Micro-simulation Urban Land Use Change Modelling: the Case of Ladprao, Bangkok, Thailand

Sirivilai Teerarojanarat
BA (Hons), Geography (Chulalongkorn University, Thailand)
MSc, Environmental Remote Sensing & Geoinfo. for Development (AIT, Thailand)

Thesis submitted for the Degree of
Doctor of Philosophy

Department of Geomatics
School of Civil Engineering and Geosciences
Newcastle University
Newcastle upon Tyne

October 2007

THESIS CONTAINS

CD

Abstract

This thesis focuses on modelling the spatial pattern of urban growth of Ladprao, a district of Bangkok, Thailand. The first part of the thesis reviews the urban growth and land use change problems in Bangkok as well as the current role of urban planning and its limitations, in order to provide the context of this study. A GIS-based cellular automata (CA) model has been developed, where the multinomial logistic regression (MNL) and multicriteria decision analysis (MCDA) methods have been integrated to identify the potential cells for development. Customized tools have been developed using a VBA macro within the ARCGIS environment to facilitate the implementation of urban simulation. The developed model has been applied to replicate the spatial pattern at the detail of the district level, focusing on the change of land from vacant to residential, commercial, and industrial during the period 1993 - 2001. Validation of the model has been undertaken through the comparison between the 2001 simulated and actual land use maps.

The simulation was unsuccessful in reproducing the actual growth. In terms of the spatial agreement, the overall accuracy was about 30% (31.59% and 32.01% with MNL and MCDA respectively). In terms of urban morphology, the results showed the emergence of urban development in a space-filling pattern. Urban growth over discrete time-steps acted as a process of building accretion, appearing as a growing cluster around the existing development. In the actual pattern, the emergence of development was dispersed over the study area. The unexpected, but interesting, results of this observation have led to the conclusion of the three possible reasons; the inappropriateness of the CA approach to simulate the pattern of urban district level growth, the inability to include all significant development factors of the study site, and finally the distinctive characteristics of Ladprao and Bangkok area itself.

Though the results are unpromising, the developed model can be considered as the first in the Bangkok area that attempts to be used as a spatial micro simulation tool operated at the district level. Future research work, if data permits, also suggests adding more development factors, adapting the agent-based modelling to the application, and extending the simulation to the growth of other areas of Bangkok both in the district and city level in order to help improve the understanding of Bangkok's growth.

Acknowledgements

First and foremost, I would like to express my deep gratitude to The Royal Thai Government for giving me a scholarship to pursue my PhD research at Newcastle University from 2003-2007, for which I feel truly grateful and appreciative.

My supervisors Dr. David Fairbairn and Dr. Stuart Barr have been incredible mentors and teachers. I would like to express my sincere thanks to them for their supervision, guidance, and help throughout this research. Their confidence in my ability and attentive support at every stage of my research have been invaluable. I have also greatly benefited from their expertise, accurate criticism, useful and constructive suggestions. The impact of our long discussions and their comments can be felt in the form and substance of this thesis.

I also express my appreciation to various agencies in Thailand, especially staff in Department of City Planning (DCP), Bangkok. Special thanks are due to Khun Thin Hongthong and Dr. Orapim Pimcharoen for providing data, information, and useful suggestions to be used for the study site throughout this research study.

Thanks, also, are owed to my colleagues at the Department of Geography, Faculty of Arts, Chulalongkorn University, Thailand, who had to work harder and had more responsibilities during my research leave. I owe a particular debt of gratitude to Dr. Dusdi Chanlikit who inspired and helped me for studying for a PhD abroad.

I am also indebted to The Royal Victoria Infirmary for giving me a part-time job. The extra earnings from working at the RVI enabled me, during my four-year study in the United Kingdom, to travel across Europe to learn and experience the different cultures and valuable heritages of various countries.

Thank you for the loving support of my friends. I appreciate everyone coming along for this ride as I pursued my doctoral degree. I have received much-appreciated assistance from my friends in Newcastle and in Thailand. Those whose love, understanding, and laughter sustained me throughout.

I would like to express my gratitude to my mother, my father and other family members, especially Na Mo, P'Bae's mother, Na Jum, and Na Tang for their encouragement and support for my sons throughout my academic career. Special thanks to both my sons, Nong Sea and Nong Earth, for their patience and understanding of my long absence. Finally, my gratitude goes to my beloved husband, P'Bae for his love, patience, and understanding throughout my study, even during the most difficult periods. Without his support, the whole thing would never have happened in the first place. I dedicate this dissertation to all of them.

Table of Contents

CHAPTER 6

CHAPTER 7

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Urban Phenomena in Developing Countries

The world's urban population is increasing at a much more rapid rate than the total population as a whole. According to *The 2005 Revision of World Urbanization Prospects* (United Nations, 2006), it is projected that all growth in the period 2000 to 2030 will be concentrated in urban areas as the urban population of the world increased from 732 million (29 per cent of the world's population) in 1950 to 3.2 billion (49 per cent) in 2005, nearly four times as many as in 1950, and is estimated to continue to 4.9 billion (roughly 60 per cent) by 2030. More strikingly, urban population growth is anticipated to be particularly increasing in the urban areas of developing countries at an average 2.2 per cent annually between 2005 and 2030. In particular, the largest percentage increases in the urban population will be in Africa and Asia. In contrast, the urban population of the developed countries is projected to increase modestly - from 900 million in 2005 to 1 billion in 2030.

Although the nature of urban growth in many countries exhibits significant diversity, there are a number of commonly critical considerations that make differences between the urban transition of developing and developed nations (Cohen, 2004; Drakakis-Smith, 2000; Lavalle et al., 2001; Potter, 1992). The first issue concerns the scale and rates of urban population growth. According to the preceding figure of the comparative rise in numbers of urban residents between developing and developed nations, it is the best remarkable indication of the unprecedented scale of growth and the rapid rates of urbanization in developing nations. Secondly, the growth rate has occurred more speedily in low-income developing areas. Thirdly, such rapid urban population growth, as a consequence, has led to the way in which the largest cities have grown at accelerating rates and reached unprecedented sizes. By 2015 the world is expected to have 480 cities with 1 million or more inhabitants. Of these, 359 cities -about two-thirds- will be in developing countries (United Nations, 1997). Such a phenomenon has been referred to as the concept of urban primacy – demographic, economic, social, cultural and political dominance of one city over all others within an urban system (Drakakis-Smith, 2000).

Current urban phenomena in developing nations, previously discussed, are a consequence of three main factors (Cohen, 2004; Drakakis-Smith, 2000; Jones, 1993; Potter, 1992; Todaro, 1984): natural population increase, rural-urban migration, and annexation and reclassification. The most important factor involves the effect of natural increase. It was the post-war periods when medical facilities and health conditions were improved. They, thus, considerably reduced the death rates while the birth rates have remained at high levels. Another factor involves the effect of rural-to-urban migration, which is due to the rapid economic booms in the late 1960s and 1970s. To promote economic growth, several cities in developing countries have experienced a shift from a predominantly agriculture-based to industry-based economy (Jones, 1993). Despite their success in industrial promotion, these changes have tremendously had an impact on the reduction of the percentage of the population employed in rural activities and the distribution of population in urban areas. The final factor is the annexation and the reclassification of previously agriculture land around the periphery as urban, which can be considered as a by-product of both previous factors (Cohen, 2004).

Unavoidably, the effect of urban growth has led to important changes in land use patterns (Beier, 1984; Darin-Drabkin, 1977; Jones, 1993; Kaothien, 1995; Potter, 1992). The emergence of the urban centre as a service-based and industry-based economy encouraging the combination of the increasing use of commercial spaces in the city centre and the improvement of new transportation systems forces the population to increasingly settle further from the city centre and/or on the periphery. Besides, it results in the conversion of the agricultural land on the urban fringe to residential and industrial development, the separation of place of work from residence, and increasing the costs in terms of personal transport and other forms of infrastructure (i.e. water supply, electricity, sewerage and drainage). In developing countries where unplanned policies and practices cannot keep up with a sudden and abrupt demand, it causes the unsuitable land use and land cover to spread widely within many metropolitan areas, especially the creation of shanty towns and squatters. Significant, uncontrollable changes result in a large number of social and physical problems; environment deterioration problems (e.g. increase of air and water pollution), traffic congestion, shortages in urban services and facilities, and major problems of urban poverty (e.g. lack of housing security, limited opportunity for education).

1.2 Urban Growth and Land Use Change Problem in Bangkok, Thailand

The Bangkok Metropolis, currently one of the fortieth largest metropolitan centre in the world, is a good example of a primary city with a total population of a mere 6.6 million in 2005 that is projected to continue to 7.4 million by the end of 2030(United Nations, 2006). At present, the city encompasses 50 districts, covering an area of 1,569 square kilometres. As the capital city, Bangkok functions as Thailand's major economic, administrative, education and health, transportation, and cultural centre. As in many other capital cities in developing countries, Bangkok experienced an economic boom during the 1970s and 1980s. Unavoidably, this has caused rapid urban population growth due to the influx of young migrants from rural areas and from provinces throughout the country in search of better job opportunities. Although such growth can be seen as a positive influence on the Thai economy as a whole, it has inevitably created negative consequences in the context of urban development and land utilization.

In terms of the physical dimensions, the Bangkok Metropolitan Area (BMA) has outgrown its administrative boundary and sprawled into the surrounding provinces of Nonthaburi, Pathumthani, Samut Prakarn, Samut Sakorn, and Nakorn Pathom and formed a region known as the Bangkok Metropolitan Region (BMR). It has spread out to the next five adjacent provinces of Ayutthaya and Saraburi in the north, Ratchaburi and Phetchaburi in the west, and all along the coast towards the east including the provinces of Chon Buri, Chachoengsao, and Rayong, the so-called the Extended Bangkok Metropolitan Region (EBMR) (Guerra, 2004; McGee, 1991; 1994). The pattern of development is, however, primarily located along major highways and roads. Opportunities for more efficient development patterns are difficult due to the lack of secondary infrastructure, such as lack of access to distributor roads, and the inadequacy of piped water and sewage systems (Guerra, 2004).

The impact of rapid population growth also results in the rapid land use conversion of agricultural land to urban area. Many canal systems, which previously played a major role for Bangkok's transportation have been filled and replaced with road transportation. In terms of land usage, Bangkok exhibits two distinct characteristics. Firstly, there is a mixed use of land activities; residential houses, commercial buildings and factories of various scales, shapes, and sizes (Krongkaew, 1996). The emergence of these land use activities increased without any effective control on land usage. They are usually built without well-prepared management of the basic facilities such as drainage and waste management systems. Secondly, most parts of Bangkok's extended area lack an efficient road network (Ross et al., 2000; Webster, 2000). Many systematic and interconnected networks were built inefficiently. Since the government could not provide sufficient roads, many streets were built by private developers for their real estate projects. These streets were built with an unplanned road layout to be connected to the road outside the projects. They caused many areas to have severe traffic problems and many areas to have poor accessibility. Inappropriate land use patterns have consequently led to many social and physical problems. The ineffective road network, for instance, has caused severe traffic congestion, which in turn increases air-pollution and consequently the deteriorating health of urban dwellers. The rapidly growing suburban areas have encountered an inadequate water supply problem, as the supply could not keep up with the abrupt increase in demand. In consequence, it has forced those who live in the suburbs to use underground water, exacerbating land subsidence, which in turn causes severe flooding during the rainy season (Krongkaew, 1996; Webster, 2004).

Hence, there is an urgent need to formulate effective land use plans, specifically under the conditions of rapid urban growth, as an attempt to control, or at least, lessen these land use problems. In turn, this will minimise the possible negative impacts of the growth of urban areas. However, to date, the Bangkok metropolis still lacks an effective land use plan, a cohesive planning policy and, management and cooperation amongst relevant agencies (Chomchan et al., 1990; Kaothien, 1995; Krongkaew, 1996; Webster, 2000). Bangkok Metropolitan Area was established under the Administration of the Bangkok Metropolis Act 1985. Following this Act, the Bangkok Metropolitan Administration was established as a local administrative body having a fully and independently administrative and management responsibility over the Bangkok Metropolitan Area, including the enforcement and implementation of the land use planning policies and plans. However, it was just after 1994 when an official city planning of Bangkok Metropolitan Administration, namely the Department of City Planning (DCP) was formally established with full responsibility for the city planning work for Bangkok.

At present, there are two major laws being enforced in land use planning which have been widely used at the implementation level for Bangkok Metropolis. They are referred to as the Town and Rural Planning Act and the Building Control Act. According to the former Act, a Comprehensive Master Plan (so-called Bangkok Comprehensive Plan for Bangkok area) and the Specific Plan were established. The Bangkok Comprehensive Plan is used as a broad plan, which is intended as a measure for general control and a provision of guidelines for land use planning. The Specific Plan involves the planning and development of a particular area regarded as a zoning ordinance. The latter Act is a more effective form of development control. It has been used in order to help make decisions regarding the granting of permission for project construction by addressing the control measure in terms of building height control, building type control and building use control. Though planning polices have been used over a long period, they are considered poorly-integrated, unable to fit in with the present dynamic urban situation and have inefficient and ineffective control over land use (Chomchan et al., 1990; Cohen, 2004; Kaothien, 1995). The contents of these laws are presently inadequate to manage the rapidly-occurring urbanisation resulting from economic, social and technological advancement since they were established in a context that intended to solve existing problems rather than prevent future land use problems (Chomchan et al., 1990). More importantly, implementation of these plans is poorly-integrated amongst related agencies (Cohen, 2004; Kaothien, 1995) and these plans bear no legal enforcement since there is no obligation to apply for planning permission for the development (Commission of the European Communities (CEC), 1995).

Bangkok's districts have been grouped into five zones: inner city, eastern transition zone (interchangeably, urban fringe), western transition zone, eastern suburban zone, and western suburban zone. They are classified based on their distance from the CBD, by the Department of City Planning (Department of City Planning (DCP), 1999), in order to administer and control the direction of Bangkok's growth. Different zones and different parts of the Bangkok area may vary in terms of urban development processes (e.g. transition from agriculture to commercial area in inner city zone) and different dominant land use patterns (e.g. a district mainly functioning for commercial purposes). In this research study, however, a particular type of land use pattern in a "transition zone" area where there are many vacant areas left for urban development is focused on. The implementation will be conducted on one district at the detail of district or neighbourhood level.

1.3 Potential Applications of GIS in Urban Land Use Change Simulation

Of spatial interest to planners is to monitor and predict future change. Information concerning the future development pattern of urban areas is useful for planners as it allows investigation of areas likely to experience urban change. Technology like GIS (Geographic Information Systems) is considered a useful tool to assist in such purposes. In spite of the provision of the functions and capabilities of GIS in handling spatial and non-spatial information (e.g. data display, data analysis), GIS have still limited capabilities in the area concerning some main parts of the analytical simulation such as dealing with complicated mathematical models and dynamic (space-time process) capabilities (Heywood et al., 2002; Longley and Batty, 2003; Longley et al., 2005; Wu, 1999).

At present, a micro-simulation approach such as cellular automata (CA) is usually integrated with GIS in order to simulate the spatial pattern of urban development. Such integration has been very popular due to the ease of its implementation (Torrens, 2000). Micro-simulation approach allows data at fine resolution (e.g. a smallest spatial unit of a grid cell) to be handled and produces detailed simulation results (Torrens, 2000). Many techniques such as the statistical regression model (e.g. the work of Almeida et al. (2003) and Wu (2002a)) and/or multi-criteria decision analysis (e.g. the work of Wu and Webster (1998)) are also incorporated to enhance the capabilities of CA and GIS. The statistical technique is the empirical estimation means, which derive the criterion weights from the relationship between land use change and development factors observed from the study area. The multi-criteria decision analysis technique allows the direct incorporation of decision-makers for setting criterion weights used for simulation areas likely to be developed. The applications that have incorporated these techniques up to now have been mainly focused in the area of developed nations. In the developing nations, however, these applications have been limited to simulating some regions of China (e.g. the work of Li and Yeh (2000) and Wu (2002a)). Limitations over the adoption of GIS in developing countries has been mainly because of the high data cost, the lack of data availability and their restriction in use, the shortage of trained and skilled staff and experts, and the lack of expertise and experience (Barredo et al., 2004; Bishop et al., 2000; Herold et al., 2003).

In the case of Thailand, applications for urban analysis to date have usually been in the area of land use change detection and urban expansion monitoring. These applications have been conducted with the integration of remote sensing imageries and aerial photographs. For example, Thomson and Hardin (2000) used GIS and remote sensing as an analytical tool for the identification of low-income housing sites. Bishop et al. (2000) introduced GIS and spatial data infrastructure as a potentiality to assist in managing human settlements. Nevertheless, until now there is no evidence of the prediction of dynamic urban land use changes or trends of urban growth for the Bangkok area using the micro-simulation approach at the detail of district or neighbourhood level. The application of a micro-simulation approach such as that of CA with the integration of techniques such as regression model and multi-criteria decision analysis can thus potentially be used to simulate the spatial pattern of urban growth in contexts that are different from those experienced by cities of developed nations.

1.4 Thesis Aim and Objectives

The urban growth and land use problem in the Bangkok area described in the preceding section has highlighted the urgent requirement to monitor and predict land development. In response to this, the principal aim of this research is to develop an urban spatial model that can be used to simulate the urban growth of Ladprao, a part of the Bangkok area, Thailand. Potentially, this model can be used to map and locate the urban development at the district level using the micro-simulation approach. Thus, to achieve the key aim, the objectives of the research are:

i). To develop an urban spatial modelling framework by the integration of a GIS-based Cellular Automata spatial model, where the multinomial logistic regression (MNL) and multi-criteria decision analysis (MCDA) methods have been integrated to identify the potential cells for development.

ii). To develop a set of customized tools using VBA (Visual basic for Applications) within the ArcGIS environment.

iii). To implement the proposed model, in order to replicate urban development for Ladprao, a part of the Bangkok area, Thailand.

iv). To evaluate the performance of the proposed GIS-based cellular automata spatial model based on historical land use data. This validation will compare the results produced by the model with the actual urban areas.

v). To test the sensitivity analysis of the model using different neighbourhood sizes and different neighbourhood thresholds. This analysis will examine the claim of Caruso et al. (2005) and Kocabus and Dragicevic (2004) that variation of the cellular automata's elements have an effect on the simulated results .

These objectives will help achieve the main aim. It is to be hoped that this innovative model can assist in creating an effective spatial urban model applicable for the Bangkok area and other parts of Bangkok that have identical land use patterns of urban land transition.

1.5 Thesis Structure

This thesis is presented in eight chapters. This chapter presents the background and rationale for the research projects, a statement of the problem and potential application, and identifies concisely the principal aim and research objectives. It concludes by providing this brief outline structure of the thesis.

Chapter 2 reviews and comments on the extensive literature works in the international context related to urban spatial modelling techniques and the information requirement for urban simulation. The related application of geographic information systems and its extension in the context that assists in the selection of appropriate tools to be used in modelling of urban development for the research study are also reviewed.

Chapter 3 provides a limited historical background of the study area and Bangkok, the related land use problems in Bangkok as well as the role of planning policy in controlling land use development. It also includes data source, data pre-processing performed prior to modelling, issues in data constraints in the context of data poor environment, as well as the selection of development factors to be used in the application.

Following on from Chapter 2, Chapter 4 describes and extends the conceptual modelling framework and theoretical methodology used to developed the GIS-based Cellular Automata model with integration of multinomial logistic regression model (MNL) and the multi-criteria decision analysis (MCDA). It also exemplifies the means to derive criterion weights using statistical logistic regression as well as the pairwise comparison method.

Chapter 5 reviews the development of the practical implementation of the model under a GIS platform using a VBA macro within ArcGIS environments. It also includes the introduction of VBA and ArcObjects being used as a basic knowledge for model development. The detailed information about graphical user interfaces and the execution of the interfaces are presented.

Chapter 6 illustrates and describes the simulated results produced from the GIS-based CA/MNL and the GIS-based CA/MCDA models developed from Chapter 5. Different development factors and weights are examined, in order to identify their effects on the simulated results. Spatial comparison between the simulated results and the real data set as a way to validate and assess the accuracy is reviewed. Also, sensitivity analysis through different neighbourhood sizes and neighbourhood thresholds are also performed and discussed.

Chapter 7 discusses the performance of these simulated results as well as suggestions to improve the performance of the model. This chapter also gives comments about the integrated techniques used in the model, in terms of the technical limitation and its contribution to the field.

Chapter 8 gives a summary of the main findings of this research work as well as suggestions for future possible research directions. This chapter also includes an assessment of the achievement of this research, revisiting the main aim and considering the objectives.

CHAPTER 2

Urban Land Use Change Simulation: Techniques and Applications

2.1 Introduction

There is a growing literature suggesting that urban land use simulation can, in part, assist in understanding how urban systems are built over time and in predicting future spatial growth (Herold et al., 2001; Torrens and O'Sullivan, 2000; Wu, 1998). Such simulation is, in appropriate circumstances, considered as one key factor for estimating and assessing future land use patterns (Barredo et al., 2004; Herold et al., 2003). Further, it can be used as a tool for creating planning scenarios which help urban and regional planners to explore effects of their decisions in order to minimize existing and future impacts (Herold et al., 2003). The discussion in this chapter firstly reviews literature on urban spatial modelling. This is followed by discussion about the information requirements for urban simulation, especially in the context of data sources in developing nations (Section 2.3). Thereafter, a review of the use of GIS for modelling the urban area, especially in the context of its integration with other approaches, and its application to developing nations is discussed (Section 2.4). Thereafter, Cellular Automata (CA) and its extension, focusing on the micro-simulation technique, on urban land use change, are discussed (Section 2.5). In Section 2.6, Multi-criteria Decision Analysis (MCDA) that can be used to extend standard GIS functionalities in terms of the decision making process is discussed. In the concluding section (Section 2.7), a review of approaches outlined above is brought together to assist in the production of simulation models which will be applied in the remainder of this thesis.

2.2 Urban Spatial Models

A wide variety of urban theories and models have been developed over time to accomplish two main purposes (Fotheringham and O'Kelly, 1988; Hester, 1970). The first purpose is to describe and explain the structure of existing urban systems, and the second is to assist in crating a predicting or forecasting model (Lee, 1973). The theories and models developed for explanation purposes only, were criticized by Forrester (1961) who suggested that they are of limited use for projection and have no success in the provision of sufficient description of system behaviour. Further, they are of little interest to urban planners since they fail both to give information about the future and direction of urban growth and to help planners in providing possible alternative programs and solutions for urban planning (Lowry, 1965). However, Lee (1973) suggested that a good descriptive theory and model will frequently assist in the development of a good forecasting model. He explained that without understanding the existing urban structure, a forecasting model fails to select and set relevant reasonable variables and their relationships, needed for driving the simulation and prediction. As a result, this can lead to the production of an unrealistic simulation result.

In this section, a brief review of spatial urban modelling mentioned above is presented. Such models can be broken down into three major groups: descriptive, predictive and prescriptive (Batty, 2001; Tomlin, 1990). They are meant to describe either growth or the results of growth. While descriptive models attempt to describe and explain the results of urban growth (e.g. explaining the existing land use patterns as a result of growth), the predictive and prescriptive models attempt to explain the growth (e.g. searching for the factors and/or processes that drive the growth). Descriptive models act as a way to describe and represent the existing urban situation (Lee, 1973). Predictive and prescriptive models include more dynamic and simulation effects. The significant difference between them is that while predictive models aim to find and predict the relationship between relevant variables and attempt to simulate the future growth, the prescriptive models are constructed with a defined set of objectives and constraints in order to answer question 'what is likely to happen as a result of certain assumptions' or 'what ought to be' (Chapin and Kaiser, 1979; Lee, 1973).

2.2.1 Descriptive Spatial Models

Early attempts to explain spatial urban structure include Hurd's central and axial urban growth concept, the concentric-zone concept, the sector concept, and the multi-nuclei concept. These descriptive models were created before the 1950s and mostly presented as architectural presentations of cities and urban environments in a physical form (Batty, 2001). These models were built mainly to describe the spatial structure of existing cities and urban areas in the western developed countries.

According to the concept of Hurd in 1930, the growth of urban areas can be represented as a pattern of central and axial growth in combination (Anderson and Egeland, 1961). Such growth can be illustrated as a roughly star-shaped pattern which is a reflection of the fact that urban development tends to take place outward from the city centre in all directions, most rapidly along the transportation routes. Despite his discussion, Hurd did not create the overall, concise descriptive generalizations to support his model's growth pattern (op. cit.).

The Burgess concentric model, developed by Burgess in 1925, was the first endeavour to investigate spatial patterns at the urban level (Rodrigue et al., 2006). According to his model, the importance of central (or concentric) growth was emphasized and exemplified by the distribution of residential areas by type (Anderson and Egeland, 1961). In his model, a central business district (CBD) is at the centre and is succeeded outwards by zones of factories, transition, working class, housing, general residential district and a commuter zone together forming a series of six concentric circles (Rodrigue et al., 2006) as shown in Figure 2.1. Each circle represents a specific socio-economic urban landscape with the rate of progression based mainly on economic and population growth within a mono-centric city (Chapin and Kaiser, 1979). Despite its simplicity, the model remains useful and is used as a broad and general concept to describe the concentric urban development of the American cities in the early-mid 20th century (Rodrigue et al., 2006).

The sector concept was developed by Homer Hoyt in 1939 as a study of residential areas in the North American context (Chapin and Kaiser, 1979). According to this model, the importance of axial patterns of growth was emphasized (Anderson and Egeland, 1961), generating wedge-shaped sectors (Briassoulis, 2000), coupled with the influential Burgess concentric model (Figure 2.2, left).

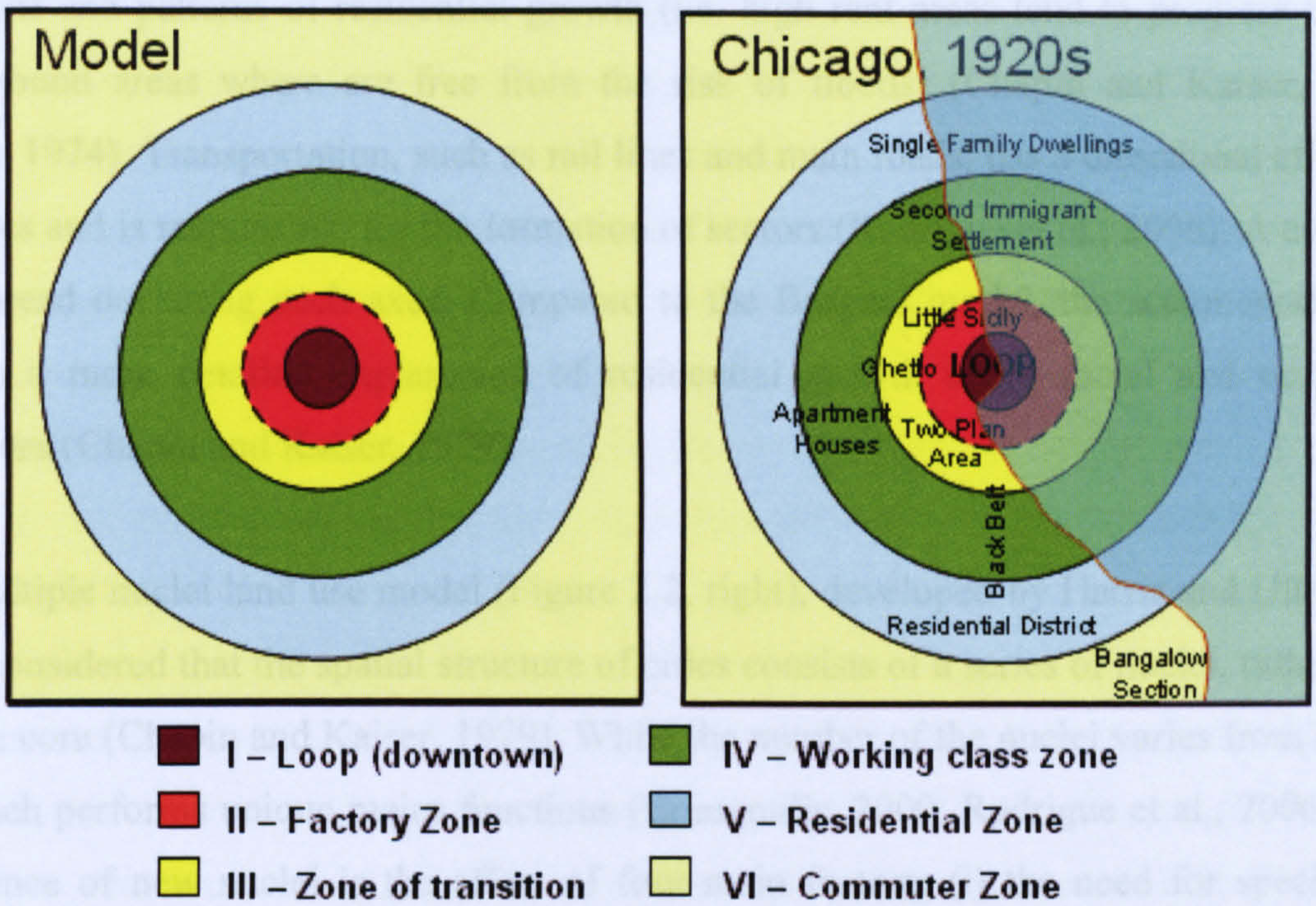


Figure 2.1: Burgess concentric model (adapted from Rodrigue et al. (2006)).

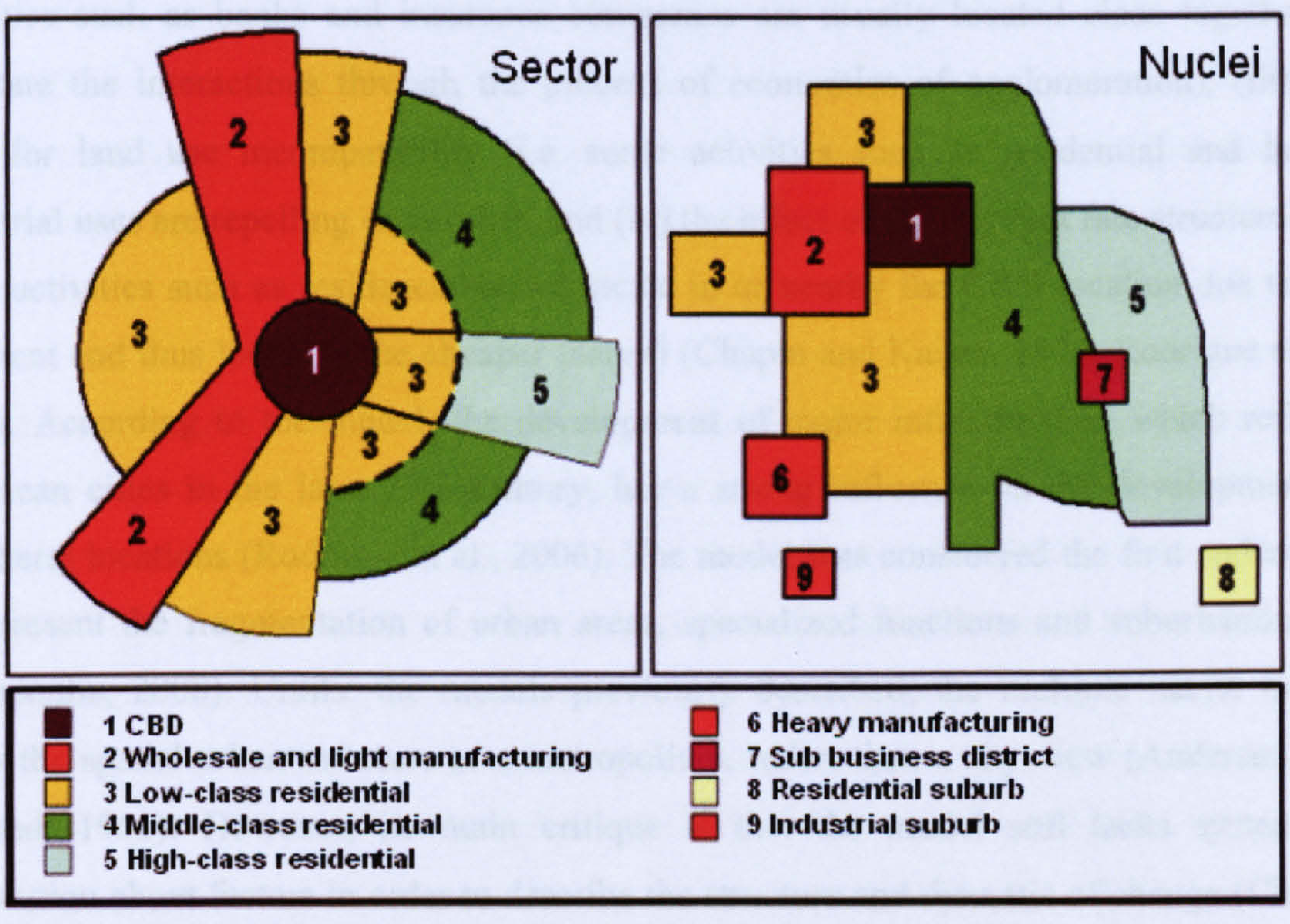


Figure 2.2: The sector concept (left) and multiple nuclei land use model (right) (adapted from H. Carter (1995)).

These sectors describe residential and commercial patterns that emerge on the basis of the income of residents spreading out from the core centre along transportation routes within a mono-centric city (Briassoulis, 2000). High-rent areas (and high-price areas) influence the directions and patterns of residential growth (i.e. high rent areas tend to progress toward high ground areas where are free from the risk of floods) (Chapin and Kaiser, 1979; Wilson, 1974). Transportation, such as rail lines and main roads, has a directional effect on land uses and is responsible for the formation of sectors (Rodrigue et al., 2006). A city will then spread out along such axes. Compared to the Burgess model, the sector model can provide a more detailed explanation of residential growth using social and economic influences (Chapin and Kaiser, 1979).

The multiple nuclei land use model (Figure 2.2, right), developed by Harris and Ullman in 1945, considered that the spatial structure of cities consists of a series of nuclei, rather than a single core (Chapin and Kaiser, 1979). While the number of the nuclei varies from city to city, each performs unique major functions (Briassoulis, 2000; Rodrigue et al., 2006). The emergence of new nuclei is the effect of four main factors; (i) the need for specialized facilities by certain activities (i.e. port terminals require specialized facilities and do not need to located close to the CBD), (ii) the need for land use compatibility (i.e. similar activities such as banks and insurance companies are usually located close together to facilitate the interactions through the process of economies of agglomeration), (iii) the need for land use incompatibility (i.e. some activities such as residential and heavy industrial uses are repelling each-other, and (iv) the effect of the city rent rate structure (i.e. some activities such as residence cannot locate in or nearby the CBD location due to the high rent and thus locate at the cheaper places) (Chapin and Kaiser, 1979; Rodrigue et al., 2006). According to the model, the development of major infrastructure, which reflects American cities in the later 20th century, has a strong influence on the development at peripheral locations (Rodrigue et al., 2006). The model was considered the first endeavour to represent the fragmentation of urban areas, specialized functions and suburbanization (Briassoulis, 2000). Unlike the models previously described, the multiple nuclei model views the spatial urban structure as a metropolitan, rather than a city view (Anderson and Egeland, 1961). However, its main critique is that the model still lacks systematic explanation about factors in order to describe the structure and dynamic of change (Chapin and Kaiser, 1979).

The Burgess concentric zone and the sector model have been tested in many American cities. For example, the work of Anderson and Egeland (1961) applied the concepts to represent the urban structure of four cities in the United States (Akron and Dayton in Ohio, Indianapolis in Indiana, and Syracuse in New York) and produced promising results as they can be used to describe the urban structure of those cities. However, all models and theories illustrated above are static and use descriptive means, primarily focusing on residential land uses (Chapin and Kaiser, 1979; Rodrigue et al., 2006). Criticized by Briassoulis (2000), all of the models bear similar criticism as none can explain the processes or mechanisms of change in the land use patterns. They do not include factors that describe the growth and decline of economic activities, the dominance of certain activities, the changes in preferences, and other constraints (e.g. institutional) on land development and use.

A further limitation was that these three descriptive models have been applied to describe cities only in the developed countries, especially in the United States. Despite the fact that the urban structures of colonial cities in the developing nations have been largely influenced by colonial expansion in a particular region, their land use patterns are varied and different from those of developed countries (Drakakis-Smith, 2000; Potter, 1992). And in some cases, individual cities tend to exhibit their own distinct land use patterns (Potter, 1992). This is as a result of the multifaceted mix of scale of colonial presence, chronological phase of urbanization, nature of the native culture, technology advancement, socio-economic aspect, and the integration of planning practice involved (Drakakis-Smith, 2000).

According to Potter (1992), the urban structure of many cities in the developing countries before the beginning of the industrialization period, especially colonial cities of Africa and Asia, was first described by the descriptive models of pre-industrial cities proposed by Sjoberg (1960). In such pre-industrial model, the wealthy and elite group lives close to the urban core (mainly for walking access), followed outwards by lower classes and outcastes, thus forming a series of three concentric circles. Such a pattern is opposite to today's urban structure in modern western cities (Potter, 1992). However, this pre-industrial pattern recently has become less typical because of rapid growth of cities, particularly large Latin American cities such as Mexico City and Bogotá. Griffin and Ford (1980), after observation of some fast growing cities of Latin America, proposed 'a Model of Latin American City Structure'. According to the model (Figure 2.3), the CBD is situated at the urban core. Extending out from the CBD is the development of a commercial spine. Next to the spine in the form of a wedge is an elite residential sector, containing upper-class and upper-middle-class housing stocks. Outside the spine/sector, there is the structure of three distinctive residential concentric zones, including a zone of maturity (better residence), a zone of in situ accretion (modest residence) and a zone of peripheral squatter settlements. These three zones represent a residential pattern that emerges on the basis of socio-economic characteristics. However, such characteristics are not too different from those of American cities (Griffin and Ford, 1980). Unlike American cities where a full range of urban services are provided before the land use development, most developing countries have limited extension of urban services. This results in the expansion of the spine and a wedge of the elite group, representing a dominant morphological feature of the typical Latin American city (op. cit.). Until now, this model has been proven to fit well with the urban structure of most developing nations that have experienced rapidly growing cities (Potter, 1992).

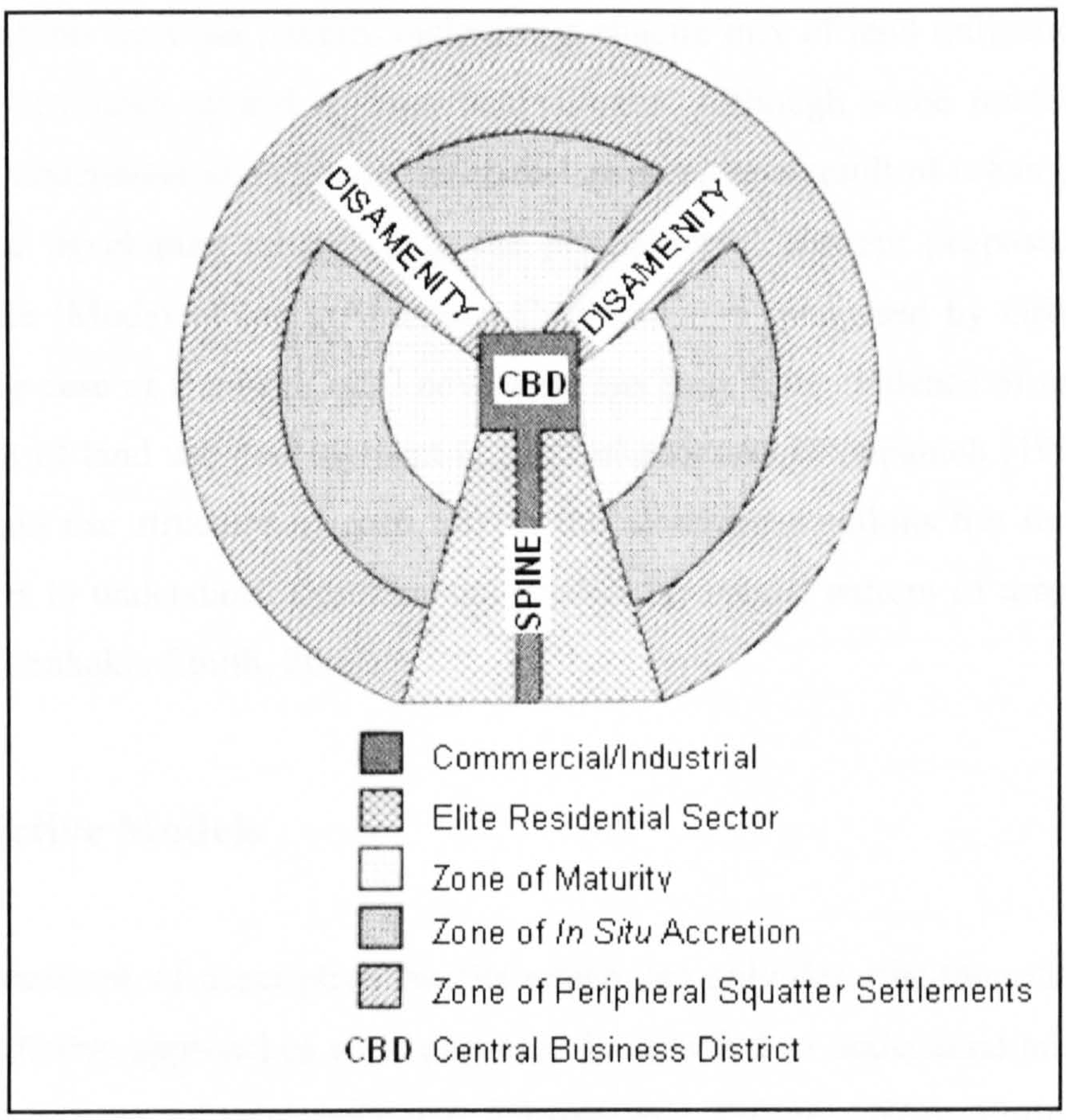


Figure 2.3: A generalized model of Latin American city structure (Adapted from Griffin and Ford (1980)).

Different to those cities of colonial developing nations, Bangkok's land use pattern has its own distinct pattern. Since 1767 the land use pattern of Bangkok's growth is described as a hybrid between multiple nuclei growth theory and Hurd's axial growth (Sharkawy and

Chotipanich, 1998). Chomchan et al. (1990) observed that the building of bridges across the Chao Phraya River and the establishment of new roads are the major factors of Bangkok's development. However, unlike cities in Latin American cities which are influenced by Spanish grid-street pattern in regular east-west and north-south directions (Griffin and Ford, 1980), Bangkok has suffered from a scantiness of secondary and distribution roads to link the axial spines with Bangkok's multiple nuclei of development (Sharkawy and Chotipanich, 1998). It results in confused road layouts and inaccessible 'blind land' in the interior (Ross et al., 2000; Sharkawy and Chotipanich, 1998; Webster, 2000). Its impact, in part, has exacerbated a misuse of land and the expansion of the city by pushing residential development farther out of the city into the suburbs (Webster, 2004).

Compared to the urban spatial pattern described by the three descriptive western models discussed above, the spatial pattern of cities in developing countries, including Bangkok, has a less uniform land use pattern, including a chaotic mix of land utilization in the city and the uncontrollable sprawl to suburban villages. Although some models have been developed to understand and explain the spatial patterns as a result of urban growth of the cities in these developing nations (e.g. the pre-industrial concept proposed by Sjoberg (1960) and the 'Model of Latin American City Structure' proposed by Griffin and Ford (1980)), in the case of Bangkok until now there has been little evidence of monitoring the spatial pattern of land use development (e.g. Sharkawy and Chotipanich (1998)). Because the spatial land use structure of each city in the developing nations has its own distinct pattern, efforts to understand, describe and predict the spatial pattern of urban growth are challenging (Drakakis-Smith, 2000).

2.2.2 Predictive Models

Due to the drawback of descriptive models which can only describe the existing land use situation, predictive approaches were proposed in an effort to understand and forecast the future spatial pattern of urban systems (Chapin and Kaiser, 1979). Various predictive means have been developed and used extensively for both planning and research objectives over the past several decades (Waddell and Ulfarsson, 2004). Here, those predictive models are classified in three categories, according to their model complexity and their operational use. Example models of the first category include two classical theories of land use modelling, the von Thunen's Agricultural Land Rent Theory, the urban land market theory of Alonso. The second category is referred to as spatial interaction modelling, being developed as the earliest operational use model and later extensively modified and applied in a number of urban application (e.g. Wilson (2002)). The third category includes the models and approaches that emerged from recent concepts of micro-simulation and dynamic modelling: the cellular automata approach (CA), the SLEUTH model, and the UrbanSim model.

2.2.2.1 Classical Theories of Land Use Modelling

One of the oldest example land use model is von Thunen's Agricultural Land Rent Theory, developed in 1826 for the analysis of agricultural land use patterns in Germany (Chapin and Kaiser, 1979). The model is considered as a micro-economic theoretical means for the analysis of land use patterns and their changes (Briassoulis, 2000). Summarized by Briassoulis (2000), von Thunen's assumption was that land was uniform, arranged on an isotropic (of equal fertility) flat plain, allowing equally possible movement in all directions around a market place (central city). Land rent, thus, differs only with distance from the centre. The theory, to explain the productivity (rent) of farmers, was expressed in a simple mathematical equation of the market price minus the transport and production costs.

Despite its model simplicity, von Thunen's theory was recently applied in the work of Stutz and Souza (1998) for the creation of an agricultural land use map over the continental United States. The agricultural land use map produced has a high level of concordance with reality. The basic principles of von Thunen's theory have been the origin of many other concepts, mainly land rent and distance-decay (so-called friction of distance) (Rodrigue et al., 2006). In addition, the model is considered the predecessor of both location theory[1] and the analysis of urban and regional spatial structure (Briassoulis, 2000). However, its main criticism is that since the model is considered a static description and explanation of land use (Briassoulis, 2000; Chapin and Kaiser, 1979; Rodrigue et al., 2006) and it is built on the basis of very restrictive and unrealistic assumptions, it gives no explicit reference to process land use change (Briassoulis, 2000; Rodrigue et al., 2006). Briassoulis (2000), nevertheless, claimed that an implicit land use change mechanism even

[1] Location theory is referred to as "a body of theories which seeks to describe, explain, and prescribe the location of economic activities in space" Briassoulis (2000).

under all the restrictive assumptions of the von Thunen's theory can be noticed. He explained that a possible land use change could be seen, if the relative prices of the products change exogenously and thus this will change the relative ability of the users of land to bid for specific locations. A relaxation of the original formulation of the von Thunen's model was carried out and modified in several subsequent applications (Briassoulis, 2000). One of these was the foundation for Alonso's (1964) urban land market theory.

The urban land market theory of Alonso in 1964, built under a mono-centric conception of the city, focuses on the problem of predicting residential locations as a function of transportation and housing costs (Waddell, 2000a). This model extends the von Thunen's concept by considering land use, rent, intensity of land use, population and employment as functions of distance to the CBD of the city (Watkins, 2006). In this model, a bid-rent function, describing the variation in land rent payable at different locations within a CBD, is applied. By overlaying the bid rent curves with a concentric land use pattern, the model reflects possible trade-off between space consumption and transportation (see Figure 2.4). As a result, land use activities are located at different distances from the CBD (i.e. the closest distance from the CBD is dominated by retail activities, and the farthest distance from the CBD is dominated by residential use). This model thus suggests that distance-decay concept and transportation through accessibility to the CBD are crucial explanatory factors of the land rent and impacts on land use (Rodrigue et al., 2006).

However, the model has some criticisms. Firstly, similar to von Thunen's theory, the Alonso model is built upon a static explanation of urban land use, giving no explicit process of land use change. Secondly, there are other influential factors that should be taken into account in the bid-rent curves such as socio-cultural (race, crime, perception, etc.), physiographic (waterfront, hills, etc.), historical (tourism) and political attributes (Briassoulis, 2000; Rodrigue et al., 2006). Despite those criticisms, Alonso's theory has been applied and modified extensively and has an influence in the analysis of urban spatial structure (Briassoulis, 2000). For example, the work of Shieh (1987) cited in Rodrigue et al. (2006)) modified the model for the presence of more than one centre. Rodrigue et al. (2006) illustrate some modifications of land rent theory to describe contemporary cities as shown in Figure 2.5. The rapid expansion of a metropolitan area, the emergence of sub-centres and huge improvements in transportation and telecommunications, result in many land use activities being pushed far from the CBD and appearing as a second peak (Figure 2.5) due to a reflection of concentration of retailing, commercial, distribution and industrial activities in the sub-centres. The urban land use pattern, as a result, tends to be far less coherent, more particular and dispersed (Rodrigue et al., 2006).

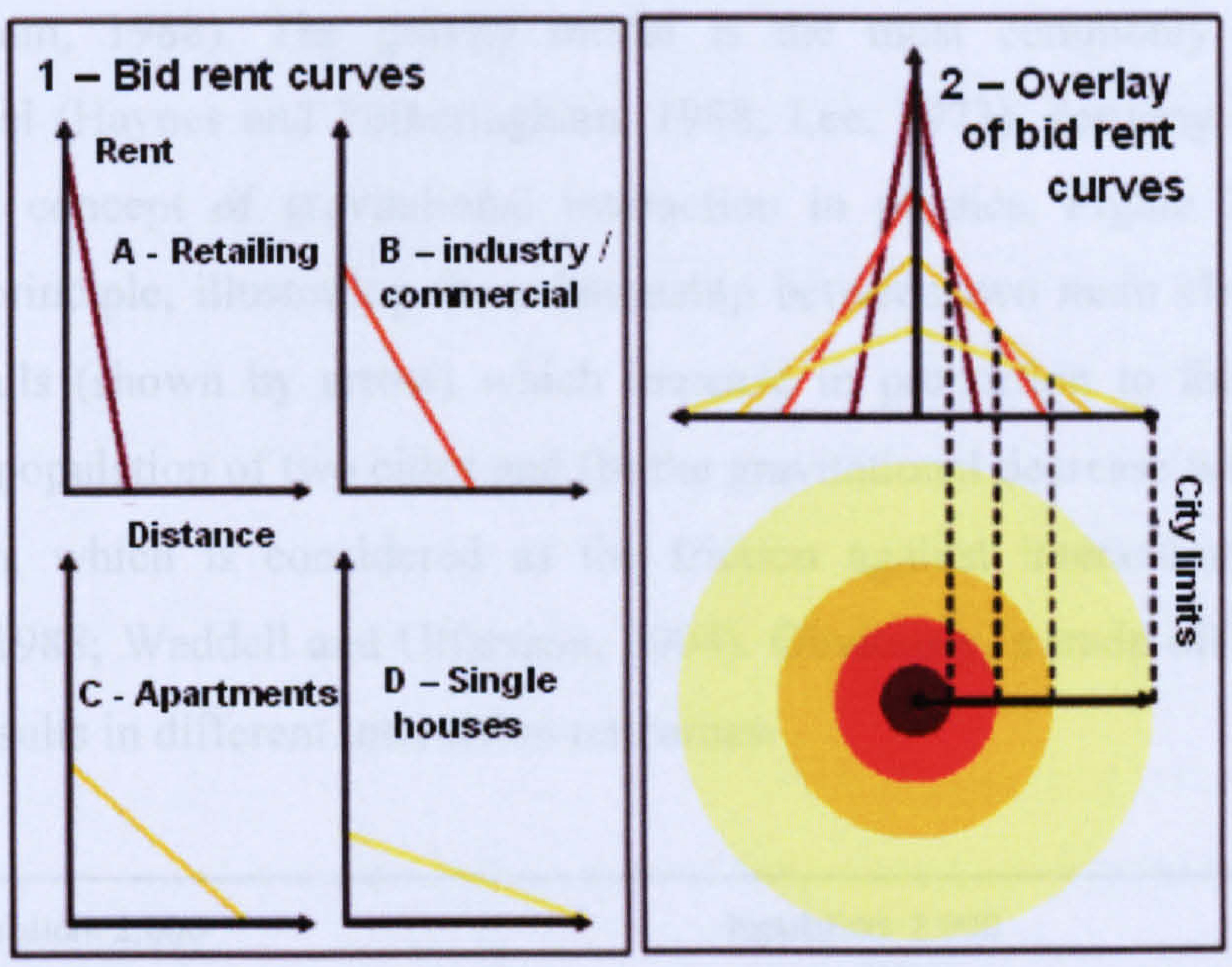


Figure 2.4: Example of bid-rent theory applied for residential location (left) and the optimal location of major urban economic activities and (right) the overlay of the bid rent curves of all the urban economic activities with a concentric land use pattern, considered as an isotropic space (adapted from Rodrigue et al. (2006)).

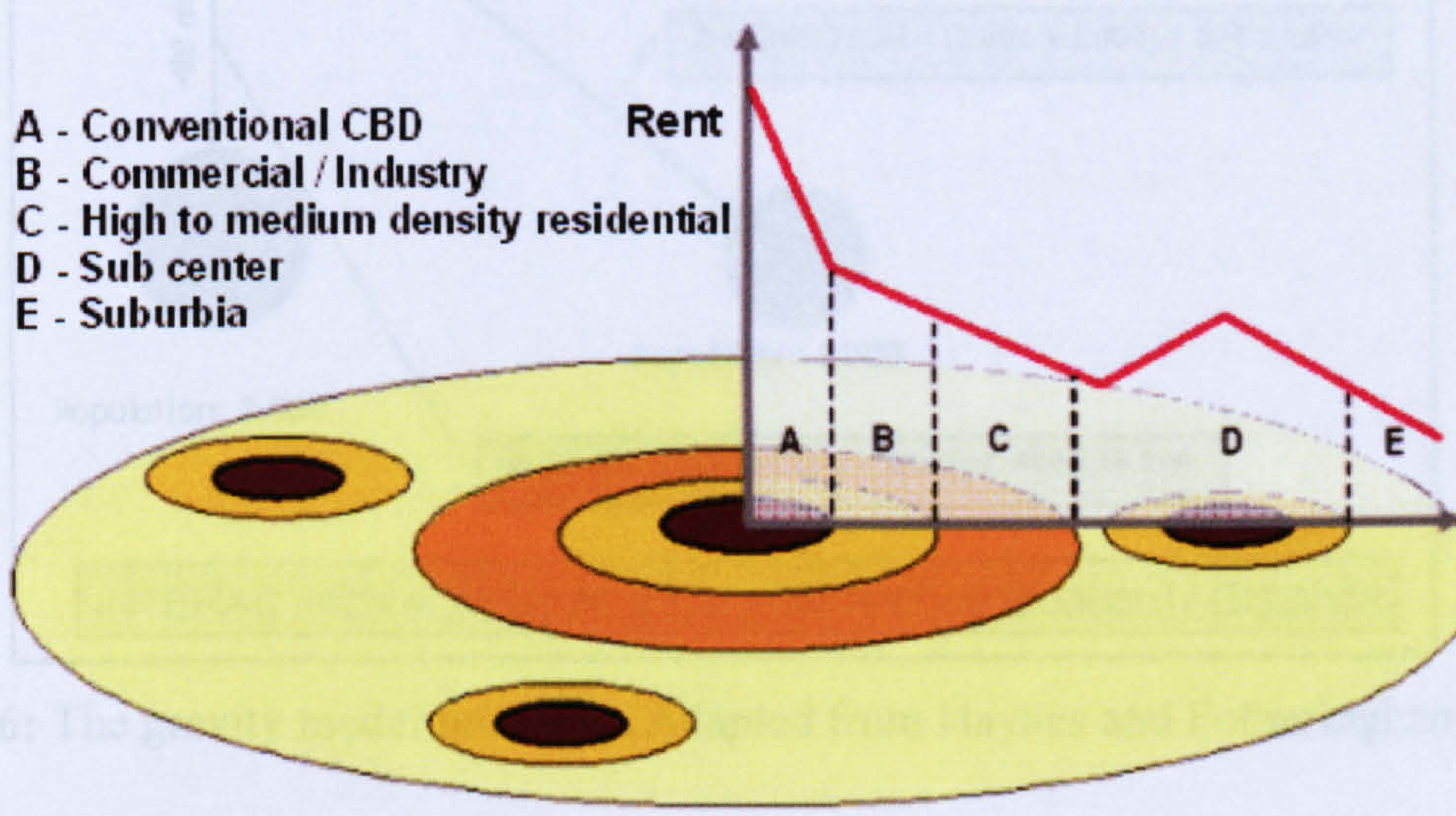


Figure 2.5: Contemporary modification of the land rent theory (adapted from Rodrigue et al. (2006)).

2.2.2.2 Spatial Interaction Model

Broadly speaking, spatial interaction refers to any movement over space resulting from humans' decision processes, and includes a wide variety of applications such as journey-to-workplace, migration, and commodity flows (Fotheringham and O'Kelly, 1988; Haynes and Fotheringham, 1988). The gravity model is the most commonly used type of interaction model (Haynes and Fotheringham, 1988; Lee, 1973), deriving its name from Isaac Newton's concept of gravitational interaction in physics. Figure 2.6 shows the gravity model principle, illustrating the relationship between two main elements: (a) the gravitational pulls (shown by arrow) which increase in proportion to the mass of two objects such as population of two cities and (b) the gravitational decrease with the distance separating them, which is considered as the friction against interaction (Haynes and Fotheringham, 1988; Waddell and Ulfarsson, 2004). Obviously, a trade-off between these two elements results in different interaction outcomes.

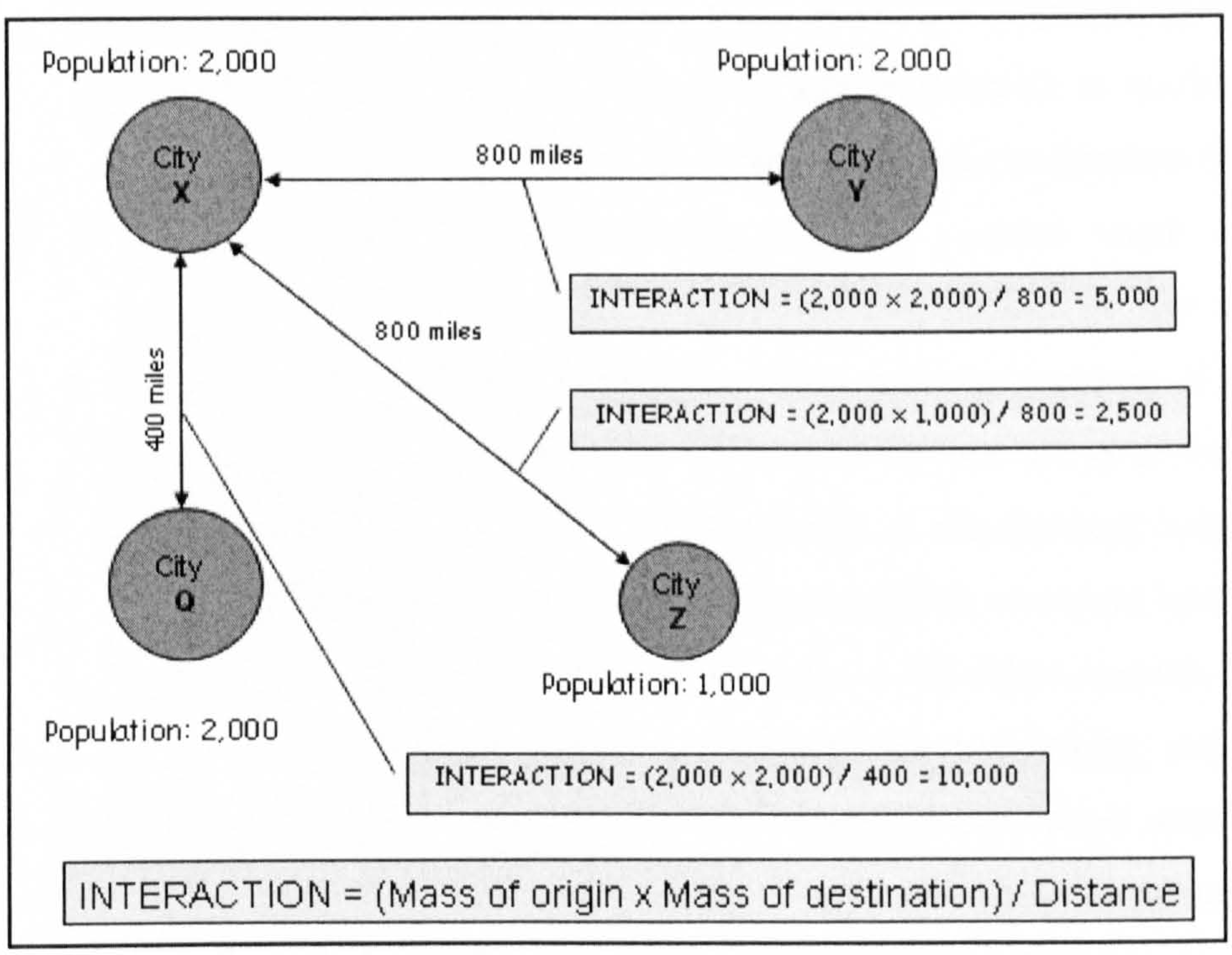


Figure 2.6: The gravity model principle (Adapted from Haynes and Fotheringham (1988)).

Based on the concept, the model was expressed in a simply modified mathematical version by Lee (1973) as:

$$I_{ij} = G\frac{P_i P_j}{d_{ij}^b} \qquad (1)$$

where I_{ij} = the interaction between area *i* and *j*,

P_i, P_j = the magnitude of area *i* and *j* (e.g. population, wealth, physical size),

d_j = the distance between area *i* and *j*,

b = a power or exponent applied to the distance between the areas,

G = a constant, equivalent to the gravitational constant, which is empirically determined, and is used to refer the relationship to actual conditions.

Summarized by Briassoulis (2000), the gravity model is fundamentally built for the purpose of computing the flows or interactions between origin and destination zones and for predicting these flows when changes occur in the origins and/or destinations and/or when changes occur in the accessibility between origins and destinations, mostly as a result of the improvement of the transportation network. Various forms and the underlying concepts of the gravity model have been applied and extended widely in many fields. In retail and market analysis, for example, Reilly's Law of Retail Gravitation developed in 1931 applying the two gravity principles of scale (such as mass of population) and distance for the purpose of calculating and drawing the market area boundaries or market shares by focusing on the competitive effects of alternative markets under consideration (Haynes and Fotheringham, 1988). Reilly's concept is considered a pioneer work for further development in market analysis (op. cit.). In another application in the urban field (employment), the Lowry model was developed in 1964 with the purpose of forecasting both the location and the distribution of the total population and service employment, given the trips generated from the flows of residential locations to employment locations (Lee, 1973). The model was promising as early development work for an urban land use model (op. cit.). The model consists of three major components of the urban system; population, employment and the means of communication, representing the linking and feedbacks between residential (so-called household sector) and retail services (retail sector) (Haynes and Fotheringham, 1988; Lee, 1973). Although the Lowry model had some modifications and improvements latter, its underlying concept has remained one of the most widely used in urban development (op. cit.).

The gravity model and its family share similar criticisms. The first and most fundamental criticism is that the model is not designed on the basis of a theory of urban system behaviour (Briassoulis, 2000; Lee, 1973). Since the model is built on a simple structure and function of interaction between two zones and a friction of distance, a few parameters are involved (Lee, 1973). The model fails to make sufficient representation of behavioural processes (Robinson, 1998) and thus it cannot represent the complex situation of urban activity patterns (Lee, 1973). Similar to the Alonso model, the gravity model cannot account for many behavioural and socio-economic factors mainly associated with the true characteristics of land use (e.g. factors influencing location choices or land prices) (Briassoulis, 2000; Waddell and Ulfarsson, 2004). A second criticism is that the model is designed to account for the observance of large groups of objects or aggregation level (e.g. administrative boundaries) (Briassoulis, 2000; Lee, 1973; Robinson, 1998), thus, it has a limitation in terms of the degree of spatial detail used (Briassoulis, 2000; Waddell, 1998). The third criticism involves the zoning problem (Briassoulis, 2000; Lee, 1973). Spatial interaction implies that the area under investigation has been divided into zones, but in a real world situation zoning systems are often applied in indiscernible shapes or with fuzzy boundaries (Lee, 1973). Further, another consideration about the suitable number and shape of zones and the effects of the zoning system remain unanswered (Briassoulis, 2000). For example, different sizes of zones can result in different model outcomes (Briassoulis, 2000). A fourth criticism is that a spatial interaction model is built on the basis of static or comparatively static description which means that it cannot take the dynamics which underlie the observed interactions into account (Briassoulis, 2000; Haynes and Fotheringham, 1988; Lee, 1973).

There is an attempt to extend the spatial interaction approach and its family to be used as a realistic forecasting model such as the extended spatial interaction and spatial choice presented in Fotheringham and O'Kelly (1988), the application of Clarke et al. (1998) that attempted to simulate dynamically structural change of retail and school development in Leeds, and the research work of Wilson (2002) that introduced the concept of dynamic spatial interaction for creating the dynamics of cities. However, one of main drawbacks is that a very limited set of characteristics of land use types can be input to the analysis, usually population of the residential areas, or the income of the population, or their floor-space factors (Briassoulis, 2000).

2.2.2.3 Cellular Automata (CA) and Its Extension

Recently, the emergence of cities is considered as the result of complexity and urban dynamics (Batty, 2003). Cities are regarded as organized complex spatial systems as they are made up of strong interdependent parts of interaction among human agents, such as individuals, households, public or private sectors, or organizations (Wilson, 2000; 2002). Further, such growth occurs as the process of change over time, so-called 'urban dynamics' (Wilson, 2000). These advances have changed urban modelling approaches from an aggregate, static, cross-sectional and spatial equilibrium aspect to a simulation pursuit that seeks to understand how cities grow and evolve in an urban system (Batty, 2001; Guhathakurta, 2001). Substantial research has been carried out in an attempt to understand the complexity and urban dynamics through the procedures of urban modelling and simulation, especially those related to self-organizing systems, neural networks and other nonlinear dynamic systems (Guhathakurta, 2001). Amongst all the research documented as dynamic models, those based on cellular automata (CA) are likely the most remarkable approach (Benenson and Torrens, 2004; Yang and Lo, 2003). This is because CA can incorporate spatial components (mostly cell-based) (Yang and Lo, 2003), include dynamism with simple rules (Torrens, 2000) and apply the bottom-up process (which allows macro-scale to emerge in the ordered pattern from the dynamics of micro-scale) (Torrens, 2000). Details about cellular automata (CA) and its extension will be given intensively in Section 2.5.

The predictive models discussed above are of a wide variety, ranging from the traditional (e.g. the urban land market theory of Alonso) to the dynamic simulation model (e.g. the CA approach). They are based on various concepts and rationales over space and time in order to explain and find the reason "why" the growth occurs. However, these predictive models share at least one similarity. They are capable of simulating spatial urban growth in order to either replicate the real world situation or forecast the future urban pattern, or both in combination.

2.2.3. Prescriptive or Normative Model

Prescriptive, also known as normative, models are considered the extensions to predictive models built to aid in supportive decision-making process in planning policies (Lee, 1973).

In planning decisions, generally, one of the main problems is as a result of the combination between a wide variety of view points (so-called multiple objectives) and a number of decision-making criteria of multiple stakeholders and planners in various fields (Feick and Hall, 2004). It thus, to some extent, can lead to multi-faceted, ill-defined and sometimes conflicting criteria among multiple interests (Feick and Hall, 2004; Makropoulos and Butler, 2006). In land use planning, for example, applications can be from many areas of interest such as the location of 'noxious' land use (e.g. waste disposal area (Sener et al., 2006)), land resource sustainable management (Banai, 2005; Wu, 1998), the future form of communities (e.g. urban transportation policies (Arampatzis et al., 2004), and designation of urban green areas (Villa et al., 1996)). Prescriptive models are those developed to help decision makers investigate and define the main objectives and alternatives, perhaps compromising them from various decision-making scenarios being generated on the basis of different policies and areas of interest (Lee, 1973; Malczewski, 1999a). Such models can be achieved by applying the certain well-defined objectives and constraints to the design of the system (Longley et al., 2005) with an attempt to predict what might happen as a result of a particular set of circumstances (Lee, 1973).

One of the well-known prescriptive approaches in urban planning applications is referred to as the optimization method (Openshaw, 1978), based on mathematical modelling, which is built to search for the best minimum or maximum solution for a given particular decision problem (Longley et al., 2005; Malczewski, 1999a). In the most general terms, an optimization model can be expressed as Malczewski (1999a):

$$\text{Minimize or maximize } f(x) \qquad (2.1)$$

$$\text{Subject to } x \in X \qquad (2.2)$$

where $f(x)$ is a criterion or objective function which is to be optimized, x is a set of decision variables which are elements of the objective function, and X is a set of feasible alternatives which is defined from a set of constraints imposed on the decision variables.

The fundamental optimization approach is referred to as linear programming (Lee, 1973), previously expressed as equation (2.1) and (2.2). The linear approach is designed to capture reality through one objective function or a unique criterion, which is commonly a function of cost or distance (Nijkamp et al., 1990 cited in Laaribi et al., 1996). For example, in the case of the travelling salesman problem, its linear function is built to minimize the distance travelled in order to obtain an optimal route solution (Laaribi et al., 1996). Such an objective function and its constraints in the model need to be linear and additive (Malczewski, 1999a).

The conventional concept of the optimization approach, such as that of linear programming, relies on only one objective function. However, most often, the real-world problems being confronted are on the basis of conflicting objectives and multiple criteria (Laaribi et al., 1996). It is in such a context that multi-criteria decision problem has been adopted for solving multiple conflicting objectives (Laaribi et al., 1996; Malczewski, 1999a). Multi-criteria analysis allows more than one criterion function, coupled with the broader sets of alternatives being considered all together (Guitouni and Martel, 1998; Janssen and Rietveld, 1990). The approach is commonly regarded as multi-criteria decision analysis (MCDA) (Guitouni and Martel, 1998; Malczewski, 1999a; 1999b). The MCDA approach analyzes such conflicting objectives in order to provide one or more satisfactory solutions that represent true compromises in which optimal solutions are not necessary (Nijkamp et al., 1990 cited in Laaribi et al., 1996; Janssen and Rietveld, 1990). Further discussion on MCDA is presented in Section 2.6.

Up to now, a number of modifications of the standard linear programming models have been developed extensively (Malczewski, 1999a) such as integer programming, 0-1 integer programming, goal programming, and dynamic programming. Integer programming is similar to standard linear programming, with the addition of constraints requiring that all decision variables are integer (op. cit.). 0-1 (or binary) integer programming, the extension of integer programming, is commonly applied to spatial decision analysis since binary variables are set to two-choice decisions (e.g. to locate or not to locate an activity to a parcel). An example is the work of Kao and Lin (1996) who applied integer programming for landfill siting decision-making. Goal programming extends standard linear programming, having ability to optimize over more than one objective (Zeleny, 1982). An example is the work of Grabaum and Meyer (1998) who applied goal programming for creation of goal-oriented "optimal landscape pattern" areas. Dynamic programming allows dynamism to be incorporated with optimization, allowing the function of previous stages to affect the output of future stages and iterations (Malczewski, 1999a). With the dynamic programming approach, the optimization problem is broken down into a sequence of stages and a smaller optimization sub-problem can be solved at each stage (op. cit.).

Despite the fact that prescriptive models, both optimization and MCDA concepts, are considered very practical as planning aids for decision-making processes (Lee, 1973; Malczewski, 1999a), the models have some criticisms. One of the main criticisms is that since the criterion weights of the model created rely mostly on the use of expert opinions and judgments. This is an unsystematic method, but is considered necessary for the model (Lee, 1973). Such unsystematic problems can be exacerbated when the decision problems involve many goals, ill-defined definitions and objectives, and intangible or non-quantifiable variables, being incorporated (Lee, 1973; Malczewski, 1999a; 1999b) which can result in the difficulty to substantiate the real-world situations (Malczewski, 2006). A further limitation is that the models are solely designed without explicitly considering spatial context (Carver, 1991; Villa et al., 1996) which make them impractical for urban applications which can involve spatial considerations (Carver, 1991). Thus, a requirement to incorporate the models with spatial handling tools such as GIS is an alternative development (Carver, 1991; Chakhar, 2003; Malczewski, 2006).

An attempt to integrate the optimization and MCDA approaches with GIS has been used in many urban applications. For example, the research work carried out by Gomes and Lins (2002) used a multi-objective linear programming approach as the multi-criteria method on the basis of the quality of urban life. Its main purpose is for the selection of the best municipal district of Rio de Janeiro State, Brazil for urban life quality. In their work, the five objectives are infrastructure, education, security, health, and work. Each objective contained many variables which functioned to maximize or minimize it. GIS were integrated in the application using a loose-coupling strategy for the purpose of data preparation, designating weights to the criteria and visualization, while the optimization process was carried out separately.

Another example is the research work of Grabaum and Meyer (1998). They applied the optimization model for the calculation of optimal land-use patterns among multiple goals using the goal programming method. In their research work, four goals based on a landscape ecology perspective were determined as functional assessments, including the groundwater regeneration function, the water discharge function, the soil erosion hazards by water flow function, and the agricultural production function. Restrictions or constraints, derived from landscape analysis for each single goal, were set. For each goal, maximum value as a means of maximization was calculated and restrictions were imposed. Finally, the optimal solution on the basis of compromising among all goals was established, with the integration of GIS, for the creation of a goal-oriented "optimal landscape pattern" map.

Recently such integration has been extended for the application of land use simulation. For such perspectives, prescriptive models, based on a set of criterion weights determined by decision makers, play roles for creation of land suitability or potentiality maps (e.g. the work of Wu and Webster (1998)). The map then will be used as input to simulate the spatial pattern of urban growth. More details and discussion about MCDA and GIS integration and applications will be described later in Section 2.4.2.1.

The prescriptive models discussed above are built with a defined set of objectives and constraints in order to answer 'what if' scenarios. In the context of urban planning, such a concept is very helpful as the models can be used to test different scenarios on the basis of different policies and areas of interest (e.g. scenarios with and without zoning). Similar to the predictive models, the prescriptive means are used as predicting or forecasting tools to simulate the patterns of urban growth. From such a perspective, they totally differ from the descriptive models. While predictive and prescriptive models attempt to explain 'what' and 'why' the growth, search for the relationship between land use activities and growth factors within the urban structure, and test these factors and their relationships through simulation, descriptive models attempt to describe and explain the structure of existing urban systems, which is considered the result of growth.

2.3 Information Requirements for Urban Simulation

Recently, models of urban growth simulation have been increasingly applied and are of particular interest as innovative tools to city planners, economists and resource managers, as these models can be used to support intelligent decisions effectively for purpose of urban planning (Herold et al., 2001). One of the main reasons behind their fast development is due to the accrued richness of information and resources in terms of multiple spatial datasets and tools for processing (e.g. the enhancement of GIS, remote sensing and aerial photogrammetry techniques) (Herold et al., 2001; Herold et al., 2003). Data and information for such simulation plays major roles in their initial input for analysis, parameterization, calibration, and model validation (Clarke et al., 1997; Herold et al., 2003). However, the application and performance of the modelling results is still limited by the availability, quality and scope of data required for model implementation and for model validation (Herold et al., 2003; Heywood et al., 2002; Longley and Mesev, 2000). In the context of data quality, there are issues concerning accurate and up-to-date data, both spatial and aspatial, required in order to increase the realistic simulation (Barredo et al., 2004; Walford, 2001).

Data and information for model simulation, despite a wide variety in simulation approaches due to differing background theoretical concepts, are recorded and collected from a range of diverse sources. These include factual (hard information) such as figures, quantitative estimates, or systematic surveys (e.g. census and household data, aerial photographs, satellite images, Global Positioning System (GPS) surveys), and particular (thematic) maps, and speculative (soft information) such as data derived from opinions and preferences including ad-hoc surveys and questionnaires (Walford, 2001). Amongst these data sources, spatial data derived from remote sensing techniques and aerial photographs has shown popularity for the analysis and modelling of urban growth and land use change (Herold et al., 2001) in many research and applications (e.g. Clarke et al. (1997), Silva and Clarke (2005), Yang and Lo (2003)). This is because they can provide historical time series data sets that cover large areas (Herold et al., 2001; Herold et al., 2003; Walford, 2001).

In the context of urban simulation, five groups of factors which influence growth of cities and urban development can be classified (Barredo et al., 2003). They are (1) environmental characteristics, (2) local-scale neighbourhood characteristics, (3) spatial characteristics of the cities (e.g. accessibility), (4) urban and regional planning policies, and (5) factors related to individual preferences, level of economic development, socio-economic and political systems. They, all in combination, are a result of human activity, acting simultaneously in time over the urban space (Barredo et al., 2003; Li et al., 2003; Wilson, 2000). As a result, these factors affect changes on the Earth's surface, lead to shape cities and results in land use change (Barredo et al., 2003; Li et al., 2003). The factors outlined above can be used separately or in some combination depending upon the objective of applications and models used in order to simulate the urban development.

Environmental characteristics are usually regarded as constraints for urban growth (Barredo et al., 2003). For example, the slope factor is used as a constraint in the SLEUTH model (Clarke et al., 1997; Silva and Clarke, 2005). Further examples of constraints include agricultural suitability areas derived from slope and soil maps (Yeh and Li, 2002), or water bodies (Ward et al., 2000; Yang and Lo, 2003).

The second factor is the effect of neighbourhood. Dynamics systems such as those applications developed based on the CA approach, distance from new features to existing land uses and the type of these land uses drives the urban dynamics at local scale. For example, new residential areas usually emerge near or adjacent to existing residential areas (Almeida et al., 2003; Barredo et al., 2004; Batty and Xie, 1997; Engelen et al., 1997a). A neighbourhood effect can be regarded as the effect of agglomeration (Wu, 1998), factor based on attraction in examining residential development. However, the neighbourhood of industrial areas can represent a repulsive factor (Barredo et al., 2003; Barredo et al., 2004).

The third group of factors are based on the spatial characteristics of cities, including positive attractions for urban land uses (Barredo et al., 2003). They include accessibility to employment and residential areas (Waddell, 2000a), accessibility to the road network (Martin and Wu, 1999), cost distance to the city centre and industrial areas (Wu, 1998), prices for properties (Wu and Webster, 1998) and other physical suitability factors (Engelen et al., 1995; Li and Yeh, 2000).

The fourth group can be regarded as institutional controls, including urban and regional planning policies. From a practical point of view, this group is represented by land use zoning status (Barredo et al., 2003). Many research projects have included this factor in order to test planning scenario perspectives (Engelen et al., 1997a; Ward et al., 2000).

The final group are factors that are related to individual preferences, levels of economic development and socio-economic and political system (Barredo et al., 2003). These factors are usually grounded in human decision-making processes, which in most cases are qualitative and unpredictable and therefore difficult or almost impossible to measure (Barredo et al., 2004; Malczewski, 1999a; Voogd, 1983). For example, a new residential area could be located in a place because it is more "beautiful" than other places (Barredo et al., 2003).

Despite the reliance on these factors for assisting urban growth simulation, however, there are further issues that need to be addressed in the context of data and information when considering differences between the cities of developed countries and those in developing nations. In developed countries, a vast array of relevant spatial data (e.g. detailed land use and land cover data with known temporal and spatial accuracy) and aspatial data (e.g. socio-economic and environmental information) tends to be available (Bishop et al., 2000; Herold et al., 2003). They are supplied in digital format, suitable for analysis and manipulation (Bishop et al., 2000). However, the availability of spatial and relevant aspatial data for cities in most developing countries is little considered, poor in quality or non existent (Barredo et al., 2004; Bishop et al., 2000; Wu, 2002b). Most socio-economic data such as census survey data are frequently outdated with poor temporal accuracy and consistency (Herold et al., 2003). Digital production of spatial data is even rarer often due to the lack of appropriate equipment and expertise (Bishop et al., 2000). Further, digital map data is available for a very few cities due to high operating costs, large and complicated technical work and administration bottlenecks (Bishop et al., 2000). The existing historical spatial data or land use maps, if available, tend to be in the form of unscaled sketches (Barredo et al., 2004; Bishop et al., 2000; Wu, 2002b). Due to urban expansion, they rarely cover the whole urban target area (Barredo et al., 2004). Further, where current maps exist, they are often of different scales or different spatial units, exacerbating the problem of sharing information and manipulation of spatial analysis (Bishop et al., 2000; Wu, 2002b). In most developing countries, including Thailand, the dissemination of classified large-scale maps and aerial photographs is restricted due to confidentiality reasons, and as a result, access by public departments or individual researchers is very difficult (Bishop et al., 2000; Herold et al., 2003; Wu, 2002b).

2.4 Geographic Information Systems (GIS): Their Role in Urban Applications

Geographic Information Systems (GIS) present a powerful tool designed for acquisition, management, and analysis of spatially-referenced data and output generation needed by a particular user (Chakhar, 2003). According to Scholten and Stillwell (1990), the popularity of GIS is due to its three main strengths. Firstly, GIS allow large amounts of data obtained from a wide variety of sources, both spatial and aspatial, such as remote sensing, aerial photographs and field investigations to be integrated, stored, converted, updated and displayed to a common spatial framework (Grossmann and Eberhardt, 1993; Heywood et al., 2002; Scholten and Stillwell, 1990). Secondly, GIS provide the means to analyze spatial data (Heywood et al., 2002; Scholten and Stillwell, 1990). A large body of methods for spatial analysis have been developed over the past century ranging from simple to sophisticated functions and including queries, measurements (e.g. distances, slope and aspect measurements), transformations (e.g. buffering and overlay functions), data mining (e.g. measures of patterns), optimization functions (e.g. point allocation, optimum paths) functions (Longley et al., 2005; Malczewski, 1999a). Finally and most importantly, GIS have abilities to handle the management of large quantities of data, multilayered and in heterogeneous databases, such that this information can be easily accessed by all users and their queries and retrievals, associated to spatial objects (e.g. locations and properties), can be accomplished in an interactive way (Fischer and Nijkamp, 1993; Scholten and Stillwell, 1990).

GIS are considered a proper tool to handle large amounts of data, with both spatial and aspatial aspects, of urban areas (Scholten and Stillwell, 1990). Urban applications based on GIS are numerous. Nowadays, GIS have been proven to be extensively used and integrated in many different fields (Malczewski, 2006; Ottens, 1990), ranging from the creation of land suitability for various land use activities (Ceballos-Silva and Lopez-Blanco, 2003), site selection (Carver, 1991; Sener et al., 2006), monitoring the urban growth (Hara et al., 2005; Thomson and Hardin, 2000), to simulation of the development of cities and change of land uses (Li and Yeh, 2000; Wu, 1998). In this section, a review of those GIS applications is presented in three parts, (1) urban applications developed and based on standard GIS functions, (2) urban applications developed based on integration of GIS with other approaches, including multi-criteria decision analysis (MCDA) and cellular automata (CA), and (3) GIS for urban simulation applications in developing countries.

2.4.1 Urban Applications Developed Based on Standard GIS Functions

GIS applications developed from an urban planning perspective are popular in the areas of land suitability, site selection and land allocation (Banai, 2005; Cromley and Hanink, 1999; Joerin et al., 2001). One example is given by the earlier work of McHarg (1971) who applied the traditional overlay routines to analyze areas suitable for land use activities including natural conservation areas, passive and active recreation areas, and residential and commercial-industrial areas for Staten Island (New York City). In his application, over thirty factors, mainly subdivided from the categories of climate, geology, soils, vegetation, wildlife habitats and land use, were taken into account. For each of these categories, data was ranked based on factors of importance to all potential land use.

This application exploited standard GIS functionalities, mainly overlay and buffer functions, to evaluate the characteristics of land from a number of layers or criteria for each location at the same time (Carver, 1991; McHarg, 1971). The traditional GIS buffer and overlay technique used, despite the fact that it has proven to aid in selection of the best site for some specific purposes on the basis of their suitability scores (Carver, 1991; Janssen and Rietveld, 1990), has been significantly criticized for being limited to descriptive and deterministic map data manipulation (Carver, 1991; Malczewski, 1999a; 1999b).

2.4.2 Urban Applications Developed Based on Integration with Other Approaches

Currently, there is much discussion over how to enhance the capability of GIS to help effective urban planning and management (Malczewski, 2006; Wagner, 1997). Approaches include coupling GIS with either collaborative decision-makers for the creation of land suitability assessment (Banai, 2005; Carver, 1991; Joerin et al., 2001) or temporal dimension for the purpose of simulation of urban development (Barredo et al., 2003; Cheng and Masser, 2004; Wagner, 1997), or both in combination (Li and Yeh, 2000; Wu, 1998).

GIS can be coupled within different approaches either a loose-coupling or a tight-coupling strategy (Heywood et al., 2002). A loose-coupling strategy refers to the model running as a separate piece of software with data exchanged to and from the GIS (Bivand and Lucas, 2000; Heywood et al., 2002). With this approach, GIS are heavily employed for data preparation and output display purposes (Malczewski, 1999a). This approach is considered difficult to handle, especially for non-technical users, because it requires knowledge of high-level programming skills (Openshaw, 1978). Besides, since data input and output need transformation and conversion, it is considered time consuming and, to some extent, can cause mechanical or transcription error due to problems in data transferring (Heywood et al., 2002). Examples of these applications are the SLEUTH model (Clarke and Gaydos, 1998; Clarke et al., 1997) and UrbanSim work (Waddell, 2000b; 2002). These models developed separately for urban simulation, while GIS are used for data preparation and visualization.

In contrast, tight-coupling strategy refers to a model that is executed as GIS scripts, or through a GIS graphical user interface (GUI), usually with a common industry-standard language such as Visual Basic for Applications (VBA), Javascript, or Python or interoperability standards such as Microsoft's .NET (Heywood et al., 2002). With this approach, the implicit transfer of data between different software packages is implemented (Bivand and Lucas, 2000; Heywood et al., 2002). For example, the work conducted by Yeh and Li (2001) coupled CA with GIS for modelling the sustainable urban development of the Pearl River Delta of China. Additionally, Wu (2002a) integrated CA and GIS to simulate rural-urban land conversion in the city of Guangzhou in South China. Both research projects were built on a GIS platform using ARC/INFO's AML (Arc Macro Language) programming language to develop the CA model. Although programming knowledge and skills are required intensively to build the model beforehand, the model built is considered easy for users to understand and utilize since data input, process and output are rigidly integrated together (Sui and Zeng, 2001). Research discussed later in this thesis uses a tight-coupling strategy.

2.4.2.1 Integration of GIS and Multi-criteria Decision analysis (MCDA)

Land suitability is determined by the fitness of land for a particular use on the basis of decision-makers' goals and interests (Bojorquez-tapia et al., 2001). A map of such land suitability is considered very useful for land planning and sustainable management (Joerin et al., 2001). Land suitability applications such as those of McHarg (1971) and Carver

(1991) discussed earlier show simple criteria handling and full implementation by the standard GIS functions. However, in real world problems, urban planners have to deal with complex decisions, usually based on various interests and judgments which result in conflicting objectives and multiple attributes in the development process. GIS have been criticized for its limited capabilities when involving such conflicting objectives and multiple attributes (Carver, 1991; Chakhar, 2003). One solution is to couple GIS with multi-criteria decision analysis. More details about MCDA model are discussed later in Section 2.6.

Numerous techniques and methods in the area of optimization and multi-criteria decision analysis (MCDA) are proposed and integrated with GIS in order to help solve complex decision problems (Malczewski, 1999a; 2006). An example (Grabaum and Meyer, 1998) integrated GIS and optimization methods using a loose-coupling strategy on the basis of goal programming technique for the calculation of an optimal land use pattern among four objectives for the test site of Jesewitz, northeast of Leipzig, Germany. The dataset for each objective was prepared and assessed using GIS, while an optimized solution among all objectives or goals was achieved on the basis of game theory by minimizing the maximum difference of each objective from its optimal value using a specially written PC software package LNOPT (LandNutzungsOPTimierung - German acronym for 'land-use optimization') outside the GIS environment.

Nowadays coupling GIS with multi-criteria evaluation (MCDA) techniques is widely used as an aid for land suitability and is accepted as a supportive tool for land planning and management (Bojorquez-tapia et al., 2001; Joerin et al., 2001). Amongst those applications that integrated GIS and MCDA for urban planning, one popular application is land use assessment (Joerin et al., 2001), which is mostly applied to determine areas suitable for specific land use activities (Wu and Webster, 1998; Yeh and Li, 2001). Different from site selection, where the goal is to isolate the best alternatives or sites, land use assessment allows multiple criteria to be considered in evaluating the potential or appropriate sites for the whole study area (Joerin et al., 2001).

Examples of land use assessment by Joerin and Musy (2000) and Joerin et al. (2001) include MAGISTER (Multicriteria Analysis and GIS for Territory), a decision support model, to assist land planners for land suitability assessment. The advantage of the model is that it allows multiple actors or participants to be incorporated for land management. According to the model, GIS manage information describing the territory and offers spatial analysis functions for computation of criteria and simulation, while MCDA (on the basis of the multi-objective decision technique) provides techniques to aggregate or weigh the information and choose the most appropriate solution based on the preference of decision-makers. In Joerin et al. (2001), a land suitability map for housing in the canton of Vaud, Switzerland was selected as a test site. Eight criteria such as noise impacts, risk of landslide etc. were evaluated among participants. A land suitability map for housing was created based on the criteria using an outranking MCDA method built upon a GIS platform.

Site selection is an application exemplified by Sener et al. (2006). They integrated GIS and MCDA (on the basis of the multi-attribute decision technique) in order to seek the appropriate landfill area for waste disposal in the vicinity of Ankara, Turkey. In their approach, 16 input map layers were used as criteria. These factors include topography, settlements, roads, railways, airport, wetlands, infrastructures, slope, geology, land use, floodplains, aquifers, and surface water. Those layers were used within the GIS analysis environment using two MCDA methods, the simple additive weighting method (SAW) and the analytic hierarchy process. The outputs of site selection maps from both methods were produced by means of multiplication of data layers, weights, and constraints.

The applications described above allow planners to select the appropriate sites for specific development on the basis of factors influencing land use development. They are, however, considered static and may not be suitable to simulate the spatial pattern of growth.

2.4.2.2 Integration of GIS and Cellular Automata (CA)

Urban simulation helps urban planners to describe, understand, predict and estimate future impacts on land use and the development of existing spatial plans (Barredo et al., 2004). It can, to some degree, help them to understand the consequences of proposed projects and planning policies (Barredo et al., 2004; White et al., 2000). Such simulation requires that GIS enhance their analytical functions to account for dynamic or time dimension in order to incorporate the process of land use change over time (Wu and Webster, 2000). However, despite its prevailing capabilities to manipulate spatial properties explicitly

(Wagner, 1997; Wu, 1998; 1999), GIS have been criticized for several limitations in simulating change of land over time (Li and Yeh, 2000). These include its poor ability to deal with the temporal dimension (Heywood et al., 2002; Longley and Batty, 2003; Longley et al., 2005; Wagner, 1997; Wu, 1999) and limited performance for many spatial operations (Engelen et al., 1997a; Wu, 1998). This is because GIS models and structures have not been designed to record, store, or visualize spatial information in association with different temporal states (Heywood et al., 2002). In dynamic modelling, those based on cellular automata are most likely popular and widely applied in simulation applications as they have functionality to handle time explicitly (Caruso et al., 2005; Cheng and Masser, 2004; Yang and Lo, 2003).

Integration of GIS and CA has been widely accepted to fulfill the strengths of both methods (Cheng and Masser, 2004; Wagner, 1997). While CA serves as an analytical engine for programming and running dynamic simulation (Cheng and Masser, 2004; Li and Yeh, 2000), GIS provide powerful functions for spatial handling, allowing real world data, factors and constraints to be incorporated (Wagner, 1997). Thus, integration between them results in a realistic simulation outcome (Engelen et al., 1997a). For example, in Engelen et al. (1997a), the modelling approach was carried out to simulate the development of land use by using raster-based GIS, which allow spaces or attributes to be manipulated and represented easily (Wagner, 1997). The model comprised 3 separate components linked in as a loose-coupling strategy; (1) growth coefficients input derived from global and regional scale data (operated at a macro level), (2) physical, environmental, technical and institutional factors input to GIS to create a suitability map and (3) the CA model (operated at a micro level). The modelling permitted the integration of growth coefficients from a global scale, coupled with a suitability map derived from GIS to be fed to a CA model to generate dynamic urban simulation for the city of Cincinnati and small islands in the Caribbean.

Another example is the work of Wu and Webster (2000) who applied a CA model on the basis of micro-economic theory supplied as transition rules to simulate artificial cities. In their model, a statistical multinomial logistic regression technique was embedded within GIS in order to calculate and predict the probability of urban development (development from vacant area to industrial area) based on the relationship of profitability and development density. Simulation of two scenarios including urban development under a free market and development under regulated regimes was generated and compared. The model was implemented using a tight-coupling strategy. CA simulation was coded as in the AML programming language, coupled with the spatial analysis function (from the ArcGIS GRID module) within a GIS environment. Such integration allows implementation to be undertaken directly within a GIS environment where the variables of the model and the output are presented as maps.

One more example is the work conducted by Wu (2002a). His work simulates rural-urban land conversions in the real city of Guangzhou, China. The model calibrates the consistent weights (coefficients) to be set as rules in CA using the sequential observation of land use change (obtained from the spatial overlay of land use layers in the two periods). According to the model, development factors are mainly based on the physical and location characteristics such as the travel distance to the edge of the city, topography, and slope. In this model, GIS, mainly using its spatial analysis functions, play the main roles in observing and investigating the pattern of change of land use conversions through the overlay of land use layers between two epochs. Variables measured by GIS operations were extracted and exported to a SPSS statistical software package to calculate the coefficients or weights of the CA rules using a binary logistic regression technique. Simulation based on the development probability was conducted using a stochastic approach – the Monte Carlo method. Simulation of three scenarios was conducted in order to test the accuracy of realistic simulation based on the effect of global (through coefficients derived) and local factors (through neighbourhood development). According to the simulated result, it was concluded that scenario based on the integration of global and local factors, considered as static and dynamic factors respectively, produce the most realistic simulation.

Recently there has been an effort to incorporate MCDA as definition rules in CA. However, urban simulation models based on the integration of GIS, CA and MCDA applications relatively few. As examples, Wu (1998) and Wu and Webster (1998) used a tight-coupling strategy for their land conversion simulation model, called SimLand. In this model, the modules of CA and MCDA (on the basis of multi-attribute decision technique) through the analytical hierarchy approach (AHP) was written in the C programming language and built into an ARC/INFO GIS. Analogous to Wu (2002a) in that its objective is to simulate land use conversion in the city of Guanzhou, the focus of this model is, however, to propose AHP as the CA's transition rules for development within neighbourhoods where development occurs (Batty and Torrens, 2001). While GIS provide a platform and a spatial database, MCDA allows a decision-maker's preference to be taken into account. According to AHP, a set of weights is extracted through a pairwise comparison technique carried out by decision-makers. The pairwise comparison technique allows the importance of different development factors to be considered against one another. Many development factors regarding socio-economic indicators were included, such as cost of traveling the distance to the city centre, cost of travelling the distance to industrial areas, cost of travelling the distance to newly built railway stations, access to the highway, development density and agglomeration in terms of neighbourhood effect. The interest in adapting MCDA is that the weights derived thus represent substantive comparisons which can be related to what we know about how land use is developed (Batty and Torrens, 2001). The simulation results, however, were carried out using weighted summation to calculate suitability scores which in turn define the probability of land conversion in simulation. Four simulation schemes were generated. They were 'centre-dominated growth', 'centre plus industrial district growth', 'compact growth around the new suburban railway station' and 'highway-promoted growth' scheme. Wu (1998) claimed that with the combination of the three components – GIS, CA and MCDA- in his model, the model has many advantages: visualization of decision-making, easier access to spatial information and more realistic definition of transition rules.

Although many applications described above have demonstrated that the integration of GIS with cellular automata (CA) and/or multi-criteria decision analysis (MCDA) can be used to model urban growth and to predict the spatial pattern of urban development for different planning scenarios, several issues need to be addressed. These include its reliability for model prediction, model calibration and validation when implementing urban simulation based on the CA approach (Wu, 2000). Another issue concerns setting weights and techniques chosen when handling decision-making processes as different results are generated with different techniques and weights (Hajkowicz and McDonald, 2000; Sener et al., 2006; Voogd, 1983). Furthermore, their applications are mostly limited to developed nations where data and its relevant information are rich and available (Barredo et al., 2004). Finally, technological experiences and expertise in GIS, CA and MCDA as well as other relevant fields for model implementation, are rare and based mainly on the experiences of cities in developed nations (Barredo et al., 2004; Bishop et al., 2000).

2.4.3 GIS for Urban Simulation Applications in Developing Nations

In the developed countries, plenty of research has been applied using GIS for urban analysis and simulation (e.g. the work of Waddell et al. (2005) and Wagner (1997)). Similar applications are, however, few and novel in most developing nations (Bishop et al., 2000; Yeh, 1999). Limitations in the adoption of GIS in developing nations are due to the high data cost, data availability and their restriction in use, the high cost of acquiring the technology, the shortage of trained and skilled staffs and experts, and the lack of expertise and experiences (Barredo et al., 2004; Bishop et al., 2000; Herold et al., 2003). Further, since the nature of growth of cities in developing nations is different from those of the developed nations (Cohen, 2004) (see discussion in Section 1.1), development factors that are commonly used for the simulation of growth of cities in the developed countries cannot truly reflect the growth of the cities in the developing nations. For example, the policy zoning system and land use plan as a product of institutional control, which are effective for the land use control for most developed countries, are not strongly applied in most cities in the developing nations (Barredo et al., 2004) such as Bangkok, Thailand (Bishop et al., 2000; Chomchan et al., 1990; Webster, 2000) and Lagos, Nigeria (Barredo et al., 2004).

Up to now some of the research applied to study sites in the developing nation were those undertaken by Wu (1998), Wu and Webster (1998), and Wu (2002a)in Guangzhou, a major city in southern China. Their work was carried out in order to identify areas likely to experience rapid urban growth in the city. Though these projects were based on the same study site they were different in terms of CA rules applied for urban simulation (see details in Section 2.4.2.2), and were carried out using the tight-coupling strategy between CA and GIS. In their work, data were delimited by the acquisition of two digital satellite imageries of 1993 and 1998 and the land use maps derived from the 1:10,000 aerial photographs at the same periods.

Another example conducted by Li and Yeh (2000) and Yeh and Li (2002) used 1993 and 1998 Landsat TM images to derive the urban areas of Dongguan city situated in the Pearl River Delta in China. The earlier work attempted to describe the spatial pattern of urban growth using differing planning scenarios based on the concept of local, regional, and global constraints. In their study, a CA approach within a grid GIS-based system were built to accommodate easy access to the GIS database for producing constraints. These were used for the development probability calculation as rules for urban simulation. The later work used one development factor, regarding the different degrees of distance to the centre, to test monocentric and polycentric scenarios of urban development. Both projects applied the concept of 'grey cell' in which the degree or percentage of urban land development increases during the iterations of modelling.

These studies share some similarities. Firstly, although the techniques applied were similar to those of the developed nations, they were confronted by limited data availability in terms of historical and high resolution datasets. In the studies, available data were mainly acquired from satellite imagery which can only reflect modern land use. Classification of land cover, thus, does not allow the creation of detailed land use categories (e.g. residential, commercial). As a result, simulation of urban growth, at best, was carried out for the conversion of developed areas (e.g. urban) from undeveloped areas (e.g. non-urban, agriculture, forest). Secondly, development factors that are used for creation of probability or potential maps were mainly based on the spatial or physical characteristics (e.g. distance to city centre, distance to roads, distance to railways, slope) acquired from the study site only.

In the case of Thailand, applications of modelling for urban analysis to date have focused on land use change detection and urban expansion monitoring. Those have been conducted with the integration of remote sensing images and aerial photographs. For example, the work conducted by Tachizuka et al. (2002) used remote sensing technology based on two types of different sensors as a tool for monitoring and detecting the land cover changes in and around Bangkok areas. With the limitation of available detailed data, however, land use change detection can only be identified for the conversion of bare land and vegetation to urban development during the ten-year period (1993 to 2002) in which satellite images are available. Another example is the work conducted by Hara et al. (2005). In their study, land use change was focused on the urban sprawl, the shifting from paddy fields to urban dwellings, in the suburbs of Bangkok area. Their research for land cover monitoring was conducted through the interpretation of aerial photographs of two epochs (1952 and 1998) and field measurements. GIS were applied for data analysis, acting as a way to detect the land cover change for both horizontal and vertical (high-rise buildings) components. Neither of these two Thai examples, however, have means for land use change simulation.

Up to now, there has been little evidence of attempts to predict dynamic urban simulation of the Bangkok area. One, conducted by Bruijn (1991), used a Bangkok development probability model to simulate the urban development of Greater Bangkok based on three spatial factors; proximity to core area, proximity to roads and proximity to development fringes. According to his work, the result was conducted solely to examine the conversion of non-developed areas. The model, however, had only a small amount of detailed data, and a grid cell of 500m x 500m. Further, the actual model is, in reality, static as development probability did not take into account local factors such as the neighbourhood effect described in the CA approach.

To recap, the applications discussed above show that GIS have been integrated with remote sensing technology and spatial models in order to monitor and simulate urban growth and development. Recent trends have seen some applications coupling GIS with a dynamic model, such as that of CA, in order to forecast the spatial pattern of urban growth, and that of MCDA in order to incorporate planners or participants for testing scenario-based land forecasts. However, they have all developed based on the constraints and limitations previously mentioned.

2.5 Micro-simulation for Urban Land Use Change: Cellular Automata and Its Extension

2.5.1 The Cellular Automata (CA) Approach

Cellular Automata (CA) is one of the promising simulation modelling approaches to depict the evolutionary dynamics of phenomena (Cecchini and Rinaldi, 1996). The CA concept was first proposed by Ulam and von Neumann in the late 1940s (Engelen et al., 1997a; Silva and Clarke, 2005) as an exercise in the philosophy of computation and then was popularized by the work of the mathematician John Conway in 1970 as "The Game of Life" (Batty and Xie, 1994; Longley et al., 2005). Despite the fact that the concept of CA was first originated in computation sciences, it has been widely used in many scientific disciplines such as Mathematics and Physics (Engelen et al., 1997a; White et al., 2000).

The basic elementary CA consists of a series of cells arranged in one-dimensional space and in regular grid-squares, a set of cell states, a neighbourhood, a transition rule(s) and a dynamic or temporal space (Silva and Clarke, 2005; Torrens, 2000; 2003). The state of cells may evolve through a number of discrete time steps according to the previous state of the cells in accordance with a set of simple transition rules, defined in terms of neighbourhood functions (Cecchini and Rinaldi, 1996; Engelen et al., 1997a; Silva and Clarke, 2005; White et al., 2000). In the geographic field, CA was first introduced by Waldo Tobler through 'cellular geography' (Tobler, 1979). However, the basic CA has been criticized as being too simplified and too restricted to represent and simulate the real-world urban and city system (Engelen et al., 1997a; Silva and Clarke, 2005; Torrens, 2000; 2003). As a result, several significant modifications of conventional CA have been applied in order to help build a practical model that can represent and capture more realistic geographical problems (Torrens, 2000; Torrens and O'Sullivan, 2000). These modifications include addressing elements such as cell space, cell state, cell neighbourhood, transition rules, and time iteration.

Cell Space (Lattice), in most cases of geographic perspective, occupies two dimensions, instead of classical one-dimension, and cells are square and have uniform regular space (Torrens, 2000). The cells are represented as the smallest, finest and irreducible units and they must have adjacency and proximity manifestation (Silva and Clarke, 2005). In urban applications, cells can be used to represent any zonal geography within a city from a parcel of land to a large administrative area (Torrens, 2000). The cell space of CA maps is set up similar to the raster data structure of geographic information systems (GIS) (Silva and Clarke, 2005), thus allowing them to cope well with the digital, raster nature of data from remotely sensed images and other sources, and also be convenient for programming and implementing grid-based structures (O'Sullivan and Torrens, 2000).

Cell state, in urban applications, allows dominant characteristics of a city to be encoded for the development of a cell through a simulation model (Torrens, 2000, Silva and Clarke, 2005). Each cell occupies one of a finite number of discrete states. In an urban context, the state of a cell can be assigned to represent attributes of the urban environment. For example, in urban growth simulation, cell state can be encoded in a binary fashion of urban (developed) and non-urban (non-developed) area (e.g. Batty (1998) and Ward et al.(2000)), classified land use types (e.g. White et al. (2000)), property price (e.g. Wu (2002b)), population densities (e.g. Yeh and Li (2001)), or the grey cells function of Li and Yeh (2000) in which the binary states have been replaced by several transient (continuous) states as a result of the degree or percentage of land development during the iteration of simulation modelling.

The neighbourhood of a cell comprises a set of adjacent cells around a considered cell or automaton (Torrens, 2003). Neighbourhoods are considered the key element of a CA since they drive the interaction between the land use and dynamics of the systems (Caruso et al., 2005). In the urban CA, neighbourhoods represent sphere of influence or activity within the cities (Torrens, 2000) such as the walking radius of commuters. In the literature about CA, the most famous neighbourhood concepts are Moore's (or Conway's) and that of von Neumann. The former (Moore) include eight neighbouring cells and the latter (von Neumann) includes four neighbouring cells (Benenson and Torrens, 2004) (Figure 2.7).

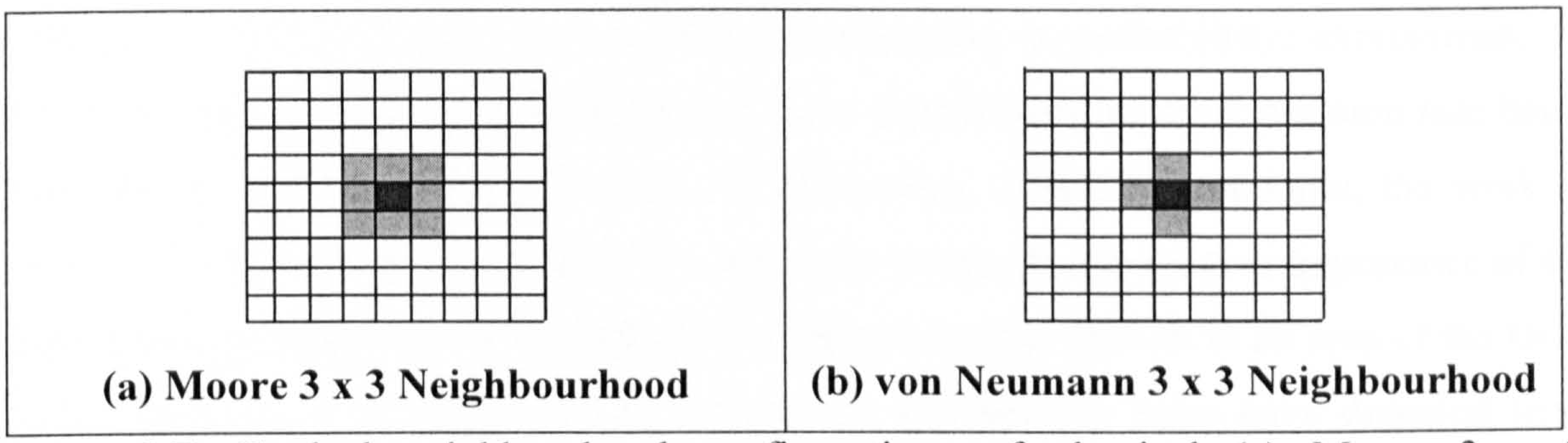


Figure 2.7: Typical neighbourhood configurations of classical (a) Moore 3 x 3 Neighbourhood (8 cells) and (b) von Neumann 3 x 3 Neighbourhood (4 cells) (adapted from Benenson and Torrens (2004), p.94).

Ward et al. (2000) suggested that when simulating urban land development by accounting for the influence of a road network, the Moore neighbourhood is best for handling cul-de-sac networks while the von Neumann neighbourhood is most suitable for a regular grid road network. He explained that this is because the Moore neighbourhood allows diagonal and perpendicular access to the transport network whilst the von Neumann neighbourhood allows only perpendicular access. In the real world, neighbourhoods have been enlarged beyond the classical Moore's and Von Neumann to accommodate action-at-a-distance (Torrens and O'Sullivan, 2000). It was suggested by Batty and Xie (1994) and Ward et al. (2000) that since space-filling during simulation using the 3x3 neighbourhood basis (local nearest neighbour) is probably inadequate to capture the real world spatial interaction of urban system, the broader or extended neighbourhood is required. For example, Engelen et al. (1995) applied the neighbourhood of 113 cells while Ward et al. (2000) used a square neighbourhood with a width of 11 cells for their study. Further, some applications modified the shape of a neighbourhood from the classical rectangle to different shapes such as the circular neighbourhood of Kocabus and Dragicevic (2004) and Yeh and Li (2001).

Transition rules are primarily used to drive dynamics of change in the model (Torrens, 2000). In an urban context, the rules, coded as a set of algorithms, are determined in order to mimic how urban phenomena in the real world operate (Torrens, 2000). A set of rules needs to be spatially uniform, applying to every cell, state, and neighbourhood and more importantly, every transition of cell from one state to another state must be local (Engelen et al., 1997a; Silva and Clarke, 2005). Rules can be deterministic or stochastic (e.g. probabilistic) (White and Engelen, 1993). In traditional CA such as that of the classical "The Game of Life", rules are set to be deterministic and considered not to change during the evolution (White and Engelen, 1993). Several studies of urban CA models are now modified to employ stochastic or probabilistic elements. In probabilistic expressions, the action of a transition function is dependent upon a probability or other decision rule being incorporated within a model (Torrens and O'Sullivan, 2000). For example, the work of Ward et al. (2000) used the probabilistic approach to specify the relative importance of the geographic constraint for the CA-based simulation of urban growth in an area of the Gold Coast, Australia. Until now, several adaptations in transition rules have departed from traditional CA in order to understand or mimic the urban phenomena. For example, (Batty and Xie (1997) simulated urban development of four land use types; residential housing, industry, commerce and transport links. Each land use transition applied different kinds of rules on the three basic ideas of reproduction (birth), mutation (survival) and exogenous change (unpredictable change from outside the system). An example of Yeh and Li (2002) applied development density using the distance decay function and the concept of 'grey cells' in the simulation of development for various urban forms of both monocentric and polycentric urban scenarios. Wu and Webster (2000) determined the transition rules based on urban micro-economic theory and statistical multinomial approach for artificial city simulation while Sui and Zeng (2001), Soares-Filho et al. (2002) and Almeida et al. (2003) applied a logistic approach and incorporated criteria defining the physical, social and economic characteristics for simulating urban development.

Time step or dynamism is one key element of CA. Dynamism with a CA needs an initial set of states over the cells and a sequence of discrete time steps (Silva and Clarke, 2005; Torrens, 2000). At each iteration, cells at each location will be updated simultaneously based on transition rules and the composition of cells in the neighbourhood (Silva and Clarke, 2005; Torrens and O'Sullivan, 2000). Torrens (2000) claimed that such interactive dynamism of CA, is considered realistic and flexible enough for modelling cities if the discrete time steps are so small enough and if data permits.

In spatial urban applications today, CA has gained attractiveness for the purpose of the simulation of urban systems with three major research objectives; the exploration of spatial complexity, urban theories testing and the development of operational urban simulation (Torrens and O'Sullivan, 2000). For example, Batty and Xie (1997) experimented using a fractal geometry concept in order to explore the spatial complexity of urban systems. Batty (1998) developed simple simulation on the basis of spontaneous growth, which is considered as a major driving force to characterize the regeneration of the location of inner and edge cities theory. Torrens (2006) simulated the suburban area in the Midwestern megalopolis region around Lake Michigan in the United States using the basis of city and urban sprawl theory. The work of Engelen et al. (1995) and Engelen et al. (1997b) applied decision support and urban planning to make urban CA models of the cities to be operational in use as a means to help specify detailed land use planning. In their model, the integration of macro (in terms of socio-economic activities demand) and micro (CA) scale, as well as GIS layers (in terms of physical suitability) were integrated for growth simulation. The research conducted by Wu and colleagues (e.g. Wu (1998); Wu and Webster (1998); Martin and Wu (1999)) applied CA and a multi-criteria decision analysis approach as well as a GIS tool to help support setting criteria and scenario testing to reflect the impacts of different decision-making in urban planning practice. Ward et al. (2000) developed a constrained CA model using physical, environmental and institutional control factors in order to simulate the control and forbidding urban growth.

The popularity of urban CA models is due to four main advantages: spatiality and affinity with GIS, dynamism, micro-simulation, and a bottom-up approach. Firstly, CA has a natural ability to handle spatial phenomena, usually in a two-dimensional plane where the urban phenomena are presented (Wagner, 1997). Further, CA is regarded as having affinity with raster data, suitable for GIS and often in the form of remotely sensed information

(Torrens, 2000; Yang and Lo, 2003). As a result, many applications have been built based on degrees of coupling between these two approaches. Secondly, cities are considered the spatial results of many agents (e.g. individuals, public and private bodies) acting simultaneously over the urban space (Barredo et al., 2003). CA has ability to implement spatial dynamic modelling and handle time flexibly and explicitly (Wagner, 1997; Torrens, 2000). Thus, it is well-suited for modelling and simulating cities, especially when the cycle of interactive events over the urban space such as the long-term economic cycles and daily commuting behaviour are varied temporally (Torrens, 2000). Thirdly, CA can be regarded as a micro-simulation approach (Silva and Clarke, 2005) since it can handle data dynamically at fine resolution, resulting in detailed simulation (Torrens, 2000). Fourthly, CA has been regarded as applying a bottom-up approach for simulation (Li and Yeh, 2000; Silva and Clarke, 2005). CA allows macro scale pattern (macro-simulation) to emerge from the interactive dynamics of local elements (micro-simulation) (Silva and Clarke, 2005; Torrens, 2000). Such behaviour is crucial for studying the complex systems of cities (Barredo et al., 2003; Torrens, 2003). These major strengths mean the CA concept is popular and widely used for modelling dynamic interaction of urban and geographical phenomena as it can model function and form, pattern and process, simultaneously and in an interactive manner (Torrens, 2000).

However, CA has some criticisms. It has limited capability to handle spatial patterns. In the spatial context, CA can provide input, output and dynamic analytical functionality, but it has no capability to store, retrieve and query a database (Wagner, 1997). In addition, CA is considered weak for many map manipulations such as map projections or map transformations (Wagner, 1997). Such limited capabilities of traditional CA are considered impractical when dealing with urban applications where the requirements include facilitating map manipulation and query functions, coupling with other spatial layers for being integrated for further analysis and involving both spatial and aspatial (attributes) aspects. Secondly, like other urban simulation models where their key purpose is to create realism, CA uses a set of transition rules and neighbourhood characteristics to simulate the growth. However, until now, many issues regarding its reliability for model calibration, verification and validation, the best set of transition rules, as well as the appropriate size and shape of neighbouring cells remain unanswered (Wu, 2000).

2.5.2 The SLEUTH Model

SLEUTH, sometimes called Clarke Urban Growth Model (UGM), is a CA-based micro-simulation urban growth model, developed to forecast urban growth and land cover change on a yearly basis (Benenson and Torrens, 2004; Herold et al., 2001; Jantz and Goetz, 2005; Silva and Clarke, 2005). The model was originally built by Keith Clarke (Clarke et al., 1997) at the University of California at Santa Barbara for the application of simulation of spatial pattern of urban development in the San Francisco Bay area (Clarke et al., 1997; Jantz and Goetz, 2005; Yang and Lo, 2003). Until now SLEUTH has been applied operationally for prediction of land cover change in many areas such as in the Middle Rio Grande Basin in central New Mexico (Hester, 1999), in the Atlanta metropolitan area (Yang and Lo, 2003) and recently for the cities of Lisbon and Porto (Silva and Clarke, 2005).

Using an underlying CA-based approach, the model works in homogeneous grid cells, with a Moore neighbourhood of eight cells (Silva and Clarke, 2005). Formation of the model is primarily based on the input data and local growth rules (Clarke et al., 1997). Input to the SLEUTH model requires six primary data layers; Slope, Land use, Excluded areas, Urbanization, Transportation and Hill shading (Clarke et al., 1997). These inputs give the model its name, SLEUTH. The Slope layer (e.g. steepness and aspect) and Excluded area layer (e.g. national park) represent the areas being constrained for urban development. Urban land use area is used initially for the creation of a seed layer that will act as a growth centre. The Transportation layer stimulates the direction of urban growth, while the Hill shading layer is for the purpose of visualization. Five transition rules are set, comprising a diffusion factor, a breed coefficient, a spread coefficient, a slope resistance factor and a road gravity coefficient (Clarke and Gaydos, 1998; Clarke et al., 1997). Each rule reflects one primary behavioural factor of urban growth dynamics in order to represent the four types of growth: spontaneous growth, new spreading centre growth, edge growth, and road-influence growth (Silva and Clarke, 2005). Explained by Clarke et al. (1997) and Silva and Clarke (2005), urbanization, at each iteration in the model comprises four sequential time steps. Firstly, spontaneous new growth simulates the occurrence of random urbanization of land. Secondly, new spreading centre growth simulates the development of new urban centres around the urbanized cells emerged from the previous growth process. Thirdly, edge growth simulates the development of cells having at least three adjacent new or existing urbanized cells. Finally, road-influence growth simulates the new urbanized cells along the roads under the influence of transportation. Based on these growth patterns, output from the model represents land cover comprising two cell states, either urban or non-urban.

SLEUTH has several advantages. Firstly, following the CA approach, it is scale independent and dynamic, considered basic needs of any dynamic urban growth simulation (Yang and Lo, 2003). Secondly, since the model is designed to incorporate multiple data sources in accordance with a long historical period of data collection, it allows the model to calibrate in a complex way a set of realistic factors or coefficients to drive the urban growth. Explained by Silva and Clarke (2005), this is due to the fact that the calibration process is carried out on the basis of the complete history of the place from all periods - rapid growth, little or no growth. As a result, it makes the simulated outcomes more realistic and operational in use. Thirdly, SLEUTH has a self-modification function, thus allowing the values of the growth coefficients to be changed when the model iterates through time (Jantz and Goetz, 2005). As a consequence, the realistic simulated results can be obtained on the basis of different rates of growth and different conditions that happen in the dynamic urban system (Jantz and Goetz, 2005; Silva and Clarke, 2005; Yang and Lo, 2003). Finally, the model's source code is open, thus permitting users and developers to update the model and develop their own components to extend the existing model's performance (Yang and Lo, 2003).

Despite these merits, the model has some drawbacks. Firstly, the model only accounts for the physical characteristics including slope, road pattern and exiting land use. In reality, as suggested by Briassoulis (2000) the process of land use change requires many non-physical factors to be incorporated, such as socio-economic, environmental, and institutional control factors. Secondly, the model requires multiple data obtained from many historical periods. Such a requirement is quite difficult to achieve in many countries, especially in the developing nations where data are quite rare and cannot date back over a long period (Bishop et al., 2000). Thirdly, the model can be significantly influenced by the topography and the road network. Thus, such a model may be not suitable for a study area that is quite flat. Finally, the model identifies just one urban land use category, thus it cannot include a detailed description of the urban land use structures (Herold et al., 2001).

2.5.3 The UrbanSim Model

UrbanSim is a newly developed operational, dynamic, urban micro-simulation model (Herold et al., 2001; Waddell, 2002; Waddell et al., 2001). Possessing a micro-simulation-based approach, UrbanSim employs land parcel or household as the basic spatial unit for implementation (Herold et al., 2001; Waddell, 2000b). The model mainly aims to integrate planning aspects for the analysis and prediction of urban development at metropolitan-scale over multiple time periods by incorporating the complex interactions between land use, transportation, and public policy (Herold et al., 2001; Waddell, 1998; 2002; Waddell et al., 2001). UrbanSim, a software-based system, was first built by Dr Paul Waddell and his team as a prototype of a metropolitan land use and transportation planning modelling for Eugene-Springfield, Oregon (Agarwal et al., 2000). The software is now being implemented operationally in many other metropolitan regions of the United States, including Honolulu, Hawaii, Houston, Texas, Salt Lake city, Utah, and Seattle, Washington. Applications to other regions in the United States (e.g. Denver, Colorado) and abroad (e.g. Amsterdam, the Netherlands, Paris, France) are being implemented at an early stage (Waddell, 2006). UrbanSim is currently being re-designed and transitioned to a new platform, named the Open Platform for Urban Simulation (Opus) in order to extend and enhance its model performance (Waddell et al., 2005). More information about the model and its transition is fully available through the UrbanSim homepage (http://www.UrbanSim.org/).

The theoretical basis of UrbanSim is heavily drawn from econometric concepts (Waddell et al., 2001). The model design is developed to reflect the outcomes of simulated decision made by key actors or agents in the urban development process; households, businesses, developers and governments (Waddell, 1998; Waddell et al., 2001). Household and business actors reflect consumer demands and preferences for different types of places and locations. Developers play a significant role in making decisions about development activity, determining where and what kind of construction to build. Government constraints, which are applied as political and environmental actions, are referred to as constrained key inputs which limit the developer alternatives and promote development of land at different locations and with different types of growth.

The main components of UrbanSim embrace several models. They include a model of accessibility (which predicts patterns of interaction), demographic and economic transition (creation or loss of households and jobs calculation), household and employment mobility (which predicts the movement of households and jobs), household and employment location choice (the location choices of households and jobs), real estate development (developer choices of what type of development to build and where) and land price (the prices of land at each location) (Noth et al., 2000; Waddell, 2002). UrbanSim's *spatial* urban simulation is carried out through *the household and employment location choice model* using techniques of discrete choice model (multinomial logit) and a Monte Carlo sampling process (Waddell, 2000a; 2006; Waddell and Ulfarsson, 2002a; 2002b). Each model is designed to work in a disaggregated approach at a household level and usually schedules to predict results on a yearly basis (Noth et al., 2000).

UrbanSim, similar to the classical CA approach, builds urban simulation on a grid-cell basis and from the bottom-up process, however, it differs in terms of the rules and factors set for urban development. While CA mostly accounts for physical and accessibility factors (e.g. development density, effect of roads, topography) and neighbourhood effects to simulate the urban development, UrbanSim examines the urban development process as dynamic interactions among key actors, notably consumer demands, land development processes, the role of land use planning, land regulations and environmental constraints (Noth et al., 2000). Factors input to the model heavily involve detailed social and economic factors, underlying the econometric concepts (Waddell et al., 2001). For example, the household (residential) location choice model (Waddell, 2006; Waddell et al., 2001) simulate the residential location choice for each classified household group. Each household group is classified and stratified by income, age of household, ethnicity, and number of car owned. Inputs to the household model include several factors notably housing characteristics (e.g. price, housing age), regional accessibility (e.g. travel time to CBD, job accessibility stratified by travel mode) and local accessibility (e.g. neighbourhood employment, neighbourhood land use mix and density) (Waddell, 2006; Waddell et al., 2001).

UrbanSim has several prominent strengths. Firstly, the model, like CA, has dynamic behaviour (Noth et al., 2000). As a result, it makes the simulation outcomes more transparent and explainable to users and decision makers, especially the complexity of the urban process in the context of econometric concepts (Agarwal et al., 2000). Secondly, the user of UrbanSim, can input data simply through an XML file, allowing the public policy issues, such as comprehensive plans, development restrictions, and environmentally sensitive lands, to be integrated and set explicitly as various scenario files (Noth et al., 2000). This is considered very useful since it enables users to evaluate the policy impacts as these constraints change with policy testing. As a result, the model can create realistic urban simulation outcomes with different sets of scenarios, making it more effective in operational use. Thirdly, UrbanSim is open source. Its software, pilot data, usage, and relevant documentations are freely available for use and for further development (Noth et al., 2000). Finally, since each model is encoded in a disaggregated approach, this enables users to run and customize each model individually such as adding new variables and constants, with no effect to other models (Teerarojanarat et al., 2004).

However, UrbanSim has some major drawbacks. Firstly, it has high data requirements (Agarwal et al., 2000). Each model needs highly detailed data, which in reality is rarely available, especially in the developing nations (Barredo et al., 2004; Bishop et al., 2000). Secondly, from the geographic point of view, UrbanSim has been criticized for not focusing enough on the spatial aspects (Benenson and Torrens, 2004). However, such 'micro' terminology is not relative to a geographical unit used. For example, data about an industrial group, in spite of being disaggregated by type of industrial activity, is considered likely to be governed over a very coarse or large spatial geographical unit. As a result, the simulation results created are probably not the production of the analysis of the smallest possible spatial unit. Thirdly, since UrbanSim has not been built based on the spatial context, it lacks some main spatial analytical function (e.g. neighbourhood effect and accessibility for each iteration) which are crucially important for dynamic simulation. Until now, GIS have been linked to UrbanSim for the purpose of map preparation and display only (Teerarojanarat et al., 2004; Waddell, 2000b; 2002), leaving the analytical part to be processed within UrbanSim itself. Finally, it has been newly developed, so it has limited experience to be applicable to other regions, especially outside the United States. Until now, there is no evidence to apply UrbanSim for the developing nations, possibly because of the high data requirements.

2.6 Extension to GIS Functionalities: Spatial Multi-criteria Decision Analysis

Until now modelling urban land use development, especially when planning policy applications on the basis of several points of views (e.g. physical, socio-economical, environmental views) are incorporated, has been of considerable interest for geographers, urban planners and urban decision-makers (Engelen et al., 1997a; Ward et al., 2000; Wu, 1998; 2002b; Wu and Webster, 2000; Yeh and Li, 2002). Despite the fact that GIS have proven to be widely used and integrated in many areas of urban and land use applications (e.g. urban simulation (Wu, 1998), site selection (Carver, 1991), land suitability (Joerin et al., 2001)), many researchers (e.g. Bojorquez-tapia et al. (2001), Chakhar (2003), Malczewski (1999a; 1999b; 2006), Merwe and Hendrik (1997), Joerin and Musy (2000), Wu (1998)), have agreed that GIS capabilities have limited functions in tackling spatial decision-making. Some critical restrictions includes the limited functionality of standard GIS overlay techniques (Janssen and Rietveld, 1990). They stated that when the number of variables (so-called layers) increases, GIS become difficult to comprehend and manage. Further, standard overlay methods consider that all variables or layers are of equal importance (Janssen and Rietveld, 1990). A second restriction is that spatial analytical functionalities of most GIS packages lie largely in the ability to perform deterministic overlay and buffer operations, and they are of limited use when involving conflicting objectives and multiple criteria (Carver, 1991; Malczewski, 1999a; 1999b).

One reliable and widely used approach to enhance the spatial decision-making capabilities of GIS is through its integration with multi-criteria decision analysis (MCDA) (Makropoulos and Butler, 2006; Malczewski, 1999a; 1999b; 2006). Multi-criteria decision analysis (MCDA), or interchangeably, multi-criteria decision making (MCDM) provides a wide range of techniques and procedures that allows a decision maker's preference to be incorporated (Malczewski, 1999a; 1999b). Integration between the two approaches, so-called GIS-based multi-criteria decision analysis (GIS-MCDA) (Malczewski, 2006), is considered useful in remedying conflicting situations both for individuals and groups involved in spatial decision-making context (Malczewski, 1999a; 1999b; 2006).

Multi-criteria decision analysis (MCDA) techniques emerged during the early 1970s as a new method for policy analysis (Zeleny, 1982). MCDA is considered a prescriptive model

(Malczewski, 1999a), which extends the capabilities of conventional evaluation programming such as the optimization method described in Section 2.2.3, in order to handle conflicting objectives and multiple criteria (Laaribi et al., 1996). MCDA mainly differs from the conventional optimization model in the latter searchs for the optimal solution, while MCDA solves conflicting problems (e.g. conflicting objectives) in order to provide one or more satisfactory solutions that represent true compromises in which optimal solutions are not necessary (Janssen and Rietveld, 1990; Nijkamp et al., 1990). Reviews of the MCDA literature can be found in Chakhar (2003), Laaribi et al.(1996), Malczewski (1999a; 2006), Voogd (1983), and Zeleny (1982).

Based on the main components of evaluation criteria, MCDA can be broken down into two general categories: (1) multi-objective decision analysis (MODA) (Malczewski, 2006) and (2) multi-attribute decision analysis (MADA) (Malczewski, 2006), which is also known as Multi-criteria Decision Making (MCE) (Malczewski, 1999a). The distinction between these two approaches relates to the question of whether objectives or attributes are the focus of the problem (Malczewski, 1999a; 1999b). The former deals with the selection of the best alternative (or choice) on the basis of a series of conflicting objectives (Malczewski, 1999a; 1999b). This technique is employed when the decision-making process involves fulfilling more than one objective (Eastman, 1999). The latter focuses on one objective having multiple attributes or criteria (Malczewski, 1999a). MADA is assumed to have a finite, usually limited, number of pre-specified alternatives and to select the best alternative on the basis of the scores of multiple attributes (Malczewski, 2006).

Conventionally, MCDA techniques do not explicitly consider a spatial (Malczewski, 1999a; 1999b; Villa et al., 1996). From the spatial point of view, they have a weakness in the sense that they are principally based on the assumption of spatial homogeneity (Malczewski, 1999a) such that impacts apply to the whole study area (Tkach and Simonvic, 1997). This assumption is considered impractical in several decision situations owing to the fact that in certain real-world situations the evaluation criteria varies across space (Malczewski, 1999a). Therefore, by linking MCDA and GIS, an improved system can be developed. This benefits from an explicit representation of the geographical dimension (op. cit.)

The GIS-MCDA approach is regarded as an analytical process that combines and transforms geographical data (input), either spatial or aspatial or both, to gain a result (output) derived from the decision-making (Malczewski, 1999a). Its procedure (Malczewski, 1999a; 1999b) involves evaluation of a set of geographically defined alternatives (events). Such alternatives are chosen based on their ranking performed on the basis of the criteria values and the decision maker's preferences with respect to a set of evaluation criteria. Consequently, the resultant information generated is the combination of the geographic distribution of the events (attributes) and the value judgments employed in the decision-making process (Malczewski, 1999a). When GIS and MCDA approaches are integrated, GIS are used to identify a set of feasible alternatives and to derive, either implicitly or explicitly, spatial attributes, whereas MCDA allows decision-makers to rank alternatives with respect to weighted criteria (Malczewski, 1999a; 1999b).

Incorporating GIS with MADA, so-called GIS-MADA (Malczewski, 2006), is considered an efficient and powerful tool for the analysis of land use evaluation and analysis tasks (Malczewski, 2006, Sener et al., 2006) as it allows not only the use of physical factors but also expert judgments, preferences (e.g. political or environmental views) to be taken into account in the decision-making process (Malczewski, 2006; Sener et al., 2006; Voogd, 1983). For example, the work conducted by Banai (2005) uses GIS-MADA as planning decision-making method by accounting for multiple indicators or criteria of sustainability in planning policies for the creation of land resource sustainability for urban land development both in the short and long term plan. In his work, multiple spatial criteria, including density variation, nodal activity location, land use mix, multimodal access, and open space/natural resource preservation, were first determined as a combination of relative weights (rating scales) and then were used in the assessment of the sustainability of land-use.

Incorporating GIS with MODA (so-called GIS-MODA (Malczewski, 2006)) is less popular than GIS-MADA (Malczewski, 2006). It was criticized by Malczewski (1999a) because it has restrictions due to the limited capabilities of GIS. He explained that MODA techniques require complex mathematical programming to handle an infinite set of alternatives which is implicitly generated within the spatial context. Such a requirement is difficult to carry out using GIS functionalities and its structure. As a result, incorporating GIS with MODA, most often, is done using a loose-coupling strategy. That is, solving the decision problems and conflicting objectives is usually incorporated using mathematical programming software such as What's Best! or LINDO outside the GIS environment while GIS are mainly used to serve as a tool for data preparation and visualization (Malczewski, 1999a). For example, the work conducted by Merwe and Hendrik (1997) on the urban edge of Cape Town, South Africa applied GIS-MADA using the IDRISI GIS software package on the basis of multi-attribute to evaluate development suitability for four land use types, including agriculture, small-farming, natural conservation and urban. The four suitability images created were then subjected to multi-objective decision-making using the IDRISI MOLAR module to produce an optimum land use allocation map.

GIS-MCDA applications have been shown to aid the evaluation of alternative policy decisions and what-if scenarios and has been applied in numerous fields (Malczewski, 2006). For example, an early application of Carver (1991) incorporated GIS with MCDA as a means to evaluate different alternatives based on multiple and conflicting criteria for the purpose of nuclear waste site selection. In the forest management domain, (Phua and Minowa (2005) applied GIS-MCDA to handle forest conservation planning with various decision-making bodies in the area of Kinabalu, Malaysia. The work conducted by Sener et al. (2006) applied GIS-MCDA in the area of Ankara, Turkey by accounting for multiple attributes to be incorporated for waste disposal site selection.

Classified by Joerin et al. (2001), applications of GIS-MCDA for land use analysis have been carried out for other areas in addition to site selection mentioned above, land suitability (Banai, 2005; Joerin et al., 2001) and collaborative decision support systems (Joerin and Musy, 2000) are applications. All follow GIS-MCDA approach are given at the end of Section 2.4.2.1. Their procedure, following the GIS-MCDA framework, can be structured as the sequence of five stages – identifying problem, defining the set of evaluation criteria, generating alternatives, setting criterion weights and selecting a decision rule (Malczewski, 1999a; 1999b). Briefly speaking, decision problems are first identified. An objective or a set of objectives are then determined and decomposed into one or more sets of decision criteria and related criteria (Feick and Hall, 2004; Malczewski, 1999b). Each criterion represents the characteristics of a given set of objectives with respect to the problem situation while the associated weight indicates the relative importance gained from the decision participants (Malczewski, 1999a; 1999b). By altering the criteria sets and the criteria weights, participants represent their own unique preference and perspective in the evaluation of decision alternatives (Feick and Hall, 2004; Malczewski, 1999a; 1999b)

Usually, each criterion needs to be aggregated into a normalized format ranging from 0 to 1 so that each is allowed to be comparable to other criteria (Malczewski, 1999a; Voogd, 1983). There are a wide variety of practical criterion-weighting methods applied for GIS-MCDA such as ranking, rating, seven-point scale, and pairwise comparison (Malczewski, 1999a; Voogd, 1983). Each method has pros and cons. For example, the ordinal ranking method is the simplest method for assessing the importance of weights in rank order and suitable for valuing preferences (e.g. the most important is assigned to 1, second important to 2, third important to 3) (Malczewski, 1999a; 1999b). However, a drawback associated with this method is that it is significantly limited by number of criteria used (Hajkowicz and McDonald, 2000; Voogd, 1983). The popular pairwise comparison method, developed by Saaty and Alexander (1981), allows only two criteria to be considered at the same time. It thus provides a meaningful way to determine the relative importance (Banai, 2005). The pairwise comparison method has been proven theoretically and empirically for various spatial decision-making applications (Malczewski, 2006) such as the work of Banai (2005), Marinoni (2004), and Sener et al.(2006). More importantly, this method has been reported to be incorporated with GIS decision-making directly such as in the IDRISI GIS software package (Ceballos-Silva and Lopez-Blanco, 2003; Cromley and Hanink, 1999; Merwe and Hendrik, 1997). However, this method is only appropriate when the number of criteria involved is less than eight (Wu, 1998).

The final stage in implementing spatial MCDA is to apply a decision rule to gain the decision output. A decision rule is a procedure that allows for ordering or ranking alternative inputs in the descending order of preference (Malczewski, 1999a). It helps identify how best to rank alternatives or to make decisions about which alternative is more preferable than another (Malczewski, 1999a; 1999b). There is a rich collection of decision rules that can be incorporated into GIS-MCDA. They include simple additive weighting (SAW), the value/utility function methods, the analytic hierarchical process (AHP), the ideal point methods, the outranking methods and the ordered weighted average (OWA) (Malczewski, 1999a). SAW, a so-called weighted linear combination (WLC) or scoring method is the most widely used technique. For example, the work of Sener et al. (2006) applied the concept of average weight, which can then be operated easily through both raster and vector format by using any GIS system having conventional overlay capabilities. However, a disadvantage associated with this method is that since a function of SAW is based on the assumption of linearity and additivity, it is considered too simple to apply in many real-world problems and lacks theoretical foundations (Sener et al., 2006). Another widely used method is AHP (Marinoni, 2004; Sener et al., 2006). With this method, a decision problem is decomposed into simple decision problems to form a hierarchical level (Sener et al., 2006). Its main advantage is that it provides ease during decision making as it allows the users to categorize the relationship between criteria or factors in a hierarchical manner (Sener et al., 2006, Malczewski, 1999a).

Integrating between GIS and MCDA has advantages. While GIS provide a powerful tool for handling spatial aspects, the MCDA procedure provides the means to tackle conflicting objectives and multiple criteria (Carver, 1991; Chakhar, 2003). MCDA helps decision-makers to understand and clarify the problem by forcing them to analyze and justify their opinions, priorities, and preferences in evaluating competing alternatives or choice possibilities (Feick and Hall, 2004; Villa et al., 1996). As a result, such integration aids in the evaluation of alternative policy decisions and the evaluation of what-if scenarios, especially in the spatial context (Malczewski, 2006). However, like other prescriptive models where decision problems are based on the judgment of individuals, the model can be considered unsystematic (Lee, 1973). Ill-defined objectives and too many criteria can lead to an ineffective output (op. cit.). Another concern involves uncertainty as a result of the varying in the criteria sets, the criteria weights and decision rules used (Voogd, 1983). The MCDA approach is highly dependent on the criteria used and the weights assigned. As different techniques and weights are applied, they result in differing output creation (Hajkowicz and McDonald, 2000; Malczewski, 1999a; Sener et al., 2006).

2.7 Conclusion

In this chapter, a review of the spatial modelling approaches (CA, GIS, and MCDA) that have been used to analyze urban system has been discussed and criticized, including advantages and disadvantages of each approach. In addition, the issues concerning data and relevant information for urban simulation have also been discussed, specifically within the context of the developing nations. The review has been made in order to select a suitable approach for modelling the spatial pattern of urban growth in the context of cities in developing countries, such as Bangkok, Thailand, where accurate, up-to-date, and historical information especially in a digital format is often difficult to obtain. Based on the review performed, a cellular automata (CA) approach and GIS will be applied in the rest of this study.

There are three reasons for selecting such integration. Firstly, the CA model is dynamic. It is developed to simulate the historical urban spatial pattern and predict future spatial growth over time. Such simulation is helpful in investigating areas having potential to experience change in urban development, and as a result it can help planners to manage and monitor urban development. Secondly, CA can be used to simulate the spatial pattern of urban growth using different planning scenarios, which allow planners to evaluate the implications of different planning programs. Thirdly, CA can be carried out by using simple rules. Such rules, with integration of GIS, can be modified and account for several variables that can represent and capture realistic geographical problems. Further, the simulated results obtained from the model are in the form of standard digital GIS maps, which allows planners to understand, share, and overlay with other GIS map layers for further analysis.

In this research study, the CA model will be integrated with GIS in order to investigate the spatial pattern of urban land use development in Ladprao, one of the district of Bangkok, Thailand. The integration will be conducted in a tight-coupling method, where CA will be programmed in VBA within the ArcGIS software package. A Multi-criteria decision analysis approach and statistical multinomial logistic regression technique will be incorporated as rules in CA.

CHAPTER 3

Study Area, Data Sources and Data Constraints

3.1 Study Area: Ladprao District, Bangkok, Thailand

The study area is Ladprao district, situated in the central-eastern part of Bangkok. Figure 3.1 shows the location of Ladprao, an area covering nearly 22 sq. km. Ladprao is one of fifty Bangkok Metropolitan Administration (BMA) units (Core Planning & Development, 2001). A number of criteria were first assigned to select a representative case study for simulating future land use changes. The area had to:

- Be in a transition zone, where adequate vacant land is left available for horizontal development during the specified period considered.
- Have digital land use data available with different land use classifications including vacant (undeveloped), residential, commercial and industrial use.
- Have the availability of digital land use data from at least two epochs to be able to detect changes and for model assessment purposes.

Ladprao was accordingly chosen for the case study. Firstly, Ladprao is classified by BMA as one of the districts being in a "transition zone" area. The area, after spatial observation, shows residential growth during the two periods studied (1993 and 2001). In 1993, Ladprao (measured from the vector map of the study site at scale 1:4,000) contained about 51 percent vacant land area (11 sq.km. out of 21.34 sq.km. total area) As such, this area stands out as one with potential for transition to urban development. The second and third requirements outlined above are due to the need for data availability. Data availability is an important issue in spatial modelling, as data input and database creation is costly and time consuming (Bishop et al., 2000; Walford, 2001). In Thailand, like many other developing countries, few useful spatial datasets exist in a digital format (Bishop et al., 2000). In the case of Bangkok, Department of City Planning, BMA is responsible for collecting and updating land use data for the Bangkok area. Unfortunately, only the current (2001) vector land use map, classified at the level that included land use type considered, at the scale of 1:4,000. The dataset includes a digital map of roads, rivers, and administrative boundaries. Land price data, during from the periods being studied is available in text format. This was later converted to a grid layer map of the same scale as the land use layer. Additionally, historical data archived in the form of aerial photographs (1: 6,000) in 1993 is available for this area. This dataset has been manually interpreted to derive land use/land cover information for 1993. The 1993 data is used as the base or initial information to obtain simulation results for a given year beyond the initial year while the actual 2001 data is for validation and assessment purposes. The availability of data from these two periods (1993 and 2001) allows change to be observed and their resultant simulation can be compared and assessed in a spatial context. This will be primarily used in Chapter 4, 5, and 6 of this thesis.

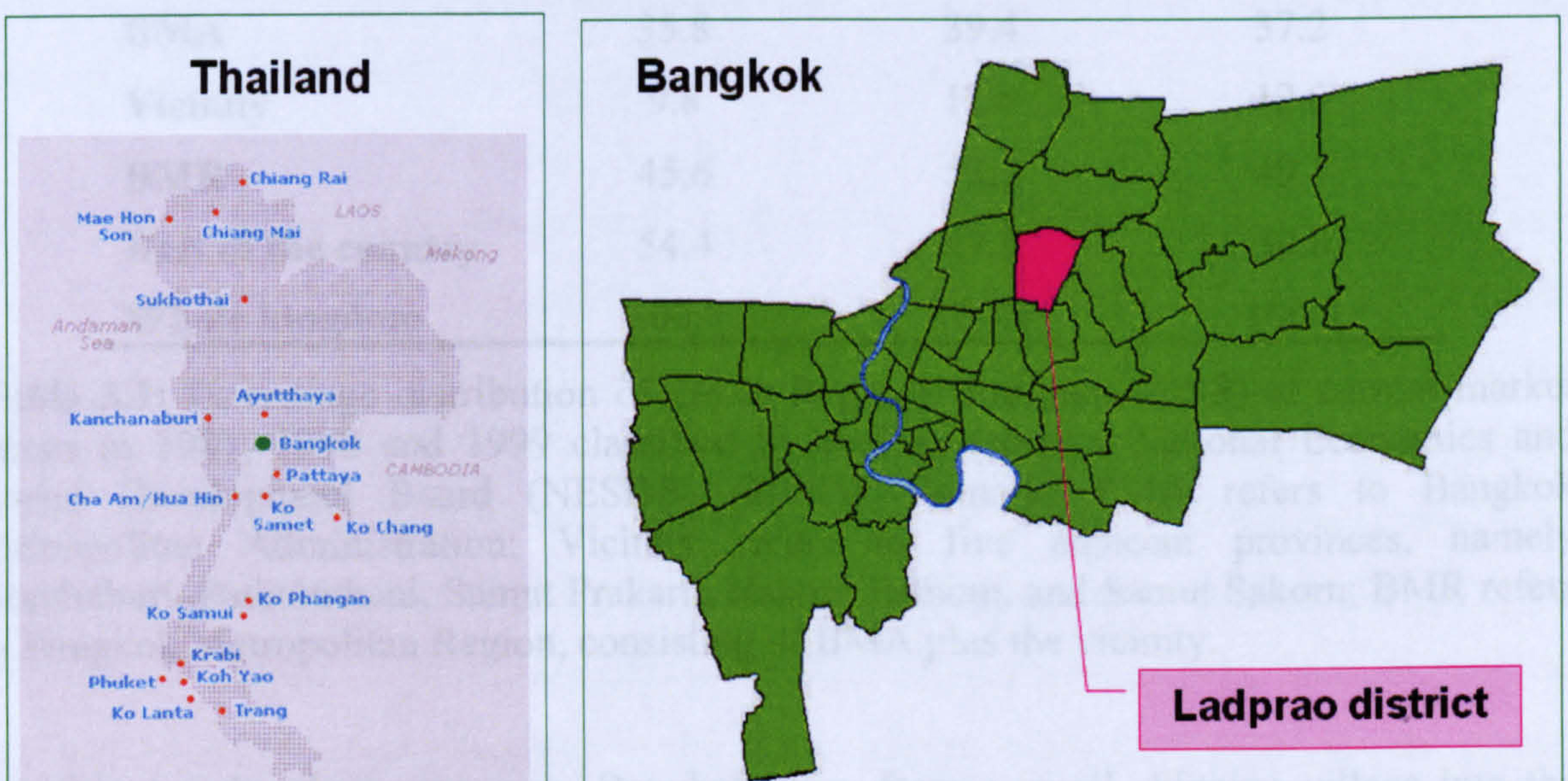


Figure 3.1: Bangkok and study site (Ladprao district).

3.1.1 The Bangkok Context

3.1.1.1 Historical Growth of Bangkok

Bangkok, Thailand's capital, was founded in 1767 and has since developed into the main administrative, economic and cultural center of the country (Sharkawy and Chotipanich, 1998). At present, the city encompasses 50 districts, covering an area of roughly 1,569 square kilometers (Department of City Planning (DCP), 1999). From the 1st to the 8th National Economic and Social Development Plans (1961 – 2001), the focal attention of the country has been on industrial development (Chomchan et al., 1990). Nowadays Bangkok is known as one of the economically fastest growing areas in South East Asia (Webster, 2004). According to NESDB's statistics shown in Table 3.1, almost half of the gross regional product (GRP) of Thailand belonged to BMA and its vicinity. Thailand experienced a high level of economic growth during the 1987 – 1995 period with an average economic growth of 10% per annum. However, the economic boom ended in 1997, when Thailand was confronted with an economic recession due to the Asian economic crisis that broke out that year (Boonchuen, 2002; Sharkawy and Chotipanich, 1998).

Regions	GRP (Percent)		
	1981	**1990**	**1999**
BMA	35.8	39.4	37.2
Vicinity	9.8	12.8	12.0
BMR	45.6	52.2	49.1
Rest of the country	54.4	47.8	50.8
Whole kingdom	100.0	100.0	100.0

Table 3.1: Percentage distribution of Gross Regional Products (GRP) at current market prices in 1981, 1990 and 1999 classified by regions (source: National Economics and Social Development Board (NESDB) (2001)). Remark: BMA refers to Bangkok Metropolitan Administration; Vicinity refers to five adjacent provinces, namely Nonthaburi, Pathumthani, Samut Prakarn, Nakorn Pathom, and Samut Sakorn; BMR refers to Bangkok Metropolitan Region, consisting of BMA plus the vicinity.

Roughly two hundred years ago, Bangkok grew from a small shipping village into the capital of the country (Chomchan et al., 1990). Back in those days, the Thai economy was agriculture-based with the settlements spreading out along both sides of the Chao Phraya river and several nearby connecting canals (Boonchuen, 2002). People usually depended on canals and Chao Phraya river as major forms of transportation. The development of the Thai economy through international market forces in accordance with the implementation of the 5th and the 6th National Economic and Social Development Plans (1982 – 1991) resulted in an economic boom and shift of the economic structure from one based on agriculture to a more industry and service-oriented (Chomchan et al., 1990; Kaothien, 1995; Webster, 2004). Unavoidably this shift brought an influx of young migrants from rural areas throughout the country to Bangkok in search of better job opportunities (Chomchan et al., 1990). Thai society has increasingly changed from rural to urban, with the trend growing over time (Boonchuen, 2002; Chomchan et al., 1990). Bangkok since then has been characterized by its central role in terms of administration, trade, services, education, health, transportation, and job opportunities, etc (Chomchan et al., 1990).

Inevitably, such growth has resulted in increasing the size of the metropolis and its administrative boundaries (Chomchan et al., 1990) and changing patterns of metropolis development (Chomchan et al., 1990; Webster, 2004). In 1960, the boundaries of the Bangkok Metropolitan Area were expanded from 96.4 square kilometers to 189.7 square kilometers in 1970, and to 1568.7 sq.km. in 1975 (still the area today). The number of districts was increased from 14 districts in 1960 to 50 districts at the present time (Department of City Planning (DCP), 1999). The growing pattern of metropolis enlargement can be characterized as the accelerating expansion of the inner city districts, outer city districts, and horizontal enlargement of the metropolis (so-called peri-urban areas) (Webster, 2004).

The urban structure in the inner city district is characterized by densely populated areas with little vacant land (Webster, 2004). Because of very limited supply of land, land prices are very high (Chomchan et al., 1990; Webster, 2004). Consequently, some areas of the inner city experienced declines in residential population as many, especially of the middle class, move outward from the core city to the suburbs in search of the better amenities (Webster, 2004). Most of the land has already been occupied for commercial purposes, government offices, educational establishments, transportation infrastructure such as elevated freeways and recently a sky-train and a subway system, in an overall environmental of mixed land use (Kaothien, 1995; Webster, 2004). Most of land utilization in the inner city has remained static since it is already occupied and has reached a saturation point (Chomchan et al., 1990). The land usage in this area can be illustrated as vertical development in the form of high-rise offices and high-rise residential buildings (op. cit.). A subway system, commencing service in early 2004, has had a profound impact on the urban structure outside the core, with the increasing number of new high-density multiple-use complexes occurring around or directly linked to key transit stations (Webster, 2004).

In contrast, the outer city districts or Bangkok's suburbs are dominated by horizontal development of two different types (Chomchan et al., 1990; Webster, 2004). One is residential – of suburban villages, so-called *mubans* in Thai, mainly developed by private housing development projects. *Mubans* can vary from "no frills" row housing developments to luxurious communities. People once traditionally settled along the banks of the canals whereas now they tend to concentrate along the major forms of road transportation, leaving the area behind (away from transport access) to either be used for agricultural purposes or be left idle. The other type of development is characterized by (increasingly uncompetitive or closing) labour-intensive factories. Interspersed can be found slum communities, in particular in low-lying areas, along canals and rail lines, and near factories. Another observable artifact of suburban areas is the regional, giant shopping malls that have a role as shopping and service centers within the urban fringe. In the past, the outer region was utilized for agricultural purposes with a portion of the land being designated as a green-belt zone (Chomchan et al., 1990). Green-belt area refers to vacant land or open space reserved for agriculture purposes, public parks, or natural areas such as swamps or woods. Traditionally, land in the eastern suburbs of the Bangkok metropolis has been cultivated as rice paddies where the western suburbs of the city have been dominated by vegetable and fruit gardens. However, due to ineffective land preservation measures, many exemptions and allowances have been provided for individual cases and also a large number of illegal buildings have been built over time (op. cit.). In addition, high land prices cause the returns from agricultural activities to be very low in comparison to other economic activities (op. cit.).

Peri-urban areas, in the Bangkok context, refer to the areas beyond suburban areas where industrialization is occurring speedily, with the coexistence of agriculture and other rural activities (Webster, 2004). During the last two decades, Bangkok extended speedily without planning. Bangkok Metropolitan Area (BMA) has outgrown its administrative boundary following major highways into the surrounding provinces of Nonthaburi, Pathumthani, Samut Prakarn, Samut Sakorn, and Nakorn Pathom and formed a region known as the Bangkok Metropolitan Region (BMR), covering the area of 7,639 sq. km. (McGee, 1994; Sharkawy and Chotipanich, 1998) as shown in Figure 3.2. Most lands in the peri-urban areas are not built up yet. Rather, it is likely to remain a mixture of rural and urban. One dominant characteristic of this area is represented by industrial estates (Webster, 2004). The pattern of development is, however, primarily located along major highways and roads, especially along extensive toll freeway systems (op. cit.). In the last several decades, BMR has spread out further to the next five adjacent provinces of Ayutthaya and Saraburi to the north, Ratchaburi and Phetchaburi to the west, and all along the coast to the east including the provinces of Chon Buri, Chachoengsao, and Rayong, making up the so-called the Extended Bangkok Metropolitan Region (EBMR) (Chomchan et al., 1990; McGee, 1994; Sharkawy and Chotipanich, 1998).

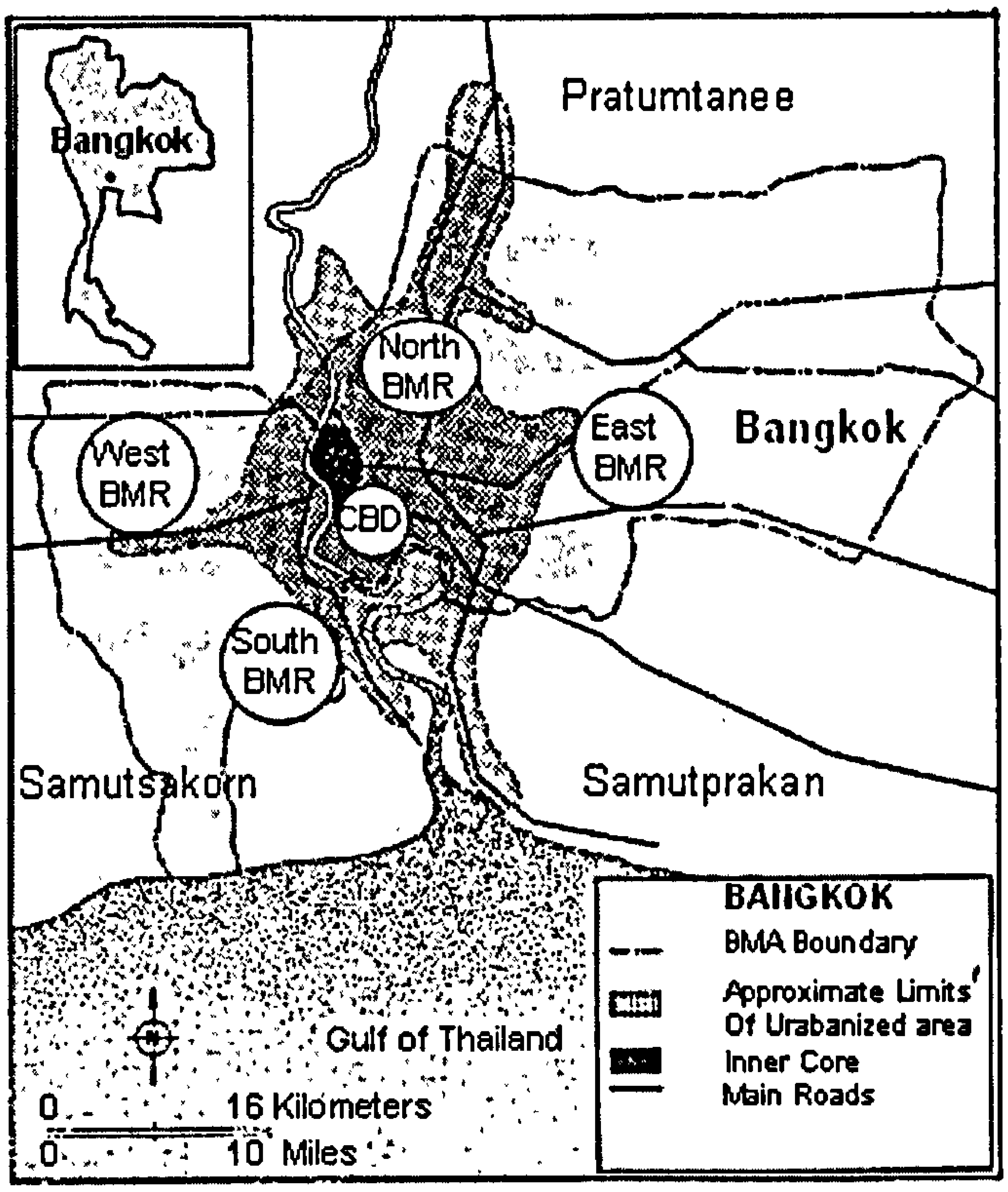


Figure 3.2: Bangkok Metropolitan Area (BMA) and its vicinity. (source: Sharkawy and Chotipanich (1998), p.32).

3.1.1.2 Demography

Population growth in Bangkok is both the result of natural population increases and, more importantly, the large number of young migrants from other provinces throughout the country moving to the Bangkok area, as a result of the economic boom during the 1970s and 1980s (Choiejit and Teungfung, 2005; Chomchan et al., 1990). Table 3.2 illustrates the population and annual growth rate between 1960 – 2000 in the area of Bangkok, its vicinity and the whole kingdom of Thailand.

	Population (in millions)					Annual growth rate (%)			
	1960	1970	1980	1990	2000	1960-1970	1970-1980	1980-1990	1990-2000
BMA	2.1	3.1	4.7	5.9	6.3	3.89	4.16	2.27	0.66
Vicinity	1.2	1.4	1.9	2.7	3.8	3.48	3.05	3.51	3.42
BMR	3.3	4.8	6.6	8.6	10.1	3.75	3.18	2.65	1.61
Whole kingdom	26.3	34.4	44.8	54.6	60.9	2.70	2.65	1.96	1.05

Table 3.2: Population figures and annual growth rate in Bangkok, its vicinity and the whole country in 1960, 1970, 1980, 1990, and 2000 (Source: Population and Housing Census, 1960, 1970, 1980, 1990 and Preliminary Report of Population and Housing Census, 2000, National Statistical Office, NSO.) Remark: BMA refers to Bangkok Metropolitan Administration; Vicinity refers to five adjacent provinces, namely Nonthaburi, Pathumthani, Samut Prakarn, Nakorn Pathom, and Samut Sakorn; BMR refers to Bangkok Metropolitan Region, consisting of BMA plus the vicinity.

According to the figures (Table 3.2), the population of Bangkok area has increased tremendously from 2.1 million in 1960 to 3.1, 4.7 and 6.3 million in 1970, 1980 and 2000 respectively with the annual growth rate being higher than 3 % per annum during 1960 – 1970, and rapidly declining during the last 20 years. The growth rate between 1980 – 1990 was approximately 2.27% per annum and it continued to decline to 0.66 % per annum during the 1990 – 2000 period. According to the latest figures (United Nations, 2006), the city contained roughly 6.6 million in 2005 and is projected to continue to grow to 7.4 million by the end of 2030. According the latest NSO official statistics of 1960 – 2000, the growth rate of BMA declined abruptly from roughly 3.9% to at roughly 0.6% while the Vicinity is growing considerably faster at about the same rate of 3.4% annually on average. It can be clearly seen, that the great number of net in-migrants (nearly 4 million people from 1960 – 2000) is an important factor contributing to the rapid growth of Bangkok's population. Rapid population growth not only occurred in Bangkok but also took place in five provinces surrounding Bangkok, including Nonthaburi, Pathumthani, Samut Prakarn, Nakorn Pathom, and Samut Sakorn, and the population growth rate of these surrounding provinces was higher than that of the core Bangkok area. In sum, of Thailand's present population of roughly 61 million inhabitants (the latest figures in 2000), approximately 17% live in the BMR, 10% in BMA and 6% in the vicinity (reffered to as five provinces surrounding BMA).

3.1.1.3 Bangkok's Land Use Patterns and Problems

Through monitoring growth of the city by means of review of satellite imageries, as conducted by the BMA over the last decades, it was found that building bridges across the Chao Phraya River and the establishment of new roads are the major factors determining city development (Chomchan et al., 1990). However, because of insufficient provision of service roads to developers, particularly since the real estate boom in the 1960s, many areas in the city have experienced the phenomenon of so-called 'superblocks' such as the one covering the Ladprao area (see Figure 3.3). This phenomenon results in inefficient road networks (Ross et al., 2000; Sharkawy and Chotipanich, 1998; Webster, 2000).The road layouts are regular crossed by circuitous lanes. This has produced large areas with low accessibility, and a disconnected assemblage of roads. Many areas in the interior are inaccessible, considered 'blind land'. Such a phenomenon makes some areas within the superblocks difficult to access and connect with new systematic and more efficient road networks. Moreover, its impact, in part, has exacerbated a misuse of land and the expansion of the city by pushing residential development farther out of the city and into the suburbs (Webster, 2004).

Apart from a land use pattern brought on by the ineffective road network in the city, there are other distinct land use patterns that should be further considered, including a chaotic mix of land utilization, the expansion of suburban villages, the expansion and the direction of growth of the city.

Bangkok is located on a low-lying area of the Chao Phraya River delta approximately 1.5 meters above sea level (Department of City Planning (DCP), 1999) that is more suitable for rice farming than residential or industrial use (Chomchan et al., 1990). Rapid expansion has brought misuse of land, in a free-for-all pattern of land usage observable as a chaotic mix of residential houses, commercial buildings and factories at all scales, shapes, and sizes (Kaothien, 1995; Krongkaew, 1996; Webster, 2000). Industrial areas are found in a random fashion along the main roads while agriculture areas fill the interstices in between (Kaothien, 1995; Webster, 2004). The number of high-rise buildings, condominiums, shopping complexes and business offices considerably increased in the inner part of the city, however, without proper preparation and management of basic ground infrastructure such as frontage access and waste management systems (Kaothien, 1995; Krongkaew, 1996; Webster, 2004)

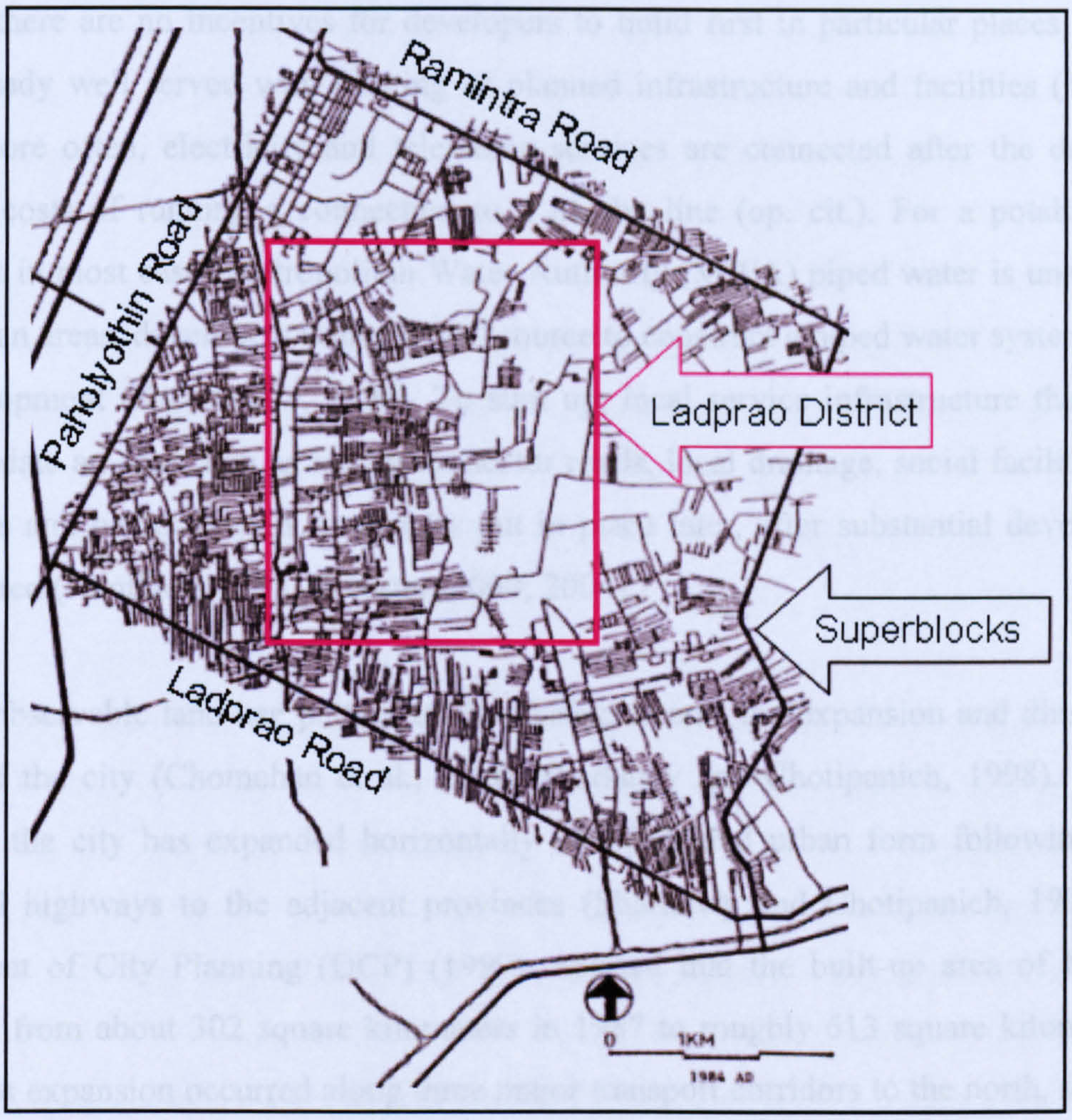


Figure 3.3: The 'superblocks' road network phenomenon (adapted from Archer (1996)).

Another distinguishable artifact is residential development growth in suburban areas (Chomchan et al., 1990; Webster, 2004). Such enlargement is largely the result of the real estate boom and the evolution of an active private housing development sector in the late 1960s (Sharkawy and Chotipanich, 1998; Webster, 2004). Prior to that time, housing was constructed on an individual basis on owners' plots (Sharkawy and Chotipanich, 1998). It was evident that in the 1970s developers accounted for 20% of new housing constructed (Webster, 2004) while in 1997 developers built about 80% of housing in Bangkok's suburbs. Residential development by private developers, of so-called *mubans* (suburban villages), has since then been the fastest-growing sector of the Bangkok housing market (Sharkawy and Chotipanich, 1998). The population of suburban Bangkok has grown at an accelerating rate since the 1970s (Webster, 2004). Despite such growth, suburban areas still have many areas of open land left as development tends to leap-frog in a linear ribbonlike pattern along the major highways (Webster, 2004). Patches of land can be found in between major developments held for speculative purposes. This situation creates a type of unplanned suburban development (op. cit.). Such a pattern is largely the reflection of the fact that there are no incentives for developers to build first in particular places, such as areas already well served with existing or planned infrastructure and facilities (Webster, 2000). More often, electricity and telephone services are connected after the developer pays the costs of running a connection to a nearby line (op. cit.). For a potable water supply, as in most cases Metropolitan Water Authority (MWA) piped water is unavailable in suburban areas, developers utilize a well source to construct a piped water system within the development (Krongkaew, 1996). To sum up, local service infrastructure that would accommodate an area such as feeder/connector roads, local drainage, social facilities such as schools and health centers, is usually put in place later, after substantial development has occurred (Kaothien, 1995; Webster, 2000; 2004).

Another observable land use pattern of Bangkok concerns the expansion and direction of growth of the city (Chomchan et al., 1990; Sharkawy and Chotipanich, 1998). Without planning, the city has expanded horizontally in a lopsided urban form following major roads and highways to the adjacent provinces (Sharkawy and Chotipanich, 1998). The Department of City Planning (DCP) (1999) reported that the built-up area of Bangkok expanded from about 302 square kilometers in 1987 to roughly 613 square kilometers in 1995. This expansion occurred along three major transport corridors to the north, southeast and southwest of the city respectively, inevitably causing degradation of agricultural areas (Chomchan et al., 1990; Sharkawy and Chotipanich, 1998). The BMR has grown farthest to the north because of the road connection between the airport (Don Muang airport) and the CBD (see Figure 3.2). However, a new Bangkok International Airport (Suvarnabhumi international airport) recently opened to the east in 2006 along with related improvements in infrastructure, a toll way, and direct access to the CBD, which is now magnetizing urban growth to the east. Growth to the west, despite a lack of vital existing infrastructure in both the southern and western parts of the BMR at the present time, is due to the completion of planned infrastructure to the west (Thong and Tsang (1990) cited in Sharkawy and Chotipanich (1998)).

3.1.1.4 Land Use Plan and the Role of City Planning in Bangkok

Apart from the fact that Bangkok's land use problems are a reflection of urban sprawl, there are some other key factors that need to be taken into account: lack of an effective land use plan, lack of enforcement of a plan, lack of detailed land planning, and lack of a cohesive planning policy (Choiejit and Teungfung, 2005; Chomchan et al., 1990; Kaothien, 1995; Krongkaew, 1996; Webster, 2000).

Of the many related land use plans, the most important is the Bangkok Comprehensive Plan, considered to be the major law controlling Bangkok's land utilization. The Bangkok Comprehensive Plan was developed under the Town Planning Act 1975 (BE 2518) and prepared by Department of City Planning, Bangkok Metropolitan Administration. Under this act plans have been released progressively, each valid for a period of five years. At present, the Bangkok Comprehensive Plan B.E. 2549, released in May 2006, is being used for the enforcement of the BMA's land use planning (Department of City Planning (DCP), 2002). The Bangkok Comprehensive Plan is divided into three plans, including the land use plan, the transportation system plan and the open space plan (Department of City Planning (DCP), 1999). The land use plan plays a major role as a guideline for control of Bangkok's land utilization (see Figure 3.4). The transportation system plan is for solving traffic problems. The open space plan is for recreation and environment conservation. To date, the Comprehensive Plan has been criticized as impractical (Chomchan et al., 1990; Kaothien, 1995). One of the major reasons is that the plan was built on the basis of an assemblage of outdated data from as far as back as 1981, and does not truly reflect the present state of economic and society. Another important reason is that the plan does not reflect the actual extent of rapid urbanization since it projects the future of the city as stable (Chomchan et al., 1990). As this strategic plan does not function very well and does not fit in with the current dynamic urban land market, coupled with the fact that the growth of the cities is often guided by market forces rather than planning policies, it has meant that the plan is often disregarded (Bishop et al., 2000; Kaothien, 1995).

Enforcement of the Comprehensive Plan and Specific Plan is considered weak since there is no obligation to obtain a *planning permit* for the development of land. (Commission of the European Communities (CEC) (1995) cited in Pimcharoen (2001)) This results in ineffective control of land use (Pimcharoen, 2001).

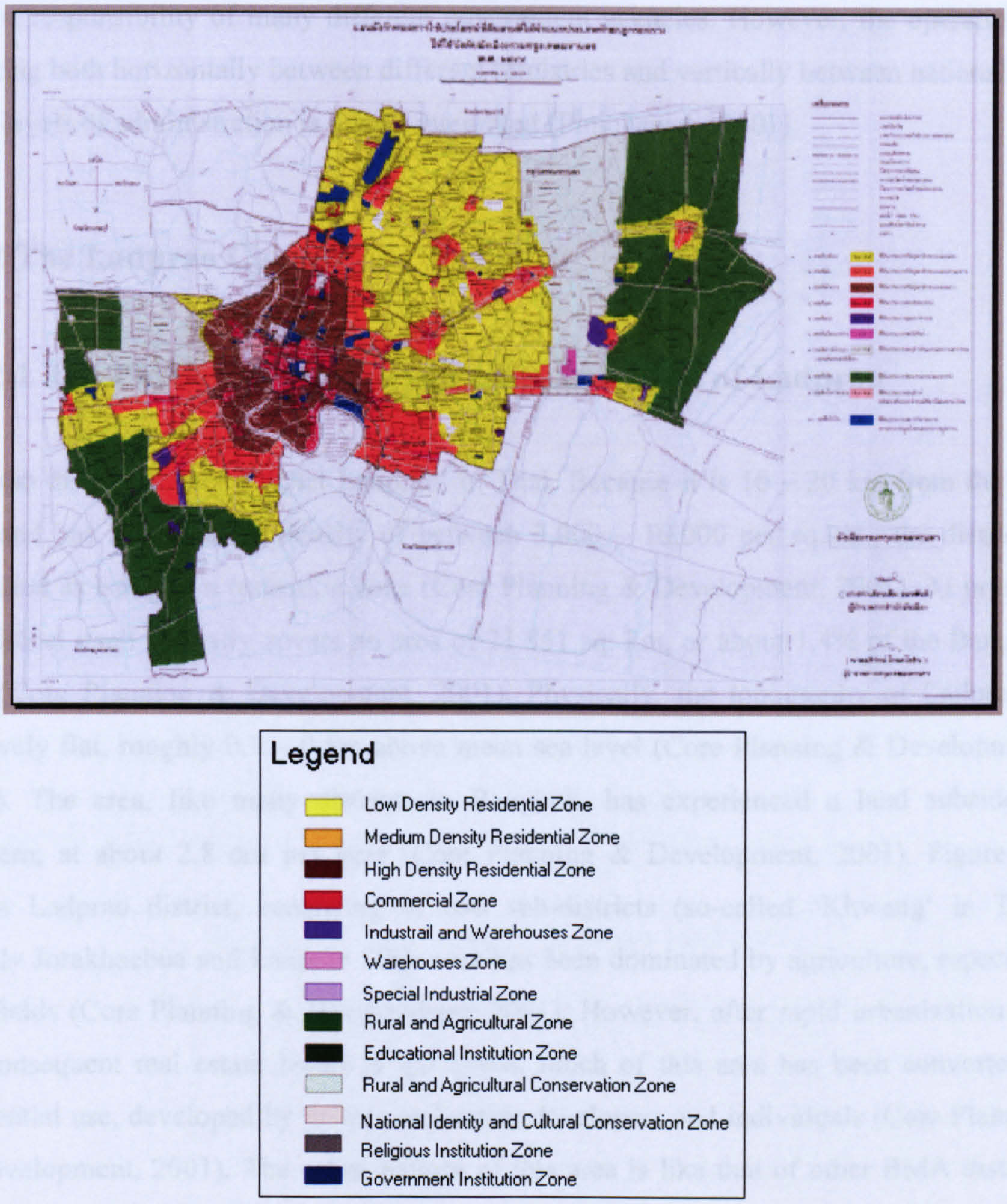


Figure 3.4: Land use plan, the Bangkok Comprehensive Plan (source: Department of City Planning (DCP) (2002)).

A third reason is that Bangkok nowadays, apart from the Comprehensive Plan which is used just for a broad view of development planning, still lacks land use development plans and enforcement at the district level, resulting in ineffective control of land use at this level (Kaothien, 1995).

The final reason is a lack of cohesion in implementation of planning policy. Like in many developing countries, laws and guidelines for planning and management are diverse and plan implementation is poorly integrated amongst local and federal authorities (Bishop et al., 2000; Cohen, 2004; Kaothien, 1995). In Bangkok, many related land-use planning laws are the responsibility of many different government agencies. However, the operation of planning both horizontally between different ministries and vertically between national and local levels of administration is poorly integrated (Pimcharoen, 2001).

3.1.2 The Ladprao Context

3.1.2.1 The Physical Characteristics and Land Use of Ladprao

Ladprao district is called 'Khet Ladprao' in Thai. Because it is 10 – 20 km from the city core and has a population density of between 3,000 – 10,000 per sq.km., the district is classified as being in a transition zone (Core Planning & Development, 2001). At present, the district itself formally covers an area of 21.851 sq. km, or about 1.4% of the Bangkok area (Core Planning & Development, 2001). Physically, the topography of Ladprao is relatively flat, roughly 0.3 – 0.5m above mean sea level (Core Planning & Development, 2001). The area, like many districts in Bangkok, has experienced a land subsidence problem, at about 2.8 cm per year (Core Planning & Development, 2001). Figure 3.5 shows Ladprao district, consisting of two sub-districts (so-called 'Khwang' in Thai) namely Jorakhaebua and Ladprao. This area has been dominated by agriculture, especially rice fields (Core Planning & Development, 2001). However, after rapid urbanization and the consequent real estate boom in the 1960s, much of this area has been converted to residential use, developed by private real estate developers and individuals (Core Planning & Development, 2001). The urban pattern of this area is like that of other BMA districts where urban activities, especially *muban* or village communities developed by private real estate developers, characterize the horizontal development running along the length of major and secondary roads, creating a back portion of land behind the villages and row houses that is inaccessible and with a very low price (Core Planning & Development, 2001; Sharkawy and Chotipanich, 1998).

According to the interpretation of land use / land cover based on aerial photographs taken in 1987, 1993, 1995 and 2000 conducted by Department of City Planning, the built-up areas of Ladprao gradually increased to 1995 and soared since then till 2000 as is shown in Figure 3.6.

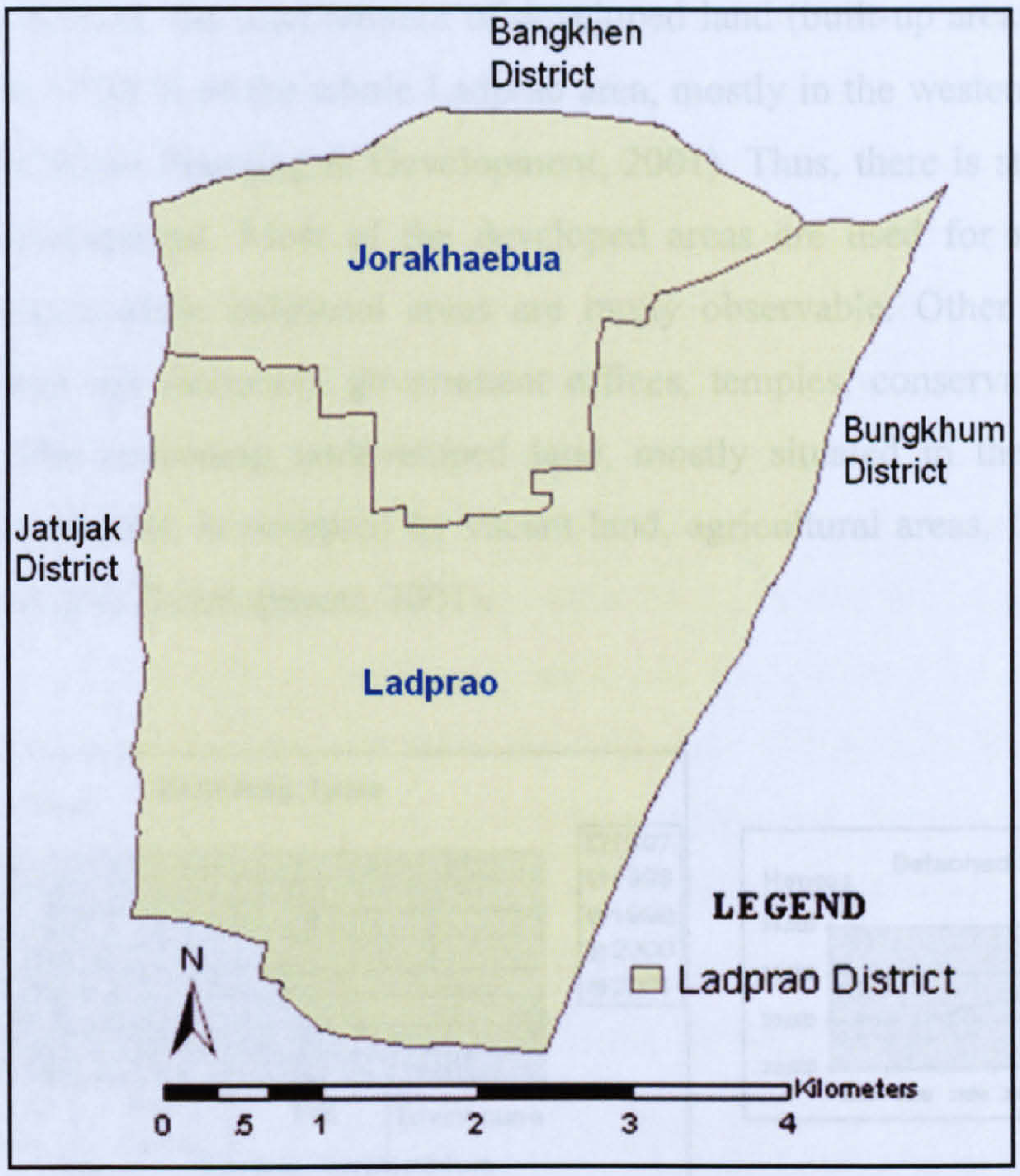


Figure 3.5: Administrative boundary of Ladprao district (Khet Ladprao).

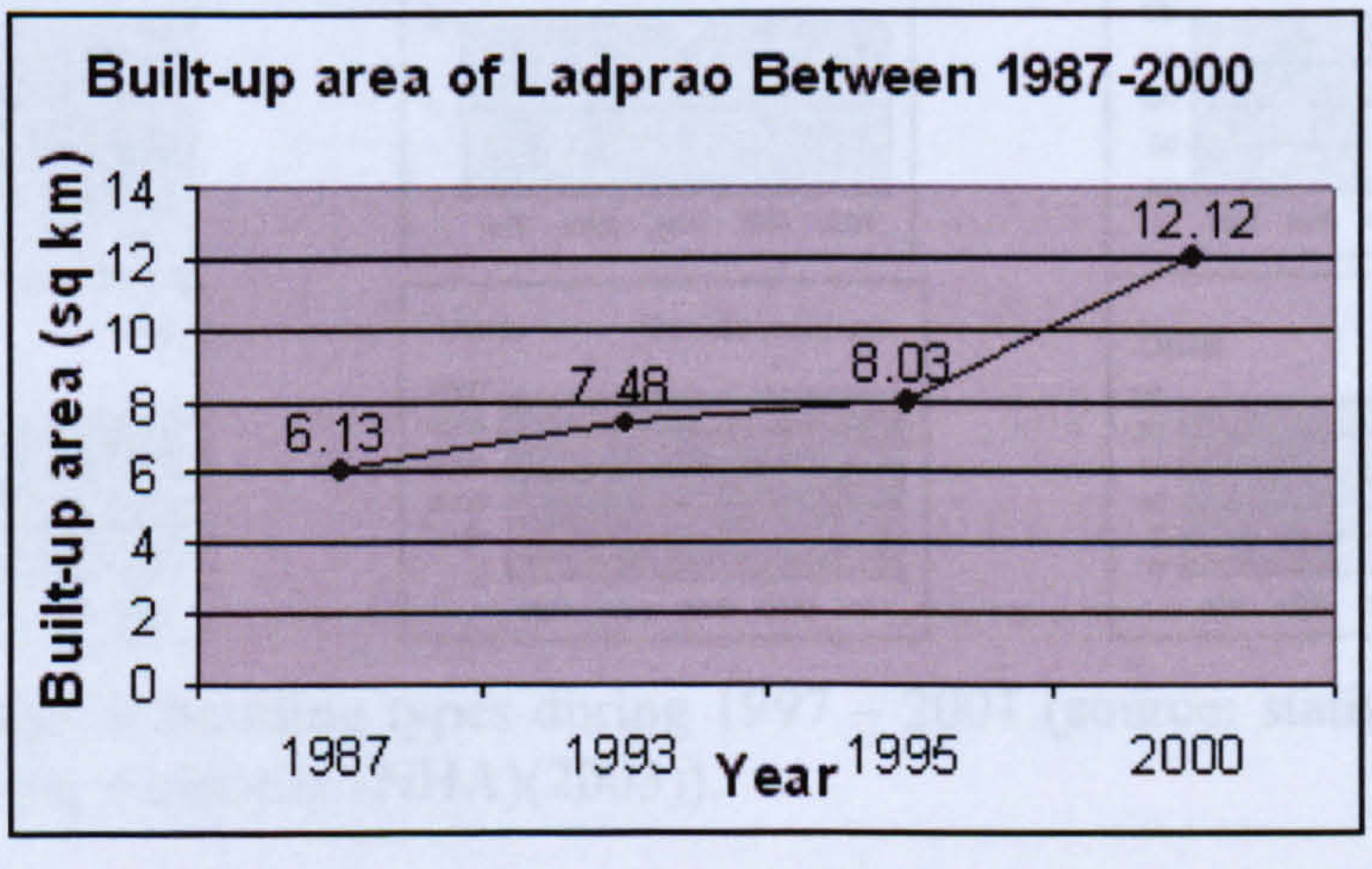


Figure 3.6: Transition of built-up areas during 1987 – 2000 (source: Department of City Planning (DCP) (2002)).

The increase of built-up areas from 1995 to 2000 is partly due to the change of government administration arrangements, coupled with the boom of housing development projects during 1985 – 1997, which later rapidly collapsed due to the economic crisis in 1997. Figure 3.7 reveals that the most significant growth since 1997 was residential, comprising detached-houses, condominiums, flats and townhouses respectively. This is probably due to demand from the middle-income group in search of a better place to live farther out from the city. At present, the total amount of developed land (built-up areas and roads) is 12.231 sq. km., or 57.38 % of the whole Ladprao area, mostly in the western and southern part of the district (Core Planning & Development, 2001). Thus, there is still vacant land left for future development. Most of the developed areas are used for residential and commercial purposes while industrial areas are rarely observable. Other built-up areas include institutional use (schools), government offices, temples, conservation areas and health services. The remaining undeveloped land, mostly situated in the northern and eastern part of the district, is occupied by vacant land, agricultural areas, fish ponds, and wells (Core Planning & Development, 2001).

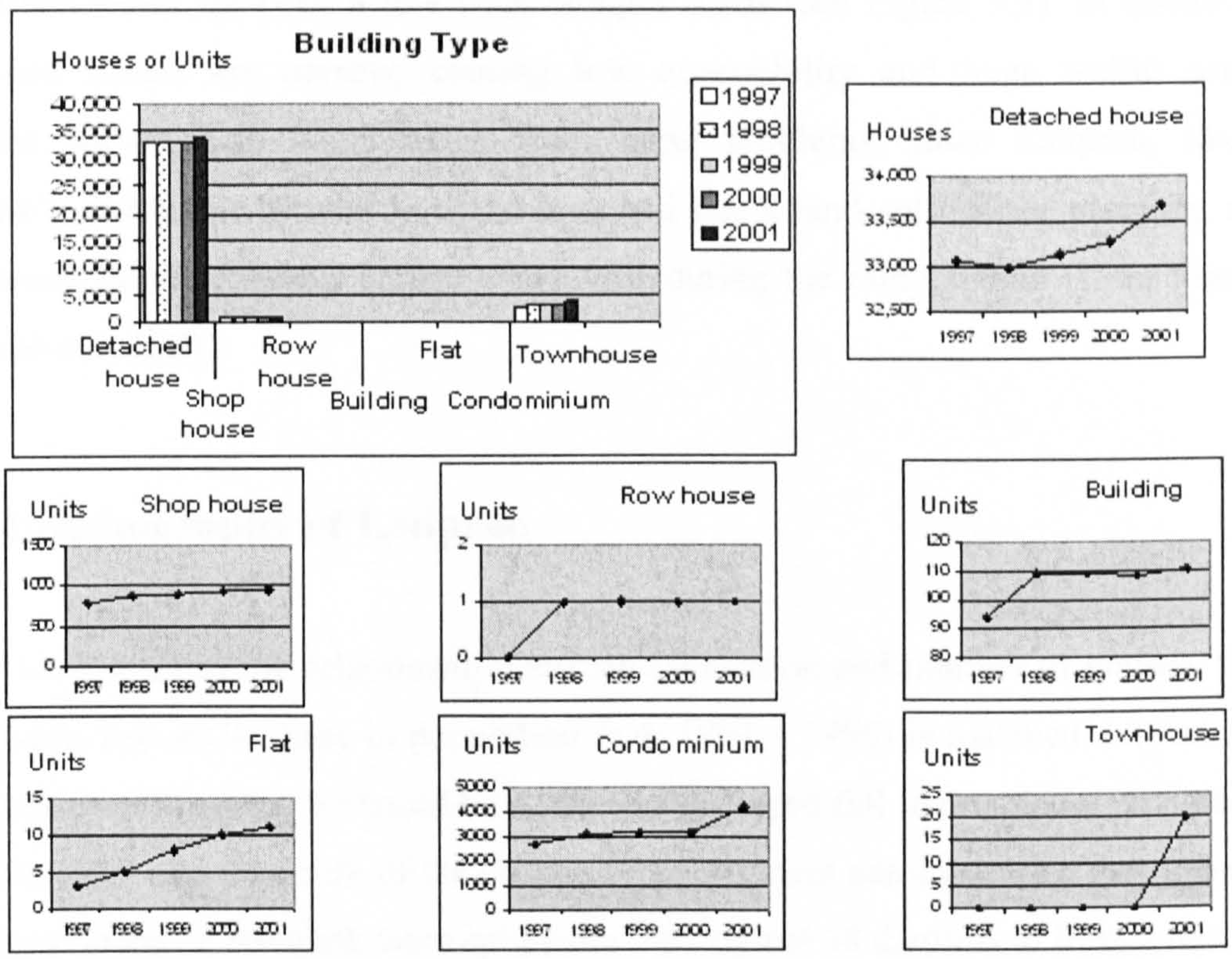


Figure 3.7: Increase of building types during 1997 – 2001 (source: statistics figures from the National Housing Authority (NHA)(2003)).

In terms of economy, until now the dominant image is residential use, which accommodates easy access to workplaces in the city center and the industrial area in the northern- and eastern-outer regions (Core Planning & Development, 2001). Employment structure in the area is largely based on the service sector. In 1999, most of the workers (92% of total workers) are from the service sector while 8% work in small (less than 15 workers) factories (Core Planning & Development, 2001).

Despite Ladprao being considered a transition zone, the area itself, in terms of accessibility, has high potential for residential development (Core Planning & Development, 2001). This is due to the fact that there are three recently built significant roads (namely Kaset-Navamin road, Ramintra-Ajnarong road, and the elevated Ramintra-Ajnarong highway), coupled with many future planned-road projects (i.e. expressway, highway), in the area and the districts nearby. Hence, the area, having great potential for residential development, benefits greatly from high accessibility, connecting the commercial area in the inner city and the industrial area in the outer region and other districts. However, the inner area of Ladprao has suffered from a road network problem, being a so-called 'superblock' as previously described in Section 3.1.1.3; an unusually large urban block between a few wide straight roads (see Figure 3.3). In addition, most roads and streets are narrow, causing low accessibility and huge traffic congestion problems during rush hour. Apart from these problems, since Ladprao, like many Bangkok's district, is located in a flat area and has a land subsidence problem, the area experiences flood problems almost every year during the rainy season (Core Planning & Development, 2001).

3.1.2.2 Demography of Ladprao

Figure 3.8 illustrates the relationship between population and number of houses. It can be clearly seen that an increase in population (e.g. 1993 – 1996) is matched with an increase in the number of houses. It should be noted that the rapid fall in population and number of houses in 1997 was the result of the change of government administration arrangements. In 1996, some parts of Bangkok were split from the original 38 districts to form a new total of 50 districts in 1997 that continue to be used up to the present. According to the (Department of City Planning (DCP), 1999), it is projected that from 2000 – 2022, the growth rate of population in Ladprao will level off or slowly decrease. However, since almost 50% of the land is still vacant, coupled with a gradual increase in the population density of Ladprao (Figure 3.9), there is potential that this vacant land may be used for future development.

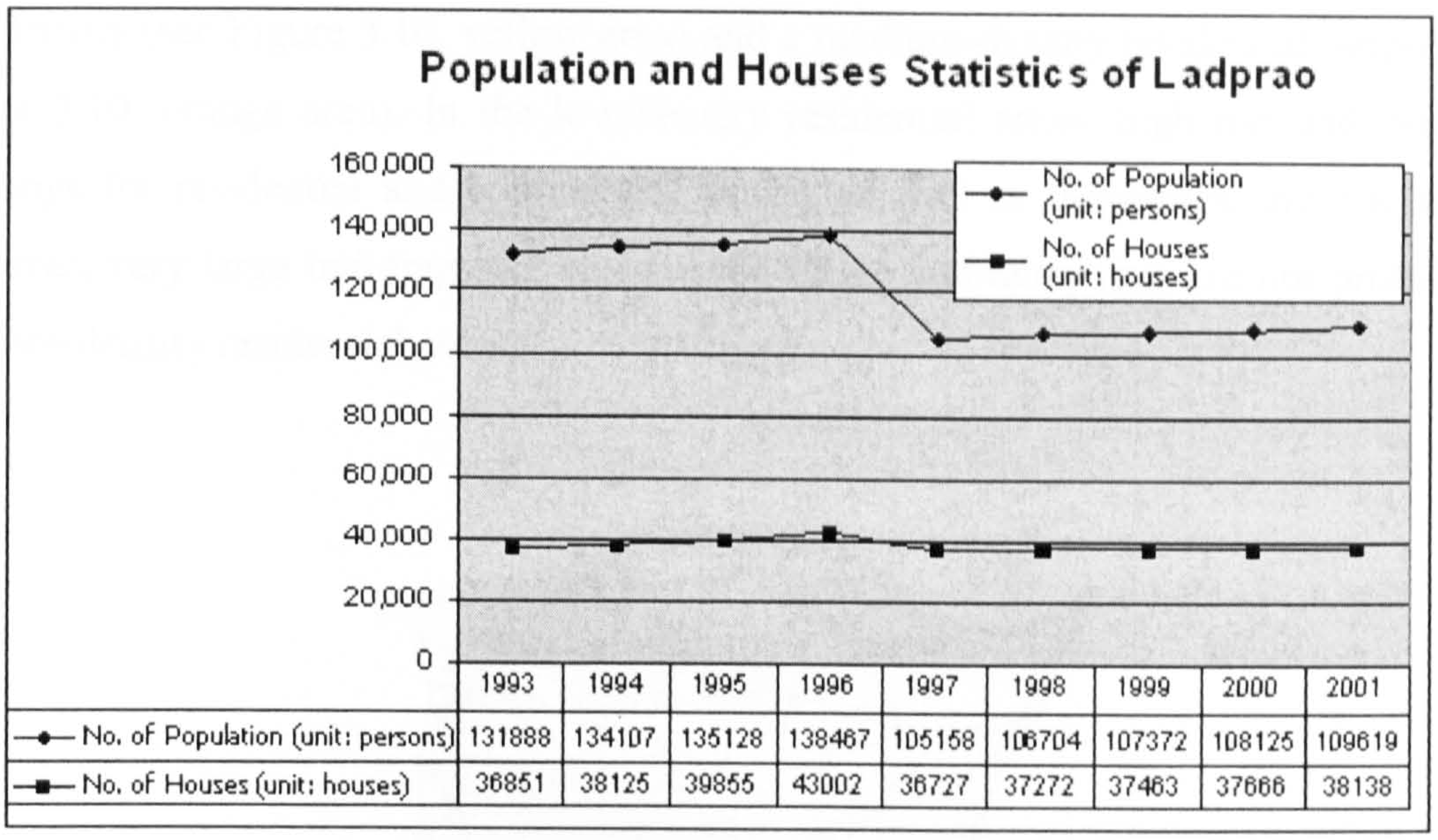


Figure 3.8: Population and number of houses in Ladprao during 1993 – 2001 (source: statistics figures from Department of Provincial Administration (DOPA)(2006)).

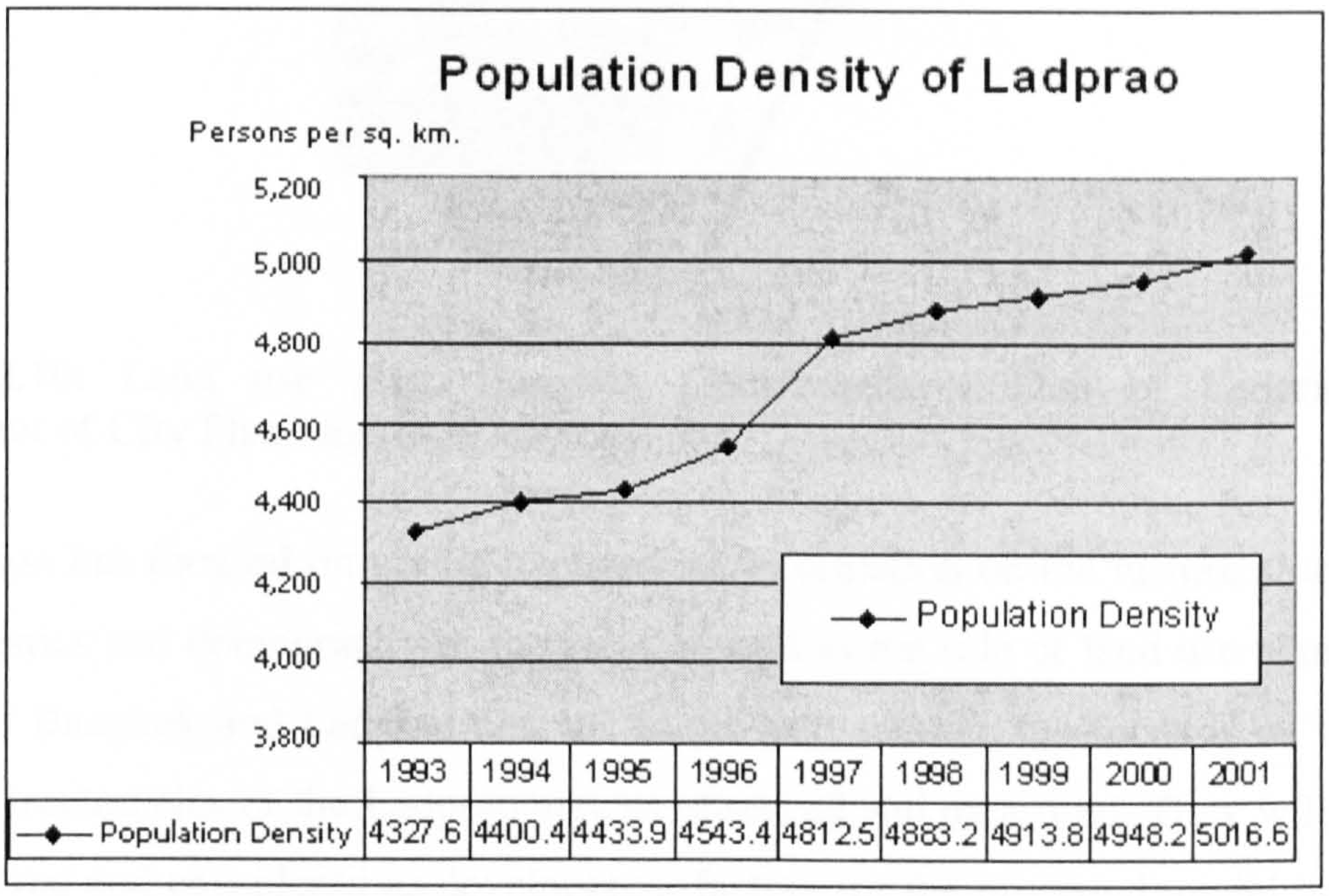


Figure 3.9: Population density of Ladprao during 1993 – 2001 (source: synthesized from statistics figures provided by Department of Provincial Administration (DOPA)(2006) and the National Housing Authority (NHA)(2003)).

3.1.2.3 District Planning Policy

Up to now, there has been no implementation of a land use plan at the district level, despite that the proposal of a development plan at the district level has been conducted. The currently enforced plan is the Bangkok Comprehensive Plan. In this plan land utilization for Ladprao has been designated for two classes, the majority of the area to be used for low-density (see Figure 3.10, yellow area) and a medium-density residential purposes (see Figure 3.10, orange area). In the low-density residential areas, high-rise and very large buildings for residential and commercial usage, as well as dumpsites, are not allowed. However, very large buildings for commercial usage and dumpsites are not prohibited in medium-density residential areas.

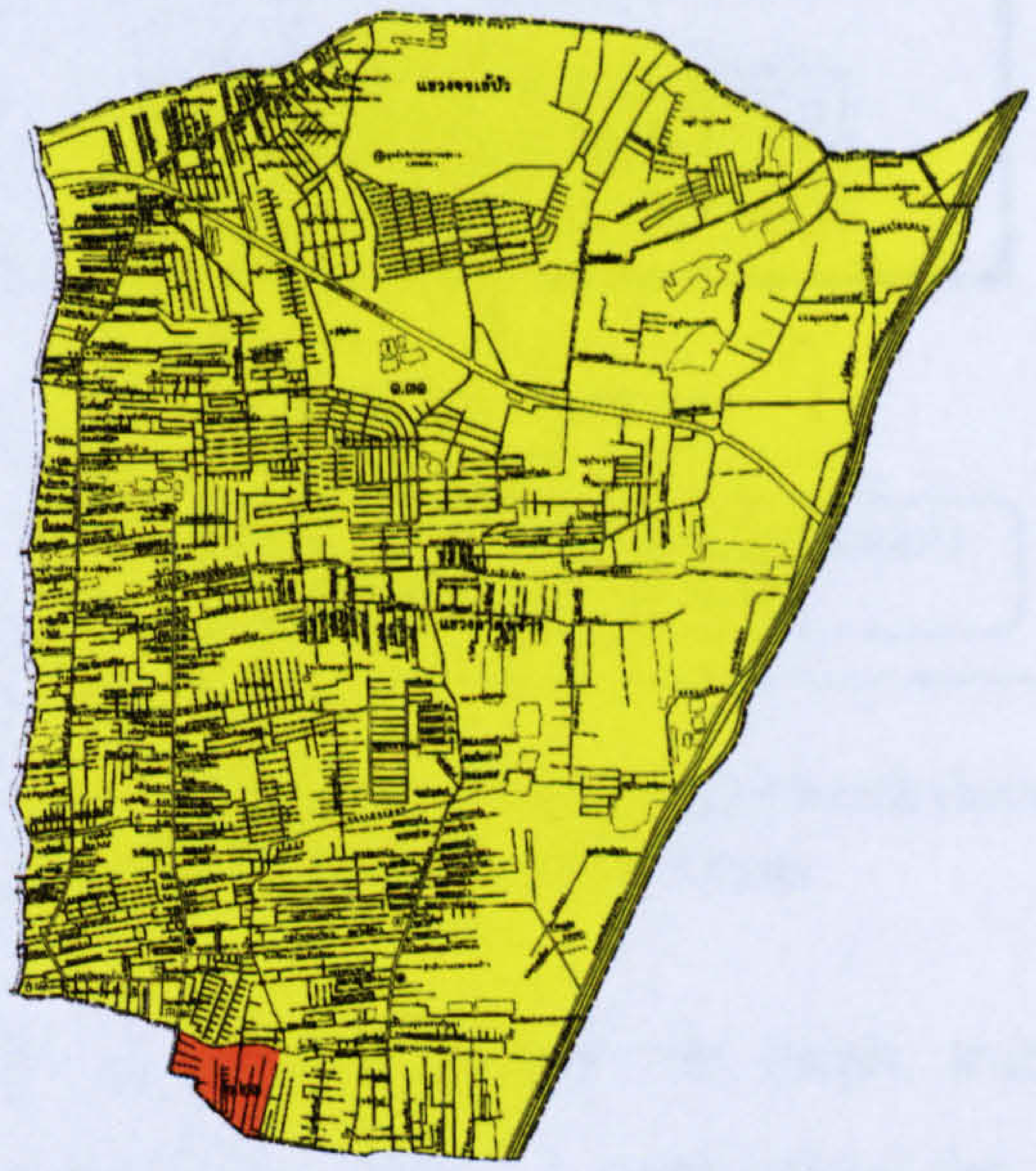


Figure 3.10: Land use plan, Bangkok Comprehensive Plan of Ladprao (source: Department of City Planning (DCP)(2002)).

This section has focused on giving background information on the historical and physical characteristics and demography of the area, as well as the role of land use planning in the context of Bangkok and Ladprao district. In the next section, the focus is on the detailed spatial characteristics of the Ladprao area as observed and measured. They will be used as a data source and considered as development factors for the proposed model implemented in Chapter 6.

3.2 Data Sources and Data Preprocessing

Figure 3.11 illustrates the schematic diagram of the overall conceptual methodology framework developed for the research study, broken down into three main stages: data input, model development and model evaluation. This figure highlights the first part, data input, which is described in this section. This includes description of the data sources and preprocessing. Here the modelling approach proposed is implemented using a raster-based data format. Thus, all data used in the study, obtained beforehand in a vector format, are rasterized into 10m cell grids.

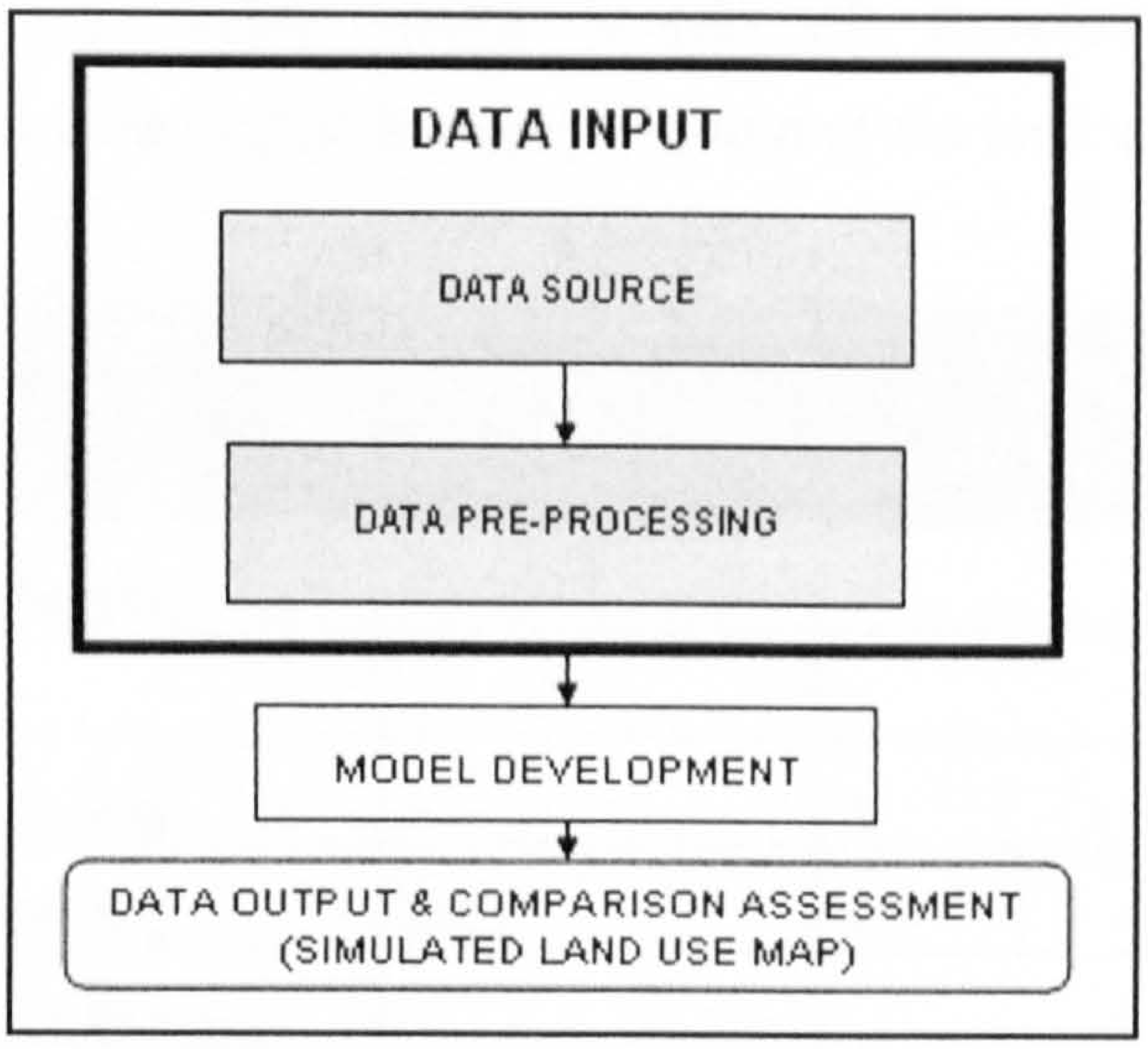


Figure 3.11: Schematic diagram of the modelling framework developed for the study, highlighting the data input being described in this section.

Map layers created for this study include land use maps, transportation maps, muban (village) maps, and land price maps. Table 3.3 summarizes the production of map layers for the model simulation. The creation of map layers was carried out using ready-to-use data and some results of preprocessing. The ready-to-use data source was obtained from the Department of City Planning (DCP) and well-prepared to be used, whereas preprocessing data sources required some preparation to derive the required map layers used for the simulation. Details about each data input map are described in the following sub-section.

3.2.1 The 1993 Aerial Imagery Covering the Entire Ladprao Area

The 1993 aerial imagery was not used for model simulation. However, it is vital in that it was used as the base map and data source for creation of the 1993 land use map being described in 3.2.2. In this study, the 22 aerial photographs at the scale of 1:6,000 in 1993 were all geo-referenced and mosaicked together to cover the entire Ladprao area as shown in Figure 3.12. The ground control points (GCP) for geo-referencing were read from a digital road map available to cover the overall area, mostly from the intersection of small roads where precise location can be pinpointed. Finally, the geo-referencing process was carried out with an overall accuracy (RMS error) of about 3.5 m by means of a geo-referencing module using the second order polynomial transformation within the ArcGIS 9.0 software package. Figure 3.13 shows an example of geo-referencing process conducted using the geo-referencing module within ArcGIS 9.0. Resultant 1993 geo-referenced imagery (Figure 3.12) was then used for the extraction of the land use map for 1993.

Description of Map/Image layer used for the simulation	Data Source Type		Description
	Ready-to-use	Preprocessing	
1.The 1993 aerial imagery		√	Section 3.2.1
2. Land use map			Section 3.2.2
2.1 The 1993 Land use map	-	√	
2.2 The 2001 Land use map	DCP	-	
3. Transportation (road) map			Section 3.2.3
3.1 The 1993 road map	-	√	
3.2 The 1998 road map	DCP	-	
3.3 The 2001 road map	DCP	-	
4. Muban (village) map			Section 3.2.4
4.1 The 1993 muban map	-	√	
4.2 The 1998 muban map	DCP	-	
4.3 The 2001 muban map	DCP	-	
5. Land price map			Section 3.2.5
5.1 The 1993 land price map	-	√	
5.2 The 1998 land price map	-	√	
5.3 The 2001 land price map	-	√	

Table 3.3: Map layers used for the simulation. DCP is the Department of City Planning, Bangkok Metropolitan Administration.

3.2.2 Land Use

In this study, the 1993 land use data was used as a base-year for simulation. It represents the areas that have been developed and those available for future development. The 2001 land use data was used as a final year for comparison assessment.

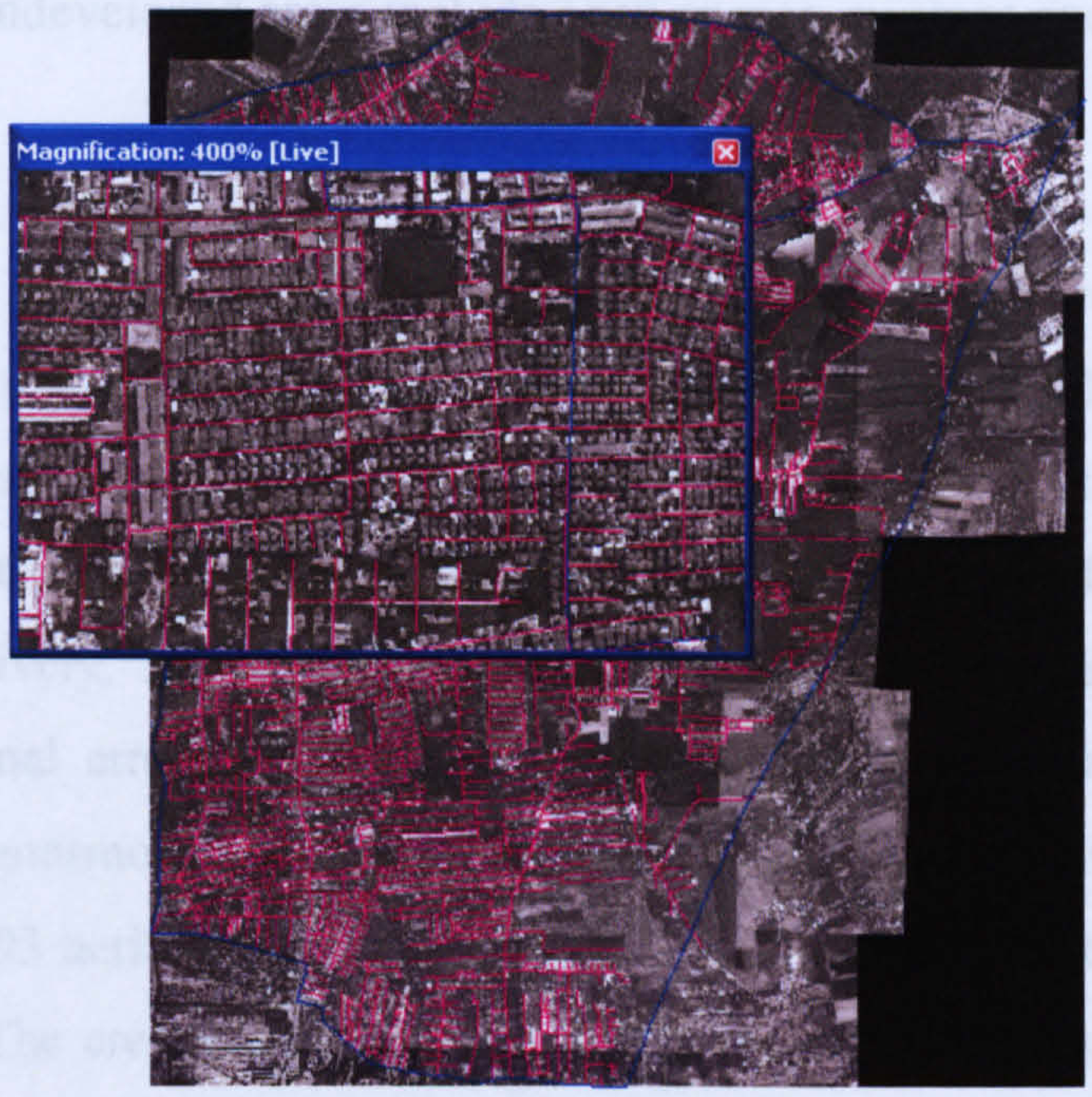


Figure 3.12: The 1993 aerial imagery draped with a vector map of the Ladprao administrative boundary (blue lines) and road map (pink lines).

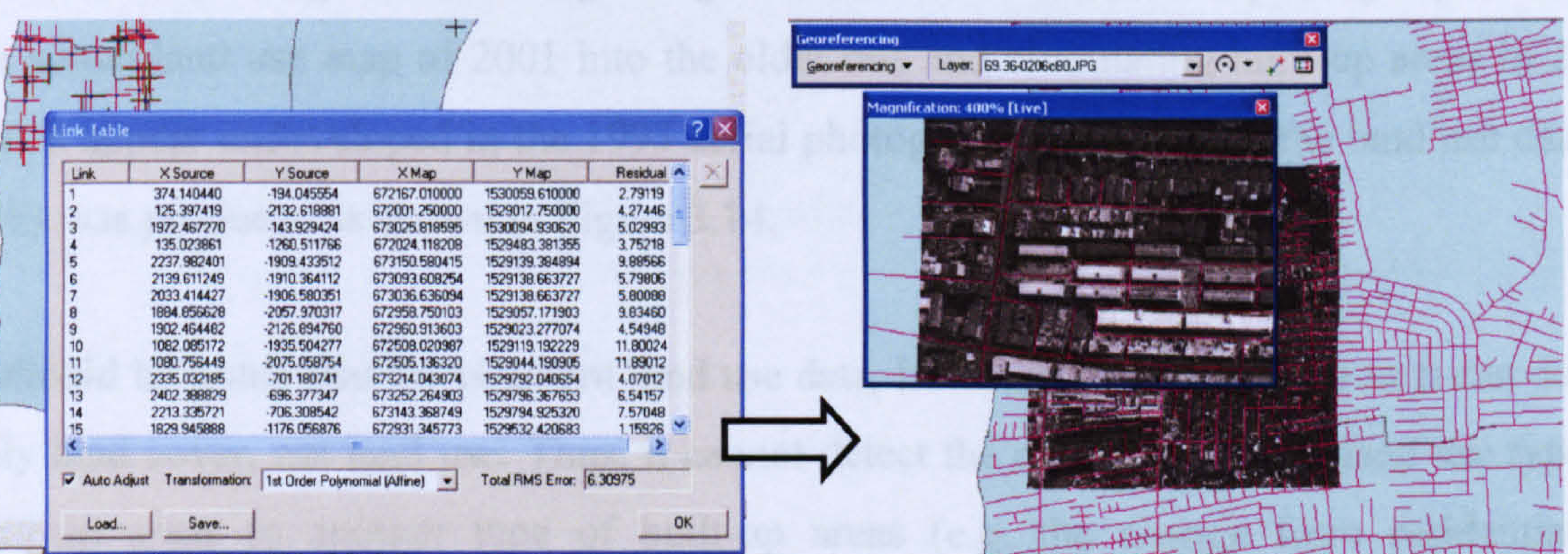


Figure 3.13: Procedure for geo-referencing aerial photographs using the geo-referencing module within ArcGIS 9.0.

Land use data for 2001, as shown in Figure 3.15, was from a digital land use map available in vector format produced by the Department of City Planning (DCP), Bangkok Metropolitan Administration at a scale of 1:4,000. Data for this map was produced in 2001 Aerial photographs at a scale of 1: 4,000, and a ground truth survey conducted in 2001 were used. The 1:4,000 scale map provides a detailed classification of different land use types at this scale distinguishing residential, commercial and industrial usage. According to this research study, the original detailed classification of land use categories was coded in ten dominant classes including: residential, commercial, industrial, government, school, conservation, agriculture, river, road, and vacant areas. In this research context, vacant areas referred to as undeveloped areas include open spaces, swamps and ponds in the study site.

The initial land use map of 1993, not available beforehand, was created based on the existing 2001 land use map mentioned above and historical aerial photographs. The available historical land use data for this study site can be dated back to 1981, 1987 and 1993. This data is in the form of paper aerial photographs at the scale of 1:30,000, 1:20,000 and 1:6,000 respectively. However, since the 1981 and 1987 aerial photographs had unacceptable positional errors when draping to the available 2001 land use map (scale 1:4,000) due to the enormous difference of resolution and scale (1: 30,000 and 1:20,000 respectively), the 1993 aerial photographs at a scale of 1:6,000 were chosen as the initial data for extraction. The creation technique was performed by mosaicking the 1993 aerial photographs together as one geo-referenced image (see the mosaicking procedure in Section 3.2.1), then draping this geo-referenced imagery to the existing 2001 land use map, and then conducting on-screen digitizing in the ArcGIS 9.0 software package by reversing the actual land use map of 2001 into the older one and eliminating built-up areas in 2001 which appear undeveloped in the 1993 aerial photographs. As a result, the land use data of 1993 was produced as shown in Figure 3.14.

It should be noted that the resultant land use data, however, has limitations as it can detect only land cover, not land use. Thus, it cannot detect the change from one land use type of built-up areas to another type of built-up areas (e.g. the change from residential to commercial use). Such limitation is also partly the result of the limited scale of available historical aerial photographs, coupled with the unavailability of a ground survey and auxiliary data from the time period considered. Consequently, only the change from undeveloped area (vacant) in 1993 to developed area in 2001 can be detected.

In result from this processing, land use maps from the two periods (1993 and 2001), were created for use in Chapter 4 and 6. The 1993 and 2001 land use maps are shown in Figure 3.14 and 3.15 respectively.

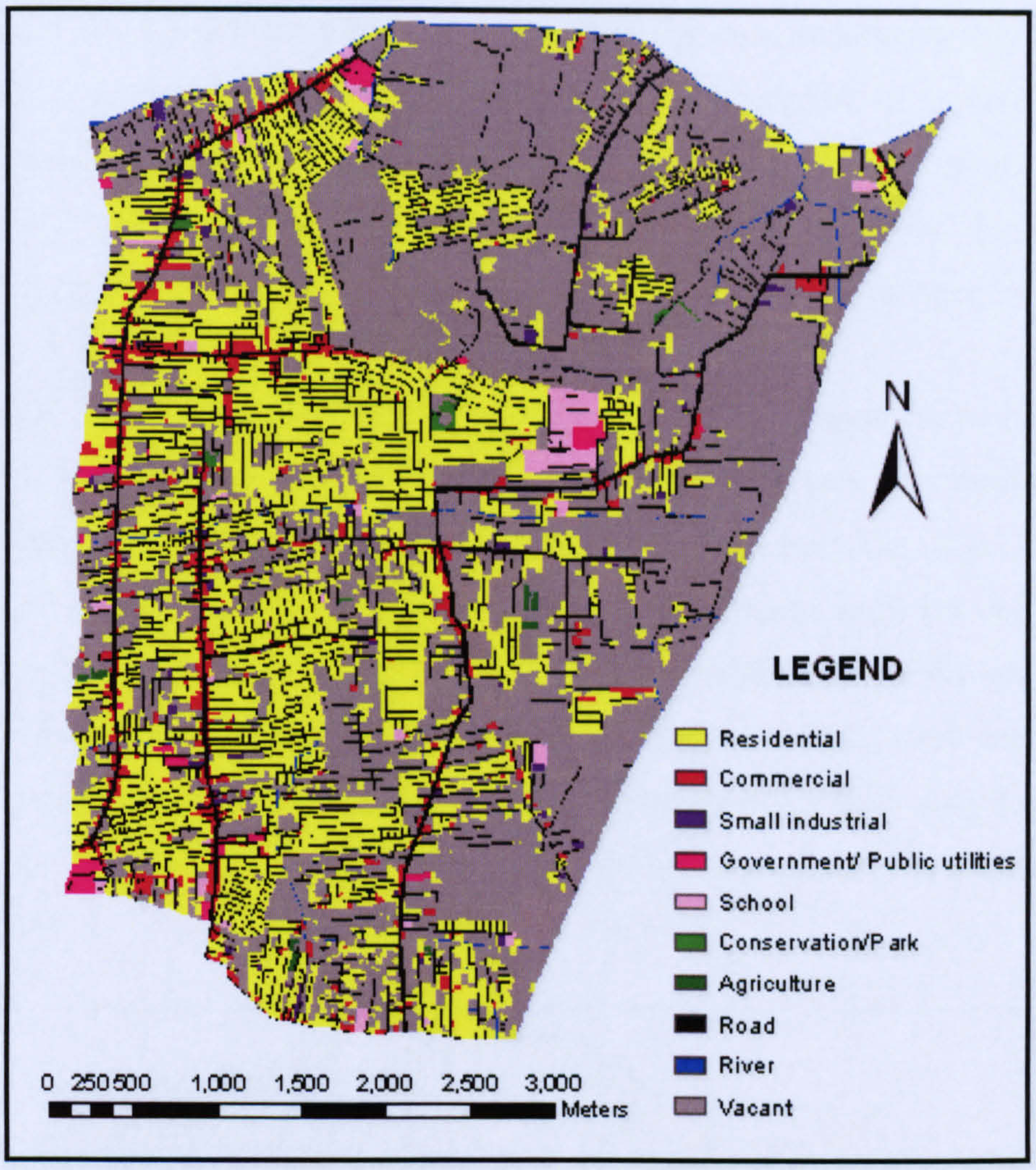


Figure 3.14: The 1993 land use map.

3.2.3 Road Transportation

The location of the transportation facilities is one of the key determinants for land suitability for different land use types. There are two main means of transportation in this area, the canal and road networks. Prior to this day, rivers, especially the Chao Phraya River and a connected canal system formed the focus of social life and transportation for Bangkokians. However, after the influence of western values brought on industrialization of the city, many viable canals were filled to support a road network and gradually transportation has shifted from canals to roads (Boonchuen, 2002; Ross et al., 2000). As assessed by means of aerial photographs and satellite imagery, (Chomchan et al., 1990) it can be observed that the establishment of new roads and the building of bridges across the river act as major determinants for Bangkok growth.

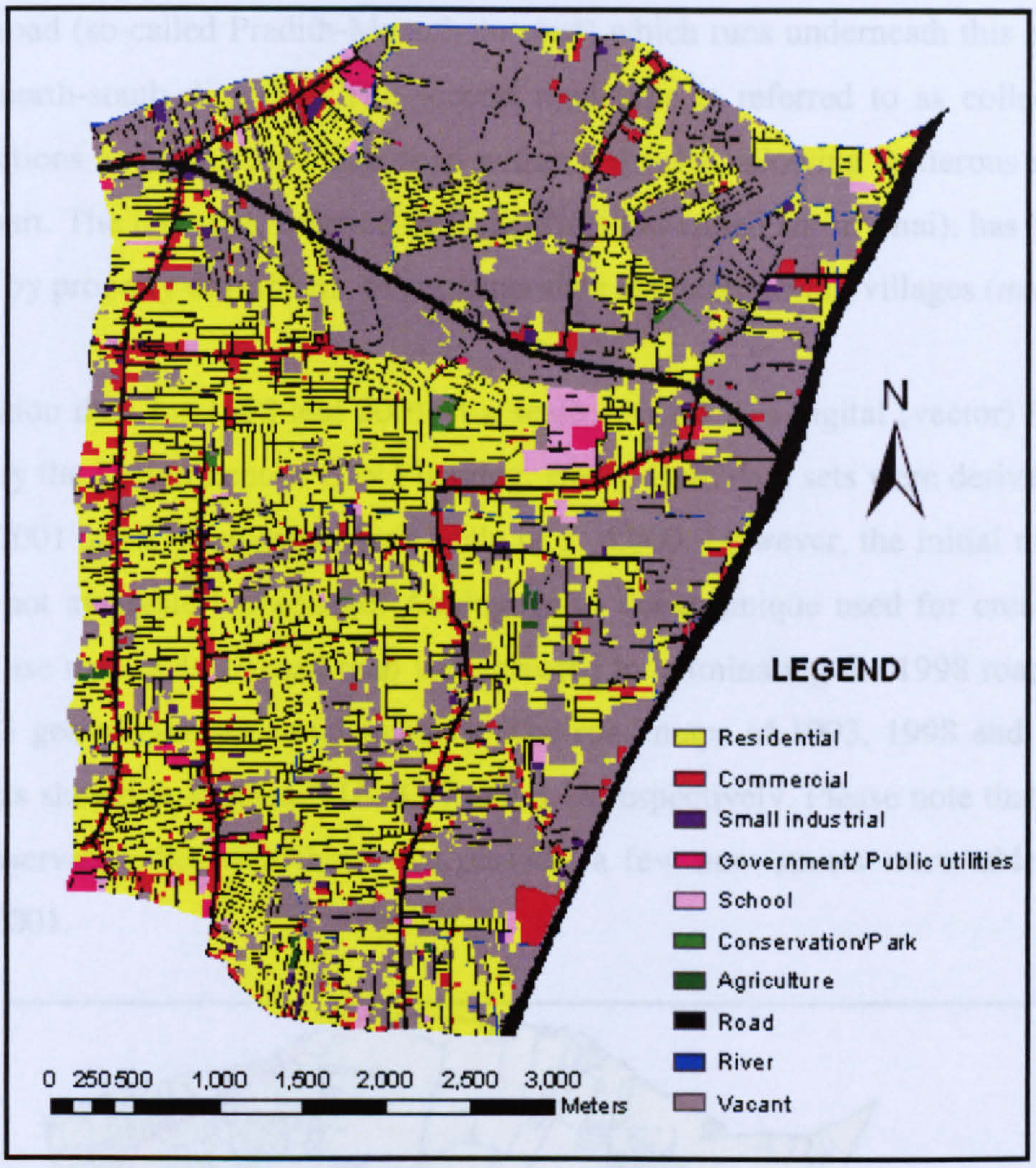


Figure 3.15: The 2001 land use map.

According to the books 'Urban Planning and Design 2nd edition' published by the Department of Highways and Public Works and 'Criteria and Standard of Comprehensive Plan Designation and Planning (BE. 2539)' published by the Department of Town and Country Planning (DTCP), roads are classified in four types on the basis of their main function which includes major roads, secondary roads, collector streets and local streets (Core Planning & Development, 2001). In this study, these four types are grouped into three. The first road type is referred to as major (highway and expressway inclusive) and secondary roads, which connect Ladprao with other adjacent districts. These two road types in this research study were grouped into one type. The reason for this is to accommodate analysis and simulation since after being observed and investigated they both performed a similar function in the study site. For this type of road, included are Chokchai 4 road and Yothinburana road running from the north to the south. However, in 1998, there were three newly built major roads and expressways including Kaset-Navamin linking in an east-west direction, the Ramindhra-Ajnarong expressway, and Ramindhra-

Ajnarong road (so-called Pradith-Manutham road) which runs underneath this expressway also in a north-south direction. The second road type is referred to as collector street, which functions as feeder/connector road within a district, allowing dexterous mobility in the inner part. The last type, referred as local street (so-called *soi* in Thai), has been partly developed by property developers, to accommodate residents within villages (*mubans*).

Transportation data for 1998 and 2001 at a scale 1:4,000 in a digital (vector) format was produced by the Department of City Planning, BMA. Both data sets were derived from the 1998 and 2001 aerial photographs at a scale of 1: 4,000. However, the initial road map of 1993 was not available beforehand. Analogous to the technique used for creation of the 1993 land use map, a 1993 road map were created by eliminating the 1998 roads not seen in the 1993 geo-referenced aerial imagery. The road maps of 1993, 1998 and 2001 were generated as shown in Figures 3.16, 3.17 and 3.18 respectively. Please note that according to road observation between these two periods, a few new streets were added between 1998 and 2001.

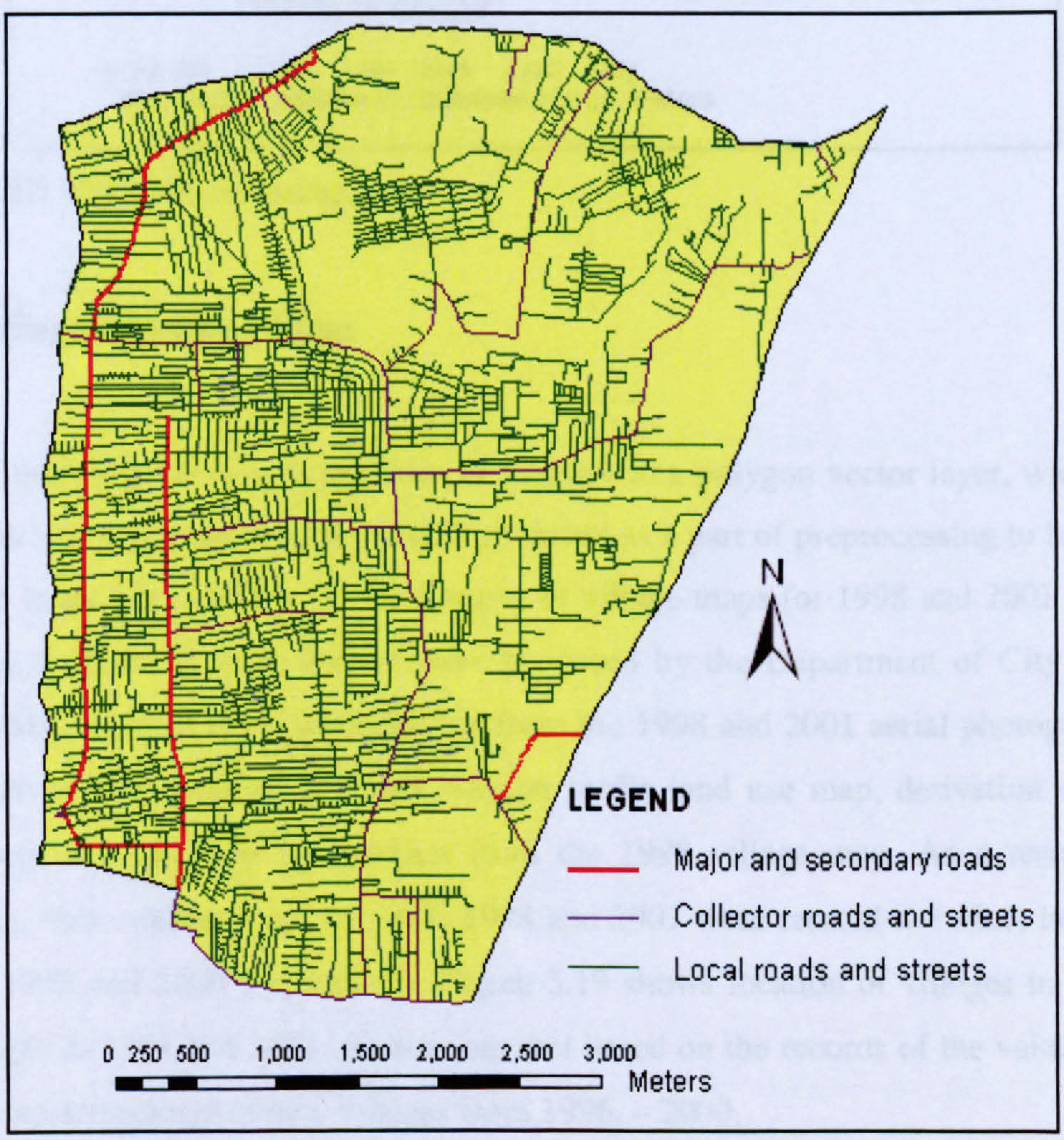


Figure 3.16: The 1993 road map.

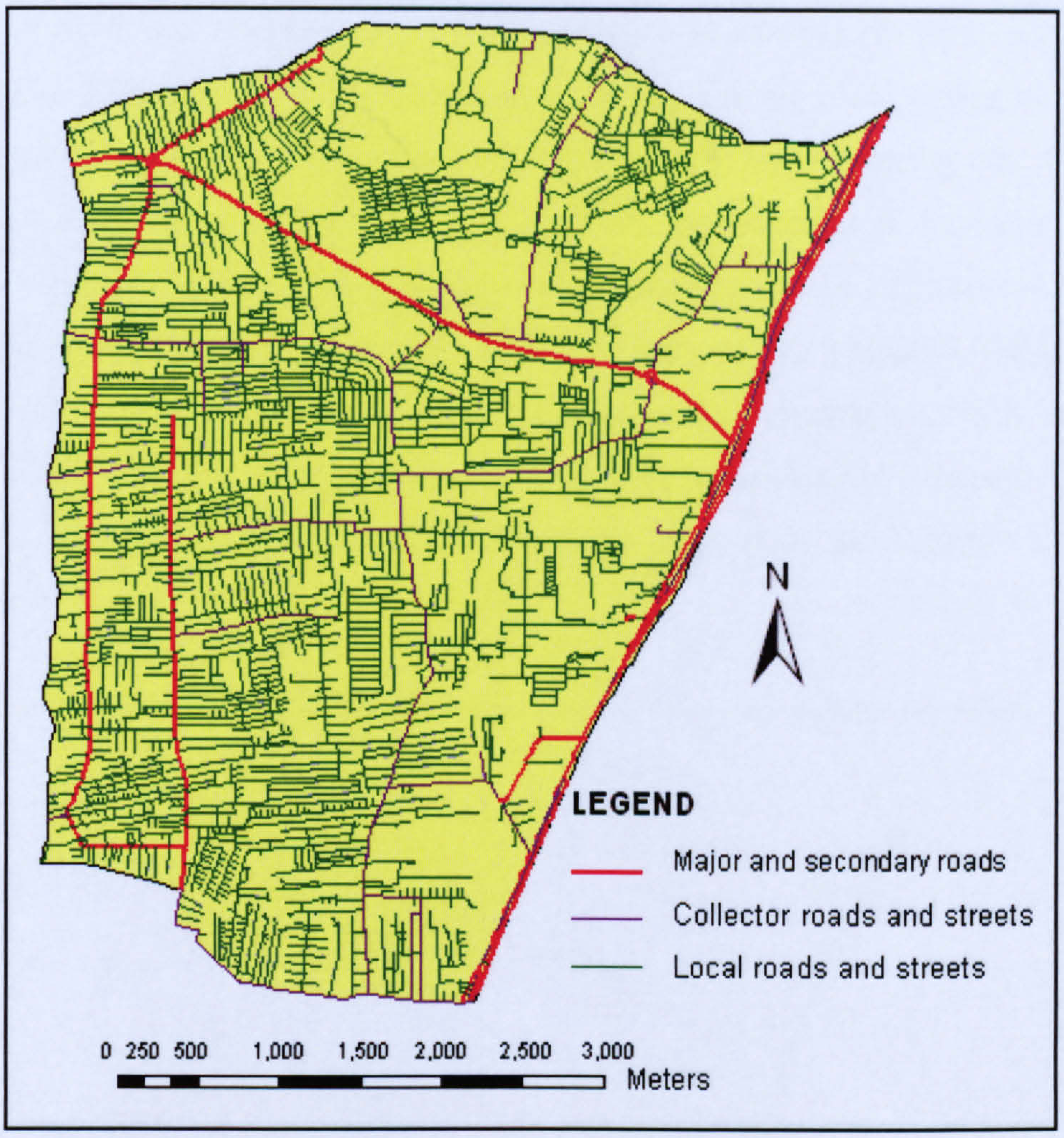


Figure 3.17: The 1998 road map.

3.2.4 Village (muban) Map

A village map, representing boundaries of villages in a polygon vector layer, was not part of the simulation process. Instead, it was produced as a part of preprocessing to help create land price maps (see Section 3.2.5). A series of village maps for 1998 and 2001 at a scale 1:4,000 in a digital (vector) format were produced by the Department of City Planning (DCP), BMA. Both of them were derived from the 1998 and 2001 aerial photographs and ground survey at a scale of 1: 4,000. Similar to the land use map, derivation of a 1993 village map was obtained by deletion from the 1998 village map. As a result of this processing, three village maps of 1993, 1998 and 2001 were created to reflect land values of 1993, 1998 and 2000 respectively. Figure 3.19 shows location of villages in 1993 and new villages in 1998 and 2001. Please note that based on the records of the valuation rolls there was no emergence of new villages from 1996. – 2000.

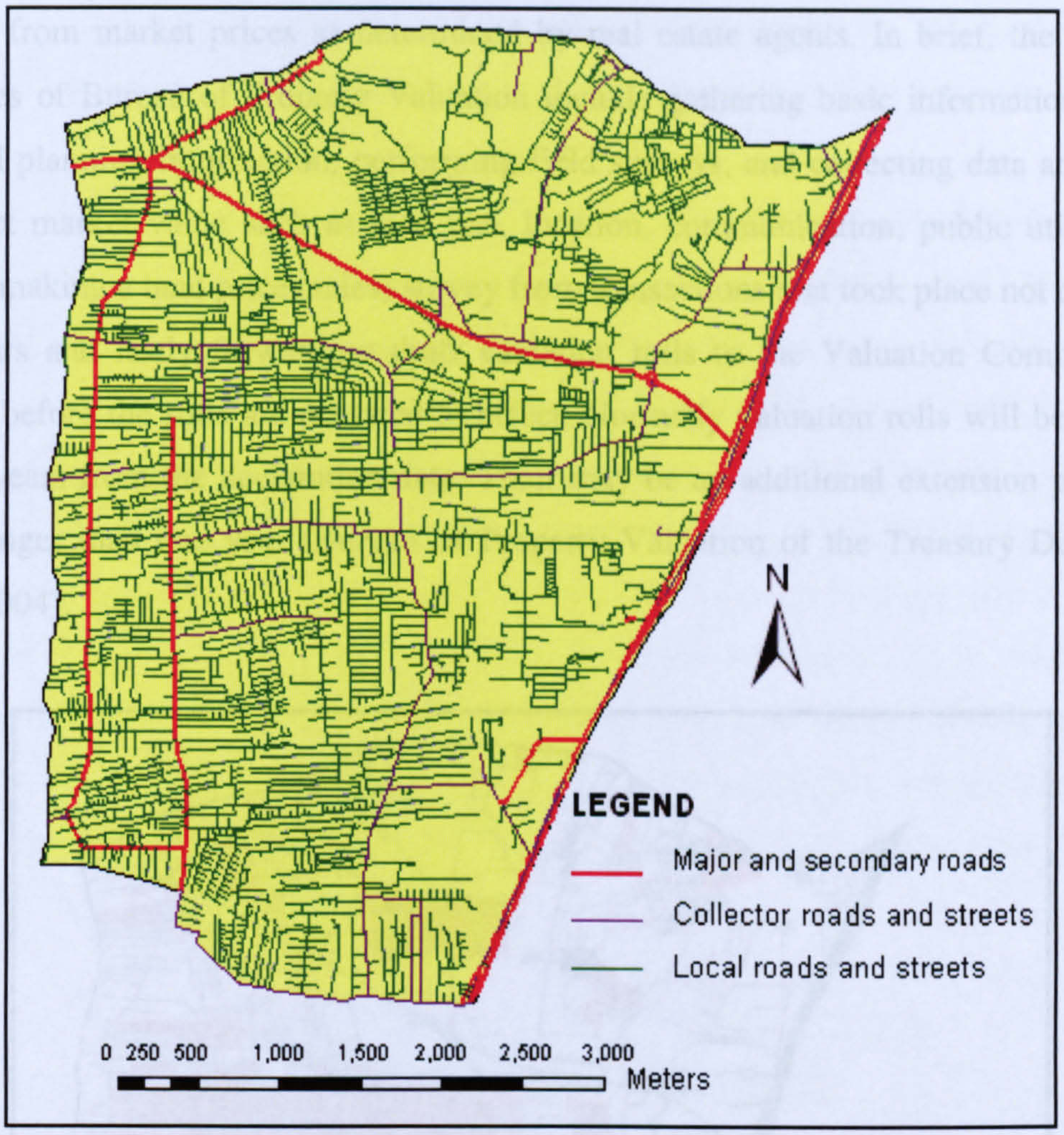


Figure 3.18: The 2001 road map.

3.2.5 Land Price

Land price from the study area was obtained from the Bureau of Property Valuation under the Treasury Department, which is mainly responsible for valuation of immovable properties including lands, buildings and condominiums in Thailand including the Bangkok Metropolitan area (Bureau of Property Valuation of the Treasury Department (TRD), 2004). In this study, land price data during the period of study was available for use as three datasets reflecting three periods (1992 – 1995, 1996 – 1999, 2000 – 2003) in a paper format of valuation rolls of Block Based Valuation. Block Based Valuation contains values groups for lands along roads, streets and paths.

These data were collected from two main sources (Bureau of Property Valuation of the Treasury Department (TRD), 2004). The first source is the average land price as appraised by their own survey team based on proximity to roads. The latter type of land price data is collected from market prices as determined by real estate agents. In brief, the valuation procedures of Bureau of Property Valuation include gathering basic information such as maps and planning information, performing field surveys, and collecting data and factors that affect market value such as land use, location, communication, public utilities and services, making a land price (sales) survey from transactions that took place not older than three years and lastly forwarding draft valuation rolls to the Valuation Committee for approval before the rolls are declared in effect. Normally valuation rolls will be in effect for four years from the declaration date. Their may be an additional extension period but for no longer than one year (Bureau of Property Valuation of the Treasury Department (TRD), 2004).

Figure 3.19: The village (muban) map.

It should be noted this data has some limitations since it probably does not reflect real market values. Firstly, the process is long (almost a year) involving land price data surveying, valuing and approval. This is then combined with the fact that the resultant land price values are used for approximate four years, making this data frequently outdated or obsolete. Moreover, since the data available only contains crude information and is paper-based, it may result in unrealistic information since it does not specify the exact land price at a detailed location as required for the micro-scale simulation used in this research study.

The digital land price maps during the study period for land use simulation described above were created based on those available paper-based formats containing crude information. To be produced, the same procedure was performed respectively for each land price map, as shown in Figure 3.20. The process starts by buffering road areas from a road map available at a scale 1:4,000 to create a buffer-road map, then overlaying this with a muban map, vector layer of village (*muban*) (see Section 3.2.4) and finally inputting land price values as land price attributes in a land price layer. Since no land use map is available for 1996, the land price map created is based on the 1998 data. As a result of this processing, three land price maps for 1993, 1998 and 2000 were created to correspond with land use and road maps. These land price maps reflect land prices during three periods (1993 – 1997, 1998 – 1999 and 2000 – 2001) respectively. These maps are shown in Figures 3.21, 3.22 and 3.23 respectively.

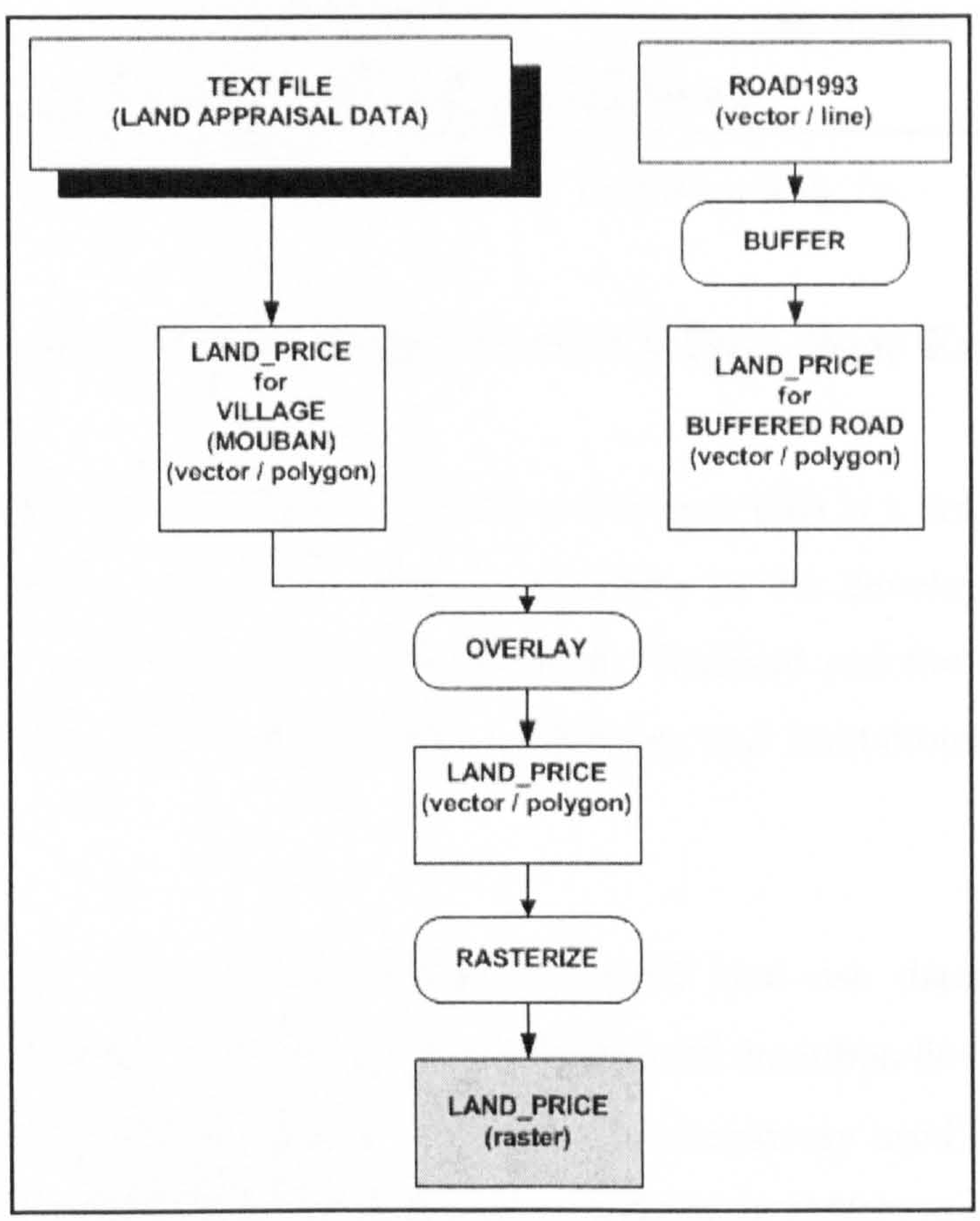


Figure 3.20: The preprocessing stage for creation of the land price maps.

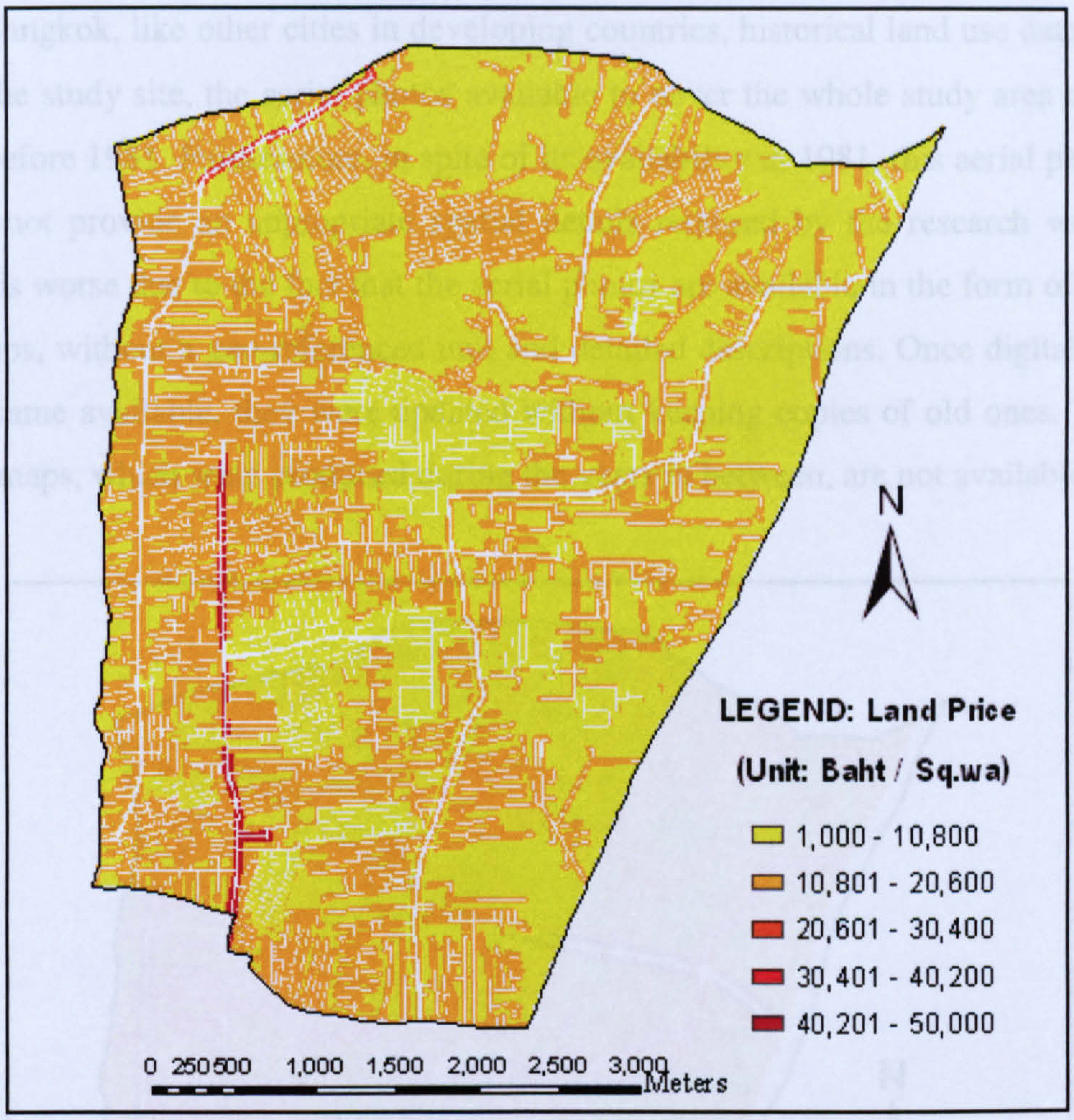


Figure 3.21: The 1993 land price map. Remark: 1 sq.wa = 4 sq. m.

3.3 Data Constraint Issues: Limitations in a Data Poor Environment

As previously mentioned in Section 2.3, the issue of poor data is a critical problem for the progress of GIS in the context of urban applications in the developing countries. This problem is also observed in the case of Bangkok, Thailand and the study site (Ladprao district). In this section, discussion of data constraints and limitations in the study site is presented.

The first limitation is due to the limited historical land use datasets. As previously described, land use data is a key source for simulation of the urban development process in this research study. Historical land use datasets are extensively used as a way to observe and understand urban growth in the study area. Clarke et al. (1997) used historical maps of urban extent over 120 years to provide a prevailing reference base for examining the urban growth in the San Francisco Bay. Herold et al. (2001) gathered the observable historical land use information over sixty years to calibrate the temporal growth of Santa Barbara. In case of Bangkok, like other cities in developing countries, historical land use data is rather rare. In the study site, the aerial photos available to cover the whole study area cannot go back to before 1981. Furthermore, in spite of its availability in 1981, this aerial photograph set does not provide an appropriate spatial details required by the research work. This problem is worse due to the fact that the aerial photos are available in the form of sketched paper maps, without a geo-referenced unit and detailed descriptions. Once digital land use maps became available, they were updated without keeping copies of old ones. Thus, the land use maps, which were produced during the years in between, are not available.

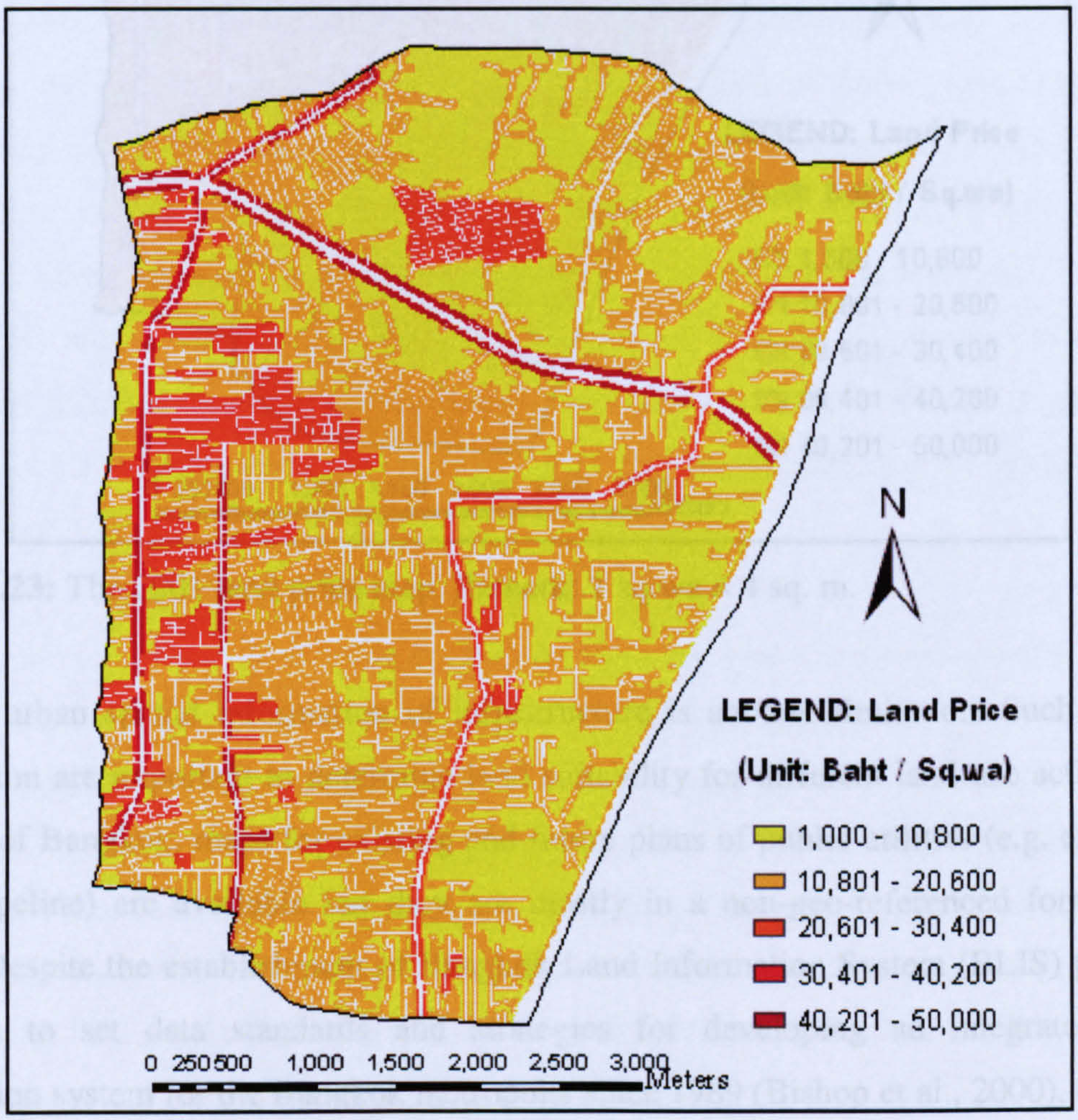


Figure 3.22: The 1998 land price map. Remark: 1 sq.wa = 4 sq. m.

The second limitation concerns land price data. As mentioned earlier, until now land price datasets mostly have been in the form of unscaled and sketched paper, containing crude information without the exact geo-referenced positions. Furthermore, they are unavailable at the parcel level. Even more problematic is that they are possibly unrealistic due to long time process for data collection and valuation.

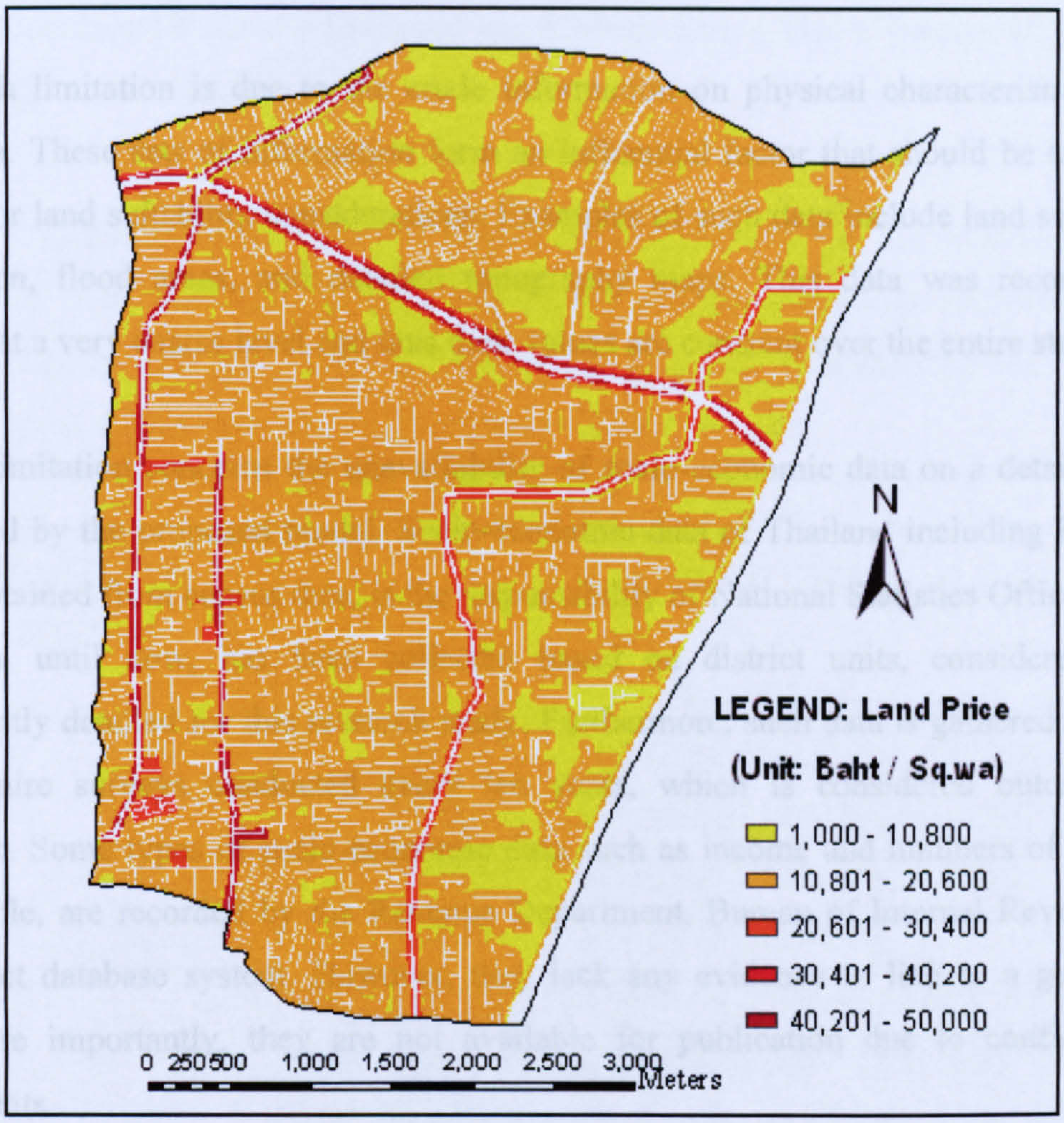


Figure 3.23: The 2001 land price map. Remark: 1 sq.wa = 4 sq. m.

Lack of urban spatial information of infrastructure is another limitation. Such kinds of information are important to determine land suitability for different land use activities. In the case of Bangkok, maps for existing and future plans of public utilities (e.g. electricity, water pipeline) are available but they are mostly in a non-geo-referenced format (blue paper). Despite the establishment of Bangkok Land Information System (BLIS) project as a means to set data standards and strategies for developing an integrated spatial information system for the Bangkok metropolis since 1989 (Bishop et al., 2000), until now the utility maps at the district level or smaller from most important utility agencies such as the Metropolitan Water Authority (MWA), responsible for potable water and pipe distribution, and the Telephone Organization of Thailand (TOT), responsible for land-line telephone installation, are unavailable. Some utility agencies such as the Metropolitan Electricity Authority (MEA), responsible for electricity distribution, and the Department of Land, responsible for land parcel registration, have geo-referenced map data at a scale 1: 1,000 or larger, however, this data is not publicly available to researchers due to confidentiality requirements.

The fourth limitation is due to the crude information on physical characteristics in the study area. These sets of information form an influential factor that should be taken into account for land suitability considerations. Examples of such data include land subsidence information, flood areas, and detailed topography maps. This data was recorded and collected at a very coarse level and thus their values are constant over the entire study site.

Another limitation concerns the unavailability of socio-economic data on a detailed level as required by the proposed model. Socio-economic data of Thailand including Bangkok, mainly obtained from census data, is the responsibility of National Statistics Office (NSO). This data, until now, has been collected based on district units, considered to be insufficiently detailed for this research study. Furthermore, such data is gathered based on questionnaire surveys conducted every ten years, which is considered outdated and unrealistic. Some types of socio-economic data such as income and numbers of workers, for example, are recorded by the Revenue Department, Bureau of Internal Revenue in a spreadsheet database system. However, they lack any evidence to link to a geographic map. More importantly, they are not available for publication due to confidentiality requirements.

A final limitation is due to the lack of detailed land use plan. Enforcement of a land use plan is assumed beforehand to help control or at least dictate the growth of urban development. However, the land use plan available for Ladprao area containing two zoning maps (see detail in Section 3.1.2.3), is considered very crude in detail and has no enforcement. Because of unavailability of a detailed land use plan, the model developed in this research study cannot account for the land use plan map as part of an institutional control factor (see details about institutional control in Section 3.4.4) for model simulation.

To recap, the limitations and constraints in the study site discussed above thus delimit the capability of the model developed in terms of limited data availability, limited accurate and up-to-date information and a lack of detailed data for inclusion in model simulation. In the research study where maps are on the scale of 1:4,000 and 1:6,000, the crude and limited data obtained, unavoidably, to some extent, affects the creation of simulation results.

3.4 Development Factors Influencing Urban Land Use Change Used in the Model

Table 3.4 illustrates the development factors used in the analysis. Development factors chosen here were delimited by data availability and constraints as addressed in the previous section (Section 3.3). Such factors may be characterized as physical, neighbourhood, environmental, and institutional factors.

Type of Factors	Development Factors
1. Physical factors	1.1 Land price (Figure 3.23)
	1.2 Proximity to roads (Figure 3.24)
	1.3 Proximity to residential use (Figure 3.25(a))
	1.4 Proximity to commercial use(Figure 3.25(b))
	1.5 Proximity to industrial use (Figure 3.25(c))
	1.6 Proximity to schools (Figure 3.26)
	1.7 Proximity to park/ conservation areas(Figure 3.27)
	1.8 Proximity to government use(Figure 3.28)
	1.9 Proximity to agriculture use (Figure 3.29)
2. Neighbourhood factors	2.1 Residential neighbourhood (Figure 3.30(a))
	2.2 Commercial neighbourhood (Figure 3.30(b))
	2.3 Industrial neighbourhood (Figure 3.30(c))
3. Environmental factors	3.1 Exclusion of water bodies, roads
4. Institutional control	4.1 Exclusion of conservation/park and agricultural areas
	4.2 Exclusion of areas constrained for development (Figure 3.31)

Table 3.4: Development factors used for model implementation in the research study.

Physical factors describe the physical characteristics of the cells themselves such as slope. They also include the spatial characteristics of cities such as the spatial relationships between cells and other functional activities within the study site (e.g. proximity to major activities). Neighbourhood factors, represented by the degree of urban development in the neighbourhood space, are used to reflect urban agglomeration and compatibility in land use (Chapin and Kaiser, 1979; Wu, 1998). Environmental factors, such as exclusion of water bodies and roads from development, are considered as constraint factors applied to the study site. Institutional factors refer to the intervention of planning policy in order to prohibit some areas from development (e.g. conservation/park).

3.4.1 Physical Factors

Physical factors are used as the main inputs to the model for the designation of land suitability areas or the potential areas for urban development. For example, an undeveloped area that is accessible by a new road is more suitable to be developed to many types of urban land use. Physical factors used in this study include land price, proximity to roads, proximity to residential use, proximity to commercial use, proximity to industrial use, proximity to government use, proximity to schools, proximity to park/ conservation areas and proximity to agricultural use.

Land price or *land value* factor (see Figure 3.21, 3.22 and 3.23) influences the suitability of an area for a variety of land use types in different locations (Chapin and Kaiser, 1979). This is the result of competition amongst land use activities, especially in urban areas where demand for land is high (Darin-Drabkin, 1977). The work conducted by Waddell and Ulfarsson (2002a) is one example that includes land value for the analysis of employment location. However, including this factor to the model should be done with caution. Firstly, Many theories (e.g. classical concepts of von Thunen's Agricultural land rent theory, Alonso's urban land market theory) and some research works such as that of Waddell (1998) have implied that there are some significant relationships between 'the valuation of land' to 'travel costs', 'access to market centers', and 'transportation improvements' (Chapin and Kaiser, 1979). Thus, an inclusive land price, incorporating other related variables such as 'proximity to roads' to the model may cause redundancy of variables. This may result in instability and unreliability in the creation of coefficients (Field, 2000; SPSS Inc., 2001) that will be used for the calculation of land suitability in the model. Secondly, prices of vacant plots in the city region in most developing nations are very high and unrealistic despite much vacant area being left (Darin-Drabkin, 1977). He explained that in some circumstances, such as that of a permanent inflation situation, land holders tend to set the price of land very high to earn profits despite low tax payments.

Proximity to roads factor measures the distance from a cell to the nearest point on the road network. This type of factor is crucial for urban development since it shows the ease of accessibility (Chapin and Kaiser, 1979). Many research studies (e.g. the work of Almeida et al. (2003) and Waddell and Ulfarsson (2002a)) include road networks for urban and landscape development. In the case of Bangkok, many studies (e.g. the work of Chomchan et al. (1990) and Webster (2004)) conclude that roads play a major role for Bangkok development, channeling residential development away from the city into the suburbs. This study is based on the assumption that each road type influences each land use activity differently (Core Planning & Development, 2001). Proximity to roads is considered in relation to the three types of roads already described in Section 3.2.3: major roads, collector streets, and local streets as shown in Figure 3.24(a), (b) and (c) respectively.

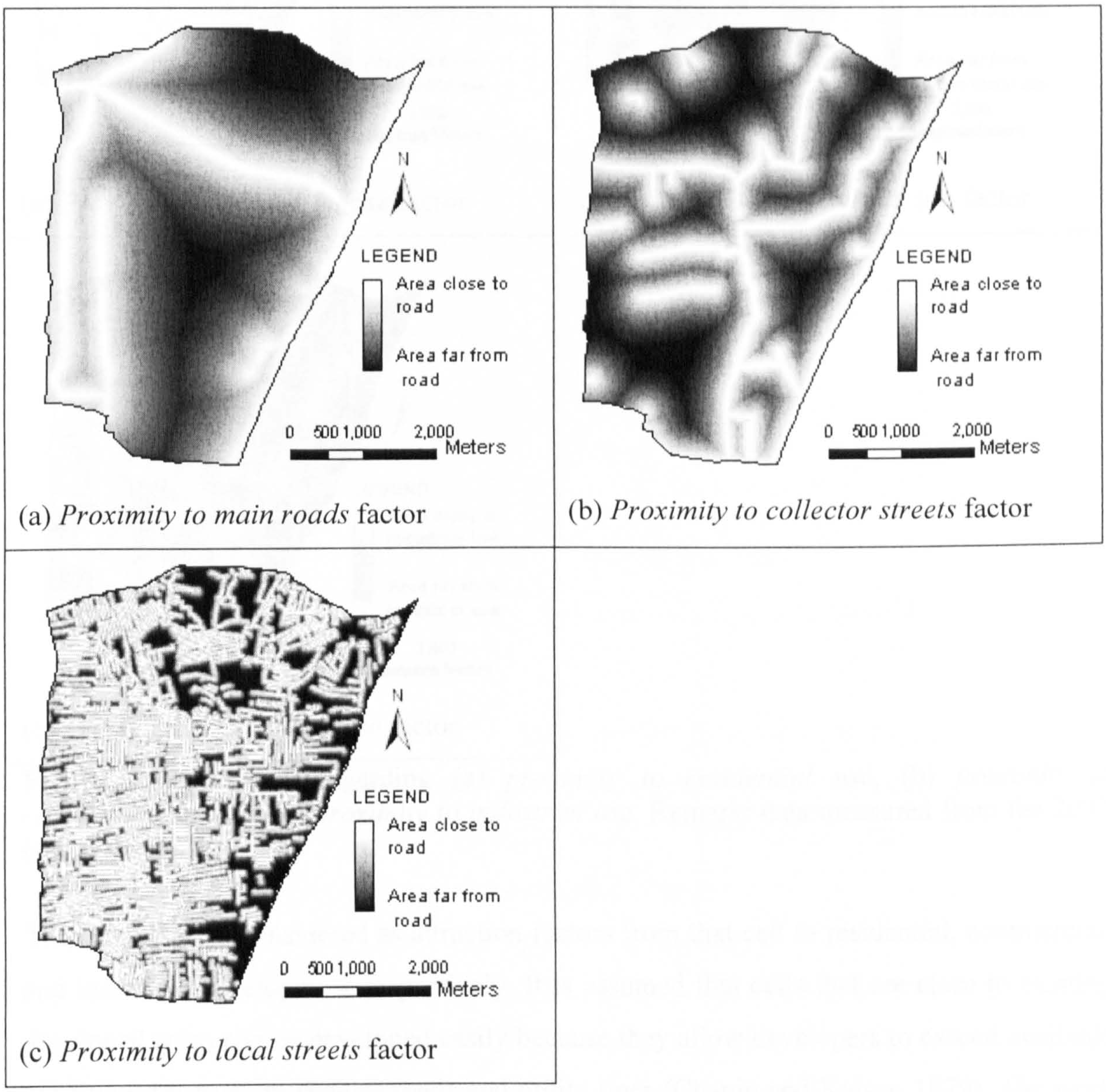
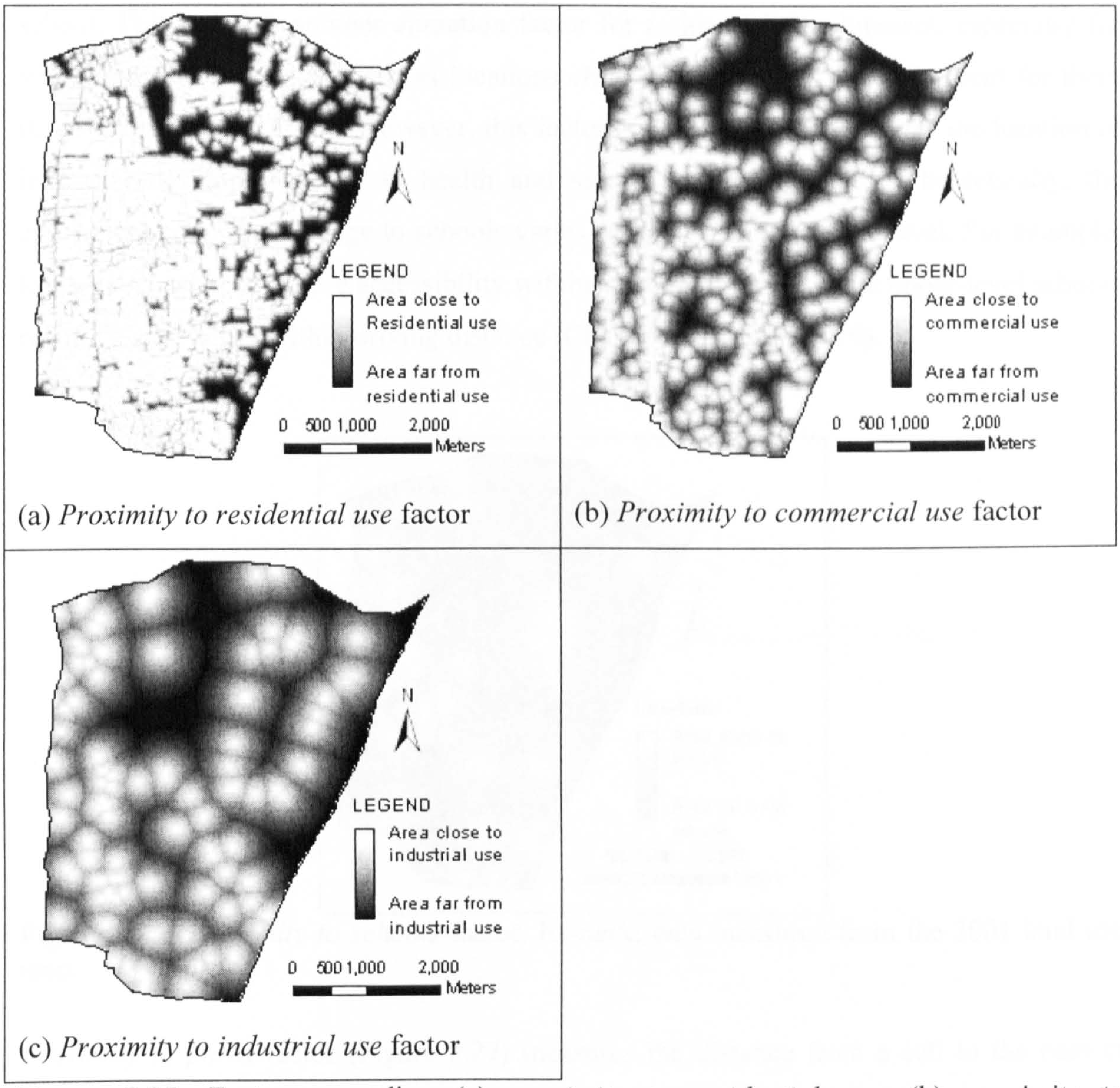


Figure 3.24: Three types of proximity to roads factors; (a) *proximity to major roads*, (b) *proximity to collector streets* and (c) *proximity to local streets*. Remark: data measured from the 2001 road map.

Proximity to residential use, proximity to commercial use and *proximity to industrial use* (see Figure 3.25(a), (b), (c) respectively) measure the distance from a cell to the nearest point with the given land use activity.



(a) *Proximity to residential use* factor

(b) *Proximity to commercial use* factor

(c) *Proximity to industrial use* factor

Figure 3.25: Factors regarding (a) *proximity to residential use*, (b) *proximity to commercial use* and (c) *proximity to industrial use*. Remark: data measured from the 2001 land use map.

These factors are considered as attraction factors from that cell to residential, commercial, and industrial development respectively. It is assumed that cells that are close to existing developed areas can be developed easily because they allow developers to extend available existing infrastructure such as roads and utility lines (Chapin and Kaiser, 1979). The work of Almeida et al. (2003), based on an empirical approach, is an example that includes proximity to residential, commercial, and industrial use for land use change simulations. *Proximity to residential use*, for example, has a positive effect for residential suitability due to health and safety reasons, however, *Proximity to industrial* use tends to have a negative effect for residential use but have a positive effect for industrial use.

Proximity to schools factor (Figure 3.26) measures the distance from a cell to the nearest school. This factor is another attraction factor for residential development, especially for young families having children as locations close to schools will be convenient for them (Chapin and Kaiser, 1979). However, this factor is considered repulsive for the location of industrial development due to health and safety reasons (op. cit.). Theoretically, the convenient commuting range to schools varies depending upon school level. For example, lower-level schools require accessibility within walking distance while upper-level schools require accessibility within driving distance (Chapin and Kaiser, 1979).

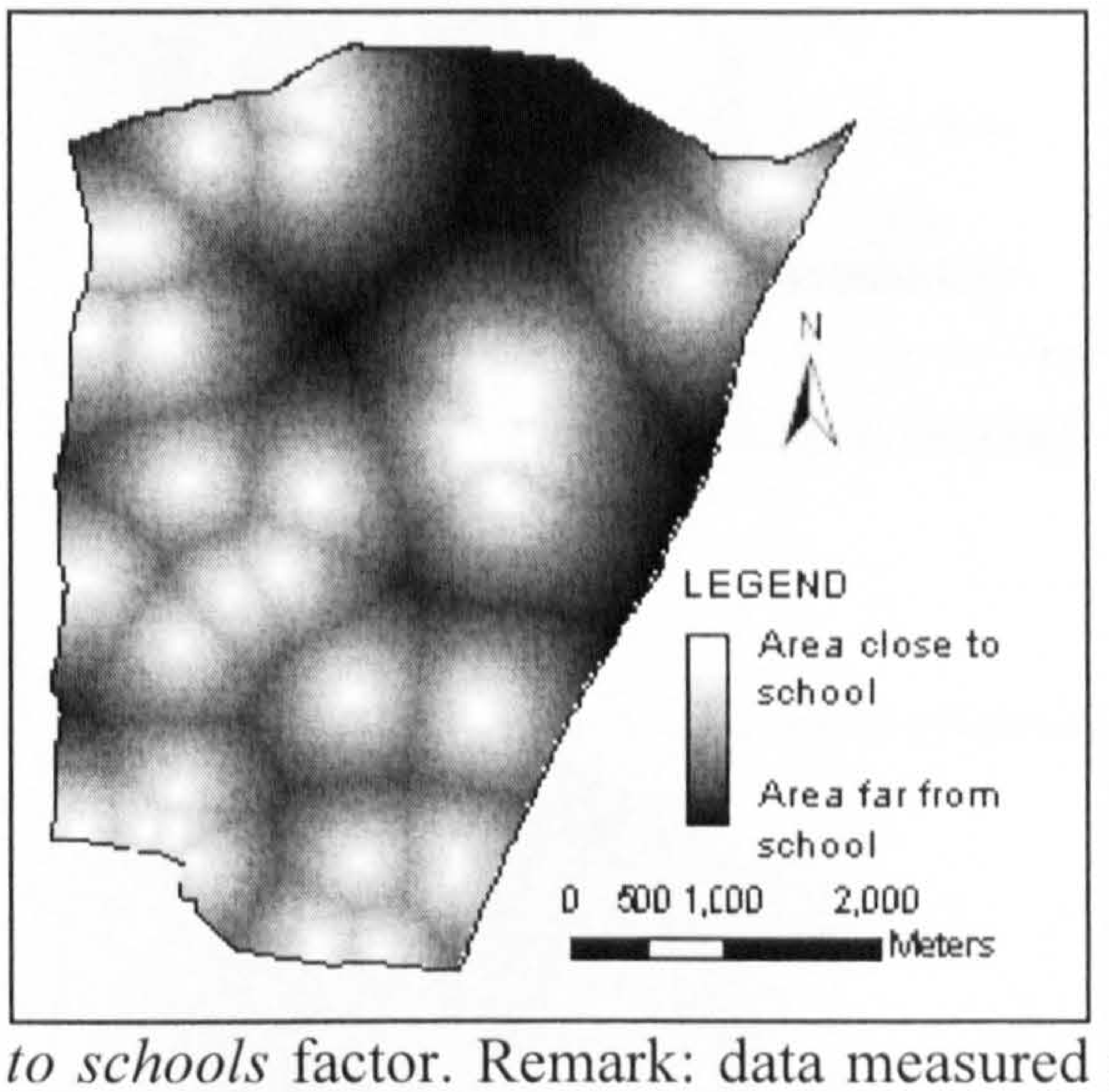


Figure 3.26: *Proximity to schools* factor. Remark: data measured from the 2001 land use map.

Proximity to parks factor (Figure 3.27) measures the distance from a cell to the nearest park, conservation or recreation area. This factor is another attraction factor for residential development. The work conducted by Almeida et al. (2003) is one example that includes this factor for the analysis of land use change simulations.

Proximity to government offices and public services factor (Figure 3.28) measures the distance from a cell to the nearest government office or public service facility. This factor is included in the model in response to the requirement of the Department of Town and Country Planning (DTCP)'s planning practice as specified in the Land Use Compatibility Matrix (see Table 4.7) in the '*Design and Development of GIS-based Planning Procedure (Executive Summary Report)*' conducted by ESRI (Thailand) Co. Ltd. (1997). According to the Land Use Compatibility Matrix, *Proximity to government offices and public services* is highly compatible to commercial use and moderately compatible to residential and industrial areas. This implies that this factor thus is considered a high attraction factor for commercial locations and a moderate attraction factor to residential and industrial locations.

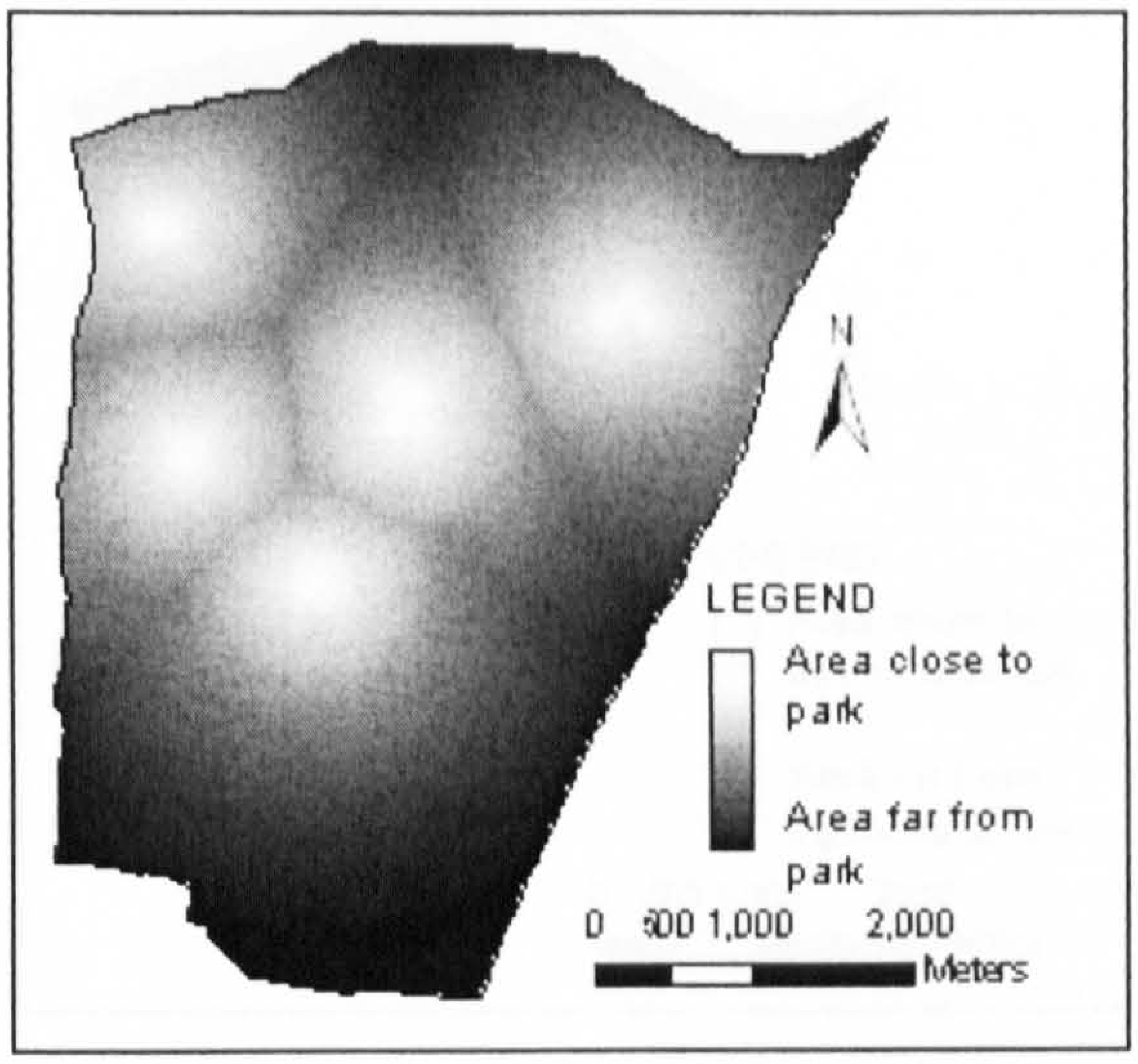


Figure 3.27: *Proximity to parks* factor. Remark: data measured from the 2001 land use map.

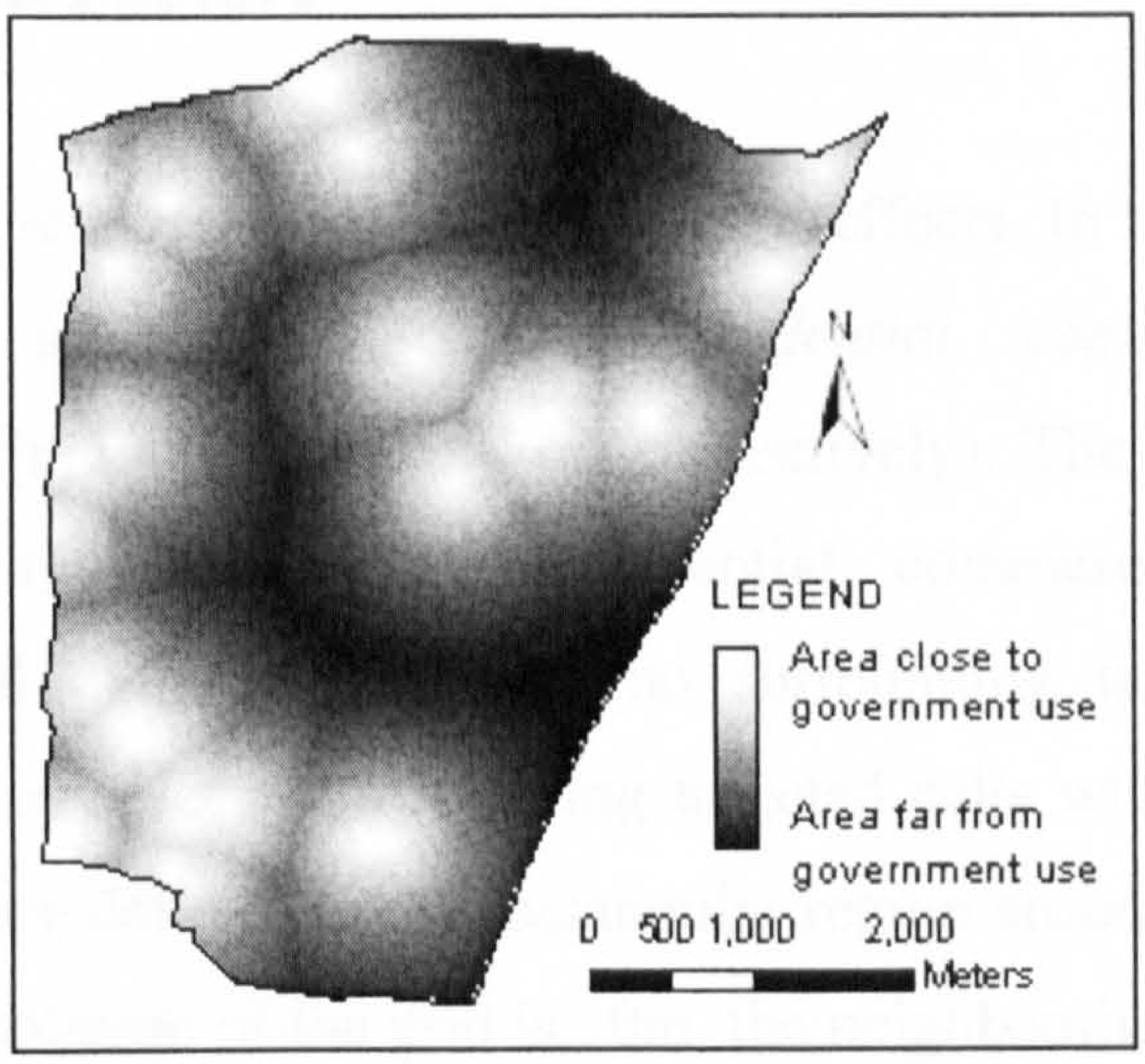


Figure 3.28: *Proximity to government offices and public services* factor. Remark: data measured from the 2001 land use map.

Proximity to agricultural use factor (Figure 3.29) measures the distance from a cell to the nearest agricultural area. This factor is required by Department of City Town and Planning (DTCP) as specified in the Land Use Compatibility Matrix (see Table 4.7) in the '*Design and Development of GIS-based Planning Procedure (Executive Summary Report)*' conducted by ESRI (Thailand) Co. Ltd. (1997). According to the Land Use Compatibility Matrix, *Proximity to Agricultural use* is moderately to highly compatible with industrial use and moderately compatible with residential and commercial areas. This implies that this factor is considered a moderately high attraction factor for industrial locations and a moderate attraction factor for residential and commercial locations.

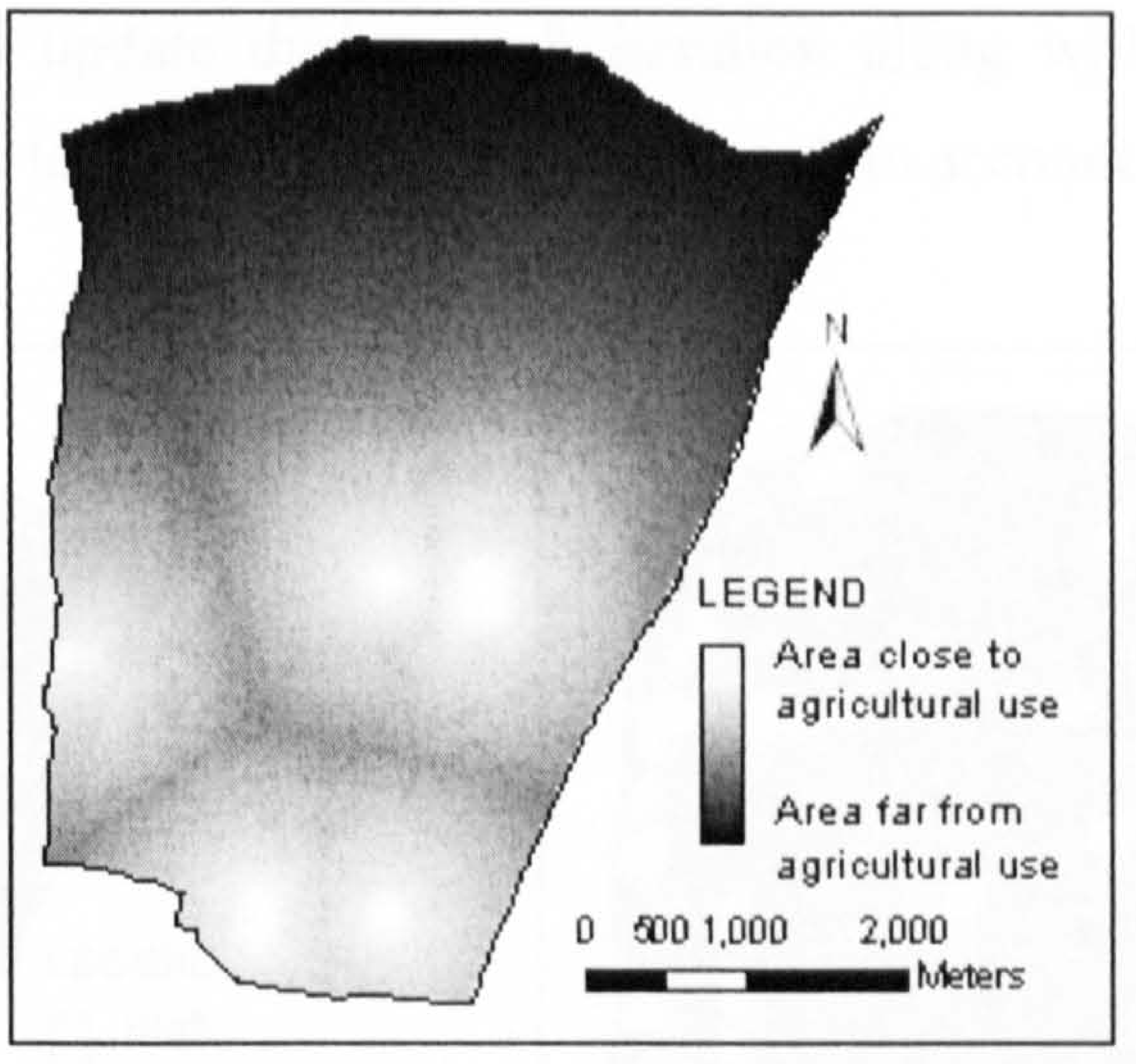
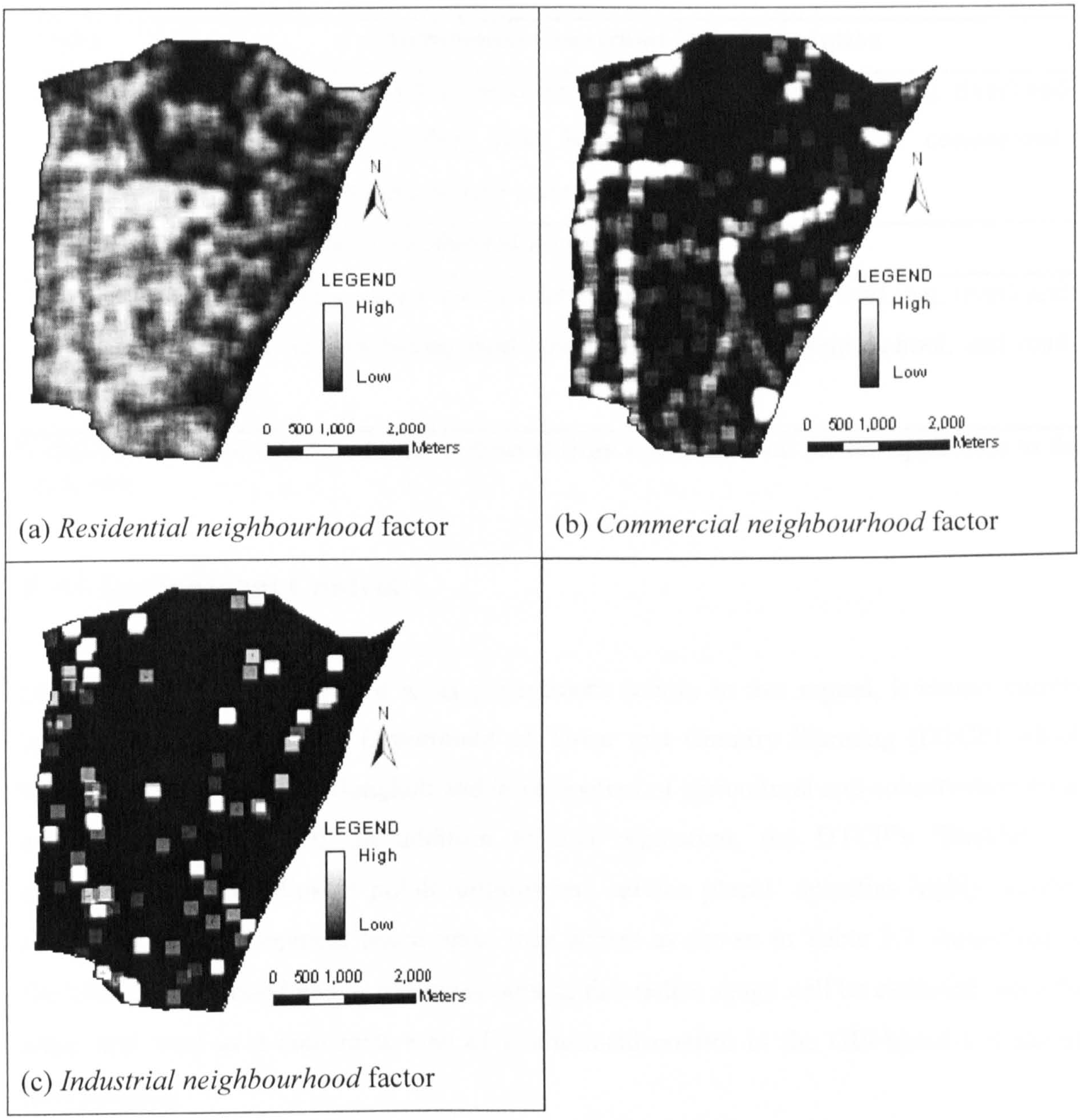


Figure 3.29: *Proximity to agricultural use* factor. Remark: data measured from the 2001 land use map.

3.4.2 Neighbourhood Factors

This group of factors is referred to as neighbourhood effects. In this study, three types of neighbourhood effects are used, including *residential, commercial* and *industrial neighbour factors* (see Figure 3.30(a), (b), (c) respectively). The neighbourhood effect is measured in terms of the proportion of residential, commercial and industrial cells surrounding a core cell ranging from 0.0 (no surrounding targeted cells within the neighbourhood) to 1.0 (completely surrounding targeted cells within the neighbourhood). Neighbourhood effects are defined as the rectangular region around the cell out to a radius of 10 cells. Since the resolution of the grid is 10m, the neighbourhood represents an area of 100 meters around a single cell, or 100m walking distance.

The neighbourhood effect encapsulates interaction among cells within a specified neighbourhood (Torrens, 2000). Neighbourhood effect can be used to reflect the effect of urban agglomeration (Chapin and Kaiser, 1979; Wu, 1998), compatibility among land use (Wu, 1998), or development density of land use change (Wu, 2002a). It represents the attraction and repulsion effect of a variety of states within the neighbourhood (Barredo et al., 2004). Cells located adjacent to existing developed areas can be developed more easily as developers may extend available infrastructure such as roads, water pipelines and electricity lines from adjacent cells. For example, a cell surrounded by many industrial cells tends to provide facilities that support industrial use rather than other uses. In this study, like most research work that applied the cellular automata for urban simulation, these factors are set to update during each iteration along with the simulation. Such operation thus allows for temporal changes to be taken into account.



(a) *Residential neighbourhood* factor

(b) *Commercial neighbourhood* factor

(c) *Industrial neighbourhood* factor

Figure 3.30: Neighbourhood factors: (a) *Residential neighbourhood* factor, (b) *Commercial neighbourhood* factor, (c) *Industrial neighbourhood* factor. Remark: data measured from the 2001 land use map.

3.4.3 Environmental Factors

In this study, environmental factors are used to set constraints for urban development. In this study, the exclusion of water bodies and roads will be set as a general constraint, these areas being considered undevelopable. Table 3.5 shows the rules used to create a general constraint map applicable to the study site for the whole simulation period (1993 – 2001). These three rules are combined together to represent general constraint areas to be used in the GIS-based CA model.

Rules	Environmental Constraint Rule Description
Rule 1	Land use change is not allowed for existing 1993 natural land (e.g. river) and developed land including land being used for residential, commercial, industrial, government, school and road purposes.
Rule 2	Land use change is not allowed for existing 1998 roads.
Rule 3	Land use change is not allowed for existing 2001 natural land (e.g. river) and developed land including land being used for government, school, and road purposes.

Table 3.5: General constraint rules extracted from environmental factors applicable to the study site.

3.4.4 Institutional Control

Institutional control is referred to as government action. In this regard, it means current planning regulation by the Department of Town and Country Planning (DTCP) which imposes a restriction over Bangkok and development of agricultural and conservation areas as shown in Table 3.6. In addition to this regulation, the DTCP's 'Standard of accessibilities and radius of public utilities and service places' specifies highly suitable area for urban development based upon such access as shown in Table 3.7. According to the restriction outlined above, the areas outside this radius range will be excluded from the study and used as a constraint map of institutional control in the GIS-based CA model development.

Rules	Institutional Control Rule Description
Rule 1	Land use change is not allowed for existing 1993 agriculture and conservation areas
Rule 2	Land use change is not allowed for existing 2001 agriculture and conservation areas

Table 3.6: Planning constraint rules extracted from institutional control factors applicable to the study site.

Public Utilities and Public Services	Distance or radius of service (units: Meters)
Schools, Colleges and Universities	≤ 2000
Park/conservation and recreation areas	≤ 2000
Commercial, market places	≤ 3000
Government offices and public utilities and services (e.g. police stations, fire brigades)	≤ 5000
Roads	
- Expressways, major roads and secondary roads	≤ 1000
- Collector streets	≤ 500
- Local streets	≤ 300

Table 3.7: Standard of accessibilities and radius of public utilities and service places (source: DTCP, extracted from ESRI (Thailand) Co. Ltd. (1997)).

According to the environmental factors (Section 3.4.3) and institutional control (Section 3.4.4) set, the constraint area map (see Figure 3.31) is created.

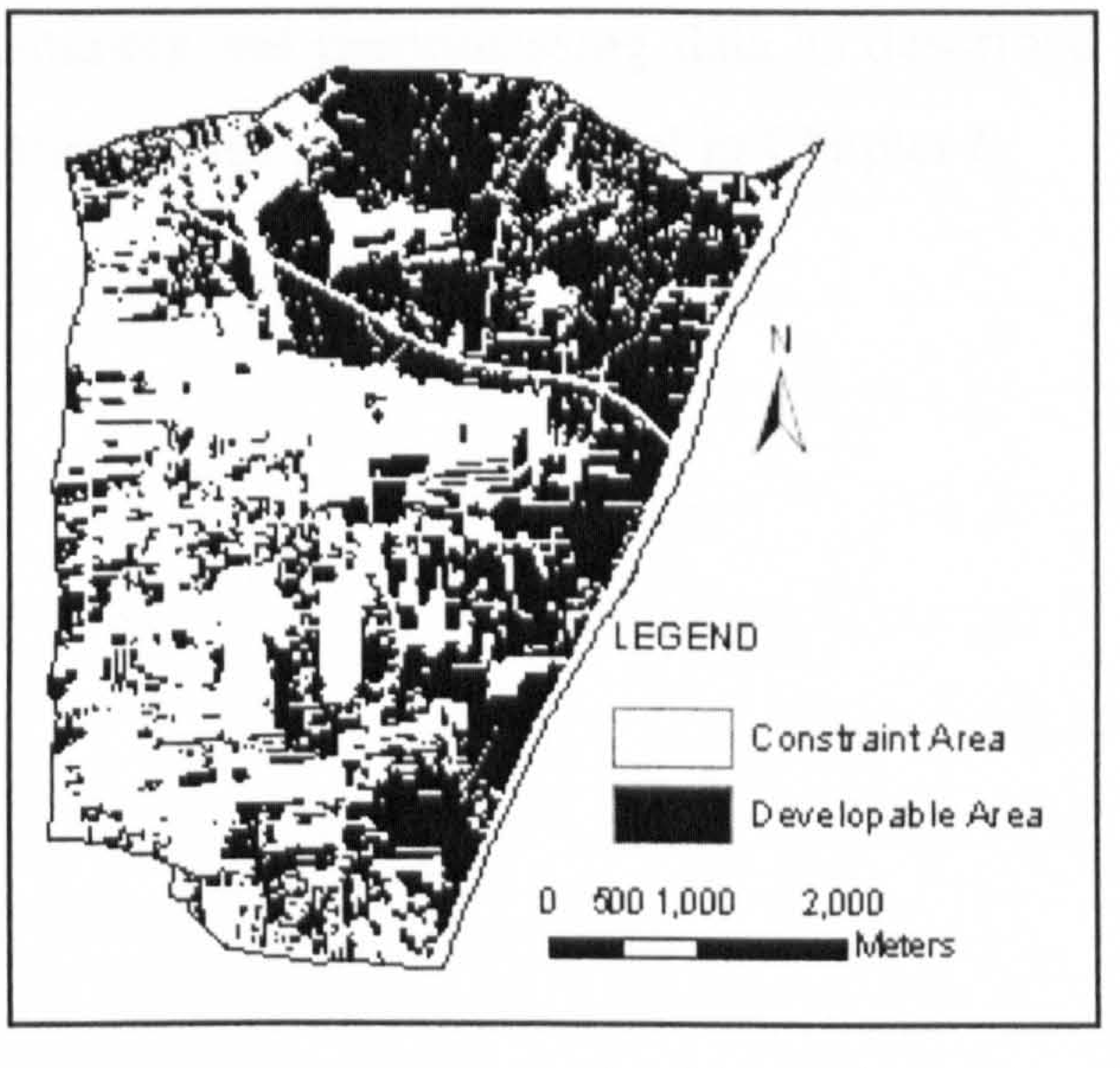


Figure 3.31: The 2001 constraint area map applicable to the study site.

This map will be used for the simulation from 1993 to 2001. According to the figure (Figure 3.31), only developable areas show in black while areas restricted for development show in white.

3.5 Conclusion

This chapter provides background information on the historical, demographic and physical characteristics as well as the role of land use planning in the context of Bangkok and Ladprao district. In the second section of this chapter, datasets used and data preprocessing manipulated are introduced and described. In the third section, data constraint issues due to limited data availability and a lack of detailed data for the study area are discussed. Data constraints have delimited the availability of development factors to be used in the model. In the final section, development factors to be used for the implementation of model in the remainder of this thesis are described.

Ladprao is chosen in this research study to represent one of Bangkok areas that has experienced uncontrollable land usage. Since Ladprao has vacant land covering an area of about 50% of total land left for horizontal development, it shows potential to be transformed and developed for various kinds of urban activities. As the agencies responsible for land planning in the Bangkok Metropolitan currently lack a planning tool to locate appropriate areas to direct the growth to specific locations, therefore, despite the data poor environment, sensible to develop planning mechanisms that can be used to understand, evaluate and monitor the dynamics of urban development applicable to the Bangkok region. Data sources and preprocessing data as described in this chapter will be used for the implementation of the proposed model in Chapter 6.

CHAPTER 4

Methodology Framework

4.1 Introduction

As discussed in Chapter 3, particular concern with urban development in Thailand focuses on the growth of Bangkok. Its continual growth has led to an unsystematic and uncontrollable pattern of land use, primarily as a result of the soaring rate of the economic boom in the 1970s and 1980s and ineffective land use planning policies (Chomchan et al., 1990; Kaothien, 1995). Despite these obvious problems, little consideration has been given to how to predict land use change for the Bangkok area (e.g. the work of Bruijn (1991)). This research was carried out as an attempt to develop an urban simulation tool that can be used to predict the future growth and, to some extent, help reflect on the consequences of previous and current planning policies in order to understand and control urban growth and developments as an aid to city planners. In this chapter, the work undertaken to develop a GIS-based cellular automata (CA) model for simulating patterns of urban land use change on the basis of micro-simulation for a part of Bangkok area is reported.

Figure 4.1 illustrates the overall conceptual methodology developed for the research, where it can be broken down into three main stages: data input, model development, simulation output and evaluation. This figure illustrates that the data input from various sources as well as the development factors chosen in Chapter 3 will be evaluated using the multinomial logistic regression (MNL) and multi-criteria decision analysis (MCDA) methods to assess land potential for development. This information will be combined with the land use data and used in a CA approach. The CA approach designed under a GIS platform will be developed to simulate the yearly land use pattern until the end year specified.

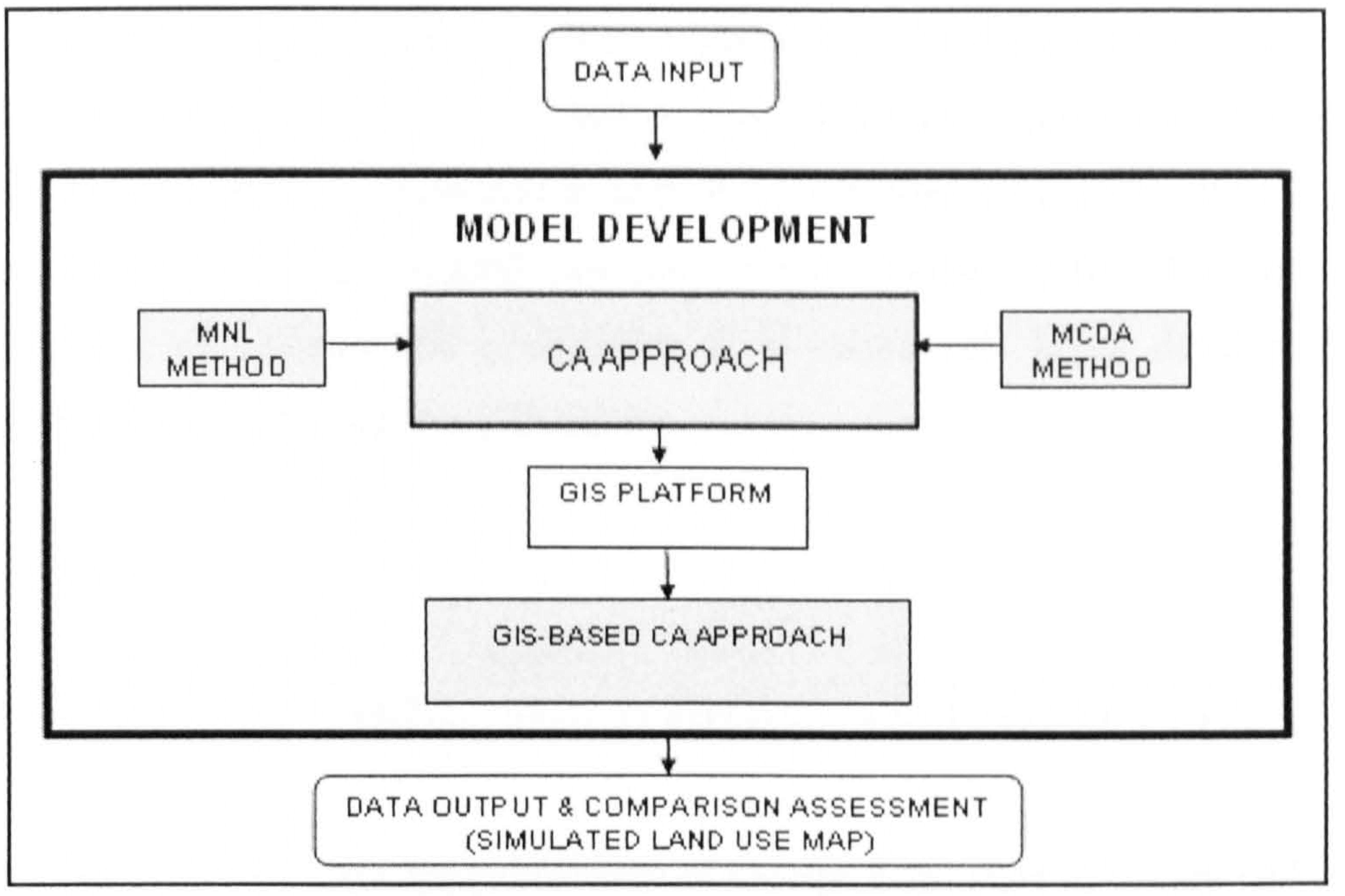


Figure 4.1: Schematic diagram of the modelling framework developed for the study.

The organization in this chapter will be as follows. The first section of this chapter (Section 4.2) describes the multinomial logistic regression (MNL) method. It is followed by the description of the multi-criteria decision analysis (MCDA) method (Section 4.3). In Section 4.4, the development of the formulation of the CA approach for the research study is presented. In the final section, the description of the derivation of weights based on both MNL and MCDA methods as well as the derivation of neighbourhood thresholds is presented.

4.2 Multinomial Logistic Regression (MNL) Method

Multinomial logistic regression (MNL) is one of the most popular methods in social science research (Powers and Xie, 2000) as it allows one to predict a discrete choice or outcome from a set of independent variables. In urban applications, MNL has been recently proven a successful method to be applied for calculation of the probability of land development. For example, the research work that developed the UrbanSim model (e.g. Waddell (1998) and Waddell and Ulfarsson (2002a)) used multinomial logistic regression as an approach for computing location choices for residential and industrial development respectively. However, their work did not undertake simulation using a dedicated graphical output and therefore no maps were created directly.

In this study, types of land use transition, regarded as the change from vacant (l_{VC}) to residential (l_{RS}), commercial (l_{CM}), industrial (l_{MA}), and vacant (no change) (l_{VC}), can be viewed as choices. MNL is applied to predict the probability of choices (that is, a vacant cell being developed to those targeted land use types). By adapting the MNL to the task of predicting land use change, the probability of a vacant cell being developed can be specified after Wu (2000) and Wu (2002a) as:

$MNL\_P(l_{VC} \rightarrow l_u) =$

$$\frac{\exp(Z(l_{VC} \rightarrow l_u))}{\exp(Z(l_{VC} \rightarrow l_{RS}) + \exp(Z(l_{VC} \rightarrow l_{CM}) + \exp(Z(l_{VC} \rightarrow l_{MA}) + \exp(Z(l_{VC} \rightarrow l_{VC})} \tag{4.1}$$

in which MNL_P($l_{VC} \rightarrow l_u$) is the probability of change from state of vacant cell only (l_{VC}) to a land use type (l_u) being considered. l_u can be residential (l_{RS}), commercial (l_{CM}), industrial (l_{MA}), or vacant (no change) (l_{VC}), and Z($l_{VC} \rightarrow l_u$) is the linear combination of independent variables (attributes) that describes the attractiveness of land transition from vacant to a particular land use type (l_u),

$$Z(l_{VC} \rightarrow l_u) = a(l_{VC} \rightarrow l_u) + \sum_k \beta_k (l_{VC} \rightarrow l_u)\, x_k \tag{4.2}$$

where a is a constant of a particular land use type (l_u), l_u can be residential (l_{RS}), commercial (l_{CM}), industrial (l_{MA}), and vacant (no change) (l_{VC}), β_k are coefficients or weights of the regression model associated with a type of land use transition ($l_{VC} \rightarrow l_u$) according to the k variables, and x_k is a set of k independent variables or attributes.

The resultant value from Equation 4.1 varies between 0.0 and 1.0 and can be interpreted in terms of the probability of transition of land use. The following example illustrates how the probability of land use transition in Equation 4.2 can be written. Based on the coefficients listed in Table 4.4, the probability of vacant transition to residential of each single grid cell, MNL_P($l_{VC} \rightarrow l_{RS}$), is calculated from Equation 4.1. (Note that the actual weight values and the factors chosen are examples taken from Table 4.4 which presents real world data). Their terms are written as follows:

$Z(l_{VC} \rightarrow l_{RS})$ = -51.625 +0.290*LP – 0.131*D_RDT1 + 1.887*D_RDT2 + 46.234*D_RDT3 - 0.678*D_GV + 2.018*D_SCH - 0.381*D_PRK

in which $Z(l_{VC} \rightarrow l_{RS})$ refers to a linear combination that describes the attractiveness of vacant transition to residential, and

$Z(l_{VC} \rightarrow l_{CM})$ = -2.519 +0.424*LP + 2.564*D_RDT1 + 5.286*D_RDT2 – 14.799*D_RDT3 - 2.620*D_GV – 0.061*D_SCH – 1.885*D_PRK

in which $Z(l_{VC} \rightarrow l_{CM})$ refers to a linear combination that describes the attractiveness of vacant transition to commercial, and

$Z(l_{VC} \rightarrow l_{MA})$ = -21.301 +0.637*LP + 0.933*D_RDT1 + 4.292*D_RDT2 + 7.403*D_RDT3 + 0.313*D_GV – 0.857*D_SCH – 1.617*D_PRK

in which $Z(l_{VC} \rightarrow l_{MA})$ refers to a linear combination that describes the attractiveness of vacant transition to industrial, and

$Z(l_{VC} \rightarrow l_{VC})$ = 0.0 +0.0*LP + 0.0*D_RDT1 + 0.0*D_RDT2 + 0.0*D_RDT3 + 0.0*D_GV + 0.0*D_SCH + 0.0*D_PRK

and refers to a linear combination that describes the attractiveness of remaining unchanged, and in this case is the base line value.

Consequently, the probabilities of each single grid cell changing from vacant to residential $P(l_{VC} \rightarrow l_{RS})$, vacant to commercial $P(l_{VC} \rightarrow l_{CM})$, vacant to industrial area $P(l_{VC} \rightarrow l_{MA})$ or vacant to vacant (no change). $P(l_{VC} \rightarrow l_{VC})$ are statistically calculated using the established multinomial logistic regression Equation 4.1.

4.3 Multi-criteria Decision Analysis (MCDA) Method

Multi-criteria decision analysis (MCDA) has recently been integrated with the cellular automata (CA) approach for urban simulation, especially in the context that participants or experts' judgements are included to help set the relative importance of weights for either multiple criteria or multiple objectives. An example work conducted by Wu and Webster (1998) in Guangzhou city of China showed the potential of simulating land development by integration of GIS-based CA approach and multi-criteria decision analysis (MCDA), which allows opinions and preferences of planners to be incorporated for testing different sets of development regimes.

In this study, MCDA is used for the task of computing land suitability. Each suitability map is created on the basis of relationship between weights derived from multi-attribute decision analysis (MADA) and weights derived from multi-objective decision analysis (MODA). Multi-attribute decision analysis (MADA) is constructed to create suitability scores based on the relative importance of multiple attributes while multi-objective decision analysis (MODA) is integrated to apply a set of weights based on the relative importance of competitive objectives. The competitive objectives, in this study, refer to the weights assigned for the competing of land transitions (e.g. the vacant change to residential land). A weighted summation technique is applied to both MADA and MODA techniques in order to compute a suitability score which in turn define the potential of land development.

Creation of a MADA suitability score in this study is adapted after Malczewski (1999b; 1999a), given by:

$$MADA\_SU(l_{VC} \rightarrow l_v) = \sum_{m=1}^{M} x_m(l_{VC} \rightarrow l_v) * W_m(l_{VC} \rightarrow l_v) \tag{4.3}$$

where,

$MADA\_SU(l_{VC} \rightarrow l_v)$ = the suitability score created for a vacant (l_{VC}) transition to a particular land use type (l_v);

$x_m(l_{VC} \rightarrow l_v)$ = the attribute value of variable m for a vacant (l_{VC})

		transition to a particular land use type (l_v);
$W_m(l_{VC} \rightarrow l_v)$	=	the weight of the relative important of variable m for a vacant (l_{VC}) transition to a particular land use type (l_v);
m	=	the variables or development factors used in the model, and
l_v	=	a particular land use type. In this equation, l_v can be residential (l_{RS}), commercial (l_{CM}), or industrial (l_{MA}).

An example in Figure 4.2 illustrates how the suitability score of residential and industrial based on the same cell can be calculated. According to the figure, three attributes regarding land price, distance to road and distance to school are considered. After normalization, these factors hold the values of 0.5, 0.6, and 0.7 respectively. Each land use type has different sets of criteria and different sets of relative weights. However, each set of weights has a sum of 1.0. In the case of residential, the suitability score is created based on three criteria; land price, distance to road, and distance to school factor.

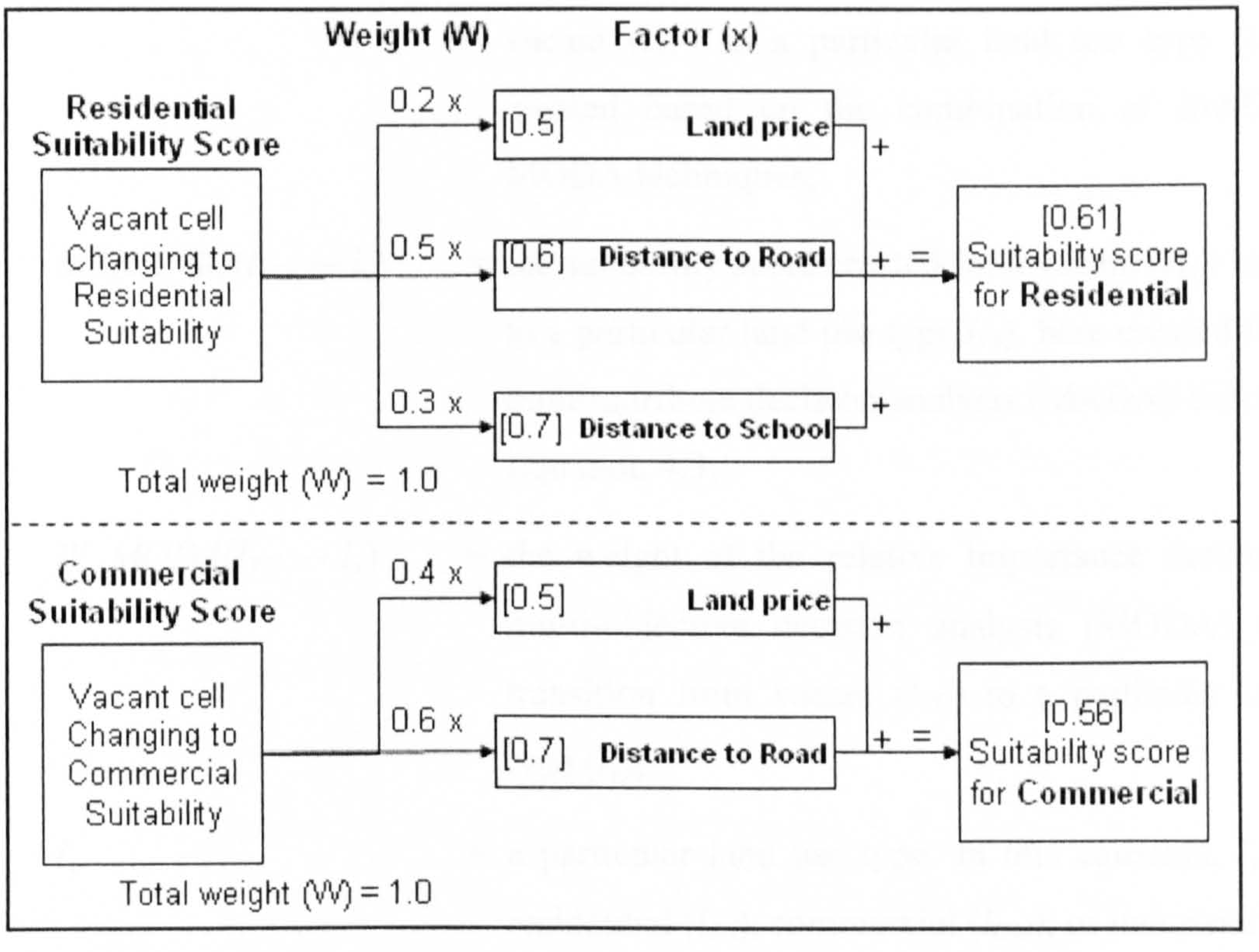


Figure 4.2: An example of suitability score calculation for residential use and commercial use using the multi-attribute technique described in Equation 4.3.

Suppose that the criterion weights $W_m(l_{VC} \rightarrow l_{RS})$, derived from decision-makers, are set as 0.2, 0.5 and 0.3 respectively. Based on Equation 4.10, these weights are multiplied by the corresponding attribute values and then those attribute scores are additively summed (0.2 x 0.5 + 0.5 x 0.6 + 0.3 x 0.7). As a result, it computes a residential suitability score of 0.61. In the case of commercial, a similar procedure is applied, however, only two development factors with the set of relative weights differing from that of residential described above are considered. These two factors can have differing weights to their use in calculating residential suitability, and these weights must total 1.0. It finally created a commercial suitability score of 0.56.

Incorporation of a MADA suitability score with the weights derived from multi-objective decision analysis (MODA) is adapted after Malczewski (1999a) and (1999b), given by:

$$MCDA\_SU(l_{VC} \rightarrow l_v) = MADA\_SU(l_{VC} \rightarrow l_v) * W\_MODA(l_{VC} \rightarrow l_v) \tag{4.4}$$

where,

$MCDA\_SU(l_{VC} \rightarrow l_v)$	= the final suitability score of land transition from vacant (l_{VC}) to a particular land use type (l_v), here created based on the combination of MADA and MODA techniques;
$MADA\_SU(l_{VC} \rightarrow l_v)$	= the suitability score created for a vacant (l_{VC}) transition to a particular land use type (l_v), here created from the multi-attribute decision analysis (MADA) described in Equation 4.3;
$W\_MODA(l_{VC} \rightarrow l_v)$	= the weight of the relative importance derived from multi-objective decision analysis (MODA) of land transition from vacant (l_{VC}) to a particular land use type (l_v);
l_v	= a particular land use type. In this equation, l_v can be residential (l_{RS}), commercial (l_{CM}), or industrial (l_{MA}).

4.4 Formulation of the CA-based Modelling Framework

The MNL and MCDA methods described in Section 4.2 and Section 4.3 are integrated to the urban CA model developed for the research study. In this section, formulation of the CA model developed is presented.

The study area (space) is represented as a rectangular grid of square cells where each cell has an area of 10m x 10m, or 100 sq.m. This cell size has been chosen in order to accommodate the minimum area of approximately one detached house in the urban areas in the study site (Department of Town and Country Planning (DTCP), 1997). Each cell is characterised by a vector of attribute layers. Each attribute layer can represent the physical characteristics (e.g. distance to main roads), environmental (e.g. exclusion of water bodies) or institutional factors (e.g. exclusion of agricultural land).

Each cell in a CA model at any point in time has a current state. In this study, the model uses 10 cell states, representing the land use types into which the study site is classified. Six of these land use types represent fixed features in the model in the sense that they are assumed not to change and they thus do not take part in the dynamic simulation (Barredo et al., 2003; White et al., 2000). They include schools, government and public utilities, conservation areas such as parks, rivers, roads and agricultural areas. The active features are the four land use types. They are residential, commercial, industrial, and vacant (non-developed) land. In this study site, however, industrial land represents the area where only small industrial units such as garages, small furniture industries are located in order to serve the local community. It thus cannot fully represent the true characteristics of the industrial land and can be considered a less important observation in the study.

Each cell location can be identified by its X,Y coordinates which are expressed in the form *i,j*. Thus, the state of each cell in terms of the ten land use types is expressed at a particular point in time as $S^t_{i,j}(l_p)$, where l_p can be residential referred to as l_{RS}, commercial referred to as l_{CM}, industrial referred to as l_{MA}, government referred to as l_{GV}, school referred to as l_{SCH}, conservation referred to as l_{PRK}, agriculture referred to as l_{AG}, road referred to as l_{RD}, river referred to as l_{RV}, or vacant (non-developed) area referred to as l_{VC}.

The cell neighbourhood comprises a window of preset size and shape that aims to capture the spatial influence on cell evolution at each iteration of the model (Li and Yeh, 2000; Yeh and Li, 1998). In this study, the neighbourhood radius is 100m. This value is often used by Thai urban planners (Department of Town and Country Planning (DTCP), 1997) to determine the maximum walking distance within an area related to city dwellers' activities in their neighbourhood. Since the grid cell has a 10m resolution, a window size of 21x21 around a centre cell is used.

When CA is applied to the urban simulation, an additional element regarded as threshold setting (so-called land use demand) is required in order to set a limit of urban growth (Barredo et al., 2004; Engelen et al., 1995). This can be done by specifying a number of cells or total area required for a particular land use type in the model (op. cit.). Such a value is determined by decision-makers or experts to control the overall growth of a city (White et al., 2000). In this study, land use demand is obtained from the total area of change observed from the available land use maps for the period of interest (1993 to 2001). The quantity of land use conversion at each iteration of the model is constrained to be the total land change between 1993 and 2001 divided by the time period. The threshold setting is mainly focused on three types of change from vacant (non-developed) area to residential, commercial, and industrial. Table 4.1 lists the actual change obtained from the study site. Based on the 10m grid cell resolution, the total area required for land transition is to ensure 18,242 vacant to residential changed cells, 2,802 vacant to commercial and 349 vacant to industrial.

Description of Change	**Total amount of area change (sq. km.)**	**Total amount of cell change between 1993-2001 based on 10x10m grid cell**
Vacant to Residential	1.82	18,242
Vacant to Commercial	0.28	2,802
Vacant to Industrial	0.03	349

Table 4.1: Thresholds setting for total amount of cell change, based on change observed between 1993-2001 at a resolution of 10x10m grid cell.

4.4.1 Formulation of Transition Rules

In this study, the transition rules that drive the model are divided into two; the constraint rule and the development probability rule. The constraint rule is used to exclude the undeveloped areas in the study site from the simulation. The constraint rule is developed based on the combination of the environmental factors and institutional control as described in Sections 3.4.3 and 3.4.4 respectively. Based on this rule, only cells without constraints on their development are developed.

The 'development probability' rule is best explained with reference to urban simulation using a CA approach, built on the linkage between global patterns and local processes (Batty and Xie, 1994). Following the CA concept, this rule is built applying the notion of Wu and Webster (1998) and Wu (2002a). Global patterns are used to capture the aggregate impact of land development. They refer to the relationship of factors of land development observed in the macroscopic view. These patterns are measured and quantified as development factors in an aggregate level in terms of a set of weights. These weights are estimated globally to represent the physical characteristics of the whole land development system under study. In this study, global factors used are physical characteristics (e.g. land price, proximity to roads) given in Section 3.4.1. The global patterns are developed by using the multinomial logistic regression (MNL) method and multi-criteria decision analysis (MCDA) method. They are used to calculate the probability and suitability of land development respectively.

Torrens (2000) and Wu (2002a) criticized that the simulation results generated from the urban spatial modelling that applies the global factors only (e.g. the traditional spatial interaction model) do not take the time aspect into account, the results thus are considered static in time. Wu (2002a) suggests incorporating another type of factor (so-called local processes) in order to capture the heterogeneity or uneven nature of land development through time process. Usually, the local process is developed through the iterative effect of neighbourhood. This type of factor is considered dynamic as it is forced to change along with the simulation over time. The integration between global factors and local processes (see Figure 4.3) plays a major role for the development of the CA transition rule in the study.

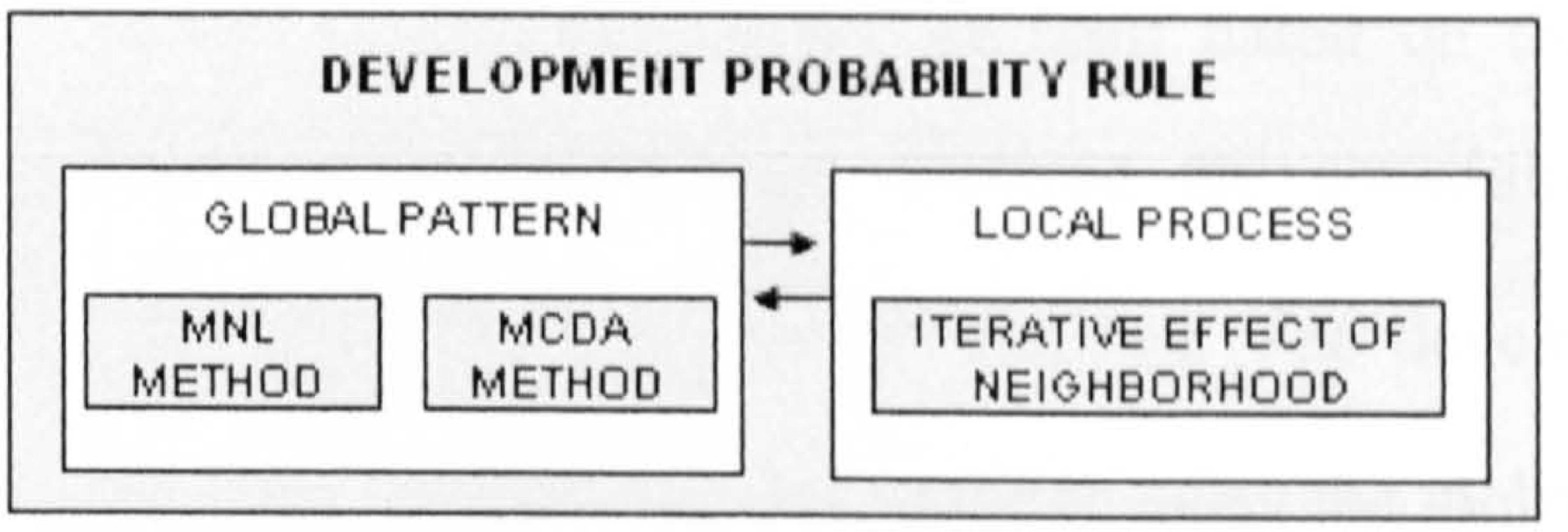


Figure 4.3: The 'development probability' rule, built under the urban CA concept, developed for the study.

Figure 4.4 illustrates the Input/Output diagram of the transition function developed in the research study. The land development simulation can be described as the state of land use at time *t+1* ($S_{i,j}^{t+1}$), determined by the state of land and the development situation at time *t* ($S_{i,j}^{t}$) in respect to a set of transition rules. Although the state of cells at time *t* ($S_{i,j}^{t}$) and *t+1* ($S_{i,j}^{t+1}$) can be of any land use type, in this study the output state of cell ($S_{i,j}^{t+1}$) is permitted for the change from the input state at time *t* ($S_{i,j}^{t}$) of vacant to only the considered land use types regarding residential, commercial, industrial, and no change. Thus, throughout this study the state of cell at time *t* ($S_{i,j}^{t}$) is considered as vacant only, while the output state at time *t+1* ($S_{i,j}^{t+1}$) can be one of the four possible land use types, where l_u refers to residential (l_{RS}), commercial (l_{CM}), industrial (l_{MA}) and vacant (l_{VC}) type.

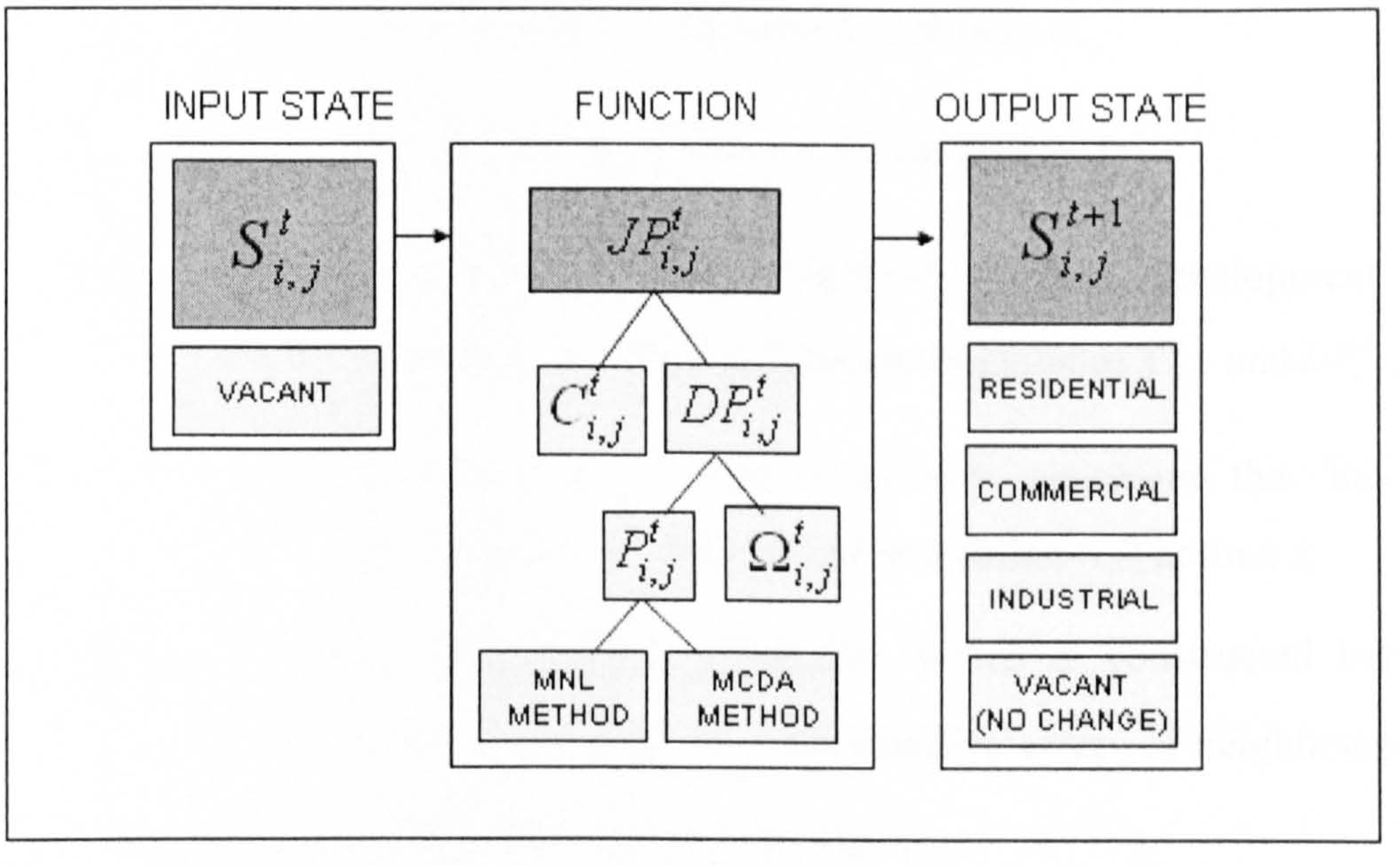


Figure 4.4: The I/O diagram of the transition function applied to the study.

In Figure 4.4, the function of transition rules are built based on the concept of joint probability ($JP^t_{i,j}$), which is the integration between the constraint rule ($C^t_{i,j}$) and development probability rule ($DP^t_{i,j}$) previously outlined. The development probability rule ($DP^t_{i,j}$) is built under the concept of global patterns using the global probability ($P^t_{i,j}$) (through MNL and MCDA method) and local processes through iterative effect of neighbourhood ($\Omega^t_{i,j}$).

A simplified rule-based structure of the transition functions described above, thus, can be expressed after Wu and Webster (1998) as:

$$S^{t+1}_{i,j} = f(S^t_{i,j}, JP^t_{i,j}) \tag{4.5}$$

$$\text{where } JP^t_{i,j} = f(C^t_{i,j}, DP^t_{i,j}) \tag{4.6}$$

$$\text{where } DP^t_{i,j} = f(P^t_{i,j}, \Omega^t_{i,j}) \tag{4.7}$$

Thus, Equation 4.5 can be equally written as

$$S^{t+1}_{i,j} = f(S^t_{i,j}, C^t_{i,j}, P^t_{i,j}, \Omega^t_{i,j}) \tag{4.8}$$

where,

$S^{t+1}_{i,j}$ = the state of cell at location (*i,j*) at time *t*;

$S^t_{i,j}$ = the state of cell at location (*i,j*) at time *t+1*;

$JP^t_{i,j}$ = the product of joint probability of land development which is constructed by the combination function of $C^t_{i,j}$ and $DP^t_{i,j}$;

$C^t_{i,j}$ = the presence or absence of any constraint that hinders new development to occur of cell at location (*i,j*) at time *t*;

$DP^t_{i,j}$ = the development probability, which is constructed between the global probability $P^t_{i,j}$ and iterative effect of neighbourhood $\Omega^t_{i,j}$ of cell at location (*i,j*) at time *t*;

$P^t_{i,j}$ = the global probability of land development of cell at location (*i,j*) at time *t*,

$\Omega^t_{i,j}$ = a neighbourhood of cells at location (i,j) at time t;

f = a transition function that defines the change of the state t to $t+1$.

4.4.1.1 Development of Joint Probability Using the MNL Method

Based on a rule-based structure of the transition functions as described in Equation 4.5, a joint probability $JP^t_{i,j}$ on the basis of the MNL method can be expressed after Wu (2002a) as:

$$JP^t_{i,j}(l_{VC} \rightarrow l_v) = P^t_{i,j}(l_{VC} \rightarrow l_v) * C^t_{i,j} * \Omega^t_{i,j}(l_{VC} \mid l_v) \tag{4.9}$$

where,

$JP^t_{i,j}(l_{VC} \rightarrow l_v)$ = the joint probability of land development from vacant (l_{VC}) to a particular land use type (l_v) at location (i,j) at time t;

$P^t_{i,j}(l_{VC} \rightarrow l_v)$ = the global development probability of land transition from vacant (l_{VC}) to a particular land use type (l_v) at location (i,j) at time t, here created based on the concept of MNL;

$C^t_{i,j}$ = the Boolean value of the presence of any constraint that hinders new development to occur at location (i,j) at time t;

$\Omega^t_{i,j}(l_{VC} \mid l_v)$ = the neighbourhood index value of the vacant cell (l_{VC}) at location (i,j) at time t where the cell itself plus the surrounding land use (l_v) cells are measured;

l_v = A particular land use type (l_v) where l_v can be residential (l_{RS}), commercial (l_{CM}), or industrial (l_{MA}).

In Equation 4.9, the global development probability $P^t_{i,j}(l_{VC} \rightarrow l_v)$ is computed using the MNL method described in the general case in Equation 4.1. Based on the MNL method, four types (l_u) of development probability of land transition from vacant to residential, commercial, industrial and vacant (no change) are created. However, since in this study the focus is on three types of transition (l_v) regarding vacant to residential, commercial and industrial, only three land transition types of the MNL products are input into Equation 4.9.

The neighbourhood function of cell transition $\Omega^t_{i,j}(l_{VC} \mid l_v)$ in Equation 4.9 presents the neighbourhood potential of the cell with a value from 0.0 (no potential) to 1.0 (highest potential). The neighbourhood function can be given after Wu (2002a) by

$$\Omega^t_{i,j}(l_{VC} \mid l_v) = \frac{\sum_{(m\times m)-1} con(s^t_{q,r} = l_v)}{(m\times m)-1} \tag{4.10}$$

where,

$\Omega^t_{i,j}(l_{VC} \mid l_v)$	=	the neighbourhood index value of the vacant cell (l_{VC}) at location (i,j) at time t where the cell itself plus the surrounding land use (l_v) cells are measured, here referred to the development density within the m x m neighbourhood of 21 x 21 on the basis of Moore neighbourhood configuration;
$s^t_{q,r}$	=	the state of cell at time t at location (q,r) where (q,r) refers to the neighbouring cells of ($i \pm 10, j \pm 10$);
con()	=	a condition function which returns 1 (TRUE value) if the state of $S_{q,r}$ is one of targeted state (l_v) considered, otherwise 0 (FALSE value);
l_v	=	a particular land use type (l_v) where l_v can be residential (l_{RS}), commercial (l_{CM}), or industrial (l_{MA}).
m	=	the width and height of the window size (neighbourhood).

The following example in Figure 4.5 illustrates how the neighbourhood effect as illustrated in Equation 4.10 can be calculated. In the Figure, the 5 by 5 grid cell has the state of four land use types, regarding residential (RS), commercial (CM), industrial (MA), and vacant (VC).

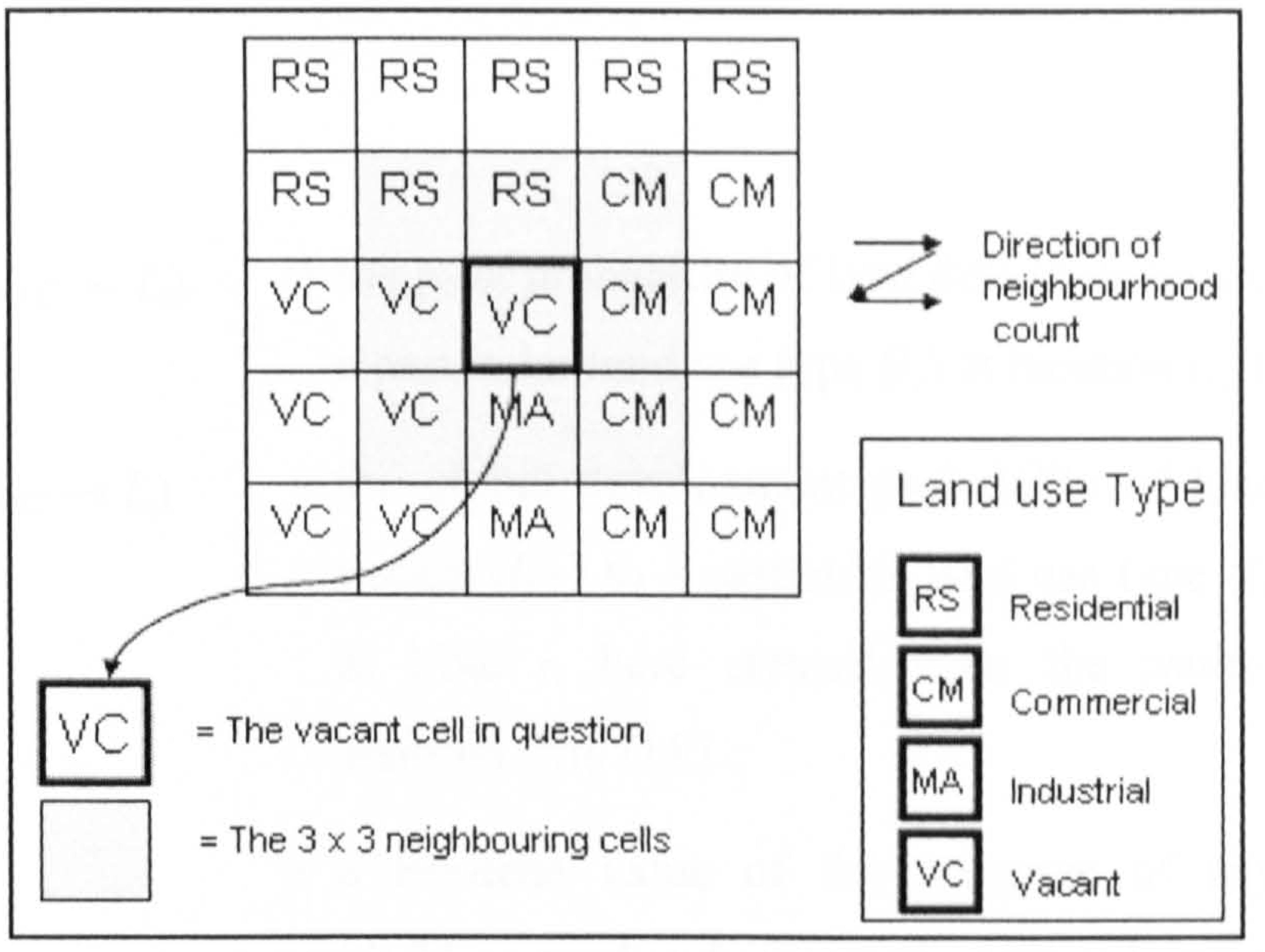


Figure 4.5: Example showing state of cell within 5 x 5 grid cell.

Based on the neighbourhood of 3 x 3 cells, the residential neighbourhood potential of the vacant cell in question $\Omega\,(l_{VC} \rightarrow l_{RS})$ can be computed as:

$$\Omega\,(l_{VC} \rightarrow l_{RS}) = (1 + 1 + 0 + 0 + 0 + 0 + 0 + 0) / (9 - 1) = 0.25$$

and the commercial neighbourhood potential of the vacant cell in question $\Omega\,(l_{VC} \rightarrow l_{CM})$ can be computed as:

$$\Omega\,(l_{VC} \rightarrow l_{CM}) = (0 + 0 + 1 + 0 + 1 + 0 + 0 + 1) / (9 - 1) = 0.38$$

and the industrial neighbourhood potential of the vacant cell in question $\Omega\,(l_{VC} \rightarrow l_{MA})$ can be computed as:

$$\Omega\,(l_{VC} \rightarrow l_{MA}) = (0 + 0 + 0 + 0 + 0 + 0 + 1 + 0) / (9 - 1) = 0.13$$

4.4.1.2 Development of Joint Probability Using the MCDA Method

Based on a rule-based structure of the transition functions as described in Equation 4.5, a joint probability $JP^{t}_{i,j}$ based on the MCDA method can be expressed as:

$$JP^{t}_{i,j}\,(\mathrm{l_{VC}} \rightarrow l_{v}) = P^{t}_{i,j}(l_{VC} \rightarrow l_{v}) * C^{t}_{i,j} * NR^{t}_{i,j}(l_{VC} \rightarrow l_{v}) \qquad (4.11)$$

where,

$JP^t_{i,j}(l_{VC} \rightarrow l_v)$ = the joint probability of land development from vacant (l_{VC}) to a particular land use type (l_v) at location (i,j) at time t;

$P^t_{i,j}(l_{VC} \rightarrow l_v)$ = the global development probability of land transition from vacant (l_{VC}) to a particular land use type (l_v) at location (i,j) at time t, here created from the multi-criteria decision analysis (MCDA) ;

$C^t_{i,j}$ = a Boolean value of the presence of any constraint that hinders new development to occur at location (i,j) at time t;

$NR^t_{i,j}(l_{VC} \rightarrow l_v)$ = a Boolean value of potential land transition from vacant (l_{VC}) to a particular land use type (l_v) at location (i,j) at time t, here derived from the thresholded neighbourhood index value being constrained by thresholding (1 implies potential for change; 0 implies no change);

l_v = A particular land use type (l_v) where l_v can be residential (l_{RS}), commercial (l_{CM}), or industrial (l_{MA}).

In Equation 4.11, the global development probability $P^t_{i,j}(l_{VC} \rightarrow l_v)$ is computed based on the formulation of the MCDA method described in Equation 4.4.

The Boolean value of $NR^t_{i,j}(l_{VC} \rightarrow l_v)$ in Equation 4.11 above is a product of the neighbourhood index value that is constrained by three sets of threshold values, regarding residential, industrial and commercial type. It is given, using IF, THEN and ELSE statements by:

FOR each land use transition type $(l_{VC} \rightarrow l_v)$ (4.12)

IF $\{\mathrm{MinNR\_RS}(l_{VC} \rightarrow l_v) \leq \Omega^t_{i,j}(l_{VC} \mid l_{RS}) \leq \mathrm{MaxNR\_RS}(l_{VC} \rightarrow l_v)$ OR

$\mathrm{MinNR\_CM}(l_{VC} \rightarrow l_v) \leq \Omega^t_{i,j}(l_{VC} \mid l_{CM}) \leq \mathrm{MaxNR\_CM}(l_{VC} \rightarrow l_v)$ OR

$\mathrm{MinNR\_MA}(l_{VC} \rightarrow l_v) \leq \Omega^t_{i,j}(l_{VC} \mid l_{MA}) \leq \mathrm{MaxNR\_MA}(l_{VC} \rightarrow l_v)\}$

THEN $NR^{t}_{i,j}(l_{VC} \rightarrow l_v) = 1$

ELSE $NR^{t}_{i,j}(l_{VC} \rightarrow l_v) = 0$

ENDFOR

where,

$NR^{t}_{i,j}(l_{VC} \rightarrow l_v)$	=	a Boolean value of potential land transition from vacant (l_{VC}) to a particular land use type (l_v) at location (i,j) at time t, here derived from the thresholded neighbourhood index value being constrained by thresholding (1 implies potential for change; 0 implies no change);
MinNR_RS($l_{VC} \rightarrow l_v$)	=	the minimum threshold value of residential neighbourhood index of land transition from vacant (l_{VC}) to a particular land use type (l_v);
MaxNR_RS($l_{VC} \rightarrow l_v$)	=	the maximum threshold value of residential neighbourhood index of land transition from vacant (l_{VC}) to a particular land use type (l_v);
MinNR_CM($l_{VC} \rightarrow l_v$)	=	the minimum threshold value of commercial neighbourhood index of land transition from vacant (l_{VC}) to a particular land use type (l_v);
MaxNR_CM($l_{VC} \rightarrow l_v$)	=	the maximum threshold value of commercial neighbourhood index of land transition from vacant (l_{VC}) to a particular land use type (l_v);
MinNR_MA($l_{VC} \rightarrow l_v$)	=	the minimum threshold value of industrial neighbourhood index of land transition from vacant (l_{VC}) to a particular land use type (l_v);
MaxNR_MA($l_{VC} \rightarrow l_v$)	=	the maximum threshold value of industrial neighbourhood index of land transition from vacant (l_{VC}) to a particular land use type (l_v);
$\Omega^{t}_{i,j}(l_{VC} \mid l_{RS})$	=	a residential neighbourhood index of the cell at location (i,j) at time t, created by Equation 4.11;

$\Omega'_{i,j}(l_{VC} \mid l_{CM})$ = a commercial neighbourhood index of the cell at location (i,j) at time t, created by Equation 4.11;

$\Omega'_{i,j}(l_{VC} \mid l_{MA})$ = an industrial neighbourhood index of the cell at location (i,j) at time t, created by Equation 4.11;

l_v = a particular land use type (l_v) where l_v can be residential (l_{RS}), commercial (l_{CM}), or industrial (l_{MA}).

In equation 4.12 above, each vacant cell has the residential, commercial and industrial neighbourhood index value, which is computed using Equation 4.11. For each land transition type, three sets of threshold values of residential, commercial, and industrial neighbourhood index are produced. The residential, industrial and commercial neighbourhood indexes of the vacant cell produced in Equation 4.11 will be compared to the corresponding sets of threshold values. If the neighbourhood index value of each vacant cell is within the range of the threshold setting, the Boolean value will be set to True (coded as 1), otherwise 0.

Table 4.2 gives an example of three sets of threshold value assigned for three land transition types – vacant to residential, vacant to commercial and vacant to industrial. Methods of calculating this value are shown in Section 4.5.3.

Based on Table 4.2, a set of threshold values applicable for vacant change to residential (as highlighted in the first column of the table) can be written in the IF, THEN, ELSE statement as:

IF $\{0.2 \le \Omega'_{i,j}(l_{VC} \mid l_{RS}) \le 0.8$ OR

$0.0 \le \Omega'_{i,j}(l_{VC} \mid l_{CM}) \le 0.3$ OR

$0.0 \le \Omega'_{i,j}(l_{VC} \mid l_{MA}) \le 0.0\}$

THEN $NR'_{i,j}(l_{VC} \rightarrow l_{RS}) = 1$

ELSE $NR'_{i,j}(l_{VC} \rightarrow l_{RS}) = 0$

Thus, if the vacant cell in question, for example, has the residential neighbourhood index value $\Omega'_{,i,j}(l_{VC} \mid l_{RS}) = 0.7$, commercial neighbourhood index value $\Omega'_{,i,j}(l_{VC} \mid l_{CM}) = 0.5$ and industrial neighbourhood index value $\Omega'_{,i,j}(l_{VC} \mid l_{MA}) = 0.1$, after substituting these values in the statement, the $NR'_{i,j}(l_{VC} \rightarrow l_{RS}) = 1$ (TRUE), which means this cell has the potential to change to residential.

Threshold description	**Threshold value for**		
	Vacant change to Residential ($l_{VC} \rightarrow l_{RS}$)	***Vacant change to Commercial*** ($l_{VC} \rightarrow l_{CM}$)	***Vacant change to Industrial*** ($l_{VC} \rightarrow l_{MA}$)
MinNR_RS	0.2	0.0	0.0
MaxNR_RS	0.8	0.0	0.0
MinNR_CM	0.0	0.0	0.0
MaxNR_CM	0.3	0.5	0.2
MinNR_MA	0.0	0.0	0.0
MaxNR_MA	0.0	0.2	0.7

Table 4.2: Example of threshold setting for three types of Neighbourhood index.

4.4.2 Handling of Conflicting Joint Probability Scores.

According to Equation 4.5 for both the MNL and MCDA method, three final joint probability maps of residential, commercial or industrial based on $JP'_{i,j}(l_{VC} \rightarrow l_v)$ are produced. However, there may be instances in which the three joint probability scores of a particular vacant cell location can possibly have an equal value (e.g. $l_{VC} \rightarrow l_{RS} = 0.80000$; $l_{VC} \rightarrow l_{CM} = 0.80000$; $l_{VC} \rightarrow l_{MA} = 0.80000$). The solution to handling this problem is that the vacant cell in question will change to the targeted land use type that has the highest neighbourhood index value.

4.4.3 State of Cell Transition and Iteration

After solving conflicting scores (if any), the final product of joint probability for three types of land use transition ($JP^{t}_{i,j}(l_{VC} \rightarrow l_v)$) derived from Equation 4.5 from both MNL and MCDA methods will contain suitability scores. By following the cell transition regime of the research work carried out by Barredo et al. (2004), Engelen et al. (1997a), and White et al. (2000), scores of those cells will be ranked by their highest potential. The rule for cell transitions is that cells with the highest scores of joint probability of each land use type will first be selected and changed to the state for which it has the highest potential, otherwise they remain vacant. Transition of cells can be given by

$$S^{t+1}_{i,j} = l_v \begin{cases} l_v, \text{if } \mathrm{JP}'_{i,j}(l_{VC} \rightarrow l_v) = \text{maximum}(\mathrm{JP}'_{i,j}(l_{VC} \rightarrow l_v)) \\ l_{VC}, \text{otherwise} \end{cases} \tag{4.13}$$

where $S^{t+1}_{i,j}$ is the state of cell transition (at time $t+1$) at location (i,j), $JP^{t}_{i,j}(l_{VC} \rightarrow l_v)$ is the final product of joint probability of land development at time t at location (i,j).

Transition starts with the highest ranked cell and continues downward until a specified number of cells based on the threshold set by land use demand for a particular land use type is reached. The CA model is designed to run in an iterative fashion to produce a new land use map at the end of each iteration (one year). In this study, the simulation is run from the initial year 1993 to 2001. Thus the *start_year* variable is set to 1993 while the *end_year* variable is set to 2001. In order to improve the computer efficiency the application is assigned to run based on a two year iteration, thus the *year_interval* is set to 2.

4.5 Derivation of Criterion Weights and Neighbourhood Thresholds

Multinomial logistic regression (MNL) and multi-criteria decision analysis (MCDA) methods are employed in a CA transition rule in order to compute development probability and suitability respectively. Both methods, however, are deterministic in the sense that they require a set of coefficients or criterion weights to drive the model. In this section, the procedure for derivation of criteria weights of both methods is presented.

4.5.1 Derivation of Criterion Weights Using the MNL Method

Multinomial logistic regression (MNL), an empirical method that helps observe and examine the relationship between the pattern of land use change and their physical and locational characteristics (Almeida et al., 2003; Mcmillen, 1989; Wu, 2002a), is used to derive coefficients. These coefficients are regarded as a set of criterion weights to be used for land transition types.

The MNL procedure for the derivation of criterion weights can be divided into two main steps; (i) statistical observation, and (ii) selection of development factors and MNL parameter estimation. They are both carried out using statistical software package SPSS 12.0.

Statistical observation is applied as a means to examine and identify a set of those development factors meaningful to explain the different types of land use changes considered. In this study, the characteristics (so–called development factors) of the 2001 data that experienced four types of land transition regarding vacant to residential, commercial, industrial land and no change, obtained from the overlaid land use layers of 1993 and 2001 using ArcGIS 9.1, are observed. These development factors (Section 3.4) are presumed to influence the land use change. They were measured by the means of GIS functions as presented in Section 3.4.1, and input into SPSS as an ASCII text file. The two statistical techniques employed, as a means for statistical observation, are boxplots and multicollinearity testing.

The analysis of boxplots is carried out to help observe whether the considered factors influence the land use transition or not. This is done by creation of a boxplot for each selected development factor versus the types of land use change. Observation is undertaken by comparing each box in the box plot of land transition types (e.g. 'vacant change to residential') with a box of 'no change' type. It is assumed that if the factor influences the land transition, the box of 'no change' type should be located in a different position from the box of the considered land transition type (e.g. 'vacant change to residential').

A boxplot diagram is used to graphically show the measure of dispersion for a given variable; the median, spread and inter-quartile range of scores (Field, 2000). Figure 4.6

shows an example of the boxplots of the transition from vacant to residential, commercial, industrial and vacant land (no change) versus main roads. The plotted data is observed based on the 2001 land use data that experienced change. Each box (e.g. box of land transition from vacant to residential) represents the inter-quartile range while the line across the box represents the median. Lines that extend from the box, so-called "whiskers", indicate the maximum and minimum values. The location of the box between the whiskers implies how the data are dispersed. If the box is in the middle of the whiskers, the data are probably rather equally distributed. If the box is located near to the lower whisker, the data tend to be skewed towards the lower end of the scale. On the contrary, if the box is located near to the upper whisker, the data are likely to be skewed towards the higher end of the scale. In the example (Figure 4.6), it seems that all land transition types tend to be located near to main roads as all boxes are close to the lower whisker. However, compared to the vacant area that did not experience change, 'main roads' factor tends to have most influence on 'vacant to industrial' transition and 'vacant to commercial' transition respectively, but has no influence on 'vacant to residential transition'.

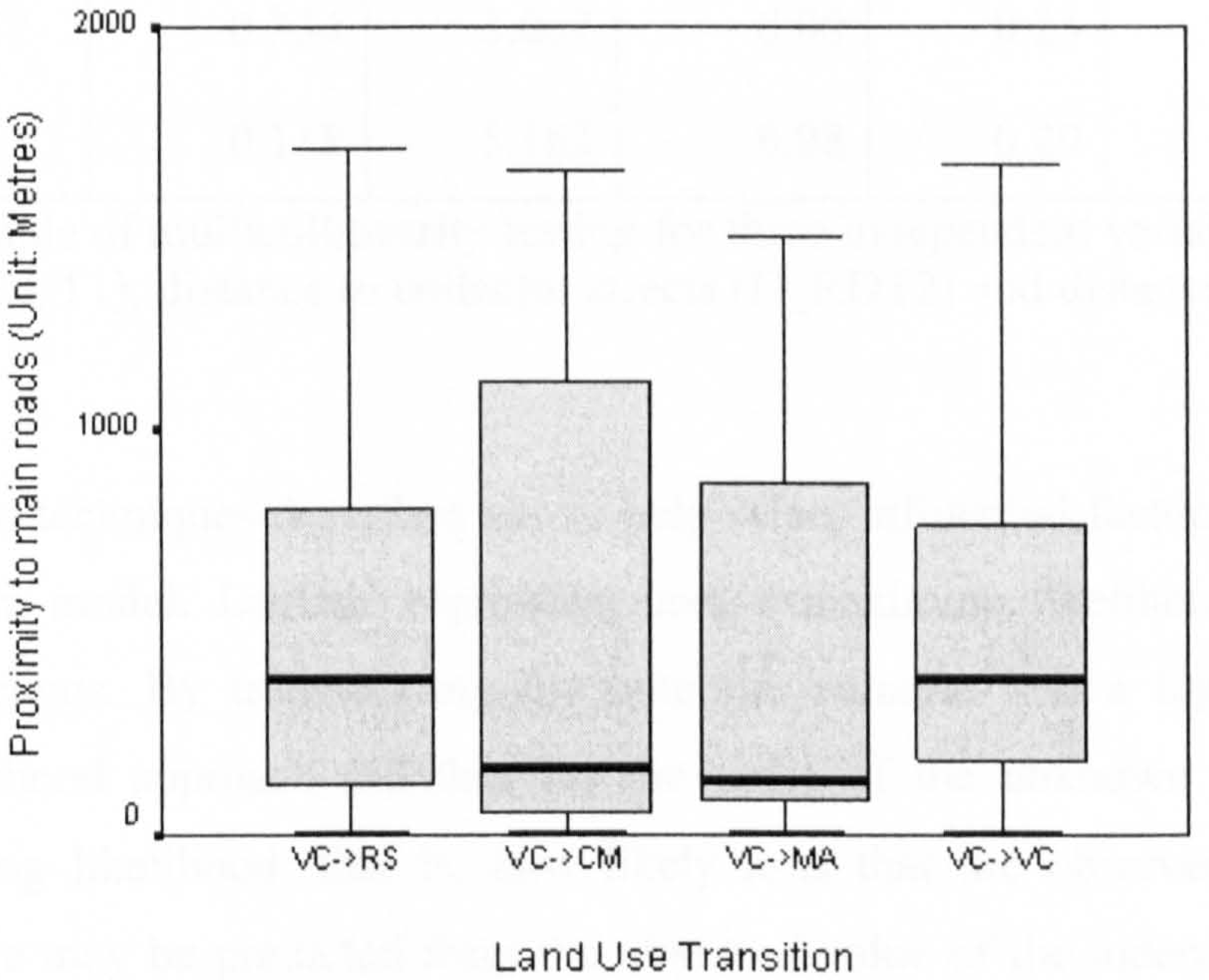


Figure 4.6: Boxplots of land transition from vacant to residential (VC->RS), commercial (VC->CM), industrial (VC->MA), and vacant land (no change) (VC->VC), versus proximity to main roads in 2001. Remark: the plotted data is based on the 2001 land use types that experienced change.

Another statistical technique employed is multicollinearity testing. This technique is carried out in order to help detect whether one variable has a perfect or near perfect linear relationship to one or more other variables. Such a situation can produce a biasing effect and result in unstable regression where a very small change to the dataset can result in considerably different coefficients (Field, 2000; SPSS Inc., 2001). The common indicator values used to measure muticollinearity are eigenvalue, condition index, and tolerance value. Table 4.3 shows an example of multicollinearity testing for three independent variables: distance to main roads, distance to collector streets and distance to local streets, based on the analysis of the SPSS software package. These results (Table 4.3) show that there was no multicollinearity among these three variables, as eigenvalues are not close to 0 and condition index values are less than 30 (SPSS Inc., 2001).

Collinearity Diagnostic (Dependent variable: Land Use Transition)

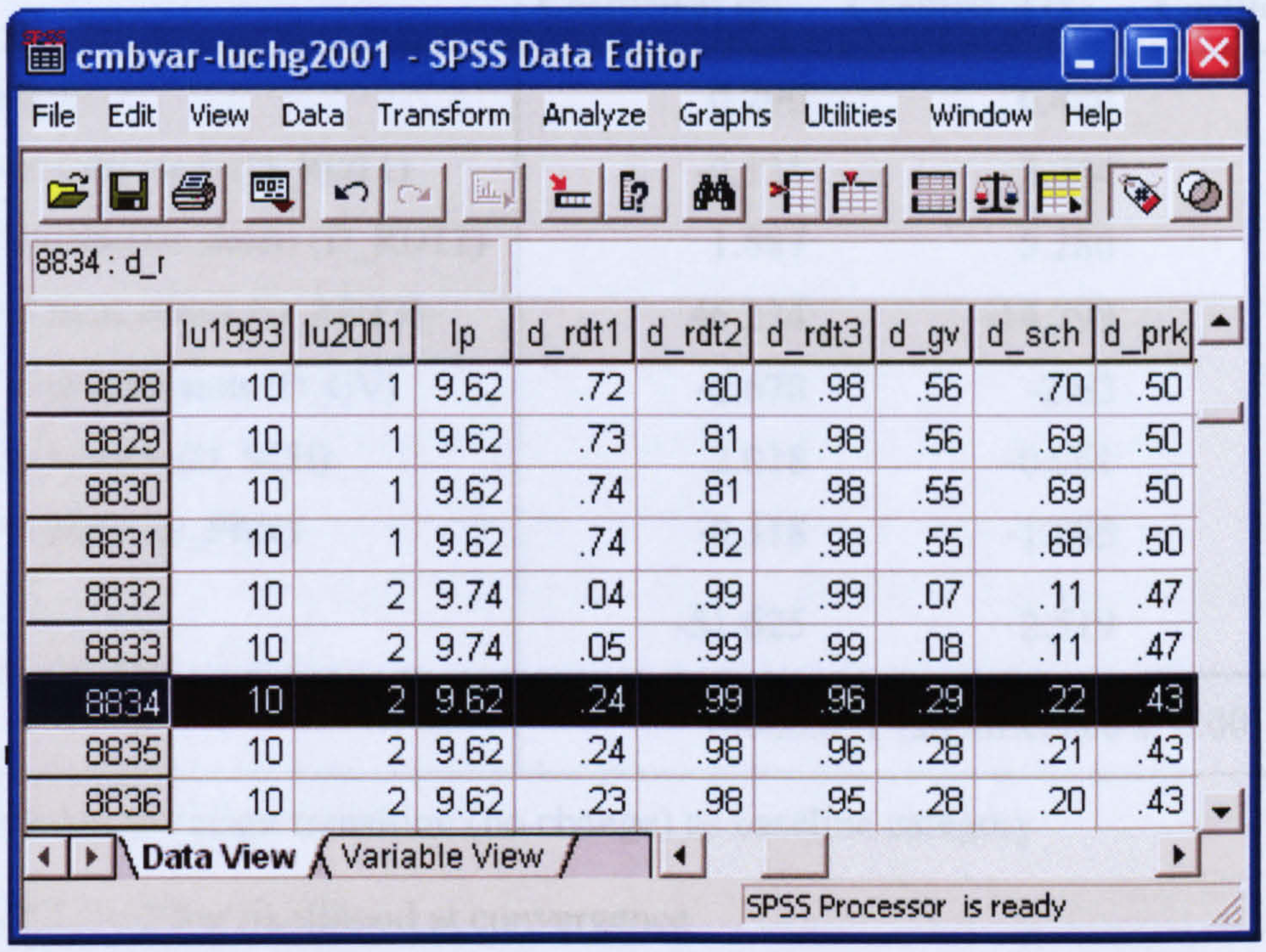
Model No.	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(constant)	D_RDT1	D_RDT2	D_RDT3
1	(constant)	3.140	1.000	0.02	0.03	0.03	0.03
	D_RDT1	0.408	2.773	0.00	0.43	0.00	0.55
	D_RDT2	0.334	3.067	0.00	0.25	0.57	0.22
	D_RDT3	0.118	5.163	0.98	0.29	0.41	0.20

Table 4.3: Example of multicollinearity testing for three independent variables; distance to main roads (D_RDT1), distance to collector streets (D_RDT2) and distance to local streets (D_RDT3).

Both observation techniques described above help select influential factors to be included in the regression model. Logistic regression uses a maximum likelihood estimation to compute coefficients. By transforming the outcome variable into a logit variable, the maximum likelihood approach searches for the value of the unknown parameters that maximize the log likelihood, that is, how likely it is that the observed values of the outcome variable may be predicted from the observed value of the independent variables (Myung, 2003).

Within the SPSS software package, the development factors in association with the considered land use types, regarding the residential, commercial, industrial and vacant use that is considered significant to the analysis, will be computed mutually using the MNL tool provided by the software to derive a set of coefficients. Figure 4.7 shows an example of the development factors transferred from GIS measurements as an ASCII text file to the

SPSS software package to be used for MNL parameter estimation. In the figure, each record contains the land use type in association with the considered development factors. For example, the highlighted row contains the 1993 land use data coded as 10 (vacant) changing to the commercial type coded as 2 in the 2001 land use in association with the development factors (e.g. proximity to main roads (valued as 0.24)). With this method, the relative importance or weights of the independent variables (e.g. land price) can be compared both within and between land transition types (Garson, 2005).



	lu1993	lu2001	lp	d_rdt1	d_rdt2	d_rdt3	d_gv	d_sch	d_prk
8828	10	1	9.62	.72	.80	.98	.56	.70	.50
8829	10	1	9.62	.73	.81	.98	.56	.69	.50
8830	10	1	9.62	.74	.81	.98	.55	.69	.50
8831	10	1	9.62	.74	.82	.98	.55	.68	.50
8832	10	2	9.74	.04	.99	.99	.07	.11	.47
8833	10	2	9.74	.05	.99	.99	.08	.11	.47
8834	10	2	9.62	.24	.99	.96	.29	.22	.43
8835	10	2	9.62	.24	.98	.96	.28	.21	.43
8836	10	2	9.62	.23	.98	.95	.28	.20	.43

Figure 4.7: An example of measured development factors used in the SPSS software package. Remark: in the table, column 'lu1993' refers to the 1993 land use type, 'lu2001' refers to the 2001 land use type, d_rdt1 refers to proximity to main roads, d_rdt2 refers to proximity to collector streets, v7 refers to proximity to local streets, v8 refers to proximity to residential and v9 refers to proximity to commercial.

Although MNL has no specific rule about the data unit and the range of value for each of the development factors to be prepared before the calculation of MNL parameter estimation, the study tested development factor with the normalized values, which gave significant and realistic estimation results compared to those with the raw data. The exception was with land price data. The study indicated that land price factor gave a better result when taking the natural logarithm. Thus, the land price value in Figure 4.7 was produced as a result of ln(land price).

Table 4.4 shows an example of a set of coefficients derived by means of MNL using the SPSS software package. In the example, land price, proximity to three road types, proximity to government, schools and parks are used to compute mutually four transition types: vacant to residential, vacant to commercial, vacant to industrial and vacant to vacant (no change).

Factors (NOTATION)	**Vacant to Residential Coefficient (β)**	**Vacant to Commercial Coefficient (β)**	**Vacant to Industrial Coefficient (β)**
Land price (LP)	0.290	0.424	0.637
Proximity to main roads (D_RDT1)	-0.131	2.564	0.933
Proximity to collector streets (D_RDT2)	1.887	5.286	4.292
Proximity to local streets (D_RDT3)	46.234	-14.799	7.403
Proximity to governments (D_GV)	-0.678	-2.62	0.313
Proximity to schools (D_SCH)	2.018	-0.061	-0.857
Proximity to parks (D_PRK)	-0.318	-1.885	-1.617
Intercept	-51.625	2.519	-21.301
-2LL	18005.031 (significance at 0.00)		

Note: 1. Vacant toVacant transition (no change) as baseline category

2. -2LL = -2 log likelihood at convergence

Table 4.4: Estimated coefficients calibrated from multinomial logistic regression.

The vacant to vacant transition in Table 4.4 is used as a baseline, which means that all coefficients for this are equal to 0. The β coefficients are referred to as logit or log (odd ratios). The odds of the event are defined as the probability of an event occurring divided by the probability of that event not occurring. If probability of an event occurring equals 0.5 and the probability of that event not occurring equals 0.5, thus the odds of the event is equal to 1 (0.5 /0.5 = 1) and the log (odd ratios), or β, is equal to 0. Overall, if β is positive, this implies that the odds of the event occurring increase. Inversely, if β is negative, this implies that the odds of the event occurring decrease. If β is zero, this means that the odds of the event are unchanged. In Table 4.4, for example, the β of *proximity to main roads* with respect to vacant to residential transition is -0.13 when compared to that of vacant to vacant transition type (β = 0). Similarly, a β of land price of 2.56 with respect to vacant to commercial transition and β of 0.93 with respect to vacant to industrial transition are derived when compared to that of vacant to vacant transition type ($\beta = 0$). This means that when including the *proximity to main roads* variable in the model, the event of 'vacant change to residential' will decrease. However, the events of 'vacant change to commercial' and 'vacant change to industrial' will increase.

More meaningful to the interpretation of logistic regression is the odds ratio, the value of exp(β), which is an indicator of the change in odds resulting from a unit change in the outcome (Field, 2000). In Table 4.4 where the β of *proximity to main roads* with respect to vacant to commercial transition is 2.56, compared to that of vacant to vacant transition type (remaining unchanged), the estimate of odds ratio computed thus equals to 12.94. This means that, when including land price in the model, land use change is about 12.94 times as likely to become commercial than remain unchanged. Overall, the larger odds ratios within a logit indicate which variables have the most effect for that logit's category of the outcome variable. An odds ratio above 1.0 means that the odds of a dependent variable (e.g. residential) is greater than for the baseline category (e.g. remaining unchanged). Conversely, an odds ratio below 1.0 means that the odds of a dependent variable is less than for the baseline category.

Two tests are usually performed for examining the significance of the coefficients produced by a logistic model (Field, 2000; Garson, 2005). The first test is used to examine the value of overall model fit by comparing the model before and after including variables. The second test is used for examining whether, or not, a variable should be included in the model. Both techniques use the likelihood-ratio statistic, -2 log-likelihood (-2LL), to indicate how much unexplained information remains after the model has been fitted (Field, 2000). If a logistic model fits completely, the -2LL value will equal 0. Because the -2LL has an approximately chi-square distribution, chi-square approach can be used to compute -2LL for the initial model (e.g. model without variables) minus -2LL for the final model (e.g. model with inclusive variables). If a chi-square value is less than the significance level (0.05), then the model will be accepted (Field, 2000).

In an example in Table 4.4, the significance value (0.00) of the -2LL statistic show the overall model fit. Since the significance value is less than 0.05, this suggests that overall the model with all the inclusive independent variables (e.g. land price (LP), proximity to schools (D_SCH)) is better to predict the outcome variable, rather than that of the model with intercept only (the model without variables). Table 4.5 shows the likelihood ratio test calculated based on the coefficients derived in Table 4.4. Similarly, the significance value of each variable in Table 4.5 is less than 0.05, it thus can be concluded that each variable gives an effect to the model and should be included along with the intercept in the model. In this table, although all variables are significant, they influence the model in different degrees. For example, proximity to local streets (D_RDT3) contributes more effect to the model than proximity to government (D_GV) since the chi-square value of D_RDT3 is higher than that of D_GV.

Effect of Variable	-2LL	Chi-Square	Degree of freedom (df)	Significance (Sig.)
Intercept	18730.358	725.327	3	0.000
LP	18155.742	150.711	3	0.000
D_RDT1	18108.739	103.707	3	0.000
D_RDT2	18216.771	211.739	3	0.000
D_RDT3	18508.871	503.839	3	0.000
D_GV	18030.189	25.158	3	0.000
D_SCH	18227.709	222.677	3	0.000
D_PRK	18098.285	93.254	3	0.000

Table 4.5: Likelihood-ratio statistic and chi-square value used for testing the significance of MNL parameter estimation, created by SPSS software package.

It should be noted that the coefficients calibrated and derived by this method have one limitation. They are limited in terms of the transferability to other locations or to other time periods. This is due to the fact that data are measured at one single point of time (Hosmer and Lemeshow, 2000) in a specific location. Different periods of time and/or different study areas can create different coefficients. In other words, a set of coefficients derived from a specific period in time and in a specific area can produce an unrealistic result if they are used for simulations of future growth that have different environments (e.g. different sets of development factors and weights). Despite its limitations, however, the main advantage of this method is that the coefficients or parameterized values derived are, to some extent, a reflection of the factual characteristics of the land use pattern obtained from the study site (Hu and Lo, 2007; Wu, 2002a). By using significant tests (e.g. the likelihood ratio test), it allows us to be confident in using these calibration values for model simulation.

4.5.2 Derivation of Criterion Weights Using the MCDA Method

Derivation of criterion weights based on the MCDA method relies considerably on decision makers' opinions, knowledge or experience (Lee, 1973; Malczewski, 1999a). As mentioned earlier (Section 4.3), MCDA can be divided into two categories; MADA (multi-attribute decision analysis) and MODA (multi-objective decision analysis). In this study, the weights sets from both MADA and MODA are used for the development suitability score. The criterion weights based on the MADA method are used to create the land suitability scores based on the multiple attributes (in this study, they refer to development factors) while criterion weights based on the MODA method are used to create the suitability scores based on the competing land use objectives. In this section, derivation of criterion weights using these two techniques is presented.

Based on both MADA and MODA, a normalized score for each criterion is always required in order to standardize the overall measurement. After normalization, the preferences of the decision makers for various degrees of the criteria can be assessed and transformed to comparable units (Malczewski, 1999a; Voogd, 1983). Usually, values of the normalized score are set ranging from 0.0 to 1.0. A high score value indicates a more favourable weight in the criterion.

One of the most often used procedures to standardize scores is given by Voogd (1983):

$$x'_{(i,j),g} = \frac{x_{(i,j)} - \min(x_g)}{\max(x_g) - \min(x_g)} \qquad (4.14)$$

where $x'_{(i,j)}$ is the standardized score of a criterion g for the cell at location (i,j), $x_{(i,j)}$ is a raw score, $\min(x_g)$ is the minimum score for the criterion g, $\max(x_g) - \min(x_g)$ is the range of a given criterion.

However, there are cases (e.g. proximity to roads) where the raw score needs to be inverted in order that the low raw value gains a high standardized score. Site selection for residential use, for example, prefers to locate close to roads. Without inversion, areas close to roads will get a low score. The reciprocal of a standardized score is given by Voogd (1983):

$$x'_{(i,j)g} = 1 - \frac{x_{(i,j)} - \min(x_g)}{\max(x_g) - \min(x_g)} \qquad (4.15)$$

where $x'_{(i,j)}$ is the standardized score of a criterion g for the cell at location (i,j), $x_{(i,j)}$ is a raw score, $\min(x_g)$ is the minimum score for the criterion g, $\max(x_g) - \min(x_g)$ is the range of a given criterion.

Amongst several techniques used for the derivation of criteria weights, the pairwise comparison matrix, developed by Saaty (Saaty and Alexander, 1981), is one of the most promising (Malczewski, 1999b; Marinoni, 2004) as it allows a comparison between every possible pair of criteria to be investigated and determined. Setting the scale of preference is often performed using a nine-point basis, ranging from 1 to 9 as shown in Table 4.6.

Intensity of importance	Definition
1	Equal Importance
3	Moderate importance
5	Strong importance
7	Very strong importance
9	Extreme importance
2,4,6,8	Intermediate values between the two adjacent judgements

Table 4.6: Scale of relative importance by means of pairwise comparison method (adapted from Saaty and Alexander (1981), p. 149).

In this study, the MADA technique was first applied to derive the weights set used for computation of land suitability. Each development factor (e.g. proximity to government areas) was first normalized using either Equation 4.14 or Equation 4.15. Through the pairwise comparison matrix (see an example in Table 4.8), assessment of the land suitability was conducted for a particular land use type by determining the level of importance of each criterion for each cell relative to the other criteria under consideration. Assigning a set of weights was undertaken mainly based on information gathered from informal interviews with planners from the Department of City Planning (DCP), BMA and the Land Use Compatibility Matrix set by the Department of Town and Country Planning (DTCP), Thailand as shown in Table 4.7.

	RS	CM	MA	GV	PRK	SCH	AGR
RS	*H*						
CM	*M*	*H*					
MA	*M*	*M*	*H*				
GV	*M*	*H*	*M*	*H*			
PRK	*H*	*H*	*M*	*H*	*H*		
SCH	*H*	*M*	*I*	*M*	*H*	*H*	
AGR	*M*	*M*	*H/M*	*M*	*H*	*M*	*H*

Table 4.7: Compatibility Matrix Standard (adapted from Department of Town and Country Planning, Thailand (source: ESRI (Thailand) Co. Ltd. (1997), p.3-15)). Remark: RS denotes residential, CM for commercial, MA for industrial, GV for government office, PRK for park and conservation area and SCH for schools. According to compatibility index; *H* = high compatibility, *H/M* = slightly high compatibility, *M* = Moderate Compatibility and *I* = Incompatibility.

Table 4.8 shows an example of the pairwise comparison matrix used for setting criteria weights for commercial suitability. Seven variables are used: close to residential, close to commercial, far from industrial, close to government offices, land price, close to main roads, and close to collector streets. The relationship between these variables can be interpreted, for example close to collector streets (C_RDT2) is considered to be 3 times more important than close to government offices (C_GV) and 5 times more important than land price (LP).

The computation of these criterion weights consists of three operations (Malczewski, 1999a). The first step is to sum up the value in this column of the pairwise comparison matrix. The second step is to divide each element in the matrix by its column total (so-called the normalized comparison matrix value). The third step is to compute the average of these elements in each row of the normalized value derived from the second step, that is, divide the sum of normalized scores for each row by the number of criteria. Table 4.9 –

Table 4.11 show the computation of criterion weights based on these three steps respectively.

Criteria	C_RS	C_CM	F_MA	C_GV	LP	C_RDT1	C_RDT2
C_RS	1						
C_CM	1	1					
F_MA	1/5	1/6	1				
C_GV	1/2	1/3	3	1			
LP	1/3	1/4	2	1/5	1		
C_RDT1	2	1	5	2	4	1	
C_RDT2	2	1	6	3	5	1	1

Table 4.8: Pairwise comparison assigned for each criterion used for commercial development. These variables refer to Close to residential (C_RS), close to commercial (C_CM), far from industrial (F_CM), close to government office (C_GV), land price (LP), close to main roads (C_RDT1), close to collector streets (C_RDT2).

	D_RS	D_CM	D_MA	D_GV	LP	D_RDT1	D_RDT2
D_RS	1.00	1.00	5.00	2.00	3.00	0.50	0.50
D_CM	1.00	1.00	6.00	3.00	4.00	1.00	1.00
D_MA	0.20	0.17	1.00	0.33	0.50	0.20	0.17
D_GV	0.50	0.33	3.00	1.00	2.00	0.50	0.33
LP	0.33	[illegible]	[illegible]	[illegible]	.00	0.25	0.20
D_RDT1	2.00	[illegible]	[illegible]	[illegible]	.00	1.00	1.00
D_RDT2	2.00	[illegible]	6.00	3.00	5.00	1.00	1.00
SUM	7.03	4.75	28.00	11.83	19.50	4.45	4.20

A sum value of D_RS factor

Table 4.9: Pairwise comparison matrix translated from Table 4.8. Here after applying the first step, that is, summing up the value (highlighted in grey).

After computation of criterion weights, these weight values need to measure the degree of consistency in order to examine the consistency of scores set in the pairwise comparison matrix. This was conducted by calculating the consistency ratio (CR). To compute the consistency ratio (CR), four steps are required (Malczewski, 1999a). To illustrate the computation, an example is given step by step. It should be noted that the example here is taken from real world data, calculated from Microsoft excel. There will be a shift in the presented value here (two digits) since real values used for computation is with more precision.

	D_RS	D_CM	D_MA	D_GV	LP	D_RDT1	D_RDT2
D_RS	0.14	0.21	0.18	0.17	0.15	0.11	0.12
D_CM	0.14	0.21	0.21	0.25	0.21	0.22	0.24
D_MA	0.03	0.04	0.04	0.03	0.03	0.04	0.04
D_GV	0.07	0.07	0.11	0.08	0.10	0.11	0.08
LP	0.05	0.05	0.07	0.04	0.05	0.06	0.05
D_RDT1	0.28	0.21	0.18	0.17	0.21	0.22	0.24
D_RDT2	***0.28***	***0.21***	***0.21***	***0.25***	***0.26***	***0.22***	***0.24***
TOTAL	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 4.10: Pairwise comparison matrix after applying step two, here the normalized comparison matrix values were generated. The row with bold italic letters refers to the normalized comparison matrix values of the total D_RDT2 factor

Factors	WEIGHT
D_RS	**0.16**
D_CM	**0.21**
D_MA	**0.03**
D_GV	**0.09**
LP	**0.05**
D_RDT1	**0.22**
D_RDT2	**0.24**
Total	**1.0**

Weight of D_RDT2 factor

= (0.28 + 0.21 + 0.21 + 0.25 + 0.26 + 0.22 +0.24) / 7

= 0.24

Table 4.11: Derivation of criterion weights based on the computation of values derived from the second step (Table 4.10) divided by the number of criteria (here, there are 7 criteria).

The first step is to determine the weighted sum vector by multiplying the weight for the first criterion (e.g. weight of D_RS in Table 4.11) by the first column of the original pairwise comparison matrix (e.g. an element D_RS in the fist column in Table 4.9) , then multiply the second weights by the second column, and so on. Then, sum these values over the row.

Step 1: computing the weighted sum vector (e.g. the weighted sum vector of D_RS)

= (**0.16** * **1.00**) + (0.21 * 1.00) + (0.03 * 5.00) + (0.09 * 2.00) + (0.05 * 3.00) + (0.22 * 0.50) + (0.24 * 0.50)

= 1.10

Remark: the first figure of each term (e.g. **0.16**) refers to the weight derived from Table 4.11 and the second number of each term (e.g. **1.00**) refers to the original value derived from the comparison matrix (Table 4.9).

The second step is to compute the consistency vector by dividing the weighted sum vector by the criterion weights previously derived in Table 4.11.

Step 2: computing the consistency vector of D_RS

= 1.10/ 0.16

= 7.11

Table 4.12 shows the consistency vector of all factors, which are calculated using a similar method. Once, the consistency vector is derived, the third step is to compute the lambda (λ) by dividing the sum of consistency vector by the number of criterion factors under consideration.

Step 3: computing the lambda (λ)

λ = (7.11 + 7.09 + 7.06 + 7.07 + 7.07 +7.13 + 10.16)/ 7

= 7.53

	consistency vector
D_RS	7.11
D_CM	7.09
D_MA	7.06
D_GV	7.07
LP	7.07
D_RDT1	7.13
D_RDT2	10.16

Table 4.12: Derivation of the consistency vector.

The fourth step is to calculate the consistency index (CI). Calculation of CI is expressed as:

$$\mathrm{CI} = \frac{\lambda - n}{n-1} \tag{4.16}$$

where n = the number of criterion weights, λ = the lambda derived from Step 3.

Step 4: computing the consistency index (CI)

$$\mathrm{CI} = \frac{7.53-7}{7-1} = 0.09$$

The final step is to compute the consistency ratio (CR). Calculation of CR is expressed as:

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{4.17}$$

where CI = the consistency index, RI= the random index value (derived from the random inconsistency indices table (Saaty (1980) cited in Malczewski (1999a)). Note that RI value depends on the number of factors being compared.

Step 5: computing CR value.

$$\mathrm{CR} = \frac{0.09}{1.32} = 0.07$$

In this example, the RI value derived from the random inconsistency indices table is equal to 1.32, based on 7 elements (n).

In general, a CR value less than 0.1 indicates a sufficient level of consistency between variables (Malczewski, 1999a; 1999b). However, if a CR value is equal to or higher than 0.1, this shows that the criterion weights set are inconsistent judgements. Thus, the rating of criterion weights need to be adjusted and re-evaluated until the CR value is less than 0.1 (Malczewski, 1999a).

Based on the same method described above, derivation of criterion weights used for residential and industrial suitability can be carried out. However, as mentioned in Section 4.3, each land suitability can have different sets of criteria.

In addition to the criterion weights for the creation of a particular suitability map, in this study another set of criterion weights based on the MODA method are derived, in order to provide competing land use objectives. Derivation of criterion weights for competing land uses can be carried out by first normalizing each criterion and then performing a pairwise comparison matrix, similar to that of the MADA method. Assessment of the goal (or objective) for competing land uses was conducted by determining the level of importance of each criterion for each cell relative to the other criteria under consideration. Assigning a set of weights was undertaken mainly based on information gathered from informal interviews with planners from the Department of City Planning (DCP), BMA. Based on the interview, however, preference scores for each land use type are determined to be set equally (see Table 4.13).

Criteria	Change from Vacant to Residential	Change from Vacant to Commercial	Change from Vacant to Industrial	**Weight**
Change from Vacant to Residential	1			**= 0.33**
Change from Vacant to Commercial	1	1		**= 0.33**
Change from Vacant to Industrial	1	1	1	**= 0.33**

Table 4.13: Pairwise comparison matrix and objective weights derived for competing land use in this study.

The criterion weights derived from both MADA and MODA methods, thus, are incorporated to use in Equation 4.4. As a result, the final transition score of land transition from vacant to residential, commercial and industrial can be carried out. Note that since the criterion weights produced from the MODA method are equal, this means that the computation of development probability of land transitions is based solely on the criterion weights produced from the MADA method.

4.5.3 Derivation of Neighbourhood Thresholds for the GIS-based CA/MCDA Model

Using the GIS-based CA/MCDA model proposed in equation 4.11, a set of neighbourhood thresholds is needed. In this section, the rationale behind this setting, derivation of neighbourhood threshold, and interpretation are given.

Table 4.14 exemplifies possible thresholded neighbourhood index values. The rationale behind this concept comes from two assumptions. Firstly, most research, which applied to the urban CA for land development, was conducted at the level of city and region (e.g. Wu and Webster (1998)), the new emergence of a particular land use type was found to be compact, mainly spreading along the existing development. Mostly, they were conducted for two types of land conversion (e.g. urban and non-urban). The emergence of one land use type was clearly found adjacent to the same existing land use type. In the study area where the detail of district or neighbourhood level was applied, it would be a challenging task to allow the influence of other land use types (e.g. residential neighbourhood) on the emergence of a particular land use type considered (e.g. commercial) to be investigated. Secondly, a mix of land use patterns mostly found in the study area and most parts of the Bangkok area as described in Section 3.1.1.3 (e.g. the mix of residential houses and commercial buildings), as well as the observation of actual land use pattern in the study itself, suggest that the emergence of one land use type may possibly be influenced by other types of land usage. Although the complete *sequence* of land development could not be investigated for all emergent locations throughout the study period, the dependent neighbourhood effect of land use types on a particular land use type should be considered.

Variables	Threshold set for land use transition		
	Vacant change to residential	**Vacant change to commercial**	**Vacant change to industrial**
Residential neighbourhood index (NI_RS)	0.20-0.75	N/A	N/A
Commercial neighbourhood index (NI_CM)	0.00-0.30	0.00-0.50	0.00-0.20
industrial neighbourhood index (NI_MA)	N/A	0.00-0.20	0.0-0.70

Table 4.14: A list of neighbourhood threshold (or range of neighbourhood index value that is used to constrain development) for the targeted land use types, regarding residential, commercial and industrial.

The range of neighbourhood index values for each land transition type (e.g. vacant change to residential) will be used for constraining each land use development. The range of neighbourhood threshold of the same type is from 0.0 to 1.0, which refers to the proportion of cells in the neighbourhood. The value of 0.0 refers to the minimum index value of surrounding cells in the neighbourhood (e.g. 0 out of 8 surrounding cells) while 1.0 refers to the maximum index value of surrounding cells in the neighbourhood (e.g. 8 out of 8 surrounding cells).

Thus, in this example, for vacant change to residential, the residential suitability cells (derived from Equation 4.11) will be checked against the neighbourhood threshold whether they have potential for development or not. If the residential neighbourhood factors of these cells are in the range of 0.20 – 0.75 or the commercial neighbourhood factors of these cells are in the range of 0.00 – 0.30, these cells will have potential for residential development. The value 'N/A' means the threshold of industrial neighbourhood index value is not taken into account. In other words, the industrial neighbourhood is ignored, as it is considered to have no effect on vacant transition to residential. For vacant change to commercial, instead, only the commercial and industrial neighbourhoods are considered. If the commercial neighbourhood factors of these cells are in the range of 0.00 – 0.50 or the industrial neighbourhood factors of these cells are in the range of 0.00 – 0.20, these cells will have potential for commercial development. Interpreting the neighbourhood for industrial transition can be carried out in a similar way.

It should be noted that setting the range of neighbourhood thresholds of a land transition type (Table 4.14) can be independent of other land use types, and make them overlap when considering other transitions. This allows neighbourhood threshold settings to be more flexible. However, they should be set based on sensible reasons and with logical consideration. For example, for 'vacant change to residential', it is possible that the cells having commercial neighbourhood index values greater than 0.3 will be ignored, if they are not in the range of residential neighbourhood threshold. For 'vacant change to commercial', it is possible that the cells having commercial neighbourhood index values greater than 0.5 will be ignored, if they are not in the range of industrial neighbourhood threshold. And for 'vacant change to industrial', it is possible that the cells having commercial neighbourhood index values greater than 0.2 will be ignored, if they are not in the range of industrial neighbourhood threshold. Setting the overlap between these threshold values or setting the neighbourhood thresholds with a very narrow range can cause competition between land use transitions as many cells are eliminated from development. Possibly, setting a wider range of neighbourhood thresholds can cause overcrowding of some land use types in the area and exaggerate the neighbourhood effect.

In this study, the thresholded neighbourhood index values were derived from the statistical observation in the study area, the compatibility matrix of land use activities (ESRI (Thailand) Co. Ltd., 1997) set by the Department of Town and Country Planning (DTCP), Thailand (Table 4.7), and an initial interview with staffs in the Department of City Planning (DCP), Bangkok. Table 4.15 shows the neighbourhood index value calculated from observation of the study area in 2001 using Equation 4.10. The thresholds produced were analyzed from the basis of a 10m grid cell resolution within the neighbourhood (window) size of 21 x 21.

The actual Data calculated from the observation of the Study Site	Land use type		
	Residential	Commercial	industrial
Residential neighbourhood index (NI_RS)	0.00*-0.88**	0.00-0.76	0.00-0.63
Commercial neighbourhood index (NI_CM)	0.00-0.54	0.00-0.96	0.00-0.45
industrial neighbourhood index (NI_MA)	0.00-0.18	0.00-0.19	0.00-.017

Table 4.15: The range of neighbourhood index values observed from the actual land use map in 2001. Remark: * refers to 2 out of 12,842 cells that have value 0.0; ** refers to 4 out of 12,842 cells that have value 0.88.

According to Table 4.15, if these thresholds were used, **all** cells would have the potential to change because threshold would cover the whole range of the neighbourhood during simulation. As a result, the model may produce a simulated result which ignores the effect of neighbourhood. Some modification was attempted in order to adjust the range of thresholds to be more realistic (e.g. narrowing or extending the threshold range, overlapping the range, ignoring the effect of neighbourhood for some land use transitions). This was carried out by investigating the degree of relationship between each type of land use transition and the range of neighbourhood, using the statistical observation (e.g. box plot, frequency count, mean) plus suggestions from the staff of the Department of City Planning (DCP), Bangkok. The result of the modifications was that it was possible to set differing thresholded neighbourhood index values for the research study. Consequently, revised neighbourhood threshold tables (e.g. Table 6.20) were derived and used for the implementation carried out in section 6.3.

4.6 Conclusion

In this chapter, the CA approach for urban growth development developed for the research study is described. The developed model focuses on the change of land from vacant to residential, commercial, and industrial. The CA-based model developed includes the integration of the statistical multinomial logistic regression model (MNL) and multi-criteria decision analysis (MCDA), applied as CA transition rules. Derivation of criterion weights based on MNL and MCDA methods as well as derivation of neighbourhood thresholds are also described. In the following chapter, the model formulated here is embedded in a GIS environment and will be used for the implementation in Chapter 6.

CHAPTER 5

Development of Urban Simulation Model under a GIS Platform

5.1 Introduction

In this chapter, the developed urban CA model and derivation of criterion weights from both MNL and MCDA methods described in the previous chapter (Chapter 4) are incorporated for the creation of an urban land use change simulation model under a GIS platform (see Figure 4.1). Figure 5.1 shows the integration strategy employed. Two types of integration strategy are used. Integration of MNL and MCDA, as a means to weights derivation, is considered a loose-coupling strategy to GIS as both methods are undertaken outside the GIS environment. Their linkage to GIS is by its transfer to an ASCII text file. Integration between CA and GIS, on the contrary, is developed under the tight-coupling strategy in that the developed urban CA model is fully embedded in a GIS environment. The role of GIS functions for model implementation in this study can be broken down into five tasks; (i) creation of cell state maps and constraint map, (ii) elimination of constraint area, (iii) measurement of development factors, (iv) calculation of transition rules using for suitability / probability map creation, (v) land use map update (updating state of transition cells) at the end of iteration.

Despite the fact that most of the GIS functions provided by existing GIS software can facilitate the implementation of the proposed modelling framework presented in Chapter 4, it involves the application of many operations in a repetitive manner. When there are many steps involved it can be hard to keep track of the operations, datasets, and other parameter values being used. Consequently, such difficulty can easily produce unexpected and incorrect output as the result of error input through the manual handling process. In addition, since the model proposed requires iteration as part of the dynamic simulation, this can be difficult to handle without dedicated programming. One of the easiest ways to automate the work flow of the model presented in Chapter 4 and keep track of a set of GIS functions implemented is to create a model.

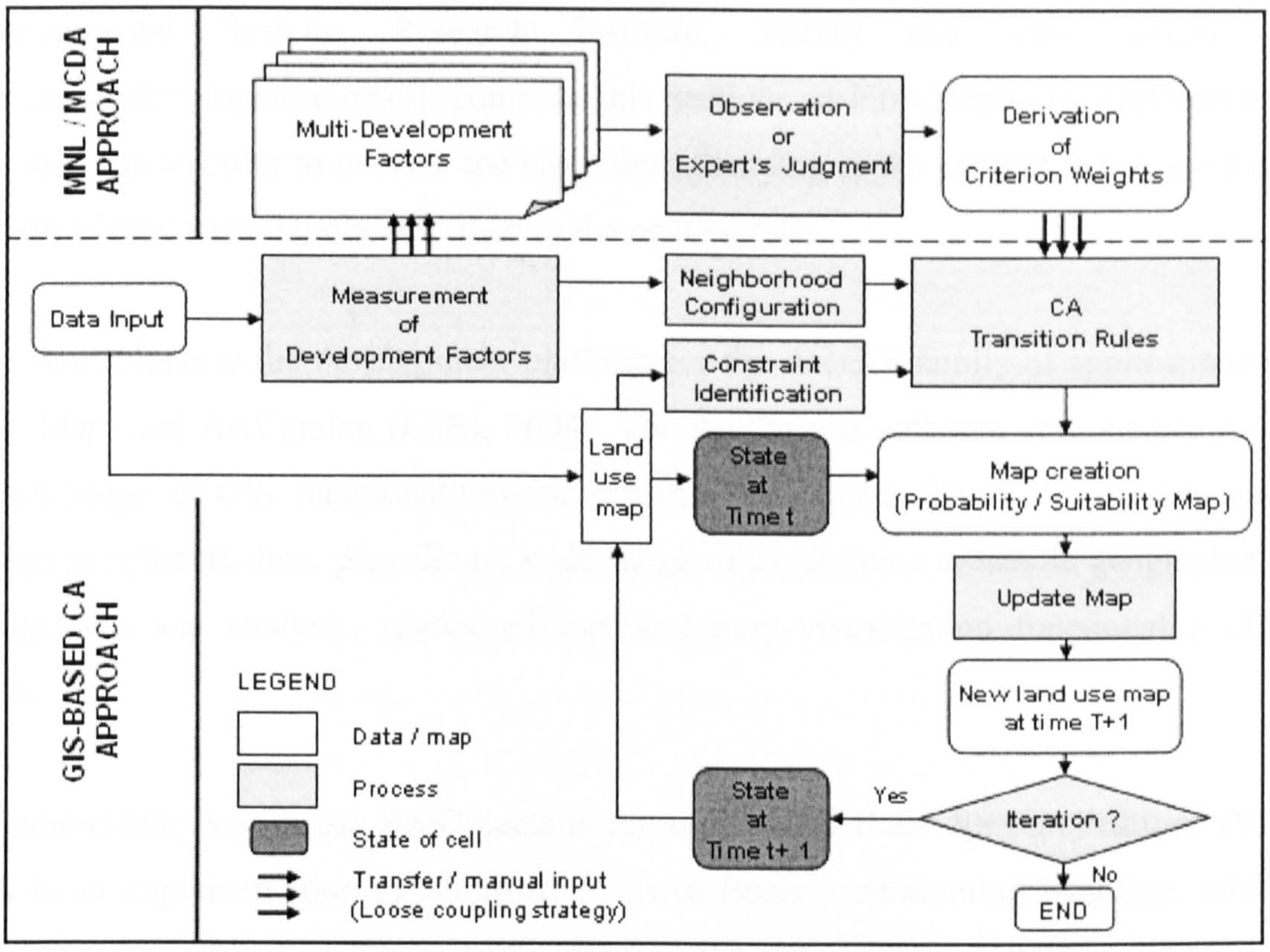


Figure 5.1: General methodology framework developed for the proposed model.

There are many programming languages, such C++, Java and Visual Basic (VB) that can be used to program a model. In this chapter, however, the proposed conceptual framework and the methodology presented in Chapter 4 is implemented using ArcGIS 9.1 (and updated in ArcGIS 9.2) VBA macro. This chapter focuses on the development, using the above framework, of a set of customized tools to enable the user to simulate land use change dynamically. The organization in this chapter is as follows. The next section of this chapter (Section 5.2) presents the concept of VBA and ArcObjects used for the model development. It is followed by the description of the user interface and execution of the tools developed in Section 5.3. Finally, a discussion that demonstrates the advantages and limitations of the tools developed is discussed in Section 5.4.

5.2 ArcObjects and Visual Basic for Applications (VBA)

ArcObjects and Visual Basic for Applications (VBA) are popular amongst ArcGIS developers as they can be used to create their own applications. Access to documentation of ArcObjects and Visual Basic for Applications (VBA) is available from Exploring ArcObjects Vol. 1 – Application and Cartography (Environmental Systems Research Institute, 2001a), Exploring ArcObjects Vol. 2 – Geographic Data Management (Environmental Systems Research Institute, 2001b) and VBA online help (http://arcgisdeveloperonline.esri.com). In this section, an introduction to ArcObjects and VBA is given in order to provide the basic understanding in the context that is used in the creation of customized tools described in the next section.

ESRI ArcObjects is the development platform for the ArcGIS family of applications such as ArcMap, and ArcCatalog (ESRI, 2004). The ArcObjects software components provide the full range of GIS functionalities for software developers. These objects are used to manage geographic data, providing a wide range of capabilities to handle geographic data management and analysis, spatial editing, and map visualization functionality (ESRI, 2004).

One convenient way to use ArcObjects is through Visual Basic for Applications (VBA). VBA is an implementation of Microsoft's Visual Basic programming language, which is built into most Microsoft Office applications. VBA comes with ArcGIS software, being embedded in ArcGIS applications, both ArcMap and ArcCatalog. VBA provides an integrated programming environment, so-called the Visual Basic Editor (VBE), which allows a Visual Basic (VB) macro to be written, debugged and tested straightforwardly in the ArcGIS environment. A VBA macro can integrate both the VB's functionalities (e.g. text boxes for input) and the ESRI ArcObjects Libraries. Through VBA, programmers can control data management and map presentation tasks, customize the ArcGIS Desktop applications as well as extend ArcGIS with their own custom commands, tools, and menus.

The underlying concept of VB and VBA is the object oriented programming. Programming with VB and VBA code, thus, means using objects. Elements like UserForms and CommandButtons are examples of VB objects while elements like maps and layers are all examples of ArcGIS objects. These objects can be manipulated by using their properties, methods, and events. Properties are characteristic (noun) of objects. When working with object properties, developers can read (get) or write (set) them. For example, the ArcObjects called RasterBand object represents a single band of a raster dataset on disk. It has properties such as *Bandname* (the name of this raster band), *Statistics* (the statistics of this raster band), and *Histogram* (the histogram of this raster band). Methods are actions (verbs) that objects know how to perform. For example, the same object, RasterBand, has a method such as *ComputeStatsAndHist* (Calculates statistics and histogram). Events are another type of action (verb). Unlike methods, events are actions to which an object responds. These event actions are usually triggered by the user (e.g. by pressing the button). For example, the Map object can respond to the user when the mouse is moved or clicked.

To work with the objects, a programmer does not communicate with properties and methods directly. Instead, these methods and properties are communicated through an interface. An object can have many interfaces, in order to access various functionalities. Figure 5.2 shows a part of the object model diagram (OMD), ESRI ArcObjects Library that is used to help work with interface. For example, the RasterBand object (see Figure 5.2) previously described uses an IRasterBand interface to work with properties and methods outlined above. The same object uses the IRasterProps interface to access and control raster properties such as *Width*, *Height*, and *Extent* properties (in order to set and get width, height, and extent of the pixel respectively).

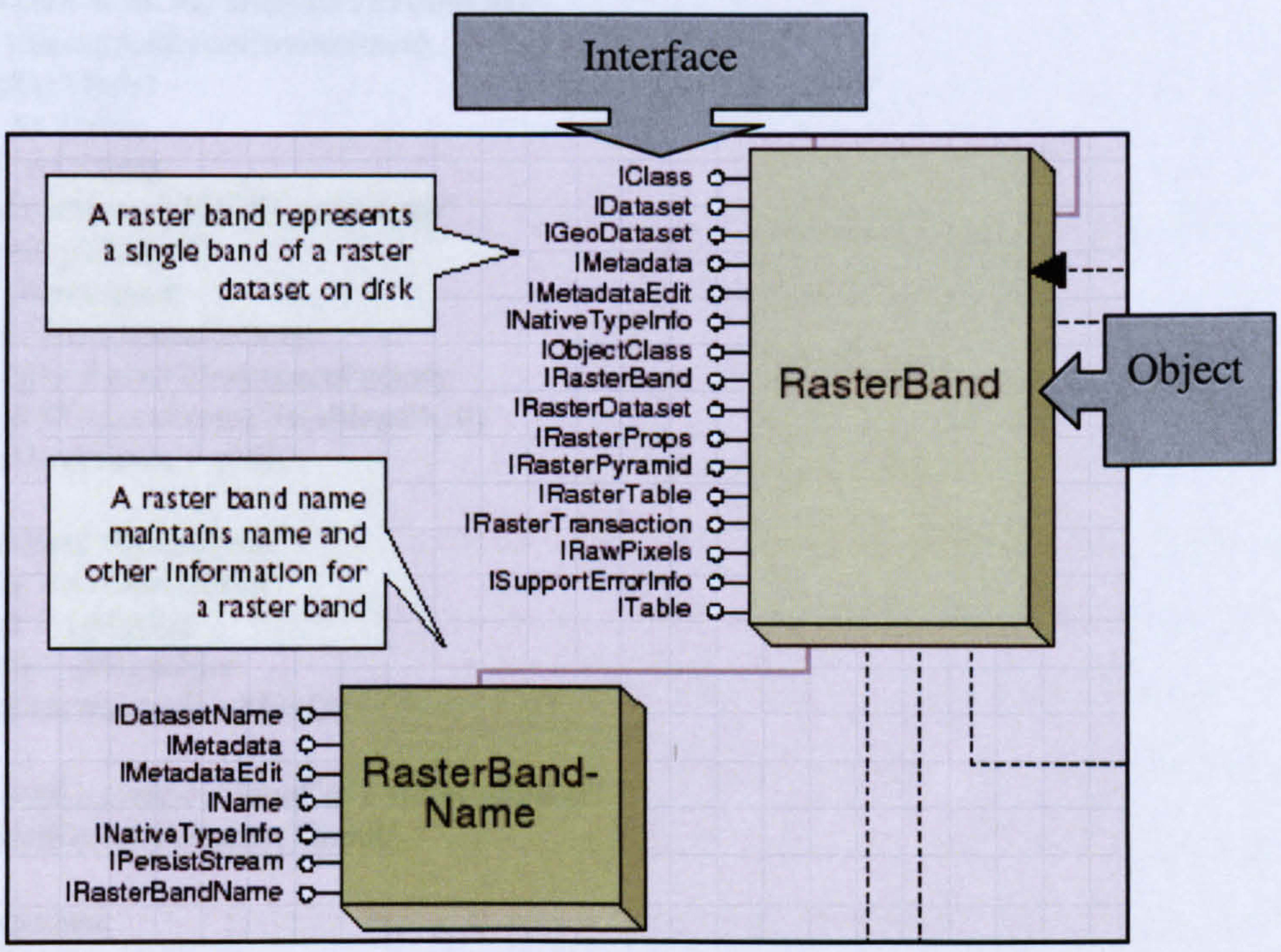


Figure 5.2: The RasterBand object and its interface, a part of the object model diagram of "Exploring ArcObjects Vol. 2 – Geographic Data Management" (Environmental Systems Research Institute, 2001b).

The box below shows an example of the VBA script that has been used in this project. This specific code was written to execute the calculation of neighbourhood effect in the study.

Algorithm of the script is given in pseudo code shown in function F:NI_MEASURE in Section 5.3.1.

```
Public Function MyFocalSum(pInraster As IRaster, iFocalsize As Integer, pMasker As IRaster) As IRaster
'Description: This program is used to calculate neighborhood effect using the FOCALSUM function.
'Input:
'  pInraster  # A raster grid layer.
'  pMasker    # A masking grid layer.
'  iFocalsize # A window size.
'  pMasker    # A masking grid layer.
'Output:
'  return MyFocalsum as IRaster.
'Remark: Interface employed in the function.
' IRaster: Provides access to members that control an in-memory raster (transient image).
' IRasterModel: Provides access to members that control the Raster Model.
' IRasterAnalysisEnvironment: Provides access to members that control the environment for raster analysis.
' IWorkspace: Provides access to members that have information about the workspace.
' IWorkspaceFactory: Provides access to members that create and open workspaces and supply
' workspace factory information.
' IGeoDataset: Provides access to members that provide information about a Geographic Dataset.

On Error GoTo EH

'Create a RasterModel object from IRasterModel interface.
  Dim pRasModel As IRasterModel
  Set pRasModel = New RasterModel

'Set output workspace from the analysis environment
  Dim pEnv As IRasterAnalysisEnvironment
   Set pEnv = pRasModel
  Dim sfilepath As String
  Dim sfilepath1 As String
   sfilepath = "c:\lumics-a3703781\scratchws"
   sfilepath1 = sfilepath & "\"
  Dim pWS As IWorkspace
  Dim pWSF As IWorkspaceFactory
   Set pWSF = New RasterWorkspaceFactory
   Set pWS = pWSF.OpenFromFile(sfilepath, 0)
   Set pEnv.OutWorkspace = pWS

'Set mask environment for analysis
  Dim pMaskEnv As IGeoDataset
   Set pMaskEnv = pMasker
   Set pEnv.Mask = pMaskEnv
  pEnv.SetExtent esriRasterEnvMaxOf

'Method used to bind a symbol "input"to a raster pInraster
  pRasModel.BindRaster pInraster, "input"

'Write script statements.
 Dim iFocalsizexsize As Integer
  iFocalsizexsize = iFocalsize * iFocalsize
 Dim sScript1 As String
 Dim sScript2 As String
  sScript1 = "[outputfocalsum] = focalsum([input],rectangle," & iFocalsize & "," & iFocalsize & ")"
  sScript2 = "[output] = [outputcon] * 100 / " & (iFocalsize * iFocalsize) - 1

'Set Property: specifying the model scripts to be executed (map algebra expression).
'Remark: use vbLf to create separate lines
```

```
pRasModel.Script = sScript1 + vbLf + _
                   "[outputcon] = con([input] == 0,[outputfocalsum],[outputfocalsum] - 1)" + vbLf + _
                   sScript2

'Method used to produce a raster by executing scripts (map algebra expression(s)).
  pRasModel.Execute

'Method used to unbind a symbol
  pRasModel.UnbindSymbol "input"

'Get output rasters created from Map algebra
  Dim pOutfocalsum As IRaster 'integer raster
  Set pOutfocalsum = pRasModel.BoundRaster("output")

'Return value
  Set MyFocalSum = pOutfocalsum


'Release memory
  Set pWS = Nothing
  Set pWSF = Nothing
  Set pEnv = Nothing
  Set pMaskEnv = Nothing
  Set pRasModel = Nothing
  Set pOutfocalsum = Nothing

Exit Function

EH:
 Set MyFocalSum = Nothing

End Function
```

5.3 Development of Customized Tools: Graphical User Interface and Execution

A set of tools developed named LUMICS (Land Use MICro-Simulation model), is a GIS-based CA model, where the multinomial logistic regression (MNL) and multicriteria decision analysis (MCDA) methods have been integrated to identify the potential cells for development, that simulates the land use change in the study. In this section, a description of the customized tools developed (see Figure 5.3) is presented. They are the Variables Observation tool, MNL (GIS-based CA approach) tool, and MCDA (GIS-based CA approach) tool. The tools developed here are designed to be operated separately. The first one, the Variables Observation tool, is designed to help observe and measure the characteristics of development factors (e.g. proximity to road) with respect to the considered land use types (e.g. residential, commercial) used in the study site. The outcome produced from this tool is used as an input for the MNL 's derivation of criteria weights described in Section 4.5.1 The second tool is referred to as MNL (GIS-based CA approach) which is developed to create a land use simulation on the basis of CA-based MNL rule. The last one is MCDA (GIS-based CA approach) which is developed to create a land use simulation on the basis of CA-based MCDA rule. Despite both the second (MNL-based method) and the third (MCDA-based method) tools being used to create a land use simulation on the basis of a set of criterion weights, they differ in terms of the algorithms employed for the calculation of development probability. While the MNL method applies the multinomial logistic regression technique, the MCDA method uses the simple additive weight technique. Details about MNL and MCDA methods used for computation of development probability are described in Section 4.2 and Section 4.3 respectively.

Note that this thesis also provides the installation guide and user's guide along with a CD-ROM to be used for running these tools, which are given as exercises in Appendix 1 and Appendix 2 correspondingly.

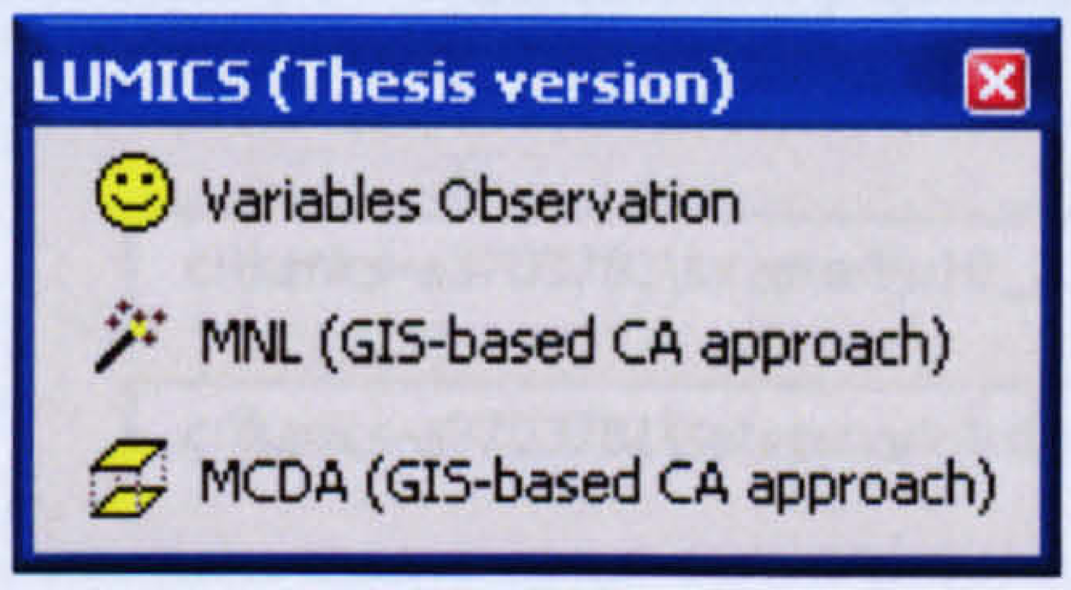


Figure 5.3: Customized tools developed for the research study.

5.3.1 Variables Observation Tool

5.3.1.1 Graphical User Interface (GUI) of Variables Observation Tool

The Variables Observation tool facilitates the creation of the combination of development factors (e.g. residential neighbourhood effect, proximity to roads) being used in the analysis. Figure 5.4 shows the graphical user interface (GUI) for the Variables Observation tool. It is separated into three sections. The first section is for layer input. This section allows users to input layers needing to be measured. It includes both the compulsory layer and the optional layer. The compulsory layers are land use layer (grid format) and road feature class (shape file format). The optional layers include planned road (shape file format), land price (grid format).

The second section refers to environment setting. This section allows users to fill in three types of setting, including masking layer, output workspace (directory), and neighbourhood setting. A masking layer is used to define the selected location to extract. For example, if only the cells experiencing the transition (e.g. between 1993 and 2001) are of interest, the land use change areas between these two epochs are used as a masking layer. The outcome of the analysis produced, thus, contains only cells that correspond with the areas defined by a mask. The second setting refers to the output workspace. This field allows users to specify directory where the output is generated. The final field, neighbourhood setting, allows users to set the window size of the neighbourhood effect regarding residential, commercial, and industrial data being generated. In this study in which land use layer is based on the 10m square grid cell, the window size of 21 means a walking radius of 10 cells by 10 cells, or 100m walking distance.

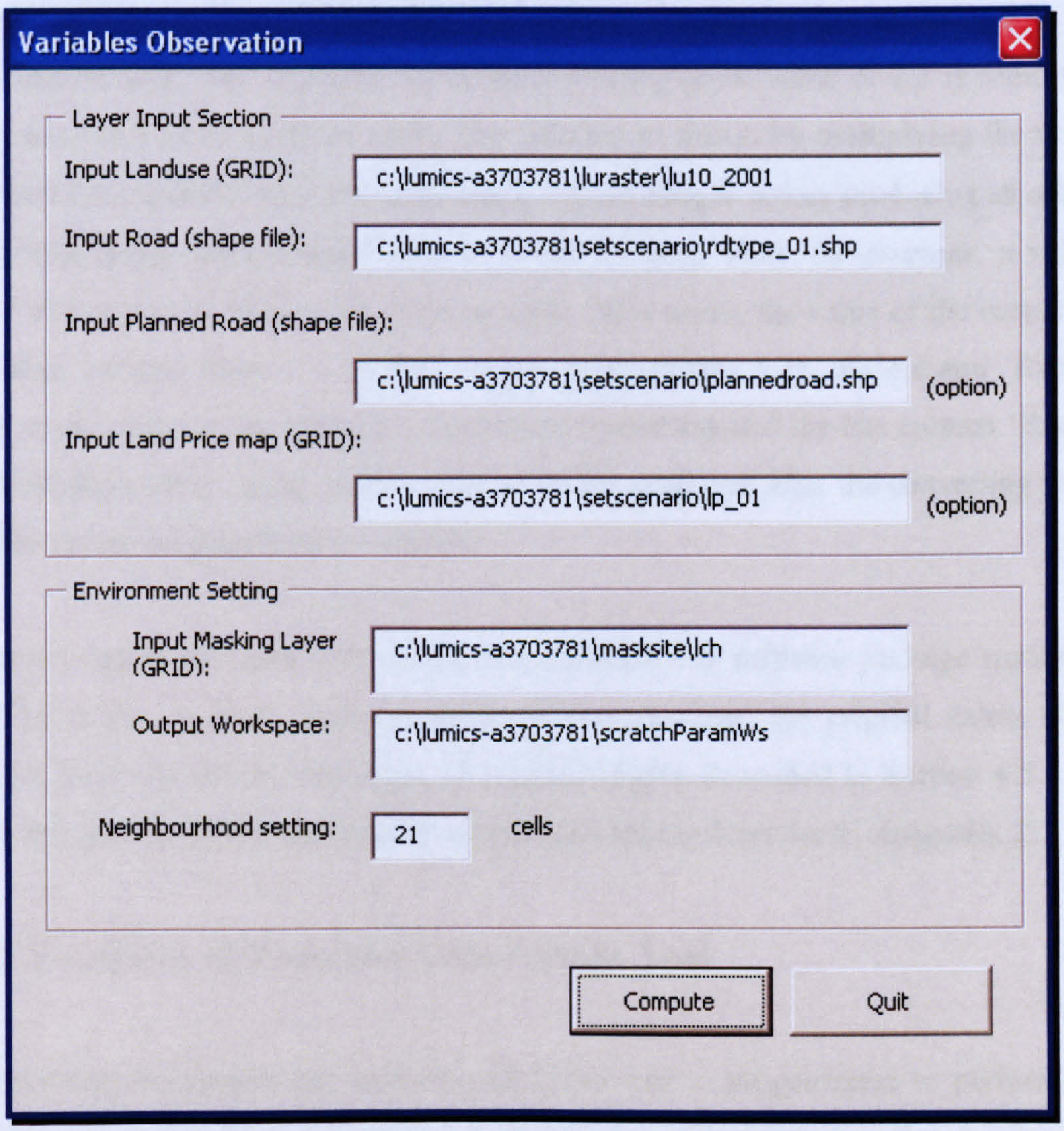


Figure 5.4: Graphical user interface: Variables Observation pop-up menu.

The last section comprises two buttons; compute and quit button. The compute button is used to execute the program using the parameters currently set in the GUI while the quit button allows users to cancel the operation. Once the compute button is clicked, the program will create an output layer with an associated attribute table. The attribute table contains a number of the measured development factors (attributes) in associated with the considered land use types regarding residential, commercial, industrial and vacant land. Table 5.1 lists a series of the development factors created.

It should be noted that due to the technical limitations of the software that a raster layer has no table associated with it. To solve this problem, changing a continuous raster layer to an integer layer is performed first so that the "attribute table" can be produced. Most output layers in relation to the development factors produced in this study contain the original value, mostly having a value ranging from 0.0 to 1.0. When those layers are converted from floating point to integer value, they can produce only the integer values of 0 or 1 in a raster attribute table. For example, the original floating point value of 0.2 is truncated to integer value of 0 in an attribute table. The solution to this is by multiplying the original values with the constant value and converting it to an integer before producing an attribute table. In this study, the constant value of 10000 is used. Thus, for example, a value of 0.34657 will appear as 3456 in an attribute table. As a result, the value of the output table has a value ranging from 0 – 10,000. In the Table (Table 5.1), the column 'Range of original value' refers to the original value before converting and the last column 'Range of value in attribute table' refers to the modified values produced after the converting process which has values ranging from 0 – 10,000.

The outcome generated here after transferring to the SPSS software package requires the conversion to the original values by users. After converting, the original values will be used as an input file for the derivation of criteria weights described in Section 4.5.1. User may run this tool by following the user's guide provided in Exercise 1, Appendix 2.

5.3.1.2 Execution of Variables Observation Tool

After receiving the parameters from the GUI, the tool is programmed to perform three main tasks as listed in Table 5.2, which can be split into nine main execution functions (see Figure 5.5).

Field no.	Attribute Description	Unit	Range of original value	Range of value in attribute table
1	Land use 1993	Land use categories*	1 – 10	1 – 10
2	Land use 2001	Land use categories*	1 – 10	1 – 10
3	Residential neighborhood effect within the specified walking distance	Percent	0 – 100**	0 – 100
4	Commercial neighborhood effect within the specified walking distance	Percent	0 – 100	0 –100
5	Industrial neighborhood effect within the specified walking distance	Percent	0 – 100	0 –100
6	Proximity to all roads	normalization	0 – 1***	0 – 10000
7	Proximity to main roads	normalization	0 – 1***	0 – 10000
8	Proximity to collector roads	normalization	0 – 1***	0 – 10000
9	Proximity to streets	normalization	0 – 1***	0 – 10000
10	Proximity to residential area	normalization	0 – 1***	0 – 10000
11	Proximity to commercial area	normalization	0 – 1***	0 – 10000
12	Proximity to industrial area	normalization	0 – 1***	0 – 10000
13	Proximity to government area	normalization	0 – 1***	0 – 10000
14	Proximity to school	normalization	0 – 1***	0 – 10000
15	Proximity to park /recreation area	normalization	0 – 1***	0 – 10000
16	Land price	Thai Baht	0 – 1***	0 – 10000
17	Proximity to planned roads	normalization	0 – 1***	0 – 10000
18	Proximity to agriculture area	normalization	0 – 1***	0 – 10000

Table 5.1: List of the attributes generated from the Variables Observation tool. Remark: * refers to ten types of land use categories where value 1 = residential, 2 = commercial, 3 = industrial, 4 = government, 5 = school, 6 = park/conservation, 7 agriculture, 8 = road, 9 = river, 10 vacant land. The meaning ** is that 0 refers to low neighbourhood proportion while 1 refers to high neighbourhood proportion. The meaning *** is that 0 refers to far from activities while 1 close to activities.

The program starts with the extraction of data used for the analysis including land use types and road types. The second step is to create the development factors using two GIS techniques: Focal (e.g. the creation of residential neighbourhood index) function and

Euclidean distance function (e.g. the creation of proximity to main roads). The final step is to combine these development factors in association with the land use types considered.

Task Description	Function	Action
1. Extraction of data	F:LUTYPE_EXTRACT	Extraction of land use type
	F:RDTYPE_EXTRACT	Extraction of road type
2. Creation of development factors	F:NR_MEASURE	Neighborhood (focal) measurement
	F:LU_EUCDIST	Euclidean distance for land use data
	F:RD_EUCDIST	Euclidean distance for road data
	F:RDTYPE_EUCDIST	Euclidean distance for road type data
	F:MINIMIZE	Searching for the smallest minimum value
	F:MAXIMIZE	Searching for the smallest maximum value
3. Combination of development factors	F:VAR_CMB	Combination of development factors

Table 5.2: GIS Functions used to build Variables Observation tool.

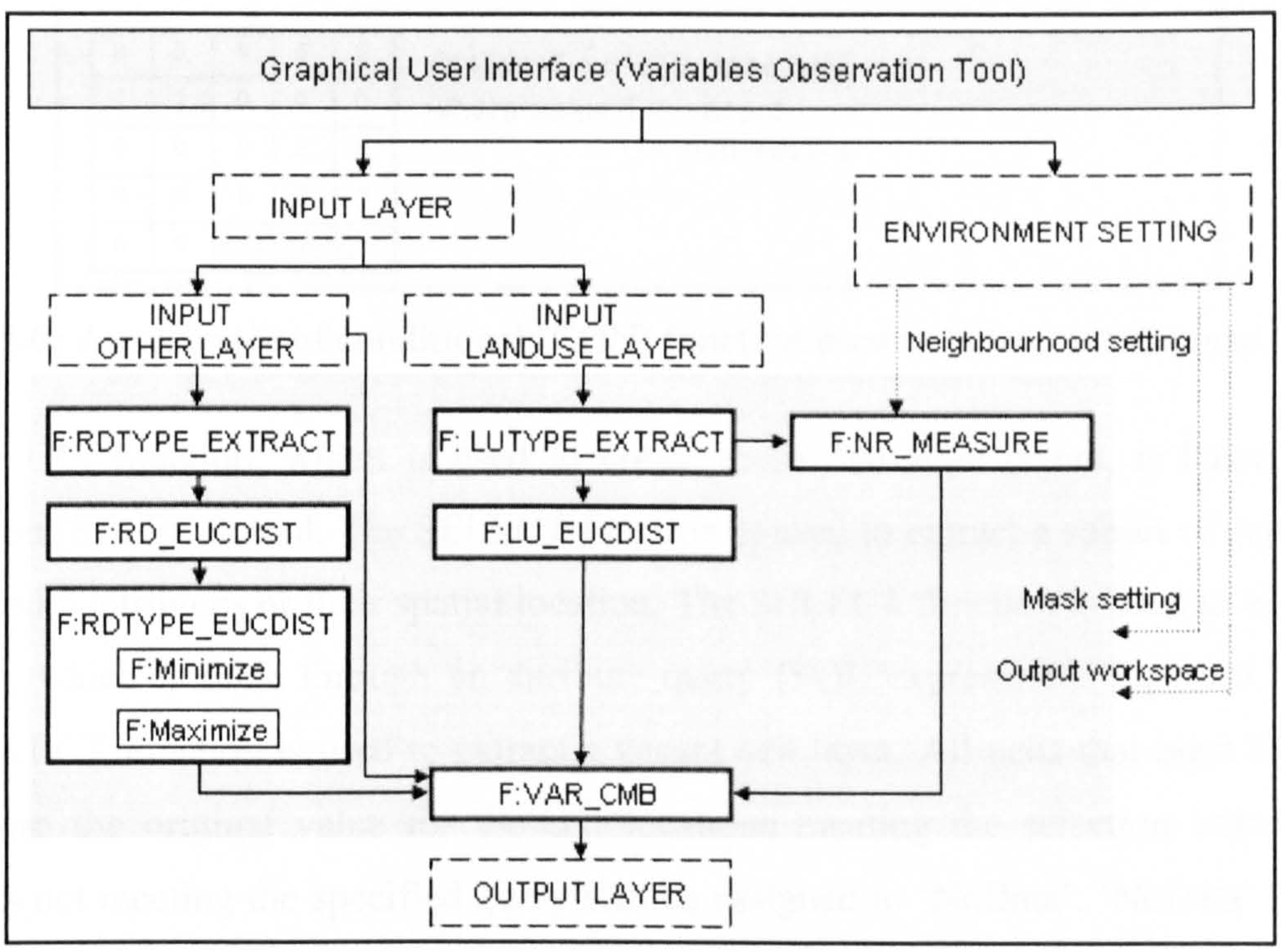


Figure 5.5: Model developed for Variables Observation tool.

Function F:LUTYPE_EXTRACT (Extraction of land use type). This function extracts land use classes from a land use grid layer into separate grid layers. The function is built based on the ArcGIS conditional function (CON) and extraction function (SELECT).

The CON function is used to control the output value for each cell based on whether the cell value evaluates to True or False in a specified conditional statement (ESRI, 2004). Figure 5.6 illustrates an example showing how the VACANT grid layer is created from the LANDUSE grid layer using the conditional CON function. In the figure, the input cell coded as '10', after performing the CON function, will evaluate to True and return a value of '1', or otherwise return a value of '0' (False value).

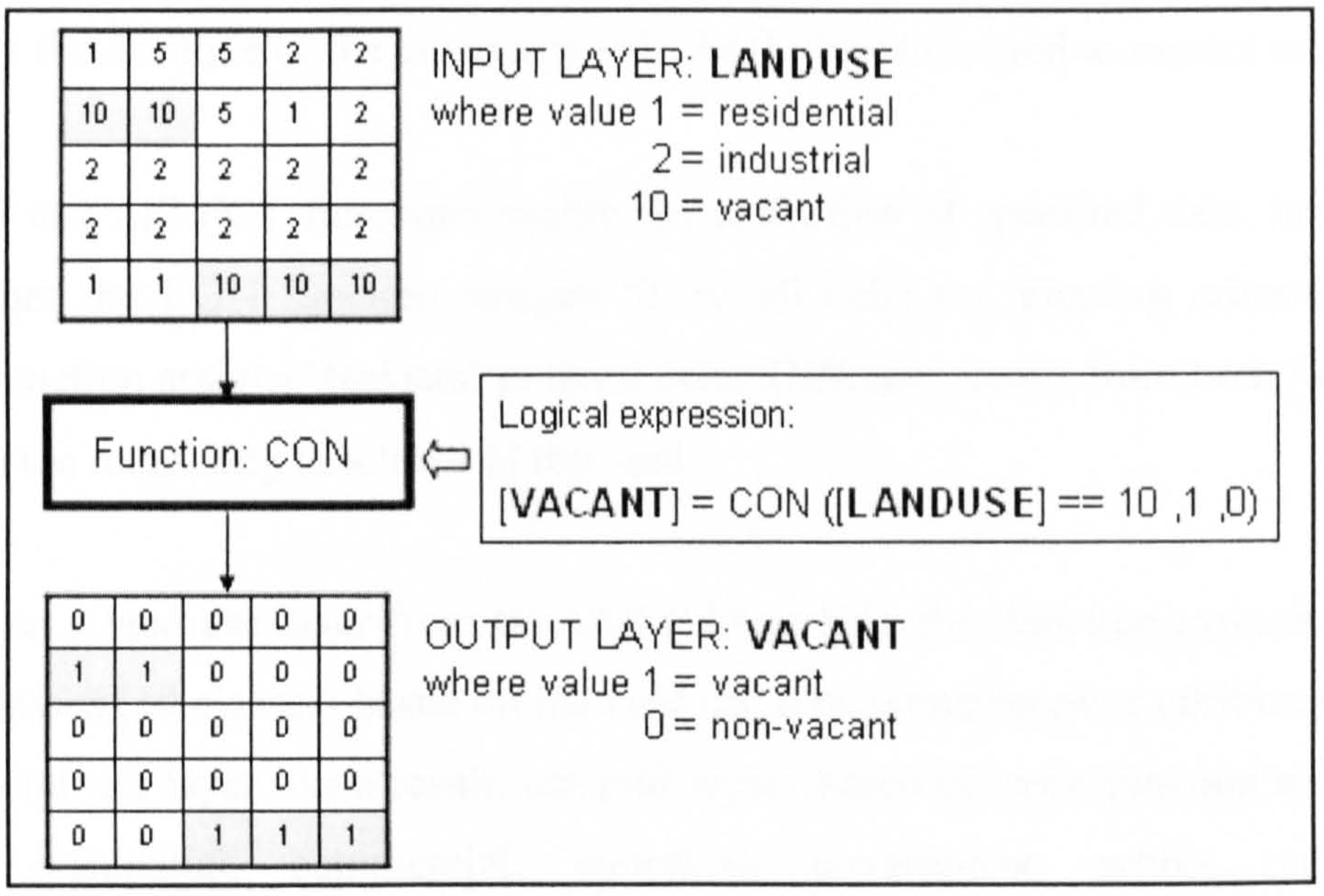


Figure 5.6: An example of conditional (CON) function used to extract vacant map.

Another GIS function, which is used to create these cell state layers, is based on the Extraction (SELECT) tool. The SELECT function is used to extract a subset of cells either by the cells' attributes or their spatial location. The SELECT function is used to extract an attribute, which is done through an attribute query (SQL expression). Figure 5.7 shows how SELECT function is used to extract a vacant cell layer. All cells that meet the query will return the original value for the cell locations meeting the selection criteria. Cell locations not meeting the specified query will be assigned to 'NoData'. 'NoData' refers to unknown cells (cells not being used for the analysis).

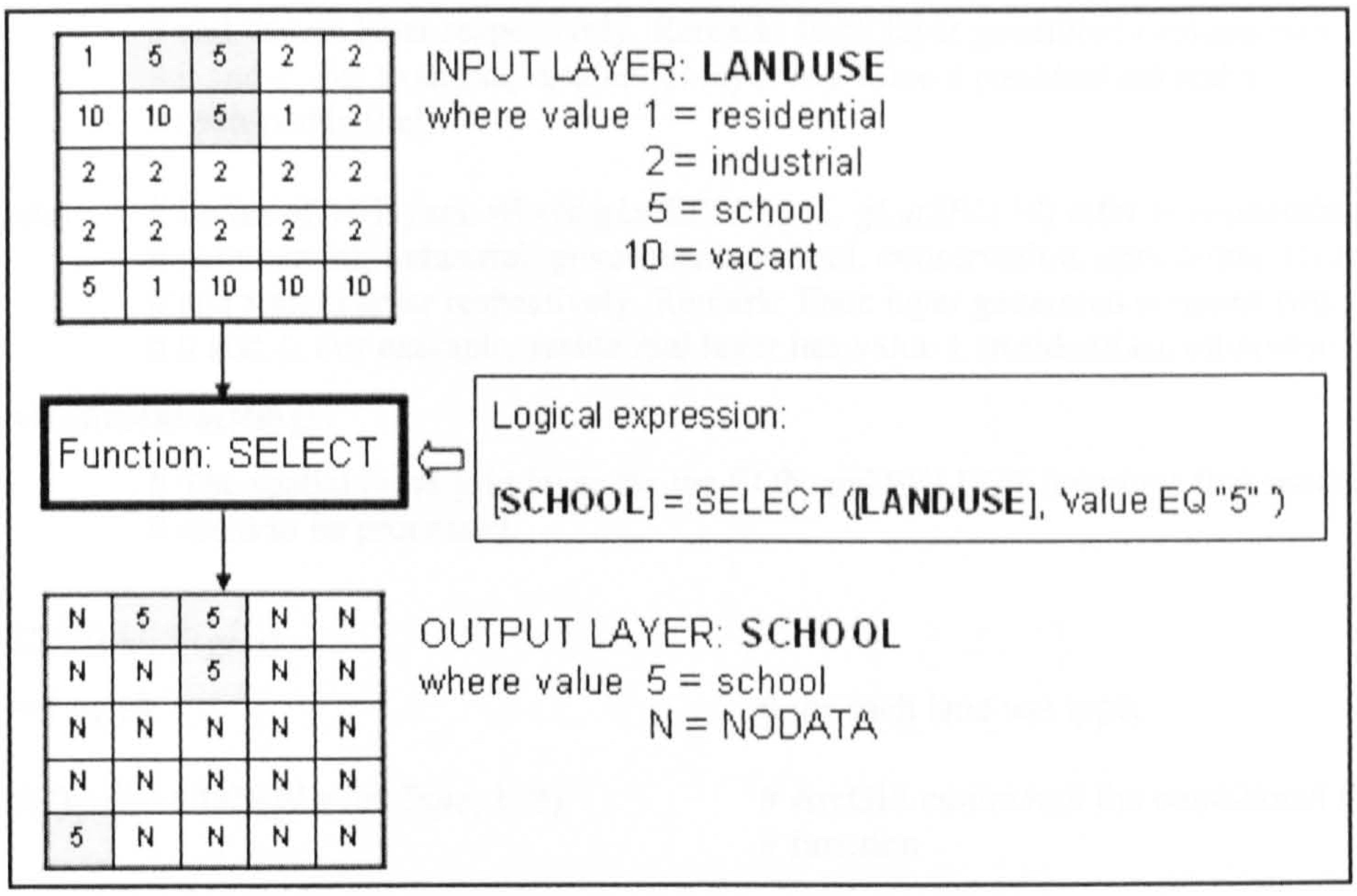


Figure 5.7: An example of the extraction (SELECT) function used to extract school map.

Both CON and SELECT functions enable the extraction of specified data, however they differ in that the CON function assigns '0' to all cells not meeting criteria while the SELECT function assigns 'NoData' to those cells. Different results from both functions are required in the remaining functions of this tool.

Having input a land use layer from the GUI (Figure 5.4), this function extracts a series of land use classes (10 classes) based on the land use type (considered as attribute) field from the input land use layer. As a result, ten grid layers based on each function are generated regarding residential, commercial, industrial, government, school, conservation, agriculture, river, road, and vacant layer.

The algorithm of function **F: LUTYPE_EXTRACT** shown in the pseudo code below is used to execute the land use data. While line 2 is conducted to create an output layer using the conditional CON function, line 3 is performed using the extraction SELECT function.

Algorithm F: LUTYPE_EXTRACT

Input:

gLu # A land use grid layer comprising 10 land use types.
Type # A land use type.

Output:

gLutCON(*Type*) # Series of 10 layers, where *gLutCON*(*1*),…., *gLutCON*(*10*) refer to residential,
commercial, industrial, government, school, conservation, agriculture, river, road

and vacant layer respectively. Remark: Each layer generated contains two values,
0 and 1. For example, residential layer has value 1 (residential) and 0
(non-residential).

gLutSEL(Type) # Series of 10 layers, where *gLutSEL(1)*,…., *gLutSEL(10)* refer to residential,
commercial, industrial, government, school, conservation, agriculture, river, road,
and vacant layer respectively. Remark: Each layer generated contains two values,
0 and 1. For example, residential layer has value 1 (residential), otherwise NoData.

Analysis environment settings:

gMask # The spatial mask grid layer for the CON and SELECT functions that specifies the
areas to be processed.

LUTYPE_EXTRACT(*gLu*)

1. **For** *Type* = 1 to 10 # for each land use type.

2. *gLutCON(Type)* = **CON**(*gLu* == *Type*, 1, 0) # ArcGIS command: the conditional CON
function.

3. *gLutSEL(Type)* = **SELECT**(*gLu*, 'value EQ "*Type*" ') # ArcGIS command: the extraction SELECT
function.

4. **EndFor**

Function F: RDTYPE_EXTRACT (Extraction of road type). This function extracts road types from a road shape file into three separate road-type shape files. The function is built based on the SELECT function (analysis toolbox). The logical expression of the SELECT function previously described in Function **F: LUTYPE_EXTRACT** is for the analysis of raster layers. Logical expression of the SELECT function here is to work with a feature class (e.g. a GIS coverage, or a shape file). However, they both are used for the extraction purpose. The SELECT function here enables the extraction of data specified through a SQL expression, in order to select a subset of features by attributes (road types). As a result, three layers of road types in a shape file format are generated regarding major roads, collector roads and streets. Algorithm **F: RDTYPE_EXTRACT** shows in a pseudo code as follows.

Algorithm F: RDTYPE_EXTRACT

Input:

sRd # A road layer in shape file (e.g. c:\workspace\road.shp)
Type # A road type.
RdField # A Field name of classified road type (e.g. RDTYPE).

Output:

sRdtSEL(Type) # Series of 3 road type layers in shape file, where *sRdSELtype(1)*, *sRdSELtype(2)*,
sRdSELtype(3) refer to the output shape file of major roads, collector, roads and
streets respectively.

Analysis environment setting:

gMask	# The spatial mask grid layer for the SELECT functions that specifies the areas to be # processed.

RDTYPE_EXTRACT(*sRd*)

1. **For** *Type* = 1 to 3 # for each road type.

2. **SELECT_ANALYSIS** *sRd sRdtSEL(Type) "RdField" = 'Type'* # ArcGIS command: the # SELECT (analysis toolbox) # function.

3. **EndFor**

According to line 2 of the algorithm **F: RD_EXTRACT**, suppose that the main road (*Type* = 1) is required to be extracted, based on the input road shape file (e.g. c:\workspace\rd.shp), the output road type file (e.g. c:\workspace\rdt1.shp) and the road filed name (ROADTYPE). An example of logical expression can be expressed as follows:

SELECT_ANALYSIS c:\workspace\rd.shp c:\workspace\rdt1.shp "ROADTYPE" = '1'

Function F: NR_MEASURE (Neighbourhood measurement). With this function, neighbourhood setting from the GUI is used to set the window size. The window size here is to determine the maximum distance of the neighbouring cells, aiming to capture the spatial influence within a neighbourhood for the analysis. The function is built based on a combination of the FOCALSUM function and arithmetic operators.

The focal or neighbourhood function creates output values for each cell location based on the value for the location and the values identified in a specified neighbourhood (ESRI, 2004). There are many choices of ArcGIS Focal Statistic operation (e.g. maximum, mean, sum). In this study, the ArcGIS FOCALSUM function is employed in order to sum up the neighbouring cells in the window and create a neighbourhood index layer for residential, commercial, and industrial areas. Figure 5.8 illustrates an example showing how FOCALSUM function is used to produce the industrial neighbourhood layer (NR_IND) from the INDUSTRIAL layer using the square neighbourhood of width and height (3 by 3) in cell units.

To create a neighbourhood index layer as described in Equation 4.10, the ArcGIS conditional CON function and the arithmetic operator are employed. The concept of the ArcGIS conditional CON function is illustrated earlier in Figure 5.6 while the concept of arithmetic operator is shown in Figure 5.9.

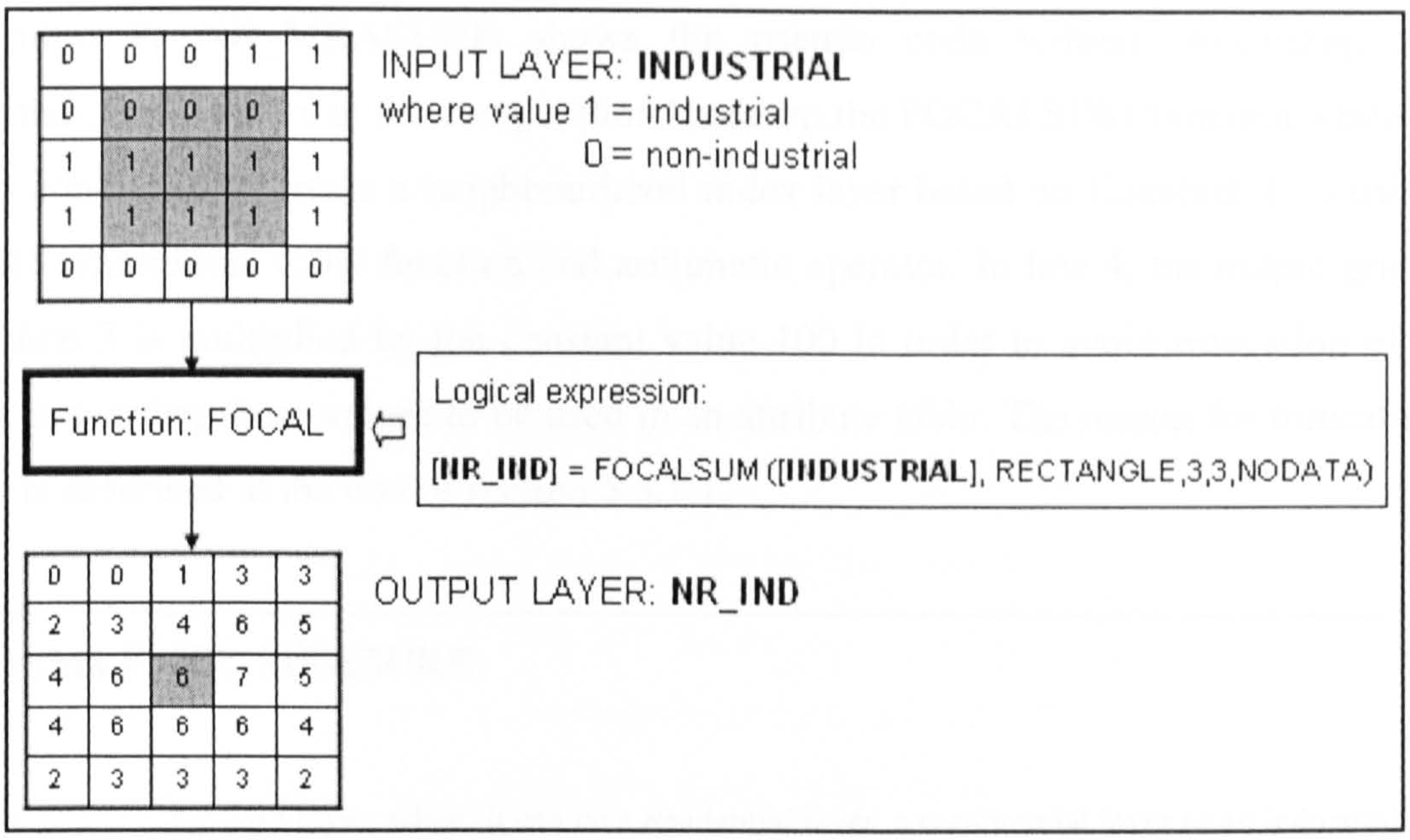


Figure 5.8: An example of the FOCALSUM function used to create the summation of industrial neighbouring cells (NR_IND) map.

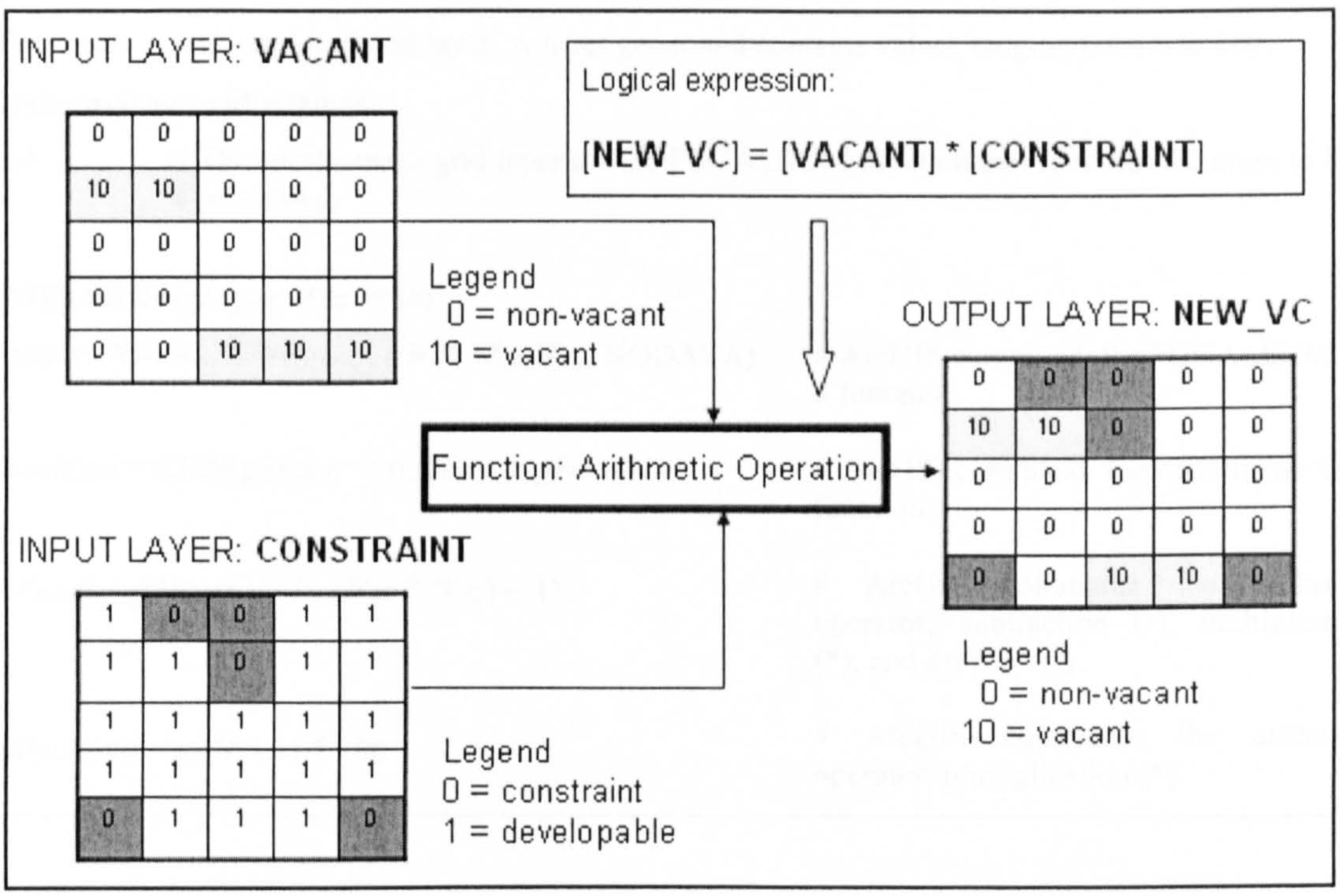


Figure 5.9: An example of arithmetic operation function used to create a new vacant map (NEW_VC).

The arithmetic mathematical operation can be used for calculation of two or more input grid maps using arithmetic operators; addition (+), subtraction (-), multiplication (*), and division (/). Figure 5.9 shows an example of the arithmetic operation function. Here a vacant map (VACANT) is multiplied by the constraint cells (CONSTRAINT) to create a new developable vacant map (NEW_VC).

Algorithm **F: NR_MEASURE** shows the pseudo code written. According to the algorithm, line 1 refers to the computation based on the FOCALSUM function while line 2 – 3 is conducted to create a neighbourhood index layer based on Equation 4.10 using the ArcGIS conditional CON function and arithmetic operator. In line 4, the output grid layer from line 3 is multiplied by the constant value 100 in order to avoid truncation of value when converting these values to be used in an attribute table. The reason for truncating the value is described at the end of section 5.3.1.1.

Algorithm F: NR_MEASURE

Input:

gLayer	# A grid layer, where it can be a residential layer, a commercial layer or an industrial layer # derived from function F:LU_EXTRACT using the CON function.
Wsz	# A window size.
Wsh	# A window shape. In this study, a RECTANGLE shape is used.

Output:

gNiPctFocal	# A neighborhood layer. A layer generated contains values ranging from 0 to 100.

Analysis environment settings:

gMask	# The spatial mask grid layer for the FOCALSUM function that specifies the areas to be # processed.

NR_MEASURE(*gLayer, Wsz, Wsh*)

1. *gFocal* = FOCALSUM(*gLayer, Wsh, Wsz, Wsz,* NODATA)	# ArcGIS command: the FOCALSUM # function.
2. *gFocalCon* = CON(*gFocal* == 0, *gFocal, gFocal*-1)	# ArcGIS command: the conditional CON function.
3. *gNiFocal* = *gFocalCon* / ((*Wsz* * *Wsz*) – 1)	# ArcGIS command: the arithmetic operator; subtraction (-), multiplication (*), and division (/).
4. *gNiPctFocal* = *gNiFocal* * 100	# ArcGIS command: the arithmetic operator; multiplication (*).

In order to create the residential, commercial, and industrial neighbourhood layers, the function (F: NR_MEASURE) is run three times. For the first run, a residential layer derived from function **F:LUTYPE_EXTRACT** is used as an input layer to produce the output, residential neighbourhood layer. For the second and the third run, input is a commercial and an industrial neighbourhood respectively. An output layer contains integer value ranging from 0 to 100. Close to 0 refers to no neighbouring cell of the considered land use type within the specified area while close to 1 refers to a high proportion of neighbouring cells.

In fact, according to the model developed, the neighbourhood measurement is not used as a part of MNL computation for weight derivation. Instead, it will be used to help observe the statistical association between neighbourhood factors and the land use type considered for derivation of neighbourhood thresholds in Section 4.5.3.

Function F: LU_EUCDIST (Euclidean distance for land use data). This function is built based on the ArcGIS EUCDISTANCE function. The function provides a distance measurement for raster (grid) layers (e.g. school, government). The layers derived from this function (e.g. proximity to schools, proximity to government areas) are the layers of development factors used as main input for model simulation.

Figure 5.10 demonstrates an example showing how the proximity to schools (D_SCH) layer is created from the SCHOOL grid layer using the Euclidean distance function. With ArcGIS 9.2, true Euclidean distance is calculated to each cell in the distance function. For each cell, the distance is calculated to each source cell by calculating the hypotenuse with the x-max and y-max as the other two legs of the triangle statement (ESRI, 2004). This calculation derives the true Euclidean, not cell, distance. The shortest distance to a source is determined and the value is assigned to the cell location on the output raster.

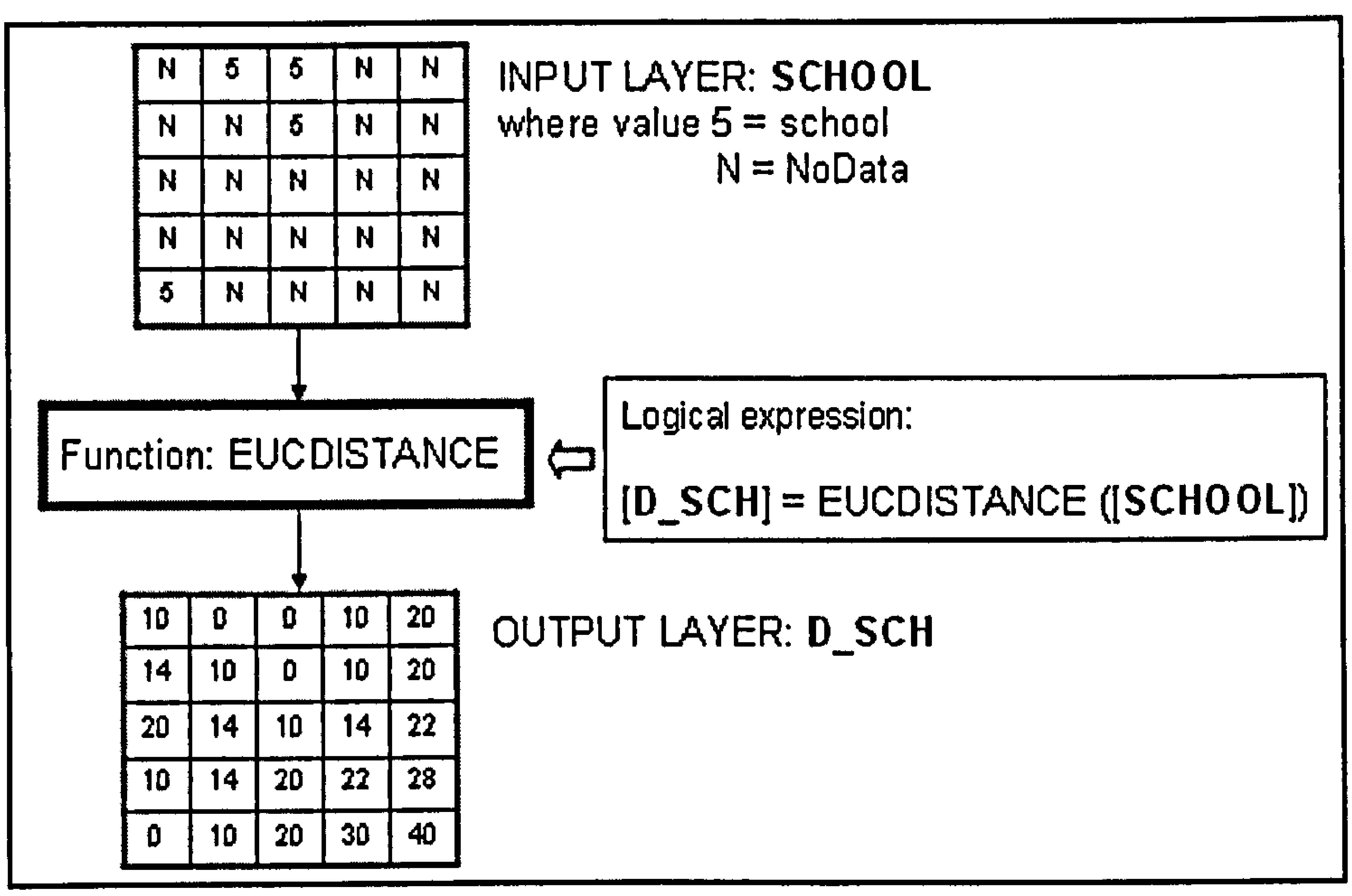


Figure 5.10: An example of euclidean distance (EUCDISTANCE) function used to extract proximity to schools (D_SCH) map.

The measurement of development factors described above is then normalized to have a value ranging from 0.0 to 1.0.

The following algorithm shows the pseudo code of Function **F: LU_EUCDIST**. According to the algorithm, line 1 refers to the computation based on the ArcGIS EUCDISTANCE function. Lines 2 and 3 are used to find the maximum and minimum distance value of the layer derived from line 1. This is achieved by using the ArcObjects interface, namely the IRasterStatistics interface. This interface allows access to raster statistics. Minimum and Maximum are properties of this interface, enabling the approximate largest and smallest value of the specified layer to be observed. Line 4 is computed using the ArcGIS arithmetic operation in order to normalize the distance using the reciprocal procedure (Equation 4.15). The result contains the normalized values ranging from 0.0 to 1.0. A value close to 1.0 means being close to the considered land use. It should be noted that the function is designed to run each time to create each development factor (e.g. using residential layer as input to create proximity to residential use factors).

Algorithm F: LU_EUCDIST

Input:

gLayer	# A grid layer, where it can be a residential layer, a commercial layer, an industrial layer, # a government layer, a school layer, an agriculture layer. Each is derived from function # **F:LU_EXTRACT** using the SELECT function.

Output:

gLuNrmEuc	# A proximity to a land use grid layer. A layer generated contains continuous (floating) # value ranging from 0.0 to 1.0.

Analysis environment settings:

gMask	# The spatial mask grid layer for the EUCDISTANCE function that specifies the areas to # be processed.

LU_EUCDIST(*gLayer*)

1. *gEuc* = **EUCDISTANCE(*gLayer*)**	# ArcGIS command: the EUCDISTANCE function.
2. *Max* = **Maximum(*gEuc*)**	# ArcObjects syntax: property of IRasterStatistics # interface.
3. *Min* = **Minimum(*Euc*)**	# ArcObjects syntax: property of IRasterStatistics # interface.
4. $gLuNrmEuc = 1 - ((gEuc - Min) / (Max - Min))$	# ArcGIS command: the arithmetic operator; # subtraction (-), and division (/).

Function F: RD_EUCDIST (Euclidean distance for road data). This function is employed to create two development factors regarding proximity to all roads and proximity to planned roads if a planned road shape file is input from the GUI. The function mainly uses the ArcGIS EUCDISTANCE function, whose concept is similar to function **F:LU_EUCDIST** (see Figure 5.9).

The following algorithm shows the pseudo code of Function **F: RD_EUCDIST**. The algorithm of this function is similar to the previous function **(F: LU_EUCDIST)**. The only difference is that the input to the function is a shape file. In addition, the cell size needs to be set for the analysis in order to specify the cell size of output generated. The outcome grid layer generated has a value ranging from 0.0 to 1.0. A value close to 1.0 means being close to roads. It should be noted that the function is designed to run each time to create each development factors (e.g. using a road shape file as input to create proximity to all roads factor).

Algorithm F: RD_EUCDIST

Input:

sLayer # A road layer in a shape file format, where it can be a road or a planned road file.

Output:

gRdNrmEuc # A proximity to road grid layer. A layer generated contains continuous (floating) value
ranging from 0.0 to 1.0.

Analysis environment settings:

gMask # The spatial mask grid layer for the EUCDISTANCE function that specifies the areas to
be processed.

gCsz # The cell size layer for the EUCDISTANCE function that specifies the cell size of the
output layer.

RD_EUCDIST(*sLayer*)

1. *gEuc* = **EUCDISTANCE(*sLayer*)** # ArcGIS command: the EUCDISTANCE function.
2. *Max* = **Maximum(*gEuc*)** # ArcObjects syntax: property of IRasterStatistics
interface.
3. *Min* = **Minimum(*gEuc*)** # ArcObjects syntax: property of IRasterStatistics
interface.
4. *gRdNrmEuc* = 1 – ((*gEuc* – *Min*) / (*Max* – *Min*)) # ArcGIS command: the arithmetic operator;
subtraction (-), and division (/).

Function F: RDTYPE_EUCDIST (Euclidean distance for road type). This function is employed to create three development factors regarding proximity to main roads, collector streets, and local streets. This function is similar to the previous function **(F:RD_EUCDIST)** in that it mainly employs the ArcGIS EUCDISTANCE function. However since each road type shape file extracted from this function **(F:RDTYPE_EXTRACT)** can have a different maximum and minimum distance value, they all are mutually compared to find the smallest minimum and largest maximum value that will be used as standard values for normalizing all road type data.

The following algorithm shows the pseudo code of function **F: RDTYPE_EUCDIST.** According to the algorithm, lines 1 – 3 refer to the computation based on the ArcGIS EUCDISTANCE function while lines 4 – 9 are used to find the maximum and minimum distance value of the layers derived from lines 1 – 3. Lines 10 – 11 are used to find and keep the smallest minimum and the largest maximum value from all input layers based on the customized VBA methods, **Maximize** and **Minimize.** Lines 12 – 14 are computed in order to normalize the distance measured to contain values ranging from 0.0 to 1.0. A value close to 1.0 means being close to roads.

Algorithm F: RDTYPE_EUCDIST

Input:

sRdT1	# A main road shape file derived from function **F:RDTYPE_EXTRACT**
sRdT2	# A collector road shape file derived from function **F:RDTYPE_EXTRACT**
sRdT3	# A street shape file derived from function **F:RDTYPE_EXTRACT**

Output:

gRdNrmEucRdT1	# A proximity to main roads grid layer
gRdNrmEucRdT2	# A proximity to collector roads grid layer,
gRdNrmEucRdT3	# A proximity to streets grid layer # Remark: All layers generated contain continuous (floating) value ranging from 0.0 # to 1.0.

Analysis environment settings:

gMask	# The spatial mask grid layer for the EUCDISTANCE function that specifies the # areas to be processed.
gCsz	# The cell size layer for the EUCDISTANCE function that specifies the cell size of # the output layer.

RDTYPE_EUCDIST(*sRdT1, sRdT2, sRdT3*)

1. *gEucRdT1* = **EUCDISTANCE(*sRdt1*)**	# ArcGIS command: the EUCDISTANCE # function.
2. *gEucRdT2* = **EUCDISTANCE(*sRdt2*)**	
3. *gEucRdT3* = **EUCDISTANCE(*sRdt3*)**	
4. *Max1* = **Maximum(*gEucRdT1*)**	# ArcObjects syntax: property of # IRasterStatistics interface.

5. *Max2* = **Maximum(***gEucRdT2***)**

6. *Max3* = **Maximum(***gEucRdT3***)**

7. *Min1* = **Minimum(***gEucRdT1***)** # ArcObjects syntax: property of # IRasterStatistics interface.

8. *Min2* = **Minimum(***gEucRdT2***)**

9. *Min3* = **Minimum(***gEucRdT3***)**

10. *Max* = **Maximize(***Max1,Max2,Max3***)** # Customized VBA method. See function # **F:Maximize**

11. *Min* = **Minimize(***Min1,Min2,Min3***)** # Customized VBA method. See function # **F:Minimize**

12. $gRdNrmEuc = 1 - ((gEucRdT1 - Min) / (Max - Min))$ # ArcGIS command: the arithmetic # operator; subtraction (-), and division (/).

13. $gRdNrmEuc = 1 - ((gEucRdT2 - Min) / (Max - Min))$

14. $gRdNrmEuc = 1 - ((gEucRdT3 - Min) / (Max - Min))$

Algorithm F: Maximize

Input:

Max1	# A value.
Max2	# A value.
Max3	# A value.

Output:

Max	# An output value.

Maximize(*Max1,Max2,Max3***)**

1. **If** *Max1* ≥ *Max2* **Then** *Max* = *Max1* # compare *Max1* and *Max2*, put the largest value to *Max*.

2. **Else** *Max* = *Max2*

3. **End If**

compare *Max* and *Max3*, put the largest value to *Max*.

4. **If** *Max* ≥ *Max3* **Then** *Max* = *Max*

5. **Else** *Max* = *Max3*

6. **End If**

7. Return *Max* # return *Max* value to the calling method.

Algorithm F: Minimize

Input:

Min1 # A value.
Min2 # A value.
Min3 # A value.

Output:

Min # An output value.

```
Minimize(Min1,Min2,Min3)
1. If Min1 ≤ Min2 Then  Min = Min1        # compare Min1 and Min2, put the smallest value to Min.
2. Else Min = Min2
3. End If
4. If Min ≤ Min3 Then  Min = Min          # compare Min and Min3, put the smallest value to Min.
5. Else Min = Min3
6. End If
7. Return Min                             # return Min value to the calling method.
```

Function F: VAR_CMB (Combination of development factors). This function is built based on the ArcGIS COMBINE function to combine the development factors used in the analysis into one grid layer with an associated attribute table. An output grid layer is stored in the output workspace (directory), and given the name CMBVAR for being used for MNL's derivation of criteria weights.

The following algorithm shows the pseudo code of function **F: VAR_CMB** which is used to combine all the development factors. It should be noted that in lines 1 – 11, all input development factor grid layers derived from function **F: LU_EUCDIST, F: RD_EUCDIST,** and **F: RDTYPE_EUCDIST** are multiplied by the constant value 10,000 in order to avoid truncating the floating value when converting these values to be used in an attribute table, as described at the end of section 5.3.1.1. Line 12 shows the GIS COMBINE function used to combine all development factors measured. The outcome of this function comprises all development factors as shown in Table 5.1. It should be noted that the line continuation character (an underscore) in line 12 is used in this pseudo code and in the remainder of the pseudo code in this chapter due to the space limitations of the page format.

Algorithm F: VAR_CMB

Input:

gLu1993	# A 1993 land use layer
gLu2001	# A 2001 land use layer
gV1	# A residential neighborhood layer, derived from function **F: NR__MEASURE.**
gV2	# A commercial neighborhood layer , derived from function **F: NR__MEASURE.**
gV3	# An industrial neighborhood layer , derived from function **F: NR__MEASURE.**
gV4	# A proximity to all roads layer, derived from function **F:RD_EUCDIST.**
gV5	# A proximity to main roads layer, derived from function **F:RDTYPE_EUCDIST.**
gV6	# A proximity to collector roads , derived from function **F:RDTYPE_EUCDIST.**
gV7	# A proximity to streets layer, derived from function **F:RDTYPE_EUCDIST.**
gV8	# A proximity to residential layer, derived from function **F:LU_EUCDIST.**
gV9	# A proximity to commercial layer, derived from function **F:LU_EUCDIST.**
gV10	# A proximity to industrial layer, derived from function **F:LU_EUCDIST.**
gV11	# A proximity to government, derived from function **F:LU_EUCDIST.**
gV12	# A proximity to school layer, derived from function **F:LU_EUCDIST.**
gV13	# A proximity to park layer, derived from function **F:LU_EUCDIST.**
gV14	# A land price layer.
gV15	# A proximity to planned roads layer, derived from function **F:RD_EUCDIST.**
gV16	# A proximity to agricultural layer, derived from function **F:LU_EUCDIST.**

Output:

gCmbVar	# An output grid layer with an associated table.

Analysis environment settings:

gMask	# The spatial mask grid layer for the COMBINE function that specifies the areas to # be processed.

VAR_CMB(*gLu1993, gLu2001, gV1, gV2, gV3, gV4, gV5, gV6, gV7, gV8, gV9, gV10, gV11, gV12, gV13, gV14, gV15, gV16*)

1. *gIV4* = *gV4* * 10000 # ArcGIS command: the arithmetic operator; multiplication (*).

2.. *gIV5* = *gV5* * 10000

3. *gIV6* = *gV6* * 10000

4. *gIV7*= *gV7* * 10000

5. *gIV8* = *gV8* * 10000

6. *gIV9*= *gV9** 10000

7. *gIV10* = *gV10* * 10000

8. *gIV11* = *gV11* * 10000

9. *gIV12* = *gV12* * 10000

10. *gIV13* = *gV131* * 10000

11. *gIV15* = *gV15* * 10000

12. *gIV16* = *gV16* * 10000

13. *gCmbvar* = **COMBINE**(*gLu1993, gLu2001, gV1, gV2, gV3, gIV4, gIV5,* _
gIV6, gIV7, gIV8, gIV9, gIV10, gIV11, gIV12, gIV13, gV14, gIV15, gIV16) # ArcGIS command: the COMBINE function.
Remark: an underscore

(_) shows a line
continuation

5.3.2 MNL (GIS-based CA Approach) Tool

5.3.2.1 Graphical User Interface (GUI) of MNL (GIS-based CA Approach) Tool

The MNL (GIS-based CA approach) tool enables land use change simulation using the MNL-based transition rule described in Equation 4.1. Figure 5.11 shows the graphical user interface (GUI) of the MNL (GIS-based CA approach) tool. The interface is divided into three sections. The first section is for layer and parameter input. The GUI starts by allowing users to set the initial year and the end year for simulation, then to set the threshold in terms of the total number of cells to change from vacant to target land use types; residential, commercial and industrial respectively. In the next section, the GUI allows users to input the layers required for simulation. Mandatory input layers are the initial land use map (grid layer) and road feature layer (shape file format). Optional layers are planned road (shape file format) and land price (grid layer).

The next section refers to update ranking. During simulation the joint probability maps for residential, commercial, and industrial areas along with the associated tables are created, the probability scores of these associated raster tables are sorted and chosen based on the highest probability. Nevertheless, when the three land use types based on the chosen score are being updated, the three attribute raster tables considered cannot be compared at the same time. This is due to the technical limitations of the raster handling of the software. This section allows users to specify the order of category update. See more details about the technical aspect of update ranking procedure in function **F:LUCHG_UPDATE**, Section 5.3.2.2. From the main MNL GUI (Figure 5.11), the default setting in the combo (combination) box option is 'User-defined RS-CM-MA', which refers to the order of category update, beginning with residential (RS), commercial (CM) and industrial (MA) respectively. The other choices that users can make from the combo box include RS-MA-CM, CM-RS-MA, CM-MA-RS, MA-RS-CM, and MA-CM-RS.

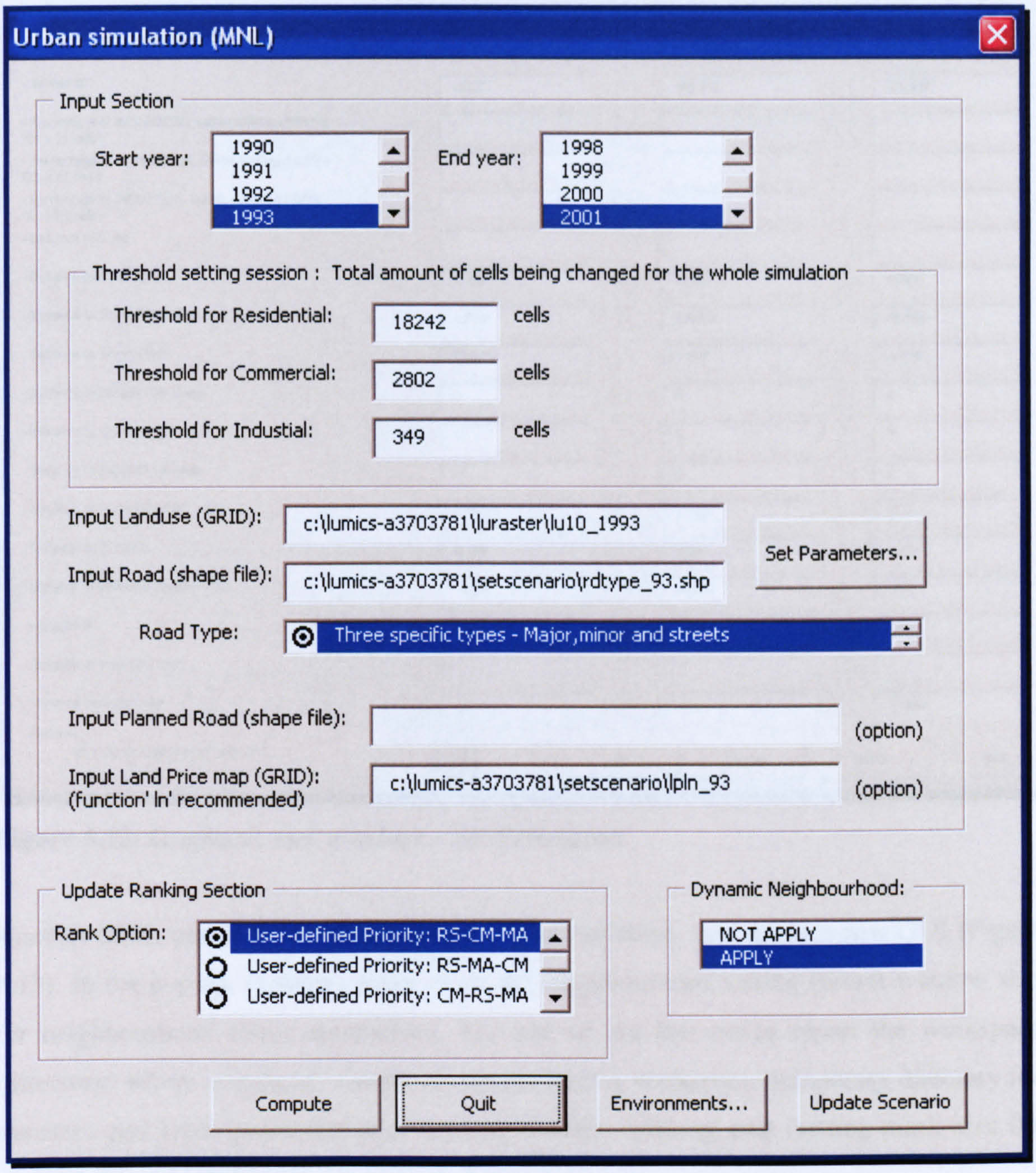


Figure 5.11: Graphical user interface: MNL (GIS-based CA approach) pop-up menu.

In addition to the main GUI described above, there are numbers of options that need to set before the simulation execution. The first button refers to 'set parameters...'. When clicking this button, a new GUI will pop up (Figure 5.12). This allows users to type in criterion weights, which are derived beforehand from the MNL parameter estimation (Section 4.5.1). The button 'LOAD' is used to load a frequently used set of weights by the user.

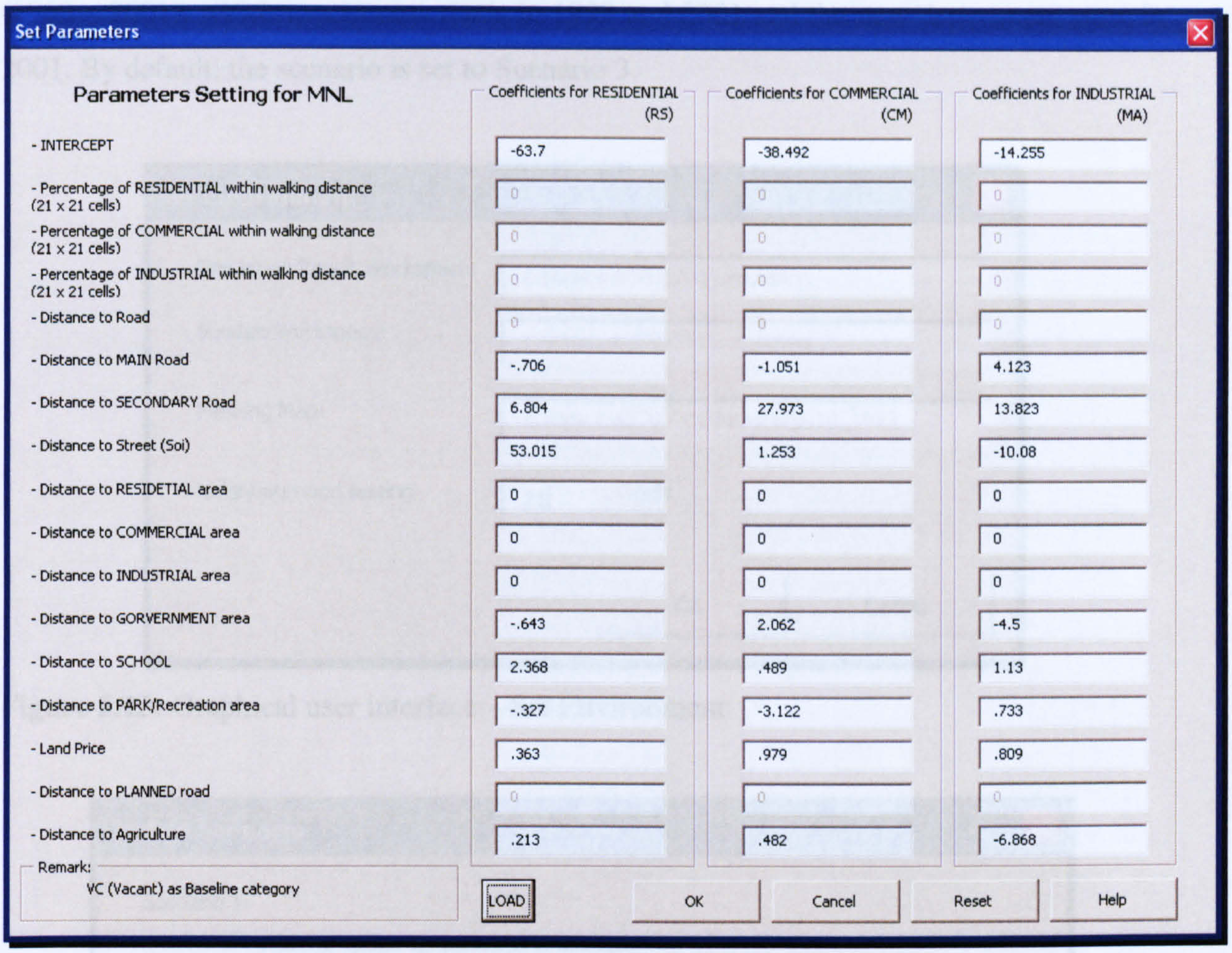


Figure 5.12: Graphical user interface – Set Parameters.

Another MNL option is 'Environments…'. When selected, this opens a new GUI (Figure 5.13). In the pop-up, it allows users to set the neighbourhood setting (preset window size for neighbourhood effect simulation). The rest of the text boxes report the workspace (directory) where simulation results are stored, scratch workspace (temporary directory for transient grid layer generated when running model), masking map (setting mask area for analysis during simulation).

The 'Update Scenario' option allows users to choose three scenarios regarding the method of update during simulation as shown in Figure 5.14. The first scenario (Scenario 1) refers to not updating all events (in this study, this means new roads in 1998 and 2001, and a new department store in 2001) during the specified period. The second scenario (Scenario 2) refers to updating some events (in this study, this means updating new roads in 1998 and 2001, but not updating new department store in 2001) during the specified period. The third scenario (Scenario 3) is to update all real events that occurred during the study period

(1993 – 2001), which means new roads in 1998 and 2001 and the new department store in 2001. By default, the scenario is set to Scenario 3.

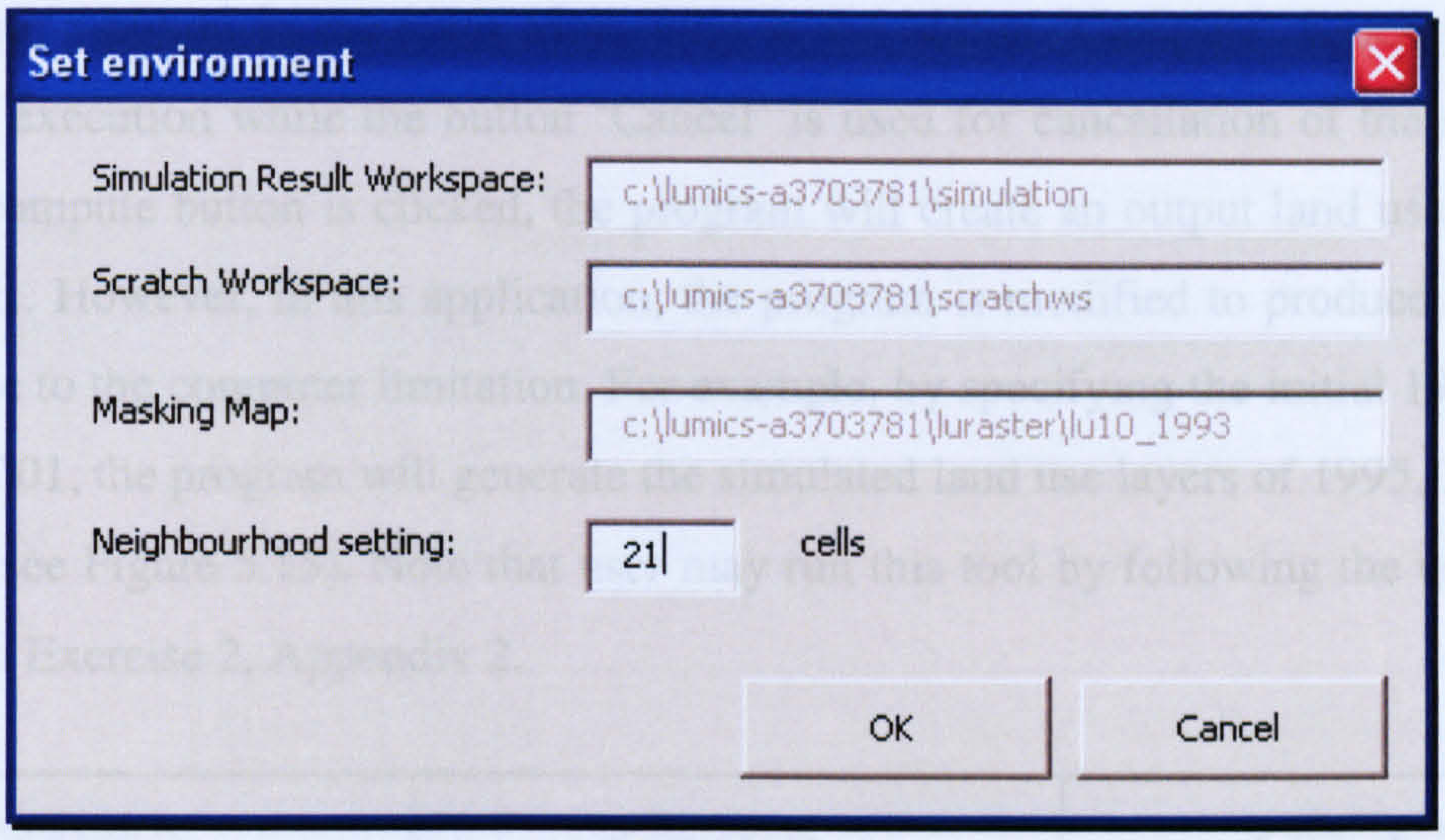


Figure 5.13: Graphical user interface – Set Environment.

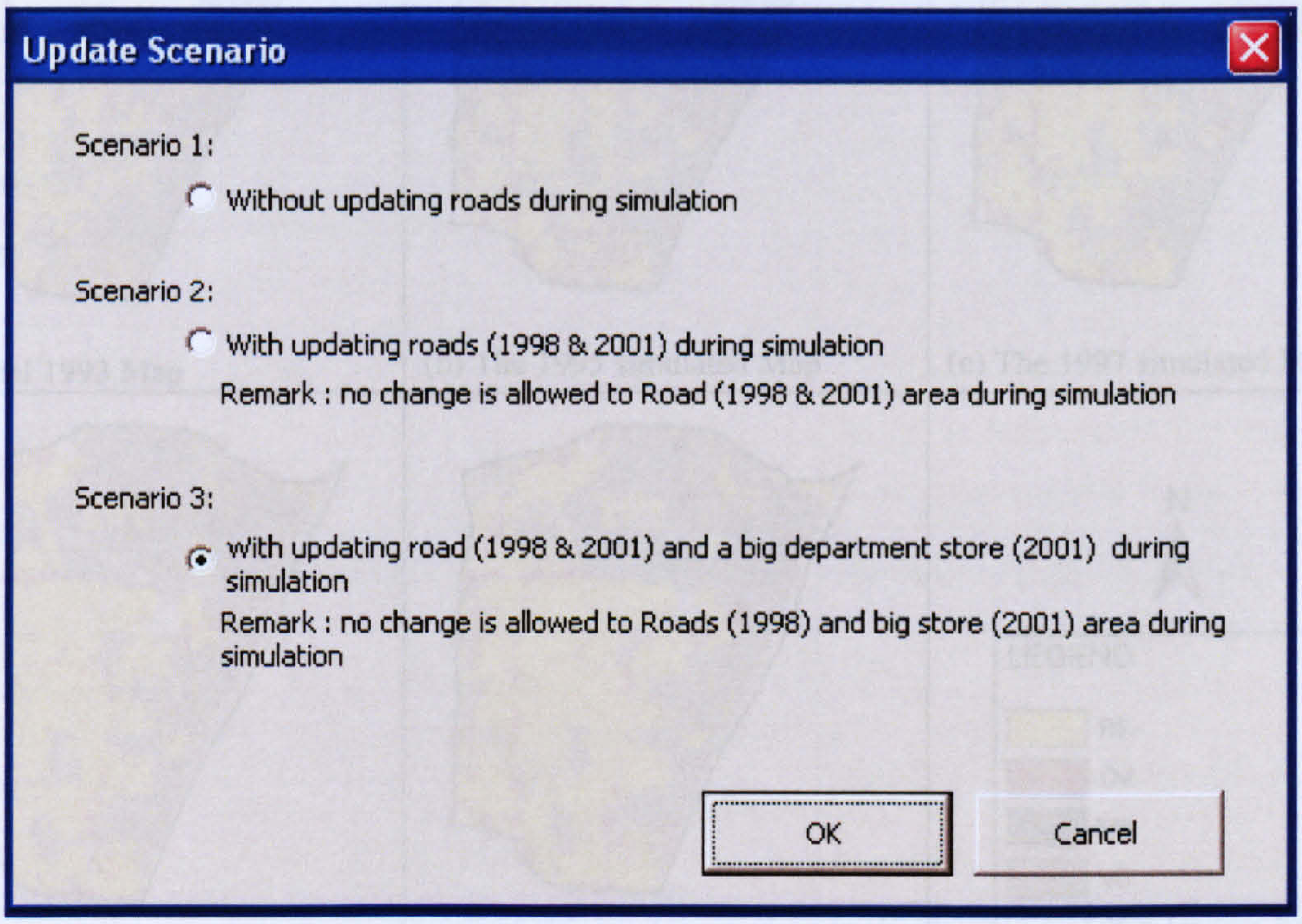


Figure 5.14: Graphical user interface – Update Scenario.

The last section refers to dynamic neighbourhood. This section allows users to define whether the dynamic or iterative effect of neighbourhood, which means residential neighbourhood, commercial neighbourhood, and industrial neighbourhood, are applied or not. According to the CA approach, users need to select 'APPLY' from the combo box, in order to generate dynamic neighbourhood according to the proposed MNL-based CA transition rules in Section 4.4.1.1. Another option is 'NOT APPLY'. This is an alternative for testing the effect of not including the neighbourhood effect for the simulation.

The remaining button in the main MNL GUI is the button 'Compute' which is used for simulation execution while the button 'Cancel' is used for cancellation of the simulation. Once the compute button is clicked, the program will create an output land use layer on a yearly basis. However, in this application, the program is modified to produce a two-year interval due to the computer limitation. For example, by specifying the initial 1993 and the end year 2001, the program will generate the simulated land use layers of 1995, 1997, 1999 and 2001 (see Figure 5.15). Note that user may run this tool by following the user's guide provided in Exercise 2, Appendix 2.

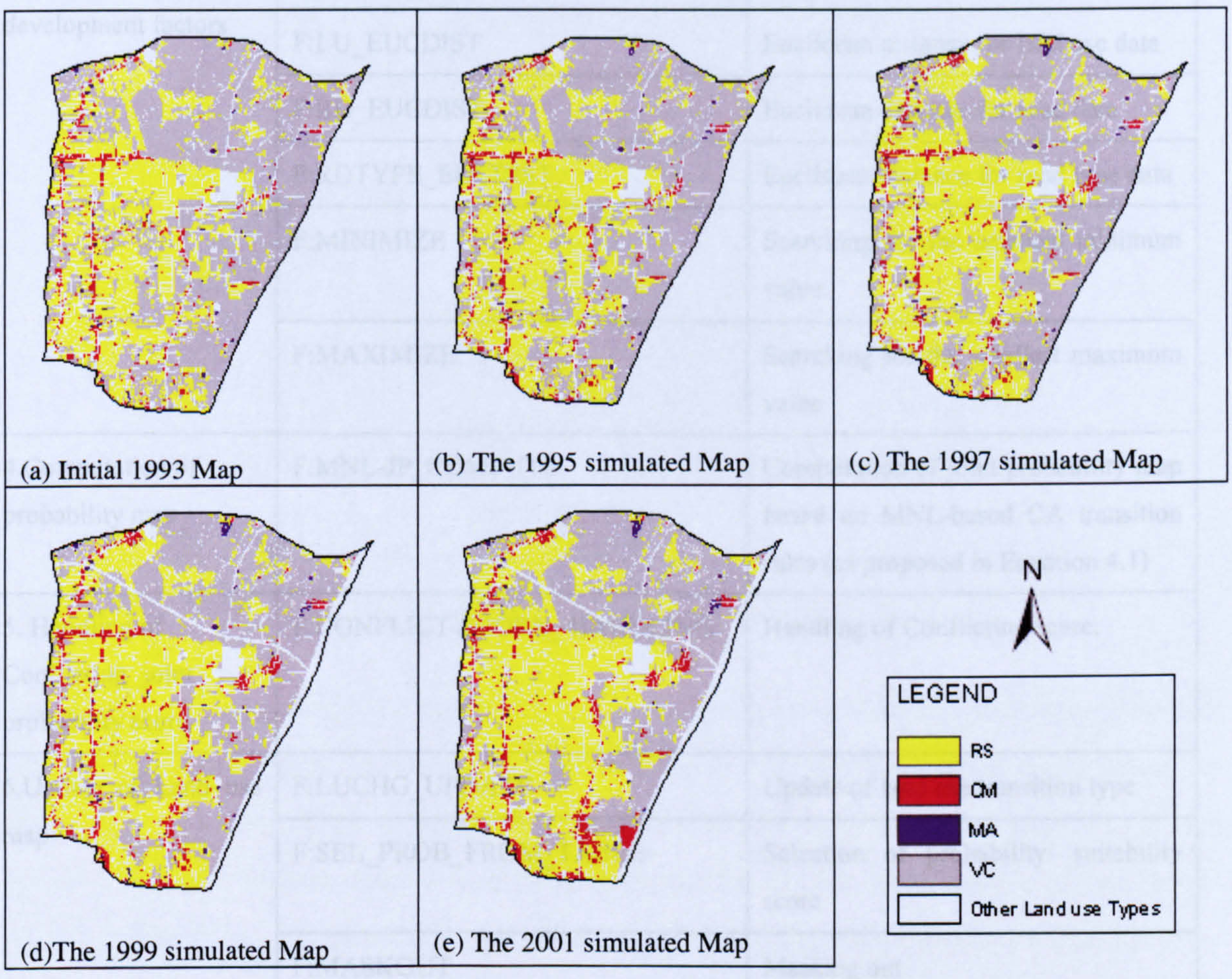


Figure 5.15: An example of simulated results from initial 1993 to 2001. The real simulated results comprise 10 land use classes (10 colours). For the sake of visual simplicity, four classes are represented here in which yellow shows residential, red shows commercial, purple shows industrial, grey shows vacant area and white shows other land use types.

5.3.2.2 Execution of MNL (GIS-based CA approach) Tool

After receiving the parameters from the GUI, the tool performs seven main tasks as listed in Table 5.3, which can be split into 17 functions (see Figure 5.16)

Task Description	Function	Action
1. Extraction of data	FLUTYPE_EXTRACT	Extraction of land use type
	F:RDTYPE_EXTRACT	Extraction of road type
2.Elimination of Constraint area	F:CONSTR_ELIMINATION	Elimination of constraint area
3.Creation of development factors	F:NR_MEASURE	Neighborhood (focal) measurement
	F:LU_EUCDIST	Euclidean distance for land use data
	F:RD_EUCDIST	Euclidean distance for road data
	F:RDTYPE_EUCDIST	Euclidean distance for road type data
	F:MINIMIZE	Searching for the smallest minimum value
	F:MAXIMIZE	Searching for the smallest maximum value
4.Computation of probability map	F:MNL-JP_COMPUTE	Computation of joint probability map based on MNL-based CA transition rules (as proposed in Equation 4.1)
5. Handling of Conflicting joint probability score	F:CONFLICT-SCORE_ HANDLING	Handling of Conflicting score.
6.Update of Land use map	F:LUCHG_UPDATE	Update of land use transition type
	F:SEL_PROB_FROM_TABLE	Selection of probability/ suitability score
	F:MASKOUT	Masking out
	F:LU_UPDATE	Update of land use change
	F:SCENARIO_UPDATE	Updating scenario
7. Iteration control	F:ITERATION	Iteration

Table 5.3: List of tasks and functions for MNL (GIS-based CA approach) tool.

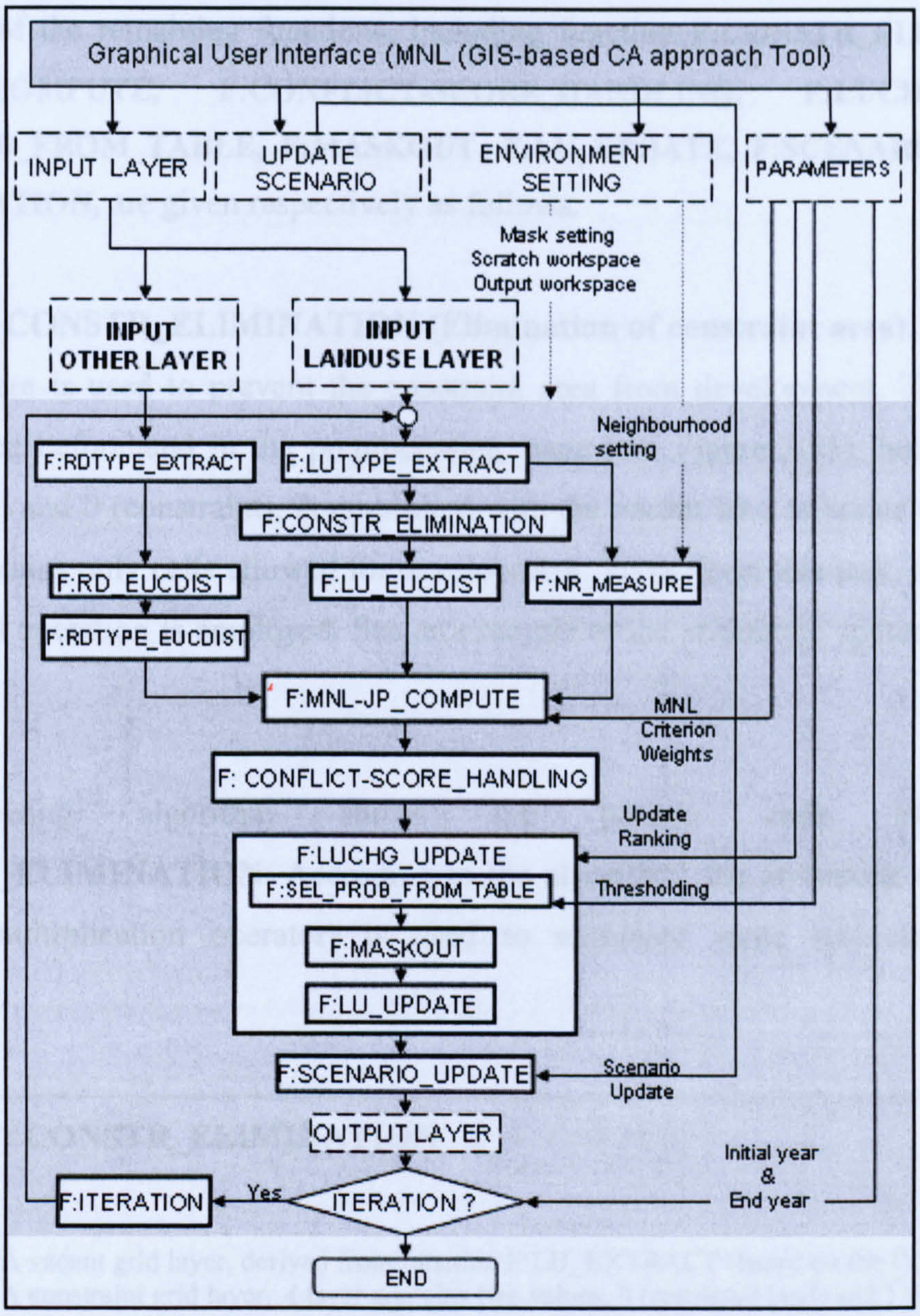


Figure 5.16: Model developed for MNL (GIS-based CA approach) Tool.

The program begins with the extraction of data used for the analysis including land use types and road types and creation of a constraint map. The second step is to eliminate the area that is constrained for development by the constraint map. The third step is to create the development factors using two GIS techniques: Focal (e.g. the creation of residential neighbourhood index) function and Euclidean distance (e.g. the creation of proximity to main roads). The fourth step is the computation of the probability map, here based on the developed MNL-based CA model as described in Section 4.4.1.1. The fifth step is to update the land use map based on the probability derived in the previous stage. In this step, it involves ranking probability scores, applying land use demand and map update. The final step concerns the iteration control, which is used to drive the simulation.

Description of the remaining functions, including function **F:CONSTR_ELIMINATION, F:MNL-JP_COMPUTE, F:CONFLICT-SCORE_HANDLING, F:LUCHG_UPDATE, F:SEL_PROB_FROM_TABLE, F:MASKOUT, F:LU_UPDATE, F:SCENARIO_UPDATE,** and **F:ITERATION,** are given respectively as follows.

Function F: CONSTR_ELIMINATION (Elimination of constraint area). The function developed here is used to prevent the constraint area from development. The constraint map produced beforehand in the preprocessing stage (see Figure 3.31), holding value 1 (developable) and 0 (constraint), is multiplied with the vacant land to create a new vacant land that contains only cells allowed for development. To perform this task, the arithmetic mathematical operation is employed. See an example of the arithmetic operation in Figure 5.9.

The following algorithm shows the pseudo code of function **F:CONSTR_ELIMINATION.** According to the algorithm, the arithmetic mathematical operation (multiplication operator) is used to eliminate some restricted areas for development.

Algorithm F: CONSTR_ELIMINATION

Input:

gVc	# A vacant grid layer, derived from function F:LU_EXTRACT (based on the CON function)
gConstr	# A constraint grid layer. *A* layer contains two values, 0 (restricted land) and 1 (developable # land).

Output:

gVcFinal	# A developable vacant grid layer. *A* layer generated contains two values, 0 (non-vacant) and 1 (vacant).

Analysis environment setting:

gMask	# The spatial mask grid layer for the arithmetic operator function that specifies the areas to be # processed.

CONSTR_ELIMINATION (*gVc,gConstr*)

1. *gVcFinal* = *gVc* * *gConstr*   # ArcGIS command: the arithmetic operator; multiplication (*).

Function F: MNL-JP_COMPUTE (Computation of joint probability map based on MNL-based CA transition rules). This function is built based on the MNL-based CA transition rule in the developed model as expressed in Equation 4.9. Within this function, it performs two tasks: the computation of global probability maps based on the MNL method and the computation of joint probability maps. The outcomes of the function, the joint probability maps of the three land use transition types, regarding vacant change to residential, commercial, industrial, will be used for further analysis. According to the function, the arithmetic mathematical operation and the exponential EXP function play a major role for computation.

The exponential EXP function is one function of ArcGIS Map algebra. Figure 5.17 shows the logical expression of the exponential EXP function employed that is used to compute the base e exponential of the input raster or number on a cell-by-cell basis.

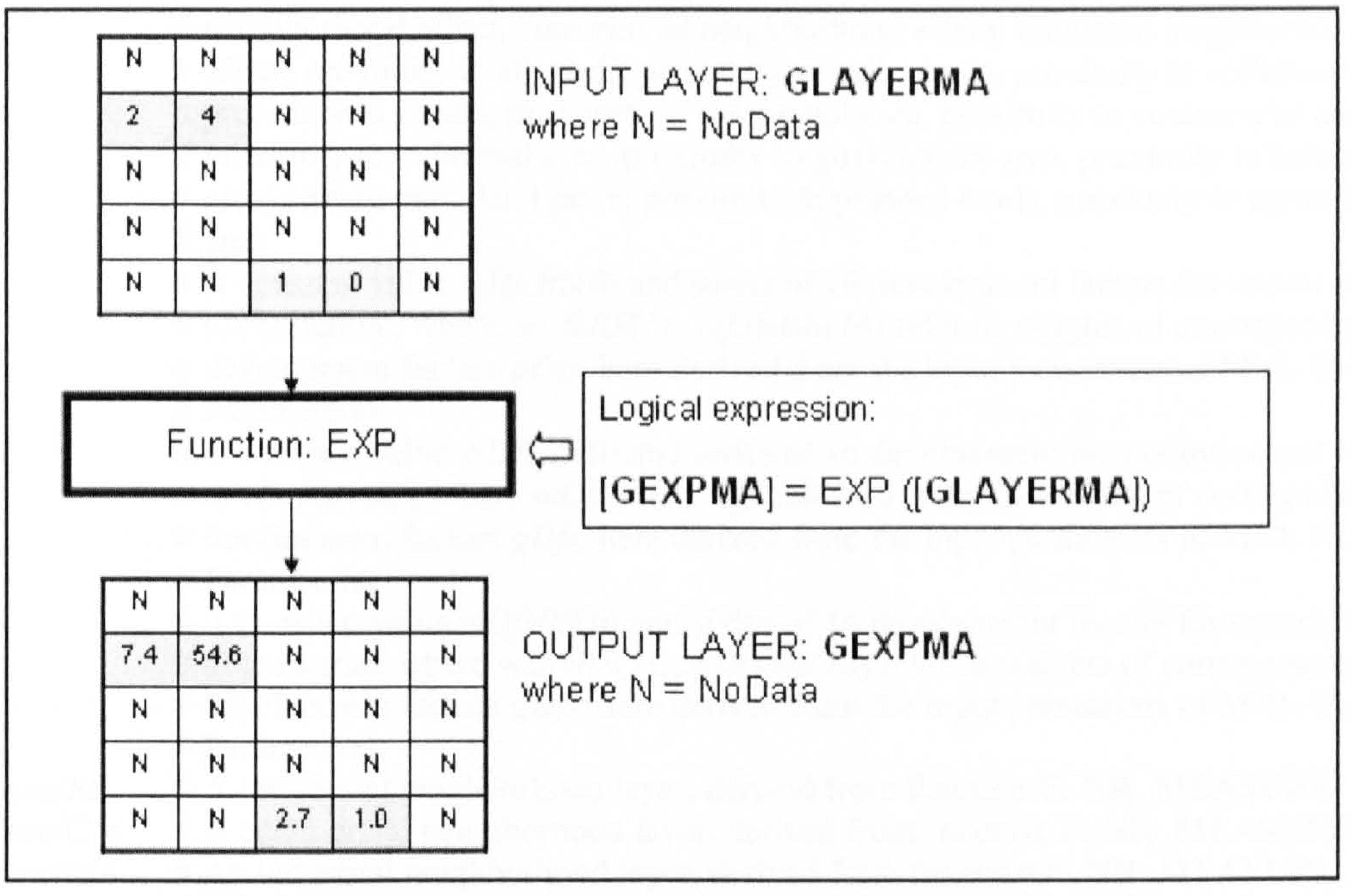


Figure 5.17: An example of the exponential EXP function used to create an output (GEXPMA).

The following algorithm shows the pseudo code of function **F:MNL-JP_COMPUTE**. According to the algorithm, lines 1 – 18 refer to the computation based on the MNL method to produce global probability as described in Equation 4.1 using the arithmetic operator (lines 3, 7 and 11) and the exponential EXP (lines 13 – 15). Based on the input development factors, derived from function **F:NR_MEASURE, F:LU_EUCDIST, F:RDTYPE_EUCDIST,** and a set of weights or coefficients of each land use type, derived from the MNL GUI - Set Parameters (Figure 5.12), it results in the creation of global probability maps for four land use transition types regarding vacant change to residential, commercial, industrial and vacant (no change). Lines 16 – 18 refers to the creation of three joint probability layers of vacant change to residential, commercial, and industrial. Each is the combination of probability layers previously created, the corresponding neighbourhood layer derived from the earlier function (**F:NR_MEASURE**) and the vacant area derived from the constraint map according to the function **F:CONSTR_ELIMINATION**. It should be noted that the neighbourhood layers, as a part of command syntax in lines 19 – 21, are divided by 100 in order to convert them into neighbourhood index layers, having values ranging from 0.0 to 1.0.

Algorithm F: MNL-JP_COMPUTE

Input:

gDfs	# Series of 16 development factors, where *gDfs*(*1*)...gDfs(*16*) refer to residential # neighborhood effect, commercial neighborhood effect, industrial neighborhood # effect, proximity to all roads, proximity to main roads, proximity to collector roads, # proximity to streets, proximity to residential area, proximity to commercial area, # proximity to industrial area, proximity to government area, proximity to school, # proximity to park, land price, proximity to planned roads, proximity to agriculture # area.
wDfRS	# A constant value w*DfsRS*(*0*) and series of 16 development factors for vacant change # to residential, where w*DfsRS*(*1*)...gDfsRS(*16*) refer to weights of corresponding 16 # development factors *gDfs*, here derived from the input parameters of MNL GUI (Set # Parameters).
wDfCM	# A constant value w*DfsRS*(*0*) and series of 16 development factors for vacant change # to commercial, where w*DfsRS*(*1*)...gDfsRS(*16*) refer to weights of corresponding 16 # development factors *gDfs*, here derived from the input parameters of MNL GUI (Set # Parameters).
wDfMA	# A constant value w*DfsRS*(*0*) and series of 16 development factors for vacant change # to industrial, where w*DfsRS*(*1*)...gDfsRS(*16*) refer to weights of corresponding 16 # development factors *gDfs*, here derived from the input parameters of MNL GUI (Set # Parameters).
gNiPctFocalRS	# A residential neighborhood layer, derived from function **F: NR_MEASURE.**
gNiPctFocalCM	# A commercial neighborhood layer, derived from function **F: NR_MEASURE.**
gNiPctFocalMA	# An industrial neighborhood layer, derived from function **F: NR_MEASURE.**
gExpVC	# A layer that all cells contain the value of 1.

Output:

gJPRS	# A joint probability grid layer of vacant change to residential.
gJPCM	# A joint probability grid layer of vacant change to commercial.
gJPMA	# A joint probability grid layer of vacant change to industrial. # Remark: Each layer generated contains (floating) probability values ranging from # 0.0 to 1.0.

Analysis environment settings:

gMask	# The spatial mask layer for the analysis function that specifies the areas to be # processed, here derived from a vacant grid layer that is created from function # F: LUTYPE_EXTRACT (based on the CON function).

F: MNL-JP_COMPUTE(*gDfs, wDfRS, wDfCM, wDfMA gNiPctFocalRS, gNiPctFocalCM, gNiPctFocalMA, gExpVC*)

1. *gLayerRS* = *wDfRS*(0)	# Initialize a residential factors # grid layer to a constant.
2. **For** i = 1 to 16	# For-loop to accumulate a #residential factors grid layer
3. *gLayerRS* = *gLayerRS* + (*gDfs*(*i*) * *wDfRS*(*i*))	# ArcGIS command: the # arithmetic operator function.

```
EndFor

gLayerCM = wDfCM(0)                                   # Initialize a commercial factors
                                                      # grid layer to a constant.
For i = 1 to 16                                       # For-loop to accumulate a
                                                      # commercial factors grid layer
gLayerCM= gLayerCM + ( gDfs(i) * wDfCM(i) )           # ArcGIS command: the
                                                      # arithmetic operator function.
EndFor

gLayerMA = wDfMA(0)                                   # Initialize an industrial factors
                                                      # grid layer to a constant.
For i = 1 to 16                                       # For-loop to accumulate an
                                                      # industrial factors grid layer
gLayerMA= gLayerMA + ( gDfs(i) * wDfMA(i) )           # ArcGIS command: the
                                                      # arithmetic operator function.
EndFor
                                                      # ArcGIS command: the
gExpRS = EXP(gLayerRS)                                exponential EXP function.

gExpCM = EXP(gLayerCM)

gExpMA = EXP(gLayerMA)

gProbRS = gExpRS / (gExpRS + gExpCM + gExpMA + gExpVC)   # ArcGIS command: the
                                                         # arithmetic operator function to
gProbCM = gExpCM / (gExpRS + gExpCM + gExpMA + gExpVC)   # calculate cell probability for
                                                         #residential, industrial and
gProbMA = gExpMA / (gExpRS + gExpCM + gExpMA + gExpVC)   #commercial respectively.

gJPRS = gProbRS * ( gNiPctFocalRS / 100 )             # ArcGIS command: the
                                                      # arithmetic operator function to
gJPCM = gProbCM * ( gNiPctFocalCM / 100 )             # calculate cell probability for
                                                      #residential, industrial and
gJPMA = gProbMA * ( gNiPctFocalMA / 100 )             #commercial respectively.
```

As noted that in Section 4.4.1.1, although four types of global development probability of land transition are created, in this study focus is on three land use transition types. Thus the outcome of the analysis needs only three types of joint probability of land development transition regarding residential, commercial and industrial use.

Function F:CONFLICT-SCORE_HANDLING (Handling of conflicting joint probability score). This function is built to handle the cases in which the joint probability of residential, industrial and commercial at the same tile have equal scores. As described in Section 4.4.2, the solution to handling this problem is that the vacant cell in question will change to the targeted land use type that has the highest neighbourhood index value. To accomplish this function, four main ArcGIS functions are employed. They are the ArcGIS conditional CON function, the ArcGIS UPOS function, the ArcGIS OVER function and, the arithmetic operator. The logical expression of the conditional CON and the arithmetic operator is previously shown in Figure 5.6 and 5.9 respectively.

The ArcGIS UPOS function assigns the position of the output raster with the maximum value in a set of input rasters on a cell-by-cell basis (ESRI, 2004). An example in Figure 5.18 shows how the UPOS function manages the position of the resultant land use map with respect to the order of map layers. It should be noted that if there is NoData at a location on any of the input rasters, the output at that location will be NoData. In addition, if two or more input raster bands contain equal maximum values for a particular cell location, the position of the first layer is returned on the output raster.

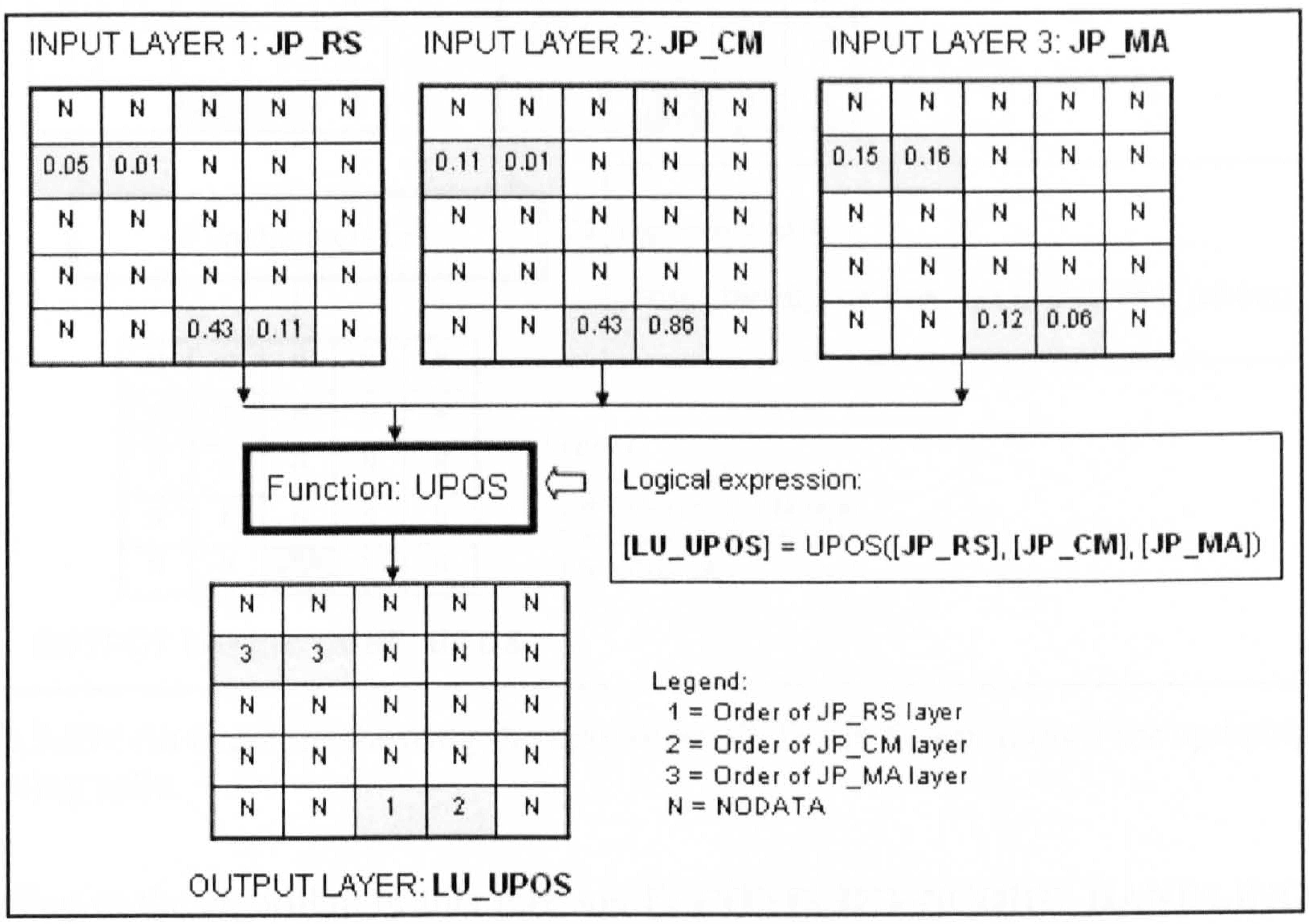


Figure 5.18: An example of the UPOS function used to create an output grid layer (LU_UPOS).

Another ArcGIS function employed is the OVER function. The OVER function is one of the map algebra (cell-by-cell analysis) functions. Its capability is mainly used to update the land use map in this study. In Figure 5.18, a particular cell location at row 5 column 3 of the first input layer JP_RS and the second input layer JP_CM has the same score (0.43). Based on the rule of the UPOS function, it will return 1 to the output layer. However, in this study the criterion of neighbourhood index value is applied for handling the conflicting cells. Suppose that, based on the criterion, the results show that that vacant cell should change to layer JP_CM. For such a case, the OVER function is performed to force updating of the cell. Figure 5.19 shows the example of the ArcGIS OVER function which is used to update the previous layer (LU_UPOS) in Figure 5.18 from a conflicting cell layer (CONF_CELL) to create a new order layer (CONF_UPOS). In the output layer (CONF_UPOS), as shown in Figure 5.19, the order of input is important in the OVER logical expression in that those values from the first input (in this case, CONF_CELL) that are nonzero will be first put into the output cells, otherwise the value from the second layer (LU_UPOS) will be assigned.

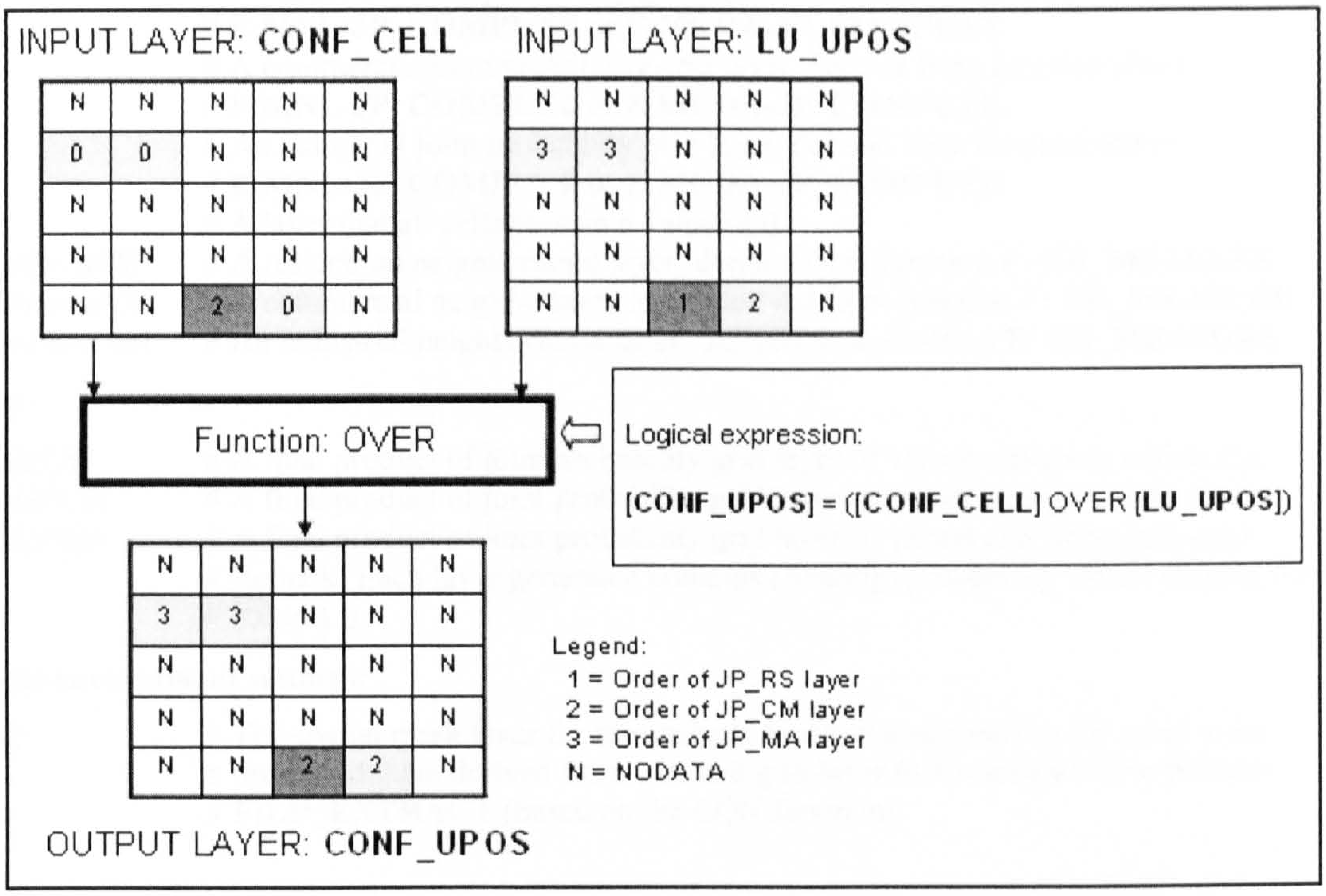


Figure 5.19: An example showing the operation of OVER function used for updating the conflicting cells.

According to the algorithm of this function **F: CONFLICT-SCORE_HANDLING,** lines 1 – 3 is used to convert the neighbourhood layers derived from function **F:NR_MEASURE** to neighbourhood index layers having values ranging from 0 to 1. In lines 4 – 12, the conditional CON function is incorporated with the UPOS function, which aims to solve the conflicting cells using neighbourhood index values. In line 6, cells that experience conflict will have a value of 9, otherwise 0. In lines 7 – 10, only cells that experience conflict (coded as 9) are conditioned using the neighbourhood index value to return values of 1 for residential, 2 for commercial, 3 for industrial. Those values are finally combined as a layer containing the order of category update, regarding 1 for residential, 2 for commercial, 3 for industrial in line 11. The OVER function in line 12 is applied to update the final order of the layers. At this stage, the conflicting cells are solved.

For example, if a particular cell location is chosen to change to commercial, that cell will get a value of 2. Finally, lines 13 – 15 are used to create the final products of three joint probability maps separately regarding residential, industrial and commercial using the ArcGIS CON function (see Figure 5.9).

Algorithm F:CONFLICT-SCORE_HANDLING

Input:

gRS	# A residential joint probability grid layer, derived from function either # **F:MNL-JP_COMPUTE** or **F:MCDA-JP_COMPUTE**.
gCM	# A commercial joint probability grid layer, derived from function either # **F:MNL-JP_COMPUTE** or **F:MCDA-JP_COMPUTE**.
gMA	# An industrial joint probability grid layer, derived from function either # **F:MNL-JP_COMPUTE** or **F:MCDA-JP_COMPUTE**.
gZero	# A layer that all cells contain a value of 0.
gNiPctFocalRS	# A residential neighborhood layer, derived from function **F: NR_MEASURE**
gNiPctFocalCM	# A commercial neighborhood layer, derived from function **F: NR_MEASURE**
gNiPctFocalMA	# An industrial neighborhood layer, derived from function **F: NR_MEASURE**

Output:

gFinalJPRS	# A final product of joint probability grid layer of vacant change to residential.
gFinalJPCM	# A final product of joint probability grid layer of vacant change to commercial.
gFinalJPMA	# A final product of joint probability grid layer of vacant change to industrial. # Remark: Each layer generated contains (floating) probability values ranging from # 0.0 to 1.0.

Analysis environment settings:

gMask	# The spatial mask layer for the analysis function that specifies the areas to be # processed, here derived from a vacant grid layer that is created from function # **F:LU_EXTRACT** (based on the CON function).

CONFLICT-SCORE_HANDLING(*gRs,gCM,gMA, gZero, gNiPctFocalRS, gNiPctFocalCM, gNiPctFocalMA*)

1. *gNiRS* = *gNiPctFocalRS* / 100 — # ArcGIS command: the # arithmetic operator function.

2. *gNiCM* = *gNiPctFocalCM* / 100

3. *gNiMA* = *gNiPctFocalMA* / 100

4. *gPos0* = **UPOS(***gRS, gCM, gMA***)** — # ArcGIS command: the # UPOS function.

5. *gPos1* = **CON(** (*gRS* == *gCM*) or (*gRS* == *gMA*) or _ (*gCM* == *MA*), 9, **UPOS(***gRS, gCM, gMA***))** — # ArcGIS command: the CON # and UPOS function. # An underscore (_) shows a # line connection.

6. *gPos2* = **CON(** *gPos1* = 9, 9, 0**)**

7. *gPos3* = **CON(** (*gPos2* == 9) and (*gPos0* == 1) and (*gRS* == *gCM*) _ and (*gRS*<> *gMA*),**UPOS(***gNiRS, gNiCM, gZero***)**, 0**)**

8. *gPos4* = **CON(**(*gPos2* == 9) and (*gPos0* == 1) and (*gRS* == *gMA*) _ and (*gRS* <> *gCM*), **UPOS(***gNiRS, gZero, gNiMA***)**, 0**)**

9. *gPos5* = **CON(**(*gPos2* == 9) and (*gPos0* == 2) and (*gCM* == *gMA*) _ and (*gRS* <> *gCM*), **UPOS(***gZero, gNiCM, gNiMA***)**, 0**)**

10. *gPos6* = **CON(**(*gPos2* == 9) and (*gPos0* == 1) and (*gRS* == *gCM*) _
and (*gRS* == *gMA*), **UPOS(***gNiRS*, *gNiCM*, *gNiMA*), 0)

11. *gPos7* = *gPos3* + *gPos4* + *gPos5* +*gPos6* # ArcGIS command: the # arithmetic operator function.

12. *gFinalPos* = *gPos7* **OVER** *gPos0* # ArcGIS command: the # OVER function.

13. *gFinalJPRS* = **CON(** *gFinalPos* == 1, *gRS*, 0) # ArcGIS command: the CON # function.

14. *gFinalJPCM* = **CON(** *gFinalPos* == 2, *gCM*, 0)

15. *gFinalJPMA* = **CON(** *gFinalPos* == 3, *gMA*, 0)

Function F:LUCHG_UPDATE (Update of land use transition type). This function is used to update the land use map. Figure 5.20 shows the logical flow of this function. Input to the model comprises (i) three final joint probability maps for residential, industrial and commercial which are derived from function **F:CONFLICT-SCORE_HANDLING**, (ii) a set of corresponding threshold values received from the threshold setting section of the main GUI (see Figure 5.11), and an update ranking option received from the update ranking section of the main GUI (see Figure 5.11). The first input, joint probability maps, are floating rasters. They need to be converted to integer raster to produce associated tables being used in the remainder of this function. A similar procedure as described in Section 5.3.1.1 is performed by multiplying each of them with a constant value of 10,000 in order to avoid truncating of floating values when converting these values to be used in an attribute table. The second input, regarded as a set of threshold values will be calculated as the number of cells allowed to change for each iteration. The third input, the ranking option, plays a major role in determining the order of update layer in the update analysis. In Figure 5.20, if users select the default update ranking (RS-CM-MA), it means the update ranking will begin with the joint probability map of vacant change to residential (RS) as a first layer (LAYER A in a diagram), the joint probability map of vacant change to commercial (CM) as a second layer (LAYER B in a diagram), and the joint probability map of vacant change to commercial (CM) as a third layer (LAYER C in a diagram).

Process of this function (see Figure 5.20) comprises four sub-tasks; **F:SEL_PROB_FROM_TABLE,** ArcGIS command: EXTRACTION, **F:MASKOUT,** and **F:LU_UPDATE**.

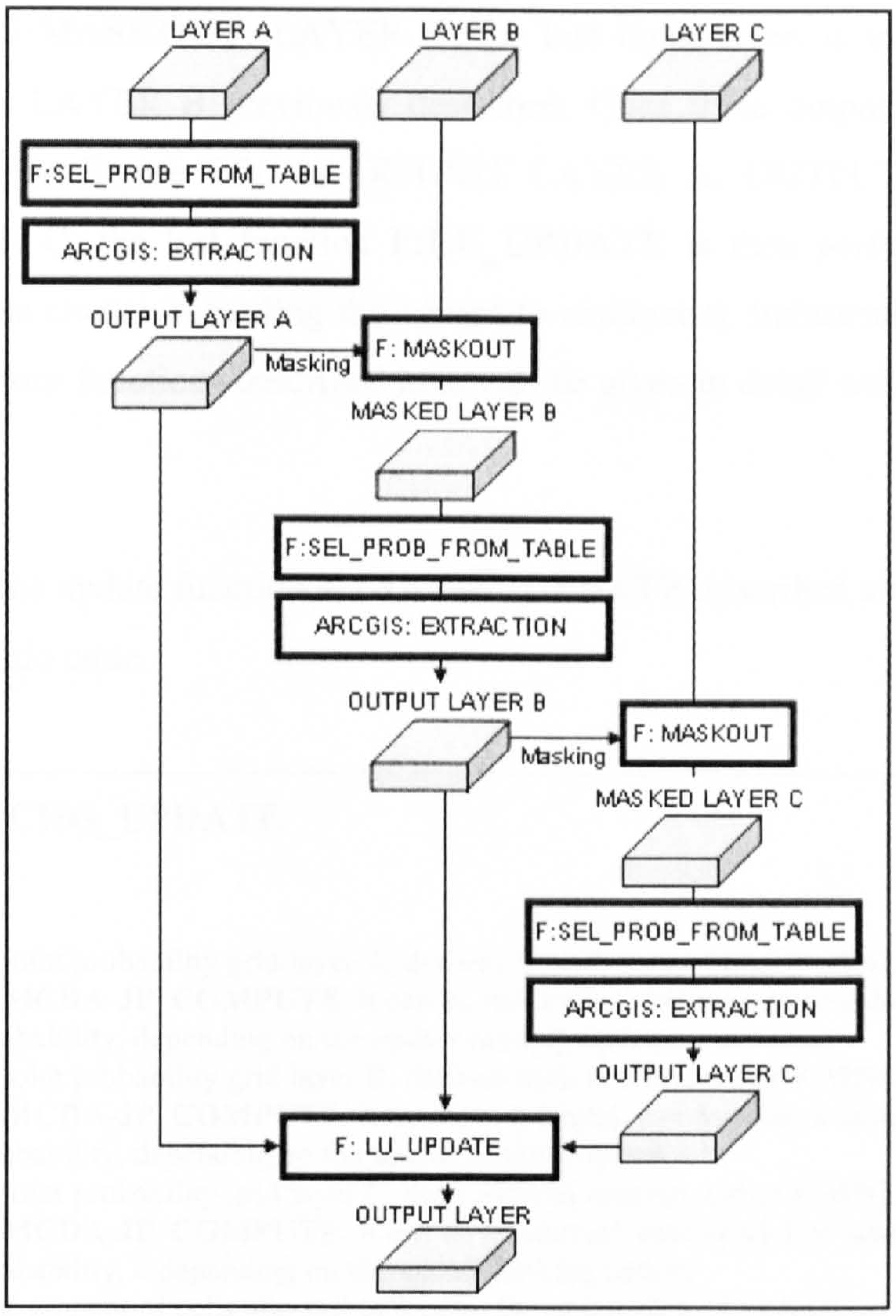


Figure 5.20: A logical flow of the update function applicable to this study.

According to the logical flow, starting with an input LAYER A, the function **F:SEL_PROB_FROM_TABLE** is used to search for the score that is used for selection of probability map in the next step. The criterion to find the score is by counting the cells until it reaches the total amount of cells allowed to change for a particular land use LAYER A. In the next step, the probability score chosen from **F:SEL_PROB_FROM_TABLE** is input to the ArcGIS command :EXTRACTION to be used to extract the probability cells. Once accomplished, an output LAYER A is created. The output is then used for undertaking two main tasks. The first task is to be used as a masking layer for LAYER B. This means that the cell locations that have occupied by LAYER A will subtract LAYER B. Function **F:MASKOUT** is used to perform this task. The second task is to be used as an input for land use update at the end of the flow. The outcome is the MASKED LAYER B that will be input and perform a similar process as that of LAYER A, including **F:SEL_PROB_FROM_TABLE,** ArcGIS command:

EXTRACTION, **F:MASKOUT**. LAYER C, the last input layer, is undertaken using a similar process as LAYER B previously described. Once three output layers have been generated, shown in Figure 5.20 as OUTPUT LAYER A, OUTPUT LAYER B and OUTPUT LAYER C, the last function **F:LU_UPDATE** is then performed in order to update the land use change regarding the vacant to residential, industrial and commercial. Details about the four functions described here will be given in detail with the remainder of this function.

The algorithm of the update function **F:LUCHG_UPDATE** described above is written as the following pseudo code.

Algorithm F: LUCHG_UPDATE

Input:

gFinalJPA	# A joint probability grid layer A, derived from function either **F:MNL-JP_COMPUTE** or # **F:MCDA-JP_COMPUTE**. It can be residential, commercial or industrial joint # probability, depending on the update ranking option.
gFinalJPB	# A joint probability grid layer B, derived from function either **F:MNL-JP_COMPUTE** or # **F:MCDA-JP_COMPUTE**. It can be residential, commercial or industrial joint # probability, depending on the update ranking option.
gFinalJPC	# A joint probability grid layer C, derived from function either **F:MNL-JP_COMPUTE** or # **F:MCDA-JP_COMPUTE**. It can be residential, commercial or industrial joint # probability, # depending on the update ranking option.
NoCellA	# The amount of cells allowed to change for an iteration of a grid layer A.
NoCellB	# The amount of cells allowed to change for an iteration of a grid layer B.
NoCellC	# The amount of cells allowed to change for an iteration of a grid layer C.
LtA	# A land use type of a grid layer A, whose value can be 1 for residential, 2 for # commercial, and 3 for industrial.
LtB	# A land use type of a grid layer B, whose value can be 1 for residential, 2 for # commercial, and 3 for industrial.
LtC	# A land use type of a grid layer C, whose value can be 1 for residential, 2 for # commercial, and 3 for industrial.

Output:

gUpdLuchg	# An output update land use map

Analysis environment setting:

gMask	# The spatial mask layer for the analysis function that specifies the areas to be processed, # here derived from a vacant grid layer that is created from function # **F:LU_EXTRACT** # (based on the CON function).

LUCHG_UPDATE (*gFinalJPA, gFinalJPB, gFinalJPC, NocellA, NocellB, NocellC, LtA, LtB, LtC*)

```
1. gIntJPA = gFinalJPA * 10000                      # ArcGIS command: the
                                                    # arithmetic operator
2. gIntJPB = gFinalJPB * 10000                      # function.

3. gIntJPC = gFinalJPC * 10000
```

4. *VProbA*= **SEL_PROB_FROM_TABLE(***TABLE:gIntJPA, NOcellA***)**	# Customized VBA method: # see function #F:SEL_PROB_FROM_TABLE.
5. *gProbA* = **CON(** *gIntJPA* >= *VProbA*, *gIntJPA***)**	# ArcGIS command: the # CON function
6. *gMskIntJPB* = **MASKOUT(***gProbA*, *gIntJPB***)**	# ArcGIS command: see function F: MASKOUT.
7. *VProbB*= **SEL_PROB_FROM_TABLE (***TABLE:gMskIntJPB, NOcellB***)**	# Customized VBA method: # see function #F:SEL_PROB_FROM_TABLE.
8. *gProbB* = **CON(** *gIntJPB* >= *VProbB*, *gIntJPB***)**	# ArcGIS command: the # CON function
9. *gMskIntJPC*= **MASKOUT(***gProbB*, *gIntJPC***)**	# ArcGIS command: see function F: MASKOUT.
10. *VProbC*= **SEL_PROB_FROM_TABLE (***TABLE:gMskIntJPC, NOcellC***)**	# Customized VBA method: # see function #F:SEL_PROB_FROM_TABLE.
11. *gProbC* = **CON(** *gIntJPC* >="*VProbC*, *gIntJPC***)**	# ArcGIS command: the # CON function
12. *gUpdLuchg* = **LU_UPDATE(** *gProbA*, *gProbB*, *gProbC*, *LtA* ,*LtB*, *LtC***)**	# Customized VBA method: see function F:LU_UPDATE.

To call this function, for example, use **F: LUCHG_UPDATE** (CM, MA, RS, 100, 60, 5, 2, 3, 1). CM refers to *gFinalJPA* (first layer) being updated as the cell value of '2', MA as *gFinalJPB* (second layer) being updated as the cell value of '3', and RS as *gFinalJPC* (third layer) being updated as the cell value of '1'. The values of 100, 60, and 5 refer to the number of cells allowed to change for an iteration of a CM, MA, and RS respectively.

It should be noted that a different order of input layers as specified in an update ranking option, to some extent, unavoidably affects the outcome layer. Obviously, the occupied cells of the first layer mask out the cells of the second and third layers. Also the occupied cells of the second layer mask the cells of the third layer. For that reason, the cells of the third layer that have the highest scores and are definitely to be chosen, if updated first, are probably masked out from the analysis as they are occupied earlier by either the first or second layers.

The logical expression concept of the CON function is illustrated earlier in Figure 5.6. It should be noted that the ArcGIS CON functions in lines 5, 8, and 11 do not specify the <false_expression>, which means NoData will be assigned to those cells. In the remainder of this part, the algorithm of function **F:SEL_PROB_FROM_TABLE, F:MASKOUT,** and **F:LU_UPDATE** will be given respectively.

Function F:SEL_PROB_FROM_TABLE (Probability / Suitability Score Selection). This function is used to select the probability score that meets the criterion. The criterion, in this regard, means the total amount of cells allowed to change for a particular land use type (e.g. 44 cells are allowed to change from vacant to industrial during 1993 – 2001 on a yearly basis). This is done using the attribute table of a probability map derived from the joint probability function (e.g. probability maps derived either from Function **F:MNL-JP_COMPUTE** or **F:MCDA-JP_COMPUTE**).

The input to the function, the joint probability map for each land transition type, has an associated table that contains the probability or suitability score multiplied by 10,000 stored in the VALUE field. The function starts by sorting the "Attribute Tables" of the probability map at each iteration based on high probability, cumulatively counting cells until it reaches the total amount of cells assigned for an iteration, and finally storing the probability score of the last record currently counted. The outcome from the function, regarded as probability / suitability score value will be used to update the cells for land use change development in further analysis.

The following algorithm shows the pseudo code of the function **F:SEL_PROB_FROM_TABLE.** According to the algorithm, line 1 is to sort probability scores of the field "VALUE" on the basis of highest probability. Line 2 is used to advance the position of the rows in the table. The while-loop is used to count and cumulate cells. Loop is terminated when the numbers of cumulative cells greater than the total amounts of cells allowed. In line 8 the probability score of a record at the end of the while-loop is stored to be used for further function. In the algorithm, three main interfaces are used to for handling with an attribute raster table. They are ITableSort, ICursor and IRow. ITableSort interface provides access to members that return and modify information to sort a table. ICursor interface provides access to members that hand out enumerated rows, field collections and allows for the updating, deleting and inserting of rows. IRow interface provides access to members that return information about the row.

Algorithm F: SEL_PROB_FROM_TABLE

Input:

tTableLayer # An associated attribute table of a joint probability grid layer.
NoCell # The amount of cells allowed to change for an iteration.

Output:

SelProbValue # An output value that is used to store the probability score.

SEL_PROB_FROM_TABLE (*tTableLayer, NOCell*)

1. *tTableSort* = **TableSort**(*tTableLayer*, Descending, VALUE*)* # ArcObjects syntax: ITableSort interface # that provides properties and methods for # sorting table *tTableLayer* based on the # VALUE field in a descending order.

2. *tRow* = **NextRow**(*tTableSort*) # ArcObjects syntax: **NextRow** as a method # of ICursor interface that is used to advance # the position to the row by one.

3. **While** *Cum* < *NoCell* **Do**

4. *Count* =**Value**(*tRow*, COUNT) + *Count* # ArcObjects syntax: **Value** as a property of # IRow interface that is used to get or set the # value of the field "COUNT".

5. *Prob* = **Value**(*tRow*, VALUE) # ArcObjects syntax: **Value** as a property of # IRow interface that is used to get or set the # value of the field "COUNT".

6. *Cum* = *Count* + *Cum*

7. *tRow* = **NextRow**(*tTableSort*) # ArcObjects syntax: **NextRow** as a method # of ICursor interface that is used to advance # the position to the row by one.

8. **EndDo**

9. *SelProbValue* = *Prob* # set value of the last probability value *Prob* # to *SelProbValue*.

Function F:MASKOUT (Masking out). This function is used to extract the cells of a raster that correspond with the areas defined by a mask layer. This function is built on the two ArcGIS functions, the SETNULL and ISNULL functions. Both are map algebra functions which operate on a cell-by-cell basis. The ISNULL function is used to create a new raster layer where the NoData areas are assigned to '1', and '0' for other (non-NoData) cells. Conversely, the SETNULL is used to create a new raster layer based on the criterion in that if the evaluation of the input condition is 'True', return NoData. If it is 'False', returns the value specified by the raster or number.

Figure 5.21 shows how the ISNULL and SETNULL work. Suppose that its aim is to eliminate the cells location of the MA layer from the INTJP_CM layer. Firstly, the

ISNULL expression is used to convert NoData to ‘1’, otherwise ‘0’. Then, after applying the SETNULL expression, any cell with a value equal to '0' will be set to NoData, and the remaining cells will retain their original value of INTJP_CM layer.

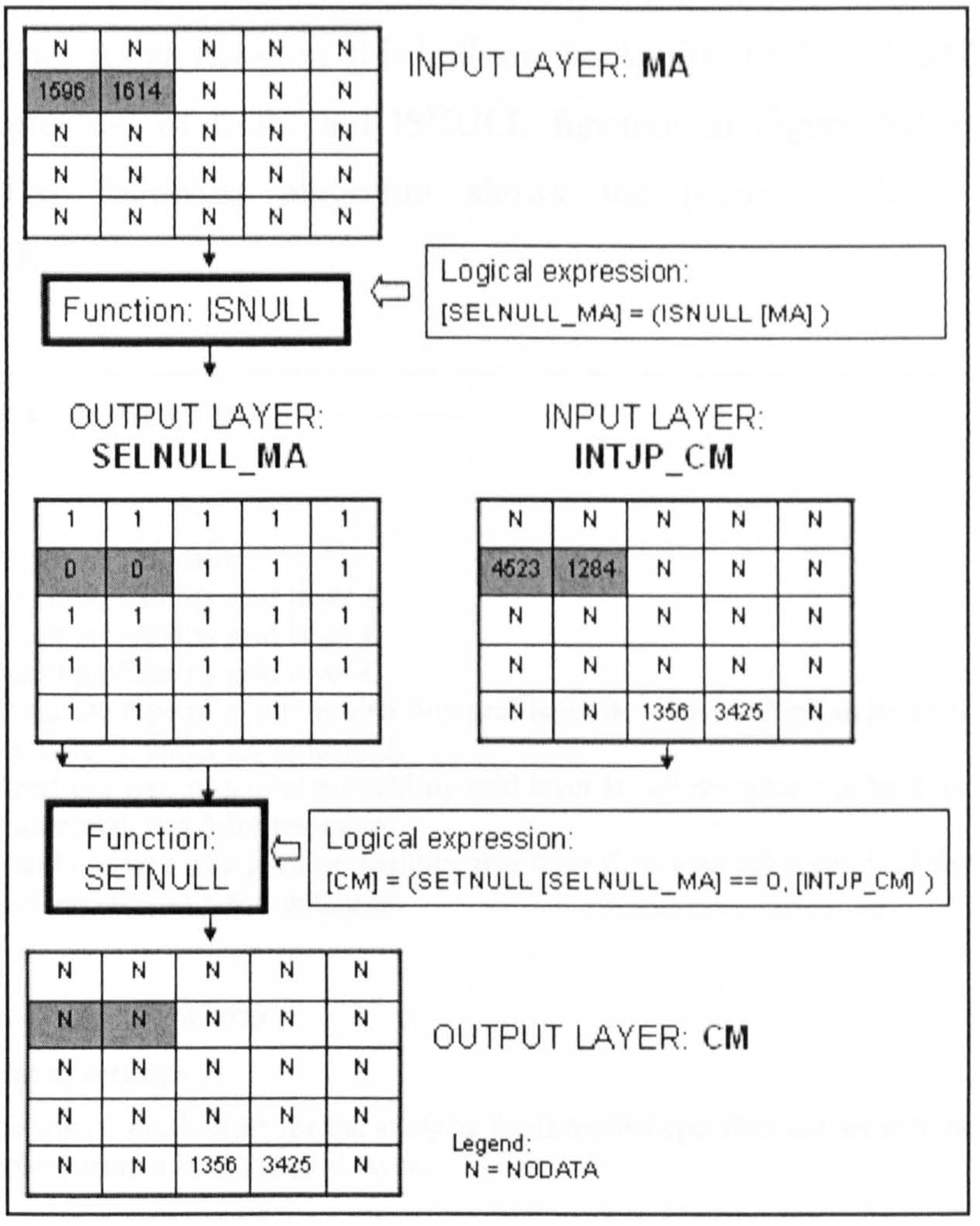


Figure 5.21: An example showing the operation of ISNULL and SETNULL function used for eliminating the cells location of one layer from another layer.

The following algorithm shows the pseudo code of the function **F:MASKOUT**.

Algorithm F: MASKOUT

Input:

gIntJP	# A joint probability grid layer.
gMask	# A masking grid layer defining areas to extract

Output:

gMskIntJP	# A masked joint probability grid layer.

MASKOUT(*gMask,gIntJP*)

1. *gIsNull* = **ISNULL(***gMask***)**	# ArcGIS command: the ISNULL function.
2. *gMskIntJP* = **SETNULL(***gIsNull* == 0, *gIntJP***)**	# ArcGIS command: the SETNULL function.

Function F:LU_UPDATE (Update of land use change). This function is used to update the land use change for an iteration. It is built on the ArcGIS CON and ISNULL functions. See logical expression of CON and ISNULL function in Figure 5.6 and Figure 5.21 respectively. The following algorithm shows the pseudo code of the function **F:LU_UPDATE**.

Algorithm F: LU_UPDATE

Input:

gLu	# A land use grid layer.
gLuA	# A joint probability grid layer A.
gLuB	# A joint probability grid layer B.
gLuC	# A joint probability grid layer C.
LucodeA	# A land use type of a joint probability grid layer A, whose value can be 1 for residential, 2 for # commercial, and 3 for industrial.
LucodeB	# A land use type of a joint probability grid layer B, whose value can be 1 for residential, 2 for # commercial, and 3 for industrial.
LucodeC	# A land use type of a joint probability grid layer C, whose value can be 1 for residential, 2 for # commercial, and 3 for industrial.

Output:

gLuchg	# An output update map

Analysis environment settings:

gMask	# The spatial mask layer for the analysis function that specifies the areas to be processed, here # derived from a land use grid layer.

LU_UPDATE (*gLu, gLuA, gLuB, gLuC, LucodeA, LucodeB, LucodeC*)

1. *gIsNullA* = **ISNULL(***gLuA***)**	# ArcGIS command: the ISNULL function.
2. *gIsNullB* = **ISNULL(***gLuB***)**	
3. *gIsNullC* = **ISNULL(***gLuC***)**	
4. *gUpdA*= **CON(** *gIsNullA* == 0, *LucodeA*, *gLu***)**	# ArcGIS command: the # CON function
5. *gUpdAB* = **CON(** *gIsNullB* == 0, *LucodeB*, *gUpdA***)**	
6. *gLuchg* = **CON(** *gIsNullC* == 0, *LucodeC*, *gUpdAB***)**	

Function F: SCENARIO_UPDATE (updating scenario). This function performs the updating of a new land use map, derived from function F:LU_UPDATE, with the scenario pre-selected in the pop-up 'Select Scenario' in the main GUI. It aims to update events occurred during the simulation period (e.g. new roads in 1998). A new final land use map created will be stored in the output directory specified by the GUI: Set Environment (Figure 5.13).

This function is built on the ArcGIS OVER function. It should be noted that the area of new events will be set as constraint areas beforehand in order to make sure that those vacant cells will not allow for development during the simulation. Figure 5.22 shows how the OVER function is applied for updating a new road event during the simulation in year 1998.

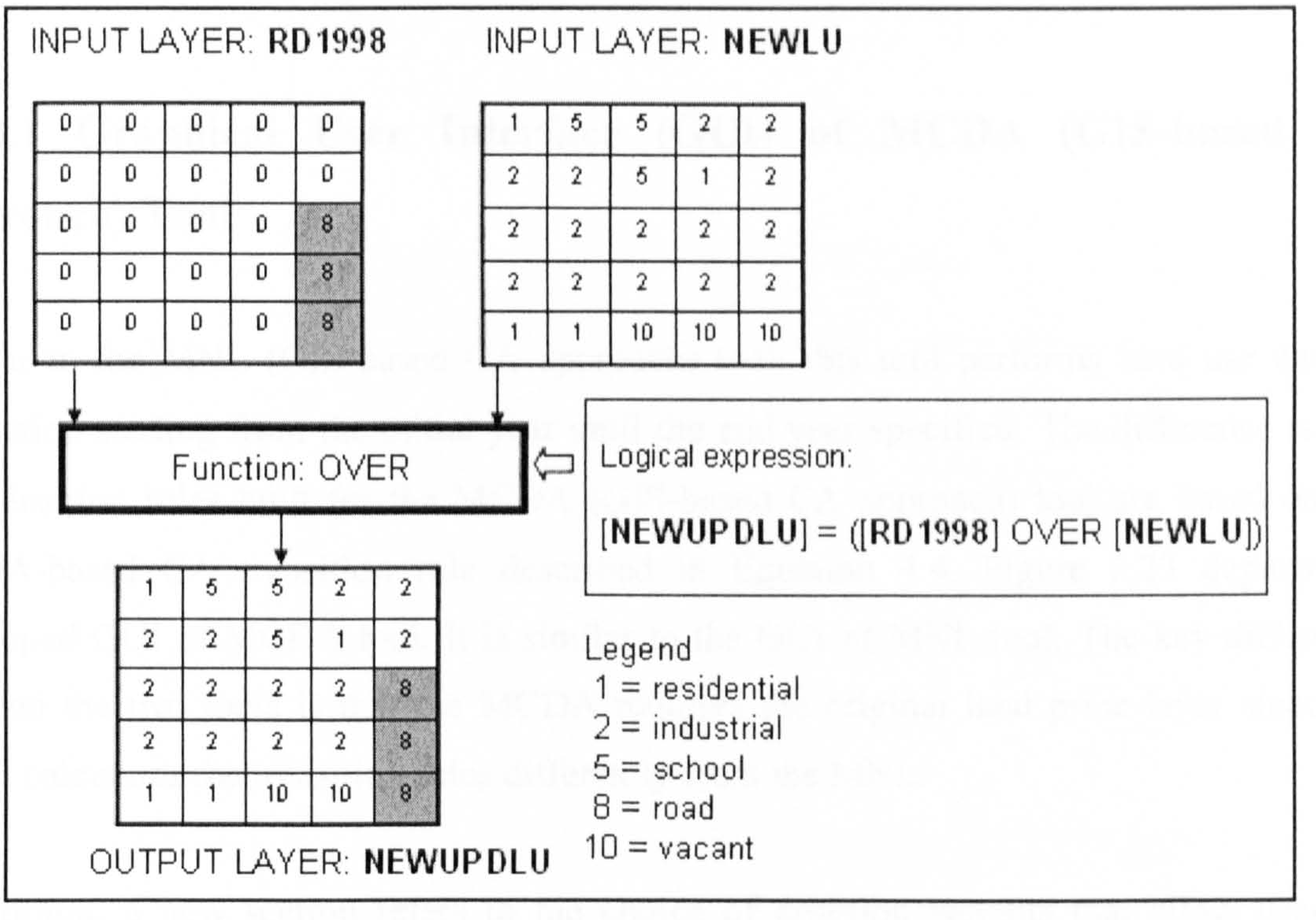


Figure 5.22: An example showing the operation of OVER function used for updating a new event (e.g. new road).

Algorithm F:SCENARIO_UPDATE

Input:

gLu	# A land use map, derived from function **F:LUCHG_UPDATE**.
gUpdateLayer	# A grid layer defining areas to update.

Output:

gNewLu	# A new land use grid layer.

Analysis environment settings:

gMask	# The spatial mask layer for the OVER function that specifies the areas to be # processed, here derived from a land use grid layer.

SCENARIO_UPDATE(*gLu*)

1. *gNewLu* = *gUpdateLayer* **OVER** *gLu*	# ArcGIS command: the # OVER function.

Function F: ITERATION (Iteration control). This function controls the number of loops (iterations) required. Once a new land use map is created, the program checks whether another iteration is required. If iteration is required, the program sets the previous iteration land use map as input for the next iteration.

5.3.3 MCDA (GIS-based CA Approach) Tool

5.3.3.1 Graphical User Interface (GUI) of MCDA (GIS-based CA Approach) Tool

Similar to the MNL (GIS-based CA approach) tool, this tool performs land use change simulation starting from the initial year until the end year specified. The difference is that the transition rules built for the MCDA (GIS-based CA approach) tool are based on the MCDA-based CA transition rule described in Equation 4.4. Figure 5.23 depicts the developed GUI of MCDA tool. It is similar to the GUI of MNL tool. The key difference between the two tools is that the MCDA requires the original land price layer since the model calculates the transition rules differently from the MNL.

In addition, a new section refers to the choice of criterion weights that allow users to choose a set of weights pre-prepared for testing different scenarios (e.g. testing road effect only). The button 'Weights Choice Description' (Figure 5.24) gives descriptions of weights and neighbourhood threshold values set for a simulation. The outcome of this tool is a series of simulated land use maps, similar to those shown in Figure 5.15. Note that user may run this tool by following the user's guide provided in Exercise 3, Appendix 2.

5.3.3.2 Execution of MCDA (GIS-based CA Approach) Tool

This tool, as in the MNL case, performs seven main tasks which split into 17 functions (see Table 5.3). The only difference is that the computation of joint probability is instead using the MCDA method. Figure 5.25 shows the flows of functions embedded in the MCDA tool. To avoid duplication of model description, in the remainder of this section description of the function **F: MCDA-JP_COMPUTE**, which is used to compute the joint probability maps, only is provided.

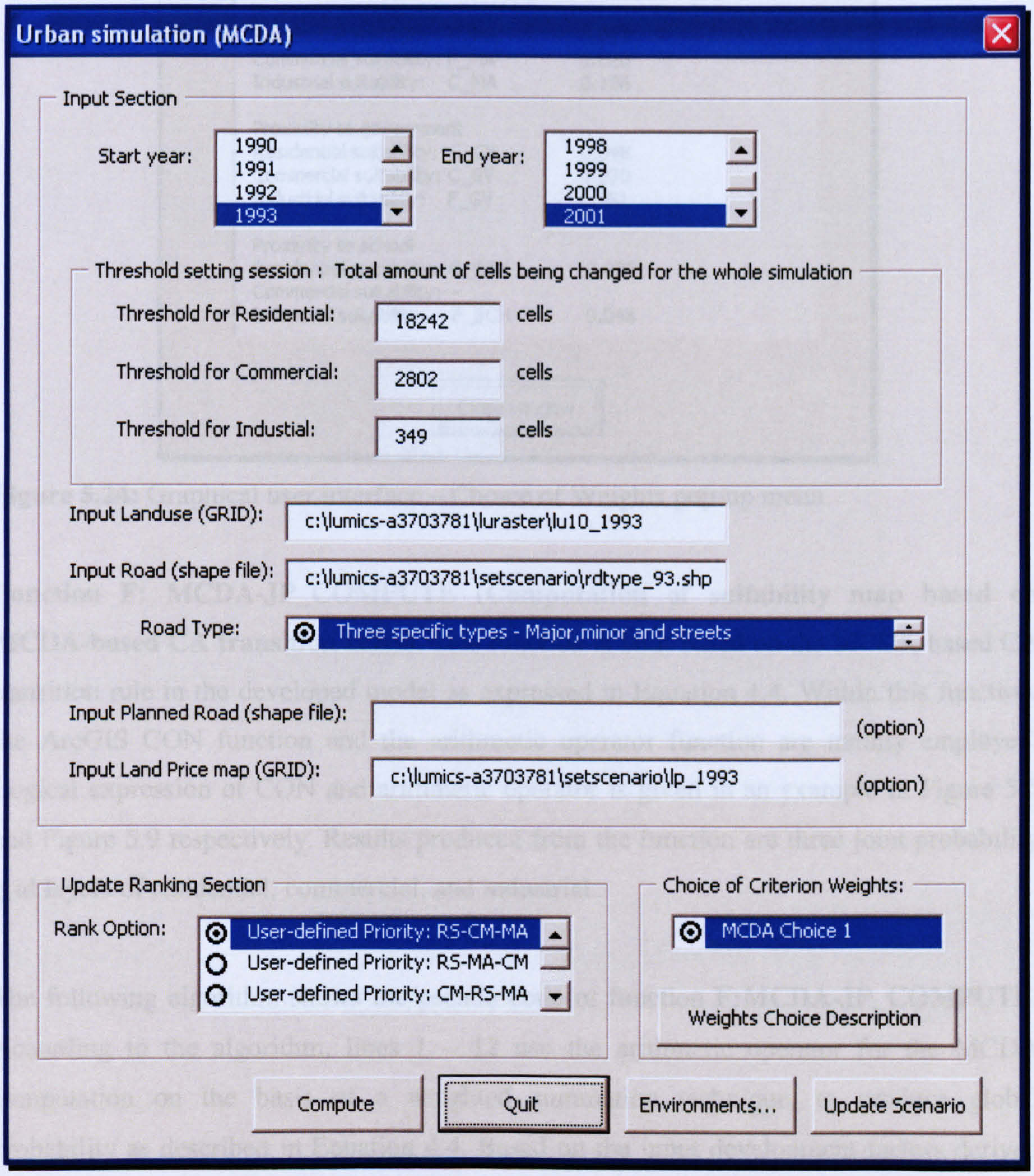


Figure 5.23: Graphical user interface – MCDA (GIS-based CA approach) Tool.

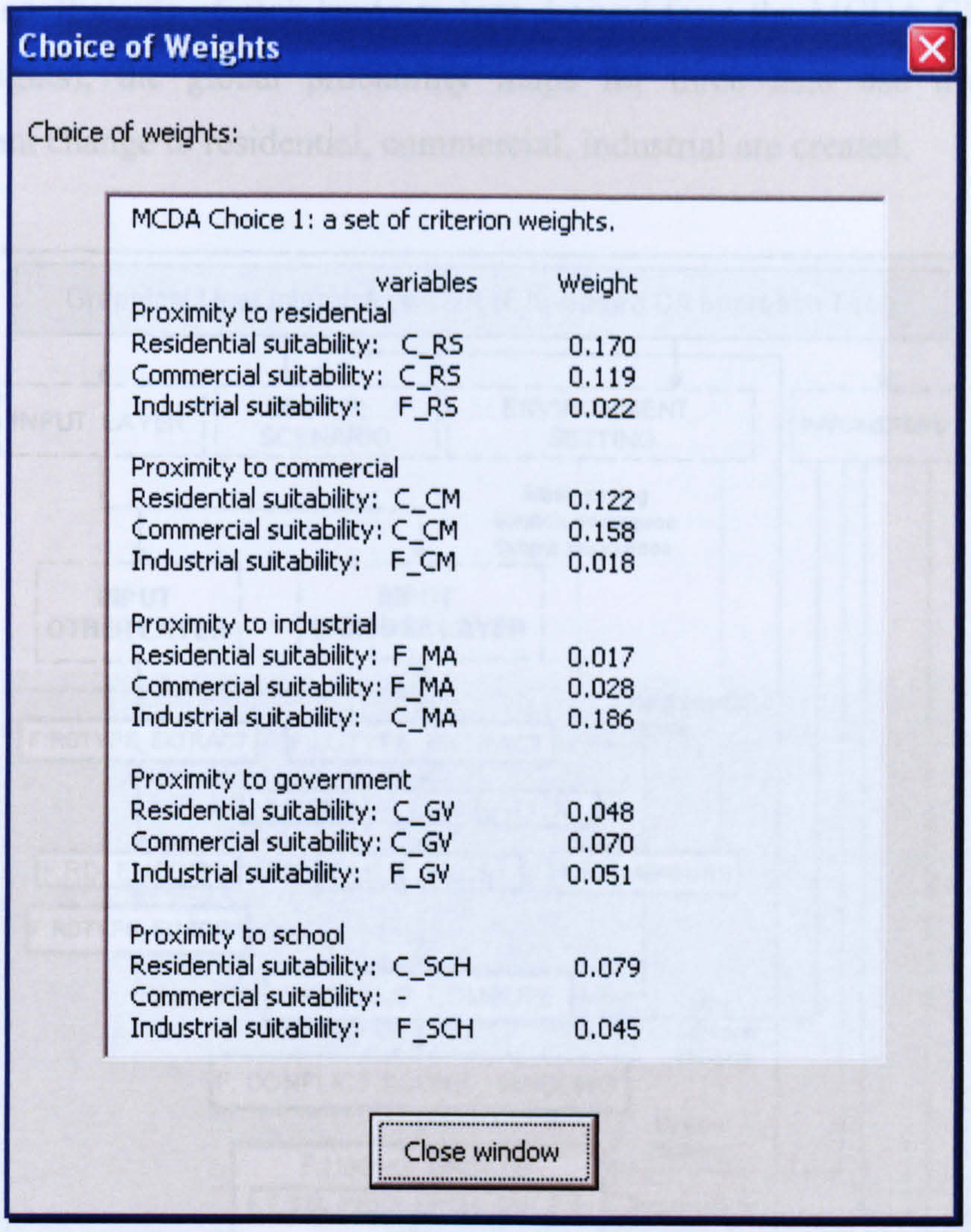


Figure 5.24: Graphical user interface – Choice of Weights pop-up menu.

Function F: MCDA-JP_COMPUTE (Computation of suitability map based on MCDA-based CA transition rules). This function is built based on the MCDA-based CA transition rule in the developed model as expressed in Equation 4.4. Within this function, the ArcGIS CON function and the arithmetic operator function are mainly employed. Logical expression of CON and arithmetic operator is given in an example in Figure 5.6 and Figure 5.9 respectively. Results produced from the function are three joint probability grid layers of residential, commercial, and industrial.

The following algorithm shows the pseudo code of function **F:MCDA-JP_COMPUTE**. According to the algorithm, lines 1 – 12 use the arithmetic operator for the MCDA computation on the basis of a weighted summation technique, to produce global probability as described in Equation 4.4. Based on the input development factors derived from functions **F:NR_MEASURE, F:LU_EUCDIST, F:RDTYPE_EUCDIST,** and a set of weights or coefficients of each land use type derived from the MCDA GUI (Choice of Criterion Weights), the global probability maps for three land use transition types regarding vacant change to residential, commercial, industrial are created.

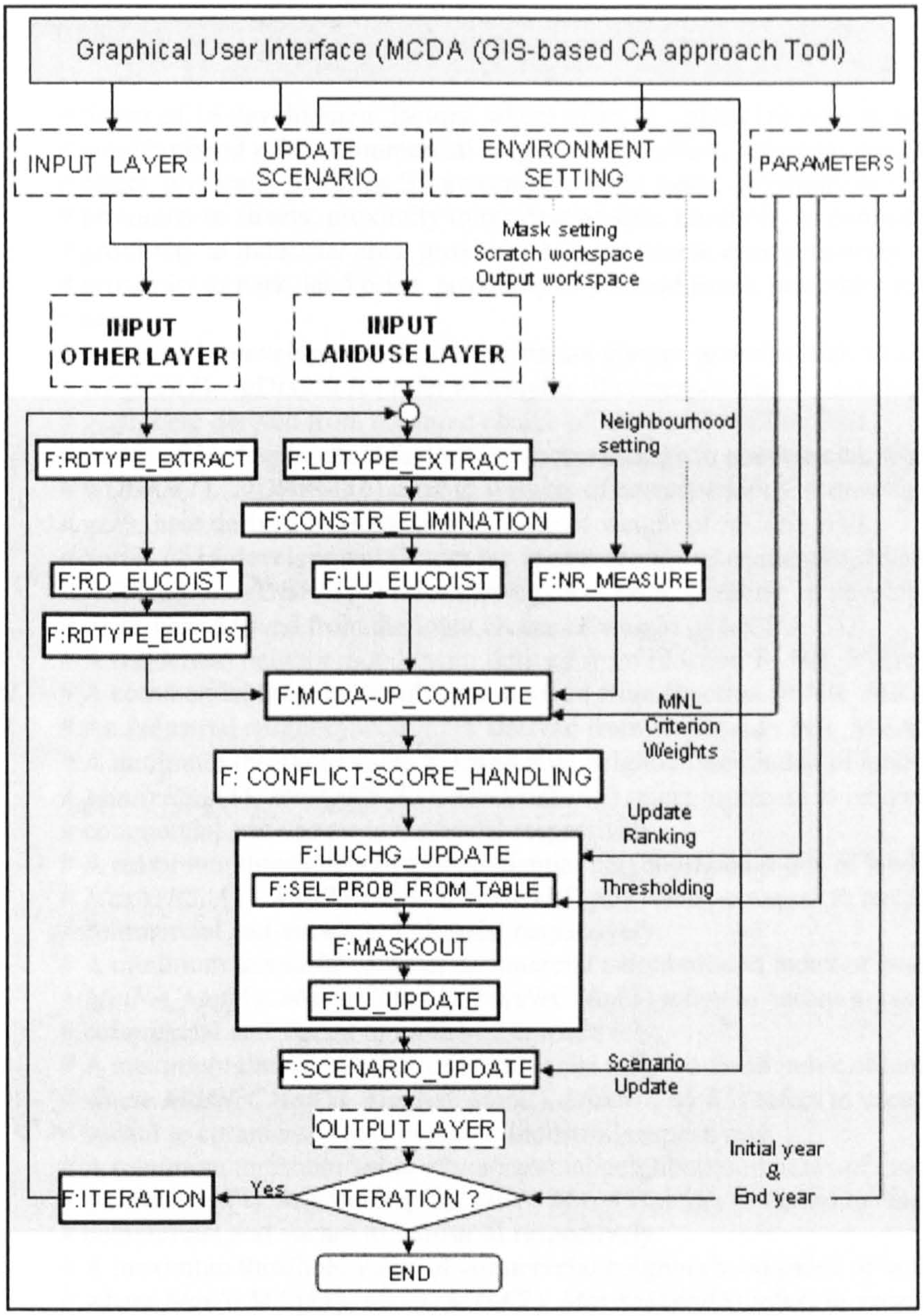


Figure 5.25: Model developed for MCDA (GIS-based CA approach) Tool.

In lines 13 – 15 the input neighbourhood layers are divided by 100 in order to convert them into neighbourhood index layers, having values ranging from 0.0 to 1.0. In lines 16 – 18, the neighbourhood thresholds using the CON function is used to limit the development for residential, industrial, and commercial layers respectively. Lines 19 – 21 refer to the creation of three joint probability layers of vacant change to residential, commercial, and industrial. Each is the combination of probability layers previously created, the corresponding constrained neighbourhood layer and the vacant area derived from the constraint map produced from function **F:CONSTR_ELIMINATION.**

Algorithm F: MCDA-JP_COMPUTE

Input:

gDfs	# Series of 16 development factors, where *gDfs*(*1*)...gDfs(*16*) refer to residential # neighborhood effect, commercial neighborhood effect, industrial neighborhood # effect, proximity to all roads, proximity to main roads, proximity to collector roads, # proximity to streets, proximity to residential area, proximity to commercial area, # proximity to industrial area, proximity to government area, proximity to school, # proximity to park, land price, proximity to planned roads, proximity to agriculture # area.
wDfRS	# Series of 16 development factors for vacant change to residential, where # w*DfsRS*(*1*)...gDfsRS(*16*) refer to weights of corresponding 16 development factors # *gDfs*, here derived from the input choice of weight of MCDA GUI.
wDfCM	# Series of 16 development factors for vacant change to commercial, where # w*DfsRS*(*1*)...gDfsRS(*16*) refer to weights of corresponding 16 development factors # *gDfs*, here derived from the input choice of weight of MCDA GUI.
wDfMA	# Series of 16 development factors for vacant change to residential, where # w*DfsRS*(*1*)...gDfsRS(*16*) refer to weights of corresponding 16 development factors # *gDfs*, here derived from the input choice of weight of MCDA GUI.
gNiPctFocalRS	# A residential neighborhood layer, derived from function **F: NR_MEASURE.**
gNiPctFocalCM	# A commercial neighborhood layer, derived from function **F: NR_MEASURE.**
gNiPctFocalMA	# An industrial neighborhood layer, derived from function **F: NR_MEASURE.**
MinNrRSof(T)	# A minimum threshold value of residential neighborhood index of land transition, where # *MinNrRSof*(1), *MinNrRSof*(2), *MinNrRSof*(3) refers to vacant to residential, vacant to # commercial and vacant to industrial respectively.
MaxNrRSof(T)	# A maximum threshold value of residential neighborhood index of land transition, where # *MaxNrRSof*(1), *MaxNrRSof*(2), *MaxNrRSof*(3) refers to vacant to residential, vacant to # commercial and vacant to industrial respectively.
MinNrCMof(T)	# A minimum threshold value of commercial neighborhood index of land transition, where # *MinNrCMof*(1), *MinNrCMof*(2), *MinNrCMof*(3) refers to vacant to residential, vacant to # commercial and vacant to industrial respectively.
MaxNrCMof(T)	# A maximum threshold value of commercial neighborhood index of land transition, # where *MaxNrCMof*(1), *MaxNrCMof*(2), *MaxNrCMof*(3) refers to vacant to residential, # vacant to commercial and vacant to industrial respectively.
MinNrMAof(T)	# A minimum threshold value of commercial neighborhood index of land transition, where # *MinNrMAof*(1), *MinNrMAof*(2), *MinNrMAof*(3) refers to vacant to residential, vacant to # commercial and vacant to industrial respectively.
MaxNrMAof(T)	# A maximum threshold value of commercial neighborhood index of land transition, # where *MaxNrMAof*(1), *MaxNrMAof*(2), *MaxNrMAof*(3) refers to vacant to residential, # vacant to commercial and vacant to industrial respectively.
T	# A land use transition type.

Output:

gJPRS	# A joint probability grid layer of vacant change to residential.
gJPCM	# A joint probability grid layer of vacant change to commercial.
gJPMA	# A joint probability grid layer of vacant change to industrial.
	# Remark: Each layer generated contains (floating) probability values ranging from # 0.0 to 1.0.

Analysis environment settings:

gMask	# The spatial mask layer for the analysis function that specifies the areas to be processed, # here derived from a vacant grid layer that is created from function **F: LU_EXTRACT** # (based on the CON function).

```
MCDA-JP_COMPUTE (gDfs, wDfRS, wDfCM, wDfMA, gNiPctFocalRS, gNiPctFocalCM, gNiPctFocalMA,
                 MinNrRSof(T), MaxNrRSof(T), MinNrCMof(T), MaxNrCMof(T), MinNrMAof(T),
                 MaxNrMAof(T))

1. For i = 1 to 16                                   # For-loop to
                                                     # accumulate a
2. gLayerRS = gLayerRS + ( gDfs(i) * wDfRS(i) )      # residential factors
                                                     # grid layer using
3. EndFor                                            # ArcGIS command:
                                                     # the arithmetic
                                                     # operator.

4. For i = 1 to 16                                   # For-loop to
                                                     # accumulate a
5. gLayerCM= gLayerCM + ( gDfs(i) * wDfCM(i) )       # commercial factors
                                                     # grid layer using
6. EndFor                                            # ArcGIS command:
                                                     # the arithmetic
                                                     # operator.

7. For i = 1 to 16                                   # For-loop to
                                                     # accumulate a
8. gLayerMA= gLayerMA + ( gDfs(i) * wDfMA(i) )       # residential factors
                                                     # grid layer using
9. EndFor                                            # ArcGIS command:
                                                     # the arithmetic
                                                     # operator

10. gProbRS = gLayerRS  * WObjRS                     # ArcGIS command: #
                                                     the arithmetic
11. gProbCM= gLayerCM  * WObjCM                      # operator to
                                                     # calculate MCDA
12. gProbMA= gLayerMA  * WObjMA                      # suitability score for
                                                     # residential,
                                                     # industrial,
                                                     # and commercial
                                                     # respectively.

13. gNiRS = gNiPctFocalRS / 100                      # ArcGIS command:
                                                     # the arithmetic
14. gNiCM = gNiPctFocalCM / 100                      # operator to
                                                     # calculate
15. gNiMA = gNiPctFocalMA / 100                      # neighborhood index
                                                     # layer for residential,
                                                     # industrial and
                                                     # commercial
                                                     # respectively.

16. For T = 1 to 3                                   # For each land use
                                                     # type

17. gThNiRS = Con( (gNiRS >= MaxNrRSof(T) and gNiRS <= MinNrRSof(T)) _        # ArcGIS command:
              or (gNiCM >= MaxNrCMof(T) and gNiCM<= MinNrCMof(T)) _           the CON function
              or (gNiMA >= MaxNrMAof(T) and gNiMA <= MinNrMAof(T)), 1, 0)     # to limit
                                                                              #development
18. EndFor                                                                    # for residential,
                                                                              # industrial and
                                                                              # commercial layer
                                                                              # respectively.
```

```
19. gJPRS = gProbRS * gThNiRS          # ArcGIS command:
                                       the arithmetic operator
20. gJPCM = gProbCM * gThNiCM          # to calculate final
                                       # cell probability for
21. gJPMA = gProbMA * gThNiMA          # residential, industrial
                                       # and commercial
                                       # respectively.
```

5.4 Conclusion: Advantages and Limitations of Customized Tools Developed.

In this chapter, the proposed model in Chapter 4 has been developed using a VBA macro in the ArcGIS 9.1 environment (and updated in the ArcGIS 9.2). In summary, there are two main reasons for this development. Firstly, VBA through ArcObjects provides access to all ArcGIS functionalities. Secondly, VBA has the same tools and many utilities as Visual Basic (e.g. VBE, custom command syntaxes), being built under the concept of object oriented programming. It thus allows programmers who are familiar with an object programming language to program with little development effort. Also in this chapter, the development of a set of customized tools, named LUMICS (Land Use MICro Simulation model), is described. They all are built by using VBA and ArcObjects, and comprise a Variables Observation tool, an MNL (GIS-based CA approach) tool, and an MCDA (GIS-based CA approach) tool. Each is presented in terms of the graphical user interface (GUI) and its execution. Its algorithms are provided in the form of pseudo code.

These customized tools contributed to spatial prediction helping to simulate urban land use change applicable to the study site. The graphical user interfaces developed are designed to be easy-to-use to the end users and specific to the user's requirements. There are many advantages. Firstly, programming through a VBA macro and ArcObjects helps filling in important gap in ArcGIS functionalities, especially in the context of repetitive tasks. Secondly, these tools help in reducing simulation time tremendously, compared to the manual operation for the whole simulation process. Thirdly, the tools prevent faulty output generation, which is likely to occur due to manual input error. More importantly, they allow several sets of parameters and weights to be easily set and altered, and thus support them as effective planning tools for alternative scenario testing.

Despite their benefits, the tools developed in this study have some limitations. Firstly, these tools were programmed with limited time and with much modification and adjustment, in order to help facilitate testing the proposed modelling framework described in Chapter 4. They are still in an early phase of development. These tools were developed without considering efficiency in terms of well-structured programming, computation time, their flexibility for model and parameter extension. Secondly, in terms of programming itself, despite many ArcObjects code samples provided in the help system and on-line on the internet, the ArcObjects themselves require thorough understanding as they involve many interfaces to interact with in order to acquire a result. However, without these tools, it is difficult to test the proposed modelling framework in Chapter 4 with many scenarios. These tools are employed to simulate many assumptions and scenarios, which are implemented and presented in the next chapter.

CHAPTER 6

Simulating the Spatial Pattern of Urban Land Use Change in Ladprao District, Bangkok, Thailand

6.1 Introduction

In this chapter, the GIS-based CA model with the integration of MNL and MCDA method described in Chapter 4 will be used to predict the spatial patterns of urban land use change for Ladprao District, Bangkok, Thailand during the period 1993 – 2001. The first method (Section 6.2) and the second method (Section 6.3) present and validate the simulation results created. Finally, a conclusion of the results generated is given in Section 6.4.

It should be noted that all simulation results reported in the next two sections (Section 6.2 and 6.3) were generated using identical settings. First, all simulated results were undertaken using linear growth temporally during the whole simulation period. Second, the simulation model was set to run iteratively in a coarser two-year interval in order to avoid overload on computer memory. For each run, the simulated results, based on the initial 1993 land use map, were generated as 1995, 1997, 1999 and 2001. Third, four types of transitions were permitted, including the vacant change to residential, commercial, industrial and vacant (no change). Fourth, as illustrated in Table 4.1, thresholds or numbers of cells allowed to change during the simulation period based on 10m grid resolution were applied to ensure 18242 vacant to residential changed cells, 2802 vacant to commercial and 349 vacant to industrial, reflecting the real world change from 1993 to 2001. Fifth, the order of category update, as shown in the update ranking section of the MNL and MCDA GUI (Figure 5.11 and Figure 5.23 respectively), was set to RS-CM-MA. This means that the order of category update starts with residential (RS), commercial (CM) and industrial (MA) respectively. Description of the order of category update is given in details in 'update ranking', Section 5.3.2.1. Finally, the update scenario, as shown in Update scenario pop-up menu of the MNL and MCDA GUI (Figure 5.13), was set to Scenario 3. By using Scenario 3, it is used to update the real events that occurred for the period of 1993 – 2001 in the study site. Figure 6.1 shows the real events that occurred during 1993 – 2001, which include roads in 1998 and 2001, and a department store in 2001. With Scenario 3, these events are updated in order as shown in Table 6.1.

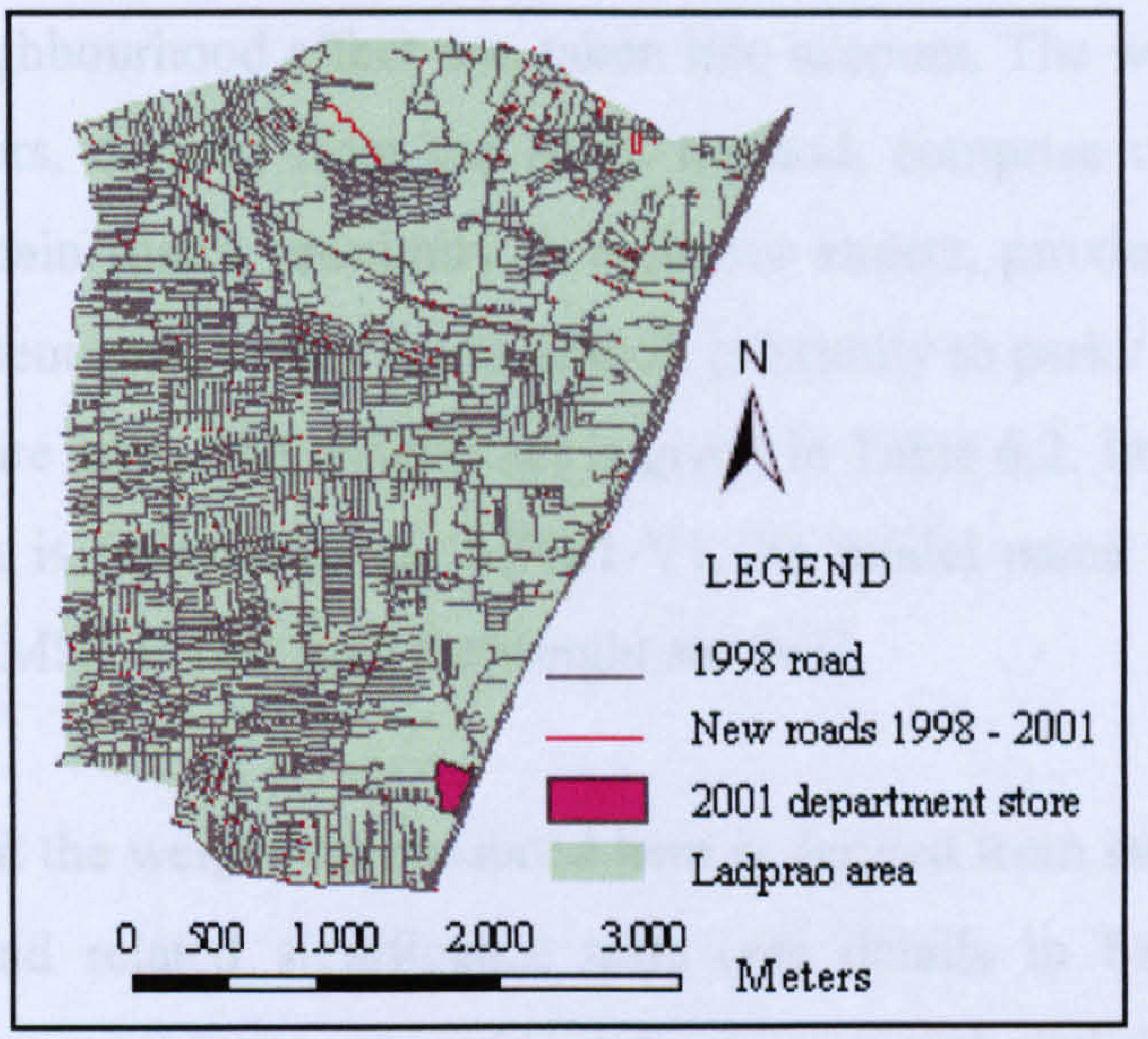


Figure 6.1: Real events in the study area during the period 1993 – 2001.

Simulation Description	*Year Schedule*								
	Initial 1993	1994	1995	1996	1997	1998	1999	2000	2001
Road 1998	■	■	■	■	■	■/▲	■	■	■
Road 2001	■	■	■	■	■	■	■	■	■/▲
New department store 2001	■	■	■	■	■	■	■	■	■/Δ

Remark: ■ refers to "setting as restricted area for development",

▲ refers to "update applied at the beginning of iteration",

and Δ refers to "update applied at the end of iteration

Table 6.1: Schedule of events set to update during simulation in the research study.

6.2 Simulated Results from the GIS-based CA/MNL Model

6.2.1 Simulated Results with a Full Set of Development Factors

The simulated results were produced using the MNL-based CA transition rule in Equation 4.9. Based on the equation, the incorporation of MNL probability transition rule, constraint rule and dynamic neighbourhood effect was taken into account. The weights of the full set of development factors, derived from the MNL method, comprise coefficients for land price, proximity to main roads, proximity to collector streets, proximity to local streets, proximity to government area, proximity to school, proximity to park / recreation area, and proximity to agriculture area. The weights set is given in Table 6.2. Implementation of the model in this section is referred to as 'MNL1-V1. Its model name refers to the model conducted by model 'MNL1' that uses the weight set 'V1'.

It should be noted that the weight set produced here is derived from statistical multinomial logistic regression and related significance tests (see details in Section 4.5.1). Three development factors – proximity to residential, commercial and industrial, were not included in this weight set because they exhibit multicollinearity when used in conjunction with residential, industrial and commercial neighbourhood factors (which were multiplied later to create joint probability in Equation 4.9).

According to the weights set in Table 6.2, the interpretation can be explained that for the vacant change to residential, proximity to local streets are the most positive influential factor while proximity to main roads is the most negative influence factor in the regression model. This means that being close to local streets will increase most the probability value of changing from vacant to residential. Conversely, being close to main roads will decrease most the probability value. For the vacant change to commercial, proximity to collector streets is the most positive influential factor while proximity to park is the most negative influence factor. For the vacant change to industrial, proximity to collector streets is the most positive influential factor while proximity to local streets is the most negative influence factor. More details about interpretation of logistic regression are given in Section 4.5.1.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Land price (LP)	0.363	0.979	0.809
Proximity to Main roads (D_RDT1)	-0.706	-1.051	4.123
Proximity to Collector Streets (D_RDT2)	6.804	27.973	13.823
Proximity to Local Streets (D_RDT3)	53.015	1.253	-10.080
Proximity to Government (D_GV)	-0.643	2.062	-4.500
Proximity to School (D_SCH)	2.368	0.489	1.130
Proximity to Park (D_PRK)	-0.327	-3.122	0.733
Proximity to Agriculture (D_AGR)	0.213	0.482	-6.868
Intercept	-63.700	-38.492	-14.255
-2LL	57977.964 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.2: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL1-V1'.

Figure 6.2 illustrates the simulated results of 1995, 1997, 1999 and 2001 land use maps. To allow focusing on the targeted land use types, in the remainder of this chapter the land use map result will be classified as five land use types (residential, commercial, industrial, vacant area, and other land use types) for the ease of visual interpretation. A visual observation of the simulated results (Figure 6.2) highlights two prominent characteristics. Firstly, the urban land development acts as a *space-filler* over time. Secondly, urban developments were through the process of *building accretion* as new emerged land use types simulated seem to enlarge the existing ones. Many developments occurred obviously because of the two key mutual effects: the expansion of existing land use types in one of both of the ways mentioned above, and being close to roads. Therefore, for example, vacant cells that were very close to existing commercial cells and roads experienced a transformation from vacant to commercial. In the simulation, new commercial cells did not emerge independently along the roads, but only next to previous commercial cells.

Further investigation was obtained by visually comparing the actual urban map of 2001 with the simulated one of 2001. In Figure 6.3, the spatial comparison of the results, magnifying two locations, reveals a very distinctively different growth pattern.

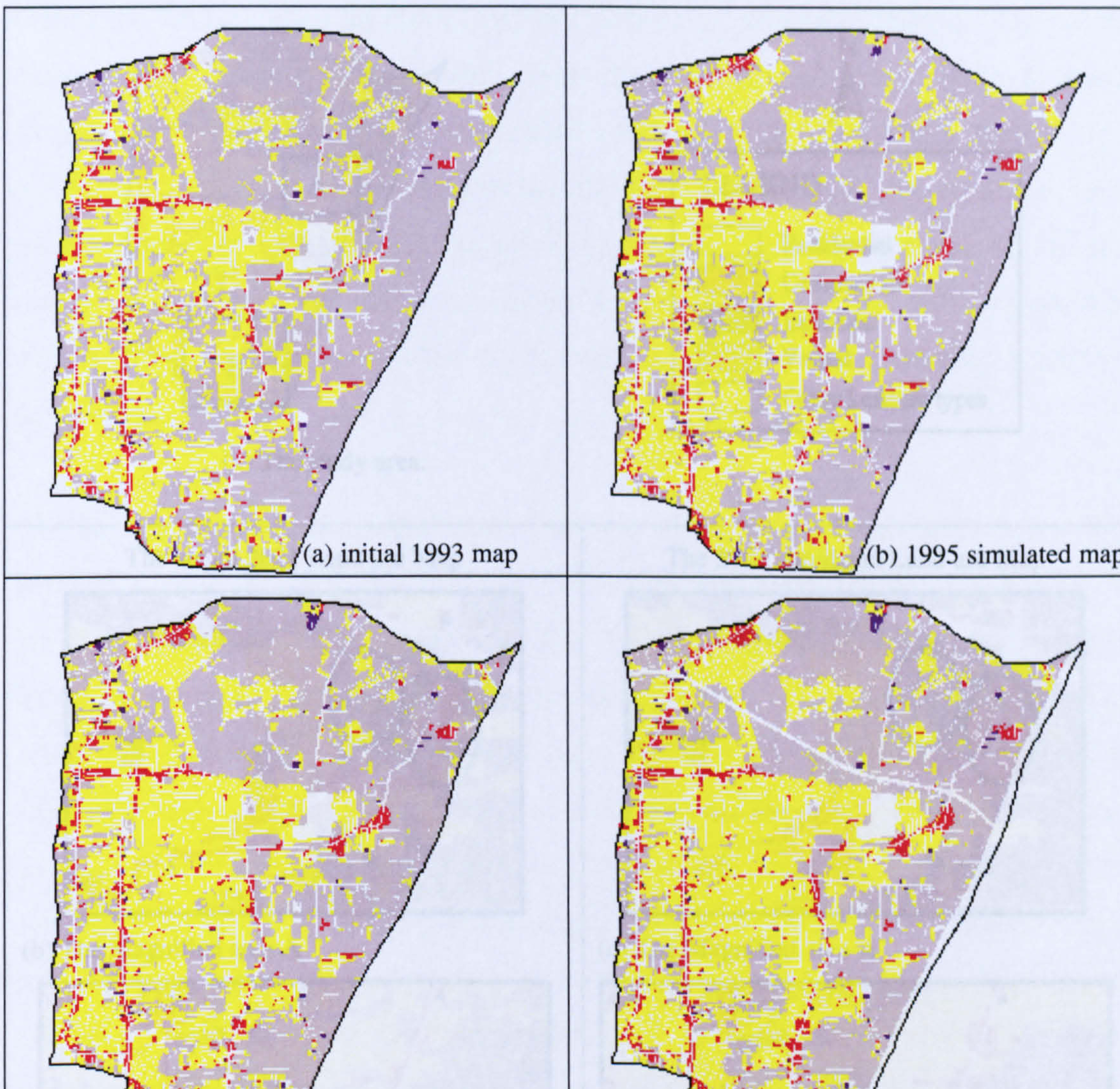
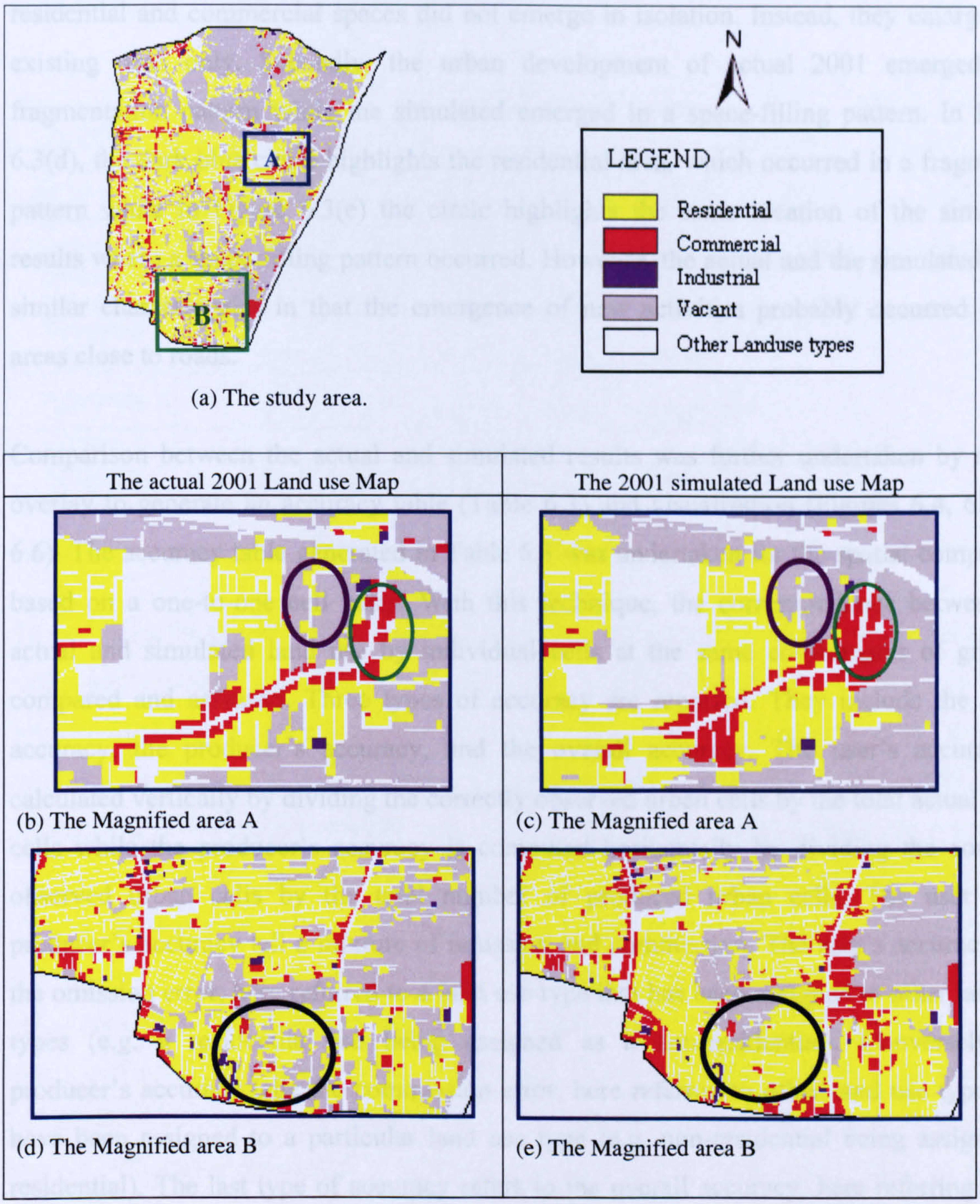


Figure 6.2: The simulated results of model 'MNL1-V1'.



(a) The study area.

(b) The Magnified area A

(c) The Magnified area A

(d) The Magnified area B

(e) The Magnified area B

Figure 6.3: Visual comparison of the actual and simulated land use map 2001 (model 'MNL1-V1).

Firstly, while the urban development of actual 2001 emerged spontaneously, the simulated 2001 occurred as an enlarging cluster around the existing development. The two highlighted areas in coloured circles in Figure 6.3 (b) and (c) support the finding. The purple and green circles highlight the locations where the emergence of new residential and new commercial spaces between the actual 2001 and simulated 2001 are compared respectively. In Figure 6.3(b), the new emergence of residential and commercial cells possibly dispersed independently along the roads while in Figure 6.3(c), these simulated residential and commercial spaces did not emerge in isolation. Instead, they enlarged the existing ones only. Secondly, the urban development of actual 2001 emerged in a fragmentation pattern while the simulated emerged in a space-filling pattern. In Figure 6.3(d), the dark blue circle highlights the residential area, which occurred in a fragmented pattern while in Figure 6.3(e) the circle highlights the same location of the simulated results where a space-filling pattern occurred. However, the actual and the simulated share similar characteristics in that the emergence of new activities probably occurred in the areas close to roads.

Comparison between the actual and simulated results was further undertaken by spatial overlay to generate an accuracy table (Table 6.3) and visualisation (Figures 6.4, 6.5 and 6.6). The accuracy table generated in Table 6.3 was undertaken for the spatial comparison based on a one-to-one cell basis. With this technique, the correspondence between the actual and simulated land use by individual cells at the same co-ordinates of grid are compared and assessed. Three types of accuracy are reported. They include the user's accuracy, the producer's accuracy, and the overall accuracy. The user's accuracy is calculated vertically by dividing the correctly observed urban cells by the total actual urban cells while the producer's accuracy is computed horizontally by dividing the correctly observed urban cells by the total number of produced urban cells. The user's and producer's accuracy tell a mixture of omission and commission. The user's accuracy tells the omission error, here referred to a land use type that has been assigned to other land use types (e.g. a residential use being assigned as a non-residential). Conversely, the producer's accuracy tells the commission error, here referred to other land use types that have been assigned to a particular land use type (e.g. non-residential being assigned as residential). The last type of accuracy refers to the overall accuracy, here referring to the correctly produced urban cells divided by the overall observed cells being considered.

It should be noted that since our focus in this study is on the change from vacant to residential, vacant to commercial, and vacant to industrial, the overall accuracy is thus computed without accounting for the producer's accuracy of vacant type, here referring to the actual non-vacant land (that is, residential, commercial and industrial) that has been assigned as vacant. In Table 6.3, the highlighted cells are used to compute the overall accuracy in the study. Overall the developed model performed poorly with a total accuracy of 31.59% being accomplished.

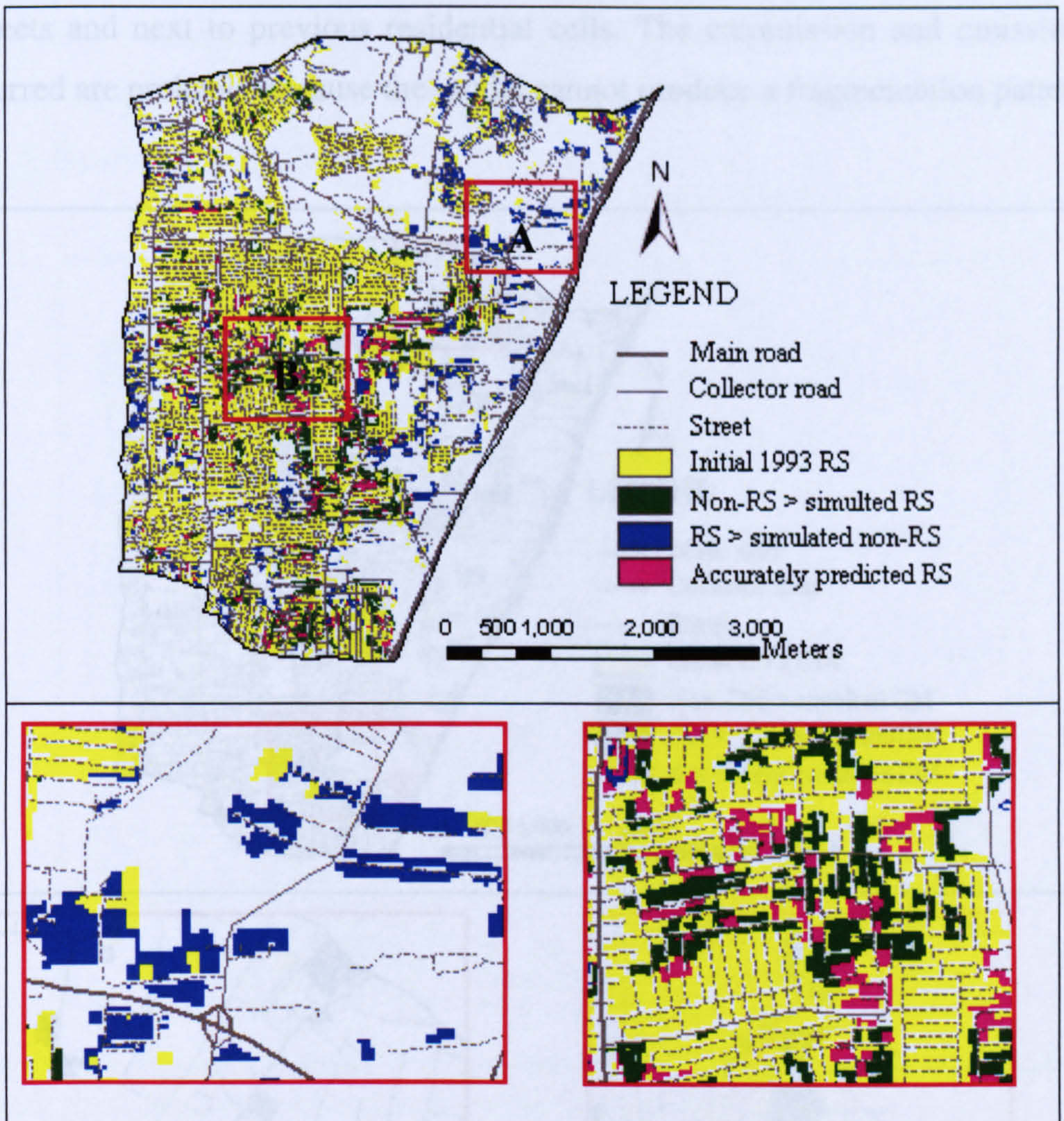
Category (no. of cells)	Residential (RS)	Commercial (CM)	Industrial (MA)	Vacant (VC)	Total cells	Producer's accuracy
Simulated Residential	6274	542	65	11376	18257	34.37
Simulated Commercial	1014	507	76	1207	2804	18.08
Simulated Industrial	0	76	0	328	404	0.00
Simulated Vacant	10954	1717	208	66650		
Total cells	18242	2842	349	79561		
User's Accuracy	34.39	17.84	0	83.77		

Total accuracy: 31.59%

Table 6.3: Evaluation of the simulated results of year 2001, based on model 'MNL1-V1'.

By excluding the consideration of vacant type, simulated residential type shows the highest user's and producer's accuracy of 34.39% and 34.37% respectively. In terms of user's accuracy, the model generated a 65.61% error of omission, which indicates the amount of land that is residential but has been assigned as other land use types (5.56% being commercial, 0.0% being industrial area, and 60.05% being vacant). In terms of producer's accuracy, the model produced a 65.63% error of commission, which indicates the amount of land that is non-residential but has been assigned as residential. The model assigned the non-residential (2.96% of commercial, 0.36% of industrial area, and 62.31% of vacant) as being residential use. The accuracy for commercial land use type is lower with the total user's and producer's accuracy of 17.84% and 18.08%. In terms of user's accuracy, the model generated an 82.16% error of omission, assigning commercial as non-commercial use (19.07% being residential, 2.67% being industrial area and 60.42% being vacant). In terms of producer's accuracy, the model assigned 81.92% of non-commercial as being commercial use (36.16% being residential, 2.71% being industrial area, and 43.05% being vacant). The worst case is the accuracy of industrial type as there was no match found between the actual and simulated maps for industrial simulation.

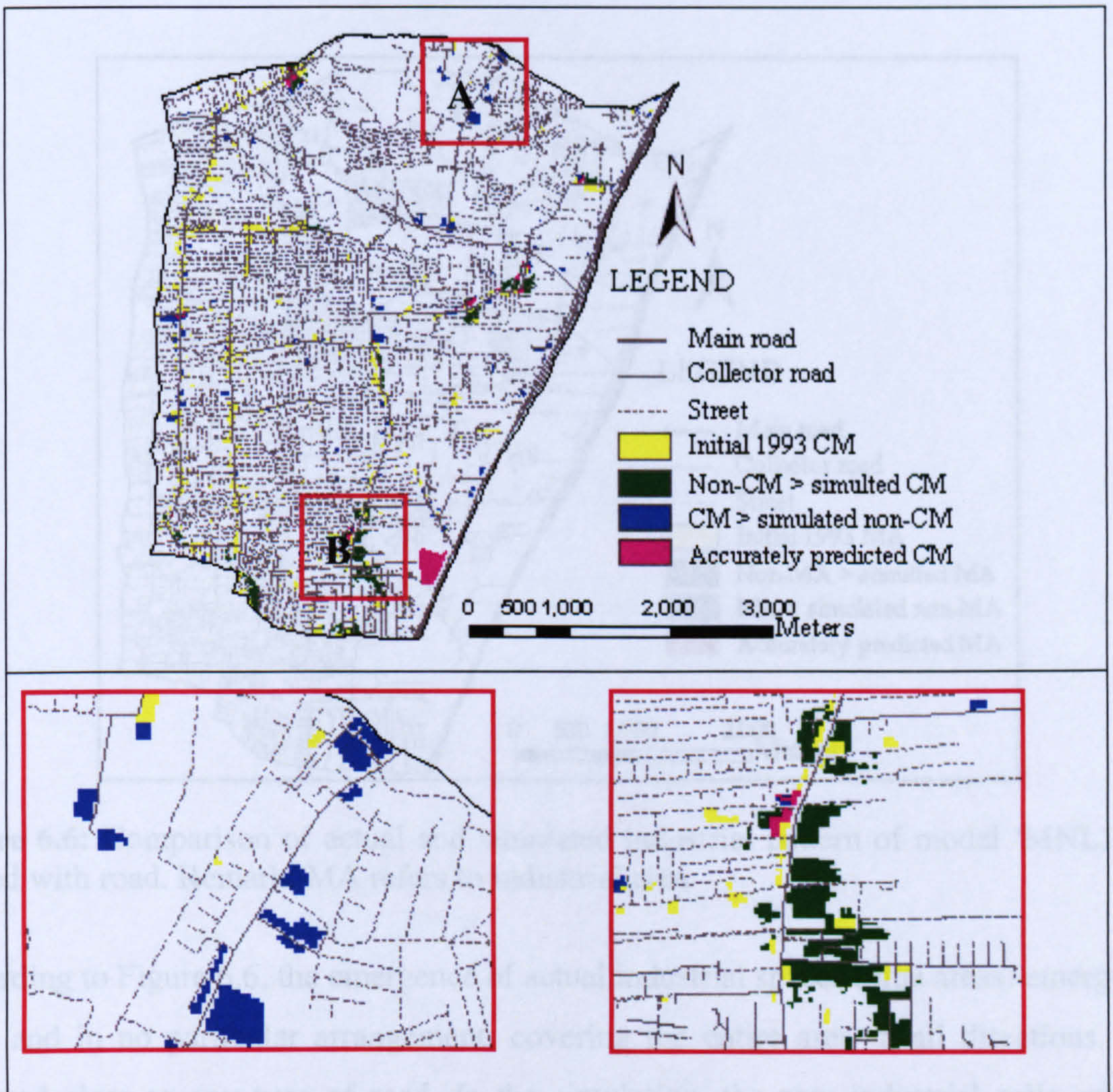
Another technique, visual observation through GIS overlay, was used to compare the spatial match of location. Figures 6.4, 6.5 and 6.6 show the spatial match of residential, commercial and industrial locations between the actual and simulated 2001 respectively. For each land use category, the symbols are classified as (i) the initial 2001 area, (ii) the commission error and (iii) the omission error, and (iv) the accurately predicted area. Having considered the relative weights (see Table 6.2), the most influential development factors were the effect of roads. Thus, these three visual assessment maps were draped with road features, in order to investigate the effect of roads in relation to the emergence of the three targeted land use types.



(a) The magnified area A (b) The magnified area B

Figure 6.4: Comparison of actual and simulated residential pattern of model 'MNL1-V1', draped with roads. Remark: RS refers to residential area.

According to Figure 6.4, the emergence of residential spaces (blue areas) of the actual 2001 which the model could not produce (omission error) mostly emerged in a fragmentation pattern out of the inner core of the district in all directions except the north-west. They were all usually dispersed along the local streets (see the magnified area in Figure 6.4 (a)). In the simulated results, new residential locations (green and pink areas) usually emerged in the area where they were both close to local streets and next to the previous residential cells (see the magnified area in Figure 6.4 (b)) such as in the inner core of the district. The green areas represent another type of the mismatch locations (commission error), here referred to actual non-residential but assigned as residential use. The pink area represents the accurately predicted locations. Based on the comparison of results shown in this figure, it can be concluded that the model could do best for the emergence of residential cells that occurred in the location where they were both close to local streets and next to previous residential cells. The commission and omission errors that occurred are probably because the model cannot produce a fragmentation pattern.



(a) The magnified area A (a) The magnified area B

Figure 6.5: Comparison of actual and simulated commercial pattern of model 'MNL1-V1', draped with roads. Remark: CM refers to commercial area.

According to Figure 6.5, the emergence of actual commercial spaces (blue areas) which the model could not produce (omission error) emerged randomly covering the entire area in all directions. They all occurred close to main and collector streets (see the magnified area in

Figure 6.5 (a)). In the simulation, the new commercial cells emerged largely in the south of the study area, appearing as an enlarging pattern next to the existing commercial cells (see the magnified area in Figure 6.5 (b)). Despite the fact that the model was successful in producing the new development areas that were close to main and collector streets, however, they rarely emerged in the same locations of the actual commercial 2001. Similar to the residential results discussed above, this is probably because of the inability of the model to produce a fragmentation pattern. It should be noted that the large pink area at the lower left of the study site was a new big department store. It was an event updated due to Scenario 3, illustrated in Table 6.1.

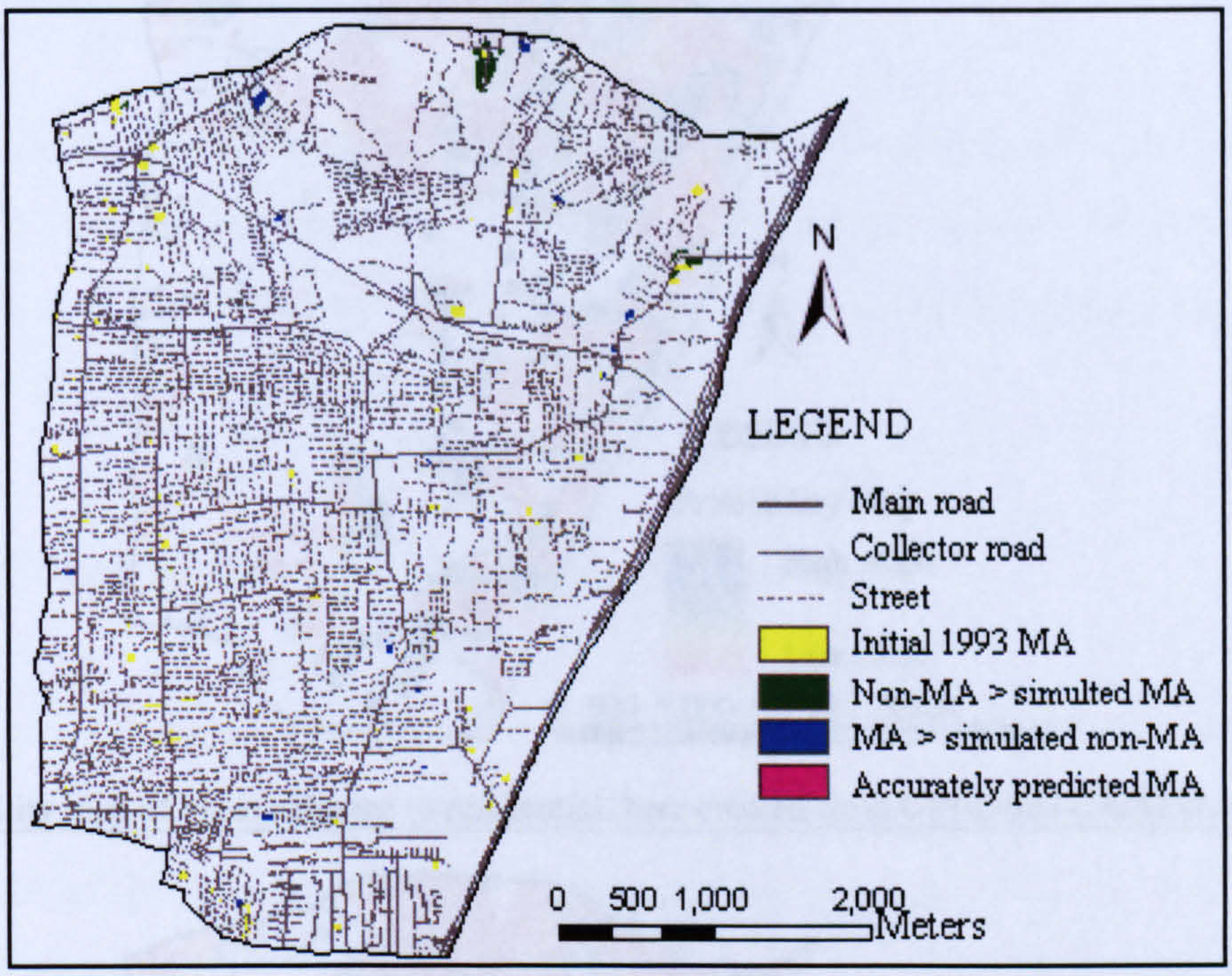
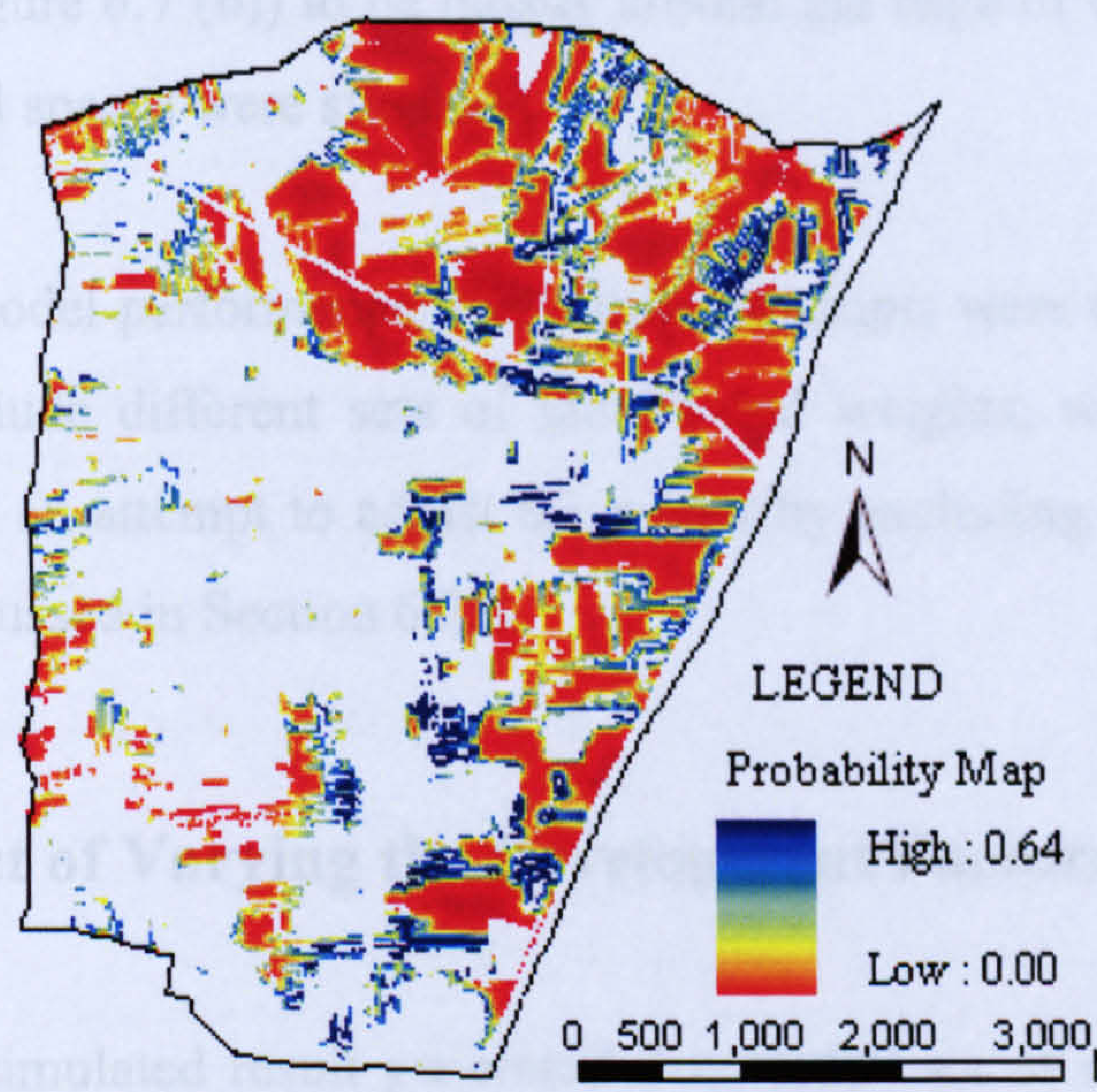


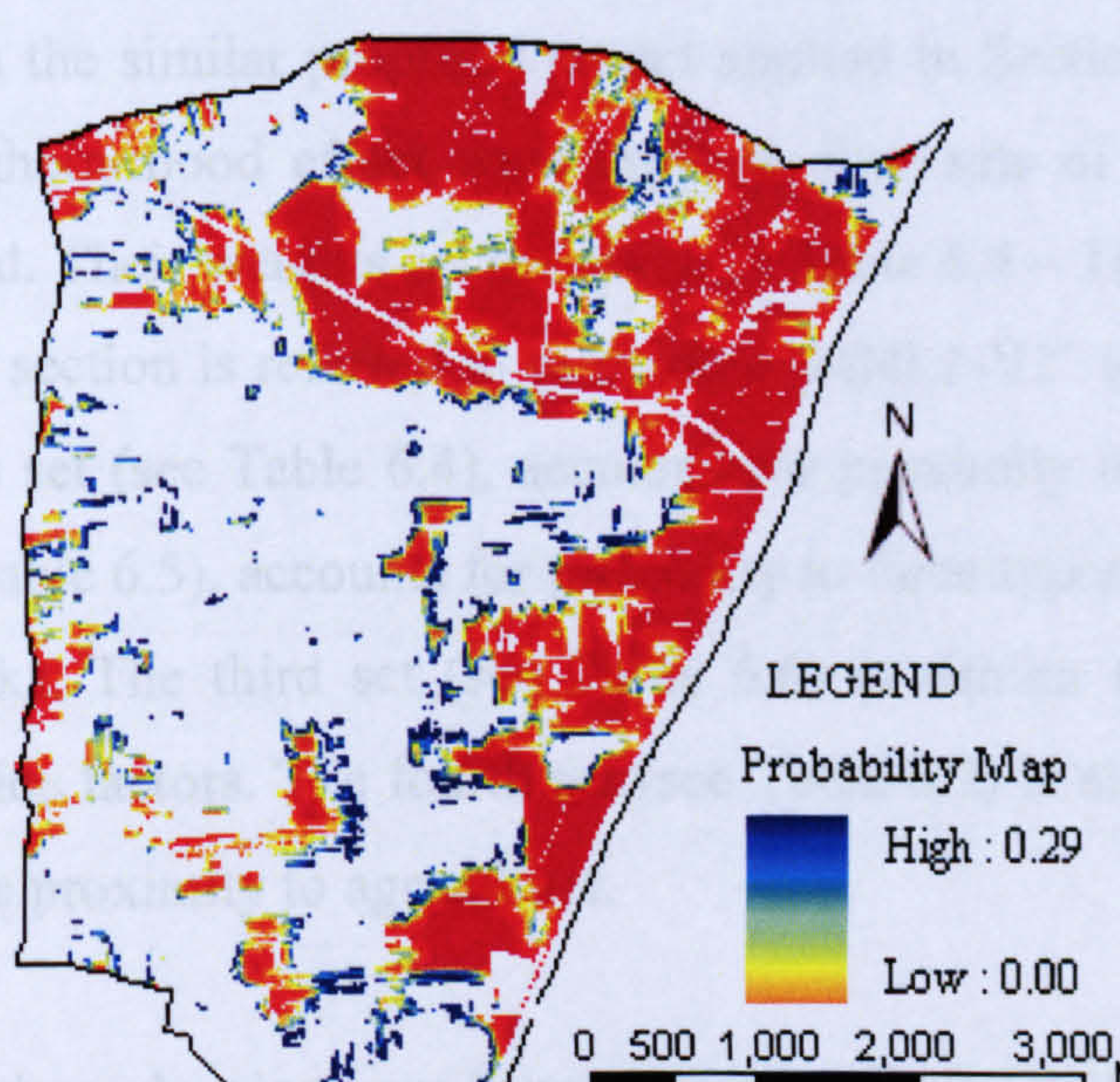
Figure 6.6: Comparison of actual and simulated industrial pattern of model 'MNL1-V1', draped with road. Remark: MA refers to industrial area.

According to Figure 6.6, the emergence of actual industrial spaces (blue areas) emerged far apart and in no particular arrangement, covering the entire area in all directions. They occurred close to any type of road. In the simulation, the new industrial cells emerged clustering in one area (green area), appearing as an enlarging pattern. No spatial match location was found for this land transition type. This is probably due to two reasons. Firstly, the development factors and their weights used for the simulation could not characterize the distinctive characteristics of the industrial pattern of the study site (minor industrial units within the urban fabric). Secondly, similar to residential and commercial cases, the model could not produce the emergence of new industrial spaces independently, but enlarged the existing industrial cells only.

The poor model performance described above suggests that the MNL1-V1 model with a set of development factors used in this section, is unrealistic as it poorly captures or simulates true urban growth that has occurred. This finding suggests that the proposed model that takes into account the neighbourhood effect has a relatively high influence on the probability values. The supportive evidence is illustrated by Figure 6.7.



(a) Probability map of vacant change to residential, here created from GIS-based CA/MNL model



(b) A Final Probability map of vacant change to residential, here created from GIS-based CA/MNL model multiplied with neighbourhood effect

Figure 6.7: The probability maps generated during an iteration (a) before applying dynamic neighbourhood effect and (b) after applying dynamic neighbourhood effect.

Figure 6.7 shows the residential probability map created based on the developed model at the last iteration of the simulation. In Figure 6.7 (a) in which the probability map is solely created from the MNL method, the high probability land (blue colour in the figure) tends to be scattered all over the entire vacant area because of the sole effects of the MNL-based probability computation, as described in Equation 4.1. However, when the MNL probability map was incorporated with the residential neighbourhood effect as addressed in Equation 4.9, the original probability values were thus multiplied by the residential neighbourhood index value. As a result, it leads to the shift of the higher probability values (blue colour in Figure 6.7 (b)) to be mostly around the edge of the vacant areas where the existing residential spaces were situated.

To improve the model performance, two further attempts were carried out. They included an attempt to include different sets of factors and weights, which will be discussed in Section 6.2.2, and an attempt to adjust the model by excluding the neighbourhood effect, which will be discussed in Section 6.2.3.

6.2.2 The Effect of Varying the Development Factors

Due to the poor simulated result generated from a full set of development factors in the previous section, this section makes further tests using different sets of development factors. Based on the similar proposed model applied in Section 6.2.1 in which the MNL method and neighbourhood effect were applied, four sets of development factors were additionally tested. Their weights set are given in Table 6.4 – Table 6.7. Implementation of the model in this section is referred to as models 'MNL1-V2' to 'MNL1-V5' respectively. The first weights set (see Table 6.4), accounts for proximity to three types of roads. The second set (see Table 6.5), accounts for proximity to three types of roads, government area, schools, and parks. The third set (see Table 6.6) is similar to the second set but with inclusive land price factors. The fourth set (see Table 6.7) is also similar to the second set but with inclusive proximity to agriculture.

Interpretation of these development factors was done in a similar way to the weights set in Section 6.2.1. Overall, each set of development factors had different sets of weights due to different combinations of factors included. However, these sets of weights corresponded in the sense that proximity to local streets was the most influential factor for the vacant change to residential while proximity to collector streets was the most influential factor for the vacant change to commercial and industrial.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Proximity to Main roads (D_RDT1)	0.108	0.231	0.289
Proximity to Collector Streets (D_RDT2)	6.687	31.251	12.507
Proximity to Local Streets (D_RDT3)	62.628	10.265	2.278
Intercept	-69.153	-42.288	-19.090
-2LL	59444.723 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.4: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL1-V2'.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Proximity to Main roads (D_RDT1)	-0.700	-0.932	2.036
Proximity to Collector Streets (D_RDT2)	7.009	31.591	13.530
Proximity to Local Streets (D_RDT3)	61.044	10.580	4.897
Proximity to Government (D_GV)	-0.590	2.027	-3.121
Proximity to School (D_SCH)	2.370	0.521	2.233
Proximity to Park (D_PRK)	-0.101	-2.349	-3.902
Intercept	-68.417	-42.099	-20.447
-2LL	58687.202 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.5: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL1-V3'.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Proximity to Main roads (D_RDT1)	-0.758	-1.171	1.793
Proximity to Collector Streets (D_RDT2)	6.940	28.053	12.018
Proximity to Local Streets (D_RDT3)	52.456	1.144	-7.429
Proximity to Government (D_GV)	-0.590	2.256	-3.076
Proximity to School (D_SCH)	2.401	0.476	2.377
Proximity to Park (D_PRK)	-0.204	-2.720	-4.333
Land price (LP)	0.370	0.983	0.952
Intercept	-63.343	-38.580	-15.688
-2LL	58201.022 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.
2. -2LL = -2 log likelihood at convergence.

Table 6.6: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL1-V4'.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Proximity to Main roads (D_RDT1)	-0.635	-0.804	4.645
Proximity to Collector Streets (D_RDT2)	6.876	31.600	14.984
Proximity to Local Streets (D_RDT3)	61.151	10.988	-0.0.056
Proximity to Government (D_GV)	-0.657	1.876	-4.593
Proximity to School (D_SCH)	2.348	0.558	0.741
Proximity to Park (D_PRK)	-0.275	-3.033	1.243
Proximity to Agriculture (D_AGR)	0.228	0.530	-7.424
Intercept	-68.729	-42.293	-17.671
-2LL	58415.996 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.
2. -2LL = -2 log likelihood at convergence.

Table 6.7: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL1-V5'.

Model performance was assessed through an accuracy table using the same technique as that described in Section 6.2.1. Tables 6.8 – 6.11 show the accuracy tables generated for model 'MNL1-V2' to 'MNL1-V5' respectively.

Category (no. of cells)	Residential (RS)	Commercial (CM)	Industrial (MA)	Vacant (VC)	Total cells	Producer's accuracy
Simulated Residential	6284	546	57	11384	18271	34.39
Simulated Commercial	716	603	75	1421	2815	21.42
Simulated Industrial	83	51	0	270	404	0.000
Simulated Vacant	11159	1602	217	66486		
Total cells	18242	2802	349	79561		
User's Accuracy	34.45	21.52	0	83.57		

Total accuracy: 32.05%

Table 6.8: Evaluation of the simulated results of year 2001, based on model 'MNL1-V2'.

Category (no. of cells)	Residential (RS)	Commercial (CM)	Industrial (MA)	Vacant (VC)	Total cells	Producer's accuracy
Simulated Residential	6176	507	56	11499	18238	33.86
Simulated Commercial	946	552	83	1229	2810	19.64
Simulated Industrial	85	2	0	265	352	0.00
Simulated Vacant	11035	1741	210	66568		
Total cells	18242	2802	349	79561		
User's Accuracy	33.86	19.70	0	83.67		

Total accuracy: 31.44%

Table 6.9: Evaluation of the simulated results of year 2001, based on model 'MNL1-V3'.

Table 6.12 summarizes the accuracy assessed from different sets of development factors. To conclude, the model, in spite of excluding some development factors, produced a correspondence accuracy rate largely similar to that of a full set of parameters ('MNL1-V1', described in Section 6.2.1). In comparison, the total accuracy of different factors is in the range of 31.4 – 32.1. Amongst these, they all correspond in the sense that residential accuracy has the highest accuracy, followed by commercial and industrial respectively. However, the model with the proximity to roads factors only ('MNL1-V2') tends to generate slightly better accuracy value than the others. Overall, it can be concluded that different factors did not significantly improve the model performance.

Category (no. of cells)	*Residential (RS)*	*Commercial (CM)*	*Industrial (MA)*	*Vacant (VC)*	*Total cells*	*Producer's accuracy*
Simulated Residential	6203	523	65	11417	18208	34.07
Simulated Commercial	1062	580	76	1086	2804	20.69
Simulated Industrial	110	8	2	238	358	0.56
Simulated Vacant	10867	1691	206	66820		
Total cells	18242	2802	349	79561		
User's Accuracy	34.00	20.70	0.57	83.99		

Total accuracy: 31.75%

Table 6.10: Evaluation of the simulated results of year 2001, based on model 'MNL1-V4'.

Category (no. of cells)	*Residential (RS)*	*Commercial (CM)*	*Industrial (MA)*	*Vacant (VC)*	*Total cells*	*Producer's accuracy*
Simulated Residential	6248	498	57	11446	18249	34.24
Simulated Commercial	969	469	83	1293	2814	16.67
Simulated Industrial	0	18	0	343	361	0.000
Simulated Vacant	11025	1817	209	66479		
Total cells	18242	2802	349	79561		
User's Accuracy	34.25	16.74	0	83.56		

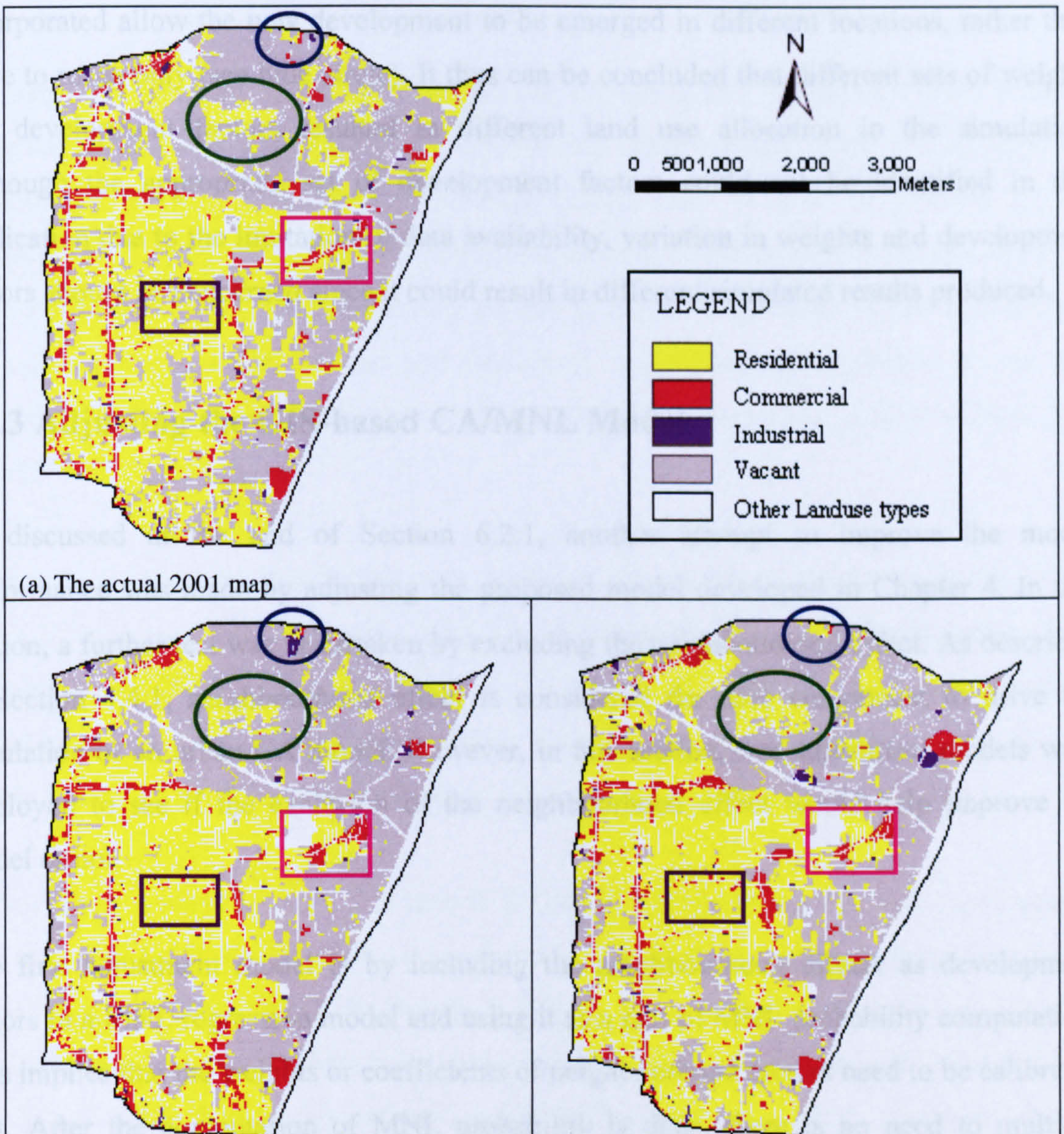
Total accuracy: 31.35%

Table 6.11: Evaluation of the simulated results of year 2001, based on model 'MNL1-V5'.

	Accuracy (%)			
Model Name	**Residential**	**Commercial**	**Industrial**	**Total**
MNL1 – V1	34.37	18.08	0.00	31.59
MNL1 – V2	34.39	21.42	0.00	32.05
MNL1 – V3	33.86	19.64	0.00	31.44
MNL1 – V4	34.07	20.69	0.56	31.75
MNL1 – V5	34.24	16.67	0.00	31.35

Table 6.12: Comparison of simulated results based on the GIS-based CA/MNL model with different set of development factors.

Further analysis was conducted through visual comparison between the 2001 actual and simulated land use map. Figure 6.8 gives an example of the visual comparison based on a full set of development factors ('MNL1-V1') described in Section 6.2.1 and the simulated map based on the inclusive proximity to three types of roads ('MNL1-V2').



(a) The actual 2001 map

(b) The simulated 2001 of model MNL1 – V1

(c) The simulated 2001 of model MNL 1 – V2

Figure 6.8: Visual comparison of the 2001 actual and simulated land use maps, model 'MNL1-V1' and 'MNL1-V2'.

Two sample locations are highlighted in coloured circles. A similar distinction between these two models can be captured in terms of morphology. It is that the development of the simulated results emerged in the pattern of space-filling (highlighted in purple box) and enlarging the existing development (highlighted in pink box). In spite of the similarities, they differed largely in terms of land development locations (highlighted in blue and green circles). The emergence of a new development of model 'MNL1-V2' occurred in the locations that were both close to roads and adjacent to the existing development (see green circle). In the model 'MNL1-V1', more development factors were incorporated (e.g. proximity to roads, land use activities). Although proximity to roads factors have higher weights when compared to other factors in the model (see Table 6.2), many factors incorporated allow the new development to be emerged in different locations, rather than close to roads only (see blue circle). It thus can be concluded that different sets of weights and development factors resulted in different land use allocation in the simulation. Although the appropriate set of development factors could not be identified in this application due to the limitation of data availability, variation in weights and development factors should be concerned since it could result in different simulated results produced.

6.2.3 Adjusting the GIS-based CA/MNL Model

As discussed at the end of Section 6.2.1, another attempt to improve the model performance was made by adjusting the proposed model developed in Chapter 4. In this section, a further test was undertaken by excluding the neighbourhood effect. As described in Section 4.4.1, neighbourhood effect is considered the core component to drive the simulation of an urban CA model. However, in this section, two adjustment models were employed to see if the exclusion of the neighbourhood effect would help improve the model or not.

The first adjustment model is by including the neighbourhood effects as development factors in a MNL regression model and using it as a part of MNL probability computation. This implies that the weights or coefficients of neighbourhood factors need to be calibrated first. After the computation of MNL probability is done, there is no need to multiply neighbourhood effects thereafter. By implicitly including the neighbourhood factors in the equation, it is expected that this, to some degree, will reduce the effect of neighbourhood. In the remainder of this section, the adjustment model developed here is referred to as 'MNL2'.

It should be noted that the adjustment model 'MNL2' was previously criticized by Wu (2002a) because it can produce unrealistic dynamic simulation since the weights of these neighbourhood effects are considered static and aggregated. Although the neighbourhood characteristics of the area considered change at the end of each iteration, the set of weights of neighbourhood factors do not change over time. He further suggests that this technique can be applied if data about the land use conversion can be observable at the fine time scale (e.g. annually). However, it is still rare to get such fine scale data in the real world situation.

The second adjustment model is by completely excluding neighbourhood effects from the proposed model. It is observed by Wu (2002a) that when simulating the growth, land use modelling that uses the global development factors only is likely to produce a more scattering land development pattern than the model that includes the neighbourhood effect. In this case in which more fragmentary pattern is required, this implies that exclusion of the neighbourhood effect in land use modelling is probably giving better results than including it in the model. Thus, this adjustment model is performed, being referred to as 'MNL3' in this section.

Implementation of the two adjustment models outlined above, so-called 'MNL2' and 'MNL3', is already provided by the MNL GUI developed in Chapter 5. To implement both adjustment models, users are required to choose "NOT APPLY' from the 'Dynamic neighbourhood' section in the main MNL GUI (Figure 5.11), in order to discard the multiplication of neighbourhood effects.

According to model 'MNL2', five sets of development factors were tested. Their weights sets are given as shown in Table 6.13 – Table 6.17 respectively. Clearly these weight sets are similar to those shown in Tables 6.2 and 6.4 – 6.7, but with the addition of neighbourhood factors. Implementation of the model is referred to as model 'MNL2-V1' to 'MNL2-V5', The first weight set (see Table 6.13) comprises coefficients for residential neighbourhood effect, commercial neighbourhood effect, industrial neighbourhood effect, proximity to main roads, proximity to collector streets, proximity to local streets, proximity to government area, proximity to school, proximity to park / recreation area, land price, and proximity to agricultural area. The second set (see Table 6.14) accounts for the neighbourhood effect of three land use types and proximity to three types of roads. The third set (see Table 6.15) accounts for the neighbourhood effect of three land use types, proximity to three types of roads, government area, schools, and parks. The fourth set (see Table 6.16), namely MNL2-V3, is similar to the third set but with inclusive land price factors. The fifth set (see Table 6.17) is similar to the third set but with inclusive proximity to agriculture.

Factors	**Vacant to Residential Coefficient (β)**	**Vacant to Commercial Coefficient (β)**	**Vacant to Industrial Coefficient (β)**
Residential Neighbourhood effect (NR_RS)	.076	.322	.960
Commercial Neighbourhood effect (NR_CM)	.055	13.630	1.884
Industrial Neighbourhood effect (NR_MA)	.041	-12.464	-26.068
Proximity to Main roads (D_RDT1)	.354	.652	-3.275
Proximity to Collector Streets (D_RDT2)	5.595	-1.020	-2.516
Proximity to Local Streets (D_RDT3)	35.250	-1.579	.844
Proximity to Government (D_GV)	.766	1.135	.973
Proximity to School (D_SCH)	-.398	-.703	.816
Proximity to Park (D_PRK)	-1.509	.035	.023
Land price (LP)	.284	.176	.118
Proximity to Agriculture (D_AGR)	-1.227	.076	.457
Intercept	-45.254	-15.319	8.205
-2LL	42915.748 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.13: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL2-V1'.

Factors	**Vacant to Residential Coefficient (β)**	**Vacant to Commercial Coefficient (β)**	**Vacant to Industrial Coefficient (β)**
Residential Neighbourhood effect (NR_RS)	0.068	0.028	0.006
Commercial Neighbourhood effect (NR_CM)	0.053	0.167	0.092
Industrial Neighbourhood effect (NR_MA)	0.039	0.081	0.421
Proximity to Main roads (D_RDT1)	0.202	0.082	-2.039
Proximity to Collector Streets (D_RDT2)	4.807	15.805	4.495
Proximity to Local Streets (D_RDT3)	42.179	-1.374	-6.927
Intercept	-49.752	-18.682	-3.413
-2LL	44063.677 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.14: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL2-V2'.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Residential Neighbourhood effect (NR_RS)	0.074	0.033	0.022
Commercial Neighbourhood effect (NR_CM)	0.052	0.171	0.111
Industrial Neighbourhood effect (NR_MA)	0.056	0.091	0.450
Proximity to Main roads (D_RDT1)	0.782	0.828	1.771
Proximity to Collector Streets (D_RDT2)	5.139	15.533	4.068
Proximity to Local Streets (D_RDT3)	42.456	-0.891	-13.529
Proximity to Government (D_GV)	0.246	0.310	-2.830
Proximity to School (D_SCH)	-0.350	-1.050	-3.260
Proximity to Park (D_PRK)	-2.383	-2.036	1.493
Intercept	-49.072	-17.619	2.843
-2LL	43572.941 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.15: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL2-V3'.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Residential Neighbourhood effect (NR_RS)	0.069	0.024	0.006
Commercial Neighbourhood effect (NR_CM)	0.002	0.040	0.019
Industrial Neighbourhood effect (NR_MA)	-0.006	-0.043	-0.001
Proximity to Main roads (D_RDT1)	0.317	-0.946	1.816
Proximity to Collector Streets (D_RDT2)	6.894	27.687	12.199
Proximity to Local Streets (D_RDT3)	45.713	-5.466	-8.267
Proximity to Government (D_GV)	0.541	2.465	-3.023
Proximity to School (D_SCH)	0.130	-0.120	2.107
Proximity to Park (D_PRK)	-2.255	-3.140	-4.713
Land price (LP)	0.018	0.977	0.923
Intercept	-53.862	-32.161	-14.735
-2LL	48154.551 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.

2. -2LL = -2 log likelihood at convergence.

Table 6.16: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL2-V4'.

Factors	Vacant to Residential Coefficient (β)	Vacant to Commercial Coefficient (β)	Vacant to Industrial Coefficient (β)
Residential Neighbourhood effect (NR_RS)	0.076	0.034	0.024
Commercial Neighbourhood effect (NR_CM)	0.054	0.171	0.115
Industrial Neighbourhood effect (NR_MA)	0.043	0.081	0.461
Proximity to Main roads (D_RDT1)	0.369	0.629	1.578
Proximity to Collector Streets (D_RDT2)	5.739	15.945	3.483
Proximity to Local Streets (D_RDT3)	41.485	-1.304	-13.844
Proximity to Government (D_GV)	0.639	0.600	-2.821
Proximity to School (D_SCH)	-0.160	-1.013	-3.297
Proximity to Park (D_PRK)	-1.531	-1.525	1.244
Proximity to Agriculture (D_AGR)	-1.249	-0.701	0.588
Intercept	-48.842	-17.732	3.655
-2LL	43298.974 (significance at 0.00)		

Note: 1. Vacant to vacant transition (no change) as baseline category.
2. -2LL = -2 log likelihood at convergence.

Table 6.17: Weights calibrated from multinomial logistic regression (MNL) method, used for model 'MNL2-V5'.

According to model 'MNL3', five sets of development factors were tested. Implementation of the model is referred to as model 'MNL3-V1' to 'MNL3-V5'. They used exactly the same development factors and weights sets as listed in Table 6.2, and Tables 6.4 – 6.7 respectively.

It should be noted that similar weight sets to those of model 'MNL1' were applied to both model 'MNL2' and 'MNL3' (with the addition of neighbourhood factors) in order that the simulated results could be comparable.

The model performance was generated through the accuracy table using the same technique as that described in Section 6.2.1. Table 6.18 summarizes the accuracy assessed from different set of development factors for both 'MNL2' and 'MNL3'. Overall, these two adjustment models give a nearly similar correspondence rate, when compared to those of model 'MNL1' (see Table 6.12). In comparison, the total accuracy of 'MNL1' (with different factors) is in the range of 31.4 – 32.1, MNL2 in the range of 29.8 – 31.2, MNL3 in the range of 31.3 – 32.0. They all correspond in the sense that residential accuracy has a highest accuracy rate, followed by commercial and industrial respectively. All in all, it can be concluded that the adjustment model did not significantly improve the spatial agreement.

Model Name	Accuracy (%)			
	Residential	Commercial	Industrial	Total
MNL2 – V1	33.37	15.31	0.00	30.46
MNL2 – V2	34.42	14.39	0.00	31.24
MNL2 - V3	34.01	11.75	0.00	30.54
MNL2 – V4	32.84	19.06	0.00	30.50
MNL2 – V5	33.16	11.50	0.00	29.78
MNL3 – V1	34.93	11.35	0.57	31.28
MNL3 – V2	33.31	23.43	0.94	31.38
MNL3 – V3	35.02	16.28	0.00	32.00
MNL3 – V4	34.92	16.51	0.00	31.94
MNL3 – V5	35.17	10.18	0.00	31.33

Table 6.18: Comparison of simulated results based on the adjusted GIS-based CA/MNL model.

Visual comparison was conducted in order to observe the urban morphology and spatial structure. Figure 6.9 illustrates the visual comparison between the three models ('MNL1', 'MNL2' and 'MNL3') that used the similar sets of development factors in Tables 6.6 and 6.16, here referring to the inclusive proximity to main roads, proximity to collector streets, proximity to local streets, proximity to government area, proximity to school, proximity to park / recreation area and land price. Clearly, they differed in terms of land development locations.

Three sample locations are highlighted by coloured circles in order to investigate the morphology. The first location, highlighted with a green circle, shows that the degree of space-filling pattern was high in both 'MNL1-V3' and 'MNL2-V3', when compared to the actual 2001 land use map. The second location and the third locations, highlighted with blue and purple circles, show that the degree of enlarging spaces around the existing development was high in both 'MNL1-V3' and 'MNL2-V3', when compared to the actual 2001 land use map. However, it seems that the enlarging development pattern of 'MNL2-V3' (the model including neighbourhood as a part of MNL probability computation) tends to be more compact than that of 'MNL1-V3'. Amongst the comparison of the three models, 'MNL3-V3' tends to be the least in terms of space-filling pattern and enlarging pattern.

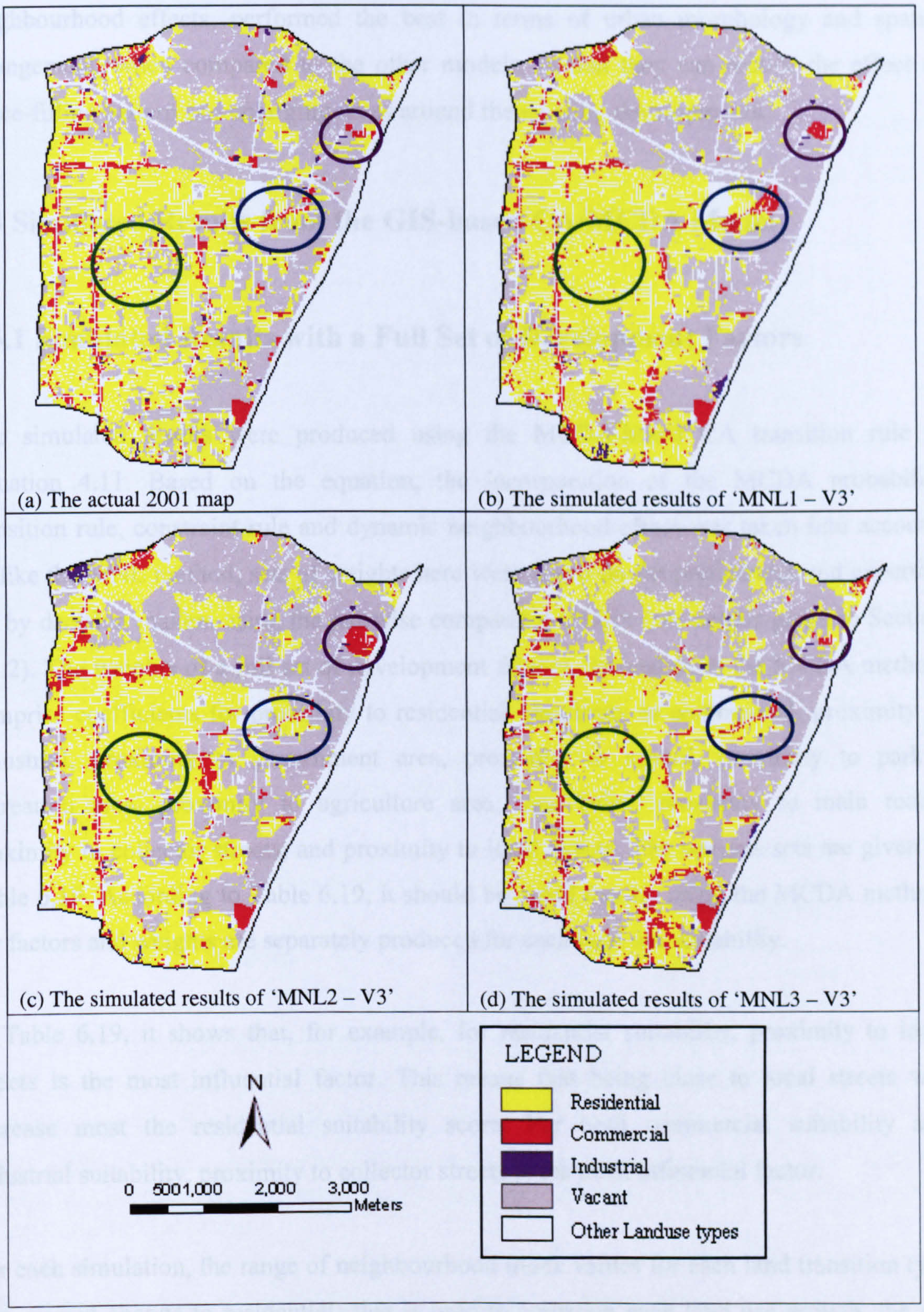


Figure 6.9: Visual comparison of the 2001 actual and simulated land use map (model 'MNL1-V3', 'MNL2-V3' and 'MNL3-V3').

Based on the simulated results examined in this figure (Figure 6.9), it is concluded that the simulated results, despite similar factors, differed in land development locations. In addition, it is likely that the mode based on 'MNL3', which completely excludes the neighbourhood effects, performed the best in terms of urban morphology and spatial arrangement, when compared to the other models, in that they can reduce the effect of space-filling pattern and enlarging areas around the existing developments.

6.3 Simulated Results from the GIS-based CA/MCDA Model

6.3.1 Simulated Results with a Full Set of Development Factors

The simulated results were produced using the MCDA-based CA transition rule in Equation 4.11. Based on the equation, the incorporation of the MCDA probability transition rule, constraint rule and dynamic neighbourhood effect was taken into account. Unlike the MNL method, sets of weights here were based on the preferences and expertise set by decision makers using the pairwise comparison matrix (see details given in Section 4.5.2). The weights of a full set of development factors, derived from the MCDA method, comprise coefficients for proximity to residential, proximity to commercial, proximity to industrial, proximity to government area, proximity to school, proximity to park / recreation area, proximity to agriculture area, land price, proximity to main roads, proximity to collector streets, and proximity to local streets. The weights sets are given in Table 6.19. According to Table 6.19, it should be noted that by using the MCDA method, the factors and weights are separately produced for each land use suitability.

In Table 6.19, it shows that, for example, for residential suitability, proximity to local streets is the most influential factor. This means that being close to local streets will increase most the residential suitability score. For both commercial suitability and industrial suitability, proximity to collector streets is the most influential factor.

For each simulation, the range of neighbourhood index values for each land transition type (e.g. vacant change to residential) that is used to constrain each land use growth, the so-called neighbourhood threshold, is given in Table 6.20. These thresholded neighbourhood index values can be set by the preferences of decision makers. In this case, the statistical observation in the study area, the compatibility of land use activities (ESRI (Thailand) Co. Ltd., 1997) set by the Department of Town and Country Planning (DTCP), Thailand, and an initial interview with staff in the Department of City Planning (DCP), Bangkok, was held to obtain some initial values for testing in the model.

	Residential suitability		*Commercial suitability*		*Industrial suitability*	
Criteria	***variables***	***Weight***	***variables***	***Weight***	***variables***	***Weight***
Proximity to residential	C_RS	**0.170**	C_RS	**0.155**	F_RS	**0.027**
Proximity to commercial	C_CM	**0.122**	C_CM	**0.213**	F_CM	**0.021**
Proximity to industrial	F_MA	**0.017**	F_MA	**0.034**	C_MA	**0.236**
Proximity to government	C_GV	**0.048**	C_GV	**0.090**	F_GV	**0.069**
Proximity to school	C_SCH	**0.079**	-	-	F_SCH	**0.060**
Proximity to park/ conservation	C_PRK	**0.038**	-	-	-	-
Proximity to agriculture	C_AGR	**0.072**	-	-	-	-
Land price	LP	**0.022**	LP	**0.053**	LP	**0.094**
Proximity to main roads	C_RDT1	**0.081**	C_RDT1	**0.216**	C_RDT1	**0.227**
Proximity to collector streets	C_RDT2	**0.150**	C_RDT2	**0.240**	C_RDT2	**0.267**
Proximity to local streets	C_RDT3	**0.201**	-	-	-	-
Consistency ratio	0.098		0.066		0.051	

Table 6.19: Weights or coefficients calibrated from multi-attribute decision analysis (MADA) method. Remark: a variable starting with "C" refers to "Close to" (e.g. C_SCH means "Close to schools) and a variable starting with "F" refers to "Far from" (e.g. F_SCH means "Far from schools").

Variables	**Threshold set for land use transition**		
	Vacant change to residential	**Vacant change to commercial**	**Vacant change to industrial**
Residential neighbourhood index (NI_RS)	0.03 – 0.63	N/A	N/A
Commercial neighbourhood index (NI_CM)	0.01 – 0.11	0.05 – 0.48	0.01 – 0.16
industrial neighbourhood index (NI_MA)	N/A	N/A	N/A

Table 6.20: A list of neighbourhood threshold (or range of neighbourhood index value that is used to constrain development) for the targeted land use types, regarding residential, commercial and industrial.

The threshold values in Table 6.20 can be interpreted, for example, for vacant change to residential: the potential cells for development depended on either the threshold (minimum and maximum) of residential neighbourhood index value in the range of 0.03 – 0.63 or the threshold (minimum and maximum) of commercial neighbourhood index value in the range of 0.01 – 0.11. In this regard, the 'N/A' means the threshold of industrial neighbourhood index value was not taken into account. More details about neighbourhood threshold have been previously given in Section 4.5.3.

Implementation of the model generated in this section is referred to as "MCDA-W21". Its model name refers to the model conducted by the MCDA method that uses the window size "W" of 21. Based on the 10m grid cell size applied in this study, this refers to the neighbourhood characteristics of 100 metres within walking distance.

Figure 6.10 illustrates the simulated results of the 1995, 1997, 1999 and 2001 land use maps. The visual observation of these simulated results shows a pattern similar to those created by the GIS-based CA/MNL model described in Sections 6.2.1 and 6.2.2 in that the urban development produces enlarging areas around the previous developments, and a space-filling pattern.

Further investigation through the visual comparison between the 2001 actual and simulated land use maps confirms this distinct growth pattern, similar to those generated from the GIS-based CA/MNL model, but different from reality. The simulated 2001 land use map illustrated in Figure 6.11 shows the enlarging clusters around the existing development. In Figure 6.11(b), the actual emergence of residential and commercial land uses is scattered independently along the roads. In Figure 6.11(c), the equivalent simulated residential and commercial spaces enlarged the existing ones only. In addition, their emergence acts in many cases as a space-filler. In Figure 6.11(d) the actual residential area of 2001 occurs in a fragmentation pattern while Figure 6.11(e) highlights the same location of simulated results where a space-filling pattern occurs.

According to the accuracy table (Table 6.21), overall the developed model performed poorly with a total accuracy of 32.01% being accomplished. The accuracy value compared to that of the best model acquired by the proposed model with the GIS-based CA/MNL model ('MNL1-V2' with the accuracy of 32.05%) is almost equal. By excluding the consideration of vacant type, the simulated residential type shows the highest user's accuracy and producer's accuracy of 34.90% and 35.04% respectively. In this case, the model assigned 65.10% of residential use as being non-residential (4.47% being commercial, 0.29% being industrial area and 60.34% being vacant).

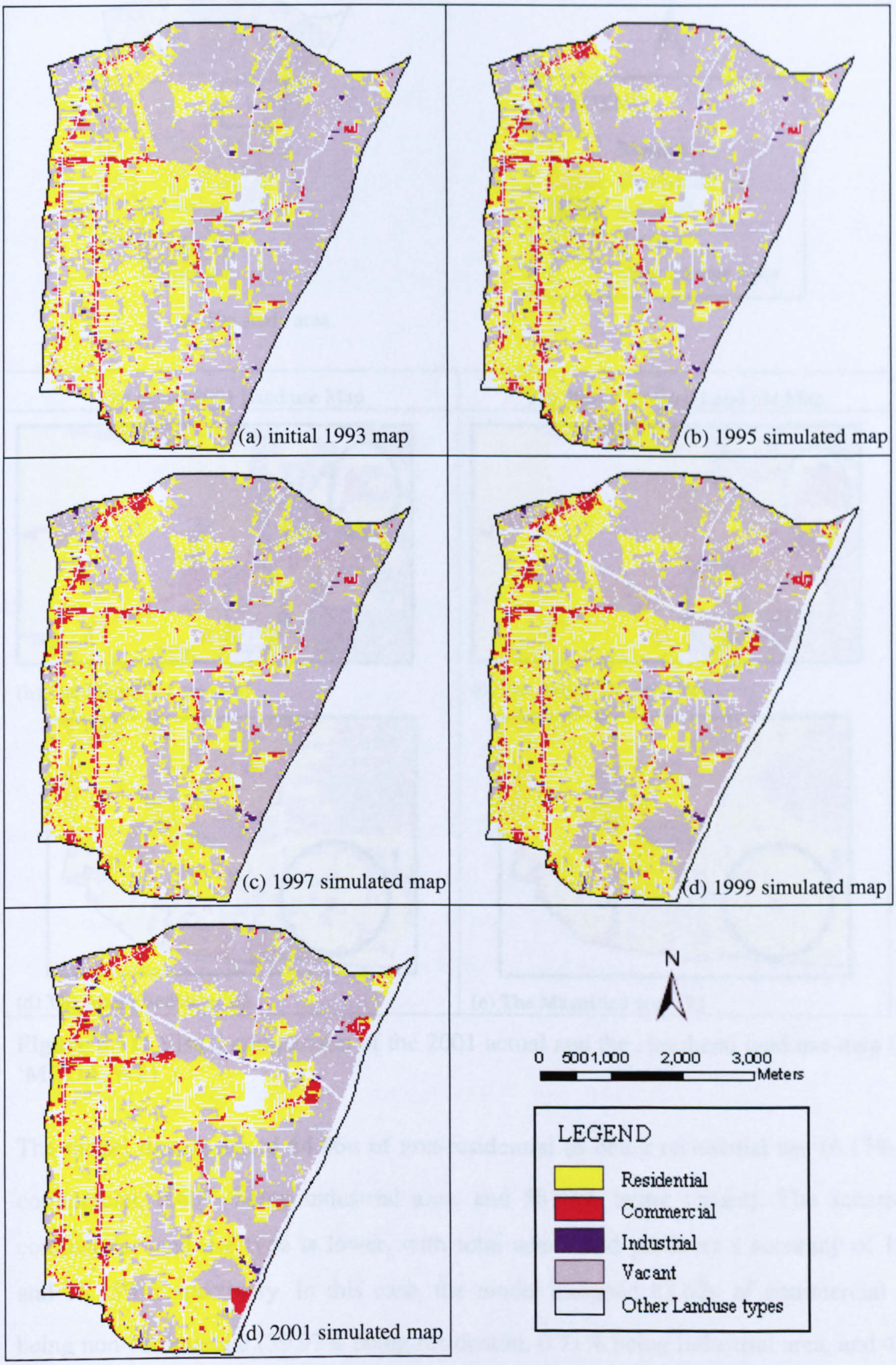


Figure 6.10: The simulated results of model 'MCDA-W21'.

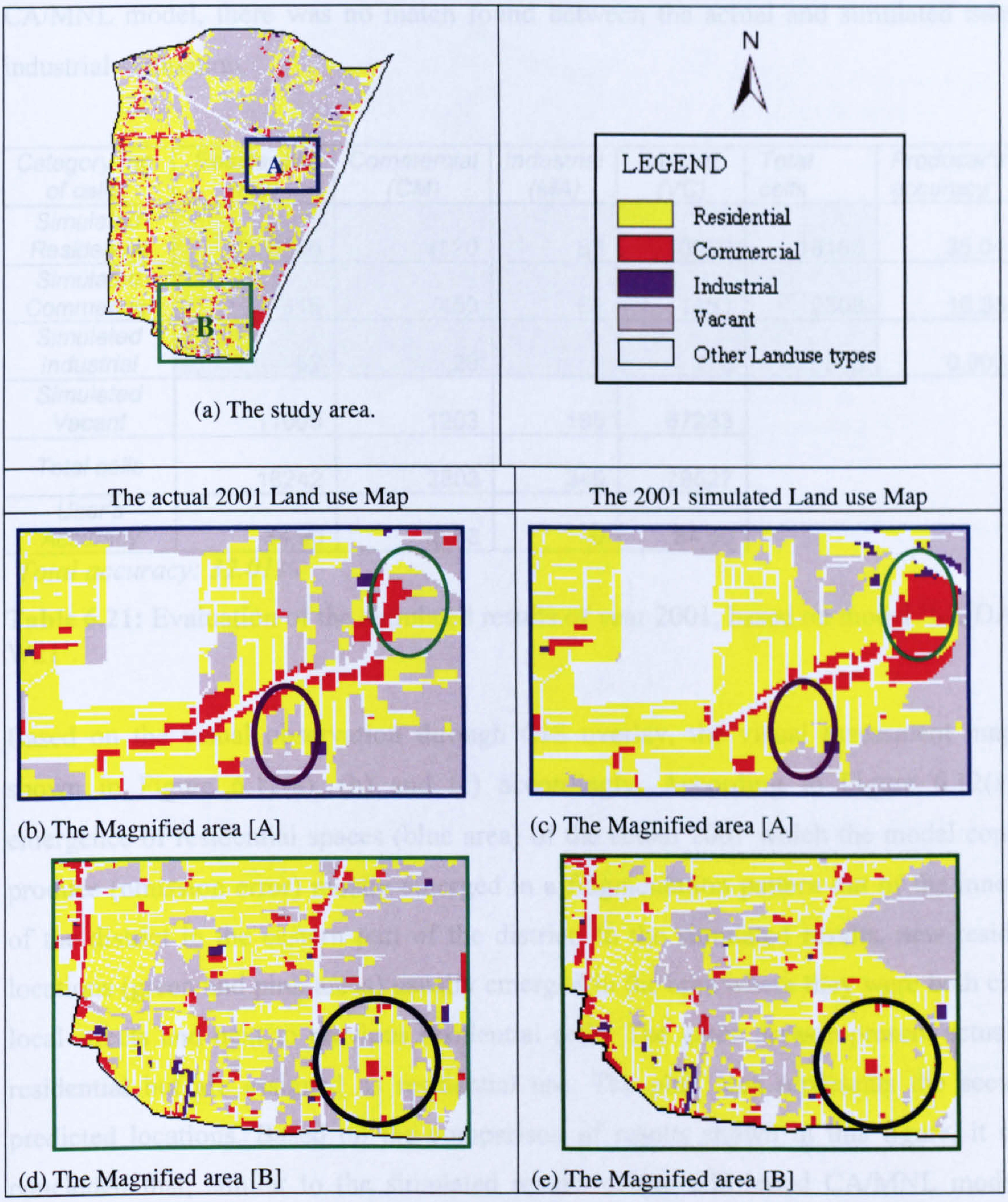


Figure 6.11: Visual comparison of the 2001 actual and the simulated land use map (model 'MCDA-W21').

The model thus assigned 64.96% of non-residential as being residential use (6.17% being commercial, 0.46% being industrial area, and 58.33% being vacant). The accuracy for commercial land use type is lower, with total user's and producer's accuracy of 16.38% and 16.35% respectively. In this case, the model assigned 83.62% of commercial use as being non-commercial (39.97% being residential, 0.71% being industrial area, and 42.94% being vacant). The model thus assigned 83.65% of non-commercial as being commercial use (29.02% being residential, 2.88% being industrial area, and 51.75% being vacant). The poorest accuracy case is of industrial type. Similar to those produced from the GIS-based

CA/MNL model, there was no match found between the actual and simulated maps for industrial simulation.

Category (no. of cells)	Residential (RS)	Commercial (CM)	Industrial (MA)	Vacant (VC)	Total cells	Producer's accuracy
Simulated Residential	6366	1120	83	10599	18168	35.04
Simulated Commercial	815	459	81	1453	2808	16.35
Simulated Industrial	52	20	0	276	348	0.000
Simulated Vacant	11009	1203	185	67233		
Total cells	18242	2802	349	79527		
User's Accuracy	34.90	16.38	0	84.50		

Total accuracy: 32.01%

Table 6.21: Evaluation of the simulated results of year 2001, based on model 'MCDA-W21'.

Based on the visual observation through GIS overlay, the visual assessment maps are shown in Figure 6.12(a), (b) and (c) accordingly. According to Figure 6.12(a), the emergence of residential spaces (blue area) of the actual 2001 which the model could not produce (omission error) mostly emerged in a fragmentation pattern out of the inner core of the district in the eastern part of the district. In the simulated results, new residential locations (green and pink areas) usually emerged in the area where they were both close to local streets and next to previous residential cells. The green areas represent actual non-residential but are assigned as residential use. The pink area represents the accurately predicted locations. Based on the comparison of results shown in this figure, it can be concluded that, similar to the simulated results of the GIS-based CA/MNL model, the model could do best for the emergence of residential cells that occurred in the location where they were both close to local streets and next to the previous residential cells.

According to Figure 6.12(b), the emergence of actual commercial spaces (blue area) which the model could not produce (omission error) emerged randomly covering the entire area in all directions. In the simulated version, the new commercial cells emerged as an enlarging pattern next to the existing commercial cells. It thus can be concluded that, similar to the simulated results of the GIS-based CA/MNL model, the model could do best for the emergence of commercial cells that occurred in the location where they were both close to collector streets and next to the previous commercial cells.

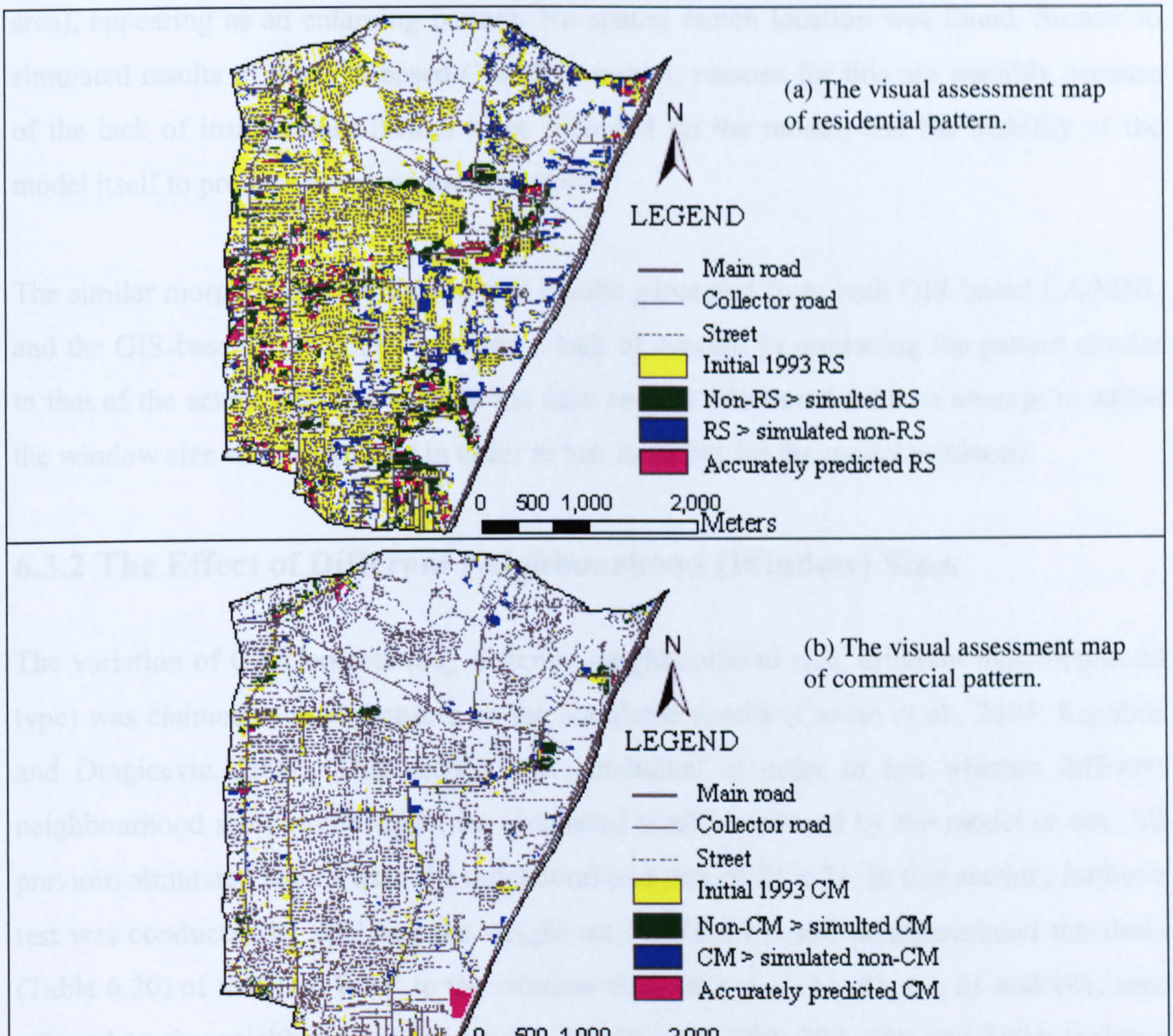


Figure 6.12: The visual assessment maps of model 'MCDA-W21'.

According to Figure 6.12 (c), the emergence of actual industrial spaces (blue area) emerged farther apart and in no particular arrangement, covering the entire area in all directions. In the simulation, the new industrial cells emerged clustering in one area (green area), appearing as an enlarging pattern. No spatial match location was found. Similar to simulated results of the GIS-based CA/MNL model, reasons for this are possibly because of the lack of insignificant factors to be included tin the model, and the inability of the model itself to produce a fragmentation pattern.

The similar morphology of the simulated results generated from both GIS-based CA/MNL and the GIS-based CA/MCDA suggests a lack of success in producing the pattern similar to that of the actual land use map. In the next section (Section 6.3.2), an attempt to adjust the window size was carried out in order to test its effect for the model produced.

6.3.2 The Effect of Different Neighbourhood (Window) Sizes

The variation of CA elements (e.g. different neighbourhood size, different neighbourhood type) was claimed to have impacts on the simulated results (Caruso et al., 2005; Kocabus and Dragicevic, 2004). This section was conducted in order to test whether different neighbourhood sizes would affect the simulated results produced by the model or not. All previous simulated results used the neighbourhood size of 21 x 21. In this section, further a test was conducted by applying the weight set (Table 6.19) and neighbourhood threshold (Table 6.20) of 'MCDA-W21' to the window sizes of 3, 5, 7, 11, 41, 61, 81 and 101, here referred to the neighbourhood of 10, 20, 30, 50, 100, 200, 300, 400 and 500m walking distance respectively. Implementation of the model was named as 'MCDA-W3', 'MCDA-W5', 'MCDA-W7', 'MCDA-W11', 'MCDA-W41', 'MCDA-W61', 'MCDA-W81', and 'MCDA-W101' respectively.

Model performance of the different neighbourhood sizes was assessed through an accuracy table using the same technique as described in Section 6.3.1. The accuracy rate of different window sizes was summarized as shown in Table 6.22. According to the table, the model, in spite of different neighbourhood sizes, produced a correspondence accuracy rate largely similar to that of model 'MCDA-W21'. In comparison, the total accuracy of different neighbourhood sizes is in the range of 28.5 – 32.0. They all correspond in the sense that residential accuracy has the highest accuracy, followed by commercial and industrial respectively. However, amongst the slight change in accuracy rates, it is likely that the model with very small neighbourhood sizes (in this case, 3 x 3 and 5 x 5) and very large neighbourhood sizes (in this case, 101 x 101) leads to the reduction of accuracy rate. Overall, it can be concluded that when using large and small neighbourhood sizes, the differences between them were not identified. Different neighbourhood sizes, thus, did not significantly affect the accuracy table of the model.

Model Name	Neighbourhood Size	Accuracy (%)			
		Residential	Commercial	Industrial	Total
MCDA-W3	3 (10m)	32.43	12.99	0.55	28.72
MCDA-W5	5 (20m)	31.37	13.4	1.98	28.51
MCDA-W7	7 (30m)	32.44	13.58	0.00	29.40
MCDA-W11	11 (50m)	34.04	16.29	0.00	31.13
MCDA-W21	21 (100m)	35.04	16.35	0.00	32.01
MCDA-W41	41 (200m)	34.35	13.8	0.00	30.88
MCDA-W61	61 (300m)	35.12	11.13	0.00	31.41
MCDA-W81	81 (400m)	33.98	7.93	0.00	29.99
MCDA-W101	101 (500m)	33.45	6.99	0.00	29.42

Table 6.22: Comparison of simulated results based on different neighbourhood sizes.

Visual observation was employed to compare the effect of different neighbourhood sizes in terms of morphology and location of developments. Figure 6.13 shows the simulated results based on the neighbourhood size 3, 11, 21, 61, and 101 cells in comparison, here referred to as models 'MCDA-W3', 'MCDA-W11', 'MCDA-W21', 'MCDA-W61', and 'MCDA-W101' respectively. Two findings can be identified. Firstly, the green circles highlighted in the model 'MCDA-W3', 'MCDA-W11', 'MCDA-W21', 'MCDA-W61' in comparison and the blue circles highlighted in the model 'MCDA-W3' , 'MCDA-W11', 'MCDA-W21' in comparison, suggest that the larger the window size, the bigger the cluster around the existing development. This finding is in line with Caruso et al. (2005) in that the size of neighbourhood has an effect on the size of the cluster.

However, this finding is considered partly true. The location highlighted in the green circle of model 'MCDA-W101', compared to that of 'MCDA-W61', as well as the location highlighted in the blue circles of 'MCDA-W61' and 'MCDA-W101', compared to that of 'MCDA-W21' share similar findings in that the increase of neighbourhood size has no effect on the size of the cluster. This is possibly because the concentration of development of each neighbourhood size moved to different locations. This finding, thus, leads to the second finding in that the simulated results based on different neighbourhood sizes created different development locations.

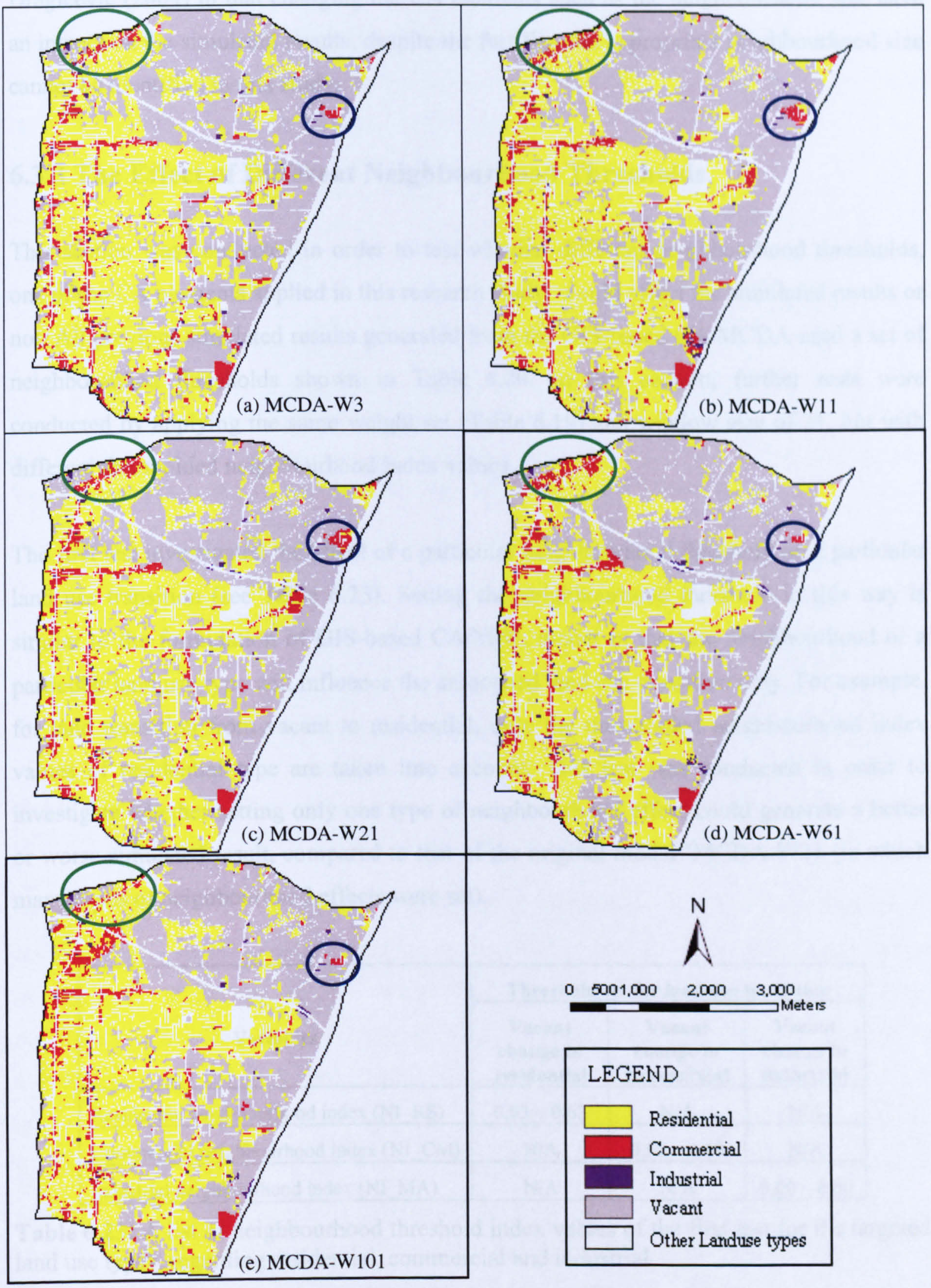


Figure 6.13: Visual comparison of land use pattern with different window sizes.

The results of this sensitivity analysis suggest that the CA approach is neighbourhood size-dependent. This finding is in line with the claim of Caruso et al. (2005) and Kocabus and

Dragicevic (2004) in that changing the CA elements such as the neighbourhood size have an impact on the simulated results, despite the fact that the appropriate neighbourhood size cannot be identified in this study.

6.3.3 The Effect of Different Neighbourhood Thresholds

This section was conducted in order to test whether different neighbourhood thresholds, one of the CA elements applied in this research study, would affect the simulated results or not. All previous simulated results generated from the GIS-based CA/MCDA used a set of neighbourhood thresholds shown in Table 6.20. In this section, further tests were conducted by applying the same weight set (Table 6.19) and window size of 21, but with different thresholded neighbourhood index values.

The first test investigated the effect of a particular neighbourhood threshold on a particular land use transition (see Table 6.23). Setting the neighbourhood threshold in this way is similar to the formulation of GIS-based CA/MNL model in that the neighbourhood of a particular land use type will influence the associated land use transition only. For example, for land transition from vacant to residential, only the thresholded neighbourhood index values of residential type are taken into account. This test was conducted in order to investigate whether setting only one type of neighbourhood effect could generate a better or worse simulated result, compared to that of the original model 'MCDA-W21 (in which many types of neighbourhood effects were set).

Variables	Threshold set for land use transition		
	Vacant change to residential	Vacant change to commercial	Vacant change to industrial
Residential neighbourhood index (NI_RS)	0.03 – 0.63	N/A	N/A
Commercial neighbourhood index (NI_CM)	N/A	0.05 – 0.48	N/A
industrial neighbourhood index (NI_MA)	N/A	N/A	0.00 – 0.60

Table 6.23: A list of neighbourhood threshold index values of the first test for the targeted land use types, regarding residential, commercial and industrial.

The second and the third tests were conducted in order to investigate the effect of reducing the range and increasing the range of neighbourhood thresholds. In the second test, the effect of neighbourhood thresholds was set with a narrower range (see Table 6.24) compared to that of the MCDA-W21 model. The third test was to test the effect of neighbourhood threshold with a bigger range (see Table 6.25) than that of the MCDA-W21 model. Implementation of the models of these three tests was named 'MCDA-THRNR1', 'MCDA-THRNR2', and 'MCDA-THRNR3' respectively.

Variables	Threshold set for land use transition		
	Vacant change to residential	Vacant change to commercial	Vacant change to industrial
Residential neighbourhood index (NI_RS)	0.00 – 0.35	N/A	N/A
Commercial neighbourhood index (NI_CM)	0.00 – 0.05	0.00 – 0.35	0.00 – 0.05
industrial neighbourhood index (NI_MA)	N/A	N/A	N/A

Table 6.24: A list of neighbourhood threshold index values of the second test for the targeted land use types, regarding residential, commercial and industrial.

Variables	Threshold set for land use transition		
	Vacant change to residential	Vacant change to commercial	Vacant change to industrial
Residential neighbourhood index (NI_RS)	0.30 – 0.50	N/A	N/A
Commercial neighbourhood index (NI_CM)	0.00 – 0.30	0.20 – 0.70	0.01 – 0.80
industrial neighbourhood index (NI_MA)	N/A	N/A	N/A

Table 6.25: A list of neighbourhood threshold index values of the third test for the targeted land use types, regarding residential, commercial and industrial.

Model performances of different neighbourhood thresholds were assessed through an accuracy table using the same technique as described in Section 6.3.1. The accuracy rates of different neighbourhood thresholds were summarized as shown in Table 6.26. According to the table, the model, in spite of different neighbourhood thresholds, produced a correspondence accuracy rate roughly similar to that of model 'MCDA-W21' (with a total accuracy of 32.01%). In comparison, the total accuracy of different neighbourhood thresholds is in the range of 29.1 – 31.3. They all correspond in the sense that residential accuracy has a highest accuracy, followed by commercial and industrial respectively. Overall, it can be concluded that the variation of thresholded neighbourhood index values did not significantly affect the accuracy table of the model.

Model Name	Accuracy (%)			
	Residential	Commercial	Industrial	Total
MCDA-THRNR1	34.52	13.94	0.00	31.25
MCDA-THRNR2	31.73	15.42	0.00	29.08
MCDA-THRNR3	33.87	7.07	0.00	30.81

Table 6.26: Comparison of simulated results based on different neighbourhood threshold setting.

Visual observation was also used to compare the effect of different neighbourhood thresholds in terms of both morphology and locations of development. Figure 6.14 compares the simulated results from the actual 2001 land use map with the original model (MCDA-W21), and the first test (MCDA-THRNR1). According to the figure, the pink circle highlights a sample location of residential development. The blue circle highlights a sample location of commercial development, and the green circle highlights a sample location of industrial development in comparison. Based on these three sample locations, the simulated result of the original model (Figure 6.14(b)) and that of the first test (Figure 6.14(c)) seem quite similar in terms of boundary morphology and location of developments, but with slightly different degrees of development (i.e. the pink circle of the simulated result created from the original model has a lesser degree of residential and industrial development than that of the first test). Also, when both simulated results (Figure 6.14(b) and Figure 6.14 (c)) are compared to that of the actual land use map (Figure 6.14 (a)), they are both similar in the sense that they are different from reality. This finding suggests that the GIS-based CA/MCDA model, which incorporated either one neighbourhood effect or more than one neighbourhood effect, did not produce much difference in the simulated results.

Further observation compared the simulated results generated from the original model with the second and the third test. Figure 6.15 shows the simulated results based on the original one ('MCDA-W21') with models 'MCDA-THRNR2' and 'MCDA-THRNR3' respectively.

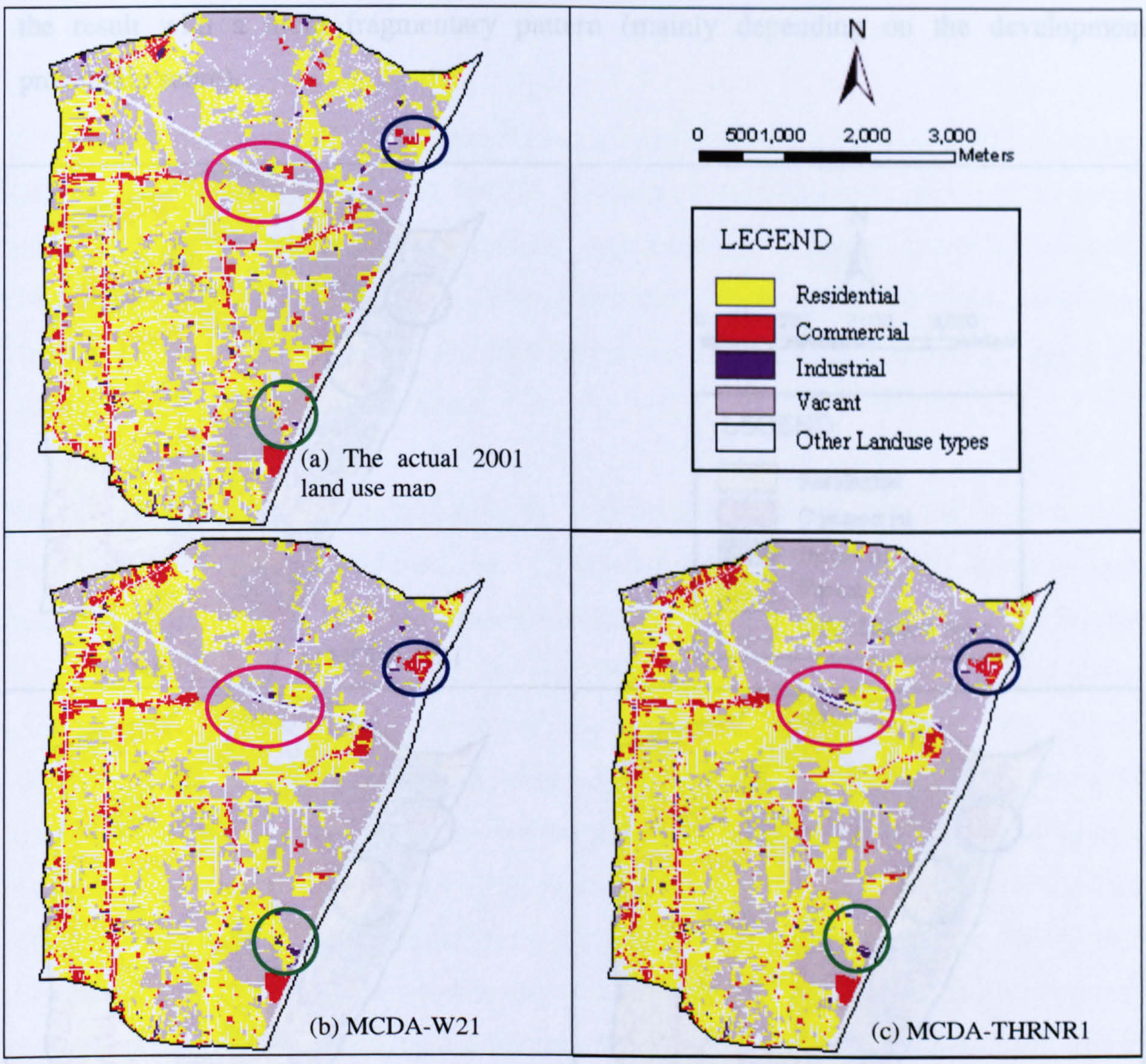


Figure 6.14: Visual comparison of land use pattern with different neighbourhood thresholds.

Three sample locations are highlighted with coloured circles. Clearly, all results differed in terms of land development locations. This is obvious because the concentration of development of these three simulated results varied. However, the simulated result of the second test (Figure 6.15(b)) in which the narrower range of neighbourhood threshold settings were set, shows a lesser degree of compact development and a more fragmentary pattern (see the highlighted blue and green circles), compared to that of the original model ((Figure 6.15(a)). On the contrary, the simulated result of the third test (Figure 6.15(c)), in which the wider range of neighbourhood threshold was set, shows a more compact pattern of development (see the highlighted blue circles), compared to that of the original model ((Figure 6.15(a)). It thus can be concluded that when extending the neighbourhood threshold, the model tends to generate the result with more compact development. In contrast, when narrowing the neighbourhood threshold setting, the model tends to generate the result with a more fragmentary pattern (mainly depending on the development probability score).

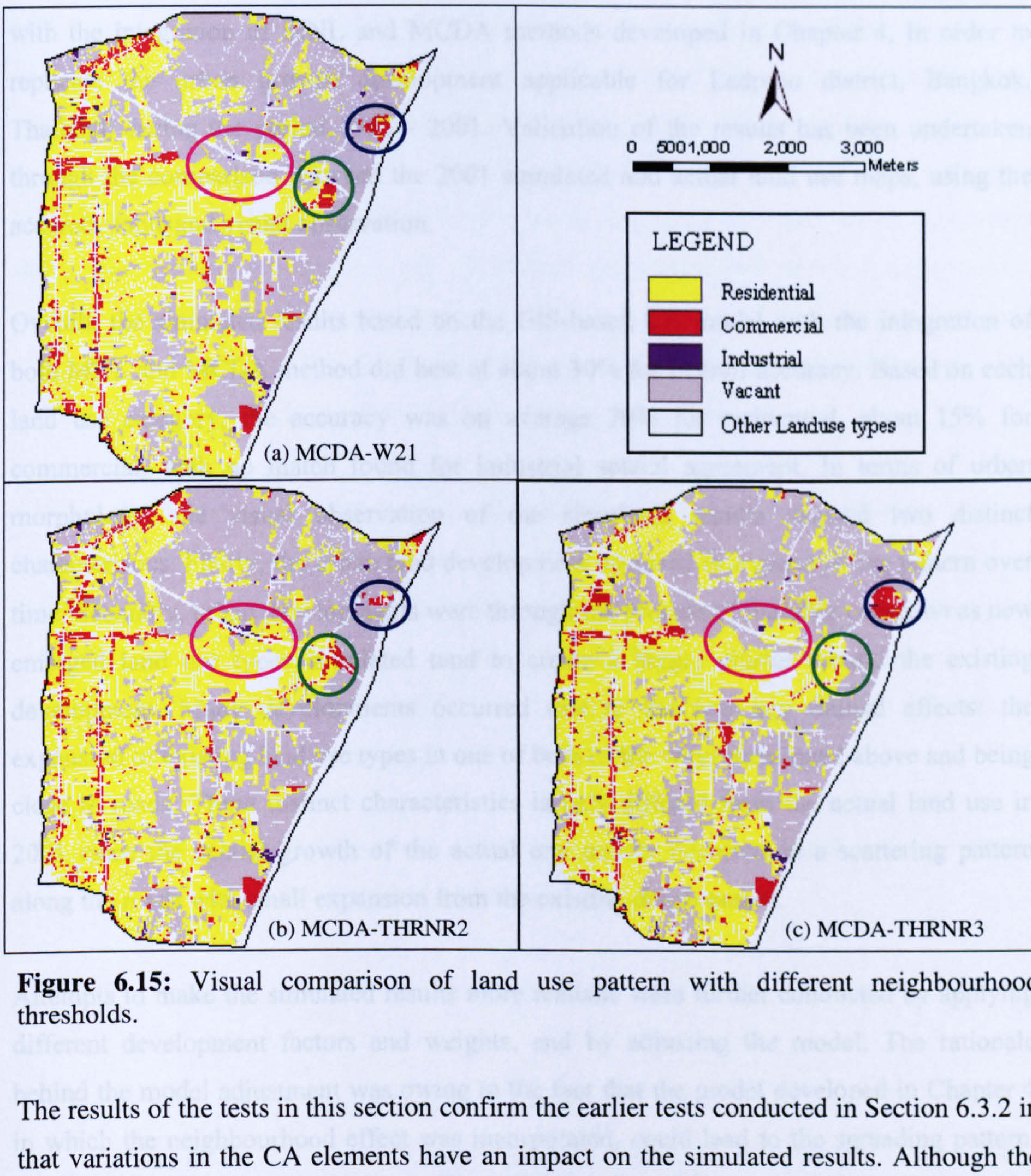


Figure 6.15: Visual comparison of land use pattern with different neighbourhood thresholds.

The results of the tests in this section confirm the earlier tests conducted in Section 6.3.2 in that variations in the CA elements have an impact on the simulated results. Although the optimum neighbourhood threshold cannot be identified in this study, investigation of the impact of the variation on the CA model results is still necessitated in order to improve the understanding of the model as well as its limitations. It is clear that modifying neighbourhood thresholds can have a significant impact on simulated results, but the derivation of an absolute value to use in MCDA testing is not possible with the datasets available. It is also suggested to expand the sensitivity analysis to other CA elements such as spatial resolution, neighbourhood type.

6.4 Conclusion

This chapter has presented the simulated results created from the GIS-based CA approach with the integration of MNL and MCDA methods developed in Chapter 4, in order to replicate the urban growth development applicable for Ladprao district, Bangkok, Thailand, during the period 1993 - 2001. Validation of the results has been undertaken through the comparison between the 2001 simulated and actual land use maps, using the accuracy table and visual observation.

Overall, the simulated results based on the GIS-based CA model with the integration of both MNL and MCDA method did best of about 30% for overall accuracy. Based on each land use category, the accuracy was on average 30% for residential, about 15% for commercial, with no match found for industrial spatial agreement. In terms of urban morphology, the visual observation of the simulated results showed two distinct characteristics. Firstly, the urban land development occurred in a *space-filling* pattern over time. Secondly, urban developments were through the process of *building accretion* as new emerged land use types simulated tend to create a bigger cluster around the existing development. Many developments occurred due to the two key mutual effects: the expansion of existing land use types in one of both of the ways mentioned above and being close to roads. These distinct characteristics largely differed from the actual land use in 2001 since the spatial growth of the actual one mainly occurred in a scattering pattern along the roads with small expansion from the existing development.

Attempts to make the simulated results more realistic were further conducted by applying different development factors and weights, and by adjusting the model. The rationale behind the model adjustment was owing to the fact that the model developed in Chapter 4 in which the neighbourhood effect was incorporated, could lead to the spreading pattern, rather than a fragmentation pattern. The adjustment model was carried out through the GIS-based CA/MNL approach by two techniques: (i) by including the neighbourhood effects as development factors in a MNL regression model and (ii) by excluding the neighbourhood effects.

Validation through the accuracy rate showed that these attempts were unsuccessful as the simulated results were produced with a slightly similar accuracy rate, when compared to the original model. Furthermore, variation in choices of development factors, and the adjustment models employed had an impact on the variation in the development locations. However, in terms of morphology, the simulated results based on the adjusted model MNL3 produced a pattern that is more fragmentary than that of the original model. The unexpected simulated results of these simulations have led to the discussion and suggestions for the model improvement which will be given in the next chapter.

Sensitivity analysis based on different neighbourhood sizes was also tested. The study shows that although the accuracy rate of different neighbourhood sizes was quite similar, the simulated results were significantly different in terms of morphology and land allocation. In addition to this, the finding suggests that the size of the neighbourhood has an effect on the size of cluster around the existing development.

Also in this chapter, different neighbourhood threshold setting was tested in order to investigate their effects on the simulated results. The results suggest that by narrowing or extending the range of thresholded neighbourhood index values, they have an effect on the simulated results in terms of both morphology and land allocation. It is suggested that expansion of the sensitivity analysis to other CA elements is also needed in order to understand the influence of these elements on the simulated results.

The simulated results produced in this chapter suggest that further explanation is necessary. This will be presented in Chapter 7.

CHAPTER 7

Discussion

7.1 Investigating the Simulated Results

In the previous chapter, the simulated results from the GIS-based CA approach with the integration of MNL and MCDA methods were generated, aimed to replicate the urban growth development applicable for Ladprao district, Bangkok, Thailand during the period 1993 - 2001. The model performance of the simulated results was evaluated through the comparison with the actual land use map using both an accuracy table and visual observation. Throughout the validation, the simulated result produced from every experiment failed to capture the realistic urban development. They all had poor model performance in terms of both the spatial agreement accuracy and urban morphology.

The spatial agreement using the accuracy table showed that the simulated results based on the developed model with both MNL and MCDA methods produced a poor correspondence accuracy rate with consistent values. Compared to that of the actual land use map, they were about 30% for overall accuracy (31.59% and 32.01% with MNL (model MNL1-V1) and MCDA (model MCDA-W21) method respectively). In spite of the poor spatial agreement, however, some research works such as those of Wu and Webster (1998) and Wu (2002a) suggest that the measurement of model performance should be considered based on the agreement of morphology rather than the accuracy of spatial matches on a cell-by-cell basis.

In terms of urban morphology, the visual observation of the simulated results each showed similar patterns in that the urban land development was modelled to result in a *space-filling* pattern and the emergence of new land use activities occurred as *bigger clusters around the existing land uses* over time. These two distinct characteristics are substantially different from the actual land use in 2001. In the actual growth, emergence of new land use activities occurred in a scattering pattern over the vacant area, largely found as small patches along areas close to roads and with some extended development areas adjacent to the existing development, as will be discussed in Section 7.1.3.

The poor model performance illustrated above is possibly due to a combination of three reasons: drawbacks in the concepts applied, data incompleteness, and the uniqueness of the study site. In this section, discussion of the poor performance of the simulated results as well as solutions to improve the model are presented

7.1.1 Discussion of the Conceptual Aspect

The major reason for poor performance is probably due to the conceptual aspect used in the research study. Here, this means the inappropriateness of the CA approach to be employed to simulate the pattern of urban district level. The CA approach is well-known for the urban simulation that produces compact morphology in the character of space-filling and the expansion of the established cell to the surroundings (Li and Yeh, 2000; Ward et al., 2000). Most of the research in the area of urban simulation using the CA approach has been based on coarser spatial units such as city, town or region, e.g. the simulation of the suburban expansion of Amherst, Buffalo, U.S. (Batty and Xie, 1994), the application of land conversion of urban / non-urban (Wu and Webster, 2000), and the application of rural and urban land conversion in the city of Guangzhou in southern China (Wu, 2002a). These research projects have been successful in capturing the characteristics of a growth pattern experiencing either the effect of urban sprawl or the encroachment of urban land use in agriculture areas.

In the study site, the CA approach was applied to simulate the growth of land use activities at the district level in the area of Ladprao, a district of Bangkok. Different development factors and weights, different neighbourhood sizes as well as different neighbourhood thresholds were tested in order to improve the simulated results so that they were more realistic. However, all results showed that the model produced a similar urban morphology: the space-filling pattern and the process of building accretion over time iterations. The developed model with CA approach tends not to allow a single cell to emerge individually as it is influenced, significantly, by the neighbourhood effect. These characteristics were very different from the actual 2001 land use map in the study area where the growth patterns were dispersed as small patches, especially along the roads.

Further work was conducted in order to investigate the effect of neighbourhood and whether, or not, it mainly causes the emergence of space-filling and enlarging pattern. This was done by adjusting the model in two different ways. The adjustment models here were referred to as model MNL2 and MNL3 in Section 6.2.3. With model MNL2, the aim was to reduce the effect of neighbourhood. By this technique, the neighbourhood effect for an iteration was included as a part of the MNL probability computation, instead of explicit multiplication. Model MNL3 aimed to discard the effect of neighbourhood completely. Thus, the neighbourhood effect was excluded from the model. In terms of spatial agreement, these two attempts were unsuccessful as the simulated results produced a slightly similar accuracy rate, when compared to the original model. However, in terms of urban morphology, the simulated results based on model MNL3 produced the fragmentary pattern, analogous to that of the actual 2001. This finding suggests that the iterative effect of neighbourhood in the original model developed in the study had a high influence on the urban morphology produced, which resulted in the poor performance of the simulated results generated. This finding corresponds to the observation given by Wu (2002a) as previously addressed in Section 6.2.3 in that simulating the growth without the effect of neighbourhood produced the land development with more scattering pattern.

The application of CA developed here used a strict CA model, which restricts allocation of new development to local neighbourhoods. The actual growth here, however, suggests the relaxation of the neighbourhood space. One alternative approach to reduce the effect of neighbourhood is to apply the concept of Multiagent System (MAS), so-called agent-based modelling (ABM). While the CA approach focuses on land and its transition over space and time, the ABM approach focuses on human actors rather than the landscape (Benenson and Torrens, 2004; Ligtenberg et al., 2001). The pioneering work in the ABM modelling approach was that of Benguigui (1995; 1998). His work adapted the concept of an aggregation model for imitating the town growth of Paris and London. According to his work, he was amongst the first to attempt to incorporate the behaviour of human actors in the model. To simulate the growth, a human performs a visit to a site (in this case, a cell) and its neighbourhood many times before making a decision to develop the 'non-urban' to 'urban' use. According to his model, setting different visiting times results in the production of different degrees of compact and fractal structure. The interesting result was that the model simulated the growth of a town as a compact aggregate in the inner-core of the town, but the periphery had a fractal structure. In spite of many factors and rules that need to be adjusted, adapting such a concept to the research study area can probably allow the emergence of a fractal pattern to be generated. This could match the visually observed change in Ladprao more closely, although reported work so far is still based on city-wide simulations.

7.1.2 Discussion of the Data Incompleteness

Another reason for the poor performance was possibly because of the inability to include all development factors significant to the study site. Firstly, this is due to the limitation of data in terms of their availability and quality. It should be noted that in this research study the data at the detailed district or ward level should be of importance for the simulation of urban development. Data constraints as discussed in Section 3.3 such as the limited historical data map, the unavailability of detailed information of physical characteristics (e.g. land subsidence, flood area information), the unavailability of socio-economic data, led to the limited detailed information included in the model. In addition to this, some development factors such as land price, in spite of its availability, are considered unrealistic, as it was unavailable in the parcel level, non-geo-referenced, and outdated. Thus, it could not be completely used to dictate land demand and supply of the study site.

Secondly, the base map used in this study was confined to cover only Ladprao district since the maps of neighbouring districts nearby were not available for the study. In other words, the influence of development factors based on the neighbouring districts was not taken into account. This potentially causes loss of influential data that is used to simulate the urban growth especially around the boundary of Ladprao area, notably on the north and south (to the east the highway is a major barrier to the neighbour effect, to the west the canal similarly acts as a barrier).

Thirdly, some types of influential development factors are considered difficult to obtain or measure. Such factors mainly account for the individual preference, unknown political, cultural, social and economic factors, which partly affect the decision of actors about land usage and locations (Wu, 2002a). For example, home-buyers probably decide to select the residential location due to the closeness to their parents and relatives, the inherited land, the reputation of housing projects, etc. The example factors given, as discussed in Section 2.3, were considered unpredictable and qualitative (Barredo et al., 2004; Malczewski, 1999a; Voogd, 1983).

Finally, most development factors available in the model were not significant for reflecting the true characteristics of land use activities likely to be developed. Most available data in this study, including the proximity to roads, the proximity to development activities (e.g. commercial area, schools, governments), and proximity to parks, were based mainly on the measurement of spatial characteristics of their locations. The measure of travel distance here could not fully reflect the true physical characteristics of land potential to be developed in Bangkok. The work conducted by Punpuing (1993) carried out a statistical survey of commuting time of Bangkokians for the whole Bangkok area in 1991. The interview results showed that the average time taken for the journey from home to workplace was about fifty minutes with an average distance of 11 km. The figures implied that commuting distance was not related to commuting time. Two roads in different areas, which have similar travel distances, may vary enormously in speed. It may take twenty minutes for one road and only five minutes for another road. Furthermore, her research finding suggested that commuting time was not a major factor in people's decisions about residential locations since not many people changed their workplaces or moved houses in order to avoid reducing commuting time.

The insignificant factors due to the reasons discussed above partly caused the failure of the model in capturing the land likely to be developed. If data is available, the application should add neighbouring districts of Ladprao in order to investigate their influences and probably improve the accuracy of the simulated result, especially the land location close to the boundary of Ladprao district. Furthermore, improvement of the model necessitates detailed information at the neighbourhood and district level. Additional information certainly requires huge data collection from the field survey and interviews in the study site. Suggestions to encompass possible influential development factors should include four sets of factors. Firstly, the additional physical characteristics of the locations should be collected through the field survey. They should include housing characteristics (e.g. housing type – detached house, townhouse, row house etc.), commercial characteristics (e.g. commercial type, building height, number of stories in a building) and industrial characteristics (e.g. industrial type, industrial size). These additional works will give information about the density and intensity of land use activities associated with their existing locations. This kind of information is likely to increase the understanding of land usage in the study area.

Secondly, additional data collection should include population and household data for the detailed level (e.g. census block group). This detailed information can be used to create the zones classified by population density and/or household density within a district. The difference in the population/household density in different locations potentially results in the variation of development density on the land. Areas with low population density tend to have a high potential for land development (Hu and Lo, 2007). Areas with high population/household density, having more intensive land use activities than those of low density, have the potential to convert to commercial areas to service the community. If data permits, this information can be used as a development factor that influences the urban development in its existing location and areas nearby (i.e. low population density attracts residential and industrial development, high population density commercial development). Furthermore, such data can help dictate the development density in the area and it can be used to characterize the emergence of urban development at a more detailed level.

Thirdly, information about the decisions of actors on land usage and locations should be carried out by an interview technique. Collection of this information should be stratified by the demographic data (e.g. age, sex, marital status) and socio-economic data (e.g. income, car ownership, family structure, education, religion, occupation, commuting time, commuting distance) of respondents living or working in the study area. The work conducted by Punpuing (1993) in the study of commuting behaviour patterns between home and workplace in the Bangkok area in 1991 showed that socio-economic data, such as income, homeownership status, education, and occupation, had a significant influence in different degrees on the decision concerning their places of living and their workplaces. For example, from her findings, educated people tended to commute for a longer time and longer distance than the people of low levels of education did. This kind of information may, to some degree, help dictate links to decisions on land use activities.

Finally, variation in speed for travelling suggests that the road network should use travel time, rather than travel distance, since travel time is more realistic than travel distance.

Collection of this information should be stratified by the travel period (e.g. peak and off-peak periods) and mode of transportation (e.g. private car, public transportation, on foot). This data is expected to help improve the accessibility factors (e.g. proximity to roads, proximity to land use activities) in the study to be more realistic.

7.1.3 Discussion of the Complexity and Uniqueness of the Study Site

The final reason for the poor model performance is possibly due to the complexity and unique characteristics of the Ladprao area itself. Unlike some districts of Bangkok, which have physical variation or have seen the effect of new subway system, the land use patterns of Ladprao have been shaped by three main factors in combination: the inadequate supply of public road networks (Dowall, 1989; Webster, 2000), the influence of private developers (Dowall, 1989; Webster, 2000), and the inefficiency of zoning (Sharkawy and Chotipanich, 1998).

Since the economic boom, most urban parts of Ladprao area have been dominated by residential development. Because the supply of infrastructure and public utilities provided by the public government could not keep up with the demands of homebuyers, most developed areas were controlled by land developers. Land developers have always provided the developed land and completed houses with the construction of local streets in the housing projects (Dowall, 1989). During 1974 – 1994, developers shared the real estate market with about 80 percent of the total land development (Sharkawy and Chotipanich, 1998). Since the official city planning work for Bangkok, namely the Department of City Planning (DCP), until now has not been able to provide a detailed land use plan and effective enforcement of zoning policy, they thus cannot control and dictate the direction of the growth of urban development. Up to the present, private developers have influenced the development of land use activities, especially the residential and commercial development projects, in Ladprao and other areas of Bangkok. They also play a major role in dictating the where, when, and how of urban development at the neighbourhood or community level of city building (Webster, 2000). Table 7.1 shows the characteristics of focused actual land patterns in the Ladprao area, which include residential, commercial, industrial and vacant land use type.

Land Use Type	Pattern Type
1. Residential	
1.1 Privately developed houses	Detached houses
1.2 Privately developed residential projects (mubans)	Detached houses, townhouses, flats, and condominiums.
2. Commercial	
2.1 Privately developed small shops	Shop houses
2.2 Privately developed big shops	Shopping malls
3.Industrial	
3.1 Privately developed small industrial	Shop houses
3.2 Privately developed large industrial	Large buildings, warehouses
4. Vacant (undeveloped land)	
4.1 Small vacant land	Small patches of land
4.2 Large vacant land	Large open space, bare land, swamp

Table 7.1: The focused land use patterns in the study area.

The first land use pattern is residential area. Two types of residential land patterns are found in the study area: privately developed houses and privately developed residential projects. Privately developed houses are usually detached houses, developed by private homeowners. The land patterns (see Figure 7.1(a)) can be developed if their locations can access roads as well as public utilities (e.g. electricity service). These types of land were usually found scattered next to local streets in a fragmentation pattern. Another type of residential pattern, private developed residential projects, is called 'muban' in Thai (see Figure 7.1(b)). This residential pattern is developed by the private real estate developers. The projects are varied in size and price level. A variety of housing units, ranging in size from 800 to 4,000 sq.m. (Sharkawy and Chotipanich, 1998), include detached houses, townhouses, condominiums and flats. The projects contain a varied range of amenities and services such as club facilities. The housing projects also include infrastructure: 'muban' streets, and the local streets connecting to the public road outside the project. Public facilities (e.g. electricity, potable water, telephone lines) have been put in place later by public government agencies (Thomson and Hardin, 2000; Webster, 2000). It should be noted that because the construction of road development is created at the same time as the house building, factors such as proximity to roads could not be used to capture the upcoming development in the simulation of the study.

(a) The land pattern of privately developed residential houses.

(b) The pattern of privately developed residential projects, called 'muban' in Thai, [A] refers to townhouses and, [B] refers to detached houses

Figure 7.1: Example of residential pattern observed from the study site (source: Aerial imagery acquired from GoogleEarth via internet, accessed on 22nd July 2007).

Two types of commercial patterns are found in the study site: shop houses and shopping malls. The first type is usually found in the study area. It is normally built as shop house. According to the survey conducted by Sharkawy and Chotipanich (1998), each shop house unit is normally constructed in blocks of eight to twenty houses. Each shop house is three to five stories high. The ground and first floors of each unit are normally used for retail or commercial activities, while the upper floors are used for residential purposes. In the study site, shop houses (see Figure 7.2(a)) are frequently found built in the front portions of large residential plots along both sides of major roads and collector streets. Another commercial use in the study is the shopping mall (see figure 7.2(b)), appearing as a large building built within a big area and surrounded by car parks and service areas.

(a) Rows of shop houses, highlighted in red.

(b) A shopping mall, highlighted in red.

Figure 7.2: Example of commercial pattern observed from the study site (source: Aerial imagery acquired from GoogleEarth via internet, accessed on 22nd July 2007).

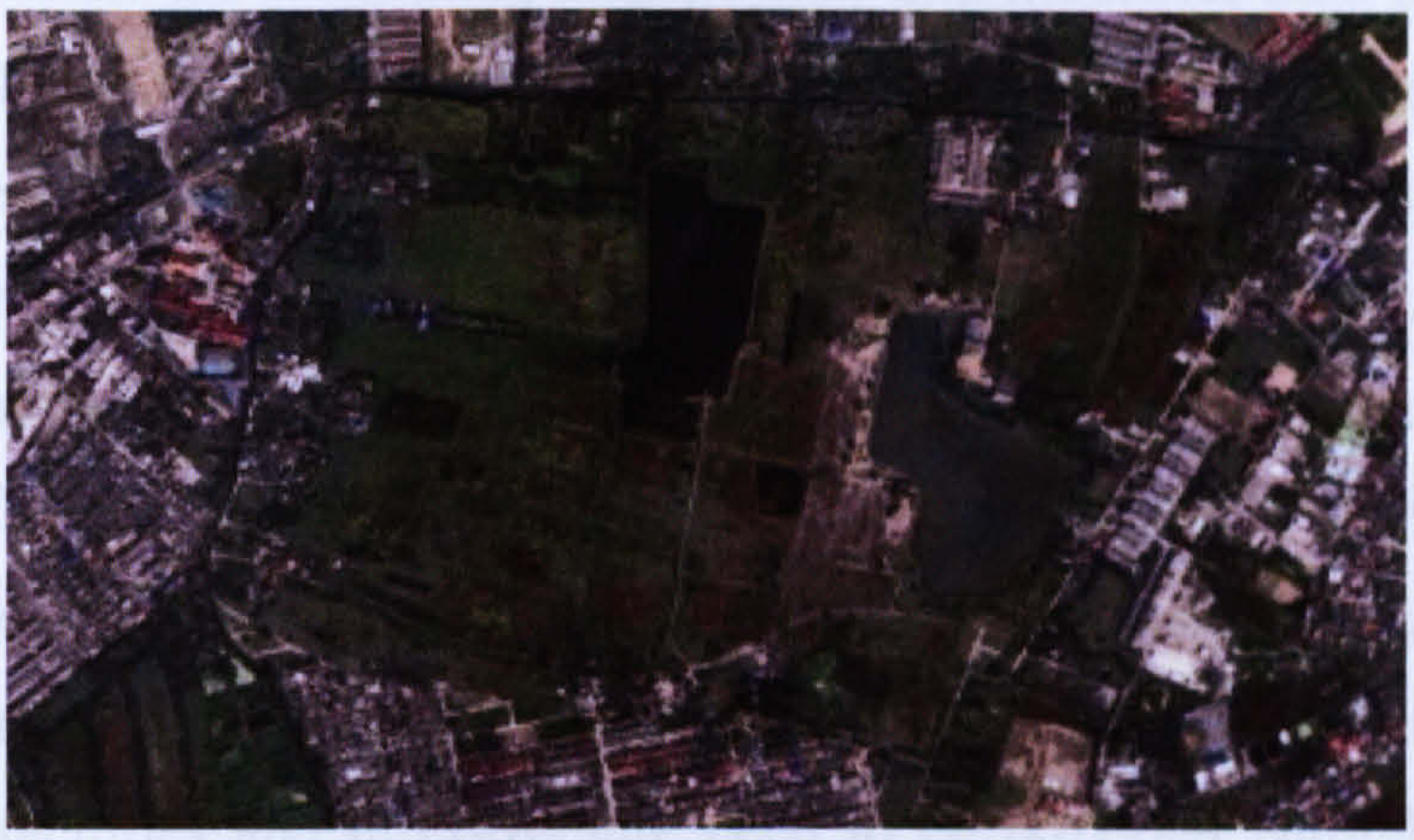
Two types of industrial use are found in the study site: small and large industry. They serve different functions in the site. Small industries (see figure 7.3(a)) such as garage repair shop occur mixed with commercial and residential activities. Normally, they are located in the shop houses or small buildings. Large industries (see figure 7.3(b)) such as factories are emergent as large buildings covering a large area, normally in areas surrounded by vacant land or dominated by less-developed areas.

(a) Pattern of small industrial units, highlighted in red.

(b) Pattern of large industrial development, highlighted in red.

Figure 7.3: Example of industrial pattern observed from the study site (source: Aerial imagery acquired from GoogleEarth via internet, accessed on 22nd July 2007).

Two types of vacant land are found in the study site: the small vacant patches and the large vacant areas. The small vacant patches usually appear as pocket areas left vacant between major developments. These small vacant plots have very limited access to the road network (see Figure 7.4(b)). Some areas appear as blind spots, inaccessible to roads. These patterns appear because of the unplanned road network described in Section 3.1.1.3. Another pattern is the large vacant area (see Figure 7.4(a)), which is usually found as open space or swamps, with no road access. These areas have potential to change to other land use activities. They are usually developed by private developers later.



(a) Large area of vacant land.

(b) Small area of vacant land, highlighted in red.

Figure 7.4: Example of vacant land observed from the study site (source: Aerial imagery acquired from GoogleEarth via internet, accessed on 22nd July 2007).

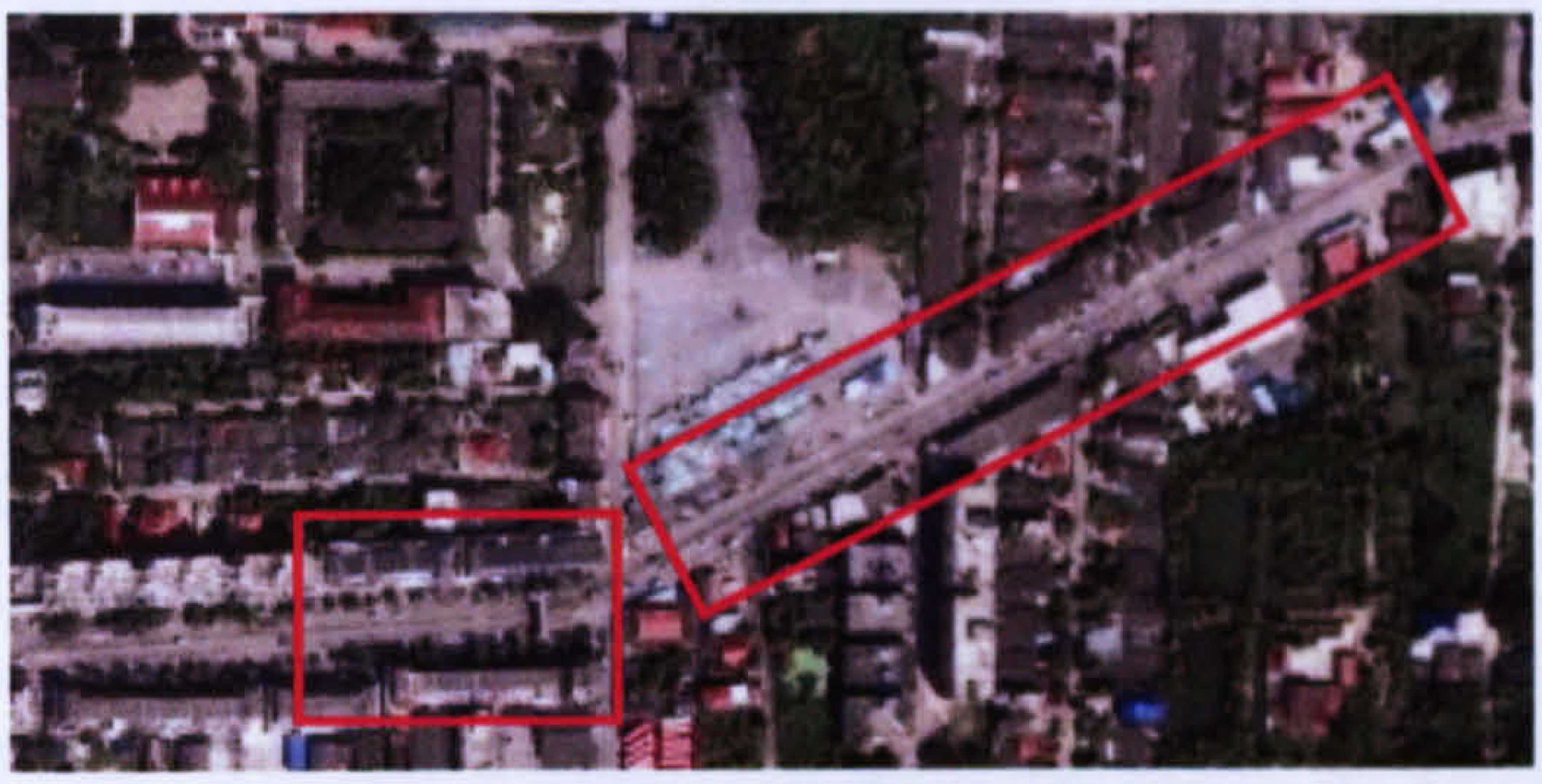
The actual land patterns observed have indicated that these patterns are different in terms of the process of emergence. As a result, this partly answers the question why the model could not replicate the land use pattern of actual growth. According to this, there are a few suggestions that potentially improve the application. Firstly, the three land use types used in this study regarding residential, commercial and industrial are not sufficiently well defined to differentiate the actual land patterns. It thus requires the detailed classification of sub-classes of land use as indicated in the actual patterns. For example, the industrial type used in this study includes both small and large industries. Both types are usually different in terms of location, size, service activities, and neighbouring land use types. They should be classified into two sub-classes, with regard to the industry pattern of small and big industries.

Secondly, since each actual land pattern has emerged due to different human actors making decisions on different land use activities, each land pattern as well as each sub-class of land pattern thus should require different development factors as well as different rules for land transition. This means the residential, commercial, and industrial land pattern potentially need different rules and different development factors. Also, each sub-class of land pattern (e.g. privately developed houses and privately developed residential projects) may require different rules and development factors. For example, for the private residential pattern, private owners are the key actors for land development. The model should mainly take into account factors behind the personal or individual decision about the land. In the land pattern of privately developed residential projects in which the developer plays a major role in developing plots, the emergence of land development could be in any location where no basic facilities and infrastructure need to be provided beforehand. Developers frequently buy small pieces of land from different landowners in a cumulative manner in many locations in advance. Then, they develop them when they decide that they can make most profit on them (e.g. after the creation of new major roads nearby). The model should mainly take into account factors behind the developer's decisions on land.

Thirdly, the application should consider the effect of different road types on the patterns of land use activities. Generally, different road types have different functions in the metropolitan transportation planning as well as their influence on land use activities (Dickey, 1975). Highways have been designed with no direct land access and no interaction with other local streets in order to move traffic quickly. Major roads move traffic and provide access to land use. Collector streets are used to filter traffic from the local streets and to connect to major roads and land use activities such as malls, schools and large industrial areas. Local streets primarily provide access to primarily residential land.

The model applied in this research used different weights for different road types for the emergence of different land use activities. For example, based on the residential suitability, the proximity to local streets factor was given higher weight compared to other factors. For the commercial suitability, proximity to collector and major roads was given higher weight compared to other factors. However, the application did not consider that different road types produced different morphologies for different land use activities and even for the same land use activity. An example in Figure 7.5 shows the observation of actual land development. The commercial land use pattern located adjacent to different road types (collector streets, major roads, and highways) has differing structures and morphologies.

While the collector streets, as well as major roads, allow the emergence of commercial spread on both sides of roads, the highway acted as barriers for the development of activities as it did not allow the emergence of growth on the opposite side due to the restricted road access. It thus suggests that the model should apply different rules based on different road types for the emergence of different land use activities.



(a) The commercial pattern located at the collector streets, highlighted in red box.

(b) The commercial pattern located at the major road, highlighted in red box.

(c) The commercial pattern located at the highway, highlighted in red box.

Figure 7.5: Example of different commercial patterns that have developed at different road types, observed in the study site (source: Aerial imagery acquired from GoogleEarth via internet, accessed on 22nd July 2007).

Finally, the involvement of human actors in the observed actual land patterns suggests that finding an approach to imitate the human actions in a spatial decision-making process on land is necessary. The CA itself has been criticized for having a limited ability for incorporating human factors in the simulation (Benenson and Torrens, 2004; Ligtenberg et al., 2001). The lack of such ability supports the introduction of the recent land use modelling approach, the agent based model (ABM), previously mentioned in Section 7.1.1. The ABM approach allows human decision-making on location to be incorporated. It thus allows us to increase the understanding of human actions in shaping the land use patterns. In an example work of Benenson (1998), agents act as imitators of the residential behaviour and development of human beings, which can interact with other agents, and change their locations as well as their own properties in the system. In the model, a two-layer structure is built. The housing infrastructure layer as a first layer represents the properties of urban housing. The second layer comprises the free agents, whose rationale decisions on residential locations are bound with the economic status and cultural identity. The interaction between the spatial processes and human actors could help simulate the residential development more realistically. By applying the ABM approach to the Ladprao area, the behaviour of different human actors (e.g. developers, commercial people) and the behaviour of different road types can be encoded to define different rules and development factors for land use development. However, this approach can be potentially applied to the study area if data incompleteness discussed in previous section can be overcome.

7.2 Investigating the Techniques Used for Urban Simulation

Techniques employed for developing the urban simulation used in this study is an endeavour to integrate many disciplines for urban development. The integration involves the fields of geographical and spatial expertise through the GIS functional capabilities, complex systems theory through the CA approach, the statistical multinomial logistic regression (MNL) method, and the multi-criteria decision analysis (MCDA) method. Implementation of these techniques, to some degree, has an effect on the simulated results produced. In this section, discussions about the limitations and advantages of these techniques in the context of technical and application aspects are presented.

In this study, GIS serve as a platform for urban growth simulation. It was developed using ArcGIS 9.1 and updated in ArcGIS 9.2 (including spatial analyst module) and a VBA scripting language as a tool to perform all operations and functions on a cell-by-cell basis. A few limitations are related to the GIS technical software applied to the study. The first one is the inability of the program to create the exact numbers of cells required to change for the whole simulation. For example, as shown in an accuracy table of Table 6.3, the total cells required (threshold set) for the residential aspect is 18,242 while the total cells actually produced for residential is 18,276, an over-production of 34 cells (or 0.2 percent). Technically, when generating the simulated land use map for each year, the program sorts the probability values based on the highest probability from an attribute table. It then accumulates the number of cells until it reaches the total number of cells required for an iteration, which usually does not exactly match the required amount. Instead, the numbers of cells that are closest to the actual amount needed are used as the stopping point. For each iteration, the program thus cumulatively increases the imbalance between the number of cells wanted and the quantity produced. This implies that more iteration will produce more difference between them. This limitation is due to the limited precision value for computer computation, which can be relaxed if the computer is more powerful.

Another technical limitation involves the limited ability of the program to choose the most appropriate probability values for map updating from three probability maps at the same time. In this study, the three probability maps are generated, regarding vacant change to residential, to commercial, to industrial. Each probability map has an associated attribute table. The technical limitation is that the program cannot read the probability values of the three tables at the same time. Instead the program starts with one attribute table, finds the amount of cells wanted based on the highest probability value, flags those values for update, then moves to the next attribute table, and so on. In this research study, the sequence of choosing the attribute table depends on the sequence of category update, set by users as described in Section 5.3.2.1. By doing this, it does not guarantee that the highest probability values of each table are chosen nor is there a guarantee that the attribute chosen is the most likely. Figure 7.6 illustrates an example of cells selection for land transition. Suppose that for each table, two cells having highest probability value need to be chosen for land conversion. Based on the order of the table (Table 1, 2 and 3 respectively) and the order of highest probability value (see column 'Rank' in the figure), Table 1 chooses cells F and B. Table 2 selects cells C and D because cell B is already occupied by Table 1. For the same reason, Table 3 chooses cells A and G. This limitation described here is due to the limitation of raster handling of the software. In the future, this limitation can be overcome if the software can provide a more powerful way to operate with raster datasets.

Table 1

Cell no.	Probability value	Rank
A	0.45	4
B	0.69	2
C	0.31	6
D	0.34	5
E	0.59	3
F	0.78	1
G	0.12	7

Table 2

Cell no.	Probability value	Rank
A	0.19	5
B	0.98	1
C	0.87	2
D	0.64	3
E	0.39	4
F	0.18	6
G	0.09	7

Table 3

Cell no.	Probability value	Rank
A	0.47	3
B	0.67	2
C	0.11	7
D	0.21	5
E	0.12	6
F	0.78	1
G	0.29	4

Cell selected based on the highest proabability order

Cell occupied by previous step and unavailable for being chosen

Figure 7.6: An example showing the means of cell selection based on the highest probability order.

Apart from the technical limitations given above, another limitation involves the computation of probability value using the MNL method employed. In this study, the probability values generated by the MNL method are computed mutually for the four transition types: vacant to residential, to commercial, to industrial and no change. The probability values produced from the multinomial logistic regression have a value ranging from zero to one. For a single cell location, it contains a probability value of all land transition to be in total equal to 1 (e.g. 0.2 for vacant change to residential, 0.5 for vacant change to commercial, 0.1 for vacant change to industrial, and 0.2 for vacant change to vacant). There is one problem related to the computation of MNL employed in this study. Suppose that a single cell location, based on its cell suitability, has a high potential to change to more than one transition type. Because the probability of all land transition is computed mutually, it will result in the reduction of probability value for both land transitions. For example, a single cell has the potential to change to both residential and commercial, it thus, after MNL computation, contains a probability value equal to 0.45 for vacant change to residential, 0.40 for vacant change to commercial, 0.10 for vacant change to industrial, and 0.05 for no change. When choosing the highest probability score for updating land use, it seems that the score of that cell will not be in the top rank. With the MNL method, thus, only cells that have the potential to change to only one land use transition will get the high probability scores. Because of this limitation, it means that only cells that have high value for one land transition are likely to be chosen.

In spite of some limitations, the integration of these techniques employed in this application demonstrates positive outcomes. Firstly, this application is considered to be one of the few efforts to establish an urban spatial simulation model for the Bangkok area. In Thailand, up to now the development of a GIS-based spatial simulation model and its applications have been hardly evident. Most of the research work in the vicinity of Bangkok has been closely focused on the area of land use change detection and urban monitoring (e.g. the work of Tachizuka et al. (2002)). One of the few efforts in this area was the application conducted by ESRI (Thailand) Co. Ltd. (1997) with the incorporation of DTCP. The work was built to incorporate GIS for land planning procedures. The final product of the application created the potential land development in the view of city planners but without an attempt for simulation prediction. In this study, with the CA approach embedded in a GIS system, a spatial micro-simulation model has been developed for the urban study. The simulated results thus predict the area for potential development for the land use activities considered.

Secondly, this application was undertaken for the simulation of the urban land use change at the detail of district level. Most research work conducted by the CA approach in the context of either developed or developing nations (e.g. the work of Batty and Xie (1994) and Wu (2002a)) have been heavily focused on the urban simulation at the city or region levels. This application was an attempt to capture the urban development at a finer level, the district level. This is considered very useful to help improve the understanding of how the city works at the district level. The study of the urban simulation through the CA approach with different details of urban level can fulfil different study purposes since different levels of urban study require different sets of development factors and interactions under the system. At the city or region level, the focus of land conversion is applicable for the state of cell where the two types (e.g. developed and non-developed, or urban and non-urban area) can be separated. The development factors used for such simulation are determined based on the characteristics under the city or region context (e.g. proximity to city centre). At the district level such as that being applied to this study, the land use activities such as vacant (non-developed), residential, commercial and industrial area can be distinguishable. Such levels permit the interaction between land use activities to be observed and investigated. In the application, thus, the simulation allowed the land development between the land use categories to be generated (e.g. the vacant transition to residential land).

Furthermore, the application demonstrated here is also considered the first of its kind, which is applied at the unit of district level in the Bangkok area. The application in which the study is focused on the spatial unit of the district level will be very useful since it can help predict and map areas prone to experience growth for each land use activity. A fine resolution can help easier implementation of land policies on a finer scale. At the present date in which the lack of the development plan at the district level, as addressed in Section 3.1.1.3, is the remaining unsolved problem for the Bangkok area, information at the district level can be used to support the work of DCP and BMA agents, the organizations which have responsibility to formulate the Bangkok planning policy, for being a guide for the creation of realistic and practical land use planning.

Thirdly, this application is an attempt to integrate the statistical multinomial logistic regression method (MNL) in an urban spatial simulation model, in order to identify the potential cells for development. Modelling with this logistic regression has an advantage in that the approach, based on an empirical estimation, allows a data-driven means to the choice of criterion weights (Hu and Lo, 2007). This is very useful for the case of Bangkok since only a few researches (e.g. Bruijn (1991)) have been conducted in the study area. As such, prior knowledge concerning development factors and their relative weights in the study area has been limited. In addition, unlike most of the early works in which logistic regression was applied on the basis of the binary outcomes of land conversion, such as the conversion of non-urban to urban land (e.g. Hu and Lo (2007), Martin and Wu (1999)), the conversion of rural to urban area (e.g. Wu (2002a)), the conversion of forest to deforestation (e.g. Soares-Filho et al. (2002)), this research study extended the regression to handle four land conversions: vacant to residential, to commercial, to industrial and no change. A major limitation due to the applied MNL method to the study previously discussed is highlighted. This allows an insight into the MNL approach used for the simulation.

Fourthly, this application is one of the few attempts to incorporate multi-criteria decision analysis (MCDA) in an urban spatial simulation model for the Bangkok area. Modelling with MCDA allows a knowledge-based set by the decision-makers and planners. This is very useful to the Bangkok area since up to the present a handful research projects (e.g. ESRI (Thailand) Co. Ltd. (1997)) in the Bangkok area have been conducted in order to incorporate the views of urban planners into the urban spatial analysis. Previously discussed in Section 3.1.1.4, one major problem concerning the ineffective planning policy is due to the lack of cohesive planning policy and coordination amongst relevant planning agents. To overcome this problem, integrating MCDA in this study opens up a way that allows those relevant agencies to test different scenarios based on different points of view. In addition to its predictive ability, this application thus can be modified to serve as a planning tool for Bangkok planners to help visualize scenarios and answer "What if" questions for several planning options.

Finally, the application was built to be user-friendly. A VBA scripting language was programmed as a set of customized tools to help facilitate model simulation, the easy-to-use graphical user interface allows users to adjust the initial conditions, and alter parameter and environment setting values. Thus, it enables different scenarios and conditions to be easily and comfortably tested and compared. This is considered useful to urban planners as it permits the investigation and the study of the urban growth dynamics through different environment settings (e.g. comparing scenarios with and without zoning control). In the context of urban planning, such capability is necessity as it allows different planning regimes to be explored (Wu and Webster, 2000). Possibly, the model designed allows the application to other regions with different datasets by changing input data layers. In this study, through the simulation tools developed, different scenarios (e.g. different sets of development factors, weights, window (neighbourhood) sizes and neighbourhood types) were easily conducted as an attempt to capture the characteristics of land development in the study site. In their current form, however, these programs still need further modification so as to be flexible enough to add more development factor options easily, to be restructured for computer efficiency.

7.3 Conclusion

Two main themes are discussed in this chapter. The first theme involves discussion about the poor simulated results generated in Chapter 6. The discussion concludes that the poor simulated results are probably because of three keys in combination: the inappropriateness of the CA approach used for the simulation, the lack of ability to include all influential development factors and the complexity and uniqueness of the study area itself. Suggestions to improve the model performance are also presented. The second theme discusses the limitations and advantages of integrated techniques employed in the study, that is, the integration of GIS technology through a VBA scripting language, the CA approach, and the MNL and MCDA methods. Despite some limitations, the application developed is the first attempt to establish an urban spatial micro-simulation model for the Bangkok area at the detail of district level. Several lessons have been learned from this study. The discussions and suggestions from this chapter have led to the comment on future research work in the next chapter.

CHAPTER 8

Summary and Recommendations

8.1 Introduction

In this research study, a combined GIS-based CA approach has been developed to simulate the spatial pattern of urban growth. Ladprao, one of Bangkok's fifty districts, was chosen to represent the area that experienced growth during the period 1993 – 2001. The developed model was implemented to replicate the spatial pattern of the targeted land use change from vacant to residential, to commercial and to industrial area during the study period on a two-year interval basis. Validation of the model was undertaken through the comparison between the 2001 simulated and actual land use map. In the following section, the findings of the thesis are discussed. This is followed by further recommendations for future research work in Section 8.3. In the last section (Section 8.4) the achievement of the research study is presented.

8.2 Main Findings of the Thesis

The work conducted in Chapter 6 presented the simulated results of a combined GIS-based CA model with the integration of multinomial logistic regression (MNL) and multi-criteria decision analysis (MCDA) methods. The implementation has been parameterised using the land use data of 1993 and evaluated on the basis of its ability to replicate the historical growth of the actual 2001 land use. Different development factors and weights were tested in Section 6.2.2 as a means to seek for the best set of parameters to replicate the actual pattern of urban development during the study period. The finding indicates that using different development factors and weights produced the variation in the development locations, but with slight equivalence in a spatial agreement. Different neighbourhood sizes and different neighbourhood thresholds were also tested in Section 6.3.2 and 6.3.3

respectively, in order to investigate the sensitivity analysis. The finding indicates that varying both neighbourhood sizes and neighbourhood thresholds had an impact on the simulation results, especially the variation in the development locations. Thus, the finding has led to the conclusion in that the model developed is sensitive to the change of neighbourhood size and neighbourhood threshold.

According to the results generated from both methods, the main finding indicates that the implementation of the developed model applicable to the Ladprao area produced unpromising results as it failed to capture the realistic pattern of land use change during the study period when compared to the actual development. The simulation created poor model performance in terms of both spatial agreement and urban morphology. In terms of the spatial agreement, when compared to that of the actual land use map, the overall accuracy was about 30% (31.59% and 32.01% with MNL and MCDA methods correspondingly). In terms of urban morphology, the results produced with both methods emerged in a space-filling pattern and their urban growth over a discrete time-step acted as a process of building accretion, appearing as a growing cluster around the existing development. The unexpected, but interesting, results of this observation have led to the conclusion of three possible reasons, which are given in detail in Section 7.1. These reasons include the inappropriateness of the CA approach to simulate the pattern of the urban district level, the inability to include all development factors significant to the study site, and finally the distinctive characteristics of Ladprao and the Bangkok area itself. This finding suggests some solutions to improve the model performance by reducing the neighbourhood effect when simulating at the urban district level, by adding significant factors if data permits in the near future, and/or by adapting the agent-based modelling approach.

8.3 Future Research Directions

This application is the first to establish an urban spatial micro-simulation model for the Bangkok area in detail at the district level. Nevertheless, the model developed here is still at the early stages of development. It has been of the simple kind. Constraints in model implementation were critically bound to budget, time, and data availability. The unpromising results produced in Chapter 6 suggest that many challenges remain. Future research work can be continued as described below.

Firstly, the urban simulation model developed here was based on the traditional cellular automata approach, which restricts the emergence of new development to a limited neighbourhood space. In this application, the influence of neighbourhood tremendously affected the allocation of new land use activities that produce a boundary morphology completely different from that of the actual growth. In response to this, the discussion in Section 7.1.3 has indicated that the involvement of human actors plays an important role in shaping the actual pattern of land use transition in this study area. Future work should investigate the possibility to extend the CA approach, which could be done by adapting the agent-based modelling (ABM) to the application. Potentially, the ABM approach may reduce the spatial error and effect of neighbourhood generated in the application of the GIS-based CA model implemented in Chapter 6. In order to model this new land use allocation, there are a few major considerations to be taken in to account. The first consideration involves the requirement for understanding the process of human actors (so–called agents) in making decisions about land. A general rule for all land transition types cannot be applied. Each land transition possibly requires different development factors as well as different rules, obtained from particular groups of human actors. The second consideration concerns the data collection. The involvement of human actors means that factors such as individual preferences, demographic data (e.g. age, marital status), socio-economic data (e.g. education, occupation, income) should also be included. It thus requires huge data collection based on interviews from the field survey, the incorporation of methods for behavioural approach and qualitative analysis.

Secondly, in this application, land use map data of the neighbouring districts is not available. Future work should be undertaken by including the neighbouring districts in the application. Such additional information can make the simulation more realistic, especially the areas near the edge of the study area. Furthermore, the comparison of the simulated results between the application with and without neighbouring districts should be examined, in order to help improve the understanding the influential effect that neighbouring districts contributed to the study area.

Thirdly, in the present form, its application has been implemented to focus on the change from vacant to residential, commercial, or industrial area. However, the model itself can potentially be modified and extended to cover the other types of land use transition. In order to do this, additional development factors and rules associated with new land transition have to be further investigated and determined.

Fourthly, the model developed in its current form produces a land use transition based on the highest probability of land development. This is to ensure that the simulated outcome generated for each run of the same model, environment and parameter setting is the same and can then be used to compare with the actual land use map for validation purposes. However, some of the urban CA research work (e.g. Ward et al. (2000), Wu (2002a), Yang and Lo (2003)) applied the integration of stochastic or random processes (e.g. Monte Carlo method) for simulation in order to increase the speed of computation (Wu, 2002a). Furthermore, since the complexity of urban land development, in reality, is unpredictable and remains unknown (Batty and Torrens, 2001), using the random process may possibly produce simulated results more sensible and more realistic to the emergence of a real-world urban development. The future version of the developed model thus should add an alternative to allow the incorporation of the stochastic process for simulation.

Fifthly, in the present application, the model has accounted for simulation based on a linear growth increase during the whole period. Future work can be carried out to modify the model to account for simulation of non-linear or irregular growth over time. Such modification can make the simulation more realistic since urban growth in reality does not follow a linear form. For example, the economic crisis in 1997 interrupted the boom of residential growth in the study area. In terms of modelling, the land allocation for each iteration needs to be adjusted to reflect the true development for land use activities.

Sixthly, future research work will be interesting to extend the simulation to the growth of other areas of Bangkok under the study of district and zone level, if data permits. At the district level, work may continue investigating other districts of the Bangkok area. Previously mentioned in Section 1.2, the Bangkok area has been classified as five zones (e.g. inner city, suburban zone) according to their distance from the CBD. At this level, work may undertake simulation based on the zone classified. For both levels, some additional factors such as additional transportation modes (e.g. sky train, river, and train network) shall be included. Potentially, this will help in seeking for the differences and similarities amongst those districts and Bangkok zones and lead to a better understanding of the characteristics of Bangkok as well as the structure and the pattern behind the land use development of the Bangkok area. To achieve this, a similar procedure has to be repeated for each district or each zone. Different sets of development factors and their weights are required to re-examine them, as each district or zone has its own distinct characteristics.

Finally, additional future work should also be done to study the overall growth of Bangkok and its region nearby, namely Greater Bangkok. With this future work, studying the urban structure of Bangkok will be focused at the city and regional level. To implement this, however, many critical modifications have to be made to the model in terms of both the cell state setting and the development factors consideration.

Future work, which will be conducted at the four levels; district, zone, city, and region, will help understand the structure of Bangkok's growth. From the viewpoint of urban planning, such integrated information can help to dictate the creation of a realistic planning policy at both levels, the land use planning and comprehensive plan.

8.4 Revisiting the Research Aim and Assessing the Achievements

The development and implementation of the GIS-based CA model demonstrated in this dissertation shows the achievement of the main aim of the research outlined in Section 1.4 in that a spatial model of urban growth for Bangkok area has been developed and implemented at the detail of district level, in order to be used to simulate the development of land use activities. The development of the model here will be useful to help map and locate the areas likely to experience urban growth. With the implementation under the GIS platform, the simulation results created can be visualized as a standard spatial map and are easily transferable to other GIS formats for further analysis. The accomplishment of the main aim was achieved through the establishment of objectives (Section 1.4), each of which is described below.

Firstly, the development of the GIS-based CA model here was integrated with multinomial logistic regression (MNL) and multi-criteria decision analysis (MCDA) methods, in order to identify the potential cells for development. While CA provides the approach for driving the simulation, GIS serve as a platform for operating urban growth simulation. The incorporation of MNL permits the observation of the true relationship between the land use characteristics and the land use categories obtained from the study area to be investigated without prior knowledge or predetermined outcomes of the system under study (Wu and Webster, 1998). The incorporation of MCDA allows a planning option from multiple points of planners and decision-makers to be tested (Malczewski, 1999a; 1999b).

Secondly, customized tools were developed using a VBA macro within the ArcGIS environment, in order to facilitate the implementation of the model. The graphical user interface and its execution are demonstrated in Chapter 5. The Customized tools developed here were intended to be user-friendly, possibly to be operated by unskilled users. Bangkok, like other the developing nations where GIS and computer knowledge of end-users are limited, the graphical user interface and a set of tools developed through a VBA macro here can enable them to interactively operate and test various scenarios in a simple way.

Thirdly, the proposed model was used to simulate urban development for Ladprao, a part of the Bangkok area, Thailand, at the detail of district level during the period 1993 – 2001 as shown in Chapter 6. Fourthly, validation of the performance of the proposed model was accomplished by comparing the simulated results produced by the model with the actual urban areas. The actual assessment allows the model to be used beyond a qualitative sense and provides confidence in using the results produced. Though the simulated results were not promising, discussion and suggestions to improve the model are given in Chapter 7.

Finally, sensitivity analysis using different neighbourhood sizes and different neighbourhood thresholds was also performed in order to investigate its impacts on the simulated results. The analysis was validated through the accuracy and morphology of urban areas. This examination allows us to improve the understanding of the model and be aware of the variation of the simulated results produced from the variation of CA elements.

8.5 Conclusion

High pressure on misuse of land development in the Bangkok area is a challenging problem for urban planners. Development of a simulation model is a solution to tackle the problem as it can help understand an urban structure and dictate the direction of growth. Such information can be incorporated with urban planning to be used to assess the consequences of previous and current planning policies which, in turn, to some degree aid in controlling or lessening the misuse of land.

In response to this issue, a combined GIS-based CA approach has been developed. The application has focused on the replication of the spatial pattern of Ladprao district, a part of the Bangkok area, Thailand during the period of 1993 – 2001. The unpromising results suggest several future research works, in order to improve the performance of the model and to improve the understanding of the Bangkok area. It is to be hoped that such spatial modelling will ensure effective spatial decision-making for the city of Bangkok in the future.

References

Agarwal, C., Green, G. L., Grove, M., Evans, T. and Schweik, C. (2000) *A Review and Assessment of Land-Use change models: Dynamics of space, Time, and Human Choice* the USDA Forest Service and Center for the Study of Institutions, Population, and Environmental Change(CIPEC), Idiana university.

Almeida, C. M. D., Monterio, A. M. V., Camara, G., Doares-Filho, B. S., Araujo, W. L. and Pantuzzo, A. E. (2003) Simulating urban land use change through CA-based modelling and logistic regression, *in* Simposio Brasileiro de Sensoriamento Remoto, Belo Horizonte, Br. Instituto Nacional de Pesquisas Espaciais, p. 1687-1694.

Anderson, T. R. and Egeland, J. A. (1961) Spatial Aspects of Social Area Analysis. *American Sociological Review*, 26(3): 392-398.

Arampatzis, G., Kiranoudis, C. T., Scaloubacas, P. and Assimacopoulos, D. (2004) A GIS-based decision support system for planning urban transportation policies. *European Journal of operational research*, 152: 465-475.

Archer, R. (1996) *Network Infrastructure for Sustainable Development: Implementing the Formula L + P + F + NI = SUD* Asian Institute of Technology, Bangkok.

Banai, R. (2005) Land Resource Sustainability for Urban Development: Spatial Decision Support System Prototype. *Environmental Management*, 36(2): 282-296.

Barredo, J. L., Demicheli, L., Lavalle, C., Kasanko, M. and McCormick, N. (2004) Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. *Environment and Planning B: Planning and Design*, 32: 65-84.

Barredo, J. L., Kasanko, M., McCormick, N. and Lavalle, C. (2003) Modelling dynamic spatial process: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning*, 64: 145-160.

Batty, M. (1998) Urban evolution on the desktop: simulation with the use of extended cellular automata. *Environment and Planning A*, 30(2): 1943-1967.

Batty, M. (2001) Models in planning: technological imperatives and changing roles. *JAG*, 3(3): 252-266.

Batty, M. (2003) *The Emergence of cities: complexity and urban dynamics Paper 64* CASA, University College London, London.

Batty, M. and Torrens, P. M. (2001) *Modeling complexity: the limits to prediction Paper 36* CASA, University College London, London.

Batty, M. and Xie, Y. (1994) From cells to cities. *Environment and Planning B*, 21: s31-s48.

Batty, M. and Xie, Y. (1997) Possible urban automata. *Environment and Planning B: Planning and Design*, 24: 175-192.

Beier, G. J. (1984) Can third World Cities Cope? In Ghosh, P. K. (Ed.), *Urban Development in the Third World*: 57-95. Westport, Connecticut: Greenwood Press.

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

Benenson, I. and Torrens, P. M. (2004) *Geosimulation: automata-based modeling of urban phenomena,* John Wiley & Sons, Chichester.

Benguigui, L. (1995) A new aggregation model. Application to town growth. *Physical A*, 219: 13-26.

Benguigui, L. (1998) Aggregation models for town growth. *Philosophical Magazine B*, 77(5): 1269-1275.

Bishop, I. D., Escobar, F. J., Karuppannan, S., Williamson, I. P., Yates, P. M., Suwarnrat, K. and Yaqub, H. W. (2000) Spatial data infrastructures for cities in developing countries: Lessons from the Bangkok experience. *Cities*, 17(2): 85-96.

Bivand, R. and Lucas, A. (2000) Integrating models and geographical information systems. In Openshaw, S. and Abrahart, R. J. (Eds.), *Geocomputation*: 331-363. London: Taylor & Francis.

Bojorquez-tapia, L. A., Diaz-mondragon, S. and Ezcurra, E. (2001) GIS-based approach for participatory decision making and land suitability assessment. *International Journal of Geographic Information Science*, 15(2): 129-151.

Boonchuen, P. (2002) Globalisation and urban design: Transformations of civic space in Bangkok. *IDPR*, 24(4): 401-417.

Briassoulis, H. (2000). Analysis of Land Use Change: Theoretical and Modeling Approaches. The Webbook of original Sciences: Regional Research Institute, West Virginia University. http://www.rri.wvu.edu/WebBook/Briassoulis/contents.htm. accessed 24 November 2006.

Bruijn, C. A. D. (1991) Modelling the Development of Asian cities using Airphotos and Remote Sensing data (Poster session), *in* The 12th Asian Conference on Remote Sensing (ACRS), Youth SEAMEO Regional Language centre 30, Singapore.

Bureau of Property Valuation of the Treasury Department (TRD) (2004). Bureau of Property Valuation webpage. Bangkok: Bureau of Property Valuation, the Treasury

Department (TRD). http://www.treasury.go.th/indexeng/Bureau/Bureau.htm. accessed 13 November 2007.

Carter, H. (1995) *The Study of Urban Geography,* Arnold, London.

Caruso, G., Rounsevell, M. and Cojacaru, G. (2005) Exploring a spatio-dynamic neighborhood-based model of residential behavior in the Brussels periurban area. *International Journal of Geographic Information Science*, 19(2): 103-123.

Carver, S. J. (1991) Integrating multi-criteria evaluation with geographic information systems. *International Journal of Geographic Information Systems*, 5(3): 321-339.

Ceballos-Silva, A. and Lopez-Blanco, J. (2003) Delineation of suitable areas for crops using a Multi-Criteria Evaluation approach and land use/cover mapping: a case study in central Mexico. *Agricultural Systems*, 77: 117-136.

Cecchini, A. and Rinaldi, E. (1996). The Multi-Cellular Automaton: A Tool To Build Sophisticated Models. A Theoretical Foundation and A Practical Implementation. www.erudit.de/erudit/events/esit99/12604.pdf. accessed 11 May 2006.

Chakhar, S. (2003) Enhancing Geographical Information Systems Capabilities with Multi-Criteria Evaluation Functions. *Journal of Geographical Information and Decision Analysis*, 7(2): 47-71.

Chapin, F. S. and Kaiser, E. J. (1979) *Urban Land Use Planning,* University of Illinois Press, Urbana.

Cheng, J. and Masser, I. (2004) Understanding spatial and temporal processes of urban growth: cellular automata modelling. *Environment and Planning B: Planning and Design*, 31: 167-194.

Choiejit, R. and Teungfung, R. (2005) Urban Growth and Commuting Patterns of the Poor in Bangkok, *in* Urban Research Symposium 2005 on "Land Development, Urban Policy and Poverty Reduction", Blue Tree Park Hotel, Brasilia, Brazil.

Chomchan, S., Silapacharanan, S., peerapun, W. and Chongpeerapien, T. (1990) *Land Use Development in Bangkok and Major Cities: Directions and Problems. Final report No. 2* Office of the National Economic and Social Development Board, Bangkok.

Clarke, G., Langley, R. and Cardwell, W. (1998) Empirical Applications of Dynamic Spatial Interaction Models. *Computers, Environment and urban Systems*, 22(2): 157-184.

Clarke, K. C. and Gaydos, L. (1998) Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12(7): 699-714.

Clarke, K. C., Hoppen, S. and Gaydos, L. (1997) A self-modifying cellular automaton model of historical urbanization in the Sanfrancisco Bay area. *Environment and Planning B: Planning and Design*, 24: 247-261.

Cohen, B. (2004) Urban Growth in Developing Countries: A Review of Current Trends and a Caution Regarding Existing Forecasts. *World Development*, 32(1): 23-51.

Commission of the European Communities (CEC) (1995) *A New Approach to Urban planning in Bangkok* Department of City Planning (DCP), BMA, Bangkok.

Core Planning & Development (2001) *District Development Plan Project: Ladprao district (in Thai) (Final Report),* Department of City Planning, Bangkok Metropolitan Administration, Bangkok.

Cromley, R. G. and Hanink, D. M. (1999) Coupling land use allocation models with raster GIS. *Journal of Geographical Systems*, 1: 137-153.

Darin-Drabkin, H. (1977) *Land Policy and Urban Growth,* Pergamon press, Oxford.

Department of City Planning (DCP) (1999) *Bangkok Comprehensive Plan (first revision) (in Thai)* Department of City Planning (DCP), Bangkok Metropolitan Administration, Bangkok.

Department of City Planning (DCP) (2002). Department of City Planning website. http://203.155.220.217/cpd/eng-about02.html. accessed 9 November 2006.

Department of Provincial Administration (DOPA) (2006) Office of Central Civil-Registration, Department of Provincial Administration, Ministry of Interior, Bangkok.

Department of Town and Country Planning (DTCP) (1997) *Standard and criteria for creation of comprehensive plan B.E. 2539 (in Thai)* Department of Town and Country Planning (DTCP),.

Dickey, J. W. (1975) *Metropolitan Transport Planning,* Scripta Book Company, Washington, D.C.

Dowall, D. E. (1989) Bangkok: A Profile of an Efficiently Perform Housing Market. *Urban Studies*, 26: 327-339.

Drakakis-Smith, D. (2000) *Third World Cities,* Routledge, London.

Eastman, J. R. (1999) *Guide to GIS and Image Processing,* Clark Lab, Worcester.

Engelen, G., White, R. and Uljee, I. (1997a) Integrating constrained cellular automata models, GIS and decision support tools for urban planning and policy-making. In H., T. (Ed.), *Decision Support Systems in Urban Planning*: 125-155. London: E&FN Spon.

Engelen, G., White, R. and Uljee, I. (1997b) The use of constrained cellular automata for high-resolution modelling of urban land use dynamics. *Environment and Planning B*, 24: 323-343.

Engelen, G., White, R., Uljee, I. and Drazan, P. (1995) Using Cellular Automata for Integrated Modelling of Socio-environmental Systems. *Environmental Monitoring and Assessment*, 34: 203-214.

Environmental Systems Research Institute (2001a) *Exploring ArcObjects Vol. 1- Application and Cartography,* Environmental Systems Research Institute, Inc., Redlands, CA.

Environmental Systems Research Institute (2001b) *Exploring ArcObjects Vol. 2- Geographic Data Management,* Environmental Systems Research Institute, Inc., Redlands, CA.

ESRI (2004). ArcGIS Desktop Help. Redlands,CA: ESRI Press. accessed 01 February 2007.

ESRI (Thailand) Co. Ltd. (1997) *Design and Development of GIS-based Planning Procedure (Executive Summary Report (in Thai)),* Department of Town and Country Planning (DTCP), Bangkok.

Feick, R. D. and Hall, G. B. (2004) A method for examining the spatial dimension of multi-criteria weight sensitivity. *International Journal of Geographic Information Science*, 18(8): 815-840.

Field, A. (2000) *Discovering Statistics using SPSS for Windows,* SAGE Publications, London.

Fischer, M. M. and Nijkamp, P. (1993) Design and use of geographic information systems and spatial models. In Fischer, M. M. and Nijkamp, P. (Eds.), *Geographic Information Systems: Spatial Modelling and Policy Evaluation*: 1-13. New York: Springer-Verlag.

Forrester, J. W. (1961) *Industrial Dynamics,* M. I. T. Press.

Fotheringham, A. S. and O'Kelly, M. C. (1988) *Spatial Interaction Models: Formulations and Applications,* Kluwer Academic Publishers, Dordrecht.

Garson, G. D. (2005). PA 765 Statnotes: An Online Textbook: Logistic Regression. Raleigh, North Carolina: North Carolina State University. http://www2.chass.ncsu.edu/garson/pa765/logistic.htm. accessed 14 March 2005.

Gomes, E. G. and Lins, M. P. E. (2002) Integrating Geographical Systems and Multi-Criteria methods: A Case Study. *Annals of Operations Research*, 116: 243-269.

Grabaum, R. and Meyer, B. C. (1998) Multicriteria optimization of landscapes using GIS-based functional assessment. *Landscape and Urban Planning*, 43: 21-34.

Griffin, E. and Ford, L. (1980) A Model of Latin American City Structure. *Geographical Review*, 70(4): 397-422.

Grossmann, W. D. and Eberhardt, S. (1993) Geographical information systems and dynamic modelling: Potential of a new approach. In Fischer, M. M. and Nijkamp, P. (Eds.), *Geographic Information Systems: Spatial Modelling and Policy Evaluation*: 169-180. New York: Springer-Verlag.

Guerra, G. A. V. (2004) Urban Development Issues in Thailand: An Insight into Bangkok. Master of Science In *Department of Infrastructure, Division of Building and Real Estate Economics,* Royal Institute of Technology, Stockholm.

Guhathakurta, S. (2001) *Urban Modelling as Storytelling: Using Simulation Models as a Narrative Paper 37* CASA, University College London, London.

Guitouni, A. and Martel, J.-M. (1998) Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, 109: 501-521.

Hajkowicz, S. A. and McDonald, G. T. (2000) An Evaluation of Multiple Objective Decision Support Weighting Techniques in Natural Resource Management. *Journal of Environmental Planning and Management*, 43(4): 505-518.

Hara, Y., Takeuchi, K. and Okubo, S. (2005) Urbanization linked with past agricultural landscape patterns in the urban fringe of a deltaic Asian mega-city: a case study in Bangkok. *Landscape and Urban Planning*, 73: 16-28.

Haynes, K. E. and Fotheringham, A. S. (1988) *Gravity and Spatial Interaction Models,* SAGE publications, Bavery Hills, California.

Herold, M., Goldstein, N. C. and Clarke, K. C. (2003) The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment*, 86: 286-302.

Herold, M., Menze, G. and Clarke, K. C. (2001) Remote Sensing and Urban Growth Models - Demands and Perspective, *in* Proceedings of the Symposium on Remote Sensing of Urban Areas, Regensburg, Germany.

Hester, D. (1999). Urban Dynamics of the Middle Rio Grande Basin. Denver, CO: US Geological Survey. http://rockyweb.cr.usgs.gov/mrgb/changetech.html. accessed 8 December 2006.

Hester, J. (1970) *Systems Models of Urban Growth and Development,* Urban Systems Laboratory, M. I. T. Press.

Heywood, I., Cornelieus, S. and Carver, S. (2002) *An Introduction to Geographical Information Systems,* Prentice Hall, New York.

Hosmer, D. W. and Lemeshow, S. (2000) *Applied Logisitc Regression,* John Wiley & Sons, Inc, New York.

Hosmer, D. W. and Lemeshow, S. (2001) *Solutions Manual to Accompany Applied Logistics Regression,* John Willey & Sons, Inc, New York.

Hu, Z. and Lo, C. P. (2007) Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and urban Systems,* doi:10.1016/j.compenvurbsys.2006.11.001.

Janssen, R. and Rietveld, P. (1990) Multicriteria Analysis and Geographical Information Systems: An Application to Agricultural Land Use in the Netherlands. In Scholten, H. J. and Stillwell, J. C. H. (Eds.), *Geographical Information Systems for Urban and Regional Planning*. Dordrecht: Kluwer Academic Publishers.

Jantz, C. A. and Goetz, S. J. (2005) Analysis of scale dependencies in an urban land-use change model. *International Journal of Geographic Information Science*, 19(2): 217-241.

Joerin, F. and Musy, A. (2000) Land management with GIS and multicriteria analysis. *International Transactions in Operational Research*, 7: 67-78.

Joerin, F., Theriault, M. and Musy, S. (2001) Using GIS and outranking multicriteria analysis for land-use suitability assessment. *Geographical Information Sicence*, 15(2): 153-174.

Jones, G. W. (1993) Industrialization and Urbanization in South-East Asia. In Brookfield, H. and Byron, Y. (Eds.), *South-East Asia's Environmental Future: the Search for Sustainability*: 47-71. Kuala Lumpur: Oxford University Press.

Kao, J.-J. and Lin, H.-Y. (1996) Multifactor Spatial for Landfill Siting. *Journal of Environmental Engineering*, 122(10): 902-908.

Kaothien, U. (1995) The Bangkok Metropolitan Region: Policies and Issues in the Seventh Plan. In McGee, T. G. and Robinson, I. M. (Eds.), *The mega-urban regions of Southeast Asia*: 328-340. Vancouver: UBC Press.

Kocabus, V. and Dragicevic, S. (2004). Sensitivity Analysis of A GIS-based Cellular Automata Model. XXth ISPRS Congress, Commission 4, 12-23 July 2004. Istanbul, Turkey: ISPRS. http://www.isprs.org/istanbul2004/comm4/papers/321.pdf. accessed 20 December 2006.

Krongkaew, M. (1996) The changing urban system in a fast-growing city and economy: The case of Bangkok and Thailand. In Lo, F.-c. and Yeung, Y.-m. (Eds.), *Emerging World Cities in Pacific Asia.*: 286-334. New York: United Nations University Press.

Laaribi, A., Chevallier, J. J. and Martel, J. M. (1996) A spatial decision aid: A multicriterion evaluation approach. *Computers, Environment and Urban Systems*, 20(6): 351-366.

Lavalle, C., Demicheli, L., Turchini, M., Casals, C. P. and Niederhuber, M. (2001) Monitoring mega-cities: the MURBANDY/MOLAND approach. *Development in Practice*, 11: 350-357.

Lee, C. (1973) *Models in Planning: An Introduction to the use of Quantitative Models in Planning,* Pergamon Press, Oxford.

Li, L., Sato, Y. and Zhu, H. (2003) Simulating spatial urban expansion based on a physical process. *Landscape and Urban Planning*, 64: 67-76.

Li, X. and Yeh, A. G. (2000) Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographic Information Science*, 14(2): 131-152.

Ligtenberg, A., Bregt, A. K. and Lammeren, R. v. (2001) Multi-actor-based land use modelling: spatial planning using agents. *Landscape and Urban Planning*, 56: 21-33.

Longley, P. A. and Batty, M. (2003) *Advanced Spatial Analysis,* ESRI Press, Redlands, CA.

Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (2005) *Geographic Information Systems and Science,* John Wiley & Sons Ltd, West Sussex.

Longley, P. A. and Mesev, V. (2000) On the measurement and generalisation of urban form. *Environment and Planning A*, 32(3): 473–488.

Lowry, I. S. (1965) A short course in model design. *Journal of the American Institute of Planners*, 31: 158-166.

Makropoulos, C. K. and Butler, D. (2006) Spatial ordered weighted averaging: incorporating spatially variable attitude towards risk in spatial multi-criteria decision-making. *Environmental Modelling & Software*, 21: 69-84.

Malczewski, J. (1999a) *GIS and Multicriteria Decision Analysis,* John Wiley & Sons, New York.

Malczewski, J. (1999b) Spatial Multicriteria Decision Analysis. In Thill, J.-C. (Ed.), *Spatial Multicriteria Decision Making and Analysis*. Aldershot: Ashgate.

Malczewski, J. (2006) GIS-based multicriteria decision analysis: a survey of the literature. *International Journal of Geographic Information Science*, 20(7): 703-726.

Marinoni, O. (2004) Implementation of the analytical hierarchy process with VBA in ARCGIS. *Computers & Geosciences*, 30: 637-646.

Martin, D. and Wu, F. (1999) Empirical CA simulation from a high resolution population surface, *in* Proceedings of the 4th International Conference on GeoComputation, Mary Washington College Fredericksburg, Virginia, USA.

McGee, T. G. (1991) The Emergence of Desakota Regions in Asia: Expanding a Hypothesis. In Ginsburg, N., Koppel, B. and McGee, T. G. (Eds.), *The Extended metropolis: settlement Transition in Asia*: 3-25. Honolulu: University of hawaii Press.

McGee, T. G. (1994) Labour Force Change mobility in the Extended Metropolitan Regions of Asia. In Fuchs, R. J., Brennan, E., Chamie, J., Lo, F.-C. and Uitto, J. L. (Eds.), *Mega-City Growth and the Future*: 62-101. New York: United Nations University Press.

McHarg, I. L. (1971) *Design with nature,* Natural History Press, New York.

Mcmillen, D. P. (1989) An empirical model of urban fringe land use. *Land Economics*, 65: 138-145.

Merwe, V. d. and Hendrik, J. (1997) GIS-aided land evaluation decision-making for regulating urban expansion: A South African case study. *GeoJournal*, 43: 135-151.

Myung, I. J. (2003) Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47: 90-100.

National Economics and Social Development Board (NESDB) (2001) NESDB, Thailand, Bangkok.

Nijkamp, P., Rietveld, P. and Voogd, H. (1990) *Multicriteria evaluation in physical planning,* North-Holland.

Noth, M., Borning, A. and Waddell, P. (2000). An extensible, modular architecture for simulating under development, transportation, and environmental impacts. http://www.urbansim.org/papers. accessed 11 December 2003.

Openshaw, S. (1978) *Using Models in Planning: a practical guide,* J. & P. Bealls, Newcastle upon Tyne.

O'Sullivan, D. and Torrens, P. M. (2000) Cellular Models of Urban Systems, *in* Proceedings of the Fourth International Conference on Cellular Automata for Research and Industry: Theoretical and Practical Issues on Cellular Automata, Karlsruhe, Germany, p. 108-116.

Ottens, H. F. L. (1990) The Application of Geographic Information Systems in Urban and regional Planning. In Scholten, H. J. and Stillwell, J. C. H. (Eds.), *Geographic Information Systems for Urban and Regional Planning*: 15-22. Dordretch: Kluwer Academic Publishers.

Phua, M.-H. and Minowa, M. (2005) A GIS-based multi-criteria decision making approach to forest conservation planning at a landscape scale: a case study in the Kinabalu Area, Sabah, Malaysia. *Landscape and Urban Planning*, 71: 207-222.

Pimcharoen, O. (2001) The planning system and environmental management in Thailand: the role of Environmental Impact assessment (EIA) in the decision making process. Doctoral Dissertation, University of Liverpool, Liverpool, UK.

Potter, R. (1992) *Urbanisation in the Third World,* Oxford University Press, Oxford.

Powers, D. A. and Xie, Y. (2000) *Statistical Methods for categorical Data Analysis,* Academic Press, London.

Punpuing, S. (1993) Correlates of Commuting Patterns: A Case-study of Bangkok, Thailand. *Urban Studies*, 30: 527-546.

Robinson, G. M. (1998) *Methods & Techniques in Human Geography,* John Wiley and Sons, Chichester.

Rodrigue, J.-P., Comtois, C. and Slack, B. (2006) *The Geography of Transport Systems,* Hofstra University, Department of Economics & Geography.

Ross, H., Poungsomlee, A., Punpuing, S. and Archavanitkul, K. (2000) Integrative analysis of city systems: Bangkok "Man and the Biosphere" programme study. *Environment&Urbanization*, 12(2): 151-161.

Saaty, T. L. (1980) *The analytic hierarchy process,* McGraw-Hill, New York.

Saaty, T. L. and Alexander, J. M. (1981) *Thinking with Models: Mathematical Models in the Physical, Biological, and Social Sciences,* Pergamon Press, Oxford.

Scholten, H. J. and Stillwell, J. C. H. (1990) Geographical Information System: The Emerging Requirements. In Scholten, H. J. and Stillwell, J. C. H. (Eds.), *Geographic Information Systems for Urban and Regional Planning*: 3-14. Dordretch: Kluwer Academic Publishers.

Sener, B., Suzen, M. L. and Doyuran, V. (2006) Landfill site selection by using geographic information systems. *Environmental Geology*, 49: 376-388.

Sharkawy, M. A. and Chotipanich, S. (1998) Housing Segmentation in Developing Countries in Transition: A Recent Case study of Residential Development in Bangkok. *Journal of Real Estate Literature*, 6: 29-42.

Shieh, Y. N. (1987) A Model of Residential Location in a Two-Dimension, Multi-Center City. *Papers of the Regional Science Association*, 63: 71-79.

Silva, E. A. and Clarke, K. C. (2005) Complexity, Emergence and Cellular Urban Models: Lessons Learned from Applying Sleuth to Two Portuguese Metropolitan Areas. *European Planning Studies*, 13(1): 93-116.

Sjoberg, G. (1960) *The Preindustrial City: Past and Present,* Free Press, New York.

Soares-Filho, S. B., Cerqueira, G. C. and Pennachin, C. L. (2002) DINAMICA - a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling*, 154(3): 217-235.

SPSS Inc. (2001) *Advanced Techniques: Regression,* SPSS Inc., Chicago, Illinois.

Stutz, R. and Souza, A. d. (1998) *The World Economy: Resources, Location, Trade and Development,* Prentice Hall, Toronto.

Sui, D. Z. and Zeng, H. (2001) Modeling the dynamics of landscape structure in Asia's emerging desakota regions: a case study in Shenzhen. *Landscape and Urban Planning*, 23(1-4): 37-52.

Tachizuka, S., Hung, T., Ochi, S. and Yasuoka, Y. (2002) Monitoring of Long-term Urban Expansion by the Use of Remote Sensing Images from Different Sensors, *in* The 23rd Asian Conference on Remote Sensing (ACRS), Birendra International Convention Centre, Kathmandu, Nepal.

Teerarojanarat, S., Fairbairn, D. J. and Chunithipaisan, S. (2004) Urban growth simulation with UrbanSim, *in* Proceedings of the FOSS/GRASS Users Conference, Bangkok, Thailand.

The National Housing Authority (NHA) (2003) The National Housing Authority (NHA), Bangkok.

Thomson, C. N. and Hardin, P. (2000) Remote sensing/GIS integration to identify potential low-income housing sites. *Cities*, 17(2): 97-109.

Thong, H. and Tsang, V. (1990) *Property Opportunities in Bangkok Report No. 3* Jones Lang Wootton, Bangkok.

Tkach, R. J. and Simonvic, S. P. (1997) A new approach to multicriteria decision making in water resources. *Journal of Geographical Information and Decision Analysis*, 1(1): 25-43.

Tobler, W. R. (1979) Cellular geography. In Gale, S. and Olsson, G. (Eds.), *Philosophy in Geography*: 379-386. Dordrecht: Reidel.

Todaro, M. P. (1984) Urbnization in Developing Nations: trends, Prospects, and Policies. In ghosh, P. K. (Ed.), *Urban Development in the Third World*: 7-26. Westport, Connecticut: Greenwood Press.

Tomlin, C. D. (1990) *Geographic Information Systems and Cartographic Modelling*, Prentice Hall, New Jersy.

Torrens, P. M. (2000) *How Cellular Models of Urban Systems Work (1. Theory) Paper 28* CASA, University College London, London.

Torrens, P. M. (2003) Automata-based models of urban-systems. In Longley, P. A. and Batty, M. (Eds.), *Advanced Spatial Analysis*. Redlands: ESRI Press.

Torrens, P. M. (2006) Simulating Sprawl. *Annals of the Association of American Geographers*, 96(2): 248-275.

Torrens, P. M. and O'Sullivan, D. (2000) Cities, cells, and complexity: developing a research agenda for urban geocomputstion, *in* Proceedings of the Fifth Annual Conference on GeoComputation, Manchester.

United Nations (1997) *World Urbanization Prospects; The 1996 Revision*, United Nations, New York.

United Nations (2006) *World Urbanization Prospects: The 2005 Revision* United Nations (Population Division of the Department of Economic and Social Affairs), New York.

Villa, F., Ceroni, M. and Mazza, A. (1996) A GIS-based method for multi-objective evaluation of park vegetation. *Landscape and Urban Planning*, 35: 203-212.

Voogd, H. (1983) *Multicriteria Evaluation for Urban and Regional Planning*, Page Bros (Norwich) Limited, London.

Waddell, P. (1998) An Urban Simulation Model for Integrated Policy Analysis and Planning: Residential Location and Housing Market Components of UrbanSim, *in* 8th World Conference on Transport Research, Antwerp, Belgium.

Waddell, P. (2000a) A behavioural simulation model for metropolitan policy analysis and planning: residential and housing components of UrbanSim. *Environment and Planning B*, 27: 247-263.

Waddell, P. (2000b) Simulating Land capacity at the Parcel Level. In Moudon, A. V. and Hubner, M. (Eds.), *Monitoring Land Supply with Geographic Information System*: 201-214. New York: John Wiley & Sons.

Waddell, P. (2002). UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. http://www.urbansim.org/papers. accessed 11 December 2003.

Waddell, P. (2006). Reconciling Household Residential Location Choices and Neighborhood Dynamics. http://www.urbansim.org/papers. accessed 12 December 2006.

Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M. and Ulfarsson, G. E. (2001). Microsimulation of urban development and location choices: Design and implementation of UrbanSim. http://www.urbansim.org/papers. accessed 11 December 2003.

Waddell, P., Sevcikova, H., Socha, D., Miller, E. and Nagel, K. (2005). OPUS: An Open Platform for Urban Simulation. accessed 12 December 2006.

Waddell, P. and Ulfarsson, G. F. (2002a). Accessibility and Agglomeration: Discrete-Choice Models of Employment Location by Industry Sector. http://www.urbansim.org/papers. accessed 10 February 2004.

Waddell, P. and Ulfarsson, G. F. (2002b). Dynamic Simulation of Real Estate Development and Land Prices within an Integrated Land Use and Transportation Model System. www.urbansim.org/papers/ waddell-ulfarsson-TRB03-real-estate.pdf. accessed 25 November 2004.

Waddell, P. and Ulfarsson, G. F. (2004) Introduction to Urban Simulation: Design and Development of Operational Models. In Stopher, B. and Kingsley, H. (Eds.), *Handbook in Transport, Volume 5: Transport Geography and Spatial Systems*: ,203-236.: Pergamon Press.

Wagner, D. (1997) Cellular automata and geographic information systems. *Environment and Planning B*, 24: 219-234.

Walford, N. (2001) *Geographical Data: Characteristics and Sources*, John Wiley & Sons, Chichester.

Ward, D. P., Murray, A. T. and Phinn, S. R. (2000) A stochastically constrained cellular model of urban growth. *Computers, Environment and Urban Systems*, 24: 539-558.

Watkins, T. (2006). Alonso's Bid Rent Function Theory. ECONOMICS DEPARTMENT, SAN JOSE STATE UNIVERSITY. http://www2.sjsu.edu/faculty/watkins/alonso.htm. accessed 11 June 2006.

Webster, D. (2000) *Financing City-Building: The Bangkok Case* Stanford Asia Pacific Research Center, Stanford University Press.

Webster, D. (2004) Bangkok: evaluation and adaptation under stress. In Gugler, J. (Ed.), *World Cities Beyond the West: Globalization, Development and Inequality*: 82-117. Cambridge: Cambridge University Press.

White, R., Engelen, G., Uljee, I., Lavalle, C. and Ehrlich, D. (2000) Developing an Urban Land Use Simulator for European Cities, *in* Proceedings of the 5th EC GIS Workshop: GIS of Tomorrow, p. 179-190.

White, R. W. and Engelen, G. (1993) Cellular automata and fractal urban form: a cellular modelling approach to the evolution of land use patterns. *Environment and Planning A*, 25(8): 1175-1199.

Wilson, A. G. (1974) *Urban and Regional Models in Geographic and Planning,* John Wiley & Sons, London.

Wilson, A. G. (2000) *Complex Spatial Systems: The Modelling Foundations of Urban and regional Analysis,* Prentice Hall, Harlow, England.

Wilson, A. G. (2002) Complex Spatial Systems: Challenges for Modellers. *Mathematical and computer Modelling*, 36: 379-387.

Wu, F. (1998) SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographic Information Science*, 12(1): 63-82.

Wu, F. (1999) GIS-based simulation as an exploratory analysis for space time process. *Journal of Geographical Systems*, 1: 199-218.

Wu, F. (2000) A Parameterised Urban Cellular Model Combining Spontaneous and Self-organising Growth. In Atkinson, P. and Martin, D. (Eds.), *Geocomputation: Innovation in GIS 7*: 73-85. London: Taylor&Francis.

Wu, F. (2002a) Calibration of stochastic cellular automata: the application to rural-urban land conversions. *International Journal of Geographic Information Science*, 16(8): 795-818.

Wu, F. (2002b) Geo-referencing social spatial data and intra-urban property price modelling in a data-poor context: the case study of Shanghai property data. In

Kidner, D., Higgs, G. and White, S. (Eds.), *Socio-Economic Applications of Geographic Information Science*: 163-178. London: Taylor & Francis.

Wu, F. and Webster, C. J. (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design*, 25: 103-126.

Wu, F. and Webster, C. J. (2000) Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. *International Journal of Geographic Information Science*, 14(7): 625-648.

Yang, X. and Lo, C. P. (2003) Modelling urban growth and landscape changes in the Atlanta metropolitan area. *Geographic Information Science*, 17(5): 463-488.

Yeh, A. G. and Li, X. (1998) Sustainable land development model for rapid growth areas using GIS. *International Journal of Geographic Information Science*, 12(2): 169-189.

Yeh, A. G.-O. (1999) Urban Planning and GIS. In Longley, P. A., Goodchild, M. F., Maguire, D. W. and Rhind, D. W. (Eds.), *Geographic Information Systems: management issues and applications*: 877-888. New York: John Wiley & Sons.

Yeh, A. G. O. and Li, X. (2001) A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. *Environment and Planning B: Planning and Design*, 28: 733-753.

Yeh, A. G.-O. and Li, X. (2002) A cellular automata to simulate development density for urban planning. *Environment and Planning B: Planning and Design*, 29: 431-450.

Zeleny, M. (1982) *Multiple Criteria Decision Making,* McGraw-Hill, New York.

APPENDIX 1

Installation Guide

This section will help you install the training data and an ArcGIS map document (*.mxd) used for implementing the application. Also, the contents in the working folder (directory) are explored.

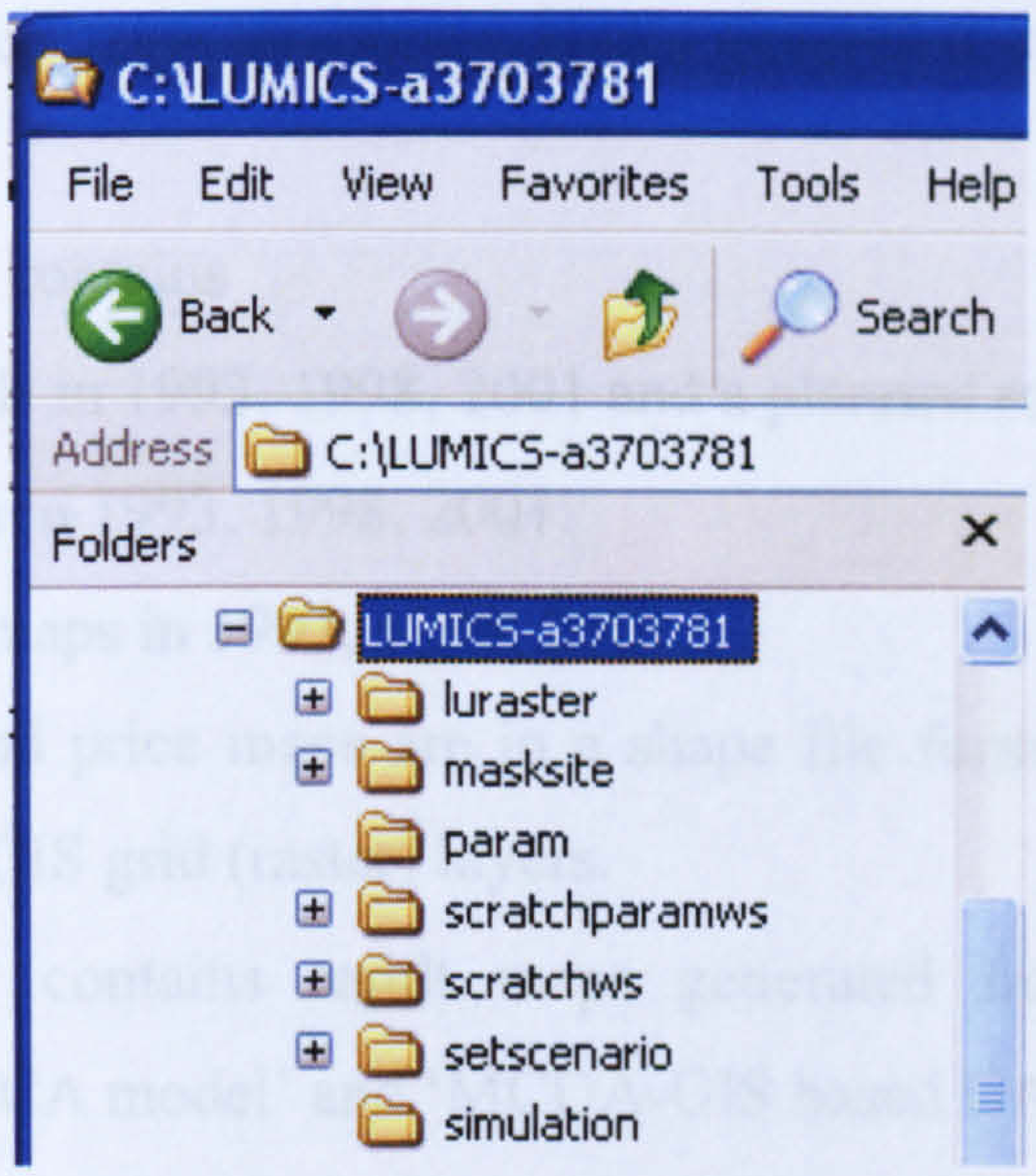
Step 1: Installing the training data and the ArcGIS map document (*.mxd)

1.1. Insert the CD-ROM into the CD-ROM drive.
1.2. From the Start menu, open Windows Explorer
1.3. Navigate through the tree structure to the CD-ROM and click on the CD-ROM
1.4. In the contents of the CD-ROM, copy folder "LUMICS-a3703781" into your hard drive "C:".

Now, the training data and ArcGIS map document (*.mxd) is installed in the destination directory: "C:\ LUMICS-a3703781" From now on, this folder, namely *workspace* (if it contains geographic data) in ArcGIS, will be referred to as your working folder.

Step 2: Explore the working directory (C:\ LUMICS-a3703781)

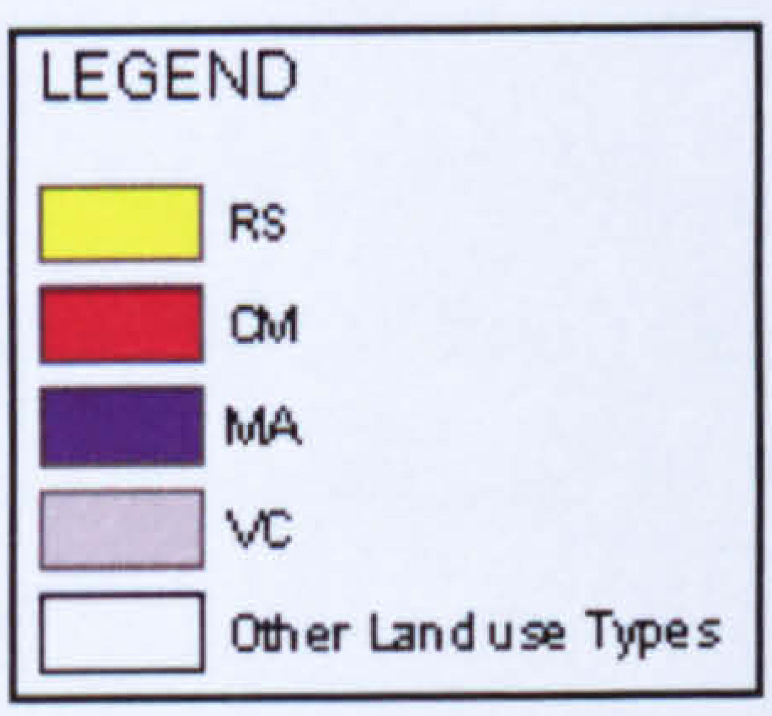
From the Windows Explorer, navigate to the working folder "C:\ LUMICS-a3703781" and expand the listings.



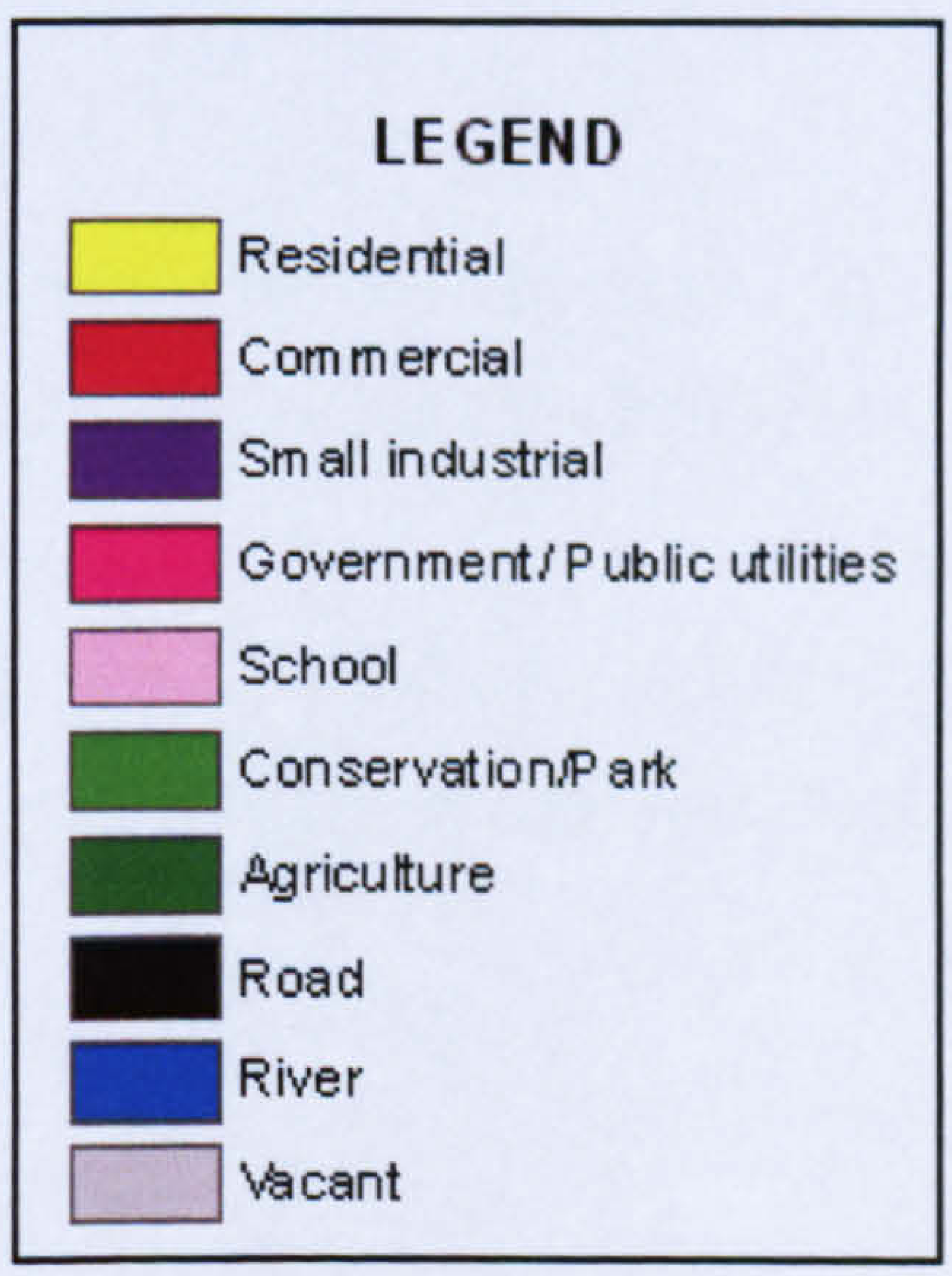
Here, they contain the following folders and files:

- Folder 'luraster' contains two land use maps in 1993 and 2001, stored in the format of ArcGIS grid (raster) layers.
- Folder 'masksite' contains a masking map that is used to extract four land use types: residential, commercial, industrial and vacant land.
- Folder 'param' contains two layer files used for symbolizing land use classes in this study as shown below.

Filename: display4luwhite.lyr



Filename: lucode10.lyr



- Folder 'scratchparamws' contains result maps generated from a customized tool: 'Derive parameter'.

- Folder 'scratchws' contains temporary layers generated during the execution of the model.
- Folder 'setscenario' contains
  - Road maps: roads in 1993, 1998, 2001 and a planned road.
  - Land price maps in 1993, 1998, 2001.
  - Constraint area maps in 1993, 1998, 2001.

    All road and land price maps are in a shape file format while all constraint areas are in ArcGIS grid (raster) layers.
- Folder 'simulation' contains result maps generated from customized tools: 'MCDA-GIS based CA model' and 'MCDA-GIS based CA model'.
- File 'lumics_thesis.mxd' is used for running the application within an ArcGIS software environment.

Now you have finished installing and exploring the necessary data used for implementing the application. To execute the application, continue to Appendix B.

End Installation Guide

APPENDIX 2

User's Guide

In this section, you will experiment with a set of tools developed (named LUMICS) for this research study. The customized tools developed consist of the Variables Observation tool, MNL (GIS-based CA approach) tool, and MCDA (GIS-based CA approach) tool. This exercise will guide you how to use these tools. Note that these tools require ArcGIS 9.2 with the extension of spatial analyst module installed in your hard drive. Furthermore, all training data and the ArcGIS software (including spatial analyst module) need to be installed beforehand. To install the program and training data, please go to the Installation Guide in Appendix 1.

Step1: Starting ArcGIS 9.2 and opening an ArcGIS map document

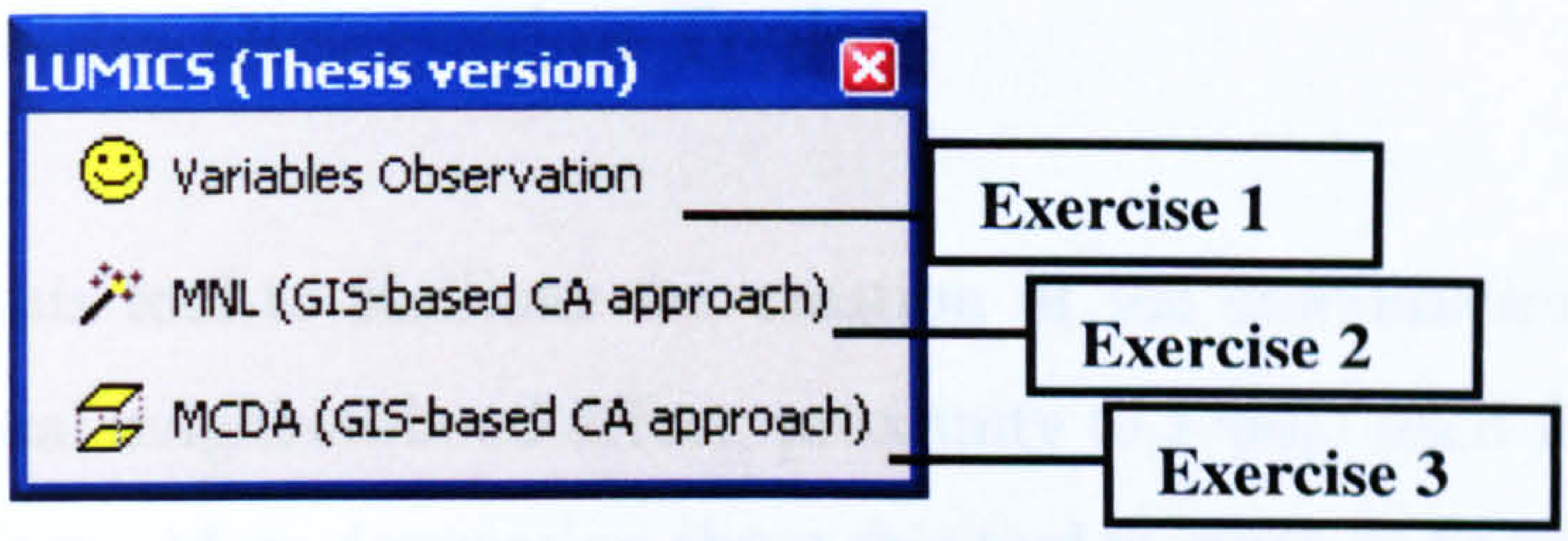
For accessing the set of tools developed, you will begin by opening LUMICS application.

1.1. Start ArcGIS 9.2.

1.2. When the startup dialog appears, choose to open an existing file "lumics_thesis.mxd" locating under "C:\ LUMICS-a3703781".

Note that, you can also directly open the LUMICS application from Windows Explorer by double-click filename "lumics_thesis.mxd" (an ArcGIS map document) locating under "C:\ LUMICS-a3703781".

At this step, a set of tools should now appear.



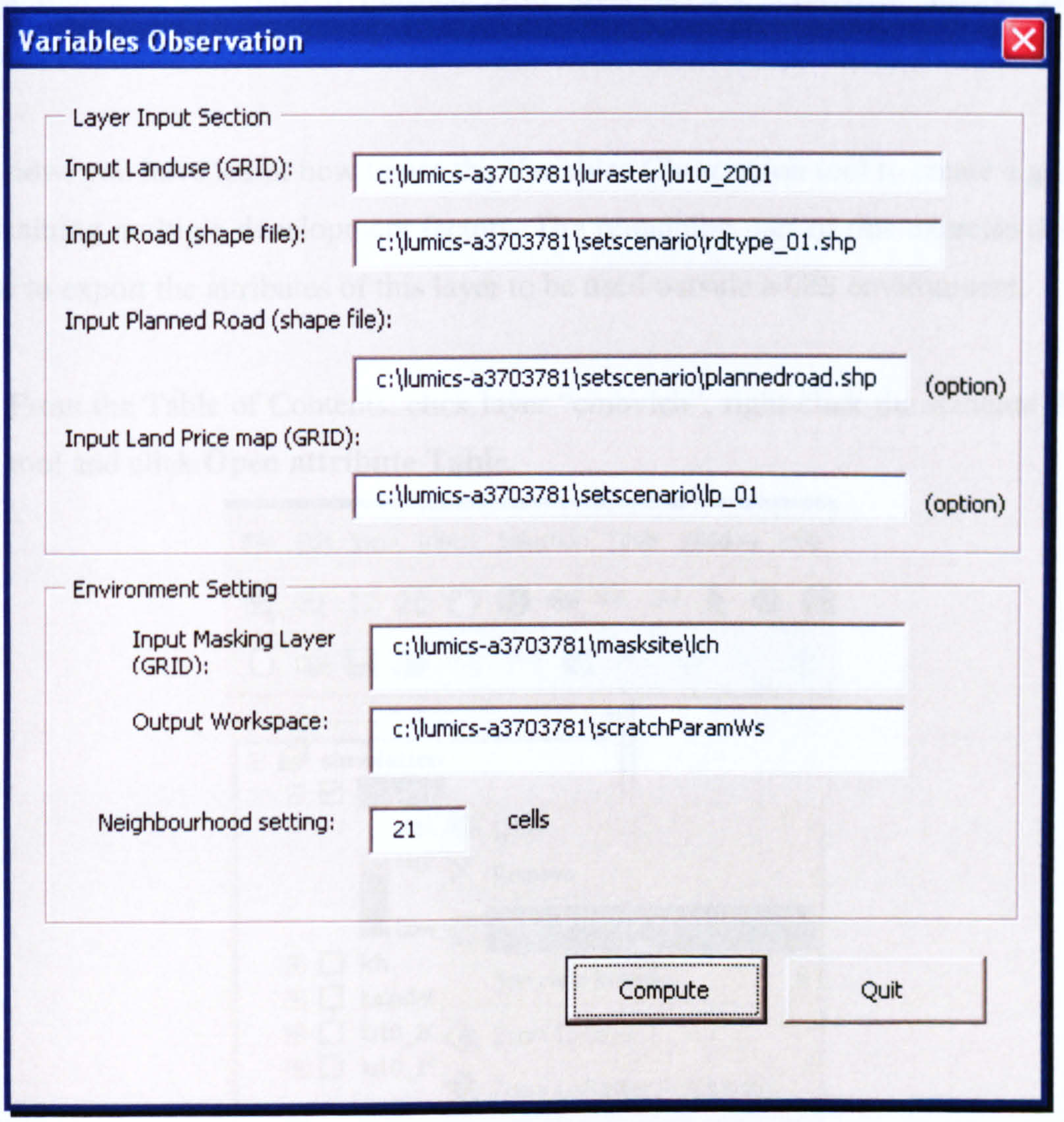
Step 2: Experimenting with these tools

To experiment with the Variables Observation tool, go to Exercise 1. For using MNL (GIS-based CA approach) tool and MCDA (GIS-based CA approach) tool, go to Exercise 2 and 3 respectively.

Exercise 1: Variables Observation Tool

Now you will use this tool to facilitate the creation of the combination of development factors (e.g. residential neighbourhood effect, proximity to roads) used for the exploration of factors in the analysis. More description about this tool is given in Section 5.3.1.1.

1. From LUMICS (Thesis version) menu, click **Variables Observation**. The Variables Observation dialog box should now pop-up.



2. The tool provides input (layer input and environment setting) ready to be executed. You may change these inputs (e.g. the neighbourhood size) by typing in a new value. In this exercise, however, you do not need to change these input values.
3. Click **Compute** to run tool.

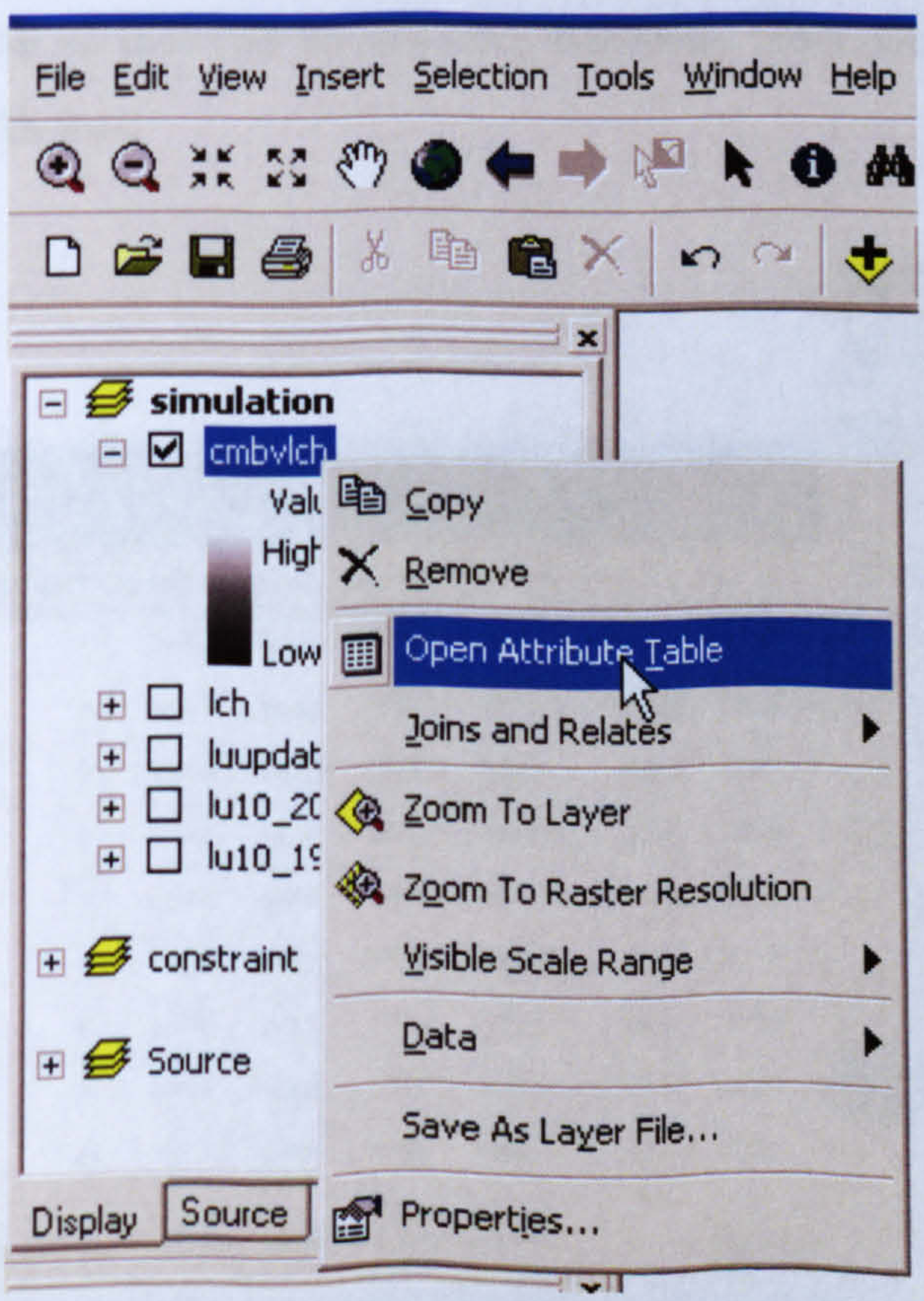
After finishing, the message box should appear as "Job about deriving parameters done...Bye Bye".

4. Click **Quit** to exit tool.

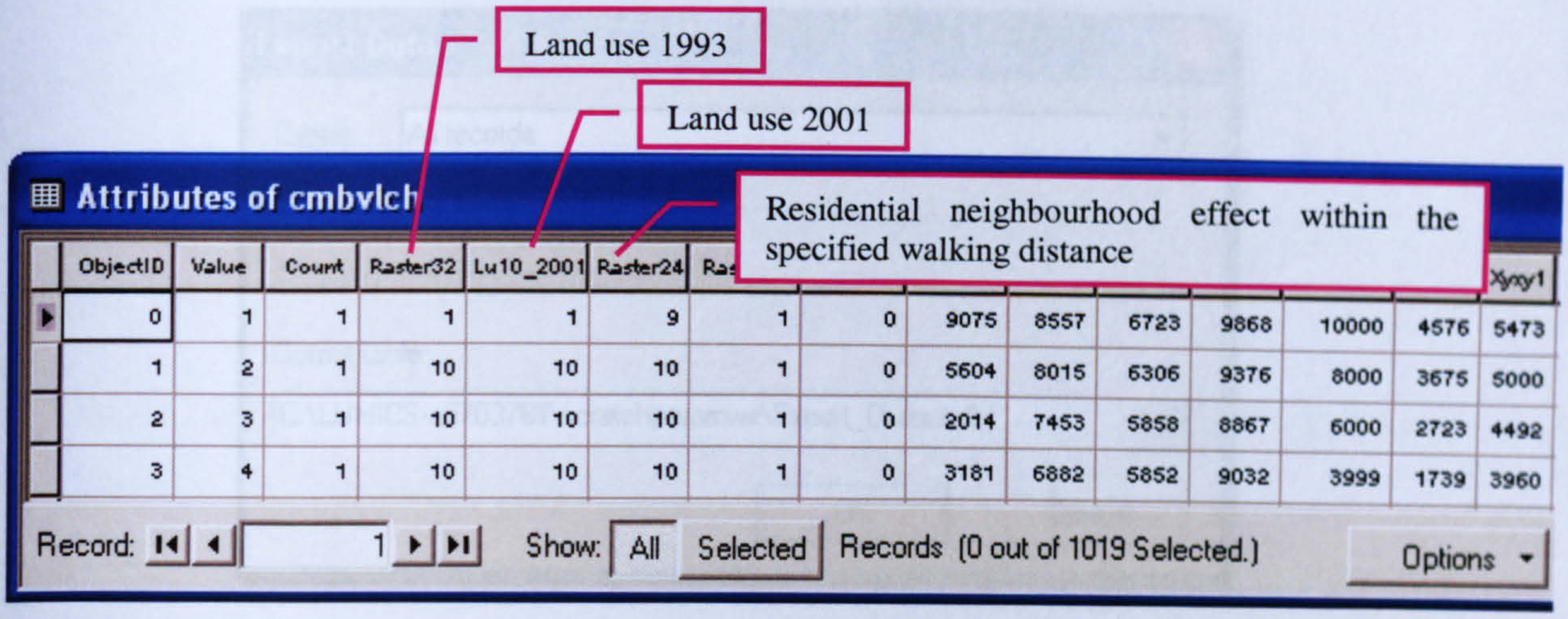
The result of this tool is a grid (raster) layer containing a series of development factors. A description of this layer is presented in Table 5.1. This resultant layer is stored in a grid layer named "cmbvlch" in the folder "C:\ LUMICS-a3703781\scratchparamws". This tool also generates a transient grid layer named "cmbvlch" in the Table of Contents.

By now, you have learnt how to use the Variables Observation tool to create a grid layer containing multiple development factors. The remaining part of this exercise describes how to export the attributes of this layer to be used outside a GIS environment.

5. From the Table of Contents, click layer "cmbvlch", right-click the standard ArcGIS tool and click **Open attribute Table**.

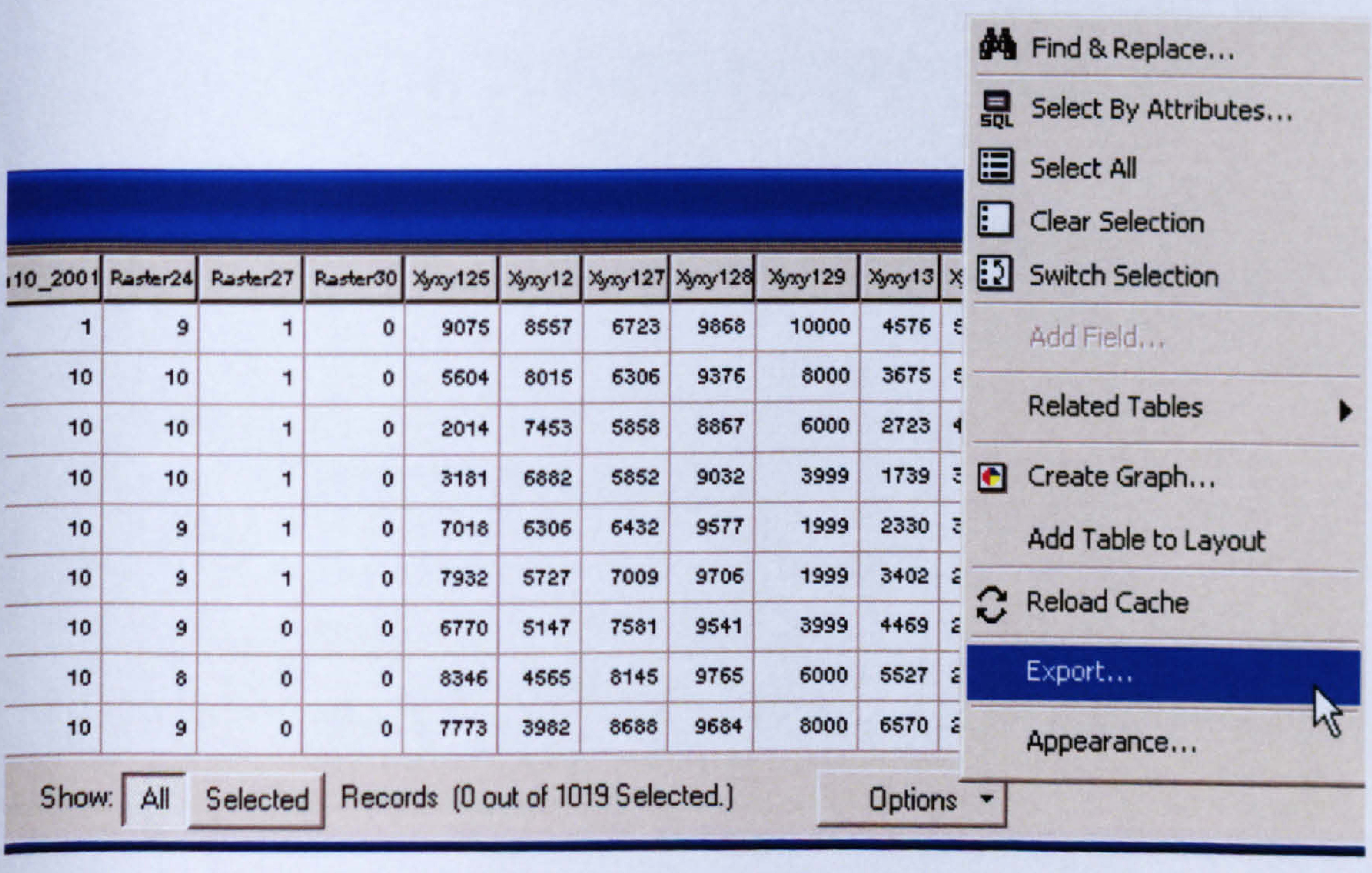


Now, the associated attribute table should appear.

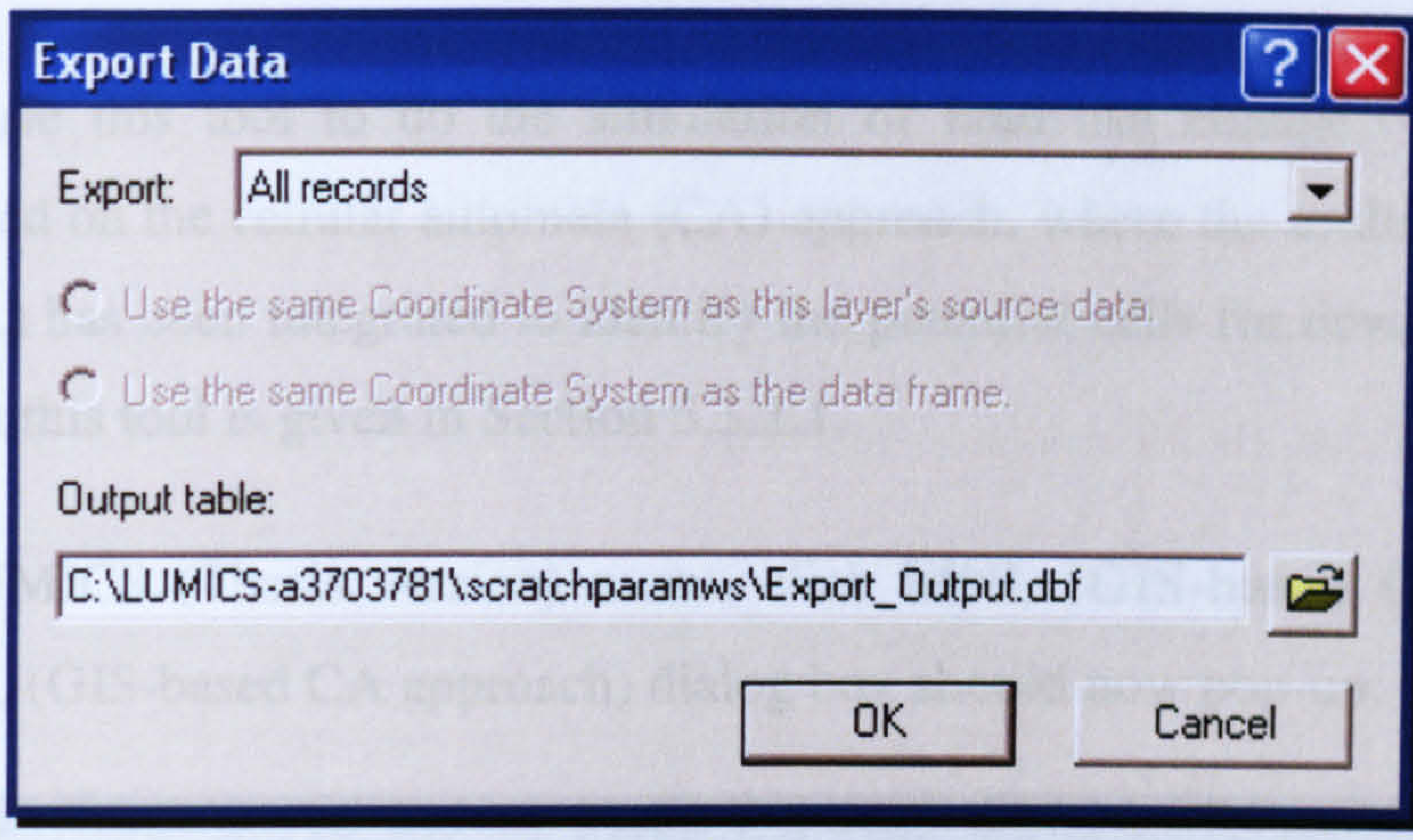


From the attribute table, many fields are generated. The first three fields (objectID, Value, and Count) refer to the standard fields provided by ArcGIS. The rest are fields generated by the Variables Observation tool. The order of these fields corresponds to that of the fields listed in Table 5.1. Note that the names of the fields in this version may change for each run, but the order of the fields corresponds to those listed in Table 5.1.

6. Export this table as text file by clicking **Options**. Then click **Export...** from the Standard ArcGIS tool.



The Export Data dialog box should appear.



7. Type in the name of database file (.dbf). Then click **OK**.

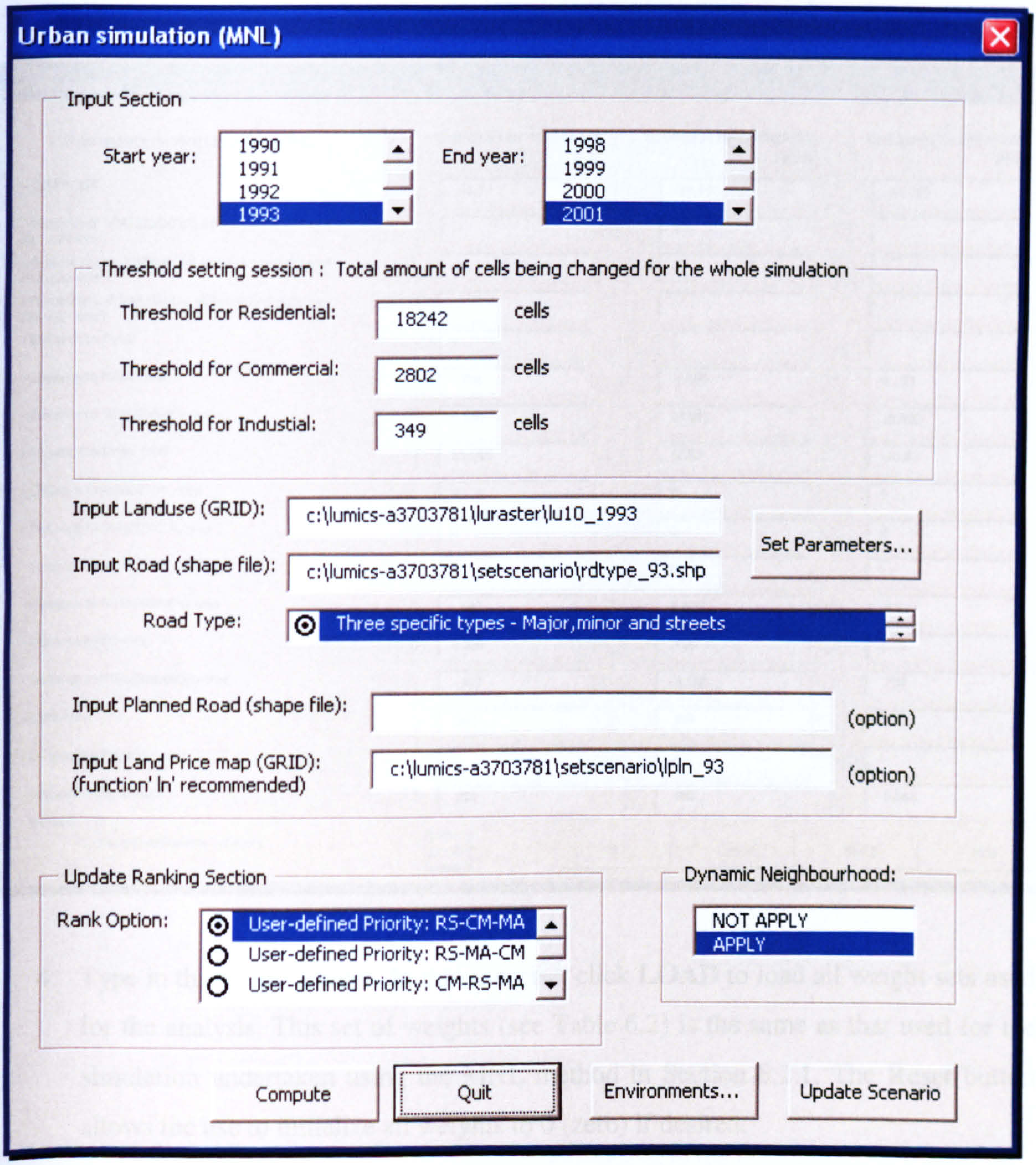
Now you have learnt how to create a text file produced from the Variables Observation tool. In this study, the text file generated will be transferred to the statistical software package (e.g. SPSS) to perform derivation of criterion weights by means of multinomial logistic regression (MNL). More details about the way to derive criterion weights using the MNL method is given in Section 4.5.1.

END Exercise 1

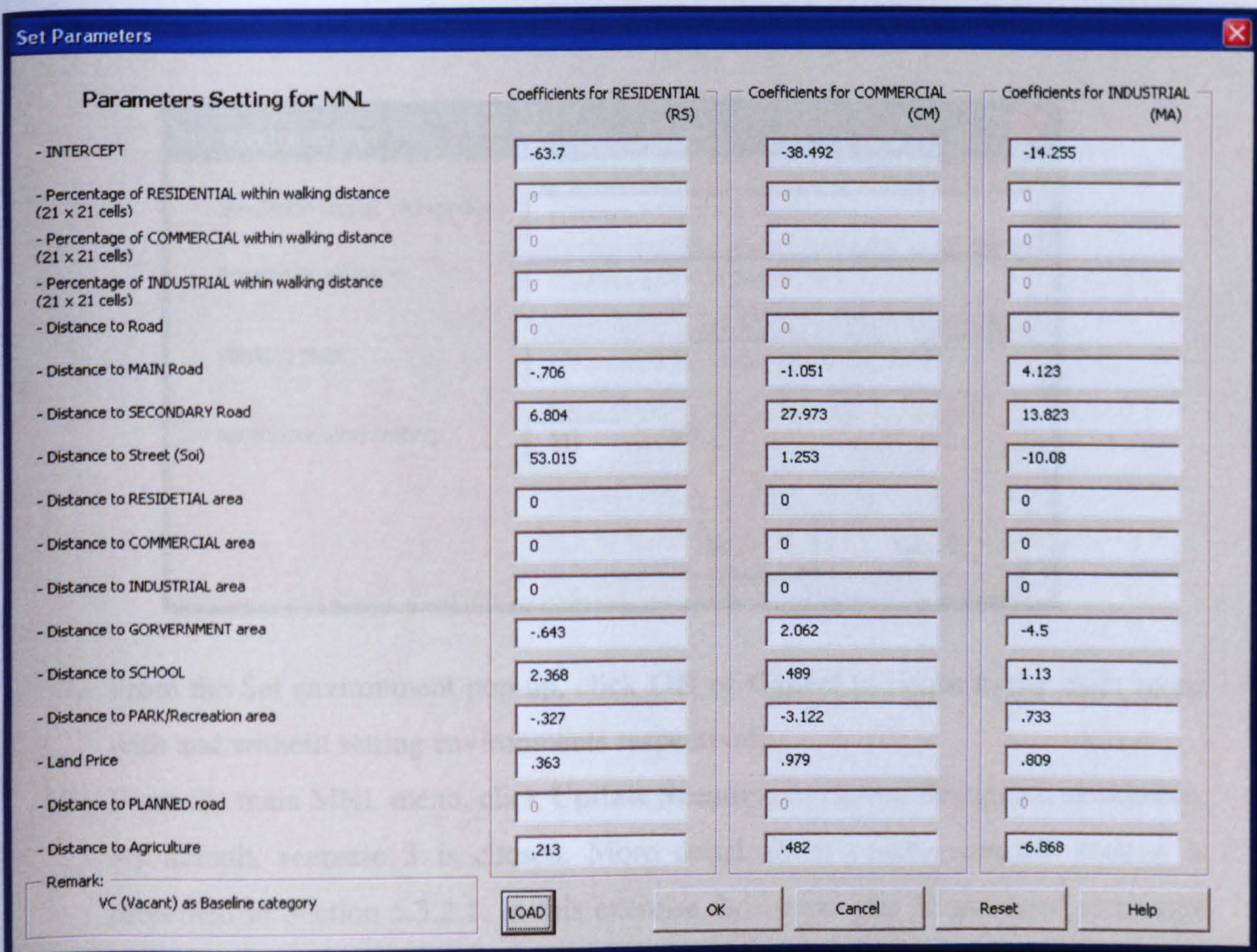
Exercise 2: MNL (GIS-based CA approach) Tool

Now you will use this tool to do the simulation of land use change. The simulation performed is based on the cellular automata (CA) approach, where the multinomial logistic regression (MNL) has been integrated to identify the potential cells for development. More description about this tool is given in Section 5.3.2.1.

1. From LUMICS (Thesis version) menu, click **MNL (GIS-based CA approach)**. The MNL (GIS-based CA approach) dialog box should now pop-up.



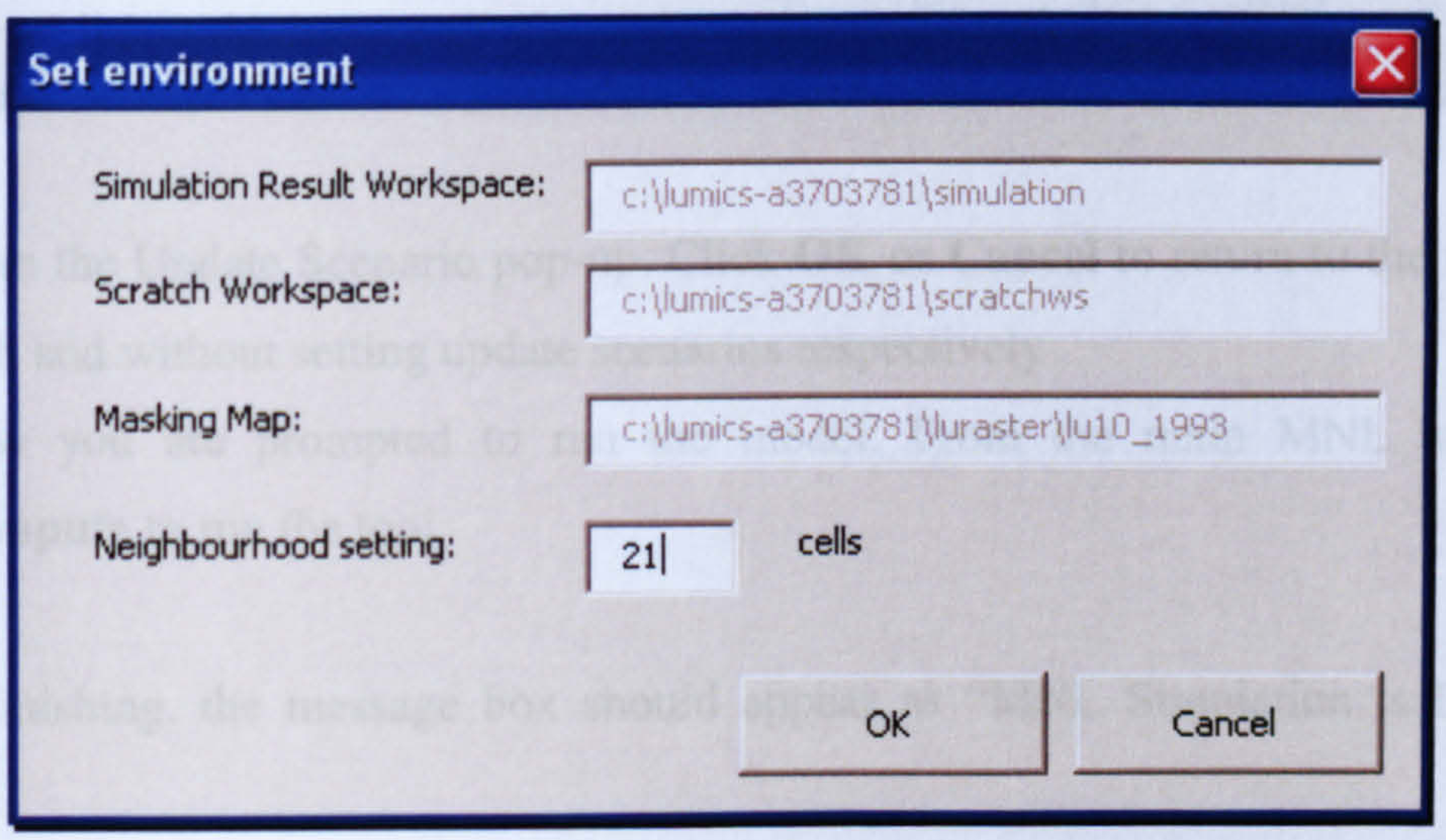
2. The tool provides input (e.g. simulation period, amount of cell allowed to change for the whole simulation, layer input) ready to be executed. You may change these inputs by typing in a new value (e.g. threshold for residential) or choose the multiple choice provided (e.g. simulation period, update ranking option). More details about sections (e.g. update ranking section, dynamic neighbourhood section) in this page are given in Section 5.3.2.1. In this exercise, however, you do not need to change these input values.
3. From the main MNL menu, click **Set Parameters...** to set weights used in the analysis. The Set Parameters pop-up should now appear.



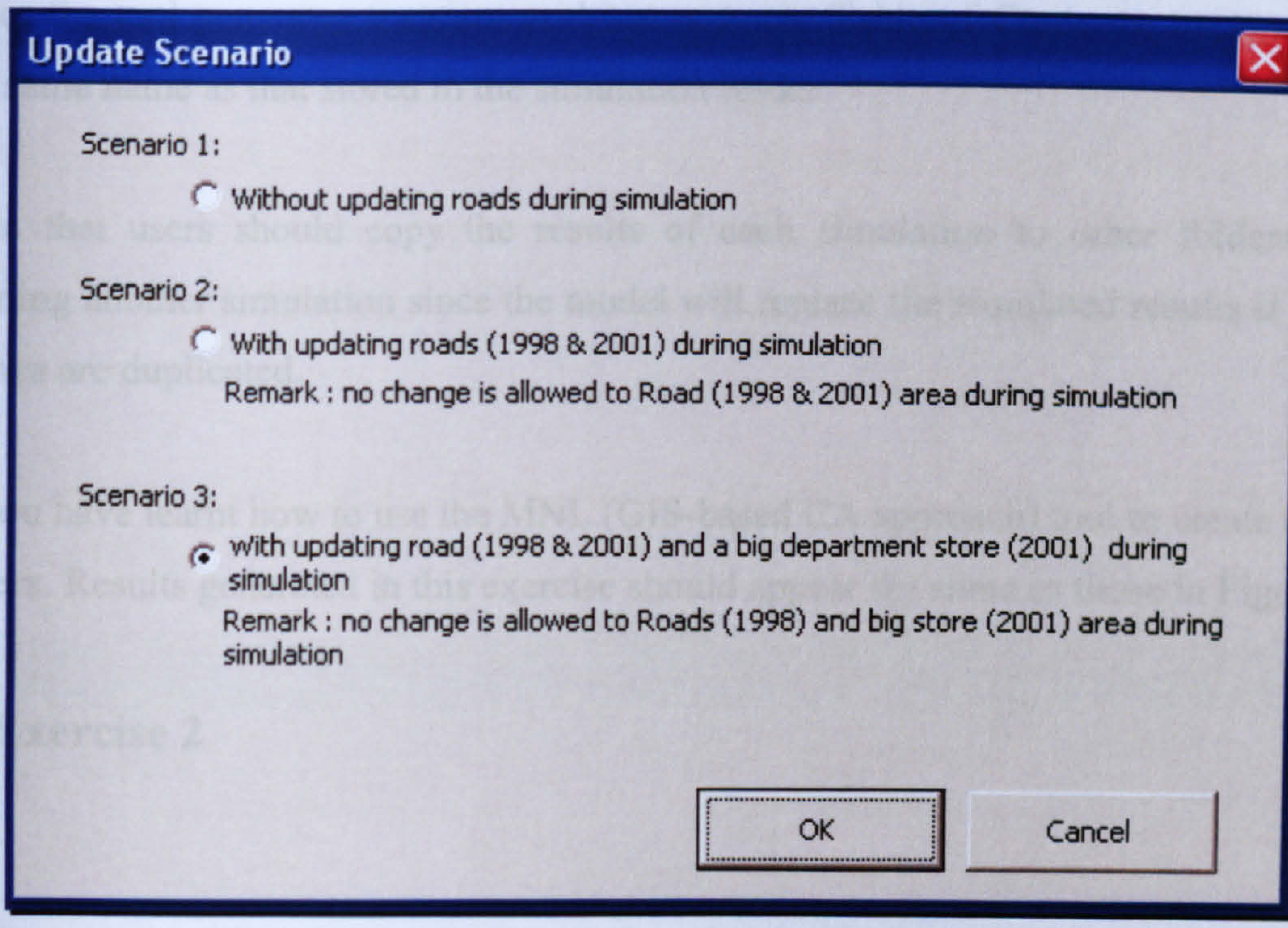
4. Type in the weight values. In this exercise, click **LOAD** to load all weight sets used for the analysis. This set of weights (see Table 6.2) is the same as that used for the simulation undertaken using the MNL method in Section 6.2.1. The **Reset** button allows the use to initialize all weights to 0 (zero) if desired.

Note that the user can also type in different sets of weights derived from the MNL method, given in Tables 6.4 – 6.7, in order to create different simulation outcomes as shown in Section 6.2.2. Otherwise, the user has to calculate a weights set by means of the MNL method described in Section 4.5.1.

5. Click **OK** or **Cancel** to return to the main menu with and without setting parameter values respectively.
6. From the main MNL menu, click **Environments…** to set neighbourhood size used in the analysis. In this exercise, however, you do not need to change these input values. If button **Environments…** is clicked, the Set environment pop-up should appear.



7. From the Set environment pop-up, click **OK** or **Cancel** to return to the main menu with and without setting environments respectively.
8. From the main MNL menu, click **Update Scenario** to choose the update alternative. By default, scenario 3 is chosen. More detail about update scenario section is presented in Section 5.3.2.1. In this exercise, however, you do not need to change scenario. If button **Update Scenario** is clicked, the Update Scenario pop-up should appear.



9. From the Update Scenario pop-up, Click **OK** or **Cancel** to return to the main menu with and without setting update scenarios respectively
10. Now you are prompted to run the model. From the main MNL menu, click **Compute** to run the tool.

After finishing, the message box should appear as "MNL Simulation is Done.. Bye Bye".

11. Click **Quit** to exit tool.

The results of this tool are a series of grid layers, stored in the folder "C:\ LUMICS-a3703781\simulation". They are in the format of ArcGIS grid layers. Since the model is set to run in two-year intervals, based on the initial year 1993 and the end year 2001, in this exercise, you should get four simulated outputs. They are named luy1995e2001, luy1997e2001, luy1999e2001, and luy2001e2001. The grid name, such as luy1995e2001, refers to the land use grid layer "lu" that is produced for the year "y" of "1995" and the end year "e" of "2001". You can use the layer files, named "lucode10.lyr" (ten land use classes) and "display4luwhite.lyr" (four land use classes) stored in the folder "C:\ LUMICS-a3703781\param" to assign colour to land use types.

Also, the tool generates transient grid layers in the Table of Contents, appearing with the same name as that stored in the simulation folder.

Note that users should copy the results of each simulation to other folders before running another simulation since the model will replace the simulated results if the grid names are duplicated.

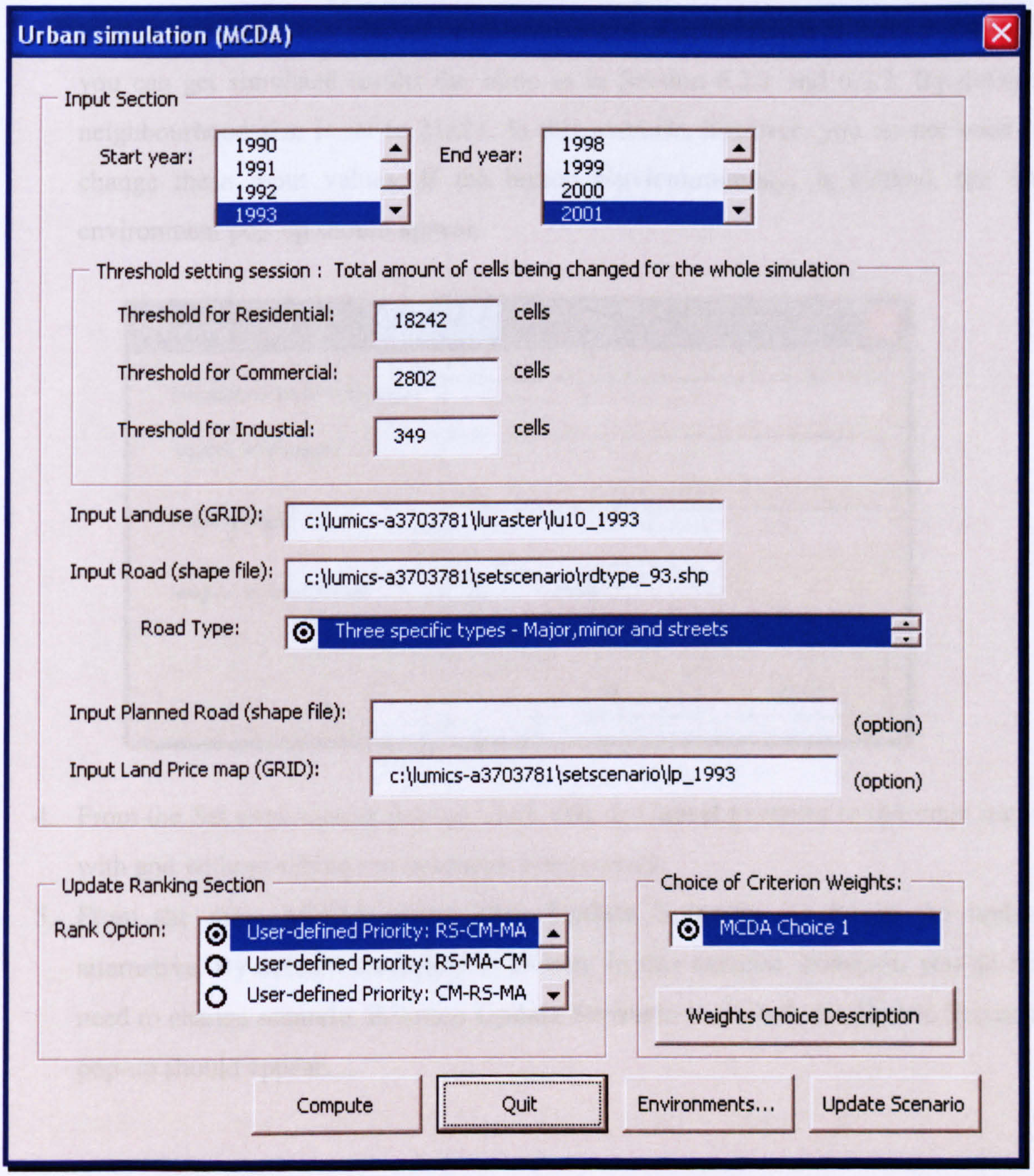
Now, you have learnt how to use the MNL (GIS-based CA approach) tool to create the land use layers. Results generated in this exercise should appear the same as those in Figure 6.2.

END Exercise 2

Exercise 3: MCDA (GIS-based CA approach) Tool

You will use this tool to do the simulation of land use change. The simulation performed is based on the CA approach, where the multi-criteria decision analysis (MCDA) has been integrated to identify the potential cells for development. More description of this tool is given in Section 5.3.3.1.

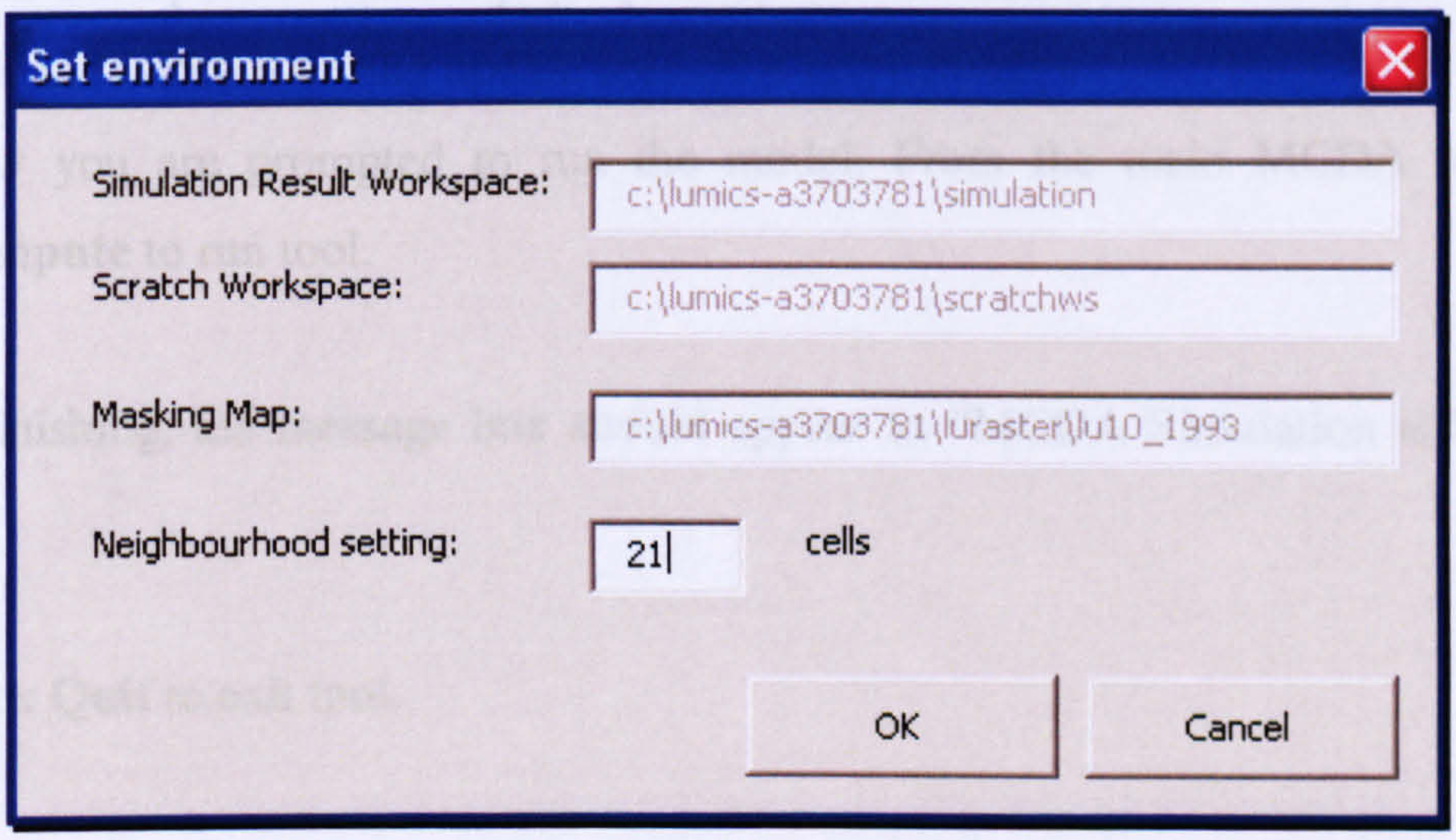
1. From LUMICS (Thesis version) menu, click **MCDA (GIS-based CA approach)**. The MCDA (GIS-based CA approach) dialog box should now pop-up.



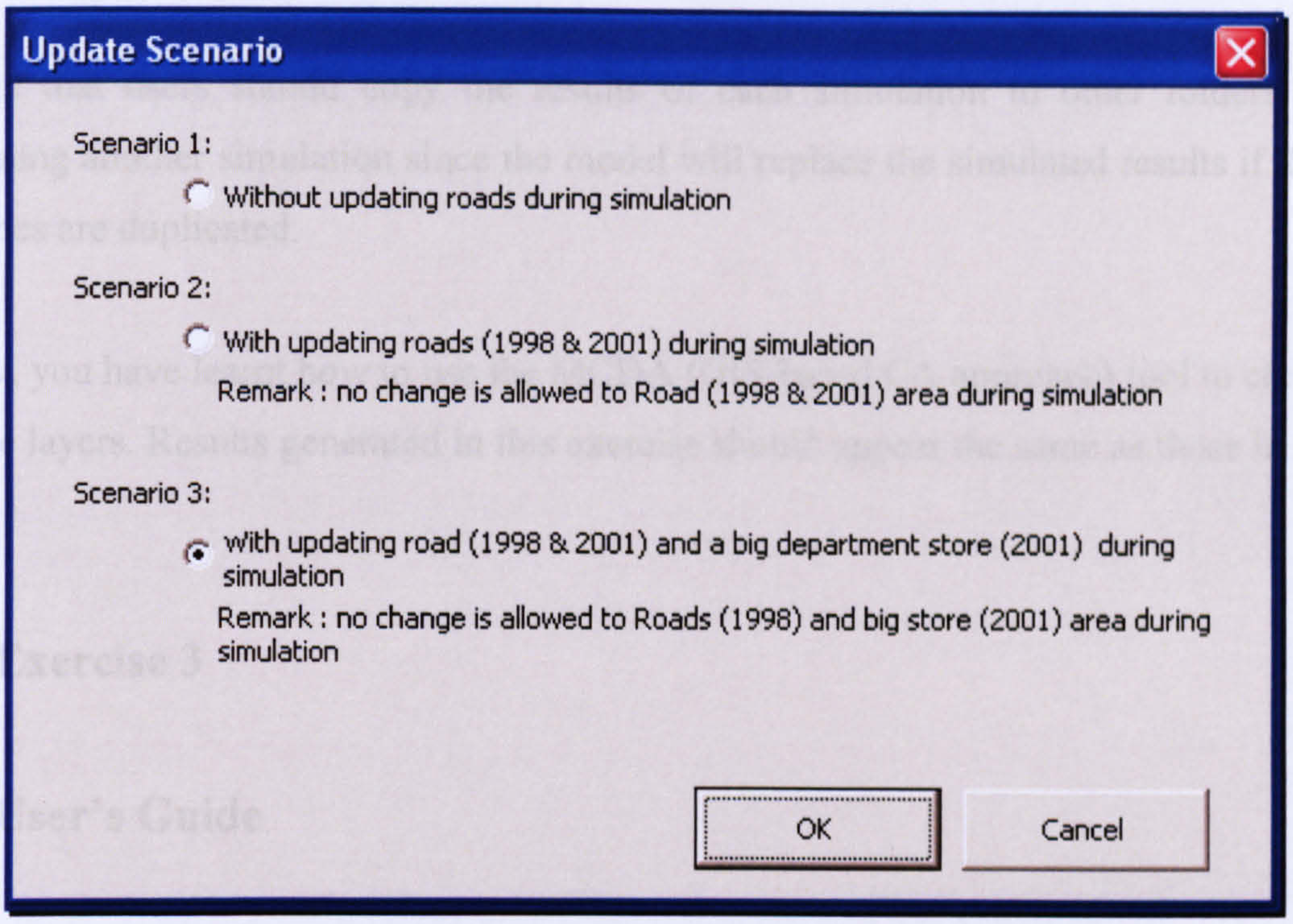
2. The tool provides input (e.g. simulation period, amount of cells set to change for the whole simulation, layer input) ready to be executed. You may change these inputs

by typing in a new value (e.g. threshold for residential) or choose the multiple choice provided (e.g. simulation period, update ranking option). More details about sections (e.g. update ranking section) in this page are given in Section 5.3.3.1. Only one set of criterion weights is available. This set of weights (see Table 6.19) and neighbourhood thresholds setting (see Table 6.20) are the same as that used for simulation based on the MCDA method undertaken in Section 6.3.1. You can click **Weight Choice Description** in Choice of criterion Weights section in order to see the criterion set and neighbourhood thresholds used for the simulation.

3. From the main MCDA menu, click **Environments...** to set the neighbourhood size used in the analysis. If you set neighbourhood size in the range of 3x3 to 101x101, you can get simulated results the same as in Section 6.3.1 and 6.3.2. By default, neighbourhood size is set to 21x21. In this exercise, however, you do not need to change these input values. If the button **Environments...** is clicked, the Set environment pop-up should appear.



4. From the Set environment pop-up, click **OK** or **Cancel** to return to the main menu with and without setting environments respectively.
5. From the main MCDA menu, click **Update Scenario** to choose the update alternative. By default, scenario 3 is chosen. In this exercise, however, you do not need to change scenario. If button **Update Scenario** is clicked, the Update Scenario pop-up should appear.



6. From the Update Scenario pop-up, Click **OK** or **Cancel** to return to the main menu with and without setting update scenarios respectively
7. Now you are prompted to run the model. From the main MCDA menu, click **Compute** to run tool.

After finishing, the message box should appear as "MCDA Simulation is Done.. Bye Bye".

8. Click **Quit** to exit tool.

The results of this tool are a series of grid layers, stored in the folder "C:\ LUMICS-a3703781\simulation". They are in the format of ArcGIS grid layers. Since the model is set to run in two-year intervals, based on the initial year 1993 and the end year 2001, in this exercise, you should get four simulated outputs. They are named luy1995e2001, luy1997e2001, luy1999e2001, and luy2001e2001. The grid name, such as luy1995e2001, refers to the land use grid layer "lu" that is produced for the year "y" of "1995" and the end year "e" of "2001". You can use the layer file, named "lucode10.lyr" stored in the folder "C:\ LUMICS-a3703781\param" to assign colour to land use types. Also, the tool generates transient grid layers in the Table of Contents, appearing with the same name as that stored in the simulation folder.

Note that users should copy the results of each simulation to other folders before running another simulation since the model will replace the simulated results if the grid names are duplicated.

By now, you have learnt how to use the MCDA (GIS-based CA approach) tool to create the land use layers. Results generated in this exercise should appear the same as those in Figure 6.11.

END Exercise 3

END User's Guide