

Machine learning for spatiotemporal PM_{2.5} estimates across the western US, 2008-2014

Wildfire smoke is often quantified in terms of PM_{2.5} (particulate matter with aerodynamic diameter less than 2.5 μm) concentrations. *In situ* monitoring is the gold standard for measuring PM_{2.5} concentrations. However, the monitoring network is sparse, especially in rural areas of the western US, and often only measures every third or sixth day. Increasingly, researchers are statistically blending information from remotely-sensed Earth observations, atmospheric models, and air quality monitoring data, all of which have strengths and weaknesses, to obtain improved spatiotemporal air pollution exposure surfaces for health studies. Many of such models that have been done for the conterminous US note that they perform worse in the western US. We previously used machine learning to blend several sources of earth observation data, including satellite-derived aerosol optical depth (AOD) data, MODIS Active Fire Products data, NDVI, meteorological reanalysis data, land use, and traffic data to create daily estimates of PM_{2.5} across multiple air basins during the 2008 California wildfires with a cross-validated R² of 0.80. We are improving on these methods by (1) updating our earth observations data to the latest with finer spatial resolution, (2) gathering monitoring data from more geographic locations using not just the US EPA monitoring data but also data from the US Forest Service and other Federal Agencies (via the Fire Cache Smoke Monitor Archive), the IMPROVE Network (via the Federal Land Manager Environmental Database), and previous field campaigns, (3) using an ensemble machine learning method which can improve model performance, and (4) by covering a larger time-space domain to create daily estimates of PM_{2.5} estimates for 11 western states from 2008-2014. These estimates of daily exposure to PM_{2.5} will be used in subsequent epidemiological analyses of the health impacts of air pollution exposures in general, as well as those due to wildfires, a source of air pollution that is increasing in the western US.