



# AIR SENSOR DASHBOARD

Graeme Carvlin

**Puget Sound Clean Air Agency**

June 8<sup>th</sup>, 2021



# Dashboard

- Overview of Current Conditions
- Forecast
- Map single site

## Air Sensor Dashboard

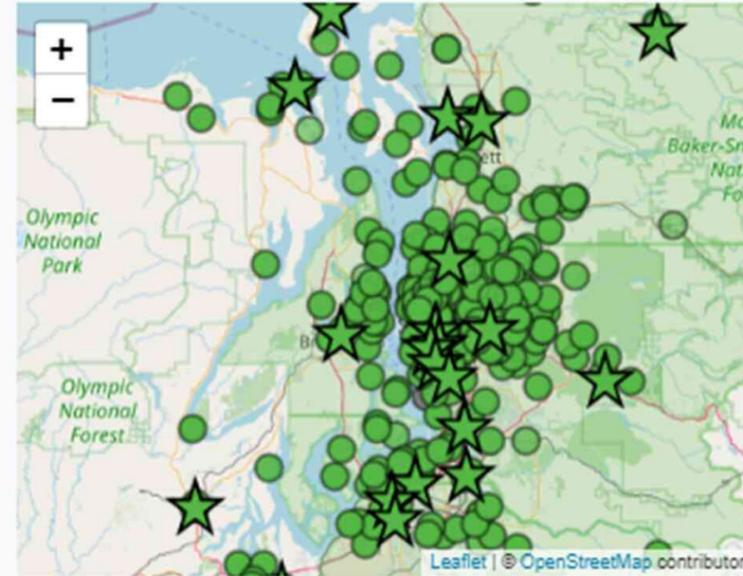
View, Download, or Analyze sensor data

Map

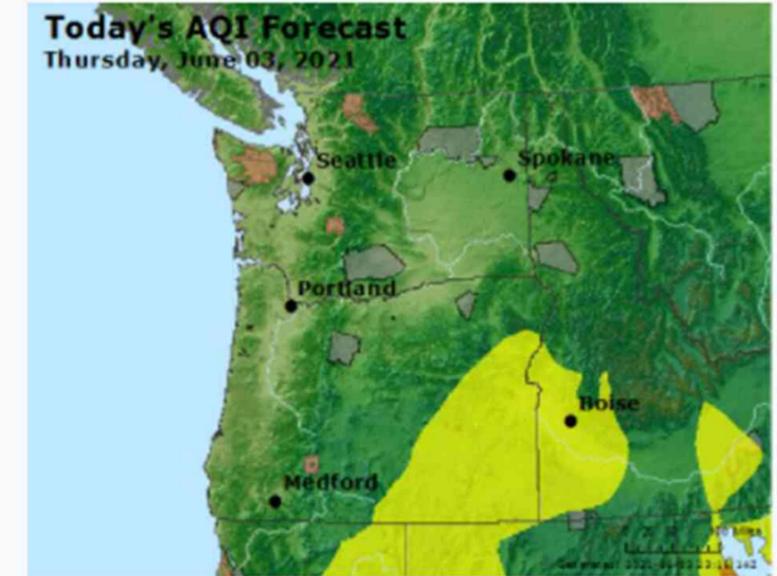
Download

Analyze

Current Conditions

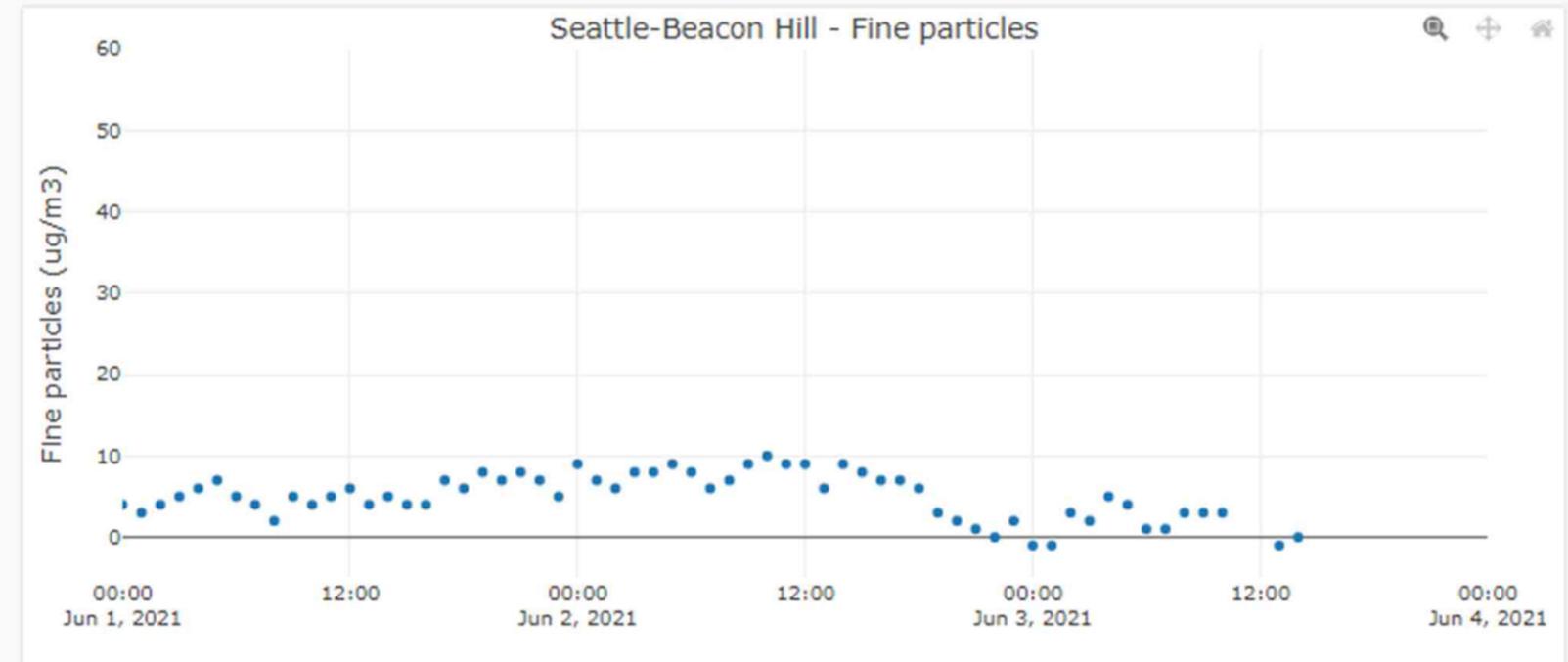


Forecast



Selected Site

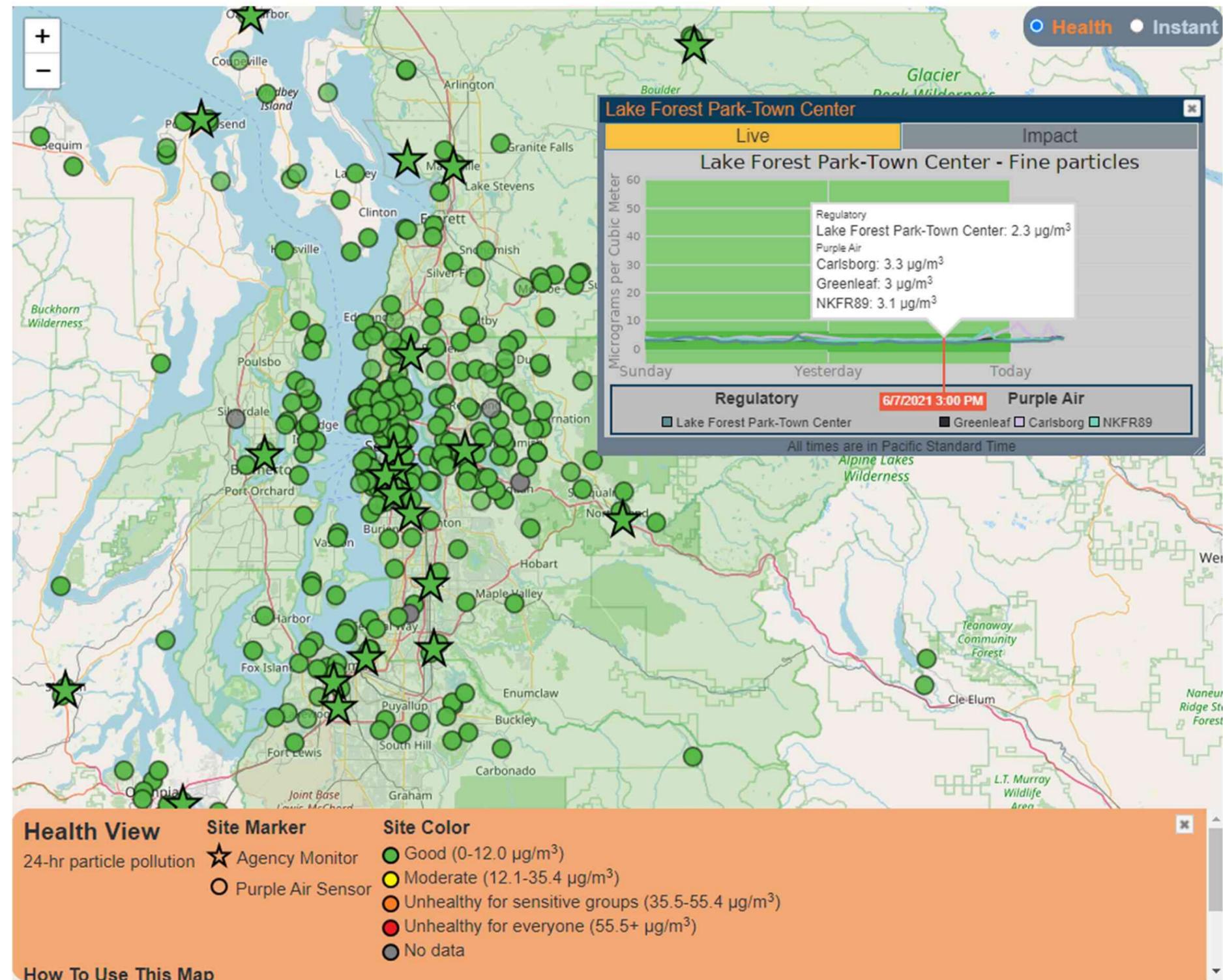
Click on a site on the Current Conditions map to graph it



[Send us feedback!](#)

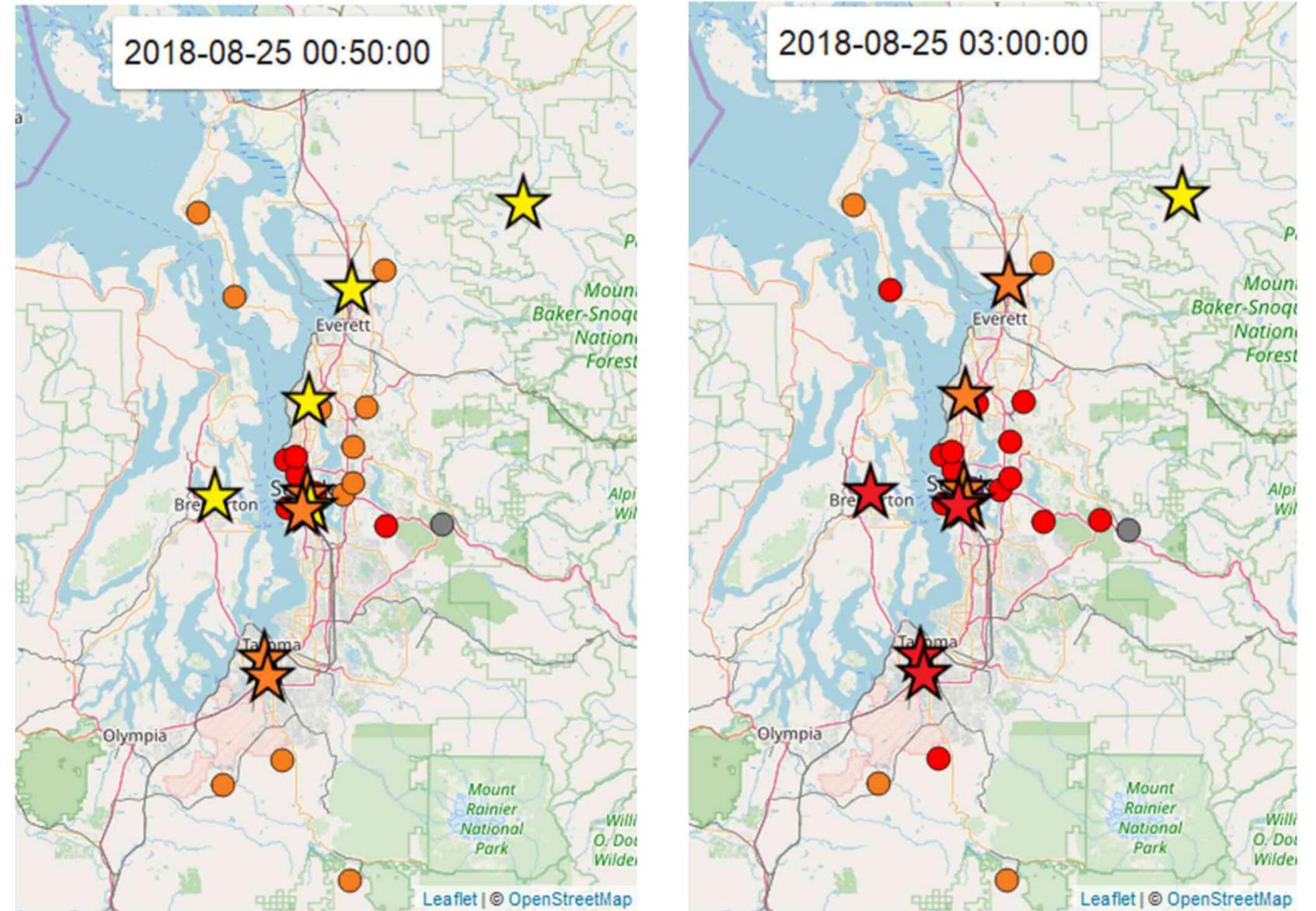
# Map

- 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

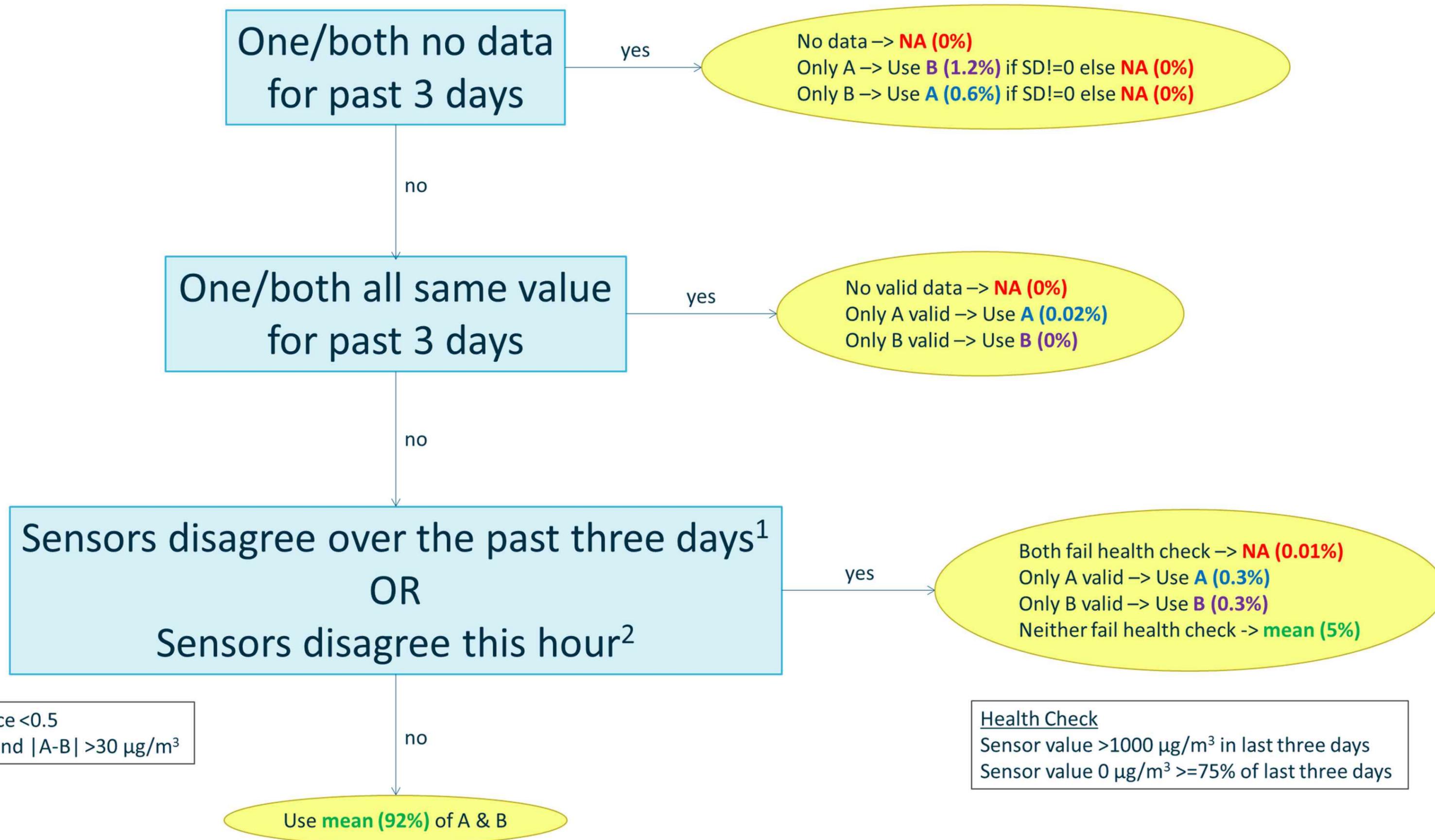
# Data Processing

- 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

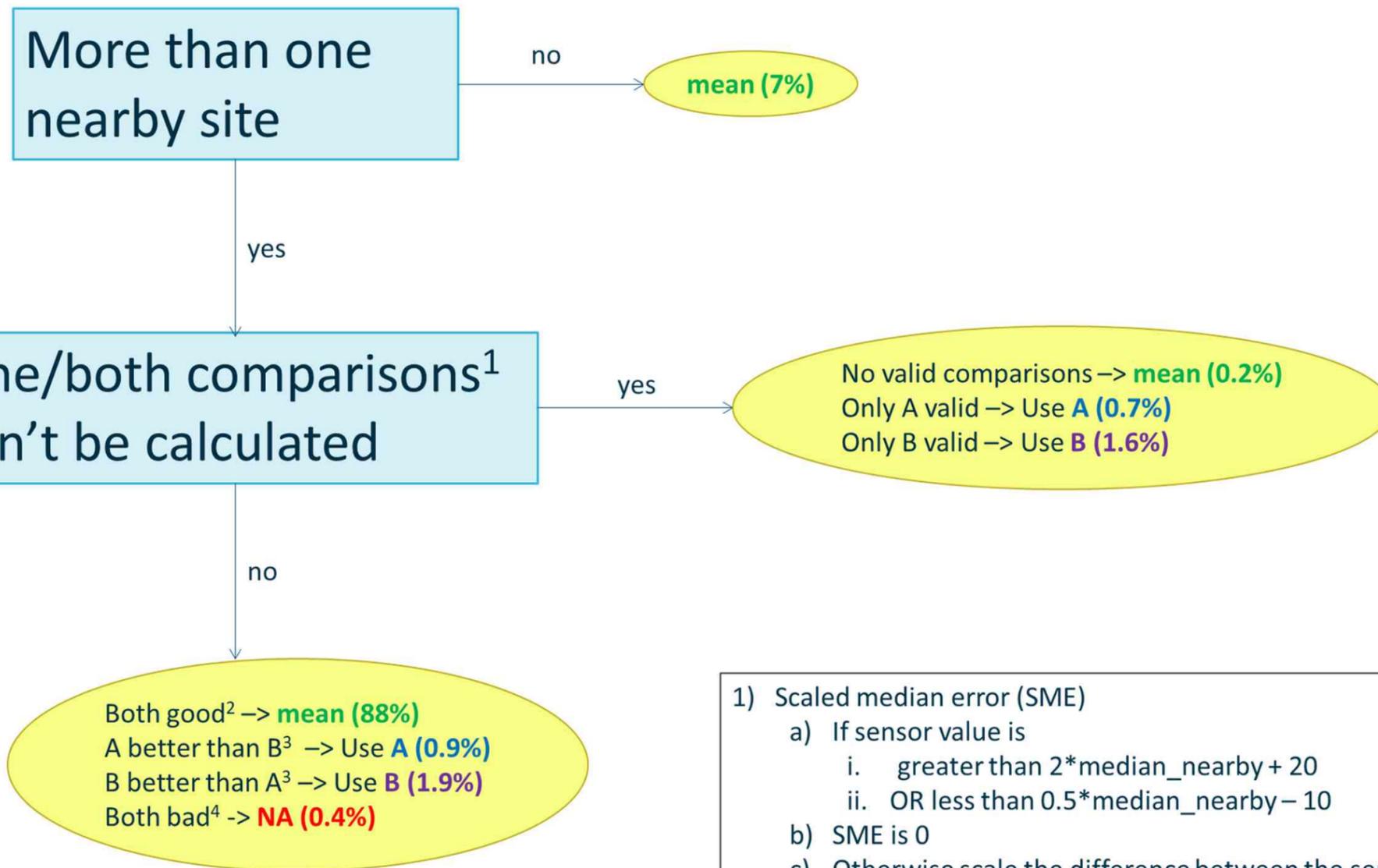
# Data Processing

- Too low when 0.3 $\mu$ m 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  $\mu$ g/m<sup>3</sup>
  - > 1000  $\mu$ g/m<sup>3</sup> possible in very high pollution environments
  - True “max” value is ~5000  $\mu$ g/m<sup>3</sup> when the plantower sensor can’t read any higher
  - 0.5% of sensor values were >1000  $\mu$ g/m<sup>3</sup> (out of these, 96% are also >3000  $\mu$ g/m<sup>3</sup>)
- 16-bit overflow - very high values are seen as 0s or low values
  - 83% of data >1000  $\mu$ g/m<sup>3</sup> also has <10 counts for 0.3 $\mu$ m particles

# Intra-monitor QC Flow Chart

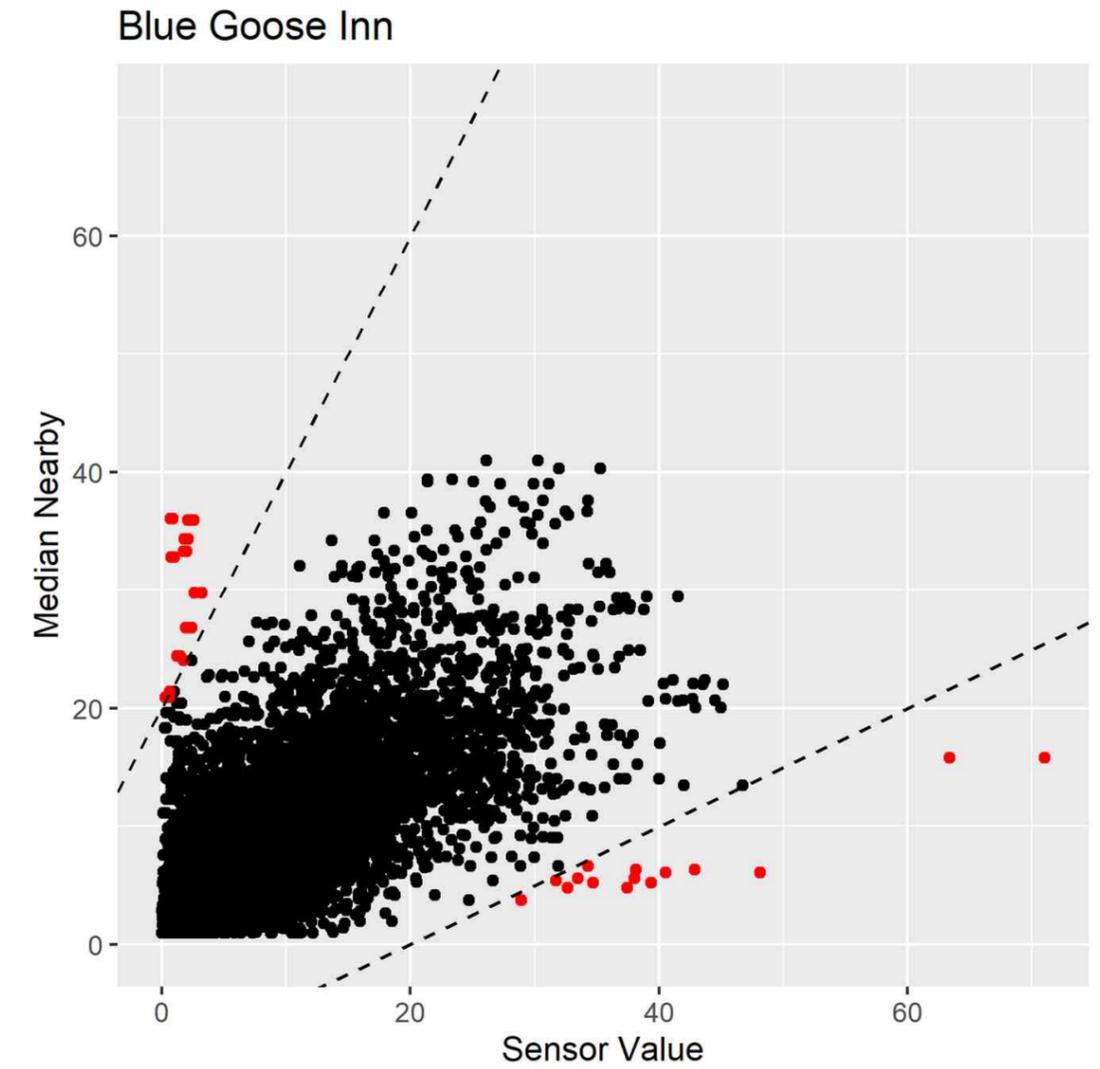


# Inter-monitor QC Flow Chart

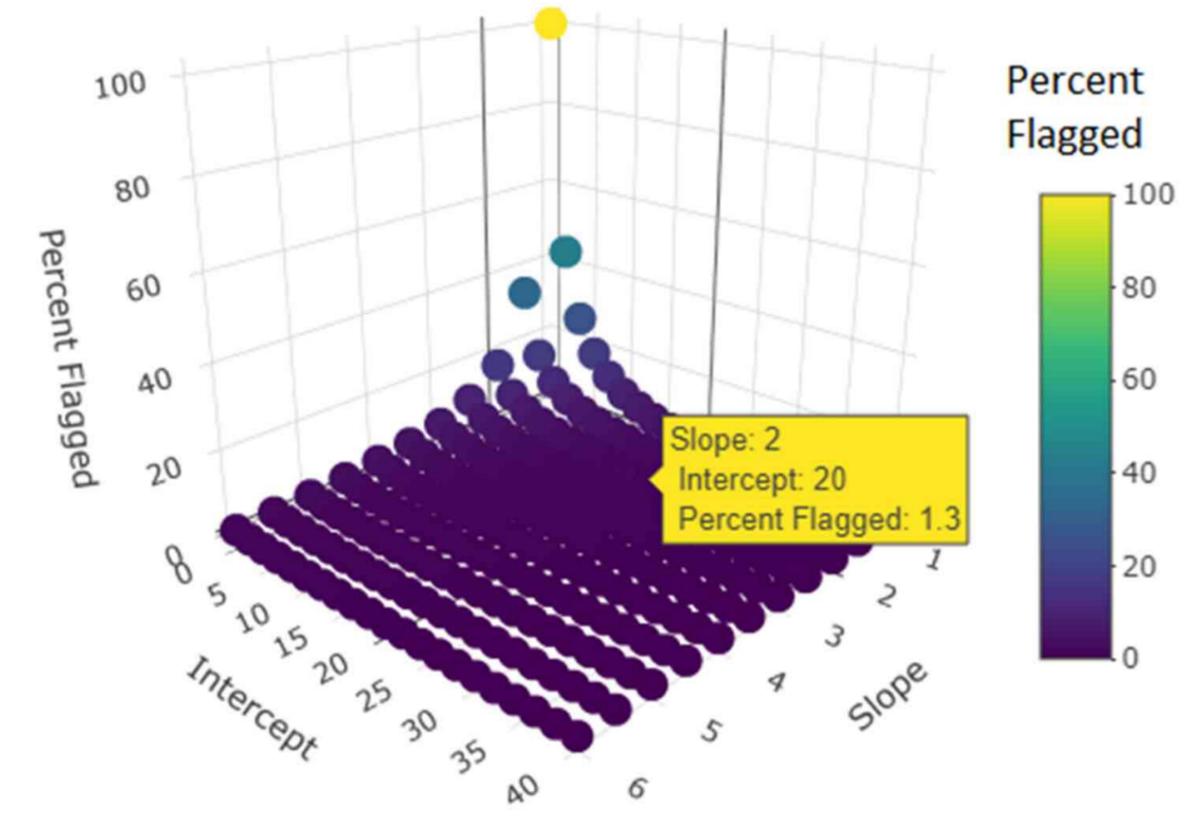
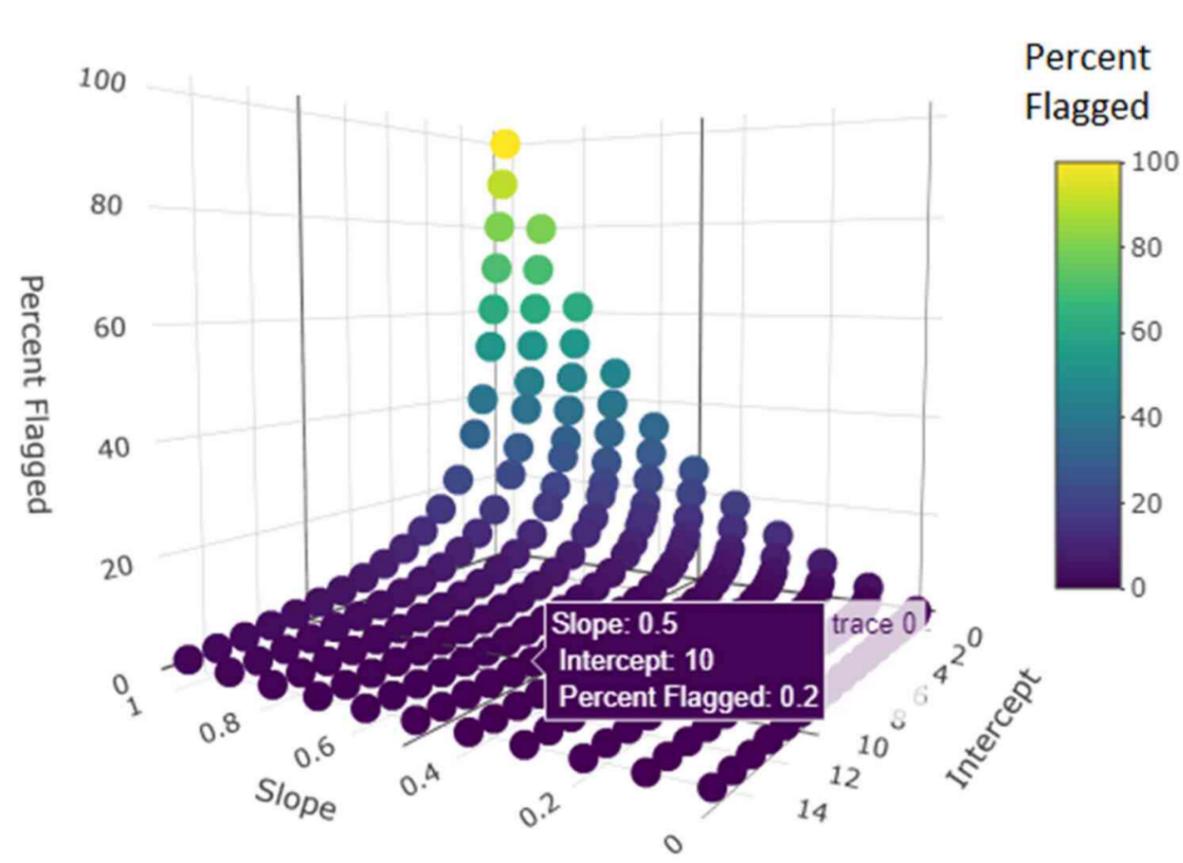
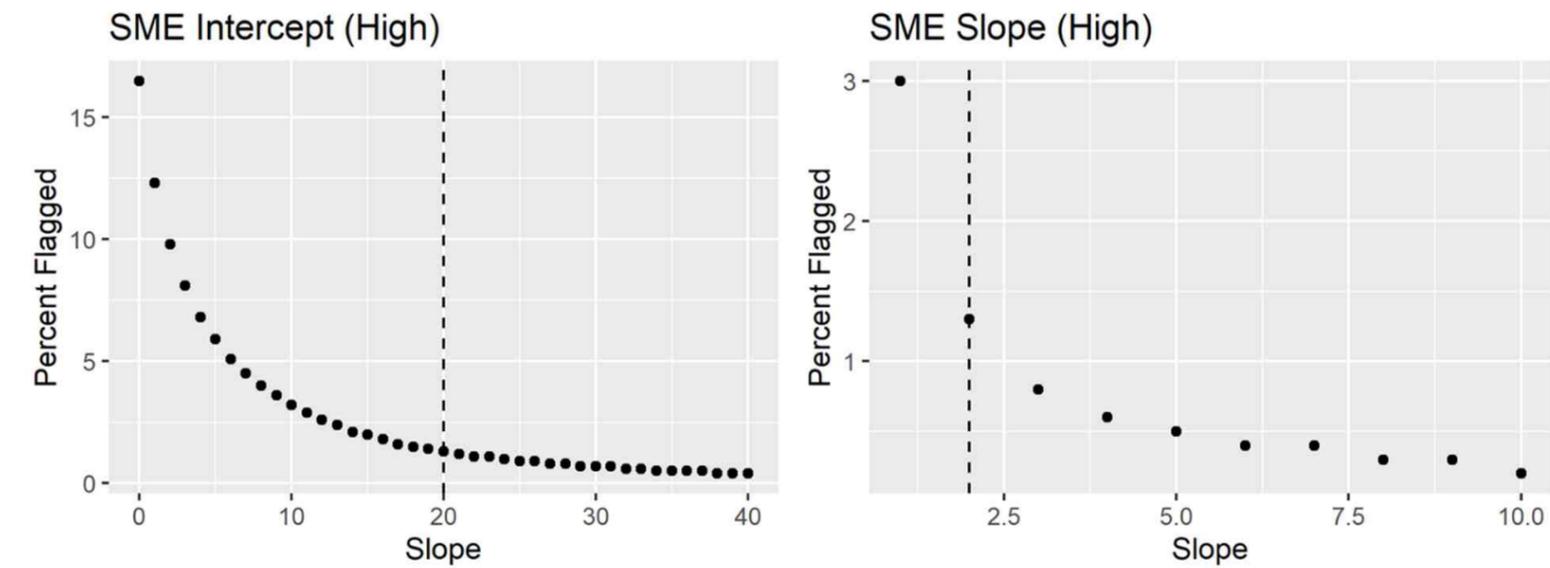
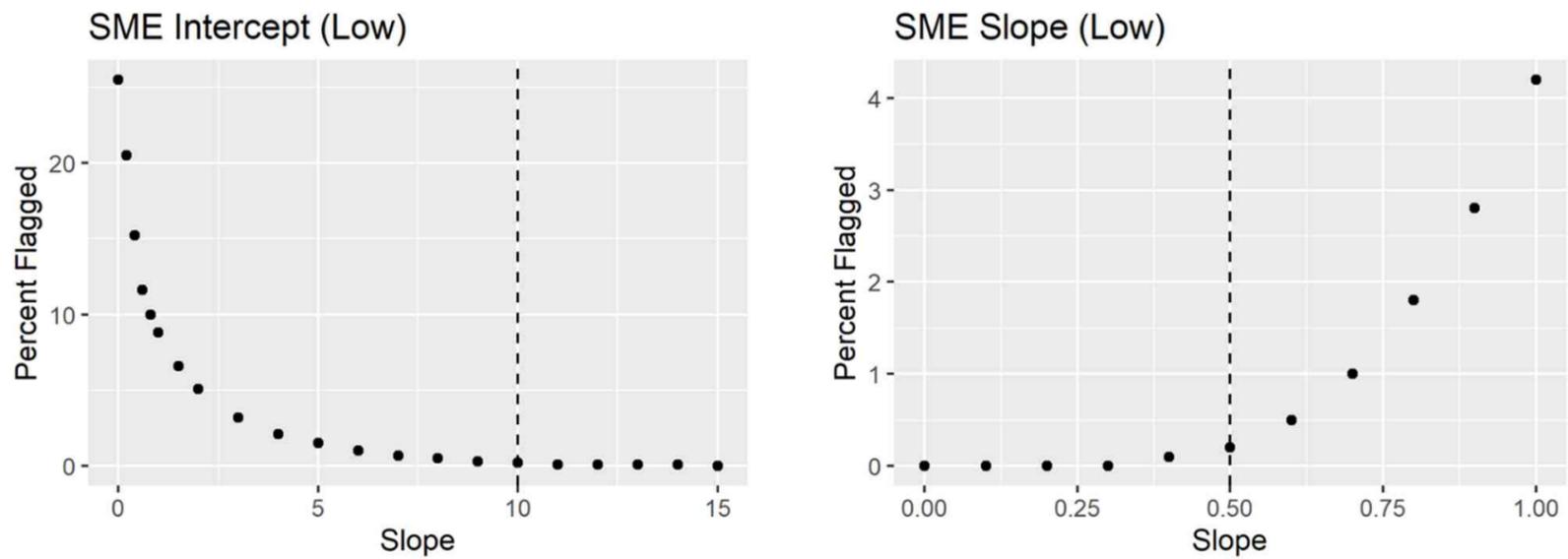


- 1) Scaled median error (SME)
  - a) If sensor value is
    - i. greater than  $2 * \text{median\_nearby} + 20$
    - ii. OR less than  $0.5 * \text{median\_nearby} - 10$
  - b) SME is 0
  - c) Otherwise scale the difference between the sensor value and median\_nearby from 0-1, where 1 is perfect agreement.
  - 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.



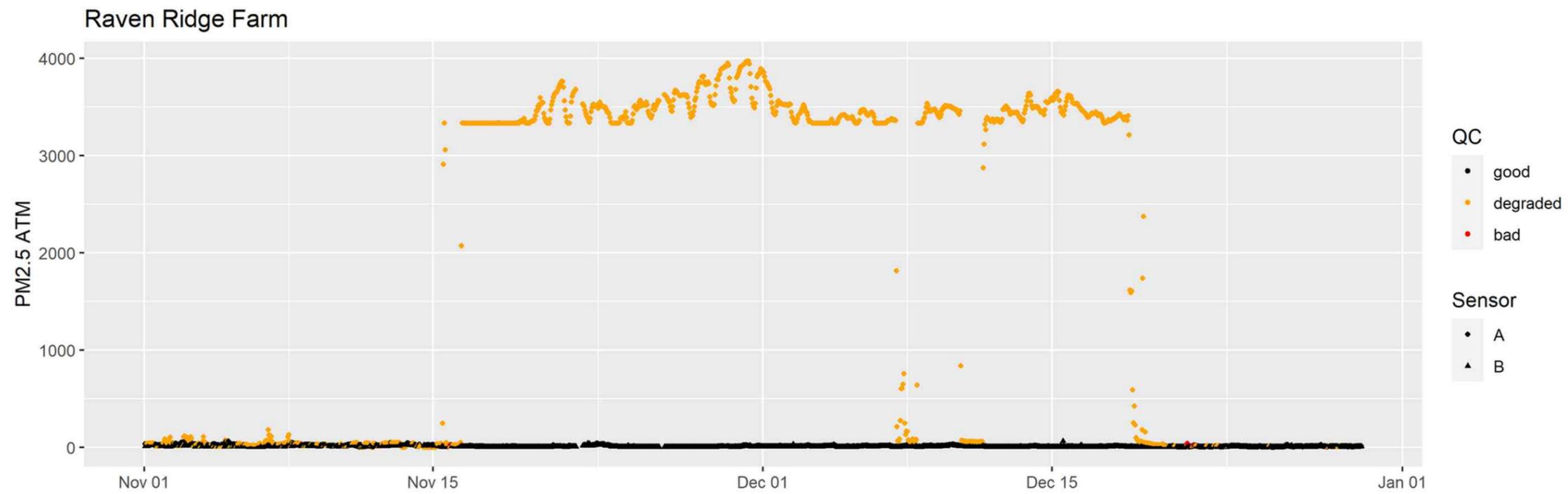
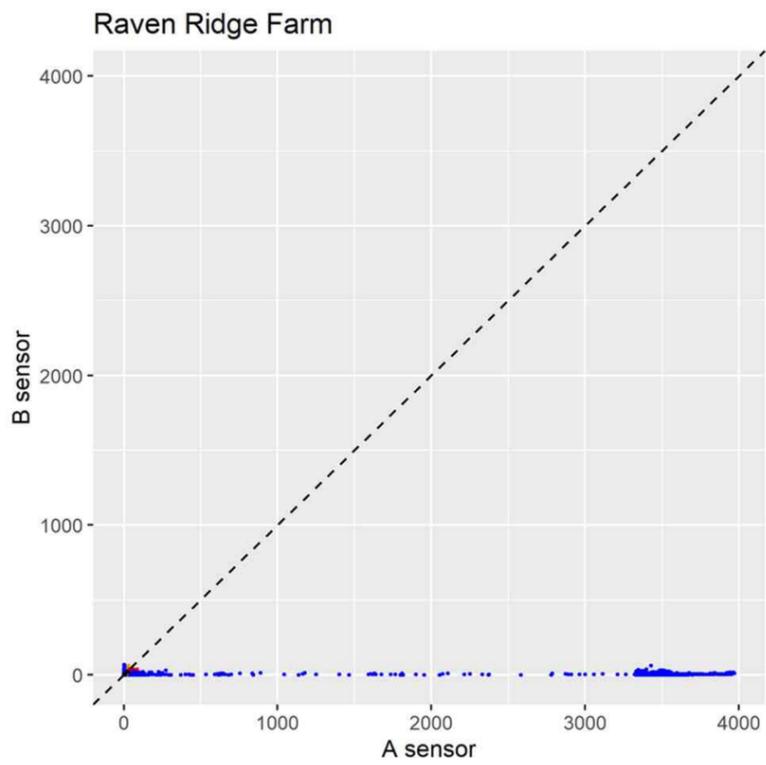
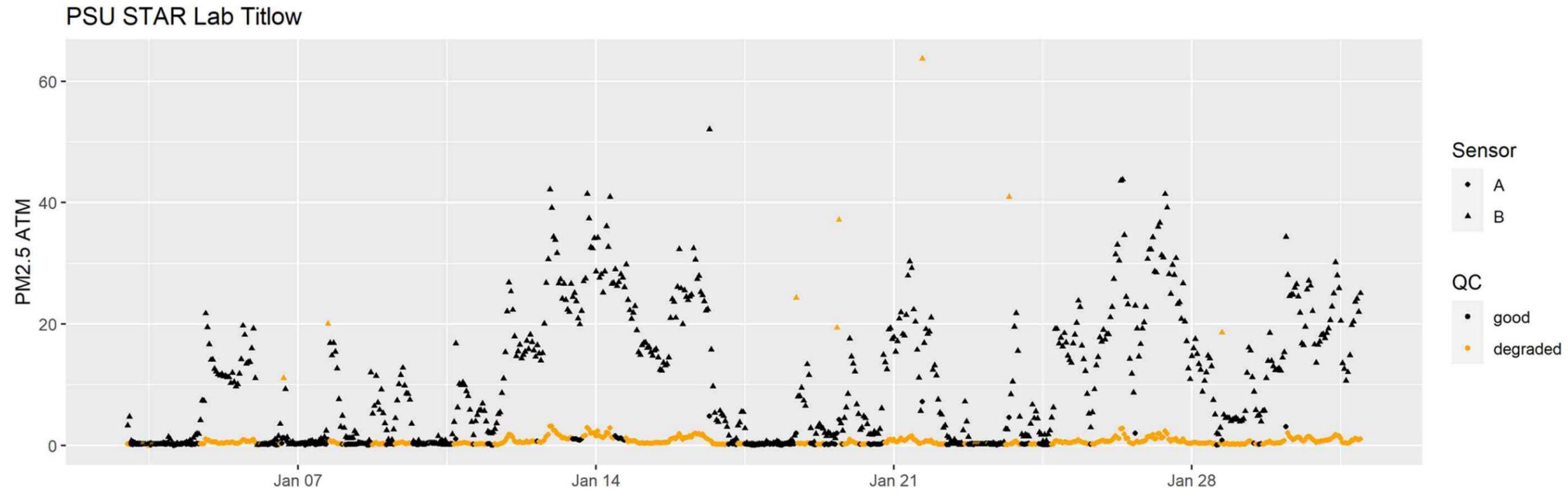
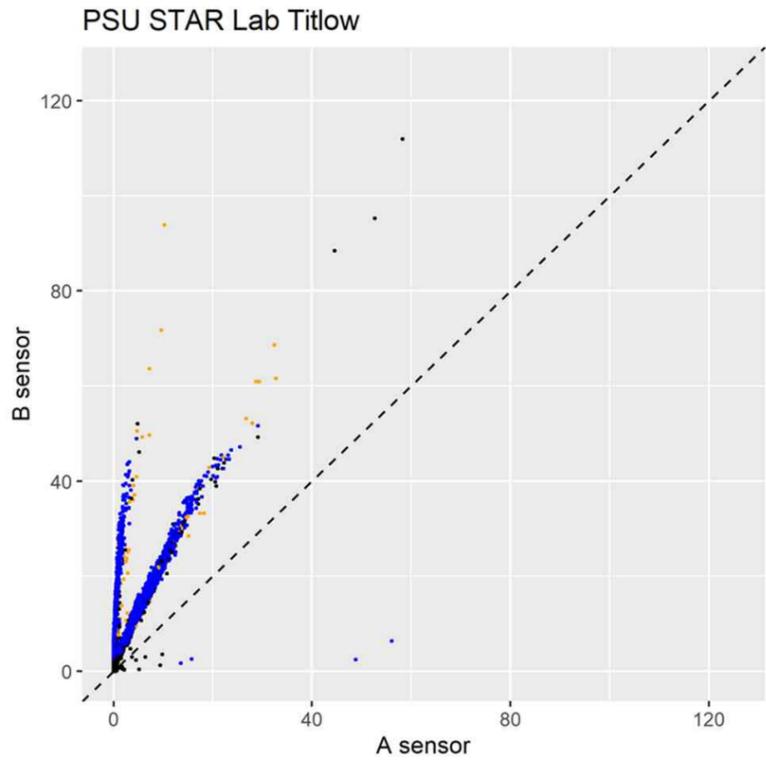
# Inter-monitor QC Cutoffs



# 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  $\mu\text{g}/\text{m}^3$  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

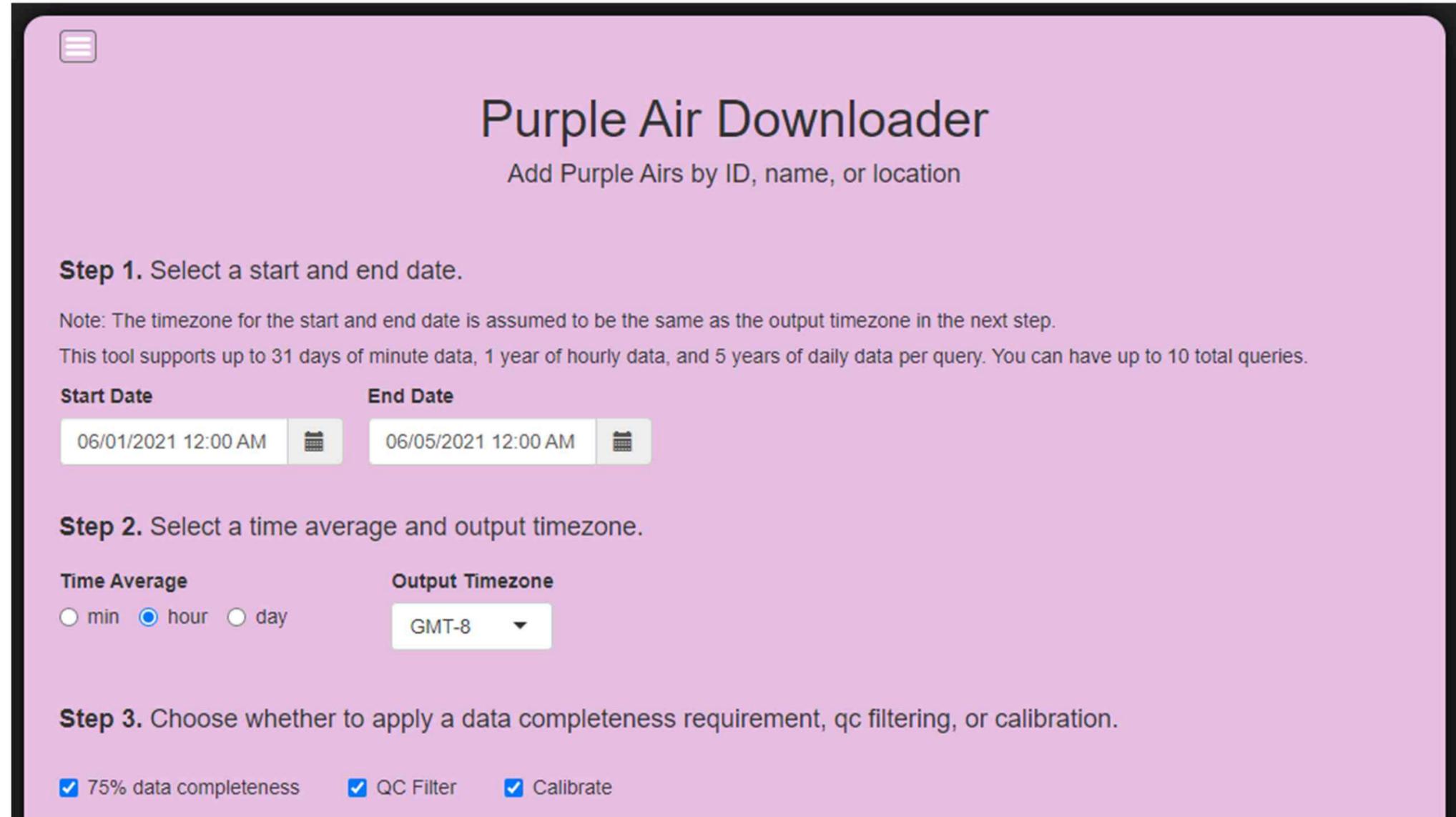
# Benefit of Inter-monitor QC





# Download

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



The screenshot shows the 'Purple Air Downloader' web interface. It features a purple header with a hamburger menu icon in the top left. The main title is 'Purple Air Downloader' with the subtitle 'Add Purple Airs by ID, name, or location'. Below the title, there are three steps: Step 1: 'Select a start and end date.' with a note about timezones and a query limit. Step 2: 'Select a time average and output timezone.' with radio buttons for 'min', 'hour', and 'day', and a dropdown for 'Output Timezone' set to 'GMT-8'. Step 3: 'Choose whether to apply a data completeness requirement, qc filtering, or calibration.' with three checked checkboxes: '75% data completeness', 'QC Filter', and 'Calibrate'.

**Purple Air Downloader**  
Add Purple Airs by ID, name, or location

**Step 1.** Select a start and end date.

Note: The timezone for the start and end date is assumed to be the same as the output timezone in the next step.  
This tool supports up to 31 days of minute data, 1 year of hourly data, and 5 years of daily data per query. You can have up to 10 total queries.

**Start Date**      **End Date**

06/01/2021 12:00 AM      06/05/2021 12:00 AM

**Step 2.** Select a time average and output timezone.

**Time Average**      **Output Timezone**

min    hour    day      GMT-8

**Step 3.** Choose whether to apply a data completeness requirement, qc filtering, or calibration.

75% data completeness    QC Filter    Calibrate

# Download

- Select by location, ID, name

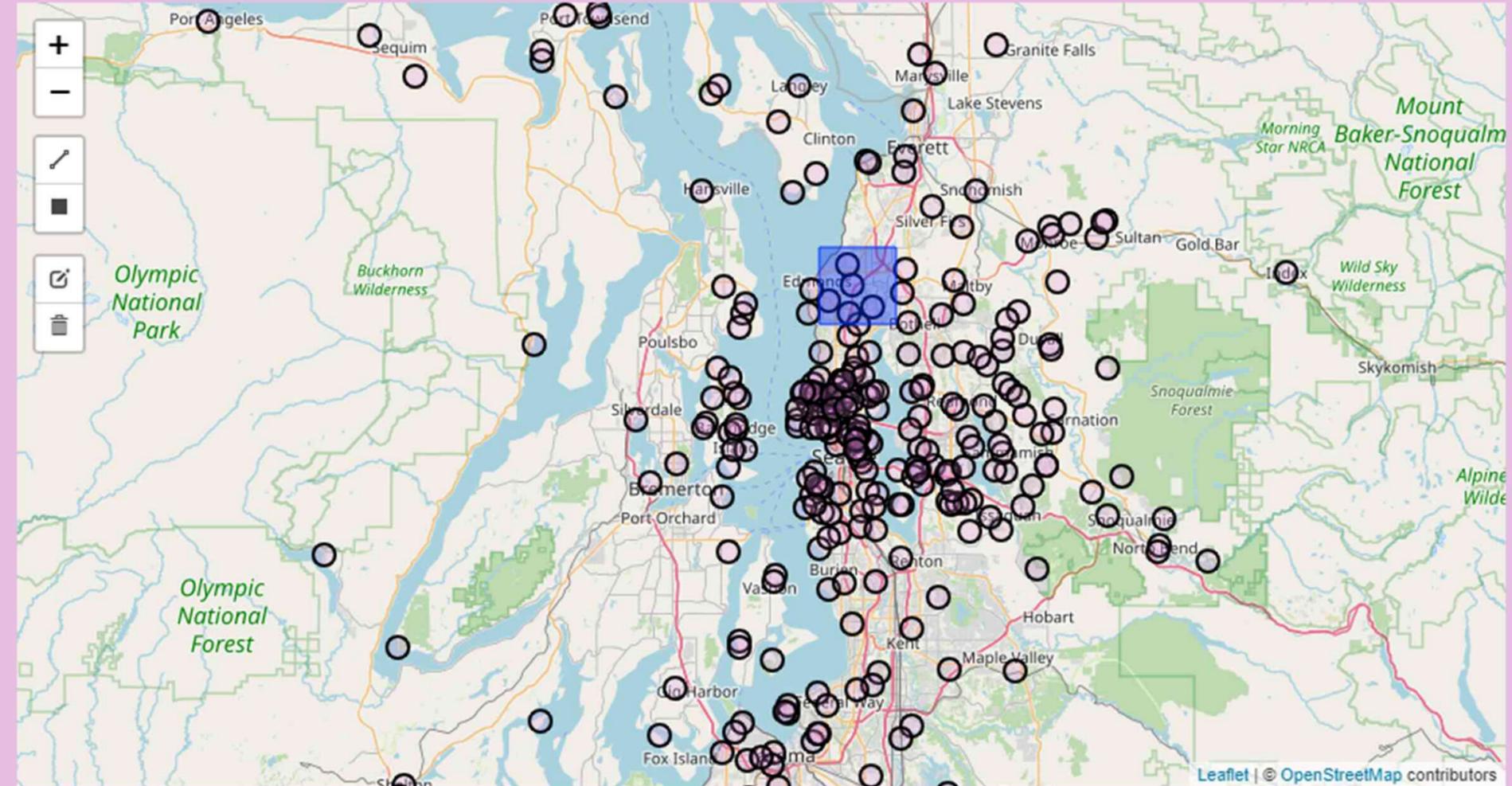
**Step 4.** Add sensors by location, ID, or name.

### Select Purple Airs by location

Add multiple sensors: Use the rectangle or polygon tool to select sensors. To delete the selection click the trash icon and Clear All.

When you are done, click the **Add** button located below the bottom right of the map.

Add a single sensor: Click on a sensor then click the **Add** button located below the bottom right of the map.



Add

Purple Air (Parent) ID

Add

Purple Air (Parent) Name

Add

# 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

[Send us feedback!](#)

# Report

- Fill in background information

## Background Information

Enter information about your air sensor experiment below.

Study Title

Darrington Study

Author

Graeme Carvlin

Study Time Zone

PST only

Note: Study time zone is the time zone where you recorded the data.

### 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.

### 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.

### Observations

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

Back

[Send us feedback!](#)

Next

# Report

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

## Upload and Label Data

**Background.** Upload your data then label it. A label is a group of measurements taken over a specific period of time in a specific place. For example, if you walked from your home to a road nearby, you could label the data recorded near your home as "home" and the data recorded near the road as "road". Then you can compare these two locations in the Hypothesis section of the app to see if pollution was higher near the road.

**Step 1.** Click Browse to choose a data file (Max 30MB). Select the sensor type and time zone. Then click Add.  
Note: For Purple Airs, add the primary and secondary data files for the A and B sensors. All 4 files can be uploaded at once. You can download Purple Air data from the Purple Air website or PSCAA Download app.

Select a file      Type of sensor      Sensor time zone

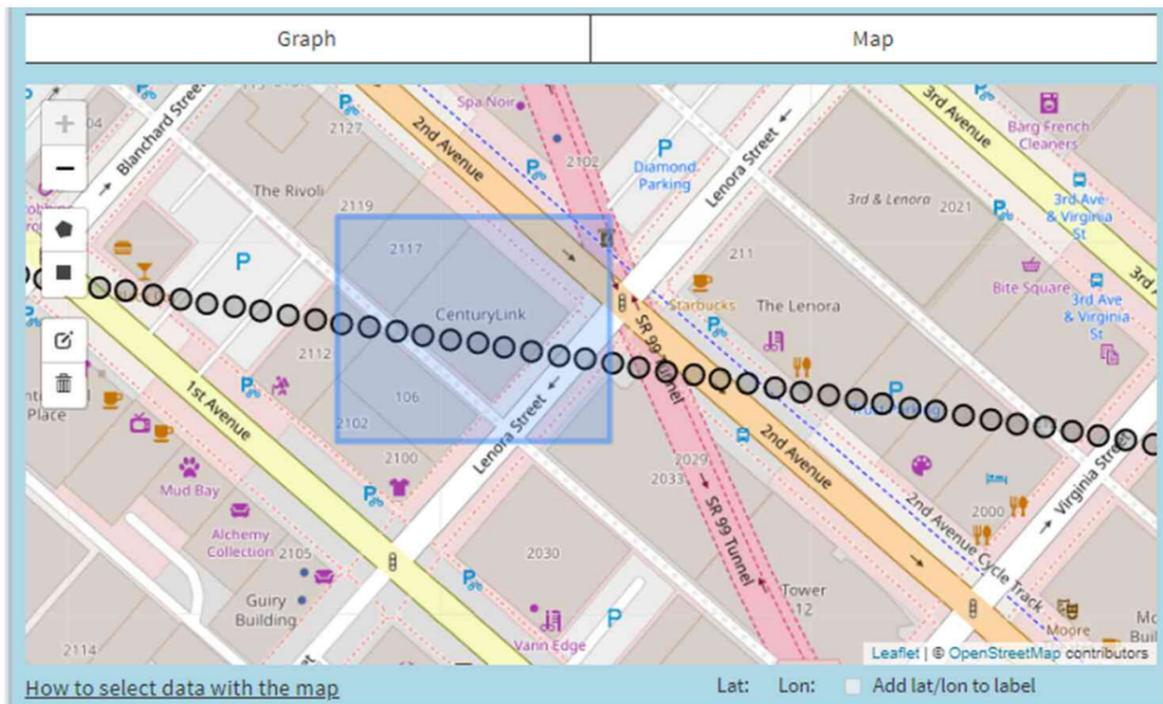
Browse...    No file selected    PurpleAir ▼    GMT/UTC ▼    Add

**Step 2.** Click on the file in the file list.

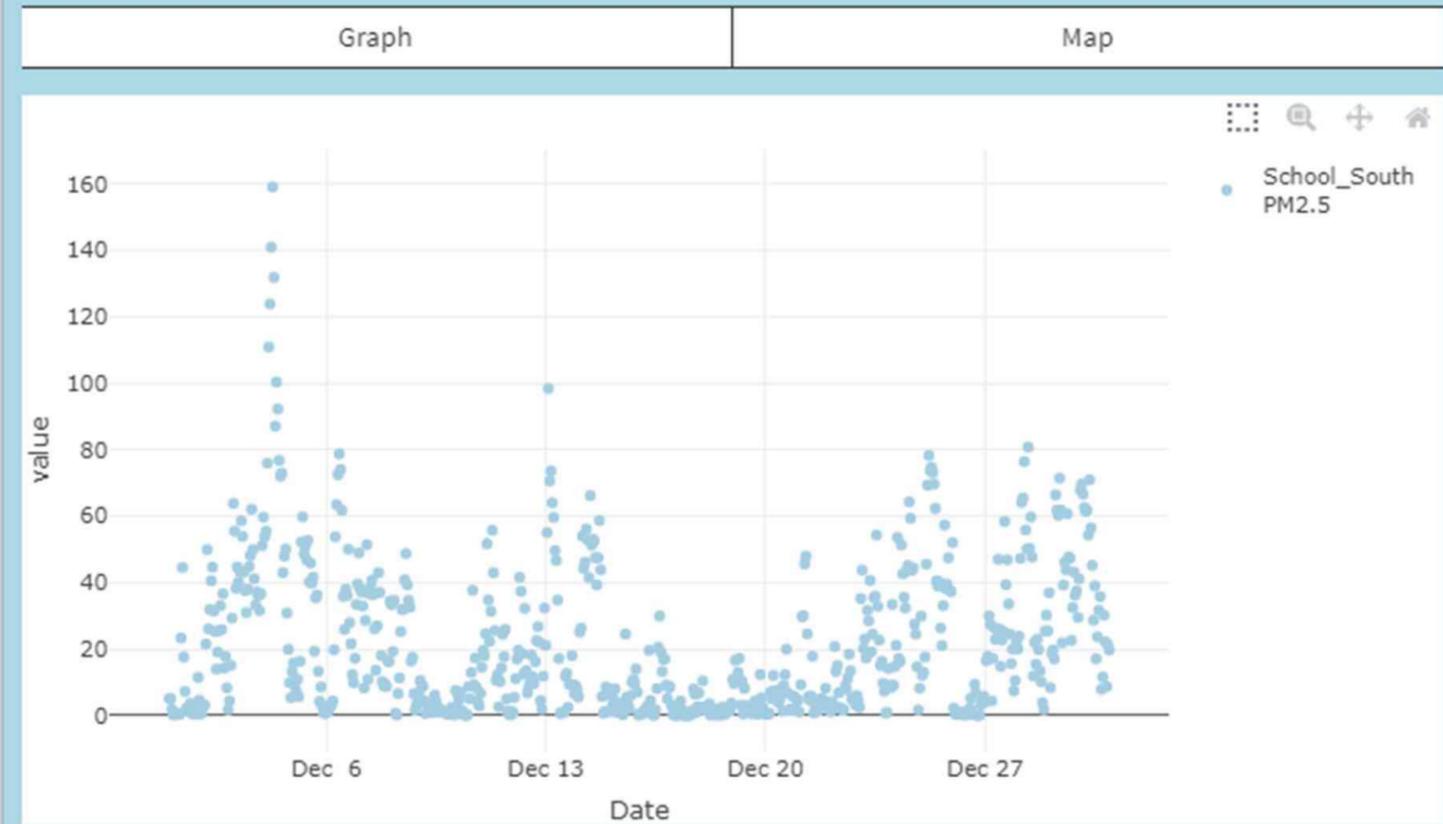
File	Sensor Type	Time Zone	Remove
Darrington Elementary School_15229_1.csv	PurpleAir	GMT/UTC	X
Darrington Library_21049_2.csv	PurpleAir	GMT/UTC	X
Darrington Airport_21011_3.csv	PurpleAir	GMT/UTC	X

# 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

# Report

- Calibrate

**Step 1.** Select the label and pollutant to use in the calibration equation.

Create calibration with this label and this pollutant

School\_South PM2.5

**Step 2.** Select the regulatory site to use in the calibration equation.

Reg State Reg County Reg Site Distance to site

WA SNOHOMISH Darrington-Fir St 0.01 miles

Note: "Distance to site" shows the distance between the selected regulatory site and label, if it has location information.

**Step 3.** Select the label(s) to calibrate and click Add.

Apply calibration to these label(s)

School\_South Library\_Central Airport\_North Add

Label	Pollutant	Reference Site	Labels to Calibrate	Remove
School_South	PM2.5	Darrington-Fir St	School_South,Library_Central,Airport_North	X

Back [Send us feedback!](#) Next

# Report

- Add hypotheses
- Site and time of day comparisons

## Create hypotheses (optional)

**Background.** Create a hypothesis to compare two labels or perform an hour of day test on a single label. For example, if you used a Dylos to measure pollution at your home (labeled "home") and a nearby road (labeled "road") you can select "road - PM Small" in the first box, "is greater than" in the second box, and "home - PM Small" in the third box. If you wanted test if the pollution at your home was the highest at night you could select "home - PM Small" in the first box, "is highest" in the second box, and " at night" in the third box. You do not need to create a hypothesis to Analyze your data and generate a report.

**Step 1.** Select a label-pollutant, comparator, and a second comparison object.

**Step 2.** Once you have added any hypotheses click Next.

Library\_Central - PM2.5 ▼ is highest ▼ in the evening (6pm - 12am) ▼ Add

Object1	Comparator	Object2	Remove
School_South - PM2.5	is greater than	Airport_North - PM2.5	X
Library_Central - PM2.5	is highest	in the evening (6pm - 12am)	X

Back

[Send us feedback!](#)

Next

# Report

- Download Report and data file

## Analyze data and download report

**Step 1.** Click **Analyze Data** to process your data.

[Analyze Data](#)

---

**Step 2.** Click **Download Report** to get a pdf report with an analysis of your data.

*Optional.* Click **Download Data** to get an excel file with the raw, QC'd, and calibrated data from your study.

[Back](#) [Send us feedback!](#) [Download Data](#) [Download Report](#)

## 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.

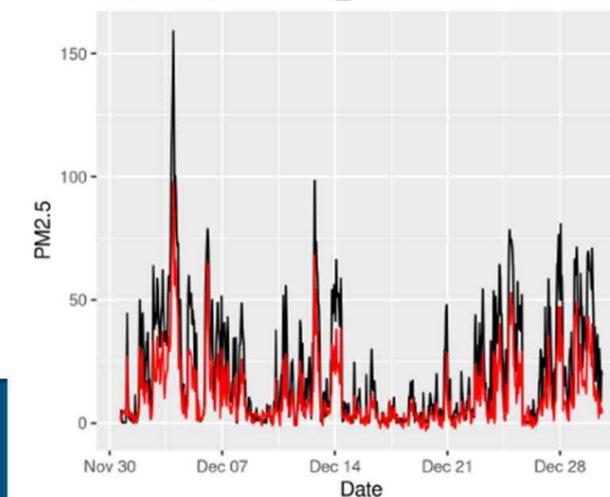
## 4 Summary Statistics

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

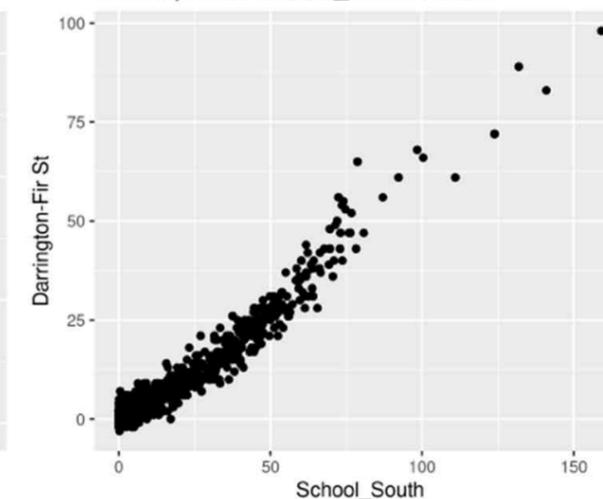
Label	Pollutant	Min	Avg	Max	Trend
Airport_North	PM2.5	-0.78	9.21	66.58	None
Library_Central	PM2.5	-0.74	16.2	139.26	None
School_South	PM2.5	-0.78	11.37	89.86	None

## 5 Calibration

Timeseries for School\_South PM2.5



Scatterplot for School\_South PM2.5

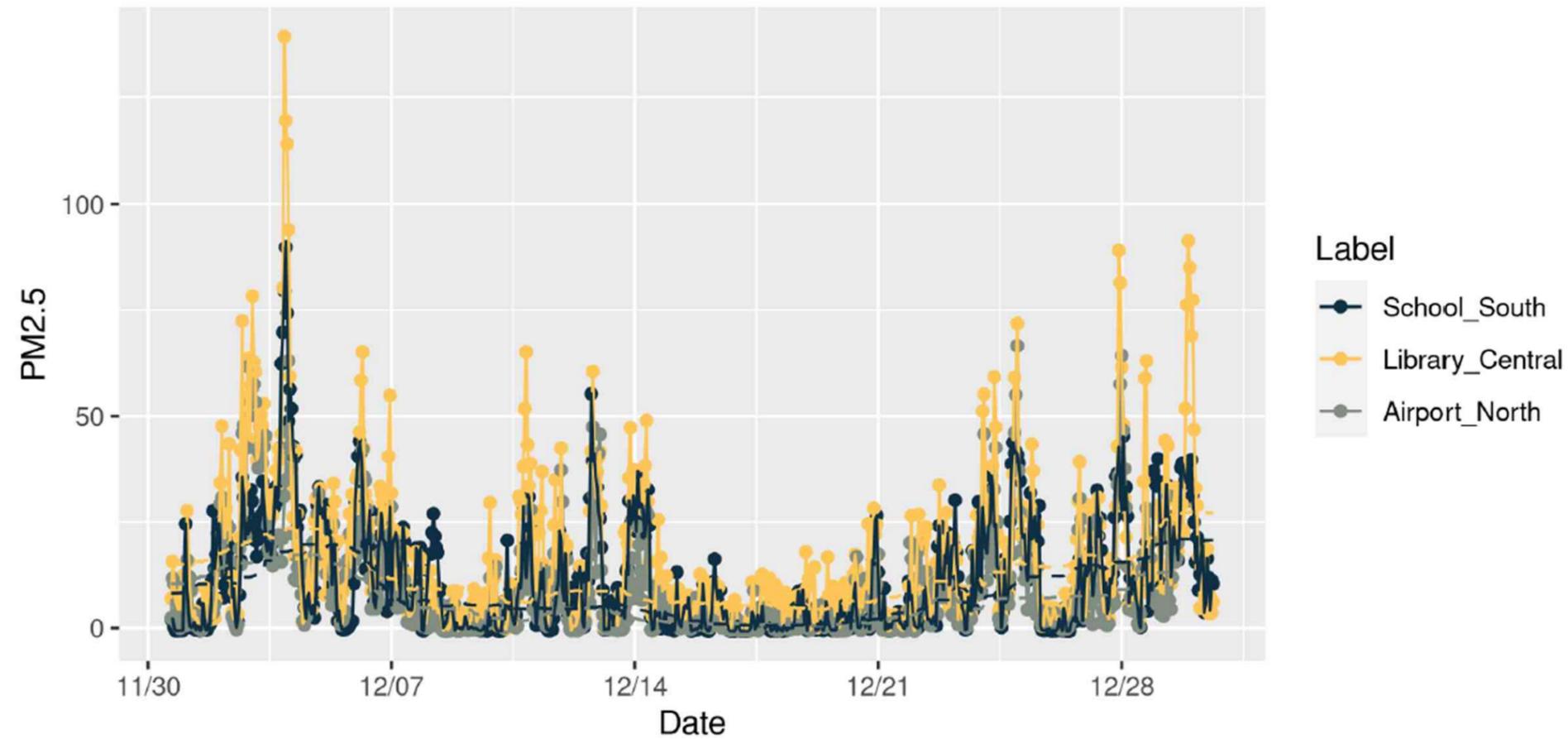


A calibration equation for PM2.5 was developed using data from School\_South and reference data from Darrington-Fir St. There were 711 overlapping observations. The slope was 0.57 and the intercept was -0.78. The performance of the calibration equation was good, with an R-squared of 0.93 and a root mean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

## 6 Figures

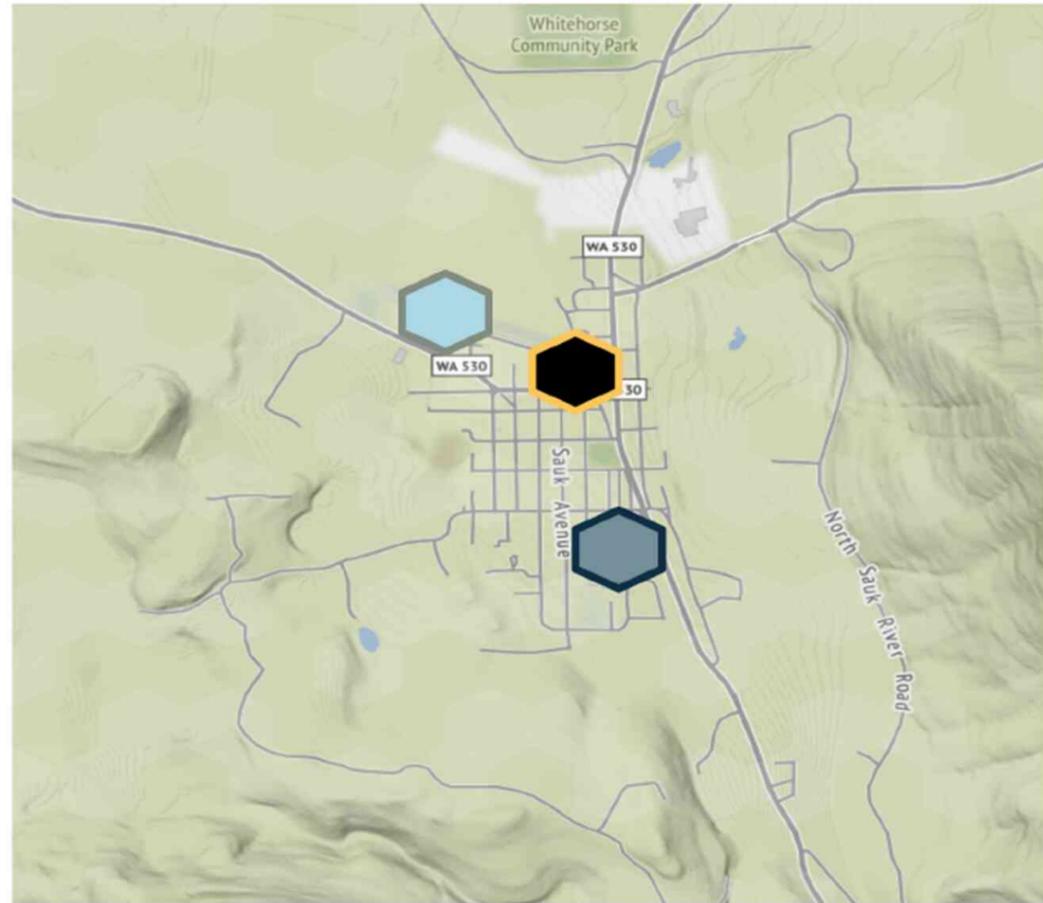
Solid lines are data, dashed lines are a running average of the data.

Figure 1a. Timeseries of PM2.5  
2020-11-30 4pm to 2020-12-30 3pm



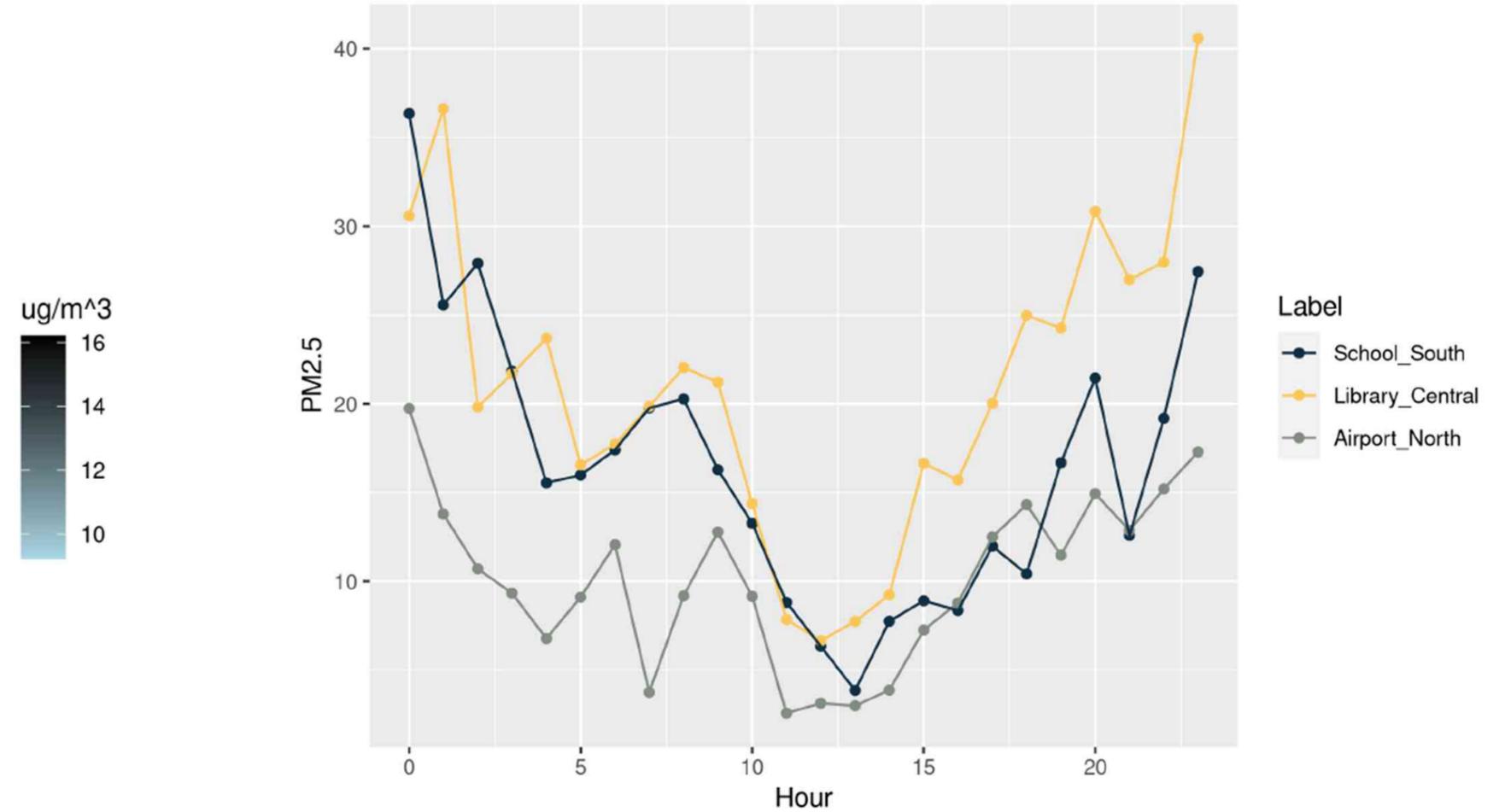
For School\_South, average PM2.5 levels were high from 12/25/20 3pm to 12/30/20 3pm. PM2.5 levels spiked from 12/03/20 8pm to 12/04/20 3am and 12/12/20 6pm. For Library\_Central, PM2.5 levels spiked from 12/03/20 9pm to 12/04/20 2am, 12/27/20 10pm to 12/28/20 12am, and 12/29/20 9pm to 12/30/20 1am. For 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.

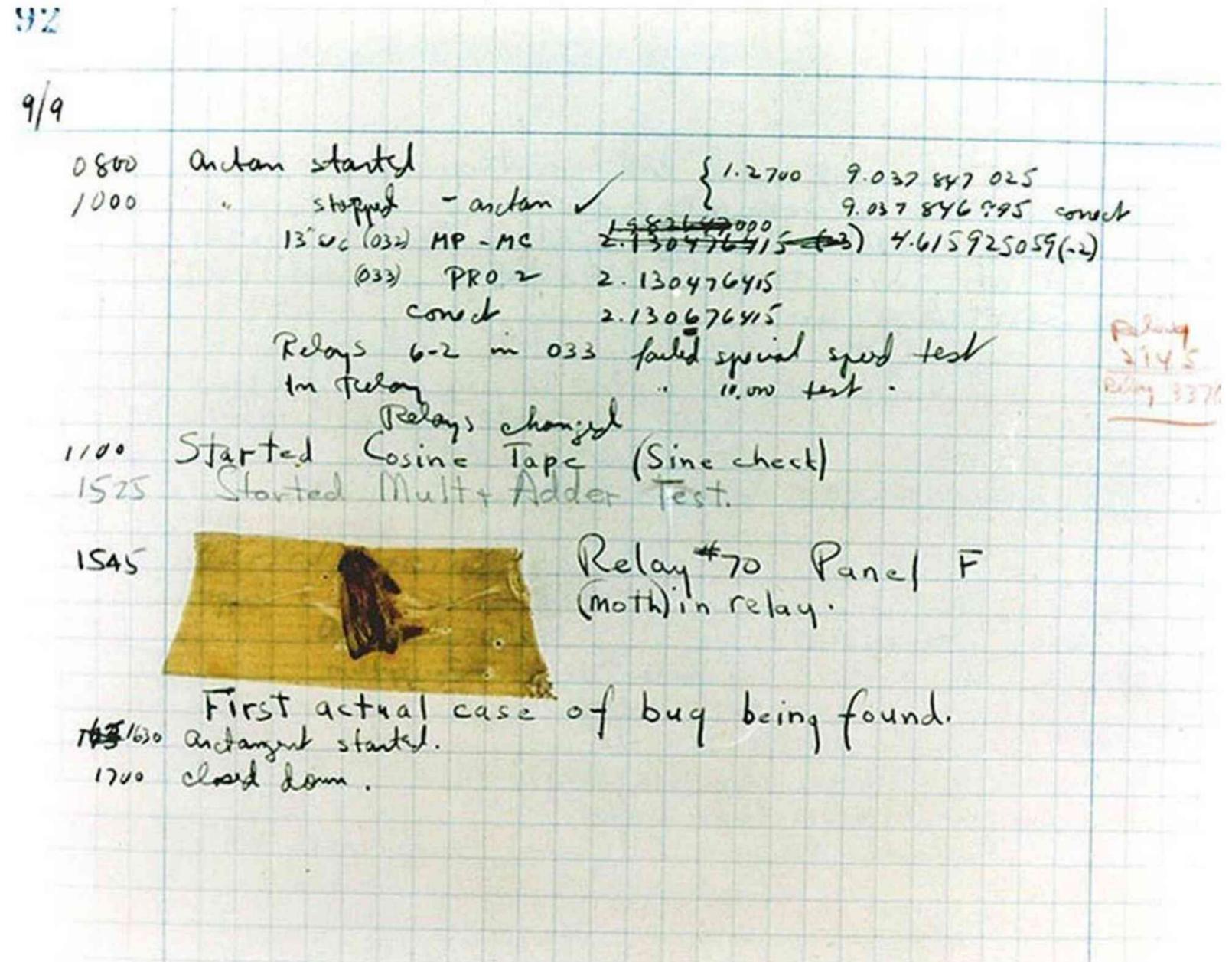
## 7 Hypotheses

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<sup>3</sup> 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?

Photo # NH 96566-KN (Color) First Computer "Bug", 1947



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

Thank you!

Questions?