


Article

A Semiparametric Bayesian Approach to Heterogeneous Spatial Autoregressive Models

Ting Liu, Dengke Xu *  and Shiqi Ke

School of Economics, Hangzhou Dianzi University, Hangzhou 310018, China; 22150107@hdu.edu.cn (T.L.); 231150084@hdu.edu.cn (S.K.)

* Correspondence: xudengke1983@163.com

Abstract: Many semiparametric spatial autoregressive (SSAR) models have been used to analyze spatial data in a variety of applications; however, it is a common phenomenon that heteroscedasticity often occurs in spatial data analysis. Therefore, when considering SSAR models in this paper, it is allowed that the variance parameters of the models can depend on the explanatory variable, and these are called heterogeneous semiparametric spatial autoregressive models. In order to estimate the model parameters, a Bayesian estimation method is proposed for heterogeneous SSAR models based on B-spline approximations of the nonparametric function. Then, we develop an efficient Markov chain Monte Carlo sampling algorithm on the basis of the Gibbs sampler and Metropolis–Hastings algorithm that can be used to generate posterior samples from posterior distributions and perform posterior inference. Finally, some simulation studies and real data analysis of Boston housing data have demonstrated the excellent performance of the proposed Bayesian method.

Keywords: heterogeneous semiparametric spatial autoregressive models; Bayesian estimate; Gibbs sampler; Metropolis–Hastings algorithm; B-spline



Citation: Liu, T.; Xu, D.; Ke, S. A Semiparametric Bayesian Approach to Heterogeneous Spatial Autoregressive Models. *Entropy* **2024**, *26*, 498. <https://doi.org/10.3390/e26060498>

Academic Editor: Ricardo Sandes Ehlers

Received: 15 April 2024

Revised: 1 June 2024

Accepted: 5 June 2024

Published: 7 June 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

In recent decades, spatial data analysis based on spatial autoregressive (SAR) models has become a very important and popular research direction in the academic research of econometricians and statisticians. Among them, in-depth research has been conducted on inference theories and methods based on linear SAR models and their extensions, including estimation, variable selection, hypothesis testing, etc., and a large amount of literature has also been produced, such as [1–3]. Specifically, there are many in-depth research achievements on linear SAR models, such as [4–7]. However, in spatial data analysis, there is often a nonlinear relationship between response variables and covariates. Therefore, in order to explore this complex phenomenon, some semiparametric SAR models have been proposed in recent years and have been thoroughly studied. For example, based on partially linear spatial autoregressive models, Su and Jin [8] proposed a profile quasi-maximum likelihood estimation method and established asymptotic theoretical properties of the obtained estimators. By using the spline approximations and instrumental variables estimation method, Du et al. [9] developed an estimation method for partially linear additive spatial autoregressive models, and derived asymptotic theoretical properties of the obtained estimators. For partially linear single-index spatial autoregressive models, Cheng and Chen [10] developed an estimation method and established consistency and asymptotic normality of the estimators under some mild assumptions. Other related research results on semiparametric SAR models can also be found in [11] [12]. Previous research on various spatial autoregressive models was mainly based on the assumption of homoscedasticity, which assumes that the variance of model errors is constant. As is well known, heteroscedasticity is a common phenomenon in spatial data analysis. Therefore, using statistical inference methods under the assumption of homogeneity may lead to erroneous inference, as seen in Lin and

Lee [13]. Therefore, it is necessary to study heterogeneous spatial autoregressive models. Especially in recent years, many researchers have conducted in-depth research on spatial autoregressive models where the error variance is heteroscedasticity. For example, SAR models with heteroscedasticity was studied by Dai et al. [14] for Bayesian local influence analysis. However, existing literature on heterogeneous spatial autoregressive models assume that the variance term is fixed and does not perform regression modeling analysis like the mean. In addition, in many application fields, such as econometrics, the assumption of equal variance may not be suitable for modeling data that exhibit heteroscedasticity. Therefore, we propose a new model, called a heterogeneous semiparametric spatial autoregressive model, in which the variance parameters are allowed to be modeled by using some covariates.

In addition, many estimation methods have recently been developed for SAR models from both frequentist and Bayesian perspectives. Specifically, due to the rapid development of advanced computing technology in the era of big data, Bayesian statistical analysis of SAR models and various other statistical models has received increasing attention, and large quantities of related research achievements have emerged in recent years. For example, within the framework of longitudinal data and for the generalized partial linear mixed models, Tang and Duan [15] studied an effective semiparametric Bayesian method. By using spline approximation, Xu and Zhang [16] introduced a Bayesian method for the partially linear model with heteroscedasticity based on the variance modelling technique. Based on the assumption that the response variables and random effects follow multivariate skew-normal distributions, a new spatial dynamic panel data model was proposed by Ju et al. [17] and a Bayesian local influence analysis method was developed to simultaneously evaluate the impact of small perturbations on the data, priors, and sampling distributions. Pfarrhofer and Piribauer [18] studied Bayesian variable selection for high-dimensional spatial autoregressive models based on two shrinkage priors. Wang and Tang [19] made Bayesian statistical inference based on a quantile regression model with nonignorable missing covariates. To capture the linear and nonlinear relationships between explanatory variables and their responses to spatially relevant data, Chen and Chen [20] developed a Bayesian sampling-based method based on the partially linear single-index spatial autoregressive models, in which it includes an efficient MCMC approach and explores the joint posterior distributions by using a Gibbs sampler. Within the framework of longitudinal data, Zhang et al. [21] proposed semiparametric mixed-effects double regression models for analysis based on spline approximation technology, in which they jointly modeled the mean and variance of the mixed-effects as a function of covariates. To our knowledge, there is not much work on semiparametric Bayesian methods for heterogeneous spatial autoregressive models due to their complex spatial correlation structures. Therefore, based on a hybrid effective algorithm that combines a Gibbs sampler and the Metropolis–Hastings algorithm and has the advantages of both algorithms, this paper develops a Bayesian method for heterogeneous semiparametric spatial autoregressive models based on variance modeling.

The outline of the paper is as follows. A new heterogeneous SSAR model is introduced in Section 2. In Section 3, we derive the full conditional distributions for implementing the sampling-based method, and develop a Bayesian method to obtain estimates by using a Gibbs sampler and the Metropolis–Hastings algorithm. Section 4 presents some simulation studies to illustrate the proposed methodology. As an application example, Section 5 analyzes the Boston house price data by using the proposed method. A brief conclusion and discussion is given in Section 6.

2. Heterogeneous Semiparametric Spatial Autoregressive Models

As is well known, the form of classical semiparametric spatial autoregressive models is as follows:

$$Y = \rho WY + X\beta + g(U) + \varepsilon, \quad (1)$$

where $Y = (y_1, y_2, \dots, y_n)^T$ is an n -dimensional response variable, $|\rho| < 1$ is an unknown spatial lag parameter that reflects spatial autocorrelation between neighbors, and W is a known spatial weight matrix with zero diagonal elements. In the mean model, $X = (x_1, x_2, \dots, x_n)^T$ is an $n \times p$ explanatory variable matrix where the i th row is $x_i^T = (x_{i1}, \dots, x_{ip})$ and $\beta = (\beta_1, \dots, \beta_p)^T$ is a p -dimensional unknown regression coefficient to be estimated; moreover, $g(\cdot)$ is an arbitrary unknown smooth function in the mean model, which needs to be estimated; $U = (u_1, u_2, \dots, u_n)^T$ is an n -dimensional vector whose i th row u_i is an univariate observed covariate; ε is an n -dimensional vector that represents the regression errors of an independent and identically distributed regression disturbances with zero mean and finite variance σ^2 .

In addition, according to Xu and Zhang [16], this paper considers the heterogeneity of the variance in the model and assumes that the variance parameters are related to other explanatory variables; thus, we establish a regression model for the variance parameters, namely

$$\sigma_i^2 = h(z_i^T \gamma), \tag{2}$$

where $z_i^T = (z_{i1}, \dots, z_{iq})$ is an explanatory variable vector related to the variance of ε_i and $\gamma = (\gamma_1, \dots, \gamma_q)^T$ is a q -dimensional unknown regression coefficient to be estimated in the variance model. Some elements in z_i may coincide with some elements in x_i . In addition, for the identifiability of the models and considering that the variance is positive, $h(\cdot)$ is a known monotonic positive function. For example, exponential functions are often used to model the variance. So, heterogeneous SSAR models are considered in this paper as follows:

$$\begin{cases} Y = \rho WY + X\beta + g(U) + \varepsilon, \\ \varepsilon \sim N(0, \Sigma), \\ \Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2) \\ \sigma_i^2 = \exp(z_i^T \gamma), \\ i = 1, 2, \dots, n. \end{cases} \tag{3}$$

3. Bayesian Inference

3.1. B-Splines for the Nonparametric Function

From model (3), we obtain the log-likelihood function

$$\ell_n(\rho, \beta, \gamma | Y, X, Z, U) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n z_i^T \gamma + \ln|A| - \frac{1}{2} e^T \Sigma^{-1} e, \tag{4}$$

where $e = AY - X\beta - g(U)$, $A = I_n - \rho W$, and I_n is an $(n \times n)$ identity matrix.

There are now a large number of nonparametric techniques for handling nonparametric functions $g(\cdot)$ in (3), such as the smoothing splines and the kernel methods. This paper considers B-spline approximation to transform $g(\cdot)$ into a linear function composed of a set of basis functions. It can be summarized as follows: $\{s_i\}, i = 1, \dots, n$ performs a partition on the interval $[0, 1]$, which is called as the internal knots and satisfies $0 = s_0 < s_1 < \dots < s_{k_n} < s_{k_n+1} = 1$. This results in $K = K_n + l$ normalized B-spline basis functions of order l , which form the basis of a linear spline space. The main reason for using B-splines here is because they have advantages such as bounded support and numerical stability. As well as we know, selection of knots is usually an important aspect that cannot be ignored in the implementation process of B-splines. In this paper, our main focus is inference on the parameters in the mean model and the variance model. Therefore, by following the idea of Zhang et al. [21], the number of internal knots is selected as the integer part of $n^{1/5}$. Thus, $\pi^T(u)\alpha$ is used to approximate $g(u)$, in which $\pi(u) = (B_1(u), \dots, B_K(u))^T$ is a basis function vector and $\alpha \in R^K$. In this way, we can linearize the nonparametric function $g(\cdot)$ in (3) as follows:

$$g(u_i) \approx \pi^T(u_i)\alpha. \tag{5}$$

Thus, based on (5), we can rewritten the likelihood function (4) as follows:

$$\ell_n(\rho, \beta, \gamma|Y, X, Z, U) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2} \sum_{i=1}^n z_i^T \gamma + \ln|A| - \frac{1}{2}(AY - X\beta - B\alpha)^T \Sigma^{-1}(AY - X\beta - B\alpha), \tag{6}$$

where $B = (\pi(u_1), \pi(u_2), \dots, \pi(u_n))^T$.

3.2. Prior Selection of Parameters

This paper will use a Bayesian approach to estimate unknown parameters ρ, β, α and γ . Thus, to obtain Bayesian estimation, the prior distributions of unknown parameters to be estimated in the model should be given first. For the convenience of algorithm implementation, normal prior distributions are often chosen as $\beta \sim N(\beta_0, b_\beta)$, $\alpha \sim N(\alpha_0, \tau^2 I_K)$, $\gamma \sim N(\gamma_0, B_\gamma)$, and $\rho \sim U(-1, 1)$, in which $\beta_0, \alpha_0, \gamma_0$ and b_β, B_γ are known hyperparameter vectors or matrices. Moreover, the prior distribution of τ^2 is the $IG(a_\tau, b_\tau)$, and its density function is

$$p(\tau^2|a_\tau, b_\tau) \propto (\tau^2)^{-a_\tau-1} \exp\left(-b_\tau/\tau^2\right),$$

in which a_τ and b_τ are assumed to be positive constants and known. In this paper, we mainly focus on the case where the prior distribution of model parameters is a normal distribution. However, the proposed computational algorithm is also applicable to other specific prior distributions.

3.3. Posterior Inference

Let $\theta = (\beta, \alpha, \gamma, \rho)$ and then we aim to estimate the unknown parameters of θ . Based on the proposed model (3) and Gibbs sampling, the specific sampling process is carried out according to the following steps by sampling from joint posterior distribution $p(\theta|Y, X, Z, U)$.

Step 1. The initial values of parameters are set as $\theta^{(0)} = (\beta^{(0)}, \alpha^{(0)}, \gamma^{(0)}, \rho^{(0)})$.

Step 2. Compute $\Sigma^{(l)} = \text{diag}\{\exp(z_i^T \gamma^{(l)})\}$ and $A^{(l)} = I_n - \rho^{(l)}W$ on the basis of $\theta^{(l)} = (\beta^{(l)}, \alpha^{(l)}, \gamma^{(l)}, \rho^{(l)})$.

Step 3. Based on $\theta^{(l)} = (\beta^{(l)}, \alpha^{(l)}, \gamma^{(l)}, \rho^{(l)})$, sample $\theta^{(l+1)} = (\beta^{(l+1)}, \alpha^{(l+1)}, \gamma^{(l+1)}, \rho^{(l+1)})$ as follows:

- Sampling $\tau^{2(l+1)}$ from the conditional distribution below:

$$p(\tau^2|\alpha^{(l)}) \propto (\tau^2)^{-\frac{K}{2}-a_\tau-1} \exp\left\{-\frac{(\alpha^{(l)} - \alpha_0)^T(\alpha^{(l)} - \alpha_0) + 2b_\tau}{2\tau^2}\right\}. \tag{7}$$

- Sampling $\alpha^{(l+1)}$ from the conditional distribution below:

$$p(\alpha|Y, X, Z, U, \gamma^{(l)}, \rho^{(l)}) \sim N(\tilde{\mu}_\alpha, \tilde{\Sigma}_\alpha), \tag{8}$$

where $\tilde{\mu}_\alpha = \tilde{\Sigma}_\alpha(\tau^{-2(l+1)}I_K\alpha_0 + B^T\Sigma^{(l)-1}(A^{(l)}Y - X\beta^{(l)}))$ and $\tilde{\Sigma}_\alpha = (\tau^{-2(l+1)}I_K + B^T\Sigma^{(l)-1}B)^{-1}$.

- Sampling $\beta^{(l+1)}$ from the conditional distribution below:

$$p(\beta|Y, X, Z, U, \alpha^{(l+1)}, \gamma^{(l)}, \rho^{(l)}) \sim N(\tilde{\mu}_\beta, \tilde{\Sigma}_\beta), \tag{9}$$

where $\tilde{\mu}_\beta = \tilde{\Sigma}_\beta(b_\beta^{-1}\beta_0 + X^T\Sigma^{(l)-1}(A^{(l)}Y - B\alpha^{(l+1)}))$ and $\tilde{\Sigma}_\beta = (b_\beta^{-1} + X^T\Sigma^{(l)-1}X)^{-1}$.

- Sampling $\gamma^{(l+1)}$ from the conditional distribution below:

$$p(\gamma|Y, X, Z, U, \beta^{(l+1)}, \alpha^{(l+1)}, \rho^{(l)}) \propto \frac{|\Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}e^{(l,l+1),T}\Sigma^{-1}e^{(l,l+1)} - \frac{1}{2}(\gamma - \gamma_0)^T B_\gamma^{-1}(\gamma - \gamma_0)\}}{\tag{10}}$$

where $e^{(l,l+1)} = A^{(l)}Y - X\beta^{(l+1)} - B\alpha^{(l+1)}$.

- Sampling $\rho^{(l+1)}$ from the conditional distribution below:

$$p(\rho|Y, X, Z, U, \beta^{(l+1)}, \alpha^{(l+1)}, \gamma^{(l+1)}) \propto |A| \exp\left\{-\frac{1}{2}e^{(l+1)T}\Sigma^{(l+1)^{-1}}e^{(l+1)}\right\}. \tag{11}$$

where $e^{(l+1)} = A^{(l+1)}Y - X\beta^{(l+1)} - B\alpha^{(l+1)}$.

Step 4. Repeating Steps 2 and 3.

According to the steps of the above algorithm, we can easily obtain the sample sequences $(\beta^{(t)}, \alpha^{(t)}, \gamma^{(t)}, \rho^{(t)}), t = 1, 2, \dots$. It is easy to find that the fully conditional distributions $p(\tau^2|\alpha^{(l)}), p(\beta|Y, X, Z, U, \alpha^{(l)}, \gamma^{(l)}, \rho^{(l)})$ and $p(\alpha|Y, X, Z, U, \beta^{(l+1)}, \gamma^{(l)}, \rho^{(l)})$ are Inverse Gamma and normal distributions, respectively, and extracting observations from these familiar standard distributions is fast and easy. Unfortunately, fully conditional distributions $p(\gamma|Y, X, Z, U, \beta^{(l+1)}, \alpha^{(l+1)}, \rho^{(l)})$ and $p(\rho|Y, X, Z, U, \beta^{(l+1)}, \alpha^{(l+1)}, \gamma^{(l+1)})$ are nonstandard distributions and appear quite complex, making it quite difficult to extract random numbers from these conditional distributions. To solve the problem of difficult sampling of these posterior distributions, the Metropolis–Hastings algorithm is used. In order to sample from (10) by Metropolis–Hastings algorithm, the normal distribution $N(0, \sigma_\gamma^2\Omega_\gamma^{-1})$ is chosen as the proposal distribution, in which we should choose a suitable σ_γ^2 so that the average acceptance rate is approximately between 0.25 and 0.45 (Gelman et al. [22]), and take

$$\Omega_\gamma = \frac{1}{2} \sum_{i=1}^n \frac{e_i^2}{\exp\{z_i^T \gamma\}} z_i z_i^T + B_\gamma^{-1}.$$

The process of implementing the Metropolis–Hastings algorithm is as follows: assuming that the current value is $\gamma^{(l)}$ at the $(l + 1)$ th iteration, we should generate a new candidate value γ^* from $N(\gamma^{(l)}, \sigma_\gamma^2\Omega_\gamma^{-1})$ and then decide whether to accept it based on the following probability:

$$\min\left\{1, \frac{p(\gamma^*|Y, X, Z, U, \beta^{(l+1)}, \alpha^{(l+1)}, \rho^{(l)})}{p(\gamma^{(l)}|Y, X, Z, U, \beta^{(l+1)}, \alpha^{(l+1)}, \rho^{(l)})}\right\}.$$

In addition, according to [18], the Bayesian estimate of ρ is obtained based on a Metropolis-within-Gibbs step and by using sampling observations from (11).

Based on the observation results generated by the above calculation algorithm, the Bayesian estimation of the unknown parameter $(\beta, \alpha, \gamma, \rho)$ can be obtained. Specifically, it is assumed that the observation value $\{(\beta^{(j)}, \alpha^{(j)}, \gamma^{(j)}, \rho^{(j)}) : j = 1, 2, \dots, J\}$ of $(\beta, \alpha, \gamma, \rho)$ generated from the joint conditional distribution $p(\beta, \alpha, \gamma, \rho|Y, X, Z, U)$ is obtained by using the hybrid algorithm proposed earlier. Then define Bayesian estimators of the unknown parameters $(\beta, \alpha, \gamma, \rho)$ as follows:

$$\hat{\beta} = \frac{1}{J} \sum_{j=1}^J \beta^{(j)}, \hat{\alpha} = \frac{1}{J} \sum_{j=1}^J \alpha^{(j)}, \hat{\gamma} = \frac{1}{J} \sum_{j=1}^J \gamma^{(j)}, \hat{\rho} = \frac{1}{J} \sum_{j=1}^J \rho^{(j)}.$$

Similar to Geyer [23], it is not difficult to prove that $(\hat{\beta}, \hat{\alpha}, \hat{\gamma}, \hat{\rho})$ are consistent estimates of their corresponding posterior means. In addition, we use the “leave-one-out” technique to obtain Bayesian estimates of posterior covariance matrices. For example, the specific formula of the estimator for $\text{Var}(\beta|\theta_{\beta-}, Y, X, Z, U)$ can be expressed as follows:

$$\widehat{\text{Var}}(\beta|\theta_{\beta-}, Y, X, Z, U) = (J - 1)^{-1} \sum_{j=1}^J (\beta^{(j)} - \hat{\beta})(\beta^{(j)} - \hat{\beta})^T,$$

in which θ_{β^-} denotes the parameter θ excluding β . A similar approach can be used to obtain the estimators of $\text{Var}(\alpha|\theta_{\alpha^-}, Y, X, Z, U)$, $\text{Var}(\rho|\theta_{\rho^-}, Y, X, Z, U)$, and $\text{Var}(\gamma|\theta_{\gamma^-}, Y, X, Z, U)$.

4. Simulation Study

This section conducts a simulation study to evaluate the performance of the proposed Bayesian method under different sample sizes, spatial parameter values, and prior information selections. According to Lee [24] and Xie et al. [7], let $n = R \times m$ and $W = I_R \otimes H_m$ be the weight matrix, in which \otimes represents the Kronecker product and $H_m = (l_m l_m^T - I_m) / (m - 1)$, l_m is an m -dimensional vector with all component elements being 1. In order to investigate whether the proposed Bayesian estimation method is sensitive to the selection of prior distributions, we consider three hyperparameter values in the prior distributions of unknown parameters $\beta, \alpha, \gamma, \rho$:

Case I: $\beta_0 = (1, -0.5, 0.5)^T, b_\beta = 0.25 \times I_3, \gamma_0 = (1, -0.5, 0.5)^T, B_\gamma = 0.25 \times I_3, \alpha_0 = (0, \dots, 0)^T, a_\tau = 1, b_\tau = 1$. This situation can be considered as having good prior information.

Case II: $\beta_0 = (0, 0, 0)^T, b_\beta = I_3, \gamma_0 = (0, 0, 0)^T, B_\gamma = I_3, \alpha_0 = (0, \dots, 0)^T, a_\tau = 1, b_\tau = 1$. This hyperparameter situation is considered to have no prior information.

Case III: $\beta_0 = 3 \times (1, -0.5, 0.5)^T, b_\beta = 10 \times I_3, \gamma_0 = 3 \times (1, -0.5, 0.5)^T, B_\gamma = 10 \times I_3, \alpha_0 = (0, \dots, 0)^T, a_\tau = 1, b_\tau = 1$. This can be seen as a situation where the previous information was inaccurate.

In this section, R is selected as 25, 50, 75 and m is set to 4, and thus, n is to be 100, 200, 300. Furthermore, we generate X and Z , respectively, from the multivariate normal distribution with zero mean vector and covariance matrix $\Sigma_0 = (c_{ij})$ where $c_{ij} = 0.5^{|i-j|}, i = 1, \dots, p(q), j = 1, \dots, p(q)$. Moreover, to reflect the different spatial dependencies between response variables, spatial parameters $\rho = -0.5, 0, 0.5$ are selected to represent different spatial dependencies; $\beta = (1, -0.5, 0.5)^T, g(u_i) = 0.5 \sin(2\pi u_i)$ where u_i follows a uniform distribution $U(0, 1)$, and the structure of the variance model is $\log(\sigma_i^2) = z_i^T \gamma$ with $\gamma = (1, -0.5, 0.5)^T$.

Based on the various parameter setting environments and generated datasets mentioned above, we use the hybrid MCMC algorithm based on 100 replications to evaluate Bayesian estimations of unknown parameters under different sample sizes. Checking whether the MCMC sampler converges in algorithm implementation is an important thing. Therefore, here the estimated potential scale reduction (EPSR) value is used to diagnose whether the MCMC algorithm converges for each dataset [25]. It can be easily observed that in all the runs we are considering, the EPSR value is very close to 1 and less than 1.1 after 3000 iterations. Therefore, after discarding the first 3000 burn-in iterations, collect the observation results of the following $J = 2000$ for statistical inference. In addition, to evaluate the performance of the nonparametric function estimation, the square root of average square errors (RASE) are used here as the criterion for evaluation,

$$\text{RASE}(\hat{g}(u)) = E \left\{ \frac{1}{n} \sum_{i=1}^n [\hat{g}(u_i) - g(u_i)]^2 \right\}^{\frac{1}{2}}.$$

The simulation results are listed in Tables 1–4. Moreover, in order to directly examine the accuracy of the estimation of function $g(u)$, we plot the true value of function $g(u)$ and its estimated curve under different cases. To save space, we only list some nonparametric estimation curve results with different spatial parameters in Figures 1–3. Figures 1–3 depict the real sine curve and its estimated curve based on B-spline approximation. It is easy to observe that all estimated curves are close to the true curve, which indicates that the estimation of $g(\cdot)$ using B-splines in the mean model performs well.

Table 1. Bayesian estimation results of unknown parameters under different sample sizes and prior information when $\rho = 0.5$.

<i>Type</i>	<i>n</i>	Para.	Bias	RMS	SD
I	100	β_1	0.0045	0.0911	0.0910
		β_2	0.0084	0.1004	0.1000
		β_3	0.0079	0.0926	0.0922
		γ_1	0.0053	0.1480	0.1480
		γ_2	0.0126	0.1459	0.1454
		γ_3	0.0215	0.1540	0.1524
		ρ	0.0633	0.0975	0.0741
	200	β_1	0.0031	0.0710	0.0709
		β_2	0.0062	0.0674	0.0672
		β_3	0.0014	0.0696	0.0695
		γ_1	0.0252	0.1215	0.1188
		γ_2	0.0093	0.1294	0.1290
		γ_3	0.0075	0.1080	0.1078
		ρ	0.0149	0.0557	0.0536
	300	β_1	0.0073	0.0554	0.0550
		β_2	0.0075	0.0568	0.0563
		β_3	0.0091	0.0597	0.0590
		γ_1	0.0049	0.0841	0.0840
		γ_2	0.0051	0.1074	0.1073
		γ_3	0.0047	0.0885	0.0884
		ρ	0.0090	0.0448	0.0439
II	100	β_1	0.0153	0.0987	0.0975
		β_2	0.0284	0.1122	0.1085
		β_3	0.0282	0.1083	0.1045
		γ_1	0.0223	0.1719	0.1704
		γ_2	0.0656	0.1866	0.1747
		γ_3	0.0540	0.1864	0.1784
		ρ	0.0532	0.0979	0.0822
	200	β_1	0.0057	0.0685	0.0683
		β_2	0.0025	0.0768	0.0768
		β_3	0.0022	0.0727	0.0727
		γ_1	0.0004	0.1174	0.1174
		γ_2	0.0122	0.1401	0.1396
		γ_3	0.0008	0.1202	0.1202
		ρ	0.0320	0.0622	0.0533
	300	β_1	0.0010	0.0658	0.0658
		β_2	0.0036	0.0625	0.0624
		β_3	0.0042	0.0593	0.0591
		γ_1	0.0018	0.1085	0.1085
		γ_2	0.0077	0.1170	0.1167
		γ_3	0.0083	0.0864	0.0860
		ρ	0.0158	0.0439	0.0409

Table 1. Cont.

Type	n	Para.	Bias	RMS	SD
III	100	β_1	0.0001	0.0994	0.0994
		β_2	0.0117	0.1106	0.1099
		β_3	0.0200	0.1073	0.1054
		γ_1	0.0330	0.1788	0.1758
		γ_2	0.0059	0.1855	0.1854
		γ_3	0.0218	0.1882	0.1869
		ρ	0.0553	0.0986	0.0816
	200	β_1	0.0013	0.0686	0.0686
		β_2	0.0100	0.0780	0.0773
		β_3	0.0062	0.0731	0.0729
		γ_1	0.0217	0.1208	0.1189
		γ_2	0.0367	0.1465	0.1418
		γ_3	0.0135	0.1246	0.1238
		ρ	0.0329	0.0624	0.0531
	300	β_1	0.0052	0.0663	0.0661
		β_2	0.0014	0.0627	0.0627
		β_3	0.0014	0.0593	0.0593
		γ_1	0.0121	0.1109	0.1102
		γ_2	0.0233	0.1205	0.1182
		γ_3	0.0160	0.0878	0.0863
		ρ	0.0160	0.0440	0.0410

Table 2. Bayesian estimation results of unknown parameters under different sample sizes and prior information when $\rho = 0$.

Type	n	Para.	Bias	RMS	SD
I	100	β_1	0.0017	0.1025	0.1025
		β_2	0.0065	0.1031	0.1029
		β_3	0.0192	0.0955	0.0936
		γ_1	0.0114	0.1689	0.1685
		γ_2	0.0089	0.1785	0.1783
		γ_3	0.0020	0.1573	0.1573
		ρ	0.0092	0.0934	0.0929
	200	β_1	0.0074	0.0623	0.0618
		β_2	0.0002	0.0833	0.0833
		β_3	0.0042	0.0665	0.0663
		γ_1	0.0001	0.0995	0.0995
		γ_2	0.0050	0.1196	0.1195
		γ_3	0.0177	0.1082	0.1067
		ρ	0.0043	0.0702	0.0701
	300	β_1	0.0038	0.0551	0.0549
		β_2	0.0093	0.0590	0.0582
		β_3	0.0102	0.0442	0.0430
		γ_1	0.0004	0.0904	0.0904
		γ_2	0.0061	0.1070	0.1068
		γ_3	0.0031	0.0977	0.0977
		ρ	0.0004	0.0540	0.0540

Table 2. Cont.

Type	n	Para.	Bias	RMS	SD
II	100	β_1	0.0180	0.1001	0.0985
		β_2	0.0027	0.1205	0.1205
		β_3	0.0111	0.1139	0.1134
		γ_1	0.0456	0.1827	0.1770
		γ_2	0.0596	0.2049	0.1960
		γ_3	0.0397	0.1852	0.1809
		ρ	0.0125	0.0975	0.0967
	200	β_1	0.0000	0.0567	0.0567
		β_2	0.0041	0.0763	0.0762
		β_3	0.0044	0.0598	0.0597
		γ_1	0.0101	0.1219	0.1215
		γ_2	0.0028	0.1599	0.1598
		γ_3	0.0041	0.1352	0.1351
		ρ	0.0012	0.0680	0.0680
	300	β_1	0.0043	0.0525	0.0523
		β_2	0.0004	0.0601	0.0601
		β_3	0.0018	0.0503	0.0503
		γ_1	0.0079	0.0910	0.0907
		γ_2	0.0019	0.1130	0.1130
		γ_3	0.0021	0.1001	0.1000
		ρ	0.0032	0.0663	0.0662
III	100	β_1	0.0016	0.0991	0.0991
		β_2	0.0157	0.1237	0.1227
		β_3	0.0013	0.1135	0.1135
		γ_1	0.0091	0.1838	0.1836
		γ_2	0.0025	0.2074	0.2074
		γ_3	0.0044	0.1888	0.1888
		ρ	0.0117	0.0964	0.0957
	200	β_1	0.0070	0.0573	0.0569
		β_2	0.0033	0.0768	0.0767
		β_3	0.0003	0.0604	0.0604
		γ_1	0.0142	0.1229	0.1220
		γ_2	0.0281	0.1658	0.1634
		γ_3	0.0096	0.1387	0.1384
		ρ	0.0017	0.0679	0.0679
	300	β_1	0.0002	0.0523	0.0523
		β_2	0.0053	0.0605	0.0603
		β_3	0.0044	0.0506	0.0504
		γ_1	0.0067	0.0923	0.0921
		γ_2	0.0179	0.1156	0.1142
		γ_3	0.0071	0.1021	0.1018
		ρ	0.0035	0.0659	0.0658

Table 3. Bayesian estimation results of unknown parameters under different sample sizes and prior information when $\rho = -0.5$.

<i>Type</i>	<i>n</i>	Para.	Bias	RMS	SD
I	100	β_1	0.0184	0.1006	0.0989
		β_2	0.0087	0.1127	0.1123
		β_3	0.0051	0.1000	0.0999
		γ_1	0.0018	0.1711	0.1710
		γ_2	0.0183	0.1643	0.1633
		γ_3	0.0107	0.1699	0.1695
		ρ	0.0364	0.1197	0.1141
	200	β_1	0.0124	0.0743	0.0733
		β_2	0.0184	0.0786	0.0764
		β_3	0.0022	0.0774	0.0773
		γ_1	0.0182	0.1269	0.1256
		γ_2	0.0091	0.1236	0.1233
		γ_3	0.0024	0.1247	0.1247
		ρ	0.0077	0.0707	0.0703
	300	β_1	0.0078	0.0592	0.0587
		β_2	0.0067	0.0618	0.0614
		β_3	0.0016	0.0549	0.0549
		γ_1	0.0026	0.1008	0.1008
		γ_2	0.0074	0.1121	0.1119
		γ_3	0.0021	0.1044	0.1044
		ρ	0.0099	0.0596	0.0588
II	100	β_1	0.0023	0.1156	0.1156
		β_2	0.0072	0.1083	0.1080
		β_3	0.0073	0.0958	0.0955
		γ_1	0.0519	0.1788	0.1711
		γ_2	0.0447	0.1885	0.1831
		γ_3	0.0173	0.1759	0.1751
		ρ	0.0103	0.1074	0.1069
	200	β_1	0.0013	0.0756	0.0756
		β_2	0.0152	0.0888	0.0875
		β_3	0.0086	0.0740	0.0735
		γ_1	0.0296	0.1097	0.1056
		γ_2	0.0352	0.1359	0.1313
		γ_3	0.0275	0.1190	0.1158
		ρ	0.0039	0.0693	0.0692
	300	β_1	0.0005	0.0520	0.0520
		β_2	0.0082	0.0668	0.0663
		β_3	0.0059	0.0627	0.0625
		γ_1	0.0224	0.1046	0.1022
		γ_2	0.0230	0.1092	0.1068
		γ_3	0.0157	0.0989	0.0977
		ρ	0.0059	0.0662	0.0660

Table 3. Cont.

Type	<i>n</i>	Para.	Bias	RMS	SD
III	100	β_1	0.0143	0.1080	0.1071
		β_2	0.0106	0.1198	0.1194
		β_3	0.0110	0.1107	0.1101
		γ_1	0.0032	0.1955	0.1954
		γ_2	0.0074	0.2199	0.2198
		γ_3	0.0143	0.2213	0.2208
	200	ρ	0.0321	0.1208	0.1165
		β_1	0.0041	0.0691	0.0690
		β_2	0.0002	0.0824	0.0824
		β_3	0.0016	0.0792	0.0791
		γ_1	0.0116	0.1109	0.1103
		γ_2	0.0117	0.1396	0.1391
	300	γ_3	0.0075	0.1269	0.1267
		ρ	0.0131	0.0798	0.0787
		β_1	0.0123	0.0608	0.0595
		β_2	0.0033	0.0655	0.0654
		β_3	0.0008	0.0553	0.0553
		γ_1	0.0055	0.0996	0.0994
		γ_2	0.0042	0.1384	0.1383
		γ_3	0.0247	0.0988	0.0956
		ρ	0.0037	0.0524	0.0522

Table 4. The estimate for the nonparametric components based on RASE.

ρ	<i>n</i>	Type = I	Type = II	Type = III
0.5	100	0.0456	0.0507	0.0490
	200	0.0302	0.0381	0.0379
	300	0.0230	0.0206	0.0204
0	100	0.0557	0.0616	0.0608
	200	0.0356	0.0379	0.0378
	300	0.0221	0.0251	0.0250
−0.5	100	0.0429	0.0630	0.0534
	200	0.0342	0.0316	0.0355
	300	0.0268	0.0251	0.0214

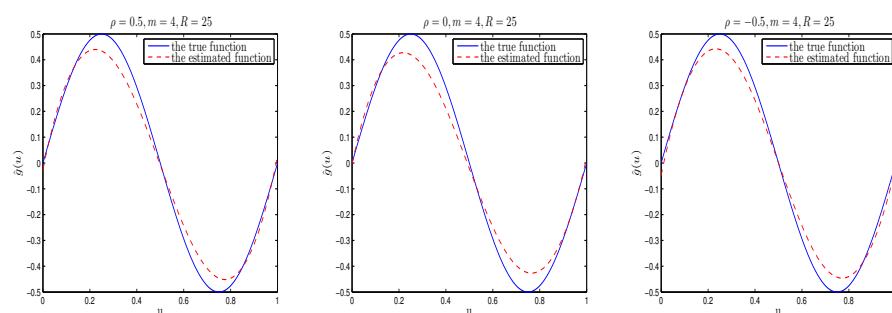


Figure 1. When $m = 4, R = 25, tpye = I$, the curve plot of the estimated function and its true values of $g(u)$ based on different ρ 's (the corresponding spatial parameters from left to right are $\rho = 0.5, 0, -0.5$).

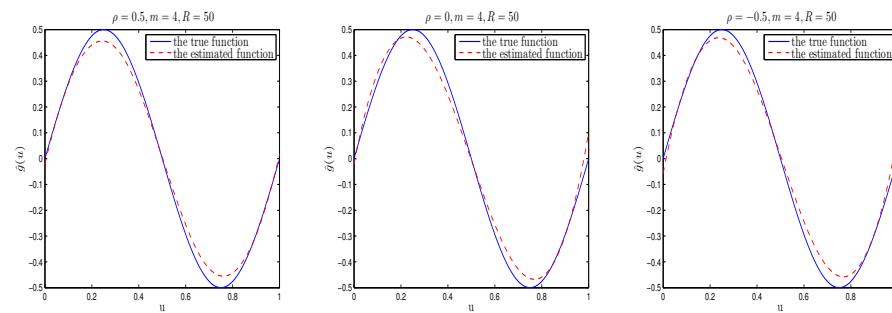


Figure 2. When $m = 4, R = 50, type = I$, the curve plot of the estimated function and its true values of $g(u)$ based on different ρ 's (the corresponding spatial parameters from left to right are $\rho = 0.5, 0, -0.5$).

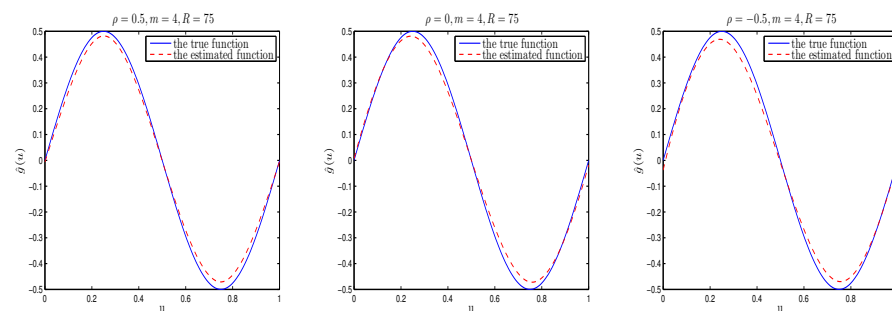


Figure 3. When $m = 4, R = 75, type = I$, the curve plot of the estimated function and its true values of $g(u)$ based on different ρ 's (the corresponding spatial parameters from left to right are $\rho = 0.5, 0, -0.5$).

In Tables 1–3, “Bias” represents the absolute difference between the true and mean values estimated by Bayesian estimation of parameters based on 100 replicates, and “SD” denotes standard deviation of the Bayesian estimates, while “RMS” is the root mean square between the estimated and true values based on 100 replicates. From Tables 1–3, we can see that (i) From the Bias, RMS and SD values of Bayesian estimation, it can be seen that regardless of the prior information input, Bayesian estimation is quite accurate; and for different prior distributions, the proposed estimation method performs well, indicating that Bayesian estimation is not sensitive to prior information input. (ii) Bayesian estimation improves as the sample size increases; (iii) The Bayesian estimation results obtained based on different spatial parameters are similar. (iv) In the same situation, the RMS and SD values of the mean parameters are smaller than that of the variance parameters, which is consistent with the fact that lower order moments are easier to estimate than higher order moments. Furthermore, from Figures 1–3, it can be seen that under the considered parameter settings, the estimated function curve is very close to its corresponding true curve, which is consistent with the phenomenon found in Table 4. In summary, the above simulation research results indicate that applying the Bayesian estimation method proposed in this paper to heterogeneous SSAR models is effective.

5. Real Data Analysis

Boston housing price data are a commonly used example, and many authors have conducted in-depth analysis based on different statistical models, such as [26,27], and so on. This section will also use the Bayesian estimation method proposed in this paper to analyze these data. This dataset can be easily obtained from R's spdep library, which includes 14 variables and 506 observations. A detailed explanation of the variables involved in the dataset is presented in Table 5.

Table 5. Detailed description of relevant variables involved in Boston housing data.

Related Variables	Detailed Description
CRIM	Per capita crime rate by town
ZN	Proportion of residential land zoned for lots over 25,000 sq.ft.
INDUS	Proportion of non-retail business acres per town
CHAS	Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
NOX	Nitric oxide concentration (parts per 10 million)
RM	Average number of rooms per dwelling
AGE	Proportion of owner-occupied units built prior to 1940
DIS	Weighted distances to five Boston employment centres
RAD	Index of accessibility to radial highways
TAX	Full-value property-tax rate per \$10,000
PTRATIO	Pupil–teacher ratio by town
B	$1000(Bk - 0.63)^2$ where Bk is the proportion of blacks by town
LSTAT	% lower status of the population
MEDV	Median value of owner-occupied homes in USD 1000's

In addition, following the variable selection results of Xie et al. [7], we take MEDV as the response variable of the model (represented by Y) and the important explanatory variables selected by Xie et al. [7] as the X - variables: CRIM (denoted by X_1), ZN (denoted by X_2), NOX (denoted by X_3), RM (denoted by X_4), DIS (denoted by X_5), RAD (denoted by X_6), TAX (denoted by X_7), PTRATIO (denoted by X_8), and B (denoted by X_9). In order to facilitate data modeling and analysis, all variables were centralized and the index variable was set to $u = \sqrt{LSTAT}$.

In addition, the Euclidean distances calculated using longitude and latitude as [27,28] are used to generate the space weight matrix $W = (w_{ij})$, where

$$w_{ij} = \max\left(1 - \frac{d_{ij}}{d_0}, 0\right),$$

d_{ij} is the Euclidean distance, and d_0 takes a value of 0.05 as the threshold distance. Thus, the spatial weight matrix contains 19.1% non-zero elements. Then, here we consider the heterogeneous SSAR model as follows:

$$\begin{cases} Y_i = \rho \sum_{j=1}^n w_{ij} Y_j + \sum_{k=1}^9 X_{ik} \beta_k + g(u_i) + \varepsilon_i, \\ \sigma_i^2 = \exp\left(\sum_{k=1}^3 Z_{ik} \gamma_k\right), \\ i = 1, 2, \dots, 506. \end{cases} \tag{8}$$

where three explanatory variables $Z_1 = X_4, Z_2 = X_5, Z_3 = X_6$ are selected in the variance model. Thus, the hybrid algorithm proposed earlier is applied to obtain Bayesian estimates of β 's, γ 's, and ρ 's, in which a B-spline with $K = 3$ and noninformative prior information are used. In order to check the convergence of the algorithm, Figure 4 shows the relationship between the EPSR values of all unknown parameters and iterations, indicating that the algorithm converges after approximately 3000 iterations due to the EPSR values of all unknown parameters being less than 1.1 in approximately 3000 iterations. The Bayesian estimates (EST) of β 's, γ 's and ρ 's and their standard deviation estimates (SD), 95% credible intervals (CI) are calculated. Results are given in Table 6. Figure 5 displays the Bayesian estimate of the nonparametric function $g(u)$, which also confirms a significant nonlinear relationship between housing prices and the variable u . Some useful conclusions can be obtained from the results of the table, which are basically consistent with the research results of other authors. For example, the regression parameter corresponding to X_1 in the mean model is estimated to be negative, indicating that housing prices will decrease as the per capita crime rate in urban areas increases. The estimated coefficient of X_4 in

the mean model is 0.3899, indicating that as the average number of rooms per dwelling increases, housing prices will also increase. $\hat{\beta}_5$ is negative, indicating that the greater the weighted distances to five Boston employment centres, the lower the housing price will be. The regression parameter corresponding to X_8 in the mean model is estimated to be negative, indicating that housing prices will decrease as the pupil–teacher ratio by town increases. The regression parameter corresponding to Z_2 in the variance model is estimated to be negative, indicating that as the weighted distances to five Boston employment centres increases, the fluctuation of housing prices will also decrease. In addition, based on the estimation of γ , we can obtain the estimated value of σ_i^2 and present the scatter plot of $\hat{\sigma}_i^2$ in Figure 6, indicating that heteroscedasticity modeling for this dataset is reasonable.

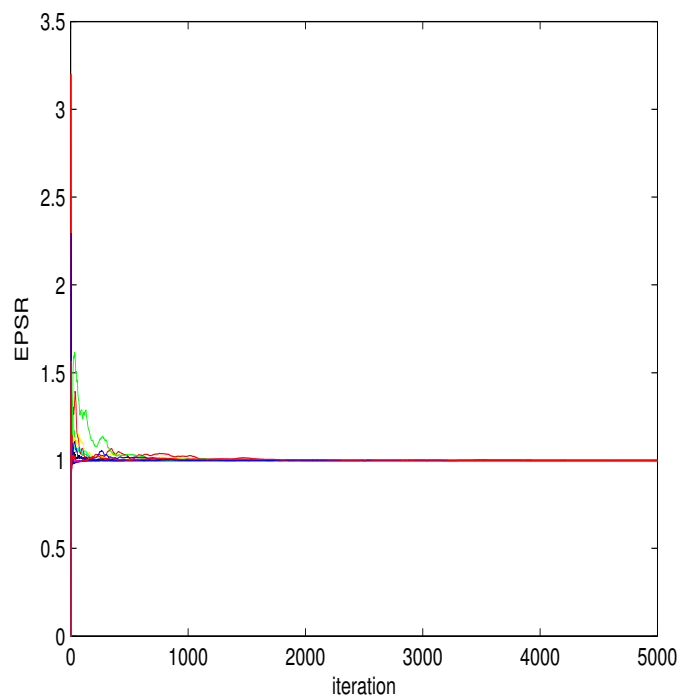


Figure 4. EPSR values for all parameters in Boston data analysis.

Table 6. Bayesian estimation results in Boston data analysis.

Parameter	EST	SD	CI
β_1	−0.1335	0.1098	(−0.3460, 0.0835)
β_2	0.0386	0.0507	(−0.0614, 0.1397)
β_3	−0.0839	0.0850	(−0.2541, 0.0803)
β_4	0.3899	0.0931	(0.2034, 0.5633)
β_5	−0.1591	0.0693	(−0.3043, −0.0305)
β_6	0.2403	0.1182	(0.0122, 0.4697)
β_7	−0.2143	0.0914	(−0.3908, −0.0346)
β_8	−0.1288	0.0530	(−0.2366, −0.0259)
β_9	0.1058	0.07521	(−0.0437, 0.2515)
γ_1	0.2132	0.1602	(−0.1234, 0.5461)
γ_2	−0.3000	0.1937	(−0.7381, 0.0715)
γ_3	0.6484	0.1974	(0.2865, 1.0527)
ρ	0.1553	0.0910	(−0.0335, 0.3302)

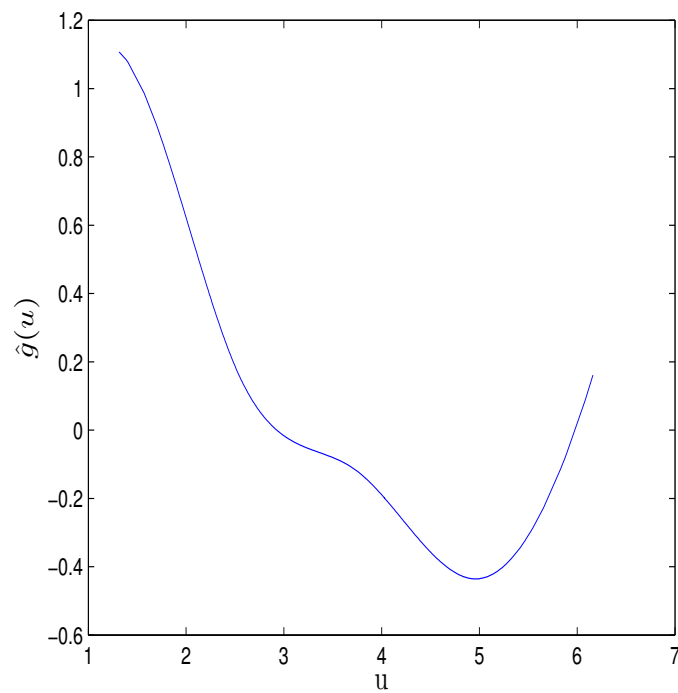


Figure 5. The solid curve is the estimated function of $\hat{g}(u)$.

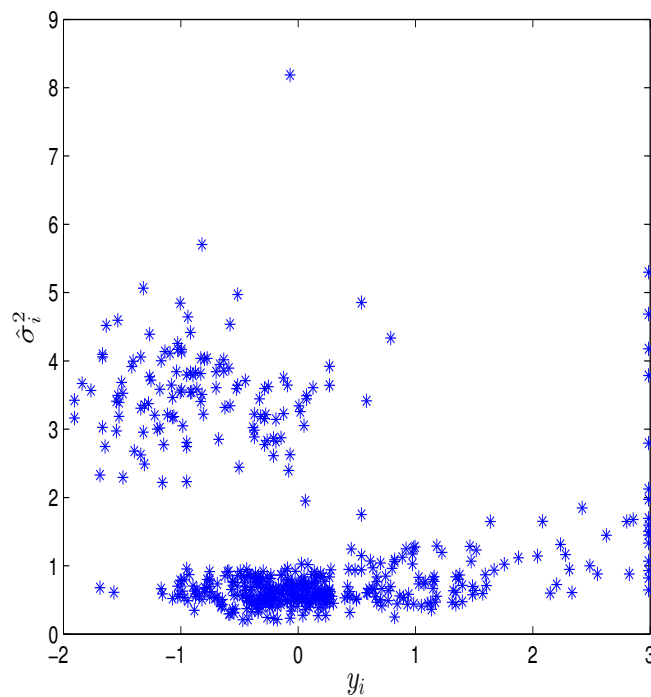


Figure 6. The scatter plot of $\hat{\sigma}_i^2$.

6. Conclusions and Discussion

Heteroscedasticity is a common phenomenon in spatial data modeling and analysis. Therefore, this paper proposes heterogeneous SSAR models, where the variance parameter is modeled as a function of the explanatory variable. Like mean regression modeling, the variance component may also depend on various explanatory variables of interest, so

estimating the joint models of the mean and variance becomes the basis for avoiding modeling bias and reducing model complexity. Then, based on the nonparametric components approximated by B-splines, we propose a complete Bayesian analysis of a heterogeneous semiparametric spatial autoregressive models. Based on the Gibbs sampler and Metropolis–Hastings algorithm, an effective MCMC sampling algorithm was developed for posterior inference by generating posterior samples from the posterior distributions. Simulation studies and real data analysis based on Boston housing data are used to illustrate the proposed method. The results show that the proposed Bayesian method has high efficiency and fast computational speed.

In addition, there are several interesting questions worth further research in the future. For example, (i) it is interesting to consider variable selection for the parametric component in the context of heterogeneous spatial autoregressive model; (ii) the model proposed in the paper is also a worthwhile issue to study when there are missing response variables.

Author Contributions: Conceptualization, D.X. and S.K.; methodology, D.X. and S.K.; software, T.L. and S.K.; data curation, T.L. and S.K.; formal analysis, T.L. and D.X.; writing—original draft, T.L. and D.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LY23A010013.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Cliff, A.; Ord, J.K. *Spatial Autocorrelation*; Pion: London, UK, 1973.
2. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, 1988.
3. Anselin, L.; Bera, A.K. *Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics*; Ullah, A., Giles, D.E.A., Eds.; Handbook of Applied Economics Statistics; Marcel Dekker: New York, NY, USA, 1998.
4. Jin, F.; Lee, L.F. GEL estimation and tests of spatial autoregressive models. *J. Econom.* **2019**, *208*, 585–612.
5. Liu, X.; Chen, J.B.; Cheng, S.L. A penalized quasi-maximum likelihood method for variable selection in the spatial autoregressive model. *Spat. Stat.* **2018**, *25*, 86–104.
6. Xie, L.; Wang, X.R.; Cheng, W.H.; Tang, T. Variable selection for spatial autoregressive models. *Commun. Stat. Theory Methods* **2021**, *50*, 1325–1340.
7. Xie, T.F.; Cao, R.Y.; Du, J. Variable selection for spatial autoregressive models with a diverging number of parameters. *Stat. Pap.* **2020**, *61*, 1125–1145.
8. Su, L.J.; Jin, S.N. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econom.* **2010**, *157*, 18–33.
9. Du, J.; Sun, X.Q.; Cao, R.Y.; Zhang, Z.Z. Statistical inference for partially linear additive spatial autoregressive models. *Spat. Stat.* **2018**, *25*, 52–67.
10. Cheng, S.L.; Chen, J.B. Estimation of partially linear single-index spatial autoregressive model. *Stat. Pap.* **2021**, *62*, 485–531.
11. Wei, C.H.; Guo, S.; Zhai, S.F. Statistical inference of partially linear varying coefficient spatial autoregressive models. *Econ. Model.* **2017**, *64*, 553–559.
12. Hu, Y.P.; Wu, S.Y.; Feng, S.Y.; Jin, J.L. Estimation in Partial Functional Linear Spatial Autoregressive Model. *Mathematics* **2020**, *8*, 1680.
13. Lin, X.; Lee, L.F. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *J. Econom.* **2010**, *157*, 34–52.
14. Dai, X.W.; Jin, L.B.; Tian, M.Z.; Shi, L. Bayesian Local Influence for Spatial Autoregressive Models with Heteroscedasticity. *Stat. Pap.* **2019**, *60*, 1423–1446.
15. Tang, N.S.; Duan, X.D. A semiparametric Bayesian approach to generalized partial linear mixed models for longitudinal data. *Comput. Stat. Data Anal.* **2012**, *56*, 4348–4365.
16. Xu, D.K.; Zhang, Z.Z. A semiparametric Bayesian approach to joint mean and variance models. *Stat. Probab. Lett.* **2013**, *83*, 1624–1631.
17. Ju, Y.Y.; Tang, N.S.; Li, X.X. Bayesian local influence analysis of skew-normal spatial dynamic panel data models. *J. Stat. Comput. Simul.* **2018**, *88*, 2342–2364.
18. Pfarrhofer, M.; Piribauer, P. Flexible shrinkage in high-dimensional Bayesian spatial autoregressive models. *Spat. Stat.* **2019**, *29*, 109–128.

19. Wang, Z.Q.; Tang, N.S. Bayesian Quantile Regression with Mixed Discrete and Nonignorable Missing Covariates. *Bayesian Anal.* **2020**, *15*, 579–604.
20. Chen, Z.Y.; Chen, J.B. Bayesian analysis of partially linear, single-index, spatial autoregressive models. *Comput. Stat.* **2022**, *37*, 327–353.
21. Zhang, D.; Wu, L.C.; Ye, K.Y.; Wang, M. Bayesian quantile semiparametric mixed-effects double regression models. *Stat. Theory Relat. Fields* **2021**, *5*, 303–315.
22. Gelman, A.; Roberts, G.O.; Gilks, W.R. Efficient metropolis jumping rules. In *Bayesian Statistics*; Oxford University Press: New York, NY, USA, 1996; Volume 5.
23. Geyer, C.J. Practical Markov Chain Monte Carlo. *Stat. Sci.* **1992**, *7*, 473–511.
24. Lee, L.F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **2004**, *72*, 1899–1925.
25. Gelman, A. *Inference and Monitoring Convergence in Markov Chain Monte Carlo in Practice*; Chapman and Hall: London, UK, 1996.
26. Pace, R.K.; Gilley, O.W. Using the spatial configuration of the data to improve estimation. *J. Real Estate Financ. Econ.* **1997**, *14*, 330–340.
27. Sun, Y.; Zhang, Y.; Huang, J.Z. Estimation of a semiparametric varying-coefficient mixed regressive spatial autoregressive model. *Econom. Stat.* **2019**, *9*, 140–155.
28. Luo, G.; Wu, M. Variable selection for semiparametric varying-coefficient spatial autoregressive models with a diverging number of parameters. *Commun. Stat. Theory Methods* **2021**, *50*, 2062–2079.

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.