WILEY | Hindawi

Research Article

Industrial Efficiency Algorithm Based on Spatio-Temporal-Data-Driven

Hongqu Lv and Wensi Cheng [ID]

Shandong University of Political Science and Law, Jinan, China

Correspondence should be addressed to Wensi Cheng; chengwensi777@163.com

Stochastic frontier model is an important and effective method to calculate industry efficiency. However, when dealing with temporal and spatial data from the industry, it is difficult to accurately calculate the industrial production efficiency due to the influence of spatial correlation and time lag effect. If the traditional spatial statistical method is used, the setting method of spatial weight matrix is often questioned. To solve this series of problems, one possible idea is to design a spatial data mining process based on stochastic frontier analysis. Firstly, the stochastic frontier model should be improved to analyze spatio-temporal data. In order to accurately measure the technical efficiency in the case of dual correlation between time and space, a more effective spatio-temporal stochastic frontier model method is proposed. Meanwhile, based on the idea of generalized moment estimation, an estimation method of spatiotemporal stochastic frontier model is designed, and the consistency of estimators is proved. In order to ensure that the most appropriate spatial weight matrix can be selected in the process of model construction, the K-fold crossvalidation method is adopted to evaluate the prediction effect under the data-driven idea. This set of spatio-temporal data mining methods will be used to measure the technical efficiency of high-tech industries in various provinces of China.

1. Introduction

Stochastic frontier analysis (SFA) is an important method to measure technical efficiency and calculate total factor productivity. The whole process is divided into two steps: the first step is the model estimation process, which can be regarded as a supervised learning process; the second step is to use the estimated model to calculate the technical efficiency, which can be regarded as an unsupervised learning process.

From the perspective of machine learning, supervised processes have three main objectives: (a) feature selection and reduction of the dimension of feature variables; (b) selecting the optimal one from multiple classifiers or prediction models; (c) model evaluation, which estimates the prediction error of the selected classifier or prediction model on the new data.

The paper found that the traditional stochastic frontier analysis method has the following defects: (a) it is not suitable for the special structure of spatial data or spatio-temporal data; (b) the modeling process lacks variety. The traditional analysis process is knowledge-driven and completely relies on a single theoretical model for estimation and testing. The above two characteristics lead to the large deviation of the traditional stochastic frontier model when analyzing the spatio-temporal data, and it is impossible to make an accurate measure of the production efficiency with spatial relationship, either. To solve the two problems above, this study considers two improvements to the industrial efficiency calculation process based on temporal and spatial data: (1) improve the existing stochastic frontier model and make it suitable for spatial data or spatio-temporal data; (2) turn the modeling process into a spatial data mining process. In view of the unique structure of spatio-temporal data, a more suitable crossvalidation method is proposed for the selection of prediction model.

Stochastic frontier analysis (SFA) was successively proposed by Aigner et al. (1977) [1], Meeusen and Broeck (1977) [2], and Battese and Corra (1977) [3]. Over the past

40 years, its theoretical system and methods have been continuously expanded and innovated; it is widely used to measure the operating efficiency of different industries.

The development of spatial statistics provides a theoretical basis for studying spatial interactions in stochastic frontier models. Druska and Horrace (2004) [4] first applied the method of spatial econometrics to the analytical framework of stochastic frontier model and started the research of spatial stochastic frontier model. Affuso (2010) [5] established the spatial stochastic frontier model and gave the maximum likelihood estimation in the empirical study. Tonini and Pede (2011) [6] applied maximum entropy method to parameter estimation of spatial stochastic frontier model. Vidolia et al. (2016) [7], Tsionas and Michaelides (2016) [8], Carvalho (2018) [9], and Adetutu et al. (2015) [10] consider SF models with local spatial dependence. Jin and Lee (2020) [11] proved the asymptotic properties of a maximum likelihood estimator of a spatial autoregressive stochastic frontier model. Kutlu et al. (2020) [12] proposed a spatial autoregressive stochastic frontier model, which allows for the endogeneity in both the frontier and environmental variables, and discussed a single-stage control function approach to estimate the parameters.

Because spatial stochastic frontier analysis methods can fully consider the impact of spatial correlation, they can obtain more accurate results in efficiency analysis of data with spatial spillover effect and thus have been more widely used in recent years. Bergantino et al. (2020) [13] analyses the potential impact of airport competition on technical efficiency by applying the spatial stochastic frontier. Graaff (2020) [14] used spatial stochastic frontier model to estimate spatially correlated technical efficiencies within a European regional production function context. At present, some literatures have studied panel spatial stochastic frontier model, for example, Druska and Horrace (2004) [4], Tonini and Pede (2011) [6], and Lin Jia-Xian (2014) [15]. These literatures all focus on the static panel space stochastic frontier model, and the model utilizes two-dimensional information from panel data; formally, the spatial lag term of the explained variable and the spatial lag term of the error are used to capture the spatial correlation of the production unit. The time lag term is not included in the model, which means that the model still cannot fit well when there is significant inertia in the research problem. In input-output analysis, current behavior is largely dependent on past behavior, for example, the adjustment of capital stock is often influenced by previous capital. Therefore, a dynamic stochastic frontier model should be established, and the model should describe the double lag effect of space and time, so as to reflect the influence relationship between economic variables more objectively. The spatial weight matrix in spatial statistics is often considered to be "subjective." Moreover, due to the various setting methods of spatial weight matrix, the selection of different spatial weight matrix may lead to the difference of model estimation results. In addition, the selection of spatial weight matrix has not formed a unified principle. Based on the above three points, the spatial weight matrix is often questioned. But in the era of "big data," such skepticism may end [16].

This paper proposes the spatiotemporal stochastic frontier model; considering that the model may be endogenous in time and space dimensions, a generalized method of moments (GMM) estimation process is designed to estimate the model. When Druska and Horrace (2004) [4] studied the static panel space stochastic frontier model, a generalized moment estimation process was proposed by referring to Kelejian and Prucha (1999) [17] for spatial error correlation. In this paper, Druska and Horrace (2004) [4] is used to deal with model's error space autocorrelation, which is different from that of the stochastic frontier model. According to the method of Kapoor et al. (2007) [18], the compound error term was processed, and the moment condition was constructed to estimate the distribution parameters of the error term. In this paper, Jacobs et al. (2009) [19] was used as a reference to construct the moment condition, and Anselin (1988) [20] was used as a reference for the selection of tool variables to obtain the generalized moment estimator. Furthermore, the consistency of the obtained structural parameter estimators is proved by using the extreme value consistency theorem and the law of uniform large numbers (ULLN). To solve the problem of selecting spatial weight matrix, we can consider a crossvalidation method suitable for spatio-temporal data. Fortunately, a series of methods such as dimensionality reduction, feature selection, and model generalization has been provided by machine learning methods. The earliest crossvalidation method was called hold-out, which relied on only one partition of the data, and there was no crossover process, so it was also called the verification method [21]. Noting that the hold-out method relies on a partition of data and is easily affected by contingency factors, Geisser (2010) [22] proposed a crossvalidation method that includes the average of multiple hold-out estimates, realizing the transition from verification estimation to crossvalidation estimation. In order to reduce the combination number of data partition in crossvalidation, Shao (1993) [23] proposed the leave-P-out crossvalidation (LPOCV) in which the number of test samples in each data partition was the same. Especially in the special case when $P = 1$, the method is evolved to leave-one-out crossvalidation (LOOCV). LOOCV is the simplest and most widely used crossvalidation in traditional analysis. Compared with the LPOCV considering all data partitioning, Geisser (2010) also proposed a crossvalidation based on only partial data partitioning, which is called RLT method. K-folded crossvalidation is proposed as an alternative to LOOCV which has a large computational overhead and relies on a basic partition of data divided into K-fold, each of which has a data capacity of N/K. In the case of limited samples, k-fold crossvalidation is the simplest and most widely used method of generalization error estimation. From the various crossvalidation methods that have appeared in the past, each method fully considers the randomness of the validation set to ensure the generalization ability of the test model. However, for the special panel data such as spatio-temporal data, there is usually an internal connection between spatial individuals, and the overall data also tends to have time trend. This problem is not taken into account by the previous crossvalidation methods, which may break the inherent regularity of spatio-

temporal data. Based on the above considerations, this paper designs a kind of crossvalidation scheme suitable for spatiotemporal data. It is used to select stochastic frontier models, especially models with different weight matrices.

Finally, the technology efficiency of China's high-tech industry is analyzed by establishing a spatiotemporal stochastic frontier model.

2. Methodology

Previous studies on panel spatial stochastic frontier models mainly involved static panel spatial stochastic frontier models. Spatial lag effect is considered in the process of model building, but the influence of time lag effect is not included. If the time lag term and time-spatial lag term are added into the model, this kind of model can be called spatiotemporal stochastic frontier model. Obviously, the time-space double lag effect will produce stronger endogeneity, and new estimation methods should be considered to solve it.

2.1. Model Specification and Assumption

2.1.1. Model Specification. The general form of the spatiotemporal stochastic frontier model can be stated in matrix form as

$$Y_t = \lambda_1 W Y_t + \lambda_2 W Y_{t-1} + \gamma Y_{t-1} + X_t B + E_t,$$

$$E_t = \rho M E_t + \varepsilon_t, \tag{1}$$

$$\varepsilon_t = v_t - u,$$

where Y_t, E_t, ε_t, and v_t are N-dimensional vectors, whose components at time $t = 1, \cdots, T$ are given by $Y_t = [y_{1t}, \cdots, y_{Nt}]'$, $E_t = [E_{1t}, \cdots, E_{Nt}]'$, $\varepsilon_t = [\varepsilon_{1t}, \cdots, \varepsilon_{Nt}]'$, and $v_t = [v_{1t}, \cdots, v_{Nt}]'$. The vector Y_t consists of the outputs of the N production units, E_t and ε_t are the composite error vectors corresponding to Y_t, v_t is the heterogeneous error vector, and $u = [u_1, \cdots, u_N]'$ is the vector of time-invariant inefficiency terms. This kind of setting is appropriate when the time span is not large. As u is time invariant, it can be regarded as the individual effect, and thus, this paper primarily considers u as a fixed effect. X_t is an $N \times K$-dimensional matrix consisting of the K exogenous input variables of the N production units at time t. W and M are $N \times N$ spatial weight matrices which are usually assumed to be different. If $W = M$, λ_1 and ρ cannot be distinguished by means of the maximum likelihood method although they can be effectively distinguished by the GMM method [19]. The variables are stacked according to the section and time series in the following matrix form:

$$Y = \lambda_1 (I_T \otimes W) Y + \lambda_2 (I_T \otimes W) Y_{-1} + \gamma Y_{-1} + XB + E,$$

$$E = \rho (I_T \otimes M) E + \varepsilon,$$

$$\varepsilon = v - (e_T \otimes I_N) u, \tag{2}$$

where $\otimes$ represents the Kronecker product of matrices, I_T and I_N are, respectively, the identity matrices of orders T and N, and e_T is a T-dimensional column vector with all the entries equal to 1. The parameter vector of the model to be estimated is $(\lambda_1, \lambda_2, \gamma, B, \rho, \sigma_v^2, \sigma_u^2)$, and its dimension is $(K + 6)$, where B is the parameters corresponding to the K-dimension exogenous explanatory variables X_t. B and λ_1, λ_2, γ together constitute the structural parameters, and ρ, σ_v^2, σ_u^2 are the error term parameters.

2.1.2. Model Assumption. The assumptions of the spatiotemporal stochastic frontier model are the following:

Assumption 1. The distribution of the error vector v is given by $v \sim N(0, \sigma_v^2 I_{NT})$.

Assumption 2. The inefficiency term u is time invariant, with distribution $u \sim N^+(0, \sigma_u^2 I_{NT})$.

Assumption 3. $0 < \sigma_v^2 < b_v < \infty$, $0 < \sigma_u^2 < b_u < \infty$, and v and u have finite fourth moment.

Assumption 4. v and u are uncorrelated with X.

Assumption 5. The spatial weight matrices W and M satisfy $w_{ii} = 0$ and $m_{ii} = 0$; $i = 1, 2, \cdots, N$. For arbitrary $|\lambda_1| < 1$, $|\lambda_2| < 1$, and $|\rho| < 1$, the matrices $1 - \lambda_1 W$, $1 - \lambda_2 W$, and $1 - \rho M$ are all nonsingular matrices. For each of the matrices W, M, $1 - \lambda_1 W$, $1 - \lambda_2 W$, $1 - \rho M$, $(1 - \lambda_1 W)^{-1}$, $(1 - \lambda_2 W)^{-1}$, and $(1 - \rho M)^{-1}$, the row sums and column sums are all absolutely uniformly bounded.

Assumption 1 is a classic assumption of the spatial error autocorrelation model. By Assumption 2, the same individual inefficiency term remains constant at different times. When using GMM to estimate the structural parameters of the model, the distribution of error terms can be ignored; nevertheless, in order to improve the efficiency of computation, the half normal distribution for the inefficiency term is usually assumed. Assumption 3 ensures the boundedness of the variance of the error term in this model, which is an important condition for the consistency of the estimator. Assumption 4 is a classical assumption commonly used in traditional regression analysis methods, and the moment condition is set according to this assumption in the generalized moment estimation of this model. Assumption 5 is set according to the space weight matrix of this model and the properties of space station autoregressive coefficient and space-time autoregressive coefficient, which also ensures the consistency of parameter estimators.

2.2. Parameter Estimation

2.2.1. Estimation strategy. There is an endogeneity problem in the model. For the spatial lag term $\lambda_1 W Y_t$ in the model, there is

$$\begin{aligned}
COV[(WY_t), E_t] &= COV\{[WA^{-1}(\lambda_2 WY_{t-1} \\
&\quad + \gamma Y_{t-1} + X_t B + E_t)], E_t\} \\
&= \lambda_2 COV\{[WA^{-1}(WY_{t-1})], E_t\} + \gamma COV \\
&\quad \cdot \{[WA^{-1}Y_{t-1}], E_t\} + COV\{[WA^{-1}]E_t, E_t\},
\end{aligned} \tag{3}$$

where $A = I - \lambda_1 W$, and $COV\{[WA^{-1}(WY_{t-1})], E_t\} = 0$, $COV\{[WA^{-1}Y_{t-1}], E_t\} = 0$, and $COV\{WA^{-1}[X_t B], E_t\} = 0$ can be obtained from the assumptions of the regression model, and $COV\{[WA^{-1}]E_t, E_t\}$ is quadratic. In spatial econometrics, W is usually not a zero matrix, and so, WA^{-1} is not a zero matrix. While taking into account the expected value of the compound error term cannot be 0, it can be considered that the quadratic form $COV\{[WA^{-1}]E_t, E_t\}$ is almost impossible to be equal to 0 (see Appendix A for proof). Therefore, in the dynamic panel spatial stochastic frontier models, there is an endogeneity problem which will lead to the inconsistency of traditional estimators. So, we considered GMM as a good way to solve the endogeneity problem.

The parameter vector to be estimated in the model is $(\lambda_1, \lambda_2, \gamma, B, \rho, \sigma_v^2, \sigma_u^2)'$, where $\lambda_1, \lambda_2, \gamma$, and B are the structural parameters of the model, and ρ, σ_v^2, and σ_u^2 are the error distribution parameter of the model. The estimation of the model is completed in three steps:

Step 1. Using the GMM to estimate the structure parameter $(\lambda_1, \lambda_2, \gamma, B)$ in the model.

Step 2. Making a moment estimation of the parameter $(\rho, \sigma_v^2, \sigma_u^2)$ that is included in the error term.

Step 3. Using the estimator obtained in Step 2 to modify the result of Step 1.

2.2.2. Estimation of Structural Parameter $(\lambda_1, \lambda_2, \gamma, B)$

(1) Difference Model and Level Model. Anderson and Hsiao (1981) [24] proposed to use $y_{i,t-2}$ as the instrumental variable of $\Delta y_{i,t-1}$, and then, 2SLS estimation is carried out. This estimator is called "Anderson-Hsiao estimator." According to the same logic, lag variables of higher order are also valid IV. Arellano and Bond (1991) [25] used all possible lag variables as IV (the number of IVs is more than the number of endogenous variables) to conduct GMM estimation. This GMM estimator is called Arellano-Bond estimator or difference GMM. The disadvantage of difference GMM is that the variable which does not change with time is eliminated, and its coefficient cannot be estimated. If the series $\{y_{i,t}\}$ has a strong persistence, that is, the first-order autoregressive coefficient is close to 1, then the correlation may be very weak and lead to the problem of weak instrumental variables. In order to solve the above two problems, Arellano and Bver (1995) [26] returned to the level equation and used $\{\Delta y_{i,t-1}, \Delta y_{i,t-2}, \cdots\}$ as IV to estimate the GMM of the level equation, which was called "level GMM." Blundell and Bond (1998) [27] combined difference GMM with level GMM and estimated the difference equation and level equation as one

equation system for GMM, which was called "system GMM." The advantage of system GMM is that it can improve the efficiency of estimation (small sample properties are better), and it can estimate the variable that does not change with time (the system GMM contains the level equation). In order to solve the endogenous problem of dynamic panel data model, Arellano and Bond (1991) [25], Arellano and Bover (1995) [26], and Blundell and Bond (1998) [27], respectively, considered from the perspective of difference model and level model, and different instrumental variables were selected.

The corresponding difference model and level model of Equation (1) are simplified as

$$\Delta Y_t = \Delta Z_t \theta + \Delta E_t, \tag{4}$$

$$Y_t = Z_t \theta + E_t. \tag{5}$$

(4) and (5) can also be collectively called spatial system model, where Equation (4) is the difference model, and Equation (5) is the level model, $Z_t = [WY_t, WY_{t-1}, Y_{t-1}, X_t]'$ is the vector composed of all explanatory variables, and $\theta = [\lambda_1, \lambda_2, \gamma, B]'$ is the vector composed of structural parameters. The expansion of Equation (5) is Equation (1); the expansion of Equation (4) can be expressed as follows:

$$\begin{aligned}
\Delta Y_t &= \lambda_1 W\Delta Y_t + \lambda_2 W\Delta Y_{t-1} + \gamma \Delta Y_{t-1} + \Delta X_t B + \Delta E_t, \\
E_t &= \rho M\Delta E_t + \Delta \varepsilon_t, \\
\Delta \varepsilon_t &= \Delta v_t.
\end{aligned} \tag{6}$$

(2) Moment Condition and Instrumental Variable. Since ΔX_t is a strictly exogenous variable, it is not related to the compound error term ΔE_t, nor is it related to E_t. The moment conditions for identifying B in the difference model and the level model are as follows:

$$\begin{aligned}
E\left(\Delta X_t{}' \Delta E_t\right) &= 0, \quad t = 3, \cdots, T, \\
E\left(\Delta X_t{}' E_t\right) &= 0, \quad t = 3, \cdots, T.
\end{aligned} \tag{7}$$

The moment condition structure for identifying λ_2 and γ in the two models is as follows: since the spatial lag term and time lag term of the dependent variable ΔY_t are both endogenous variables, therefore, it is necessary to find a set of instrumental variables that is related to time lag and space lag and exogenous explanatory variables, but not related to the difference error term $\Delta E_t (t = 3, \cdots, T)$. Arellano and Bond (1991) [25] uses all possible level lag variables $(y_{t-2}, \cdots, y_1)$ of Y_t as instrumental variables for the time-lag first-order difference term (ΔY_{t-1}) of the dependent variable. These instrumental variables are related to (ΔY_{t-1}), but not to ΔE_t. The moment conditions corresponding to the difference model and the level model are as follows:

$$E\left(Y_{t-s}{}'\Delta E_t\right) = 0, \quad t = 3, \cdots, T; s = 2, \cdots, t-1, \quad (8)$$

$$E\left(\Delta Y_{t-s}{}'E_t\right) = 0, \quad t = 3, \cdots, T; s = 2, \cdots, t-1. \quad (9)$$

The moment conditions for identifying λ_1 in the two models are as follows.

Construct a spatial lag item WY_t as follows; Jacobs et al. (2009) [19] provided a method of finding instrumental variables, that is time lag terms of spatial lag dependent variables, who also proved that the moment condition obtained by this method was as valid as Equation (8). So, corresponding to the difference model and the level model, the following moment conditions can be listed:

$$E\left(\left\{W^l Y_{t-s}\right\}'\Delta E_t\right) = 0, \quad t = 3, \cdots, T; s = 2, \cdots, t-1; l = 1, \cdots, L,$$

$$E\left(\left\{W^l \Delta Y_{t-s}\right\}'E_t\right) = 0, \quad t = 3, \cdots, T; s = 1, \cdots, t-2; l = 1, \cdots L,$$

$$(10)$$

where l is the exponential of matrix W and the integer L is the maximum order of spatial lag that can be used as the instrumental variable.

In addition, based on the method provided by Kelejian and Robinson (1993) [28], formula (1) shows that WY_t depends on WX_t, so the instrumental variable $W\Delta X_t$ can be selected by the first-order difference method for $W\Delta Y_t$.

Since ΔX_t is a strictly exogenous variable, it is not related to the compound error term ΔE_t, so corresponding to the difference model and the level model, the instrumental variables satisfy the following moment conditions:

$$E\left(\left\{W^l \Delta X_t\right\}'\Delta E_t\right) = 0, \quad t = 3, \cdots, T,$$

$$E\left(\left\{W^l \Delta X_t\right\}'E_t\right) = 0, \quad t = 3, \cdots, T. \quad (11)$$

(3) GMM Estimation. When we estimate the parameters of the spatio-temporal stochastic frontier model, we use the system GMM method similar to the general dynamic panel model to construct the spatial system GMM estimation. Unlike the system GMM, the IVs of the spatial system GMM are composed of time lag variable and spatial lag variable.

For each period of t, the moment condition of $J \geq K + 2$ can be given. The moment conditions corresponding to the difference model and the level model can be abbreviated as

$$E\left(H_{N,ABt}{}'\Delta E_t\right) = 0,$$

$$E\left(H_{N,Lt}{}'E_t\right) = 0. \quad (12)$$

The matrices $H_{N,ABt}$ and $H_{N,Lt}$ are expressed as follows

$$Y_{t-s} = \begin{bmatrix} y_1 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & y_2 & y_1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & y_3 & y_2 & y_1 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & y_{t-2} & y_{t-3} & \cdots & y_1 \end{bmatrix}, X_t = \begin{bmatrix} x_3 \\ x_4 \\ \vdots \\ x_t \end{bmatrix},$$

$$\Delta Y_{t-s} = \begin{bmatrix} \Delta y_1 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \Delta y_2 & \Delta y_1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \Delta y_3 & \Delta y_2 & \Delta y_1 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \Delta y_{t-2} & \Delta y_{t-3} & \cdots & \Delta y_1 \end{bmatrix}, \Delta X_t = \begin{bmatrix} \Delta x_3 \\ \Delta x_4 \\ \vdots \\ \Delta x_t \end{bmatrix}.$$

$$(13)$$

Then, $H_{N,ABt} = [Y_{t-s}, WY_{t-s}, W^2 Y_{t-s}, \cdots, W^L Y_{t-s}, W\Delta X_t, \Delta X_t]$ and

$$H_{N,Lt} = \left[\Delta Y_{t-s}, W\Delta Y_{t-s}, W^2\Delta Y_{t-s}, \cdots, W^L\Delta Y_{t-s}, WX_t, X_t\right](t = 3, \cdots, T).$$

$$(14)$$

That is, $H_{N,ABt}$ and $H_{N,Lt}$ are matrices with instrumental variables as column vectors, and the subscript N means that the matrix depends on the unit number of individuals. Let

$H_{N,AB}$ and $H_{N,L}$ be block diagonal matrices composed of block $H_{N,ABt}$ and $H_{N,Lt}$, respectively. In order to define the GMM (Spatial Blundell Bond, SBB) estimator of the spatial dynamic panel stochastic frontier model, the difference variables and level variables are combined to define the matrix as follows:

$$Y*_N = \left[\Delta Y'_N, Y'_N\right], Z*_N = \left[\Delta Z'_N, Z'_N\right], E*_N = \left[\Delta E'_N, E'_N\right].$$

$$(15)$$

The instrument variable matrix is

$$H_{N,BB} = \text{diag}\left\{H_{N,AB}, H_{N,L}\right\} = \begin{bmatrix} H_{N,AB} & 0 \\ 0 & H_{N,L} \end{bmatrix}, \qquad (16)$$

where $H_{N,AB}$ is the instrumental matrix of spatial difference GMM estimation, and $H_{N,L}$ is the instrumental matrix of spatial level GMM estimation. The weight matrix is

$$G_{N,BB} = \text{diag}\left\{G_{N,AB}, I_{T-2} \otimes I_N\right\} = \begin{bmatrix} G_{N,AB} & 0 \\ 0 & I_{T-2} \otimes I_N \end{bmatrix}. \qquad (17)$$

This diagonal of the matrix is composed of the weight matrix defined in the process of spatial difference GMM estimation and an identity matrix, where $G_{N,AB} = I_N \otimes G$ is $N(T-2) \times N(T-2)$ weight matrix which elements are

$$G_{ij} = \begin{cases} 2, & i = j, \\ -1, & i = j+1, \\ -1, & j = i+1, \\ 0, & others. \end{cases} \qquad (18)$$

This weight matrix is proposed by Arellano and Bond (1991) [25] which is further define the weight matrix:

$$A_{N,BB} = \left[\frac{H'_{N,BB}G_{N,BB}H_{N,BB}}{N}\right]^{-1}. \qquad (19)$$

Through the above process, combining the spatial difference equation with the spatial level equation, we get the spatial system GMM estimation process. Get the objective function of generalized moment estimation for spatial system as follows:

$$\frac{1}{N}\left(H'_{N,BB}\Delta E*_N\right)'A_{N,BB}\left(H'_{N,BB}\Delta E*_N\right)$$
$$= \frac{1}{N}\left[H'_{N,BB}(\Delta Y*_N - \Delta Z*_N\theta)_N\right]'A_{N,BB} \qquad (20)$$
$$\cdot \left[H'_{N,BB}(\Delta Y*_N - \Delta Z*_N\theta)_N\right].$$

The one-stage SBB estimator of θ can be obtained by minimizing Equation (20):

$$\widehat{\theta}_{SBB} = \left(Z*'_N H'_{N,BB}A*_N H'_{N,BB}Z*_N\right)Z*'_N H'_{N,BB}A*_N H'_{N,BB}Y*_N. \qquad (21)$$

Equation (21) can also be called the spatial system GMM estimator.

(4) Improvement of Instrumental Variable Matrix. The instrumental variable matrix constructed in accordance with the above method has a high dimension and grows exponentially as the values of T and L increase. In order to reduce the dimension of the instrumental variable matrix and avoid overfitting the instrumental variable, we can simplify it by using the "condensing instrumental variable matrix" proposed by Beck and Levine (2004) [29].

We still set $s = 2$ and $L = 1$ in the GMM instrumental variable matrix of the space system, and the corresponding condensed instrumental variable matrix is

$$H'_{N,BB} = \begin{bmatrix} H^1_{AB} & \cdots & H^N_{AB} \\ H^1_L & \cdots & H^N_L \end{bmatrix}, \qquad (22)$$

where H^i_L $(i = 1, \cdots, N)$ is the instrumental variable quantum matrix of the level model corresponding to the i individual.

$$H^i_L = \begin{bmatrix} \Delta y_1 & 0 & 0 & \cdots & W\Delta y_1 & 0 & 0 & \cdots & W\Delta X_3 & \Delta X_3 \\ \Delta y_2 & \Delta y_1 & 0 & \cdots & W\Delta y_2 & W\Delta y_1 & 0 & \cdots & W\Delta X_4 & \Delta X_4 \\ \Delta y_3 & \Delta y_2 & \Delta y_1 & \cdots & W\Delta y_3 & W\Delta y_2 & W\Delta y_1 & \cdots & W\Delta X_5 & \Delta X_5 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{bmatrix}. \qquad (23)$$

2.3. Estimation of Error Term Parameters

2.3.1. Estimation of ρ, σ_v^2. After the estimator of the structural parameter $\theta_N = (\lambda_1, \lambda_2, \gamma, B)'$ in model (1) is obtained in the first stage, the model residual $\widehat{E}_t = Y_t - Z_t\widehat{\theta}_N$ can be further obtained, in which the $Z_t = (WY_t, WY_{t-1}, Y_{t-1}, X_t)$ is the vector set composed of all explanatory variables in model

(1). Consistent GMM estimation can be obtained by using residual $\widehat{E}_t$ and modifying the moment condition proposed by Kapoor et al. (2007) [18]. The specific process is as follows.

According to the assumptions of the model, the individual effect of the model is the inefficiency term. According to the covariance structure of the compound error term, it can be known that

$$Cov\left(\varepsilon\varepsilon'\right) = \sigma_u^2(I_N \otimes J_T) + \sigma_v^2 I_{NT}. \qquad (24)$$

Introduce transformation matrix:

$$Q_0 = \left(I_T - \frac{J_T}{T}\right) \otimes I_N, \ Q_1 = \left(\frac{J_T}{T}\right) \otimes I_N, \qquad (25)$$

where I_T and I_N are the identity matrix of order T and order N, $J_T = e_T e_T'$ is the matrix of order $T \times T$, and the elements of that are all 1. Properties of transformation matrix Q_0: (i) $tr(Q_0) = N(T-1)$; (ii) $Q_0(e_T \otimes I_N) = 0$; (iii) $Q_0(I_T \otimes R_N) = (I_T \otimes R_N)Q_0$ (where R_N is any $N \times N$ matrix). From the properties (ii), we can further deduce the special properties (iv) $Q_0\varepsilon = Q_0 v$ of matrix Q_0 in this paper, where ε and v are the corresponding error terms in model (1).

Let

$$\bar{E} = (I_T \otimes M)E \ ; \ \bar{\bar{E}} = (I_T \otimes M)\bar{E} \ ; \ \bar{\varepsilon} = (I_T \otimes M)\varepsilon. \qquad (26)$$

Then,

$$\varepsilon = E - \rho\bar{E}, \bar{\varepsilon} = \bar{E} - \rho\bar{\bar{E}}. \qquad (27)$$

Based on the above transformation and referring to the first three of the six moment conditions given by Kapoor et al. (2007) [18] and related properties, the following three moment conditions are given in this paper:

$$E\left(\varepsilon' Q_0 \varepsilon\right) = \sigma_v^2 N(T-1), E\left(\bar{\varepsilon}' Q_0 \bar{\varepsilon}\right)$$
$$= \sigma_v^2 N(T-1) tr\left(M'M\right), E\left(\bar{\varepsilon}' Q_0 \varepsilon\right) = 0. \qquad (28)$$

To further integrate the above moment conditions, we can get that

$$E \begin{pmatrix} \dfrac{1}{N(T-1)} \varepsilon' Q_0 \varepsilon \\[2ex] \dfrac{1}{N(T-1)} \bar{\varepsilon}' Q_0 \bar{\varepsilon} \\[2ex] \dfrac{1}{N(T-1)} \bar{\varepsilon}' Q_0 \varepsilon \end{pmatrix} = \begin{pmatrix} \sigma_v^2 \\[2ex] \dfrac{1}{N} \sigma_v^2 tr\left(M'M\right) \\[2ex] 0 \end{pmatrix}. \qquad (29)$$

Substitute Equations (26) and (27) into Equation (29) to obtain that

$$\frac{1}{N} E \begin{pmatrix} \dfrac{2}{(T-1)} \bar{E}' Q_0 E & -\dfrac{1}{(T-1)} \bar{E}' Q_0 \bar{E} & 1 \\[2ex] \dfrac{2}{(T-1)} \bar{E}' Q_0 \bar{\bar{E}} & -\dfrac{1}{(T-1)} \bar{\bar{E}}' Q_0 \bar{\bar{E}} & tr\left(M'M\right) \\[2ex] \dfrac{1}{(T-1)}\left(\bar{E}' Q_0 \bar{\bar{E}} + \bar{E}' Q_0 \bar{\bar{E}}\right) & -\dfrac{1}{(T-1)} \bar{E}' Q_0 \bar{\bar{E}} & 0 \end{pmatrix} \begin{pmatrix} \rho \\[2ex] \rho^2 \\[2ex] \sigma_v^2 \end{pmatrix}$$
$$= \frac{1}{N} \begin{pmatrix} \dfrac{1}{(T-1)} E' Q_0 E \\[2ex] \dfrac{1}{(T-1)} \bar{E}' Q_0 \bar{E} \\[2ex] \dfrac{1}{(T-1)} E' Q_0 \bar{E} \end{pmatrix}. \qquad (30)$$

The residual $\widehat{E}_t = Y_t - Z_t \widehat{\theta}_N$ estimated in the first stage is substituted into $\bar{E}$ and $\bar{\bar{E}}$ in Equation (30) to obtain the sample moment equation. In the sample moment equation, the estimated value $\widehat{\rho}, \widehat{\sigma}_v^2$ of ρ and σ_v^2 can be solved by the following objective function:

$$\left(\widehat{\rho}, \widehat{\sigma}_v^2\right) = \arg\min\left\{\left[G_N\left(\rho, \rho^2, \sigma_v^2\right) - g_N\right]'\left[G_N\left(\rho, \rho^2, \sigma_v^2\right) - g_N\right]\right\},$$
$$\rho \in [-a, a], \sigma_v^2 \in [0, b_v], \qquad (31)$$

where

$$G_N = \frac{1}{N} \begin{pmatrix} \dfrac{2}{(T-1)} \bar{E}\wedge' Q_0 \widehat{E} & -\dfrac{1}{(T-1)} \bar{E}\wedge' Q_0 \bar{\bar{E}} & 1 \\[2ex] \dfrac{2}{(T-1)} \bar{E}\wedge' Q_0 \bar{\bar{E}} & -\dfrac{1}{(T-1)} \bar{\bar{E}}\wedge' Q_0 \bar{\bar{E}} & tr\left(M'M\right) \\[2ex] \dfrac{1}{(T-1)}\left(\bar{E}\wedge' Q_0 \bar{\bar{E}} + \bar{E}\wedge' Q_0 \widehat{E}\right) & -\dfrac{1}{(T-1)} \bar{E}\wedge' Q_0 \bar{\bar{E}} & 0 \end{pmatrix}, g_N$$
$$= \frac{1}{N} \begin{pmatrix} \dfrac{1}{(T-1)} E\wedge' Q_0 \widehat{E} \\[2ex] \dfrac{1}{(T-1)} \bar{E}\wedge' Q_0 \bar{E} \\[2ex] \dfrac{1}{(T-1)} E\wedge' Q_0 \bar{E} \end{pmatrix}. \qquad (32)$$

2.3.2. Estimation of σ_u^2. The fourth moment condition given by Kapoor et al. (2007) [18] is

$$E\left(\varepsilon' Q_1 \varepsilon\right) = NT\sigma_u^2 + N\sigma_v^2 = N\sigma_1^2, \qquad (33)$$

where $\sigma_1^2 = T\sigma_u^2 + \sigma_v^2$. However, considering the characteristics of the stochastic frontier model, the compound error term ε obeys the asymmetric distribution of the expected nonzero, so the moment condition (33) cannot be directly applied, and the following formula can be proved:

$$COV\left(\varepsilon' Q_1 \varepsilon\right) = NT\sigma_u^2 + N\sigma_v^2 = N\sigma_1^2. \qquad (34)$$

The estimator of parameter σ_1^2 can be obtained by combining Equation (34) with Equation (31):

$$
\begin{aligned}
\widehat{\sigma}_1^2 &= \frac{1}{N} COV\left[\left(E\wedge - \rho\wedge\bar{E}\wedge\right)' Q_1\left(\widehat{E} - \widehat{\rho}\widehat{\bar{E}}\right)\right] \\
&= \frac{1}{(T-1)} COV\left(E\wedge' Q_1\widehat{E}\right) - \frac{2}{(T-1)} COV\left(\bar{E}\wedge' Q_1\widehat{E}\right) \\
&\quad \cdot \widehat{\rho} - \frac{1}{(T-1)} COV\left(\bar{E}\wedge' Q_1\bar{\widehat{E}}\right)\rho\wedge^2.
\end{aligned}
$$

(35)

Substitute $\widehat{\sigma}_v^2$ and $\widehat{\sigma}_1^2$ into (33) to obtain the moment estimator $\widehat{\sigma}_u^2$ of σ_u^2.

2.4. *Spatial Correction of Estimators.* Although it can be proved that the estimator (21) is a consistent estimator, it can also be proved that the consistency of the GMM estimator can be guaranteed even if the model has spatial error autocorrelation. However, the estimator (21) cannot solve the spatial dependence of the error term, and the variance of the estimator is relatively large. After obtaining the consistent estimator of ρ by Equation (31), the consistent estimator can be obtained by a correcting transformation. According to the spatial correction method given by Jacobs et al. (2009) [19], the estimator obtained in the first step was corrected.

The estimator $\widehat{\rho}$ obtained from Equation (31) is used to construct matrix $I - \widehat{\rho}M$, and left the difference GMM and the explained variables and the instrumental variables matrix of the system GMM estimation, if

$$
\begin{aligned}
\tilde{Y} &= (I - \widehat{\rho}M)Y, \Delta\tilde{Y}_N = (I - \widehat{\rho}M)\Delta Y_N, \tilde{Y}_{-1} \\
&= (I - \widehat{\rho}M)Y_{-1}, \tilde{W}Y = (I - \widehat{\rho}M)WY, \tilde{X} = (I - \widehat{\rho}M)X.
\end{aligned}
$$

(36)

The corresponding explanatory variable set $Z_t = [WY_t, WY_{t-1}, Y_{t-1}, X_t]$ is corrected as

$$
\tilde{Z} = (I - \widehat{\rho}M)Z*_N, \Delta\tilde{Z} = (I - \widehat{\rho}M)\Delta Z_N.
$$

(37)

The instrumental variable matrix and weight matrix corresponding to the GMM estimation of the spatial system are corrected as follows:

$$
\begin{aligned}
\tilde{H}_{SBB} &= (I - \widehat{\rho}M)H_{N,BB}, \\
\tilde{A}_{SBB} &= \left(\tilde{H}'_{SBB}\tilde{H}_{SBB}\right)^{-1}.
\end{aligned}
$$

(38)

Then, the corrected system GMM estimator is

$$
\widetilde{\theta}_{SBB} = \left(\tilde{Z}'_{SBB}\tilde{H}_{SBB}\tilde{A}_{SBB}\tilde{H}'_{SBB}\tilde{Z}_{SBB}\right)\tilde{Z}'_{SBB}\tilde{H}_{SBB}\tilde{A}_{SBB}\tilde{H}'_{SBB}\tilde{Y}_{SBB}.
$$

(39)

3. Results and Discussion

3.1. Properties of the Estimators

3.1.1. Properties of the Estimator $\widetilde{\theta}_{SBB}$.
According to the Extremum Consistency Theorem [30] (see Appendix B), the estimators $\widehat{\theta}_{SBB}$ and $\widetilde{\theta}_{SBB}$, obtained by Equations (21) and (27), are consistent.

Proof. see Appendix C. □

3.1.2. Properties of the Estimators $\widehat{\rho}$, $\widehat{\sigma}_v^2$, $\widehat{\sigma}_u^2$, and $\widetilde{\theta}_{SBB}$.
It can be proved that the estimators $\widehat{\rho}$, $\widehat{\sigma}_v^2$, $\widehat{\sigma}_u^2$, and $\widetilde{\theta}_{SBB}$ are consistent. The proof of the consistency of $\widehat{\rho}$, $\widehat{\sigma}_v^2$, and $\widehat{\sigma}_u^2$ is similar to that in Kapoor et al. (2007) [18], and it is omitted here. The consistency of $\widetilde{\theta}_{SBB}$ can be derived from the consistency of $\widehat{\rho}$, $\widehat{\sigma}_v^2$, and $\widehat{\theta}_{SBB}$. Similar to the consistency of GLS estimates, the correcting transformation does not affect the consistency of the estimator $\widetilde{\theta}_{SBB}$.

3.2. *Crossvalidation Scheme and Selection of Spatial Weight Matrix.* In order to avoid affecting the accuracy of model estimation due to the choice of spatial weight matrix, the optimal spatial weight matrix was selected by crossvalidation. This is a widely used model selection and generalization method in machine learning. However, since the data used is panel data and the model used is spatio-temporal model, the structural features of spatio-temporal data may be destroyed if the training set and validation set are generated by hold-out method and LOOCV or K-folded crossvalidation. Therefore, this paper considers a stratified crossvalidation approach. For the spatio-temporal data, if N and T are assumed to be the number of spatial individuals and the number of periods contained in the observed samples, respectively, and the rest of the conventions on independent variables and dependent variables are the same as Equation (1), stratified crossvalidation can include the following three forms.

3.2.1. *Leave-One-Out Crossvalidation for the Time Dimension (TLOOCV).* Select the date t as the validation set and the rest $T - 1$ of the date as the training set. Let $C_1, C_2, \cdots, C_T$ denote, respectively, the index values of the observations contained in period t ($t = 1, 2, \cdots, T$), and $N_1, N_2, \cdots, N_T$ the number of observations contained in period t ($t = 1, 2, \cdots, T$). Let $n_1, n_2, \cdots, n_T$ denote the number of the observations in part t. Do the above for each period $t = 1, 2, \cdots, T$ in turn and calculate

$$
CV_T = \frac{1}{T}\sum_{t=1}^{T} \text{MSE}_t,
$$

(40)

where $\text{MSE}_t = \sum_{i=1}^{N_t}\left(y_{it} - y\wedge_{it}\right)^2/N_t$, and $\widehat{y}_{it}$ is the fitting value of the ith observed value in period Ty_{it}.

This crossvalidation method is suitable when the number of periods T is not too large.

3.2.2. *K-Fold Pooled Crossvalidation for the Spatial Dimension (SK-Fold PCV).* When the total number of
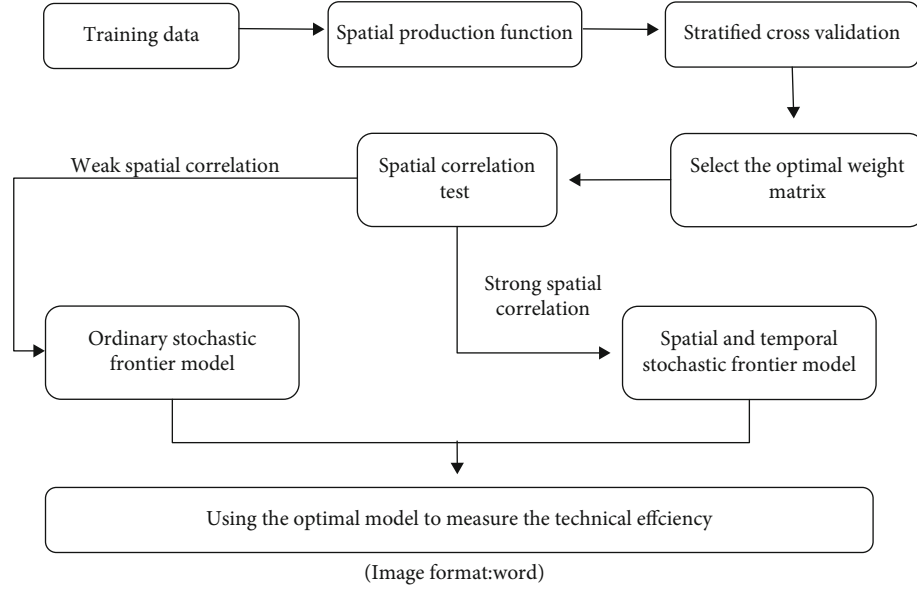
(Image format:word)

FIGURE 1: Spatio-temporal industrial efficiency measurement process.

periods T is large and the number of spatial individuals N is also large, this method is suitable for use. All observed values in each period t ($t = 1, 2, \cdots, T$) were randomly divided into K groups of equal size (i.e., the subsample size of each group was N/K), and a group was randomly selected from each period t ($t = 1, 2, \cdots, T$) to obtain NT/K observed values combined as the validation set and $N(K-1)T/K$ remaining observed values in each period combined as the training set. Do the above for each period sequence and calculate

$$CV_K = \frac{1}{K} \sum_{k=1}^{K} \text{MSE}_k, \qquad (41)$$

where $\text{MSE}_k = \sum_{i=1,t=1}^{N_t, T} K(y_{it} - y\wedge_{it})^2 / TN_t$, and $\hat{y}_{it}$ is the fitting value of the ith observed value in period Ty_{it}.

3.2.3. Leave-One-Out Crossvalidation for the Spatial Dimension (SLOOCV). When the crossvalidation method presented in Section 3.2.2 and the condition $K = N$ attached, it can be called K-fold pooled crossvalidation for the spatial dimension (SN-fold PCV) or stratified leave-one crossvalidation (SLOOCV). When the total number of periods T is large and the number of spatial individuals N is small, this method is suitable for use.

3.2.4. Determination of Weight Matrix and Industrial Efficiency Measure. To discuss industrial efficiency from the perspective of spatial statistics or spatial data mining, a good spatial weight matrix should be determined first. In this paper, the spatial lag production function is chosen as the basic model, and the spatial weight matrix involved in the construction of the model can take various alternative forms. In a data-driven way, the training samples were imported into the model for parameter estimation, and then, the most appropriate spatial weight matrix was determined by stratified crossvalidation. To determine whether the spa-

tial model is selected for analysis, the spatial correlation test is further carried out. If there is a strong spatial correlation, a spatial stochastic frontier model or a spatio-temporal stochastic frontier model will be established; if the spatial correlation is weak, an ordinary panel stochastic frontier model will be selected. After the estimation is completed, the best performing model is used to measure the technical efficiency. The flow chart of the entire analysis process is shown in Figure 1.

4. The Efficiency of the High and New Technology Industries in China

China has developed high and new technology industries for many years in order to transform the economic growth mode, cultivating knowledge and technology intensive new companies with great growth potential and low resources consumption that provide a sustainable development. As the technology of such industries disseminates, some issues emerge, such as spatial technology spillover, continuity of technological upgrade, and delay from research and development to market acceptance. In this paper, we analyze the efficiency of this kind of industries by the above analysis process.

4.1. Introduction to the Model and Data. In the framework of spatio-temporal model analysis, the spatial lag production function model based on the Cobb-Douglas production function is chosen as the basic model for weight selection. The matrix form of the model is as follows:

$$\ln Y = \rho W_N \ln Y + [\ln k, \ln l]B + u, \qquad (42)$$

where $\ln Y$, $\ln k$, and $\ln l$ are, respectively, the logarithm vectors of the main business income, the assets investment, and the mean number of employees of the high and new technology industries in every province in China; W_N is the spatial

Expression	Matrix name	Meaning
W_1	Position adjacency weight matrix	The weight matrix is constructed by rook, bishop, and queen position adjacence, and queen adjacence matrix is selected in this paper. Matrix element $w_{ij} = \begin{cases} 0 & \text{Region } i \text{ and } j \text{ are not adjacent} \\ 1 & \text{Region } i \text{ and } j \text{ are adjacent} \end{cases}$
W_2	Geographical distance weight matrix	The weight matrix is constructed by geographical distance between regions, and this paper constructs the weight matrix by reciprocal distance between the centers of provincial capitals in China. Matrix element $w_{ij} = 1/d_{ij}$, where d_{ij} is the geographical distance between regions i and j
W_3	Economic distance weight matrix	The weight matrix is constructed by the difference of economic level among different regions, that is, the smaller the economic gap, the stronger the spatial correlation. In this paper, the GDP of each province in China is used as the proxy variable of economic development level to construct the matrix Matrix element $w_{ij} = \begin{cases} 1/\left|\overline{\text{GDP}}_i - \overline{\text{GDP}}_j\right| & i \neq j \\ 0 & i = j \end{cases}$ ($\overline{\text{GDP}}_i$ represents the annual average GDP of region i)
W_4	Adjacency and distance combination weight matrix	Matrix element $w_{ij} = \begin{cases} 0 & \text{Region } i \text{ and } j \text{ are not adjacent} \\ 1/d_{ij} & \text{Region } i \text{ and } j \text{ are adjacent} \end{cases}$
W_5	Adjacency and economic combination weight matrix	Matrix element $w_{ij} = \begin{cases} 0 & \text{Region } i \text{ and } j \text{ are not adjacent} \\ 1/\left|\overline{\text{GDP}}_i - \overline{\text{GDP}}_j\right| & \text{Region } i \text{ and } j \text{ are adjacent} \end{cases}$
W_6	Weight matrix of distance and economy combination	Matrix element $w_{ij} = \left(1/d_{ij}\right) \cdot \left(\overline{\text{GDP}}_i/\overline{\text{GDP}}_j\right)$ (where d_{ij} is the geographical distance between region i and j; $\overline{\text{GDP}}_i$ represents the annual average GDP of region i)

weight matrix of the 31 provinces in China. Among many spatial weight matrices, we choose the three most commonly used spatial weight matrices in economic problems and their combination forms. Various weight matrices and their interpretations are shown in Table 1.

Before the fitting, each weight matrix was row-standardized, and the optimal weight matrix determined by crossvalidation was implemented to establish the stochastic frontier model. Starting from the ordinary panel stochastic frontier model and considering the spatial correlations, we construct the static panel spatial stochastic frontier model and the spatiotemporal stochastic frontier model. In order to determine if the model variables present spatial correlation, we let them go through a spatial correlation test, and to determine if there should be a time lag term in the model, we test the significance. Comparing the results of the three models and selecting the one that provides the best fit, we estimate the technology efficiency of the high and new technology industries in every province.

The matrix forms of the three models are as follows:
Ordinary panel stochastic frontier model:

$$\ln Y = [\ln k, \ln l]B + v - u. \tag{43}$$

Static panel space stochastic frontier model:

$$\ln Y = \lambda(I_T \otimes W_N)\ln Y + [\ln k, \ln l]B + v - u,$$
$$v = \rho(I_T \otimes M_N)v + \xi. \tag{44}$$

Spatiotemporal stochastic frontier model:

$$\ln Y = \lambda(I_T \otimes W)\ln Y + \pi(I_T \otimes W)\ln Y_{-1} + [\ln k, \ln l]B + E,$$
$$v = \rho(I_T \otimes M_N)v + \xi, \tag{45}$$

where v is a general vector of stochastic error, u is the inefficiency term, I_T is the identity matrix of order T, λ and ρ are the spatial correlation coefficients of the corresponding equations, π is the spatiotemporal time lag coefficient of the spatiotemporal stochastic frontier model, and B is the regression coefficient vector.

The development plan of high and new technology industry in China started in 1988, but due to the relatively slow progress in the beginning, the scale development of this industry did not start until the beginning of the twenty-first century. For this reason, we have chosen as research sample the panel data of the high and new technology industries of the 31 provinces in China from 2001 to 2018. The data of capital and labor input factors have been taken from "China high-tech industry yearbook." Descriptive statistics are shown in Table 2.

Figure 2 is the histogram drawn by taking the intragroup mean of data in each region according to year.

Figure 2 shows the difference in investment and average development level of high-tech industries in different provinces of China from 2001 to 2018. As can be seen from the figure, Guangdong, Jiangsu, and Shandong provinces have

TABLE 2: Descriptive statistics of output value and factor input index of high-tech industry.

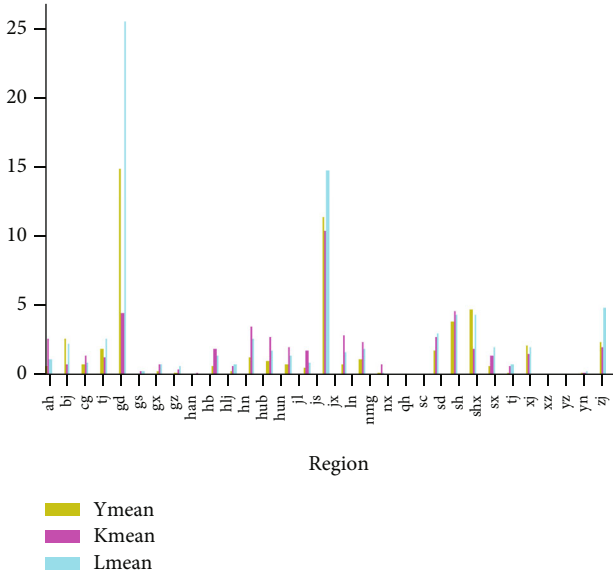| Variable	Average	Standard deviation	Median	Maximum
Y	30.283	53.712	13.615	237.170
k	31.682	54.757	15.884	241.980
l	39.477	75.530	18.817	372.000

FIGURE 2: Histogram of the mean value of output value and factor input of high-tech industries in each province.

the highest input and output levels of high-tech industries, and the differences among these three provinces are also very large. In the past 18 years, the average output value of the high-tech industry of Guangdong, which ranks first, reached 23.7 billion yuan, while that of Shandong, which ranks third, reached only 6.997 billion yuan, less than one third of that of Guangdong. In terms of the development and distribution of high-tech industries nationwide, the gap between provinces is even more obvious. The average output value of the high-tech industries in Tibet, which ranks the last, is only 0.085 billion yuan, less than 1/1000 of that of Guangdong.

4.2. Crossvalidation Results and the Selection of Weight Matrix. According to the characteristics of the data obtained, the time limit contained in the data is 18 years, which is relatively short and smaller than the number of regions. Therefore, the leave-one-out crossvalidation for the time dimension (TLOOCV) method was chosen. Each weight matrix in Table 1 was introduced into model (31), the training set data were imported into the model one by one for fitting, and then, Equation (40) was calculated to obtain the CV statistics corresponding to each weight matrix. The calculation results are shown in Table 3.

By comparing the calculation results of CV statistics of validation set, it can be found to be the optimal spatial weight matrix required by this paper.

4.3. Empirical Results. To ensure that spatial econometrics is applicable to the problem we are studying, we need to test the spatial correlation of the variables we are interested in. The most popular method to measure spatial autocorrelation is Moran's index I (Moran's I):

$I = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})/S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$, where S^2 is the sample variance, w_{ij} is the (i, j) element of the spatial weight matrix (used to measure the distance between region i and region j), and $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$ is the sum of all spatial weights.

The value of Moran's I is generally between -1 and 1, and its greater than 0 indicates positive autocorrelation. That is, the high value is adjacent to the high value and the low value is adjacent to the low value. Less than 0 means negative autocorrelation. That is, a high value is adjacent to a low value. If the Moran's I is close to 0, then the spatial distribution is random, and there is no spatial autocorrelation.

To test the existence of spatial correlations in the variables of the high and new technology industry, we calculate the global Moran's I indices of the production values of the industry from 2011 to 2018 (Table 4).

From the results of Moran's I index calculation, we found that the P value of the index is smaller than 0.01 for every year, demonstrating that the index is significant below 1% for every year, and the average Moran's I index is also significant for every year. The Moran's I index reached the minimum value 0.286 in 2011 and the maximum value 0.340 in 2013. We observe that the production value of the high and new technology industries of every province shows significant spatial correlation for every year and conclude that the production values of these industries of the provinces in China have apparent spatial aggregating effects. Furthermore, the testing of spatial correlation and the location quotient calculation both demonstrated that the high and new technology industries in different regions of China have apparent spatial correlation. We therefore choose the panel space stochastic frontier model for the analysis.

Table 5 presents the estimation results of the static, spatiotemporal stochastic frontier model, where the static model was further analyzed by considering fixed and stochastic effects.

The spatial autoregressive coefficients λ and ρ of the three models can all pass the significance test. From this and the spatial correlation test, one can conclude that the spatial stochastic frontier model is more reasonable. The estimated spatial autoregressive coefficients of the three models are all positive, which implied that the spatial effects have a positive impact on the development of the high and new technology industries. The negative value of the Hausman statistics of the static panel space model implies that the random effect model should be chosen. The random effect σ_u^2 value of the static panel space stochastic frontier model is far greater than the σ_v^2 value, and the value of γ is 0.719, implying that there apparently exists technical inefficiency. Comparing the spatiotemporal stochastic frontier model with the estimation of the static panel spatial random effect, the spatiotemporal lag coefficient of the former can

TABLE 3: CV statistics corresponding to various weight matrices.

Weight matrices	W_1	W_2	W_3	W_4	W_5	W_6
CV statistics	526.932	403.9296	333.7308	157.5527	320.589	104.8576

TABLE 4: Moran's I indices among the production values of the high and new technology industry.

	Index	P value	Index	P value	Index	P value	Index	P value
Year		2011		2012		2013		2014
Moran I	0.286	$P \leq 0.01$	0.321	$P \leq 0.001$	0.340	$P \leq 0.001$	0.302	$P \leq 0.01$
Year		2015		2016		2017		2018
Moran I	0.312	$P \leq 0.01$	0.298	$P \leq 0.01$	0.305	$P \leq 0.01$	0.317	$P \leq 0.01$

TABLE 5: Estimation results by the static panel space and spatiotemporal stochastic frontier model.

Dependent variable			Ln y			
Model		Static panel spatial stochastic frontier			Spatiotemporal stochastic frontier	
Variable		Fixed effects		Random effects		
	Coefficients	P value	Coefficients	P value	Coefficients	P value
ln k	0.397***	0.001	0.401**	0.001	0.405**	0.008
ln l	0.686***	0.008	0.713**	0.017	0.718***	0.005
λ	0.435**	0.012	0.576***	0.002	0.369***	0.005
π	—	—	—	—	0.205**	0.011
ρ	0.551***	0.007	0.698**	0.013	0.585**	0.017
σ_u^2	2.65E-05		0.337		0.656	
σ_v^2	0.387		0.132		0.180	
γ	6.850E-04		0.719		0.786	
$Hausman$ statistics	-39282.77				D statistics	10.668**

In the table, ** indicates those that can pass the test with 5% significance level.

TABLE 6: Average technological efficiency of high-tech industries in 31 provinces.

Province	TE	Province	TE	Province	TE	Province	TE
Anhui	0.817	Guizhou	0.818	Hunan	0.812	Ningxia	0.786
Beijing	0.985	Hainan	0.876	Jilin	0.835	Qinghai	0.733
Chongqing	0.912	Hebei	0.786	Jiangsu	0.951	Sichuan	0.851
Fujian	0.936	Heilongjiang	0.835	Jiangxi	0.792	Shandong	0.876
Gansu	0.932	Henan	0.752	Liaoning	0.931	Shanghai	0.966
Guangdong	0.755	Hubei	0.851	Neimenggu	0.872	Shănxi	0.795
Guangxi	0.785	Shānxi	0.787	Tianjin	0.975	Xinjiang	0.651
Tibet	0.668	Yunnan	0.815	Zhejiang	0.952		

Data source: calculated based on pattern II stochastic effects model of a function.

pass the 10% significance test, obtaining that the spatiotemporal lag term in the model has a significant function. The distance statistic of the spatiotemporal stochastic frontier model pass the 5% significance distance test, proving that the spatiotemporal stochastic frontier model is globally significant. Moreover, the estimation of the γ value of the spatiotemporal stochastic frontier model is higher than that of the static panel spatial stochastic frontier model, demonstrating that the inefficiency term of the spatiotemporal stochastic frontier model has a more significant function. All the impact factor variables of the technical inefficiency terms of the analysis of the two models can at least pass the 5% significance test, and the signs of the regression coefficient of the two models are consistent, the numerical values are relatively close. Taken together, the above results all demonstrate that the analysis of the spatiotemporal stochastic frontier model is more reasonable and the development of the high and new technology industry has positive correlations in space and time. Due to space limitation, this paper does not report the annual technical efficiency of the high-

tech industries of each province. The spatial-temporal stochastic frontier model is used to calculate the average technical efficiency of the high-tech industries of each province from 2001 to 2018 as follows.

It can be seen from the calculation results in Table 6 that the average technical efficiency value of the high-tech industries in all provinces in China is less than 1, which indicates that the actual output of the high-tech industries in all provinces has not reached the most effective output level, and there is technological inefficiency in production. The five-year national average technical efficiency level was 0.837, and there are obvious regional differences in the technical efficiency values presented in Table 4. Nine provinces (Beijing, Chongqing, Fujian, Gansu, Jiangsu, Liaoning, Shanghai, Tianjin, and Zhejiang) achieved an average technical efficiency of more than 0.9, seven provinces are located in the eastern region, one in the central region, and one in the western region. There are 11 provinces with average technical efficiency below 0.8, namely, Guangdong, Guangxi, Guizhou, Hebei, Henan, Jiangxi, Ningxia, Qinghai, Shaanxi, Xinjiang, and Tibet. Only one of the provinces is in the east, six in the central region, and four in the west.

5. Conclusion

In this paper, taking into account that the variables to be explained might be affected by the time lag term and the space-time interaction, we develop a dynamic model within the framework of the panel spatial stochastic frontier model. Due to the apparent endogeneity of the model, we use the systematic GMM method to estimate the parameters, choose suitable tool variables according to the model assumptions and variable characteristics, and construct the suitable spatiotemporal stochastic frontier model. We use the extreme value consistence theorem and the uniform law of large numbers (ULLN) to prove the consistency of the structural parameter estimators and of the estimators of the error term distribution parameters. Aiming at the selection of spatial weight matrix of spatio-temporal model, a stratified crossvalidation method is designed to select the most appropriate spatial weight matrix in a data-driven way according to the characteristics of spatio-temporal data. Although the spatial weight matrix selected by supervised learning may not be suitable for analyzing all problems, this data-driven model selection method is undoubtedly valuable and efficient.

From the analysis of the stochastic frontier model of high-tech industries in China and the measurement of their technical efficiency, we can draw the following conclusions.

There is a spatial positive correlation in the development of high-tech industries between different regions of China. The positive correlation between the output values of these industries in different regions has been obtained by calculating the Global Moran's I index in each year. The estimation of the spatial panel stochastic frontier model also indicates that the spatial autoregressive coefficient is positive, proving the existence of such a positive correlation which has a positive impact on the development of high-tech industries. There is also a spatial agglomeration effect and a spatial and temporal lag effect in these industries, illustrating that both static spill over

and dynamic continuity occur in the development of the high-tech industries in China. The technical efficiency of high-tech industries is relatively low. The strategic emerging industries started earlier in eastern region, but developed more slowly than in the central and western regions.

The Chinese economy is at a critical stage of replacing old drivers of growth with new ones and transforming and upgrading industries. The new round of technological and industrial revolution 5.0 has given rise to new technologies, new industries, new forms of business, and new models. In this study, the data mining algorithm based on stochastic frontier is used to calculate industrial efficiency, which is not only suitable for high-tech industry but also helpful to further enrich the research on the efficiency of new industry and new mode and has certain practical significance to promote the steady development of the new round of scientific and technological revolution of industry 5.0.

Appendix

A. Proof

$E_t = [E_{1t}, \cdots, E_{Nt}]'$ is an n-dimensional nonzero composite random error term vector, and E_t is set as interindividual nonautocorrelation according to classical econometric assumptions for simplicity.

$$COV\{[WA^{-1}]E_t, E_t\} = \sum_i \sum_j \delta_{ij} \text{cov}(E_{it}, E_{jt}) \\ = \sum_i \delta_{ii} \text{var}(E_{it}), \tag{A.1}$$

where δ_{ij} is the element of the matrix WA^{-1}. For any E_{it}, var $(E_{it}) > 0$, then the sufficient and necessary condition for the above formula to be 0 is $WA^{-1} = 0$, which is obviously inconsistent with the assumption of spatial weight matrix in this paper. Therefore, it is proved that

$$COV\{[WA^{-1}]E_t, E_t\} \neq 0. \tag{A.2}$$

B. Extremum Consistency Theorem

If (1) (identification) $Q(\theta)$ is uniquely maximized at θ_0, (4) (compactness) the parameter space Θ is compact, (8) (continuity) $Q(\theta)$ is continuous, and (9) (uniform convergence in probability) $\sup_{\theta \in \Theta} | \widehat{Q}(\theta) - Q(\theta) | \longrightarrow^p 0$, then $\widehat{\theta} = \arg\max_{\theta \in \Theta} \widehat{Q}(\theta) \longrightarrow^p \theta_0$.

C. Proof

From (25), we let $\widehat{H}(\theta) = (1/N)[H'_{N,BB}(\Delta Y*_N - \Delta Z*_N \theta)]$ and $H(\theta) = E[H'_{BB}(\Delta Y* - \Delta Z* \theta)]$, which are continuous functions of the corresponding parameters constructed by sample and global moment conditions, respectively. $\widehat{A} = A_{N,BB}$ and $A = A_{BB} = E[H'_{BB}G_{BB}H_{BB}]^{-1}$ are the weight matrix sample constructed and global constructed, respectively,

where $\theta = (\lambda_1, \lambda_2, \gamma, B)'$ is the structural parameter vector of the model. The above settings satisfy the following conditions:

(i) When $\theta = \theta_0$, $H(\theta) = E[H'_{BB}(\Delta Y * - \Delta Z * \theta)] = 0$, and $\widehat{A}$ and A are both positive definite matrices

(ii) Θ, the parameter space of θ, is compact, and $\theta_0 \in \Theta$

(iii) For an arbitrary point $\theta \in \Theta$, $H(\theta)$ is a continuous function

(iv) $E \sup_{\theta \in \Theta} \|H(\theta)\| < \infty$. Let

$$
\begin{aligned}
\widehat{Q}(\theta) &= -\widehat{H}(\theta)' \widehat{A} \widehat{H}(\theta), \\
Q(\theta) &= -H(\theta)' A H(\theta)
\end{aligned} \tag{C.1}
$$

We first prove uniqueness of the maximum value of $Q(\theta)$.

As $H(\theta_0) = 0$, one has $Q(\theta_0) = 0$, which is the maximum value of $Q(\theta)$ given that A is a positive definite matrix. From the uniqueness of the true parameter value, for $\theta \neq \theta_0$, it is satisfied that

$$
H(\theta) \neq 0, Q(\theta) = -H(\theta)' A H(\theta) < 0, \tag{C.2}
$$

and therefore, $\theta = \theta_0$ is the unique maximum of $Q(\theta)$.

According to the uniform law of large numbers (ULLN), when conditions (ii), (iii), and (iv) hold, $E[H(\theta)]$ and $Q(\theta)$ are continuous functions, and $\sup_{\theta \in \Theta} \|\widehat{H}(\theta) - H(\theta)\| \longrightarrow^p 0$. Noticing the structure of the weight matrices $\widehat{A}$ and A, and the fact that the instrumental variable matrices $H'_{N,BB}$ and H'_{BB} can be regarded as $\widehat{H}(\theta)$ and $H(\theta)$ under the condition $\Delta Y * - \Delta Z * \theta \equiv 1$, one has $\widehat{A} \longrightarrow^p A$.

$$
\begin{aligned}
\left|\widehat{Q}(\theta) - Q(\theta)\right| &= \left|\widehat{H}(\theta)' \widehat{A} \widehat{H}(\theta) - H(\theta)' A H(\theta)\right| \\
&= \left|\widehat{H}(\theta)' (\widehat{A} - A + A) \widehat{H}(\theta) - H(\theta)' A H(\theta)\right| \\
&= \left|\widehat{H}(\theta)' (\widehat{A} - A) \widehat{H}(\theta) + \widehat{H}(\theta)' A \widehat{H}(\theta)\right. \\
&\quad \left. - H(\theta)' A H(\theta)\right| = \left|\widehat{H}(\theta)' (\widehat{A} - A) \widehat{H}(\theta)\right. \\
&\quad + [H \wedge (\theta) - H(\theta) + H(\theta)]' A [\widehat{H}(\theta) \\
&\quad - H(\theta) + H(\theta)] - H(\theta)' A H(\theta) \Big| \\
&= \left|\widehat{H}(\theta)' (\widehat{A} - A) \widehat{H}(\theta) + [H \wedge (\theta)\right. \\
&\quad - H(\theta)]' A [\widehat{H}(\theta) - H(\theta)] + 2[H \wedge (\theta) \\
&\quad - H(\theta)]' A H(\theta) + H(\theta)' A H(\theta) \\
&\quad - H(\theta)' A H(\theta) \Big| = \left|\widehat{H}(\theta)' (\widehat{A} - A) \widehat{H}(\theta)\right. \\
&\quad + [H \wedge (\theta) - H(\theta)]' A [\widehat{H}(\theta) - H(\theta)] \\
&\quad + 2[H \wedge (\theta) - H(\theta)]' A H(\theta) \Big|,
\end{aligned} \tag{C.3}
$$

where in the last line, we have applied the triangle inequality.

$$
\begin{aligned}
&\leq \left|\widehat{H}(\theta)' (\widehat{A} - A) \widehat{H}(\theta)\right| + \left|[H \wedge (\theta) - H(\theta)]' A [\widehat{H}(\theta)\right. \\
&\quad \left. - H(\theta)]\right| + 2\left|[H \wedge (\theta) - H(\theta)]' A H(\theta)\right|
\end{aligned} \tag{C.4}
$$

Taking supreme on both sides of the above inequality, we obtain

$$
\begin{aligned}
\sup_{\theta \in \Theta} \left|\widehat{Q}(\theta) - Q(\theta)\right| &\leq \sup_{\theta \in \Theta} \left|\widehat{H}(\theta)' (\widehat{A} - A) \widehat{H}(\theta)\right| \\
&\quad + \sup_{\theta \in \Theta} \left|[H \wedge (\theta) - H(\theta)]' A [\widehat{H}(\theta) - H(\theta)]\right| \\
&\quad + 2 \sup_{\theta \in \Theta} \left|[H \wedge (\theta) - H(\theta)]' A H(\theta)\right| \\
&\leq \sup_{\theta \in \Theta} \left\|H \wedge (\theta)'\right\|^2 \|\widehat{A} - A\| \\
&\quad + \sup_{\theta \in \Theta} \|H \wedge (\theta) - H(\theta)\|^2 \|A\| \\
&\quad + \sup_{\theta \in \Theta} \|\widehat{H}(\theta) - H(\theta)\| \\
&\quad + 2 \sup_{\theta \in \Theta} \|\widehat{H}(\theta) - H(\theta)\| \cdot \|A\| \\
&\quad \cdot \sup_{\theta \in \Theta} \|H(\theta)\| \xrightarrow{p} 0
\end{aligned} \tag{C.5}
$$

According to the Extremum Consistency Theorem, we have $\widehat{\theta}_{\text{SBB}} \longrightarrow^p \theta$.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. Aigner, K. C. A. Lovell, and P. Schmidt, "Formulation and estimation of stochastic frontier production function models," *Journal of Econometrics*, vol. 6, no. 1, pp. 21–37, 1977.

[2] W. Meeusen and J. van den Broeck, "Technical efficiency and dimension of the firm: some results on the use of frontier production functions," *Empirical Economics*, vol. 2, no. 2, pp. 109–122, 1977.

[3] G. E. Battese and G. S. Corra, "Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia," *Australian Journal of Agricultural and Resource Economics*, vol. 21, no. 3, pp. 169–179, 1977.

[4] V. Druska and W. C. Horrace, "Generalized moments estimation for spatial panel data: Indonesian rice farming," *American Journal of Agricultural Economics*, vol. 86, no. 1, pp. 185–198, 2004.

[5] E. Affuso, "Spatial Autoregressive Stochastic Frontier Analysis: An Application to an Impact Evaluation Study," 2010, SSRN 1740382.

[6] A. Tonini and V. Pede, "A generalized maximum entropy stochastic frontier measuring productivity accounting for spatial dependency," *Entropy*, vol. 13, no. 11, pp. 1916–1927, 2011.

[7] F. Vidoli, C. Cardillo, E. Fusco, and J. Canello, "Spatial nonstationarity in the stochastic frontier model: an application to the Italian wine industry," *Regional Science and Urban Economics*, vol. 61, pp. 153–164, 2016.

[8] E. G. Tsionas and P. G. Michaelides, "A spatial stochastic frontier model with spillovers: evidence for Italian regions," *Scottish Journal of Political Economy*, vol. 63, no. 3, pp. 243–257, 2016.

[9] A. Carvalho, "Efficiency spillovers in Bayesian stochastic frontier models: application to electricity distribution in New Zealand," *Spatial Economic Analysis*, vol. 13, no. 2, pp. 171–190, 2018.

[10] M. Adetutu, A. J. Glass, K. Kenjegalieva, and R. C. Sickles, "The effects of efficiency and TFP growth on pollution in europe: a multistage spatial analysis," *Journal of Productivity Analysis*, vol. 43, no. 3, pp. 307–326, 2015.

[11] F. Jin and L.-f. Lee, "Asymptotic properties of a spatial autoregressive stochastic frontier model," *Journal of Spatial Econometrics*, vol. 1, no. 1, pp. 1–40, 2020.

[12] L. Kutlu, K. C. Tran, and M. G. Tsionas, "A spatial stochastic frontier model with endogenous frontier and environmental variables," *European Journal of Operational Research*, vol. 286, no. 1, pp. 389–399, 2020.

[13] A. S. Bergantino, M. Intini, and N. Volta, "Competition among airports at worldwide level: a spatial analysis," *Transportation Research Procedia*, vol. 45, no. 1, pp. 621–626, 2020.

[14] T. D. Graaff, "On the estimation of spatial stochastic frontier models: an alternative skew-normal approach," *The Annals of Regional Science*, vol. 64, no. 2, pp. 267–285, 2020.

[15] L. Jia-Xian, L. Zhi-He, and L. Guang-Ping, "Spatial Panel Stochastic Frontier Model and Technical Efficiency Estimation," *Journal of Business Economics*, vol. 223, no. 5, pp. 71–78.

[16] N. Deepa, Q. V. Pham, D. C. Nguyen et al., "A survey on blockchain for big data: approaches, opportunities, and future directions," 2020, https://arxiv.org/abs/2009.00858.

[17] H. H. Kelejian and I. R. Prucha, "A generalized moments estimator for the autoregressive parameter in a spatial model," *International Economic Review*, vol. 40, no. 2, pp. 509–533, 1999.

[18] M. Kapoor, H. H. Kelejian, and I. R. Prucha, "Panel data models with spatially correlated error components," *Journal of Econometrics*, vol. 140, no. 1, pp. 97–130, 2007.

[19] J. P. Jacobs, J. E. Ligthart, and H. Vrijburg, "Dynamic panel data models featuring endogenous interaction and spatially correlated errors," *SSRN Electronic Journal*, vol. 92, pp. 1–37, 2009.

[20] L. Anselin and D. A. Griffith, "Do spatial effecfs really matter in regression analysis?," *Papers in Regional Science*, vol. 65, no. 1, pp. 11–34, 1988.

[21] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna et al., "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, pp. 54776–54788, 2020.

[22] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40–79, 2010.

[23] S. Jun, "Linear model selection by cross-validation," *Journal of the American Statistical Association*, vol. 88, no. 422, pp. 486–494, 1993.

[24] T. W. Anderson and C. Hsiao, "Estimation of Dynamic Models With Error Components," *Journal of American Statistical Association*, vol. 76, no. 375, pp. 598–606, 1981.

[25] M. Arellano and S. Bond, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic studies*, vol. 58, no. 2, pp. 277–297, 1991.

[26] M. Arellano and O. Bover, "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, vol. 68, no. 1, pp. 29–51, 1995.

[27] R. Blundell and S. Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, vol. 87, no. 1, pp. 115–143, 1998.

[28] H. H. Kelejian and D. P. Robinson, "A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model," *Papers in Regional Science*, vol. 72, no. 3, pp. 297–312, 1993.

[29] T. Beck and R. Levine, "Stock markets, banks, and growth: Panel evidence," *Journal of Banking & Finance*, vol. 28, no. 3, pp. 423–442, 2004.

[30] W. Newey and D. McFadden, "Chapter 36 Large sample estimation and hypothesis testing," in *Handbook of Econometrics, vol. 4*, R. Engle and D. McFadden, Eds., Elsevier B.V, New York: North-Holland, 1994.