Can crowdsourcing increase the durability of an urban meteorological network?

Chapman, Lee; Bell, Simon; Randall, Sophie

DOI:
10.1016/j.uclim.2023.101542

Urban Climate 49 (2023) 101542

Contents lists available at ScienceDirect

Urban Climate

journal homepage: www.elsevier.com/locate/uclim

Can crowdsourcing increase the durability of an urban meteorological network?

Lee Chapman *, Simon Bell, Sophie Randall

University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

A R T I C L E I N F O

Keywords:
Crowdsourcing
Opportunistic sensing
Transient sensing
Urban meteorological network
Sustainability
Citizen science
Responsible innovation
Urban observatory
Living lab

A B S T R A C T

Dense networks of weather stations in urban areas are now becoming an increasingly common, but expensive way of monitoring the urban climate. The expense is not only related to the initial deployment of the network, but also the ongoing maintenance which can be increasingly problematic if monitoring is needed for extended periods of time. As an alternative, crowdsourcing weather data provides an effectively free option to acquire hyperlocal weather data, but is not yet fully accepted by the user community due to data quality concerns. This paper explores an approach between these two options to assess whether a more durable (i.e. longer term, less transient) means of monitoring can be achieved by using a controlled deployment of 'low-cost' Netatmo weather station maintained by citizens. As a result, the novelty of this paper resides in the first assessment of the durability of crowdsourcing as opposed to just the data quality alone. The results show that expert input at the deployment stage in terms of device modification and siting improves data quality, but the length and completeness of the record of unsupported weather stations is highly variable with less than 50% of devices surviving the measurement campaign, and far fewer providing full data records. Better durability, albeit at the expense of data quality, was exhibited by pre-existing Netatmo devices procured by the public in the study area with many units producing useable data throughout and beyond the measurement campaign. Overall, these are significant findings which need to be taken into account at the design stage of an urban meteorological network and prompt a rethink in the way that equipment is deployed via living lab approaches. Fundamentally, existing approaches are all largely built on a paradigm of transient sensing methodologies which are difficult to reconcile with a science that uses thirty-year baselines.

1. Background

The last decade has seen the rise of Urban Meteorological Networks (UMN's) as an emerging standard for the high-resolution monitoring of urban climate (WMO, 2023). Low-cost sensing approaches, underpinned by the Internet of Things (IoT), have paved the way for a transformation in the resolution of which the urban climate can be monitored, both spatially and temporally (Chapman et al., 2015). Previously, the use of traditional weather stations represented a compromise in spatial monitoring, which being too expensive to densely deploy, often resulted in a maximum of one or two weather stations per city. This meant that increased spatial resolution could only be provided by mobile transects or satellite imagery, albeit at the expense of temporal resolution (WMO, 2023).

* Corresponding author.
E-mail address: l.chapman@bham.ac.uk (L. Chapman).

https://doi.org/10.1016/j.uclim.2023.101542
Received 7 December 2022; Received in revised form 14 March 2023; Accepted 17 April 2023
Available online 21 April 2023

The step-change in monitoring afforded by UMN's has been embraced in cities across the globe. Muller et al. (2013) provided an early review into the proliferation of the approach, showcasing a range of demonstrator UMN's operational at that time. Since then, the IoT has become increasingly mainstream, enabling further growth of UMN's as urban areas embraced broader smart city initiatives (Dunjic et al., 2022). For many cities, the inclusion of hyperlocal weather monitoring, along with its applications (e.g. links to public transport, energy demand, public health, etc) appeared a logical compliment to other forms of monitoring being introduced as standard (Erbel and Brune, 2022). Capital funding for such schemes has been abundant in many countries across the world, allowing researchers to now acquire urban climate data at an unprecedented resolution (see DoE, 2022 as a recent example).

However, this ongoing capital investment rarely heeds the warnings from the early pioneers of the UMN approach with many deployments eventually suffering the same fate of being 'mothballed' or decommissioned at the end of the original tranche of funding – it has now frequently been shown that a UMN is unfortunately a luxury that many cities can't afford to maintain (Chapman et al., 2017). One issue stems from the 'low-cost' sensors used in the networks where 'low-cost' is effectively a misnomer. It is true that the sensor itself is 'low-cost', but its upkeep in terms of maintenance, technical support and data management are not. This brings into question the wider sustainability, or durability, of the UMN approach. It may be that a well-designed measurement campaign, featuring a UMN, only needs to span the period of original funding. However, as climatology is a science that works to a 30 year baseline, it is worthwhile to consider approaches which have the potential to achieve a longitudinal record of that stature. Failure to consider this will lead to more limited smart-city initiatives or 'demonstrators', often funded by public money, achieving only an incremental advance in scientific research. Fundamentally, viewing this from a responsible innovation perspective (see Stilgoe et al., 2013), this is far from satisfactory. One approach proposed by Chapman et al. (2017) was to limit funding of new UMN's in favour of investing in existing infrastructure which could then act as a testbed to explore new, more durable, forms of monitoring which could then be rolled out more broadly. If a more sustainable approach to monitoring could be found, then this would open-up the benefits of UMN's to cities across the world, whilst freeing up research funding to be used for less incremental means.

One such approach, which has found extensive favour in the literature, is the concept of crowdsourcing urban weather data (sometimes known as opportunistic sensing). This approach harnesses existing sensors and instrumentation in cities in a near human-out-the-loop citizen science approach, thus removing the need for long term investment into maintenance and personnel. There exists a range of potential sources of crowdsourced weather data (see Muller et al., 2015 for an early review), but the one area of focus in the scientific literature are citizen weather stations. Here, commercial organisations and national weather organisations have sought to collate weather data routinely collected by weather enthusiasts into platforms for further use (e.g. UK Met Office Weather Observation Website: WOW, Kirk et al., 2021). Such platforms bring together a global community of amateur weather observers to voluntarily share their data, with the need to 'opt in' ensuring that higher quality data is more likely.

Some citizen weather stations feature their own data platform, and one in particular (the Netatmo) has resulted in an abundance of interest in the scientific literature. Netatmo provides a fascinating use case. It is commercially designed, built and distributed, making it widely available and represents a cost-efficient means for citizens to collect weather data (Meier et al., 2017). The device is relatively basic, consisting of an indoor and (unshielded) outdoor module to measure air temperature and relative humidity. Further outdoor modules are also available to measure precipitation and wind, based on a tipping bucket rain gauge and ultrasonic anemometer respectively. Modules are easy to install, harnessing the citizens Wi-Fi network for data transfer to the cloud. Although the device does not present anything exceptionally innovative in terms of the way observations are made, it has been the scale of uptake that has proved so potentially transformational. Originally launched in 2011, the device made an immediate impact on science (e.g. Stranda et al., 2011). Netatmo then saw prolific growth with stations installed in cities across the world and by 2015, conference papers had started to emerge documenting the potential of the device for urban climate monitoring (e.g. Meier et al., 2015). This rapid rise coincided with a period of intense proliferation of UMN's, but the number of Netatmo devices appearing in cities around the world was already starting to dwarf the efforts of scientists deploying their own meteorological monitoring equipment (Chapman, 2018). This subsequently led to an extensive body of research into the potential of the device for monitoring the urban climate, with studies covering the individual parameters measured by the device, including air temperature (e.g. Chapman et al., 2017; Meier et al., 2017), precipitation (e.g. de Vos et al., 2017), wind (e.g. Droste et al., 2020), as well as their broader collective utility (e.g. de Vos et al., 2020). However, by far the most common application is the use of the device to investigate the canopy (i.e. 2 m) Urban Heat Island. Here, Netatmo has provided unprecedented spatial insights into the phenomenon (Chapman et al., 2017). Previously, only satellite imagery has come close to this being achieved, but with its inherent limitation that it only provides estimates of canopy temperatures from the observed land surface temperature (e.g. Feng et al., 2019).

Despite the enormous potential of crowdsourcing approaches, the atmospheric science community has been wary of adopting the techniques (Chapman et al., 2017). This hesitance is mostly underpinned by concerns about data quality, but as numerous studies have now shown, quality assurance is possible resulting in meaningful data for scientific studies (e.g. Napoly et al., 2018; de Vos et al., 2019; Coney et al., 2022). As this new-found confidence increases the useability of crowdsourced data, attention now needs to switch to whether the approach can yield the longer-term records needed to produce the climate baselines that UMN's have so far struggled to achieve. To this end, this paper focuses on exploring the durability of an augmented network of Netatmo devices to collect weather and climate data. It aims to determine whether delegating the management of devices to schools and citizen scientists can result in longer term measurement records than that can be achieved through non-recurrent grant funding. In doing so it raises fundamental questions relating to the future role of transient sensing approaches in the discipline.

2. Methods

This research was conducted within the broader research infrastructure of the Birmingham Urban Observatory located in

Birmingham, UK. Urban Observatories represents a significant extension of smart city initiatives, grounded in governance by supporting evidence-based decision-making (UN Habitat, 2021). Analytics are at the heart of this missive, with the aim of developing longitudinal datasets measuring a vast range of variables across the city, which can be combined and processed to provide new insights into the functioning of the urban environment. Urban Observatories often include a living lab element, relying on citizens, businesses and public sector organisations to come together for collaboration, co-creation and ideation (Hossain et al., 2019). As an example, this can be of the form of hosting sensors and sharing data for the greater good. Living labs are often citizen-centric (Bergvall-Kareborn and Stahlbrost, 2009) and hence, the idea of using citizen weather data (e.g. Bell et al., 2015) lends itself well to the principle. The approach is often further embedded into UMN deployments with schools seen as ideal hosts for weather monitoring equipment (e.g. Davies et al., 2011; Hung and Wo, 2012; Chapman et al., 2015).

The origins of the Birmingham Urban Observatory are based on the UMN deployment of the Birmingham Urban Climate Laboratory (BUCL: see Chapman et al., 2015 for more details). As a demonstrator project, BUCL suffered the same fate of many UMN's, being mothballed after the original research funding ran out leaving a useable dataset spanning just 2 years (see Warren et al., 2016). However, as part of the development of the Birmingham Urban Observatory, BUCL was revitalised through new investment. This enabled a network of 24 automatic weather stations to be reinstated across the city, joined by a range of other new sensor networks (e.g. noise, road surface temperature, air quality etc). It also recreated the opportunity to use the observatory as a meteorological testbed and to further explore the role of crowdsourced data, i.e. Netatmo.

In the time between the deployment of BUCL and the Birmingham Urban Observatory, a number of citizens across Birmingham had procured Netatmo devices and were allowing their data to be shared in real-time in the cloud. These, along with more traditional citizen weather stations available on platforms like WOW, meant that even in the absence of a facility such as BUCL, crowdsourced data could be readily quality controlled (e.g. by Napoly et al., 2018) to effectively provide a ready-made, effectively free, UMN for the city. Indeed, given the scale of the global Netatmo community, this can now realistically be achieved for many cities worldwide. However,

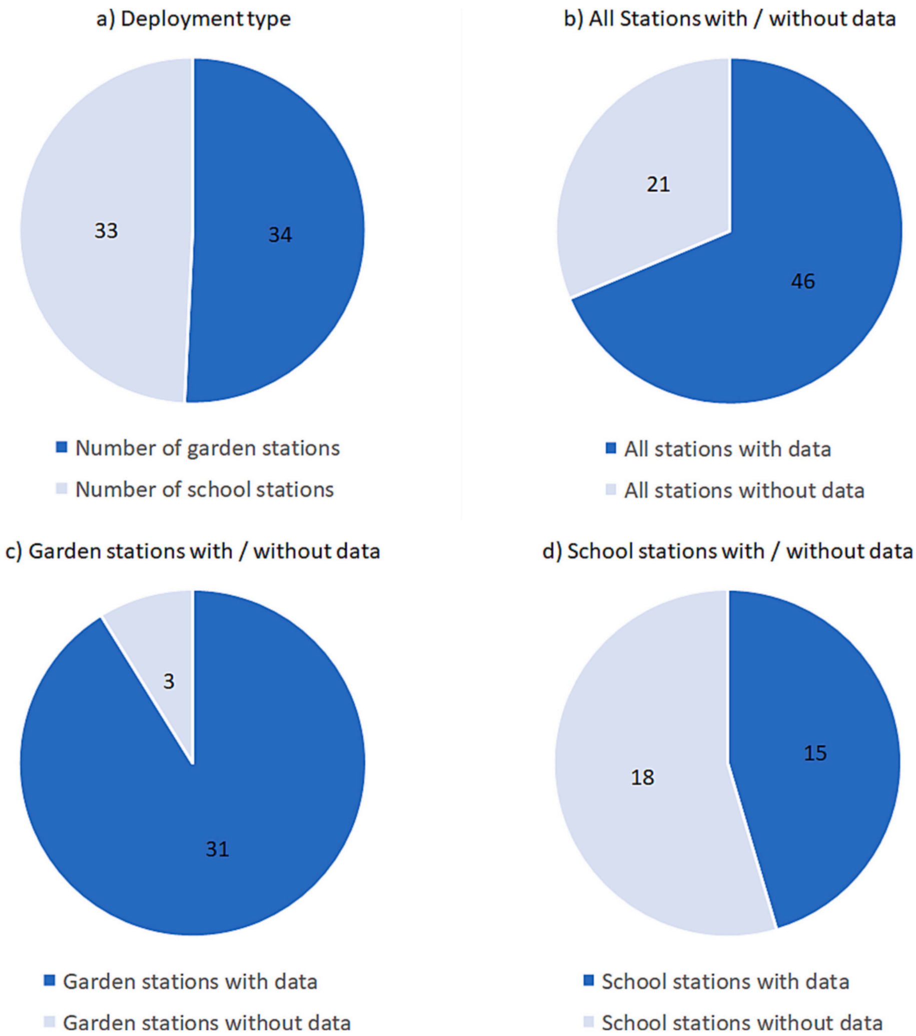

Fig. 1. Availability of stations and data across the two gatekeeper networks at the census point.

the added presence of a meteorological testbed, such as the Birmingham Urban Observatory, means high levels of quality control can also be potentially applied (e.g. buddy matching: Bell et al., 2015) using the existing dense network of weather stations in the observatory.

The research design of the Birmingham Urban Observatory allowed for the procurement and deployment of additional Netatmo weather stations to augment the existing network of citizen weather stations, enabling both increased network densification and a further element of quality control. Until recently, a major failing of the Netatmo weather station was that the air temperature observations were unshielded (and remain so unless the Netatmo radiation shield is purchased separately), representing a major source of data quality concerns. This was overcome in this study by using a radiation shield developed and tested by Young et al. (2014) with the procured Netatmo weather stations. The procured devices were then installed at sites across the city by the scientific team from the University, ensuring a further element of control in terms of the exposure of the weather station. Referred to hereafter as gatekeepers, these higher quality observation sites were deployed using two strategies. Firstly, as is common for many UMNs (see Muller et al., 2013), schools were contacted across the city to enquire whether they were willing to host a sensor. Secondly, an alternative approach was trialled where academic staff at the University were given the option to adopt a sensor to locate in their garden. The end-result was two additional networks of Netatmo weather data which can be compared against the existing network of Netatmo's hosted by the citizens of Birmingham.

Given the existing body of research into the quality control of Netatmo devices, an assessment of data quality is not the primary aim

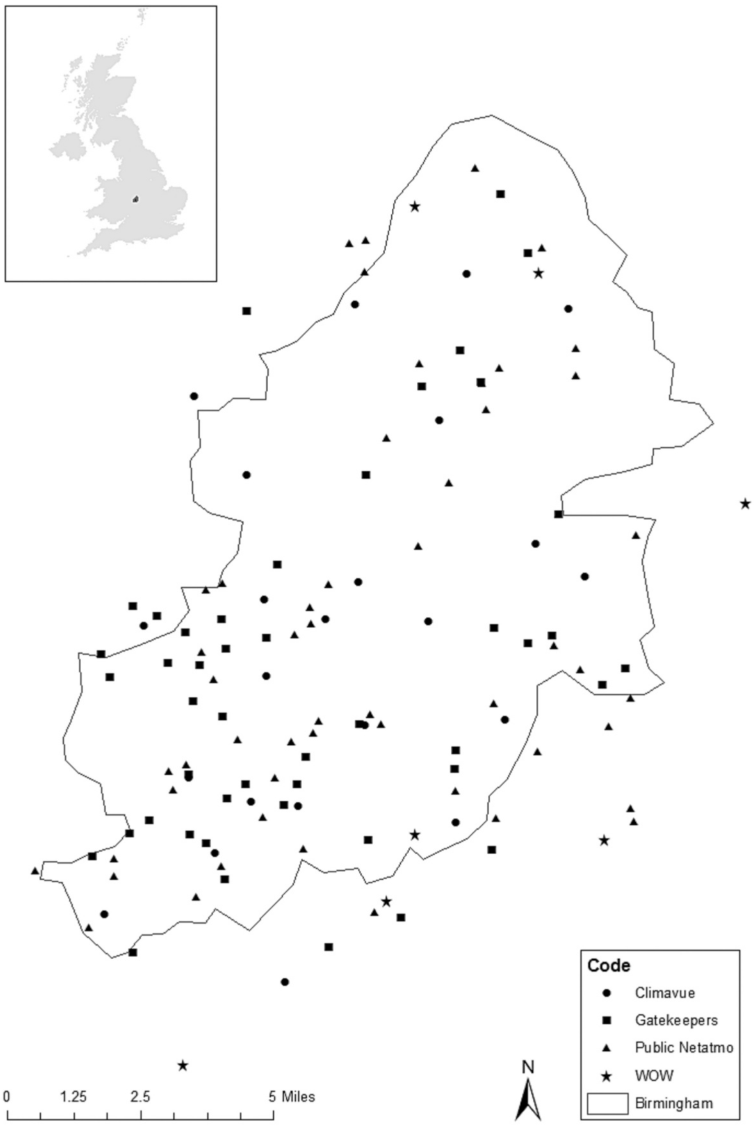

Fig. 2. The Birmingham Urban Observatory – a map showing the location of Climavue weather stations, WOW sites, gatekeeper sites, and public Netatmo. Inset map shows location of Birmingham within the UK.

of this paper (however, an example is provided for context). Instead, network sustainability is the key focus. It primarily seeks to explore whether an augmented network of citizen science sensors is more durable than a network of weather stations managed via short term research funding. The benchmark for comparison was the timeframe of the funding which allowed for the automatic weather stations of the Birmingham Urban Observatory to operate (Jun 2020 – Mar 2022). Crucially, other than initial installation, no further support (unless specifically requested) was offered to gatekeeper hosts. If a sensor went down, it stayed down until the host realised and intervened. This was done to closely mimic the cliff-edge of technical support after the money has run out on a UMN reliant on research funding.

3. Results

3.1. Station availability

A total of 33 schools expressed an interest in hosting a gatekeeper and were supplied with an installed unit. A further 34 additional gatekeepers were adopted by staff members at the university (Fig. 1a), creating two networks of school and garden sites. A census date of operational stations was held in July 2020 (chosen to remove the worse impact of the initial Covid-19 lockdown). For a station to contribute to this study, the Netatmo must have relayed data at some point during this month. Unfortunately, almost a third of deployed gatekeeper devices failed to send data to the Birmingham Urban Observatory cloud server in July 2020 (Fig. 1b). In all cases, this actually represented a permanent failure of the station with no data being recorded thereafter. There may be numerous causes of this failure, ranging from vandalism, unplugging of the indoor unit or changes to WiFi connectivity. Notably, the failure rate was significantly higher on the school network than on the garden network (Fig. 1c and d). The initial failure rate was high which can probably be attributed to Covid-19 lockdown (i.e. school closures) and justifies the choice of census date.

This led to a final total of 46 operational gatekeeper stations at the census point. This still provides a good sample size for comparison with the 53 'public Netatmo' units active on the census within the boundary of the Birmingham Urban Observatory and operated independently by citizens (note: a further 15 public Netatmo stations became operational post census, but these are excluded from this analysis). The distribution of available Netatmo weather stations on the census date are shown in Fig. 2. For completeness, the 24 full weather stations deployed as part of the Birmingham Urban Observatory are also displayed (hereafter referred to as Climavue stations) along with the location of the seven weather stations available from WOW at the census date. The latter includes six citizen weather stations and one official Met Office monitoring site.

3.2. Example application: the hottest day of 2020

The hottest day of 2020 in the UK occurred at the end of the census period (31st July) and provided an opportunity to demonstrate the utility of quality controlled crowdsourced air temperature data whilst all networks were operating at (near) full capacity. At the time, this was the UK's 3rd hottest day on record with temperatures peaking at 37.8 °C at Heathrow Airport (UKMO, 2020). The maximum temperature recorded by the UK Met Office in Birmingham was 34.4 °C, with a slightly higher maximum of 34.8 °C recorded at nearby Birmingham Airport.

Fig. 3 shows boxplots of the range of values recorded across the Climavue network of weather stations, WOW sites and the Netatmo networks. Data was quality controlled before inclusion in this analysis using a methodology based on the work of Meier et al. (2017) and Napoly et al. (2018). Quality control began with a manual approach to investigate the time stamps of the maximum recorded temperatures at public Netatmo sites. Most sites reached a peak at around 14:30UTC and any significant deviation from this was flagged. Sites recording a maximum later than this time were likely to be units that were located indoors resulting in lower maximum temperatures. This can be confirmed by exploring the standard deviation of the longer-term record compared to other sites (Meier et al., 2017). In contrast, maximum temperatures occurring at earlier times of the day can also be indicative of poor exposure where the

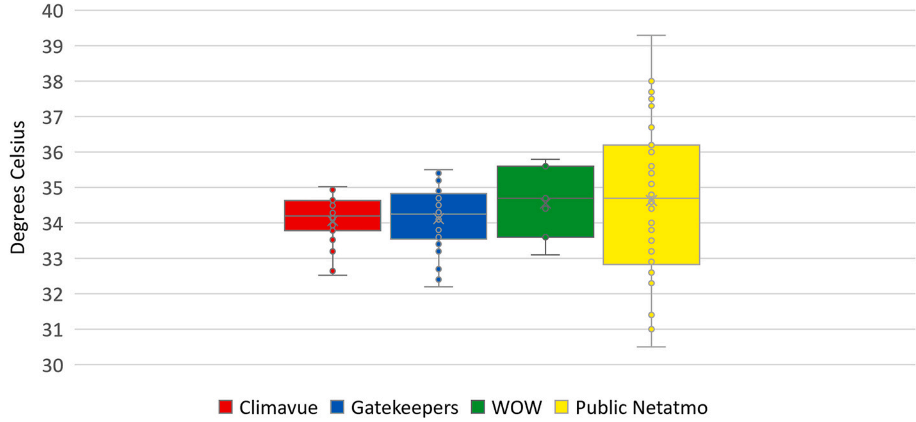

Fig. 3. Range of maximum temperature values recorded by the four measurement networks on the 31st July 2020.

unshielded unit is exposed to direct morning sun, resulting in higher AM maximum temperatures. This manual QC resulted in the removal of 6 public Netatmo units (4 indoor, 2 poorly exposed). The final stage of the QC was the use of z-scores as per Napoly et al. (2018). Any site not within $-2.32 < Z < 1.64$ was considered an outlier and removed from the analysis. This resulted in the removal of

Fig. 4. Distribution of cumulative lengths of air temperature / relative humidity records for a) all sites, b) garden sites and c) school sites.

a further two public Netatmo site and surprisingly, one of the gatekeeper sites.

The Climavue network of automatic weather stations have a small range of values with a median of 34.2 °C, close to the official maximum recorded for Birmingham by the UK Met Office. The gatekeeper stations also operate within a plausible range of values with a median of 34.3 °C. As expected, the citizen science networks of WOW and especially the public Netatmo network show the greatest range of readings. WOW values are broadly acceptable but public Netatmo report values beyond what is realistically plausible. Interestingly, this happens at both ends of the distribution. High temperatures can be readily explained by the lack of a radiation shield and poor exposure, but temperatures at the lower end are more difficult to explain. Although these are most probably indoor units, these do not occur in isolation, hence why they were not flagged by quality control. Indeed, this remains one of the challenges of crowdsourced data – i.e. whether readings in this range are actually true or not (Napoly et al., 2018). It is here where 'buddy checking' between standard and crowdsourced stations could potentially help, but a more robust solution is to not consider data from crowdsourced weather stations in isolation, but instead look at them as an ensemble. Overall, Fig. 3 shows the general decline in data quality through the move from managed networks to partially managed to unmanaged.

Finally, it is worth considering the presence of Netatmo units available for the case study. The case study occurred at the end of the census period when data presence should be at its peak. However, even at this stage of the experiment, 7 out of 46 gatekeepers and 11 out of 53 public Netatmo did not actually have data available for this particular date. This underpins the need to have large numbers of sites available for analysis / averaging. It also highlights the acute importance of the need to better understand network durability.

3.3. Network durability

The gatekeeper Netatmo supplied consist of three peripheral units. These each house different sensors pertaining to air temperature / relative humidity, precipitation and windspeed. As each unit is subject to different failure rates, the durability of each will be presented separately. Results are presented in terms of the number of months operational throughout the campaign. This is up to a maximum of 21 months (March 2022) when the initial funding for the Birmingham Urban Observatory ended and support of the Climavue network ceased. An operational month is defined as any month where an observation was relayed to the Birmingham Urban Observatory Server. Month by month data availability of each parameter is provided in the supplementary material.

3.3.1. Air temperature

Fig. 4a shows a clear bimodal distribution in the length of record provided for air temperature / relative humidity. Less than half of the sensors (15) managed to log for the majority of the measurement campaign (defined as $\geq$ 18 months), with only 3 sensors providing a full record (<7%) and 10 sensors still live and sending data in the final month of the campaign. Worryingly, seven sensors failed to provide observations beyond the first 3 months. Inspection of Fig. 4b and c indicate that the bimodal distribution can clearly be attributed to the network of sensors in schools falling short of the typical length of records provided in gardens. This follows the earlier trend highlighted at the census date which indicates that garden-based deployments were far more durable than school-based deployments (Fig. 1). This issue is likely exacerbated due to the regular closure of schools for vacations which mean power and communications could be turned off. Indeed, the longest available records are all obtained from the garden network where power and communications would be more consistently available.

Fig. 5 show the comparable statistics for the public Netatmo units. Here, 21 out of the 49 sites available at the census point (43%) managed to provide data for the majority of the measurement campaign (18 months>). Furthermore, although there were significant gaps in some of the records, 29 sites sent data in the final month of the campaign (59%). Overall, while there was still a natural attrition

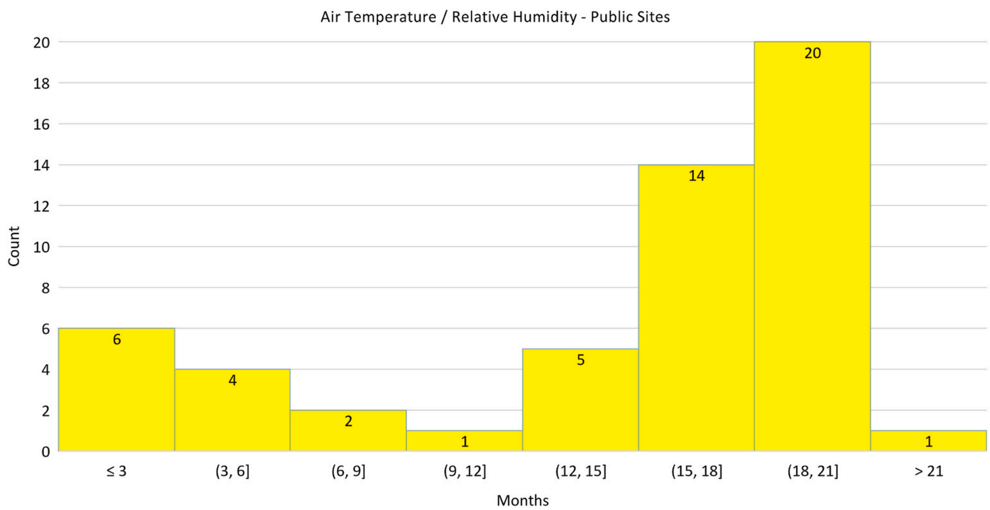

Fig. 5. Distribution of cumulative lengths of air temperature / relative humidity records Public Netatmo sites.

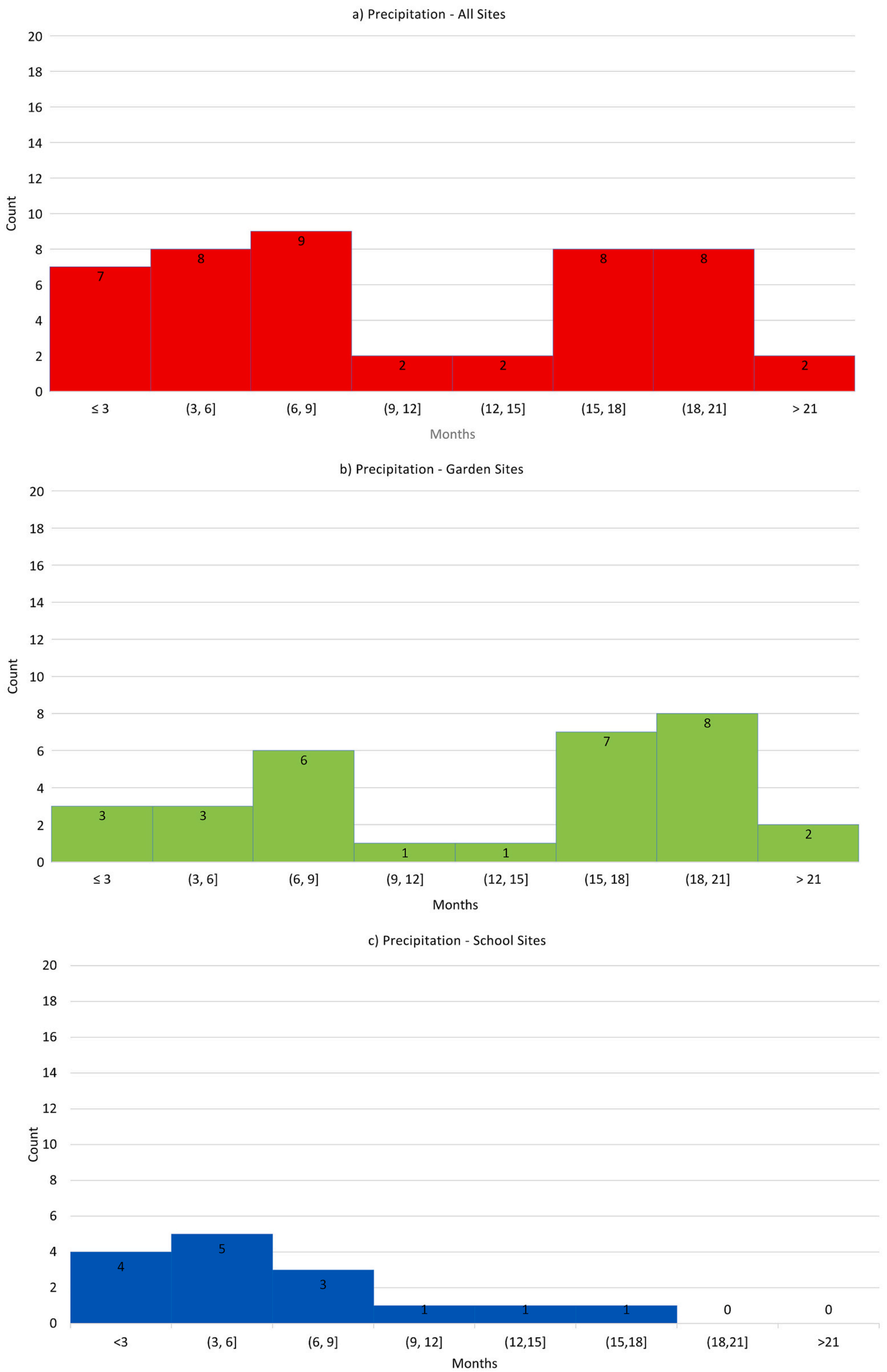

Fig. 6. Distribution of cumulative lengths of precipitation records for a) all sites, b) garden sites and c) school sites.

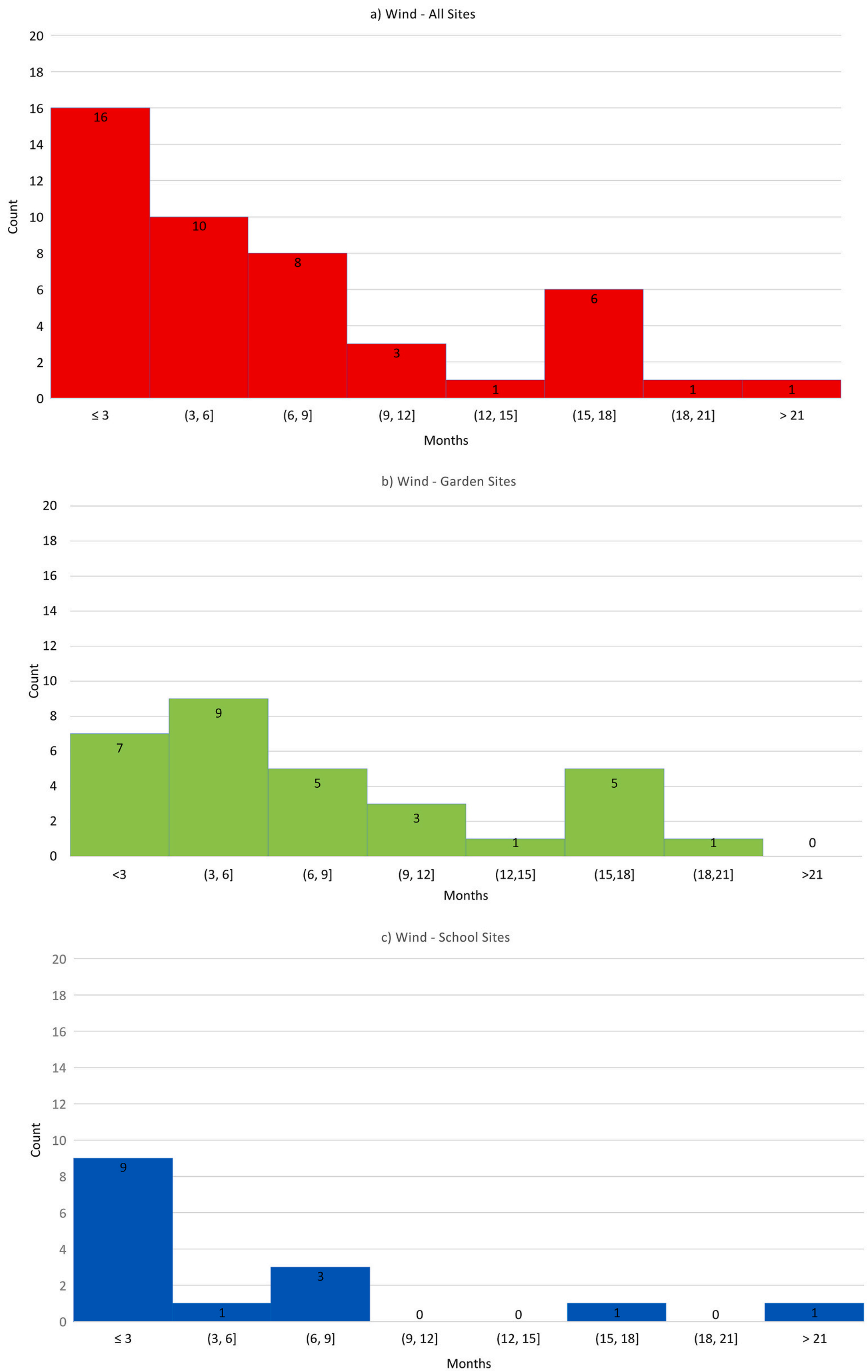

Fig. 7. Distribution of cumulative lengths of wind records for a) all sites, b) garden sites and c) school sites.

of units, this rate was far lower than that experienced with the gatekeeper devices and underpins the better motivation of citizens to maintain their stations.

3.3.2. Other parameters

The bimodal distribution also holds for precipitation records (Fig. 6a). The bimodal distribution in record length is particularly evident at the garden sites (Fig. 6b), where as failure of school sites shows a negative skew (Fig. 6c). The likely explanation for these patterns of failure is the need for battery replacement through the campaign. If signal strength between the gauge and the unit is weak, then the supplied batteries will need replacing sooner and explains the range of values on display. It is hypothesised that the garden sites fell into two categories – those who changed the batteries when required and those who didn't. It does also appear that battery maintenance at the school sites was largely absent. Overall, a total of 9 sites (<20%) were still sending data at the end of the campaign with only 4 sites provided near complete records.

In contrast to the other parameters, the wind data is far more skewed with only 4 sites left sending data at the end of the campaign and a single continuous record evident (Fig. 7). Again, the likely explanation for this failure rate is battery life. Another factor is the utility of the data. There are two possible explanations for this. Firstly, the batteries are much easier to change on the raingauge and this may have a big part to play. However, it is also hypothesised that wind data is perceived less useful / interesting to citizens than the other parameters and this will impact on the motivation to perform battery maintenance.

Due to the small number of precipitation gauges and anemometers at public Netatmo sites, this data was not routinely collected by the Birmingham Urban Observatory servers and is excluded from this analysis.

4. Discussion

In line with other studies, this work has shown that given adequate quality control, crowdsourced data can provide useful, indicative, urban meteorological data (Fig. 3), but it has also provided mixed results in the use of crowdsourced data for improving the durability of UMNs.

Firstly, the rate of attrition of devices supplied to schools and citizen scientists is somewhat alarming with nearly a third lost even before the census date. It is accepted that the Covid-19 pandemic will have accelerated the early decline in numbers, but the decline thereafter in the number of devices is equally dramatic and the chances of obtaining a complete record over the course of a measurement campaign are slim. This was further evidenced by the number of non-reporting stations of the data quality control case study in this paper where seven sites had already been lost even though this was just a few weeks into the measurement campaign. Although continuous records were obtained in a handful of cases, no single device managed to provide continuous records of all three parameters of air temperature, precipitation and wind. Overall, it is evident that if a gatekeeper approach is to be used to improve the quality assurance of a network, then device monitoring and technical support is still a necessity – even if it is just to alert a host to change the battery.

Secondly, the public Netatmo devices significantly out-performed the durability of gatekeeper devices in this study. With over half (55%) of the units logged on the initial census date still sending data at the end of campaign, this provided the most consistent source of Netatmo data in this study. This is before considering the added advantage of this data source with new devices coming online as others are lost. This means that, despite the transcient nature of the approach, this network can provide a relatively steady stream of data for further analysis and with new sites appearing over time providing a natural refresh in the instrumentation meaning that issues such as sensor drift are less problematic. However, the issues with these unregulated public devices still remain. The exposure of the device is unknown and large numbers of the devices will need to be removed from further analysis following the application of a QC algorithm (e.g. Napoly et al., 2018). This trade-off in data quality is clearly seen in Fig. 3 and underlines the need to use averaging approaches rather than considering each site in isolation. However, this then undermines the hyperlocal application of the data which is fundamentally driving the desire for UMNs.

The nature of the gatekeeper deployment strategies used in this study is also worthy of discussion. Both deployments were targeted at what should be potentially engaged parties. The weather station had educational value for the schools involved whereas academics would be both interested and understand why it was important to maintain the station. Bearing in mind they had openly volunteered, there would also be interest in the data from themselves. These factors clearly have a role in the length of the record obtained. This study has shown that although schools are frequently chosen as sites to host scientific equipment, due to the mutual benefits already outlined, they are actually problematic. An engaged teacher is essential to a successful outcome as the educational value will rarely extend beyond a small period / module of study. Maintaining power and communications will be beyond the remit of the teacher and there will always be periods of time where access to the site is limited. The durability of the garden network outperformed the school network in all cases. However, as previously noted, independent public Netatmo deployments were far superior in terms of durability and illustrate the importance of having an engaged audience for successful citizen science. This is further emphasised by an inspection of the UK Met Office WOW data where all six citizen weather stations active in the Birmingham area on the census date persisted for the full monitoring period (with a further two coming online during the campaign). Interestingly, from inspection of the site metadata, one of these sites is a Netatmo unit. The survival rate of these citizen weather stations, diligently maintained by weather enthusaists, is in stark contrast to the gatekeeper approaches explored in this paper.

5. Conclusions

A sustainable UMN using crowdsourcing is a possibility but as this study has showed, there is considerable doubt about how durable the entire crowdsourcing approach is. Indeed, there is a chance that society may have already reached 'peak netatmo' and the number of sites available for studies of this nature using that particular device may naturally start to decline. If this is the case, then alternative data sources will need to be sought (connected vehicles as an example). This again highlights the transient nature of crowdsourcing, algorithms need to be able to take data, perform quality control and spatial averaging on a wide range of inputs.

With respect to quality control, confidence is improving in the data for scientific use, but there is still large potential to tackle this at source. For example, Netatmo data would instantly have more value if radiation shields were supplied as standard, greater instruction on siting and exposure was provided and metadata were routinely collected from purchasers. This would allow 'real world' gate-keepers to be identified. It is also logical to assume that the keepers who willingly supply this information, will also be the most likely to maintain their station and provide the longest, most durable, records. This is the already case for platforms such as WOW, which as highlighted in this study, provide durable data of an acceptable quality.

Despite these ongoing concerns on data quality and durability, crowdsourcing still arguably provides the best option of a genuine low-cost means of producing a sustainable UMN. However, this paper has clearly demonstrated that if a longer-term record is desired, then deployment and management strategies need very careful consideration. Realistically, only the most enthusiastic of citizen scientists will be engaged beyond the short-term. This greatly reduces the number of stations that can be expected to function in an unsupported UMN, but it does mean that the data that is obtained is likely to be of the best quality possible from the crowdsourcing approach.

Overall, the previous 'deploy and decommission' – 'boom and bust' of UMNs has already proved to be unsustainable and the folly of repeating this mistake in city after city does need to stop. However, as urban observatories overtake smart city initiatives and grow in popularity as a basis to provide the high-resolution data needed for digital twins, then the hunger to densify observations is only going to increase. The question is whether the use-cases for hyperlocal weather data in these digital ecosystems are economically viable to support managed networks, if not, then compromises will be needed. At the moment, that compromise is using transient data through techniques such as crowdsourcing.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.uclim.2023.101542.

CRediT authorship contribution statement

Lee Chapman: Funding acquisition, Supervision, Conceptualization, Investigation, Data curation, Writing – original draft. **Simon Bell:** Software, Data curation, Writing – review & editing. **Sophie Randall:** Resources, Data curation, Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Lee Chapman reports financial support was provided by University of Birmingham. Lee Chapman reports a relationship with University of Birmingham that includes: employment and funding grants.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Engineering & Physical Sciences Research Council [grant number EP/P016782/1].

References

Bell, S., Cornford, D., Bastin, L., 2015. How good are citizen weather stations? Addressing a biased opinion. Weather 70 (3), 75–84.
Bergvall-Kareborn, B., Stahlbrost, A., 2009. Living lab: an open and citizen-centric approach for innovation. Int. J. Innov. Region. Dev. 1 (4), 356–370.
Chapman, L., 2018. Crowdsourcing and opportunistic sensing: The technology Isn't going away. In: So how Are We Going to Make it Work? 10th International Conference on Urban Climate, August 5th–9th, New York, USA.
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. Int. J. Climatol. 37 (9), 3597–3605.
Chapman, L., Muller, C.L., Young, D.T., Warren, E.L., Grimmond, C.S.B., Cai, X.M., Ferranti, E.J., 2015. The Birmingham urban climate laboratory: an open meteorological test bed and challenges of the smart city. Bull. Am. Meteorol. Soc. 96 (9), 1545–1560.
Coney, J., Pickering, B., Dufton, D., Lukach, M., Brooks, B., Neely III, R.R., 2022. How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom. Meteorol. Appl. 29 (3), e2075.
Davies, L., Bell, J.N.B., Bone, J., Head, M., Hill, L., Howard, C., Hobbs, S.J., Jones, D.T., Power, S.A., Rose, N., Ryder, C., Seed, L., Stevens, G., Toumi, R., Voulvoulis, N., White, P.C.L., 2011. Open air laboratories (OPAL): a community-driven research programme. Environ. Pollut. 159 (8–9), 2203–2210.
DoE, 2022. https://www.energy.gov/science/articles/urban-integrated-field-laboratories-will-equitably-address-critical-scientific accessed 25/10/2022.
Droste, A.M., Heusinkveld, B.G., Fenner, D., Steeneveld, G.-J., 2020. (2020) assessing the potential and application of crowdsourced urban wind data. Q. J. R. Meteorol. Soc. 146, 2671–2688. https://doi.org/10.1002/qj.3811.
Dunjic, J., Savic, S., Milosevic, D., Secerov, I., 2022. Urban meteorological networks in Europe: A review. EMS2022, (EMS2022-267).
Erbel, R., Brune, P., 2022, August. Measure it yourself-why smart cities need custom weather data sources. In: 2022 9th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE, pp. 221–226.

Feng, J., Cai, X.-M., Chapman, L., 2019. Impact of atmospheric conditions and local climate zone on the relationship between nocturnal surface and urban canopy heat islands. Q. J. R. Meteorol. Soc. 145, 3284–3299.
Hossain, M., Leminen, S., Westerlund, M., 2019. A systematic review of living lab literature. J. Clean. Prod. 213, 976–988.
Hung, T.K., Wo, O.C., 2012. Development of a community weather information network (co-WIN) in Hong Kong. Weather 67 (2), 48–50.
Kirk, P.J., Clark, M.R., Creed, E., 2021. Weather observations website. Weather 76 (2), 47–49.
Meier, F., Fenner, D., Grassmann, T., Jänicke, B., Otto, M., Scherer, D., 2015, July. Challenges and benefits from crowd sourced 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, vol. 7.
Meier, F., Fenner, D., Grassmann, T., Otto, M., Scherer, D., 2017. Crowdsourcing air temperature from citizen weather stations for urban climate research. Urban Clim. 19, 170–191.
Muller, C.L., Chapman, L., Grimmond, C.S.B., Young, D.T., Cai, X., 2013. Sensors and the city: a review of urban meteorological networks. Int. J. Climatol. 33 (7), 1585–1600.
Muller, C.L., Chapman, L., Johnston, S, Kidd, C., Illingworth, S., Foody, G, Overeem, A., Leigh, R.R., 2015. Crowdsourcing for climate and atmospheric sciences: current status and future potential. Int. J. Climatol. 35 (11), 3185–3203.
Napoly, A., Grassmann, T., Meier, F., Fenner, D., 2018. Development and application of a statistically-based quality control for crowdsourced air temperature data. Front. Earth Sci. 6, 118.
Stilgoe, J., Owen, R., Macnaghten, P., 2013. Developing a framework for responsible innovation. Res. Policy 42 (9), 1568–1580.
Stranda, R.D., Stokkenesb, S., Kristensenb, L.M., Loga, T., 2011. Fire risk prediction using cloud-based weather data services. J. Ubiquit. Syst. Pervas. Netw. 16 (1), 37–47.
UKMO, 2020. https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/weather/learn-about/uk-past-events/interesting/2020/2020_05_july_temperature.pdf. Accessed 22/11/2022.
UN Habitat, 2021. https://unhabitat.org/urban-observatories-a-comparative-review. Accessed 04/11/2022.
de Vos, L., Leijnse, H., Overeem, A., Uijlenhoet, R., 2017. The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. Hydrol. Earth Syst. Sci. 21 (2), 765–777.
de Vos, L.W., Leijnse, H., Overeem, A., Uijlenhoet, R., 2019. Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring. Geophys. Res. Lett. 46 (15), 8820–8829.
de Vos, L.W., Droste, A.M., Zander, M.J., Overeem, A., Leijnse, H., Heusinkveld, B.G., Steeneveld, G.J., Uijlenhoet, R., 2020. Hydrometeorological monitoring using opportunistic sensing networks in the Amsterdam metropolitan area. Bull. Am. Meteorol. Soc. 101 (2), E167–E185.
Warren, E.L., Young, D.T., Chapman, L., Muller, C., Grimmond, C.S.B., Cai, X.M., 2016. The Birmingham urban climate laboratory—a high density, urban meteorological dataset, from 2012–2014. Sci. Data 3 (1), 1–8.
WMO, 2023. Guidance to Measuring, Modelling and Monitoring the Canopy Layer Urban Heat Island. https://library.wmo.int/index.php?lvl=notice_display&id=22236#.ZAc9gnbP2Uk accessed 07/03/2023.
Young, D.T., Chapman, L., Muller, C.L., Cai, X.M., Grimmond, C.S.B., 2014. A low-cost wireless temperature sensor: evaluation for use in environmental monitoring applications. J. Atmos. Ocean. Technol. 31 (4), 938–944.