# *Thunderstorm occurrence at ten sites across Great Britain over 1884–1993*

Received: 28 February 2019 | Revised: 14 May 2019 | Accepted: 20 June 2019

DOI: 10.1002/gdj3.75

DATA SERVICE PAPER

Geoscience Data Journal RMetS WILEY

# Thunderstorm occurrence at ten sites across Great Britain over 1884–1993

Maria Valdivieso | Mathew J. Owens | Chris J. Scott | Ed Hawkins | Stephen Burt | Philip Craig

Department of Meteorology, University of Reading, Reading, UK

**Correspondence**
Mathew J. Owens, Department of Meteorology, University of Reading, Reading, UK.
Email: m.j.owens@reading.ac.uk

**Abstract**

The UK Met Office's Daily Weather Reports (DWR) contain extensive logs of UK thunderstorm activity. To date, only a very small fraction of these data have been digitized as part of the MIDAS dataset, and exclusively after 1950. Using the recently-scanned UK Met Office Monthly Weather Reports (MWR), which are based on a subset of the observations that form the DWR, we here provide digitized data and a summary of thunderdays from 10 long-running British stations over the period 1884–1993. The data are presented 'as is', with no attempt to provide any corrections or calibration. For 4 of the 10 stations, thunderday observations were discontinued at various times between 1949 and 1964, and it is necessary to switch to a neighbouring station in order to continue the series. Approximately half the series exhibit sharp drops in thunderdays at various points between 1960 and 1990, although none are coincident with known station changes. Comparison with nearby MIDAS stations suggests the low thunderdays are the result of changes in observing practice, rather than genuine changes in thunderstorm occurrence. These potential data issues limit interpretation of the long-term trends. However, it can nevertheless be concluded that none of the stations show the expected increase in thunderdays as a result of the rise in surface temperature over the 20th century. In order to provide more quantitative determination of the long-term trends in thunderstorm occurrence, we advocate further digitization efforts to recover the data from the numerous stations in the MWRs, and subsequent analysis of the common signals across neighbouring stations.

**KEYWORDS**

climate, lightning, thunderstorms

**Dataset**
Identifier: https://dx.doi.org/10.5285/2503df629a8a485e8582e5150876b210
Creators: Valdivideso, M., Owens, M.J., Scott, C.J., Hawkins, E., Burt, S. and Craig, P.
Title: Thunderstorm occurrence at tens British sites over 1884-1993
Publisher: The Natural Environment Research Council's Data Repository for Atmospheric Science and Earth Observation
Publication year: 2019
Resource type: Dataset
Version: 1.0

# 1 | INTRODUCTION

Thunderstorms are responsible for some of the most extreme precipitation events, while the associated lightning strikes result in numerous fatalities, forest fires and infrastructure damage (Aich *et al.*, 2018). Thunderstorm occurrence and intensity are generally expected to increase with climate change through increased convective available potential energy (e.g. Romps *et al.*, 2014), though decreases from changes in cloud ice fluxes have also been projected (Finney *et al.*, 2018). It has also been suggested that lightning feeds back on climate, through the production of nitrogen oxides, which are strong greenhouse gases (Price *et al.*, 1997; Price, 2013). Interest in thunderstorm occurrence also extends to putative external (solar) drivers, which have been suggested on the basis of correlations over limited spatial regions and relatively short temporal scales (Stringfellow, 1974; Owens *et al.*, 2014).

Thus, it is important to monitor long-term trends in thunderstorm occurrence. There are a number of observational datasets which can be used for this purpose (Nag *et al.*, 2015). Radio networks, such as the UK Met Office Arrival Time Distance network (ATDnet; Lee, 1989) and the World Wide Lightning Network (WWLN; Rodger *et al.*, 2006), are able to determine individual stroke timing, position and intensity over a large portion of the globe. From a climatological perspective, however, these data have limitations. The automated forms of the networks were generally initiated in the 1990s or later, and as the networks have been expanded their sensitivity has increased, often resulting in an order of magnitude increase in number of detected flashes over the same geographic location. Alternatively, the optical signatures of lightning can be detected from space-based platforms (Nag *et al.*, 2015; Albrecht *et al.*, 2016). However, limited mission lifetime, changing instrument sensitivity, and even, the growth of urban lighting limit studies of long-term variability.

Alternatively, measurements of thunder can be used. 'Thunderdays' (Lewis, 1991; Burt, 2012; Prichard, 2016) are constructed by a human observer simply recording a 1 or a 0 depending on whether or not they have heard thunder that day (Brooks, 1925; Changnon, 1985; Kitagawa, 1989). These observations are independent of lightning measurements, be it visual identification of a flash or more recent radio-based observations. In many respects, thunderday observations are crude: They are local (an estimated 10–30 km audible radius of thunder), dependent on the diligence and hearing of the observer, and are low dynamic range, with storms producing 1 or 1,000 lightning strokes both registering a 1 in thunderdays. Anecdotally, thunderdays are sometimes criticized as

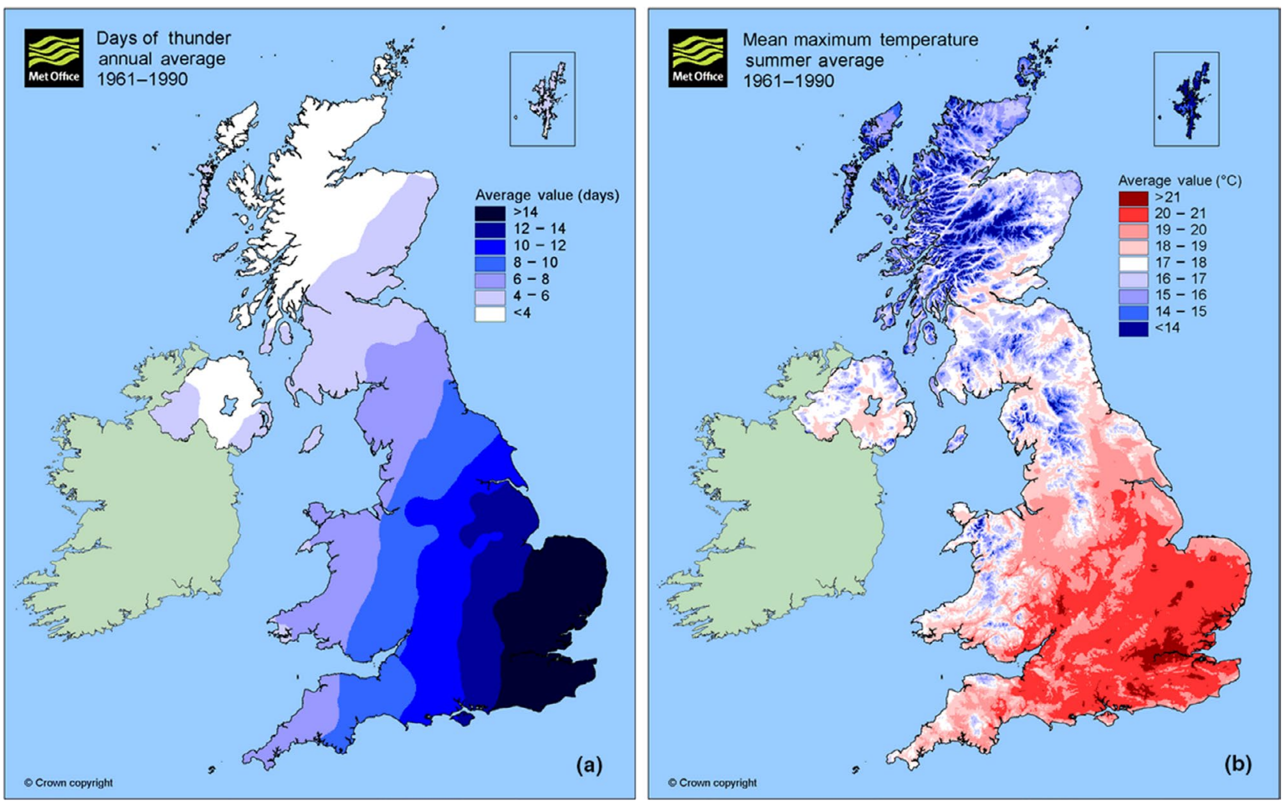


**FIGURE 1** Spatial distribution of mean climate parameters over the period 1961–1990, from gridded Met Office station data. (a) Thunderdays, annually averaged. (b) Mean maximum daily temperature, averaged over the summer months (JJA). Figures reproduced from: https://www.metoffice.gov.uk/public/weather/climate/

being subject to false positives from explosives or industrialization, but there is little evidence for this. Indeed, periods of expected false positives, such as from widespread firework displays on 5 November, so no such bias and may even lead to an under-reporting (Owens, 2016). Thus, the dominant dataset bias is likely to be false negatives from a lack of reporting of actual storms. Despite the limitations, thunderdays remain invaluable for being reported in approximately the same way for more than 100 years, providing the only possible means to quantify thunderstorm variability on decadal to centennial time scales. A number of long-term (here taken to be >50 years) thunderday records exist and have been analysed for distinct regions (Brooks, 1925; Changnon, 1985; Kitagawa, 1989; Enno *et al.*, 2014) with both positive and negative long-term trends reported.

For the UK, thunderdays from a limited number of Met Office stations have been digitized from the Met Office Daily Weather Reports (DWR) as part of the 'Met Office Integrated Data Archiving System (MIDAS) land and marine surface stations (1853–current)' dataset, made available by the NCAS British Atmospheric Data Centre (http://badc.nerc.ac.uk). Figure 1a shows the spatial distribution of annual mean thunderdays from the MIDAS dataset, over the period 1961–1990. There is a strong south-east to north-west trend, with more than 14 thunderdays per year observed on average in the south east, down to fewer than 4 thunderdays per year in north-west Scotland. Figure 1b shows the mean maximum daily temperature over the summer months (JJA), when the majority of UK thunder occurs. There's clearly a good deal of correspondence between maximum summer temperatures and thunderstorm occurrence, as expected. However, the east-to-west gradient appears stronger in the thunderdays, suggestive of an additional contribution from atmospheric dynamics.

In the MIDAS dataset, the earliest thunderday observations are 1950, where data are available from 9 stations, increasing to around 100 until 1970, before peaking around 400 stations through 2000 and gradually declining to around 100 by 2010 (Owens, 2016). The decline in number of stations reporting thunderdays is largely due to the adoption of automation at stations which were previously staffed by a human observer. It should be noted that no single station within the MIDAS dataset reports thunderdays contiguously through the 1950–2010 period, meaning construction of composite records is necessary for long-term studies.

The UK Met Office recently scanned the Monthly Weather Reports (MWR), which are based on the climatological data from selected stations which also appear in the DWR, over the period 1884–1993. Publication of the MWR ceased in 1993, though thunderdays are still recorded at some stations in the DWR. Comparison of the MWR with the current MIDAS dataset makes it clear that MIDAS only contains a small subset of the total thunderday data present in the DWR. Note that for reasons of data homogeneity, throughout this study we only consider thunderdays from MIDAS and MWR records (and not information about lightning via automated systems such as radio sferics, though that information will also be available to some stations after around 1957), and thus, all data here are based solely on manual human observers.

We select and digitize thunderdays from 10 long-running stations in the scanned MWR. It is hoped that this study will demonstrate the potential scientific value of these data, though also highlight issues with individual stations and hence provide the necessary motivation to rescue the remainder of the data. The study is structured as follows. Section 2 gives a summary of the thunderday station data digitized from MWR, along with details of the digitization procedure and additional MIDAS data used for validation. Section 3 provides an overview of the data in terms of annual and seasonal variations. Section 4 makes comparisons between MWR and MIDAS thunder data for the period of overlap (after 1957 for the stations considered). The discussion of issues with the data and possible ways forward is presented in Section 5.

## 2 | DIGITIZED THUNDER DATA FROM MWR

The scanned Monthly Weather Reports (MWRs) are publicly available from the Met Office Digital Library and Archive: https://digital.nmla.metoffice.gov.uk/collection_75a68cd2-cabe-43a8-98bb-3919f51e59a9/. MWRs contain a general summary of the monthly weather, including monthly maps and monthly-averaged station data. These data are based on a subset of the observations which form the DWR. Some of this data already been digitized and made available as part of the Met Office MIDAS database archived at the British Atmospheric Data Centre (BADC) (Met Office, 2012), available from http://catalogue.ceda.ac.uk/uuid/220a656152 18d5c9cc9e4785a3234bd0. In particular, temperature, rainfall, and so on, are well represented within MIDAS in terms of spatial and temporal coverage. However, only a small fraction of the thunderday observations have been digitized, and exclusively after 1950 (e.g. Owens, 2016).

In this study, we describe digitization of thunderday data from the scanned MWRs. After 1883, the MWR contain tabulated data as monthly totals of thunderdays at each station. These data are not easily machine readable, thus all digitization is done manually. Due to the labour intensive nature of digitizing and checking the data, we here focus only on 10 long-running stations which provide sufficient data to calculate averages for the full MWR time frame 1884–1993. Table 1 summarizes the 10 stations. Thunderday data for approximately 99% of the 110-year interval were reported at Wick, Nairn, Aberdeen, Spurn Head, Holyhead, Yarmouth, Oxford and Scilly St Mary's (though station replacement is required

**TABLE 1** Summary of the 10 British stations at which thunderdays have been digitized from the Monthly Weather Reports

| Original station | | | Replacement station | | | | Total coverage |
|---|---|---|---|---|---|---|---|
| Station name | MWR latitude (MIDAS) | MWR longitude (MIDAS) | Start date | Station details | Latitude | Longitude | |
| Wick MIDAS ID: 32 | 58.27 (58.45) | −3.05 (−3.09) | – | – | – | – | 99.17% (1,309) |
| Nairn MIDAS ID: 126 | 57.35 (57.59) | −3.52 (−3.89) | – | – | – | – | 99.62% (1,315) |
| Aberdeen MIDAS ID: 14,930 | 57.08 (57.16) | −2.08 (−2.10) | 1948 | Dyce MIDAS ID: 161 | 57.21 | −2.20 | 99.77% (1,317) |
| York MIDAS ID: 350 | 53.57 (53.96) | −1.05 (−1.09) | 1964 | Heslington MIDAS ID: 349 | 53.95 | −1.04 | 97.20% (1,283) |
| Spurn Head MIDAS ID: 16,595 | 53.35 (53.58) | 0.07 (0.12) | 1965 | Hull MIDAS ID: 369 | 53.76 | −0.36 | 98.56% (1,301) |
| Holyhead MIDAS ID: 1,145 | 53.19 (53.32) | −4.37 (−4.62) | – | – | – | – | 99.70% (1,316) |
| Yarmouth MIDAS ID: 4,926 | 52.37 (52.79) | 1.43 (1.43) | 1950 | Cromer MIDAS ID: 426 | 52.93 | 1.29 | 99.09% (1,308) |
| Cambridge Bot Gdns MIDAS ID: 454 | 52.12 (52.19) | 0.08 (0.13) | – | – | – | – | 96.89% (1,279) |
| Oxford MIDAS ID: 606 | 51.46 (51.76) | −1.16 (−1.26) | – | – | – | – | 98.79% (1,304) |
| Scilly St Mary's MIDAS ID: N/A | 49.56 | −6.18 | – | – | – | – | 98.86% (1,305) |

in some instances and there are issues of actual observational coverage, both are discussed further below). Data for approximately 97% of the 110-year sequence were reported at York and Cambridge Botanic Garden. In both cases, the largest gaps are at the end of the first decade, from 1898 through 1900. The available MWR metadata indicates that the governing authorities of these stations have changed numerous times during the period of consideration, particularly in the early 20th century.

Observations from 4 of the 10 stations ceased at various times between 1949 and 1965, shown as dashed vertical lines in Figure 2 and detailed in Table 1. For each of these four stations, we identified the nearest station in the MWR that was reporting thunderdays at that time, and continued the series using this substitute. In the remainder of the study, we refer to these composite stations by the original MWR station name only. Aberdeen ceases to appear in the MWR in 1948, but the Dyce station is station is close and so is a natural replacement (Geddes, 1955). Similarly, the replacement of York with Heslington in 1964 would not be expected to result in large systematic changes in thunder occurrence. For Spurn Head and Yarmouth, the replacement stations (Hull and Cromer, respectively) are more than 30 km away and thus should not necessarily be expected to experience the same thunderstorm conditions. The change from Spurn Head to Hull also constitutes a change from a coastal site to an inland site. Furthermore, the Spurn Head/Hull, Aberdeen/Dyce and the Yarmouth/Cromer switch overs could also produce changes in observing practice, with Spurn Head, Dyce and Cromer being either coastguard or airfields, and thus more likely to provide well-staffed observational coverage. Despite these concerns, Figure 2 shows that none of these substitutions result in obvious changes in the reported annual mean thunderdays. Indeed, the total number of thunderdays in the 5-year intervals before and after each station change agrees to within 10\% in all cases, which is within the 5-year variability of thunderdays reported at the original stations prior to 1950. While this certainly does not rule out the possibility that a systematic bias has been introduced, it must be small in comparison to the stepwise changes seen later in the series and discussed in more detail below.

We here note that there is a known issue with thunderday reporting at one of these stations. At Oxford, routine thunderday recording ceased in 1985/1986, though it was still sporadically reported after that time (Burt and Burt, 2019). Thus, from the MWR alone, this station appears to continue as normal through this transition. Of course, similar changes of which we are unaware may also have also occurred at other stations. Thus, it is useful to consider Oxford, with its known thunderday reporting issues, as one of the MWR in order to determine the degree to which unknown data problems can be detected.

The red dots in Figure 3 show the locations of the 10 digitized MWR stations. Black dots show the MIDAS station network, with each dot representing a station which contributes at least one positive thunderday measurement within the 1950–2010 intervals. The circle surrounding each MWR station represents a range of 30 km from the station, where additional thunderday data taken from the MIDAS dataset may be

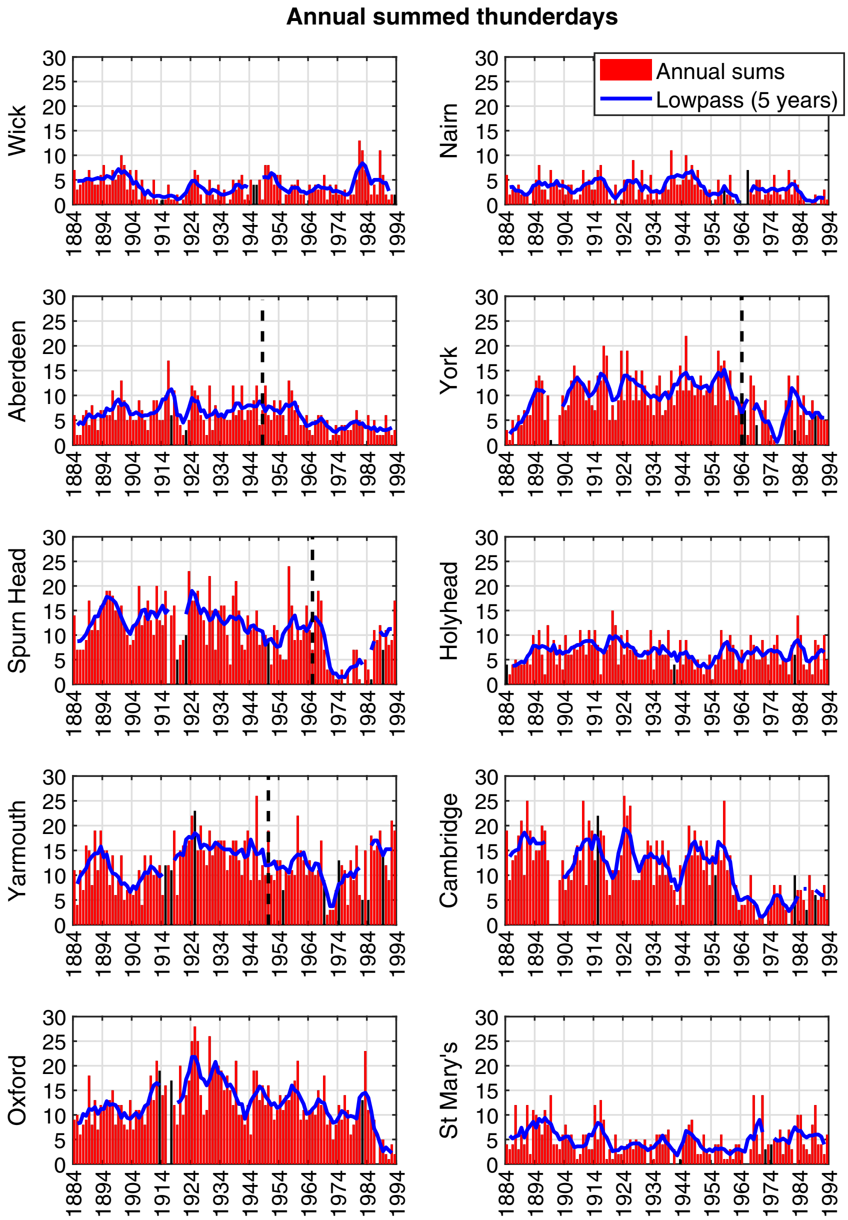


**FIGURE 2** Time series of annual summed thunderdays from the 10 individual MWR stations (listed in Table 1) over the entire MWR period 1884–1993. The red bars show years of complete data coverage, whereas black bars show years with one or more months missing. Thick blue lines show 5-year running averages, with the requirement for at least 4 complete years of data. The black dashed vertical lines indicate times when substitute stations were introduced

available for validation (Section 4). These additional MIDAS neighbouring stations are not used in our monthly or annual thunder series, only for assessing potential issues with the MWR-based thunderday estimates.

## 3 | OVERVIEW OF THE MWR THUNDER DATA

Figure 4 shows mean thunderday at the 10 British stations over the periods 1884–1993 and 1961–1990, along with yearly standard deviations about the means, representing the interannual thunderday variability for these two periods. The stations have been ordered along the *x*-axis according to their latitude, from north to south. Over the whole 1884–1993 period, the latitudinal trend is in agreement with that expected from the MIDAS data (Figure 3). Over the modern period (1961–1990), the latitudinal trend is significantly weaker, owing to suspected data issues discussed below. Over the 1884–1993, the MWR annual mean thunderdays range from 11–12 in East England to 3–4 days a year in Scotland. Western England/Wales sites, closest to the Atlantic, have considerably fewer thunderdays (<7 days a year on average at the Holyhead site) than that

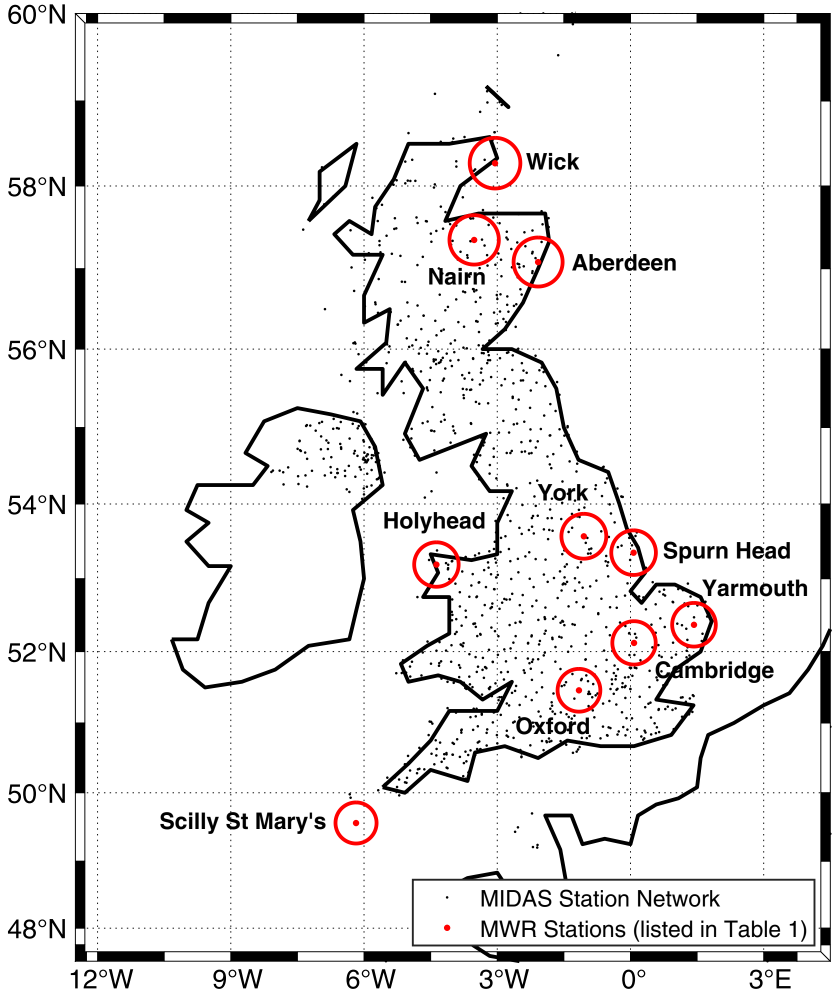


**FIGURE 3** Locations of the MWR stations (Table 1), shown in red, superimposed on the MIDAS station network, shown in black. Red circles indicate a 30-km radius about the MWR station position

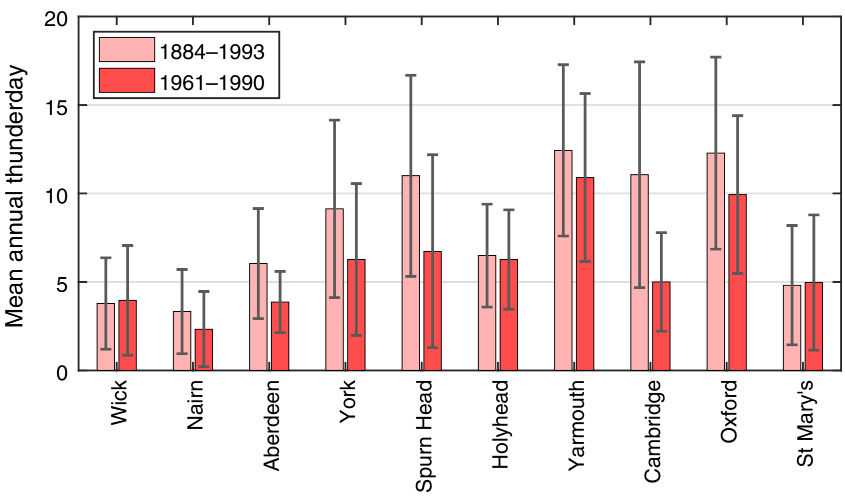


**FIGURE 4** Mean annual thunderdays from the 10 British stations digitized from the MWR, averaged over the periods 1884–1993 and 1961–1990. Error bars represent interannual standard deviations. Stations are ordered from highest to lowest latitude

of East England stations, and also have small interannual variability. These values are slightly lower than the annual mean values in Figure 3a. This could be the result of the MIDAS dataset including stations with only partial thunderday coverage through the year: For the almost complete MWR data series (in terms of no missing months within the MWR), annual mean values are essentially identical to annual sums. This is not always the case for the existing MIDAS data, where the annual mean can be significantly greater than the annual sum, suggesting a bias towards reporting in the summer months.

The reduced mean annual thunderdays in the period 1961–1990 seems to imply a long-term downward trend in thunderstorm occurrence. However, it is more likely to be a

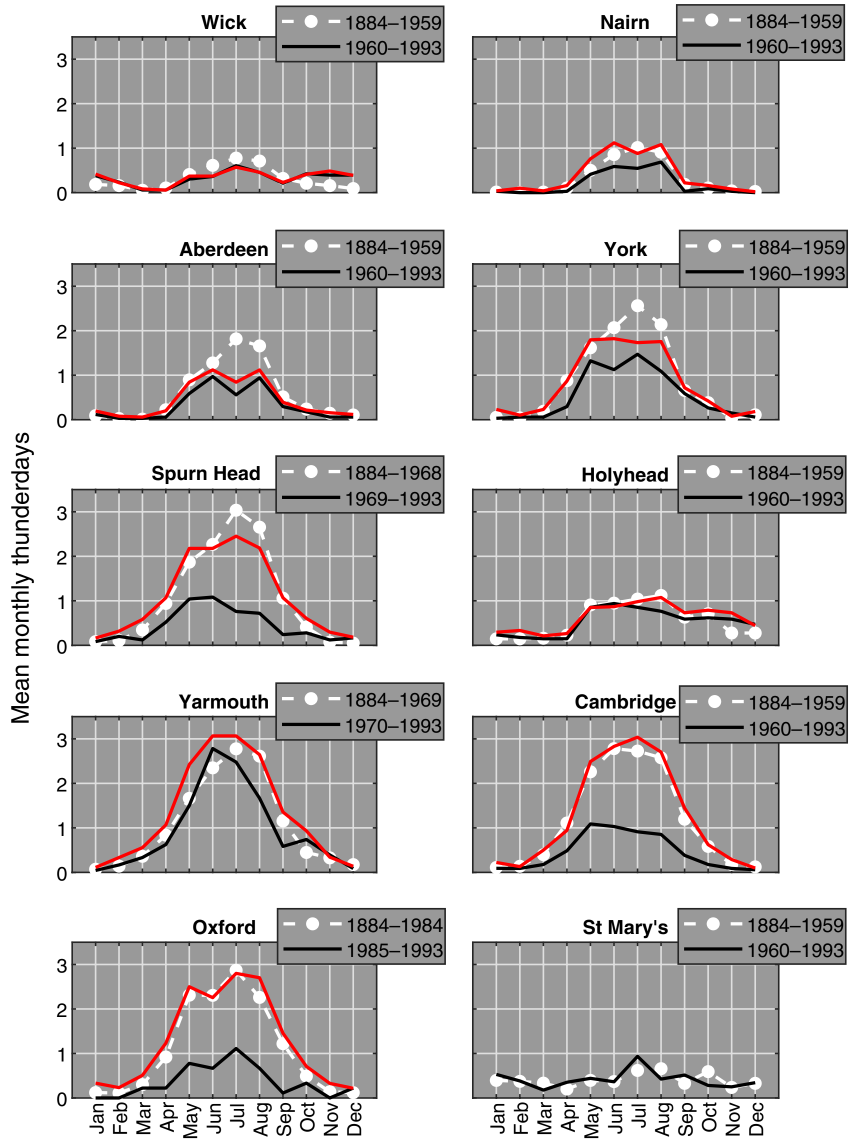


**FIGURE 5** Monthly mean thunderdays at the 10 MWR stations for two periods of averaging. The period from the start of the data sequence in 1884 up to the first stepwise drop seen in the time series (or 1960, if no stepwise drop has been identified) is shown in white. The period from the stepwise drop to the end of the sequence in 1993 is shown in black. The red lines show the mean of the regional maximum series, constructed from monthly maximum thunderdays from MIDAS stations within a 30-km radius of the MWR station

reflection of changing observing practice in the late 20th century. This can be more clearly seen by inspection of the time series, shown in Figure 2.

Annual summed thunderdays for the 10 individual stations are presented in red (for years with complete data coverage) and black (for years with one or more months of missing data). Thick blue curves show 5-year running means, with the requirement that at least 4 of the 5 years are complete. The known issue with Oxford thunderdays in 1985/1986 can be seen in the large decline in the 5-year average thunderdays. However, there is no obvious change in terms of the number of reported months (e.g. 1987–1993 are complete years in terms of data coverage), making it difficult to algorithmically identify similar problems in the other MWR series. Visual inspection of the time series strongly suggests a change in thunderday observing/recording practice at Cambridge around 1960–1962. As with Oxford, this is not associated with fewer reported months of data. At Spurn Head, there is a steep decline in reported thunderdays in 1969, though unlike Oxford and Cambridge, the number of thunderdays appears to recover by 1989. Shorter drops in reported thunderdays are present at Yarmouth and York around 1971 and 1976 respectively. From these MWR data alone, it is not possible to conclude whether these features are real or issues with reporting and data collection. Similarly, we cannot conclusively determine the origin of an approximately 25% decrease in annual thunderdays reported at Aberdeen around 1960. Comparison with the available thunderdays at nearby MIDAS stations is presented in the next Section.

Holyhead, Wick, Nairn, Holyhead and St Mary's do not show any obvious data issues, though the generally lower occurrence of thunder at these sites may mean any issues are more difficult to identify.

Figure 5 shows the annual cycles in thunderdays, split into two periods: Prior to and after potential issues identified in the time series, above. For stations where there is no obvious step-change in the time series (i.e. Wick, Nairn, Holyhead, St Mary's), annual cycles are considered before and after 1960, the earliest potential issue at the other stations. As expected, the late spring, summer and early autumn thunderdays are greatly reduced at Aberdeen, York, Spurn Head, Cambridge and Oxford after the identified data issues. Yarmouth, however, shows a similar annual cycle before and after the identified problem, suggesting it was a short, transient issue.

In the period prior to any potential data issues, the annual cycles predominantly follow expected seasonal changes in convective activity, peaking in the summer months for the central and eastern England. The Scottish sites, also situated toward in central and eastern regions, show a similar pattern though at greatly reduced amplitude. Holyhead and St Mary's, the only western sites, show a qualitatively different pattern. Less convective activity in the summer months means the summer peak in thunderstorm activity is greatly reduced. In the winter months, however, convective/showery activity in polar maritime air, particularly near active cold fronts or troughs means thunderstorm activity continues at higher levels than central/eastern regions where such activity is often quickly suppressed over land as the source of heating (relatively warm seas) is replaced by cooler land (Barrett, 1976).

## 4 | COMPARISON WITH MIDAS DATA

Validation of the MWR thunderday station data is possible from the mid-1950s onwards, when additional thunderday data are available from multiple neighbouring meteorological stations as part of the existing MIDAS dataset. For 9 of the 10 MWR stations, we have been able to identify the associated station ID in the MIDAS dataset, the exception being Scilly St Mary's. The coordinates of stations are not reported in each volume of the MWR, but collated approximately every 10 years in a summary report. As shown in Table 1, the latitudes and longitudes of the MWR stations are often slightly different to the equivalent stations in the MIDAS datasets. While this may indicate slight movements of the exact station locations over the years, there are known issues with the quality of the MWR metadata, which is the more likely source of the discrepancies.

As expected, there is a high level of agreement over the common period 1957–1993 between our MWR-based thunderday data (shown as white lines in Figure 6) and the DWR-based MIDAS data (red dots in Figure 6), at both the interannual (Figure 6) and seasonal (Figure 5) timescales.

In this section, we further compare thunderdays at each of the 10 MWR stations with all MIDAS stations (a) within

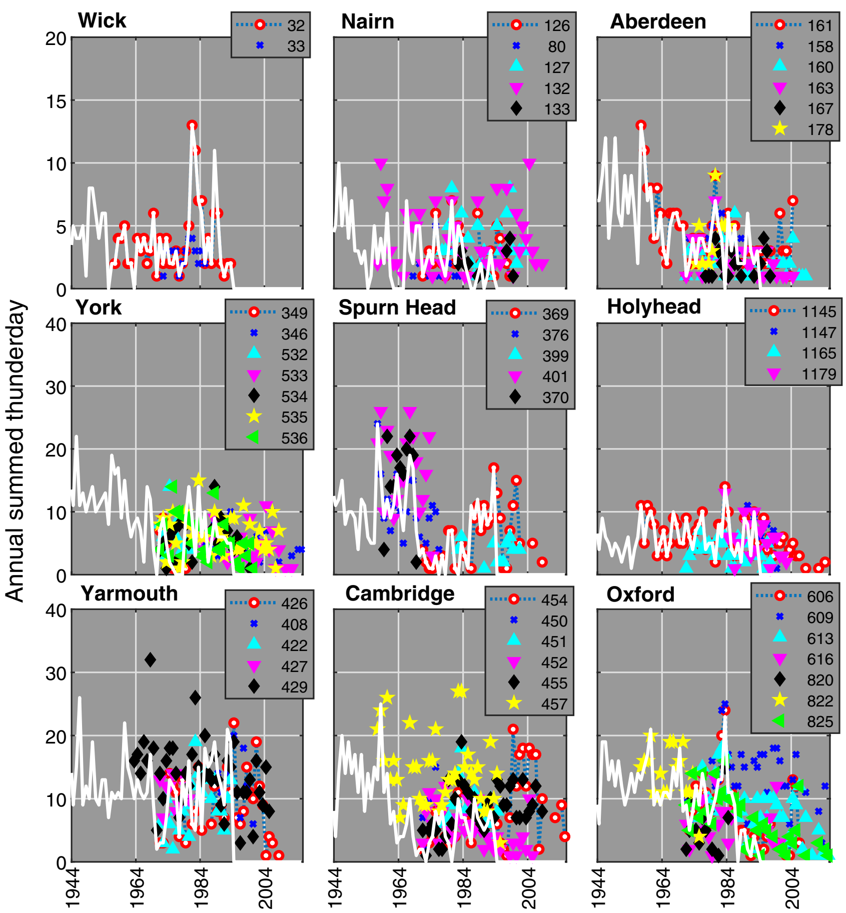


**FIGURE 6** Time series of annual summed thunderdays from individual MWR stations (white) and available MIDAS neighbours within 30 km (coloured symbols). The red symbols represent the MIDAS data for the same MWR station (listed in Table 1). MIDAS station IDs are given in the figure legends. Scilly St Mary's is omitted as no MIDAS stations are available for comparison. Only the period of overlap (i.e. after 1950) is shown

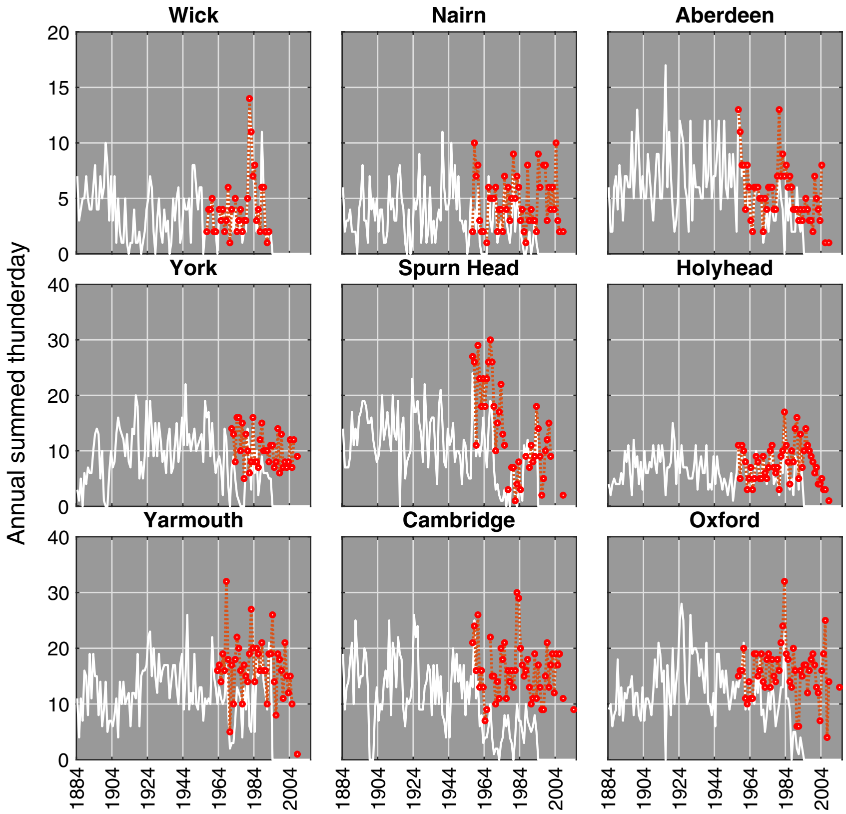


**FIGURE 7** Time series of annual summed thunderdays from individual MWR stations (white) and the MIDAS regional maximum, constructed from the highest monthly thunderday count from any available MIDAS neighbours within 30 km (red)

30 km range (denoted by red circles in Figure 3) and (b) with at least 1 complete year of data between 1950 and 1993. For the purpose of computing MIDAS neighbour stations, we use the MIDAS coordinates of the MWR stations. This comparison may help highlight station-specific issues. In particular, it is important to determine whether the sudden drops in thunderdays at various times in the late 20th century at Spurn Head, Cambridge and Oxford and, to a lesser extent, York and Aberdeen, are the result of genuine changes in thunderstorm occurrence or of (unknown) changes in measurement or recording practice at those stations. (The known Oxford 1985 issue serves as a general test of this methodology.)

Figure 6 shows time series of annual-mean thunderdays from MWR station data (except Scilly St Mary's), and the corresponding estimates derived from the MIDAS neighbour stations. There are two key points of note. First, the low thunderday counts at Cambridge after 1961, Oxford after 1985 and Spurn Head after 1969 are not unique to these stations; Other independent stations within 30 km show similar low thunderday counts. However, we cannot rule out the possibility that these other stations are also under-reporting. Second, in all three cases, there is at least one neighbouring station reporting thunderdays at a level comparable to the MWR station prior to the identified issue.

Given false positives in thunderdays are rare (e.g. Owens, 2016), a reasonable approximation is that the highest reporting station in a given region produces the most compete record. This is shown in Figure 7, where the red line shows the annual regional maximum thunderday time series, constructed from the maximum monthly thunderday count from any neighbouring MIDAS station. Figure 5 shows the mean monthly counts in the same format. The time series of the difference between the MWR and maximum MIDAS annual summed thunderdays is shown in Figure 8.

Even a perfectly reporting individual station would be expected, on average, to result in thunderdays slightly below the regional maximum level, as the latter can sample from a wider area. But the issues with Cambridge and Oxford are nevertheless immediately apparent. Prior to 1961 and 1985, Cambridge and Oxford, respectively, are within >50% of the regional maximum values. But they drop to below 50% after these dates. At Spurn Head, the under-reporting is apparent for 1960–1974, but after this period there are insufficient MIDAS neighbour stations to reliably identify issues with this kind of analysis. The one Spurn Head neighbour appears to be showing comparably low thunderday counts through the 1980s and 1990s. This highlights the need for a dense enough network of observations to ensure at least one diligent observation site throughout the period of interest. At York, the drop in thunderday counts around 1976 is not seen in the regional maximum and the recovery in thunderdays in the late 1980s remains low. At Aberdeen, there is good agreement between the MWR record and the regional maximum, but as with Spurn Head, the number of MIDAS stations is limited. Thus it is unclear whether the under-reporting issues mask a genuine decline in thunderdays through the late 20th century or not.

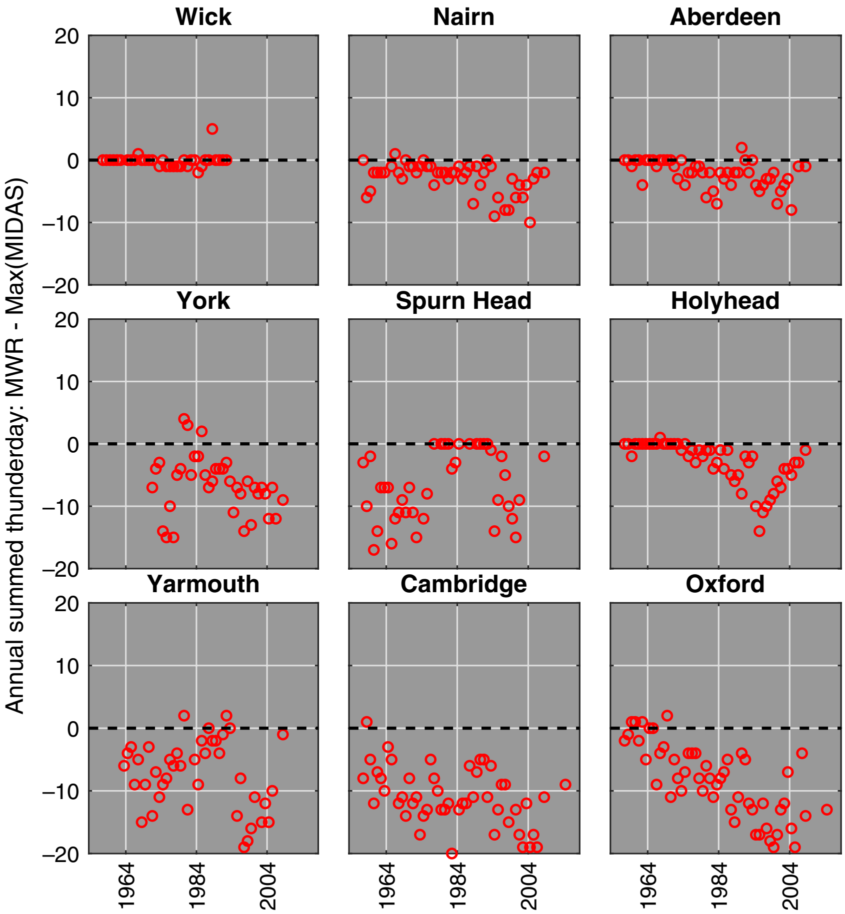


**FIGURE 8** Time series of the difference between annual summed thunderdays from individual MWR stations and the MIDAS regional maximum, constructed from the highest monthly thunderday count from any available MIDAS neighbours within 30 km. Only the period of overlap (i.e. after 1950) is shown

## 5 | DISCUSSION

In this study, we have presented a brief summary of newly digitized thunderday data from 10 stations from the UK Met Office Monthly Weather Reports (MWRs) over the period 1884–1994. For 4 of the 10 stations, it was necessary to switch to a substitute station at various points between 1949 and 1965. This does not appear to have introduced any significant biases, but as discussed below, digitization of a larger number of stations would enable better quantification of this.

The digitized data are available from <on acceptance, data to be deposited at BADC and link inserted here>. While the data are presented ‘as is’, there are a number of features that are likely the result of changes in observing and reporting practice, rather than genuine changes in local thunderstorm occurrence. In particular, at various times between the 1960 and 1990, there are clear stepwise declines in thunderdays reported at Cambridge, Spurn Head and Oxford, and to a lesser extent at Aberdeen and York. Yarmouth shows a decline limited to just a few years. At Oxford, the change in observational practice can be directly identified through additional records and station meta-data (Burt and Burt, 2019) which results in a lack of routine reporting of thunder. For other stations, the provenance is not known and the cause can be more subtle. The veracity of thunderday data does not necessarily require a change in official observing practice. A change in observer or even the observer's living personal arrangements (e.g. whether they are situated in the vicinity of the observing site, and thus able to make effective 24/7 observers or not) can have a significant effect. Such effects are extremely difficult to identify and attribute. Instead, we have made comparisons with thunderdays recorded at neighbouring stations available as part of the MIDAS dataset after 1950. Of course, the definition of ‘neighbouring’ is critical here. It must be large enough that a sufficient number of additional stations are available, but small enough that stations are expected to experience the same synoptic weather conditions. In this study, we have used a radius of 30 km, though as more UK thunderday data are digitized, this could be relaxed. With greater observation density, it may be possible to take a more sophisticated approach to identifying proxy stations, for example similar synoptic conditions, than a simple geographic distance criterion. We also note that as there are no longer any 24 × 7 manual observing sites in the UK Met Office observing network. Thus, high-quality observational records of thunder occurrence from private citizens may be able to address this problem, particularly for the modern period.

In the cases of Cambridge, Spurn Head, Oxford, York and Aberdeen, other neighbouring stations report comparably low values in the late 20th century. However, with the shift towards automated weather recording, it is perhaps not surprising that multiple stations would under-report thunder during this period. A better proxy for accurate thunderdays may be the highest reporting station in the neighbourhood, as false positives are rare (Owens, 2016). Using this measure, it is clear that the declines at Cambridge, Oxford, Spurn Head and York are not genuine. At Aberdeen and Yarmouth, further neighbouring stations are required. Clearly, these issues merit further investigation and more in-depth validation against either other (as yet undigitized) stations present in the MWR or other independent thunderday records (e.g. the TORRO network The TORRO group, 2018). The digitization of a large number of stations from the MWR does not seem amenable to machine-based approaches and would require significant person hours to achieve manually, thus citizen science projects may provide the best route.

An alternative to the labour-intensive 'digitize everything' approach may be to more critically select stations for their likely reporting quality, rather than length of time series in the MWR. These stations could be used to establish baselines by which to compare other stations. Obvious choices are Kew Observatory, which extends back to 1843, but closed in 1980. There is a long overlap with Heathrow (1949–2000), which could be used to extend the record. Eskdalemuir and Lerwick have been Met Office Observatory stations since 1910 and 1921, respectively, and thus may be expected to be more likely to provide a consistent reporting level.

Detailed analysis of long-term trends in these 10 MWR thunderday series is inadvisable until further validation and calibration has been performed. However, by combining the MWR records and the highest-reporting MIDAS neighbours, broad conclusions can nevertheless be reached. Figure 5 compares the highest reporting record for the modern era (1960–present) with the MWR records before any identifiable data issues (typically 1884–1959) and after identified issues (typically 1960–1993). This clearly suggests the modern declines seen in most of the MWR series are not physical. But the agreement between the highest reporting MIDAS neighbours and the early MWR series also suggests that there has been no significant increase in UK thunderdays, as might be expected from thermodynamic arguments under climate change (Romps *et al.*, 2014). Thus competing effects, such as changing atmospheric dynamics and changing UK air mass origins (Jones *et al.*, 2013) or changes in the cloud ice flux (Finney *et al.*, 2018) must also be at play.

## ACKNOWLEDGEMENTS

We are grateful to the UK Met Office for making the scanned Monthly Weather Report data publicly available. They can be obtained from the Met Office Digital Library and Archive: https://digital.nmla.metoffice.gov.uk/collection_75a68cd2-cabe-43a8-98bb-3919f51e59a9/. We have also benefited from the availability of the MIDAS data via the British Atmo-spheric Data Centre (BADC): http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0.

## ORCID

*Mathew J. Owens* https://orcid.org/0000-0003-2061-2453
*Ed Hawkins* https://orcid.org/0000-0001-9477-3677
*Philip Craig* https://orcid.org/0000-0001-9213-4599

## REFERENCES

Aich, V., Holzworth, R., Goodman, S., Kuleshov, Y., Price, C. and Williams, E. (2018)Lightning: a new essential climate variable, Eos 99. Available at: https://eos.org/project-updates/lightning-a-new-essential-climate-variable.

Albrecht, R.I., Goodman, S.J., Buechler, D.E., Blakeslee, R.J., Christian, H.J., Albrecht, RI *et al*. (2016) Where are the lightning hotspots on earth? *Bulletin of the American Meteorological Society*, 97(11), 2051–2068. Available at: http://journals.ametsoc.org/doi/10.1175/BAMS-D-14-00193.1.

Barrett, E. (1976), Cloud and thunder. In: Chandler, T.J. and Gregory, S. (Eds.) *The Climate of the British Isles*. London, New York: Longman, pp 390. Available at: https://www.worldcat.org/title/climate-of-the-british-isles/oclc/603596109.

Brooks, C.E.P. (1925) *The Distribution of Thunderstorms Over the Globe*, number 24. London: HM Stationery Office.

Burt, S. (2012) *The Weather Observer's Handbook*. Cambridge: Cambridge University Press.

Burt, S. and Burt, T. (2019) *Oxford Weather and Climate Since 1767*. Oxford: Oxford University Press.

Changnon, S.A. (1985) Secular variations in thunderday frequencies in the twentieth century. *Journal of Geophysical Research*, 90(D4), 6181–6194.

Enno, S.-E., Post, P., Briede, A. and Stankunaite, I. (2014) Long-term changes in the frequency of thunder days in the Baltic countries. *Boreal Environment Research*, 19, 452–466. Available at: https://helda.helsinki.fi/bitstream/handle/10138/228651/ber19-5-452.pdf?sequence=1.

Finney, D.L., Doherty, R.M., Wild, O., Stevenson, D.S., MacKenzie, I.A. and Blyth, A.M. (2018) A projected decrease in lightning under climate change. *Nature Climate Change*, 8(3), 210–213. Available at: http://www.nature.com/articles/s41558-018-0072-6.

Geddes, A.E.M. (1955) The development of the study and practice of meteorology at Aberdeen. *Weather*, 10(11), 385–389. Available at: http://doi.wiley.com/10.1002/j.1477-8696.1955.tb00116.x.

Jones, P.D., Harpham, C. and Briffa, K.R. (2013) Lamb weather types derived from reanalysis products. *International Journal of Climatology*, 33(5), 1129–1139. Available at: http://doi.wiley.com/10.1002/joc.3498.

Kitagawa, N. (1989) Long-term variations in thunderday frequencies in Japan. *Journal of Geophysical Research*, 94(D11), 13183–13189.

Lee, A.C.L. (1989) Ground truth confirmation and theoretical limits of an experimental VLF arrival time difference lightning flash locating system. *Quarterly Journal of the Royal Meteorological Society*, 115, 1147–1166. Available at: http://10.0.3.234/qj.49711548908.

Lewis, R.P.W. (1991) *Meteorological Glossary*, 6th edition. London: Her Majesty's Stationary Office.

Met Office (2012) Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current). Available at: http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0.

Nag, A., Murphy, M.J., Schulz, W. and Cummins, K.L. (2015) Lightning locating systems: insights on characteristics and validation techniques. *Earth and Space Science*, 2(4), 65–93. Available at: http://doi.wiley.com/10.1002/2014EA000051.

Owens, M.J. (2016) Remember, remember the 5th of November: was that thunder I heard or not? *Weather*, 76(6), 134–137.

Owens, M.J., Scott, C.J., Lockwood, M., Barnard, L., Harrison, R.G., Nicoll, K *et al*. (2014) Modulation of UK lightning by heliospheric magnetic field polarity. *Environmental Research Letters*, 9(11), 115009.

Price, C.G. (2013) Lightning applications in weather and climate research. *Surveys in Geophysics*, 34(6), 755–767. Available at: https://doi.org/10.1007/s10712-012-9218-7.

Price, C., Penner, J. and Prather, M. (1997) NOX from lightning: 2. Constraints from the global atmospheric electric circuit. *Journal of Geophysical Research: Atmospheres*, 102(D5), 5943–5951. Available at: http://link.springer.com/10.1007/s10712-012-9218-7.

Prichard, B. 2016. *Thunderstorm Observing in the United Kingdom, In 'Extreme Weather'*, Chichester, UK: John Wiley & Sons Ltd, pp. 135–154. Available at: http://doi.wiley.com/10.1002/9781118949986.ch8.

Rodger, C.J., Werner, S., Brundell, J.B., Lay, E.H., Thomson, N.R., Holzworth, RH *et al*. (2006) Detection efficiency of the VLF World-Wide Lightning Location Network (WWLLN): initial case study. *Annales Geophysicae*, 24(12), 3197–3214. Available at: https://www.ann-geophys.net/24/3197/2006/.

Romps, D.M., Seeley, J.T., Vollaro, D. and Molinari, J. (2014) Projected increase in lightning strikes in the United States due to global warming. *Science*, 346(6211), 851–854.

Stringfellow, M.F. (1974) Lightning incidence in Britain and the solar cycle. *Nature*, 249, 332–333. Available at: http://10.0.4.14/249332a0.

The TORRO group. (2018) TORRO - Thunderstorm Observations. Available at: http://www.torro.org.uk/.

**How to cite this article:** Valdivieso M, Owens MJ, Scott CJ, Hawkins E, Burt S, Craig P. Thunderstorm occurrence at ten sites across Great Britain over 1884–1993. *Geosci. Data J*. 2019;00:1–12. https://doi.org/10.1002/gdj3.75