AWS Data Correction Description

Federico Covi - 3/16/2020
Project: Firn Greenland

This file contains an accurate description of all the corrections applied to the AWS data. For each variable every step of the correction procedure is described stating what and why needs correction and how it is corrected.

For every AWS an html file (e.g. AWS_SiteJ_datacorrection.html) is provided. This file is a Python Jupyter Notebook that is used to correct the RAW data. It can be opened in any web browser and show all the details of the corrections applied for every site (e.g. some particular corrections are site specific). Furthermore it has interactive figures of corrected vs raw data as well as general statistics about the corrections (e.g. how many times one variable has been corrected).

Air Temperature

Sensor: Vaisala HMP60
Instrument accuracy: ±0.6°C
Measurement Range: -40°C to +60°C

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.
2. **-40C Issue**: the temperature range in which the HMP60 probe can work is stated to be -40C to +60C. When the temperature drops below -40C the sensor reads values very close to -40C (e.g. -39.76, -39.73, -39.80).
   Correction: replace the data below a certain threshold with data recorded by the Logger Temperature probe. The newly corrected data is then visually checked, and whenever needed ad hoc modifications are made to obtain a smooth curve.
3. **Unreasonable Spikes**: sometimes unreasonably high spikes in the temperature data happens (e.g. jump of over 60C in a single time step). These spikes are considered to be unreasonable.
   Correction: manually remove these spikes and interpolate from neighboring values.

Relative Humidity

Sensor: Vaisala HMP60
Instrument accuracy: ±5% (0 to 90% RH) & ±7% (90 to 100% RH)
Measurement Range: 0% to 100%

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.
2. **0% <= RH <= 100%**: check whether RH is > 100% or < 0%.
   Correction: set all values greater than 100% and smaller than 0% to 100% and 0%, respectively.
3. **Unreasonable Spikes**: sometimes unreasonably high spikes in the data happens (e.g. jump of over 80% in a single time step. These spikes are considered to be unreasonable.
   Correction: manually remove these spikes and interpolate from neighboring values.

Shortwave Radiation

Sensor: Hukseflux NR01
Instrument accuracy: 20% for hourly sums

Calibration Coefficients
When the two CR1000 at EKT and Site J were programmed in Spring 2017 to read the two radiometers (NR01) there was a major mistake: calibration coefficients for both shortwave and longwave radiation were switched between upfacing and downfacing sensors. This is mostly the result of uncareful reading and confusing descriptions in the manual and calibration sheets.
Just to clarify, the manual refers to:
-) upfacing sensors measuring downward radiation with SR01↓ and IR01↓ (INCOMING RADIATION)
-) downfacing sensors measuring upward radiation with SR01↑ and IR01↑ (OUTGOING RADIATION)
Calibration sheets use this notation as well (SR01↓, SR01↑, IR01↓, IR01↑). The confusing part is that the example program provided by Campbell to perform measurements of the sensors using a CR1000 uses the following notation:
-) SR01Up_W_m2 and IR01Up_W_m2 for INCOMING radiation (SR01↓ and IR01↓)
-) SR01Dn_W_m2 and IR01Dn_W_m2 for OUTGOING radiation (SR01↓ and IR01↓) This is probably done because the program refers to upfacing and downfacing sensors but it lead me into mistake.

Fortunately the calibration coefficients for the EKT SR01 sensor are pretty different and allowed us to realize the error. To avoid any further confusion I will refer to the component as incoming and outgoing.

Correction: switch the calibration coefficients

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.
2. **SWin & SWout >= 0**: shortwave radiation is always positive.
   Correction: first check if there is a bias in the sensor. If there is no bias correct the data setting all negative values to 0 for both SWin and SWout.
3. **SWout = 0 then SWin = 0**: if SWout = 0 then also SWin = 0, else it is not consistent (albedo = 0).
   Correction: set SWin = 0 whenever SWout = 0. It may still be the case that then SWin = 0 but SWout is not equal to 0 in this case, next correction (max albedo = 0.9) is gonna take care of it.
4. **SWout <= 0.9*SWin**: max albedo allowed = 0.9. before correcting this check how frequently it happens that SWout > SWin during the day (good indicator for anomalies, use 20 W m-2 to define day time).
   Correction: Replace unreasonable values assuming that SWout is the most accurate measurement (no issue of snow on the dome) and assume a maximum albedo of 0.9.

Longwave Radiation

Sensor: Hukseflux NR01
Instrument accuracy: 20% for hourly sums

For the longwave radiation correction use the corrected values for the instrument temperature NR01TC_degC_Avg. These values are already logged by the datalogger.

Calibration Coefficients
When the two CR1000 at EKT and Site J were programmed in Spring 2017 to read the two radiometers (NR01) there was a major mistake: calibration coefficients for both shortwave and longwave radiation were swicthed between upfacing and downfacing sensors. This is mostly the result of uncareful reading and confusing descriptions in the manual and calibration sheets.
Just to clarify, the manual refers to:
-) upfacing sensors measuring downward radiation with SR01↓ and IR01↓ (INCOMING RADIATION)
-) downfacing sensors measuring upward radiation with SR01↑ and IR01↑ (OUTGOING RADIATION)
Calibration sheets use this notation as well (SR01↓, SR01↑, IR01↓, IR01↑). The confusing part is that the example program provided by Campbell to perform measurements of the sensors using a CR1000 uses the following notation:
-) SR01Up_W_m2 and IR01Up_W_m2 for INCOMING radiation (SR01↓ and IR01↓)
-) SR01Dn_W_m2 and IR01Dn_W_m2 for OUTGOING radiation (SR01↓ and IR01↓) This is probably done because the program refers to upfacing and downfacing sensors but it lead me into mistake.

Fortunately the calibration coefficients for the EKT SR01 sensor are pretty different and allowed us to realize the error. To avoid any further confusion I will refer to the component as incoming and outgoing.

Correction: switch the calibration coefficients

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.
2. **LWin**: not much can be done to correct longwave incoming radiation, just check for unreasonable spikes and print min and max values.
3. **LWout <= 315.6574 W m-2**: longwave outgoing radiation is capped by the blackbody radiation of ice at the melting temperature.
   LWout = $\varepsilon\sigma T^4$, with $\varepsilon=1$, $\sigma=5.670367\times10^{-8}$ W m−2 K−4 and T=273.15 K
   LWoutMAX = 1 * 5.670367 * (10 ** (-8)) * (273.15 ** 4) = 315.65740930067255
   Correction: Before correcting check if the values > 315.6574 W m-2 look like a bias or just noise. Replace all LWout values > 315.6574 with 315.6574

Tiltmeter

Sensor: Turck B2N45H
Instrument accuracy: 0.1°

Not much can be really done with the tiltmeter, not sure it is the right sensor, and not sure why it stops working and then restarts again. However at any visit the station was always almost perfectly leveled.

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.

Wind Speed and Direction

Sensor: RM Young 05103
Instrument accuracy: wind speed = 0.3 m s-1 | wind direction = 3°

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.
2. **Wspd >= 0**: wind speed is always positive. Check some statistics before correction.
   Correction: set all negative values of wind speed to 0.
3. **Wspd Rimed Sensor**: Multiple consecutive measurements of 0 m s-1 wind speed might be an artifact. While 0 m s-1 wind speed is plausible it is very unlikely that it occurs recursively at our sites. In these situations it is more likely that the wind propeller is rimed and not allowed to move.
   Correction: you can either manually check these events (for rather small dataset) or apply an automated procedure (for bigger dataset like in this case).
   Automated Procedure: set to NaN all the wind speed measurements that are = 0 for 1 hour interval or more.
4. **Wdir Rimed Sensor**: Sometimes, especially during fall (November, e.g.), exactly the same values (or very similar) are shown in consecutive measurements, this is typically an artifact due to the wind sensors blocked by rime.
   Correction: you can either manually check these events (for a rather small dataset) or apply an automated procedure (for a bigger dataset like in this case).
   Automated Procedure: set to NaN all 3 consecutive (or more) measurements that fall within a defined threshold.

Surface Height Change

Sensor: Campbell SR50A
Instrument accuracy: 0.01 m

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: interpolate from neighboring values if it is just a couple of NaN here and there.
2. **Semi-Automatic Procedure**: to homogenize the data and get rid of unreasonable spikes.
   SR50 data can be very noisy, due to several reasons. E.g. low accumulation rates, high wind speeds, particles (like snowflakes) below the sensor during snowfall events. First of all the data (both SR50 distance and Quality Number) have been plotted and visually checked.
   Quality Number Range Quality Range Description:
   -) 0 Not able to read distance
   -) 152 to 210 Good measurement quality numbers
   -) 210 to 300 Reduced echo signal strength
   -) 300 to 600 High measurement uncertainty
   Procedure Description:
   1) first flag as NaN all the values that have a Quality Number > 210
   2) calculate the difference between each consecutive value (this is done to SR50 raw data, prior to any correction step), flag as NaN all the values for which the difference is > +-0.02 m in this case, so 2 cm. I have been playing around with this threshold and I am not totally convinced by it, but it seems to produce reasonable results. Few thougths about the 2 cm number:
   -) 1 cm is the declared SR50 accuracy, lowering the threshold below 2 cm would make no sense
   -) snowfall at Site J have typically a low accumulation rate (due to wind) so it would make sense to see a pretty low hourly threshold
   3) interpolate the whole series to fill the NaNs. At the moment I am using a linear interpolation but other methods can be explored
3. **Station Rised**: when the station was raised and or the cross arm was moved.
   Correction: use manual height measurements before and after the station mast was extended to correct the data.
4. **Surface Height Change Conversion**: the sensor measures the distance from itself and the ground. Transform this data to surface height change with 0 reference on the day of installation.

Firn Temperature

Thermistor String: Geokon 3810-2
Instrument accuracy: 0.2°
Snow thermistors: custom build (check model)
Instrument accuracy: 0.2°

In Spring 2019 the two or three lower thermistors have been disconnected from the datalogger and 2 new thermistors have been added at the interface between firn and seasonal snow.

1. **Check NaN**: first look for NaN values in the Raw data.
   Correction: NaN values in the firn temperature are currently not corrected.
2. **Installation and Station Reboot**: after the thermistor string is lowered in the borehole it takes some days for the sensors to reach thermal balance with the surrounding firn. Furthermore every time the station was rebooted the first measurements were completely off (not sure way).
   Correction: remove first days after installation and exclude first measurements after the station is rebooted.
3. **Unreasonable Spikes**: the whole data series of firn temperatures from the thermistor string is affected but some sort of instability causes unreasonable spikes to happen. These spikes mostly affect the thermistor closer to the surface although at times it affects all of the thermistors. The fact that these spikes exhibit no evidence of signal lag with depth as well as the amplitude of spikes (up to over 0.5 C even in the middle of the winter when no meltwater is present in the system) suggests that these are not real. We are unclear on the nature of what is causing them, maybe some electrical instability in the logger when the thermistors are read? Despite this we do believe that the base of the measurements (spikes removed) is correct.
   Correction: in order to remove the spikes we do apply a low pass filter to the daily mean of firn temperature data. Only daily means are provided in the final daily file. The thermistor closer to the surface is not filtered since it doesn't show any unreasonable spike. Also the snow thermistors installed in Spring 2019 since they do not exhibit this behavior as well.
4. **Thermistor at the surface**: if a thermistor melts out to the surface, or very close to the surface where shortwave radiation penetration in the snow affects the measurements, the measurements itself it's not reflecting anymore the actual firn temperature. This is clearly visible in the 2 snow thermistors added in Spring 2019: first of all these thermistors were installed close to the surface and in summer 2019 there was enough surface lowering to cause the sensors to reach the surface. Once the thermistors are out the daily signal of the sun is clearly visible in the data. This is of no interest and hence removed from the series.
   Correction: manually remove all data when thermistors are out.