stated in each panel is the time of the British and Irish DWR observations, with observations from Europe within 3 h of the DWR observations also included. The grey crosses in panels (a) and (b) indicate the locations of land and sea pressure observations in ISPDv4 within 3 h of the DWR observations. Darker crosses indicate multiple observations at these locations within 3 h of the DWR observations. Units are hPa.

6 feet, roughly 2 m). Letters from several locations such as Dover, Shields and Liverpool were published in the newspaper that detailed damage to buildings and ships, thunder and lightning, railway disruption and high tides from the strong winds (The Times, 1863). The 1863 Met Office Annual Report also documented "violent and universal gales" for 3 days from 19th January along Western Scotland and the Irish Sea. There were storm warnings for all coasts and a ship was driven into Belfast by the "stress of weather."

The 20CRv3 ensemble mean mslp for this storm is greater than the DWR observations over Scotland and Eastern England by up to 4 hPa when gravity corrections (Craig & Hawkins, 2020) are applied to the observations (Figure 7a), highlighting that the existing observations used in 20CRv3 (crosses) are not dense enough to accurately represent the severity of the storm. However, the ensemble mean mslp is closer to the DWR observations over South-West England (specifically Penzance and Portsmouth). The ensemble standard deviation is greatest north of Scotland with values in excess of 12 hPa (Figure 7c).

We also examine the reanalysis using the standard z-scores,

$$z = \frac{\overline{p_R} - p_0}{\sigma}$$

where $\overline{p_R}$ is the 20CRv3 ensemble mean mslp, p_0 is the observed DWR value, and σ is the ensemble spread (standard deviation). This approach calculates how many standard deviations separate the observation from the ensemble mean. For this case study, the largest z-scores at Lorient (0.97) and The Helder (0.79) with the smallest at Queenstown and Rochefort where $|z| < 0.1$ (Figure 7c). For a "reliable" ensemble, we would expect around one-third of the observations to be outside ±1 standard deviations but, in this

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

example, none are, suggesting that the ensemble spread is too large and so underconfident.

The 20CRv3 ensemble mean mslp was also compared to additional undigitized data from Shetland and Orkney. These data were obtained from Scottish Meteorological Society documents that are available online from the UK National Meteorological Library & Archive (NMLA). Kirkwall and Sandwick Manse on Orkney have mslp of 960.6 hPa when corrected for gravity and temperature, and Bressay Manse on Shetland has 957.1 hPa. This supports the comparison between the DWR observations and 20CRv3 that the ensemble mean mslp is ~4 hPa too high compared to the observations, but does suggest that the mslp gradient is approximately correct so the centre of the storm may be close to the correct location. The tightly packed isobars in Figure 6a indicate a strong northwesterly flow across GBI which is consistent with the anecdotal evidence of strong winds and reports of flooding (Zong & Tooley, 2003). Considering the sparsity of pressure observations in ISPDv4 (Figure 4) and size of the ensemble spread (Figure 7c), 20CRv3 represents this storm well, but that representation would likely improve if these additional observations were assimilated (Hawkins, Brohan, et al., 2023).

On 18th January 1872, there was a low pressure off North-West Scotland (Figure 7b) that was not mentioned by Lamb and Frydendahl (2005) in their catalogue of storms in this region. The 20CRv3 ensemble mean mslp suggests that the centre of the storm was between 956 and 952 hPa and the tightly packed isobars indicate strong westerly winds across GBI. Notes on the second page of the DWR support this with a mention of a "stiff gale from SW at almost all stations last night" that "veered to WNW at the Irish stations and Ardrossan" and was "blowing hard in the NW." The pressure gradient between Thurso and Rochefort (48 hPa, Figure 7b) is also noted on the second page of the DWR.

Gravity-corrected mslp from the DWR at Thurso and Nairn (952 and 953 hPa, Figure 7b) indicate that, similar to the previous case study, the ensemble mean 20CRv3 mslp is ~4 hPa too high. Nairn's mslp observation of 952 hPa suggests that the centre might be too far north. Further south, ensemble mean mslp at stations around the English channel are much closer to the DWR observations with the 20CRv3 values at Plymouth and Portsmouth within 2 hPa of the observations. The DWR observations at Nairn, Aberdeen, Leith, Scarborough, Oxo, The Helder & Rochefort are more than one standard deviation away from the 20CRv3 ensemble mean (Figure 7d). Valentia, Lyon, Yarmouth and Portrush have very small z-scores by comparison (less than 0.3), indicating that these observations are much closer to ensemble mean. Approximately 75% of the z-scores are within ± 1 which is closer to what would be expected, although the ensemble spread is probably slightly too large. In this case study, there are more observations used in the construction of 20CRv3 and the ensemble spread decreases from west to east with a region of 8 hPa ensemble spread over the Atlantic Ocean west of Scotland (Figure 7d).

4.2.3 | Temperature observations

We can also compare some of the rescued temperature observations with existing digitized records and with the 20CRv3 reanalysis. In Figure 8, we compare the DWR temperature observations for an example year (1873) to

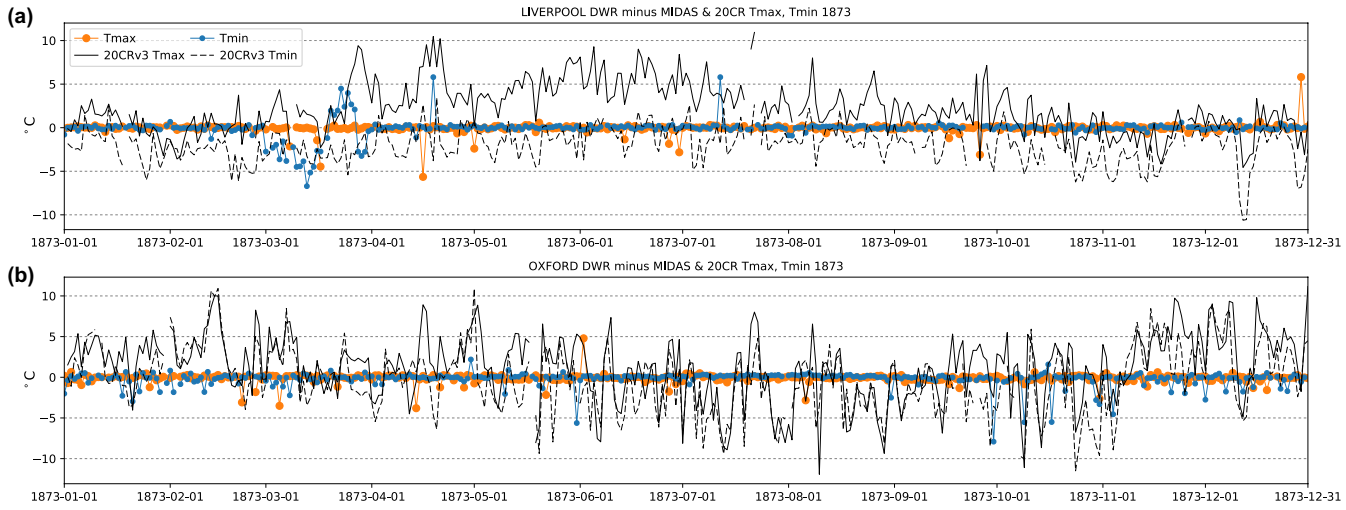


FIGURE 8 Comparison between DWR temperatures, MIDAS temperatures, and 20CRv3 ensemble mean. In (a) and (b) the differences between DWR and MIDAS T_x (orange) and T_n (blue) at Liverpool and Oxford in 1873 are shown along with the daily differences between the DWR values and 20CRv3 T_x (solid line) and T_n (dashed line). Breaks in the lines for the 20CRv3 temperatures indicate days where there are no DWR observations.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

data from the Met Office Integrated Data Archive System (MIDAS; Met Office, 2023) and the 20CRv3 ensemble mean for Liverpool and Oxford.

At Liverpool, the DWR and MIDAS observations are almost exactly the same for most of 1873 (Figure 8a) with only small differences or some notable outliers. However, throughout March the T_n differences change from negative (DWR < MIDAS) to positive (DWR > MIDAS) and back to negative before becoming approximately equal again in April. It is not clear what has caused these differences and we cannot confidently provide an explanation. There does not appear to be any common patterns between the MIDAS and 20CRv3 differences, although there is a strong warm bias in T_x from April to July. Since the 20CRv3 values are estimated using bilinear interpolation and the Liverpool site is close to the coast, there may be some influence from temperatures at grid points above the ocean in the calculation.

The differences in T_x and T_n at Oxford are approximately zero throughout 1873 (Figure 8b). There are some small negative T_n differences of less than 5°C in January, February, November and December with only four differences exceeding −5°C in the entire year. These larger differences in T_n also appear to coincide with some notable differences between the DWR and 20CRv3 in October which suggests that the DWR observations may be in error and too cold compared to MIDAS and 20CRv3. The T_x differences between DWR and MIDAS are also approximately zero throughout 1873 with one positive difference close to 5°C in June that occurs when the 20CRv3 T_x also exceeds the MIDAS value by a similar amount, suggesting that the DWR observation may be too warm for that day.

4.2.4 | Rainfall observations

A final comparison can be made with previously digitized daily rainfall observations. The only two stations in MIDAS that overlap with the DWR data are Oxford and Liverpool. Figure 9 compares the daily rainfall for 1873 for the two sites from the two sources. Overall, the agreement is good, with the root mean square differences (RMSD) between the series being around 1.4 mm, with similar values for 1874 (not shown). We are unsure what may cause differences as the original source for the MIDAS data is uncertain, but it is possible that the time of observation may be slightly different. There are also gaps in the DWR data and the MIDAS data is shown with black bars with circles on top to indicate those days.

5 | SUMMARY

We have described the citizen scientist transcription and quality control of around 570,000 instrumental weather observations taken between 1861–1875 at 71 locations, mainly across Great Britain & Ireland, and extending across north-western Europe with some limited data from Canada. This project has again demonstrated how volunteers can effectively recover detailed meteorological observations from archive material (Craig & Hawkins, 2020; Gergis et al., 2022; Hawkins et al., 2019; Lakkis et al., 2022; Sieber & Slonosky, 2019). Using two case studies, we have also highlighted that the rescued pressure data will likely allow improved reconstructions of severe storms during this period.

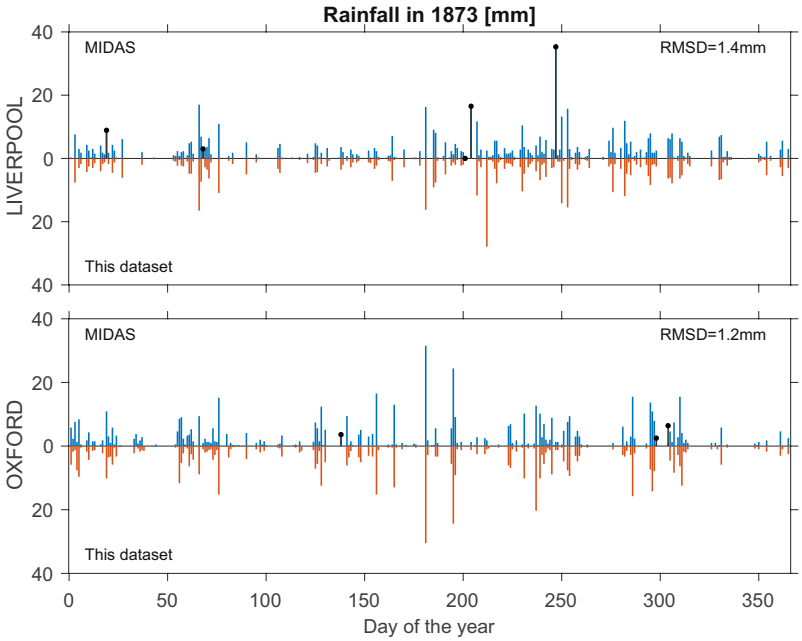


FIGURE 9 Daily rainfall in mm during 1873 for Liverpool (top) and Oxford (bottom). Both panels contain data from MIDAS (blue) and this dataset (orange). Days where there are observations for MIDAS but are missing from the DWRs are shown with black bars and circles. The RMSDs between the series are indicated.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Limited resources mean that the "Additions and Corrections," which are available from July 1868 onwards (see Section 2), have not yet been transcribed or applied to the data so there will be a small fraction of potentially erroneous data. However, the comparisons with existing data in Section 4 suggest that the newly rescued data are of high quality. The data recovered from this effort agree well with a small number of overlapping series rescued in other projects, but we have identified Galway as a location where more effort is needed to ensure the reliability of the existing and newly recovered data. Further improvements could be made to this dataset with regard to data that we were unable to digitize. For example, only ~50% of the Heart's Content pressure observations were digitized because they were transmitted too late to be published in the DWR so were later recorded in the additions section. These pressure data could be important for generating improved representations of the downstream cyclogenesis phase of North Atlantic cyclones in 20CRv3 and other long-term reanalyses.

ACKNOWLEDGEMENTS

We thank the more than 3500 volunteer citizen scientists who transcribed the pages, and without whom the digitization of this dataset would not have been possible. We also thank the UK National Meteorological Archive for scanning and providing access to the DWRs. PC and EH were both funded by the NERC GloSAT project, and EH was additionally supported by the UK National Centre for Atmospheric Science. We also thank Scarborough History Society, Great Yarmouth History Society and the Shetland Amenities Trust for providing valuable information on the location of some of the weather stations. This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and a grant from the Alfred P. Sloan Foundation.

OPEN RESEARCH BADGES

This article has earned Open Data, Open Materials and Preregistered Research Design badges. Data, materials and the preregistered design and analysis plan are available at [https://doi.org/10.5281/zenodo.5940391].

DATA AVAILABILITY STATEMENT

The dataset is openly available at https://doi.org/10.5281/zenodo.5940391.

ORCID

Ed Hawkins https://orcid.org/0000-0001-9477-3677

REFERENCES

Ansell, T.J., Jones, P.D., Allan, R.J., Lister, D., Parker, D.E., Brunet, M. et al. (2006) Daily mean sea level pressure reconstructions for the European-North Atlantic region for the period 1850–2003. *Journal of Climate*, 19, 2717–2742.

Brandt, M.L. (1994a) *The North Atlantic climatological dataset: Dokumenteret stationhistorie for 25140 Nordby, Fanø 1871–1994*. Technical Report. Danish Meteorological Institute.

Brandt, M.L. (1994b) *The North Atlantic climatological dataset: Dokumenteret stationhistorie for 30380 Landbohøjskolen, København 1861–1994*. Technical Report. Danish Meteorological Institute.

Brunet, M., Brugnara, Y., Noone, S., Stephens, A., Valente, M.A., Ventura, C. et al. (2020) *Best practice guidelines for climate data and metadata formatting, quality control and submission of the Copernicus climate change service data rescue service [Ref: C3S_DC3S311a_Lot1.3.4.2_2020_v4 - contract: 2020/C3S_311a_Lot1_Met Office/SC3]*. https://doi.org/10.24381/kctk-8j22

Burt, S. & Burt, T. (2019) *Oxford weather and climate since 1767*. Oxford: Oxford University Press.

Cilento, S.N. (2017) *From cod to conservation: the intersection of tourism, the cod moratorium, and telecommunications history in Newfoundland*. PhD thesis, Cornell University.

Compo, G.P., Slivinski, L.C., Whitaker, J.S., Sardeshmukh, P.D., McColl, C., Brohan, P. et al. (2019) *The International Surface Pressure Databank version 4, Research Data Archive. National Center for Atmospheric Research, Computational and Information Systems Laboratory*. https://doi.org/10.5065/9EYR-TY90

Craig, P.M. & Hawkins, E. (2020) Digitizing observations from the met Office daily weather reports for 1900–1910 using citizen scientist volunteers. *Geoscience Data Journal*, 7(2), 116–134. Available from: https://doi.org/10.1002/gdj3.93

Fitzroy, R. (1860) Remarks on the late storms of October 25-26 and November 1, 1859. *Proceedings of the Royal Society of London*, 10, 222–224. Available from: https://doi.org/10.1098/rspl.1859.0047

Galvin, J.F.P. (2003) Kew observatory. *Weather*, 58, 478–483.

Gergis, J., Baillie, Z., Ingallina, S., Ashcroft, L. & Ellwood, T. (2022) A historical climate dataset for southwestern Australia, 1830–1875. *International Journal of Climatology*, 41, 4898–4919. Available from: https://doi.org/10.1002/joc.7105

Hawkins, E., Alexander, L.V. & Allan, R.J. (2022) Millions of digitized historical sea-level pressure observations rediscovered. *Geoscience Data Journal*, 10, 385–395. Available from: https://doi.org/10.1002/gdj3.163

Hawkins, E., Brohan, P., Burgess, S., Burt, S., Compo, G., Gray, S. et al. (2023) Rescuing historical weather observations improves quantification of severe windstorm risks. *Natural Hazards and Earth Systems Sciences*, 23, 1465–1482. https://doi.org/10.5194/nhess-23-1465-2023

Hawkins, E., Burt, S., Brohan, P., Lockwood, M., Richardson, H., Roy, M. et al. (2019) Hourly weather observations from the Scottish highlands (1883–1904) rescued by volunteer citizen scientists. *Geoscience Data Journal*, 6, 160–173. Available from: https://doi.org/10.1002/gdj3.79

Hawkins, E., Burt, S., McCarthy, M., Murphy, C., Ross, C., Baldock, M. et al. (2022) Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists.

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Geoscience Data Journal, 10, 246–261. Available from: https://doi.org/10.1002/gdj3.157

Hickey, K.R. (2003) An assessment of the meteorological archives from N. U. I. Galway 1861-1966. *Journal of the Galway Archaeological and Historical Society*, 55, 117–126.

Hollis, D., McCarthy, M., Kendon, M., Legg, T. & Simpson, I. (2019) HadUK-Grid—A new UK dataset of gridded climate observations. *Geoscience Data Journal*, 6, 151. Available from: https://doi.org/10.1002/gdj3.78

Kremser, V. (1891) The climate of Heligoland [translated from German]. *Ann. d. Hydr. u. Mar. Met*, 5, 175–188.

Lakkis, S.G., Canziani, P.O., Rodriquez, J.O. & Yuchechen, A.E. (2022) Early meteorological records from Corrientes and Bahía Blanca, Argentina: initial ACRE-Argentina data rescue and related activities. *Geoscience Data Journal*, 10, 328–346. Available from: https://doi.org/10.1002/gdj3.176

Lamb, H. & Frydendahl, K. (2005) *Historic storms of the North Sea, British Isles and Northwest Europe*. Cambridge: Cambridge University Press.

Lowe, E.J. (1846) *A treatise on atmospheric phenomena*. Nottingham: Longman, Brown, Green and Longmans. Available from: https://www.google.co.uk/books/edition/A_Treatise_on_Atmospheric_Phenomena/NPUvAAAAYAAJ?hl=en&gbpv=0. [Accessed 1st December 2023].

Mellish, H. (1893) Rainfall of Nottinghamshire, 1861-90. *Quarterly Journal of the Royal Meteorological Society*, 19, 46–57.

Met Office. (2023) *Met Office MIDAS open: UK land surface stations data (1853-current)*. Available from: https://catalogue.ceda.ac.uk/uuid/dbd451271eb04662beade68da43546e1. [Accessed 13th February 2023].

Muller, S.M. (2016) *Wiring the world: the social and cultural creation of global telegraph networks*. New York, NY: Columbia University Press.

Paul, D.B., Stenhouse, J. & Spencer, H.G. (2013) The two faces of Robert Fitzroy, captain of HMS beagle and governor of New Zealand. *The Quarterly Review of Biology*, 88(3), 219–225.

Rowe, T. (2009) *Connecting the continents: Heart's content and the Atlantic cable*. St. John's, NL: Creative Publishers.

Scott, R.H. (1873) On some results of weather telegraphy. *Quarterly Journal of the Royal Meteorological Society*, 1(7), 181–188.

Shaw, F. (1877) The climate of Scarborough. *Quarterly Journal of the Royal Meteorological Society*, 3(20), 225–229.

Sieber, R. & Slonosky, V. (2019) Developing a flexible platform for crowdsourcing historical weather records. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 52(3), 164–177. Available from: https://doi.org/10.1080/01615440.2018.1558138

Sieber, R., Slonosky, V., Ashcroft, L. & Pudmenzky, C. (2022) Formalizing trust in historical weather data. *Weather, Climate, and Society*, 14, 993–1007. Available from: https://doi.org/10.1175/WCAS-D-21-0077.1

Slivinski, L.C., Compo, G.C., Whitaker, J.S., Sardeshmukh, P.D., Giese, B.S., McColl, C. 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, 2876–2908.

Slivinski, L.C., Compo, G.P., Sardeshmukh, P.D., Whitaker, J.S., McColl, C., Allan, R.J. et al. (2021) An evaluation of the performance of the twentieth century reanalysis version 3. *Journal of Climate*, 34(4), 1417–1438.

Smith, T. (2014) Gael or gall? Music and identity in early 1970s Cape Clear Island. In: Fitzgerald, D.M. & O'Flynn, D.J. (Eds.) *Music and identity in Ireland and beyond*. London: Ashgate Publishing Limited, pp. 205–218.

The Times. (1863) Violent weather. https://link.gale.com/apps/doc/CS202021941/TTDA?u=rdg&sid=bookmark-TTDA&xid=9918f423 [Accessed 18th August 2023].

Thurau, K. & Kaufeld, L. (1990) Die Wetterstation auf Helgoland [in German]. *Wettlotse*, 42(522), 200–204.

Walker, M. (2011) *History of the meteorological office*. Cambridge: Cambridge University Press.

WMO. (1993) *Chapter 6 - quality control procedures. Guide on the global data-processing system*. Geneva: World Meteorological Organization, No. 305, VI.1-VI.27.

Zong, Y. & Tooley, M.J. (2003) A historical record of coastal floods in Britain: frequencies and associated storm tracks. *Natural Hazards*, 29, 13–36.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Craig, P.M. & Hawkins, E. (2024) Digitizing observations from the 1861–1875 Met Office Daily Weather Reports using citizen scientist volunteers. *Geoscience Data Journal*, 00, 1–15. Available from: https://doi.org/10.1002/gdj3.236

20496060, 0, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/gdj3.236 by Test, Wiley Online Library on [30/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License