Reply to Referee 1

December 19, 2025

Dear Referee,

Thank you for reviewing our manuscript and for this constructive and helpful suggestion. Your comments are shown in blue, followed by our responses in black.

General comments:

I appreciate the work conducted by the author team to update the manuscript in response to the provided comments.

In general, I believe that the team has done a good job with this. One point that would still improve the manuscript, in my opinion:

1) For the data QC, the authors now state that "Data were marked for exclusion when any of the following conditions held: (i) non-physical or out-of-range values for the measured quantity, (ii) packet corruption, (iii) timestamp discontinuities or non-monotonic sequences, and (iv) sensor dropouts or communication faults recorded by the logger.". It would still be nice if more detail was included. For example, what is "non-physical or out-of-range" for each variable measured? State this explicitly. How often did you encounter packet corruption? Timestamp discontinuities? Sensor dropouts? Provide percentages, relative to the total length of the dataset.

Thank you for this suggestion. To address the request for more quality-flag detail, we implemented and documented variable-specific QC directly in the released NetCDF files.

"Out-of-range" limits for temperature, relative humidity, and pressure (T/RH/P) were defined and explicitly stated in the attributes for each quantity. Furthermore, we added a quality-flag variable for each of these quantities for each sensor. Metadata also includes the percentage of data excluded due to out-of-range values. The quality flag is encoded as: 0 = valid sample, 1 = invalid due to parsing failure, or communication fault and 2 = invalid due to an "out-of-range value".

All samples with quality flags 1 or 2 are stored as fill values (NaN) in the data arrays. The NetCDF metadata reports the number of valid samples and the total number of samples (valid + NaN).

Beyond quality flags, the NetCDF metadata also includes standard dataset descriptors and links to the relevant sensor documentation/datasheets. We have added additional text in the manuscript that points readers to the NetCDF metadata fields for the quality-flag definitions and percentages, since those are provided as part of the published dataset.

Were there specific conditions that resulted in more or less of this behavior? This manuscript should help the reader understand the basic facts of the dataset as it is published, and this information would be helpful.

We did not observe a systematic dependence on environmental or operational conditions. The flagged periods were associated with occasional data logger "jams" and sensor-inherent behaviour. These events occurred sporadically and were handled via the QC procedures described above.

To the manuscript we added:

For WinDarts, explicit out-of-range limits for temperature, relative humidity, and pressure are stated in the NetCDF attributes for each quantity, together with per-sensor quality-flag variables denoted as, *varaiable_name*_qc_flag, where 0 indicates valid data, 1 denotes a parsing, timestamp, or communication fault, and 2 indicates values outside the specified valid range. The variable attributes also include the percentage of data excluded due to out-of-range values, i.e. quality-flag 2, which for most cases is 0%. For all variables, samples affected by parsing failures, timestamp issues, or sensor dropouts/communication faults are stored as NaN. The attributes of the variables include the number of valid samples and the total sample count. For WinDart data, excluding GPS-related time series, the percentage of valid data is 100% for all quantities. For ground-weather-station data, again excluding GPS/GNSS-related time series, the percentage of valid data is >99% for all time series. The GPS/GNSS-related time series are parsed from NMEA 0183 sentences, in which null fields may occur and indicate invalid data or that no data are available at a given epoch. During conversion to NetCDF, we have replaced null fields with fill values (i.e. NaN) and are counted as 'invalid' data points, which explains the large number of invalid data points in GPS/GNSS-related time series. We did not observe a systematic dependence on environmental or operational conditions for parsing failures, timestamp issues, or sensor dropouts/communication faults. Possible systematic errors can be identified owing to the WinDarts' redundant measurement design, as illustrated and discussed in the examples below.