Validation of CSN Data

Sean Raffuse
Nicholas Spada
Nicole Hyslop

Air Quality Group
UC Davis

2016 National Ambient Air Monitoring Conference
St. Louis, MO
August 8-11, 2016

Outline

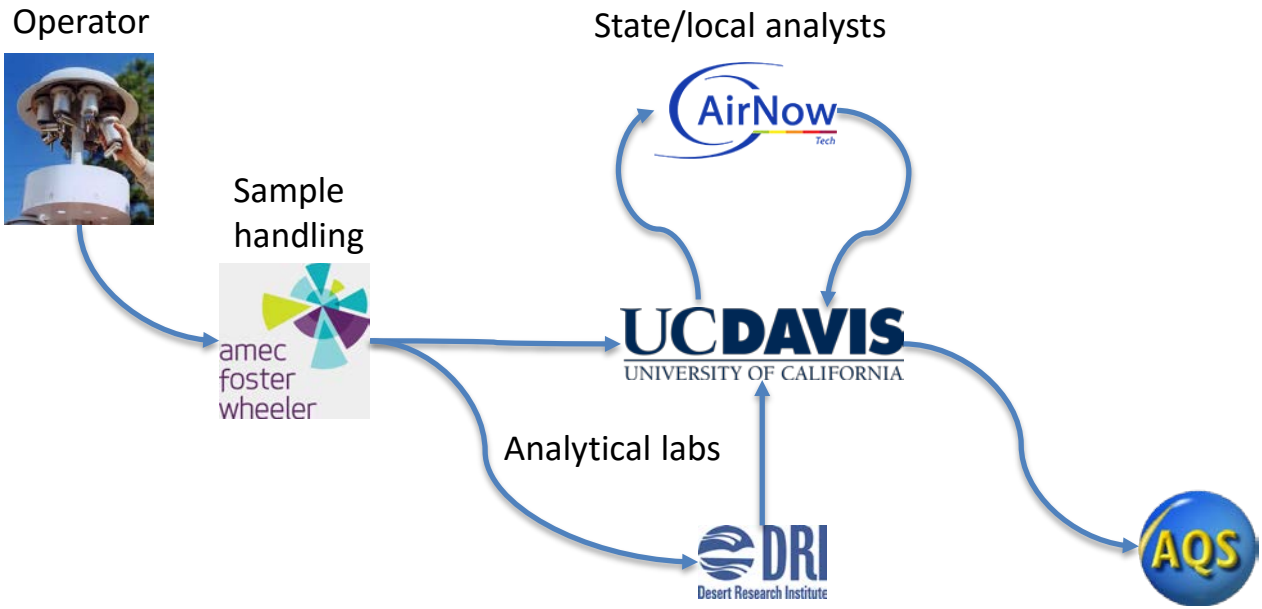
1. Introduction
2. Current CSN Routine Validation
3. Case Studies
4. Future Developments

CSN Measurements

- 33 elements by x-ray fluorescence (XRF) of PTFE filters
  - S, K, Cl*
  - Soil elements (Fe, Al, Si, …)
  - Trace metals (Ni, V, Mg, …)
- Ions by ion chromatography (IC) of nylon filters
  - Cations
    - Ammonium, Sodium, Potassium
  - Anions
    - Nitrate, Sulfate, Chloride*
- Carbon by thermal optical reflectance (TOR) of quartz filters
  - Organic carbon
  - Elemental carbon

* Chloride ion and Cl are being examined for suspected contamination
Cl is invalidated and chloride is not yet being reported

The CSN Validation Process – Many Involved



Our Data Validation Philosophy

- All data should be validated
- Definitive evidence is required to invalidate records
- Do not censor the data
- Revisit and improve checks over time

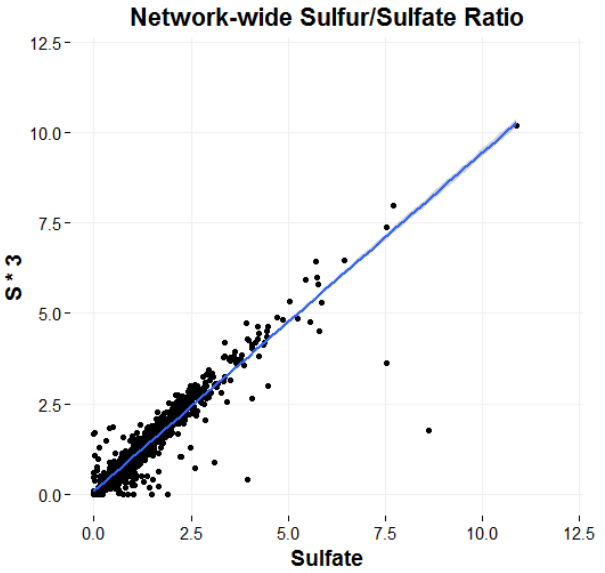
Checks

- Data integrity
- Automated sampler operating data
- Sample shipping conditions
- **Comparison of measurements by different analytical techniques**
- **Filter swaps between sequential dates**

Ratio Analysis – Sulfur/Sulfate

Sulfur (by XRF) correlates with Sulfate (by IC)

- If (3*S)/Sulfate > 1.5 OR (3*S)/Sulfate < 0.66 then flag as outlier (code 5) and examine
- An indication that either XRF or IC may be invalid on this sample date
- However, if no reason for discrepancy can be found, the data are not invalidated



UCDAVIS

Ratio Analysis – Sulfur/Sulfate

Data Month

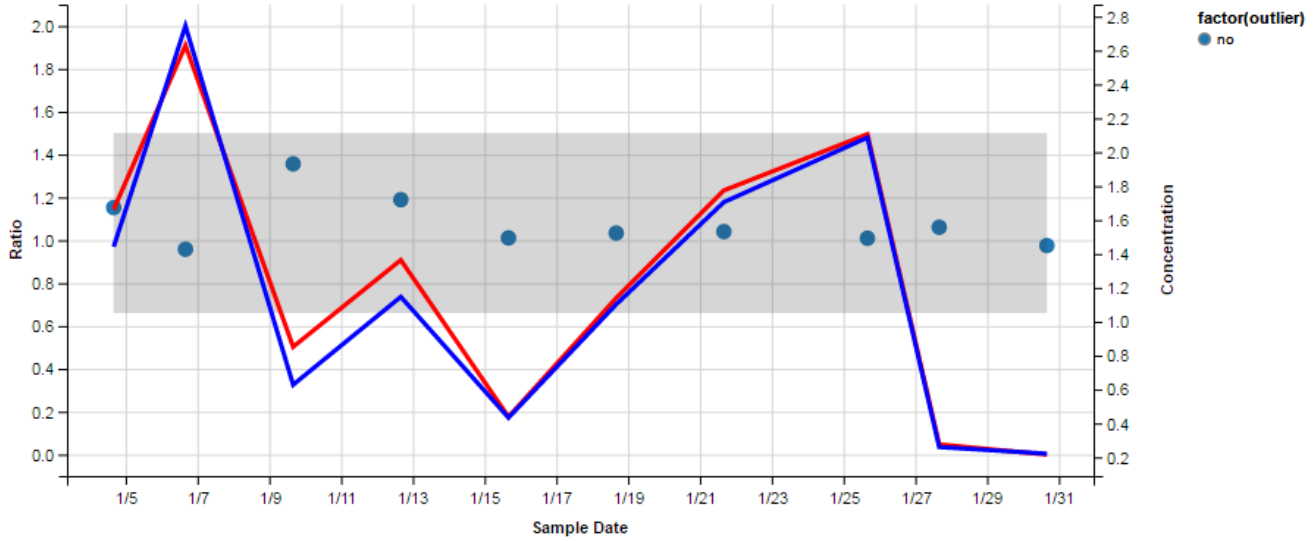
2016-01

Site

46-099-0008-5

Ignore invalid records

Red is S*3, Blue is Sulfate



UCDAVIS

Ratio Analysis – Sulfur/Sulfate

Data Month

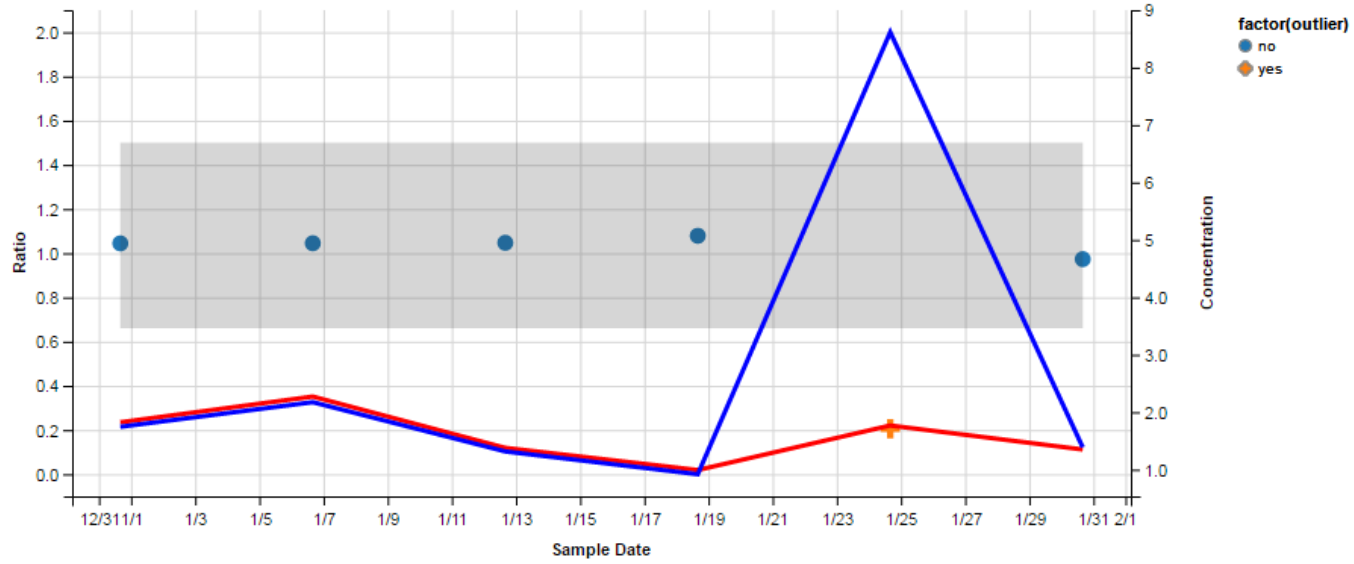
2016-01

Site

42-101-0055-5

Ignore invalid records

Red is S*3, Blue is Sulfate

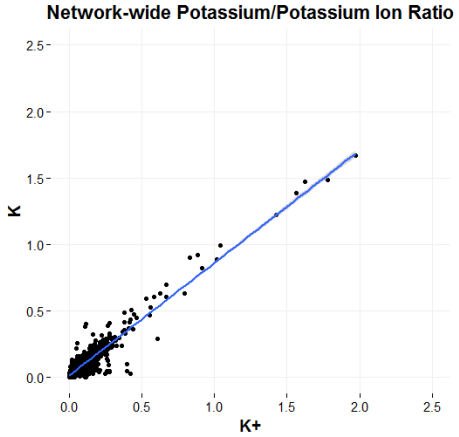


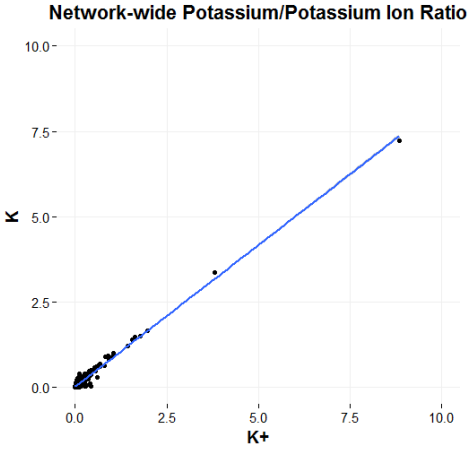
UCDAVIS

Ratio Analysis – K/K^+

When well above detection limits, Potassium/Potassium ion ratios also show good agreement thus far

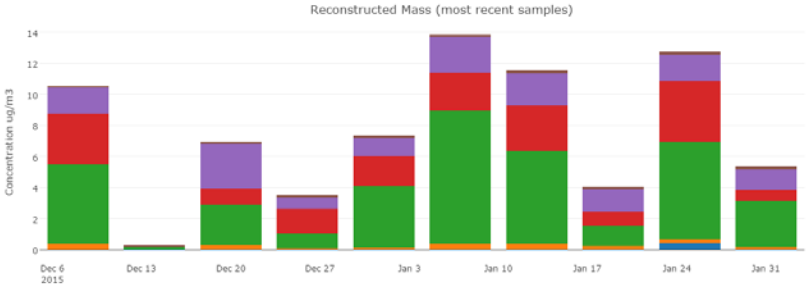
- $K^+ = 1.16 * K$
- $R^2 = 0.96$
- No flagging criteria established yet

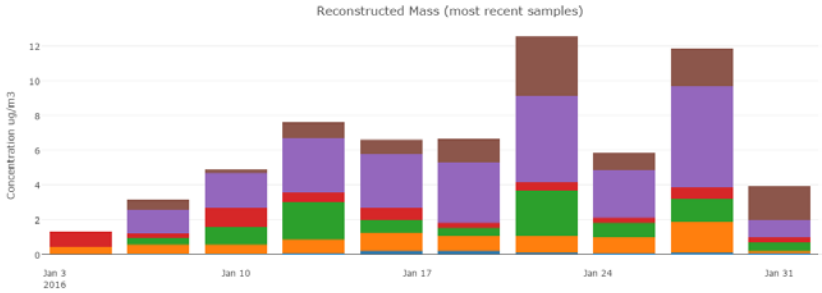




Reconstructed Mass

An estimate of total $PM_{2.5}$ mass based on speciated measurements and assumptions of chemical composition





- Soil
- Organic matter
- Ammonium sulfate
- Ammonium nitrate
- Elemental carbon
- Sea salt

Reconstructed Mass - Calculation

We reconstruct total $PM_{2.5}$ mass from chemical speciation measurements as:

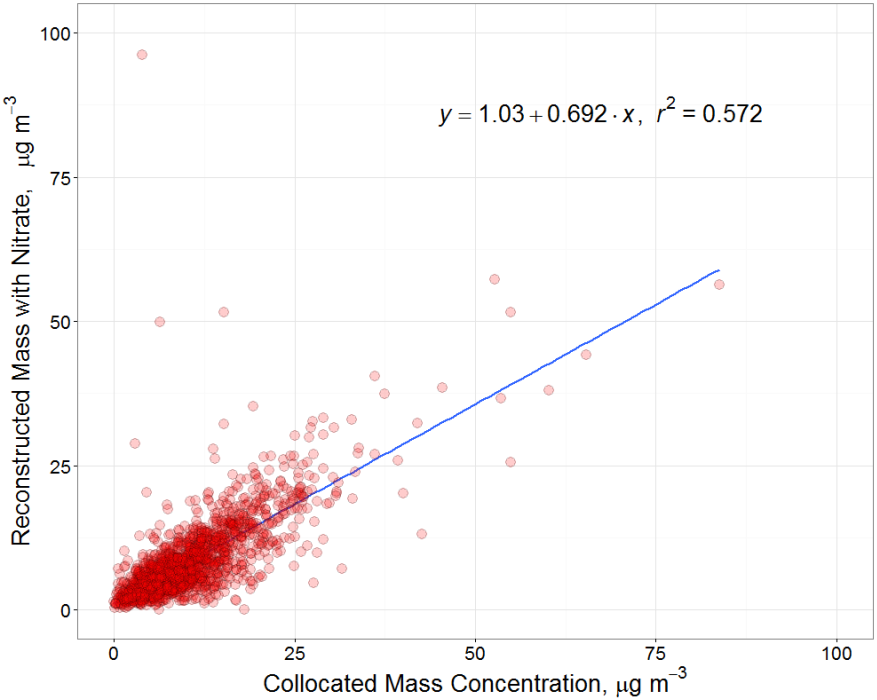
$PM_{2.5}$ = Ammonium sulfate + Ammonium Nitrate + Elemental Carbon + Organic Mass + 1.8 Chlorine + Soil

where

- Ammonium Sulfate = 4.125 * Sulfur
- Ammonum Nitrate = 1.29 * Nitrate
- Organic Mass = 1.4 * Organic Carbon
- Soil = 2.2 Al + 2.49 Si + 1.63 Ca + 2.42 Fe + 19.4 Ti

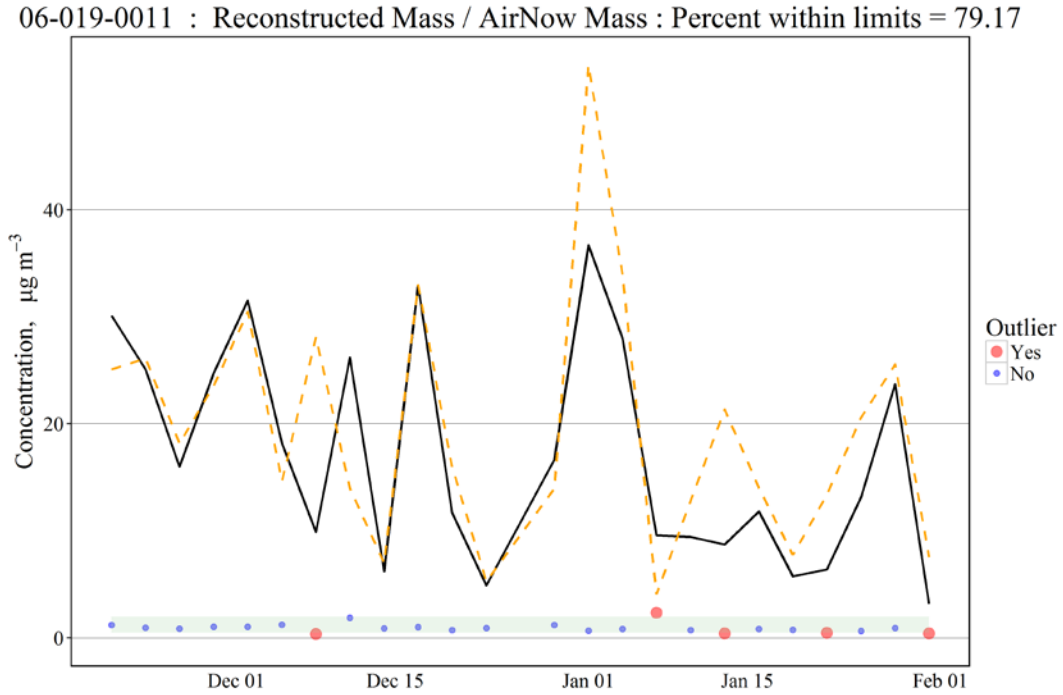
This reconstructed mass is compared to $PM_{2.5}$ mass from nearby FEM and FRM monitors

Reconstructed Mass – Network Wide



Reconstructed Mass

Yellow = AirNow
Black = Reconstructed

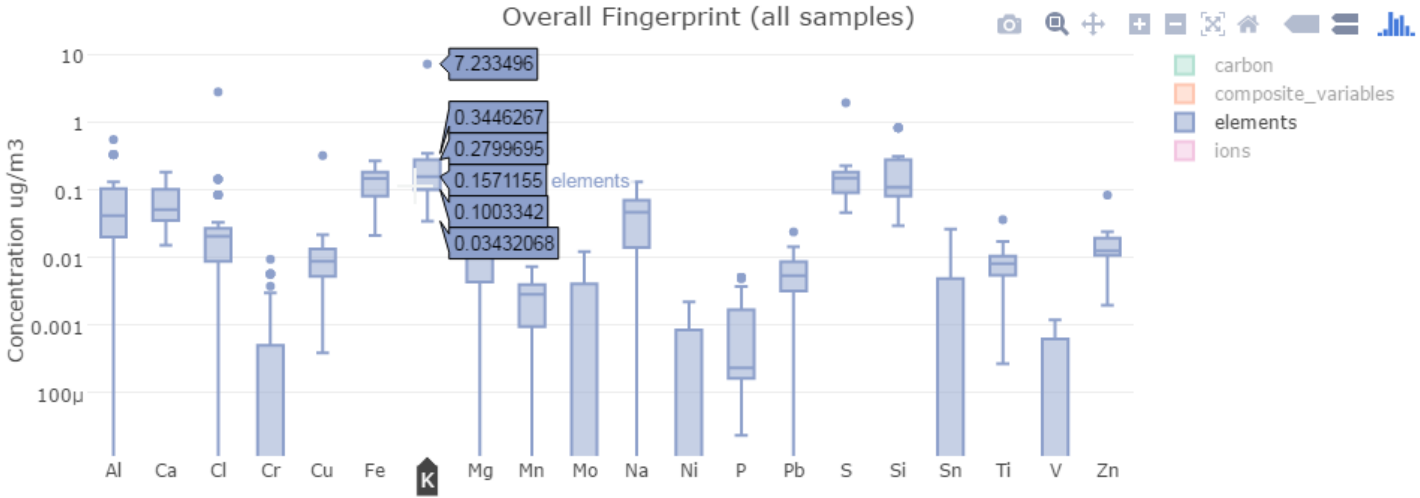


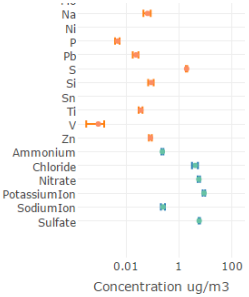
Case studies

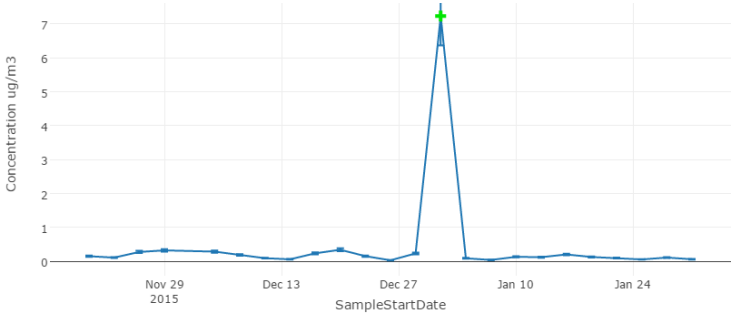
High potassium in Phoenix

High sulfate in Philadelphia

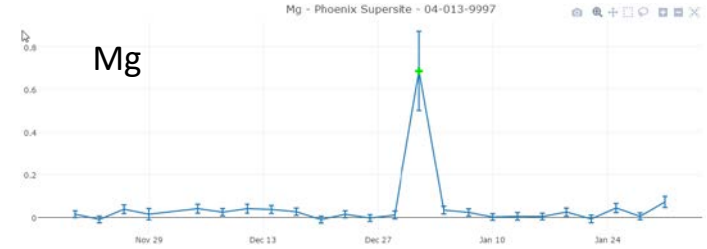
Phoenix – High Potassium

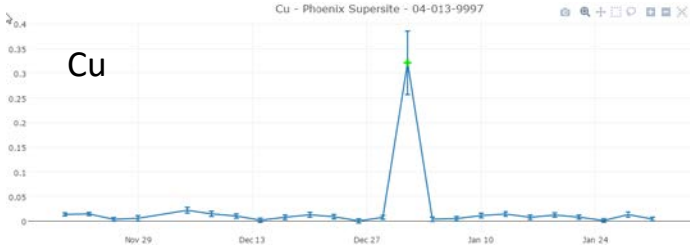


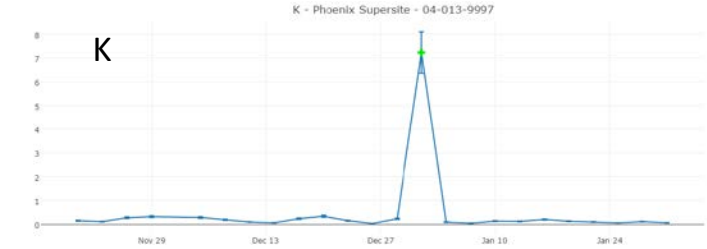


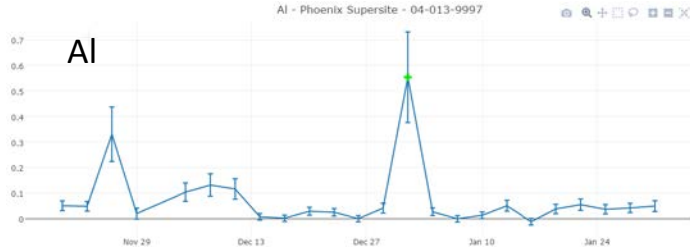


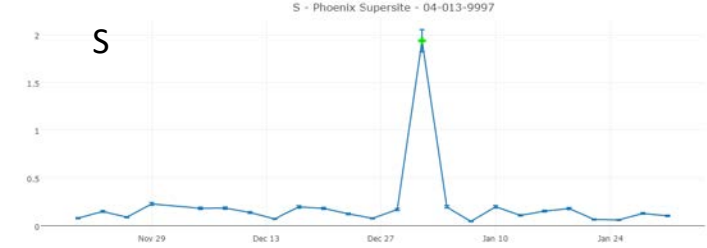
Phoenix – January 1

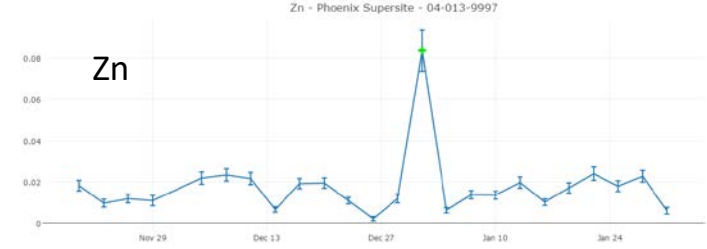


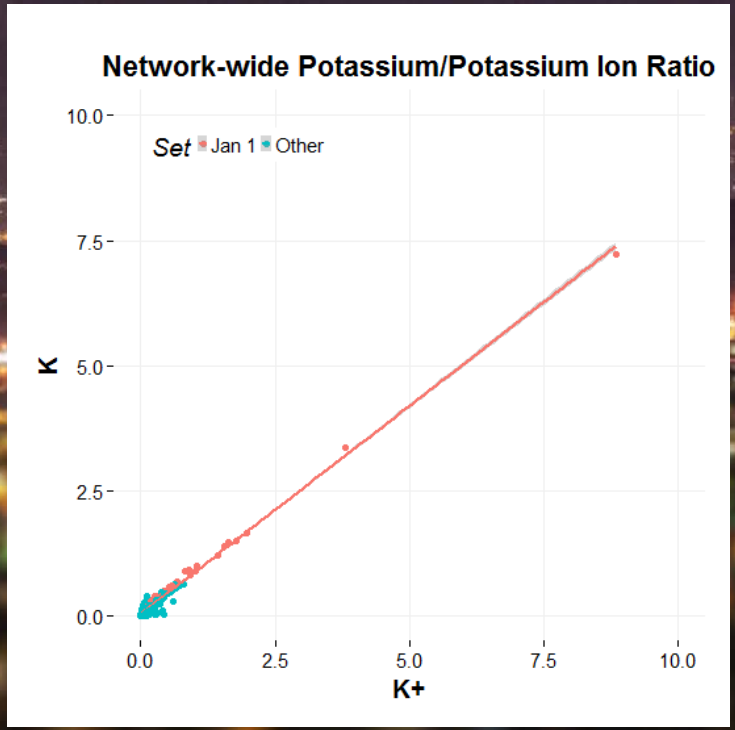








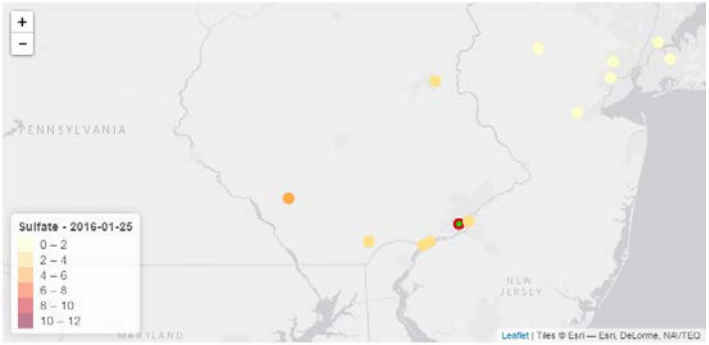


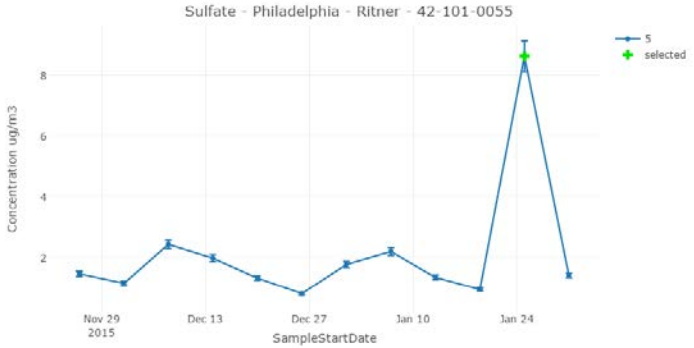


Courtesy of Imgur

Philadelphia – High Sulfate

- Sulfate is high compared to neighbors and nearby days
- Is this reasonable?





Ratio Analysis – Sulfur/Sulfate

Data Month

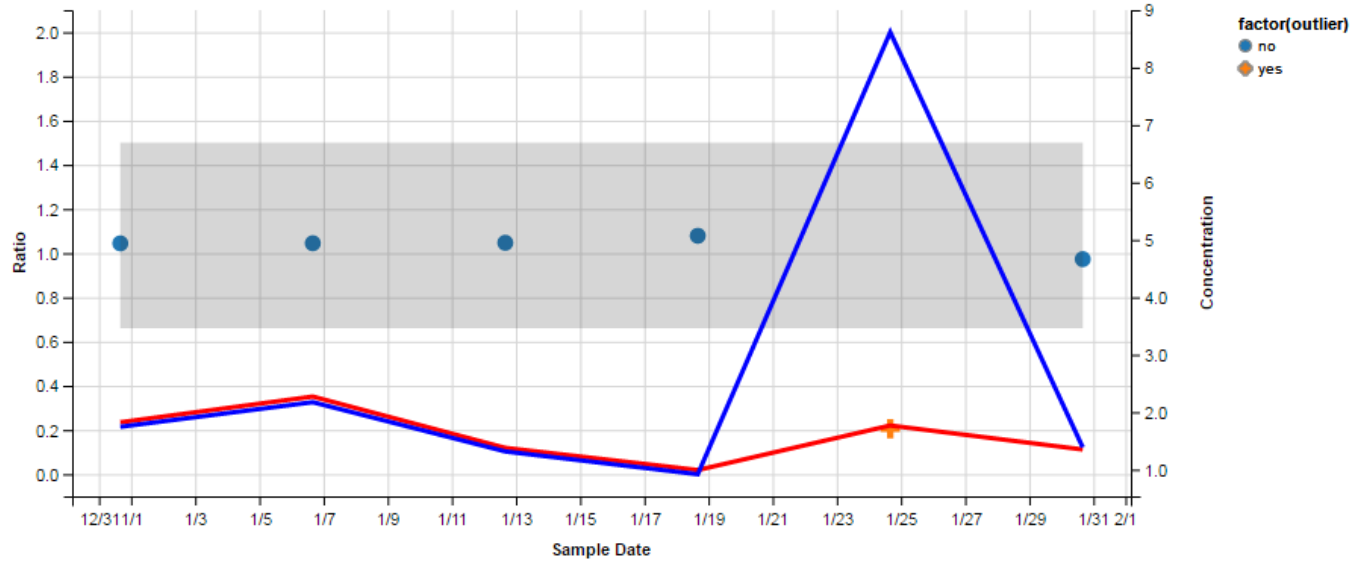
2016-01

Site

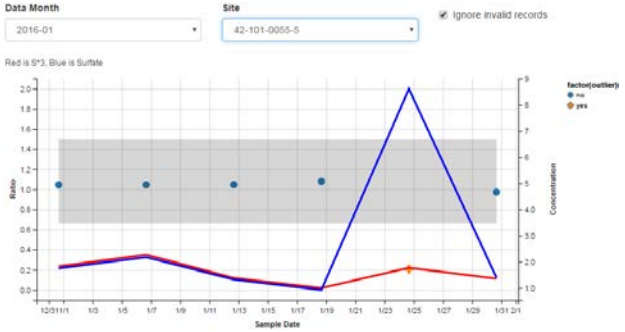
42-101-0055-5

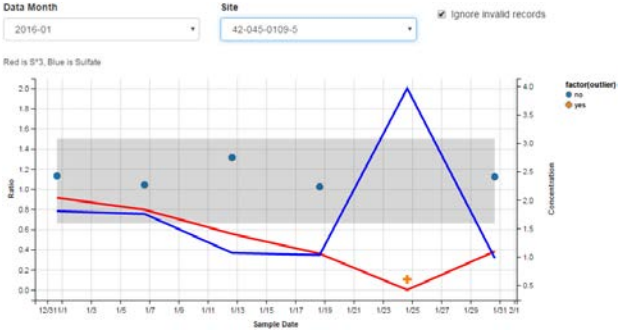
Ignore invalid records

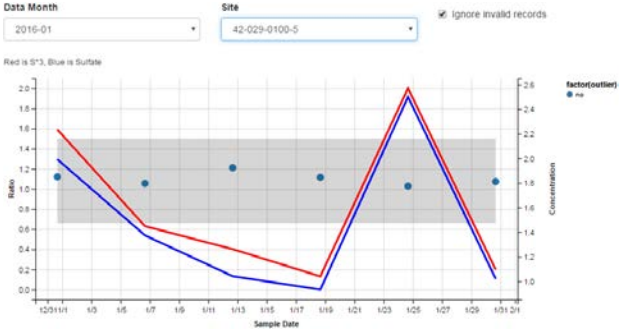
Red is S*3, Blue is Sulfate

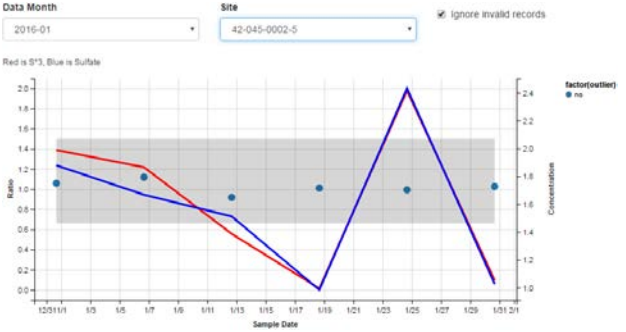


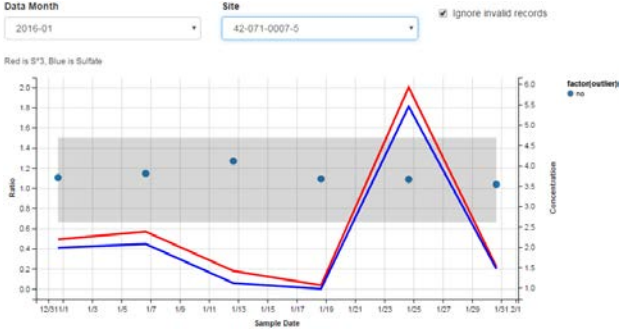
Southeastern Pennsylvania – Sulfur/Sulfate Ratios

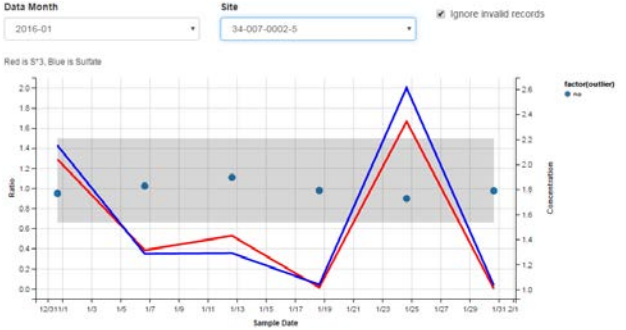






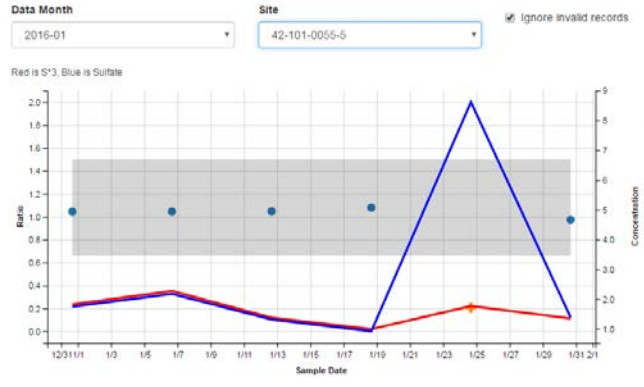


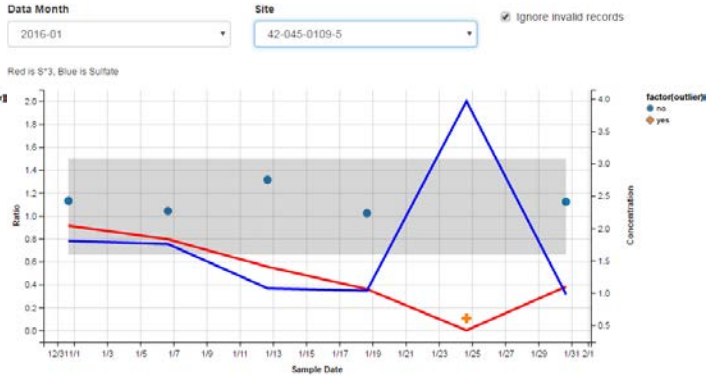




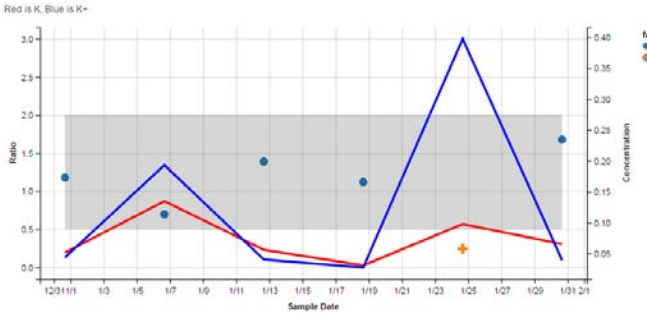
Two Odd Sites – Sulfur and Potassium Ratios

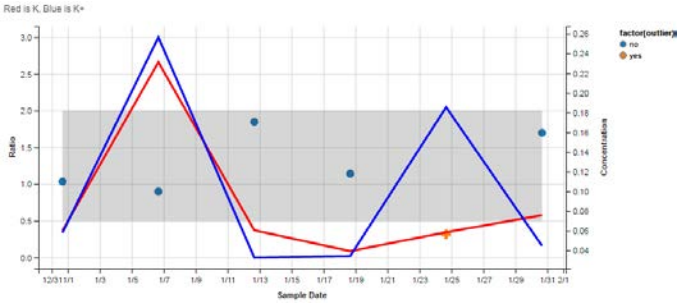
Sulfur/Sulfate





Potassium/Potassium Ion





Philadelphia – High Sulfate

- It would appear that elements were low for two sites on January 25
- We requested reanalysis for both XRF (elements) and IC (ions)
- Reanalysis returned very similar results

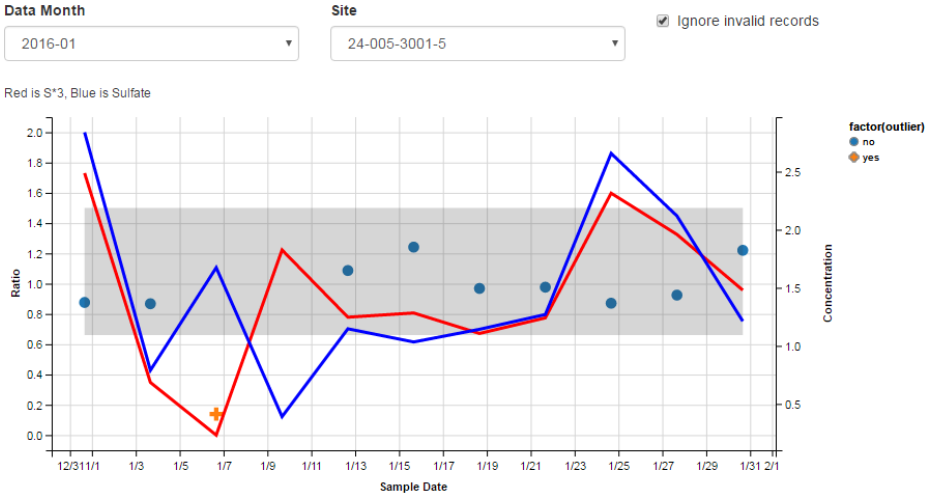
Actions

- Added qualifier flag for values outside of sulfur/sulfate ratio
- Did not invalidate results
- Continued investigation
    - How frequently does this occur?
    - Is this isolated to specific sites or conditions?
    - What are the potential mechanisms?

Possible Filter Swap between Consecutive Dates

Ratio plots can indicate a swap may have occurred

- Swaps can occur in the field or the lab
- By swapping dates, data may look better
- But need evidence to adjust values
- Instances are examined as potential swaps
- Comments added to samples

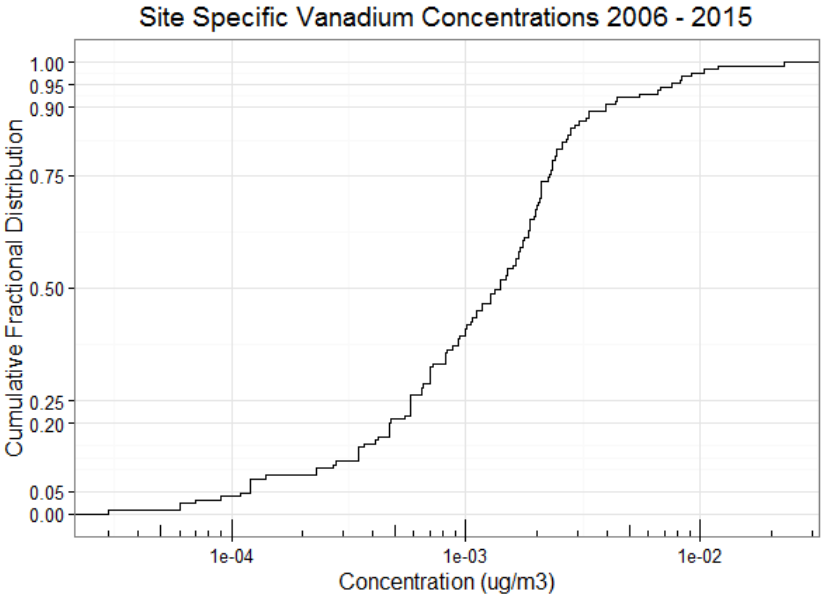


Improvements in Progress

- Historical context – percentile checks
- Spatial outliers
- Optical Transmissometer measurements vs. TOR carbon
- Feedback with state and local validators

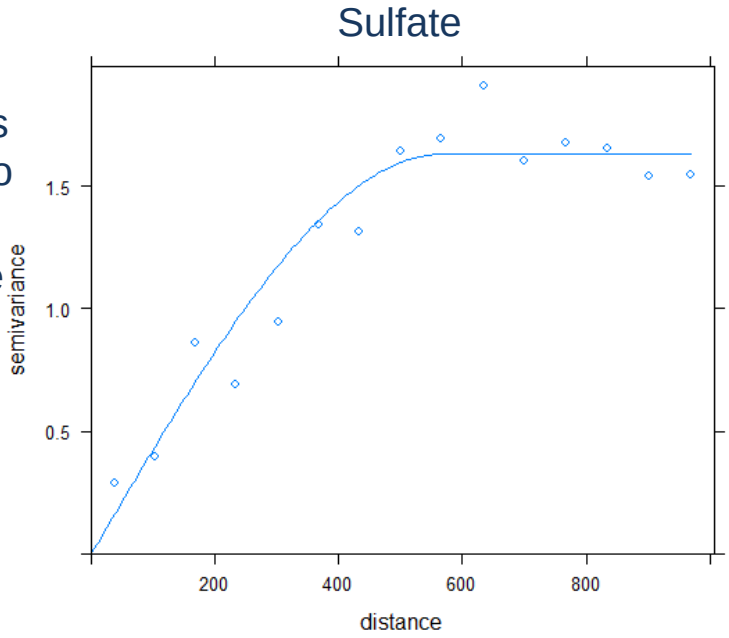
Historical Outliers

- We are developing a database of the CSN historical archive
- Will allow us to routinely compare measured values with the historical range
- Can automatically flag and examine samples that are outside the norm for the site



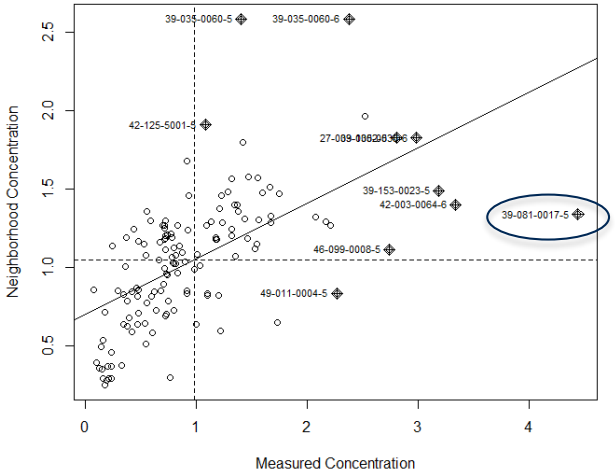
Spatial Outliers

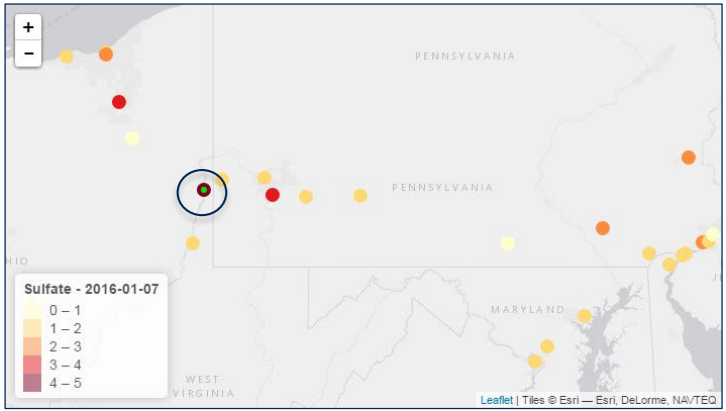
- Many pollutants are spatially autocorrelated (concentrations at one site tend to be similar to neighboring sites)
- We can highlight and examine sites that are different from their neighbors for autocorrelated species



Spatial Outliers

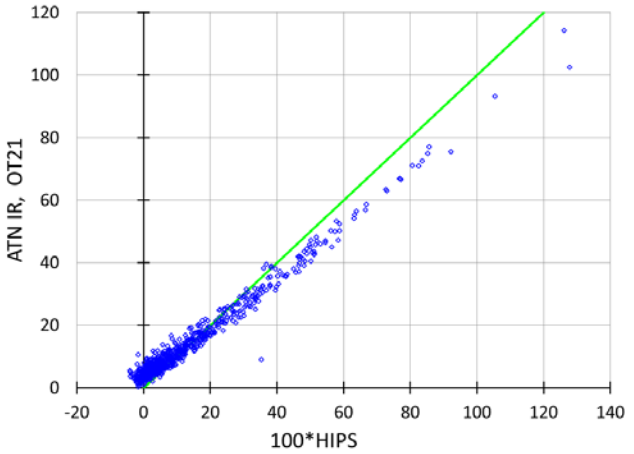
We can highlight and examine sites that are different from their neighbors for autocorrelated species





Optical Measurement vs. TOR Carbon

- We can corroborate elements with ions using sulfur/sulfate and potassium/potassium ion ratios
- Currently, we cannot corroborate carbon with anything
- A new Transmissometer is now in testing and will begin operational analysis soon
- For details, see Warren White's talk:
  - Filter Transmittance Measurements: Experience with IMPROVE and Plans for CSN
  - 9:30am Wednesday - Chemical Speciation Technical Session



Feedback from State and Local Experts

Many of you have site specific, state specific, CSN specific or general knowledge that could improve our validation process.

We're interested in hearing your ideas.

sraffuse@ucdavis.edu – Sean Raffuse – Data Management Lead

njspada@ucdavis.edu – Nick Spada – CSN Data Validator

CSNSupport@sonomatech.com – Support list watched by STI, EPA, and UC Davis

Thank you!

Challenges Unique to CSN

- We are removed from sample operators
  - Local conditions or sampler issues may not be known
- No data history with the current process
  - Although we have historical data, it was handled by a different lab
- No optical Transmissometer™ measurements yet reported
  - Optical measurements added with the new contract so we don't have much data to compare against the TOR carbon measurements
- Sample loadings are often low
  - Many measurements are below detection

Data Integrity – Import Validation

When electronic data are ingested, the import script performs several checks

- Is the file well formed (correct columns found)?
- Do all filter records match with existing sites?
- Are number columns numbers; date columns dates?
- Do records with the same ID already exist in our database?

Data Integrity – Typo Checks

Operational data from field data sheet are manually entered and mistakes are inevitable (~ 1%)

- Date Checks
  - Sample Start Date – Intended Use Date = 0
  - Sample End Date – Sample Start Date = 24 hours
  - For inconsistent dates, request sample handling lab to check original sheets for data entry typos
    - Sample Start Date = 2015-12-20
    - Sample End Date = 2015-11-21
    - Intended Use Date = 2015-11-20
- Flow Rates
  - Flow rates outside of the normal range on samples still marked as ‘valid’ are cross-checked with sample handling lab

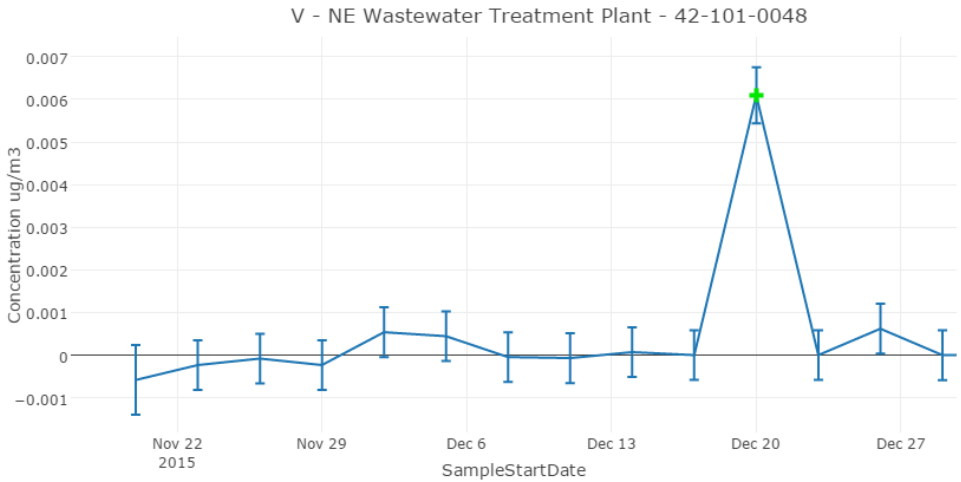
Automated Flagging

Qualifier flags are added to records that fail certain criteria

- Sample delivery temperature < 4°C
  - Flag as TT (Transport temperature out of spec.)
- Measured concentration below MDL
  - Flag as MD (Value less than MDL)
- Negative measured concentration
  - Flag as 9 (Negative value detected – zero reported)
- Flow rate
  - Flag as AH (Flow rate average out of spec.)
  - This is terminal (i.e., the sample is invalid)

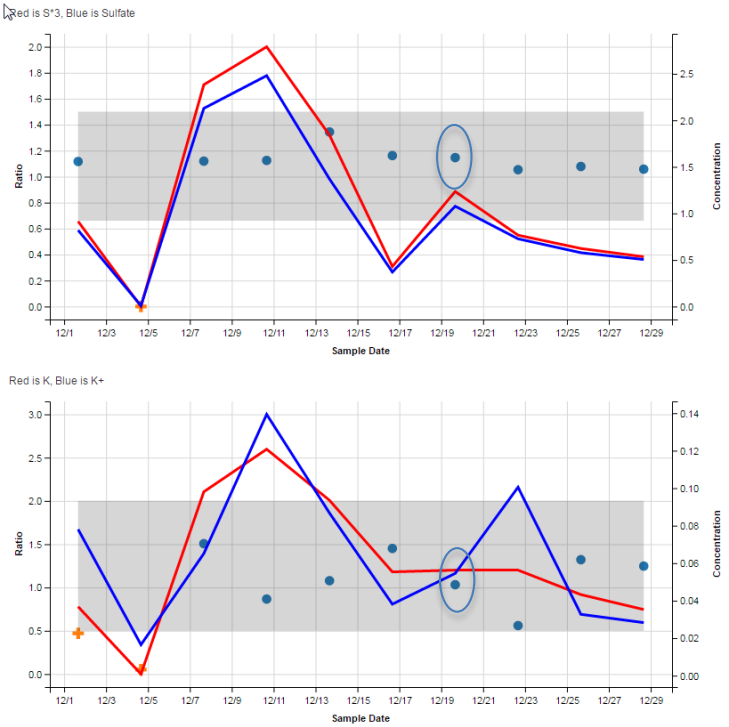
High Vanadium

- Here we see a single V concentration, well above the others
- Is it real?



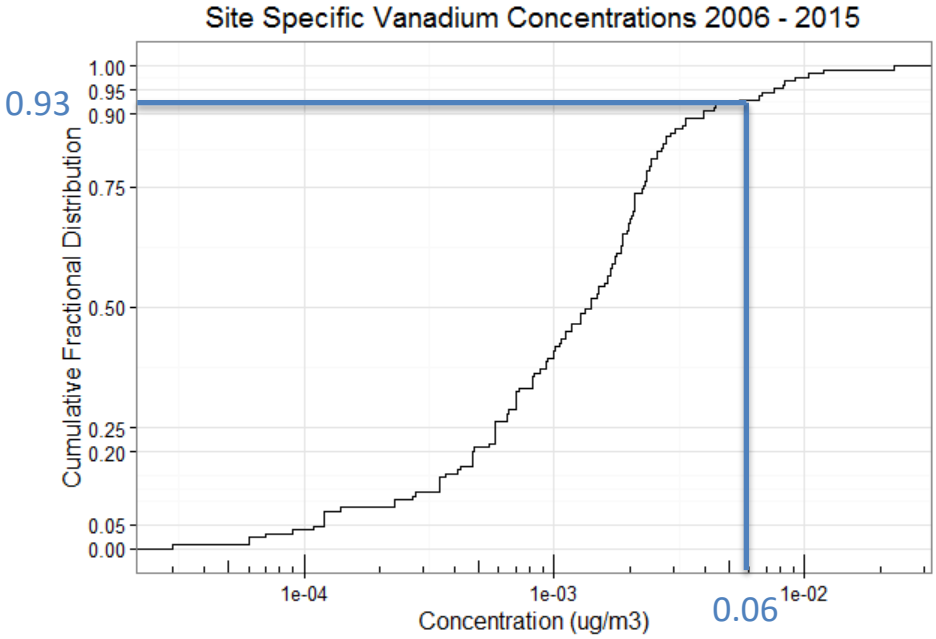
High Vanadium – Checking the Ratios

- Both sulfur/sulfate and K/K^+ are as expected for the date in question (Dec 20)
- No sample-wide problem detected



High Vanadium – Historical Range

Does the measured value have precedence at the site?



The measured value is at about the 93rd percentile for this site over the last 10 years.

High, but within the normal range.

High Vanadium – Broader Context

Adding January data makes the high value in December look more reasonable.