

Article

Bayesian Estimation of the Semiparametric Spatial Lag Model

Kunming Li ¹ and Liting Fang ^{2,*}¹ College of Economics and Management, Fujian Agriculture and Forestry University, Fuzhou 350002, China; likm@fafu.edu.cn² School of Economics and Management, Fuzhou University, Fuzhou 350108, China

* Correspondence: fzugyflt@fzu.edu.cn

Abstract: This paper proposes a semiparametric spatial lag model and develops a Bayesian estimation method for this model. In the estimation of the model, the paper combines Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm, random walk Metropolis sampler, and Gibbs sampling techniques to sample all the parameters. The paper conducts numerical simulations to validate the proposed Bayesian estimation theory using a numerical example. The simulation results demonstrate satisfactory estimation performance of the parameter part and the fitting performance of the nonparametric function under different spatial weight matrix settings. Furthermore, the paper applies the constructed model and its estimation method to an empirical study on the relationship between economic growth and carbon emissions in China, illustrating the practical application value of the theoretical results.

Keywords: semiparametric spatial lag model; Bayesian estimation; polynomial spline; RJMCMC

MSC: 62G05; 62G08; 62F15



Citation: Li, K.; Fang, L. Bayesian Estimation of the Semiparametric Spatial Lag Model. *Mathematics* **2024**, *12*, 2289. <https://doi.org/10.3390/math12142289>

Academic Editor: Manuel Alberto M. Ferreira

Received: 7 June 2024

Revised: 4 July 2024

Accepted: 20 July 2024

Published: 22 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Econometric models can be divided into parametric regression models and nonparametric regression models based on the structural aspect of the regression object. Parametric models assume that the regression object follows a known parameter structure, while nonparametric models do not make specific structural assumptions about the regression object but instead rely on mild smoothness conditions to effectively avoid the risk of misspecification. However, in practical applications, researchers often have limited knowledge about the exact functional form of the impact of explanatory variables on the dependent variable. In such cases, combining both models into a single model, known as a semiparametric regression model, is necessary to address real-world problems effectively. Therefore, the semiparametric regression model can be seen as a bridge between parametric and nonparametric regression models, allowing for the utilization of information contained in the data. It establishes a parametric model for the part of the explanatory variables that have a clear relationship with the response variable and a nonparametric model for the parts with less explicit relationships (such as spatial factors and temporal factors). The advantage of this model lies in its ability to combine the interpretability of linear regression with the robustness of nonparametric regression, while addressing the “curse of dimensionality” issue associated with nonparametric regression, making it easier to estimate. Additionally, as the model contains both parametric and nonparametric components, it provides a closer approximation to reality and better utilizes the information provided by the data when describing real-world problems. Therefore, the semiparametric model is practical in both theoretical research and practical applications, and it holds significant importance.

Due to the rapid development of spatial econometrics in recent years, spatial econometric models have gained widespread attention and extensive application in various fields (see [1–3]). Particularly in the fields of economics, management, and environmental science,

variables are not only influenced by regional factors (exhibiting spatial correlation) but also often exhibit both linear and nonlinear relationships. Therefore, this paper considers incorporating spatial factor and spatial weight matrix into the semiparametric model, expanding it into a spatial semiparametric model.

Given the practicality of the spatial semiparametric model, there have been increasing applications and theoretical advancements in this area in recent years. Su [4] applied the quasi-maximum likelihood method to estimate partially linear spatial autoregressive model, and subsequently, Su [5] further discussed the estimation issues of the semiparametric spatial autoregressive model. Malikov and Sun [6] discussed the estimation and testing problems of the smooth coefficient semiparametric spatial autoregressive model, and later, Malikov et al. [7] applied the semiparametric spatial quantile autoregressive model to the study of residential housing values. Li and Chen [8] employed the maximum likelihood method to estimate the semiparametric varying coefficient spatial lag model. Sun [9] discussed the estimation and application of the semiparametric spatial lag model for panel data.

From the existing research results, it can be observed that in terms of theory, the main research interest in the semiparametric spatial model lies in the estimation of the coefficients in the parametric part and the unknown functions in the nonparametric part. Since this model contains both parametric and nonparametric components, the methods used to study this model combine the approaches used in linear regression and nonparametric regression, but it is not a simple combination of the two methods. The existing literature mainly employs traditional frequentist methods such as (quasi-) maximum likelihood estimation and generalized method of moments to estimate the semiparametric spatial model. In comparison to traditional methods, the Bayesian estimation method has the distinctive feature of utilizing information. It not only uses sample information but also incorporate prior information (even though acquiring precise prior information can be challenging), whereas traditional methods rely solely on sample information for estimation. Additionally, the Bayesian estimation method can obtain consistent and efficient estimates under more relaxed conditions. Particularly, the application of Bayesian methods, especially simulation-based Bayesian inference in the semiparametric model can effectively alleviate the “curse of dimensionality” issue. Therefore, in recent years, the Bayesian method has been successfully applied in semiparametric regression and spatial models. Gelfand et al. [10] applied Bayesian methods to discuss the estimation of the nonparametric spatial model. Han and Lee [11] studied the Bayesian estimation and model selection problems of the spatial durbin error model. Kang and Cressie [12] investigated Bayesian inference for the spatial random effects model. Fang and Qian [13] and Fang [14] applied Bayesian methods to estimate the nonparametric spatial lag model and general spatial lag model respectively. Lesage [15] used Bayesian method to study statistical inference problems in the panel data spatial econometric model. Louzada et al. [16] focused on the progressive development and content analysis of the Bayesian spatial models. Thapa et al. [17] employed Bayesian statistics using the nested-sampling algorithm to compare and rank multiple models of ergodic diffusion (including anomalous diffusion) as well as to assess their optimal parameters for in silico-generated and real time-series.

Due to the advantages of the Bayesian method, this paper will explore the Bayesian estimation problem of the spatial semiparametric model. Depending on the treatment of spatial dependence, spatial econometric models can be divided into spatial lag model and spatial error model. However, the spatial error model can essentially be viewed as a traditional econometric model with a special heteroscedastic structure, and it can be transformed into a generalized form of the spatial lag model known as the spatial durbin model. Therefore, in a sense, the spatial lag model is a more general spatial model. Hence, the following content of this paper will apply the Bayesian method to estimate the semiparametric spatial lag model.

The rest of this paper is organized as follows. Section 2 presents a model analysis. Section 3 develops prior distributions. Section 4 describes posterior distributions. Section 5

introduces the sampling process. Section 6 provides numerical simulation. Section 7 shows an application example. We conclude the paper in Section 7 with a conclusion.

2. Model Analysis

For a cross-sectional semiparametric spatial lag model with n spatial units:

$$\mathbf{Y} = \rho \mathbf{WY} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{H} + \boldsymbol{\varepsilon} \quad (1)$$

where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, $\mathbf{Y} = (y_1, y_2, \dots, y_n)'$, $\mathbf{H} = (h(x_1), h(x_2), \dots, h(x_n))'$, $h(x_i)$ are unknown link functions, $\mathbf{Z}$ is an $n \times l$ matrix, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_l)'$, and $\mathbf{W}$ represents the spatial weight matrix. To estimate model (1), it is necessary to estimate the spatial lag coefficient ρ , the parameter coefficient vector $\boldsymbol{\beta}$, the variance of the random disturbance term σ^2 , and the unknown nonparametric function $h(x_i)$, $i = 1, 2, \dots, n$.

Regarding the estimation of the nonparametric function $h(x_i)$, this paper adopts a m spline method, which involves a linear combination of given spline bases and spline coefficients. The spline function estimation of the unknown function $h(x_i)$ is given by:

$$\hat{h}(x_i) = \sum_{j=1}^{1+m+k} B_j(x_i) \hat{\gamma}_j = \mathbf{B}(x_i)' \cdot \boldsymbol{\gamma} \quad (2)$$

where $\mathbf{B}(x_i) = (B_1(x_i), B_2(x_i), \dots, B_{1+m+k}(x_i))'$ is a given set of spline bases, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_{1+m+k})'$ is the spline coefficient, and $\mathbf{r} = (r_1, r_2, \dots, r_k)'$ are the k spline nodes with a, b as the two boundary nodes. Therefore, to estimate the unknown function $h(x_i)$, it is necessary to estimate the number of spline nodes k , the positions of the nodes $\mathbf{r} = (r_1, r_2, \dots, r_k)'$, and the spline coefficients $\boldsymbol{\gamma}$. Finally, the unknown function $h(x_i)$ is estimated using the average value of the Markov chain $E(h(x_i)|k, \mathbf{r}, \mathbf{Y})$, which is denoted as $\hat{h}(x_i) = E(h(x_i)|\mathbf{Y}) = E(E(h(x_i)|k, \mathbf{r}, \mathbf{Y}))$.

Based on the above analysis, using matrix equations, model (1) can be written in the following form:

$$\mathbf{Y} = \rho \mathbf{WY} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{X}_r \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (3)$$

where

$$\mathbf{X}_r = \begin{pmatrix} B_1(x_1) & B_2(x_1) & \cdots & B_{1+m+k}(x_1) \\ B_1(x_2) & B_2(x_2) & \cdots & B_{1+m+k}(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ B_1(x_n) & B_2(x_n) & \cdots & B_{1+m+k}(x_n) \end{pmatrix}_{n \times (1+m+k)}$$

If we let $\mathbf{X} = (\mathbf{Z}, \mathbf{X}_r)$, $\boldsymbol{\zeta} = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}$, then Equation (3) can be further written as:

$$\mathbf{Y} = \rho \mathbf{WY} + \mathbf{X}\boldsymbol{\zeta} + \boldsymbol{\varepsilon} \quad (4)$$

Therefore, assuming that $\boldsymbol{\varepsilon}$ follows a normal distribution, given the number and positions of the spline nodes k and $\mathbf{r}$, the likelihood function of model (4) is as follows:

$$\begin{aligned} L(\mathbf{Y}|k, \mathbf{r}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \rho) \\ \propto (2\pi\sigma^2)^{-n/2} |\mathbf{A}| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{AY} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{X}_r \boldsymbol{\gamma})^T (\mathbf{AY} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{X}_r \boldsymbol{\gamma}) \right\} \end{aligned} \quad (5)$$

where $\mathbf{I}_n$ is the n matrix and $\mathbf{A} = (\mathbf{I}_n - \rho \mathbf{W})$.

3. Prior Distributions

Since the number k and positions $\mathbf{r}$ of the spline nodes are both random variables and unknown, they also need to be estimated. First, the prior distributions of the number k and positions $\mathbf{r}$ of the spline nodes are set as follows:

$$\pi(k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \dots \quad (6)$$

$$\pi(\mathbf{r}|k) = \frac{k!}{(b-a)^k} \Delta_k \quad (7)$$

where $\Delta_k = I\{a < r_1 < r_2 < \dots < r_k < b\}$ is the indicator function.

Second, the number of spline nodes k , the prior distributions of the parameter vector ζ , and the variance σ^2 are set as multivariate normal distribution and inverse gamma distribution, respectively:

$$\sigma^2|k \sim IG\left(\frac{v_1}{2}, \frac{v_2}{2}\right) \quad (8)$$

$$\zeta|\sigma^2, k \sim N_{1+l+m+k}(\mathbf{0}, \sigma^2 \Sigma_0) \quad (9)$$

That is,

$$\pi(\sigma^2|k) = \frac{1}{\Gamma(\frac{v_1}{2})} \left(\frac{v_2}{2}\right)^{v_1/2} (\sigma^2)^{-(\frac{v_1}{2}+1)} e^{-\frac{v_2}{2\sigma^2}} \quad (10)$$

$$\pi(\zeta|k, \sigma^2) = (2\pi\sigma^2)^{-\frac{1+l+m+k}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} \zeta^T (\Sigma_0)^{-1} \zeta\right\} \quad (11)$$

where

$$\Sigma_0 = \text{diag}(\mathbf{I}_l, \mathbf{V}_{1+m+k}) \quad (12)$$

Finally, the prior distribution of the spatial lag coefficient ρ is set as a uniform distribution:

$$\rho \sim U\left(\lambda_{\min}^{-1}, \lambda_{\max}^{-1}\right)$$

That is,

$$\pi(\rho) = \frac{1}{\lambda_{\max}^{-1} - \lambda_{\min}^{-1}} \quad (13)$$

where $\lambda_{\min}^{-1}$ and $\lambda_{\max}^{-1}$ are the minimum and maximum eigenvalues of the spatial weight matrix $\mathbf{W}$.

4. Posterior Distributions

First, we combine the prior distribution of all parameters with the likelihood function to obtain the joint distribution of all variables:

$$\begin{aligned} P(k, \mathbf{r}, \zeta, \sigma^2, \rho, \mathbf{Y}) \\ = L(\mathbf{Y}|k, \mathbf{r}, \zeta, \sigma^2, \rho) \cdot \pi(k) \cdot \pi(\mathbf{r}|k) \cdot \pi(\sigma^2|k) \cdot \pi(\zeta|k, \sigma^2) \cdot \pi(\rho) \end{aligned} \quad (14)$$

By Bayes' theorem, the joint posterior density function of all parameters $(k, \mathbf{r}, \zeta, \sigma^2, \rho)$ can be obtained according to the above formula:

$$P(k, \mathbf{r}, \zeta, \sigma^2, \rho|\mathbf{Y}) = \frac{P(k, \mathbf{r}, \zeta, \sigma^2, \rho, \mathbf{Y})}{\iint \iiint P(k, \mathbf{r}, \zeta, \sigma^2, \rho, \mathbf{Y}) d\zeta d\sigma^2 d\rho dk d\mathbf{r}} \quad (15)$$

The denominator $\iint \iiint P(k, \mathbf{r}, \zeta, \sigma^2, \rho, \mathbf{Y}) d\zeta d\sigma^2 d\rho dk d\mathbf{r}$ in the above formula is independent of the parameters $(k, \mathbf{r}, \zeta, \sigma^2, \rho)$, so in application, only the numerator is considered,

and it is called the kernel of the posterior density. It can be seen in the following formula, where the symbol “ $\propto$ ” means proportional to.

$$P(k, \mathbf{r}, \zeta, \sigma^2, \rho | \mathbf{Y}) \propto P(k, \mathbf{r}, \zeta, \sigma^2, \rho, \mathbf{Y}) \quad (16)$$

Then, the prior distribution density function and likelihood function of all variables are substituted into the above formula to obtain the joint posterior distribution of all variables. It is shown in the following formula:

$$\begin{aligned} & P(k, \mathbf{r}, \zeta, \sigma^2, \rho | \mathbf{Y}) \\ & \propto (2\pi\sigma^2)^{-n/2} |\mathbf{A}| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{A}\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{X}_r\boldsymbol{\gamma})^T (\mathbf{A}\mathbf{Y} - \mathbf{Z}\boldsymbol{\beta} - \mathbf{X}_r\boldsymbol{\gamma}) \right\} \\ & \times \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{k!}{(b-a)^k} \Delta_k \cdot \frac{1}{\Gamma(\frac{v_1}{2})} \left(\frac{v_2}{2}\right)^{v_1/2} (\sigma^2)^{-(\frac{v_1}{2}+1)} e^{-\frac{v_2}{2\sigma^2}} \\ & \times (2\pi\sigma^2)^{-\frac{1+l+m+k}{2}} |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \zeta^T (\boldsymbol{\Sigma}_0)^{-1} \zeta \right\} \cdot \frac{1}{\lambda_{\max}^{-1} - \lambda_{\min}^{-1}} \\ & \propto \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{k!}{(b-a)^k} \Delta_k \cdot (2\pi)^{-\frac{1+l+m+k}{2}} \cdot (\sigma^2)^{-(n+v_1+l+m+k+3)/2} \cdot |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \cdot \frac{1}{\lambda_{\max}^{-1} - \lambda_{\min}^{-1}} \\ & \times \exp \left\{ -\frac{1}{2\sigma^2} \left[(\mathbf{A}\mathbf{Y} - \mathbf{X}\zeta)^T (\mathbf{A}\mathbf{Y} - \mathbf{X}\zeta) + \zeta^T (\boldsymbol{\Sigma}_0)^{-1} \zeta + v_2 \right] \right\} \cdot |\mathbf{A}| \\ & \propto \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{k!}{(b-a)^k} \Delta_k \cdot S^{-\frac{n+v_1}{2}} \cdot |\boldsymbol{\Sigma}|^{\frac{1}{2}} \cdot |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \\ & \times S^{\frac{n+v_1}{2}} \cdot (\sigma^2)^{-(n+v_1+2)/2} \cdot \exp \left\{ -\frac{S}{2\sigma^2} \right\} \\ & \times (2\pi\sigma^2)^{-\frac{1+l+m+k}{2}} \cdot |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \left[\left(\zeta - \bar{\zeta}_r \right)^T (\boldsymbol{\Sigma})^{-1} \left(\zeta - \bar{\zeta}_r \right) \right] \right\} \\ & \times |\mathbf{A}| \cdot \frac{1}{\lambda_{\max}^{-1} - \lambda_{\min}^{-1}} \end{aligned} \quad (17)$$

where $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_0^{-1} + \mathbf{X}^T \mathbf{X})^{-1}$, $S = v_2 + \mathbf{Y}^T \mathbf{A}^T \mathbf{A} \mathbf{Y} - \mathbf{Y}^T \mathbf{A}^T \mathbf{X} (\boldsymbol{\Sigma}_0^{-1} + \mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{A} \mathbf{Y}$, $\bar{\zeta}_r = (\boldsymbol{\Sigma}_0^{-1} + \mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{A} \mathbf{Y}$, and note that $\mathbf{X} = (\mathbf{Z}, \mathbf{X}_r)$, $\zeta = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}$, $\boldsymbol{\Sigma}_0 = \text{diag}(\mathbf{I}_l, \mathbf{V}_{1+m+k})$.

Furthermore, based on Equation (14), the conditional posterior distributions of each parameter can be obtained:

$$\begin{aligned} & P(k, \mathbf{r} | \zeta, \sigma^2, \rho, \mathbf{Y}) \\ & \propto \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{k!}{(b-a)^k} \Delta_k \cdot S^{-\frac{n+v_1}{2}} \cdot |\boldsymbol{\Sigma}|^{\frac{1}{2}} \cdot |\boldsymbol{\Sigma}_0|^{-\frac{1}{2}} \end{aligned} \quad (18)$$

$$P(\sigma^2 | k, \mathbf{r}, \zeta, \rho, \mathbf{Y}) \propto IG \left(\frac{n+v_1}{2}, \frac{S}{2} \right) \quad (19)$$

$$P(\zeta | k, \mathbf{r}, \sigma^2, \rho, \mathbf{Y}) \propto N_{1+l+m+k} \left(\bar{\zeta}_r, \sigma^2 \boldsymbol{\Sigma} \right) \quad (20)$$

$$\begin{aligned} & P(\rho | k, \mathbf{r}, \zeta, \sigma^2, \mathbf{Y}) \\ & \propto |\mathbf{A}(\rho)| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{A}(\rho)\mathbf{Y} - \mathbf{X}\zeta)^T (\mathbf{A}(\rho)\mathbf{Y} - \mathbf{X}\zeta) \right\} \frac{1}{\lambda_{\max}^{-1} - \lambda_{\min}^{-1}} \end{aligned} \quad (21)$$

where $\mathbf{A}(\rho) = (\mathbf{I}_n - \rho \mathbf{W})$.

5. Sampling

We first make a description of the Gibbs sampling method and the Metropolis–Hastings algorithm as well as the RJMCMC algorithm. The details are as follows.

Markov Monte Carlo simulation involves simulating a Markov process on a set of parameters, which converges to a stationary distribution, and this method to obtain a stationary distribution is called the MCMC method. Therefore, the basic idea of the MCMC method is to obtain samples of $\pi(x)$ by establishing a Markov chain of stationary distribution $\pi(x)$, based on which various statistical inferences can be made.

There are two popular MCMC methods in Bayesian analysis: Gibbs sampling (sampler) method and Metropolis–Hastings (MH) algorithm.

Gibbs sampling is the simplest and most widely used MCMC method. It was proposed by Geman in 1984. It was initially used for the analysis of large and complex data such as image processing analysis, artificial intelligence, and neural networks, and then introduced into the Bayesian model research by Gelfand and Smith [18]. This has a profound impact on the practical application of Bayesian methods. The success of Gibbs sampling is that it uses a fully conditional distribution to reduce a complex problem with multiple related parameters to a simpler problem with only one parameter at a time. Therefore, the most essential function of Gibbs sampling is to generate random numbers of joint posterior density, so it is the key of the Gibbs sampler to derive all conditional posterior density.

Another MCMC method that is more general than Gibbs sampling is the Metropolis–Hastings method. Metropolis et al. [19] proposed a method for transferring kernels, and Hastings [20] later extended it to form the Metropolis–Hastings method.

The core of the MH method is to find a suitable transfer probability density function, and then construct a nonperiodic irreducible normal return Markov chain, so that the stationary distribution of the Markov chain is the joint posterior distribution we need. The purpose of approximate sampling from joint posterior distribution is realized by using the property that limits distribution as stationary distribution.

In fact, the Metropolis–Hastings sampling method is convenient for numerical operations in lower dimensions. However, at higher dimensions, it is not easy to select the appropriate proposed distribution, whereas Gibbs sampling only needs to know the full conditional distribution, so it has an advantage for integration in higher dimensions. Gibbs sampling is essentially a special case of Metropolis–Hastings sampling when the acceptance rate is all 1. In addition, it should be noted that if all conditional distributions are easy to sample directly, the Gibbs sampling method is the most suitable, because there is no need to subjectively select the proposed distribution, and the acceptance rate is 100%. If the fully conditional distribution is not easy to sample directly, the Metropolis–Hastings algorithm can be used. It is more common to use the Gibbs sampler and the Metropolis–Hastings algorithm in a fully conditional distribution where one part is easy to sample directly and the other part is not.

The traditional MCMC sampling algorithm introduced above is usually carried out on the basis of the same probability density dimension of parameter joint distribution, while the RJMCMC (Reversible Jump Markov Chain Monte Carlo) sampling algorithm is a complex MCMC algorithm. It is extended based on the Metropolis–Hastings algorithm to allow Markov chains to jump in variable-dimensional parameter space models. This method is often applied to Bayes models and variable selection problems where parameter spaces may not be consistent.

Therefore, according to the specific characteristics of all parameters, this paper combined the RJMCMC algorithm, random walk Metropolis sampler, and Gibbs sampling technology to design the entire sampling process.

Regarding the estimation of parameters ζ and σ^2 , we use the Gibbs sampling, and for the estimation of ρ , the Metropolis–Hastings algorithm with random walk is employed. For the sampling of $(k, \mathbf{r})$, Reversible Jump Markov Chain Monte Carlo (MCMC) method is used, which constructs a mixture sampler by removing a node, adding a node, and moving a node. The following section will introduce the updating process of each parameter.

5.1. Updating the Number k and Positions $\mathbf{r}$ of Spline Nodes

Assuming that the positions of the current k spline nodes are $\mathbf{r} = (r_1, r_2, \dots, r_k)'$, a node is randomly selected from the current nodes to be deleted, added, or to change its position with probability $d_k, b_k, m_k (m_k = 1 - d_k - b_k)$, respectively. Let $(k', \mathbf{r}')$ denote the proposed number and positions of the nodes after the change.

When adding a node, we first randomly select a sub-interval (r_{l-1}, r_l) , and then randomly generate a value (e.g., ζ) from the distribution function $\text{Beta}(\alpha, \alpha)$ on the interval

(r_{l-1}, r_l) as the position of the new node. Because the location of the new node has a specific interval range, the proposed distribution of the location of the new node is generally set to uniform distribution or Beta distribution $Beta(\alpha, \alpha)$. In particular, when $\alpha = 1$, the Beta distribution becomes uniform, i.e., $Beta(1, 1) = U(0, 1)$. Therefore, this paper sets the proposed distribution for the location of the new node as Beta distribution.

If a node is to be removed, we randomly select one from the existing k nodes (e.g., r_l , $l = 1, 2, \dots, k$), and delete it.

When moving a node, we randomly select one (e.g., r_l) from the current nodes $r_1, r_2, \dots, r_k$, change its position, and the new position r'_l is generated from the Beta distribution function $Beta(\alpha, \alpha)$ on the interval (r_{l-1}, r_{l+1}) .

Therefore, the total acceptance rate from state $(k, \mathbf{r})$ to state $(k', \mathbf{r}')$ is:

$$Ratio = \min \left\{ 1, R_k \cdot \left(\frac{S'}{S} \right)^{-\frac{n+v_1}{2}} \cdot \left(\frac{|\Sigma'|}{|\Sigma|} \right)^{\frac{1}{2}} \cdot \left(\frac{|\Sigma'_0|}{|\Sigma_0|} \right)^{-\frac{1}{2}} \right\} \quad (22)$$

where,

$$R_k = \begin{cases} \frac{b_{k-1}}{d_k} \cdot \frac{(b-a)}{\lambda Beta(\alpha, \alpha)} \cdot \frac{(r_l - r_{l-1})^{\alpha-1} (r_{l+1} - r_l)^{\alpha-1}}{(r_{l+1} - r_{l-1})^{2\alpha-1}}, & k' = k - 1 \\ \frac{d_{k+1}}{b_k} \cdot \frac{\lambda Beta(\alpha, \alpha)}{(b-a)} \cdot \frac{(r_l - r_{l-1})^{2\alpha-1}}{(\xi - r_{l-1})^{\alpha-1} (r_l - \xi)^{\alpha-1}}, & k' = k + 1 \\ \left[\frac{(r_l - r_{l-1})(r_{l+1} - r_l)}{(r'_l - r_{l-1})(r_{l+1} - r'_l)} \right]^{\alpha-1}, & k' = k \end{cases}$$

5.2. Updating Parameter ρ

According to the expression of the conditional posterior distribution of parameter ρ , it is known that the density function of the conditional posterior distribution of parameter ρ is complex and cannot be directly sampled. Therefore, the random walk Metropolis–Hastings algorithm is used to handle it. Assuming the current value of ρ is $\rho^{(t)}$ ($P(\rho^{(t)} | k, \mathbf{r}, \zeta, \sigma^2, \mathbf{Y}) > 0$), the candidate value ρ^* is generated from the proposal distribution $F(\rho^* | \rho^{(t)}) = f(\rho^* - \rho^{(t)})$ (where $f(\cdot)$ represents the density function), and the transition process is: $\rho^* = \rho^{(t)} + cz$, $z \sim N(0, 1)$. c is the transition parameter. The acceptance probability of ρ^* is then $A(\rho^* | \rho^{(t)}) = \min\{1, R\}$, where

$$\begin{aligned} R &= \frac{P(\rho^* | k, \mathbf{r}, \zeta, \sigma^2, \mathbf{Y}) F(\rho^{(t)} | \rho^*)}{P(\rho^{(t)} | k, \mathbf{r}, \zeta, \sigma^2, \mathbf{Y}) F(\rho^* | \rho^{(t)})} \\ &= \frac{|\mathbf{A}(\rho^*)| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{A}(\rho^*) \mathbf{Y} - \mathbf{X} \zeta)^T (\mathbf{A}(\rho^*) \mathbf{Y} - \mathbf{X} \zeta) \right\}}{|\mathbf{A}(\rho^{(t)})| \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{A}(\rho^{(t)}) \mathbf{Y} - \mathbf{X} \zeta)^T (\mathbf{A}(\rho^{(t)}) \mathbf{Y} - \mathbf{X} \zeta) \right\}} \end{aligned}$$

Finally, based on the conditional posterior distributions of the other parameters ζ, σ^2 in the model, Gibbs sampling is directly used to sample them.

Therefore, the estimation process of parameter $(k, \mathbf{r}, \rho, \zeta, \sigma^2)$ can be summarized into the following four steps:

Step 1: Assign initial values to all parameters $(k, \mathbf{r}, \rho, \zeta, \sigma^2)$, denoted as $(k_0, \mathbf{r}_0, \rho_0, \zeta_0, \sigma_0^2)$.

Step 2: Fix the initial value of ρ, ζ, σ^2 , update parameter $k, \mathbf{r}$ using the RJMCMC algorithm, and denote the updated value as $k_1, \mathbf{r}_1$.

Step 3: For each updated state $k_1, \mathbf{r}_1$ of $k, \mathbf{r}$, fix the value of ζ, σ^2 , update parameter ρ using the Metropolis–Hastings algorithm, and denote the updated value as ρ_1 .

Step 4: Fix the values of $k_1, \mathbf{r}_1$, and ρ_1 , use Gibbs sampling to sequentially update parameters ζ, σ^2 , and denote the updated parameters as ζ_1, σ_1^2 .

After these steps, the transition from the initial state $(k_0, \mathbf{r}_0, \rho_0, \zeta_0, \sigma_0^2)$ to the next state $(k_1, \mathbf{r}_1, \rho_1, \zeta_1, \sigma_1^2)$ is completed. Then, we use $(k_1, \mathbf{r}_1, \rho_1, \zeta_1, \sigma_1^2)$ as the new initial state and repeat steps 2, 3, and 4 until convergence. Finally, we estimate all unknown parameters using the posterior sample mean. The estimation of the unknown function

$h(x_i)$ can be approximated by the average of the Markov chain $E(h(x_i)|k, \mathbf{r}, \mathbf{Y})$. That is $\hat{h}(x_i) = E(h(x_i)|\mathbf{Y}) = E(E(h(x_i)|k, \mathbf{r}, \mathbf{Y}))$.

6. Numerical Simulation

In the following, a numerical example is used to simulate the estimation method proposed in the previous sections. In the numerical simulation, this study uses truncated power spline functions to fit the unknown function, with $m = 1$. In the prior distribution of β , $\mathbf{V} = n \cdot \mathbf{I}_{1+m+k}$, where n and $\mathbf{I}_{1+m+k}$ are the sample capacity and the $(1 + m + k)$ order identity matrix, respectively. In the Poisson prior, $\lambda = 1$, and in the prior distribution of σ^2 , $v_1 = v_2 = 0.001$. In the sampling process, $b_k = c_0 \cdot \min\left\{1, \frac{\pi(k+1)}{\sqrt{n\pi(k)}}\right\}$, $d_k = c_0 \cdot \min\left\{1, \frac{\sqrt{n\pi(k-1)}}{\pi(k)}\right\}$, and $m_k = 1 - b_k - d_k$, where $c_0 = 0.4$, $b_0 = 1$, and $d_0 = 0$. The hyperparameter is $\alpha = 1$ and transition parameter is $c = 0.5$.

For the initial values of the Markov chain, this study makes the following selections: the number of spline nodes $k = 2$, $k_{\max} = 10$, and the initial values of the node positions $\mathbf{r}$ are directly generated from the corresponding prior distribution $Dirichlet(1, 1, \dots, 1)$. The initial values of the spline coefficients are $\beta_0 = (0.1, 0.1, \dots, 0.1)'_{1+m+k}$. The initial values of the parameter ρ and variance σ^2 are set to their true values.

Regarding the estimation of the unknown function, the goodness of fit of the fitted function is measured by the mean square error (MSE) on 50 grid points within the interval $[-2, 2]$, i.e.,

$$MSE_h = \frac{1}{50} \sum_{i=1}^{100} \left(\hat{h}(u_i) - h(u_i) \right)^2$$

Considering the semiparametric spatial lag model:

$$\mathbf{Y} = \rho \mathbf{WY} + \mathbf{Z}\beta + \mathbf{H} + \varepsilon, \varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

where $\mathbf{Y} = (y_1, y_2, \dots, y_n)'$, $\mathbf{H} = (h(x_1), h(x_2), \dots, h(x_n))'$, $\mathbf{Z}$ is an $n \times l$ matrix, and $\beta = (\beta_1, \beta_2, \dots, \beta_l)'$. In this example, we set $\beta = (1, 1/\sqrt{2}, 1/\sqrt{3})'$, $h(x_i) = \cos(2x_i - 1)$, $\rho = 0.3$, $\sigma^2 = 0.5$, and the spatial weight matrix $\mathbf{W}$ is the Rook adjacency spatial weight matrix, which is randomly generated.

Using random methods and with the aid of MATLAB software, 625 sample points are generated, and the estimation method proposed in the previous sections is iterated 7000 times. The average of the posterior values from the last 5000 observations is taken as the estimate of the posterior mean.

Table 1 presents the true values, posterior means, posterior medians, standard deviations, and 95% highest posterior density intervals (HPD) of the model parameters. The simulation processes and results for each parameter are detailed in Figure 1. The left side of the figure shows the boxplots of the estimates of the components of parameter β , and the right side shows the boxplots of parameter ρ and the MSE between the function $h(\cdot)$ and its true value. Observing Table 1 and Figure 1, it can be seen that the estimated values of each parameter in the numerical example are very close to the true values, and the corresponding confidence intervals are relatively small and contain the true values, indicating accurate estimation results.

Table 1. Estimated results of the model parameters.

Parameters	True	Mean	Median	Standard Deviation	95% HPD
ρ	0.3	0.2743	0.2763	0.0521	(0.1723, 0.3764)
β_1	1	1.0328	1.0289	0.0982	(0.8404, 1.2251)
β_2	$1/\sqrt{2}$	0.7505	0.7475	0.0930	(0.5682, 0.9328)
β_3	$1/\sqrt{3}$	0.5607	0.5598	0.0905	(0.3834, 0.7380)

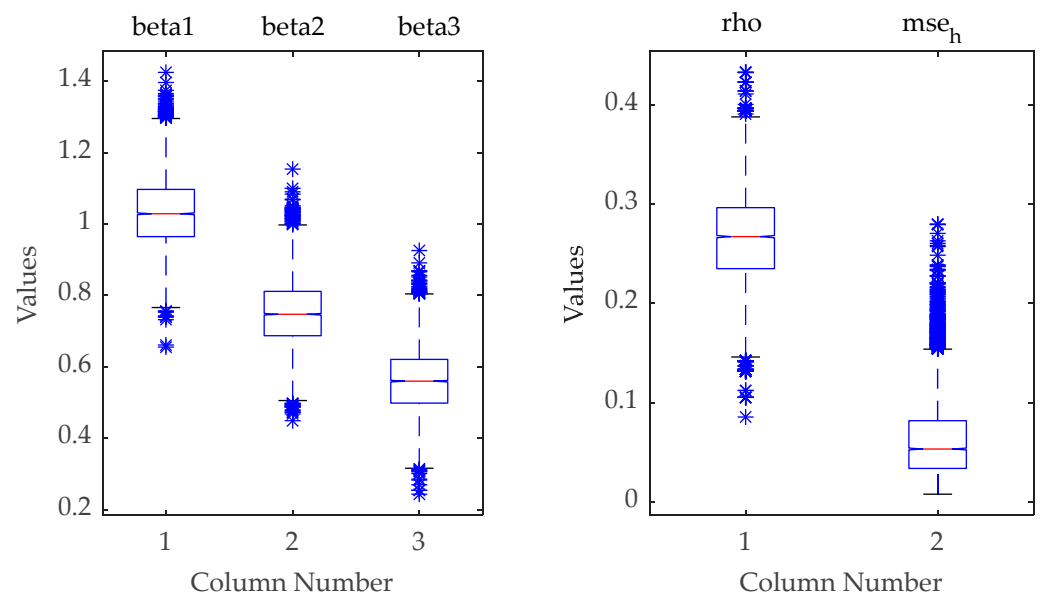


Figure 1. Simulation results for Example 1. The left side shows the boxplots of the estimates of the components of parameter β , and the right side shows the boxplots of parameter ρ and the MSE between the function $h(\cdot)$ and its true value.

To examine the performance of the proposed algorithm, this study also presents the running status of each parameter and the MSE of function $h(\cdot)$ in Figure 2. Figure 3 depicts the fitting of the function $h(\cdot)$, where the solid line represents the actual curve, the dotted line represents the fitted function, and the two dashed lines are the 95% confidence limit curves. The mean MSE between the function $h(\cdot)$ and its true value is 0.0509, indicating a very good fit.

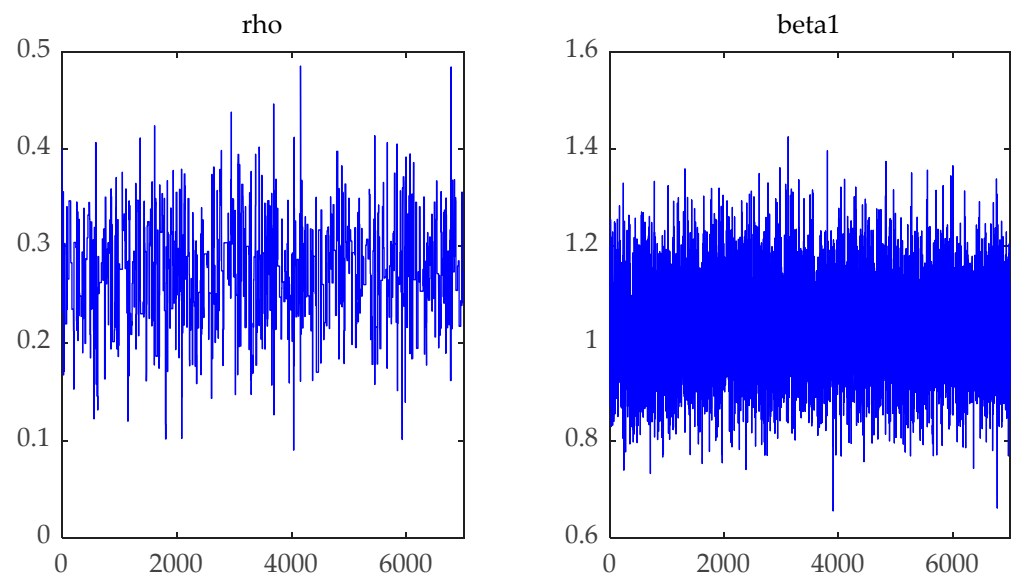


Figure 2. Cont.

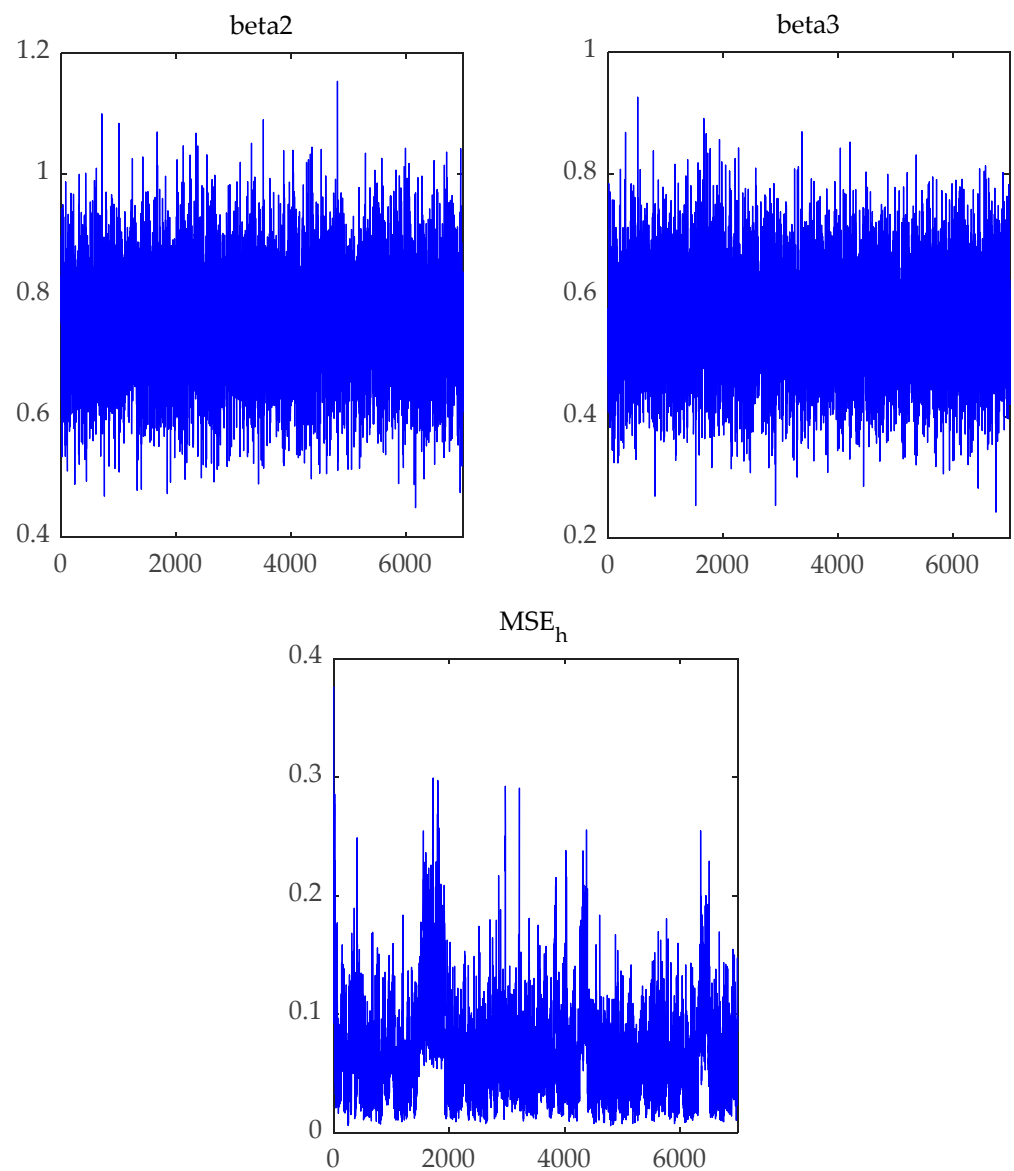


Figure 2. Running status of the parameters ρ , β , and MSE_h in Example 1.

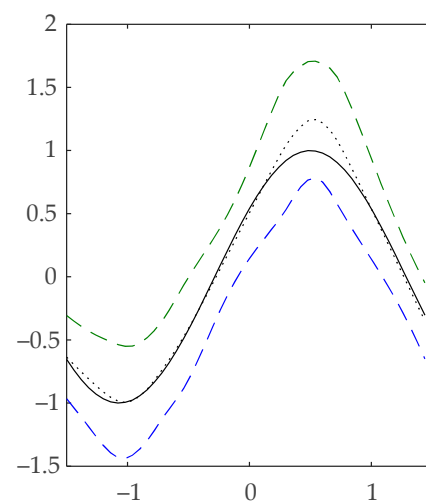


Figure 3. Fitting of the function $h(\cdot)$ in Example 1, where the solid line represents the actual curve, the dotted line represents the fitted function, and the two dashed lines are the 95% confidence limit curves.

7. Application Examples

China's economy is currently undergoing a transition from extensive development to high-quality development, and greening and low carbonization are basic requirements and main manifestations of high-quality development. In this context, it is particularly important to explore the relationship between China's economic growth and the environment, and the study of the interaction mechanism between the two has outstanding practical significance. Based on these considerations, this study will empirically examine the relationship between urban economic growth and carbon emissions in China using samples of prefecture-level and above cities.

7.1. Empirical Research Design

7.1.1. Empirical Model Specification

The *IPAT* equation is a commonly used method for assessing the environmental impact of human activities, first proposed by Commoner [21] and Ehrlich and Holdren [22]. It channels the impact of human activities on the environment (I) into three aspects: population (P), affluence (A), and technology (T), i.e.,

$$I = PAT \quad (23)$$

The *IPAT* equation is just an accounting identity and cannot be used for hypothesis testing. To overcome this limitation, Dietz and Rosa [23] extended the *IPAT* equation into a stochastic form and proposed the *STIRPAT* equation, whose mathematical definition is as follows:

$$I = aP^b A^c T^d e \quad (24)$$

where a , b , c , and d are the coefficients to be estimated, and e represents the random disturbance term. Taking the logarithm of both sides of the equation yields a multivariate linear regression model:

$$\ln I = \alpha + b \ln P + c \ln A + d \ln T + \ln e \quad (25)$$

At this point, the parameters b , c , and d represent the ecological elasticity of population, affluence, and technology, respectively.

Since there are many technological factors that affect the environment, it is generally difficult to find suitable proxy variables in practical applications. York et al. [24] believed that, without violating the product principle of the *IPAT* equation (i.e., T can be expressed as the product of its influencing factors), any variable that affects the level of technology can be added to Equation (25). In other words, variables other than population and affluence can be added to the regression model according to the concept of the technological multiplier, as long as these additional variables meet the concept of the technological multiplier.

To explore the impact of China's urban economic growth on carbon emissions, this study extends the *STIRPAT* model mentioned above. In addition to population and affluence factors, we also consider energy consumption factors and incorporate industrial structure and innovation input as technological factors into the model. At the same time, considering that carbon emissions between neighboring cities may affect each other and that there may be a nonlinear relationship between economic growth and carbon emissions (such as the Environmental Kuznets Curve, which indicates a U-shaped nonlinear relationship between economic growth and environmental quality), we construct the following empirical model based on the semiparametric spatial lag model proposed earlier:

$$\begin{aligned} \ln CO_{2it} = & \rho \sum_{j=1}^n w_{ij} \ln CO_{2jt} + \beta_1 \ln pop_{it} + \beta_2 instr_{it} \\ & + \beta_3 inov_{it} + f(\ln pgdp_{it}) + \varepsilon_{it} \end{aligned} \quad (26)$$

where the subscripts i, t represent cities and years, respectively; CO_2 is carbon emissions; $pgdp$ represents per capita income; $pop, ener, instr, inov$ represent population, energy con-

sumption, industrial structure, and innovation input, respectively; $f(\cdot)$ is an unknown function; ε_{it} is a random disturbance; ρ is the spatial correlation coefficient; w_{ij} is the (i, j) element of the spatial weight matrix. We choose the distance threshold weight matrix as the spatial weight matrix; that is, first determine a minimum distance based on the criterion that each city has “neighbors” and then find all “neighbors” of each city based on this distance. If two cities are neighbors, their weight relative to each other is 1; otherwise, it is 0. In the empirical analysis, we normalize the weight matrix by rows.

7.1.2. Variable Measurement and Data Sources

The measurement methods for each variable in Equation (26) are as follows:

- (1) Carbon emissions ($\ln CO_2$): Measured by the logarithm of the total carbon emissions of a city. First, estimate the national total carbon emissions CO_2 using the following method:

$$CO_2 = \sum_{j=1}^n c_j \times \lambda_j \quad (27)$$

where c_j is the total consumption of the j type of fossil fuel, and λ_j is the carbon emission coefficient of the j type of fossil fuel. Then, following the approach of Fan et al. [25], we match the total carbon emissions CO_2 to each city using the light-matching method to obtain the total carbon emissions of each city CO_2 .

- (2) Population ($\ln pop$): Measured by the logarithm of the ratio of the total population of a prefecture-level city to the total area of its administrative region. It represents the population factor affecting the environment.
- (3) Per capita income ($\ln pgdp$): Measured by the logarithm of the per capita GDP of a city. It represents the wealth factor affecting the environment.
- (4) Industrial structure ($instr$): Measured by the proportion of the value added of the secondary industry to GDP in a prefecture-level city. It represents the industrial factor affecting the environment.
- (5) Innovation input ($inov$): Measured by the proportion of scientific expenditure in fiscal expenditure of a prefecture-level city to GDP. It represents the technological factor affecting the environment.

In the empirical analysis, this study uses panel data from Chinese prefecture-level cities spanning from 2009 to 2019. The original data are sourced from the annual *China City Statistical Yearbook*. Descriptive statistics of the main variables are presented in Table 2, where N, Mean, Median, SD, Min, and Max, respectively, denote sample size, sample mean, sample median, sample standard deviation, minimum value, and maximum value.

Table 2. Descriptive statistics of the main variables.

Variable	N	Mean	Median	SD	Min	Max
lnc	3100	1.357	1.366	0.323	0.463	2.363
lnpop	3100	2.561	2.567	0.289	1.308	3.538
lnpgdp	3100	4.596	4.584	0.270	3.652	5.670
instr	3100	47.485	47.815	10.342	10.680	82.240
ino	3100	0.255	0.181	0.261	0.013	6.334

7.2. Model Estimation Results and Analysis

We use the Bayesian estimation method constructed earlier to estimate the parameters and unknown function in Equation (26). Additionally, for comparison purposes, we also estimate the nonspatial linear model, nonspatial semiparametric model, and linear spatial model. The parameter results are shown in Table 3, and the nonparametric function estimation results are shown in Figures 4 and 5.

Table 3. Model parameter estimation results.

Variables	Nonspatial Model	Linear Spatial Model	Nonspatial Semiparametric Model	Semiparametric Spatial Model	
				Parameter Estimation Results	95% Confidence Interval
lnpop	0.719 *** (0.012)	0.671 *** (0.011)	0.723 *** (0.012)	0.675	(0.653, 0.697)
lnpgdp	0.571 *** (0.014)	0.456 *** (0.014)			
instr	0.001 *** (0.000)	0.001 * (0.000)	0.001 *** (0.000)	0.001	(−0.000, −0.0012)
ino	−0.143 *** (0.014)	−0.127 *** (0.013)	−0.140 *** (0.015)	−0.127	(−0.153, −0.101)
Constant	−3.130 *** (0.068)	−3.126 *** (0.062)			
W*lnC		0.495 *** (0.019)		0.495	(0.458, 0.532)

Standard errors in parentheses *** $p < 0.01$, * $p < 0.1$.

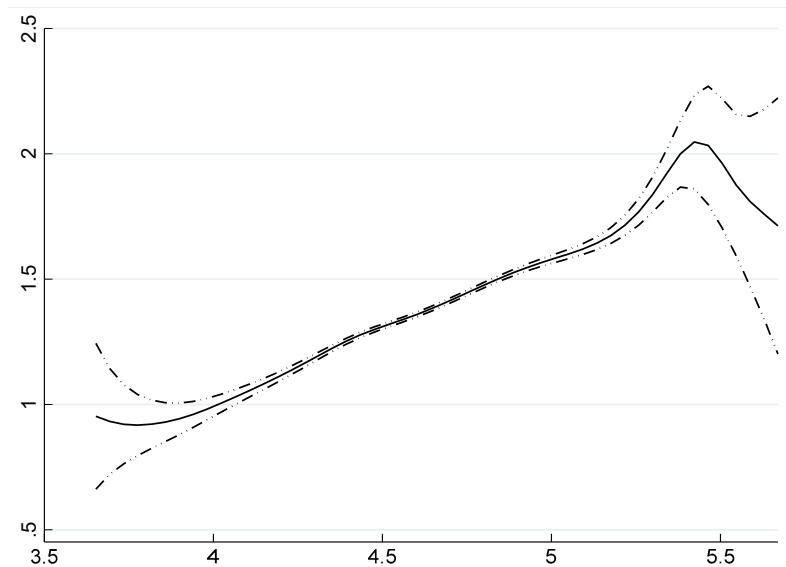


Figure 4. Estimation result of the unknown function $f(\cdot)$ in the nonspatial semiparametric.

From the estimation results in Table 3, it can be observed that in all four types of models, factors such as population, wealth, and industrial proportion have significant positive effects on carbon emissions. An increase in population size, wealth level, and industrial proportion will all promote an increase in carbon emissions. On the other hand, the impact of innovation input is the opposite; an increase in innovation input level helps suppress carbon emissions. Furthermore, from the estimation results of the variable coefficients, it can be concluded that population and wealth factors remain the core factors exacerbating carbon emissions, while improving the innovation level is an important means of mitigating carbon emissions. The impact of industrial structure on carbon emissions is relatively limited.

The estimation results of the linear spatial model reveal significant positive spatial clustering effects and spatial spillover effects in carbon emissions, consistent with existing literature. This also indicates that spatial correlation is an important factor that must be considered in the empirical model. Compared to the nonspatial linear model, the linear spatial model shows a significant decrease in the impact of factors such as population, wealth, and innovation input on carbon emissions, indicating that omitting spatial correlation can exaggerate the impact of these factors on carbon emissions.

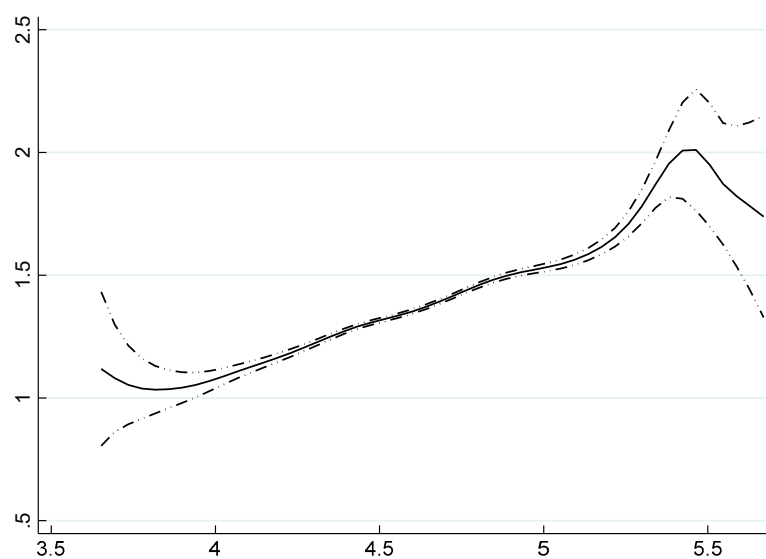


Figure 5. Estimation result of the unknown function $f(\cdot)$ in the semiparametric spatial model.

From the estimation results of the nonspatial semiparametric model and the semiparametric spatial model, it can be observed that compared to the corresponding linear model, the estimation results of the parameter part in the nonlinear model are relatively close to the linear model. This suggests that the nonparametric structure of the model does not affect the identification of the parameter part and is therefore a robust setting. Based on the above analysis, it can be concluded that the semiparametric spatial lag model, which considers both spatial correlation and semiparametric structure, is a relatively robust model setting in empirical research.

Figures 4 and 5 show the estimation results of the nonparametric component of two types of semiparametric models. The solid line represents the estimation of the unknown function, while the two dashed lines in between denote the 95% confidence band. From the figures, it can be observed that the impact of wealth level on carbon emissions exhibits an “inverted U” shape, indicating the presence of an Environmental Kuznets Curve effect for carbon emissions. Moreover, this effect has a “tail” characteristic, especially before reaching the turning point, where the promoting effect of wealth level on carbon emissions appears to be relatively flat. This suggests that using a simple linear model for estimation would miss the identification of the Environmental Kuznets Curve effect for carbon emissions. Additionally, when comparing the estimation results of the unknown function in the two types of models, it can be observed that the shapes of the unknown functions in both models are highly similar. This indicates that considering spatial correlation does not affect the identification of the nonparametric part of the model. However, the overall unknown function in the spatial model is larger than that in the nonspatial model, suggesting that ignoring spatial correlation would underestimate the unknown function. This further emphasizes that the semiparametric spatial model is a more robust model setting.

The empirical results above indicate that in China, there exists a complex and close relationship between economic growth and carbon dioxide emissions. The empirical research conclusions demonstrate the significant impact of factors such as population, wealth, industrial proportion, and innovation input on carbon emissions, revealing direct connections between certain aspects of economic growth and carbon emissions.

First, the increase in population size and wealth level has been found to have a promoting effect on carbon emissions, closely related to China’s current economic development stage and industrial structure. As the economy grows rapidly, the processes of industrialization and urbanization accelerate, leading to increased energy consumption and carbon emissions.

Second, the industrial proportion has a significant impact on carbon emissions, indicating that the industrial sector is one of the main sources of carbon emissions. This also reflects the environmental pressure posed by China's industrial-based economic development model.

Third, the increase in innovation input helps to restrain the growth of carbon emissions, indicating that technological innovation plays a crucial role in slowing down the increase in carbon emissions. This demonstrates that China, while promoting economic growth, is actively seeking ways to reduce carbon emissions through technological innovation.

Furthermore, there is a nonlinear "inverted U" relationship between wealth level and carbon emissions, known as the Environmental Kuznets Curve effect. This suggests that with further economic growth, carbon emissions may reach a turning point and tend to decline, providing hope for the future decoupling of economic development and carbon emissions.

8. Conclusions

This paper proposes a semiparametric spatial lag model and develops a Bayesian estimation method for this model by using polynomial spline functions to fit the unknown link function. In the Bayesian estimation approach, the paper combines the Reversible Jump Markov Chain Monte Carlo method, Metropolis–Hastings algorithm with random walk, and Gibbs sampling techniques to sample all the parameters. The paper also conducts numerical simulations to demonstrate the Bayesian estimation theory using a numerical example, where the spatial weight matrix W is constructed using both Rook adjacency and Case adjacency spatial weight matrices. The numerical simulation results show that under both types of spatial weight matrices, the estimation performance of the parameter part and the nonparametric function is satisfactory, indicating that the proposed Bayesian estimation method has high estimation accuracy and robustness. Finally, the paper applies the proposed theoretical method to an empirical study on the nonlinear relationship between per capita GDP and carbon emissions in China, further demonstrating the applicability and value of the proposed Bayesian estimation method in empirical research.

The semiparametric spatial lag model and its Bayesian estimation method developed in this paper, by comprehensively considering the nonlinear relationships between variables and spatial correlations, provide a more accurate and robust analytical framework for understanding the complex relationship between China's economic growth and carbon emissions, offering strong empirical support for policy-making. Firstly, the primary drivers of economic growth such as population growth, wealth accumulation, and industrial development significantly contribute to the increase in China's carbon emissions, highlighting a direct link between economic development and environmental pressure. However, increased investment in innovation effectively mitigates this trend, indicating that technological innovation is key to balancing economic growth and environmental protection. Secondly, the paper reveals spatial agglomeration and spillover effects of carbon emissions, emphasizing the mutual influence of carbon emissions behavior between regions and suggesting the importance of considering regional collaborative governance in policy-making to address cross-regional environmental challenges. Furthermore, the study finds a nonlinear "inverted U" relationship between wealth levels and carbon emissions, known as the Environmental Kuznets Curve, suggesting that as economic growth progresses further, carbon emissions may reach a turning point and begin to decline, offering hope for decoupling future economic development from carbon emissions.

In comparison to traditional frequentist methods, the advantage of Bayesian estimation lies in its ability to better handle situations with small sample sizes, more complex models, and diverse parameters. The superior estimation capability of the Bayesian method in small sample situations allows the model and its estimation method proposed in this paper to be applicable to a wider range of empirical analyses. However, in Bayesian analysis, the construction of prior distributions is not a straightforward task, making the specification of prior distributions for model parameters generally challenging. Furthermore, in Bayesian

analysis, typically only the kernel of the posterior distribution density function is known, making it difficult to obtain the specific density function and to find numerical quantiles of the cumulative distribution function. Therefore, the derivation of the posterior distribution is often challenging.

Author Contributions: Conceptualization, K.L. and L.F.; methodology, K.L. and L.F.; software, K.L.; validation, K.L. and L.F.; formal analysis, L.F.; investigation, K.L. and L.F.; resources, L.F.; data curation, K.L.; writing—original draft preparation, K.L. and L.F.; writing—review and editing, K.L. and L.F.; visualization, L.F.; supervision, L.F.; project administration, L.F.; funding acquisition, K.L. and L.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Social Science Fund of China, grant number 22BTJ026; National Natural Science Foundation of China, grant number 72003034; Natural Science Foundation of Fujian Province in China, grant number 2021J01113; Science and Technology Innovation Special Fund Project of Fujian Agricultural and Forestry University, grant numbers CXZX2022026, CXZX2023026, and KFb22106XA; and Outstanding Young Research Talent Program of Fujian Agriculture and Forestry University, grant number xjg201821.

Data Availability Statement: The data will be made available by the authors on request.

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

References

- Long, X.; Zhu, Y.; Cai, W.; Li, S. An empirical analysis of spatial tax competition among Chinese counties based on spatial econometric models. *Econ. Res. J.* **2014**, *49*, 41–53. (In Chinese)
- Li, K.; Fang, L.; He, L. How urbanization affects China's energy efficiency: A spatial econometric analysis. *J. Clean. Prod.* **2018**, *200*, 1130–1141. [\[CrossRef\]](#)
- Li, K.; Fang, L.; He, L. The impact of energy price on CO₂ emissions in China: A spatial econometric analysis. *Sci. Total Environ.* **2020**, *706*, 135942. [\[CrossRef\]](#)
- Su, L.; Jin, S. Profile Quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econom.* **2010**, *157*, 18–33. [\[CrossRef\]](#)
- Su, L. Semiparametric estimation of spatial autoregressive models. *J. Econom.* **2012**, *167*, 543–560. [\[CrossRef\]](#)
- Malikov, E.; Sun, Y. Semiparametric estimation and testing of smooth coefficient spatial autoregressive models. *J. Econom.* **2017**, *199*, 12–34. [\[CrossRef\]](#)
- Malikov, E.; Sun, Y.; Hite, D. (Under) Mining local residential property values: A semiparametric spatial quantile autoregression. *J. Appl. Econom.* **2019**, *34*, 82–109. [\[CrossRef\]](#)
- Li, K.; Chen, J. Profile maximum likelihood estimation of semi-parametric varying coefficient spatial lag model. *J. Quant. Technol. Econ.* **2013**, *30*, 85–98. (In Chinese)
- Sun, L. Estimation and Application of Panel Data Semiparametric Spatial Lag Models. Ph.D. Thesis, Xiamen University, Xiamen, China, 2015. (In Chinese)
- Gelfand, A.; Kottas, A.; MacEachern, S. Bayesian nonparametric spatial modeling with dirichlet process mixing. *J. Am. Stat. Assoc.* **2005**, *100*, 1021–1035. [\[CrossRef\]](#)
- Han, X.; Lee, L. Bayesian estimation and model selection for spatial durbin error model with finite distributed lags. *Reg. Sci. Urban Econ.* **2013**, *43*, 816–837. [\[CrossRef\]](#)
- Kang, E.; Cressie, N. Bayesian inference for the spatial random effects model. *Publ. Am. Stat. Assoc.* **2011**, *106*, 972–983. [\[CrossRef\]](#)
- Fang, L.; Qian, Z. Bayesian estimation of non-parametric spatial lag model. *J. Quant. Technol. Econ.* **2013**, *30*, 72–84. (In Chinese)
- Fang, L. Bayesian Estimation of the spatial lag model. *Stat. Res.* **2014**, *31*, 102–106. (In Chinese)
- Lesage, J. Spatial econometric panel data model specification: A Bayesian approach. *Soc. Sci. Electron. Publ.* **2014**, *9*, 122–145. [\[CrossRef\]](#)
- Louzada, F.; Diego, C.; Osafu, A. Spatial statistical models: An overview under the Bayesian approach. *Axioms* **2021**, *10*, 307. [\[CrossRef\]](#)
- Thapa, S.; Lomholt, M.; Krog, J.; Cherstvy, A.; Metzler, R. Bayesian analysis of single-particle tracking data using the nested-sampling algorithm: Maximum-likelihood model selection applied to stochastic-diffusivity data. *Phys. Chem. Chem. Phys.* **2018**, *20*, 29018–29037. [\[CrossRef\]](#)
- Gelfand, A.; Smith, A. Sampling-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* **1990**, *85*, 398–409. [\[CrossRef\]](#)
- Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Augusta, H.; Edward, T. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [\[CrossRef\]](#)
- Hastings, W. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [\[CrossRef\]](#)
- Commoner, B. *The Closing Circle*, 1st ed.; Random House Inc.: New York, NY, USA, 1971.

22. Ehrlich, P.; Holdren, J. A Bulletin dialogue on the 'closing circle': Critique: One dimensional ecology. *Bull. At. Sci.* **1972**, *28*, 16–27. [[CrossRef](#)]
23. Dietz, T.; Rosa, E. Rethinking the environmental impacts of population, affluence and technology. *Hum. Ecol. Rev.* **1994**, *1*, 277–300.
24. York, R.; Rosa, E.; Dietz, T. STIRPAT, IPAT and ImPACT: Analytic tools for unpacking the driving forces of environmental impacts. *Ecol. Econ.* **2003**, *46*, 351–365. [[CrossRef](#)]
25. Fan, Z.; Peng, F.; Liu, C. Political connections and economic growth: Evidence from the DMSP/OLS satellite data. *Econ. Res. J.* **2016**, *51*, 114–126. (In Chinese)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.