mathematics

MDPI

Article

Semiparametric Integrated and Additive Spatio-Temporal Single-Index Models

Hamdy F. F. Mahmoud [1,2] and Inyoung Kim [1,*]

[1] Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; ehamdy@vt.edu
[2] Department of Statistics, Mathematics, and Insurance, Faculty of Commerce, Assiut University, Assiut 71515, Egypt
[*] Correspondence: inyoungk@vt.edu; Tel.: +1-540-231-5366

Abstract: In this paper, we introduce two semiparametric single-index models for spatially and temporally correlated data. Our first model has spatially and temporally correlated random effects that are additive to the nonparametric function, which we refer to as the "semiparametric spatio-temporal single-index model (ST-SIM)". The second model integrates the spatially correlated effects into the nonparametric function, and the time random effects are additive to the single-index function. We refer to our second model as the "semiparametric integrated spatio-temporal single-index model (IST-SIM)". Two algorithms based on a Markov chain expectation maximization are introduced to simultaneously estimate the model parameters, spatial effects, and time effects of the two models. We compare the performance of our models using several simulation studies. The proposed models are then applied to mortality data from six major cities in South Korea. Our results suggest that IST-SIM (1) is more flexible than ST-SIM because the former can estimate various nonparametric functions for different locations, while ST-SIM enforces the mortality functions having the same shape over locations; (2) provides better estimation and prediction, and (3) does not need restrictions for the single-index coefficients to fix the identifiability problem.

Keywords: Markov chain expectation maximization; semiparametric regression models; nonparametric regression models; single-index model; spatio-temporal correlated data

MSC: 62P12; 62G05

check for updates

Citation: Mahmoud, H.F.F.; Kim, I. Semiparametric Integrated and Additive Spatio-Temporal Single-Index Models. *Mathematics* **2023**, *11*, 4629. https://doi.org/10.3390/math11224629

Academic Editors: Antonio Musolesi and Alexandra Soberon

Received: 27 September 2023
Revised: 31 October 2023
Accepted: 5 November 2023
Published: 13 November 2023

1. Introduction

Epidemiology has a long history of studying factors that affect the variability of mortality. These factors include geographical or spatial variations, which play a crucial role in evaluating healthcare distribution. Spatio-temporal analysis offers additional benefits over spatial analysis by enabling researchers to simultaneously study patterns over time. Due to recent advancements in computational methods for analyzing spatio-temporal data, spatio-temporal data analysis has emerged as a prominent research area.

Numerous studies [1–16] have introduced parametric statistical methods for modeling spatio-temporal data, such as the generalized linear mixed model, generalized linear additive model, and the spatio-temporal auto-regressive model. On the other hand, mixed-effects models and spatio-temporal models differ in terms of the structure of the variance-covariance matrix, which defines the type of correlation between time random effects and spatial dependencies. In spatio-temporal data, the covariance matrix structure of spatial effects depends on the distance between any two locations (or cities) and follows a parametric function form. The estimation of covariance functions and the prediction of spatial effects have also been studied [9,11,12,15,17]. Both parametric and nonparametric models have been developed. Parametric models provide convenient and interpretable results; however, they lack flexibility due to strong parametric assumptions. These assumptions are often not satisfied in real data applications. Nonparametric and semiparametric models relax these assumptions, making them more suitable for analyzing spatio-temporal data.

Hence, in this article, we focus on semiparametric modeling for spatio-temporal data and introduce two semiparametric models. These two models are built based on the single-index model (SIM), which incorporates spatial and time effects into the model.

The authors of Ichimura [18] introduced the single-index model, and many articles have extensively studied the estimation and inference of SIM [19–23]. SIM offers several advantages over parametric models: (1) it assumes that the function describing the relationship between the response variable and the explanatory variables is unknown, thereby avoiding the misleading results of misspecifying the link function [24]; (2) it does not assume a specific type of error distribution, and (3) it mitigates the curse of dimensionality problem by reducing the p-dimensional explanatory variables to a single dimension using the single-index linear combination.

In SIM estimation, the parameters require restrictions to address the identifiability problem. One possible solution to this problem is to set the norm of the parameters equal to 1 [25,26], while another solution is to set the coefficient of the first explanatory variable equal to 1 [2,18]. This article employs the second approach.

In Pang and Xue [27], random effects were considered as additive effects to the single-index function; however, these random effects were assumed to be independent. In contrast, in this article, we introduce correlated spatial and temporal effects and incorporate these random effects into the single-index function. We develop two models: the spatio-temporal single-index model (ST-SIM), in which correlated spatial and temporal effects are additive to the single-index function, and the integrated spatio-temporal single-index model (IST-SIM), in which the spatial effect is integrated into the single-index function, and spatial effects are added to the unknown function. To the best of our knowledge, the IST-SIM model is not reported in the statistical literature. This article primarily focuses on IST-SIM.

The organization of this article is as follows: Section 2 presents the proposed models, ST-SIM and IST-SIM. Section 3 describes the two estimation algorithms. Section 4 contains the simulation studies we conducted. Section 5 describes the real data application and presents the results of applying the two models to the motivating data. Conclusions are provided in Section 6.

2. Semiparametric Models

In this section, we introduce two semiparametric models: ST-SIM and IST-SIM. In ST-SIM, the spatial effects are additive to the single-index function, while in IST-SIM, the spatial effects are non-additive to the single-index function.

Let $Y_{s,t}$ be the observed value of the response variable at location s and time point t, u_s ($s = 1, \ldots, r$) be a spatial random effect following Gaussian process (GP), w_t ($t = 1, \ldots, n$) be the time effect, and $\mathbf{x}_{1(s,t)}, \mathbf{x}_{2(s,t)}, \ldots, \mathbf{x}_{p(s,t)}$ be the p observed values of the explanatory variables at location s and time point t.

2.1. Semiparametric Spatio-Temporal Single-Index Model

The additive ST-SIM model is defined as follows:

$$
\begin{aligned}
Y_{s,t}|\mu_{s,t} &\sim P_d(\mu_{s,t}|u_s, \omega_t), \\
\mu_{s,t}|u_s, \omega_t &= g(X_{s,t}\boldsymbol{\beta}) + u_s + \omega_t, \\
u_s &\sim GP(0, \sigma_u^2 \Sigma(s, s')),\ s \neq s' \\
\omega_t &\sim P_d(0, \sigma_\omega^2 \Omega),
\end{aligned}
\tag{1}
$$

where $g(\cdot)$ is a smoothed unknown function, $\boldsymbol{\beta}$ are the single-index coefficient parameters, $\Sigma(s, s')$ is the (s, s')th component of the covariance function Σ of the spatial effects, Ω is the covariance function of a probability distribution (P_d) of the time effects, and σ_u^2 and σ_ω^2 are the variances of the spatial and time effects, respectively. Given the effects u_s and ω_t, $Y_{s,t}$ follows a probability distribution (P_d) with mean $\mu_{s,t}$.

This model assumes that the effects u_s and ω_t are additive to the single-index function $g(\cdot)$. The single-index coefficient parameters ($\boldsymbol{\beta}$), spatial and time effects (u_s, ω_t), and the unknown function ($g(\cdot)$) must be estimated simultaneously. A restriction on $\boldsymbol{\beta}$ is needed to resolve the identifiability problem. Two possible approaches include setting one of the parameters of $\boldsymbol{\beta}$ equal to 1 [2,18] or to use $||\boldsymbol{\beta}|| = 1$ [25,26,28]. The first approach is employed in this article.

2.2. Semiparametric Integrated Spatio-Temporal Single-Index Model

The IST-SIM is defined as follows:

$$\begin{aligned} Y_{s,t}|\mu_{s,t} &\sim P_d(\mu_{s,t}|u_s, \omega_t), \\ \mu_{s,t}|u_s, \omega_t &= g(u_s + X_{s,t}\boldsymbol{\beta}) + \omega_t, \\ u_s &\sim GP(0, \sigma_u^2 \Sigma(s,s')),\ s \neq s' \\ \omega_t &\sim P_d(0, \sigma_\omega^2 \Omega), \end{aligned} \tag{2}$$

where $g(\cdot)$ is a smoothed unknown function, $\boldsymbol{\beta}$ represents the single-index coefficient parameters, $\Sigma(s,s')$ is the (s,s')th component of the covariance function Σ of the spatial effects, Ω is the covariance function of a probability function (P_d) of the time effect, and σ_u^2 and σ_ω^2 are the variances of the spatial and time effects, respectively. Given u_s and ω_t, $Y_{s,t}$ follows some probability distribution (P_d) with a mean of $\mu_{s,t}$.

One of the advantages of this model is that it does not have an identifiability problem, unlike the additive model (1). Consequently, all of the parameters can be estimated. This is because the spatial effect u_s has a coefficient equal to 1, allowing all of the single-index coefficient parameters to be estimated without requiring additional assumptions for the unknown function. This model assumes that the spatial effect u_s is integrated into the single-index function, while the time effect ω_t is additive to the mean of the unknown function $g(\cdot)$. Additionally, $\boldsymbol{\beta}$, u_s, ω_t, and $g(\cdot)$ must be estimated simultaneously.

2.3. Covariance Functions of Spatial and Temporal Effects

The spatial covariance function that describes the spatial correlation between any two locations, u_s and $u_{s'}$, can be described as

$$\begin{aligned} \mathbf{u} &\sim MN(\mathbf{0}, \sigma_u^2 \Sigma_\rho), \\ Cov(u_s, u_s') &= \sigma_u^2 \Sigma_\rho(s,s'), \end{aligned} \tag{3}$$

where MN represents a multivariate normal distribution, σ_u^2 is the variance of the spatial effects, $\Sigma_\rho(s,s')$ is the (s,s')th component of the spatial covariance matrix Σ_ρ, which is assumed to have a known parametric covariance function to guarantee positive definite, and ρ is the dependence range that can be estimated by the semivariogram. The dependence range, ρ, represents the distance between two locations such that the spatial effects within that range are correlated and, outside of that range, the correlation is assumed to be negligible.

In time series, it is common to use the autoregressive moving average, autoregressive integrated moving average, or a random walk to model the time effects. In random walk, the relationship between two consecutive time points, say ω_t, and ω_{t-1}, takes the form:

$$\omega_t = \omega_{t-1} + \epsilon_t, \tag{4}$$

where ϵ_t is the random noise term accounting for the difference between two consecutive time points within the location s. When ω_t follows a normal distribution [29], where n

locations exist, the vector of the temporal effects $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n)^T$ takes the following joint probability distribution:

$$f(\boldsymbol{\omega}|\sigma^2_\omega) \propto exp\left(-\frac{\sigma^2_\omega}{2}\omega^T\Omega^{-1}\boldsymbol{\omega}\right), \tag{5}$$

where Ω^{-1} is the inverse of the temporal covariance matrix. For example, when we have four-time points, the covariance matrix, Ω, based on the description above, takes the form

$$\Omega = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

This precision matrix is singular, so the covariance matrix Ω^{-1}, cannot be obtained. Hence, we propose a modified version to overcome this problem

$$\Omega = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

In this form, all of the diagonal elements are equal to 2, which means ω_1 and ω_4 are also random values and the other time points follow the first-order random walk. Other possible functions can be used, such as the Gaussian function with $\rho = 2$ [30], which means the time effect ω_t depends only on the following and previous time effects, ω_{t+1} and ω_{t-1}. The Gaussian process of the temporal effects, in this case, can be described as

$$\begin{aligned} \boldsymbol{\omega} &\sim MN(\mathbf{0}, \sigma^2_\omega\Omega_{\rho=2}), \\ Cov(\omega_t, \omega_{t'}) &= \sigma^2_\omega\Omega_{\rho=2}(t, t'), \end{aligned} \tag{6}$$

where ω_t and $\omega_{t'}$ are two temporal effects at two different time points, t and t', at location s.

3. Model Estimation

In this section, two estimation algorithms are introduced to estimate the ST-SIM and IST-SIM. The Monte Carlo expectation maximization (MCEM) algorithms are implemented.

3.1. Estimating Spatial and Time Effects

The expectation maximization (EM) algorithm consists of two steps: expectation (the E-step) and maximization (the M-step). Estimates are obtained by iteratively performing these until convergence is achieved. This algorithm is commonly used for estimating generalized linear mixed models [31–35]. However, for our proposed models, there is no closed form that is available for the expectation part. Therefore, we incorporated Markov chain Monte Carlo (MCMC) to generate random samples from the full conditional distributions of $\mathbf{u}$ and $\boldsymbol{\omega}$. In the E-step, we employed the Metropolis–Hastings (M-H) algorithm. As a result, we developed an MCEM algorithm for estimating the proposed models.

The Metropolis–Hastings (M-H) algorithm employs the conditional multivariate normal distribution to generate random samples from the spatial effects $\mathbf{u}$ and time effects $\boldsymbol{\omega}$, as dictated by the following complete-data log-likelihood function:

$$\log f(\mathbf{Y}, \mathbf{u}, \boldsymbol{\omega}|\boldsymbol{\mu}, \sigma^2_u\Sigma, \sigma^2_\omega\Omega)\& = \&\log f_{Y|u,\omega}[\mathbf{Y}|\boldsymbol{\mu}, \mathbf{u}, \boldsymbol{\omega}] + \log f_u[\mathbf{u}|\sigma^2_u\Sigma] + \log f_\omega[\boldsymbol{\omega}|\sigma^2_\omega\Omega], \tag{7}$$

where $\mathbf{Y} \sim \text{Pois}[\boldsymbol{\mu}|\mathbf{u}, \boldsymbol{\omega}]$, $\boldsymbol{\mu}|\mathbf{u}, \boldsymbol{\omega} = g(X\boldsymbol{\beta}) + Z_1\mathbf{u} + Z_2\boldsymbol{\omega}$ for the additive model and $\boldsymbol{\mu}|\mathbf{u}, \boldsymbol{\omega} = g(Z_1\mathbf{u} + X\boldsymbol{\beta}) + Z_2\boldsymbol{\omega}$ for the integrated spatial effects model; $\mathbf{u} \sim GP(\mathbf{0}, \sigma^2_u\Sigma)$, $\boldsymbol{\omega} \sim GP(\mathbf{0}, \sigma^2_\omega\Omega)$, $\sigma^2_\omega\Omega = Cov(\omega_t, \omega_{t+\delta}) = \sigma^2_\omega exp(||\delta||^2/\rho_\omega = 2)$ for all $t, \delta \in R$; and $\sigma^2_u\Sigma = Cov(u_s, u_{s+d}) = \sigma^2_u exp(||d||^2/\rho_u)$ for all $s, d \in R^2$.

The single-component Metropolis–Hastings algorithm is employed, where, at each iteration, only a single component of the spatial or time effects is updated. The proposed conditional distribution of the time effect ω_t and its parameters can be derived as follows:

Let $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n) = \left[\omega_1\ \omega_2\right]^T$ have a multivariate normal distribution with mean $\mathbf{0}$ and a variance-covariance matrix $\sigma_0^2\Omega$, where

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}.$$

Then, the proposal distribution of ω_1 conditioned on $\boldsymbol{\omega}_2 = \mathbf{a}$ is a normal $(\omega_1|\boldsymbol{\omega}_2 = \mathbf{a}) \sim N(\bar{\mu}_{\omega_1}, \sigma_0^2\bar{\Omega})$, where σ_0^2 is the proposed variance of the time effects, $\bar{\mu}_{\omega_1} = \Omega_{12}\Omega_{22}^{-1}\mathbf{a}$, and the covariance matrix $\bar{\Omega} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$.

With this proposed conditional distribution, and given a vector of spatial effects $\mathbf{u}$, and assuming that the spatial and time effects are independent, the acceptance probability rule takes the form

$$\min\left\{\frac{f(\mathbf{Y}|\boldsymbol{\mu}, \mathbf{u}, \omega_t^*) f_\omega(\omega_t^*|\bar{\mu}_{\omega_t^*}, \sigma_\omega^2\bar{\Omega}) f_u(\mathbf{u}|\sigma_u^2\Sigma)}{f(\mathbf{Y}|\boldsymbol{\mu}, \mathbf{u}, \omega_t) f_\omega(\omega_t|\bar{\mu}_{\omega_t}, \sigma_\omega^2\bar{\Omega}) f_u(\mathbf{u}|\sigma_u^2\Sigma)}, 1\right\}, \tag{8}$$

where $f_\omega(\omega_t|\bar{\mu}_{\omega_t}, \sigma_\omega^2\bar{\Omega})$ is the conditional distribution of ω_t given the other time effects, $\omega_1, \omega_2, \ldots, \omega_{t-1}, \omega_{t+1}, \ldots, \omega_n$. Similarly, one can drive the proposed conditional distribution of the spatial effect u_t given the other spatial effects as $(u_1|\mathbf{u_2} = \mathbf{a}) \sim N(\bar{\mu}_{u_1}, \sigma_0^2\bar{\Sigma})$, where σ_0^2 is the proposed variance of the spatial effects, $\bar{\mu}_{u_1} = \Sigma_{12}\Sigma_{22}^{-1}\mathbf{a}$, and the covariance matrix $\bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. With this proposed conditional distribution and a given vector of time effects $\boldsymbol{\omega}$, the acceptance probability rule takes the form

$$\min\left\{\frac{f(\mathbf{Y}|\boldsymbol{\mu}, u_s^*, \boldsymbol{\omega}) f_\omega(\boldsymbol{\omega}|\sigma_\omega^2\Omega) f_u(u_s^*|\bar{\mu}_{u_s^*}, \sigma_u^2\bar{\Sigma})}{f(\mathbf{Y}|\boldsymbol{\mu}, u_s, \boldsymbol{\omega}) f_\omega(\boldsymbol{\omega}|\sigma_\omega^2\Omega) f_u(u_s|\bar{\mu}_{u_s}, \sigma_u^2\bar{\Sigma})}, 1\right\}.$$

The following steps make up a subroutine for simulating random samples from the spatial and time effects. This subroutine is used within the algorithms to estimate the model parameters of the proposed models.

Step 0: Initialize $\mathbf{u}^{(\mathbf{0})}$ and $\boldsymbol{\omega}^{(0)}$, and, given the estimates of σ_u^2, σ_ω^2, and $\boldsymbol{\mu}_{s,t}$, set $s = 1$, $t = 1$, and $m = 1$.

Step 1: Let $\boldsymbol{\omega}$ = mean($\boldsymbol{\omega}^{[0:(m-1)]}$), generate a value from the spatial effect proposal distribution at location s, u_s^*, and generate a value from uniform(0,1), U.

$$\text{If U} < \min\left\{\frac{f(\mathbf{Y}|\boldsymbol{\mu}, u_s^*, \boldsymbol{\omega}) f_\omega(\boldsymbol{\omega}|\sigma_\omega^2\Omega) f_u(u_s^*|\bar{\mu}_{u_s^*}, \sigma_u^2\bar{\Sigma})}{f(\mathbf{Y}|\boldsymbol{\mu}, u_s, \boldsymbol{\omega}) f_\omega(\boldsymbol{\omega}|\sigma_\omega^2\Omega) f_u(u_s|\bar{\mu}_{u_s}, \sigma_u^2\bar{\Sigma})}, 1\right\}, \text{set } \mathbf{u}^{(m)} = (u_s^*, u_2, \ldots, u_S);$$

otherwise, $\mathbf{u}^{(m)} = \mathbf{u}$. Set $s = s + 1$ and repeat this step until all locations are visited.

Step 2: Given $\mathbf{u}$= mean($\mathbf{u}^{[0:(m-1)]}$), generate a value from the proposal time effects distribution of time t, ω_t^*, and generate a value from uniform (0,1), U.

$$\text{If U} < \min\left\{\frac{f(\mathbf{Y}|\boldsymbol{\mu}, \mathbf{u}, \omega_t^*) f_\omega(\omega_t^*|\bar{\mu}_{\omega_t^*}, \sigma_\omega^2\bar{\Omega}) f_u(\mathbf{u}|\sigma_u^2\Sigma)}{f(\mathbf{Y}|\boldsymbol{\mu}, \mathbf{u}, \omega_t) f_\omega(\omega_t|\bar{\mu}_{\omega_t}, \sigma_\omega^2\bar{\Omega}) f_u(\mathbf{u}|\sigma_u^2\Sigma)}, 1\right\}, \text{set } \boldsymbol{\omega}^{(m)} = (\omega_t^*, \omega_2, \ldots, \omega_n);$$

otherwise, $\mathbf{u}^{(i)} = \mathbf{u}$. Set $t = t + 1$ and repeat this step until all time points are visited.

Step 3: Repeat Steps 1–2, a large number of times, M-, and, based on Geyer [36], discard a percentage of burn-in between 1% and 2% of the total number of iterations, M, and use the rest, N_0, to estimate the spatial and time effects.

3.2. ST-SIM Estimation Algorithm

The following steps comprised the proposed algorithm for estimating the additive ST-SIM's parameters and the spatial and time effects:

I-step Initialize parameters:

(a) Obtain initial values for $\mathbf{u}$, $\boldsymbol{\omega}$, σ_u^2, and σ_ω^2 ($\mathbf{u}^{(0)}$, $\boldsymbol{\omega}^{(0)}$, $\sigma_u^{2(0)}$, $\sigma_\omega^{2(0)}$).

(b) Calculate $\mathbf{Y}^* = \mathbf{Y} - Z_1\mathbf{u}^{(0)} - Z_2\boldsymbol{\omega}^{(0)}$, obtain the single-index coefficient estimate $\boldsymbol{\beta}^{(0)}$, and estimate the unknown function using a smoothing method to obtain $\hat{g}(\cdot)^{(0)}$.

E-step Given the initial values from the I-step, simulate random samples for each spatial effect u_s, $(u_s^1, u_s^2, \ldots, u_s^N)$, and for each time effect ω_t, $(\omega_t^1, \omega_t^2, \ldots, \omega_t^N)$, from $f(\mathbf{Y}, \mathbf{u}, \boldsymbol{\omega}|\boldsymbol{\mu}, \sigma_u^2\Sigma, \sigma_\omega^2\Omega)$, where $\boldsymbol{\mu} = g(X_{s,t}\boldsymbol{\beta}) + Z_1\mathbf{u} + Z_2\boldsymbol{\omega}$ via the subroutine described in Section 3.1.

M-step Maximize $\sum_k \log f(\mathbf{u}^k|\sigma_u^2\Sigma)$ and $\sum_k \log f(\boldsymbol{\omega}^k|\sigma_\omega^2\Omega)$,

(a) Obtain $\sigma_u^{2(1)}$ and $\sigma_\omega^{2(1)}$, and calculate $\mathbf{u}^{(1)} = \frac{1}{N_0}\sum_{k=1}^{N_0}\mathbf{u}^k$, $\boldsymbol{\omega}^{(1)} = \frac{1}{N_0}\sum_{k=1}^{N_0}\boldsymbol{\omega}^k$ and $\mathbf{Y}^* = \mathbf{Y} - Z_1\mathbf{u}^{(1)} - Z_2\boldsymbol{\omega}^{(1)}$.

(b) Using the Ichimura method, estimate $\boldsymbol{\beta}$ $\boldsymbol{\beta}^{(1)}$, and smooth the unknown function $g(\cdot)$ to obtain $\hat{g}(\cdot)^{(1)}$.

Do iteration of the E-step and M-step until convergence is achieved. The stopping rule for the EM algorithm is $|\frac{LogLikelihood^{(t)} - LogLikelihood^{(t-1)}}{LogLikelihood^{(t-1)}}| < 0.001$, where $LogLikelihood^{(t)}$ is the log of the likelihood function at the iterated step t, and $LogLikelihood^{(t-1)}$ is the log of the likelihood function at the iterated step $t-1$.

3.3. IST-SIM Estimation Algorithm

The algorithm described in Section 3.2 does not work for the IST-SIM due to two main issues: (1) the long computation time resulting from the intensive calculations (specifically, when using the Metropolis–Hastings algorithm with only 1000 iterations, the single-index function needs to be estimated 36,000 times to run the MCEM algorithm just once); and (2) the inability to separate the impact of spatial random effects on the acceptance ratio from the effect of the single-index coefficient parameters when comparing the current and previous spatial effect values. In the following, we provide a detailed explanation of these issues.

The acceptance ratio takes the form

$$\min\left\{\frac{f(\mathbf{Y}|\mathbf{u}^*, \boldsymbol{\omega}, \boldsymbol{\mu} = \hat{g}^*[X\hat{\boldsymbol{\beta}}^* + Z_1\mathbf{u}^*] + Z_2\boldsymbol{\omega})f_u(\mathbf{u}^*|\sigma_u^2\Sigma)f_\omega(\boldsymbol{\omega}|\sigma_\omega^2\Omega)}{f(\mathbf{Y}|\mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\mu} = \hat{g}[X\hat{\boldsymbol{\beta}}^* + Z_1\mathbf{u}] + Z_2\boldsymbol{\omega})f_u(\mathbf{u}|\sigma_u^2\Sigma)f_\omega(\boldsymbol{\omega}|\sigma_\omega^2\Omega)}, 1\right\}.$$

If this ratio is greater than 1, it is unclear whether this is because $\mathbf{u}^*$ is better than $\mathbf{u}$, $\hat{g}^*$ is superior to $\hat{g}$, or $\hat{\boldsymbol{\beta}}^*$ is more accurate than $\hat{\boldsymbol{\beta}}$. In such cases, determining which of the two proposed spatial random effects represents a generated value from the true spatial effect distribution becomes challenging. One possible solution to this problem is to employ a Taylor series approximation for the unknown function at a specific value $(X\boldsymbol{\beta}^{(0)} + Z\mathbf{u}^{(0)})$ to isolate the spatial effect from the model-parameter effect on the ratio, as follows:

$$\boldsymbol{\mu} = g(X\boldsymbol{\beta} + Z_1\mathbf{u}) + Z_2\boldsymbol{\omega} = \hat{g}(\cdot) + \hat{g}'(\cdot)(X\boldsymbol{\beta} + Z_1\mathbf{u} - X\boldsymbol{\beta}^{(0)} - Z_1\mathbf{u}^{(0)}) + Z_2\boldsymbol{\omega}.$$

Here, $\hat{g}(\cdot)$ represents the unknown function, estimated using a smoothing method, like p-spline [37] or kernel smoothing [38], and $\hat{g}'(\cdot)$ is the estimate of the first derivative of this unknown function. In this paper, we employ the local linear kernel method to estimate both the unknown function and its derivative.

The following steps comprise the proposed algorithm for estimating the IST-SIM's parameters and the spatial and time effects:

I-step Initialize the parameters:

(a) Obtain initial values for $\mathbf{u}$, $\boldsymbol{\omega}$, σ_u^2, and σ_ω^2 ($\mathbf{u}^{(0)}$, $\boldsymbol{\omega}^{(0)}$, $\sigma_u^{2(0)}$, $\sigma_\omega^{2(0)}$);

(b) Calculate $\mathbf{Y}^* = \mathbf{Y} - Z_2\boldsymbol{\omega}^{(0)}$, obtain the single-index coefficient estimates $\boldsymbol{\beta}^{(0)}$, and estimate the unknown function using a smoothing method to obtain $\hat{g}(\cdot)^{(0)}$ and its derivative $\hat{g}'(\cdot)^{(0)}$.

E-step Given the initials from the I-step and the Taylor approximation of the unknown function, simulate random samples for each spatial effect u_s, $(u_s^1, u_s^2, \ldots, u_s^N)$, and for each time effect ω_t, $(\omega_t^1, \omega_t^2, \ldots, \omega_t^N)$, from $f(\mathbf{Y}, \mathbf{u}, \boldsymbol{\omega}|\boldsymbol{\mu}, \sigma_u^2\Sigma, \sigma_\omega^2\Omega)$, where $\boldsymbol{\mu} \approx \hat{g} + \hat{g}'(X\boldsymbol{\beta} + Z_1\mathbf{u} - X\boldsymbol{\beta}^{(0)} - Z_1\mathbf{u}^{(0)}) + Z_2\boldsymbol{\omega}$, via the subroutine described in Section 3.1

M-step Maximize $\sum_k \log f(\mathbf{u}^k|\sigma_u^2\Sigma)$ and $\sum_k \log f(\boldsymbol{\omega}^k|\sigma_\omega^2\Omega)$,

(a) Obtain $\sigma_u^{2(1)}$ and $\sigma_\omega^{2(1)}$, and calculate $\mathbf{u}^{(1)} = \frac{1}{N_0}\sum_{k=1}^{N_0}\mathbf{u}^k$, $\boldsymbol{\omega}^{(1)} = \frac{1}{N_0}\sum_{k=1}^{N_0}\boldsymbol{\omega}^k$ and $\mathbf{Y}^* = \mathbf{Y} - Z_2\boldsymbol{\omega}^{(1)}$.

(b) Using the Ichimura method, estimate $\boldsymbol{\beta}$ $\boldsymbol{\beta}^{(1)}$, and smooth the unknown function $g(\cdot)$ to obtain $\hat{g}(\cdot)^{(1)}$.

Do iteration of the E-step and M-step until convergence is achieved. The stopping rule for the EM algorithm is $|\frac{LogLikelihood^{(t)} - LogLikelihood^{(t-1)}}{LogLikelihood^{(t-1)}}| < 0.001$, where $LogLikelihood^{(t)}$ is the log of the likelihood function at iteration step t, and $LogLikelihood^{(t-1)}$ is the log of the likelihood function at iteration step $t-1$.

4. Simulation

In this section, we evaluate the performance of the two proposed model algorithms in terms of estimating the model parameters, fitting the data, and predicting under correct and misspecified model specifications. The performance of the IST-SIM estimation algorithm is assessed through the simulation of 100 data sets based on the following integrated model:

$$
\begin{aligned}
y_{is}|\mu_s, u_s, \omega_t &\sim p_d(y_s|\mu_s, u_s), \\
\mu_s|u_s, \omega_t &= g(\beta_1 x_{1is} + \beta_2 x_{2is} + u_s) + \omega_t, \\
u_s &\sim GP(0, \sigma_u^2\Sigma(s, s')) \text{ and } \omega_t \sim GP(0, \sigma_\omega^2\Omega(t, t')),
\end{aligned}
$$

where $t = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, r$, with six locations ($r = 6$) and 12 time points at each location ($n = 12$). In addition, x_1 is generated from uniform (5, 20), and x_2 is generated from a standard normal distribution. The true model parameters are $(\beta_1, \beta_2, \sigma_u^2, \sigma_\omega^2) = (1, 1, 0.5, 0.5)$, and two cases of the dependence range are considered ($\rho_u = 1$ and $\rho_u = 3$) in the domain $[0, 3] \times [0, 3]$. In addition, the performance of the ST-SIM algorithm was studied by simulating 100 data sets from the same setting, with model mean equal to $\mu_s|u_s, \omega_t = g(\beta_1 x_{1is} + \beta_2 x_{2is}) + u_s + \omega_t$. The IST-SIM and ST-SIM algorithms are used for estimating the model parameters. For the spatial effects and temporal effects, the initial values are generated from $N(0, 1)$, and to generate strictly positive starting points for σ_u^2 and $\sigma^2\omega$, an inverse-gamma distribution was used, $IG(1, 1)$. Table 1 shows that the two proposed algorithms worked fine in estimating the model parameters, with the mean of the 100 estimates being close to the true parameter values and small standard error values. The mean square error (MSE) of the IST-SIM was greater than that of the ST-SIM model. A possible reason is because of the Taylor approximation used for the unknown function. These findings were the same under both dependence ranges ($\rho_u = 1$ and $\rho_u = 3$). In addition, all parameters or the single-index coefficients can be estimated for the IST-SIM, but for the ST-SIM, one of the parameters is set to 1 to address the identifiability problem, $\beta_1 = 1$.

We conducted another simulation study to assess the performance of the proposed models in fitting and predicting data under both correct and misspecified model settings.

For each model, we used the mean square error (MSE) to evaluate the fitting quality and the predicted residual sum of squares (PRESS) to assess the prediction accuracy. We evaluated each model's performance in two scenarios: when the data were generated from the true model and when the data were not generated from the true model.

Table 1. Results of 100 simulated data sets: mean ± standard error (SE), median, and interquartile range (IQR) of the 100 model parameters estimates and MSE of the two proposed models (IST-SIM and ST-SIM) at different values of the dependence range ($\rho = 1$ and $\rho = 3$).

	Model		True	Mean ± SE	MSE	Median	IQR
$\rho = 1$	ST-SIM	β_2	1	0.993 ± 0.007	0.006	0.998	0.071
		σ_u^2	0.5	0.451 ± 0.005	0.006	0.439	0.066
		σ_t^2	0.5	0.455 ± 0.001	0.006	0.541	0.017
		R^2	-	0.989 ± 0.000	-	0.989	0.003
	IST-SIM	β_1	1	1.002 ± 0.043	0.119	0.923	0.286
		β_2	1	1.015 ± 0.044	0.123	0.937	0.325
		σ_u^2	0.5	0.486 ± 0.077	0.127	0.310	0.370
		σ_t^2	0.5	0.457 ± 0.007	0.006	0.452	0.047
		R^2	-	0.989 ± 0.000	-	0.989	0.003
$\rho = 3$	ST-SIM	β_2	1	0.989 ± 0.006	0.004	0.988	0.088
		σ_u^2	0.5	0.437 ± 0.004	0.006	0.428	0.061
		σ_t^2	0.5	0.457 ± 0.001	0.002	0.452	0.021
		R^2	-	0.9888 ± 0.000	-	0.988	0.003
	IST-SIM	β_1	1	0.994±0.069	0.089	0.866	0.233
		β_2	1	0.989 ± 0.066	0.101	0.856	0.252
		σ_u^2	0.5	0.407 ± 0.042	0.088	0.298	0.027
		σ_t^2	0.5	0.476 ± 0.005	0.004	0.461	0.029
		R^2	-	0.989 ± 0.000	-	0.989	0.004

For instance, in the context of the IST-SIM described earlier, we generated 100 data sets from the model and calculated the MSE and PRESS for both models. We followed the same approach for the ST-SIM. The results in Table 2 demonstrate that when the true model was IST-SIM (i.e., when the data were simulated from the IST-SIM), the IST-SIM significantly outperformed the ST-SIM in terms of the mean MSE (123.8 vs. 266.4) and the mean PRESS (371.8 vs. 587.8), with smaller standard errors compared to the ST-SIM. Conversely, when the true model was the ST-SIM (i.e., when the data were simulated from the ST-SIM), the IST-SIM still performed better in terms of the mean MSE (124.5 vs. 135.7) and exhibited smaller standard errors, with their mean PRESS values being comparable (388.6 vs. 372.78). The code used for generating the simulation results can be provided upon receiving a reasonable request.

Table 2. The mean square error (MSE) and the predicted residual sum of squares (PRESS) results of the proposed models of 100 simulated data sets from each true model.

		True Model			
		ST-SIM		IST-SIM	
Criterion	Fitted Model	Mean ± SE	Median	Mean ± SE	Median
MSE	ST-SIM	135.7 ± 5.11	135.0	266.4 ± 16.6	236.2
	IST-SIM	124.5 ± 4.69	125.5	123.8 ± 6.17	130.1
PRESS	ST-SIM	372.7 ± 36.2	275.8	587.8 ± 45.9	446.8
	IST-SIM	388.6± 40.7	267.4	371.8 ± 25.0	312.9

5. Application

These two models are applied to real data from South Korea, covering the period from 2000 to 2007. The data set includes multiple daily recorded variables, such as mortality, temperature, humidity, pressure, and time, for six major cities in South Korea (Busan, Seoul, Daejeon, Incheon, Gwangju, and Daegu). Figure 1 shows the locations of these six cities in South Korea.

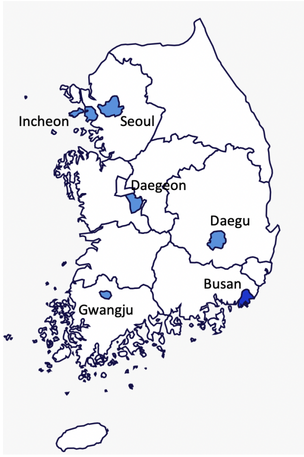

Figure 1. The 6 major cities locations in South Korea: Seoul, Busan, Daegu, Incheon, Gwangju, and Daejeon.

5.1. Data and Models

The non-accidental mortality (the number of deaths excluding deaths due to accidents), mean temperature, mean humidity, mean pressure, and time were recorded daily for six major cities in South Korea: Busan, Seoul, Daejeon, Incheon, Gwangju, and Daegu. However, we utilized monthly data by calculating the means of the weather variables per month. We opted for monthly data instead of daily data to avoid the 'big-N' problem [39], which could have significantly increased the computing time due to the rank of the variance-covariance matrix. Additionally, using monthly data allowed us to study the patterns of the mortality functions throughout each year.

The time component (consisting of 12 time points) and the city locations (six in total) represent the temporal and spatial effects, respectively. It is important to note that the population sizes of the six cities differ, potentially influencing the relative mortality rates among them. To account for this, we calculated the number of deaths per one hundred thousand people for each city.

In this paper, we applied the two spatio-temporal models to the South Korea data set to determine which model was more suitable for describing the data, based on the model selection criteria. For each model, we needed to simultaneously estimate the single-index function, model parameters, and spatial and time effects. The dependence range was estimated using the variograms of the models. In both models, the response variable was mortality (Y), and the explanatory variables included the temperature (x_1), pressure (x_2), and mean humidity (x_3).

Figure 2a–c depict the relationships between mortality and temperature (negative), humidity (negative), and pressure (positive). Figure 2d reveals that mortality was highest at the beginning and end of the year, with a minimum in the middle of each year. The highest mortality was observed in Busan, while Seoul, Daejeon, and Gwangju had the lowest mortality rates.

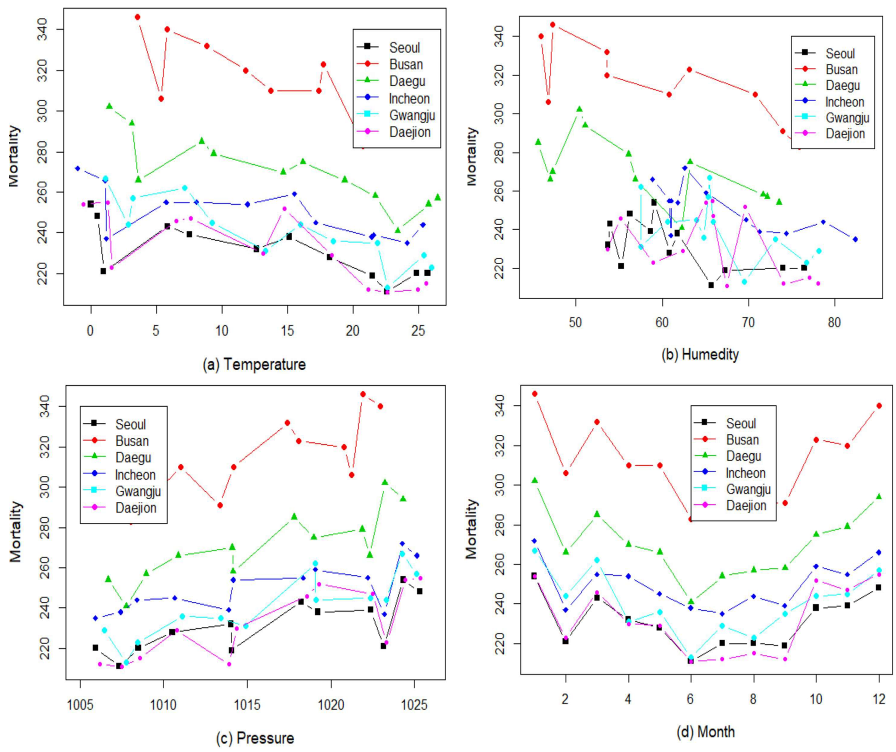

Figure 2. The relationship between temperature and mortality (**a**), the relationship between humidity and mortality (**b**), the relationship between pressure and mortality (**c**), and the relationship between month and mortality (**d**) for the six cities in South Korea.

In addition to applying the two models to the mortality data, we evaluated which model was most appropriate for the data based on fitting and prediction criteria. The following are the two proposed models for our motivating data:

- IST-SIM

$$\begin{aligned} \mathbf{y}|\boldsymbol{\mu} &\sim \mathrm{Pois}(\boldsymbol{\mu}|\mathbf{u},\boldsymbol{\omega}), \\ \boldsymbol{\mu} &= g(Z_1\mathbf{u} + \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{x}_3\beta_3) + Z_2\boldsymbol{\omega}, \\ \mathbf{u} &\sim MN(0,\sigma_u^2\Sigma), \text{ and } \boldsymbol{\omega} \sim MN(0,\sigma_\omega^2\Omega). \end{aligned}$$

- ST-SIM

$$\begin{aligned} \mathbf{y}|\boldsymbol{\mu} &\sim \mathrm{Pois}(\boldsymbol{\mu}|\mathbf{u},\boldsymbol{\omega}), \\ \boldsymbol{\mu} &= g(\mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{x}_3\beta_3) + Z_1\mathbf{u} + Z_2\boldsymbol{\omega}, \\ \mathbf{u} &\sim MN(0,\sigma_u^2\Sigma), \text{ and } \boldsymbol{\omega} \sim MN(0,\sigma_\omega^2\Omega). \end{aligned}$$

5.2. Models Estimation

The proposed algorithms were employed to estimate the parameters and spatial and temporal effects of both the ST-SIM and IST-SIM models. The Metropolis–Hastings (M-H) algorithm was utilized to obtain 10,000 samples from the spatial and temporal effects, with the first 2% of the MCMC samples discarded. The Markov chain Monte Carlo expectation maximization (MCEM) algorithm was executed until convergence was achieved. Using the variogram, the dependence range was estimated and found to be equal to 1.8, as illustrated in Figure 3.

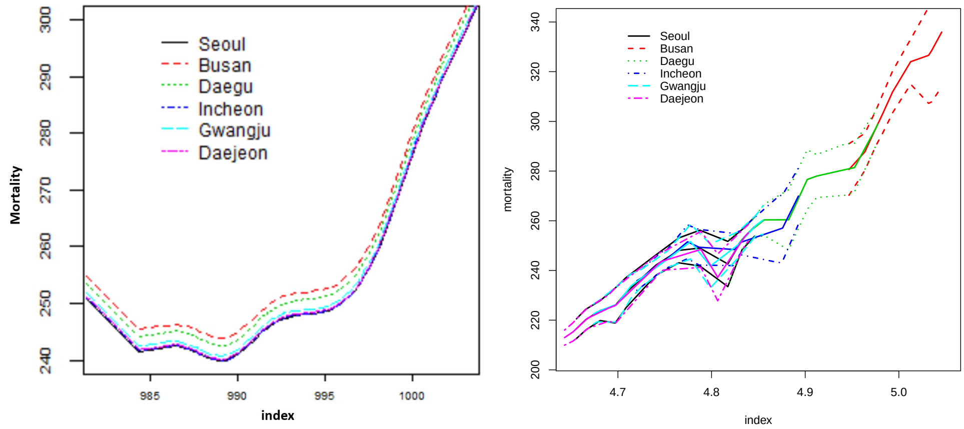

Figure 3. Smoothed mortality functions from the ST-SIM (**left**) and smoothed mortality functions from the IST-SIM along with 95% confidence intervals (**right**) of the six major cities in South Korea.

Table 3 presents the two model parameters along with their 95% confidence intervals and fitting criteria values, including the MSE, R^2, and log-likelihood. The IST-SIM outperformed the ST-SIM in several aspects. Specifically, the IST-SIM exhibited a smaller MSE (131.7 versus 131.7) compared to the ST-SIM, a higher R^2 (0.87 versus 0.41), and a superior log-likelihood value (-297.3 versus -363.2). In the ST-SIM, the coefficient parameter of the pressure variable, x_2, was set to 1 to address the identifiability problem. In contrast, the IST-SIM did not encounter this issue and successfully estimated all the parameters. Table 4 highlights that both models identified Busan as having the highest mortality rate and Seoul as having the lowest mortality rate.

Table 3. The ST-SIM = $g(Z_1\mathbf{u} + X\boldsymbol{\beta}) + Z_2\boldsymbol{\omega}$ and IST-SIM = $g(X\boldsymbol{\beta}) + Z_1\mathbf{u} + Z_2\boldsymbol{\omega}$ parameter estimates along with the 95% confidence intervals, R^2 values, and the log-likelihood values.

	ST-SIM		**IST-SIM**	
	Estimate	**95% CI**	**Estimate**	**95% CI**
x_1	-0.05	$(-0.049, -0.052)$	0.66	(0.60, 0.72)
x_2	0.06	(0.055, 0.066)	-	-
x_3	0.02	(0.011, 0.032)	-0.49	$(-0.41, -0.52)$
σ_u^2	2.13		0.97	
σ_ω^2	0.47		0.52	
log Likelihood	-297.3		-363.2	
R^2	0.87		0.41	
MSE	637.1		131.7	

Table 4. Spatial random effects estimates of the two introduced models [ST-SIM = $g(Z_1\mathbf{u} + X\boldsymbol{\beta}) + Z_2\boldsymbol{\omega}$ and IST-SIM = $g(X\boldsymbol{\beta}) + Z_1\mathbf{u} + Z_2\boldsymbol{\omega}$].

	ST-SIM	**IST-SIM**
Busan	1.254	2.377
Daegu	0.715	1.092
Gwangju	-0.422	-0.637
Daejeon	-0.697	-1.258
Incheon	-0.876	-1.479
Seoul	-0.905	-1.577

Figure 3 shows that the mortality functions of the six cities had the same form when using the ST-SIM but not with the IST-SIM. This highlights one of the advantages of the

IST-SIM over the ST-SIM: the former is more flexible, allowing mortality functions to vary by location.

5.3. Model Selection

Multiple criteria were employed to evaluate the fitting and prediction performance of both ST-SIM and IST-SIM. These included the MSE, R^2, and log-likelihood to assess the models' suitability for describing the data. Additionally, the means and medians of the predicted mean square error (PMSE) and predicted log-likelihood were utilized. These evaluation criteria were calculated using the following steps:

1. Select n observations randomly from each city, consider the selected observations as the evaluation data, and consider the remaining observations as the training data.
2. For the training data, fit the model and compute the R^2, MSE, and log-likelihood fits. For the evaluation data, calculate the log-likelihood prediction and PMSE $= \sum_{i=1}^{6n}(\mathbf{y}_i - \hat{\mathbf{y}}_i)^2/6n$, where $\mathbf{y}_i$ and $\hat{\mathbf{y}}_i$ are the actual and predicted response values, respectively.
3. Repeat Steps 1–2 for a large number of times; compute the mean and median of the estimated PMSE values and calculate the means of the R^2, MSE, log-likelihood predictions, and log-likelihood fits.

These steps were repeated for different values of n ($n = 2, 4,$ and 6) for the two proposed models. Table 5 demonstrates that the IST-SIM outperformed the ST-SIM consistently, exhibiting higher R^2, lower MSE, and higher log-likelihood at all values of n. For example, at $n = 2$, the IST-SIM had an R^2 of 0.88 compared to 0.45 for the ST-SIM, an MSE of 120.8 versus 600.3, and a log-likelihood of -248.6 versus -300.2.

Table 5. Predicted mean square error (PMSE), mean square error (MSE), log likelihood prediction and fits, and R^2 of ST-SIM = $g(Z_1\mathbf{u} + X\boldsymbol{\beta}) + Z_2\boldsymbol{\omega}$ and IST-SIM = $g(X\boldsymbol{\beta}) + Z_1\mathbf{u} + Z_2\boldsymbol{\omega}$ summary results of 500 estimates of the fitting and prediction criteria for ST-SIM and IST-SIM at different sizes of the evaluating data sets ($n = 2, 4, 6$).

		PMSE		**Log Likelihood**			
	Model	**Mean**	**Median**	**Prediction**	**Fits**	**MSE**	R^2
n = 2	ST-SIM	922.9	884.0	−75.6	−300.2	600.3	0.45
	IST-SIM	274.2	210.5	−63.3	−248.6	120.8	0.88
n = 4	ST-SIM	967.0	930.1	−235.3	−235.3	531.0	0.51
	IST-SIM	352.3	219.3	−134.3	−201.6	122.59	0.88
n = 6	ST-SIM	1094.8	985.7	−212.6	−174.7	467.3	0.56
	IST-SIM	633.4	250.2	−192.7	−153.8	117.9	0.89

In terms of prediction, once again, the IST-SIM outperformed the ST-SIM consistently, with smaller mean and median values for the PMSE and higher values for the log-likelihood prediction. For example, at $n = 2$, the mean PMSE was 274.2 versus 922.9, the median PMSE was 210.5 versus 884.0, and the log-likelihood prediction was -63.3 versus -75.6.

Figure 4 displays boxplots representing 500 estimates of the PMSE from each model at different data set sizes ($n = 2, 4, 6$). The boxplots indicate that the variability (interquartile range) of the PMSE estimates for the IST-SIM is lower than that of the ST-SIM at $n = 2$ and $n = 4$. However, when $n = 6$, with half of the data used for training and the other half for evaluation, the PMSE estimates for the IST-SIM exhibit greater variability than those for the ST-SIM. Nevertheless, the median PMSE for the IST-SIM is lower than that for the ST-SIM. Overall, the IST-SIM outperforms the ST-SIM in terms of describing the mortality data and making predictions from the South Korean mortality data. Upon receiving a reasonable request, we are able to furnish the code utilized to analyze the real data application.

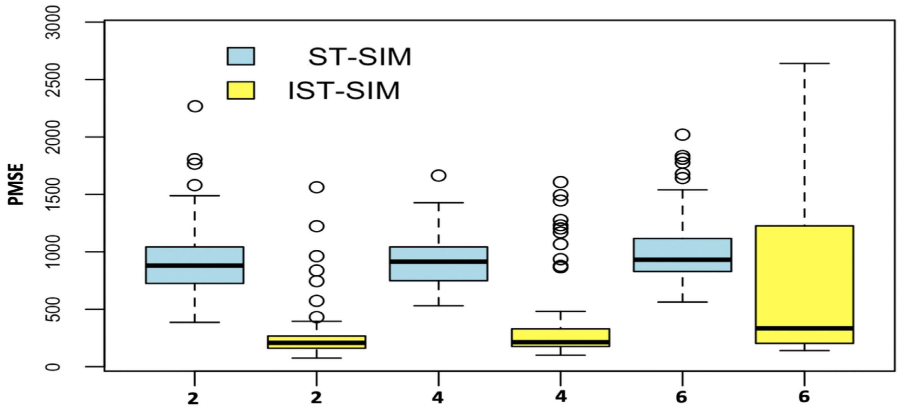

Figure 4. Boxplots of the prediction mean square error (PMSE) of the proposed two models (ST-SIM and IST-SIM) at different evaluation data set sizes ($n = 2, 4, 6$).

6. Conclusions

This article introduces two semiparametric models for modeling correlated spatio-temporal data. In the ST-SIM, the spatial effect is treated as an additive component to the single-index model, while in the IST-SIM, the spatial effect is integrated into the single-index function. For each model, we proposed an algorithm to simultaneously estimate the unknown function, single-index coefficient parameters, variance of the spatial effects, spatial effects, and time effects, using an MCEM algorithm that we developed. To the best of our knowledge, the IST-SIM has not yet been documented in the statistical literature.

We conducted several simulation studies and found that the IST-SIM outperformed the ST-SIM based on various criteria. Both models were also applied to South Korean mortality data for comparison. The IST-SIM demonstrated superior performance in fitting and prediction for the motivating data. The analysis revealed that Busan had the highest nonaccidental mortality among the six major cities, while Seoul had the lowest mortality, with the other cities falling in between.

The IST-SIM offers several advantages over the ST-SIM: (1) it does not necessitate restrictions on the coefficient parameters like the ST-SIM, allowing estimation of all parameters, and (2) it does not require mortality functions to have the same form across each location, a requirement of the ST-SIM. There exist several avenues for enhancing the IST-SIM model in future research. For instance, potential improvements may encompass: (1) incorporating temporal effects into the unknown nonparametric function, (2) exploring alternative estimation methods for the nonparametric function beyond the Ichimura method, and (3) exploring the utilization of the Bayesian approach for the model estimation, which has the potential to streamline parameter estimation and expedite the process in comparison to the EM algorithm.

Author Contributions: Conceptualization, H.F.F.M. and I.K.; methodology, H.F.F.M. and I.K.; software, H.F.F.M.; validation, H.F.F.M.; formal analysis, H.F.F.M.; writing—original draft preparation, H.F.F.M.; writing—review and editing, H.F.F.M. and I.K.; visualization, H.F.F.M.; supervision, I.K.; project administration, H.F.F.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and computing code are available for replication from the corresponding author upon reasonable request.

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

Abbreviations

The following abbreviations are used in this manuscript:

SIM	Single-index model
ST-SIM	Spatio-temporal single-index model
IST-SIM	Integrated spatio-temporal single-index model
MCEM	Monte Carlo expectation maximization
EM	Expectation maximization
MCMC	Markov chain Monte Carlo
M-H	Metropolis–Hastings
MSE	Mean square error
PRESS	Predicted residual sum of squares
SE	Standard error
PMSE	Predicted mean square error

References

1. Cressie, N.A.C. *Statistics for Spatial Data*; Wiley: New York, NY, USA, 1993.
2. Sherman, R.P. U-process in analysis of a generalized semi-parametric regression estimator. *Econ. Theory* **1994**, *10*, 372–395. [CrossRef]
3. Cressie, N.; Huang, H.C. Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Am. Stat. Assoc.* **1999**, *94*, 1330–1340. [CrossRef]
4. Kanevski, M.; Maignan, M. *Analysis and Modeling of Spatial Environmental Data*; EPFL Press: Lausanne, Switzerland, 2004.
5. Genton, M.G.; Butry, D.T.; Gumpertz, M.L.; Prestemon, J.P. Spatio-temporal analysis of wildfire ignitions in the St. Johns River Water Management District, Florida. *Int. J. Wildland Fire* **2006**, *15*, 87–97. [CrossRef]
6. Landagan, E.B.; Barrios, O.Z. An estimation procedure for a spatial-temporal model. *Stat. Probab. Lett.* **2007**, *77*, 401–406. [CrossRef]
7. Li, B.; Genton, M.G.; Sherman, M. A nonparametric assessment of properties of space-time covariance functions. *J. Am. Stat. Assoc.* **2007**, *102*, 736–744. [CrossRef]
8. Nelson, T.A.; Duffus, D.; Robertson, C.; Laberfee, K.; Feyrer, L.J. Spatial-temporal analysis of marine wildlife. *J. Coast. Res.* **2009**, *56*, 1537–1541.
9. Hayn, M.; Beirle, S.; Hamprecht, F.A.; Platt, U.; Menze, B.H.; Wagner, T. Analysing spatio-temporal patterns of the global NO2-distribution retrieved from GOME satellite observations using a generalized additive model. *Atmos. Chem. Phys.* **2009**, *9*, 6459–6477. [CrossRef]
10. Sherman, M. *Spatial Statistics and Spatio-Temporal Data*; Wiley: New York, NY, USA, 2011.
11. Arcuti, S.; Calculli, C.; Pollice, A.; D'Onghia, G.; Maiorano, P.; Tursi, A. Spatio-temporal modeling of zero-inflated deep-sea shrimp data by Tweedie generalized additive. *Statistica* **2013**, *73*, 103–122.
12. Lekdee, K.; Ingsrisawang, L. Generalized linear mixed models with spatial random effects for spatio-temporal data: An application to dengue fever mapping. *J. Math. Stat.* **2013**, *9*, 137–143. [CrossRef]
13. Barzegar, Z.; Rivaz, F. A scalable Bayesian nonparametric model for large spatio-temporal data. *Comput. Stat.* **2021**, *35*, 153–173. [CrossRef]
14. Harper, A.; Baker, P.N.; Xia, Y.; Kuang, T.; Zhang, H.; Chen, Y.; Han, T.-L.; Gulliver, J. Development of spatiotemporal land use regression models for PM2.5 and NO2 in Chongqing, China, and exposure assessment for the CLIMB study. *Atmos. Pollut. Res.* **2021**, *12*, 101096. [CrossRef]
15. Ibañez, M.V.; Martínez-Garcia, M.; Simó, A. A Review of Spatiotemporal Models for Count Data in R Packages. A Case Study of COVID-19 Data. *Mathematics* **2021**, *9*, 1538. [CrossRef]
16. Feng, C. Spatial-temporal generalized additive model for modeling COVID-19 mortality risk in Toronto, Canada. *Stat. Sci.* **2022**, *49*, 100526. [CrossRef]
17. Cressie, N.; Hawkins, D.M. Robust estimation of the variogram: I. *J. Int. Assoc. Math. Geol.* **1980**, *12*, 115–125. [CrossRef]
18. Ichimura, H. Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *J. Econom.* **1993**, *58*, 71–120. [CrossRef]
19. Hridtache, M.; Juditski, A.; Spokoiny, V. Direct estimation of the single coefficients in a single-index model. *Ann. Stat.* **2001**, *29*, 595–623.
20. Wang, J.L.; Xue, L.G.; Zhu, L.X.; Chong, Y.S. Extension for a partial-linear single-index model. *Ann. Stat.* **2010** *38*, 246–274.
21. Chang, Z.Q.; Xue, L.G.; Zhu, L.X. On asymptotically more efficient estimation of the single-index model. *J. Multivar. Anal.* **2010**, *101*, 1898–1901. [CrossRef]
22. Mahmoud, H.F.; Kim, I.; Kim, H. Semiparametric single index multi change points model with an application of environmental health study on mortality and temperature. *Environmetrics* **2016**, *27*, 49–506. [CrossRef]
23. Mahmoud, H.F.; Kim, I. Semiparametric spatial mixed effects single index models. *Comput. Stat. Data Anal.* **2019**, *136*, 108–122. [CrossRef]

24. Horowitz, J.L.; Hardle, W. Direct semiparametric estimation of single-index models with discrete covariates. *J. Am. Stat. Assoc.* **1996**, *91*, 1623–1629. [CrossRef]
25. Xia, Y.; Tong, H.; Li, W. K.; Zhu, L.X. An adaptive estimation of dimension reduction space. *J. R. Stat. Soc. Ser. B* **2002**, *64*, 363–410. [CrossRef]
26. Lin, W.; Kulasekera, K.B. Identifiability of single index models and additive index models. *Biometrika* **2007**, *94*, 496–501. [CrossRef]
27. Pang, Z.; Xue, L. Estimation of the single-index models with random effects. *Comput. Stat. Data Anal.* **2012**, *56*, 1837–1853. [CrossRef]
28. Hardle, W.; Hall, P.; Ichimura, H. Optimal smoothing in single-index models. *Ann. Stat.* **1993**, *21*, 157–178. [CrossRef]
29. Knorr-Held, L. Bayesian modeling of inseparable space-time variation in disease risk. *Stat. Med.* **2000**, *19*, 2555–2567. [CrossRef]
30. Liu, H.; Davidson, R.A.; Apanasovich, T.V. Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms. *Reliab. Eng. Syst. Saf.* **2008**, *93*, 875–890. [CrossRef]
31. McCulloch, C.E. Maximum likelihood variance components estimation for binary data. *J. Am. Stat. Assoc.* **1994**, *89*, 330–335. [CrossRef]
32. Booth, J.G.; Hobert, J.P. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo em algorithm. *J. R. Stat. Soc. Ser. B* **1999**, *61*, 265–285. [CrossRef]
33. Caffo, B.S.; Jank, W.; Jones, G.L. Ascent-based Monte Carlo expectation maximization. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 235–251. [CrossRef]
34. Tan, M.; Tian, G.-L.; Fang, H.-B. An efficient MCEM algorithm for fitting generalized linear mixed models for correlated binary data. *J. Stat. Comput. Simul.* **2007**, *77*, 929–943. [CrossRef]
35. An, X.; Bentler, P.M. Efficient direct sampling MCEM algorithm for latent variable models with binary responses. *Comput. Stat. Data Anal.* **2012**, *56*, 231–244. [CrossRef]
36. Geyer, C.J. Practical Markov chain Monte Carlo. *Stat. Sci.* **1922**, *7*, 473–483. [CrossRef]
37. Ruppert, D.; Wand, M.P.; Carroll, R.J. *Semiparametric Regression*; Cambridge Press: New York, NY, USA, 2003.
38. Wand, M.P.; Jones, M.C. *Kernel Smoothing*; Chapman and Hall: London, UK, 1995.
39. Banerjee, S.; Carlin, C. P.; Gelfand, A.E. *Hierarchical Modeling and Analysis for Spatial*; Chapman and Hall: London, UK, 2004.