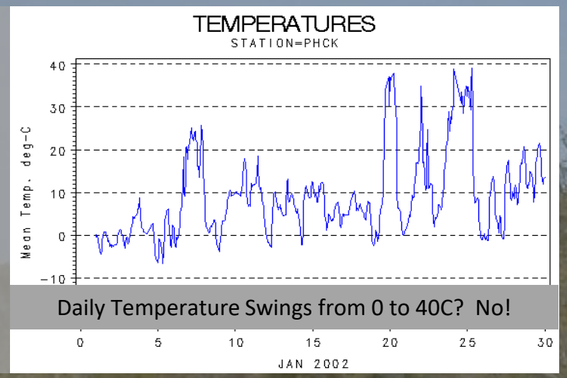
Approaches to Quality Control and Assurance in a Meteorological Station Network

John H. Porter, Nina Groleger & Jonah Morreale

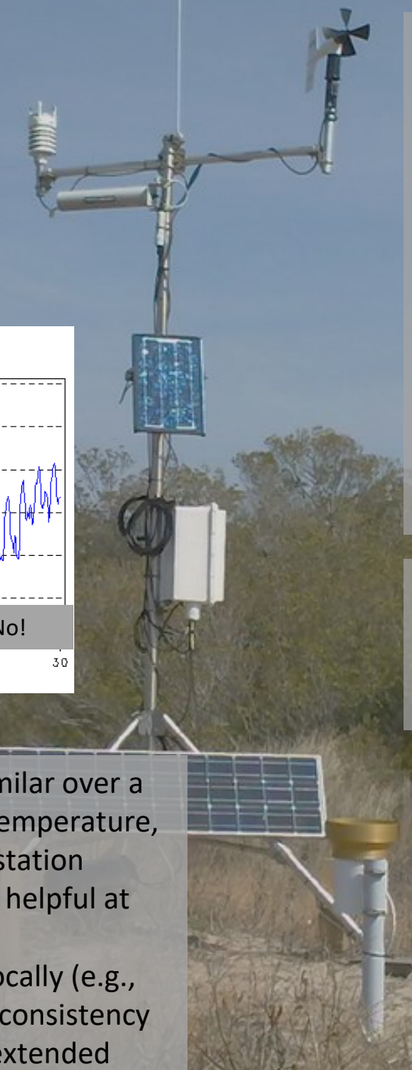
During the summer of 2021 we have been going back over 30+ years of data from a small network of meteorological stations to resolve existing quality issues caused by partial or complete sensor failures. Once detected, issues are input into a database that then uses code-generation to flag, remove or comment on issues. Here we present some of the automated or semi-automated methodologies we have found to be most effective, as well as future plans for gap filling.

Issues:

- Sensor Failure (complete)
  - Easy to detect – no data recorded
- Sensor Failure (intermittent)
  - Harder to detect
- Anomalous Sensor Readings
  - Consistent – “stuck” sensors
  - Intermittent - Bad values
  - Longterm Drift



Daily Temperature Swings from 0 to 40C? No!

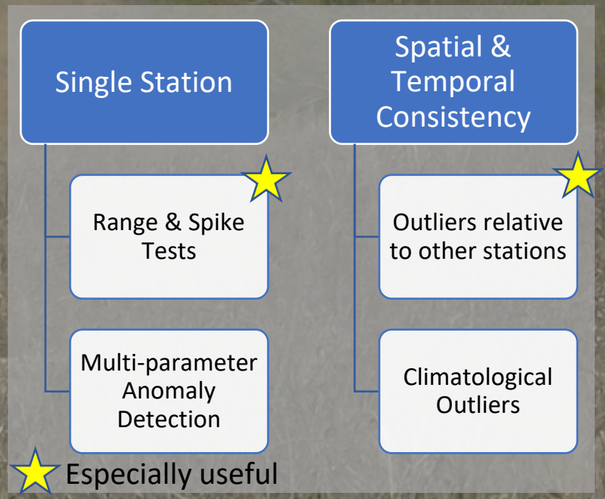


Automated Detection is accomplished using R code to estimate periods of sensor failure. Depending on the severity of the apparent failure, data are flagged, or in extreme cases, set to missing.

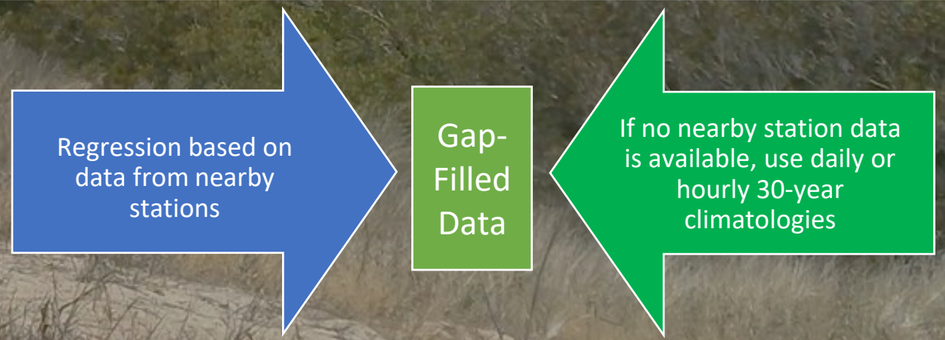
Instructions for data processing are recorded into a database (SensorProb) and code-generation services are used to transform level 0 (raw) data into level 1 (processed) data.

For example, automated code identified over 1,600 periods out of 30+ years of data at 3 stations where air temperature represented an anomaly when compared with other stations at the same time. A web service was used to “flag” as questionable the identified data in the SensorProb database.

To aid in eventual production of gap-filled data we are developing inter-station regression models so that one station can be used to predict values at another, and daily and hourly annual climatologies for each station based on 30+ years of data



- For parameters that are similar over a broad spatial range (e.g., Temperature, Pressure, Humidity), inter-station comparisons are especially helpful at identifying issues
- For parameters that vary locally (e.g., summer rainfall) temporal consistency may be more useful (e.g., extended periods of no rain, or apparent rainfall every day)



This work was supported by NSF Grant 1832221 and a Fellowship to N. Goleger from the Environmental Data Initiative (NSF Grants 1931143 and #1931174)