RESEARCH ARTICLE

A hyperlocal hybrid data fusion near-road PM$_{2.5}$ and NO$_2$ annual risk and environmental justice assessment across the United States

Alejandro Valencia[1,2], **Marc Serre**[1], **Saravanan Arunachalam**[2]*

1 Department of Environmental Sciences and Engineering, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **2** Institute for the Environment, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America

* sarav@email.unc.edu

Abstract

Exposure to traffic-related air pollutants (TRAPs) has been associated with numerous adverse health effects. TRAP concentrations are highest meters away from major roads, and disproportionately affect minority (i.e., non-white) populations often considered the most vulnerable to TRAP exposure. To demonstrate an improved assessment of on-road emissions and to quantify exposure inequity in this population, we develop and apply a hybrid data fusion approach that utilizes the combined strength of air quality observations and regional/local scale models to estimate air pollution exposures at census block resolution for the entire U.S. We use the regional photochemical grid model CMAQ (Community Multiscale Air Quality) to predict the spatiotemporal impacts at local/regional scales, and the local scale dispersion model, R-LINE (Research LINE source) to estimate concentrations that capture the sharp TRAP gradients from roads. We further apply the Regionalized Air quality Model Performance (RAMP) Hybrid data fusion technique to consider the model's nonhomogeneous, nonlinear performance to not only improve exposure estimates, but also achieve significant model performance improvement. With a R^2 of 0.51 for PM$_{2.5}$ and 0.81 for NO$_2$, the RAMP hybrid method improved R^2 by ~0.2 for both pollutants (an increase of up to ~70% for PM$_{2.5}$ and ~31% NO$_2$). Using the RAMP Hybrid method, we estimate 264,516 [95% confidence interval [CI], 223,506–307,577] premature deaths attributable to PM$_{2.5}$ from all sources, a ~1% overall decrease in CMAQ-estimated premature mortality compared to RAMP Hybrid, despite increases and decreases in some locations. For NO$_2$, RAMP Hybrid estimates 138,550 [69,275–207,826] premature deaths, a ~19% increase (22,576 [11,288 – 33,864]) compared to CMAQ. Finally, using our RAMP hybrid method to estimate exposure inequity across the U.S., we estimate that Minorities within 100 m from major roads are exposed to up to 15% more PM$_{2.5}$ and up to 35% more NO$_2$ than their White counterparts.

Introduction

Traffic related emissions are a significant source of urban air pollution [1]. Exposure to traffic related air pollutants (TRAPs) such as PM$_{2.5}$ and NO$_2$ has been associated with a myriad of

adverse health effects [2–8]. Additionally, concentrations from TRAPs are the highest in the vicinity of heavily trafficked roads [9, 10]. In the U.S., around 19% of the population resides near heavily trafficked roads and in states with higher population e.g., California, up to 40% of the population live close to TRAPs [11]. These estimates not only reinforce how widespread the exposure to TRAPs is in the U.S., but it also shows that the risk is highly variable and potentially greater in areas with higher population and increased road densities. In addition, many studies have shown that a disproportionate amount of the health risk is placed on Minority populations [12–22], a group which is already found to be more susceptible to air pollution exposure.

Assessing the burden of disease from TRAPs usually involves air quality and health modeling. Typically, these nationwide modeling efforts have been conducted with chemical transport models (CTMs) at coarse (12 km x 12 km or even 36 km x 36 km) spatial resolutions [23–31]. Dedoussi et al. [31] estimated 30,800 annual premature deaths in 2011 across the U.S. due to road transportation attributable to O$_3$ and PM$_{2.5}$ using 55-km x 55-km grid resolution using the GEOS-Chem CTM but did not consider secondary organic aerosols in their estimate. Davidson et al. [30] used a 12 km x 12 km (12-km) resolution model and estimated 9,666 annual premature deaths across the U.S. that were attributable to on-road PM$_{2.5}$. This corresponded to 9.6% of the total PM$_{2.5}$ attributable deaths for 2011. Additionally, Davidson et al. (2020) project that in 2025 the on-road premature deaths attributable to PM$_{2.5}$ across the U.S. will correspond to 5.6% of the total premature deaths. These estimates at coarse resolutions, however, lack the spatial gradients necessary to assess the relationship between population exposure to TRAPs and health impacts. A few studies have explored the effect of grid resolution on estimates of air pollution from CTMs on premature mortality [27, 32–34]. But even at 1 km x 1 km (1-km) resolution, these types of studies might not fully capture the fine-scale TRAP concentration gradients. This can be attributed to the fact that most TRAP concentrations are highest near roads and decrease to background levels within 150 m to 200 m from the road [10, 35, 36]. Thus, capturing variability within a grid cell at fine spatial scales is critical for risk assessment. In risk assessment, higher population density is usually positively associated with air pollution, thus a method that does not capture within-cell variability will suffer from exposure misclassification. Therefore, diverse methods have been developed to address the uncertainty due to the dramatic spatial variation of TRAPs at even smaller spatial resolutions. Some studies have developed statistical regression-based models that estimate TRAPs at a spatial resolution of < 1 km [37–44]. These models, however, use surrogate indicators of emissions that are specific to an area and ignore the physical and chemical processes (e.g., dispersion, advection/diffusion, chemical reactions in gaseous and aerosol phases, and deposition) that can better quantify road source contribution.

To overcome these challenges, hybrid approaches have been developed that take coarse scale predictions of TRAPs from CTMs and combines them with high resolution dispersion model estimates [45–51]. As opposed to CTMs which allocate emission to a grid, dispersion modeling can represent emissions in a much finer spatial resolution that retains the physical characteristics and shape of the emitted plume. Because most dispersion models lack the ability to predict secondary pollutants such as secondary PM$_{2.5}$, combining these models avoids underestimating the total impact of a pollutant. Thus, by combining results from CTMs with dispersion models, predictions will include regional background, detailed chemical mechanisms, and emission sources that might not be included in a dispersion model. Using a hybrid air quality model, Parvez and Wagstrom [51] studied the effect of resolution (census block group vs 12-km grid resolution) on health impacts attributable to PM$_{2.5}$ in three locations in Connecticut. Their findings showed an underestimation of health impacts at the coarser resolution. Their study, however, upscaled hybrid PM$_{2.5}$ concentrations from 0.04 km x 0.04 km

grid resolution to census block groups (losing the fine-scale variability necessary to capture elevated air pollution from roads). The study by Chang et al. [47], which estimated TRAP concentrations using a hybrid model at census block level, concluded that their hybrid approach estimated 24% more on-road PM$_{2.5}$-related premature mortality than a CTM at a 36 km x 36 km resolution in central North Carolina. However, neither of these studies considered the biases in either model (i.e., dispersion model or hybrid model), and results from the modeling were localized to a small region. Recently, Shukla et al. [50] developed a hybrid modeling approach that combines C-TOOLS [52], a local-scale dispersion model for primary PM$_{2.5}$ with the CoBenefits Risk Assessment (COBRA) Screening model [53] with secondary PM$_{2.5}$ components. This model was used to characterize ZIP-code level air pollution estimates of PM$_{2.5}$ in New York City for performing rapid assessment for evaluating health benefits of emissions reduction measures.

Continuing these efforts, our work describes and applies a novel hybrid data fusion method that combines model predictions from a dispersion model specially designed to predict on-road emission and a CTM at 12-km resolution with observations to produce fine-scale air quality characterization of PM$_{2.5}$ and NO$_2$ at census block levels across the continental U.S. In this approach, we also consider the nonhomogeneous, nonlinear behavior of model performance and thus have the ability to correct biases at fine-scale. Additionally, by estimating air pollution at high resolution our model will be able to assess whether our method can capture the variability necessary to better characterize the spatial gradient from TRAPs to provide more accurate assessments of the health risk overall. This approach will aid in further identifying vulnerable populations, quantify their exposure inequity near and away from roads, and prevent potential exposure misclassification.

Methods

Hybrid data fusion overview for 2016

The hybrid data fusion approach used in this study combines the gaussian Research LINE source dispersion model (R-LINE) [54] developed for near-roadway assessments, with the Community Multiscale Air Quality Model (CMAQ) CTM [55] version 5.2.1 [56] with Carbon Bond 6 version r3 at 12-km horizontal grid resolution with aero6 treatment of secondary organic aerosols (SOA) set up for standard cloud chemistry for the year 2016. The data fusion approach then applies the Regionalized Air quality Model Performance (RAMP) method that uses observations from the Air Quality System (AQS) network to correct biases in the model [57]. This hybrid data fusion approach models total hourly average PM$_{2.5}$ and NO$_2$ concentration that consider the fine-scale resolution of on-road emissions for 2016 for each of the 11,063,509 Census blocks for a full year from January 2016 to December 2016 across the continental U.S.

Measurement data processing

Hourly, daily, and annually observed NO$_2$ and PM$_{2.5}$ from EPA's AQS database [58] for each space/time location during 2016 were collected from monitoring stations.

Emissions and meteorological data processing

Emissions for CMAQ and R-LINE are based on the 2016v7.2 (beta) platform [59]. This platform was produced by the National Emissions Inventory Collaborative and contains the emission inventories for 2016. For R-LINE, specifically, this platform provided emissions factors, fleet mix, and temporal allocation tables. These datasets were combined with road network

data like geometry and annual average daily traffic (AADT) for major (or primary) road types, including interstates, principal arteries, minor arteries, and major collectors from the United States Federal Highway Administration (FHWA)'s Highway performance monitoring system (HPMS). The data set consists of ~7.5 million major road segments also from 2016 and was used to calculate hourly emissions using an approach similar to previous studies [60–62]

The gridded meteorological data for 2016 were obtained from version 3.8 of the Weather Research and Forecasting Model (WRF) [63]. For CMAQ, WRF meteorological outputs were processed using the Meteorology-Chemistry Interface Processor (MCIP) package [64], version 4.3 for the continental U.S. at a 12-km resolution. The meteorological inputs for R-LINE were also based on the same 12-km resolution WRF data used for CMAQ. Given that our hybrid data fusion approach will rely on R-LINE estimates for all major roadway sources within every CMAQ grid cell, R-LINE ready meteorological files were created using the Mesoscale Model Interface Program (MMIF) [65] for every grid cell. MMIF outputs have been shown to compare well against observed meteorological data and have been applied to dispersion models to address sparse observations networks [49, 66, 67].

Hybrid model

To combine CMAQ and R-LINE predictions we follow approaches similar to previous studies [45, 48, 49, 68] as discussed below. Thus, to create PM$_{2.5}$ estimates, we first remove primary roadway contributions from CMAQ to avoid double counting (This is also done for NO$_2$ estimates). It is of note, that R-LINE can only estimate primary PM$_{2.5}$ and thus we rely on CMAQ to provide secondary PM$_{2.5}$. Additionally, the effect of resolution shows greater biases for primary PM$_{2.5}$ species than secondary species [27]. To begin, we run R-LINE for all primary roads (i.e., Interstates, principal/minor arteries, and major collectors) in each CMAQ 12-km grid cell to estimate concentrations at a spatially distributed 1-km grid of 144 receptors. Then, we average this R-LINE grid of 1-km grid resolution within every CMAQ 12-km grid cell to obtain the R-LINE$_{GAVG}$. We then subtract this R-LINE$_{GAVG}$ from the total CMAQ estimate (CMAQ$_{TOT}$) that includes impacts from all sources in each grid cell to remove on-road contributions. Note that CMAQ$_{TOT}$ represents an average value for the entire grid cell. This computed difference then corresponds to an estimate of CMAQ without primary roadway attributable PM$_{2.5}$ at a 12-km resolution. The next step is to linearly interpolate this difference to Census Block Centroids (CBCs). These interpolated values are then added to the primary PM$_{2.5}$ estimates from R-LINE at a CBC resolution (R-LINE$_{CBC}$). Because we are combining models with different biases, there is the potential of negative estimates of CMAQ without primary roadways. If this is the case, we rely only on R-LINE$_{CBC}$ for the hybrid estimate (i.e., we zero-out CMAQ without primary roads). The resulting concentration represents the hybrid PM$_{2.5}$ at the CBCs as shown in Eq 1.

$$HYB_{(PM_{2.5}, NO_2)} = [(CMAQ_{TOT} - R{-}LINE_{GAVG})_{interpolate}] + R{-}LINE_{CBC} \tag{1}$$

The hybrid estimate refines the variability of the coarse CMAQ estimate to capture variability within each CMAQ grid described by R-LINE (i.e., hyperlocal variability of primary PM$_{2.5}$ from mobile emissions occurring within each CMAQ grid). However, note that the hybrid estimate does not refine the variability of the coarse CMAQ estimate arising from the variability of non-mobile emissions within each CMAQ grid. A similar hybrid approach is used for NO$_2$, a pollutant that has been linked to increased mortality in several epidemiological studies [69]. However, in this approach we model NO$_x$ using R-LINE and then use the polynomial method described by Valencia et al. [62] that transforms NO$_x$ to NO$_2$. This is an empirical approach based upon fitting a 4th order polynomial to existing near-road observations across

the continental U.S. Using this polynomial, we can calculate a yield of NO$_2$ to NO$_x$ with which we are able to obtain NO$_2$ estimates.

Regionalize Air quality Model Performance (RAMP) bias correction

Once we have combined CMAQ with R-LINE to create the "Hybrid" estimate, we apply the RAMP method [57]. This method tries to account for the space/time variability in model performance which has been shown to fluctuate significantly across the national domain in other studies [57, 70, 71]. In this study, the main goal of RAMP is to correct the hybrid model biases using measurements. These biases have been shown to be nonhomogeneous and nonlinear [57, 70, 71], and which RAMP can account for.

To achieve this, the first step in the RAMP method consists of selecting the most relevant observation and prediction pairs $(z_k, \tilde{z}_k)$ in time and space at each CMAQ 12-km grid cell $G(\boldsymbol{p})$ with the space/time coordinate p. In this study, we chose to pair daily averaged model/observed values for the entire year. For PM$_{2.5}$, we obtained the 50 closest monitoring sites while for NO$_2$, we obtained the closest 10 sites. The number of stations was chosen to be as spatially specific to correlate with spatial trend while still ensuring a consistent pattern not highly influenced by outliers. Additionally, more sites were necessary for PM$_{2.5}$ given that most of the daily measurements are reported every 3 days as opposed to NO$_2$ sites where most sites provide daily measurements throughout the year.

The second step is to stratify the pairs in 10 equal percentile bins of increasing predicted values $\tilde{z}_k$, and for each bin $\tilde{z}_l$ we calculate the mean $\lambda_1(\tilde{z}_l, G(\boldsymbol{p}))$ of the observed values (see detailed equations in Reyes et al. [57] and Xu et al. [71]). The third step is to create interpolation/extrapolation lines that connect the mean of each bin $\lambda_1(\tilde{z}_l, G(\boldsymbol{p}))$. This piecewise linear function is then used to obtain bias corrected Hybrid (RAMP Hybrid) value $\lambda_1(\tilde{z}_k, G(\boldsymbol{p}))$ for the pertinent Hybrid value. Fig 1C shows this RAMP function and how the mean at each bin changes as a function of the average modeled value bin as well as how the method accounts for the nonlinear performance of the Hybrid model.

In addition to the steps laid out by Reyes et al. [57], we improve on the RAMP method by safeguarding that the slope between each $\lambda_1(\tilde{z}_l, G(\boldsymbol{p}))$ is never negative. In other words, we make sure that the slope is always monotonically increasing [72]. To do so, we first calculate the mean of all observed values z_K pertinent to the specific $G(\boldsymbol{p})$ as $\lambda_{1,M}$. We then compare $\lambda_{1,M}$ with $\lambda_1(\tilde{z}_5, G(\boldsymbol{p}))$ the λ_1 in the 5th decile bin. If $\lambda_{1,M} < \lambda_1(\tilde{z}_5, G(\boldsymbol{p}))$, $\lambda_1(\tilde{z}_5, G(\boldsymbol{p}))$ is set to $\lambda_{1,M}$. In a similar manner, we compare the 5th and the 4th bin making sure that $\lambda_1(\tilde{z}_k, G(\boldsymbol{p})) < \lambda_1(\tilde{z}_{k-1}, G(\boldsymbol{p}))$ and set them equal if necessary. The same algorithm is applied to the other half of the RAMP curve (k = 6 to k = 10). In this case, we compare $\lambda_{1,M}$ to 6th bin and set $\lambda_1(\tilde{z}_6, G(\boldsymbol{p}))$ to $\lambda_{1,M}$ if $\lambda_{1,M} > \lambda_1(\tilde{z}_6, G(\boldsymbol{p}))$. This improvement safeguards the ordinality of estimates of the original model with the same $G(\boldsymbol{p})$.

Health impact evaluation

To evaluate the health impact of air pollution from the model, we combine population data, background health data, and concentration response functions (CRFs) from epidemiological literature that provide the relationship between PM$_{2.5}$ and NO$_2$ to health impacts. For estimating premature mortalities attributable to PM$_{2.5}$, we apply a CRF from a published meta-analysis by Vodonos et al. [73]. In this study, findings show that there was a 1.29% (95% CI 1.09–1.5) increase in all-cause mortality for every 10 μg/m^3 increase of PM$_{2.5}$. For NO$_2$-attributable premature mortalities, we use a CRF associating all-cause mortality to long-term effect of NO$_2$ exposure with a hazard ratio of 1.04 (95%CI 1.02–1.06) per 10 μg/m^3 increase in NO$_2$ [74]. This approach is similar to what was used in a recent study that focused on NO$_2$-related premature mortalities from commercial aviation [75].
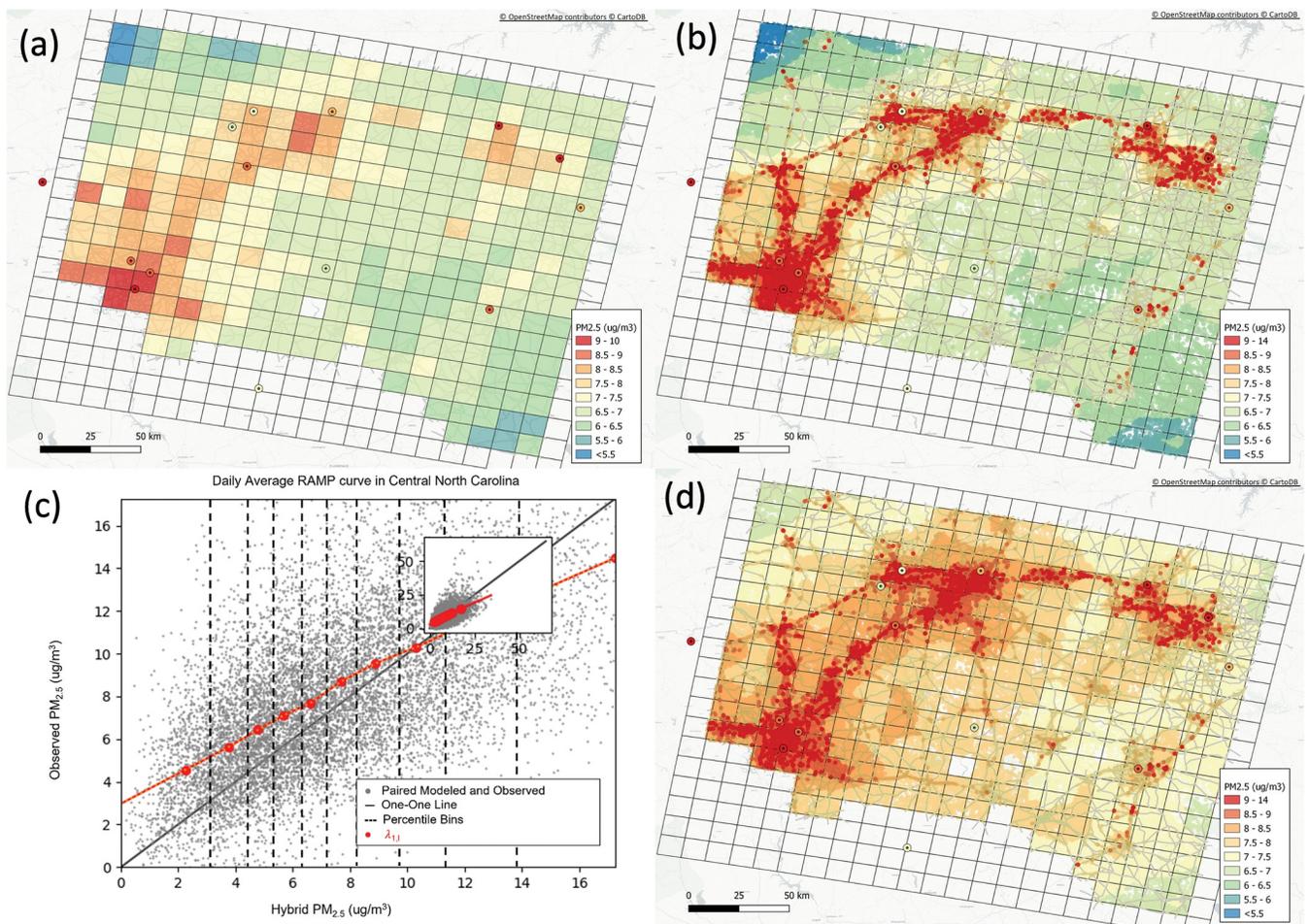
Fig 1. Spatial map of 2016 annual PM$_{2.5}$ model concentrations for (a) CMAQ 12 km x 12 km grid resolution for a subdomain in central North Carolina (b) CMAQ + R-LINE Hybrid, and (d) RAMP Hybrid at census block centroids. Each map includes measurements represented as circles. The gray lines represent the major roads. The gradient bar in figures a, b, and d represents concentration levels in μg/m^3. Also shown is one of the RAMP curves (c) used to correct biases from the Hybrid model. The grey dots represent paired observed and modeled daily PM$_{2.5}$ concentrations for all of 2016 consisting of the 50 closest AQS measurement stations to the centroid of the CMAQ grid cell $G(p)$. The dashed vertical lines represent the 10 equally divided bins used to stratify all the paired data where each bin includes one decile of all the paired points. The solid black line is the one-to-one line between model and observed. The red dots in each bin identify $\lambda_1(\tilde{x}_1, G(p))$ the average of paired observed values within each decile bin. The red dots are linearly interpolated to obtain $\lambda_1(p)$ corresponding to the hybrid modeled data $\tilde{x}(p)$ with $G(p)$.

Population data at census block level were obtained from the U.S. Census Bureau for 2010 and projected to 2016. The data set also included population by race/ethnicity at census block level [76]. The data was split by age with data from the U.S. Census American Community Survey for the year 2016. Given that this dataset was at census tract resolution, we estimated the percentage of the population over 25 years old for each census tract and applied that fraction to each census block inside the census tract. Average county level baseline mortalities from 1999 to 2016 (the most recent available years) were obtained from the U.S. Centers for Disease Control Wide-ranging Online Database for Epidemiologic Research (CDC-WONDER) [77].

Results

Hybrid data fusion

Fig 1 shows the progression of the hybrid data fusion model for 2016 annual PM$_{2.5}$ in the central North Carolina region as an illustration. Fig 1A shows PM$_{2.5}$ concentrations at 460 12-km

grid cells. Fig 1B shows the hybrid combination of CMAQ and R-LINE at 144,276 census block centroids. Results show how the hybrid method captures fine-scale gradients near major roads that CMAQ averages out as shown in Fig 1A. Specifically, hybrid captures elevated concentrations surrounding I-40 from Winston-Salem to Statesville in the northwest and it also captures hotspots from I-95 near Fayetteville in the southeast of the domain. Please refer to Supporting Information (S1 File) for a map of central North Carolina city locations. Fig 1C shows the shape of the RAMP curve used to correct biases from the Hybrid model. From this RAMP curve, we can see that the hybrid model is underpredicting up until ~10 μg/m^3 after which it begins to overpredict concentrations. Thus, the RAMP adjustment will increase low hybrid concentrations and decrease high hybrid concentrations. This is an example of how the RAMP methods consider the nonlinear model performance at varying concentrations of the model. Fig 1D shows the final RAMP hybrid product where hybrid concentrations have been adjusted. The bias correction that stands out the most occurs away from the big metropolitan areas where concentrations were increased by ~1.5 to ~2 μg/m^3 to better represent measurement surrounding the area. A harder to detect correction (that happens on the other side of the RAMP) occurs on the I-285 from Winston-Salem to Lexington where concentrations also increased by ~2 μg/m^3.

The spatial maps from Fig 2 show annual PM$_{2.5}$ for CMAQ, Hybrid, and RAMP Hybrid at census block level focusing on smaller domains in Boston, MA and Chicago, IL. These are also considered the most congested cities in the U.S. according to the Global Traffic Scorecard released by INRIX, a data analytics company that studies traffic globally [78]. These maps reinforce that by combining R-LINE and CMAQ the hybrid results can capture sub-grid scale (within each 12-km grid) variability, which most traditional grid-scale models cannot. For Boston, because we are attempting to capture concentrations for all models, the color scale does not highlight the near-road gradients in the domain. Thus, while concentrations of the 10th and 90th percentile for CMAQ and Hybrid range from 6.9 to 8.9 μg/m^3, the RAMP hybrid concentrations range from 6.1 to 7.2 μg/m^3. This difference in concentration range is due to the RAMP regional correction, where this area was found to be biased high across the board. Please refer to the S1 File for spatial maps of Boston that have a color scale that highlight the near-road gradients for the RAMP method. In Chicago, the spatial maps highlight the near-road concentration gradients where both Hybrid and RAMP Hybrid show better distribution of high and lows. Once again, CMAQ and Hybrid both show significantly higher concentrations than RAMP Hybrid. Concentrations of the 10th and 90th percentile for CMAQ and Hybrid range from 10.5 to 13.9 μg/m^3, while the RAMP hybrid concentrations range from 8.4 to 9.6 μg/m^3.

For annual NO$_2$ shown in Fig 2, Hybrid and RAMP are capturing higher concentrations near roads in Boston. One can see the impact from the Concord Turnpike (US-2) and US-3 coming from the north into downtown Boston from both Hybrid and RAMP hybrid. Additionally, these models show peaks >14 ppb near downtown Boston not captured by CMAQ. In Chicago, the evolution from CMAQ to Hybrid to RAMP Hybrid is drastic. Both CMAQ and Hybrid have higher overall concentrations surrounding downtown. Hybrid estimates significantly high values that we truncated in this area where the 90th percentile concentrations for Hybrid were ~26 ppb. Thus, the range of the color scale only shows the 10th and 90th percentile of CMAQ and RAMP Hybrid. These high values only seen in Hybrid emphasize the need for the RAMP bias correction. The concentration adjustments due to the RAMP method in both these areas also showcase the ability of this method to correct regional and nonlinear biases, a capability that is lacking from other linear regression methods. Additionally, this method keeps the steep near-road gradients (as seen the most in the Chicago area) and captures the within-cell spatial variability of NO$_2$. Please refer to the S1 File for spatial maps of
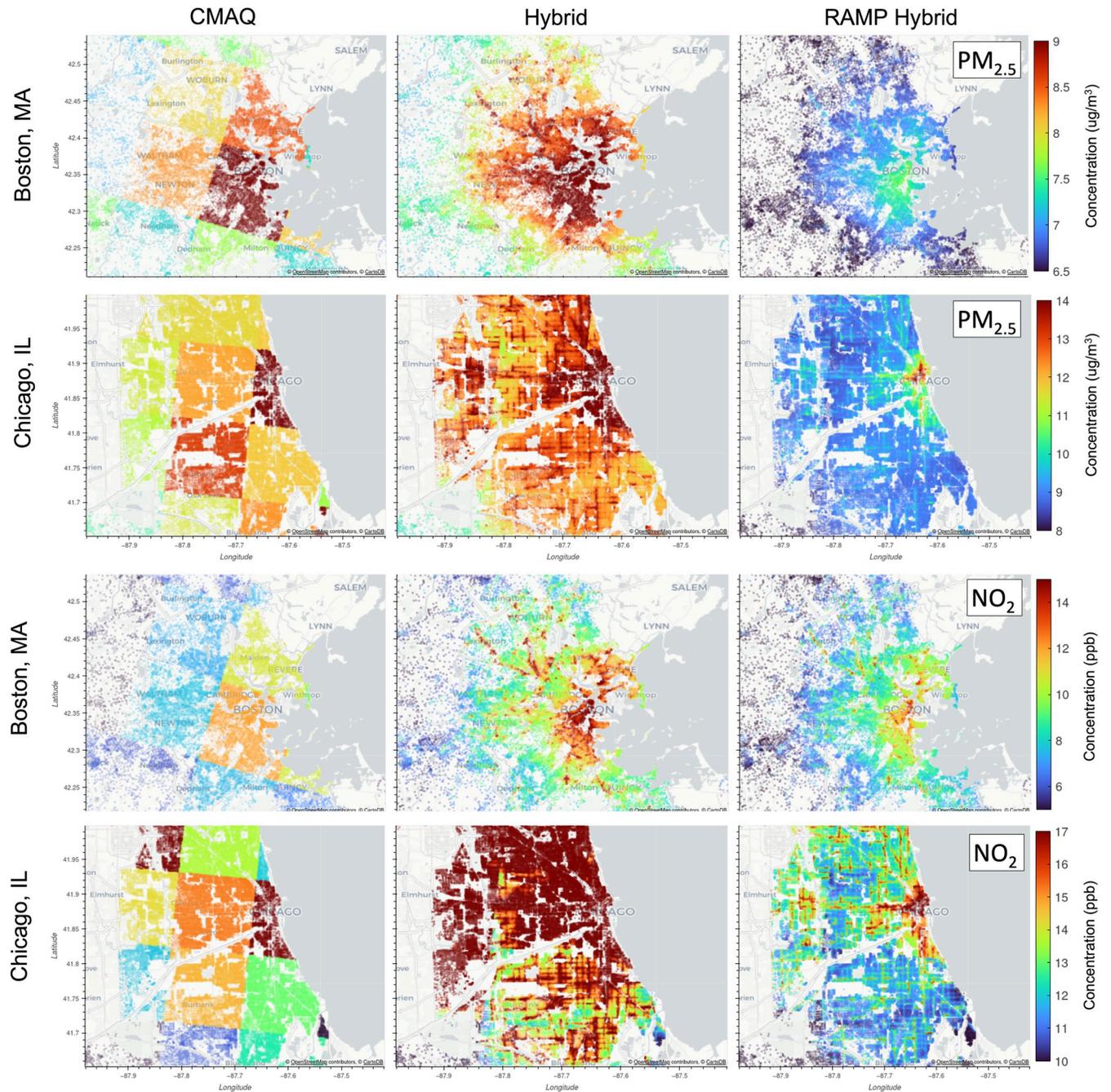
Fig 2. Spatial map of 2016 annual PM$_{2.5}$ and NO$_2$ model concentrations for CMAQ, CMAQ R-LINE Hybrid, and RAMP Hybrid at census block centroids at Boston, Massachusetts and Chicago, Illinois. The gradient bar represents concentration levels in μg/m^3 that range from the smallest 10th percentile of the 3 three models to the highest 90th percentile of the three models for the corresponding domain.

https://doi.org/10.1371/journal.pone.0286406.g002

nine additional U.S. metropolitan areas (e.g., Atlanta, Houston, Los Angeles, New York, Philadelphia, Portland, San Francisco, Seattle, and Washington DC).

Overall, the RAMP Hybrid and Hybrid methods capture variability at significantly finer scales (∼0.01 km^2) than CMAQ (12-km). To our knowledge, only a few studies have estimated PM$_{2.5}$ and NO$_2$ at census block scale across the U.S [39, 79]. These studies used

Table 1. Model performance statistics for annual PM$_{2.5}$ and NO$_2$ for 2016 of CMAQ, Hybrid, and RAMP hybrid evaluated at AQS Sites across the continental United States. [PM$_{2.5}$ (μg/m^3): MO = 7.64, SDO = 2.06; NO$_2$ (ppb): MO = 8.38, SDO = 5.73]. The best performing model metric has been highlighted when possible.

	Model	# Of Sites	FAC2 (%)	ME	SDE	RMSE	MZ	SDZ	Pearson R^2	Spearman ρ^2
PM$_{2.5}$ (μg/m^3)	CMAQ	928	95	**0.02**	2.29	2.29	7.67	2.64	0.30	0.35
	Hybrid		95	0.29	2.30	2.32	7.93	2.71	0.32	0.36
	RAMP Hybrid		99	0.08	**1.44**	**1.45**	7.72	1.64	**0.51**	**0.48**
NO$_2$ (ppb)	CMAQ	444	85	-1.18	3.47	3.66	7.20	4.82	0.63	0.72
	Hybrid		85	2.61	5.36	6.31	10.99	9.04	0.62	0.76
	RAMP Hybrid		95	**0.05**	**2.51**	**2.51**	8.43	4.89	**0.81**	**0.84**

Note: FAC2 is percentage of modeled values between a factor of 2 of the observations; ME is the Mean of the Error; SDE is the Standard Deviation of the Error; RMSE is the Root Mean Squared Error; SDZ is the standard deviation of the modeled value; SDO is the standard deviation of the observed value; MZ is the mean of the modeled value; MO is the mean of the observed value

https://doi.org/10.1371/journal.pone.0286406.t001

statistical models and did not perform a health impact assessment. Please refer to S1 File, where we have quantified and compared to what resolution these methods capture within-cell variability.

As shown in Table 1, model performance between CMAQ and Hybrid is comparable. For PM$_{2.5}$, the Pearson R^2 is slightly higher for the Hybrid (0.32) method than CMAQ (0.30), while for NO$_2$ CMAQ (0.63) is slightly higher than Hybrid (0.62). The mean error (ME) is higher for Hybrid for both pollutants, but Hybrid also shows that it captures higher variability having a higher standard deviation of the model values (SDZ). RAMP Hybrid shows better model performance for nearly every metric for both pollutants. Through bias correction RAMP Hybrid increases Pearson R^2 by up to ~0.2 for both pollutants. One of the biggest successes from the RAMP Hybrid method is correcting the root squared mean error (RMSE) where CMAQ shows an RMSE of 2.29 and hybrid shows RMSE of 2.32 while RAMP Hybrid shows RMSE of 1.45 for PM$_{2.5}$ and for NO$_2$, CMAQ shows an RMSE of 3.66 and hybrid shows RMSE of 6.31 while RAMP Hybrid shows RMSE of 2.51.

Health risk assessment

To estimate the difference between RAMP hybrid and CMAQ in premature mortality attributable to PM$_{2.5}$ and NO$_2$, we aggregated all-cause premature mortality for each county in Fig 3 in space. Results from this figure show that for PM$_{2.5}$ the difference in premature mortality varies regionally. For PM$_{2.5}$, several regions in the east of the U.S. estimate overall lower premature mortality for RAMP Hybrid than CMAQ while regions in the west show higher premature mortalities for RAMP Hybrid when compared to CMAQ. These differences between east and west are mainly due to biases in the CMAQ model (see S4 Fig in Model Performance section of S1 File for spatial maps of error). The hybrid method alone usually increases concentrations in urban areas near roads. So, the bulk of the regional differences are due to the RAMP correction rather than the hybrid method. For NO$_2$, RAMP Hybrid estimates show lower premature mortalities than CMAQ mostly in counties from significantly large metropolitan areas (e.g., Boston, Chicago, Dallas, Houston, Miami, New York, Orlando, Pittsburg, Seattle, Tampa, and Portland).

The bar plots in Fig 3 shows the 40 counties that had the most difference in premature mortalities between RAMP Hybrid and CMAQ. In this figure, the red bars show the counties where RAMP hybrid estimated more premature mortalities when compared to CMAQ, and the blue bars represent the counties where RAMP Hybrid showed less premature mortalities when compared to CMAQ. For PM$_{2.5}$, 15 out of the 20 counties that showed more premature
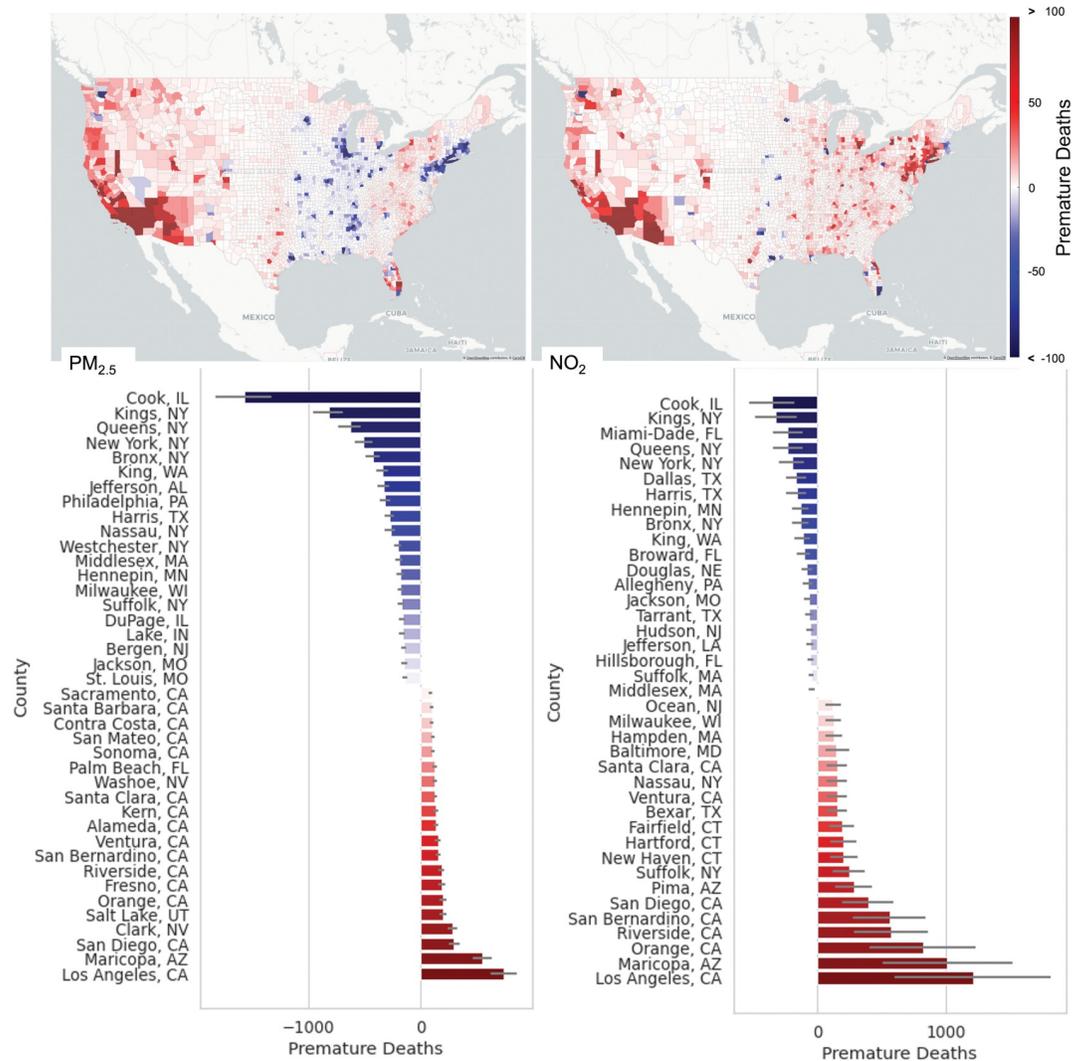
Fig 3. Premature mortality difference between RAMP Hybrid and CMAQ attributable to PM$_{2.5}$ (left) and NO$_2$ (right) aggregated by county. The Top Panel Show spatial differences across the continental United States. In the Bottom Panel, the blue bars represent the top 20 counties where RAMP Hybrid shows less premature mortality than CMAQ. The red bars represent the top 20 counties where RAMP Hybrid shows more premature mortality than CMAQ. The lines in each bar correspond to the 95% confidence intervals.

https://doi.org/10.1371/journal.pone.0286406.g003

mortalities in RAMP Hybrid when compared to CMAQ were in CA. Los Angeles County, CA was at the top where RAMP Hybrid showed 736 [622 856] more premature mortalities than CMAQ. 10 out of 20 counties where CMAQ showed the most premature mortalities when compared to RAMP Hybrid were in the Northeast. Cook County, IL (not in the northeast) was at the top where CMAQ estimated 1,574 [1,330 1,830] more premature mortalities than RAMP Hybrid. For NO$_2$, regional patterns in difference in premature mortalities are not as evident. The counties where RAMP Hybrid showed the most premature mortalities when compared to CMAQ are almost evenly split between the east coast and the west coast. Los Angeles county still showed the most difference with 1,212 [95% CI, 605–1,827]. On the opposite side of the graph, CMAQ estimated that Cook County showed 354 [95% CI, 177–532] more premature mortalities than RAMP Hybrid. Additionally, most of the counties in the top

20 where CMAQ showed more premature mortalities were major metropolitan areas with dense populations (> 800 people/km^2) and high road density.

The sum of differences in premature mortalities attributable to PM$_{2.5}$ and NO$_2$ between RAMP Hybrid and CMAQ show different findings depending on the pollutant and the distance from road. For PM$_{2.5}$, CMAQ estimated 18,079 [95% CI, 15,276–21,023] more premature deaths than RAMP Hybrid (or 7% more premature mortalities than RAMP Hybrid) in some regions, and RAMP Hybrid estimated 15,310 [95% CI, 12,936–17,802] more premature mortalities than CMAQ (or 6% more premature mortalities than CMAQ) in other parts of the U.S. If we aggregate these values, we can estimate that overall CMAQ estimated 2,769 [95% CI, 2,340–3,221] more premature mortalities than RAMP hybrid. This corresponds to an overall ~1% decrease of estimated premature mortalities when estimating premature mortality with RAMP Hybrid when compared to CMAQ. This would seem counterintuitive, but at the core of the Hybrid method, we essentially displace concentrations in the grid away from roads to near the roads. This causes an increase in concentrations near roads while causing a decrease in concentrations away from roads. In the case of PM$_{2.5}$, these competing effects result in a modest net difference. For NO$_2$, CMAQ estimated 6,301 [95% CI, 3,150–9,451] more premature deaths than RAMP Hybrid in some regions, while RAMP Hybrid estimated 28,876 [95% CI, 14,438–43,315] more premature deaths than CMAQ in other regions. The net difference between these estimates correspond to RAMP Hybrid estimating 22,576 [95% CI, 11,288–33,864] more premature mortalities attributable to NO$_2$ than CMAQ. This corresponds to an overall ~19% increase in estimated premature mortalities due to NO$_2$ using RAMP Hybrid when compared to CMAQ.

To assess the effect of only fine-scale variability on risk assessment, we re-gridded our RAMP Hybrid estimates to match CMAQ's 12-km resolution. This type of analysis comparing RAMP Hybrid vs RAMP Hybrid at 12-km resolution (RAMP Hybrid 12-km) shows a different distribution of premature mortalities. Across the U.S. for PM$_{2.5}$, RAMP Hybrid 12-km estimated 3,203 [95% CI, 2,706–3,725] more premature deaths than RAMP Hybrid, while RAMP Hybrid estimated 1,975 [95% CI, 1,669–2,297] more premature deaths than RAMP Hybrid 12-km (or ~1% more premature mortality than the coarser model). Again, this might seem counterintuitive, but because on-road primary PM$_{2.5}$ corresponds to a small fraction of all the ambient PM$_{2.5}$, the difference between coarse scale and fine-scale RAMP Hybrid is modest. If we aggregate these values, we can estimate that the coarser RAMP Hybrid estimated 1,228 [95% CI, 1,037–1,428] more premature mortalities than the finer scale RAMP hybrid, but these changes represent less than 1% of the total premature mortality calculated. This implies that the difference between fine-scale and coarse scale RAMP Hybrid are minor. For NO$_2$, RAMP Hybrid 12-km estimated 10,065 [95% CI, 5,033–15,098] more premature mortalities than the fine-scale RAMP Hybrid. This is equivalent to 7% more premature mortalities estimated by the coarser resolution RAMP Hybrid. Alternatively, the fine-scale RAMP Hybrid estimated 4,854 [95% CI, 2,427–7,281] more premature deaths than RAMP Hybrid 12-km. Overall, the coarser scale model estimated 5,211 [95% CI, 2,606–7,817] more premature deaths attributable to NO$_2$ than fine-scale RAMP Hybrid. Furthermore, focusing on the premature deaths near major roads, we find that most of the excess premature deaths shown by the RAMP Hybrid model are within 500 m from the road (92% for PM$_{2.5}$ and 98% for NO$_2$). At distances of less than 125 m, the net premature differences between coarse and fine-scale RAMP hybrid flip, where the fine-scale RAMP hybrid estimates more premature mortality than the coarse RAMP hybrid for both PM$_{2.5}$ and NO$_2$. This is the outcome that is expected near roads. Please refer to S1 File for detailed tables of differences in premature mortality between the methods at varying distances from the road.
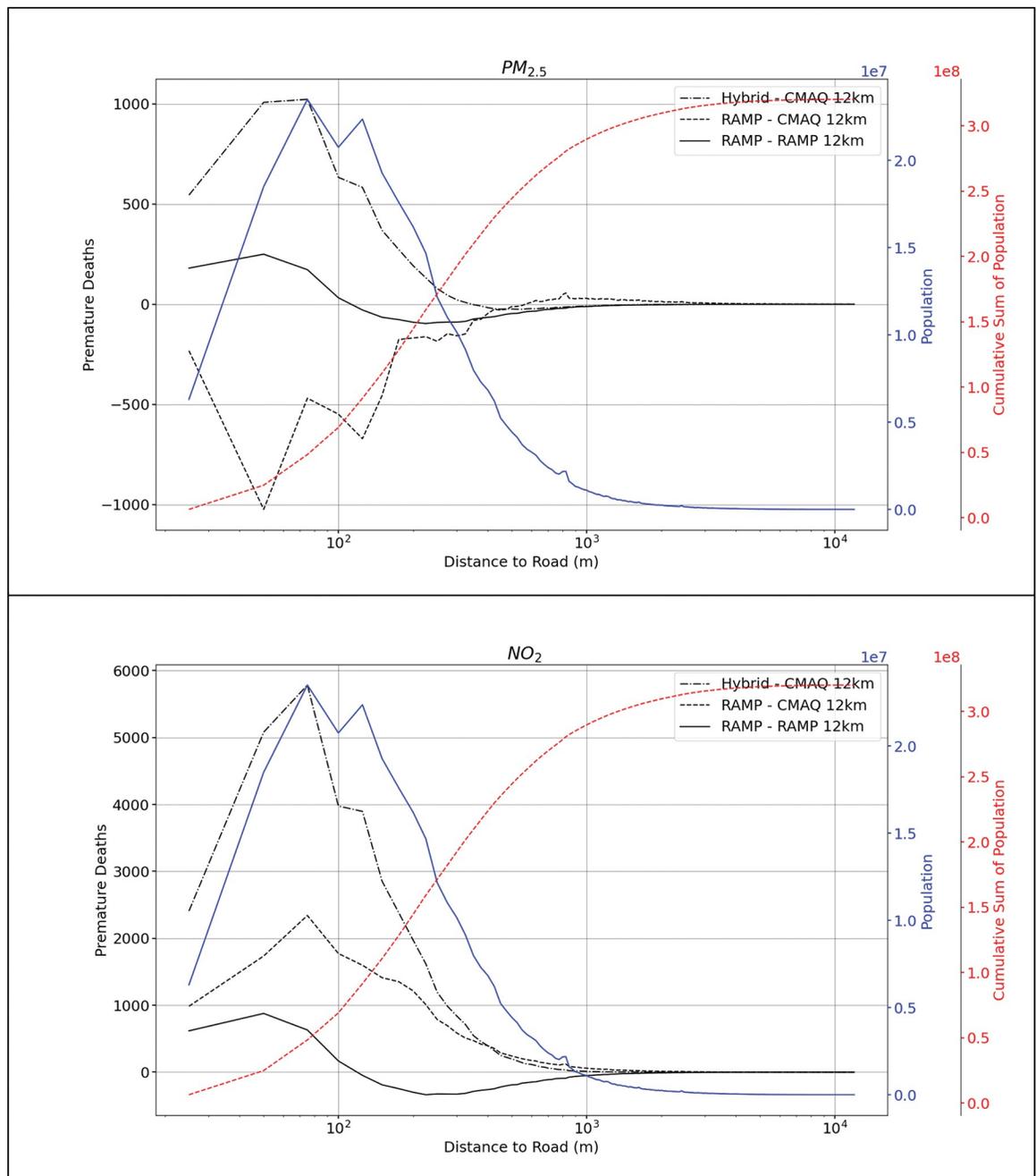
Fig 4. Net difference in premature mortality for PM$_{2.5}$ and NO$_2$ between models with varying resolution vs distance from primary road. Net premature mortality was aggregated at every 25 m from a primary road. The blue line represents the population aggregated at every 25 m from a primary road and the red line corresponds to the cumulative sum of the population.

Fig 4 compares the difference of premature mortality attributable to PM$_{2.5}$ and NO$_2$ between our models (e.g., Hybrid, CMAQ, and RAMP Hybrid) across the U.S. as a function of distance from road. The dashed-dot curves show net premature mortality difference between Hybrid-only minus CMAQ. The dashed curves show net premature mortality difference between RAMP Hybrid minus CMAQ. We have also added a blue line that represents the population aggregated at every 25 m from the road and a red line corresponds to the cumulative

sum of the population. Most of the differences in premature mortality occur at census blocks that are less than 500 m from the road, as mentioned previously. Additionally, we can see that RAMP correction from Hybrid vs CMAQ (dashed-dot) to RAMP Hybrid vs CMAQ (dashed) causes a downward shift of the premature mortality difference for both pollutants. PM$_{2.5}$ shows negative net premature mortality at census blocks less than 500 m from a major road which corresponds to more premature deaths estimated with CMAQ compared to RAMP Hybrid. For NO$_2$, there is also less premature deaths difference between Hybrid and RAMP Hybrid, but the net difference is still positive. For both pollutants, these changes are mostly driven by RAMP rather than the effect of finer resolution due to hybrid as can be seen from the solid curve (RAMP Hybrid vs coarse RAMP Hybrid) comparison which shows smaller net differences and shows only the effect of resolution. Additionally, the solid curve shows positive net differences at census blocks closer to the road between the fine-scale and the coarse-scale RAMP hybrid which correspond to more premature mortality estimated by the fine-scale RAMP-hybrid near the road. This positive net difference occurs for ~80 million people for both PM$_{2.5}$ and NO$_2$, as seen from the cumulative population curve. Additionally, the positive net difference flips (to negative) at ~125 m from major roads. Thus, at census blocks at a distance > 125 m from major roads (where most of the population resides), we estimate higher premature mortality with coarse-scale RAMP Hybrid. For concentration differences between models instead of premature mortality differences, please refer to S1 File.

Exposure inequity

Using the RAMP hybrid method, we calculated population weighted exposure for the entire population in the continental US, and we also stratified population-weighted exposure by race. With these, we calculated an exposure inequity ratio (EIR) which corresponds to the population weighted exposure of the Minority group (corresponds to all the non-White population) over the population weighted exposure of the White group for the continental US. The EIR ratio allows us to quantify to what degree Minorities are affected by PM$_{2.5}$ and NO$_2$ pollution when compared to their White counterparts. The EIR shows Minorities are exposed to 11% more PM$_{2.5}$ and 39% more NO$_2$ than the White population. See S1 File for details. To our knowledge this is the first time that this ratio has been applied using census block level resolution from a hybrid model that incorporates a CTM, dispersion model, and bias correction through RAMP.

Exposure to PM$_{2.5}$ and NO$_2$ varies spatially. Thus, we aggregated population weighted exposure by county across the U.S (see S1 File for map). For PM$_{2.5}$, the highest concentrations are clustered in California with levels of ~12 μg/m^3 or higher. These concentrations are much higher than in the rest of the country where concentrations gradually change across space with a maximum of about ~9 μg/m^3 in the Midwest. Similarly, for NO$_2$, the highest concentrations are predicted in southern California, Colorado (specifically the Denver area), and New York/New Jersey with concentrations that are at least 15 ppb, while in the rest of the country concentrations max out at ~10 ppb in other big cities or across sizeable highways. To evaluate exposure inequity, we must determine if the population is distributed across the country in a stratified way across racial groups, since that could lead to inequity when large proportions of Minorities reside in areas with high air pollution. Analyzing the proportion of Minority groups in space, it becomes apparent that the Asian population and the Hispanic population are clustered in California where we saw higher concentrations of PM$_{2.5}$ and NO$_2$. This disproportionate distribution of Minorities residing in areas with high air pollution also extends to the Black population near the New York City area when considering NO$_2$. The combination of localized minorities and localized high air pollution begins to show an expected geographical inequity

Table 2. Distribution of exposure inequity ratio estimated at county level, census tract level, and census block group level for PM$_{2.5}$ and NO$_2$.

Pollutant	Metric	County	Census Tract	Census Block Group
PM$_{2.5}$	Count	3,107	71,916	215,071
	Mean	1.0074	1.0010	1.0006
	STD	0.0146	0.0059	0.0049
	50%	1.0044	1.0005	1.0002
	95%	1.0324	1.0088	1.0067
	99%	1.0562	1.0200	1.0162
	99.9%	1.1109	1.0493	1.0400
NO$_2$	Count	3,107	71,916	215,071
	Mean	1.0669	1.0104	1.0060
	STD	0.0923	0.0494	0.0410
	50%	1.0502	1.0038	1.0015
	95%	1.2358	1.0712	1.0582
	99%	1.3544	1.1659	1.1398
	99.9%	1.6124	1.4048	1.3566

https://doi.org/10.1371/journal.pone.0286406.t002

nationwide because of the distribution of these populations across the nation. Please refer to S1 File for a spatial map of the proportion of minorities by county.

Using the hybrid RAMP model, we can explore EIRs at scales even finer than the county level. Table 2 shows the distribution of EIRs for PM$_{2.5}$ and NO$_2$ at County, Census Tract, and Census Block Group levels. As expected, when looking at the mean of EIRs overall, the finer the geographic unit, the smaller the exposure inequity ratio is for both PM$_{2.5}$ and NO$_2$. This is expected given that as we consider smaller geographic groupings, we expect less changes in population proportions or in air concentrations. Nonetheless, there is still exposure inequity in some counties, census tracts and census block groups as shown in the higher percentiles of Table 2. EIRs can be as high as 1.05 for PM$_{2.5}$ and 1.4 for NO$_2$ at some census tracts and 1.04 for PM$_{2.5}$ and 1.36 for NO$_2$ in some census block groups.

Zooming into the effect of exposure inequity in a particular county, one can focus on counties with highest EIR like Madera, CA. This county has the highest EIR for both PM$_{2.5}$ and NO$_2$ among counties with a population greater than 100,000 (see S1 File for the Table listing the top 10 counties by EIR). As shown in the top panel of Fig 5, Madera County has strong exposure inequity patterns that explain within county EIR. Madera County has a high EIR of 1.2 for PM$_{2.5}$ and 1.44 for NO$_2$, at the county aggregation level, but also has high EIR within the county at the census tract and census block group aggregation levels. Focusing on PM$_{2.5}$, the two census block groups that show the highest EIR and the lowest EIR in the middle of Madera County can help explain the effect of air pollution and distribution of the Minority population on EIR. The bottom panel of Fig 5 shows that for PM$_{2.5}$, there are higher concentrations near the city of Madera. Furthermore, even though there is a significant amount of the population residing away from the city, the percent of Minority map shows that this population is predominantly White. Thus, when looking at the two census block groups mentioned previously, the census block group in the north with Minorities closer to Madera city has an EIR greater than 1 while the census block group in the south with a higher proportion of the White population away from Madera city has an EIR less than 1. For NO$_2$, there is a similar pattern for these two census block groups. The magnitude of the EIR is larger and the effect of NO$_2$ is more localized and closer to the roads as is shown in the bottom panel of Fig 5, where the highest concentrations follow Highway 99.
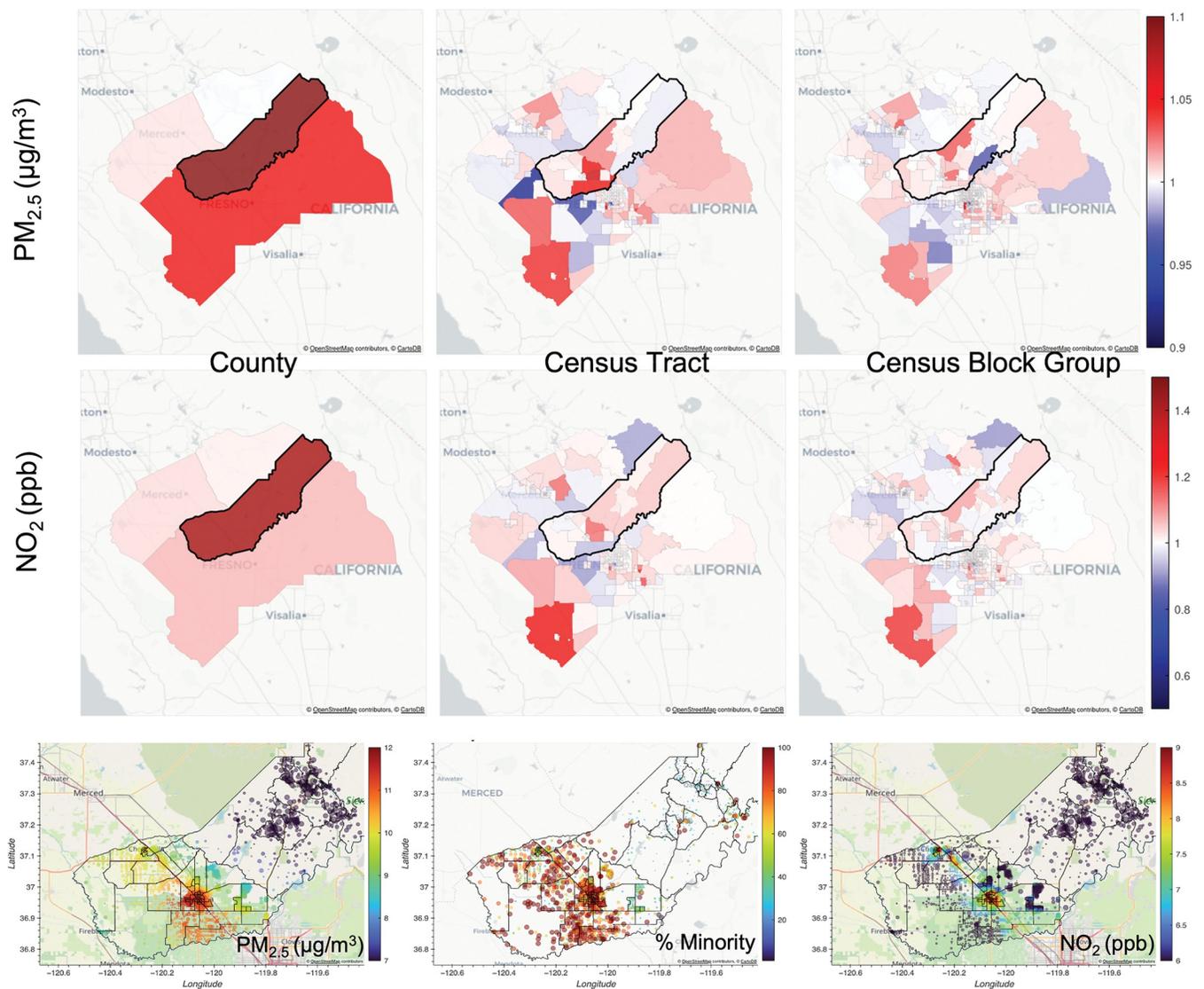
Fig 5. Exposure inequity ratio (EIR) at Madera County, CA. Top panel shows Exposure Inequity Ratio (EIR) at Madera County, CA (highlighted) and surrounding counties for PM$_{2.5}$ and NO$_2$ at County (left), Census Tract (middle), and Census Block Group level (right). Bottom panel shows RAMP Hybrid concentration as circles for PM$_{2.5}$ (left) and NO$_2$ (right) at census block centroids where the size of the census block centroid is proportional to population, as well as the percent of Minority population (middle) at census block centroids where size of census block is proportional to percent Minority at Madera County, CA.

https://doi.org/10.1371/journal.pone.0286406.g005

To follow the near-road effect nationwide, we have created Fig 6 which shows the population-weighted exposure using RAMP aggregated for every census block at every 10m from primary roads across the continental U.S. As mentioned previously, we can see how the air pollution gradients differ between PM$_{2.5}$ and NO$_2$ near road, where NO$_2$ has a steeper decline of concentration when moving away from the road. This figure also shows that not only are Minorities exposed to higher air pollution than their White counterparts in the corridor very near (i.e., within 10 m of) roads, but this exposure inequity persists even at distance of 100m to 1km of roads, where Minorities can be exposed to up to 15% more PM$_{2.5}$ than the White population and up to 35% more NO$_2$ than the White population (bottom panel of Fig 6). This shows that even when controlling for distance to road, the EIR is greater than 1 for hundreds
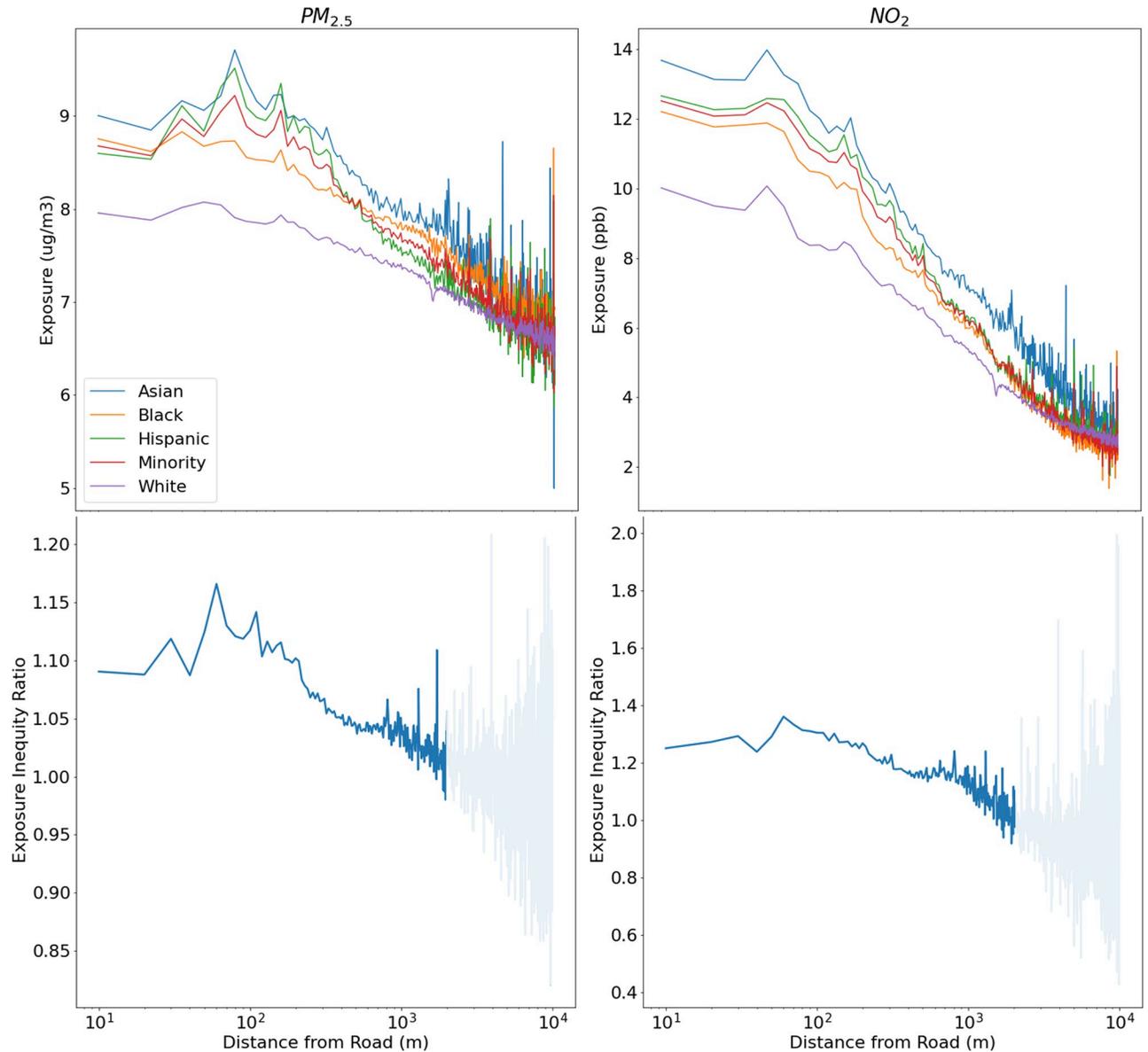
Fig 6. Population-Weighted Exposure using RAMP (top) for PM$_{2.5}$ (left) and NO$_2$ (right) and Exposure Inequity ratio (bottom) aggregated at 10m from major roads across the continental U.S. In the bottom two frames, EIRs has been blurred at distances greater than 2 km from the road to convey that there is high noise in the EIR results.

of meters away from roads, which indicates that exposure inequity is pervasive and entrenched across the nation.

Discussion and conclusions

Health assessment comparison with previous studies

Across the continental U.S., we estimate the total burden of premature mortalities attributable to all sources of ambient PM$_{2.5}$ (anthropogenic and nonanthropogenic) ranges between 223,506 and 307,577 using the RAMP Hybrid approach (225,846 and 310,797 for CMAQ). These premature mortality estimates fall within those of other studies. For example, Davidson

et al. [30] estimate between 100,000 and 220,000 premature mortalities were associated with mobile and non-mobile sources of PM$_{2.5}$ in 2011. Tessum et al. [80] estimated that 102,000 premature mortalities were associated with anthropogenic PM$_{2.5}$ in 2015. Dedoussi et al. [31] estimated that the U.S. combustion emissions of PM$_{2.5}$ caused premature mortalities between 62,400 and 104,200 in 2011, and 49,300 and 82,900 in 2018. There are several differences between our study and the previously mentioned studies. These differences include the air quality model used, model resolution, source types considered, health incidence data resolution, causes of mortality (all cause vs specific causes), and, more importantly, varying CRFs.

In our study, the health impact estimates were based on premature mortality estimates from the most updated PM$_{2.5}$ mortality CRF from a recent meta-analysis [73]. This CRF was based on newer PM$_{2.5}$ studies conducted at lower and at higher concentrations of PM$_{2.5}$ than previous studies. This CRF reported higher slopes at low concentrations which estimated higher mortality. Furthermore, a study by Burnett et al. [81], compared their earlier integrated exposure–response (IER) function [82] to a newer IER including newer studies. Their findings suggest that the newer IER produces similar results to ours (~210,000 all-cause premature mortalities attributable to PM$_{2.5}$ across the U.S.). Burnett et al. compared the updated estimates with those of their older IER [82] and found that the updated CRF estimates were 2.22 times larger than with the older IER. They also compared the updated IER that considered all-cause mortality to the updated IER that considered only five separate causes. The IER that considered all-causes estimated 1.75 times more premature mortality than the one with five separate causes. This study [81] emphasizes a potential underestimation of premature mortality when using older CRFs or not considering all-cause mortality.

Providing a more even comparison, the Vodonos et al. [34] study which applies a similar CRF as our study estimated that premature mortalities would fall by 104,786 [57,16 135,726] attributable to a reduction of 40% of PM$_{2.5}$. If we estimate the proportion of total deaths without reduction for the Vodonos et al. study, this is equivalent to 261,965 [142,540 339,315] premature mortalities for 2015 which compares well to our estimate of 264,516 [223,506 307,577] for RAMP Hybrid in 2016. This last estimate provides us with the most confidence in the magnitude of the effects estimated from our study for PM$_{2.5}$.

For NO$_2$, most studies have only estimated the total burden of diseases in smaller domains (e.g., Hong Kong [83], San Francisco Bay Area [21]) and to our knowledge no studies have performed a health risk assessment based upon premature mortality due to NO$_2$ across the U.S. This might be in part due to the active discussion on the independent causal relationship between long-term NO$_2$ and mortality. Nonetheless, meta-analysis has determined that there are likely causal relationships for long-term exposure to NO$_2$ and premature mortality [74, 84–86]. Overall, in our current study, we estimate 138,550 [69,275 207,826] premature mortalities attributable to NO$_2$ using RAMP Hybrid compared to 115,975 [57,987 173,961] using CMAQ.

RAMP hybrid data fusion findings

We correct biases in the model through the RAMP approach and overall, our findings show that the difference in premature mortality between CMAQ at 12-km resolution and RAMP Hybrid at census blocks are mostly due to bias correction. When we compare RAMP Hybrid at fine-scale resolution vs RAMP hybrid 12-km, we find that the difference in premature mortality is <1%, suggesting that at a 12-km resolution annual PM$_{2.5}$ can be captured adequately. Nonetheless, we do see a noticeable increase of exposure near the road. For NO$_2$, given that on-road emission contributions are much higher than PM$_{2.5}$ to the total anthropogenic emissions, the comparison between RAMP at fine-scale resolution and coarse-scale resolution shows bigger differences. The RAMP Hybrid shows around ~4% more premature mortality

attributable to NO$_2$ across the U.S. than RAMP Hybrid 12-km. RAMP Hybrid 12-km estimates ~7% more premature deaths than the fine-scale RAMP Hybrid at all distances from road. So, the overall net differences in premature mortality for NO$_2$ is higher for the coarse RAMP Hybrid (~4%) than at with the fine-scale RAMP Hybrid. This finding is supported by Batterman et al.' s (2014) Detroit, MI fine-scale study. In this local scale study, they conclude that near-road dispersion modeling at coarser resolution than census block (e.g., census tract and ZIP code levels) have the tendency to overestimate average exposures because at this scale the range of concentrations would be compressed, a smaller amount of individuals reside "very near major roads", most exposures would be misclassified, and higher concentrations near roads would be excluded. It is of note that the Batterman et al. (2014) study compares R-LINE exposure (population-weighted) concentrations of NOx at census blocks to R-LINE at census tracts. Our analysis compares RAMP Hybrid premature mortality attributable to PM$_{2.5}$ and NO$_2$ at 12-km resolution to RAMP Hybrid at census block. However, our overall findings would not change significantly when comparing RAMP Hybrid premature mortality at census block to RAMP Hybrid at census tract. The magnitude of overestimation from the coarser scale averaged grid results would decrease by 5% for NO$_2$ and 10% for PM$_{2.5}$ (see S1 File for details). Another significant difference highlighted here is that the Batterman et al.'s study focuses only on NO$_x$ concentrations/exposures from on-road emissions and does not perform a health risk assessment. Thus, their conclusions are based on population-weighted concentrations.

In general, our findings show that the RAMP Hybrid method can provide reliable estimates for both PM$_{2.5}$ and NO$_2$ that capture the near-road gradient in urban environments. These estimates produce reliable exposure estimates that can reduce exposure misclassification and can ultimately provide reliable inputs to individual-level exposure models which will improve risk assessment for epidemiological studies.

Exposure inequity comparison with previous studies

We have shown how a coarse model (12-km resolution) can overestimate NO$_2$ and PM$_{2.5}$ concentration and mortality associated with these pollutants away from the road while underestimating these near major roads when compared to the fine-scale RAMP hybrid method. This phenomenon affects a significant amount of the population that are exposed to these pollutants and can only be captured with the fine-scale RAMP model. This has implications regarding exposure inequity given that, as mentioned previously, a significant amount of the Minority population resides near major roads. In fact, the proportion of Minorities that live near (~200 m) major roads is higher than the proportion of the White population (See S1 File). Other studies have shown that the Minorities are exposed to more PM$_{2.5}$ and NO$_2$ than their White counterparts [13–17, 20, 22, 87]. For example, a study by Clark et al. (2014) suggested that Minorities were exposed to 38% more NO$_2$ than their White counterparts using estimates at census tract for 2006. A more recent study that estimates several pollutants at census blocks from 1990 to 2010 showed that nationwide in 2010, the Minority populations were exposed to 13% more PM$_{2.5}$ and 54% more NO$_2$ than the White population [13]. A similar study that looked at the effect of data aggregation at different spatial resolutions on exposure inequity reinforced that exposure inequity estimated at state and county scales underestimates disparity when compared to census tract or finer scales [79]. A nationwide study at ZIP code level estimated that in 2016, Minority populations had ~12% higher exposure of PM$_{2.5}$ when compared to the White population [14]. Most of these studies implement national empirical statistical models which do not consider all local sources of air pollution to obtain air pollution exposure. Thus, they overlook the physical/chemical processes and the fine-scale effect of road

sources that can better quantify emission contribution at hyperlocal scales. Using our RAMP hybrid method which considers these processes, we have estimated that Minorities are exposed to 11% more PM$_{2.5}$ and 39% more NO$_2$ than the White population across the US. The hyper local resolution of our model on a nation-wide basis allows us to examine and, for the first time, visualize these inequities meters away from the road. Our method reveals that within 100 meters from major roads, Minority populations can be exposed to up to 15% more PM$_{2.5}$ and up to 35% more NO$_2$ than their White counterparts.

Furthermore, recent studies have shown that fine-scale air pollution estimates of NO$_2$ and PM$_{2.5}$ can be used to identify discriminatory policies that impacted environmental exposure inequity [17]. This study leverages land use regression models to show how redlining in the U. S. disproportionately impacts Minority populations. Likewise, we saw from our results, that the fine-scale RAMP hybrid which leverages the chemical/transport process in the atmosphere showed exposure inequity that persists at the census block level. Moreover, not only are we aware that there is an increased proportion of Minorities near the road, but our results also show exposure inequity continues to persist within close vicinity of the road (~10 m) and up to 100 m from the road. Thus, other factors such as traffic magnitude, weather patterns, and certain policies can contribute to this exposure inequity. Using our novel method will allow us to better estimate, identify, and address the exposure inequity near roads that occurs at the finest scales.

Limitations and future work

The present work is subject to several uncertainties and limitations. Both the dispersion model and the CTMs involve several assumptions/parameters and input data (e.g., emissions inventory, meteorological modeling) that have considerable uncertainties especially at fine-scale resolution. For example, we are using CMAQ v5.2.1 (latest available at the time of this study) which does not include additional aerosol updates with updated secondary aerosol chemistry. As mentioned before, R-LINE tends to overestimate under low dispersion conditions. Thus, both models have different biases that the Hybrid approach does not consider individually. Other studies have tried to treat the biases before combining these models [45]. However, for correcting R-LINE biases, this required detailed collocated near-road monitoring data for NO$_2$, NOx, CO, and SO$_2$ that is not readily available across the continental U.S. [88]. By correcting model biases before combining CMAQ and R-LINE, another study developed a multiplicative hybrid method [45]. In this study, this method was recommended for gaseous pollutants to avoid negative estimates of CMAQ without major roads. We attempted to implement this method for NO$_2$. But since we could not correct biases before applying the hybrid method, this resulted in unrealistic hybrid estimates. Thus, we relied on RAMP to correct biases after combining the models with our additive method as other studies have also done [68]. Our RAMP method is an effective (shows good model performance) and computationally efficient method to account for the biases after the Hybrid method has been applied. Furthermore, for our analysis we are predicting concentrations at census block levels which also has uncertainties. This is the smallest geographic unit where population data is available. However, even at this scale, Batterman et al. (2014) found some effects of exposure misclassification (but less than at census tract and ZIP code level).

For our health impact assessment, we assume that an individual's exposure is mainly representative of census block where they reside and only represents exposure when they are in their census block. A significant amount of the population travels outside of their census block (e.g., commuting for work or school) and this can change their level of exposure. Additionally, when applying our CRF, we use county-based health incidence data. But other studies have

found that using county level health incidence data underestimates the risk when compared to ZIP code level [34]. Future work might explore health outcomes using higher resolution health incidence data. Additionally, given that our modeling approach can estimate short-term exposure, future work might also explore daily exposure of PM$_{2.5}$ and NO$_2$ to expand our analysis to finer temporal scales, and assess acute health impacts.

Finally, an aspect of the RAMP method not considered in this analysis is that it can consider the non-homoscedasticity between modeled and observed data. By considering the non-homoscedastic behavior of model performance we can assess how the error variance changes among predicted values, and ultimately assign a variance (i.e., uncertainty) to our predictions. Furthermore, once we apply a RAMP framework that considers the non-linear, non-homogeneous, and non-homoscedastic behavior of model performance, we can further improve model performance through the full RAMP Bayesian Maximum Entropy (BME) data fusion method [57].

Supporting information

S1 File.
(DOCX)

S2 File.
(ZIP)

Acknowledgments

Author Contributions

Conceptualization: Marc Serre, Saravanan Arunachalam.

Data curation: Alejandro Valencia.

Formal analysis: Alejandro Valencia.

Funding acquisition: Saravanan Arunachalam.

Investigation: Alejandro Valencia.

Methodology: Alejandro Valencia, Marc Serre, Saravanan Arunachalam.

Project administration: Saravanan Arunachalam.

Software: Alejandro Valencia.

Supervision: Marc Serre, Saravanan Arunachalam.

Validation: Alejandro Valencia.

Visualization: Alejandro Valencia.

Writing – original draft: Alejandro Valencia.

Writing – review & editing: Marc Serre, Saravanan Arunachalam.

References

1. HEI. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. Health Effects Institute. 2010.

2. Gan WQ, Koehoorn M, Davies HW, Demers PA, Tamburic L, Brauer M. Long-term exposure to traffic-related air pollution and the risk of coronary heart disease hospitalization and mortality. Environ Health Perspect. 2011. https://doi.org/10.1289/ehp.1002511 PMID: 21081301

3. Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Xun WW, Katsouyanni K, et al. Long-term exposure to air pollution and cardiovascular mortality: An analysis of 22 European cohorts. Epidemiology. 2014; 25: 368–378. https://doi.org/10.1097/EDE.0000000000000076 PMID: 24589872

4. Jerrett M, Burnett RT, Beckerman BS, Turner MC, Krewski D, Thurston G, et al. Spatial analysis of air pollution and mortality in California. Am J Respir Crit Care Med. 2013; 188: 593–599. https://doi.org/10.1164/rccm.201303-0609OC PMID: 23805824

5. Urman R, McConnell R, Islam T, Avol EL, Lurmann FW, Vora H, et al. Associations of children's lung function with ambient air pollution: Joint effects of regional and near-roadway pollutants. Thorax. 2014; 69: 540–547. https://doi.org/10.1136/thoraxjnl-2012-203159 PMID: 24253832

6. Dell SD, Jerrett M, Beckerman B, Brook JR, Foty RG, Gilbert NL, et al. Presence of other allergic disease modifies the effect of early childhood traffic-related air pollution exposure on asthma prevalence. Environ Int. 2014. https://doi.org/10.1016/j.envint.2014.01.002 PMID: 24472824

7. McConnell R, Islam T, Shankardass K, Jerrett M, Lurmann F, Gilliland F, et al. Childhood incident asthma and traffic-related air pollution at home and school. Environ Health Perspect. 2010. https://doi.org/10.1289/ehp.0901232 PMID: 20371422

8. Wilhelm M, Ghosh JK, Su J, Cockburn M, Jerrett M, Ritz B. Traffic-related air toxics and term low birth weight in Los Angeles County, California. Environ Health Perspect. 2012. https://doi.org/10.1289/ehp.1103408 PMID: 21835727

9. Jeong C-H, Evans GJ, Healy RM, Jadidian P, Wentzell J, Liggio J, et al. Rapid physical and chemical transformation of traffic-related atmospheric particles near a highway. Atmos Pollut Res. 2015; 6: 662–672. https://doi.org/10.5094/APR.2015.075

10. Karner AA, Eisinger DS, Niemeier DA. Near-roadway air quality: Synthesizing the findings from real-world data. Environ Sci Technol. 2010; 44: 5334–5344. https://doi.org/10.1021/es100008x PMID: 20560612

11. Rowangould GM. A census of the US near-roadway population: Public health and environmental justice considerations. Transportation Research Part D. 2013; 25: 59–67. https://doi.org/10.1016/j.trd.2013.08.003

12. Clark LP, Millet DB, Marshall JD. National Patterns in Environmental Injustice and Inequality: Outdoor NO2 Air Pollution in the United States. PLoS One. 2014; 9: e94431. https://doi.org/10.1371/journal.pone.0094431 PMID: 24736569

13. Liu J, Clark LP, Bechle MJ, Hajat A, Kim SY, Robinson AL, et al. Disparities in Air Pollution Exposure in the United States by Race/Ethnicity and Income, 1990–2010. Environ Health Perspect. 2021;129. https://doi.org/10.1289/EHP8584 PMID: 34908495

14. Jbaily A, Zhou X, Liu J, Lee TH, Kamareddine L, Verguet S, et al. Air pollution exposure disparities across US population and income groups. Nature 2022 601:7892. 2022; 601: 228–233. https://doi.org/10.1038/s41586-021-04190-y PMID: 35022594

15. Goodkind AL, Tessum CW, Coggins JS, Hill JD, Marshall JD. Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. Proceedings of the National Academy of Sciences. 2019; 116: 8775–8780. https://doi.org/10.1073/pnas.1816102116 PMID: 30962364

16. Colmer J, Hardman I, Shimshack J, Voorheis J. Disparities in PM2.5 air pollution in the United States. Science (1979). 2020; 369: 575–578. https://doi.org/10.1126/SCIENCE.AAZ9353/SUPPL_FILE/AAZ9353_COLMER_SM.PDF

17. Lane HM, Morello-Frosch R, Marshall JD, Apte JS. Historical Redlining Is Associated with Present-Day Air Pollution Disparities in U.S. Cities. Environ Sci Technol Lett. 2022; 9: 345–350. https://doi.org/10.1021/acs.estlett.1c01012 PMID: 35434171

18. Kerr GH, Goldberg DL, Anenberg SC. COVID-19 pandemic reveals persistent disparities in nitrogen dioxide pollution. Proc Natl Acad Sci U S A. 2021; 118: e2022409118. https://doi.org/10.1073/pnas.2022409118 PMID: 34285070

19. Chambliss SE, Pinon CPR, Messier KP, LaFranchi B, Upperman CR, Lunden MM, et al. Local- And regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. Proc Natl Acad Sci U S A. 2021; 118: e2109249118. https://doi.org/10.1073/pnas.2109249118 PMID: 34493674

20. Tessum CW, Paolella DA, Chambliss SE, Apte JS, Hill JD, Marshall JD. PM2.5 polluters disproportionately and systemically affect people of color in the United States. Sci Adv. 2021; 7: 4491–4519. https://doi.org/10.1126/SCIADV.ABF4491/SUPPL_FILE/ABF4491_SM.PDF

21. Southerland VA, Anenberg SC, Harris M, Apte J, Hystad P, Donkelaar A van, et al. Assessing the Distribution of Air Pollution Health Risks within Cities: A Neighborhood-Scale Analysis Leveraging High-Resolution Data Sets in the Bay Area, California. Environ Health Perspect. 2021;129. https://doi.org/10.1289/EHP7679 PMID: 33787320

22. Clark LP, Millet DB, Marshall JD. Changes in Transportation-Related Air Pollution Exposures by Race-Ethnicity and Socioeconomic Status: Outdoor Nitrogen Dioxide in the United States in 2000 and 2010. Environ Health Perspect. 2017;125. https://doi.org/10.1289/EHP959 PMID: 28930515

23. Caiazzo F, Ashok A, Waitz IA, Yim SHL, Barrett SRH. Air pollution and early deaths in the United States. Part I: Quantifying the impact of major sectors in 2005. Atmos Environ. 2013; 79: 198–208. https://doi.org/10.1016/j.atmosenv.2013.05.081

24. Fann N, Fulcher CM, Baker K. The recent and future health burden of air pollution apportioned across u.s. sectors. Environ Sci Technol. 2013; 47: 3580–3589. https://doi.org/10.1021/es304831q PMID: 23506413

25. Grabow ML, Spak SN, Holloway T, Brian SS, Mednick AC, Patz JA. Air quality and exercise-related health benefits from reduced car travel in the midwestern United States. Environ Health Perspect. 2012; 120: 68–76. https://doi.org/10.1289/ehp.1103440 PMID: 22049372

26. Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature. 2015; 525: 367–371. https://doi.org/10.1038/nature15371 PMID: 26381985

27. Punger EM, West JJ. The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the USA. Air Qual Atmos Health. 2013; 6: 563–573. https://doi.org/10.1007/s11869-013-0197-8 PMID: 24348882

28. Zawacki M, Baker KR, Phillips S, Davidson K, Wolfe P. Mobile source contributions to ambient ozone and particulate matter in 2025. Atmos Environ. 2018; 188: 129–141. https://doi.org/10.1016/j.atmosenv.2018.04.057 PMID: 30344445

29. Zwack LM, Paciorek CJ, Spengler JD, Levy JI. Modeling spatial patterns of traffic-related air pollutants in complex urban terrain. Environ Health Perspect. 2011; 119: 852–859. https://doi.org/10.1289/ehp.1002519 PMID: 21262596

30. Davidson K, Fann N, Zawacki M, Fulcher C, Baker KR. The recent and future health burden of the US mobile sector apportioned by source. Environmental Research Letters. 2020; 15: 75009.

31. Dedoussi IC, Eastham SD, Monier E, Barrett SRH. Premature mortality related to United States cross-state air pollution. Nature. 2020; 578: 261–265. https://doi.org/10.1038/s41586-020-1983-8 PMID: 32051602

32. Jiang X, hye Yoo E. The importance of spatial resolutions of Community Multiscale Air Quality (CMAQ) models on health impact assessment. 2018; 627: 1528–1543. https://doi.org/10.1016/j.scitotenv.2018.01.228 PMID: 30857114

33. Liu Y, Zhao N, Vanos JK, Cao G. Revisiting the estimations of PM2.5-attributable mortality with advancements in PM2.5 mapping and mortality statistics. Science of The Total Environment. 2019; 666: 499–507. https://doi.org/10.1016/j.scitotenv.2019.02.269 PMID: 30802665

34. Vodonos A, Schwartz J. Estimation of excess mortality due to long-term exposure to PM2.5 in continental United States using a high-spatiotemporal resolution model. Environ Res. 2021; 196: 110904. https://doi.org/10.1016/j.envres.2021.110904 PMID: 33636186

35. Barzyk TM, George BJ, Vette AF, Williams RW, Croghan CW, Stevens CD. Development of a distance-to-roadway proximity metric to compare near-road pollutant levels to a central site monitor. Atmos Environ. 2009; 43: 787–797. https://doi.org/10.1016/j.atmosenv.2008.11.002

36. Hagler GSW, Baldauf RW, Thoma ED, Long TR, Snow RF, Kinsey JS, et al. Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants. Atmos Environ. 2009; 43: 1229–1234. https://doi.org/10.1016/j.atmosenv.2008.11.024

37. Bechle MJ, Millet DB, Marshall JD. National Spatiotemporal Exposure Surface for NO2: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–2010. Environ Sci Technol. 2015; 49: 12297–12305. https://doi.org/10.1021/acs.est.5b02882 PMID: 26397123

38. Novotny E V, Bechle MJ, Millet DB, Marshall JD. National satellite-based land-use regression: NO2 in the United States. Environ Sci Technol. 2011; 45: 4407–4414. https://doi.org/10.1021/es103578x PMID: 21520942

39. Wang Y, Bechle MJ, Kim SY, Adams PJ, Pandis SN, Pope CA, et al. Spatial decomposition analysis of NO2 and PM2.5 air pollution in the United States. Atmos Environ. 2020; 241: 117470. https://doi.org/10.1016/J.ATMOSENV.2020.117470

40. Hankey S, Marshall JD. Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, PM2.5, Particle Size) Using Mobile Monitoring. Environ Sci Technol. 2015; 49: 9194–9202. https://doi.org/10.1021/acs.est.5b01209 PMID: 26134458

41. Reyes JM, Serre ML. An LUR/BME framework to estimate PM2.5 explained by on road mobile and stationary sources. Environ Sci Technol. 2014; 48: 1736–1744. https://doi.org/10.1021/es4040528 PMID: 24387222

42. Di Q, Amini H, Shi L, Kloog I, Silvern R, Kelly J, et al. Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging. Environ Sci Technol. 2019; 54: 1372–1384. https://doi.org/10.1021/ACS.EST.9B03358 PMID: 31851499

43. Achakulwisut P, Brauer M, Hystad P, Anenberg SC. Global, national, and urban burdens of paediatric asthma incidence attributable to ambient NO 2 pollution: estimates from global datasets. Lancet Planet Health. 2019; 3: e166–e178. https://doi.org/10.1016/S2542-5196(19)30046-4 PMID: 30981709

44. Mohegh A, Goldberg D, Achakulwisut P, Anenberg SC. Sensitivity of estimated NO2-attributable pediatric asthma incidence to grid resolution and urbanicity. Environmental Research Letters. 2020; 16: 14019.

45. Bates JT, Pennington AF, Zhai X, Friberg MD, Metcalf F, Darrow L, et al. Application and evaluation of two model fusion approaches to obtain ambient air pollutant concentrations at a fine spatial resolution (250m) in Atlanta. Environmental Modelling and Software. 2018; 109: 182–190. https://doi.org/10.1016/j.envsoft.2018.06.008

46. Beevers SD, Kitwiroon N, Williams ML, Carslaw DC. One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions. Atmos Environ. 2012; 59: 47–58. https://doi.org/10.1016/j.atmosenv.2012.05.034 PMID: 23471172

47. Chang SY, Vizuete W, Serre M, Vennam LP, Omary M, Isakov V, et al. Finely Resolved On-Road PM2.5and Estimated Premature Mortality in Central North Carolina. Risk Analysis. 2017; 37: 2420–2434. https://doi.org/10.1111/risa.12775 PMID: 28244115

48. Lefebvre W, Vercauteren J, Schrooten L, Janssen S, Degraeuwe B, Maenhaut W, et al. Validation of the MIMOSA-AURORA-IFDM model chain for policy support: Modeling concentrations of elemental carbon in Flanders. Atmos Environ. 2011; 45: 6705–6713. https://doi.org/10.1016/j.atmosenv.2011.08.033

49. Scheffe RD, Strum M, Phillips SB, Thurman J, Eyth A, Fudge S, et al. Hybrid modeling approach to estimate exposures of hazardous air pollutants (HAPs) for the National air Toxics Assessment (NATA). Environ Sci Technol. 2016; 50: 12356–12364. https://doi.org/10.1021/acs.est.6b04752 PMID: 27779870

50. Shukla K, Seppanen C, Naess B, Chang C, Cooley D, Maier A, et al. ZIP Code-Level Estimation of Air Quality and Health Risk Due to Particulate Matter Pollution in New York City. Environ Sci Technol. 2022; 56: 7119–7130. https://doi.org/10.1021/acs.est.1c07325 PMID: 35475336

51. Parvez F, Wagstrom K. Impact of regional versus local resolution air quality modeling on particulate matter exposure health impact assessment. Air Qual Atmos Health. 2020; 13: 271–279.

52. Isakov V, Barzyk T, Arunachalam S, Naess B, Seppanen C, Monteiro A, et al. A web-based screening tool for near-port air quality assessments. Environmental Modelling & Software. 2017; 98: 21–34. https://doi.org/10.1016/j.envsoft.2017.09.004 PMID: 29681760

53. US EPA. User's Manual for the Co-Benefits Risk Assessment Health Impacts Screening and Mapping Tool (COBRA) Developed for State and Local Climate and Energy Program. 2021. Available: https://www.epa.gov/system/files/documents/2021-11/cobra-user-manual-nov-2021_4.1_0.pdf

54. Snyder MG, Venkatram A, Heist DK, Perry SG, Petersen WB, Isakov V. RLINE: A line source dispersion model for near-surface releases. Atmos Environ. 2013; 77: 748–756. https://doi.org/10.1016/j.atmosenv.2013.05.074

55. Byun D, Schere KL. Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. Appl Mech Rev. 2006; 59: 51. https://doi.org/10.1115/1.2128636

56. US EPA. CMAQ. Zenodo; 2018. https://doi.org/10.5281/zenodo.1212601

57. Reyes JM, Xu Y, Vizuete W, Serre ML. Regionalized PM2.5 Community Multiscale Air Quality model performance evaluation across a continuous spatiotemporal domain. Atmos Environ. 2017; 148: 258–265. https://doi.org/10.1016/j.atmosenv.2016.10.048 PMID: 28848374

58. US EPA. Air Quality System (AQS). 2020 [cited 11 Mar 2021]. Available: https://www.epa.gov/aqs

59. National Emissions Inventory Collaborative. 2016beta Emissions Modeling Platform. 2019.

60. Chang SY, Vizuete W, Valencia A, Naess B, Isakov V, Palma T, et al. A modeling framework for characterizing near-road air pollutant concentration at community scales. Science of the Total Environment. 2015; 538: 905–921. https://doi.org/10.1016/j.scitotenv.2015.06.139 PMID: 26363146

61. Snyder M, Arunachalam S, Isakov V, Talgo K, Naess B, Valencia A, et al. Creating locally-resolved mobile-source emissions inputs for air quality modeling in support of an exposure study in Detroit, Michigan, USA. Int J Environ Res Public Health. 2014; 11: 12739–12766. https://doi.org/10.3390/ijerph111212739 PMID: 25501000

62. Valencia A, Venkatram A, Heist D, Carruthers D, Arunachalam S. Development and evaluation of the R-LINE model algorithms to account for chemical transformation in the near-road environment. Transp Res D Transp Environ. 2018; 59: 464–477. https://doi.org/10.1016/j.trd.2018.01.028 PMID: 29780271

63. Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Wang W, et al. A description of the Advanced Research WRF version 3. 2008.

64. Otte TL, Pleim JE. The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ modeling system: Updates through MCIPv3.4.1. Geosci Model Dev. 2010; 3: 243–256. https://doi.org/10.5194/gmd-3-243-2010

65. Brashers B, Emery C. The Mesoscale Model Interface Program (MMIF) Version 3.4.1. 773 San Marin Drive Novato, CA 94998; 2019.

66. Isakov V, Irwin JS, Ching J. Using CMAQ for exposure modeling and characterizing the subgrid variability exposure estimates. J Appl Meteorol Climatol. 2007; 46: 1354–1371. https://doi.org/10.1175/JAM2538.1

67. Parvez F, Wagstrom K. Comparing estimates from the R-LINE near road dispersion model using model-derived and observation-derived meteorology. Atmos Pollut Res. 2018; 9: 483–493. https://doi.org/10.1016/j.apr.2017.10.007

68. Oh I, Hwang MK, Bang JH, Yang W, Kim S, Lee K, et al. Comparison of different hybrid modeling methods to estimate intraurban NO2 concentrations. Atmos Environ. 2021; 244: 117907. https://doi.org/10.1016/j.atmosenv.2020.117907

69. Huang S, Li H, Wang M, Qian Y, Steenland K, Caudle WM, et al. Long-term exposure to nitrogen dioxide and mortality: A systematic review and meta-analysis. Science of The Total Environment. 2021; 776: 145968. https://doi.org/10.1016/j.scitotenv.2021.145968 PMID: 33640547

70. Xu Y, Serre ML, Reyes JM, Vizuete W. Impact of temporal upscaling and chemical transport model horizontal resolution on reducing ozone exposure misclassification. Atmos Environ. 2017; 166: 374–382. https://doi.org/10.1016/j.atmosenv.2017.07.033

71. Xu Y, Serre ML, Reyes J, Vizuete W. Bayesian Maximum Entropy Integration of Ozone Observations and Model Predictions: A National Application. Environ Sci Technol. 2016; 50: 4393–4400. https://doi.org/10.1021/acs.est.6b00096 PMID: 26998937

72. Becker JS, West JJ, Serre ML, Vizuete W. Using Regionalized Air Quality Model Performance and Bayesian Maximum Entropy Data Fusion to Map Global Surface Ozone Concentration and Associated Uncertainty. University of North Carolina at Chapel Hill. 2021.

73. Vodonos A, Awad YA, Schwartz J. The concentration-response between long-term PM2.5 exposure and mortality; A meta-regression approach. Environ Res. 2018; 166: 677–689. https://doi.org/10.1016/j.envres.2018.06.021 PMID: 30077140

74. Faustini A, Rapp R, Forastiere F. Nitrogen dioxide and mortality: review and meta-analysis of long-term studies. European Respiratory Journal. 2014; 44: 744–753. https://doi.org/10.1183/09031936.00114713 PMID: 24558178

75. Arter CA, Buonocore JJ, Moniruzzaman C, Yang D, Huang J, Arunachalam S. Air quality and health-related impacts of traditional and alternate jet fuels from airport aircraft operations in the U.S. Environ Int. 2022; 158: 106958. https://doi.org/10.1016/J.ENVINT.2021.106958 PMID: 34710732

76. Manson S, Schroeder J, van Riper D, Kugler T, Ruggles S. IPUMS National Historical Geographic Information System: Version 16.0 [Data set]. IPUMS. 2021.

77. Friede A, Reid JA, Ory HW. CDC WONDER: A comprehensive on-line public health information system of the Centers for Disease Control and Prevention. Am J Public Health. 1993;83. https://doi.org/10.2105/AJPH.83.9.1289 PMID: 8395776

78. INRIX. Global Traffic Scorecard. INRIX Research. 2020.

79. Clark LP, Harris MH, Apte JS, Marshall JD. National and Intraurban Air Pollution Exposure Disparity Estimates in the United States: Impact of Data-Aggregation Spatial Scale. Environ Sci Technol Lett. 2022 [cited 12 Sep 2022]. https://doi.org/10.1021/acs.estlett.2c00403 PMID: 36118958

80. Tessum CW, Apte JS, Goodkind AL, Muller NZ, Mullins KA, Paolella DA, et al. Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure. Proceedings of the National Academy of Sciences. 2019; 116: 6001–6006. https://doi.org/10.1073/pnas.1818859116 PMID: 30858319

81. Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proceedings of the National Academy of Sciences. 2018; 115: 9592–9597. https://doi.org/10.1073/pnas.1803222115 PMID: 30181279

82. Burnett RT, Arden Pope C, Ezzati M, Olives C, Lim SS, Mehta S, et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ Health Perspect. 2014; 122: 397–403. https://doi.org/10.1289/ehp.1307049 PMID: 24518036

83. Hossain MS, Frey HC, Louie PKK, Lau AKH. Combined effects of increased O3 and reduced NO2 concentrations on short-term air pollution health risks in Hong Kong. Environmental Pollution. 2021; 270: 116280. https://doi.org/10.1016/j.envpol.2020.116280 PMID: 33360064

84. COMEAP (Committee on the Medical Effects of Air Pollutants). Associations of long-term average concentrations of nitrogen dioxide with mortality. 2018.

85. Forastiere F, Peters A. Invited Perspective: The NO2 and Mortality Dilemma Solved? Almost There! Environ Health Perspect. 2021; 129: 121304. https://doi.org/10.1289/EHP10286 PMID: 34962423

86. Qian Y, Li H, Rosenberg A, Li Q, Sarnat J, Papatheodorou S, et al. Long-Term Exposure to Low-Level NO2 and Mortality among the Elderly Population in the Southeastern United States. Environ Health Perspect. 2021;129. https://doi.org/10.1289/EHP9044 PMID: 34962424

87. Paolella DA, Tessum CW, Adams PJ, Apte JS, Chambliss S, Hill J, et al. Effect of Model Spatial Resolution on Estimates of Fine Particulate Matter Exposure and Exposure Disparities in the United States. Environ Sci Technol Lett. 2018; 5: 436–441. https://doi.org/10.1021/ACS.ESTLETT.8B00279/SUPPL_FILE/EZ8B00279_SI_001.PDF

88. Zhai X, Russell AG, Sampath P, Mulholland JA, Kim BU, Kim Y, et al. Calibrating R-LINE model results with observational data to develop annual mobile source air pollutant fields at fine spatial resolution: Application in Atlanta. Atmos Environ. 2016; 147: 446–457. https://doi.org/10.1016/j.atmosenv.2016.10.015

89. CartoDB/basemap-styles/LICENSE.md · GitHub. [cited 16 Apr 2023]. Available: https://github.com/CartoDB/basemap-styles/blob/master/LICENSE.md

90. CartoDB. Copyright of the Positron Basemap. [cited 16 Apr 2023]. Available: https://qms.nextgis.com/geoservices/488/