# Prediction of daily mean and one-hour maximum $PM_{2.5}$ concentrations and applications in Central Mexico using satellite-based machine-learning models

Iván Gutiérrez-Avila[1], Kodi B. Arfer[1], Daniel Carrión[2,3], Johnathan Rush[1], Itai Kloog[1,4], Aaron R. Naeger[5], Michel Grutter[6], Víctor Hugo Páramo-Figueroa[7], Horacio Riojas-Rodríguez[8], Allan C. Just[1]

[1]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, United States
[2]Department of Environmental Health Sciences, Yale University School of Public Health, New Haven, CT, USA
[3]Center on Climate Change and Health, Yale University School of Public Health, New Haven, CT, USA
[4]Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer Sheva, Israel
[5]Earth System Science Center, University of Alabama in Huntsville, Huntsville, AL, USA
[6]Instituto de Ciencias de la Atmósfera y Cambio Climático, Universidad Nacional Autónoma de México, Ciudad de México, México
[7]Comisión Ambiental de la Megalópolis, Ciudad de México, México
[8]Dirección de Salud Ambiental, Instituto Nacional de Salud Pública, Cuernavaca Morelos, México

## Abstract

Machine-learning algorithms are becoming popular techniques to predict ambient air $PM_{2.5}$ concentrations at high spatial resolutions (1×1 km) using satellite-based aerosol optical depth (AOD). Most machine-learning models have aimed to predict 24h-averaged $PM_{2.5}$ concentrations (mean $PM_{2.5}$). Over Mexico, none has been developed to predict subdaily peak levels, such as the maximum daily one-hour concentration (max $PM_{2.5}$). We present a new modeling approach based on extreme gradient boosting (XGBoost) and inverse-distance weighting that uses AOD data, meteorology, and land-use variables to predict mean and max $PM_{2.5}$ in Central Mexico (including the Mexico City Metropolitan Area) from 2004 through 2019. Our models for mean and max $PM_{2.5}$ exhibited good performance, with overall cross-validated mean absolute errors (MAE) of 3.68 and 9.21 $\mu g/m^3$, respectively, compared to mean absolute deviations from the median (MAD) of 8.55 and 15.64 $\mu g/m^3$. We also investigated applications of our mean $PM_{2.5}$ predictions that can aid local authorities in air-quality management and public-health surveillance, such as the co-occurrence of high $PM_{2.5}$ and heat, compliance with local air-quality standards, and the relationship of $PM_{2.5}$ exposure with social marginalization.

**Correspondence**
Iván Gutiérrez-Avila, Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1057, New York, NY 10029, USA. Email: ivan_2c@hotmail.com

## 1. Background

Fine particulate matter with aerodynamic diameter ≤ 2.5 microns ($PM_{2.5}$) affects more people than any other pollutant, and has been consistently associated with mortality and morbidity from cardiovascular and respiratory causes.[1] Over the last decade, epidemiological evidence has related $PM_{2.5}$ to many other health outcomes, such as cardio-metabolic diseases (including diabetes, hypertension, metabolic syndrome), neurological disorders (stroke, dementia, Alzheimer's disease, autism, Parkinson's disease), and perinatal outcomes (premature birth and low birth weight).[2–4] At the same time, exposure scientists have developed new modeling approaches for air-pollution epidemiology, moving away from the use of data from ground monitors alone. Interest has grown in models using remote-sensing products, particularly aerosol optical depth (AOD) for the prediction of ground level $PM_{2.5}$ concentrations at high spatial resolutions, such as 1 x 1 km.[5–12] AOD is a measure of the amount of light absorbed and scattered throughout the atmospheric vertical column by the collection of suspended particles (e.g., urban haze, smoke, desert dust, sea salt) in the atmosphere. Thus, it is related to $PM_{2.5}$ concentrations most relevant to health as measured by ground monitors, but the relationship is complex and depends on a number of other factors.[13] Popular approaches to predicting ground-level $PM_{2.5}$ concentration using AOD include chemical-transport models, as well as geostatistical approaches such as mixed-effect models, geographically weighted regression, and land-use regression, which use additional $PM_{2.5}$ predictors and modifiers of the $PM_{2.5}$–AOD relation such as weather and land use.[13,14] Among the most comprehensive efforts to reconstruct ground concentrations of $PM_{2.5}$ is NASA's Global Modeling Initiative (GMI) chemistry transport model integrated with Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2 GMI), which estimates the global distribution of $PM_{2.5}$ mass concentrations with a spatial resolution of 0.5° × 0.625° (about 50 km in the latitudinal direction), and temporal resolution as fine as 1 hour (https://acd-ext.gsfc.nasa.gov/Projects/GEOSCCM/MERRA2GMI/).[15]

Predicting ground-based $PM_{2.5}$ from satellite AOD retrievals is difficult. AOD is strongly influenced by particles above the surface layer, which have different characteristics from ground-level particles. Also, AOD retrieval algorithms assume consistent particle size distributions within large regions, such as Mexico and Central America.[16] Furthermore, AOD often has gaps in spatial coverage due to clouds, snow, or ice. Thus, researchers must often impute missing AOD,[17] and the complex relationship between AOD and $PM_{2.5}$, along with the use of additional $PM_{2.5}$ predictors, has motivated machine-learning approaches such as neural networks, random forests, and gradient boosting.[14,18–26]

These recently developed AOD-based $PM_{2.5}$ (AOD-$PM_{2.5}$) models and predictions have allowed epidemiologists to move away from traditional exposure-assessment methods that rely on proximity to sparse ground monitors. With sufficient spatiotemporal resolution, AOD-$PM_{2.5}$ models may further improve exposure assessment in epidemiologic research by picking up the effects of

microenvironments. Few AOD-$PM_{2.5}$ models exist for middle-income countries. Our group developed one of the first AOD-$PM_{2.5}$ models using daily Multi-Angle Implementation of Atmospheric Correction (MAIAC) spectral AOD derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on NASA's Aqua satellite at a 1x1 km spatial resolution, along with data from ground monitors, land use, and meteorological features.[27,28] Our previous model for the Mexico City region provides daily $PM_{2.5}$ predictions from 2004–2014, and those predictions have been used in several epidemiologic studies in this region.[9,29–46] However, model improvements are needed to better characterize the spatiotemporal distribution of $PM_{2.5}$,[47,48] particularly since the Mexico City Metropolitan Area, like other middle-income regions, has undergone considerable urban sprawl. $PM_{2.5}$ emitted and produced in large metropolitan areas affects not only people in the city center but also people in its suburban and rural outskirts. People in the outskirts, where air-quality information is scarce, may face disproportionate health risks due to lower socioeconomic status and less access to healthcare. This environmental injustice can be even more pronounced in low- and middle-income regions.[49,50]

AOD-$PM_{2.5}$ models covering large urban areas have great value for epidemiology, but also for public-health surveillance (as in quantifying mortality and morbidity attributable to $PM_{2.5}$),[51] environmental regulation (as in assessment of compliance with air quality standards),[52] and risk communication (as in designing air-quality indices).[53] Furthermore, AOD-$PM_{2.5}$ models can help administrators in air-quality management to see trends in the spatiotemporal distribution of $PM_{2.5}$, map hotspots in regions with few monitors, and identify emissions sources to consider for abatement actions (e.g. prediction and surveillance of air pollution contingencies).[54,55] Overall, AOD-$PM_{2.5}$ models can be powerful aids for decision-making.

Most of the satellite-based $PM_{2.5}$ models yield predictions of 24-hour mean concentrations, perhaps driven by traditional approaches in epidemiology that have focused on this exposure metric, which in turn support standards for daily $PM_{2.5}$ levels. There is growing interest in identification of specific sub-daily $PM_{2.5}$ exposures (e.g., peak concentrations) that may trigger the onset of adverse health outcomes and harm vulnerable people, compared to the use of 24-hour averages alone. To our knowledge, this is the first model reconstructing sub-daily $PM_{2.5}$ concentrations in Mexico.

In this study, we present a new model based on extreme gradient boosting (XGBoost) and inverse-distance weighting (IDW) that uses satellite and land-use variables to predict daily mean and max $PM_{2.5}$ concentrations in Central Mexico. We also present some example applications of our model and $PM_{2.5}$ predictions.

## 2. Method

We constructed and evaluated two models: one for daily mean ambient $PM_{2.5}$, spanning 2004 through 2019, and one for daily max ambient $PM_{2.5}$ (more precisely, the greatest hourly concentration of $PM_{2.5}$ observed each day), spanning 2011 through 2019. We restricted our max $PM_{2.5}$ predictions to 2011 onwards because of greater coverage of ground monitoring stations. Days were defined according to UTC−6, which coincides with the local time of the study region (Mexico's Zona Centro) when daylight-saving time is not in effect (namely, before the first Sunday of April and after the last Sunday of October).

### 2.1. Study region

We modeled an irregularly shaped area of 6,650 $km^2$, 127 km in diameter, around Mexico City. The model used a grid of 7,745 square cells, 927 m on a side, in a global sinusoidal projection (the same one used for NASA's MODIS products). This study area and its grid was a subset of that considered in our ambient temperature model for Central Mexico.[56] We built the subset by finding the largest connected set of cells in the Valley of Mexico with all cells ≤ 3 km above sea level (Figure 1). The Valley of Mexico is a plateau with a mean elevation of 2,250 m above sea level, and is surrounded on three sides by mountain ranges, preventing the dispersion of air pollutants.[57]

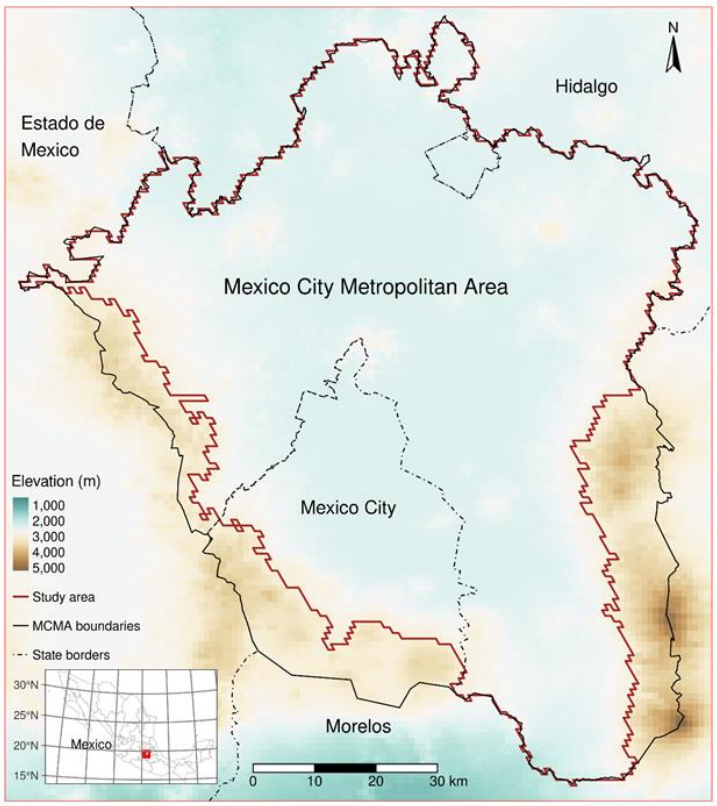


Figure 1. The study area used for our $PM_{2.5}$ models in the Mexico City Metropolitan Area (MCMA)

### 2.2. Data

We used $PM_{2.5}$ data from ground monitoring stations organized by the Instituto Nacional de Ecología y Cambio Climático de México (INECC) including records from the Automated Atmospheric Monitoring Network (RAMA, Spanish acronym) from the Mexico City's Atmospheric Monitoring System (SIMAT, website http://www.aire.cdmx.gob.mx/). We downloaded observations from INECC's website (http://scica.inecc.gob.mx). We discarded one extreme one-hour value that was greater than 900 $\mu g/m^3$ because it appeared to be implausible when compared with the previous and subsequent one-hour concentrations recorded at the same monitoring station. For each station in the study area and day of $PM_{2.5}$ observations, we computed the mean and max $PM_{2.5}$ among the hourly observations, so long as there were at least 18 hours of observations in the day. Other station-days were discarded. The result was a total of 60,365 station-days from 25 stations for mean $PM_{2.5}$ and 40,819 station-days from the same 25 stations for max $PM_{2.5}$.

Our models used the following 14 predictors:

* Longitude and latitude in degrees
* The date, as an integer count of days
* The IDW mean (exponent 2) of all observations of the same dependent variable on the given day
* MAIAC AOD from NASA's Terra and Aqua satellites,[58] whose local overpass times range from 10:40 to 15:15 and 13:10 to 15:05, respectively
* $PM_{2.5}$ ($\mu g/m^3$) as predicted by MERRA-2 GMI (https://acd-ext.gsfc.nasa.gov/Projects/GEOSCCM/MERRA2GMI/), either the mean of the day's 24 hourly values (for modeling mean $PM_{2.5}$) or the value at 10:00 UTC−6 (for max $PM_{2.5}$)
* Temperature (K), precipitation (mm), and vapor pressure (Pa) from Daymet,[59] the temperature being computed as the mean of Daymet's maximum and minimum temperature
* The height of the planetary boundary layer (m) and meridional and zonal wind speeds (m/s) from Climate Data Store,[60] using the mean of the day's 24 hourly values (for mean $PM_{2.5}$) or the value at 10:00 UTC−6 (for max $PM_{2.5}$)
* The density of roads ($m/km^2$) from OpenStreetMap,[61] considering only primary, secondary, residential, and tertiary roads

We selected the midmorning time of day 10:00 UTC−6 in constructing some of the predictors for the max $PM_{2.5}$ model because it was the most frequent hour of greatest daily per-station $PM_{2.5}$ concentration in our sample.

### 2.3. Model evaluation

We evaluated models with leave-one-station-out cross-validation (CV). There are 25 stations, so for each dependent variable, we fit the model 25 times, leaving out one station from training and then testing the model's predictions on the left-out station. We evaluated models with absolute loss rather than squared loss so as not to overweight the importance of a minority of very high observed concentrations of $PM_{2.5}$. Absolute loss leads to mean absolute error (MAE) as a natural measure of predictive accuracy (in place of root mean square error, RMSE, for squared loss), and mean absolute deviation from the median (MAD) as a measure of baseline variation in place of the standard deviation (SD) for squared loss. Note that $R^2$, which is often used for model assessment, is defined as a squared-loss metric. For our study, we compute $R^2$ as 1 minus the MSE divided by the variance, and we show $R^2$, RMSE, and SD in tables for completeness, although the models are more properly judged in terms of absolute loss.

When computing the IDW predictor during CV, we excluded the held-out station to avoid data leakage.

### 2.4. Models

We predicted $PM_{2.5}$ with XGBoost,[62] a scheme for fast boosted decision trees. We used a log-cosh objective function to approximate absolute loss. Instead of providing $PM_{2.5}$ as the dependent variable to XGBoost directly, we provided $PM_{2.5}$ minus the IDW interpolation and added the IDW back to XGBoost's predictions. This method partly smooths out the otherwise discrete predictions produced by decision trees. We tuned XGBoost with twofold station-wise CV; during the larger CV discussed above, this CV was nested within each fold. Tuning adjusted four hyperparameters:

* The number of trees, which could be 10, 25, 50, or 100
* The maximum tree depth, which could be 3, 6, or 9
* The learning rate $\eta$, which could range from 0.01 to 0.5
* A ridge penalty $\lambda$, which could range from $2^{-10}$ to $2^{10}$

We preselected a set of 25 random vectors from this space with a maximin Latin-hypercube sample.

Once the outer CV was done, to make new predictions, we trained the two models (one for mean $PM_{2.5}$ and one for max $PM_{2.5}$) on all the data, with one more tuning CV apiece. These final models had the following hyperparameters: for mean $PM_{2.5}$, 10 trees, max depth 3, $\eta = 0.047$, $\lambda = 10$; for max $PM_{2.5}$, 25 trees, max depth 9, $\eta = 0.073$, $\lambda = 260$.

### 2.5. Applications

We present three applications of our $PM_{2.5}$ predictions for the Mexico City Metropolitan Area. We examined co-occurring exposures to high $PM_{2.5}$ concentrations and high temperatures from our published spatiotemporal model.[56] Person-time estimates of exposure relied on population density estimates for 2010. We estimated the population density within each of our grid cells using the R package exactextractr[63] to calculate the area-weighted mean of the population density in the intersecting Gridded Population of the World (GPWv4) ~1-km raster cells.[64] The GPWv4 used data from the 2010 census in Mexico at the level of Área Geoestadistica Básicas (AGEBs; the Mexican equivalent of US Census tracts). When comparing exposures to permissible annual limits, we computed "yearly" means as the means of four 3-month means, per the Mexican standard.[65] Finally, we examined how AGEB-level $PM_{2.5}$ exposure varied with social marginalization within the study region.[66] Every AGEB was assigned the mean $PM_{2.5}$ prediction of all 1x1 km grid cells whose centroids fell within the AGEB polygon.

## 3. Results

Overall, the observed $PM_{2.5}$ that we trained and tested on had a median of 23 $\mu g/m^3$ (MAD = 8.55, IQR = 14.08) for mean $PM_{2.5}$, and a median of 44 $\mu g/m^3$ (MAD = 15.64, IQR = 25.00) for max $PM_{2.5}$.

### 3.1. Cross-validation

The model for mean $PM_{2.5}$ achieved a MAE of 3.68 $\mu g/m^3$ (compared to a MAD of 8.55 $\mu g/m^3$), and the model for max $PM_{2.5}$ achieved a MAE of 9.21 $\mu g/m^3$ (compared to a MAD of 15.64 $\mu g/m^3$). These differences indicate a substantial improvement in accuracy compared to assigning the median exposure to all places and times throughout the study domain. The much greater MAE for max $PM_{2.5}$ than mean $PM_{2.5}$ is to be expected, because maxima are inherently more difficult to predict than means. Tables 1 and 2 show the performance of these models broken down by year.

Table 1. Assessment of cross-validated predictions from the daily mean $PM_{2.5}$ model by year.

| Year | Number of stations | Observations | $R^2$ | SD | RMSE | MAD | MAE |
|---|---|---|---|---|---|---|---|
| 2004 | 8 | 2,751 | 0.76 | 12.02 | 5.86 | 9.12 | 3.91 |
| 2005 | 8 | 2,701 | 0.81 | 14.80 | 6.43 | 11.28 | 4.38 |
| 2006 | 8 | 2,685 | 0.68 | 13.19 | 7.48 | 9.55 | 5.04 |
| 2007 | 9 | 2,855 | 0.71 | 10.87 | 5.85 | 8.16 | 4.28 |
| 2008 | 9 | 3,040 | 0.64 | 12.16 | 7.29 | 9.27 | 4.61 |
| 2009 | 9 | 2,670 | 0.75 | 10.14 | 5.09 | 7.71 | 3.61 |
| 2010 | 9 | 2,844 | 0.79 | 11.70 | 5.41 | 8.83 | 3.64 |
| 2011 | 12 | 3,019 | 0.77 | 11.53 | 5.56 | 8.90 | 3.88 |
| 2012 | 13 | 4,025 | 0.76 | 10.10 | 4.95 | 7.63 | 3.59 |
| 2013 | 13 | 4,362 | 0.80 | 11.75 | 5.25 | 8.85 | 3.87 |
| 2014 | 14 | 4,203 | 0.73 | 9.87 | 5.10 | 7.50 | 3.86 |
| 2015 | 19 | 5,194 | 0.77 | 10.78 | 5.11 | 7.90 | 3.76 |
| 2016 | 17 | 5,307 | 0.83 | 11.44 | 4.73 | 8.56 | 3.37 |
| 2017 | 17 | 4,901 | 0.80 | 10.79 | 4.78 | 8.42 | 3.15 |
| 2018 | 17 | 4,633 | 0.84 | 9.91 | 3.94 | 7.19 | 2.83 |
| 2019 | 20 | 5,175 | 0.86 | 11.50 | 4.26 | 7.98 | 2.85 |

Standard deviation (SD), Root mean squared error (RMSE), Mean absolute deviation (MAD), and Mean Absolute Error (MAE)

Table 2. Assessment of cross-validated predictions from the daily one-hour maximum $PM_{2.5}$ model by year

| Year | Number of stations | Observations | $R^2$ | SD | RMSE | MAD | MAE |
|---|---|---|---|---|---|---|---|
| 2011 | 12 | 3,019 | 0.47 | 24.26 | 17.65 | 16.65 | 10.38 |
| 2012 | 13 | 4,025 | 0.45 | 21.80 | 16.11 | 15.18 | 10.19 |
| 2013 | 13 | 4,362 | 0.57 | 23.78 | 15.52 | 17.28 | 10.29 |
| 2014 | 14 | 4,203 | 0.52 | 19.54 | 13.57 | 14.46 | 9.77 |
| 2015 | 19 | 5,194 | 0.63 | 25.30 | 15.33 | 16.37 | 9.97 |
| 2016 | 17 | 5,307 | 0.62 | 25.33 | 15.59 | 16.48 | 8.69 |
| 2017 | 17 | 4,901 | 0.56 | 23.86 | 15.75 | 15.85 | 8.49 |
| 2018 | 17 | 4,633 | 0.63 | 19.68 | 11.96 | 13.74 | 7.84 |
| 2019 | 20 | 5,175 | 0.66 | 21.39 | 12.49 | 14.08 | 8.04 |

Standard deviation (SD), Root mean squared error (RMSE), Mean absolute deviation (MAD), and Mean Absolute Error (MAE)

We also compared model performance by season: cold dry (spanning November through February), warm dry (March to May), and rainy (June to October).[67] Table 3 shows that the largest improvement in prediction accuracy (MAD minus MAE) was observed during the cold dry season for both mean and max $PM_{2.5}$ models, although this season still had the largest MAE.

Table 3. Assessment of cross-validated predictions for the mean and max $PM_{2.5}$ models by season from 2004-2019

| $PM_{2.5}$ model | Season | Number of stations | Observations | $R^2$ | SD | RMSE | MAD | MAE |
|---|---|---|---|---|---|---|---|---|
| **Mean $PM_{2.5}$** | Cold-Dry | 25 | 20,135 | 0.75 | 12.97 | 6.42 | 9.50 | 4.42 |
| | Warm-Dry | 25 | 10,294 | 0.66 | 9.94 | 5.79 | 7.62 | 3.93 |
| | Rainy | 25 | 29,936 | 0.80 | 9.54 | 4.27 | 7.13 | 3.10 |
| **Max $PM_{2.5}$** | Cold-Dry | 25 | 13,669 | 0.58 | 28.46 | 18.44 | 18.31 | 11.03 |
| | Warm-Dry | 24 | 7,048 | 0.38 | 20.07 | 15.76 | 14.07 | 9.50 |
| | Rainy | 24 | 20,102 | 0.60 | 17.96 | 11.39 | 13.47 | 7.86 |

Standard deviation (SD), Root mean squared error (RMSE), Mean absolute deviation (MAD), and Mean Absolute Error (MAE)

Table 4 shows the Pearson correlations among observed and predicted $PM_{2.5}$ for both models. As would be expected, all four variables are positively related. Predictions are more associated with the kind of observation they are meant to predict than the other kind, but there are also strong correlations between mean and max $PM_{2.5}$.

Table 4. Correlation coefficients between observed and predicted mean and max $PM_{2.5}$ concentrations

| $PM_{2.5}$ | Mean observed | Maximum observed | Mean predicted | Maximum predicted |
|---|---|---|---|---|
| Mean observed | 1.00 | | | |
| Maximum observed | 0.85 | 1.00 | | |
| Mean predicted | 0.89 | 0.73 | 1.00 | |
| Maximum predicted | 0.82 | 0.77 | 0.90 | 1.00 |

### 3.2. Diagnostics of new predictions

After making predictions for every grid cell and day with both models, we mapped the per-cell mean $PM_{2.5}$ and max $PM_{2.5}$ averaged over 2019 (Figure 2). Discontinuities in the prediction surfaces evident in our maps are the result of model-based splits selected in the longitude and latitude predictors. Although we also include an IDW interpolation that adds some smoothness, XGBoost selects for the most predictively accurate model. Smoothing our predictions more aggressively could make for more realistic-looking maps, but would not necessarily improve predictive accuracy. As expected, the highest concentrations (shown in dark purple) are in the center-north and center-east subregions of the Mexico City Metropolitan Area (north and east of Mexico City, respectively), with the highest population density and industrial land use. This pattern is also visible in the max $PM_{2.5}$ map, but is most pronounced in the center-north. The lowest $PM_{2.5}$ concentrations (shown in light purple and yellow) are in the southwest, corresponding to the least populated and the most vegetated subregion.

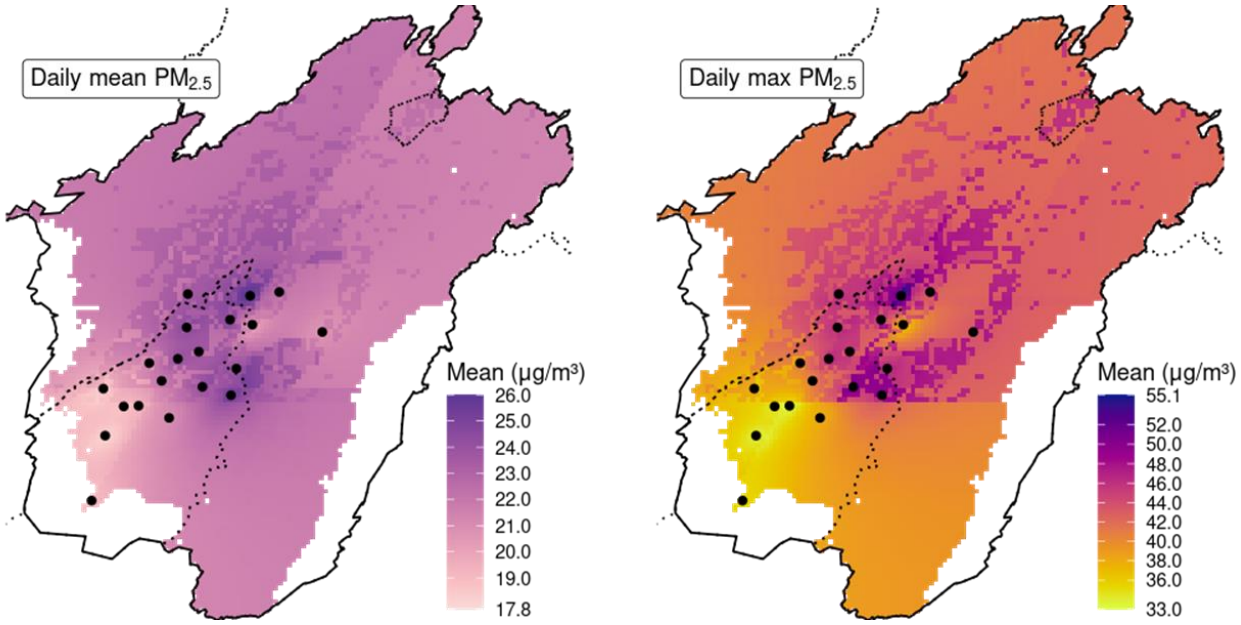


Figure 2. Maps of the averaged annual daily mean and daily max $PM_{2.5}$ concentrations for 2019 in the Mexico City Metropolitan Area. Solid and dotted lines indicate the Mexico City Metropolitan Area and Mexican states boundaries, respectively. Black dots indicate ground monitors.

We examined the per-day ratio (collapsing across all cells) of mean and max $PM_{2.5}$. Figure 3 shows this ratio for each day in 2019. Generally, the max is about twice the mean, but the ratio decreases in the first half of the year and increases in the second. During the rainy season (June to October), we examined how the ratio differed between days with and without a mean per-cell precipitation of at least 1 mm, and found little difference: the mean ratio was 2.03 on dry days and 2.13 on rainy days.

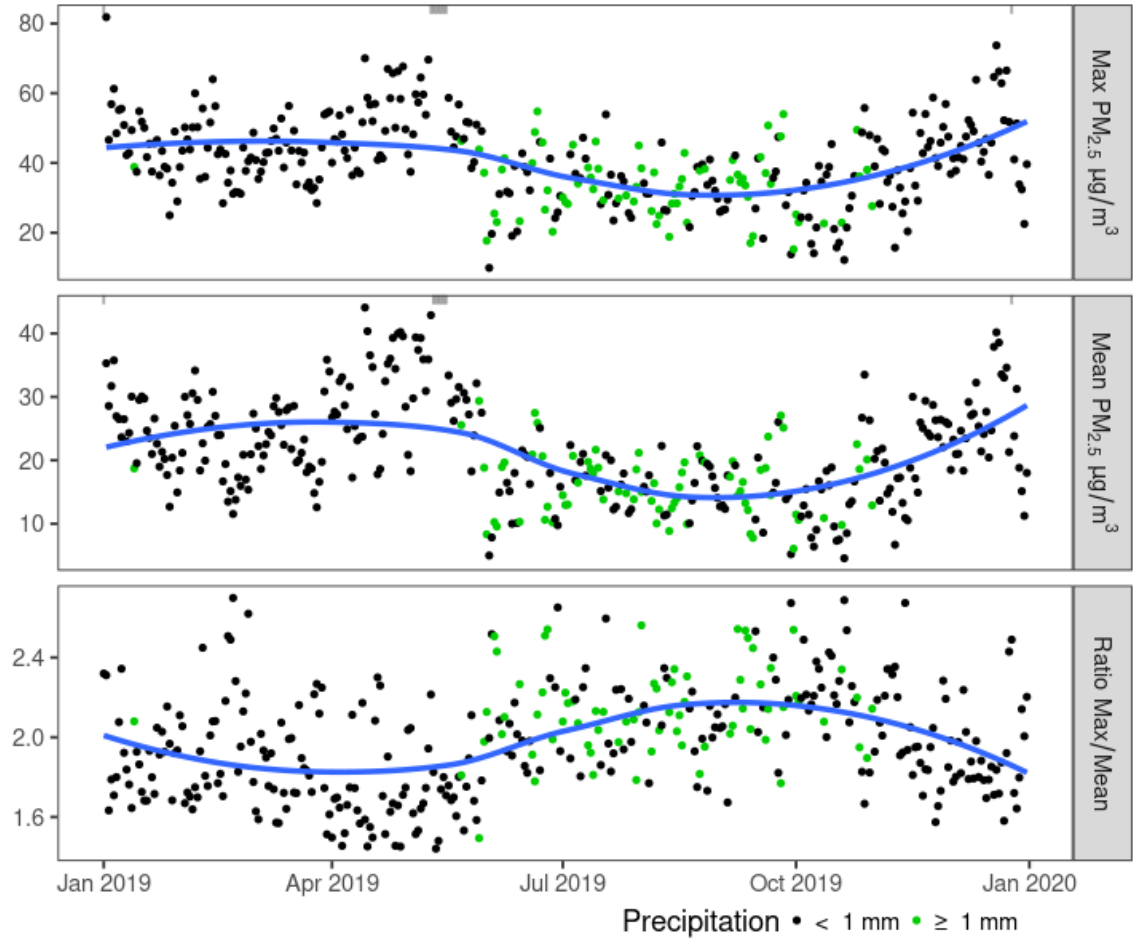


Figure 3. For each day in 2019, the mean of three predicted quantities (max $PM_{2.5}$, mean $PM_{2.5}$, and max divided by mean) across all cells. A locally estimated scatterplot smoothing (LOESS) trendline is shown for each panel. Days are colored according to the mean precipitation across all cells. For legibility, 9 especially high points (> 90 µg/m$^3$) are excluded in the max panel and 8 especially high points (> 45 µg/m$^3$) are excluded in the mean panel (indicated by ticks on the top border); the corresponding ratios are still included in the bottom panel.

With our temperature model,[56] we examined the relationship between mean daily $PM_{2.5}$ and mean daily temperature. The Kendall correlation between the two over the whole study period was 0.05, indicating a very weak positive relationship overall. Figure 4 breaks this relationship down by season. It can be observed that the $PM_{2.5}$ concentrations are more stable and remain high during the cold dry season, which has been related to the stable atmospheric conditions and frequent thermal inversions in the study region. For the warm dry season and rainy season, there is a clearer tendency for higher $PM_{2.5}$ concentrations on hotter days.

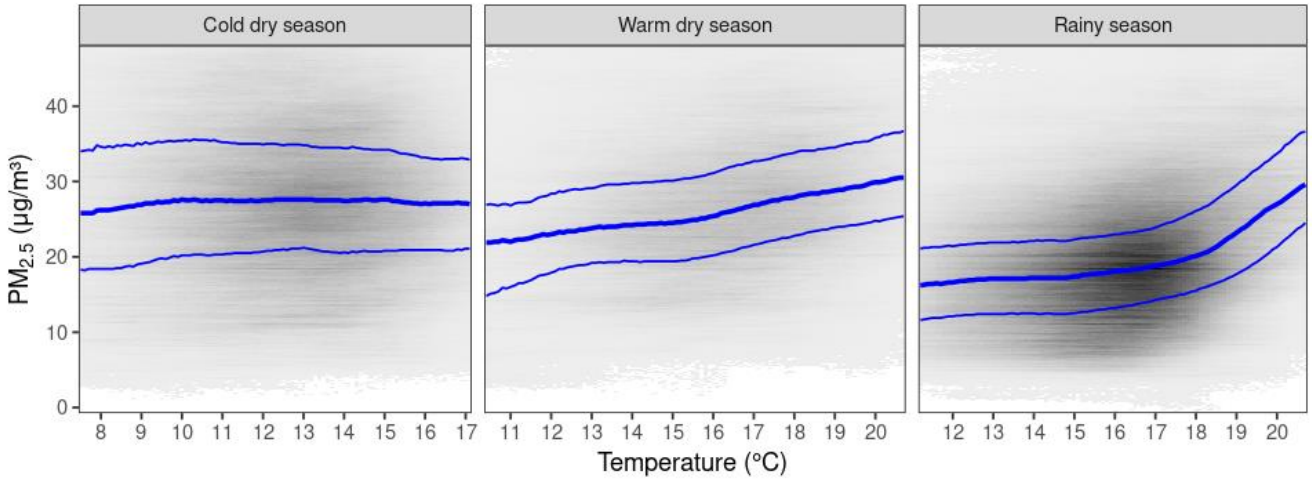


Figure 4. Heatmaps of mean temperature and mean $PM_{2.5}$, counting all grid cells and days equally. Darker areas indicate more grid cells, more days, or both. Temperature and $PM_{2.5}$ predictions are already rounded to the nearest tenth, so no further grouping is needed for a heatmap. For legibility, the temperature scale only shows the middle 95% of the data for each season, and the $PM_{2.5}$ scale only goes up to the 98th percentile for all seasons. Blue lines show the quartiles of $PM_{2.5}$ conditional on temperature.

Considering the 88,399 cell-days in which mean $PM_{2.5}$ exceeded Mexico's permissible daily limit of 41 $\mu g/m^3$,[65] the median temperature was 19.2 °C, somewhat warmer than the median in all other cell-days, 15.9 °C. Considering the 173,170 cell-days with a mean temperature of at least 20 °C, we found substantially higher median $PM_{2.5}$, 30.2 $\mu g/m^3$, than in all other cell-days, 19.7 $\mu g/m^3$.

We used population density from GPWv4 in every prediction cell of the study area to estimate person-days of $PM_{2.5}$ exposure in 2010, referring to Mexico's standards for annual and daily ambient concentrations of $PM_{2.5}$.[65] We compared the exposure estimated by our XGBoost-with-IDW model to that estimated by IDW alone, a $PM_{2.5}$ interpolation technique that has been used for a health-impact assessment in this region.[68] The study area contained 20,279,491 people in 2010. According to both our model and the IDW-only model, every single person in the Mexico City Metropolitan Area experienced a yearly mean $PM_{2.5}$ worse than the permissible limit of 10 $\mu g/m^3$. The large majority of people (97%, or more than 99% according to IDW) experienced a yearly mean more than twice the limit. Similarly, all people experienced at least one day with a mean $PM_{2.5}$ worse than the daily permissible limit of 41 $\mu g/m^3$. People experienced a mean of 21.6 (23.7 according to IDW) days exceeding the limit. The total number of exceeded person-days was 439 million (481 million according to IDW). Overall, we find widespread exposure to worse-than-permissible air pollution, although our full model suggests slightly less exposure than an IDW-only model. To show population exposure distributions over time, we also calculated the annual average concentration for each populated grid cell for each year, using more than 45 million model predictions. Figure 5 shows the empirical cumulative distribution functions for these annual concentrations calculated with 2010 census population densities. As observed in Figure 5, there has been an overall reduction in the annual

exposure to $PM_{2.5}$ since the earliest years (2004 and 2005); however, there is considerable variability in the estimated annual exposures, with less clear recent trends.

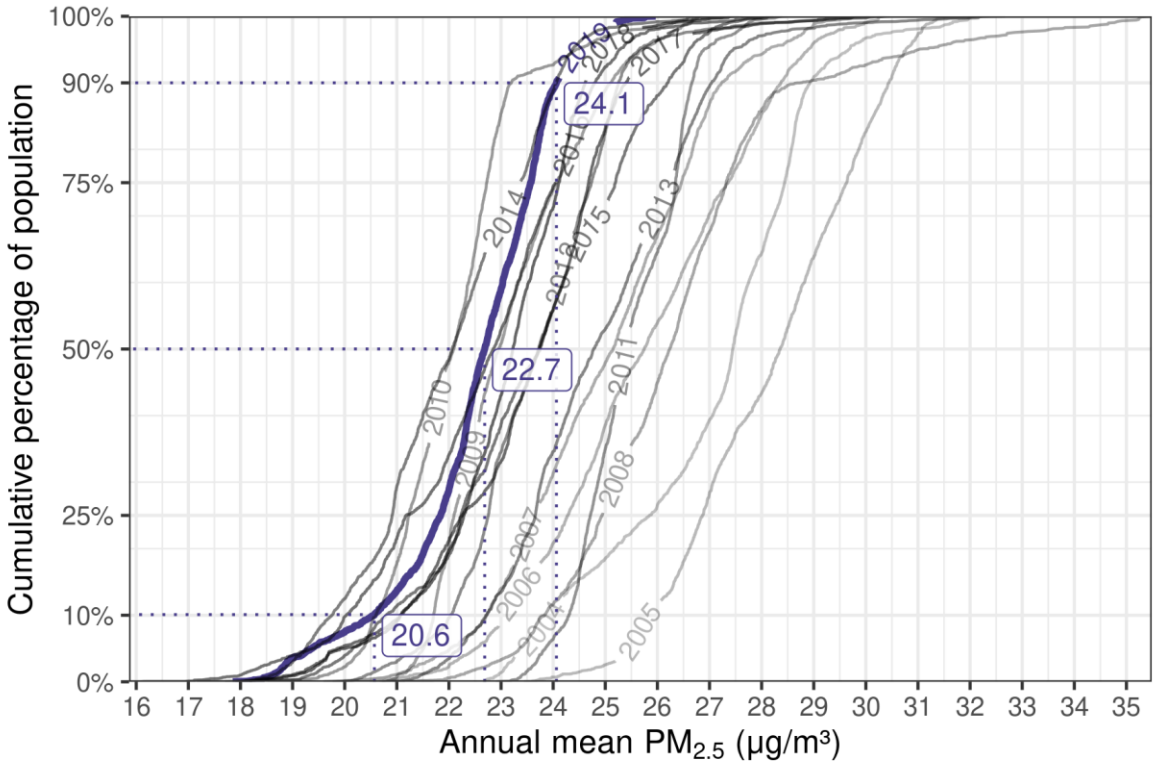


Figure 5. Population estimated annual average exposures. The figure shows an empirical cumulative distribution curve for each year from 2004 to 2019, generated from our daily mean model and using the 2010 census population density. Specific quantiles are labeled for the year 2019, where only 10% of the population in the study region had an annual average exposure below 20.6 µg/m³.

We used an index of social marginalization developed by the Consejo Nacional de Población (CONAPO), which considers access to education and health, housing characteristics, and possession of goods,[66] to compare urban marginalization in 2010 to mean $PM_{2.5}$. There were 2,065 AGEBs for which marginalization scores were available, with one score per AGEB and year, so we summarized mean $PM_{2.5}$ in 2010 by AGEB. Overall, marginalization and $PM_{2.5}$ were Kendall-correlated 0.024, which is a relationship in the expected direction (i.e., in the direction of more AGEBs with marginalized populations being exposed to more air pollution), but very weak. Breaking the AGEBs into 0.5-unit groups of marginalization (with one group for marginalization -2 to -1.5, one for -1.5 to -1, etc.), we find a small range of mean per-group $PM_{2.5}$, from 21.78 to 22.56 $\mu g/m^3$.

## 4. Discussion

We constructed and validated models to predict mean and max $PM_{2.5}$ in the Mexico City Metropolitan Area, and examined potential applications in air-pollution epidemiology and air-quality management. Our machine-learning-based model is the first of its kind in Mexico, although previously, our team used mixed-effects models with AOD to predict mean $PM_{2.5}$ in this region.[9] Also new is our consideration of max $PM_{2.5}$, an exposure metric that is becoming relevant to address subdaily health effects from peak exposures to $PM_{2.5}$.[69] Overall, our models exhibited good performance, with prediction errors that decreased over time, as the number of ground monitoring stations increased. Our per-year $R^2$ for mean $PM_{2.5}$ ranged from 0.64 to 0.86, similar to the $R^2$ values for our team's XGBoost model in the Northeastern US, which ranged from 0.64 to 0.80.[26] Our new modeling approach to predict $PM_{2.5}$ is not limited to one specific area; it could be extended to other regions with low or intermediate density of ground monitoring stations.

$PM_{2.5}$ predictions from AOD-$PM_{2.5}$ models have been used in epidemiology to reduce exposure measurement error, but may also be useful for applications such as air-quality management, particularly in sparsely monitored regions. Our maps of annual means of $PM_{2.5}$ (Figure 2) show wide variation in both $PM_{2.5}$ metrics across the Mexico City Metropolitan Area. More $PM_{2.5}$ has historically been observed in the center-north and center-east of the Mexico City Metropolitan Area (in the densely populated limits between Mexico City and the State of Mexico), where there are substantial emissions from industry and traffic.[70] Our $PM_{2.5}$ predictions allowed us to assess exposure to $PM_{2.5}$ in the entire Mexico City Metropolitan Area, unlike previous studies that could only partly cover this region with data from ground monitoring stations alone.[68] The estimated annual mean concentrations from our model exceeded the current annual $PM_{2.5}$ Mexican permissible limits across the entire study region, supporting previous results pointing out that despite significant improvements in the air quality of Mexico City for $PM_{10}$ and ozone since the 1990s, there remain substantial obstacles to reducing emissions of $PM_{2.5}$ and its precursors.[71] Using spatiotemporally-resolved exposure data in the study region (e.g., our $PM_{2.5}$ predictions) should improve future health impact assessments and support targeted exposure reduction strategies.[72]

Seasonally, there is a well-defined pattern of higher $PM_{2.5}$ concentrations during the two dry seasons, due to frequent thermal inversions and stable atmospheric conditions, which favors the accumulation of $PM_{2.5}$.[73] The lowest $PM_{2.5}$ concentrations occur during the rainy season, due to wet deposition.[74] We hypothesized that the pattern we saw in the daily ratios of mean and max $PM_{2.5}$ (Figure 2) reflects the influence of seasonal meteorological conditions. We checked whether higher ratios observed during the rainy season could be explained by precipitation, since late-afternoon showers can reduce $PM_{2.5}$.[74] However we found that days with at least 1 mm of daily precipitation had only a 5% greater ratio than other days.

In the context of climate change, it is important to characterize the increasingly common joint occurrence of extreme air pollution and extreme temperatures.[75] We found that while $PM_{2.5}$ and temperature are only weakly related overall, higher $PM_{2.5}$ concentrations tended to occur on warmer days, particularly in the rainy season (Figure 4), and conversely, days with mean temperatures of at least 20° C had a substantially worse median $PM_{2.5}$ concentration than cooler days. It has been reported that co-occurring extreme $PM_{2.5}$ and extreme temperatures may increase the acute risk of illness,[76] and that the influence of $PM_{2.5}$ on mortality rates may be stronger in warmer cities.[77] Previous studies in the Mexico City Metropolitan Area have suggested stronger associations with mortality on days with high $PM_{2.5}$ and extreme temperatures,[78] but they may have estimated effects imprecisely, given their citywide approach for estimating exposure. We expect that our $PM_{2.5}$ predictions can improve exposure assessment and air-pollution epidemiology, including studies addressing the interactive effects of $PM_{2.5}$ with temperature.[79]

Although air-quality standards are observed over geographic regions, and not at the individual level, it is important to quantify the extent of human exposure to unhealthy concentrations of air pollution during specific periods of time. This can aid in communicating the impacts of air pollution to stakeholders and decision-makers. To put into perspective the human cost of $PM_{2.5}$ exposure, we found that in 2010, every person in the study region was exposed to unhealthy air quality according to the Mexican standards for annual (10 $\mu g/m^3$) and daily (41 $\mu g/m^3$) concentrations, which are several times the recently enacted World Health Organization Guidelines of 5 and 15 $\mu g/m^3$, respectively.[80] Overall, in 2010 the population of the study region experienced a mean of nearly three weeks of $PM_{2.5}$ above the current daily Mexican permissible limit. For epidemiologic research, the distribution of continuous exposures is more relevant for health studies than the dichotomous assessment or duration of compliance with a particular standard. We combined our $PM_{2.5}$ predictions with population data to calculate annual empirical cumulative distributions for all inhabited areas in the study region, a summary of the population distribution of our exposure estimates that is suitable for assessment of long-term ambient $PM_{2.5}$ exposures and related chronic health effects.

Assessment of compliance with $PM_{2.5}$ standards is made with regulatory ground monitoring stations. Concentrations of $PM_{2.5}$ measured in a single monitoring station are used to represent the pollution conditions over large spatial domains (up to tens of kilometers) for a specific amount of time, such as one day or one year. However, $PM_{2.5}$ levels can be rapidly influenced by local sources, increasing not only concentrations between monitoring sites, but also the risks of acute health effects. $PM_{2.5}$ measurements from ground monitoring stations alone might be insufficient to represent air quality in large areas, especially when the objective is to protect human health. A distinguishing feature of our model is that we also generated a sub-daily metric of $PM_{2.5}$ concentrations, namely, max $PM_{2.5}$. There are not yet any air-quality standards for sub-daily $PM_{2.5}$ concentrations, but new research into the

health impacts from such exposures could eventually support standards for shorter exposure timeframes.[81–84] The US Environmental Protection Agency states that *"*Because a focus on annual average and 24-hour average $PM_{2.5}$ concentrations could mask sub-daily patterns, and because some health studies examine PM exposure durations shorter than 24-hours, it is useful to understand the broader distribution of sub-daily $PM_{2.5}$ concentrations*"*.[69] Because it's more challenging to reconstruct extrema (e.g., max $PM_{2.5}$) versus measures of central tendency (e.g., mean $PM_{2.5}$), future work on estimating health impacts from max $PM_{2.5}$ exposures could particularly benefit from estimating and propagating prediction uncertainty into downstream analyses.[85, 86]

Our comparison of $PM_{2.5}$ exposure across levels of social marginalization did not suggest meaningful differences between groups. However, the 2010 Mexican index of social marginalization was only available for urban AGEBs: those with a total population of more than 2,500. Without data for rural AGEBs or irregular settlements, it is naturally more difficult to assess the influence of socioeconomic status. Also, because the methods employed in the construction of the Mexican index of social marginalization have changed over time, it would be difficult to analyze multiple years and make sense of the differences between them.

A limitation of our model arises from the limited temporal resolution in the AOD data. Each satellite passes over the Central Mexico region only once during each period of daylight, possibly missing sudden episodes of intense $PM_{2.5}$. However, the overpass time of the Terra satellite is similar to the daily peak of $PM_{2.5}$ according to ground monitoring stations, so in general, Terra AOD should be representative of max $PM_{2.5}$. Future work will utilize AOD data from the Advanced Baseline Imager (ABI) aboard NOAA's Geostationary Operational Environmental Satellite - R Series (GOES-16 and GOES-17) with temporal resolution as high as 5 minutes over Mexico City. Synergistic AOD products developed from the ABI and upcoming NASA geostationary Tropospheric Emissions: Monitoring of Pollution (TEMPO) mission, planned for launch in December 2022, will further enhance capabilities to predict and monitor $PM_{2.5}$ concentrations in the region. TEMPO will advance exposure science in North America, particularly by providing hourly observations of aerosols and gaseous pollutants for supporting air-pollution models.[87, 88]

**Acknowledgments:**

*Data used in the preparation of this article were obtained from the Sistema Nacional de Información de Calidad del Aire (SINAICA) from the Instituto Nacional de Ecología y Cambio Climático (INECC). We also obtained ground monitoring stations' geographic location from INECC's Coordinación General de Contaminación y Salud Ambiental and advice from Maria Guadalupe Tzintzun Cervantes, Sub-Directorate of Air Quality.*

**Funding information**

*This work was supported by grants from the National Institutes of Health (NIH): R01 ES013744, R01 ES014930, R01 ES021357, R01 ES031295, R01 ES032242, R24 ES028522, P30 ES023515. D.C. was supported by T32 HD049311.*

**References**


1. World Health Organization. Fact sheet on ambient (outdoor) air pollution. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health (2021).
2. Bowe, B., Xie, Y., Yan, Y. & Al-Aly, Z. Burden of Cause-Specific Mortality Associated With PM2.5 Air Pollution in the United States. *JAMA Netw Open* **2**, e1915834 (2019).
3. Gu, J. *et al.* Ambient air pollution and cause-specific risk of hospital admission in China: A nationwide time-series study. *PLoS Med.* **17**, e1003188 (2020).
4. Bekkar, B., Pacheco, S., Basu, R. & DeNicola, N. Association of Air Pollution and Heat Exposure With Preterm Birth, Low Birth Weight, and Stillbirth in the US: A Systematic Review. *JAMA Netw Open* **3**, e208243 (2020).
5. van Donkelaar, A. *et al.* Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect.* **118**, 847–855 (2010).
6. Ma, Z., Hu, X., Huang, L., Bi, J. & Liu, Y. Estimating ground-level PM2.5 in China using satellite remote sensing. *Environ. Sci. Technol.* **48**, 7436–7444 (2014).
7. Hu, X. *et al.* Estimating ground-level PM2.5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sens. Environ.* **140**, 220–232 (2014).
8. Kloog, I. *et al.* A new hybrid spatio-temporal model for estimating daily multi-year PM2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. *Atmos. Environ.* **95**, 581–590 (2014).
9. Just, A. C. *et al.* Using High-Resolution Satellite Aerosol Optical Depth To Estimate Daily PM2.5 Geographical Distribution in Mexico City. *Environ. Sci. Technol.* **49**, 8576–8584 (2015).
10. Kloog, I. *et al.* Estimating daily PM2.5 and PM10 across the complex geo-climate region of Israel using MAIAC satellite-based AOD data. *Atmos. Environ.* **122**, 409–416 (2015).
11. Sorek-Hamer, M. *et al.* Assessment of PM2.5 concentrations over bright surfaces using MODIS satellite observations. *Remote Sens. Environ.* **163**, 180–185 (2015).
12. Zhang, Y. & Li, Z. Remote sensing of atmospheric fine particulate matter (PM2.5) mass concentration near the ground from satellite observation. *Remote Sensing of Environment* vol. 160 252–262 (2015).
13. Chu, Y. *et al.* A Review on Predicting Ground PM2.5 Concentration Using Satellite Aerosol Optical Depth. *Atmosphere* **7**, 129 (2016).
14. Sorek-Hamer, M., Chatfield, R. & Liu, Y. Review: Strategies for using satellite-based products in modeling PM2.5 and short-term pollution episodes. *Environment International* vol. 144 106057 (2020).
15. Strode, S. A. *et al.* Global changes in the diurnal cycle of surface ozone. *Atmos. Environ.* **199**, 323–333 (2019).
16. Lyapustin, A., Wang, Y., Korkin, S. & Huang, D. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* **11**, 5741–5765 (2018).
17. Li, L. *et al.* Spatiotemporal Imputation of MAIAC AOD Using Deep Learning with Downscaling. *Remote Sens. Environ.* **237**, (2020).
18. Lary, D. J. *et al.* Estimating the global abundance of ground level presence of particulate matter (PM2.5). *Geospat. Health* **8**, S611–30 (2014).
19. Hu, X. *et al.* Estimating PM2.5 Concentrations in the Conterminous United States Using the Random Forest Approach. *Environ. Sci. Technol.* **51**, 6936–6944 (2017).
20. Zhan, Y. *et al.* Spatiotemporal prediction of continuous daily PM2.5 concentrations across China using a spatially explicit machine learning algorithm. *Atmos. Environ.* **155**, 129–139 (2017).

21. Chen, G. *et al.* A machine learning method to estimate PM2.5 concentrations across China with remote sensing, meteorological and land use information. *Sci. Total Environ.* **636**, 52–60 (2018).
22. Stafoggia, M. *et al.* Estimation of daily PM10 and PM2.5 concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. *Environ. Int.* **124**, 170–179 (2019).
23. Di, Q. *et al.* An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environment International* vol. 130 104909 (2019).
24. Mhawish, A. *et al.* Estimation of High-Resolution PM2.5 over the Indo-Gangetic Plain by Fusion of Satellite Data, Meteorology, and Land Use Variables. *Environ. Sci. Technol.* **54**, 7891–7900 (2020).
25. Schneider, R. *et al.* A Satellite-Based Spatio-Temporal Machine Learning Model to Reconstruct Daily PM2.5 Concentrations across Great Britain. *Remote Sens (Basel)* **12**, 3803 (2020).
26. Just, A. C. *et al.* Advancing methodologies for applying machine learning and evaluating spatiotemporal models of fine particulate matter (PM2.5) using satellite data over large regions. *Atmos. Environ.* **239**, (2020).
27. Lyapustin, A., Martonchik, J., Wang, Y., Laszlo, I. & Korkin, S. Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. *J. Geophys. Res.* **116**, (2011).
28. Lyapustin, A. *et al.* Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *J. Geophys. Res.* **116**, (2011).
29. Bansal, E. *et al.* Prenatal PM2.5 exposure in the second and third trimesters predicts neurocognitive performance at age 9–10 years: A cohort study of Mexico City children. *Environmental Research* vol. 202 111651 (2021).
30. Prada, D. *et al.* Long-term PM2.5 exposure before diagnosis is associated with worse outcome in breast cancer. *Breast Cancer Research and Treatment* vol. 188 525–533 (2021).
31. Hurtado-Díaz, M. *et al.* Prenatal PM2.5 exposure and neurodevelopment at 2 years of age in a birth cohort from Mexico city. *Int. J. Hyg. Environ. Health* **233**, 113695 (2021).
32. Tamayo-Ortiz, M. *et al.* Exposure to PM2.5 and Obesity Prevalence in the Greater Mexico City Area. *Int. J. Environ. Res. Public Health* **18**, (2021).
33. Rivera Rivera, N. Y. *et al.* Prenatal and early life exposure to particulate matter, environmental tobacco smoke and respiratory symptoms in Mexican children. *Environ. Res.* **192**, 110365 (2021).
34. McGuinn, L. A. *et al.* Prenatal PM2.5 exposure and behavioral development in children from Mexico City. *Neurotoxicology* **81**, 109–115 (2020).
35. Wu, H. *et al.* Association of ambient PM2·5 exposure with maternal bone strength in pregnant women from Mexico City: a longitudinal cohort study. *Lancet Planet. Health* **4**, e530–e537 (2020).
36. Rosa, M. J. *et al.* Identifying critical windows of prenatal particulate matter (PM2.5) exposure and early childhood blood pressure. *Environmental Research* vol. 182 109073 (2020).
37. Niedzwiecki, M. M. *et al.* Particulate air pollution exposure during pregnancy and postpartum depression symptoms in women in Mexico City. *Environ. Int.* **134**, 105325 (2020).
38. Téllez-Rojo, M. M. *et al.* Children's acute respiratory symptoms associated with PM2.5 estimates in two sequential representative surveys from the Mexico City Metropolitan Area. *Environ. Res.* **180**, 108868 (2020).
39. McGuinn, L. A. *et al.* Fine particulate matter exposure and lipid levels among children in Mexico city. *Environ. Epidemiol.* **4**, e088 (2020).
40. Moody, E. C. *et al.* Association of Prenatal and Perinatal Exposures to Particulate Matter With Changes in Hemoglobin A1c Levels in Children Aged 4 to 6 Years. *JAMA Netw Open* **2**, e1917643 (2019).
41. Rosa, M. J. *et al.* Association between prenatal particulate air pollution exposure and telomere length in cord blood: Effect modification by fetal sex. *Environ. Res.* **172**, 495–501 (2019).

42. Bose, S. *et al.* Prenatal particulate air pollution exposure and sleep disruption in preschoolers: Windows of susceptibility. *Environ. Int.* **124**, 329–335 (2019).
43. Gutiérrez-Avila, I. *et al.* Cardiovascular and Cerebrovascular Mortality Associated With Acute Exposure to PM 2.5 in Mexico City. *Stroke* vol. 49 1734–1736 (2018).
44. Rosa, M. J. *et al.* Prenatal particulate matter exposure and wheeze in Mexican children: Effect modification by prenatal psychosocial stress. *Ann. Allergy Asthma Immunol.* **119**, 232–237.e1 (2017).
45. Rosa, M. J. *et al.* Prenatal exposure to PM 2.5 and birth weight: A pooled analysis from three North American longitudinal pregnancy cohort studies. *Environment International* vol. 107 173–180 (2017).
46. Chilian-Herrera, O. L. *et al.* PM exposure as a risk factor for type 2 diabetes mellitus in the Mexico City metropolitan area. *BMC Public Health* **21**, 2087 (2021).
47. Lome-Hurtado, A., Touza-Montero, J. & White, P. C. L. Environmental Injustice in Mexico City: A Spatial Quantile Approach. *Exposure and Health* **12**, 265–279 (2020).
48. Rincón, V. R., Martínez-Alier, J. & Mingorria, S. Environmental Conflicts Related to Urban Expansion Involving Agrarian Communities in Central Mexico. *Sustainability* vol. 11 6545 (2019).
49. Bravo, M. A. *et al.* Airborne fine particles and risk of hospital admissions for understudied populations: Effects by urbanicity and short-term cumulative exposures in 708 U.s. counties. *Environ. Health Perspect.* **125**, (2016).
50. Guo, W., Chai, Y. & Kwan, M.-P. Chapter 13 - Travel-related exposure to air pollution and its socio-environmental inequalities: Evidence from a week-long GPS-based travel diary dataset. in *Spatiotemporal Analysis of Air Pollution and Its Application in Public Health* (eds. Li, L., Zhou, X. & Tong, W.) 293–309 (Elsevier, 2020).
51. Southerland, V. A. *et al.* Global urban temporal trends in fine particulate matter (PM) and attributable health burdens: estimates from global datasets. *Lancet Planet Health* (2022) doi:10.1016/S2542-5196(21)00350-8.
52. Andreão, W. L. & Toledo de Almeida Albuquerque, T. Avoidable mortality by implementing more restrictive fine particles standards in Brazil: An estimation using satellite surface data. *Environ. Res.* **192**, 110288 (2021).
53. Fowlie, M., Rubin, E. & Walker, R. Bringing Satellite-Based Air Quality Estimates Down to Earth. *AEA Papers and Proceedings* **109**, 283–288 (2019).
54. Diao, M. *et al.* Methods, availability, and applications of PM2.5 exposure estimates derived from ground measurements, satellite, and atmospheric models. *Journal of the Air & Waste Management Association* vol. 69 1391–1414 (2019).
55. Hammer, M. S. *et al.* Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998-2018). *Environ. Sci. Technol.* **54**, 7879–7890 (2020).
56. Gutiérrez-Avila, I. *et al.* A spatiotemporal reconstruction of daily ambient temperature using satellite data in the Megalopolis of Central Mexico from 2003 to 2019. *Int. J. Climatol.* **41**, 4095–4111 (2021).
57. Jáuregui, E. The Climate of the Mexico City Air Basin: Its Effects on the Formation and Transport of Pollutants. in *Urban Air Pollution and Forests: Resources at Risk in the Mexico City Air Basin* (eds. Fenn, M. E., de Bauer, L. I. & Hernández-Tejeda, T.) 86–117 (Springer New York, 2002).
58. Lyapustin, A. & Wang, Y. MCD19A2 MODIS/Terra+aqua land aerosol optical depth daily L2G global 1km SIN grid V006. (2018) doi:10.5067/MODIS/MCD19A2.006.
59. Thornton, M. M. *et al.* Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4. (2020) doi:10.3334/ORNLDAAC/1840.
60. Hersbach, H. *et al.* ERA5 hourly data on single levels from 1979 to present. Copernicus Climate

Change Service (C3S) Climate Data Store (CDS). doi:10.24381/cds.adbb2d47.
61. OpenStreetMap Wiki contributors. Main page. https://wiki.openstreetmap.org/w/index.php?title=Main_Page&oldid=1060762.
62. Chen, T. & Guestrin, C. XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016) doi:10.1145/2939672.2939785.
63. Baston, D. exactextractr: Fast extraction from raster datasets using polygons. *R package version 0. 5. 0* (2020).
64. Center for International Earth Science Information Network-CIESIN-Columbia University. Gridded population of the world, version 4 (GPWv4): population density. *NASA Socioeconomic Data and Applications Center (SEDAC)* (2016).
65. SSA. Norma Oficial Mexicana NOM-025-SSA1-2021, Salud ambiental. Valores límite permisibles para la concentración de partículas suspendidas PM10 y PM2.5 en el aire ambiente y criterios para su evaluación. *Secretaría de Salud. Diario Oficial de la Federación* (2021).
66. Consejo Nacional de Población. Datos Abiertos del Índice de Marginación. http://www.conapo.gob.mx/es/CONAPO/Datos_Abiertos_del_Indice_de_Marginacion (2013).
67. USAID. Actualización al diagnóstico de la Megalópolis del centro de México. Mexico Low Emissions Development Program (MLED). http://www.plataformaleds.org/productos-programa-mled.php (2014).
68. Trejo-González, A. G. *et al.* Quantifying health impacts and economic costs of PM2.5 exposure in Mexican cities of the National Urban System. *Int. J. Public Health* **64**, 561–572 (2019).
69. Office of Air Quality Planning and Standards Health and Environmental Impacts Division Research Triangle Park, NC. *Policy Assessment for the Review of the National Ambient Air Quality Standards for Particulate Matter*. https://www.epa.gov/sites/default/files/2020-01/documents/final_policy_assessment_for_the_review_of_the_pm_naaqs_01-2020.pdf (2020).
70. Instituto Nacional de Ecología y Cambio Climático, INECC. Informe Nacional de Calidad del Aire 2019, México. https://sinaica.inecc.gob.mx/archivo/informes/Informe2019.pdf (2020).
71. Secretaría del Medio Ambiente. *Historical Analysis of Population Health Benefits Associated with Air Quality in Mexico City during 1990 and 2015*. http://www.data.sedema.cdmx.gob.mx/beneficios-en-salud-por-la-mejora-de-la-calidad-del-aire/descargas/analisis-ingles.pdf (2018).
72. Martenies, S. E., Wilkins, D. & Batterman, S. A. Health impact metrics for air pollution management strategies. *Environ. Int.* **85**, 84–95 (2015).
73. Molina, L. T., Velasco, E., Retama, A. & Zavala, M. Experience from Integrated Air Quality Management in the Mexico City Metropolitan Area and Singapore. *Atmosphere* **10**, 512 (2019).
74. Molina, L. T., de Foy, B. & Martínez, O. V. Air quality, weather and climate in Mexico City. *WMO Bulletin 58 (1) - January 2009*.
75. Kinney, P. L. Interactions of Climate Change, Air Pollution, and Human Health. *Curr Environ Health Rep* **5**, 179–186 (2018).
76. Yitshak-Sade, Maayan, Jennifer F. Bobb, Joel D. Schwartz, Itai Kloog, and Antonella Zanobetti. The association between short and long-term exposure to PM2.5 and temperature and hospital admissions in New England and the synergistic effect of the short-term exposures. *Sci. Total Environ.* **639**, 868–875 (2018).
77. Kioumourtzoglou, M.-A., Schwartz, J., James, P., Dominici, F. & Zanobetti, A. PM2.5 and Mortality in 207 US Cities: Modification by Temperature and City Characteristics. *Epidemiology* **27**, 221–227 (2016).
78. Godwin, W. *Assessment of interactive effects of temperature and air pollution on mortality in Mexico City*. (2018).
79. Gutiérrez Avila, I. *et al.* Daily particulate matter and temperature from satellite-hybrid models and 1.5 million deaths: A time-stratified case-crossover analysis in Central Mexico. *Environ.*

*Health Perspect.* **2021**, (2021).
80. World Health Organization. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide: executive summary. (2021).
81. Darrow, L. A. *et al.* The use of alternative pollutant metrics in time-series studies of ambient air pollution and respiratory emergency department visits. *J. Expo. Sci. Environ. Epidemiol.* **21**, 10–19 (2011).
82. Son, J.-Y. & Bell, M. L. The relationships between short-term exposure to particulate matter and mortality in Korea: Impact of particulate matter exposure metrics for sub-daily exposures. *Environ. Res. Lett.* **8**, 014015 (2013).
83. Lin, H. *et al.* Hourly peak PM2.5 concentration associated with increased cardiovascular mortality in Guangzhou, China. *J. Expo. Sci. Environ. Epidemiol.* **27**, 333–338 (2017).
84. Link, M. S. *et al.* Acute Exposure to Air Pollution Triggers Atrial Fibrillation. *J. Am. Coll. Cardiol.* **62**, 816–825 (2013).
85. Chang, H. H., Hu, X. & Liu, Y. Calibrating MODIS aerosol optical depth for predicting daily PM2.5 concentrations via statistical downscaling. *J. Expo. Sci. Environ. Epidemiol.* **24**, 398–404 (2013).
86. Keller, J. P., Chang, H. H., Strickland, M. J. & Szpiro, A. A. Measurement Error Correction for Predicted Spatiotemporal Air Pollution Exposures. *Epidemiology* **28**, 338–345 (2017).
87. Zoogman, P., X. Liu, R. M. Suleiman, W. F. Pennington, D. E. Flittner, J. A. Al-Saadi, B. B. Hilton, et al. Tropospheric emissions: Monitoring of pollution (TEMPO). *J. Quant. Spectrosc. Radiat. Transf.* **186**, 17–39 (2017).
88. Naeger, A. R. *et al.* Revolutionary Air-Pollution Applications from Future Tropospheric Emissions: Monitoring of Pollution (TEMPO) Observations. *Bull. Am. Meteorol. Soc.* **102**, E1735–E1741 (2021).