https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Performance and longevity of compact all-in-one weather stations – the good, the bad and the ugly

Authors:
Christopher. Walker. Brown[1], Marin. Schadee[2], Marijn. de Haij[1] & Theo. Brandsma[1]
[1] Royal Netherlands Meteorological Institute, De Bilt, The Netherlands
[2] Meteorology and Air Quality, Wageningen University, Wageningen, The Netherlands

chriswalkerbrown@gmail.com/ brown@knmi.nl

marinschadee@live.nl

Marijn.de.haij@knmi.nl

Theo.brandsma@knmi.nl

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Abstract

We provide a long-term evaluation of compact, all-in-one automatic weather stations (AiOWS) compared to professional-grade Automatic Weather Stations (AWS). We examine the performance, longevity, and degradation of six AiO WS models over several years of non serviced use. The objective was to determine how closely these low-cost stations meet World Meteorological Organization (WMO) performance standards for temperature, humidity, wind, and precipitation, and to identify their weaknesses and maintenance needs.

Previous studies show the potential value of AiOWS when data are properly quality-controlled, yet long-term reliability remains uncertain. To address this we deployed six AiO WS units—Davis VVue, Davis VP2, METER ATMOS41, Lufft WS601, and Vaisala WXT520, alongside two collocated reference AWS meeting WMO standards. Before field installation, each unit was tested in (KNMI's) calibration lab for baseline validation. The stations were then operated in open terrain for multiple years without any servicing, simulating typical end-user neglect.

Initially, all AiO WS met manufacturer specifications. After long-term exposure, however, sensors displayed varied durability. The Vaisala unit operated continuously for over 13 years, while others failed between four and seven years due to corrosion, component wear, and sensor drift. The METER and Davis VVue remained mostly functional but with degraded performance, whereas both Davis VP2 rain gauges failed early due to reed switch damage.

Temperature measurements were the most robust. In climate chamber tests, new and aged sensors maintained accuracy within ±0.3 °C across -15 °C to 30 °C, drifting slightly (underestimating by 0.5–0.7 °C) above 30 °C. Field data confirmed these results, though strong solar radiation caused overestimations during summer. The Vaisala and Davis VVue units remained within WMO Class B limits after a decade. Relative humidity showed consistent deterioration. Most sensors overestimated low humidity and underestimated above 90%, particularly the METER unit, whose bias grew markedly after five years. Wind speed accuracy degraded due to mechanical wear. Cup anemometers underreported low winds and failed completely in some cases. Sonic sensors (Vaisala, METER) produced erratic readings after several years, highlighting their fragility outdoors. Precipitation performance was weakest across all models. Tipping bucket designs suffered from clogging, internal corrosion, and undercatch errors, while haptic or drip-based sensors became inaccurate as components aged or fouled.

We concluded that compact AiO WS can provide scientifically useful temperature data if properly managed but fall short for humidity, wind, and particularly precipitation unless regularly serviced. Long-term unattended operation severely limits reliability, yet moderate maintenance can potentially restore performance close to WMO Class A/B standards, extending their utility for dense observation networks.

1. Introduction

Over the past decades, compact all-in-one Automatic Weather Stations (AiO WS) have become widespread, offering meteorological measurements at low costs compared to conventional Automatic Weather Stations (AWS) operated by National Meteorological Services (NMS).

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Indeed, NMS, such as the Royal Netherlands Meteorological Institute (KNMI) are using this data in tiered networks, where it is combined with AWS data to improve numerous products ranging from rainfall to temperature now-casting. Their affordability and ease of deployment have enabled a rapid expansion of observational networks, particularly in data-sparse regions and citizen science initiatives. Many peer-reviewed studies have demonstrated the scientific value of AiO WS in varied research applications. These low-cost, rapidly deployable instruments have been used for urban climate monitoring (Meier et al., 2017), air quality and heat stress exposure assessments (Chapman et al., 2020), and enhancements of meteorological networks beyond a backbone of AWS sites (Overeem et al., 2013). De Vos et al. (2017) showed that crowdsourced data from thousands of AiO WSs across the Netherlands could be statistically corrected to produce high-resolution rainfall maps comparable to those from official networks. Similarly, Meier et al. (2017) validated temperature data from citizen-operated stations in Zurich, finding them suitable for detecting spatial variations in urban heat islands. These studies highlight that, when properly calibrated and quality-controlled, data from AiO WS can meaningfully contribute to scientific research, particularly in applications where high spatial density is prioritised over individual sensor quality.

The World Meteorological Organization (WMO) provides guidance for the deployment and operation of AiO WS through its Guide to Instruments and Methods of Observation (WMO-No. 8). While AiO WS systems are not typically formally certified to meet WMO performance standards, their measurement principles, and operational concepts often align closely with those applied to Automatic Weather Stations (AWS) in professional networks. These include instrument characteristics such as measurement accuracy, long-term reliability, structural durability and ease of maintenance. The WMO also emphasises the importance of routine inspection and preventive maintenance, particularly for sensors exposed to outdoor conditions. Proper siting (away from buildings, trees and shading) is important, with problematic AiO WS placement locations requiring to be noted in metadata. Furthermore, the WMO promotes quality control procedures, metadata management, and intercomparison testing to ensure interoperability and scientific utility. For AiO WS systems to contribute meaningfully to research or operational monitoring, alignment with an agreed global set of standards is (in our opinion) the best option to encourage greater use of AiO observations amongst the science community.

Although well-managed networks of AiO WS are useful for research and can adhere to WMO guidance on siting, operation and servicing, many networks or individual AiO WS are poorly maintained/ sited and hence underperforming with respect to manufacturers specifications. As a result, this has prompted numerous studies evaluating their accuracy and reliability. Vučković and Schmidt (2023) analysed city-wide data and found a “substantial amount of erroneous occurrences” (gaps and faulty signals), especially in humidity readings. This shows broader concerns with the equipment when compared to official Automatic Weather Stations (AWS); AiO WSs often have suboptimal siting, and exhibit unmonitored and uncorrected instrument drifts. Hahn et al. (2022) similarly note that some AiO WSs have temperature biases in excess of WMO guidelines often caused by solar heating or non-standard placement such as indoors or shielded locations. These issues have motivated the development of rigorous quality control (QC) methods (e.g., statistical bias corrections) to vet AiO WS observations (Nipen et al., 2020; Alerskans et al., 2022) prior to use. Nevertheless, confidence in raw data quality is lowered by these issues, underscoring the need to benchmark AiO WS against reference instruments under real-world conditions.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

The aim of this work is to quantify how closely compact AiO WS approach WMO performance guidelines, and to assess their sensor longevity and failure modes observed both in our experiment and also reported in the literature. This paper publishes data from an opportunistic AiO WS evaluation experiment conducted at the test field of KNMI in De Bilt, the Netherlands using the following AiO WS; Davis VVue (called TX7 and TX8 in this paper), Davis VP2 (instruments TX1 and TX2), METER ATMOS41, Lufft WS601, Vaisala WXT520. By directly benchmarking compact AiO WS against standard meteorological instruments, our aim was to understand the practical accuracy and longevity limits of a selection of commonly used AiO WS models. We report observations on sensor accuracy and long-term drift from the literature and from our long-term field experiment, focusing on temperature, humidity, wind, and precipitation measurements.

2. Data and Methods

This experiment was conducted at KNMI's test field in De Bilt, the Netherlands (Figure 1). We evaluated AiO WS models mounted in open terrain at heights between 1.2 and 1.6 m, alongside two collocated reference Automatic Weather Stations (WMO AWS code 06260 and 06261). The AWS 06261's temperature and humidity were measured at 1.5 m above ground level 43 m North of the AiO WSs and rainfall at 1.0 m above ground level 15 m North East of the AiO WSs. All three observations are fully calibrated (to WMO AWS standards) and continuously cross validated with the observations made at DeBilt AWS (06260) sited 200 m east of the test field by KNMI's technical observations team (Figure 1).

Each AiO WS was initially assessed in the KNMI calibration laboratory—using a wind tunnel, climate chamber, and rain simulator—before being deployed in the field with minimal maintenance for multiple years.

The study represents a 'worst-case' scenario typical of many end users: no instrument maintenance, cleaning, or mid-deployment recalibration was performed. Throughout the multi-year deployment, we collocated AiO WS data with high-quality AWS measurements to detect biases, sensor drift, and data gaps.

This approach follows prior field comparisons of AiO WS, such as Jenkins (2014) and Bell et al. (2015 and 2017) who co-located Davis and Fine Offset stations for one year, and Droździoł and Absalon (2023), whose multi-month comparisons tested amateur rain gauges against certified Hellmann gauges.

3. Results

3.1 Laboratory Comparison

Prior to deployment, each AiO WS underwent calibration laboratory testing to quantify baseline accuracy, operational range, and to verify consistency with manufacturer specifications. All AiO WS tested initially met manufacturer specifications (Table 1). After being deployed in the field for a minimum of 5 years, the AiO WS were removed and reassessed in the calibration laboratory. Table 1 summarises the manufacturer's specifications and figure 2 shows the observed measurement ranges of the aged units that remained functional.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Both Davis VVue units (TX7 and TX8), and the METER ATMOS41 remained functional at the end of deployment (10 years TX7 and TX8, 5 years for the METER). The Vaisala remained active for >13 years in the field, eventually failing in July 2024 and prompting the finalisation of this research.

In contrast, both Davis VP2 units (after 7 years, 4 months TX1, 6 years 8 months TX2) and the Lufft WS601 (after 4 years and 2 months) ceased transmitting data. The Davis VP2 both had raingauge failures very early on in deployment, and no reliable rainfall information was collected for the entire deployment. Both Davis VP2s (TX7 and TX8) solar panel covering plastic had become discolored, and upon removal of this plastic covering and insertion of a new battery, both became became partially functional again. The Lufft WS601 partially recovered functionality in the lab, transmitting only rain gauge data. Evidence of substantial internal corrosion was found in the Lufft WS601 (Figure 3, panel 4), possibly due to moisture intrusion. The Vaisala remained functional for temperature and humidity until the final month of operation on the test field. The unit failed, and could not be revived for subsequent the calibration laboratory. However, we note that the Vaisala had operated for more than 13 years without maintenance, showing the potential lifespan of AiOWS, as well as the risk of 'zombie' instruments (i.e. those that have been uncalibrated for many years), potentially increasing uncertainty in data quality for networks of AiOWS.

Both Davis VP2 units continued to transmit temperature, humidity, and wind speed data, but their tipping bucket rain gauges' magnetic reed switches failed (Figure 3 panel 3), a known fragile component in outdoor environments (Saraf and Ivan, 1978). Newer Davis VVue and updated VP2 models utilise a more robust Hall-effect sensors in their tipping bucket rain gauges, perhaps resulting in the longer life spans observed in TX7 and TX8 . Although the METER provided temperature, relative humidity, rainfall, and wind speed at deployment's end, only temperature and humidity measurements were considered reliable; wind speed and rainfall data were implausible.

3.1.1 Air Temperature

Temperature sensors were tested in a climate chamber using 5 °C incremental steps from -15 °C to +50 °C under 20% relative humidity, holding at each step for one hour for full equilibrium. All new AiO WS units adhered closely to specifications within the 0–30 °C band (table 1). The aged AiO WS showed best performance within -15 to 30 °C, but above 30 °C, temperature was typically underestimated (figure 2). All Davis AiOWS (TX1,2,7 & 8) had less than 0.3 °C errors between -15 to 30 °C, whilst above 30 °C all Davis and the METER AiOWS underestimated by 0.5–0.7 °C. Although laboratory tests for the aged Lufft AiO WS were unavailable, Fenner et al. (2018) reported good precision (±0.1 °C) for Lufft WS700UMB, consistent with the new unit performance tested at KNMI prior to deployment. Due to non-operation status, the Vaisala (and aforementioned Lufft) were not tested post-deployment for air temperature.

3.1.2 Relative Humidity (rH)

RH testing involved making 10% rH step increments between 17% and 97% rH at a constant 20 °C conducted in a climate chamber. At equilibrium at each step, rH was held for one hour to allow for full equilibration. All AiOWS demonstrated similar trends: measured rH started to drop relative to the reference above ~60-70% rH. The METER performed poorest at low rH but accuracy improved at higher rH values (figure 2).

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Davis TX 1,2,7 and 8 rH sensors transition from slight overestimation at low rH to underestimation above 70% aligns with Bell et al. (2017) in their two-year drift study.

Due to non operation status, again, the Lufft and Vaisala were not tested post-deployment.

3.1.3 Wind Speed

Assessment of wind speed accuracy was completed within a wind tunnel, where the windspeed was adjusted from 1 m/s to 19 m/s in 2 m/s increments. All Davis VVUE (TX7,8) and VP2 (TX1,2) cup underestimated wind speed, particularly below 3 m/s. One of the METER sonic transducer pairs malfunctioned, recording constant 2 m/s easterly wind at zero real wind speed, and was excluded from further tests (figure 2).

Davis VVue TX7 and TX8 improved accuracy at higher wind speeds (<10ms). The two Davis VP2 units performed differently compared to each other: TX2 worsened with speed, attributed to a rusted bearing evident in run-up test (Figure 3, panel 2), while TX1 slightly improved, similarly to the TX7 and TX8, indicating that aging of the cup and vane does not result in a consistent drift or offset that can be corrected at a later data processing stage. Lufft, METER and Vaisala AiOWS were not tested, due to aforementioned malfunction.

Low-wind bias agrees with Droste et al. (2020), who found AiOWS cup anemometers often fail to register light breezes. Davis VVue and VP2 devices slightly improve at mid-range speeds (5–12 m/s), whereas Vaisala WXT520 showed a steady –0.5 m/s offset. After eight years, no aged units provided reliable wind data, confirming mechanical wear and bearing corrosion degrade cup/vane accuracy, while sonic anemometers suffer from transducer failures.

3.1.4 Precipitation

Ten mm of simulated rainfall was applied three times to each of the functioning AiOWS. The two Davis VVue units showed excellent performance, with the VVue TX7 performing with only -1% error (TX7 measurement vs simulated rainfall amount), and the TX8 performed with a -5% error. Both Davis VP2 had poor performance in tests; although the reed-switch mechanism worked moderately well when assessed initially (–5% error) but both units failed almost immediately in the test field (figure 3, panel 3).

The Lufft AiO WS displayed good accuracy (3% error), but we did observe some sticking within the tipping bucket mechanism during the test; where water overflowed the bucket and hence was not counted as part of the measurement. Interior corrosion of the tipping bucket was evident upon disassembly (Figure 3, panel 4).

The METER AiO WS uses a drip counting mechanism; with the number of drips counted per second relating directly to liquid precipitation entering the funnel. Initial tests conform to WMO standards. However, upon removing the aged METER from the testfield, we observed that the funnel was clogged with dust and insect debris, resulting in a delay in rain being fed into the drip counting mechanism (figure 3, panel 1). We attempted to clear the debris, and lightly clean the contacts within the drip counting mechanism. However, this did not resolve the issue, and the unit continued to report inaccurate rainfall. Due to the design of the drip counting mechanism the mechanism is highly reliant on a calibrated volume of water within each drip to be formed. If the drip is too small or large a volume, the number of drips compared to the rainfall rate will move outside of calibration specification, and a under/overestimation will occur. We suggest that upon a correctly calibrated volume drip of water being attained, the METER would be

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

performant again, highlighting the need for regular recalibration/ replacement of the rain gauge components.

3.2 Field Performance

Personal weather stations showed systematic temperature biases versus the AWS reference (Figure 4). All AiOWS apart from the METER underpredicted at low temperatures, and over predict at high temperatures (with likely strong solar radiation, figure 4). We also noticed drift in some of the AiOWS, with Vaisala initially overestimating temperature by ~+0.3 °C in 2016, increasing to ~+0.45 °C by 2024. The Meter performed exceptionally well, with very slight underreading of temperatures across all temperature bands.

Mean Absolute Error (MAE) reflected similar trends. Vaisala and VVue errors increased over deployment; summer MAE exceeded winter MAE- again highlighting issues in the small radiation shield designs common in AiOWS. Lufft and VP2 MAE increased over time with higher summer errors; METER MAE remained steady with seasonal oscillations. Error magnitudes (~0.2–0.5 °C) align with the aforementioned previous studies.

Relative humidity aligned closely with climate chamber results- underrecording of rH >90% and overrecording of humidities <70% for all instruments. We also noted a worsening of the METER rH observations across all rH bands after 2023, indicating that the deterioration within the sensor may accelerate with age (figure 4.)

Precipitation monthly totals showed large undercatch (Figure 5). Vaisala consistently under-collected relative to AWS with no seasonal or drift pattern detectable. Lufft performed best; monthly totals within ±25 mm of AWS. Undercatch is consistent with known tipping bucket limitations where wind can cause 2–10% loss in liquid precipitation, and potentially higher with snow (Segovia-Cardozo et al., 2023) and poorly maintained gauges tend to under-read (De Vos et al., 2019). As mentioned above, both Davis VP2 models had very early failure of the reed switch rain gauge, and no useful data was collected during the deployment. Research into event detection probability (PoD) and false alarm rate (FAR) showed that none of the AiOWS gauges detected every rainfall event. PoD averaged well below 1; e.g., only one summer 2018 month saw Vaisala register 100% of AWS rainfall events. In that month, Lufft detected none. FAR was low (<0.1) for all except METER, whose spurious tips raised FAR. Our results underline rain gauge quality control's importance: even well-performing AiO WS can miss light/short events (low PoD) and produce false triggers (FAR) without careful calibration and sheltering.

Windspeed observations are not possible to precisely compare with data from the AWS, owing to the AWS anemometer being mounted at 10m height. Surface roughness is not consistent on the testfield (due to other instruments, fences and buildings within the surrounding 500 m), and hence an algorithmic method of calculating true 1.5m wind velocity from the 10m observations will introduce uncertainty. We note in our comparison with the 10m AWS wind data that the sonic anemometer equipped METER had the values closest to the AWS data, and the cup and vane equipped Davis VVUE had the most different (figure 5).

4. Discussion

4.1 Sensor drift and degradation mechanisms due to environmental exposure

The long unattended deployment across multiple AiOWS revealed sensor-specific degradation pathways that explain the differences seen in Figures 2–5. Degradation for all variables (aside

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

from rH) is seemingly governed less by instrument age than by instrument design coupled with cumulative environmental stressors (such as solar loading, humidity cycling, particulates, biofouling and corrosion from wind-driven rain/ condensation.) This is demonstrated by the early damage to the Davis VP2 rain gauge system resulting in no data being collected, whilst the Vaisala system worked for more than 13 years, whilst operating in exactly the same environmental conditions. Relative humidity did however show a potential drift that worsened with age of the sensor.

Temperature displayed the strongest long-term stability. Laboratory tests showed that aged units remained close to specifications in the 0–30 °C band, with a consistent underestimation of ≈0.6 °C above 30 °C relative to the reference. In the field, monthly biases were typically within ≈±0.2–0.5 °C. The Vaisala trended from ≈+0.3 °C toward ≈+0.45 °C by 2024, while the two Davis VVue sensors exhibited seasonality—underestimation in winter and overestimation in summer (Figure 4). The patterns are indicative of radiation-shield issues in AiOWS under high insolation conditions rather than thermistor drift. This is demonstrated by the high temperature bias (> 30 °C) on the test field being in the opposite direction to the bias observed in the climate chamber. The Netherlands is too cold to have >30 C air temperatures without solar radiation (i.e. at night time). However, the climate chamber does not heat using insolation, and thus is representative of nighttime temperature and humidity conditions. Therefore, a laboratory calibration of the AiOWS temperature sensors under conditions of high temperature and zero radiation (and correction of the temperature measured) would actually result in a larger positive temperature bias being recorded by the instruments at peak summertime solar radiation for all instruments apart from the METER.

Data from the field tests confirmed the laboratory observations on relative humidity (Figure 4). Here, all AiOWS overreported rH at low rH conditions, and underreported rH at high rH conditions (>90% rH). Vaisala's rH bias started near –6%, trending towards zero over ~10 years. METER's rH bias grew rapidly from –1% to +8% within 4–5 years. VVue sensors followed similar patterns: bias rose until 2020, then returned towards zero by 2021. Lufft's rH bias increased from ~0% to +3% in the first three years. All except Lufft showed largest underestimation at 90–100% RH and overestimation below 60%. rH MAE trends mirrored the biases. METER's MAE rose steeply (5–7%), Vaisala's declined towards ~2%, VVue and Lufft remained ~3–4%. These confirm that long-term sensor calibration drift and bias trends observed in other work (e.g. Bell et al., 2017; Ingleby et al., 2012). As discussed in the literature, these behaviors match also failure mechanisms of polymer-film capacitive sensors such as moisture ingress, contamination, UV/thermal aging, producing low-rH positive bias and high-rH saturation; as seen in both laboratory tests and field time series (Bell et al., 2017; Ingleby et al., 2012).

Wind speed saw all anemometers under-reporting relative to the 10 m AWS (primarily due to the height of the AiOWS (1.5 m) being influenced by surface roughness of the ground.) For this reason, it is not logical to directly compare the wind data with the AWS data. Our binned AiOWS in figure 5 shows this, with substantial under recording of windspeeds across all sensors. This highlights the necessity of using wind tunnel testing of the AiOWS to understand degradation of windspeed observations, rather than imperfectly co-located observations.

Precipitation. Rain was the least reliable and durable variable. Substantial detritus was found in all AiO after field deployments (figure 6), comprising of insects, leaves and pine cones. Considering the open grassland location of the study, we would predict that AiOWS situated

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

closer to woodland or agricultural areas would suffer more from such fouling. Tipping buckets under-collected in the field and deteriorated mechanically; both VP2 units suffered reed-switch failure (no tips after mid-2018).

4.2 Performance relative to WMO standards and international guidance

The WMO Guide (WMO-No. 8) provides accuracy classes that serve as practical benchmarks for research and operations. The combined laboratory and field results indicate initial compliance with Class B for temperature and portions of wind, but rapid degradation toward Class C/D for humidity and rainfall.

Temperature. New units routinely achieved Class B (±0.5–0.6 °C) and sometimes approached Class A (±0.2 °C) in laboratory conditions. Field MAE typically remained ≈0.2–0.5 °C (Figure 2; Table 2). The systematic ≈0.6 °C underestimation above 30 °C across multiple platforms suggests overheating caused by the non ventilated compact radiation shields found on the tested AiO WS, mitigable by aspiration, improved siting, or post-hoc radiation corrections (Cornes et al., 2020; WMO-No. 8).

Relative humidity showed that initial performance was often within Class B (±5% rH) and occasionally near Class A (±2%) for sensors, but field drift drove many instruments into Class C/D within 3–5 years—consistent with polymer-film aging and contamination. This suggests that annual recalibration or sensor replacement is necessary for unattended deployments (Bell et al., 2017; Ingleby et al., 2012).

Wind speed showed that in laboratory conditions, anemometers typically met Class B thresholds (±10% or ±2 m s^{-1}), but aging of bearings reduced effective class to C for the Davis VP2 and VVUE AiOWS especially at low winds where friction from aged bearing and run-up thresholds dominate.

Precipitation showed that none of the AiO WS achieved reliable WMO Class B compliance. Wind-induced undercatch, intensity-dependent tipping bias, clogging, and component failure led to persistent negative biases. Using a reference of 10 minutes NI and Event detection metrics (PoD well below 1; low FAR except spurious METER tips) quantify reliability limits. Even where monthly aggregates appeared plausible (e.g., Lufft within ≈±25 mm), failures emerged with age and contamination. As deployed, precipitation sensing in AiO WS should be treated as qualitative unless supported by frequent inspection, leveling, shielding, and calibration checks (Segovia-Cardozo et al., 2023; Droździoł & Absalon, 2023).

4.3 Recommendations for AiO WS networks and users

Quality control and assurance should be designed around known degradation/ sensor weaknesses rather than nominal specifications from the manufacturer. We recommend to implement quarterly visual inspections, funnel cleaning, and re-leveling for gauges. At semiannual checks; annual cup/vane bearing need inspection and replacement as needed, and given the drift seen in figure 4, the capacitive rH sensors needs to be viewed as a consumable item, and renewed frequently.

There is potentially a need to improve radiation shielding/aspiration design in AiO WS, particularly those deployed in high-insolation sites. The failure of the METER and Vaisala sonic anemometers suggest that attempts by the manufacturer to strengthening these components against damage and environmental degradation would also be beneficial. For precipitation,

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

maintaining level AiO WS mounting and the use of bird guards (where feasible) to reduce undercatch and ingestion of debris into the funnel would likely improve observational quality.

Maintaining metadata on installations, maintenance, component swaps, and calibration results is vital to improve confidence in AiOWS data. There is a need to train AiOWS operators in how to perform functional tests (manual bucket tips, anemometer spin tests) and complete basic QC interpretation, particularly when the AiOWS are operated by non-experts/ citizen scientists.

4.4 Implications for applications and cost-benefit considerations

For temperature and mid-range wind measurements, AiO WS offer value, particularly for spatially dense nowcasting, urban climatology, and micro-meteorological applications where biases of a few tenths of a degree or ≈5–10% are acceptable or correctable. For humidity and precipitation, AiO WS generally require more frequent servicing to remain useful in operational systems (such as required by a national meteorological service). Operational plans should budget for recurring sensor replacements (rH), routine gauge cleaning/calibration (raingauge and cup and vane anemometers), and redundancy in the form of a higher quality raingauge system for rainfall where this parameter is of high importance.

4.5 Limitations and future work

The study mimicked worst-case user behavior (minimal maintenance) to test intrinsic durability limits. While this reveals fundamental constraints, networks with preventive maintenance and calibration schedules will likely outperform the results shown here. However, as this was an opportunistic study, we can see suggest improvements for future studies within this field. The most logical improvement is to concurrently test multiple examples of AiOWS, so we can assign greater confidences to issues and trends observed across multiple units, rather than a single/ two individual weather stations. We would also like future work to quantify improvements from scheduled maintenance regimes, assess newer solid-state rain sensors and low-cost sonics under extended deployments, and further investigate radiative heat related air temperature over-reading caused by high solar insolation heating the comparatively (vs a AWS) small radiation shields found on most AiO WS.

5. Conclusions

Six AiO WSs were assessed through laboratory calibration and multiple years of unattended deployment at KNMI's testfield site. The results show that, despite a prolonged period without maintenance, certain variables—particularly air temperature— remained within acceptable accuracy limits for scientific applications. Across both the laboratory and field settings below temperatures of 30 °C, all AiO WSs aligned with WMO Class B standards even after environmental exposure and zero maintenance.

In contrast, precipitation measurements proved to be generally unreliable. Tipping bucket mechanisms—particularly those in the Davis Vantage Vue and Lufft—performed best in both lab and field tests, but even these showed considerable degradation over time when compared with rain gauge data from an AWS. Haptic and drip-counting sensors (e.g., Vaisala WXT520 and METER Atmos41) were especially prone to underreporting, likely from sensor fouling and material degradation.

Relative humidity sensors showed the clearest evidence of long-term drift. The AiO WSs met Class B performance thresholds under high-humidity conditions in the lab, real-world data

© Author(s) 2025. CC BY 4.0 License.

showed increasing bias over time. The Vaisala remained within Class B tolerances in the field, but most others, particularly the METER Atmos41, degraded into Class C or D performance bands.

The long-term field deployment revealed operational challenges as well. Data availability was severely impacted by yellowing of solar panel covers, clogging of rain gauges with detritus, and failure of instrument components, such as reed switches. Of the six stations, only the Vaisala WXT520 delivered a near-complete dataset. However, we conclude that with moderate effort in cleaning the AiOWS, replacement of humidity sensors and wind tunnel calibration, the data collected from these devices will become increasingly valuable to end users.

6. Author contribution:

CB and MdH designed the experiment, with CB and MS completing the post deployment laboratory testing. CB prepared the manuscript, with figures 4,5 and 6 produced by MS. Review and scientific improvements were suggested by TB and MdH, along with extensive rewording.

7. Competition interests:
   The authors declare that they have no conflict of interest.
8. Acknowledgements:
   We would like to acknowledge the work of Jos Verbeek, of KNMIs' calibration laboratory, without whom, we would not have been able to assess the AiOWS performance pre and post deployment, and Sanne Vega who assisted with the initial calibration and field installation.
9. Financial support:
   This work was entirely funded by KNMI directly.
10. Figures

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

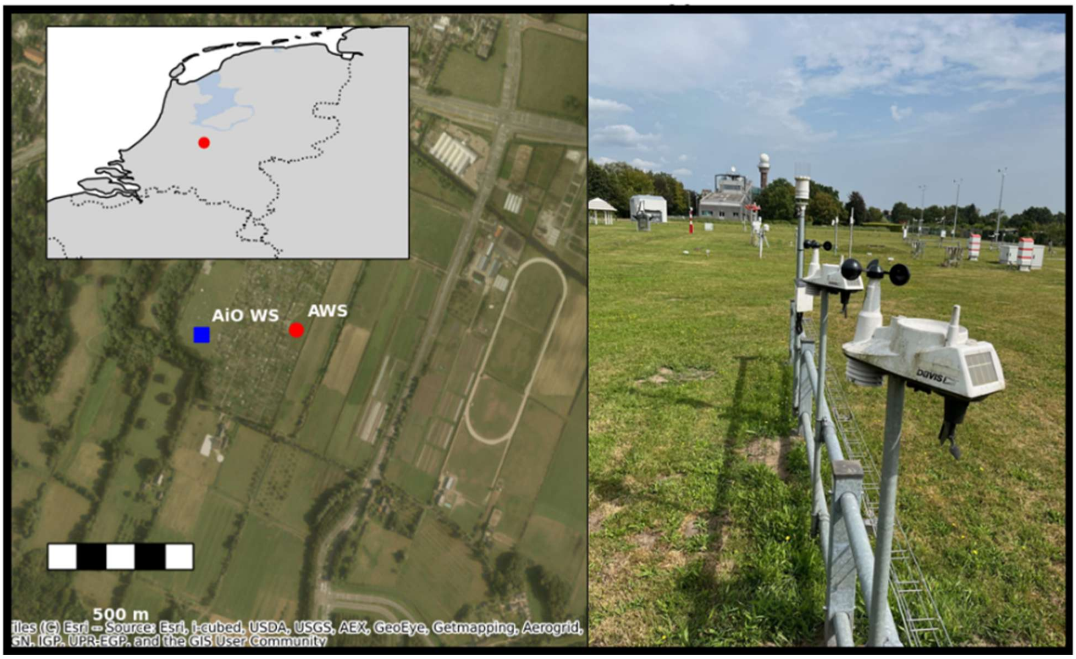

Figure 1.

Left- Location of the KNMI testfield (N 52.099, E 5.176) with the AiOWS and AWS 06261 site indicated with a blue square, and the AWS 06260 indicated with a red circle. Right, photograph of the aged AiOWS installed at the testfield

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

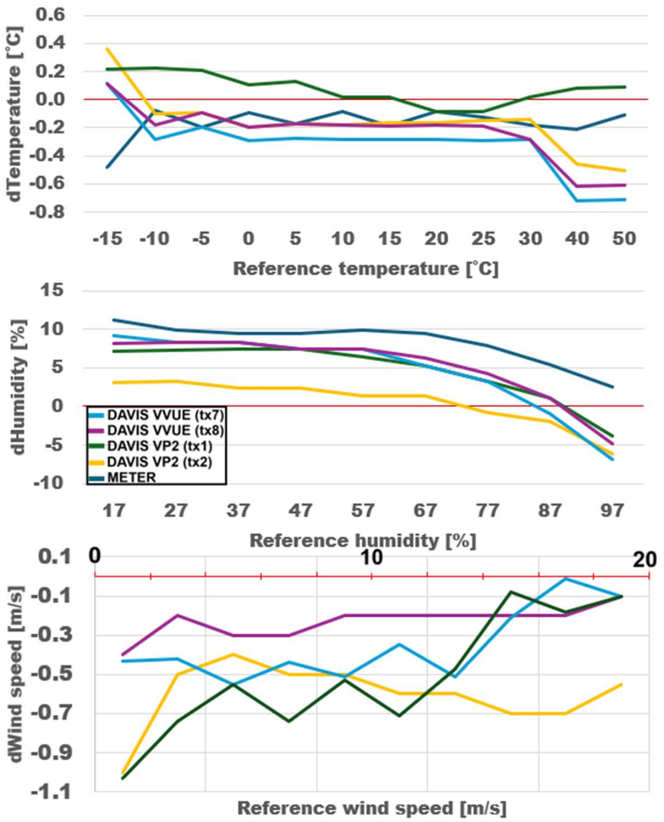

Figure 2.

Top: Temperature difference of each Personal Weather Station (PWS) from the Reference, at 20% relative humidity with no wind or irradiation. At each step, the 1 minute average measured temperature is taken from the PWS 1 minute before the climate chamber moved on to the next step.

Middle: Relative humidity difference of each Personal Weather Station (PWS) from the Reference at constant temperature (20°C) and no wind or irradiation. At each step, the 1 minute average measured relative humidity is taken from the PWS 1 minute before the climate chamber moved on to the next step.

Bottom: Wind speed difference of each Personal Weather Station from Reference in the wind tunnel. Measurement was taken after one to two minutes at stable reference speed and measured speed.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

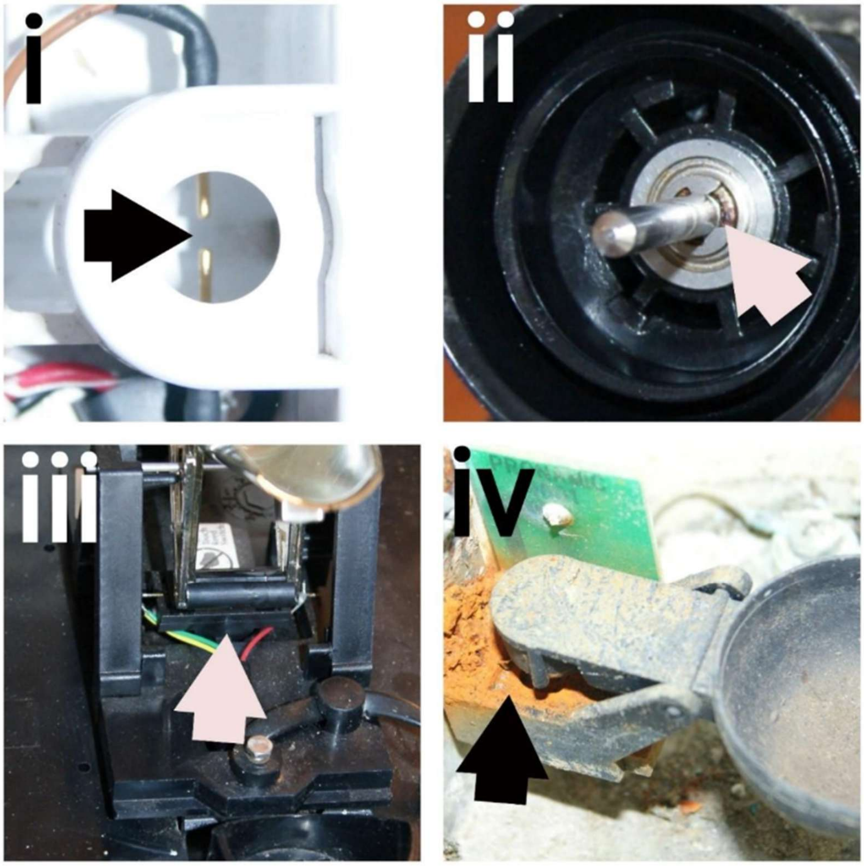

Figure 3.

i METER's drop counter system, using two gold plated electrodes

ii Corrosion of the sealed ball bearing supporting the cup anemometer on the Davis VP2 instrument

iii Failed reed switch on the Davis VP2

iv Corrosion on circuitboards and tipping bucket rain gauge in Lufft

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

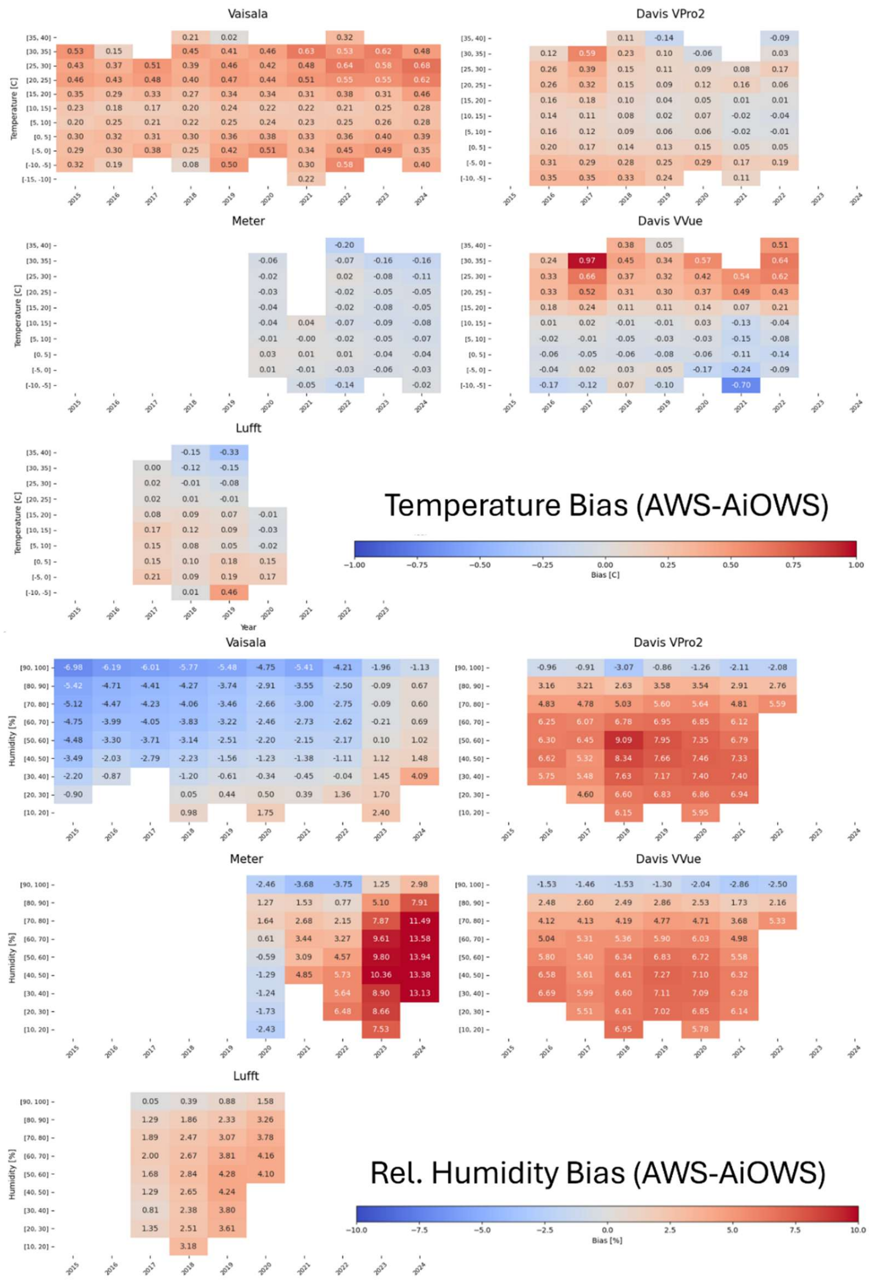

Figure 4 Temperature Bias and relative humidity bias from the AiO WS deployed at the testfield.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

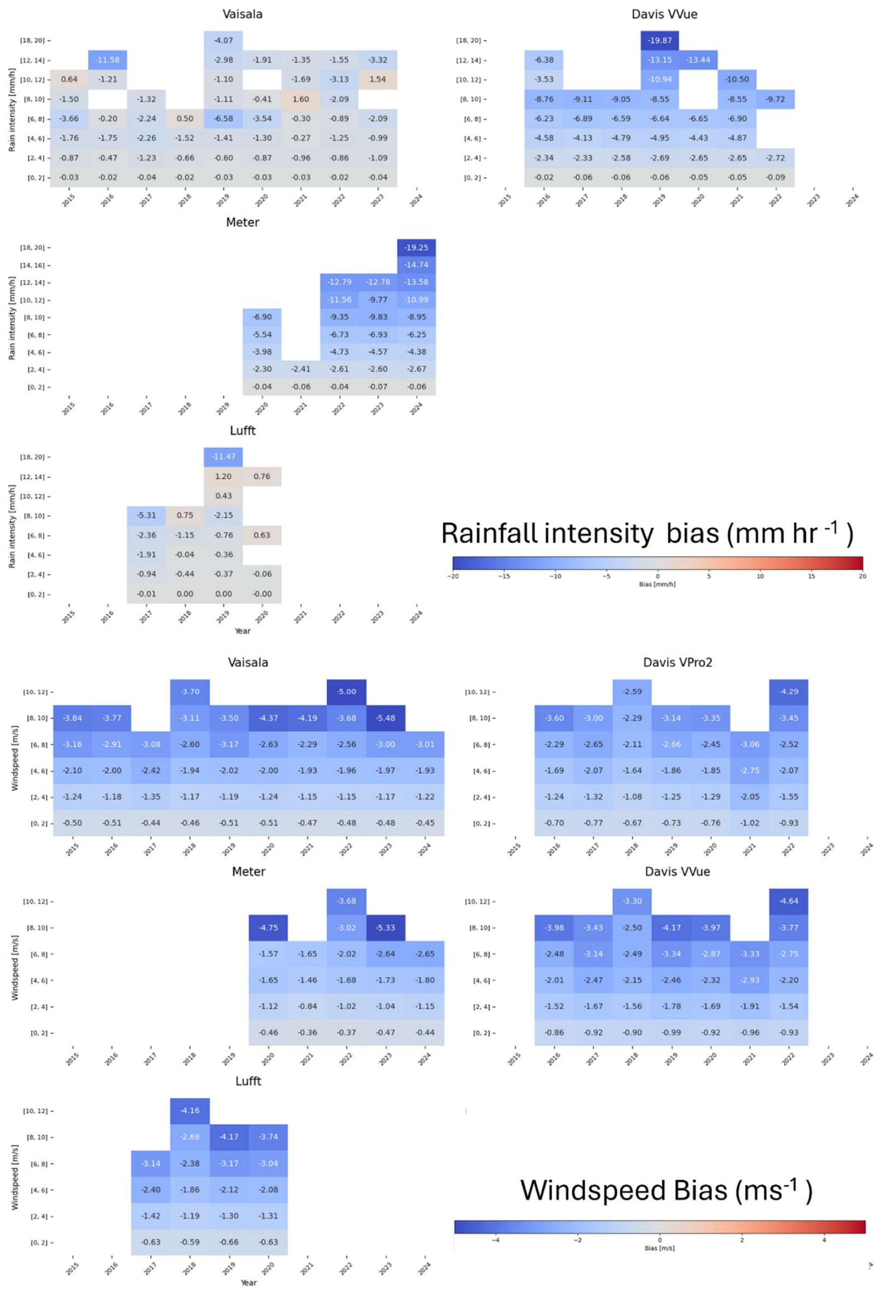

Figure 5 Rainfall intensity bias and windspeed bias from the AiO WS deployed at the testfield.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Figure 6. Detritus found in Davis VP2 tipping bucket rain gauge after 7 years of operation.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

	Class A	Class B	Class C	Class D
Temperature (°C)	< 0.2°C	< 0.5°C	1°C	> Class C
Relative Humidity	< 2%	< 5%	< 10%	> Class C
Wind speed	< 1 m/s or < 5%	< 2 m/s or < 10%	< 3 m/s or < 15%	> Class C
Liquid Precipitation Intensity (mm/hr)	< 0.2 mm/hr or < 5%	< 0.5 mm/hr or < 10%	< 2 mm/hr or < 15%	> Class C

PWS Manufacturers' Specifications confirmed at initial calibration

	Vaisala WXT520	METER ATMOS41	Lufft WS700-UMB	Davis Vantage Pro-2 (both)	Davis Vantage Vue (both)
Temperature (°C)	±0.3°C	±0.2°C	±0.2°C ±0.5°C >30 °C	±0.3°C	±0.3°C
Relative Humidity	±3%	±2%	±2%	±2%	±2%
Wind Speed	±0.3 m/s	maximum: ±3% or ±0.3 m/s	maximum: ±3% or ±0.3 m/s	maximum: ±5% or ±0.9 m/s	maximum: ±5% or ±0.9 m/s
Liquid Precipitation Intensity (mm/hr)	±2%	±5%	±2%	maximum: ±3% of total or ±0.2mm	maximum: ±5% of total or ±0.9mm

Table 1. WMO classification of measurement accuracy from AiO WS and specifications versus observed stated and laboratory confirmed measurement accuracy for new AiO WS units.

11. References

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Bell, S., Cross, J., and Parry, D.: A methodology for study of in-service drift of meteorological humidity sensors, Metrologia, 54, 583–591, https://doi.org/10.1088/1681-7575/aa7f2b, 2017.

S. Bell, D. Cornford en L. Bastin, „How good are citizen weather stations? Addressing a biased opinion," Weather, nr. 70, pp. 75-84, 2015.

Chapman, L., Bell, C., Bell, S., Litchfield, A., Stoke, A., Neely, A., and Hemming, D.: Can citizen science data be used to predict urban heat island exposure? A case study from Birmingham, UK, Theor. Appl. Climatol., 141, 949–965, https://doi.org/10.1007/s00704-020-03199-0, 2020.

Cornes, R., Fenner, D., Jones, C., and Lussana, C.: A correction method for short-wave radiation bias in citizen temperature observations, Q. J. R. Meteorol. Soc., 146, 2430–2446, https://doi.org/10.1002/qj.3794, 2020.

De Vos, L. W., Droste, A. M., Zander, M. J., Overeem, A., Leijnse, H., Heusinkveld, B. G., Steeneveld, G. J., and Uijlenhoet, R.: Hydrometeorological monitoring using opportunistic sensing networks in the Amsterdam metropolitan area, Bull. Am. Meteorol. Soc., 101, E167–E185, https://doi.org/10.1175/BAMS-D-18-0237.1, 2020.

Dombrowski, O., Klaus, J., Gisler, M., and Niessner, F.: Performance of the ATMOS41 all-in-one weather station for weather monitoring, Sensors, 21, 741, https://doi.org/10.3390/s21030741, 2021.

Droste, A. M., Heusinkveld, B. G., Fenner, D., and Steeneveld, G. J.: Assessing the potential and application of crowdsourced urban wind data, Q. J. R. Meteorol. Soc., 146, 2671–2688, https://doi.org/10.1002/qj.3829, 2020.

Fenner, D., Chapman, L., Bechtel, B., and Demuzere, M.: Correcting citizen-science air temperature measurements across urban landscapes, Int. J. Climatol., 38, 539–552, https://doi.org/10.1002/joc.5362, 2018.

Hahn, C., Garcia-Martí, I., Sugier, J., Emsley, F., Beaulant, A.-L., Oram, L., Strandberg, E., Lindgren, E., Sunter, M., and Ziska, F.: Observations from personal weather stations—EUMETNET interests and experience, Climate, 10, 192, https://doi.org/10.3390/cli10120192, 2022.

Ingleby, B., Moore, D., Sloan, C., and Dunn, R.: Evolution and accuracy of surface humidity reports, J. Atmos. Oceanic Technol., 30, 2025–2043, 2012.

Jenkins, T.: A comparison between two types of widely used weather stations, Weather, 69, 100–107, https://doi.org/10.1002/wea.2292, 2014.

KNMI TR296: Wind Tunnel and Field Test of Three 2D Sonic Anemometers, Royal Netherlands Meteorological Institute Technical Report, 2020.

Meier, F., Fenner, D., Demuzere, M., and Kadygrov, N.: CrowdQC+: A quality control method for crowdsourced meteorological data, Environ. Model. Softw., 96, 188–200, https://doi.org/10.1016/j.envsoft.2017.06.021, 2017.

Meier, F., Fenner, D., Grassmann, T., Otto, M., and Scherer, D.: Crowdsourcing air temperature from citizen weather stations for urban climate research, Urban Clim., 19, 170–191, https://doi.org/10.1016/j.uclim.2017.01.006, 2017.

https://doi.org/10.5194/egusphere-2025-5194
Preprint. Discussion started: 27 October 2025
© Author(s) 2025. CC BY 4.0 License.

Napoly, A., Grassmann, T., Meier, F., and Fenner, D.: Development and application of a statistically-based quality control for crowdsourced air temperature data, Front. Earth Sci., 6, 118, https://doi.org/10.3389/feart.2018.00118, 2018.

NOAA PMEL Technical Note: Wind Speed Variability of Vaisala WXT520, Pacific Marine Environmental Laboratory, NOAA, 2011.

Overeem, A., Robinson, J. C., Leijnse, H., Steeneveld, G. J., Horn, B. K. P., and Uijlenhoet, R.: Crowdsourcing urban air temperatures from smartphone battery temperatures, Geophys. Res. Lett., 40, 4081–4085, https://doi.org/10.1002/grl.50786, 2013.

Saraf, R., and Ivan, G.: Reed Switch Degradation in Environmental Applications, IEEE Trans. Device Mater. Reliab., 1978.

Segovia-Cardozo, D. A., Bernal-Basurco, C., and Rodríguez-Sinobas, L.: Tipping bucket rain gauges in hydrological research: Summary on measurement uncertainties, calibration, and error reduction strategies, Sensors, 23, 5385, https://doi.org/10.3390/s23125385, 2023.

Vučković, M., and Schmidt, J.: On the importance of data quality assessment of crowdsourced meteorological data, Sustainability, 15, 6941, https://doi.org/10.3390/su15086941, 2023.

World Meteorological Organization: Guide to Instruments and Methods of Observation (WMO-No. 8), 2018 edition, WMO, Geneva, Switzerland, 2018.