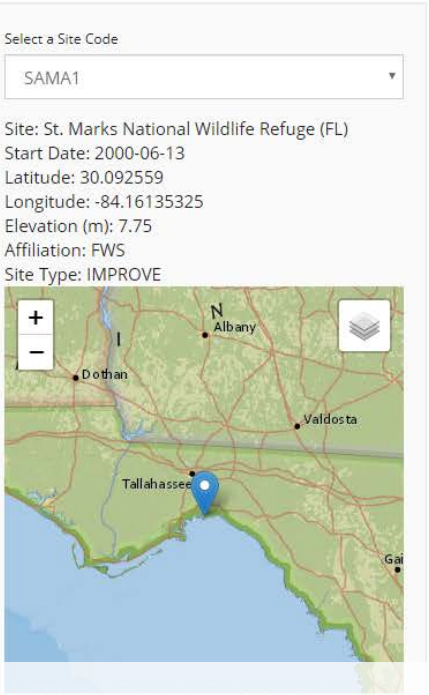
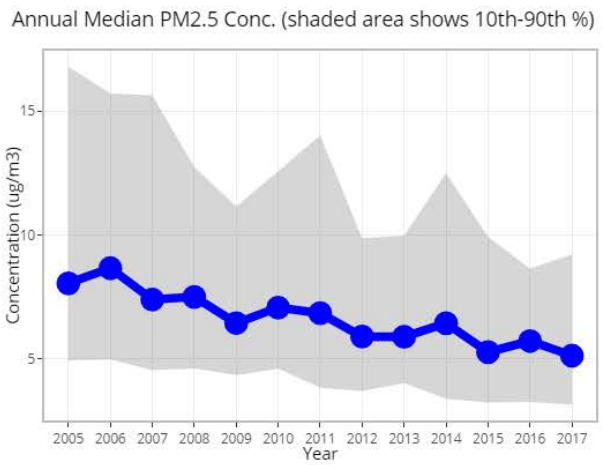
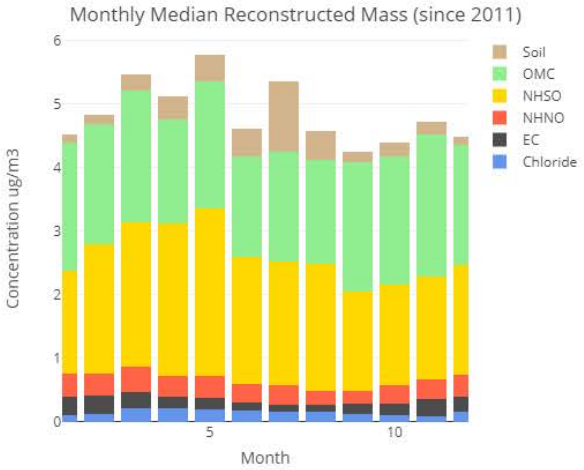
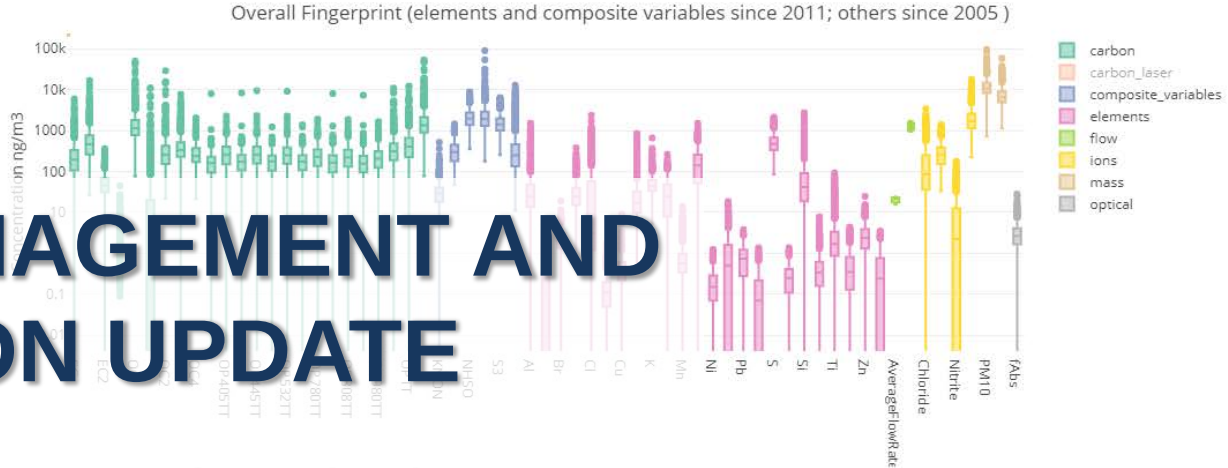








DATA MANAGEMENT AND VALIDATION UPDATE

SEAN RAFFUSE
UC DAVIS AIR QUALITY RESEARCH CENTER

IMPROVE Steering Committee Meeting
Fort Collins, CO October 16, 2018

What’s New in Data Management and Validation

Last year I presented our transition to new databases and software.

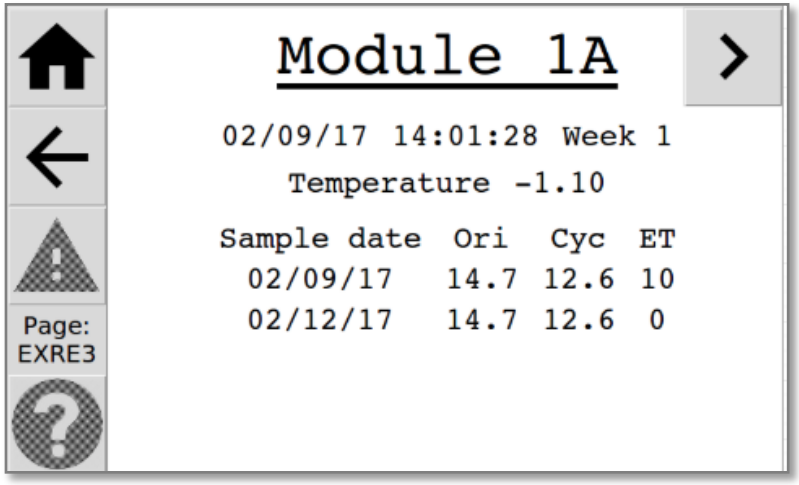
Today I will discuss some of the larger new features to support

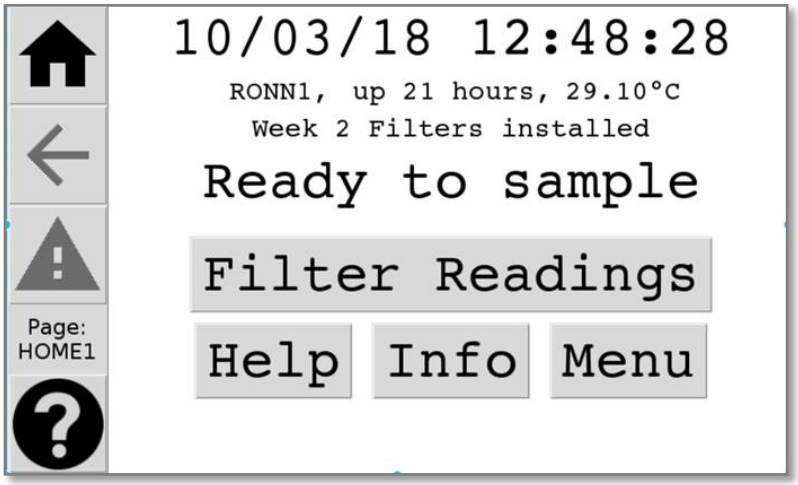
- The redesigned controller
- The new automated weighing system
- Improvements in data validation
- Improvements in data analysis/exploration

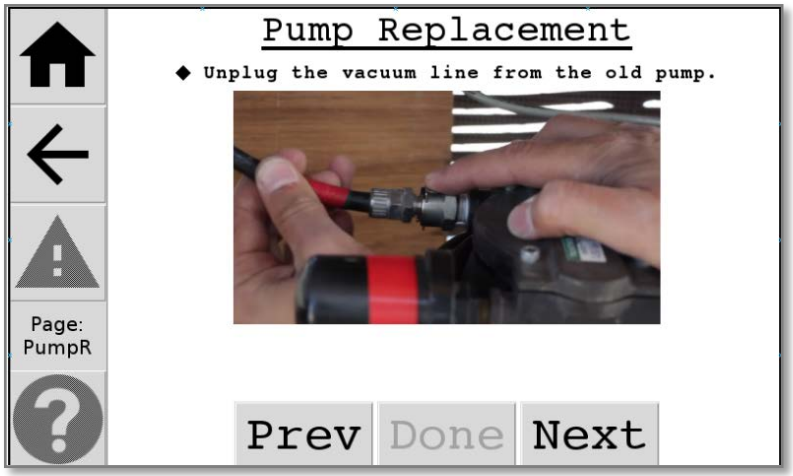
New Controller Software

Software for the new controller was designed in house from the ground up

Touch screen UI with guidance for the operator



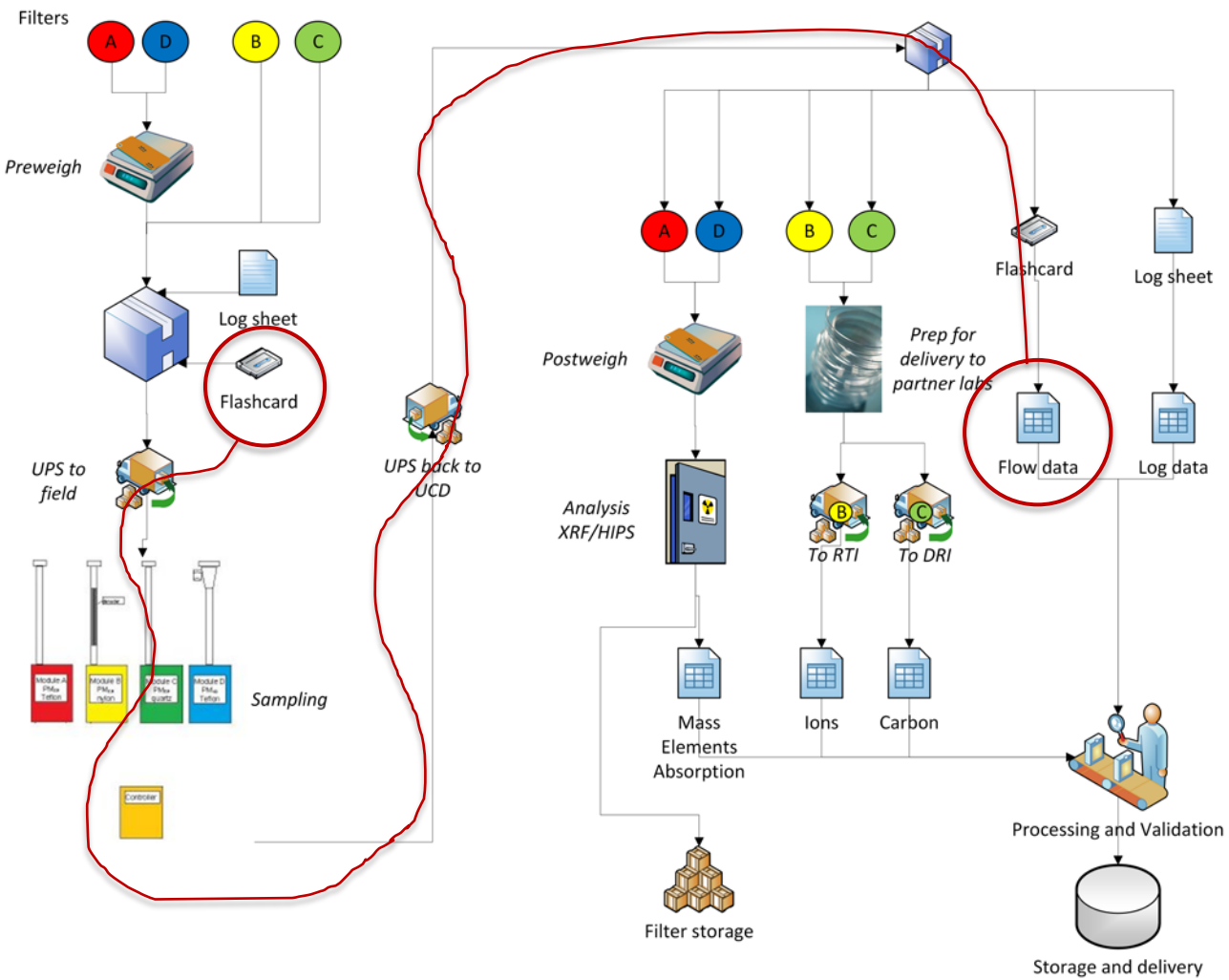




New Controller Data Management

The old way:

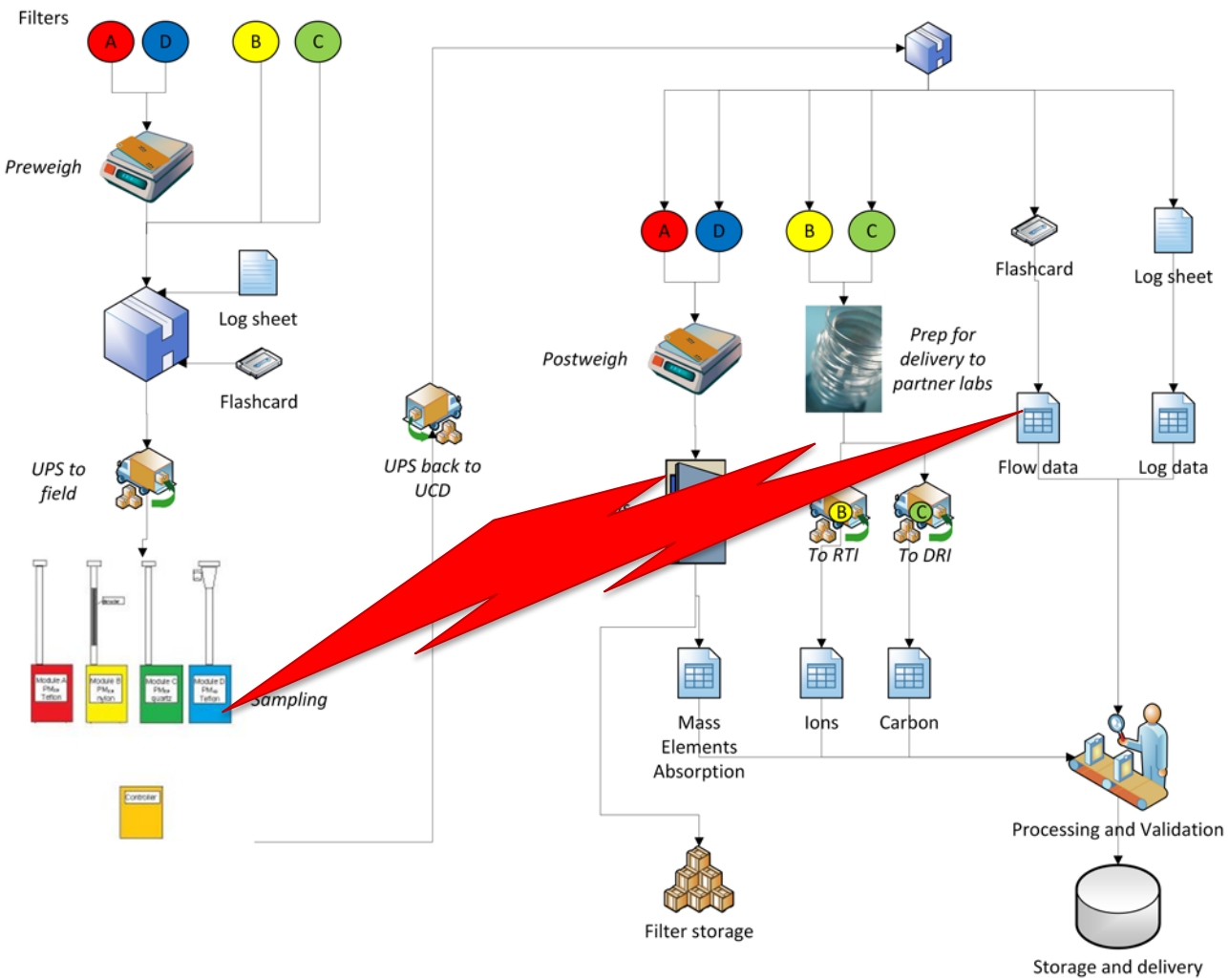
- Flow data collected on flashcards
- Flow data not seen by the lab until weeks (or more!) after sampling
- Potential for many lost samples before issues are caught



New Controller Data Management

The new way:

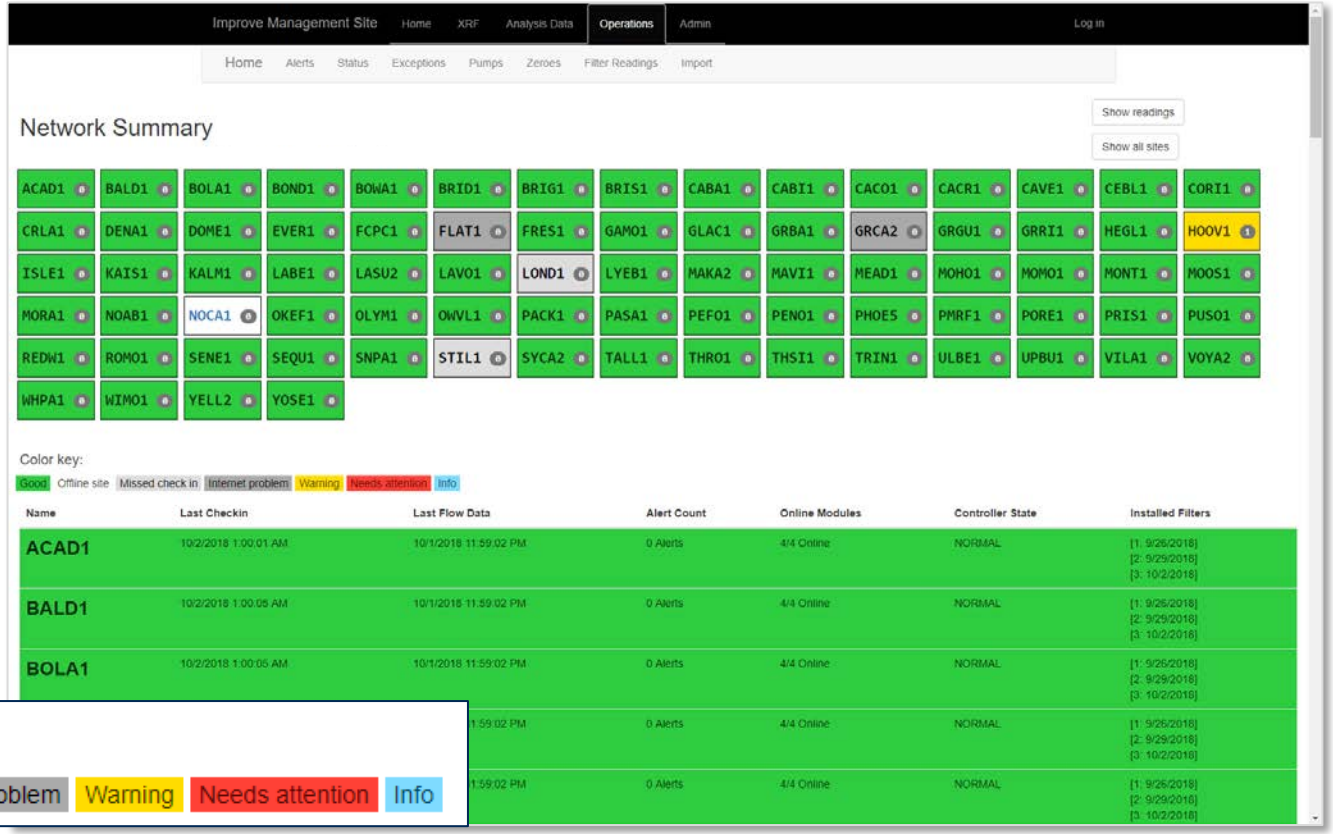
- Sites connected via Wi-Fi, cellular, or satellite
- Flow data delivered overnight
- Many tools to review flow information and respond quickly



Tracking Sampler Health – Network Summary

At-a-glance view of all sites with the new controller

Color-coded to show where attention may be needed



Color key:
Good
Offline site
Missed check in
Internet problem
Warning
Needs attention
Info

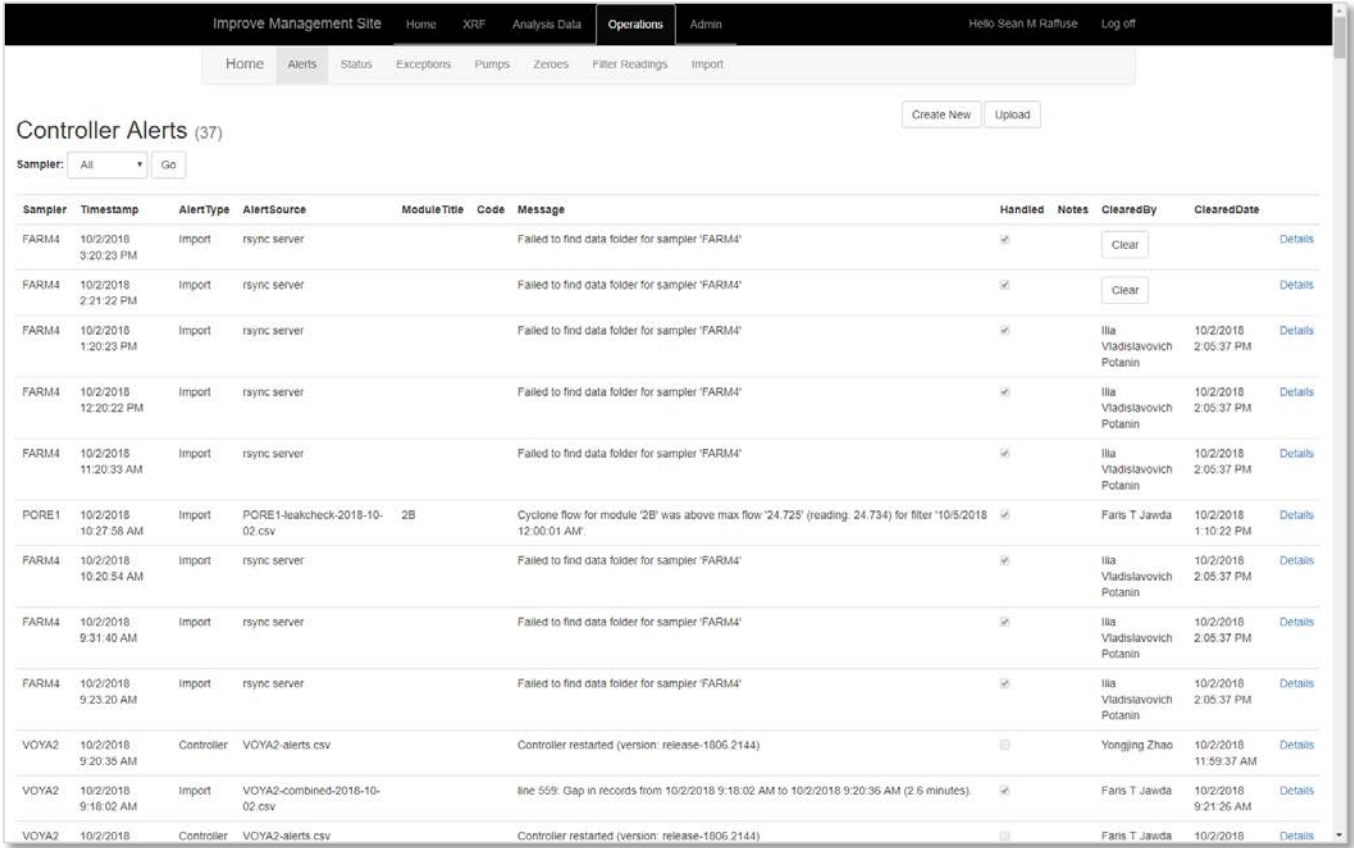
Tracking Sampler Health – Network Summary

Thumbnails of flow readings and temperature for quick scan

Tracking Sampler Health – Controller Alerts

Nightly (or faster) harvesting of logs allows for review by the field group each morning

Responders can mark an alert as cleared and add notes for further follow up



Tracking Sampler Health – Controller Alerts

Generated on controller:

- Controller restarted
- Orifice min pressure below 5.0 during initial test
- Ebox failed to communicate on CAN
- Ebox went offline
- Out of range module ids
- Ebox had sensor read errors

Generated on Sawmill (log processing server):

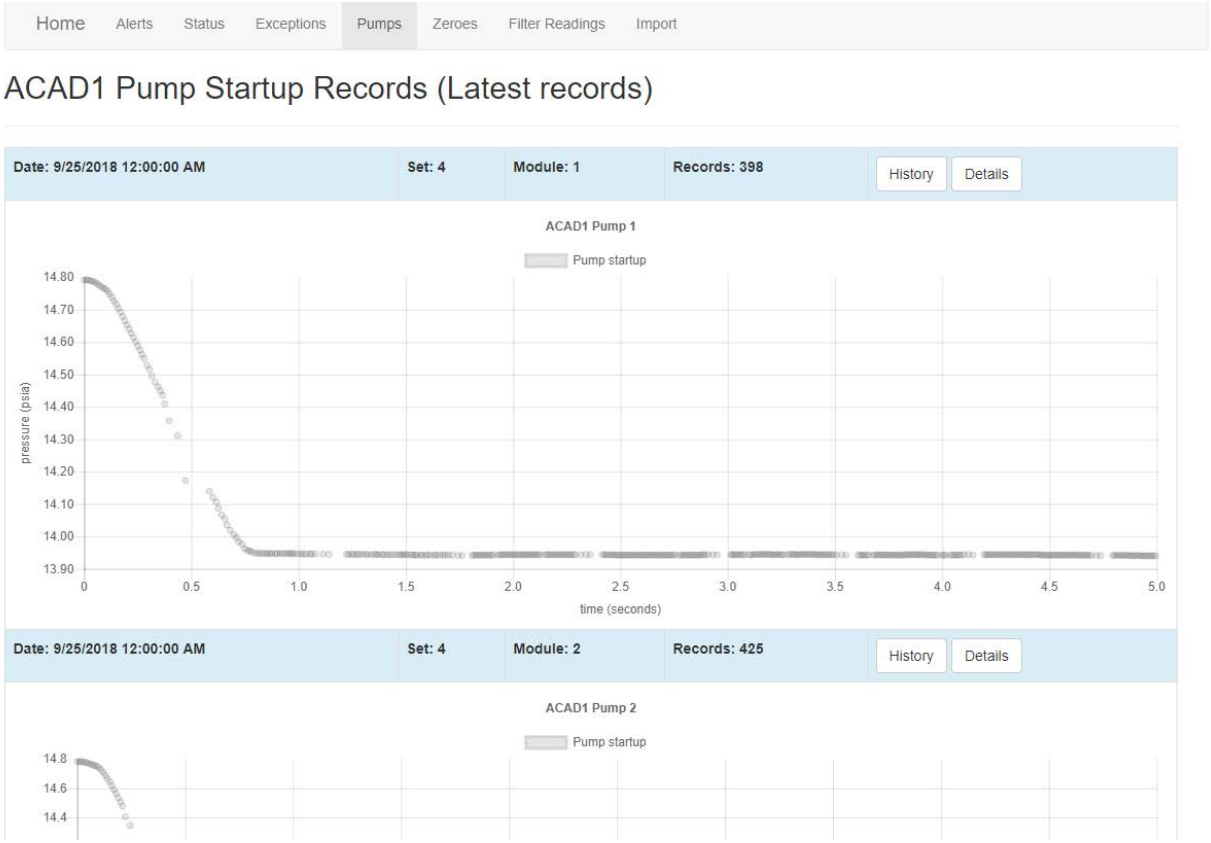
- Failed to find data folder for sampler
- Found two folders for sampler

Generated on import to database:

- Ebox is offine
- Controller timezone offset does not match database
- Ebox position was removed/added to controller config file
- Hostname changed
- Cyclone/Orifice was stuck for [#] hours
- Gap in flow records
- Cyclone/Orifice pressure below threshold
- Cyclone/Orifice pressure above threshold
- Time went backward in flow data
- Max vacuum below min/above max flow
- Cyclone/Orifice flow/pressure below min/above max

Tracking Sampler Health – Pump Starts

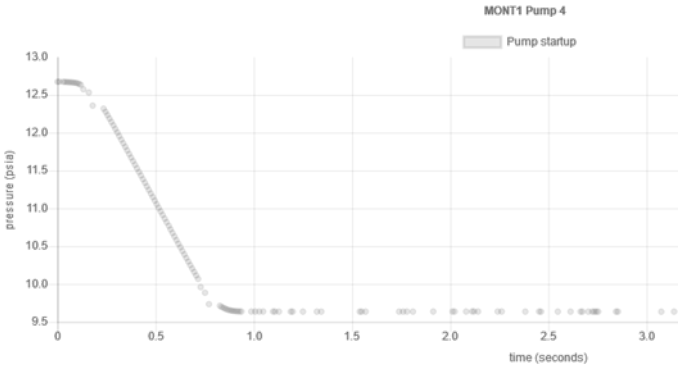
- On startup, pressure is recorded at high time resolution
- Slow response is an indication of poor pump health needing attention
- Plots now available on demand



Pump Health Case Study – Monture (MONT)

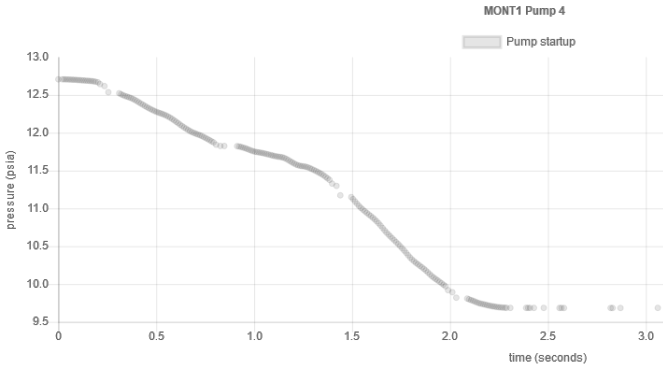
9/11

Normal



9/18

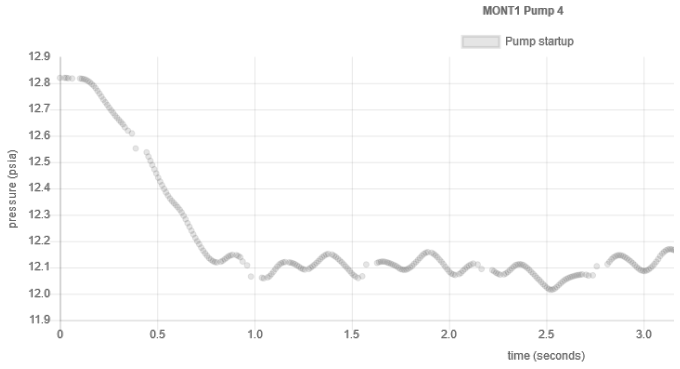
Slow start



9/25

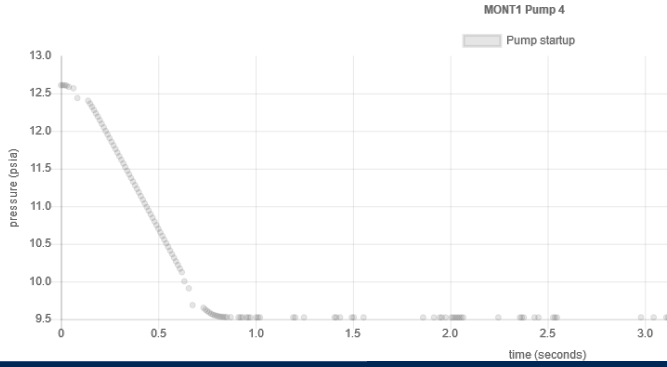
Weak pump

Alerted by system



10/2

Pump replaced



Automated Weighing System

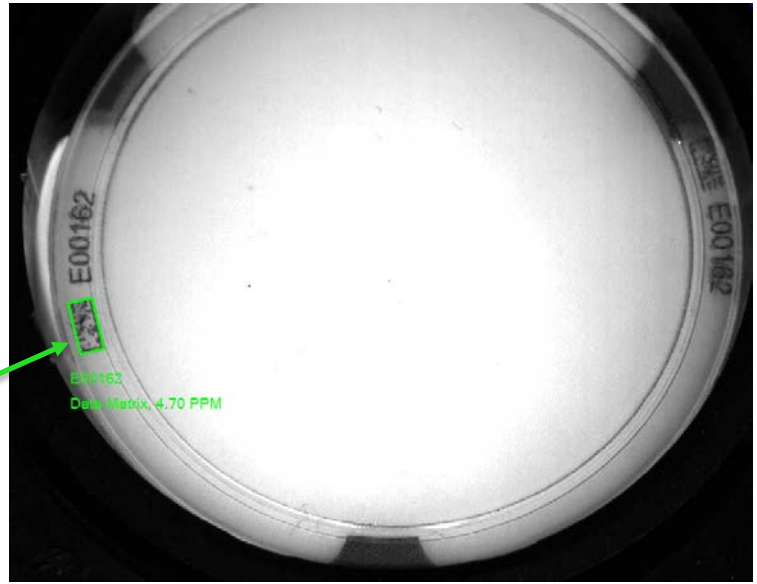
We’ve wanted to move to barcode scanning for a long time. The automated weighing system (Luna), which can load 400 filters in a batch, made it a necessity.

Barcodes printed directly on the filter media are incompatible with HIPS and XRF.

Thus, we needed a code printed on the ring, but that means the “matrix code” is tiny.

Automated Weighing System

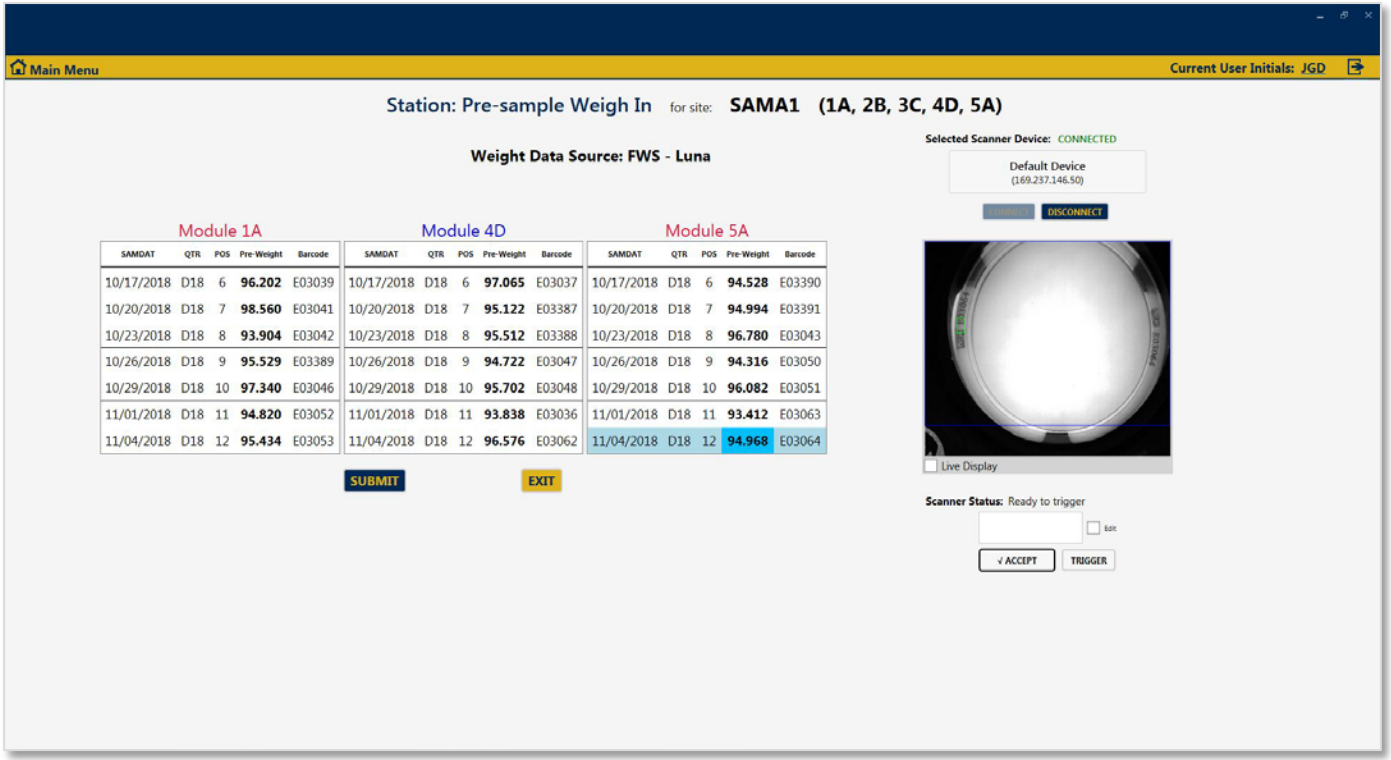
- Modern, high-end scanner
- Scanner trained to conditions of our setup
- Custom-built jig to center filter in tray



1 x 2 mm (!!)

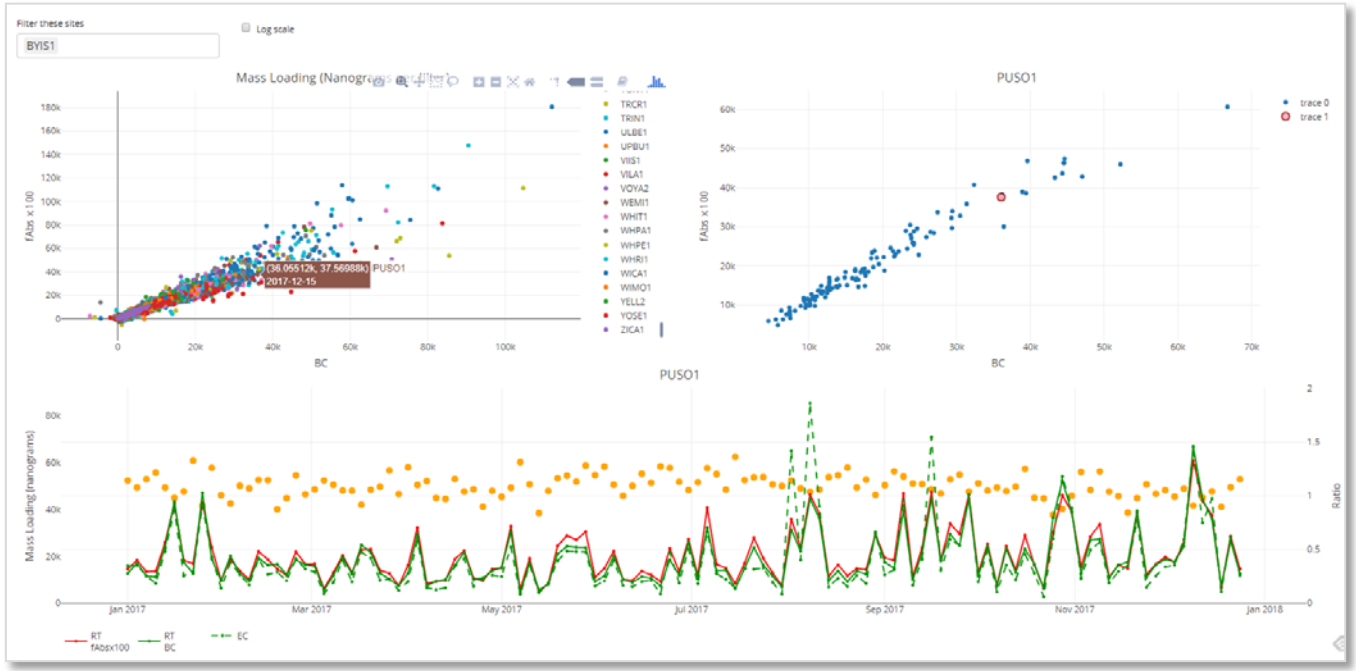
Automated Weighing System

- Modifications to our Lab application to support Luna
- Additional support screens currently in development



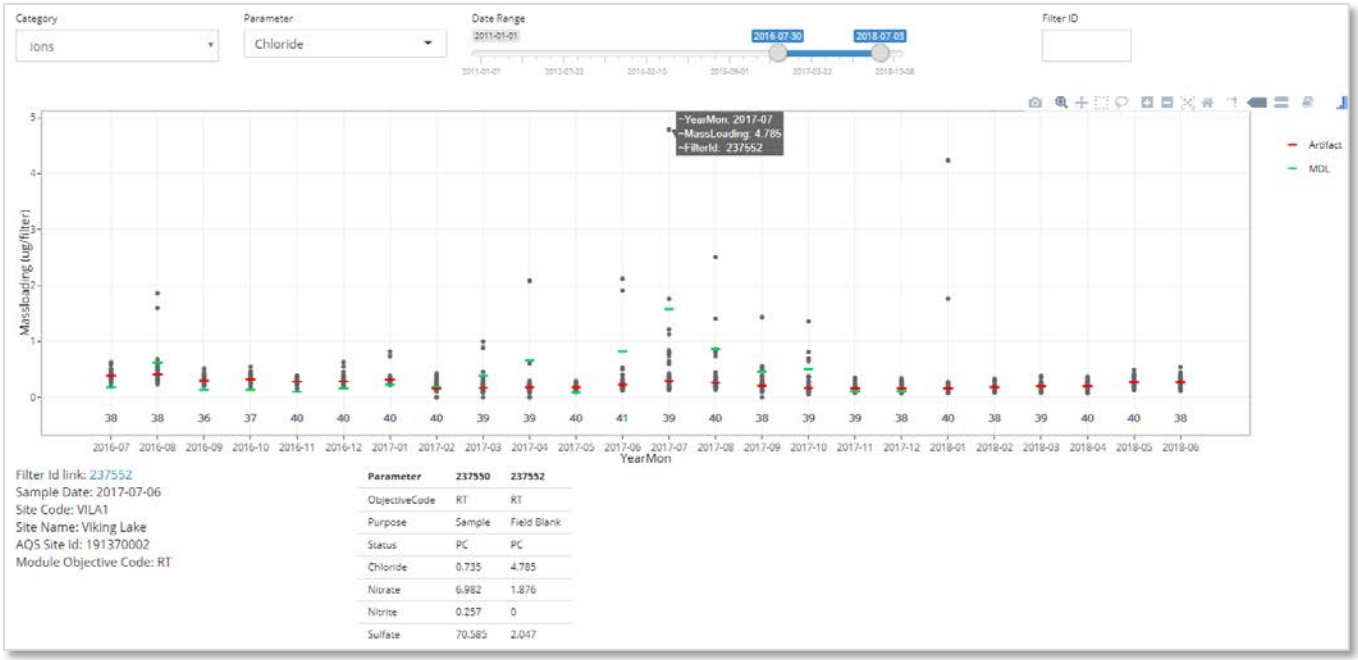
New Data Validation Views – BC from TOR vs. fAbs

- BC can be estimated from laser readings provided by TOR
- BC_{TOR} is often better correlated with fAbs from HIPS than EC
- Interactive view for cross-filter comparison



New Data Validation Views – Field Blanks

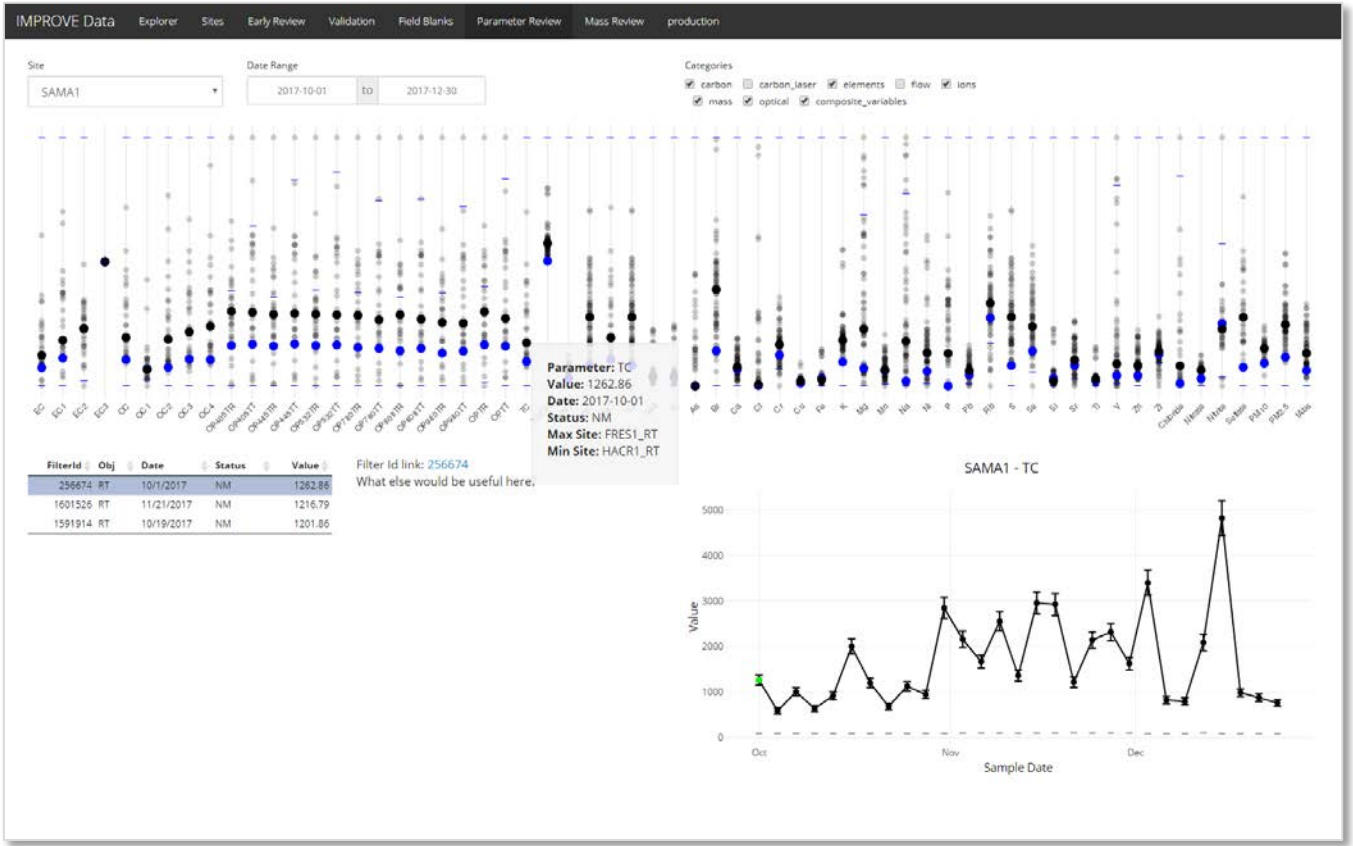
- Automated view of all field blanks and resulting artifact and MDL
- Compare loadings between FB and sample to assess swaps
- Track FB loadings over time



New Data Validation Views – Parameter Review

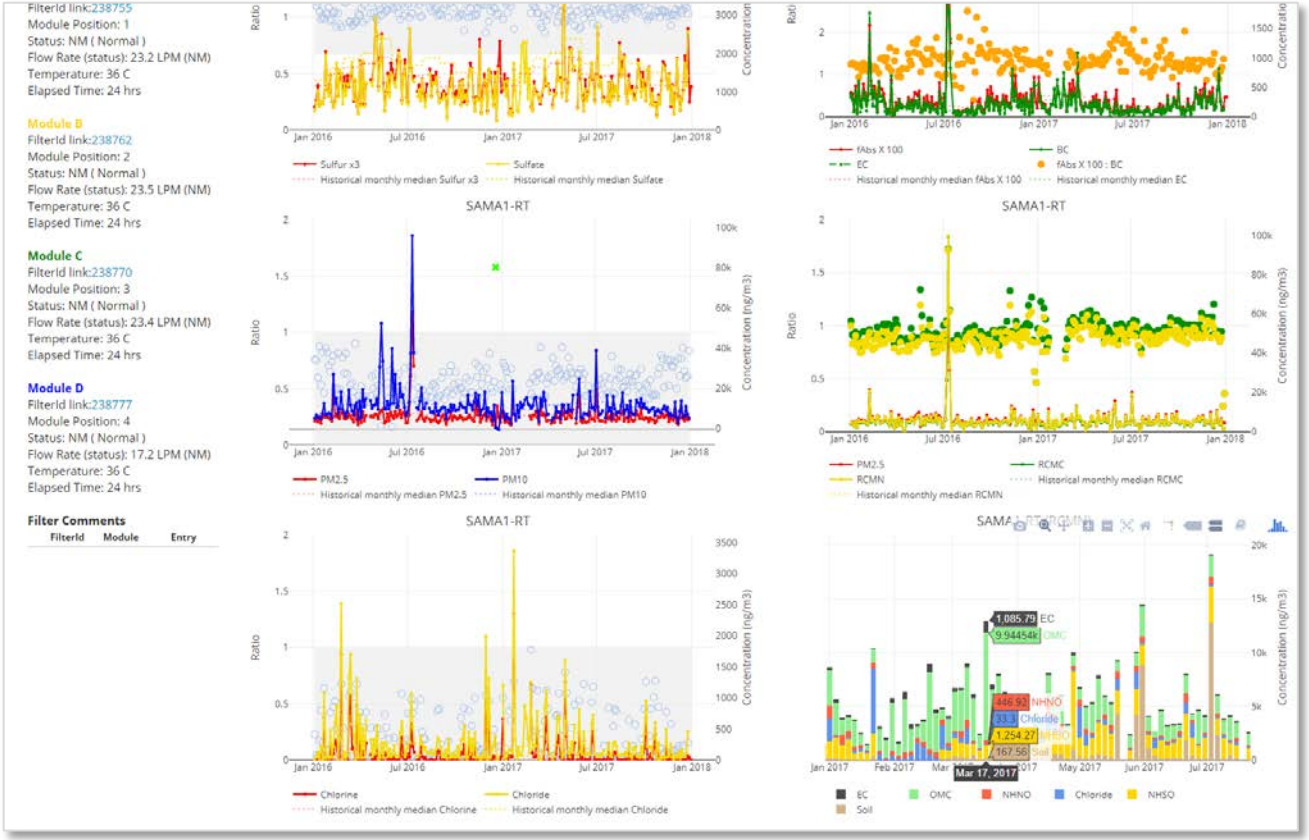
A way to assess all parameters at a site against network norms

Shows individual values, site median, and network median



New Data Validation Views – Reconstructed Mass

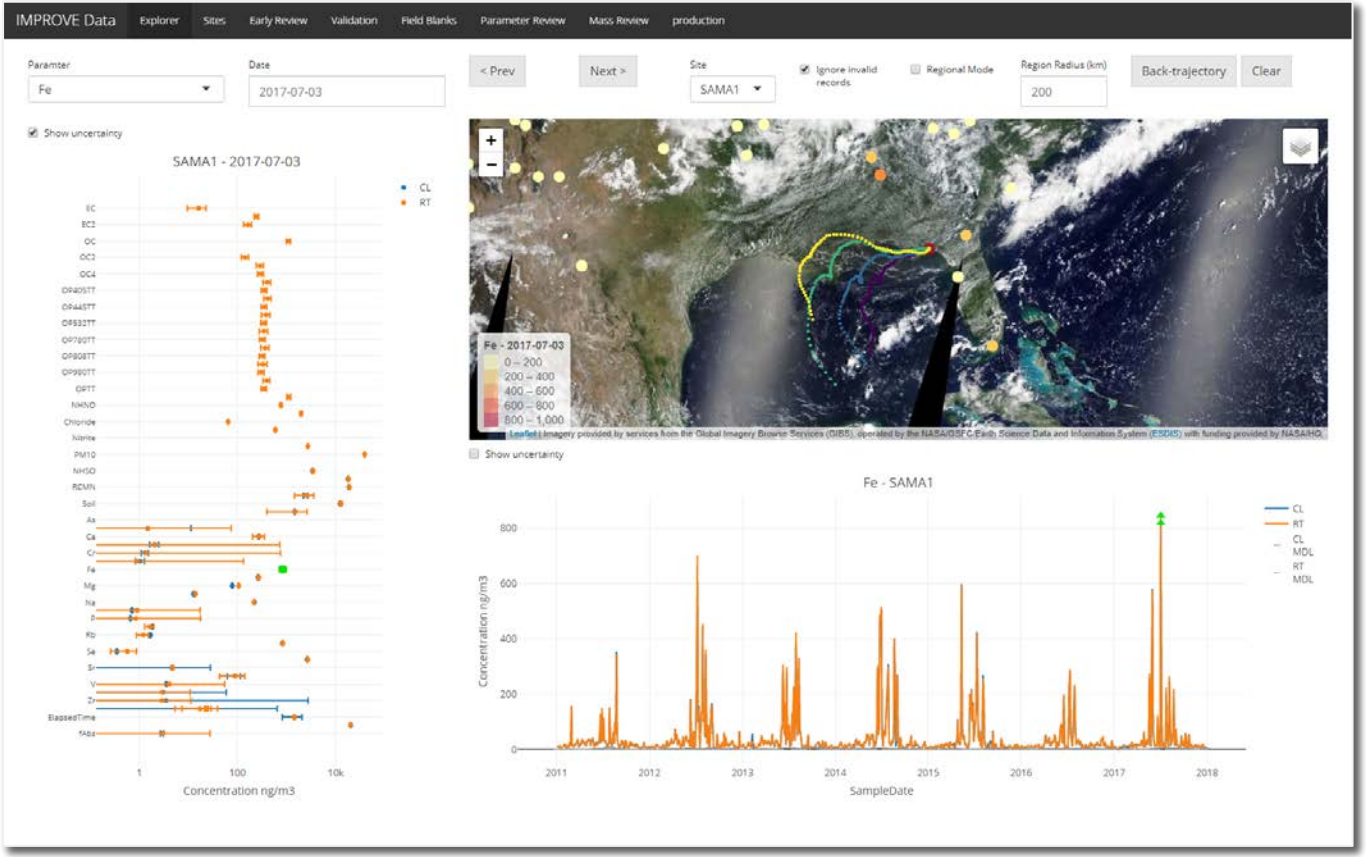
Daily site reconstructed mass components



New Exploratory Views – Backward Trajectories

Data explorer links views across space, time, and parameter

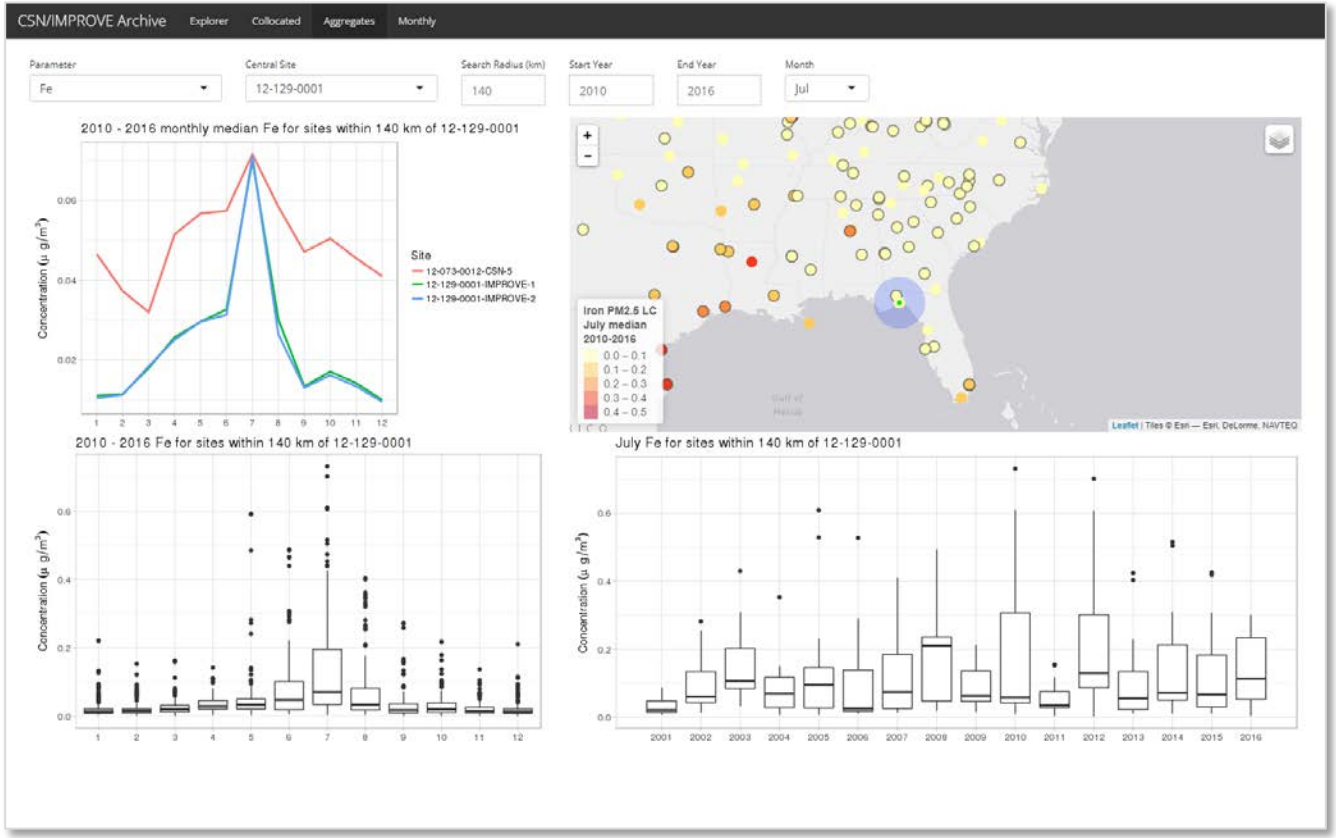
Now run back trajectories from any site on a sample day to assess transport

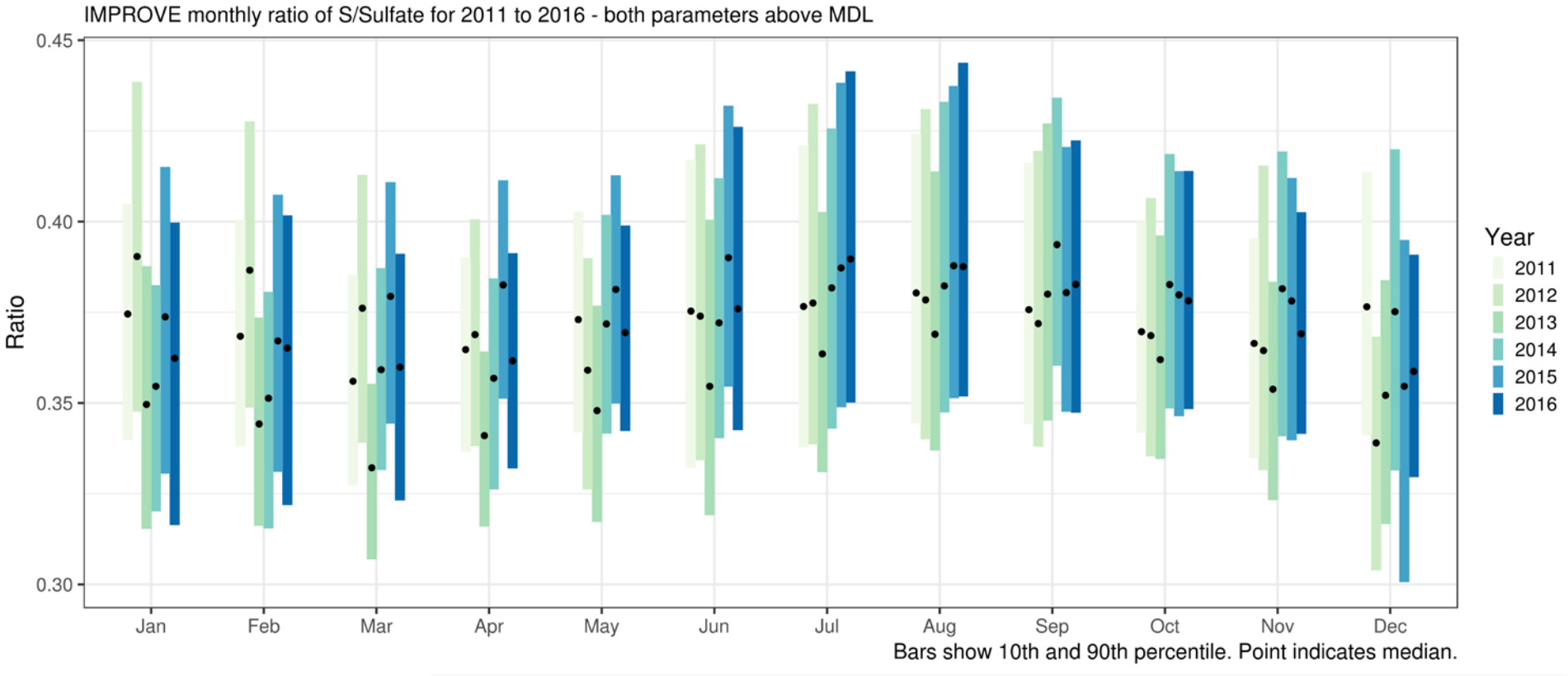


New Exploratory Views – Aggregates

Aggregates multiple sites and dates

Includes IMPROVE and CSN sites



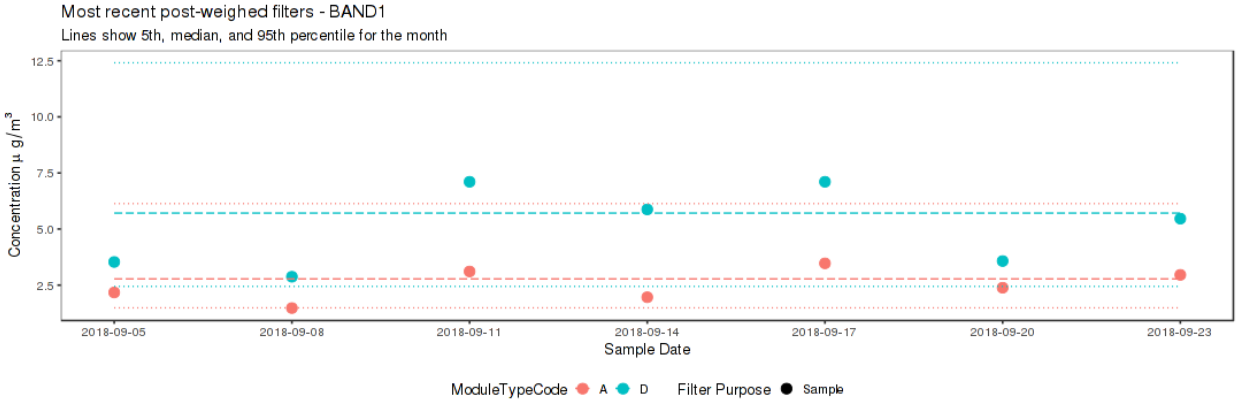

IMPROVE monthly ratio of S/Sulfate for 2011 to 2016 - both parameters above MDL
Ratio
0.45
0.40
0.35
0.30
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Bars show 10th and 90th percentile. Point indicates median.
Year
2011
2012
2013
2014
2015
2016

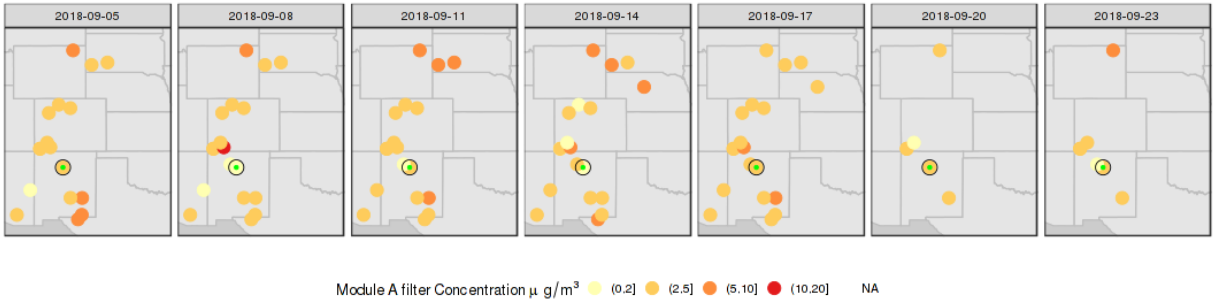
BAND1 - Bandelier (NM)

Start Date: 1988-03-02
Latitude: 35.780
Longitude: -106.266
Elevation (m): 1988
Affiliation: NPS
Site Type: IMPROVE

Bandelier National Monument is a 33,677-acre (13,629 ha) United States National Monument near Los Alamos in Sandoval and Los Alamos Counties, New Mexico. The monument preserves the homes and territory of the Ancestral Puebloans of a later era in the Southwest. Most of the pueblo structures date to two eras, dating between 1150 and 1600 AD.

The Monument is 50 square miles (130 km^2) of the Pajarito Plateau, on the slopes of the Jemez Volcanic field in the Jemez Mountains. Over 70% of the Monument is wilderness, with over one mile elevation change, from about 5,000 feet (1,500 m) along the Rio Grande to over 10,000 feet (3,000 m) at the peak of Cerro Grande on the rim of the Valles Caldera, providing for a wide range of life zones and wildlife habitats. There are three miles of road, and more than 70 miles of hiking trails. The Monument protects Ancestral Pueblo archeological sites, a diverse and scenic landscape, and the country's largest National Park Service Civilian Conservation Corps National Landmark District.





* Concentration equals the mass of material on the filter divided by the volume. The volume is the air volume passing through the filter for the whole sampling period, normally 24 hours. For $PM_{2.5}$ module (A module), it's about 33.1 m^3/day; for PM_{10} module (D module), it is about 24.3 m^3/day.

Recap

Major new systems

- New controller software and data management
- New automated weighing system management

Additions to existing tools

- New views for data validation
- New views for exploration and analysis

On the Horizon

- Further improvement of controller software, data handling, alerting and response
- Continued refinement of data validation tools and practices
- Restructure of sample handling lab software to fully integrate the automated weighing system into the processing stream
- Integration of FTIR analysis into the IMPROVE operational database
- Simplification of XRF process automation and integration of scanned filter bar codes