Data Quality Assurance and Control Methods for Weather Observing Networks

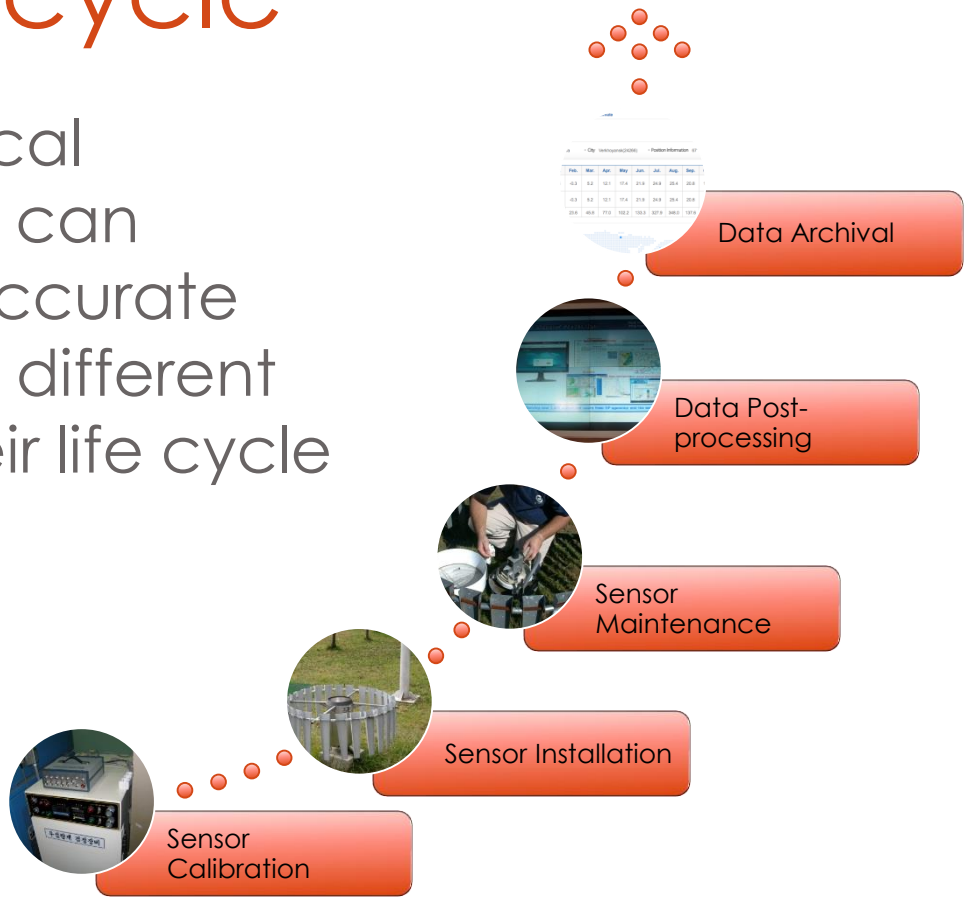
Cindy Luttrell
University of Oklahoma
Oklahoma Mesonet

Quality data are

- Trustworthy
- Reliable
- Accurate
- Precise
- Accessible

Data life cycle

- Meteorological observations can become inaccurate during many different stages of their life cycle



Unavoidable problems

Unavoidable problems

Unavoidable problems

Unavoidable problems

Designing QA/QC processes

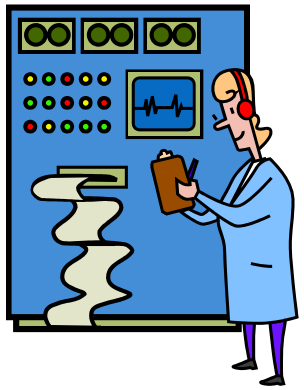
- Identify network goals
- Sensor dependent
- As simple or complex as you desire

End-to-end QA system

- Incorporation of
  - sensor calibrations
  - maintenance information
  - automated and manual quality control
- is essential for producing **trusted, high-quality data**.

General QA Considerations

- Station siting
- Routine site maintenance
- Routine calibration of sensors
- Archival of original observations
- Use of Coordinated Universal Time and standard units
- Use of similar instruments and instrument configurations
- Installation of redundant sensors for core variables



Station siting

- Representative of area that measurements are intended to characterize
- Secure location
- Easy to access by maintenance personnel

Routine site maintenance

- Vegetation may encroach site over time
- Sensors eventually coated in
  - Dust
  - Mold
  - Debris

Routine site maintenance

Routine site maintenance

SPRING 2009 MESONET SITE MAINTENANCE PASS FORM. (Pass to be completed by 06-30-2009.)
(Please use BLACK INK for this form to facilitate scanning.)

Site ID	Date	ARRIVAL TIME (GMT)	DEPARTURE TIME (GMT)	Tech(s)
WYNO	4-21-09	19:10	20:15	RH

1. Site Photos Taken Upon Arrival
- ☑ Soil temperature plots | **SP1**; **BP1** ☑ Soil moisture plots | **SS1**; **SN1** ☑ Net rad footprint | **NF1**
- ☑ Inside SE to NW photo across site showing veg height gauge placed 6ft S and 6 ft E of tower taken @ veg height | **IH1**
- ☑ Outside photo from the south showing vegetation height gauge placed 10 ft S of south fence taken @ veg height | **OH**

2. Site Vegetation Maintenance
Bare plot ___0___% covered with veg. upon arrival | Vegetation removed: ☐ Yes ☑ No | Sterilant applied: ☑ Yes ☐ No
Bare plot edging installed correctly: ☑ Yes ☐ No | Corrected: ☐ Yes ☑ No
Bare depth indicator correct on arrival: ☑ Yes ☐ No->_____cm ☐ Too Deep ☐ Too Shallow | Corrected: ☐ Yes ☑ No
Sod depth indicator correct on arrival: ☑ Yes ☐ No->_____cm ☐ Too Deep ☐ Too Shallow | Corrected: ☐ Yes ☑ No
☑ Site vegetation at required height upon departure **(See reverse side for guidelines)**

3. Conduct Pre- and Post-Cleaning Rain Gauge Drip Tests (Location 45 on Keypad with Rain Shunt)

PRIMARY METONE RG DRIP TEST	Test TIPS	Tips Expected	Error Tips
Pre-Cleaning	50.8	50	+0.8
Post-Cleaning	50.9	50	+0.9
Extra- as needed		50	

SECONDARY METONE RG DRIP TEST	Test TIPS	Tips Expected	Error Tips
Pre-Cleaning	n/a	50	n/a
Post-Cleaning	n/a	50	n/a
Extra- as needed	n/a	50	n/a

4. General Maintenance (Enter Y(es) or (n)O in every box) **[*NOTE* Clean IRT lens with Methanol and Q-Tip]**

Cleaning/ Leveling	Clean on Arrival	Cleaned	Level on Arrival	Leveled
SRAD	Y	O	Y	O
RNET	Y	Y	Y	O
IRT	Y	Y	Y	O
RAIN_PRIM	O	Y	Y	O
RAIN_SEC	n/a	n/a	n/a	n/a

MesoTower Cleaning	Clean on Arrival	Cleaned
T&RH Shelter	Y	Y
TAIR Shelter	O	Y
Solar Panel	Y	O
Battery Terminals	Y	O
Volt Reg Terminals	Y	O

Aux Pwr Cleaning	Clean on Arrival	Cleaned
Solar Panel	Y	Y
Battery Terminals	Y	O
Volt Reg Terminals	Y	O
TA9M	NA	NA
WS2M	Y	Y

5. Sensor Install, Replace, Remove, etc.

EQUIPMENT TYPE/PARM	S/N Removed or Repaired	S/N Installed	TT# (if any)	Ops Notified	Comment (ROT,.RPL, RMV,OSR, INI, NAT)
PYRA			N/A	YES	ROT
T&RH			N/A	N/A	ROT

6. Site Photos Taken Upon Departure
- ☑ Soil temperature plots | **SP2**; **BP2** ☑ Soil moisture plots | **SS2**; **SN2** ☑ Net rad footprint | **NF2**
- ☑ Inside SE to NW photo across site showing veg height gauge placed 6ft S and 6 ft E of tower taken @ veg height | **IH2**
- ☑ Inside photo from NW fence corner to SE showing extent of 3-inch cut | **CH**

7.
- ☑ **Desiccant Replaced** (At least every 6 months, replace 4 packs of desiccant and write date on new packs at edge)
- ☑ **Check ALL datalogger LOCs for "normal" readings (*6 mode)**
- ☑ **Station departure time from data logger (* 5 mode) recorded at top of form and logger placed in * 0 mode just before securing data logger enclosure**

Tech_RH_ Date Submitted to Field Manager:___4-22-09_____ Date Received by Field Mgr: 4-22-09

Routine site maintenance

- Routine maintenance provides an efficient means of conducting sensor inspections and tests, as well as documentation of stations with digital pictures

Southeast South Southwest

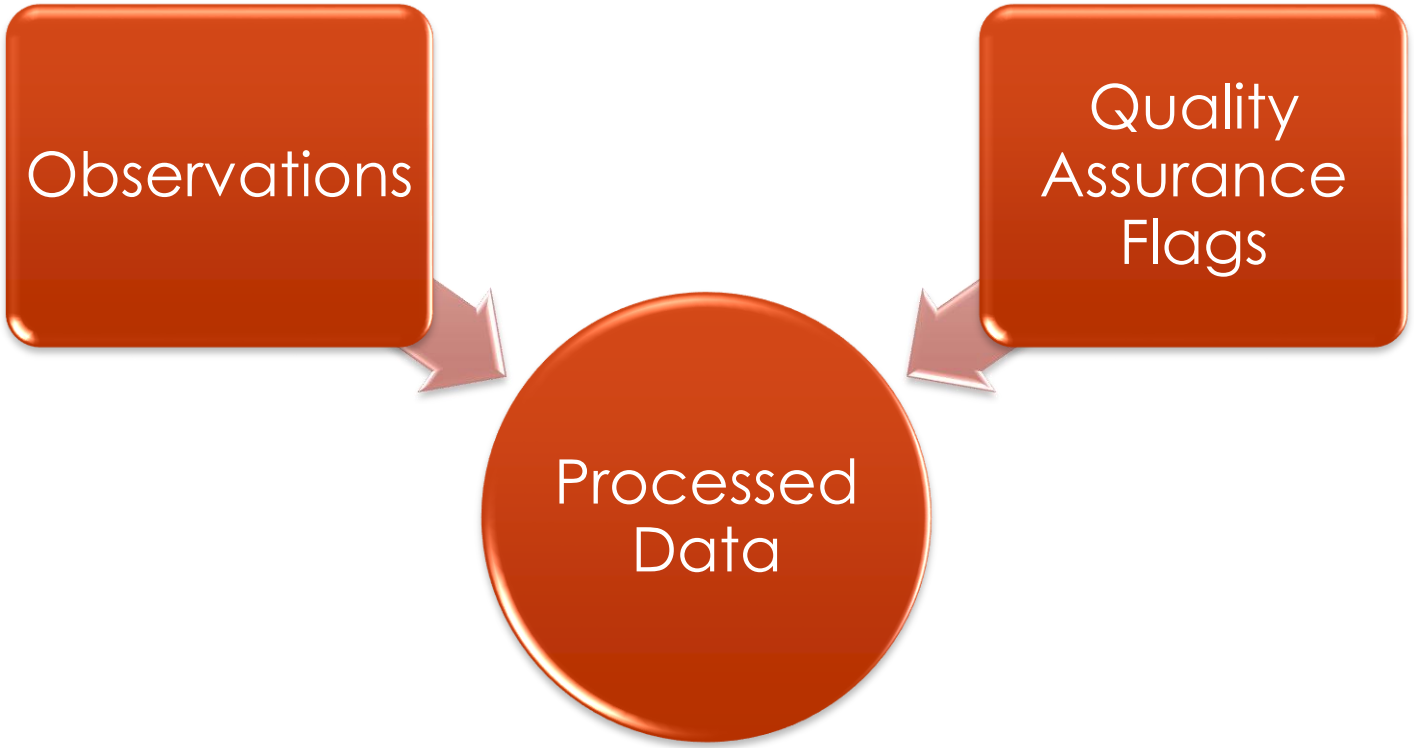
Routine sensor calibration

- Testing of sensor prior to deployment to station (pre-field calibration)
- Testing of sensor after removal from station, prior to cleaning (post-field calibration)
- Helpful to document how long sensor is at a station

Archival of original data

- Never change data
- Data quality flags can be linked to each datum, identifying the quality of the observation
- Flags can be adjusted as data are re-evaluated (you might change your mind)

Data and QA Flags



Standard units

- For intercomparison of data across time zone boundaries, it is imperative that raw observations adhere to standard time
- UTC eliminates confusion during transition to/from daylight saving time
- Routine verification of datalogger clocks critical to avoid clock drift
- Conversion to local time or other units may be applied during post-processing (after QA)

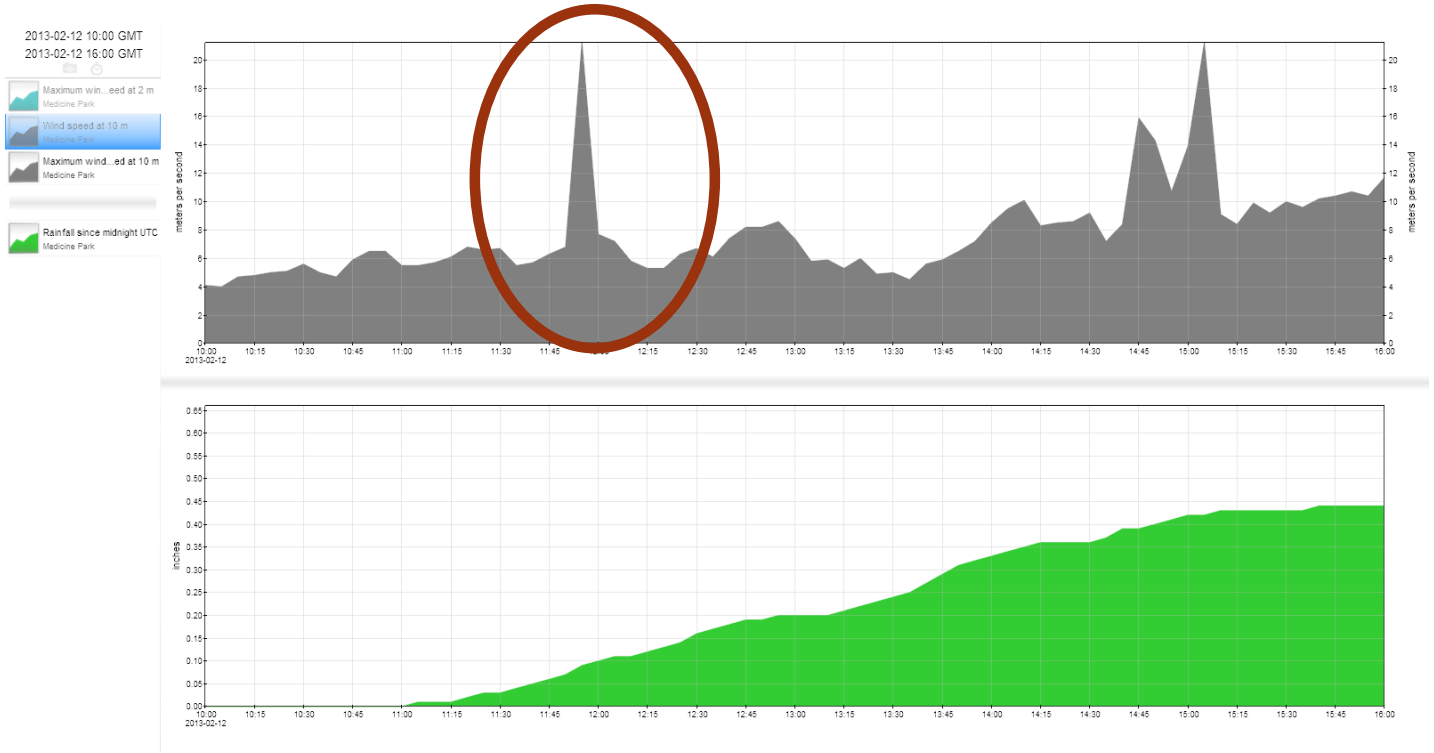
Similar instruments and configurations

- Use of similar instruments and site configurations allow for efficient troubleshooting
- Networks consisting of several different datalogger, sensor, and mounting configurations produce abundant combinations of potential problems
- Use of multiple types of sensors, perhaps with different time constants or measurement methods, present obstacles to making objective comparisons

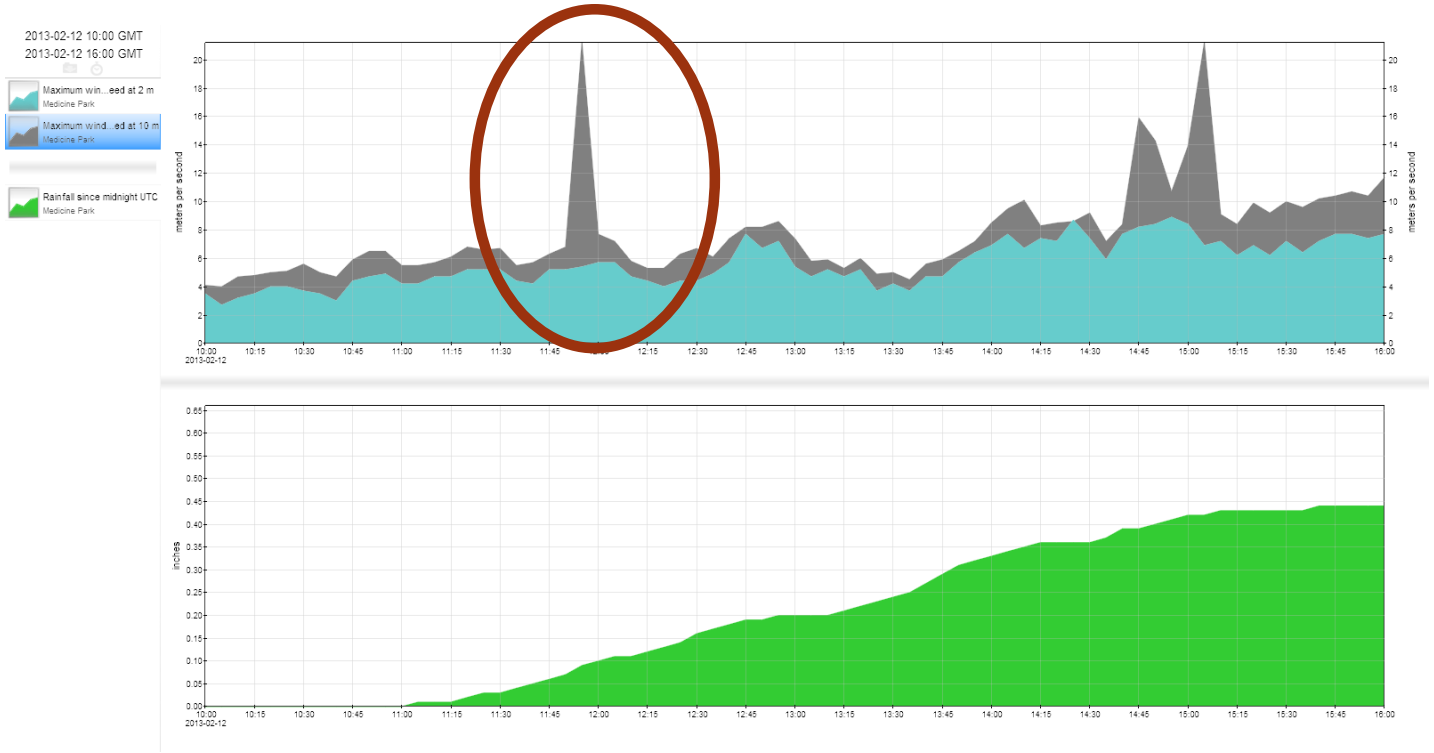
Redundant Sensors

- Most straight forward QA test involves comparison of two or more identical sensors at same station and height
- Should be considered during planning of station configurations
- If very accurate temperature data needed, but funds are limited, better to install two temperature sensors and forgo an ancillary measurement (e.g., pressure or solar radiation)

Redundant Sensors



Redundant Sensors



Sensor Behavior

- Must understand how sensors function
- Determine common causes of failure
- Known problems are not advertised by manufacturers and are infrequently documented in journals
- Share experiences with other weather networks

Automated QA

- Frequent weather observations and limited staff make data quality assurance difficult
- Computer programs easily identify most suspicious observations

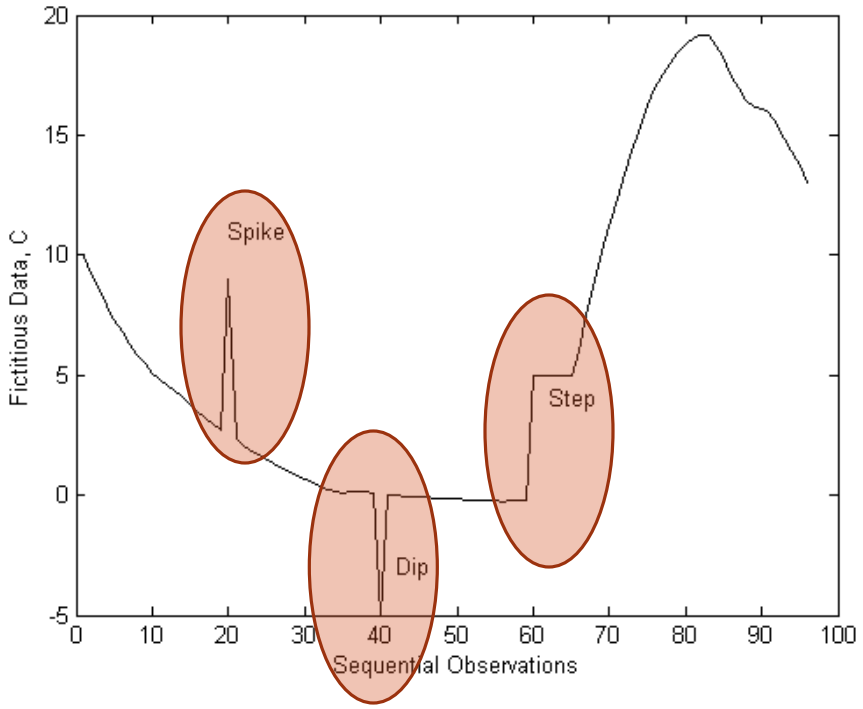
Automated QA

- **Range Tests**
  - **Sensor-based range tests: detect observations that are outside the range of sensor hardware or theoretical limits**
  - Climate-based range tests: typically use archived data to calculate thresholds by variable, station (or region), and date (or set of dates) to account for seasonal variation of observations

Automated QA

- **Temporal Checks**
  - **Step Tests: Compare the change in magnitude between sequential observations**
    - Threshold values dependent on station location (i.e., climate regime), time interval (e.g., 5-min, hourly, and daily), variable, and tendency
    - Spike and dip tests may be more successful at finding erroneous data

Automated QA



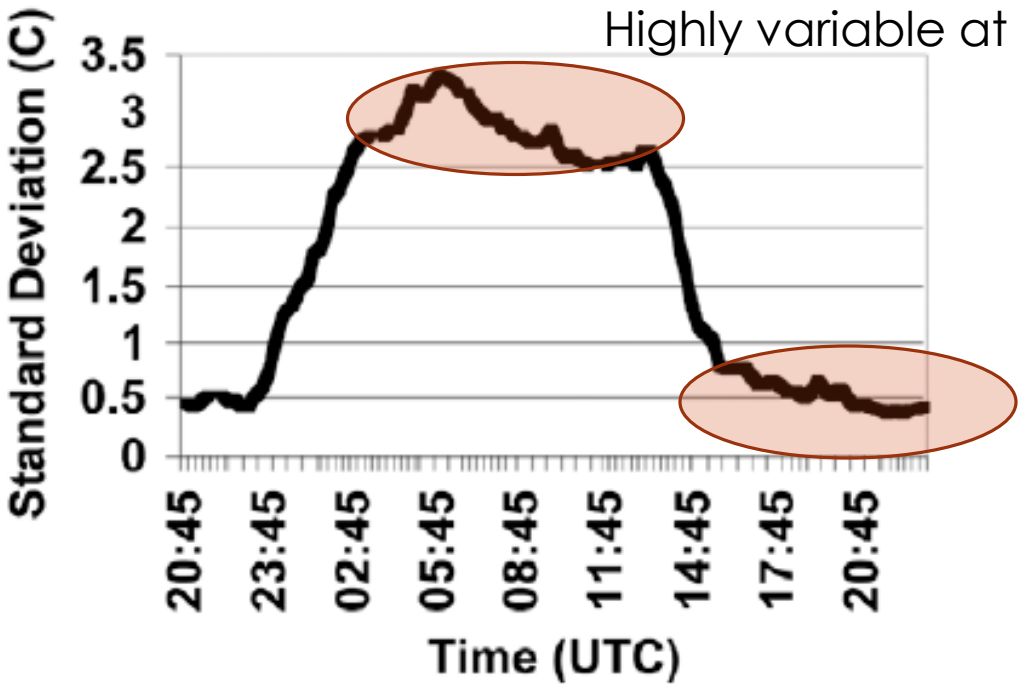
Automated QA

- **Temporal Checks**
  - **Persistence Tests: Assess whether observations vary minimally with time, possibly indicating a physical problem with either the sensor or its wiring**
  - Variable dependent and compare length of time a variable has repeated the same observation to its persistence threshold

Automated QA

- **Spatial Tests: Identify observations that are inconsistent with data from nearby stations**
  - Typically, data from site evaluated are compared to expected values (calculated using a spatial objective analysis algorithm)
  - Observations that differ by more than a predefined threshold from the expected value flagged as suspect
  - Thresholds depend on variable, locations of nearby stations, and distance to neighboring stations

Standard deviation of temperature across a micronet



Automated QA

- **Like-instrument** and internal consistency test
  - Compare air temperature at different heights
  - Compare temperature with dewpoint, sea level pressure with station pressure, etc.

Manual QA

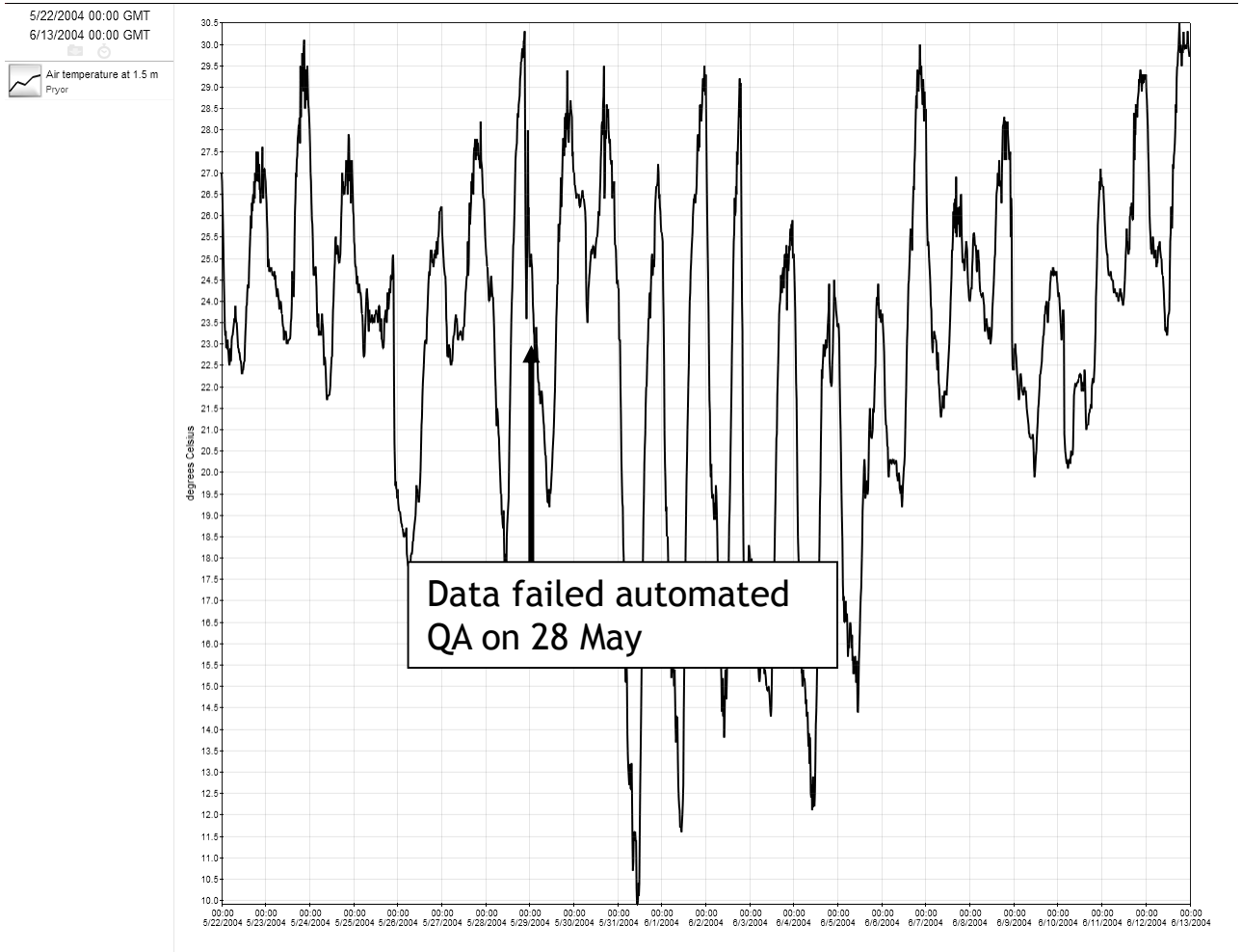
- Impossible to implement automated quality assurance system that identifies every bad observation yet never inadvertently flags good data as erroneous
- Outputs from QA system provide crucial pieces of evidence to help you determine which data need further analysis

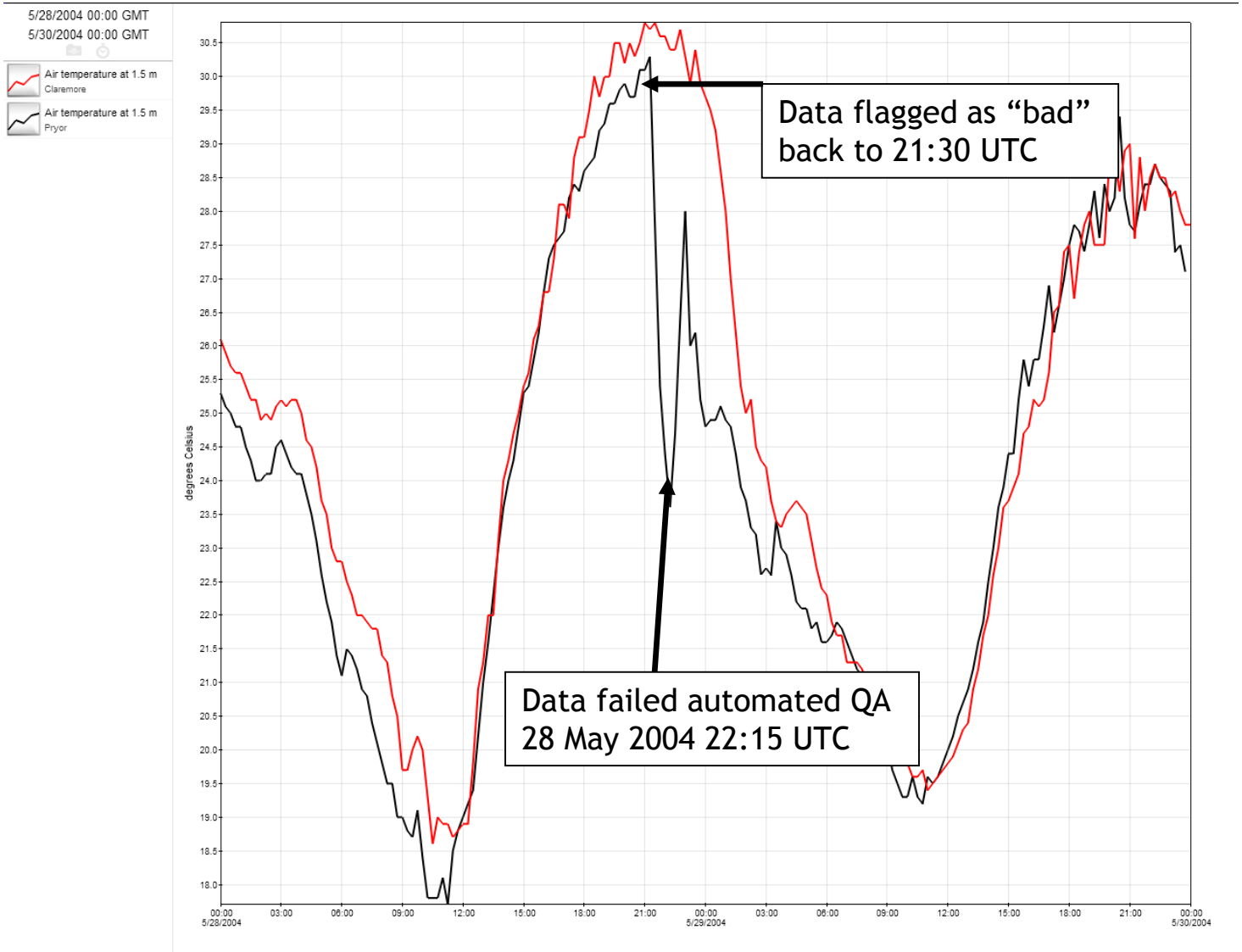
Manual QA

- **Review all observations that fail automated QA tests**
- Trace true start of problem (i.e., typically a time before automated software detected problem)
- Data flagged from true trace time of problem until sensor replaced or repaired by technician

Manual QA

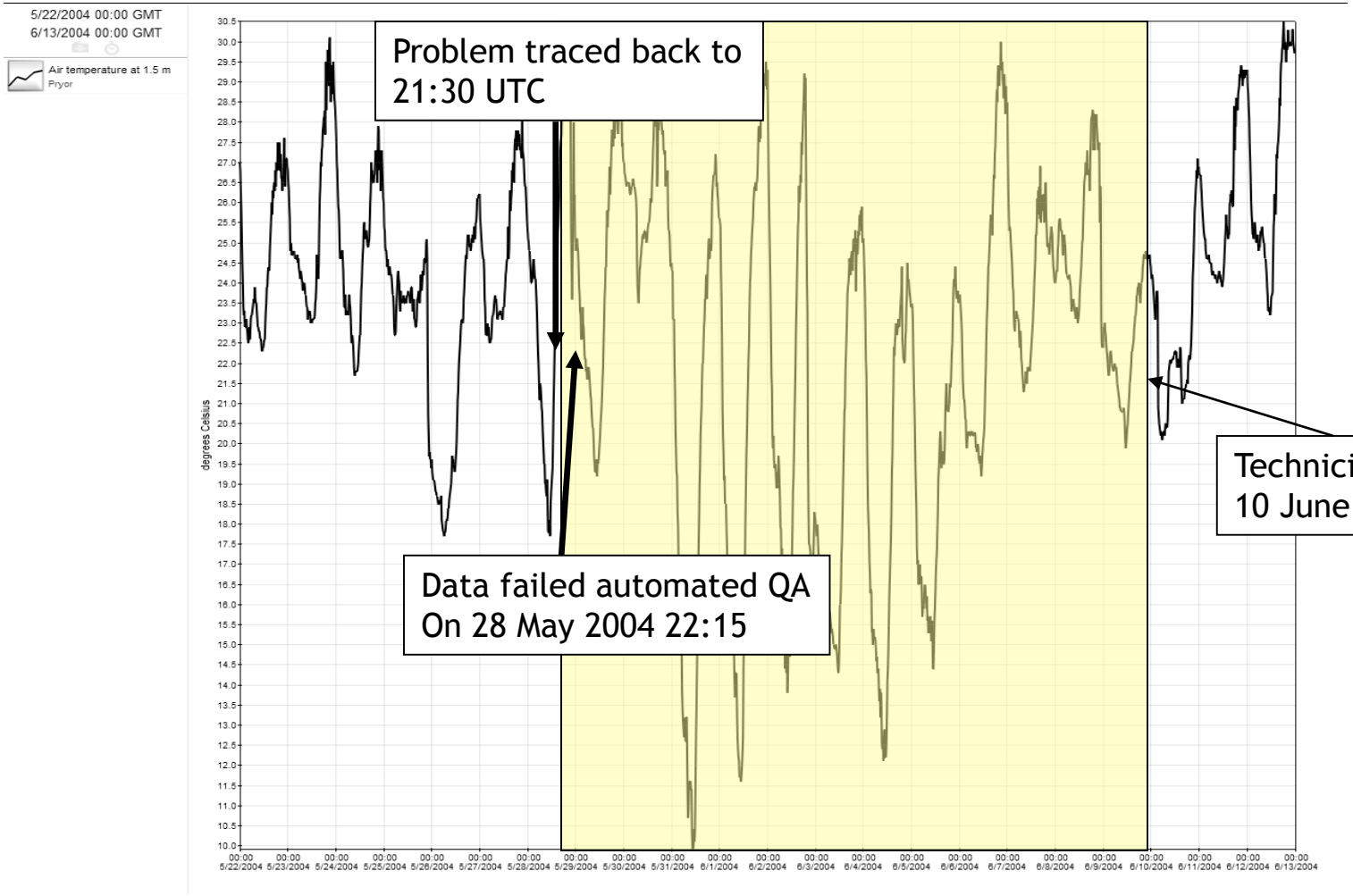
Flagging the Appropriate Data




5/28/2004 00:00 GMT
5/30/2004 00:00 GMT
Air temperature at 1.5 m
Claremore
Air temperature at 1.5 m
Pryor
degrees Celsius
Data flagged as "bad"
back to 21:30 UTC
Data failed automated QA
28 May 2004 22:15 UTC
5/28/2004
5/29/2004
5/30/2004

Manual QA

Flagging the Appropriate Data



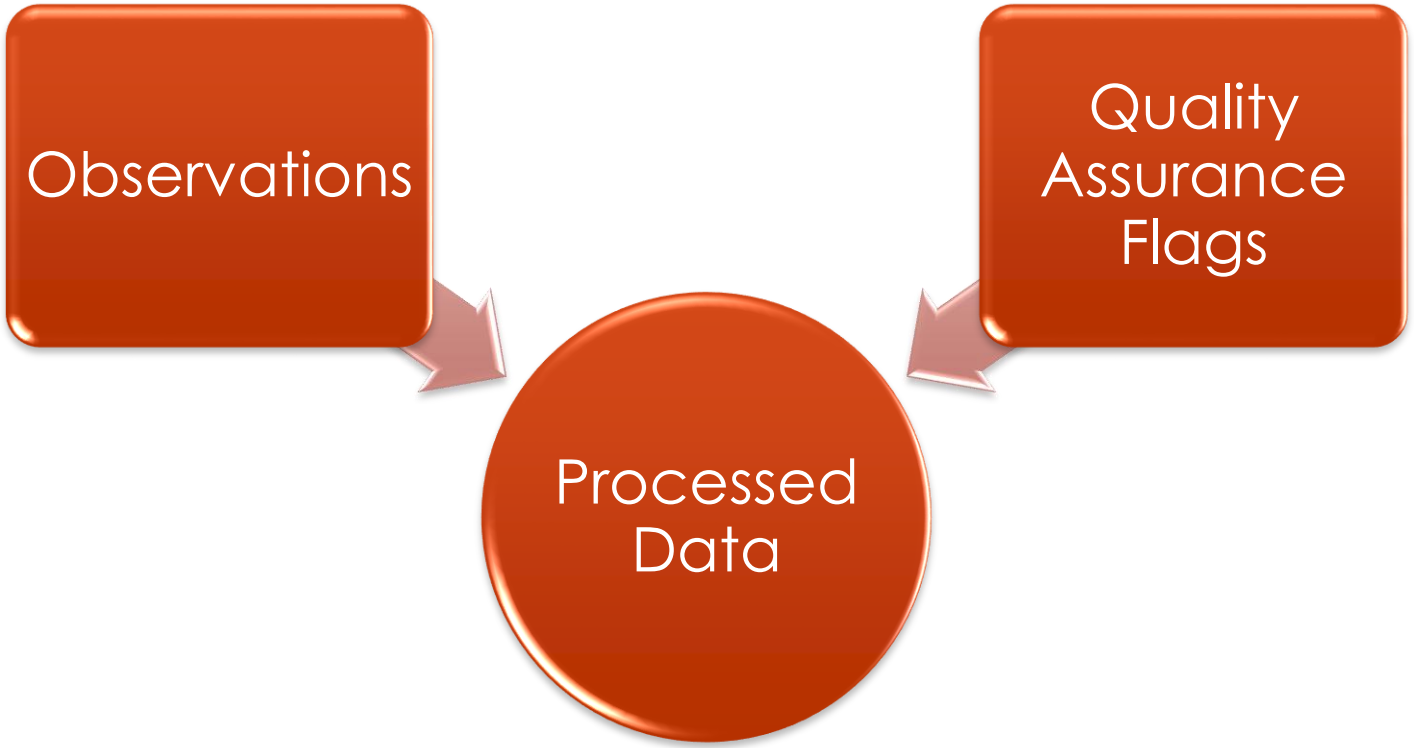
Manual QA

- Isolated or extreme weather events (e.g., hurricanes, sharp fronts) sometimes fail automated QA tests
- **Remove automated QA flags on data if they are determined to represent real meteorological phenomena**

Decision-maker for final QA flag

- Combines manual QA flags (from you) with automated QA flags
- Determines final flag for each observation

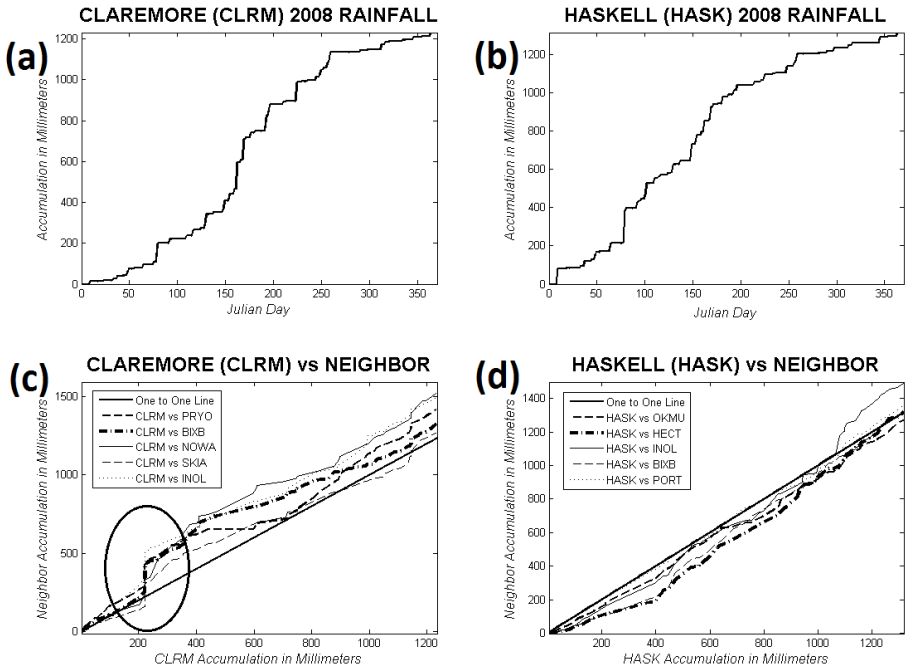
Data and QA Flags



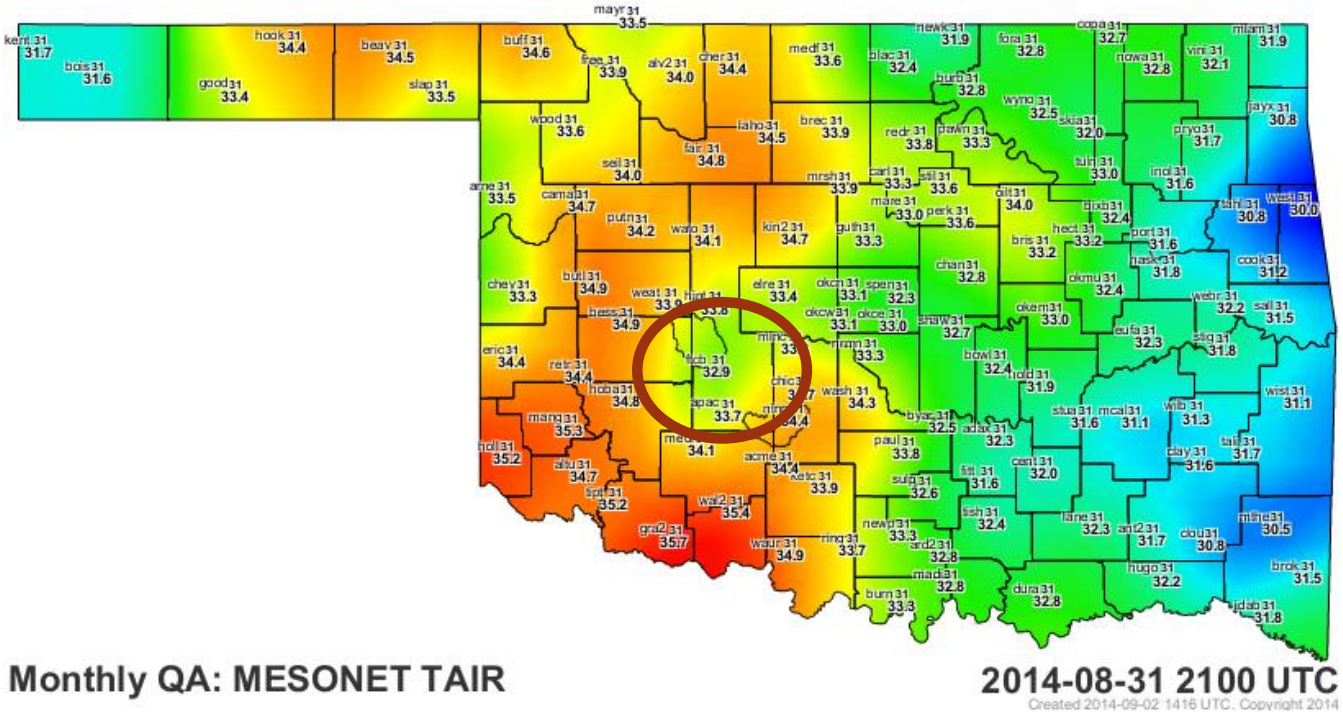
Additional Manual Analysis

- Double Mass Analysis
- Monthly Averages or Accumulations
- Monthly Extremes
- Time Series Graphs (meteograms)

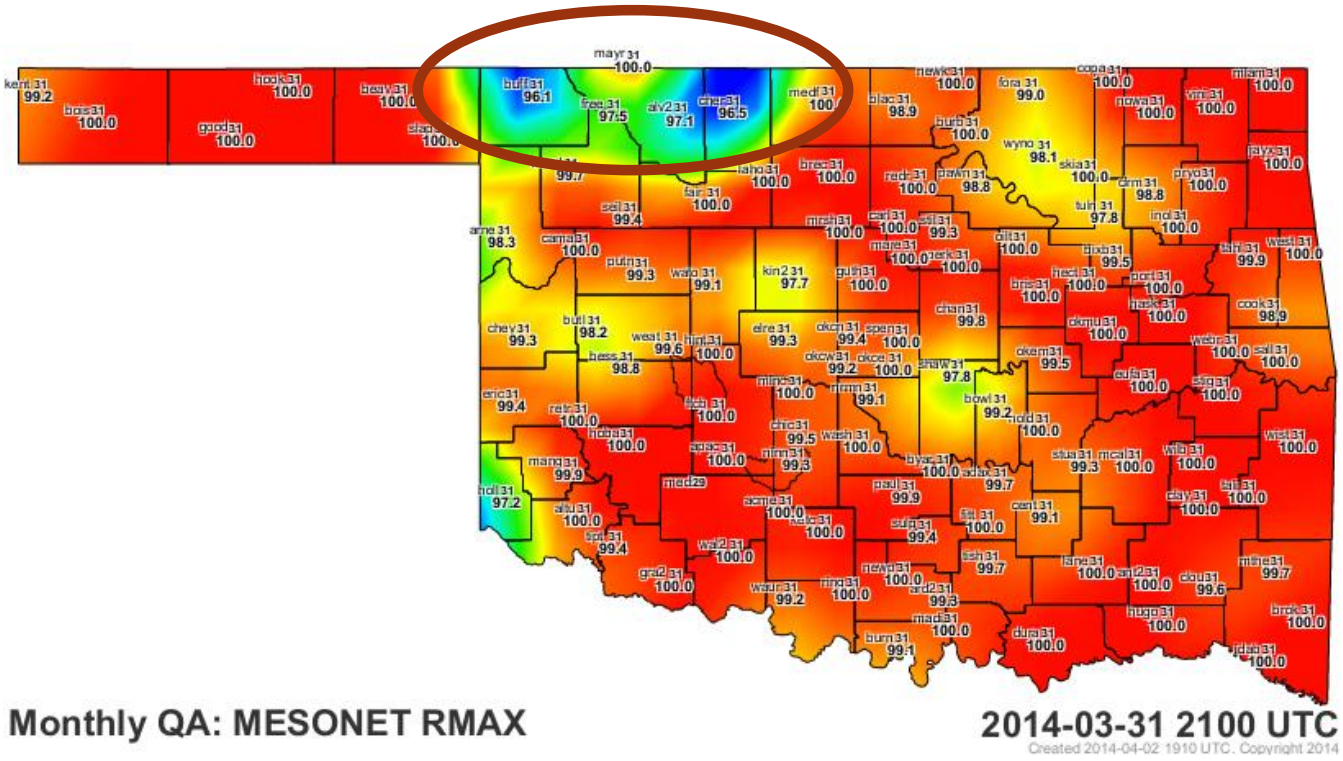
Doublemass Analysis



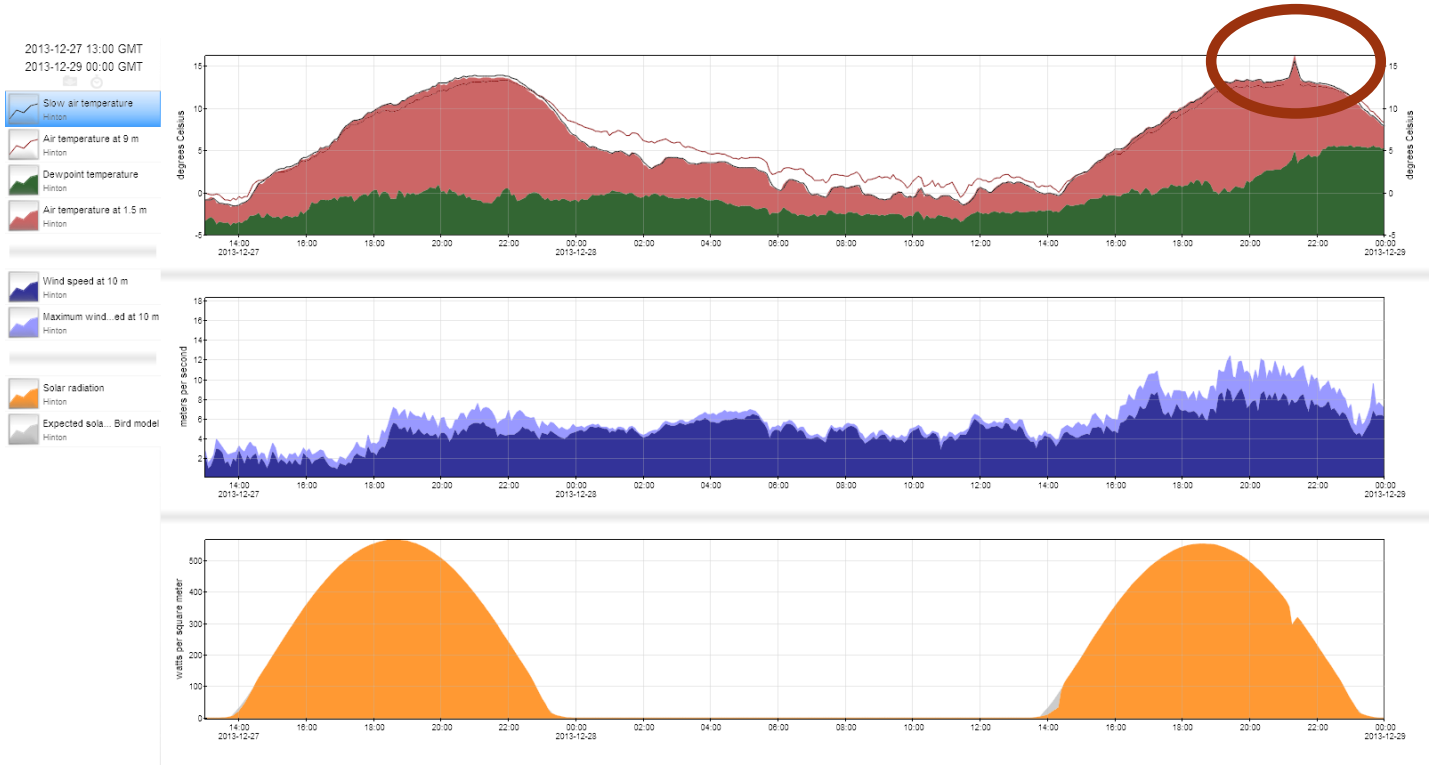
Monthly Averages/Accumulations



Monthly Extremes



Meteograms


2013-12-27 13:00 GMT
2013-12-29 00:00 GMT
Slow air temperature
Hinton
Air temperature at 9 m
Hinton
Dewpoint temperature
Hinton
Air temperature at 1.5 m
Hinton
Wind speed at 10 m
Hinton
Maximum wind...ed at 10 m
Hinton
Solar radiation
Hinton
Expected sola... Bird model
Hinton
degrees Celsius
meters per second
watts per square meter
2013-12-27
2013-12-28
2013-12-29

Summary

- Quality assuring meteorological data requires an evolving, dynamic system
- Adherence to network defined standards in siting, maintenance, and calibration ensures a strong foundation for data quality
- A set of core, automated algorithms are useful for identifying suspicious observations
- Automated tests must always be complimented with manual analysis to ensure high quality, trustworthy data
- Data observations should never be changed
- Use data quality flags to supplement archived observations

Questions?