



# UNLOCKING THE VALUE IN SENSOR DATA

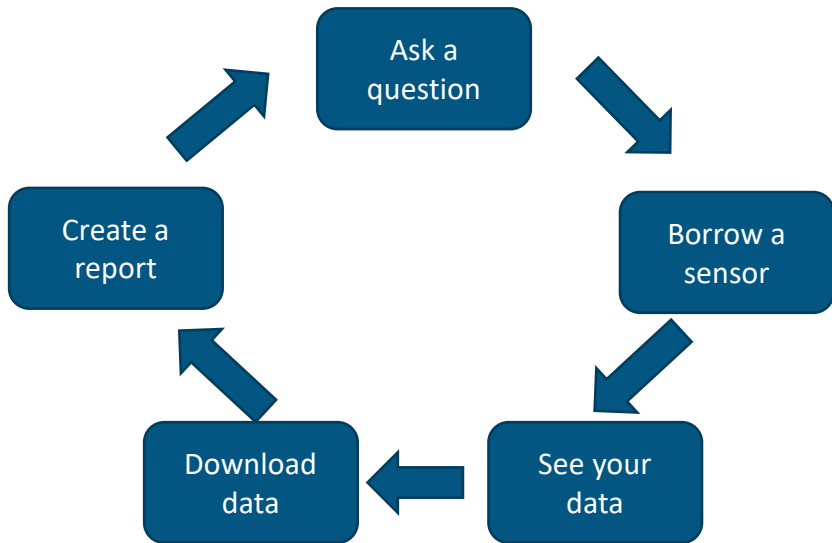
Graeme Carvlin

**Puget Sound Clean Air Agency**

May 11, 2022



# Our sensor outreach process



# A few statistics from the program so far

- Pilot launched in **2018!**
- **2020:**
  - ~8 applications received
  - **12** sensors distributed (5 Purple Air & 7 Dylos)
- **2021:**
  - **10** applications received
  - **15** sensors distributed (7 Purple Air & 8 Dylos)
- Spike of applications after wildfire smoke events (Sept 2020 and Aug 2021)
- Currently 19 Purple Air + 12 Dylos deployed

## Air Sensor Fact sheet available for the public:

### Air Sensors + Your Community

#### Checking Your Air Quality Is Easier Than Ever

Did you know this location is measuring real-time, hyper-local air quality?

Attached to this building is a device called a PurpleAir sensor. The sensor measures fine particles or PM<sub>2.5</sub>, a type of air pollution that contributes to asthma, heart attacks, strokes, cancer, and premature death.

#### HOW TO VIEW YOUR AIR QUALITY

You can check your community's air quality at: [map.pscleanair.gov](http://map.pscleanair.gov).

Zoom in to find your location. Each dot shows a PurpleAir sensor and the stars show the monitoring sites operated by government agencies.

You can click on a dot or star to bring up a graph of hourly levels of fine particle pollution over the last two days. With that graph still up, you can click on other dots to see how their pollution levels compare.



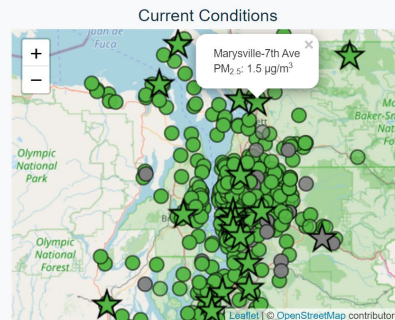
# Dashboard

- Current Conditions
- Forecast
- Timeseries
- <http://apps.pscleanair.gov/>

## Air Sensor Dashboard

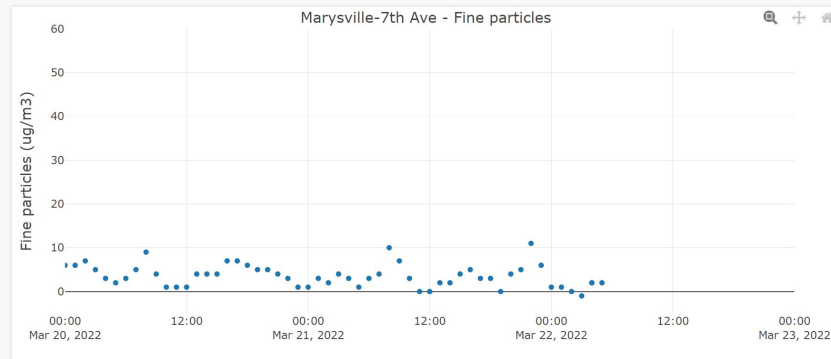
View, Download, or Analyze sensor data

Map	Download	Analyze
-----	----------	---------



### Selected Site

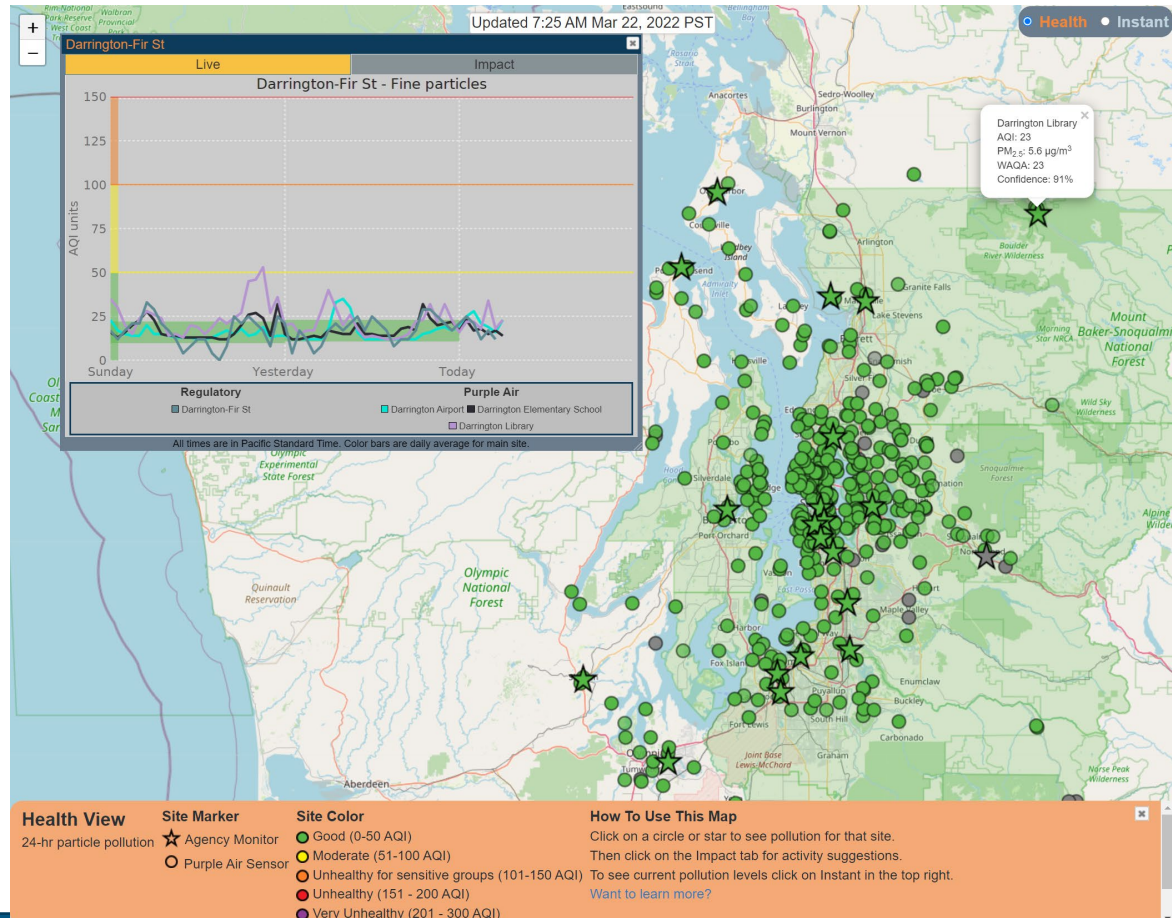
Click on a site on the Current Conditions map to graph it





# Map

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



# Surface

**Options** ⬆

**Choose Date Range**

2021-11-01 to 2022-02-28

- Include wildfire days
- Include outliers
- Include July 3rd-5th

**Function to apply**

98th ▾

**Kriging distance**

5000 ▾

**Name for time period**

Winter '21-22

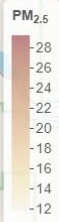
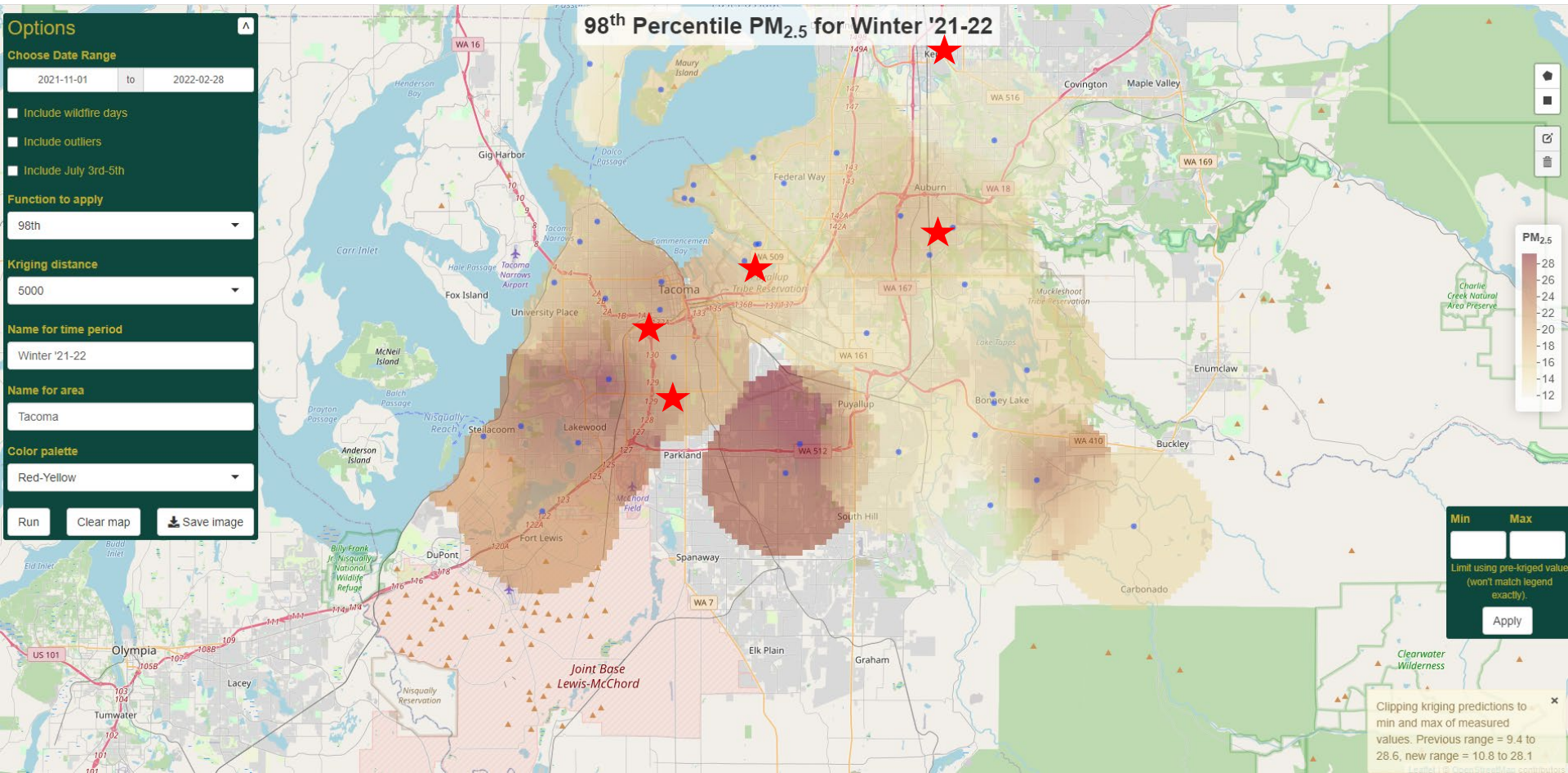
**Name for area**

Tacoma

**Color palette**

Red-Yellow ▾

## 98th Percentile PM<sub>2.5</sub> for Winter '21-22



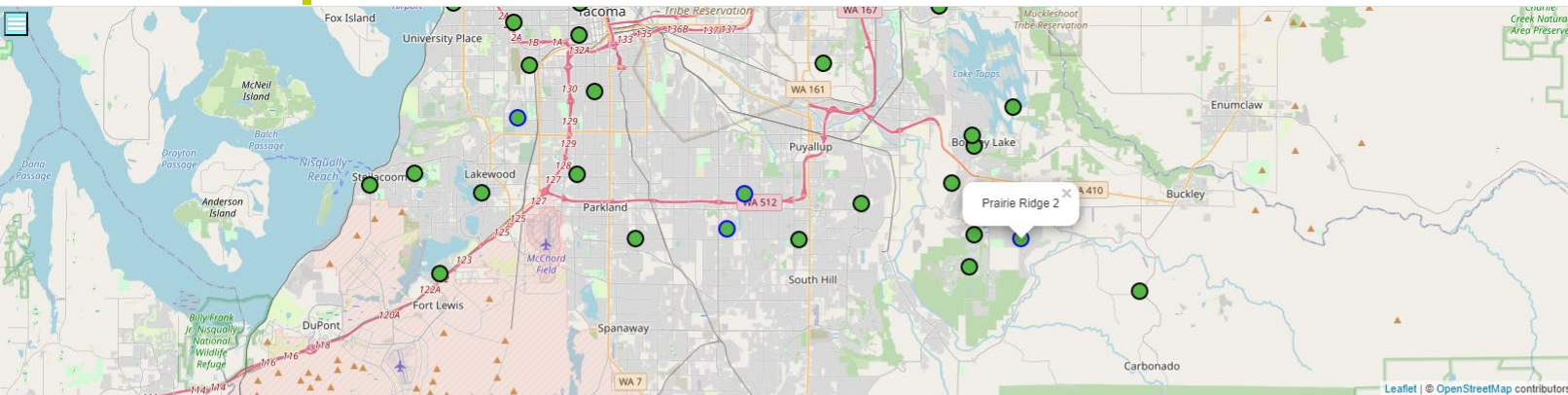
**Min**  **Max**

Limit using pre-kriged values (won't match legend exactly).

Clipping kriging predictions to min and max of measured values. Previous range = 9.4 to 28.6, new range = 10.8 to 28.1



# Compare



**Prairie Ridge 2**

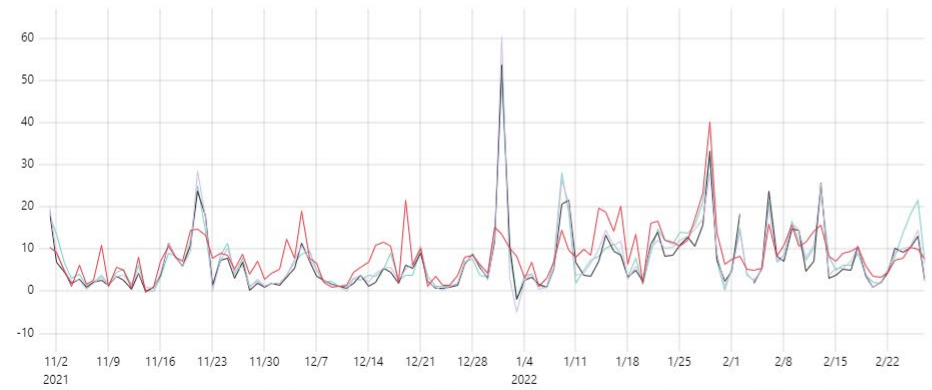
4/4

Hour Day Month Year Custom

Start: Mon Nov 01 2021  
End: Mon Feb 28 2022

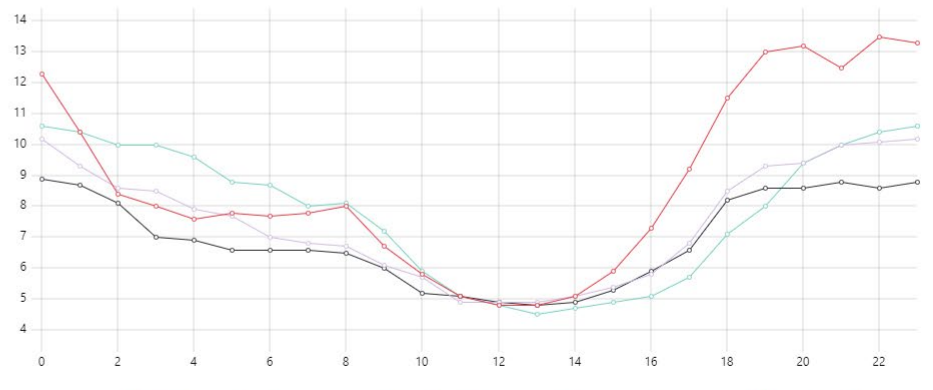
This site: 8.7  $\mu\text{g}/\text{m}^3$   
Other sites: 7.4  $\mu\text{g}/\text{m}^3$   
Range: 11/1/2021 to 2/28/2022

Timeseries



Time: -- ■ PSU Star Lab South Tacoma: -- ■ Parker Deck Summit: -- ■ Woodland/Summit: -- ■ Prairie Ridge 2: --

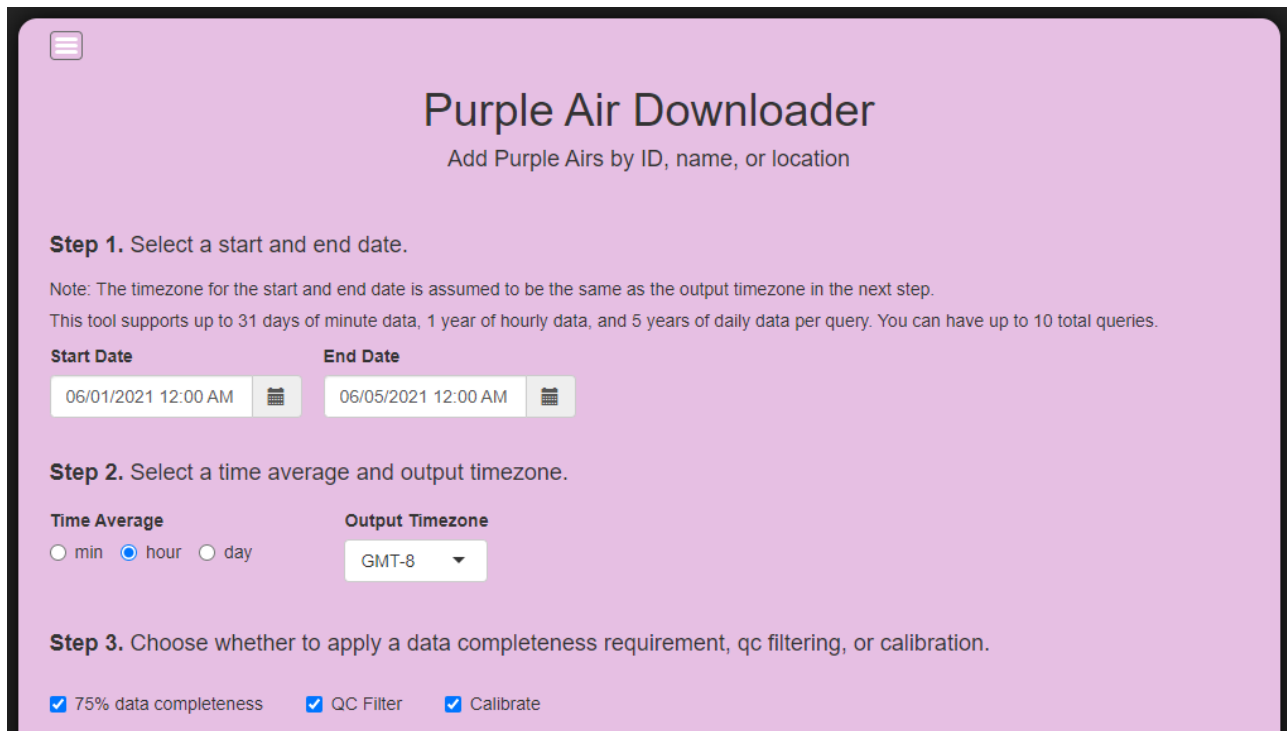
Hour of Day



Hour: -- ■ PSU Star Lab South Tacoma: -- ■ Parker Deck Summit: -- ■ Woodland/Summit: -- ■ Prairie Ridge 2: --

# Download

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



The screenshot shows the 'Purple Air Downloader' web interface. At the top, it says 'Add Purple Airs by ID, name, or location'. Below this, there are three steps:

**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

**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.

Purple Air (Parent) ID  Add

Purple Air (Parent) Name  Add

# Report

- Fill in background information

## Background Information

Enter information about your air sensor experiment below.

Study Title	Author	Study Time Zone
<input type="text" value="Darrington Study"/>	<input type="text" value="Graeme Carvlin"/>	<input style="border: 1px solid #ccc;" type="text" value="PST only"/>

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

**Project Description**

**Site Information**

**Observations**

[Send us feedback!](#)

# Report

## •Upload

### 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	
<input type="button" value="Browse..."/> <input type="text" value="No file selected"/>	<input type="text" value="AirBeam"/> ▼	<input type="text" value="GMT/UTC"/> ▼	<input type="button" value="Add"/>

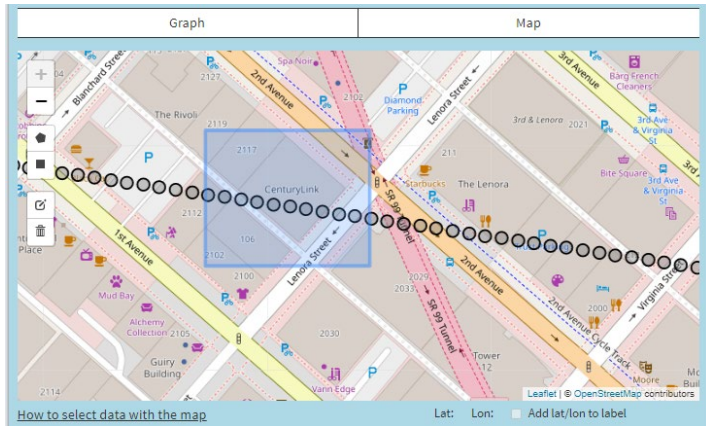
**Step 2.** Click on the file in the file list.

File	Sensor Type	Time Zone	Remove
PSU Star Lab South Tacoma_7384_1.csv	PurpleAir	Pacific	<input type="button" value="X"/>
WoodlandSummit_71029_2.csv	PurpleAir	Pacific	<input type="button" value="X"/>
Parker Deck Summit_81091_3.csv	PurpleAir	Pacific	<input type="button" value="X"/>
Prairie Ridge 2_48783_4.csv	PurpleAir	Pacific	<input type="button" value="X"/>

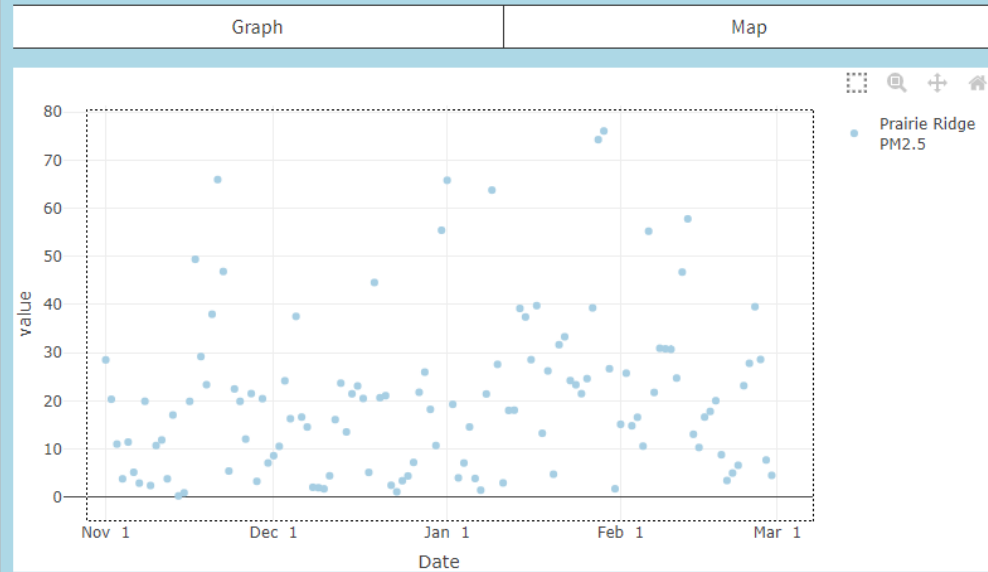


# Report

•Label



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



Step 4. Name and add the label.

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

Pollutant:

Label name:

# Report

## •Calibrate

### Calibrate (optional)

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

Create calibration with this label and this pollutant

South Tacoma  PM2.5

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

Reg State	Reg County	Reg Site	Distance to site
WA <input type="text"/>	PIERCE <input type="text"/>	Tacoma-L St <input type="text"/>	2.37 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)

Default to EPA calibration for Purple Air data.

Label	Pollutant	Reference Site	Labels to Calibrate	Remove
No data available in table				

[Send us feedback!](#)

# 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.

South Tacoma - PM2.5 ▼ is greater than ▼ North Summit - PM2.5 ▼ Add

Object1	Comparator	Object2	Remove
Prairie Ridge - PM2.5	is greater than	All other labels - PM2.5	X
North Summit - PM2.5	is greater than	All other labels - PM2.5	X
South Tacoma - PM2.5	is highest	in the evening (6pm - 12am)	X
North Summit - PM2.5	is highest	in the evening (6pm - 12am)	X
South Summit - PM2.5	is highest	in the evening (6pm - 12am)	X

Back

[Send us feedback!](#)

Next



## 1 Project Description

Investigating the sites near Tacoma with the highest 98th percentile PM<sub>2.5</sub> from Nov 2021 through Feb 2022.

## 2 Site Information

Sites near Tacoma with high wintertime PM.

## 3 Observations

In the Compare app I noticed a high spike at New Year's. I also noticed that the Woodland/Summit site doesn't pick up the same spikes as other sites - is it at a high elevation or incorrectly sited?

## 4 Summary Statistics

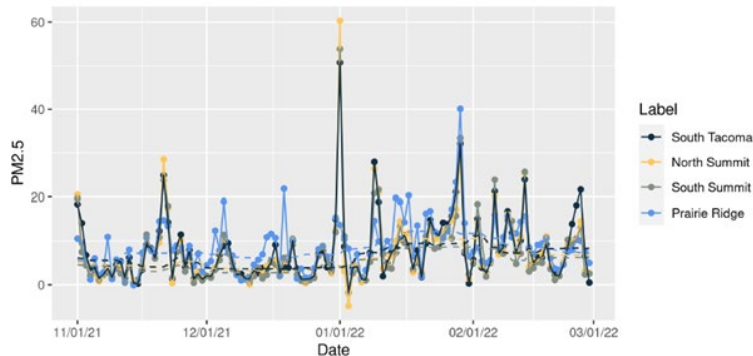
Data were collected from 2021-11-01 to 2022-02-28.

Label	Pollutant	Min	Avg	Max	Trend
North Summit	PM <sub>2.5</sub>	-4.86	7.46	60.32	None
Prairie Ridge	PM <sub>2.5</sub>	-0.08	8.72	40.12	None
South Summit	PM <sub>2.5</sub>	-1.7	7.07	53.81	None
South Tacoma	PM <sub>2.5</sub>	-1.76	7.76	50.7	None

## 5 Figures

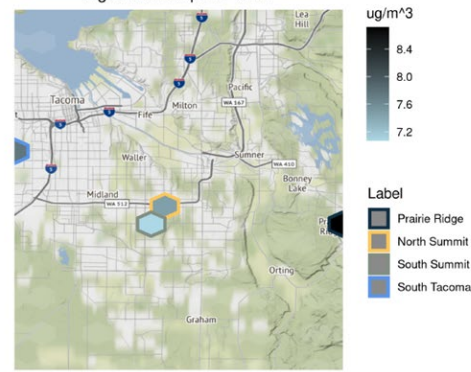
Solid lines are data, dashed lines are a running average of the data.

Figure 1a. Timeseries of PM<sub>2.5</sub>  
2021-11-01 to 2022-02-28



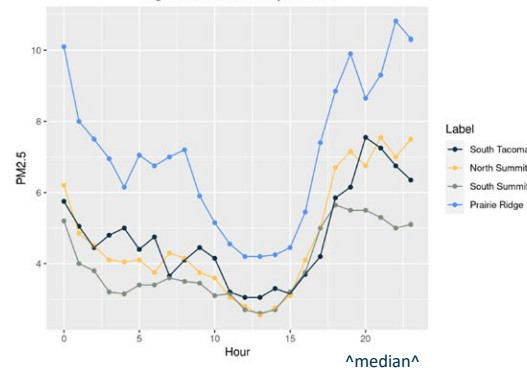
For South Tacoma, PM<sub>2.5</sub> levels spiked on 01/01/22, 01/09/22, and 01/28/22 to 01/29/22. For North Summit, average PM<sub>2.5</sub> levels were high from 01/15/22 to 02/08/22. PM<sub>2.5</sub> levels spiked on 11/21/21, 01/01/22, and 01/29/22. For South Summit, PM<sub>2.5</sub> levels spiked on 01/01/22, 01/29/22, and 02/13/22. For Prairie Ridge, average PM<sub>2.5</sub> levels were high from 01/10/22 to 02/09/22. PM<sub>2.5</sub> levels spiked on 12/19/21, 01/17/22, and 01/28/22 to 01/29/22.

Figure 1b. Map of PM<sub>2.5</sub>



This map shows the average PM<sub>2.5</sub> for each label by location.

Figure 1c. Hour of Day, PM<sub>2.5</sub>



For South Tacoma, PM<sub>2.5</sub> was higher on average from 6pm-11pm. For North Summit, PM<sub>2.5</sub> was higher on average from 6pm-11pm. For South Summit, PM<sub>2.5</sub> was higher on average at 12am, from 6am-9pm, and at 11pm. For Prairie Ridge, PM<sub>2.5</sub> was higher on average at 12am, from 6pm-7pm, and from 9pm-11pm.

## 6 Hypotheses

Hypothesis 1, Prairie Ridge - PM<sub>2.5</sub> is greater than All other labels - PM<sub>2.5</sub>. A t-test shows that Prairie Ridge - PM<sub>2.5</sub> is 1.3 ug/m<sup>3</sup> greater than All other labels - PM<sub>2.5</sub> and this result is not significant. Hypothesis 2, North Summit - PM<sub>2.5</sub> is greater than All other labels - PM<sub>2.5</sub>. A t-test shows that North Summit - PM<sub>2.5</sub> is 0.39 ug/m<sup>3</sup> less than All other labels - PM<sub>2.5</sub> and this result is not significant. Hypothesis 3, South Tacoma - PM<sub>2.5</sub> is highest in the evening (6pm - 12am). Result: South Tacoma - PM<sub>2.5</sub> is highest in the evening (6pm - 12am). Hypothesis 4, North Summit - PM<sub>2.5</sub> is highest in the evening (6pm - 12am). Result: North Summit - PM<sub>2.5</sub> is highest in the evening (6pm - 12am). Hypothesis 5, South Summit - PM<sub>2.5</sub> is highest in the evening (6pm - 12am). Result: South Summit - PM<sub>2.5</sub> is highest in the evening (6pm - 12am).

“Prairie Ridge – PM<sub>2.5</sub> is 1.3 ug/m<sup>3</sup> greater than all other sites”

“North Summit – PM<sub>2.5</sub> is highest in the evening (6pm – 12am)”

# Site location



- Darrington, WA has its own meteorology
- Steep surroundings limit smoke dispersion in winter



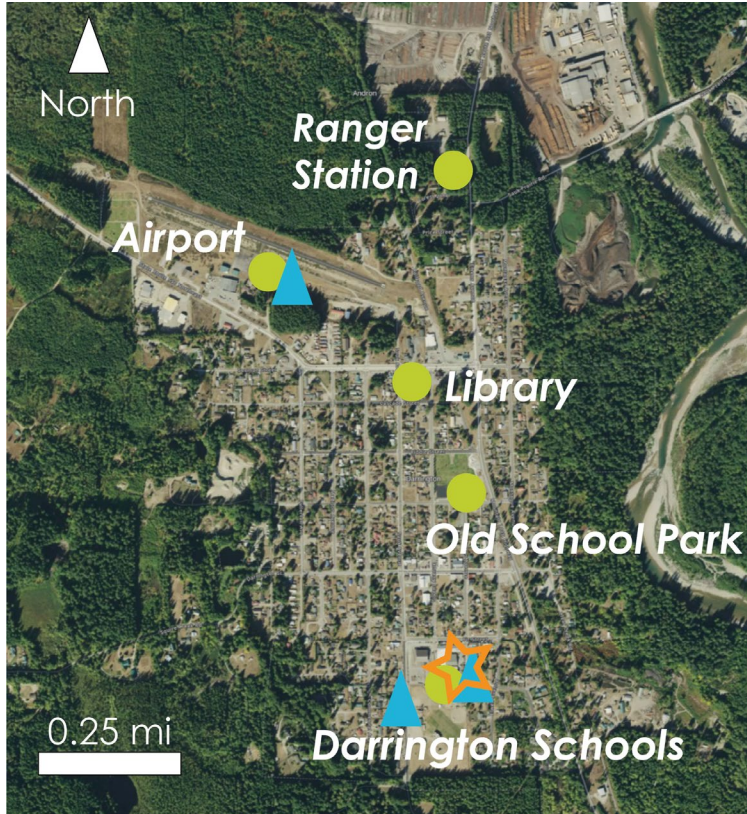


# Instrument locations

Orange Star = Monitoring Station

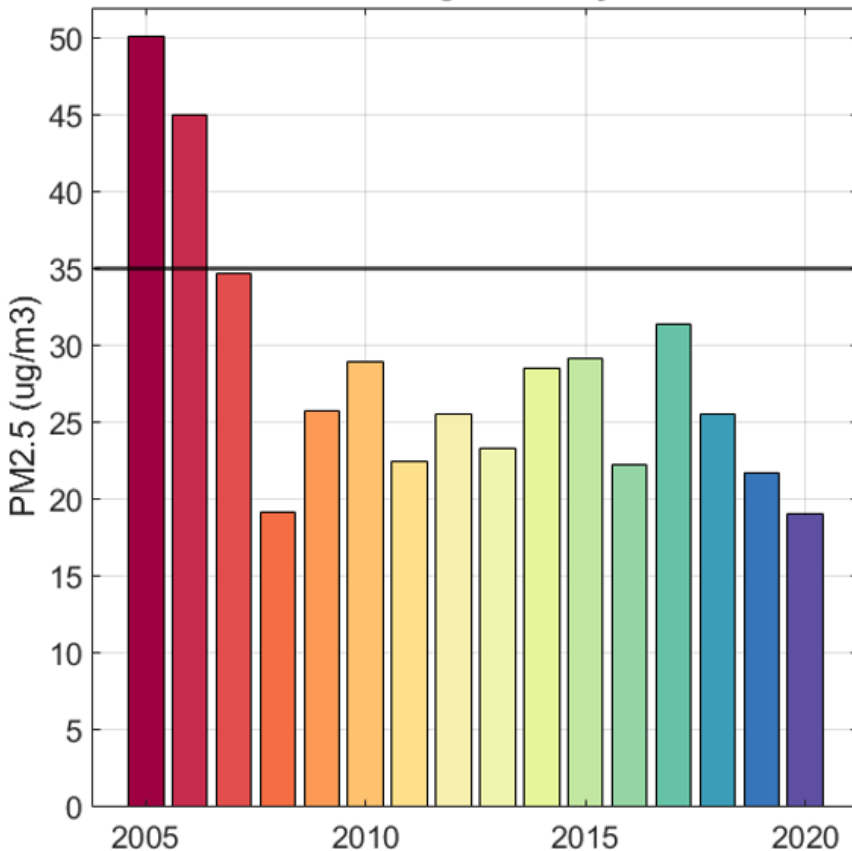
Blue triangles = Wind Sensors

Green circles = Purple Airs



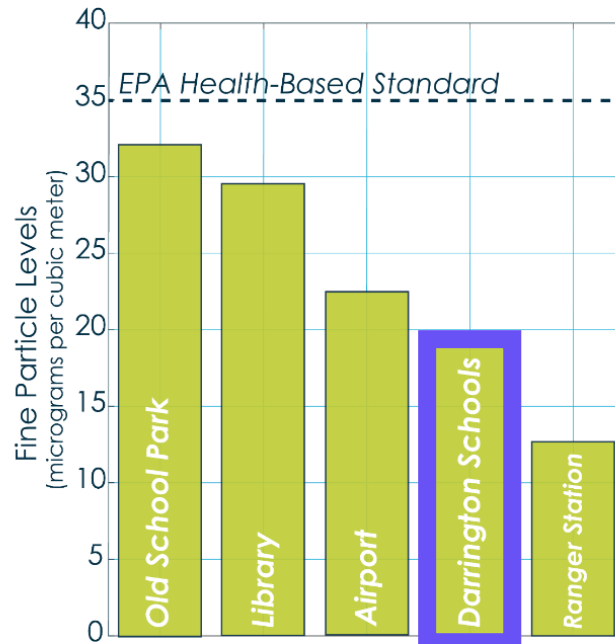


### Winter 8th Highest Daily PM2.5



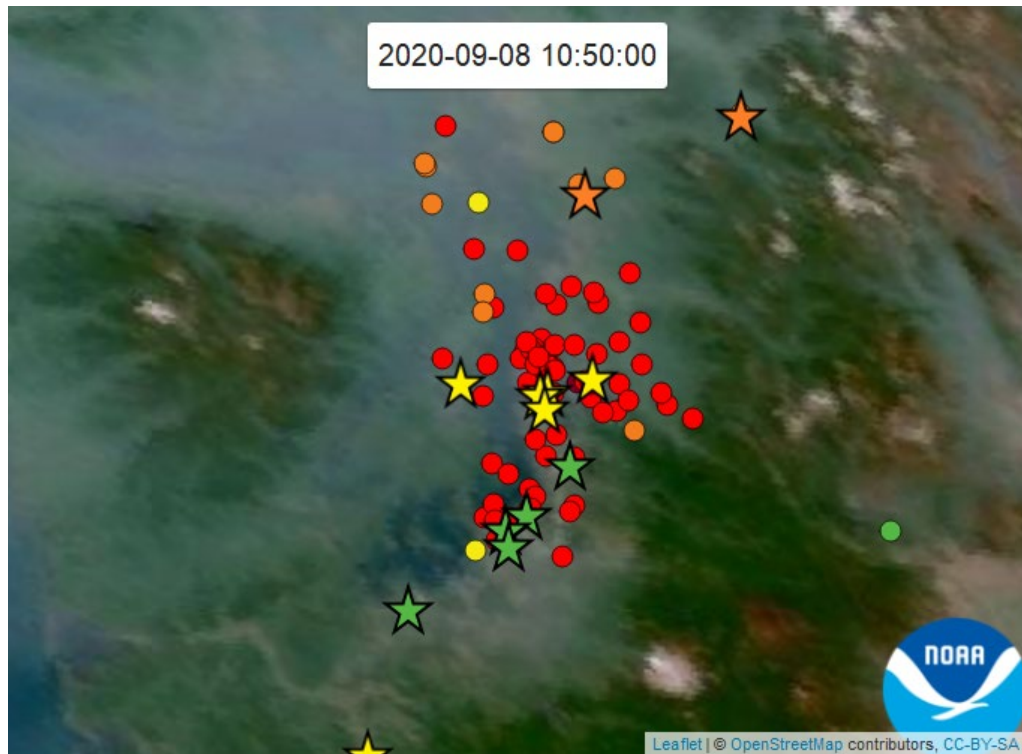
Levels put in context with previous winters

Estimated 8th-highest 24-hr fine particle levels across Darrington, Winter of 2019-20



# Forecasting

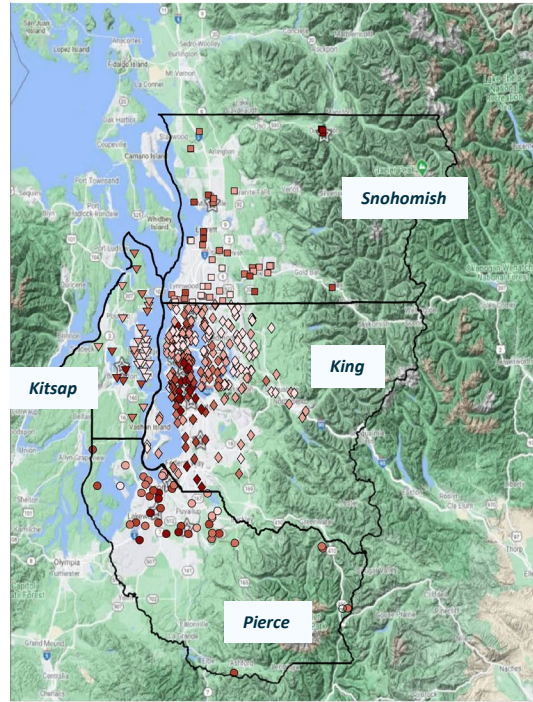
- Minute data helpful during “shorter-term” events
  - Wildfire smoke, wood smoke
- Number of visitors:
  - Sept 2020: 11 consecutive days of Unhealthy or worse AQ – 100k visitors
  - Aug 2021: 2 days of Unhealthy for Sensitive Groups – 19k visitors



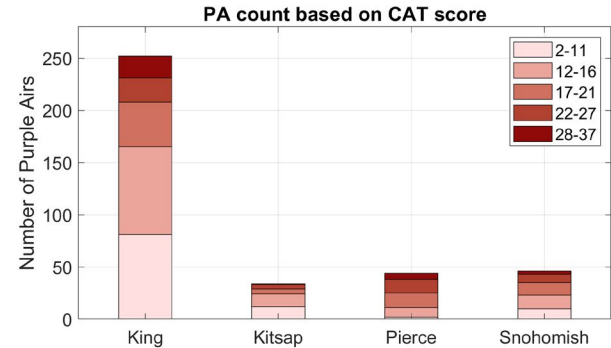
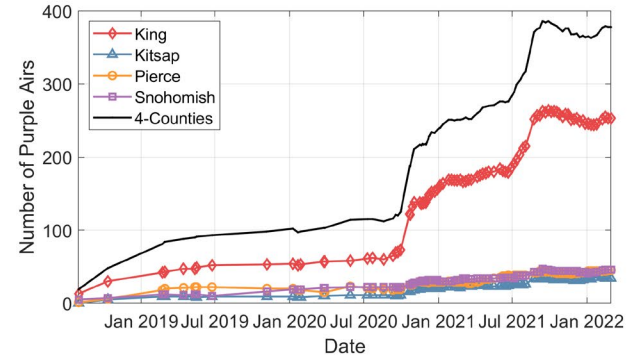
Purple Air sensors pick up wildfire smoke before regulatory sites

# Equity

- More affluent areas end up with more sensors
- Air agencies can step in to ensure more equitable sensor placement



Purple Air Locations for: 01-Mar-2022



# Next steps

1. We need beta testers! (<http://apps.pscleanair.gov/>)
2. Expanded use cases?



# Thank you!

[GraemeC@pscleanair.gov](mailto:GraemeC@pscleanair.gov)