



## RESEARCH ARTICLE

# How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom

Jonathan Coney<sup>1</sup>  | Ben Pickering<sup>1,2,3</sup>  | David Dufton<sup>2</sup> | Maryna Lukach<sup>2</sup> | Barbara Brooks<sup>2</sup> | Ryan R. Neely III<sup>1,2</sup>

<sup>1</sup>School of Earth and Environment,  
University of Leeds, Leeds, UK

<sup>2</sup>National Centre for Atmospheric Science,  
University of Leeds, Leeds, UK

<sup>3</sup>Menapia Ltd., Leeds, UK

## Correspondence

Jonathan Coney, School of Earth and  
Environment, University of Leeds,  
Leeds, UK.

Email: [mm16jdc@leeds.ac.uk](mailto:mm16jdc@leeds.ac.uk)

## Funding information

Natural Environment Research Council,  
Grant/Award Numbers: NE/N008359/1,  
NE/S007458/1; UKRI

## Abstract

Observations of the real-time state of the atmosphere are required in order to initialize numerical weather prediction (NWP) models. As NWP resolution improves, more observations are needed, to better capture regional variations in atmospheric conditions. In particular, surface observations are necessary to reflect conditions experienced on the surface. One proposed opportunity to increase the number of surface observations available for assimilation into NWP is to crowdsource the data from home weather stations. This study investigates the outdoor air temperature measurements made by Netatmo home weather stations, through validation against a calibrated laboratory chamber and by evaluating quality control schemes that are applied to a UK-wide network of Netatmo stations. In a series of controlled lab experiments, it was found that the Netatmo temperature sensor was accurate to 0.3°C. The response to fluctuations in temperature is lagged, with  $\tau$  (the time taken for 63% of the change to be measured) calculated as 12.7 min for a near-instantaneous decrease in temperature. Netatmo temperature observations were compared with Met Office MIDAS hourly weather observations. A warm bias in excess of 1°C was present in the Netatmo temperature observations, which was lessened by the three quality control schemes tested, but still in excess of 0.5°C. Hence, Netatmo temperature measurements have potential to be assimilated in NWP in the United Kingdom, but work is required to find a suitable agreed quality control scheme to filter out anomalous observations in the United Kingdom.

## KEYWORDS

crowdsourcing, evaluation, Netatmo, observations, temperature, verification

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. *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

## 1 | INTRODUCTION

Surface observations are an integral part of initializing weather models and verification of forecasts. In the United Kingdom, the Met Office operates and maintains around 309 surface stations for record-keeping and for numerical weather prediction (NWP) (Met Office, 2012). While the observations made by Met Office surface stations are subject to calibration checks and are quality controlled (Chapman et al., 2017), the mean distance between Met Office observation stations is 17.6 km. This is a significantly longer distance than the grid spacing employed by the Met Office's UKV forecasting model of 1.5 km (Met Office, 2019). Good initial conditions are crucial to ensuring an accurate forecast is produced, so spatially dense surface observations become more necessary to initialize NWP models as resolution of forecast models increases in the future (Inness & Dorling, 2013). The increase in skill of NWP, along with improved resolution due to increased computing power and additional observation sources (such as satellites), now mean accurate forecasts can be provided on a very local scale (Bauer et al., 2015; Benjamin et al., 2019; ECMWF, 2020).

Crowdsourcing has grown in popularity over the last decade as a method to utilize small efforts from many people or devices and combine them into a useful output (Chapman et al., 2017; Elmore et al., 2014; Hawkins et al., 2019; Lackstrom et al., 2017). Thousands of crowdsourced observations are made every day from a wide range of automatic weather stations globally, and these observations are not yet widely utilized by meteorological agencies to forecast the weather. The Met Office's Weather Observations Website (WOW) (<https://wow.metoffice.gov.uk/>) and Weather Underground (<https://www.wunderground.com/>) are two examples of websites allowing users and organizations to submit data from their own weather stations to a larger database. However, observations on WOW and Weather Underground are made by a large variety of weather stations from a range of manufacturers, which have different biases and error characteristics (S. Bell et al., 2015).

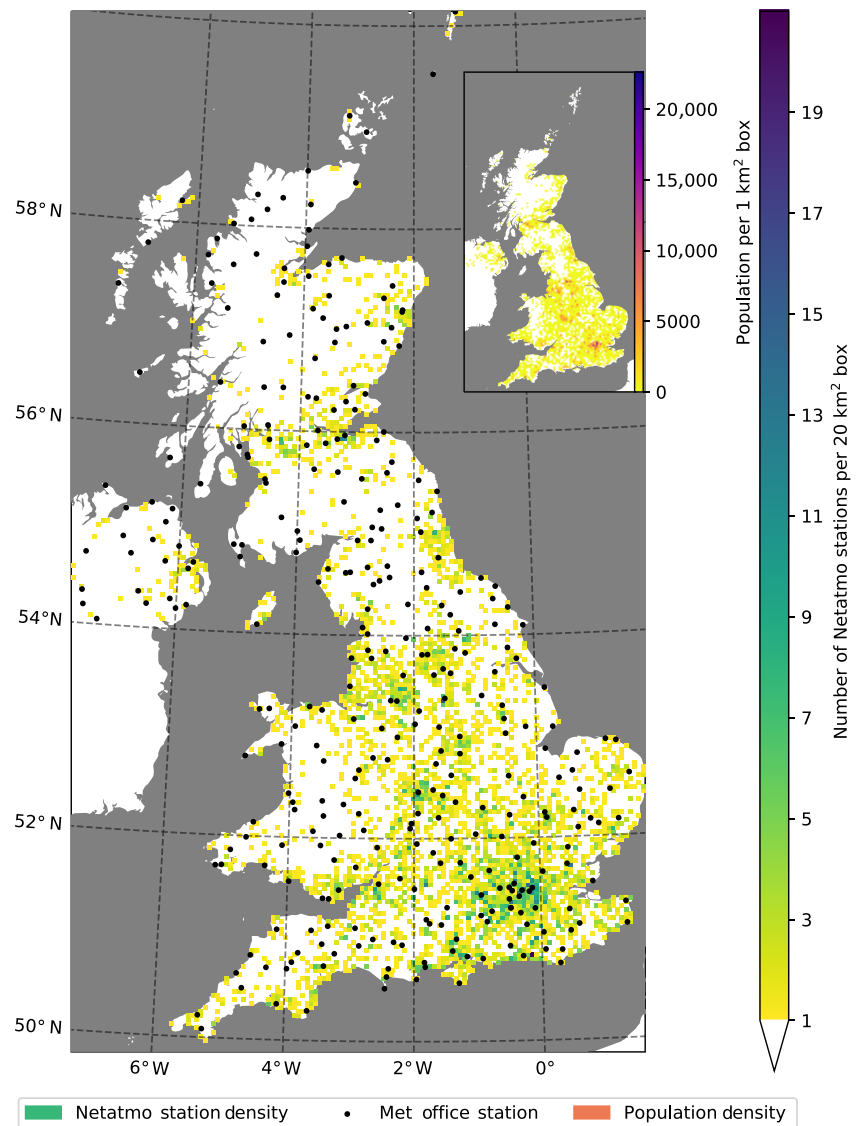
The Netatmo Smart Home weather station network (<https://www.netatmo.com/en-gb/weather/>) overcomes this issue as all data are created from the same type of sensors. This reduces the biases between different sensor types, which can be accounted for when using the observations. There are over 5000 Netatmo stations making observations in the United Kingdom, more than 15 times the number of Met Office surface stations in the Meteorological Monitoring System (MMS) (Green, 2010). However, the data quality of Netatmo stations is expected to be less than that of the Met Office stations, which this study will attempt to quantify.

While the Met Office stations are approximately equidistant, the Netatmo stations are correlated to the population density of the United Kingdom. Figure 1 shows this distribution: there are far more Netatmo stations in London and the southeast of England than the Highlands and Islands of Scotland. Meanwhile, the Met Office surface observation stations are fewer in number than Netatmo stations, but offer more even coverage across the country.

Observations from Met Office stations are trusted because they are regularly calibrated and are well-maintained, with quality control procedures in place in accordance with World Meteorological Organisation (WMO) guidelines (Inness & Dorling, 2013; WMO, 2018). For example, the platinum resistance thermometers at Met Office surface stations are calibrated every 8 years (Met Office, 2020b). Quality controlling data is the 'greatest challenge' faced by network managers of automatic weather stations (Fiebrich, 2009). In contrast, there is no centralized way for owners of Netatmo stations to supply metadata about how often their stations are calibrated and maintained, if at all. Hence, there is no way to tell whether an anomaly in Netatmo observations is due to some microscale effect or due to station placement or instrument error (Chapman et al., 2017). Some owners may maintain their stations, others may not, such as checking that the Netatmo temperature sensor is not in direct sunlight, which contributes to a warm bias in the Netatmo data (Nipen et al., 2020). Identifying and quantifying the uncertainties within measurements made by Netatmo stations would make assimilation of Netatmo data into an NWP model far more valuable. Table 1 summarizes the differences between observations from Met Office surface stations and Netatmo home weather stations.

Supplementing observations from traditional networks with crowdsourced data has proved useful for studies where traditional observations are sparsely located, such as the following examples. Evidence of a nighttime urban heat island (UHI) effect of up to 5.5°C was shown in London using 287 Netatmo and 7 Met Office stations, utilizing the Met Office data to remove anomalous Netatmo observations (Chapman et al., 2017). Netatmo stations were also used by Mandement and Caumont (2019) to analyse deep convection in France, showing that rapid changes in atmospheric pressure were detected around squall lines. Data from 63 citizen rain gauges around Amsterdam were correlated with rain radar observations (De Vos et al., 2017). However, limitations specific to Netatmo stations are outlined, including the lack of a radiation shield on the Netatmo temperature module, contributing to a lagged response time to temperature changes (Büchau, 2018; Chapman et al., 2017).

**FIGURE 1** Map showing the location and density of Netatmo stations (yellow/green to purple) and Met Office surface stations (black) in the United Kingdom. Each box represents approximately 20 km<sup>2</sup> of area. There are far more Netatmo stations in highly populated areas than there are in rural areas, as demonstrated by the population density map (inset). Population data supplied by the Natural Environment Research Council (Reis et al., 2017).



Netatmo data are used operationally by MET Norway to correct near-future temperature forecasts due to features such as cold pools and UHIs (Nipen et al., 2020). These examples show that crowdsourced data can be used for studies of phenomena missed by conventional observations, and aid forecasting.

This study characterizes the uncertainties in temperature observations made by Netatmo weather stations. This is completed in two stages. A two-staged approach separates the influence of the sensors themselves, and the placement of sensors. Firstly, a set of calibration experiments in a climate chamber quantifies the sensor uncertainty characteristics without external influencing factors. The second part of this study quantifies the bias in temperature observations in the UK Netatmo network, and involves assessing potential quality control schemes for crowdsourced Netatmo temperature observations in the United Kingdom. Quality control schemes smooth the data, which remove some real events. However, it is

advantageous to NWP assimilation to receive a smoothed field (Dunne & Entekhabi, 2005). Several different schemes have already been proposed in the literature (Clark et al., 2018; Meier et al., 2017; Nipen et al., 2020), three of which are tested here to determine which scheme best reduces the warm biases for the UK network. There is scope for meso- $\gamma$  and denser observation networks to add value to forecasting and environmental monitoring (Chapman et al., 2015). For Netatmo sensors in particular, there is evidence to show that there are biases and a lagged response within the observations made by the temperature sensor (Büchau, 2018; Chapman et al., 2017; Meier et al., 2015).

## 2 | DATA

This study makes use of three data sources: Netatmo citizen weather station observations; laboratory calibration

**TABLE 1** Summary of the characteristics of both the Netatmo stations and Met Office surface stations in the United Kingdom as of April 2020 (Büchau, 2018; Green, 2010; Met Office, 2018b; WMO, 2018)

Description	Met Office (MIDAS)	Netatmo
Number of UK observations	259 (daily median); 309 (number of unique stations)	5067 (daily median); 6098 (number of unique stations)
Mean distance to nearest observation station	17.6 km	2.72 km
Time interval between observations	1 h	5 min
Temperature measurement	Two Platinum Resistance Thermometers in Stevenson Screen	Sensirion SHT20 sensor chip in aluminium casing
Ongoing calibration	In accordance with World Meteorological Organisation guidelines	None

Abbreviation: MIDAS, Met Office Integrated Data Archive System.

**TABLE 2** Summary of the Netatmo code regime to collect observations from nearly all UK stations

Time of day	Task description
12:00 AM	The file containing Netatmo station MAC addresses is updated with any new stations which came online for the first time in the last 24 h
12:30 AM	All Netatmo MAC addresses are split up into smaller groups, to prevent reaching the hourly limit of the Netatmo API
From 1:00 AM onwards	Every 30 min, a new group of MAC addresses is read by a script and is called from the Netatmo server. In each API call, observations from a single station in the last 24 h are retrieved from the server and are stored into netCDF files

and reference equipment at the National Centre for Atmospheric Science (NCAS) Atmospheric Measurement and Observation Facility (AMOF); and Met Office MMS data. The following subsections describe these data sources.

## 2.1 | Netatmo data

The data from Netatmo weather stations are collected using the Netatmo Weather Application Programming Interface (API). A series of Python scripts were written to collect data operationally on high-performance computing (Lawrence et al., 2013). There are limits imposed on the API by Netatmo (of no more than 50 requests every 10 s, and no more than 500 requests per hour [Netatmo, 2012]), so the scripts are designed to collect as much data as possible, with a lag of 1 day to allow slow to upload stations catch up (Table 2).

**TABLE 3** Table showing the difference between the Netatmo sensors and the equipment in the AMOF laboratory

Characteristic	Netatmo sensor	PTUCal measurement
Temperature measurement	Sensirion SHT20 sensor chip, in aluminium casing	Platinum resistance thermometer connected to the Dewpointmeter
Length of time between temperature measurements	5 min	2 s
Stated temperature accuracy	$\pm 0.3^{\circ}\text{C}$	$\pm 0.1^{\circ}\text{C}$
Response speed	To be calculated	$1^{\circ}\text{C/s}$ , plus settling time
Temperature range	$-40^{\circ}\text{C}$ to $+65^{\circ}\text{C}$	$-40$ to $+120^{\circ}\text{C}$

Note: Data sourced from STMicroelectronics Pty Ltd (2012); Büchau (2018); Michell Instruments (2018); Netatmo (2020).

## 2.2 | Pressure, temperature, humidity calibration facility calibration

The calibration in the pressure, temperature, humidity calibration (PTUCal) chamber at AMOF in NCAS utilizes two main pieces of equipment: a Votsch VT-4011 temperature test chamber containing a high accuracy thermometer and a Michell S8000 Remote Precision Dewpointmeter. The test chamber holds each Netatmo sensor being tested, the temperature of which is regulated by the dewpointmeter. Table 3 summarizes the differences between the laboratory instruments and the Netatmo sensors being tested.

## 2.3 | Met Office observations

The Met Office observation data are used to compare the effectiveness of the quality control schemes for removing anomalous results in the United Kingdom. The data are sourced from the Centre for Environmental Data Analysis (CEDA) archive, and are from the Met Office Integrated Data Archive System (MIDAS) (Met Office, 2012). The two data sets are: *MIDAS: UK Hourly Weather Observation Data* data set, which contains temperature observations from the surface observation stations in the United Kingdom; and the *MIDAS: Global Radiation Observations* data set, which is required by the Meier quality control scheme. Data from 309 surface stations are utilized in this study. Temperature observation at these stations is automated using platinum resistance thermometers (however, at some stations, a meteorologist does add extra information not recorded by the automatic station) (Met 96 Office, 2018a). These observations are stated to be accurate to within  $\pm 0.1^\circ\text{C}$  for mean temperatures, and  $\pm 0.3^\circ\text{C}$  for extremes (Clark et al., 2014).

## 3 | METHODS

This study characterizes the uncertainties and behaviours of the Netatmo station network in the United Kingdom for temperature measurements, using two approaches. Firstly, the bias contribution from the sensors themselves is identified by using ideal conditions within the PTUCal chamber, removing external factors like solar or anthropogenic heating. In addition, the bias from the placement of a sensor can be quantified once the sensor bias is known from the chamber tests. Three quality control schemes to remove or correct anomalous observations from Netatmo temperature sensors in the United Kingdom are compared.

### 3.1 | PTUCal climate chamber experiment

Seven Netatmo outdoor temperature sensors (one owned by the University of Leeds [UoL], six loaned from the University of Birmingham [UoB1–6], see Acknowledgements) are tested using the PTUCal chamber described in Section 2.2 at AMOF at NCAS in Leeds to determine biases within the sensors. The specification of the instruments is outlined in Table 3.

Each Netatmo outdoor temperature sensor (UoL, UoB1–6) is tested individually using the same experimental set-up, shown in Figure 2, using the equipment outlined in Section 2.2. In the study by Büchau (2018), some



FIGURE 2 A Netatmo temperature sensor in the climate chamber. The thermocouple is attached by a cable tie so that the same parcel of air is measured by both sensors during the test.

experiments were conducted by removing the outer case of the temperature sensor. Since there is a very low probability that any Netatmo owners in the network remove the aesthetic casing, it is retained for the experiment to realistically capture the characteristics of a networked sensor. The climate chamber is a highly controlled environment, which allows assessment of the Netatmo temperature measurements against a well-calibrated dewpointmeter, but is not a proxy for the real world. The thermocouple was placed directly next to the Netatmo outdoor module so that the same parcel of air is measured by both sensors. The chamber contains a fan for ventilation.

The chamber cycles through the same programme of temperatures for each different Netatmo sensor: the temperature in the chamber is set to  $40^\circ\text{C}$ , and then after 3 h the temperature is decreased by  $2.5^\circ\text{C}$ , until the test at  $-10^\circ\text{C}$  is completed. In each 3-h period, the first 2 h are a delay time to allow the chamber and the Netatmo sensor to reach a stable equilibrium, so no lag time effect exists. The range in temperatures tested reflects the range of air temperatures commonly observed in the United Kingdom (Parker et al., 1992). The  $2.5^\circ\text{C}$  step change is reasonable, as it is a realistic change in temperature observed in the United Kingdom. For example, Clark et al. (2018) give an example of a  $3.2^\circ\text{C}$  change in temperature, in 8 min.

### 3.2 | Comparison of Netatmo and Met Office observations

Having investigated the existence of any biases in the Netatmo temperature measurements, using the PTUCal

climate chamber, comparisons are made to Met Office observations from Netatmo stations located throughout the United Kingdom to test its performance in the real world. Met Office observations are accurate to  $\pm 0.1^\circ\text{C}$ , while Netatmo temperature measurements are stated to be accurate to  $\pm 0.3^\circ\text{C}$  (Clark et al., 2014; Netatmo, 2020). The Netatmo observations are gridded using the `scipy.interpolate.griddata` function in the SciPy library in Python (Virtanen et al., 2020), which uses Delaunay triangulation on the data before barycentric linear interpolation. Clark et al. (2018) use Delaunay triangulation for gridding Netatmo data. Linear barycentric interpolation avoids points of discontinuity, which is the case when using bilinear interpolation (the approach used by Hollis et al. (2019)). The interpolated Netatmo temperatures co-located over the Met Office station locations are then compared with the observations made directly by the Met Office surface stations. Interpolating Netatmo observations in this way may cause some error, but the error is assumed to be unbiased when the mean of the whole United Kingdom is taken.

### 3.3 | Quality control schemes

Three quality control schemes, from Meier et al. (2017), Nipen et al. (2020), and Clark et al. (2018), are implemented on the UK Netatmo data from April 2020, to quantify the bias in the Netatmo temperature sensor from its placement within the environment.

#### 3.3.1 | Meier scheme

The Meier scheme has four stages:

1. Remove Netatmo stations with identical latitude and longitude (as this suggests the location has defaulted to one based upon their IP address). A day (month) of observations are removed if a station records for less than 80% of the time within that day (month).
2. Daily averages of minimum air temperature from all Netatmo and reference stations (in our case, Met Office sites) are compared. If a Netatmo station has a monthly mean minimum temperature that is outside  $5\sigma$  of the average Netatmo temperature, or SD outside  $5\sigma$  of the reference data, then the Netatmo station is removed.
3. For each hour, the Pearson linear correlation between incoming solar radiation and temperature difference between a Netatmo station and the mean reference observed temperature are calculated. If  $p < 0.01$  and Pearson  $r > 0.5$  (Wilks, 2011) between the Global Solar

Irradiation Amount and  $t_{\text{Netatmo}} - t_{\text{Reference}}$ , then it is concluded that solar heating is directly affecting the sensor and it is therefore removed.

4. For each station, the mean and SD  $\sigma_{\text{obs}}$  of all available observations over the entire month are calculated. If any single observation deviates from the mean by more than  $3\sigma_{\text{obs}}$  (for example, if a sensor was temporarily moved to have its batteries changed), then this observation is removed.

#### 3.3.2 | Nipen scheme

The Nipen scheme utilizes the spatial distribution of stations and has three sections:

1. A ‘buddy check’ to remove observations within 3 km and 30 m of altitude for each observation. If an observation has a deviation from the mean neighbourhood observation of greater than  $2\sigma$ , then this observation is removed.
2. Stations without five neighbouring stations within 15 km and 200 m of altitude are removed to avoid an isolated site having an erroneous value that cannot be cross-referenced.
3. The majority of the computational expense is required for the spatial consistency test (outlined in Lussana et al., 2010): The United Kingdom is grouped into approximately 100 km by 100 km squares. For each region at each time, the mean temperature and the variance of the error compared with the mean are calculated. Any observation with a warm deviation that is more than four times the error variance will be removed, and any observation with a cold deviation that is more than eight times the error variance will be removed also. A less strict threshold is used for cold deviations because Nipen et al. (2020) explain that ‘most error sources’, for example, direct sunlight or proximity to walls, ‘contribute to a warm bias’.

#### 3.3.3 | Clark scheme

The final scheme used here is described in Clark et al. (2018), which adjusts or ‘nudges’ observations with anomalous measurements, based on the nearest Met Office site, which is in contrast to the other schemes outlined above that remove anomalous observations. The amount to adjust each Netatmo observation is calculated as the difference between the 6-h mean temperature from the nearest Met Office site ( $T - 3$  to  $T + 3$ ), to the same 6-h mean temperature from the Netatmo station, using the observations made at the same times used for the Met

Office observations. This correction varies from day to day. For example, for 12:00 UTC, the observations from 9:00, 10:00, 11:00, 12:00, 13:00, 14:00 and 15:00 UTC are used to calculate these averages. Originally, the Clark scheme was used to analyse the events during an evening, when biases due to solar heating are smaller. Hence, the Clark scheme may perform poorly during daylight hours when biases in Netatmo readings due to solar heating have more of an effect (Clark et al., 2018, p. 640). There are no limits on the distance to the nearest Met Office site, which means that Netatmo sites isolated large distances from the nearest Met Office sites may be measuring different meteorological features and will be nudged unnecessarily.

For each Netatmo temperature observation  $T_{\text{Netatmo}}$  at some time  $H : 00$ :

1. The mean of the observations on the hour from  $(H - 3) : 00$  until  $(H + 3) : 00$ ,  $\bar{T}_{\text{Netatmo}}$  is found.
2. The nearest Met Office site with observations at  $H : 00$  is found, and its mean,  $\bar{T}_{\text{MetOffice}}$ , is calculated from the hourly observations from  $(H - 3) : 00$  until  $(H + 3) : 00$ .
3. The adjusted Netatmo observation is calculated as  $T_{\text{Clark}} = T_{\text{Netatmo}} - (T_{\text{Netatmo}} - T_{\text{MetOffice}})$ .

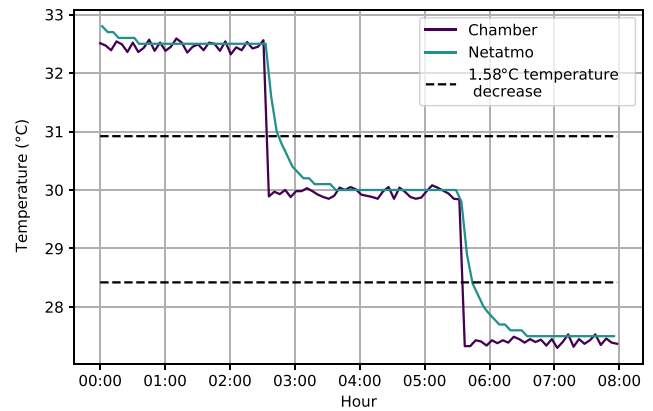
## 4 | RESULTS

In this section, the results from the temperature tests from the PTUCal climate chamber and the UK comparisons of Netatmo and Met Office observations are presented. In addition, the lag in Netatmo observations is studied by comparing a co-located Netatmo and Met Office station.

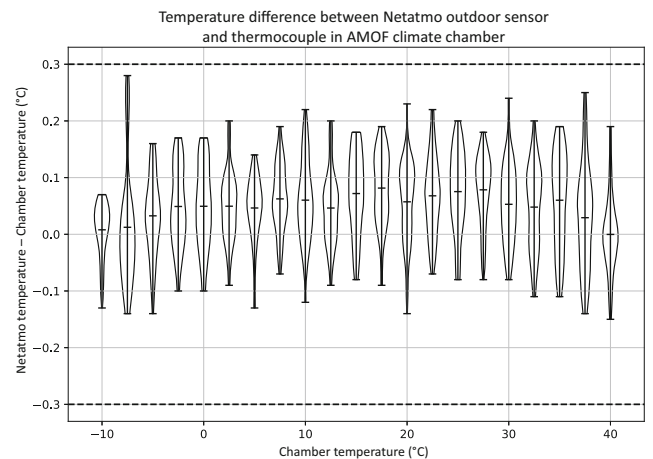
### 4.1 | PTUCal climate chamber experiment

As discussed in Section 3.1, seven Netatmo stations (UoL, UoB1–6) are tested using the PTUCal climate chamber in AMOF at NCAS in Leeds. The validation of the manufacturer-claimed performance in a climate chamber allows the quantification of biases within the sensors themselves, so uncertainties in temperature measurements can be separated into those as a result of the Netatmo manufacture/calibration process and those as a result of a user placing the sensors inadequately.

Figure 3 shows a short section of the temperature experiment on the UoB1 outdoor sensor. The response by the Netatmo temperature sensor to changes in temperature is significantly lagged (quantified later in Figure 5).



**FIGURE 3** Section of a temperature chamber test, showing the lagged response time in the Netatmo temperature sensor compared with the response by the thermocouple in the chamber. The dashed horizontal lines are plotted at  $y = 30.92$  and  $y = 28.42$ , to represent a temperature decrease of  $1.58^{\circ}\text{C}$  from the initial chamber temperature ( $\tau$ ).



**FIGURE 4** Violin plot showing the difference between the Netatmo temperature measurement and the measurement made by the thermocouple in the climate chamber, for all sensors. The readings are from the final hour of each 3-h temperature test, when the Netatmo sensor has fully responded to the change in temperature from the previous value. Thus, 12 measurements (5-min frequency) from seven stations (84 total data points) feed into each individual violin. The bar represents the mean temperature anomaly. The stated accuracy of Netatmo measurements is also plotted ( $\pm 0.3^{\circ}\text{C}$ ). AMOF, Atmospheric Measurement and Observation Facility

However, after the Netatmo readings have adjusted to the new temperature, the Netatmo measurements stay within  $0.3^{\circ}\text{C}$  of the chamber thermocouple observations (and in most cases better than  $0.3^{\circ}\text{C}$ ). The same behaviour is witnessed across all temperature ranges ( $-10$  to  $40^{\circ}\text{C}$ ) and all seven sensors used in the chamber experiments. Netatmo states on their website that the

temperature measurements are accurate to  $0.3^{\circ}\text{C}$  (Netatmo, 2020). Figure 4 shows that observations from Netatmo stations in the final hour of each test are well within  $0.3^{\circ}\text{C}$  of the chamber observations, and in most cases, the Netatmo measurements are within  $0.2^{\circ}\text{C}$  of the measurement from the chamber observations, across all temperatures  $-10$  to  $40^{\circ}\text{C}$  and all seven sensors tested. Any fluctuations in the temperature of the climate chamber will be captured by the thermocouple, but the lagging of the Netatmo response means that some noise in the data may be due to this. The study by Meier et al. (2017) included an investigation of the accuracy of Netatmo temperature observations at seven distinct air temperatures in the range  $0$ – $30^{\circ}\text{C}$  in a climate chamber. Our study investigated the accuracy of Netatmo observations at a greater range of temperatures, and used a reference sensor with a stated temperature accuracy of  $0.1^{\circ}\text{C}$  (see Table 3), in contrast to  $0.28^{\circ}\text{C}$  in the Meier study. The Meier study found all sensors were within the Netatmo accuracy of  $0.3^{\circ}\text{C}$ , except at  $0^{\circ}\text{C}$ , and some sensors showed a slight warm bias of around  $0.5^{\circ}\text{C}$ . In our study, all observations were within  $0.3^{\circ}\text{C}$  of the reference observations, but no station showed a warm bias of a magnitude close to  $0.5^{\circ}\text{C}$ .

The Netatmo temperature sensors only make an observation every 5 min compared with the thermocouple in the chamber, which makes an observation every 2 s (Figure 3 includes thermocouple observations at 5 min intervals to match the gap between Netatmo observations). Figure 5 shows the time between the Netatmo sensor and chamber thermocouple recording a temperature

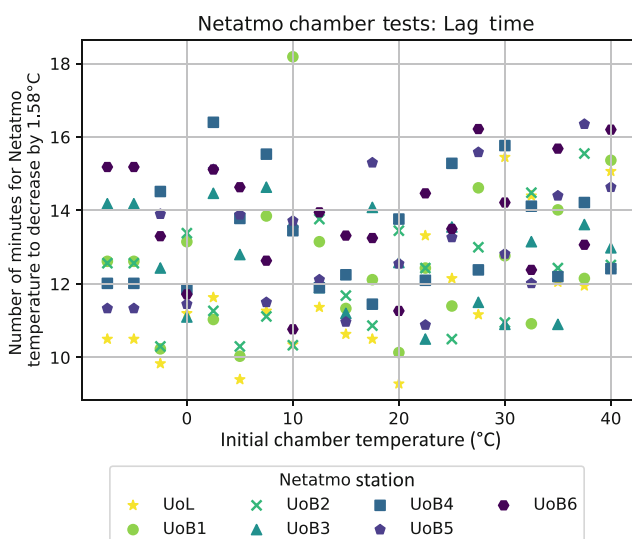
decrease of  $1.58^{\circ}\text{C}$  (63% of  $2.5^{\circ}\text{C}$ :  $\tau$ ). The mean lag time was 12.7 min. At least two Netatmo observations had been made in the intervening time. There is no visible dependence between the temperature of the chamber and the length of time for the Netatmo measurement to have decreased by  $1.58^{\circ}\text{C}$ , which suggests that the lagged response is a systematic bias within the Netatmo sensor. Büchau (2018) also included an investigation of the lagged response to temperature changes, but over a greater step (in excess of  $15^{\circ}\text{C}$ ) than studied here. Our aim here was to have a step change of temperature that was realistic in the United Kingdom. Büchau (2018) used a non-linear least square method to calculate a mean time constant of 22.46 min.

## 4.2 | National observations and quality control schemes

After quantifying the uncertainty characteristics of the sensors in a controlled environment, 1 month of air temperature data from the Netatmo stations within the United Kingdom are analysed, from April 2020. The data are collected using the Netatmo API, and the number of captured stations depends upon the status of the local batch cluster where jobs are run, and the Netatmo API (see Section 2.1). The number of stations varies between days, with a mean of 4667 over the month. The smallest number of stations captured was 1384, and the largest 5381. The Netatmo data are used to compare the performance of the Netatmo stations against Met Office stations, and to characterize any uncertainties exhibited in the Netatmo data.

When originally implemented around Berlin, the Meier scheme (Section 3.3.1) removed around 54% of observations. In this study, only 30% of observations are removed using this scheme. In Norway, the Nipen QC scheme (Section 3.3.2) removed approximately 21% of observations, while in this study, around 11% of UK observations are removed. Unlike the other two schemes, which removed anomalous observations, the Clark QC scheme (Section 3.3.3) adjusted temperatures if they were inaccurate. Exactly 26% of observations are adjusted by more than  $2^{\circ}\text{C}$  by the Clark scheme in this study.

Figure 6 shows the comparison between the effects of each quality control scheme on the UK Netatmo temperature data at midday on 1 April 2020. During this day, the south of England was experiencing a ridge of high pressure, while a cold front moved slowly southwards through Scotland, before decaying (Wetter, 2020). In the unfiltered data plot, there are hot spots (with deviations from the mean in excess of  $10^{\circ}\text{C}$  in cases). The hot spots indicate that there are many Netatmo temperature

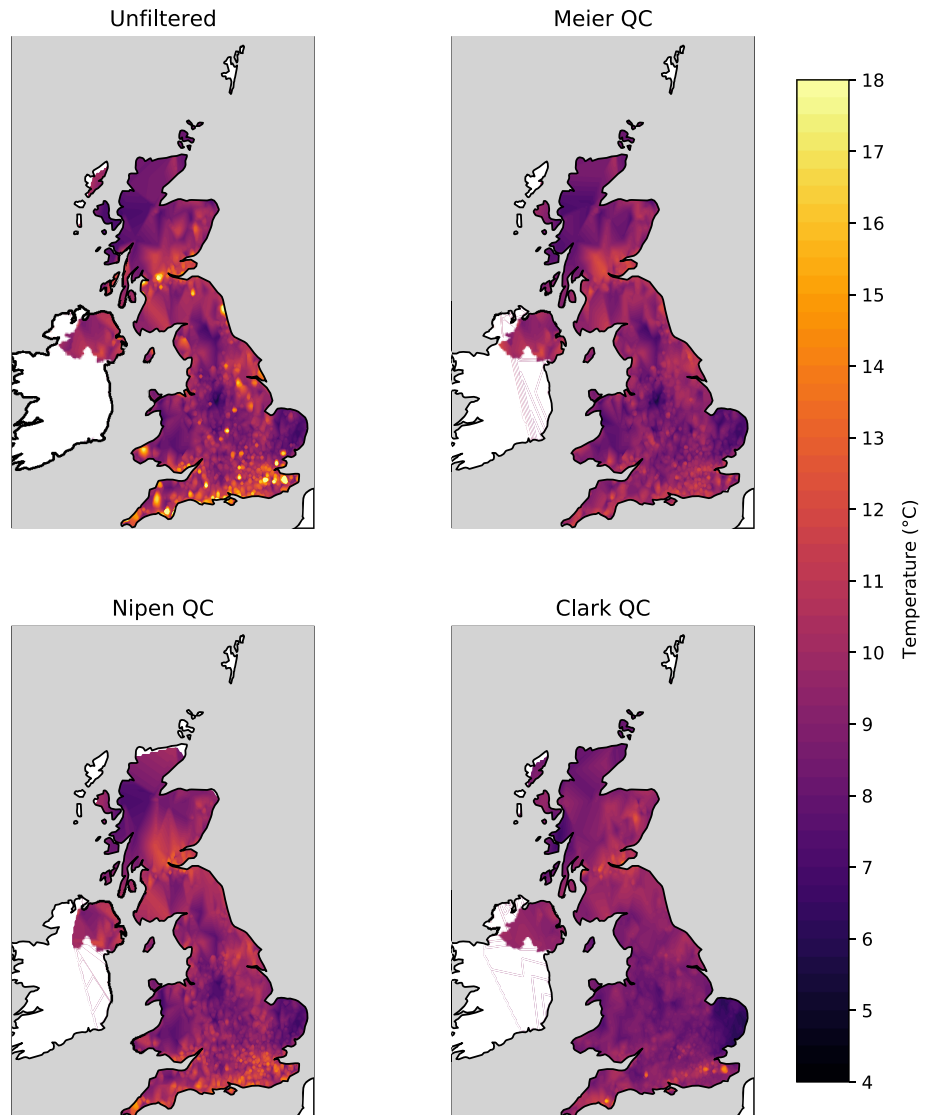


**FIGURE 5** Time taken in minutes for the Netatmo temperature to record a decrease in temperature of  $1.58^{\circ}\text{C}$ , after the chamber had already recorded such a decrease. The mean lag was 12.7 min.



**FIGURE 6** Midday interpolated temperatures on 1 April 2020, with the unfiltered data and the quality controlled data plotted for each of the three schemes. The QC schemes remove or reduce the majority of hot spots in the unfiltered data. There are data from 4933 stations plotted.

UK Quality Control Scheme Comparison 2020-04-01 1200

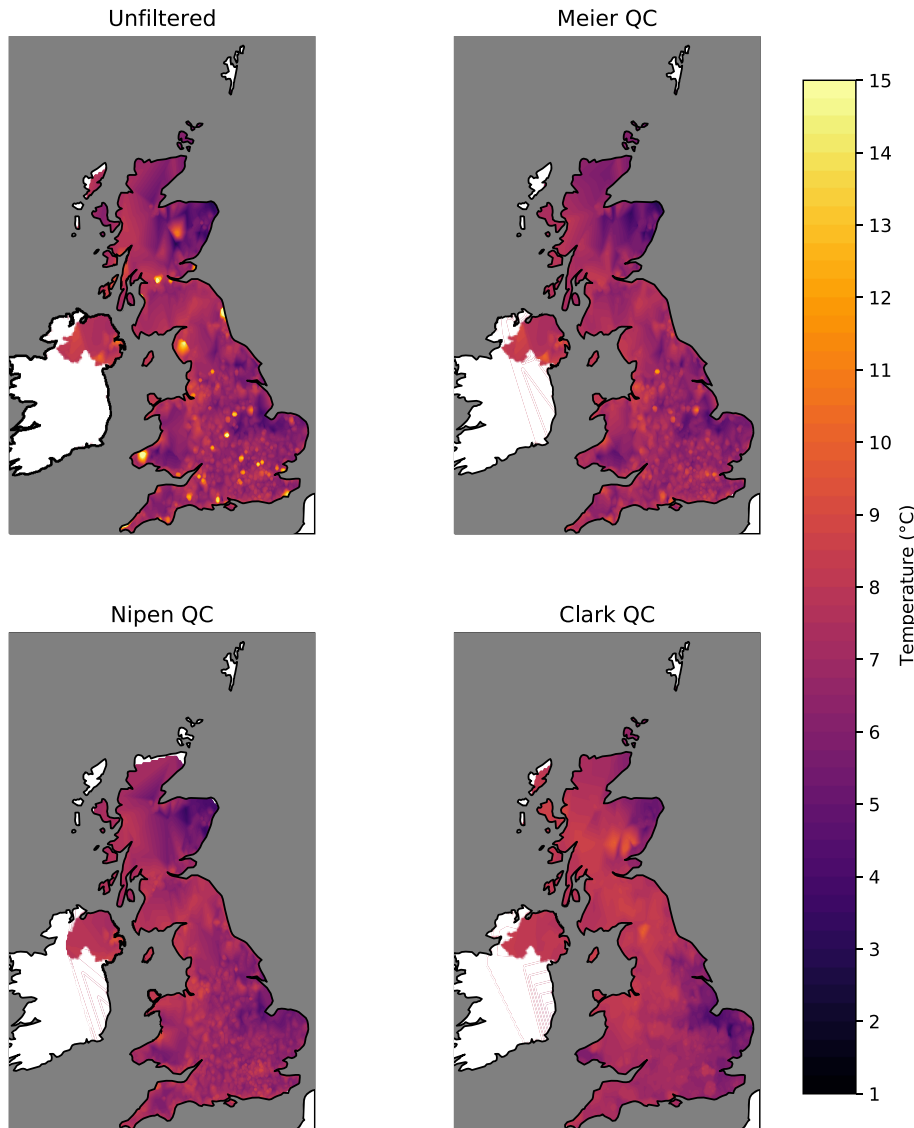


sensors recording unreasonably warm temperatures for outdoor sensors. This is most likely caused by these sensors either being located indoors, being affected by heat from buildings such as near an exhaust vent, or being heated from direct sunlight. A hot spot present during both daytime and nighttime would suggest that there is a station that is left indoors, while a hot spot only present during the day indicates a station has been left outdoors, but is affected by direct sunlight. Figure 6 demonstrates qualitatively the removal of daytime anomalous temperatures by each of the quality control schemes. Due to the interpolation used, there are some areas with missing data (such as Shetland and the Western Isles). The Meier and Nipen schemes both perform similarly, by removing the most extreme hot spots from the data, while the Clark scheme has the smoothest texture of the three schemes.

Qualitatively, all three quality control schemes reduce most of the hot spots in the data. However, there are a few hot spots remaining in the Clark map for midday (Figure 6), which are removed in the Meier and Nipen plots.

Figure 7 shows the interpolated Netatmo temperatures for the United Kingdom at midnight on 2 April 2020. Again, there are significant hot spots in the unfiltered data where Netatmo sensors are recording unrealistically warm temperatures. There are more hot spots in the unfiltered data for the midday temperatures compared with midnight, 12 h later. This indicates that the majority of stations with anomalous temperatures are being affected by incoming solar radiation, while others are likely to be located indoors. The Meier map in Figure 7 shows some warmer pools in the midnight data,

## UK Quality Control Scheme Comparison 2020-04-02 0000



**FIGURE 7** Midnight interpolated temperatures on 2 April 2020, with the unfiltered data and the quality controlled data plotted for each of the three schemes. Again, the QC schemes remove or reduce the majority of hot spots in the unfiltered data. There are data from 4931 stations plotted.

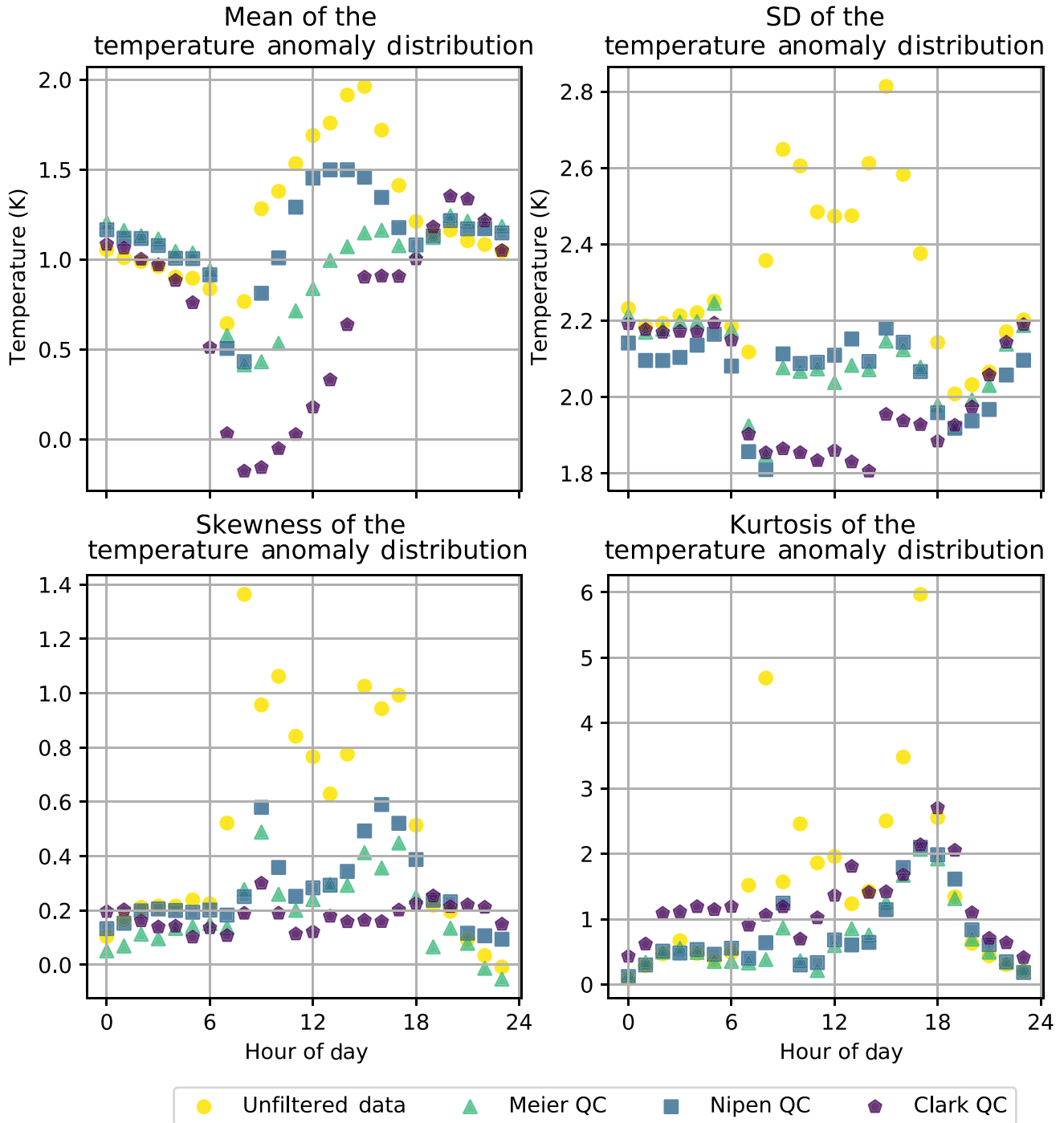
which may be evidence of a UHI effect in cities, while the Nipen and Clark maps show some variation throughout the country.

Figure 8 shows statistical metrics (mean, SD, skewness and kurtosis) for the differences between the gridded Netatmo and point Met Office data, by hour of the day, for April 2020 in southeastern England (as defined in Table 5). This restriction is to ensure that there are Netatmo stations within a reasonable distance to reference stations so that there are no large errors due to the interpolation over data-sparse regions such as the Highlands and Islands of Scotland, where there are large distances between some Met Office sites and the nearest Netatmo station (as shown by Figure 1).

The mean temperature plot shows a diurnal effect in the unfiltered data with the warmest bias during the afternoon, where the Netatmo stations are recording a

warmer temperature than the Met Office: most likely because there are Netatmo stations positioned in direct sunlight or indoors. Indoor measurements are likely to have a greater impact on the temperature difference overnight. The stated accuracy of the Netatmo temperature measurements is  $\pm 0.3^{\circ}\text{C}$ , and the mean temperature is in excess of  $0.3^{\circ}\text{C}$  for each hour of the day for the unfiltered data, the Meier and the Nipen schemes. In the unfiltered data and all the quality controlled data, there is a decrease in the warm bias during the morning, between 0000 and 07:00 UTC. Only the Clark quality controlled data have mean biases less than  $0.3^{\circ}\text{C}$ , between 07:00 and 12:00 UTC. If the Netatmo data set were temporally shifted to account for any lag, this may make the warm bias more uniform throughout the day but would not remove the warm bias. However, the decrease in the warm bias during the morning is likely because of the

Hourly statistical metrics for interpolated Netatmo stations compared with Met Office stations during April 2020 (SE England)



**FIGURE 8** Hourly mean, SD, skewness and kurtosis of the temperature anomaly distribution between interpolated Netatmo data and Met Office data in southeastern England (as defined in Table 5). The unfiltered data and quality controlled data using the three schemes discussed are plotted.

lagged response to changes in temperature, as noted in Section 4.1. The change in SD between the quality control schemes is indicative of their effectiveness, but the actual SD value has a fundamental minimum due to the approach used here of gridding Netatmo stations onto

Met Office sites. The interpolation does increase the spread of the data in Figure 8 because of the natural variability in the interpolation to a new point. If Netatmo stations were co-located with the Met Office sites, then the SD in Figure 8 would be smaller.

**TABLE 4** Summary of the statistics for Netatmo and Met Office temperature differences during April 2020, for the unfiltered data and each QC scheme

Scheme	Mean	SD	Skewness	Kurtosis	RMSE
Unfiltered	1.23	2.36	0.66	2.09	2.66
Meier	0.98	2.11	0.20	0.67	2.32
Nipen	1.12	2.08	0.29	0.78	2.36
Clark	0.70	2.06	0.21	1.09	2.18

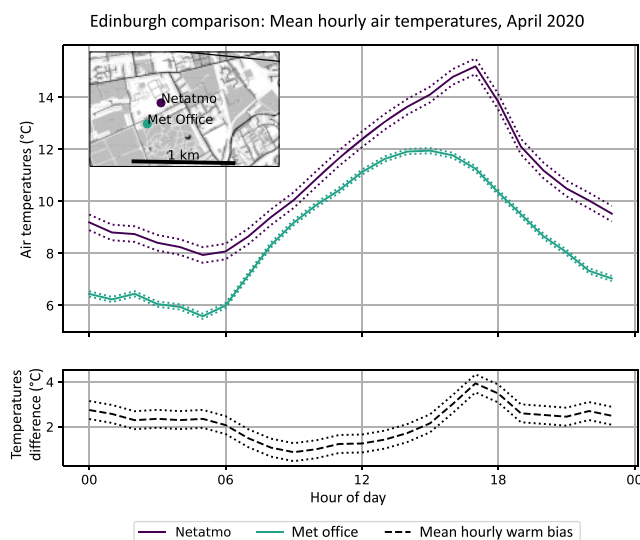
Abbreviation: RMSE, root mean square error.

The overall statistics for the temperature differences between the interpolated Netatmo data (unfiltered and the quality controlled data) and Met Office data are shown in Table 4. As shown in Figure 8, the SD of the temperature differences in the unfiltered data is at its maximum during the afternoon, at the same time as the peak mean difference in temperature. The QC schemes reduce the SD of the warm bias, most notably during daylight hours. This indicates that the schemes remove data from stations making observations in direct sunlight. The warm biases are positively skewed, with the unfiltered data notably skewed during daylight hours, indicating that solar heating is occurring. All three schemes reduce the skewness of the warm bias during daylight hours, although the biases remain positively skewed in all cases except for late evening in the Meier scheme. The kurtosis of the quality controlled data varies throughout the day, but does not exceed 3. This indicates that the temperature differences, while positively skewed, have fewer outliers than would be expected if the differences were normally distributed. The root mean square error of the unfiltered temperature differences is  $2.66^{\circ}\text{C}$ , which was reduced (but still in excess of  $2^{\circ}\text{C}$ ) by the three QC schemes.

### 4.3 | Lag time

The mean temperature difference between the Netatmo and Met Office observations decreases during the morning. The warmup and cooldown experiments with and without the temperature sensor casing performed by Büchau (2018) demonstrated a marked decrease in the time constant when the Netatmo cover was removed, with the lag time decreasing from 26.89 (22.46) min to 15.40 (14.02) min in the warmup (cooldown) tests. This is consistent with the hypothesis that the lagged response to temperature changes by Netatmo sensors is because of the casing. The lagged response may explain the decrease in the mean temperature difference during the morning in Figure 8.

One hypothesis is that the Netatmo temperature sensor experiences a lag when the air temperature changes



**FIGURE 9** Mean air temperature for each hour of the day in April 2020, for a Netatmo station and Met Office station in Edinburgh, 250 m apart. Note that the difference between mean Netatmo and Met Office observations decreases during the morning. Stated uncertainties for each measurement are shown as dotted lines. Map tiles from Stamen Design and OpenStreetMap contributors (2021)

(as shown by Figure 3 and the tests in Section 4.1). As the ambient air temperature increases after sunrise, the Netatmo temperature sensor records an artificially low temperature because of this lag, yet still higher than the reference observations because of the general warm bias. There is variation in the reduction in the warm bias between different days that is hypothesized to be due to the changing intensities of solar heating, although the evidence was not found in this study to support such a hypothesis. Future studies should examine the Netatmo biases with co-located radiance measurements to answer this hypothesis.

In Edinburgh, there is a Netatmo station located approximately 250 m away from the Met Office surface station. This Netatmo station is just one station out of the network, so may not be representative of the placement uncertainty observed across the whole crowdsourced

network. Figure 9 shows the mean hourly air temperature during all hours in April 2020 for both stations. While similar air temperatures are recorded, the difference between the mean Met Office and Netatmo observations decrease in the morning, between 06:00 and 12:00 UTC. The decrease in the temperature difference may be explained by the rate of solar heating being maximum during the morning coupled with the lagged response of Netatmo observations. Since this warm bias was present throughout the entire day, and not shown during the PTUCal chamber tests (Section 4.1), this anomaly must be positional. Therefore, studies using Netatmo data for quantification of UHI effects (such as Chapman et al., 2017) must be aware of the fine line between a UHI effect and observations that are artificially warm due to their placement.

## 5 | DISCUSSION

In the United Kingdom, there is a relatively dense network of reference Met Office stations. In many countries however, observation networks are less dense, so the potential benefits of crowdsourced observations are greater. For example, the Oklahoma Mesonet in the United States has an average station spacing of 29 km (Ziolkowska et al., 2017). The Met Office surface stations in the United Kingdom, with an average spacing of 17.6 km, are at a greater spatial resolution than the Oklahoma Mesonet. Therefore, Netatmo sensors will add the most value to locations with sparse surface observations, and little infrastructure to maintain costly surface stations.

Table 5 shows the density of Netatmo and Met Office stations in the United Kingdom, as well as in southeastern England, and northern Scotland. The distance between Netatmo weather stations is short enough to theoretically enable observations of meso- $\gamma$  phenomena that the Met Office surface sites may miss. However, the UK Netatmo temperature observations, even those that have been quality controlled, still exhibit mean warm biases in excess of 1°C. The mean distance between Netatmo stations in northern Scotland is almost 3.5 times the mean distance for stations in the southeast of England. The spread between Met Office stations in northern Scotland is only 1.3 times the spread in the southeast of England. Hence, Netatmo observations are more likely to add value to forecasts around urban areas. The 17.6 km spacing of Met Office stations means that there are large gaps that crowdsourced observations could fill. The spread of Netatmo stations around the United Kingdom is not even, and there are far more stations, densely packed, in the southeast of England than in northern Scotland, as demonstrated in Table 5. If an appropriate quality control scheme is applied, then Netatmo observations may prove effective at forecasting UHI effects in urban areas.

The isolation test from the Nipen et al. (2020) quality control scheme removes stations that are outlying in very rural areas. The radiation correlation technique from the Meier et al. (2017) scheme is useful to remove stations influenced by solar heating, and the section to remove stations with identical latitude and longitude is important, since otherwise, the location of a Netatmo sensor may be completely incorrect, adversely affecting observations. All

TABLE 5 Comparison of the distances in km between the nearest Netatmo stations for different QC schemes and also the distances between the nearest Met Office stations in the United Kingdom

Distance between stations (km)	Unfiltered Netatmo data	Meier QC	Nipen QC	Clark QC	Met Office (MIDAS)
UK mean	2.14	2.67	2.00	2.14	17.6
UK minimum	0.00	0.00157	0.00	0.00	0.0787
UK maximum	78.3	78.3	37.8	78.3	76.1
SE England mean	1.49	1.94	1.51	1.49	16.5
SE England minimum	0.00	0.00157	0.00	0.00	0.0787
SE England maximum	12.9	16.5	12.9	12.9	35.9
Northern Scotland mean	5.10	5.64	3.91	5.10	21.6
Northern Scotland minimum	0.0110	0.0110	0.0110	0.0110	2.11
Northern Scotland maximum	43.6	43.6	37.8	43.6	56.6

Note: Southeastern England (for these purposes, defined as the part of the United Kingdom between -2° and 2° E in longitude, and 50.5°-53° N in latitude) and northern Scotland (defined as the portion of the UK north of 56° N) for the unfiltered Netatmo data and each quality control scheme, as well as the distances between Met Office stations for the same regions.

Abbreviation: MIDAS, Met Office Integrated Data Archive System.

the schemes examined here improved on the warm bias present throughout the data, but a scheme that fully addresses the lagged response to temperature changes (discussed in Section 4.3) may prove more effective at removing some of the warm bias in the data. However, since the bias is present throughout the entire day (not just during daylight hours), a stronger approach to removing anomalous data from Netatmo stations is required in a quality control scheme, perhaps by combining features from all three schemes. The hot spots shown in the maps in Figures 6 and 7 are larger than any individual mesoscale effect that may be present (such as urban heat islands). Drawing conclusions from this data must take into account warm biases due to station placement. Although the three quality control schemes discussed in Section 3.3 did remove the hot spots, there still remained a warm bias

in the data throughout the day. However, this warm bias was not present in the results of the PTUCal chamber tests, presented in Section 4.1. Hence, there must still be some results from incorrect placements in the data, despite the largest deviations being removed. Quality control schemes, while effective at removing perceived hot spots in the data, could be stripping out genuine variations in temperature, especially where direct comparison to reference data is made by the scheme (such as in the Meier and Clark schemes). The Meier scheme was originally used for a single city and so using all stations to compute averages may not be entirely appropriate when considering an entire country, as in our case. The linear bias correction used in the Clark scheme assumes that the bias between the two instruments is also linear over time. Figure 9 shows that the bias is non-linear over the diurnal

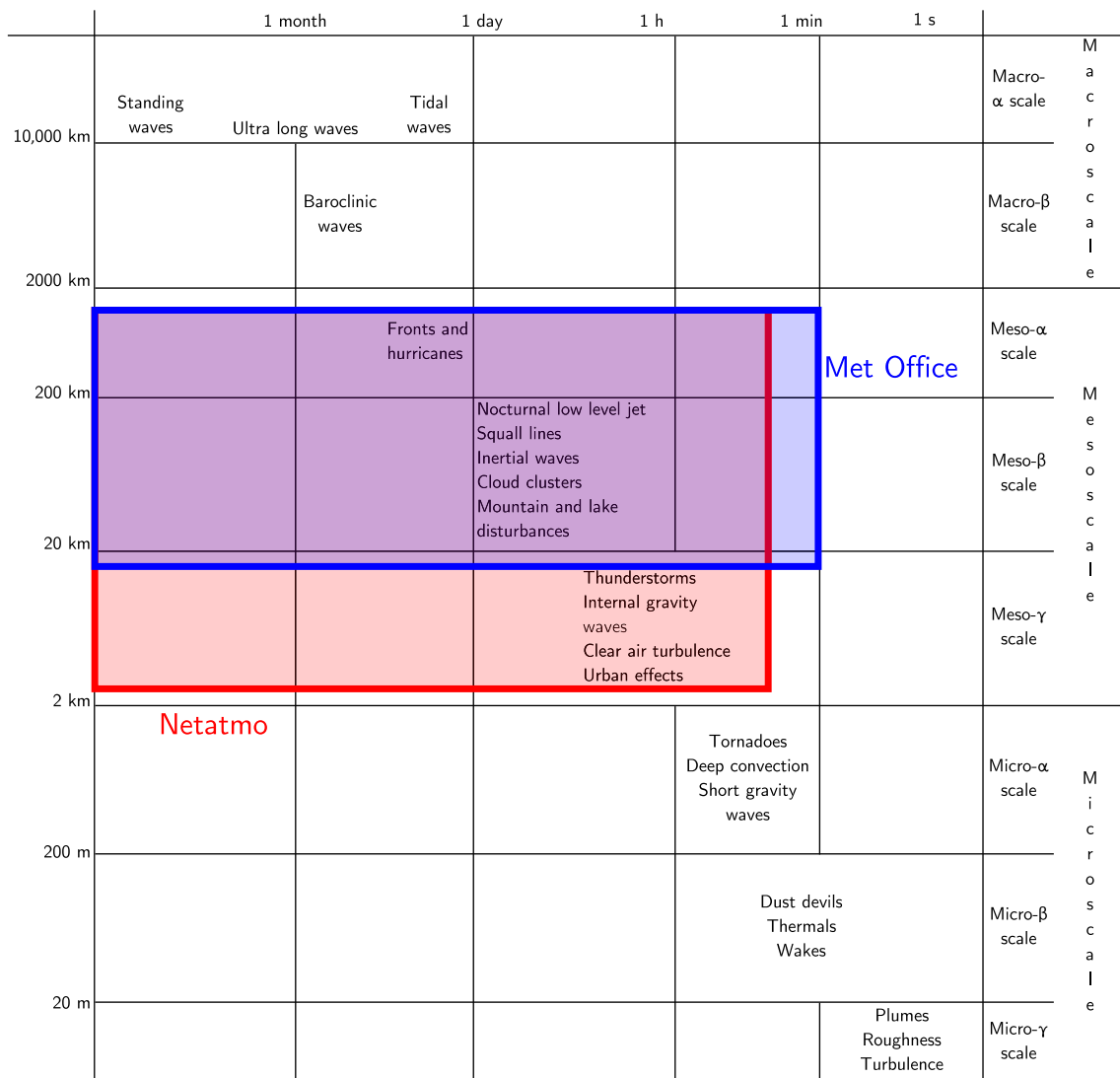


FIGURE 10 Schematic describing various meteorological phenomena, adapted from Orlandi (1975). The blue box represents events that Met Office surface stations can currently observe, and the red box represents the additional features that Netatmo sensors may be able to resolve once anomalous stations are removed. Note that these classifications are not strict.

cycle, due to the significant lag time (and possibly also solar heating) that the Netatmo suffers from. The result is that the Clark linear bias correction does not fully compensate for the difference between the individual Netatmo measurements with a 6-hourly bias correction. A shorter time window, such as 1- or 3-hourly bias correction, would take into account more of the non-linearity of the diurnal bias between the Netatmo and the Met Office stations—but may then suffer locale biases if the two sites are not co-located.

In addition, the lagged response to temperature changes by Netatmo sensors will result in any quick increase then decrease in temperature being severely smoothed in the Netatmo data. The chamber testing in Section 4.1 shows that the Netatmo stations are slow to react to a rapid change in temperature, which Clark et al. (2018) showed can occur in the United Kingdom with an example of a 3.2°C temperature change in 8 min. Note that the Netatmo sampling time of 5 min would be inadequate in that scenario. A lag time of 12.7 min to record a 63% decrease in temperature is too coarse to resolve the finest mesoscale (meso- $\gamma$ ) and for any microscale phenomena. The significant lag time, and to a lesser extent the sample frequency of 5 min, reduces the potential impact of using Netatmo observations to capture events not currently observed by traditional meteorological sensor networks.

From this work, a schematic describing scales of meteorological phenomena (Figure 10) has been constructed. It shows features that a properly calibrated and quality checked network of Netatmo weather stations are capable of resolving, compared with the Met Office observation network. While Met Office observations every minute are available directly through the MMS, Netatmo stations only make an observation every 5 min. However, the long lag time observed in Netatmo temperature readings, as shown in Section 3.1, means that Netatmo stations are not capable of resolving phenomena on a 5-min timescale, since the mean temperature lag time was 12.7 min in the chamber. Figure 10 shows Netatmo stations make sufficient observations and are close enough together to be able to resolve urban meteorological effects, albeit on timescales longer than 15 min.

The observations studied here are from just 1 month, April 2020, which was the sunniest April on record in the United Kingdom (Met Office, 2020a). An analysis of the observations from Netatmo stations over a longer time period would prove useful to investigate how the warm bias differs throughout the year. With April 2020 being as sunny as it was, the Netatmo temperature sensors may have been influenced by solar heating more than during the rest of the year, potentially contributing to the warm bias observed in the data.

## 6 | CONCLUSION

The validation and evaluation of temperature observations from Netatmo weather stations in this study means that recommendations can be made on how to use data from Netatmo stations. Netatmo observations are more likely to add value in urban areas than rural areas, where anomalies are more obvious and meso- $\gamma$  phenomena are more likely to be observed due to the higher density of observations. Netatmo stations observe more than temperature, so an investigation into how other variables compare to a trusted network of observations may prove more successful, and may not experience the lag time that Netatmo temperature measurements do. However, there remains uncertainty within the observations because of their placement. As noted above, the time period of this study was the sunniest April on record in the United Kingdom, and the resultant solar heating influences on Netatmo temperature observations can be seen in the data in Figure 8. The unfiltered data show a large SD and positively skewed data during daylight hours, which are mostly removed by the quality control schemes. This shows that the QC schemes do have a good effect on the data, but an overall warm bias remains in the quality controlled data throughout the day.

Additionally, the lagged response to changes in temperature discussed in Section 4.3 means that Netatmo data are unsuitable for observing temperature on a timescale of less than 15 min because any phenomenon that lasts for less than 15 min would be severely smoothed in the temperature observations.

Netatmo stations are popular in the United Kingdom (Table 1 shows that 6098 UK Netatmo stations were releasing data publicly during some of April 2020), and any competing sensor would take time to achieve a similar number of stations to Netatmo ones. One option to help improve trust in Netatmo observations would be to provide more guidance to owners of Netatmo stations on how best to place their stations outside: currently, there is little information provided within the instruction booklet. Alternatively, other quality control schemes could be tested, which also account for the lagged response in observations. Lastly, other sensors or accessories may be produced to account for biases, such as a Stevenson screen for the temperature sensor to reduce solar heating effects. The lagged response to temperature observations by Netatmo sensors is unfortunate, as it becomes difficult to account for sensor biases. Technological trends change over time, and a new popular sensor may be invented, which observes more frequently than at 5-min intervals, that is less affected by incoming solar radiation and suffers little lag in its observations. The meteorological community will need to keep up with current trends within

consumer weather stations when deciding how to get optimal data from crowdsourcing. Long-term drift is a concern for consumer weather stations that do not receive regular recalibration like national meteorological service stations do. To test this, a future study could either keep a station on a site for many years and then redo a calibration chamber test, or examine a historical data set of public Netatmo stations and compare the newer stations with the old. Future evaluation of the QC schemes should refine the gridding and interpolation to only resolve a reference station if there are sufficient nearby Netatmo stations.

There is great potential offered by crowdsourcing meteorological observations to supplement existing surface observations, especially in parts of the world that have surface stations at greater distances than those used by the Met Office in the United Kingdom (17.6 km on average). As the resolution of numerical weather prediction (NWP) models increases, a denser network of observations will be required if processes on the scale of 1 km need to be resolved, and crowdsourcing observations are an already existing set of observations to fill the gap. The potential offered by crowdsourcing observations, from home weather stations as discussed here, or vehicle-based observations (Z. Bell et al., 2022; Mahoney & O'Sullivan, 2013), mean that there is a wealth of high-resolution surface data to help forecasting centres address the challenges of observing and modelling weather into the future.

Amateur meteorologists have kept valuable records from their personal weather stations for decades, yet it is the invention of 'smart home', internet-connected products that has allowed their observations to be shared in real time. The number of household objects, not just home weather stations, with the capability to automatically upload data to the Web is increasing. Automation of society is becoming more commonplace, and relies on local measured data for contextual awareness. It is unlikely that consumers will stop owning such 'smart' products. Sensors are getting smaller, cheaper and more connected, so it is likely that forecasting centres will continue to have a wealth of crowdsourced data available to them through the inclusion of sensors in consumer products: for example, many smartphones contain barometers (Hintz et al., 2019). Continued improvements to computing power will increase the ability of forecasting centres to synthesize large volumes of data through assimilation; the successful advent of satellite data assimilation being a prime beneficiary thus far. Finally, the resolution of NWP models is perpetually increasing, and will require high-resolution observation networks that are uneconomical for budget-constrained national meteorological

services to deploy and maintain. Crowdsourced observation networks show some promise to be a scalable solution. Given the refinement of sensor technology and schemes to effectively assimilate crowdsourced data, these observations will increase in quality and therefore in usefulness and economic value.

## AUTHOR CONTRIBUTIONS

**Jonathan Coney:** Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); project administration (equal); software (lead); validation (lead); writing – original draft (lead); writing – review and editing (lead). **Ben Pickering:** Conceptualization (lead); data curation (supporting); funding acquisition (lead); methodology (equal); project administration (equal); resources (lead); software (supporting); supervision (lead); validation (supporting); writing – review and editing (equal). **David Dufton:** Conceptualization (equal); methodology (equal); project administration (equal); supervision (equal); validation (supporting); writing – review and editing (supporting). **Maryna Lukach:** Conceptualization (equal); methodology (equal); project administration (equal); supervision (equal); validation (supporting); writing – review and editing (supporting). **Barbara Brooks:** Methodology (supporting); resources (lead); software (supporting); supervision (supporting); validation (supporting); writing – review and editing (supporting). **Ryan Neely:** Funding acquisition (equal); supervision (supporting); writing – review and editing (supporting).

## ACKNOWLEDGEMENTS

Thanks go to Sophie Randall and Simon Bell at the Birmingham Urban Observatory, University of Birmingham, for lending the project six Netatmo weather stations and answering our questions. Thanks also to the National Centre for Atmospheric Science's Atmospheric Measurement and Observation Facility (AMOF) for enabling this research. Finally, thank you to the anonymous referees whose feedback, comments and suggestions improved the manuscript. This work used JASMIN, the United Kingdom's collaborative data analysis environment (<http://jasmin.ac.uk>).

## FUNDING INFORMATION

Jonathan Coney was funded by a NERC DTP award (NE/S007458/1), and Ben Pickering was funded by a NERC Industrial CASE Studentship (NE/N008359/1). Publication fees were funded by the University of Leeds Library through the UKRI Open Access grant. David Dufton, Maryna Lukach, Barbara Brooks and Ryan R. Neely III were funded through the NCAS underpinning science programme through UKRI.



## CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

## DATA AVAILABILITY STATEMENT

The UK-wide Netatmo data used in this study was from weather stations whose owners had opted in to sharing their weather station data publicly. Data from the Netatmo weather stations are publicly available via the Netatmo API: <https://dev.netatmo.com/apidocumentation/weather>. The Met Office MIDAS weather observation data are available from the NCAS British Atmospheric Data Centre: <https://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0>. The population data from Figure 1 is available from the UK Centre for Ecology & Hydrology under the Open Government Licence. ©NERC (Centre for Ecology & Hydrology). Contains National Statistics data. ©Crown copyright and database right 2011: <https://doi.org/10.5285/0995e94d-6d42-40c1-8ed4-5090d82471e1>.

## ORCID

Jonathan Coney  <https://orcid.org/0000-0001-7310-8002>

Ben Pickering  <https://orcid.org/0000-0002-8474-9005>

## REFERENCES

- Bauer, P., Thorpe, A. & Brunet, G. (2015) The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47–55. <https://doi.org/10.1038/nature14956>
- Bell, S., Cornford, D. & Bastin, L. (2015) How good are citizen weather stations? Addressing a biased opinion. *Weather*, 70(3), 75–84. <https://doi.org/10.1002/wea.2316>
- Bell, Z., Dance, S.L. & Waller, J.A. (2022) Exploring the characteristics of a vehicle-based temperature dataset for kilometrescale data assimilation. *Meteorological Applications*, 29(3), e2058. <https://doi.org/10.1002/met.2058>
- Benjamin, S.G., Brown, J.M., Brunet, G., Lynch, P., Saito, K. & Schlatter, T.W. (2019) 100 years of progress in forecasting and NWP applications. *Meteorological Monographs*, 59, 13.1–13.67. <https://doi.org/10.1175/amsmonographs-d-18-0020.1>
- Büchau, Y.G. (2018) Modelling shielded temperature sensors an assessment of the Netatmo citizen weather station. MSc Thesis, University of Hamburg.
- Chapman, L., Muller, C.L., Young, D.T., Warren, E.L., Grimmond, C. S.B., Cai, X.M. et al. (2015) The Birmingham urban climate laboratory: an open meteorological test bed and challenges of the Smart city. *Bulletin of the American Meteorological Society*, 96(9), 1545–1560. <https://doi.org/10.1175/BAMS-D-13-00193.1>
- Chapman, L., Bell, C. & Bell, S. (2017) Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat Island of London quantified using Netatmo weather stations. *International Journal of Climatology*, 37(9), 3597–3605. <https://doi.org/10.1002/joc.4940>
- Clark, M.R., Lee, D.S. & Legg, T.P. (2014) A comparison of screen temperature as measured by two Met Office observing systems. *International Journal of Climatology*, 34(7), 2269–2277. <https://doi.org/10.1002/joc.3836>
- Clark, M.R., Webb, J.D.C. & Kirk, P.J. (2018) Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations. *Meteorological Applications*, 25, 472–492. <https://doi.org/10.1002/met.1715>
- De Vos, L., Leijnse, H., Overeem, A. & Uijlenhoet, R. (2017) The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrology and Earth System Sciences*, 21(2), 765–777. <https://doi.org/10.5194/hess-21-765-2017>
- Dunne, S. & Entekhabi, D. (2005) An ensemble-based reanalysis approach to land data assimilation. *Water Resources Research*, 41(W02013), 1–18. <https://doi.org/10.1029/2004WR003449>
- ECMWF. (2020) *ECMWF forecasts*. Available at: <https://apps.ecmwf.int/webapps/opencharts> [Accessed 22nd September 2020].
- Elmore, K.L., Flaming, Z.L., Lakshmanan, V., Kaney, B.T., Farmer, V., Reeves, H.D. et al. (2014) MPING: crowdsourcing weather reports for research. *Bulletin of the American Meteorological Society*, 95(9), 1335–1342. <https://doi.org/10.1175/BAMS-D-13-00014.1>
- Fiebrich, C.A. (2009) History of surface weather observations in the United States. *Earth-Science Reviews*, 93(3), 77–84. <https://doi.org/10.1016/j.earscirev.2009.01.001>
- Green, A. (2010) From observations to forecasts – part 7. A new meteorological monitoring system for the United Kingdom's Met Office. *Weather*, 65(10), 272–277.
- 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(2), 160–173. <https://doi.org/10.1002/gdj3.79>
- Hintz, K.S., Vedel, H. & Kaas, E. (2019) Collecting and processing of barometric data from smartphones for potential use in numerical weather prediction data assimilation. *Meteorological Applications*, 26(4), 733–746. <https://doi.org/10.1002/met.1805>
- 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(2), 151–159.
- Inness, P. & Dorling, S. (2013) *Operational weather forecasting*. Oxford: Wiley-Blackwell.
- Lackstrom, K., Farris, A., Eckhardt, D., Doesken, N., Reges, H., Turner, J. et al. (2017) CoCoRaHS observers contribute to “condition monitoring” in the Carolinas: a new initiative addresses needs for drought impacts information. *Bulletin of the American Meteorological Society*, 98(12), 2527–2531. <https://doi.org/10.1175/bams-d-16-0306.1>
- Lawrence, B.N., Bennett, V.L., Churchill, J., Juckes, M., Kershaw, P., Pascoe, S. et al. (2013) Storing and manipulating environmental big data with JASMIN. In: *2013 IEEE international conference on big data*. IEEE. Available at: <https://doi.org/10.1109/bigdata.2013.6691556>
- Lussana, C., Uboldi, F. & Salvati, M.R. (2010) A spatial consistency test for surface observations from mesoscale meteorological networks. *Quarterly Journal of the Royal Meteorological Society*, 136(649), 1075–1088. <https://doi.org/10.1002/qj.622>
- Mahoney, W.P. & O'Sullivan, J.M. (2013) Realizing the potential of vehicle-based observations. *Bulletin of the American Meteorological Society*, 94(7), 1007–1018. <https://doi.org/10.1175/bams-d-12-00044.1>
- Mandement, M. & Caumont, O. (2019) Contribution of personal weather stations to the observation of deep-convection features near the ground. *Natural Hazards and Earth System Sciences Discussions*, 20(1), 1–35. <https://doi.org/10.5194/nhess-20-299-2020>

- Meier, F., Fenner, D., Grassmann, T., Jänicke, B., Otto, M. & Scherer, D. (2015) Challenges and benefits from crowdsourced atmospheric data for urban climate research using Berlin, Germany, as testbed. In: *ICUC9–9th international conference on urban climate jointly with 12th symposium on the urban environment*. Available at: <http://www.cocorahs.org/>
- Meier, F., Fenner, D., Grassmann, T., Otto, M. & Scherer, D. (2017) Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, 19, 170–191.
- Met Office. (2012) *Met Office integrated data archive system (MIDAS) land and marine surface stations data (1853–current)*. Available at: <https://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0> [Accessed 20th May 2020].
- Met Office. (2018a) *The UK land observation network*. Available at: <https://www.metoffice.gov.uk/weather/guides/observations/uk-observations-network> [Accessed 21st September 2020].
- Met Office. (2018b) *How we measure temperature*. Available at: <https://www.metoffice.gov.uk/weather/guides/observations/how-we-measure-temperature> [Accessed 27th July 2020].
- Met Office. (2019) *Numerical weather prediction models*. Available at: <https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model/weather-forecasting> [Accessed 20th July 2020].
- Met Office. (2020a) *Record-breaking April sunshine*. Available at: <https://www.metoffice.gov.uk/about-us/press-office/news/weather-and-climate/2020/2020-april-stats> [Accessed 8th March 2021].
- Met Office. (2020b) *MIDAS data user guide for UK land observations*. Available at: [http://cedadocs.ceda.ac.uk/1488/7/MIDAS\\_User\\_Guide\\_for\\_UK\\_Land\\_Observations\\_Version20200921.pdf](http://cedadocs.ceda.ac.uk/1488/7/MIDAS_User_Guide_for_UK_Land_Observations_Version20200921.pdf) [Accessed 8th April 2021].
- Michell Instruments. (2018) *S8000 Remote high precision chilled mirror hygrometer*. Available at: [http://www.michell.com/uk/products/s8000\\_remote.htm](http://www.michell.com/uk/products/s8000_remote.htm) [Accessed 30th July 2020].
- Netatmo. (2012) *Netatmo connect APIs*. Available at: <https://dev.netatmo.com/en-US/resources/technical/reference> [Accessed 4th December 2019].
- Netatmo. (2020) *Smart home weather station's technical specifications*. Available at: <https://www.netatmo.com/en-us/weather/weatherstation/specifications> [Accessed 4th December 2019].
- Nipen, T.N., Seierstad, I.A., Lussana, C., Kristiansen, J. & Hov, Ø. (2020) Adopting citizen observations in operational weather prediction. *Bulletin of the American Meteorological Society*, 101(1), E43–E57. <https://doi.org/10.1175/BAMS-D-18-0237.1>
- Orlanski, I. (1975) A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, 56(5), 527–530.
- Parker, D.E., Legg, T.P. & Folland, C.K. (1992) A new daily central England temperature series 1772–1991. *International Journal of Climatology*, 12(4), 317–342. <https://doi.org/10.1002/joc.3370120402>
- Reis, S., Liska, T., Steinle, S., Carnell, E., Leaver, D., Roberts, E. et al. (2017) *UK gridded population 2011 based on census 2011 and land cover map 2015*. Available at: <https://doi.org/10.5285/0995e94d-6d42-40c1-8ed4-5090d82471e1> [Accessed 13th April 2021].
- Stamen Design and OpenStreetMap Contributors. (2021) *Stamen maps*. Available at: <http://maps.stamen.com/> [Accessed 19th February 2021].
- STMicroelectronics Pty Ltd. (2012) *MEMS pressure sensor: 260–1260 mbar absolute digital output barometer*. Available at: <https://media.digikey.com/pdf/DataSheets/STMicroelectronicsPDFS/LPS331AP.pdf> [Accessed 30th July 2020].
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D. et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wetter. (2020) *Wetter Met Office synoptic chart archive*. Available at: [http://www1.wetter3.de/archiv\\_ukmet\\_dt.html](http://www1.wetter3.de/archiv_ukmet_dt.html) [Accessed 27th October 2020].
- Wilks, D.S. (2011) *Statistical methods in the atmospheric sciences*, 3rd edition. Oxford: Academic Press.
- WMO. (2018) *Guide to instruments and methods of observation. Volume I – measurement of meteorological variables*. Geneva: World Meteorological Organization.
- Ziolkowska, J.R., Fiebrich, C.A., Carlson, J.D., Melvin, A.D., Sutherland, A.J., Kloesel, K.A. et al. (2017) Benefits and beneficiaries of the Oklahoma Mesonet: a multisectoral ripple effect analysis. *Weather, Climate, and Society*, 9(3), 499–519. <https://doi.org/10.1175/WCAS-D-16-0139.1>

**How to cite this article:** Coney, J., Pickering, B., Dufton, D., Lukach, M., Brooks, B., & Neely, R. III (2022). How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom. *Meteorological Applications*, 29(3), e2075. <https://doi.org/10.1002/met.2075>