

Article

# Extraction of Continuous and Discrete Spatial Heterogeneities: Fusion Model of Spatially Varying Coefficient Model and Sparse Modelling

Ryo Inoue \*  and Koichiro Den

Department of Human-Social Information Sciences, Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan; koichro.den.r2@alumni.tohoku.ac.jp

\* Correspondence: rinoue@tohoku.ac.jp

**Abstract:** Geospatial phenomena often have spatial heterogeneity, which is caused by differences in the data generation process from place to place. There are two types of spatial heterogeneity: continuous and discrete, and there has been much discussion about how to analyze one type of spatial heterogeneity. Although geospatial phenomena can have both types of spatial heterogeneities, previous studies have not sufficiently discussed how to consider these two different types of spatial heterogeneity simultaneously and how to detect them separately, which may lead to biased estimates and the wrong interpretation of geospatial phenomena. This study proposes a new approach for the analysis of spatial data with both heterogeneities by combining the eigenvector spatial filtering-based spatially varying coefficient (ESF-SVC) model, which assumes the continuous spatial heterogeneity and generalized lasso (GL) estimation, which assumes discrete spatial heterogeneity and proposes the ESF-GL-SVC model. The performance of ESF-GL-SVC was evaluated through experiments based on a Monte Carlo simulation and confirms that the ESF-GL-SVC showed better performance in estimating coefficients with both types of spatial heterogeneity than the previous two models. The application of the apartment rent data showed that the ESF-GL-SVC outputs the result with the smallest BIC value, and the estimated coefficients depict continuous and discrete spatial heterogeneity in the dataset. Reasonable coefficients were estimated using the ESF-GL-SVC, although some coefficients by ESF-SVC were not.

**Keywords:** spatial heterogeneity; eigenvector spatial filtering-based spatially varying coefficient model; generalized lasso



**Citation:** Inoue, R.; Den, K. Extraction of Continuous and Discrete Spatial Heterogeneities: Fusion Model of Spatially Varying Coefficient Model and Sparse Modelling. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 358. <https://doi.org/10.3390/ijgi11070358>

Academic Editors: Wolfgang Kainz and Huayi Wu

Received: 28 April 2022

Accepted: 19 June 2022

Published: 23 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 years, due to the proactive disclosure of data by government agencies and private companies, detailed spatial data with high spatial resolution have become available, allowing us to quantitatively understand the reality of socioeconomic activities in greater detail.

Geospatial data commonly have a property known as spatial autocorrelation, also known as the ‘first law of geography’ [1], where data from locations closer to each other have a stronger correlation. One approach for the analysis of geospatial data with spatial autocorrelation is to assume that the data generation process is common regardless of location and to build models that represent the spatial autocorrelation of dependent variables, disturbances, etc. The spatial regression models in the field of spatial econometrics (e.g., [2]), the universal kriging model in the field of spatial statistics (e.g., [3]), and the eigenvector spatial filtering (ESF) model in the field of quantitative geography [4,5] are examples of this approach.

Another approach is to assume the existence of spatial heterogeneity, which means that the data generation process differs by location. This study focuses on this approach, which can be broadly divided into two types depending on assumptions about the structure of

spatial heterogeneity [6]. One assumption is that the influence of the data formation factors varies continuously with respect to spatial location. The competing assumption is that the influence of the data formation factors differs discontinuously at certain spatial boundaries.

Models which use the first assumption, that data formation factors vary continuously, are known as spatially varying coefficient (SVC) models, and the most applied model is geographic weighted regression (GWR) [6–8]. A weighted least-squares method, which uses a distance decay function to give large weights to geospatial data in the vicinity of an analysis target point, is used to obtain point-specific estimates of coefficients that vary smoothly in space. The eigenvector spatial filtering-based spatially varying coefficient (ESF-SVC) model [9–11] is an extension of ESF [4,5], which represents the spatial correlation structure using the eigenvectors of the spatial weight matrix and estimates the coefficients that vary continuously in space. ESF-SVC is a highly useful analysis method with several advantages over GWR, such as the ability to represent the structure of spatial heterogeneity more flexibly, easier estimation, and improved applicability to large scale data [11]. It has been applied to various regional analyses [12–15].

Models which use the second assumption, that data formation factors vary discretely at certain spatial boundaries, have been used for the analysis of the geographical segmentations of the real estate market in regional sciences and the detection of point event clusters, such as geographical clusters of infectious disease outbreaks in epidemiology and criminal concentration areas in criminology. In the analysis of geographic real estate market segmentations, the scale of spatial heterogeneity in property valuations is analyzed by comparing models with different geographic tessellation, such as school districts and neighborhoods (e.g., [16,17]). In the detection of point event clusters, spatial scan statistics [18], which are representative methods developed in the field of spatial epidemiology, and agglomeration analysis [19,20], which apply the false discovery rate controlling method [21], have been used to determine whether the frequency of point events in a particular region differs from that of other regions, according to predetermined regional divisions. As both approaches predefine regional divisions, they have limitations in analyzing a series of multiple regions with spatial heterogeneity. To solve the above-mentioned limitations, the generalized lasso (GL) [22] has been applied for analyses of discrete spatial heterogeneity [23–29]. The GL is an extension of the lasso [30] by introducing  $\ell_1$  regularization to the difference between the coefficients of adjacent regions. Using a model in which each region has a coefficient that represents the difference from the overall trend, a series of regions with the same level of spatial heterogeneity can be extracted by regularizing the difference between the coefficients of neighboring regions.

The above-mentioned approaches use different assumptions on spatial heterogeneity; the former assumes the heterogeneity that varies continuously in space, while the latter assumes the heterogeneity that varies discretely at specific geographical borders. However, which assumption is adequate to represent the spatial heterogeneity of geospatial data is still unclear.

Using the real estate market analysis as an example, those who prefer to live in the city center and those who prefer to live in the suburbs evaluate real estate properties differently; the former value proximity to urban services and the latter value proximity to natural environment and broad space; the coefficients of the explanatory variables expressing proximity to those services differ by location. If there are no clear boundaries between urban and suburban areas, as is the case in Japan's metropolitan areas, it is reasonable to assume that coefficients are continuously changing in space. In addition, if there are specific neighborhoods recognized as high-end residential areas, the valuation may differ from the surrounding areas and real estate prices will show discontinuity at the border of those neighborhoods. Considering the above, there might exist both continuous and discrete spatial heterogeneity in real estate price data. However, previous studies have not sufficiently discussed how to consider these two different types of spatial heterogeneity simultaneously and how to detect them separately, which may lead to biased estimates and the wrong interpretation of geospatial phenomena.

This paper proposes a new approach for the analysis of geospatial data with both continuous and discrete spatial heterogeneity by fusing ESF-SVC and GL and evaluates its performance to verify its effectiveness through the application of geospatial data analysis.

The remainder of this paper is organized as follows. In Section 2, after the ESF-SVC and GL are outlined, the fusion model, the eigenvector spatial filtering, and generalized lasso-based spatially varying coefficient model (ESF-GL-SVC) are presented and are evaluated using simulation data. In Section 3, we apply the ESF-GL-SVC, ESF-SVC, and GL models to the residential rent data in the southwestern part of the Tokyo metropolitan area, interpret the estimated results, and discuss the effectiveness of the ESF-GL-SVC model. Finally, Section 4 concludes the study.

## 2. ESF-GL-SVC: A Model to Analyze Continuous and Discrete Spatial Heterogeneity

We first explain two previous models that deal with spatial heterogeneity, introduce the ESF-GL-SVC, which is a fusion of the two, and then represent the performance evaluation.

### 2.1. Previous Models for Spatial Heterogeneity

#### 2.1.1. Eigenvector Spatial Filtering-Based Spatially Varying Coefficient (ESF-SVC) Model

The ESF-SVC model utilizes the common statistical test for spatial autocorrelation, Moran's I, to express the spatial heterogeneity of coefficients. Here we consider the analysis that the geospatial data at location  $i$ ,  $y_i$ , is regressed on the attributes of the  $K$  types of its attributes  $x_{ij}$ . Let  $N$  denote the number of locations,  $\mathbf{y}$  denote an  $N$  by 1 vector of dependent variables,  $\mathbf{X}$  denote an  $N$  by  $K$  matrix of explanatory variables whose first column is an  $N$  by 1 vector of ones.  $\beta$  is a  $K$  by 1 coefficient vector of the regression of  $\mathbf{y}$  on  $\mathbf{X}$  and the first element is an intercept.

Let  $\mathbf{C}$  denote an  $N$  by  $N$  spatial proximity matrix between  $N$  locations,  $\mathbf{1}$  denote an  $N$  by 1 vector of ones,  $\mathbf{I}$  denote an  $N$  by  $N$  identity matrix, and  $\mathbf{M}$  denote an  $N$  by  $N$  centering matrix for  $N$  by 1 vector, which is  $\mathbf{M} = (\mathbf{I} - \mathbf{1}\mathbf{1}'/N)$ . Then the Moran's I statistics  $MC$  of dependent variables  $\mathbf{y}$  observed at  $N$  locations is:

$$MC(\mathbf{y}) = \frac{N}{\mathbf{1}'\mathbf{M}\mathbf{1}} \frac{\mathbf{y}'\mathbf{M}\mathbf{C}\mathbf{M}\mathbf{y}}{\mathbf{y}'\mathbf{M}\mathbf{y}} \quad (1)$$

The term  $\mathbf{M}\mathbf{C}\mathbf{M}$  represents the autocorrelation structure of  $\mathbf{y}$  that have spatial proximity expressed by  $\mathbf{C}$ .

Griffith [4] proposed an ESF, which utilizes the eigenvectors of matrix  $\mathbf{M}\mathbf{C}\mathbf{M}$  as explanatory variables in linear regression models to express the spatial correlation of dependent variables and remove the correlation from disturbances. The eigenvectors that correspond to large eigenvalues represent the continuous global spatial correlation patterns and the eigenvectors with small and positive eigenvalues represent the continuous local spatial correlation patterns.

Let  $\mathbf{E}$  denote an  $N$  by  $M$  matrix, which consists of the first  $M$  eigenvectors, whose eigenvalues are the largest, of matrix  $\mathbf{M}\mathbf{C}\mathbf{M}$  and whose  $i$ -th column is a  $i$ -th eigenvector  $\mathbf{e}_i$ ,  $\gamma$  denote an  $M$  by 1 coefficient vector for the explanatory matrix  $\mathbf{E}$ . The basic linear model of ESF is given by:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{E}\gamma + \varepsilon \quad (2)$$

where  $E(\varepsilon) = 0$ ,  $Var(\varepsilon) = \sigma^2\mathbf{I}$ . It is assumed that the disturbance  $\varepsilon$  have homoscedasticity and no correlation.

Griffith [9] expanded ESF to an ESF-SVC model to estimate spatially varying coefficients that represent the heterogeneity of coefficients of explanatory variables. The location-based coefficient for explanatory variables  $\mathbf{x}_k$  and  $\beta_k^{ESF}$  is modelled by:

$$\beta_k^{ESF} = \beta_k\mathbf{1} + \mathbf{E}\gamma_k \quad (3)$$

where  $\boldsymbol{\gamma}_k = (\gamma_{k1}, \dots, \gamma_{km})'$  represents the coefficients that vary by location, and the ESF-SVC model is:

$$\mathbf{y} = \sum_{k=1}^K \mathbf{x}_k \circ \boldsymbol{\beta}_k^{ESF} + \boldsymbol{\varepsilon} \quad (4)$$

where  $\circ$  denotes the Hadamard product and  $E(\boldsymbol{\varepsilon}) = 0$ ,  $Var(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$ . It assumes that the spatial autocorrelation of dependent variables is represented by the heterogeneity of  $\boldsymbol{\beta}_k^{ESF}$  and there is no spatial autocorrelation left on disturbances.

The ESF-SVC model that utilized many eigenvectors might cause the overfitting problem. One solution is to utilize only eigenvectors whose Moran's I statistics are larger than one-fourth of Moran's I statistics of the first eigenvector [31]. It turns out to be only eigenvectors that correspond to large eigenvalues and represent continuous global spatial heterogeneity are utilized to express the spatial heterogeneity of spatially varying coefficients, then their continuous local spatial heterogeneity cannot be considered in the model.

The ESF-SVC model is used in this study because it can be described as a linear regression model and has high compatibility with GL, which performs  $\ell_1$  regularization for a linear regression model.

### 2.1.2. Generalized Lasso (GL)

The lasso (least absolute shrinkage and selection operator) [30] is a most common sparse modelling method that adds  $\ell_1$  penalty terms to the objective function of estimation to obtain the sparse estimates of coefficients. It is often utilized in the variable selection. The estimation of the linear regression model by lasso is given by:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left[ \frac{1}{2} \sum_{i=1}^N \left( y_i - \sum_{j=1}^K \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^K |\beta_j| \right] \quad (5)$$

where  $\lambda$  denotes a weight for the penalty term.

The GL [22] is an expansion of lasso that adds  $\ell_1$  penalty term on the differences between 'adjacent' coefficients in addition to the  $\ell_1$  penalty term on the coefficients themselves. It can detect change points when it is applied to time-series analyses and borders when it is applied to spatial analyses. The estimation of the linear regression model with GL is represented by:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left[ \frac{1}{2} \sum_{i=1}^N \left( y_i - \sum_{j=1}^K \beta_j x_{ij} \right)^2 + \delta \lambda \sum_{j=1}^K |\beta_j| + \lambda \sum_{(i,j) \in C} |\beta_i - \beta_j| \right] \quad (6)$$

where  $\lambda$  and  $\delta$  are hyperparameters of weights on penalty terms and  $C$  is a set of pairs of adjacent coefficients.

The GL has been applied to detect discrete spatial heterogeneity. It is applied not only to linear regression analyses [22–24] but also to many types of analyses, such as the estimation of the region-specific spatial covariance function [25], the estimation of the spatial and temporal quantile function [26], and the spatial cluster detection of point events based on Poisson regression models [27–29].

### 2.2. ESF-GL-SVC

We propose an ESF-GL-SVC model to analyze continuous and discrete spatial heterogeneity simultaneously by combining the ESF-SVC and GL.

Continuous heterogeneity is represented by coefficients of ESF-SVC. The ESF-GL-SVC model utilizes the eigenvectors whose Moran's I statistics are larger than one-fourth of Moran's I statistics of the first eigenvector [31]. Discrete heterogeneity is represented by coefficients of dummy variables  $\boldsymbol{\alpha}$  that are set for all subregions in the study area. Let



$\mathbf{D}$  denote the subregion dummy variable matrix; the element  $D_{ir}$  is one if point  $i$  is in subregion  $r$  and zero otherwise. Then, the ESF-GL-SVC model is expressed as:

$$\mathbf{y} = \sum_{k=1}^K \mathbf{x}_k \circ \boldsymbol{\beta}_k^{ESF} + \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\beta}_k^{ESF} = \beta_k \mathbf{1} + \mathbf{E}\boldsymbol{\gamma}_k \quad (7)$$

The estimation of the ESF-GL-SVC model with the GL regularization term can be represented by Equation (8), where  $C$  denotes the set of combinations of adjacent subregions, and  $\lambda$ ,  $\delta_1$ , and  $\delta_2$  are the hyperparameters.

$$\left( \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}} \right) = \underset{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}}{\operatorname{argmin}} \left[ \frac{1}{2} \sum_{i=1}^N \left( y_i - \sum_{j=1}^K \left( \beta_j + \sum_{l=1}^L e_{il} \gamma_{lj} \right) x_{ij} - \sum_{m=1}^M d_{im} \alpha_m \right)^2 + \lambda \sum_{(p,q) \in C} |\alpha_p - \alpha_q| + \delta_1 \lambda \sum_{m=1}^M |\alpha_m| + \delta_2 \lambda \sum_{l=1}^L \sum_{j=1}^K |\gamma_{lj}| \right] \quad (8)$$

The second term, the  $\ell_1$  penalty term on the differences between coefficients of adjacent subregions, enables the extraction of a series of subregions with the same level of spatial discrete heterogeneity; it can mitigate the scale issues associated with the preset segmentation of subregions. The third term, the  $\ell_1$  penalty term on the coefficients of subregion dummy variables, and the fourth term, the  $\ell_1$  penalty term on the ESF-SVC coefficients, have the effect of inducing a sparse solution. When data have only continuous spatial heterogeneity, the third term functions to lead to an estimation result where the values of the coefficients of subregion dummies are zero, and when data have only discrete spatial heterogeneity, the fourth term functions to lead to an estimation result where the values of the ESF-SVC coefficients are zero. The regularization for the coefficients themselves can reduce the occurrence of parameter identification problems between the ESF-SVC and the subregion dummy variable coefficients.

Three hyperparameters should be selected according to the fitness of the models. The Bayesian information criterion (BIC) might be an option. However, since the estimated coefficients would be biased due to the regularization terms of lasso [32], the model selection utilized by the estimates of Equation (8) is not appropriate. Therefore, we proposed the estimation of coefficients first by Equation (8), building a model that corresponds to the estimated results by removing coefficients estimated as zero and setting a common dummy subregion coefficient if adjacent subregions have the same estimates, estimate the model without regularization terms, and evaluating the results by BIC. These procedures avoid biased estimates, and the fitness of model would be evaluated appropriately. The estimation of the ESF-GL-SVC model is executed by the ‘genlasso’ package [33] in R. This package estimates coefficients by gradually varying hyperparameter  $\lambda$  under a fixed ratio of weights  $\delta$ . This study set both  $\delta_1$  and  $\delta_2$  in the estimation of fusion model and  $\delta$  for the estimation of GL as {0.1, 1, and 10} and searched for the estimation result with the minimum BIC value. The BIC-based evaluation may also be useful in reducing the possibility of parameter identification problems by selecting models with fewer non-zero coefficients.

### 2.3. Performance Evaluation by Simulation Experiments

We evaluated the performance of ESF-GL-SVC applied to simulated data with continuous and discrete spatial heterogeneity. The settings for the simulation data generation are outlined below.

We set a square with side length 1 as a study area and generated a predetermined number of points in it. The coordinates of points are generated by the uniform distribution between 0 and 1. Simulated data were generated from:

$$\mathbf{y} = \mathbf{1} \circ \boldsymbol{\beta}_1^{ESF} + \mathbf{x} \circ \boldsymbol{\beta}_2^{ESF} + \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\beta}_k^{ESF} = \beta_k \mathbf{1} + \mathbf{E}\boldsymbol{\gamma}_k, \quad \boldsymbol{\varepsilon} \sim N(0, \sigma_\varepsilon^2 \mathbf{I}) \quad (9)$$

where  $\mathbf{x}$  is a vector of explanatory variables whose elements are generated from the uniform distribution between 0 and 1.

To generate the data values with continuous spatial heterogeneity at each point, we set a spatial proximity matrix **C** by the gaussian kernel whose width is 0.2, obtained eigenvectors of the matrix **MCM** by the approximate calculation by [34], selected eigenvectors whose eigenvalues were larger than one-fourth of the largest eigenvalues according to [31], and generated simulation data at each point by setting coefficients  $\beta$ s and  $\gamma$ s. To simulate spatially varying coefficients with different structures of spatial heterogeneity for each trial, one eigenvector was randomly selected for each explanatory variable. The coefficient  $\gamma$  of selected eigenvectors was set to one and other coefficients were set to zero. The non-spatially varying coefficients,  $\beta_1$  and  $\beta_2$ , were set to one.

To generate the data value with discrete spatial heterogeneity, we divided the study area into 10 by 10 square subregions with side length 0.1 and assigned a dummy variable to each subregion. Let  $\alpha$  denote a vector of subregion dummy variables and **D** denote a matrix to assign subregion dummy variables to each point. The coefficients of dummy variables of the four subregions in the center of study area in fifth and sixth rows and fifth and sixth columns were set to one and the other coefficients were set to zero.

Then the simulated value at each point was generated adding disturbances independent and identically distributed according to a normal distribution.

Table 1 summarizes the simulation data generation settings for the following three experiments. The first two experiments evaluate the performance of ESF-GL-SVC by changing the amount of data and the size of variance of disturbances. The last experiment compares the performance of ESF-GL-SVC with those of ESF-SVC and GL. Simulation data are generated 1000 times for all experiments. The codes for the simulation experiments are available in the Supplementary Materials.

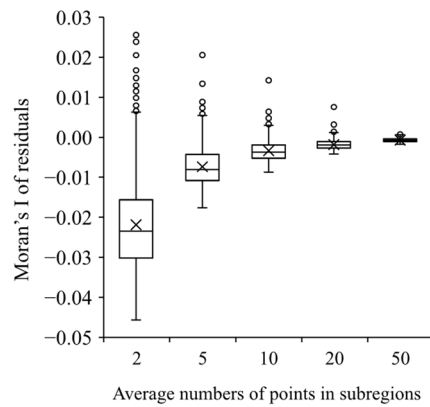
**Table 1.** Settings for simulation data generation.

		Effect of Amount of Data	Effect of Magnitude of Variances of Disturbances	Performance Comparison with Previous Models
$\beta_1^{ESF}$	$\beta_1$ Coefficients of eigenvectors	A randomly selected coefficient of eigenvector is set to 1.0 and other coefficients are set to zero.		1.0
$\beta_2^{ESF}$	$\beta_2$ Coefficients of eigenvectors	A randomly selected coefficient of eigenvector is set to 1.0 and other coefficients are set to zero.		1.0
$\alpha$	Coefficients of subregion dummies	$\alpha_{45} = \alpha_{46} = \alpha_{55} = \alpha_{56} = 1.0$ and other coefficients are set to zero.		
Other settings	Standard deviation of disturbances	0.1	0.1, 0.2, 0.4, 0.8 and 1.2	0.1
	Average number of points in a subregion	2, 5, 10, 20, and 50	20	20
	Bandwidth of kernel density function to build proximity matrix <b>C</b>		0.2	

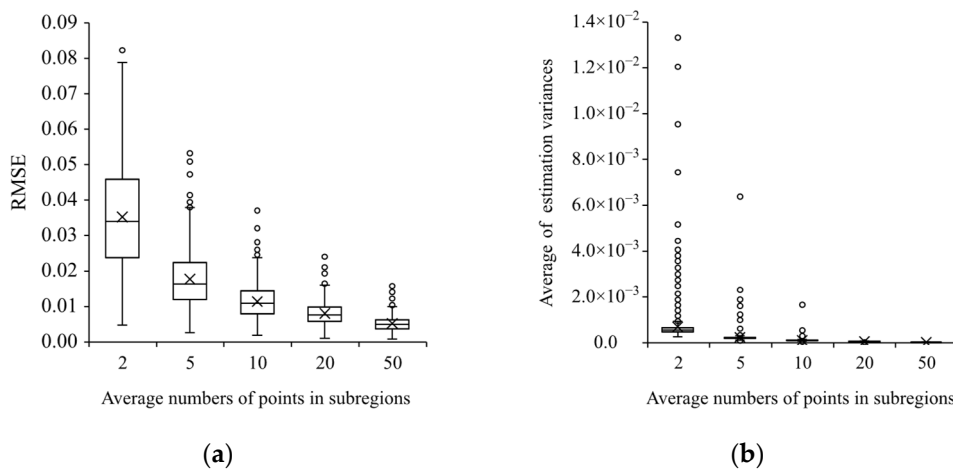
### 2.3.1. Effect of Amount of Data on Performance of ESF-GL-SVC

Five different settings were set for the total number of points to check the effect on model estimation. When generating data whose average numbers of points in each subregion were set to two, five, and ten, the numbers of points in each subregion were controlled to be the same and the position of points were randomly set inside each subregion. When generating the data by other settings, points were randomly distributed in the whole area and then the number of points in each subregion was not the same.

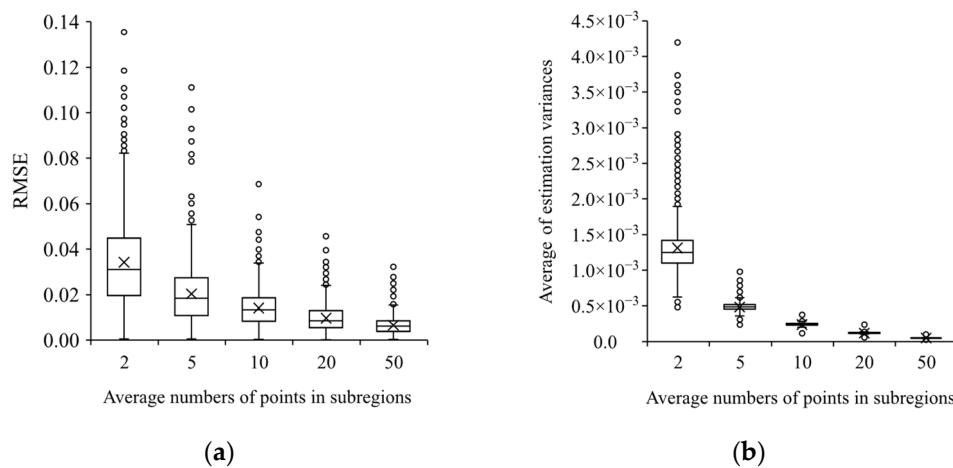
Figure 1 shows the relationship between amount of data and Moran’s I statistics of residuals, and Figures 2 and 3 show the root mean square error (RMSE) between simulated and estimated coefficients and the average of estimation variances of coefficients of the intercept  $\beta_1^{ESF}$  and the explanatory variable  $\beta_2^{ESF}$ , respectively.



**Figure 1.** Average numbers of points in subregions and Moran’s I statistics of residuals.



**Figure 2.** Average numbers of points in subregions and estimated  $\beta_1^{ESF}$ . (a) RMSEs between simulated and estimated coefficients; (b) average of estimation variances.



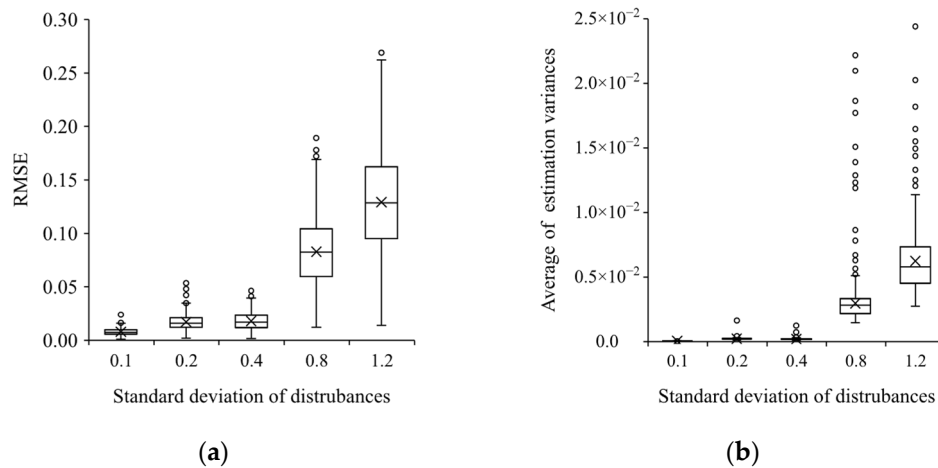
**Figure 3.** Average numbers of points in subregions and estimated  $\beta_2^{ESF}$ . (a) RMSEs between simulated and estimated coefficients; (b) average of estimation variances.

The precision of estimation increases as the number of points increases. When the numbers of points in each subregion are two and five, the RMSEs between simulated and estimated coefficients and estimation variances are quite large; however, when the number of points in each subregion exceeded 10, it was confirmed that the estimation precision was high and the spatial correlation in residuals were removed. The experiment suggests that

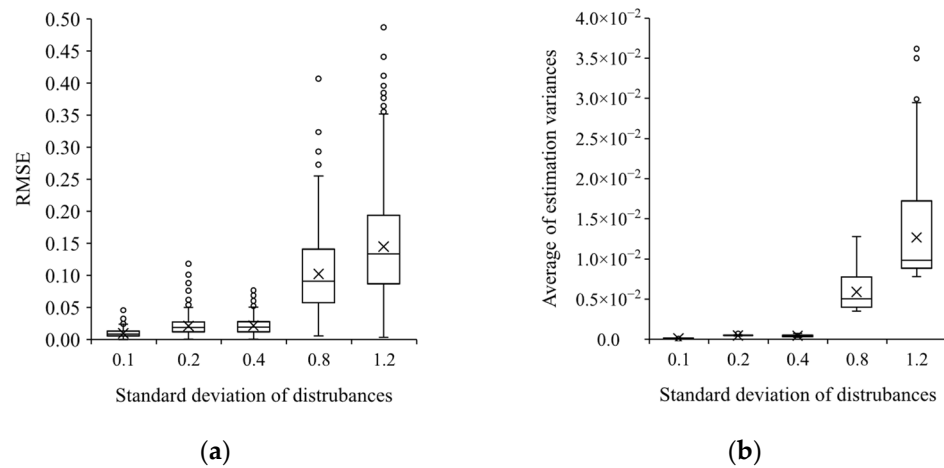
the estimated results are stabilized when the number of points in each subregion is larger than 10.

### 2.3.2. Effect of Magnitude of Variances of Disturbances on Performance of ESF-GL-SVC

Five different settings for the standard deviations of disturbances were tested to determine their effect on model estimation. Figures 4 and 5 show the RMSEs between the simulated and estimated coefficients and the estimation variances of coefficients of the intercept  $\beta_1^{ESF}$  and the explanatory variables  $\beta_2^{ESF}$ , respectively.



**Figure 4.** Variances of disturbances and estimated  $\beta_1^{ESF}$ . (a) RMSEs between simulated and estimated coefficients; (b) average of estimation variances.



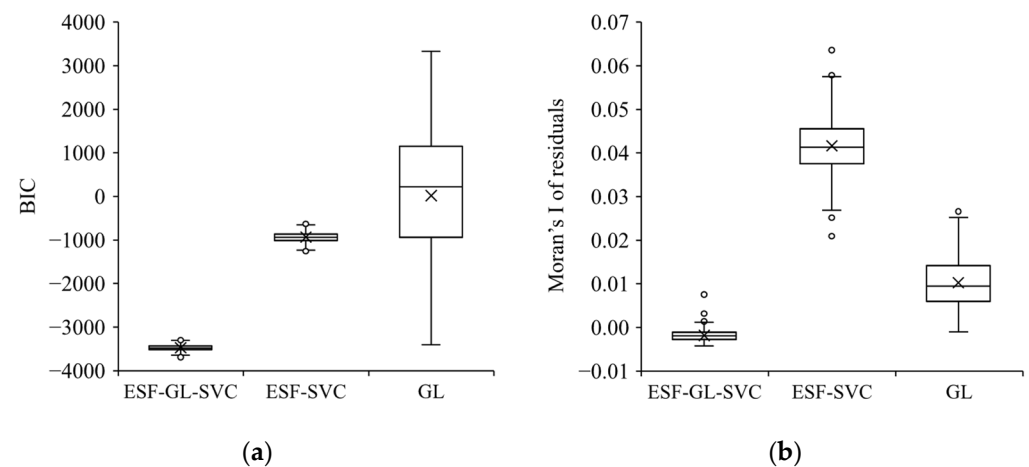
**Figure 5.** Variances of disturbances and estimated  $\beta_2^{ESF}$ . (a) RMSEs between simulated and estimated coefficients; (b) average of estimation variances.

When the standard deviations of disturbances were set to 0.8 and 1.2, the spatially varying coefficients were not estimated properly. The RMSEs between simulated and estimated coefficients and the estimation variances of coefficients were large at these settings. Considering that the values of  $1 \circ \beta_1^{ESF} + x \circ \beta_2^{ESF} + D\alpha$ , the simulated values before adding disturbances, had a standard deviation of 0.710 in the average of simulations, it is confirmed that the spatial heterogeneity of coefficients cannot be specified if the standard deviation of disturbances exceeds that of the effect of explained variables.

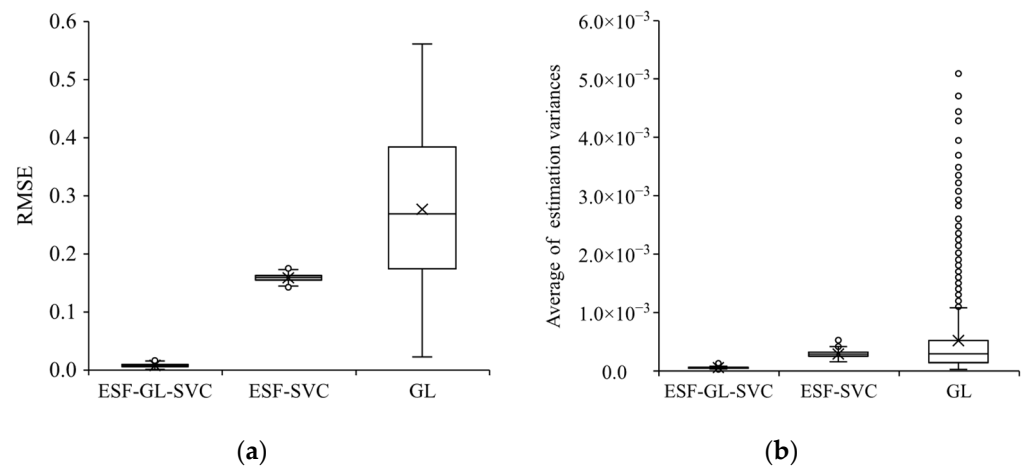
### 2.3.3. Comparison with Previous Models

We estimated the coefficients of ESF-GL-SVC, ESF-SVC, and GL models and evaluated their performances by the BIC calculated by the model without regularization terms, Moran's I statistics of residuals, and the RMSEs between simulated and estimated coefficients of each location. The plot of spatial distribution of estimated coefficients were also considered for the evaluation.

The distributions of BIC and Moran's I statistics of residuals for data with continuous and discrete spatial heterogeneity are shown in Figure 6, the distribution of RMSEs between simulated and estimated coefficients and the average of estimation variances of coefficients and are shown in Figures 7 and 8, respectively, and the spatial distributions of simulated and estimated coefficients by three models at one simulation are shown in Figure 9.



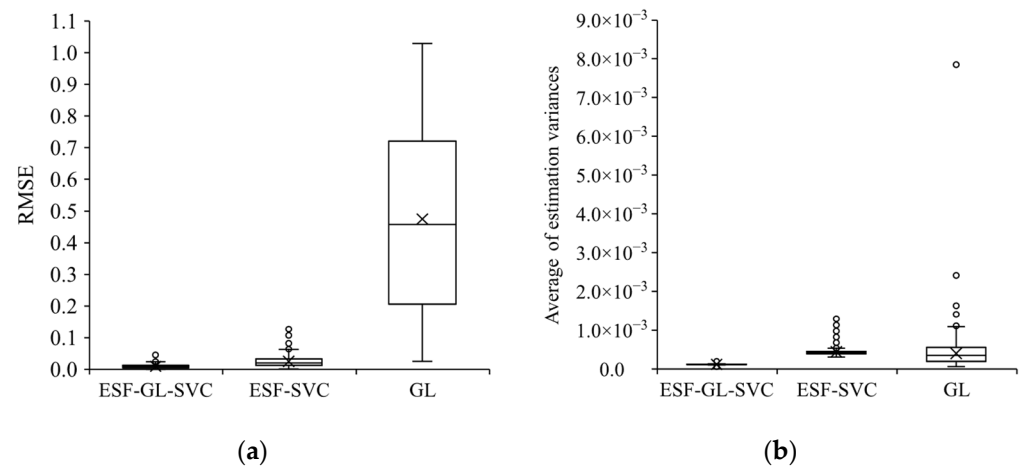
**Figure 6.** BIC and Moran's I statistics of residuals. (a) BIC; (b) Moran's I of residuals.



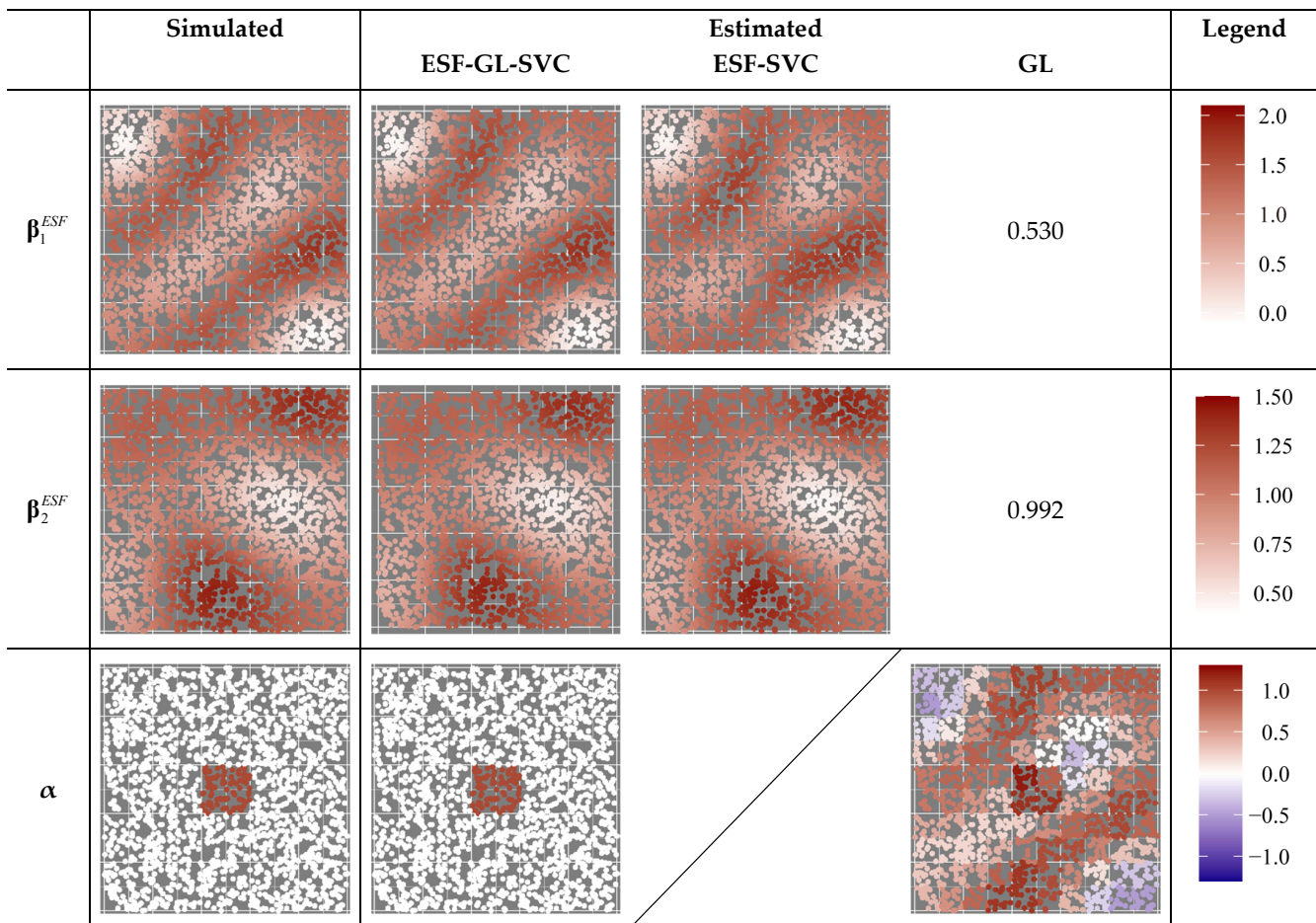
**Figure 7.** Estimated  $\beta_1^{ESF}$ . (a) RMSEs between simulated and estimated coefficients; (b) average of estimation variances.

The distribution of BICs and Moran's I statistics of residuals reveals that the ESF-GL-SVC outputs the results with the smallest BIC, removing the spatial autocorrelation from the residuals. The differences between the estimators by the ESF-GL-SVC and the simulated coefficients are very small, representing the continuous heterogeneity by the estimates of spatially varying coefficients and the discrete heterogeneity by the estimates of subregion dummy coefficients. On the other hand, the ESF-SVC and GL models failed to remove the spatial autocorrelation from the residuals, and the estimators were different from the simulated coefficients with large variances.





**Figure 8.** Estimated  $\beta_2^{ESF}$ . (a) RMSEs between simulated and estimated coefficients; (b) average of estimation variances.



**Figure 9.** Spatial distribution of simulated and estimated coefficients at one simulation.

Figure 9 indicates that the ESF-SVC represents the discrete heterogeneity by the continuous spatially varying coefficients and the GL represents the continuous heterogeneity by the discrete subregion-based coefficients. Even though the coefficients of subregion dummy variables of the proposed model could potentially be used to represent continuous spatial heterogeneity, only discrete heterogeneity is extracted in this estimation by them. This may indicate that the regularization for the coefficients of the subregion dummy variables and

the model selection based on BIC are effective in avoiding parameter identification problems between continuous and discrete spatial heterogeneity. If both continuous and discrete spatial heterogeneity exists, they would fail to extract the structure of spatial heterogeneity.

The experiments confirmed that the ESF-GL-SVC model showed better performance than the others for the data with continuous and discrete spatial heterogeneity.

### 3. Application to Apartment Rent Data

#### 3.1. Data and Models

We applied the ESF-GL-SVC, ESF-SVC, and GL models to the apartment rent data in the Shibuya, Setagaya, and Meguro wards in the southwestern part of the Tokyo Metropolitan area in 2017. The data was collected by 'At Home Co., Ltd. '; the company publishes real estate price information.

High-rise condominiums whose number of floors exceed 15 were excluded as their rents have a different pricing structure from the pricing of other apartments. If the rent data from one building make up most data for that neighborhood, it might cause difficulties in estimating coefficients because the most explanatory variables have the same or similar values. Thus, to avoid such issues, only one apartment was randomly selected from each floor of each building. Consequently, the total number of records was 13,748.

The dependent variable is the logarithm of rent per square meter in Japanese Yen (JPY), and explanatory variables are the logarithm of 'floor level,' 'building age,' 'property size,' 'walking time to the nearest train station' (hereafter, 'time to train service'), and 'average of travel time by train service to five major stations located in central business districts (CBD)' (hereafter, 'time to CBD'). As there are properties whose building ages are zero, we added one when we calculated the logarithm of building age. The selected major stations were Shinjuku, Ikebukuro, Shibuya, Tokyo, and Shinagawa, which serve the largest numbers of passengers in the Tokyo Metropolitan area. We surveyed the travel time that includes the half of train headway to five major stations when leaving the station at noon on weekday on the public transport route planning service of Yahoo! Transit and calculated the weighted average by the numbers of passengers at major stations. Table 2 summarizes the descriptive statistics of dependent and explanatory variables. Since the scales of explanatory variables affect the regularization terms on the estimation of each coefficient, the explanatory variables were standardized to zero mean and unit variance.

**Table 2.** Descriptive statistics of dependent and independent variables.

	Average	Maximum	Minimum	Standard Deviation
<b>Rent (JPY/month)</b>	121,000	2,000,000	10,000	94,100
<b>Rent per square meter (JPY/square meter/month)</b>	3340	35,900	505	759
<b>Floor level (floors)</b>	2.59	14.0	1.00	1.59
<b>Building age (years)</b>	20.2	35.0	0.00	9.30
<b>Property size (m<sup>2</sup>)</b>	37.2	445	10.0	24.2
<b>Travel time to train service (min)</b>	7.33	28.0	1.00	4.10
<b>Travel time to CBD (min)</b>	24.8	37.4	9.64	5.39

To represent continuous heterogeneity by spatially varying coefficients in the ESF-GL-SVC and ESF-SVC models, the elements of spatial proximity matrix **C** were set by inputting distances between properties into the gaussian kernel function with the band at 500 m intervals from 1 km to 5 km. For each band setting, the eigenvectors of **MCM**, whose eigenvalues are larger than the one-fourth of the largest eigenvalue, were selected as the explanatory variables in the two models. In this application, we selected the band of 4.5 km for the ESF-GL-SVC and 2 km for the ESF-SVC that output the smallest BIC by each model. In the GL estimation, ward-based coefficients for explanatory variables were estimated to represent the heterogeneity of valuation in global scale.

To represent discrete spatial heterogeneity by subregion dummies in the ESF-GL-SVC and GL models, the study region was divided into 445 neighborhoods (*cho* and *cho-me* in Japanese) and at least one real estate property existed in 431 neighborhoods. A subregion dummy was set for each neighborhood except the reference neighborhood, *Hachimanyama 3 cho-me*. The selected reference neighborhood is the neighborhood where the root mean square of residuals by ESF-SVC is the minimum. The adjacency of the coefficients of subregion dummies is set when neighborhood polygons share borders.

This analysis used the same search process of the best hyperparameter setting for each model as the simulation experiments;  $\delta_1$  and  $\delta_2$  in the ESF-GL-SVC model and  $\delta$  in the GL are set to {0.1, 1, 10}, and the estimation was performed by changing  $\lambda$  using the R package of ‘genlasso’. The results with the  $\ell_1$  regularization were used only to select variables whose estimators are non-zero; then, BICs were calculated via OLS estimation of models with selected variables. The code and sample data for the analysis are available in the Supplementary Materials.

### 3.2. Estimation Results

Table 3 summarizes the estimated results of the three models. First, all three models effectively consider the spatial autocorrelation of dependent variables, as Moran’s I statistics of residuals indicate that residuals do not have spatial autocorrelation. Second, it is confirmed that the ESF-GL-SVC model shows the best performance as the coefficients of determination is the maximum and the BIC is the minimum.

**Table 3.** Estimated results of three models.

	Selected Hyperparameters	Band for Spatial Proximity Matrix C (meters)	Number of Utilized Eigenvectors	BIC	Adjusted Coefficient of Determination	Moran’s I of Residuals
ESF-GL-SVC	$\{\delta_1, \delta_2, \lambda\} = \{1, 1, 0.427\}$	4500	5	−18.015	0.6540	−0.00016
ESF-SVC	$\{\lambda\} = \{0.576\}$	2000	11	−17.290	0.6278	0.00131
GL	$\{\delta, \lambda\} = \{1, 0.652\}$			−17.122	0.6280	0.00171

The estimated intercepts and coefficients of explanatory variables by the ESF-GL-SVC and ESF-SVC models are summarized in Tables 4 and 5, respectively, and the estimated ward-based coefficients by GL is summarized in Table 6. The spatial distribution of intercepts is shown in Figure 10, the spatial distributions of the subregion coefficients by the ESF-GL-SVC and GL model are shown in Figure 11, and the spatial distributions of coefficients of explanatory variables are shown in Figures 12–16.

**Table 4.** Statistics of estimated coefficients by ESF-GL-SVC.

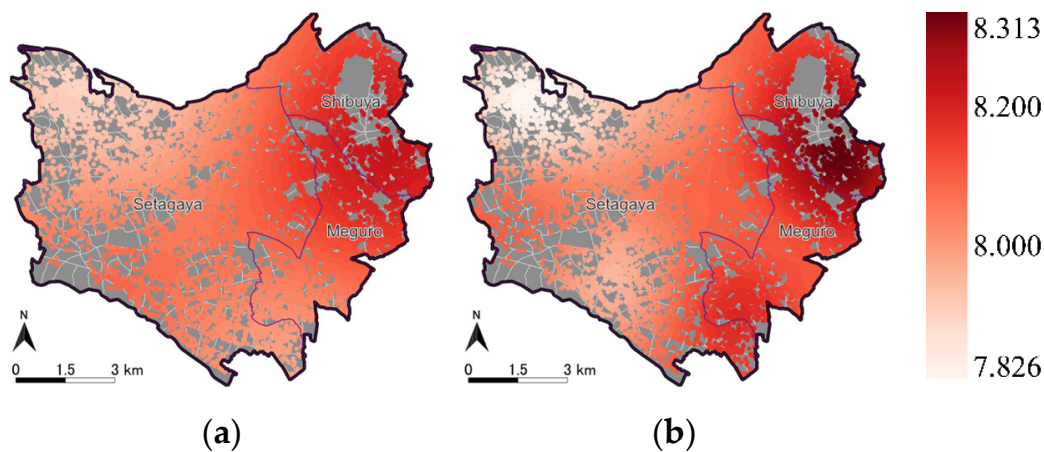
	Average	Maximum	Minimum	Standard Deviation
<b>Intercept</b>	8.077	8.210	7.915	0.07134
<b>Floor level</b>	0.02577	0.03185	0.02089	0.00231
<b>Building age</b>	−0.05642	−0.04493	−0.07519	0.00729
<b>Property size</b>	−0.08240	−0.01915	−0.13630	0.02801
<b>Time to train service</b>	−0.02520	−0.00586	−0.05727	0.01316
<b>Time to CBD</b>	−0.03331	−0.00823	−0.04780	0.00921

**Table 5.** Statistics of estimated coefficients by ESF-SVC.

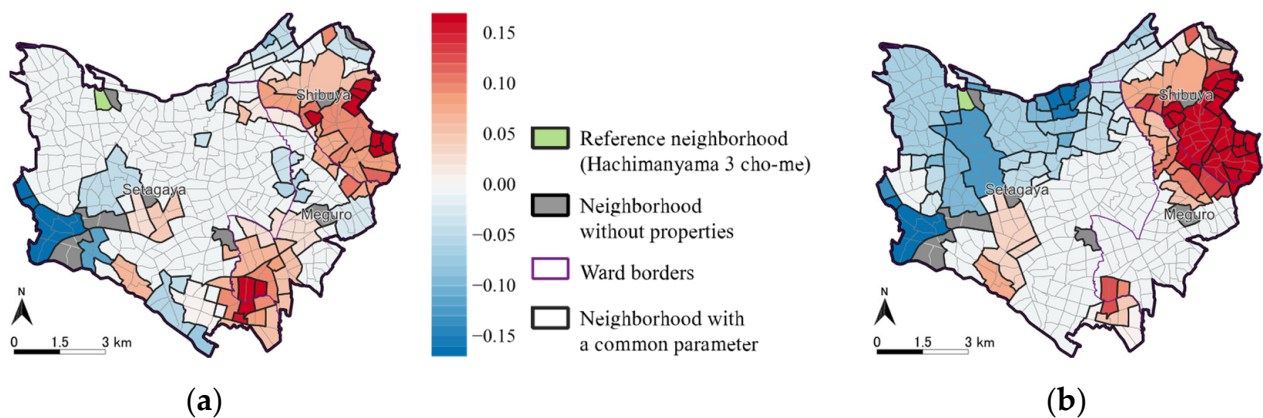
	Average	Maximum	Minimum	Standard Deviation
Intercept	8.088	8.313	7.827	0.09753
Floor level	0.02512	0.03560	0.01374	0.00461
Building age	−0.05684	−0.04557	−0.07106	0.00598
Property size	−0.08165	0.01688	−0.13348	0.03389
Time to train service	−0.02661	−0.00061	−0.05945	0.01088
Time to CBD	−0.04206	0.04879	−0.16013	0.04536

**Table 6.** Estimated ward-based coefficients by GL.

Ward	Shibuya	Meguro	Setagaya
Intercept	8.100 (Common for Three Wards)		
Floor level	0.02838	0.02513	0.02519
Building age	−0.05068	−0.05581	−0.05890
Property size	−0.03989	−0.07641	−0.10063
Time to train service	−0.02424	−0.01816	−0.03097
Time to CBD	−0.03093	−0.04049	−0.08261



**Figure 10.** Estimated intercepts. (a) ESF-GL-SVC; (b) ESF-SVC.



**Figure 11.** Estimated coefficients for subregion dummies. (a) ESF-GL-SVC; (b) GL.



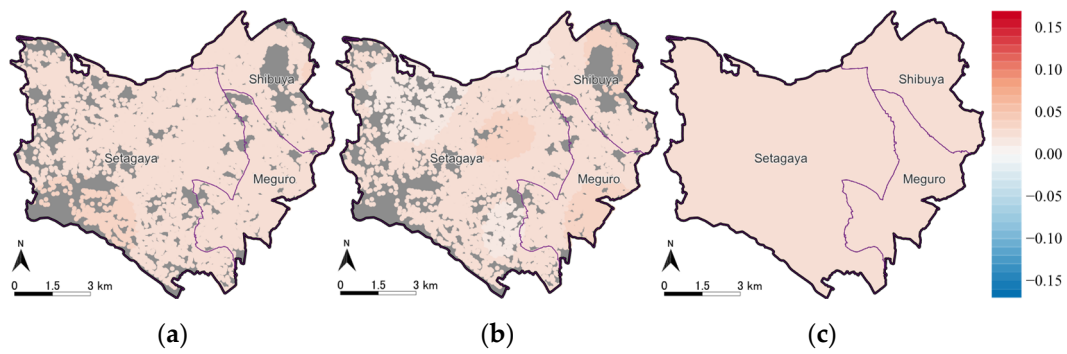


Figure 12. Estimated coefficients for floor level. (a) ESF-GL-SVC; (b) ESF-SVC; (c) GL.

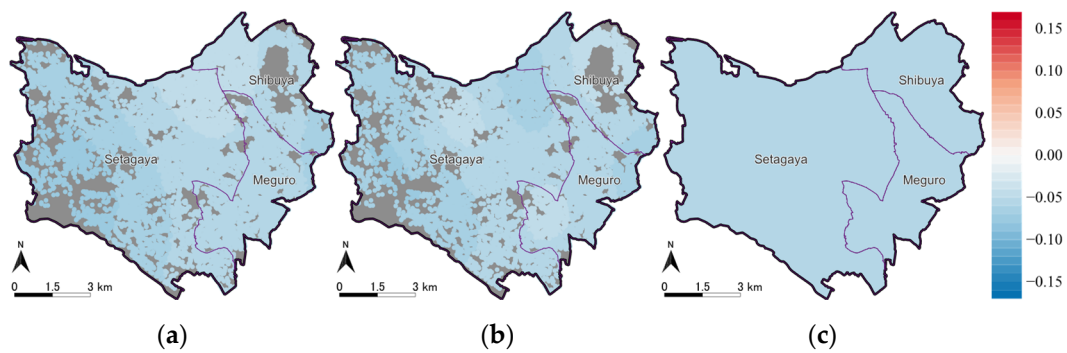


Figure 13. Estimated coefficients for building age. (a) ESF-GL-SVC; (b) ESF-SVC; (c) GL.

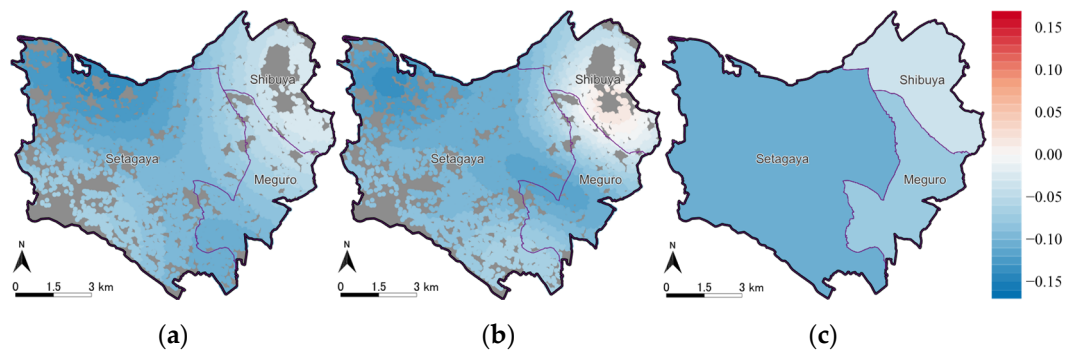


Figure 14. Estimated coefficients for property size. (a) ESF-GL-SVC; (b) ESF-SVC; (c) GL.

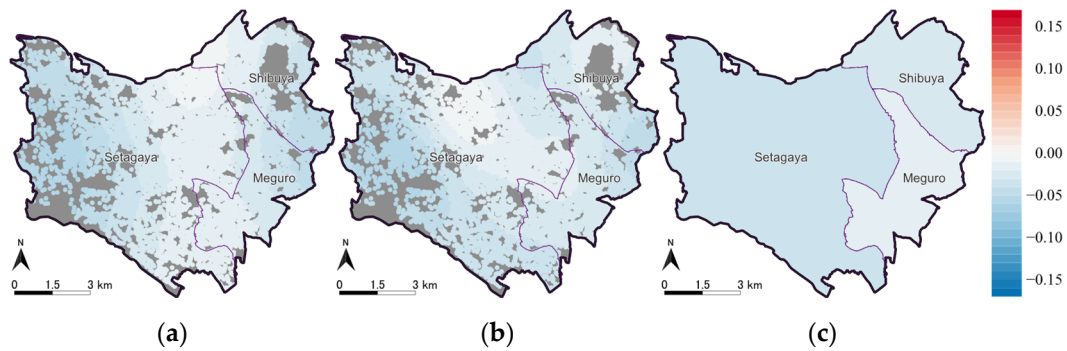
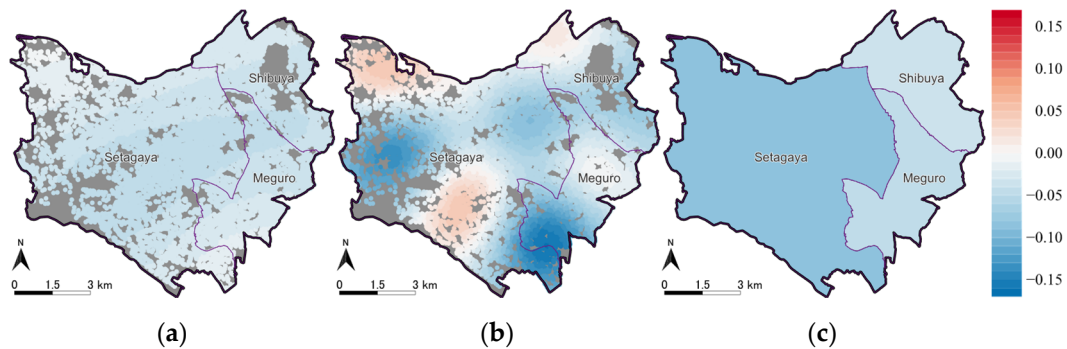


Figure 15. Estimated coefficients for time to train service. (a) ESF-GL-SVC; (b) ESF-SVC; (c) GL.





**Figure 16.** Estimated coefficients for time to CBD. (a) ESF-GL-SVC; (b) ESF-SVC; (c) GL.

### 3.3. Discussion through Comparison of Results by Three Models

#### 3.3.1. Intercepts and Coefficients of Subregion Dummies

First, we focused on the intercepts by the ESF-GL-SVC and ESF-SVC models (Figure 10) and subregion dummies using the ESF-GL-SVC model (Figure 11a). The sum of estimated results of intercepts and subregion dummies by the ESF-GL-SVC model and the estimated results of intercepts by ESF-SVC model have higher values in the eastern part of the study area, which is closer to the CBD, indicating that the rents are expensive in the area. However, the maximum intercept shown in Table 4 is smaller than that in Table 5, as the discrete spatial heterogeneity is represented by the coefficients of the subregion dummies by the ESF-GL-SVC model.

The south of Meguro ward is a well-known high class residential area; the ESF-GL-SVC model extracted the discrete spatial heterogeneity represented by positive subregion coefficients and the ESF-SVC model extracted the continuous spatial heterogeneity by larger values of spatially varying intercept. In the next section, we will comment on which identification of spatial heterogeneity is appropriate with the results of the estimation of the coefficients of the explanatory variables.

Second, we compared the estimates by the GL with those by the two other models. The estimated intercept by GL is close to the averages of the coefficients by the ESF-GL-SVC and ESF-SVC models. The number of subregions with non-zero coefficients by the GL is larger than that by the ESF-GL-SVC model; the GL model represents the spatial heterogeneity of rent by the subregion-based heterogeneity (Figure 11b). The result of GL indicates similar spatial patterns to those of the ESF-GL-SVC and ESF-SVC models; however, the GL model has higher BIC and lower adjusted coefficient of determination than the ESF-GL-SVC model; it seems that the discrete spatial heterogeneity only is not appropriate for the spatial heterogeneity representation of rent.

#### 3.3.2. Coefficients of Explanatory Variables

First, it is confirmed that all three models can distinguish whether the coefficients of the explanatory variables have spatial heterogeneity. The spatial patterns of estimates reveal that the coefficients of 'floor level' and 'building age' have weak heterogeneity, but the coefficients of 'property size', 'time to train service', and 'time to CBD' have strong heterogeneity.

Second, the estimates by the ESF-GL-SVC and ESF-SVC models are similar, except for the coefficients of 'time to CBD'. The estimates for 'property size' and 'time to train service' are almost the same for the ESF-GL-SVC and ESF-SVC models, and the estimates for the GL also show similar spatial patterns, although the spatial resolution of the ward-based coefficients is quite low.

If the neighborhood-based coefficients are set for all explanatory variables in the GL model, it might be possible to represent the spatial heterogeneity of coefficients more precisely; however, as the order of calculation of GL is estimated as  $O(mn^2 + Tm^2)$ , where  $m$  is the number of constraints,  $n$  is the number of explanatory variables, and  $T$  is the

number of iteration and should be larger or equal to  $m$  [22], the feasibility of estimation of high resolution model with large  $m$  and  $n$  is low. Thus, the ESF-GL-SVC model has advantages over the GL model in utilizing the spatially varying coefficient model to express the continuous spatial heterogeneity.

Third, the coefficients of 'time to CBD' reveal the advantage of ESF-GL-SVC model over ESF-SVC. It is natural that the coefficients of 'time to CBD' are negative, but the estimates by ESF-SVC have both positive and negative values, especially indicating large positive values in the northwest and southwest corner of the study area. On the other hand, the estimates by the ESF-GL-SVC model are all negative using the same eigenvector setting.

This indicates that ESF-SVC overfits to the dataset. It is caused by the utilization of a shorter band for spatial proximity matrix  $C$  and more eigenvectors that represent local spatial heterogeneity patterns in the ESF-SVC model. The cause of overfit is likely because 'time to CBD' has strong spatial autocorrelation on a global scale, although other explanatory variables do not. By multiplying the coefficients and explanatory variables, both of which have spatial autocorrelation, it is possible to express the spatial autocorrelation of dependent variables. The rents in the south of Meguro ward have strong local and discrete spatial heterogeneity, represented by the estimated coefficients of subregion dummies shown in Figure 11. The ESF-SVC model tried to express the local heterogeneity by the intercept (Figure 10b) and the coefficient of 'time to CBD' (Figure 16b); as a consequence, the coefficients of 'time to CBD' in the surrounding area were highly variable and difficult to interpret. This misspecification by ESF-SVC is avoidable if the only longer band is set for spatial proximity matrix  $C$ ; however, when the longer band was set, the eigenvectors represent only global spatial patterns, and the ESF-SVC model would lose explanatory power for the local heterogeneity.

### 3.4. Summary of Application to Apartment Rent Data Analysis

The application to apartment rent data reveals that the ESF-GL-SVC model outperforms the two previous models; it has advantages in extracting both continuous and discrete spatial heterogeneity, which apartment rent data have.

The ESF-GL-SVC model was able to estimate the spatially varying coefficients that are interpretable, although the ESF-SVC model was not. The analysis by the ESF-GL-SVC model clarifies the structure of spatial heterogeneity of the effect of explanatory variables. The coefficients of property size have the largest spatial heterogeneity; the size does not affect the rent per square meter near the CBD but decreases in suburban areas. It also depicts the existence of discrete spatial heterogeneity by neighborhood-based coefficients.

## 4. Discussion

This study proposed an analysis to extract both continuous and discrete spatial heterogeneity by ESF-GL-SVC combining ESF-SVC and GL. Through the analysis of simulated data and apartment rent data, it is confirmed that the ESF-GL-SVC model can separate the continuous and discrete spatial heterogeneity of dataset. It is possible to avoid the overfitting the dataset and output the interpretable estimates.

There are three ways to improve the ESF-GL-SVC model. The first improvement would be the mitigation of the effect caused by biased estimates of lasso. It was pointed out that the lasso estimator is biased toward zero and does not have an oracle property, which consists of the consistency in variable selection and the asymptotic normality [32]. The minimax concave penalty (MCP) was proposed to mitigate the bias [35] and was expanded to the fused MCP [36], which can reduce the bias of GL. We analyzed the geographic market segmentation of apartment rent [37] and the extraction of spatio-temporal changes in real estate market prices [38] by fused-MCP and confirmed that the fused-MCP outputs better estimates than the GL. However, there is an issue of computational complexity. The construction of a fusion model of ESF-SVC and fused-MCP with efficient estimation is an effective extension. The second improvement would be the mitigation of the overfitting problem of ESF-SVC. Murakami et al. [10] proposed a random effects ESF-SVC (RE-ESF-

SVC) model based on the random effects specification of ESF [34], and Murakami et al. [39] confirmed that RE-ESF-SVC is one of the models that can estimate the structure of spatially varying coefficients accurately and is the most computationally efficient. The application of RE-ESF-SVC would improve the estimation of spatially varying coefficients; however, as the estimation of RE-ESF requires the restricted maximum likelihood method, the introduction of regularization with the  $\ell_1$  norm or MCP functions might be challenging. The third improvement would be to enhance the capability of the method to analyze spatial heterogeneity at various scales. The proposed method is structured to analyze phenomena consisting of global and continuous spatial heterogeneity and local and discrete spatial heterogeneity. There might be local and continuous spatial heterogeneity, such as the impact of a small park on the neighborhood environment, and global and discrete spatial heterogeneity, such as the impact of regional boundaries in interconnected urban areas. RE-ESF-SVC and multi-scale GWR [40] have been proposed as methods to consider various scales in continuous spatial heterogeneity analysis, and group lasso [41] and tree structured group lasso [42] are expected to consider multiple scales of discrete spatial heterogeneity. Considering multiscale heterogeneity for both continuous and discrete spatial heterogeneity is an important development direction for this research.

**Supplementary Materials:** The code for simulation and rent analysis and the sample rent data are available at: <https://www.mdpi.com/article/10.3390/ijgi11070358/s1>.

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

**Funding:** This work was supported by JSPS KAKENHI, grant number 18H01552 and 21H01447.

**Data Availability Statement:** The original real estate rent data provided by At Home Co., Ltd., used in Section 3 cannot be made publicly available due to the data-disclosing policy of the data owner. We disclose a part of the dataset recorded on 3000 properties as Supplementary Materials. The dependent variables, rent per square meter of properties, were transformed logarithmically and the explanatory variables, the attributes of properties, were also transformed logarithmically and then normalized to mean 0 and variance 1.

**Acknowledgments:** Part of this research is the result of the joint research with the Center for Spatial Information Science, University of Tokyo (No. 815), and uses the 'Real Estate Database 2013–2017' provided by At Home Co., Ltd.

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

## References

1. Tobler, W. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [[CrossRef](#)]
2. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988.
3. Cressie, N.A.C. *Statistics for Spatial Data*; Wiley: New York, NY, USA, 1991.
4. Griffith, D.A. Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying georeferenced data. *Can. Geogr.* **1996**, *40*, 351–367. [[CrossRef](#)]
5. Griffith, D.A. *Spatial Autocorrelation and Spatial Filtering*; Springer: Berlin/Heidelberg, Germany, 2003.
6. Fotheringham, A.S.; Sachdeva, M. On the importance of thinking locally for statistics and society. *Spat. Stat.* **2022**, *in press*. [[CrossRef](#)]
7. Brunson, C.; Fotheringham, S.; Charlton, M. Geographically weighted regression: A method for exploring spatial non-stationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [[CrossRef](#)]
8. Fotheringham, A.S.; Brunson, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; Wiley: West Sussex, UK, 2002.
9. Griffith, D.A. Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environ. Plan. A Econ. Space* **2008**, *40*, 2751–2769. [[CrossRef](#)]
10. Murakami, D.; Yoshida, T.; Seya, H.; Griffith, D.A.; Yamagata, Y. A Moran coefficient-based mixed effects approach to investigate spatially varying relationships. *Spat. Stat.* **2017**, *19*, 68–89. [[CrossRef](#)]

11. Murakami, D.; Griffith, D.A. Spatially varying coefficient modeling for large datasets: Eliminating N from spatial regressions. *Spat. Stat.* **2019**, *30*, 39–64. [[CrossRef](#)]
12. Tan, H.; Chen, Y.; Wilson, J.P.; Zhang, J.; Cao, J.; Chu, T. An eigenvector spatial filtering based spatially varying coefficient model for PM<sub>2.5</sub> concentration estimation: A case study in Yangtze River Delta region of China. *Atmos. Environ.* **2020**, *223*, 117205. [[CrossRef](#)]
13. Murakami, D.; Kajita, M.; Kajita, S. Scalable model selection for spatial additive mixed modeling: Application to crime analysis. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 577. [[CrossRef](#)]
14. Peng, Z.; Inoue, R. Specifying multi-scale spatial heterogeneity in the rental housing market: The case of the Tokyo metropolitan area. In Proceedings of the GIScience 2021 Short Paper Proceedings, Poznań, Poland, 29 September 2021. [[CrossRef](#)]
15. Peng, Z.; Inoue, R. Identifying Multiple scales of spatial heterogeneity in housing prices based on eigenvector spatial filtering approaches. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 283. [[CrossRef](#)]
16. Goodman, A.C.; Thibodeau, T.G. Housing market segmentation. *J. Hous. Econ.* **1998**, *7*, 121–143. [[CrossRef](#)]
17. Goodman, A.C.; Thibodeau, T.G. Housing market segmentation and hedonic prediction accuracy. *J. Hous. Econ.* **2003**, *12*, 181–201. [[CrossRef](#)]
18. Kulldorff, M.; Nagarwalla, N. Spatial disease clusters: Detection and inference. *Stat. Med.* **1995**, *15*, 707–715. [[CrossRef](#)] [[PubMed](#)]
19. Castro, M.C.; Singer, B.H. Controlling the false discovery rate: A new application to account for multiple and dependent tests in local statistics of spatial association. *Geogr. Anal.* **2006**, *38*, 180–208. [[CrossRef](#)]
20. Brunsdon, C.; Charlton, M. An assessment of the effectiveness of multiple hypothesis testing for geographical anomaly detection. *Environ. Plan. B Plan. Des.* **2011**, *38*, 216–230. [[CrossRef](#)]
21. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [[CrossRef](#)]
22. Tibshirani, R.J.; Taylor, J. The solution path of the generalized lasso. *Ann. Stat.* **2011**, *39*, 1335–1371. [[CrossRef](#)]
23. Inoue, R.; Ishiyama, R.; Sugiura, A. Identification of geographical segmentation of the rental apartment market in the Tokyo Metropolitan Area. In Proceedings of the 10th International Conference on Geographic Information Science 2018, Melbourne, Australia, 30 August 2018. [[CrossRef](#)]
24. Inoue, R.; Ishiyama, R.; Sugiura, A. Identification of geographical segmentation of the rental housing market in the Tokyo Metropolitan Area by generalized fused lasso. *J. Jpn. Soc. Civ. Eng. Ser. D3 (Infrastruct. Plan. Manag.)* **2020**, *76*, 251–263. (In Japanese) [[CrossRef](#)]
25. Parker, R.J.; Reich, B.J.; Eidsvik, J. A fused lasso approach to nonstationary spatial covariance estimation. *J. Agric. Biol. Environ. Stat.* **2016**, *21*, 569–587. [[CrossRef](#)]
26. Sun, Y.; Wang, H.J.; Fuentes, M. Fused adaptive lasso for spatial and temporal quantile function estimation. *Technometrics* **2016**, *58*, 127–137. [[CrossRef](#)]
27. Wang, H.; Rodríguez, A. Identifying pediatric cancer clusters in Florida using log-linear models and generalized lasso penalties. *Stat. Public Policy* **2014**, *1*, 86–96. [[CrossRef](#)] [[PubMed](#)]
28. Choi, H.; Song, E.; Hwang, S.S.; Lee, W. A modified generalized lasso algorithm to detect local spatial clusters for count data. *ASA Adv. Stat. Anal.* **2018**, *102*, 537–563. [[CrossRef](#)]
29. Masuda, R.; Inoue, R. Point-event cluster detection via the Bayesian generalized fused lasso. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 187. [[CrossRef](#)]
30. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **1996**, *58*, 267–288. [[CrossRef](#)]
31. Griffith, D.A.; Chun, Y. Spatial autocorrelation and eigenvector spatial filtering. In *Handbook of Regional Science*; Fischer, M., Nijkamp, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 1477–1507. [[CrossRef](#)]
32. Fan, J.; Li, R. Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.* **2001**, *95*, 1348–1360. [[CrossRef](#)]
33. Arnold, T.B.; Tibshirani, R.J. Path Algorithm for Generalized Lasso Problems. 2020. Available online: <https://cran.r-project.org/web/packages/genlasso/genlasso.pdf> (accessed on 20 March 2022).
34. Murakami, D.; Griffith, D.A. Random effects specifications in eigenvector spatial filtering: A simulation study. *J. Geogr. Syst.* **2015**, *17*, 311–331. [[CrossRef](#)]
35. Zhang, C.H. Nearly unbiased variable selection under minimax concave penalty. *Ann. Stat.* **2010**, *38*, 894–942. [[CrossRef](#)]
36. Jing, B.; Yang, G.; Yu, X.; Zhang, C. Fused-MCP with application to signal processing. *J. Comput. Graph. Stat.* **2018**, *27*, 872–886. [[CrossRef](#)]
37. Inoue, R.; Ishiyama, R.; Sugiura, A. Identifying local differences with fused-MCP: An apartment rental market case study on geographical segmentation detection. *Jpn. J. Stat. Data Sci.* **2020**, *3*, 183–214. [[CrossRef](#)]
38. Den, K.; Inoue, R. Extracting area and period of influence of new rail service on real estate market using fused-MCP. In Proceedings of the GeoComputation 2019, Queenstown, New Zealand, 19 September 2019. [[CrossRef](#)]
39. Murakami, D.; Lu, B.; Harris, P.; Brunsdon, C.; Charlton, M.; Nakaya, T.; Griffith, D.A. The importance of scale in spatially varying coefficient modelling. *Ann. Am. Assoc. Geogr.* **2019**, *109*, 50–70. [[CrossRef](#)]
40. Fotheringham, A.S.; Yang, W.; Kang, W. Multi-scale geographically weighted regression. *Ann. Am. Assoc. Geogr.* **2017**, *107*, 1247–1265. [[CrossRef](#)]

- 
41. Yuan, M.; Lin, Y. Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2006**, *68*, 49–67. [[CrossRef](#)]
  42. Zhao, P.; Rocha, G.; Yu, B. The composite absolute penalties family for grouped and hierarchical variable selection. *Ann. Stat.* **2009**, *37*, 3468–3497. [[CrossRef](#)]