

# Air pollution estimation under air stagnation—A case study of Beijing

Ying Zhang<sup>1</sup> | Song Xi Chen<sup>2,3</sup> | Le Bao<sup>1</sup> 

<sup>1</sup>Department of Statistics, The Pennsylvania State University, State College, Pennsylvania, USA

<sup>2</sup>Guanghua School of Management, Peking University, Beijing, China

<sup>3</sup>Center for Statistical Science, Peking University, Beijing, China

## Correspondence

Song Xi Chen, Guanghua School of Management, Peking University, Beijing, 100871, China.

Email: [songxichen@pku.edu.cn](mailto:songxichen@pku.edu.cn)

Le Bao, Department of Statistics, The Pennsylvania State University, Thomas Building 323, State College, PA, USA.

Email: [lebao@psu.edu](mailto:lebao@psu.edu)

## Funding information

LMEQF at Peking University; National Natural Science Foundation of China, Grant/Award Numbers: 92046021, 12071013, 12026607; NIH/NIAID, Grant/Award Numbers: R01AI136664, R01AI170249

Air pollution continues to be a major environmental concern in China. The wind-driven transmission poses difficulties in understanding the air pollution patterns at the local level. The main objective of this study is to offer a straightforward approach for investigating the temporal trends and meteorological effects on the air pollutant concentrations during the generation process without being confounded by the complex wind-driven transmission effect. We focus on the hourly data of the three most common air pollutants: PM<sub>2.5</sub>, NO<sub>2</sub>, and CO under air stagnation in Beijing, China, during 2014–2017. We find that the local pollution levels under air stagnation in Beijing have decreased over the years; winter is the severest month of the year; Sunday is the clearest day of the week. Our model also interpolates the air pollutant concentrations at sites without monitoring stations and provides a map of air pollution concentrations under air stagnation. The results could be used to identify locations where air pollutants easily accumulate.

## KEYWORDS

air pollution, air stagnation, meteorology, spatio-temporal model

## 1 | INTRODUCTION

In recent years, awareness of the threat posed by air pollution has increased among the international community. Several studies have observed associations between air pollution and respiratory diseases in the past (Ferkol & Schraufnagel, 2014; Kampa & Castanas, 2008). Epidemiological studies suggest that exposure to air pollution adversely affects human health, promoting the risk of asthma, lung cancer, heart diseases, and cardiovascular diseases (Kampa & Castanas, 2008). In 2019, 99% of the world's population was living in places where the WHO air quality guidelines levels were not met. Ambient (outdoor) air pollution is estimated to have caused 4.2 million premature deaths worldwide in 2019 (World Health Organization, 2022).

Air pollution-related deaths account for around 17% of all deaths in China (Rohde & Muller, 2015). In the past few years, air pollution data in China has become more readily available, including satellite imaging and hourly air pollution measurements from a ground monitoring network in China published by China National Environmental Monitoring Center (CNEMC, <http://www.cnemc.cn/>) since 2013. Particulate matter with aerodynamic diameters below 2.5 micrometers (PM<sub>2.5</sub>) serves as an indicator of air pollution level. These ultra-fine particles are fatal as they can translocate from

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Environmetrics* published by John Wiley & Sons Ltd.

the lungs and penetrate deep into the circulatory system (Xing et al., 2016). The long-term exposure to PM<sub>2.5</sub> contributed to an estimated 1.6–2.2 million premature deaths annually during 2013–2015 in China (Li et al., 2018). Besides, nitrogen dioxide (NO<sub>2</sub>) and carbon monoxide (CO) are also of great concern, and they are pervasive and primarily human-caused through fossil fuel consumption.

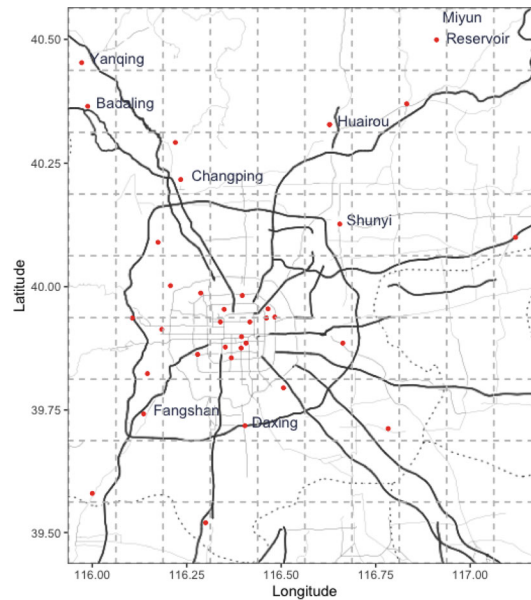
Estimating and predicting air pollution levels is challenging due to complex atmospheric processes, various emission sources, data availability and quality, rapid urbanization, and industrialization (U.S. Environmental Protection Agency, 2021). Despite these challenges, advancements in monitoring technologies and modeling techniques continue to improve our understanding of air pollution dynamics (Lelieveld et al., 2015). The recently developed modeling approaches include non-linear prediction with neural networks (Wang & Cao, 2023), machine learning method for air pollution prediction (Krylova & Okhrin, 2022), Bayesian hierarchical models (Szpiro et al., 2010), geographically weighted regression (Hu et al., 2013), non-parametric regression (Liang et al., 2015), the kriging models (Sampson et al., 2013), and the spatio-temporal models (Fioravanti et al., 2022; Wan et al., 2021). Covariates commonly used in the past air pollution studies include meteorological variables (Wan et al., 2021; Zheng et al., 2021), land-use related (Henderson et al., 2007; Hoek et al., 2008) and road-traffic related variables (Henderson et al., 2007). More recently, satellite data became popular information to model and estimate air pollution concentrations (Ma et al., 2016; Procházka et al., 2000). Procházka et al. (2000) provided general methods to process satellite images for detecting and predicting air pollution. Ma et al. (2016) utilized remote sensing satellite retrieved aerosol optical depth data to predict air pollution concentrations. This paper includes meteorological variables and nighttime light intensity from satellite images as covariates to help estimate air pollution concentrations.

However, wind effects inevitably dominate air pollution levels due to their strong transmission effects, which vary across land cover types. Many analysis results show that a higher wind speed can help disperse and dilute air pollutants and decrease their concentration (Li et al., 2014). However, wind under different meteorological conditions could have different effects. For instance, a strong wind can blow surface pollutants into the air and thus increase air pollution levels (Kukkonen et al., 2005; Vardoulakis & Kassomenos, 2008). For Beijing, China, in particular, a study has shown that the south wind brings pollutants while the north wind guarantees clean air (Liang et al., 2015). This paper aims to study the air pollutants' levels under air stagnation to eliminate the noises caused by the wind effects and better estimate the temporal and meteorological effects. Our model considers the spatial heterogeneity and dependence at the grid level and the temporal pattern at different scales such as year, month, weekday, and hour. We also investigate the relationship between air pollutants and meteorological factors, including temperature, boundary layer height, pressure, dew point, evaporation, rain, and nighttime light intensity from satellite images. On the northern edge of the severest air-polluted region in China, we take Beijing as an illustrative example with the hourly PM<sub>2.5</sub>, NO<sub>2</sub>, and CO data from 36 monitoring stations between February 1, 2014, and May 31, 2017.

The rest of the paper is organized as follows. Section 2 describes air pollution data, the explanatory variables, and the air stagnation condition. Section 3 presents the spatio-temporal models for kriging and propose two evaluation metrics for their kriging performances. We also discuss the methods for inference of associations between air pollutants and other variables including meteorological variables and night-time light intensity. Section 4 provides the kriging maps and results of temporal and meteorological associations for three pollutants. Section 5 concludes with discussions on future works.

## 2 | DATA DESCRIPTION

The hourly air pollution data in Beijing from February 1, 2014, to May 31, 2017, for this study were obtained from 36 monitoring stations, where 35 were from the Beijing Municipal Environmental Monitoring Center (BMEMC) and one from the US Embassy site. Figure 1 shows the locations of the 36 monitoring stations in Beijing used in the study. The meteorological variables and the wind variables are the hourly re-analysis datasets from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>) and TIGGE (<https://apps.ecmwf.int/datasets/data/tigge/>) with a grid resolution of  $0.125^\circ \times 0.125^\circ$ . The nightlight intensity data are obtained from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) produced by the Earth Observation Group, Payne Institute for Public Policy (<https://payneinstitute.mines.edu/eog/nighttime-lights/>), with resolution 15 arc seconds (around 500 m) in the geographic grid.



**FIGURE 1** Map of monitoring sites for air pollution concentrations across Beijing, China, from February 1st, 2014, to May 31st, 2017. The red points represent the monitoring sites, and the grids with the dashed line are the spatial resolution of the meteorological variables from ECMWF ERA5 hourly data.

## 2.1 | Variables description

We focus on three pollutants: PM<sub>2.5</sub>, NO<sub>2</sub>, and CO. Table 1 summarizes the continuous variables used in our analysis. The meteorological variables in Table 1 are mainly based on the past studies (Cai et al., 2017; Holzworth, 1967; Liang et al., 2015; Tan et al., 2018; Wan et al., 2021; Zhang et al., 2017; Zheng et al., 2021). For example, the air temperatures at 2 m, 250 hPa, and 850 hPa are correlated with the air pollution level in (Cai et al., 2017; Holzworth, 1967). The pressure (PRES) is associated with air pollution in (Liang et al., 2015; Wan et al., 2021; Zhang et al., 2017; Zheng et al., 2021). We also added a night-time light intensity as a variable, which reflected the population density across different locations (Tan et al., 2018).

Besides the instant measurements of wind and precipitation, their cumulative amounts are also expected to affect the air pollution level (Liang et al., 2015). Therefore, we create the integrated (cumulative) version of those variables, IHour, IRAIN, IWS10m: IHour (Integrated Stagnant Hour) stands for the percentage of hours under stagnant air status over the past 24 h, where the air stagnation is defined in subsection 2.2; IRAIN stands for the average precipitation over the past 24 h; IWS10m stands for the integrated wind speed at 10 m above the surface, where the wind speed is accumulated if the wind direction remains the same as the wind direction in the previous hour.

The spatial resolutions of monitoring stations, meteorological variables, and night-time light intensity differ, and it is known as the change of support problem (Gelfand et al., 2001). Since most of the variables are meteorological, we decide to use the centroids of their spatial grid as the standards for the location in modeling and inference. We set each grid's hourly air pollution levels as the average air pollution levels at the monitoring stations within that grid. After the aggregation of monitoring stations, there are 23 grid cells with observed air pollutant concentrations. Similarly, averaging night-time light intensity is taken in an 8-km buffer zone within each grid to match the night-time light intensity with the spatial resolution of the meteorological variables.

In addition to variables described in Table 1, we include the categorical variables such as year, month, weekday, and hour of the day and a spatial indicator for being at the center of Beijing or not. Here we define the center of Beijing as the region within Beijing's Fifth Ring Road.

Lastly, we also derive the probability of air stagnation using the ensemble wind forecast from the European Centre for Medium-Range Weather Forecasts (ECMWF) TIGGE Data (<https://apps.ecmwf.int/datasets/data/tigge/>) to study and control for the dependence between pollutant concentrations and air stagnation. The probability of air stagnation is calculated as the proportion of instant wind forecast ensemble members that satisfy the air stagnation definition provided in the next subsection.

TABLE 1 Description for continuous variables.

Name	Type	Range	Mean	Description
PM2.5	Pollutant	[3, 696]	79.65	PM2.5 concentration level ( $\mu\text{g} \cdot \text{m}^{-3}$ )
NO <sub>2</sub>	Pollutant	[2, 300]	49	NO <sub>2</sub> concentration level ( $\mu\text{g} \cdot \text{m}^{-3}$ )
CO	Pollutant	[100, 10,000]	1,126	CO concentration level ( $\mu\text{g} \cdot \text{m}^{-3}$ )
Longitude	Spatial	[116.0, 117.1]	116.4	Longitude
Latitude	Spatial	[39.5, 40.5]	40.02	Latitude
Temp at 2m	Meteorological	[261.7, 314.1]	293.6	Temperature at 2 m (unit:K)
t250	Meteorological	[208.4, 241.7]	225.05	Temperature at 250 hPa (unit:K)
t850	Meteorological	[261.5, 300.1]	286.7	Temperature at 850 hPa (unit:K)
blh	Meteorological	[10.27, 3795.74]	362.39	Boundary layer height (unit:m)
Evapor	Meteorological	[-6.67e-04, 1.28e-05]	-7.48e-05	Evaporation (unit:m)
DEWP	Meteorological	[265.4, 298.3]	280.9	Dew point at 2 m (unit:K)
HUMI	Meteorological	[18.30, 81.17]	54.51	Relative humidity (unit: %)
Geopotential	Meteorological	[13,249, 15,618]	14,541	Geopotential height: (unit:m <sup>2</sup> s <sup>-2</sup> )
PRES	Meteorological	[90,978, 103,532]	98,868	Surface pressure (unit:Pa)
Radiation	Meteorological	[0, 3,839,385]	1,276,707	Surface net solar radiation (unit:J/m <sup>2</sup> )
IRAIN	Meteorological	[0, 4.14e-03]	9.94e-05	Average precipitation over the past 24 h (unit:m)
Wind10m	Meteorological	[0, 3.2]	1.25	Current wind speed at 10 m (unit:m·s <sup>-1</sup> )
Wind500	Meteorological	[0, 13]	7.6	Current wind speed at 500 hPa (unit:m·s <sup>-1</sup> )
IWS10m	Meteorological	[0, 3.2]	1.49	Integrated wind speed at 10 m (unit:m·s <sup>-1</sup> )
IHour	Other	[1.39%, 95.83%]	37%	Integrated air stagnant hour (unit: %)
Nightlight	Other	[39.5, 40.5]	40.02	Night-time light intensity (unit: nanowatts · sr <sup>-1</sup> · cm <sup>-2</sup> )
Prob at stagnant	Meteorological	[0, 100]	33.42	Probability of air stagnation (unit: %)

## 2.2 | Air stagnation data

The air stagnation days were firstly defined by the National Oceanic and Atmospheric Administration (NOAA) in the United States (Wang & Angell, 1999). The criteria of air stagnation days are determined by three factors: lower-air (10 m) wind, upper-air (500 hPa) wind, and precipitation. An air stagnation day is when 10 m wind speed (lower-air wind) is less than 3.2 m per second ( $\text{m} \cdot \text{s}^{-1}$ ), the 500 hPa wind (upper-air wind) is less than  $13 \text{ m} \cdot \text{s}^{-1}$ , and it is precipitation-free.

The National Oceanic and Atmospheric Administration (NOAA) in the United States (Wang & Angell, 1999) defined the air stagnation days with three factors: lower-air (10 m) wind, upper-air (500 hPa) wind, and precipitation. An air stagnation day is when 10 m wind speed (lower-air wind) is less than 3.2 m per second ( $\text{m} \cdot \text{s}^{-1}$ ), the 500 hPa wind (upper-air wind) is less than  $13 \text{ m} \cdot \text{s}^{-1}$ , and it is precipitation-free. To accommodate for hourly data, we similarly defined the air stagnation hours: the integrated 10 m wind speed and daily 10 m (lower-air) wind speed are less than  $3.2 \text{ m} \cdot \text{s}^{-1}$ , the current and daily 500 hPa (upper-air) wind speed is less than  $13 \text{ m} \cdot \text{s}^{-1}$ . Since we are also interested in the effects of precipitation on PM2.5, NO<sub>2</sub>, and CO under air stagnation, precipitation is not used as the data exclusion criteria.

Our air stagnation data range from February 1, 2014, to May 31, 2017. We include 87,261 time-location combinations that satisfy the air stagnation conditions in our study. The proportion of air stagnation data takes up to 14% of the total hourly data. Although it is often preferable to incorporate all existing data in a model, we believe the air stagnation data themselves are more suitable for revealing the spatial and temporal patterns of air pollution concentrations and for investigating the associations between three air pollutants and auxiliary variables, including temporal variables, meteorological variables, and nightlight intensity. The large number of data points satisfying the air stagnation condition is also sufficient for deriving reliable estimates. The detailed patterns for the percentages of hours not under air stagnation across locations and months are shown in Appendix S1.

### 3 | MODELING OF POLLUTION CONCENTRATION

We take the grid-level air stagnation data as point-referenced data for modeling, using their centroids as the location of each grid cell to calculate the distance between two grid cells. We first establish a Gaussian spatio-temporal process model for the air pollution concentrations in subsection 3.1. We then introduce two parametric spatio-temporal models and the variogram approach for estimating the covariance parameters in subsection 3.2. The model parameters are estimated by using only air stagnation data. We resolve the issue of selection bias with the covariate adjustment using propensity score in subsection 3.3 and discuss the association studies in subsection 3.4. In subsection 3.5, the kriging model is used to interpolate pollution levels at the unobserved locations and times. In subsection 3.6, we present our strategy for model selection and the metrics for evaluating the prediction accuracy.

#### 3.1 | Gaussian spatio-temporal process model

The explanatory variables being considered are the temporal variables (year, month, weekday, and hour of the day, all of which are modeled as dummy variables), the meteorological variables (rain, temperatures at 2 m, 250 hPa, and 850 hPa, surface pressure, dew point temperature, boundary layer height, integrated precipitation, integrated air stagnant hour up to time  $t$ , solar radiation, geopotential height, and evaporation), nightlight and spatial location (latitude, longitude, and the indicator for whether being at the center of Beijing). In addition to those main effects, we consider two-way interactions between significant meteorological effects and the temporal dummy variables for possible temporally varying meteorological effects.

Let  $s \in S = \{1, \dots, 90\}$  denote a grid cell for all 90 grid cells of interest,  $t \in \mathcal{T} = \{1, \dots, T\}$  denote a time point during our study period where  $T = 29,184$  is the total number of hours, and  $n = 90 \times T$  is the number of all  $(s, t)$  pairs. The matrix  $\mathbf{X}$  represents the covariates selected from the candidate exploratory variables as mentioned above and their two-way interactions. It is a  $n \times p$  design matrix where  $p$  is the number of covariates. One row of  $\mathbf{X}$ ,  $X(s, t)$ , corresponds to the covariates measured at grid cell location  $s$  and time  $t$ . The  $n$ -dimensional vector,  $\mathbf{Y}$ , represents the cubic-root transformed air pollution levels **under the air stagnation** for all  $(s, t)$  pairs. The cubic-root transformation provides better normality of the residuals. Note that  $\mathbf{Y}$  are not fully observed for all the locations and time points. 86% of elements in  $\mathbf{Y}$  are missing due to the absence of monitoring stations or the unmet air stagnation condition. The rows of  $\mathbf{X}$  and the elements of  $\mathbf{Y}$  are sorted by location first and then by timestamp.

Our modeling framework can be written as two parts: an unobservable process model and an observable data model (Cressie & Wikle, 2015). The underlying process model consists of the stationary fixed effect explained by the covariates,  $x(s, t)$  and the spatio-temporal random field,  $\delta_z(s, t)$ , that capture the spatio-temporal dependence. The process model is

$$Z(s, t) = X(s, t)\boldsymbol{\beta} + \delta_z(s, t), \quad (1)$$

where  $Z(s, t)$  represents the underlying pollution level at location  $s$  and time  $t$ , the  $p$ -dimensional vector  $\boldsymbol{\beta}$  represents the fixed effect coefficients for the covariates  $X(s, t)$ , and  $\delta_z, \delta_z(s, t) \in \delta_z$ . The  $n$ -dimensional vector,  $\delta_z$ , are the Gaussian spatio-temporal random effects with mean  $\mathbf{0}$  and covariance  $\boldsymbol{\Sigma}$  which will be specified in subsection 3.2. The data model is

$$Y(s, t) = Z(s, t) + \delta_y(s, t), \quad (2)$$

where  $Y(s, t)$  is the cubic-root transformed air pollution level measured at the time point  $t \in \mathcal{T}$  and the grid cell  $s \in S$ .  $\delta_y(s, t) \in \delta_y$  is the associated measurement error that is independent of  $\delta_z(s, t)$ . The  $n$ -dimensional vector of measurement errors,  $\delta_y$ , follows multivariate Gaussian distribution with mean  $\mathbf{0}$  and covariance  $\sigma_{\text{nugget}}^2 \mathbf{I}_{n \times n}$ , where the variance  $\sigma_{\text{nugget}}^2$  is also known as the nugget variance.  $\delta_y$  is independent of  $\delta_z$ .

We combined the two processes to have

$$Y(s, t) = X(s, t)\boldsymbol{\beta} + \epsilon(s, t), \quad (3)$$

where  $\epsilon(s, t) \in \epsilon$  and the  $n$ -dimensional vector  $\epsilon \sim N(\mathbf{0}, \boldsymbol{\Sigma} + \sigma_{\text{nugget}}^2 \mathbf{I}_{n \times n})$ . Let  $\mathbf{Y}$  and  $\mathbf{X}$  be the overall observation vector of pollutant and the covariate matrix pooling all available sites and observational times together, which gives rise to the summarizing model equation:  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$ .

We use a two-stage estimation procedure to expedite the computing time of parameter estimation. We first estimate the fixed effects  $\beta$  using the ordinary least square (OLS) estimator as

$$\hat{\beta}_{\text{OLS}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (4)$$

Then, we obtain the ordinary least square residuals  $\hat{\epsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta}_{\text{OLS}}$  and use the residual to estimate the parameters in covariance models with the details in subsection 3.2. Lastly, using the estimated covariance models to obtain the generalized least square (GLS) estimator that we will describe in subsection 3.4.

### 3.2 | The covariance models

We assume the spatio-temporal covariance,  $\Sigma$ , to follow either the Gneiting model (Gneiting, 2002) or the Product-Sum model (De Iaco et al., 2001). These two parametric spatio-temporal covariance models allow for the interactions between space and time. They have been proven to be flexible in describing a wide range of covariance surfaces in various studies.

In spatio-temporal covariance models, the dependence structure is often characterized by the spatio-temporal variogram. The variogram between a pair of  $i^{\text{th}}$  residuals  $\epsilon_i$  and  $j^{\text{th}}$  residuals  $\epsilon_j$  is defined as

$$2\gamma_{ij} = \text{Var}(\epsilon_i - \epsilon_j) = \text{Var}(\epsilon_i) + \text{Var}(\epsilon_j) - 2\text{Cov}(\epsilon_i, \epsilon_j), \quad (5)$$

and  $\gamma_{ij}$  is known as the semivariogram (Cressie & Wikle, 2015). The Product-Sum model assumes a spatio-temporal covariance structure:

$$\text{Cov}_{s,t}(h, k) = k_1 \text{Cov}_s(h) + k_2 \text{Cov}_s(h) \text{Cov}_t(k) + k_3 \text{Cov}_t(k),$$

where  $\text{Cov}_s(h)$  and  $\text{Cov}_t(k)$  are valid temporal and spatial covariance structures, respectively. It can also be written in terms of semivariograms:

$$\gamma_{s,t}(h, k) = \gamma_{s,t}(h, 0) + \gamma_{s,t}(0, k) - m\gamma_{s,t}(h, 0)\gamma_{s,t}(0, k), \quad (6)$$

where  $\gamma_{s,t}(h, 0)$  and  $\gamma_{s,t}(0, k)$  are the spatial and temporal semivariograms, respectively, and  $m$  is a parameter that controls for the degree of space-time interaction. We use spherical  $c_{\text{sill},1} \text{sph}(\frac{h}{r})$  for  $\gamma_{s,t}(h, 0)$ , where parameter  $r$  is the range parameter, and exponential semivariograms  $c_{\text{sill},2} \exp(-\frac{h}{\alpha_{\text{scale}}})$  for  $\gamma_{s,t}(0, k)$ .

The Gneiting model assumes an alternative spatio-temporal correlation structure:

$$C_{s,t}(h, k) = \begin{cases} (ak^{2\alpha} + 1)^{-(\delta+\eta)}, & \text{if } h = 0, \\ p_0(ak^{2\alpha} + 1)^{-(\delta+\eta)} \times \exp\left(-\frac{bh}{(ak^{2\alpha} + 1)^{\frac{\eta}{2}}}\right), & \text{otherwise,} \end{cases}$$

where  $a$  and  $b$  are non-negative scaling parameters of time and space respectively,  $\alpha$  is a smoothness parameter, the exponent  $\delta$  is the parameter for temporal correlation,  $\eta$  is a space-time interaction parameter,  $\sigma_{\text{sill}}^2$  is the variance of residuals (or global sill variance), and  $\sigma_{\text{sill}}^2(1 - p_0)$  is the nugget effect. The corresponding semivariogram structure is

$$\gamma_{s,t}(h, k) = \sigma_{\text{sill}}^2 (1 - C_{s,t}(h, k)). \quad (7)$$

To estimate  $\sigma_{\text{nugget}}^2$  and  $\Sigma$ , we first obtain the empirical variogram by using the ordinary least squares residuals  $\hat{\epsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta}_{\text{OLS}}$ . Under the isotropic assumption, that is, that the semivariogram  $\gamma_{ij}$  depends on neither the location nor time, the spatio-temporal semivariogram can be written as a function of distance  $h_{ij}$  and time lag  $k_{ij}$  between two points,

$$\gamma(h_{ij}, k_{ij}) = \frac{1}{2} \text{Var}(\epsilon_{s_i, t_i} - \epsilon_{s_j, t_j}),$$

where  $h_{ij} = \text{Distance}(s_i, s_j)$  using the longitude and latitude of site  $s_i$  and  $s_j$ , and  $k_{ij} = \text{Time Lag}(t_i, t_j) = |t_i - t_j|$  (in hours). The empirical semivariogram is

$$\hat{\gamma}(h, k) = \frac{1}{2|A_{h,k}|} \sum_{i,j \in A_{h,k}} (\hat{\epsilon}_i - \hat{\epsilon}_j)^2,$$

where  $\hat{\epsilon}$  is the estimated ordinary least square residuals,  $A_{h,k}$  is the set of pairs of data points that have distance ( $h - 0.5, h + 0.5$ ) kilometers and time lag  $k$  hours respectively, and  $|A_{h,k}|$  is the cardinality of  $A_{h,k}$ . In practice, we only compute the empirical semivariogram at different integer values of time lag  $k$  and distance  $h$ . We refer to the upper bound of the semivariogram as the global sill variance  $\sigma_{\text{sill}}^2$ , which is defined as  $\sigma_{\text{sill}}^2 = \text{Var}(\epsilon_i) = \sigma_{\text{nugget}}^2 + \text{Var}(\delta_z(s_i, t_i))$  for any  $i$ .

The spatial semivariogram is the spatio-temporal semivariogram when the time lag  $k = 0$ , which is  $\gamma_{s,t}(h, 0)$ , and the temporal semivariogram is the spatio-temporal semivariogram when the distance  $h = 0$ , which is  $\gamma_{s,t}(0, k)$ . We separately fit the spatial and temporal semivariograms to expedite the optimization process. Given the estimated parameters for spatial and temporal semivariograms, we can then estimate space and time interaction parameters ( $m$  in the Product-Sum model and  $\eta$  in Gneiting model) by the weighted least squares with empirical spatio-temporal semivariograms.

### 3.3 | Covariate adjustment using the propensity score

We aim to investigate the temporal trends and meteorological effects on air pollutant concentrations during air stagnation hours and recover the air pollution maps under air stagnation. However, using only air stagnation data could introduce selection bias if a dependence exists between the pollutant concentration and the air stagnation status. To resolve this issue, we borrow the concepts from causal inference. For the location,  $s$  and time  $t$ ,  $Y(s, t)$  is the air pollution level after cubic-root transformed as we defined before, and let  $T_{s,t}$  be a binary air stagnation indicator (one corresponds to air stagnant),  $\mathbf{X}(s, t)$  be the covariates to forecast wind speed. The outcome of interest is air pollution under air stagnation, that is,  $Y(s, t)(T_{s,t} = 1)$  or  $Y_{s,t}(1)$ . Here  $\mathbf{Y}(1)$  means the “potential outcome”, a term from causal inference, that is, air pollutant levels if it is under air stagnation. Conceptually, if we could have “forced” the wind to be stagnant every hour, we would have observed the “complete data” of  $\mathbf{Y}(1)$ . Since we cannot force the wind to be stagnant, we shall work on a subset of the complete data, that is, the observed data of  $Y_{s,t}(1)$  under the air stagnation condition—when  $T_{s,t} = 1$ .

Fitting the Gaussian spatio-temporal process model to the observed data, we found that the residuals correlate with the air stagnant probability. It indicates that the observed data of  $Y_{s,t}(1)$  is not a random sample of the complete data. So, the potential outcome  $Y_{s,t}(1)$  is not independent of  $T_{s,t}$ , that is, some covariates affect both  $Y_{s,t}(1)$  and  $T_{s,t}$ .

We consider the covariate adjustment using the propensity score,  $P(T_{s,t} = 1)$ , to recover the map air pollution levels during stagnant air hours (Rosenbaum & Rubin, 1983). Since the wind speed has been well studied in weather forecasting, instead of modeling the propensity score on our own, we calculate  $P(T_{s,t} = 1)$  from the probabilistic estimates of wind speed from the European Centre for Medium-Range Weather Forecasts (ECMWF) TIGGE Data (<https://apps.ecmwf.int/datasets/data/tigge/>). So the propensity scores are treated as known in our study. In addition, since air pollution has rarely been considered an important predictor in wind speed modeling, we believe the “strongly ignorable” assumption is valid, that is,  $Y_{s,t}(1)$  and  $T_{s,t}$  are independent conditional on the propensity score  $P(T_{s,t} = 1)$ .

The inference for association study in subsection 3.4 is conditioning on the same value of the propensity score so that  $Y_{s,t}(1)$  and  $T_{s,t}$  are conditionally independent under the “strongly ignorable” assumption. After controlling the probability of being air stagnant, we create the spatial maps of air pollution levels in four seasons for three pollutants in subsection 3.5.

### 3.4 | Inference for associations

Our primary goal is to estimate each pollutant’s yearly, monthly, weekly, and hourly trends. In addition, we want to study each meteorological variable’s effect on pollutant concentrations. With the association inference results, we can enhance our understanding of the pattern of air pollution concentrations under varying weather conditions.

The parametric structure of the covariance  $\Sigma + \sigma_{\text{nugget}}^2 \mathbf{I}_{n \times n}$  for the residuals  $\epsilon$  in the model (3) is chosen from the Gneiting model (Gneiting, 2002) or the Product-Sum model (De Iaco et al., 2001) by their corresponding cross-validation performance, with the details given in subsection 3.6. To estimate the regression coefficients of temporal and other meteorological effects in  $\beta$ , one could use the Generalized Least Squares estimator (GLS), accounting for the spatio-temporal

dependence:

$$\hat{\boldsymbol{\beta}} = \left( \mathbf{X}^T (\boldsymbol{\Sigma} + \sigma_{\text{nugget}}^2 \mathbf{I})^{-1} \mathbf{X} \right)^{-1} \mathbf{X}^T \left( \boldsymbol{\Sigma} + \sigma_{\text{nugget}}^2 \mathbf{I} \right)^{-1} \mathbf{Y}, \quad (8)$$

whose variance is

$$\text{Var}(\hat{\boldsymbol{\beta}}) = \left( \mathbf{X}^T (\boldsymbol{\Sigma} + \sigma_{\text{nugget}}^2 \mathbf{I})^{-1} \mathbf{X} \right)^{-1}. \quad (9)$$

The presence of strong multicollinearity among meteorological variables makes the association results hard to interpret. For instance, the boundary layer height (blh) and the pressure are negatively correlated, the former is negatively correlated with air pollution, and the latter is positively correlated with air pollution. However, including both blh and pressure in the model could change the direction of one coefficient due to the multicollinearity. Although the coefficient is interpreted as the expected change of air pollution conditional on all other covariates in the model, such a coefficient can still be confusing and not very useful in practice. On the other hand, the year, month, weekday, and hour effects are almost perpendicular to each other, and it is more intuitive to ask how the air pollutant level would change with a meteorological factor, conditional on a certain year, month, day, hour, and at a given location. Therefore, we consider the model that includes all spatial and temporal effects and only one additional meteorological variable at a time. It will avoid the multicollinearity among meteorological variables.

### 3.5 | Kriging and linear coregionalization model

Kriging was first proposed by Matheron (1963) and has been used to interpolate unsampled locations in geostatistics based on the Gaussian process. The interpolation of air pollution concentration levels at location  $s'$  and time  $t'$  requires estimating two components, the global mean,  $\mathbf{x}^T(s', t')\boldsymbol{\beta}$ , and the Gaussian spatio-temporal random field,  $\epsilon(s', t')$ . The estimated regression coefficients  $\hat{\boldsymbol{\beta}}$  are presented in subsection 3.4, and the distribution of  $\epsilon(s', t')$  given  $\epsilon$ , the residuals of observed data points, is

$$\epsilon(s', t') | \epsilon \sim N \left( \text{Cov}(\epsilon, \epsilon(s', t'))^T (\sigma_{\text{nugget}}^2 \mathbf{I} + \boldsymbol{\Sigma})^{-1} \epsilon, \sigma_*^2 \right),$$

where

$$\sigma_*^2 = \sigma^2 - \text{Cov}(\epsilon, \epsilon(s', t'))^T (\sigma_{\text{nugget}}^2 \mathbf{I} + \boldsymbol{\Sigma})^{-1} \text{Cov}(\epsilon, \epsilon(s', t')).$$

The estimated air pollution level at location  $s'$  and time  $t'$  given,  $\mathbf{x}^T(s', t')$ ,  $\hat{\boldsymbol{\beta}}$ , and  $\hat{\epsilon}$  is

$$\hat{Y}(s', t') = \mathbf{x}^T(s', t')\hat{\boldsymbol{\beta}} + \text{Cov}(\epsilon, \epsilon(s', t'))^T (\sigma_{\text{nugget}}^2 \mathbf{I} + \boldsymbol{\Sigma})^{-1} \hat{\epsilon}, \quad (10)$$

whose variance is

$$\mathbf{x}^T(s', t') \text{Var}(\hat{\boldsymbol{\beta}}) \mathbf{x}^T(s', t')^T + \sigma_*^2.$$

We refer to the previously defined model as the univariate kriging model. Next, we develop a multivariate kriging model that allows correlations across three types of air pollutants, and we examine whether the multivariate model can improve the kriging accuracy. The univariate Product-Sum spatio-temporal variogram can be easily extended to the multivariate joint model using Linear Coregionalization Model (LCM) (De Iaco et al., 2012). Let  $\mathbf{E}(s, t) = [\epsilon_1(s, t), \epsilon_2(s, t), \epsilon_3(s, t)]^T$  be the residual vector of three pollutants (PM<sub>2.5</sub>, NO<sub>2</sub>, and CO) at site  $s$  and time  $t$ . We model the residuals of three pollutants jointly and assume that

$$\mathbf{E}(s, t) = \mathbf{D}\mathbf{M}(s, t),$$

where  $\mathbf{M}(s, t) = [M_1(s, t), \dots, M_P(s, t)]^T$  is a vector of  $P$  uncorrelated Gaussian process, and  $\mathbf{D}$  is a  $3 \times P$  matrix. We use the following auto and cross-variogram of the residual pairs  $(\epsilon_l, \epsilon_s)$ .

$$g_{lq}(h, k) = 0.5 \text{Cov}(\epsilon_l(s+h, t+k) - \epsilon_l(s, t), \epsilon_q(s+h, t+k) - \epsilon_q(s, t)), \quad \text{for } l, q = 1, \dots, 3.$$

We can then model the matrix auto and cross-variogram as an LCM:

$$\mathbf{G}(h, k) = \sum_{j=1}^P \gamma_j(h, k) \mathbf{B}_j,$$

where  $\mathbf{G}(h, k)$  is a  $3 \times 3$  matrix, whose entries are  $g_{lq}(h, k)$ , and the  $\mathbf{B}_j$ 's are positive definite  $P \times P$  matrices. The basic variograms  $\gamma_j(h, k)$ 's are modeled as a Product-Sum (De Iaco et al., 2001). Here, we set  $P = 3$ , the spatial variograms to be  $\text{sph}(\frac{h}{r})$ , and the temporal variograms to be  $c_{\text{sill}} \exp(-\frac{k}{\alpha_{\text{scale}}})$ . We follow the fitting procedure by De Iaco et al. (2012) to estimate the parameters in  $B_j$ 's, spherical and exponential variograms, and the spatio-temporal interaction terms  $m$  in each Product-Sum model. To identify the covariance structures of  $M_1(s, t), \dots, M_P(s, t)$ , we firstly compute the sample auto and cross variograms:

$$\hat{\gamma}_{lq}(h, t) = \frac{1}{2|A_{h,k}|} \sum_{i,j \in A_{h,k}} ((\hat{\epsilon}_l(s_i, t_i) - \hat{\epsilon}_l(s_j, t_j)) (\hat{\epsilon}_q(s_i, t_i) - \hat{\epsilon}_q(s_j, t_j))),$$

where  $A_{h,k}$  is the set of pairs of data points that have distance  $h$  kilometers and time lag  $k$  hours, respectively, and  $|A_{h,k}|$  is the size of  $A_{h,k}$ . We can fit the basic structures of  $M_1(s, t), \dots, M_P(s, t)$  by simultaneous diagonalization (De Iaco et al., 2012). Lastly, we compute the coregionalization matrices  $\mathbf{B}_l$  given the fitted basic structures of  $M_1(s, t), \dots, M_P(s, t)$ .

With the modeled variograms, we deploy the same kriging method and cross-validation strategies for the interpolations and evaluations. We compare the performance of the joint model with separate models in our result.

### 3.6 | Model selection and evaluation

We assess the prediction accuracy of different proposed models via cross-validation. The prediction accuracy is stratified by grid cells.

1. At grid cells with monitoring stations, we impute the PM2.5, NO<sub>2</sub>, and CO that are missing due to either not being collected or not qualifying the stagnant air condition. The time-series data at a grid cell are randomly divided into  $P$  equal-sized subsamples, and we set  $P = 5$  in this paper. We evaluate the prediction accuracy by repeatedly withholding one subsample at a grid cell as the test data. We refer to this evaluation approach as the Leave-Partial-Time-Out (LPTO) cross-validation. It assesses the interpolation accuracy at a location with historical air pollutant data.
2. To evaluate the prediction accuracy for grid cells without any monitoring stations, we take one observed grid cell as the test data repeatedly for all 23 observed grid cells. This evaluation approach will be referred to as the Leave-One-Location-Out cross-validation (LOLO). Such a target-oriented cross-validation design was recommended by Meyer et al. (2018) for spatio-temporal data to prevent over-fitting—the model very well predicts subsets of the time series of the locations used for training but fails in the extrapolation of unknown locations.

The LOLO cross-validation selects more variables that capture the spatial variation, while the LPTO cross-validation favors variables that capture temporal variation. The MAE of LOLO cross-validation is

$$MAE_{\text{LOLO cross-validation}} = \frac{1}{\sum_{s=1}^S \mathcal{T}_s} \sum_{s=1}^S \sum_{t \in \mathcal{T}_s} |y_{st} - \hat{y}_{st}|,$$

for each removed location  $s$ ,  $\{y_{st}\}$  are the observed response from test data, and  $\hat{y}_{st}$  are the predicted value of test data with training data excluding location  $s$ .  $\mathcal{T}_s$  is the time points under air stagnation at location  $s$  with observed air pollutant

levels  $\{y_{st}\}$ . And the MAE is averaged over the entire spatial location  $s$ . The MAE of LPTO cross-validation is

$$MAE_{LPTO \text{ cross-validation}} = \frac{1}{\sum_{s=1}^S \sum_{p=1}^P \mathcal{T}_{sp}} \sum_{s=1}^S \sum_{p=1}^P \sum_{t \in \mathcal{T}_{sp}} |y_{st} - \hat{y}_{st}|,$$

for each removed location  $s$  and partial  $p^{th}$  time set,  $\{y_{st}\}$  are the observed response from test data, and  $\hat{y}_{st}$  are the predicted value of test data with training data excluding the partial  $p^{th}$  time set at location  $s$ .  $\mathcal{T}_{sp}$  is the partial  $p^{th}$  time set under air stagnation at location  $s$  with observed air pollutant levels  $\{y_{st}\}$ . And the MAE is averaged over the entire spatial location  $s$  and time  $p$ . Other metrics (MAE, RMSE,  $R^2$ ) are calculated in a similar way.

We calculate the mean absolute value (MAE) and the square root of the mean square error (RMSE), with the data points being equally weighted for each approach. LPTO RMSE and LOLO RMSE are both aggregated across all locations and time points. We provide these two types of cross-validation metrics to show our kriging performance for locations with partial observations and new locations. We selected the predictors and covariance structures based on the mean absolute error (MAE) by combining these two cross-validation strategies.

The covariance structures of the final models were selected first based on the average of LPTO MAE and LOLO MAE for each pollutant. After we fixed the covariance structure, we obtained the variables in the final models by combining the predictors selected by both LOLO and LPTO under selected covariance structures. We also considered the interactions between selected main effects, but no interaction term has been selected in the forward selection process.

The candidate predictors are first screened by the Bayesian Information Criterion (BIC) (Schwarz, 1978) to avoid the high computational cost and the forward selection procedure without taking the spatio-temporal effect into account. Within the candidate predictors for each pollutant, we use the forward selection cross-validation to determine the best set of predictors, including both main effects and two-way interactions of candidate predictors. These cross-validation variable selection procedures are conducted for each proposed covariance structure: Product-Sum and Gneiting.

## 4 | RESULTS

In this section, we report the results of empirical analysis using the models and inference methods as reported in Section 3. We first present the model selection results in subsection 4.1 and then analyze the temporal trends of three pollutants under air stagnation in subsection 4.2 and the associations with meteorological and nightlight variables in subsection 4.3. Lastly, in subsection 4.4, we present the spatial distributions of three air pollutants in four seasons after interpolations using the model (3) with the selected covariance and variables by cross-validation methods described in subsection 3.6.

### 4.1 | Model selection and prediction accuracy

We select the predictors and covariance structures based on the mean absolute error (MAE) under overall cross-validation performances. According to the LOLO MAE and LPTO MAE average, the product-sum model is used for PM<sub>2.5</sub>, NO<sub>2</sub>, and CO. The intermediate results of model selection are provided in Appendix S2. The covariance parameters estimates are shown in Appendix S3. Table 2 summarizes the cross-validation performance of our final models for three pollutants in MAE, RMSE, and  $R^2$  at the transformed scale (cubic root). The cross-validation  $R^2$  is defined as

$$R^2_{\text{cross-validation}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_{test})^2},$$

where  $\{y_i\}_{i=1}^n$  are the observed response from test data and  $\{\hat{y}_i\}_{i=1}^n$  are the predicted value of test data.  $\bar{y}_{test}$  is the sample mean of the test data. The cross-validation  $R^2$  does not depend on the scale of response, and thus it is more comparable among different air pollutants than MAE and RMSE.

The Leave-Partial-Time-Out (LPTO) cross-validation  $R^2$ 's are above 0.8 for all three pollutants. With the product-sum covariance, LPTO  $R^2$ 's are 0.938 for PM<sub>2.5</sub>, 0.862 for NO<sub>2</sub>, and 0.811 for CO. It indicates that our proposed model could accurately impute the missing values at the monitoring stations. The Leave-One-Location-Out (LOLO) cross-validation  $R^2$ 's are lower than LPTO  $R^2$ 's due to the lack of historical air pollution data. The LOLO cross-validation  $R^2$ 's are 0.873

**TABLE 2** Evaluation metrics under two types of cross-validation methods for the final models: Leave-one-location-out (LOLO) and leave-partial-time-out (LPTO) for univariate models and multivariate models.

Pollutant	Selected variables	Type of CV	MAE (multivariate)	RMSE (multivariate)	R <sup>2</sup> (multivariate)
PM <sub>2.5</sub>	Month, Latitude, Hour, Geopotential, HUMI	LPTO	0.172 (0.170)	0.273 (0.261)	0.938 (0.943)
		LOLO	0.271 (0.356)	0.389 (0.467)	0.873 (0.814)
NO <sub>2</sub>	Month, Weekday, Hour, Nightlight, Temp at 2m, t850, Geopotential, Latitude, Longitude	LPTO	0.185 (0.192)	0.266 (0.279)	0.862 (0.848)
		LOLO	0.455 (0.493)	0.608 (0.614)	0.526 (0.518)
CO	Month, Nightlight, t850, t250, PRES, Geopotential	LPTO	0.393 (0.386)	0.904 (0.623)	0.811 (0.914)
		LOLO	0.763 (0.845)	1.1 (1.22)	0.726 (0.665)

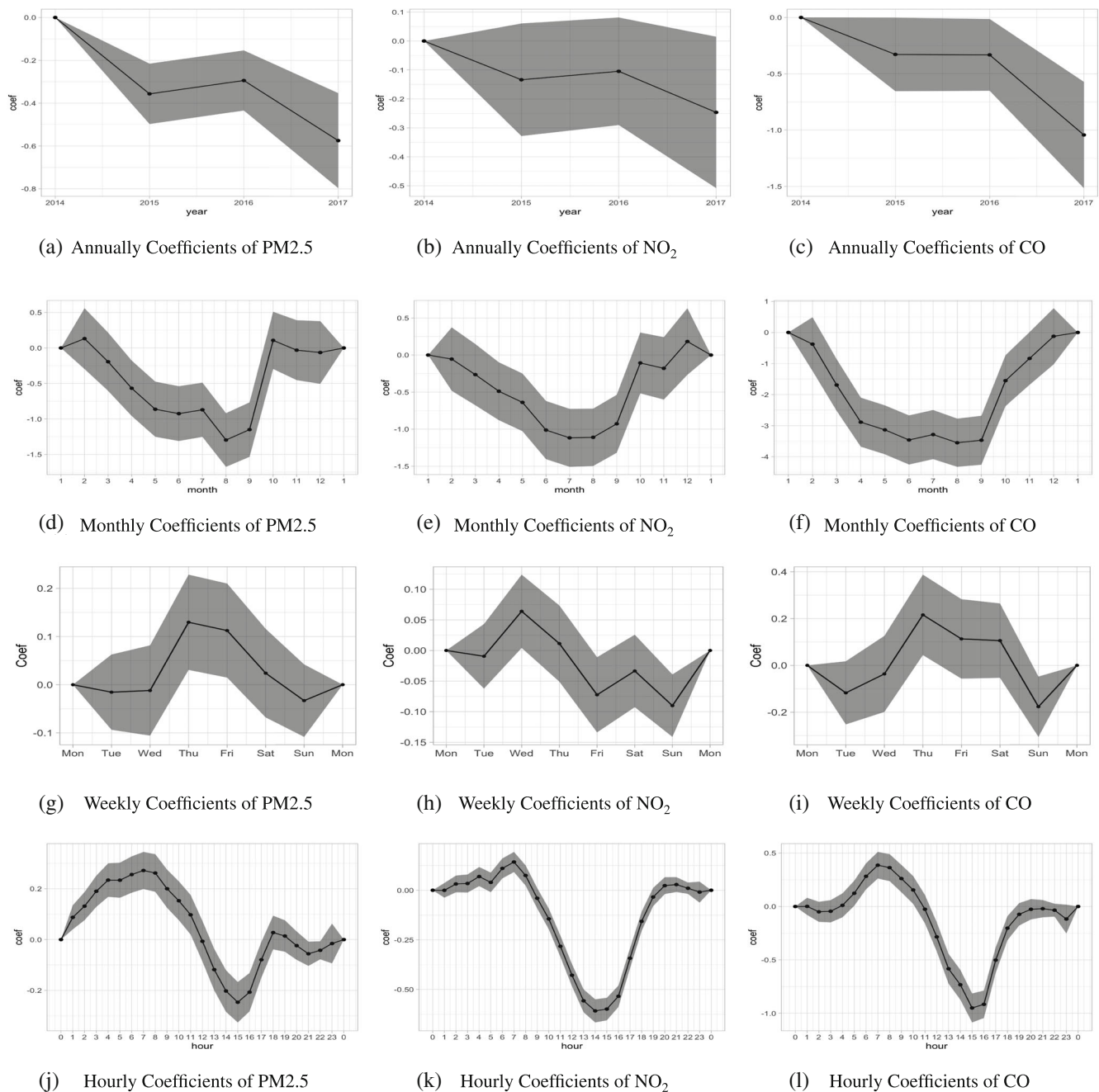
Note: The covariance structure uses the product-sum model.

for PM<sub>2.5</sub>, 0.526 for NO<sub>2</sub>, and 0.726 for CO. NO<sub>2</sub> is relatively harder to predict because it is more reactive than PM<sub>2.5</sub> and CO in the atmosphere and hence has a weaker spatio-temporal dependence. We also present cross-validation results of multivariate kriging models in Table 2. The multivariate performs better for LPTO cross-validations but worse for LOLO cross-validations than the univariate models. In all cases, except the CO model with LPTO cross-validation strategy, the difference of R<sup>2</sup> between univariate and multivariate models is less than 0.06. Regarding the variables selected for kriging models, month and geopotential height are selected for all three pollutants. Using nightlight intensity and temperature at different levels seems to improve the prediction accuracy for NO<sub>2</sub> and CO but not PM<sub>2.5</sub>.

We use the empirical variograms and residual diagnostic plots to check the validity of model assumptions. We first look at the empirical variogram based on five spatial lags up to half the maximum possible distance (60 km) and 24-time lags (from 0- to 23-h). The temporal variogram has more variations for a smaller spatial lag than for a larger spatial lag, suggesting a nonseparable feature for the spatio-temporal covariance. We then define a local space-time ‘cylinder’ Haas (1995) that consists of air-stagnant data from the nearest ten monitoring locations in space to each grid location of interest and within a specific month of the observing period. For each space-time ‘cylinder,’ we fit the Produce-Sum and Gneiting models to the empirical variogram and examine the heterogeneity of estimated parameters across cylinders. We present the estimates’ maximum, minimum, mean, and standard deviation across cylinders in Appendix Table 3. The parameter estimates were fairly close. So, assuming spatio-temporal stationarity of the air pollution level under air stagnation seems reasonable. Residuals of three pollutants models under air stagnation are all tested trend-stationary using Dickey-Fuller test (Dickey & Fuller, 1981). Finally, we plot the residuals versus fitted values to verify the homoscedasticity assumption for error terms. It also shows that the fitted values capture most trends, and our variables have little non-linearity. The normal Q-Q plots of the three models show that the normality assumptions of our models are satisfied. See details of those diagnoses in Section 4 of the Appendix.

## 4.2 | Temporal trends of air pollutants

Figure 2 shows the year effects (2014 as the baseline) and the month effects (January as the baseline) of PM<sub>2.5</sub><sup>1/3</sup>, NO<sub>2</sub><sup>1/3</sup> and CO<sup>1/3</sup>. The mean estimates of all three pollutants decreased over the years in general, except for a mild bump in 2016. The monthly trends suggest the severest air pollution is in winter. One of the reasons is the heating effect during winter, given that the central heating in Beijing usually turns on from November 15 to March 15. Comparing the monthly trends across three pollutants, CO stays at a low concentration level for much longer, from April to September. In contrast, PM<sub>2.5</sub> remains at a low concentration level only from August to September. For all three pollutants, Sunday is the cleanest day in Beijing, as more people stay home and fewer human activities occur during that day. The weekly peaks are different for the three pollutants, but they are all concentrated from Wednesday to Friday as it approaches the end of the weekday. PM<sub>2.5</sub> remains higher on Thursday and Friday, while it can stay lower from Sunday to Wednesday, and NO<sub>2</sub> remains low from Friday to Sunday. The different weekly patterns could be attributed to the different characteristics of the three



**FIGURE 2** Temporal trends (regression coefficients) of PM<sub>2.5</sub>, NO<sub>2</sub>, CO in Beijing under air stagnation from February 1, 2014 to May 31, 2017, with 95% confidence intervals. The gray bands are the 95% confidence interval using GLS inference. The concentration levels are under the cubic root transformation. The first row presents the annual trends of (a) PM<sub>2.5</sub>, (b) NO<sub>2</sub>, and (c) CO; the second row presents the monthly trends of (d) PM<sub>2.5</sub>, (e) NO<sub>2</sub>, and (f) CO; the third row shows the weekly patterns of (g) PM<sub>2.5</sub>, (h) NO<sub>2</sub>, and (i) CO; the last row shows the hourly patterns of (j) PM<sub>2.5</sub>, (k) NO<sub>2</sub>, and (l) CO.

pollutants. PM<sub>2.5</sub> and CO are more stable and could accumulate over days, while NO<sub>2</sub> is not. The hourly trend of the three pollutants all show a dip around 3 p.m.–5 p.m. during the afternoon, which could be attributed to the temperatures, sunlight, and human activities. The boundary layer height reaches its maximum in the afternoon because of the high temperature and the sunshine amount; thus, the air pollution is diluted near the ground. Most people in Beijing end their office hours at 5 p.m. or later, and a typical commute in Beijing can take more than an hour. So the evening rush hour starting at 5 p.m. contributes to the increasing the air pollution level between 5 p.m. and 8 p.m.

### 4.3 | Associations of meteorological and nightlight effects

We control for the temporal effects but investigate the meteorological variables and night-time light one at a time to resolve the issue of multicollinearity as described in subsection 3.4. Here, we summarize the regression coefficients of meteorological variables and night-time light intensity. We also compare the results with the ones controlling for the probability of air stagnation as described in subsection 3.3 in which the air stagnation status does not confound the results.

Table 3 presents the regression coefficients for PM<sub>2.5</sub>, NO<sub>2</sub>, and CO models, respectively. It summarizes results from three approaches: generalized least squares (GLS), generalized least squares with propensity score adjustment (GLS-PS), and ordinary least squares with propensity score adjustment (OLS). The GLS point estimates and standard errors are provided in Equations (8) and (9), where the Product-Sum (PS) model is used for the covariance. The OLS estimates are provided in Equation (4). The point estimates among the three methods always have the same sign, and the relative differences are less than 25% for most coefficients. The standard errors between GLS and GLS-PS are similar and less than the OLS standard errors. Given the similarity of estimates obtained from different methods, we next focus on the GLS estimates and sort the coefficients by their GLS-PS absolute t-values.

Boundary layer height (blh) showed the most significant associations with NO<sub>2</sub> and CO and the second most significant association with PM<sub>2.5</sub>. We quantify that with a 100-meter increase of blh, PM<sub>2.5</sub><sup>1/3</sup> will decrease 0.256, NO<sub>2</sub><sup>1/3</sup> will decrease 0.295, and CO<sup>1/3</sup> will decrease 0.652 for a particular year, month, weekday, and time.

The concentration levels of all three air pollutants increase with the temperature at 2 m, 850 hPa (about 1457 m), and 250 hPa (about 10,363 m). Among the three temperature measures, the temperature at 2 m has the most significant positive association with PM<sub>2.5</sub> and NO<sub>2</sub>, and the temperature in the low-pressure area (250 hPa) has the most significant positive association with CO.

Evaporation is among the top six most significant associations for all three pollutants. The higher the evaporation rate is, the more severe air pollution is. A related meteorological indicator is the humidity level which is also positively associated with all three pollutants, but the associations are less significant than the evaporation rate.

The night-time light intensity reflects the population density at different locations (Tan et al., 2018; Zhuo et al., 2009) and thus the spatial heterogeneity of human activities. The night-time light intensity is the second most significant factor for both NO<sub>2</sub> and CO; it is the tenth most significant factor for PM<sub>2.5</sub>. Air pollution is more severe in areas with more human activities: a one-hundred unit increase of the night-time light intensity attributes to a 2.8 increase of NO<sub>2</sub><sup>1/3</sup>, a 2.8 increase of CO<sup>1/3</sup>, and a 0.4 increase of PM<sub>2.5</sub><sup>1/3</sup>.

For IRAIN (integrated rain), the PM<sub>2.5</sub> and NO<sub>2</sub> concentration levels significantly decrease after heavy rain because PM<sub>2.5</sub> and NO<sub>2</sub> can be dissolved in water while CO is not dissolvable. IHour indicates how long the air has been stagnant, and it is significantly associated with PM<sub>2.5</sub> and CO and marginally associated with NO<sub>2</sub>. Among all covariates investigated, the dew point at 2 m, the geopotential height, and the surface net solar radiation show no significant associations with any air pollutant after controlling for the temporal effects.

### 4.4 | Spatial and seasonal distributions of air pollutants

The resulting maps were rendered using ggmap package in R (Kahle & Wickham, 2013). Geo-location variables such as Latitude, Longitude, and Geopotential frequently appear in the PM<sub>2.5</sub> and NO<sub>2</sub> kriging models but not in the CO model. It is because CO is stable and quickly travels for a long time and distance. As a result, its spatial heterogeneity is not as substantial as PM<sub>2.5</sub> and NO<sub>2</sub>. PM<sub>2.5</sub> and NO<sub>2</sub> concentration levels decrease from south to north and west to east. Several cities with heavy industries in the south of Beijing, like Baoding and Langfang in Hebei province, explained why air pollution was severe in the south. The central area of Beijing lies slightly west within our studying range. Hence, the negative coefficient of longitude in the NO<sub>2</sub> model captures Beijing's central effects, as the central area of Beijing showed higher air pollution levels than the rural area.

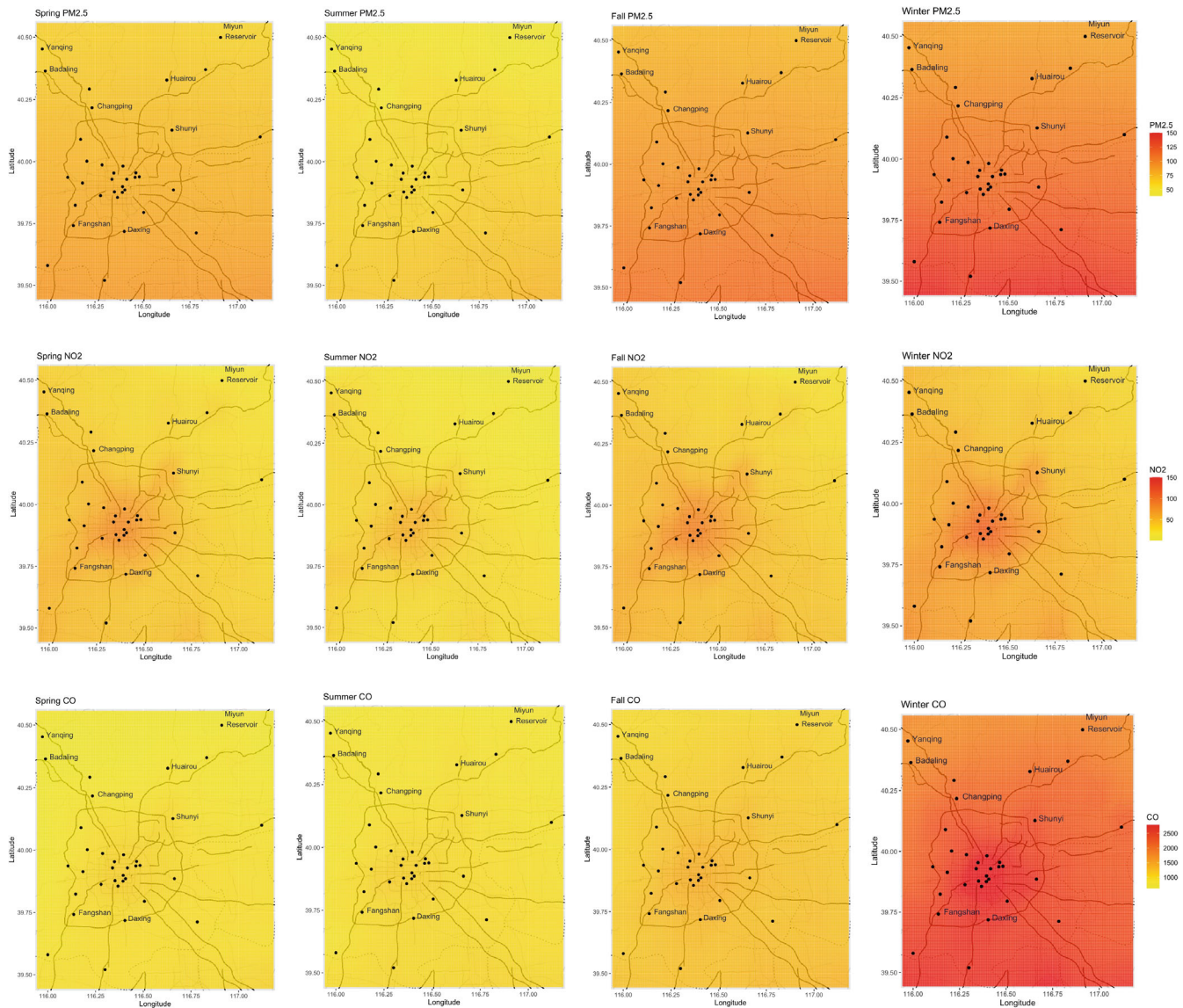
Figure 3 presented the spatial distributions of PM<sub>2.5</sub>, NO<sub>2</sub>, and CO concentration levels under air stagnation at their original scales and averaged by season. Darker colors indicate higher concentration levels. All three pollutants have the lowest concentration levels in the summer and are severest in the winter under stagnant air. Besides, the pollution level of NO<sub>2</sub> is worse in the central area of Beijing City than in rural areas. It is likely attributed to the traffic emission on the roads of intensive transportation in the urban area.

Figure 4 shows the standard errors of cubic root transformed PM<sub>2.5</sub>, NO<sub>2</sub> and CO concentration levels at all grid cells (a darker color indicates a higher standard error), and the Leave-One-Location-Out RMSE at the monitoring stations (a

TABLE 3 Coefficients of meteorological and nightlight variables for PM<sub>2.5</sub>, NO<sub>2</sub>, and CO models.

Variables	Estimate (GLS)	SE (GLS)	Estimate (GLS-PS)	SE (GLS-PS)	Estimate (OLS)	SE (OLS)	t-value (GLS)	p-value (GLS)
PM <sub>2.5</sub> model								
TEMP	0.068	5.99e-04	0.069	5.54e-04	0.069	1.13e-03	113.08	<0.0001
blh	-2.56e-04	2.76e-06	-2.54e-04	2.53e-06	-2.50e-04	1.22e-05	-92.78	<0.0001
t250	0.048	5.26e-04	0.047	4.82e-04	0.047	9.00e-04	91.20	<0.0001
IHour	1.38	0.016	1.37	0.017	1.36	0.019	85.55	<0.0001
t850	0.069	1.47e-03	0.071	1.36e-03	0.071	1.17e-03	47.06	<0.0001
Evapor	791.53	20.55	827.01	19.15	823.22	55.29	38.52	<0.0001
IRAIN	-156.05	10.07	-209.36	10.01	-209.67	14.45	-15.49	<0.0001
HUMI	2.94e-04	2.27e-05	2.59e-04	2.08e-05	2.49e-04	2.27e-04	12.91	<0.0001
PRES	3.71e-05	4.05e-06	4.19e-05	4.05e-06	4.21e-05	1.53e-06	9.17	<0.0001
Nightlight	4.16e-03	5.05e-04	5.48e-03	5.04e-04	5.49e-03	3.14e-04	8.23	<0.0001
Solar	3.24e-10	4.51e-10	1.91e-09	4.13e-10	1.54e-09	3.30e-09	0.72	0.473
Geopotential	-2.67e-05	4.84e-05	-3.00e-06	4.70e-05	-3.09e-05	1.30e-05	-0.55	0.581
DEWP	-3.71e-06	3.37e-05	2.39e-05	3.08e-05	3.49e-05	2.55e-04	-0.11	0.912
NO <sub>2</sub> model								
blh	-2.95e-04	5.26e-06	-2.52e-04	5.26e-06	-2.50e-04	9.00e-06	-55.99	<0.0001
Nightlight	0.028	5.89e-04	0.032	5.89e-04	0.032	2.07e-04	47.43	<0.0001
Evapor	1401.93	36.39	1318.44	36.38	1317.01	40.75	38.52	<0.0001
IRAIN	-460.30	14.39	-374.86	14.38	-375.94	10.64	-32.00	<0.0001
TEMP	0.027	1.09e-03	0.034	1.09e-03	0.034	8.48e-04	24.71	<0.0001
PRES	8.69e-05	4.78e-06	9.94e-05	4.78e-06	9.93e-05	1.09e-06	18.19	<0.0001
t850	0.028	2.44e-03	0.035	2.44e-03	0.035	8.79e-04	11.67	<0.0001
t250	7.06e-03	9.83e-04	8.26e-03	9.83e-04	8.29e-03	6.76e-04	7.18	<0.0001
HUMI	1.32e-04	4.48e-05	1.42e-04	4.48e-05	1.39e-04	1.68e-04	2.94	0.003
IHour	0.048	2.03e-02	0.067	2.03e-02	0.066	1.45e-02	2.38	0.017
Geopotential	1.04e-04	4.54e-05	1.05e-04	4.54e-05	1.05e-04	9.59e-06	2.28	0.022
Solar	1.08e-09	8.74e-10	2.08e-09	8.74e-10	1.97e-09	2.45e-09	1.24	0.421
DEWP	5.64e-05	6.54e-05	-4.13e-05	6.54e-05	-3.81e-05	1.89e-04	0.86	0.389
CO model								
blh	-6.52e-04	1.50e-05	-6.45e-04	1.52e-05	-6.38e-04	2.12e-05	-43.46	<0.0001
Nightlight	0.026	9.75e-04	0.029	9.89e-04	0.030	5.40e-04	26.91	<0.0001
t250	0.060	2.76e-03	0.058	2.80e-03	0.058	1.59e-03	21.66	<0.0001
PRES	1.24e-04	6.93e-06	1.25e-04	6.99e-06	1.37e-04	2.66e-06	17.85	<0.0001
Evapor	1785.63	108.50	2008.42	109.19	1682.80	96.67	16.45	<0.0001
TEMP	0.050	3.13e-03	0.052	3.18e-03	0.052	2.01e-03	16.07	<0.0001
IHour	1.16	9.05e-02	1.05	9.12e-02	1.20	3.40e-02	12.80	<0.0001
t850	0.068	6.87e-03	0.072	6.91e-03	0.081	2.07e-03	9.95	<0.0001
IRAIN	214.07	54.43	40.63	54.58	80.90	25.30	3.93	<0.0001
HUMI	4.48e-04	1.27e-04	2.85e-04	1.29e-04	4.17e-04	3.97e-04	3.53	0.0004
Geopotential	2.34e-04	1.23e-04	3.01e-04	1.23e-04	2.37e-04	2.26e-05	1.91	0.056
DEWP	-3.27e-04	1.84e-04	-3.56e-04	1.86e-04	-3.47e-04	4.46e-04	-1.78	0.075
Solar	2.86e-09	2.45e-09	5.49e-09	2.48e-09	7.18e-09	5.78e-09	1.16	0.244

Note: The table summarizes the point estimates and the standard errors (SE) from three approaches: generalized least squares (GLS), generalized least squares with propensity score adjustment (GLS-PS), and ordinary least squares (OLS). All coefficients are controlled for temporal effects and sorted by their GLS-PS absolute t-values.

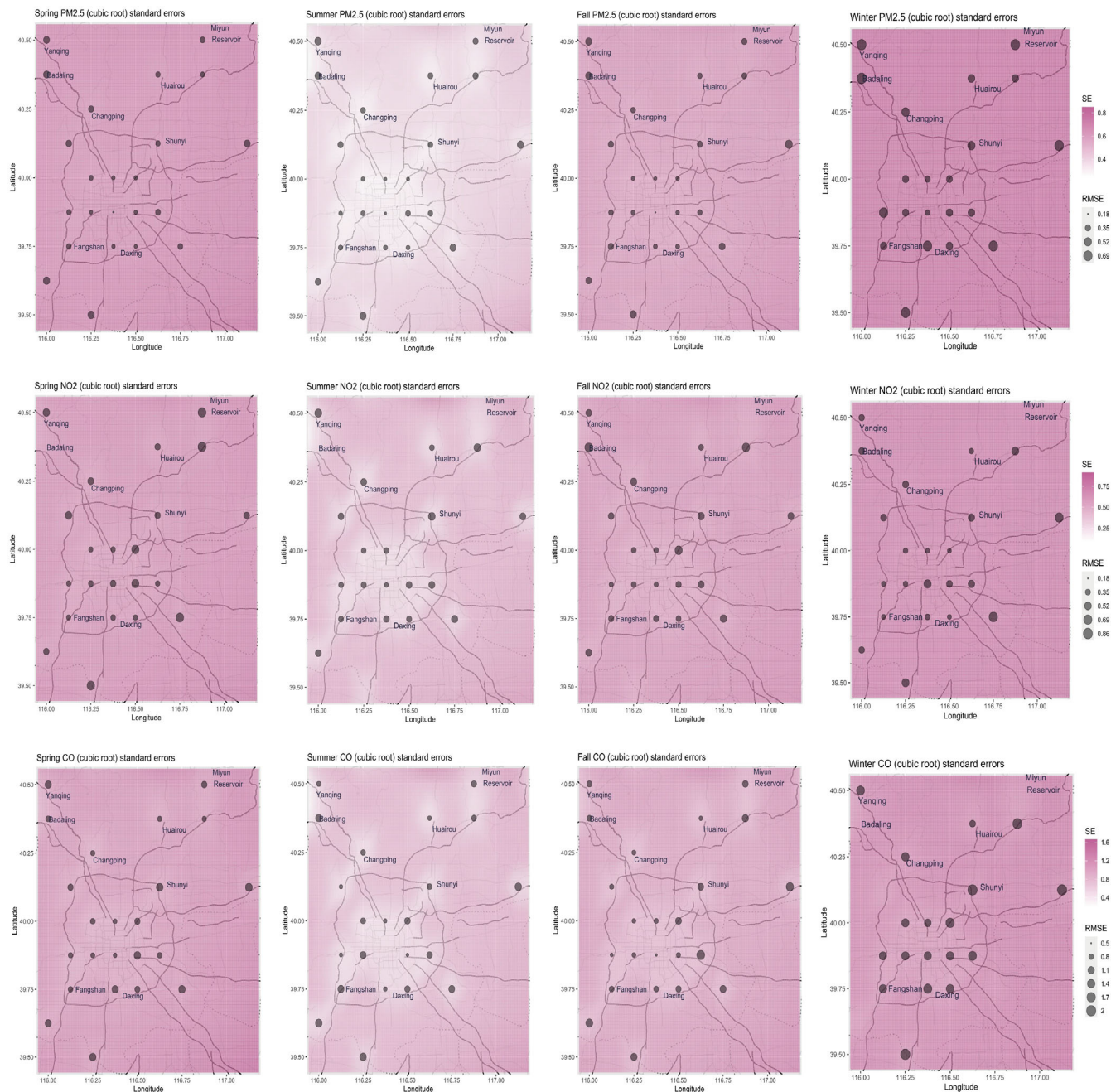


**FIGURE 3** Maps of average pollutant concentrations for PM<sub>2.5</sub> (first row), NO<sub>2</sub> (second row), and CO (third row) in Beijing from spring (left) to winter (right) from February 1, 2014 to May 31, 2017. Concentrations are shown using color gradients from yellow to red, where yellow represents a lighter polluted level, and red represents a more severe polluted level. Monitoring stations are located in grids with black dots. Missing pollution concentrations were interpolated using available hourly air stagnation data from February 2014 to May 2017, and the hourly concentration fields were averaged over each season. The resulting maps were rendered using the ggmap package in R (Kahle & Wickham, 2013).

larger dot indicates a larger RMSE). The central Beijing area has relatively lower standard errors and cross-validation RMSE because of the excellent coverage of monitoring stations. The standard errors are more prominent in spring and winter since there are fewer air-stagnant hours, or in other words, more windy days in spring and winter. The large cross-validation RMSEs often appear on the map's boundary, and they are likely to be improved by enlarging the study area to include more monitoring stations.

## 5 | DISCUSSION

The behavior of air pollutants under air stagnation conditions has been studied as an important index for air pollution severity. Wang and Angell (1999) showed that air pollution is significantly severer during air stagnation days in the United



**FIGURE 4** Standard errors of estimated pollutant hourly concentrations for PM<sub>2.5</sub> (first row), NO<sub>2</sub> (second row), and CO (third row) in Beijing from spring (left) to winter (right) from February 1, 2014, to May 31, 2017. The standard errors at each grid are averaged over each season and shown using color gradients from white (low value) to purple (high value). Monitoring stations are located in grids with black dots. The size of the dots represents the average leave-one-location-out RMSE. The resulting maps were rendered using the ggmap package in R (Kahle & Wickham, 2013).

States. Cai et al. (2017) and Feng et al. (2018) focused on creating the air stagnant index related to the heavily polluted period in Northern China and Beijing, China, respectively. Huang et al. (2017) further interpolated the mean air stagnant days in China. However, all the previous studies have yet to model or quantify the spatial and temporal trend of air pollutants under air stagnation status. We fill this gap by developing a spatio-temporal model for three air pollutants during the stagnant air period.

It is an initial attempt to separate the meteorological and temporal effects from the wind-driven transmission effects for the air pollution problem. When many stagnant air hours are available (over 80,000 data points in our case), the

air stagnation analysis leads to reliable and interpretable estimates of the associations between air pollution levels and auxiliary variables of interest. We found a dependence between air pollution level and the probability of air stagnant, which suggests missing-not-at-random (MNAR). We consider the covariate adjustment using the propensity score, the probability of being air stagnant, to mitigate the potential bias due to MNAR. The investigation of the empirical variograms suggests non-separable spatio-temporal effects. So, we considered two parametric spatio-temporal covariance models. We also found dependence among air pollutants and developed multivariate non-separable stationary spatio-temporal models. To our best knowledge, the multivariate non-separable stationary spatio-temporal model has yet to be developed in the literature.

Our results indicate that Beijing's pollution levels have decreased over the years, which is consistent with China's emission control efforts in recent years. The monthly trends suggest that winter's severe air pollution is likely due to the need for heating. Sunday is the cleanest day as more people stay home and fewer human activities occur. The hourly trend of the three pollutants all show a dip around 3 p.m.–5 p.m., which could be attributed to the temperatures, sun lights, and human activities during that time. We also compare our spatial mapping with the maps, including windy days in Qi et al. (2017). Our maps show the same spatial pattern as Qi et al. (2017) that the south part of Beijing has worse air quality than the north. In addition, the spatio-temporal models uncover the relationships between the meteorological effects and air pollution levels under air stagnation. These association estimates are considered stable and reliable after eliminating the noises caused by winds.

Our results have the following limitations. The variable selection procedure is based on cross-validation error, which cannot imply any causal effects between the weather condition and the polluted levels. A future research direction is to combine our air stagnation models with other wind-driven transmission models, which will provide estimates for both calm and windy days. The wind-driven transmission can be better estimated if we know the source of emissions in the surrounding areas or have monitoring data in a larger geographical area. We also acknowledge the change of support problem (Gelfand et al., 2001) in the data. We use a simple average to solve this problem. More advanced approaches (Stein, 2012) may be utilized to solve this problem.

## ACKNOWLEDGMENTS

Ying Zhang and Le Bao were supported by NIH/NIAID R01AI136664 and R01AI170249. Song Xi Chen was supported by the National Natural Science Foundation of China grants 92046021, 12071013, and 12026607, and LMEQF at Peking University. The authors are grateful to Xun Cao for discussing the sources of covariates, Lei Chen for advising the collection of meteorological data, and Kaiyi Wu and Jinjun Gao for processing the nightlight data. The authors thank the associate editor and reviewers for their constructive suggestions and comments.

## ORCID

Le Bao  <https://orcid.org/0000-0003-0191-5918>

## REFERENCES

- Cai, W., Li, K., Liao, H., Wang, H., & Wu, L. (2017). Weather conditions conducive to Beijing severe haze more frequent under climate change. *Nature Climate Change*, 7(4), 257–262.
- Cressie, N., & Wikle, C. K. (2015). *Statistics for spatio-temporal data*. John Wiley & Sons.
- De Iaco, S., Maggio, S., Palma, M., & Posa, D. (2012). Towards an automatic procedure for modeling multivariate space–time data. *Computers & Geosciences*, 41, 1–11.
- De Iaco, S., Myers, D. E., & Posa, D. (2001). Space–time analysis using a general product–sum model. *Statistics & Probability Letters*, 52(1), 21–28.
- Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 49(4), 1057–1072.
- Feng, J., Quan, J., Liao, H., Li, Y., & Zhao, X. (2018). An air stagnation index to qualify extreme haze events in northern China. *Journal of the Atmospheric Sciences*, 75(10), 3489–3505.
- Ferkol, T., & Schraufnagel, D. (2014). The global burden of respiratory disease. *Annals of the American Thoracic Society*, 11(3), 404–406.
- Fioravanti, G., Cameletti, M., Martino, S., Cattani, G., & Pisoni, E. (2022). A spatiotemporal analysis of no 2 concentrations during the Italian 2020 covid-19 lockdown. *Environmetrics*, 33(4), e2723.
- Gelfand, A. E., Zhu, L., & Carlin, B. P. (2001). On the change of support problem for spatio-temporal data. *Biostatistics*, 2(1), 31–45.
- Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97(458), 590–600.
- Haas, T. C. (1995). Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *Journal of the American Statistical Association*, 90(432), 1189–1199.

- Henderson, S. B., Beckerman, B., Jerrett, M., & Brauer, M. (2007). Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology*, *41*(7), 2422–2428.
- 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. *Atmospheric Environment*, *42*(33), 7561–7578.
- Holzworth, G. C. (1967). Mixing depths, wind speeds and air pollution potential for selected locations in the United States. *Journal of Applied Meteorology*, *6*(6), 1039–1044.
- Hu, X., Waller, L. A., Al-Hamdan, M. Z., Crosson, W. L., Estes, M. G., Jr., Estes, S. M., Quattrochi, D. A., Sarnat, J. A., & Liu, Y. (2013). Estimating ground-level PM<sub>2.5</sub> concentrations in the southeastern US using geographically weighted regression. *Environmental Research*, *121*, 1–10.
- Huang, Q., Cai, X., Song, Y., & Zhu, T. (2017). Air stagnation in China (1985–2014): Climatological mean features and trends. *Atmospheric Chemistry and Physics*, *17*(12), 7793–7805.
- Kahle, D., & Wickham, H. (2013). Ggmap: Spatial visualization with ggplot2. *The R Journal*, *5*(1), 144–161.
- Kampa, M., & Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, *151*(2), 362–367.
- Krylova, M., & Okhrin, Y. (2022). Managing air quality: Predicting exceedances of legal limits for PM<sub>10</sub> and O<sub>3</sub> concentration using machine learning methods. *Environmetrics*, *33*(2), e2707.
- Kukkonen, J., Pohjola, M., Sokhi, R. S., Luhana, L., Kitwiroon, N., Fragkou, L., Rantamäki, M., Berge, E., Ødegaard, V., & Slørdal, L. H. (2005). Analysis and evaluation of selected local-scale PM<sub>10</sub> air pollution episodes in four European cities: Helsinki, London, Milan and Oslo. *Atmospheric Environment*, *39*(15), 2759–2773.
- Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, *525*(7569), 367–371.
- Li, J., Liu, H., Lv, Z., Zhao, R., Deng, F., Wang, C., Qin, A., & Yang, X. (2018). Estimation of PM<sub>2.5</sub> mortality burden in China with new exposure estimation and local concentration-response function. *Environmental Pollution*, *243*, 1710–1718.
- Li, L., Qian, J., Ou, C.-Q., Zhou, Y.-X., Guo, C., & Guo, Y. (2014). Spatial and temporal analysis of air pollution index and its timescale-dependent relationship with meteorological factors in Guangzhou, China, 2001–2011. *Environmental Pollution*, *190*, 75–81.
- Liang, X., Zou, T., Guo, B., Li, S., Zhang, H., Zhang, S., Huang, H., & Chen, S. X. (2015). Assessing Beijing's PM<sub>2.5</sub> pollution: Severity, weather impact, APEC and winter heating. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *471*(2182), 20150257.
- Ma, Z., Hu, X., Sayer, A. M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., & Liu, Y. (2016). Satellite-based spatiotemporal trends in PM<sub>2.5</sub> concentrations: China, 2004–2013. *Environmental Health Perspectives*, *124*(2), 184–192.
- Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, *58*(8), 1246–1266.
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, *101*, 1–9.
- Procházka, A., Kolinova, M., Fiala, J., Hampl, P., & Hlavaty, K. (2000). *Satellite image processing and air pollution detection*. In *2000 IEEE international conference on acoustics, speech, and signal processing. Proceedings (cat. No. 00CH37100)* (Vol. 4, pp. 2282–2285). IEEE.
- Qi, J., Zheng, B., Li, M., Yu, F., Chen, C., Liu, F., Zhou, X., Yuan, J., Zhang, Q., & He, K. (2017). A high-resolution air pollutants emission inventory in 2013 for the Beijing-Tianjin-Hebei region, China. *Atmospheric Environment*, *170*, 156–168.
- Rohde, R. A., & Muller, R. A. (2015). Air pollution in China: Mapping of concentrations and sources. *PLoS One*, *10*(8), e0135749.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.
- Sampson, P. D., Richards, M., Szpiro, A. A., Bergen, S., Sheppard, L., Larson, T. V., & Kaufman, J. D. (2013). A regionalized national universal kriging model using partial least squares regression for estimating annual PM<sub>2.5</sub> concentrations in epidemiology. *Atmospheric Environment*, *75*, 383–392.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.
- Stein, M. L. (2012). *Interpolation of spatial data: Some theory for kriging*. Springer Science & Business Media.
- Szpiro, A. A., Sampson, P. D., Sheppard, L., Lumley, T., Adar, S. D., & Kaufman, J. D. (2010). Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics*, *21*(6), 606–631.
- Tan, M., Li, X., Li, S., Xin, L., Wang, X., Li, Q., Li, W., Li, Y., & Xiang, W. (2018). Modeling population density based on nighttime light images and land use data in China. *Applied Geography*, *90*, 239–247.
- U.S. Environmental Protection Agency. (2021). Air quality modeling.
- Vardoulakis, S., & Kassomenos, P. (2008). Sources and factors affecting PM<sub>10</sub> levels in two European cities: Implications for local air quality management. *Atmospheric Environment*, *42*(17), 3949–3963.
- Wan, Y., Xu, M., Huang, H., & Chen, S. X. (2021). A spatio-temporal model for the analysis and prediction of fine particulate matter concentration in Beijing. *Environmetrics*, *32*(1), e2648.
- Wang, H., & Cao, J. (2023). Nonlinear prediction of functional time series. *Environmetrics*, e2792. <https://doi.org/10.1002/env.2792>
- Wang, J. X., & Angell, J. K. (1999). Air stagnation climatology for the United States. *NOAA/Air Resource Laboratory ATLAS*, *1*, 8.
- World Health Organization. (2022). Ambient (outdoor) air quality and health.
- Xing, Y.-F., Xu, Y.-H., Shi, M.-H., & Lian, Y.-X. (2016). The impact of PM<sub>2.5</sub> on the human respiratory system. *Journal of Thoracic Disease*, *8*(1), E69.
- Zhang, S., Guo, B., Dong, A., He, J., Xu, Z., & Chen, S. X. (2017). Cautionary tales on air-quality improvement in Beijing. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2205), 20170457.

- Zheng, X., Guo, B., He, J., & Chen, S. X. (2021). Effects of corona virus disease-19 control measures on air quality in North China. *Environmetrics*, 32(2), e2673.
- Zhuo, L., Ichinose, T., Zheng, J., Chen, J., Shi, P., & Li, X. (2009). Modelling the population density of China at the pixel level based on DMSP/OLS non-radiance-calibrated night-time light images. *International Journal of Remote Sensing*, 30(4), 1003–1018.

## SUPPORTING INFORMATION

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

**How to cite this article:** Zhang, Y., Chen, S. X., & Bao, L. (2023). Air pollution estimation under air stagnation—A case study of Beijing. *Environmetrics*, e2819. <https://doi.org/10.1002/env.2819>