Atmospheric Environment 364 (2026) 121637

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Bias corrections for species- and source-resolved $PM_{2.5}$ chemical transport model simulations using a geographically weighted regression

Ksakousti Skyllakou [a], Carlos S. Hernandez [b], Pablo Garcia Rivera [c], Brian Dinkelacker [c], Julian D. Marshall [d], C. Arden Pope III [e], Allen L. Robinson [f], Spyros N. Pandis [a,g,*], Peter J. Adams [b,h]

[a] Institute of Chemical Engineering Sciences (FORTH/ICE-HT), Patras, Greece
[b] Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
[c] Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
[d] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA
[e] Department of Economics, Brigham Young University, Provo, UT, USA
[f] College of Engineering, Colorado State University, Fort Collins, CO, USA
[g] Department of Chemical Engineering, University of Patras, Patras, Greece
[h] Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA

HIGHLIGHTS

- Novel dataset of $PM_{2.5}$ exposure estimates across the contiguous U.S. at census tract-level resolution.
- Geographically weighted regression predicted and corrected biases in source-resolved chemical transport model predictions.
- Geographically weighted regression generally outperforms ordinary least squares in bias predictions.

ARTICLE INFO

Keywords:
Air quality
Exposure estimates
Exposure assessments

ABSTRACT

Chemical transport models (CTMs) can provide information — albeit with biases — that is lacking in observations, including $PM_{2.5}$ composition and source. Correcting biases in simulated $PM_{2.5}$ could facilitate their use in exposure assessment, environmental epidemiology, environmental justice analyses, and other applications. Here we develop a novel set of species- and source-resolved $PM_{2.5}$ estimates across the contiguous United States for 2001 and 2010. We use geographically weighted regressions (GWR) to predict and then correct for the bias in CTM $PM_{2.5}$ concentration predictions for five chemical species (elemental carbon, organic aerosol, ammonium, nitrate, and sulfate). The GWR models are trained using speciated measurements, empirical $PM_{2.5}$ exposure estimates, CTM predictions, and other geographic information. A 10-fold cross-validation shows minimal bias across each simulated $PM_{2.5}$ species (−25 to 39 % before; 0–3 % after) and improved correlations with ground-level monitors for elemental carbon, nitrate, and organic aerosol (R^2: 0.30 to 0.53 before; 0.53 to 0.87 after). GWR outperforms ordinary least squares (OLS) corrections for all $PM_{2.5}$ species except elemental carbon, where performance is comparable. Corrected fields also show improved performance in predicting fractional composition. Tract-level species- and source-resolved exposure estimates developed in this study are publicly available at www.caces.us.

1. Introduction

Chronic exposure to fine particulate matter ($PM_{2.5}$) causes adverse human health outcomes, including pulmonary and cardiovascular disease (Pope and Dockery, 2006; Pope et al., 2020), and is the leading contributor to morbidity and mortality among air pollutants (Cohen et al., 2017; Lefler et al., 2019). Air quality management in the United States has reduced concentrations of $PM_{2.5}$ in recent decades (Zhang

* Corresponding author. Institute of Chemical Engineering Sciences (FORTH/ICE-HT), Patras, Greece.
E-mail address: spyros@chemeng.upatras.gr (S.N. Pandis).

https://doi.org/10.1016/j.atmosenv.2025.121637
Received 14 June 2025; Received in revised form 13 October 2025; Accepted 26 October 2025
Available online 30 October 2025

et al., 2018), with health benefits consistently exceeding the cost of regulations several-fold (U.S.E.P.A. 1999, 2011, 2012). Studies suggest that further reductions of $PM_{2.5}$ in the United States would lead to improved health and increased life expectancy (Apte et al., 2018; Bennett et al., 2019; Correia et al., 2013; Pope III et al., 2009). To date, air quality regulations have targeted $PM_{2.5}$ total mass, with the implicit assumption that all $PM_{2.5}$ is equally toxic. However, $PM_{2.5}$ is a complex mixture with varying properties, including but not limited to size, phase, acidity, chemical composition, and source. Given the various pathophysiological pathways for $PM_{2.5}$-induced morbidity and mortality (Pope and Dockery, 2006), $PM_{2.5}$ toxicity could be a function of its properties; however, prior work has been unable to conclusively identify key drivers of $PM_{2.5}$ toxicity (Harrison and Yin, 2000; Kelly and Fussell, 2012). Investigating this topic in epidemiological studies requires accurate exposure estimates of $PM_{2.5}$ resolved by composition, sources, and/or other properties.

Conventional $PM_{2.5}$ exposure estimates for health studies are primarily derived from observations (Diao et al., 2019; Hoek et al., 2008, 2017). For example, data from ground-level monitors and satellites have been used with land-use and land-cover data to build empirical models that provide reliable and consistent estimates of ambient $PM_{2.5}$ levels; those estimates have been used in epidemiological studies as measures of exposure. Empirical models for $PM_{2.5}$ typically predict total $PM_{2.5}$ mass rather than composition and sources, reflecting in part the fact that observational data on $PM_{2.5}$ species, sources, and other properties are often limited or unavailable.

In contrast, chemical transport models (CTMs) can readily provide information about the character of $PM_{2.5}$, such as chemical composition and source. However, CTMs alone are unable to match the accuracy of empirical estimates. Several sources of errors – for example, errors in emission inventories and in chemical mechanisms, coarse grid resolution, large emissions gradients, complex terrain, and complex meteorology – lead to error and bias in CTM predictions. Spatial biases in uncorrected CTM estimates can make them unsuitable for epidemiological analyses. The underlying causes of CTM errors are generally non-random in space, resulting in different regions having different characteristic biases and errors. Therefore, applying a statistical technique that can identify and remediate regionally varying biases in CTM estimates could facilitate their use, including in epidemiological analyses, and provide much needed information on $PM_{2.5}$ properties.

Geographically weighted regression (GWR) is a local spatial analysis technique that models the spatially varying relationships between independent and dependent variables (Brunsdon et al., 1996). Regression coefficients in GWR are determined locally and can allow targeted identification of spatial biases in simulated $PM_{2.5}$. GWR's predictions of spatial biases can then be used to correct the CTM predictions throughout the domain. Several studies have used GWR to develop $PM_{2.5}$ exposure estimates, primarily as a predictive tool to correct satellite Aerosol Optical Depth (AOD) measurement to ground-level monitors (Franklin et al., 2017; Hammer et al., 2020; Hu, 2009; Hu et al., 2013; Li et al., 2017a; Ma et al., 2014; Meng et al., 2019; Song et al., 2014; van Donkelaar et al., 2015, 2016, 2019; You et al., 2016; Zhai et al., 2018). A subset of these studies has incorporated information from CTMs, such as spatiotemporal extent of $PM_{2.5}$ (Hammer et al., 2020; Li et al., 2017b; Meng et al., 2019; van Donkelaar et al., 2015, 2016, 2019; Zhang et al., 2020a). Only a few studies used CTMs to predict the chemical composition of $PM_{2.5}$ (Di et al., 2016; Geng et al., 2020; Jin et al., 2024; Rahman and Thurston, 2022; van Donkelaar et al., 2019). In the broader literature, studies using techniques other than GWR have also incorporated CTMs in $PM_{2.5}$ exposure estimates (Berrocal et al., 2020; Brauer et al., 2012; de Hoogh et al., 2016; Evans et al., 2013; Geng et al., 2015, 2017; Huang et al., 2021; Lee et al., 2012; Lyu et al., 2019; van Donkelaar et al., 2006, 2010; Wang et al., 2016; Zhang et al., 2020b). However, most of these studies used CTMs to obtain spatiotemporal estimates of total $PM_{2.5}$ mass; a few used information on chemical composition (Di et al., 2016; Geng et al., 2017; Jin et al., 2024; Li et al., 2017b; Meng et al., 2018a, 2018b; Philip et al., 2014). Zhang et al. (2022) extended the approach of Zhang et al. (2020b) to include also AOD data for greater coverage in regions where ground level observations are not available. Jin et al. (2024) combined the predicted WRF-Chem $PM_{2.5}$ composition with the AOD observations from the Multi-angle Imager for Aerosols (MAIA) instrument using Bayesian model averaging methods.

Data assimilation methods produce optimal estimates of a system state which is usually dynamical (e.g. temporal variation of pollutant concentrations), combining prior knowledge from model simulations with observations to produce an updated posterior estimate with techniques including Kalman filtering, Ensemble Kalman filtering, etc. If temporal variations are not considered GWR is more efficient, since data assimilation methods are more computationally demanding than GWR models.

Moreover, other studies used observations to treat algorithms in order to improve gas phase species predictions like O_3, using different methods like decision trees (Ivatt and Evans, 2020), bias correction (Skipper et al., 2021, 2024) or even machine learning (Keller et al., 2021). Most of the previous studies (Meng et al., 2018b; Ivatt and Evans, 2020; Zhang et al., 2020a, Skipper et al., 2024) rely on only one type of observations, usually ground-based observations for their analysis. A few studies (Di et al., 2016; Geng et al., 2020; Jin et al., 2024) have combined different types of data like ground-based observations and/or satellite data and/or land use data in order to improve CTM predictions of $PM_{2.5}$ and its composition. These studies were applied to a specific region of US and used a limited number of land use parameters to describe the details of each certain point.

In this work, we propose a novel approach for correcting biases in species- and source-resolved CTM $PM_{2.5}$ predictions by combining for the first time Geographically Weighted Regression (GWR) with the Integrated Empirical Geographic (IEG) regression (Kim et al., 2020). The IEG model leverages a comprehensive dataset—including land use regression and satellite observations—that characterizes well each area within the domain, ensuring rich contextual information. While CTM predictions may suffer from both spatial and temporal biases, our focus is on correcting spatial biases in long-term predictions, which are particularly critical for exposure assessment. By integrating diverse data sources (ground-level monitors, land use information, satellite retrievals, and CTM predictions), the method reduces uncertainties and strengthens prediction reliability for spatially resolved $PM_{2.5}$ sources and composition. A key advantage of our approach lies in the role of GWR, which explicitly accounts for spatial heterogeneity by performing local regressions around each point and weighing nearby observations. When local observations are unavailable, the IEG model provides robust fallback estimates, thereby maintaining high prediction accuracy. This dual mechanism offers a clear improvement over traditional machine learning algorithms, which are generally not inherently spatial and require spatial features to be added explicitly. As a result, our method can deliver more accurate and spatially consistent source-resolved $PM_{2.5}$ estimates. These improved predictions not only enhance the reliability of source apportionment analysis but also provide a strong foundation for epidemiological studies across large regions, such as the contiguous US (Pond et al., 2022). Tract-level species- and source-resolved $PM_{2.5}$ exposure estimates developed through this method are publicly accessible at www.caces.us.

2. Methods

2.1. Overview

Our approach combines observations from speciated ground-level monitors, empirical model estimates of total $PM_{2.5}$ (Kim et al., 2020), CTM predictions (Skyllakou et al., 2021), and other geographic information. First, GWR is used to predict the bias in speciated CTM predictions using training data at speciated monitor locations, similar to

van Donkelaar et al. (2019). The bias for each species predicted by GWR is then used to correct annually averaged CTM simulations across the contiguous United States. Species included in this study are elemental carbon (EC), organic aerosol (OA), ammonium (NH_4^+), nitrate (NO_3^-) and sulfate (SO_4^{2-}). Since we lack national-scale observations of source apportionment, we hold fractional source apportionment constant as we correct each species.

2.2. PMCAMx chemical transport model

PMCAMx (Karydis et al., 2010; Murphy and Pandis, 2010; Posner et al., 2019; Tsimpidi et al., 2010) is used to simulate $PM_{2.5}$ over the contiguous United States for 2001 and 2010 using a methodologically consistent set of emissions inventories (Xing et al., 2013). The Particulate Source Apportionment Technology (PSAT) algorithm (Skyllakou et al., 2014, 2017, 2021; Wagstrom et al., 2008, Wagstrom and Pandis, 2011a,b) is used to track the apportionment of emissions sources in a computationally efficient manner. The model domain covers the contiguous United States, portions of Canada and Mexico, and nearby offshore regions. The horizontal resolution is 36 km. Six source categories are resolved here: electricity-generating units (EGUs; industrial point sources included in EPA's Integrated Planning Model), other industrial point sources (non-EGU), on-road mobile, off-road mobile, biogenic volatile organic compounds (VOCs) from vegetation, and "other" which includes on-road emissions from Canada and Mexico, wildfire emissions, and all other emissions. This work is based on the results of Skyllakou et al. (2021). In Skyllakou et al. (2021) the wildfires were included into the "other" source category. The PM emissions from biomass burning were similar for 2001 and 2010. Wildfires were the dominant PM source of the "other" category, contributing 80 % in 2001 and 83 % in 2010 to the total emissions of this category. $PM_{2.5}$ concentrations attributed to the model's boundary and initial conditions are tracked as separate categories. Species predicted by the model and used in this work include sulfate, nitrate, ammonium, EC, primary organic aerosol (POA) and secondary organic aerosol (SOA). The apportionment of secondary PM components is based on the apportionment of their precursors. As a result, the apportionment of sulfate, nitrate, ammonium, and SOA is based on the apportionment of SO_2, NO_x, NH_3, and VOCs respectively. PMCAMx uses an advanced treatment of OA (using the "volatility basis set") that accounts for the semi-volatile nature of primary organic emissions and recent advances in our understanding of SOA chemistry (Donahue et al., 2006; Murphy and Pandis, 2009; Robinson et al., 2007). Model predictions of sodium, chloride, and mineral dust are excluded from this analysis due to large uncertainties in their emissions and because there are no direct measurements of ambient dust concentrations. Meteorological data input to PMCAMx are taken from simulations performed using the Weather Research Forecasting model (WRF v3.6.1) for these time periods. Initial and boundary conditions were generated from the ERA-Interim global climate re-analysis database. The CTM simulations are described further in Skyllakou et al. (2021).

2.3. Ground-level speciated $PM_{2.5}$ observations

Observations of $PM_{2.5}$ species (EC, OC, NH_4^+, NO_3^-, SO_4^{2-}) from the EPA Chemical Speciation Network (CSN) and the IMPROVE monitoring network for 2001 and 2010 were downloaded from the Federal Land Manager Environmental Database (http://views.cira.colostate.edu/fed/). Prior to the CSN transition period from 2007 to 2009, CSN and IMPROVE used different analytical and sampling protocols for carbon measurements, requiring harmonization across the datasets (Malm et al., 2011; Solomon et al., 2014; Spada and Hyslop, 2018). Most notably, pre-transition CSN monitors used the thermal optical transmittance (TOT) analytical protocol for carbon measurements, while IMPROVE and post-transition CSN monitors use the thermal optical reflectance (TOR) protocol. We adjust 2001 CSN carbon measurements to match post-transition CSN protocols following the approach in Lordo et al. (2016). Additionally, a filter blank correction of 0.4 μg m^{-3} is applied to organic carbon (OC) measurements in 2010. To account for differences in the aging of organic aerosol in urban and rural areas, an OA:OC ratio of 1.4 and 1.8 was applied to OC measurements collected at CSN and IMPROVE sites, respectively. There are uncertainties in the estimation of these constants, especially about the one used for IMPROVE sites because it could be higher or lower if the OA is very aged or not (El-Zanan et al., 2005; Ruthenburg et al., 2014). The value of 1.8 applied for IMPROVE sites is an assumption for the regulatory sites based on Pitchford et al. (2007). On average, 23 CSN and 92 IMPROVE monitors were used in 2001, and 165 CSN and 146 IMPROVE monitors were used in 2010 (Table 1).

2.4. Geographically weighted regression

We use GWR to correct bias in speciated PMCAMx predictions. Here, GWR is a spatial extension of ordinary least squares (OLS) regression. In GWR, regression coefficients vary in space, and observations are weighted in the regression according to their proximity to a desired prediction point in space. A consequence of this formulation is that there is no global model. Instead, the model is solved locally for every prediction point in space such that:

$$X^T W(i) X\beta(i) = X^T W(i) Y \quad (1)$$

where X is a matrix containing predictor variables, W is a weighting matrix (kernel) at location i, β is a vector of regression coefficients at location i, and Y is the dependent variable. The weighting matrix is a diagonal matrix, where each diagonal element is the weight assigned to an observation and is calculated by a user-defined weighting function (Eq. (2)). Selection of the weighting function depends largely on the nature of the dataset. Weighting functions are typically calibrated to an optimal bandwidth, which controls the rate observations are down-weighed with distance. Weighting functions can also have cut-offs, which exclude observations past a certain distance threshold. Common weighting functions include inverse distance weighting and a Gaussian function. Several weighting functions were considered and are discussed in the Supporting Information (SI). The primary results presented in this paper use a Gaussian weighting function:

$$w_{ij} = \exp\left(-\alpha d_{ij}^2\right) \quad (2)$$

where w_{ij} is the weight assigned to an observation in location j for predictions in location i, α is the decay coefficient or bandwidth, and d is the distance between location i and j. The bandwidth (α) is calibrated by minimizing the root mean square error in the GWR model. In Eq. (2), a

Table 1
Number of speciated monitors used in observational data by year, $PM_{2.5}$ species, and monitoring network. The number of co-located monitors reflects the instances where all 5 species are measured at the same monitoring location.

Year	Species	CSN	IMPROVE	Total
2001	EC	28	92	120
2010	EC	151	147	298
2001	OM	27	92	119
2010	OM	153	145	298
2001	NH_4^+	20	91[a]	111
2010	NH_4^+	174	146[a]	320
2001	NO_3^-	20	91	111
2010	NO_3^-	168	145	313
2001	SO_4^{2-}	21	91	112
2010	SO_4^{2-}	175	146	321
2001	Co-located	18	91	109
2010	Co-located	140	144	284

[a] Ammonium values from IMPROVE monitors are not directly measured and are instead inferred assuming full neutralization of sulfate and nitrate.

bandwidth of zero would lead to equal weighting for all observations, making the GWR model equivalent to OLS. The decay coefficient describes how the influence of nearby data points decreases with distance in the model. Variations in dispersion characteristics and atmospheric lifetimes across $PM_{2.5}$ components can result in differing estimated decay coefficients.

GWR is used to predict the bias in simulated $PM_{2.5}$ species, with GWR models being trained for each species and simulation year. Bias predictions are made at the centroids of U.S. census tracts and used to correct CTM simulations projected to census tracts. Bias predictions were performed at monthly and annual resolutions and are evaluated in the SI. However, the primary results of this paper are corrected with annually averaged bias predictions due to generally better performance. The GWR model form, modified from the model form used in van Donkelaar et al. (2019) is:

$$(sim\ SPEC - obs\ SPEC) = \beta_{ED} ED + \beta_{IDU} IDU + \beta_{IEGb} IEGb + \sum_c \beta_c\ sim\ SPEC_c \quad (3)$$

The left side of Eq. (3) represents the CTM bias relative to speciated ground-level monitors (i.e., the outcome variable); the right side contains predictor variables and regression coefficients. *SPEC* represents $PM_{2.5}$ species (EC, OA, NH_4^+, NO_3^-, SO_4^{2-}). *ED* represents the sub-grid elevation difference, which is the difference between the elevation of a prediction point and the mean elevation of the overlying CTM grid cell and is used to represent topographic effects unresolved by the CTM. Elevation is taken from the NOAA ETOP1 global relief model (Amante and Eakins, 2009; NOAA, 2009). *IDU* is the inverse distance to the nearest urban land cover as determined by using year-specific MODIS MCD12Q1 land cover data (Friedl and Sulla-Menashe, 2019). Higher spatial variability is expected in urban areas, potentially contributing to model error given the relatively coarse horizontal resolution of 36 km in the CTM. A maximum limit of 2 km^{-1} is set for *IDU* (i.e., when the distance to the nearest land cover is less than 0.5 km, the distance is replaced with the value 0.5 km prior to calculating the inverse distance) to avoid excessively high values for *IDU*. *IEGb* is the bias in total $PM_{2.5}$ concentration from the CTM relative to predictions of the Integrated Empirical Geographic (IEG) model by Kim et al. (2020). The *IEG* model is considered reliable, since it estimates annual averages of total $PM_{2.5}$ based on ground-level monitors, universal kriging and partial least squares of many geographic variables including land use variables and satellite-derived estimates of pollution levels and land-cover. *IEGb* serves as a useful initial guess of the bias, which we hypothesize is useful especially in locations that are far from an ambient monitor. The final predictor variables, *SPECc*, are a subset of simulated species from the CTM: when predicting the bias of carbonaceous species, *SPECc* represents carbonaceous species (EC, POA, SOA); when predicting the bias of a non-carbonaceous species, *SPECc* represents non-carbonaceous species (NH_4^+, NO_3^-, SO_4^{2-}). This specification for *SPECc* is meant to limit the number of predictor variables in the model and avoid model over-fitting. While the CTM provides predictions of POA and SOA, observations do not, so it is not possible to separately model biases in POA and SOA directly. Instead, the GWR is trained to model biases in total OA; after total OA is corrected, primary and secondary fractions as predicted by the CTM are applied proportionally to corrected OA. While corrected estimates of OA, POA and SOA are generated by this method, only those for total OA are evaluated here. The impact of individual regressors on the predicted bias for all simulation years and species is shown in Figs. S2–S11 in the SI. Each term of equation (3) can be a different physical quantity, but similar scales are ensured by the definition of the different coefficients β. Because the regression coefficients β are calculated locally and for each type of regression those values are a function also of the selection of the bandwidth (equation (2)).

The density of census tracts varies: smaller (denser) tracts in urban areas, larger (less dense) tracts in rural and remote areas. Predicting CTM biases at a census tract resolution accomplishes several objectives. Predicted biases can be used to downscale (i.e., to the tract level) the 36 km -resolution CTM estimates, including in urban and population-dense areas, where $PM_{2.5}$ experiences often greater spatial variability. Conversely, rural and low-populations areas are given lower resolution corrections. Finally, using tracts facilitates population-weighted averaging to coarser census geographies, such as counties or metropolitan statistical areas (MSAs). This aspect allows GWR predictions and therefore epidemiological analyses to be easily performed at the desired spatial resolution (if they are tract or coarser). GWR-corrected estimates at various census geographies are evaluated in the SI.

GWR models are evaluated here using three cross-validation methods: 1) leave-one-out, 2) 10-fold cross-validation (CV), and 3) a regional holdout CV. The regional holdout CV is designed to inform model errors in regions far from monitors, where performance is likely to be worse than that suggested by 10-fold CV. It functions similarly to the leave-one-out CV except that all monitors within a 400 km radius are also excluded from model training. Model training, prediction and evaluation are performed on the R open-source software with community-developed packages (R Core Team, 2022).

3. Results and discussion

3.1. CTM and GWR evaluation

We compared estimates of total uncorrected $PM_{2.5}$ from the CTM to those from the IEG model (Kim et al., 2020) and found significant regional biases in the raw CTM predictions (Fig. 1). Spatial trends in biases are problematic for long-term epidemiological studies because differences in health responses cannot be properly related to differences in chronic $PM_{2.5}$ exposures. This underscores the importance of correcting biases in CTM estimates with the GWR model.

In the Eastern U.S., CTM estimates are on average (population-weighted) 4.2 and 2.9 μg m^{-3} higher – in 2001 and 2010, respectively – when compared to the IEG model (Fig. 1, top panel). In the Western U.S., CTM estimates are on average 3.4 and 1.5 μg m^{-3} lower – in 2001 and 2010, respectively – when compared to the IEG model. We observed that much of the CTM's over-prediction in the Eastern U.S. could be attributed to high concentrations of crustal and sea salt PM. On average, crustal PM accounts for 6.1 and 5.4 μg m^{-3} in the Eastern U.S., in 2001 and 2010, respectively. In the Western U.S., crustal PM accounts for 3.0 and 2.0 μg m^{-3}, in 2001 and 2010, respectively. Previous studies have noted large uncertainties associated with crustal PM in emission inventories (Appel et al., 2013; Xu et al., 2019), making their predicted source mixtures potentially unreliable. Due to the fact that crustal and sea salt mostly exist in the coarse mode we make the assumption of removing them from our analysis. However, different regional biases persist after removing crustal PM and sea salt (Fig. 1, middle panel). In 2001, the bias is on average −1.7 μg m^{-3} in the Eastern U.S. and −6.3 μg m^{-3} in the Western U.S. In 2010, the bias is −1.9 μg m^{-3} in the Eastern U.S. and −3.6 μg m^{-3} in the Western U.S. In California, where CTM underpredictions are most severe, biases are on average −9.0 and −4.3 μg m^{-3}, in 2001 and 2010, respectively. Underlying CTM predictions corresponding to panels in Fig. 1 are illustrated in Fig. S1 in the SI.

GWR corrections address spatial patterns in total $PM_{2.5}$ bias and improve the performance of all species against available observations across several evaluation metrics (Fig. 1, bottoms panel). GWR corrections improve performance of all species, reducing overall bias and improving the predicted spatial pattern as measured by the R^2 (Fig. 2). Reducing bias improves predictions, lowers systematic residual error, and thereby decreases the unexplained variance. This directly translates into a higher R^2 because more of the variance in the data is now being accounted for by the model. Uncorrected OA and NO_3^- tend to be severely underpredicted in the West, often by a factor of 2 or more. OA in the East tends to be slightly overpredicted, particularly in the Southeast. GWR corrections significantly improve R^2 coefficients for simulated OA and NO_3^- from 0.30 to 0.50 (without GWR) to 0.53–0.87

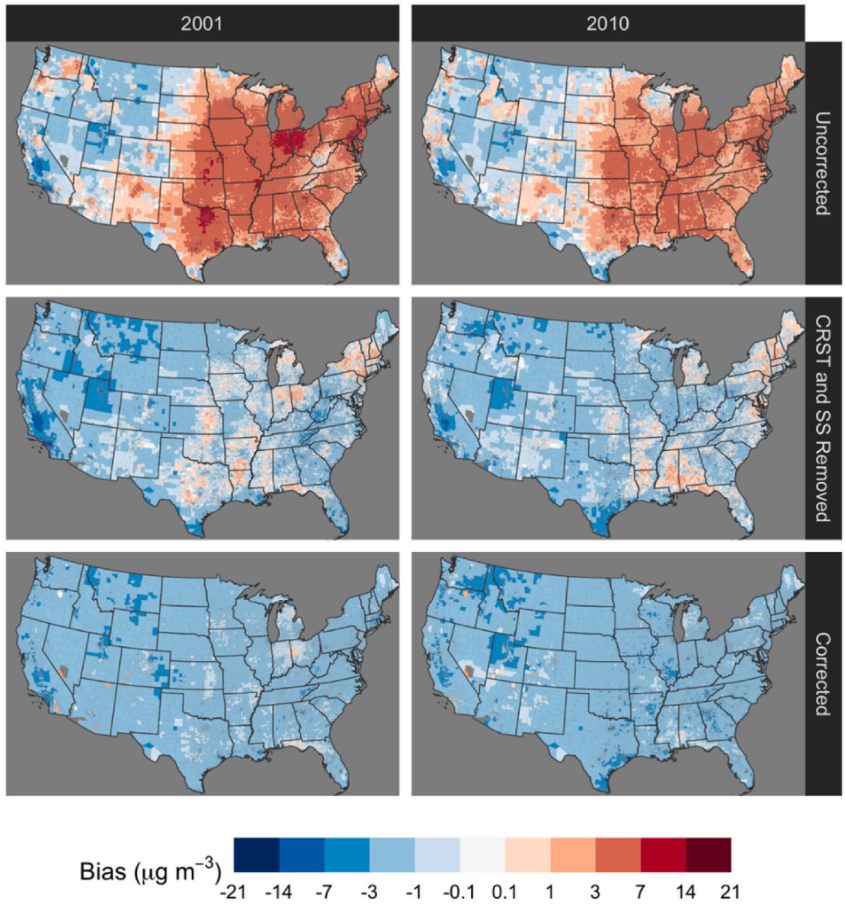

Fig. 1. Bias of CTM-predicted $PM_{2.5}$ relative to IEG-predicted $PM_{2.5}$ (i.e., CTM – IEG). Top row: uncorrected CTM predictions. Middle row: uncorrected CTM predictions with crustal PM (CRST) and sea salt (SS) removed. Bottom row: corrected CTM predictions without crustal PM and sea salt. Left and right columns show annually averaged biases for 2001 and 2010 predictions, respectively. Results are displayed here at the tract level.

(with GWR). While overall error and bias are improved for OA and NO_3^-, GWR is unable to reduce the error in Western concentrations, particularly for OA. A combination of lower monitor density, complex meteorology and terrain, and high spatial variability in urban areas could be responsible for the difficulty in correcting biases in the Western U.S. Additionally, this could reflect a need to find predictor variables that spatially correlate with the bias for OA and NO_3^- in the Western U.S. For NH_4^+ and SO_4^{2-}, GWR corrections improve upon the already good performance of the original CTM results. In the West, NH_4^+ tends to be underpredicted while SO_4^{2-} is slightly overpredicted. GWR addresses both biases and improves R^2 coefficients from 0.70 to 0.97 (without GWR) to 0.84–0.97 (with GWR). Successfully modeling EC is challenging due to its high spatial variability and the CTM's coarse resolution. However, GWR corrections make significant improvements to simulated EC: uncorrected EC simulations tend to be noisy and correlate poorly with monitors, with R^2 coefficients of 0.38 (in 2001) and 0.52 (in 2010). GWR corrections significantly reduce error and bias and increase R^2 coefficients to 0.62 (2001) and 0.71 (2010). GWR-corrected estimates of $PM_{2.5}$ species are mapped in Figs. S12–S19. Despite geographic differences in the biases, GWR models successfully improve model performance in regionally specific ways.

GWR corrections are largely robust across the three cross-validation methods used. Evaluation metrics like fractional error and fractional bias were calculated for the evaluation of the CTM's predictions against observations. All three CV results show significant reductions in fractional bias, moderate reductions in fractional error, and improved correlations for all species (Fig. 3). Results from leave-one-out (LOO) and 10-fold CV are nearly identical. Random folding may not provide additional insights beyond the LOO CV, because ground-level monitors tend to be clustered. In contrast, the regional holdout CV is helpful in evaluating model performance in locations far from a monitor, which tend to be remote. As expected, GWR performance is lower for regional holdout CV than for LOO or 10-fold CV. However, the regional holdout CV still yields significant improvements over the uncorrected CTM despite the 400 km radius holdout. In general, the regional holdout CV shows a fair bit of robustness of the improvements despite the challenging nature of the test. Of the models trained, the GWR performance for NO_3^- in 2001 is the most sensitive to having nearby monitors.

In general, the spatial patterns of observed biases (Fig. 4) are replicated by the GWR models (Fig. 5). For reference, uncorrected speciated CTM predictions are illustrated in Fig. S20 in the SI. Biases for EC are generally small and tend to be positive. In 2001, select metropolitan biases exhibit a stronger positive bias, while in 2010 the bias does not exhibit a clear spatial pattern. Biases for OA are negative in the Western U.S., particularly in the California Central Valley, and positive in the East Coast. For 2010, OA is overpredicted in the Southeastern U.S., where biogenic emissions contribute significantly to OA production. The biogenic emissions are in general higher in the Southeastern part compared to the other parts of US (Dinkelacker et al., 2024; Murphy and Pandis, 2010; Sindelarova et al., 2014; Skyllakou et al., 2021). In 2001, the OA bias in the Southeast may not be as severe in part because of higher anthropogenic emissions that partially offset the bias, or in part

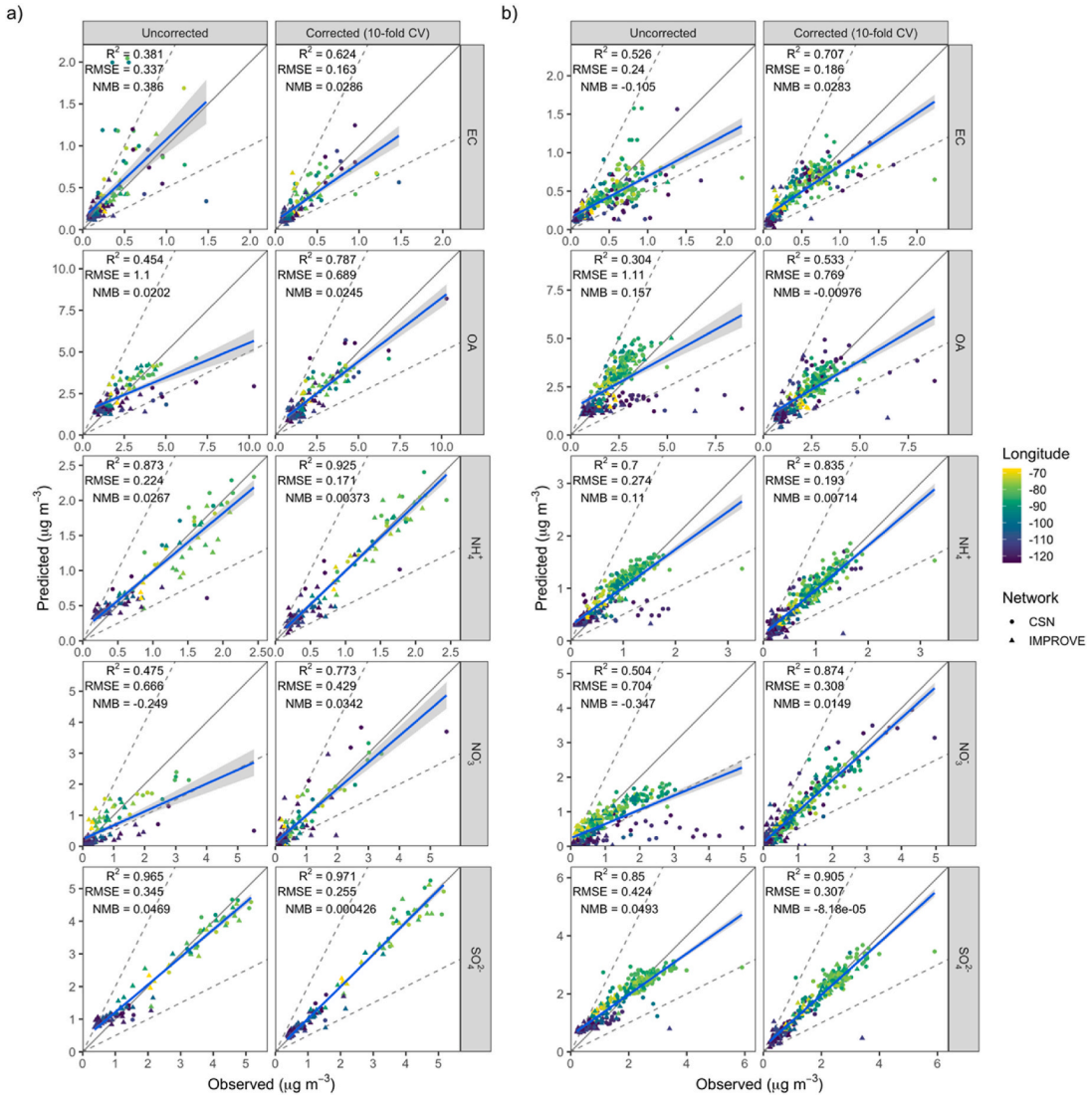

Fig. 2. Evaluation of CTM-predicted $PM_{2.5}$ species against observations from speciated ground-level monitors for simulation years a) 2001 and b) 2010. Values shown are annual averages at monitor locations. RMSE represents the root mean square error. NMB represents the normalized mean bias. Solid lines denote a 1:1 slope. Dashed lines denote a 1:2 or 2:1 slope. Points are color-coded by longitude to illustrate potentially differing biases in Eastern (color: yellow) versus Western (color: blue) U.S. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

because of the relatively few monitors available. Biases for NO_3^- are negative in the Western U.S., particularly in California, and positive in the Eastern U.S. In the Midwest, biases for NO_3^- tend to be positive in 2001, except for some urban areas, and tend to be negative in 2010. The change in sign may reflect the lack of monitors in the Midwest in 2001, which causes GWR to rely heavily on positively biased eastern monitors. A similar pattern in the Midwest, although opposite in sign, is also observed for SO_4^{2-}. Biases of SO_4^{2-} are overall relatively small but tend to be negative in the Eastern U.S. and California Central Valley, and positive in the Western U.S. Likewise, biases for NH_4^+ are relatively small and tend to mirror the patterns observed for SO_4^{2-}, except in the Southeast. Since NH_4^+ neutralizes both NO_3^- and SO_4^{2-}, positive biases for NO_3^- in the Southeast likely lead to positive biases for NH_4^+. Uncertainties based on the size distributions of the inorganic species arise because the semivolatile nature of ammonium nitrate increases the complexity (Bauer et al., 2007; Seinfeld and Pandis, 2016). Corrected species are reconstructed in the bottom panel of Fig. 1, with crustal PM and sea salt still omitted. For both simulation years, the bias-corrected CTM estimates continue to be biased low, by approximately 2 and 1.5 μg m^{-3} in 2001 and 2010, respectively. However, this is likely due to the omission of crustal PM and sea salt as residual biases in species treated here are too small to account for this remaining bias. Additionally, the bias is relatively uniform and no longer characterized by severe regional variations. GWR corrections improve the spatial consistency of CTM estimates and make them viable for exposure assignment in epidemiological studies.

3.2. *Changes to source mixtures*

The nation-wide network of speciated ground-level $PM_{2.5}$ monitors provide critical information on chemical composition, which is the basis for the corrections performed in this study. However, there are no observational data for source apportionment of $PM_{2.5}$ on a national scale. Therefore, analogous corrections to the source-resolution of $PM_{2.5}$

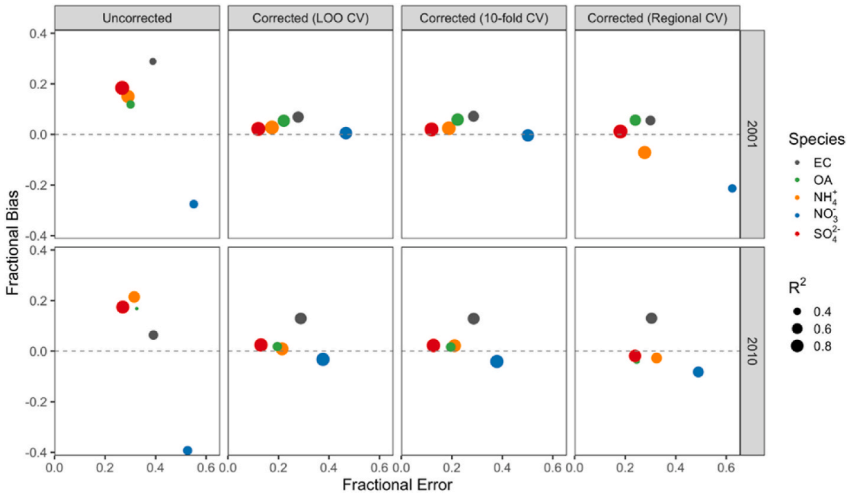

Fig. 3. CTM evaluations for uncorrected and GWR-corrected simulations trained under leave-one-out (LOO), 10-fold and regional cross-validation (CV) methods. Evaluation metrics are calculated for individual $PM_{2.5}$ species across the contiguous United States.

are not possible. Instead, as described above, for each species, we preserve the original source mixture, as predicted by the CTM, in our corrected estimates. The fractional source mixture is calculated for each simulated $PM_{2.5}$ species and then re-applied to the corrected estimates.

$$f_{i,s} = \frac{C_{i,s}}{C_{i,tot}} \tag{4}$$

$$C'_{i,s} = f_{i,s} C'_{i,tot} \tag{5}$$

where $f_{i,s}$ is the fractional source mixture for species i and source s, $C_{i,s}$ is the uncorrected concentration of species i and source s, $C_{i,tot}$ is the total uncorrected concentration of species i, C'_{is} is the corrected concentration of species i and source s, and C'_{itot} is the total corrected concentration of species i. Because the quantity of each species has been adjusted, the source mixture for total $PM_{2.5}$ is altered by GWR corrections (Fig. S21). In general, the changes to source mixture are minimal, except for the on-road and off-road source categories. GWR correction increases the contribution from on-road and off-road emissions in the West by approximately 50 %. This outcome is primarily due to the increase in NO_3^- in corrected CTM simulations. On-road and off-road emissions tend to be concentrated near specific locations (e.g., roadways). Coupled with complex terrain in the Western U.S., NO_3^- emissions gradients tend to be large and poorly represented in coarse 36 km resolution grids. As a result, NO_3^- from mobile sources are under-predicted by the CTM, an error that is indirectly adjusted by GWR corrections. GWR-corrected estimates of source-specific $PM_{2.5}$ are mapped in Figs. S22–S27.

3.3. OLS and GWR comparison

We compare the performance of GWR versus OLS at predicting CTM biases by (1) comparing evaluation metrics for GWR- and OLS-corrected estimates and (2) quantifying the spatial autocorrelation in GWR and OLS residual biases. In the latter evaluation, the model that is best suited to addresses region-specific biases should exhibit the least spatial autocorrelation in their respective residual biases.

Those results (Fig. S28) indicate that model performance is better for OLS than uncorrected CTM output; GWR-corrected estimates modestly outperform than OLS-corrected estimates across all these evaluation metrics, except for EC, where the performance for both models is comparable. That result reflects that EC is primary and lacks speciated monitors needed to fully characterize the high intra-urban variability of EC. Based on this evaluation, GWR does provide appreciable improvements over OLS, given that spatial patterns of $PM_{2.5}$ species are adequately characterized in the training data.

To further evaluate GWR improvements over OLS, residual biases are calculated at each monitor location as the difference between the observed and predicted CTM bias. Moran's I is used to quantify the spatial autocorrelation in the models' residual biases. A Moran's I greater than 0 indicates positive autocorrelation (i.e., clustered values), a Moran's I less than 0 indicates negative autocorrelation (i.e., dispersed values), and a Moran's I equal to 0 indicates no autocorrelation (i.e., randomized values). Positive autocorrelation values are undesirable here because they indicate spatially coherent – rather than random – biases remaining in the corrected estimates, which could bias subsequent epidemiological work. Moran's I is calculated at cumulatively increasing lag distances of 100–2000 km from monitor locations. Calculation of Moran's I is discussed in greater detail in the SI. For NH_4^+, NO_3^- and SO_4^{2-}, residual biases from the OLS model exhibit a statistically significant positive autocorrelation across all lag distances (Fig. 6). Residual biases that are spatially autocorrelated suggest that region-specific CTM biases are unidentified and untreated in the subsequent corrections. In contrast, residual biases for NH_4^+, NO_3^-, and SO_4^{2-} from the GWR show no statistically significant autocorrelation, highlighting the value added from the GWR model. In 2010, residual biases for OA from the GWR-model exhibit a statistically significant negative autocorrelation across all lag distances. Likewise, residual biases from EC tend to be negatively autocorrelated, although not statistically significant. This negative autocorrelation (i.e., dissimilarity between nearby data points) is likely attributable to the high spatial variability of EC and OA and the lack of monitors needed to fully characterize said spatial variability in urban areas. In 2001, no statistically significant autocorrelation is exhibited for EC and OA, however this is due to a monitoring network that is dominated by rural IMPROVE monitors in the Western U.S. The increase of urban CSN monitors from 2001 to 2010 likely increased the degree of dissimilarity in GWR residual biases for EC and

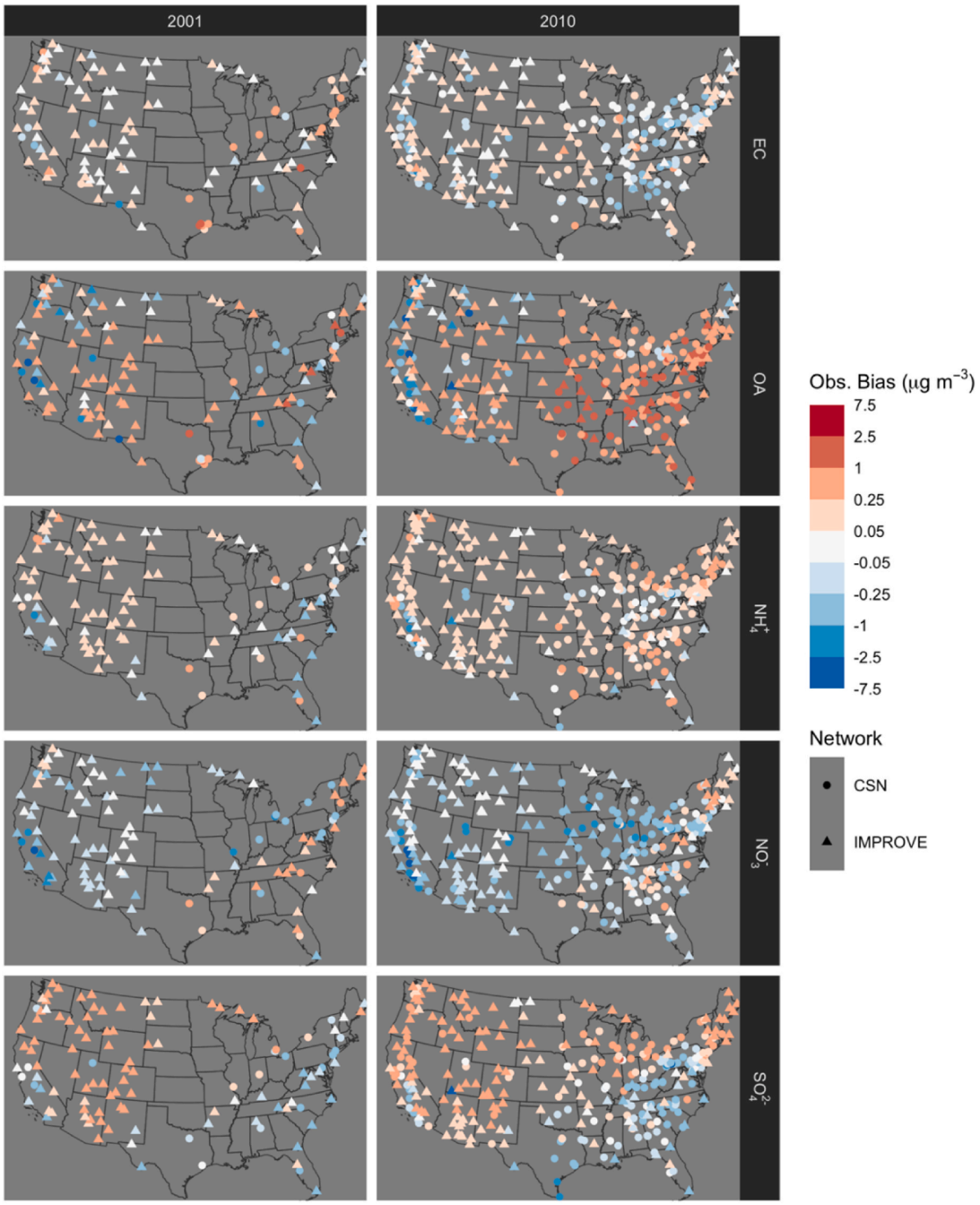

Fig. 4. Bias of uncorrected CTM $PM_{2.5}$ species relative to observations at speciated ground-level monitors (observed bias) expressed as (uncorrected CTM prediction – observation).

OA. The analysis of spatial autocorrelation, therefore, presents strong evidence that OLS alone leaves regionally specific biases that are problematic for use in epidemiological analysis, but that GWR is able to address.

3.4. Compositional evaluation

Some applications of the predicted composition fields may use the fractional composition of $PM_{2.5}$ rather than the speciated concentrations themselves (Pond et al., 2022). Therefore, as an additional point of evaluation, the fractional composition of simulated $PM_{2.5}$ is compared to that observed in speciated ground-level monitors. Here, we are interested in examining the CTM's performance in predicting species mixtures on a fractional basis and improvements made by GWR corrections. Fractional composition can be represented as a vector, where each vector component corresponds to the fractional contribution of a species. The degree of dissimilarity between species mixtures can be quantified by calculating the angle between their corresponding vectors. For reference, Fig. S29 compares example monitor and CTM species mixtures with vector angles of 5, 10, 20 and 30° between them. A vector

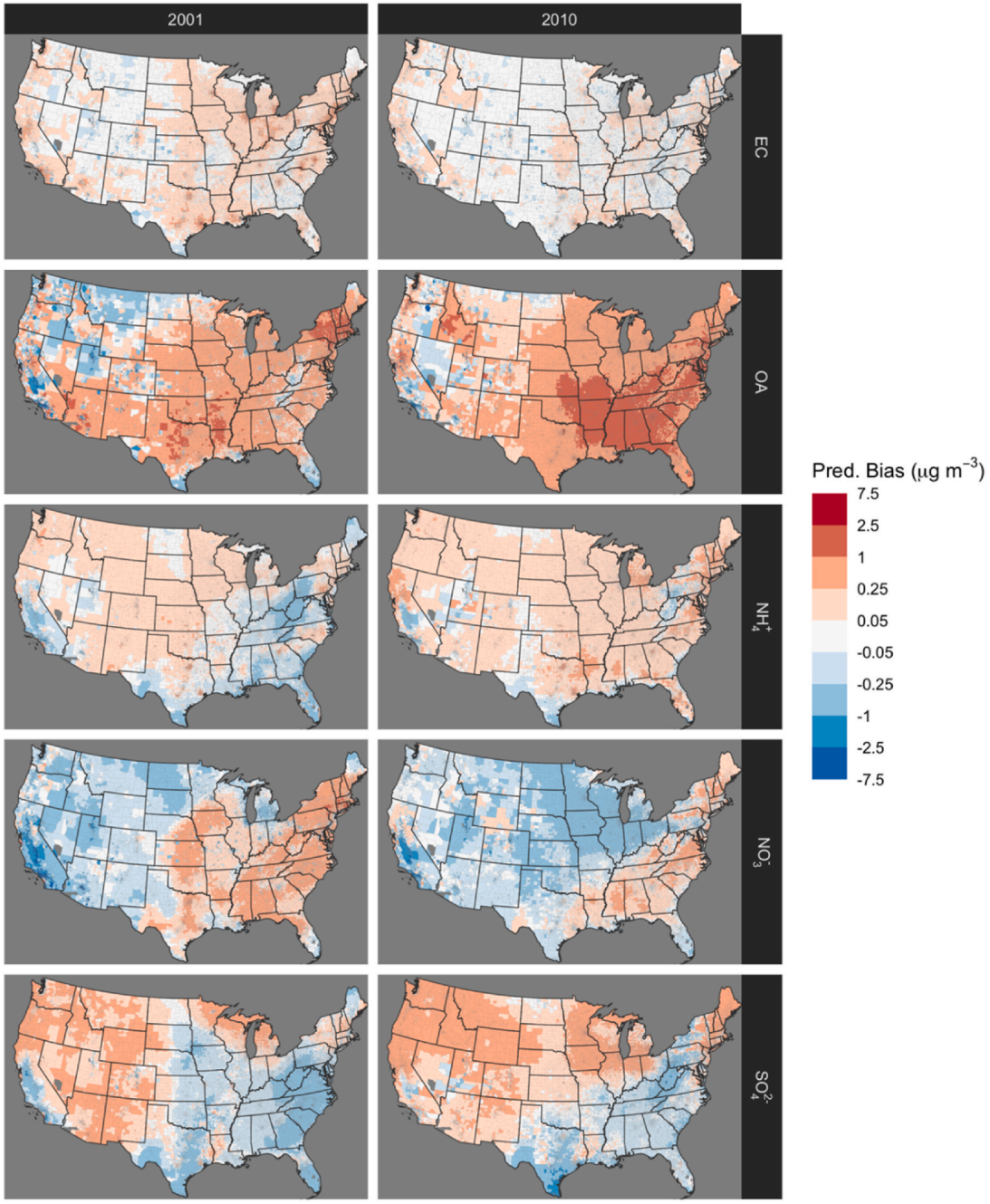

Fig. 5. Bias of CTM $PM_{2.5}$ species as predicted by GWR models (predicted bias) expressed as (uncorrected CTM – corrected CTM prediction). Results are displayed here at the tract level.

angle below 10° is classified as showing good agreement between simulated and observed species mixtures. Changes in vector angles before and after GWR corrections are shown in Fig. 7, with 10-fold CV results being presented. Fig. 7 highlights a clear shift in the distribution of vector angles, with an average decrease of 2.3 and 5.5° in 2001 and 2010, respectively. In 2001, the number of monitor locations with vector angles under 10° increases from 56 to 70 % as a result of GWR corrections. In 2010, the increase is greater, from 32 to 80 %. In general, corrected estimates in eastern locations agree better with monitors than those in western locations. GWR corrections improve agreement between CTM and CSN monitors. However, there appears to be a degradation in performance at some western IMPROVE locations. This is consistent with earlier results showing persistent error after corrections are applied to western OA and NO_3^- estimates, and regional holdout CV results that suggest weaker predictive ability in remote regions.

3.5. *Species- and source-resolution in four metropolitan areas*

GWR corrections improve the spatial consistency and performance of CTM simulations by leveraging valuable information from monitors,

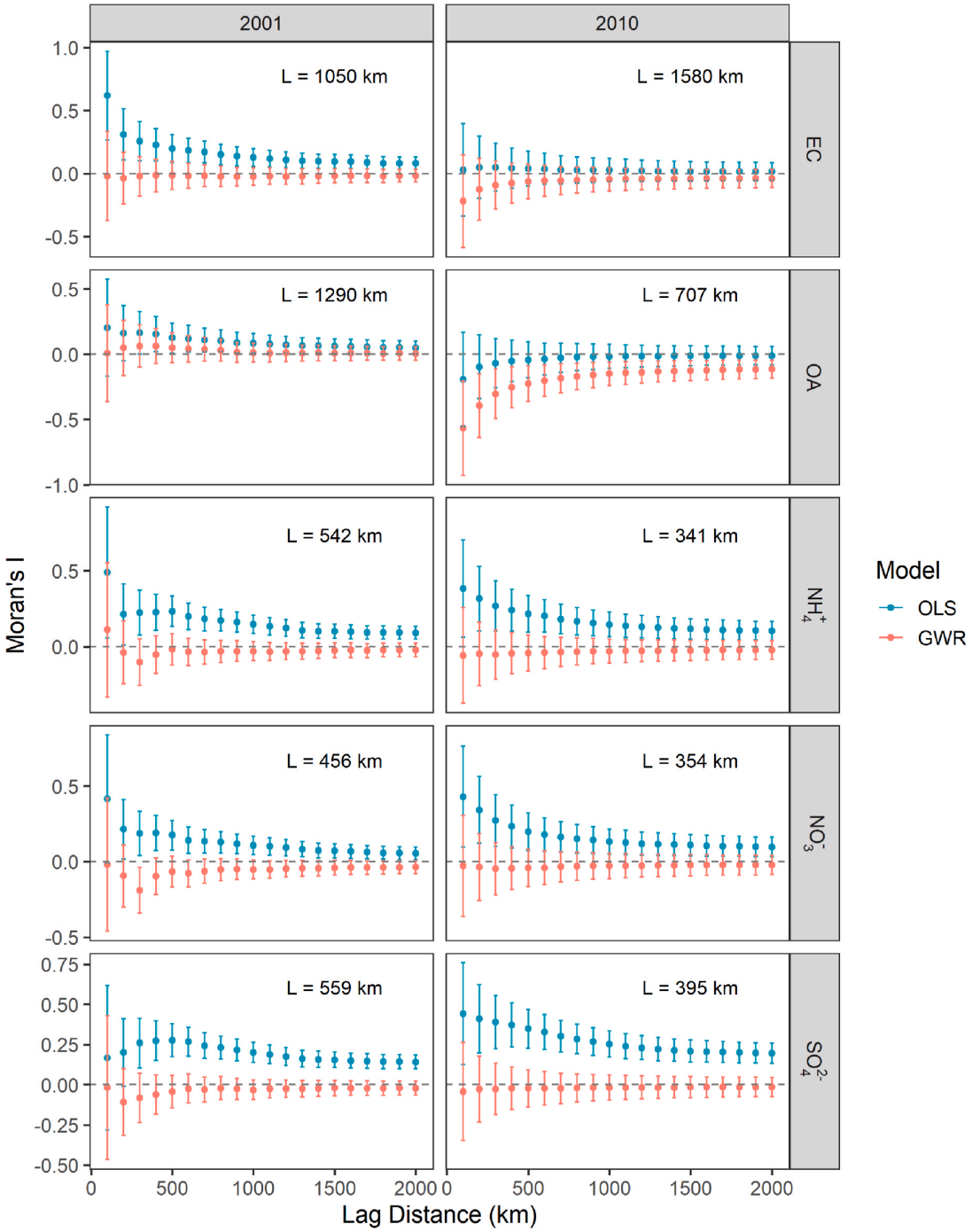

Fig. 6. Spatial autocorrelation (Moran's I) of residual biases in OLS and GWR models across cumulatively increasing lag distances. Error bars correspond to the 95 % confidence interval. For reference the calibrated *e*-folding length ($L = \sqrt{1/\alpha}$), the distance at which observation points are weighted e^{-1}, is shown for each GWR model.

empirical model estimates, and other geographic information. With these corrections, the species- and source-resolution of the $PM_{2.5}$ exposure estimates provided by the CTM greatly enriches their utility in exposure assessment and environmental epidemiology. The impact of the GWR corrections (Fig. 8) illustrates the absolute and compositional changes in population-weighted $PM_{2.5}$ for four metropolitan statistical areas. The greatest changes are observed in the western metropolitan areas, where corrections to OA and NO_3^- result in large increases to $PM_{2.5}$. In the Los Angeles metropolitan area, $PM_{2.5}$ increases by 76–82 %, with OA increasing by 2.9 and 2.2 μg m^{-3}, in 2001 and 2010 respectively, and NO_3^- increasing 3.4 and 2.1 μg m^{-3}, in 2001 and 2010 respectively. In the Denver metropolitan area, $PM_{2.5}$ increases by 43–46

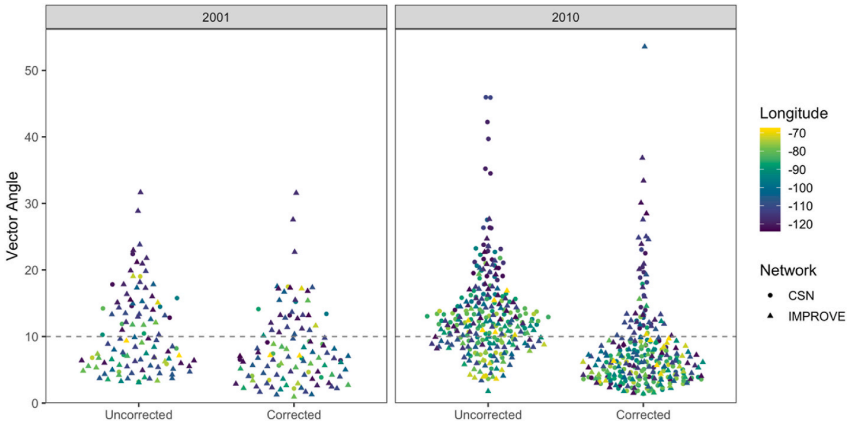

Fig. 7. Change in vector angle distributions before and after GWR correction. Vector angles are calculated when speciated monitors for EC, OC, NH_4^+, NO_3^-, and SO_4^{2-} are all present. Corrected vector angles are based on results from the 10-fold cross-validation.

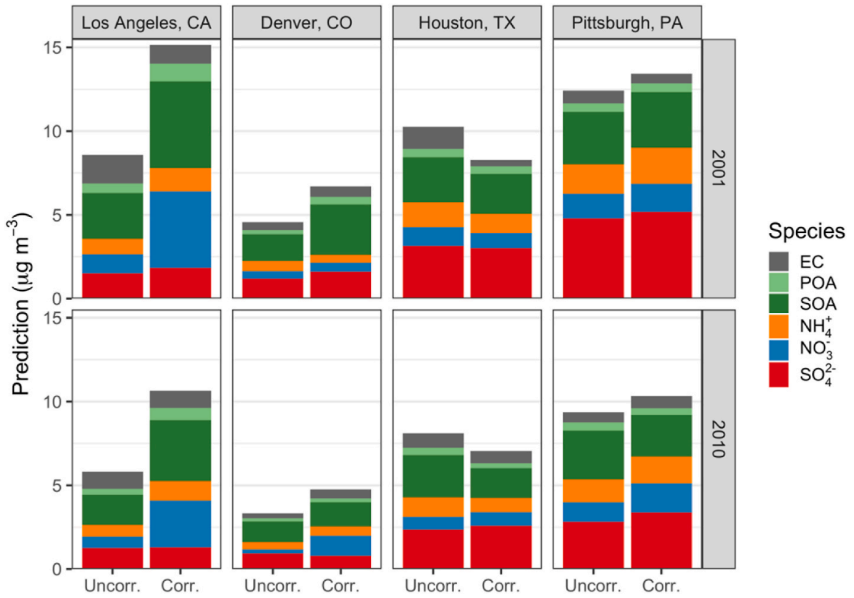

Fig. 8. Comparison of compositional changes to population-weighted $PM_{2.5}$ across four representative metropolitan statistical areas. Uncorrected and corrected CTM estimates are shown for each metropolitan statistical area and simulation year.

%, with OA increasing 1.6 μg m^{-3} in 2001, and NO_3^- increasing 1.0 μg m^{-3} in 2010. For eastern metro areas, the changes are relatively modest, since CTM predictions generally perform better in this region. Unlike the Western U.S., $PM_{2.5}$ changes in the Eastern U.S. are not always driven by corrections to a consistent set of species. In the Houston metropolitan area, $PM_{2.5}$ decreases by 2.0 and 1.1 μg m^{-3}, in 2001 and 2010 respectively. In the Pittsburgh metropolitan area, $PM_{2.5}$ increases by 1.0 and 0.9 μg m^{-3}, in 2001 and 2010 respectively.

While the dataset presented here was generated with exposure assessment in mind, researchers and others may find the national-scale source resolution useful for other applications. Fig. 9 illustrates the source apportionment of corrected $PM_{2.5}$ predictions for the same set of metropolitan statistical areas discussed previously. Mobile sources tend to be most significant in western metros, while industrial sources tend to be most significant in eastern metros. Additionally, the contribution from these traditional emissions sources has decreased from 2001 to 2010. The fraction of total $PM_{2.5}$ coming from mobile sources decreased from 35 % (in 2001) to 30 % (in 2010) in Los Angeles and from 25 % (in 2001) to 20 % (in 2010) in Denver. The proportion of total $PM_{2.5}$ from industrial sources (i.e., EGU plus non-EGU) decreased during 2001–2010, from 30 % to 27 % in Houston and from 37 % to 30 % in Pittsburgh. As air quality regulations in the United States continue to target traditional industrial and mobile sources, contributions from long-range transport and "other" sources – which include emissions from Canada and Mexico, agriculture, and wildfires – are rising in prominence. Sources previously unresolved in emissions inventories will likely require greater attention.

4. Conclusions

In this study, the GWR method was extended using a tailored formulation that incorporated both available observational data and Integrated Empirical Geographic (IEG) estimates. The extended GWR was applied to a large dataset of PMCAMx-PSAT CTM predictions for the

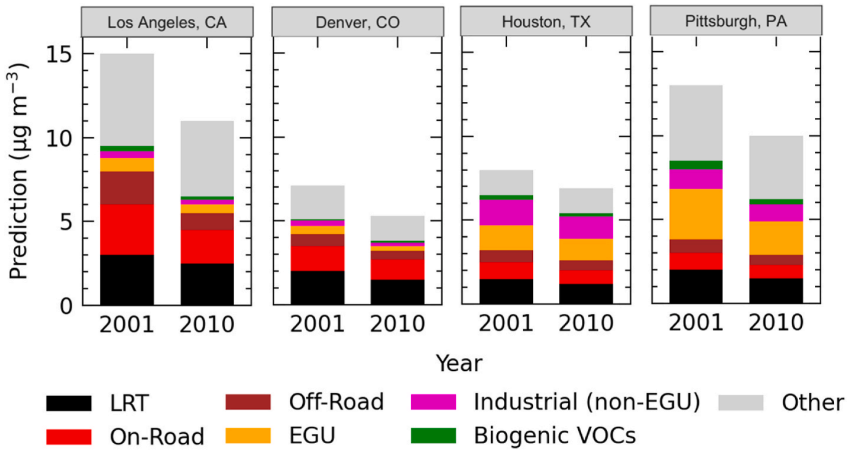

Fig. 9. Source apportionment of population-weighted $PM_{2.5}$ across four representative metropolitan statistical areas. Contributions from the CTMs initial and boundary conditions are grouped together as long-range transport (LRT).

contiguous United States for the years 2001 and 2010. The integration of GWR and IEG is a key advantage of the proposed approach, as it allows multiple data sources to be combined to improve CTM predictions. The approach resulted in substantial improvements in predictive accuracy ($R^2 = 0.53–0.97$), with notable increases in R^2, and reductions in RMSE and NMB for all $PM_{2.5}$ components. The most significant gains were observed for OA and nitrate, with up to a 77 % increase in R^2 and close to a 100 % reduction in NMB.

Going forward, we hope the approach developed and applied here enables CTM results to be more widely used for predictions, such as $PM_{2.5}$ source apportionment, that are unavailable elsewhere at the same scale. This method is flexible and transparent as the user can define each parameter separately and may add more parameters to achieve better performance. The use of geostatistical methods, including but not limited to GWR, should also be considered when processing CTM simulations, with the aim of improving and further evaluating those estimates.

CRediT authorship contribution statement

Ksakousti Skyllakou: Writing – review & editing, Software, Methodology, Investigation, Formal analysis. **Carlos S. Hernandez:** Writing – original draft, Software, Methodology, Formal analysis. **Pablo Garcia Rivera:** Writing – review & editing, Software, Investigation. **Brian Dinkelacker:** Writing – review & editing, Software, Investigation, Formal analysis. **Julian D. Marshall:** Writing – review & editing, Supervision, Conceptualization. **C. Arden Pope:** Writing – review & editing, Supervision, Conceptualization. **Allen L. Robinson:** Writing – review & editing, Supervision, Conceptualization. **Spyros N. Pandis:** Writing – review & editing, Supervision. **Peter J. Adams:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This publication was developed as part of the Center for Air, Climate, and Energy Solutions (CACES), which was supported under Assistance Agreement No. R835873 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by EPA. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2025.121637.

Data availability

Data will be made available on request.

References

Amante, C., Eakins, B.W., 2009. ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. https://doi.org/10.7289/V5C8276M.

Appel, K.W., Pouliot, G.A., Simon, H., Sarwar, G., Pye, H.O.T., Napelenok, S.L., Akhtar, F., Roselle, S.J., 2013. Evaluation of dust and trace metal estimates from the Community Multiscale Air Quality (CMAQ) model version 5.0. Geosci. Model Dev. (GMD) 6, 883–899. https://doi.org/10.5194/gmd-6-883-2013.

Apte, J.S., Brauer, M., Cohen, A.J., Ezzati, M., Pope, C.A., 2018. Ambient $PM_{2.5}$ reduces global and regional life expectancy. Environ. Sci. Technol. Lett. 5, 546–551. https://doi.org/10.1021/acs.estlett.8b00360.

Bauer, S.E., Koch, D., Unger, N., Metzger, S.M., Shindell, D.T., Streets, D.G., 2007. Nitrate aerosols today and in 2030: a global simulation including aerosols and tropospheric ozone. Atmos. Chem. Phys. 7, 5043–5059. https://doi.org/10.5194/acp-7-5043-2007.

Bennett, J.E., Tamura-Wicks, H., Parks, R.M., Burnett, R.T., Pope, C.A., Bechle, M.J., Marshall, J.D., Danaei, G., Ezzati, M., 2019. Particulate matter air pollution and national and county life expectancy loss in the USA: a spatiotemporal analysis. PLoS Med. 16, e1002856. https://doi.org/10.1371/journal.pmed.1002856.

Berrocal, V.J., Guan, Y., Muyskens, A., Wang, H., Reich, B.J., Mulholland, J.A., Chang, H. H., 2020. A comparison of statistical and machine learning methods for creating national daily maps of ambient $PM_{2.5}$ concentration. Atmos. Environ. 222, 117130. https://doi.org/10.1016/j.atmosenv.2019.117130.

Brauer, M., Amann, M., Burnett, R.T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S. B., Krzyzanowski, M., Martin, R.V., Van Dingenen, R., van Donkelaar, A., Thurston, G.D., 2012. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. Environ. Sci. Technol. 46, 652–660. https://doi.org/10.1021/es2025752.

Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. 28, 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet 389, 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6.

Correia, A.W., Pope, C.A., Dockery, D.W., Wang, Y., Ezzati, M., Dominici, F., 2013. Effect of air pollution control on life expectancy in the United States: an analysis of 545 U.

S. counties for the period from 2000 to 2007. Epidemiology 24, 23–31. https://doi.org/10.1097/EDE.0b013e3182770237.

de Hoogh, K., Gulliver, J., Donkelaar, A. van, Martin, R.V., Marshall, J.D., Bechle, M.J., Cesaroni, G., Pradas, M.C., Dedele, A., Eeftens, M., Forsberg, B., Galassi, C., Heinrich, J., Hoffmann, B., Jacquemin, B., Katsouyanni, K., Korek, M., Künzli, N., Lindley, S.J., Lepeule, J., Meleux, F., de Nazelle, A., Nieuwenhuijsen, M., Nystad, W., Raaschou-Nielsen, O., Peters, A., Peuch, V.-H., Rouil, L., Udvardy, O., Slama, R., Stempfelet, M., Stephanou, E.G., Tsai, M.Y., Yli-Tuomi, T., Weinmayr, G., Brunekreef, B., Vienneau, D., Hoek, G., 2016. Development of West-European $PM_{2.5}$ and NO_2 land use regression models incorporating satellite-derived and chemical transport modelling data. Environ. Res. 151, 1–10. https://doi.org/10.1016/j.envres.2016.07.005.

Di, Q., Koutrakis, P., Schwartz, J., 2016. A hybrid prediction model for $PM_{2.5}$ mass and components using a chemical transport model and land use regression. Atmos. Environ. 390–399. https://doi.org/10.1016/j.atmosenv.2016.02.002.

Diao, M., Holloway, T., Choi, S., O'Neill, S.M., Al-Hamdan, M.Z., Van Donkelaar, A., Martin, R.V., Jin, X., Fiore, A.M., Henze, D.K., Lacey, F., Kinney, P.L., Freedman, F., Larkin, N.K., Zou, Y., Kelly, J.T., Vaidyanathan, A., 2019. Methods, availability, and applications of $PM_{2.5}$ exposure estimates derived from ground measurements, satellite, and atmospheric models. J. Air Waste Manag. Assoc. 69, 1391–1414. https://doi.org/10.1080/10962247.2019.1668498.

Dinkelacker, B.T., Rivera, P.G., Skyllakou, K., Adams, P.J., Pandis, S.N., 2024. Predicted and observed changes in summertime biogenic and total organic aerosol in the southeast United States from 2001 to 2010. Atmos. Environ. 316, 120186. https://doi.org/10.1016/j.atmosenv.2023.120186.

Donahue, N.M., Robinson, A.L., Stanier, C.O., Pandis, S.N., 2006. Coupled partitioning, dilution, and chemical aging of semivolatile organics. Environ. Sci. Technol. 40, 2635–2643. https://doi.org/10.1021/es052297c.

El-Zanan, H.S., Lowenthal, D.H., Zielinska, B., Chow, J.C., Kumar, N., 2005. Determination of the organic aerosol mass to organic carbon ratio in IMPROVE samples. Chemosphere 60, 485–496. https://doi.org/10.1016/j.chemosphere.2005.01.005.

Evans, J., van Donkelaar, A., Martin, R.V., Burnett, R., Rainham, D.G., Birkett, N.J., Krewski, D., 2013. Estimates of global mortality attributable to particulate air pollution using satellite imagery. Environ. Res. 120, 33–42. https://doi.org/10.1016/j.envres.2012.08.005.

Franklin, M., Kalashnikova, O.V., Garay, M.J., 2017. Size-resolved particulate matter concentrations derived from 4.4km-resolution size-fractionated multi-angle imaging SpectroRadiometer (MISR) aerosol optical depth over Southern California. Remote Sens. Environ. 196, 312–323. https://doi.org/10.1016/j.rse.2017.05.002.

Friedl, Mark, Sulla-Menashe, Damien, 2019. MCD12Q1 MODIS/Terra+Aqua land cover type yearly L3 global 500m SIN grid V006. https://doi.org/10.5067/MODIS/MCD12Q1.006.

Geng, G., Zhang, Q., Martin, R.V., van Donkelaar, A., Huo, H., Che, H., Lin, J., He, K., 2015. Estimating long-term $PM_{2.5}$ concentrations in China using satellite-based aerosol optical depth and a chemical transport model. Rem. Sens. Environ. 166, 262–270. https://doi.org/10.1016/j.rse.2015.05.016.

Geng, G., Zhang, Q., Tong, D., Li, M., Zheng, Y., Wang, S., He, K., 2017. Chemical composition of ambient $PM_{2.5}$ over China and relationship to precursor emissions during 2005–2012. Atmos. Chem. Phys. 17, 9187–9203. https://doi.org/10.5194/acp-17-9187-2017.

Geng, G., Meng, X., He, K., Yang Liu, Y., 2020. Random forest models for $PM_{2.5}$ speciation concentrations using MISR fractional AODs. Environ. Res. Lett. 15, 034056.

Hammer, M.S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A.M., Hsu, N.C., Levy, R.C., Garay, M.J., Kalashnikova, O.V., Kahn, R.A., Brauer, M., Apte, J.S., Henze, D.K., Zhang, L., Zhang, Q., Ford, B., Pierce, J.R., Martin, R.V., 2020. Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). Environ. Sci. Technol. 54, 7879–7890. https://doi.org/10.1021/acs.est.0c01764.

Harrison, R.M., Yin, J., 2000. Particulate matter in the atmosphere: which particle properties are important for its effects on health? Sci. Total Environ. 249, 85–101. https://doi.org/10.1016/S0048-9697(99)00513-6.

Hoek, G., 2017. Methods for assessing long-term exposures to outdoor air pollutants. Curr. Environ. Health Rep. 4, 450–462. https://doi.org/10.1007/s40572-017-0169-5.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 42, 7561–7578. https://doi.org/10.1016/j.atmosenv.2008.05.057.

Hu, X., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes, M.G., Estes, S.M., Quattrochi, D.A., Sarnat, J.A., Liu, Y., 2013. Estimating ground-level $PM_{2.5}$ concentrations in the southeastern U.S. using geographically weighted regression. Environ. Res. 121, 1–10. https://doi.org/10.1016/j.envres.2012.11.003.

Hu, Z., 2009. Spatial analysis of MODIS aerosol optical depth, $PM_{2.5}$, and chronic coronary heart disease. Int. J. Health Geogr. 8, 27. https://doi.org/10.1186/1476-072X-8-27.

Huang, C., Hu, J., Xue, T., Xu, H., Wang, M., 2021. High-resolution spatiotemporal modeling for ambient $PM_{2.5}$ exposure assessment in China from 2013 to 2019. Environ. Sci. Technol. 55, 2152–2162. https://doi.org/10.1021/acs.est.0c05815.

Ivatt, P.D., Evans, M.J., 2020. Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees. Atmos. Chem. Phys. 20, 8063–8082. https://doi.org/10.5194/acp-20-8063-2020.

Jin, Z., Pu, Q., Janechek, N., Zhang, H., Wang, J., Chang, H., Liu, Y., 2024. A MAIA-like modeling framework to estimate $PM_{2.5}$ mass and speciation concentrations with uncertainty. Rem. Sens. Environ., 113995 https://doi.org/10.1016/j.rse.2024.113995.

Karydis, V.A., Tsimpidi, A.P., Fountoukis, C., Nenes, A., Zavala, M., Lei, W., Molina, L.T., Pandis, S.N., 2010. Simulating the fine and coarse inorganic particulate matter concentrations in a polluted megacity. Atmos. Environ. 44, 608–620. https://doi.org/10.1016/j.atmosenv.2009.11.023.

Keller, C.A., Evans, M.J., Knowland, K.E., Hasenkopf, C.A., Modekurty, S., Lucchesi, R.A., Oda, T., Franca, B.B., Mandarino, F.C., Díaz Suárez, M.V., Ryan, R.G., Fakes, L.H., Pawson, S., 2021. Global impact of COVID-19 restrictions on the surface concentrations of nitrogen dioxide and ozone. Atmos. Chem. Phys. 21, 3555–3592. https://doi.org/10.5194/acp-21-3555-2021.

Kelly, F.J., Fussell, J.C., 2012. Size, source and chemical composition as determinants of toxicity fattributable to ambient particulate matter. Atmos. Environ. 60, 504–526. https://doi.org/10.1016/j.atmosenv.2012.06.039.

Kim, S.-Y., Bechle, M., Hankey, S., Sheppard, L., Szpiro, A.A., Marshall, J.D., 2020. Concentrations of criteria pollutants in the contiguous U.S., 1979 – 2015: role of prediction model parsimony in integrated empirical geographic regression. PLoS One 15, e0228535. https://doi.org/10.1371/journal.pone.0228535.

Lee, S.-J., Serre, M.L., van Donkelaar, A., Martin, R.V., Burnett, R.T., Jerrett, M., 2012. Comparison of geostatistical interpolation and remote sensing techniques for estimating long-term exposure to ambient $PM_{2.5}$ concentrations across the Continental United States. Environ. Health Perspect. 120, 1727–1732. https://doi.org/10.1289/ehp.1205006.

Lefler, J.S., Higbee, J.D., Burnett, R.T., Ezzati, M., Coleman, N.C., Mann, D.D., Marshall, J.D., Bechle, M., Wang, Y., Robinson, A.L., Arden Pope, C., 2019. Air pollution and mortality in a large, representative U.S. cohort: multiple-pollutant analyses, and spatial and temporal decompositions. Environ. Health 18, 101. https://doi.org/10.1186/s12940-019-0544-9.

Li, C., Martin, R.V., van Donkelaar, A., Boys, B.L., Hammer, M.S., Xu, J.-W., Marais, E.A., Reff, A., Strum, M., Ridley, D.A., Crippa, M., Brauer, M., Zhang, Q., 2017b. Trends in chemical composition of global and regional population-weighted fine particulate matter estimated for 25 years. Environ. Sci. Technol. 51, 11185–11195. https://doi.org/10.1021/acs.est.7b02530.

Li, T., Shen, H., Zeng, C., Yuan, Q., Zhang, L., 2017a. Point-surface fusion of station measurements and satellite observations for mapping $PM_{2.5}$ distribution in China: methods and assessment. Atmos. Environ. 152, 477–489. https://doi.org/10.1016/j.atmosenv.2017.01.004.

Lordo, R., Landis, E., Rice, J., Triplett, C., 2016. Assessing the Statistical Relationship in Carbon Measurements Between Old and New Sampling and Analysis Protocols in the Chemical Speciation Network (CSN).

Lyu, B., Hu, Y., Zhang, W., Du, Y., Luo, B., Sun, X., Sun, Z., Deng, Z., Wang, Xiaojiang, Liu, J., Wang, Xuesong, Russell, A.G., 2019. Fusion method combining ground-level observations with chemical transport model predictions using an ensemble deep learning framework: application in China to estimate spatiotemporally-resolved $PM_{2.5}$ exposure fields in 2014–2017. Environ. Sci. Technol. 53, 7306–7315. https://doi.org/10.1021/acs.est.9b01117.

Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating ground-level $PM_{2.5}$ in China using satellite remote sensing. Environ. Sci. Technol. 48, 7436–7444. https://doi.org/10.1021/es5009399.

Malm, W.C., Schichtel, B.A., Pitchford, M.L., 2011. Uncertainties in $PM_{2.5}$ gravimetric and speciation measurements and what we can learn from them. J. Air Waste Manag. Assoc. 61, 1131–1149. https://doi.org/10.1080/10473289.2011.603998.

Meng, X., Garay, M.J., Diner, D.J., Kalashnikova, O.V., Xu, J., Liu, Y., 2018a. Estimating $PM_{2.5}$ speciation concentrations using prototype 4.4 km-resolution MISR aerosol properties over Southern California. Atmos. Environ. 118, 70–81. https://doi.org/10.1016/j.atmosenv.2018.03.019.

Meng, X., Hang, J.L., Schichtel, B.A., Liu, Y., 2018b. Space-time trends of $PM_{2.5}$ constituents in the conterminous United States estimated by a machine learning approach, 2005-2015. Environ. Int. 121, 1137–1147. https://doi.org/10.1016/j.envint.2018.10.029.

Meng, J., Li, C., Martin, R.V., van Donkelaar, A., Hystad, P., Brauer, M., 2019. Estimated long-term (1981–2016) concentrations of ambient fine particulate matter across North America from chemical transport modeling, satellite remote sensing, and ground-based measurements. Environ. Sci. Technol. 53, 5071–5079. https://doi.org/10.1021/acs.est.8b06875.

Murphy, B.N., Pandis, S.N., 2010. Exploring summertime organic aerosol formation in the eastern United States using a regional-scale budget approach and ambient measurements. JGR Atmos. 115. https://doi.org/10.1029/2010JD014418.

Murphy, B.N., Pandis, S.N., 2009. Simulating the formation of semivolatile primary and secondary organic aerosol in a regional chemical transport model. Environ. Sci. Technol. 43, 4722–4728. https://doi.org/10.1021/es803168a.

NOAA, 2009. ETOPO1 1 Arc-Minute Global Relief Model.

Philip, S., Martin, R.V., van Donkelaar, A., Lo, J.W.-H., Wang, Y., Chen, D., Zhang, L., Kasibhatla, P.S., Wang, S., Zhang, Q., Lu, Z., Streets, D.G., Bittman, S., Macdonald, D.J., 2014. Global chemical composition of ambient fine particulate matter for exposure assessment. Environ. Sci. Technol. 48, 13060–13068. https://doi.org/10.1021/es502965b.

Pitchford, M., Malm, W., Schichtel, B., Kumar, N., Lowenthal, D., Hand, J., 2007. Revised Algorithm for estimating light extinction from IMPROVE particle speciation data. J. Air Waste Manag. Assoc. 57 (11), 1326–1336. https://doi.org/10.3155/1047-3289.57.11.1326.

Pond, Z.A., Hernandez, C.S., Adams, P.J., Pandis, S.N., Garcia, G.R., Robinson, A.L., Marshall, J.D., Burnett, R., Skyllakou, K., Garcia Rivera, P., Karnezi, E., Coleman, C.J., Pope, C.A., 2022. Cardiopulmonary mortality and fine particulate air pollution by species and source in a national U.S. cohort. Environ. Sci. Technol. 56, 7214–7223. https://doi.org/10.1021/acs.est.1c04176.

Pope, C.A., Coleman, N., Pond, Z.A., Burnett, R.T., 2020. Fine particulate air pollution and human mortality: 25+ years of cohort studies. Environ. Res. 183, 108924. https://doi.org/10.1016/j.envres.2019.108924.

Pope, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 56, 709–742. https://doi.org/10.1080/10473289.2006.10464485.

Pope III, C.A., Ezzati, M., Dockery, D.W., 2009. Fine-particulate air pollution and life expectancy in the United States. N. Engl. J. Med. 11.

Posner, L.N., Theodoritsi, G., Robinson, A., Yarwood, G., Koo, B., Morris, R., Mavko, M., Moore, T., Pandis, S.N., 2019. Simulation of fresh and chemically-aged biomass burning organic aerosol. Atmos. Environ. 196, 27–37. https://doi.org/10.1016/j.atmosenv.2018.09.055.

R Core Team, 2022. A language and environment for statistical computing. https://www.R-project.org/.

Rahman, M.M., Thurston, G., 2022. A hybrid satellite and land use regression model of source-specific $PM_{2.5}$ and $PM_{2.5}$ constituents. Environ. Int. 163, 107233. https://doi.org/10.1016/j.envint.2022.107233.

Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking organic aerosols: semivolatile emissions and photochemical aging. Science 315, 1259–1262. https://doi.org/10.1126/science.1133061.

Ruthenburg, T.C., Perlin, P.C., Liu, V., McDade, C.E., Dillner, A.M., 2014. Determination of organic matter and organic matter to organic carbon ratios by infrared spectroscopy with application to selected sites in the IMPROVE network. Atmos. Environ. 86, 47–57. https://doi.org/10.1016/j.atmosenv.2013.12.034.

Seinfeld, J.H., Pandis, S.N., 2016. In: Atmospheric Chemistry and Physics, third ed. John Wiley and Sons, New Jersey, USA.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., Knorr, W., 2014. Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. Atmos. Chem. Phys. 14, 9317–9341. https://doi.org/10.5194/acp-14-9317-2014.

Skipper, T.N., Hu, Y., Odman, M.T., Henderson, B.H., Hogrefe, C., Mathur, R., Russell, A. G., 2021. Estimating US background ozone using data fusion. Environ. Sci. Technol. 55, 4504–4512. https://doi.org/10.1021/acs.est.0c08625.

Skipper, T.N., Hogrefe, C., Henderson, B.H., Mathur, R., Foley, K.M., Russell, A.G., 2024. Source-specific bias correction of US background and anthropogenic ozone modeled in CMAQ. Geosci. Model Dev. (GMD) 17, 8373–8397. https://doi.org/10.5194/gmd-17-8373-2024.

Skyllakou, K., Fountoukis, C., Charalampidis, P., Pandis, S.N., 2017. Volatility-resolved source apportionment of primary and secondary organic aerosol over Europe. Atmos. Environ. 167, 1–10. https://doi.org/10.1016/j.atmosenv.2017.08.005.

Skyllakou, K., Murphy, B.N., Megaritis, A.G., Fountoukis, C., Pandis, S.N., 2014. Contributions of local and regional sources to fine PM in the megacity of Paris. Atmos. Chem. Phys. 14, 2343–2352. https://doi.org/10.5194/acp-14-2343-2014.

Skyllakou, K., Rivera, P.G., Dinkelacker, B., Karnezi, E., Kioutsioukis, I., Hernandez, C., Adams, P.J., Pandis, S.N., 2021. Changes in $PM_{2.5}$ concentrations and their sources in the US from 1990 to 2010. Atmos. Chem. Phys. 21, 17115–17132. https://doi.org/10.5194/acp-21-17115-2021.

Solomon, P.A., Crumpler, D., Flanagan, J.B., Jayanty, R.K.M., Rickman, E.E., McDade, C. E., 2014. U.S. National $PM_{2.5}$ chemical speciation monitoring Networks—Csn and IMPROVE: description of networks. J. Air Waste Manag. Assoc. 64, 1410–1438. https://doi.org/10.1080/10962247.2014.956904.

Song, W., Jia, H., Huang, J., Zhang, Y., 2014. A satellite-based geographically weighted regression model for regional $PM_{2.5}$ estimation over the Pearl River Delta region in China. Rem. Sens. Environ. 154, 1–7. https://doi.org/10.1016/j.rse.2014.08.008.

Spada, N.J., Hyslop, N.P., 2018. Comparison of elemental and organic carbon measurements between IMPROVE and CSN before and after method transitions. Atmos. Environ. 178, 173–180. https://doi.org/10.1016/j.atmosenv.2018.01.043.

Tsimpidi, A.P., Karydis, V.A., Zavala, M., Lei, W., Molina, L., Ulbrich, I.M., Jimenez, J.L., Pandis, S.N., 2010. Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area. Atmos. Chem. Phys. 22. https://doi.org/10.5194/acp-10-525-2010.

U.S. Environmental Protection Agency, 2012. Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U.S. Environmental Protection Agency, 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020. Office of Air and Radiation, Washington D.C.

U.S. Environmental Protection Agency, 1999. The Benefits and Costs of the Clean Air Act, 1990 to 2010. Office of Air and Radiation, Washington D.C.

van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., Winker, D.M., 2016. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. 50, 3762–3772. https://doi.org/10.1021/acs.est.5b05833.

van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., Villeneuve, P.J., 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environ. Health Perspect. 118, 847–855. https://doi.org/10.1289/ehp.0901623.

van Donkelaar, A., Martin, R.V., Li, C., Burnett, R.T., 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. 53, 2595–2611. https://doi.org/10.1021/acs.est.8b06392.

van Donkelaar, A., Martin, R.V., Park, R.J., 2006. Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing. JGR Atmos. 111. https://doi.org/10.1029/2005JD006996.

van Donkelaar, A., Martin, R.V., Spurr, R.J.D., Burnett, R.T., 2015. High-resolution satellite-derived $PM_{2.5}$ from optimal estimation and geographically weighted regression over North America. Environ. Sci. Technol. 49, 10482–10491. https://doi.org/10.1021/acs.est.5b02076.

Wagstrom, Kristina M., Pandis, S.N., 2011a. Source–receptor relationships for fine particulate matter concentrations in the Eastern United States. Atmos. Environ. 45, 347–356. https://doi.org/10.1016/j.atmosenv.2010.10.019.

Wagstrom, K.M., Pandis, S.N., 2011b. Contribution of long-range transport to local fine particulate matter concerns. Atmos. Environ. 45, 2730–2735. https://doi.org/10.1016/j.atmosenv.2011.02.040.

Wagstrom, K.M., Pandis, S.N., Yarwood, G., Wilson, G.M., Morris, R.E., 2008. Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model. Atmos. Environ. 42, 5650–5659. https://doi.org/10.1016/j.atmosenv.2008.03.012.

Wang, M., Sampson, P.D., Hu, J., Kleeman, M., Keller, J.P., Olives, C., Szpiro, A.A., Vedal, S., Kaufman, J.D., 2016. Combining land-use regression and chemical transport modeling in a spatiotemporal geostatistical model for ozone and $PM_{2.5}$. Environ. Sci. Technol. 50, 5111–5118. https://doi.org/10.1021/acs.est.5b06001.

Xing, J., Pleim, J., Mathur, R., Pouliot, G., Hogrefe, C., Gan, C.-M., Wei, C., 2013. Historical gaseous and primary aerosol emissions in the United States from 1990 to 2010. Atmos. Chem. Phys. 13, 7531–7549. https://doi.org/10.5194/acp-13-7531-2013.

Xu, J.-W., Martin, R.V., Henderson, B.H., Meng, J., Öztaner, Y.B., Hand, J.L., Hakami, A., Strum, M., Phillips, S.B., 2019. Simulation of airborne trace metals in fine particulate matter over North America. Atmos. Environ. 214, 116883. https://doi.org/10.1016/j.atmosenv.2019.116883.

You, W., Zang, Z., Zhang, L., Li, Y., Pan, X., Wang, W., 2016. National-scale estimates of ground-level $PM_{2.5}$ concentration in China using geographically weighted regression based on 3 km resolution MODIS AOD. Remote Sens. 8, 184. https://doi.org/10.3390/rs8030184.

Zhai, L., Li, S., Zou, B., Sang, H., Fang, X., Xu, S., 2018. An improved geographically weighted regression model for $PM_{2.5}$ concentration estimation in large areas. Atmos. Environ. 181, 145–154. https://doi.org/10.1016/j.atmosenv.2018.03.017.

Zhang, Y., West, J.J., Mathur, R., Xing, J., Hogrefe, C., Roselle, S.J., Bash, J.O., Pleim, J. E., Gan, C.-M., Wong, D.C., 2018. Long-term trends in the ambient $PM_{2.5}$- and O_3-related mortality burdens in the United States under emission reductions from 1990 to 2010. Atmos. Chem. Phys. 18, 15003–15016. https://doi.org/10.5194/acp-18-15003-2018.

Zhang, T., Geng, G., Liu, Y., Chang, H.H., 2020a. Application of bayesian additive regression trees for estimating daily concentrations of $PM_{2.5}$ components. Atmosphere 11, 1233. https://doi.org/10.3390/atmos11111233.

Zhang, H., Wang, J., García, L.C., Ge, C., Plessel, T., Szykman, J., Murphy, B., Spero, T.L., 2020b. Improving surface $PM_{2.5}$ forecasts in the United States using an ensemble of chemical transport model outputs: 1. Bias correction with surface observations in nonrural areas. JGR Atmos. 125, e2019JD032293. https://doi.org/10.1029/2019JD032293.

Zhang, H., Wang, J., Garcia, L.C., Zhou, M., Ge, C., Plessel, T., Szykman, J., Levy, R.C., Murphy, B., Spero, T.L., 2022. Improving surface $PM_{2.5}$ forecasts in the United States using an ensemble of chemical transport model outputs: 2. Bias correction with satellite data for rural areas. JGR Atmos. 127, e2021JD035563. https://doi.org/10.1029/2021JD035563.