Spatio-temporal Analysis of Land Use Change: Shenzhen as a Case Study

LIU, Biao

A Thesis Submitted Partial Fulfillment
of the Requirement for the Degree of
Doctor of Philosophy
in
GeoInformation Science

The Chinese University of Hong Kong
September 2010

Thesis/Assessment Committee

Professor Tung Fung (Chair)

Professor Bo Huang (Thesis Supervisor)

Professor Xinyuan Song (Committee Member)

Professor Jianquan Cheng (External Examiner)

ABSTRACT

Abstract of thesis entitled:

Spatio-temporal Analysis of Land Use Change: Shenzhen as a Case Study

Submitted by LIU, Biao

for the degree of Doctor of Philosophy

at The Chinese University of Hong Kong in July 2010

Research focusing on land use change analysis is of tremendous importance in global change studies. Land use change modeling, which is a prerequisite to understanding the complexity of land use dynamics, is an effective way to describe the change patterns and delve into the causes for the changes. Despite the development of many models in the past, several important issues still remain to be addressed such as spatio-temporal non-stationarity, spatio-temporal correlation, and individual effect. The primary objective of this research is to make improvements on.the traditional logistic models to suit the characteristics and requirements of land use change modeling. Specifically, three enhancements have been made. The first enhancement deals with spatio-temporal non-stationarity, the second improvement aims to incorporate spatio-temporal autocorrelation, and the third includes individual effect.

Three spatio-temporal logit models for land use change analysis, namely, geographically and temporally weighted logit model (GTWLM), spatio-temporal panel logit model (ST-PLM) and generalized spatio-temporal logit model (GSTLM), are proposed accordingly to deal with the aforementioned issues. GTWLM, which considers spatio-temporal non-stationarity, includes temporal data in a spatio-temproal framework by proposing a spatio-temporal distance. ST-PLM incorporates the spatio-temporal correlation and individual effect in one model, where spatio-temporal correlation is considered in the random individual effect with an assumption that the correlation between such components is inversely

proportional to the spatio-temporal distance. By integrating GTWLM and ST-PLM, the GSTLM explores spatio-temporal non-stationarity and correlations simultaneously, whilst considering their individual effects to construct an integrated model.

Based on the models, a case study is performed on multi-temporal land use change analysis in the Special Economic Zone (SEZ), Shenzhen. The results show that all the proposed models outperform the traditional logistic regression model: multinomial logit model (MNLM), which overlooks the aforementioned issues. Compared with MNLM, GTWLM and ST-PLM increased the percentage of correctly predicted (PCP) values from 74.1% to 82.3% and 79.4%, respectively. McNamara's test shows that the differences between those models are significant. The kappa coefficients reveal that the GTWLM and ST-PLM are better than MNLM. In particular, the GSTLM yields a considerably higher PCP of 85.9%. The Kappa coefficients also indicate that the GSTLM is the most optimal model. Furthermore, the GTWLM allows the model parameters to vary across space and time, which provides deep insights into the spatio-temporal variations of the land use pattern. Assisted with the visual results, the spatio-temporal land use distribution patterns in Shenzhen are analyzed and the results presented thereafter.

摘要

在全球变化研究中，土地利用变化分析一直被广泛关注。土地利用变化建模致力于理解土地利用动态变化的复杂性，是描述土地利用变化和探求变化原因的一种有效方式。虽然以往的研究中涌现出大量土地利用变化模型，但在土地利用变化分析建模中，还是存在一些重要的、尚未解决的问题。本研究的主要目标是开发一系列模型来处理这些重要问题。具体地说，本研究考虑土地利用建模的特性和需求，对传统的逻辑回归模型进行改进，使之能处理土地利用变化建模中的时空非平稳性，时空自相关性，以及个体效应。

本文提出三个时空逻辑回归模型来分析土地利用变化：时空地理加权逻辑回归模型(GTWLM)、时空面板逻辑回归模型(ST-PLM)及广义时空逻辑回归模型(GSTLM)。GTWLM 通过构造时空距离来考虑时空非平稳性。该模型能够更细致地考察因变量和自变量关系的时空变化。ST-PLM 则在随机效应面板数据模型的基础上，通过赋予个体效应相关性来解决土地利用变化的时空自相关性和个体效应。GSTLM 通过结合前面两个模型，同时解决土地利用变化的时空非平稳性、时空自相关性及个体效应。

基于以上提出的模型，本文分别对深圳经济特区的时空土地利用展开分析。结果显示本研究提出的所有模型都优于传统的逻辑回归模型(MNLM)。与 MNLM 相比，GTWLM 和 ST-PLM 的 PCP 分别从 74.1%上升到 82.3%和 79.4%。McNamara 检验显示这些模型间存在显著差别。Kappa 系数表明 GTWLM 和 ST-PLM 优于 MNLM。值得注意的是，同时考虑三个问题的 GSTLM 得到了相当高的正确预测率(PCP: 85.9%)。Kappa 系数也显示 GSTLM 是最优模型。此外，GSTLM 还提供了系数的时空变化图，以支持土地利用形态的时空分析。

Table of Content

List of Tables

List of Figures

Acknowledgements

There are many people who deserve my deep appreciation for their roles in making this thesis a reality. Principally I want to express the gratitude to my supervisor, Professor Bo Huang. He patiently read and revised my drafts, and gave me a lot of valuable suggestions. Without his guidance and assistançe, this thesis would never be completed.

I would also like to give my thankfulness to my thesis committee members, Professor Tung Fung, Professor Xinyuan Song and Professor Jianquan Cheng for their precious comments on my research work. I learned a lot from their courses and publications and through discussions.

Special thanks go to Dr. Li Hongga for his help on data preparation.

Finally, I want to express my thanks to all the professors, students, and supporting staff in ISEIS. They have been supporting me throughout the period of my study at CUHK.

Chapter 1: Introduction

1.1 Background

The 20^{th} century has witnessed significant changes in the land use patterns. During the last three centuries, almost 1.2 million km^2 of forest and woodland and 5.6 million km^2 of grassland and pasture have been converted to other uses, globally (Ramankutty and Foley 1999). Meanwhile, the impacts of land use change have escalated from significant to threatening proportions around the globe. Lots of environmental problems could be traced back to land use change: such as desertification, eutrophication, acidification, climate change, eustatic sea-level rise, greenhouse effect, and biodiversity loss. The study of land use change has, therefore, received a great deal of attention from policy makers, planners, and developers for developing a sustainable land use plan (Turner and Meyer, 1991; Lambin *et al.*, 2003; Rindfuss *et al.*, 2004; Pontius *et al.*, 2007; Huang *et al.*, 2009a). Both the International Geosphere-Biosphere Programme (IGBP) and the International Human Dimensions Programme (IHDP) have proposed a Land Use and Cover Change (LUCC) Core Project/Research Programme.

Considering the current land use change that is occurring all over the world, China is undergoing the most rapid and largest land use change. During the past century, $2.06*10^7$ hm^2 of land in China was urbanized (Ge *et al.*, 2000). These rapid land use changes have been accompanied by numerous related problems in recent years. (e.g., snow disaster in 2007, drought in 2010, food shortage, etc.). At present, the urban population constitutes 30 % of the total population of China. According to Paulussen (2003), this will increase to more than 50 % in the next 20 years, which implies more rapid and larger land use change in the future.

In China, land use change is mostly noticeable in some big cities. Shenzhen is one of the most developed cities in China and this city has grown at a dramatic speed. In less than 30 years, Shenzhen, a tiny border town of 30,000 people in 1979, has developed into a modern metropolis. The land use pattern of Shenzhen has changed drastically in the past three decades. A careful study of the land use change in Shenzhen will benefit the development of other cities in China. Special Economic Zone (SEZ) is an ideal choice for the study due to its 'one city, two systems' framework in Shenzhen.

In order to study the land use change, considerable research in the domain of spatial databases and GIS has been undertaken to track spatio-temporal change, e.g., land cell fabric change. A well-designed spatio-temporal data model is fundamental to spatio-temporal data management even analysis. Several spatio-temporal data models have hence been developed to facilitate data representation and perform both spatial and temporal queries; for details see Huang and Claramunt (2005), and Yuan and Stewart (2008). Meanwhile, a number of empirical models have been devised to simulate and model land use change using Cellular Automata (CA) and statistical regressions; see reviews in Agarwal *et al.* (2002) and Parker *et al.* (2003). These models aid interpreting and predicting land use change patterns. The efforts by different researchers over the years have yielded varying levels of success. However, due to the shortage of reasonable methods to represent the complicated spatial and temporal effect in the spatio-temporal process of land use change, the modeling accuracy and the validation of theoretical models continue to pose challenges (Huang *et al.*, 2009b). Thus, there is clearly a need to develop more sophisticated models for understanding and analyzing land use change, wherein the most compelling issue is the dynamics of land use change within the spatio-temporal framework.

1.2 Problem Statement

Even though earlier land use change studies have demonstrated varying levels of success in their specific domains, several problems still need to be addressed.

First of all, with the limitation of land resource, not all the land use change is from rural to urban, more and more land use changes are occurring within the urban area. The modeling methodology should be capable of accommodating multi-class land use change (e.g. residential to commercial and industrial, etc. in support of urban redevelopment studies) rather than binary changes only (e.g. rural to urban in support of urban sprawl studies).

Secondly, due to different administrative, political, or other contextual issues, the relationships between land use change (or multi-temporal land use pattern) and causal factors in different locations and time may be intrinsically different. Hence, spatio-temporal non-stationarity, which indicates the different relationships exist at different locations and time, need to be considered in land use change analysis.

Thirdly, substantial amounts of spatial variables and spatial land-use data tend to be self-dependent. This phenomenon is referred to as spatial autocorrelation (Legendre and Fortin, 1989). When dealing with spatio-temporal data, specific approaches should be considered to incorporate both spatial and temporal autocorrelation. Otherwise, inefficient parameter estimates and inaccurate measures of statistical significance will result.

Fourthly, each land cell is characterized by its own individual effect (e.g. If the land type of a cell is urban, the probability of land use change on this cell will be lower than rural.). Under such circumstances, a more detailed analysis of the individual effect in modeling land use change should be incorporated.

Finally, all the problems mentioned above are not totally disparate (e.g. spatio-temporal non-stationarity can be partly considered in spatio-temporal autocorrelation). Thus, the land use change model should be capable of generalizing a series of changes in spatio-temporal framework into a unified model for better projection of land use distribution. This should be done with due consideration of multi-class land use type, spatio-temporal non-stationarity, spatio-temporal autocorrelations, and individual effect.

1.3 Research Objectives

This research aims to develop a set of innovative models within the logistic regression framework to address the challenges identified and apply it for analyzing the land use change in SEZ, Shenzhen. Specifically, the objectives are to:

- Extend the logit model to support multi-class land use change analysis, which implies a more complicated spatio-temporal effect;
- Extend the logit model to include spatio-temporal non-stationarity in multi-class land use change;
- Extend the logit model to include spatio-temporal autocorrelation in multi-class land use change;
- Extend the logit model to incorporate the individual effect for each land cell;
- Extend the logit model to generate a unified model for a time series of land use distributions, whilst considering spatio-temporal non-stationarity, spatio-temporal autocorrelations, and the individual effect in multi-class land use change;
- Evaluate the performance and the reliability of the proposed approaches using the land use data in SEZ, Shenzhen;
- Analyze and interpret the land use change in the SEZ, Shenzhen.

1.4 Research Significance

This study aims to develop a set of statistical models to effectively address land use change data. Land use change data has both spatial and temporal characteristics as land use change is a complex spatio-temporal process. Statistically modeling quantitative relationships between response variable and explanatory variables for spatio-temporal data is a two-fold problem. The first issue is that spatial or temporal non-stationarity occurs in the relationships being modeled; the second is that spatial or temporal autocorrelation exists between the observations. Traditional models such as multinomial logit model used in land use change modeling have largely ignored these two issues that violate the basic assumptions. Spatio-temporal non-stationarity violates the assumption that a single relationship exists across the sample data observations, while spatio-temporal autocorrelation violates the assumption that explanatory variables remain unchanged during repeated sampling. Though the two problems are theoretically distinct, many evidences evince that there exists non-stationarity in the presence of autocorrelation. An inadequate model that fails to capture non-stationarity will result in residuals that exhibit autocorrelation. Lesage (2005) argued that it is better to cover both spatial and temporal heterogeneity and autocorrelation effects in a mixed model. Therefore, building a model that integrates spatio-temporal non-stationarity with autocorrelation effects is of considerable importance. Besides, other features of land use change data, such as multi-class land use type and individual effect, should also be considered. This will aid modeling the land use change in a precise manner.

The models proposed in this research will build up the relationship between various factors and land use patterns over time while addressing the aforementioned issues in a logistic regression framework. In order to incorporate spatio-temporal non-stationarity, the geographically weighted logit model (GWLM) that can deal with spatial non-stationarity is extended to geographically and temporally weighted logit model (GTWLM). Then, based on the panel data analysis framework, the spatio-temporal panel logit model (ST-PLM), which incorporates the covariance of a cell to other cells into the model formulation to involve spatio-temporal autocorrelation, considers the individual effect of a land use cell. Finally, the two aforementioned models are combined to a generalized spatio-temporal logit model (GSTLM). Notably, GSTLM can handle spatio-temporal non-stationarity, spatio-temporal autocorrelation, and individual effects in a unified manner. The effectiveness of the proposed models is finally tested using a study of the multi-temporal land use change in the SEZ, Shenzhen. The results demonstrate that our models significantly outperform the conventional model.

This research will aid urban planners and policy makers to effectively understand the land use change in a spatial and temporal framework, make more precise land use projections, and hence facilitate generating systematic land use plans.

1.5 Research Framework

In this research, three spatio-temporal logit models are developed and employed to analyze land change in Shenzhen. The framework is shown in Figure 1.1. First of all, the data from different sources should be collected and processed. Since the total number of observations in the SEZ, Shenzhen is too large, a set of randomly selected samples are used. The geographically and temporally weighted logit model (GTWLM) is developed to andle spatio-temporal non-stationarity. The spatio-temporal panel logit model (STPLM) is developed to manage the individual effect and spatio-temporal autocorrelation. Then, STPLM and GTWLRM are combined to construct a generalized spatio-temporal logit model (GSTLM) for land use change modeling. Finally, based on the results of generalized spatio-temporal logit model, the spatio-temporal land use distribution patterns in the SEZ, Shenzhen are analyzed.

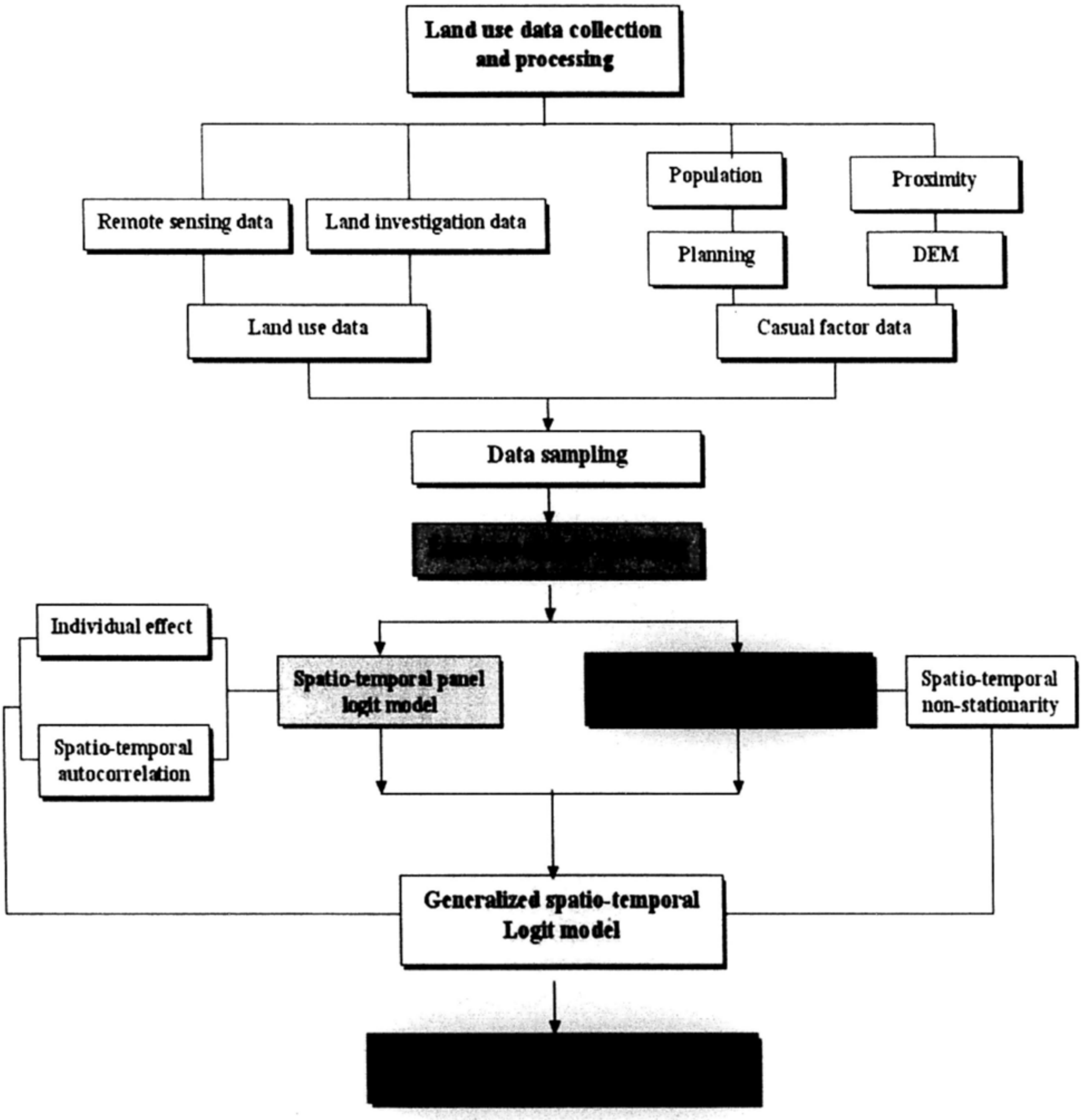


Figure 1.1: Spatio-temporal analysis of land use change in SEZ, Shenzhen

1.6 Thesis Organization

This thesis is divided into seven chapters. The remainder of this thesis is organized as follows:

Chapter 1 lays out the background, the problems, and the objectives of the research. It also outlines the research significance, research framework, and thesis organization.

Chapter 2 reviews the earlier work on land use change modeling. The introduction of land use modeling is presented at first. Then, the causal factors driving land use change are discussed comprehensively. Finally, the literature review discusses the common methodologies for land use change modeling, such as cellular automata, multi-agent system, Markov chain analysis, logistic regression model, and other models. Specifically, spatio-temporal models for land use change will be focused on. The limitations and advantages of the techniques for land use change modeling are also analyzed.

Chapter 3 describes the study area (SEZ) and delineates the land use change situation in Shenzhen over the past decades. Then, data preparation and sampling are discussed. Subsequently, the preliminary analysis of land use change in Shenzhen is carried out using the multinomial logit model.

Chapter 4 deals with the spatio-temporal non-stationarity in geographically and temporally weighted logit model. At first, a brief of introduction to spatio-temporal non-stationarity in land use change modeling is provided. Then, GWLM model is extended to a spatio-temporal framework to consider spatio-temporal non-stationarity. Subsequently, the performances of the geographically and temporally weighted logit model are evaluated. Finally, the conclusion and the chapter summary are provided.

Chapter 5 deals with the spatio-temporal autocorrelation and the individual effect to enhance the land use change modeling process. Firstly, a brief introduction to spatio-temporal autocorrelation in land use change modeling is provided. Then, based on the panel data model, improvements on individual effect and spatio-temporal autocorrelation are developed. Following this, the performance of the improved panel data model (spatio-temporal panel logit model) is evaluated to demonstrate its efficiency. Finally, the conclusion and chapter summary are provided.

Chapter 6 integrates GTWLM and STPLM to deal with spatio-temporal non-stationarity, spatio-temporal autocorrelation, and individual effect in a unified model: generalized spatio-temporal logit model. After a brief review, the formulation of the generalized spatio-temporal model is presented. Following that, the results and discussion are provided and the spatio-temporal land use distribution pattern in Shenzhen is analyzed. Finally, the conclusion and chapter summary are given.

Lastly, chapter 7 provides the conclusions, along with the findings of the study with reference to the aforementioned research objectives. Finally, this chapter discusses the limitations of this research and provides recommendations for further research.

Chapter 2: Land Use Change Modeling

2.1 Introduction

Over the past decades, many nations have witnessed significant changes in land use patterns. The impacts of land use change have escalated from significant to threatening proportions all around the globe. Numerous regions around the globe have been afflicted by deforestation, flooding, food shortage, green house effect, urban extension etc, as a result of disproportionate land use changes. Increasing attention has been paid by planners and decision-makers around the globe to develop sustainable land use plans.

Land use change models aim to understand the causes and consequences of land use dynamics. Currently, several modeling approaches with different outcomes for land use change simulation and exploration are available. Such approaches offer possibilities for experiments to test our understanding of the key process, thus allowing the model to describe the sensitivity of quantitative changes and provide alternative pathways into the future.

Earlier studies in land use modeling involve dividing the study area into grid cells and describing it by a pre-determined set of biophysical and socio-economic variables. Based on this, methods such as cellular automata (Clarke and Gaydos 1998; Wu 1998; Wu 2002), agent (Kohler and Gumerman 2000; Gimblett 2002), Markov chain analysis (Lopez *et al*, 2001) and logistic regression (Wu and Yeh 1997, Cheng and Masser 2003) are employed. The causal factors and the main models (CA, agent, regression, etc.) used for LULC change analysis are described in the following sections.

2.2 Causal Factors

Since land use change modeling endeavors to explore the relationship between causal factors and land use pattern or change (multi-temporal land use pattern), the analysis of causal factors is an indispensable part in land use change modeling.

Land use change is a complex process influenced by a number of factors. As is evident from earlier studies, no single set of factors can explain the changes. Researchers (Turner, 1995; Bicik *et al.*, 2001) have provided a summarization of the causal factors. They have mentioned two kinds of factors: 1) Natural factors, such as climate change, soil, hydrology and nature disaster; 2) Human factors, such as population, technology level, economy growth, and technology level. The IHDP (Nunes *et al.*, 1999) report also summarizes the causal factors into natural and human factors.

Initially, most researchers focused on natural factors, which were treated as the determinants to land use change. Crow *et al.* (1999) and Naveh (1995) considered the impact from climate change, soil, and DEM. Climate change was considered as a major causal factor to land use change in a large spatial and temporal scale. The influence of natural factors on land use change is realized over a long time scale. However, human factors play a key role in land use change in a short time scale. Hence, more attention has been paid to the study of human factors. Many researchers consider economic growth as a very important factor (Bingham *et al.*, 1995; Hougtopn, 1994; Fischer *et al.*, 2001; Bicik *et al.*, 2001). Also the impact from population growth, political regimen, and technology development is studied (Reid *et al.*, 2000; Bicik *et al.*, Houghton, 1994). Nowadays, due consideration is given to both human and natural factors.

The important parameters that influence land use changes are as follows.

(1) Demographic factors (population size, population growth, and population density) are widely treated as major causal factors of land use change (Verburg *et al.*, 2001).
(2) Accessibility is also often viewed as a significant driver for land use change through its effect on the cost of transportation and ease of settlement (Geist and Lambin, 2001).
(3) Spatial details play an important role in land use change process (White *et al.*, 1997).

(4) A causal force analysis conducted by Chen (2000) found that policy and economy generally influence land use change.

After examining relevant citations, Xie *et al.* (2006) provide a summary of causal factors commonly used in different land use change models as follows:

Table 2.1: Causal factors for land use change

Category	**Causal factor**
Demography	Population size
	Population growth
	Population density
Proximity	Distance to road
	Distance to town/market
	Distance to settlement
	Distance to shopping mall
	Proximity to the urban structure
Economic	Investment structure
	Industry structure
	Housing commercialization
	Returns to land use (costs and prices)
	Job growth
	Cost of land use change
	Rent
Social	Affluence
	Human attitudes and values
Collective rule making	Zoning
	Tenure
Site characteristics	Soil quality
	Slope
Constraints	Water body
	Environment sensitive area
Neighborhoods	Availability of exploitable sites

	Agglomeration of developed areas
Others	Technology level

2.3 Land Use Change Models

2.3.1 Cellular Automata

Tobler (1979) proposed using cellular automata (CA) as a tool for modeling spatial dynamics. Couclelis (1985, 1988, and 1997) explored the implications of the idea in a series of theoretical papers. The approach has been implemented by others in a wide range of applications (e.g. Batty and Xie, 1994; Benenson, 1998; Ceccini and Viola, 1990; Clark *et al.*, 1997; Papini and Rabino, 1997; Phipps, 1989; Portugali and Benenson, 1995; White and Engelen, 1993, 1999; White, Engelen and Uljee, 1997). Recently, the approach has been linked to GIS and applied to land use change study.

CA (Cellular Automata) are dynamic spatial systems in which the state of each cell in an array depends on the previous state of the cells within the cell's neighborhood, based on a set of state transition rules. As the system is discrete and iterative and involves interactions only within local regions rather than between all pairs of cells, it is hence possible to work with grids containing substantial cells. The very fine spatial resolution that can be attained is an important advantage when modeling land use dynamics. A conventional cellular automaton consists of

1) a *Euclidean space* divided into an array of identical cells;
2) a *neighborhood* of defined cell size and shape;
3) a set of discrete *cell states;*
4) a set of *transition rules,* which determine the state of a cell as a function of the states of cells in the neighborhood; and
5) *discrete time steps*, wherein all cell are states updated simultaneously.

However, these defining characteristics can be interpreted broadly, or relaxed in response to the requirements of a particular modeling problem. Hence, many types of CA could exist.

Cellular automata offer a number of advantages. As already mentioned, they permit extreme spatial detail. Hence, they are able to reproduce the actual complexity in nature. Besides, owing to the high resolution and raster nature, they are compatible with GIS databases, and can be linked with them in a relatively simple manner. At the other end of the spatial scale, CA can be linked through their transition rules to other, macro-scale models that limit or drive the CA dynamics. This facilitates comprehensively modeling integrated environmental-human systems. Finally, CA are defined and calibrated in a single operation, since calibration corresponds to finding the optimal transition rules. This means that CA models are typically implemented much faster and simply than traditional spatial models.

Cellular automata also have some disadvantages. Many CA models inductively assume that the human impact is important, but do not explicitly model decisions. Others explicitly hypothesize a set of agents coincident with lattice cells and use transition rules as proxies to decision making. These efforts succeed when the unit of analysis is tessellated, decision-making strategies are fixed, and heterogeneous actors are affected by local neighbors in a simple, well-defined manner. However, when actors are not tied to location in the intrinsic manner of CA cells, the problem of spatial orientedness might arise (Hogeweg 1988). 'Spatial orientedness' refers to the extent to which neighborhood relationships do not reflect actual spatial relationships. This can be remedied using techniques that have non-uniform transition rules and can dynamically change the strength and configuration of the connections between cells. As these characteristics are beyond the capacities of the rigidly defined CA, the pure, traditional CA method may not be generally suited to model land-use/cover change. Furthermore, CA models focus on the simulation of spatial patterns rather than the interpretation of the spatio-temporal land use change process.

2.3.2 Multi-agent System

While cellular automata models focus on transitions, agent-based models focus on human actions. Agents are the basic component in these models. Several characteristics of agents are described as follows:

1) Agents are autonomous;

2) Agents share an environment through agent communication and interaction;
3) Agents make decisions that link behavior to the environment.

Agents are used to represent a lot of entities, including animals, people, and organizations. Agents must act according to some rules, which link their autonomous goals to the environment. An autonomous agent needs to enable it react to environmental changes. Beyond pure reaction, some of the models are based on rational choice theory. These models generally assume that the actors are perfectly rational optimizers. Consequently, these agents are able to maximize their well-being and can balance long-run vs. short-run payoffs.

A large number of scholars attempt to employ agent-based models in land use/cover change study. In the context of a land use change model, an agent may represent a land manager who can decide a land-use. The model agents also may represent entities or social organizations such as a village assembly, local governments, or a neighboring country. Agent-based models in land use/cover change study consist of the decision rules, such as income maximization, of each human actor, their environmental feedbacks. Land markets, social networks, and resource management institutions may provide other significant environments for interaction. When coupled with a cellular model representing the landscape on which agents act, these models are truly compatible for explicitly representing the land use change processes.

However, the appropriateness of models of perfect rationality for agent-based modeling applied to land-use/cover change remains questionable, given the importance of spatial interdependencies and feedbacks in these systems. Recognition of the complex environment in which human decision-making occurs has resulted in a movement toward agent-based models that employ some variant of bounded rationality (Gigerenzer and Todd 1999). Typically, bounded rational agents have goals that relate their actions to the environment. Rather than implementing an optimal solution that fully anticipates all future states of the system they constitute, they make inductive, discrete, and evolving choices that move them toward achieving goals (Bower and Bunn 2000).

2.3.3 Markov chain analysis

In mathematical terminology, a markov chain is a discrete random process. A discrete random process refers to a system that exists in various states and changes randomly in discrete steps. Since the system changes randomly, it is generally impossible to predict the exact future state of such a system. However, the statistical properties of the system in the future can often be described.

In land use change, different land use categories represent the states of a chain. The markov chain states that the probability of land use category at time t, X_t, (and in fact during all future steps) only depends on the land use category at time t-1, X_{t-1}, and not additionally on the land use category of the system during the previous steps. The changes in the state of land use are called transitions, and the probabilities associated with various state-changes are called transition probabilities. The probability of a land use change from a land use category (state) a_i to a land use category (state) a_j in one time period, 'one step transition probability', is $P(X_t = a_j \mid X_{t-1} = a_i)$. In case of a homogeneous markov chain, the transition probability can be expressed as:

$$P(X_t = a_j \mid X_{t-1} = a_i) = P_{ij} \tag{2.1}$$

The transition probability can be estimated using the following equation:

$$P_{ij} = \frac{n_{ij}}{n_i} \tag{2.2}$$

Where n_{ij} is the number of times the land use changed from state i to j; n_i is the number of times that land use category a_i occurred.

Combining all the transition probabilities between all states $a_1, a_2, ..., a_m$ gives the following transition matrix:

$$\mathbf{P} = (P_{ij}) = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & & \vdots \\ P_{m1} & P_{m2} & \cdots & P_{mm} \end{bmatrix} \qquad (2.3)$$

According to the Chapman-Kolomogorov equation, the n steps transition matrix can be easily deduced from the one step transition matrix:

$$\mathbf{P}^{(n)} = \mathbf{P}^{n} \qquad (2.4)$$

Therefore, the land use state of a period after n steps can be easily predicted.

In markov chain analysis, the transition probabilities are estimated as proportions of cells that have changed state from one point in time to another. It is a useful way of estimating these probabilities despite the development of procedures for estimating transition probabilities on the basis of more complex scientific consideration. However, the Markov chain model lacks explanatory power as the causal relationships underlying the transition studies have been left unexplored.

2.3.4 Logistic Regression

In statistics, regression analysis is the collective term for techniques used for modeling and analyzing numerical data consisting of values of a dependent variable and of one or more independent variables. In the aforementioned process, the dependent variable is also called the response variable or measurement and the independent variable is called the explanatory variable or predictor. The dependent variable in the regression equation is modeled as a function of the independent variables, corresponding parameters ("constants"), and an error term. The error term is treated as a random variable. It represents the unsubstantiated variation in the dependent variable. The parameters are estimated to give a "best fit" of the data. Typically, the best fit is evaluated using the least squares method, but other criteria have also been used. For land use change analysis, the logistic regression model is widely applied to estimate the outcomes of a categorical dependent variable, while the independent variables can be a combination of both continuous and categorical variables. It is a suitable approach to estimate the coefficients of casual factors from the observation of land use change, as the determinants of land use change usually consist of both continuous and categorical variables.

Considering urbanization, for instance, a collection of m independent variables will be denoted by the vector $X' = (x_1, x_2, ..., x_m)$. Let the conditional probability that the conversion of rural lands to urban and other built-up uses is present be denoted by $P(y=1 \mid X)$, or simply $P(X)$. The linear logistic model is represented as follows:

$$\operatorname{logit}\left[P(X)\right] = \ln\left(\frac{P(X)}{1-P(X)}\right) = \alpha + \beta' X + \varepsilon \tag{2.5}$$

where α is the intercept, $\beta = (\beta_1, \beta_2, \cdots, \beta_m)'$ is the corresponding factor coefficient vector, and ε is the independent error term, that is, $E(\varepsilon_i, \varepsilon_j) = 0$.

To fit the logistic regression model in Equation (2.5) to a set of data requires that the value of α and β_i s, the unknown parameters, are estimated. Unlike ordinary linear regression, the method of ordinary least squares (OLS) estimation cannot be applied to a model with a dichotomous outcome. Instead, maximum likelihood techniques are used to maximize the value of a function, the log-likelihood function. The work by Hosmer and Lemeshow (1989) provides more detailed information. Finding the solution of the likelihood equations entails special software, which may be found in several packaged programs. Different combinations of explanatory variables are available for the regression model, such as stepwise regression, "best subset" models, and predefined conceptual models.

Earlier studies considered the spatial dependence of the land use data when employing logistic regression to model land use change analysis. This can generally be done by building a model that incorporates an autoregressive structure (Anselin, 1988). For example, a spatial lag model can be used as an extension of the traditional linear regression model:

$$y = a + \sum_{i=1}^{m} x_i b_i + \rho W y + \varepsilon \tag{2.6}$$

where ρ is a coefficient on the spatially lagged dependent variable and W is a spatial weight matrix. It should be noted that the maximum likelihood estimator (MLE) is usually employed to solve the parameters of such a model that best fits the data:

$$L = [y \ln(\frac{\exp(a + X\beta + \rho W y)}{1 + \exp(a + X\beta + \rho W y)} - (1 - y) \ln(1 + \exp(a + X\beta + \rho W y)] \tag{2.7}$$

The likelihood can be maximized using a simplex unvaried optimization routine (LeSage, 1998). The logistic version of Model (2.6), which incorporates spatial autocorrelation, called the Spatial Autologistic Regression model has been devised (Dubin 1995; Dubin 1997; LeSage, 1998). . This model, also called Spatial AutoLogit (SAL) model, has been proven to be effective in regression involving spatial autocorrelation (Páez and Suzuki, 2001). Similar to the spatial lag model, a spatial error model (Anselin, 1988) can also be utilized for regression models involving spatial autocorrelation. Nonetheless, they are primarily used for diagnostic analysis rather than for extrapolation-like prediction (Jetz *et al.*, 2005).

Logistic regression improves gradually and allows the causal factors to be a mixture of continuous and categorical variables, which suits land use change data well. However, the assumption of logistic regression, such as uncorrelation of error part, is not always valid. If the assumption cannot be satisfied, the generalization performance of logistic regression may degrade drastically.

2.3.5 Other models

In addition to the aforementioned four types of models, there are also some other models, such as Expert Models, Evolutionary Models, etc.

2.3.5.1 Expert Models

Expert models combine expert judgment with nonfrequentist probability techniques such as Bayesian probability, Dempster-Schaefer theory (Eastman 1999), or symbolic artificial intelligence approaches such as expert systems and rule-based knowledge systems (Gordon and Shortliffe 1984; Lee *et al.* 1992). These methods express qualitative knowledge in a quantitative fashion that enables the modeler to determine where specified land uses are likely to occur. However, it can be difficult to incorporate all aspects of the problem domain, which leaves room for gaps and inconsistencies.

2.3.5.2 Evolutionary Models

Within the field of artificial intelligence, symbolic approaches such as expert systems are complemented using the biologically inspired evolutionary paradigm. Artificial neural networks and evolutionary programming have already found their way into LUCC models (e.g., Balling *et al.* 1999; Mann and Benwell 1996). In brief, neural networks are silicon analogs of neural structure that are trained to associate outcomes with stimuli. Evolutionary programming mimics the techniques of Darwinian evolution by employing computational algorithms/programs that iterate over many generations to generate the solutions for a problem.

2.4 Spatial-temporal models of land use change

Land use change is essentially a spatio-temporal process. It is thus reasonable to conduct land use change study within a spatio-temporal framework. Recent studies in land use change analysis have attempted to incorporate spatial and temporal effects in their statistical models. An and Brown (2008) pioneer in studying the concepts and models in survival analysis, and their potential applications in land change science. Their model seeks to determine the optimal switching time and considers the option value of a choice. It could avoid the assumption of independence from irrelevant alternatives and make explicit use of the time dimension. While survival analysis has proven to be effective in addressing temporal complexities, their model does not consider spatial non-sationarity that commonly exists in land use data. As their model employs vector data, there might be questions pertaining to the prediction of change patterns. This is especially true when there are heterogeneous changes even within the same land cell. Also, model validation and the improvement of the modeling accuracy are other factors give rise to concerns under such circumstances.

The land use model generated by Huang *et al.* (2009a) considers various factors, such as population density, slope, proximity to roads, neighboring land use, and their influence on land use change. This model accounts for spatial neighbors in the last cross-section and employs a modified exponential smoothing technique to produce a smoothed one from a series of bi-temporal spatial models for different time periods. While spatial and temporal autocorrelation is partly considered in this model, it does not identify the spatio-temporal non-stationarity and individual effect in each cell.

Alessandro and Gerald (2009) provide a dynamic approach to supporting land use change modeling. In their discrete dynamic model, land use is assumed to be the result of an ongoing optimization process. At each location in time the agent chooses a land use with the highest expected net present value. Compared with other models of land use, the irreversibility of some decisions, expectations about future prices, and forward-looking behavior of an agent could all be accounted for by this approach. Besides, this approach could improve upon the existing models in terms of prediction accuracy. However, spatial dependencies and feedbacks are ignored by this approach.

Panel data model, typically handling all the time series and cross-sectional data, can be exploited for spatio-temporal analysis (Greene, 2000; Wooldrige, 2002). The grid cells in the panel data involve a minimum of two dimensions (a cross-sectional dimension, indicated by subscript *i*, and a time series dimension, indicated by subscript *t*). The formula of the panel data model is:

$$y_{it} = X_{it}^{\prime}\beta + u_{it}, i = 1,2,\cdots,N, t = 1,2,\cdots,T \tag{2.8}$$

where X_{it} and β are the vectors of explanatory variables and parameters respectively, and u_{it} is an error component whose assumption is different from the classic assumption in linear regression. Based on the components of u_{it}, panel data could be thought of as having fixed effect and/or random effect. The key advantage of panel data modeling is the possibility of pooling information from all series. Thus, this advantage is related to generating more accurate predictions for individual outcomes by pooling the data rather than using the data on the individual being considered. However, spatio-temporal non-stationariy and spatio-temporal autocorrelation have seldom been considered in the panel data framework for land use change analysis.

2.5 Chapter Summary

This chapter summarized the causal factors that influence land use pattern and the change reported in the literature. Also, a variety of techniques for land use change modeling were briefly reviewed. Especially, the spatio-temporal model of land use change was introduced. Although many land use change models have been developed, Parker *et al.* (2003) conclude that none of approaches yet "dominates this nascent field".

Owing to the complexity of the land development process and the differences in modeling objectives, an obviously superior approach was hard to find. Each method had its strengths, weaknesses, and particular application domains. Consequently, selecting the methods for land use change modeling should depend on the demands of the analysis, the feasibility of the techniques, and the availability or limitation of the data framework.

Recent research focuses especially on the spatial-temporal aspect of land use change analysis. Some notable challenges have been introduced by the temporal dimension. With due consideration to the pros and cons of the above-mentioned models, the logistic regression framework has been selected in this research. Despite the fundamental assumptions problem, the regression model is still considered to be a powerful tool for land use change analysis. This framework can easily incorporate spatio-temporal non-stationarity, spatio-temporal autocorrelation, and individual effect.

Chapter 3: Analysis of Land Use Change in Special Economic Zone, Shenzhen Using Multinomial Logit Model

Earlier studies may have established the relationship between land use change and causal factors in a straightforward manner. Issues still exist in studies involving 'multi-class change' by using a multinomial logit model. Errors may be caused by the inadequacy of some kind of land use change in the estimation process. However, the study of the relationship between land use distribution and causal factors may be complemented by using multi-temporal data.

3.1 Study Area

Shenzhen City, which is one of the important cities of the Guangdong province, China, is located between 113.46 and 114.37 degrees east, and its latitude is between 22.27 and 22.52 degrees north. The total area is approximately 1,952.84 square kilometers. As seen from figure 3.1, it lies on the eastern edge of the Daya and Dapeng Bay, west of the Pearl River, northeast of Dongguan and south of Hong Kong.

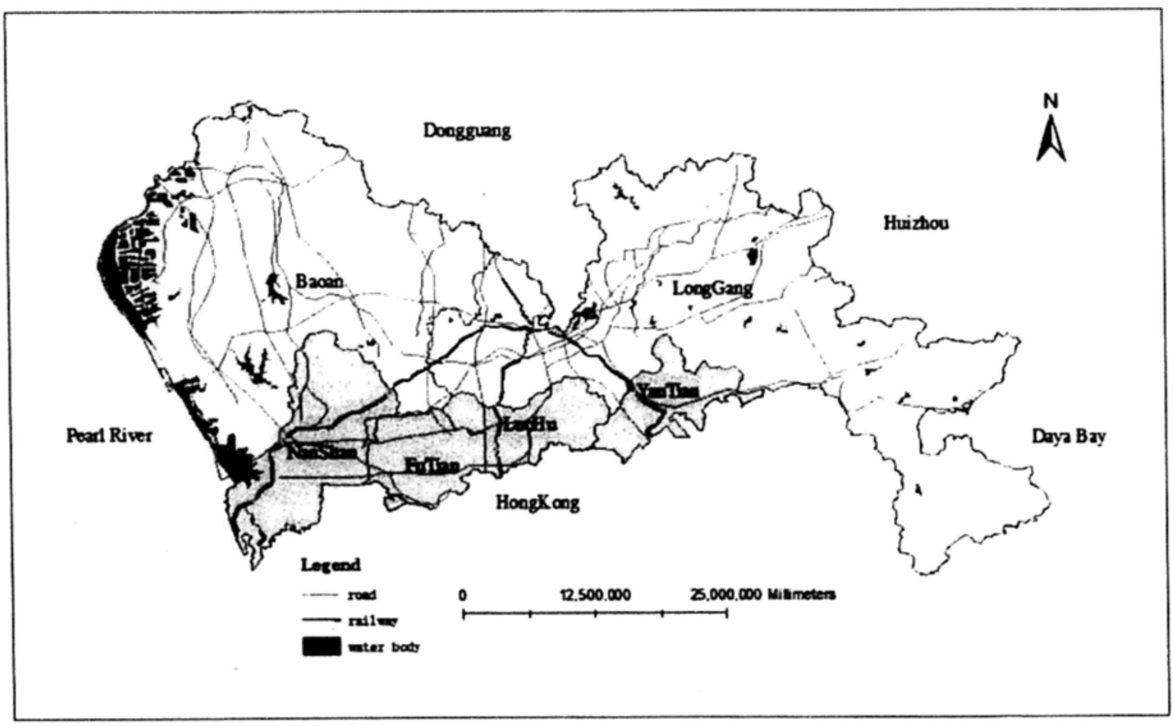


Figure 3.1: A summary map of the study area

Shenzhen consists of six administrative districts: Luohu, Futian, Nanshan, Yantian, Bao'an and Longgang. In May 1980, the first special economic zone (SEZ), consisting of Luohu, Futian, Nanshan, and Yantian (but not Bao'an and Longgang) was established in Shenzhen by the Chinese Government. Luohu, which is the financial and trading centre, locates in the centre of the SEZ and adjacent to Hong Kong. It covers an area of 78.89 km^2. Futian, where the Municipal Government is situated, is at the heart of the SEZ and covers an area of 78.04 km^2. Covering an area of 164.29 km^2, Nanshan is the centre for high-tech industries and is situated in the western part of the SEZ. Yantian (75.68 km^2) is well-known for logistics. Yantian Port is the second largest deepwater container terminal in China and the fourth largest in the world. Outside the SEZ, Bao'an (712.92 km^2) and Longgang (844.07 km^2) are located to the north-west and north-east of Shenzhen respectively.

As a gateway to the world for China, Shenzhen is one of the most developed cities in China and is a city that grew at a tremendous pace. In less than 30 years, Shenzhen, a tiny border town of 30,000 people in 1979, has grown into a modern metropolis. By the end of 2008, the permanent resident population reached 8.7683 million, among which 2.2807 million (26%) had a permanent hukou and 6.496 million (74%) were non-hukou households. The GDP has increased from 0.196 billion in 1979 to 780.7 billion in 2008. Shenzhen ranks first among major cities in China in terms of per capita GDP. The economic and population growth in Shenzhen has been accompanied by rapid land use change over the past two decades. As a result, Shenzhen has experienced rapid rate of urbanization, almost 80% now. The urbanization indicates a transformation of vacant and agricultural land use to construction of urban fabrics including residential, industrial, commercial, and transportation developments.

Due to the 'one city, two system' framework in Shenzhen, there is a big difference between SEZ and the region outside of SEZ. Most of the financial expenditure occurs in the SEZ. Also the urban planning can be better performed in SEZ. However, the situation is bound to change in the future. In 2009, a new reform plan has been approved by the central government. According to this plan, the special economy will be extended to Bao'an and Longgang. Hence, a careful study of land use change in SEZ will benefit the development of the new SEZ.

3.2 Land Use Change in Shenzhen

As mentioned already, Shenzhen is the fastest growing city in China, with the land use experiencing a dramatic change. The primary type of land use change is from non built-up to built-up. As can be observed from figures 3.2, 3.3, and 3.4, the built-up area in Shenzhen has increased from 41613.24 ha to 92050.32 ha in 18 years. Most of this change has occurred in Bao'an and Longgang (outside the SEZ). In these two districts, the built-up area has increased from 28269.32 ha to ha 71383.28. The increase in the built-up areas in SEZ is only 7323.12 ha. That's because most of land in SEZ has already been used. As can be seen from figure 3.2, the curve representing the increasing trend of built-up area in SEZ has become flattened. The land use change, which may happen in residential, industrial, commercial and transportation, is more complex. This study is performed in SEZ and will take into account the multi-class land use change.

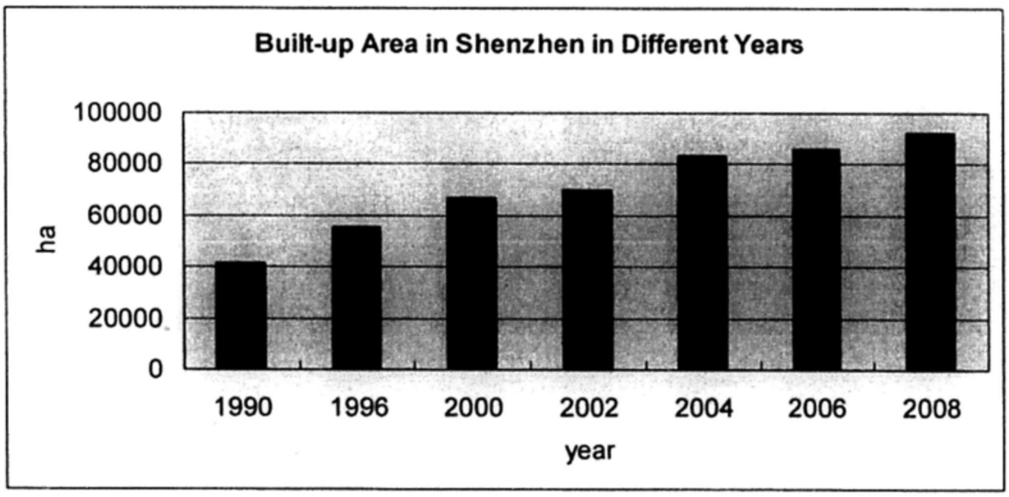


Figure 3.2: Built-up area in Shenzhen in different years (Source: urban planning, land and resources commission of Shenzhen Municipality).

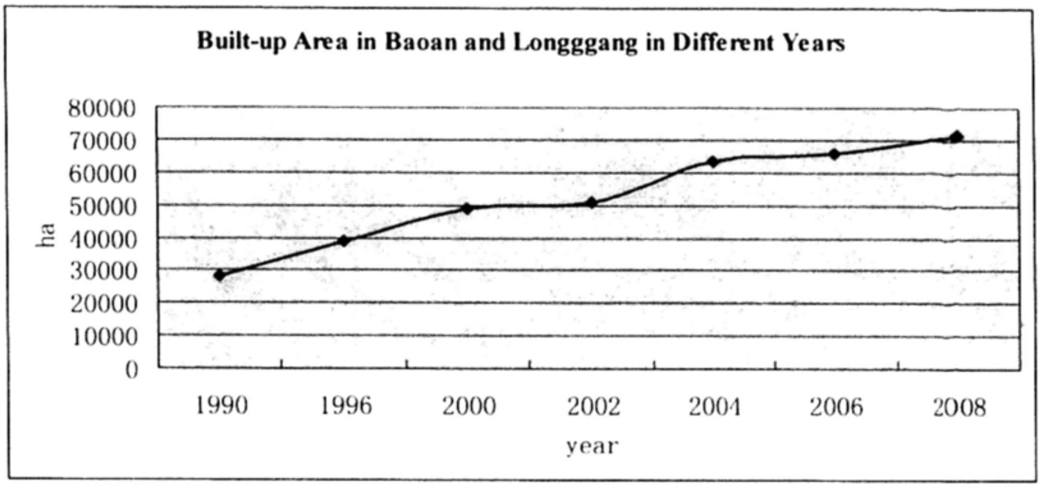


Figure 3.3: Built-up area in Baoan and Longgang in different years (Source: urban planning, land and resources commission of Shenzhen Municipality).

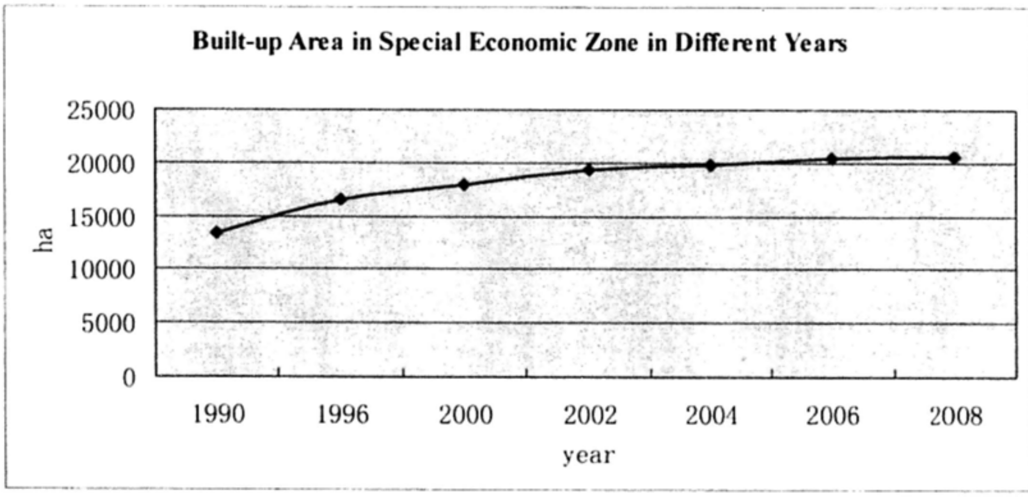


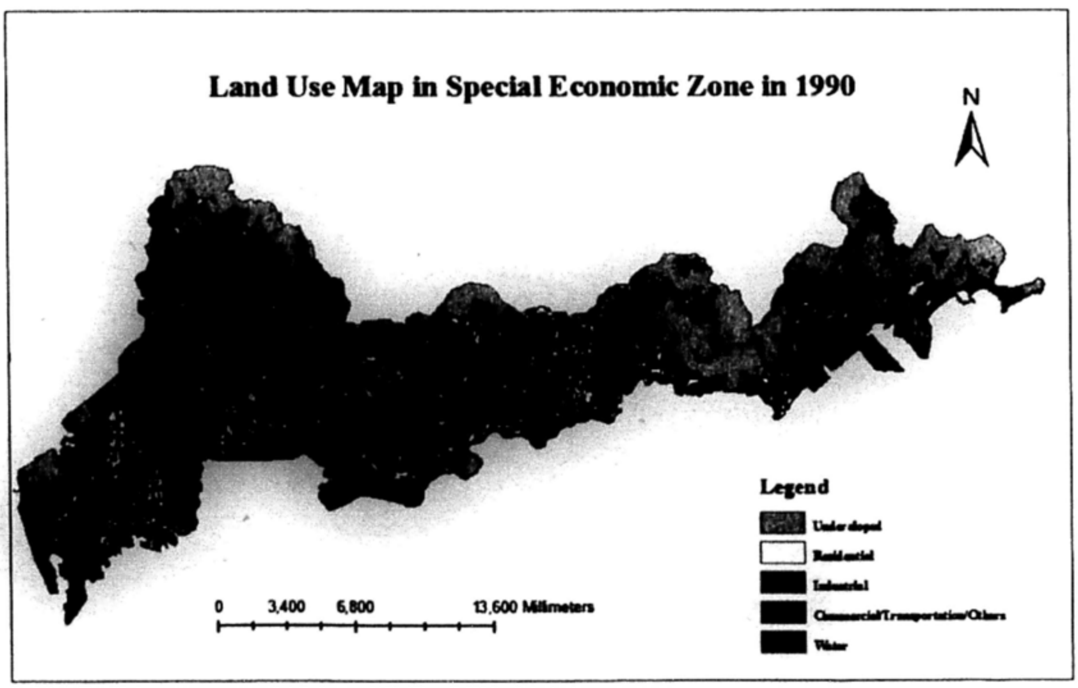
Figure 3.4: Built-up area in special economic zone in different years (Source: urban planning, land and resources commission of Shenzhen Municipality).

3.3 Data Preparation

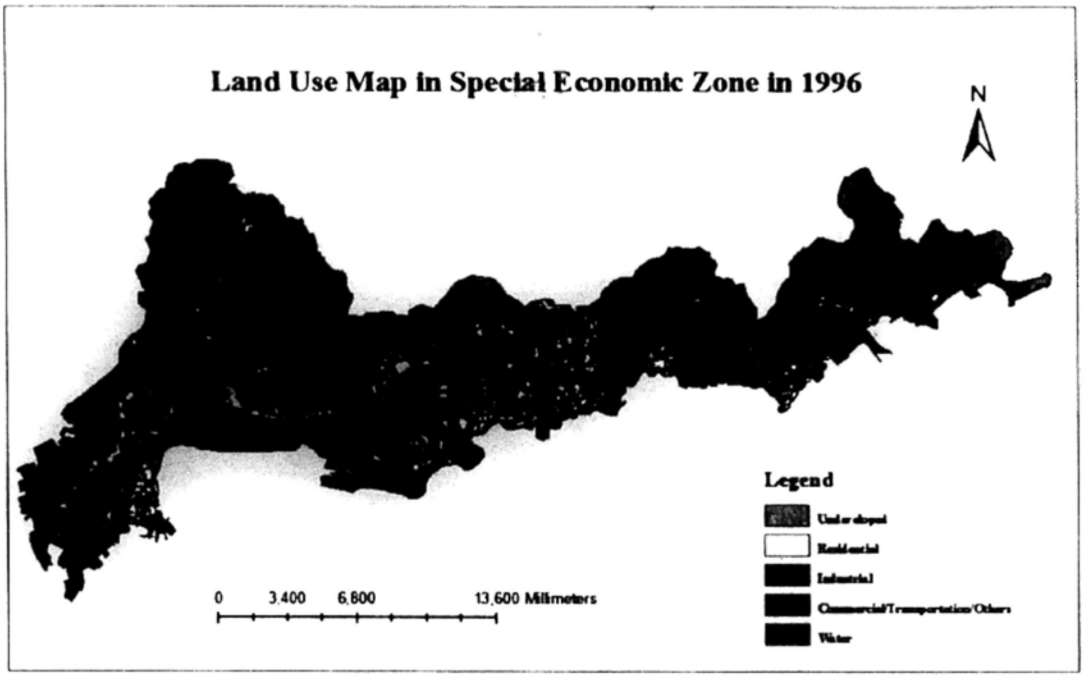
The focus of this study is to analyze the relationship between explanatory (or independent) and response (or dependent) variables. Thus, the data in this study is twofold: one is response data: land use data; the other is explanatory variables that are the driving factors considered as influencing the land use.

3.3.1 Land Use Data

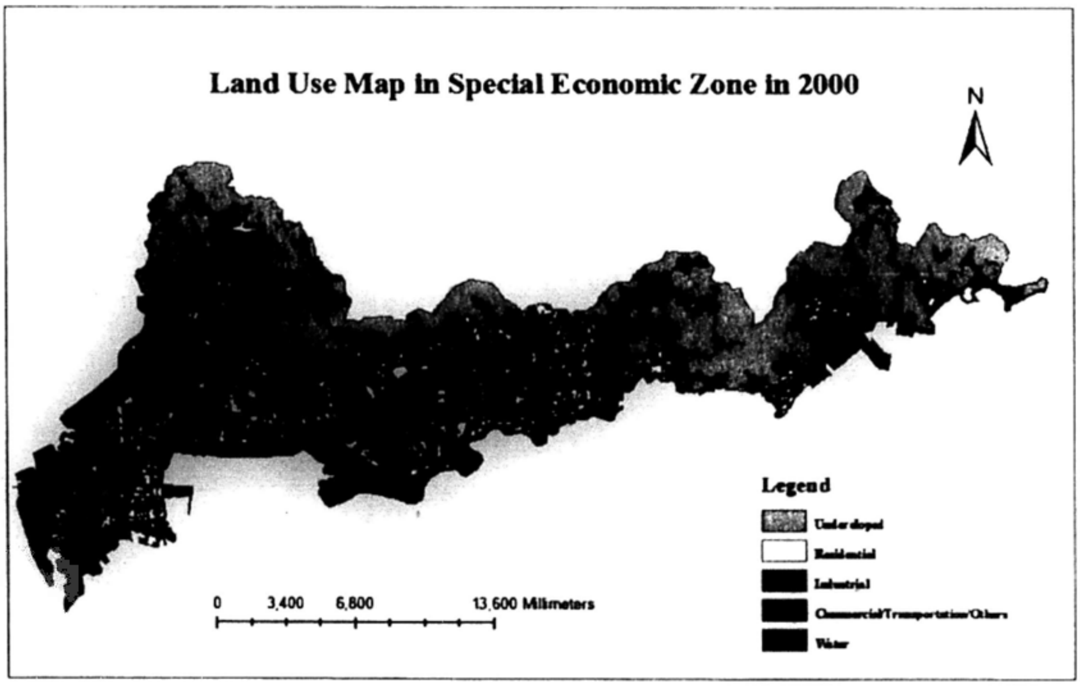
Land use maps for the years 1990, 1996, 2000, 2002, 2004, 2006 and 2008 were acquired from land investigation and digital orthophotos obtained from the urban planning, land, and resources commission of Shenzhen Municipality. Each land cell is classified into one of the eleven land use types: farmland, woodland, garden Land, grass land, residential, industrial, commercial, financial, education, transportation and water land use type. Since the land use of commercial, financial, and education is too small in their proportions, the eleven land use types were aggregated into four categories in order to focus on the human development of land. These four categories include undeveloped, residential, industrial, and commercial/transportation/others land use type (excluding water). Figures 3.1 (a) ~ (g) represent the land use maps corresponding to 1990, 1996, 2000, 2002, 2004, 2006 and 2008 respectively. All maps have the same size of 1801 columns by 1002 rows and a common spatial resolution of 50m. All the maps follow the same legend: green is for undeveloped, yellow is for residential, purple is for industrial, red is for commercial/transportation/others, blue is for water, and white is for 'no data'. It can be found from Figure 3.5 that most of the industrial land use is concentrated in Nanshan district while undeveloped land use types are prominent in Yantian district. The residential and commercial/transportation/others land use are distributed randomly in Luohu, Futian and Nanshan. Water bodies are treated as constraints that are inappropriate for development.



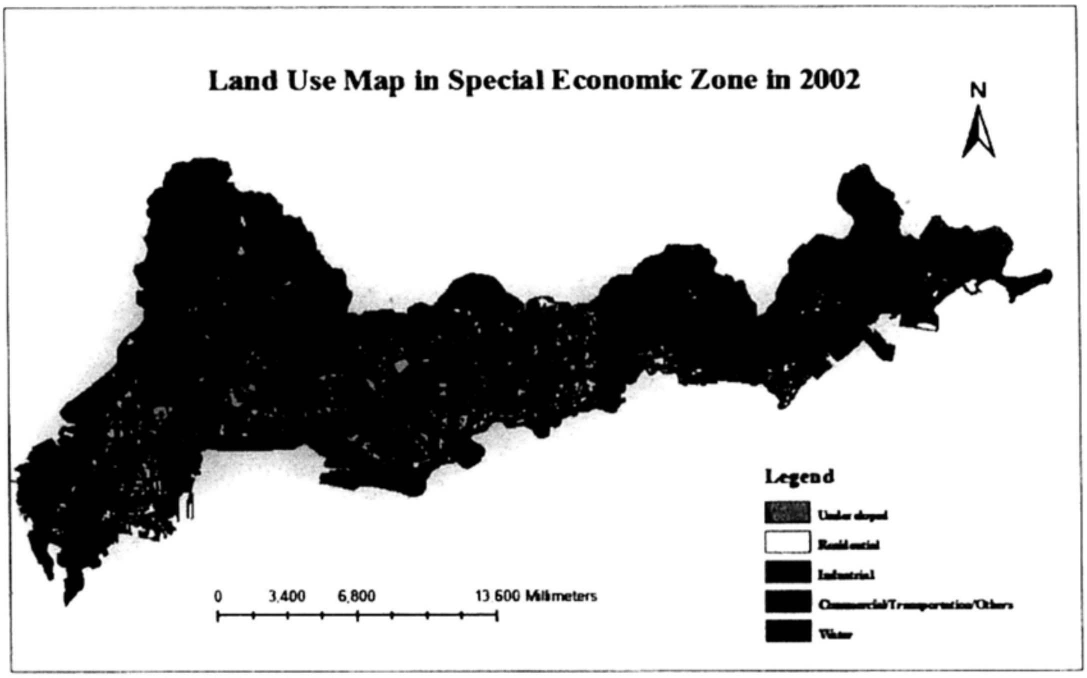
(a) 1990


Land Use Map in Special Economic Zone in 1996
N
Legend
Residential
Industrial
Commercial/Transportation/Others
Water
0 3,400 6,800 13,600 Milimeters

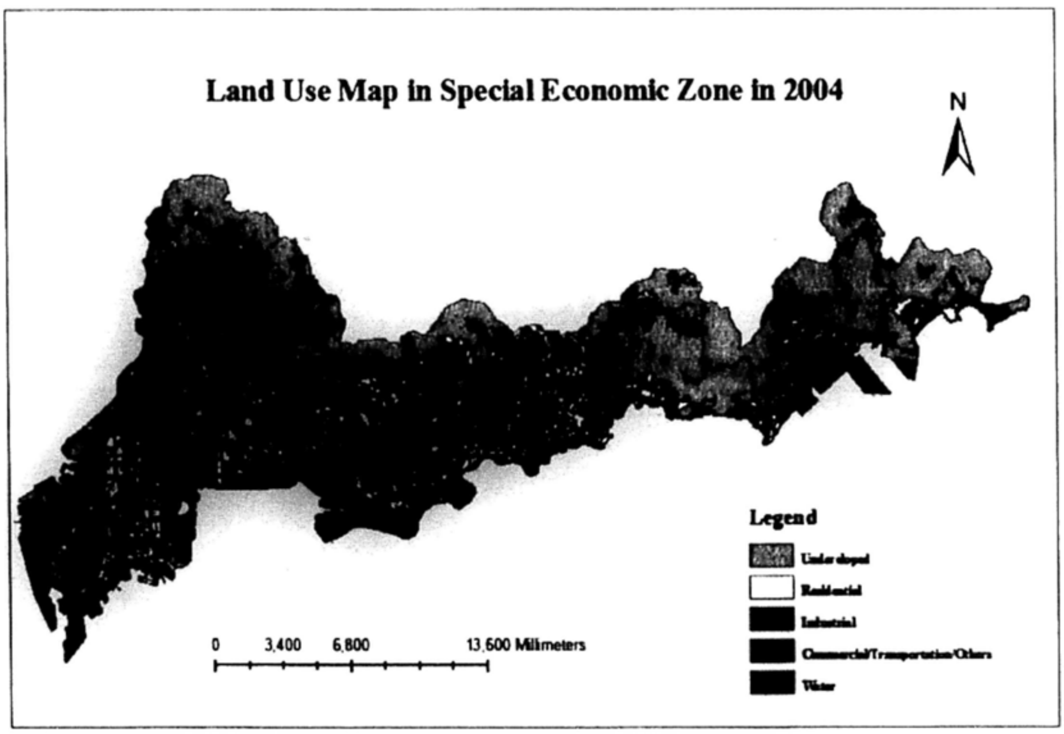
(b) 1996


Land Use Map in Special Economic Zone in 2000
N
Legend
Residential
Industrial
Commercial/Transportation/Others
Water
0 3,400 6,800 13,600 Milimeters

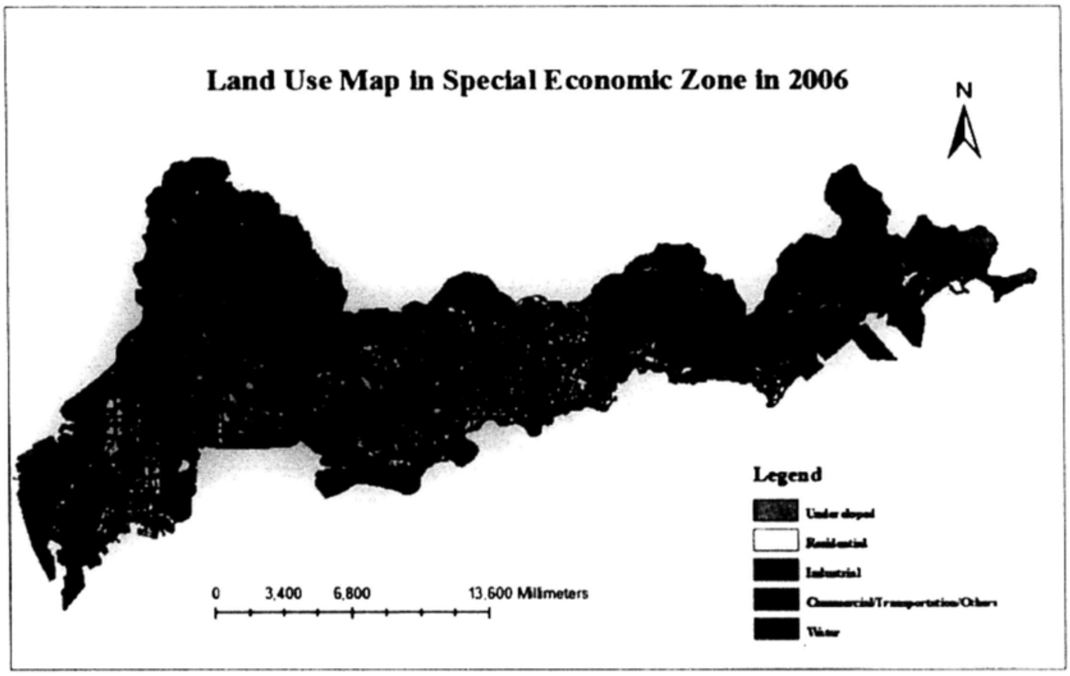
(c) 2000


Land Use Map in Special Economic Zone in 2002
N
Legend
0 3,400 6,800 13,600 Milimeters

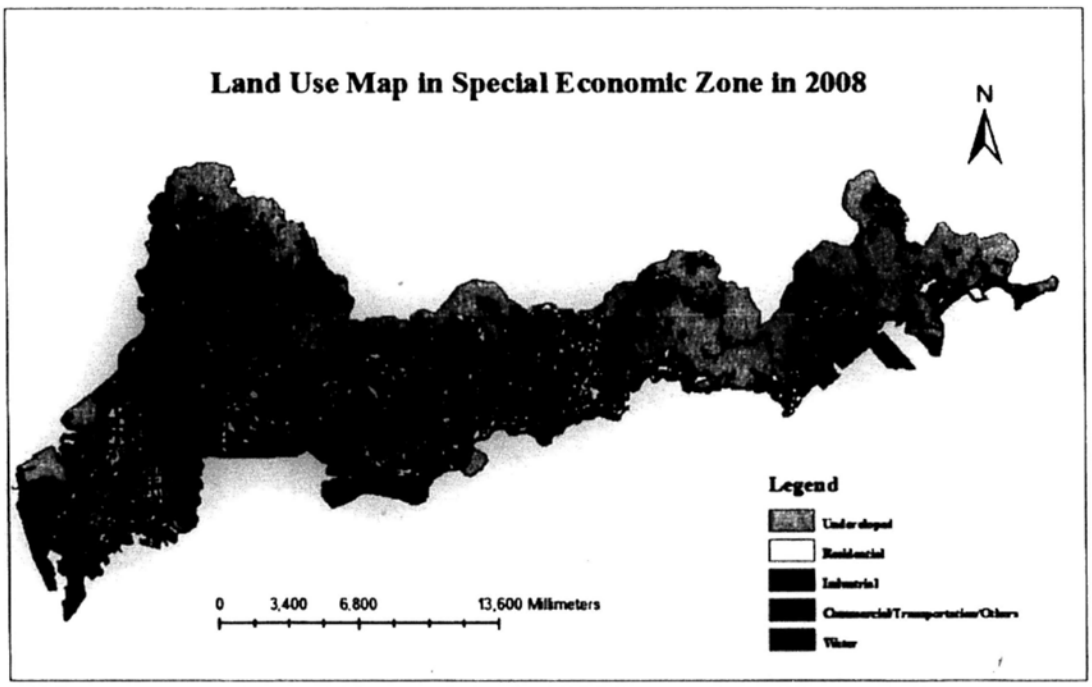
(d) 2002


Land Use Map in Special Economic Zone in 2004
N
Legend
0 3,400 6,800 13,600 Milimeters

(e) 2004



(f) 2006



(g) 2008

Figure 3.5: Special economic zone land use maps (Source: urban planning, land and resources commission of Shenzhen Municipality).

3.3.2 Causal Factors

Considering previous literature, the context of SEZ, and the data availability, this study includes the following causal factors data in the proposed model: elevation data, transportation data (major roads and rail roads), commercial map, financial map, industrial map, educational facilities map, demographic data and planning map (1996-2010). All these shape files/raster layers were compiled in ESRI ArcMap v9.1®. All layers were projected to the *WGS_1984_W114* coordinate system. Raster layers were re-sampled using a cell size of 50 meters to be consistent with other land use data.

Slope raster was generated from the elevation data using the ArcMap spatial analyst extension. Sequential shape files for the road networks, rail lines/stations, commercial centers, financial centers, industrial centers, educational facilities were generated from land investigation data by feature selection. Subsequently, the ArcMap spatial analyst extension was used to generate distance raster layers to these utilities in each modeled year according to the Euclidian distance.

The demographic data corresponding to each town of Shenzhen was obtained from the statistical year book at the Shenzhen Library. In order to use the available demographic data, spatialization of demographic data was done by Dr. Li Hongga from Institute of Remote Sensing Applications, CAS. Then, spatial interpolation method: kriging was employed to obtain the population density distribution.

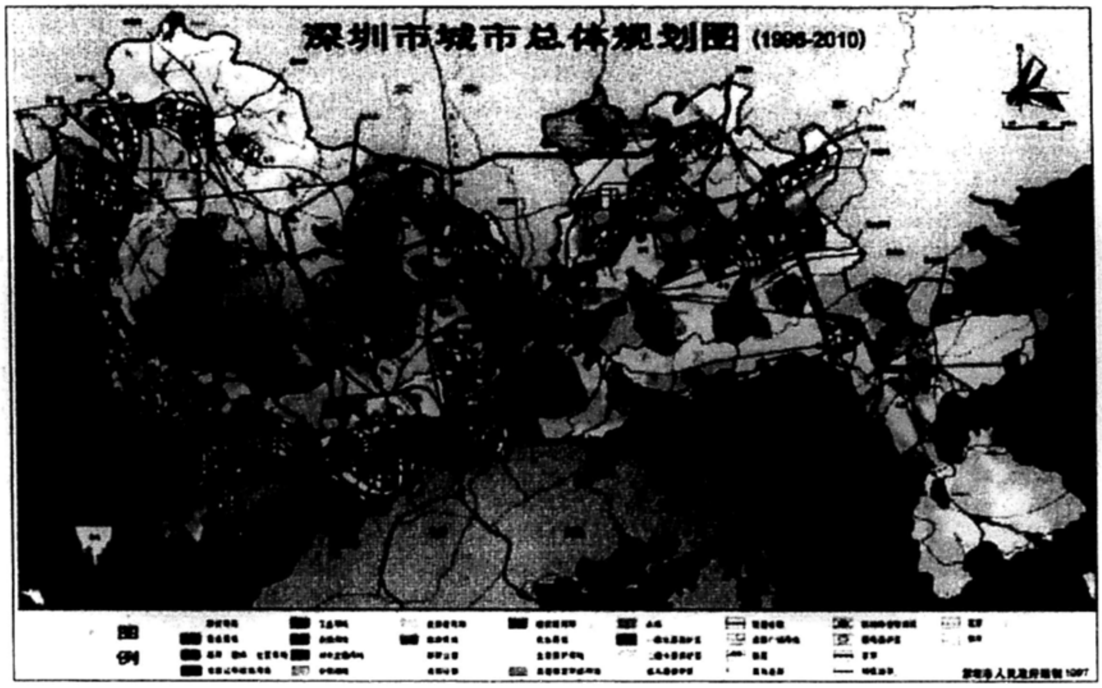


Figure 3.6: Shenzhen planning map (Source: urban planning, land and resources commission of Shenzhen Municipality).

The planning map (1996-2010) in the CAJ format was obtained from the urban planning, land and resources commission of Shenzhen Municipality. As can be found from figure 3.6, there are many layers in the CAJ map and the road network is not closed. Thus, the CAJ map could not be converted to shapefile directly. The land use layer for each land use type planning should be extracted from the CAJ map and the road work should be closed. The land use planning is also classified to five types in line with the land use map. Then land use planning map and road network planning can be made. Both the maps are shown in figure 3.7 and 3.8 respectively. For our research, the two maps are clipped to the SEZ shape (figure 3.9 and 3.10).

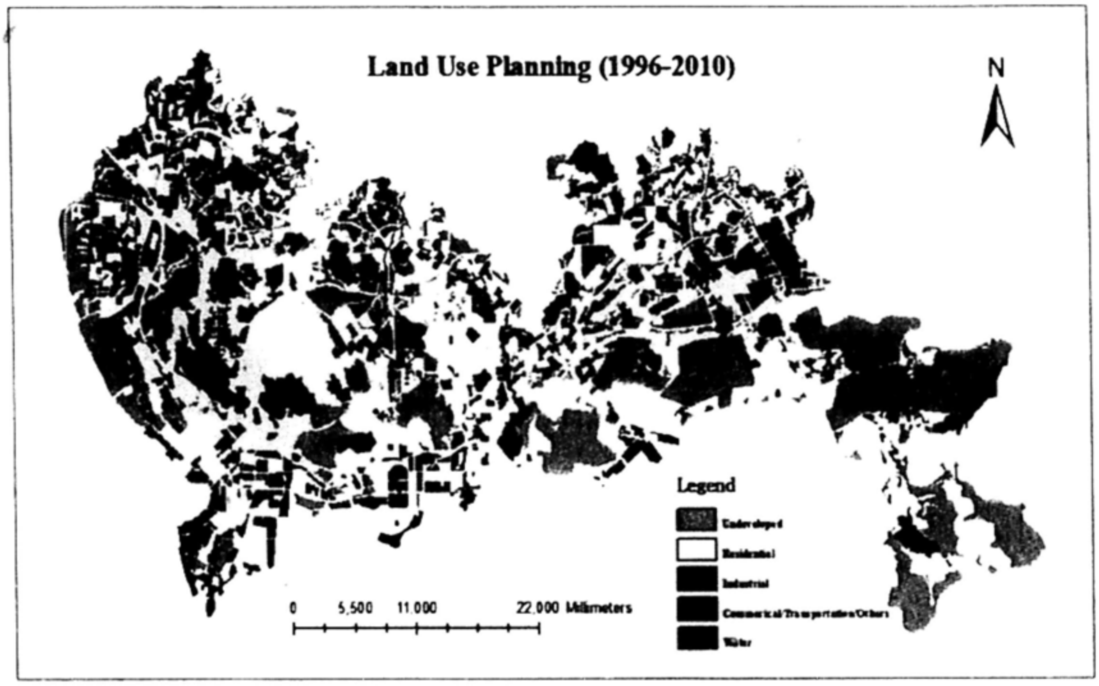


Figure 3.7: Land use planning in Shenzhen

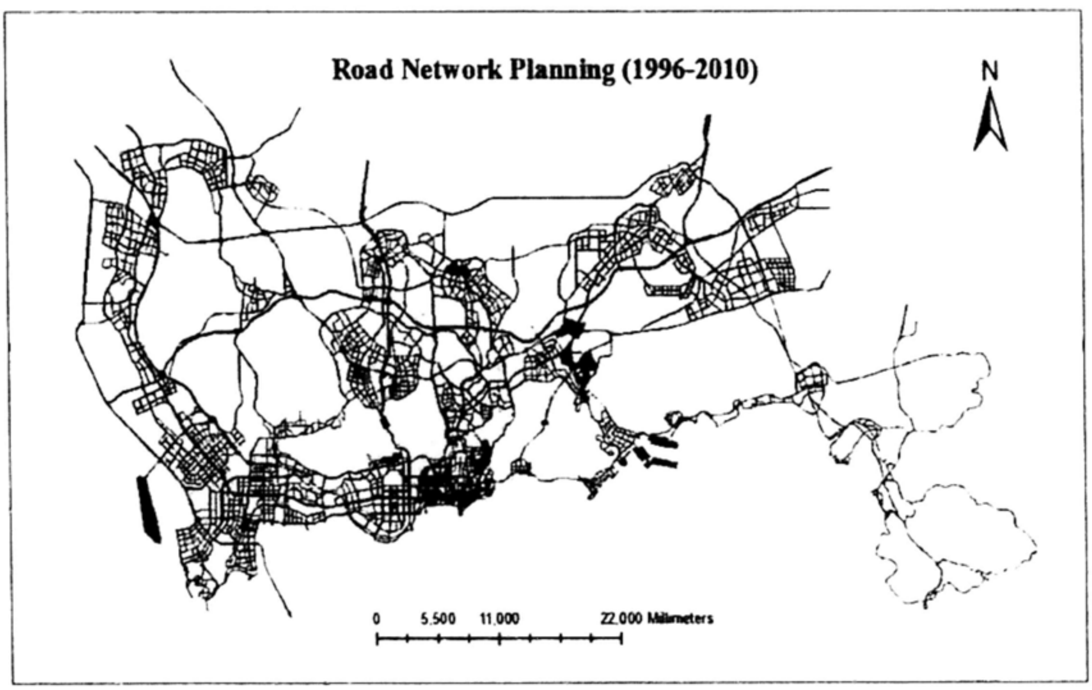


Figure 3.8: Road network planning in Shenzhen

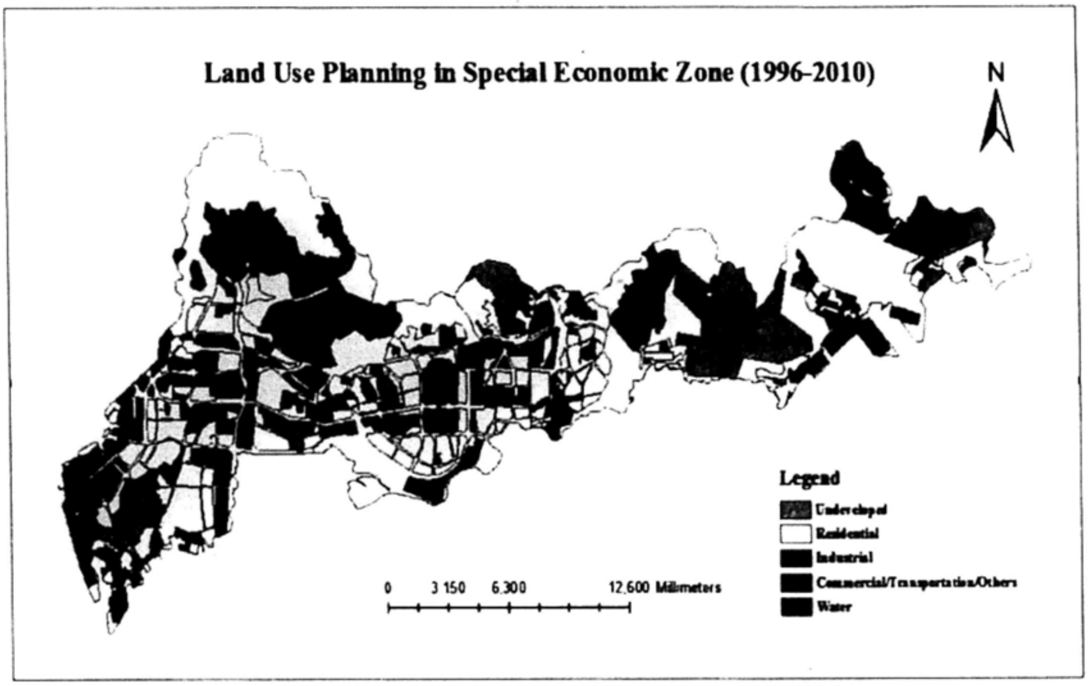


Figure 3.9: Land use planning in special economic zone

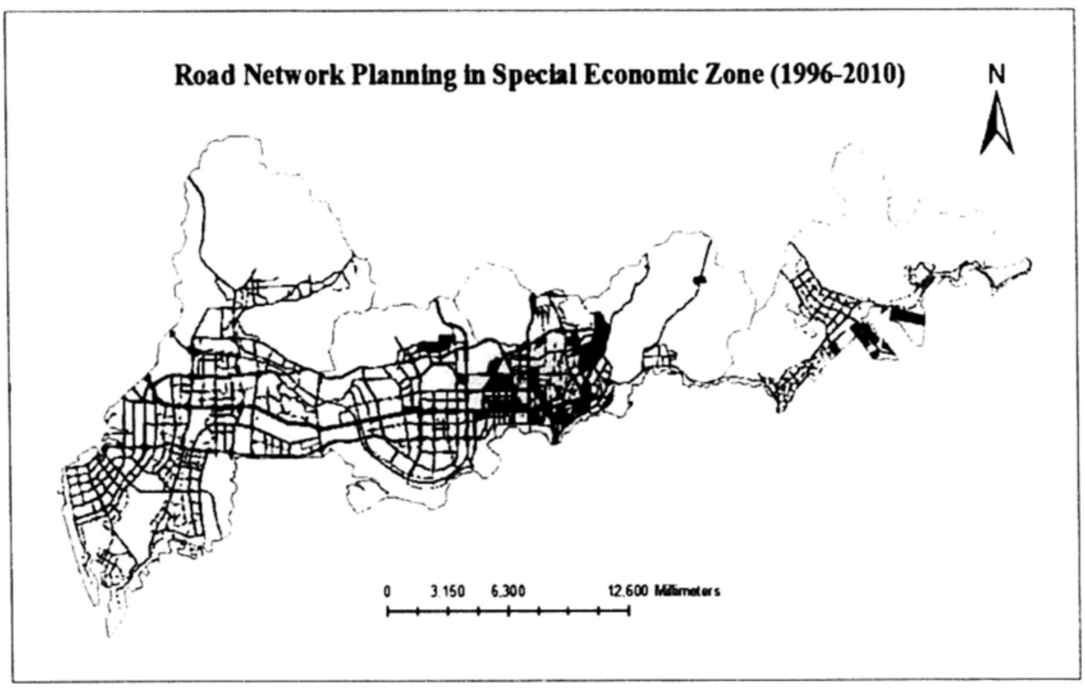


Figure 3.10: Road network planning in special economic zone

Overall, the framework of driving factors preparation is shown in figure 3.11.

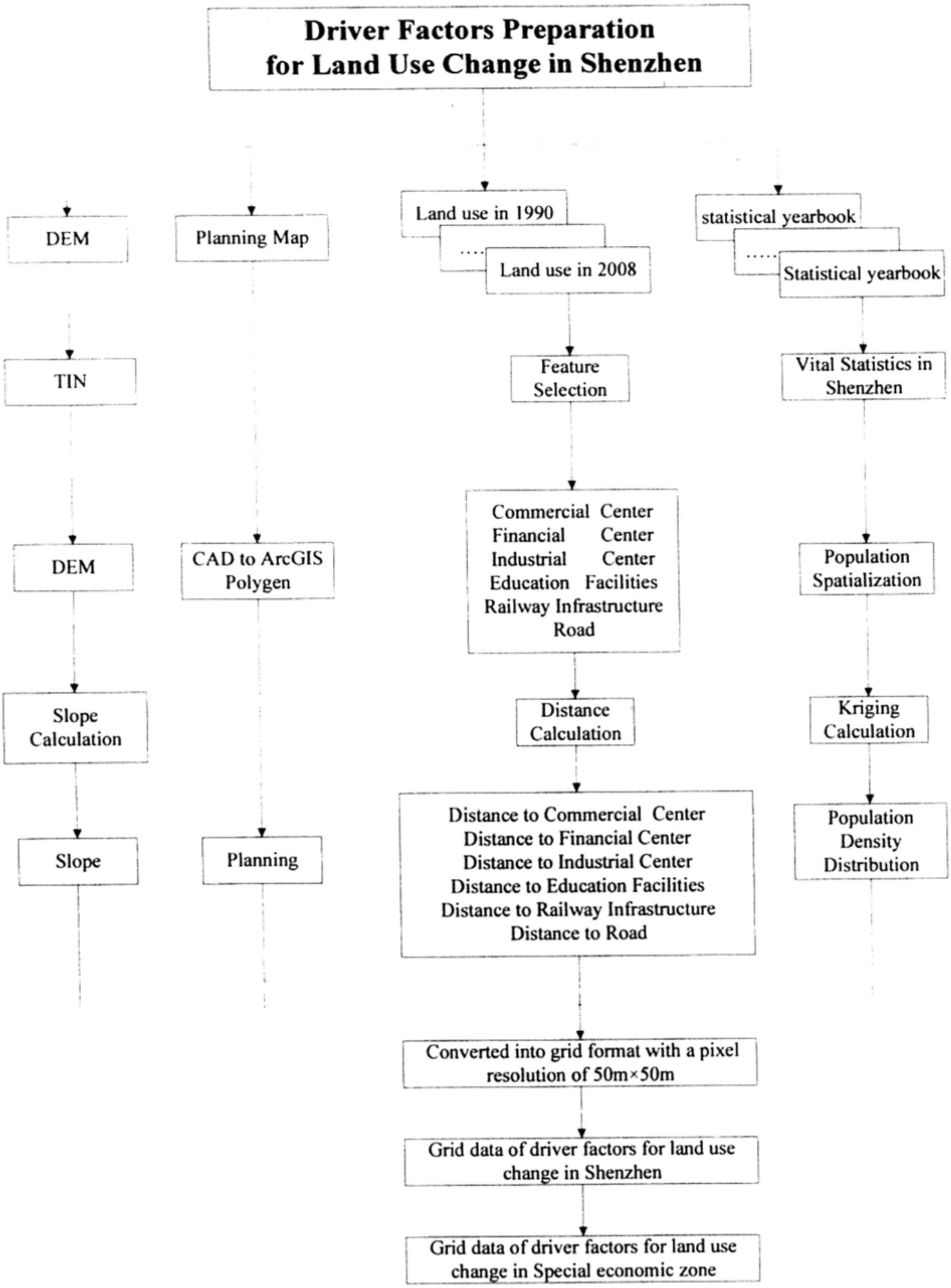


Figure 3.11: The framework of driver factor preparation for land use change in SEZ, Shenzhen

On the whole, ten causal factors were considered in this study. A summary of these factors is shown in Table 3.1.

Table 3.1: Summary of Explanatory Variables for the Land Use Change Model

Variables	Definition
Distance to Commercial Centre	Distance from the cell to the commercial centre
Distance to Financial Centre	Distance from the cell to the financial centre
Distance to Industrial Centre	Distance from the cell to the Industrial Centre
Distance to Educational Facilities	Distance from the cell to the nearest Educational Facilities
Distance to Railway Infrastructure	Distance from the cell to the Railway Infrastructure
Distance to Road	Distance from the cell to the nearest road
Population	Population density of the cell
DEM	DEM
Slope	Measurement of the degree of slope
Planning	Land use planning

3.4 Data Sampling

A random sampling method was employed in our study to obtain the samples from the study area for each year. The samples across different time periods were used to construct the land use change data set. Considering the heavy computational intensity, 700 samples were selected from different years, i.e., 1996, 2000, 2002, 2004, 2006 and 2008 with 1990 as the initial year (i.e., t =0). Thus, a total of 4200 samples were obtained. Figures 3.12 and 3.13 illustrate the sample distribution over different years and space respectively. Figure 3.14 provides the spatio-temporal distribution of the samples.

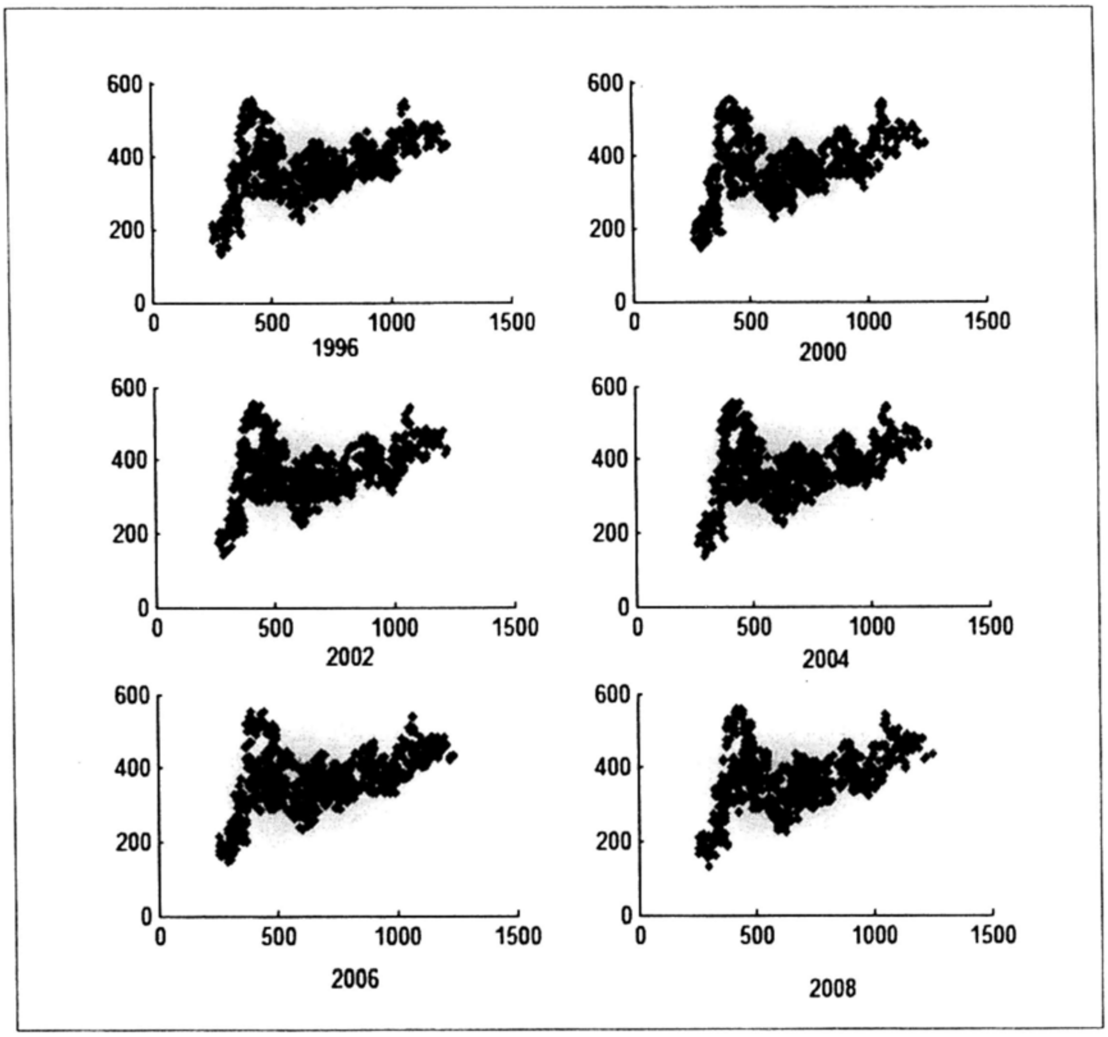


Figure 3.12: Distribution of Samples in Different Years

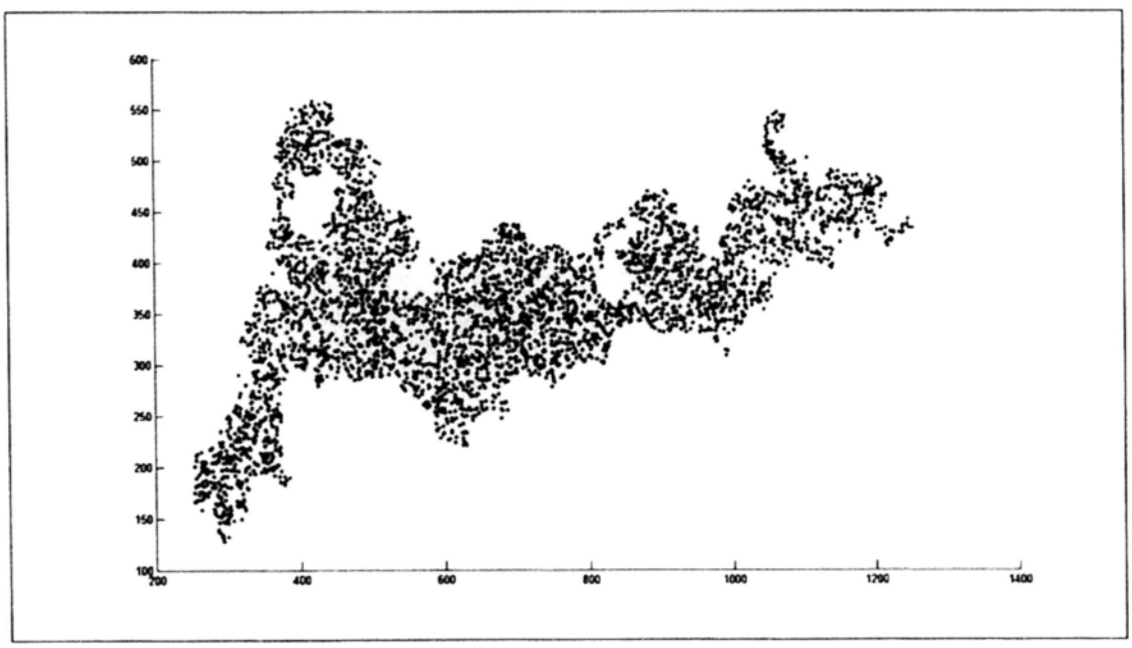


Figure 3.13: Distribution of samples in space

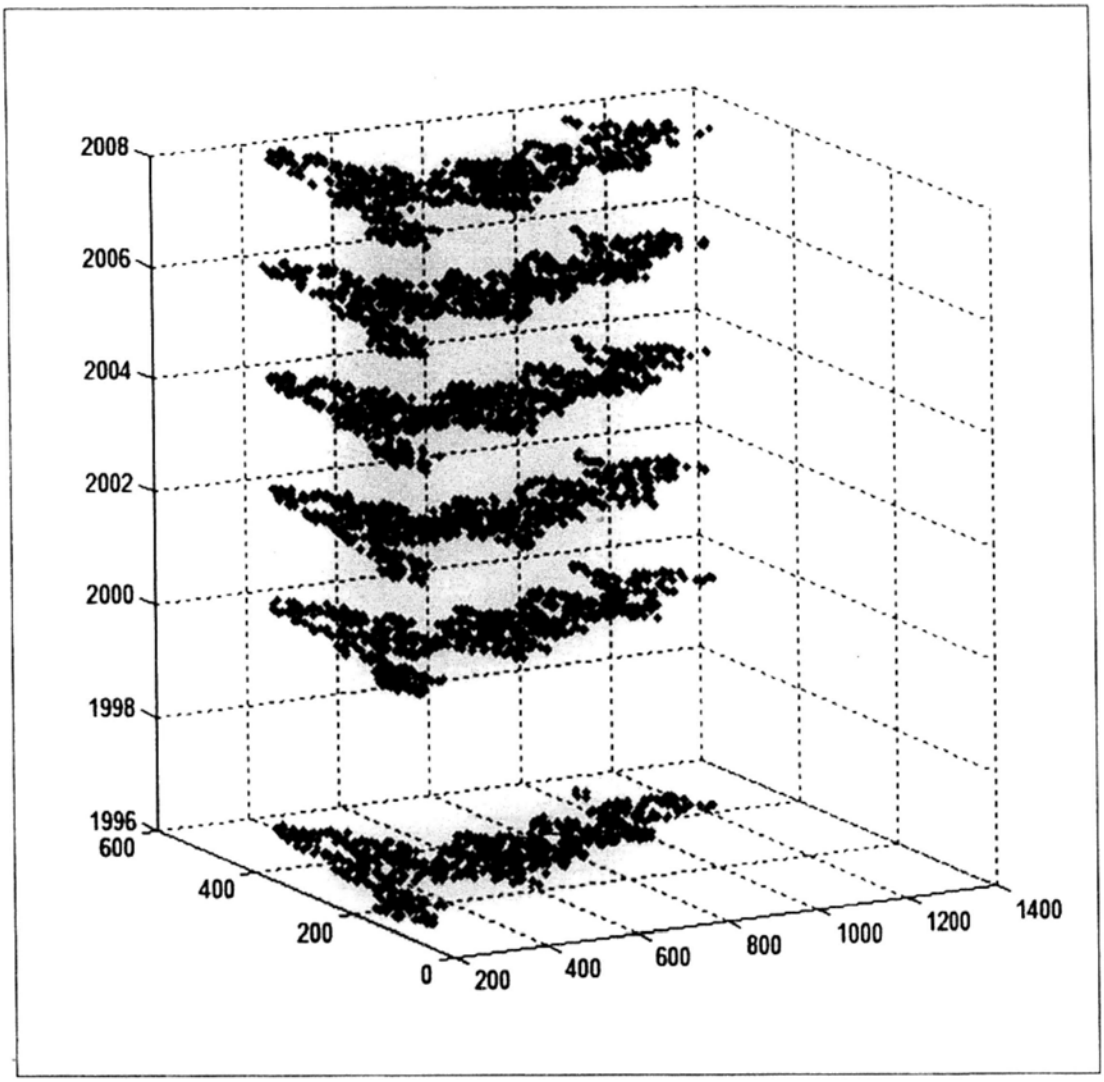


Figure 3.14: Sptaio-temporal Distribution of Samples

3.5 Methodology

3.5.1 Multinomial Logit Model (MNLM)

A multinomial logit model is used for data in which the dependent variable is discrete and unordered, and independent variables are continuous or categorical predictors. Unlike a binary logistic model, wherein a dependent variable has only a binary choice (e.g., persist/change), the dependent variable in a multinomial logit model can have more than two choices that are coded categorically, and one of the categories is assumed as the reference category.

Assuming that y_i, the dependent variable for the i^{th} observation, has $J+1$ categories and is indexed as j (j = 0,1,...,J). If y_i is j, let $y_i^j = 1$, otherwise $y_i^j = 0$. Let x_i denote the exploratory variables for the i-th observation. Category 0 is considered the reference category. The logits, which compare any category $j=1,...,J$ with the reference category, are described as follows:

$$\log\left(\frac{prob(y_i^j = 1)}{prob(y_i^0 = 1)}\right) = \log\left(\frac{\exp(\beta^j x_i)}{\exp(\beta^0 x_i)}\right) = \beta^j x_i,\ j = 1,...,J \tag{3.1}$$

where β^j is the coefficient vector of length p, corresponding to p covariates. β^0 is set to be zero. $\beta^1,\ldots,\beta^J$ are to be estimated. The parameter, β^j, represents the additive effect of a one-unit increase in the independent variable, x, on the log-odds of being in category j, rather than the reference category. An alternative way to interpret the effect of an independent variable, x_i, is to use predicted probabilities $prob(y_i^j = 1)$ for different values of x_i:

$$prob(y_i^j = 1) = \frac{\exp(\beta^j x_i)}{\sum_{j=0}^{J} \exp(\beta^j x_i)}, j = 1,...,J \tag{3.2}$$

Then, the probability of being in the reference category, "0" (*normal*), can be calculated by subtraction:

$$prob(y_i^0 = 1) = 1 - \sum_{j=1}^{J} prob(y_i^j = 1) \tag{3.3}$$

In this model, the same independent variable appears in each of the j categories, and $\beta^1,\ldots,\beta^J$, is usually estimated for each contrast.

Maximum likelihood method was employed to estimate multinomial logit model. Firstly, the probabilities were calculated by formulation 3.2 and 3.3 for each cell's category and time each other. Then, the parameters can be obtained by maximizing the product.

3.5.2 Evaluation of the Model's Classification Accuracy

3.5.2.1 Percentage of Correct Prediction (PCP)

The percentage of correct prediction (PCP), which measures the overall concordance between a classification and the actual land use type, is employed to assess the goodness-of-fit of the models. An efficient way to assess the goodness-of-fit of classification is utilized in this study, which cross tabulates predictions with observations and calculates the overall concordance. Table 3.2 provides an example of the cross evaluation table used in this study. 0, 1, 2 and 3 represent undeveloped, residential, industrial and commercial/transportation/others land use type respectively.

Table 3.2: An example of the cross evaluation table

Observed	Predicted				Total
	0	1	2	3	
0	Num(0-0)	Num(0-1)	Num(0-2)	Num(0-3)	Num(O-0)
1	Num(1-0)	Num(1-1)	Num(1-2)	Num(1-3)	Num(O-1)
2	Num(2-0)	Num(2-1)	Num(2-2)	Num(2-3)	Num(O-2)
3	Num(3-0)	Num(3-1)	Num(3-2)	Num(3-3)	Num(O-3)
Total	Num(P-0)	Num(P-1)	Num(P-2)	Num(P-3)	Num(Total)

Let us consider land use type 0 and 1 as an example. $Num(0-0)$ is the number of cells with a label of 0 (undeveloped) and is classified as 0 in table 3.2. $Num(0-1)$ is the number of cells with a label of 0 (undeveloped) but is classified as 1 (residential). $Num(1-0)$ is the number of cells with a label of 1 (residential) but is classified as 0. $Num(1-1)$ is the number of cells with a label of 1 (residential) and is classified as 1.

$Num(P-0)$, $Num(P-1)$, $Num(P-2)$ and $Num(P-3)$ are the number of cells with a label of 0, 1, 2 and 3 respectively. $Num(O-0)$, $Num(O-1)$, $Num(O-2)$ and $Num(O-3)$ are the number of cells that are classified as 0, 1, 2 and 3 respectively. $Num(Total)$ is the size of the training set/evaluation set. Based on the cross evaluation table, some important indicators can be calculated. For example:

$$PCP = \frac{Num(0-0) + Num(1-1) + Num(2-2) + Num(3-3)}{Num(Total)} \tag{3.4}$$

$$PCPo0 = \frac{Num(0-0)}{Num(O-0)} \tag{3.5}$$

$$PCPp0 = \frac{Num(0-0)}{Num(P-0)} \tag{3.6}$$

PCP indicates the overall accuracy of the classifier. $PCPo0$ is the probability that the land use category 0 can be accurately predicted by the model. It reveals the capacity of the classifier to detect the land use categorized as 0. The greater the value of $PCPo0$, the more the land use 0 can be correctly predicted. $PCPp0$ is measured by the percentage of correctly predicted land use 0 over the total land cells classified as 0. It exhibits the efficiency of the classifier in detecting the land use category 0. A higher $PCPp0$ implies that the model can predict the land use 0 with a higher accuracy.

3.5.2.2 Kappa

Kappa, which is usually attributed to Cohen (1960), is a member of a family of indices to measure the agreement between cellular maps. The indices (Cohen's kappa) are given by equation 3.7:

$$Kappa = \frac{PCP - p_c}{1 - p_c} \tag{3.7}$$

where PCP is the observed proportion agreement among raters, and p_c is the expected proportion agreement due to chance. Kappa has the following properties: if the observed correct proportion is greater than the expected correct proportion due to chance, then Kappa >0; if observed correct proportion is equal to the expected correct proportion due to chance, then Kappa = 0; and if observed correct proportion is less than the expected correct proportion due to chance, then Kappa < 0. If the raters are in complete agreement, then Kappa = 1.

The Cohen's Kappa is appropriate for contingency tables when the scientist does not have control over the marginal distributions. In contrast, usually the goal of a spatially explicit model is to obtain similar marginal distributions. Thus, Pontius and Gilmore (2000) suggest the use of Kno to evaluate a model's overall success, Klocation to evaluate the model's ability to specify location, and Kquantity to evaluate the model's ability to specify quantity. SUPPOSE there are J categories, table 3.3 give the proportion of correct classification according to a model's ability to specify quantity and location accurately.

Table 3.3: Proportion correct classification according to a model's ability to specify accurately quantity and location

Ability to Specify Quantity	**Ability to Specify Location**		
	No(NL)	Medium(ML)	Perfect(PL)
No(NQ)	1/J	(1/J)+Klocation(NQPL-(1/J))	$\sum_{j=0}^{J-1}\min(1/J, O_j)$
Medium(MQ)	$\sum_{j=0}^{J-1}(P_j, O_j)$	Percentage of Correct Prediction	$\sum_{j=0}^{J-1}(P_j, O_j)$
Perfect(PQ)	$\sum_{j=0}^{J-1}(O_j^2)$	PQNL+Klocation(1-PQNL)	1

Where $O_j = \dfrac{Num(O-j)}{Num(Total)}$ $j = 0,1,2,3$ and $P_j = \dfrac{Num(P-j)}{Num(Total)}$ $j = 0,1,2,3$.

Following that, the definitions of Kappa coefficients are given as follows.

$$\text{Kno} = \frac{PCP - NQNL}{1 - NQNL} \tag{3.8}$$

$$\text{Klocation} = \frac{PCP - MQNL}{MQPL - MQNL} \tag{3.9}$$

$$\text{Kquantity} = \frac{PCP - NQML}{PQML - NQML} \tag{3.10}$$

The model's performance is corroborated from the values of the Kappa coefficients, which are closer to 1.

3.6 Results and Discussion

The regression result generated by MNLM, estimated using the maximum likelihood algorithm, is shown in table 3.4. In this study, the negative effect of the distance variable for residential land use reflects the fact that residential zone is increasingly dispersed in the locations where living environment is more attractive. Almost all the distance variables have negative effect on the industrial land use type, except the distance to the educational facilities and the distance to road. For transportation/commercial/others, the transportation factors (distance to the rail infrastructure and the distance to road) have significant negative effect. This reflects that transportation infrastructure plays a crucial role in the expansion of transportation/commercial/others.

The positive signs on population density for all built-up land use types suggest that population is a chief factor influencing urbanization. Meanwhile, the signs on DEM and slope are significantly negative, which implies that the probability of built-up land use type will decrease as the elevation and slope increase.

Industrial zoning regulations are not estimated to have any significant effect on any land use types. Residential and transportation/commercial/others zoning appear to have a positive impact on such development. Since the 'no planning' factor is not significant for all land use types, such (no planning) places appear in a random development pattern. In all, zoning appears to mainly facilitate development of corresponding land uses, but does not necessarily impede other land use types. Zoning policy continues to play a pivotal role in controlling the urban development in the special economic zone.

Table 3.4: Estimation results of MNLM

Parameters	Coef.(Residential)	t-stats	Coef.(Industrial)	t-stats	Coef.(Commercial / Transportation /others)	t-stats
Constant	3.07e01	9.27e01	9.87e01	2.53	1.10	4.97
Distance to Commercial Centre	-1.45e-04	-6.93e-01	-8.13e-05	-4.05e-01	2.70e-04	2.25
Distance to Financial Centre	-1.67e-04	-1.30	-4.70e-05	-4.11e-01	-9.25e-05	-1.32
Distance to Industrial Centre	-2.63e-05	-1.36e-01	-9.89e-03	-1.39e01	1.60e-04	1.07
Distance to Educational Facilities	-6.66e-04	-2.45	1.17e-03	4.68	6.28e-05	4.05e-01
Distance to Railway Infrastructure	-4.89e-05	-1.32	-1.55e-04	-3.51	-1.03e-04	-3.61
Distance to Road	-2.24e-03	-3.81	1.24e-04	3.53e-01	-1.82e-03	-6.76
Population	4.07e-05	3.16	3.93e-05	2.48	4.54	4.22
DEM	-1.09e-02	-3.08	-1.72e-02	-2.96	-1.23e-02	-7.29
Slope	-6.56e-02	-5.02	-8.66e-02	-5.19	-7.3e-02	-1.01e01
Planning for Residential	2.30	8.15	1.20	3.45	9.70e-01	4.52
Planning for Industrial	1.32e01	6.37e-02	1.33e01	6.44e-02	1.31e01	6.33e-02
Planning for Transportation/commercial/others	-1.31e-01	-4.86e-01	1.08	3.55	8.23e-01	5.06
No Planning	-1.23e-01	-4.14e-01	9.82e-02	2.97e-01	-1.49e-02	-8.88e-02
Log likelihood function	-2.6980e+003					
PCP	74.1%					

Table 3.4 presents the accuracy of multinomial logit model. The PCP is 74.1%.

Table 3.4: The accuracy of multinomial logistic regression model

Observed	Predicted				Total	PCP
	0	1	2	3		
0	1726	36	9	206	1997	86.4%
1	33	352	8	153	546	64.5%
2	27	24	152	97	300	50.7%
3	178	274	42	883	1377	64.1%
Total	1964	686	211	1339	3113	74.1%

The Kappa coefficients are also calculated and Table 3.5 shows that the model is good at specifying quantity.

Table 3.5 MNLM's Kappa coefficients

	MNLM
Kno	0.65
Klocation	0.63
Kquantity	0.90

3.7 Chapter Summary

This chapter presented a brief introduction to the study area: special economic zone. This was followed by a detailed discussion pertaining to data preparation and processing procedures. Then a random sampling was employed to obtain the spatio-temporal data set for use in the following study. Following that, the formation of multinomial logit was covered. PCP and Kappa coefficients, which are used to evaluate classification accuracy and ability of the model, were discussed. Finally, the land use change in SEZ was analyzed and the results evince that planning plays an important role in the urbanization of SEZ.

Chapter 4: Geographically and Temporally Weighted Logit Model

Typically, non-stationarity is a characteristic of spatial data. Non-stationarity refers to the variations of relationship between response variable and explanatory variables in a regression model. Fotheringham and Brunsdon *et al.* (1996) proposed geographically weighted regression (GWR) to consider spatial non-stationarity by estimating the parameters in each location. This chapter aims to explore the possibility of extending GWR to support a discrete model in a spatio-temporal setting.

4.1 Introduction

Applied work in spatial analysis always relies heavily on the sample data that are collected with reference to location. When the sample data have a locational component, spatial heterogeneity may occur in the relationships being modeled. This could violate the basic assumption of statistical independence of observations, which typically is required for unbiased and efficient estimation (Huang *et al.*, 2009). A thorough understanding of non-stationarity is inevitable to utilize the literature for considering spatial and temporal non-stationarity in land use change modeling. The term 'stationarity' is often taken to refer to the outcome of some process that has similar properties at all points of interest. In other words, the statistical properties (e.g., mean and variance) of the variable or variables do not vary over the area of interest. Irrespective of the location, a stationary model has the same parameters, whereas with a non-stationary model the parameters are allowed to vary locally.

The traditional approach for land use change modeling is to estimate the underlying function globally. However, it is obvious that spatio-temporal non-stationarity exists in land use change models. Herein, acknowledgement of the stationary assumption is perhaps likely to be invalid, as parameters tend to vary over space and time. For example, the expansion of the European Union with the consequent likely restructuring of the Common Agricultural Policy (CAP) will lead to the shifts in the political and economic forces driving LUCC in the region. The implication is that some factors, which are contributors to land use change models (e.g. socio-economic factors), are likely to interpret land use change differently at varying locations and points of time.

Meanwhile, local models which consider the non-stationarity have been a source of immense interest. This is due to the better results such local models could provide than the global models. Numerous local models have been developed, such as spatial expansion method, spatially adaptive filtering, multilevel modeling, Geographically Weighted Regression (GWR), dynamic model, and coefficient smoothing. However, due attention has not been paid to consider spatial and temporal non-stationarity within the same framework. Thus, based on the logistic regression model, this chapter aims to construct geographically and temporally weighted logit Model (GTWLM) including both spatial and temporal non-staionarity to assist the spatio-temporal land use change analysis. The emphasis on the logistic regression model is due to the fundamental role it plays in examining the relationship quantitatively between land use changes and explanatory variables. A study on spatio-temporal land use change in SEZ, Shenzhen will be carried out to evaluate the performance and reliability of the proposed model. A significant improvement achieved by GTWLR will be demonstrated by comparing the estimation results of the models.

4.2 Local Models

Interest in local forms of spatio-temporal analysis and spatio-temporal modeling is not new. The local models can be constructed in various ways. This section reviews several local models, which are capable of considering both spatial and temporal non-stationarity.

4.2.1 Models considering spatial non-stationarity

Four methods are described here: Spatial expansion method, spatially adaptive filtering, multilevel modeling, and geographically weighted regression. The following passages explain the rationale behind the limited application of these techniques to the analysis of spatial non-stationarity.

4.2.1.1 The spatial expansion method

The spatial expansion model considering spatial non-stationarity was introduced by Casetti (1972):

$$y_i = \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i \tag{4.1}$$

where y denotes a dependent variable, the x are independent variables, $\beta_1, \beta_2, \ldots, \beta_p$ represent parameters to be estimated, ε represents an error term. This global model can be expanded by allowing each of the parameters to be functions of coordinates. For instance, the parameters are allowed to vary geographically as follows.

$$\beta_{1i} = \alpha_{10} + \alpha_{11} u_i + \alpha_{12} v_i \tag{4.2}$$

$$\beta_{2i} = \alpha_{20} + \alpha_{21} u_i + \alpha_{22} v_i \tag{4.3}$$

......

$$\beta_{pi} = \alpha_{p0} + \alpha_{p1} u_i + \alpha_{p2} v_i \tag{4.4}$$

where u_i and v_i represent the spatial coordinates of location i. equations (4.2) to (4.4) represent simple linear expansions of the global parameters. This expansion method is very important in promoting the awareness of spatial non-stationarity. However, the form of the expansions equations needs to be assumed to follow a priori.

4.2.1.2 The Spatially Adaptive Filtering

Another approach that allows parameters to vary is the spatial adaptive filtering (Widrow and Hoff 1960).

$$y_i = \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i \quad (4.5)$$

As a rule, when a new observation occurs at another location j, the existing regression parameters, $\beta_1, \beta_2, \ldots, \beta_p$, are used to predict the dependent variable y_j. In spatially adaptive filtering, the values of the regression coefficient $\hat{\beta}$ are adjusted to $\hat{\beta}'$ to improve the estimate. To avoid problems of overcompensation, the degree of adjustment applied could be 'damped' by the following update rule:

$$\hat{\beta}' = \hat{\beta} + |\hat{\beta}| \alpha (y_j - \hat{y}_j) / |\hat{y}_j| \quad (4.6)$$

where $\hat{y}_j$ is the predicted value of y_j based on $\hat{\beta}$ and α is the damping vector controlling the extent to which the correction is applied for parameters. In spatial analysis, the flow of coefficient estimates updating is a collaborative process between a pair of neighboring zones. This requires iterating between coefficient estimates until some form of convergence is achieved. Then, the results should be a unique estimate of the regression coefficient vector β for each case. The fact remains that the case-wise correction procedure is some degree of spatial smoothing of the estimates of individual elements of β. Thus, the method only tends to allow the parameters to 'drift' slowly across the geographical space.

4.2.1.3 Multilevel Modeling

Multilevel modeling tries to combine an individual-level model with a macro-level model. The model has the following form:

$$y_{im} = \alpha_m + \beta_m x_{im} + \varepsilon_{im} \quad (4.7)$$

where y_{im} represents the value of individual i living place m; x_{im} is the *ith* observation of attribute x at place m. α_m and β_m are rewrote as follows.

$$\alpha_m = \alpha + \mu_m^{\alpha} \tag{4.8}$$

$$\beta_m = \beta + \mu_m^{\beta} \tag{4.9}$$

where μ_m^{α} and μ_m^{β} are random place-specific variable. Substituting (4.8) and (4.9) into (4.7) yields the multi-level model,

$$y_{im} = \alpha + \beta x_{im} + (\varepsilon_{im} + \mu_m^{\alpha} + \mu_m^{\beta} x_{im}) \tag{4.10}$$

Adding place attributes μ_m^{α} and μ_m^{β} into α_m and β_m, and extending the number of levels in the hierarchy are the salient advantages of this model. However, the reliability on a priori definition of discrete set of spatial units at each level is a disadvantage.

4.2.1.4 Geographically Weighted Regression (GWR)

The GWR model allows regression parameters to be estimated locally. The model can be expressed as follows:

$$Y_i = \sum_p \beta_p(u_i, v_i) X_{ip} + \varepsilon_i \; i = 1, \cdots, n \tag{4.11}$$

where (u_i, v_i) denotes the coordinates of the point i in space, $\beta_p(u_i, v_i)$ represents a set of parameters at point i.

The estimation of parameters $\beta_p(u_i, v_i)$ is given by the following equation:

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y \, . \tag{4.12}$$

where $W(u_i, v_i)$ is a $n \times n$ diagonal matrix computed for each point i. The closer the observation is to i, the greater the weight.

$$W(u_i, v_i) = \begin{pmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{pmatrix} \tag{4.13}$$

Especially, GWR based on the regression model has powerful explanatory capability and facilitates exploring the variation of the parameters by visualization. This implies that the local parameter estimates are capable of being mapped. However, the estimations required for each point increase the computational burden of this method.

4.2.2 Models considering temporal non-stationarity

Models built during each time period can partly consider temporal non-stationarity. Besides, as can be seen below, there are some other methods to deal with this issue.

4.2.2.1 Dynamic Model

The dynamic model is used to articulate and model the behavior of the system over time. In such a model, the state in time t is usually derived from the state before time t. This framework facilitates carefully considering the temporal non-stationarity.

For example, in the article titled as 'Modeling Land Use Decisions with Aggregate Data: Dynamic Land Use', a Markov decision process is used to represent the conditional land-use decisions made by landowners. The decision process is assumed to be non-stationary so that the probability that the cell transforms from one land use to another changes over time. Also, it is dependent on the output and input prices as well as other relevant decision variables. The dynamic modeling framework could be used to recover non-stationary transition probabilities for land-use allocations.

The principal advantage of using dynamic models is the ability to identify the land use change components in the future. Often, this information is needed to accurately represent responses to and impacts of land use policies. Summarily, although estimating dynamic models is complex and time-consuming, such models have powerful capabilities for prediction.

4.2.2.2 Coefficient smoothing

In this method, for each of explanatory variable, the relationship between its influence coefficient and time is investigated using a certain function. The function could be a quadratic polynomial or exponential smoothing and so on. Huang *et al.* (2009) use a model coupled with a logistic regression model and an exponential smoothing technique for exploring the effects of various factors on land use change. The modified exponential smoothing technique is employed to generate a smoothed model from a series of bi-temporal models obtained from different time periods.

By integrating the bi-temporal models, this method generalizes the details of land use change over different phases during the long period. Smoothing technique can be implemented at the coefficient level or the utility function level. The challenge here is that smoothing at the coefficient level entails a specific relationship, e.g. linear, which may not be present.

4.3 Methodology

4.3.1 Overview of geographically weighted logit model (GWLM)

Assuming that there is a set of n observations x_{ip} with the spatial coordinates (u_i, v_i), $i = 1, 2, ..., n$, on p predictor variables, $p = 1, 2, ..., P$, and a set of n observations on a dependent variable y_i^j, which is an indicator of the observed land use type j for cell i:

if yes, y_i^j equals 1; if no, y_i^j equals 0. j is the index of land use type in this study. The underlying model for GWLM is as follows:

$$\log\left(\frac{prob(y_i^j=1)}{prob(y_i^0=1)}\right)=\sum_p \beta_p^j(u_i,v_i)X_{ip}+\varepsilon_i \tag{4.14}$$

where j=0 denotes undeveloped land use type as basic land use type in this study, $\beta_p^j(u_i,v_i)$ are p continuous functions of the location (u_i,v_i) in the study area. Unlike the multinomial logit model, Geographically Weighted Logit Model allows the parameters to vary across space, and hence is more likely to capture positional effects.

Parameter estimation is a moving window process. A region or window is drawn around a location i, and all the data points within this region or window are used to estimate the parameters. The process is repeated for each observation in the data, and finally, a set of parameter estimates is obtained for each location. The estimator of $\beta_p^j(u_i,v_i)$ is given at each location i by using the multinomial logistic regression with X's transform as follows:

$$X'=W(u_i,v_i)^{1/2}X \tag{4.15}$$

where $W(u_i,v_i)$ is an n by n diagonal matrix:

$$W(u_i,v_i)=\begin{pmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{pmatrix} \tag{4.16}$$

In global weighted regression models, the values of $W(u_i,v_i)$ remain constant. However, it is assumed that the observed data close to point i have a greater influence in

$W(u_i, v_i)$ than the data located farther from the point of observation i in GWLM.

Basically, there are two kernels to construct weighting regimes: fixed kernel and adaptive kernel. For the fixed and adaptive kernel, the distance and the number of neighbors remain constant, respectively. Frequently used weighting functions include the bi-square function, the tri-cube kernel function, Gaussian and exponential distance decay based functions. Exponential distance decay based function, which is employed in this case, is expressed as follows:

$$w_{im} = \exp(-d_{im}^2 / h^2) \tag{4.17}$$

where h is bandwidth, and d_{im} , which is the spatial distance between location i and m, can be calculated by equation 4.18.

$$d_{im} = \sqrt{(u_i - u_m)^2 + (v_i - v_m)^2} \tag{4.18}$$

where (u_i, v_i) and (u_m, v_m) refer to the point coordinates. According to equations (4.17) and (4.18), the weighting of other data decreases in accordance with the exponential form when the distance between i and m increases. If i and m coincide, the weighting value at this point will be equal to 1.

There is an issue involved in the estimation of GWLM. On the one hand, compared to OLS, multinomial logistic regression, which involves non-linear optimization, is time-consuming. Employing the entire sample will greatly increase the computational load. Hence, when parameters are estimated in every point, sub-samples are chosen using a computational ruse that ignores observations with negligible weight. On the other hand, a sub-sample cannot be produced with the values of all single types (e.g. all ones). Hence, weighting method employed here should guarantee the estimation of multinomial logistic regression. Adaptive weighting functions are used to adapt themselves in size to ensure that the same and enough number of non-zero weights are used for each multinomial logistic regression point being analyzed. In this manner, the kernels have larger bandwidths where the data points are sparsely distributed and have smaller ones where the data are abundant. In this study, the specific adaptive weighting function is expressed as follows:

$$w_{im} = \begin{cases} \exp(-d_{im}^2 / h^2), & \textit{if } m \text{ is the } q \text{ nearest points around } i \\ 0, & \textit{otherwise} \end{cases} \tag{4.19}$$

where h stands for the bandwidth, q represents the proportion of observations to consider in the estimation of regression at each location.

The choice of an appropriate bandwidth h and sub-sample size q is very important for the GWR based model. Then, the weighting matrix could be decided by cross-validation. Actually, it can be automatically obtained with an optimization technique by maximizing PCP (Hurvich *et al.*, 1998; Fotheringham *et al.*, 2002).

4.3.2 Geographically and temporally weighted logit model (GTWLM)

The form of GTWLM is as follows:

$$\log\left(\frac{prob(y_i^j = 1)}{prob(y_i^0 = 1)}\right) = \sum_p \beta_p^j(u_i, v_i, t_i) X_{ip} + \varepsilon_i \tag{4.20}$$

where t_i is temporal coordinate, $\beta_p^j(u_i, v_i, t_i)$ are p continuous functions of the spatio-temporal coordinates (u_i, v_i, t_i) in the study area.

In order to estimate $\beta_p^j(u_i, v_i, t_i)$ in each point, X should be transferred into X' in GTWLM as follows.

$$X' = W^{ST}(u_i, v_i, t_i)^{1/2} X \tag{4.21}$$

$W^{ST}(u_i, v_i, t_i)$ is also a diagonal matrix, whose elements w_{im}^{ST}, should be spatio-temporal distance decay functions of (u, v, t) when calibrating weight between point m adjacent to observation point i. Consequently, the estimation of GTWLM relies on the appropriate specification of the spatio-temporal distance decay function. Nevertheless, space and time are usually measured in different units and have different scale effects. To handle the scale problem, all the coordinate data are standardized firstly. For the difference about scale effects, an ellipsoidal coordinate system could be employed to combine spatial effect and temporal effect into the same framework. This is done to construct the spatio-temporal distance between the point, which will be estimated, and its surrounding observed points (Figure 4.1).

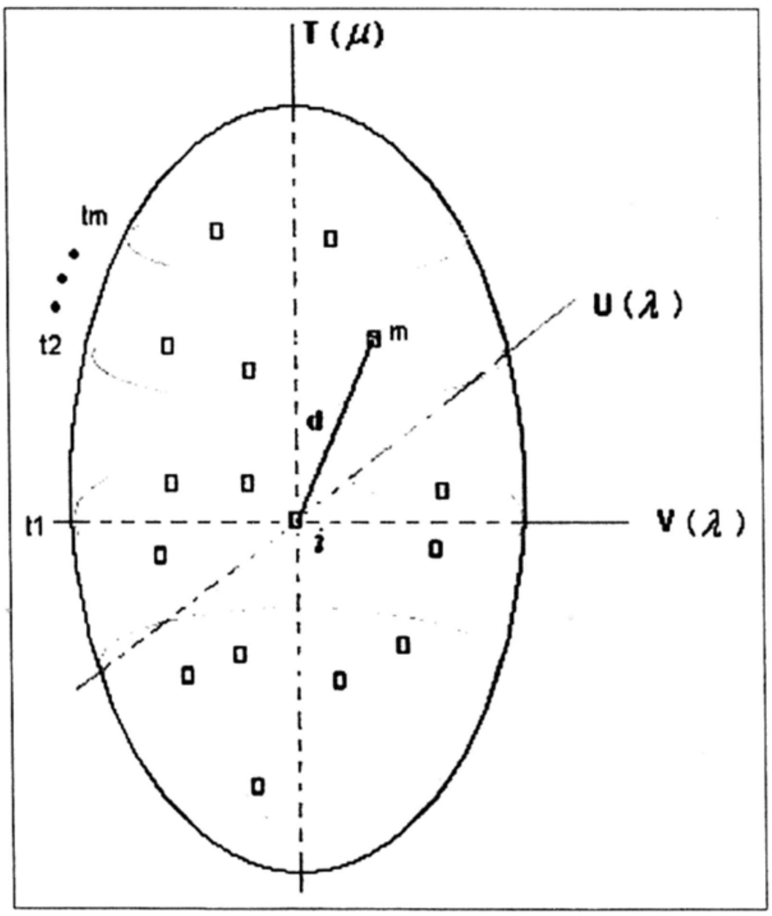


Figure 4.1: An illustration of spatio-temporal distance

where i is regression point, m is the nearby point (u_m, v_m, t_m), $m = \{1, 2, \cdots, n\}$ $(m \neq i)$. The formula of the spatio-temporal distance is as follows:

$$d_{im}^{ST} = \sqrt{\lambda((u_i - u_m)^2 + (v_i - v_m)^2) + \mu(t_i - t_m)^2} \tag{4.22}$$

where t_i and t_m are the temporal coordinates, λ and μ are the scale factors to balance the different effects used to measure the spatial and temporal distance in their respective metric systems. The d_{im}^{ST} will be adopted to compute the spatial-temporal distance in GTWLR.

Based on d^{ST}, the elements of $W^{ST}(u_i, v_i, t_i)$, w_{im}^{ST} is defined.

$$w_{im}^{ST} = \begin{cases} \exp\left\{-(\frac{(d_{im}^{ST})^2}{h_{ST}^2})\right\}, & \textit{if } m \text{ is the } q \text{ nearest points around } i \\ 0, & \textit{otherwise} \end{cases} \tag{4.23}$$

where h_{ST} is spatio-temporal bandwidth. Specifically, we substitute 4.22 in 4.23:

$$\begin{aligned} w_{im}^{ST} &= \exp\left\{-(\frac{\lambda[(u_i - u_m)^2 + (v_i - v_m)^2] + \mu(t_i - t_m)^2}{h_{ST}^2})\right\} \\ &= \exp\left\{-(\frac{(u_i - u_m)^2 + (v_i - v_m)^2}{h_S^2} + \frac{(t_i - t_m)^2}{h_T^2})\right\} \\ &= \exp\left\{-((d_{im}^S)^2 / h_S^2 + (d_{im}^T)^2 / h_T^2)\right\} \\ &= \exp\left\{-(d_{im}^S)^2 / h_S^2\right\} \times \exp\left\{-(d_{im}^T)^2 / h_S^T\right\} \\ &= w_{im}^S \cdot w_{im}^T \end{aligned} \tag{4.24}$$

where $w_{im}^{S} = \exp\left\{-\frac{(d_{im}^{S})^2}{h_S^2}\right\}$, $w_{im}^{T} = \exp\left\{-\frac{(d_{im}^{T})^2}{h_S^T}\right\}$, $(d_{im}^{S})^2 = (u_i - u_m)^2 + (v_i - v_m)^2$,

$(d_{im}^{T})^2 = (t_i - t_m)^2$, $h_S^2 = \frac{h_{ST}^2}{\lambda}$ and $h_T^2 = \frac{h_{ST}^2}{\mu}$ are the spatial and temporal bandwidths, respectively. Since the weighting function is a diagonal matrix, whose diagonal elements are multiplied by $w_{im}^{S} \cdot w_{im}^{T}$ $(1 \le m \le n)$, W^{ST} can be seen as the combination of spatially weighted matrix W^S and temporally weighted matrix W^T: $W^{ST} = W^S \times W^T$. Thus, if μ is set to 0, GTWLM becomes GWLM, which only consider spatial non-stationarity. If λ is set to 0, GTWLM becomes the temporal weighted logit model (TWLM), which only consider temporal non-stationarity.

Similar to GWLM, the parameters are also estimated by CV method. In order to reduce the parameters to be estimated in the model, Equation 4.22 is divided by $\sqrt{\lambda}$ ($\lambda \neq 0$).

$$d_{im}^{ST} / \sqrt{\lambda} = \sqrt{[(u_i - u_m)^2 + (v_i - v_m)^2] + \tau (t_i - t_m)^2} \tag{4.25}$$

where τ represents the parameter ratio μ/λ and can be seen as balance ratio between spatial distance effect and temporal distance effect. Thus,

$$\begin{aligned} w_{im}^{ST} &= \exp\left\{-\left(\frac{\lambda[(u_i - u_m)^2 + (v_i - v_m)^2] + \mu (t_i - t_m)^2}{h_{ST}^2}\right)\right\} \\ &= \exp\left\{-\left(\frac{[(u_i - u_m)^2 + (v_i - v_m)^2] + \tau (t_i - t_m)^2}{h_{ST}^2 / \lambda}\right)\right\} \end{aligned} \tag{4.26}$$

Here, setting $\lambda = 1$ will not change the results estimated by λ, μ, and h_{ST} using cross validation in terms of PCP. The estimation results will give τ and h_{ST} instead of λ, μ and h_{ST} in this study.

The Nelder-Mead Simplex Method is employed to obtain τ and h_{ST}, which enable the model to attain the maximum PCP. This is a direct search method that does not use numerical or analytic gradients. If n is the length of x, a simplex in n-dimensional space is characterized by the $n+1$ distinct vectors that are its vertices. In two-dimensional, a simplex is a triangle. During each step of the search, a new point in or near the current simplex is generated. The function value at the new point is compared with the function's values at the vertices of the simplex. Typically, one of the vertices is replaced by the new point, giving a new simplex. This step is repeated until the diameter of the simplex is less than the specified tolerance.

4.3.3 McNamara's test

To demonstrate the effectiveness of the proposed model, the estimation results are compared with those obtained from the MNLM. In addition to the Percentage of Correctly Predicted (PCP), which measures the goodness-of-fit of the models, the difference between the accuracies achieved by the two methods (GTWLM and the conventional MNLM) is also assessed by McNamara's test (see also Foody, 2004). This test is based on the standardized normal test statistic:

$$Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \tag{4.27}$$

where f_{12} denotes the number of samples classified correctly and wrongly by the first and second models, respectively. Accordingly, f_{12} and f_{21} are the counts of classified samples, which the first and second models disagree. A lower prediction error (higher accuracy) is identified by the value of Z. A negative value of Z indicates that the results from f_{12} are more accurate than the results from model f_{21}. At the commonly used 5% level of significance, the difference in the accuracies between the first and the second models is evaluated to be statistically significant if $|Z| > 1.96$.

4.4 Results and discussion

The TWLM, GWLM, and GTWLM are performed using the data of SEZ and the estimates are reported in Tables 4.1, 4.2 and 4.3, respectively. As the output of the local parameter estimates from WLM, GWLM, and GTWLM would be voluminous, only the minimum, median, and maximum value for each parameters are provided to give a summary of the distribution. In TWLM, GWLM, or GTWLM, the magnitude of all the parameters in the global models are between the minimum and maximum, and the signs of median for all the parameters are almost the same as MNLM. The balance factor between the spatial distance effect and the temporal distance effect τ being 0.356 implies the leading role played by the spatial effect.

Table 4.1: TWLM parameter estimate summaries

TWLR (bandwidth=1.180; q=700)									
	Residential			**Industrial**			**Transportation/commercial/others**		
Parameter	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Medi**	**Max**
Constant	-1.97	0.905e-01	1.13	-3.59e-01	4.97e-01	2.46	2.18e-01	8.37e-01	2.66
Distance to Commercial Centre	-1.40e-03	-4.95e-04	7.44e-04	-1.13e-03	-1.75e-04	7.26e-04	-2.03e-04	2.00e-04	1.05e-03
Distance to Financial Centre	-4.49e-04	-1.26e-04	2.42e-04	-2.77e-04	-6.79e-05	7.26e-04	-4.79e-04	1.43e-05	3.43e-04
Distance to Industrial Centre	-6.15e-04	-1.38e-04	1.22e-03	-1.79e-02	-1.17e-02	-4.62e-03	-5.60e-04	3.26e-04	5.36e-04
Distance to Educational facilities	-1.14e-03	-6.59e-04	2.19e-04	-1.96e-04	1.65e-03	2.54e-03	-2.53e-04	9.21e-05	6.73e-04
Distance to Railway Infrastructure	-1.02e-04	-2.41e-05	3.25e-05	-3.21e-04	-1.87e-04	-6.09e-05	-2.14e-04	-1.26e-04	8.25e-05
Distance to Road	-5.33e-03	-2.80e-03	2.66e-04	-1.68e-03	7.30e-04	7.39e-03	-9.54e-03	-3.42e-03	-3.97e-04
Population	-3.12e-06	3.69e-05	1.05e-04	-4.06e-06	6.97e-05	1.11e-04	1.89e-05	5.10e-05	1.08e-04
DEM	-2.47e-02	-1.19e-02	-9.51e-03	-3.76e-02	-2.19e-02	-5.23e-03	-2.52e-02	-1.19e-02	-6.57e-03
Slope	-1.37e-01	-6.87e-02	-4.56e-02	-2.66e-01	-8.16e-02	-6.85e-02	-1.04e-01	-6.52e-02	-3.79e-02
Planning for Residential	1.49	2.45	3.83	7.06e-01	1.47	2.26	6.97e-01	8.00e-01	2.58
Planning for Industrial	9.94	1.13e01	1.37e01	1.28	1.13e01	1.30e01	9.93	1.12e01	1.26e01
Planning for Transportation/commercial/others	-1.40	-3.99e-02	9.35e-01	8.57e-01	1.30	2.18	3.81e-01	6.36e-01	2.08
No Planning	-9.85e-01	-2.30e-01	1.60	-4.48e-01	1.47e-01	1.73	-3.07e-01	-7.34e-02	8.57e-01
PCP	75.4%								

Table 4.2: GWLM parameter estimate summaries

GWLR (bandwidth=0.5916; q=1323)									
	Residential			**Industrial**			**Transportation/commercial/others**		
Parameter	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Medi**	**Max**
Constant	-26.8e01	4.09	2.53e01	-3.29e01	10.1e01	6.22e01	-9.97	3.70	4.29e01
Distance to Commercial Centre	-3.85e-02	-3.62e-04	6.20e-03	-2.35e-02	-3.68e-04	1.83e-02	-1.92e-02	1.06e-03	1.27e-02
Distance to Financial Centre	-8.67e-03	-9.33e-04	1.38e-02	-8.79e-03	2.11e-04	2.41e-02	-6.77e-03	-5.21e-04	1.02e-02
Distance to Industrial Centre	-1.55e-02	-2.59e-04	2.55e-02	-3.06	-7.09e-02	-4.16e-03	-8.77e-03	1.04e-03	1.75e-02
Distance to Educational facilities	-5.35e-02	-1.58e-03	1.02e-02	-2.95e-02	4.28e-03	3.91e-02	-8.08e-03	2.40e-04	9.23e-03
Distance to Railway Infrastructure	-5.77e-03	-4.93e-04	6.78e-03	-2.81e-02	-1.17e-03	1.46e-03	-4.74e-03	-6.52e-04	7.22e-04
Distance to Road	-8.11e-02	-9.96e-03	6.44e-03	-3.72e-02	3.26e-03	6.47e-02	-5.75e-02	-1.01e-02	9.74e-03
Population	-4.72e-04	2.52e-04	3.51e-03	-1.23e-03	1.69e-04	7.76e-03	-4.08e-04	4.16e-04	3.50e-03
DEM	-1.15	-1.08e-01	3.31e-03	-9.35e-01	-9.08e-02	5.24e-02	-3.48e-01	-6.92e-02	-1.50e-02
Slope	-3.16	-3.17e-01	3.74e-01	-5.49	-4.83e-01	2.03e-01	-2.87	-2.80e-01	7.39e-03
Planning for Residential	-5.69e-01	9.64	1.05e02	-3.00e01	3.76	6.05e01	-1.71	3.34	6.21e01
Planning for Industrial	-5.36	1.77e02	9.19e02	8.20	1.65e02	6.96e02	-1.28e01	1.73e02	9.85e02
Planning for Transportation/commercial/others	-4.01e01	-5.18e-01	3.87e01	-9.59	4.60	3.69e01	-3.71	3.27	1.77e01
No Planning	-5.05e01	-4.81e-01	3.12e01	-1.24e02	3.41	2.74e01	-6.17	1.40	1.77e01
PCP	79.2%								

Table 4.3: GTWLM parameter estimate summaries

GTWLM (bandwidth=0.6671; τ =0.356; q=920)									
	Residential			**Industrial**			**Transportation/commercial/others**		
Parameter	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**
Constant	-8.84e01	7.23	5.30e01	-5.48e01	1.38e01	6.99e01	-1.34e01	5.45	4.75e01
Distance to Commercial Centre	-3.50e-02	-4.86e-04	1.24e-02	-2.88e-02	-2.95e-04	1.72e-02	-2.42e-02	2.93e-03	1.89e-02
Distance to Financial Centre	-1.56e-02	-1.43e-03	2.22e-02	-1.76e-02	-2.06e-04	4.50e-02	-1.48e-02	-7.90e-04	2.04e-02
Distance to Industrial Centre	-2.06e-02	-7.79e-04	3.30e-02	-5.10	-1.02e-01	-6.94e-03	-1.57e-02	7.31e-04	2.77e-02
Distance to Educational facilities	-6.50e-02	-3.54e-03	2.62e-02	-3.49e-02	7.47e-03	6.90e-02	-1.23e-02	-1.79e-04	2.17e-02
Distance to Railway Infrastructure	-7.25e-03	-5.13e-04	8.75e-03	-2.99e-02	-2.20e-03	1.78e-03	-5.98e-03	-9.27e-04	1.58e-03
Distance to Road	-9.41e-02	-1.63e-02	1.01e-02	-6.93e-02	8.17e-03	1.31e-01	-1.11e-01	-1.64e-02	6.78e-03
Population	-2.86e-03	3.15e-04	4.22e-03	-1.90e-03	5.91e-04	1.80e-02	-9.61e-04	6.01e-04	4.79e-03
DEM	-1.28	-1.47e-01	3.17e-02	-7.23e-01	-1.44e-01	2.10e-01	-5.09e-01	-8.26e-02	7.68e-02
Slope	-5.11	-5.02e-01	1.08	-5.77	-6.27e-01	9.54e-01	-3.73	-3.96e-01	1.28e-01
Planning for Residential	-2.05e01	1.27e01	1.34e02	-4.04e01	3.81	7.43e01	-1.06e01	5.45	7.20e01
Planning for Industrial	-7.74e01	1.56e02	6.91e02	-9.01	1.68e02	1.210e03	-5.99e01	1.67e02	7.46e02
Planning for Transportation/commercial/others	-4.60e01	-1.11	7.99e01	-1.39e01	5.21	6.34e01	-1.50e01	5.06	2.38e01
No Planning	-4.66e02	-1.06	9.60e01	-6.53e03	1.02	7.19e01	-1.10e01	1.64	3.26e01
PCP	82.3%								

Table 4.4 provides the comparison of PCPs between MNLM and GTWLM. It should be noted that the PCP has increased from 74.1% in the MNLM to 82.3% in GTWLM. As seen from table 4.5, the PCP is 75.4% in TWLM, and 79.2% in GWLM. A comparison with PCP indicates that GTWLM gives a better fit of data than the TWLM, GWLM, and MNLM because GTWLM can handle both spatial and temporal non-staionarity. Moreover, Table 4.5 shows that the GWLM achieved a better goodness-of-fit than that of TWLM model. It reveals that the temporal non-stationary effect is less significant than that of spatial non-stationarity. A possible reason is that the experimental data only includes seven time spots, which provide less information.

Table 4.4: Comparison of PCPs between MNLM and GTWLM

Observed	MNLM						GTWLM					
	Predicted				Total	PCP	Predicted				Total	PCP
	0	1	2	3			0	1	2	3		
0	1726	36	9	206	1997	86.4%	1798	13	5	161	1977	90.9%
1	33	352	8	153	546	64.5%	25	370	7	144	546	67.8%
2	27	24	152	97	300	50.7%	2	3	248	47	300	82.7%
3	178	274	42	883	1377	64.1%	136	178	28	1035	1377	75.2%
Total	1964	686	211	1339	3113	74.1%	1961	564	288	1387	3451	82.3%

Table 4.5: PCPs of TWLM and GWLM

Observed	TWLM						GWLM					
	Predicted				Total	PCP	Predicted				Total	PCP
	0	1	2	3			0	1	2	3		
0	1738	27	14	198	1977	87.9%	1768	19	8	182	1977	89.4%
1	34	302	10	200	546	55.3%	27	355	15	149	546	65.0%
2	16	16	168	100	300	56.0%	4	3	236	57	300	78.6%
3	174	197	46	960	1377	69.7%	151	220	38	968	1377	70.3%
Total	1962	542	238	1458	3168	75.4%	1950	597	297	1356	3327	79.2%

TWLM, GWLM and GTWLM show improvements over the MNLM in terms of PCP. However, it is still necessary to investigate whether those models perform significantly better than MNLM from a statistical viewpoint. The improvement on TWLM and GWLM by GTWLM should be also tested.

Table 4.6: Significance comparison for GWR-based models

Models Comparison	Z values		
	TWLM	**GWLM**	**GTWLM**
MNLM	-2.64	-9.57	-14.16
TWLM	-	-6.73	-12.08
GWLM	-	-	-7.75

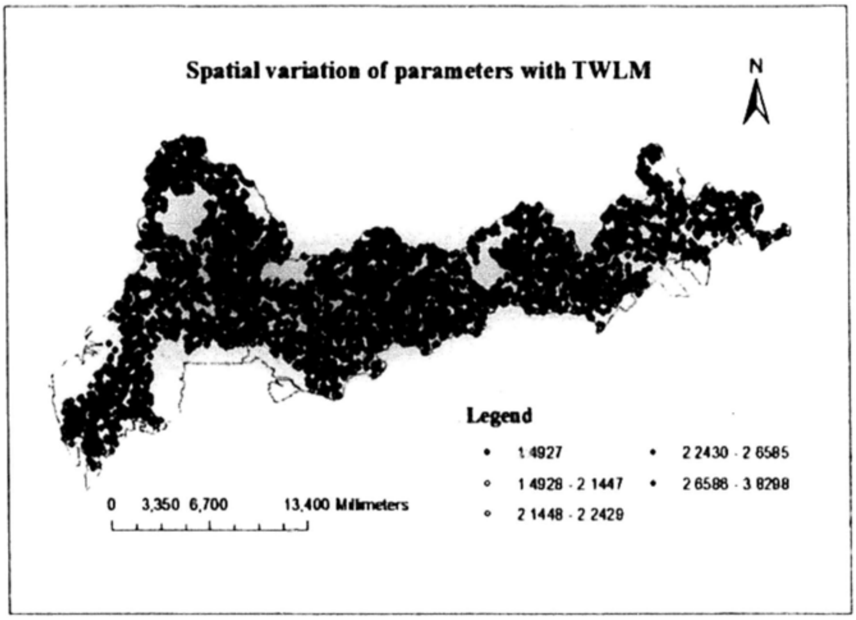
For comparison purposes, McNamara's test is employed to test the significant difference between MNLM, TWLM, GWLM, and GTWLM, and the results are shown in Table 4.6. They clearly indicate (negative value) that the GTWLM model performed better than GWLM and MNLM. The Z values between TWLM and GWLM are -5.33 and -13.67, respectively, indicating that GWLM substantially outperformed TWLM. Also, the Z values between TWLM and MNLM is -2.64, which is less than -1.96. These results demonstrate a significant difference between MNLM, TWLM, GWLM, and TWGR models at the 95% confidence level. These comparisons evince that GTWLM outperforms MNLM, GWLM, and TWLM in the model accuracy.

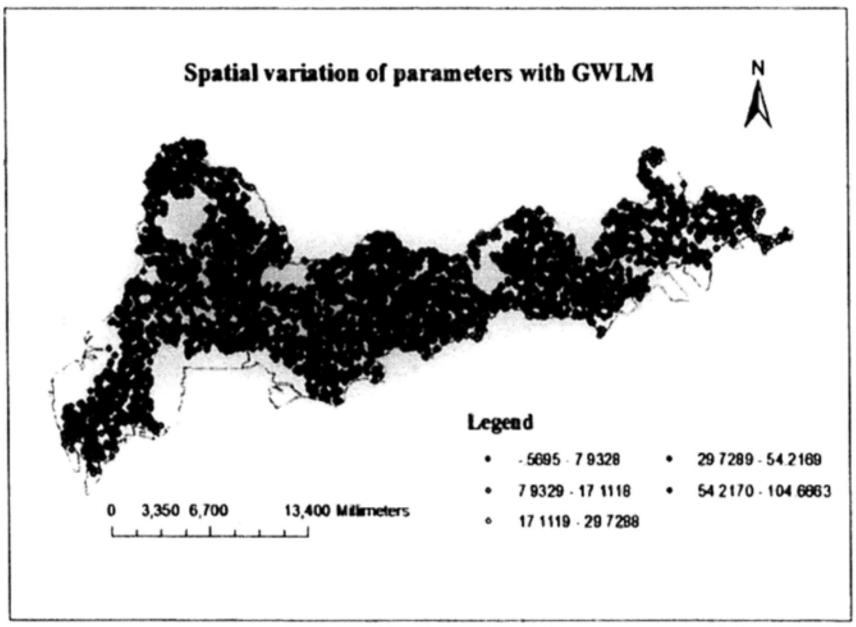
Their Kappa measures are also calculated and shown in table 4.7. All the kappa coefficients show that GTWLM provides a better result than the other models.

Table 4.7: Comparison of MNLM, TWLM, GWLM and GTWLM with Kappa coefficients

	MNLM	TWLM	GWLM	GTWLM
Kno	0.65	0.67	0.72	0.76
Klocation	0.63	0.64	0.69	0.73
Kquantity	0.90	0.94	0.97	0.98

Furthermore, the local parameter estimates of TWLM, GWLM, and GTWLM, which denote local relationships, are able to be displayed visually. Considering the coefficients of "Planning for Residential in land use residential" as an example, they can be grouped into several groups and each group can be colored to visualize the spatial variation patterns of this parameter.





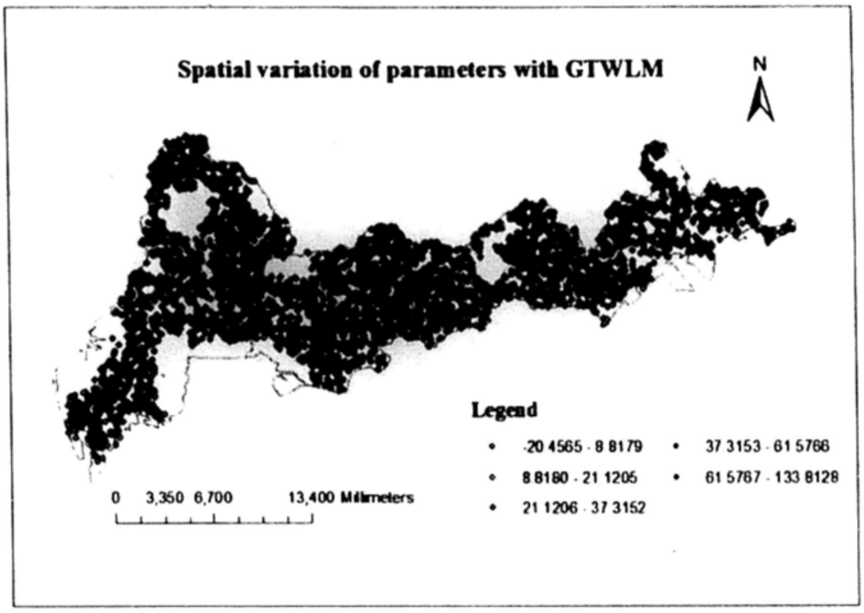


Figure 4.2: Spatial variation of the Planning for Residential coefficient in residential

The spatial distributions of the parameter estimates for "Planning for Residential in land use residential" of TWLM, GWLM, and GTWLM are shown in Figure 4.2. For the TWLM, no significant spatial variation is observed over time. The spatial variation of "Planning for Residential in land use residential" in both GWLM and GTWLM share analogous distributions. This, however, excludes the spatial variation in GTWLM, which portrays heterogeneity in greater detail. It can be also inferred that the spatio-temporal non-stationarity of the GTWLM is dominated by the spatial effect for this land use dataset.

Figure 4.2 illustrates that the parameter "Planning for Residential in land use residential" in GTWLM is smaller in Nanshan and Yantian district. This holds true especially in the northwest of Nanshan district, which is next to Bao'an, and bigger in Futian and Luohu district. This serves to indicate that the planning for residential areas is implemented better in the centre of SEZ. Owing to the rapid development, planning for residential almost does not work in some places for e.g. the region in the northwest of Nanshan district,.

4.5 Conclusion

The local models that consider non-stationarity refute the criticism that those adopting a quantitative approach to investigate land use change process are only concerned with the search for broad generalizations and have little interest in identifying local exceptions. Local forms of analysis also provide the relation between the outputs of the spatial techniques and the powerful visual display capabilities of GIS. Notably, they provide detailed information on spatial and temporal relationship, which enhances the model development process and facilitates understanding the land use change.

In this chapter, GWLM considering spatial non-stationarity has been extended to GTWLM in a spatio-temporal framework. The land use change in SEZ, Shenzhen has been modeled. Compared with the other models, GTWLM model shows a significant improvement in the percentage of correctly predicted. Compared with the global MNLM, the PCP values of TWLM and GWLM increase from 74.1% to 75.4% and 79.2%, respectively. The GTWLM yielded a considerably higher PCP of 82.3%. Statistical tests evince that a significant difference exists among MNLM, TWLM, GWLM, and GTWLM. The kappa coefficients also indicate that the GTWLM is better than the other models. Consequently, it can be safely concluded that it is meaningful to incorporate spatial and temporal non-stationarity into a land use change model.

However, several limitations remain to be studied in future. For instance, only seven temporal periods are available. The inadequacy in temporal information can be expected to degrade the model performance of TWLM and GTWLM. There exists a possibility that unobservable factors (not covered by our data set) and individual effect of land cell could influence the estimation results. More efficient spatio-temporal distance and weighting functions should be designed to constitute the weight. Besides, the spatio-temporal autocorrelation also should be considered in a more flexible spatio-temporal framework. Hence, a more elaborate version of GTWLM is covered in chapter 6.

4.6 Summary

In the chapter, non-stationarity was introduced at first. Following that, various ways in which earlier research dealt with spatial or temporal non-stationarity were described and analyzed. Then, a basic framework for GWLM was presented and it was extended to include temporal data as GTWLM, which was applied to the study of land use change in SEA, Shenzhen. The results for different models were compared and analyzed and it was concluded that the proposed model (GTWLM) performed better than the other models.

Chapter 5: Spatial-Temporal Panel Logit model

Anselin (1988, 1998) provides three principal methods for addressing the spatial autocorrelation: use of spatial stochastic processes, a direct representation of autocorrelations, and a non-parametric framework. The second method is preferred in our context and extended to consider spatio-temporal autocorrelation in land use change models. Meanwhile, the individual effect can be considered in a panel data framework. The combination of these two methods can consider both spatio-temporal autocorrelation and individual effect in land use change modeling.

5.1 Introduction

Spatial autocorrelation refers to the tendency of spatial variables and spatial land-use data to be dependent (Legendre and Fortin, 1989). Spatial autocorrelation, exhibited by substantial amounts of spatial variables and land use data, has been studied by numerous researchers. Spatial Autologistic Regression model or Spatial AutoLogit (SAL) model, the logistic regression model incorporating spatial autocorrelation, has been devised (Dubin 1995; Dubin 1997; LeSage, 1998). Such models have proven to be effective in regression analysis with consideration of spatial autocorrelation (Páez and Suzuki, 2001). Similar to the spatial lag model, a spatial error model (Anselin, 1988) can be employed to deal with spatial autocorrelation in a regression model.

On the other hand, less attention has been paid to temporal autocorrelation. Land use change modeling entails considering both spatial and temporal autocorrelation. Otherwise, inefficient parameter estimates and inaccurate measures of statistical significance will result. Two recent studies in land use change analysis have taken into consideration spatial and temporal autocorrelation in their statistical models. An and Brown (2008) introduce the concepts and models in survival analysis, and their potential applications in land use change. While survival analysis has proven to be effective in addressing temporal complexities, their model does not consider spatial autocorrelation usually based on the spatial weight structure. The land use model generated by Huang *et al.* (2009a) accounts for spatial neighbors in the last cross-section and employs a modified exponential smoothing technique to produce a smoothed model from a series of bi-temporal spatial models for different time periods. While spatial and temporal autocorrelation is partly considered in this model, it does not identify the individual specific effect in each cell.

In view of the shortcomings of the aforementioned models, this chapter aims to construct an innovative model to delineate the spatial-temporal pattern changes of land use. Panel data model, covered in detail in section 5.2, can be used to represent the individual effect of a land cell. Thus, panel data framework is employed to model land use change in this study. Based on the panel data analysis framework, the proposed spatio-temporal panel logit model (ST-PLM) considers the random effects of a land use cell with reference to both space and time and establishes the relationship among various factors and land use patterns over time. Considering spatial and temporal autocorrelation, the model incorporates the covariance of a cell to other cells into the model formulation. To incorporate the individual effect, the model accounts for the land use type of a cell in the initial year 1990. Notably, the proposed model integrates spatial autocorrelation with temporal autocorrelation in the same framework and performs the maximum simulated likelihood estimations. A study on spatio-temporal land use change in the special economy zone (SEZ) has been carried out to demonstrate the superior performance of the proposed model over the multinomial logit model.

5.2 Panel data model

A panel that is a cross-section or group of objects periodically observed over a given time span (Anselin, 1988; Greene, 2000) can be used to represent multi-temporal land use distributions of a region. Panel data analysis thus serves as a suitable tool for spatio-temporal land use change modeling.

5.2.1 Panel data

A panel data set follows a given sample of individuals over time, and subsequently provides multiple observations on each individual in the sample. A sample panel data set, consisting of a pool of observations on a cross-section over four time periods, is presented in Figure 5.1.

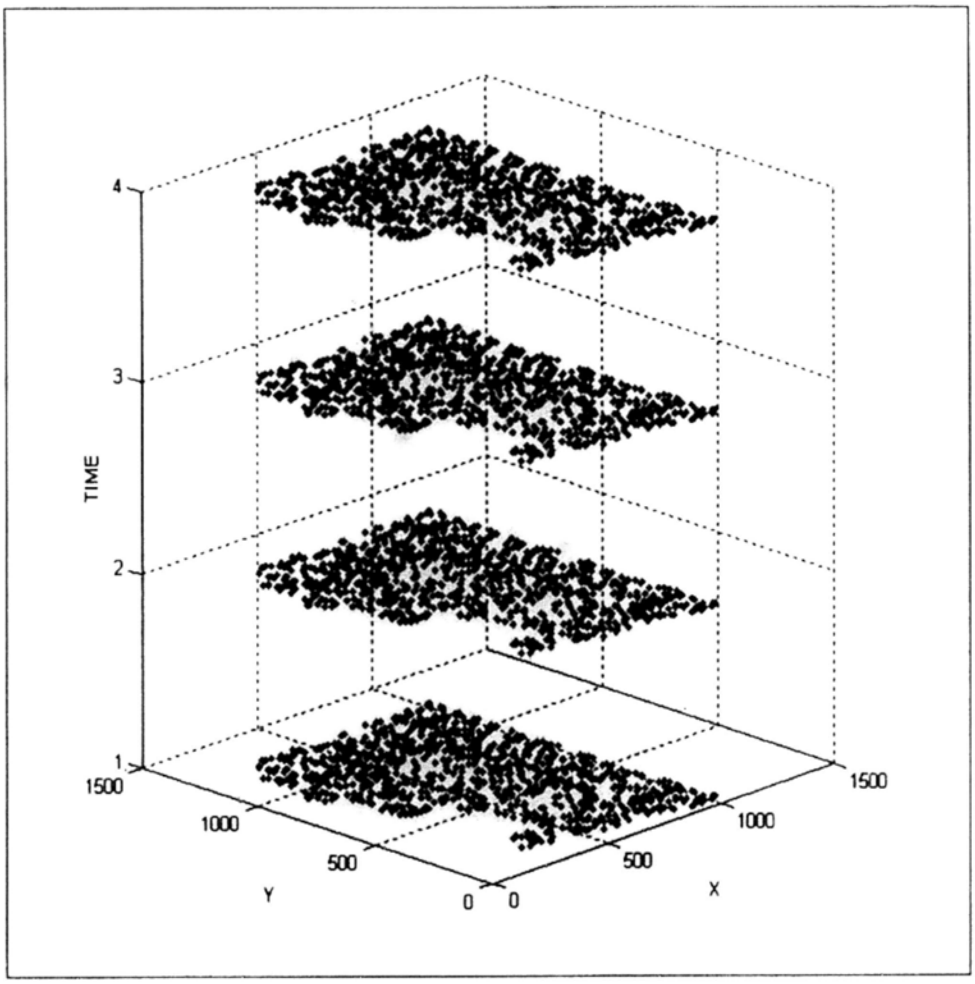


Figure 5.1: A sample of panel data set

Hsiao (2003) describes several major advantages of panel data sets over conventional cross-sectional or time-series data sets. Panel data usually offer the researchers a large number of data points. This serves to increase the degrees of freedom and reduce the collinearity among explanatory variables, hence improving the efficiency of econometric estimates. More importantly, panel data allows a researcher to analyze individual effect.

5.2.2 Panel data model

Panel data model, typically handling all the time series and cross-sectional data, can be employed for spatio-temporal analysis (Greene, 2000; Wooldrige, 2002). Grid cells in panel data involve a minimum of two dimensions (a cross-sectional dimension, indicated by subscript i, and a time series dimension, indicated by subscript t). The formula of the panel data model is:

$$y_{it} = X_{it}'\beta + u_{it}, i = 1,2,\cdots,N, t = 1,2,\cdots,T \tag{5.1}$$

where X_{it} and β are vectors of explanatory variables and parameters respectively and u_{it} is an error component whose assumption is different from the classic assumption in linear regression. The panel data model in one way error component assumes the following form:

$$u_{it} = \mu_i + v_{it} \tag{5.2}$$

where $\mu_i \sim IID(0,\sigma_\mu^2)$, $v_{it} \sim IID(0,\sigma_v^2)$. μ_i is used to identify unobserved individual effect. If μ_i is assumed to be a constant parameter to be estimated, the model will become panel data model with fixed effect. To account for spatio-temporal autocorrelation by using a direct representation in error component in land use change analysis, the error part should be allowed to vary randomly across land use choices. This idea led to panel data model with random effects where choice probabilities for repeated observations on the same individual share the same unobserved random effects (Train, 2003).

5.2.3 Estimation of panel data model with random effect: Maximum Simulated Likelihood Estimation (MSLE)

As there is no closed form solution to the marginal likelihood for the panel data model with random effect, simulation based methods are used to integrate the random latent heterogeneity term: error part. This procedure is similar to the maximum likelihood (ML) except that simulated probabilities are used instead of the exact probabilities.

In maximum likelihood, the log-likelihood function is as follows:

$$LL(\beta) = \sum_i \ln P_i(\beta) \tag{5.3}$$

where β is a vector of parameters, $P_i(\beta)$ is the (exact) probability of the observed choice of observation i, and the summation is over a sample. β can be estimated by maximum the log-likelihood function. If let $\hat{P}_i(\beta)$ be a simulated approximation to $P_i(\beta)$, the simulated log-likelihood function is:

$$SLL(\beta) = \sum_i \ln \hat{P}_i(\beta) \tag{5.4}$$

β can also be estimated by maximum the simulated log-likelihood function. The properties of MSL were derived by Gourieroux and Monfort,(1993). Train (2003) and Greene (2000) provide the technical details of the likelihood function and methods to maximize the likelihood function based on simulation techniques.

5.3 Methodology

5.3.1 Formulation of STLM

The general form of the multinomial logit model (MNLM) used for land use change analysis is as follows:

$$prob(y_i^j = 1) = \frac{\exp(\beta^j x_i)}{\sum_{j=0}^{J} \exp(\beta^j x_i)} \quad ,j = 0,\ldots,J \tag{5.5}$$

where y_i^j is an indicator of the observed land use type j for cell i: if yes, y_i^j equals 1; if no, y_i^j equals 0. J is the set of all the land use types. In this case, j=0 denotes undeveloped, 1 denotes residential, 2 denotes industrial, and 3 denotes transportation /commercial/others. β^j and x_i are the vectors of parameters for type j and explanatory variables for cell i respectively. x_i refers to the variables of driver factors listed in table 3.1 for cell i in this analysis.

$$\beta^j x_i = \beta_1^j \text{distance to Commercial Center}_i + \beta_2^j \text{distance to Financial Center}_i + \cdots + \beta_{13}^j \text{No Planning}_i$$

(5.5)

In order to reduce the number of parameters of the model, a value of 0 is assigned to all the variables of undeveloped land (base category), i.e., $\beta^0 = 0$.

The proposed model also aims to incorporate individual effect. Following the framework of panel data model, Equation (5.5) is revised as follows to identify the individual effect of a sampled cell:

$$prob(y_i^j = 1) = \frac{\exp(\beta^j x_i + \mu_i^j)}{\sum_{j=0}^{J} \exp(\beta^j x_i + \mu_i^j)}, j = 0, \ldots, J \tag{5.7}$$

where μ_i^j represents an individual random effect. y_i^j is an indicator of the observed land use type for cell *i*.

Considering the context of land use change, a difference in land use exists between the built-up area and undeveloped area. μ_i^j is assumed to be normally distributed with mean $\alpha_0^j + \alpha_1^j\left(y_{i0}^{j=1} + y_{i0}^{j=2} + y_{i0}^{j=3}\right)$ and a standard deviation δ^j as the following formula:

$$\mu_i^j = \alpha_0^j + \alpha_1^j\left(y_{i0}^{j=1} + y_{i0}^{j=2} + y_{i0}^{j=3}\right) + \delta^j \varepsilon_i^j \tag{5.8}$$

where ε_i^j is assumed to be independent and identically distributed (*iid*) *Standard Normal Distribution* across the land use types and observations. α_0^j, α_1^j and δ^j are the parameters to be estimated. y_{i0}^j is the indicator of the observed land use type *j* for cell *i* in the first year 1990. In Equation (5.8) , if ε_i^j is assumed to be *iid* across land use types and observations, this model will then be similar with an MNLM for panel data with random effects (Greene, 2000).

Spatio-temporal autocorrelation exists in the change process and it should be incorporated into the proposed model. Thus, individual random effect is used to identify the spatio-temporal autocorrelation among cells, which means the covariance matrix of ε_i^j is no longer assumed to be independent and identically distributed (*iid*) *Standard Normal Distribution*. The autocorrelation between cells *i* and *m* is assumed to be inversely proportional to the spatio-temporal distance between them. This is expressed as follows:

$$corr\left(\varepsilon_i^j, \varepsilon_m^j\right) \propto \exp\left(\frac{-d_{im}^{ST}}{\eta}\right) \tag{5.9}$$

where d_{im}^{ST} is the spatio-temporal distance between cells *i* and *m*. η is a scale factor. If there are a total of *n* observations, then the covariance matrix can be defined as follows:

$$corr(\varepsilon^j, \varepsilon^j) = \begin{pmatrix} \exp(\frac{-d_{11}^{ST}}{\eta}) & \exp(\frac{-d_{12}^{ST}}{\eta}) & \cdots & \exp(\frac{-d_{NN}^{ST}}{\eta}) \\ \exp(\frac{-d_{21}^{ST}}{\eta}) & \exp(\frac{-d_{22}^{ST}}{\eta}) & \cdots & \exp(\frac{-d_{2n}^{ST}}{\eta}) \\ \vdots & \vdots & \ddots & \vdots \\ \exp(\frac{-d_{n1}^{ST}}{\eta}) & \exp(\frac{-d_{n2}^{ST}}{\eta}) & \cdots & \exp(\frac{-d_{nn}^{ST}}{\eta}) \end{pmatrix} \tag{5.10}$$

The spatio-temporal weighted function used in GTWLM is adopted to identify $\exp(\frac{-d_{im}^{ST}}{\eta})$. Thus, $\exp(\frac{-d_{im}^{ST}}{\eta})$ can be divided into to part: $\exp(\frac{-d_{im}^{S}}{h_S^2})$ and $\exp(\frac{-d_{im}^{T}}{h_T^2})$.

$$\exp(\frac{-d_{im}^{ST}}{\eta}) = \exp(\frac{-d_{im}^{S}}{h_S^2}) * \exp(\frac{-d_{im}^{T}}{h_T^2}) \tag{5.11}$$

Where d_{im}^{S} and d_{im}^{T} are the spatial and temporal distance respectively. The same h_S^2 and h_T^2 as GTWLR are used here.

Overall, the ST-PLM's log-likelihood function is formulated as follows:

$$\ln(L)=\sum_{i=1}^{n}\sum_{j=0}^{J} y_i^j \ln(p(y_i^j=1|x_i;u_i^j,\beta^j)f(\mu_i^j|y_{i0},\varepsilon_i^j)) \tag{5.12}$$

According to equation (5.9), μ_i^j can be obtained by parameter ω^j, which consists of α_0^j, α_1^j and δ^j. Equation (5.12) can be rewritten as follows:

$$\ln(L(\beta^j,\omega^j))=\sum_{i=1}^{n}\sum_{j=0}^{J} y_i^j \ln\left[\int_{\mu_i^j}\left\{\frac{\exp(\beta^j x_i+\mu_i^j)}{\sum_{j=0}^{J}\exp(\beta^j x_i+\mu_i^j)}\right\}g(\mu_i^j|\omega^j)\partial\mu_i^j\right] \tag{5.13}$$

Theoretically, the parameters of the model are estimated by maximizing the log-likelihood function. The computation of the probabilities for each land use type has posed significant problems in this context. The log-likelihood function above involves the estimation of an integral, which cannot be evaluated analytically for its lack of a closed-form solution. Considering the speed and precision, simulation techniques (Train, 2003) have been used to estimate the log likelihood function. The integrals in the choice probabilities are approximated by Monte Carol technique. Then, the resulting simulated log-likelihood function is maximized. Therefore, maximum simulated likelihood (MSL) is adopted (same as maximum likelihood (ML) except that the simulated probabilities are used instead of the exact probabilities).

Initially, 3 (for j=1, 2 and 3) n-dimensional normally distributed random vectors should be generated with the autocorrelation matrix given by Equation (5.9). In order to obtain the vectors, Equation (5.9) should be decomposed to a lower-triangular matrix to be multiplied by n-dimensional normally distributed random vectors. Commonly, the Cholesky decomposition is adopted for simulating systems with multiple correlated variables (the inter-variable autocorrelations matrix is decomposed to obtain the lower-triangular L. Applying this to a vector of uncorrelated simulated shocks, v produces a shock vector L*v with the covariance properties of the system modeled). However, the autocorrelation matrix among observations can only be guaranteed to be symmetric, not positive definite. This renders the application of Cholesky decomposition to be impossible. Instead, the generalized Cholesky decomposition method devised by Gill and King (2004) is used. In the Cholesky decomposition, when a symmetric matrix is not positive definite, in order to produce a positive definite matrix, a non-negative diagonal matrix with element values as small as possible is added to the original matrix. For more details about generalized Cholesky decomposition, please refer to Appendix A.

The simulated probabilities for this case are obtained from the following steps:

Step 1) Decompose the covariance matrix into a triangular matrix by the generalized Cholesky decomposition method.
Step 2) Obtain *3* (for j=1, 2 and 3) sequences of draws following the *iid* normal distribution by shuffled Halton draws (please check Appendix B for details). Each sequence consists of n number for each cell.
Step 3) Multiply the lower-triangular matrix by n *iid* normal values and obtain 2 new sequences.
Step 4) Calculate the equation (5.8) according to ε_i^j obtained from the new sequences.
Step 5) Repeat steps 2-4 for 100 times and average the results.

The integrals in the choice probabilities are approximated by the steps. Then, the resulting simulated log-likelihood function is maximized. The model's parameters are estimated finally.

5.3.2 Evaluation of the ST-PLM's improvement: Akaike's Information Criterion (AIC)

Akaike's information criterion is grounded in the concept of entropy, effectively offering a relative measure of the information lost when a given model is used to describe reality. This describes the tradeoff between bias and variance in model construction or that between precision and complexity of the model.

In the general case, the AIC could be calculated as follows:

$$AIC = 2k - 2\ln(L) \tag{5.14}$$

where k is the number of parameters in the model, and L is the maximized value of the likelihood function for the estimated model. Given a data set, several competing models may be ranked according to their AIC (lowest AIC representing the best). The AIC methodology attempts to find the model that best explains the data with a minimum of free parameters. In this case, AIC is employed to decide which model is better (i.e., MNLM vs. ST-PLM).

5.4. Results and discussions

The following section explains the model validation in this case. The last five years map data (1996, 2000, 2002, 2004, 2006 and 2008) are used for the estimation of the model. The data of the year 1990 are employed to construct the mean of μ_{it}^{l}. For each sample cell, the probabilities on all the land use types are computed with the spatio-temporal panel logit model. Then, the probabilities are compared and the land use type with the highest probability is selected for the cell. Table 5.1 presents the parameters generated by ST-PLM using the MSL estimation developed in this study. The log likelihood is -2323.4 and PCP is 79.4%.

Table 5.1: Estimation results of ST-PLM

Parameters	Coef.(Residential)	t-stats	Coef.(Industrial)	t-stats	Coef.(Commercial/Transportation/others)	t-stats
Distance to Commercial Centre	-2.08e-4	-9.62e-1	-2.37e-4	-1.12	3.13e-4	2.43
Distance to Financial Centre	-2.20e-4	-1.60	-7.41e-05	-6.09e-1	-4.74e-05	-5.95e-1
Distance to Industrial Centre	5.20e-05	2.55e-1	-9.65e-3	-1.36e01	1.35e-4	0.824
Distance to Educational Facilities	-7.89e-4	-2.73	1.39e-3	5.13	-1.21e-4	-0.697
Distance to Railway Infrastructure	-7.07e-05	-1.81	-1.79e-4	-3.84	-9.06e-05	-2.90
Distance to Road	-2.44e-3	-3.90	2.17e-4	5.81e-1	-1.20e-3	-4.64
Population	4.04e-05	3.03	4.08e-05	2.50	5.10	4.40
DEM	-1.01e-2	-2.81	-1.89e-2	-3.09	-9.07e-3	-5.33
Slope	-6.67e-2	-5.01	-8.91e-2	-5.16	-5.62e-2	-7.37
Planning for Residential	2.24	7.87	1.10	3.12	8.91e-1	3.75
Planning for Industrial	1.29e01	7.22e-2	1.28e01	7.17e-2	1.29e01	7.22e-2
Planning for Transportation/commercial/others	-1.21e-1	-4.43e-1	1.10	3.53	8.23e-1	4.47
No Planning	-9.23e-3	-3.06e-2	1.25e-1	3.67e-1	3.41e-1	1.83
α_0	4.09e-1	1.20	1.10	2.70	-2.51e-1	-1.00
α_1	5.62e-1	2.65	2.21e-1	8.95e-1	2.89	1.64e01
δ'	1.44	5.39e-1	-6.12	-1.81	4.73	2.06
Log likelihood function	-2.3234e+03					
PCP	79.4%					

It can be found that δ', which represents spatio-temporal autocorrelation, is significant for industrial and Commercial/Transportation/others with a large absolute t statistics. The positive parameters of individual effect for all the built-up area reveals built-up land use types have persistently feature.

The comparison between MNLM and ST-PLM is used to examine whether the proposed ST-PLM offers any improvement. Table 5.2 compares the accuracies achieved by MNLM and ST-PLM. ST-PLM incorporates the spatio-temporal autocorrelation and individual effect in one land use data set, whereas MNLM ignores it. Consequently, the spatio-temporal model achieves a higher overall PCP (79.4%) than MNLM (74.1%). Also, the accuracies for undeveloped, industrial, and commercial/ transportation/ others achieved by ST-PLM are better (i.e., 86.4% vs. 90.4%; 50.7% vs. 56.0%; 64.1% vs. 76.6%). Since the distribution of residential is dispersed, the consideration of spatio-temporal autocorrelation is not viable for this land use type. Thus, the accuracy for residential in MNLM is lower than that for ST-PLM.

Table 5.2: Comparison of PCPs between MNLM and ST-PLM

Observed	MNLM (Multinomial logit model)						ST-PLM (Spatio-temporal panel logit model)					
	Predicted				Total	PCP	Predicted				Total	PCP
	0	1	2	3			0	1	2	3		
0	1726	36	9	206	1997	86.4%	1806	35	15	121	1997	90.4%
1	33	352	8	153	546	64.5%	54	305	26	161	546	55.9%
2	27	24	152	97	300	50.7%	28	22	168	82	300	56.0%
3	178	274	42	883	1377	64.1%	187	86	49	1055	1377	76.6%
Total	1964	686	211	1339	3113	74.1%	2075	448	258	1419	3334	79.4%

To further examine if ST-PLM demonstrates significant improvements over MNLM, McNamara's test and AIC test are performed. For McNamara's test, the Z value calculated is -8.3590 (its absolute value is greater than 1.96). Hence, it can be concluded that at 5% level of significance, the difference in the accuracies between the two models is statistically significant. In other words, ST-PLM outperforms MNLM notably. For AIC test, the values of MNLM and ST-PLM are 5424 and 4678.8 respectively. This reveals that the ST-PLM better explains the data with lesser free parameters.

The Kno in table 5.3 evinces that the ST-PLM achieves a better result than MNLM consistently. Specifically, ST-PLM's ability to identify location is better than MNLM's, and the two models have the same ability when it comes to specifying quantity.

Table 5.3: Comparison of MNLM and ST-PLM with Kappa coefficients

	MNLM	ST-PLM
Kno	0.65	0.73
Klocation	0.63	0.71
Kquantity	0.90	0.90

5.5 Conclusion

Over the years, spatial autocorrelation has been seriously considered in the context of land use change, which is inherently a spatio-temporal process. A natural step forward involves incorporating both spatial and temporal autocorrelation to analyze the change.

This chapter has presented such an endeavor using ST-PLM, which provides a powerful option to establish the relationship between the explanatory variables identified and the land use type. Notably, the proposed model effectively explores the spatial and temporal autocorrelation. In the proposed model, the spatio-temporal autocorrelation is considered in the random effect component ε_{it}^{j} with an assumption that the autocorrelation between ε_{it}^{j} is inversely proportional to the spatio-temporal distance between them. The proposed model has been validated using the special economy zone (SEZ), Shenzhen land use data. The case study demonstrates that the proposed model can improve the PCP as well as the accuracy. McNamara's test and AIC test were performed, which corroborate the superior performance of ST-PLM over the MNLM. Besides, the

Kno also shows that ST-PLM consistently provides a better result than MNLM. Specifically, ST-PLM's ability to specify location is better than MNLM's, and the two models have the same ability in specifying quantity.

As demonstrated by the model developed in this study, individual effect and spatio-temporal autocorrelation should be integrated when studying the dynamics of land use. A rigorous structure for spatio-temporal autocorrelation in conjunction with a solid statistical model can serve as an effective alternative for understanding land use change patterns.

5.6 Chapter Summary

In this chapter, two improvements (spatio-temporal autocorrelation and individual effect) in the multinomial logit model were discussed initially. Then, the panel concept was introduced. Based on the panel data analysis framework, ST-PLM was developed. The implementation and the evaluation of proposed model were discussed in detail. Subsequently, results showing the enhanced performance of the model were discussed. Finally, the conclusion summarizing the proposed model was provided.

Chapter 6 Generalized Spatio-temporal logit Model

Non-stationarity and autocorrelation are two closely-related issues. Many evidences point to the existence of non-stationarity in the presence of autocorrelation and vice-versa. It is hence customary to consider both of them in a single model. Besides, individual effect is also very important issue in land use change modeling. An ideal model for land use change study should solve the existing problems as much as possible. As discussed earlier, the GTWLM and ST-PLM can handle those issues respectively. Thus, the integration of GTWLM and ST-PLM hold immense potential in solving the major challenges in this domain.

6.1 Introduction

Statistically modeling quantitative relationships between response variable and explanatory variables for spatio-temporal data involves two important problems. The first problem is that spatial or temporal non-stationarity occurs in the relationships being modeled. The second problem involves the spatial or temporal autocorrelation that exists between the observations. Traditional models such as multinomial logit model used in land use change modeling have largely ignored these two issues that violate the basic assumptions. Spatio-temporal non-stationarity violates the assumption that a single relationship exists across the sample data observations, while spatio-temporal autocorrelation violates the assumption that explanatory variables are fixed in repeated sampling.

Therefore, alternative approaches have been developed to model aforementioned. For non-stationarity, one of the well known models is the nonparametric local linear regression introduced by Mcmillen (1996) and Mcmillen and McDonald (1997). Brunsdon *et al.* (1996) labeled them geographically weighted regression (GWR). GWR allows the exploration of the variation of the parameters, and therefore received considerable attention (Brunsdon *et al*, 1996; Fotheringham and Brunsdon, 1999;

Fotheringham *et al*, 2001; Huang and Leung ,2002; Yu and Wu, 2004). For autocorrelation, Ord (1975) proposed the use of autoregressive and moving average terms in regression models to account for spatial autocorrelation in the response variable and the residuals, respectively. Cressie (1993) addressed spatial effects by modeling the residual variance-covariance matrix directly. Chapters 5 and 6 discussed the extension of the models into a spatio-temporal framework and developed geographically and temporally weighted logit model (GTWLM) and spatio-temporal panel logit model (ST-PLM) to deal with spatio-temporal non-stationarity and autocorrelation in land use change modeling, respectively.

However, most of previous work deals with aforementioned problems on an individual basis. Even though the two problems are theoretically distinct, many evidences show there exists non-stationarity in the presence of autocorrelation, and an inadequate model that fails to capture non-stationarity will result in residuals that exhibit autocorrelation. Lesage (2005) argued for covering both spatial and temporal heterogeneity and autocorrelation effects in a mixed model. In fact, non-stationarity and autocorrelation share some similarities. For instance, both depend on the definition of spatio-temporal weighting matrices and both deal with spatio-temporal dependencies in data. Besides, the spatio-temporal weighting function in GTWLM and spatio-temporal error autocorrelation matrix adopt the same form to quantify the notion of proximity. Although spatio-temporal non-stationarity and autocorrelation are often related in the context of modeling, only few studies that jointly construct and estimate both of them simultaneously could be found. Therefore building an integrated model with spatio-temporal non-stationarity and autocorrelation effects assumes research significance. Brunsdon *et al.* (1998) attempted to model both spatial non-stationarity and spatial autocorrelation in a complicated process. They proposed a geographically weighted regression with spatially lagged objective variable mode (GWRSL) to exploit the spatial variations. But their model ignores temporal information and individual effect.

This chapter ascertains the links between the two problems and focuses on integrating

GTWLM and ST-PLM. A generalized spatio-temporal logit model (GSTLM) is developed to utilize the spatio-temporal effect in land use change in SEA, Shenzhen.

6.2 Generalized Spatio-Temporal Logit Model (GSTLM)

The GSTLM is expressed as follows:

$$\log\left(\frac{prob(y_i^j=1)}{prob(y_i^0=1)}\right)=\sum_p \beta_p^j(u_i,v_i,t_i)X_{ip}+\mu_i^j, j=0,\dots,J \quad (6.1)$$

where y_i^j is an indicator of the observed land use type j for cell i: if yes, y_i^j equals 1; if no, y_i^j equals 0. J is the set of all the land use types. Herein, j=0 denotes undeveloped, 1 denotes residential, 2 denotes industrial, and 3 denotes transportation /commercial/others. $\beta_p^j(u_i,v_i,t_i)$ are p continuous functions of the spatio-temporal coordinates (u_i,v_i,t_i) in the study area. X_{ip} refers to the p variable of driver factors listed in table 3.1 for cell i in this analysis. Following the form of ST-PLM, μ_i^j, which represents the individual random effect, is assumed to be normally distributed. The mean $\alpha_0^j+\alpha_1^j\left(y_{i0}^{j=1}+y_{i0}^{j=2}+y_{i0}^{j=3}\right)$ and standard deviation δ^j are represented as follows:

$$\mu_i^j=\alpha_0^j+\alpha_1^j\left(y_{i0}^{j=1}+y_{i0}^{j=2}+y_{i0}^{j=3}\right)+\delta^j\varepsilon_i^j \quad (6.2)$$

where y_{i0}^j is the indicator of the observed land use type j for cell i in the first year 1990. To identify the spatio-temporal autocorrelation among cells, the covariance matrix of ε_i^j is no longer an identity matrix. The autocorrelation between cell i and cell m is assumed to be inversely proportional to the spatio-temporal distance between them. This can be expressed as follows:

$$corr\left(\varepsilon_i^j, \varepsilon_m^j\right) \propto \exp(\frac{-d_{im}^{ST}}{h_{ST}}) \tag{6.3}$$

where h_{ST} is bandwidth to be estimated, d_{im}^{ST} is the spatio-temporal distance between cells i and m,

$$d_{im}^{ST} = \sqrt{\lambda((u_i - u_m)^2 + (v_i - v_m)^2) + \mu(t_i - t_m)^2} \tag{6.4}$$

λ and μ are scale factors to balance the different effects used to measure the spatial and temporal distance in their respective metric systems. If there are n observations in total, then the covariance matrix can be expressed as follows:

$$corr(\varepsilon^j, \varepsilon^j) = \begin{pmatrix} \exp(\frac{-d_{11}^{ST}}{h_{ST}}) & \exp(\frac{-d_{12}^{ST}}{h_{ST}}) & \cdots & \exp(\frac{-d_{1n}^{ST}}{h_{ST}}) \\ \exp(\frac{-d_{21}^{ST}}{h_{ST}}) & \exp(\frac{-d_{22}^{ST}}{h_{ST}}) & \cdots & \exp(\frac{-d_{2n}^{ST}}{h_{ST}}) \\ \vdots & \vdots & \ddots & \vdots \\ \exp(\frac{-d_{n1}^{ST}}{h_{ST}}) & \exp(\frac{-d_{n2}^{ST}}{h_{ST}}) & \cdots & \exp(\frac{-d_{nn}^{ST}}{h_{ST}}) \end{pmatrix} \tag{6.5}$$

In line with GTWLM, α_0^j, α_1^j and δ^j , which are the fixed parameters to be estimated, is also allowed to vary over space and time as shown below:

$$\mu_{it}^j = \alpha_0^j(u_i, v_i, t_i) + \alpha_1^j(u_i, v_i, t_i)\left(y_{i0}^{j=1} + y_{i0}^{j=2} + y_{i0}^{j=3}\right) + \delta^j(u_i, v_i, t_i)\varepsilon_i^j \tag{6.6}$$

Thus, equation 6.1 can be revised as follows:

$$\log\left(\frac{prob(y_i^j = 1)}{prob(y_i^0 = 1)}\right) = \sum_p \beta_p^j(u_i, v_i, t_i)X_{ip} + \alpha_0^j(u_i, v_i, t_i) + \alpha_1^j(u_i, v_i, t_i)\left(y_{i0}^{j=1} + y_{i0}^{j=2} + y_{i0}^{j=3}\right)$$
$$+ \delta^j(u_i, v_i, t_i)\varepsilon_i^j, \qquad j = 0, \ldots, J$$

(6.7)

$\alpha_0^j(u_i, v_i, t_i)$ can been see perceived as the parameter of constant. Let $X^{\bullet}_{ip}$ denotes the vector [X_{ip}, 1, $\left(y_{i0}^{j=1}+y_{i0}^{j=2}+y_{i0}^{j=3}\right)$, ε_{it}^{j}]. In order to estimate $\beta_p^j(u_i, v_i, t_i)$, $\alpha_0^j(u_i, v_i, t_i)$, $\alpha_1^j(u_i, v_i, t_i)$ and $\delta^j(u_i, v_i, t_i)$ for each point, $X^{\bullet}$ should be transferred into $\tilde{X}^{\bullet}$ in GSTLM as follows.

$$\tilde{X}^{\bullet}=W^{ST}\left(u_i, v_i, t_i\right)^{1/2} X^{\bullet} \tag{6.8}$$

where $W^{ST}\left(u_i, v_i, t_i\right)$ is an *n* by *n* diagonal matrix:

$$W^{ST}\left(u_i, v_i, t_i\right)=\begin{pmatrix} w_{i1}^{ST} & 0 & \cdots & 0 \\ 0 & w_{i2}^{ST} & \cdots & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{i3}^{ST} \end{pmatrix} \tag{6.9}$$

Based on d^{ST}, the elements of $W^{ST}\left(u_i, v_i, t_i\right)$, w_{im}^{ST} are defined. The same issue arises in the estimation of GSTLM as GTWLM. On the one hand, the estimation of ST-PLM, involving non-linear optimization and MSL estimation, is computationally intensive. . Furthermore, if the GSTLM parameters are to be computed at *n* regression points, then the above procedures must be repeated *n* times! Thus, when parameters are estimated for every point, sub-samples are chosen by using a computational ruse that ignores observations with negligible weight to reduce the computational burden. On the other hand, a sub-sample with values of all single types cannot be produced. Hence, adaptive weighting functions are employed to adapt themselves in size to ensure that the same and sufficient non-zero weights are used for each in the analysis. The specific adaptive weighting function is as follows:

$$w_{im}^{ST} = \begin{cases} \exp\left\{-(\frac{(d_{im}^{ST})^2}{h_{ST}^2})\right\}, & if\ m \text{ is the } q \text{ nearest points around } i \\ 0, & otherwise \end{cases} \quad (6.10)$$

where h_{ST} is same spatio-temporal bandwidth as equation 6.5. q represents the number or proportion of observations to consider in the estimation of regression at each location to reduce the number of sub-samples and make enough so that the regression can work. Substituting 6.4 in 6.5 and 6.10 results in the following:

$$\begin{aligned} w_{im}^{ST} &= \exp\left\{-(\frac{\lambda[(u_i-u_m)^2+(v_i-v_m)^2]+\mu(t_i-t_m)^2}{h_{ST}^2})\right\} \\ &= \exp\left\{-(\frac{(u_i-u_m)^2+(v_i-v_m)^2}{h_S^2}+\frac{(t_i-t_m)^2}{h_T^2})\right\} \\ &= \exp\left\{-((d_{im}^S)^2/h_S^2+(d_{im}^T)^2/h_T^2)\right\} \\ &= \exp\left\{-(d_{im}^S)^2/h_S^2\right\}\times\exp\left\{-(d_{im}^T)^2/h_S^T\right\} \\ &= w_{im}^S\cdot w_{im}^T \end{aligned} \quad (6.11)$$

where $w_{im}^S = \exp\left\{-(d_{im}^S)^2/h_S^2\right\}$, $w_{im}^T = \exp\left\{-(d_{im}^T)^2/h_S^T\right\}$, $(d_{im}^S)^2 = (u_i-u_m)^2+(v_i-v_m)^2$, $(d_{im}^T)^2 = (t_i-t_m)^2$, $h_S^2 = \frac{h_{ST}^2}{\lambda}$ and $h_T^2 = \frac{h_{ST}^2}{\mu}$ are the spatial and temporal bandwidths, respectively. Since the weighting function is a diagonal matrix, whose diagonal elements are multiplied by $w_{im}^S\cdot w_{im}^T$ $(1\le m\le n)$, W^{ST} can be seen as the combination of spatially weighted matrix W^S and temporally weighted matrix W^T: $W^{ST} = W^S\times W^T$. Hence, if μ was set to 0, GSTLM becomes a generalized spatial logit model (GSLM) which only considers spatial effect. If λ was set to 0, GSTLM becomes a generalized temporal logit model (GTLM) which only considers temporal effect.

In order to reduce the parameters to be estimated in the model, the parameter ratio $\tau = \mu/\lambda$ is introduced. It can be considered as the balance ratio between spatial distance effect and temporal distance effect. As discussed in chapter 4, setting $\lambda = 1$ will not change the results estimated by optimizing λ, μ and h_{ST} using cross validation in terms of PCP. Hence, w_{im}^{ST} can be expressed as follows:

$$w_{im}^{ST} = \exp\left\{-\left(\frac{[(u_i - u_m)^2 + (v_i - v_m)^2] + \tau(t_i - t_m)^2}{h_{ST}^2}\right)\right\} \tag{6.12}$$

Here, the estimation results will provide τ and h_{ST} instead of λ, μ and h_{ST} in this study. In GSTLM, CV method is used to choose an appropriate bandwidth h_{ST} and sub-sample size q by maximum PCP. The weighting matrix is decided subsequently. Following that, the parameters in each point can be estimated by the local log-likelihood function formulated as follows:

$$\ln(L(u_i,v_i,t_i)) = \sum_{i=1}^{N}\sum_{j=0}^{J} y_{it}^{j} \ln(p(y_i^j = 1 \mid \tilde{X}^*_i;\ \beta_p^j(u_i,v_i,t_i),\ \alpha_0^j(u_i,v_i,t_i), \alpha_1^j(u_i,v_i,t_i), \delta^j(u_i,v_i,t_i))) \tag{6.13}$$

Theoretically, the parameters of the local model are estimated by maximizing the log-likelihood function. However, that ε_i^j in $\tilde{X}^*_i$ is assumed to be random complicated the computation of the probabilities for each land use type. The local log-likelihood function above involves the estimation of an integral, which cannot be evaluated analytically owing to its lack of a closed-form solution. Therefore, maximum simulated likelihood (MSL) is adopted (same as maximum likelihood (ML), except that the simulated probabilities are used instead of the exact probabilities). Generalized Cholesky decomposition is also employed to decompose the matrix given by equation (6.5).

The estimations for GSTLM in each point are obtained from y the following steps:

Step 1) Decompose the covariance matrix into a triangular matrix by the generalized Cholesky decomposition method.

Step 2) Obtain 3 (for j=1, 2 and 3) sequences of draws following the *iid* normal distribution by shuffled Halton draws. Each sequence consists of n number for each cell.

Step 3) Multiply the lower-triangular matrix by n *iid* normal values and get 2 new sequences.

Step 4) Calculate the equation (6.13) according to ε_i' given by the new sequences.

Step 5) Repeat steps 2-4 100 times and average the results.

Step 6) Maximize the simulated log-likelihood function and get parameters.

6.3 Results and Disscussion

To check the applicability of GSTLM and analyze the land use change in SEZ, Shenzhen, a study was implemented using the spatio-temporal sample set between 1990 and 2008.

The accuracy and parameter estimate result are reported in Table 6.1 and 6.2 respectively. The PCP of GSTLM reaches a value of 85.9%, which is higher than MNLN, GTWLM and ST-PLMB. However, the PCP improvement of GSTLM on MNLM is less than the sum of the improvement on MNLM by GTWLM and ST-PLM. (MNLM: 74.1%; GTWLM: 82.3%; ST-PLM: 79.4%) That's because allowing for non-stationarity in the regression parameters can account for at least some, and possibly a large part, of the autocorrelation in error terms in a global model calibrated with spatio-temporal data. Also, the accuracies for all the land use types achieved by GSTLM are balanced. As the output of local parameter estimates from GSTLM would be voluminous, Table 6.2 provides a 3-number summary of the distribution of each parameter to indicate the extent of the variability. In this case, the impact of those variables varies spatially, and it indicates that local effects do exist.

Table 6.1: The accuracy of GSTLM

Observed	Predicted				Total	PCP
	0	1	2	3		
0	1844	19	7	107	1977	93.3%
1	31	354	14	147	546	64.8%
2	4	6	255	35	300	85.0%
3	124	77	21	1155	1377	83.9%
Total	2003	456	297	1444	4200	85.9%

Table 6.2: GSTLM parameter estimate summaries

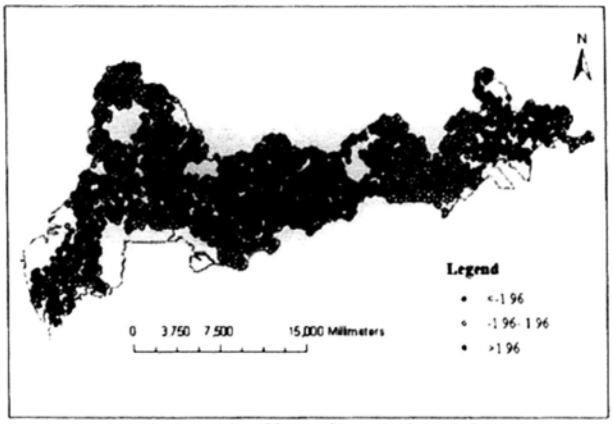
GSTLM (bandwidth= 0.6633; τ = 0.1826; q=1130)									
	Residential			**Industrial**			**Transportation/commercial/others**		
Parameter	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**
Distance to Commercial Centre	-3.79e-02	-1.13e-03	1.22e-02	-2.62e-02	-2.16e-03	9.95e-03	-1.87e-02	1.86e-03	1.09e-02
Distance to Financial Centre	-1.87e-02	-1.59e-03	1.92e-02	-1.65e-02	6.84e-05	2.69e-02	-5.11e-03	1.70e-04	1.50e-02
Distance to Industrial Centre	-1.85e-02	-5.57e-04	2.34e-02	-3.49	-8.48e-02	-9.09e-03	-1.42e-02	2.96e-04	1.42e-02
Distance to Educational Facilities	-5.34e-02	-2.55e-03	9.70e-03	-2.50e-02	6.78e-03	6.42e-02	-1.40e-02	-6.95e-04	1.79e-02
Distance to Railway Infrastructure	-6.41e-03	-6.09e-04	6.99e-03	-2.27e-02	-2.01e-03	1.05e-03	-8.49e-03	-7.38e-04	1.08e-03
Distance to Road	-7.01e-02	-1.45e-02	8.94e-03	-6.61e-02	7.26e-03	8.09e-02	-7.73e-02	-1.11e-02	1.14e-02
Population	-9.09e-04	2.17e-04	3.16e-03	-8.47e-04	3.01e-04	7.27e-03	-3.80e-04	4.19e-04	3.24e-03
DEM	-1.12	-1.18e-01	1.90e-02	-8.30e-01	-1.24e-01	1.28e-01	-3.99e-01	-6.11e-02	-2.80e-03
Slope	-3.45	-4.31e-01	5.70e-01	-4.53	-5.56e-01	4.15e-01	-2.64	-2.88e-01	5.45e-01
Planning for Residential	-1.14e01	1.10e01	9.60e02	-4.06e01	3.83	6.53e01	-1.10e01	7.16	5.53e01
Planning for Industrial	-2.13e02	1.27e02	5.16e02	-3.61e01	1.32e02	6.35e02	-2.53e01	1.34e02	5.46e02
Planning for Transportation/commercial/others	-2.91e01	-1.31e-01	3.85e01	-8.78	5.63	3.50e01	-7.44	4.70	2.81e01
No Planning	-4.31e02	-5.32e-01	5.83e01	-5.50e02	1.36	5.69e01	-4.09	3.42	3.12e01
α_s^l	-2.71e01	5.56	3.49e01	-3.83e01	13.8e01	6.21e01	-15.7e01	-1.10	1.65e01
α_t^l	-1.71e01	8.55	2.65e02	-2.30e01	4.96	2.69e02	6.19	2.22e01	2.99e02
δ^l	-1.28e02	4.79	2.33e02	-4.84e02	-5.25e01	3.69e02	-1.83e02	1.91e01	5.09e02
ll	-2.3898e+003								

Further, two problems remain to be considered. Firstly, it needs to be ascertained if it is necessary to consider individual effect and spatio-temporal autocorrelation in the model. Secondly, it needs to be found whether the GSTLM is truly better than MNLM, GTWLM, and ST-PLM.

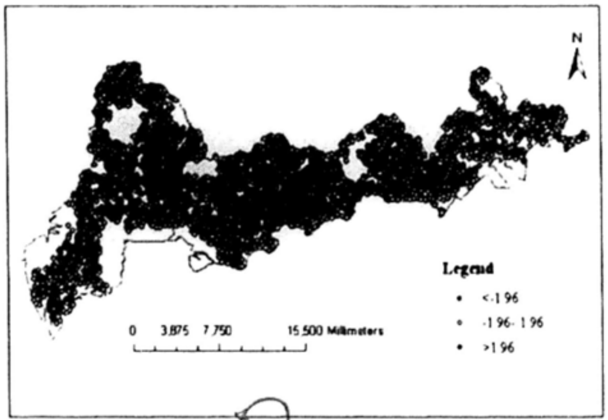
For the first problem, a pseudo t-statistic is calculated to indicate the significance of individual effect and spatio-temporal autocorrelation. This is obtained by dividing a parameter estimate by its standard error (Fotheringham *et al.*, 2001). If the absolute t-statistic value is greater than 1.96, it can be considered significant at the 95% confidence level. Table 6.3 lists the absolute t-stats for all the parameters. Figure 6.1 provides the spatial variation of the t- statistic for individual effect and spatio-temporal autocorrelation.

From table 6.3 it can be seen that the median value of absolute t-statistic for residential and transportation/commercial/others are 2.39 and 7.27 respectively. This implies that the individual effect is significant for most of the observations for residential and transportation/commercial/others, especially in Luohu and Futian districts. For the industrial land use type, the individual effect is primarily significant in the northwest of Yantian district.

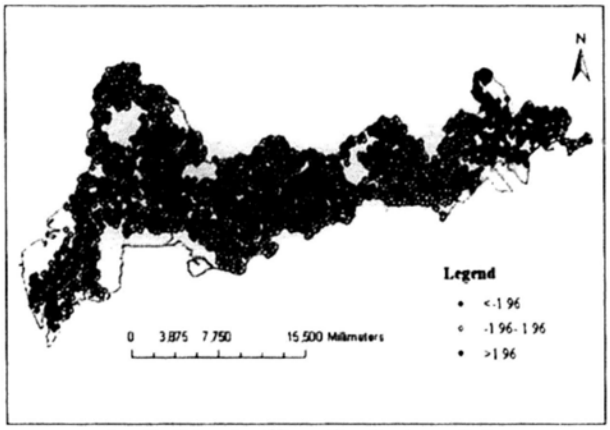
The smaller median value of the absolute t-statistic indicates that only part of the observations have significant spatio-temporal autocorrelation, whereas the others have weak spatio-temporal autocorrelation with the absolute t-statistic values (< 1.96). The median value of absolute t for residential is 0.82 (Table 6.3). This shows that the spatio-temporal autocorrelation of the tested data for residential is rather weak. Besides, the median values of absolute t for the other two land use types are about 1, which evinces that the spatio-temporal autocorrelation for this two land use types is not significant on some observations. One important reason is that only 700 observations were collected in each year and the temporal distribution of land use data is thin. Hence, the sample set is sparsely scattered among the spatio-temporal space. However, it is still reasonable to account for spatio-temporal autocorrelation in some observations where the t-statistic values are greater than 1.96. It can be deduced from figure 6.1 that significant spatio-temporal autocorrelation exists in the southeastern part of Nanshan district and east of Luohu for industrial land use type, and in the northeast of Yantian district. Furthermore, the spatio-temporal autocorrelation is also significant for transportation/commercial/others.



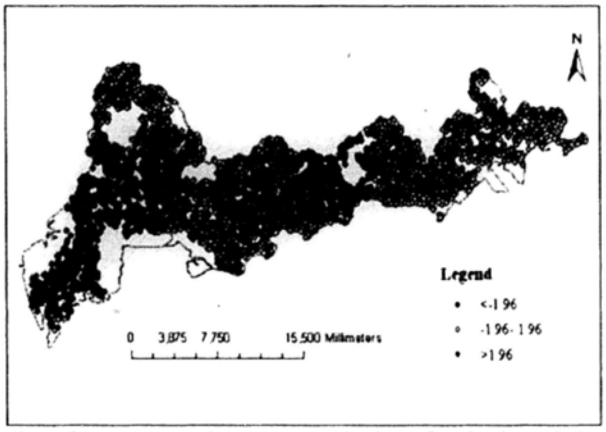
Individual effect for residential



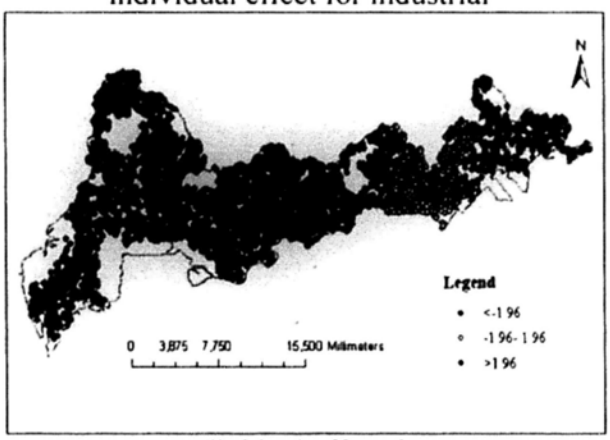
Sptiao-temporal autocorrelation for residential



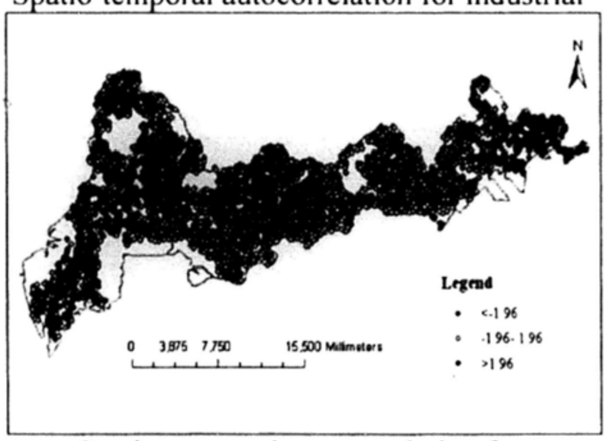
Individual effect for industrial



Spatio-temporal autocorrelation for industrial



Individual effect for Transportation/commercial/others



Spatio-temporal autocorrelation for Transportation/commercial/others

Figure 6.1: Spatial variation of t-stats for individual effect and spatio-temporal autocorrelation

Table 6.3: GSTLM absolute t estimate summaries

	Residential			Industrial			Transportation/commercial/others		
Parameter	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**
Distance to Commercial Centre	3.09e-03	1.19	5.68	1.96e-03	8.51e-01	3.22	6.74e-04	1.81	5.38
Distance to Financial Centre	1.12e-03	1.55	4.75	8.53e-04	1.10	4.08	1.38e-03	1.25	5.40
Distance to Industrial Centre	5.25e-04	1.11	4.95	2.05	7.05	10.0	3.37e-04	9.73e-01	4.07
Distance to Educational Facilities	1.39e-04	8.79e-01	6.09	7.85e-03	1.60	5.33	3.20e-03	1.54	4.10
Distance to Railway Infrastructure	1.10e-03	1.26	5.32	2.90e-04	2.01	6.99	3.68e-04	1.54	6.86
Distance to Road	4.03e-03	2.03	4.71	1.84e-04	1.52	5.14	5.59e-03	3.40	7.27
Population	3.68e-05	1.19	4.53	1.19e-03	1.21	5.73	2.59e-03	2.02	6.54
DEM	1.81e-03	2.30	5.47	1.12e-03	1.16	4.87	3.62e-02	2.52	6.67
Slope	3.94e-03	2.57	6.25	1.59e-03	2.06	4.26	2.30e-03	3.34	7.57
Planning for Residential	1.49e-03	3.10	6.80	1.06e-03	1.38	4.63	5.10e-04	2.00	5.94
Planning for Industrial	**4.65e-07**	**4.81e-02**	**8.02e-02**	**1.72e-04**	**4.87e-02**	**9.99e-02**	**2.93e-07**	**4.85e-02**	**7.61e-02**
Planning for Transportation/commercial/others	7.80e-05	1.07	3.86	1.19e-03	1.25	3.89	7.34e-03	2.58	4.91
No Planning	2.34e-04	1.10	4.04	1.28e-03	1.28	3.66	1.28e-05	1.39	5.63
α_{s}^{j}	1.01e-03	1.48	4.60	1.92e-04	1.71	5.34	7.05e-05	8.08e-01	4.56
α_{t}^{j}	4.10e-04	2.39	5.54	2.07e-03	1.40	4.56	3.50e-02	7.27	9.85
δ^{j}	1.26e-04	8.23e-01	3.33	3.20e-04	1.37	4.13	4.47e-04	1.04	6.40

Additionally, a *t*-statistic is calculated to indicate the significance of other parameters. As is evident from table 6.3, planning for industrial is not significant on every observation (Highlighted).

With regards to the second question mentioned earlier, McNamara's test is employed to test the significant difference between the MNLM, GTWLM, ST-PLM and GSTLM, and the results are provided in Table 6.4

Table 6.4: Significance comparison for different models

Models Comparison	Z values GTWLM	ST-PLM	GSTLM
MNLM	-14.16	-8.36	-17.30
GTWLM	-	4.31	-6.88
ST-PLM	-	-	-13.84

The negative values clearly indicate that the GSTLM model outperforms the MNLM, GTWLM, and ST-PLM. The Z values between MNLM, GTWLM, ST-PLM and GSTLM are -17.30, -6.88 and -13.84 respectively, indicating that GSTLM substantially outperforms the other models. Also, the Z values between GTWLM and ST-PLM is 4.31, which is more than 1.96. This result demonstrates that a significant difference exists between the GTWLM and ST-PLM at the 95% confidence level. This comparison further demonstrates that the GTWLM outperforms the ST-PLM in terms of model accuracy.

Table 6.5 provides the kappa coefficients and their comparison with the other models. The Kno shows that the GSTLM achieves a better result than all the other models. Specifically, the GSTLM is the most optimal model for specifying the location and the GTWLM is the best model to specify quantity.

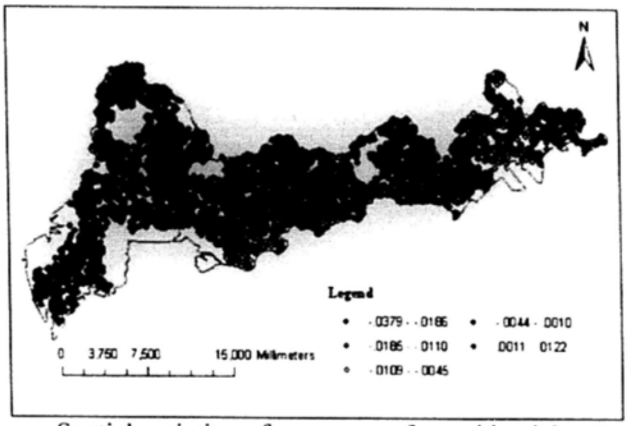
Table 6.5: Comparison of MNLM, GTWLM, ST-PLM and GSTLM with coefficients

	MNLM	GTWLM	ST-PLM	GSTLM
Kno	0.65	0.76	0.73	0.81
Klocation	0.63	0.73	0.71	0.81
Kquantity	0.90	0.98	0.90	0.93

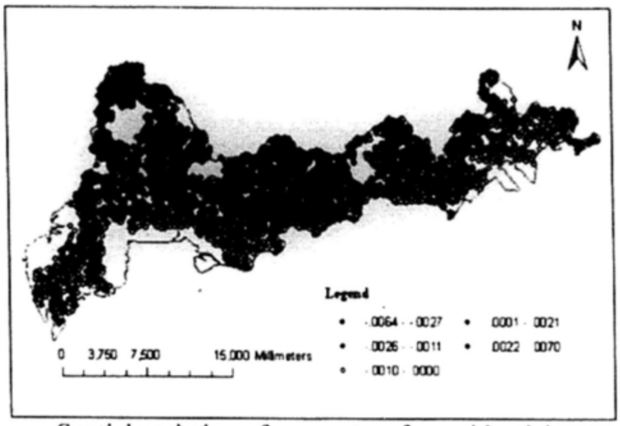
6.4 Analysis of spatio-temporal land use distribution pattern in SEZ, Shenzhen

Unlike the global models (MNLM, ST-PLM), which have unified parameters across space and time, the GSTLM generates a set of spatio-temporal parameter estimates on each land use sample observation. These can be used to analyze spatial and temporal variations of the effects of land use pattern determinants. Based on the sample points with parameter estimates, a set of pictures (Figures 6.2-6.4) are generated to reveal the spatial variations of explanatory factors for each land use type with generally regular spatial patterns. Then, several parameter estimates, which have remarkable temporal variation, are selected and spatio-temporal analysis is performed. The spatial and temporal variations of those factors are shown in figure 6.5.

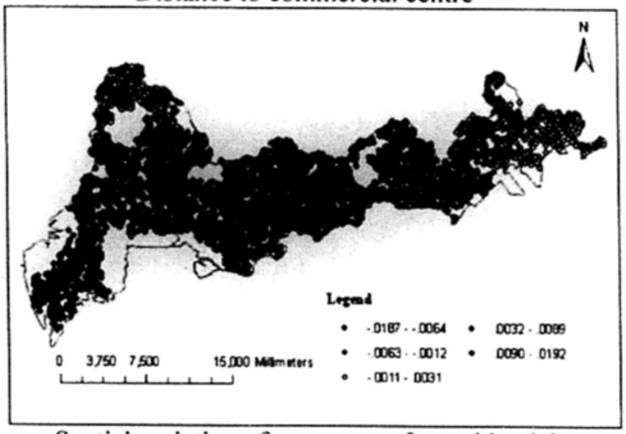
Fig. 6.2 presents the spatial variation of parameter estimates for residential land use. The distance to commercial centre has a greater negative effect on residential land use in the east of Luohu district and west of Yantian district. Almost all the positive effects from distance to financial centre occur in the Luohu district. Distance to the educational facilities has more negative effect in the east of SEZ. This is quite logical as the educational facilities are scarce resources in the east and hence have greater effect. Both the influence of distance to railway infrastructure and road are higher negative values. This signifies that the residential land use is distributed in the areas that are relatively closer to the transportation networks. Especially, the effect of proximity to transportation networks in Futian district is less or positive. Population has positive effect, whereas DEM and slope present a negative influence on the residential land use. However, negative influence of population can be found in some regions in the south of Nanshan district. The planning for residential has a greater positive influence on the residential land use, which implies that the planning for residential land use plays a very crucial role in the study area.



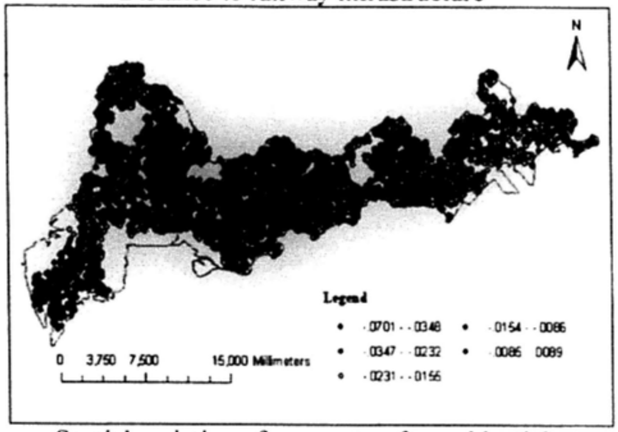
Spatial variation of parameters for residential:
Distance to commercial centre



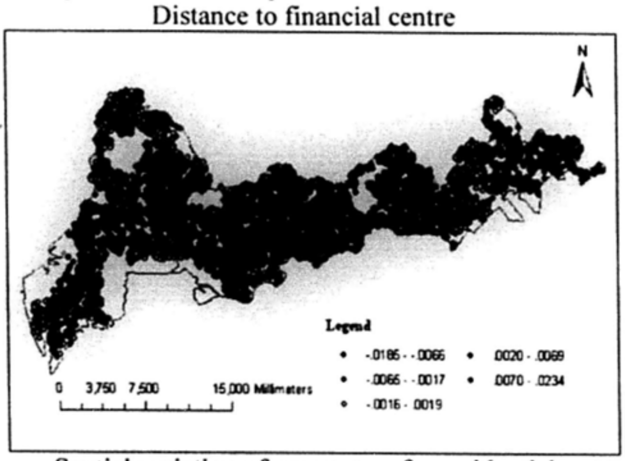
Spatial variation of parameters for residential:
Distance to railway infrastructure



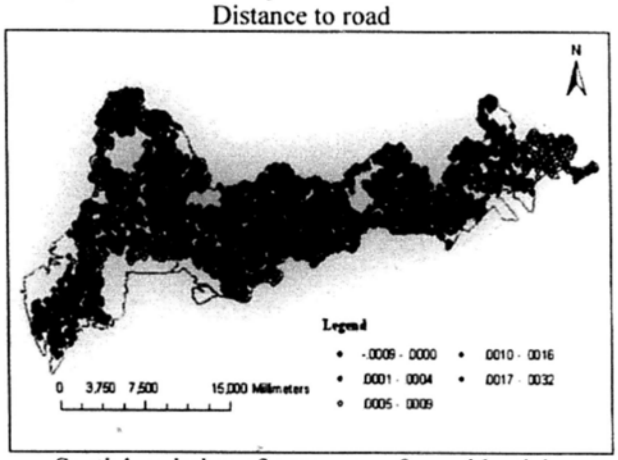
Spatial variation of parameters for residential:
Distance to financial centre



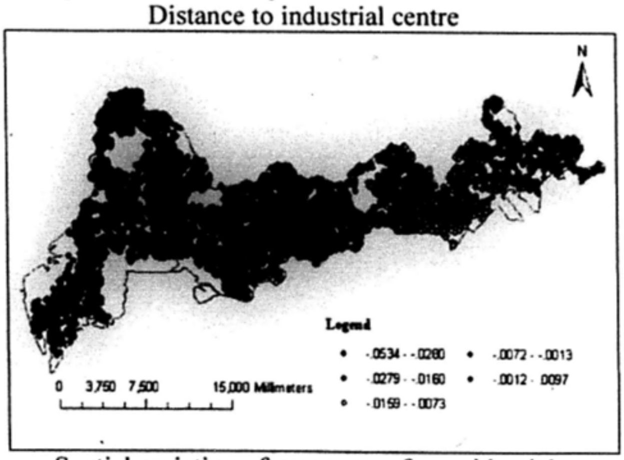
Spatial variation of parameters for residential:
Distance to road



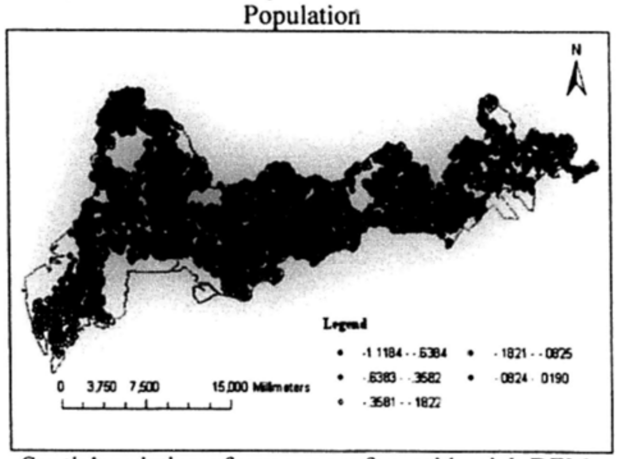
Spatial variation of parameters for residential:
Distance to industrial centre



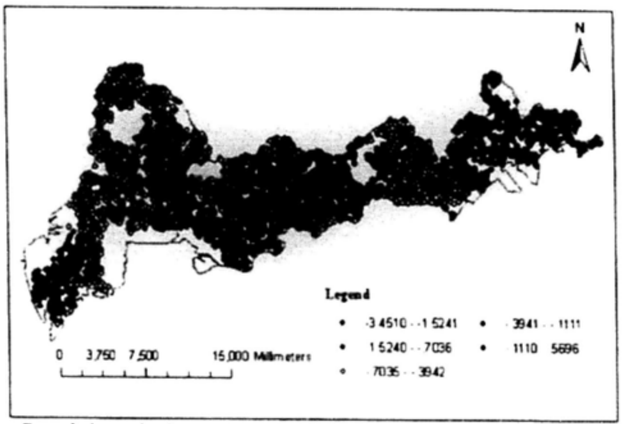
Spatial variation of parameters for residential:
Population



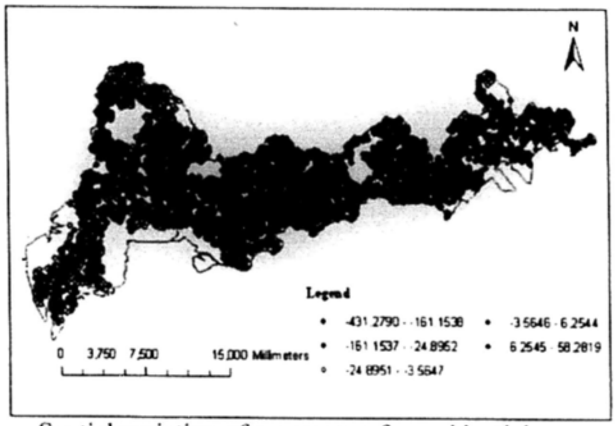
Spatial variation of parameters for residential:
Distance to educational facilities



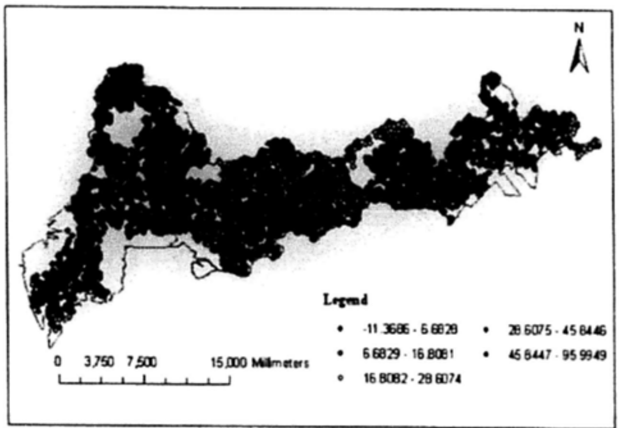
Spatial variation of parameters for residential: DEM



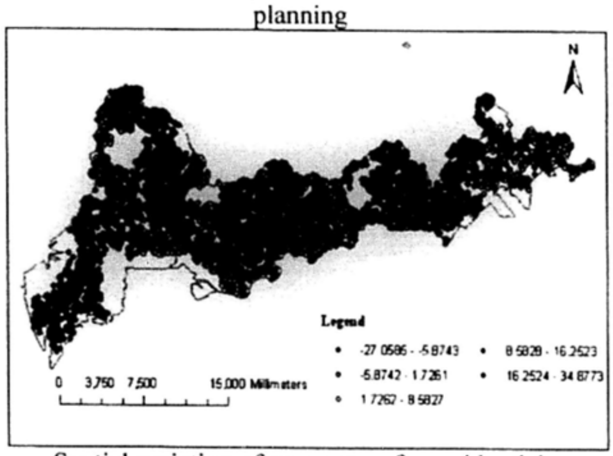
Spatial variation of parameters for residential: Slope



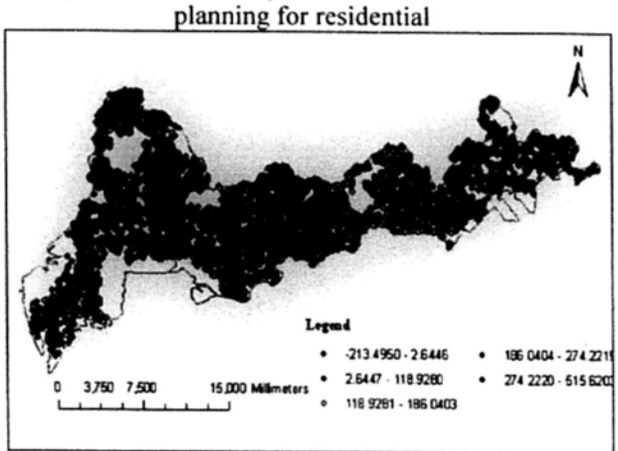
Spatial variation of parameters for residential: no planning



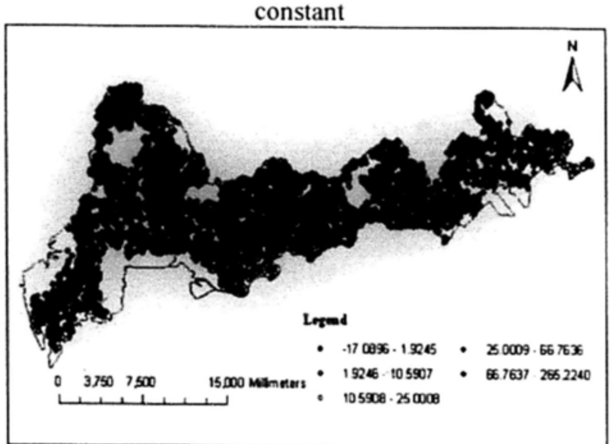
Spatial variation of parameters for residential: planning for residential



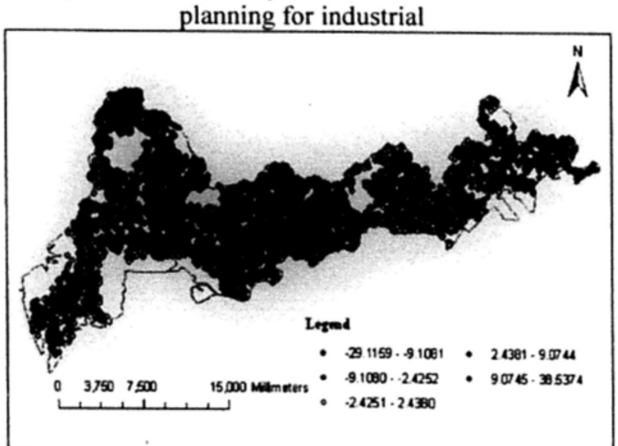
Spatial variation of parameters for residential: constant



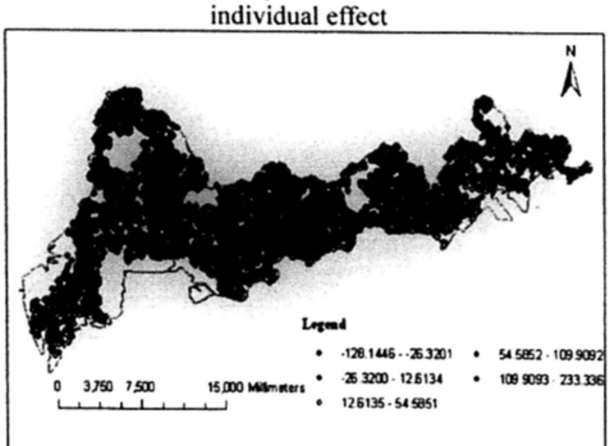
Spatial variation of parameters for residential: planning for industrial



Spatial variation of parameters for residential: individual effect



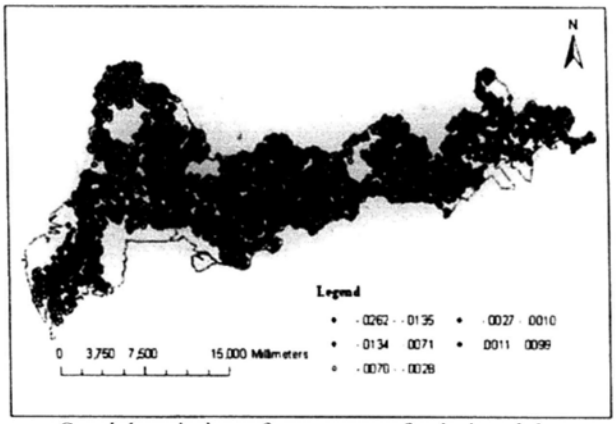
Spatial variation of parameters for residential: planning for transportation/commercial/others



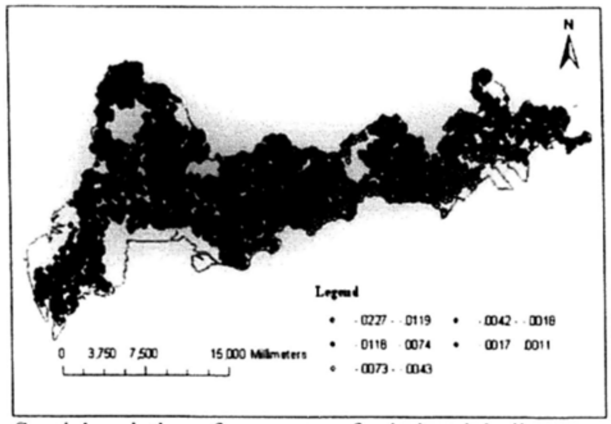
Spatial variation of parameters for residential: spatio-temporal autocorrelation

Figure 6.2: Spatial variation of parameters for residential

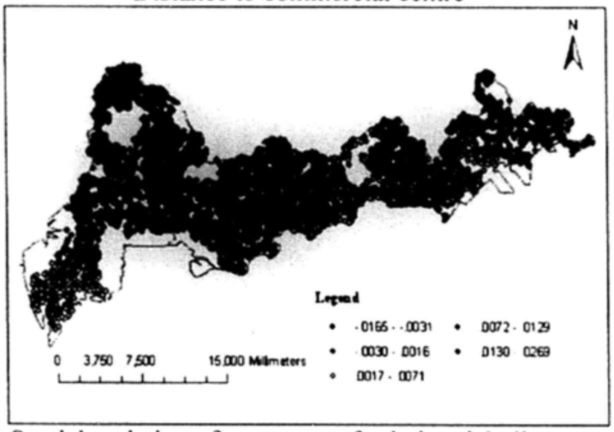
Figure 6.3 provides the spatial variations of parameter estimates for industrial land use. It can be seen that the distance to the commercial/financial centre has an additional negative effect on the industrial land use in the western part of the study area than the eastern part. The reason for the positive effects on the industrial land use is that most of the commercial and financial centres are located in the middle of the study area (Luohu and Futian district). As expected, the distance to the industrial centre has a negative influence throughout the study area, but more so in the north of Nanshan district. This is where the biggest industrial park, Nanshan high technology park is located. With the influence on residential land use, the distance to educational facilities has positive effect in the east and negative effect in the west. The distance to the railway infrastructure and distance to road chiefly affect the industrial land use negatively. Especially, a greater negative effect is found in the east. The influence from population has almost the same pattern as it in the residential land use. DEM and slope still have negative influence for industrial land use. The planning for industrial land use has a stronger influence with large parameter estimates. The picture also reveals that most of the industrial areas are built in accordance with the planning except in the middle east of Yantian district.



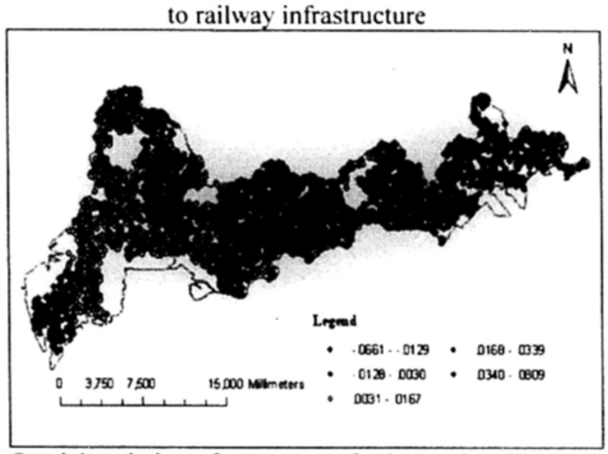
Spatial variation of parameters for industrial: Distance to commercial centre



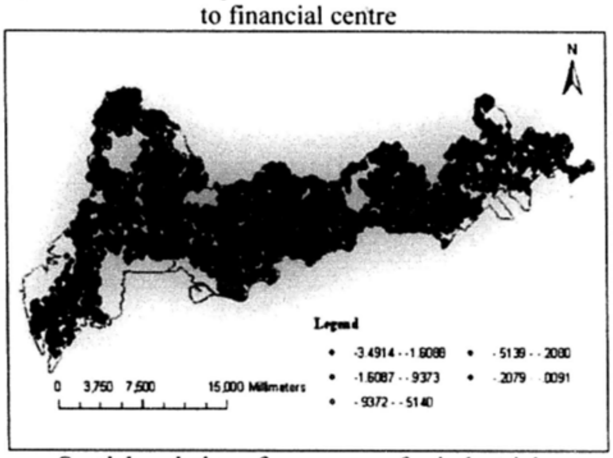
Spatial variation of parameters for industrial: distance to railway infrastructure



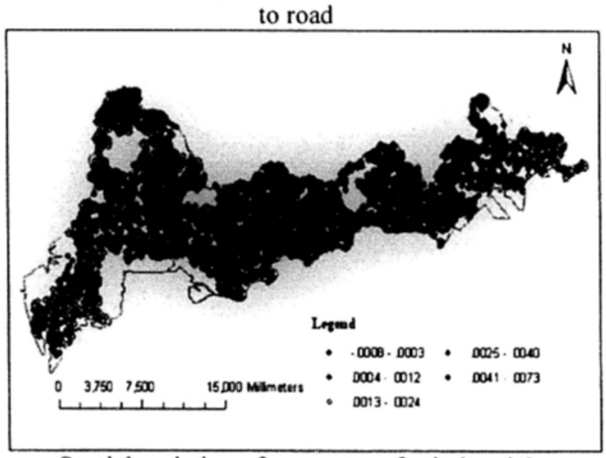
Spatial variation of parameters for industrial: distance to financial centre



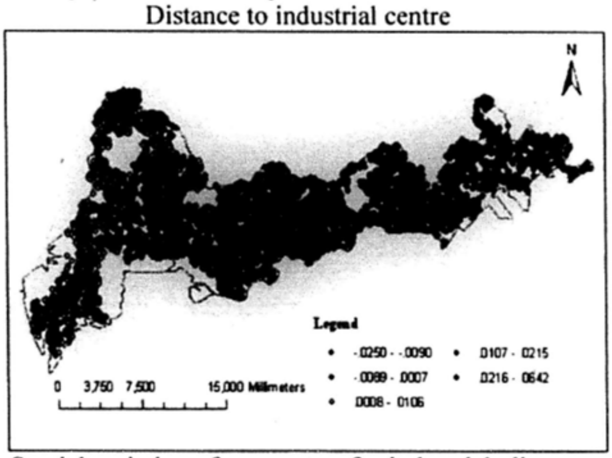
Spatial variation of parameters for industrial: distance to road



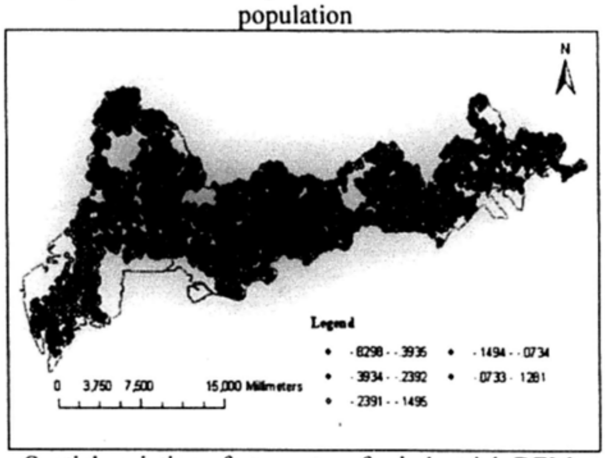
Spatial variation of parameters for industrial: Distance to industrial centre



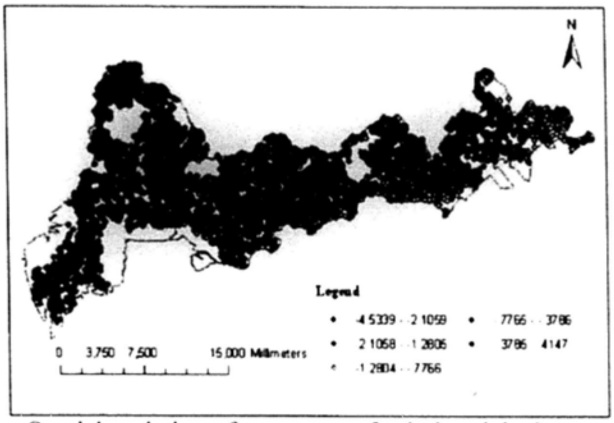
Spatial variation of parameters for industrial: population



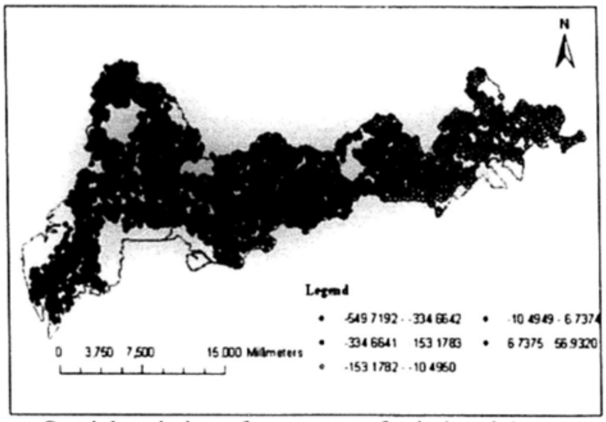
Spatial variation of parameters for industrial: distance to educational facilities



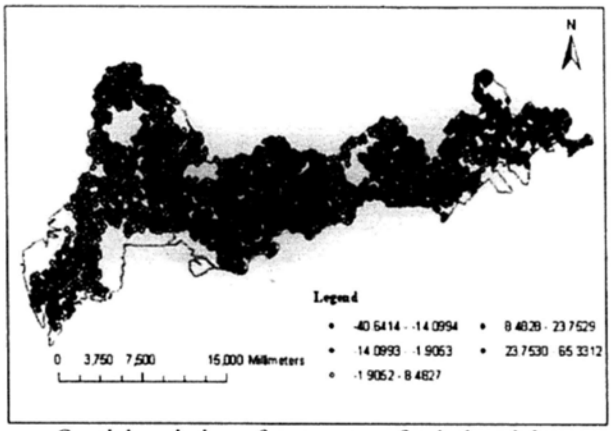
Spatial variation of parameters for industrial: DEM



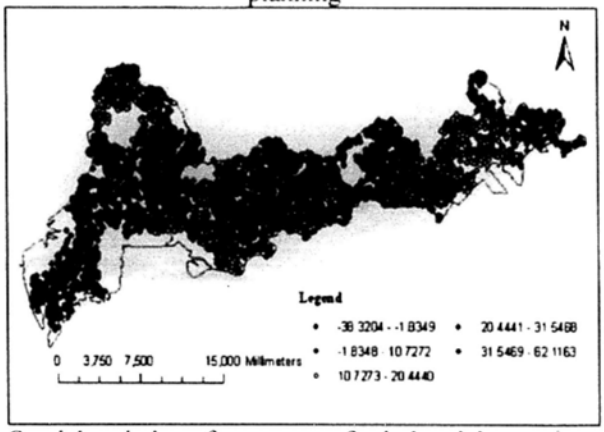
Spatial variation of parameters for industrial: slope



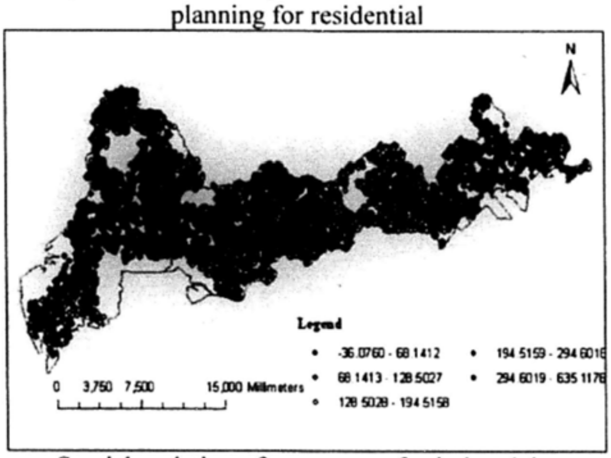
Spatial variation of parameters for industrial: no planning



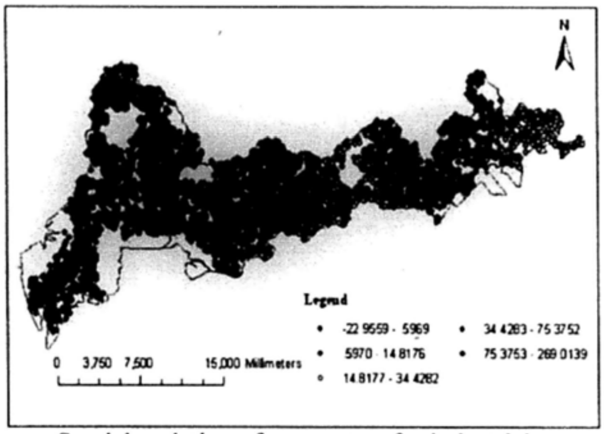
Spatial variation of parameters for industrial: planning for residential



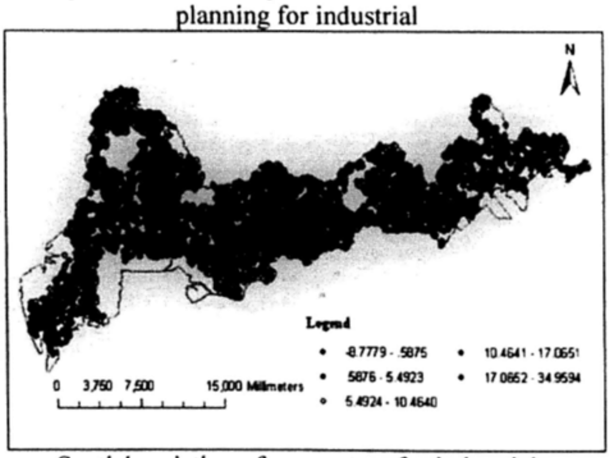
Spatial variation of parameters for industrial: constant



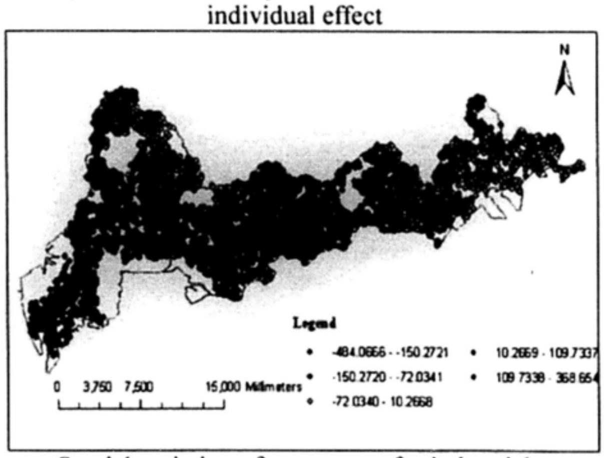
Spatial variation of parameters for industrial: planning for industrial



Spatial variation of parameters for industrial: individual effect



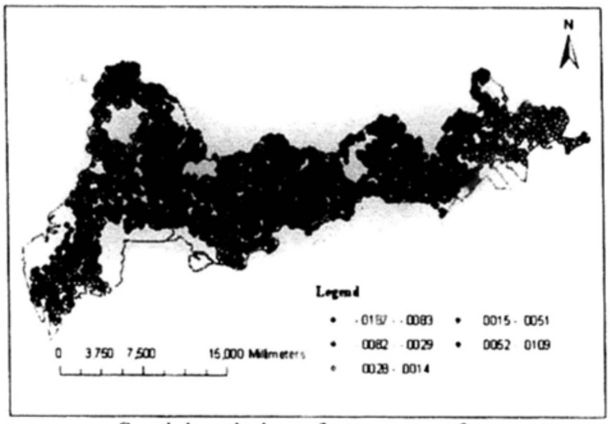
Spatial variation of parameters for industrial: planning for transportation/commercial/others



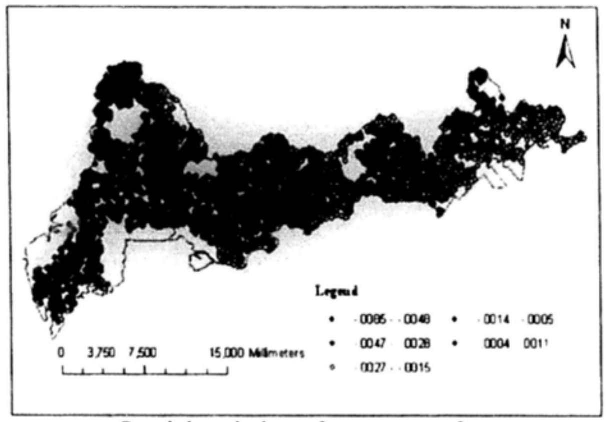
Spatial variation of parameters for industrial: spatio-temproal autocorrelation

Figure 6.3: Spatial variation of parameters for industrial

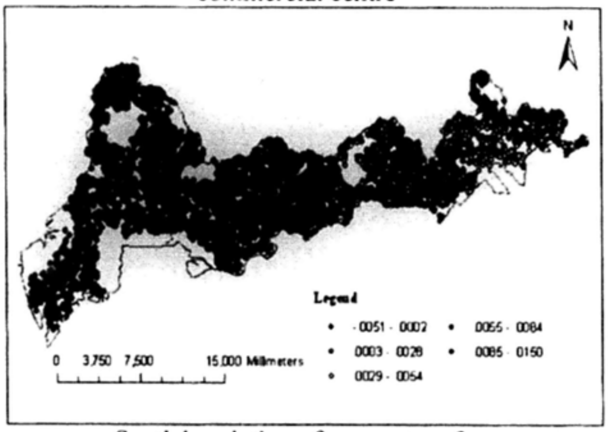
Figure 6.4 illustrates that the distance to commercial has an added negative influence for Transportation/commercial/others land use type. Also, the distance to the financial centre and educational facilities have more positive effect at the junction of Luohu district and Futian district. Distance to the railway infrastructure and road have negative influence. In particular, it can be seen that the positive influence of the distance to transportation network is mainly concentrated in the west of the study area. Distance to the industrial centre has a negative effect in the study area. However, positive influence of the distance to the industrial centre can still be found in some parts. Similar to residential and industrial land use, DEM and slope have negative influence for Transportation/commercial/others. For some part, population has a negative effect on Transportation/commercial/others in the Study area. Planning for transportation/commercial/others performs well except in the old district, Luohu.



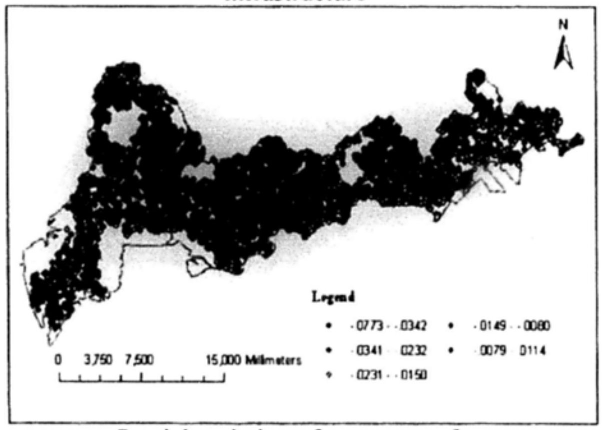
Spatial variation of parameters for transportation/commercial/others: Distance to commercial centre



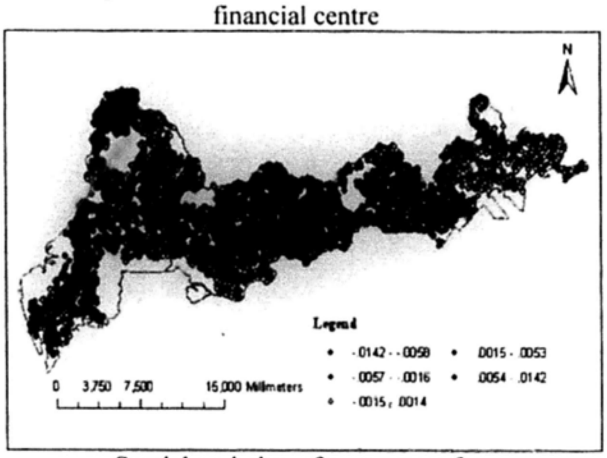
Spatial variation of parameters for transportation/commercial/others: distance to railway infrastructure



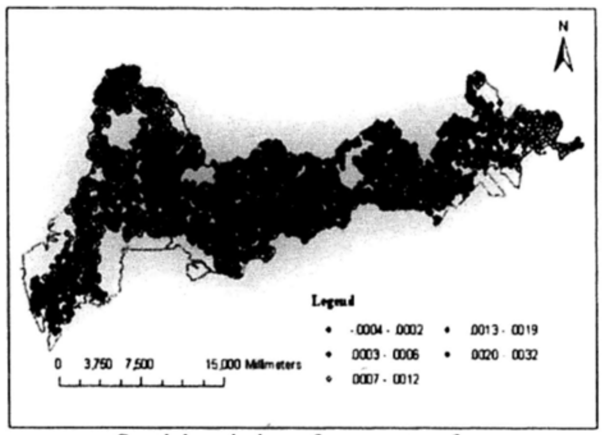
Spatial variation of parameters for transportation/commercial/others: distance to financial centre



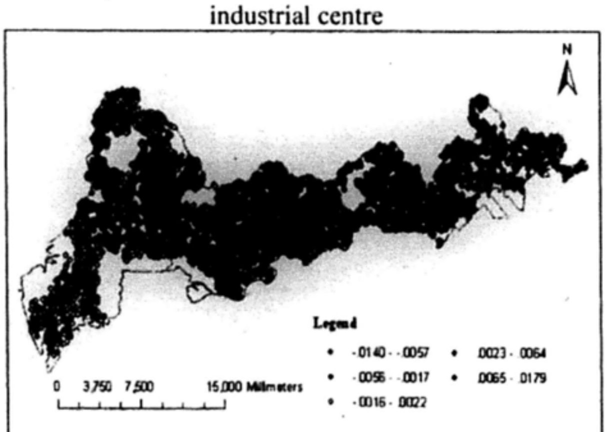
Spatial variation of parameters for transportation/commercial/others: distance to road



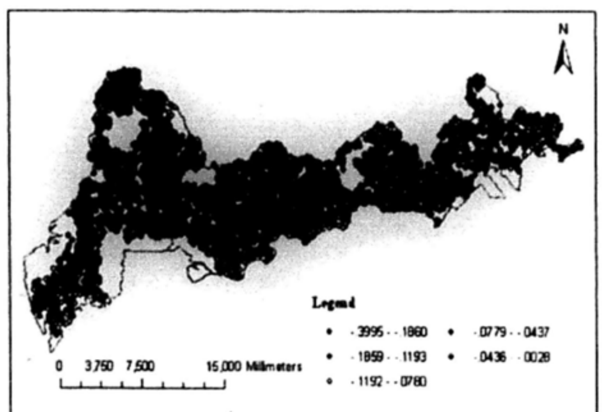
Spatial variation of parameters for transportation/commercial/others: distance to industrial centre



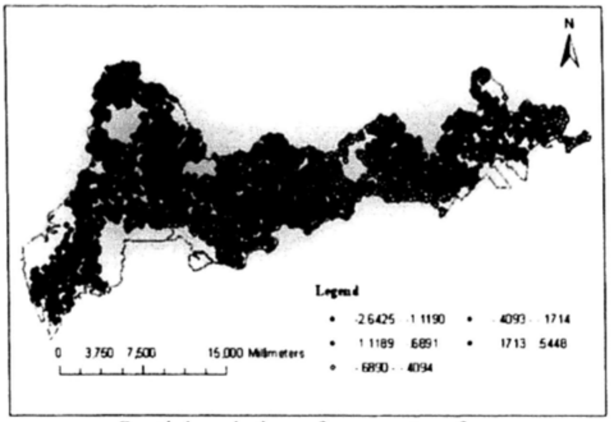
Spatial variation of parameters for transportation/commercial/others: population



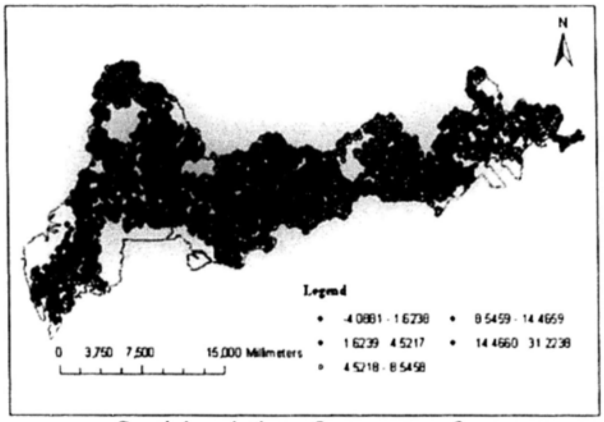
Spatial variation of parameters for transportation/commercial/others: distance to educational facilities



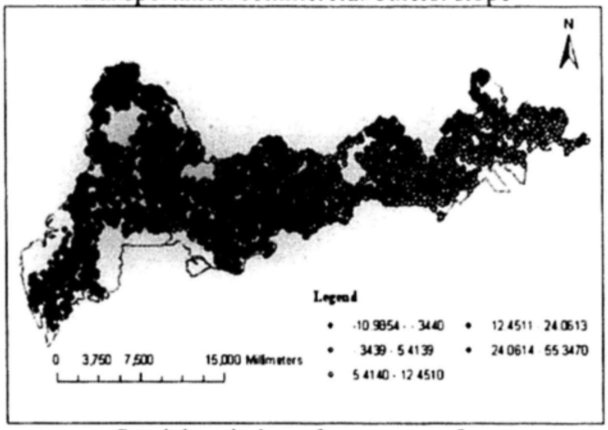
Spatial variation of parameters for transportation/commercial/others: DEM



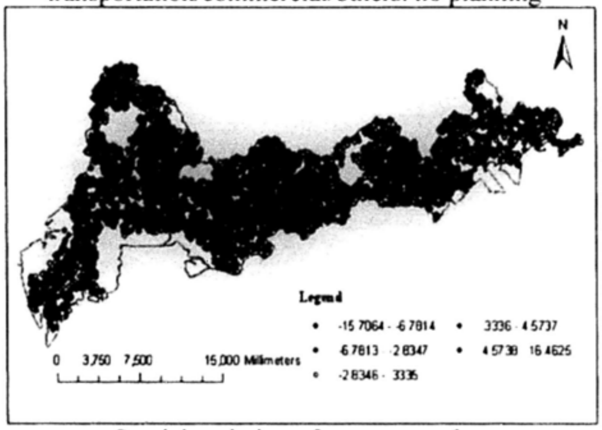
Spatial variation of parameters for transportation/commercial/others: slope



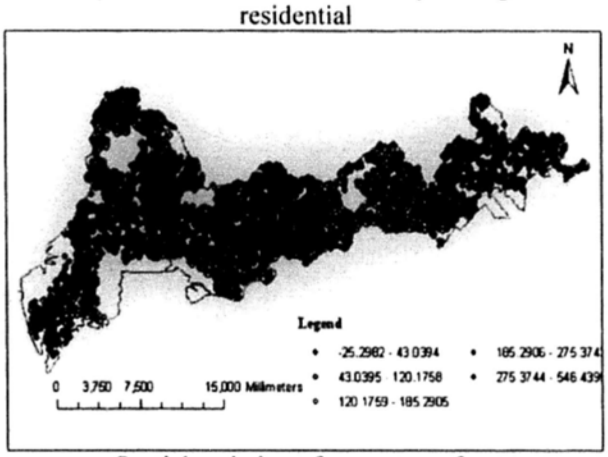
Spatial variation of parameters for transportation/commercial/others: no planning



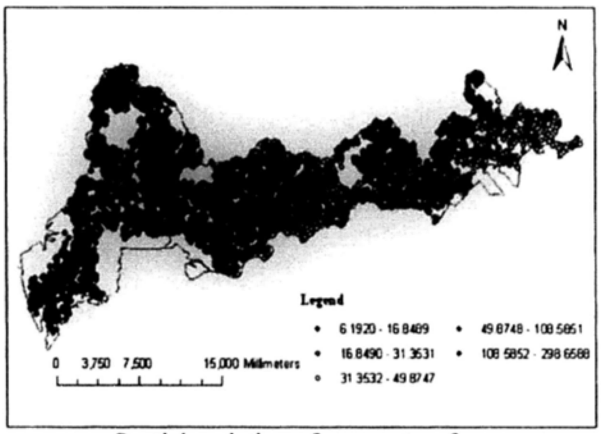
Spatial variation of parameters for transportation/commercial/others: planning for residential



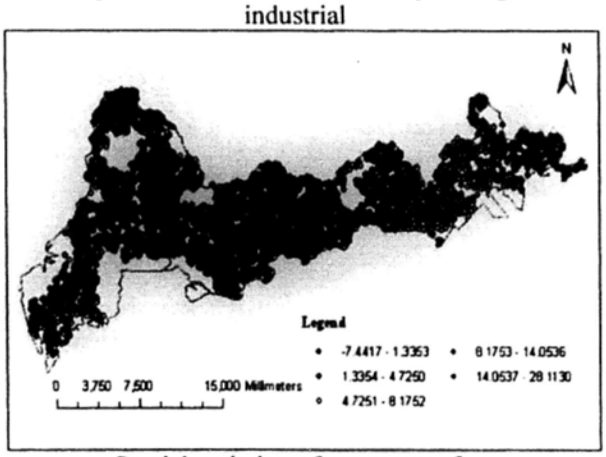
Spatial variation of parameters for transportation/commercial/others: constant



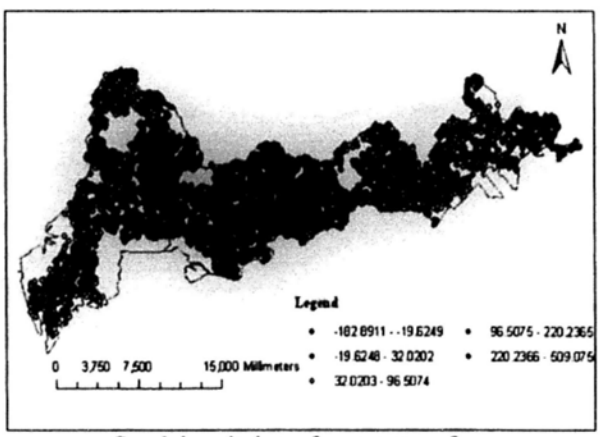
Spatial variation of parameters for transportation/commercial/others: planning for industrial



Spatial variation of parameters for transportation/commercial/others: individual effect



Spatial variation of parameters for transportation/commercial/others: planning for transportation/commercial/others



Spatial variation of parameters for transportation/commercial/others: spatio-temporal autocorrelation

Figure 6.4: Spatial variation of parameters or transportation/commercial/others

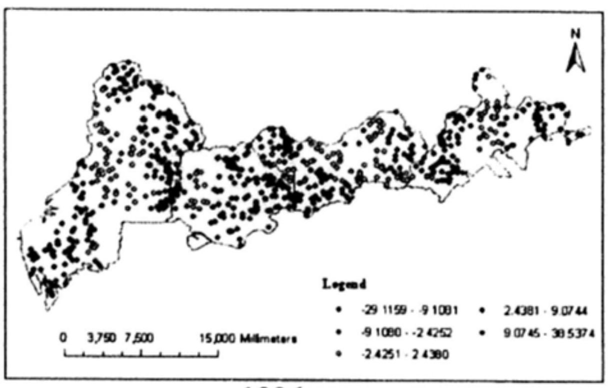
Table 6.7: GSTLM median estimate in different years

Residential						
Time	**1996**	**2000**	**2002**	**2004**	**2006**	**2008**
Distance to Commercial Centre	-1.65e-03	-3.70e-04	-4.09e-04	-1.51e-03	-1.34e-03	-1.13e-03
Distance to Financial Centre	-1.31e-03	-1.94e-03	-1.59e-03	-1.66e-03	-1.62e-03	-1.44e-03
Distance to Industrial Centre	-1.85e-05	-5.03e-04	-1.04e-03	-1.45e-04	-9.71e-05	-1.38e-03
Distance to Educational Facilities	-1.81e-03	-2.33e-03	-3.38e-03	-2.72e-03	-2.40e-03	-1.60e-03
Distance to Railway Infrastructure	-4.67e-04	-4.94e-04	-6.65e-04	-7.11e-04	-6.20e-04	-6.24e-04
Distance to Road	-1.09e-02	-1.37e-02	-1.43e-02	-1.56e-02	-1.59e-02	-1.69e-02
Population	3.31e-04	2.10e-04	1.08e-04	1.53e-04	1.76e-04	2.16e-04
DEM	-1.54e-01	-1.30e-01	-1.21e-01	-1.12e-01	-9.71e-02	-9.78e-02
Slope	-4.34e-01	-4.52e-01	-4.85e-01	-4.47e-01	-3.93e-01	-4.23e-01
Planning for Residential	1.17e+01	1.12e+01	1.02e+01	1.05e+01	1.06e+01	1.14e+01
Planning for Industrial	1.59e+02	1.19e+02	1.18e+02	1.08e+02	1.27e+02	1.36e+02
Planning for Transportation/commercial/others	-3.89e-01	-8.33e-01	-5.33e-01	3.35e-01	8.97e-01	1.07
No Planning	8.53e-01	-2.57	-1.30	-2.64e-01	-6.30e-01	2.59e-01
α_0^j	4.52	5.19	6.45	6.42	5.63	5.02
α_1^j	4.37	4.22	4.98	8.78	1.19e+01	1.32e+01
δ^j	3.53e+01	2.19e+01	2.94	-8.85	-1.31e+01	-1.16e+01
Industrial						
Time	**1996**	**2000**	**2002**	**2004**	**2006**	**2008**
Distance to Commercial Centre	-1.97e-03	-1.24e-03	-2.45e-03	-2.71e-03	-2.16e-03	-2.39e-03
Distance to Financial Centre	-9.80e-05	-5.57e-04	-3.36e-04	2.04e-04	1.32e-03	1.76e-04
Distance to Industrial Centre	-8.84e-02	-7.39e-02	-1.01e-01	-9.78e-02	-7.93e-02	-8.24e-02
Distance to Educational Facilities	6.77e-03	7.19e-03	8.09e-03	7.69e-03	4.97e-03	5.67e-03
Distance to Railway Infrastructure	-2.39e-03	-2.47e-03	-1.80e-03	-1.59e-03	-2.05e-03	-1.64e-03
Distance to Road	7.81e-03	5.51e-03	7.05e-03	8.44e-03	6.86e-03	7.67e-03
Population	7.64e-04	4.25e-04	4.20e-04	2.40e-04	2.24e-04	2.57e-04
DEM	-1.48e-01	-1.39e-01	-1.45e-01	-1.36e-01	-1.00e-01	-1.04e-01

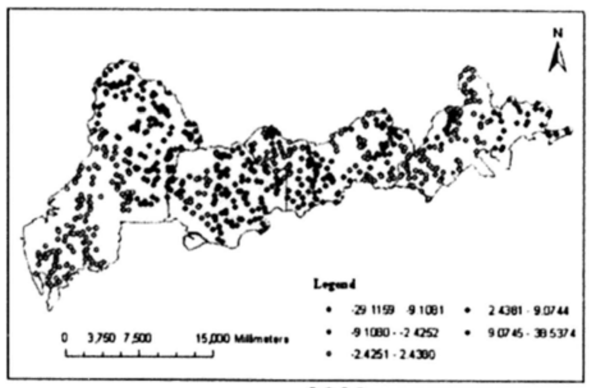
Slope	-7.72e-01	-5.55e-01	-4.79e-01	-5.28e-01	-5.40e-01	-5.56e-01
Planning for Residential	6.44	5.58	3.02	2.89	3.04	3.02
Planning for Industrial	1.46e+02	1.36e+02	1.30e+02	1.28e+02	1.23e+02	1.37e+02
Planning for Transportation/commercial/others	1.03e+01	7.91	4.87	3.64	3.94	5.16
No Planning	3.63	7.32e-01	1.03	6.72e-01	5.28e-01	1.74
α_0^j	1.21e+01	9.38	1.35e+01	1.56e+01	1.40e+01	1.35e+01
$\boldsymbol{\alpha_1^j}$	**-2.96**	**-7.25e-01**	**1.83**	**3.91**	**7.34**	**9.33**
δ^j	-8.45e+01	-6.5 8e01	-4.89e+01	-4.31e+01	-4.22e+01	-3.55e+01
Transportation/commercial/others						
Time	**1996**	**2000**	**2002**	**2004**	**2006**	**2008**
Distance to Commercial Centre	1.09e-03	1.75e-03	2.28e-03	2.08e-03	1.73e-03	1.60e-03
Distance to Financial Centre	**6.08e-04**	**-1.25e-05**	**1.01e-04**	**-2.15e-04**	**-7.67e-05**	**2.07e-04**
Distance to Industrial Centre	**1.01e-03**	**1.29e-03**	**7.70e-04**	**-2.24e-04**	**-4.77e-04**	**-1.26e-03**
Distance to Educational Facilities	-9.86e-04	-1.20e-03	-4.74e-04	-5.29e-04	-7.86e-04	-3.59e-04
Distance to Railway Infrastructure	-4.80e-04	-7.07e-04	-7.62e-04	-8.16e-04	-8.07e-04	-7.91e-04
Distance to Road	-6.44e-03	-7.05e-03	-8.22e-03	-1.24e-02	-1.61e-02	-1.76e-02
Population	6.27e-04	5.00e-04	4.37e-04	3.41e-04	2.94e-04	3.38e-04
DEM	-8.22e-02	-7.14e-02	-6.98e-02	-5.14e-02	-4.48e-02	-4.64e-02
Slope	-2.87e-01	-2.76e-01	-2.80e-01	-3.06e-01	-2.86e-01	-3.07e-01
Planning for Residential	5.32	7.23	7.51	8.53	7.51	7.12
Planning for Industrial	1.61e+02	1.39e+02	1.36e+02	1.22e+02	1.24e+02	1.29e+02
Planning for Transportation/commercial/others	4.76	5.37	4.87	5.09	4.02	3.50
No Planning	3.88	4.31	3.91	3.06	2.40	2.47
$\boldsymbol{\alpha_0^j}$	**-4.27**	**-3.85**	**-3.13**	**-7.77e-01**	**8.44e-01**	**9.84e-01**
α_1^j	2.37e+01	2.08e+01	1.92e+01	2.10e+01	2.35e+01	2.41e+01
δ^j	2.83e+01	4.18e+01	2.82e+01	1.40e+01	7.94e+00	1.18e+01

Table 6.7 shows that the median parameter estimates change over time. It can be found that some parameter estimates have a remarkable change in different years. In order to analyze the spatio-temporal relationship between the land use pattern and the explanatory factors, only those explanatory factors which show changes on sign are selected. Figure 6.5 illustrates the spatio-temporal variations of the selected parameter estimates over different years. The parameter, distance to the financial

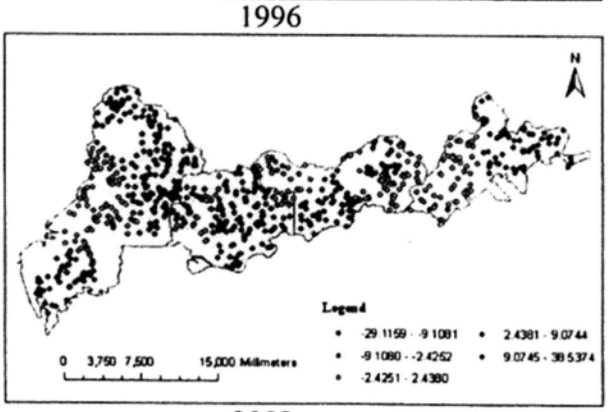
centre for industrial land use, is considered for example. In SEZ, high technology industry is the major component of industry. It is a kind of capital-intensive industry, and largely depends on financial market. Hence, the distribution of financial centre has a strong relationship with industry land use. At the early days of city development, most industry land use is built near the financial centre. For Shenzhen, almost all the financial organizations are located in Futian at the beginning. Thus, the parameter estimates in Futian are negative over the time. As time goes by, the land near the financial centre is more expensive. Meanwhile, with the development of finance, the financial organizations can cover the whole SEZ. The new industry land use less depends on the distance to the financial centre located in Futian. More and more industry land use can be built in the east of Luohu district and west of Yantian district, which are undeveloped. As seen from the pictures, the colors of the parameter estimates gradually change from yellow to red from 1996 to 2008 in the east of Luohu district and west of Yantian district. This implies that the distance between the financial centre and the industrial land use is farther in this part.



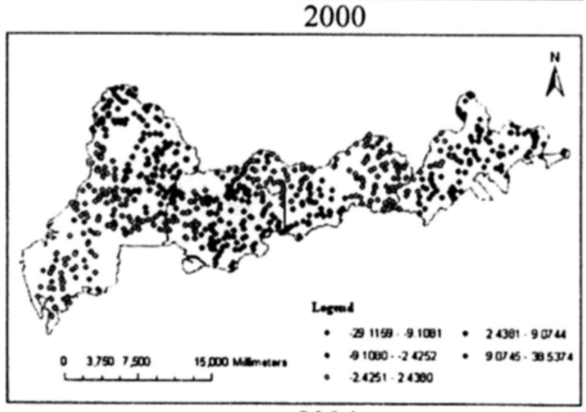
1996



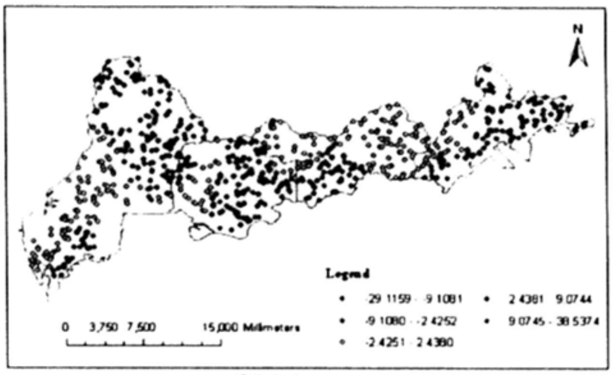
2000



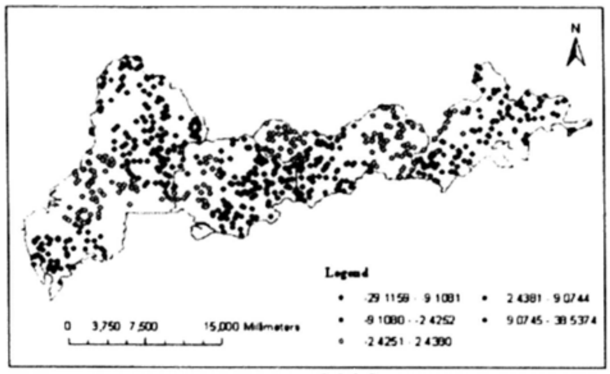
2002



2004

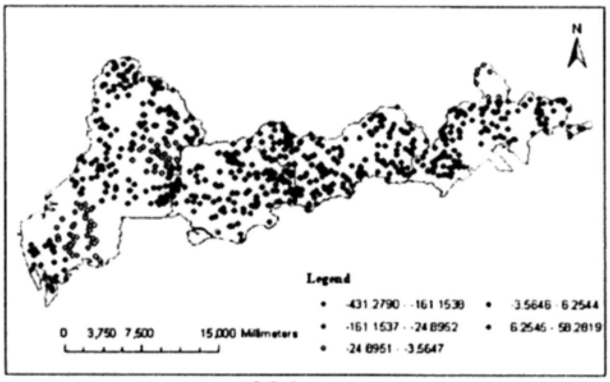


2006

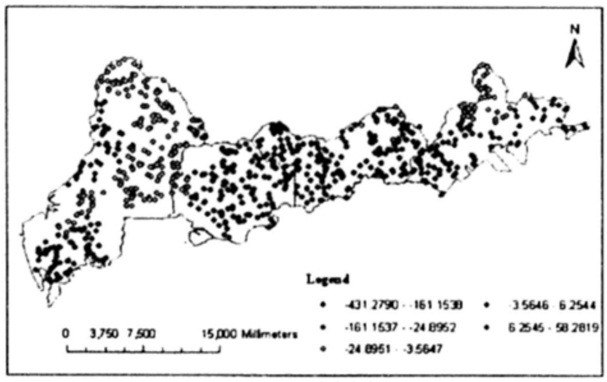


2008

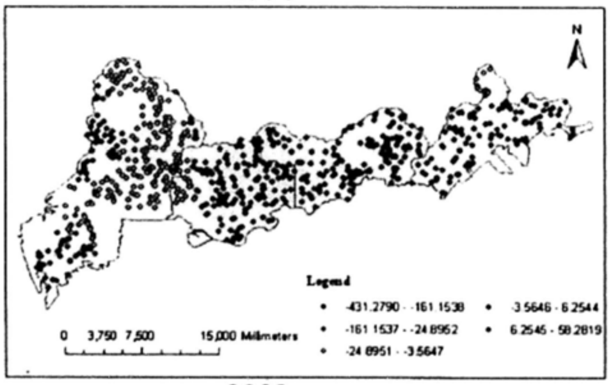
Spatio-temporal variation of parameters for residential: Planning for Transportation/commercial/others



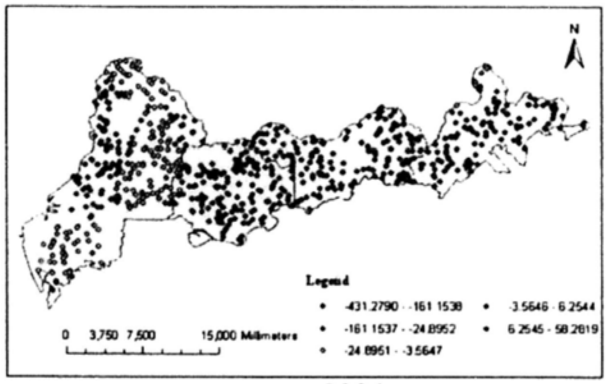
1996



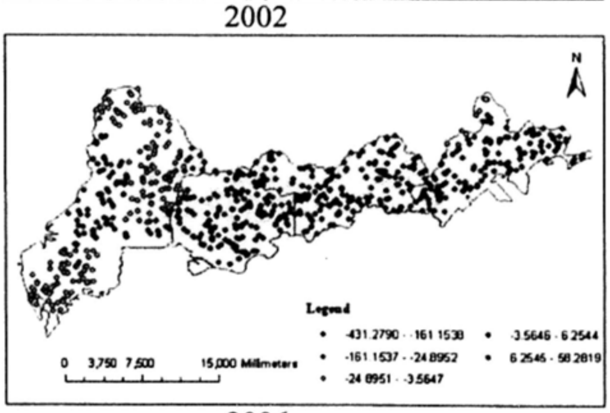
2000



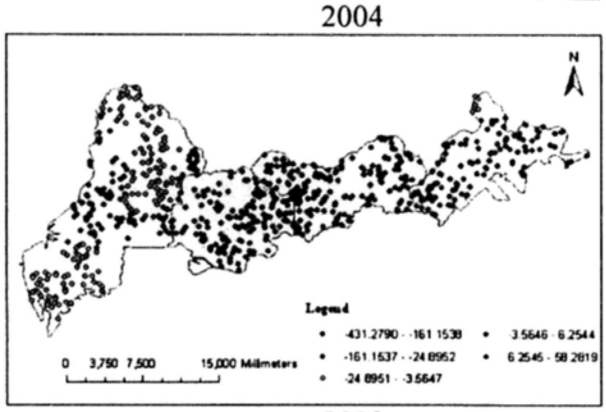
2002



2004

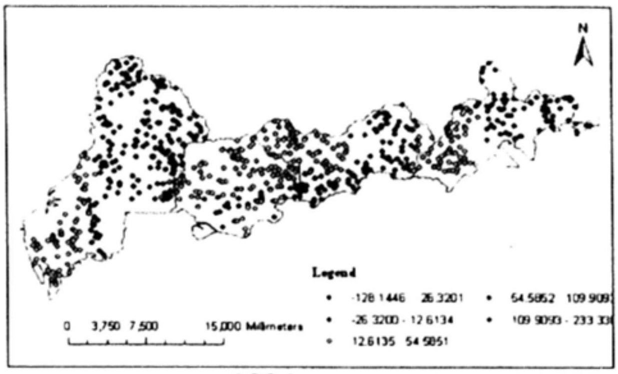


2006

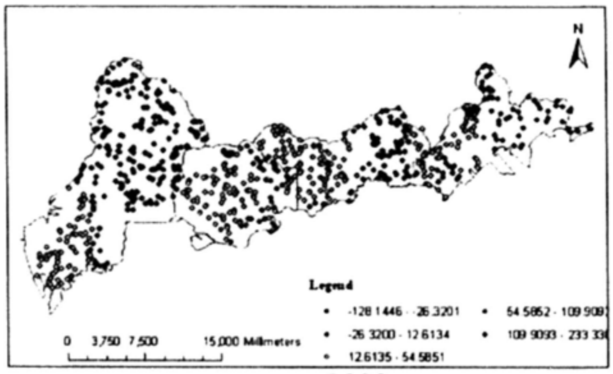


2008

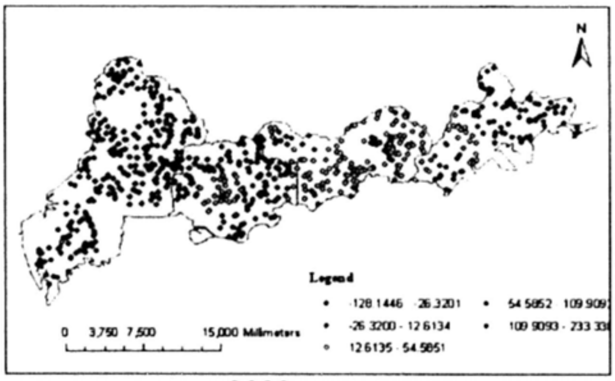
Spatio-temporal variation of parameters for residential: No Planning



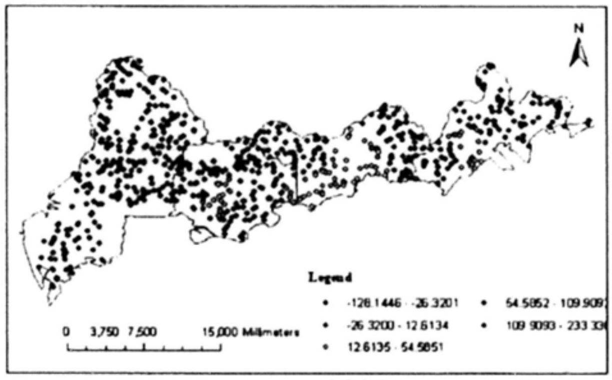
1996



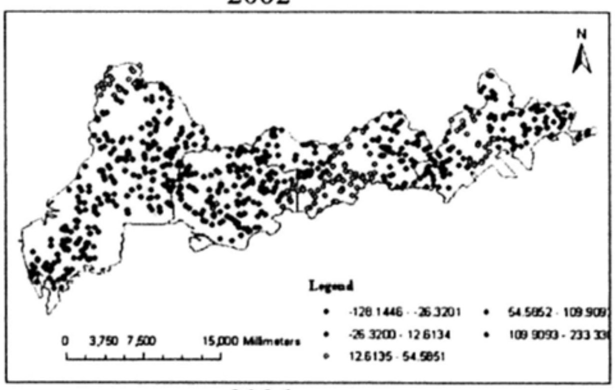
2000



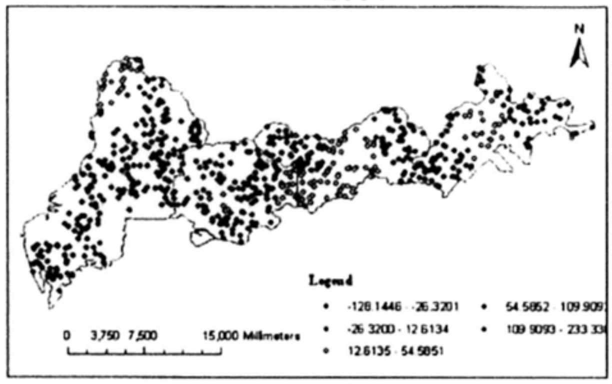
2002



2004

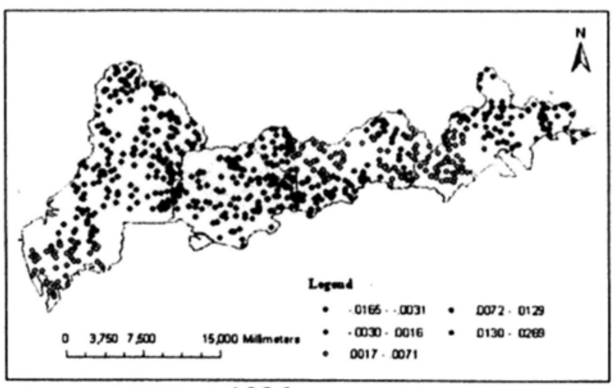


2006

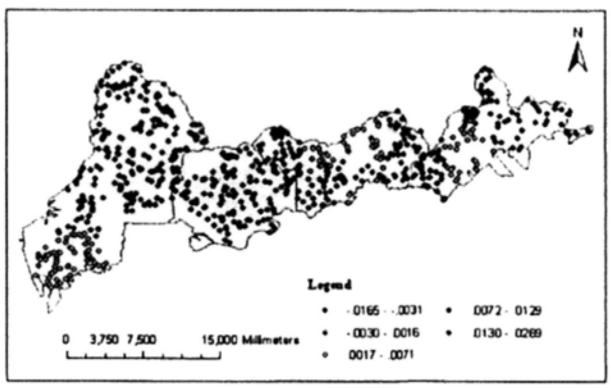


2008

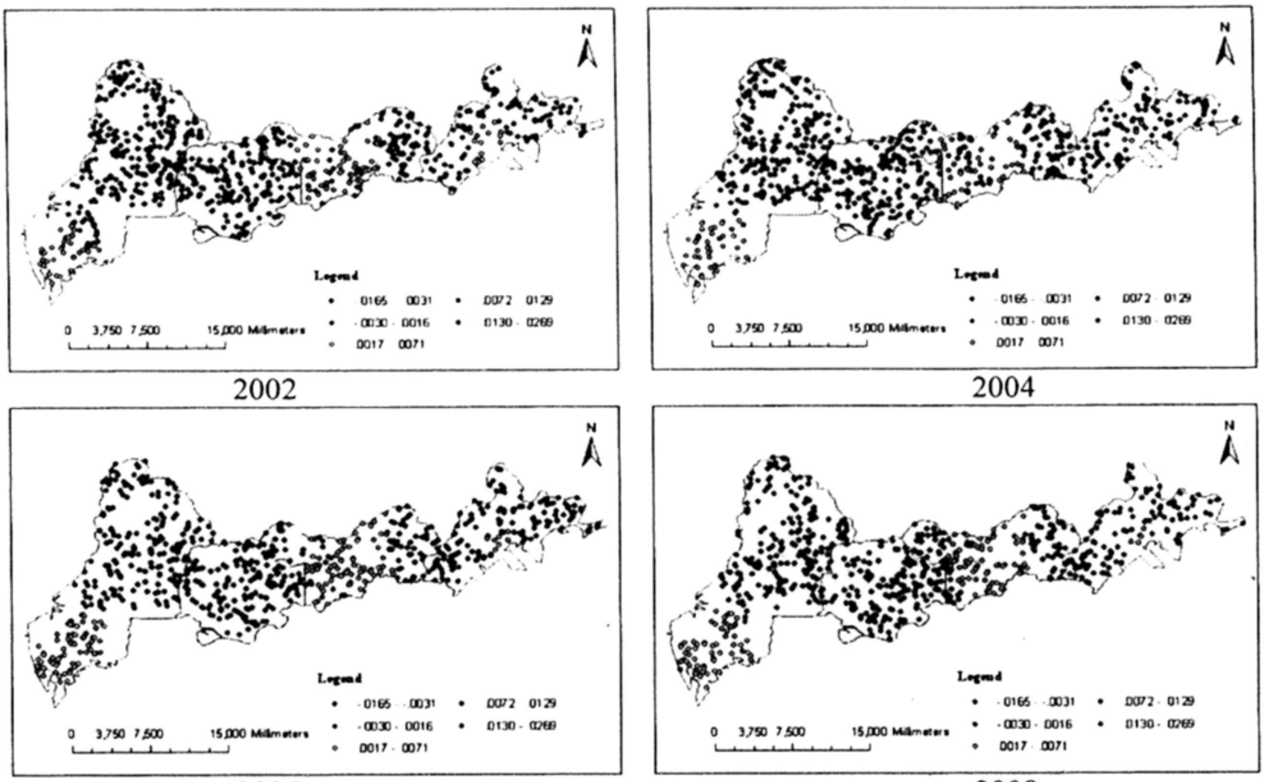
Spatio-temporal variation of parameters for residential: Spato-temproal autocorrelation



1996



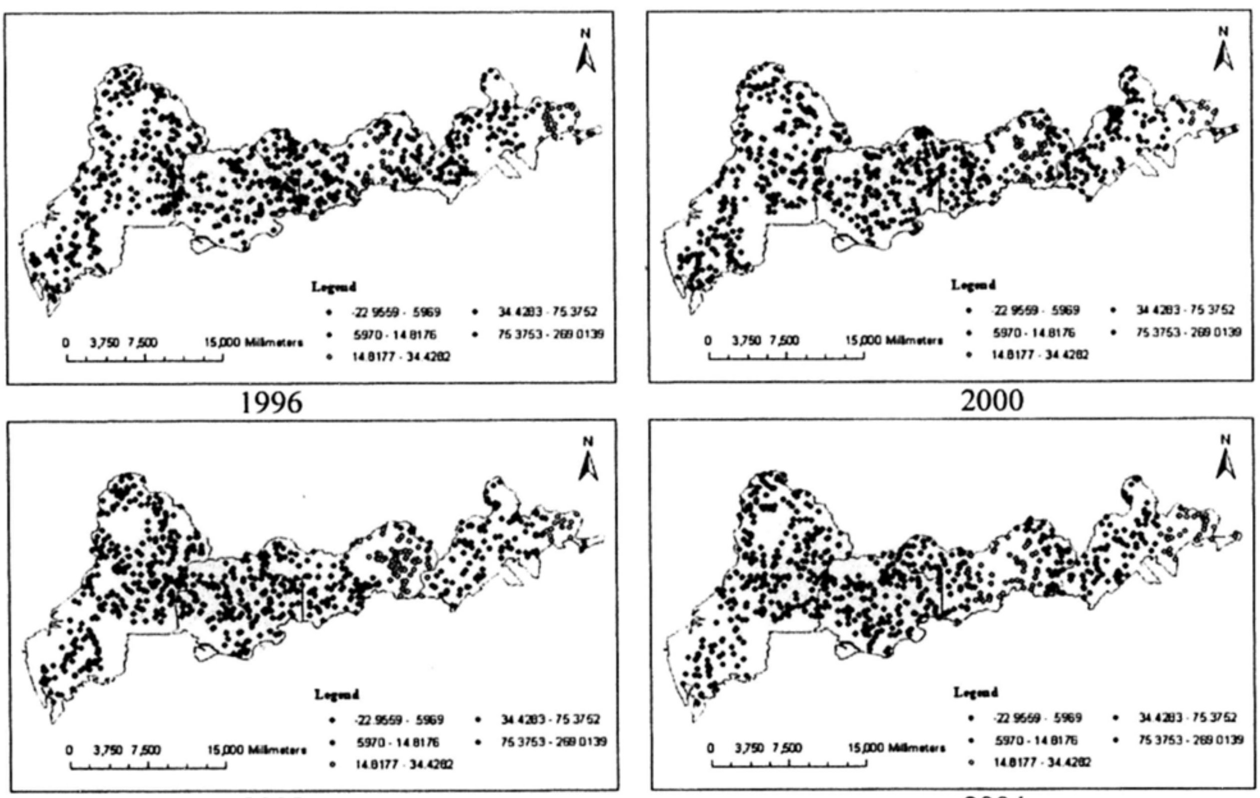
2000



2002 2004

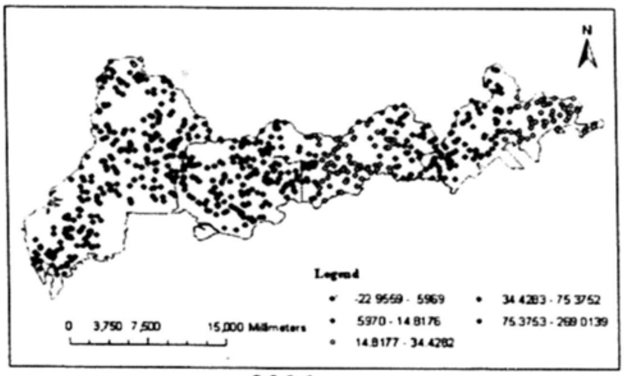
2006 2008

Spatio-temporal variation of parameters for industrial: Distance to Financial Centre

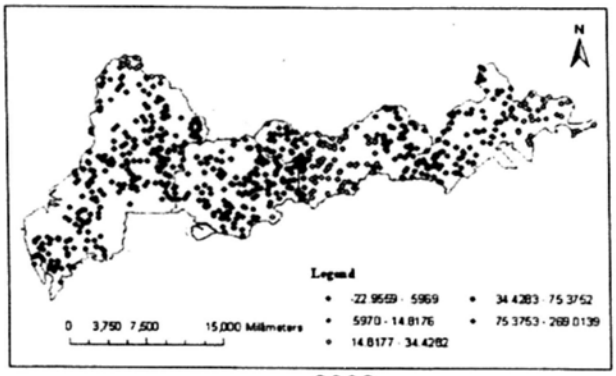


1996 2000

2002 2004

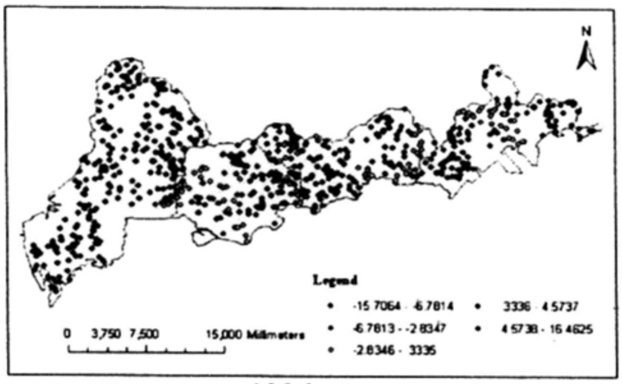


2006

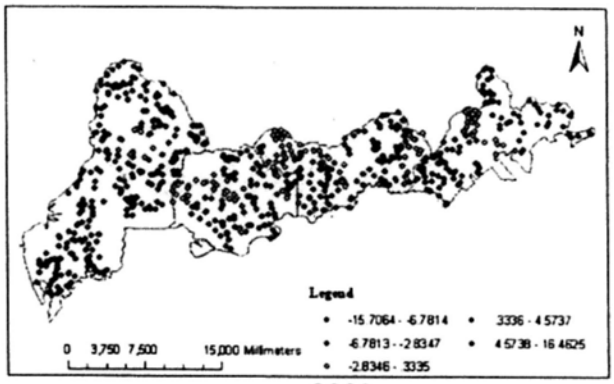


2008

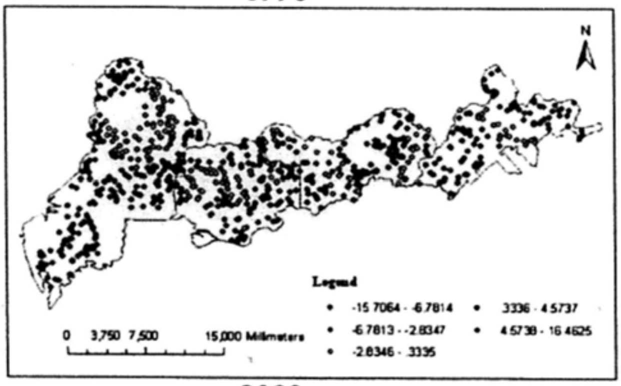
Spatio-temporal variation of parameters for industrial: individual effect



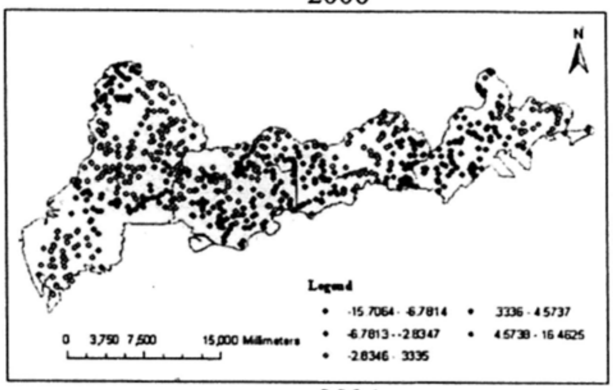
1996



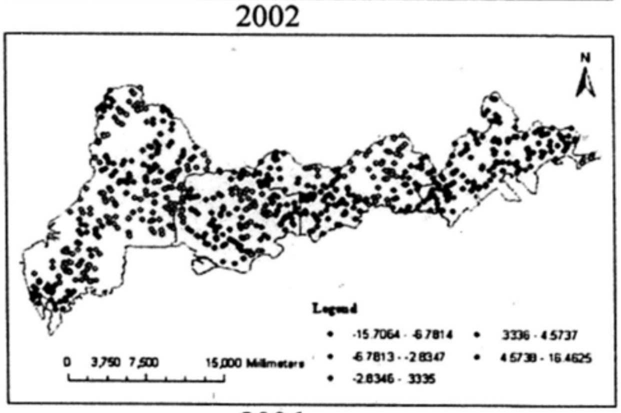
2000



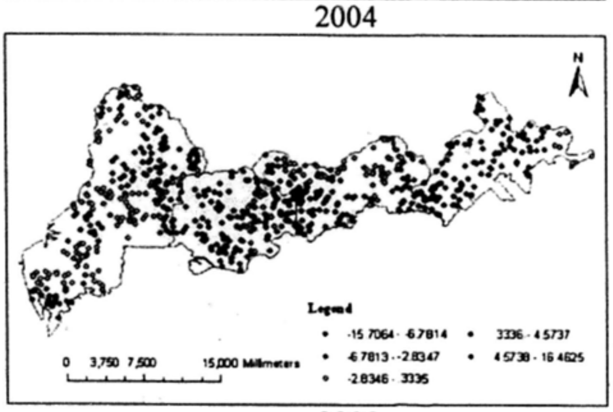
2002



2004

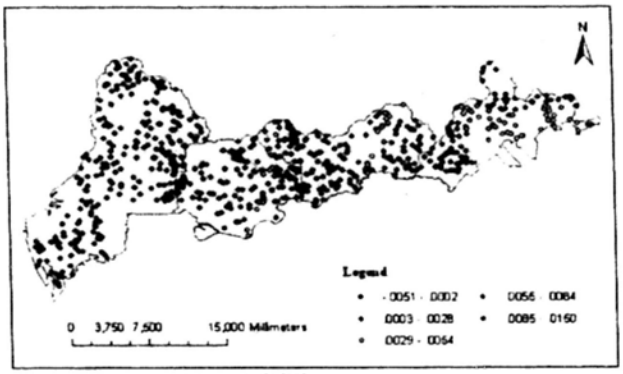


2006

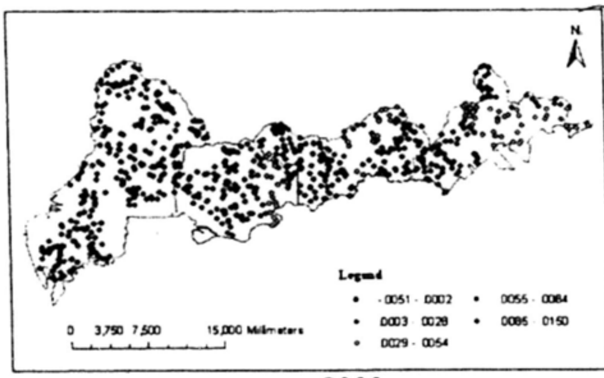


2008

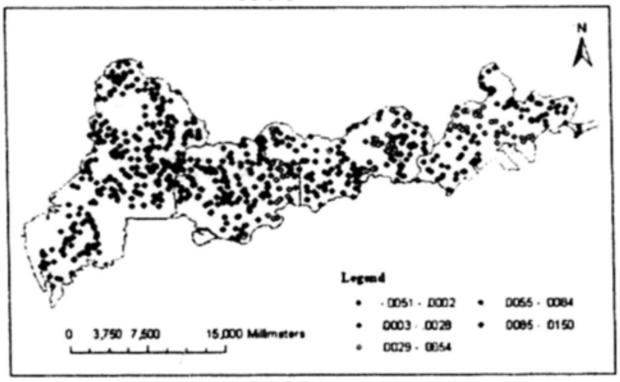
Spatio-temporal variation of parameters for transportation/commercial/others: Constant α_0^j



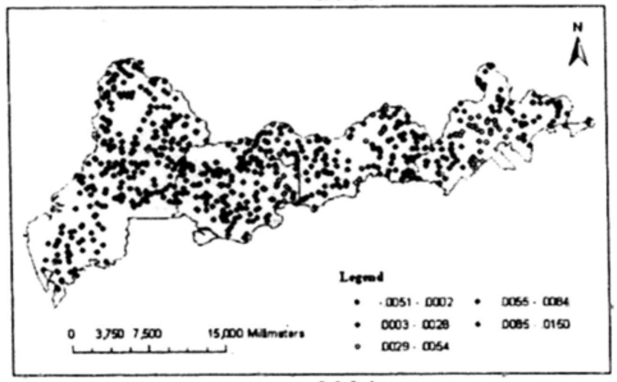
1996



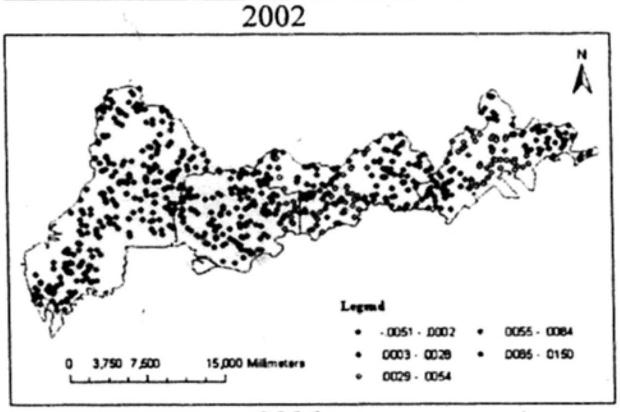
2000



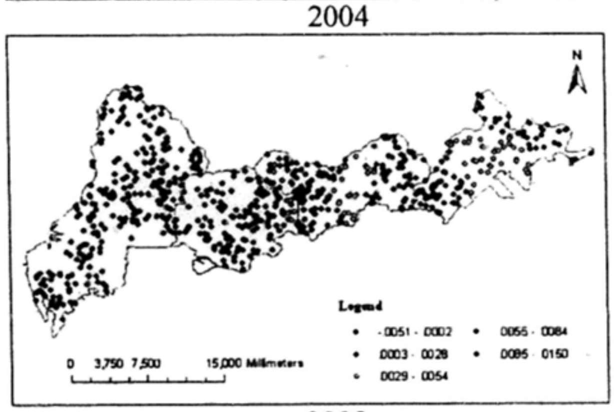
2002



2004

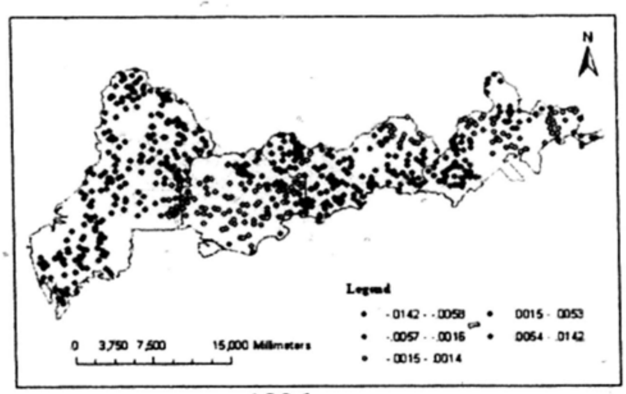


2006

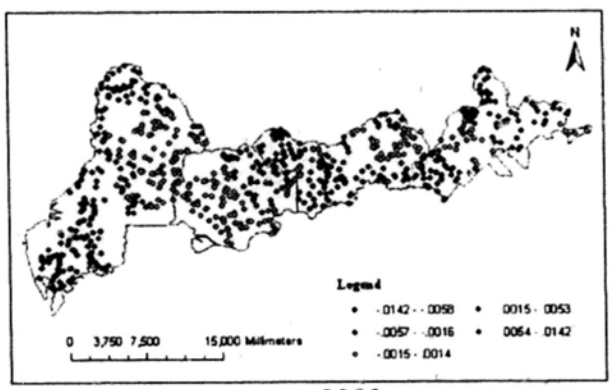


2008

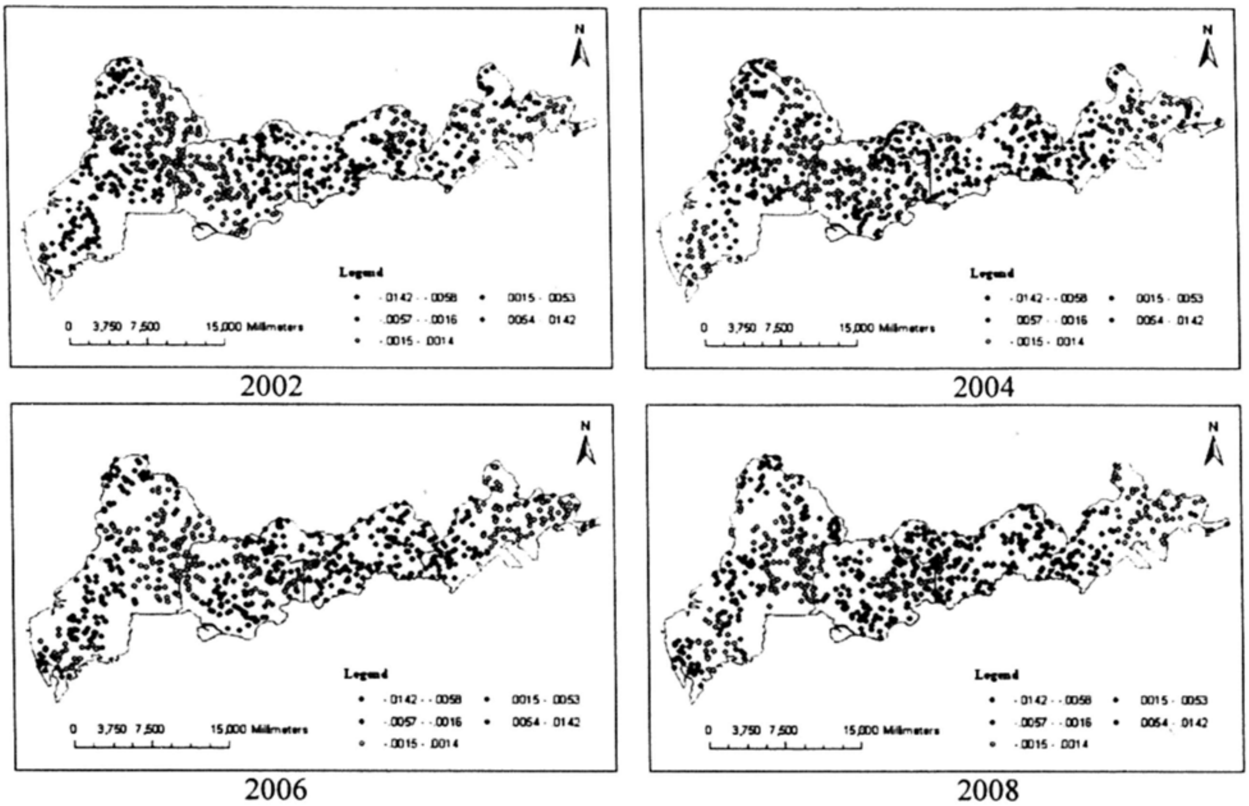
Spatio-temporal variation of parameters for transportation/commercial/others: Distance to Financial Centre



1996



2000



Spatio-temporal variation of parameters for transportation/commercial/others: Distance to Industrial Centre

Figure 6.5: Spatio-temporal variation of parameters

6.5 Conclusion

This chapter proposed a Generalized Spatio-temporal logit Model (GSTLM) to simultaneously explore spatio-temporal non-stationarity and autocorrelations normally found in spatio-temporal data whilst considering the individual effect.

The experimental results from the case study of SEZ in Shenzhen evince that the GSTLM achieved a better modeling accuracy than both the GTWLM, which deals with spatio-temporal non-stationarity only and the ST-PLM, which deals with spatio-temporal autocorrelation only. Compared with the local models, the GSTLM can improve the PCP of MNLM, GTWLM, and ST-PLM from 74.1%, 82.3% and 79.4% to 85.9%, respectively. McNamara's test demonstrated that there was a significant difference among MNLM, GTWLM, ST-PLM, and GSTLM. Besides, the Kno, which is used to evaluate the model's overall success, indicates that the GSTLM is the most successful model. The results of the t test also show that the individual effect is significant in most of the observations and spatio-temporal autocorrelation exists in some observations. Consequently, it can be concluded that it is meaningful to incorporate spatio-temporal non-stationarity and autocorrelations simultaneously.

More importantly, the GSTLM also allows the model parameters to vary across space and time, which provides deep insights into the spatio-temporal variations of the land use pattern. It has been shown that the spatio-temporal variability of each factor influencing land use pattern presents varying patterns. It has been found that some driver factors even have opposite signs for space and time combinations. In general, the GSTLM analysis reveals different effects of the land use determinants over varying space and time combinations.

Some other limitations also exist. For instance, only seven temporal periods have been considered due to the unavailability of data. The inadequacy in temporal effect may degrade the model performance of GSTLM. A random sampling had to be performed to reduce the regression points to ease the computational load. Besides these, some other factors need to be studied further, for e.g., the statistical properties of the estimators.

6.6 Chapter Summary

This chapter introduced two important problems in spatio-temporal data modeling: spatio-temporal non-stationarity and spatio-temporal autocorrelation and briefly summarized the earlier work. Then, to construct and estimate both the problems simultaneously, the GTWLM was integrated with the ST-PLM to be GSTLM. Then the chapter offered the detailed estimated technique for the GSTLM with maximum simulated likelihood estimation and Generalized Cholesky decomposition procedure. In section 6.4, the study of land use change by GSTLM in SEZ, Shenzhen was performed and reported. Subsequently, a spatio-temporal analysis of land use in SEZ, Shenzhen was conducted and the conclusions were drawn accordingly.

Chapter 7: Conclusions and Recommendations

7.1 Summary and Conclusions

The study of land use change is a prerequisite to understanding the complexity of land use change, forecasting future trends of land use change, and evaluating the ecological impacts of land use change. In order to better address some important issues in land use change study, this research aims to construct novel statistical models to establish spatio-temporal relationship between land use and explanatory factors. The outcome of this research, which focuses on the analysis, could benefit urban planners and policy makers in their efforts to effectively understand the land use change process from a spatio-temporal perspective.

A review of the causal factors driving land use changes reported in the literature is provided in this study. Furthermore, a variety of techniques for land use change modeling are briefly reviewed and discussed. In particular, studies on spatio-temporal model of land use change are introduced. Advantages and limitations of those techniques are also provided. It is found that some challenges still exist within the domain of 'spatio-temporal modeling' for land use change. With due consideration to the powerful explanatory capacity, the logistic regression framework is selected in this research to address three major problems: spatio-temporal non-stationarity, spatio-temporal autocorrelation and individual effect.

Ever since the economic reform, the city of Shenzhen has undergone dramatic growth and development. During the past decade, Shenzhen has witnessed drastic changes in the land use change pattern. Considering its 'one city, two system' framework, Special Economic Zone (SEZ) is selected as the study area for this research. Since the data involved in the study area is huge, a random sampling from different years is performed to obtain the data set. In view of the aforementioned problems, this research proposes three spatio-temporal logit models for studying the land use change. These are the GTWLM, ST-PLM, and GSTLM, which have the potential to effectively address the research challenges.

The GTWLM, which is able to consider spatio-temporal non-stationarity, includes temporal data in a spatio-temproal framework by proposing a spatio-temporal distance. Also, TWLM and GWLM, which consider temporal non-stationarity and spatial non-stationarity respectively, are introduced. Compared with the global model (MNLM), TWLM and GWLM increase the PCP values from 74.1% to 75.4% and 79.2%, respectively, and GTWLM yields a considerably higher PCP of 82.3%. McNamara's test shows that the differences between those models are significant. The kappa coefficients also reveal that the GTWLM is better than MNLM, TWLM and GWLM. More importantly, TWLM, GWLM, GTWLM can provide much more information on spatial and temporal relationship, which facilitates model development and leads to a better understanding of land use change process.

In ST-PLM, spatio-temporal autocorrelation is considered in the random individual effect component ε_i^j with an assumption that the autocorrelation between ε_i^j is inversely proportional to the spatio-temporal distance between them. The ST-PLM incorporates the spatio-temporal autocorrelation and individual effect in one model, which is ignored by the MNLM. As a result, the ST-PLM model achieves a higher overall PCP (79.4%) than MLNM (74.1%). Also, the accuracies for undeveloped, industrial and commercial/transporation/others achieved by ST-PLM are better (i.e., 86.4% vs. 90.4%; 50.7% vs. 56.0%; 64.1% vs. 76.6%). Both the McNamara's test and the AIC test corroborate the superior performance of the ST-PLM. Besides, the Kno evinces that the ST-PLM achieves a better result than MNLM consistently. Specifically, ST-PLM's ability to specify location is better than MNLM's, and the two models have the same ability to specify quantity.

Considering all the challenges in one model is an attempt to concentrate on the integration of GTWLM and ST-PLM. Thus, the GSTLM has been proposed to explore spatio-temporal non-stationarity and autocorrelations simultaneously, whilst considering the individual effect. The experimental results of the case study in the SEZ demonstrate that the GSTLM achieves a better modeling accuracy than both the GTWLM, which deals with spatio-temporal non-stationarity only and the ST-PLM, which deals with spatio-temporal autocorrelation and individual only. Compared with the other models, GSTLM can improve the PCP of MNLM, GTWLM, and ST-PLM from 74.1%, 82.3% and 79.4% to 85.9%, respectively. The McNamara's test shows that there is a significant difference among MNLM, GTWLM, ST-PLM and GSTLM. Here, the PCP improvement of GSTLM on MNLM is less than the sum total of the improvement on MNLM by GTWLM and ST-PLM. This is so because allowing for non-stationarity in the regression parameters can account for at least some, and possibly a large part of the autocorrelation in terms of error in a global model calibrated with spatio-temporal data. Kappa coefficients also indicate that the GSTLM is the most successful model. The results of t test also demonstrate that the individual effect is significant in most of the observations and spatio-temporal autocorrelation exists in some observations. Similar to the GTWLM, the GSTLM also allows the model parameters to vary across space and time, which provides deep insights into the spatio-temporal variations of the land use pattern. It has been shown that the spatio-temporal variability of each factor influencing land use pattern leads to different patterns.

Overall, the research has led to the following findings:

1. Individual effect, non-stationarity, and autocorrelation in space and time do exist in this study area. The proposed models, which consider those problems, can greatly improve the accuracy and reliability of land use change study especially in a spatio-temporal framework.

2. The form of spatio-temporal distance, which is obtained by combining by spatial distance and temporal distance, is case based. For land use change study in the SEZ, Shenzhen, the balance ratio between spatial distance effect and temporal distance effect (τ) is 0.1826.

3. The small value of τ suggests that the spatial effect is greater than temporal effect in this study area. The main reason is that only seven temporal periods are available. The inadequacy in temporal information can be expected to degrade the model performance of TWLM which considers temporal non-stationarity only.

4. The selection of a particular spatio-temporal weighting function significantly influences the results. In this study, the exponential function with adaptive bandwidth is adopted.

5. The parameter estimates demonstrate a remarkable spatial and temporal change in SEZ, Shenzhen. Most of the parameters show a significant difference between Yantian districts and others. It indicates that the determinants of land use pattern have varying effects in different locations. Although the temporal space of the study is only ten years, the temporal variation of parameter estimates still exists.

7.2 Recommendations for future work

While the GTWLM, ST-PLM, and GSTLM have been applied for land use change modeling in SEZ, Shenzhen in this study, there are still a number of unanswered questions concerning the application of those models.

Although the spatio-temporal distance has been proposed to extend land use change models to spatio-temporal framework, more flexible combination of spatial and temporal distances need to be considered in future studies. In this study, a simple "+" operator is adopted to measure the spatio-temporal distance with a linear combination between spatial distance and temporal distance. Future work should focus on constructing more operators "$\otimes$"to obtain the spatio-temporal distance.

$$(d^{ST})^2 = (d^s)^2 \otimes (d^T)^2 \tag{7.1}$$

The results of the spatio-temporal models are, in part, dependent on the weighting function selected. All the three models employ exponential distance-decay-based weighting functions to construct a spatial-temporal weight matrix. Many other distance-decay-based weighting functions forms are available in reality, for e.g. Gaussian weighting function and tri-cube weighting function. The theory to guide the selection of a particular spatio-temporal weighting function should be studied in detail.

Prior to the application of the proposed models to analyze land use change, it is important to consider the following questions carefully:

- Does a particular set of local parameter estimates exhibit significant spatio-temporal variation?
- Does the land use data exhibit significant spatio-temporal correlation and indicidual effect?

Future work also should introduce the formal statistical test for spatio-temporal autocorrelation, spatio-temporal non-stationarity, and individual effect. This is important as they are inherently present in spatio-temporal land use data. Besides, the statistical properties of the estimators in proposed models remain to be discovered.

Besides describing and explaining land use change, an important purpose for land use change analysis is to predict future land use. The models proposed in this study for the spatio-temporal analysis of land use change focus more on description and explanation rather than prediction. Based on the proposed models, both spatial and temporal prediction should be studied in the future research. Alternative solutions to use GSTLM to predict future land use type involves employing spatio-temporal interpolation method, which can find the parameters in the future and calculate the probability for each land use type.

One of the primary challenges in future studies involves the problem of computational burden. In the proposed models, the GTWLM has several parameters to be determined, leading to the optimal parameters selection with CV procedure involving heavy computation. The proposed algorithm for ST-PLM to decompose the correlation matrix and obtain the simulated likelihood function is also a computationally-intensive task. Additionally, the GSTLM, which integrates GTWLM and ST-PLM, needs to estimate the ST-PLM in each cell for each iterative calculation. Considering the huge land use data, it's impossible to employ the models for the whole data set. A sampling method is usually employed to reduce the data. However, some spatial and temporal information will be lost in the sampling process. The optimal method is to draw as much samples as possible and optimize the estimation program. The possible solution is to employ intelligent algorithms, such as Genetic algorithms and ANT algorithms, or enhance performance using distributed parallel technique, which needs to be studied further.

Bibliography

Agarwal, C., Green, G.M., Grove, J.M., Evans, T.P., Schweik, C.M., 2002. A review and assessment of land-use change models: dynamics of space, time, and human choice. Gen. Tech. Rep. NE-297. Newton Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 61 p.

Alessandro, P., and Gerald, C. Nelson. 2009. Land Use Change with Spatially Explicit Data: A Dynamic Approach, *Environmental and Resource Economics*, 43(2): 209-229.

Alonso, D., and R. V. Sole. 2000. The DivGame Simulator: A stochastic cellular automata model of rainforest dynamics. *Ecological Modelling* 133 (1/2): 131-141.

Alig, R. J., J. D. Kline, M. Lichtenstein. 2004. Urbanization on the US landscape: looking ahead in the 21st century. *Landscape and Urban Planning*, 69: 214-234.

An, L. and Brown, D.G., 2008, Survival analysis in land-change science: integrating with GIScience to address temporal complexities. *Annals of Association of American Geographers*, 98(2), pp. 1–22.

Andreoni, J. A., and J. H. Miller. 1993. Rational cooperation in the finitely repeated Prisoner's Dilemma: Experimental evidence. *Economic Journal* 103 (418): 570-85.

Anselin, L., 1988, *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic, 284 pp.

Anselin, L. and A. Bera. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics." In A. Ullah and D. Giles (eds.), *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, 237–289.

Anselin, L., 1998. GIS research infrastructure for spatial analysis of real estate markets. J. *Housing Res.*, 9, 113–133.

Auyang, S. Y. 1998. Foundations of Complex Systems Theories: in economics, evolutionary biology, and statistical physics. Cambridge University Press, Cambridge, New York, Melbourne.

Axtell, R. L., and J. M. Epstein. 1994. Agent-based modeling: Understanding our creations. *The Bulletin of the Santa Fe Institute*: 28-32.

Baker, W. L. 1989. A review of models in landscape change. *Landscape Ecology* 2

(2): 111-133.

Balling, R. J., J. T. Taber, M. Brown, and K. Day. 1999. Multiobjective urban planning using a genetic algorithm. *ASCE Journal of Urban Planning and Development* 125 (2): 86-99. 53

Balzter, H., P. W. Braun, and W. Kohler. 1998. Cellular automata models for vegetation dynamics. *Ecological Modelling* 107 (2/3): 113-125.

Başkent, E. Z. and A. İ. Kadıoğulları. 2007. Spatial and temporal Dynamics of land use pattern in Turkey: A case study in İnegöl. *Landscape and Urban Planning*, 81(4): 316-327.

Batty, M., Y. Xie, and Z. Sun. 1999. Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems* 23 (3): 205-233.

Batty, M. 2001. Agent-based pedestrian modeling. *Environment and Planning B* 28 (3): 321-326.

Benenson, I. 1998. Multiagent simulations of residential dynamics in the city. *Computers, Environment and Urban Systems* 22 (1): 25-42.

Berger, T. 2001. Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes, and policy analysis. *Agricultural Economics* 25 (2-3): 245-260.

Bockstael, N. 1996. Modeling economics and ecology: The importance of a social perspective. *American Journal of Agricultural Economics*, 78(5): 1168-1180.

Brown, D. G., B. C. Pijanowski, and J. D. Duh, 2000. Modeling the relationships between land use and land cover on private lands in the Upper Midwest, USA. *Journal of Environmental Management*, 59: 247-263.

Brown, D. G., P. Goovaerts, A. Burnicki, and M. Y. Li. 2002. Stochastic simulation of land-cover change using geostatistics and generalized additive models. *Photogrammetric Engineering & Remote Sensing*, 68(10): 1051-1061.

Brunsdon, C., Fotheringham, A. S., and Charlton, M., 1996. Geographically weighted regression: a method for exploring spatial non-stationarity, *Geographical Analysis*, 28, 281-298.

Brunsdon, C., Fotheringham, A. S., and Charlton, M., 1999. Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Sciences*, 39, 3, 497-524.

Can, A., 1992. Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22, 453-474.

Can, A. and Megbolugbe, I., 1997. Spatail dependence and house price index construction, *Journal of Real Estate Finance and Economics*, 14, 203-222.

Chen, S. H., and C. H. Yeh. 2001. Evolving traders and the business school with genetic programming: A new architecture of the agent-based artificial stock market. *Journal of Economic Dynamics and Control* 25 (3-4): 363-93.

Cheng, J., and Masser, I., 2003, Urban growth modeling: a case study of Wuhan city, PR China. *Landscape and Urban Planning*, 62, 199-217.

Clarke, K. C., and Gaydos, L. J., 1998, Loose-coupling a CA model and GIS: long-term urban growth prediction for San Franciso and Washington/Baltimore. *International Journal of Geographical Information Science*, 12, 699-714.

Dombrow, J., Knight, J.R., and Sirmans, C.F., 1997. Aggregation bias in repeat sales indices. *Journal of Real Estate Finance and Economics*, 14, 75-88.

Dubin, R.A., 1992. Spatial correlation and neighborhood quality. *Regional Science and Urban Economics*, 22, 432–452.

Dubin, R. 1995. Estimating Logit Models with Spatial Dependence. in New Directions in Spatial Econometrics. Eds L. Anselin, R. Florax (Springer-Verlang, New York).

Eckert, J. K. (ed.), 1990. Property Appraisal and Assessment Administration. Chicago, Illinois: International Association of Assessing Officers.

Foody, G. M., 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering & Remote Sensing*, 70(5): 627–633.

Fotheringham, A.S., Charlton M.E., and Brunsdon, C., 1996. The geography of parameter space: an investigation of spatial non-stationarity, *International Journal of Geographical Information Science*, 10, 605 – 627.

Fotheringham, A.S., Charlton M.E., and Brunsdon, C., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30, 1905-1927.

Fotheringham, A.S, Brunsdon C., and Charlton, M.E. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: Wiley.

Gelfand, A. E., Ecker, M. D., Knight, J. R., and Sirmans, C. F., 2001. The dynamics of location in home prices. Working Paper, University of Connecticut.

Gelfand, A.E., Kim, H-J., Sirmans, C.F., and Banerjee, S., 2003. Spatial modeling with spatially varying coefficients processes. *Journal of the American Statistical Association*, 98, 387-396.

Geoghegan, J., S.C. Villar, P. Klepeis *et al*. 2001. Modeling tropical deforestation in the southern Yucatan peninsular region: comparing survey and satellite data. Agriculture, *Ecosystems and Environment*, 85: 25-46.

Gill, J. and King, G., 2004. What to do when your Hessian is not invertible: alternatives to model respecification in nonlinear estimation. *Sociological Methods and Research*, 32(1): 54-87.

Gilruth, P. T., S. E. Marsh, and R. Itami. 1995. A dynamic spatial model of shifting cultivation in the highlands of Guinea, West Africa. *Ecological Modelling* 79: 179-197.

Gimblett, H. R., ed. 2002. *Integrating Geographic Information Systems and Agent-Based Modeling Techniques for Simulating Social and Ecological Processes*. New York: Oxford University Press.

Goodman, A. C. and Thibodeau, T. G., 1998, *Housing market segmentation. Journal of Housing Economics*, 7, 121–143

Greene, W., 2000. *Econometric Analysis*. Upper Saddle River: Prentice-Hall.

Gumpertz, M. L., C. Wu, and J. M. Pye. 2000. Logistic Regression for Southern Pine Beetle Outbreaks with Spatial and Temporal Correlation. *Forest Science*, 46(1): 95-107.

Hanjin Shi, Lianjun Zhang, and Jianguo Liu. 2006. *Anew spatial-attribute weighting function for geographically weighted regression*. NRC press.

Hanjin Shi, Lianjun Zhang, and Jianguo Liu. 2006. *A new spatial-attribute weighting function for geographically weighted regression*. NRC press.

He, F. L., J. Zhou, and H. Zhu. 2003. Autologistic regression model for the distribution of vegetation. *Journal of Agricultural, Biological, and Environmental Statistics*, 8(2): 205-222.

Hosmer, D. and S. Lemeshow. 1989. *Applied logistic regression*. New York: John Wiley and Sons, Inc.

Huang, B. and Claramunt, C., 2005. Spatiotemporal data model and query language for tracking land use change. *Transportation Research Record: Journal of the Transportation Research Board*, 1902: 107–113.

Huang, B., Zhang, L., and Wu, B., 2009a. Spatio-temporal analysis of rural-urban land conversion. *International Journal of Geographical Information Science*, 23(3): 379 – 398.

Huang, B., Xie, C. L., and Tay, R., 2009b. Land use change modeling using unbalanced support vector machines. *Environment and Planning B*, 36(3): 398 – 416

Huang, B., Wu, B., and Barry, M., 2010. Geographically and Temporally Weighted Regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science,* 24(3): 383 – 401.

Huffer, F. W., and H. L. Wu. 1998. Markov Chain Monte Carlo for Auto-logistic Regression Models with Application to the Distribution of Plant Species. *Biometrics*, 54, 509-524.

Hurvich, C.M., Simonoff, J.S., and Tsai C.L., 1998. Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society (Series B)*, 60, 271-293.

Irwin, E. J., and Geohegan J., Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems, and Environment*, 85: 7-24.

Journal, A., and Huijbregis C.H., 1978. *Mining geostatisticals*. London, UK: Academic Press.

Kafadar K (1994) Choosing among two-dimensional smoothers in practice. *Comp Stat & Data Analysis* 18:419–439

Kaimowitz, D., and A. Angelsen. 1998. *Economic Models of Tropical Deforestation: A Review*. Centre for International Forestry Research, Jakarta, Indonesia.

Kollman, K., J. H. Miller, and S. E. Page. 1997. Political institutions and sorting in a Tiebout model. *American Economic Review* 87 (5): 977-92.

Krider, R. E., and C. B. Weinberg. 1997. Spatial competition and bounded rationality: retailing at the edge of chaos. *Geographical Analysis* 29 (1): 16-34.

Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," *SIAM Journal of Optimization*, Vol. 9, Number 1, pp. 112–147, 1998.

Lambin, E.F., Geist, H.J., and Lepers, E., 2003. Dynamics of land-use and landcover change in tropical regions. *Annu. Rev. Environ. Resour.* 28, 205–241.

Langville, A. N., and Steward W.J., 2004. The Kronecker product and stochastic automata networks. *Journal of computational and applied mathematics*, 164, 429-447.

Leung, Y., Mei, C. L., and Zhang, W. X., 2000. Statistical tests for spatial non-stationarity based on the geographically weighted regression model. *Environment and Planning A*, 32, 9–32

Longley, P., G. Higgs, and D. Martin. 1994. The predictive use of GIS to model property valuations. *International Journal of Geographical Information Systems* 8 (2): 217-235.

Ludeke, A. K., R. C. Maggio, and L. M. Reid. 1990. An analysis of anthropogenic deforestation using logistic regression and GIS. *Journal of Environmental Management* 31: 247-259. 66

Manard, Scott. 1995. *Applied Logistic Regression Analysis*. Series: Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage.

Manel S., J.M. Dias, S.J. Ormerod. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species' distribution: a case study with a Himalayan river bird. *Ecological Modelling*, 120: 337-347.

Marimon, R., E. McGratten, and T. J. Sargent. 1990. Money as a medium of exchange in an economy with artificially intelligent agents. *Journal of Economic Dynamics and Control* 14: 329-373.

Marks, R. E. 1992. Breeding hybrid strategies: Optimal behavior for oligopolists. *Journal of Evolutionary Economics* 2: 17-38.

McMillen, D. P., 1996. One hundred fifty years of land values in Chicago: a nonparametric approach. *Journal of Urban Economics*, 40, 100-124.

McMillen, D. P., McDonald J.F., 1997. A nonparametric analysis of employment density in a polycentric city. *Journal of Regional Science*, 37, 591-612.

Miller, J. H., and M. Shubik. 1992. *Some Dynamics of a Strategic Market Game with a Large Number of Agents*. Santa Fe, NM: Santa Fe Institute Publication 92-11-057.

Munroe, D.K., Southworth, J., and Tucker, C. M., 2004. Modelling Spatially and Temporally Complex Land-Cover Change: The Case of Western Honduras. *The Professional Geographer*, 56(4): 544-559.

Natural Resources Conservation Service, United States Department of Agriculture. 1999. Summary Report 1997 Natural Resources Inventory. Web http://www.nhq.nrcs.usda.gov/NRI, accessed on March 20, 09.

Nunes C, Auge J I. 1999. Land use and land cover change: implementation strategy. In: IHDP report No. 10.

Openshaw S, Taylor P. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley N. (ed) Statistical applications in the spatial sciences. Pion, London, pp 127–144

O'Sullivan, D. 2001. Graph-cellular automata: a generalised discrete urban and regional model. *Environment and Planning B* 28 (5): 687-706.

Pace, R. K., Barry, R. and Sirmans, C. F. 1998, Spatial statistics and real estate, *Journal of Real Estate Finance and Economics*, 17, 5– 13.

Paez, A., Uchida, T., and Miyamoto K., 2002. A general framework for estimation and inference of geographically weighted regression models: location-specific kernel bandwidths and a test for local heterogeneity. *Environmental and Planning A*, 34, 733–754.

Parker, D. C. 2000. Edge-effect externalities: *Theoretical and empirical implications of spatial heterogeneity.* Ph D. diss. University of California at Davis, Davis, CA

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., and Deadman, P., 2003. Multi-agent system models for the simulation of land-use and land-cover change: a review, *Annals of the Association of American Geographers,* 93(2): 314–337.

Pavlov, A., 2000. Space varying regression coefficients: a semi-parametric approach applied to real estate markets. *Real Estate Economics*, 28, 249-283.

Pontius, Jr., and Gilmore, R., 2000. Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering & Remote Sensing*, 66(8): 1011-1016.

Pontius, Jr., Gilmore, R., Walker, R., Yao-Kumah, R., Arima, E., Aldrich, S., Caldas, M., and Vergara, D., 2007. Accuracy assessment for a simulation model of Amazonian deforestation. *Annals of the Association of American Geographers*, 97(4): 677-695.

Portugali, J., I. Benenson, and I. Omer. 1997. Spatial cognitive dissonance and sociospacial emergence in a self-organizing city. *Environment and Planning B* 24: 263-285.

Ramankutty, N., and Foley, J. A., 1999. Estimating historical changes in global land cover: croplands from 1700 to 1992. *Global Biogeochemical Cycles*, 13(4): 997-1027.

Rindfuss, R. R., Walsh, S. J., Turner II, B. L., Fox, J., and Mishra, V., 2004. Developing a science of land change: challenges and methodological issues. *PNAS*, 101(39): 13976–13981.

Sanders, L., D. Pumain, H. Mathian, F. Guérin-Pace, and S. Bura. 1997. SIMPOP - a multiagent system for the study of urbanism. *Environment and Planning B*: Planning and Design 24: 287-305.

Silvertown, J., S. Holtier, J. Johnson, and P. Dale. 1992. Cellular automaton models of interspecific competition for space: The effect of pattern on process. *Journal of Ecology* 80 (3): 527-534.

Sirmans, G., Macpherson, D., and Zietz, A., 2005. The composition of hedonic pricing models. *Journal of real Estate Literature*, 13, 1, 3-41.

Talbot TO, Kulldorff M, Forand SP, Haley VB (2000) Evaluation of spatial filters to create smoothed maps of health data. *Stat in Med* 19:2399–2408

Theobald, D. M. and Hobbs, N. T., 1998. Forecasting rural land use change: a comparison of regression and spatial transition-based models. *Geographical and Environmental Modeling*, 2, 65-82.

Torrens, P. M., and D. O'Sullivan. 2001. Cellular automata and urban simulation: Where do we go from here? *Environment and Planning B* 28 (2): 163-168.

Train, K., 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Turner, B. L. and Meyer, W. B., 1991. Land use and land cover in global environmental change: considerations for study. *International Social Sciences Journal*, 130: 667-679.

Turner, M. G., R. Costanza, and F. Sklar. 1989. Methods to evaluate the performance of spatial simulation models. *Ecological Modelling* 48 (1/2): 1-18.

Upton, G. J. G., and B. Fingleton. 1985. *Spatial Data Analysis by Example*, Volume 1: Point Pattern and Quantitative Data (New York: John Wiley and Sons).

Veldkamp, A., and L. O. Fresco. 1996. CLUE: A conceptual model to study the conversion of land use and its effects. *Ecological Modelling* 85 (2/3): 253-270.

Veldkamp, A., and E. F. Lambin. 2001. Predicting land-use change. *Agriculture, Ecosystems, and Environment* 85 (1-3): 1-6.

Verburg, P. H., W. Soepboer, A. Veldkamp, R. Limpiada, V. Espaldon, and S. A. Sharifah Mastura. 2002. Modeling the spatial dynamics of regional land use: the CLUE-S model. *Environmental Management*, 30(3): 391-405.

Wang, P., 2006. *Exploring Spatial Effects on Housing Price: the Case Study of the City of Calgary*. Master dissertation, University of Calgary, Canada.

Weinberg, M., C. L. Kling, and J. E. Wilen. 1993. Water markets and water quality. *American Journal of Agricultural Economics* 75 (2): 278-91.

Weng, Q., 2002. Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS, and stochastic modeling. *Journal of Environmental Management*, 64, 273-284.

Wooldrige, J. M., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge.

Wu, F., and Yeh, A. G., 1997, Changing spatial distribution and determinants of land development in Chinese cities in the transition from a centrally planned economy to a socialist market economy: a case study of Guangzhou. *Urban Studies*, **34**, 1851-1879.

Wu, F., 1998, SimLand: a prototype to simulate land conversion through the ntegrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science*, 12, 63-82

Yu, D-L, 2006. Spatially varying development mechanisms in the Greater Beijing Area: a geographically weighted regression investigation, *The Annals of Regional Science*, 40, 1, 173-190.

Yuan, M. and Stewart, K., 2008. *Computation and Visualization for the Understanding of Dynamics in Geographic Domains: A Research Agenda*. CRC/Taylor and Francis.

Zhu, J., H. Huang, J. Wu. 2005. Modelling Spatial-Temporal Binary Data Using Markov Random Fields. Journal of Agricultural, Biological & Environmental Statistics, 10(2): 212-2

APPENDIX A: GENERALIZED CHOLESKY DECOMPOSITION

The correlation matrix is commonly decomposed by cholesky decomposition to give the lower-triangular L in the Monte Carlo method for simulating systems with multiple correlated variables. Applying this to a vector of uncorrelated samples, v, produces a sample vector $L*v$ with the covariance properties of the system being modeled.

Classic cholesky decomposition is employed to obtain the lower-triangular L from a positive-definite matrix C (for any x vector except $x = 0$, $x\ C\ x>0$). It proceeds via the following decomposition:

$$C = L_o\ D\ L'_o \tag{A.1}$$

where D is a positive diagonal matrix, L_o, D, and C are $k*k$ matrixes. There are k stages in the algorithm corresponding to the k-dimensionality of the input matrix C. The j th step ($1 \leq j \leq k$) is characterized by two operations:

$$D_{j,j} = C_{j,j} - \sum_{\ell=1}^{j-1} L_{j,\ell}^2 D_{\ell,\ell} \tag{A.2}$$

$$L_{i,j} = \left[C_{i,j} - \sum_{\ell=1}^{j-1} L_{j,\ell} L_{i,\ell} D_{\ell,\ell} \right] \Big/ D_{j,j} \qquad i = j+1,\ldots,k \tag{A.3}$$

So that upon the completion of the algorithm, the square root of D is multiplied by L_o to obtain L for cholesky decomposition.

In this study, the matrix ($corr(\varepsilon^{i},\varepsilon^{j})$) may be non-positive-definite input matrices, a divide-by-zero problem will be caused in (A.3), and non-positive-definite matrices cause the sum in (A.2) to be greater than $C_{j,j}$. Hence, there is a need to find a nonnegative diagonal matrix, E, such that $C + E$ is positive definite and the diagonal values of E are as small as possible. Gill *et al.*'s (1981) proposed an approach to handle non-positive-definite matrices. The Cholesky algorithm provided as (A.2) and (A.3) are rewritten in matrix notation. The j th submatrix of its application at the j th step is

$$C_j = \begin{bmatrix} c_{j,j} & c'_j \\ c_j & C_{j+1} \end{bmatrix} \qquad \text{(A.4)}$$

where $c_{j,j}$ is the j th pivot diagonal; c'_j is the row vector to the right of $c_{j,j}$, which is the transpose of the c_j column vector beneath $c_{j,j}$; and C_{j+1} is the *(j + 1)*th submatrix. The j th row of the L matrix is calculated by $L_{j,j} = \sqrt{c_{j,j}}$, and $L_{(j+1):k,j} = c_{(j+1):k,j} / L_{j,j}$.

The *(j + 1)*th submatrix is then updated by

$$C^{\bullet}_{j+1} = C_{j+1} - \frac{c_j c'_j}{L^2_{j,j}} \qquad \text{(A.5)}$$

Suppose that at each iteration, we defined $L_{j,j} = \sqrt{c_{j,j} + \delta_j}$. Therein, the size of δ_j, which is a small positive integer sufficiently large to ensure $C_{j+1} > \frac{c_j c'_j}{L^2_{j,j}}$, can be calculated by the method proposed by Schnabel and Eskow (1990).

$$\delta_j = \max(\in_m,\ -C_{j,j} + \max(\|\alpha_j\|,\ (\in_m)^{1/3} \max(diag(C))),\ E_{j-1,j-1}) \qquad \text{(A.6)}$$

where $\in_m$ is the smallest positive number that can be represented on the computer used to implement the algorithm.

APPENDIX B: SHUFFLED HALTON SEQUENCE

The process of generating a random shuffled Halton sequence has two steps as follows:

(1) Generate Halton sequence to cover the space;

(2) For each dimension, randomly order the numbers in the original Halton sequence.

Halton sequence for a standard normal density

Halton sequence, which can covers the interval uniformly, is defined in terms of a given number called as a prime. For instance, the Halton sequence for prime 3 is created by dividing the unit interval into three parts with breaks at 1/3 and 2/3. The first terms in the sequence are these breakpoints: 1/3 and 2/3. Then each of the three segments is divided into thirds, and the breakpoints for these segments are added to the sequences in a particular way (Figure B.1).

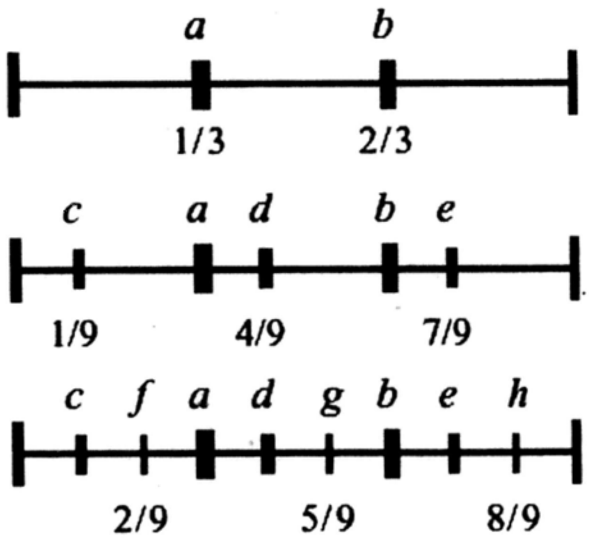


Figure B.1: Halton sequence for prime 3

Thus, the sequence based on prime 3, is

{1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9, 1/27, 10/27, 19/27, 4/27, 27/13 ....}

It is Halton draw for a uniform density. To obtain a sequence of points for a standard normal density, the inverse cumulative normal is evaluated at each element of the Halton sequence. The resulting sequence is

$$\begin{aligned}
&\phi^{-1}(1/3)=-0.43,\\
&\phi^{-1}(2/3)=0.43,\\
&\phi^{-1}(1/9)=-1.2, \qquad \text{(B.1)}\\
&\phi^{-1}(4/9)=-0.14,\\
&\phi^{-1}(7/9)=0.76,\\
&\quad\vdots
\end{aligned}$$