Master’s Thesis

RITSUMEIKAN ASIA PACIFIC UNIVERSITY

GRADUATE SCHOOL OF ASIA PACIFIC STUDIES

ESDA Techniques in identifying the Spatial Structure of the Tokyo Metropolitan Area:

Preliminary Research

By

TAWHID MONZUR

(Student Number: 51212606)

Date

2014/07/18

Supervisor:

PROFESSOR YAN LI

Contents

Acknowledgement

The challenges that I faced in preparing the master's thesis could not have been overcome without proper and affable guidance, encouragement and support from some special persons.

Firstly, my warmhearted gratitude and appreciation to Professor Yan LI, whose sincere supervision, monitoring and generosity helped me to complete this research. Moreover, I express my heartfelt gratefulness to her for helping me selecting the research topic and providing necessary information in conducting the research which broadened my views and thoughts in the research field. Nevertheless, unconditional contributions and sincerity of Professor Yan Li inspired me to choose this research.

Also, my indebted thankfulness to the professors; especially Professor Isoda Yuzuru (Tohoku University) and Professor Steven Farber (The University of Utah) for their pragmatic views and thoughts besides, valuable pieces of advice, guidance, comments and recommendations in pursuing my research work. I am also grateful to Professor Sanga-Ngoie whose important suggestions really helped me to specify my research view.

Furthermore, I am thankful to my parents, sisters and all my friends whose inspiration and enthusiastic encouragement physically and mentally supported me in conducting the research study.

Lastly, my unending gratitude to all of the special persons for their contribution, generosity, hospitality and cooperation in guiding me to complete my master's thesis on time.

Thank you all very much indeed and wish you a prosperous life….

Abstract

Urban spatial structure studies are becoming an important research field area because of the current pace of urban growth change. Tokyo Metropolitan Region (TMR), the largest mega regional area of the world is poorly studied from the urban studies point of view.

This research focused on analyzing the urban spatial structure of the Tokyo metropolitan area by using Exploratory Spatial Data Analysis (ESDA) techniques and aimed to observe the effectiveness and acceptability of ESDA in identifying the Spatial Clustering. The ESDA techniques that have been selected for this research are Global Moran's I; Local Moran's I and Local G Statistic. The Local statistics are called Local Indicators of Spatial Association (LISA)

A household distribution variable from the 2000 Population census database has been extracted for the spatial structure analysis of Tokyo which has been manipulated and tabulated by using MS - Excel statistical software. ArcGIS 10.2 and GeoDa software have been used to project and analyze the manipulated data. A null hypothesis (*the household distribution in Tokyo is showing no spatial clustering pattern*) has been tested through the analyses for the urban spatial structure study which is a pre-requisite. The analysis showed strong evidence against the null hypothesis which means the household distribution in Tokyo is showing a spatial clustering pattern.

The Global and Local Moran's I analyses identified specific spatial pattern types of the household distribution in Tokyo. The Local Moran's I divided all the household neighboring values into five clustering groups (HH, LL, LH, HL, NS) where most of the values are observed located in the HH clustering group which means that the density of household in one observed area is very high and are surrounded by the areas with high density of household. The analysis also located the outliers (LH and HL types), different from HH and LL clustering groups which means high density of household areas are surrounded by low density of household areas. Local G statistic analysis has been used which calculates by using a significance level test. The statistical significance level test of 95% to 99.9% level of confidence identified the true locations of the spatial clustering of household density in Tokyo.

The ESDA analysis pointed out specific locations of clustering types that cannot be observed only from the normal distribution map. Different significance level tests pinpoint the exact locations of the clustering of high population concentration besides the outliers of the area. Moreover, this study proved that huge data sets can be analyzed through the ESDA techniques.

Keyword: Spatial structure; ESDA; LISA; Tokyo; Global Moran's I; Local G Statistic; Local Moran's I; ArcGIS 10.2; GeoDa; EXCEL

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

1. *CBD- Central Business District*
2. *ESDA- Exploratory spatial data analysis.*
3. *GIS- Geographic information system.*
4. *GISA- Global Indicators of Spatial Association*
5. *LISA- Local Indicators of Spatial Association.*
6. *MIC - Ministry of Internal Affairs and Communications.*
7. *TMA- Tokyo Metropolitan Area.*

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Cities can be regarded as the complex of socio-economic and physical spaces. Understanding the spatial structure or the framework of land use and socio-economic patterns of the complex yields insights about economy-wide growth processes and hence provides the knowledge and tools for policymakers in city management.

Studying the spatial structure of urban areas, especially that of large urban areas is becoming more important than ever because of the current pace of urban growth. Half of the population in the world is now living in the urban areas. Urban regions are becoming edgeless and endless. Mega urban areas around the world have increased and 70% are located in the Asia Pacific region where population concentration is high in the outside of the urban core areas (Uchiyama and Okabe, 2012). Many researchers from diverse fields are trying to understand and identify the spatial pattern of the urban areas.

The Tokyo metropolitan region (henceforward abbreviated to "Tokyo") is the urban area centered in the 23 inner wards and stretched to the neighboring prefectures including Saitama, Kanagawa, Chiba, Gunma, Ibaraki, and Tochigi. It is known as the largest mega urban area of the world with improved transportation system and efficient urban management policy despite a huge population. This research has selected Tokyo. Firstly, it is the heart of the country, sharing less than 8.5% of the land in Japan but yields 40% of the national Gross Domestic Product (GDP). Secondly, the urban management plan of Tokyo is praised because of its uniqueness and effectiveness from the side of urban planning (Hein, 2010). Thirdly, the urban development process of Tokyo experienced multiple political and environmental upheavals; but it still carries the fame of sustainable urban form attainment (Cho, 2011; Hein, 2010; Morita et al., 2012; Nakabayashi, 1986). Henceforth, understating

Tokyo could not only contribute to the urban and regional planning of this region, but also will provide a role model for the emerging Asian cities.

1.2 Research Aim

This research aims to explore the methods in unveiling the spatial structure of Tokyo. It uses Geographic Information System (GIS) and applies ESDA techniques for the purpose. These methods have been used for studying the urban spatial pattern of many western cities but still weak for understanding Tokyo. For the preliminary research household distribution has been analyzed. The objectives of this research are as follows:

- *Implement Global Moran's I, local Moran's I and local G statistic to analyze the spatial structure of Tokyo metropolitan area.*

- *To demonstrate the performance and acceptability of the ESDA techniques in studying spatial pattern of Hugh area.*

1.3 Structure of the Thesis

Chapter 2 gives a thorough review of urban spatial structure studies from the side of global perspective and also Tokyo. The review also illustrates the methods that have been used to study the spatial structure as well as the factors that influenced the urban spatial structure.

Chapter 3 illustrates the study area selected for the research study which is analyzed in chapters 4, 5 and 6. The data processing and manipulation of the data, GIS software and analysis process have been explained.

Chapter 4 discusses the first technique, Global Moran's I for analyzing the spatial pattern of Tokyo. The chapter explains the calculation procedure and also implements the technique to study the spatial structure of Tokyo. The findings are also provided in this chapter.

Chapter 5 deals with the 2^{nd} technique; Local Moran's I for analyzing the spatial structure of Tokyo. The differences can also be observed between the global and local Moran's I computation procedure as well as importance of using the local Moran's I. Moreover, the findings are also explained in this chapter.

C**hapter 6** explains the 3^{rd} technique implemented for the analysis, Local G statistic. Besides using the global and local Moran's I techniques, it's important to use local G statistic to find out whether the evidence is true or false. The findings are also included in this chapter.

C**hapter 7** summarizes the thesis findings and significance of the outcome. Moreover, the chapter also explains the contribution of the research. Besides, research limitations and future directions for further studies are also provided.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Urban spatial structure of Tokyo lacks thorough analysis and investigation. Existing literature regarding the spatial structure study gives a blurred view and unsatisfactory conclusions about the spatial pattern. Moreover, a thorough review revealed the fact that despite Tokyo's having a good urban management plan and transportation system, it fails to answer simple questions regarding spatial pattern ("What, how and where").The next section provides a definition of urban spatial structure and an over view of the urban spatial structure studies from the global perspective. The following section explains the spatial structure studies regarding the Tokyo metropolitan area.

2.2 Urban Spatial Structure

The urban spatial structure varies over time and place. Classical urban model or "Monocentric urban model" introduced by Alonso (1964), Muth (1969) and Mills (1972) is no longer applicable for emerging cities in developing regions because of segregation of the employment and population which can also be termed as the extended metropolitan region (EMR) phenomena (Jones, 2001; Hakim and Parolin, 2009; Hepp, 2011; Bertaud and Malpezzi, 1999). Megacities in the developing nations especially in Asian region are structured and evolving in more complex patterns. These new urban forms (i.e. polycentric city formation, Sprawling) are influencing the travel behavior, transportation, economy, environment quality, land usage, etc. of an urban area and creating a complex urban system (Uchiyama and Okabe, 2012; Lewis et al., 2013; Hepp, 2011).

2.2.1 Definition of urban spatial structure

Urban structure or urban spatial structure is the framework of land use of a city. It is the basic understanding of how people, economy and society occupy the space. Spatial structure of a city can determine the future efficiency from the side of economy, environment and society. Therefore, learning the spatial structure can be helpful in the way to implement policies with better performances.

2.2.2 Urban Spatial Structure Studies

Studying urban spatial structure is becoming increasingly important because of the accelerating urban expansion process. Current pace of population growth, bringing environmental threats, space shortage, health dilemma, resource depletion, etc. all require proper attention which are more or less related to an efficient structure of the cities (Smith, 2011; Lewis et al., 2013 ; Bertaud, 2001, 2004; Rode, 2008; Doytsher et al., 2010; Lee, 2007; Glaeser and Khan, 2001; Mieszkowski and Mills, 1993; Dieleman and Wegener, 2004).

Moreover, Mono-centric city is becoming polycentric; resulting in difficulties to study urban spatial structure because of traditional based urban methods is no more applicable. Recent trend of urban growth is difficult to understand and demands for improved methods or techniques to analyze the spatial structure of urban areas.

Urban spatial structure has been studied keeping in view the employment distribution, job accessibility, population distribution and growth, transportation influence and sub-center formation (building sub centers). (For example, see Baum-Snow, 2007; Helsley and Sullivan, 1991; Pan & Ma, 2006; SaadAllah et al., 2013; Smith, 2003; Kawabata, 2003).

Urban spatial structure has been studied from the sense that it can promote the cohesion among the urban growth patterns. Many scholars focused on the suburban studies to identify the spatial structure of urban areas (Guiliano and Small, 1999; Redfearn, 2007; McMillen, 2003; Pan & Ma, 2006). Emerging megacities around the world are taking a form that includes many centers. Employment sub-centers as explained in the papers of (Bertaud and Malpezzi, 1999; McMillen, 2003) act as isolated centers or independent Central Business District (CBD) areas because it can attract people to move in and reduce the time and cost of the production.

2.2.3 Factors that influence the urban spatial structure

The spatial structure of an urban area as stated by Ibrahim (1997) is the “interaction between the urban form and its functions”. Stanislav (2011) referred urban spatial structure as “unpredictable consequence” that has no constructive shape but structured based on the policies and regulations.

Various scholars have pointed out the factors that influenced the spatial structure is as follows:

-“Spatial structure is the physical outcome of the subtle interactions over centuries between land market, and topography, infrastructure, regulations and taxation” (Bertaud and Malpezzi, 2003).

-“Structure of a city is the outcome of the economic interactions between firms and households, which favor spatial concentration by reason of agglomeration economics”(Garcia and Muniz, 2011).

-“The distribution pattern of urban economic activities and residents along with the existing transport network.”(Sohn, 2005).

-“The spatial structure of the large cities is shaped by the advancement of the transportation and communication” (Anas et al., 1998).

Spatial structure of a city is influenced by the population size and their activities whose structure is formed based on transportation, develops suburban centers and creates sub-urbanization phenomena (Helsley and Sullivan, 1991).

Suburbanization, though a problem of the US cities now can be observed in other parts of the world because of emerging mega cities specifically in the Asian region (Uchiyama and Okabe, 2012). China and India are two of the Asian countries owners of most of the emerging megacities. Megacity growth occurs when a country's economic condition improves. However, in the case of emerging cities specially in the developing countries beset with economic, environmental, social, political and health problems which is referred as "the developing country phenomena" rather than "Engines of growth" (Gavin, 2000). Moreover, distribution of the employment depends on the government regulation and policy implementation. In the papers of (Bollinger and Ihlanfeldt, 2000; Mills and Price, 1984), the policy regarding housing, tax, infrastructure development, etc. are found to have influenced greatly in distribution of the employment where locations outside the city core areas are regarded as the best option to construct industrial zones. Sub-centers formation is also referred as a transport led urban growth centers. Transportation has played a vital role in distribution of the employment and population in suburban areas. Besides, negative urban externalities also played a key role in population distribution within cities and suburbs (Mills and Price, 1984). Suburban areas are cheap and homogenous in nature and fit to healthy living which pulls population to migrate. Despite the fact that the gentrification pours in, central city areas also concentrates with population referred as "counter urbanization" or "shrinking cities phenomena" which can be observed in the European Union (EU) and United states (US) cities (Speare, 1991; Champion, 1989; Sexto, 2009; Wiechmann and Pallagst , 2012).

Spatial structure of cities can be well managed if the austerity policies are lessened for the sake of urban regeneration. Employment decentralization has created a situation where the central areas are dominated by the middle class and the suburbs by the rich class which has increased the usage of transportation in the case people commutes every day to the workplace by using public rather than private mode (Ross and Bartolome, 2004;

Glaeser et al., 2008). On the other hand, Japan especially Tokyo which has an improved nodal system (Railway network) gives the priority to the people living in long distant areas (Shen, 2006).

2.2.4 Different methods in use to study Urban spatial structure

The study of urban spatial structure usually starts with analyzing the distribution of population and employment (For example, see Palumbo, 1990; Boarnet, 1994; Glaeser and Khan, 2001). Several methods have been introduced to study the urban spatial structure. Typical methods are presented in Table 2 from where it can be easily understood about the different approaches that have been taken to study the urban spatial structure of the megacities. The methods have been used in studying specially identifying the sub-centers, population density and employment distribution to study the spatial structure of the urban area.

Table 2: Different methods in use studying the spatial structure of urban areas

Researchers	Model type	Methods
McMillen (2001)	Non- parametric Approach	Locally weighted regression (LWR)
		Semi-parametric Regression
Lee (2007)	Non-parametric Approach	Modified Wheaton index (MWI);
		Area based centralization Index (ACI);
		Weighted average distance;
		Gini- Coefficient;
		Delta Index;
Uchiyama and Okabe (2012)	ESDA Approach	Global Moran's I
Lewis et al. (2013)	Cluster type Approach	Gini coefficient, Lorenz curve,
		Density gradient
Hakim and Parolin (2009)	ESDA & Clustering Approach	Global and local Moran's I & Getis ordGi*

The first two types of methods are implemented in several researches related to identify the sub-centers and distribution of the population and employment within metropolitan areas and suburban areas. The 3rd and 5th types of methodological approach have been used for identification of urban spatial structure but are limited in identifying centers and sub-centers. The 4th type of method especially the density gradient is a traditional model in analyzing the urban growth.

GIS and Spatial structure based studies can be found in the papers of (Barredo and Demicheli , 2003; Samat, 2007, 2009; Xi et al., 2009; Li and Gar, 2004; Estiri, 2012; Gamez and Dallerba, 2012; Hakim and Parolin, 2009; Baumont et al., 2004; Arribas-Bel, 2009). Mostly used GIS-based methods are *Cellular Automata spaital model or CA* and the point pattern analysis (Anas et al., 1998) which is the *ESDA techniques*. Both of the GIS-based methods are used for understanding the spatial structure of urban areas. Despite the fact that the *Cellular Automata Spatial Model* has been criticized in the papers of (Samat, 2007, 2009) as traditional and expensive, the second type of GIS-based method has been used which is easy to impliment and manipulate in conducting research (Arribas-Bel, 2009; Mitra et al., 2010). In this study, the latter has been selected.

The ESDA techniques are explained in the Chapters 4, 5 and 6.

2.3 Studies on Tokyo’s Spatial Structure

The spatial structure of Tokyo has been studied from different perspectives. Many scholars have focused on the environmental and political and historical events. Alongside, development of GIS takes the way to analyze the urban growth studies of Tokyo which has revolutionized the understanding of the urban structure. However, extensive review revealed that literature regarding analyzing spatial structure of Tokyo is rather limited compared with US and European cities.

Suburbanization and employment decentralization process have become urban issues of Tokyo after the introduction of western culture and modernization of transportation network (Ichikawa, 1994). Since 1968, implementation of "new city planning law" can be observed in formulating urban policy of Tokyo metropolitan area. The most prominent land readjustment though praised worldwide for its effectiveness in controlling urban sprawl phenomena failed to stop the suburban growth in Tokyo region (Sorensen,1999, 2001a; Okata and Murayama, 2011). Moreover, the capital regional plan, green space plan, etc. failed to control the population growth outside the CBD area (Morita et al., 2012; Sorensen, 2004). The population concentration or in other words inner and outer migration pushed city to grow outward where employment is influenced by the space shortage in the CBD area and dispersed. The dispersion of employment is further influenced by the improvement of transportation which is introduced in 1872 onwards (Hirooka, 2000; Chorus and Bertolini, 2011). Transportation, especially introduction of railway played a crucial role in population distribution and in formation of suburban areas outside Tokyo region (Alpkokin et al., 2007; Calimente, 2012; Suzuki and Muromachi, 2010; Doi, 1975). Several studies based on the urban planning have stated that the Tokyo urban growth change is influenced by the policy of the government which prioritize in economic development over social improvement has imposed a society which is referred as "Rich Japan, Poor Japanese" (Sorensen, 1999, 2001a, 2004; An, 2008).

Spatial structure of Tokyo has been analyzed mainly based on historical review; categorization of the important events; location and amenities preference; employment distribution and inter-urban population distribution and migration (For example, see, Ichikawa, 1994; Watanabe, 1972, Watanabe et al., 1980; Kikuchi and Obara, 2004; Sorensen , 2001a, 2001b; Okata and Murayama, 2011; Pernice, 2007; Tonuma,1998; Hein, 2010). Tokyo has a good transportation network specially railway which connects the urban and suburban locations and enables city population to commute from long distant areas. Decentralization, as stated in the paper of An (2008) is one of the main governmental policies which has been emphasized for reducing the concentration within CBD areas. To get rid of the negative urban externalities, government plans seemed successful for the development of the CBD areas but the concentration of population and employment outside the CBD areas get worse and become one of the prominent issues the government of Tokyo has to deal with (Sorensen, 2001a, 2004). Several urban policies have been implemented but population concentration surpasses all the regulations.

Tokyo spatial structure studies can be found in the papers where the main focus was personal income, transportation cost, land price fluctuation and land use change for agricultural purpose which has influenced the spatial pattern change (Fujita and Kashiwadani, 1989; Zheng, 1990, 1991; Inoue et al., 2007; Kikuchi and Obara, 2004). Moreover, A Global Moran's I technique based analysis has been performed in Tokyo to find out the urban land use pattern (Zhao and Murayama, 2005, 2006). However, it needs further research because Global Moran's I alone failed to identify the distribution of population in local level as to be explained in chapter 4. A local Moran's I with k-order neighbors used to analyze the distribution of elderly people in Ichikawa City (Murayama, 2011), where it is only a small part of Tokyo metropolitan area. Besides, a ranking based analysis on world megacities by using Global Autocorrelation identified the distribution of population in Tokyo which lacks further investigation (Uchiyama and Okabe, 2012). A grid cell based analysis on the population distribution seems incomplete and rather restricted from the side it showed only the normal distribution of the population distribution and its dependency over land use changes (Bagan and Yamagata, 2012).

In conclusion, the articles related to Tokyo spatial structure are large in number but most of which focus on modeling the factors that affect the distribution and lack of visual understanding. Only very small number of research tried to understand the spatial structure itself, however, they failed to study the whole metropolitan region. This research, though only focusing on the household distribution, will explore the methods for understanding the spatial structure of the entire Tokyo metropolitan area by applying ESDA techniques on GIS which can give a clearer view of the urban spatial structure. By further research, the spatial structure of Tokyo can be unveiled and it will contribute to a better urban policy-making for the region and serve as a reference to the emerging megacities of Asia which needs proper urban planning and policy as well (Uchiyama and Okabe, 2012).

CHAPTER 3

STUDY AREA AND DATA PROCESSING

3.1 General introduction

This chapter will elaborate the study area and the manipulation of data that have been used for the empirical analysis conducted in chapter 4, 5 and 6.

3.2 The study area

This research focuses on the Tokyo metropolitan area (we simply call it Tokyo), including the 23 inner wards and the neighboring prefectures: Saitama, Kanagawa, Chiba, Gunma, Ibaraki, and Tochigi. It is an area of 15,930km2 and is referred as the largest metropolitan region in the world. Tokyo shares 8.5% of the land in Japan and has a population of 37.5 million with a density of 4400 per square kilometers. It is the economic hub of Japan with a high rate of population concentration. The figure 3.1 shows the population density shows the household density of the Tokyo metropolitan area where in can be observed that the density is distributed from central areas to about 70 kilometers outside.

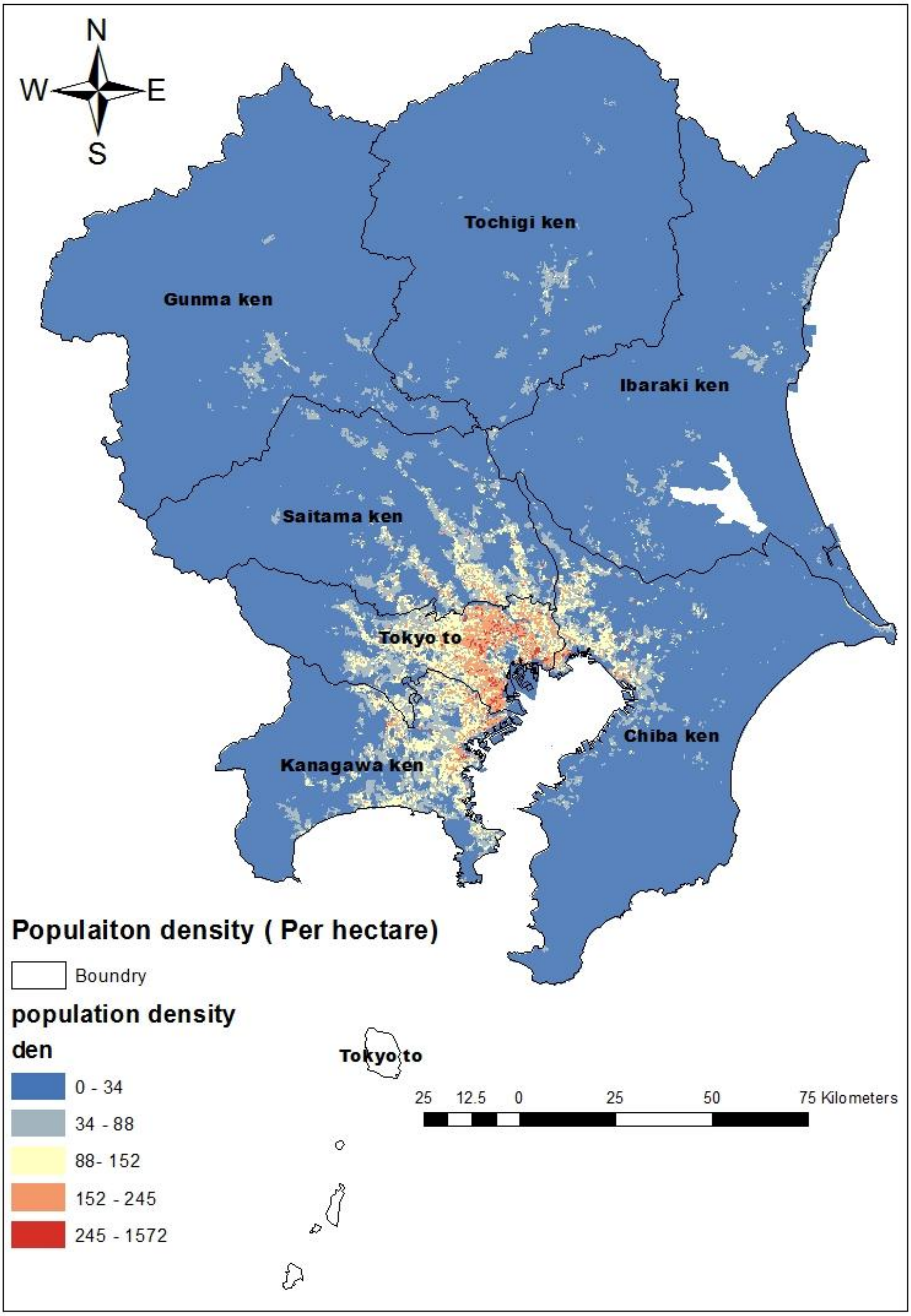

Figure 3.1: Population Density Map (Natural Breaks)

From the Figure 3.1, it can be observed that outside the 70 kilometer range, population density and household density is really low. To avoid unnecessary heavy calculation, this research selected only high density areas which are within 70 kilometers of range as the figures 3.2 shows below.

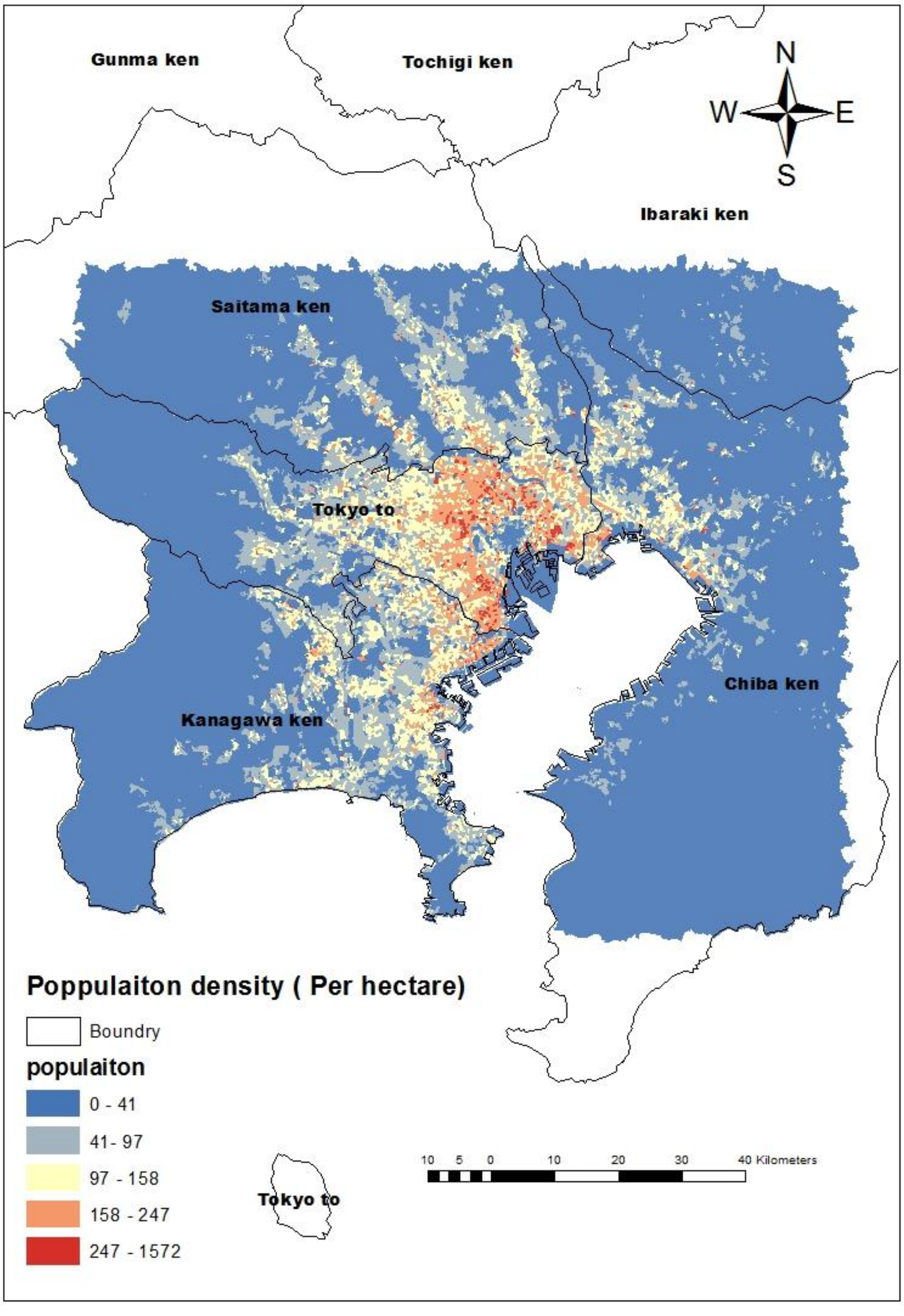

Figure 3.2: Population density map up to 70 km range

3.3 Data processing methodology

This research is focused on implementing the ESDA techniques in understanding the spatial pattern of the study area. The household distribution dataset has been extracted from the Population census data and manipulated by using MS-Excel statistical software and projected by using ArcGIS 10.2 and GeoDa software.

3.3.1 Population census data

This research uses population census data which is carried out by the Ministry of Internal Affairs and Communications (MIC) in every 5 years. The population census dataset contains more than one hundred of variables, but for the research purpose, only the household data has been manipulated and tabulated. The boundaries of census area units are provided in an ArcGIS shape file as also can been seen from Figure 3.1 .

3.3.2 Household distribution dataset

Since the area of census units (polygons) are different, this study is further divided into 200 meter by 200 meter cells with population evenly allocated to each cell according to the total household number within the polygon. The figure 3.3 shows the distribution of the household in Tokyo after the data treatment by using ArcGIS 10.2. It contains 2, 61,171 cells in total.

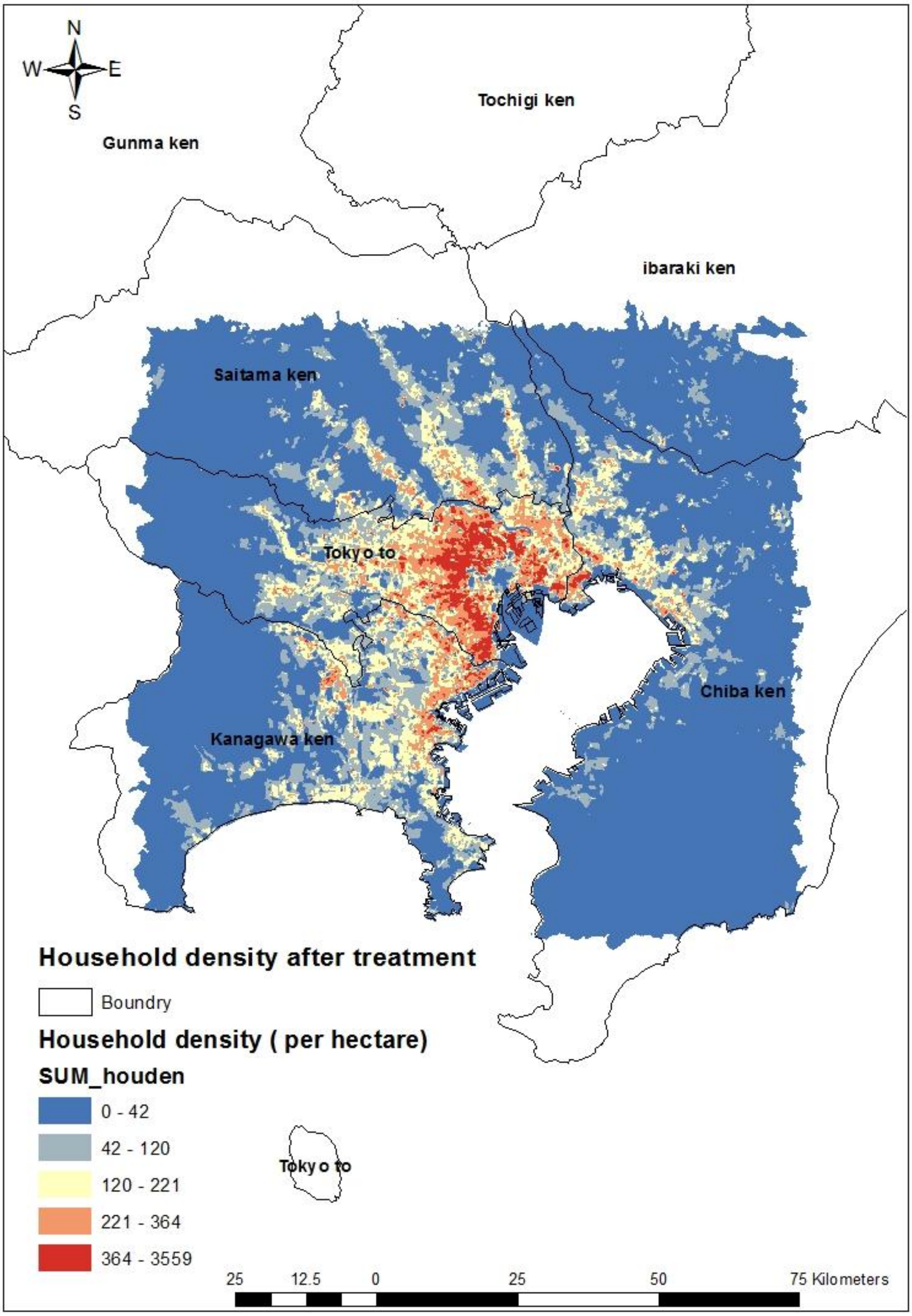

Figure 3.3: Household distribution after treating the data by using MS-Excel and GIS

3.3.3 GIS and spatial statistical toolset

For the tabulation and manipulation of the extracted data, MS-Excel statistical software has been used. For the selected research, ESDA techniques have been used. The two types of spatial statistical methods of ESDA techniques have been used for this research study. From the Global spatial statistical methods or also called the Global Indicator of Spatial Association (GISA), Global Moran's I has been selected. From the Local spatial statistical methods, Local Indicator of Spatial Association (LISA) which includes Local Moran's I and Local G statistic toolsets have been selected for this research study.

The spatial statistical toolsets have been analyzed by using the ArcGIS 10.2 and GeoDa software.

CHAPTER 4

GLOBAL MORAN'S I IN ANALYZING SPATIAL STRUCTURE

4.1 Introduction

Global Spatial Autocorrelation or Global Moran's I, introduced by Patrick Moran in 1948 is used to measure the degree of the selected feature's distribution patterns (i.e. Cluster, dispersed, random) across the study area (Moran, 1948). In this chapter, the Global Moran's I or Global Spatial Autocorrelation is explained besides its practical use in studying the spatial structure of Tokyo.

4.2 Distribution Pattern

The distribution of a certain phenomenon can have 3 typical types of spatial pattern: Dispersed; Random and Clustered spatial pattern. The Dispersed spatial pattern means that each value from its neighboring values is located far from each other in a uniformed manner. The Random spatial pattern means the distribution of the values is homogenous or independent in nature. The Clustered spatial pattern means most of the values are concentrated to nearby locations or adjacent together (Goodchild, 1986). The figure 4.1 shows the image of the 3 types of spatial distribution pattern.

(1)

Figure 4.1: Clustering patterns

However, in the real world, distribution of a feature usually does not fall in any of these three types. Rather, they are in a continuum of dispersed to clustered pattern as shown in Figure 4.2.

Source: ArcGIS Resources[1]

Figure 4.2: Spatial Autocorrelation Pattern Analysis

[1] http://resources.arcgis.com/en/help/main/10.1/index.html#//005p0000000n000000

In order to identify the degree of clustering, for example, the distribution like Table 4.1, Global Moran's I Index provides numerical answer.

Table 4.1: An example of spatial distribution

9	*3*	*1*	*3*	*7*	*0*	*2*	*9*	*4*	*5*
0	*6*	*0*	*3*	*0*	*1*	*8*	*2*	*2*	*8*
7	*7*	*0*	*2*	*8*	*5*	*9*	*10*	*7*	*9*
8	*10*	*8*	*5*	*0*	*4*	*3*	*3*	*5*	*10*
5	*0*	*9*	*1*	*0*	*6*	*4*	*10*	*7*	*3*
8	*7*	*4*	*9*	*2*	*3*	*2*	*10*	*2*	*4*
0	*7*	*10*	*10*	*6*	*9*	*9*	*8*	*8*	*5*
9	*4*	*6*	*10*	*10*	*2*	*10*	*8*	*2*	*8*
5	*3*	*3*	*3*	*10*	*7*	*4*	*4*	*10*	*2*
2	*9*	*2*	*6*	*10*	*3*	*9*	*5*	*2*	*0*

4.3 Calculation of Moran's I Index value

According to **ArcGIS Resources**[2], the Moran's I Index value is given as:

The Moran's I statistic for spatial autocorrelation is given as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \quad (1)$$

where z_i is the deviation of an attribute for feature i from its mean $(x_i - \bar{X})$, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features, and S_0 is the aggregate of all the spatial weights:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \quad (2)$$

The z_I-score for the statistic is computed as:

$$z_I = \frac{I - \mathrm{E}[I]}{\sqrt{\mathrm{V}[I]}} \quad (3)$$

where:

$$\mathrm{E}[I] = -1/(n-1) \quad (4)$$
$$\mathrm{V}[I] = \mathrm{E}[I^2] - \mathrm{E}[I]^2 \quad (5)$$

2 http://resources.arcgis.com/en/help/main/10.1/index.html#/How_Spatial_Autocorrelation_Global_Moran_s_I_works/005p0000000t000000/

Additional calculations are as follows:

$$\mathrm{E}[I^2] = \frac{A - B}{C} \tag{6}$$

$$A = n\left[(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2\right] \tag{7}$$

$$B = D\left[(n^2 - n)S_1 - 2nS_2 + 6S_0^2\right] \tag{8}$$

$$C = (n-1)(n-2)(n-3)S_0^2 \tag{9}$$

$$D = \frac{\sum_{i=1}^{n} z_i^4}{\left(\sum_{i=1}^{n} z_i^2\right)^2} \tag{10}$$

$$S_1 = (1/2)\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{i,j} + w_{j,i})^2 \tag{11}$$

$$S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{i,j} + \sum_{j=1}^{n} w_{j,i}\right)^2 \tag{12}$$

As seen in formula (1), the Global Moran's I or Global Spatial Autocorrelation is identical to Pearson's correlations despite the fact the calculation is different (Paradis, 2009; Griffith, 1992). In A Pearson's correlation, Z_i and Y_i values are used whereas the Global Autocorrelation only have Z_i value. The cross-product term of the Pearson's correlation can be written as $(Z_i - \bar{Z})(Y_i - \bar{Y})$ which finds out whether the two values $Z_i \& Y_i$ are above or below at the same time from the mean while in Spatial Autocorrelation it is $(Z_i - \bar{Z})(Z_j - \bar{Z})$. The cross-product term of the spatial autocorrelation tries to find out the two locations of the Z_i value are above or below at the same time from the mean in space. This means if the Z_i value and its neighboring Z_j values are adjacent together and above the mean then the cross-product term will be positive. If the two locations are dissimilar, which means Z_i value is high but neighboring Z_j values are low, then the cross-product term will be negative.

To know the relationship among the $Z_i \& Z_j$ values, a weight W_{ij} (1, 0) is added which gives the locational relationship of the $Z_i \& Z_j$ values, which can be written as $W_{ij}Z_i\,Z_j\ or\ W_{ij}(Z_i - \bar{Z})\big(Z_j - \bar{Z}\big)$. for each adjacent locations of Z_i and neighboring Z_j values, a 1 is given for W_{ij}, otherwise 0.

The total observed values $\boldsymbol{n}$ is divided by the total weight $\sum_i \sum_j w_{ij}$ or S_o and sum of the variance ($\sum_i^n Z^2$) of Z_i value is added to the cross product term so that the output (which is called Global Moran' I Index) values falls in range between -1 to +1.

The Global Moran's I or Global Spatial Autocorrelation also calculates a Z score for hypothesis tests as seen in formula (3). Z scores are normally the standard deviation. The range of the Z score and associated P values determine whether to keep the null hypothesis or reject. A P-value is a probability which identifies whether the provided Z score is showing strong or weak evidence against null hypothesis. A high Z score with small P value indicates a positive spatial autocorrelation, which means a clustered pattern, (i.e. high values are related to high values and low values related to low values). A negative Z score with small P value indicates a negative spatial autocorrelation or dispersed pattern. A Z-score with high P value indicates no spatial autocorrelation or random pattern **(ArcGIS Resources[3])**.

4.4 Result of Tokyo and Interpretation

The Global Moran's I or Global Spatial Autocorrelation is implemented in studying the degree of spatial clustering of Tokyo by using ArcGIS 10.2. As introduced in Chapter 3, we selected household distribution in Tokyo. The figure 4.3 illustrates the output

---

3 http://resources.arcgis.com/en/help/main/10.1/index.html#/How_Spatial_Autocorrelation_Global_Moran_s_I_works/005p0000000t000000/

of the analysis where on the upper right corner the Moran's I Index, Z score and P value is provided.

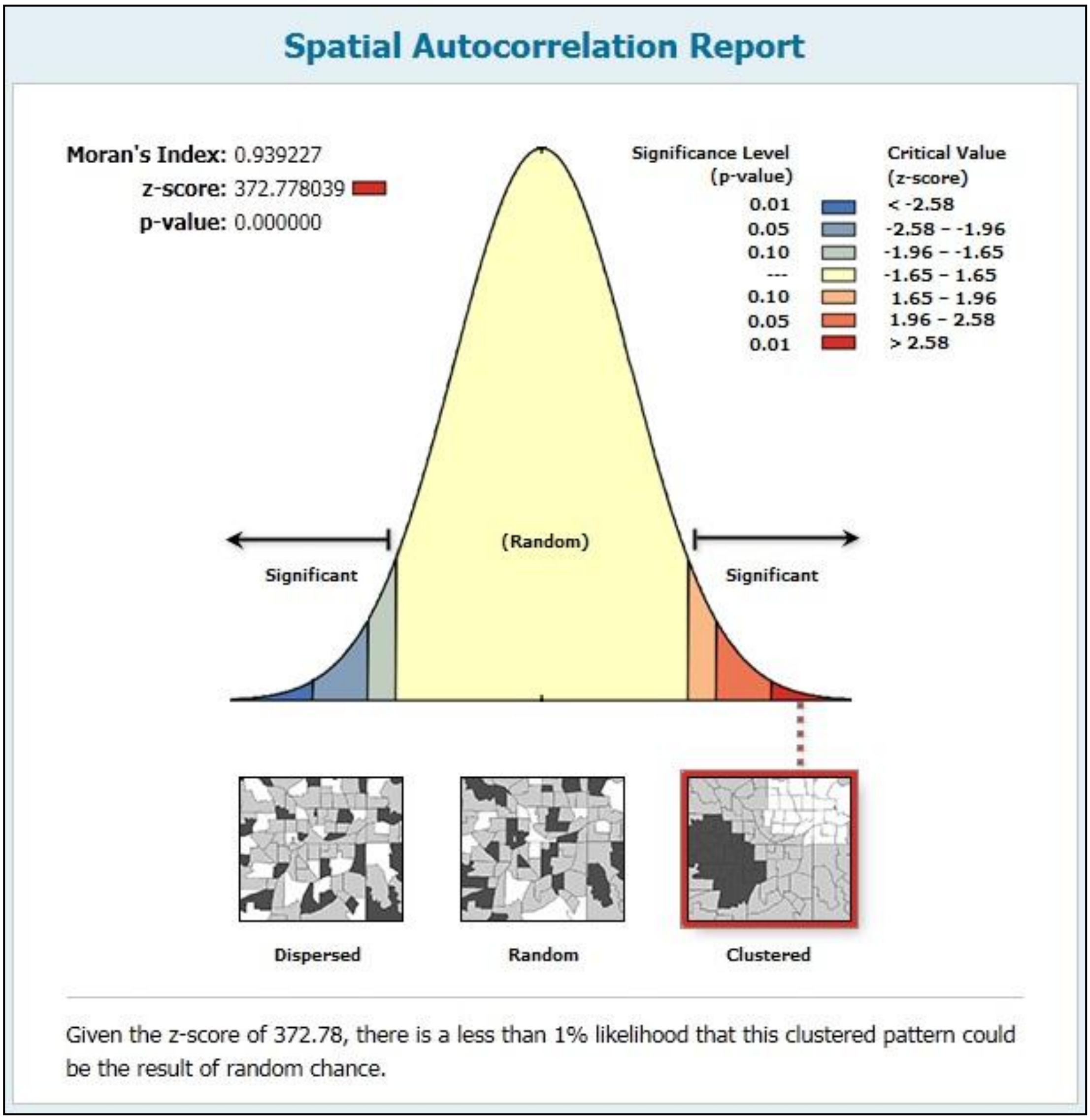

Figure 4.3: The Global Moran's I output

The Global Moran's I Index for the selected household distribution is 0.939 which is close to +1 range. The Z-score 372.778 and P-value 0.00000 indicates strong evidence that the selected data is showing a positive Spatial Autocorrelation. The null hypothesis that has been selected for this research is the household distribution in Tokyo area is having no specific spatial pattern. The Global Moran's I analysis pointed out that the household distribution in Tokyo area is having a clustered spatial pattern which in this case rejects the null hypothesis.

The upper left corner of the figure shows the different ranges of standard deviation based on the confidence level. The Z-score higher than +2.58 implies a 99% chance that the data is taking a clustered pattern which is summarized at the bottom of the figure.

4.5 Summary

Global Moran’s I or Global Spatial Autocorrelation analysis gives an overall result of what kind of pattern the data is imposing. Global Moran’s I or Global Spatial Autocorrelation analysis alone cannot provide enough evidence to identify the specific patterns of spatial differences. The output fails to provide evidence against whether the high density locations and low density locations are separately located or not. To find out the high and low density locations and also the dissimilar locations, LISA techniques have been used which is further discussed in chapter 5 and 6.

CHAPTER 5

LOCAL MORAN, S I IN ANALYZING SPATIAL STRUCTURE

5.1 Introduction

In chapter 4, by applying Global Moran's I technique, we found a strong evidence that a kind of clustered pattern is existing in Tokyo in terms of household density. However, the technique fails to identify what and where the clusters are. Local Moran's I, one of the techniques of LISA (one of the branches of ESDA techniques), introduced by Anselin in 1995, analyses and identifies the heterogeneity or difference in spatial patterns as seen in Figure 5.1[4]. In this chapter, the Local Moran's I have been discussed and applied in analyzing and finding out the location of different spatial patterns in Tokyo's household distribution.

Figure 5.1: Local Moran's I Output

5.2 Local Moran's I Calculation and Interpretation

Local Moran's I calculates an Index for each location in the area and an associated Z score as well. The calculation details are as below (ArcGIS Resources[5])

[4] http://resources.arcgis.com/en/help/main/10.1/index.html#//005p0000000z000000

[5] http://resources.arcgis.com/en/help/main/10.1/index.html#/How_Cluster_and_Outlier_Analysis_Anselin_Local_Moran_s_I_works/005p00000012000000/

The Local Moran's I statistic of spatial association is given as:

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(x_j - \bar{X}) \qquad (1)$$

where x_i is an attribute for feature i, $\bar{X}$ is the mean of the corresponding attribute, $w_{i,j}$ is the spatial weight between feature i and j, and:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_j - \bar{X})^2}{n - 1} - \bar{X}^2 \qquad (2)$$

with n equating to the total number of features.

The z_{I_i}-score for the statistics are computed as:

$$z_{I_i} = \frac{I_i - \mathrm{E}[I_i]}{\sqrt{\mathrm{V}[I_i]}} \qquad (3)$$

where:

$$\mathrm{E}[I_i] = -\frac{\sum_{j=1, j \neq i}^{n} w_{ij}}{n - 1} \qquad (4)$$

$$\mathrm{V}[I_i] = \mathrm{E}[I_i^2] - \mathrm{E}[I_i]^2 \qquad (5)$$

Additional calculations are as follows:

$$\mathrm{E}[I^2] = A - B \qquad (6)$$

$$A = \frac{(n - b_{2_i}) \sum_{j=1, j \neq i}^{n} w_{i,j}^2}{n - 1} \qquad (7)$$

$$B = \frac{(2b_{2_i} - n) \sum_{k=1, k \neq i}^{n} \sum_{h=1, h \neq i}^{n} w_{i,k} w_{i,h}}{(n-1)(n-2)} \qquad (8)$$

$$b_{2_i} = \frac{\sum_{i=1, i \neq j}^{n} (x_i - \bar{X})^4}{\left(\sum_{i=1, i \neq j}^{n} (x_i - \bar{X})^2 \right)^2} \qquad (9)$$

The Local Moran's I, derived from the Global Moran's I identifies the high and low values in local level which means specific locations of spatial clustering. The deviation of the X_ith value $(X_i - \bar{X})$ is multiplied by the sum of all the neighboring values deviations $\sum_j W_{ij}(X_j - \bar{X})$ value. The sum of the all neighboring values' deviations are selected based on the values with positive connectivity (W_{ij} = 1) with the X_i values which means the weight W_{ij} with value 1. In Local Moran's I calculation the (I_i) denotes the different measures of (I) which is the local Moran's I for each sub *i-th locations* on the map. The Local Moran's I is also referred as the decomposition of the Global Moran's I which means if the cross product term of Local Moran's I is summed up over all *i-th* locations then the summed values will be correlative to Global Moran's I. For example, a map is created by using the Global Moran's I. despite the fact the Global Moran's I gives an overall assumption, it's nearly impossible to do further calculation. For further analysis, the Global Moran's I can be distributed by separating into little pieces (i.e. the map into different locations) by using the Local Moran's I. After that, the separated values are added up to get the Global Moran's I. The Local Moran's I identifies the specific locations that are having significant positive and negative Autocorrelation (Anselin, 1995).

The method separates all the observed locations into 5 groups based on the Local Moran's I value and the significance level. This is explained in next section using a scatter plot. The output of the data than can be illustrated by providing a Cloropleth map.

5.3 Scatterplot and Cluster Types

If we plot the Z score of a location at the X-axis and the average Z scores of the locations surrounding it, we will get a scatter plot like figure 5.2:

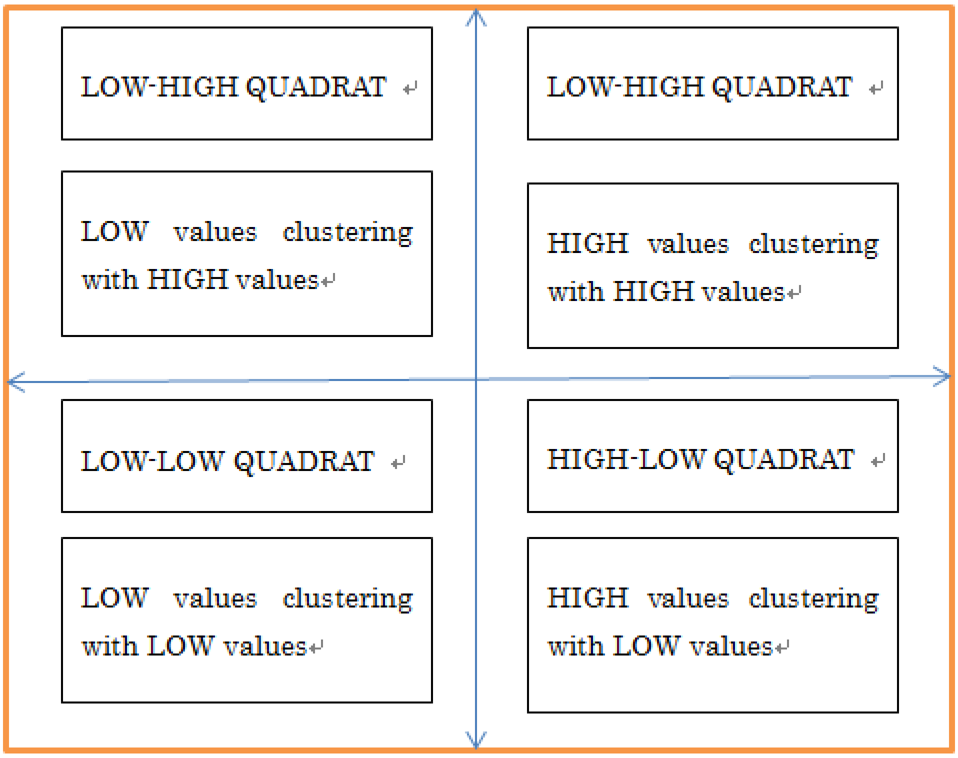

Figure 5.2: The 4 quadrat of the Scatter plot (Clustering Type)

According to the location on the scatterplot, the locations can be divided into 5 types:

- ***High-High group*** (i.e. high values are clustering with other neighboring high values and are above the mean);
- ***Low-Low group*** (i.e. low values are clustering with other low values and below the mean);
- ***High-Low group*** (i.e. high values clustering with low values and are below the mean)
- ***Low-High group*** (i.e. low values are clustering with high values and are below the mean).
- ***Insignificant*** (i.e. no significant clustering is existed);

Local Moran's I analysis divides all the location into 5 separate groups which can give a clear understanding of where the similar and dissimilar values are clustering. It not only identifies the high and low values but also addresses the outliers of the values. The first 2 groups (HH and LL) signify a positive Spatial Autocorrelation or in other words the high values are clustering with high values and low values are clustering with low values (Baumont et al., 2004; Anselin, 1995; Griffith, 2007). The 3rd and 4th group (HL and LH) signifies a negative Spatial Autocorrelation or dissimilar values or outliers.

5.4 Local Moran's I Calculation in Analyzing Tokyo Household Spatial Pattern

Local Moran's I has been used in studying the spatial pattern of Tokyo by focusing on the household distribution. For the analysis, ArcGIS 10.2 and GeoDa software are being used. The analysis of the Local Moran's I give 4 types of output: *1. Local Moran's I Index; 2. Local Moran's I Z-score; 3. Local Moran's I P- value and 4. Clustering type.* The Local Moran's I is used to find out the ***similar*** and ***dissimilar*** values. Similar values are the spatial clustering of high values which can be identified by the positive *Z-values* and the small *P-values*. For the dissimilar values which mean the values that are spread out from the neighboring values or also referred as the outliers are considered as imposing a negative spatial autocorrelation. Figure (5.3) shows the 5 types of clustering types. The *insignificant clustering group or pattern* contains the highest number of 164550 areas (cells), the *high-high clustering group or pattern* contains about 28255 areas, *low –low clustering group or pattern* contains 63207 areas, *low high clustering group or pattern* contains 125 areas and *high-low clustering group or pattern* contains 1 area.

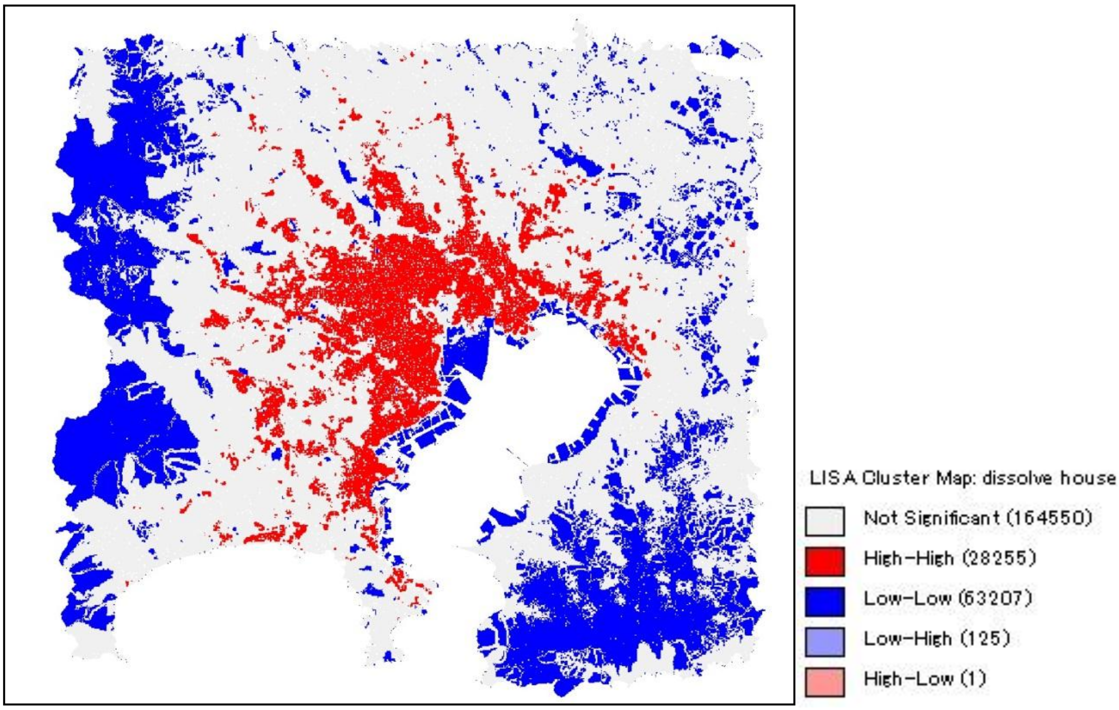

Figure 5.3: The 5 types of Clustering

The different clustering groups or pattern are explained in the next section separately. Since the HL clustering type having only 1 value or very small, further explanation is not provided.

5.4.1 HH type

The first quadrat (HH) contains the high values or positive values of the analysis which means the observed high household values (X) are surrounded by the high neighboring (Y) household values. It also can be said all the similar high observed household values are located nearby high neighboring household values. The analysis identified about 28255 locations as a high clustering of household distribution. The figure 5.4 illustrates the locations of the high clustering which can be seen located in the CBD areas of Tokyo.

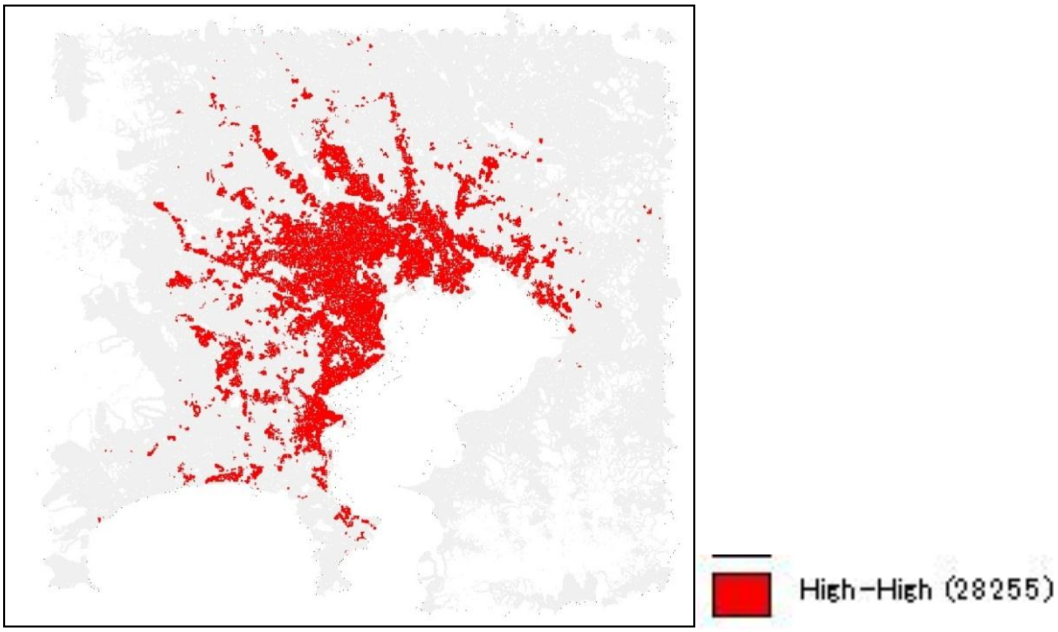

Figure 5.4: HIGH-HIGH Clustering Type

5.4.2 LL type

The third quadrat from the right of scatter plot is also positive in nature but contains the low values which mean the low household values (x) are surrounded by low household values. Unlike HH, the LL clustering pattern interprets the location of low density areas.

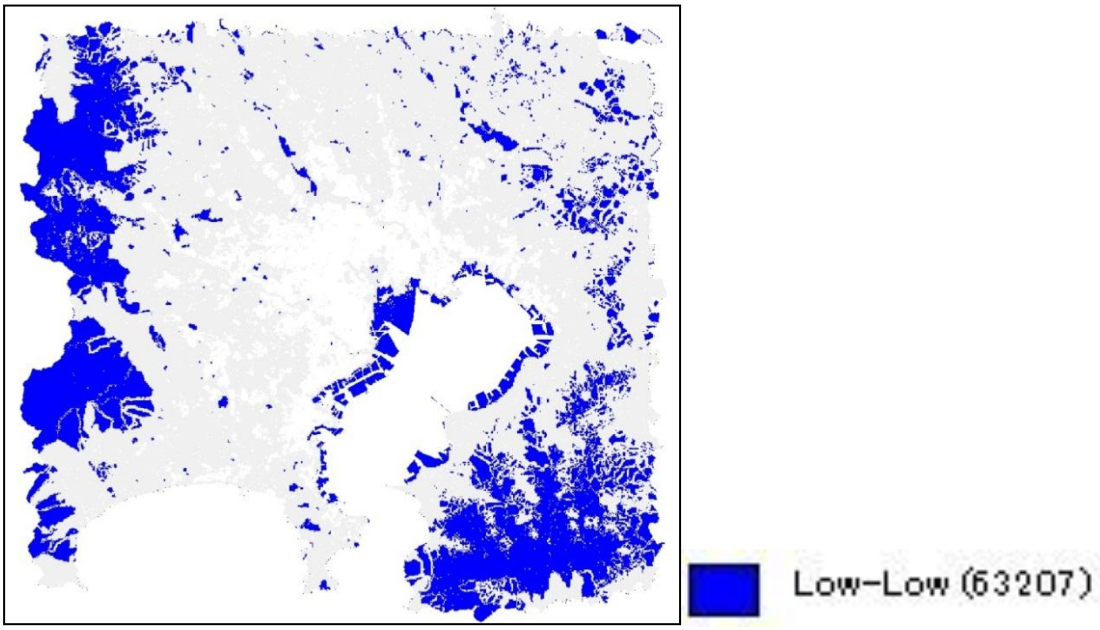

Figure 5.5: LOW-LOW Clustering Type

The analysis identified about 63207 locations as low density areas. The figure 5.5 shows the distribution of the low-low density areas in Tokyo which are located outside the central areas.

5.4.3 LH type

The 2nd quadrat of the scatter plot is different than the LL and HH quadrats. The LH clustering pattern imposes a pattern where the low and high values are clustering together which means low density of household values are neighboring with high density household values. The LH clustering pattern values are called the outliers or dissimilar values. In this case the locations are having a negative local spatial autocorrelation where observed household values are neighboring with high and low household values. The analysis identifies about 125 locations having a LH clustering pattern. The figure 5.6 illustrates the distribution of the LH clustering pattern values in Tokyo which can be seen existing within the HH household density values.

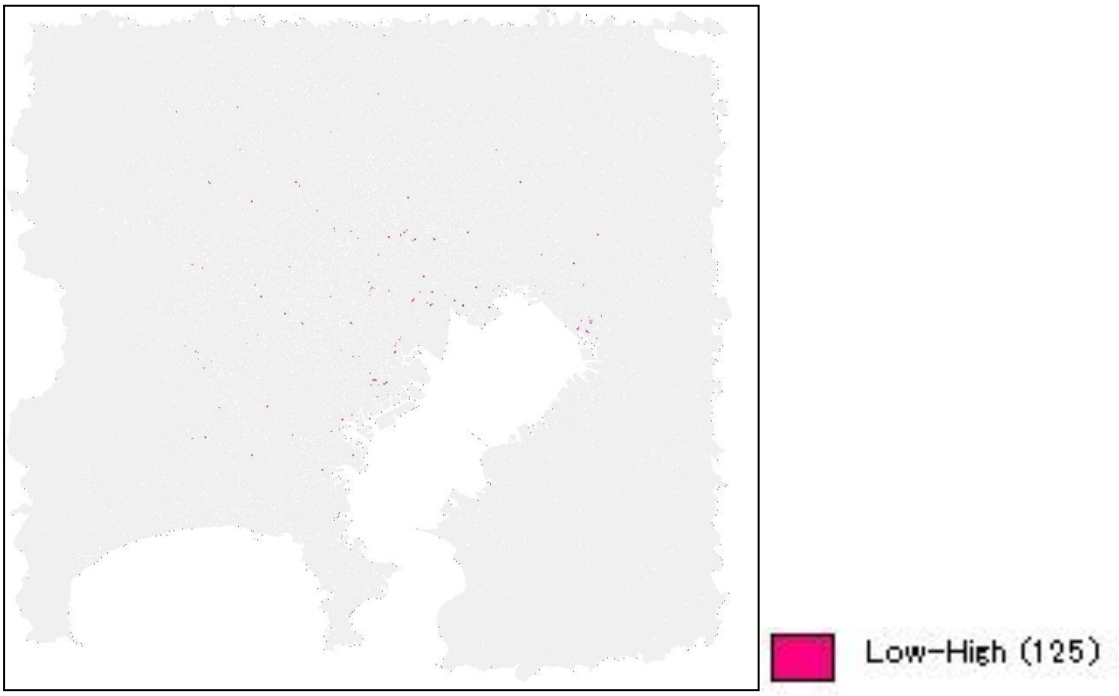

Figure 5.6: LOW-HIGH Clustering Type

5.4.4 No significance type (NS)

The 5th type of pattern that the analysis provides is the values which are showing no spatial clustering patterns. In this case the locations are imposing a random distribution. The analysis identifies about 164550 locations which are having no spatial autocorrelation or insignificant in nature. The figure 5.7 shows the distribution of the insignificant values in Tokyo area.

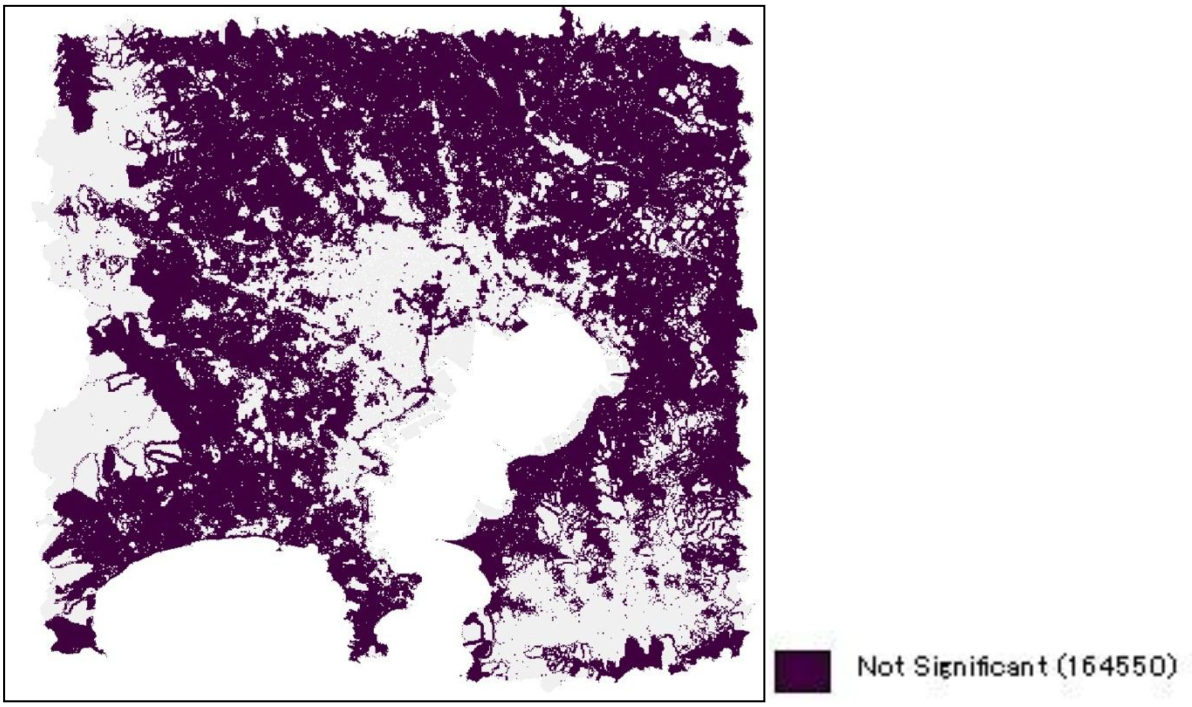

Figure 5.7: Not Significant Clustering Type

5.5 Summary

To summarize the chapter, the Local Moran's I or Local Spatial Autocorrelatoin analysis found out the specific locations of the high density houshold areas and low density houshold areas in Tokyo. Besides, the analysis also identified the outliers or dissimilar locaitons adjacent to high household density locations. The Local Moran's I identified clustering patterns in local level. To evaluate the Local Moran's I outputs another LISA technique has been used which is discussed in the next chapter.

CHAPTER 6

LOCAL G STATISTIC IN ANALYZING SPATIAL STRUCTURE

6.1 Introduction

The Global Moran's I performed for the research identified a clustered characteristic for the study area. The Local Spatial Autocorrelation (Local Moran's I) showed specific locations of spatial differences in the study area by spatial pattern types. Here, the 3rd spatial statistical method called Local G Statistic or Getis-ord Gi* or hotspot analysis has been implemented for a clearer understanding of the patterns.

The Local G Statistic is done alongside with the Local Moran's I or Local Spatial Autocorrelation to know the spatial association among the values whether high values are surrounded by high values and low values surrounded by low values. Both of the local versions of the statistics impose same characteristics but the difference is in illustrating the output of the analysis. The Local Moran's I categorized the observed values by four spatial pattern types or clustering types whereas the Local G Statistic tries to find out the association among the values. Another distinctive feature of the Local Moran's I is that it calculates the outliers and the analysis depends on the Local Moran's I Index values and excludes the observed values while in Local G Statistic includes all the values including the target values during the analysis.

6.2 Local G Statistic Calculation and Interpretation

Local G-Statistic or Getis-Ord Gi* or hotspot analysis can be written as follows ***(ArcGIS Resources[6])***:

The Getis-Ord local statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \qquad (1)$$

where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features and:

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n} \qquad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \qquad (3)$$

The G_i^* statistic is a z-score so no further calculations are required.

Local G statistic is computed to find out significant high and low values locations. The Local Moran's I identifies statistically significant high and low values locations on the map. The Local G statistic uses that information to finds out whether all the high and low values in the map are clustering together which means statistically significant hot spots and cold spots, because only a place can be referred as dense when the same high or low values are clustering with same high and low values. The result (Gi*) which is also pronounced a G-I –Star also can be written as $Gi *_{(d)}$ which means within a given distance (d), the *i-th* value will be summed with the neighboring *j-th* values. Suppose, within a specific distance (d), X is neighboring with P, Q, and R, S. here X is the i-*th* value and neighboring P,

[6] http://resources.arcgis.com/en/help/main/10.1/index.html#//005p00000011000000

Q, R, and S are j-th values. In the case of $Gi *_{(d)}$ both the X_i and the neighboring values X_j(P, Q, R, S) will be summed which can be written as $Gi *_{(d)} = (X + P + Q + R + S)$. The W_{ij} of the local G statistic will be the values within a specific distance (d) also can be written as $W_{ij}(d)$ or $W_{ij}\ (X + P + Q + R + S)$. The $W_{ij}X_j$ or $W_{ij}(d)X_j$ is implying the same meaning where the W_{ij} or $W_{ij}(d)$ *is* the total weight and the X_j values are the neighboring locations to the *i-th* value. The total sum of the weight values then subtract from the mean $\bar{X}$ which is multiplied by the total weight. In The denominator of the equation, the weight (W_{ij}=1 for each neighboring value) is multiplied by the total observed values then subtracted from the total $weight^2$ which is later divided by (n-1) (Getis and Ord, 1996).

Differing from the Local Moran's I which shows the evidence of spatial pattern types, here, and the Local G statistic produces the level of clustering. The analysis provides 2 types of output: 1. Z-score and 2. P-value. The z-score and p -values then determine whether the high clustering of value and low clustering of values are statistically significance or showing a high clustering or low clustering. A high Z score value with small significant P-value interprets the location is having a high clustering of values or also called the hot spots and a negative low Z-score with high P value implies a low clustering of values or cold spots. The higher the Z score, the intensity of the clustering is high, and smaller Z score means the intensity of the low clustering of values in space.

6.3 Local G statistic in Analyzing Tokyo household distribution pattern

6.3.1 Cluster map in different significance level

The local G statistic analysis provided a cluster map which is separated by the Z- scores and P-values to pinpoint the exact locations of high and low values of clustering. The figure (6.1) shows the clustering of high and low values or high and low density areas in

Tokyo. From the cluster map it can be observed that most of the high clustering values are near the central business district areas.

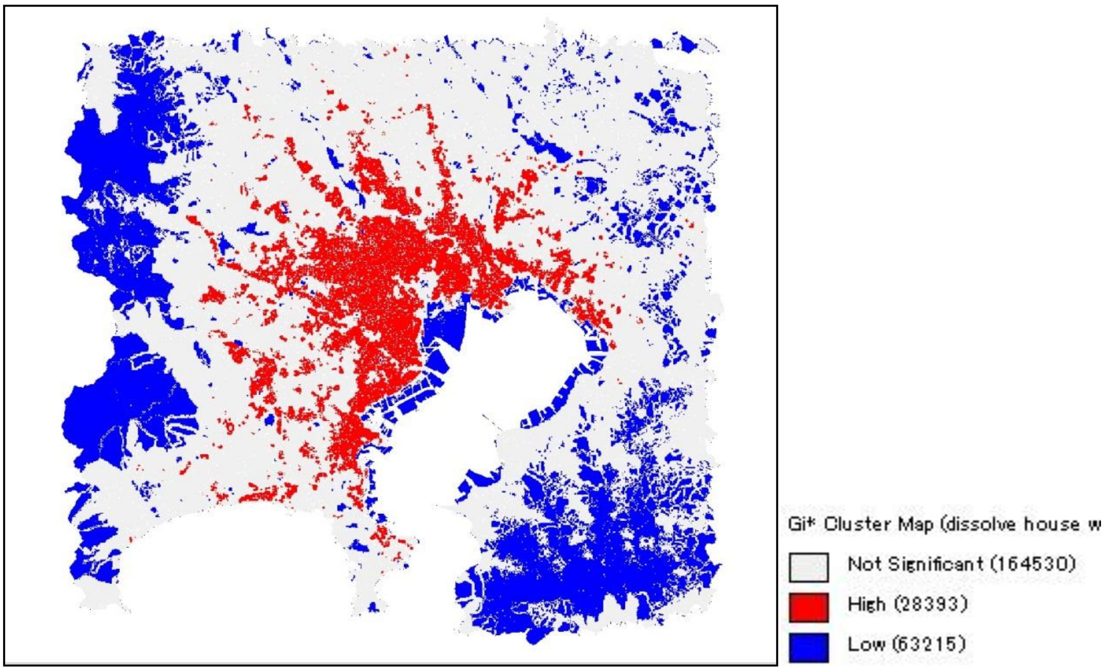

Figure 6.1: Cluster Map at 0.05 significance Level

The cluster map that has been given here is in 0.05 significance level or in 95% level of confidence. The figure 6.1 illustrates the high and low values locations where it can be observed that the Locations of the high values are 28393 and low values are 63215. Also from the analysis it is proved that the household distribution in Tokyo is showing a clustered spatial pattern.

The Local Moran’s I and Local G statistic provided maps are quite identical from the sense that both of the techniques have identified the clustering of high values and low values. However, Local Moran’s I output is based on the Local Moran’s I Index and Local G statistic output is provided based on the positive and negative Z scores. To be more specific, the Local Moran’s I only includes the neighboring values during the computation and Local G statistic includes all the values including the observed values in the computation to point out the exact locations of high and low clustering of values. For example, Local

Moran's I pointed out one location as outlier based on the Local Moran's I Index (-2.55) in table 6.2 where the household density is 473 (table 6.1). From the table it can be easily seen that the selected location is surrounded by dissimilar values or in other words having a negative Spatial Autocorrelation. But after applying the Local G statistic (table 6.3) the selected location is showing a high clustering of values with a Z-score of 2.98 and P value of 0.00. So, only by using the Local Moran's I for analysis could give arbitrary and misleading results. Henceforth, to evaluate the Global and Local Moran's I outputs Local G statistic is performed.

Table 6.1: Household density and clustering types

87	362	454	175
51	473	201	141
45	2	16	52
94	72	58	34

NS	HH	HH	HH
NS	HL	HH	NS
NS	NS	NS	NS
NS	NS	NS	NS

Table 6.2: Local Moran's I Index

1.48	5.05	17.48	3.50
-0.00	-2.55	8.11	1.30
-0.00	0.20	0.20	0.01
0.12	0.12	0.02	0.08

Table 6.3: Z scores of the selected locations

2.78	3.50	5.93	2.75
-0.02	2.98	4.57	1.62
-0.02	-0.64	-0.64	0.36
0.12	0.53	0.25	0.47

To test if there is a probability that the density of the household might be more clustered in more specific locations, this research also used different significance level up to 0.001 level or 99.9% level of confidence to see the changes in high and low values of clustering.

6.3.2 Confidence level of 99% or 0.01 significance level

The following figure 6.2 is the output of the Local G statistic with significance level of 0.01 or 99% level of confidence. The figure illustrates the high and low values locations where it can be observed that the Locations of the high values are 16011 and low values are 36509. From the map it can be observed that in the 95% level of confidence the analysis identified all possible clustering areas of high and low values. But in 99% level of confidence the analysis picked up denser locations based on the high Z scores which in this case having a P value of 0.01.

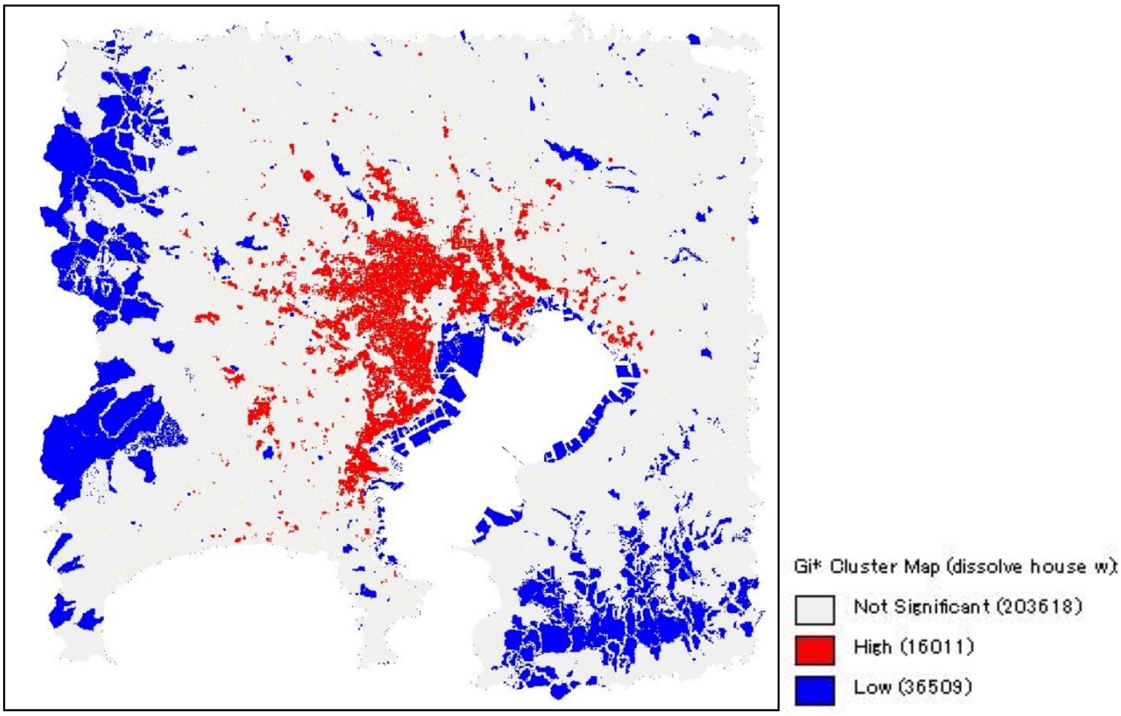

Figure 6.2: Cluster Map at 0.01 Significance Level

6.3.3 Confidence level of 99.9% or 0.001 significance level

The following figure 6.3 is the output of the Local G statistic with significance level of 0.001 or 99.9% level of confidence. The figure illustrates the high and low values location where it can be observed that the Location of the high values are 8065 and low values are 19172. In 95% level of confidence the analysis identified every possible high and low density locations. But the 99% level of confidence showed that not all areas are high or low density areas by identifying much more denser locations. The 99.9% level of confidence or 0.001 significance level only picked up the higest Z-scores which means the densest locations of houshold in the Tokyo.

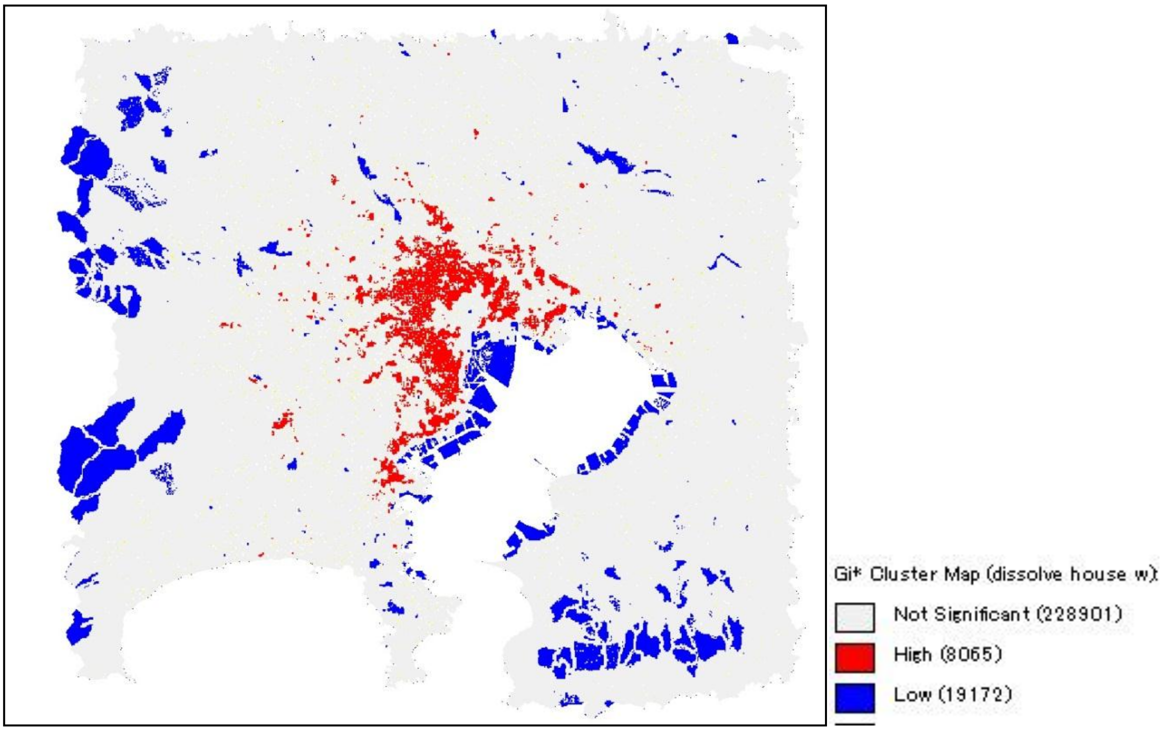

Figure 6.3: Cluster Map at 0.001 significance Level

6.4 Advantages of Using the ESDA techniques compare with Cloropleth map

The normal household distribution map shows the locations of high density areas (figure 6.4). But only by looking at the map it's difficult to determine the exact locations of household density in the case which places are highly dense areas and low dense areas. After analyzing the same household distribution data by using the ESDA techniques revealed locations which cannot be identified only from the normal distribution map. The figure 6.4 and 6.5 are provided to show the differences of before applying the ESDA techniques and after. Both Cloropleth maps have been overlaid in figure 6.6 to show the locations identified after using the ESDA techniques. The red colored locations are identified as high density areas after using the ESDA techniques.

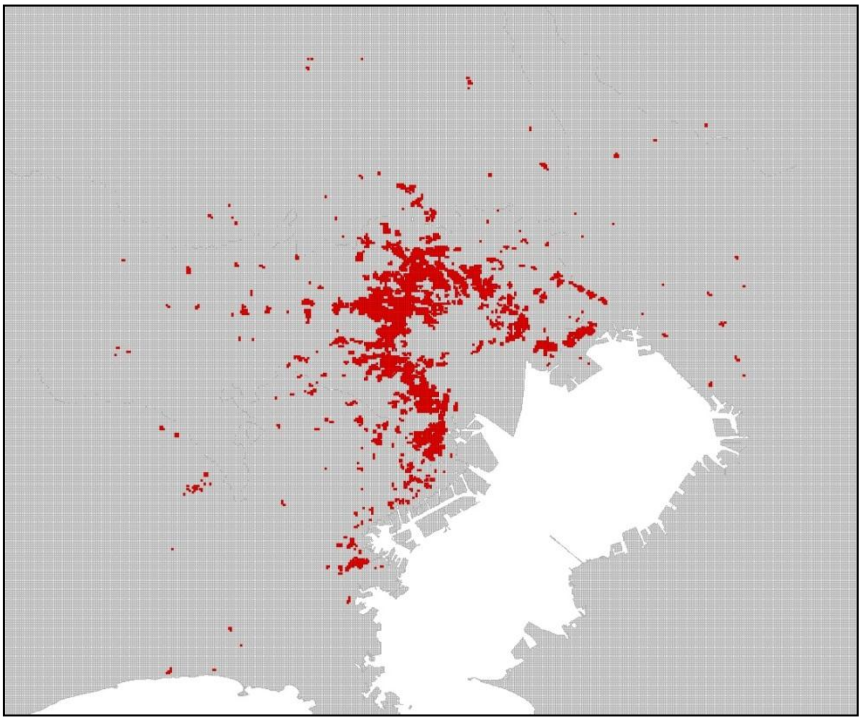

Figure 6.4: Normal Cloropleth map of household distribution

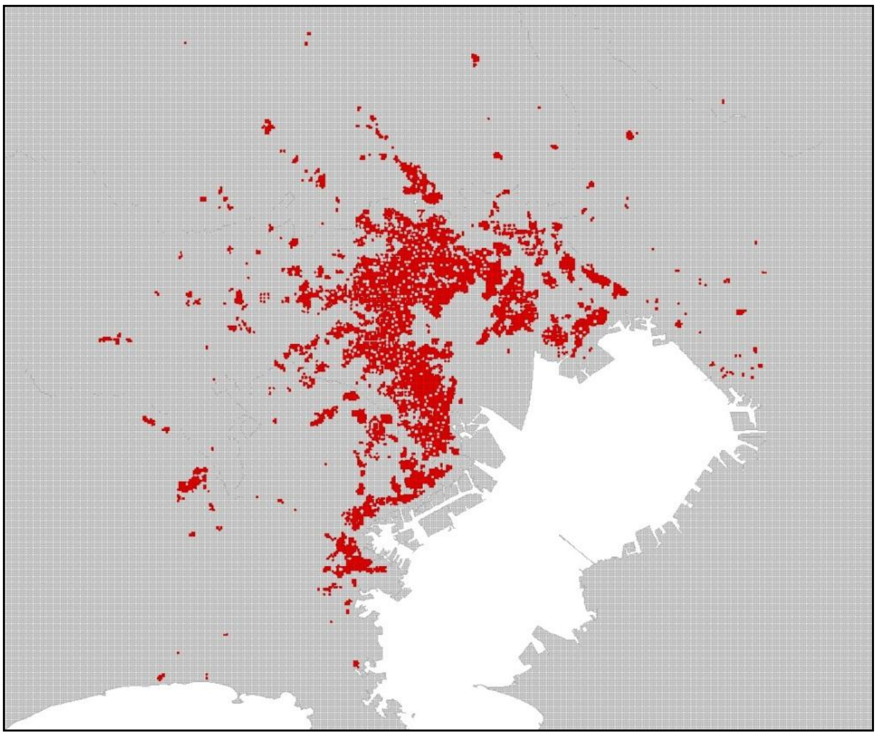

Figure 6.5: Cloropleth map by using the ESDA techniques

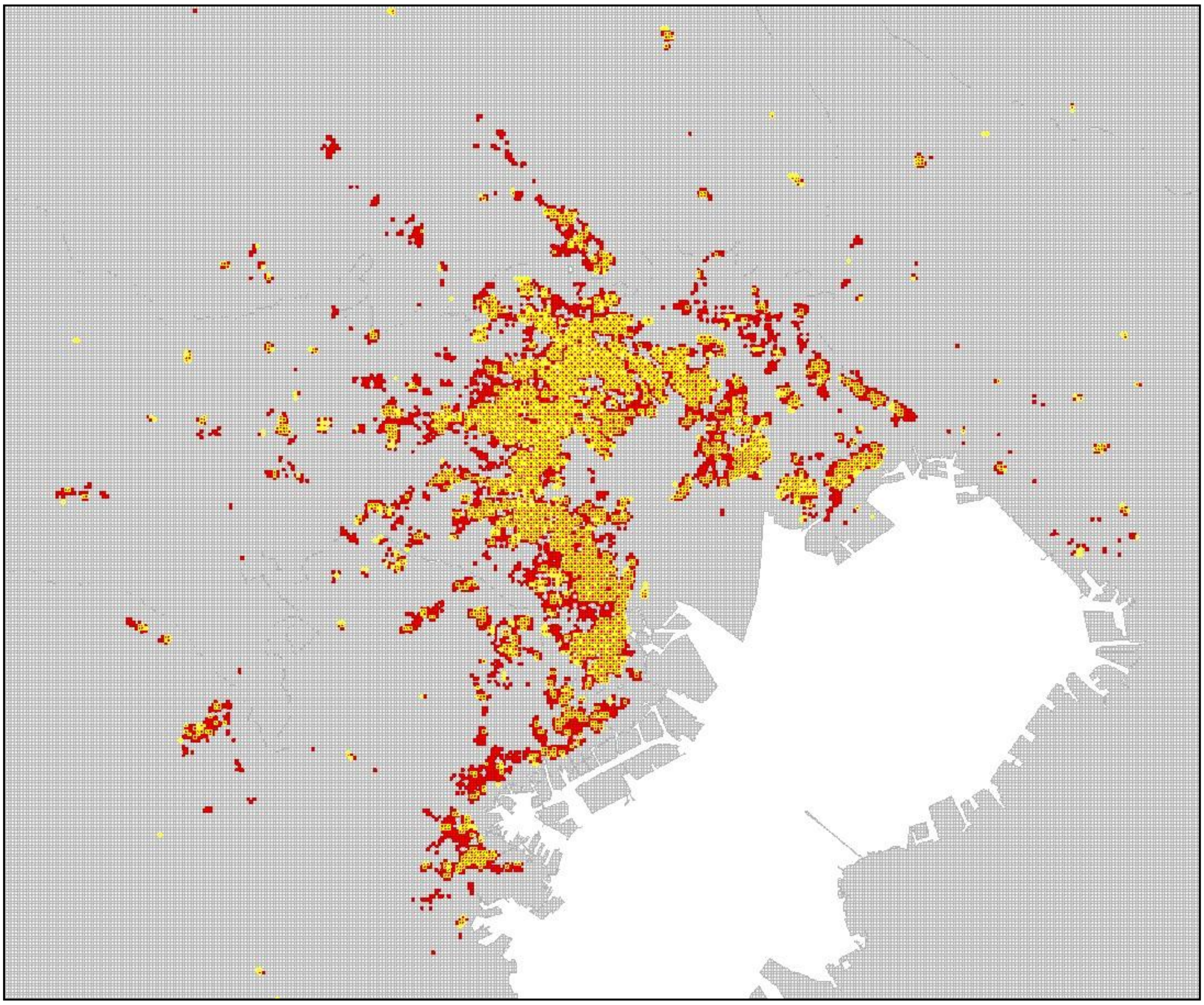

Figure 6.6: Overlaid map of the household distribution

6.5 Summary

To summarize, from the above findings it can be concluded that the analysis of Local G Statistic identifies the exact location of the high and low clustering of values. In this case the household distribution analysis has been chosen to observe the specific locations of the density and distribution of the household. The analysis revealed the exact locations of clustering which can be seen outside the Tokyo central areas. The significance level test proves against null hypothesis that the household density is clustered rather than dispersed in Tokyo area. Moreover, the significance test gives a real representation of the data which can provide the true geographic framework of a distribution pattern. The table 6.4 is given below to show the different P values based on the Z scores and the number of the locations under each different P values.

Table 6.4: Different P-values and the Z scores

Cluster	**Z-score**	**P-values**			
		P < 0.001	P < 0.01	0.01< P < 0.05	0.05 < P < 0.10
High Density areas	**38**	1			
	14	2			
	12	3			
	11	4			
	10	14			
	9	46			
	8	248			
	7	561			
	6	1133			
	5	2002			
	4	3475			
	3	3790	1977		
	2		3604	6722	
	1			553	4832
					Low density areas
Total Locations	**261171** (Including all the insignificant areas)				

P-value < 0.001 (Very strong evidence against null hypothesis)
P-value < 0.01 (Very strong evidence against null hypothesis)
0.01 < P-value < 0.05 (Strong evidence against null hypothesis)
0.05 < P- value< 0.10 (Weak evidence against null hypothesis)

Chapter 7

CONCLUSION

7.1 Findings

The research study started by explaining the importance of understanding the urban spatial structure which is important for an urban area's progress and prosperity from economic, environmental, social and political point of view. A spatial statistical- based approach has been used which can define, estimate and provide viable information in studying the urban spatial structure and pattern, characteristics and can be applied in studying entire urban areas for analyzing the urban spatial structure.

The main research aim is to understand the spatial structure of the Tokyo metropolitan area. For that purpose, a statistical-based approach which is called ESDA has been implemented. The main research objectives imposed for this research study are- *Implement Global Moran's I, Local Moran's I and Local G statistic to analyze the spatial structure of Tokyo metropolitan area* and *to demonstrate the performance and acceptability of the ESDA techniques in studying spatial pattern of Tokyo*. For addressing the research objectives 3 types of ESDA techniques are explained and discussed and analyzed in chapters 4, 5 and 6 respectively. Based on the outcome provided by the chapters 4, 5 and 6 the conclusions can be drawn as follows:

1. ***ESDA(exploratory spatial data analysis) method can be used in analyzing the spatial structure of Tokyo***

The research study has proved the acceptability of the ESDA method in analyzing the spatial structure of Tokyo. Moreover, the ESDA techniques in analyzing the household distribution identified locations of high and low density areas beside specific spatial patterns which helped to understand the distribution of household.

2. Only understanding by distribution will not give a perfect picture of the urban spatial structure but should also include location based analysis

Several papers have used the ESDA techniques for understanding the spatial pattern of Tokyo area, but they have used only the global version of the ESDA techniques. A conclusion cannot be made based on only the general distribution of data to understand the spatial structure of the Tokyo. It is also recommended to understand the spatial pattern in local level means. The analysis used global version to identify the distribution type and later local version of the ESDA techniques to understand the locational spatial patterns.

3 Significance level test more specifically identified the locations of high and low density areas

Another advantage that ESDA techniques provide is the test of the intensity of the situation by separating the locations based on different clustering types. These clustering types help to understand the high and low density areas besides the areas with shared boundaries. The significance test identified exact locations of high and low clustering of household density which from a normal distribution map cannot be seen.

7.2 Contribution of the Study

The main purpose of the research study is to introduce spatial statistical methods in analyzing the spatial pattern of the Tokyo metropolitan area besides, understanding the spatial structure more clearly.

The thorough literature review pointed out that scholars have selected a specific location to understand the spatial distribution of the area. It is very likely that they lacked either advanced software or databases. Only analyzing a small area cannot provide enough evidence to draw a conclusion on the spatial structure of an urban area. In this case a method that has also been found seemed applicable to identify the entire spatial structure of Tokyo.

An extensive review on the methods that are in use to analyze and identify the urban spatial structure also addressed a statistical based approach called ESDA which has been used for Western and European cities to identify urban spatial structure. However, The ESDA techniques have not been properly implemented for studying Tokyo spatial structure. The unique characteristic of these techniques is that a huge amount of data input and analysis can be done which is the main reason behind selecting GIS based ESDA techniques for this research.

This research proves that ESDA techniques can be applied in studying the entire Tokyo region and also perform well in identifying the exact locations of clustering. The GIS based ESDA techniques in analyzing the spatial structure of Tokyo gave an overall picture of the spatial structure of household distribution and clustering type.

7.3 Limitations

This research conducted only based on population census dataset of year 2000. For the preliminary research purpose, the research is focused on household distribution dataset which in this case provided efficient and acceptable outcomes. The outcomes would be much better if they could be projected based on not only for 1 year but for several years. Besides, the research only focused on the household distribution in Tokyo metropolitan area and it is not sufficient to prove a certain spatial pattern. Moreover, this study only visualized the household distribution but did not explain the reason behind clustering of household in specific areas.

7.4 Future Direction

This research is conducted based on a statistical approach called ESDA which can perform well in analyzing and identifying the spatial pattern of the Tokyo spatial structure. Though the research picked only household distribution for preliminary research, it gave interesting outcomes regarding the spatial structure pattern of the Tokyo. Despite the fact that only one variable cannot provide enough evidence in identifying a spatial structure of an urban area, datasets of employment distribution and transport distribution could give a satisfactory outcome and understanding the overall pattern of Tokyo. For further analysis, the population census data which includes dozens of variables will be joined with the employment distribution datasets along with transportation datasets to identify the change of the patterns and understand the factors that determine the patterns of spatial structure of Tokyo.

REFERENCES

Alonso, W. (1964). *Location and land use. Toward a general theory of land rent.* London: Harvard University Press.

Alpkokin, P., Komiyama, N., Takeshita, H., & Kato, H. (2007). Tokyo metropolitan area employment cluster formation in line with its extensive rail network. *Journal of the Eastern Asia Society for Transportation Studies, 7*, 1403-1416.

An, S. K. (2008). Recentralization of Central Tokyo and Planning Responses. *Journal of Regional Development Studies*, 1-20.

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of economic literature*, 1426-1464.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis, 27*(2), 93-115.

Arribas-Bel, D. (2009). *How many? Using ESDA to evaluate polycentricity in the US cities.* University of Zaragoza, Department of Economic Analysis. Zaragoza: Department of Economic Analysis.

Bagan, H., & Yamagata, Y. (2012). Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40 years. *Remote sensing of Enviornment, 127*, 210-222.

Barredo, J. I., & Demicheli, L. (2003). Urban sustainability in developing countries' megacities: modelling and predicting future urban growth in Lagos. *Cities, 20*(5), 297-310.

Baumont, C., Ertur, C., & Gallo, J. l. (2004). Spatial Analysis of Employment and population Density: The Case of the Agglomeration of Dijon 1999. *Geographical Analysis, 36*(2), 147-176.

Baum-Snow, N. (2007). Suburbanization and transportation in the monocentric model. *Journal of Urban Economics*, 405-423.

Bertaud, A. (2001). *Metropolis: A measure of the spatial organization of 7 large cities.* Unpublished working paper.

Bertaud, A. (2004). *The spatial organization of cities:Deliberate outcome or unforeseen consequence?* Institute of Urban and Regional Development. UC Berkeley: Institute of Urban and Regional Development.

Bertaud, A., & Malpezzi, S. (1999). *The spatial distribution of population in 35 World Cities: the role of markets, planning and topography.* The University of Wisconsin.

Bertaud, A., & Malpezzi, S. (2003). *The spatial distribution of population in 48 world Cities: implications for Economies in Transition.* Madison: The Center for urban land Economics Research.

Boarnet, M. G. (1994). An empirical model of intrametropolitan population and employment growth. *Regional Science, 73*(2), 135–152.

Bollinger, C. R., & Ihlanfeldt, K. R. (2000). *Intrametropolitan Locational Patterns of People and Jobs: Which Government Interventions Make a Difference?* Lincoln Institute of Land Policy.

Calimente, J. (2012). Rail integrated communities in Tokyo. *The journal of transport and land use, 5*(2), 19-32.

Champion, A. G. (1989). Counterurbanization in Britain. *The Geographical Journal, 155*(1), 52-59.

Cho, S. (2011). Urban transformation of Seoul and Tokyo by logal redevelopment project. *ITU A/Z, 8*(1), 169-183.

Chorus, P., & Bertolini, L. (Spring 2011). An application of the node place model to explore the spatial development dynamics of station areas in Tokyo. *The journal of Transport and land use, 4*(1), 45-58.

Dieleman, F., & Wegener, M. (2004). Compact city and Urban Sprawl. *Built Environment, 30*(4), 308-323.

Doi, S. (1975). The development of spatial structure and transportation network in japan. *Geographical reports of tokyo metropolitan univeristy*, 77-82.

Doytsher, Y., Kelly, P., Khouri, R., Mclaren, R., & Potsiou, C. (2010). *Rapid urbanization and mega cities: The need for spatial information management.* FIG Commission 3. FIG Publication.

Estiri, H. (2012). Tracking Urban Sprawl: Applying Moran's I Technique in Developing Sprawl Detection Models. *Emergent Placemaking*, 47-53.

Fujita, M., & Kashiwadani, M. (1989). Testing the Efficiency of Urban Spatial Growth: A Case Study of Tokyo. *Journal of Urban Economics, 25*, 156-192.

Gamez, L. R., & Dallerba, S. (2012). Spatial Distribution of employment in Hermosillo, 1999-2004. *Urban studies, 49*(16), 3663-3678.

Garcia-Lopez, M., & Muniz, I. (2011). Urban spatial structure,agglomeration economics, and economic growth in barcelona: An intra-metropolitan perspective. *Papers in Regional Science, 92*, 515-534.

Gavin, W. (2000). Megacities in the Asia-Pacific region. *Biennial conference*, 1-13.

Getis, A., & Ord, J. K. (1996). Spatial forcusting: Local Spatial Statistics An overview. In p. l. Batty, *Spatial analysis: modeling in a GIS environment* (pp. 261-278). John Wiley & Sons.

Glaeser, E. L., & Kahn, M. E. (2001). *Decentralized employment and the transformation of the American city.* National Bureau of Economic Research.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics, 63*(1), 1-24.

Goodchild, M. F. (1986). *Spatial Autocorrelation.* Norwich: Geo Books.

Griffith, D. A. (1992). What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. *In: Espace géographique*, 265-280.

Guiliano, G., & Small, K. A. (1999). The determinants of growth of employment subcenters. *Journal of transport geography, 7*, 189-201.

Hakim, I., & Parolin, B. (2009). Spatial Structure and spatial impacts of the Jakarta metropolitan Area: A southeast Asian EMR Perspective. *International Journal of Human and Social Science*, *4*(6), 397-405.

Hein, C. (2010). Shaping Tokyo: Land Development and Planning Practice in the Early Modern Japanese Metropolis. *Journal of Urban History*, *36*(4), 447-484.

Helsley, R. W., & Sullivan, A. M. (1991). Urban subcenter formation. *Regional Science and Urban Economics*, 255-275.

Hepp, S. (2011). *Metropolitan Spatial Structure: Measuring the Change.* University of Maryland. Unpublished doctoral Dissertation .

Hirooka, H. (2000). The developement of tokyo,s Rail Network. *Japan railway & transport review, 23*, 22-31.

Ibrahim, A. (1997). *Investigation of the Relationship between Urban Spatial Structure and Travel Demand in the GTA.* University of Toronto. ûttawa: Unpublished Doctoral Dissertation .

Ichikawa, H. (1994). the evolutionary process of urban form in Edo/Tokyo to 1900. *The town planning review, 65*(2), 179-196.

Inoue, R., Kigoshi, N., & Shimizu, E. (2007). Visualization of spatial distribution and Temporal change of land prices for residential use in Tokyo 23 wards using Spatio-Temporal Kriging. *10th international conference on computers in urban planning and urban management*, *63*, pp. 1-11. Tokyo.

Jones, G. W. (2001). *Studying extended metropolitan regions in South-East Asia.* Salvador de Bahia: General Conference of the International Union for the Scientific Study of Population (IUSSP).

Kawabata. (2003). *A Gis-based analysis of jobs, workers,and job access in tokyo.* Center for spatial information science . Tokyo: CSIS.

Kikuchi, T., & Obara, N. (2004). Saptio-temporal changes of urban frindge in tokyo metropolitan area. *Geographical reports of tokyo metropolitan university*, 57-69.

Lee, B. (2007). Edge and Edgeless cities? urban spatial structure in US metropolitan areas. *Journal of regional scince, 47*(3), 479-515.

Lewis, R., Knaap, G., & Schindewolf. (2013). The Spatial Structure of Cities in the United States. *Smart growth America*.

Li, X., Yeh, & Gar-On, A. (2004). Analyzing spatial restructuring of land use patterns in a fast growing region using remote sensing and GIS. *Landscape and urban planning, 69*, 335–354.

McMillen, D. P. (2001). Nonparametric Employment Subcenter Identification. *Journal of Urban Economics, 50*, 448-473.

McMillen, D. P. (2003). The number of subcenters in large urban areas. *Journal of urban Economics, 53,* 321-338.

Mieszkowski, P., & Mills, S. E. (1993). The causes of Metropolitan Suburbanization. *The Journal of Economic Perspectives, 7*(3), 135-147.

Mills, E. S. (1972). *Studies in the Structure of the Urban Economy.* Baltimore: Johns Hopkins University press.

Mills, E. S., & Price, R. (1984). Metropolitan Suburbanizaiton and Central City problems. *Journal of Urban Economics, 15,* 1-17.

Mitra, R., Buliung, R. N., & Faulkner, G. E. (2010). Spatial clustering and the temporal mobility of walking school trips in the Greater Toronto Area, Canada. *health & Place, 16,* 646-655.

Moran, P. A. (1948). The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society. Series B (Methodological), 10*(2), 243-251.

Morita, T., Nakagawa, Y., Morimoto, A., Maruyama, M., & Hosokawa, Y. (2012). Changes and Issues in Green Space Planning in the Tokyo Metropolitan Area: Focusing on the "Capital Region Plan". *International Journal of Geomate, 2*(1), 191-196.

Murayama, Y. U. (2011). Testing local spatial Autocorrelation using k-order neighbours. In &. R. Yuji Murayama, *Spatial Analysis and Modeling in Geographical Transformation Process: GIS based application* (pp. 45-56). London: Springer.

Muth, R. F. (1969). *Cities and Housing* (Vol. 65). Chicago: University of Chicago Press.

Nakabayashi, I. (1986). Spatial Structure of damage intensity by disasters in Japan of the 1970s. *Geographical reports of tokyo metropolitan univeristy*, 275-292.

Okata, J., & Murayama, A. (2011). Tokyo's Urban Growth, Urban Form and Sustainability. *In Megacities*, 15-41.

Palumbo, G. (1990). Population decentralization within metropolitan Areas: 1970-1980. *Journal of Urban Economics, 27*, 151-167.

Pan, Q., & Ma, L. (2006). Employment Subcenter Identification: A GIS-Based Method. *Science and Urban Economics, 21*(2), 63-82.

Paradis, E. (2009). *Moran's Autocorrelation Coefficient in Comparative Methods.* Vienna: R Foundation for Statistical Computing.

Pernice, R. (2007). Urban Sprawl in Postwar Japan and the Vision of the City based on the Urban Theories of the Metabolists′. *Journal of Asian Architecture and builidng Engineering, 6*(2), 237-244.

Redfearn, C. L. (2007). the topography of metropolitan employment: identifying centers of employment in a polycentric urban area. *Journal of urban Economics, 61*, 519-541.

Rode, P. (2008). *Integrated city making Governace planning and transport.* London: Urban age programme.

Ross, S. L., & Bartolome, C. A. (2004). Who,s in charge of the central city? the conflict between efficiency and equity in the design of a metropolitan area. *Journal of urban Economics, 56,* 458-483.

SaadAllah, D. M., Ibtehal Y. El Bastawissi, & Ayad, H. M. (2013). Identification of evolving metropolitan sub-centers: a gis based method. *World academy of science*, *76*, 474-481.

Samat, N. (2007). Integrating Gis and cellular Automata Spatial Model in evaluating Urban Growth: Prospects and Challenges. *Jurnal Alam Bina*, *9*, 79-93.

Samat, N. (2009). Integrating GIS And CA-MARKOV Model In Evaluating Urban Spatial Growth. *Malaysian Journal of Environmental Management, 10*(1), 83-99.

Sexto, C. F. (2009). Is the Counterurbanization Process a Chaotic Concept in Academic Literature? *Geographica Pannonica, 13*(2), 53-65.

Shen, M. k. (2006). job accessibility as a n indicator of auto-oriented urban structure: a comparison of Boston and los Angeles with Tokyo. *Environment and plannign B: planning and design*, *33,* 115-130.

Smith, D. A. (2003). The number of subcenters in large urban areas. *Journal of urban economics, 53,* 321-338.

Smith, D. A. (2011). Polycentricity and Sustainable Urban Form: An Intra-Urban Study of Accessibility, Employment and Travel Sustainability for the Strategic Planning of the London Region .

Sohn, J. (2005). Are commuting patterns a good indicator of urban spatial strucutre? *Journal of Transport Geography*, *13*(4), 306-317.

Sorensen, A. (1999). Land Readjustemnt, urban planning and urban sprawl in the tokyo metropolitan Area. urban studies. *Urban studies*, *36*(13), 2333-2360.

Sorensen, A. (2001a). Building suburbs in Japan: continuous unplanned change on the urban frindge. *TPR*, *72*(3), 247-273.

Sorensen, A. (2001b). Subcentres and Satellite Cities: Tokyo's 20th Century Experience of Planned Polycentrism. International Planning Studies. *International Planning Studies, 6*(1), 9-32.

Sorensen, A. (2004). Major issues of land management for sustainable urban regions in Japan. Towards Sustainable Cities. *Towards sustainable cities*, 197-216.

Speare, A. (1991). Counterurbanization: the changing pace and nature of population deconcentration. *Population and Development Review, 17*(3), 539-543.

Stanislav, K. (2011). Urban spatial structure,Evoluiton of cities and urban population. *Seria Stiinte Economice*, 34-43.

Suzuki, T., & Muromachi, Y. (2010). Empirical analysis on the railroad development impact on local population density in Japan. *Journal of the Eastern Asia Society for Transportation Studies, 8*, 1039-1052.

Tonuma, K. (1998). Tokyo: Policies toward the 21st century. *Ekistics*, 388-390.

Uchiyama, Y., & Okabe, A. (2012). Categorization of 48 Mega-Regions by Spatial Patterns of Population Distribution: The Relationship between Spatial Patterns and Population Change. *48th ISOCARP congress*.

Watanabe, Y. (1972). Some aspects of recent japanese metropolitan growth. *Geographical reports of tokyo metropolitan university*, 51-62.

Watanabe, Y., Takeuchi, K., Nakabayashi, I., & Kobayashi, A. (1980). Urban growth nad landscape change in the tokyo metropolitan area. *Geographical Reports of Tokyo metropolitan University*, 1-26.

Wiechmann, T., & Pallagst, k. m. (2012). Urban shrinkage in Germany and the USA:A Comparison of Transformation Patterns and Local Strategies. *International Journal of Urban and Regional Research, 36*(2), 261–280.

Xi, F., He, H. S., Hu, Y., Wu, X., Bu, R., Chang, Y., et al. (2009). Simulate urban growth based on RS, GIS, and SLEUTH model in Shenyang-Fushun metropolitan area northeastern China. *Urban Remote Sensing Joint Event*, 1-10.

Zhao, Y. A., & Murayama, Y. U. (2005). Effect characteristics of spatial resolution on the analysis of urban land use pattern. *Cities in global perspective: Diversity and transition*, 585-594.

Zhao, Y. A., & Murayama, Y. U. (2006). Effect of spatial scale on urban land-use pattern analysis in different classification systems: An empirical study in the CBD of Tokyo. *Theory and Applications of GIS, 14*(1), 29-42.

Zheng, X.-p. (1990). The spatial Structure of Hierarchical inter-urban system: equlibrium and optimum. *Journal of Regional Science, 30*(3), 375-392.

Zheng, X.-P. (1991). Metropolitan Spatial Structure and its Determinants : A Case-study of Tokyo. *Urban Studies, 28*(1), 87-104.