WILEY

Nonstationary spatiotemporal Bayesian data fusion for pollutants in the near-road environment

O. Gilani[1] | V. J. Berrocal[2] | S. A. Batterman[3]

[1]Department of Mathematics, Bucknell University, Lewisburg, Pennsylvania

[2]Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, Michigan

[3]Department of Environmental Health Sciences, University of Michigan School of Public Health, Ann Arbor, Michigan

Correspondence
O. Gilani, Department of Mathematics, Bucknell University, Lewisburg, PA 17837.
Email: owais.gilani@bucknell.edu

Funding information
Health Effects Institute, Grant/Award Number: R-82811201

Abstract

Concentrations of near-road air pollutants (NRAPs) have increased to very high levels in many urban centers around the world, particularly in developing countries. The adverse health effects of exposure to NRAPs are greater when the exposure occurs in the near-road environment as compared to background levels of pollutant concentration. Therefore, there is increasing interest in monitoring pollutant concentrations in the near-road environment. However, due to various practical limitations, monitoring pollutant concentrations near roadways and traffic sources is generally rather difficult and expensive. As an alternative, various deterministic computer models that provide predictions of pollutant concentrations in the near-road environment, such as the research line-source dispersion model (RLINE), have been developed. A common feature of these models is that their outputs typically display systematic biases and need to be calibrated in space and time using observed pollutant data. In this paper, we present a nonstationary Bayesian data fusion model that uses a novel data set on monitored pollutant concentrations (nitrogen oxides or NO_x and fine particulate matter or $PM_{2.5}$) in the near-road environment and, combining it with the RLINE model output, provides predictions at unsampled locations. The model can also be used to evaluate whether including the RLINE model output leads to improved pollutant concentration predictions and whether the RLINE model output captures the spatial dependence structure of NRAP concentrations in the near-road environment. A defining characteristic of the proposed model is that we model the nonstationarity in the pollutant concentrations by using a recently developed approach that includes covariates, postulated to be the driving force behind the nonstationary behavior, in the covariance function.

KEYWORDS

covariates in covariance, data fusion, fine particulate matter, nitrogen oxide, nonstationarity, numerical model output

1 | INTRODUCTION

Concentrations of near-road air pollutants (NRAPs) have increased to very high levels in many urban centers around the world, particularly in developing countries. Due to the well-established adverse effects of exposure to air pollutants

on various health outcomes, a number of environmental regulations have been established to monitor concentrations of pollutants. Episodes of exceedance over compliance thresholds have occurred in some urban environments, which have led to drastic measures to reduce pollutant concentration levels within cities (Garfield, 2018). Evidence suggests that the adverse health impacts of exposure to NRAPs are greater when the exposure occurs closer to major roads and highways as compared to background levels of pollutants (Hoek, Brunekreef, Goldbohm, Fischer, & van den Brandt, 2002; Lewis et al., 2005). An important societal implication of this result is that the greatest burden of this exposure is carried by populations belonging to lower socioeconomic status, since they constitute a higher proportion of residents living in close proximity to major roads (Gunier, Hertz, Von Behren, & Reynolds, 2003).

Accurate maps of various pollutant concentrations on a fine spatial and temporal resolution, particularly within the near-road environment, are important both for regulatory purposes and for assigning exposure to subjects in environmental epidemiologic studies. However, due to practical and monetary limitations, good-quality data on concentrations of pollutants on a dense spatial and temporal scale are generally not available. Instead, to monitor pollutant concentrations in real time, environmental agencies often rely on a network of fixed site monitors, most commonly those operated by states and local offices of the United States Environmental Protection Agency (US Environmental Protection Agency, 2017). There are a number of limitations to using such monitors for both regulatory purposes and for estimating concentrations at unsampled locations. First, the network of monitors is generally spatially very sparse. Studies conducted within the near-road environments have observed that fixed/central site monitors do not capture the spatial variability exhibited by NRAPs (Vardoulakis, Gonzalez-Flesca, Fisher, & Pericleous, 2005), and the hourly maxima recorded at these sites are often well below the concentrations measured near major highways (Wallace, Corr, Deluca, Kanaroglou, & McCarry, 2009). Second, due to practical limitations, these monitors are placed at considerable heights (between 3 and 15 m above ground; Office of Air Quality Planning and Standards, 1998), which is generally not the height of air inhaled at ground level. Third, monitors are often preferentially located, so that the concentrations measured by these monitors may not be representative of the entire population's exposure. Finally, due to maintenance downtime and mechanical failures, data are sometimes not recorded. Therefore, while these monitors contain valuable information about urban and background levels of pollutant concentrations, they do not provide very accurate information about concentrations closer to highways or the specific contribution of traffic sources to air pollution.

In light of these limitations, there is growing interest in utilizing outputs from deterministic computer models that apply numerical simulation techniques in various ways to estimate air pollution concentrations. Two examples of such models are the Community Multiscale Air Quality model (Byun & Schere, 2006), which numerically simulates the underlying pollutant chemistry, and the research line-source dispersion model (RLINE; Snyder et al., 2013), which simulates the dispersion of air pollutants from traffic and other sources. While such models may provide estimates of pollutant concentrations on a dense temporal resolution over the geographic region of interest, they often have systematic biases associated with them due to multiple reasons. Therefore, these model outputs need to be calibrated in space and time. To address the joint goals of calibrating numerical model outputs (i.e., estimating their bias) and providing more accurate predictions of atmospheric pollutants at unsampled locations, various statistical modeling techniques have been developed and successfully applied to a number of pollutants, including sulfur dioxide (Fuentes & Raftery, 2005), ground-level ozone (Berrocal, Gelfand, & Holland, 2010b, 2012; Reich, Chang, & Foley, 2014; Zidek, Le, & Liu, 2012), nitrogen dioxide (Gilani et al., 2016b), and fine particulate matter (Berrocal, Gelfand, & Holland, 2010a; Choi, Fuentes, & Reich, 2009; McMillan, Holland, Morara, & Feng, 2010; Crooks & Özkaynak, 2014; Rundel, Schliep, Gelfand, & Holland, 2015). Although these spatial and spatiotemporal models perform well when calibrating regional-scale deterministic model outputs over an extended spatial domain, they have some limitations for modeling pollutant concentrations within a localized urban setting in a near-road environment. In addition, most of these models assume a stationary covariance function for the residual spatial dependence in pollutant concentrations after accounting for the spatial correlation that comes from the numerical model output, that is, these models commonly express the residual correlation between pollutant concentrations at two locations as a function only of their distance, which might not be appropriate in an urban setting.

Modeling spatial and spatiotemporal processes using nonstationary covariance functions has received much attention over the last few years (Fuentes, 2001; Higdon, 1998; Holland, Saltzman, Cox, & Nychka, 1999; Lindgren, Rue, & Lindström, 2011; Paciorek & Schervish, 2006; Sampson & Guttorp, 1992). Recent work has further extended some of the nonstationary covariance methods to include spatial covariates in the covariance function, thus aiding in identifying factors that drive nonstationarity (Calder, 2008; Gilani, Berrocal, & Batterman, 2016a; Neto, Schmidt, & Guttorp, 2014; Reich, Eidsvik, Guindani, Nail, & Schmidt, 2011; Risser & Calder, 2015; Schmidt, Guttorp, & O'Hagan, 2011).

In this paper, we present a nonstationary spatiotemporal data fusion model for two NRAPs: nitrogen oxides (NO_x) and fine particulate matter ($PM_{2.5}$). The goals of this model are (a) to model the contribution of traffic sources to NRAP

concentrations, particularly in the near-road environment, and provide concentration predictions at unsampled locations; (b) to evaluate whether incorporating additional information on NRAPs in the near-road environment provided by the RLINE dispersion model output improves the estimation of pollutant concentrations; and (c) if incorporating additional information does improves the estimation of pollutant concentrations, to assess whether the RLINE model output captures the spatial dependence structure of NRAP concentrations in near-road environments. The proposed model follows the joint modeling approach, with a couple of innovations. First, since exposure to NRAPs in the near-road environment has a potentially significantly worse adverse impact on human health as compared to exposure to background concentration levels, instead of modeling the background concentration of pollutants, we model the near-road increment (NRI) in pollutant concentrations due to vehicular sources. This is achieved using data obtained from a novel pollutant concentration monitoring approach that used a mobile air pollution lab (MAPL) to monitor concentrations across highways within an urban setting. Second, we capture the potential nonstationarity in the residual spatial correlation between sites that are upwind versus downwind of the highway using a recently developed model (Gilani et al., 2016a) that includes covariates in the covariance function. To our knowledge, this is the first time a nonstationary data fusion model is presented in which the covariance function includes spatial covariates.

A few recent studies have evaluated the RLINE model output for various pollutant concentrations. However, these studies have either just visually compared observed concentrations versus the RLINE model output (Saha, Khlystov, Snyder, & Grieshop, 2018) or compared the correlation between the RLINE model output against various forms of observed concentrations: across regional air quality monitors (Milando & Batterman, 2018; Yu et al., 2018) or smoothed averages based on MAPL measurements (Patton, Milando, Durant, & Kumar, 2017). They do not attempt to quantify the potential biases in the RLINE model output. While Zhai et al. (2016) do quantify the calibration bias in the RLINE output for NO_x, $PM_{2.5}$, and carbon monoxide (CO), they use a linear regression approach with independent error terms, without accounting for the spatial/spatiotemporal structure of the residual process.

This paper is organized as follows. Section 2 provides details on the monitored data and the RLINE computer model output. Section 3 gives details of the proposed nonstationary joint Bayesian data fusion model. Section 4 presents results from fitting the model to our data as well as a comparison with various special cases of our model including stationary models. Finally, Section 5 provides concluding remarks and future directions.

2 | DATA

2.1 | Monitor data

A MAPL, a recreational vehicle equipped with air monitoring instruments, measured ambient concentrations of nitrogen oxides (NO_x) and fine particulate matter ($PM_{2.5}$) along nine transects that crossed major roadways in Detroit, Michigan, on seven consecutive days in December 2012 during the morning and afternoon rush hour periods (Figure 1a). Details of the sampling are described in another report analyzing factors that affect pollutant concentrations in the near-road environment (Baldwin et al., 2015). In short, at each of the nine "areas," pollutant concentrations were sampled in residential neighborhoods adjacent to the major roadways. In areas 1 through 8, the sampling sites were located at nominal distances of 50 m (two sites), 150 m (two sites), and 500 m (one site) from both edges of the roads. In these areas, the 10 sampling sites formed roughly two parallel transects across the roads. On the other hand, in area 9, six sampling sites were used at distances of 50, 150, and 500 m, respectively, from both edges of the road.

Areas 1 through 8 were monitored during both the morning (07:15–09:45) and afternoon (15:45–18:15) rush hour periods, whereas area 9 was monitored once or twice each day, just after the morning rush hour or before the afternoon rush hour. Sites were monitored from the morning of Friday, December 14, 2012, to the morning of Thursday, December 20, 2012. On any given day, up to three different areas were monitored. The MAPL visited the monitoring sites and collected air samples for 5 min at each site. NO_x concentrations were measured using conventional federal reference monitors (Model 42i, Thermo, MA, USA), whereas particle number concentrations in multiple size bins were measured with a GRIMM Model 1.109 Spectrometer (GRIMM Aerosol Technik, Germany). The particle number concentrations were converted into mass concentrations using the protocol described by Grimm and Eatough (2009), assuming a particle density of 1.67 g/cm³. $PM_{2.5}$ concentrations were calculated as the sum of mass concentrations for particles with diameters less than or equal to 2.5 μm. For each pollutant, the "background" concentration for a transect area was determined as the lowest pollutant concentration observed in that area during the monitoring period. While different methods could be used to estimate background concentration, including using observations from regulatory monitors, we selected our approach
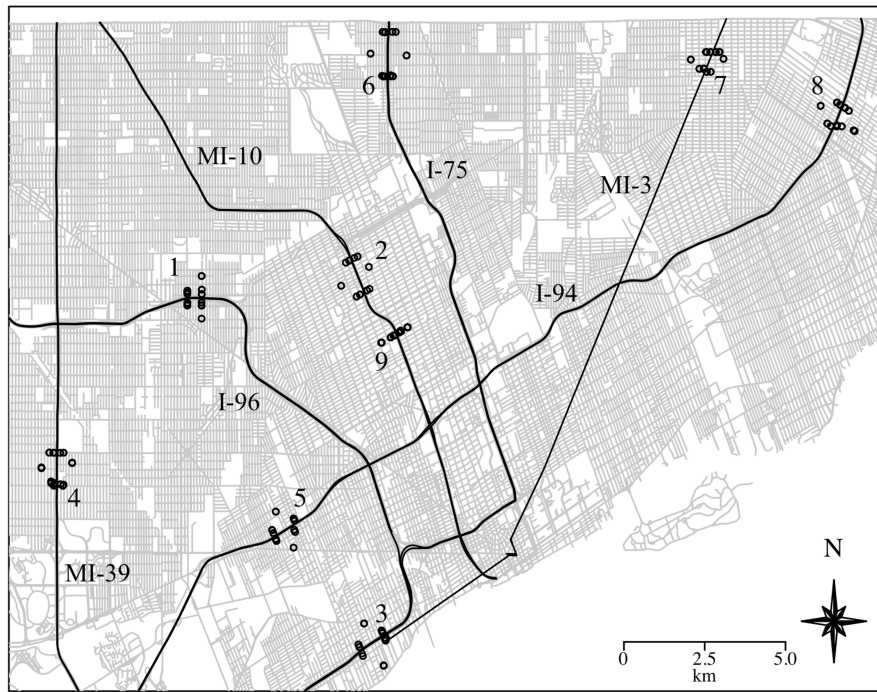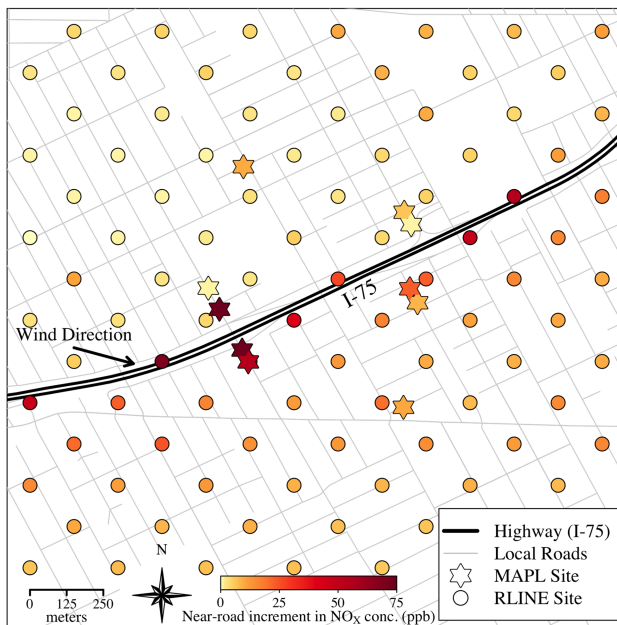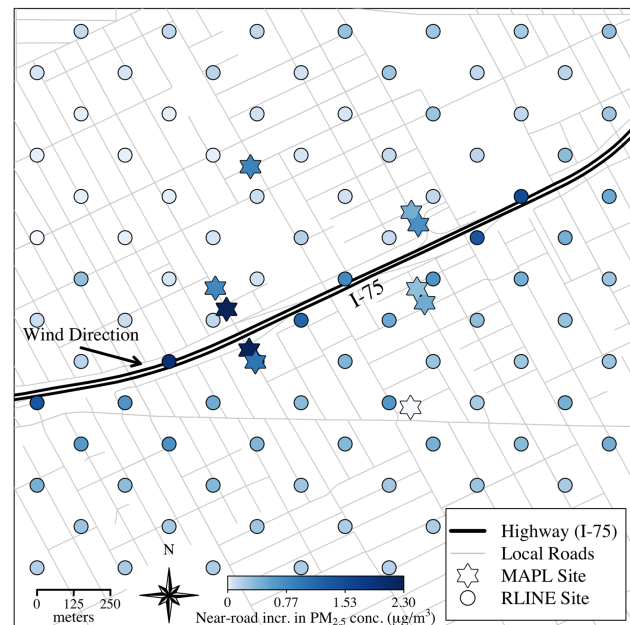
FIGURE 1 (a) Nine transect areas and monitoring sites in Detroit, Michigan. (b) Near-road increment in NO_x concentrations (in parts per billion [ppb]) at mobile air pollution lab (MAPL) and research line-source dispersion model (RLINE) sites in area 5 on the morning of December 18, 2012. (c) Near-road increment in $PM_{2.5}$ concentrations (in $\mu g/m^3$) at MAPL and RLINE sites in area 5 on the same day

so as to account for urban local sources near the transect. Regulatory monitors are typically designed to capture regional sources of pollution and, thus, might not be representative of the urban environment.

The NRI in pollutant concentration at a monitoring site was defined as the difference between the observed and background pollutant concentrations. By this definition, for NO_x, across the nine areas, a total of 254 5-min average NRI in concentrations were collected, whereas 235 average NRI in concentrations were sampled for $PM_{2.5}$. Figures 1b and 1c show NO_x and $PM_{2.5}$ NRI in pollutant concentrations, respectively, at sites within area 5 on the morning of December 12, 2012.

Initial exploratory data analysis revealed that there was considerable variability in the NRI in concentrations of the two pollutants between the different areas. The mean NRI in concentrations of both pollutants was lowest at area 9 (9.6 ppb [parts per billion] for NO_x ($SD = 8.5$ ppb) and 0.5 μg/m^3 ($SD = 0.4$ μg/m^3) for $PM_{2.5}$). This was expected since area 9 sites were sampled either just after the morning rush hour or before the afternoon rush hour period, whereas the other sites were monitored during the morning and afternoon rush hour periods. The highest mean NRI in concentrations of NO_x was recorded at area 5 (21.5 ppb ($SD = 18.7$ ppb)) and at area 1 for $PM_{2.5}$ (2.2 μg/m^3 ($SD = 2.2$ μg/m^3)). Histograms of the NRI in pollutant concentrations for both pollutants displayed right-skewed distributions, as shown in Figure 1 in the Supplementary Material, which suggests that modeling the NRI in pollutant concentrations on the log scale would be more appropriate. A time-series plot of the mean of the log of NRI in pollutant concentrations for sites monitored within the same time period (before or after 12 noon on the same day) is given in Figure 2a for NO_x and in Figure 2b for $PM_{2.5}$.

2.2 | RLINE data

Concentrations of NO_x and $PM_{2.5}$ in the near-road environment attributable to on-road vehicular traffic were modeled using a modified version of RLINE (Snyder et al., 2013). The dispersion mechanisms were unchanged, but the model was modified to incorporate Motor Vehicle Emission Simulator emission factors (grams per vehicle mile) that vary by vehicle class, ambient temperature and month, and used reported traffic volumes for each modeled roadway adjusted by hourly temporal allocation factors (Batterman, Cook, & Justin, 2015) for commercial and noncommercial traffic.

At each transect area, RLINE model estimates for average hourly pollutant concentrations were obtained at 96 regularly spaced point locations within 2 km^2 centered around the major road, during the same time period that the MAPL data were recorded. This selection of RLINE prediction sites allowed for some buffer around the MAPL monitoring sites, in addition to providing data at sites relatively far from the highway to better estimate the spatial range for the spatially varying additive calibration bias (detailed in Section 3.1). Since in data fusion models, the number of observations from each data source can control the relative influence of the data source on the estimation of the underlying true process, 96 RLINE sites for 10 MAPL sites provided a reasonable balance between the relative influence of each data source, the maximum distance from the highway, and the distance between neighboring points for the RLINE prediction. Across the nine areas, a total of 3,038 1-hr average NRI in concentrations were estimated for both NO_x and $PM_{2.5}$. Figures 1b and 1c show NO_x and $PM_{2.5}$ NRI in pollutant concentrations, respectively, at both the MAPL sites and RLINE estimation locations within area 5 on the morning of December 12, 2012.

Similar to the MAPL data, there was considerable variability observed in the concentrations of the two pollutants from the RLINE model output between the different areas. However, unlike the MAPL observations, the lowest mean NRI in concentrations for both pollutants were observed at area 7 (8.0 ppb ($SD = 10.1$ ppb) for NO_x and 0.23 μg/m^3 ($SD = 0.30$ μg/m^3) for $PM_{2.5}$), whereas the highest concentrations were observed at area 6 (42.4 ppb ($SD = 52.1$ ppb) for NO_x and 1.2 μg/m^3 ($SD = 1.5$ μg/m^3) for $PM_{2.5}$). Figure 3a shows a comparison of the log of NRI in monitored NO_x concentrations against RLINE model predictions at the nearest monitor site (the dotted line has slope 1 and intercept 0), while Figure 3b gives a similar plot for $PM_{2.5}$ concentrations. These plots suggest that the RLINE model tends to provide higher predictions for NRI in NO_x concentrations as compared to the MAPL data, whereas the predictions for $PM_{2.5}$ tend to be slightly lower than monitored concentrations.

2.3 | Covariate data

In addition to the monitored and estimated pollution concentrations, we also obtained data on traffic and meteorological factors that may help explain some of the variability in the pollutant concentrations at these areas. Wind speed and wind direction data were obtained from the National Oceanic and Atmospheric Administration (Quality Controlled Local Climatological Data: http://cdo.ncdc.noaa.gov/qclcd/QCLCD) for four airport sites in the study area: Detroit City Airport, Detroit Metro Airport, Grosse Pointe Farms, and St. Clair Shores. To obtain representative data, hourly wind direction and wind speed were averaged across the four airport sites. Additionally, we obtained annual average daily traffic (AADT) and commercial AADT (CAADT) volumes for major roads in Detroit from the Michigan Department of Transportation Traffic Monitoring Information System (2014). AADT and CAADT were adjusted to hourly values using Detroit-specific temporal allocation factors (Batterman et al., 2015), with noncommercial AADT (NCAADT) defined as AADT − CAADT. Commercial traffic, mostly using heavy-duty diesel, tends to contribute a much greater proportion to atmospheric pollutant levels compared to noncommercial traffic, which operates mostly on light-duty gasoline
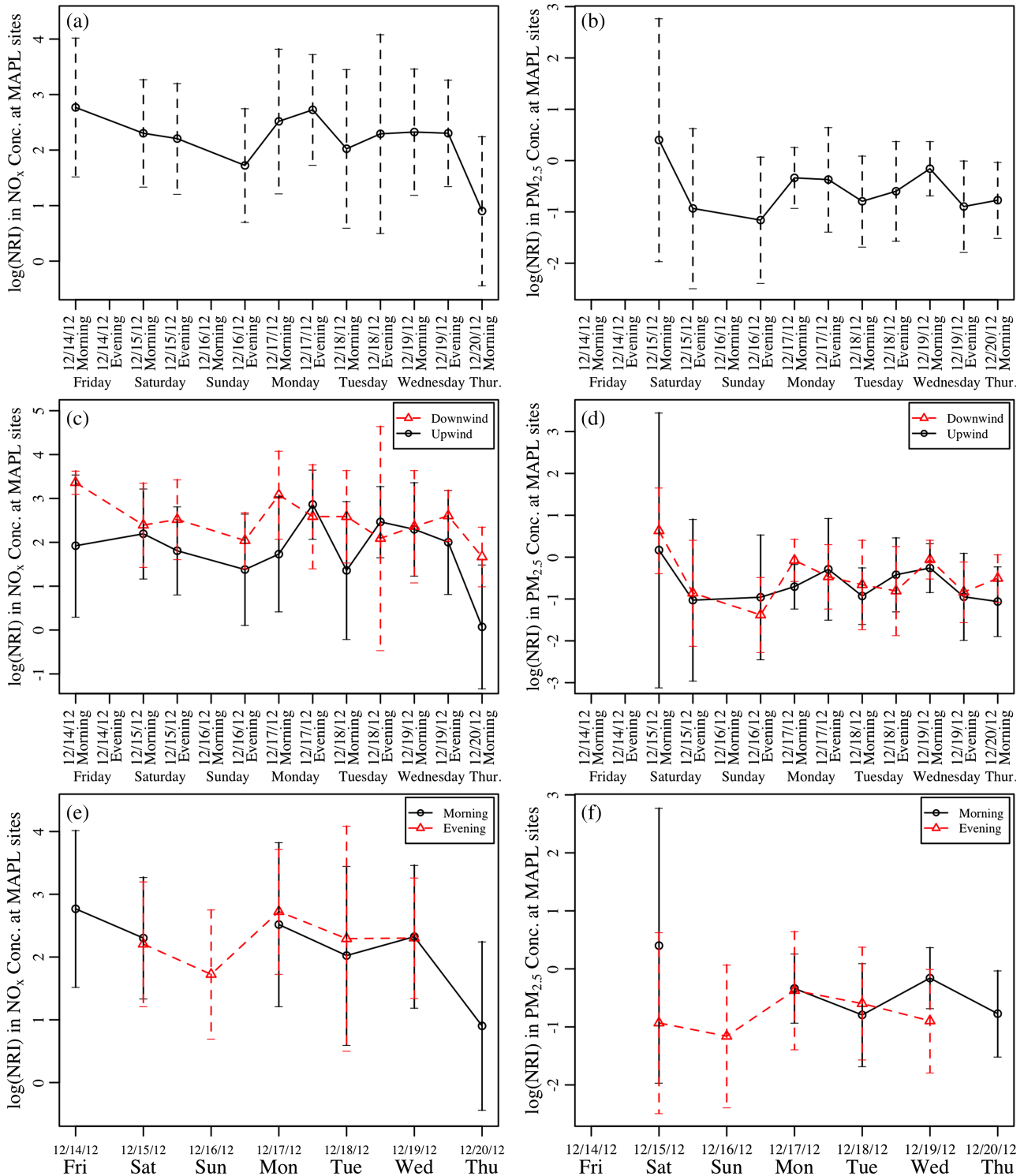
FIGURE 2 Mean log(NRI) in (a) NO_x concentration (in log ppb) and (b) $PM_{2.5}$ concentration (in log $\mu g/m^3$) at MAPL sites at each time period. The mean log(NRI) concentrations are stratified by downwind versus upwind in (c) for NO_x concentration (in log ppb) and in (d) for $PM_{2.5}$ concentration (in log $\mu g/m^3$) and by morning versus evening in (e) for NO_x concentration (in log ppb) and in (f) for $PM_{2.5}$ concentration (in log $\mu g/m^3$). The vertical bars represent ± 1 standard deviation from the mean. NRI = near-road increment; MAPL = mobile air pollution lab

(Batterman et al., 2015; Watkins & Baldauf, 2012). To more accurately account for the relative contribution of these vehicles to air pollution, a composite "traffic volume" was calculated as NCAADT + c·CAADT. In our study, following Baldwin et al. (2015), we set parameter c to 10 for NO_x and to 50 for $PM_{2.5}$.
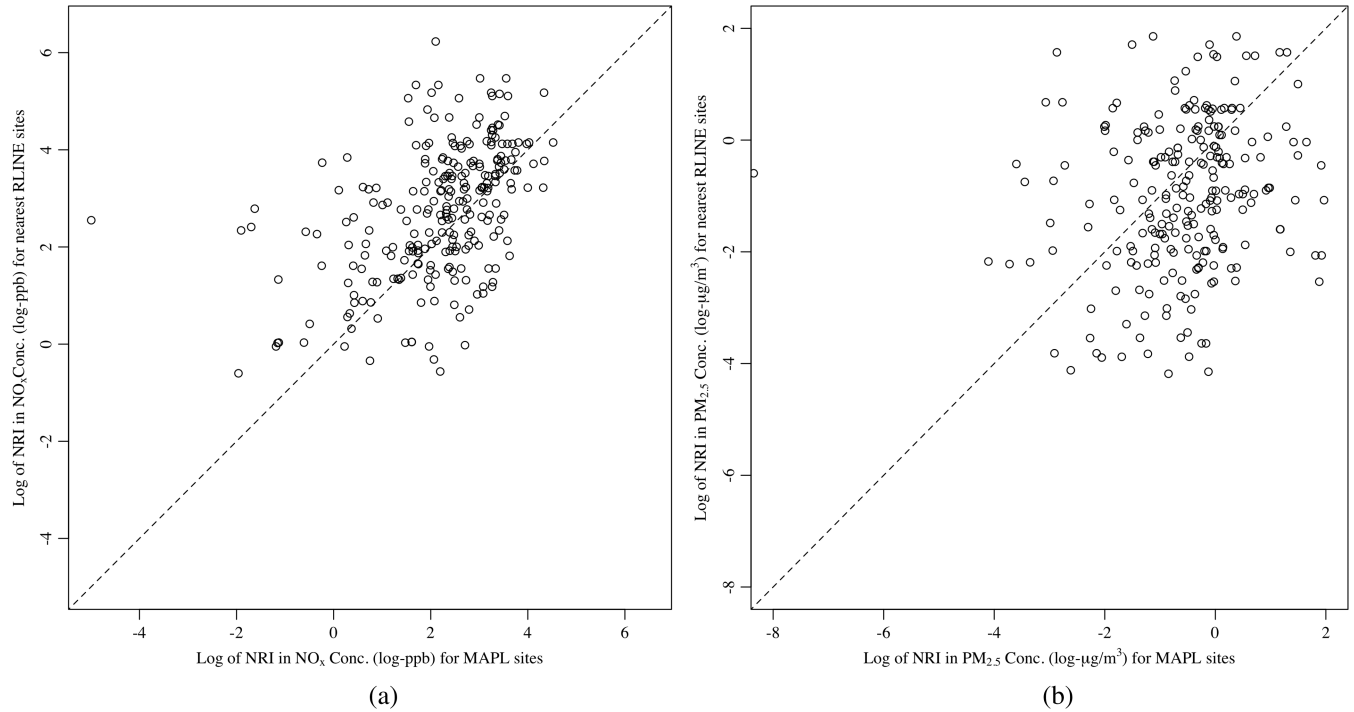
FIGURE 3 Scatterplot of natural log of near-road increment (NRI) in monitored (a) NO_x concentrations (in log ppb) and (b) $PM_{2.5}$ concentrations (in log $\mu g/m^3$) against research line-source dispersion (RLINE) model prediction at the nearest site. MAPL = mobile air pollution lab

3 | MODEL

In this section, we present the proposed nonstationary Bayesian data fusion model for NRIs in pollutant concentrations of NO_x and $PM_{2.5}$, which allows us to both evaluate the RLINE model output and provide predictions at unsampled locations along with a measure of uncertainty associated with the predictions. Our proposed model utilizes the joint modeling approach, whereas we capture the nonstationary behavior observed in the concentration of pollutants in the near-road environment using a recently developed approach by Gilani et al. (2016a), which expresses pollutant concentrations as a mixture of two independent spatial processes, each equipped with a nonstationary covariance function with covariates driving the nonstationarity and the mixture weights.

3.1 | Nonstationary spatiotemporal Bayesian data fusion model

For each pollutant $k, k \in \{NO_x, PM_{2.5}\}$, let t represent "time periods," where a time period is defined by day and time of day (before or after 12:00) of pollution monitoring. By this definition, there were $\mathcal{T} = 11$ distinct time periods during the study data collection. As is customary in the joint modeling approach, we do not consider the monitored MAPL measurements or the RLINE model output as the *truth* since these observations/estimates have measurement errors and/or biases associated with them. Instead, for each pollutant k, we assume that there exists an underlying (unobserved) spatiotemporal field $Y_t^{(k)}(\mathbf{s})$ that represents the *true* NRI in concentration at location $\mathbf{s} \in S$ and time period $t = 1, \dots, \mathcal{T}$. Working on the log scale, let $Z_t^{(k)}(\mathbf{s}) = \log(Y_t^{(k)}(\mathbf{s}))$ denote the natural log of the *true* NRI in the concentration of pollutant k, at location $\mathbf{s}$ and time period t. For each pollutant k, we decompose $Z_t^{(k)}(\mathbf{s})$ in the sum of two terms, that is,

$$Z_t^{(k)}(\mathbf{s}) = \mu_t^{(k)}(\mathbf{s}) + \eta_t^{(k)}(\mathbf{s}), \qquad (1)$$

where $\mu_t^{(k)}(\mathbf{s})$ accounts for the large-scale spatial trend in the kth pollutant log NRI concentration at location $\mathbf{s}$ and time period t, and $\eta_t^{(k)}(\mathbf{s})$ is a spatially correlated, mean-zero error term.

At a MAPL monitoring site $\mathbf{s}$ and time period t, we denote the measured log NRI concentration of pollutant k by $\hat{Z}_t^{(k)}(\mathbf{s})$ and assume that it is measured with some error, that is,

$$\hat{Z}_t^{(k)}(\mathbf{s}) = Z_t^{(k)}(\mathbf{s}) + e_t^{(k)}(\mathbf{s}), \qquad e_t^{(k)}(\mathbf{s}) \stackrel{\text{iid}}{\sim} N\left(0, \tau_e^{2(k)}\right), \tag{2}$$

where $e_t^{(k)}(\mathbf{s})$ represents the measurement error at site $\mathbf{s}$ and time period t, independent of the true underlying process $Z_t^{(k)}(\mathbf{s})$.

For the RLINE model output, we assume that the log NRI concentration estimate for pollutant k at location $\mathbf{s}$ and time t, denoted by $\widetilde{Z}_t^{(k)}(\mathbf{s})$, has both additive and multiplicative biases associated with it. These biases may be spatially varying or constant; while previous studies have shown evidence of spatially varying additive bias, the multiplicative bias is generally modeled as constant in space and time (Berrocal et al., 2010b). Exploratory data analysis (such as Figure 3) also suggests that while there does appear to be some multiplicative bias for the RLINE model output, it does not vary across space and time. We therefore model the log NRI concentration from the RLINE model output as

$$\widetilde{Z}_t^{(k)}(\mathbf{s}) = a_0^{(k)} + a_t^{(k)}(\mathbf{s}) + b^{(k)} \cdot Z_t^{(k)}(\mathbf{s}) + \delta_t^{(k)}(\mathbf{s}), \tag{3}$$

where $a_0^{(k)}$ is a constant (over space and time) additive bias, $a_t^{(k)}(\mathbf{s})$ is a time series of serially independent Gaussian spatial processes, $b^{(k)}$ is a multiplicative bias that is constant over space and time, and $\delta_t^{(k)}(\mathbf{s})$ is a mean-zero, independent across time and space, error process, that is, $\delta_t^{(k)}(\mathbf{s}) \stackrel{\text{iid}}{\sim} N(0, \tau_\delta^{2(k)})$. At each time period t, the spatial covariance for the spatially correlated additive bias $a_t^{(k)}(\mathbf{s})$ is modeled using an exponential covariance function so that the covariance between two sites $\mathbf{s}_i$ and $\mathbf{s}_j$ at time period t is given by $\text{Cov}(a_t^{(k)}(\mathbf{s}_i), a_t^{(k)}(\mathbf{s}_j)) = \sigma_a^{2(k)} \exp\{-\phi_a^{(k)}|\mathbf{s}_i - \mathbf{s}_j|\}$. With this modeling choice, we assume that the additive bias of the RLINE output is similar at nearby locations, but less so at distant ones, with a correlation that decays exponentially with distance.

Earlier work on modeling pollutant concentrations in the near-road environment in Detroit (Baldwin et al., 2015; Gilani et al., 2016a) has revealed that observed pollutant concentrations are dependent on day of the week, time of the day, being upwind versus downwind of major highways, meteorology, traffic activity and composition, and distance to the highway. Figure 2 shows a comparison of log NRI in MAPL concentrations between upwind versus downwind sites and between morning versus evening for both pollutants. Boxplots comparing the distribution of log NRI in MAPL concentrations between upwind and downwind sites, between morning and evening, and between weekday and weekend for both pollutants are given in Figure 2 in the Supplementary Material. Although these plots do not reveal much difference in the log NRI pollutant concentrations between downwind and upwind sites, between morning and evening, and between weekdays and weekends, for consistency with earlier work (Gilani et al., 2016a), we model the mean spatial trend $\mu_t^{(k)}(\mathbf{s})$ at location $\mathbf{s}$ and time period t for each pollutant k as

$$\begin{aligned}
\mu_t^{(k)}(\mathbf{s}) = {} & \beta_0^{(k)} + \beta_1^{(k)}\text{Down}_t(\mathbf{s}) + \beta_2^{(k)}\text{Weekday}_t + \beta_3^{(k)}\text{Morning}_t + \beta_4^{(k)}\text{Weekday}_t * \text{Morning}_t \\
& + \beta_5^{(k)}\left(\frac{\text{Traffic}_t(\mathbf{s})}{\text{Wind Speed}_t(\mathbf{s})}\right) + \beta_6^{(k)}\left(\frac{\text{Traffic}_t(\mathbf{s})}{\text{Wind Speed}_t(\mathbf{s})}\right) * \text{Down}_t(\mathbf{s}) + \beta_7^{(k)}\text{Distance}(\mathbf{s}) \\
& + \beta_8^{(k)}\text{Distance}(\mathbf{s}) * \text{Down}_t(\mathbf{s}),
\end{aligned} \tag{4}$$

where $\text{Down}_t(\mathbf{s})$ is an indicator for site $\mathbf{s}$ at time t being downwind (versus upwind) of the highway, Weekday_t is an indicator for time period t falling on a weekday (versus weekend), Morning_t is an indicator for morning (before 12:00) versus evening at time period t, $\frac{\text{Traffic}_t(\mathbf{s})}{\text{Wind Speed}_t(\mathbf{s})}$ is a scalar that quantifies the amount of traffic on the nearest highway around $\mathbf{s}$ during time period t normalized by the wind speed at site $\mathbf{s}$ during time period t, and $\text{Distance}(\mathbf{s})$ records the distance of $\mathbf{s}$ from the nearest highway. This formulation allows different intercepts at downwind and upwind sites and different gradients at downwind versus upwind sites for the covariates related to traffic normalized by wind speed ($\frac{\text{Traffic}_t(\mathbf{s})}{\text{Wind Speed}_t(\mathbf{s})}$) and distance from the nearest highway ($\text{Distance}(\mathbf{s})$). We also assume that the temporal variability of log NRI concentration $Z_t^{(k)}(\mathbf{s})$ is adequately captured by the time-varying covariates in $\mu_t^{(k)}(\mathbf{s})$, so that the residual error term $\eta_t^{(k)}(\mathbf{s})$ is purely spatial in nature and can be modeled as independent realizations over time periods of a mean-zero spatial Gaussian process.

Previous studies modeling concentrations of atmospheric pollutants have observed differing patterns of spatial dependence depending on the pair of locations considered (see Calder, 2007; Fuentes, 2001; Higdon, 1998; Paciorek & Schervish, 2006, among others), which indicates that we should model the residual process $\eta_t^{(k)}(\mathbf{s})$ as a nonstationary spatial process. Recent work in modeling concentrations of various pollutants in the near-road environment (Gilani et al., 2016a) has further indicated that certain spatiotemporally varying covariates, such as wind speed, might be important covariates that help account for the different spatial dependency patterns observed in a localized spatial domain that includes sites that

are both upwind and downwind of major highways. Figure 3 in the Supplementary Material shows the empirical semi-variogram of the residuals obtained from a linear regression model fitted to log NRI concentrations for the two pollutants separately using the mean function given in Equation (4), stratified by upwind and downwind sites. The Figure provides evidence of a nonstationary residual process. Therefore, in this study, for each pollutant k, we model the term $\eta_t^{(k)}(\mathbf{s})$ as a nonstationary spatial process using a recently developed approach by Gilani et al. (2016a), which builds upon the kernel mixing approach of Fuentes (2001) and a method proposed by Schmidt et al. (2011) that allows inclusion of covariates in a covariance function. The new approach expresses pollutant concentrations as a mixture of two independent spatial processes, each equipped with a nonstationary covariance function with covariates driving the nonstationarity and the mixture weights.

Following the specification of Gilani et al. (2016a), for each pollutant k, we model $\eta_t^{(k)}(\mathbf{s})$, $\quad t = 1, \dots, \mathcal{T}$, as independent realizations in time of a weighted mixture of two mutually independent mean-zero Gaussian spatial processes $\eta_{1,t}^{(k)}(\mathbf{s})$ and $\eta_{2,t}^{(k)}(\mathbf{s})$, independent over time, with nonstationary covariance functions $C_{\boldsymbol{\theta}_1^{(k)}}(\cdot, \cdot)$ and $C_{\boldsymbol{\theta}_2^{(k)}}(\cdot, \cdot)$, respectively. That is, for each time period t and site $\mathbf{s} \in \mathcal{S}$,

$$\eta_t^{(k)}(\mathbf{s}) = w_{1,t}(\mathbf{s}) \cdot \eta_{1,t}^{(k)}(\mathbf{s}) + w_{2,t}(\mathbf{s}) \cdot \eta_{2,t}^{(k)}(\mathbf{s}), \tag{5}$$

where the spatially and temporally varying weights $w_{1,t}(\mathbf{s})$ and $w_{2,t}(\mathbf{s})$ add up to 1, which is achieved by setting

$$w_{i,t}(\mathbf{s}) = \frac{\widetilde{w}_{i,t}(\mathbf{s})}{\widetilde{w}_{1,t}(\mathbf{s}) + \widetilde{w}_{2,t}(\mathbf{s})}, \qquad i = 1, 2.$$

The un-normalized weights $\widetilde{w}_{1,t}(\mathbf{s})$ and $\widetilde{w}_{2,t}(\mathbf{s})$ are specified in such a way that when site $\mathbf{s}$ is downwind of the highway at time t, the spatial process $\eta_{1,t}(\mathbf{s})$ receives a larger weight as compared to $\eta_{2,t}(\mathbf{s})$, and vice versa. One weighting scheme that achieves this is given by

$$\widetilde{w}_{i,t}(\mathbf{s}) = \begin{cases} \exp\left(-\psi \cdot (1 - \text{Down}_t(\mathbf{s})) \cdot |\mathbf{s} - \mathbf{s}^*|\right), & i = 1 \\ \exp(-\psi \cdot (\text{Down}_t(\mathbf{s})) \cdot |\mathbf{s} - \mathbf{s}^*|), & i = 2, \end{cases} \tag{6}$$

where $\mathbf{s}^*$ is defined as the projection of site $\mathbf{s}$ onto the further edge of the highway.

To model the two mutually independent spatial processes $\eta_{1,t}^{(k)}(\mathbf{s})$ and $\eta_{2,t}^{(k)}(\mathbf{s})$ for each pollutant k as nonstationary processes, where their nonstationarity is driven by a spatial covariate (e.g., wind speed), we define a nonlinear 1–1 smooth transformation $\widetilde{\mathbf{f}}_t(\cdot)$ from the *geographical* space $\mathcal{S} \subset \mathbf{R}^2$ to the *deformation* space $\mathcal{D} \subset \mathbf{R}^3$ that, for each time period t, maps a site $\mathbf{s} = (\mathbf{s}_x, \mathbf{s}_y) \in \mathcal{S}$ onto $\widetilde{\mathbf{f}}_t(\mathbf{s}) = (\mathbf{s}_x, \mathbf{s}_y, h_t(\mathbf{s}))$, where $h_t(\mathbf{s})$ is a spatial covariate. In this formulation, for any pair of sites $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$,

$$\text{Cov}\left(\eta_{i,t}^{(k)}(\mathbf{s}), \eta_{i,t}^{(k)}(\mathbf{s}')\right) = C_{\mathcal{D},i}^{(k)}\left(\widetilde{\mathbf{f}}_t(\mathbf{s}), \widetilde{\mathbf{f}}_t(\mathbf{s}')\right) = C_{\mathcal{D},i}^{(k)}\left(|\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')|\right), \quad i = 1, 2,$$

where $C_{\mathcal{D},i}^{(k)}(\cdot, \cdot)$ is a stationary, isotropic covariance function (e.g., a Matérn covariance function) on $\mathcal{D}$, and $|\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')|$ is the Mahalanobis distance between $\widetilde{\mathbf{f}}_t(\mathbf{s})$ and $\widetilde{\mathbf{f}}_t(\mathbf{s}')$. Schmidt et al. (2011) show that while the spatial processes $\eta_{1,t}^{(k)}(\widetilde{\mathbf{f}}_t(\mathbf{s}))$ and $\eta_{2,t}^{(k)}(\widetilde{\mathbf{f}}_t(\mathbf{s}))$ in the deformation space $\mathcal{D} \subset \mathbf{R}^3$ are stationary and isotropic (since the correlation between any two vectors $\widetilde{\mathbf{f}}_t(\mathbf{s})$ and $\widetilde{\mathbf{f}}_t(\mathbf{s}')$ is dependent only on the distance $|\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')|$), the spatial processes $\eta_{1,t}^{(k)}(\mathbf{s})$ and $\eta_{2,t}^{(k)}(\mathbf{s})$ in the geographic space $\mathcal{S} \subset \mathbf{R}^2$ are nonstationary.

Since we believe that the different spatial dependence behavior in the log NRI concentrations exhibited between sites that are upwind and downwind of major highways is driven by local meteorological factors, we define $h_t(\mathbf{s})$ as signed wind speed at $\mathbf{s}$ at time period t, that is, wind speed with a positive sign if $\mathbf{s}$ is downwind at time t and with a negative sign otherwise. Finally, for each pollutant k and time period t, we model the covariance function of the underlying process $\eta_{i,t}^{(k)}(\mathbf{s})$ using an exponential covariance function with distance given by the Mahalanobis distance, so that for any pair of sites $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$,

$$\text{Cov}\left(\eta_{i,t}^{(k)}(\mathbf{s}), \eta_{i,t}^{(k)}(\mathbf{s}')\right) = \sigma_i^{2^{(k)}} \cdot \exp\left\{-\sqrt{\left(\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')\right)^T \cdot \boldsymbol{\Phi}_i^{(k)} \cdot \left(\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')\right)}\right\}, \quad i = 1, 2. \tag{7}$$

For each $i = 1, 2$ and each pollutant k, the matrix $\boldsymbol{\Phi}_i^{(k)}$ is a 3×3 diagonal matrix with diagonal elements $(\phi_i^{(k)}, \phi_i^{(k)}, \varphi_i^{(k)})$ and off-diagonal elements equal to 0, where $\phi_i^{(k)}$ controls the range or smoothness of the covariance function along the x–y direction and $\varphi_i^{(k)}$ controls it along the signed wind speed direction.

Under the above specification, the covariance function of the residual spatial process $\eta_t^{(k)}(\mathbf{s})$ is a mixture of two nonstationary covariance functions with nonstationarity driven by a covariate (i.e., signed wind speed) and with spatiotemporally

varying weights that also depend on a covariate (i.e., being upwind or downwind). Thus, for each pollutant k, at time period t and sites $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$,

$$\text{Cov}\left(\eta_t^{(k)}(\mathbf{s}), \eta_t^{(k)}(\mathbf{s}')\right) = \sum_{i=1}^{2} w_{i,t}(\mathbf{s}) w_{i,t}(\mathbf{s}') \cdot \sigma_i^{2(k)} \cdot \exp\left\{-\sqrt{\left(\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')\right)^T \cdot \mathbf{\Phi}_i^{(k)} \cdot \left(\widetilde{\mathbf{f}}_t(\mathbf{s}) - \widetilde{\mathbf{f}}_t(\mathbf{s}')\right)}\right\}. \tag{8}$$

The hierarchical Bayesian data fusion model for each pollutant is completed by assigning prior distributions to the various parameters in the process model (Equation (1)) and the data models (Equations (2) and (3)), as detailed below. Denoting the vector of process model parameters for pollutant k by $\theta_p^{(k)} = \{\boldsymbol{\beta}^{(k)}, \theta_1^{(k)}, \theta_2^{(k)}\}$, the vector of data model parameters by $\theta_{\hat{Z}}^{(k)}$ and $\theta_{\widetilde{Z}}^{(k)}$, and the vector of realizations of the true process, observed data, and model output collected over all time periods by $\mathbf{Z}^{(k)}, \hat{\mathbf{Z}}^{(k)}$, and $\widetilde{\mathbf{Z}}^{(k)}$, respectively, and noting that given $\mathbf{Z}^{(k)}, \hat{\mathbf{Z}}^{(k)}$, and $\widetilde{\mathbf{Z}}^{(k)}$ are independent, the joint distribution of all of the unknowns and available data (MAPL measurements and RLINE model output) can be decomposed as

$$p\left(\hat{\mathbf{Z}}^{(k)}, \widetilde{\mathbf{Z}}^{(k)}, \mathbf{Z}^{(k)}, \theta_p^{(k)}, \theta_{\hat{Z}}^{(k)}, \theta_{\widetilde{Z}}^{(k)}\right) = p\left(\hat{\mathbf{Z}}^{(k)} \mid \mathbf{Z}^{(k)}, \theta_{\hat{Z}}^{(k)}\right) \cdot p\left(\widetilde{\mathbf{Z}}^{(k)} \mid \mathbf{Z}^{(k)}, \theta_{\widetilde{Z}}^{(k)}\right) \cdot p\left(\mathbf{Z}^{(k)} \mid \theta_p^{(k)}\right) \cdot p\left(\theta_{\hat{Z}}^{(k)}\right) \cdot p\left(\theta_{\widetilde{Z}}^{(k)}\right) \cdot p\left(\theta_p^{(k)}\right). \tag{9}$$

Posterior estimates of the model parameters $\theta_p^{(k)}, \theta_{\hat{Z}}^{(k)}$, and $\theta_{\widetilde{Z}}^{(k)}$ as well as prediction of the true underlying process $Z_t^{(k)}(\mathbf{s}_0)$ at an unsampled location $\mathbf{s}_0$ at time period t can be obtained using Markov chain Monte Carlo (MCMC) approaches.

3.2 | Prior specification

For each pollutant k, the full model contains 21 parameters: nine regression coefficients $\beta_0^{(k)}, \ldots, \beta_8^{(k)}$ as given in Equation (4); constant additive and multiplicative biases $a_0^{(k)}$ and $b^{(k)}$, respectively, as given in Equation (3); a nugget effect $\tau_e^{2(k)}$ for the MAPL data model given in Equation (2); a nugget effect $\tau_\delta^{2(k)}$ for the RLINE data model given in Equation (3); two spatial covariance parameters $\sigma_a^{2(k)}$ and $\phi_a^{(k)}$ for the spatially correlated additive bias $a_t^{(k)}(\mathbf{s})$ given in Equation (3); and six spatial covariance parameters $\sigma_1^{2(k)}, \phi_1^{(k)}, \varphi_1^{(k)}, \sigma_2^{2(k)}, \phi_2^{(k)}, \varphi_2^{(k)}$. On all of these parameters except for $\tau_e^{2(k)}$, we specify non-informative prior distributions. Following Gilani et al. (2016a), the parameter ψ in the definition of the mixture weights $w_{1,t}(\mathbf{s})$ and $w_{2,t}(\mathbf{s})$ given in (6) was kept constant and determined through empirical considerations. Specifically, we chose a value of ψ such that if a site $\mathbf{s}$ is the furthest downwind site during time period t, the mixture weight $w_{2,t}(\mathbf{s})$ is equal to about 0.05.

The nugget effect $\tau_e^{2(k)}$ provides the variability observed in the MAPL data due to measurement error. The precision of the monitoring equipment used to record concentrations of NO_x and $PM_{2.5}$ was therefore used to provide informative priors on these parameters. We place inverse gamma priors such that the prior mean of $\tau_e^{2(k)}$ is equal to 10% of the observed mean in NRI concentration and the prior variance is very small. To specify the priors on the remaining 20 pollutant-specific model parameters, we use an empirical Bayes approach. Specifically, on each of the $\beta_l^{(k)}, l = 0, \ldots, 8$, we place independent normal priors with large variances and means set equal to the estimates of the regression coefficients obtained by fitting a linear regression model on the MAPL data for each pollutant with the nine variables in Equation (4) as the covariates. On the variance component parameters $\sigma_1^{2(k)}$ and $\sigma_2^{2(k)}$, we place vague inverse gamma priors with shape and scale parameters chosen so that the prior variance is infinite and the prior mean is equal to 40% of the estimated residual variance from the linear regression model fit to get the prior means for the regression coefficients. On the variance component parameters $\sigma_a^{2(k)}$ and $\tau_\delta^{2(k)}$ for the RLINE data model given in Equation (3), we place vague inverse gamma priors with shape and scale parameters chosen such that the prior variance is infinite and the prior mean is equal to, respectively, 40% and 60% of the estimated residual variance from a linear regression model on the RLINE data for each pollutant with the same nine covariates. On the parameters $\phi_1^{(k)}, \phi_2^{(k)}, \varphi_1^{(k)}, \varphi_2^{(k)}$ controlling the roughness/smoothness of the covariance function of the underlying spatial processes $\eta_{1,t}^{(k)}$ and $\eta_{2,t}^{(k)}, t = 1, \ldots, \mathcal{T}$, in the x–y and signed wind speed direction, we place gamma priors with shape and scale parameters chosen so that the variance of the gamma distribution is large whereas the prior mean is such that the correlation is equal to 0.05 at a distance equal to 3 divided by the prior mean. Specifically, we set the prior means to be approximately 3 divided by the maximum inter–monitoring site distance and 3 divided by the maximum signed wind speed "distance." Finally, for the overall additive and multiplicative bias terms of the log RLINE output, $a_0^{(k)}$ and $b^{(k)}$, we used two vague independent normal priors with means 0 and 1, respectively, assuming a priori no bias in the log RLINE output.

For each pollutant k, we fit the full model by running an MCMC algorithm with Gibbs sampling for all the regression coefficients and variance component parameters and with Metropolis–Hastings steps to update the spatial decay parameters.

3.3 | Model comparison

For each pollutant separately, we compared the predictive performance of six joint Bayesian data fusion models that differed in (a) the approach used to model the additive bias of the log RLINE output and (b) the type of spatial dependence structure hypothesized for the two latent processes, $\eta_{1,t}^{(k)}(\mathbf{s})$ and $\eta_{2,t}^{(k)}(\mathbf{s})$. Table 1 provides a summary of the models considered. For the additive bias of the log RLINE output, three different approaches were used. The simplest model, called "Model 1," assumes that the additive error of the RLINE output in representing the true, unobserved field is constant in space; thus, $a_t^{(k)}(\mathbf{s}) \equiv 0$ in Equation (3). "Model 2" postulates that although the RLINE output has a spatially varying additive bias on the local scale (where local means a small region on the order of a couple of kilometers around the urban highway), overall, the RLINE output does not have a global additive bias (where global means over the entire study region), that is, $a_0^{(k)} \equiv 0$. "Model 3" is the full model as described in Section 3.1. For each of these three models, we contrast two cases: The first case assumes that $\eta_t^{(k)}(\mathbf{s})$ is a stationary Gaussian process independent in time with mean 0 and with an exponential covariance function (with marginal variance $\sigma_Z^{2(k)}$ and spatial decay parameter $\phi_Z^{(k)}$), whereas the second case assumes that $\eta_t^{(k)}(\mathbf{s})$ is equipped with the nonstationary covariance function described in Equation (5). These cases are distinguished by appending "S" or "NS" to the model labels, respectively.

We evaluated the predictive performance of each of the six models for both pollutants through comparisons between observations and predictions at out-of-sample sites. Out of the 254 observed NRI in NO_x concentrations from the MAPL data, we randomly selected 221 observations for model fitting and held out 33 observations for model validation. Similarly, for $PM_{2.5}$, we randomly selected 202 observations in NRI concentrations from the MAPL data for model fitting and held out 33 observations for model validation. In each case, the held out observations were sampled so that we had three observations held out per time period. Since the modeling was conducted on the log scale, we assessed the predictive performance by back-transforming predictions to the original scale. Using the median of the posterior predictive distribution as the predicted value at each site, we evaluated the predictive performance of each model in terms of the *mean absolute prediction error* (MAPE), the *average length of the* 90% *prediction interval* (PI), the *empirical coverage of the* 90% *PI*, and the *average bias*. The empirical coverage of the 90% PI is used to assess whether the uncertainty in the prediction is correctly quantified: If empirical coverage is below the nominal level, assuming no bias in the predictions, the model is underestimating the variability/uncertainty in the predictions. Conversely, empirical coverage of the PIs above the nominal level indicates that the model overestimates the variability/uncertainty, assuming no bias in the predictions.

For each fitted model for both pollutants, we examined trace plots of the posterior samples for the model parameters to verify that the MCMC algorithm had converged. For all six models for both pollutants, trace plots indicated that the posterior samples for all model parameters had converged within the first 10,000 iterations. We therefore treated the first 10,000 iterations as burn-in. We also examined the autocorrelation within posterior samples for each parameter. Posterior samples for the stationary models (with fewer model parameters) exhibited weaker autocorrelation than the nonstationary models. For each model, a thinning factor was used to ensure that the posterior samples used to determine the posterior distribution had very low autocorrelation. For the stationary models for both pollutants, a thinning factor of

TABLE 1 Summary of the joint Bayesian data fusion models for log NRI of NO_x and $PM_{2.5}$

Model name	Form of additive bias for RLINE	Covariance structure of $\eta_t^{(k)}(\mathbf{s})$
Model 1-S	$a_t^{(k)}(\mathbf{s}) \equiv 0$	Stationary: exponential
Model 2-S	$a_0^{(k)} \equiv 0$	Stationary: exponential
Model 3-S	Full	Stationary: exponential
Model 1-NS	$a_t^{(k)}(\mathbf{s}) \equiv 0$	Nonstationary, as in Equation (5)
Model 2-NS	$a_0^{(k)} \equiv 0$	Nonstationary, as in Equation (5)
Model 3-NS	Full	Nonstationary, as in Equation (5)

Note. NRI = near-road increment; RLINE = research line-source dispersion model.

12 was sufficient to ensure near independence of the thinned posterior samples. For the nonstationary models for NO_x, thinning factors of 64, 29, and 31 were used for Models 1, 2, and 3, respectively, whereas for $PM_{2.5}$, thinning factors of 60, 62, and 53 were used for Models 1, 2, and 3. Each MCMC algorithm was run for a number of iterations large enough so that the posterior inference was based on posterior samples of size 1,000 after thinning for each model. While some models needed considerably long chains, we did not observe any identifiability issue with any model.

The second goal of the study, that is, evaluating whether incorporating the RLINE output in our model improves the estimation of pollutant concentration, can be assessed by examining the posterior distribution of $b^{(k)}$ in Equation (3). When $b^{(k)} \equiv 0$, the conditional distribution of $\mathbf{Z}^{(k)}$, given the MAPL and RLINE data and model parameters, is equivalent to the conditional distribution of $\mathbf{Z}^{(k)}$, given the MAPL data and model parameters only. That is, when $b^{(k)} \equiv 0$, $p(\mathbf{Z}^{(k)} | \hat{\mathbf{Z}}^{(k)}, \widetilde{\mathbf{Z}}^{(k)}, \theta_p^{(k)}, \theta_{\hat{Z}}^{(k)}, \theta_{\widetilde{Z}}^{(k)}) \equiv p(\mathbf{Z}^{(k)} | \hat{\mathbf{Z}}^{(k)}, \theta_p^{(k)}, \theta_{\hat{Z}}^{(k)})$. Therefore, if the posterior distribution of $b^{(k)}$ is centered at 0 and has a small standard deviation, there is not sufficient evidence that including the RLINE output in the model helps in improving the estimates of the true concentration of pollutant k. In fact, when $b^{(k)} = 0$, Equation (3) is simply an ordinary kriging model fitted to the RLINE predictions, whereas Equations (1) and (2) together specify a universal kriging model for the MAPL pollution concentration process. Similarly, the third goal of the study, that is, evaluating whether the RLINE output captures the spatial dependence structure of NRAP concentrations in the near-road environment, can be assessed by comparing Model 1 to Models 2 and 3. If Model 1 outperforms Models 2 and 3, then the additive bias does not have any spatial pattern, so that the RLINE model output would appear to capture the spatial variability exhibited by the true underlying pollutant concentration process. However, if Model 2 or 3 outperforms Model 1, then there is evidence of some residual spatial structure that is not captured by the RLINE model output.

4 | RESULTS

Summaries of the predictive performance statistics of the joint Bayesian data fusion models for NO_x and $PM_{2.5}$ are shown in Table 2. No single model was clearly dominant with respect to all criteria for either pollutant, that is, no single model had the lowest MAPE and the shortest PI with empirical coverage close to the nominal level.

4.1 | Nitrogen oxides

For NO_x, the stationary models provided better predictive performance than the nonstationary counterparts, and Model 2-S yielded the lowest MAPE and 90% PIs that have an empirical coverage that was the second closest to the nominal coverage, although, on average, they were wider than the 90% PIs from Model 1-S (the estimated parameters for Model 2-S are shown in Table 3). While earlier studies have suggested nonstationarity in the concentrations of pollutants such as NO_x across major highways (Gilani et al., 2016a), these results suggest that the residual process for NRI in NO_x concentrations for sites across major highways is stationary.

The estimated coefficient for the multiplicative bias $\hat{b}$ is 1.25 (95% CI [credible interval]: 1.11, 1.35), which is statistically significantly greater than 1. This estimate suggests that including the RLINE output in our model does help improve our estimate of the true NRI concentrations and confirms our initial belief that the RLINE model tends to provide higher predictions for NRI in NO_x concentrations than the true NRI concentrations.

TABLE 2 Predictive performance of the joint Bayesian data fusion models averaged across validation sites and time periods. The Table shows the mean absolute prediction error (MAPE), the average length of the 90% prediction interval (PI), the empirical coverage of the 90% PI, and the average bias. Models are defined in Table 1

	Criteria	Mod 1-S	Mod 2-S	Mod 3-S	Mod 1-NS	Mod 2-NS	Mod 3-NS
NO_x	MAPE	6.59	6.35	7.00	6.95	6.69	8.13
	Avg. length of 90% PI	23.81	28.31	32.49	108.25	33.98	300.41
	Emp. coverage of 90% PI	0.84	0.77	0.77	0.84	0.74	0.74
	Avg. bias	4.98	4.34	4.17	3.08	4.68	2.94
$PM_{2.5}$	MAPE	1.11	0.72	0.78	0.48	0.42	0.44
	Avg. length of 90% PI	4.72	7.10	7.66	3.30	4.87	4.22
	Emp. coverage of 90% PI	0.55	0.90	0.86	0.69	0.66	0.69
	Avg. bias	−0.77	−0.48	−0.53	−0.10	−0.07	−0.09

TABLE 3 Posterior medians and 95% credible interval (CI) of regression coefficients $\beta_0, \ldots, \beta_8$ and covariance parameters estimated by the best fitting models (Model 2-S for NO_x and Model 2-NS for $PM_{2.5}$). Traffic is in adjusted number of vehicles per hour, wind speed is in meters per second, and distance is in kilometers

Parameter		NO_x Median	95% CI	$PM_{2.5}$ Median	95% CI
β_0	Intercept	1.65	(1.26, 2.10)	−0.89	(−2.06, 0.51)
β_1	Down	0.96	(0.83, 1.11)	1.37	(−0.65, 3.36)
β_2	Weekday	0.57	(0.03, 1.15)	−0.25	(−1.57, 1.10)
β_3	Morning	0.32	(−0.43, 1.07)	1.06	(−0.88, 3.05)
β_4	Weekday*Morning	−0.77	(−1.67, 0.06)	−0.63	(−2.79, 1.59)
β_5	Traffic/WS (10^{-5})	5.78	(2.71, 8.63)	14.1	(4.79, 24.3)
β_6	Traffic/WS*Down (10^{-5})	−4.54	(−6.04, −3.13)	−15.7	(−27.0, −5.22)
β_7	Distance (10^{-4})	−15.1	(−17.8, −12.6)	−12.9	(−19.2, −6.88)
β_8	Distance*Down (10^{-4})	−3.01	(−6.28, 0.15)	−1.49	(−8.44, 5.40)
a_0	Additive bias (RLINE)	–	–	–	–
b	Multiplicative bias (RLINE)	1.25	(1.11, 1.35)	0.66	(0.59, 0.75)
τ_e^2	Nugget (MAPL)	0.42	(0.42, 0.42)	0.06	(0.06, 0.06)
τ_δ^2	Nugget (RLINE)	0.07	(0.05, 0.09)	0.07	(0.06, 0.09)
σ_1^2	Sill - Downwind	–	–	6.34	(2.70, 13.0)
ϕ_1	Downwind	–	–	0.005^a	(0.001, 0.030)
φ_1	Downwind, Wind speed	–	–	65.8^b	(6.18, 353)
σ_2^2	Sill - Upwind	–	–	8.90	(5.06, 14.3)
ϕ_2	Upwind	–	–	0.07^c	(0.02, 0.19)
φ_2	Upwind, Wind speed	–	–	256^d	(64.5, 765)
σ_Z^2	Sill	0.49	(0.35, 0.69)	–	–
ϕ_Z	Decay	1.01^e	(0.65, 1.52)	–	–
σ_a^2	Sill (Additive bias - RLINE)	0.60	(0.43, 0.92)	1.01	(0.68, 1.69)
ϕ_a	Decay (Additive bias - RLINE)	0.78^f	(0.44, 1.26)	0.15^g	(0.07, 0.26)

Note. Effective range for spatial correlation: [a] 42.4 m/s; [b] 0.37 m/s; [c] 11.3 km; [d] 0.19 m/s; [e] 2.97 km; [f] 3.85 km; [g] 7.75 km. RLINE = research line-source dispersion model; MAPL = mobile air pollution lab.

The superior performance of Model 2 (with a mean-zero spatially varying additive bias term for the RLINE model) suggests that while the additive bias, on average, is negligible, there is some spatial structure to the bias term, so that the RLINE output does not fully capture the spatial dependence structure of NRI in NO_x concentrations. The estimate of the marginal variance of the spatially varying error $a_t(\mathbf{s})$ of the log RLINE output $\hat{\sigma}_a^2$ is 0.60 (0.43, 0.92), whereas the spatial correlation of this bias term decays fairly slowly, with an effective range of 3.85 km.

The estimated nugget effect $\hat{\tau}_e^2$ is 0.42 (0.42, 0.42), which provides a measure of the variability in the measurement error of the MAPL measurements. This is greater than the estimated nugget effect for the RLINE model output ($\hat{\tau}_\delta^2$ is 0.07 (0.05, 0.09)), suggesting smaller residual variability in the RLINE model output after accounting for multiplicative bias as compared to MAPL measurements.

The effective range of the stationary model error $\eta_t(\mathbf{s})$ is 2.97 km, so that the errors for sites within the same area (2 km × 2 km) are fairly correlated. This provides evidence that while the model covariates, MAPL measurements, and RLINE output, when combined together, account for the large-scale spatial variation in log NRI concentrations for NO_x, there is still some small-scale residual spatial variability that they do not account for.

As Table 3 shows, the estimated coefficient for downwind indicator ($\hat{\beta}_1$) is positive; controlling for other variables in the model, downwind NRI NO_x concentrations are, on average, exp(0.96) = 2.61 ppb greater than at upwind sites. Similarly, the estimated coefficient for weekday indicator ($\hat{\beta}_2$) is positive, whereas the morning indicator ($\hat{\beta}_3$) and the interaction term between morning and weekday ($\hat{\beta}_4$) are not significantly different from 0. This means that controlling for other variables, there is no difference in the NRI concentrations of NO_x between mornings and evenings, whereas NRI concentrations on weekdays tend to be 1.77 ppb higher, on average, than on weekends. As expected, the estimated coefficient for
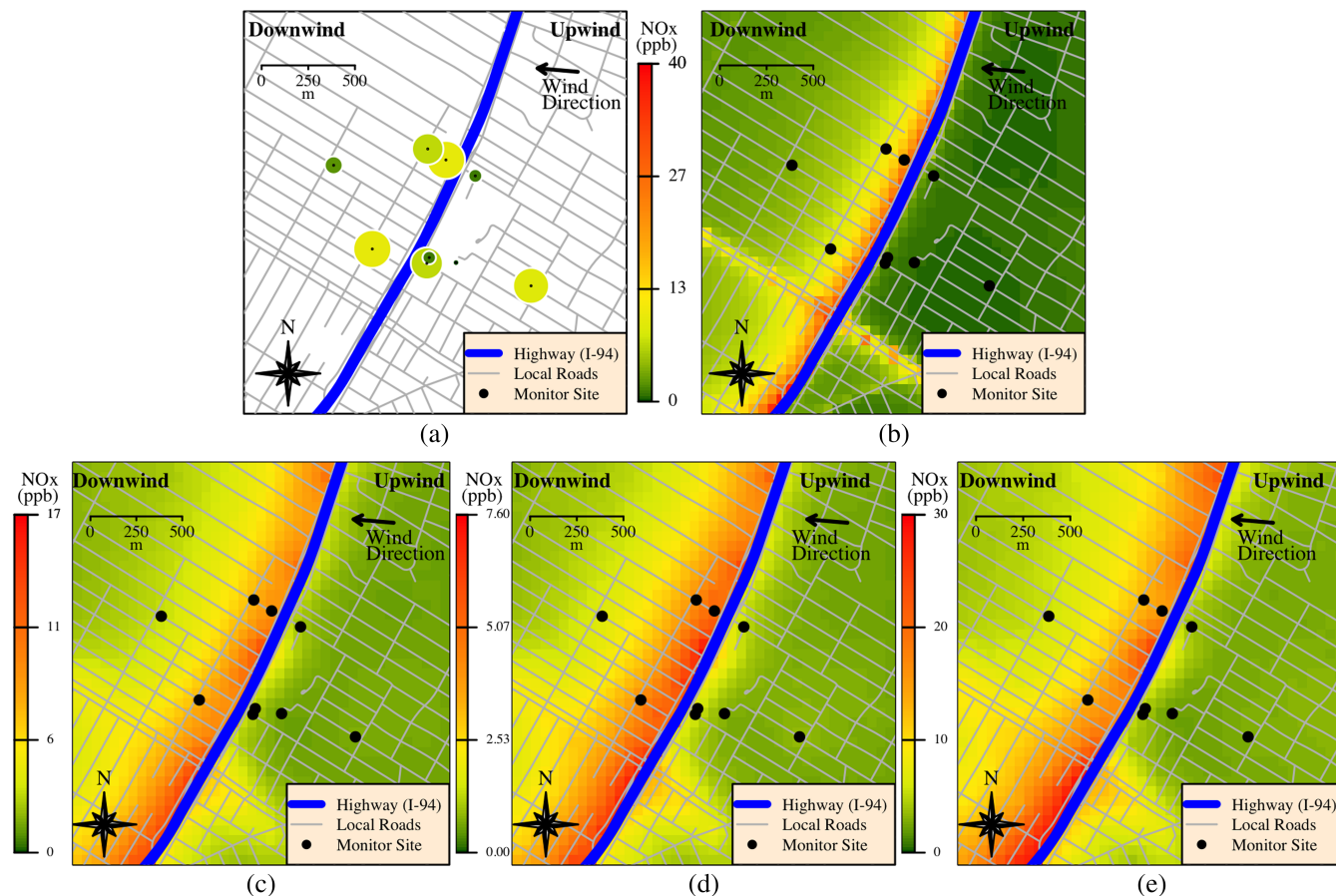
FIGURE 4 Near-road increment (NRI) in NO$_x$ concentration (in ppb) (a) monitored at mobile air pollution lab sites and (b) output from the research line-source dispersion model at area 8 near I-94 in Detroit, Michigan, at 8:00 a.m. on Thursday, December 20, 2012. (c) Predictions of hourly NRI in NO$_x$ concentration (in ppb) at area 8 at 8:00 a.m. on Thursday, December 20, 2012, as yielded by Model 2-S. (d) Lower and (e) upper bounds of the 90% prediction interval

distance to highway ($\hat{\beta}_7$) is negative, and the effect of distance is not different between upwind and downwind sites ($\hat{\beta}_8$ is not significantly different from 0).

Figure 4a shows NRI in NO$_x$ concentrations as monitored at MAPL sites at area 8 near I-94 on the morning of December 20, 2012, whereas Figure 4b shows the output from the RLINE model for the same area. The relative size of the bubbles in Figure 4a is proportional to the magnitude of the observed concentration. As noted earlier, the RLINE model output provides much higher estimates of the NRI in NO$_x$ concentrations, particularly at downwind sites close to the highway as compared to the monitored MAPL observations. Figure 4c shows predictions of hourly NRI in NO$_x$ concentrations obtained using Model 2-S for the same region and time period, whereas Figures 4d and 4e show the lower and upper bounds, respectively, of the 90% PIs. Predictions from the best fitting Bayesian data fusion model are much closer to the monitored MAPL concentrations, while they also capture the influence of secondary roads provided by the RLINE model output that was not captured by MAPL observations.

4.2 | Fine particulate matter

For PM$_{2.5}$, the situation was different: The joint Bayesian data fusion models with nonstationary covariance functions provided a strikingly better predictive performance than the stationary models. However, just as for NO$_x$, no model was better in terms of all three criteria (MAPE, average length of the 90% PI, and empirical coverage of the 90% PI). Weighting the MAPE criterion the most, Model 2-NS yielded the best predictions. The estimated coefficient for the multiplicative bias of the RLINE output $\hat{b}$ is 0.66 (95% CI: 0.59, 0.75), which suggests that for PM$_{2.5}$, just as for NO$_x$, including the RLINE output in the model helps in estimating the log NRI in true pollutant concentrations, while again confirming

our initial belief that the RLINE model tends to provide lower predictions for NRI in PM$_{2.5}$ concentrations than the true NRI concentrations.

Similar to NO$_x$, Model 2-NS hypothesized a null overall additive bias, but localized additive error at each location. The spatially varying additive bias of RLINE had a small marginal variance ($\hat{\sigma}_a^2$ is 1.01 (0.68, 1.69)) and, similar to NO$_x$, a fairly large effective range of 7.75 km (Table 3). The RLINE output for PM$_{2.5}$ does not appear to fully capture the spatial dependence structure of NRI concentrations, and the residual spatial correlation seems to persist for fairly large distances.

The estimated nugget effects for the MAPL measurements and the RLINE output are very similar ($\hat{\tau}_e^2$ is 0.06 (0.06, 0.06), and $\hat{\tau}_\delta^2$ is 0.07 (0.06, 0.09)), suggesting similar residual variability after the multiplicative error has been controlled for. The decay parameters ϕ_1 and ϕ_2 of the underlying spatial processes $\eta_{1,t}^{(k)}(\mathbf{s})$ and $\eta_{2,t}^{(k)}(\mathbf{s})$ relative to the x–y direction were fairly different, with effective ranges of 42.4 km for sites that are downwind and 11.3 km for upwind sites. Spatial correlation decays extremely slowly with an increasing distance for sites that are upwind, and even slower for downwind sites. Similarly, the decay parameters in the signed wind speed direction, φ_1 and φ_2, are also fairly different for the two underlying spatial processes, with effective ranges of 0.37 m/s ($\eta_{1,t}^{(k)}(\mathbf{s})$) and 0.19 m/s ($\eta_{2,t}^{(k)}(\mathbf{s})$). Correlation between upwind sites decays rather rapidly with increasing signed wind speed, but more slowly for downwind sites.

Examining the estimated regression coefficients in Table 3, we observe that indicators for downwind, weekday, morning, and interaction for morning and weekday are all not significantly different from 0. As expected, NRI in PM$_{2.5}$ concentrations decreases with an increasing distance from the highway, and the effect is similar between upwind and downwind sites. It is interesting to note that while NRI concentrations do not appear to differ between upwind and downwind sites (as also seen in Figure 2d), the spatial dependence structure is nonstationary, driven by wind speed and direction.

Figure 5a shows NRI in PM$_{2.5}$ concentrations as monitored at MAPL sites at area 8 near I-94 on the morning of December 20, 2012, whereas Figure 5b shows the output from the RLINE model for the same area. The relative size of the
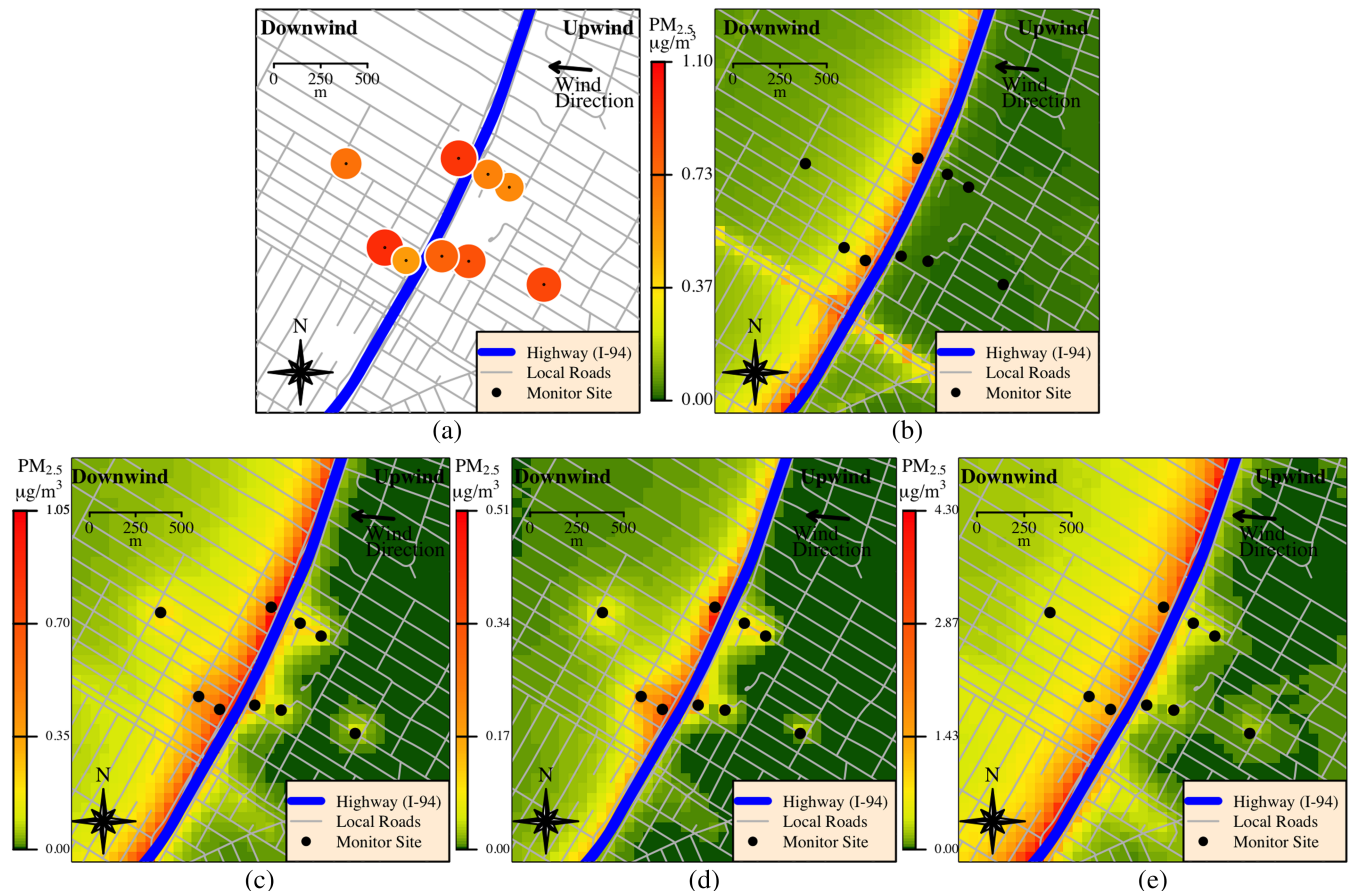

(a) (b)

(c) (d) (e)

FIGURE 5   Near-road increment (NRI) in PM$_{2.5}$ concentration (in μg/m^3) (a) monitored at mobile air pollution lab sites and (b) output from the research line-source dispersion model at area 8 near I-94 in Detroit, Michigan, at 8:00 a.m. on Thursday, December 20, 2012. (c) Predictions of hourly NRI in PM$_{2.5}$ concentration (in μg/m^3) at area 8 at 8:00 a.m. on Thursday, December 20, 2012, as yielded by Model 2-NS. (d) Lower and (e) upper bounds of the 90% prediction interval

bubbles in Figure 5a is proportional to the magnitude of the observed concentration. As noted earlier, the RLINE model output provides lower estimates of the NRI in $PM_{2.5}$ concentrations, particularly at upwind sites further away from the highway as compared to the monitored MAPL observations. Figure 5c shows predictions of hourly NRI in $PM_{2.5}$ concentrations obtained using Model 2-NS for the same region and time period, whereas Figures 5d and 5e show the lower and upper bounds, respectively, of the 90% PIs. Predictions from the best fitting Bayesian data fusion model are much closer to the monitored MAPL concentrations, while, just as for the predictions for NO_x, they also capture the influence of secondary roads provided by the RLINE model output that was not captured by MAPL observations.

5 | DISCUSSION

In this paper, we have presented a nonstationary spatiotemporal Bayesian data fusion model for NRIs in concentrations of two NRAPs (NO_x and $PM_{2.5}$) with the goals of (i) estimating the true underlying NRI pollutant concentration process and predicting concentrations at unsampled sites, (ii) evaluating whether the RLINE dispersion model output provides additional useful information in estimating the true process, and (iii) if the RLINE dispersion model output does provide additional useful information, evaluating whether the RLINE model adequately captures the spatial dependence structure of NRI in these pollutant concentrations. The data fusion model is specified using the joint modeling framework, with nonstationarity modeled using an approach that allows the incorporation and identification of covariates that might drive nonstationarity. Results from fitting the Bayesian data fusion models suggest that while the RLINE output provides additional information that is helpful in estimating the true unobserved NRI field for both NO_x and $PM_{2.5}$, it needs to be calibrated to be considered a good representation of the true unobserved NRI field for these pollutants. However, the bias adjustment mostly consists of multiplicative correction (or rescaling) of the RLINE output for both pollutants since the best performing model for both pollutants (in terms of MAPE) were models that did not include a spatially constant additive bias. For both pollutants, RLINE displayed a spatial error that has a small-to-moderate marginal variance and a fairly slow-decaying spatial correlation.

Predictions from the final model can provide improved estimates of NRAP concentrations for use in health studies. As noted earlier, adverse health effects due to exposure to NRAPs close to traffic sources, such as major roads and highways, differ significantly from exposure at background levels. Therefore, there is increasing interest in monitoring NRIs in pollutant concentrations. However, monitored data on the NRI in pollutant concentrations are difficult to obtain. In our study, we model NRIs in pollutant concentrations using a novel dataset on observed concentrations of pollutants in the near-road environment collected via a mobile air pollution lab (MAPL). Other studies have also attempted to obtain data on pollutants in the near-road environment. Similar spatiotemporal models might be applied to data collected in such mobile monitoring campaigns, especially if repeated measurements for the same season and time of day (e.g., rush hour) are obtained. Recently, an initiative has been launched to use a Google Street View car equipped with air pollution monitors to collect such data in the Oakland, California area (Apte et al., 2017). While such programs can provide useful data, our study was specifically designed to investigate the NRI during the critical rush hour period. Additionally, a key advantage of our modeling approach over using the RLINE model directly (which readily provides estimates of NRI in pollutant concentration) is that, in addition to accounting for the multiplicative bias in the RLINE output, predictions from our model are accompanied by appropriate quantification of uncertainty in the prediction, which is highly relevant when improving estimates of NRAP exposures in epidemiological studies given the nature of exposure measurement errors.

Our results are based on short-term measurements on transects across a variety of major roads in Detroit, collected during the wintertime morning and afternoon rush hour periods. Transect locations were selected to represent predominantly residential areas across the city, and the selected roads differ significantly with respect to traffic volume and fleet mix. In these ways, the collected data are representative and relevant to health studies using cohorts based on residence location. Areas with large industrial emission sources, which are common in the Detroit area, were avoided. Still, NO_x and $PM_{2.5}$ are emitted from numerous sources in addition to the local roads studied, including point sources, and non–road mobile area and regional sources. Results of our analysis should, therefore, be interpreted and generalized cautiously given the limited spatiotemporal nature of the observed data.

Fitting our model using additional datasets on observed pollutant concentrations would be useful for extending our findings. In particular, applications using long-term (e.g., seasonal to annual average) observations might be particularly revealing, especially since it is difficult for dispersion models to attain good performance at short averaging periods (Milando & Batterman, 2018). However, spatiotemporal models require a sufficient number of sampling locations to

estimate parameters. Unfortunately, few datasets provide the sufficient spatial coverage needed to develop and assess long-term urban-scale spatiotemporal models, and current monitoring networks do not capture the small-scale gradients of NRAP concentrations. There may be applications using satellite data, other mobile monitoring platforms, or possibly low-cost monitoring sensors if sufficient spatial resolution, accuracy, sensitivity, and selectivity for NRAPs can be achieved. In particular, new tools have been developed to cheaply monitor atmospheric pollutants, but with relatively low accuracy (Kumar et al., 2015). If the range of measurement errors for these new gauges is known, our approach can be readily used to monitor NRI in pollutant concentrations across transects more frequently, so that the RLINE model output can be calibrated over longer time durations and wider spatial domains.

REFERENCES

Apte, J. S., Messier, K. P., Gani, S., Brauer, M., Kirchstetter, T. W., Lunden, M. M., … Hamburg, S. P. (2017). High-resolution air pollution mapping with Google street view cars: Exploiting big data. *Environmental Science & Technology*, *51*(12), 6999–7008. PMID:28578585.

Baldwin, N., Gilani, O., Raja, S., Batterman, S., Ganguly, R., Hopke, P., … Hoogterp, S. (2015). Factors affecting pollutant concentrations in the near-road environment. *Atmospheric Environment*, *115*, 223–235.

Batterman, S., Cook, R., & Justin, T. (2015). Temporal variation of traffic on highways and the development of accurate temporal allocation factors for air pollution analyses. *Atmospheric Environment*, *107*, 351–363.

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2010a). A bivariate space-time downscaler under space and time misalignment. *The Annals of Applied Statistics*, *4*(4), 1942.

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2010b). A spatio-temporal downscaler for output from numerical models. *Journal of Agricultural, Biological, and Environmental Statistics*, *15*(2), 176–197.

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2012). Space–time data fusion under error in computer model output: An application to modeling air quality. *Biometrics*, *68*(3), 837–848.

Byun, D., & Schere, K. L. (2006). Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale air quality (CMAQ) modeling system. *Applied Mechanics Reviews*, *59*(2), 51–77.

Calder, C. A. (2007). Dynamic factor process convolution models for multivariate space–time data with application to air quality assessment. *Environmental and Ecological Statistics*, *14*(3), 229–247.

Calder, C. A. (2008). A dynamic process convolution approach to modeling ambient particulate matter concentrations. *Environmetrics*, *19*(1), 39–48.

Choi, J., Fuentes, M., & Reich, B. J. (2009). Spatial–temporal association between fine particulate matter and daily mortality. *Computational Statistics & Data Analysis*, *53*(8), 2989–3000.

Crooks, J. L., & Özkaynak, H. (2014). Simultaneous statistical bias correction of multiple $PM_{2.5}$ species from a regional photochemical grid model. *Atmospheric Environment*, *95*, 126–141.

Fuentes, M. (2001). A high frequency kriging approach for non-stationary environmental processes. *Environmetrics*, *12*(5), 469–483.

Fuentes, M., & Raftery, A. E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics*, *61*(1), 36–45.

Garfield, L. (2018). 13 cities that are starting to ban cars. Retrieved from https://www.businessinsider.com/cities-going-car-free-ban-2017-8.

Gilani, O., Berrocal, V. J., & Batterman, S. A. (2016a). Non-stationary spatio–temporal modeling of traffic-related pollutants in near-road environments. *Spatial and Spatio-Temporal Epidemiology*, *18*, 24–37.

Gilani, O., McKay, L. A., Gregoire, T. G., Guan, Y., Leaderer, B. P., & Holford, T. R. (2016b). Spatiotemporal calibration and resolution refinement of output from deterministic models. *Statistics in Medicine*, *35*(14), 2422–2440.

Grimm, H., & Eatough, D. J. (2009). Aerosol measurement: The use of optical light scattering for the determination of particulate size distribution, and particulate mass, including the semi-volatile fraction. *Journal of the Air & Waste Management Association*, *59*(1), 101–107.

Gunier, R. B., Hertz, A., Von Behren, J., & Reynolds, P. (2003). Traffic density in California: Socioeconomic and ethnic differences among potentially exposed children. *Journal of Exposure Science & Environmental Epidemiology*, *13*(3), 240.

Higdon, D. (1998). A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, *5*(2), 173–190.

Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., & van den Brandt, P. A. (2002). Association between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. *The Lancet*, *360*(9341), 1203–1209.

Holland, D. M., Saltzman, N., Cox, L. H., & Nychka, D. (1999). Spatial prediction of sulfur dioxide in the eastern United States. In *geoENV II – Geostatistics for Environmental Applications* (pp. 65–76). Dordrecht, the Netherlands: Springer.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., … Britter, R. (2015). The rise of low-cost sensing for managing air pollution in cities. *Environment International*, *75*, 199–205.

Lewis, S. A., Antoniak, M., Venn, A. J., Davies, L., Goodwin, A., Salfield, N., … Fogarty, A. W. (2005). Secondhand smoke, dietary fruit intake, road traffic exposures, and the prevalence of asthma: A cross-sectional study in young children. *American Journal of Epidemiology*, *161*(5), 406–411.

Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(4), 423–498.

McMillan, N. J., Holland, D. M., Morara, M., & Feng, J. (2010). Combining numerical model output and particulate data using Bayesian space–time modeling. *Environmetrics*, *21*(1), 48–65.

Milando, C. W., & Batterman, S. A. (2018). Operational evaluation of the RLINE dispersion model for studies of traffic-related air pollutants. *Atmospheric Environment*, *182*, 213–224.

Neto, J. H. V., Schmidt, A. M., & Guttorp, P. (2014). Accounting for spatially varying directional effects in spatial covariance structures. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *63*(1), 103–122.

Office of Air Quality Planning and Standards. (1998). SLAMS/NAMS/PAMS: Network Review Guidance. In United States Environmental Protection Agency, Research Triangle Park, North Carolina 27711, EPA-454/R-98-003 edition.

Paciorek, C. J., & Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, *17*(5), 483–506.

Patton, A. P., Milando, C., Durant, J. L., & Kumar, P. (2017). Assessing the suitability of multiple dispersion and land use regression models for urban traffic-related ultrafine particles. *Environmental Science & Technology*, *51*(1), 384–392. PMID:27966909.

Reich, B. J., Chang, H. H., & Foley, K. M. (2014). A spectral method for spatial downscaling. *Biometrics*, *70*(4), 932–942.

Reich, B. J., Eidsvik, J., Guindani, M., Nail, A. J., & Schmidt, A. M. (2011). A class of covariate-dependent spatiotemporal covariance functions. *The Annals of Applied Statistics*, *5*(4), 2265–2687.

Risser, M. D., & Calder, C. A. (2015). Regression-based covariance functions for nonstationary spatial modeling. *Environmetrics*, *26*(4), 284–297.

Rundel, C. W., Schliep, E. M., Gelfand, A. E., & Holland, D. M. (2015). A data fusion approach for spatial analysis of speciated $PM_{2.5}$ across time. *Environmetrics*, *26*(8), 515–525.

Saha, P. K., Khlystov, A., Snyder, M. G., & Grieshop, A. P. (2018). Characterization of air pollutant concentrations, fleet emission factors, and dispersion near a North Carolina interstate freeway across two seasons. *Atmospheric Environment*, *177*, 143–153.

Sampson, P. D., & Guttorp, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, *87*(417), 108–119.

Schmidt, A. M., Guttorp, P., & O'Hagan, A. (2011). Considering covariates in the covariance structure of spatial processes. *Environmetrics*, *22*(4), 487–500.

Snyder, M. G., Venkatram, A., Heist, D. K., Perry, S. G., Petersen, W. B., & Isakov, V. (2013). RLINE: A line source dispersion model for near-surface releases. *Atmospheric Environment*, *77*, 748–756.

US Environmental Protection Agency (2017). Outdoor air quality data. Retrieved from https://www.epa.gov/outdoor-air-quality-data/interactive-map-air-quality-monitors

Vardoulakis, S., Gonzalez-Flesca, N., Fisher, B. E., & Pericleous, K. (2005). Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. *Atmospheric Environment*, *39*(15), 2725–2736.

Wallace, J., Corr, D., Deluca, P., Kanaroglou, P., & McCarry, B. (2009). Mobile monitoring of air pollution in cities: the case of Hamilton, Ontario, Canada. *Journal of Environmental Monitoring*, *11*(5), 998–1003.

Watkins, N., & Baldauf, R. (2012). Near-road NO2 monitoring technical assistance document (No. EPA-454/B-12-002). EPA, Research Triangle Park, NC.

Yu, H., Russell, A., Mulholland, J., Odman, T., Hu, Y., Chang, H. H., & Kumar, N. (2018). Cross-comparison and evaluation of air pollution field estimation methods. *Atmospheric Environment*, *179*, 49–60.

Zhai, X., Russell, A. G., Sampath, P., Mulholland, J. A., Kim, B.-U., Kim, Y., & D'Onofrio, D. (2016). Calibrating R-LINE model results with observational data to develop annual mobile source air pollutant fields at fine spatial resolution: Application in Atlanta. *Atmospheric Environment*, *147*, 446–457.

Zidek, J. V., Le, N. D., & Liu, Z. (2012). Combining data and simulated data for space–time fields: Application to ozone. *Environmental and ecological statistics*, *19*(1), 37–56.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.