International Journal of Engineering Sciences &Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

THE SPATIAL AUTOCORRELATION ANALYSIS FOR VOLUME OF FREIGHT IN DIFFERENT REGIONS: A CASE OF VIETNAM

Truong Cong Bac[*1] & Phan Quan Viet[2]
[*1&2]Faculty of Commerce, Van Lang University, Vietnam

DOI: 10.5281/zenodo.3577227

ABSTRACT

The transportation plays an important role in economy so the issue of development of transport has attracted much attention for every country and region in the world. There are lots of investigations about the transportation from regional to global scale, however, in these research, the spatial autocorrelation has not been mentioned while spatial autocorrelation related to the independence of variables which challenges the validity of the hypotheses. This study explores and analysis the occurrence of spatial autocorrelation in volume of freight which is investigated in 6 economic regions individually in Vietnam and whole country also. In order to evaluate the phenomena, this article used classic Moran I statistic based on the geographic information of research subjects in Vietnam which extracted from GADM (Database of Global Administrative Areas) and the data of freight volume collected from Statistical Yearbook of Vietnam 2018 published by General Statistics Office of Vietnam. The results show that there are existed spatial autocorrelation of transportation activities in Vietnam. Based on the findings, some research commendations are given for future investigation of transportation.

KEYWORDS: Spatial autocorrelation, freight volume, regional economics, Vietnam.

1. INTRODUCTION

Transportation has always been fundamental component of human societies, it has led to political, social, economic and even cultural development of the country. Nowadays, with the dynamic economic and technology changes result in the fact that societies have become increasingly dependent on their transport development to support a wide variety of activities. From economic perspective, transportation provides a vital link between different geographic regions, it enables the market to access to the material and end products easily so it is widely acknowledged as an important catalyst for economic development, at the various level from region, nation, as well as inter-nation. Contemporary economic processes have been accompanied by significant increase in mobility and higher levels of transport demand. These are also the reasons that investigation of transportation has been a continuous challenge for policy-makers, who need have deep understanding in order to satisfy mobility needs, to support economic development.

A number of studies have explored the relationships of transportation by looking into panel data, but less attention has been paid to analysis occurrence of spatial autocorrelation. From a statistical point of view, some analyses such as: analysis of correlation, linear regressions… are based on the hypothesis of independence of variables. In the case, a variable is autocorrelation, the independence hypothesis is no longer respected, thus challenging the validity. Even though, autocorrelation can come from the temporal autocorrelation or spatial autocorrelation, the existing relative literature almost considered with a time lag (temporal autocorrelation) and less attention was paid to space (spatial autocorrelation). Therefore, the analysis occurrence of spatial autocorrelation in transportation is really important. In order to seek to make up this gap, the main objective of this study is the evaluation of spatial autocorrelation in transportation in Vietnam by combining the hypothesis test of Moran I statistic and Moran's diagram (scatter graph).

The remainder of the paper is organized as follows: Section 2 will discuss brief literature which are relevant this study and provides a background for this research. In the Section 3, methodology and data resource will be introduced. Analysis results are presented in Section 4. Section 5 will discuss the findings, conclude the paper and give some potential directions for future research.

2. LITERATURE REVIEW

Tracing the early development of the concept of spatial autocorrelation, it will be deficient without mentioning about the first law of geography was developed by Waldo Tobler [1]: “Everything is related to everything else. But near things are more related than distant things”. This first law is the foundation of the fundamental concepts of spatial dependence and spatial autocorrelation however, the first indicators which described the spatial autocorrelation appeared a long time ago by some another words. Three statisticians: Moran [2], Krishna Iyer [3], and Geary [4]laid out the mathematical characteristics of spatial autocorrelation by using the term contiguity ratio to describe their work. They developed join count statistics based on the probability that neighboring spatial units were of the same type (black or white) more than chance would have it. Their work was extended to take into account interval data. Cliff and Ord [5]shedded light on the problem spatial autocorrelation and demonstrated anew statistically how researchers can test residuals of regression analysis for spatial randomness by explicating and generalizing Moran's earlier work. Getisand Ord [6]introduced a family of measures of spatial association called G statistics by a relatively simple variation on a basic autocorrelation statistic. These statistics have a number of attributes that make them attractive for measuring association in a spatially distributed variable. Legendre, P. [7] proposed two approaches of statistical testing in the presence of autocorrelation. The spatial structure took the form of a polynomial of geographic coordinates of sampling stations in raw-data approach. As for matrix approach, the spatial structure is introduced in the form of a geographic distance matrix among locations. Anselin [8] introduced local Moran’s I and local Geary’s c statistics which were able to show that the local values were proportional to their global values. These statistics were named by Anselin as LISA (Local Indicators of Spatial Association) and they identified hot spots or possible centers of statistically significant clustering. Ord and Getis[9]realized that local statistic outcomes are influenced by the degree of spatial autocorrelation in the global statistic therefore they developed a statistic which is called an O statistic.

As for transport fields, the spatial autocorrelation has realized early. Bolduc, D and Santarossa, G [10] noted that socio-econmic and network variables adjacent to market would also be incorporated in the relationship explaining passenger’s travel flow. The omission of those variables and the frequent lack of data describing the geographic structure (distance between zones, size of zones, length of frontier between adjacent zones, etc.) of the region leads to some spatial autocorrelation in the regression errors. Kwan, M.-P. [11] argued that new analytical methods that can handle the location of activities and trips in real geographic space are needed in dealing with the spatial and temporal dimensions of human activity-travel patterns. Czado and Prokopenko[12] identified areas of low or high utilization of public transport by included spatial components in models which based on logit formulation after adjusting for explanatory factors. After considering cluster effects using group and individual approaches in some different models, the results showed that there was correlation between cluster effects and public transport demand. In recent years, some scholars consider using the Moran’s I index in testing the spatial autocorrelation in transportation. Górniak[13] assessed and analyzed the occurrence of spatial autocorrelation in connection with the transport accessibility (measured by density of a motorway network) by Moran’s I statistics. Wang et al.[14] used Moran’s I Index with three spatial weight features to measure the spatial autocorrelation of macro-level traffic crashes between each regional unit and the adjacent regional unit in the Traffic Analysis Zones scale. Wandaniet al.[15] measured the spatial dimensions of automobile and motorcycle trips on national roads with using simultaneously Moran’s I, LM, Robust LM tests for spatial correlation. Zhang et al.[16] presented a study on the accessibility of major cities by using the Space Syntax technology to quantitatively measure the accessibility of each urban road network. In his study also used the Moran’s I index a for calculating spatial autocorrelation. Gao et al.[17] studied about spatial fairness and changes in transport infrastructure while Moran’s I index was used as statistic reflected the overall correlation and difference degree of the observed values of spatially adjacent regional units.

Han et al.[18] evaluated the spatial deprivation of public transportation resources in areas of rapid urbanization in which Moran’s I index was used as main spatial clustering analysis statistic. In the same way, Rybarczyk[19] examined and visualized linkages among various contextual and compositional factors and active travel (bicycling, walking, and mass-transit modes) among a commuter-university population with using Moran’s I

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

index in testing for autocorrelation of the dependent variable, independent variables. Ma et al.[20] used the Moran's I index to test the spatial autocorrelation of the integrated transport efficiency of provinces. Zhao et al.[21] assesses the economic influence of intercity multimodal transport demand at the prefecture level in country. Spatial autocorrelation test with Moran's Index was used for transport demand (land, air, and water) at both national and regional levels.

This study seeks to fill in some of research gaps and shed light on the association of volume of freight with economic regions. The findings from this study can provide researchers a clear understanding about this relationship as well as potential direction in the deep study in transportation.

3. METHODOLOGY AND DATA RESOURCE

Methodology

Moran's I index

Spatial autocorrelation is defined as the correlation of a variable with itself across geo referenced space and this autocorrelation can be positive or negative. It will be positive when similar values of the variable to be studied are grouped geographically and be negative when the dissimilar values of the variable to be studied come together geographically or in another word, nearby locations are more different than remote locations. In order to express this relationship, there are some ways in which statistics are particular examples. Spatial autocorrelation indices are used not only to test hypotheses of no spatial autocorrelation but also to assess the degree of spatial autocorrelation extant in the spatial data. Among many statistics of spatial association, Moran's I index is one of the best known and the most widely used measure of and test for spatial autocorrelation [22].
Following Cliff and Ord [23], in this paper Moran's I is defined as follows:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}\left[w_{ij}\left(X_i - \overline{X}\right)\left(X_j - \overline{X}\right)\right]}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right)\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

Where: n is number of observations; X_i is the value of variable at location *i;* $\overline{X}$ is the mean value of the variables; w_{ij} is a weight indexing location of *i* relative to *j*.

Spatial weight matrix W=(w_{ij} : i,j =1,…n) summarizing spatial relations between n spatial units. Here each spatial weight, w_{ij} ,typically reflects the "spatial influence" of unit *j* on unit *i*. Following standard convention, we here exclude "self-influence" by assuming that w_{ij} = 0 for all *i*=1,..,n (so that W has a zero diagonal).Spatial weight matrix can be established by some ways, however, in many cases the boundaries shared between spatial units play in important role in determining degree of "spatial influence". Therefore, in this study I established the spatial weight matrix bases on boundaries (Spatial contiguity weights). The simplest types of these of weights simply indicate whether spatial units share a boundary or not. If the set of boundary points of unit *I* denoted by bound(*i*) then the so-called queen contiguity weights are defined by:

$$w_{ij} = \begin{cases} 1, & bound(i)\cap bound(j)\neq\varnothing \\ 0, & bound(i)\cap bound(j)=\varnothing \end{cases}$$

In order to remove dependence on extraneous scale factors (such as the particular units of distance employed in exponential and power weights), I normalized spatial weights. Recall that the row i^{th} of W contains all spatial weights influencing spatial unit i, namely (w_{ij} : i ≠ j). So the spatial weights in each row are normalized to have unit sum:

$$\sum_{j=1}^{n} w_{ij} = 1, \qquad i = 1,\ldots, n$$

The value of Moran's Index varies on a scale between -1 (high negative spatial autocorrelation) through 0 (no spatial autocorrelation)to +1 (high positive spatial autocorrelation). The hypothesis will be used for Moran's index test of absence of spatial autocorrelation:

H_0 : no spatial autocorrelation

H_1 : spatial autocorrelation

If the null hypothesis is true, it means there is no spatial autocorrelation. The attribute being analyzed is randomly distributed among the features in study area; said another way, the data observed are only one of the many outcomes possible. If the null hypothesis is rejected, it means spatial autocorrelation occurred, we can answer the question as to the signals and strength of the spatial autocorrelation.

Moran Scatter Plots

Moran's scatter plot allows a rapid reading of the spatial structure. This is a useful visual tool for exploratory analysis, because it enables to assess how similar an observed value is to its neighboring observations. The horizontal axis (x-axis) of scatter graph is based on the values of the observations y and is also known as the response axis. The vertical axis (y-axis) is based on the weighted average Wy (averaged values received in neighboring observations) or spatial lag of the corresponding observation on the horizontal x-axis. The slope of a least-squares regression line through the points is the value of Moran's I. In the Moran scatter plot, there are 4 quadrants:

- Quadrant I (The upper right) – states that values of the variable are higher than average, in a neighborhood similar to it — positive spatial autocorrelation and high index value; high-high structure.
- Quadrant II (The bottom right) – states that values of the variable are higher than average, in a neighborhood not similar to it — negative spatial autocorrelation and high index value; high-low structure.
- Quadrant III (The bottom left)– states that values of the variable are lower than average, in a neighborhood similar to it — positive space autocorrelation and low index value; low-low structure.
- Quadrant IV (The top left) – states that values for the variable are lower than the average, in a neighborhood not similar to it — negative spatial autocorrelation and low index value; low-high structure.

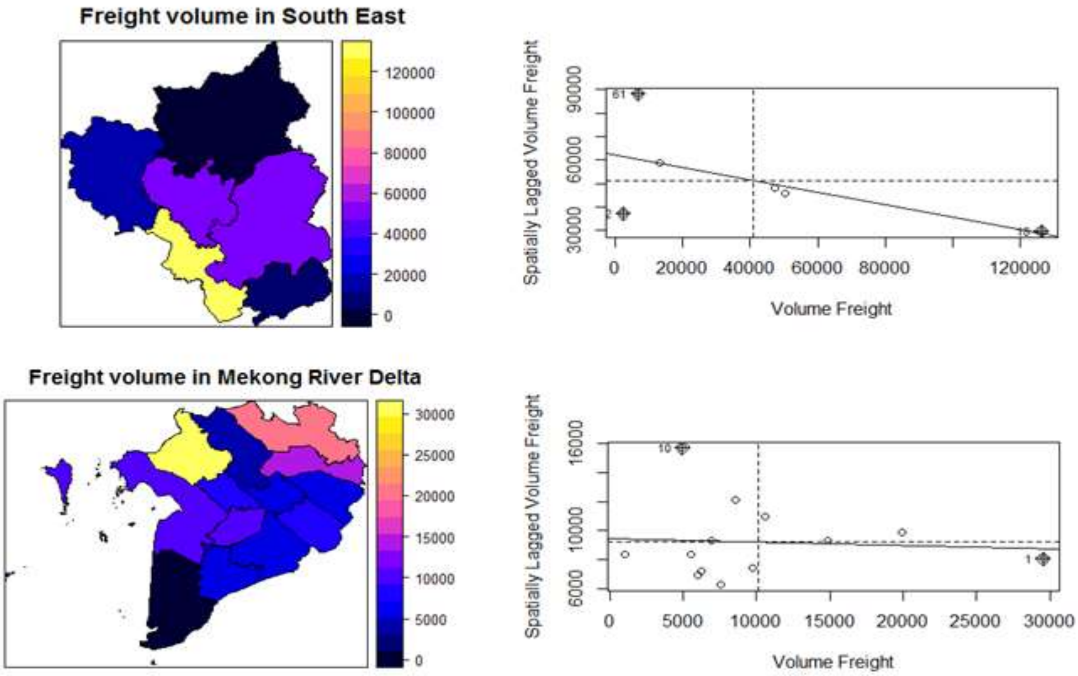
Data Source

This study takes all of provinces and cities in Vietnam which are 63 prefecture-level units as the study objects. The information about freight volume of these provinces is collected from Statistical Yearbook of Vietnam 2018 published by General Statistics Office of Vietnam. In this research, volume of freight is the volume of cargoes transported by transportation establishments and others operating in transportation business activities regardless of travel distance. Volume of freight is calculated in thousand tons by the actual weight of goods carried (including packing). It is only measured after the completion of transportation to the destination as mentioned in the contracts and finishing delivery procedure. As for spatial data, the geographic information of prefecture-level units in Vietnam is taken from GADM (Database of Global Administrative Areas). In some types of data which are accessed, the author chosen data in (.rds) type which is analyzed by R software (Version3.6.1).

4. RESULTS

Figure 1 represents the spatial distribution of freight volume in different regions of Vietnam visually. It can be seen from maps that almost regions in Vietnam have unequal values and random spatial distribution of freight volume. In the Moran Scatter plot, it allows more clear view of spatial structure in each regions respectively. As we can see, the average of freight volume in 3 regions: (Northern midlands and mountain areas, Central highlands and Mekong River Delta) are the same at 10 000 thousand tons while these value of 2 other regions

(Red River Delta and South East) are very high at 40 000 (four times).From the graphs, the slopes of regression lines are different with 4 downward slopes and 2 upward slopes, that means the nearby observations tend to have dissimilar values in 4 regions and tend to have similar values in 2 regions. In another word, the Moran's I Index could be negative or positive in different regions in Vietnam. However, also in these graphs, in detail, almost observations are unclear distributed in quadrants, only values of observations in North Central and Central Coastal Areas, Central highlands are distributed in clear quadrants which show significant slope of regression lines. Therefore, the initial findings just that the average of freight volume is different among economic regions in Vietnam.



ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

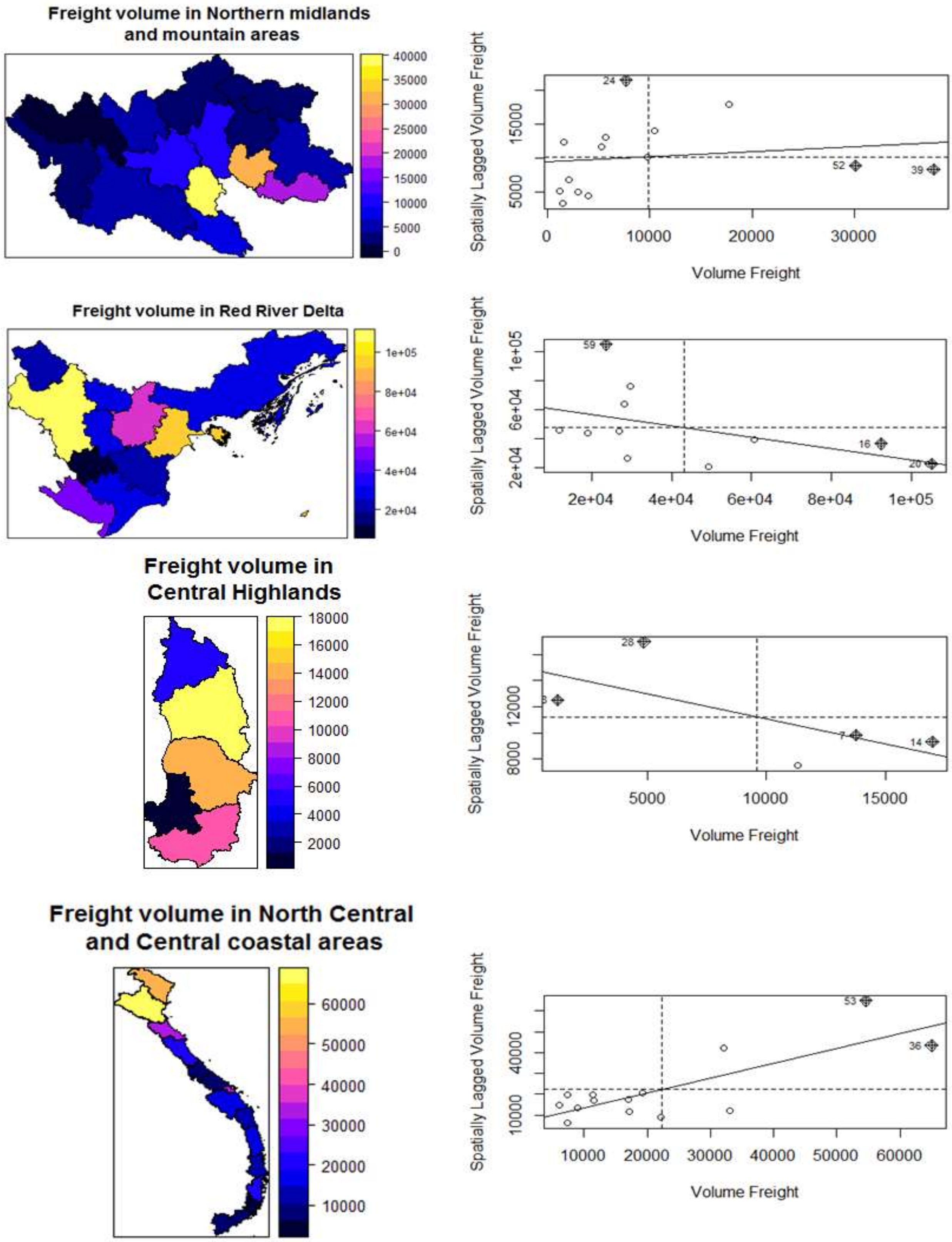


Figure 1: Quantile map of freight volume and Moran Scatter plot respectively in different economic regions in Vietnam.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

The Moran's Test for freight volume in different economic regions in Vietnam is shown in Table 1. The test results show that we can reject H_0 (no spatial autocorrelation) just in North Central and Central Coastal Areas (significant at the 0.01 level) while we do not reject the H_0 in 5 rest of regions in Vietnam. In another words, there is sufficient evidence in the sample in favor spatial autocorrelation of freight volume in North Central and Central Coastal Areas as well as there are insufficient evidences to support the spatial autocorrelation for rest of regions.

Table 1: Moran's Test of the freight volume in regions in Vietnam

Regions	**Moran's I Test**			
	Moran's I index	**Standard deviate**	**Variance**	**p-value**
Northern midlands and mountain areas	0.074	0.985	0.023	0.162
Red River Delta	-0.391	-1.488	0.038	0.931
North Central and Central Coastal Areas	0.717	3.312	0.058	0.000
Central highlands	-0.386	-0.351	0.149	0.637
South East	-0.264	-0.425	0.023	0.665
Mekong River Delta	-0.023	0.415	0.021	0.339

The connection between neighbor provinces is described at Contiguity Neighbors Map in Figure 2. Each red link represents the connection of two provinces with contiguous boundaries that are sharing one or more boundary point. It can be seen that in the Northern and Southern of Vietnam, there are more connections among provinces than Middle. As for quantile map of freight volume, the distribution of freight volume in whole country is more clear. By the rapid reading of the spatial structure, it is clear that high values of freight volume are clustering in central of Northern and Southern of Vietnam while the rest of regions are low values.

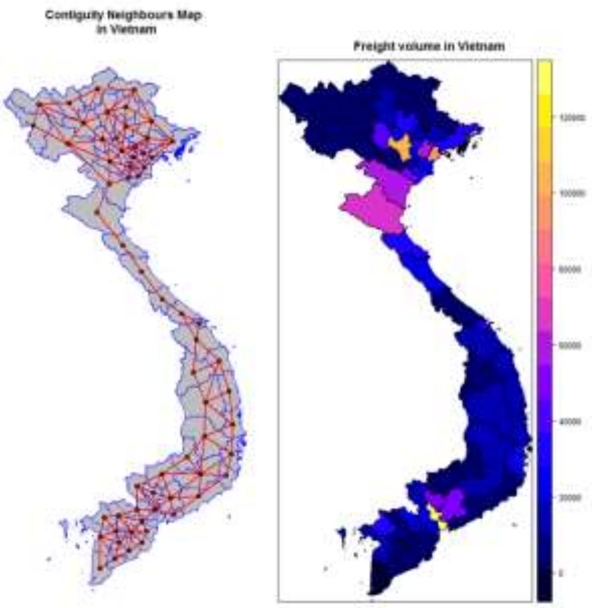


Figure 2: Contiguity Neighbours Map and Freight volume in Vietnam

Figure 3 illustrates the spatial density distribution of freight volume in axial graph from which we can assess the cluster easily. Visually, the values of freight volume of provinces concentrate almost in quadrants I and III

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

(upper right and bottom left) that represents the positive spatial autocorrelation in freight volume of Vietnam.

This is confirmed when we look at the Table 2, P-value is 0.002, it means we can reject H_0 (significant at the 0.01 level): No spatial autocorrelation in freight volume of Vietnam while A positive Moran's I Index (0.224) indicates that the value of the freight volume tends to be similar to its neighbors.

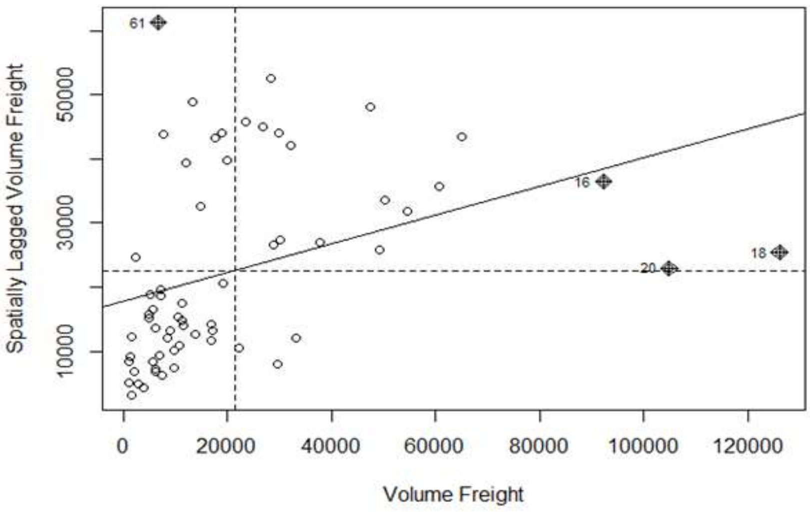


Figure 3: Moran Scatter plot of Freight volume in Vietnam

Table 2: Moran's Test of the freight volume in Vietnam

Moran's I index	Standard deviate	Variance	P-value
0.224	2.915	0.007	0.002

5. DISCUSSION AND CONCLUSION

This study seeks to analyze the spatial autocorrelation in volume of freight in different economic regions of Vietnam and whole country also. Data of freight volume in all provinces are used to explore the spatial autocorrelation while Moran's I Index and Moran scatter plot are performed as methodology.

Findings from this study indicate that at the regional level, the spatial autocorrelation becomes less significant or even absent. There is one region (North Central and Central Coastal Areas) shows the signal about occurrence of spatial autocorrelation. In reversely, at the nation level, values of freight volume demonstrate significantly spatial cluster pattern. This is positive autocorrelation, in another word, there is similarity value of volume freight between provinces with high-high value and low-low value and these groups of country create clusters. This is reasonable since cluster evidence generally becomes less clear at a smaller geographical scale.

Based on the results from this study, in future research, some researches which relate to transportation, specially freight volume need to consider about spatial autocorrelation when they are study objectives. In additionally, the size of sample will be also important issue in which scholars can confirm the spatial autocorrelation reasonably. Besides that, this study has a limitation, which also provides potential directions for future research. The study focused on measure of how similar locations are to their neighborhoods while it did not dig a little deeper and understand exactly which objects were similar or different to the objects in their neighborhoods. With the advancement of econometric and data availability, this gap is expected to filled.

REFERENCES

[1] Tobler, W. R. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography*, vol. 46, pp. 234–240,June. 1970.
[2] Moran, P. A. P. "The Interpretation of Statistical Maps." Journal of the Royal Statistical Society. Series B (Methodological), Vol. 10, no. 2, pp. 243–251, 1948

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[3] P. V. Krishna Iyer. "The First and Second Moments of Some Probability Distributions Arising from Points on a Lattice and Their Application." Biometrika, Vol. 36, no. 1/2, pp. 135–141, June. 1949.
[4] Geary, R. C. "The Contiguity Ratio and Statistical Mapping." The Incorporated Statistician, Vol. 5, no. 3, pp. 115–146, November. 1954.
[5] Cliff, A. D., and J. K. Ord. Spatial Autocorrelation. London: Pion, 1973.
[6] Getis, A., and J. K. Ord. ''The Analysis of Spatial Association by Distance Statistics.'' Geographical Analysis, vol. 24, Issue. 3, pp. 189–206, July.1992.
[7] Legendre, P. "Spatial autocorrelation: trouble or new paradigm?" Ecology, Vol. 74, Issue. 6, pp. 1659–1673, September. 1993.
[8] Anselin, L. ''Local Indicators of Spatial Association–LISA.'' Geographical Analysis, Vol. 27, Issue. 2, pp. 93–115, April. 1995.
[9] Ord, J. K., and A. Getis. ''Testing for Local Spatial Autocorrelation in the Presence of Global Autocorrelation.'' Journal of Regional Science, Vol. 41, Issue. 3, pp. 411–432, August.2001
[10] Bolduc, D., R. Laferrière, and G. Sanatrossa. "Spatial autoregressive error components in travel flow models" Regional Science and Urban Economics, vol. 22, Issue. 3, pp. 371-38, September. 1992.
[11] Kwan, M.-P. "Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set." Transportation Research Part C: Emerging Technologies, Vol. 8, Issue. 1-6, pp.185–203, September. 2000.
[12] Czado, Claudia, and SergijProkopenko. "Modelling Transport Mode Decisions Using Hierarchical Logistic Regression Models with Spatial and Cluster Effects." Statistical Modelling, Vol. 8, Issue. 4, pp. 315–345, Dec. 2008.
[13] Górniak, J. "The Spatial Autocorrelation Analysis for Transport Accessibility In Selected Regions Of The European Union.", Comparative Economic Research, Vol. 19, Issue. 5, pp. 25–42, December. 2016.
[14] Wang, S., Chen, Y., Huang, J., Chen, N., and Lu, Y. "Macrolevel Traffic Crash Analysis: A Spatial Econometric Model Approach." Mathematical Problems in Engineering, vol. 2019, pp.1–10, April. 2019
[15] Wandani, F. P., Siti, M., Yamamoto, M., and Yoshida, Y. "Spatial econometric analysis of automobile and motorcycle traffic on Indonesian national roads and its socio-economic determinants: Is it local or beyond city boundaries?" IATSS Research, Vol.42, Issue 2, pp. 76–85, July. 2018.
[16] Zhang, X., Ren, A., Chen, L., and Zheng, X. "Measurement and Spatial Difference Analysis on the Accessibility of Road Networks in Major Cities of China." Sustainability, Vol. 11, Issue 15, pp. 4209, August. 2019.
[17] Gao, X., Li, T., and Cao, X. "Spatial Fairness and Changes in Transport Infrastructure in the Qinghai-Tibet Plateau Area from 1976 to 2016". Sustainability, Vol. 11, Issue. 3, pp. 1-16, January. 2019.
[18] Chunyang, Han, Xinquan Liu, Xiaojing Shen, Ling Zhang, and Nana Feng." Evaluating the Spatial Deprivation of Public Transportation Resources in Areas of Rapid Urbanization: Accessibility and Social Equity." Discrete Dynamics in Nature and Society, vol. 2019, pp. 1-11, February. 2019.
[19] Greg Rybarczyk. "Toward a spatial understanding of active transportation potential among a university population," International Journal of Sustainable Transportation, Vol. 12, Issue. 9, pp. 625-636, January. 2018.
[20] Fei Ma,Wenlin Wang, Qipeng Sun, Fei Liu and Xiaodan Li. "Integrated Transport Efficiency and Its Spatial Convergence in China's Provinces: A Super-SBM DEA Model Considering Undesirable Outputs." Applied sciences, Vol. 8, Issue. 9, pp. 1698, September. 2018.
[21] Zhao, Jinbao&Guo, Dong & Wang, Jian & Yang, Zhao & Zhang, Hefang. "Examining the Association of Economic Development with Intercity Multimodal Transport Demand in China: A Focus on Spatial Autoregressive Analysis." ISPRS International Journal of Geo-Information, Vol.7, Issue 2, pp. 56, February 2018.
[22] Getis, A. "A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective." Geographical Analysis, Vol. 40, no. 3, pp. 297–309, July. 2008.
[23] Cliff, A. D., and J. K. Ord. Spatial Processes, Models and Applications. London: Pion, 1981.
[24] Moran, P. A. P. "Notes on Continuous Stochastic Phenomena." Biometrika, Vol. 37, no. 1/2, pp. 17–23, June. 1950.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[25] Whittle, P. "On Stationary Processes in the Plane." Biometrika, Vol. 41, no. 3/4, pp. 434–449, December. 1954.