International Journal of
Environmental Research and Public Health

MDPI

Article

The Functional Spatio-Temporal Statistical Model with Application to O_3 Pollution in Beijing, China

Yaqiong Wang [1,†], Ke Xu [2,*,†] and Shaomin Li [1]

1 Guanghua School of Management, Peking University, Beijing 100871, China; yaqiongwang@pku.edu.cn (Y.W.); lsmjim@pku.edu.cn (S.L.)
2 School of Statistics, University of International Business and Economics, Beijing 100029, China
* Correspondence: xk@uibe.edu.cn
† These authors contributed equally to this work.

Received: 28 March 2020; Accepted: 28 April 2020; Published: 2 May 2020

Abstract: In recent years, with rapid industrialization and massive energy consumption, ground-level ozone (O_3) has become one of the most severe air pollutants. In this paper, we propose a functional spatio-temporal statistical model to analyze air quality data. Firstly, since the pollutant data from the monitoring network usually have a strong spatial and temporal correlation, the spatio-temporal statistical model is a reasonable method to reveal spatial correlation structure and temporal dynamic mechanism in data. Secondly, effects from the covariates are introduced to explore the formation mechanism of ozone pollution. Thirdly, considering the obvious diurnal pattern of ozone data, we explore the diurnal cycle of O_3 pollution using the functional data analysis approach. The spatio-temporal model shows great applicational potential by comparison with other models. With application to O_3 pollution data of 36 stations in Beijing, China, we give explanations of the covariate effects on ozone pollution, such as other pollutants and meteorological variables, and meanwhile we discuss the diurnal cycle of ozone pollution.

Keywords: spatio-temporal statistical model; functional data analysis; O_3 pollution

1. Introduction

As one of the major pollutants, ground-level ozone (O_3) has received a lot of public attention. Lots of studies have shown that O_3 could have detrimental effects on human health, including exacerbation of cardiovascular and respiratory dysfunction, and even premature mortality [1,2]. Additionally, tropospheric ozone, as a greenhouse gas, plays an important role in climate change, and further affects, for example, agricultural crop production [3,4]. In recent years, as the consequence of rapid industrialization and alarmingly increasing energy consumption, China has encountered severe air pollution [5–8]. Particularly, ozone becomes one of the serious and worsening pollutants in major areas of China, such as Beijing–Tianjin–Hebei urban agglomeration, and the Pearl River delta [9,10]. With a population of over 20 million, Beijing is one of the world's largest mega cities. Due to coal burning, fugitive dust, and more recently a rapid increase in vehicular emissions, Beijing faces serious air pollution problems, and especially, studies regarding photochemical ozone pollution are attracting more and more attention [11,12].

The Chinese government identifies the urgency for air quality assessment and emission control, and has built a large monitoring network since 2013. Now, there are over 1500 national pollution monitoring stations in over 300 cities. Hourly readings of air pollutants are regularly recorded and directly transferred to China National Environmental Monitoring Center (CNEMC). The real-time observation and recording of the air pollution data provide a solid basis for studying the dynamic changes of pollutants and the underlying causes. Air quality data are collected over space and

time; thus, the amount of data are large, and the analysis is complex. One important and common statistical characteristic of such data worthy of our notice is that the nearby (both in space and time) observations tend to be more alike than those far apart. Consequently, an assumption that spatio-temporal data follows the "independent and identically distributed" (iid) statistical paradigm should typically be avoided. Based on the underlying spatio-temporal structure of the pollution data, spatio-temporal statistical model, which simultaneously considers both the spatial covariance and temporal dependence, is thus a sensible and reasonable choice [13]. Moreover, O_3 data show a clear diurnal cycle. It peaks during the day and reaches a minimum at night. Since ozone data are sampled at a high frequency in time, it provides an overview of the daily cycle of pollutant concentrations.

A spatio-temporal statistical model is powerful to reveal spatial correlation structure and temporal dynamic mechanism in data. Huang and Cressie (1996) [14] introduced a dynamic random field with a separable spatio-temporal covariance structure, which is widely used in the environmental field. When the spatio-temporal dependencies become complicated, the power of the hierarchical statistical modeling (HM), which is capable of decomposing an uncertainty source of data, becomes apparent. The HM's strength is well discussed in Cressie et al. [15]. Moreover, the daily pattern of ozone pollution needs more exploration. To do this, we divided the collection time into two parts, one related to intra-day fluctuations and the other related to intra-day changes. Geographic space is defined by latitude and longitude, with the date being the third dimension, and the intra-day hour is regarded as the fourth dimension, which gives a four-dimensional representation of the data. In this way, the functional data analysis (FDA) approach [16] is used to model the intra-day variation of the measurement data, and the remaining dimensions are processed according to the classic spatio-temporal data modeling. To summarize, in addition to the dynamic random field and the hierarchical modeling, the third building block is based on the functional representation of daily profiles of atmospheric pollution through a functional data analysis approach, which is the main innovation of the method.

In the present study, we propose a functional spatio-temporal statistical model, which is also a two-level hierarchical spatio-temporal model. A fruitful approach is based on the representation of random functional objects as linear combinations of the basis functions with Gaussian random coefficients. This allows for representing a functional model as a random components model and inheriting the related inferential machinery, e.g., Wood [17]. Based on the Kalman filter and expectation–maximization (EM) algorithm, a model inference for parameter estimates is implemented [18,19]. In addition, from the marginal likelihood function, an information matrix is obtained to measure the uncertainty of the model parameters [20]. The proposed model has the following advantages: (i) the dynamic random field is used to describe the spatio-temporal characterization of emissions of air pollution; (ii) and covariate effects are incorporated to analyze the underlying formation mechanism of atmospheric pollutants; (iii) in addition, the main innovation is the introduction of the functional data analysis approach, which is performed to explore the daily pattern of pollutants. In the paper, we show the capability of the model by using O_3 pollution data from 36 pollution monitoring stations in Beijing, China.

The paper is organized as follows. In Section 2, we describe the data in the study region, and introduce the Fourier basis functions, and the functional spatio-temporal statistical model, including the implementation of model estimation and cross-validation. In Section 3, we first show the selection of covariates and basis numbers. After comparing our model with others, we show the outstanding model capability, and finally give a comprehensive interpretation of the results. Conclusions are in Section 4.

2. Material and Methods

In this section, we first describe the data in the study region. Then, we introduce the Fourier basis function, and describe the functional spatio-temporal statistical model. In particular, model equations, model estimation, and cross-validation are discussed.

2.1. Data Description

The World Health Organization (WHO) set a guideline of 100 $\mu g/m^3$ for a maximum daily 8-h average exposure to ground-level O_3; otherwise, adverse impacts on human health may occur [21]. Considering the increasing public concern on ozone, we attempt to analyze the effects from other pollutants and meteorological variables on ozone pollution, and provide some insight into the diurnal cycle of O_3, which peaks in the mid-day and reaches minimum at night-time.

In this study, we collect hourly concentration of the ground-level ozone in spring, summer, and autumn of year 2017, from thirty-six pollution monitoring stations in Beijing, China, which are directly managed by the Ministry of Environment and Protection (MEP). We also collect four other pollutant gases—particulate matter (PM_{10}), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and carbon monoxide (CO). All of the pollutant gases are measured in $\mu g/m^3$. The oxides of nitrogen (NO_x) and the volatile organic components (VOC) constitute are known to be the important precursors of the ground ozone generation [22]. However, components of VOC are not measured by the air quality monitoring network.

We also collect meteorological data: barometric pressure ($PRES$, in hectopascal), air temperature ($TEMP$, in degree celsius), dew point temperature ($DEWP$, in degree celsius), integrated rainfall ($IRAIN$, in millimeter), and integrated wind speed (Iws, in meter per second) from nine weather stations of China Meteorological Administration (CMA). All the measurements are recorded hourly. We match between air quality stations and meteorological stations by the geodesic distance. Figure 1 displays the spatial locations of the air quality stations with red dots as well as the meteorological stations with blue triangles [23]. In addition to these meteorological variables, ultraviolet radiation is also a significant meteorological factor that influences O_3 generation. Therefore, we download the data of UVB (in J/m^2) with wavelengths between 200 and 440 nanometers from the European Centre for Medium-Range Weather Forecasts (ECMWF, https://cds.climate.copernicus.eu). The UVB data are provided at a grid size of $0.25^\circ \times 0.25^\circ$ at hourly frequency available over the study region. Since the UVB data vary greatly during the day and night, we take their log-transform before adding to the model. Note that the integrated rainfall and integrated wind speed are respectively calculated by:

$$IWS_t = \begin{cases} WS_t, & WD_t != WD_{t-1}, \\ IWS_{t-1} + WS_t, & WD_t == WD_{t-1}. \end{cases} \quad (1)$$

$$IRAIN_t = \begin{cases} RAIN_t, & RAIN_t = 0, \\ IRAIN_{t-1} + RAIN_t, & RAIN_t != 0. \end{cases} \quad (2)$$

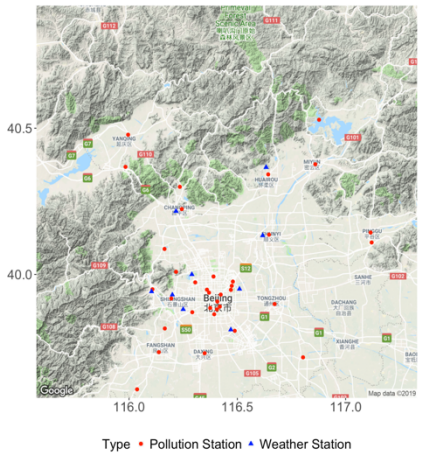

Figure 1. Thirty-six air quality monitoring stations with red dots and nine meteorological stations with blue triangles.

2.2. Fourier Basis

The basic philosophy of functional data analysis is to think of observed data functions as single entities, rather than merely as a sequence of individual observations. In practice, functional data are usually observed and recorded discretely as n pairs (t_j, y_j), and y_j is a snapshot of the function at time t_j, possibly blurred by measurement error. Time is so often the continuum over which functional data are recorded that we may slip into the habit of referring to t_j as such, but certainly other continua may be involved, such as spatial position, frequency, weight, and so forth:

$$y_j = x(t_j) + \epsilon_j \tag{3}$$

In functional data analysis, we need a strategy for constructing functions, which balances the model fitting and complexity. We built a set of functions where $\phi_k, k = 1, ..., K$ are called basis functions, and their linear combination is defined as a function:

$$x(t) = \sum_{k=1}^{K} c_k \phi_k(t) = \boldsymbol{c'\phi}(t), \tag{4}$$

the expansion of the basis function, where the parameters $c_k, k = 1, ..., K$ are the expansion coefficients to be estimated. In effect, basis expansion methods represent the potentially infinite dimensional world of functions within the finite-dimensional framework of vectors like $\boldsymbol{c}$. The functional data analysis is simplified to multivariate data analysis.

The basis functions used for data modeling mostly belong to two categories: periodic and non-periodic. Most functional data analyses involve either a Fourier basis for periodic data, or a B-spline basis for non-periodic data. Since we are interested in the diurnal variations of ozone, we introduce the Fourier basis functions in detail. In order to express the repeated pattern in long-term sequences, basis functions need to be repeated within a certain time period T. The famous basis function extension for periodic data provided by the Fourier series is:

$$\hat{x}(t) = c_0 + c_1 sin(\omega t) + c_2 cos(\omega t) + c_3 sin(2\omega t) + c_4 cos(2\omega t) + ... \tag{5}$$

where $\omega = 2\ \pi/T$. Defining a Fourier basis system requires two pieces of information: the number of basis functions K and the period T. Figure 2 shows the Fourier basis system with $K = 5$ and $T = 1$. Followed by the constant, the Fourier basis functions are arranged in consecutive sine/cosine pairs:

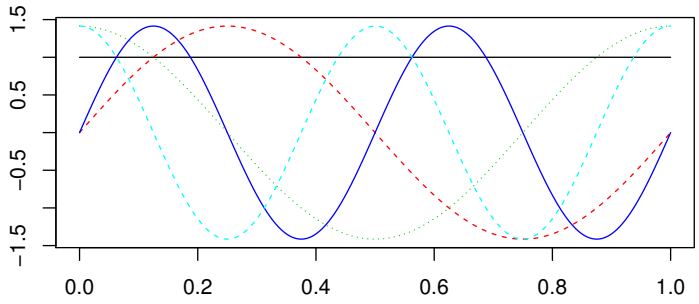

Figure 2. Fourier basis function system with $K = 5$ and $T = 1$.

We select the ozone data from one of the pollution stations—Wanliu Monitoring Station, which is located at Haidian District, Beijing, for preliminary analysis. The time span is one week from 21 May 2017 to 27 May 2017. We capture the daily variation of ozone data by using five Fourier basis functions. The mean square error (MSE) of fitted residuals is 14.79 $\mu g/m^3$. As shown in Figure 3, the predicted value at hour 24 matches the predicted value at hour 0 in the next day, guaranteeing the periodic nature of the daily cycle.

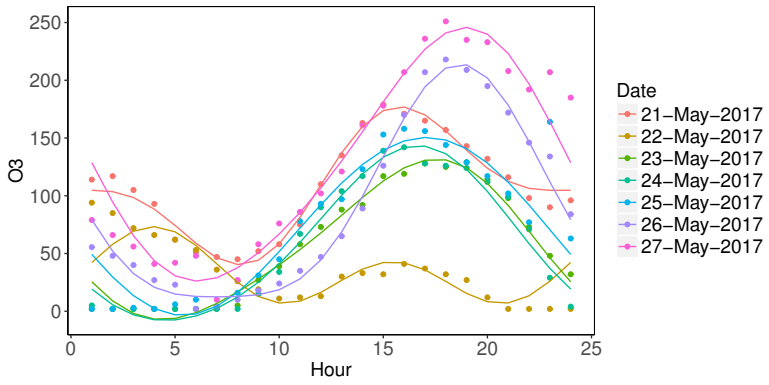

Figure 3. Ozone data fitting by using five Fourier basis functions.

2.3. Model Equation

Let $\boldsymbol{s} = (s_{lat}, s_{lon})$ be the generic spatial location on the Earth's sphere with sample size n, and $t = 1, ..., T$ the day index, and domain $\mathcal{H} = [h_1, h_2] \subset \mathbb{R}$ the time within the day expressed in hours. The model for ozone observations $O_3(\boldsymbol{s}, t, h)$ is:

$$O_3(\boldsymbol{s}, t, h) = \boldsymbol{x}(\boldsymbol{s}, t, h)' \boldsymbol{\beta}(h) + \boldsymbol{\phi}(h)' \boldsymbol{z}(\boldsymbol{s}, t) + \varepsilon(\boldsymbol{s}, t, h), \quad (6)$$

$$\boldsymbol{z}(\boldsymbol{s}, t) = \boldsymbol{G}\boldsymbol{z}(\boldsymbol{s}, t-1) + \boldsymbol{\eta}(\boldsymbol{s}, t). \quad (7)$$

This model is referred to as the functional dynamic spatio-temporal model. In Equation (6), ε is a zero-mean Gaussian measurement error independent in space and time with functional variance $\sigma^2_\varepsilon(h)$, which implies that ε is heteroskedastic across the domain $\mathcal{H}$. The variance is modeled as

$$\log(\sigma^2_\varepsilon(h)) = \boldsymbol{\phi}(h)' \boldsymbol{c}_\varepsilon, \quad (8)$$

where $\boldsymbol{\phi}(h)$ is a $p \times 1$ vector of basis functions evaluated at h while $\boldsymbol{c}_\varepsilon$ is a vector of coefficients to be estimated. In Equation (6), $\boldsymbol{x}(\boldsymbol{s}, h, t)$ is a $b \times 1$ vector of covariates while $\boldsymbol{\beta}(h) = (\beta_1(h), ..., \beta_b(h))'$ is the vector of functional parameters modeled as

$$\beta_j(h) = \boldsymbol{\phi}(h)' \boldsymbol{c}_{\beta,j}, \quad (9)$$

and $\boldsymbol{c}_\beta = \left(\boldsymbol{c}'_{\beta,1}, ..., \boldsymbol{c}'_{\beta,b}\right)'$ is the $pb \times 1$ vector of coefficients to be estimated. Additionally, $\boldsymbol{z}(\boldsymbol{s}, t)$ is a $p \times 1$ latent space-time variable with Markovian dynamics given in Equation (7). Matrix $\boldsymbol{G}$ is a diagonal transition matrix with diagonal elements in the $p \times 1$ vector $\boldsymbol{g}$. The vector $\boldsymbol{\eta}$ is described by a multivariate Gaussian process independent in time but correlated across space with matrix spatial covariance function given by

$$\boldsymbol{\Gamma}(\boldsymbol{s}, \boldsymbol{s}'; \boldsymbol{\theta}) = diag\left(v_1 \rho(\boldsymbol{s}, \boldsymbol{s}'; \boldsymbol{\theta}_1), ..., v_p \rho(\boldsymbol{s}, \boldsymbol{s}'; \boldsymbol{\theta}_p)\right), \quad (10)$$

and $\boldsymbol{v} = (v_1, ..., v_p)'$ is the vector of scale coefficients while $\rho(\boldsymbol{s}, \boldsymbol{s}'; \boldsymbol{\theta}_j)$ is a valid spatial correlation function for locations $\boldsymbol{s}, \boldsymbol{s}' \in \mathbb{S}^2$ parametrized by $\boldsymbol{\theta}_j$, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_p)'$. The unknown model parameter vector is $\boldsymbol{\psi} = \left(\boldsymbol{c}'_\varepsilon, \boldsymbol{c}'_\beta, \boldsymbol{g}', \boldsymbol{v}', \boldsymbol{\theta}'\right)$.

In Figure 4, we summarize the methodology. The main innovation is to incorporate the function data analysis approach to the classic spatio-temporal statistical model, which facilitates exploring the intra-day fluctuations of ozone pollution as well as the functional effects of covariates. Note that, in order to ease the notation, the same p-dimensional basis functions $\boldsymbol{\phi}(h)$ are used to model σ^2_ε, β_j and $\boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s}, t)$ in Equations (6) and (7). In the empirical analysis, we choose different numbers of basis functions for modeling according to the model criteria, such as the mean square error (MSE), and R^2 (see Section 2.5 for details).

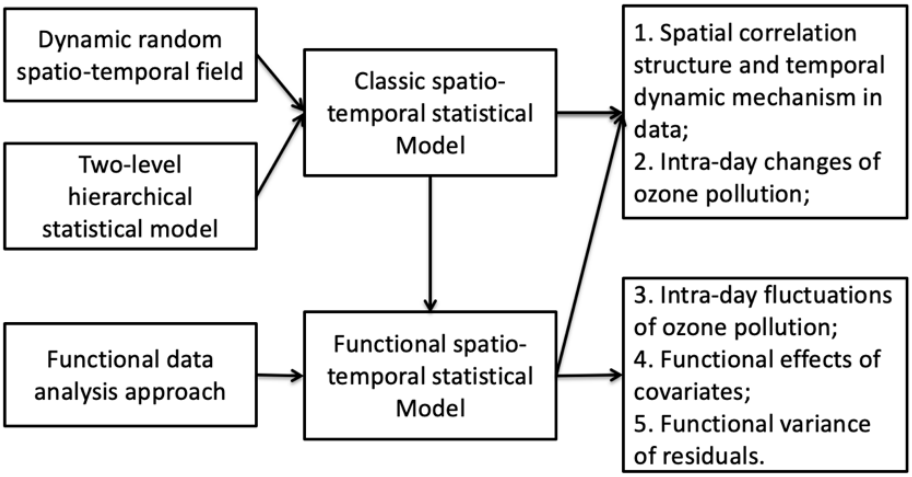

Figure 4. Methodology summary.

2.4. Model Estimation

The estimation of $\boldsymbol{\psi}$ and the latent space-time variable $\boldsymbol{z}(\boldsymbol{s},t)$ is based on the maximum likelihood approach and Kalman filter. At a specific location $\boldsymbol{s}_i$ and time t, q measurements are taken at hour points $\boldsymbol{h} = (1,2,...,q)'$ and collected in the vector

$$\boldsymbol{y}_{\boldsymbol{s}_i,t} = (O_3(\boldsymbol{s}_i,t,1),...,O_3(\boldsymbol{s}_i,t,q))', \tag{11}$$

where $q = 24$ as pollutants are hourly recorded. Daily profiles of ozone data observed at time t across spatial locations $\mathcal{S}$ are then stored in the vector $\boldsymbol{y}_t = (\boldsymbol{y}'_{s_1,t},...,\boldsymbol{y}'_{s_n,t})'$. Accordingly, Equations (6) and (7) are rewritten as

$$\boldsymbol{y}_t = \tilde{\boldsymbol{X}}_t\boldsymbol{c}_\beta + \boldsymbol{\Phi}_{z,t}\boldsymbol{z}_t + \boldsymbol{\varepsilon}_t, \tag{12}$$

$$\boldsymbol{z}_t = \tilde{\boldsymbol{G}}\boldsymbol{z}_{t-1} + \boldsymbol{\eta}_t, \tag{13}$$

where $\tilde{\boldsymbol{X}}_t = \boldsymbol{X}_t\boldsymbol{\Phi}_{\beta,t}$ is a $nq \times bp$ matrix, with $\boldsymbol{X}_t$ the matrix of covariates and $\boldsymbol{\Phi}_{\beta,t}$ the basis matrix for $\boldsymbol{\beta}$. $\boldsymbol{\Phi}_{z,t}$ is the $nq \times np$ basis matrix for the latent $np \times 1$ vector $\boldsymbol{z}_t = (\boldsymbol{z}(\boldsymbol{s}_1,t)',...,\boldsymbol{z}(\boldsymbol{s}_n,t)')'$. $\boldsymbol{\eta}_t = (\boldsymbol{\eta}(\boldsymbol{s}_1,t)',...,\boldsymbol{\eta}(\boldsymbol{s}_n,t)')'$ is the $np \times 1$ innovation vector, while $\boldsymbol{\varepsilon}_t$ is the $nq \times 1$ vector of measurement errors. Additionally, $\tilde{\boldsymbol{G}} = \boldsymbol{G} \otimes \boldsymbol{I}_n$ is the $np \times np$ diagonal transition matrix.

The complete-data likelihood function $L(\boldsymbol{\psi};\boldsymbol{Y},\boldsymbol{Z})$ can be written as

$$L(\boldsymbol{\psi};\boldsymbol{Y},\boldsymbol{Z}) = L(\boldsymbol{\psi}_{z_0};z_0)\prod_{t=1}^{T} L(\boldsymbol{\psi}_y;\boldsymbol{y}_t|\boldsymbol{z}_t)L(\boldsymbol{\psi}_z;\boldsymbol{z}_t|\boldsymbol{z}_{t-1}), \tag{14}$$

where $\boldsymbol{Y} = (\boldsymbol{y}_1,...,\boldsymbol{y}_T)$, $\boldsymbol{Z} = (z_0,\boldsymbol{z}_1,...,\boldsymbol{z}_T)$, $\boldsymbol{\psi}_z = (\boldsymbol{g}',\boldsymbol{v}',\boldsymbol{\theta}')$, $\boldsymbol{\psi}_y = \left(\boldsymbol{c}'_\varepsilon,\boldsymbol{c}'_\beta\right)$, and z_0 is the Gaussian initial vector with parameter $\boldsymbol{\psi}_{z_0}$. The model parameter set $\boldsymbol{\psi}$ is initialized with starting values $\boldsymbol{\psi}^{\langle 0\rangle}$ and then updated at each iteration ι of the EM algorithm. The algorithm terminates if any of the conditions is satisfied

$$\left\|\boldsymbol{\psi}^{\langle\iota\rangle} - \boldsymbol{\psi}^{\langle\iota-1\rangle}\right\| / \left\|\boldsymbol{\psi}^{\langle\iota\rangle}\right\| < \epsilon, \tag{15}$$

$$\left|L(\boldsymbol{\psi}^{\langle\iota\rangle};\boldsymbol{Y}) - L(\boldsymbol{\psi}^{\langle\iota-1\rangle};\boldsymbol{Y})\right| / \left|L(\boldsymbol{\psi}^{\langle\iota\rangle};\boldsymbol{Y})\right| < \epsilon, \tag{16}$$

where $\|\|$ is the l_2 norm, $\boldsymbol{\psi}^{\langle \iota \rangle}$ is the parameter set at the ι-th iteration, $L(\boldsymbol{\psi}^{\langle \iota \rangle}; \boldsymbol{Y})$ is the observed-data likelihood function evaluated at $\boldsymbol{\psi}^{\langle \iota \rangle}$, and ϵ is a small positive number (e.g., 10^{-3}).

The EM algorithm provides a point estimate of the parameter vector $\boldsymbol{\psi}$ but without uncertainty information. Note that $\boldsymbol{Y}$ is a vector with dimension $N = nqT$. Generally speaking, inverting the full variance–covariance matrix of the N-dimensional data vector $\boldsymbol{Y}$ has a computational complexity in the order of $O(N^3)$, which is clearly unfeasible. Thanks to the state space representation of model, we estimate the variance–covariance matrix $\hat{\Sigma}_{\boldsymbol{\psi}} = \mathbb{V}(\boldsymbol{\psi} \mid \boldsymbol{Y})$ from the marginal likelihood, which may be used for model selection and inference.

2.5. Cross-Validation

We implement a 2-fold cross-validation by partitioning the original spatial locations $\mathcal{S}$ into subsets $\mathcal{S}_{est}$ and $\mathcal{S}_{xval}$. Data related to $\mathcal{S}_{est}$ are used for model estimation while data related to $\mathcal{S}_{xval}$ are used for cross-validation. The cross-validation mean squared errors are then computed by

$$MSE_{\boldsymbol{s}} = \frac{1}{B} \sum_{t=1}^{T} \sum_{h \in \boldsymbol{h}_{\boldsymbol{s},t}} \left(O_3(\boldsymbol{s}, h, t) - \hat{O}_3(\boldsymbol{s}, h, t) \right)^2, \tag{17}$$

where $\hat{O}_3(\boldsymbol{s}, h, t) = E_{\hat{\phi}}(O_3(\boldsymbol{s}, h, t) \,|\boldsymbol{Y})$ is the prediction of ozone data at the cross-validation stations, and B is the number of terms in each sum. We also obtain the cross-validation R^2 with respect to station $\boldsymbol{s}$:

$$R^2_s = 1 - \frac{MSE_s}{VAR(\{O_3(\boldsymbol{s},h,t),t,h\})} \tag{18}$$

The choice of the numbers of basis functions is very essential for model estimation. Here, based on the cross-validated mean square error and other model criteria, we choose the reasonable numbers of basis functions to estimate σ^2_ε, β_j and $\boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s}, t)$ respectively. After implementing leave-one-station-out cross-validation, we take the average $\overline{MSE}$ and $\overline{R^2}$ as our criteria:

$$\overline{MSE} = \frac{1}{n} \sum_{i=1}^{n} MSE_{\boldsymbol{s}_i}, \tag{19}$$

$$\overline{R^2} = \frac{1}{n} \sum_{i=1}^{n} R^2_{\boldsymbol{s}_i}. \tag{20}$$

3. Analysis of O_3 Pollution in Beijing

In the paragraph, we first show the selection of covariates and basis numbers with application to ozone data in Beijing, and then focus on the summertime modeling. By comparing our proposed model with other models, we show the outstanding advantage of the functional spatio-temporal statistical model. Finally, we show the model results and interpret the parameter estimates, especially the functional effects of covariates.

3.1. Selection of Covariates and Basis Numbers

In the following text, we select the covariates in $\boldsymbol{x}(\boldsymbol{s}, t, h)$ by using the Akaike information criterion (AIC). Table 1 displays the results of forward selection based on AIC, which means starting with no covariates, and iteratively adding the most contributive covariates. For instance, in the summertime modeling, at the beginning (Iter 0), we select the variable NO_2, which results in the best model performance with maximum AIC. Then, at the next iteration (Iter 1), the variable particulate matter

(PM_{10}) is further selected. Table 1 shows that the importance of covariates varies among seasons, but the most important variables are SO_2, NO_2, and PM_{10}.

Table 1. The selection of model covariates according to AIC.

	Iteration									
Season	**Iter 0**	**Iter 1**	**Iter 2**	**Iter 3**	**Iter 4**	**Iter 5**	**Iter 6**	**Iter 7**	**Iter 8**	**Iter 9**
Spring	NO2	SO2	PM10	CO	PRES	UVB	Iws	IRAIN	TEMP	DEWP
Summer	NO2	PM10	SO2	TEMP	IRAIN	UVB	DEWP	PRES	CO	Iws
Autumn	NO2	SO2	TEMP	DEWP	PM10	UVB	CO	Iws	PRES	IRAIN

The ozone concentrations display a significant seasonal pattern, being pretty high in summer, while meanwhile being moderate in winter [24,25]. Therefore, we focus on the analysis of O_3 pollution in summer. Figure 5 shows the maximum AIC at each iteration for summertime modeling.

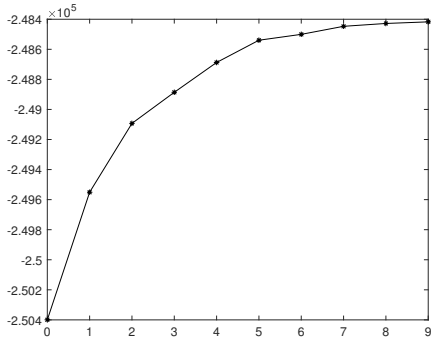

Figure 5. Improvement of AIC at each iteration for summertime modeling.

The improvement of model AIC is no longer significant after five iterations; therefore, we find the optimal subset of covariates—NO_2, PM_{10}, SO_2, $TEMP$, $IRAIN$, and UVB. Hence, the measurement equation for ozone data is

$$\begin{aligned} O_3(\boldsymbol{s},h,t) \quad = \quad & \beta_0(h) + x_{NO2}(\boldsymbol{s},h,t)\beta_{NO2}(h) + x_{PM_{10}}(\boldsymbol{s},h,t)\beta_{PM_{10}}(h) \\ & + x_{SO_2}(\boldsymbol{s},h,t)\beta_{SO_2}(h) + x_{TEMP}(\boldsymbol{s},h,t)\beta_{TEMP}(h) \\ & + x_{IRAIN}(\boldsymbol{s},h,t)\beta_{IRAIN}(h) + x_{UVB}(\boldsymbol{s},h,t)\beta_{UVB}(h) \\ & + \boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t) + \varepsilon(\boldsymbol{s},h,t), \end{aligned} \tag{21}$$

where data are available at $h = 1,...,24$, $s \in \{s_1,...,s_{36}\}$, and $t = 1,...,92$. Moreover, due to the circularity of time, Fourier basis functions are adopted. This implies that $\beta_j(h)$, $\sigma_\varepsilon^2(h)$ are periodic functions. Under the different combinations of the numbers of basis functions, the model criteria $\overline{MSE}$ and $\overline{R^2}$ in Section 2.5 are obtained and shown in Table 2.

Table 2. Criteria $\overline{MSE}$, $\overline{R^2}$, and AIC under different numbers of Fourier basis.

$\boldsymbol{\phi}(h)'\boldsymbol{z}(s,t)$	$\boldsymbol{\beta}(h)$	σ_ε^2	$\overline{\boldsymbol{MSE}}$	$\overline{\boldsymbol{R^2}}$	$\boldsymbol{AIC}$
5	3	3	357.58	0.9206	−255,385
5	3	5	356.32	0.9209	−254,607
5	3	7	356.47	0.9208	−254,599
5	5	3	352.61	0.9215	−254,235
5	5	5	352.25	0.9216	−253,459
5	5	7	352.39	0.9215	−253,454
5	7	3	352.14	0.9215	−254,116
5	7	5	351.62	0.9217	−253,309
5	7	7	351.76	0.9217	−253,306
7	3	3	332.95	0.9259	−249,514
7	3	5	331.88	0.9261	−248,723
7	3	7	331.99	0.9261	−248,713
7	5	3	330.06	0.9264	−248,848
7	5	5	329.80	0.9264	−248,066
7	5	7	329.90	0.9264	−248,062
7	7	3	329.09	0.9266	−248,656
7	7	5	328.97	0.9266	−247,879
7	7	7	329.05	0.9266	−247,876
9	3	3	324.08	0.9278	−246,673
9	3	5	323.13	0.9280	−245,937
9	3	7	323.19	0.9280	−245,928
9	5	3	322.07	0.9281	−246,056
9	5	5	321.80	0.9282	−245,327
9	5	7	321.86	0.9282	−245,322
9	7	3	321.28	0.9283	−245,879
9	7	5	321.12	0.9283	−245,152
9	7	7	321.13	0.9283	−245,150

From the table, when the number of basis functions for estimating $\boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t)$ increases, it significantly reduces the $\overline{MSE}$. When the number of basis functions increases from 5 to 7, the $\overline{MSE}$ is reduced more than that from 7 to 9. Considering such enormous calculation stress, we choose seven basis functions to estimate the latent component $\boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t)$. However, increasing the number of basis functions for the variance $\sigma_\varepsilon^2(h)$ of the residual $\varepsilon(\boldsymbol{s},h,t)$ does not significantly reduce $\overline{MSE}$ but is helpful to improve the AIC. We find that an increase from 3 to 5 has an improvement in AIC, but the improvement becomes very minor from 5 to 7. Thus, we choose five basis functions to estimate the variance $\sigma_\varepsilon^2(h)$. Finally, we choose five basis functions to estimate the effects from covariates $\beta_j(h)$, considering the trade-off between the model interpretation and over-fitting problem. Based on the analysis above, the number of basis functions for $\beta_j(h)$, $\sigma_\varepsilon^2(h)$ and $\boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t)$ is chosen to be 5, 5, and 7, respectively.

3.2. Model Comparison

In the paragraph, we compare the five models, namely Equations (22), (23), (24), (25), and (26). Equation (22) is an ordinary regression model; Equation (23) is a regression model with functional $\beta(h)$ estimates; Equation (24) introduces the latent spatio-temporal variable $\boldsymbol{z}(\boldsymbol{s},t)$ to characterize the spatio-temporal correlation; Equation (25) is a simplified version of the proposed functional spatio-temporal statistical model that is $\beta(h) \equiv \beta$, $\sigma_\varepsilon^2(h) \equiv \sigma_\varepsilon^2$; Equation (26) is the functional spatio-temporal statistical model:

$$O_3 = X\beta + \epsilon \tag{22}$$

$$O_3(h) = X(h)\beta(h) + \epsilon(h) \tag{23}$$

$$O_3(s,t) = X(s,t)\beta + z(s,t) + \epsilon(s,t) \tag{24}$$

$$z(s,t) = Gz(s,t-1) + \eta(s,t)$$

$$\begin{aligned} O_3(s,h,t) &= X(s,h,t)\beta + \boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t) + \epsilon(s,h,t) \\ z(s,t) &= Gz(s,t-1) + \eta(s,t) \end{aligned} \tag{25}$$

$$\begin{aligned} O_3(s,h,t) &= X(s,h,t)\beta(h) + \boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t) + \epsilon(s,h,t) \\ z(s,t) &= Gz(s,t-1) + \eta(s,t) \end{aligned} \tag{26}$$

Similar to the selection of the numbers of basis functions, the average $\overline{MSE}$ and $\overline{R^2}$, and AIC are used to assess the model performance. As shown in Table 3, our model Equation (26) is the optimal among the five models in view of the three model criteria. Equation (23) is much improved from Equation (22) in terms of $\overline{MSE}$ and $\overline{R^2}$, which means a better model forecast in general. Benefiting from the latent spatio-temporal variable $\boldsymbol{z}(\boldsymbol{s},t)$, Equation (24) has an unbeatable advantage over the ordinary regression models, accessing much smaller $\overline{MSE}$ and much larger $\overline{R^2}$ and AIC. Equation (25) introduces the functional data analysis approach, and characterizes the latent component as a linear combination of the basis functions and the latent random spatio-temporal variable $\boldsymbol{z}(\boldsymbol{s},t)$. Although the AIC is only a little increased, a smaller $\overline{MSE}$ and larger $\overline{R^2}$ are achieved. Eventually, when Equation (26) adds the functional covariate effects $\beta(h)$ and the functional variance of the residuals $\sigma_\epsilon(h)$, $\overline{MSE}$, and $\overline{R^2}$ is not improved much. However, AIC is further improved, which benefits from the more capable interpretation of covariates and the flexibility of the residual variance.

Table 3. $\overline{MSE}$, $\overline{R^2}$, and AIC for the five models.

	Number of Basis			Model Criteria				
	β	$\boldsymbol{\phi}(h)'z(s,t)$	σ_ϵ	$\overline{MSE}$	$\overline{R^2}$	AIC	logL [1]	Npar [2]
Equation (22)	0	0	0	1880.54	0.5863	−414,714	−414,700	7
Equation (23)	5	0	0	1171.55	0.7423	−395,874	−395,812	31
Equation (24)	0	0	0	552.7	0.879	−256,960	−256,940	10
Equation (25)	0	7	0	336.88	0.925	−252,426	−252,370	28
Equation (26)	5	7	0	329.8	0.9264	−248,066	−247,954	56

[1] log likelihood, [2] number of parameters.

In Equation (26), firstly, the latent hidden variable $\boldsymbol{z}(\boldsymbol{s},t)$ captures the spatial correlation by range parameter θ, and variance parameter v, which shows that an average standard deviation of 48 $\mu g/m^3$ of ozone data are explained by $\boldsymbol{z}(\boldsymbol{s},t)$ (refer to Table 5). Secondly, the functional $\beta(h)$ shows that the covariate effects are both significant and nonlinear, indicating the complicated formation of ozone pollution by using the functional representation (refer to Figure 6). In summary, the hierarchical spatio-temporal statistical model, combined with functional data analysis approach, contributes to the high amount of $\overline{R^2}$.

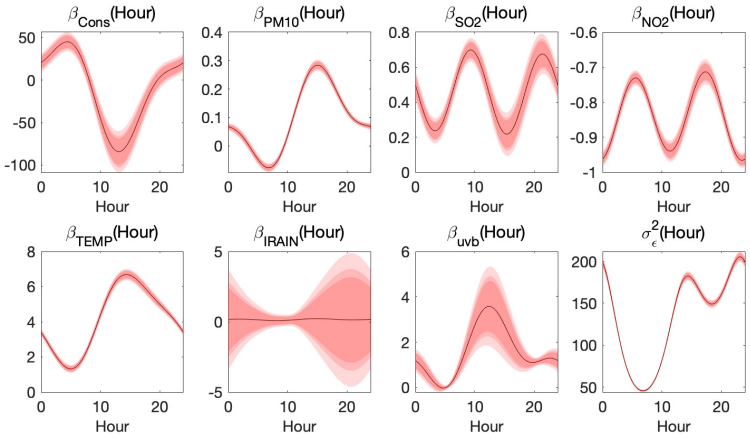

Figure 6. Estimated $\beta_{cons}(hour), \beta_{PM10}(hour), \beta_{SO2}(hour), \beta_{NO2}(hour), \beta_{TEMP}(hour), \beta_{IRAIN}(hour), \beta_{UVB}(hour)$ and $\sigma^2_\epsilon(hour)$, with 90%, 95%, *and*99%- confidence bands.

3.3. Model Result

Figure 6 shows the estimated $\boldsymbol{\beta}(h)$ and $\sigma^2_\epsilon(h)$ for model Equation (21). Thanks to Fourier basis functions, the estimation result at the end of the day matches the beginning of the next day.Since, in general, the confidence bands of estimated $\boldsymbol{\beta}(h)$ may contain zero, it may be useful to test the significance of covariates. The χ^2 tests are introduced as follows:

$$\beta_j(h) = \boldsymbol{\phi}(h)'\boldsymbol{c}_{\beta,j},\ \boldsymbol{c}_{\beta,j} \sim N(0, \Sigma_{\boldsymbol{c}_{\hat{\beta},j}}). \tag{27}$$

Thus, $\boldsymbol{c}'_{\beta,j}\Sigma^{-1}_{\boldsymbol{c}_{\hat{\beta},j}}\boldsymbol{c}_{\beta,j} \sim \chi^2(rank(\Sigma_{\boldsymbol{c}_{\hat{\beta},j}}))$. In Figure 6, $IRAIN$ fluctuates around zero. The results of χ^2 tests for the significance of covariates are reported in Table 4, and indicate that the effect of variable $IRAIN$ is not jointly significant.

Table 4. χ^2 tests for significance of fixed effects.

Covariate	χ^2 Statistic	p-Value
Cons	282.77	0
PM_{10}	2114.06	0
SO_2	1048.50	0
NO_2	29,032.23	0
$TEMP$	5554.16	0
$IRAIN$	0.91	0.96
UVB	30,934	0

Therefore, it comes to the final model equation by excluding the $IRAIN$ variable:

$$\begin{aligned} O_3(\boldsymbol{s},h,t) \quad = \quad & \beta_0(h) + x_{NO2}(\boldsymbol{s},h,t)\beta_{NO2}(h) + x_{PM_{10}}(\boldsymbol{s},h,t)\beta_{PM_{10}}(h) \\ & + x_{SO_2}(\boldsymbol{s},h,t)\beta_{SO_2}(h) + x_{TEMP}(\boldsymbol{s},h,t)\beta_{TEMP}(h) \\ & + x_{UVB}(\boldsymbol{s},h,t)\beta_{UVB}(h) + \boldsymbol{\phi}(h)'\boldsymbol{z}(\boldsymbol{s},t) + \varepsilon(\boldsymbol{s},h,t). \end{aligned} \tag{28}$$

In Table 5, we show the estimates and standard deviation of parameters relevant to the latent spatio-temporal variable $\mathbf{z}(s,t)$, which are the transition coefficient $\mathbf{g}$, range parameter $\boldsymbol{\theta}$, and variance vector $\mathbf{v}$. Most estimates of $\mathbf{g}$ parameter are positive, and the absolute values are all within one, which guarantees the stability of the 7-variate spatio-temporal process $\mathbf{z}(s,t)$. Compared with the geodesic distance of Beijing (around 50 km), the values of $\boldsymbol{\theta}$ parameter, ranging from 31.92 km to 63.12 km, indicate a strong spatial correlation within the city. The average $\mathbf{v}$ estimate is around 2313 (with standard deviation of 48 $\mu g/m^3$), and shows that the latent variable $\mathbf{z}(s,t)$ accounts for much

more proportion of original O_3 variance than the unexplained term $\sigma_\epsilon^2(h)$. Hence, introducing the latent spatio-temporal variable $\mathbf{z}(s,t)$ guarantees the advantage of the proposed model.

Table 5. Estimates and standard error of parameter g, θ, and v.

	Transition g		θ [km]		Variance v	
	Est	**Std.err**	**Est**	**Std.err**	**Est**	**Std.err**
Basis 1	0.739	0.018	63.12	4.57	8422.14	549.12
Basis 2	0.229	0.026	50.94	0.96	3799.47	176.33
Basis 3	0.179	0.03	36.98	1.02	2027.63	106.59
Basis 4	0.034	0.032	36.34	0.54	896.86	50.61
Basis 5	0.106	0.034	39.75	0.84	702.64	41.65
Basis 6	0.043	0.043	31.92	0.87	191.09	13.53
Basis 7	−0.210	0.042	37.10	0.35	151.80	10.78

Finally, in Figure 7, we show the estimated $\boldsymbol{\beta}(h)$ and $\sigma_\epsilon^2(h)$. The last figure is the plot of functional variance $\sigma_\epsilon^2(h)$, which represents the unexplained portion of O_3 variance. The plot shows that the model is more capable when explaining the situation during the daytime [26].

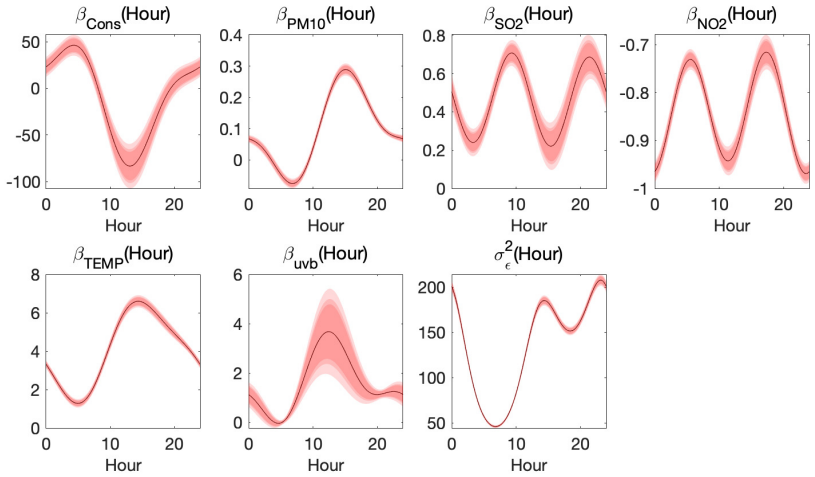

Figure 7. Estimated $\beta_{cons}(hour)$, $\beta_{PM10}(hour)$, $\beta_{SO2}(hour)$, $\beta_{NO2}(hour)$, $\beta_{TEMP}(hour)$, $\beta_{UVB}(hour)$ and $\sigma_\epsilon^2(hour)$, with 90%, 95%, and 99%- confidence bands.

As shown in Figure 7, the coefficient curves of $TEMP$, uvb and PM_{10} are similar, increasing from early morning and attending the peak at 12:00 p.m.–2:00 p.m., then falling down. Focusing on daytime, we see that the three curves are consistent with the trend of temperature (or uvb), which implicates that the relationship between ozone and temperature (or uvb) might be quadratic [27], or there were interactions between temperature and uvb, that is, the ozone concentrations were dependent on $TEMP^2$, uvb^2, or $TEMP \times uvb$. The coefficients of $TEMP$ and uvb in daytime are positive, which is consistent with the present research [28]. While the coefficient of PM_{10} is negative at 5:00 a.m.–10:00 a.m. and positive during other time periods. The positive correlation between PM_{10} and ozone may be caused by their common sources, secondary nature, and interactions of their precursors [29], and the negative correlation could be explained by PM's consumption of hydroperoxy (HO2) radicals, which would otherwise react with NO for ozone generation [30]. Furthermore, the positive correlation becomes the strongest at 3:00 p.m., at which time the ozone concentration attains the largest.

In addition, the coefficient curves of NO_2 and SO_2 both have two spikes, while the coefficient of NO_2 is negative and the other is positive. The negative relationship between NO_2 and ozone is consistent with results in many studies [31,32], and the positive correlation between SO_2 and ozone could be explained by their common dependences on meteorology [33]. The strongest correlation between NO_2 and ozone in daytime appears at about 11:00 a.m., and the weakest correlation appears

at 5:00 a.m. and 6:00 p.m. In contrast, the correlation between SO_2 and ozone is the strongest at 9:00 a.m. and 8:00 p.m., in other words, approximately the end of morning/evening rush hours in Beijing, respectively, and such correlation is the weakest at 3:00 p.m.

4. Conclusions

In this paper, we propose a functional spatio-temporal statistical method to analyze air quality data, and explore the mechanism of pollution formation.

- The method has several advantages. First, as a hierarchical spatio-temporal statistical model, it is flexible enough to handle latent variable while capturing spatio-temporal dynamics. Second, the proposed model also takes covariates into consideration, thereby being efficient in discovering relational patterns from chemical reaction, and meteorological factors on the formation of O_3 pollution. Third, in the framework of the spatio-temporal models, we are the first to explore the intra-day variation of ozone through the functional data analysis approach, which is the most innovative part of the model.
- The model has made the following progresses. First of all, our model outperforms other models in many ways, as shown in Section 3.2. Second, the latent spatio-temporal variable $\mathbf{z}(s,t)$ well captures the temporal dynamic and spatial structure of ozone data. Third, from the functional effects of the covariates, we explore the possible effects of air pollutants and meteorological variables on ozone data.
- Our model is flexible enough to model any kind of data with spatio-temporal structure; therefore, it can be applied in many fields, such as economy and agriculture, apart from the environment. The introduction of the functional data analysis approach in the functional spatio-temporal model is not restricted to model the daily pattern of the data, and provides us more capability to explore the nature of the data of our interest.

Author Contributions: Conceptualization, Y.W. and K.X.; Methodology, Y.W. and K.X.; Software, Y.W. and K.X.; Validation, Y.W. and K.X.; Writing, Y.W. and K.X.; Writing-Review & Editing, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by China's National Key Research Special Program Grant 2016YFC0207701. Ke Xu is supported by "the Fundamental Research Funds for the Central Universities" in UIBE (No. 19QD22).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lippmann, M. Health effects of ozone a critical review. *Japca* **1989**, *39*, 672–695. [CrossRef] [PubMed]
2. Bell, M.L.; McDermott, A.; Zeger, S.L.; Samet, J.M.; Dominici, F. Ozone and short-term mortality in 95 US urban communities, 1987–2000. *JAMA* **2004**, *292*, 2372–2378. [CrossRef] [PubMed]
3. Ashmore, M.R.; Marshall, F. Ozone impacts on agriculture: An issue of global concern. In *Advances in Botanical Research*; Elsevier: Amsterdam, The Netherlands, 1998; Volume 29, pp. 31–52.
4. Simpson, D.; Arneth, A.; Mills, G.; Solberg, S.; Uddling, J. Ozone—The persistent menace; interactions with the N cycle and climate change. *Curr. Opin. Environ. Sustain.* **2014**, *9*, 9–19. [CrossRef]
5. Chan, C.K.; Yao, X. Air pollution in mega cities in China. *Atmos. Environ.* **2008**, *42*, 1–42. [CrossRef]
6. Hu, H.; Yang, Q.; Lu, X.; Wang, W.; Wang, S.; Fan, M. Air pollution and control in different areas of China. *Crit. Rev. Environ. Sci. Technol.* **2010**, *40*, 452–518. [CrossRef]
7. Liang, X.; Zou, T.; Guo, B.; Li, S.; Zhang, H.; Zhang, S.; Huang, H.; Chen, S.X. Assessing Beijing's PM2.5 pollution: Severity, weather impact, APEC and winter heating. *Proc. R. Soc. A* **2015**, *471*, 20150257. [CrossRef]
8. Lin, J.; Zhang, A.; Chen, W.; Lin, M. Estimates of Daily PM2.5 Exposure in Beijing Using Spatio-Temporal Kriging Model. *Sustainability* **2018**, *10*, 2772. [CrossRef]
9. Wang, T.; Xue, L.; Brimblecombe, P.; Lam, Y.F.; Li, L.; Zhang, L. Ozone pollution in China: A review of concentrations, meteorological influences, chemical precursors, and effects. *Sci. Total Environ.* **2017**, *575*, 1582–1596. [CrossRef]

10. Li, J.; Lu, K.; Lv, W.; Li, J.; Zhong, L.; Ou, Y.; Chen, D.; Huang, X.; Zhang, Y. Fast increasing of surface ozone concentrations in Pearl River Delta characterized by a regional air quality monitoring network during 2006–2011. *J. Environ. Sci. China* **2014**, *26*, 23–36. [CrossRef]
11. Xue, L.K.; Wang, T.; Gao, J.; Ding, A.J.; Zhou, X.H.; Blake, D.R.; Wang, X.F.; Saunders, S.M.; Fan, S.J.; Zuo, H.C.; et al. Ground-level ozone in four Chinese cities: Precursors, regional transport and heterogeneous processes. *Atmos. Chem. Phys.* **2014**, *14*, 13175–13188. [CrossRef]
12. Wang, W.N.; Cheng, T.H.; Gu, X.F.; Chen, H.; Guo, H.; Wang, Y.; Bao, F.W.; Shi, S.Y.; Xu, B.R.; Zuo, X.; et al. Assessing Spatial and Temporal Patterns of Observed Ground-level Ozone in China. *Sci. Rep.* **2017**, *7*, 3651. [CrossRef]
13. Cressie, N.; Wikle, C.K. *Statistics for Spatio-Temporal Data*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
14. Huang, H.C.; Cressie, N. Spatio-temporal prediction of snow water equivalent using the Kalman filter. *Comput. Stat. Data Anal.* **1996**, *22*, 159–175. [CrossRef]
15. Cressie, N.; Calder, C.A.; Clark, J.S.; Hoef, J.M.V.; Wikle, C.K. Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecol. Appl.* **2009**, *19*, 553–570. [CrossRef]
16. Ramsay, J.O.; Silverman, B.W. *Applied Functional Data Analysis: Methods and Case Studies*; Springer: Berlin, Germany, 2007.
17. Wood, S.N. *Generalized Additive Models: An Introduction with R*; CRC Press: Boca Raton, FL, USA, 2017.
18. Kalman, R.E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **1960**, 35–45. [CrossRef]
19. Krishnan, T.; McLachlan, G. The EM algorithm and extensions. *Wiley* **1997**, *1*, 58–60.
20. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*; Springer: Berlin, Germany, 2017.
21. WHO Regional Office for Europe. *Air Quality Guidelines Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide*; World Health Organization: Geneva, Switzerland, 2006.
22. Sillman, S. The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments. *Atmos. Environ.* **1999**, *33*, 1821–1845. [CrossRef]
23. Kahle, D.; Wickham, H. Ggmap: Spatial Visualization with ggplot2. *R J.* **2013**, *5*, 144–161. [CrossRef]
24. Pf, C.; Q, Z.; Jn, Q.; Y, G.; My, H. Temporal and Spatial Distribution of Ozone Concentration by Aircraft Sounding over Beijing. *Environ. Sci.* **2012**, *33*, 4141.
25. Dufour, G.; Eremenko, M.; Orphal, J.; Flaud, J.M. IASI observations of seasonal and day-to-day variations of tropospheric ozone over three highly populated areas of China: Beijing, Shanghai, and Hong Kong. *Atmos. Chem. Phys.* **2010**, *10*, 3787–3801. [CrossRef]
26. Dohan, J.; Masschelein, W. The photochemical generation of ozone: Present state–of–the–art. *Ozone Sci. Eng.* **1987**, 315–334. [CrossRef]
27. Belan, B.D.; Savkin, D.E.; Tolmachev, G.N. Air-Temperature Dependence of the Ozone Generation Rate in the Surface Air Layer. *Atmos. Ocean. Opt.* **2018**, *31*, 187–196. [CrossRef]
28. Awang, N.R.; Ramli, N.A.; Yahaya, A.S.; Elbayoumi, M. Multivariate methods to predict ground level ozone during daytime, nighttime, and critical conversion time in urban areas. *Atmos. Pollut. Res.* **2015**, *6*, 726–734. [CrossRef]
29. Lamarque, J.; Kiehl, J.T.; Hess, P.G.; Collins, W.D.; Emmons, L.K.; Ginoux, P.; Luo, C.; Tie, X. Response of a coupled chemistry-climate model to changes in aerosol emissions: Global impact on the hydrological cycle and the tropospheric burdens of OH, ozone, and NOx. *Geophys. Res. Lett.* **2005**, *32*. [CrossRef]
30. Li, K.; Jacob, D.J.; Liao, H.; Zhu, J.; Shah, V.; Shen, L.; Bates, K.H.; Zhang, Q.; Zhai, S. A two-pollutant strategy for improving ozone and particulate air quality in China. *Nat. Geosci.* **2019**, *12*, 906–910. [CrossRef]
31. Chou, C.C.K.; Liu, S.C.; Lin, C.Y.; Shiu, C.J.; Chang, K.H. The trend of surface ozone in Taipei, Taiwan, and its causes: Implications for ozone control strategies. *Atmos. Environ.* **2006**, *40*, 3898–3908. [CrossRef]
32. Geng, F.; Tie, X.; Xu, J.; Zhou, G.; Peng, L.; Gao, W.; Tang, X.; Zhao, C. Characterizations of ozone, NOx, and VOCs measured in Shanghai, China. *Atmos. Environ.* **2008**, *42*, 6873–6883. [CrossRef]
33. National Research Council. *Rethinking the Ozone Problem in Urban and Regional Air Pollution*; National Academies Press: Washington, DC, USA, 1992.