SDM Model Based on Stata to Analyze the Impact of Industrial Agglomeration on Logistics Efficiency

Min Qian[1, a], Meiyuan Jin [2, b]
[a]e-mail: 1529956548@qq.com,*Corresponding author: [b]e-mail: 2218495706@qq.com

[1]School of Management, Xi'an University of Science and Technology, Xi’an, China

[2]School of Management, Xi’an University of Science and Technology, Xi’an, China

Abstract—In this paper, the data of 12 provinces and cities in western China from 2010 to 2020 are taken as samples. Firstly, super-SBM model is used to evaluate logistics efficiency and analyze its current situation. Stata software and SDM model are used to study the impact of industrial agglomeration on logistics efficiency. The results show that the western region of China logistics efficiency has obvious spatial unbalance. There is significant spatial correlation between industrial agglomeration and logistics efficiency.

Keywords—SDM model; Super - SBM model; Spatial correlation analysis; Industrial agglomeration; Logistics efficiency

1 INTRODUCTION

With the Silk Road economic belt to the north and the new land and sea passage to the south, the logistics industry in the western region, as a bridge and hub connecting China and many neighboring countries, has a unique position and role. But as a result of western provinces geographical location, natural conditions, economic development level, and the differences degree of opening to the outside world, and the gap is still evident in eastern and middle, the level of logistics industry development is not balanced.

The development level of logistics industry is an important factor affecting the development of trade, and logistics efficiency is an important indicator to judge the overall development of logistics industry. Improving logistics efficiency is crucial to economic development. With the continuous improvement of infrastructure in western China, various kinds of port, airport, bonded and comprehensive logistics parks have emerged. The logistics industry shows a trend of agglomeration in space, and industrial agglomeration contributes to the improvement of logistics efficiency. This paper studies the development of industrial agglomeration and logistics efficiency in western China and explores the impact of logistics industry agglomeration on logistics efficiency in order to deeply understand the status of logistics industry in western China.

2 RELEVANT THEORETICAL FOUNDATIONS

2.1 Spatial econometric analysis

According to the first law of geography, all things and phenomena do not exist independently,

ICICA 2022, December 02-04, Chongqing, People's Republic of China

DOI 10.4108/eai.2-12-2022.2328018

but are associated with other things, with inherent attributes such as spatial dependence or spatial correlation. Traditional statistical analysis methods treat each unit as an independent sample and ignore the essential characteristics of cross section units, so there may be errors in the calculation results[1].Without considering the spatial effect, Panel data model is a common econometric model considering the characteristics of time and section. With the development of spatial econometrics, scholars organically combine the advantages of panel data model and spatial econometrics model, and the spatial panel model is proposed and has been widely used, mainly used to solve the problem of complex space effect. Spatial econometric analysis is a method for spatial data statistical analysis. Different from the traditional method, it includes spatial factors. It believes that a certain geographical phenomenon or attribute on each unit is related to the same phenomenon or attribute on adjacent units. In the calculation, the geographical location factors are comprehensively considered, which makes the research and analysis more comprehensive and practical significance.

2.2 Literature Review

For the study of logistics efficiency, domestic and foreign scholars have some differences in the selection of research objects. Foreign scholars mainly focused on aviation industry and port logistics, and studied logistics efficiency from the industry level[2][3].While domestic scholars focus on logistics efficiency in some regions, generally using DEA or SFA model to research logistics efficiency[4][5].However, there is little research on the impact of agglomeration on logistics efficiency. The research of T. Tao (2016)[6] shows that there is a positive linear relationship between professional agglomeration and logistics efficiency, and diversified agglomeration has a non-linear impact on logistics efficiency, which decreases first and then increases. The research of J. Leng, (2019)[7] shows that there is a significant threshold effect between industrial agglomeration and logistics efficiency. With the improvement of industrial agglomeration, its positive impact on logistics efficiency continues to weaken. According to the literature review, there is still a gap in the research on the relationship between logistics industry agglomeration and logistics efficiency.

This paper analyzes the impact of logistics industry agglomeration on logistics efficiency in 12 western provinces and cities. In this paper, the space measurement model into the logistics industrial agglomeration effect on the efficiency of the logistics research, and to explore its spatial effect, so as to western efficient logistics industry coordinated development to provide decision-making reference.

2.3 Theoretical analysis

2.3.1 The promotion effect of industrial agglomeration on logistics efficiency

The profits brought by economies of scale attract the flow of production factors and promote a large number of homogeneous enterprises to tend to the same place when making location choices. As a transmission mechanism, the externalities generated by agglomeration affect all links of the industrial development process and logistics efficiency. To be specific: Firstly is labor market sharing. With the spatial agglomeration of logistics industry, it will naturally attract more labor employment, so as to provide a large-scale labor market for the agglomeration area or its affiliated area, and indirectly improve the labor skills and quality of the region. Secondly is value chain dynamic matching. The continuous improvement of logistics industry agglomeration level will promote the continuous refinement of industrial division of labor and further enhance the relevance of regional vertical industries. At the same time, with the continuous refinement of the division of labor, it will give full play to the industrial advantages, improve the added value of each link, and then improve the overall logistics benefits of the region. Third is knowledge spillovers among industries. With the continuous improvement of spatial agglomeration level, the knowledge factor spillover effect of logistics industry will increase effectively, which will further reduce the knowledge factor cost of vertically related industries in the region, so as to improve the logistics efficiency in the region.

2.3.2 The spatial spillover effect of industrial agglomeration on logistics efficiency

The impact of logistics industry agglomeration on logistics efficiency in adjacent areas. According to the spatial interaction theory, the regional logistics industry agglomeration is not independent of each other. First, industrial chain extension. The regional logistics industry agglomeration will further extend its own industrial chain and optimize the industrial layout in the agglomeration area, so as to transfer some industries to adjacent areas. Second, the circulation of elements is complementary. The local logistics industry cluster forms a large-scale trading market. In order to reduce production costs, relevant enterprises or departments will pursue lower cost production factors, and then absorb the resource factors and labor factors of adjacent areas. At the same time, the region will also transport excess commodity factors to adjacent areas, so as to reduce the commodity prices of adjacent areas, to maximize the full use of regional factors.

3 MODEL AND DATA

3.1 Model Building

3.1.1 Super-SBM model

Traditional DEA models measure the efficiency of decision-making units from both radial and angle, and can not evaluate the problem that multiple DMUs are 1.In order to solve the problem of distinguishing effective DMU in the model, Tone[8] further proposed the Super-SBM model based on the modified relaxation variable, and the efficiency value of effective DMU would be greater than 1, which could effectively solve the defect that it could not be evaluated and sorted.

$$\rho^* = min\frac{\frac{1}{m}\sum_{i=1}^{m}\overline{x_i}/x_{io}}{\frac{1}{s}\sum_{r=1}^{s}\overline{y_r}/y_{ro}} \tag{1}$$

$$\text{s.t. } \overline{x_{io}} \geq \sum_{j=1,j\neq 0}^{n} \lambda_j x_{ij}, i = 1,2,\cdots,\boldsymbol{m},$$

$$\overline{y_{ro}} \geq \sum_{j=1,j\neq 0}^{n} \lambda_j y_{rj}, i = 1,2,\cdots,\boldsymbol{s},$$

$$\overline{x_{io}} \geq x_{io}, i = 1,2,\cdots,m,$$

$$\overline{y_{ro}} \geq y_{ro}, i = 1,2,\cdots,s,$$

$$\lambda_j \overline{y_{ro}} \geq 0, j = 1,2,\cdots,n, j \neq 0$$

$$\overline{y_{ro}} \geq 0, r = 1,2,\cdots,s.$$

In Formula (1), $\overline{x_i}$ & $\overline{y_r}$ are the relaxation variables of input and output respectively, *m&s* represents the number of input &output indicators respectively, λ_j is the weight vector of the decision-making unit. When ρ^*≥0, DMU indicates effective; When $\rho^* < 1$, DMU indicates that it is invalid and needs to improve input and output.

3.1.2 Spatial correlation analysis

Global spatial autocorrelation is used to test the level of global spatial correlation of variables and verify whether the subsequent empirical research can adopt spatial econometric model. The specific calculation formula is as follows:

$$Moran's\ I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \quad (2)$$

In Formula(2), W_{ij} is the spatial weight matrix, x_i&x_j are the data of the research variable, $\bar{x}$&S are the mean and standard deviation respectively, *i&j* all indicate provincial numbers, *n* represents the number of samples. Moran′s I∈(-1,1), given the significance criteria, when Moran′s I > 0,it shows that there is a spatial positive correlation; Moran′s I < 0 is spatial negative correlation; Moran′s I = 0 is no spatial correlation with random distribution.

3.1.3 Spatial econometric model

SDM model nested SAR model and SEM model, which has a more general form. In this paper, the SDM model was first set and then tested to determine the use of the model.

SDM model:

$$LE_{i,t} = \alpha_{i,t} + \rho wLE_{i,t} + \beta_1 LQ_{i,t} + \beta_2 wLQ_{i,t} + \beta_k \sum_{3,5,\cdots 2k+1}^{n} X_{control} + \gamma_{k+1} \sum_{k+1}^{n} wX_{control} + \varepsilon_{i,t} \quad (3)$$

In Formula (3), $LE_{i,t}$ is logistics efficiency of *t-th(t=1,2,⋯,11)*year in *i(i=1,2,⋯,12)* province$\alpha_{i,t}$ is Constant term;ρ is the spatial autoregressive coefficient, which indicates the impact of the logistics efficiency ($wLE_{i,t}$) of spatially related provinces and urban areas on the logistics efficiency of urban areas of the province. $\rho > 0$,it indicates that the logistics efficiency of spatially related provinces and cities has a positive spillover effect on the logistics efficiency of urban areas of the province, on the contrary, there is a negative spillover effect. $w$is the spatial weight matrix (geographic distance spatial weight matrix is selected in this paper). $LQ_{i,t}$ is the level of logistics industry agglomeration, β_1 is elasticity coefficient of logistics industry agglomeration degree to logistics efficiency, $X_{control}$ is a series of control variables, β_k is elasticity coefficient of control variable; $\mu_{i,t}$ is Random error term; β_2 represents the elasticity coefficient of logistics industry agglomeration in spatially related provinces and urban areas to the logistics efficiency of the province, γ_{k+1} represents the elasticity coefficient of the spatially related provincial and urban control variable $wX_{control}$ to the logistics efficiency of the provincial and urban areas.

3.2 Selection of Indicators

In this paper, the three departments of transportation, warehousing and postal industry are used to represent the logistics industry. The data are from the China Statistical Yearbook and the statistical bulletin of national economic and social development over the years. The missing data are processed by interpolation.

3.2.1 Explained variable

Logistics efficiency (*LE*). According to the principles of scientificity, systematization and data availability, this paper is constructed from the perspective of input and output. The input includes the number of employees in the logistics industry (*X1*), the total fixed investment in the logistics industry (*X2*) and the density of transportation network (*X3*); Output includes added value of logistics industry (*Y1*) and freight turnover (*Y2*).

3.2.2 Core explanatory variable

Logistics industry agglomeration(*LQ*).This paper decides to use location entropy to measure the degree of logistics industry agglomeration in Western China, and improves it. The economic development level and labor population of each province are included in the index to comprehensively reflect the degree of industrial agglomeration.

$$lq_{ij} = \frac{q_{ij}/q_j}{q_i/q} \times \frac{p_{ij}/q_{ij}}{p_i/q_i} = \frac{p_{ij}/q_j}{p_i/q} \quad (4)$$

In Formula(4), *i* represents the logistics industry and j represents each province.lq_{ij} represents the agglomeration of logistics industry; q_{ij} represents the employees of industry i in j,q_i represents employees in the whole region of the i industry,q_j represents total employees in j,q represents number of employees in the whole region; p_{ij} represents output value of the *i* industry in j,p_i represents total regional output value of the *i* industry.

3.2.3 Control variable

In addition to the logistics industrial agglomeration, there are many factors that affect the efficiency of logistics, this paper introduced control variables include: The economic level (PGDP) is expressed in per capita GDP; The urbanization rate (URB) is expressed by the proportion of urban population in the total resident population; The process of industrialization (IND) is expressed by the proportion of the output value of the secondary industry in GDP; The degree of marketization (MAR) is expressed by the number of Enterprises Owned per million people; Government support (GOV) is expressed by the proportion of local financial transportation expenditure in local financial expenditure; Opening to the outside world (OPEN) is expressed in terms of total imports and exports.

4 EMPIRICAL ANALYSIS

4.1 Logistics efficiency evaluation results

In this paper, MaxDEA software is used to measure logistics efficiency. The DMU consists of 12 provinces and cities in western China, with 3 input indicators and 2 output indicators. It is evaluated based on super-SBM model. For reasons of length, the specific analysis of each province has been omitted and only regional efficiency is shown, as shown in Figure 1.

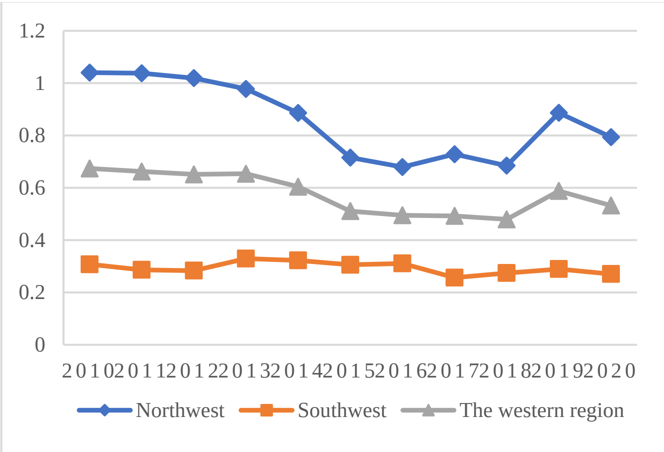

Figure 1 logistics efficiency in Western China (2010-2020)

In terms of the overall comprehensive efficiency value in southwest China is not effective, and the northwest China also fluctuates near 1. Specifically, in 2010, the overall efficiency of logistics industry was "Northwest > Southwest", and it will remain the same in 2020.In recent years, China has entered a critical period of structural adjustment, transformation and development. Domestic demand drives the development of e-commerce. The western region seizes development opportunities, actively develops modern logistics services, increases employment, increases investment in fixed assets, and improves the level of informatization. The differences in logistics efficiency between provinces and cities have been reduced.

4.2 Spatial correlation test results

Use Stata15.0 software to obtain the global Moran's I value and normal statistics Z value of logistics efficiency in the western region from 2010 to 2020, as shown in Table 1.

Table 1 Global Moran index of logistics efficiency

Year	Moran'sI	Z	p-value
2010	0.365	3.381	0.000
2011	0.340	3.305	0.000
2012	0.432	3.655	0.000
2013	0.500	3.797	0.000
2014	0.423	3.499	0.000
2015	0.105	2.396	0.008
2016	0.070	1.764	0.039
2017	0.008	1.798	0.036
2018	0.007	1.720	0.026
2019	0.356	2.673	0.004

2020	0.265	2.994	0.001

As can be seen from Table 1, the overall Moran's I index of logistics efficiency in western China has experienced a fluctuating development process of "rise-decline-rising", gradually changing from a scattered state to a concentrated state. Although Moran's I fluctuates, it is always positive, that is, there is a positive spatial correlation.

4.3 Model testing

The model is first set as SDM model. The choice of fixed effect model or random effect model in this study needs to be determined by Hausman test. The test result is: the Hausman statistical value is 5200, and the corresponding p-value is less than 0.01, indicating that it has passed the significance level test at the significance level of 1%, so the fixed effect should be selected for empirical test.

Then, Wald test and LR test need to be compared to test whether SDM can be degraded into SAR and SEM models.

Table 2 Wald test and LR test results

Test	Statistic	p-value
Wald Spatial Lag test	5.51	0.0189
LR Spatial Lag test	21.13	0.0036
Wald Spatial Error test	5.17	0.0229
LR Spatial Error test	23.64	0.0013

It can be seen from Table 2 that both Wald test and LR test passed the test at the significance level of 5%, indicating that they rejected the two original assumptions. SDM model cannot be degraded into SAR model or SEM model, and it is most appropriate to establish SDM model.

4.4 Benchmark regression

After a series of tests, SDM model was selected for analysis in this paper. The analysis results of Stata15.0 are shown in Table 3.

Table 3 SDM model regression results

Variables	SDM model		
	TE	FE	Both
LQ	0.752*** (0.00)	0.295*** (0.01)	0.356*** (0.00)
lnPGDP	0.288** (0.04)	0.128* (0.08)	0.033* (0.09)
URB	-0.021* (0.09)	-0.003 (0.82)	0.010 (0.19)
IND	0.002* (0.07)	0.018** (0.02)	0.020*** (0.01)
lnMAR	0.01 (0.49)	0.001 (0.26)	0.002 (0.13)
GOV	0.082** (0.04)	-0.013 (0.11)	-0.004 (0.70)
LnOPEN	0.346**	0.314*	0.020*

	(0.01)	(0.09)	(0.07)
W*LQ	1.085*** (0.00)	0.246 (0.68)	0.574** (0.02)
W*lnPGDP	0.288* (0.04)	0.019 (0.41)	0.167* (0.09)
W*URB	-0.001* (0.08)	0.058*** (0.00)	-0.045 (0.39)
W*IND	0.039* (0.07)	0.004** (0.05)	0.077*** (0.00)
W*lnMAR	0.056 (0.49)	-0.001 (0.96)	0.006* (0.09)
W*GOV	0.082** (0.04)	-0.047 (0.60)	0.029* (0.09)
W*LnOPEN	0.346** (0.01)	0.314* (0.09)	0.017 (0.09)*
R-squared	0.6776	0.5547	0.7016
sigma2_e	0.034 *** (0.00)	0.023*** (0.00)	0.021*** (0.00)

a.*, **, *** are significant at the level of 10%, 5% and 1% respectively

In Table 3, According to the R^2 value of SDM model under three different effects: time fixed, space fixed and time-space double fixed, it shows that the space-time double fixed effect is more suiTable for the result analysis. Therefore, the regional time double fixed effect result of SDM model is selected for analysis. The specific analysis is as follows:

As a core explanatory variable, LQ has a significant positive effect on logistics efficiency, with a direct influence coefficient of 0.356. Passing the significance test of 1%, it shows that logistics industry agglomeration has a positive promoting effect on regional logistics efficiency. Every 1% increase in the agglomeration level of the logistics industry in a province directly drives the increase of the logistics efficiency of the province by 0.356%.

According to the regression results of control variables, it shows: PGDP has a positive impact on logistics efficiency, which proves that a certain economic foundation will provide infrastructure guarantee for logistics enterprise agglomeration, and regional economic development will accelerate the refinement of social division of labor, so as to promote the intensification, integration and scale of logistics industry.IND has a positive impact on logistics efficiency. As a productive service industry, logistics industry serves the secondary industry. The optimization and upgrading of industrial structure will promote the more reasonable flow and allocation of factors under market conditions, and become the "lubricant" for the development of logistics industry. GOV on logistics efficiency is positive, which shows that the government in the western region is more efficient in resource allocation to a certain extent, and the dividend of industrial policy promotes the development of logistics industry. OPEN has a positive impact on logistics efficiency. The western region is connected with the two major economic circles of Asia Pacific and Europe. By giving full play to its resource endowment and comparative advantages, it can drive the expansion and structural upgrading of foreign trade, and become an important growth pole for the improvement of logistics efficiency in the region.

The impact of the urbanization rate (URB) and The degree of marketization (MAR) on logistics efficiency is positive, but not significant. This may be due to the low level of urbanization development in the western region at this stage, the lack of effective cooperation and communication between regions, the improvement of logistics parks or logistics bases and other facilities, and the degree of marketization has not reached a reasonable structural state. Therefore, at present, the driving effect of Urbanization and marketization on logistics efficiency is not significant.

4.5 Spillover effect analysis

The partial differential method is used to decompose the regression results of the spatial econometric model in the above, and the impact of each explanatory variable on logistics efficiency is decomposed into direct effect and indirect effect, so as to clarify the impact of regional industrial agglomeration on the logistics efficiency of surrounding areas. The effect decomposition results processed by Stata15.0 are shown in Table 4.

Table 4 Results of spatial effect decomposition

Variables	Direct	Indirect	Total
LQ	0.336*** (0.00)	0.426** (0.01)	0.762*** (0.00)
lnPGDP	0.189* (0.08)	-0.086* (0.09)	0.103* (0.08)
URB	0.006 (0.16)	-0.039 (0.38)	-0.033 (0.41)
IND	0.018*** (0.04)	0.066*** (0.00)	0.084*** (0.00)
lnMAR	0.002* (0.09)	0.005 (0.20)	0.007 (0.18)
GOV	0.006* (0.09)	0.025* (0.06)	0.019* (0.05)
LnOPEN	0.024** (0.05)	0.025* (0.08)	0.050* (0.07)

a. *, ** and *** are significant at the level of 10%, 5% and 1% respectively

It can be seen from the effect decomposition results that the spillover effect of logistics industry agglomeration on surrounding areas is 0.426, passing the significance test of 5%. It shows that logistics industry agglomeration can not only promote the improvement of local logistics efficiency, but also promote the improvement of logistics efficiency in neighboring provinces and cities through labor market sharing, value chain dynamic matching and inter-industry knowledge spillover of logistics industry agglomeration. At the same time, the spatial spillover effect accounts for 55.90% of the total effect, indicating that the spillover effect of logistics industry agglomeration on logistics efficiency is more significant than the direct effect.

5 CONCLUSIONS

In this paper, MaxDEA software is first used to measure logistics efficiency and analyze its spatial correlation by super-SBM model. On the basis of Stata software, the SDM model is established to study the impact of logistics industry agglomeration on logistics efficiency. The research finds that:

1) The degree of logistics industry agglomeration and logistics efficiency in Western China have

obvious spatial imbalance and volatility. On the whole, the northwest is slightly higher than the southwest. The logistic efficiency in western China has a significant positive spatial correlation, indicating that the logistic efficiency in western China has a robust and obvious spatial dependence.

2) The agglomeration of logistics industry has a positive impact on the regional logistics efficiency. The spillover effect of logistics industry agglomeration is very significant, indicating that logistics industry agglomeration can not only promote the improvement of local logistics efficiency, but also promote the improvement of logistics efficiency in neighboring areas.

Therefore, it is necessary to promote the industrial agglomeration of logistics industry in different ways, make full use of the advantages of logistics agglomeration, innovatively improve the logistics efficiency of logistics industry, constantly optimize the spatial layout of logistics industry in western China, and then improve the logistics efficiency in western China.

Ordinary panel model estimation does not consider the spatial effect, and can not measure the impact of explanatory variables on adjacent provinces and cities. The results are prone to bias, which will overestimate or underestimate the impact of explanatory variables on logistics industry agglomeration. Therefore, it is necessary to introduce spatial factors, use spatial econometric model to analyze the impact of logistics industry agglomeration on logistics efficiency, and test its impact from the perspective of time and space.

References

[1] A D. Cliff, J K. Ord. (1981) Spatial Processes: Models and Applications. London: Pion.

[2] N. Chamier-Gliszczyński, J. Fiuk.(2018)Logistic aspects in the Methodology of efficiency evaluation of transport system on the example of Air SAR service. Logistics & Transport, vol.40(4):137-144.

[3] AE. Périco, GRD. Silva.(2020)Port performance in Brazil: A case study using data envelopment analysis. Case Studies on Transport Policy, vol.12(1):123-135.

[4] G. Wei, X. Le, L. Kong.(2018)Research on logistics industry efficiency and its influencing factors in the Yangtze River economic belt.Mathematics in Practice and Theory,vol.48(09):41-47.

[5] X. Gong, L. Jing.(2020)Research on efficiency evaluation and temporal and spatial differences of China's logistics industry under the background of cost reduction and efficiency increase.Journal of Commercial Economics.vol.(14):100-103.

[6] T. Tao.(2017)Can industrial agglomeration promote the efficiency of logistics industry—Empirical analysis of provincial panel data from China. Commercial Research,vol.(1):75-83.

[7] J. Leng, Y. Liu.(2019)Research on the threshold effect of industrial agglomeration on logistics efficiency.Prices Monthly,vol.(12):47-54.

[8] K. Tone, M. Tsutsui.(2010)Dynamic DEA: A slacks-based measure approach[J]. Omega, vol.38(3/4):145-156.