Overview of Models

These models provide estimates of outdoor concentrations for six pollutants (four gases: O3, CO, SO2, NO2; Two aerosols: PM10, PM2.5) throughout the contiguous U.S. Model estimates, with the exception of ozone are annual-average values, for years with available monitoring data. Ozone model estimates are the average during May through September of the daily maximum 8-hour moving average, for years with available monitoring data.

Pollutant	Units	Available Years
O_3	ppb	1979-2015
SO_2	ppb	1979-2015
NO_2	ppb	1979-2015
PM_{10}	$\mu g/m^3$	1988-2015
CO	ppm	1990-2015
$PM_{2.5}$	$\mu g/m^3$	1999-2015

Data are available at national, state, county, census tract*, and census block group* levels, based on population-weighted averages of census block level predictions using 2010 census geographies.

*Our models only provide estimates for tracts and block groups with a population greater than zero.

These models are derived from publicly available concentration measurements from U.S. EPA regulatory monitors, and use information about land use (for example, locations of major and minor roads; elevation; and whether an area is urban or rural) and satellite-derived estimates of air pollution to predict concentrations at locations without measurements. Researchers generally do not use estimates of the types given here to draw conclusions about air quality at one location or a small number of specific locations; as such, we do not recommend that type of use.

File Format

Parameter	Description
Fips	Unique code for geographic unit with the following structure: SSCCCTTTTTTG[1] [1] S = 1-2 digit state code; C = 3 digit county code; T = 6 digit tract code; G = 1 digit block group code This code can be used to join these data to Census data and shapefiles
pollutant	Pollutant for a given observation (co, no2, o3, pm10, pm25, so2)
year	Prediction year for a given observation
pred_wght	Population-weighted concentration based on block level centroid predictions, units vary by pollutant (see above)
state_abbr	State abbreviation for a given observation
lat/lon	Population-weighted latitude and longitude based on block level centroid

Acknowledgment

Data are free and publicly available for non-commercial purposes, data may not be used for commercial purposes without prior permission. Please cite as follows:

"This article includes concentration estimates developed by the Center for Air, Climate and Energy Solutions using v1 empirical models as described in Kim et al., 2018."

Suggested citation: Kim S.-Y.; Bechle, M.; Hankey, S.; Sheppard, L.; Szpiro, A. A.; Marshall, J. D. 2018. "Concentrations of criteria pollutants in the contiguous U.S., 1979 – 2015: Role of prediction model parsimony in integrated empirical geographic regression." PLoS ONE 15(2), e0228535. DOI: 10.1371/journal.pone.0228535

Overview of Ultrafine (Particle Number Concentration) Model

Model predictions: Our model provides estimates of outdoor concentrations for particle number concentrations (PNC; unit: particles number / cm^3) throughout the contiguous U.S. *Model estimates are annual-average values for 2017*. The model is derived using particle number concentration measurements from mobile monitoring and fixed sites across the U.S. and a land-use regression modeling framework to predict concentrations at locations without measurements. It is noted that we only made predictions at census blocks with predictor variable values within the 1st and 99th percentile range of the measurement dataset used for model development. This means that model predictions do not extrapolate in covariate space. After applying this constraint, the model predicts outdoor concentration at 6,056,703 residential census block centroids in the contiguous US, which corresponds to 98.6% of the population. This constraint censors predictions for 1.9% of census blocks (1.4% of the population) in the contiguous US. Censored census blocks are generally in extremely urban (e.g., in Manhattan, New York City) and rural locations.

Population-weighted average concentrations: Estimates available here are at national, state, county, census tract, and census block group levels, based on population-weighted averages of census block level predictions using 2010 census geographies. You may get the census block level estimates by contracting Albert A. Presto (albert.presto@gmail.com) or Provat K. Saha (sahaprovat@gmail.com).

File format

fips: Unique code for geographic unit with the following structure: SSCCCTTTTTTG

(S = 1–2-digit state code; C = 3-digit county code; T = 6-digit tract code; G = 1 digit block group code).

PNC: Population-weighted particle number concentration based on block level centroid predictions, unit (particle number /cm^3)

lat/lon: Population-weighted latitude and longitude based on block level centroid

Suggested citation: Saha, P. K.; Hankey, S.; Marshall, J. D.; Robinson, A. L.; Presto, A. A. High-Spatial-Resolution Estimates of Ultrafine Particle Concentrations across the Continental United States. Environ. Sci. Technol. 2021, 55 (15), 10320–10331. https://doi.org/10.1021/acs.est.1c03237.