IHS
Making cities work

IHS is the international institute of urban management
of Erasmus University Rotterdam

MSc Programme in Urban Management and Development

Rotterdam, The Netherlands

September 2017

Thesis

Title: The relationship between street network and land use

Name: Chenxue Yang

Supervisor: Luca D'Acci

Specialization:

MASTER'S PROGRAMME IN URBAN MANAGEMENT AND DEVELOPMENT

(October 2016 – September 2017)

The relationship between street network and land use

Chenxue Yang

China

Supervisor: Luca D'Acci

UMD 13 Report number:1006

Rotterdam, September 2017

Summary

Since the opening-up policy in 1980s, there has been a fast expansion of China's transport facilities, and the construction of a transport network covering the whole country. With the high-speed development and huge amount of financial investments, it is important to better understand urban development and make policy based on these objective regularities, so to improve the resources use efficiency and achieve the goal of sustainable development.

Based on this background, this research aims to investigate the relationship between land use and street network, so to find whether street properties can capture the land development patterns of different land-use types. This paper investigates this relationship in Ningbo, China. Street property is measured by centrality indicator: closeness, betweeness and straightness on the street network. Six land use types are collected in the study area: commercial land, residential land, industrial land, public services, grass land and forest. Kernel density estimation is used to convert datasets into a basic raster unit. The relationship between each land use and each street network centrality is analyzed by Pearson's correlation. In addition, geographically weighted regression was used to show how this relationship varied across space.

The results indicate that commercial land, public services land and residential land are highly correlated with betweeness and straightness centralities. For the industrial, forest and grass land, there is almost no correlation between three centrality indices and these land uses. Furthermore, the relationships vary not only among different land-use types but also in a single land-use type. Spatial heterogeneity also exists in these relationships. This finding confirms that street centrality can capture the land development of different land uses and plays a important role in shaping urban fabrics.

Keywords:

Street centrality; Land use; Kernel density estimation; Spearman correlation; Geographically weighted regression.

Acknowledgements

First of all, I want to thank my parents for their love, support and forgive in the last 23 years. They provide me so much that I can grow up with happiness. Especially they give me the opportunity to study abroad in Netherlands. This year will become a cherished memory in my life.

Secondly, I want to thank my instructor, Luca D'Acci. Every time when I met problems, I know he would be there and gave me all supports and helps. Also he gave me a lot of encourage and insights, without his guidance I cannot finish my thesis.

I further extend my sincere thanks to my second reader Lei Qu for giving me patient guidance in the process of thesis writing. I really learned a lot from her. Without her help, my thesis can never be in a good structure.

I further extend my appreciation to Mr. Ronald Wall and Mr. Jan Fransen for the instruction in the study process. They are so kind and always gave me a lot of suggestions and instructions in the course period.

Last but not least, I thank all the classmates. I made a lot of friends from different countries in this year. And most importantly I became more vibrant, and learned how to share, to inspire, and to embrace each other in the diverse community.

Abbreviations:

HSR	high-speed rail
CBD	central business district
MCA	multiple centrality assessment
GIS	geographic information system
OSM	Open-Street Map
VGI	volunteered geographic information
GPS	global positioning system
KDE	kernel density estimation
GWR	geographically weighted regression
StdResid	standardized residual

Table of Contents:

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Background:

Since the 1980s, there is a rapid urbanization in China because of the opening-up policy. China's economy and international trade also developed rapidly and resident incomes increased, it reflecting and driving a fast expansion of China's transport facilities, and the construction of a transport network covering the whole country. Urban street network construction is one of urban infrastructure projects with huge investments in many countries. "By the end of 2012, China had 173,000 km of national-level roads, and accounted for 4 percent of China's road network at all levels, including national-level, provincial and rural roads" (Wang, 2016). According to the "National Road Network Plan (2013-2030)" in China, the national road will increase from 10.4 to 26.5 ten thousand kilometers, state highway will increase from 7.1 to 11.8 ten thousand kilometers by 2030. "This plan, compiled by the National Development and Reform Commission, will require an investment of 4.7 trillion yuan ($761.8 billion), according to information provided at a joint press conference. The network expansion aims to bring the length of national non-toll roads to 265,000 km and that of low-toll expressways to 118,000 km by 2030" (Wang, 2016).

Another important development is the high-speed rail (HSR) construction. "By the 1990s and 2000s, China's railroad construction was widely considered well-funded, with operating rail transit lines increased to around 66,000 km and connecting all major cities and countries along the railway networks" (Shaw, Fang, et al., 2014). "The development of HSR at the national scale began with the announcement of the Mid-long Term Railway Network Plan by the State Council in 2005, with the goal of expanding high-speed railroad length to 120,000 km HSR lines by the end of 2020 with a budget of around 4000 billion yuan"(Shaw, Fang, et al., 2014). By the end of 2012, there were about 17,000 km of HSR lines in operation, carrying trains of an average speed of above 200km/h.

With that high-speed development and the huge amount of financial investments, it is important to construct sustainable and resilience city through rational planning. However at the same time, there are lots of urban problems existed in this high-speed development progress. Congestion is one of the most prevalent problems in the large urban area, also it is a major factor in air pollution problems. At the same time, urban street network has a huge impact on urban land value and land rent. Nowadays some new town construction results in a severe lack of urban vitality and sometimes even deserted "ghost towns".

To solve and avoid these urban problems, it is important to better understand urban development and make policy and planning based on these objective regularities, so to improve the resources use efficiency and achieve the goal of sustainable development.

1.2 Problem statement:

This research intends to investigate the interaction between land use and street network, which is one of the natures of urban development. This issue becomes more and more important nowadays and long debated by experts, also it is at the core of the evaluation for the impact of many projects and policies in planning.

Urban street network construction is one of urban infrastructure projects with huge investments in many countries. To maximize the benefits from infrastructure investments in both intra city and intercity levels, one of the most important tasks is to evaluate different possible plans by simulating the regional and location influences brought by these new adding roads in each plan, and choose the optimal one. This research will provide a method to evaluate and simulate the location impact created by the street network. Meanwhile, because there are several different ways of transportation: walking on foot, traveling by bike, by car or by train, the location properties of the network will be analyzed by different means of transportation.

What's more, many studies showed that different categories of economic activities have different location preferential in the street network (Porta, Strano, et al., 2009a) (Wang, Antipova, et al., 2011), however, these studies limited in single land use type. Different land use preference of location in street network was poorly studied. This study will find this issue in the intra-urban context. If the results show significant preference by different land use in the street network, we can put up with advice for urban land use planning, so to realize the land value in a better way.

In conclusion, this research intends to contribute with policy advices in two ways. First is to quantify location properties in the street network. Second is to find out different land use location tendency in the street network. Finally, it can be found if there are relationships between the spatial distribution of land uses and street network, and try to answer whether street properties can capture the land development of different land-use types.

1.3 Research objective:

This research objective is to explore the relationship between the urban street network and land use. By understanding these relationships, this research aims to answer whether street network properties can capture the land development of different land use types? So to provide with policy advice for both urban street network and urban land use planning. To be specific, what are the spatial properties of a certain street network configuration? And what are the spatial distributions of different land use? By analyzing these two spatial characteristics, we can answer whether there is a relationship of spatial distribution between different land use and street network. This will give support and foundation for land use location choices.

1.4 Research questions:

According to the problem statement and research objective, the main research

question is:
To what extant are street network correlated with different land use?
To answer this main question, two sub-questions should be answered first:

1. What are the spatial properties of the street network?
2. What are the spatial distributions of different land uses in the study area?

1.5 Scope and limitation:

This study is an intra-urban study. The study region in this research should be a rapid development area with high-speed infrastructure construction nowadays. In recent years, a batch of new emerging cities rose in China in particular regions such as Yangzi River Delta and Pearl River Delta. Along the Chinese coastline, there is a city named Ningbo, located in Yangzi River Delta. Since the predominant geological location, this city experiences overwhelming socioeconomic development. Specifically, this metropolitan study area is polycentric as we can see in figure 1, which means there are two city centers in this region. What's more, few studies investigate this issue in polycentric cities, so in this research, I choose this region as the study area.
The data of street network and different land use in the study area are included in this research.

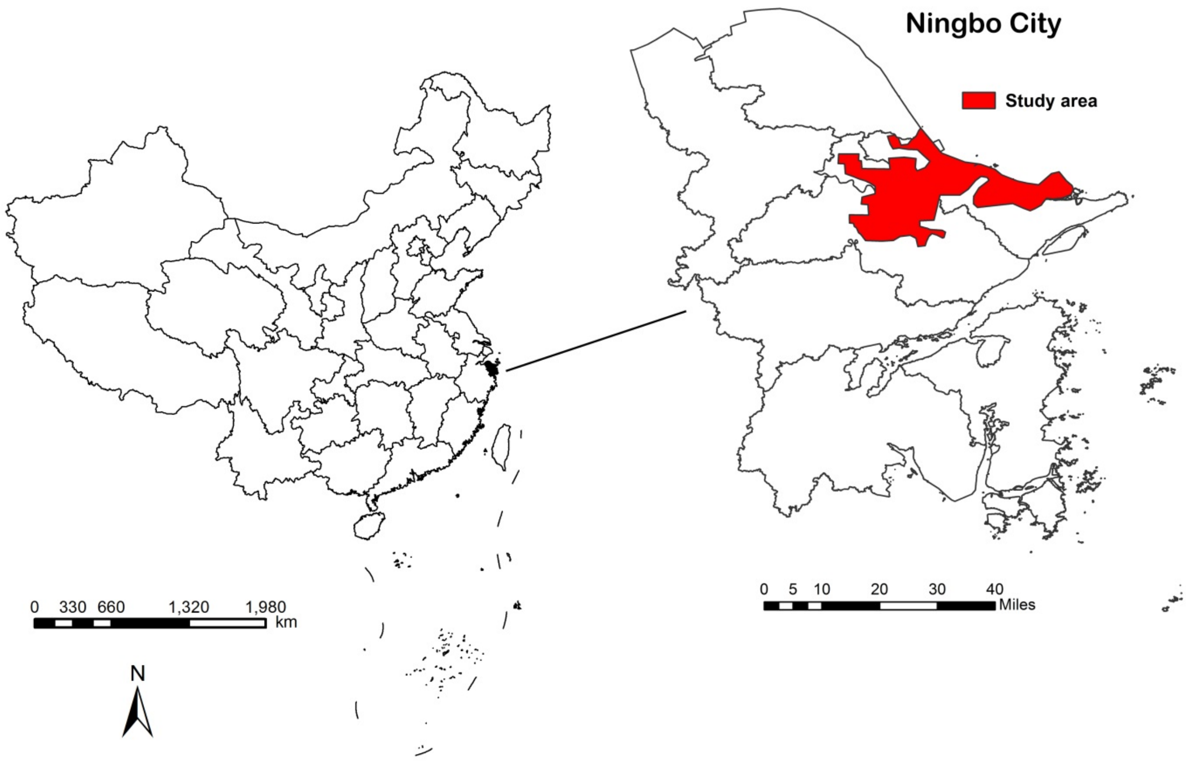

Figure 1 Location of the study area

There are also some limitations in this research. In many cases, land use types are determined by the government. That means the relationships between these two systems are also determined by many other factors. However, this study aims to find the relationship but not to find the driving forces. We can also find the relationship under all these factors.

There are also many indicators that can describe spatial properties of networks, such as accessibility, density, centrality etc. However this study will not quantify all these kinds of indicators. This study will just use one measurement to quantify one sort of indicators which are most related to land use development, this will be clarified in next part.

Chapter 2: Literature Review / Theory

2.1 Introduction

This chapter reviews literature on key concepts which are relevant to this research. This review will be organized based on three parts. First two parts introduce the basic definition of these two important city components: land use and street network. Last part is to introduce the literature of the relationship between street networks and land use.

2.2 Land use

"Land use refers to the activities accommodated with urban form, their distribution across buildings and open space, their intensity, or rate of change" (Sevtsuk, 2010). "Land uses are typically categorized into loose groupings, such as residential, commercial, industrial, and other activities" (Sevtsuk, 2010).

There are three main models of urban land use: "the concentric zone model, sector model and multiple nuclei model" (Bulmer, 1981). Concentric city radiates outward from a common center which is a business area. "Different land uses are distributed like concentric rings around city center". "Sector model represents the city that a series of sectors and radiating out from the CBD along major transportation lines" (Hoyt, 1939). "Multiple nuclei model means cities grow from a number of points (ports, universities, airports, park), rather than only one central that expand and merge in a single urban area" (Harris and Ullman, 1945). So apart from the CBD, there are several secondary, separated centers. Similar parts may group together for agglomeration economies. Figure 2 shows these three models. The study area in this paper belongs to multiple nuclei model.

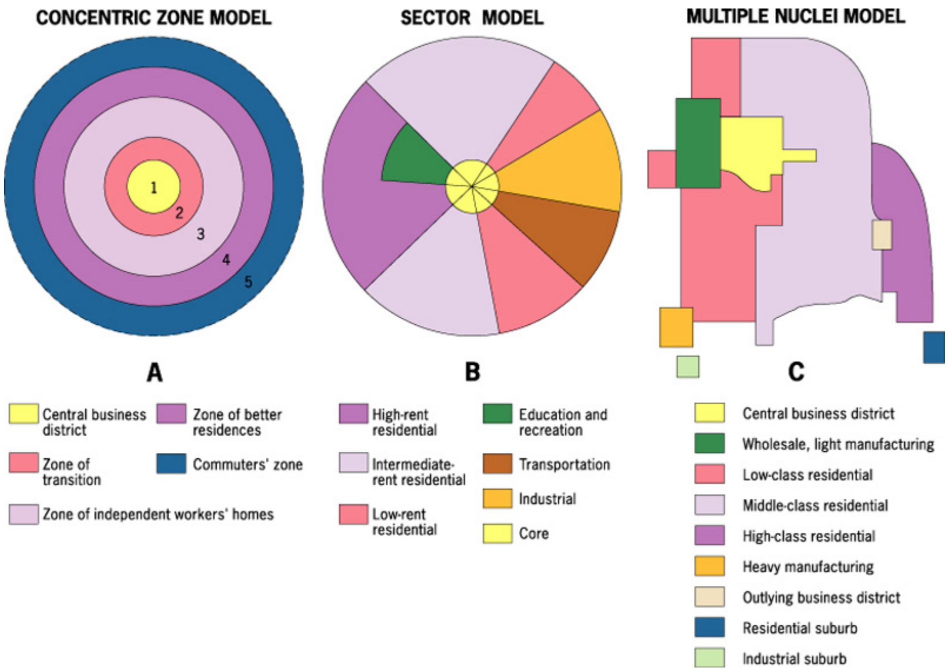

Figure 2 Three models of urban form (source: http://lewishistoricalsociety.com)

2. 3 Street network

Street network is the backbone of our built environment. “A more compact street network designs provide greater accessibility with a more direct route, reduce land consumptions and increase network efficiency”(Hickman, Bonilla, et al., 2015) . It also has effects on sustainable communities such as road safety, public health and travel patterns.

Street networks connect people to destinations and to each other. They provide public space for human interaction and foster economic activity. Also, they support a robust mix of commerce and culture. There are multiple mode-specific networks such as pedestrian, bike, cars and rails and sometimes they overlap, in other places they are separate. “A sustainable street network respects, protects and enhances ecological systems and the natural features of its urban environment” (Venckauskaite and Skrodenis, 2007).

“Cities are large collections of buildings held together by a network of space: the street network” (Al-Sayed, Turner, et al., 2014). “It is what holds city all together and it is a certain geometry and a certain topology and a certain scaling” (Al-Sayed, Turner, et al., 2014). During the course of the last century, there has been a dramatic change in street networks. Figure 3 shows this development of the street network in U.S. from the medieval pattern in the left of the figure.

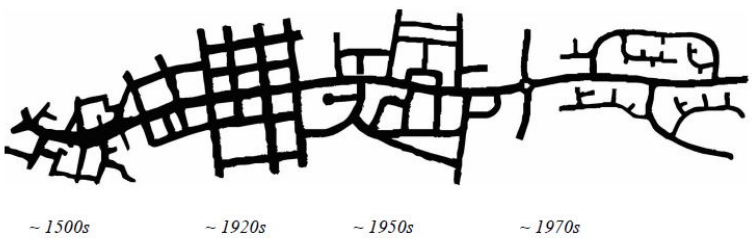

Figure 3 Street Patterns Evolution (Marshall, 2004)

2.4 The relationship between land use and street network

“Urban planners and researchers have long been interested in quantifying urban environments and analyzing the function, structure, and dynamics of the urban system” (Batty, 2008). “Recently, the issues involved in a complex urban structure have been the subject of a large number of theoretical and empirical inquiries” (Cardillo, Scellato, et al., 2006). Among these issues, “economic models (Casetti, 1993) and other studies (Erath, Löchl, et al., 2009) have highlighted the importance of transportation in shaping the urban spatial structure”.
The classic economic model proposed by (Mills, 1972) and (Muth, 1969), “often referred to as the monocentric model, assumes that all employment is concentrated in the city center” (Wang, Antipova, et al., 2011). “As everyone commutes to the city center for work, a household farther away from the center spends more on commuting and is compensated by living in a larger- lot house (also cheaper in terms of price per area unit)” (Wang, Antipova, et al., 2011). As we can see in figure 4, “the resulting population density exhibits a declining pattern with distance from the city center”. The economic model is ‘‘simplification and abstraction that may prove too limiting and confining when it comes to understanding and modifying complex realities”, but “highlights the important role of transportation costs in shaping urban structure” (Casetti, 1993).

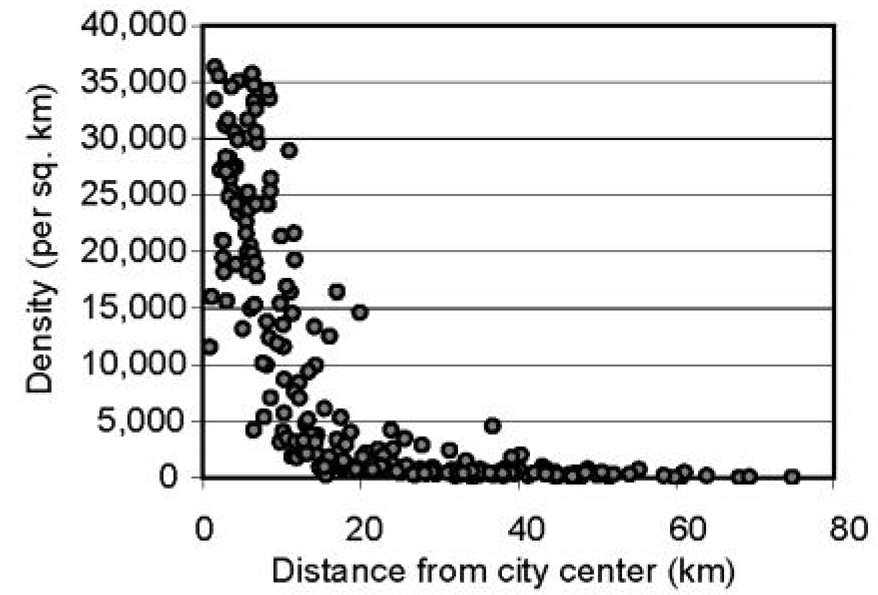

Figure 4 Population density in Beijing’s sub-districts (the year of 2000) (Wang, Antipova, et al., 2011)

Urban geographers and planners tend to be more recognizant of the complexity of urban structure, and have developed several models to capture the relationship between land use and transportation networks.

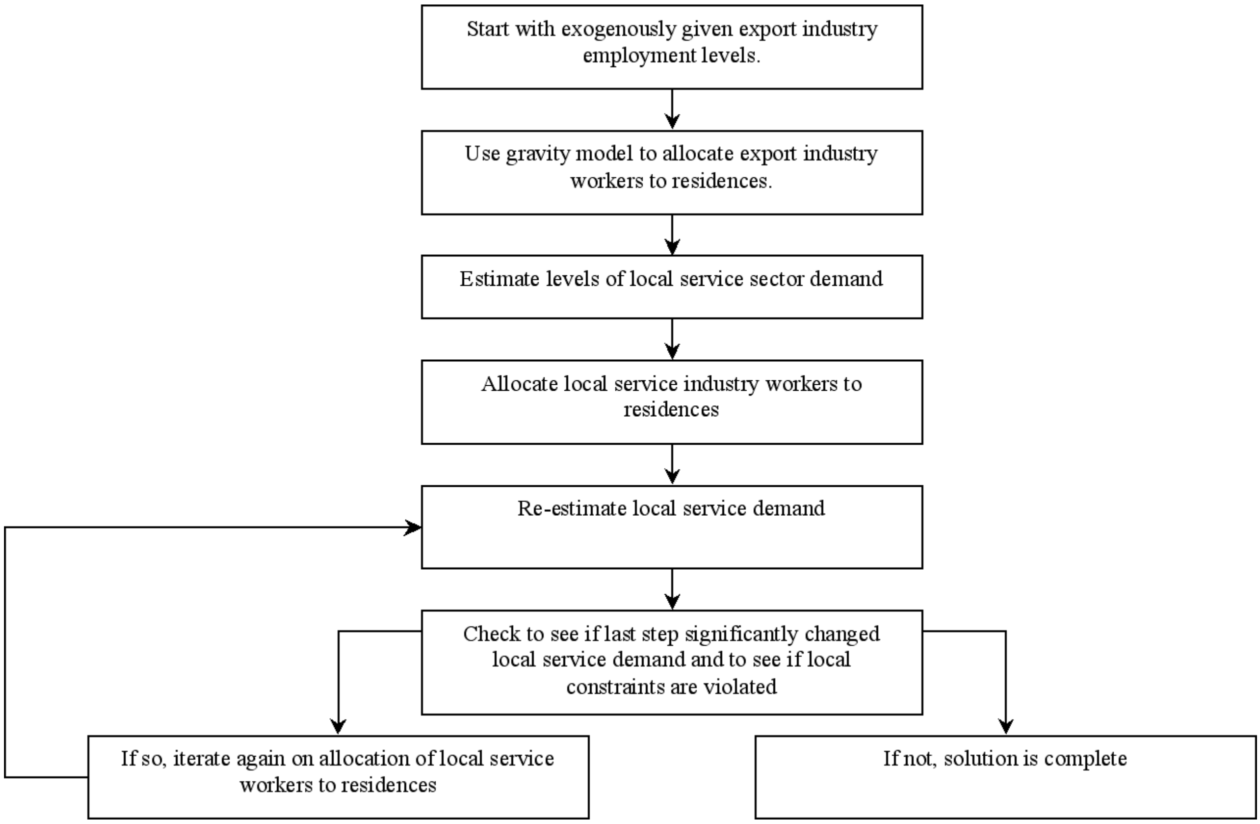

Figure 6 Lowry Model (Lowry, 1964)

"Transportation system and land use have been two most key subsystems in the complex urban systems. Transportation network and land-use development are assumed to mutually influence each other" (Wegener, 1994).

(Lowry, 1964) developed a model, which is shown in figure 5, emphasizes the interactions between population and employment distributions. This model predicts well the relationship between land use and transportation. It forecasts future land use changes and allocations and incorporates those changes into the transportation demand models. "The interactions between employment and population decline with distances, which are defined by a transportation network. The model has the flexibility of simulating population and employment distribution patterns corresponding to a given road network" (Wang, Antipova, et al., 2011).

(Wang and Guldmann, 1996) used this model to examine "how the population and employment density patterns respond to changes in transportation networks". However, this model divided employment into different sectors, which is infeasible in practice.

(Wang and Guldmann, 1996) proposed a "gravity-based model to simulate urban densities (in general, perhaps a combination of population and employment densities) given a road network". The model assumes that "density at a location is proportional to its accessibility to all other locations in a city, measured as a gravity potential". It emphasizes that "location determined by the road network is the force that shapes the variation of land use intensity" (Wang and Guldmann, 1996). Like any gravity models, the Wang–Guldmann model needs a value for the distance friction coefficient, which is not conveniently available and requires additional data and model calibration to derive.

Most recently, urban planners benefited from the advancement of network science. "A street network can be characterized by a well-defined geometric structure consisting of nodes and segments" (Batty, 2008). Graph theory characterizes the planimetric properties of the street network. This application on spatial networks suggests that geometric layout could affect the suitability of a land use location. Two currently popular street network models based on graph theory, which is a dual representation approach, exemplified in Space Syntax method, and the primal representation exemplified by *Multiple Centrality Assessment* (MCA) model.

"Networks of streets and intersections can be represented by spatial graphs in which intersections are turned into nodes, and streets are turned into edges or links. Because of the coherence between the graph entities, this kind of representation is hereby termed 'direct', or primal: analogously, representations in which streets are turned into nodes and intersections are turned into edges, are hereby defined 'indirect', or dual, that is, the case of conventional space syntax analysis"(Porta, Crucitti, et al., 2006b).

The major differences between these two approaches lie in the particular representation of the underlying graph environment and in the metrics that are used to measure the relationships between graph objects. Whereas MCA applications of graph theory use geometrical distances along the network to describe inter-relationships between graph elements, Space Syntax measure topological distances such as the number of turns and nodes to represent the shortest path.

"The space syntax model, which is based on the notions of visibility and integration, has been widely used in a substantial number of studies and helped establish correlations between street centrality and social dynamics" (Hillier, Hanson, et al., 1984). "The core of the space syntax method is the integration index, which is stated to be so fundamental that it is probably in itself the key to most aspects of the human spatial organization" (Croxford, Penn, et al., 1996). The integration value of street is defined as "how easy it is to get to that segment from all other segments; it is a measure of the to-movement potential having the segment as destination" (D'Acci, 2016). During the last several decades, their work has proposed that "street network configuration is related to diverse social phenomena including the flow of pedestrians, the geography of crime rates, and the distribution of retailers" (Croxford, Penn, *et al.*, 1996). This Space Syntax method helps to revealed social forces that can shape urban form. Figure 6 below shows the basic analysis steps in space syntax method.

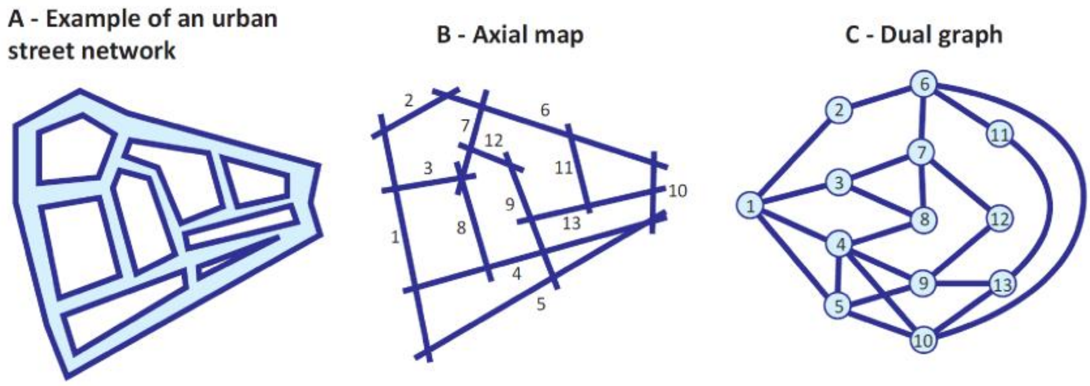

Figure 7 Space Syntax used in city layout

Figure 7 shows "axial maps of four geometrically different cities in different parts of the world from the most to least geometric: Atlanta in the USA, The Hague in Holland, Manchester in the UK and Shiraz in Iran" (Ratti, 2004).

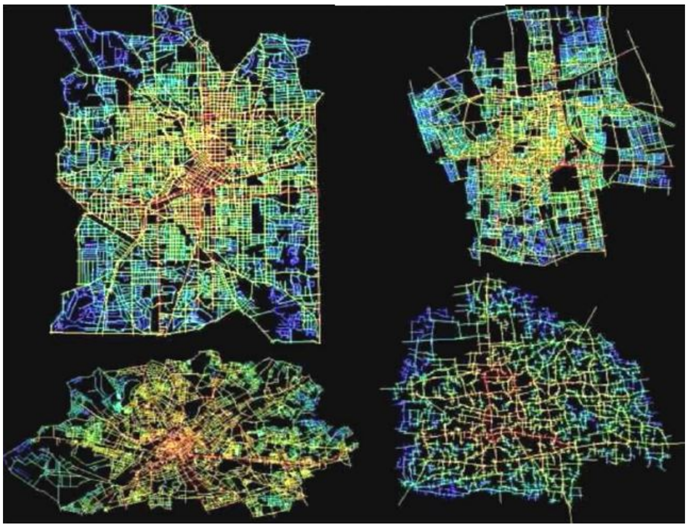

Figure 8 street network's integration values for 4 cities. "Clockwise from top left: Atlanta, The Hague, Shiraz, Manchester" (Hillier, 2005).

However, another model termed "multiple centrality assessment (MCA)" has offered an alternative and new way to define centrality. "The MCA uses a primal graph approach and computes the distance between nodes metrically along the network rather than topologically in terms of the number of turns, similar to that in space syntax" (Porta, Crucitti, et al., 2006a). With analogy to space syntax, this model uses a series of different centrality measures based on graph theory. "Three common metrics employed in MCA are betweenness, closeness, straightness centrality" (Porta, Strano, et al., 2009b).

(Porta, Crucitti, et al., 2006a) proposed that: "MCA provides a different perspective from space syntax in that: (1) it is based on primal, rather than dual, street graphs; (2) it works within a metric, rather than topological, framework; (3) it investigates a plurality of peer centrality indices rather than a single index".

The betweenness centrality of a street segment *i* is defined as "the fraction of shortest paths between pairs of vertices in a network that pass through *i*" (Freeman, 1977). First, the shortest path connection is calculated between all nodes in the graph. "Given a matrix of shortest paths between all node pairs, a particular node's betweenness index is then calculated as the number of times that the node is traversed in this set of shortest paths"(Pinho and Oliveira, 2009). The formulation is shown below.

$$Betweenness_i = \frac{1}{(\mathrm{N}-1)(\mathrm{N}-2)} \sum_{j=1;k=1;j\neq k\neq i}^{N} \frac{n_{jk}(i)}{n_{jk}}$$

N is the number of nodes, "n_{jk} is the number of shortest paths between nodes *j* and *k*, and $n_{jk}(i)$ is the number of these shortest paths that pass through the node i". "When applied on street networks, betweenness measure can be intuitively thought of as the potential amount of traffic on each street segment that results if one person were to travel from each intersection to each other intersection in the given road network along shortest paths" (Pinho and Oliveira, 2009).

"The closeness centrality of a node is defined as the inverse of distance required to reach from one node to all other nodes in the system along shortest paths" (Sabidussi, 1966). The method calculates the sum of distances at each node and normalized by the count of nodes in the system. This indicator means how far each location is from all other locations.

$$Closeness_i = \frac{\mathrm{N}-1}{\sum_{j=1;j\neq i}^{N} d_{ij}}$$

N is the number of nodes, d_{ij} is the geodesic shortest path distance between nodes *i* and *j*. This indicator is similar with the integration measure in Space Syntax, but each uses a different impedance measure. "Space Syntax uses the count of topological turns as the distance metric, and closeness centrality uses metric distance" (Porta, Strano, et al., 2009b).

The straightness metric illustrates "the extent to which the shortest paths from a node of interest to all other nodes in the system resemble straight Euclidian paths" (Porta, Strano, et al., 2009b). "The straightness metric captures the positive deviations in travel distances that result from the geometry of the road network in comparison to ideal straight-line distances in a featureless plane" (Porta, Strano, et al., 2009b).

$$straightness_i = \frac{1}{\mathrm{N}-1} \sum_{j=1;j\neq i}^{N} \frac{d_{ij}^{Euclidian}}{d_{ij}}$$

$d_{ij}^{Euclidian}$ is the straight-line Euclidian distance between node *i* and *j*, and d_{ij} is the geodesic network distance between the same nodes.

The leaves graph below shows these three measures for the same graph. Left is Betweenness, middle is Closeness, right is Straightness. Red color means high value and blue color means low value.

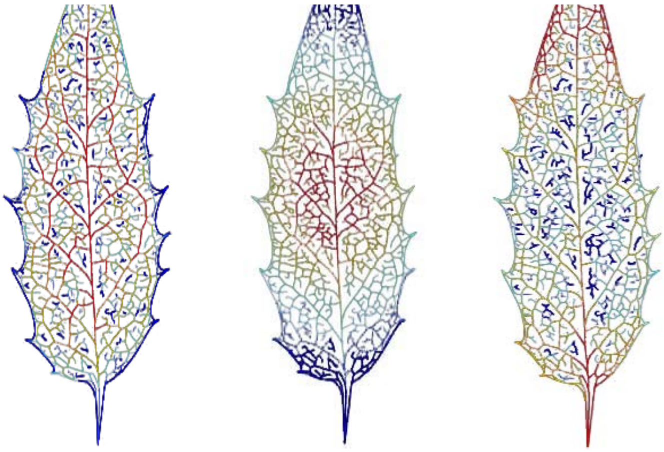

Figure 9 Betweenness, Closeness and Straightness measures (source: http://www.humanspacelab.com)

MCA model defines centrality of “a place not only as being central in terms of closeness to other places but also being the intermediary, straight and critical to others” (Porta, Strano, et al., 2009b). Therefore this model is a more comprehensive assessment for locations. In previous empirical researches, centrality has been widely explored and some similarities were found for different street patterns.

(Wang, Antipova, et al., 2011) calculated these three network centralities in Baton Rouge, Louisiana, which is shown in figure 9.

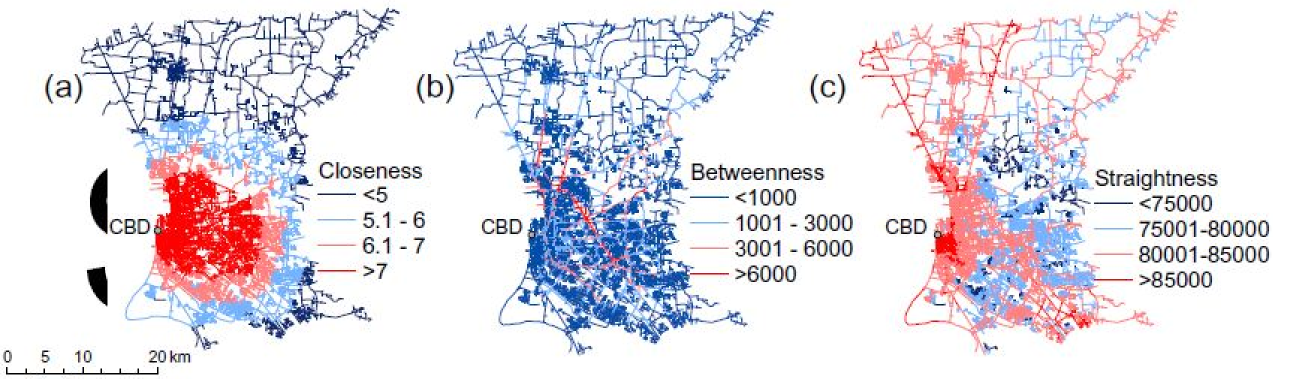

Figure 10 Spatial distributions of global (a) closeness, (b) betweenness, and (c) straightness (Wang, Antipova, et al., 2011)

This study examined the “relationship between street centrality and land use intensity which is measured by population and employment densities in census tracts”. “Two GIS-based methods are used to transform data sets of centrality and densities to one unit for correlation analysis, which is the kernel density estimation, converts both measures to raster pixels and floating catchment area method, computes average centrality values around census tracts” (Wang, Antipova, et al., 2011).Results indicate that “population and employment densities are highly correlated with street centrality values. Among the three centrality indices, closeness exhibits the highest correlation with land use densities, straightness the next and betweeness the last” (Wang, Antipova, et al., 2011).

(Porta, Latora, et al., 2012) examines “the geography of three street centrality indices and their correlations with various types of economic activities in Barcelona, Spain. The focus is on what type of street centrality (closeness, betweeness and straightness) is more closely associated with which type of economic activity (primary and secondary)” (Porta, Latora, et al., 2012). Centralities are calculated by a multiple centrality assessment model, and a kernel density estimation method is applied to both street centralities and economic activities to permit correlation analysis between them. Results indicate that “street centralities are correlated with the location of economic activities and that the correlations are higher with secondary such as retail, hotel and restaurant, than primary activities, for example, the manufacture and fishery industry. The research suggests that, in urban planning, central urban arterials should be conceived as the cores, not the borders, of neighborhoods” (Porta, Latora, et al., 2012).

(Kang, 2016) investigates “the effect of accessibility and centrality to walking volume on retail sales in Seoul by considering pedestrian volumes and street network configuration concurrently. Multilevel regression models confirm that spatial access to pedestrians has differing effects on retail sales according to the type of retail sector” (Kang, 2016). Specifically, “the sales of the Medical Services and Education sectors are remarkably sensitive to the combined effects of pedestrian and street configuration, unlike those of the three other sectors, namely Food, Retail, and Services” (Kang, 2016).

(Liu, Wei, et al., 2015) also investigated “the relationship between street centrality and land-use intensity in the main urban area of Wuhan, China”. “Land-use intensity is measured based on the building and economic activity density in different land-use types”. Also, they use “kernel density estimation to convert the measures to a basic raster unit”. This study also explained “the disparities in the global relationships, by using geographically weighted regression method which is used to explore the spatial heterogeneity” (Liu, Wei, et al., 2015). The results show there are strong relationships between street centrality and land-use intensity. Furthermore, “the relationships vary not only among the different land-use types but also in the different categories of a single land-use type. Spatial heterogeneity also exists in these relationships” (Liu, Wei, et al., 2015).

However, these researches are all limited in one type of land use but not all types of land use. Also, a few of studies figure out how this relationship varies across space. What’s more, few studies calculated centrality indicators in different search radius which means different means of transportation.

In conclusion, the first contribution of the thesis is that, correlations between street centralities and different land use types are fully investigated to study the effects of human activities on land development, especially in the polycentric area. The second contribution is to calculate street centrality metrics by different means of transportation. The third contribution is to investigate how these relationships vary across space.

2.5 Conceptual framework

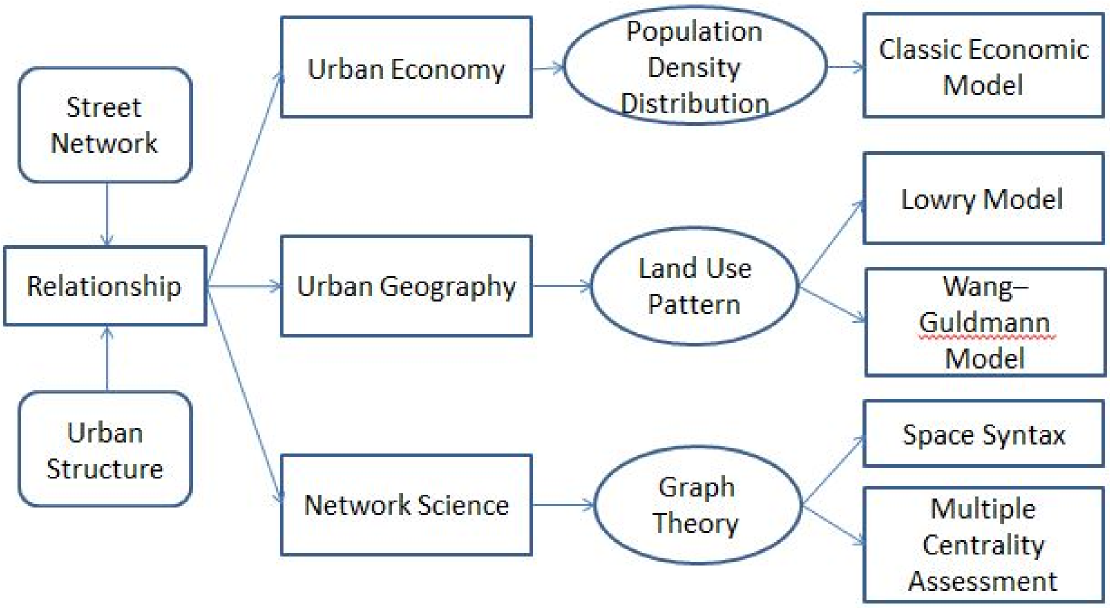

Figure 11 Conceptual Framework

The conceptual framework is shown in figure 10. In the literature review, there are several schools about the relationship between street network and urban structure. Based on the conceptual framework, this research intends to investigate street network and different land uses, by using multiple centrality assessment model.

Chapter 3: Research Design and Methods

3.1 Introduction

This chapter focuses on the approaches that were undertaken in this research, the data collection methods, the variable and indicator of this research, data analysis methods, also the validity and reliability of the study.

3.2 Revised Research Questions

The main question of this research is:
To what extant are street network correlated with different land use?
The research sub-questions are as follows:
What are the spatial properties of the street network?
The first step is to know spatial distribution of street network before the following steps. The spatial characteristics of this network can be quantified by three aspects in the perspective of network centralities: betweenness, closeness and straightness of every street. These spatial indicators can be obtained from MCA model. The network centralities will be calculated in different search radius which represents different means of transportations.
What are the spatial distributions of different land uses in the study area?
Because the main question is about the spatial relationship of street network and land use, land use distribution should be found beforehand. This sub-question will be answered by different land use maps. After obtaining these maps, we can found whether the streets configuration coordinated with land use in the study area in the next step, and if particular land use types are distributed along the most accessible streets or which indicators of the network matter most.

3.3 Research Strategy

In this research, case study research strategy was used. Case study "is a research strategy in which the researcher concentrates on one or two cases of the research subject, which are studied in their everyday setting" (Van Thiel, 2014). One important characteristic of the case study is phenomenon is studied in very great detail although it only concentrates on a few situations. As a result, "within the unique context of the case in question, the researcher can also try to arrive at an explanation of the research subject" (Van Thiel, 2014). Since the research intends to find the relationship between sub-system: land use and street network specifically in the context of Ningbo, a city of China, therefore single case study is the most suitable research strategy.

3.4 Research Methodology

Case study often tends to take a rather one-sided methodological view. In this research, different research methods such as survey and desk research should be were used. When it refers to quantitative data, research consists of three phases: data collection, data quantifying and data analysis. In the first phase, this study used secondary data got from remote sensing images and proceeds field survey to make sure data accuracy. In data quantifying process, all vector data were quantified in software ArcGIS. In the data analysis phase statistic models were applied to answer the research questions. Details of data will be clarified in the following parts.

3.5 Operationalization: Variables, Indicators

Question	Variables	Indicators	Data Source
What are the spatial characteristics of street network?	Street network centrality	• Betweeness • Closeness • Straightness	- vector street network data are downloaded from OpenStreetMap website - Urban Network Analysis Toolbox
What is the spatial distribution of different land uses in the study area?	Different land use	• Residential • Industrial • Commercial • Public services • Grass land • Forest	- vector land use maps are obtained by field survey and interpreted from remote sensing images - Remote sensing images are downloaded from USGS(https://glovis.usgs.gov/)

Figure 12 Operationalization of Variables

3.6 Research Scope and Selection

The scope of this research was focused on the metropolitan area in the city of Ningbo, which is a polycentric area. This area includes several administrative regions, as follows: Haishu and Jiangdong districts, a part of Jiangbei, Zhenhai, Beilun and Yinzhou districts. All data in this research are vector data collected from OpenStreetMap and remote sensing images, and going through editing work before analysis.

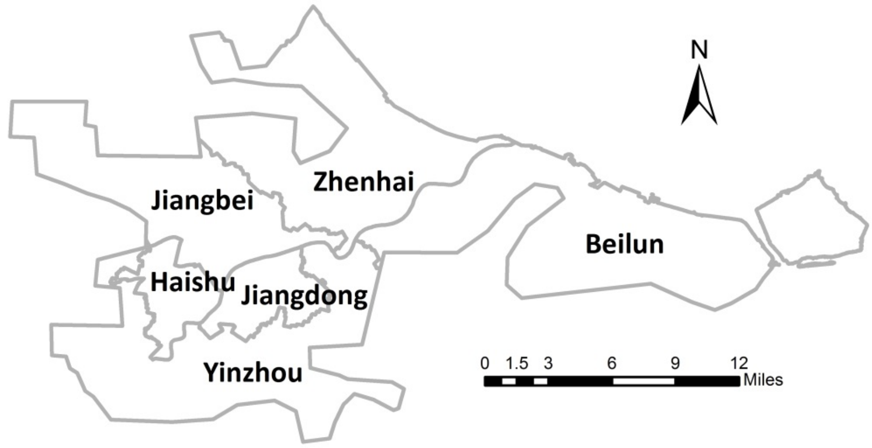

Figure 13 Location of the study area

3.7 Data Collection Methods

Street network data were export from OpenStreetMap, which is a collaborative project to create a free editable map database of the work. “Recent years have seen the widespread engagement of large numbers of private citizens in the creation of geographical or spatial data, this paradigm of user generated spatial data is commonly referred to as volunteered geographic information (VGI)” (Corcoran and Mooney, 2013). This OSM spatial database contains over 100 million way features which predominantly correspond to streets and over 500,000 contributors registered with this project. “Contributors predominately capture geographical data through the use of global positioning system (GPS) or tracing over aerial imagery” (Corcoran and Mooney, 2013). In this research, transportation layers were exported from OSM which includes different hierarchy routes such as railways routes, main road, branch road and main pedestrian. The street database of the study area consisted of edges and nodes. Editing work was implemented to ensure correctness of the topology. The resulting network contained links and nodes, with attributes for example identifiers, lengths, classes, corresponding coordinates and IDs of the two ends nodes.

There are several ways to get land use data, for example from OpenStreetMap, however, the data size of land use patch is not sufficient. So in this study, land use data was interpreted from remote sensing images: Landsat Thematic Mapper(TM) image. The Landsat Program is a series of Earth-observing satellites co-managed by USGS& NASA, and provides the longest continuous space-based record of Earth’s land in existence. The image can be downloaded from the United States Geological Survey website (https://earthexplorer.usgs.gov/). In this research, the image was captured by Landsat 8 which spatial resolution is 15-meter and it was captured on the day April 1, 2017.

“Photointerpretation is the detection, identification, description, and assessment of the significance of objects and patterns on a photograph. With respect to land use mapping, photointerpretation requires that land areas be identified and classified to a category in accordance with a land use classification scheme” (Corcoran and Mooney, 2013, Loelkes, 1983). Based on typical characteristics such as patterns, size, shapes, tones, colors and textures, three land use types were extracted including built-up area,

forest and grass land. The image was first geometrically corrected within 0.5 pixels using the quadratic method. The pre-processing also included false color composition. Then this image was imported into the ArcGIS 10.2 platform and set the photogrammetric scale at 1:50000. Images are then visually interpreted.
Inside the built-up area, field survey was implemented to obtain what the building is used for. The buildings were classified into four general categories based on "the code for classification of urban land use and planning standards of development land" [GB50137-2011]. The four general categories were as follows: residential, industrial, commercial and public services area. Public services area is the combination of administration, municipal utilities and public services area.

3.8 Data Analysis Techniques

Street centrality measure

Street network property was investigated by MCA model, "which is a primal graph approach and computes the distance between nodes metrically along the network" (Liu, Wei, et al., 2015). The previous mentioned three centrality indicators: betweeness, closeness and straightness were calculated by the "Urban Network Analysis toolbox" of ArcGIS. The formula and definition of each centrality indicator are shown below. The centrality value for each edge is calculated as the average value of its two end nodes.

Centrality	Formula	Definition
$Betweenness_i$	$\frac{1}{(N-1)(N-2)}\sum_{j=1;k=1;j\neq k\neq i}^{N}\frac{n_{jk}(i)}{n_{jk}}$	The fraction of shortest paths between pairs of vertices in a network that pass through *i*
$Closeness_i$	$\frac{N-1}{\sum_{j=1;j\neq i}^{N} d_{ij}}$	How far each location is from all other locations
$straightness_i$	$\frac{1}{N-1}\sum_{j=1;j\neq i}^{N}\frac{d_{ij}^{Euclidian}}{d_{ij}}$	The extent to which the shortest paths from a node of interest to all other nodes in the system resemble straight Euclidian paths

Figure 14 Formula and definition of centrality indicator in MCA model

N is the number of nodes, n_{jk} is the number of shortest paths between nodes *j* and *k*, and $n_{jk}(i)$ is the number of these shortest paths that pass through the node i. d_{ij} is the geodesic shortest path distance between nodes *i* and *j*. $d_{ij}^{Euclidian}$ is the straight-line Euclidian distance between node *i* and *j*, and d_{ij} is the geodesic network distance between the same nodes.

Betweeness centrality measures "how often a node is traversed by the shortest paths connecting all pairs of nodes in the network". *Straightness centrality* measures "how much the shortest paths from a node to all others deviate from the virtual straight lines connecting them". *Closeness centrality* measures "how close a node is to all the other nodes along the shortest paths of the network" (Liu, Wei, et al., 2015).

These three indices were also calculated in a buffer area. On the other hand, each indicator was calculated in different search radius: 500, 1500, 5000 and 8000 meters which were chosen by (Mashhoodi and Berghauser Pont, 2011), and represent people's walking choice, riding bike choice, driving car choice and railway choice.

Kernel Density Estimation

As explained earlier, the kernel density estimation (KDE) method is used to transform both the street centrality and land use into a new framework (i.e., a raster system) so that the relationship between them can be assessed at the same scale. Data transformations from one scale or analysis unit to another utilize spatial smoothing or spatial interpolation techniques. There are rich choices for this task (Wang, 2006). While the choice of a particular smoothing or interpolation technique should not significantly affect the outcome of this research, this research uses the KDE.

"The KDE estimates the density within a range (window) of each observation to represent the value at the center of the window. Within the window, the KDE weighs nearby objects more than far ones based on a kernel function. Among various kernel functions, popular choices include the standard Gaussian and quartic functions" (Wang, 2006). Epanechnikov (1969) finds that "the choice among the various kernel functions does not affect significantly the outcomes of the process".

The KDE generates a density of the events (discrete points) as a continuous field (e.g., raster). "By using the density (or average attributes) of nearby objects to represent the property at the middle location, the KDE captures the very essence of location: it is not the place itself but rather its surroundings that make it special and explains its setting" (Wang, 2006). Therefore, using the KDE here is not only convenient with a built-in tool available in ArcGIS but also a necessity of accurately capturing the true intention of analysing the relationship between two neighbourhood features. The choice of bandwidths was based on (Liu, Wei, et al., 2015). Considering the area of this region, two search bandwidths h were tested for every variable: 500and 1500 meters.

Pearson's correlation

"At present, the relationships between land use and street centrality are examined using traditional statistical methods Pearson's correlation analysis, which assumes that the relationships are constant across space" (Liu, Wei, et al., 2015).

In statistics, the Pearson correlation coefficient is a measure of the linear correlation between two variables. It has a value between +1 and −1, where 1 is a total positive linear correlation, 0 is no linear correlation, and −1 is a total negative linear correlation. According to (Evans, Heath, et al., 1996), 0.00-.19 means "very weak";

0.20-.39 means “weak”; 0.40-.59 means “moderate”; 0.60-.79 means “strong”; 0.80-1.0 means “very strong”.

The formula for Pearson’s correlation coefficient was shown below:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

Where Cov is the covariance; $\sigma_X$is the standard deviation of X; $\sigma_Y$is the standard deviation of Y.

Geographically Weighted Regression

To explore how this relationship varies across the space, GWR model was used in this research. This model can capture the spatial variations of relationship across space.

The formula of GWR model is introduced below:

$$y_i = \beta_0(\mu_i, v_i) + \sum_k \beta_k(\mu_i, v_i)x_{ik} + \epsilon_i$$

“(μ_i, v_i) means spatial position of location i; $\beta_0(\mu_i, v_i)$means model intercept; $\beta_k(\mu_i, v_i)$ means slope coefficient of the *k*th independent variable in location i”. The coefficients can be calculated by the equation below:

$$\hat{\beta}(\mu_i, v_i) = [X^T W(\mu_i, v_i)X]^{-1} X^T W(\mu_i, v_i)y$$

“$\hat{\beta}(\mu_i, v_i)$ means unbiased estimate of β; $W(\mu_i, v_i)$means weighting matrix which ensures that observations near the specific location have larger weight” (Liu, Wei, et al., 2015).

GWR is also an ArcGIS based analysis tool, to meet the requirement of this model, 2000 points randomly selected and the corresponding value of these variable were used for the GWR model.

3.9 Validity and Reliability

In order to ensure the quality of research, reliability and validity are critical parts of the research process and structure. In order for quality and credibility of research, accuracy assessment refers to field survey was implemented for both street network data and land use data.

Accuracy assessment means the comparison of a classification with ground truth data to evaluate how well the classification represents the real world. There are several sampling methods, such as simple random sampling, system sampling, stratified random sampling and systematic non-aligned sampling. Stratified random sampling which means a minimum number of observations are randomly placed in category was used in this research. In this way, accuracy is determined by 100 points collected on field trips, and made comparison with land use data in the same location. Then the results were summarized using an error matrix. There are several ways to quantify accuracies such as total accuracy means, user's accuracy and producer's accuracy. In this research, Kappa statistic was applied which reflects the difference between actual agreement and the agreement expected by chance. The equation of this method is shown below: observed accuracy determined by diagonal in error matrix, chance agreement incorporates off-diagonal sum of product of row and column totals for each class.

$$\widehat{K} = \frac{obseved\ accuracy - chance\ agreement}{1 - chance\ agreement}$$

In the research methods, two regression models were applied in this research to see the relationship from the global and local scale. These data assessment and analysis methods create reliability.

3.10 Time Schedule

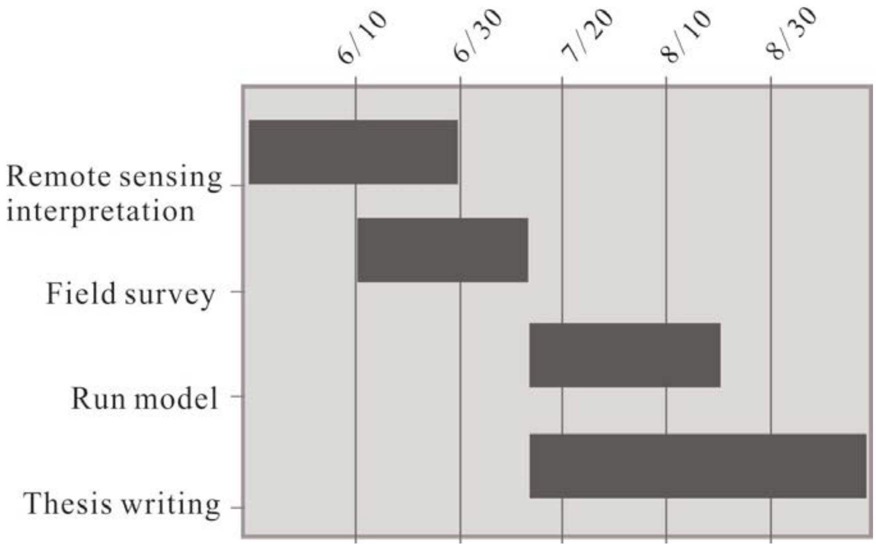

Figure 15 Thesis Research Timeline

Chapter 4 research results

4.1 Introduction

This chapter will show the results of research. The first section is about the spatial analysis of street network by the perspective of network centrality. What is the betweeness, closeness and straightness value of each street, and what is the spatial distributed pattern of these values.

The second section is to show the land use distribution of Ningbo. Interpolation maps of different land use types will be shown. From these maps, we can see what is the city functional layout and urban land use structure.

The next part moves towards a comparative study of the relationships between land use and street network. There are two regression analysis methods: Pearson's regression and geographically weighted regression.

4.2 Street Network Centrality

"The MCA model uses a primal graph approach and computes the distance between nodes metrically along the network rather than topologically in terms of the number of turns"(Porta, Strano, et al., 2009a). In this research, three major indices of street centrality were chosen.

As is shown in figure 15, local centrality indicators can be calculated in different search radius. In this research, four values for the network distance were set as follows: 500, 1500, 5000 and 8000 meters, which represent people's walking choice, biking choice, driving choice and the railway choice.

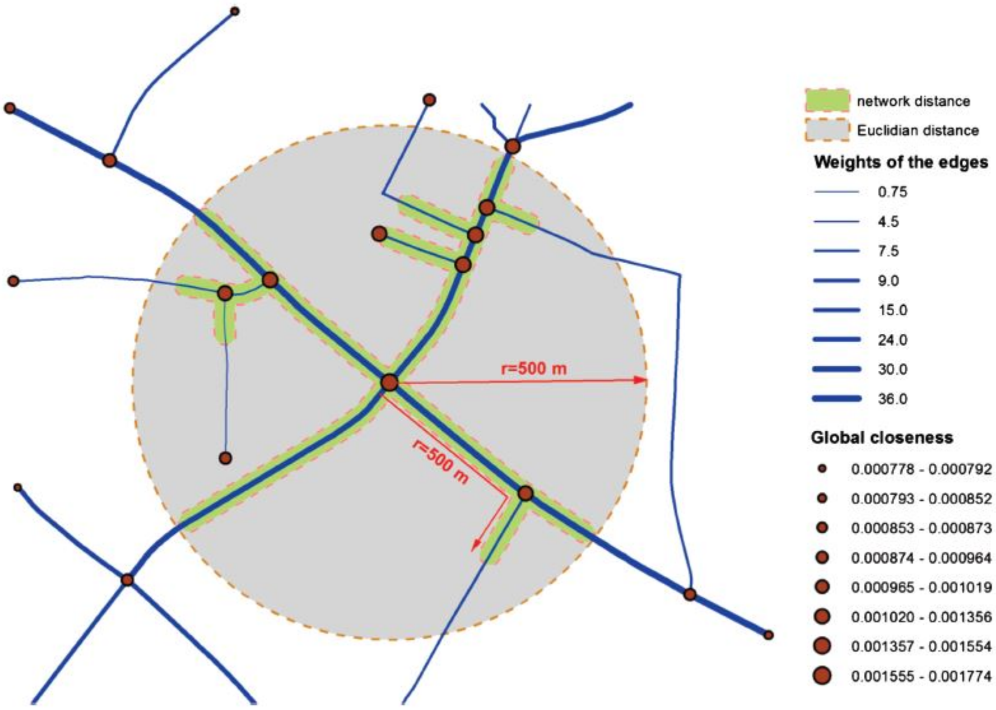

Figure 16 Illustration of network distance, Euclidian distance (Liu, Wei, et al., 2015)

To avoid boundary effect in this model, study area in this step was enlarged by 20 kilo-meters as shown in figure 16. The choice of buffer distance was based on (Mashhoodi and Berghauser Pont, 2011). That means when applied this model, streets which in the outside of study area but within 20 kilo-meters far from the area were included. Inside this region, the network has 19164 nodes and 28811 edges with attributes such as lengths, updated edges IDs, two end nodes' IDs and their corresponding coordinates.

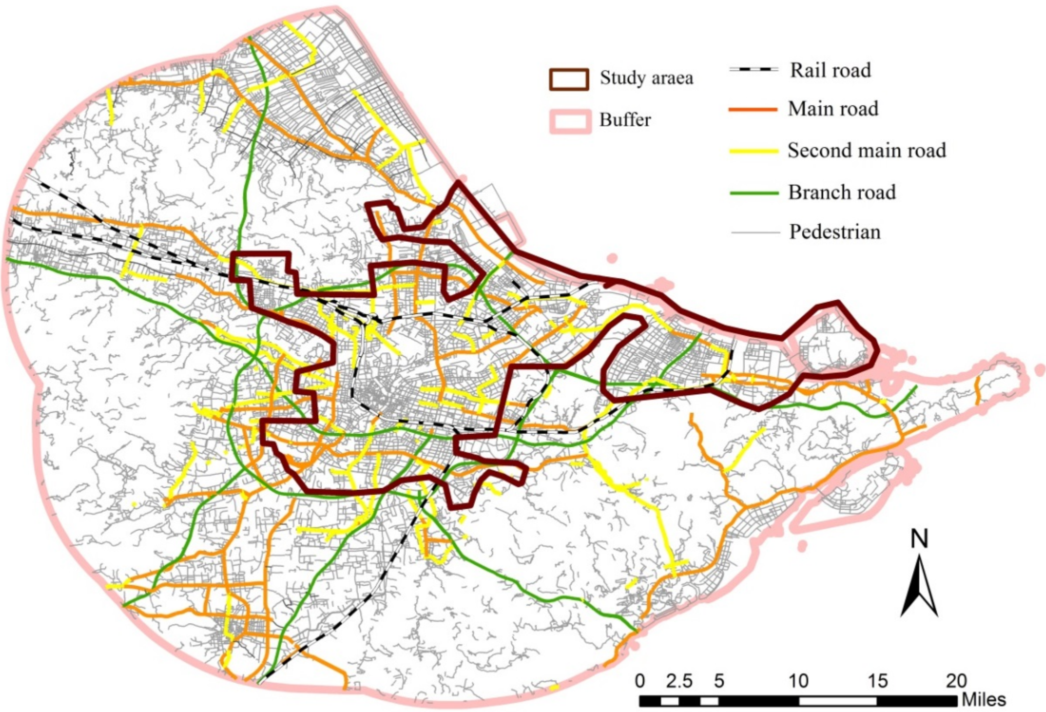

Figure 17 Street network calculated in MCA model

4.2.1 Betweeness

"Betweeness centrality is based on the idea that a node is central if it is located between many other nodes, in the sense that the node is traversed by the shortest paths between pairs of other nodes in the network"(Liu, Wei, et al., 2015).

Betweeness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. Figure 17 displays several clusters of high values and shows sub-centers.

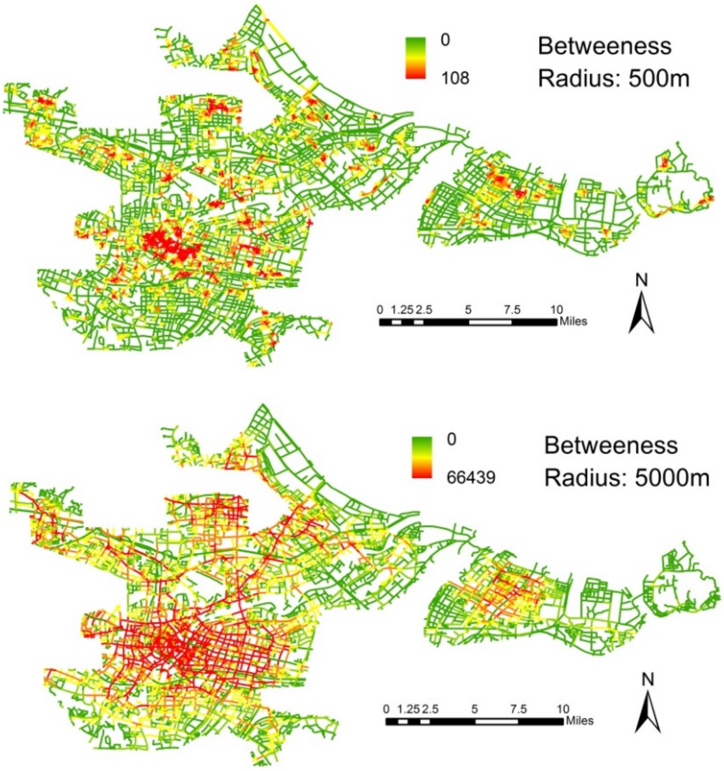

Figure 18 Betweeness value in various search radius

For 500 meters radius, high betweeness values are aggregated in the center of the study area and scatter distributed in other places. The range of betweeness is from 0 to 108. In the map of radius 1500 meters, more streets have high betweeness value and aggregated not only in the center area, but also in many other places, especially in Beilun district. Within the radius of 5000 and 8000, more roads were selected, some of them connect several districts, therefore the betweeness values were higher than others. From figure 17, it can be concluded that with the increase of search radius, more streets with high betweeness value can be selected.

4.2.2 Closeness

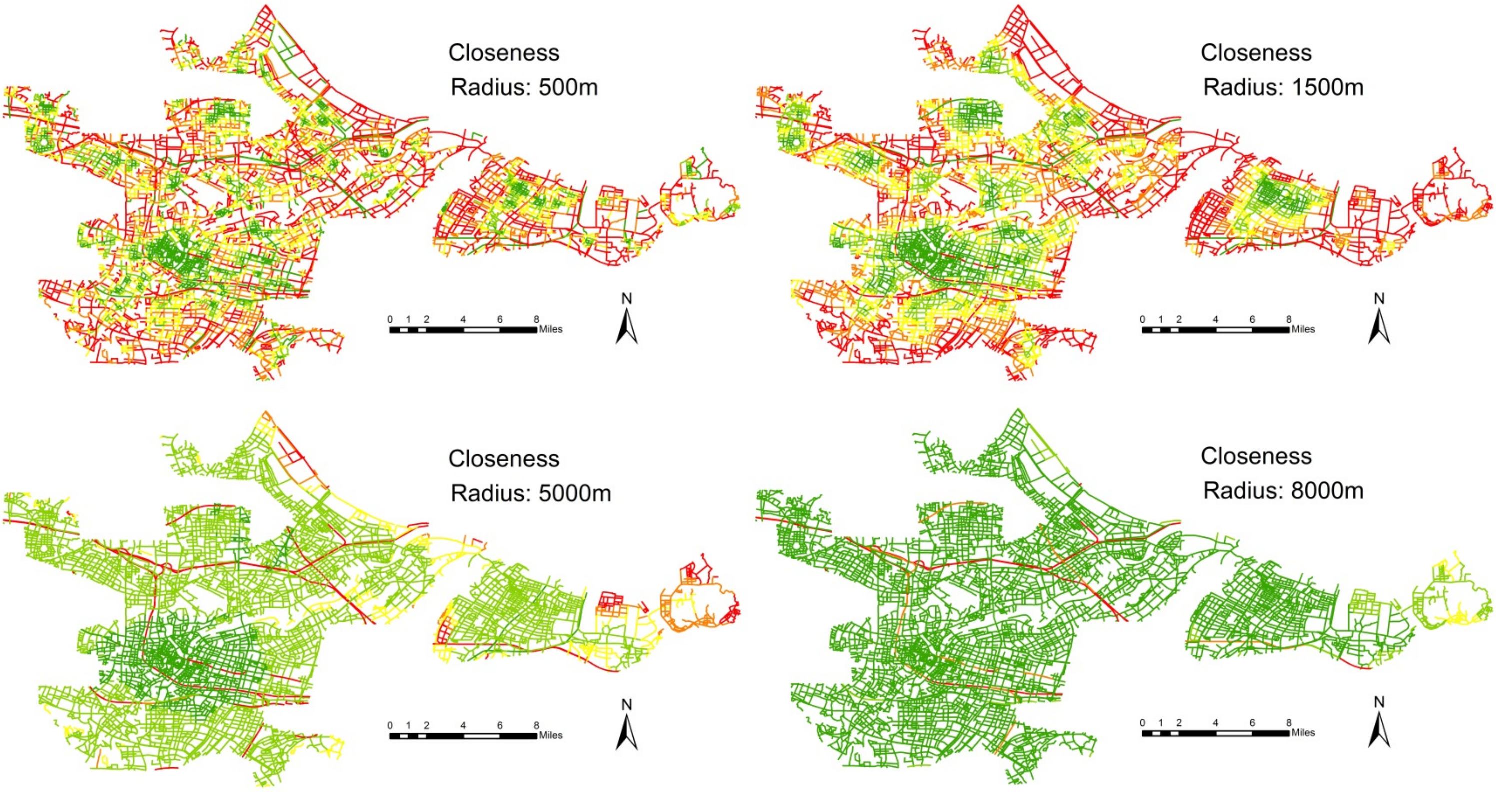

Figure 19 Closeness value in various search radius

“The closeness centrality of a node is defined as the inverse of distance required to reach from one node to all other nodes in the system along shortest paths” (Sabidussi, 1966). It measures how long it takes to spread information from a node to all other nodes. Within the radius of 500 and 1500 meters, places with a high value of betweeness always have low closeness value. That means the “bridge” roads also have the advantage to reach other nodes in a short distance. It also shows that high value of closeness exists in the outlier of this area, which means it will take longer distance from these roads to the others. However, within the radius of 5000, there are only a few roads with high closeness value. That means with the increase of search radius, the number of low closeness roads also increase. It will be easier to reach the far area in high-speed vehicle.

4.2.3 Straightness

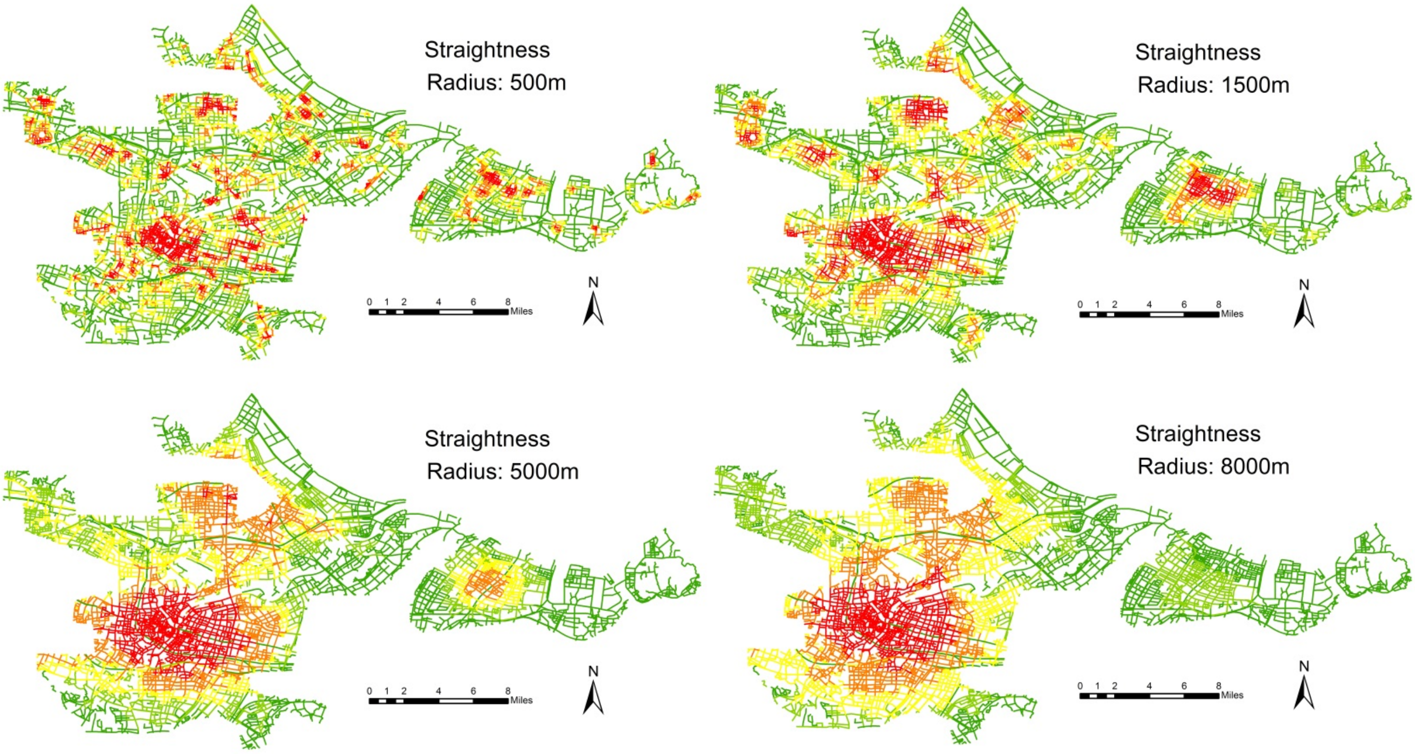

Figure 20 Straightness value in various search radius

The straightness metric illustrates “the extent to which the shortest paths from a node of interest to all other nodes in the system resemble straight Euclidian paths” (Porta, Strano, et al., 2009b). “The straightness metric captures the positive deviations in travel distances that result from the geometry of the road network in comparison to ideal straight-line distances in a featureless plane” (Porta, Strano, et al., 2009b). Within the radius of 5000 and 8000meters, low straightness streets distributed in the outlier of the study area. That means when departing from these outliers to other places, average straight line distance is much shorter than the geometry distance.

4.3 Density distributions of centralities and land use

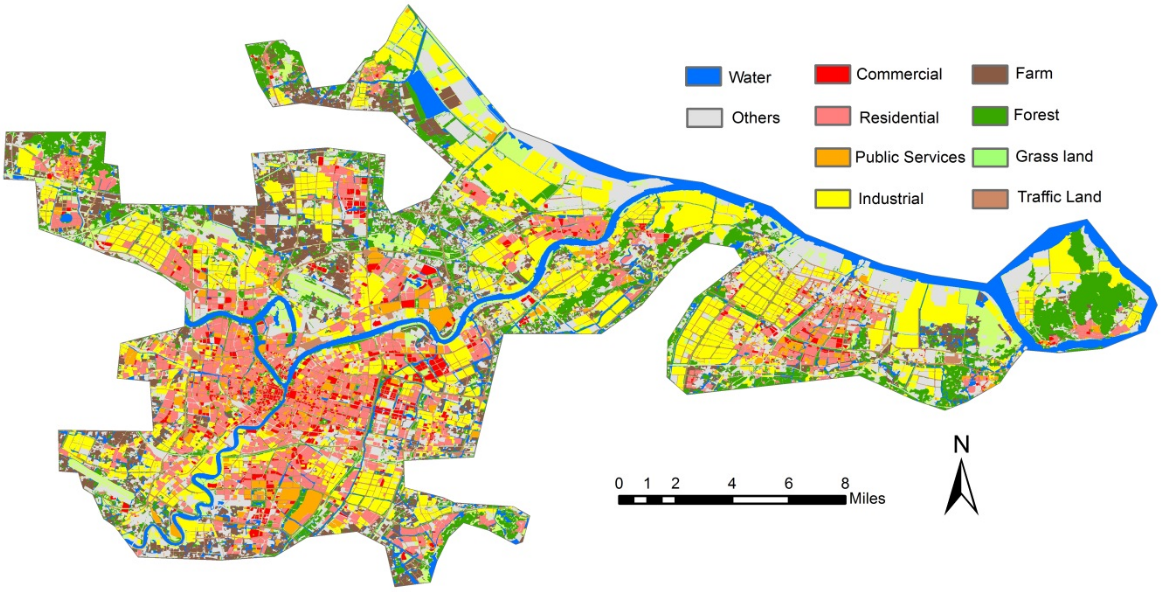

Figure 21 Major land use types in the study area

This land cover dataset in Ningbo comprises several land use classes. What is interesting to study are these seven categories: commercial, residential, public services and industrial area, also the farm, forest and grass land. Commercial land includes retails, restaurants, shops, markets, hotels, bars, theaters and cinemas. Public services area includes administrations, libraries, museums, schools, gyms, post offices and hospitals. The Kappa statistic for all types of land uses data accuracy is all above 95%. The land uses map was shown in figure 20.

A spatial smoothing tool should be used to get the spatial patterns more evidently before measuring the relationship between land use and street centralities. Considering this interpolation method should not influence the outcome of the research significantly, kernel density estimation was chosen. Then this method was applied to both land use and street network centrality data and created a so-called density map which was raster data and captured the spatial distribution of the original data. This method transformed both polygon and polyline data to the same data framework, therefore, they could be compared in next step.

"Normally, to perform KDE for polygon land use, centroids are created for polygons. Each centroid has a value representing the polygon area as the weight in KDE process" (Wang, 2016). This method ignores the shapes of the polygon, so I transformed the land-use polygon data and street network polyline data with a fine resolution.

It is important to select an appropriate bandwidth in the interpolation method to get a fine resolution. Two search bandwidths or radii *h*: 500m, 1500m were tested for every variable of interest. Every density maps are resampled to the same resolution as 100m×100m, and all data were overlaid in one framework, for example, each cell in the framework has the attributes of all types of land use and all the three streets centrality values. The choice of a bandwidth is based on (Brunsdon, 1995). In the following parts, I will first discuss the density distribution maps. Then the correlation between these two datasets will be investigated based on these density maps.

4.3.1 Density distributions of land use

Figure 21 shows the KDE results of different land uses: commercial, residential, public services, industrial area, grass land and forest. We can see from density maps, the high value of commercial area mainly distributed in the city center, as well as the residential area. Also, there are high-value pixels gathered in the edge of Jiangbei district. In another city center, there is also commercial and residential land distributed. Industrial land patches distributed along the study area but aggregated in the Beilun district which is another center of this study region. And there are also some industrial lands in the margin of the city, but a few in the city center. Public services land patches aggregated in both centers and also in Jiangbei and Zhenhai region. We can see distribution patterns of commercial, residential and public services area are alike from the density map, and public services area distributed more widely. Grass lands are evenly distributed in the area, as well as the forest lands. But forest land patches mostly distributed in Zhenhai district and also the peripheral region of the study area. The spatial smoothing effects of KDE help sketch the spatial patterns clearly so that we can know the spatial distributions of different land uses.

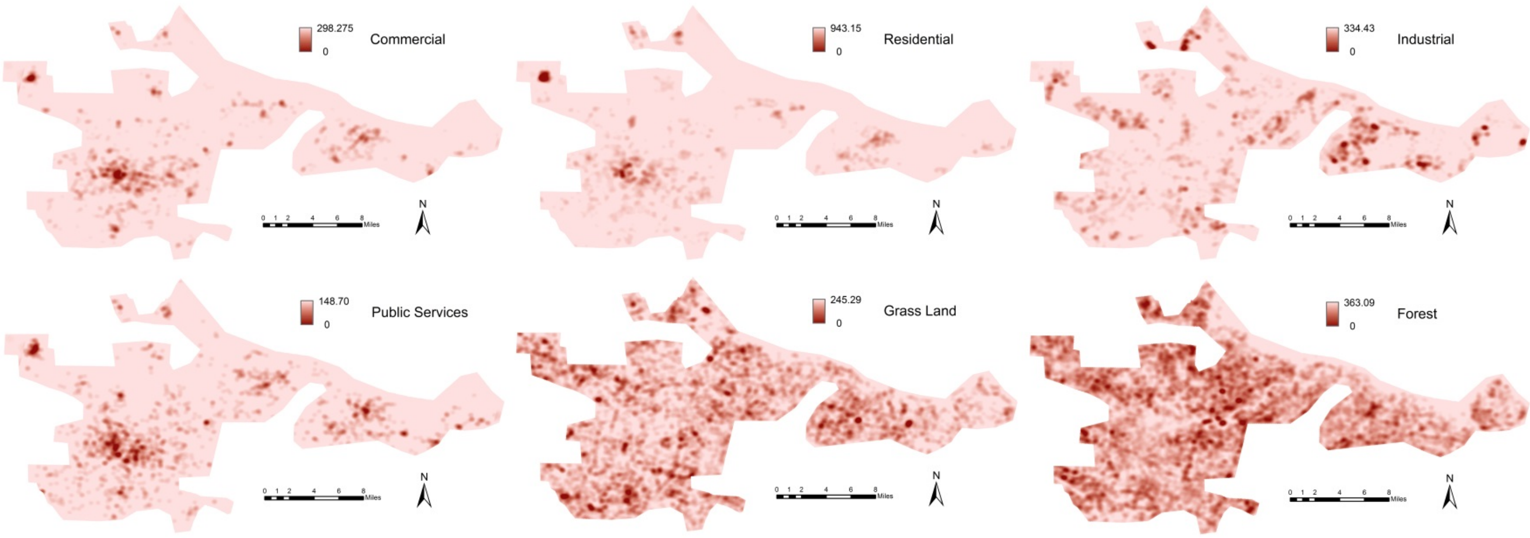

Figure 22 Density Maps of the Land uses (h=500)

Figure 22 shows the statistical distribution of graphs in figure21. The horizontal axis is the pixel values of density map and the vertical axis represents the relative frequency of the cells' numbers. Both axes are on a linear scale. From figure 8 we can see that the latter parts of all graphs are well-estimated by a decreasing exponential curve. Therefore, the lower-density areas are numerous, but the higher-density areas are few in number. The commercial, residential, public services and industrial land shows more unchanging patterns, with rapidly declining tails. The distributions of forest and grass land are characterized by a single peak. We can see that, for previous land uses, density value are aggregated in small scopes, but there are also lots of high values ranged in large scopes. However, for the grass and forest land, density values aggregated in the front part of the plots. These findings are similar to the results of the previous study (Liu, Wei, et al., 2015).

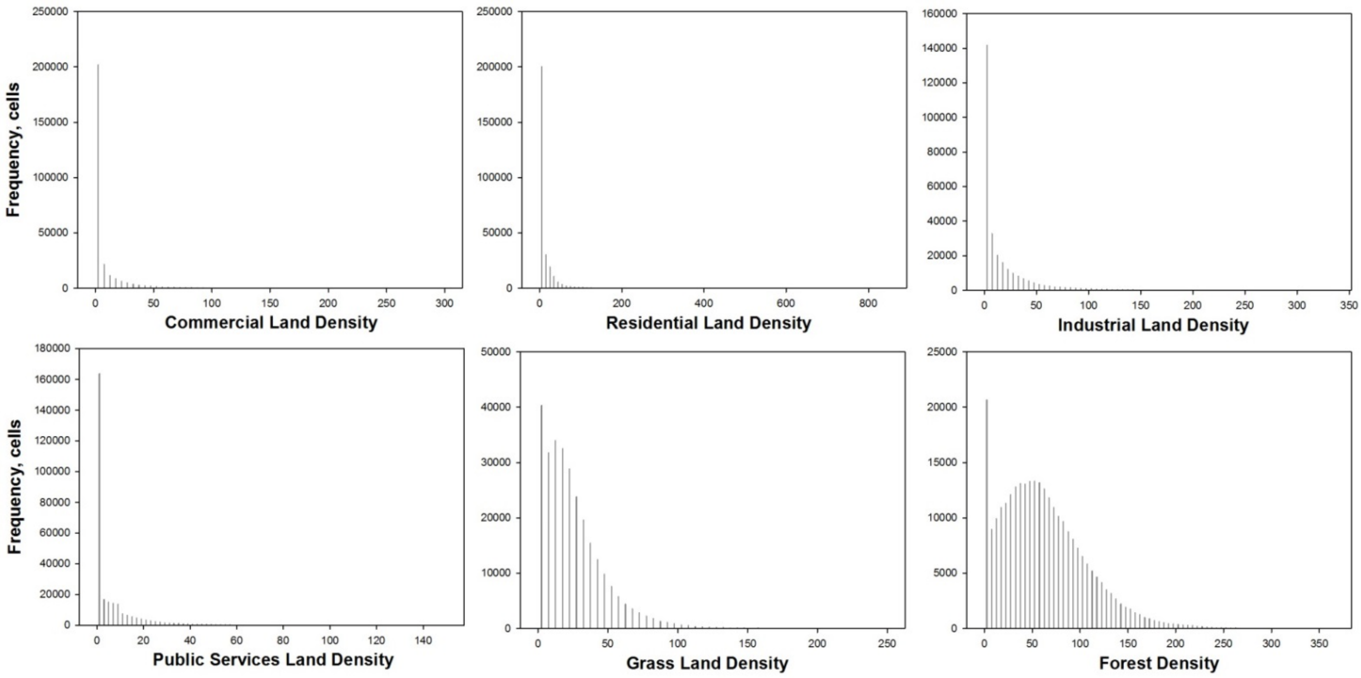

Figure 23 Statistical distribution of land uses density maps

4.3.2 Density distributions of street network centrality

Figure 23 shows the density distributions of street network centrality indictors in different scales. The bandwidth h=500m, the size of each cell are 100m×100m. This method can capture the spatial distribution effectively.

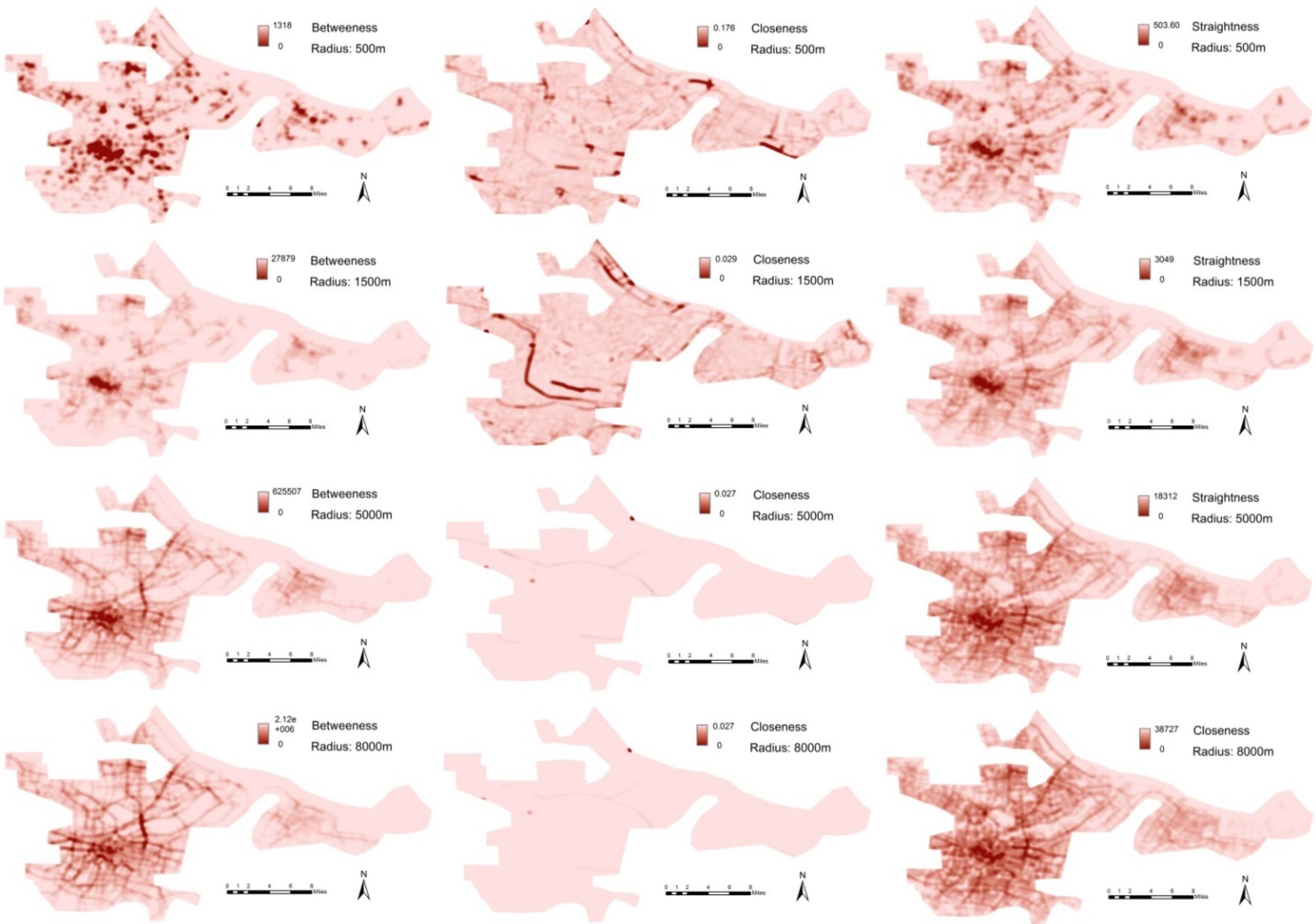

Figure 24 Density of the street centrality measures (h=500)

We can see from figure 23 that, all streets with a high value of centrality were highlighted and we can know exactly where the high value aggregated. Figure 24 is the statistical distribution of graph in figure 23.

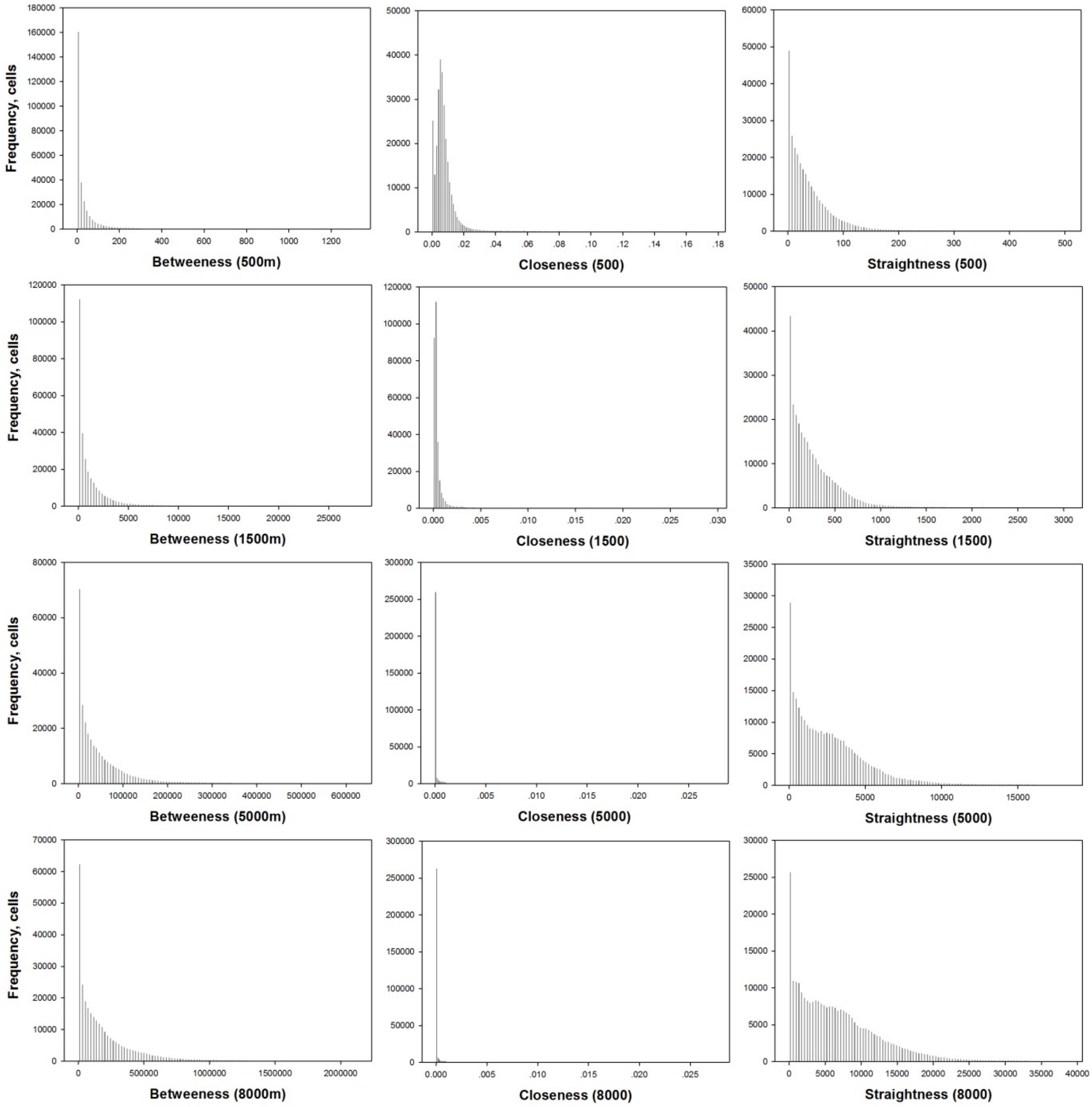

Figure 25 Statistical distribution of the street centralities (h=500)

In general, the high centrality values are concentrated in a few places, so the best accessibility is only enjoyed by a small portion of the whole population in the city. Specifically, density distributions show different patterns for different centrality indices. Betweeness density decreases quickly and exhibits an exponential distribution. And we can also see that betweeness density decrease more slowly when increasing the search radius. Closeness density shows with more rapidly decreasing tails. Also, closeness density in 500 meters radius has two peaks. Compared with closeness and betweeness density, the density distributions of stratightness are much flatter, also with several peaks within the search radius of 5000 and 8000 meters.

4.4 Correlation analysis

Correlation analysis is used to calculate the statistical relationship between different centralities and land use types on a cell-by-cell raster in the same framework. When the Pearson's correlation was calculated, to ensure the most precise results, cells with zero value in each of the raster layers were excluded. Then data normality was checked and make sure all of them are positively skewed. All of the correlation coefficients reported in following parts were significant at the 0.01 level.

In the first part, coefficient under the bandwidth 500 meter will be showed. The second part will be the comparison of the results under different bandwidth, so to understand the impact brought by interpolation methods in the correlation process.

Before analysis, there is a brief explanation of the coefficients of Pearson correlation. In statistics, the coefficient has a value between +1 and −1, where 1 is a total positive linear correlation, 0 is no linear correlation, and −1 is a total negative linear correlation. According to (Evans, Heath, et al., 1996), 0.00-.19 means "very weak"; 0.20-.39 means "weak"; 0.40-.59 means "moderate"; 0.60-.79 means "strong"; 0.80-1.0 means "very strong".

4.4.1 Correlation analysis under 500 meters bandwidth

All kinds of land uses have positive relationships with different centralities. Commercial land is the highest positively correlated area and is most correlated with betweeness within search radius 500 meters, which means commercial land intend to distributed along the most connected streets under the walking distance. Meanwhile, commercial land is also correlated to betweeness within biking, driving distance.

When it refers to public services land, the correlation coefficients are lower than commercial land and the highest correlation appeared in 1500 meters, which means public services land is most correlated with street betweeness within the biking distance. People are easy to get to the public services area when they riding bikes.

For the residential land, there are no strong correlations with street network betweeness. The coefficients are ranged from 0.3 to 0.4 and that means they are correlated in a medium level. Among different search radius, residential lands are most correlated with betweeness within 1500 meters.

Forest, industrial and grass land are weakly correlated with street network betweeness, which means the locations of these land doesn't have a strong relationship with street betweeness, they are not distributed in the most betweeness streets. The coefficients of these relationships are ranged from 0.02 to 0.2.

Table 1 Pearson's Correlations between Land use and Street Betweeness (Bandwidth:500m)

Betweeness-Land use	500m	1500m	5000m	8000m
Commercial Land	0.61056	0.5683	0.5379	0.4232
Public Services Land	0.46739	0.52954	0.46728	0.36684
Residential Land	0.4096	0.46678	0.38154	0.30028
Forest	0.10309	0.13994	0.18673	0.20724
Industrial Land	0.06625	0.08737	0.04734	0.02369

Grass Land	0.04613	0.07244	0.0919	0.08902

Table 2 shows the correlation between closeness indicator and different land uses. In general, the absolute values of coefficients are low and many of them are minus. Residential lands have negative relationship with street network closeness, but the coefficients are low value and there are little correlations between residential land and closeness centrality.

There are positive relationships between grass land and closeness centrality. However, the values are so low and there are little correlations between them. Commercial land and public services land have similar relationships with the street network, in the search radius of 500, 5000 and 8000 meters, the coefficients are negative, but in the search radius of 1500, there are positive coefficients. Forest and industrial land also have similar relationships with street network. The coefficients are negative within the search radius 1500, 5000 and 8000 radius, but positive within 500 meters. However the coefficients among these relationships with closeness are all in low values, so there are no high degrees of correlation in this situation.

Table 2 Pearson's Correlations between Land use and Street Closeness (Bandwidth:500m)

Closeness-Land use	500m	1500m	5000m	8000m
Commercial Land	-0.04705	0.01511	-0.0163	-0.01793
Public Services Land	-0.02626	0.03227	-0.00947	-0.01245
Residential Land	-0.05617	-0.00826	-0.01721	-0.01927
Forest	0.10843	-0.03018	-0.00664	-0.00887
Industrial Land	0.05174	-0.00027	-0.02463	-0.02773
Grass Land	0.08579	0.05318	0.05794	0.03671

Table 3 shows the relationships different land uses and different straightness. In general, all coefficients are positive under this situation. Commercial and public services land have similar relationships with straightness centrality. Commercial lands are most correlated with straightness within the radius of 1500 meters. That means when people riding bikes, commercial lands are mostly distributed along the most straightness streets which are similar with the straight-lines. Public services lands also have similar distributed pattern however all the coefficients are lower than commercial lands.

The range of coefficients for the residential land is from 0.35 to 0.43, for forest it is ranged from 0.22 to 0.27. Industrial land and grass land have little correlations with straightness centrality because the coefficients are around 0.1.

Table 3 Pearson's Correlations between Land use and Street Straightness (Bandwidth: 500m)

Straightness-Land use	500m	1500m	5000m	8000m
Commercial Land	0.55375	0.58963	0.51612	0.44836
Public Services Land	0.48562	0.53164	0.47028	0.40503
Residential Land	0.43488	0.4655	0.39725	0.35229
Forest	0.2213	0.23106	0.25291	0.27842
Industrial Land	0.15808	0.13235	0.08029	0.0698
Grass Land	0.10153	0.11497	0.11244	0.12176

Table 4 shows all coefficients which are higher than 0.3 in the case of 500 meters bandwidth. Two important findings are shown in this table. First, among all these three land uses, commercial lands are most correlated with both betweeness and straightness centrality. And public services lands are more correlated with these two centrality indicators than residential lands.

Second, for commercial land, a larger search radius leads to lower coefficients, it means commercial land distributions are more correlated with people's walking choices in the street network. However, for public services and residential lands, coefficients in radius 1500 meters are the highest value, which means public services and residential area are more correlated with people's biking choices in the street network than other ways of transportation.

Table 4 Pearson's Correlations between Land use and Street Centrality (>0.3)

Radius	Centrality	Commercial	Public Services	Residential
500m	Betweeness	0.61056	0.46739	0.4096
	Straightness	0.55375	0.48562	0.43488
1500m	Betweeness	0.5683	0.52954	0.46678
	Straightness	0.58963	0.53164	0.4655
5000m	Betweeness	0.5379	0.46728	0.38154
	Straightness	0.51612	0.47028	0.39725
8000m	Betweeness	0.4232	0.36684	0.30028
	Straightness	0.44836	0.40503	0.35229

4.4.2 Correlation analysis under 1500 meters bandwidth

Table 5 Pearson's Correlations between commercial, public services, residential and Street Centralities (Bandwidth: 1500m)

	Radius	Commercial	Public Services	Residential
Betweeness	500	0.79054	0.71424	0.63865
	1500	0.80522	0.73643	0.65848
	5000	0.75659	0.69606	0.60972
	8000	0.6393	0.59201	0.5255
Closeness	500	-0.06961	-0.04438	-0.08916
	1500	0.12282	0.14396	0.05523
	5000	-0.01787	-0.01578	0.00227
	8000	-0.03729	-0.03308	-0.03293
Straightness	500	0.74268	0.69178	0.62673
	1500	0.77316	0.72204	0.64554
	5000	0.69686	0.65449	0.57698
	8000	0.60444	0.56381	0.50887

Table 5 shows Pearson's correlations between three kinds of land uses, commercial, public services, residential and street centralities indicators within the bandwidth 1500 meters. We can see from that, the coefficients for commercial and betweeness centralities are ranged from 0.639 to 0.8, which means there are highly correlated relationships between these two indicators. For the public services land, the coefficients are less than commercial land which ranged from 0.59 to 0.736, but it also means highly correlation between public services land and betweeness centralities. For the residential land, the coefficients are ranged from 0.52 to 0.65. Among the three kinds of land uses, commercial exhibits the highest correlation with betweeness under 1500 meters bandwidth.

For the straightness indicators, the coefficients for each land use are similar but a bit less than coefficients of betweeness indicators. We can say that commercial lands are most correlated with straightness and public services, residential land are highly related with straightness. However, there are low degree of correlations between closeness and all these three land uses.

Table 6 Pearson's Correlations between forest, industrial land, grass land and Street Centralities (Bandwidth: 1500m)

	Radius	Forest	Industrial Land	Grass Land

Betweeness	500	0.15946	0.14912	0.10704
	1500	0.20075	-0.0154	0.13684
	5000	0.25874	-0.02002	0.13629
	8000	0.29683	-0.04165	0.12447
Closeness	500	0.13722	0.12794	0.17284
	1500	-0.14405	0.10171	0.026
	5000	0.05174	0.03053	0.09979
	8000	0.03948	0.00381	0.05853
Straightness	500	0.32742	0.0158	0.21541
	1500	0.32742	0.0422	0.21152
	5000	0.34825	0.0422	0.19358
	8000	0.38394	0.0422	0.20781

We can see from table 6, the relationships between the other three types of land uses: forest, industrial land, grass land, and street centralities within 1500 meters bandwidth. There are low coefficients which represent almost no correlation of spatial distribution between street network centralities and the location of forest, industrial land, and grass land.

Table 7 Comparison between bandwidth 500 and 1500 meters

Bandwidth(m)	Centrality	Radius	Commercial	Public Services	Residential
500	Betweeness	500	0.61056	0.46739	0.4096
		1500	0.5683	0.52954	0.46678
		5000	0.5379	0.46728	0.38154
		8000	0.4232	0.36684	0.30028
	Straightness	500	0.55375	0.48562	0.43488
		1500	0.58963	0.53164	0.4655
		5000	0.51612	0.47028	0.39725
		8000	0.44836	0.40503	0.35229
1500	Betweeness	500	0.79054	0.71424	0.63865
		1500	0.80522	0.73643	0.65848
		5000	0.75659	0.69606	0.60972

	8000	0.6393	0.59201	0.5255
Straightness	500	0.74268	0.69178	0.62673
	1500	0.77316	0.72204	0.64554
	5000	0.69686	0.65449	0.57698
	8000	0.60444	0.56381	0.50887

Several important findings are shown in table 7, as follows:

First, a larger bandwidth leads to a stronger smoothing effect and the correlation coefficients are therefore higher. We can see that every coefficient within bandwidth 500 meters is lower than the corresponding coefficient within 1500 meters bandwidth. Also, the coefficients of public services and residential land are obviously increased with a wider bandwidth. It shows that smoothing technology has some effect on the correlation results. Wider bandwidth will lead to higher correlation.

Second, for commercial land, betweeness correlated relatively higher compared with straightness in both bandwidths, which means the distribution of commercial land highly correlates with betweeness of the street network. Betweeness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. It can be concluded that commercial activities are tended to distribute along the "bridge" streets which connected other streets better. It is easier for people get reach of these streets which benefit commercial activities such as retails, hotels and restaurants.

Third, with a larger search band, the result of betweeness and straightness declines in both bandwidths. However, the highest correlation exists in 1500 meter search radius for each group of centrality. That means the distribution of commercial, public services and residential land are more correlated with network biking centrality then other means of transportation.

In the next part, these three types of land uses will be merged together, and the correlation between street network centrality and all these three land uses will be analyzed.

4.4.3 Correlation analysis for three land uses

In this part, commercial land, public services land and residential land are merged together to see the relationship between the whole area (all these three land uses) and centrality in spatial distribution. And we will see if the correlations between the merged area and centralities have some differences with the correlation of a single land use.

From the table 8, several findings can be found. First, betweenss and straightness centrality are correlated with the merged area (these three land uses), the coefficients are ranged from 0.4 to 0.6. Second, there is no big difference between the bandwidth 500 and 1500 meters. Also, the difference is smaller than the same comparison of each land use in table 7. Third, the coefficients in table 8 are lower than the respective results in table 7, which means when considering this three land uses together, the correlation with the street network centrality will not increase, and there are no big difference with the correlation for a single land use.

Table 8 Correlation between the merged area (three land uses) and centrality

Centrality	Radius (m)	500m	1500m
Betweeness	500	0.539	0.60774
	1500	0.60134	0.61743
	5000	0.51415	0.55143
	8000	0.40413	0.46558
Closeness	500	-0.05235	-0.06436
	1500	0.00768	0.06529
	5000	-0.01768	-0.01115
	8000	-0.01993	-0.03124
Straightness	500	0.55424	0.58933
	1500	0.59597	0.59879
	5000	0.51833	0.52143
	8000	0.45373	0.45723

In conclusion, we can see from table 9, the highest correlations value among land uses and street network centralities.

Table 9 Highest correlation between land uses and centrality

Bandwidth(m)	Centrality	Radius	Commercial	Public Services	Residential
1500	Betweeness	1500	0.80522	0.73643	0.65848
	Straightness	1500	0.77316	0.72204	0.64554

4.5 Spatial Heterogeneity

"Unlike the statistical relationships between street centrality and land-use intensity, the GWR model shows these relationships are not constant across the space and explore spatial heterogeneity in such relationships"(Liu, Wei, et al., 2015). Spatial heterogeneity refers to the uneven distribution of a trait, event, or relationship across a region (Anselin and Getis, 2010). Spatial heterogeneity is also sometimes referred to as sub-regional variation. In previous parts, Pearson's correlation has been analyzed to find the "average relationships" of the whole area, which can be seen as a global model, and in this part, local relationships will be analyzed.

“The GWR can help examine the spatial pattern of the local estimates to obtain an understanding of the hidden possible causes of the spatial pattern”(Liu, Wei, et al., 2015). In this research, “given that multicollinearity exists when street centrality acts as an explanatory variable and high positive correlations exist between street centrality and land uses” (Liu, Wei, et al., 2015), only one street centrality indicator was used in each GWR model to analyze the correlation of street centrality indicator and a single indicator of land use.

The formula of GWR model is introduced below:

$$y_i = \beta_0(\mu_i, v_i) + \sum_k \beta_k(\mu_i, v_i)x_{ik} + \epsilon_i$$

The major difference between GWR and other regression model is, GWR includes the location as a variable. “(μ_i, v_i) means spatial position of location i; $\beta_0(\mu_i, v_i)$means model intercept; $\beta_k(\mu_i, v_i)$ means slope coefficient of the *k*th independent variable in location i”. So based on this function, the coefficients can be calculated by the equation below:

$$\hat{\beta}(\mu_i, v_i) = [X^T W(\mu_i, v_i)X]^{-1} X^T W(\mu_i, v_i)y$$

$W(\mu_i, v_i)$ means space weighting matrix which “ensures that observations near the specific location have larger weight” (Liu, Wei, et al., 2015). This matrix is based on the first law of Geography, which is: everything is related to everything else, but near things are more related than distant things. There are many ways to calculate the space weighting matrix, in this research the following Gaussian weighting kernel function was used.

$$w_{ij} = \exp(-(d_{ij}/b)^2)$$

b is the kernel bandwidth. d_{ij} is the Euclidean distance between points i and j. w_{ij} will decrease faster with a larger bandwidth, vice versa.

We can see from this model, if the matrix equals to 1, then this model turns to be an ordinary least squares (OLS model).

Maps of slope parameters (β coefficients), standardized residuals (StdResid), and local R^2 provide a simple visual depiction of the spatial variations in the relationships between street centrality and land uses.

Local R^2 range between 0.0 and 1.0 and indicate how well the local regression model fits observed y values. Very low values indicate that the local model is performing poorly. Mapping the Local R2 values to see where GWR predicts well and where it predicts poorly can tell us about the goodness of fit of the regression model across the space. To obtain the StdResidual, the fitted y values are subtracted from the observed y values. Standardized residuals have a mean of zero and a standard deviation of 1. The coefficient is the $\beta_k(\mu_i, v_i)$ for the independent variable in location i in the model. The function of the coefficient has been shown in above.

Figure 25 shows the spatial variation of the regression results from the GWR model for betweeness (in search radius 1500 meter) and commercial, residential land.

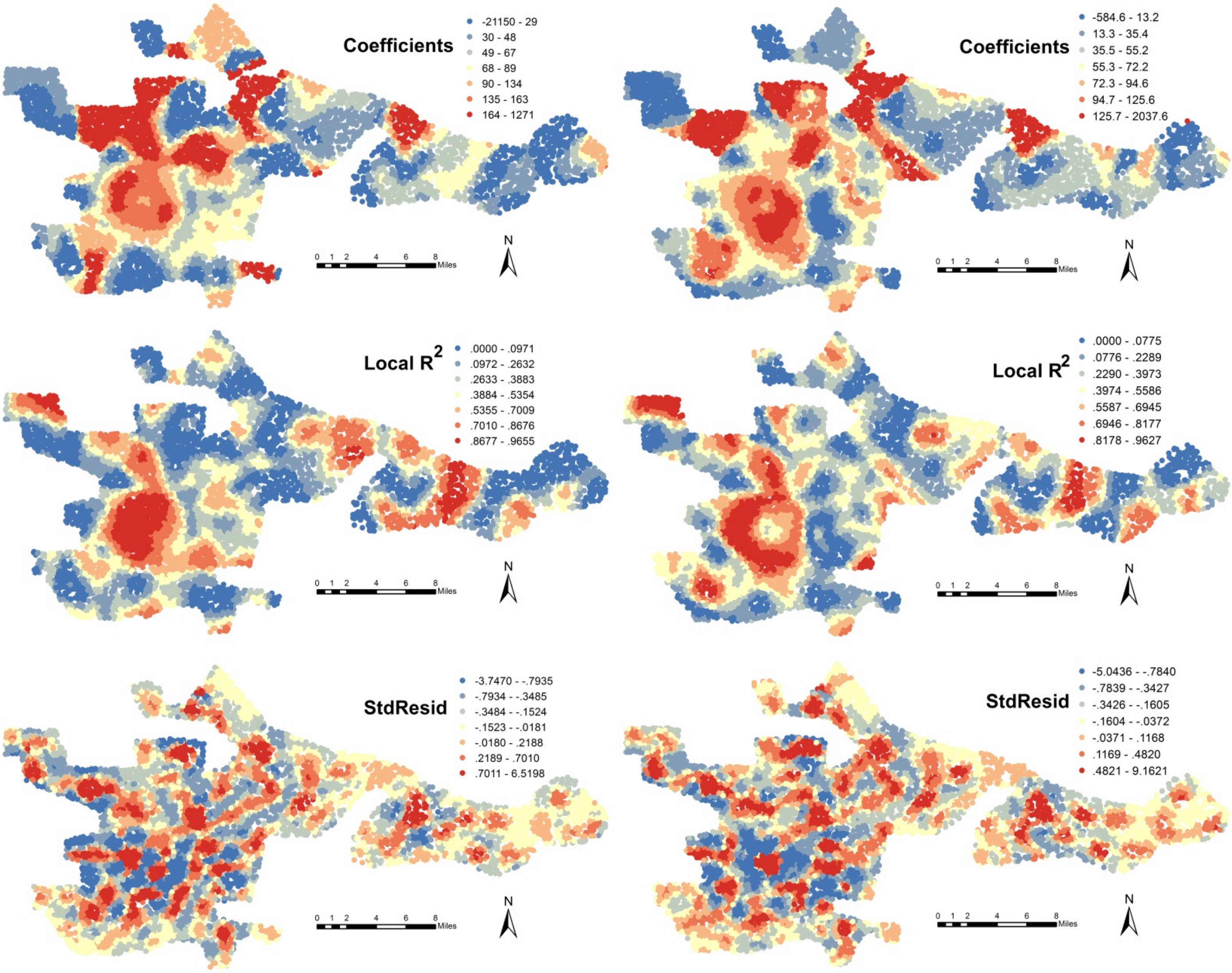

Figure 26 Spatial variations of the regression outputs from the GWR model for betweeness and commercial, residential land.

Left side column of figure 25 are the slope parameter; local R^2; standardized residuals for commercial land. Right side column are the slope parameter; local R^2; standardized residuals for residential land.

For the commercial land in the left side, significant positive correlations are found in the center and north part of the city, whereas higher local R^2 values and significant correlations are located in the central part of the study area. In addition, the spatial pattern of the standardized residuals shows a scattered high and low value. Also compared with the other two maps, the spatial pattern of the standardized residuals shows a relatively random pattern in the whole study area.

For the residential land in the right side, high positive correlations are also found in the center and northern part of this area, but high negative correlations are found in the outskirt of the study area. The spatial pattern of local R^2 values is similar with commercial land correlation. Also, the range of standardized residuals is wider than commercial land.

It is worth mentioning the coefficients of GWR are different from the coefficient of Pearson's correlation in previous parts. Correlation is described as the analysis which can show the association or the absence of the relationship between two variables x and y. It only quantifies the degree to which two variables are related. And the results are ranged from -1 to1. On the other hand, regression analysis predicts the value of the dependent variable based on the know value of the independent variable, and the cause and effect as the regression model is determined as the best way to predict Y from X. That is why there is different order of magnitude for these two kinds of coefficients. Also there is magnitude difference existed in the KDE interpolation raster of these two variables in each GWR model analysis.

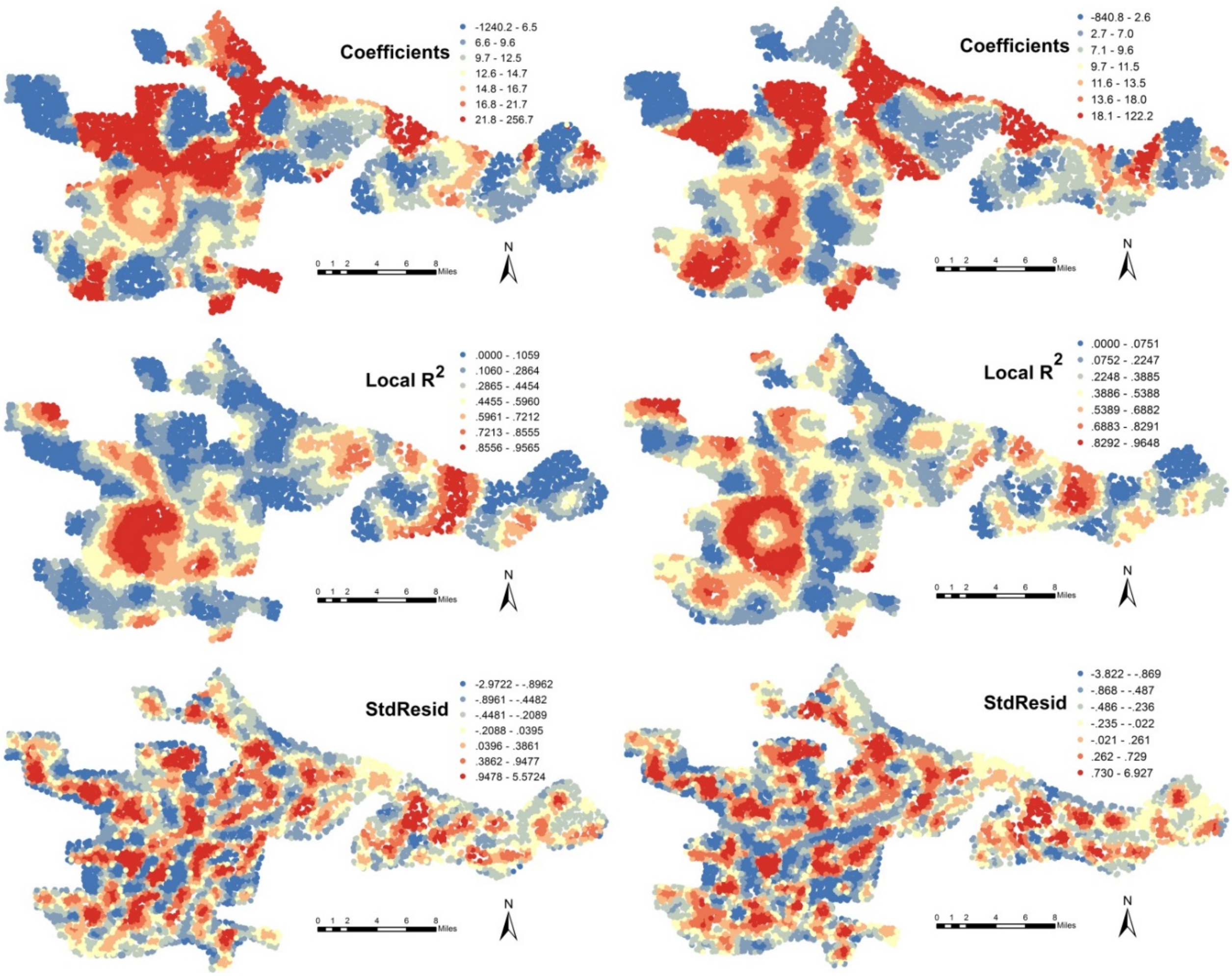

Figure 27 Spatial variations of the regression outputs from the GWR model for straightness and commercial, residential land.

Figure 26 shows the maps of straightness in 1500 meters and commercial, residential land. Maps of local R^2 values and the slope parameters are very similar to those of betweeness and these two lands. By contrast, the coefficients are more concentrated in the northern part in figure 26 rather than the center area in figure 25.

From the coefficients map, we can see that high correlation exists in the districts of Jiangbei and Zhenhai, which means in these two regions, commercial land and residential land are more correlated with straightness centrality.

The patterns of the maps of StdResid are highly similar to those of betweeness and commercial, residential land, which shows a random distribution.

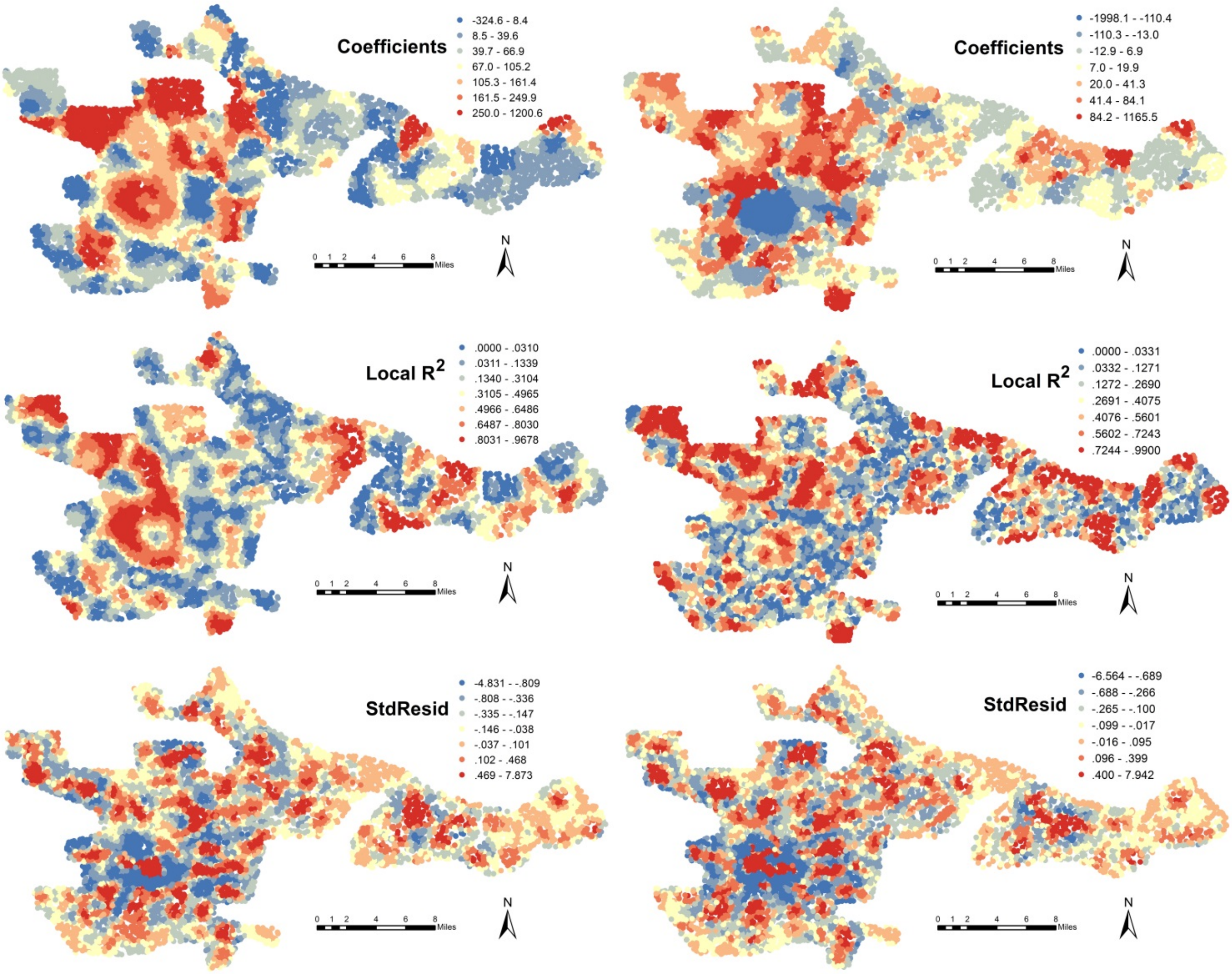

Figure 28 Spatial variations of the regression outputs from the GWR model for betweeness and public services land and industry land.

Figure 27 shows the maps of spatial variations of the regression outcomes from the GWR model for betweeness and public services land and industrial land. The spatial patterns of the correlation between public services land and betweeness, are similar with commercial land, but there is significant difference for the industrial land.

We can see that, the coefficients for industrial land are significant negative in the center of the study area, which is opposite to the previous results. That means, for industrial land, there are negative relations with betweeness centrality in the city center. We can also found there are indeed few industrial lands in this area. But for the region around this area, the relationships are highly positive.

Also for the local R^2, the spatial distribution is more dispersed and the high values are aggregated in Jiangbei district. On the other hand, in the region where the coefficients are not high or low, the local R^2 is high, which means there is no significant correlation for the industrial land and betweeness in Beilun region.

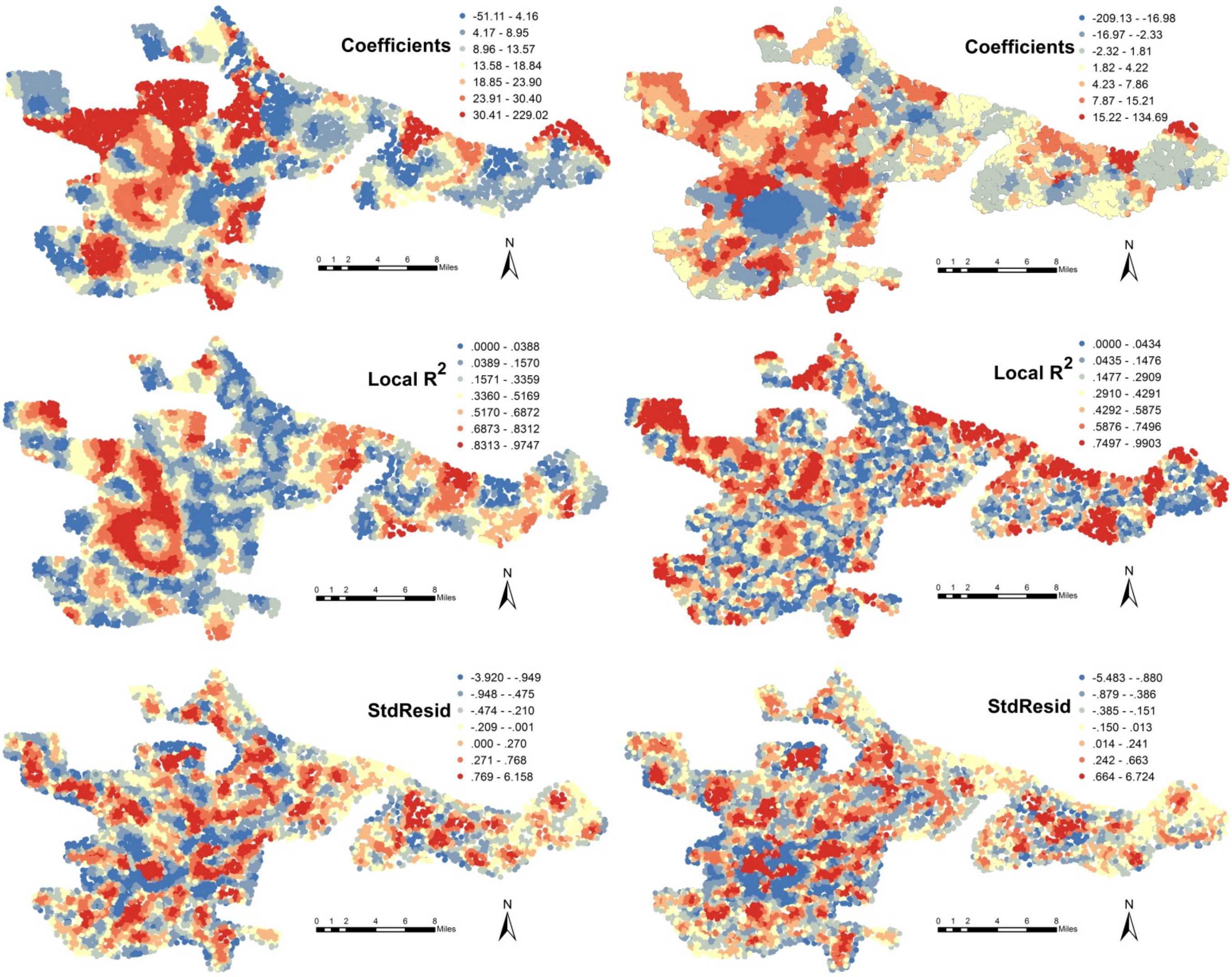

Figure 29 Spatial variations of the regression outputs from the GWR model for straightness and public services land and industry land.

We can see from figure 28, the spatial patterns of straightness are similar with betweeness. For the public services, high-value coefficients are more aggregated in the north part than betweeness and public services.

Also, industrial lands are more correlated with the straightness in the area around the city center. The StdResid values also show a random spatial pattern, which is very similar to that in figure 27.

In the spatial regression part, even though the three centrality measures cannot be considered as the only driving forces of land use distribution, the results verified that the street centrality indices can explain a substantial amount of the land use spatial distribution.

It is obvious that spatial heterogeneity exists in this relationship and refers to the uneven distribution relationship across the region.

Chapter 5 Conclusion and recommendation

This chapter first reviews the background and the problem statement of this research. Then, all the research questions are answered. In the last part, recommendations about relevant policies and further studies are provided.

5.1 Retrospect

Since the opening-up policy begins from the 1980s, there is a rapid urbanization in China, it reflecting and driving a fast expansion of China's transport facilities, and urban street network construction is one of urban infrastructure projects with huge investments in many countries. With that huge amount of financial investments, it is important to understand urban development based on objective regularities, so to improve the resources use efficiency and achieve the goal of sustainable development in a long term.

The research objective is to explore the relationship between urban street network and land use. This paper investigates this relationship in Ningbo, China. Based on the street network composed of links (streets) and nodes (intersections of the streets), street centrality was quantified by three indices. Betweeness centrality measures how often a node is traversed by the shortest paths connecting all pairs of nodes in the network. Closeness centrality measures how close a node is to all the other nodes along the shortest paths of the network. Straightness centrality measures how close the shortest paths from a node to all other lines connecting them. These three centrality indices capture and evaluate the location impact created by the street network. Also, the edge effects were considered in the research, and the street centrality indices were analyzed in a buffer area. Different land use patches were collected: commercial land, residential land, industrial land, public services, grass land and forest. The kernel density estimation (KDE) was used to convert the data sets of centrality indices at polylines, and land use polygons to the same unit in order to investigate the relationship between them. Two search bandwidths: 500m, 1500m were tested for every variable. The statistical distribution of the KDE values of both centrality and land use data indicating that high values are concentrated in a few places, and thus the best centrality is only enjoyed by a small part of the whole population in the city.

To analyze the relationships, Pearson's correlation was used to show the global relationship between each land use and each street network centrality. In addition, how this relationship varied across space was analyzed by GWR model. This model was applied based on a interpolation raster and only one explanatory variable was used. The maps of standardized residuals and local R^2 that were obtained through the GWR models showed spatial heterogeneity.

Academically, this study showed different land use preference of location in street network, which was poorly studied in previous researches. Practically, it provided a method to evaluate different possible plans by simulating the regional and location influences brought by these new adding roads in each plan and helps to choose the optimal one. In conclusion, by using this method, we can put up with advice for urban land use planning and also realize the land value in a better way.

5.2 Answer to research question

5.2.1 The answer of sub-question 1

What are the spatial properties of the street network?

For the betweeness indicator, high values are aggregated in the center of study area and scatter distributed in other places. Betweeness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes.With the increase of search radius, more streets have high betweeness level.

From the closeness indicator, we can see that "bridge" roads also have the advantage to reach other nodes in a short distance. With the increase of search radius, the number of low closeness roads increases, which means it will be easier to reach a far area in the high-speed vehicle.

Within the radius of 5000 and 8000meters, low straightness streets distributed in the outlier of the study area. That means when departing from these outliers to other places, average straight line distance is much shorter than the geometry distance.

5.2.2 The answer of sub-question 2

What are the spatial distributions of different land uses in the study area?

Ningbo is a polycentric area. Commercial and residential area are mainly distributed in the city center, also gathered in another city center in Beilun district. Public services area has a similar spatial pattern with the commercial and residential land, but distributed more widely. Industrial land distributed along the whole study area but also aggregated in Beilun district. Grass lands are evenly distributed, as well as the forest lands. But forest land patches mostly distributed in Zhenhai district and also the peripheral region of the study area.

5.2.3 The answer of the main question

To what extant are street network correlated with different land use?

Spearman correlation coefficient was computed to investigate the relationship between each pair of land use and centrality. Commercial land, public services land and residential land are highly correlated with betweeness and straightness centralities. The highest correlation coefficient is 0.8 between commercial land and betweeness with a 1500 meter bandwidth based on the KDE results, and 0.77 between commercial land and straightness with the same bandwidth. For the industrial, forest and grass land, there is almost no correlation between three centrality indices and these land uses.

Using the GWR model, the relationships were proven to be nonstationary and spatial. Spatial non-stationarity is a condition in which a simple "global" model cannot explain the relationships between some sets of variables. There are significant positive correlations of commercial, residential land and betweeness in the city center area. In addition, commercial and residential land are more correlated with straightness centrality in the northern part, which is Jiangbei and Zhenhai districts. However, for industrial land, there are negative relations with betweeness and straightness centralities in the city center.

The results indicate strong relationships between street centrality and commercial, residential, public services area. Furthermore, the relationships vary not only among different land use types but also in different categories of a single land use type. This result confirms that street centrality can capture the land development of different land uses and plays an important role in shaping urban fabrics.

5.3 Recommendations

There are some future research recommendations.

First, street class and street width were not considered in this research because of the data limitation. However, in practice, it would have an impact on the location properties in the street network. Therefore, the MCA model can be improved with a street width-based weight. Consequently, the correlation between the weighted street centrality and land uses would be more precise.

Second, this research only selected one city Ningbo, which is a polycentric city. But this geographic scale is not sufficient for a thorough analysis of the relationship between urban street network and land use. Each city has its own characteristics and different planning schemes. Therefore, if further research expects to conduct a full and detailed analysis, it would be better to choose more than one polycentric city to study.

Third, this research did not consider other factors' impact on the relationships. However, the land use pattern could not be determined only by the urban street network. For example, top-down urban planning evolution affects land use pattern directly. Thus, it is also important for the further study to figure out other factors which would have the impact on land use patterns.

Bibliography:

Al-Sayed, K., Turner, A., Hillier, B., Iida, S., et al., 2014. Space syntax methodology.

Anselin, L. and Getis, A. 2010. Spatial statistical analysis and geographic information systems. Spatial statistical analysis and geographic information systems. 2010. Perspectives on spatial data analysis. Springer. pp. 35-47.

Batty, M., 2008. The size, scale, and shape of cities. *Science (New York, N.Y.),* 319 (5864), pp. 769-771.

Brunsdon, C., 1995. Estimating probability surfaces for geographical point data: an adaptive kernel algorithm. *Computers & Geosciences,* 21 (7), pp. 877-894.

Bulmer, M., 1981. Quantification and Chicago social science in the 1920s: A neglected tradition. *Journal of the History of the Behavioral Sciences,* 17 (3), pp. 312-331.

Cardillo, A., Scellato, S., Latora, V. and Porta, S. 2006. Structural properties of planar graphs of urban street patterns. *Physical Review E,* 73 (6), pp. 066107.

Casetti, E., 1993. Spatial analysis: perspectives and prospects. *Urban Geography,* 14 (6), pp. 526-537.

Corcoran, P. and Mooney, P. 2013. Characterising the metric and topological evolution of OpenStreetMap network representations. *The European Physical Journal Special Topics,* 215 (1), pp. 109-122.

Croxford, B., Penn, A. and Hillier, B. 1996. Spatial distribution of urban pollution: civilizing urban traffic. *Science of the Total Environment,* 189 pp. 3-9.

D'Acci, L., 2016. Do Pedestrians Prefer Straight Routes?: Route Angularity and Complexity-Minimizing Path versus Cosy Sinuosity in Urban Street Layout Geometry.

Epanechnikov, V. A., 1969. Non-parametric estimation of a multivariate probability density. *Theory of Probability & its Applications,* 14 (1), pp. 153-158.

Erath, A., Löchl, M. and Axhausen, K. W. 2009. Graph-theoretical analysis of the Swiss road and railway networks over time. *Networks and Spatial Economics,* 9 (3), pp. 379-400.

Evans, G., Heath, A. and Lalljee, M. 1996. Measuring left-right and libertarian-authoritarian values in the British electorate. *British Journal of Sociology,* pp. 93-112.

Freeman, L. C., 1977. A set of measures of centrality based on betweenness. *Sociometry,* pp. 35-41.

Harris, C. D. and Ullman, E. L. 1945. The nature of cities. *The Annals of the American Academy of Political and Social Science,* 242 (1), pp. 7-17.

Hickman, R., Bonilla, D., Givoni, M. and Banister, D., 2015. International handbook on transport and development. Edward Elgar Publishing.

Hillier, B., 2005. The art of place and the science of space. *World Architecture,* 185 pp. 96-102.

Hillier, B., Hanson, J. and Peponis, J., eds., 1984. What do we mean by building function? E & FN Spon Ltd.

Hoyt, H., 1939. The pattern of movement of residential rental neighborhoods. *The Structure and Growth of Residential Neighborhoods in American Cities, Federal Housing Administration.Washington DC,* pp. 112-122.

Kang, C., 2016. Spatial access to pedestrians and retail sales in Seoul, Korea. *Habitat International,* 57 pp. 110-120.

Liu, Y., Wei, X., Jiao, L. and Wang, H. 2015. Relationships between Street Centrality and Land Use Intensity in Wuhan, China. *Journal of Urban Planning and Development*, 142 (1), pp. 05015001.

Loelkes, G. L., 1983. Land use/land cover and environmental photointerpretation keys. US Dept. of the Interior, Geological Survey.

Lowry, I. S., 1964. A model of metropolis. Rand Corporation Santa Monica, CA.

Marshall, S., 2004. Streets and patterns. Routledge.

Mashhoodi, B. and Berghauser Pont, M., eds., 2011. Studying land-use distribution and mixed-use patterns in relation to density, accessibility and urban form, [ISUF 2011: 18th International seminar on urban form: Urban Morphology and the Post-Carbon City, Montreal, Canada, 26-29 August 2011].

Mills, E. S., 1972. Studies in the Structure of the Urban Economy.

Muth, R., 1969. Cities and housing: The spatial patterns of urban residential land use. *University of Chicago, Chicago,* 4 pp. 114-123.

Pinho, P. and Oliveira, V. 2009. Cartographic analysis in urban morphology. *Environment and Planning B: Planning and Design,* 36 (1), pp. 107-127.

Porta, S., Crucitti, P. and Latora, V. 2006a. The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications,* 369 (2), pp. 853-866.

Porta, S., Crucitti, P. and Latora, V. 2006b. The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design,* 33 (5), pp. 705-725.

Porta, S., Latora, V., Wang, F., Rueda, S., et al., 2012. Street centrality and the location of economic activities in Barcelona. *Urban Studies,* 49 (7), pp. 1471-1488.

Porta, S., Strano, E., Iacoviello, V., Messora, R., et al., 2009a. Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design,* 36 (3), pp. 450-465.

Porta, S., Strano, E., Iacoviello, V., Messora, R., et al., 2009b. Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design,* 36 (3), pp. 450-465.

Ratti, C., 2004. Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design,* 31 (4), pp. 487-499.

Sabidussi, G., 1966. The centrality index of a graph. *Psychometrika,* 31 (4), pp. 581-603.

Sevtsuk, A., 2010. *Path and Place: A Study of Urban Geometry and Retail Activity in Cambridge and Somerville, MA,* .

Shaw, S., Fang, Z., Lu, S. and Tao, R. 2014. Impacts of high speed rail on railroad network accessibility in China. *Journal of Transport Geography,* 40 pp. 112-122.

Van Thiel, S., 2014. Research methods in public administration and public management: an introduction. Routledge.

Venckauskaite, J. and Skrodenis, E. 2007. The principles of sustainable street network and safe traffic system in Lithuanian cities. *Slovak Journal of Civil Engineering,* 15 (1), pp. 45-50.

Wang, F., 2006. Quantitative methods and applications in GIS. CRC Press.

Wang, F., Antipova, A. and Porta, S. 2011. Street centrality and land use intensity in Baton Rouge, Louisiana. *Journal of Transport Geography,* 19 (2), pp. 285-293.

Wang, F. and Guldmann, J. 1996. Simulating urban population density with a gravity-based model. *Socio-Economic Planning Sciences,* 30 (4), pp. 245-256.

Wang, H., 2016. Analysis of the Sustainable Development Strategy of the China Highway Policy. *Journal of Highway and Transportation Research and Development (English Edition),* pp. 103-110.

Wegener, M., 1994. Operational urban models state of the art. *Journal of the American Planning Association,* 60 (1), pp. 17-29.

Annex 1: IHS copyright form

In order to allow the IHS Research Committee to select and publish the best UMD theses, participants need to sign and hand in this copy right form to the course bureau together with their final thesis.

Criteria for publishing:

A summary of 300 to 500 words should be included in the thesis.

The number of pages for the thesis is about 60.

The thesis should be edited.

Please be aware of the length restrictions of the thesis. The Research Committee may choose not to publish very long and badly written theses.

By signing this form you are indicating that you are the sole author(s) of the work and that you have the right to transfer copyright to IHS, except for items cited or quoted in your work that are clearly indicated.

I grant IHS, or its successors, all copyrights to the work listed above, so that IHS may publish the work in *The IHS thesis series*, on the IHS web site, in an electronic publication or in any other medium.

IHS is granted the right to approve reprinting.

The author(s) retain the rights to create derivative works and to distribute the work cited above within the institution that employs the author.

Please note that IHS copyrighted material from *The IHS thesis series* may be reproduced, up to ten copies for educational (excluding course packs purchased by students), non-commercial purposes, providing full acknowledgements and a copyright notice appear on all reproductions.

Thank you for your contribution to IHS.

Date : ____

Your Name(s) : ____

Your Signature(s) : ____

Please direct this form and all questions regarding this form or IHS copyright policy to:

The Chairman, IHS Research Committee Burg. Oudlaan 50, T-Building 14th floor, 3062 PA Rotterdam, The Netherlands	j.edelenbos@ihs.nl Tel. +31 10 4089851