Ensemble averaging based assessment of spatiotemporal variations in ambient PM2.5 concentrations over Delhi, India, during 2010–2016

Link: https://www.sciencedirect.com/science/article/abs/pii/S1352231020300510?via%3Dihub

What did we do?

Assessing air pollution exposures at high spatiotemporal resolution is critical to understand it's effects on health. However the sparsity of monitoring networks makes it difficult to obtain such estimates in India. Although the number of monitoring stations has grown over time, it is still inadequate in terms of coverage, while existing modeling methods based on land use regressions, chemical transport models and dispersion models each have their limitations. In this paper, we developed a hybrid model to retrospectively assess daily $PM_{2.5}$ concentrations at 1-km x 1-km spatial resolution over Delhi over a seven year period, using multiple sources of data and machine learning based techniques.

What data and methods did we use?

Ambient air pollution is determined by several factors, such as meteorology, land use patterns, and emission sources. In addition, satellite observations provide high resolution imagery of aerosol absorption with a 1 sqkm coverage. We utilized publicly available data from satellites (for aerosol optical depth, light intensity at night and fire emissions), meteorology, vegetation, land use maps, population density and emissions inventories from UrbanEmissions.info to develop these statistical predictive models. The model was built by calibrating against data from ground monitoring stations in Delhi maintained by the Central Pollution Control Board.

We used multiple machine learning based techniques to capture the complex interactions between the above mentioned variables to explain variations in the ground monitoring based $PM_{2.5}$ data. To borrow strength from these techniques, we utilized an ensemble averaging method and cross-validated our predictions using test datasets.

What were our findings?

We obtained daily $PM_{2.5}$ concentrations over Delhi from 2010-2016 and observed that the annual average concentrations ranged from 87 to 138 µg/m^3. There was a clear seasonal pattern with concentrations peaking around October-December each year. However the pollution patterns varied across years with enough spatial variability, indicating some regions in Delhi were more polluted than others. In addition, we found that meteorological variables were crucial in distinguishing most and least polluted grids across all years and seasons, whereas land use patterns were important discriminatory factors in specific seasons such as Fall.

What is the significance of the findings? This study conducted by researchers at Public Health Foundation of India and Centre for Chronic Disease Control in collaboration with Harvard School of Public Health provides a detailed exposure assessment for ambient $PM_{2.5}$ that can be used to study its effects on any health outcome during this period. Given the sparsity of monitoring data, especially pre 2015, this model provides critical coverage both spatially and temporally. In addition, this model can be scaled up to the national level thus opening up vast opportunities to generate evidence on health effects of ambient air pollution in the Indian scenario.

For queries, please reach out to the lead author of the study Dr. Siddhartha Mandal (siddhartha@ccdcindia.org) or Prof. Dorairaj Prabhakaran (dprabhakaran@ccdcindia.org).