Weather Data Quality Assurance and Control (QA/QC) Summary

Manitoba Metis Federation and the Centre for Earth Observation Science – Dawson Bay Li Taan Aen Staansyoon 51490

UM | Centre for Earth Observation Science

Table of Contents

Document Control

Version History

Version	Author(s)	Type	Date Modified	Comments
1.0	Friesen, K. L.	Working Copy	2022/06/14	Working copy
1.1	Friesen, K. K.	Working Copy	2023/04/06	Updated QA/QC

Document Location

A digital copy of the document can be found in the Manitoba Métis Federation (MMF) repository on Gitlab. This repository is accessible by the MMF and its designees.

Link: https://canwin-datahub.ad.umanitoba.ca/data/dataset/dawbay-metdata

License

Manitoba Métis Federation(2023). Weather Data QA/QC Summary, Version 1.1. Manitoba Métis Federation and Centre for Earth Observation Science. University of Manitoba.

QA Summary for Weather Data

This document is meant to inform data users of the data cleaning steps to treat possible errors or faulty data measurements collected from the Weather Stations in the Weather Keeper Program managed; owned and maintained by the MMF. Weather Keepers in this program are specifically MMF Red River Métis citizens who maintain the weather stations on their personal or community property.

Data cleaning for these stations does not remove or delete any lines of data, but associates a metadata field beside each variable, titled "<variable name>_result_value_qualifier" with a code to identify error prone data or faulty sensors.

QA Assessment

Analyzed Weather Data

Weather data analyzed in this document pertains to the St Laurent Li Taan Aen Stansyoon weather station and the versions of the data are provided in Table 1 along with comments on changes.

Table 1. Current Archived Datasets on CanWIN's DataHub.

Filename	IMEI	Comment
DawsonBay_historical_2023-04-04	300534061454190	Curating archived data.
DawsonBay_historical_2023-04-19	300534061454190	Converted NA values to white spaces.

Changes and Corrections

The following section describes in general terms the changes of corrections made to the datasets listed in Table 2. At no point are data deleted from the datasets, only values are removed if there was a degree of uncertainty with the observed measurement. For example, due to the weather station project starting in 2021 rain measurements during seasonal changes and wintering months have been removed due to the inability to discern whether the observation is from rain, snow, or snow melt. Further description is provided below.

Step 1: Start time

The start time of the data set is formatted to the ISO8601 Coordinated Universal Time (UTC), YYYY-MM-DD hh:mm:ss.

Step 2: Standardize variable names

Each of the variable names are standardized from their raw data format using Climate Forecast (CF), or if no name is found the British Oceanographic Data Centre (BODC) controlled vocabularies (Table 2). You can review the controlled terms in the Weather Data Cookbook and Codebook located on CanWIN's DataHub.

Step 3: Adding result value qualifier

To identify problematic or errors in weather measurements, a result value qualifier field is associated with each variable and is positioned to the right of each respective variable column/field. If a value is above the highest value a sensor can measure accurately, the result value qualifier field will indicate that the measurement is Above the Detection Limit (ADL). If it is below the lowest value a sensor can measure, the field will indicate the measurement is Below the Detection Limit (BDL). Limits for each variable are shown in Table 2.

Table 2. Weather Parameter Detection Limits.

Standardized Variable	Above Detection Limit	Below Detection Limit	Units
Air pressure	>1070	<660	Millibar (mbar)
Photosynthetically Active Radiation (PAR)	>2500	<0	MicroEinsteins (µE)
Air temperature	>75	<-40	Celsius (°C)
Relative humidity	>100	<0	Percent (%)
Precipitation	>127	<0	Millimeters (mm)
Wind speed	>50	<0	Kilometers per hour (Km/h)
Gust speed	>50	<0	Kilometers per hour (Km/h)

Sensor calibration information and instrument details can be found on the Dawson Bay Li Taan Aen Staansyoon Instrument Details: https://canwin-datahub.ad.umanitoba.ca/data/instrument_details/dawbay-met-sensors.

If measurements are known to be impacted by an environment change or sensor failure, then a result value qualifier field of “prob_bad” is applied. The affected variables are listed in Table 3.

Table 3. Sensors with probably bad result value qualifier.

Standardized Variable	Reason	Time Range
Wind speed	Gust speed is the same speed.	Whenever phenomena are observed.
Gust speed	Wind speed is the same speed.	Whenever phenomena are observed.
Wind direction	Direction is greater than 355° and less than 360°.	Whenever phenomena are observed.
All variables except Battery output	Battery failed as well as backup batteries (noted in “FEF” section), hence all other variables are recording in non-optimal conditions.	2022-10-06 to 2022-10-10

When sensors experience failures or are below optimal ranges a result value qualifier field of “FEF” is applied, indicating Field Equipment Failed. If no data is observed when the sensor if failed, then null values are represented by “NA”. If a sensor failure impacts multiple sensors, the field “FEF” is given to the sensor that failed and the measurements from the other sensors are

marked with the probably bad code “prob_bad”, within the specific range of measurements that the failure occurred. Affected variables are listed in Table 4.

Table 4. Sensors with field equipment failure result value qualifier.

Standardized Variable	Reason	Time Range
Wind speed, Gust speed, and Wind direction	Battery failure causes issues in measurements	2022-10-05 to 2022-10-10
Battery output	Battery and backup battery failure	2022-10-06- to 2022-10-10

Lastly, the code “NC” identifies measurements that have been not collected or have been removed due to uncertainty in accuracy. Uncertain measurements, thus far, only have been recorded for rain measurements during the seasonal changes and wintering months. The weather keeper program started in 2021 with rain buckets potentially being exposed to the elements during seasonal changes and wintering months. Hence, the data produce is unable to be discerned from rain, snow or snow melt. In the future we hope to mitigate this by covering the rain bucket during winter and giving weather keepers explicit instructions to cover the rain bucket before the first snowfall in fall and after the last snowfall in spring. The variables that utilize NC are listed in Table 6 below.

Table 5. Sensors that use not collected result value qualifier.

Standardized Variable	Reason	Time Range
Precip	Applied to months of October to March. Unable to confirm if rain measurement is rain, snow, or snow melt.	20XX-10-01 to 20XX-04-31

Comments

Currently data shown on our near-real time dashboards have not undergone any data QA QC measures. Corrections and changes have only been applied to archive datasets from Weather Stations that are available on the CanWIN DataHub (links above).

QC Measures

It is recommended by the Centre of Earth Observation (CEOS) that MMF weather keepers cover the rain gauges during the winter months to minimize the introduction of artificial data. Additionally, maintenance checks are required by weather station operators (Weather Keepers) to ensure that sensors are not producing inaccurate data due to snow, dust, extreme moisture, or damage from environmental factors.

To data users, please review the entire document so you are fully informed on the state of the dataset. In addition, each site has specific biases that are outside of the Weather Keepers control and should be noted when using these datasets (Appendix A).

Dawson Bay Li Taan Aen Staansyoon Data Bias

Wind and gust speeds reported at Dawson Bay Li Taan Aen Staansyoon indicate speeds that would be felt nearby on the lake only when the wind blows from SE. The anemometer is sheltered from the full force of the wind by trees and a building in other directions. Wind and gust wind speeds are measured 3 m above the ground surface. Environment Canada and Manitoba Department of Agriculture weather stations report faster speeds because they are recorded on higher, 10 m towers

Weather Station Terms of Access and Terms of Use

Each site has a specific Terms of Access that includes data bias descriptions. Please review these terms with the links provided below.

By accessing this data you agree to CanWIN's Terms of Use.

Map and Dashboard of Stations

Map of weather stations can be accessed here:
https://canwinmap.ad.umanitoba.ca/dashboards/weather-stations/

Near-real time dashboards of weather data.

St Laurent Li Taan Aen Staansyoon: https://geoconnections.ad.umanitoba.ca/d/JbY1plZ7z/st-laurent-mb-weather-dashboard?orgId=2&refresh=1m

Dawson Bay Li Taan Aen Staansyoon:
https://geoconnections.ad.umanitoba.ca/d/HX66kVc7k/dawson-bay-mb-weather-dashboard?orgId=2&refresh=30s

References

1. Government of Canada. (2023, March 29). Daily Data Report for July 2022. https://climate.weather.gc.ca/climate_data/daily_data_e.html?hlyRange=%7C&dlyRange=1987-06-01%7C2022-07-03&mlyRange=1987-01-01%7C2007-02-01&StationID=1309&Prov=BC&urlExtension=_e.html&searchType=stnProv&optLimit=yearRange&StartYear=1840&EndYear=2022&selRowPerPage=25&Line=2&Month=7&Day=4&lstProvince=BC&timeframe=2&Year=2022

Appendix

A: Weather Data Caveats for Data Users

There are currently data errors within the near real-time data being streamed on the dashboards (links provided above). These data issues will be corrected for once data during this time period has been archived. These issues are expanded on in the Weather Station Common Errors manual available on the Weather Keeper Program page.

Column	Date Range	Caveat
All fields	Past to 2021-10-27	Testing sensor and not actual weather data at current location. Deployment date is also the end of specified date range.
Datetime	2022-03-13 to 2022-04-04	Datetime was set to correct time zone (UTC-6); however, after daylight savings measurements have been recorded two hours behind actual location time.
Rain	2021-10-XX to 2022-04-XX	Rain measurement removed because measurements cannot be distinguished from rain or snow melt.
Battery failure	2022-10-05 to 2022-10-10	Main battery likely failed and subsequent backup batteries in Onset logger failed. After replacing the backup batteries they failed again and the entire system was sent out for repairs.
Logger box and sensors removed	2022-10-28	Logger box and RH and PAR were removed from station location for repairs.

B: Environment and Climate Change Canada Climate Data Quality Codes

Table 6. ECCC Data Quality Codes.

Code	Definition
A	Accumulated
C	Precipitation occurred, amount uncertain
E	Estimated
F	Accumulated and estimated
L	Precipitation may or may not have occurred
M	Missing
N	Temperature missing but known to be > 0
S	More than one occurrence
T	Trace
Y	Temperature missing but known to be < 0
[empty]	Indicates an unobserved value
^	The value displayed is based on incomplete data
†	Data that is not subject to review by the National Climate Archives

C: Standardized Variable Names and Description

The following is a tabulated list of the terms used to describe weather data parameters in the archived file available on the CanWIN DataHub. These terms can also be found in the data dictionary for the resource.

Table 7. Weather Data Variable Names and Descriptions.

Column	Label	Description
Date_and_time	Datetime	String corresponding to format 'YYYY-MM-DDThh:mm:ss.sssZ' or other valid ISO8601 string.
air_pressure	Barometric Pressure	Measured using a barometer mounted inside the enclosure that houses the air temperature sensor. It is reported in millibars, where 1,000 millibars is the average air pressure at sea

		level. Changing pressure often indicates a coming change in the weather. Increasing pressure is associated with clearing skies. Decreasing pressure is associated with increasing cloudiness, and possibly an approaching storm.
air_pressure _result_value_qualifier	Pressure result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control

		that forms part of a feature that is probably inconsistent with real phenomena.
Photosynthetically _Active_Radiation	PAR	The intensity of the part of sunlight that plants can use to support new growth, and also the wavelengths of light that our eyes are sensitive to. It is measured on a small white disc that records the sum of light falling directly from the sun, and light scattered by the sky and clouds. It is reported in a unit called a micro-Einstein.
Photosynthetically_Active_Radiation_result_value_qualifier	PAR result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing

		problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
air_temperature	Temperature	We report temperature in the shade, in Centigrade degrees. It is measured using an electronic thermometer shielded from direct sunlight, mounted inside an enclosure with louvered walls to allow free air flow past the sensor.
air_temperature _result_value_qualifier	Temperature result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the

		sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
relative_humidity	RH	The amount of water vapour in the air reported as a percentage of the amount that would saturate it at the air temperature. Warmer air can hold more water vapour than cooler air.
relative_humidity _result_value_qualifier	RH result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is

		faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
Precip	Rain	Reported in millimetres accumulated every 15 minutes in a gauge set about 1 m above the ground. Rain falling into a 6-inch diameter funnel drips onto a one of a pair of “buckets”, mounted on a teeter-totter. When one bucket is filled, the teeter-totter flips, the first bucket is emptied and rain then drips into the other “bucket” until it is filled, and so on. Each flip is equals 0.2 mm of rain.
Precip _result_value_qualifier	Precipitation result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated

		detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
wind_speed	Wind speed	Measured using a spinning anemometer mounted on the wind vane at the top of the tower, about 3 m above the ground. We report the average speed for every 15 minute period, in meters per second.
wind _speed_result_value_qualifier	Wind speed result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated

		detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
wind_speed_of_gust	Gust Speed	The highest wind speed recorded in each 15 minute period, and are also recorded in kilometres per hour.
wind_speed_of_gust_result_value_qualifier	Gust speed result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and

		"BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
wind_from_direction	Wind direction	Measured using a wind vane mounted at the top of the tower. As we do with wind speed, we report the average direction for every 15 minutes, in compass degrees from true north.
wind_from_direction_result_value_qualifier	Wind direction result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and

		"BDL" is below the detection limit. Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
battery_output	Battery voltage	Power provided by an electric battery, in units of volts (V). Note, name is standardized as 'battery output'.
battery_output_result_value_qualifier	Battery output result value qualifier	A result value qualifier is a code field to identify measurements outside of the sensor calibrated detection range. "ADL" is a measurement above the detection limit and "BDL" is below the detection limit.

		Additionally, field equipment error code "FEF" is given when the sensor or station is faulty and producing problematic data along with probably bad data code "prob_bad" when data is recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.