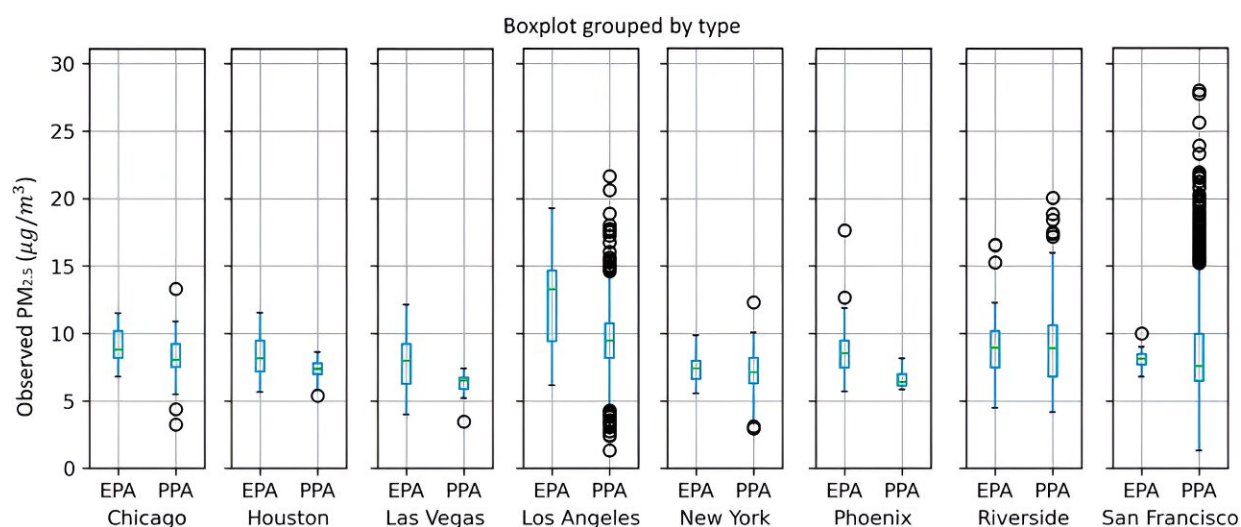


Researchers improve air pollution exposure models using artificial intelligence and mobility data

August 16 2023, by Francisco Tutella



Boxplots of PM_{2.5} concentrations measured by EPA monitors and PurpleAir monitors in eight metropolitan areas. Each box extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a line at the median. The whiskers extend from the box by 1.5x the interquartile range (IQR). Outlier points are those past the end of the whiskers.

Americans in the Northeast paid greater attention to air quality alerts this summer as wildfire smoke thickened skies with an orange-tinted haze. Smoke and other sources of air pollution contain tiny particles, called fine particulate matter (PM 2.5). Smaller than the width of a human hair,

PM 2.5 pose health dangers when inhaled, especially to people with pre-existing heart and lung conditions.

To assess exposure to PM 2.5 and help [public health officials](#) develop strategies, a Penn State-led research team has designed improved models using artificial intelligence and mobility data.

"Our research shows that incorporating artificial intelligence and mobility data into [air quality](#) models can improve the models and help [decision makers](#) and public health officials prioritize areas that need extra monitoring or safety alerts because of unhealthy air quality or a combination of unhealthy air quality and high pedestrian traffic," said Manzhu Yu, assistant professor of geography at Penn State and first author of the study.

As reported in the journal *Frontiers in Environmental Science*, the researchers examined PM 2.5 measurements across eight large metropolitan areas in the continental United States. Air quality data came from Environmental Protection Agency (EPA) monitoring stations and low-cost sensors usually purchased and distributed by local community organizations. They used the data to find hourly PM 2.5 averages in each region.

The scientists input the air quality data into a land use regression model. The model uses local geographical factors like satellite-measured aerosol levels, also called aerosol optical depth; distance to nearest road or stream; elevation; vegetation; and meteorological conditions such as humidity and wind speed to examine how the factors affect air quality.

Past models have taken a linear approach to assessing air pollution, meaning that they assigned a fixed importance to each geographic factor and its impact on air quality, Yu explained. Certain factors like vegetation and meteorological conditions, however, cannot be

represented this way because they change hourly or seasonally and may have [complex interactions](#) with other factors that affect air quality.

Yu and her colleagues took a nonlinear approach to better account for these changing or complex factors by incorporating automated [machine learning](#)—a type of [artificial intelligence](#) that automatically performs time-consuming tasks such as data preparation, parameter selection, and model selection and deployment—into the land use regression model.

The automated machine learning approach used an ensemble method, which allows the machine to run and combine multiple models, to identify the best-performing model for each region. The researchers also examined anonymized cell phone mobility data to pinpoint areas with unhealthy air quality and high visitor numbers.

The researchers found that their automated machine learning method with integrated data from low-cost sensors and EPA monitoring stations improved the accuracy of air pollution exposure models by an average of 17.5%, offering greater spatial variation than using regulatory monitors alone.

Yu credited the improved accuracy to the method's ability to better account for the dynamic variables of aerosol optical depth and meteorological factors, which consistently proved to be the most important across all study regions. The [mobility data](#) component allowed the team to map potential hotspots within regions and times during the day and year when large numbers of people may be exposed to high PM 2.5 levels in these areas.

"Many areas may have consistently high air pollution levels, like those near factories and major transportation hubs, but that is not enough information to make a prioritized list of places needing extra monitoring or health alerts," she said.

"Our mobility-based exposure maps show public health officials and decision makers hotspots that have unhealthy air quality levels plus high visitor traffic. They can use this information to send alerts to people's mobile phones when they enter an area with really high PM 2.5 levels to reduce their exposure to unhealthy air quality."

More information: Manzhu Yu et al, Developing high-resolution PM2.5 exposure models by integrating low-cost sensors, automated machine learning, and big human mobility data, *Frontiers in Environmental Science* (2023). [DOI: 10.3389/fenvs.2023.1223160](https://doi.org/10.3389/fenvs.2023.1223160)

Provided by Pennsylvania State University

Citation: Researchers improve air pollution exposure models using artificial intelligence and mobility data (2023, August 16) retrieved 27 April 2024 from <https://phys.org/news/2023-08-air-pollution-exposure-artificial-intelligence.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.