AIR SENSOR DASHBOARD

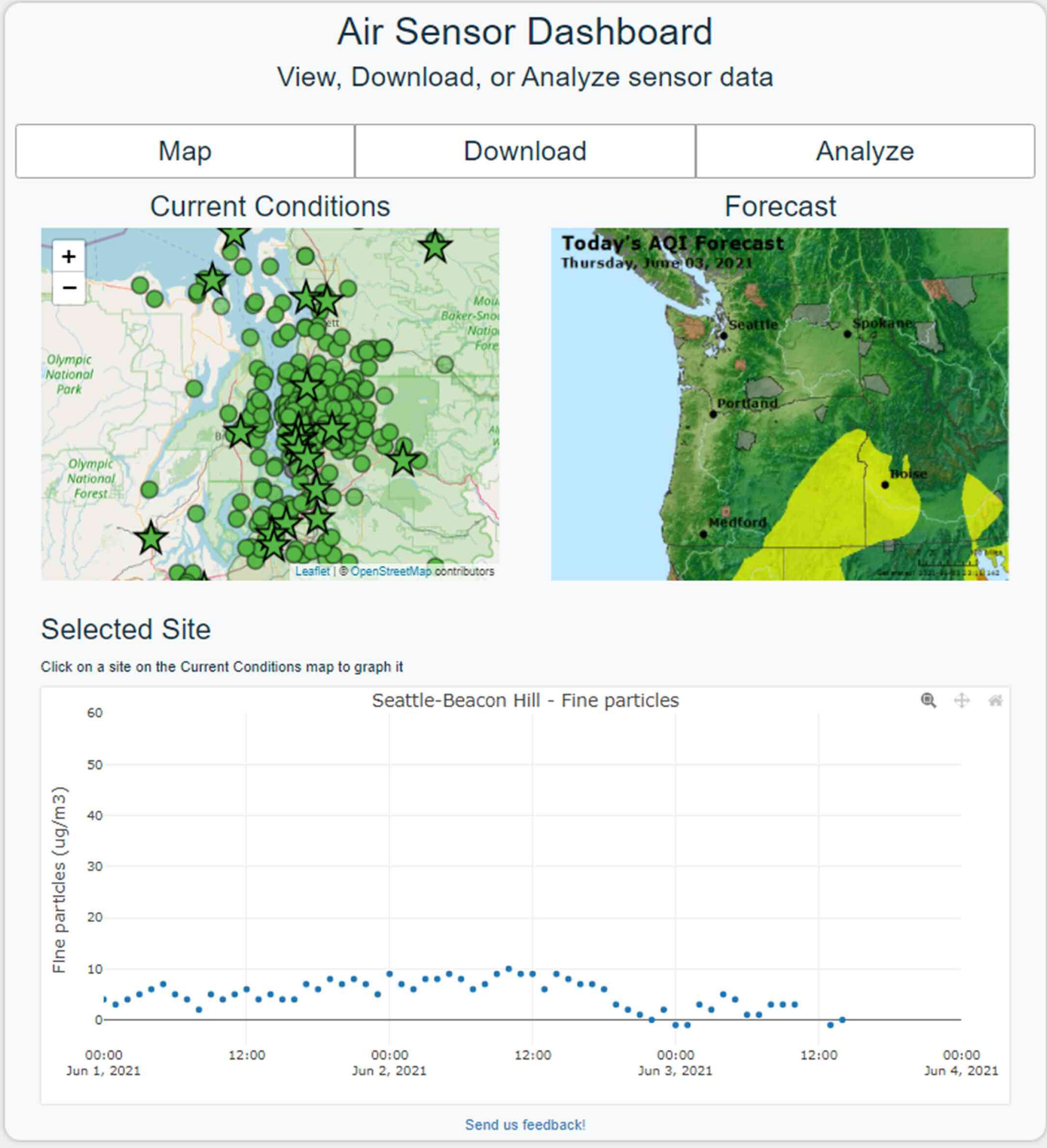
Graeme Carvlin

Puget Sound Clean Air Agency

June 8th, 2021

Dashboard

- Overview of Current Conditions
- Forecast
- Map single site

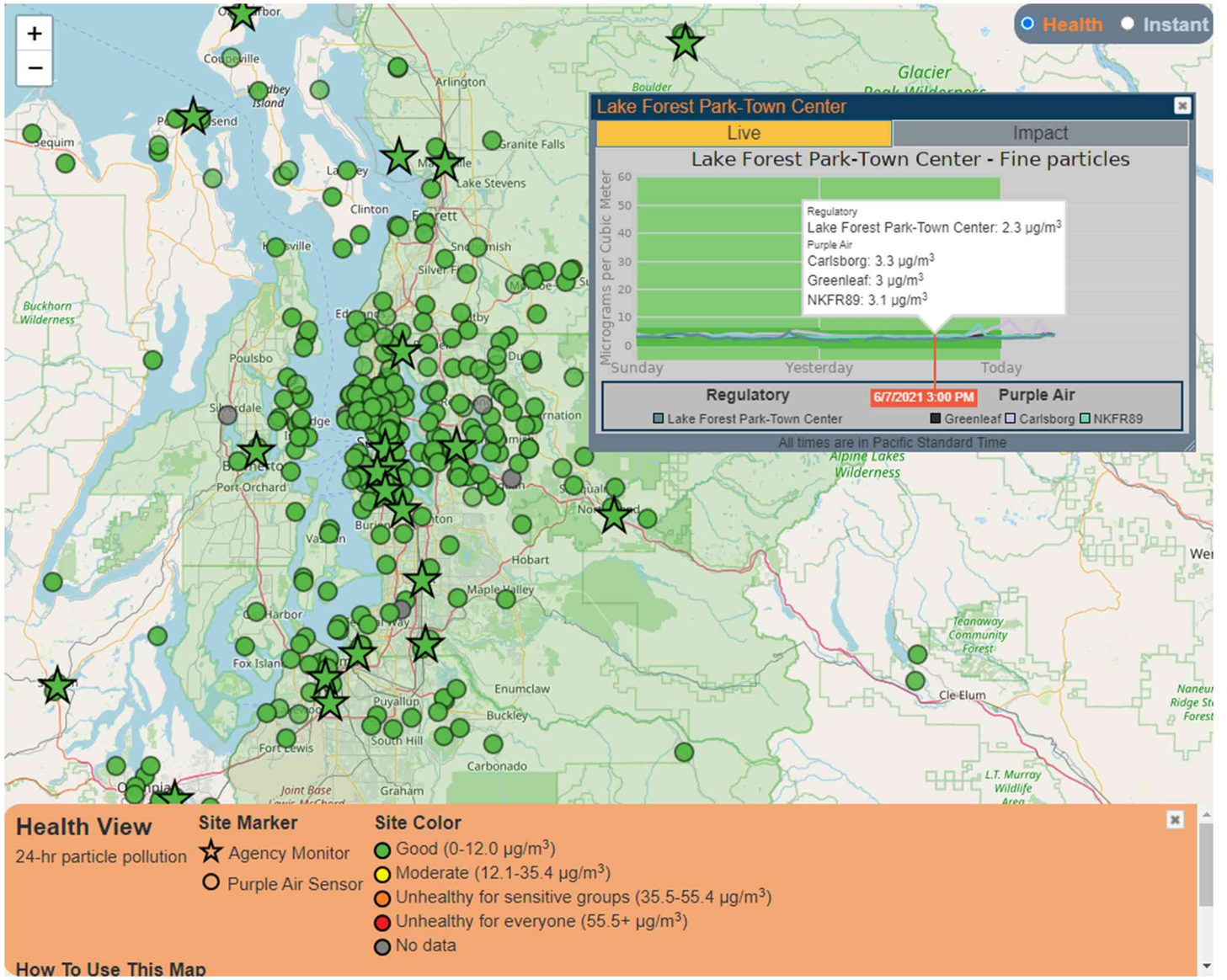


Air Sensor Dashboard

June 8, 2021

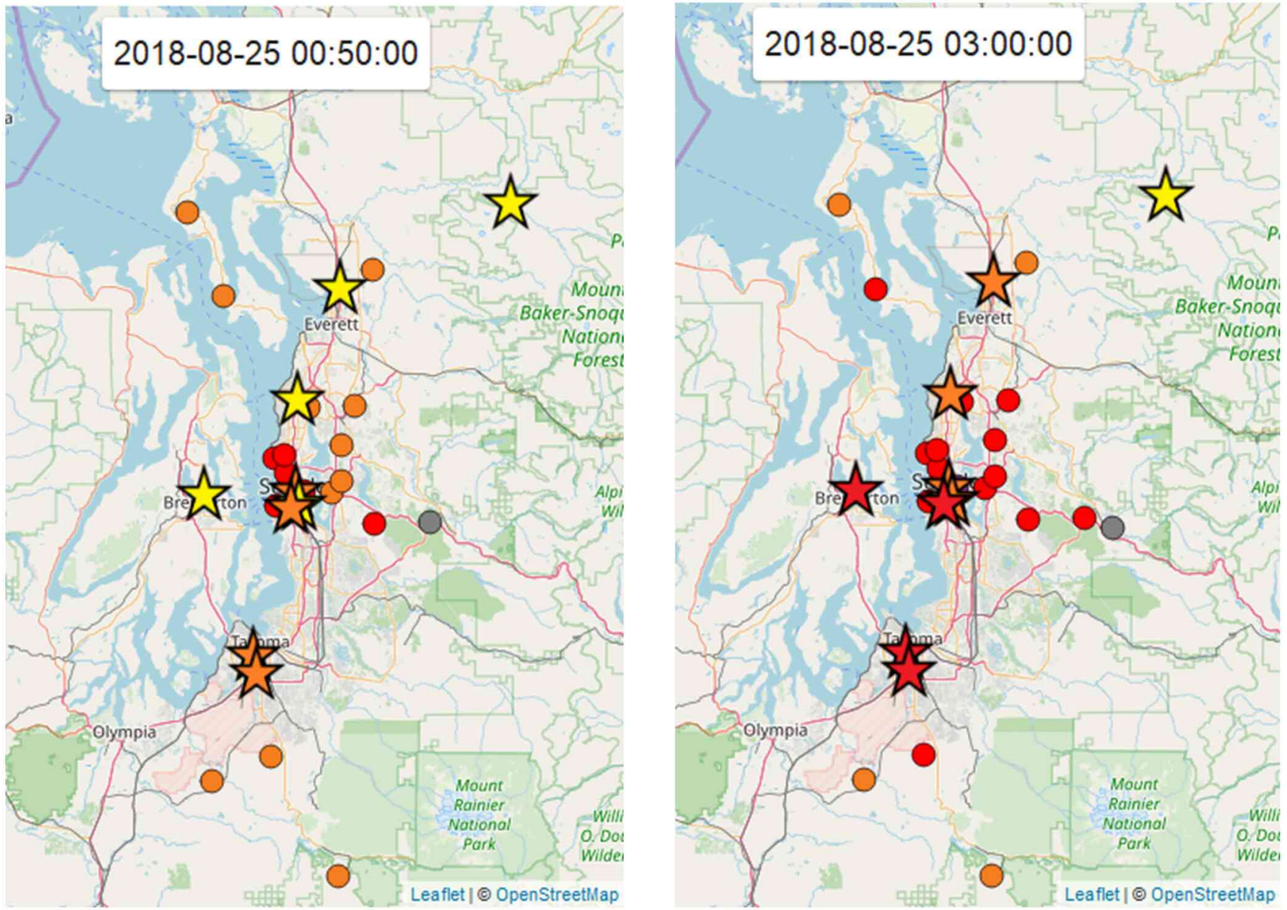
Map

- Regulatory monitors and calibrated + QC'd sensors
- Good coverage, minute data
- http://map.pscleanair.gov/



Map

- Minute data helpful during events



Purple Air sensors pick up wildfire smoke before regulatory sites

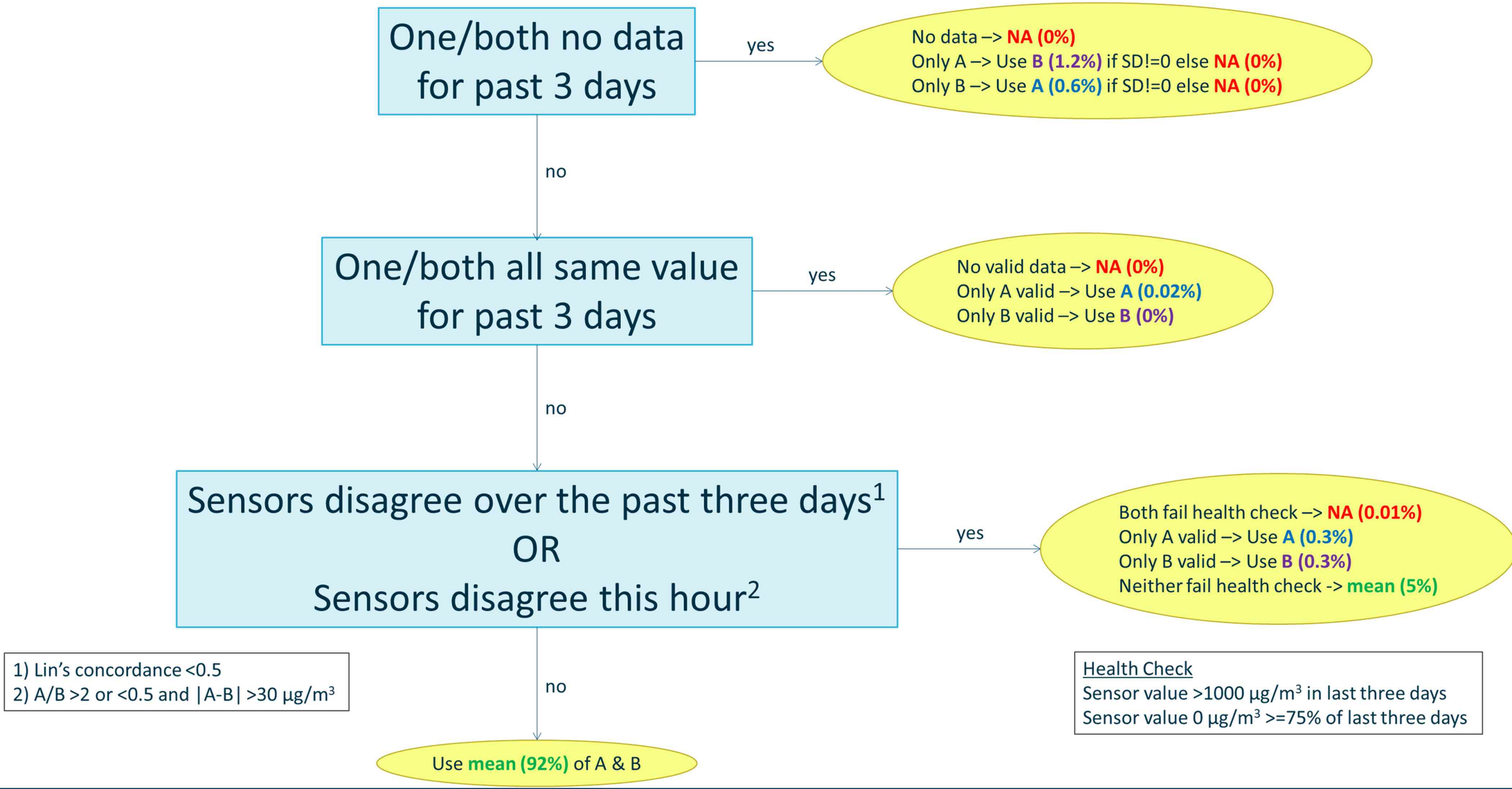
Data Processing

- Following slides use all data from 2019 in PSCAA 4 counties
- Remove values that are too low and too high
- Intra-monitor QC
- Inter-monitor QC
- Calibration

Data Processing

- Too low when 0.3µm count is <10 (now 2)
    - This value should never be 0. Very clear conditions have led to valid measures as low as 2.2 (hourly avg). Error states can be up to ~10, but it is worth capturing the true low values. The PM2.5 value can still be 0 up to ~100-200 counts of this particle size.
    - 1.1% of sensor values were too low
- Too high when PM2.5 ATM is >1000 µg/m3
    - > 1000 µg/m3 possible in very high pollution environments
    - True “max” value is ~5000 µg/m3 when the plantower sensor can’t read any higher
    - 0.5% of sensor values were >1000 µg/m3 (out of these, 96% are also >3000 µg/m3)
- 16-bit overflow - very high values are seen as 0s or low values
    - 83% of data >1000 µg/m3 also has <10 counts for 0.3µm particles

Intra-monitor QC Flow Chart



Inter-monitor QC Flow Chart

More than one nearby site

no → **mean (7%)**

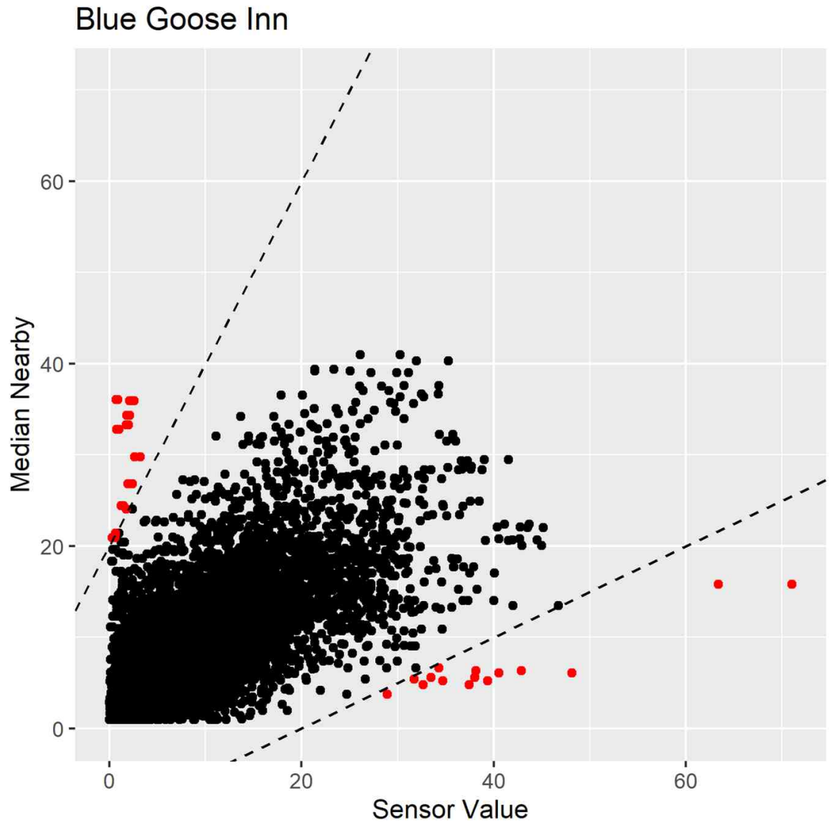
yes

One/both comparisons[1] can't be calculated

yes → No valid comparisons –> **mean (0.2%)**
Only A valid –> Use **A (0.7%)**
Only B valid –> Use **B (1.6%)**

no

Both good[2] –> **mean (88%)**
A better than B[3] –> Use **A (0.9%)**
B better than A[3] –> Use **B (1.9%)**
Both bad[4] -> **NA (0.4%)**

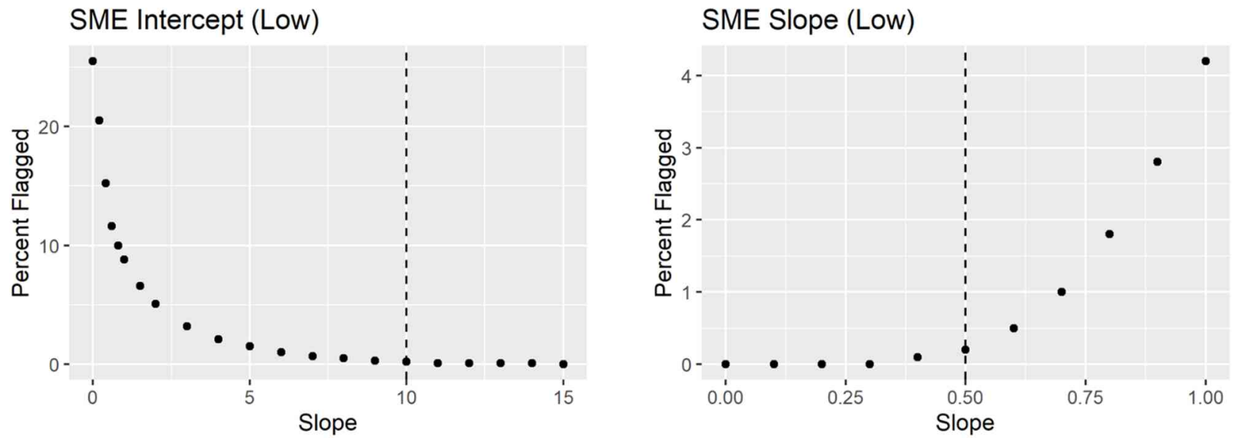


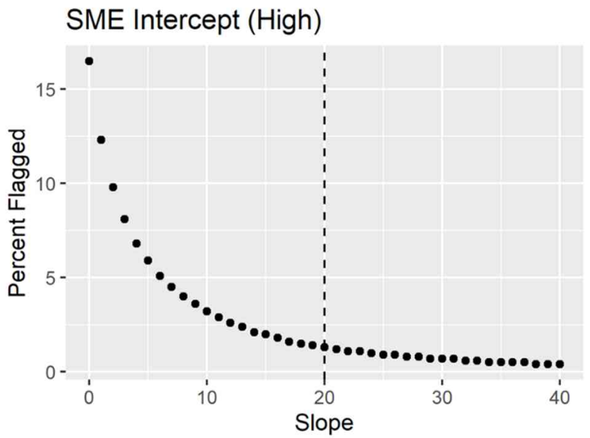
1) Scaled median error (SME)
   a) If sensor value is
      i. greater than 2*median_nearby + 20
      ii. OR less than 0.5*median_nearby – 10
   b) SME is 0
   c) Otherwise scale the difference between the sensor value and median_nearby from 0-1, where 1 is perfect agreement.
   d) median_nearby is the median of all A and B values from all sensors within range.
2) Both good
   a) The difference between A and B SME is < 0.2
   b) AND A SME is not zero
   c) AND B SME is not zero
3) SME A > SME B; SME B > SME A
4) A and B SME are 0

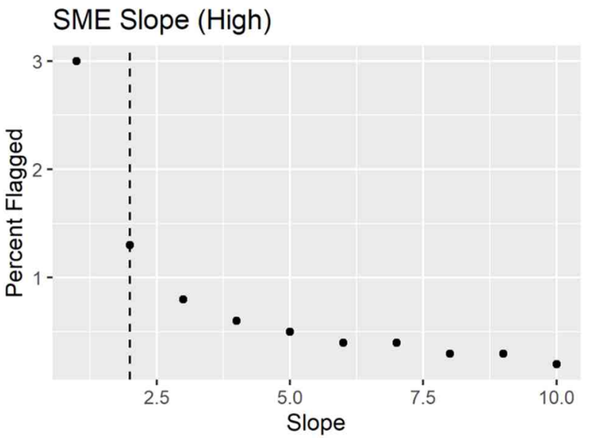
Note: Percentages are based on application of inter-monitor QC to entire dataset.

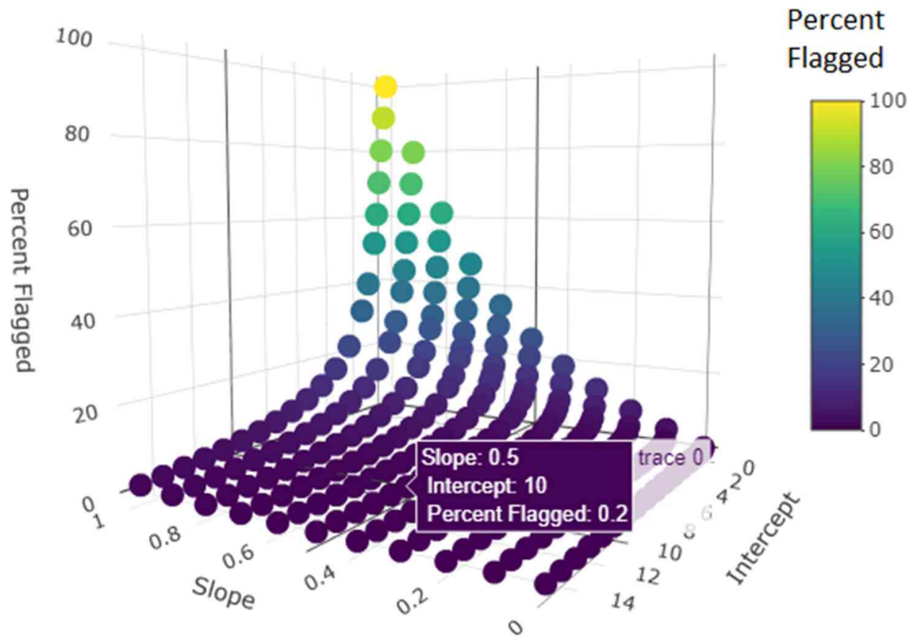
PUGET SOUND
Clean Air Agency

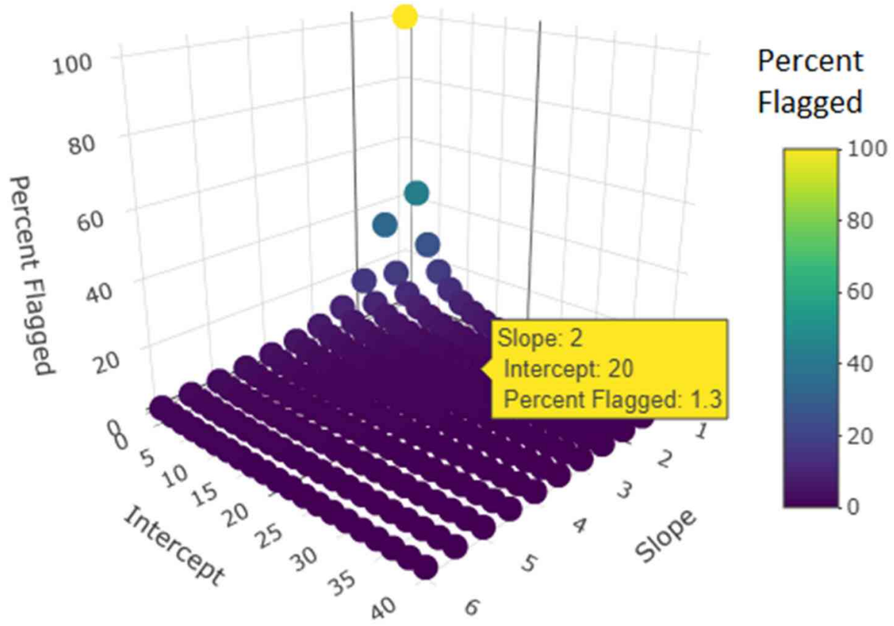
Inter-monitor QC Cutoffs







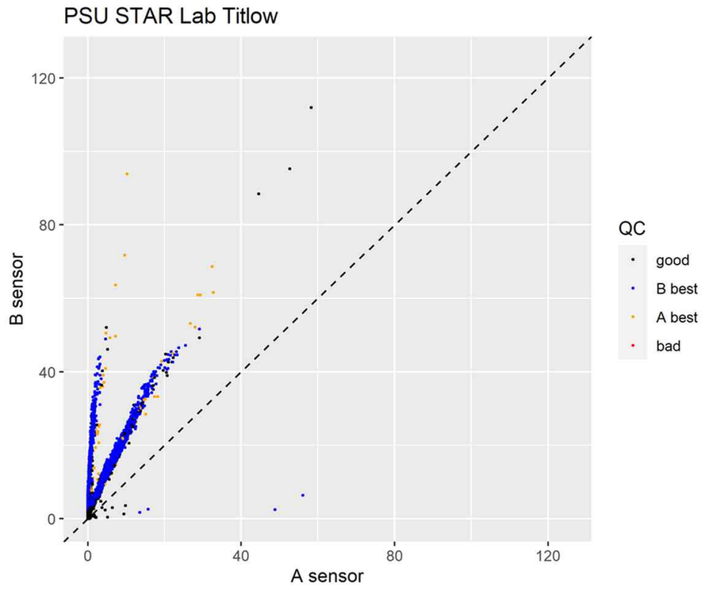


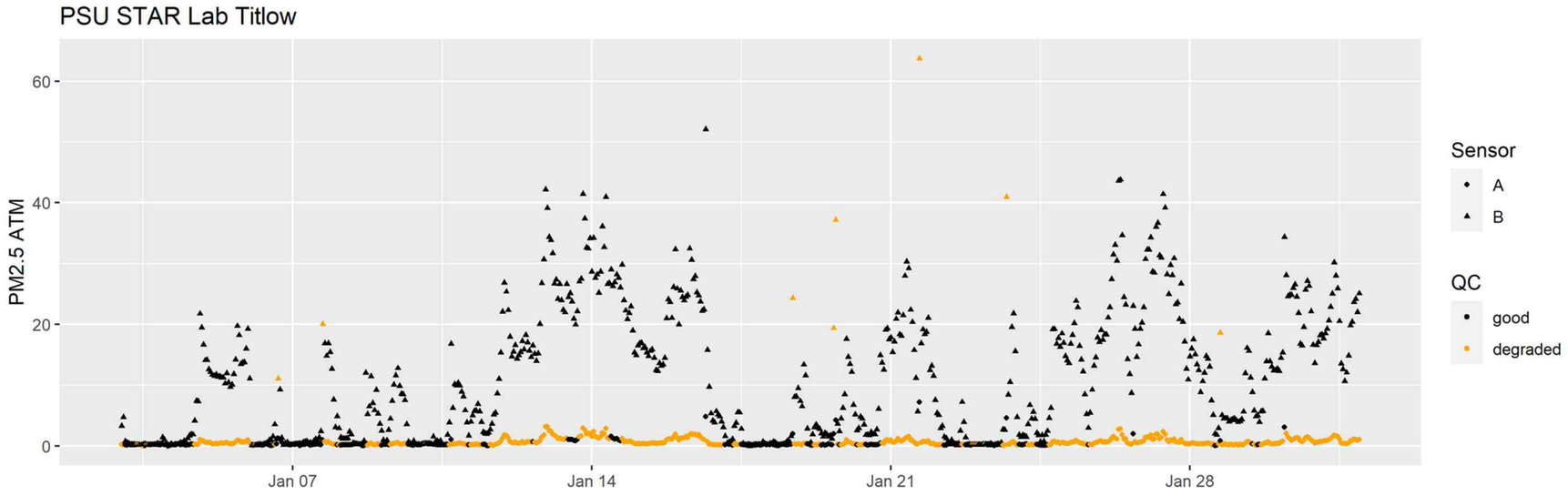


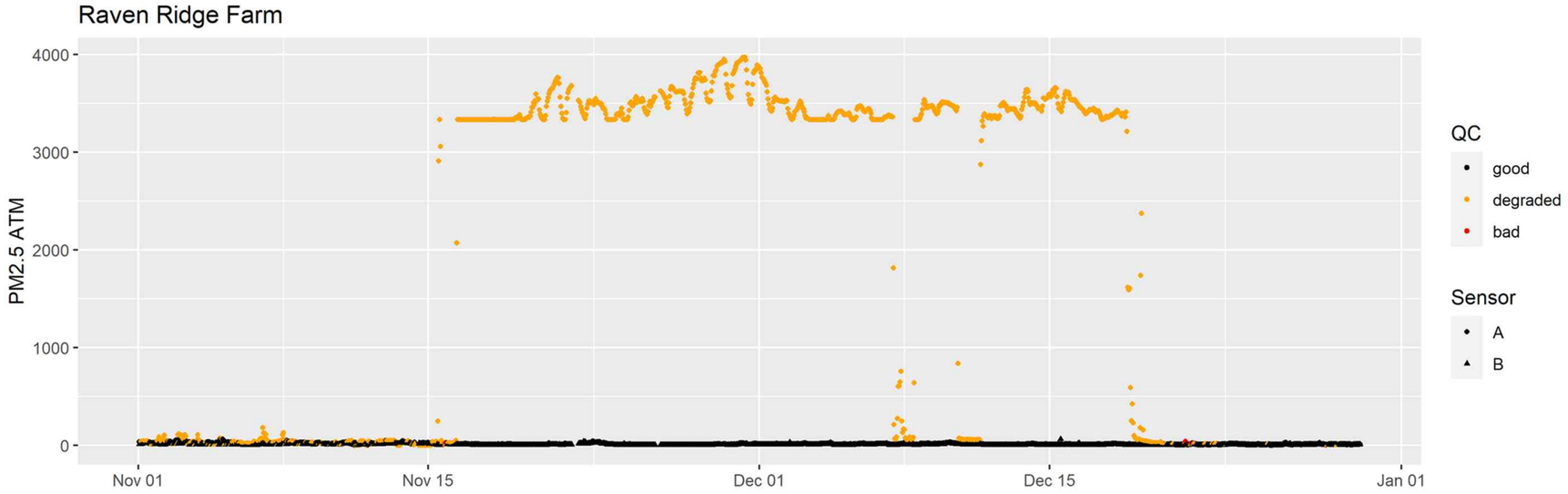
Benefits and Drawbacks of Inter-monitor QC

- Only applies when intra-monitor QC fails
    - This will leave in valid points that are not like nearby monitors (i.e. pollution may really be 300 µg/m3 at “Outdoor Firepit”, but that is not representative beyond microscale) – about 6% of all data
    - Of the 5% of data that fail intra-monitor QC, inter-monitor QC reports 64% good, 29% A or B, 7% both bad or no nearby sites
- Dependent on correct range. Can be too aggressive if range is too large and “nearby” monitors don’t have similar PM2.5 composition.
    - Range calculated by semi-variogram
    - Min of 6km, max of 50km

Benefit of Inter-monitor QC

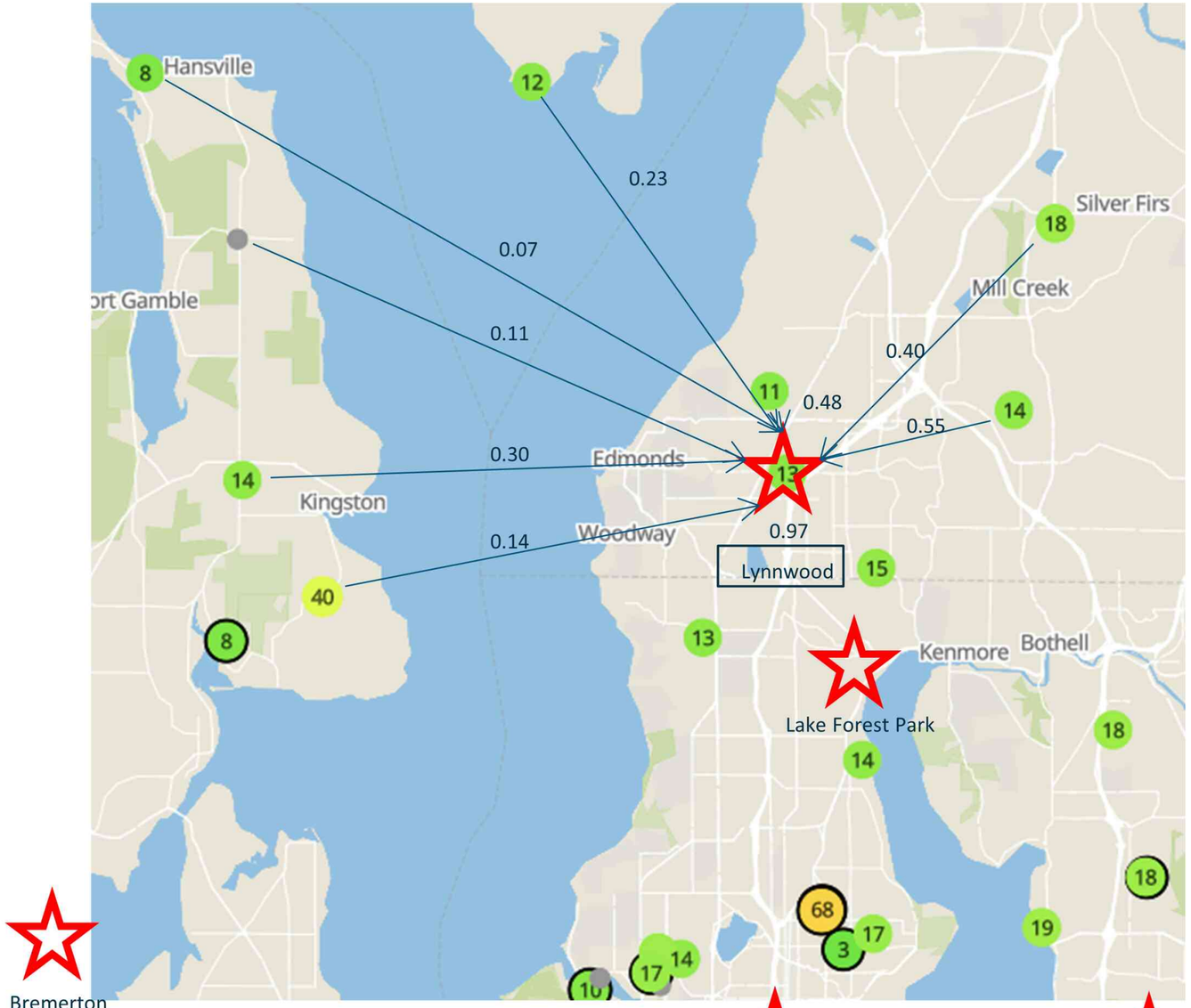






Calibration

Example of Colocation R^2 to Nearest Regulatory Monitor (Lynnwood)



Colocated Sensor Model Performance

Model	RMSE (µg/m³)	R^2
Nearest Site (33%)	2.45	0.82
Regional[a]	2.92	0.81
EPA (67%)	3.05	0.82
EPA with temp & RH[b]	3.10	0.81

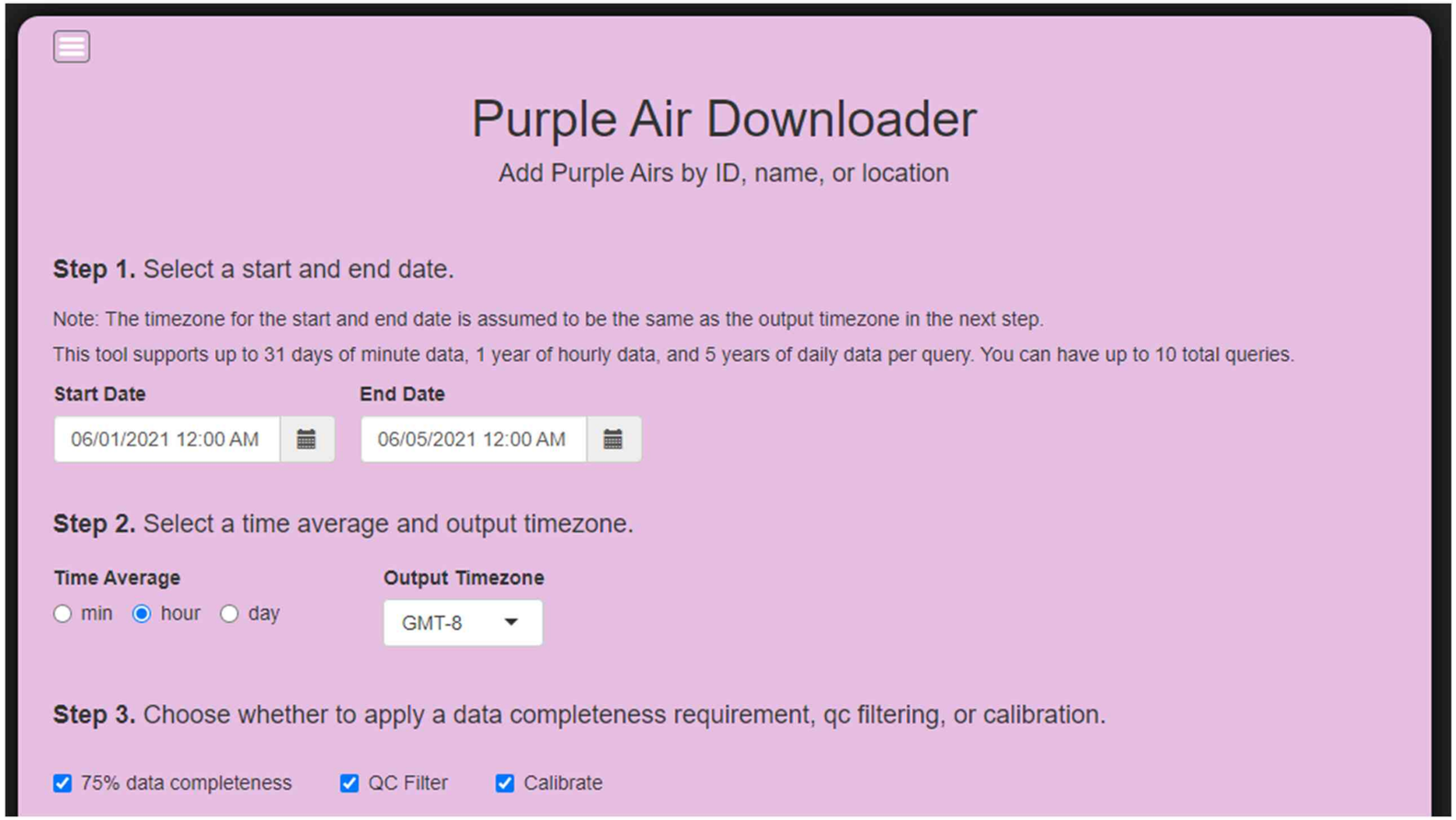
[a]The sensors used to create the regional model were removed from this table so as not to bias the results

[b]This model is likely worse than the base EPA model because the temperature and RH relationship in the Puget Sound is different from the national average

PUGET SOUND
Clean Air Agency

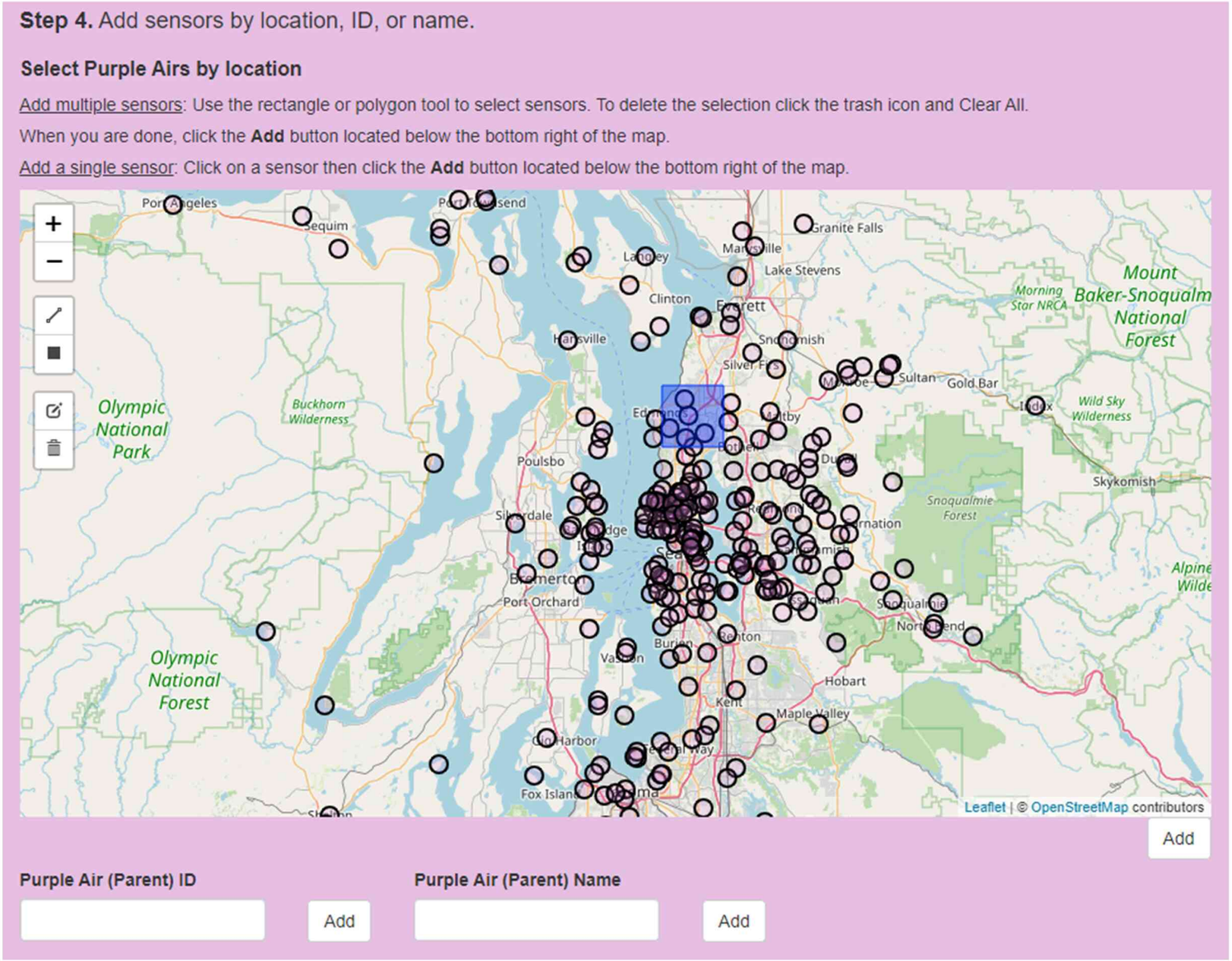
Download

- Download Purple Air data
- QC, calibration, data completeness
- Minute, hourly, daily



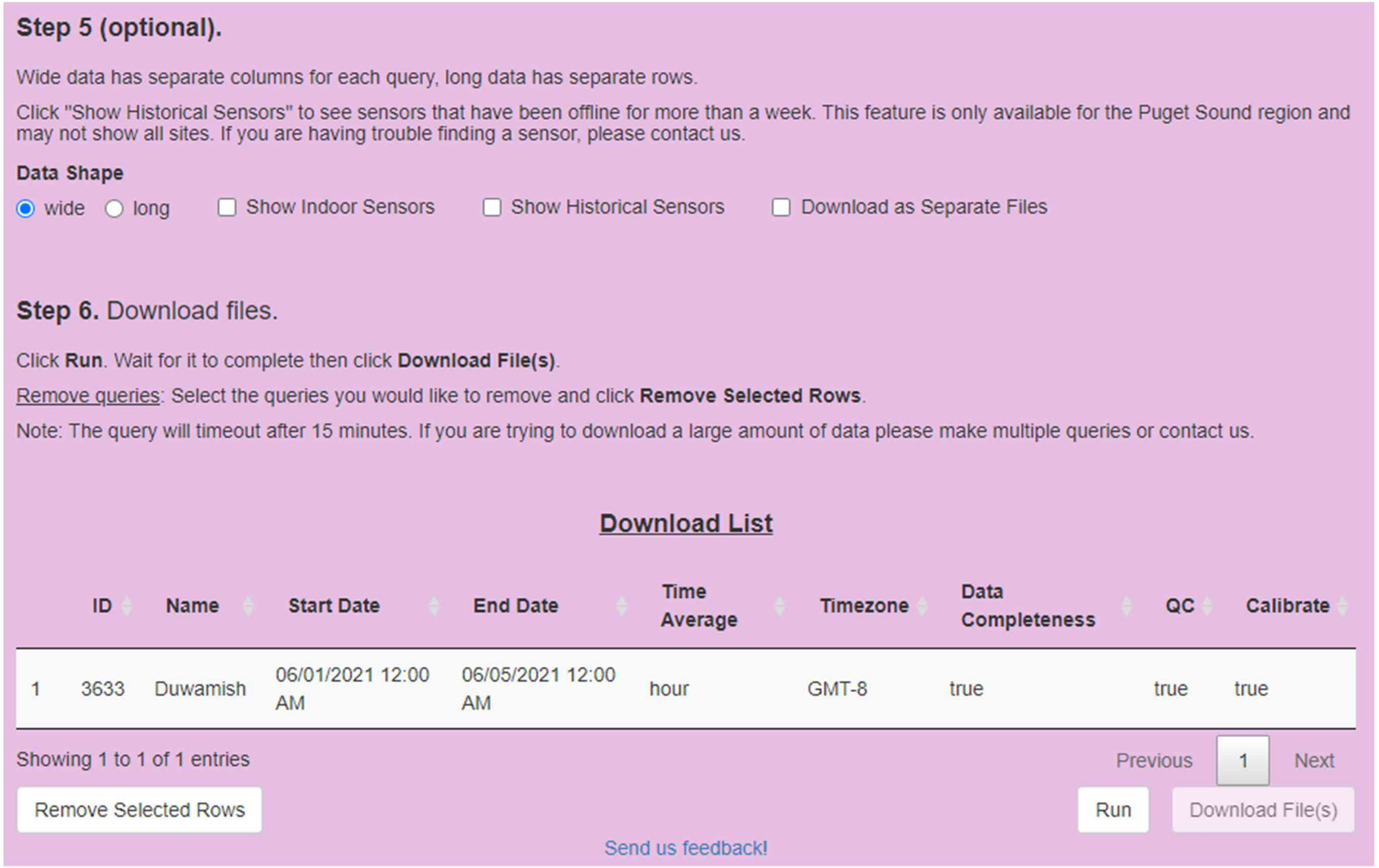
Download

- Select by location, ID, name



Download

- Output wide or long, separate or single file
- Show Historic sensors
- Run query then download
- AWS Lambda



Step 5 (optional).

Wide data has separate columns for each query, long data has separate rows.

Click "Show Historical Sensors" to see sensors that have been offline for more than a week. This feature is only available for the Puget Sound region and may not show all sites. If you are having trouble finding a sensor, please contact us.

Data Shape

wide long Show Indoor Sensors Show Historical Sensors Download as Separate Files

Step 6. Download files.

Click **Run**. Wait for it to complete then click **Download File(s)**.

Remove queries: Select the queries you would like to remove and click **Remove Selected Rows**.

Note: The query will timeout after 15 minutes. If you are trying to download a large amount of data please make multiple queries or contact us.

Download List

	ID	Name	Start Date	End Date	Time Average	Timezone	Data Completeness	QC	Calibrate
1	3633	Duwamish	06/01/2021 12:00 AM	06/05/2021 12:00 AM	hour	GMT-8	true	true	true

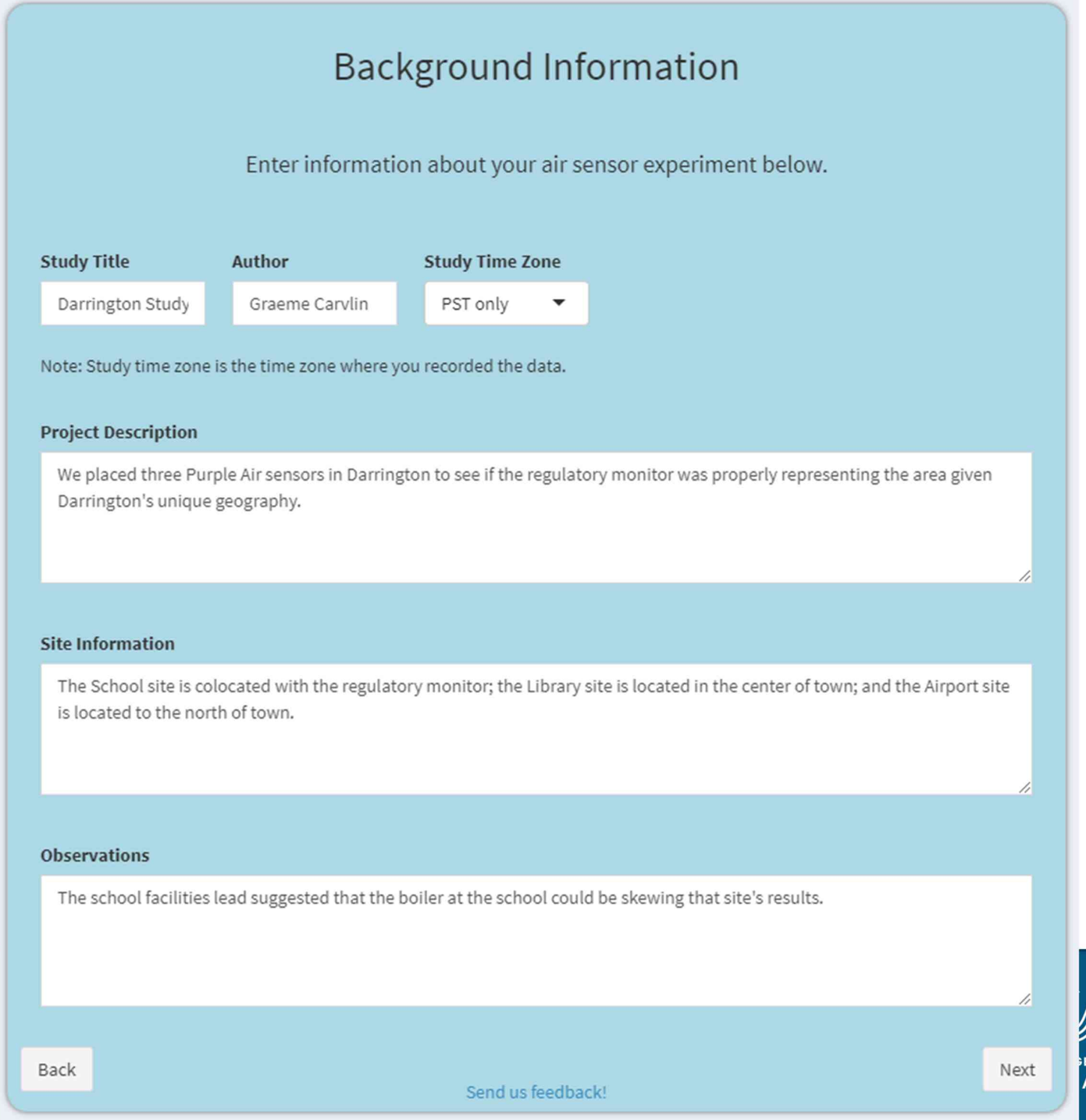
Showing 1 to 1 of 1 entries Previous 1 Next

Remove Selected Rows Run Download File(s)

Send us feedback!

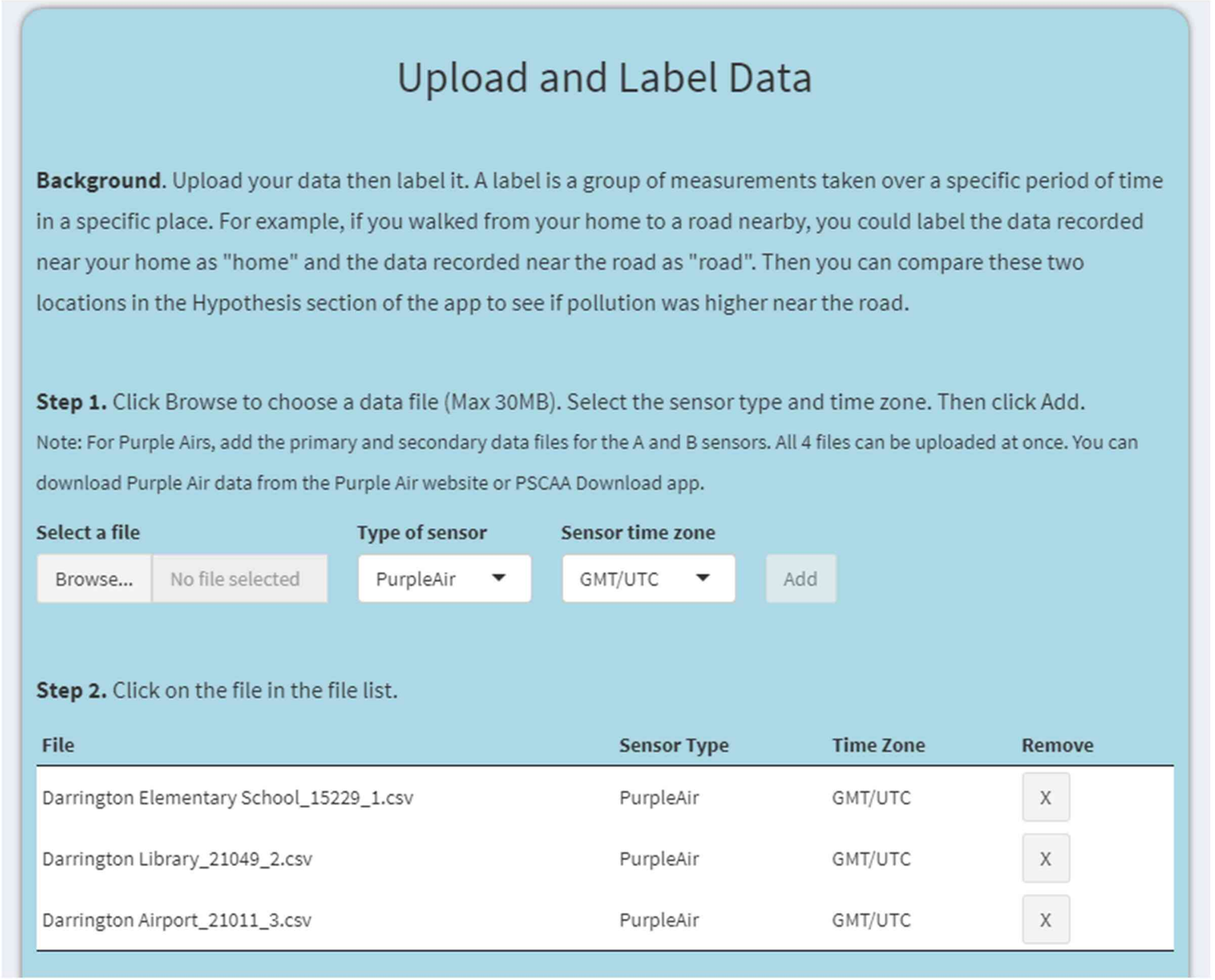
Report

- Fill in background information



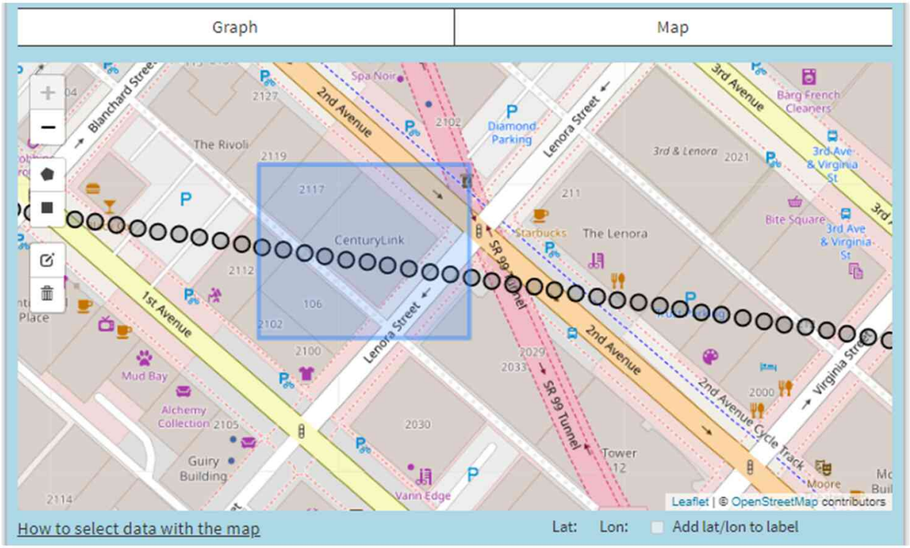
Report

- Upload files for AirBeam, Dylos, or Purple Air (PA website or Download tool)

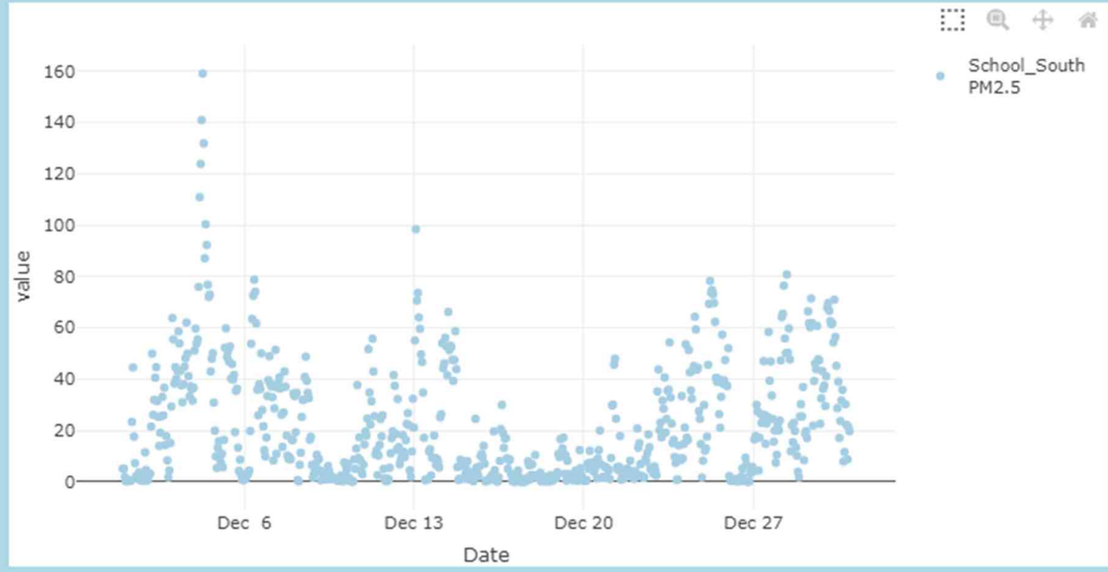


Report

- Label your data
- Select by timeseries or map



Step 3. Select the data from your study period using the graph or map.



How to select data with the graph

When a file is selected in the file list the data will show up in the graph. And four icons will show up in the top right of the graph.

(1) The dotted square is a box selection tool. Click and drag to highlight data. Double click to clear your selection.

(2) The magnifying glass is a zoom tool. Click and drag to zoom in on the data. Double click to return to normal zoom.

(3) The arrow cross is a pan tool. Click and drag to move left/right and up/down on the graph. Double click to recenter the graph.

(4) The home button resets the zoom and pan.

Click on a legend name to hide those data. Once hidden they won't be included in your selection.

Step 4. Name and add the label.

(optional) Use the Pollutant selection to only add a label to specific pollutants.

Pollutant PM2.5

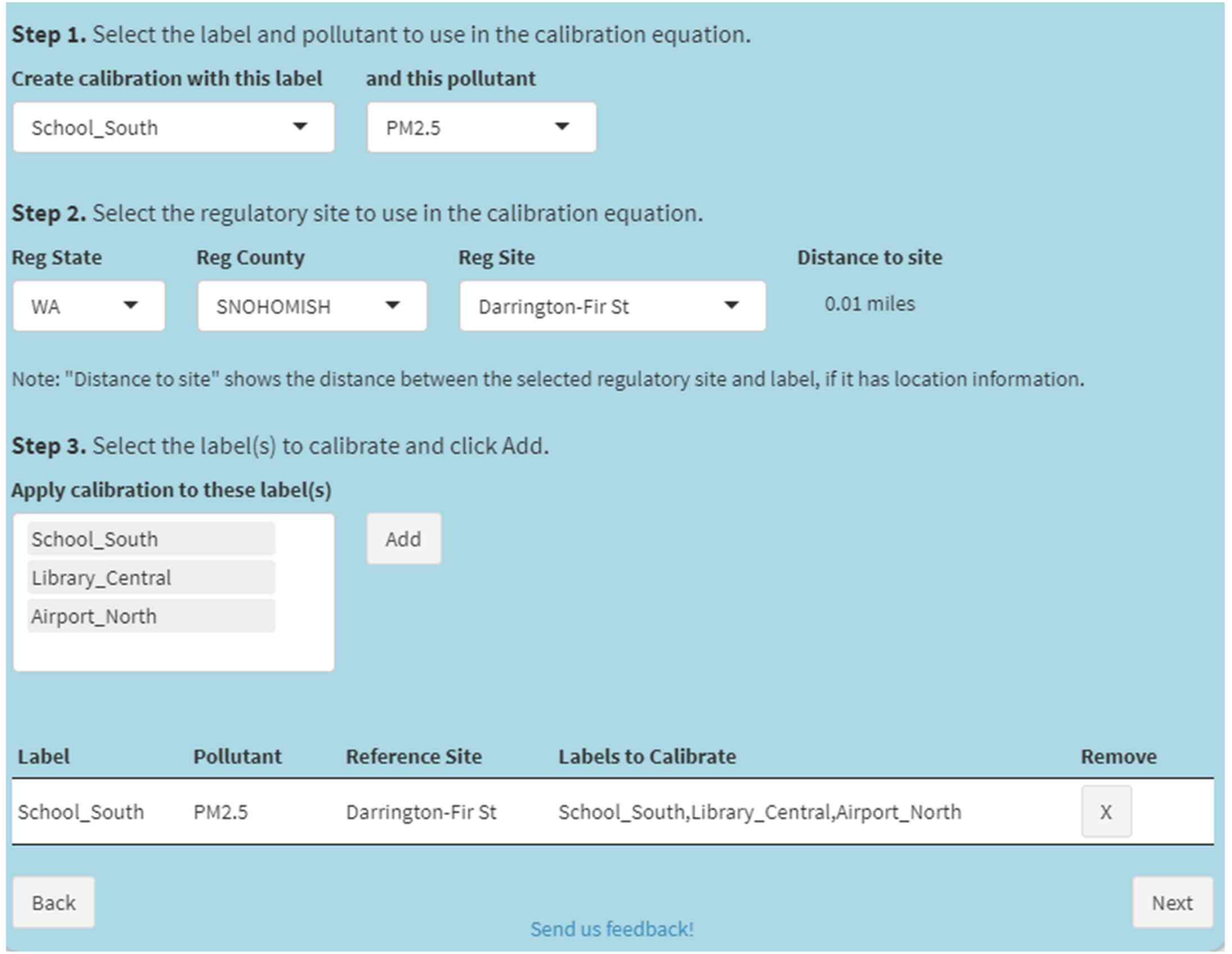
Label name School_South Add

Air Sensor Dashboard

June 8, 2021

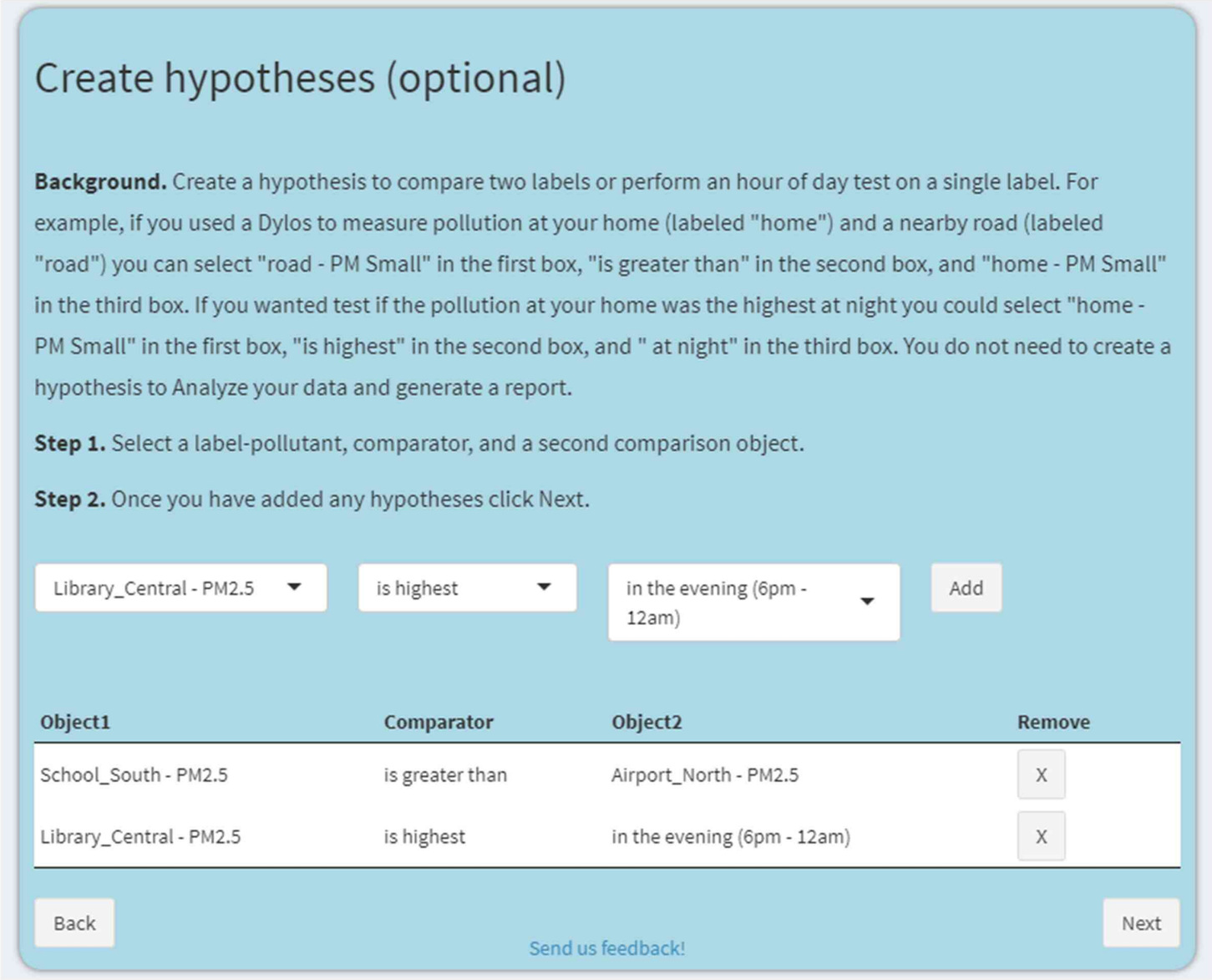
Report

- Calibrate



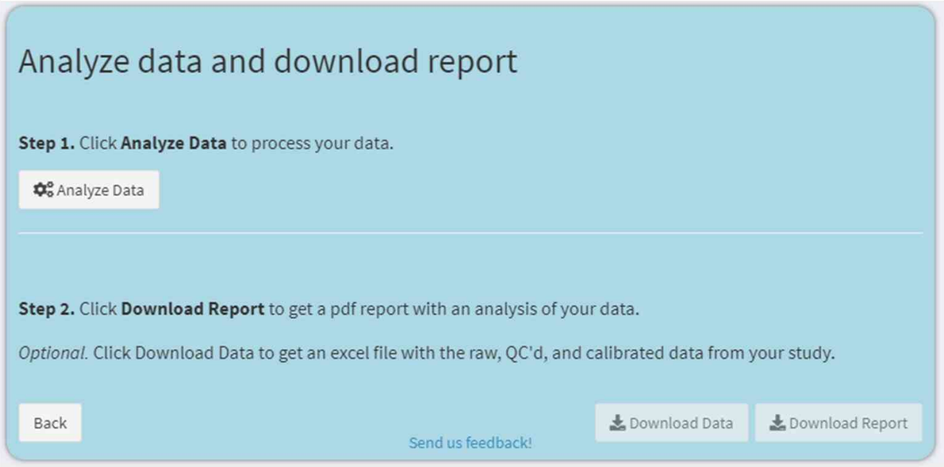
Report

- Add hypotheses
- Site and time of day comparisons



Report

- Download Report and data file



Darrington Study

Graeme Carvlin

06/07/2021

1 Project Description

We placed three Purple Air sensors in Darrington to see if the regulatory monitor was properly representing the area given Darrington's unique geography.

2 Site Information

The School site is colocated with the regulatory monitor; the Library site is located in the center of town; and the Airport site is located to the north of town.

3 Observations

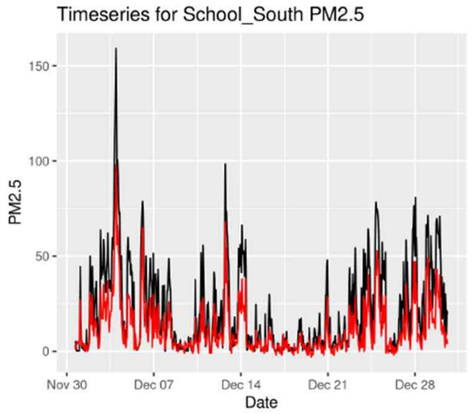
The school facilities lead suggested that the boiler at the school could be skewing that site's results.

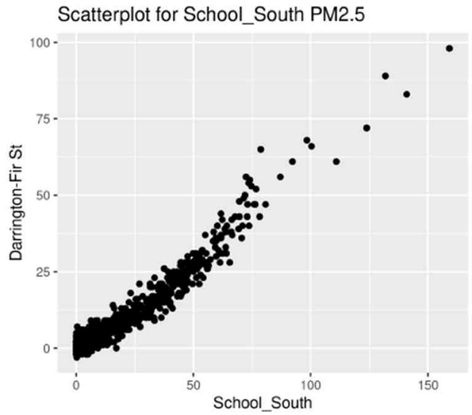
4 Summary Statistics

Data were collected from 2020-11-30 4pm to 2020-12-30 3pm.

Label	Pollutant	Min	Avg	Max	Trend
Airport_North	PM2.5	-0.78	9.21	66.58	None
Library_Central	PM2.5	-0.74	16.2	139.26	None
School_South	PM2.5	-0.78	11.37	89.86	None

5 Calibration

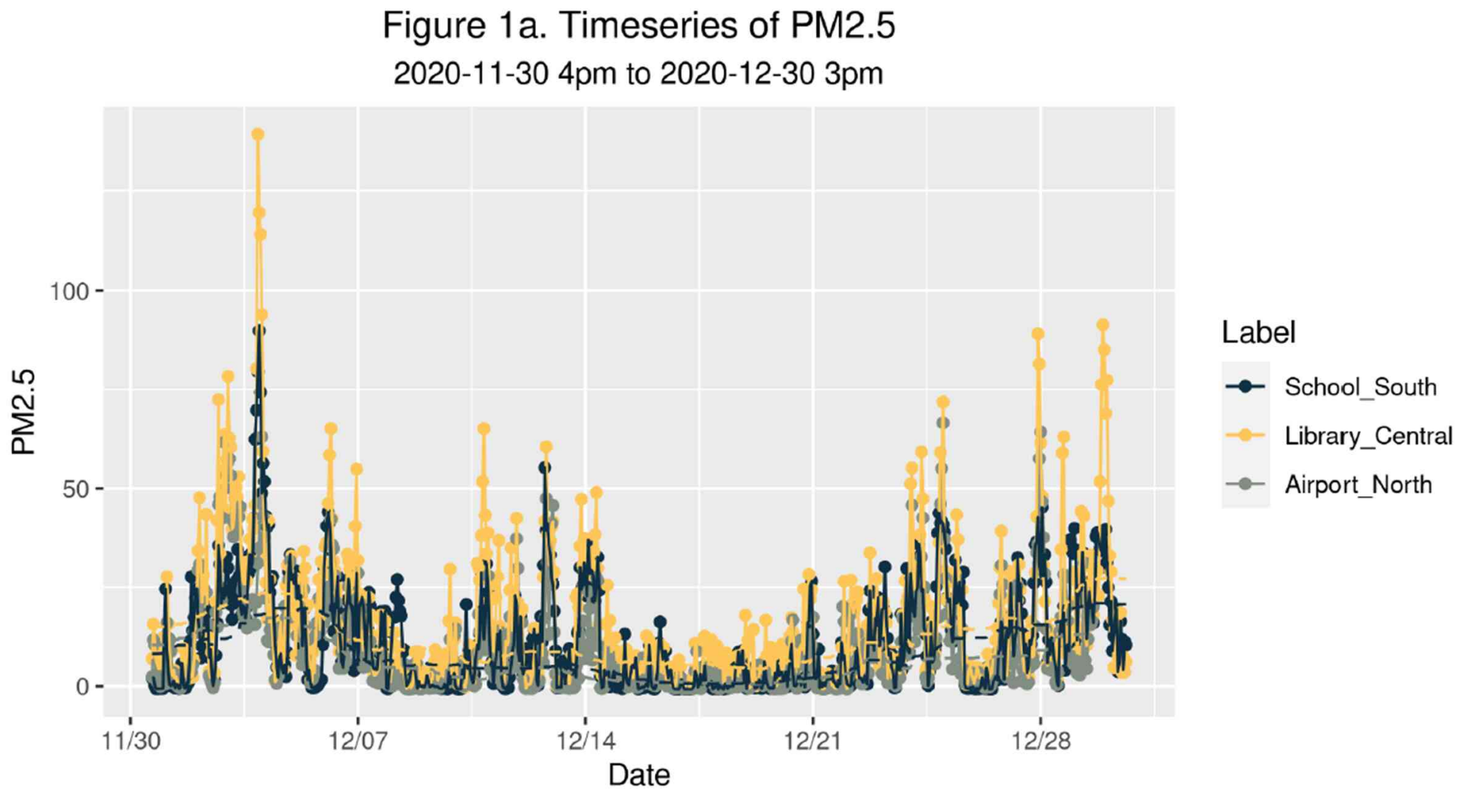




A calibration equation for PM2.5 was developed using data from School_South and reference data from Darrington-Fir St. There were 711 overlapping observations. The slope was 0.57 and the intercept was -0.78. The performance of the calibration equation was good, with an R-squared of 0.93 and a root mean square error of 3.51.

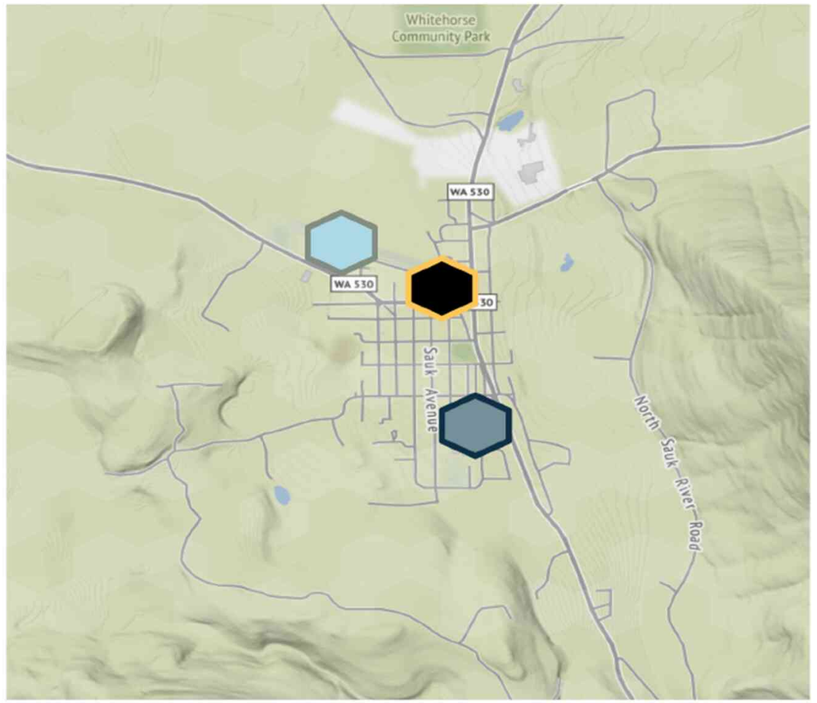
6 Figures

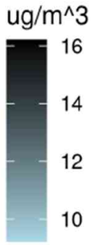
Solid lines are data, dashed lines are a running average of the data.



For School_South, average PM2.5 levels were high from 12/25/20 3pm to 12/30/20 3pm. PM2.5 levels spiked from 12/03/20 8pm to 12/04/20 3am and 12/12/20 6pm. For Library_Central, PM2.5 levels spiked from 12/03/20 9pm to 12/04/20 2am, 12/27/20 10pm to 12/28/20 12am, and 12/29/20 9pm to 12/30/20 1am. For Airport_North, average PM2.5 levels were high from 11/30/20 4pm to 12/06/20 5pm. PM2.5 levels spiked from 12/02/20 8pm to 12/03/20 9am, 12/04/20 12am to 12/04/20 3am, and 12/24/20 10pm to 12/25/20 12am.

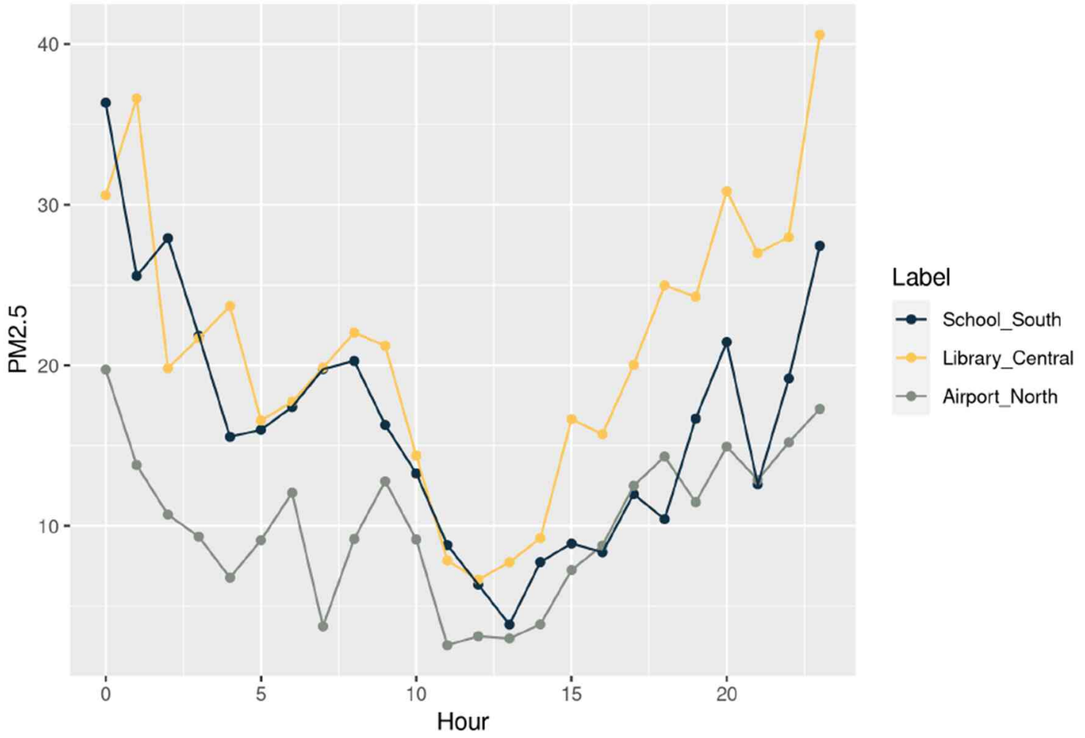
Figure 1b. Map of PM2.5





This map shows the average PM2.5 for each label by location.

Figure 1c. Hour of Day, PM2.5



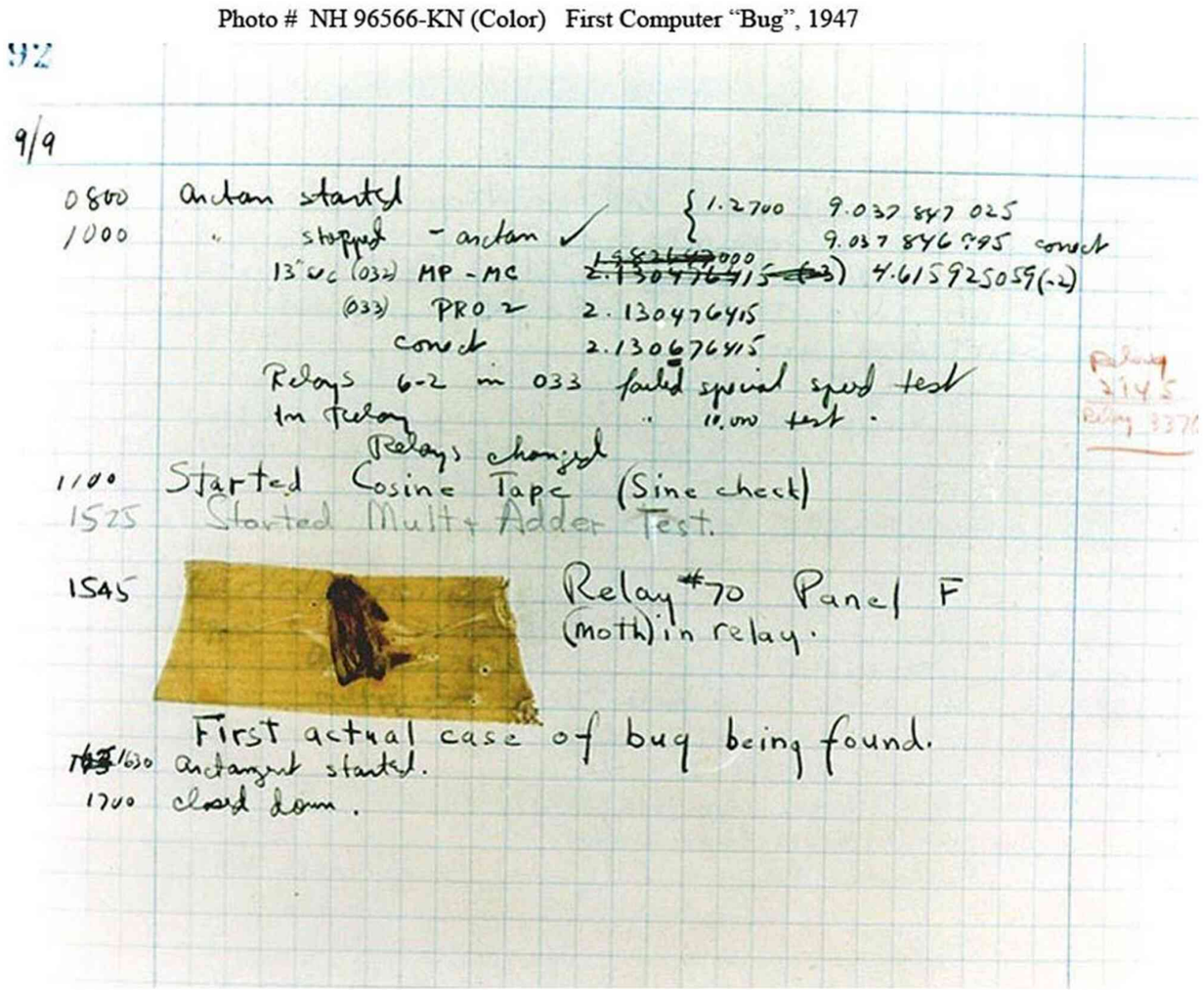
For School_South, PM2.5 was higher on average at 8pm, at 11pm, and from 12am-3am. For Library_Central, PM2.5 was higher on average from 8pm-11pm and from 12am-1am. For Airport_North, PM2.5 was higher on average from 12am-1am, at 6pm, and from 10pm-11pm.

7 Hypotheses

Hypothesis 1, School South - PM2.5 is greater than Airport North - PM2.5. A t-test shows that School South - PM2.5 is 2.16 ug/m^3 greater than Airport North - PM2.5 and this result is significant. Hypothesis 2, Library Central - PM2.5 is highest in the evening (6pm - 12am). Result: Library Central - PM2.5 is highest in the evening (6pm - 12am).

Next steps

- Open beta
- There will be bugs
- http://apps.pscleanair.gov/
- Expanded use cases?



https://www.nationalgeographic.org/thisday/sep9/worlds-first-computer-bug/

Thank you!

Questions?