

# *Millions of digitized historical sea-level pressure observations rediscovered*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Hawkins, E. ORCID: <https://orcid.org/0000-0001-9477-3677>, Alexander, L. V. and Allan, R. J. (2023) Millions of digitized historical sea-level pressure observations rediscovered. *Geoscience Data Journal*, 10 (3). pp. 385-395. ISSN 2049-6060 doi: <https://doi.org/10.1002/gdj3.163> Available at <https://centaur.reading.ac.uk/105061/>

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/gdj3.163>

Publisher: Royal Meteorological Society

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](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online





## DATA ARTICLE

# Millions of digitized historical sea-level pressure observations rediscovered

Ed Hawkins<sup>1</sup> | Lisa V. Alexander<sup>2</sup> | Rob J. Allan<sup>3</sup>

<sup>1</sup>National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK

<sup>2</sup>ARC Centre of Excellence for Climate Extremes and Climate Change Research Centre, University of New South Wales, Sydney, New South Wales, Australia

<sup>3</sup>Met Office Hadley Centre, Exeter, UK

**Correspondence**

Ed Hawkins, National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK.

Email: [e.hawkins@reading.ac.uk](mailto:e.hawkins@reading.ac.uk)

**Funding information**

National Centre for Atmospheric Science, Grant/Award Number: S015574; Natural Environment Research Council, Grant/Award Number: GloSAT; University of Exeter; University of Southern Queensland; Newton Fund, Grant/Award Number: C3S2; Australian Research Council, Grant/Award Number: FT210100459 and CE170100023

**Abstract**

Millions of sub-daily sea-level pressure observations taken between 1919 and 1960 over the British and Irish Isles were transcribed from paper records in the early 2000s but were not published and subsequently forgotten. A chance discussion led to the rediscovery of the transcribed data and 5.47 million observations from 160 locations are now made available, although the data have not been fully quality-controlled. Much of the data are 3-hourly, allowing for detailed examinations of synoptic weather variations for this region and time period, and will be invaluable for constraining future reanalyses. We illustrate the value of the data using a stormy period during October and November 1928 and discuss the remaining quality-control issues.

**KEYWORDS**

data rescue, pressure, Ireland, UK

**Dataset**

Identifier: [10.5281/zenodo.6424344](https://doi.org/10.5281/zenodo.6424344)

Creator: Ed Hawkins

Title: Weather Rescue Data v1.0

Publisher: Zenodo

Resource type: Dataset

Publication year: 2022

Version: 1.0

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.

© 2022 The Authors. *Geoscience Data Journal* published by Royal Meteorological Society and John Wiley & Sons Ltd.

## 1 | INTRODUCTION

In 1887, John Venn used daily observations of atmospheric pressure from a single weather station in the United Kingdom to demonstrate that the natural world could exhibit non-Gaussian behaviour. Specifically, he highlighted that low pressures had a longer tail than high pressures (Venn, 1887). The significant impacts on society that can result from both extreme low and high pressure events emphasizes the importance of better quantifying the variations of atmospheric pressure on timescales from hours to decades. Long records of atmospheric pressure will help monitor whether the properties of such events are changing as the world continues to warm (e.g. Allan et al. 2009; Woollings et al. 2018).

Measurements of surface or sea-level pressure also allow a picture to be constructed of the variations of the atmosphere from the surface upwards and, due to the relative simplicity of the measurement, this quantity has been observed for centuries. These pressure observations enable a detailed reconstruction of the variations in the weather as far back as 1836 (the 20th Century Reanalysis; Slivinski et al. 2021), in contrast to reanalyses of the modern era which use a much wider range of observations, but only extend back to about 1950 (e.g. ERA5; Hersbach et al., 2020). However, huge quantities of meteorological observations, including of atmospheric pressure, remain unavailable to science as they are still only recorded on paper in various archives or, if we are lucky, in scanned images of the original documents. Many recent projects have added to our databases of surface pressure observations (e.g. Freeman et al. 2016; Ashcroft et al. 2018), including through citizen science activities (e.g. Hawkins et al. 2019; Craig & Hawkins 2020), but gaps remain. It is essential to ensure that the data from any new transcription of observations to digital formats becomes widely available for use and scientific analysis.

This study briefly discusses a set of more than 5 million sea-level pressure observations which were manually transcribed from paper documents many years ago but were subsequently forgotten and remained unpublished until a fortuitous scientific discussion led to the rediscovery of the data. The necessary tasks to make the data available have now been completed.

## 2 | DATASET DESCRIPTION

### 2.1 | The Daily Weather Reports

The UK Met Office produced a Daily Weather Report (DWR) every day from September 1860 until 1980. These DWRs contain detailed weather observations, mainly

taken around the British and Irish Isles, but also including data from stations across Europe and beyond; these were transmitted to the Met Office each day, initially by telegraph. Scanned copies of the DWRs are now all online and freely available (Met Office Digital Library and Archive, 2022) and are a valuable source of historical weather data. For example, the temperature, rainfall and pressure observations from the 1900–1910 DWRs were recently digitized using citizen scientist volunteers (Craig & Hawkins, 2020). The format and information contained within the DWRs changed many times, from once-per-day observations taken at around 15 stations in the 1860s, to twice-per-day observations taken at around 50 stations in the early 1900s, and six- or eight-times-per-day observations taken at tens of stations in the 1920s onwards. Several stations are included for large fractions of this period (see Section 3.1).

Figure 1 shows an example DWR page from 5th April 1919, showing the stations from which eight sea-level pressure observations per day can be derived. Each station has a listing for 01Z, 07Z, 13Z and 18Z, with a pressure observation converted to sea-level (given to a precision of 0.1 mb) and a change in pressure over the previous 3 hr. This allows the pressures for 22Z, 04Z, 10Z and 15Z to be calculated, but with a small uncertainty as the change is only given with a precision of 1mb. Note that the rows are not always complete, highlighting missing data, especially for 01Z, and therefore also for 22Z the day before.

Figure 2 shows another DWR page showing observations from 28th December 1960, listing many stations with complex codes describing the various weather variables recorded. These variables include pressure observations every 6 hr and a pressure tendency over the previous 3 hr, allowing 3-hourly data to be similarly produced. After 1945, the times of observation included were a regular 3-hourly schedule (00Z, 03Z, 06Z, 09Z, 12Z, 15Z, 18Z and 21Z).

### 2.2 | Transcription and dataset rediscovery

In the early 2000s, the Met Office funded a commercial company to undertake the transcription from the paper records of the 3-hourly pressure observations taken at the British and Irish stations contained within the 1919–1960 DWRs. This transcription project did not include other types of weather observation or any data from the many ‘foreign’ stations also listed in the DWRs; those data, consisting of many millions of individual weather observations, remain largely unrescued. In addition, for 1919–1921, there is a separate section listing a group of around 25 UK and Ireland stations



**METEOROLOGICAL OFFICE, LONDON:**  
**DAILY WEATHER REPORT, BRITISH SECTION.**  
No. 20,953 B ... SATURDAY, APRIL 5<sup>TH</sup> 1919.

OBSERVATORIES AND STATIONS CONTRIBUTING REPORTS AT 1st, 7th, 13th AND 19th.

STATIONS.	Barom. at M.S.L.		WIND.		TEMP.		Weather.	% Humidity.	Visibility.	Height of Low Cloud.	Low Cloud.		Upper Cloud.		Sea.	Barom. at M.S.L.		WIND.		TEMP.		Weather.	% Humidity.	Visibility.	Height of Low Cloud.	Low Cloud.		Upper Cloud.		Sea.
	Change in 3 hours.	m. b.	Dir.	Force 0-12.	'F.	'F.					Form.	Am't 0-10.	Form.	Am't 0-10.		0-3.	0-9.	Change in 3 hours.	Dir.	Force 0-12.	'F.					'F.	Form.	Am't 0-10.	0-3.	
Houton B. (Orkney)	1029.6	-2	W	1	44	od	98	3	4500	St	10	-	-	6	1020.3	+2	WNW	3	43	47	98	0	4500	St.	5	A.Cu	2	-		
Eskdalemuir	1021.7	-1	SW	4	44	om	82	5	2-5000	St.	8	-	-	1	1022.7	-1	WSW	3	46	om	87	5	1 1/2-2000	St.	10	-	-	-		
Stranraer	1022.9	-2	WSW	1	51	cu	75	8	1 1/2-2000	St.	8	-	-	1	1022.2	-1	NW	3	48	od	92	5	1 1/2-2000	St.	10	-	-	-		
Lough Foyle	1024.3	0	W	5	45	od	98	0	4500	Nb	10	-	-	3	1023.6	0	W	5	44	of	98	0	-	-	-	-	-	-	3	
Valencia (Cahire'n)	1026.8	0	NW	3	52	bc	75	8	2-3000	Cu	6	-	-	3	1026.4	0	NW	2	49	0	98	4	1/2-1000	St.	10	-	-	-	3	
Anglesey (Llangefno)																														
Falmouth	1023.8	+1	SSE	1	49	c	65	6	-	Cu-Nb	8	-	-	x	1023.0	0	N	3	48	c	-	-	-	-	-	-	-	-	x	
Calshot (Hants)	1022.8	0	-	0	53	om	65	7	1-1500	Fr. Cu	5	-	-	9	1022.4	0	NE	2	50	om	75	4	5-7000	St.	10	-	-	-	1	
Capel (Folkestone)	1022.2	0	SSE	2	47	om	75	4	-	-	-	-	-	2	1022.7	+1	E	2	43	b	75	5	3-7000	St.	1	-	-	-	2	
Aberdeen	1016.1	-3	SW	2	56	bc	55	7	-	St. Cu	6	-	-	2	1015.5	-1	WNW	4	53	c	65	7	-	St. Cu	7	-	-	-	2	
East Fortune	1019.9	-2	WSW	6	51	c	65	4	5-7000	St. Cu	2	-	-	x	1018.9	0	WSW	6	49	c	87	4	2-3000	Nb	6	A. St.	4	x		
Howden	1021.3	-2	WSW	2	53	b	65	4	5-7000	Fr. Cu	1	-	-	2	1022.2	0	SW	4	50	c	65	4	5-7000	St.	4	-	-	x		
Felixstowe	1023.6	0	-	0	48	0	87	5	5-7000	St.	5	-	-	2	1023.1	0	SSE	1	43	bc	92	5	-	St. Cu	4	Cc	1	2		
Cranwell (Lines)	1022.3	-1	-	0	52	c	55	4	-	-	-	-	-	x	1022.4	0	-	0	49	b	75	4	3-7000	St.	2	Fr. Cu	1	x		
Benson	1022.8	-1	-	0	52	c	45	5	-	-	-	-	-	x	1022.2	0	N	1	49	0	65	5	-	St. Cu	8	-	-	x		
London (Kew Ob.)	1022.8	-1	NNE	2	52	c	45	5	-	St.	3	-	-	6	1022.5	0	NNE	2	53	0z	55	3	-	St.	5	Cc	4	x		
Houton B. (Orkney)	1011.0	-1	W	4	42	bc	98	6	1 1/2-2000	St.	10	-	-	-	999.7	-12	W	11	49	f, g	98	3	1-1000	Nb	10	-	-	-	6	
Eskdalemuir	1019.0	-4	WSW	4	44	om	92	5	-	St.	10	-	-	-	1016.0	-2	WSW	5	44	0	92	5	1 1/2-2000	Nb	10	-	-	-	x	
Stranraer	1017.8	-4	W	5	45	0	87	8	-	-	-	-	-	1017.8	-2	W	5	45	0	87	8	1 1/2-2000	St.	10	-	-	-	-		
Lough Foyle	1018.4	0	W	6	44	0	92	5	-	-	-	-	-	1018.4	0	W	6	44	0	92	5	1 1/2-2000	St.	10	-	-	-	-	3	
Valencia (Cahire'n)	1027.0	0	WNW	3	46	om	-	-	-	-	-	-	-	1025.1	-3	-	0	46	0	92	8	2-3000	St. Cu	10	-	-	-	-	3	
Anglesey (Llangefno)	1023.5	-2	SSW	2	44	0	82	2	4-5000	St.	10	-	-	5																
Falmouth	1024.4	-1	NNW	2	41	b	98	5	-	-	-	-	-	x	1023.0	-1	NNW	1	41	bz	98	5	-	-	-	-	-	-	x	
Calshot (Hants)	1024.2	+1	ESE	2	37	om	98	3	-	-	-	-	-	3	1021.9	-3	-	0	35	f, i	92	3	-	-	-	-	-	-	2	
Capel (Folkestone)	1021.1	-1	N	2	41	om	75	4	-	-	-	-	-	2	1021.1	-1	N	2	41	om	75	4	-	-	-	-	-	-	2	
Aberdeen	1012.8	-2	WNW	4	46	od	87	3	2-3000	Nb	10	-	-	-	1008.7	-2	WSW	2	50	c	75	7	-	St. Cu	6	A. Cu	1	2		
East Fortune	1016.1	-5	WSW	6	47	0	92	-	-	-	-	-	-	x	1013.0	-5	WSW	6	47	0	87	5	1 1/2-2000	Nb	7	A. St.	3	x		
Howden	1018.4	-2	W	1	40	bc	92	3	-	-	-	-	-	x	1018.4	-2	W	1	40	bc	82	3	1 1/2-2000	St.	4	Cc	2	x		
Felixstowe	1022.9	-1	-	0	39	b	98	-	-	-	-	-	-	2	1020.7	-1	WNW	2	39	0z	98	3	-	St. Cu	9	-	-	-	2	
Cranwell (Lines)	1021.0	-1	WSW	2	37	b	92	4	4-5000	Cu	2	-	-	1	1019.2	-3	WSW	2	42	bc	92	4	3-7000	St. Cu	2	Cc	3	x		
Benson	1020.5	-2	-	0	0	0	33	em	98	2	-	-	-	x	1020.5	-2	-	0	0	33	em	98	2	-	-	-	-	-	x	
London (Kew Ob.)	1021.2	-2	-	0	0	0	36	f, i	92	2	-	-	-	x	1021.2	-2	-	0	0	36	f, i	92	2	-	-	-	-	-	-	x

FIGURE 1 An example page from the DWRs for 5th April 1919 showing the locations with observations four times per day, but which also include a change in pressure over the previous 3 hr (columns 1 and 2 in each of the four sections). Note that the names of the locations are repeated for the top and bottom halves of the table

with observations four times per day, but these were not transcribed (around 70,000 observations). After transcription, the resulting dataset was partly processed, but not published or included in any national or international databases.

One of the main reasons the dataset production process was not completed in the early 2000s was that the 'additions and corrections' component of the DWRs, published once per month until the 1930s, were not transcribed. This is because the original paper copies of the DWRs had not been scanned at that time and could not be sent out for digitizing; the duplicate copies that were sent did not include the 'additions and corrections' section. All the DWRs are now scanned and online (Met Office Digital Library and Archive, 2022). This means that any observations that were initially missing (but subsequently

arrived after the daily publication of the DWR) could not be included in the dataset without considerable additional transcription. In addition, any observation found to be transmitted or written on the DWR incorrectly could not be corrected in the dataset. It is a significant undertaking to complete these quality-control tasks and the resources were not available at the time. The dataset was then largely forgotten.

In 2019, author EH gave a seminar, describing the recovery of the observations contained in the 1900-1910 DWRs using citizen science (Craig & Hawkins, 2020). Author LVA was in the audience and recalled leading the project at the Met Office many years earlier to digitize pressure data contained in the DWRs. Subsequent discussions led to the fortunate rediscovery of the 1919-1960 DWR pressure data on an old laptop owned by author



DAILY WEATHER REPORT OF THE BRITISH METEOROLOGICAL OFFICE

No. 36197

Date of Issue Thursday 29 December 1960

Table with columns for Station, Observations at 12h, Observations at 18h, and Observations during Day. Includes sub-columns for wind, pressure, temperature, and weather codes.

FIGURE 2 An example page from the DWRs showing observations for 28th December 1960 for the locations with pressure observations (columns PPP) four times per day, and which also include a barometric tendency (columns app) for the change in pressure over the previous 3 hr. These observations are also included when using specific codes described in the DWRs

LVA. The data have now been reprocessed into Station Exchange Format (SEF) files so that they can finally be included in updates to the International Surface Pressure Databank (ISPD, Compo et al. 2019) and developing databases such as C3S and GHCN-h.

2.3 | Quality control

The issues around the quality-control of the data remain, that is, that the 'additions and corrections' are not applied. It is the view of the authors that this issue is less relevant now than when the dataset was initially transcribed, mainly due to the subsequent development of centennial reanalyses such as the 20th Century Reanalysis (20CRv3, Slivinski et al. 2021). Such reanalyses will be a major user of the data, especially as 20CRv3 only assimilates historical pressure observations to produce dynamical reconstructions of past weather variations. These scientific

developments mean that pressure data are more valuable to climate science than previously, making it important to ensure data availability. In addition, the reanalysis assimilation process naturally down-weights or rejects observations that are likely to be erroneous.

The issues that will exist with small fractions of the dataset include:

- 1. Missing data which were never taken, or not transmitted to the Met Office
2. Late arriving data which are in the additions pages but not added here
3. Errors which are listed in the corrections pages but not corrected here
4. Measurement or writing errors which were not identified at the time
5. Errors made during the modern transcription from the hand-written sheets
6. Systematic biases

Some observations within (1) could be found in other sources such as the original logbooks of the station, if they can be located. In principle, issues (2) and (3) could be addressed with a significant time investment using the ‘additions and corrections’ pages; targeted efforts to address such issues for certain significant weather events may be worthwhile in future. Even if this process was undertaken, errors of types (4) and (5) will still exist, and these are harder to find and potentially more numerous. Examining individual timeseries for ‘jumps’ in pressure (inhomogeneities) may allow some of these errors to be identified through manual checking of the original DWRs; for example, this time-consuming process was performed by Alexander & Power (2009) for a station in Australia. Examples of probable type (4) errors are shown later, along with examples of type (2) missing observations. Systematic biases (type 6) could exist due to, for example, an incorrect station elevation being used for the correction to sea level, or an incorrect calibration.

The use of these data within a reanalysis framework could pick up many of the errors by flagging individual observations which are rejected by a reanalysis assimilation process, but that is only possible after all the observations have been added to databases such as ISPD. An example of this identification for a type (5) issue from a different dataset is discussed in Craig & Hawkins (2020). Systematic biases are also estimated and removed when pressure observations are assimilated in 20CRv3 to account for type (6) issues.

For these reasons, the authors believe it is far better to produce a pressure dataset that is (say) 97% correct, 2% missing and 1% erroneous than no dataset at all, even

though we cannot be certain about the percentages that are missing or erroneous. We note that some users of the data may want to undertake their own quality-control procedures depending on the application.

### 3 | PRESSURE OBSERVATIONS: LOCATIONS, EXAMPLES AND COMPARISONS

#### 3.1 | Observations and locations

Figure 3 summarizes the number of stations and observations included in this dataset. A total of 5.47 million observations are made available from 160 locations covering different time periods. The first data available are in April 1919 and the last data are in December 1960. From April 1921 onwards, more than 40 stations have 3-hourly pressure data available, although this is often actually six times per day with 22Z and 01Z missing. From December 1943 to December 1948, around 70–80 stations are available, before this number drops to around 60. Three locations (Eskdalemuir, Valentia and Aberdeen-Dyce) have largely complete data for the whole time period with several other stations largely complete from April 1921 onwards. Figure 4 shows the locations where sub-daily pressure data are available in this dataset during example years between 1919 and 1960, and Figure 5 labels the locations mentioned in the text. We note that we have not identified precise coordinates for every station due to lack of metadata available

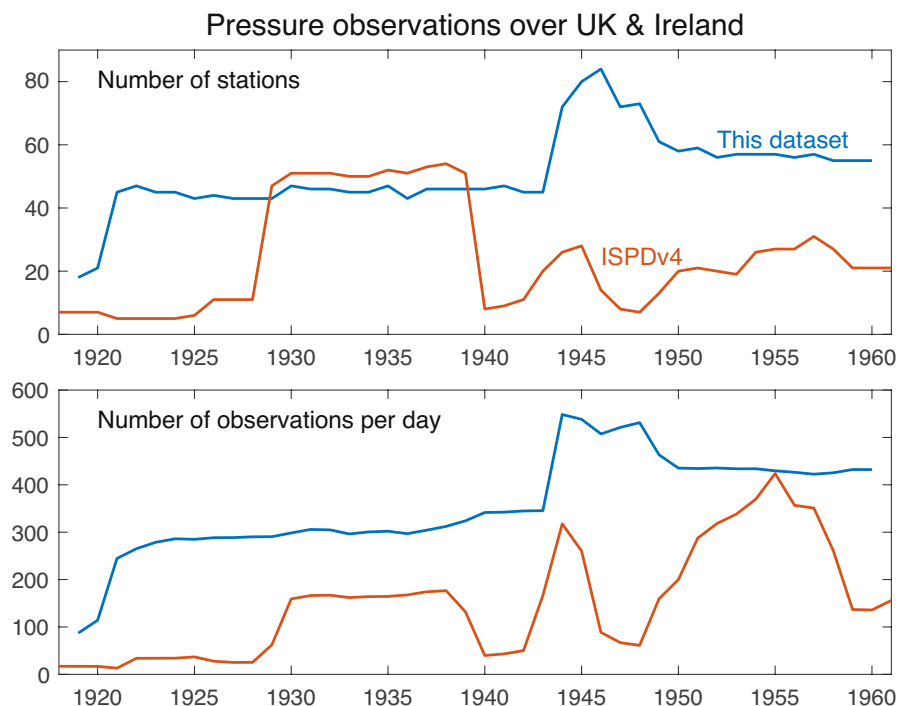
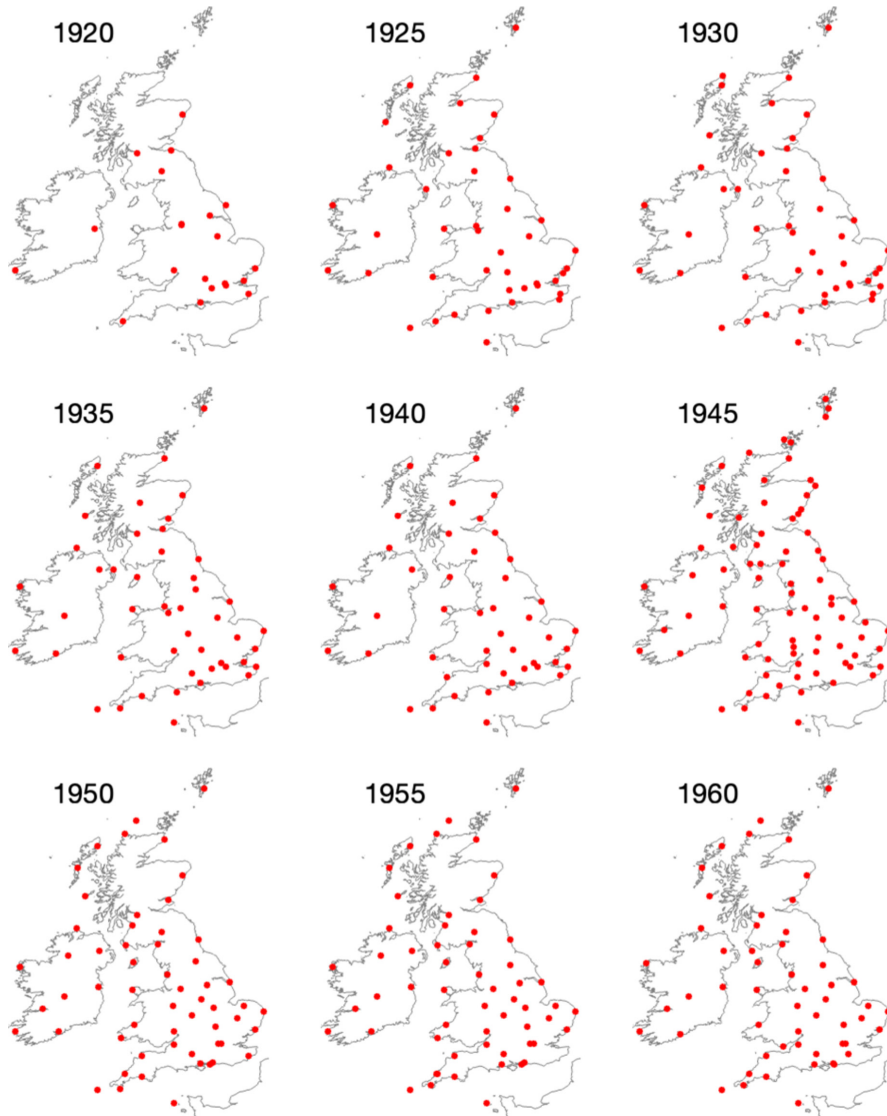


FIGURE 3 Number of stations for which observations have been recovered, and the average number of available observations per day in each year, for both ISPDv4 (existing; orange) and this dataset (blue). There is a considerable overlap between stations in the 1929–1939 period, although higher frequency observations are available in this new dataset



**FIGURE 4** Locations of the rescued data from example years between 1919 and 1960. Each red point represents a location with at least 200 observations during the year

on the DWR sheets. But, given that pressure is relatively insensitive to the site details, we do not consider this to be a serious issue.

Figure 3 also shows the number of stations and observations available for the United Kingdom and Ireland in the same period from ISPDv4. This dataset will increase the total number of observations available considerably. For the 1929–1939 period, there is overlap between the two datasets, but this dataset has higher frequency data from the stations that are already in ISPDv4.

### 3.2 | Example of observations available in 1928

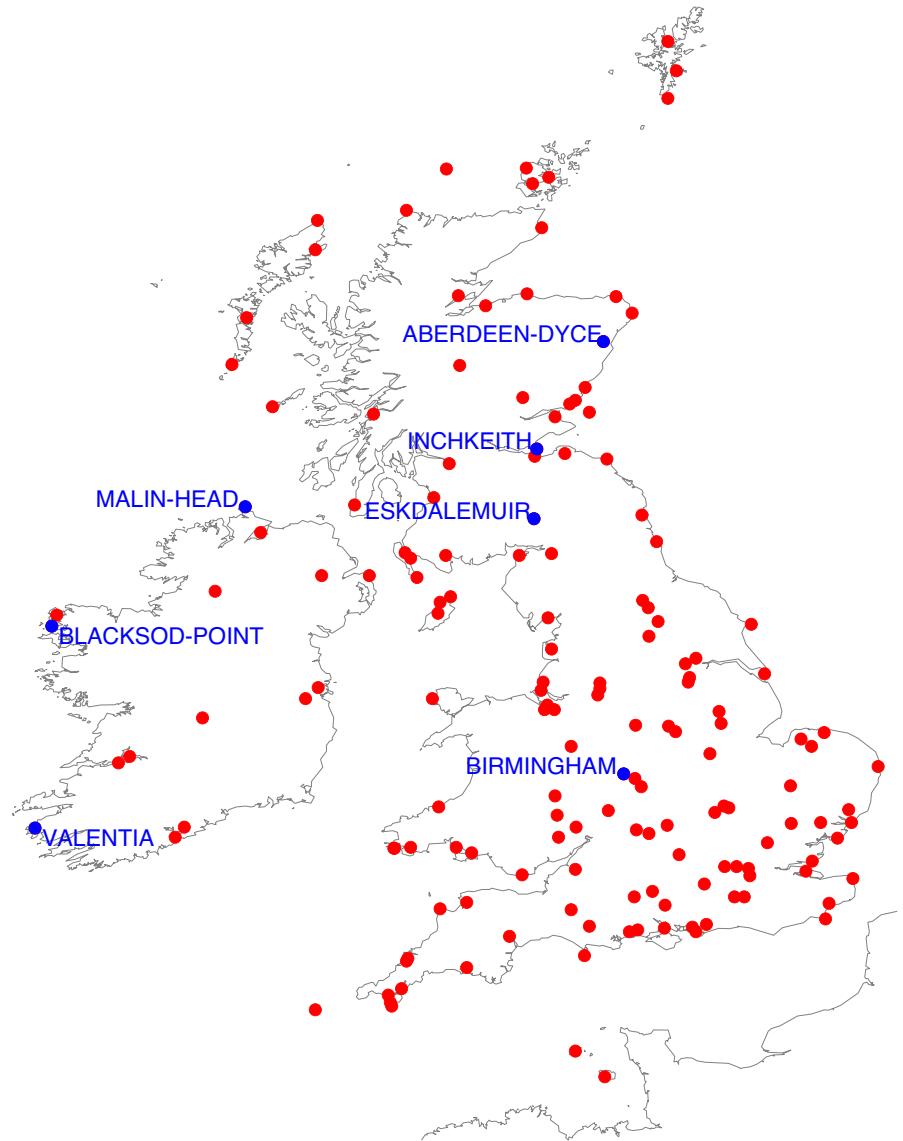
To highlight the ability of this dataset to provide information about synoptic variations, Figure 6 shows all the observations during October and November 1928. This was a stormy period, with at least three low pressure systems (<960 mb) moving over the British and Irish Isles. The top

row of Figure 7 maps the individual observations at the peak of those storms with darker blue colours used to denote lower sea-level pressures.

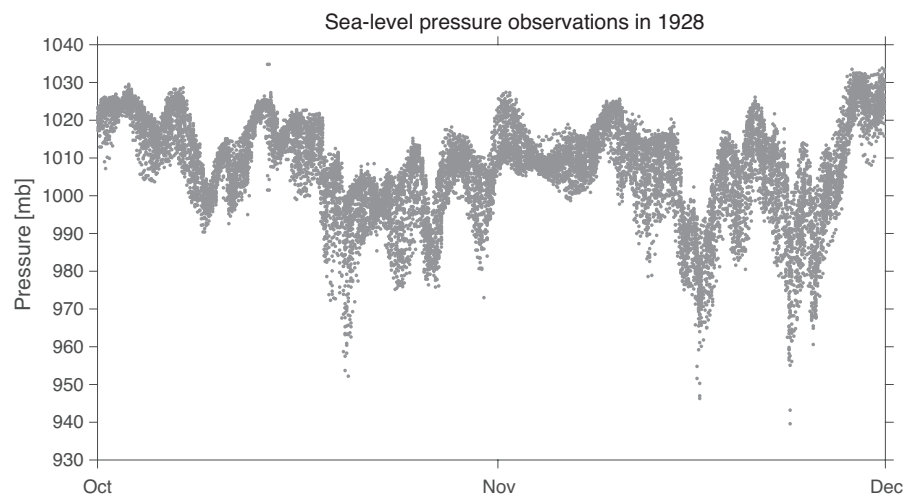
Note there are missing data for some of the times shown in Figure 7, that is, there are stations which are reporting for some events but not others. This is frequently seen around historical severe storms which caused delays to the transmission of the data. Some of these missing observations will be in the monthly ‘additions’ pages of the DWRs, and some appear written in red ink on the on-line copies of the DWRs, but these were not transcribed because the duplicate copy of the DWRs being used the earlier stage of the digitization process did not include them (see Section 2.2). For example, the missing observation at Eskdalemuir in southern Scotland at 15Z on 23rd November is 946 mb, with other missing observations in Ireland from Malin Head at 968 mb and Blacksod Point at 982 mb. Recovering such individual missing observations may be worthwhile if analysing case studies of particular severe storms.



**FIGURE 5** All locations appearing in the dataset (dots), with those mentioned in the text shown in blue and labelled

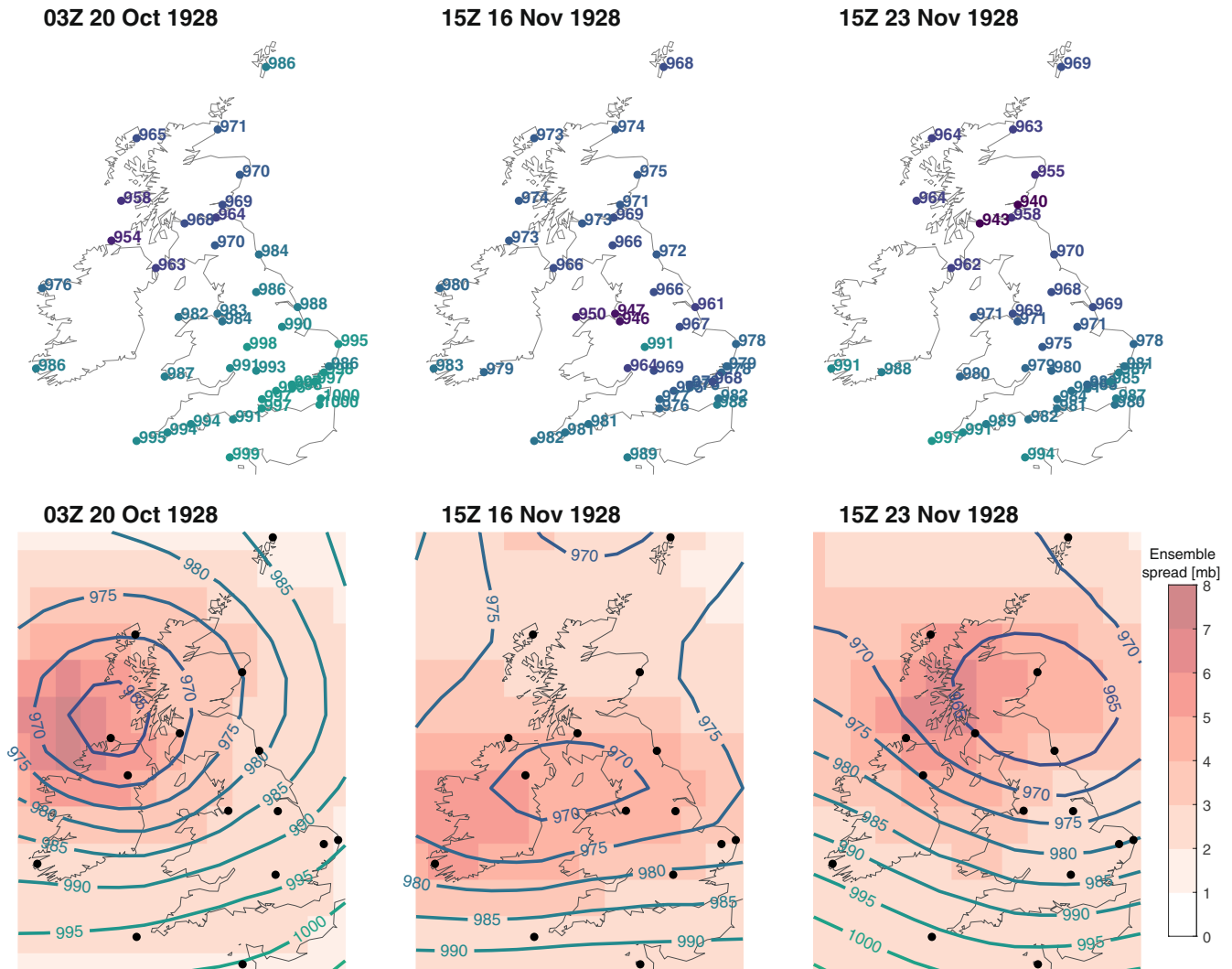


**FIGURE 6** Sea-level pressure observations for October and November 1928, highlighting a series of low pressure events over the British and Irish Isles

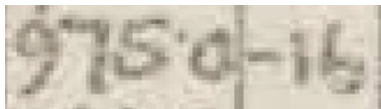


Note one almost certainly erroneous observation in the middle panel of the top row of [Figure 7](#). The 991 mb observation for Birmingham (south-east of the lowest pressure

values) at 15Z on 16th November 1928 has no correction listed in the DWRs and is correctly transcribed from the original DWR sheets. The 18Z observation is 975 mb, and



**FIGURE 7** (top) Map of locations and sea-level pressure observations from this new dataset for three low pressure events during October and November 1928. (bottom) Maps showing ensemble spread (red shades) and the ensemble mean of sea level pressure (contours) from 20CRv3 for the same times. The black dots indicate stations where some data exist for the same day in ISPDv4 and so were used to produce 20CRv3



**FIGURE 8** The 18Z pressure observation for Birmingham on 16th November 1928. Note that this is the resolution of the online scanned copies of the DWRs. Higher resolution images exist but are not online

this is indicated to be minus 16 mb from 3 hr earlier, resulting in a 991 mb observation for 15Z (Figure 8). It seems highly likely that the handwritten ‘-16’ should be ‘+16’, and that the 15Z observation was actually 959 mb, rather than 991 mb; this would fit the other available observations of the synoptic situation. There will be other

examples such as this in the dataset, but they would likely be rejected in a reanalysis assimilation. This is an example of issue (4) listed above and suggests that the data at times derived from both a transcribed observation and a change in pressure will contain more errors. The 958 mb at Inchkeith at 15Z on 23rd November also looks too high but is similarly transcribed correctly with no correction reported.

The bottom row in Figure 7 shows isobars for the same times from 20CRv3 (Slivinski et al. 2021) which only assimilates previously digitized pressure observations contained within ISPDv4. Just 16 observations from 11 locations (black dots) were available during each day for this time period (around half at 06Z, and none at 03Z or 15Z). These are enough data to place the low pressure centres in roughly the correct positions, but the severity of the

storms is not well represented. The ensemble spread (red shades) highlight that the uncertainty can reach  $\sim 7$  mb for some parts of the domain shown. With  $\sim 250$  observations per day from  $\sim 40$  locations now available for the same spatial region, subsequent reanalyses of these events will be much improved with a much reduced uncertainty.

The value of having spatially diverse and high-frequency data is clear, especially for storms which cross the country at night when traditional once- or twice-per-day climatological observations are not made.

### 3.3 | Observation comparisons

Comparing the locations of observations in the top and bottom rows of Figure 7 highlights that some of the newly digitized data come from places where observations already exist in ISPDv4 for at least some of the period. This allows a comparison of the two different digitizations of the same data to check accuracy. Figures 9 and 10 show comparisons for Eskdalemuir and Valentia, comparing ISPDv4 (grey) and this dataset (blue).

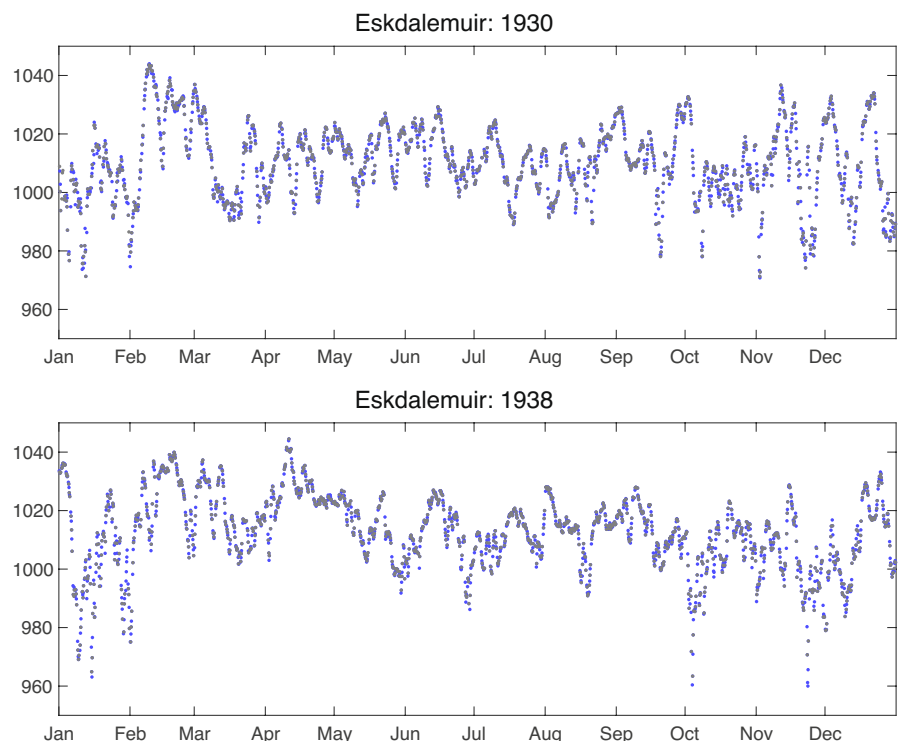
Overall, there is high agreement between the different digitizations. For Eskdalemuir, there are just 35 differences out of 10,790 overlapping observations covering 1929–1939, with an overall standard deviation between the two series of 0.2 mb. For Valentia, there are numerous small differences during 1919–1960, with an overall standard deviation of 0.7 mb between the time series. The standard deviation of differences is larger during

1922–1929 at 1.3 mb. It is unclear why, but we speculate that this could be due to variations in the conversion from station pressure to sea-level pressure in different sources. There could also be some errors in the ISPDv4 version of the data. There are 362 differences out of 94,392 overlapping observations which are larger than 5 mb for Valentia. Note that some of the Valentia data are already hourly within ISPDv4 so this dataset will not add much new information for this site.

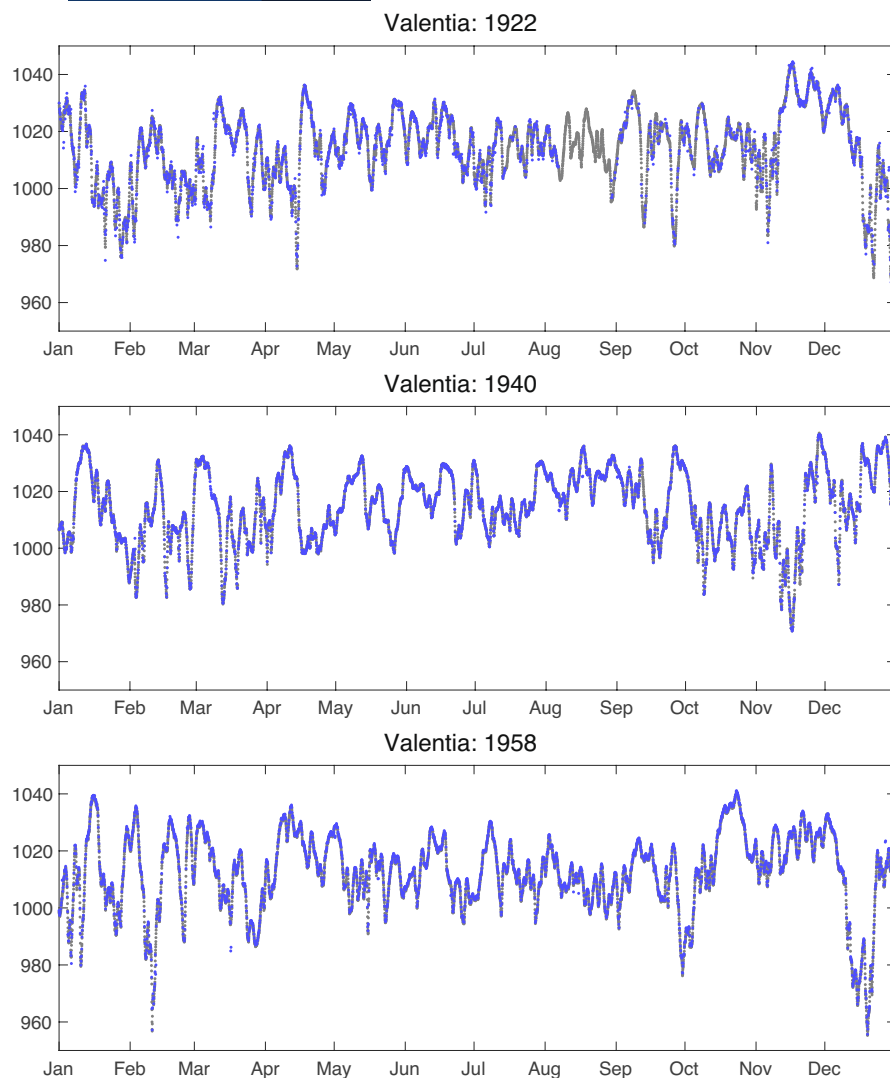
## 4 | SUMMARY

More than 5 million sub-daily sea level pressure observations are made available, from 160 locations around the British and Irish Isles between 1919 and 1960. If using these data, it is important to recognize that some quality-control procedures may need to be applied, depending on the specific use planned.

These data will be submitted to global datasets and therefore be available for use in projects such as future reanalyses of the historical period. Experiments are planned to include these data within dedicated simulations with both ERA5 (Hersbach et al. 2020) and 20CRv3 (Slivinski et al. 2021) reanalysis systems to demonstrate the value of such data for reconstructing the atmospheric circulation during particular extreme weather events. These additional data also have potential to augment time series of pressure for distinct locations in the United Kingdom, to understand storminess via gridded datasets or ‘pressure



**FIGURE 9** Comparing sea-level pressure data for Eskdalemuir during two example years from ISPDv4 (grey) and this dataset (blue). Note that this dataset contains 3-hourly data and ISPDv4 is 6-hourly and appears plotted on top of the data from this dataset



**FIGURE 10** Comparing sea-level pressure data for Valentia Observatory during three example years from ISPDv4 (grey) and this dataset (blue). Note that ISPDv4 contains hourly observations for Valentia for some of this period, but this dataset contains 3-hourly data and appears plotted on top of the ISPDv4 data

triangle' approaches (e.g. Wang et al. 2009; Cornes & Jones 2011), and for classification of weather patterns (e.g. Lamb weather types; Jones et al. 2014).

#### AUTHOR CONTRIBUTIONS

**Ed Hawkins:** Data curation (equal); formal analysis (lead); project administration (equal); visualization (lead); writing – original draft (lead); writing – review and editing (supporting). **Lisa Alexander:** Conceptualization (lead); data curation (equal); funding acquisition (lead); project administration (equal); writing – review and editing (supporting). **Rob Allan:** Data curation (supporting); project administration (supporting); writing – review and editing (supporting).

#### ACKNOWLEDGEMENTS



EH was funded by the UK National Centre for Atmospheric Science and the NERC GloSAT project (NE/S015574/1). LVA is supported by Australian Research Council grants CE170100023 and FT210100459. RJA is supported by funding from the UK Newton Fund [which is managed by

the UK Department for Business, Energy and Industrial Strategy (BEIS)], under its CSSP China and WCSSP South Africa projects, plus the EU Copernicus C3S2 311 Lot 1 Data Rescue Service. He also acknowledges the University of Southern Queensland, Toowoomba, Australia, and the Centre for Maritime Historical Studies, University of Exeter, Exeter, United Kingdom, where he is an Adjunct and Honorary Professor respectively.

#### OPEN PRACTICES STATEMENT

This article has earned an Open Data badge for making publicly available the digitally shareable data necessary to reproduce the reported results. The data is available at <https://doi.pangaea.de/10.1594/PANGAEA.936921>. Learn more about the Open Practices badges from the Center for Open Science: <https://osf.io/tvyxz/wiki>.

#### ORCID

*Ed Hawkins*  <https://orcid.org/0000-0001-9477-3677>  
*Lisa V. Alexander*  <https://orcid.org/0000-0002-5635-2457>



## REFERENCES

- Alexander, L. & Power, S. (2009) Severe storms inferred from 150 years of sub-daily pressure observations along Victoria's 'shipwreck coast'. *Australian Meteorological and Oceanographic Journal*, 58, 129–133. <https://doi.org/10.22499/2.5802.005>
- Allan, R., Tett, S. & Alexander, L. (2009) Fluctuations in autumn-winter severe storms over the British Isles: 1920 to present. *International Journal of Climatology*, 29, 357–371. <https://doi.org/10.1002/joc.1765>
- Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Brunet, M., Aguilar, E. et al. (2018) A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. *Earth System Science Data*, 10, 1613–1635. <https://doi.org/10.5194/essd-10-1613-2018>
- 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 at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/10.5065/9EYR-TY90>.
- Cornes, R.C. & Jones, P.D. (2011) An examination of storm activity in the Northeast Atlantic region over the 1851–2003 period using the EMULATE gridded MSLP data series. *Journal of Geophysical Research*, 116, D16110. <https://doi.org/10.1029/2011JD016007>
- Craig, P. & Hawkins, E. (2020) Digitizing observations from the met Office daily weather reports for 1900–1910 using citizen scientist volunteers. *Geoscience Data*, 7, 116–134. <https://doi.org/10.1002/gdj3.93>
- Freeman, E., Woodruff, S.D., Worley, S.J., Lubker, S.J., Kent, E.C., Angel, W.E. et al. (2016) ICOADS release 3.0: a major update to the historical marine climate record. *International Journal of Climatology*, 37, 2211–2232. <https://doi.org/10.1002/joc.4775>
- 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*, 6, 160–173. <https://doi.org/10.1002/gdj3.79>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) the ERA5 global reanalysis. *QJRMS*, 146, 1999–2049. <https://doi.org/10.1002/qj.3803>
- Jones, P.D., Osborn, T.J., Harpham, C. & Briffa, K.R. (2014) The development of lamb weather types: from subjective analysis of weather charts to objective approaches using reanalyses. *Weather*, 69, 128–132. <https://doi.org/10.1002/wea.2255>
- Met Office Digital Library and Archive (2022), 'Daily Weather Reports', Available at [https://digital.nmla.metoffice.gov.uk/SO\\_86058de1-8d55-4bc5-8305-5698d0bd7e13/](https://digital.nmla.metoffice.gov.uk/SO_86058de1-8d55-4bc5-8305-5698d0bd7e13/) (Accessed January 7, 2022)
- 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, 1417–1438. <https://doi.org/10.1175/JCLI-D-20-0505.1>
- Venn, J. (1887) The law of error. *Nature*, 36, 411–412. <https://doi.org/10.1038/036411c0>
- Wang, X.L., Zwiers, F.W., Swail, V.R. & Feng, Y. (2009) 'Trends and variability of storminess in the Northeast Atlantic region', 1874–2007. *Climate Dynamics*, 33, 1179–1195. <https://doi.org/10.1007/s00382-008-0504-5>
- Woollings, T., Barriopedro, D., Methven, J., Son, S.W., Martius, O., Harvey, B. et al. (2018) Blocking and its response to climate change. *Current Climate Change Reports*, 4, 287–300. <https://doi.org/10.1007/s40641-018-0108-z>

**How to cite this article:** Hawkins, E., Alexander, L.V. & Allan, R.J. (2022) Millions of digitized historical sea-level pressure observations rediscovered. *Geoscience Data Journal*, 00, 1–11. Available from: <https://doi.org/10.1002/gdj3.163>