R Package for Air Sensors

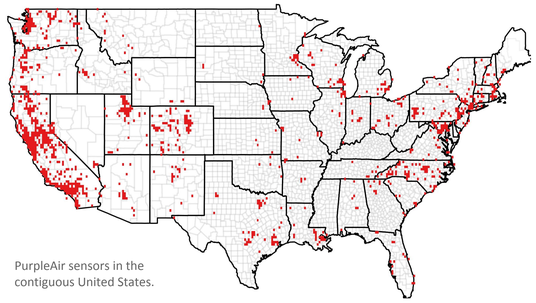
Hans Martin[1], Jonathan Callahan[1], Ashley Monika Collier-Oxandale[2], Brandon Feenstra[2], Vasileios Papapostolou[2], Emma Ranhiem[2]

Goal

Increased public interest in low cost air quality sensors creates a need for software that provides high quality analysis and visualization of sensor data in a manner that is open, transparent and reproducible. The **AirSensor** R package addresses part of this need with functionality focused on data from Purple Air sensors.

Introduction

Low-cost air quality sensors are providing an increasingly dense network of high resolution air quality measurements throughout North America. This is particularly true of laser particle counters measuring PM1.0, PM2.5 and PM10 such as those produced by PurpleAir. Much of the careful analysis of PurpleAir data by academic institutions and government agencies is performed with in-house or proprietary tools that are not freely available to members of the general public. Mazama Science has partnered with the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) at California's South Coast Air Quality Management District to create an open source R package that provides a full suite of data download and analysis capabilities for PurpleAir PM2.5 data.



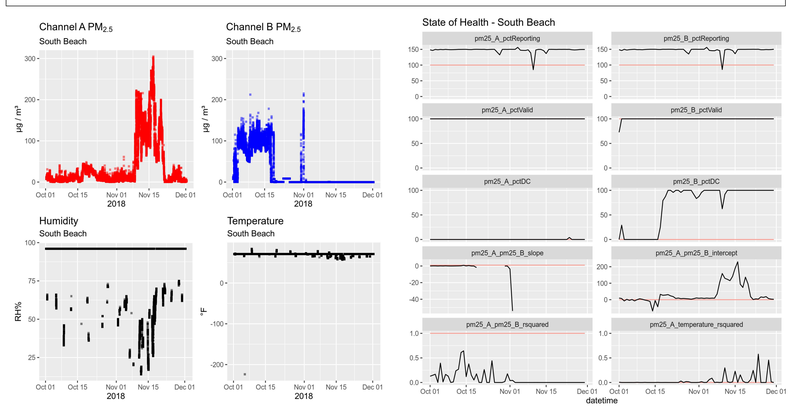
PurpleAir sensors in the contiguous United States.

Data Access

The **AirSensor** R package provides functions to access, regularize and QC data from over 9000 PurpleAir sensors with a simple, readable syntax. With **AirSensor**, analysts can access up-to-the-hour PurpleAir data as well as historical time series.

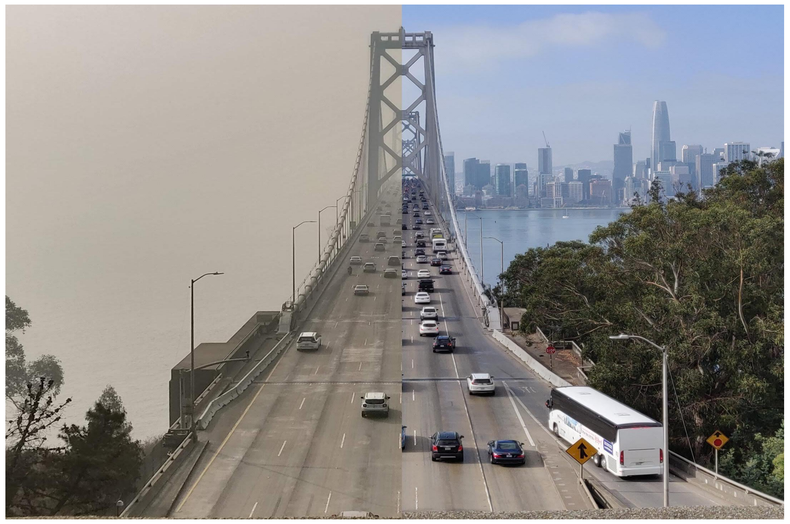
Sensor State-of-Health

An important aspect in the analysis of low-cost sensor data is determining whether or not the device itself is "healthy". The **AirSensor** package provides functions that calculate "State-of-Health" metrics to help assess the day-to-day reliability of data from a PurpleAir sensor. Below is an example of a San Francisco PurpleAir sensor performing poorly and providing inaccurate results during the Camp Fire smoke intrusion in October-November, 2018.

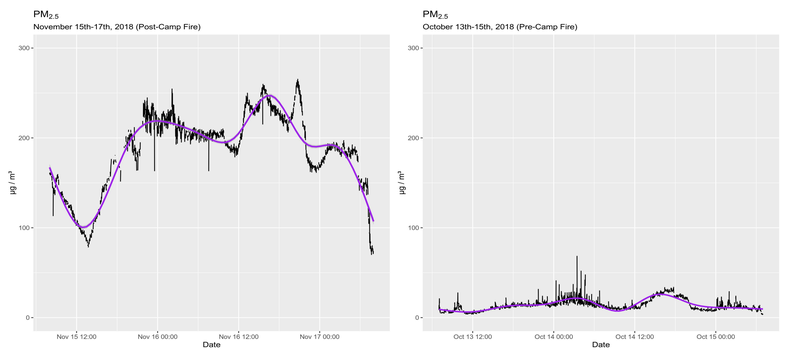


AirSensor R Package
mazamascience.github.io/AirSensor

Mazama Science
mazamascience.com



The Bay Bridge in San Francisco, California. The photo on the left was taken November 16, 2018, and the one on the right October 14, 2018.



Powerful Air Quality Analytics in R

The R statistical programming language and the RStudio IDE provide powerful tools for air quality researchers and analysts. The **AirSensor** R package is part of a suite of open source R packages focused on making high quality analysis and visualization more accessible to anyone working with air quality data. Other R packages of note include: **openair**, **PWFSLSmoke**, **AirMonitorPlots**. Taken together, these packages make it straightforward to thoroughly interrogate air quality data. In the example above, PurpleAir data is used to quantify the particulate levels seen in two images of the Bay Bridge before and during the 2018 Camp Fire smoke episode.

Mazama Science

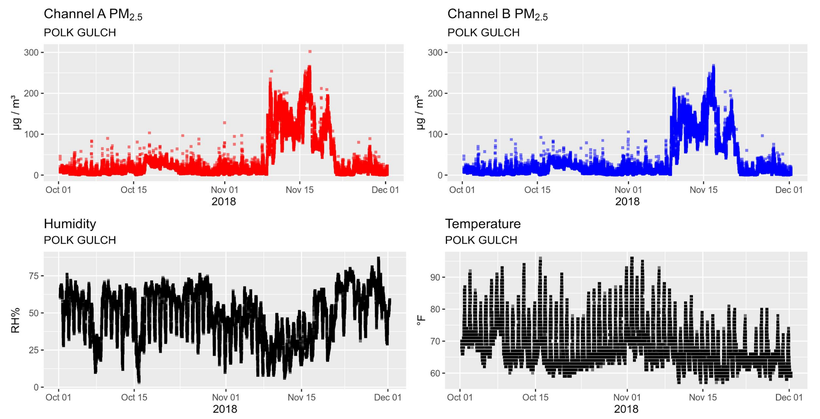
Mazama Science provides modern, open source approaches to data management, analysis and visualization. Our suite of open source R packages is focused on spatial and environmental time series data and analysis with an emphasis on air quality data. For more information see:

smoke.mazamascience.com -- examples of Purple Air data analysis
github.com/MazamaScience -- additional open source R packages

Easy Data Pipeline Syntax

The **AirSensor** R package provides many intuitively named functions and arguments and harnesses R's pipeline syntax so that analysts can write R code that is easy to understand. Several functions provide high level analysis and visualization in a single line of code. The code below demonstrates how straightforward it is to work with data from a PurpleAir sensor in downtown San Francisco. The plot clearly shows the decline in air quality from smoke that drifted in from Butte County's catastrophic Camp Fire in 2018.

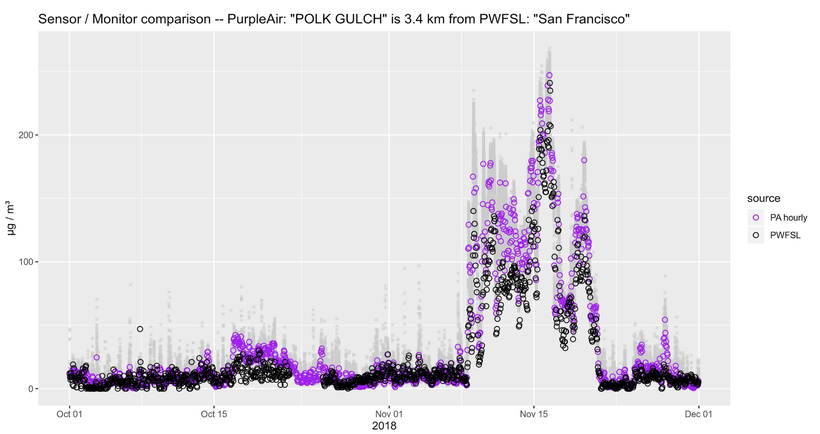
```
pas_load() %>%
  pat_createNew('POLK GULCH', startdate = 20181001, enddate = 20181201) %>%
  pat_multiplot()
```



Sensor/Monitor Comparison

PurpleAir sensors generate data using laser particle counters. With assumed values of scattering coefficient and particle density, particle counts are converted into mass concentration (µg/m^3). Federal reference monitors measure mass concentration more directly and accurately. It is thus important to compare sensor data with federal reference data to validate and scale the sensor data. The **AirSensor** R package provides functions to compare a PurpleAir sensor with the nearest federal reference monitor. The comparison below shows a sensor performing quite well during the 2018 Camp Fire.

```
pas_load() %>%
  pat_createNew('POLK GULCH', startdate = 20181001, enddate = 20181201) %>%
  pat_monitorComparison()
```



*Prepared as part of a project funded through a Science to Achieve Results (STAR) grant award (RD83618401) from the U.S. Environmental Protection Agency to the South Coast Air Quality Management District (South Coast AQMD). The opinions, findings, conclusions, and recommendations are those of the author and do not necessarily represent the views of the U.S. EPA or the South Coast AQMD, nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The U.S. EPA, the South Coast AQMD, their officers, employees, contractors, and subcontractors make no warranty, expressed or implied, and assume no legal liability for the information in this package. The U.S. EPA and South Coast AQMD have not approved or disapproved this package, and neither have passed upon the accuracy or adequacy of the information contained herein.

[1]Mazama Science, Seattle, WA
[2]South Coast Air Quality Management District, Diamond Bar, CA