https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

A multi-stage quality control framework for automated weather station networks in semi-arid environments: Application to SASSCAL WeatherNet in Botswana

Thuso H. Tlhalerwa[1,2], Kgakgamatso M. Mphale[2], Kago E. Maabong[1], Douglas Letsholathebe[2]

[1]Department of Chemical and Physical Sciences, Botswana University of Agriculture and Natural Resources, Gaborone, Botswana
[2]Department of Physics, University of Botswana, Gaborone, Botswana

Correspondence to: Thuso H. Tlhalerwa (thusotlhalerwa@gmail.com)

Abstract. Automated weather stations (AWS) are increasingly used to fill gaps in ground-based observations in data-sparse regions, but their reliability depends on robust quality control (QC). This study presents a multi-stage QC framework that integrates basic integrity, range checks, climatic outlier detection, and internal, temporal, and external consistency assessments. The framework was applied to twelve AWS of the Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) WeatherNet network in Botswana. A flagging system distinguishes critical errors, which are removed, from suspicious observations, which are retained, so that genuine extreme values are not mistaken for errors. The network had high data availability, with a mean uptime of 96 % (87.7-100 % across stations), and fewer than 2 % of temperature observations were flagged as climatic outliers. Spatial and seasonal variations in data quality were detected, including site-dependent performance differences and sensor-specific anomalies. Comparison with ERA5-Land temperature identified three stations (Shakawe, Sowa, and Tubu) with differences that warrant further investigation. The proposed framework is transferable to other data-sparse regions and provides a structured approach to improving data reliability in semi-arid environments.

1 Introduction

Reliable, accurate and timely meteorological observations are essential for climate research, agricultural planning, and water resource management. The need is greater in southern Africa, where vast areas lack adequate ground-based weather stations or the collected data is not easily accessible with Botswana facing the same problem. The country's weather stations are unevenly distributed, and some areas lack long climatological records. Accurate observations are particularly needed because the country is predominantly semi-arid and climate-variable, which could negatively affect agricultural initiatives and water resources (Batisani & Yarnal, 2010; Juana et al., 2016; Mphale et al., 2014). Molosiwa et al. (2026) identified major research gaps in Botswana in forecasting, early drought warning systems, and mitigation efforts.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Without adequate ground-based observations, it is difficult to validate satellite products and climate models (New et al., 2006). According to Molosiwa et al. (2025) and Nkoni et al. (2024), adaptation efforts in Botswana cannot be accurately assessed because too few historical climate records are available. Establishing and maintaining weather stations to address this is expensive, especially in remote and harsh environments. Automated weather stations (AWS) have been established to help fill this gap.

1.1 Automated Weather Stations and Quality Control

AWS are fitted with sensors, dataloggers, and communication systems that measure, record, and transmit ground-based meteorological data autonomously. Because they are automated, they can be deployed in locations where staffing, or even building conventional weather stations, would be impractical or expensive. They can operate for extended periods with minimal maintenance because they are solar powered and transmit data via satellite or cellular networks. This makes them well suited to data-sparse regions such as Botswana and to rough terrains (Muche et al., 2018).

However, AWS have limitations, mainly in data quality. They are prone to sensor drifts and can be buried by the sand or flood waters, fail in extreme temperatures, suffer communication failure, or have components damaged by animals (Muche et al., 2018). All these can affect the quality of the recorded observations. Unlike conventional, manually operated stations, AWS malfunctions may remain undetected for a long time, until someone reviews the data or visits the site (Shafer et al., 2000). It is therefore vital to apply quality control to data from AWS networks.

Several quality control frameworks (QCF) have been developed to address these challenges. Durre et al. (2010) established comprehensive procedures that involve range checks, internal consistency tests, temporal checks for spikes and flatlines, and spatial consistency assessments with neighbouring stations. Fiebrich et al. (2010) documented quality assurance procedures showing how real-time automated checks combined with expert review can preserve data quality in operational networks. Estévez et al. (2011) provided practical guidelines for validating AWS data in agricultural and hydrological applications. In addition, World Meteorological Organization (WMO), (2018) gave guidance on instrument standards, measurement uncertainty, and quality management systems for operational networks, establishing international benchmarks for acceptable data quality. These frameworks use sequential validation stages that separate physically implausible values from statistically unusual ones that may be actual extreme events. This preserves potentially valid climate extremes while protecting users from obvious measurement errors (Hubbard et al., 2005; World Meteorological Organization (WMO), 2011).

1.2 The SASSCAL WeatherNet initiative

Recognizing the need for reliable climate data in southern Africa, the Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) established a WeatherNet network of AWS. With support from Germany, the initiative installed AWS in Angola, Botswana, Namibia, South Africa, and Zambia. The AWS are equipped with WMO-certified sensors that measure various weather variables at 15-minute intervals. The data is then transmitted near

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

real-time to a central database, where it is freely available to download through a SASSCAL online portal (Muche et al., 2018).

In Botswana, the WeatherNet AWS expanded ground-based observation coverage. By 2024, the network comprised 20 AWS in diverse environments, from the Okavango Delta wetlands in the northwest, to the Kalahari sands in the southwest, to the eastern hard veldt. Most of the stations have operated since February 2014 or March 2015, accumulating a decade of observations.

Kaspar et al. (2015) and Muche et al. (2018) reported that SASSCAL WeatherNet data already undergo automated QC tests. However, there is no documentation that describes the tests, thresholds applied or the results of these checks. This lack of transparency raises questions about the reliability of the recorded data, given that Estévez et al. (2011) noted that automated checks alone can fail to uncover subtle quality issues.

1.3 Research objectives

The main aim of this study is to address the detected gaps by presenting and documenting a five-stage QCF designed for the SASSCAL WeatherNet AWS in Botswana. The AWS data will be checked for completeness, range validity, internal consistency, temporal stability, and external coherence. Furthermore, the framework will also try to characterize performance at station and network level and identify problematic stations. The result will be a detailed and replicable template for similar AWS networks in data-scarce, semi-arid regions (Fig. A1 in the appendix).

2 Study area and data

2.1 Study area

Botswana is a semi-arid, landlocked country in southern Africa, characterized by strong seasonal and high interannual rainfall variability. The mean annual precipitation ranges from approximately 330 mm in the southwest to 600 mm in the north-northeast (Moses, 2019). The country has two eco regions, the sand veldt, and the hard veldt. The sand veldt lies along the western part of the country and forms part of the Kalahari desert. The country's elevations range from about 800 m to 1300 m (see Fig. 1) and were generated from ETOPO 2022 15 Arc Second Global Relief Model (NOAA National Centers for Environmental Information, 2022).

2.2 AWS Network

The meteorological data used in this study come from the SASSCAL WeatherNet AWS across Botswana. At present, these are the only freely accessible, near real-time ground observations available in Botswana (SASSCAL WeatherNet, 2026). Since inception, 20 AWS have been deployed across remote areas of Botswana. At the time of the study, two were not active, one had been inactive for an extended period before being active again, and five had been replaced after a period of inactivity. This left 12 active stations and a large spatial gap. These 12 stations were also selected because their records

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

90 exceeded five years at the time of download. Figure 1 shows the selected stations and their elevations. The data used start from the first full month after each station was activated and end on the 31st December 2024 for all the stations.

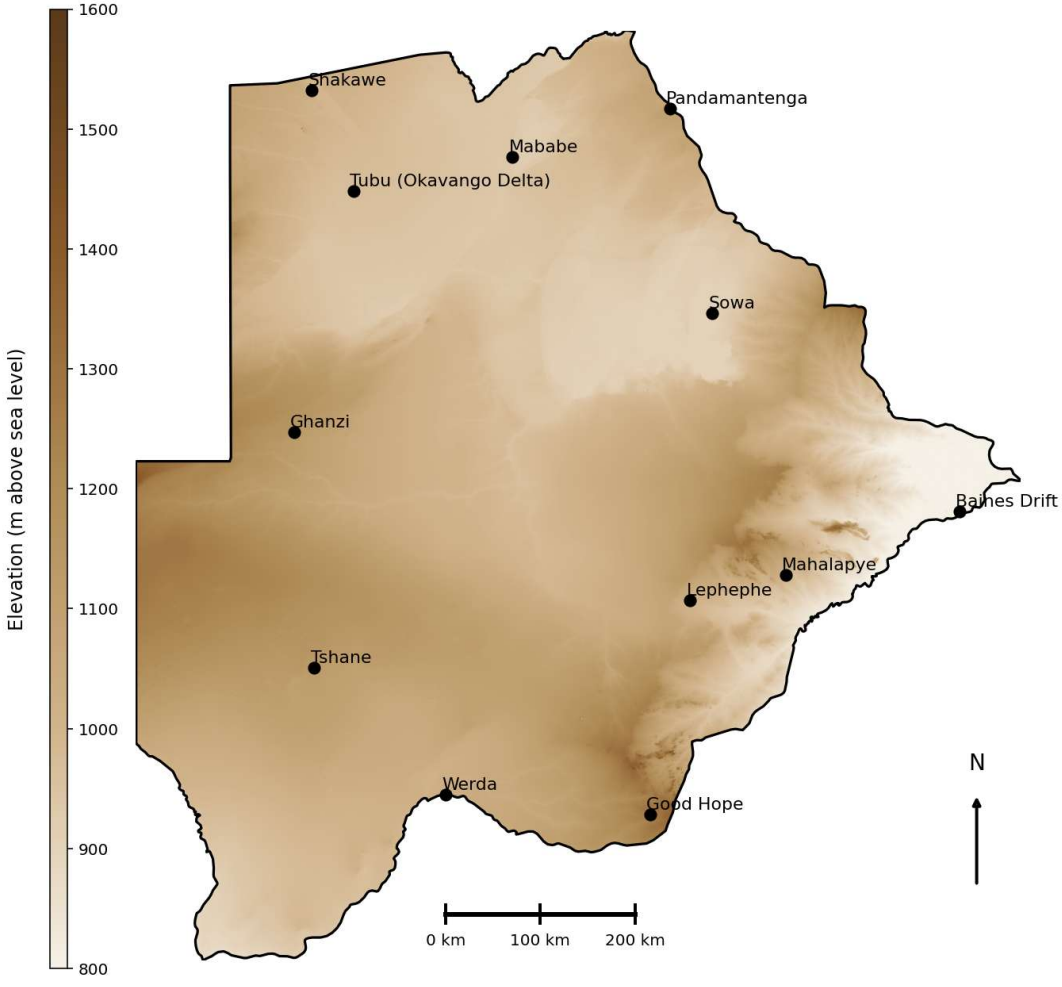


Figure 1: The 12 SASSCAL WeatherNet AWS in Botswana used for the study.

The 12 stations represent diverse environments in Botswana. The locations ranged from the Okavango delta wetlands
95 (Shakawe and Tubu), through the semi-arid Kalahari desert (Ghanzi, Tshane and Werda), to the eastern hard veldt (Mahalapye and Goodhope). Each AWS is equipped with WMO-certified sensors (see Table 1). Although each AWS records multiple meteorological variables, the QCF focused on six variables, chosen for their importance in climate monitoring in semi-arid environments and their relevance to agricultural and hydrological applications. These variables were

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

minimum (T_n), maximum (T_x) and mean (T_a) air temperature, total precipitation (Precip), mean wind speed (W_m) and mean relative humidity (RH).

Table 1: Instrumentation and measurement specifications for SASSCAL AWS stations in Botswana

Parameter/Variable	Sensor Type	Operating Range	Measurement Height
Temperature/ Humidity	Vaisala HMP155	-80 to +60 ºC 0 to 100 % RH	3.1 m
Wind Speed/ Direction	Vaisala WA15	0.4 to 75 m s^{-1} 0 to 360 º	10 m
Rain	Young 52202 Tipping Bucket	0 to 50 mm hr^{-1}	1.2 m
Data Logger	Campbell CR1000		1.7 m

The WeatherNet records 15-minute observations in real-time and then aggregates to daily and monthly values. This study used daily data for all analyses because this temporal resolution matches the requirements of most climate applications and aligns with established QC methodologies (Durre et al., 2010). Daily aggregation helps to reduce data volume while preserving information on extreme events and temporal variability that is essential for climate monitoring.

2.3 Reference dataset

Independent reference datasets were used for external consistency assessment of the AWS precipitation and temperature variables. ERA5-Land hourly data was used for the temperature variables (T_n and T_x) while CHIRPS v2 (Climate Hazards Group InfraRed Precipitation with Stations data version 2) for Precip (Funk et al., 2015; Muñoz-Sabater et al., 2021). Their role was not to represent absolute truth but to provide a spatially consistent benchmark against which AWS observations can be assessed for plausibility and reliability.

3 Methodology

A five-stage QCF was applied to the SASSCAL AWS observations to identify potential errors and quality issues. The sequential stages were: (1) basic integrity checks, (2) range checks and climatic outlier detection, (3) internal consistency checks, (4) temporal consistency checks, and (5) external consistency checks against independent reference datasets. Each observation was assigned quality control flags. The flagging system distinguished severe errors, which required removal, from suspicious values, which were retained. The flagging system allowed that a record which failed more than one stage could receive multiple flags. This approach in turn retains plausible extreme events while giving users the information they need to make application-specific decisions about data quality (Zahumenský, 2004).

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.1 Basic integrity checks

The first stage applied basic integrity checks to the preliminary observations to assess their structural quality before performing value-based checks. AWS data can suffer from time-formatting errors and from improperly combined datasets (Faybishenko et al., 2022). There were four types of basic integrity issues assessed, each with its own flag: (1) missing dates in the daily sequence (flag 1a), (2) duplicate timestamps (flag 1b), (3) chronological order violations (flag 1c), and (4) missing values for individual variables (flag 1d).

3.2 Range checks and climatic outlier detection

The second stage employed two complementary approaches to detect questionable values, being range checks and climatic outlier detection.

3.2.1 Range checks

Range checks are used to detect values that exceed physical limits for the Earth's atmosphere (Brunet et al., 2020). The limits, adopted from Patro and Bartakke (2025), are extreme boundaries beyond which values are considered instrument errors and not valid observations. The limits were $0 \leq RH \leq 100$ %, $-40 \leq T_{n/x/a} \leq 100$ °C, and $0 \leq W_m \leq 75$ m s^{-1}. Observations that failed the range test were flagged (flag 2a) and removed from the dataset. Precipitation was excluded from the range checks because of its natural variability in Botswana, making a fixed upper threshold inappropriate (Matenge et al., 2023).

3.2.2 Climatic outlier detection

Range checks detect physically impossible values, while climatic outlier detection flags values that are physically possible but highly unlikely for the region and season. It was applied to T_n, T_x, T_a, and W_m (Brunet et al., 2020). Because regionally specific climatological thresholds have not been established for Botswana, the method from Durre et al. (2010) was adopted, which uses seasonally stratified percentile thresholds computed within a 15-day moving window. The 0.1st and 99.9th percentile thresholds were calculated for each calendar day from data within ±7 days of that day across all available years, so the test date is at the centre of a 15-day window. A minimum of 100 valid observations in the window was required to ensure a robust seasonal threshold estimate. The days with insufficient data were not assessed. Thresholds were computed separately for each station to accommodate spatial climate variation across Botswana. The values below the lower threshold or above the upper thresholds were flagged as climatic outliers (flag 2b). Unlike range-check failures, climatic outliers were retained to preserve possible genuine extreme events. Rainfall was excluded for the same reason as in the range checks.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.3 Internal consistency checks

The third stage applied internal consistency checks to examine whether related meteorological variables maintain physically logical relationships. This helps to identify sensor malfunctions or data transmission errors (Patro and Bartakke, 2025). There were four internal consistency tests applied to the daily data. First, the temperature hierarchy was verified, see Equation (1):

$$T_n \leq T_a \leq T_x \quad (1)$$

Any violation of this relationship indicates a clear instrument error, so such values were flagged (flag 3a) and removed from the dataset. Secondly, a calculated mean temperature $T_{a(cal)}$ was computed using Equation (2) and compared to the measured mean temperature (T_a) using Equation (3):

$$T_{a(cal)} = \frac{T_n + T_x}{2} \quad (2)$$

$$\left|T_a - T_{a(cal)}\right| > 1\ ^{o}C \quad (3)$$

If the difference exceeded 1 °C, the values were flagged (flag 3b) yet retained. This is consistent with the tolerance used in operational quality control systems by Durre et al. (2010). That threshold was derived from GHCN-Daily stations, which are in temperate climates, so the thresholds could be conservative for semi-arid Botswana. Moreover, modest discrepancies are likely to arise from measurement timing differences or natural diurnal temperature patterns, not instrument failure.

Thirdly, the measured relative humidity (RH) was compared to a calculated relative humidity (RH_{cal}), computed from dew point temperature (T_d) using the Magnus-Tetens approach (Lawrence, 2005). This checks the internal consistency among T_a, T_d and RH, see Equation (4):

$$RH_{cal} = 100 \times 10^{\frac{7.5T_d}{237.3+T_d} - \frac{7.5T_a}{237.3+Ta}} \quad (4)$$

Discrepancies of more than 10 % were flagged but retained. This threshold was deliberately wide compared with the HMP155 sensor accuracy of ±1.0 to 1.7 % RH under normal conditions (Vaisala, 2012).

Finally, the logical consistency between wind speed (W_m) and direction (W_d) was evaluated. The rule was to identify and remove cases (flag 3a) that had $W_m = 0$ but $W_d \neq 0$ or $W_m > 1$ m s^{-1} but a missing W_d (Degaetano, 1997). The method recognized that W_d is only physically meaningful when the W_m exceeded the WA15 sensor's beginning threshold of about 0.4 m s^{-1}.

3.4 Temporal consistency checks

The fourth stage employed temporal consistency checks to identify unrealistic rates of change and suspicious patterns inside the time series. These checks identified sudden spikes or jumps, as well as consecutive identical values (flatlines). These could be caused by sensor malfunction or data recording errors. This stage assessed the coherence of observations across sequential days.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.4.1 Spike detection

The spike test was applied to the temperature variables. This was because temperature spikes can expose abrupt or unrealistic changes between consecutive days, which may stem from instrument malfunction or data transmission errors. Following the procedure of Durre et al. (2010), a temperature value on day i was flagged as a spike if the condition in Equation (5) were met:

$$|T_i - T_{i-1}| > \text{threshold and } |T_i - T_{i+1}| > \text{threshold} \quad (5)$$

where T_i is the temperature on day i, and the threshold was set at 10 °C (flag 4a soft, retained) and 20 °C (flag 4a hard, removed).

3.4.2 Flatline detection (persistence test)

Flatline checks detect sequences of consecutive identical values, which can be caused by stuck sensors or instrument malfunction (Durre et al., 2010). Flatlines were assessed for the temperature variables (T_n, T_x, T_a) and for non-zero precipitation values. Sequences of 10 or more consecutive identical values were flagged but retained (flag 4b soft). Sequences of 15 or more were flagged, with the first value retained and the subsequent values removed (flag 4b hard). These thresholds are stricter than the 20-day threshold of GHCN daily (Durre et al., 2010) and were chosen to enhance error detection in Botswana's variable climate.

3.5 External consistency checks

The final stage comprised external consistency checks, which add a further layer of quality assurance by comparing AWS observations to independent estimates (Duan et al., 2016). Temperature variables were compared to ERA5-Land hourly data, and precipitation data were compared with CHIRPS. This comparison can identify issues such as instrument drift, calibration problems, and site-specific anomalies that internal checks may miss. It is known that both reference datasets have limitations, coarse spatial resolutions and can miss localized phenomena. On the other hand, AWS measurements are at a single point that may not represent the surrounding grid cell. For each AWS location, the corresponding grid cell was identified, and the mean bias (MB), standard deviation (SD), root mean square error (RMSE), and Pearson correlation (r) were computed for each variable. In addition, the daily difference (AWS minus satellite) was calculated for all paired observations, with the threshold set at three standard deviations from the mean. This captures true outliers while accommodating typical point-to-grid mismatches. The values that exceeded the thresholds were flagged (flag 5) and retained to preserve valid data while allowing end users to apply their own criteria based on application-specific tolerance for AWS-satellite disagreement.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

4 Results

4.1 Basic integrity checks

The basic integrity checks on the 12 AWS detected no duplicate dates and the chronological sequence was consistent throughout. Missing dates were detected at two stations, Mababe and Tubu, with violation percentages of 0.62 % and 0.10 %, respectively. The dataset was in good condition with regard to dates.

4.1.1 Station uptime

The SASSCAL WeatherNet achieved an average uptime of 96 % across all the AWS (Fig. 2), a good result for a remote AWS network across varied climate settings. Mean annual uptime at individual stations ranged from 87.7 % at Mababe to 100 % at Sowa. Based on this range, the stations were ranked by mean annual uptime and divided into three performance tiers.

Tier 1 stations, the most dependable, had a mean annual uptime of 98 % or more. These stations include Sowa (100.0 %), Ghanzi (99.9 %), Shakawe (99.3 %), Tubu (99.2 %), Baines Drift (98.9 %), Goodhope (98.5 %), and Tshane (98.0 %) as shown in Fig. 2. These stations accounted for 58 % of the network and showed near-continuous data availability. Their geographical dispersion suggests that uptime does not depend solely on climate zone. Factors such as accessibility, infrastructure quality and maintenance frequency may also contribute to performance (Muche et al., 2018).

Tier 2 stations showed good but marginal performance with mean annual uptime of 90-98 %. They make up 25 % of the network, Werda (96.6 %), Mahalapye (93.9 %) and Lephephe (92.7 %). They were broadly functional, but their downtime warrants closer monitoring and targeted maintenance.

Tier 3 stations, the lowest performers, had a mean annual uptime below 90 %, Pandamatenga (87.8 %) and Mababe (87.7 %). They are clear outliers and need priority attention, including constant monitoring and maintenance. Both are situated in northern Botswana, within the broader Okavango-Chobe region. Mababe and Pandamatenga also raise data quality concerns. At Mababe, more than one in eight daily records across the full period is missing. Such gaps limit the reliability of any downstream climate or hydrological application of the data.

Figure 2 further shows seasonal uptime patterns that extend beyond individual station characteristics and point to network-wide environmental drivers. The seasonal network means were 95.4 % (DJF), 96.6 % (MAM), 95.6 % (JJA), and 96.3 % (SON).

DJF had the lowest of the four seasons. It coincides with Botswana's summer rainfall period, when thunderstorms, lightning, flooding, and high humidity are common (Shafer et al., 2000). These conditions may increase operational stress on AWS electronics and communication systems and thus contribute to the seasonal differences in data quality. Estévez et al. (2011) similarly noted that QC performance in semi-arid climates is sensitive to environmental stressors. The most affected stations in DJF are Pandamatenga and Mababe, both areas of high rainfalls.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

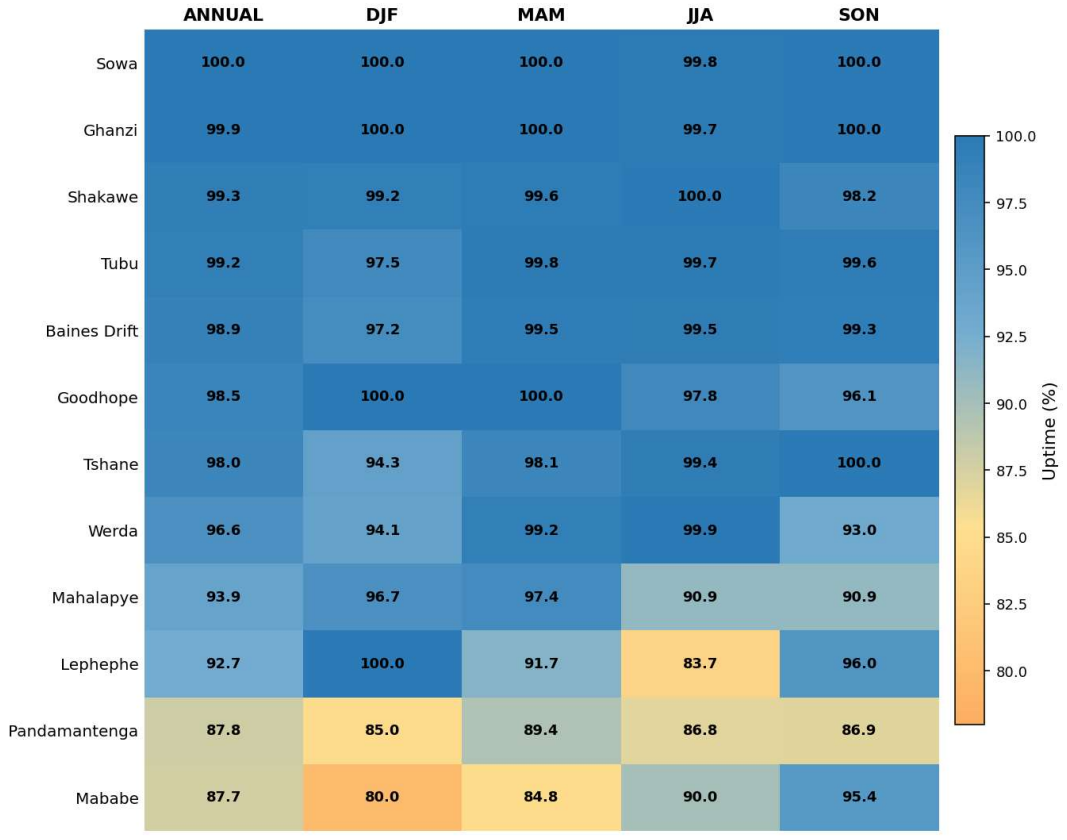


Figure 2: Annual and seasonal uptime (%), DJF = December, January, February, MAM = March, April, May, JJA = June, July, August, SON = September, October, November.

4.1.2 Seasonal performance patterns

MAM had the highest mean network uptime. This autumn season follows the peak of the summer rains when environmental conditions calm and maintenance is easier. JJA, the Botswana winter, also showed a decline in uptime, which was slightly smaller than in DJF (95.6 % versus 95.4 %). The failure could be caused by reduced solar irradiance from shorter days, or by cold nights that stress batteries and other components. These failure mechanisms remain hypotheses unless supported by power-system or equipment diagnostics. Uptime improved again in SON (spring). This pattern points to a modest recovery outside the summer and winter seasons, which bring weather extremes. Together, the results suggest that uptime was influenced by both station-specific characteristics and broader seasonal conditions.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Three of the stations showed extreme seasonal degradation, Lephephe, Mababe, and Pandamatenga. Despite similar low annual means (87-93 %), they had different seasonal failure patterns. Lephephe had 100 % uptime in DJF, 91.7 % in MAM, 83.7 % in JJA and 96 % in SON. The station recovered in SON and fully in DJF. This recovery shows that its seasonal stress is reversible rather than permanent. Mababe's DJF uptime of 80.0 % was the lowest seasonal value in the network. It recovered progressively through MAM (84.8 %), JJA (90.0 %) and SON (95.4 %). The low DJF uptime may be related to the station's location in the Mababe depression. The station may get flooded during the wet summer or damaged by animals (Muche et al., 2018). Pandamatenga showed the most concerning pattern. Its uptime stayed between 85 and 89 % in all seasons and never reached 90 %. Unlike Lephephe and Mababe, which recovered during favourable seasons, it underperformed chronically. This implies that the station needs extra maintenance attention and that infrastructure deficiencies may need to be addressed.

The seasonal results show that network reliability cannot be assessed based on uptime alone. Stations with similar annual means can have different failure patterns. Hence, combining annual and seasonal uptime provides a more informative basis for maintenance prioritization and QC decision making.

4.1.3 Outage characteristics

Outage analysis provides a review of reliability beyond mean uptime. It distinguishes chronic data loss from catastrophic single-event failures (Fig. 3). The relationship between outage frequency and maximum outage duration shows four distinct patterns. The first pattern is exceptionally reliable stations, Sowa and Ghanzi. They had few outages and short maximum outage durations, consistent with their ranking as the two best-performing stations by mean annual uptime. Their performance may reflect limited exposure to environmental stressors, effective communications, or timely and effective maintenance, but the available outage data cannot distinguish among these explanations.

The second pattern is frequent but short outages, seen at Baines Drift, Tubu and Shakawe. Their maximum outage durations did not exceed 17 days. All three stations are Tier 1, so they form the second most reliable group. Their outage profiles suggest recurrent minor disruptions instead of prolonged system failures.

The third pattern is isolated catastrophic events, seen at Goodhope, Mahalapye, Tshane and Werda. These stations had fewer outages but long maximum durations. Goodhope and Tshane are Tier 1 stations, and Mahalapye and Werda are Tier 2. The fourth pattern is chronic instability as seen at Lephephe, Mababe, and Pandamatenga. These stations had frequent outages combined with exceedingly long outage durations and are the most problematic in the network. Pandamatenga's longest outage lasted for 332 days, almost an entire year. This matches its poor performance in all seasons. The combination of frequent and long outages indicates a systematic reliability issue, not isolated events.

Figure 3 also shows that outages longer than 30 days were few and occurred at seven of the stations. Their spatial distribution is markedly highly uneven and chronic long-duration failures are not a network-wide issue but affect only some vulnerable stations. Using a seven-day threshold showed a similar pattern.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

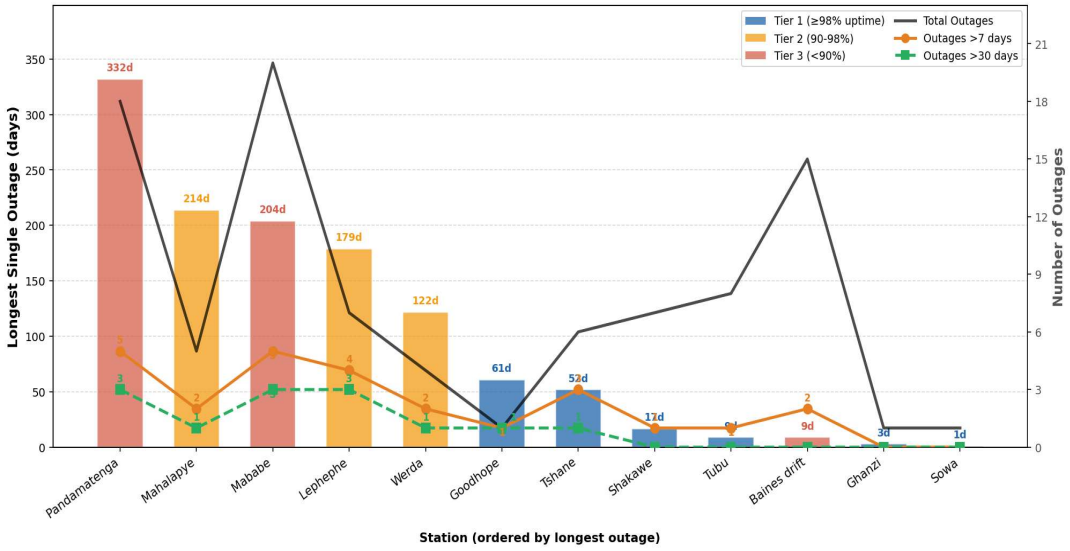


Figure 3: Outage frequency and severity across the SASSCAL WeatherNet in Botswana.

4.1.4 Missing data by variable

Missing data analysis was conducted because an AWS can be online while failing to record or transmit one or more variables. Variable-level analysis can therefore reveal sensor-specific problems that complete station outages do not capture. The percentage of missing observations for each variable at each station is shown in Fig. 4a. Two distinct patterns emerge in the network. The first is uniform missing data across variables, as at Ghanzi, Goodhope, Mahalapye, and Werda. This is consistent with station-wide outages affecting all sensors at once, from power failure, communication loss, or datalogger malfunction. These stations recorded less than 7 % missing data for the assessed variables, showing comparatively high completeness despite occasional outages.

The second pattern is differential missing data among variables at the remaining stations (67 % of the network). This is consistent with sensor-specific failures or variable-specific recording problems that occur independently of complete station outages. The temperature variables (T_n, T_x, T_a) had identical missing data percentages because they are derived from the same temperature sensor. They were combined into a single variable, T, for subsequent analysis. In this second pattern, temperature was the most affected variable, with four stations exceeding 10 % missing data, Mababe (21.80 %), Shakawe (14.01 %), Pandamatenga (12.15 %) and Tubu (10.61 %). Mababe lost more temperature data than any other station, consistent with its low seasonal uptime and high outage frequency.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

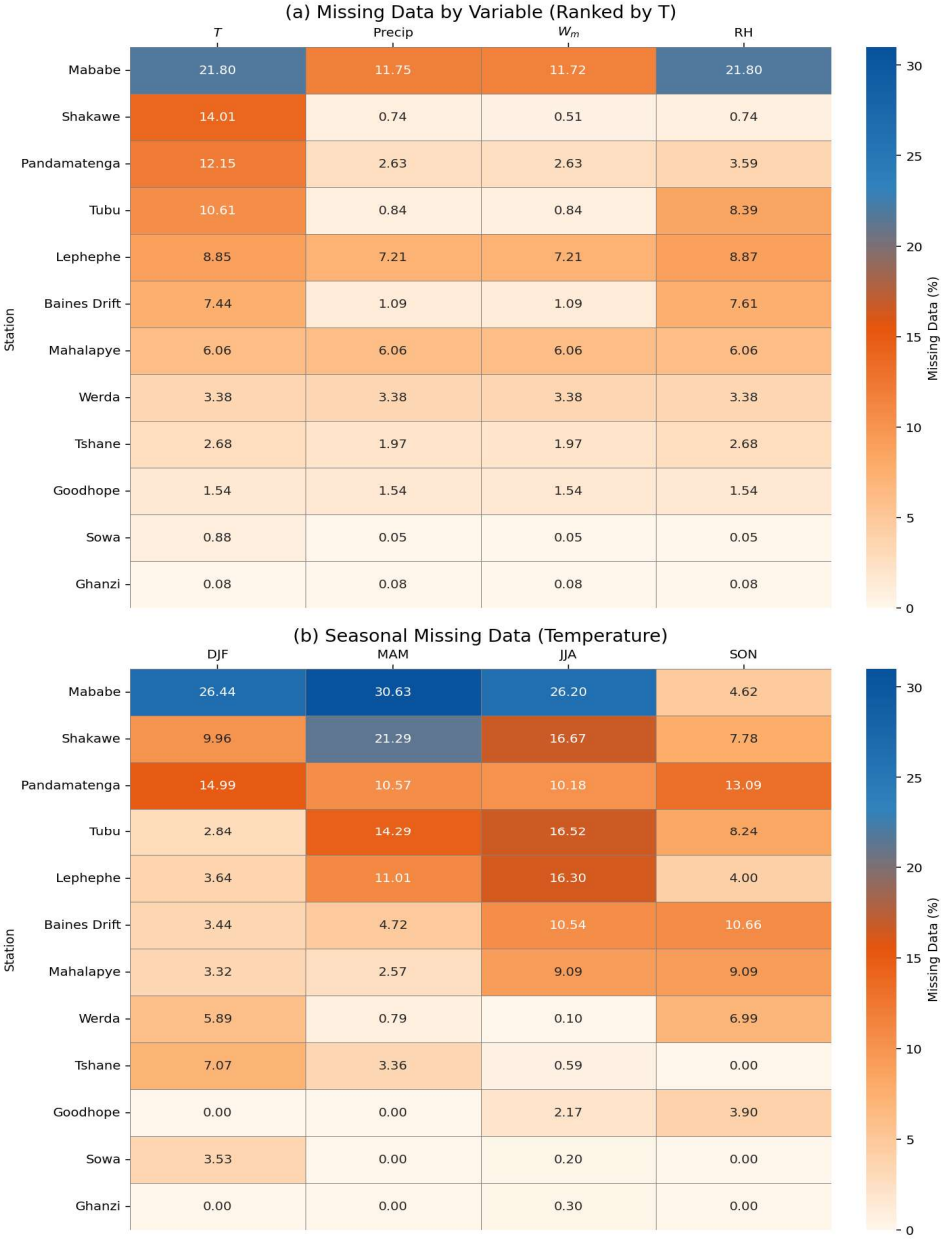


Figure 4: Data completeness across the SASSCAL WeatherNet in Botswana: (a) missing data (%); (b) seasonal missing temperature (%). Both heatmaps are ordered by descending missing T.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

At Baines Drift, Mababe and Pandamatenga, missing temperature and relative humidity data exceeded missing precipitation and wind data. This points to the temperature-humidity sensor being more failure-prone than the rain gauge and anemometer. Tubu had 10.61 % missing temperature data and 8.39 % missing humidity data. This difference may result from partial probe failure or different recovery behaviour between temperature and humidity measurement channels. Mababe was the only station with more than 11 % missing data for every variable, including precipitation and wind.

By contrast, five stations had less than 5 % missing data for all variables, Ghanzi (0.08 %), Sowa (0.05-0.88 %), Goodhope (1.54 %), Tshane (1.97-2.68 %) and Werda (3.38 %). They are practical benchmarks for data completeness and show that near-complete data capture is achievable across Botswana's varied environments.

Because temperature was the most affected variable, its seasonal missingness was examined in more detail (Fig. 4b). The results show that environmental conditions affect temperature sensor performance. Lephephe was most vulnerable in winter (JJA, 16.30 %). This corresponds closely to its JJA uptime of 83.70 %, which shows that most of the seasonal temperature loss resulted from the station being offline. Reduced solar charging or cold temperature may contribute but this needs to be supported by station maintenance records.

Mababe had persistently high missing temperature data in DJF, MAM, and JJA (26-31 %) before recovering to 4.62 % in SON. Its highest temperature loss occurred in MAM (30.63 %) whereas its lowest station uptime was in DJF (79.98 %). This points to different failure processes in the two seasons. DJF may have involved a station-wide outage, and MAM a temperature sensor failure. These results also show why station uptime and variable-level completeness should be analysed separately. Shakawe and Tubu had peak missing temperature data in MAM and JJA (14-21 %). Tshane and Werda showed the opposite pattern with JJA missingness below 1 %. Different environmental stressors thus may drive temperature sensor failures at the different stations.

The variables, T, Precip, W_m, and RH, were then combined into a mean missingness per station which was used to rank the stations (Fig. 5). Mababe was the poorest performer with mean missing data of 16.8 %. Ghanzi and Sowa were the top performers, with mean missing data below 1 %. This study shows that data availability should be assessed at both station and variable level, because station-wide outage metrics identify interruptions that affect the whole AWS, whereas variable-level missingness reveals sensor-specific degradation that may otherwise remain hidden. Combining the two provides a stronger basis for diagnosing failure modes and prioritizing maintenance across the WeatherNet.

4.2 Range and climatic outlier detection

No observations failed the physical range checks. The variables checked (T_n, T_x, T_a, Wm, RH) stayed within physically plausible limits throughout the observation period, implying acceptable sensor operation and data transmission. However, range compliance alone cannot confirm calibration accuracy or exclude more subtle systematic errors.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

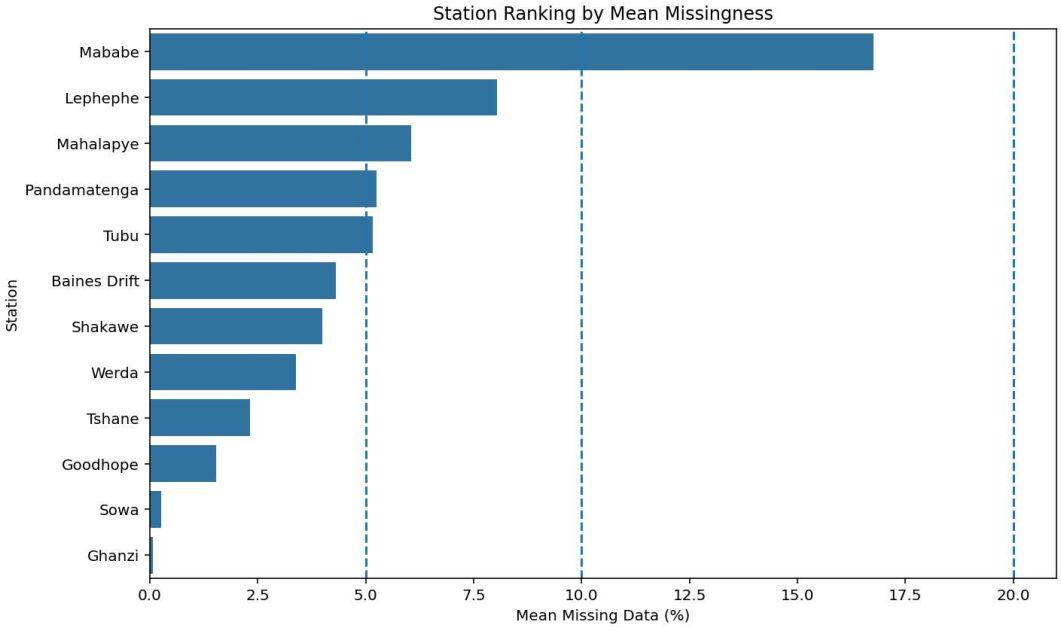


Figure 5: Mean missing data by variables by station.

Climatic outlier rates were low across the network, with all values below 2 % (Table 2). This indicates that the network captured a broad range of meteorological conditions without excessive production of statistically extreme values. The station-level outlier rates ranged from 0.92 % (T_a at Werda) to 1.69 % (T_n Mababe). The network mean outlier rates were 1.21 % for T_n, 1.12 % for T_x and 1.14 % for T_a. This narrow range suggests consistent temperature variability across the geographically dispersed stations and climate zones. It also implies that the 0.1st and 99.9th percentile thresholds flagged only a small proportion of extreme observations without extensively flagging natural climate variability. All flagged outliers were retained in the dataset to preserve actual extreme events, which are important for climate variability studies in semi-arid environments.

4.3 Internal consistency test

The observations that passed the range and climatic outlier checks were subjected to internal consistency tests. The first test, the temperature hierarchy ($T_n < T_a < T_x$), was passed at all the stations, meaning that the temperature variables kept their physical ordering throughout the observation period. Next, the comparison of T_a and $T_{a(cal)}$ detected differences at all stations (Fig. 6), with violation rates of 13-26 %, i.e. differences exceeding 1 °C. These may reflect real meteorological processes rather than sensor failure. The rates were highest at Werda (26.01 %), Mababe (22.75 %), Sowa (22.20 %) and Shakawe (21.22 %). This may reflect greater diurnal asymmetry in their environments, whether in desert margins (Werda, Sowa,

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Tshane) or wetland-influenced zones (Mababe, Shakawe). Spatially, the elevated discrepancies between T_a and $T_{a(cal)}$ form localized clusters (Fig. 6). This indicates that the inconsistencies are influenced by regional climatic regimes.

Table 2: Climatic outlier percentages by variable and station

Station	T_n	T_x	T_a
Baines Drift	1.49	1.30	1.12
Ghanzi	1.16	1.06	1.06
Goodhope	0.95	1.10	1.08
Lephephe	1.05	1.19	1.17
Mababe	1.69	1.34	1.30
Mahalapye	1.16	1.10	1.13
Pandamatenga	1.29	1.04	1.18
Shakawe	0.94	1.08	1.18
Sowa	1.10	1.02	0.99
Tshane	1.17	1.09	1.17
Tubu	1.44	1.10	1.32
Werda	1.12	1.07	0.92

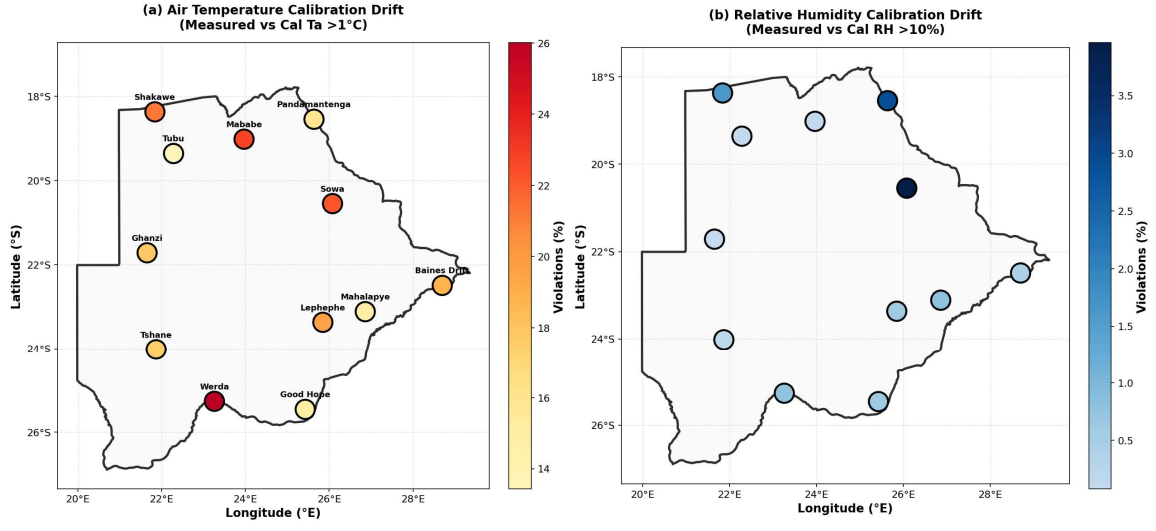


Figure 6: Spatial distribution of internal consistency violations across the AWS network in Botswana: (a) T_a vs $Ta_{(cal)}$ (>1 °C) and (b) RH vs RH_{cal} (>10 %).

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

The next internal consistency check compared measured RH with calculated RH (RH_{cal}). The agreement was strong network-wide, with violation rates below 1 % at ten of twelve stations. Only Sowa (3.96 %) and Pandamatenga (2.91 %) exceeded a 1 % violation rate. These elevated rates may point to a calibration issue or dew-point measurement inconsistencies. The evidence for Sowa is its high violation rate combined with its elevated T_a discrepancy (22.20 %). The station could benefit from a sensor recalibration if this has not already been done.

In contrast to temperature, the spatial distribution of RH discrepancies is weak and localized (Fig. 6). Elevated values were mainly in the northeastern part of the country. This indicates that RH sensor performance is stable across the network, with deviations more likely to be site-specific instead of systematic.

The last test checked the logic of wind speed and direction. The violations were negligible across the network (<1 %), so the wind sensor appears to have worked well. Only Baines Drift (0.65 %) is worth noting.

4.4 Temporal consistency tests

The data that passed internal consistency tests were assessed for temporal anomalies, beginning with the spike test. Temperature spike rates were low network-wide (Fig. 7). T_a was the most temporally stable, with violation rates below 0.2 % at all stations.

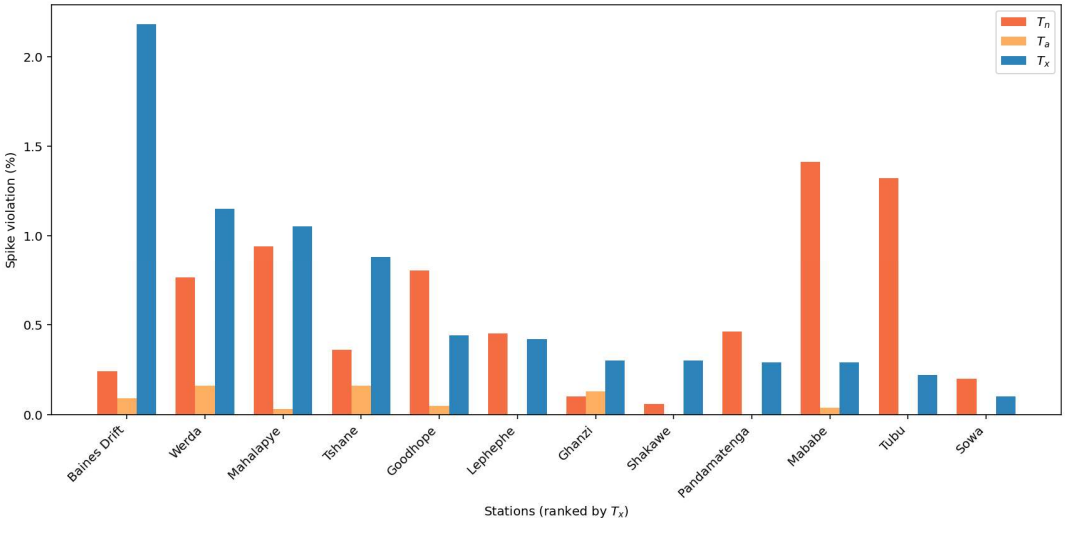


Figure 7: Temporal spike rates by station for the temperature variables.

T_x had the highest spike rates, with Baines Drift (2.18 %), Werda (1.15 %), and Mahalapye (1.05 %) exceeding 1 %. Because these stations did not show comparable T_a and T_n spike rates, the T_x spikes may represent short-lived daytime heating not sensor glitches. T_n spike rates were elevated at Mababe (1.41 %) and Tubu (1.32 %), consistent with their high climatic outlier rates (Sect. 4.2). The pattern may reflect localized nocturnal temperature variability, but sensor or sampling

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

effects are also possible. The absence of corresponding T_a spikes (<0.05 % at both stations) points to physical temperature events rather than drifting sensors. The daily mean also appears to smooth out the short-term jumps in T_n and T_x, which suggests that extremes offset each other. The flatline test found almost no violations (<0.03 %). Only three isolated cases were detected at Ghanzi (T_a, 0.03 %), Shakawe (T_n, 0.03 %) and Mahalapye (Precip, 0.03 %). This shows that stuck sensors, frozen loggers, and repeated identical values were hardly ever present.

4.5 External consistency with reference datasets

AWS temperature observations correlated strongly with ERA5-Land temperature across the network (Fig. 8). The correlation (r) for T_n and T_x exceeded 0.80 at all stations, indicating strong temporal agreement between the station observations and the reanalysis. Tubu was the only station with r below 0.85 and had the lowest correlation for both T_n and T_x. It was also the only station where the correlations for T_x were lower than for T_n. At most other stations, Fig. 8 shows that T_x correlated more strongly with ERA5-Land than T_n.

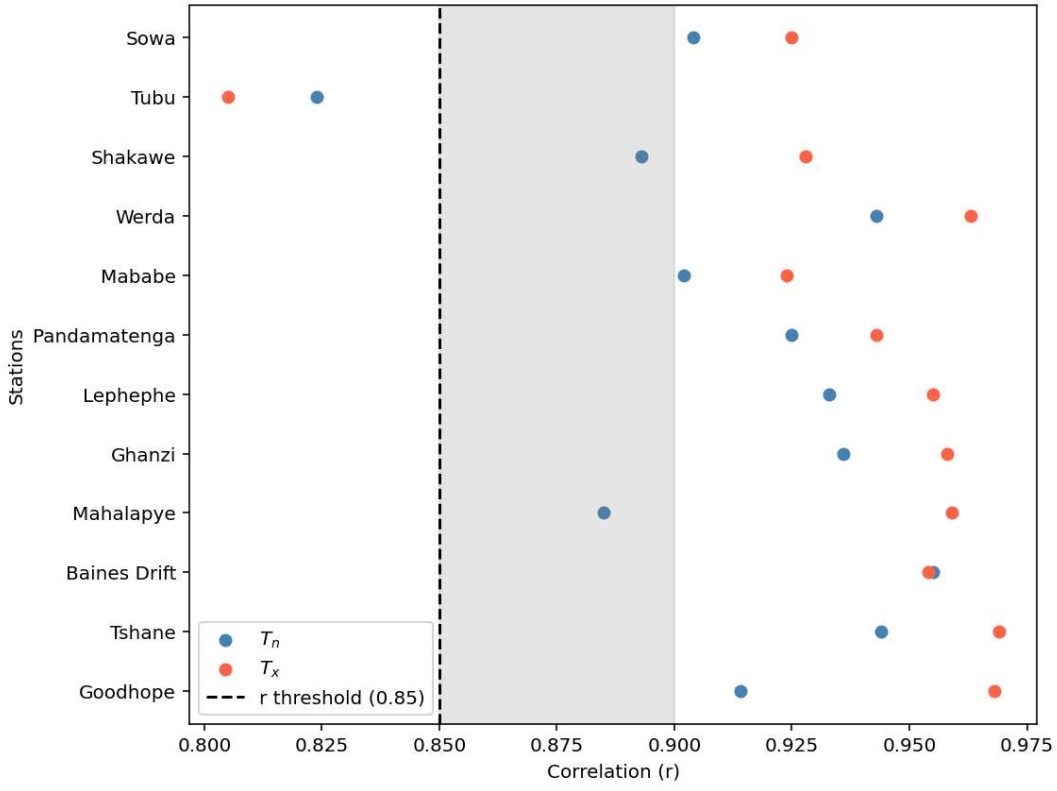


Figure 8: Correlations (r) between AWS and ERA5-Land for T_n and T_x.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

The bias in T_n ranged from -3.71 ºC to +1.42 ºC, and the bias in T_x from +0.35 ºC to +4.37 ºC, indicating systematic differences among stations (Fig. 9). The predominantly negative T_n biases mean that AWS night-time temperatures were generally lower than ERA5-Land at several stations. The positive T_x biases mean that AWS daytime maxima were higher than ERA5-Land. These differences may reflect how point observations and gridded reanalysis fields capture diurnal extremes as well as differences in elevation and land-surface representation, sensor exposure, and treatment of daily extremes. The biases nevertheless indicate partial agreement between AWS and ERA5-Land temperature.

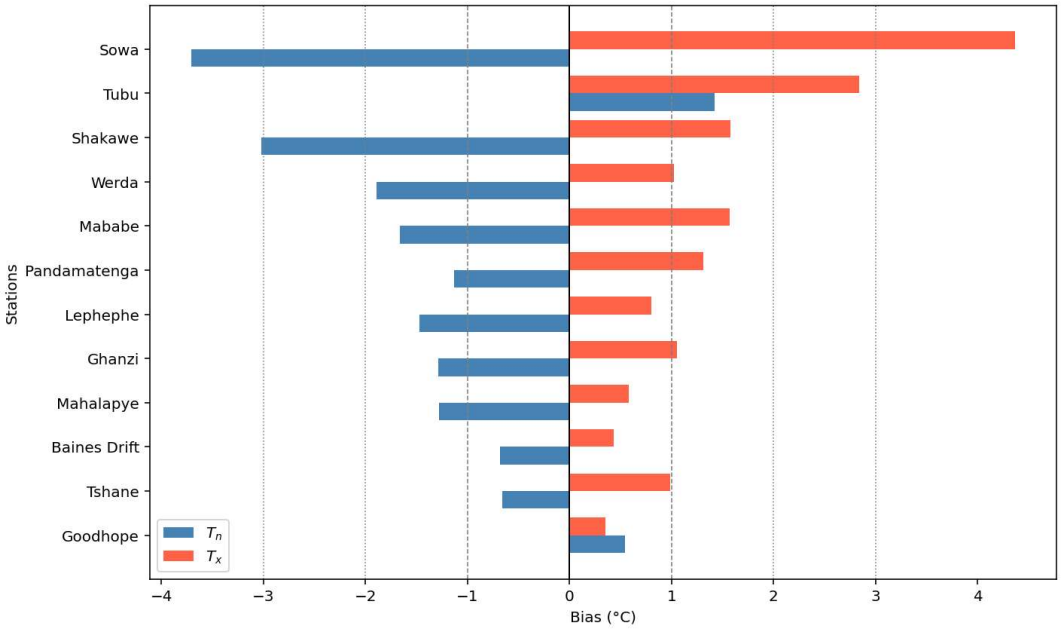


Figure 9: Bias (AWS minus ERA5-Land) for T_n and T_x.

The standard deviation (SD) of the AWS and ERA5-Land differences ranged from 0.9 ºC to 2.5 ºC for T_n and T_x, and the root mean square errors (RMSE) from 1.4 ºC to 4.7 ºC (Fig. 10). This indicates moderate but variable agreement among stations. Stations with larger departures warrant further quality assessment using station sensor history, exposure, and seasonal bias information.

Flags from the differences between AWS (T_n and T_x) and ERA5-Land (T_n and T_x) using thresholds of three standard deviations from the mean for each station are shown as percentage violations in Fig. 11. Most stations had exceedance rates below 10 %, indicating good agreement with ERA5-Land. Only Sowa had a higher violation rate, 33.64 % for T_x, a large difference.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

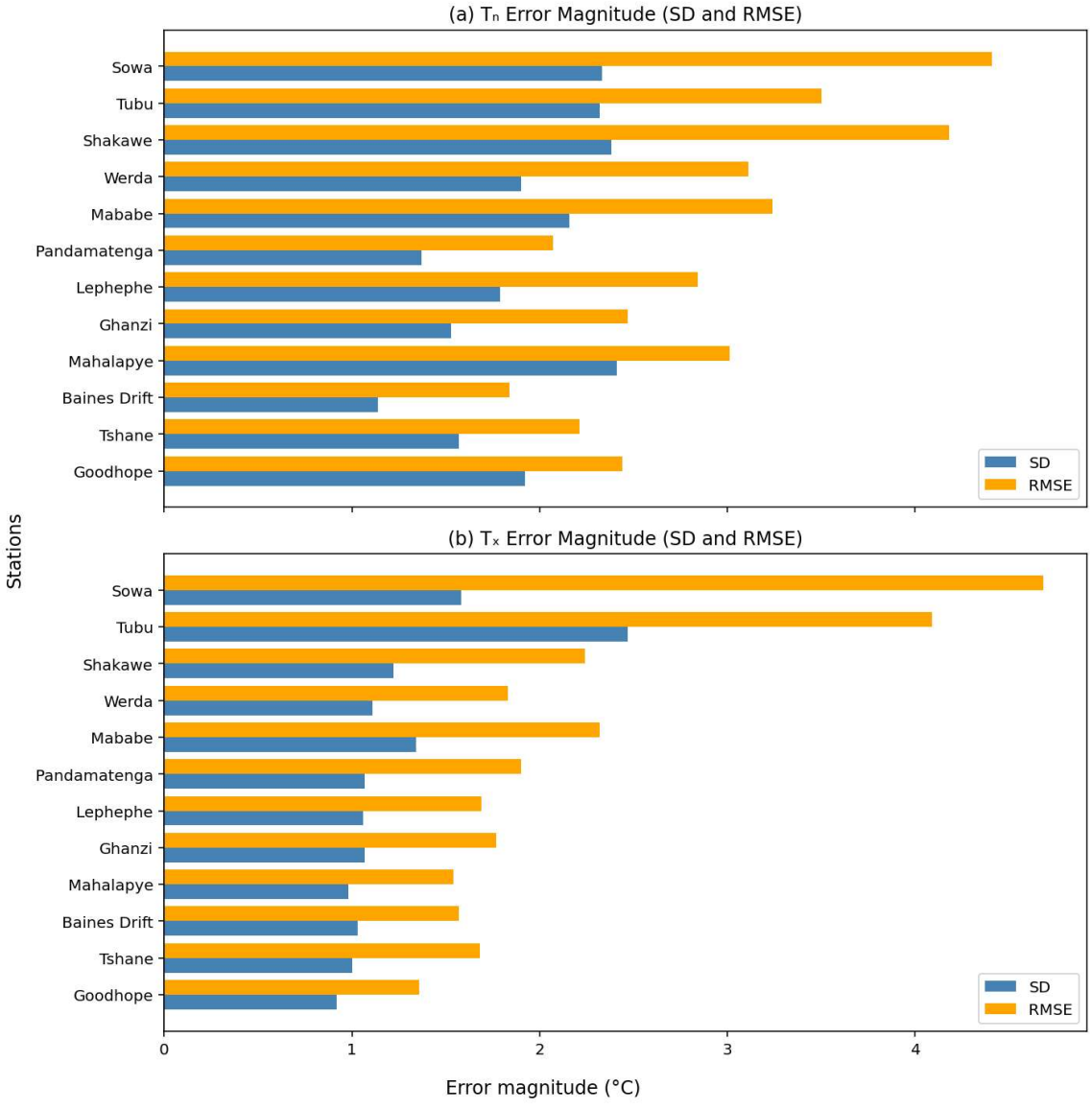


Figure 10: SD and RMSE of the AWS- ERA5-Land difference;(a) T_n, (b) T_x.

Three stations deviated from the network norm, namely Sowa, Tubu and Shakawe. Sowa had the largest biases in the network with a T_n bias of -3.71 ºC and T_x bias of 4.37 ºC and the highest flagged rate (33.64 % for T_x). This pattern could suggest sensor calibration drift rather than an environmental offset, given its elevated RH violation rate of 3.96 %. Tubu had

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

the weakest correlations in the network (T_n: 0.824, T_x: 0.805) and consistent positive biases (T_n: 1.42 °C, T_x: 2.84 °C). This matches Tubu's temperature sensor issues, its 10.61 % missing data and its elevated temporal spikes. Shakawe had a T_n bias of -3.02 °C and 7.80 % flagged observations. Its T_n bias was the second largest after Sowa and its T_x comparison stayed within typical ranges (bias 1.58 °C, r 0.928).

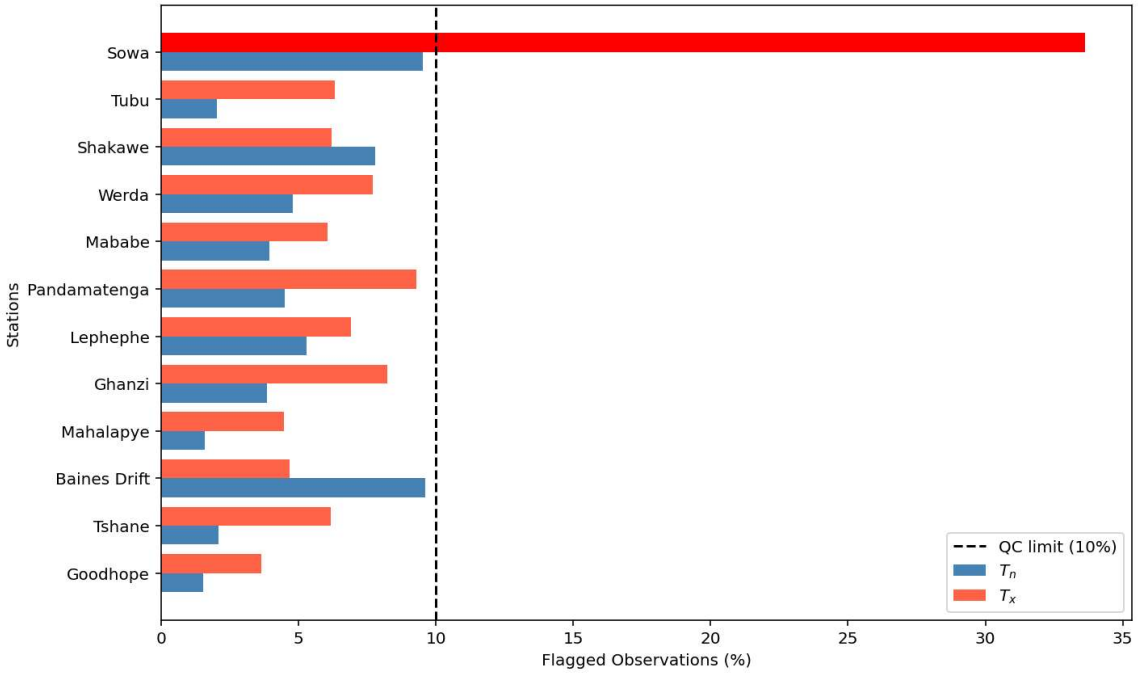


Figure 11: Percentage observations flagged from the difference between AWS and ERA5-Land.

In contrast to temperature, the correlation between AWS and CHIRPS precipitation ranged from 0.35-0.60 (Fig. 12). Correlations exceeded 0.40 at 75 % of the stations. This implies that AWS and CHIRPS capture some of the same rainfall patterns. At Ghanzi, Shakawe and Tshane (25 % of the stations), however, r was between 0.37 and 0.40. These correlations are weak and imply that the two datasets track actual rainfall together only marginally. Because most stations show intermediate correlations, precipitation can be considered to have passed the correlation test, although further comparison of the two datasets is needed for validation.

Figure 13 shows the bias between AWS and CHIRPS precipitation. Station biases were small and centred near zero, from -0.59 to 0.10 mm. Most stations had slightly negative biases, meaning that AWS rainfall was lower than the CHIRPS estimate. The largest negative bias was at Werda, followed by Shakawe, and only four stations had a small positive bias. Despite these differences, no station showed unusually large deviations from CHIRPS.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

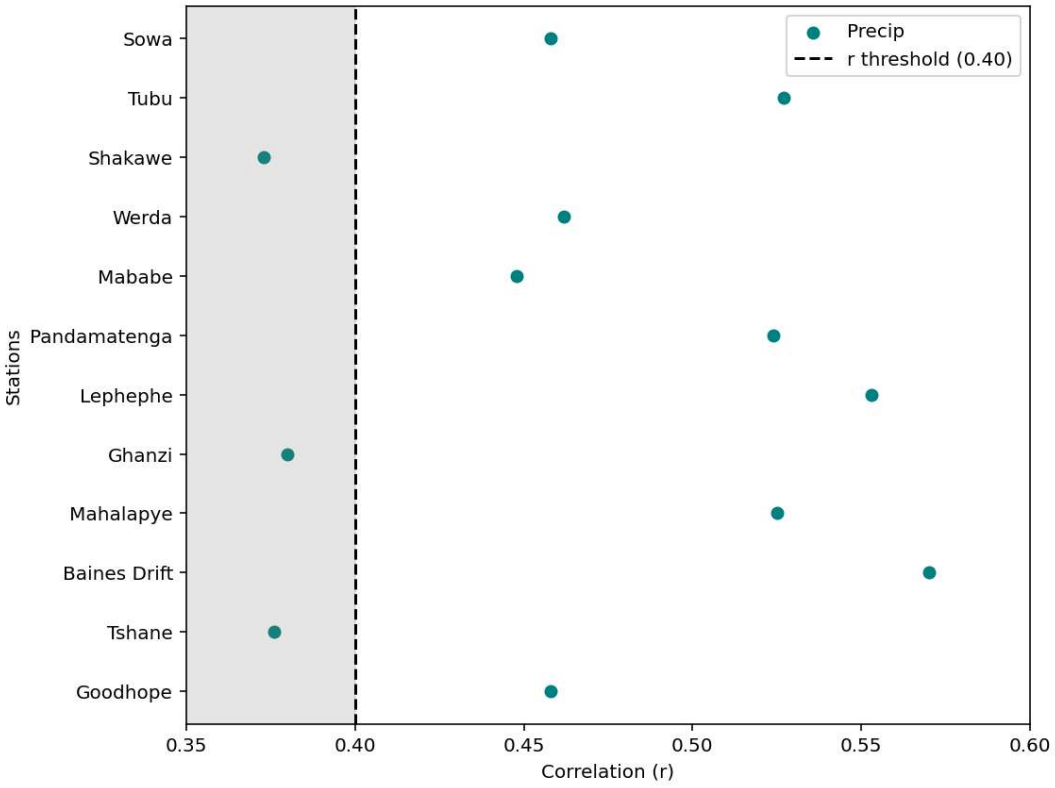


Figure 12: Correlations (r) between AWS and CHIRPS precipitation.

The RMSE ranged between 3.5 to 7.0 mm and the SD from 3.6 to 6.5 mm, (Fig. 14). The stations with higher RMSE also tended to have higher variability, reflecting larger day-to-day differences between the AWS and CHIRPS. The RMSE values indicate differences within a reasonable range for daily comparisons in semi-arid regions. The RMSE and SD results support the consistency of the AWS Precip observations and do not reveal widespread anomalous behaviour in the network.

Figure 15 shows the percentage of observations flagged from the AWS-CHIRPS differences, using a threshold of three deviations from the mean. Flagged percentages were low at all stations, from 1.1 % (Mababe) to 2.7 % (Werda), which implies that AWS precipitation is consistent with CHIRPS.

Shakawe had the highest precipitation variability (SD of 6.52 mm and RMSE of 6.66 mm) but small bias. This could reflect actual Okavango rainfall variability rather than gauge malfunction. There was no station that showed systematic rainfall measurement biases. These metrics are typical for point-to-grid rainfall comparisons in semi-arid regions (Kimani et al., 2017). Overall, AWS T_x, T_n, and Precip show consistency with ERA5-Land T_x and T_n and CHIRPS Precip, and are suitable

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

for further analysis within the expected uncertainty bounds of the region and datasets. For QC purposes, CHIRPS was used as an independent reference.

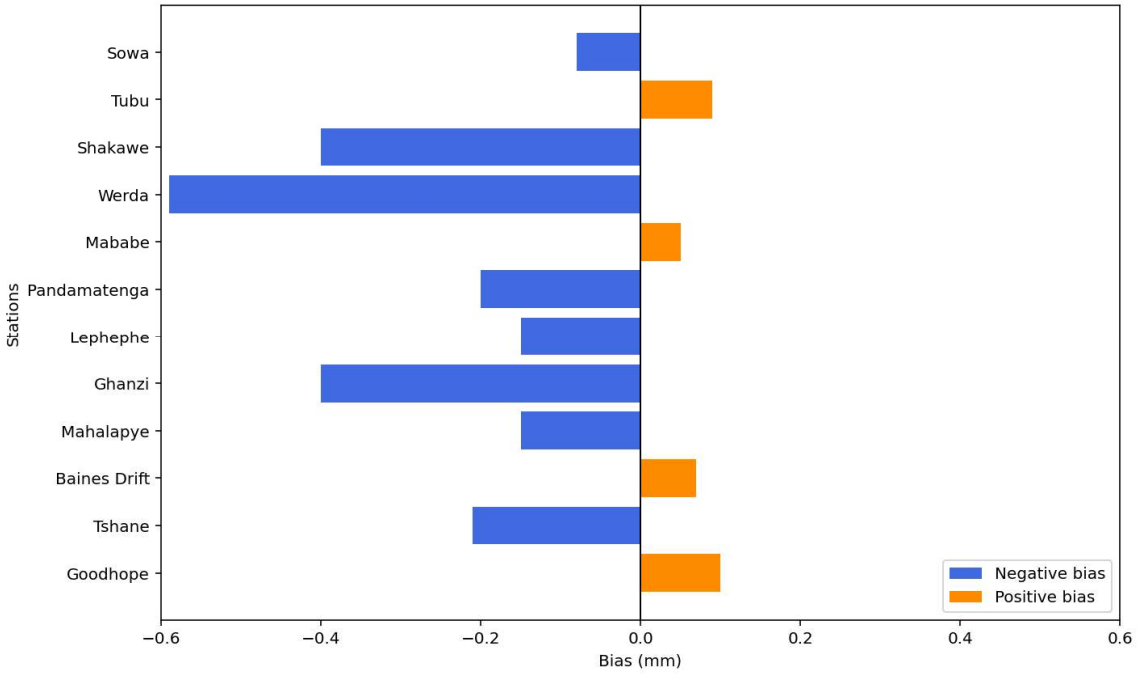


Figure 13: Bias (AWS minus CHIRPS) for precipitation.

4.6 Data retention after quality control

Data retention after quality control was high across the network. Table 3 shows that most stations retained more than 99 % of their observations of T_n, T_x, T_a, Precip, W_m and RH. Ghanzi and Werda showed near-complete retention for all variables, which reflects strong data continuity and minimal quality issues. Data retention was lower only at Pandamatenga, at approximately 91 % across all variables, which indicates a more persistent data quality problem at that station that needs extra attention. Nevertheless, the high retention rates show that the AWS data are of excellent quality.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

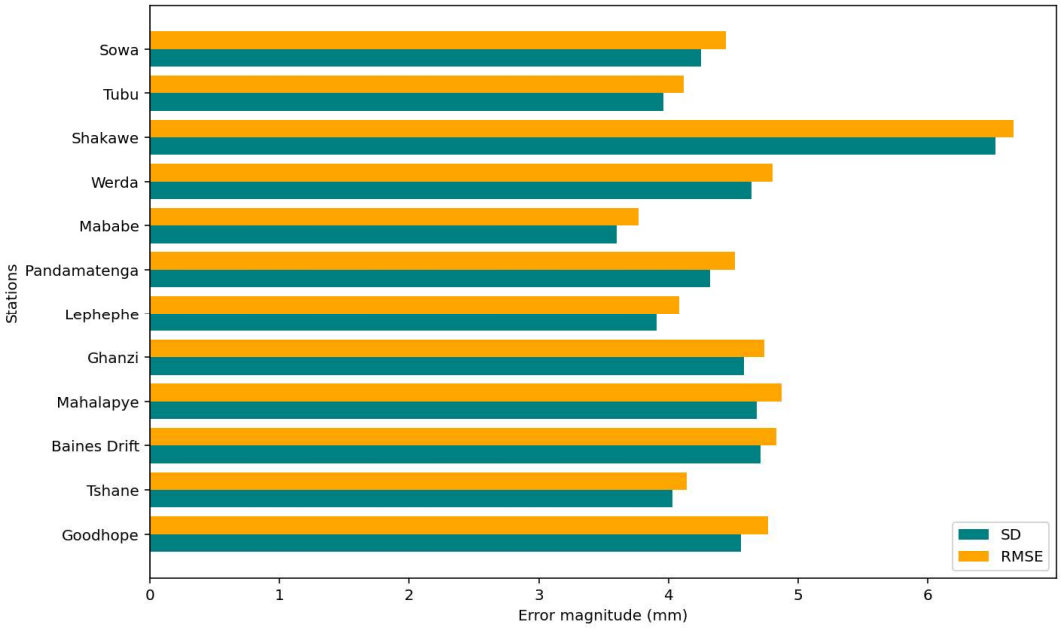


Figure 14: SD and RMSE between AWS and CHIRPS precipitation.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

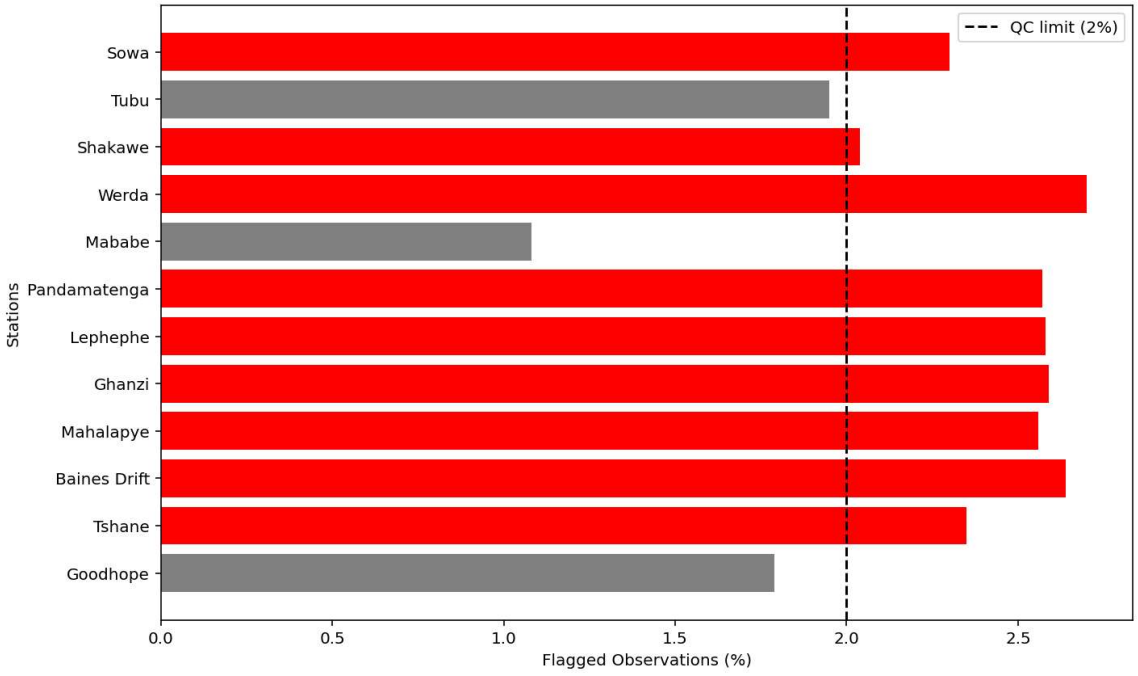


Figure 15: Percentage observations flagged from the difference between AWS and CHIRPS.

Table 3: Percentage data retention after quality control for the variables per station

Station	T_n	T_x	T_a	Precip	W_m	RH
Baines Drift	99.97	100	100	100	99.35	100
Ghanzi	100	100	100	100	100	100
Goodhope	99.92	100	100	100	100	100
Lephephe	100	100	100	100	100	100
Mababe	99.78	100	100	100	99.97	100
Mahalapye	99.73	100	100	100	99.95	100
Pandamatenga	91.00	91.00	91.00	91.81	91.76	91.73
Shakawe	100	100	100	100	100	100
Sowa	100	100	100	100	99.95	100
Tshane	99.97	100	100	100	99.97	100
Tubu	99.84	100	100	100	100	100
Werda	100	100	100	100	99.97	100

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

5 Conclusion

This study presented a five-stage QCF for SASSCAL WeatherNet AWS in Botswana. It addresses a serious gap in operational AWS data validation for semi-arid environments. The framework was applied to twelve AWS with records spanning 2014/15-2024. It detected data quality issues, distinguished natural climate variability from sensor errors, and identified the stations that need maintenance intervention, notably Mababe and Pandamatenga.

Overall, the network demonstrated robust performance, achieving a mean uptime of 96.04 %. All the stations passed the range checks in full, while climatic outlier rates were low (0.9-1.7 %) across all temperature variables. Violations of internal and temporal consistency checks were rare. The main exception was the daily mean temperature check, which flagged 13-26 % of days at all stations. External consistency correlations for temperature exceeded 0.80 at all stations, indicating that the data is of high quality and suitable for climate research, agricultural planning, and hydrological applications.

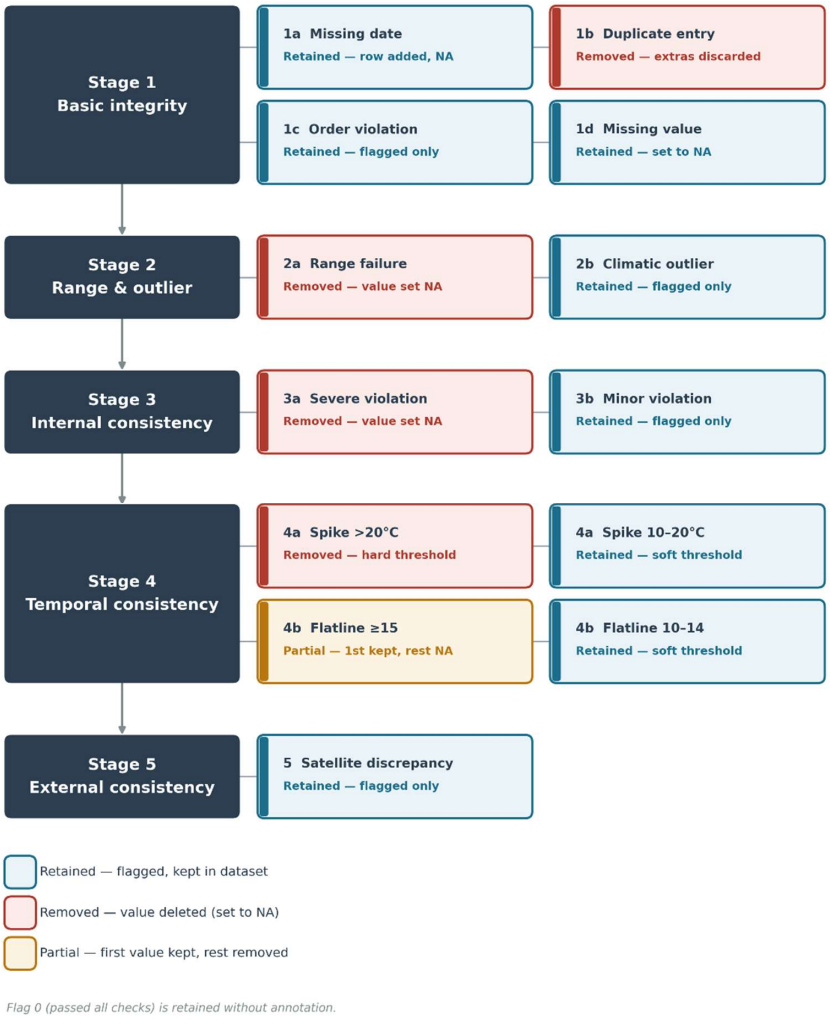
Despite this strong overall performance, station and seasonal variability was evident. Three distinct performance tiers were detected: seven stations recorded uptime of at least 98 % with minimal data gaps (Tier 1); three stations achieved 90–98 % uptime but showed seasonal vulnerabilities (Tier 2); and two stations fell below 90 % uptime and experienced chronic reliability issues (Tier 3). The outage analysis revealed three main failure modes, catastrophic single-event failures, chronic instability, and frequent short interruptions, which may be linked to communication issues. Temperature sensors emerged as the most vulnerable component, while seasonal patterns suggest that some failures may be environmentally driven.

External consistency comparisons also revealed station-specific concerns. Sowa showed large opposing temperature biases and an exceptionally high flag rate, suggesting possible sensor calibration drift that requires immediate attention. Tubu recorded the weakest temperature correlations in the network, alongside sensor failures, and therefore needs verification against independent observations. In contrast, Shakawe showed stronger nighttime cooling relative to ERA5-Land, which may reflect local microclimate of the Okavango Delta rather than measurement error. Importantly, the external consistency results aligned with issues identified in earlier QC stages, such as Sowa's elevated RH violations and Tubu's temporal spikes, reinforcing the value of the multi-stage approach.

The framework effectively separated operational failures from environmental stressors and supported evidence-based maintenance prioritization. The targeted interventions, such as flood mitigation at Mababe, winter infrastructure upgrades at Lephephe, a comprehensive assessment at Pandamatenga, and sensor recalibration at Sowa are likely to be more effective than generic network-wide maintenance. Overall, AWS data that passed all five QC stages met acceptable performance criteria, given the high retention rates, and can be considered fit for climatological and hydrological applications. The proposed methodology therefore provides a practical and replicable template for operational AWS quality control in semi-arid environments.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix A



490

Figure A1: A five-stage quality control hierarchy applied to the SASSCAL WeatherNet automated weather station (AWS) data in Botswana.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Data availability

The SASSCAL WeatherNet automated weather stations data are publicly available via the SASSCAL online portal (https://sasscalweathernet.org/). The quality-controlled data has been uploaded to Zenodo (https://zenodo.org/records/22815337). The ERA5-Land hourly temperature data is available at Climate Data Store online portal (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview) and CHIRPS data v2 is available at the Climate Hazards Center online data portal (https://www.chc.ucsb.edu/data/chirps).

Author contributions

Thuso H. Tlhalerwa wrote the concept layout, performed the analysis, and wrote the original manuscript. Kago E. Maabong along with Thuso H. Tlhalerwa designed the methodology and data curation. Kgakgamatso M. Mphale and Douglas Letsholathebe oversaw the study and provided valuable contributions at various points. All coauthors contributed to the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank SASSCAL for providing the AWS and making the dataset freely available.

Financial support

This research did not receive any funding.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Batisani, N., & Yarnal, B.: Rainfall variability and trends in semi-arid Botswana: Implications for climate change adaptation policy. Applied Geography, 30(4), 483–489. https://doi.org/10.1016/j.apgeog.2009.10.007, 2010.

Brunet, M., Brugnara, Y., Noone, S., Stephens, A., Valente, M. A., Ventura, C., Jones, P., Gilabert, A., Brönnimann, S., Luterbacher, J., Allan, R., Brohan, P., & Compo, G. P.: Best Practice Guidelines for Climate Data and Metadata Formatting, Quality Control and Submission Best Practice Guidelines for Climate Data Rescue. https://doi.org/10.24381/kctk-8j22, 2020.

Degaetano, A. T.: A Quality-Control Routine for Hourly Wind Observations. Journal of Atmospheric and Oceanic Technology, 14, 308–317. https://doi.org/10.1175/1520-0426(1997)014<0308:AQCRFH>2.0.CO;2, 1997.

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., & Disse, M.: Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. Science of the Total Environment, 573, 1536–1553. https://doi.org/10.1016/j.scitotenv.2016.08.213, 2016.

Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G., & Vose, R. S.: Comprehensive automated quality assurance of daily surface observations. Journal of Applied Meteorology and Climatology, 49(8), 1615–1633. https://doi.org/10.1175/2010JAMC2375.1, 2010.

Estévez, J., Gavilan, P., & Giraldez, J. V.: Guidelines on validation procedures for meteorological data from automatic weather stations. Journal of Hydrology, 144–154. https://doi.org/10.1016/j.jhydrol.2011.02.031, 2011.

Faybishenko, B., Versteeg, R., Pastorello, G., Dwivedi, D., Varadharajan, C., & Agarwal, D.: Challenging problems of quality assurance and quality control (QA/QC) of meteorological time series data. Stochastic Environmental Research and Risk Assessment, 36(4), 1049–1062. https://doi.org/10.1007/s00477-021-02106-w, 2022.

Fiebrich, C. A., Morgan, Y. R., Mccombs, A. G., Hall, P. K., & Mcpherson, R. A.: Quality assurance procedures for mesoscale meteorological data. Journal of Atmospheric and Oceanic Technology, 27(10), 1565–1582. https://doi.org/10.1175/2010JTECHA1433.1, 2010.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J.: The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. Scientific Data, 2. https://doi.org/10.1038/sdata.2015.66, 2015.

Hubbard, K. G., Goddard, S., Sorensen, W. D., Wells, N., & Osugi, T. T.: Performance of Quality Assurance Procedures for an Applied Climate Information System. Journal of Atmospheric and Oceanic Technology, 22, 105–112. https://doi.org/https://doi.org/10.1175/JTECH-1657.1, 2005.

Juana, J. S., Okurut, F. N., Makepe, P. M., & Kahaka, Z.: Climate change perceptions and adaptations for livestock farmers in Botswana. International Journal of Economic Issues, 9(1), 1–21. https://doi.org/10.5539/jas.v5n4p121, 2016.

Kaspar, F., Helmschrot, J., Mhanda, A., Butale, M., de Clercq, W., Kanyanga, J. K., Neto, F. O. S., Kruger, S., Castro Matsheka, M., Muche, G., Hillmann, T., Josenhans, K., Posada, R., Riede, J., Seely, M., Ribeiro, C., Kenabatho, P., Vogt,

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

R., & Jürgens, N.: The SASSCAL contribution to climate observation, climate data management and data rescue in Southern Africa. Advances in Science and Research, 12(1), 171–177. https://doi.org/10.5194/asr-12-171-2015, 2015.

Kimani, M. W., Hoedjes, J. C. B., & Su, Z.: An assessment of satellite-derived rainfall products relative to ground observations over East Africa. Remote Sensing, 9(5). https://doi.org/10.3390/rs9050430, 2017.

Lawrence, M. G.: The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. Bulletin of the American Meteorological Society, 86(2), 225–233. https://doi.org/10.1175/BAMS-86-2-225, 2005.

Matenge, R. G., Parida, B. P., Letshwenyo, M. W., Ditalelo, G.: Impact of Climate Variability on Rainfall Characteristics in the Semi-Arid Shashe Catchment (Botswana) from 1981–2050. Earth, 4(2), 398-441. https://doi.org/10.3390/earth4020022, 2023.

Molosiwa, R., Mphale, K. M., Matlhodi, B., & Mbangiwa, N. C.: Spatiotemporal characterization of drought in Botswana: a regionalization study using long-term chirps satellite data. Theoretical and Applied Climatology, 157(4). https://doi.org/10.1007/s00704-026-06154-6, 2026.

Molosiwa, R., Mphale, K., & Mbangiwa, N.: Validation of Satellite Rainfall Estimates against Rain Gauge Data for Enhanced Drought Monitoring and Early Warning Systems in Botswana. In Preprint. https://doi.org/10.21203/rs.3.rs-6196056/v1, 2025.

Moses, O.: Weather systems influencing Botswana rainfall: the of 9 December 2018 storm in Mahalapye, Botswana. Modeling Earth Systems and Environment, 5, 1473-1480. https://doi.org/10.1007/s40808-019-00640-9, 2019.

Mphale, K. M., Dash, S. K., Adedoyin, A., & Panda, S. K.: Rainfall regime changes and trends in Botswana Kalahari Transect's late summer precipitation. Theoretical and Applied Climatology, 116(1–2), 75–91. https://doi.org/10.1007/s00704-013-0907-z, 2014.

Muche, G., Kruger, S., Hillmann, T., Josenhans, K., Ribeiro, C., Bazibi, M., Seely, M., Nkonde, E., de Clercq, W., Strohbach, B., Kenabatho, P., Vogt, R., Kaspar, F., Helmschrot, J., & Jürgens, N.: SASSCAL WeatherNet: present state, challenges, and achievements of the regional climatic observation network and database. Biodiversity & Ecology, 6, 34–43. https://doi.org/10.7809/b-e.00302, 2018.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. In Copernicus Climate Change Service (C3S). https://doi.org/10.5194/essd-2021-82, 2021.

New, M., Hewitson, B., Stephenson, D. B., Tsiga, A., Kruger, A., Manhique, A., Gomez, B., Coelho, C. A. S., Masisi, D. N., Kululanga, E., Mbambalala, E., Adesina, F., Saleh, H., Kanyanga, J., Adosi, J., Bulane, L., Fortunata, L., Mdoka, M. L., & Lajoie, R.: Evidence of trends in daily climate extremes over southern and west Africa. Journal of Geophysical Research Atmospheres, 111(14). https://doi.org/10.1029/2005JD006289, 2006.

https://doi.org/10.5194/egusphere-2026-2862
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Nkoni, G., Mphale, K., Mbangiwa, N., Samuel, S., & Molosiwa, R.: Use of multivariate techniques to regionalize rainfall patterns in semiarid Botswana. Discover Environment, 2(1). https://doi.org/10.1007/s44274-024-00111-8, 2024.

NOAA National Centers for Environmental Information.: ETOPO 2022 15 Arc-Second Global Relief Model. https://doi.org/https://doi.org/10.25921/fd45-gt74, 2022.

Patro, B. S., & Bartakke, P. P.: Quality Control (QC) and Quality Assurance (QA) Procedures for Meteorological Data from Automatic Weather Stations. 2025 4th International Conference on Range Technology, ICORT 2025. https://doi.org/10.1109/ICORT64008.2025.11115447, 2025.

SASSCAL WeatherNet.: SASSCAL WeatherNet. SASSCAL. https://www.sasscalweathernet.org, last access 6 January 2026.

Shafer, M. A., Fiebrich, C. A., Arndt, D. S., Fredrickson, S. E., & Hughes, T. W.: Quality Assurance Procedures in the Oklahoma Mesonetwork. Journal of Atmospheric and Oceanic Technology, 17(4), 474–494. https://doi.org/10.1175/1520-0426(2000)017<0474:qapito>2.0.co;2, 2000.

Vaisala.: HMP155 Humidity and Temperature Probe. www.vaisala.com, 2012.

World Meteorological Organization (WMO).: Guide to Meteorological Instruments and Methods of Observation (WMO-No. 8). www.wmo.int, (2018).

World Meteorological Organization (WMO). Guide to Climatological Practices (WMO-No. 100). www.wmo.int, 2011.

Zahumenský, I.: Guidelines on Quality Control Procedures for Data from Automatic Weather Stations, World Meteorological Organization, Geneva, 2004.