Metadata

Name of the program: program_09.py

Name of program creator: Karoll Quijano – Kquijano

Description:
The script program_09.py uses the file 'DataQualityChecking.txt' as input and performs a data quality check. The variables to import are: date, precipitation (mm), maximum air temperature (°C), minimum air temperature (°C), and wind speed (m/s), using date as index. The quality check to follow are:

- Check 1: Removes No Data values, replacing all values of -999 in the file with the NumPy NaN values (e.g., numpy.NaN or np.NaN).
- Check 2: Check for gross errors, applying the error thresholds: $0 \leq P \leq 25$; $-25 \leq T \leq 35$, $0 \leq WS \leq 10$. Replacing values outside this range with NaN.
- Check 3: Swap Max Temp and Min Temp when Max Temp is less than Min Temp.
- Check 4: Check for daily temperature range exceedence. Identifying days with temperature range (Max Temp minus Min Temp) greater than 25°C, replacing both Tmax and Tmin with NaN when range is exceeded.

The output file 'DataQualityCheckingCorrected.txt' saves the data that passes the quality check.

The tab delimited output file 'ReplacedValuesDF.txt' records the number of values replaced for each data quality check as: "1. No Data", "2. Gross Error", "3. Swapped", and "4. Range Fail" (Table 1.)

Table 1. Replaced values in dataframe

1.	No Data	2.0	2.0	2.0	0.0
2.	Gross Error	15.0	14.0	2.0	2.0
3.	Swapped	0.0	4.0	4.0	0.0
4.	Range Fail	0.0	5.0	5.0	0.0

The following graphs are shown as results:

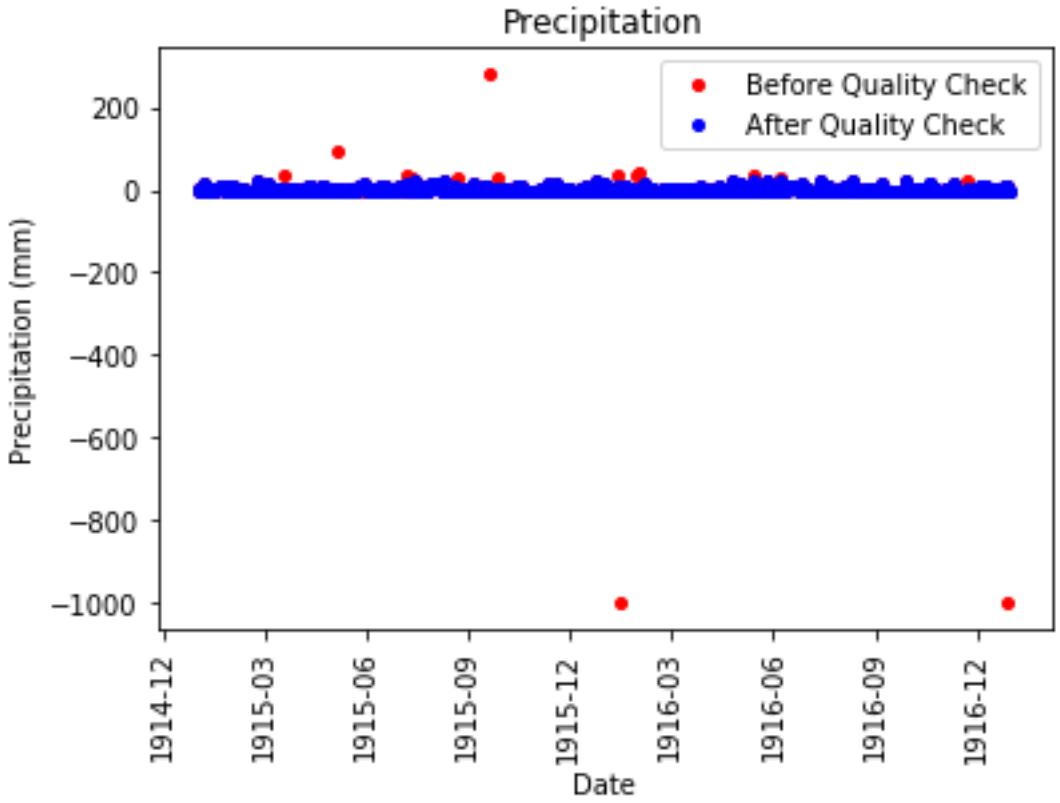

Figure 1. Plot of precipitation data before (red) and after (blue) quality check.

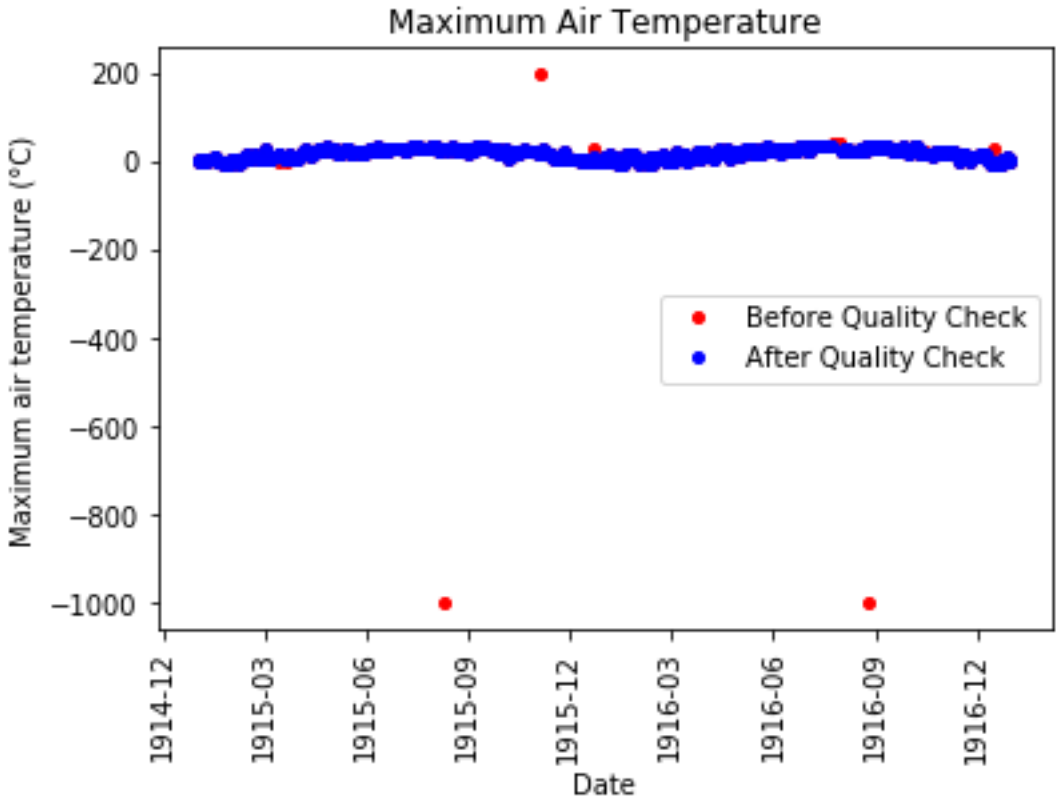

Figure 2. Plot of maximum air temperature data before (red) and after (blue) quality check.

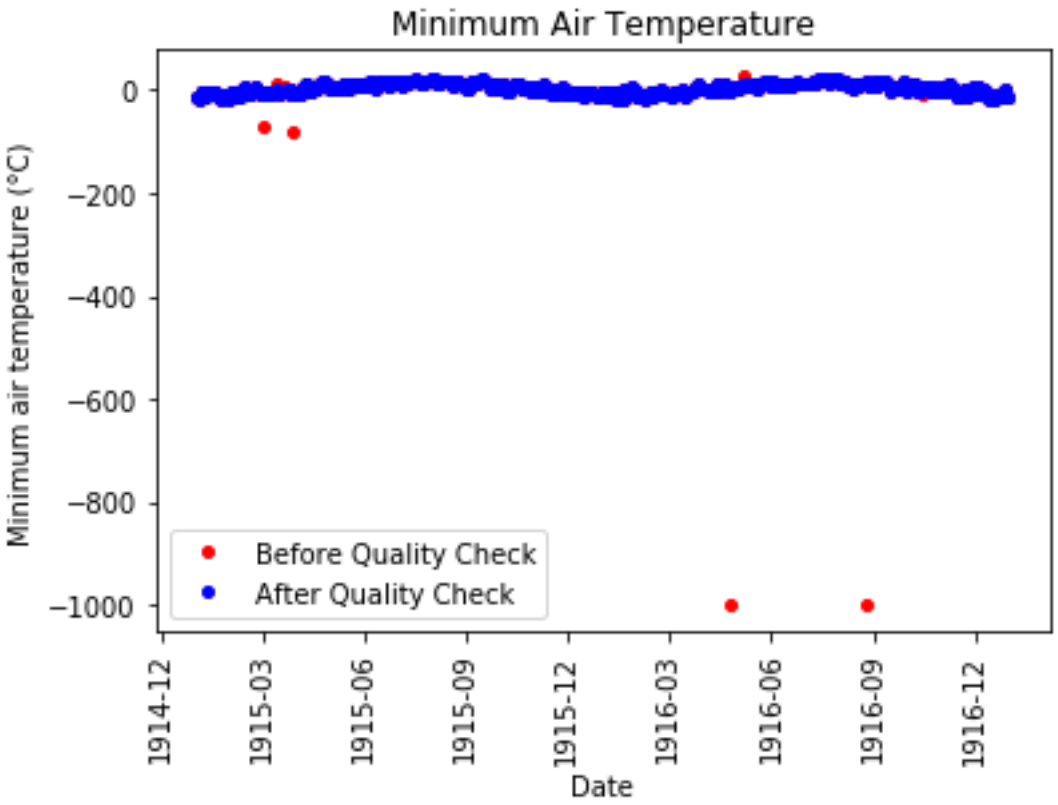

Figure 3. Plot of minimum air temperaturdata before (red) and after (blue) quality check.

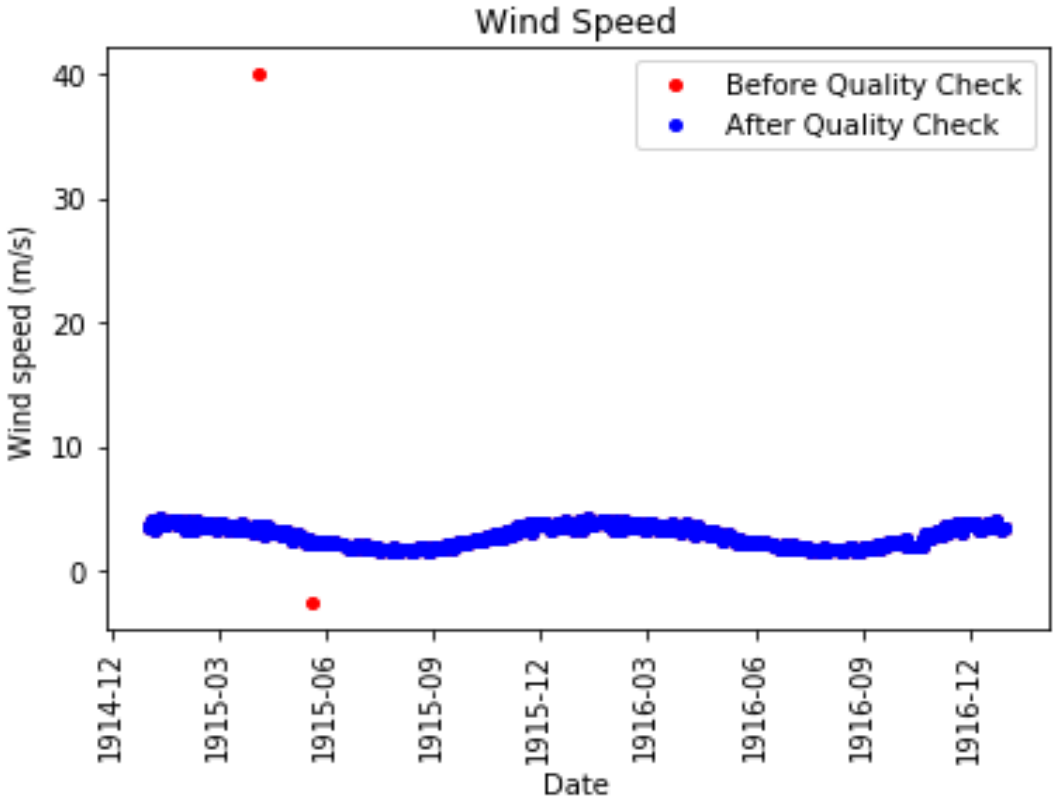

Figure 4. Plot of wind speed data before (red) and after (blue) quality check.