11-24-2017 10:00 AM

Geosimulation and Multicriteria Modelling of Residential Land Development in the City of Tehran: A Comparative Analysis of Global and Local Models

Hossein Hosseini, *The University of Western Ontario*

Supervisor: Malczewski, Jacek, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Geography

Abstract

Conventional models for simulating land-use patterns are insufficient in addressing complex dynamics of urban systems. A new generation of urban models, inspired by research on cellular automata and multi-agent systems, has been proposed to address the drawbacks of conventional modelling. This new generation of urban models is called geosimulation. Geosimulation attempts to model macro-scale patterns using micro-scale urban entities such as vehicles, homeowners, and households. The urban entities are represented by agents in the geosimulation modelling. Each type of agents has different preferences and priorities and shows different behaviours. In the land-use modelling context, the behaviour of agents is their ability to evaluate the suitability of parcels of land using a number of factors (criteria and constraints), and choose the best land(s) for a specific purpose. Multicriteria analysis provides a set of methods and procedures that can be used in the geosimulation modelling to describe the behaviours of agents.

There are three main objectives of this research. First, a framework for integrating multicriteria models into geosimulation procedures is developed to simulate residential development in the City of Tehran. Specifically, the local form of multicriteria models is used as a method for modelling agents' behaviours. Second, the framework is tested in the context of residential land development in Tehran between 1996 and 2006. The empirical research is focused on identifying the spatial patterns of land suitability for residential development taking into account the preferences of three groups of actors (agents): households, developers, and local authorities. Third, a comparative analysis of the results of the geosimulation-multicriteria models is performed. A number of global and local geosimulation-multicriteria models (scenarios) of residential development in Tehran are defined and then the results obtained by the scenarios are evaluated and examined. The output of each geosimulation-multicriteria model is compared to the results of other models and to the actual pattern of land-use in Tehran. The analysis is focused on comparing the results of the local and global geosimulation-multicriteria models. Accuracy measures and spatial metrics are used in the comparative analysis. The results suggest that, in general, the local geosimulation-multicriteria models perform better than the global methods.

Keywords: geosimulation, local multicriteria analysis, residential land development, the City of Tehran.

Acknowledgments

I would first like to express my sincere gratitude to my supervisor Dr. Jacek Malczewski for advising this project. He always gave me useful advice and pointed me in the right direction.

I would like to thank the members of my advisory committee– Dr. Milford Green and Dr. Diana Mok– to provide valuable insights and suggestions. I also want to thank the members of my PhD defense committee– Dr. Isaac Luginaah, Dr. Jinfei Wang, Dr. Zack Taylor, Dr. Claus Rinner, and Dr. Jean-Louis Schaan– who agreed to evaluate the thesis in terms of academic merit and provided intellectual comments.

I am grateful to the second reader of my dissertation, Dr. Milford Green, who spent his time reading through the dissertation and provided constructive feedback.

I am grateful to all geography department staff, especially Lori Johnson, Joe Smrekar, and Karen VanKerkoerle. I would like to extend my gratitude to all colleagues in the program, especially Luis Silva who helped me in the process of making power point for the public presentation.

Sincere thanks go to "DigitalGlobe Foundation" for the provision of an imagery grant. The present research was not concluded without this imagery grant.

Last but not least, I want to express my deep sense of gratitude to my family for their patience and support. Especial thanks go to my Mom and Dad who always encouraged me to pursue the PhD program.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

1 General introduction

1.1 Introduction

A number of approaches for simulating the process of urban development have been proposed over the last thirty years or so (Wu, 2005). Most of the approaches are based on the complex system theory (Allen, 1997; Benenson, Aronovich, and Noam, 2005). A complex system is made up of many distinct and autonomous elements which are interdependent and interrelated (Wolfram, 1984; Benenson, Aronovich, and Noam, 2005). When these autonomous elements are connected, they can create complex phenomena, patterns, and behaviours (Benenson and Torrens, 2004b; Wolfram, 1984). A complex system is characterized by three attributes: heterogeneity, interdependencies, and nested hierarchies (Arthur, Durlauf, and Lane, 1997; Epstein, 1999; Kohler and Gummerman, 2000). Individuals' behaviours and the features of the landscape over which individuals interact generate complexities in an urban system (Parker et al., 2003). According to Parker et al. (2003), heterogeneity is embodied in both individuals (agents) and landscape. Agents may be classified into different groups based on their preferences, capabilities, knowledge, power, and so on. Furthermore, the physical landscape is heterogeneous in that there is an uneven distribution of various species over space; also, the land surface and landscape characteristics (e.g., temperature and precipitation) are different from one location to another. In addition, one can recognize interdependencies between individuals and between individuals and landscape (Parker et al., 2003; Benenson and Torrens, 2004b). An individual learns from his/her previous experiences and creates knowledge, then uses his/her and others' knowledge to improve his/her decision behaviours. Accordingly, there are interdependencies between individuals and these interdependencies also affect the landscape; for example, the landscape is subject to changes in land cover and land-use type due to individuals' actions. Also, one can identify physical and social systems with hierarchical and nested structures (Parker et al., 2003). For instance, individuals communicate to establish families which in turn interact with other families via some economic and political systems.

Geosimulation is a fast-growing area of research in Geographic Information Science (GISci) and complex system theory (Benenson and Torrens, 2004a). The main objective of geosimulation is to understand the dynamics of complex human-driven spatial systems based on computer simulations (Parker et al., 2003; Benenson and Torrens, 2004a). Computer-based technologies help researchers to simulate behaviours (e.g., actions and interactions) of individual entities in a complex system. Geosimulation aims to understand how these actions and interactions affect the underlying landscape. Multicriteria analysis can be integrated into geosimulation to provide a framework for simulating actions and interactions of individual entities (Malczewski and Rinner, 2015). Geosimulation and the role of multicriteria analysis and GIS in geosimulation models will be discussed in more detail in the next sections.

1.2 What is geosimulation?

Two general categories of models have been introduced to examine urban dynamics and spatial patterns: macro-scale and micro-scale models (Irwin, Jayaprakash, and Munroe, 2009). Macro-scale urban models consider urban dynamics as the result of exogenous factors, such as political, socio-economic, or biophysical driving forces, and simulate urban dynamics in aggregated spatial units, like zones or regions (Li, Wu, and Zang, 2014). Panel data analysis, econometric models, and systems dynamic models are all among macro-scale models that can be applied to understand urban dynamics. On the other hand, micro-scale models examine urban dynamics at the individual level (Li, Wu, and Zang, 2014). These types of models simulate the behaviours of individuals and scale up these behaviours to explain urban dynamics and spatial patterns (Berger, 2001; Parker et al., 2003).

According to Benenson and Torrens (2004a) the conventional models for simulating urban patterns were insufficient in addressing complex dynamics inherent in urban systems. Urban models faced severe criticism in the 1970s. Lee (1973) and Sayer (1979) questioned the efficiency of macro-scale urban models, such as the Lowry model (Lowry, 1964), as a supporting tool for land-use planning. Accordingly, a new wave of urban models, inspired by research on cellular automata (CA) and multi-agent systems (MAS), superseded the conventional models (Batty, Couclelis and Eichen, 1997; Benenson and Torrens, 2004b; O’Sullivan and Torrens, 2001; Torrens, 2003). This new generation of urban models is based on micro-scale urban entities such as pedestrians, residents, and homeowners (Benenson, 1998, 1999; Benenson, Omer, and Hatna, 2002). Indeed, these models aim at simulating macro-level systems at micro-scale and entity-level units (Moulin et al., 2003). With the advent of new computer-based technologies, it was possible to model the urban dynamics by taking into account important details that could not be considered earlier due to the capacities of previous tools. The new class of urban models was coined "geosimulation" by Benenson and Torrens (2004a). The focus of geosimulation models is on the spatial and geographical nature of urban systems. Geosimulation deals with human individuals and any spatial entities and the interactions that exist between them. Torrens (2006) defined geosimulation models by referring to these three characteristics. First, while conventional models for simulating urban dynamics are based on aggregated geographical units which are spatially modifiable, non-modifiable spatial entities form the simulation space in geosimulation modelling procedures. Second, contrary to conventional models, geosimulation tries to understand spatial patterns and dynamics by simulating the behaviours at the entity-level. Geographical entities exhibit autonomous and independent behaviours in geosimulation and their behaviours are heterogeneous across the space. Geographical entities can be humans, vehicles, and other moving objects as well as non-moving objects like parcels of land. In the land-use/cover change context, individuals, interest groups, and/or parcels of land can be regarded as influential geographical entities. These entities are usually referred to as agents in geosimulation models. Third, unlike conventional urban models, geosimulation methods are more event-driven instead of time-driven.

1.3 Geosimulation and multicriteria analysis

As mentioned in the previous section, the simulation of behaviours of geographical entities is at the core of geosimulation modelling. Irrespective of the type of geographical entities, there is a need for methodology to simulate the behaviours of these entities in the urban system. Multicriteria analysis provides a set of methods and procedures that can be used in geosimulation modelling to describe the behaviours of individual entities. One can identify two general classes of multicriteria analysis: global and local (Malczewski and Rinner, 2015). Global multicriteria analysis deems that the behaviours of geographical entities are the same across the study area. The premise behind the local multicriteria modelling is that the behaviours of the geographical entities may change based on landscape characteristics. The behaviours of spatial entities in a situation depend on three elements: (i) evaluating the situation based on a set of attributes (or evaluation criteria) that is assumed to be important in the process of entities' decision making, (ii) a set of constraints that limit their behaviours, and (iii) entities' preferences with respect to the contributing attributes. The behaviours of entities in different situations are the result of the combination of these three elements. The combination procedure is operationalized through decision rules. Defining the decision rule is a key part of any multicriteria analysis (Malczewski, 1999). Indeed, actions of an entity may vary in a situation by applying different decision rules (or multicriteria methods). A wide range of decision rules exists in multicriteria analysis, such as analytic hierarchy process (Saaty, 1980), ideal point method (Lotfi, Stewart, and Zionts, 1992), weighted linear combination (Einhorn and Hogarth, 1975), and ordered weighted averaging (Yager, 1988).

1.4 Role of GIS in geosimulation

The modelling of an urban system depends on the integration of two technologies: computer-based simulation procedures and Geographic Information Systems (GIS). GIS has contributed substantially to apply geosimulation models for analysing real-world phenomena and urban processes. Simulating urban dynamics by micro-scale models requires spatial micro-scale data (Irwin, Jayaprakash, and Munroe, 2009). The advent of GIS and remote sensing has enabled researchers to simulate real-world urban processes. However, it was not until the late 1980s that GIS tools and capabilities have been employed in urban studies (Anselin and Getis, 2010; Fotheringham and Rogerson, 1994). GIS serve as platforms for implementing geosimulation and also as a spatial database to store, retrieve, manipulate, analyse, and display data. The database and geosimulation procedures can be connected to other data sources, such as remote sensing software or online maps, to get required data. A wide range of GIS tools can be utilized in geosimulation to manipulate the relevant dataset (Benenson and Torrens, 2004a). Furthermore, the outcomes of the modelling can be presented in a GIS environment and in different scales with the help of visualization capabilities of GIS. Output maps can be stored in the GIS environment and restored later as the initial state in multi-stage simulations.

1.5 Importance of the research

1.5.1 Limitation of earlier models

A wide range of models have been developed to simulate land-use/cover change by integrating geosimulation and multicriteria analysis (e.g., Wu, 1998; Loibl and Toetzer, 2003; Ligtenberg et al., 2004; Manson, 2005; Ligmann-Zielinska and Jankowski, 2010; Sabri, Ludin, and Ho, 2012; Nourqolipour et al., 2015; Singh et al., 2015). One can identify two major limitations of the previous geosimulation-multicriteria modelling approaches (see Chapter 2). First, a large number of these studies employed weighted linear combination (WLC) as the decision rule to model the urban dynamics. A very few geosimulation studies used other multicriteria methods such as the ordered weighted averaging (OWA), which is a generalization of the most often used GIS-based multicriteria procedures including WLC and Boolean operations (Jiang and Eastman, 2000). Specifically, OWA is a class of multicriteria operators that involves two sets of weights: criterion weights and order weights (Yager, 1988). A family of OWA operators can be obtained by changing the order weights; that is, OWA can be used to change the type of combination of attribute (criterion) maps from the logical AND combination through all intermediate types (including conventional WLC) to the logical OR combination (Yager, 1988; Jiang and Eastman, 2000). A central element of the OWA procedure is the process of assigning criterion weights and selecting a set of order weights. The criterion weights can be determined by using the ranking/rating methods; and subsequently, the order weights can be inferred from the criterion weights using the concept of linguistic quantifiers (Yager, 1996; Malczewski, 2006b). This approach is referred to as the linguistic quantifiers-based OWA or linguistic quantifiers-OWA model.

Second, all previous geosimulation-multicriteria studies use multicriteria models at a 'global' level, meaning that one set of results is generated from the analysis and these results are assumed to apply equally across the study area (Malczewski, 2011). In practice, however, it is often unreasonable to make an assumption about homogeneous behaviour of geographical entities; that is, the behaviour is assumed to be the same irrespective of the entities' locations and conditions of their neighbourhood. This limitation of the global geosimulation-multicriteria methods can be overcome by local multicriteria models (Malczewski, 2011; Carter and Rinner, 2014; Malczewski and Liu, 2014; Cabrera-Barona et al., 2015). Specifically, this research aims at integrating the local form of linguistic quantifiers-OWA into the GIS-based geosimulation procedure.

1.5.2 Rapid urban development

As a developing country, Iran is struggling with continual large-scale residential developments (Rafiee et al., 2009). Urban growth and land-use/cover changes continue to occur in large urban areas by sacrificing a great amount of farmlands and losing open lands (Hosseinali, Alesheikh, and Nourian, 2013). Residential development in Iran is mostly taking place in a few big cities, such as Tehran, Esfahan, Tabriz, and Mashad (Rafiee et al., 2009). With a population in excess of thirteen million, Tehran metropolitan region is one of the fastest-growing urban areas in the world. Rapid urban growth brings many environmental and social challenges. For example, as the built-up areas continue to expand, local authorities need to provide more residents with services like clean water and electricity, which consumes more resources. Land-use/cover change models can be employed to simulate, understand, and project urban dynamics and their consequences (Rafiee et al., 2009). In fact, developing a robust model aids urban planners and policy makers to examine the pattern of land-use/cover change and urban development.

1.6 Research objectives and questions

There are three main objectives of this research: (i) *developing a framework for integrating local multicriteria models into geosimulation procedures to simulate residential land development in the City of Tehran*, (ii) *testing the framework in the context of residential land development in Tehran between 1996 and 2006*, and (iii) *analysing the results of the geosimulation-multicriteria procedure.* Forty-two scenarios (global and local geosimulation-multicriteria models) of residential development in Tehran are defined and then the results obtained by the scenarios are evaluated and examined. The output of each scenario is compared to the results of other scenarios and to the actual pattern of land-use in Tehran. The analysis focuses on comparing the results generated by the different scenarios in terms of the two components of the geosimulation-multicriteria models: the linguistic quantifiers (or associated order weights) and the size of the neighbourhood (or the order of contiguity) used for the local multicriteria modelling. A series of hypotheses is tested to address the following research questions: (i) *are there significant differences between the results of local and global geosimulation-multicriteria models for different linguistic quantifiers*? and (ii) *are there significant differences between the results of local and global geosimulation-multicriteria models for different neighbourhood sizes*? There is some evidence to show that the results generated by the OWA models depend on linguistic quantifiers (e.g., Rinner and Malczewski, 2002; Eldrandaly, 2013; Malczewski and Liu, 2014; Cabrera-Barona et al., 2015). One can hypothesize that for a given linguistic quantifier the results of local geosimulation-multicriteria models provide us with a better (more accurate) description of the spatial pattern of residential land development than global models. Also, the previous studies suggest that the results of multicriteria analysis depend on the spatial scale at which the analysis is performed (e.g., Can, 1992; Lopez Ridaura et al., 2005). Given a study area, any change in the neighbourhood size affects the results generated by the local models (Malczewski and Rinner, 2015). One can hypothesize that more localized geosimulation-multicriteria models provide us with a more accurate description of the spatial pattern of land-use; that is, one would expect that the smaller the neighbourhood's size, the greater differences between the results of local and global geosimulation-multicriteria models.

1.7 Thesis organization

The thesis is organized in seven chapters as follows:

- Chapter 1 General introduction. This chapter provides an introduction to the geosimulation-multicriteria modelling of spatial pattern of residential land development in Tehran.
- Chapter 2 Literature review. This chapter presents a review of previous studies on the integration of multicriteria analysis and geosimulation.
- Chapter 3 Study area. This chapter describes different aspects of the study area, such as geographical and social characteristics.
- Chapter 4 Methods: Geosimulation-multicriteria models. This chapter discusses the theoretical foundation of geosimulation and multicriteria analysis. It also provides more details as to what are the actors in the process of residential land development in the study area and how the decision behaviours of these actors are simulated. This chapter also explains how the suggested framework is developed and how different components of the framework work.
- Chapter 5 Input data. This chapter describes all spatial and non-spatial input data required for executing the geosimulation-multicriteria procedure for the study area. It also explains how non-spatial data were collected for the model.
- Chapter 6 Results and discussion. In this chapter, 42 scenarios (models) are defined to simulate the residential land development in Tehran. Then, the geosimulation-multicriteria procedure is executed for each scenario and the results are summarized in both tabular and cartographic formats. The main focus of this Chapter is on a comparative analysis of the

results of global and local geosimulation-multicriteria models. The results of the 42 scenarios are evaluated and compared using different accuracy measures and spatial metrics.

- Chapter 7 Conclusion. This chapter gives a summary of the geosimulation-multicriteria procedure and research findings. Also, the limitations of the current study are discussed and some aspects for future research are suggested.

Chapter 2

2 Literature review

2.1 Background

Human activities are responsible for most dynamics taking place in urban ecosystems, such as land-use/cover change, land-use intensification, and land degradation (Lambin, 1997). Land-use/cover changes have been accelerating in recent years due to socio-economic and biophysical factors (Lambin et al., 2001). In the wake of growing awareness among urban researchers about the need for developing new models for simulating urban dynamics, a large number of land-use/cover change models have been introduced (Verburg et al., 2002).

Earlier approaches to modelling urban dynamics tried to simulate the urban system by concentrating on coarse urban structures and considering urban areas as homogenous geographical entities (Loibl and Toetzer, 2003). The models introduced in the 1960s, 1970s and 1980s to simulate urban dynamics were mostly grounded on the theories developed by Forrester (1969) and Lowry (1964) (Loibl and Toetzer, 2003). Bid rent theory and equilibrium notion were at the core of those models. According to Alonso (1964), an urban system is in economic equilibrium if it meets the following four conditions at the same time: (i) location equilibrium for residents, (ii) location equilibrium for businesses, (iii) equilibrium of labor market, and (iv) competition in land market. However, these models ignore the micro-scale dynamics behind changes in urban systems. Since the mid-1990s, there has been a significant improvement in urban models inspired by cellular automata (Torrens and O'Sullivan, 2001) and multi-agent system models (Clifford, 2008; Parker et al., 2003).

2.2 Literature survey procedure

To gain a better insight into the research about integrating geosimulation and multicriteria analysis and understand recent progresses and main challenges, a research review of relevant literature was carried out. The review involved a systematic search for publications about integrating geosimulation and multicriteria modelling approaches. The following web-based databases and electronic libraries were used to search for relevant papers published so far: IEEE Xplore®, Pion Publications Ltd., Project MUSE®, ProQuest®, ScienceDirect, Scopus, and SpringerLink.

The objective of searching for relevant publications was to find any papers on integrating geosimulation modelling with multicriteria analysis. The search was done using Boolean operators and research keywords. Six key terms were selected, three associated with geosimulation fields of research and three other terms to cover multicriteria studies. Geosimulation techniques include cellular automata and multi-agent systems. Here, the term "agent" was used to cover studies related to both agent-based models and multi-agent systems. "multicriteria" covers all related terms, such as multicriteria analysis, multicriteria evaluation, multicriteria decision analysis, and multicriteria decision making. Moreover, all papers within multiattribute and multiobjective fields of research are related to the field of multicriteria analysis. Table 2.1 shows all keywords used to generate the primary results. A five-step search was performed to find all relevant articles from the selected databases. Table 2.2 contains a summary of each step. The primary search included any combination of one or more key terms in the first column with one or more key terms in the second column. For example, the following query: [(agent OR "cellular automata" OR geosimulation) AND (multicriteria OR multiobjective OR multiattribute)] was used in step 1 to find the relevant articles using web-based databases (Table 2.2).

Table 2.1: Keywords used to generate the initial results for relevant literature

Key terms used for geosimulation	Key terms used for multicriteria analysis
geosimulation	multicriteria
cellular automata	multiattribute
agent	multiobjective

Table 2.2: Searching procedure: Boolean operators and keywords

Searching procedure	Boolean operators and keywords (if applicable)
Step 1: basic keywords	TITLE-ABS-KEY (*agent* OR *"cellular automata"* OR *geosimulation*) AND TITLE-ABS-KEY (*multicriteria* OR *multiattribute* OR *multiobjective*)
Step 2: basic keywords and 'land'	(TITLE-ABS-KEY (*agent* OR *"cellular automata"* OR *geosimulation*) AND TITLE-ABS-KEY (*multicriteria* OR *multiattribute* OR *multiobjective*) AND TITLE-ABS-KEY (*land*))
Step 3: basic keywords and 'land' and 'urban'	(TITLE-ABS-KEY (*agent* OR *"cellular automata"* OR *geosimulation*) AND TITLE-ABS-KEY (*multicriteria* OR *multiattribute* OR *multiobjective*) AND TITLE-ABS-KEY (*land*) AND TITLE-ABS-KEY (*urban*)
Step 4:	Screening Articles
Step 5:	Removing duplicates

The search was limited to title, abstract, and keywords or just to title and abstract based on the database engine capabilities. The results also were filtered based on the document type and language. In this study, only documents published in English and refereed journals were considered. Figure 2.1 shows the details of each of the five steps for each database. For instance, according to Figure 2.1, the total of 5074 articles was found in step 1. In the second step of the systematic searching procedure, the results of the first search were refined by including the term "land" to exclude articles not related to land-use/cover context (Table 2.2). The number of articles at the end of step 2 was 648 (Figure 2.1). In step 3, the term "urban" was used to exclude those articles carried out in rural or other non-urban areas (Table 2.2). The total of 359 articles was found to be relevant at the end of step 3 (Figure 2.1). In step 4, the final articles were reviewed to exclude irrelevant ones. In the final step, the results of step 4 for all databases were merged and duplicates were removed. By the end of the searching process, 53 articles were found to be in alignment with the nature of the current research. The details of the relevant literature are summarized in Table A1 (see Appendix A).

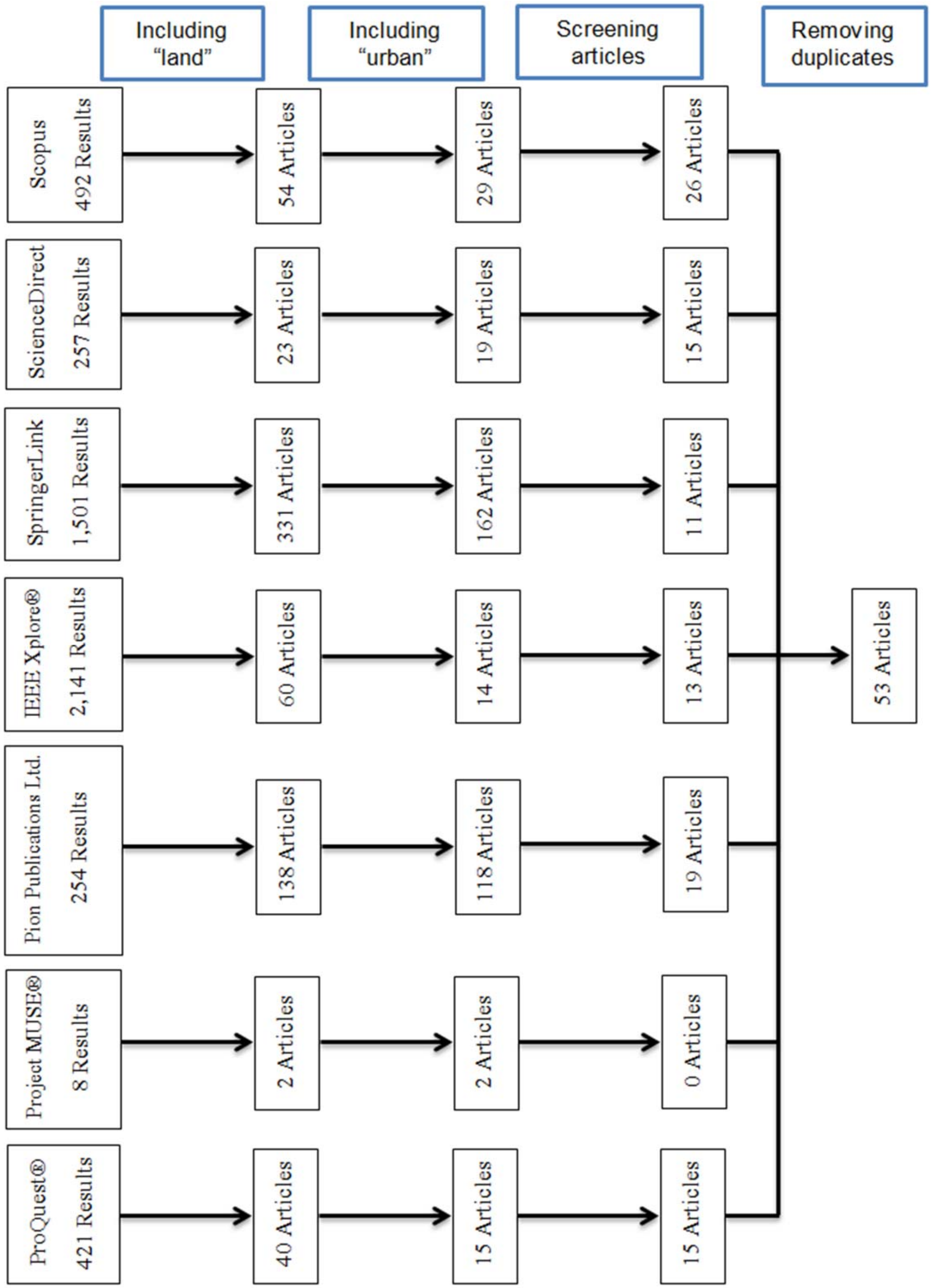


Figure 2.1: The results of searching online databases

2.3 Relevant literature

There is a considerable volume of literature on applying multicriteria analysis methods in geosimulation modelling. Figure 2.2 shows the number of articles about the integration of geosimulation and multicriteria analysis published in refereed journals by April 30, 2017. The first relevant article was published in 1998. In the last five years or so, there has been a significant increase in the number of geosimulation-multicriteria analysis articles published in refereed journals. Of 53 papers, about 70% have been published since 2011. There are only seventeen relevant articles published between 1998 and 2010. This rapid growth of research on integrating geosimulation and multicriteria analysis can be attributed to three main factors. First, a number of open-source or low-cost frameworks and software have been developed enabling academics and practitioners to integrate multicriteria analysis and geosimulation. For example, SLEUTH (Clarke, Hoppen, and Gaydos, 1997), is a tightly coupled package developed in Clark lab (Worcester, MA) as a result of the studies on simulating the spread of wildfire (Clarke, Olsen, and Brass, 1993; Clarke, Riggan, and Brass, 1995). SLEUTH provides researchers with an open-source framework to simulate land-use/cover changes and predict future urban growth. The advantage of an open-source framework is that any researcher can modify the framework based on the requirement of the study or even improve the framework for future uses. Although some of these software packages were developed prior to 1998, they started to receive more recognition in geosimulation-multicriteria analysis modelling in recent years. I will shed light on the software and frameworks applied in the selected studies later in the chapter. Second, there has been an exponential growth of studies (and publications) on GIS-based multicriteria analysis over the last two decades (see Malczewski, 2006a; Malczewski and Rinner, 2015). Third, there is a growing interest among academics and practitioners to use geosimulation methods for modelling spatial processes (Verburg et al., 2002).

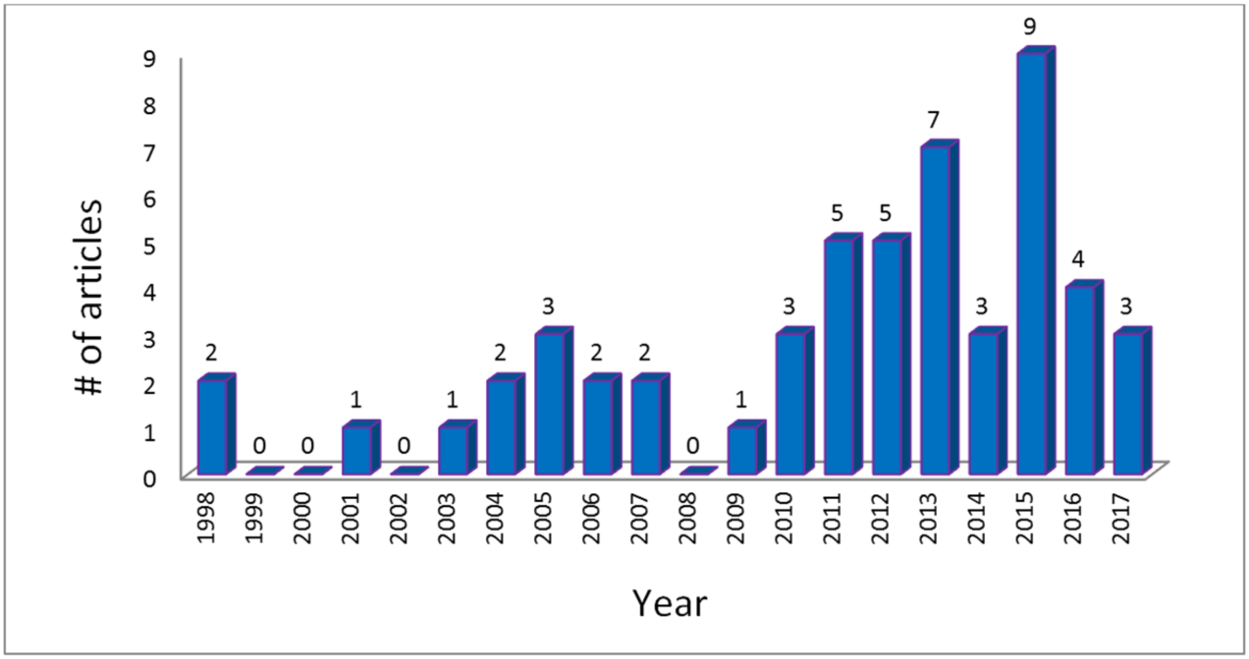


Figure 2.2: Number of geosimulation-multicriteria articles published in 1998-2017

Another interesting point is the location of the study areas used by the authors to test or implement their models (see Figure 2.3). About one third (32%) of the geosimulation-multicriteria studies selected a study area in China or Iran. According to the authors of these articles, most big cities in China are experiencing rapid urban development due to population and economic growth (Cheng and Masser, 2004; Li and Liu, 2007; Liu et al., 2007; Liu et al., 2014). In the case of Iran, in addition to fast urban growth, most major cities are suffering from uncontrolled development and urban sprawl (Hosseinali, Alesheikh, and Nourian, 2013; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013). The case study was not applicable in the research performed by Sabri, Ludin, and Ho (2012), and two studies used hypothetical landscapes to implement their model (see Ligmann-Zielinska, 2009; Ligmann-Zielinska and Sun, 2010).

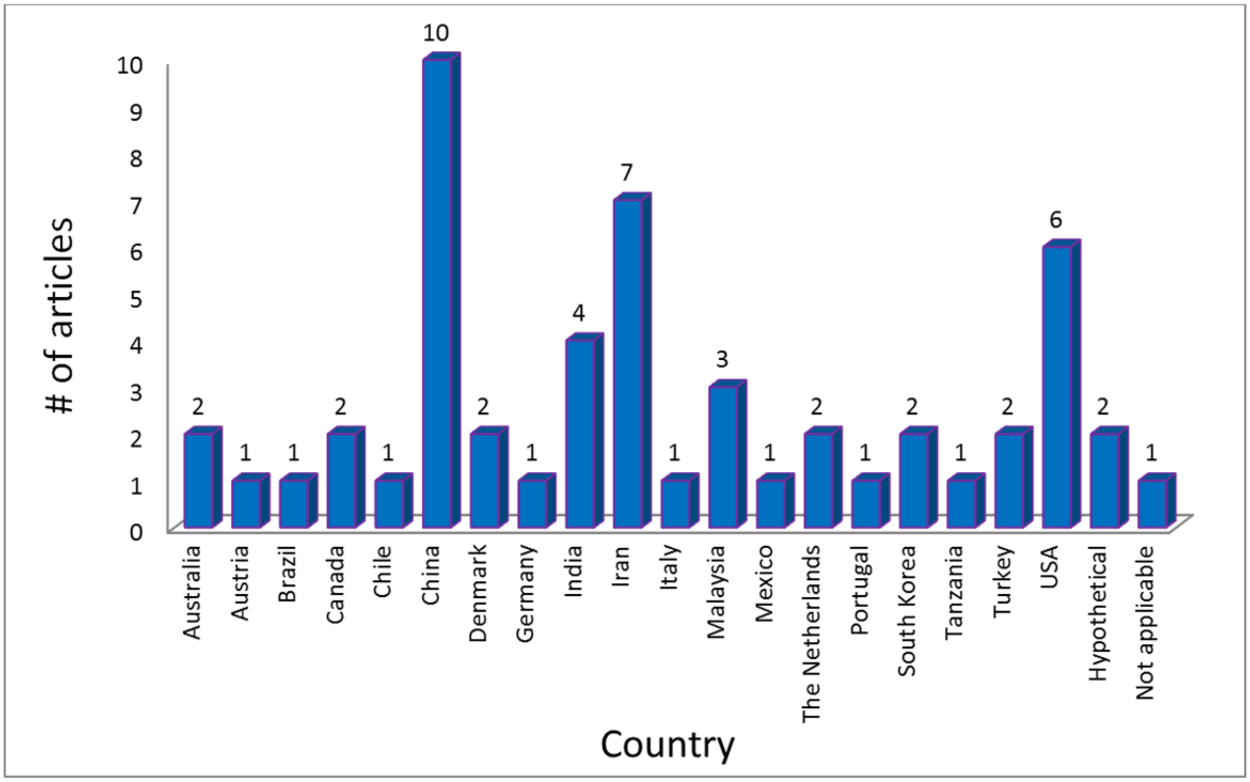


Figure 2.3: Geosimulation-multicriteria articles by country

2.4 Classification of articles

Four criteria were considered for classifying papers about integrating geosimulation and multicriteria analysis: (i) the type of geosimulation approach applied for modelling urban dynamics, (ii) the type of multicriteria analysis method used as a decision rule, (iii) the characteristics of data inputs for geosimulation modelling, and (iv) the software packages used for geosimulation-multicriteria analysis.

2.4.1 Geosimulation model

From the geosimulation perspective, any spatial dynamics are the result of micro-level spatial processes (Benenson and Torrens, 2004a). Cellular automata (CA) and multi-agent systems (MAS) are among the geosimulation models that attempt to understand spatial patterns by modelling individual processes and their interactions. CA considers the urban landscape as a grid of cells with a specific shape. Each cell holds a value that shows the state of the corresponding land parcel. The set of the feasible values depends on the number of land-use types required to be considered in the modelling process. For example, if there are just two land-use classes, built-up and non-built-up, then binary cells would be sufficient to represent the urban landscape. The state of each cell can change over time based on some transition rules. These rules define a new state for each cell based on the state of its neighbourhood cells. CA ability to incorporate the spatiotemporal nature of a process into the model is one of the reasons behind its popularity in simulation of urban phenomena. CA-based models are flexible and can be coupled with other models and tools to simulate urban development (Clarke et al., 1997). Specifically, cellular automata can be integrated into GIS to understand how a spatial process develops over time.

Multi-agent systems offer another approach to simulate spatial patterns by considering the actions and interactions of micro-level dynamics (Ligtenberg, Bregt, and Van Lammeren, 2001). In this approach, the first step is to understand how actors (agents) are engaged in a specific process. By simulating the actions and behaviours of micro-scale entities, a macro-scale pattern can be modeled. CA can be assumed as a special case of MAS in which agents are the cells that cannot move. MAS can be integrated into GIS to simulate spatial processes at different scales (Ligtenberg et al., 2004). MAS/GIS methodology has been used extensively to address many spatial problems ranging from developing a supporting tool for park management (Itami and Gimblett, 2001) to simulating people-environment interactions (Deadman and Gimblett, 1994).

If the relevant literature are categorized based on the methodology they applied, a majority of them used either CA or an integration of CA and MAS or agent-based models (ABM). In 43 out of 53 articles the authors explicitly talked about cellular concepts in the modelling process which account for almost 81% of articles (e.g., Akın, Sunar, and Berberoğlu, 2015; Hansen, 2012; Loibl and Toetzer, 2003; and Manson 2005). The reason behind this is related to the ability of cellular-based models to represent the landscape in a simple way and also the large number of software packages that support raster based inputs. In some MAS studies, cellular automata was not explicitly mentioned (e.g., Ligtenberg et al., 2004; Li and Liu, 2007; Hosseinali et al., 2013; Jokar Arsanjani et al., 2013; Ghavami and Taleai, 2016); nevertheless, the landscape was presented using a grid of cells and the state of each cell changes as an outcome of agents' actions and interactions. In all CA or MAS/CA studies the state of each cell indicates the land-use type. The type of land-use will be discussed in more details later in the chapter.

2.4.2 Multicriteria analysis

Multicriteria analysis (MCA) is a collection of methods for comparing different decision alternatives or evaluating scenarios using several evaluation criteria to help the decision-maker(s) in the process of assessing and choosing the best choice(s) (Roy, 1996). MCA is also referred to as multicriteria decision making (MCDM), multicriteria decision analysis (MCDA), or multicriteria evaluation (MCE). Malczewski (1999) suggested that a multicriteria decision (or evaluation) problem consists of six elements: (i) a decision goal or a number of decision goals that aims at determining the desirable state, (ii) a set of evaluation criteria among which different decision choices are assessed, (iii) a decision-maker or group of stakeholders who assess the decision choices, (iv) a set of decision alternatives that along with predefined evaluation criteria form a decision matrix, (v) factors over which the decision-makers have no control, and (vi) a set of evaluation outcomes for each element of the decision matrix.

Based on the six elements of MCA, one can recognize different classes of multicriteria decision problems and multicriteria methods including: multiobjective decision making (MODM) and multiattribute decision making (MADM), individual and group decision analysis, and decision problems under certainty and uncertainty (crisp and fuzzy decision making) (Malczewski, 1999; Chen, 2005). MADM is applied to select the best decision alternative among a finite number of alternatives based on decision-makers' priorities. In MADM, alternatives are defined explicitly by their attributes. In MODM, alternatives are specified implicitly by a multiobjective optimization mathematical model. Both MADM and MODM methods can be used to tackle multicriteria decision problems under the conditions of certainty or uncertainty as well as in the decision situation involving a single decision-maker or a group of decision-makers. MADM and MODM are usually referred to as multicriteria evaluation and multiobjective optimization in the literature.

The models presented in studies about urban dynamics and planning can be classified based on the type of multicriteria analysis integrated into geosimulation procedures. Multicriteria analysis can be used to generate suitability maps to elicit transition potential from one cell state to another (e.g., Bozkaya et al., 2015; Gong et al., 2015; Henríquez, Azócar, and Romero, 2006; Sakieh et al., 2015). From the cellular automata modelling point of view, land-use cells are the agents and the state of each cell can change over time based on some physical and socio-economic conditions. Therefore, a suitability surface in a given time is generated to assess the probability of transition from one state to another for each cell. Since in CA modelling agents cannot change their positions, it is not possible to model behaviours like way-finding or commuting using CA (Benenson and Torrens, 2004a). Moreover, unlike MAS, agents cannot be recognized with autonomous characteristics and behaviours in CA-based models. In MAS/ABM, agents are usually interest groups or individuals who can make decisions and interact with other individuals (e.g., Loibl and Toetzer, 2003; Hosseinali et al., 2013). For example, Li and Liu (2007) used a grid of cells to represent the landscape and agent-based models to define six autonomous groups of individuals based on family structure and income level. From the perspective of MAS/ABM, change occurs on the landscape as a result of the actions of individuals or interest groups. Individuals change the landscape according to some objectives. Individuals calculate the suitability of a cell for a specific use, such as residential or commercial areas. Whether it is CA or MAS/ABM, a method needs to be used to combine all important factors into a single number for each location, which shows the overall suitability of the location to be converted to another land-use type.

The first attempt to integrate MCA into cellular automata to generate the suitability maps for each land-use type was made by Wu and Webster (1998). They used multicriteria methods to elicit some rules that indicate the probability of transition from one land-use type to another. In the agent-based context, Ligtenberg, Bregt, and Van Lammeren (2001) applied weighted summation in their suggested MAS model for spatial planning. Because of the simplicity, multicriteria evaluation methods have more often been integrated to MAS or CA compared to multiobjective optimization methods. The multiobjective optimization approach was used in studies by Bone, Dragicevic, and White (2011). Zhang et al. (2011), Surabuddin Mondal et al. (2013), and Nourqolipour et al. (2015) employed a combination of multicriteria evaluation and multiobjective optimization methods in their modelling procedures. Other studies used multicriteria evaluation methods to create the suitability map (e.g., Hyandye and Martz, 2017; Keshtkar and Voigt, 2016; Lau and Kam, 2005; Wu, 1998). One of the interesting findings of this survey is the type of decision rule applied to evaluate the suitability/utility of a location for a specific purpose. Almost two-thirds of studies used weighted linear combination (WLC) or other types of weighted summation to combine the decision criteria (e.g., Mokadi, Mitsova and Wang, 2013; Singh et al., 2015; and Terra, dos Santos, and Costa, 2014). This can be attributed to the simplicity of WLC model. Six studies used analytic hierarchy process (AHP) (Ghavami and Taleai, 2016; Liu et al., 2007; Manganelli et al., 2016; Park, Jeon, and Choi, 2012; Park et al., 2011; Wu and Webster, 1998); there was a single paper presenting an application of fuzzy-AHP (Pooyandeh and Marceau, 2013), and one paper presented a study involving analytic network process (ANP) (Sabri, Ludin, and Ho, 2012). Ideal point method was applied in four studies (Ligmann-Zielinska, 2009; Ligmann-Zielinska and Sun, 2010; Ligmann-Zielinska and Jankowski, 2010; Liu et al., 2014), and ordered weighted averaging operator was used in a procedure proposed by Liu et al. (2014).

The suitability of a cell or parcel of land is assessed on the basis of several factors, which can be operationalized using the concept of criterion and constraint. Constraints define a set of rules or conditions based on which the transition from one state to another state in some part of the landscape (study area) is not feasible either because of the government regulations or physical conditions. Moreover, criteria represent a set of factors that are assumed to have impact on the transition probability of a cell. There is no consensus over the number of criteria that needs to be considered. For example, Hosseinali et al. (2013) suggested that a large number of criteria increases the interdependencies among criteria and decreases the accuracy of the model, and they considered as few as three criterion maps. On the other hand, some researchers considered any factors that they assumed are important; Jokar Arsanjani, Helbich, and de Noronha Vaz (2013), and Mahiny and Clarke (2012) provide an example since they used as many as 17 and 15 criteria, respectively, to create suitability maps.

One of the most important elements of multicriteria analysis is the procedure of assigning weights to different criteria based on their degree of importance. In the relevant literature there are two approaches for weighting: some authors used experts' knowledge and some authors used data to approximate the criterion weight. In the data-driven approach, the criterion weights are calculated based on the previous data – usually for several different times. Statistical methods are the most often used data-driven methods. In statistical methods, the regression analysis is performed on data to find if a criterion is important in an existing pattern and at what level. In knowledge-based approaches, the importance of a criterion is assessed more qualitatively, using experts, policy makers, or stakeholders' priorities or by performing an interview among the interest groups. In nine studies (17%), the data-driven approach was applied to weight the criteria (e.g., Akın, Sunar, and Berberoğlu, 2015; Bozkaya et al., 2015; Cheng and Masser, 2004), while 35 studies (66%) employed experts' or stakeholders' knowledge (e.g., Li and Liu, 2007; Manganelli, et al., 2016; Nourqolipour et al., 2015; Yu et al., 2011). Li and Zhao (2017) applied both the data-driven and knowledge-based methods to calculate more accurate criterion weights. In two studies carried out by Ligmann-Zielinska (2009), and Ligmann-Zielinska and Sun (2010), the analysis was operationalized using hypothetical weights. Pair-wise comparison was found to be the most popular weighting method in the knowledge-based approach. In this approach, two criteria are compared at a given time in terms of their relative importance. It seems the theoretical background is the reason behind the popularity of the pair-wise comparison. However, it may be too difficult for stakeholders or non-experts to express their preferences in the pair-wise comparison method. In many studies, authors seem to not recognize the difference between AHP and pair-wise comparison procedure (e.g., Chowdhury and Maithani, 2014; de Noronha et al., 2012). Pair-wise comparison is only a part of AHP method. In the AHP approach, a hierarchical structure of goal, objectives, and attributes needs to be developed and then the pair-wise comparison procedure is used for assessing the relative importance of the elements of the hierarchical structure (Saaty, 1980).

2.4.3 Input data characteristics

2.4.3.1 Data model

There are two general classes of data models in GIS: raster data models and vector data models (Burrough, 1986). Accordingly, one can distinguish between two classes of GIS-MCA: vector-based GIS-MCA and raster-based GIS-MCA (Malczewski, 1999). The spatial data model is a very important component in geosimulation-multicriteria analyses because it represents the landscape within which land-use/cover changes occur. If the landscape is represented by a raster data model then each cell can be seen as a land parcel. If the vector data model is used, then land parcels, which are the basic units in the modelling process, are represented by polygons. Although vector data models display geographical objects more accurately, raster data models have been applied more often to represent landscape. Out of 53 articles, 51 of them (96%) used the raster data model (e.g., Akın et al., 2015; de Noronha Vaz et al., 2012; Hansen, 2012; Liu et al., 2014; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013), and one study employed vector data model (Bone et al., 2011), and one did not include any details about the spatial data model (Pooyandeh and Marceau, 2013). This tendency towards raster data models can be attributed to the availability of satellite images ranging from high spatial resolution to low spatial resolution. Satellite images are also updated frequently and usually have high temporal resolution. These images are sometimes available at no cost – for instance, Landsat images. Moreover, raster data structure is simple compared to the complex vector data model and the computation time in the raster data model can decrease significantly by downgrading the spatial resolution of base maps.

2.4.3.2 Base map properties

The result of geosimulation-multicriteria analysis heavily depends on the accuracy of input data. Table A1 shows that the geosimulation-multicriteria analysis studies used either satellite imageries or land-use/cover maps (see Appendix A). Although there are other satellites with higher spatial resolution, Landsat images were directly used in 20 out of 53 studies (38%) as the base map (e.g., Mahiny and Clarke, 2012; Moghadam and Helbich, 2013; Singh et al., 2015; Wu, 1998). This can be attributed to the Landsat historical archive (high temporal resolution) and its availability at no cost through the U.S. Geological Survey website. The resolution of the base maps in the geosimulation-multicriteria analysis models ranges from 3 meters (Terra et al., 2014) to 1 kilometer (Chowdhury and Maithani, 2014; Li and Zhao, 2017). Downgrading the spatial resolution can significantly decrease the processing time; however, the accuracy of classification will decrease because if each cell is assigned to just one land-use type then there would be a vast amount of information loss. 30 meters and 100 meters are the most frequent spatial resolution (58% of articles). 30 meters is the typical spatial resolution associated with Landsat images used in 18 articles (e.g., Bozkaya et al., 2015; Ghavami and Taleai, 2016; Moghadam and Helbich, 2013). 100 meters is a popular resolution because the execution time of the model is substantially less than 30 meters and the spatial resolution is still fine. The 100-meter resolution was applied in 13 papers (e.g., Cheng and Masser, 2004; de Noronha Vaz et al., 2012; Hansen, 2010; Hosseinali, Alesheikh, and Nourian, 2015). Execution time is even more important for geographically large areas. Due to this fact, some authors selected images with coarse spatial resolution (e.g., Cheng and Masser, 2004; Liu et al., 2014). Also, it seems that information loss by considering 100 meter-cells is not very large with respect to the size of the city (e.g., Liu et al., 2014). This is the reason why some authors (e.g., Li and Zhao, 2017; Li and Liu, 2007) resampled the images and decreased the spatial resolution before inputting data into the model. However, choosing a proper spatial resolution is critical; if a very coarse resolution is used in a model, the results can be misleading. For example, a land-use change model with 1 kilometer cell size may show higher accuracy than another model with 10 meter cell size; however, the result of the accuracy assessment heavily depends on the number of cells in the area. Therefore, it is highly recommended to consider the spatial resolution of models while comparing their performances.

2.4.3.3 Classification of land-use types

The number of land-use types considered in a study depends on the objective of the modelling process. If a study aims to model built-up areas, then using two land-use classes (built-up and non-built-up lands) can be sufficient. This is because it does not make any difference whether the built-up area is commercial, industrial, or residential. This classification approach was employed in some studies, for example, research carried out by Chowdhury and Maithani (2014), Cheng and Masser (2004), Jokar Arsanjani, Helbich, and de Noronha Vaz (2013), and Pontius and Malanson (2005). The two land-use classification strategies can also be applied when the focus of the research is on simulating residential growth (e.g., Hosseinali et al., 2013). Ligmann-Zielinska and Jankowski (2010) considered a separate class for restricted area along with the two aforementioned classes. The highest number of land-use classes used in the geosimulation-multicriteria studies is 20 (Ligtenberg et al., 2004). There was no information about the number of land-use classes in three studies conducted by Pooyandeh and Marceau (2013), Sakieh et al. (2015), and Liu et al. (2014). 37 out of 53 studies (70%) used seven land-use classes or fewer (e.g., Pontius and Malanson, 2005; Sun, et al., 2013; Wu and Webster, 1998; Zhang, et al., 2011). One conclusion emerging from the survey is that if it is not feasible to make accurate transition rules from one cell state to another or there is no need to simulate the transition from a specific type to another, it is better to reclassify images to have as small a number of classes as possible. This strategy was adopted by some authors to reduce the complication of the modelling process (e.g., Cheng and Masser, 2004; Chowdhury and Maithani, 2014; Hosseinali et al., 2013).

2.4.3.4 Classification of agents

Whether the geosimulation model is CA or MAS, agents are the driving force behind urban dynamics. As explained, CA-based models are special cases of agent-based models (see Section 2.4.1). In CA-based models, cells can be recognized as the agents that act and interact to create a pattern. Their action is associated with changing states over time based on some transition rules. The interactions of cells are related to the fact that the state of any cell in the future is a function of the state of its neighbouring cells. In all CA-based multicriteria studies, land-use/cover cells were considered as the agent (e.g., Cheng and Masser, 2004; Sun et al., 2013; Surabuddin Mondal et al., 2013; Bozkaya et al., 2015). In MAS/ABM multicriteria models, individuals or interest groups were considered as agents. For example, Li and Liu (2007) considered three types of interest groups that have impact on the land-use changes: residents, real estate developers, and governments. They went further and categorized residents based on the income and structure into six groups: low-income without children, middle-income without children, high-income without children, low-income with children, middle-income with children, and high-income with children. Bone, Dragicevic, and White (2011) recognized households and commercial enterprises as the influential actors behind the land-use change. Hosseinali, Alesheikh, and Nourian (2013, 2015) identified five groups who are engaging in the urban development process in the city of Qazvin, Iran: young person, high-income developers, rich people, low-income people, and moderate to low-income people. Jokar Arsanjani, Helbich, and de Noronha Vaz (2013) considered three interest groups that are actively involving in the residential growth process in Tehran: residents, real estate developers, and governments.

2.4.4 Software

Most of the geosimulation-multicriteria studies are based on the modelling capabilities of raster-based software such as IDRISI (Eastman, 1997), and SLEUTH (Clarke, Hoppen, and Gaydos, 1997), or multi-agent based packages such as NetLogo (Wilensky, 1999), and REPAST (Collier, Howe, and North, 2003). IDRISI was the most often used software in geosimulation-multicriteria studies. According to the survey, 15 studies (~28%) used IDRISI (e.g., Henríquez, Azócar, and Romero, 2006; Hyandye and Martz, 2017; Pontius and Malanson, 2005; Singh et al., 2015). In 7 articles, authors developed their own framework using application programmer interfaces (API) or software libraries, such as ArcObjects (Burke, 2003), REPAST libraries (Collier, Howe, and North, 2003), or SWARM libraries (Hiebeler, 1994). For instance, Ligtenberg et al. (2001) employed JAVA programming language to extend their model using SWARM library. Ghavami and Taleai (2016) developed their own framework using the C++ programming language and there is no indication if they used APIs or software libraries. In some studies, authors prepared data in remote sensing software and then transferred the output to another software package to analyse it. For example, Cheng and Masser (2004) used an integration of ERDAS and ArcView. Park et al. (2011) and Park et al. (2012) applied ERDAS/ArcGIS/SPSS to prepare and process the data and analyse the outputs. 13 out 53 studies did not report the type of software they used to implement a geosimulation-multicriteria model. Also, there was a conceptual framework without implementation in an article by Sabri, Ludin, and Ho (2012).

2.5 Conclusions

The quality and quantity of studies in the field of geosimulation-multicriteria has increased substantially in the past six years. Academics and practitioners who are working on urban dynamics and land-use/cover changes recognize the benefits of integrating multicriteria analysis into geosimulation methods. One of the benefits of geosimulation-multicriteria analysis is its ability to help decision-makers or urban planners develop future scenarios for an urban area based on their priorities and judgments. The results of geosimulation-multicriteria studies can help in the process of designing infrastructures, such as transportation networks, based on future urban structure and demands.

This review has revealed some gaps in the literature that need to be addressed in the future. First, most research applied WLC or AHP as the decision rule and there is not enough research in geosimulation-multicriteria using other decision rules or procedures such as the ideal point method, OWA, and fuzzy operators. Second, all previous studies used global decision rules to generate suitability/utility maps. Using global methods, researchers seem to implicitly assume that the parameters of decision models are the same across the study area. Malczewski (2011) suggested that the parameters of decision models vary from one location to another based on the characteristic of the location (see also Malczewski and Liu, 2014). Third, most studies used pair-wise comparison to evaluate the degree of importance associated with the driving forces behind urban dynamics. However, pair-wise comparison seems to be overcomplicated for non-experts. Fourth, there is a lack of enough research using participatory GIS – one of the fastest-growing disciplines within GIScience – in the modelling process.

Chapter 3

3 Study area

3.1 Introduction

It is estimated that Asia contains half of the population living in urban centres of more than 500,000 people in the world (Cox, 2015). There are 34 megacities in the world (a megacity is defined as metropolitan area with a population of more than ten million people). The metropolitan area of Tehran with a population in excess of thirteen million is such an urban area and ranks as the 22nd largest urban area (Table 3.1). However, the physical size of Tehran metropolitan area is not that large in comparison with other megacities – only being the 65th largest in the world in terms of the total area. It is worth mentioning that the definition of an urban area is different from the definition of a city. A city is usually a part of an urban area that is distinguished by administrative boundaries. In the case of Tehran, the urban area consists of the city of Tehran and a number of its satellite cities and towns. In the following sections some of the most important characteristics of Tehran are discussed. All statistics presented here are taken from Statistical Center of Iran (2017a) and Tehran Municipality (2017a).

Table 3.1: Largest metropolitan areas in the world (Cox, 2015)

Rank	Geography	Urban Area	Population Estimate	Year	Land Area (Km^2)	Population Density
1	Japan	Tokyo-Yokohama	37,843,000	2015	8,547	4,400
2	Indonesia	Jakarta	30,539,000	2015	3,225	9,500
3	India	Delhi	24,998,000	2015	2,072	12,100
4	Philippines	Manila	24,123,000	2015	1,580	15,300
5	South Korea	Seoul-Incheon	23,480,000	2015	2,266	10,400
6	China	Shanghai	23,416,000	2015	3,820	6,100
7	Pakistan	Karachi	22,123,000	2015	945	23,400
8	China	Beijing	21,009,000	2015	3,820	5,500
9	United States	New York	20,630,000	2015	11,642	1,800
10	China	Guangzhou	20,597,000	2015	3,432	6,000
11	Brazil	Sao Paulo	20,365,000	2015	2,707	7,500
12	Mexico	Mexico City	20,063,000	2015	2,072	9,700
13	India	Mumbai	17,712,000	2015	546	32,400
14	Japan	Osaka	17,444,000	2015	3,212	5,400
15	Russia	Moscow	16,170,000	2015	4,662	3,500
16	Bangladesh	Dhaka	15,669,000	2015	360	43,500
17	Egypt	Cairo	15,600,000	2015	1,761	8,900
18	United States	Los Angeles	15,058,000	2015	6,299	2,400
19	Thailand	Bangkok	14,998,000	2015	2,590	5,800
20	India	Kolkata	14,667,000	2015	1,204	12,200
21	Argentina	Buenos Aires	14,122,000	2015	2,681	5,300
22	Iran	Tehran	13,532,000	2015	1,489	9,100
23	Turkey	Istanbul	13,287,000	2015	1,360	9,800
24	Nigeria	Lagos	13,123,000	2015	907	14,500
25	China	Shenzhen	12,084,000	2015	1,748	6,900

3.2 Geography

3.2.1 Location and administrative districts

Tehran province along with thirty other provinces constitutes the territory of Iran. It is located in the north-central part of the Iranian plateau (see Figure 3.1). The province of Tehran is the most populous province in Iran by a large margin. According to the 2016 census data, the population of Tehran province accounts for more than 16% of the total population of Iran (Statistical Centre of Iran, 2017b). However, with an area of 13640.30 km^2, Tehran province is the third smallest province by total area. The province is home to the capital and most populous city of Iran, the city of Tehran (see Figure 3.1).

The city of Tehran has a population of 8.154 million and an administrative area of 730 km^2. Tehran extends from 35°34' north to 35°50' north latitude and from 51°2' east to 51°36' east longitude. Tehran shares borders with Karaj and Shahriar to the west, Kan to the north-east, Shemiranat to the north, Damavand to the east, and Rey, Pakdasht, and Eslamshahr to the south.

The administrative borders of the city have been changed a few times over its history. However, the administrative borders of Tehran have remained unchanged since 1999. Twenty-two administrative districts form the political boundary of Tehran (Figure 3.2). District 4 in the north-east of the city is the largest district by the total area and district 10 in the centre is the smallest one.

3.2.2 Topography

The territory of Tehran is formed by three types of landscape: mountainous, mountainside, and desert. Tehran is surrounded by the Alborz Mountains on the north, north-west, north-east, and part of east. The Alborz Mountains separates the Iranian plateau from the Caspian plain. The Sorkhe Hesar forest is located to the east and south-east of the city (see Figure 3.3). The Varamin desert and swath of farmlands lie in the south of Tehran and make it unsuitable for residential development. Accordingly, based on these geographic conditions, the west part of the city is more conducive to residential growth. Unlike some other big cities in Iran, no major river runs through the city of Tehran. The altitude in the residential areas varies from 1800 meters above sea level in mountainous lands in the north of the city to 900 meters in the southern parts. This difference in elevation has some impacts on the physical and social characteristics of Tehran. For example, the northern part of the city has been occupied by more affluent families and the southern part by less affluent residents.

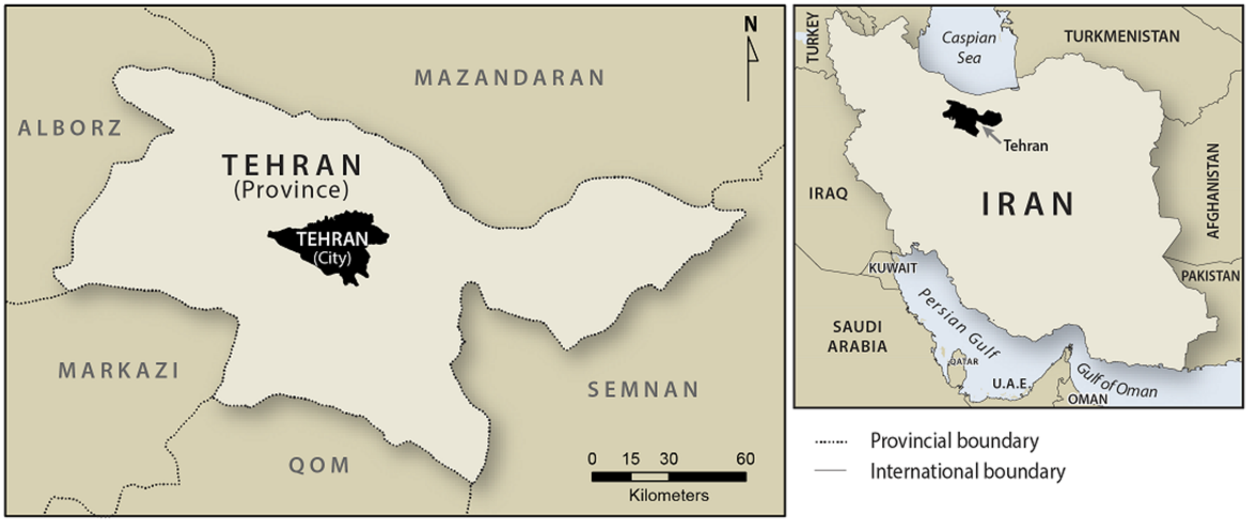


Figure 3.1: The location of Tehran province and Tehran city

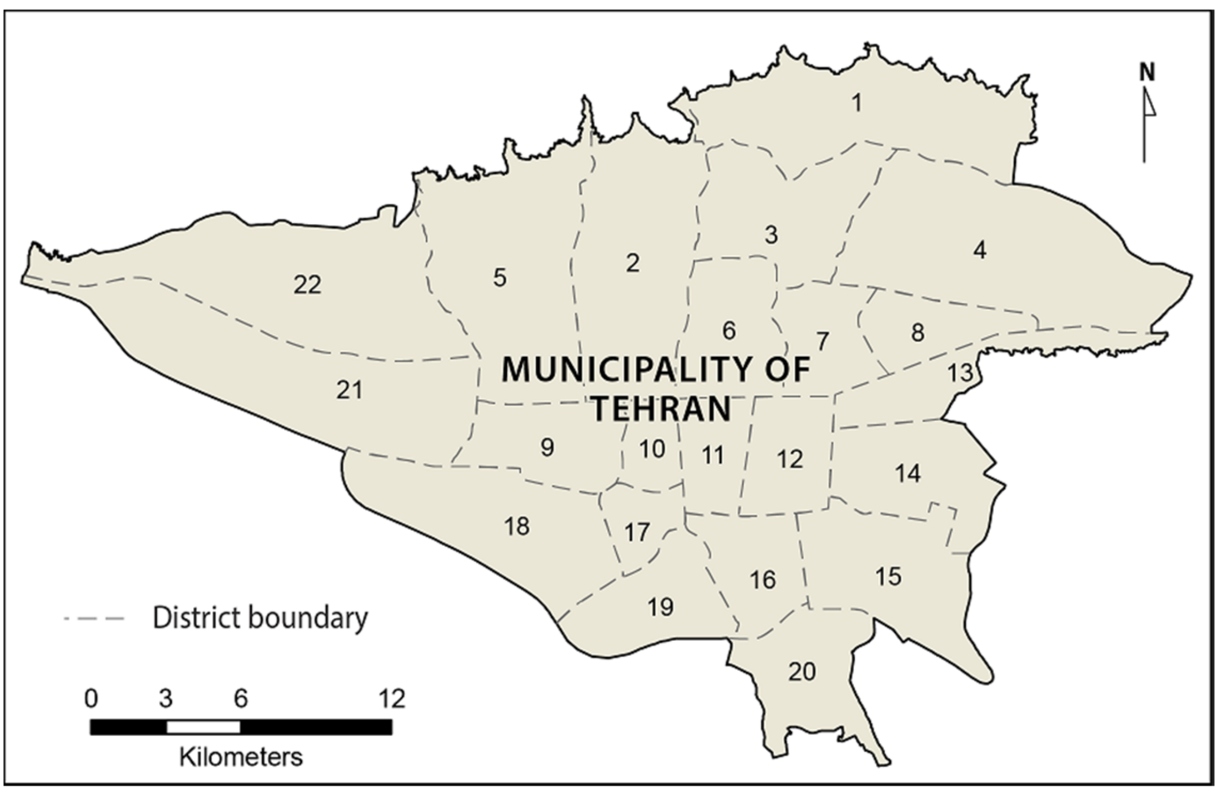


Figure 3.2: The administrative districts of Tehran city

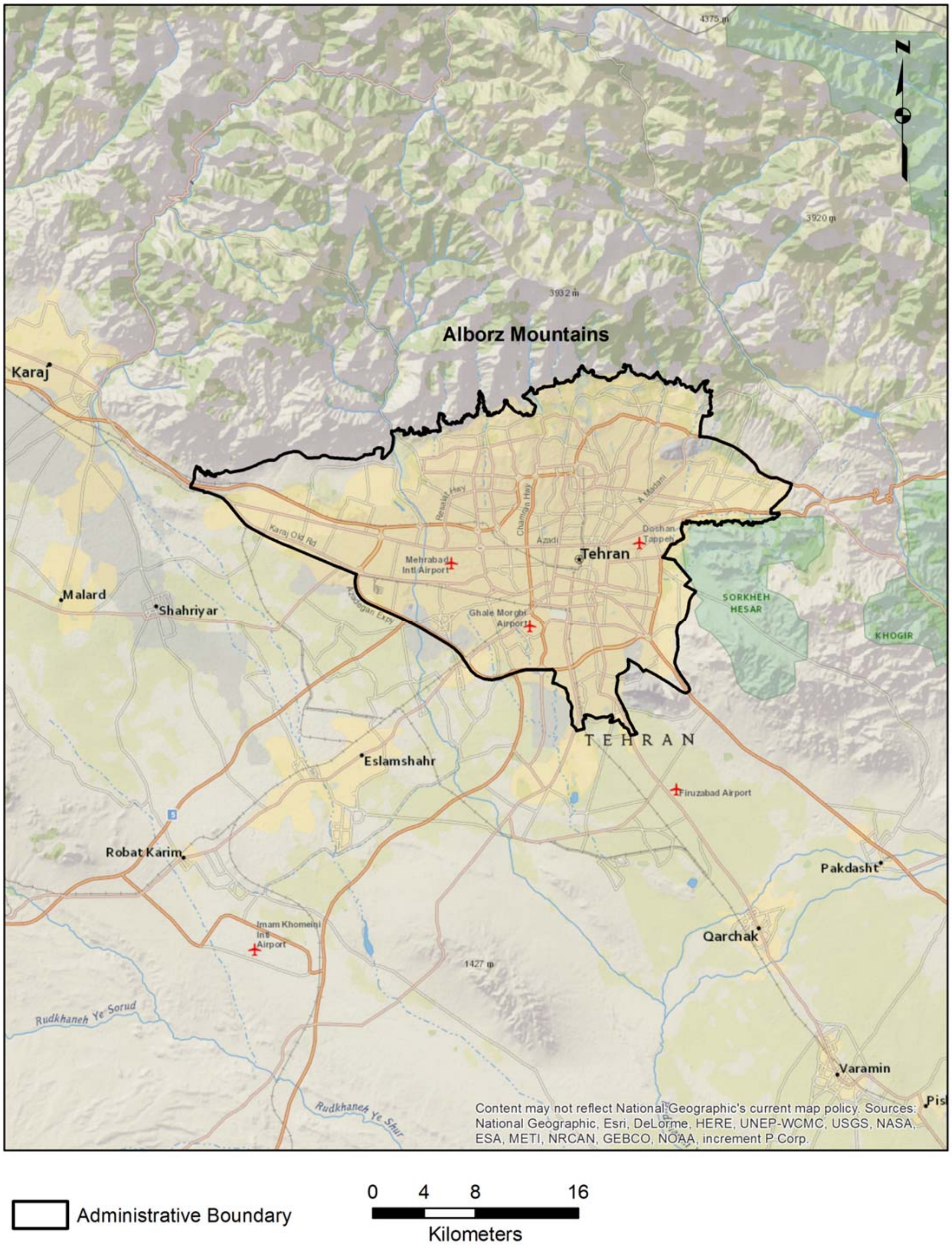


Figure 3.3: The landform of Tehran

3.2.3 Climate

The climate of the city is affected by large differences in elevation. Northern and north-western parts of the city are characterized by more temperate weather; other parts of the city, especially the southern parts, are characterized with a semi-arid climate. Statistics show that the average temperature in a long term for the northern part of the city is 15.4 °C, while the average temperature at the same time was 17.3 °C in the south-western, 17.4 °C in the western, and 17.8 °C in the eastern part (Tehran Municipality, 2017b). Tehran's weather as a whole can be described as very hot in the summer, not very cold in the winter, and moderate in the autumn and spring. There is also considerable spatial variability in precipitation in Tehran. The average annual precipitation varies from 422 mm in the foothills of the Alborz Mountains to 145 mm in the southern part of the city (Tehran Municipality, 2017b).

3.3 Population

3.3.1 Population growth

Census data shows that the population of Tehran has increased significantly in the past century (see Table 3.2). The city used to be a small town in 1905 with around 147,000 residents. In the late 1920s, it started to grow both in size and population. The city of Tehran accounted for less than 2% of the total population of Iran in 1905 (Figure 3.4). In the late 1970s and the early 1980s it reached more than 13% and now it stands at almost 11%. Nowadays, the city of Tehran with more than eight million people and a population density in excess of 9000 people per square kilometer is considered a dense city by the world standards. In the early 1990s, Tehran Municipality and the Iranian government took some measures to slow down the trend of population growth in the city and freeze the population of the city at around seven million. However, Figure 3.5 indicates that these efforts were without any significant success. In recent years, the rate of growth is even worse than it was thought to be in some parts of the city. To name a few, District 22 was developed as a major tourist and recreation centre and the limit of 500,000 residents that had been established by the central government for 2025 has been exceeded a decade earlier. The pace of the development in the region was somehow out of control in the past few years. A large number of skyscrapers and commercial centres have been constructed in the region recently. According to the new estimations, there will be over 1,000,000 residents in the district by 2025, which is twice as what was once planned.

Table 3.2: Changes in the population of Tehran (Source: Tehran Municipality, 2017c; Ranji, et al., 2013)

Year	Population (million)	Growth rate (%)
1905	0.15	2.9
1930	0.25	2.4
1940	0.54	6.6
1956	1.56	5.5
1966	2.72	5.1
1976	4.53	2.9
1986	6.06	1.3
1991	6.5	.78
1996	6.76	1.3
2006	7.71	1.1
2011	8.15	1.4
2016	8.74	

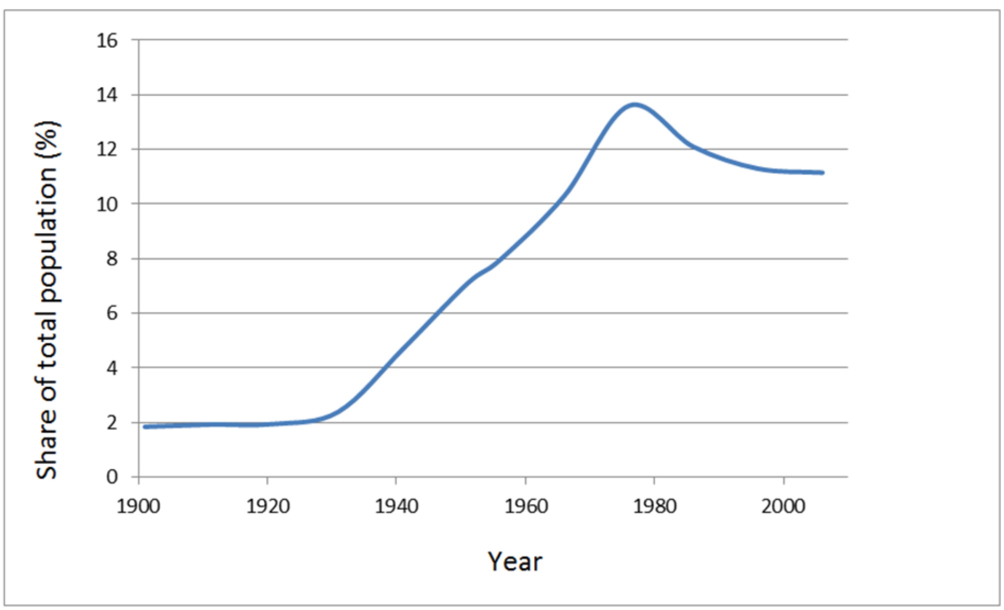


Figure 3.4: Changes in Tehran population with respect to the total population of Iran (Source: Tehran Municipality, 2017c)

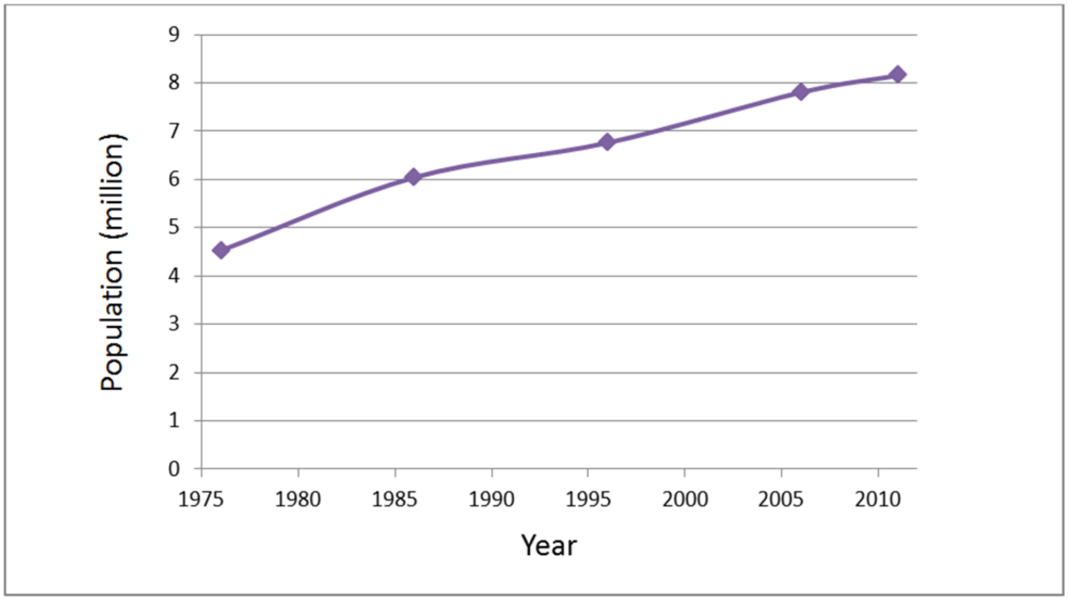


Figure 3.5: Changes in the population of Tehran (Source: Tehran Municipality, 2017c)

3.3.2 Population density

Statistics about population density show that there has been a large difference between the population density of the northern and southern parts of Tehran (Tehran Municipality, 2017d). Figures 3.6 and 3.7 represent the population density of Tehran administrative districts for 1996 and 2006. In 1996, the population density in the southern part of the city was between 300 and 412 persons per hectare; however, the population density in the northern part varied between 40 to 90 persons per hectare (Tehran Municipality, 2017d). According to the 2006 statistics, the gap between northern and southern population density has narrowed.

In 1996, the lowest population density could be found in the west and southwest sections of Tehran. In some western parts of the city, the population density is as low as one person per hectare. The reason for this is that the Tehran municipality started to provide urban facilities and services to those parts only in the early 2000s (Tehran Municipality, 2017d). Having low population density and offering urban facilities and services made the western part of the city more suitable for residential development in the past two decades. Although the population density of some counties in the western parts of the city has increased from 1996 to 2006, western districts still possess the lowest population density.

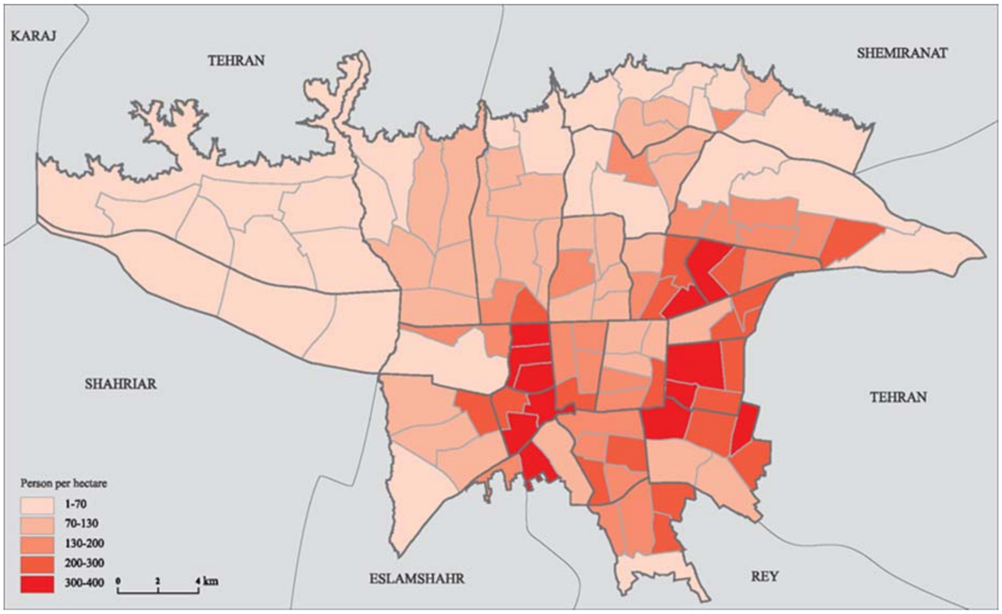


Figure 3.6: Population density of Tehran in 1996 (Source: Tehran Municipality, 2017d)

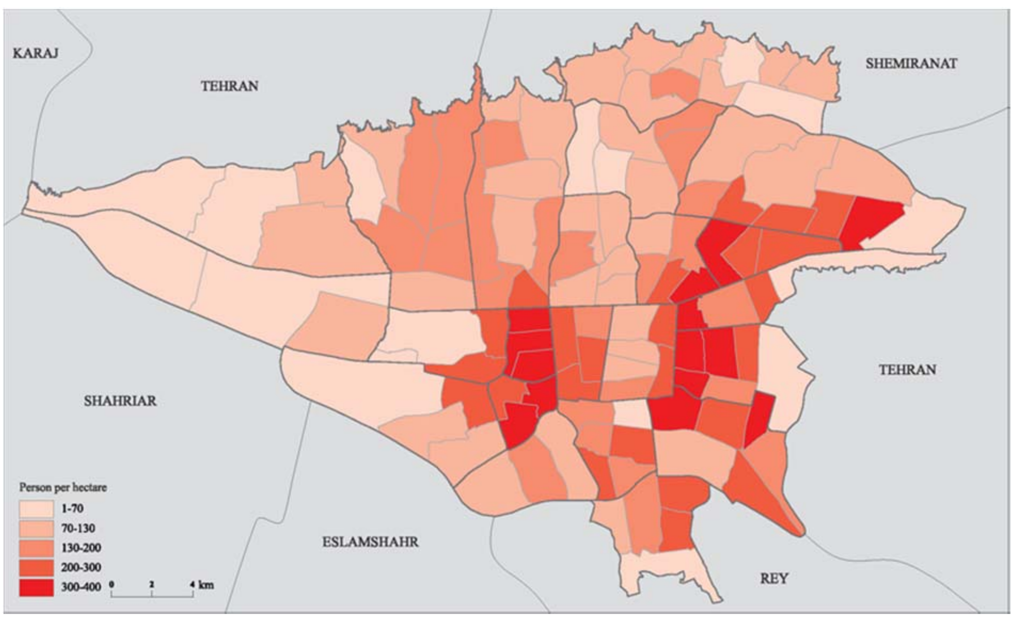


Figure 3.7: Population density of Tehran in 2006 (Source: Tehran Municipality, 2017d)

3.4 Spatial structure

Urban structure and land-use patterns are formed primarily based on job opportunities and market forces of the city core or cores (Bertaud, 2003). Although most retail activities, wholesale trading, light manufacturing, and financial activities concentrate in the central part of

Tehran, it does not exactly follow the traditional concept of monocentric cities. However, Tehran had witnessed the growth of several internally functioning core areas that are somehow dependent on the central business district (CBD). If we examine the population density gradient over the city it is quite different from what is common in a monocentric city. Population density gradient is a measure used to describe how population density changes with distance. In cities with a strong CBD, as one moves away from the central city the population density drops gradually. However, it is not the case for Tehran. Figure 3.8 indicates how the population density of built-up areas changes with distance from the CBD (Bertaud, 2003). The population density of Tehran increases with distance to the CBD and in 6 kilometers it reaches its maximum value, then it decreases. This suggests that jobs and retail activities do not concentrate in the CBD of the city and therefore, Tehran lacks a dominant CBD and its structure can be generally described as mildly polycentric (Bertaud, 2003).

Moreover, there is other evidence to support the claim that Tehran has a weak CBD. The pattern of land price in Tehran cannot be described by referring to proximity to the CBD. The price of land is much higher in the northern part of the city than the central parts (Bertaud, 2003). In the foothills of the Alborz Mountains, the most expensive land parcels can be found. The pattern of land price in Tehran is heavily related to environmental quality (Bertaud, 2003). Since northern districts are located in a higher elevation, the weather is more temperate. Also, the level of Tehran's notorious air pollution reaches its minimum in northern parts.

A set of driving forces attract residents to the peripheral and less-central districts, mostly in the northern, western, and north-western sections. The most important factors are:

(i) Reduction in travel costs: The improvement of transportation systems, construction of broad highways, and extending subway routes reduced both the monetary and time costs of transport from suburbs to the city centre (Bertaud, 2003).

(ii) Social and economic problems: Although Tehran is not completely a monocentric city, according to data from Tehran municipality the central district of the city is recognized with some urban ills such as a high rate of crime, dilapidated houses, traffic congestion, and air and noise pollution. Moreover, peripheral sections have better access to recreational and sport facilities. Therefore, some households prefer to settle in less populated peripheral areas to avoid the difficulties of living in the central areas (see Figure 3.9).

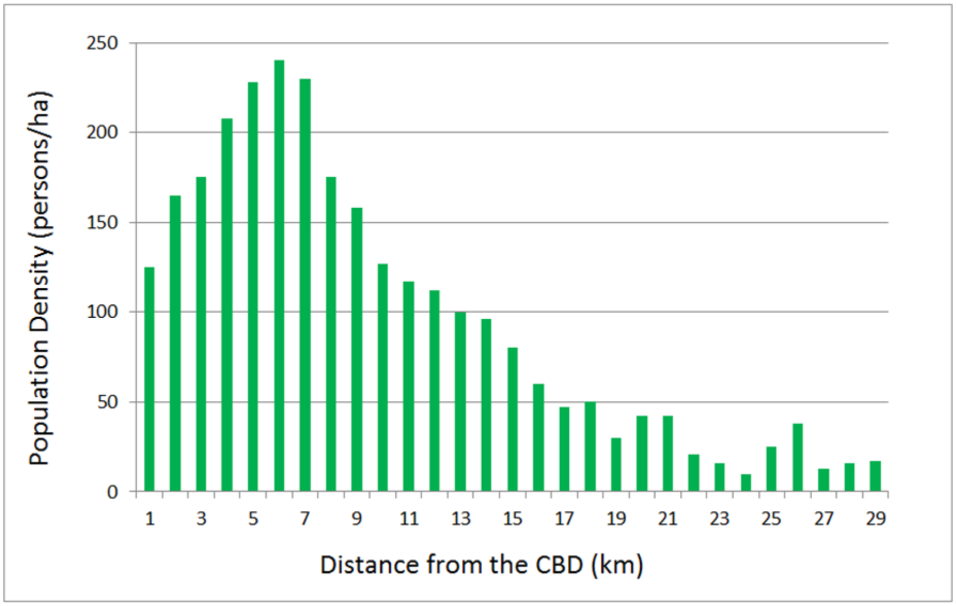


Figure 3.8: Changes in the population density of built-up areas (source: Bertaud, 2003)

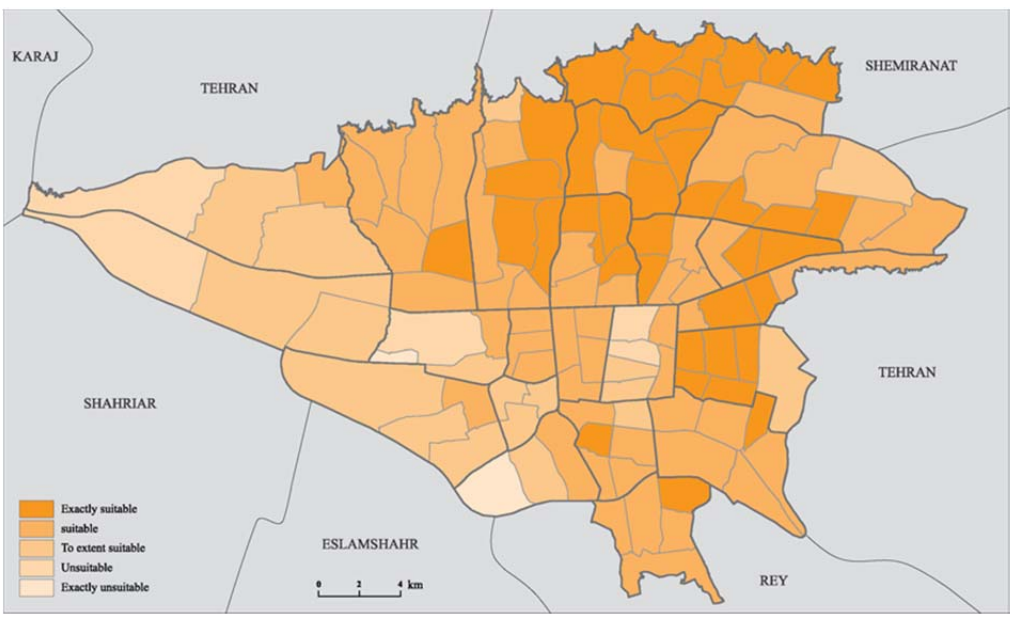


Figure 3.9: The quality of life of Tehran in 2006 (Source: Tehran Municipality, 2017e)

Chapter 4

4 Methods: Geosimulation-multicriteria models

4.1 Introduction

This chapter describes the geosimulation and multicriteria models and discusses the integration of these two approaches. Also, the actors engaging in the residential growth process in Tehran are introduced and their roles in the geosimulation-multicriteria procedure are explained.

4.2 Geosimulation

Geosimulation tries to understand macro-scale spatial dynamics and patterns by modelling actions and interactions of individual entities, such as local governments, stakeholders, land owners, and households (Benenson and Torrens, 2004a). According to Benenson and Torrens (2004a), cellular automata (CA) and multi-agent systems (MAS) are two major classes of geosimulation models. Applying these two models, macro-scale spatial patterns can be examined by simulating the actions and interactions of individual entities at the micro-scale level. In Sections 4.2.1 and 4.2.2, the fundamentals of cellular automata and multi-agent systems/agent-based models will be discussed.

4.2.1 Cellular Automata

CA-based models were introduced by Neumann and Burks (1966) to provide a framework for examining the behaviour of complex systems. CA can model complex patterns based on some simple, local rules (Liu, 2008). It describes a complex system by simulating interactions among simple entities of the system. In this approach, the space is divided into a grid of cells, each of which interacts with its neighbours, in which time advances in discrete steps. A basic CA model consists of five components: a grid of cells, a neighbourhood, transition rules, cell state, and time.

4.2.1.1 Space

A grid of cells provides the space within which CA models are implemented. In the simplest situation, the grid of cells can be one-dimensional that corresponds to a line of cells (Figure 4.1). The dimension of a grid of cells can theoretically be any finite number. The shape of cells is usually defined by a regular polygon, such as a square or a hexagon.

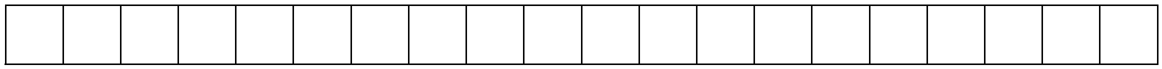


Figure 4.1: One-dimensional grid of cells

In urban studies a grid of cells can be applied to model the geographical landscape. A two-dimensional grid of square cells is the most common representation of space in urban studies. In this approach, the urban area is considered as a set of cellular automaton. A cellular automaton resembles a parcel of land that is associated with a finite number of possible states (Liu, 2008). The set of possible states can be based on land-use types, development status, the probability of development, and so on. The behaviour of each single parcel of land is controlled by transition rules (Wu, 1998).

4.2.1.2 Neighbourhood

A neighbourhood is composed of a target cell and its surrounding cells. Every cell is only interacting with its neighbours. Therefore, CA-based models are most suitable in situations where an interaction between entities and their immediate neighbours generates the current pattern, such as diffusion (Liu, 2008). In one-dimensional CA-based models, the neighbours of any cell are identified by considering d cells on each side of the target cell. d is called order of the contiguity. Figure 4.2 illustrates a situation where $d = 2$ in a one-dimensional CA model.

In two-dimensional CA models, a number of methods for neighbourhood definition can be applied; the von Neumann neighbourhood and Moore neighbourhood are the most often-used approaches (see Figure 4.3). In the first approach, the target cell, together with its four immediate non-diagonal surrounding cells, form the neighbours (Figure 4.3a). In the second one, the target cell and its eight surrounding cells define a neighbourhood (Figure 4.3b). von Neumann and Moore neighbourhoods are also referred to as Rook and Queen contiguity in the literature, respectively. These two types of neighbourhood can be extended to consider the influence of a greater number of cells on the target cell. Figure 4.3c shows the extended Moore neighbourhood where d=2. The size of the neighbourhood is $2d$+1, which is equal to 5 in Figure 4.3c. Moreover, the range of the neighbourhood can be obtained by multiplying d and cell size.

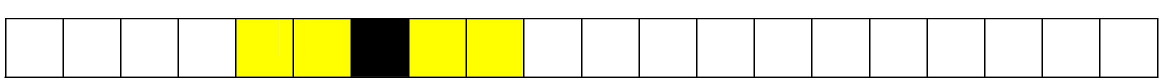


Figure 4.2: Neighbourhood definition in a one-dimensional grid of cells (d = 2)

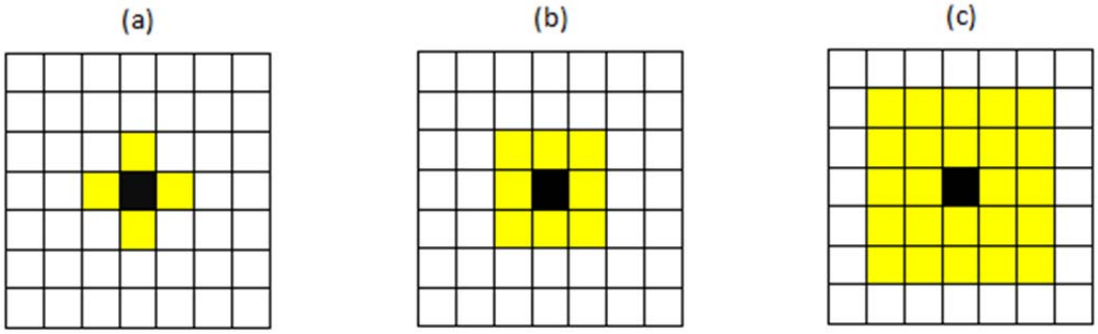


Figure 4.3: Neighbourhood definition: (a) von Neumann (or first-order Rook contiguity); (b) Moore (or first order Queen contiguity); (c) extended Moore (or second-order Queen contiguity)

4.2.1.3 Cell State

Cell state shows the possible values that a single cell can take. It can be expressed in one of the following ways: (i) binary (or 0 and 1) values (e.g., developed and non-developed parcel of land) (Figure 4.4), (ii) quantitative values (e.g., development probability), or (iii) qualitative values (e.g., residential, commercial, and industrial land-uses). In urban studies all three of these methods have been applied frequently. For example, Yu et al. (2011) applied the transition probability of each cell from the current land-use type to other types to model land-use changes. In another study conducted by Li et al. (2011), the development probability was assigned to each cell to simulate an urban expansion.

0	0	0	1	0	1	1	0	0	1	0	1	1	1	0	1	0	0	1	0

Figure 4.4: One-dimensional grid of cells with binary values

4.2.1.4 Time

Time represents the temporal scale in CA modelling. Each cell (parcel of land) is assigned a single value that denotes its state at a given time. The state of all cells is subject to changes at the next time depending on the transition rules. For example, the state of some cells may change from non-developed at time t = 0 to developed at time t = 1 and vice versa. Since time is discrete in CA modelling, choosing suitable time steps has a huge impact on the performance of a CA-based model. It is better to select time steps in a way that the system under examination shows a significant change with respect to its previous state. In most urban research, it is assumed that all transition rules are applied at each time step and also the state of each cell either remains the same or completely changes in the next iteration (Liu, 2008). Cecchini and Rizzi (2001) developed an urban model by considering two types of transition rules; one applied in all time steps and the other operationalized just at certain time steps. Stevens and Dragićević (2007) suggested that not all changes in land-use types begins at the same time and occurs at the same pace. They believe that some developments take few months and some others take several years based on the size of the project. Accordingly, they introduced a high temporal resolution CA-based urban model to incorporate the amount of time that it takes for each parcel of land to be fully developed.

4.2.1.5 Transition rules

Transition rules are the most crucial aspect of CA-based models. These rules define the basic algorithms to simulate real-world processes in cellular environments. Transition rules specify the behaviour of cells between time steps based on the current cell state and the state of its surrounding cells. Indeed, transition rules provide developers with a tool to realize what would be the new state of cells after any changes or what would be the conversion probability of each single cell from the current state to other states during the process. The notion that local interactions in the previous state have influence on the future state of the landscape provides the basis for extracting transition rules (Liu, 2008). CA-based modelling processes can be described using the following formula (Wu, 1998):

$$S_{ij}^{t+1} = f\left(S_{ij}^{t}, \Omega_{ij}^{t}, T^{t}\right) \tag{4.1}$$

where S_{ij}^{t+1} and S_{ij}^{t} are the state of the *ij*-th cell *ij* at time *t*+1 and *t*, respectively; Ω_{ij}^{t} is the state of the cells in the neighbourhood of *ij*-th cell; and T^{t} is a set of transition rules.

Transition rules can be implemented in a number of ways. The simplest method is to explicitly define the outcome of each transition rule based on a possible configuration state of neighbouring cells. For example, in a one-dimensional grid of cells and by considering $d = 1$, the following transition rules can be defined for a binary value state.

(0,0,0) → 0; (0,0,1) → 0; (1,0,0) → 0; (1,0,1) → 1;
(1,1,0) → 1; (0,1,1) → 1; (0,1,0) → 0; (1,1,1) → 1;

0 can be seen as undeveloped cells and 1 as developed cells. These rules can be interpreted in this way: an undeveloped cell remains undeveloped in the next time, unless it has two developed neighbours; also, a developed cell remains developed in the next time, unless it has two undeveloped neighbours.

In situations where there are a large number of possible states, this approach is very tedious and inefficient. The better approach is then to use “IF … THEN” statements. In this approach, the definition of transition rules is more efficient; for example:

IF “Distance to main roads” < 1km AND “Land-use type” = ‘farmland’
THEN “The probability of development” = 0.9

4.2.1.5.1 Defining transition probability using multicriteria analysis

In urban CA-based models, transition rules define how cities work through a set of iterative rules (Torrens and O'Sullivan, 2001). Transition rules in urban practices can be extracted in different ways, such as regression analysis (Sui and Zeng, 2001; Wu, 2000), artificial neural networks (Li and Yeh, 2001, 2002), and multicriteria methods (Wu and Webster, 1998). Contrary to a formal CA method that only uses the current state of a parcel of land and the state of its neighbours to extract the transition rules, in urban studies, some external forces should be taken into account as well. For example, factors such as socioeconomic measures, price of land acquisition, and accessibility can be of significant importance to the future state of a parcel of land. Also, there might be some restrictions on future land state, such as slope of the land, and government regulations. For example, development in urban areas cannot occur without government approval. Therefore, in urban development modelling there are two sets of factors: factors that contribute to urban development, and restrictive factors (or constraints) (Wu and Webster, 1998). Wu (1998) proposed an approach that integrates multicriteria analysis into CA (see also Wu and Webster, 1998). Specifically, to model urban dynamics, the CA formula (see Equation 4.1) can be modified as follows:

$$S_{ij}^{t+1} = f\left(p_{ijs}^{t}, T^{t}\right) \tag{4.2}$$

where S_{ij}^{t+1} is the state of the cell *ij* at time *t*+1; p_{ijs}^{t} is the probability of cell *ij* to be converted to the state *s*; and T^{t} is a set of transition rules; ; p_{ijs}^{t} can be calculated using the following equation (Wu and Webster, 1998):

$$p_{ijs}^{t} = \phi\left(E_{ijs}^{t}\right) = \phi\left[\omega\left(F_{ijz}^{t}, w_z\right)\right] \tag{4.3}$$

where E_{ijs} is the suitability of the ij-th cell to be converted to the state s; F_{ijz}^{t} is the evaluation score of the ij-th cell with respect to the development factor (criterion) z; w_z is the importance weight associated with criterion z; ϕ is the function that converts the suitability value into the probability of development (if a deterministic approach is applied, there is no need to convert the suitability values to probability); and ω is a combination function (multicriteria decision rule or method) that aggregates evaluation scores and their associated weights.

Combining the scores of development factors by taking their importance into account is the contribution of multicriteria analysis in defining transition probability (suitability). In this approach, a set of factors that generate urban growth patterns and also restrictive factors must be identified first. Next, the importance weights associated with the contributing factors need to be determined. Finally, a combination function is applied to aggregate the contributing factors and their associated weights. Wu and Webster (1998) used weighted summation as the combination function and determined the importance weights based on the experts' judgments:

$$\omega\left(F_{ijz}^{t}, w_z\right) = \left(\sum_{z=1}^{m} F_{ijsz}^{t} \; w_{sz}\right) \prod_{z=m+1}^{n} F_{ijsz}^{t} \tag{4.4}$$

Where F_{ijsz}^{t} is the evaluation score of the ij-th cell at time t with respect to the development factor z to be converted to state s; $1 \le z \le m$ are non-restrictive development factors and $m+1 \le z \le n$ are restrictive development factors. Equation 4.4 shows how development factors and their associated weights are combined by considering development restrictions or constraints, which determine a set of infeasible cells (parcels of land).

The output of the combination function is a single value that shows the suitability of each cell to be converted to another state in the next step of the geosimulation-multicriteria procedure. Suitability scores can be used directly in transition rules if the deterministic approach is chosen. Accordingly, any cell ij with suitability higher than a threshold value will be converted to a new state at t+1. If a non-deterministic approach is used, the suitability scores must be transformed into probability values. In a simple situation the transformation of suitability scores to probability values can be performed by the following equation (Wu and Webster, 1998):

$$p_{ij} = \exp\left[\theta\left(\frac{E_{ij}}{E_{max}} - 1\right)\right] \tag{4.5}$$

where E_{ij} is the suitability score of cell ij; E_{max} is the highest suitability score in a study area; and θ is the dispersion parameter ranging from 1 to 10.

The main advantage of the multicriteria analysis approach is that it enables researchers to create a wide range of urban growth scenarios based on different decision situations (Jiao and Boerboom, 2006). The disadvantage of using multicriteria procedures to define suitability scores (or transition probabilities) is related to the fact that the resulted values are sensitive to the criterion weights (Jiao and Boerboom, 2006).

4.2.2 Multi-agent systems

4.2.2.1 What is an agent?

Agents are software programs, which are capable of autonomous actions within their environment in order to meet their design objectives (Wooldridge and Jennings, 1995). Agents aim at solving a problem or simulating a scenario by acting on the environment and interacting with other agents and the environment. An agent interacts with other agents in a system through some types of agent-communication language (Wooldridge and Jennings, 1995). Agents are representation of micro-scale entities in the computer modelling. Humans, animals or any other dynamic object can be considered agents (Benenson, Omer, and Hatna, 2002). The way in which intelligent agents interact and cooperate with one another to achieve a common goal is similar to the way that individuals or interest groups collaborate with each other to carry out a particular task. According to Benenson and Torrens (2004a) agents are: (i) goal-directed and they change their behaviour to reach their goals, (ii) autonomous, i.e., they can act independently and produce reaction over the landscape, (iii) flexible in that they can learn from their experiences and adjust their future actions, (iv) able to interact and collaborate with each other in the environment in the way that humans cooperate with each other to fulfill a particular task, (v) are located within a specific environment, and (vi) self-interested, i.e., each agent has its own view about the desirable state of a system.

4.2.2.2 Agent-based models and multi-agent systems

Agent-based models (ABM) attempt to simulate the actions and interactions of individual agents, each of which representing a single actor or a group of actors, to explore their influences on the current state of a system. ABM is an effective tool for modelling macroscopic phenomena using individual behaviours. This class of models has typically been employed in social science studies to substantiate or represent social theories, and to simulate the behaviour of actors engaging in social interactions (Basu and Pryor, 1997). Although in conventional ABM agents are restricted to moving objects, there are a number of pioneering studies in which fixed objects with important characteristics have been regarded as agents as well. For example, in a study conducted by Chen et al. (2010) any single parcel of land is recognized as an agent. ABM can be described by four major characteristics (see Fagiolo, Windrum, and Moneta, 2006):

(i) *Microscopic perspective*. Each agent in an ABM model embodies an individual entity in the real-world. This feature facilitates the model design procedure and makes it easier to interpret the results of the model (Gilbert, 2008). For example, in an urban expansion process involving three different groups, residents, developers, and governments, each group of actors can be represented by a single agent or each individual in the real-world can be associated with one agent. It is also possible to use several agents to show a group of actors. The number of agents can be proportional to the size of the group. There is not a general answer for the question of which approach is more sensible. The number of agents is directly related to the level of decomposition, which can be defined by the requirements of a model.

(ii) *Heterogeneity*. Conventional models are based on the assumption that all agents within a group are identical in all characteristics. However, this assumption is not realistic in most cases. In an ABM model, each single agent can be defined based on the preferences and priorities of its associated entity.

(iii) *Representation of the environment*. Agents are acting over an environment with which they are interacting. It is feasible to examine how agents' actions change the environment and how the environment affects the agents' behaviours.

(iv) *Bounded rationality*. According to a bounded rationality notion, individuals are subject to three rational restrictions while making a decision (Simon, 1957): (a) limited and sometimes unreliable information regarding the decision's situations and the possible outcomes of different scenarios; (b) an individual's mind has limited cognitive capacity to assess the information; and (c) limited time is available for the decision-making process. Hence, individuals involved in a decision-making situation can only search for a satisfactory rather than optimal solution.

Although ABM and multi-agent systems (MAS) are very similar, there are some subtle differences between these two approaches. An MAS comprises a large number of agents collaborating to solve a problem that is highly complex and beyond the capabilities and knowledge of a single agent. It provides a tool for incorporating different urban actors into the simulation of urban dynamics. An MAS model for the simulation of urban dynamics is composed of two elements: a cellular space and an agent-based model (Parker et al., 2003).

4.2.2.3 Agents in urban systems

MAS and ABM are appropriate tools to simulate existing patterns in an urban area at different scales. Agents in MAS or ABM embody different individuals or interest groups, who have various roles in creating existing patterns (e.g., residential growth, segregation, or deforestation). The number of agents in the model is equal or proportional to the population of the urban area or members of interest groups. Heterogeneity is one of the major characteristics of these agents. Each group of agents has different preferences and these preferences form fundamental aspects of agents' behaviours. The interactions between these groups are the driving force behind urban patterns. Therefore, to simulate urban patterns, different types of actors that contributed to create a pattern in a region should be recognized first, and then the characteristics and preferences of each type must be modeled.

4.2.3 Cellular-based vs. agent-based models

CA can be considered as a special case of ABM\MAS. Li et al. (2011) suggest that if agents in an ABM\MAS are fixed, they function as cells in a CA model. This suggestion gives rise to some important distinctions between the two approaches: (i) the main weakness of CA modelling compare to ABM\MAS is its inability to simulate moving objects, like relocating firms, vehicles, migrating households, or pedestrians (Benenson, Omer, and Hatna, 2002; Benenson and Torrens, 2004a), (ii) CA is easier to implement, (iii) ABM\MAS needs more data for the simulation process in comparison with CA, (iv) if social and economic data are not available, CA is the better method for the simulation process, (v) determining, computing, and updating physical parameters that control the change of states, like surrounding land-use types and distance to different facilities and centres, is easier in CA modelling, (vi) although CA and ABM\MAS have been extensively applied in modelling urban dynamics, ABM\MAS have more strength in simulating the behaviours of individuals and the interactions between them (Van

Dyke Parunak et al., 2006) , and (vii) standard CA-based models have some limitations to grasp the complexity inherent in urban processes, because the homogenous cellular structure and synchronous time advancement are too inflexible (Costanza, Sklar, and White, 1990; Sklar, Costanza, and Day, 1985).

4.3 Multicriteria analysis

Urban dynamics, including urban development and land-use change, are the result of actions and interactions of different types of agents. Therefore, to simulate these dynamics, one should first simulate the decision behaviour of different groups of agents. Each group of agents has different preferences and priorities, and makes its evaluation according to these preferences. Geosimulation methods provide a platform for integrating multicriteria analysis (MCA) into group decision making processes (Malczewski and Rinner, 2015). In geosimulation, agents are regarded as decision-makers and MCA procedures are applied to describe the agents' evaluations and predict the consequences of their decisions. Using MCA methods, it is possible to model the behaviour of different groups of agents by considering both the criteria contributing to agents' evaluations, and the different preferences associated with different types of agents. Agents' preferences are mirrored in the different importance weights assigned to each criterion. The importance weights vary within different groups of agents according to their preferences and beliefs. Based on these preferences, each group of agents evaluates the parcel of land for development. The result of the evaluation is reflected in a single number that represents the suitability (probability) of urban growth or the suitability (probability) of land-use change for each land parcel.

4.3.1 GIS-based multicriteria analysis

GIS is a set of tools that helps in capturing, storing, manipulating, retrieving, managing, analysing, and displaying spatial information (Longley et al., 2001). GIS-based multicriteria decision analysis extends the concept of multicriteria analysis by placing emphasis on spatial aspects of decision alternatives and evaluation criteria (Malczewski, 1999). In fact, decision alternatives are characterized by their geographical coordinates. For instance, some attributes such as the distance from natural forests or proximity to main roads are spatial in nature and can be measured using geographical data. To equip MCA techniques with effective tools to deal with geographical data, these techniques are integrated with GIS. GIS-based multicriteria decision analysis as a process includes a series of activities. The most important steps are:

(i) *Defining evaluation criteria*. An important facet of any decision is its objectives. To quantify the level to which these objectives are satisfied, a number of attributes associated with each objective are specified. The performance of decision alternatives within different objectives are evaluated by these attributes. The set of objectives and their associated attributes is called evaluation criteria.

(ii) *Defining constraints*. All alternatives are subject to a different level of natural or artificial limitations. If a decision alternative does not comply with at least one of these limitations, it is regarded as an infeasible alternative and it will be removed from the set of decision alternatives.

(iii) *Defining decision alternatives*. This step requires generating a range of potential choices such that the predefined objectives are best attained (Keeney, 1992). A decision-maker should identify a range of alternatives and then remove some of them based on resource limitation or other constraints.

(iv) *Defining decision-makers' preferences*. Different decision-makers or interest groups have various preferences for evaluation criteria. Their preferences and priorities are reflected in different importance weights that are assigned to each criterion.

(v) *Defining value function.* Values of each alternative among evaluation criteria must be standardized before the combination procedure. The reason behind standardization is that every decision criterion is measured on the basis of different scales. For instance, the scale for measuring temperature is not comparable to the scale for measuring distance. In order to make criterion scores comparable, standardization must be performed. Value function transforms the raw criterion scores into a value that ranges from 0 (the least-desirable outcome) to 1 (the most-desirable outcome). The value function can be mathematically represented as follows:

$$a_{lz} = \vartheta(c_{lz}) = \begin{cases} \left(\dfrac{c_{lz} - \min\limits_{z} c_z}{r_z}\right)^{\rho}, & \text{if higher values are desirable (benefit attributes)} \\ \left(\dfrac{\max\limits_{z} c_z - c_{lz}}{r_z}\right)^{\rho}, & \text{if lower values are desirable (cost attributes)} \end{cases} \tag{4.6}$$

where c_{lz} indicates the raw score of alternative l in the z-th criterion; and a_{lz} represents the standardized score of alternative l in the z-th criterion, which equals $\vartheta(c_{lz})$; ϑ is a value function. ρ is a parameter that defines the shape of the value function; if 0< ρ<1, the shape of the value function is concave; if ρ>1, the value function has a convex shape; and for ρ=1, the value function reduced to a linear function; r_z is the global range of values for criterion z and calculated as follows:

$$r_z = \max_z c_z - \min_z c_z \tag{4.7}$$

where $max_z\, c_z$ and $min_z\, c_z$ are the global maximum and minimum of raw scores for the z-the criterion.

(vi) *Decision rules*. Decision rules are methods or procedures that aid decision-maker(s) in choosing the best alternative(s) among a large number of potential alternatives. Decision rules combine evaluation scores of an alternative to a single value that shows the overall performance of the alternative. Since the concept of the decision rule is one of the most crucial steps in a GIS-based multicriteria decision analysis, it will be discussed in more details in the next section.

4.3.1.1 Decision rules

Applying a proper decision rule, which is usually referred to as combination function, underpins any multicriteria analysis. Decision rules provide a platform to rank decision choices and select the best choice(s). Therefore, the final output of the decision making process strongly relates to the type of decision rule employed in order to aggregate the evaluation criteria and decision-makers' preferences. A wide range of decision rules for MCA are available, including analytic hierarchy process (AHP), ideal point method, weighted linear combination (WLC), concordance method, and ordered weighted averaging (OWA) (Malczewski and Rinner, 2015). Since WLC and OWA were selected for the present study, they are explained in more detail in the following sections.

4.3.1.1.1 Weighted linear combination

The weighted linear combination (or simple additive weighting) method is one of the most often used combination functions in GIS (Eatsman et al., 1993; Malczewski, 2000, 2006a). According to the WLC method, the overall suitability score of an alternative can be formulated as follows:

$$WLC_l(a_{l1}, a_{l2}, \dots, a_{ln}) = \sum_{z=1}^{n} w_z a_{lz} \tag{4.8}$$

$$s.t. \quad \sum_{z=1}^{n} w_z = 1$$

where WLC_l is the overall suitability of alternative l; the higher WLC_l is, the more desirable the alternative l would be with respect to the decision objectives. w_z is the importance weight associated with the z-th criterion, and a_{lz} is the standardized evaluation score of the l-th alternative with respect to the z-th criterion (see Equation 4.6).

4.3.1.1.2 Ordered weighted averaging

The OWA procedure is an extension and generalization of the most often used GIS-MCA models, including WLC and Boolean AND and OR operations (Jiang and Eastman, 2000). It consists of the following elements: (i) reordering the input data (criterion values), (ii) defining the OWA order weights, and (iii) performing an aggregation (Yager, 1996). An OWA function of dimension n is a mapping $I^n \rightarrow I$ and can be stated as follows:

$$OWA_l(a_{l1}, a_{l2}, \dots, a_{ln}) = \sum_{z=1}^{n} \lambda_z b_{lz} \tag{4.9}$$

$$s.t. \quad \sum_{z=1}^{n} \lambda_z = 1$$

$$0 \leq \lambda_z \leq 1, \quad z = 1, \dots, n$$

where OWA_l is the overall suitability of alternative l; b_{lz} represents the z-th largest elements of the input data for alternative l obtained by reordering ($a_{l1}, a_{l2}, \dots, a_{ln}$), such that $b_{l1} \geq b_{l2} \geq \cdots \geq b_{ln}$; λ_z is an order weight associated with a particular ordered position of the input data. It means, the first order weight, i.e., λ_1, is allocated to the highest input argument, λ_2 is assigned to the second highest input, and in the similar way, λ_n is allocated to the lowest input data.

There are two main indices derived from the OWA order weights that indicate the distribution of order weights and the behaviour of the OWA function in the combination process. The first index is the degree of optimism (degree of ORness), which reflects the extent to which the OWA function displays behaviour similar to the logical operator OR (Yager, 1988). It can also be deemed as an index to quantify the optimism degree of a decision-maker. Degree of optimism ranges from zero to one and can be calculated for different set of order weights by the following equation (Yager, 1988):

$$Optimism(\lambda) = \frac{1}{n-1}\sum_{z=1}^{n}(n-z)\lambda_z \tag{4.10}$$

in this equation, n indicates the number of input arguments, and λ_z is an order weight associated with the z-th highest criterion value. If the first element of the order weighting vector is one and all other elements are zero, ($\lambda_z = 0 \text{ if } z \neq 1, \text{and } \lambda_1 = 1$), then the OWA function exhibits behaviour like the logical OR operator. Under this condition, the degree of optimism reaches its highest value at one. In fact, the decision-making is on the basis of the maximum value of the input arguments (or decision criteria). This attitude towards the decision situation is known as an optimistic attitude where the decision-maker(s) concentrates on the positive aspects of decision alternatives. However, if all elements of the order weighting vector are zero except for the last element which is assigned one, ($\lambda_z = 0 \text{ if } z \neq n, \text{and } \lambda_n = 1$), then the OWA function behaves like the logical AND operator. In this case, the degree of optimism reaches its lowest value at zero. In other words, the decision-making process focuses on the minimum value of the input arguments. This attitude towards the decision situation is recognized as a pessimistic attitude in which the decision-makers emphasize the negative features of decision alternatives. In a multicriteria decision problem, an optimistic decision is made based on the criterion that achieves the maximum value for each alternative, while a pessimistic decision is made according to the criterion that has the lowest value for each alternative.

The second important index to describe a set of order weights is the measure of dispersion (or entropy). This measure defines the degree to which all input arguments (a_{l1}, a_{l2}, . . . , a_{ln}) are used equally (Yager, 1996). It describes the entropy of distribution of order weights. The measure of dispersion lies in [0, *ln n*] interval and is calculated as follows (Yager, 1988):

$$Disp(\lambda) = -\sum_{z=1}^{n}\lambda_z \, ln(\lambda_z) \tag{4.11}$$

where, λ_z represents an order weight associated with the z-th highest criterion value. The measure of dispersion reaches its minimum value when $\lambda_z = 1$ for one z and $\lambda_z = 0$ for other z values, and maximum when $\lambda_z = {}^{1}\!/_{n}$ for all z.

4.3.1.1.2.1 How to derive order weights

At the core of any OWA function is the definition of the order weights. A large variety of techniques for producing the order weights of OWA function have been introduced so far, such as maximum entropy method (O'Hagan, 1988), maximum variance method (Fullér and Majlender, 2003), maximum disparity approach (Wang and Parkan, 2005), and linguistic quantifier approach (Yager, 1996). Since the linguistic quantifier approach applies linguistic statements to derive order weights, it is more descriptive of the risk of the evaluation process. Therefore, this method was selected in this research to elicit order weights.

The theory of linguistic quantifiers was presented by Zadeh (1983) to provide an approach to translate the natural language arrangements into formal mathematical formulations (Munda, 1998). Two general categories of linguistic quantifiers can be recognized: the relative linguistic quantifiers, and the absolute linguistic quantifiers (Boroushaki and Malczewski, 2008). Statements like, almost zero, at least about one, about ten, and more than hundred are some instances of absolute quantifiers. On the other hand, relative linguistic quantifiers specify relative quantities such as few, about half, many, most, and almost all. In the OWA context, the emphasis is on a class of relative linguistic quantifiers (Yager, 1996).

According to Yager (1996), a single α value can be defined corresponding to each relative linguistic quantifier. By changing the value of α, it is possible to generate a wide range of linguistic quantifiers from “At least one” quantifier to “All” quantifier. The connection between the linguistic quantifiers and different values of α is depicted in Table 4.1. In the table, “Half” quantifier generates equal order weights for all criteria that corresponds to a situation where the risk of the evaluation is neutral. This quantifier behaves like weighted linear combination function. “At least one”, “Few”, and “Some” are associated with risky evaluation results (optimistic scenarios); whereas, “Many”, “Most”, and “All” are associated with cautious evaluation processes (pessimistic scenarios). For example, if “At least one” quantifier is used, any decision alternative that satisfies at least one of the criteria is acceptable; however, if “All” quantifier is employed, all criteria need to be satisfied by an acceptable decision alternative.

Yager (1996) suggested an approach to calculate the order weights based on importance weights of criteria and linguistic quantifiers. According to this approach, if the number of criteria is *n* and the importance weight of criterion *z* (w_z) after reordering is denoted by u_z, then the order weight of a criterion, which is *x*-th largest criterion after reordering, can be calculated as follows:

$$\lambda_x = \left(\frac{\sum_{z=1}^{x} u_z}{\sum_{z=1}^{n} u_z}\right)^{\alpha} - \left(\frac{\sum_{z=1}^{x-1} u_z}{\sum_{z=1}^{n} u_z}\right)^{\alpha} \tag{4.12}$$

Since in GIS-MCDA $\sum_{z=1}^{n} w_z = 1$, then $\sum_{z=1}^{n} u_z = 1$. Therefore, Equation 4.12 can be simplified as (Boroushaki and Malczewski, 2008):

$$\lambda_x = \left(\sum_{z=1}^{x} u_z\right)^{\alpha} - \left(\sum_{z=1}^{x-1} u_z\right)^{\alpha} \tag{4.13}$$

Table 4.1: Linguistic quantifiers and the corresponding values of α (Malczewski and Rinner, 2005)

Linguistic quantifier	At least one	Few	Some	Half	Many	Most	All
α	0.0001	0.1	0.5	1	2	10	1000

4.3.1.1.2.2 Integrating criterion weights into OWA function

The conventional form of the OWA function does not consider the decision-makers' preferences regarding different attributes or evaluation criteria in the combination process. To overcome this weakness, Malczewski (2006b) suggested an approach to incorporate the importance weights of the evaluation criteria into the OWA function. Based on this approach, the outcome of the OWA function for the alternative *l* is calculated as follows:

$$OWA_l(a_{l1}, a_{l2}, \dots, a_{ln}) = \sum_{z=1}^{n} \left(\frac{u_z \lambda_z}{\sum_{z=1}^{n} u_z \lambda_z}\right) b_{lz} \tag{4.14}$$

where OWA_l denotes the overall score of the decision alternative l; b_{lz} represents the evaluation score of alternative l in the z-th criterion obtained by reordering the criterion scores of alternative l; u_z is the criterion weight (after reordering), and λ_z is the order weight. OWA function can be reduced to WLC if all order weights are equal, i.e. $\lambda_z = \frac{1}{n}$ for all z. It means that WLC is just one special case that can be generated using OWA function.

4.3.1.1.2.3 Local ordered weighted averaging function

The conventional or global OWA function (see Section 4.3.1.1.2) is based on the notion that there is no heterogeneity in the input arguments and criteria weights (Malczewski, 2011; Malczewski and Liu, 2014). However, this premise may not be acceptable in a spatial context where we deal with a highly heterogeneous landscape (Malczewski, 2011; Malczewski and Liu, 2014). By definition, spatial heterogeneity is an uneven distribution of features, events within an area (Anselin, 2010). A heterogeneous geographical space has uneven terrain features and environmental characteristics, such as temperature and rainfall. To deal with this shortcoming of the global multicriteria analysis, Malczewski (2011) introduced the idea of local multicriteria analysis. Accordingly, one can consider two general classes of the multicriteria analysis to address spatial multicriteria decision problems: global and local (Anselin, 2010; Malczewski, 2011). The first class deems that the parameters of multicriteria analysis remain the same across the study area. Based on the global spatial multicriteria analysis, the parameters of a value function and decision-makers' preferences with respect to the evaluation criteria are the same over a geographical area. In local multicriteria analyses, the heterogeneity inherent in landscape characteristics is considered in the modelling process. Therefore, decision-makers' preferences and value function parameters are modified based on spatial properties. In other words, the combination function can be different from one location to another based on the local characteristics (Makropoulos and Butler, 2006; Makropoulos et al., 2007). Malczewski (2011) and Malczewski and Liu (2014) conceptualized this idea by highlighting the effect of the criterion range on the importance weight of a criterion. Their approach is based on the range sensitivity theory, according to which, the larger the range of the values for a specific criterion, the higher importance should be attached to that criterion (Fischer, 1995; Keeney and Raiffa, 1976).

To apply this theory to the multicriteria combination function, the study area should be divided into several zones (or neighbourhoods). Malczewski (2011) suggested two methods to describe a neighbourhood for any geographical phenomena in local multicriteria analyses. First, one can divide the study area into different zones based on administrative districts, land-use zones, economic regions, etc. This is referred to as the non-moving window approach, where the neighbourhood of each object is the zone within which the object falls. Second, a neighbourhood can be determined for each location based on the moving window concept (Fotheringham, Brunsdon, and Charlton, 2000, 2003; Lloyd, 2010). In this approach, a threshold distance or a neighbourhood size is defined around each object and any geographic phenomena that fall within this range constitute the object neighbourhood.

Each zone or neighbourhood has a different value of the criterion range that can be quantified by subtracting the minimum value of a specific criterion from the maximum value of the same criterion in that zone. Instead of evaluating locations with respect to the whole study area, each object is assessed within the neighbourhood it belongs to. Therefore, the value function and criterion importance weights need to be modified for each location based on neighbourhood characteristics. The value function in Equation 4.6 should be modified as below:

$$a_{lz}^{q} = \vartheta(c_{lz}^{q}) = \begin{cases} \dfrac{c_{lz}^{q} - \min\limits_{q} c_{z}^{q}}{r_{z}^{q}}, & \text{if higher values are desirable (benefit attributes)} \\ \dfrac{\max\limits_{q} c_{z}^{q} - c_{lz}^{q}}{r_{z}^{q}}, & \text{if lower values are desirable (cost attributes)} \end{cases} \tag{4.15}$$

where $\vartheta(c_{lz}^{q})$ shows the local value function for neighbourhood q; c_{lz}^{q} represents the raw score of alternative l, which is located in neighbourhood q, in the z-th criterion; $min_q\, c_z^q$ and $max_q\, c_z^q$ are the minimum and maximum values of the z-th criterion in neighbourhood q, respectively. r_z^q is the local range of the z-th criterion in neighbourhood q and is obtained as follows:

$$r_z^q = \max_q c_z^q - \min_q c_z^q \tag{4.16}$$

Moreover, the global importance weight of criterion z must be modified with reference to the local range of the z-th criterion in neighbourhood q (Malczewski, 2011):

$$w_z^q = \frac{w_z \frac{r_z^q}{r_z}}{\sum_{z=1}^{n} w_z \frac{r_z^q}{r_z}} \tag{4.17}$$

where w_z^q is the local weight of the z-th criterion associated with neighbourhood q; r_z is the global range of values for the z-th criterion (see Equation 4.7). By substituting the local standardized criterion values and local importance weights into the global OWA function (Equation 4.14), the local OWA function can be determined as follows (Malczewski and Liu, 2014):

$$LOWA_l(a_{l1}, a_{l2}, \dots, a_{ln}) = \sum_{z=1}^{n} \left(\frac{u_z^q \lambda_z}{\sum_{z=1}^{n} u_z^q \lambda_z} \right) b_{lz}^q \tag{4.18}$$

where $LOWA_l$ is the score of alternative l obtained by the local OWA function; b_{lz}^q is the local standardized criterion values obtained by reordering a_{lz}^q from highest to lowest; λ_z is the order weight associated with z-th highest criterion value; u_z^q is the local weight of the z-th criterion reordered according to the criterion values.

4.4 Developing geosimulation-multicriteria model

From this section onwards, the focus is to design and develop a model for the study area based on the geosimulation-multicriteria analysis. Accordingly, the landscape and different types of agents engaging in the residential growth process need to be recognized first. Involving agents in the residential growth process in Tehran and their behaviours are identified based on the previous studies and the experts' opinions. In the following sections, the discussion is based on the literature. The results of the group discussion will be presented in Chapter 5.

4.4.1 Landscape

The suggested geosimulation-multicriteria model uses the raster data model because of its benefits (see Section 2.4.3.1). In this model, the landscape is represented by a grid of square cells over which actors (individuals or groups of individuals) act and interact. Satellite images or land-use maps can be applied to simulate the landscape (see Section 2.4.3.2); however, satellite images need to be classified before inputting into the model. The number of land-use classes may vary based on the study area and application (see Section 2.4.3.3).

4.4.2 Actors engaging in the process of residential growth

The outcome of the actions and interactions of individuals is a changing pattern of land-use. These individuals can be seen as agents and classified into different groups according to the type of involvement in a specific process (see Section 2.4.3.4). This research adopts the classification of agents proposed by Jokar Arsanjani (2012), and Jokar Arsanjani, Helbich, and de Noronha Vaz (2013). They identified three major groups of agents in the process of urban development in Tehran: households, real estate developers, and local authorities. Each group of agents plays a different role in the residential growth process. An agent may interact with other agents in its group or with agents in other groups to fulfil its objective(s). In a residential development context, all three types of agents interact to make a final decision. Households and real estate developers evaluate each undeveloped parcel of land for future residency based on different objectives. The local authorities control the process by setting out a set of rules and the other two types of agents need to adhere to those rules.

4.4.2.1 Household agents

Households of Tehran represent the fundamental units in simulating household agents' behaviours. Each household in the city can be represented by an agent or any agent can represent a group of households that seems to have more or less the same interests and priorities. In the next sections, the role of households in residential development is described and representation of households in a multi-agent model is discussed.

4.4.2.1.1 Households impact on residential development

Population growth is the major driving force behind residential growth in Tehran (Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013). The amount of land that needs to be developed by a certain time directly depends on households' demands. Different undeveloped parcels of land have different levels of suitability for residential development from households' perspective. Since it is not feasible to simulate the decision behaviours of all households in the city, each agent represents a group of households that is considered to have the same preferences. Therefore, it is required to categorize households of the city into a number of groups based on common preferences.

4.4.2.1.2 Classification of households

According to Li and Liu (2007), two main factors affect the location evaluation of households: household income and structure. These two factors are applied to categorize households in this research. Studies show that different stages in the life cycle coincide with a specific household structure and it has an effect on households' location preferences (Clark and Dieleman, 1996). However, there is no general consensus over the best classification of households based on the life cycle (Clark and Dieleman, 1996). Lansing and Morgan (1955) suggested a linear approach to categorize households based on the basic life cycle, from young and single to older stages. Clark and Dieleman (1996) designed a diagram based on the life cycle to demonstrate changing in location preferences with the household structure (Figure 4.5). They considered the following classes of households: single male/female adult, young couple, family with one child, family with three children, and older couples. In this research, Clark and Dieleman's (1996) classification has been adopted with small modifications. Here, four classes of household are considered based on the structure: young singles, couples without children, couples with children, old couples.

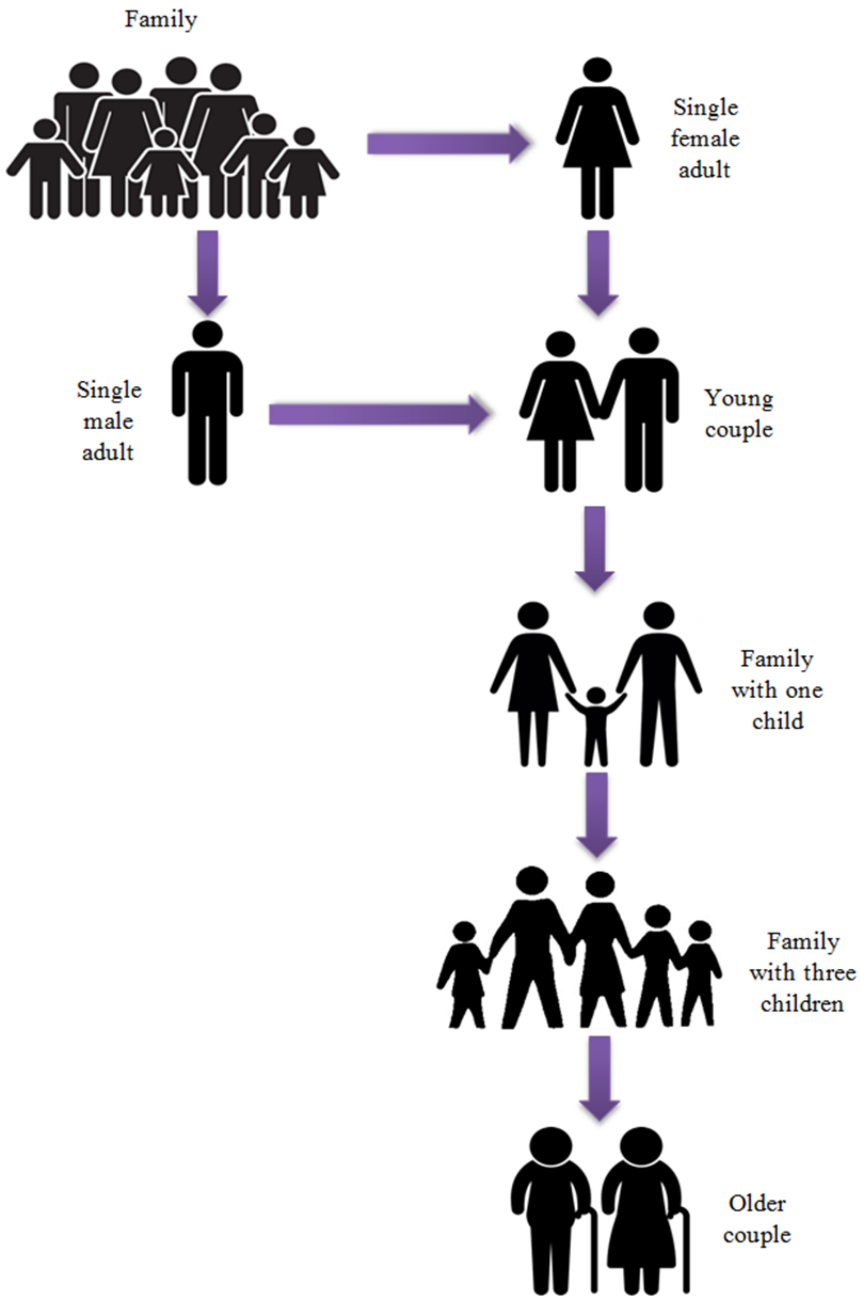


Figure 4.5: Household structure that affects the location preferences

Iranian households are divided into ten groups by the government based on their after-tax family income, which are called income decile groups (Statistical Center of Iran, 2017a). These groups indicate the relative economic situation of a household compared to other households. Households in the first decile have the lowest after-tax income compared to other households and those in the last decile are recognized as the richest households. According to the Statistical

Centre of Iran's definition, the first four deciles are considered as low-income households. The next three deciles constitute middle-income households. The last three deciles form high-income households. Accordingly, we assumed that 40% of Tehran households belong to low-income category, 30% to middle-income category, and 30% to high-income category.

Given household income and structure, twelve classes of households were developed in this research. Each of these twelve groups has different preferences and makes their evaluations based on these preferences. These twelve groups are as follows: Low-income young singles; Medium-income young singles; High-income young singles; Low-income old couples; Medium-income old couples; High-income old couples; Low-income couples with children; Medium-income couples with children; High-income couples with children; Low-income couples without children; Medium-income couples without children; and High-income couples without children. The preferences of these twelve groups are estimated based on the experts' judgments. The procedure and results of the experts' judgments will be presented in Section 5.3.

4.4.2.1.3 Households' decision behaviours

Some parcels of land (cells) are more suitable for residential development from households' point of view. Multicriteria analysis has been applied extensively to quantify the suitability of lands for development in the context of geosimulation modelling (Jokar Arsanjani, 2012; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013; Li et al., 2011; Li and Liu, 2007; Myint and Wang, 2006; Wu, 1998). MCA provides a framework to simulate agents' behaviours by considering social, economic, and physical factors (Li et al., 2011). In the land development context, MCA is used to calculate the suitability (probability) of each undeveloped land parcel for growth through the trade-off of several development factors (Wu, 1998). To calculate the suitability of each parcel of land for development from households' perspective, their objectives and evaluation criteria need to be identified first.

The process of household agents' decision making can be represented by a hierarchy (Saaty, 1980) that includes: goal, objectives, and attributes (see Figure 4.6). The achievement of the overall goal is measured by evaluation criteria; that is, objectives and related attributes (criteria). Several objectives and attributes are considered in this research based on the literature and group discussion (see also Section 5.3.1). The objectives associated with households include:

(i) *Maximizing accessibility*. Accessibility represents the connection between type of land-use and transportation (Waddell, 2000). Access to workplaces and shopping opportunities are amongst the factors which affect households' location preferences (Waddell, 2000). It involves two set of attributes: (i) those that indicate how easy a parcel of land can be reached by major roads, e.g., expressways and highways, and (ii) how easily the residents of that land parcel can access public facilities and services.

(ii) *Maximizing neighbourhood quality.* This objective indicates the configuration of a neighbourhood in terms of environmental and aesthetic conditions. The state of the neighbouring lands (cells) affects the future state of an undeveloped cell (Wu and Webster, 1998).

The objectives are operationalized by underlining quantifiable attribute(s) (see Figure 4.6). In this section, the justification of selected attributes on the basis of previous literature is discussed. Experts' opinions and judgments about contributory attributes will be presented in Section 5.3.1.

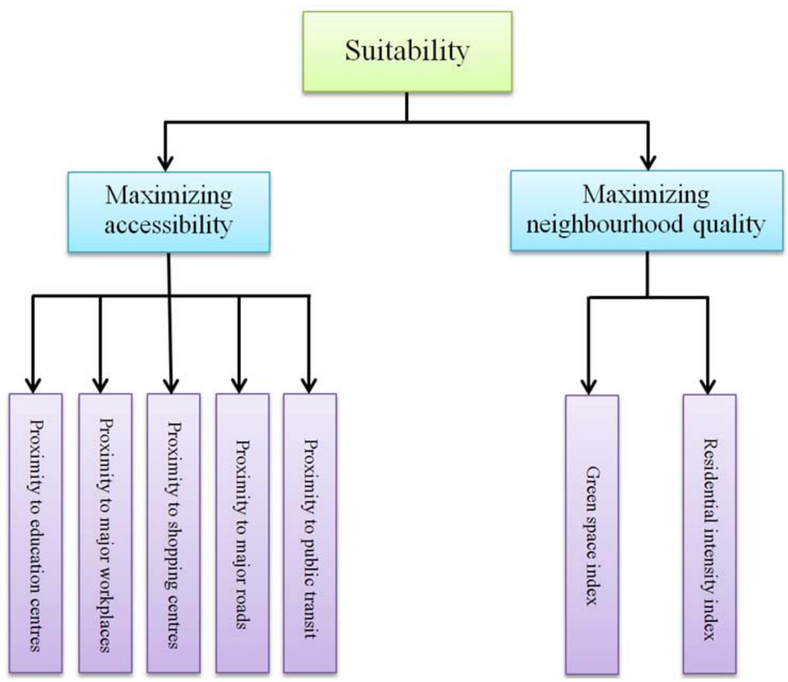


Figure 4.6: Hierarchical structure for household agents

4.4.2.1.3.1 Accessibility

Households perform actions (e.g., shopping and going to school) and there are some places that households visit more frequently, such as shopping centres. Consequently, locations close to shopping centres are more suitable for residency. Accessibility is a measure to indicate how close a location to these centres is and how easily they can be accessed. It is operationalized using five attributes (see Figure 4.6): (i) proximity to education centres or public schools (undeveloped cells surrounding education centres is more suitable for residency from the perspective of families with children), (ii) proximity to the major workplaces (locations close to workplaces are more likely to be developed because of lower commuting costs – see Huu Phe and Wakely, 2000; Waddell, 2000), (iii) proximity to shopping centres (this attribute is one of the most important factors that affect the suitability of a cell to be converted to residential – Hinshaw and Allott, 1972), (iv) proximity to major roads (one of the factors that determine the accessibility of a location is how it can be reached by highways and expressways; studies have shown that there is a link between transportation networks and land-use change – see Frazier and Kockelman, 2005, and Silva and Clarke, 2002. According to Zhou and Kockelman (2008), distance to the nearest highway is one of the factors that determines the type of land-use), and (v) proximity to public transit (locations with better access to transportation services have higher possibility for development – see Hinshaw and Allott, 1972; Zhou and Kockelman (2008) suggest that access to transit stops along with proximity to closest highways are two of the variables associated with the state of transportation services; having easy access to public transport helps households minimize their travel costs. To measure the attributes related to accessibility, network distance was used in this study.

It is worth mentioning that distance to the central business district has also been considered in some studies as an attribute to measure accessibility (e.g., Li and Liu, 2007; Loibl and Toetzer, 2003; Nourqolipour et al., 2016; Wu, 1998). However, studies carried out by Bertaud (2003) revealed that Tehran is a city without a dominant CBD (see Section 3.4). He claimed that distance to the centre of the city is a weak factor for residential development. Therefore, distance to the CBD was not considered in this study as an important factor having an impact on households' evaluation.

4.4.2.1.3.2 Neighbourhood quality

The quality of a neighbourhood is measured using some environmental and physical attributes. These attributes are different in nature compared to accessibility attributes. While accessibility attributes are distance-based, neighbourhood attributes quantify conditions of a neighbourhood based on the amount of desirable features in the vicinity of a location. Two indices were used: green space index, and residential intensity index. The green space index shows the percentage of green spaces in the vicinity of a location. People prefer to live in locations with more green space and water in the surrounding areas (Li and Liu, 2007). Since there is no major water body in Tehran, it has not been considered in the model. The residential intensity index shows how much of the landscape around each cell belongs to the residential type. Undeveloped cells adjacent to already developed cells are more suitable for residency from households' points of view (Hosseinali, Alesheikh, and Nourian, 2013). If a large part of the cells around an undeveloped cell belongs to the residential type, it is more likely the undeveloped cell is chosen for residential development in the near future. In addition, this index controls the compactness of the model output. The map for residential intensity is dynamic; i.e., it needs to be updated after any changes on the landscape.

To quantify the neighbourhood quality of a cell, a moving window was considered around each cell. Although the size and shape of the moving window is an important factor in the process, no theoretical justification has been presented so far to advise which neighbourhood configuration should be adopted in a specific situation (Liu 2008). Li and Liu (2007) employed a 9-by-9 moving window to quantify the environmental quality. However, the spatial resolution in their model is 100m, which means the moving window covers 900m by 900m of the landscape. Since the spatial resolution in the current study is 30m, a 29-by-29 moving window is employed around each cell to measure the green space and residential intensity indices. The green space index is the percentage of the cells in the window that belongs to public parks, to the total number of cells in the window. The residential intensity index is the percentage of the cells in the window that belong to residential areas, to the total number of cells in the window.

4.4.2.2 Real estate developer agents

Real estate developers play a crucial role in developing residential areas in Tehran (Jokar Arsanjani, 2012; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013). They have to both consider households' location preferences and follow the local authorities' policies and regulations. Real estate developers invest money to produce new housings. Those parcels of land that generate more profits are more likely to be developed by real estate developers (Jokar Arsanjani, 2012; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013; Li and Liu, 2007; Tian et al., 2011). From real estate developers' points of view, the investment profit can be estimated as follows (Jokar Arsanjani, 2012; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013; Li and Liu, 2007; Tian et al., 2011):

$$IP_{ij}^{t} = HP_{ij}^{t} - CLA_{ij}^{t} - DC_{ij}^{t} \tag{4.19}$$

where IP_{ij}^{t} is the investment profit of the ij-th cell (parcel of land), HP_{ij}^{t} is the housing price of the ij-th cell, CLA_{ij}^{t} and DC_{ij}^{t} are the cost of land acquisition and development cost of the ij-th cell, respectively. Land parcels that generate higher profit are more suitable to be developed by developer agents. Since there is no reliable data for development cost for the study area, it has not been included as an attribute in the land suitability procedure. Moreover, the panel of experts agreed that housing price and cost of land acquisition are the most important factors that control real estate developers' behaviour (see Section 5.3.1). Therefore, developers use two attributes to quantify the suitability of each land parcel for residential development: housing price and cost of land acquisition. Figure 4.7 shows the hierarchical structure of the real estate developer agent, which consists of the overall goal, one objective and two quantifiable attributes.

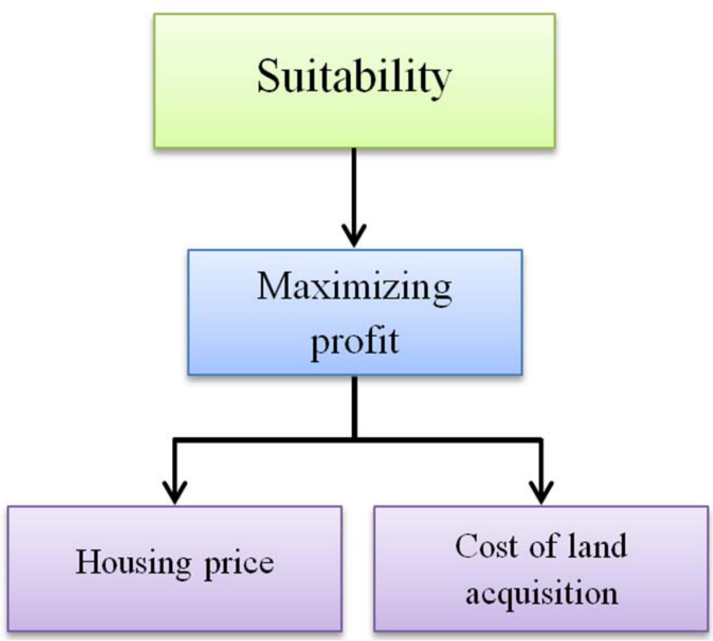


Figure 4.7: Hierarchical structure for real estate developer agents

4.4.2.3 Local authority agent

Any development and land-use change in the study area should be permitted by the local authorities (Jokar Arsanjani, 2012; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013). Local authorities' rules and policies impose constraints on some areas. Although it is not clear what procedures are exactly adopted to grant permission to developers, some general rules that are applicable in the region are considered based on the experts' opinions (see Section 5.3.1) and previous studies (Jokar Arsanjani, 2012; Jokar Arsanjani, Helbich, and de Noronha Vaz, 2013): (i) conservation land-use policy should be taken into account; that is, development in some land-use types, such as public parks, is not permitted. Since the focus of this research is to simulate the conversion of undeveloped lands to residential areas, it is also assumed that constructing new buildings within any developed land parcel is not permitted whether it is residential or other land-use types, (ii) development in areas characterized by steep slopes is not allowed (in this research any slope greater than 10 degrees is considered steep), (iii) development near military zones is not permitted (within a 100-meter buffer around military zones), and (iv) development near airports is not permitted (within a 100-meter buffer around airports). Figure 4.8 shows the hierarchical structure of the local authority agent, which consists of the overall goal, an objective and a set of four constraints.

The condition of parcels of land for development from local authorities' points of view is determined in a raster format. Parcels of land are assigned a binary value: 0 or 1. Development in parcels that are assigned by 0 is not allowed as far as the local authority agent is concerned (restrictive areas). Since the local authority agent is the final decision-maker, those parcels will not be developed in the model even if they are highly suitable for the residential area. Indeed, household and developer agents evaluate land parcels based on some residential growth factors, while local authorities impose some constraints to the growth areas.

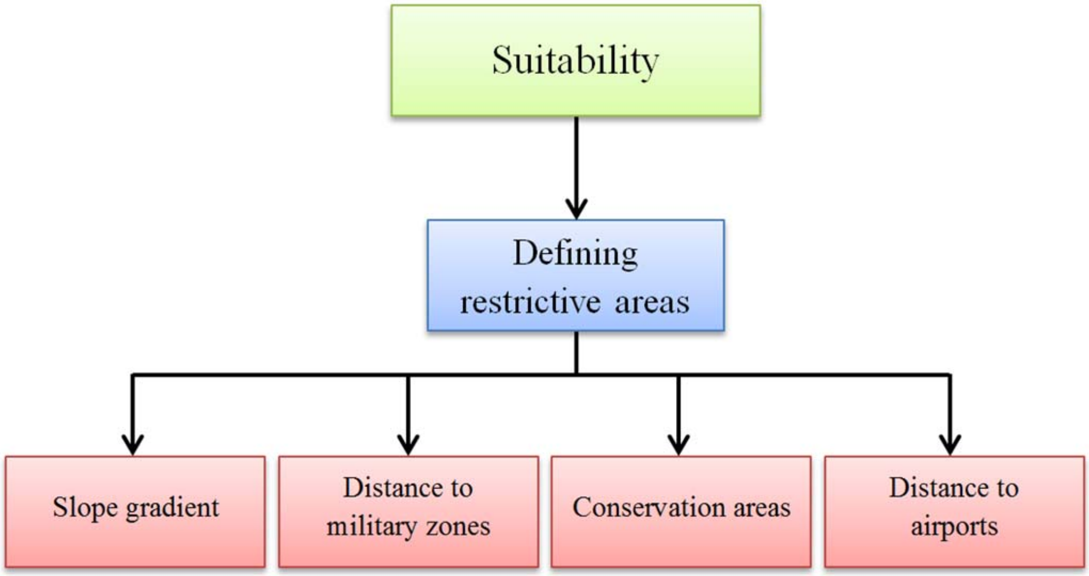


Figure 4.8: Hierarchical structure for local authority agents

4.4.3 Decision making process

In order to make the final decision, preferences of all three types of agents must be considered. While household and developer agents generate a land suitability pattern that indicates the suitability of each cell for residential development from their perspective, local authorities approve or reject the request for development. The final land suitability scores should be calculated by combining the preferences of all involving actors. The combination process embodies the interactions that exist among different interest groups. In this research, MADM approach is used like in many other geosimulation studies (see Section 2.4.2).

The combination process of preferences is performed using the group hierarchical decision method (Saaty, 1980; Dyer and Forman, 1992). Therefore, three sub-hierarchical structures are developed associated with three engaging actors: households, real estate developers, and local authorities. To represent the decision problem, household agents develop a sub-hierarchy that consists of two objectives and seven associated attributes (see Section 4.4.2.1.3 and Figure 4.6). The sub-hierarchy for the real estate developer agent is made up of one objective and two attributes (see Section 4.4.2.2 and Figure 4.7). The sub-hierarchy for the local authority agent includes one objective and four constraints (see Section 4.4.2.3 and Figure 4.8). The hierarchy structure of the group for the land suitability analysis is shown in Figure 4.9, which consists of the overall goal, objectives and attributes/constraints. The goal is finding the most suitable cells for residential growth. There are three quantifiable objectives and one restrictive objective. There are nine attributes and four constraints. The preferences of households and developers are reflected in the importance attribute (or criteria) weights that they assign to each attribute. According to the principle of hierarchical structure, the sum of attribute weights must be equal to 1 (Saaty, 1980).

There are nine criterion maps and four constraint maps. Criterion maps are GIS layers that represent the quantifiable attributes associated with each objective (Malczewski, 1999). In a criterion map, each cell is assigned a single number (criterion evaluation score). All criterion maps must have same spatial resolution as the base map (see Section 4.4.1) and cover the same area.

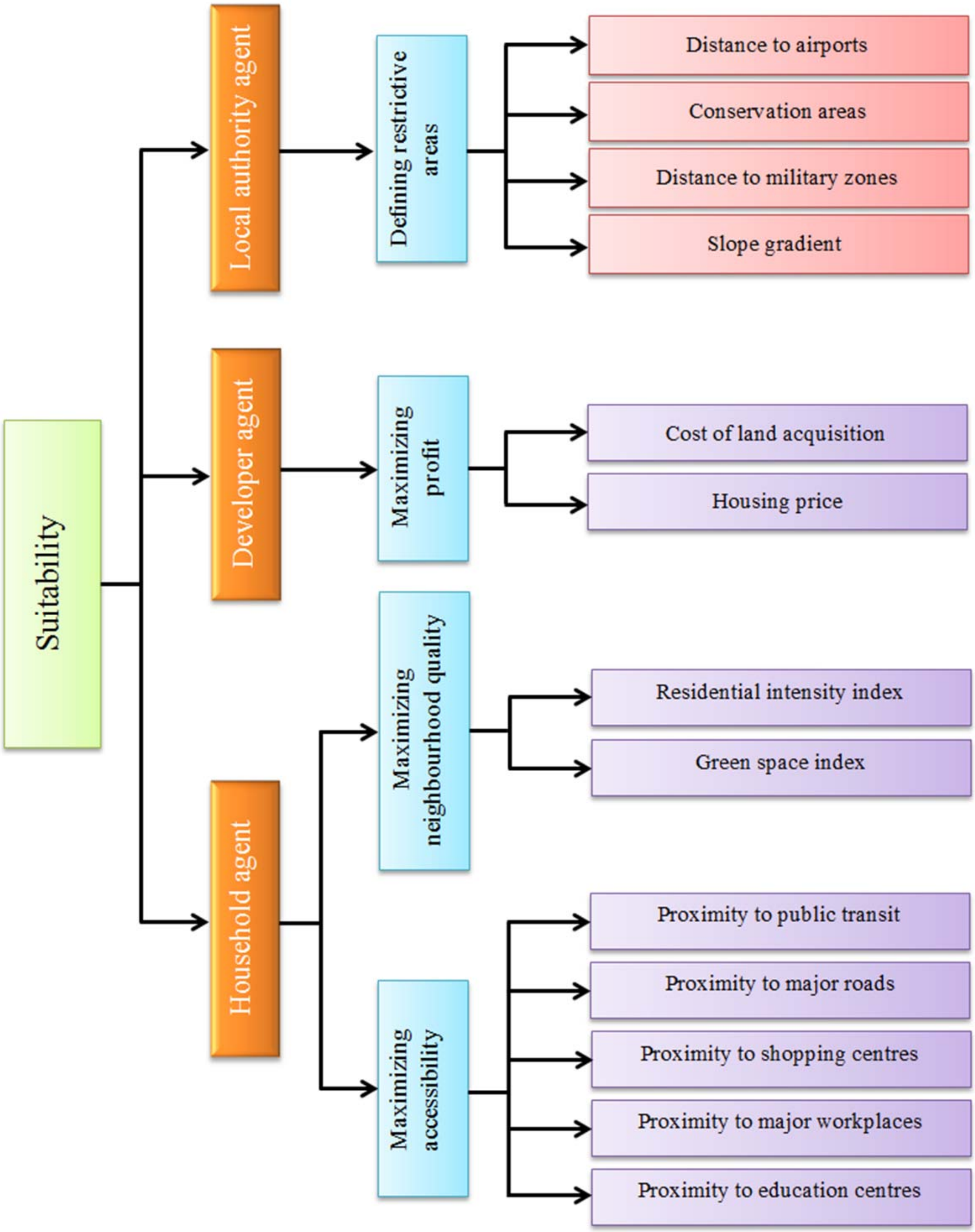


Figure 4.9: Hierarchical structure of group decision making (see Figures 4.6, 4.7, and 4.8)

Formally, the overall suitability of a cell at a specific time is a function of households, developers, and local authorities' suitability analysis. This relationship can be formulated as follows:

$$E_{ij}^{t} = f\left(SH_{ijk}^{t}, SD_{ij}^{t}, LAR_{ij}^{t}\right) \tag{4.20}$$

where E_{ij}^{t} is the overall suitability of the *ij*-th cell at time *t*; SH_{ijk}^{t} is the suitability score of the *ij*-th cell to be converted to the residential area by household agent type *k*; SD_{ij}^{t} indicates the suitability of the *ij*-th cell for residential development with respect to real estate developers' preferences. LAR_{ij}^{t} indicates the suitability of the same cell for development according to local authorities' rules. Since LAR_{ij}^{t} contains a set of constraints that take a binary value of 0 or 1, Equation 4.20 can be rewritten as follows:

$$E_{ij}^{t} = f\left(SH_{ijk}^{t}, SD_{ij}^{t}\right) \prod LAR_{ij}^{t} \tag{4.21}$$

since SH_{ijk}^{t} is a function of seven corresponding attributes, it can be formulated as follows:

$$SH_{ijk}^{t} = f_1\left(EC_{ijk}^{t}, MW_{ijk}^{t}, SC_{ijk}^{t}, MR_{ijk}^{t}, PT_{ijk}^{t}, GS_{ijk}^{t}, RI_{ijk}^{t}\right) \tag{4.22}$$

where EC_{ijk}^{t} is proximity to education centres for the *ij*-th cell at time *t*, MW_{ijk}^{t} is proximity to major workplaces, SC_{ijk}^{t} is proximity to shopping centres, MR_{ijk}^{t} is proximity to major roads, PT_{ijk}^{t} is proximity to public transit, GS_{ijk}^{t} is the green space index, and RI_{ijk}^{t} is the residential intensity index.

Also, SD_{ij}^{t} can be defined as:

$$SD_{ij}^{t} = f_2\left(HP_{ij}^{t},\ CL_{ij}^{t}\right) \tag{4.23}$$

where HP_{ij}^{t} and CL_{ij}^{t} are housing price and cost of land acquisition for the *ij*-th cell at time *t*, respectively.

Given Equations 4.21 to 4.23, Equation 4.20 can be expanded as:

$$E_{ij}^{t} = f\left(EC_{ijk}^{t}, MW_{ijk}^{t}, SC_{ijk}^{t}, MR_{ijk}^{t}, PT_{ijk}^{t}, GS_{ijk}^{t}, RI_{ijk}^{t}, HP_{ij}^{t}, CL_{ij}^{t}\right) \prod LAR_{ij}^{t} \tag{4.24}$$

where f can be any combination function. A number of combination functions have been used in land-use/cover change models so far (see Section 2.4.2). In this study the combination function is the local OWA function (see Section 4.3.1.1.2.3). Therefore, using Equation 4.18, the overall suitability of each cell is obtained by the following formula:

$$E_{ij}^{t} = LOWA_{ij}^{t} = \left[\sum_{z=1}^{9}\left(\frac{u_{kz}^{tq}\lambda_{z}}{\sum_{z=1}^{9} u_{kz}^{tq}\lambda_{z}}\right) b_{ijz}^{tq}\right]\prod LAR_{ij}^{t} \tag{4.25}$$

where z indicates a criterion; b_{ijz}^{tq} represents the evaluation score of the ij-th cell at time t in the z-th criterion and q-th neighbourhood obtained after reordering the scores of cell ij among all nine criteria; u_{kz}^{tq} is the local weight of the z-th criterion in neighbourhood q with respect to the household type k; and λ_z is the order weight associated with z-th highest criterion value. The higher the overall suitability of a cell is, the more desirable it would be for residential development.

4.4.4 Geosimulation-multicriteria workflow

Figure 4.10 shows the workflow of the framework for simulating the pattern of residential growth. The framework aims at identifying the most desirable cells to be converted to residential-type by considering the preferences of the actors involved in the process. The procedure begins with collecting required data at time t and ends with producing a map for land-use pattern at time t+1 (see Sections 4.4.1 to 4.4.3). The framework consists of the following elements:

(i) A set of spatial and non-spatial data must be collected for the study area. Spatial data, including: satellite images for the study area at different time steps, land-use maps and aerial photos; physical layers, including: road network, education centres, shopping centres, subway stations, digital elevation models (DEM), cost of land acquisition, and housing price; and socio-economic data, including: households' structure and income. Non-spatial data comprises: contributory attributes to residential growth in the region (including criteria and constraints), households' preferences about the evaluation criteria, and real estate developers' preferences about evaluation criteria. Non-spatial data are obtained by the focus group approach (see Chapter 5).

(ii) Satellite images for time *t* are processed and classified to provide the base map for the geosimulation-multicriteria model (see Section 4.4.1).

(iii) Geosimulation techniques are applied to generate a set of agents (interest groups) that represent households, real estate developers, and local authorities (see Section 4.4.2); the developers and local authorities are represented by one agent; the total number of household agents in the model is equal to the total number of cells that are needed to be developed to meet the residential demand; there are twelve classes of household agents (see Section 4.4.2.1.2); the number of household agents in each class is proportional to the number of households belonging to that class in reality.

(iv) Multicriteria analysis is incorporated into geosimulation to mimic agents' behaviour. In this study, agents' behaviour is related to their ability to evaluate the suitability of each cell for residential development (see Section 4.4.2). Households start assessing the suitability of each cell for development using a set of evaluation criteria (see Section 4.4.2.1.3); thus, each household agent generates seven criterion maps (see Figure 4.6). Simultaneously, all cells are evaluated by the real estate developer agent. Two criterion maps are generated that represent the evaluation of the developer agent (see Section 4.4.2.2 and Figure 4.7). At the same time, the local authority agent provides a set of regulations and policies for new development (see Section 4.4.2.3 and Figure 4.8). These regulations and policies divide the study area into two categories: feasible lands (construction is permitted) and infeasible lands (construction is restricted).

(v) The combination procedure is applied to aggregate the suitability evaluation of all involving interest groups. The suitability evaluations of households and the real estate developer and the constraints produced by the local authority agent are combined to take the preferences and opinions of all interest groups into consideration (see Section 4.4.3 and Figure 4.9). The output of the combination process is an overall suitability map. By the end of this stage, all infeasible cells for residential development take value of 0 and all feasible cells are assigned a suitability value.

(vi) From this step onwards, geosimulation capabilities are used. In this step, each household agent sorts all feasible cells in descending order based on their overall suitability score obtained in step v; that is, the most suitable cell for development tops the list and least suitable cell is at the bottom.

(vii) For each of the twelve classes of households (see Section 4.4.2.1.2), a feasible cell with the highest overall suitability is selected as the most desirable parcel of land for residential area.

(viii) If a cell is selected by two or more types of agents at the same time, then there is a conflict of interests between household agents. The conflict of interests takes place because there is a competition between different types of agents. In this case, the cell will be assigned to the agent with higher loss index. The loss index is the difference between the suitability value of a cell and the next suitability value in the ordered list. Therefore, those household agents who are not able to select the most suitable cell are forced to select the next most suitable cell in the list.

(ix) The selected cells are converted to residential land-use type.

(x) Those cells that are developed in each model run are removed from the set of potential alternatives by the local authority agent (updating restrictive map).

(xi) Finally, it is checked to determine whether the demand for residential areas has been met; if yes, then the model execution stops and an output map presenting the pattern of land-use in the final time step is created; if no, $t \rightarrow t+1$ and the program execution jumps to step iv.

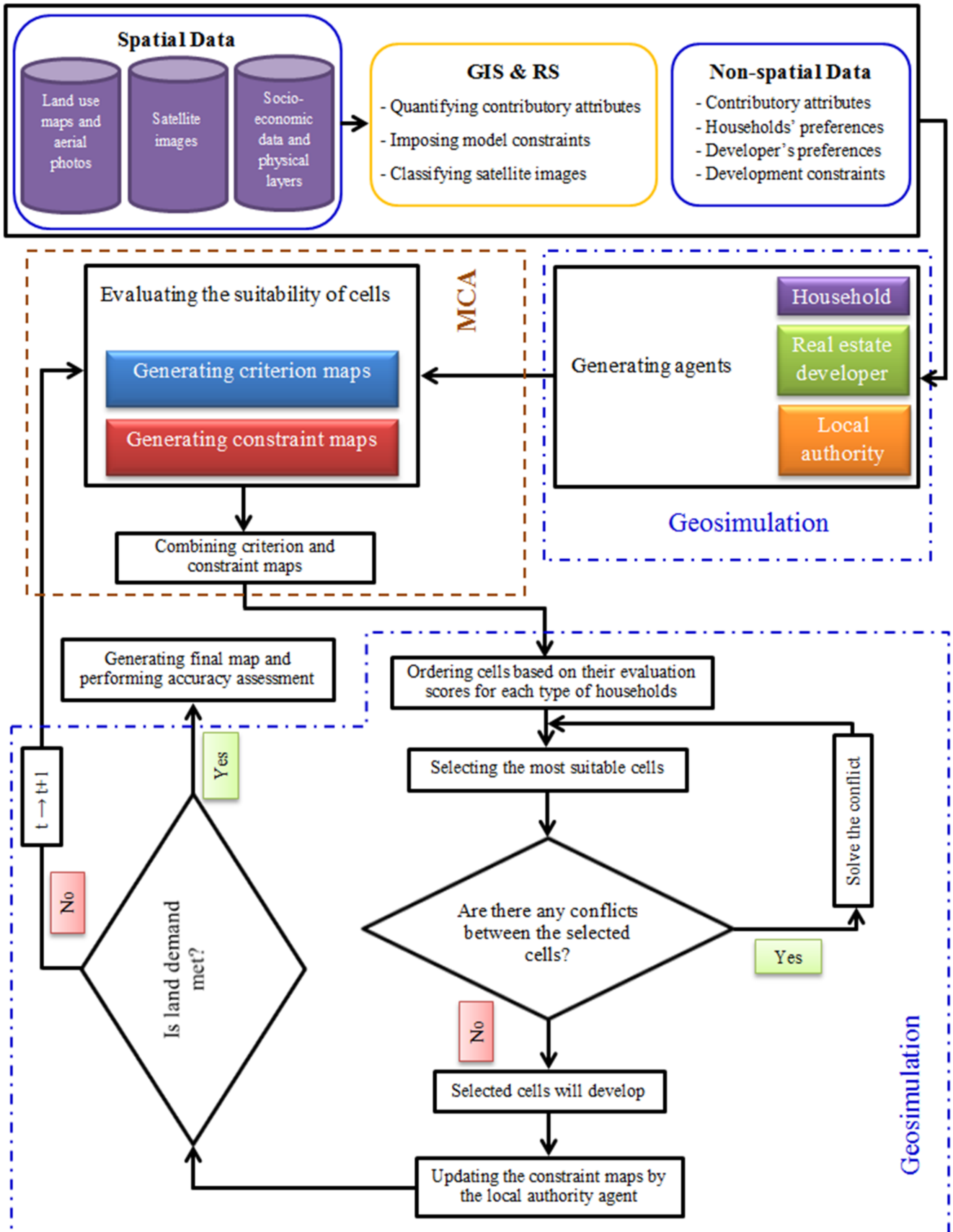


Figure 4.10: Workflow of the simulation process

Some raster-based software packages can be applied to implement the geosimulation-multicriteria model (see Section 2.4.4). ESRI ArcMap 10.3 was used as the platform for implementing the model in this study. ESRI ArcGIS has a set of software components that offers tools to developers to access ArcGIS functionalities for implementing their models, which is called ArcObjects (Burke, 2003). A number of programming languages (e.g., VB.NET, Java, C#, and Python) can be used to access ArcObjects libraries. Python programming language was selected in this research to implement the suggested framework because of its compatibility with ArcMap 10.3 and abundance of resources (e.g., Pimpler, 2015; Zandbergen, 2013). Python is a free and open-source programming language. PyScripter was chosen as the software development environment (SDE) to do Python scripting.

4.4.5 Illustrative example

A hypothetical situation is used to illustrate the framework procedure. Specifically, the intention here is to find the most suitable cells for residential development using the geosimulation-multicriteria model. The first step is to prepare the input data (see Section 4.4.4, and Figure 4.10). It is assumed that the study area (urban landscape) is represented by a 6-by-6 grid of square cells. The cell size is 1 km by 1 km and each cell is described by the row number (i) and column number (j). A single cell is denoted by δ. For example, δ_{11} is in the top-left corner and δ_{66} is in the right-bottom corner. There are five types of land-uses on the landscape: public park, residential, open land, farmland/orchard, and non-residential. Figure 4.11 shows the land-use pattern (or landscape).

Figure 4.11: Hypothetical land-use pattern

The goal is to identify the most suitable cell(s) for residential development from households' and real estate developers' points of view according to local authorities' regulations regarding feasible lands (see Figure 4.9). To simplify the situation, it is assumed that two types of households are involved in the procedure: high-income single and low-income couple with children. First, the demand for residential area needs to be determined. Suppose that two cells need to be converted to the residential type to meet the demand for residential areas. All involving agents start evaluating the suitability of each cell for residential development simultaneously (see Figure 4.10). It is assumed that the households have two objectives: maximizing accessibility and maximizing neighbourhood quality. It is also assumed that the first objective is quantified using two criteria: proximity to shopping centres (in km) and proximity to education centres (in km), and the second objective is operationalized using the residential intensity index (in %). Simultaneously, the real estate developer agent evaluates cells to identify the most suitable one(s) for residential development. From the real estate agent perspective the most suitable cell is the one that brings in more profit (see Section 4.4.2.2). Therefore, the real estate developer has one objective (maximizing profit). It is assumed that this objective is quantified using one attribute, i.e., housing price.

By assuming that δ_{44} is a shopping centre and δ_{62} is an education centre (see Figure 4.11), distance to shopping centres and education centres for each cell can be calculated. The results are shown in Figures 4.12a and 4.12b. For the sake of simplicity, the Euclidean metric is used as distance definition. A 3-by-3 moving window is applied to measure the residential intensity in the vicinity of each cell (see Section 4.4.2.1.3.2). Figure 4.12c shows the residential intensity index for each cell. Also, Figure 4.13 shows the spatial pattern of housing price.

(a)

4.2	3.6	3.2	3	3.2	3.6
3.6	2.8	2.2	2	2.2	2.8
3.2	2.2	1.4	1	1.4	2.2
3	2	1	0	1	2
3.2	2.2	1.4	1	1.4	2.2
3.6	2.8	2.2	2	2.2	2.8

(b)

5.1	5	5.1	5.4	5.8	6.4
4.1	4	4.1	4.5	5	5.6
3.2	3	3.2	3.6	4.2	5
2.2	2	2.2	2.8	3.6	4.5
1.4	1	1.4	2.2	3.2	4.1
1	0	1	2	3	4

(c)

50	33	67	50	50	25
33	22	56	56	56	33
50	33	44	44	44	33
50	33	33	33	33	17
50	33	33	22	22	0
25	17	33	33	33	0

Figure 4.12: Maps of (a) Euclidean distance to shopping centres (in km); (b) Euclidean distance to education centres (in km); and (c) residential intensity index (in %)

200	180	183	160	158	142
173	155	180	150	147	136
158	159	148	140	130	125
164	168	175	164	128	111
119	145	138	130	110	103
106	100	98	88	90	87

Figure 4.13: Housing price (in 1000 $)

The local authority agent produces a map that shows the restrictive areas for residential development according to a set of rules and policies (see Section 4.4.2.3). Let us assume that six cells are feasible to be converted to residential areas, three of them are open lands and the other three are farmlands/orchards (Figure 4.14).

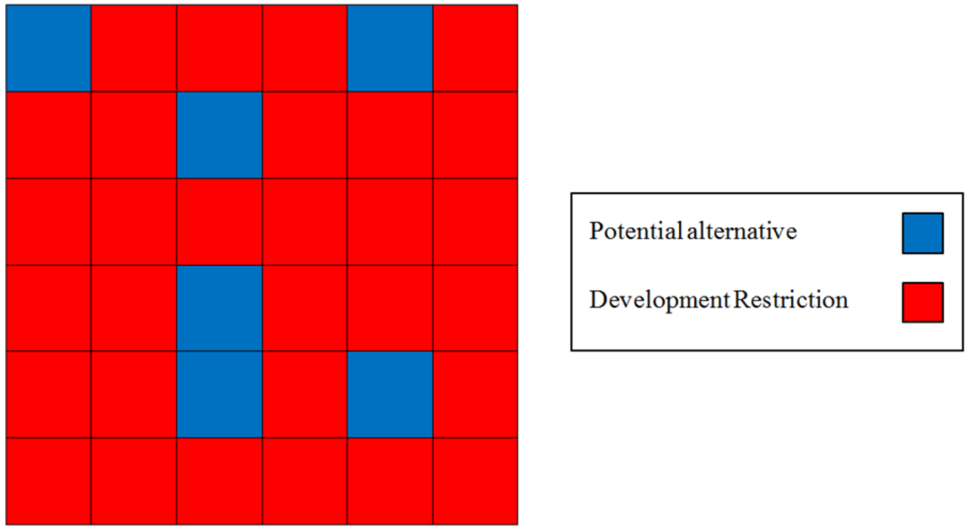


Figure 4.14: Feasible and infeasible cells according to local authorities' rules

Having had all three involved agents evaluate all cells for residential development, now the suitability of each cell needs to be calculated by combining all preferences and constraints. The combination process is carried out as explained in Section 4.4.3 using a local OWA function. The four evaluation maps associated with household and developer agents and the constraint map associated with local authority must be combined. However, the four evaluation maps have different scales and are not comparable. Therefore, the standardization procedure using local value function needs to be performed to map the evaluation values in [0, 1] interval (see Equation 4.15). To apply a local value function, a neighbourhood must be defined around each cell or the landscape must be divided into several zones (see Section 4.3.1.1.2.3). In this example, a 3-by-3 moving window is used as the neighbourhood of each cell. Also, it is assumed that a linear function can be applied to standardize criterion values. Two distance-based criteria are cost attributes, i.e., lower values are desirable, while residential intensity and housing price are benefit attributes, i.e., higher values are desirable. Therefore, according to Equation 4.15, the formula for cost attributes is used for distance-based attributes and the formula for benefit attributes is employed for residential intensity and housing price. Figure 4.15 illustrates the neighbourhood around δ_{23} for the map of distance to shopping centres (see also Figure 4.12a). The minimum and maximum values in the neighbourhood are 1 and 3.6, respectively. The standardized value for δ_{23} in distance to shopping centres is calculated as $\frac{(3.6-2.2)}{(3.6-1)} = 0.5$ (see Equation 4.15).

4.2	3.6	3.2	3	3.2	3.6
3.6	2.8	2.2	2	2.2	2.8
3.2	2.2	1.4	1	1.4	2.2
3	2	1	0	1	2
3.2	2.2	1.4	1	1.4	2.2
3.6	2.8	2.2	2	2.2	2.8

Target Cell

Neighborhood

Figure 4.15: Map of distance to the shopping centre: Neighbourhood for the target cell δ_{23}

The same procedure is applied to other cells in the map for distance to shopping centres and other criterion maps, i.e., distance to education centres, residential intensity index, and housing price. Figure 4.16 demonstrates the standardized criterion maps associated with household agents. The standardized criterion map of housing price is shown in Figure 4.17. Since lower distance to shopping and education centres is desirable, the maps associated with these two criteria are called proximity to shopping and education centres after standardization.

(a)

0	.3	.25	.17	.25	0
.3	.5	.54	.55	.54	.36
.25	.54	.5	.55	.5	.33
.17	.55	.55	1	.55	.17
.25	.54	.5	.55	.5	.33
0	.36	.33	.17	.33	0

(b)

0	.09	.21	.24	.32	0
.48	.52	.53	.5	.5	.36
.43	.52	.52	.5	.5	.3
.45	.55	.54	.5	.5	.28
.36	.55	.5	.54	.52	.27
.29	1	.55	.55	.52	.09

(c)

1	.24	1	0	.81	0
.39	0	.76	.52	1	.26
1	.32	.65	.48	.69	.41
1	0	.5	.5	.75	.39
1	.48	1	0	.67	0
.24	0	1	1	1	0

Figure 4.16: Standardized criterion maps: (a) proximity to shopping centres; (b) proximity to education centres; (c) residential intensity index

1	.56	1	.36	.92	.27
.4	.13	.93	.38	.63	.33
.17	.34	.2	.23	.36	.39
.92	.88	1	.83	.41	.3
.28	.61	.57	.48	.3	.39
.13	.04	.18	0	.07	0

Figure 4.17: Standardized criterion map for housing price

The evaluation criteria have different degrees of relative importance based on the agents' preferences. The importance of those criteria that are related to the household agents depends on the household income and structure (see Section 4.4.2.1.2). However, the importance weight of housing price is independent of type of households. Let us assume that Table 4.2 shows the hypothetical global criterion weights for both household and developer agents.

Table 4.2: Hypothetical global weights of attributes

	Household agent			Developer agent
Attribute / Household type	Proximity to shopping centres	Proximity to education centres	Residential intensity index	Housing price
High-income single	0.21	0.05	0.35	0.39
Low-income couple with children	0.17	0.32	0.12	0.39

Figure 4.18 shows the hierarchical structure of the land suitability problem (it is a modified hierarchical structure of Figure 4.9). At the top of the hierarchy is the goal (or main objective): finding the most suitable cell(s) for residential development. There are four objectives, i.e., maximizing accessibility, maximizing neighbourhood quality, maximizing profit, and defining restrictive areas. At the bottom of the hierarchy there are five elements (maps): four criterion maps and a constraint map. These five maps and associated criterion weights must be combined to produce the final suitability map.

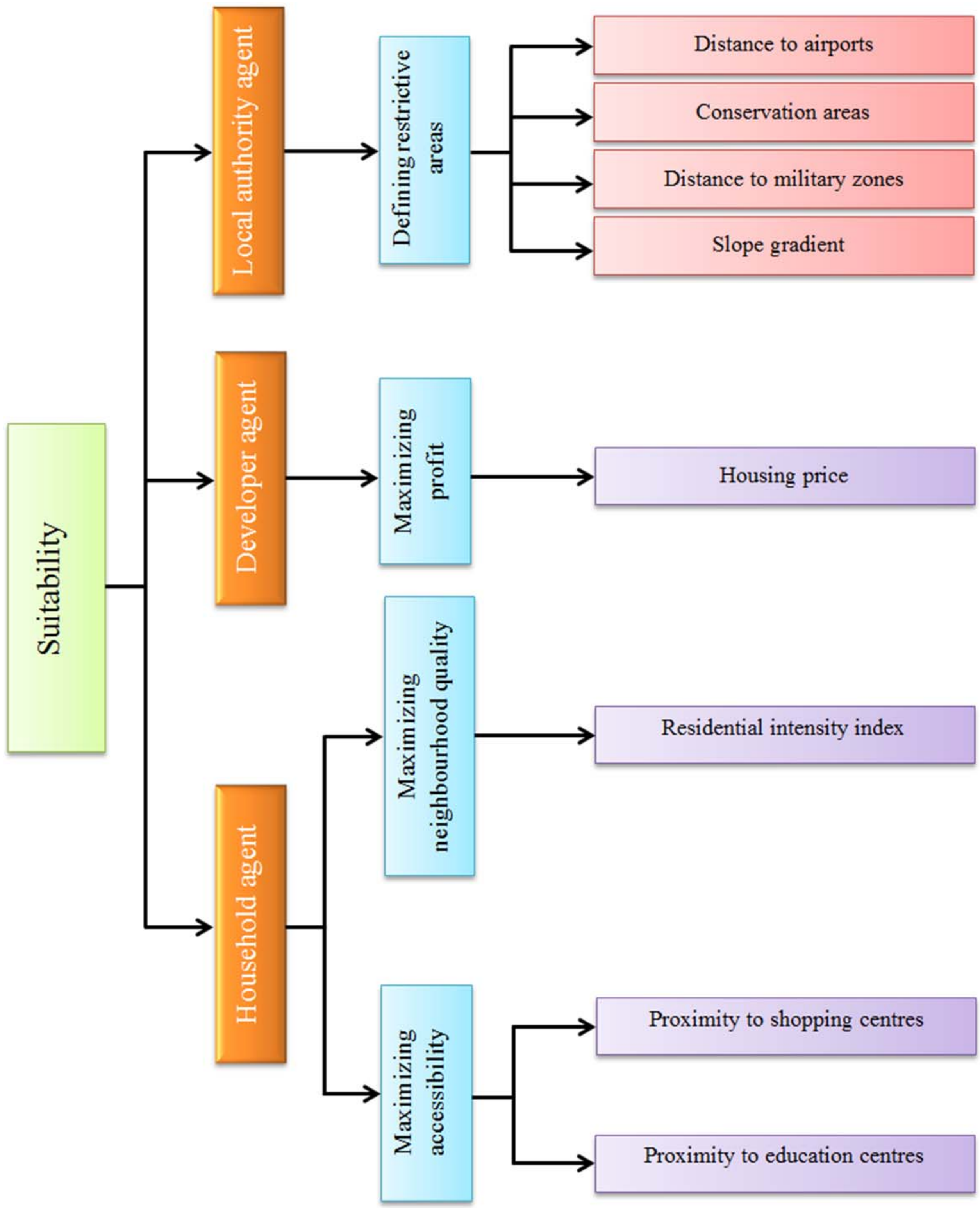


Figure 4.18: Hierarchical structure of the hypothetical land suitability problem

The method for generating the final suitability map is illustrated by the computational procedure for one cell. It is assumed that the high-income single agent, real estate developer agent, and local authority agent are evaluating the suitability of δ_{23} for residential development. The criterion maps generated by the household and developer agents and the constraint map are combined by considering criterion weights. Table 4.2 shows the global criterion weights, which are used for generating local weights for each cell based on the neighbourhood structure (see Section 4.3.1.1.2.3). A 3-by-3 moving window is applied to determine the neighbourhood of each cell (see Figure 4.15). Accordingly, the criterion weights for δ_{23} are defined as follows (see Equation 4.17):

$$w_{z_1}^{q_{23}} = \frac{w_{z_1}\frac{r_{z_1}^{q_{23}}}{r_{z_1}}}{\sum_{n=1}^{4} w_{z_n}\frac{r_{z_n}^{q_{23}}}{r_{z_n}}} \tag{4.26}$$

$$s.t. \quad r_{z_1}^{q_{23}} = \max_{q_{23}} c_{z_1}^{q_{23}} - \min_{q_{23}} c_{z_1}^{q_{23}}\ , \quad r_{z_n}^{q_{23}} = \max_{q_{23}} c_{z_n}^{q_{23}} - \min_{q_{23}} c_{z_n}^{q_{23}}$$

$$s.t. \quad r_{z_1} = max\, c_{z_1} - min\, c_{z_1}\ , \quad r_{z_n} = max\, c_{z_n} - min\, c_{z_n}$$

where $w_{z_1}^{q_{23}}$ is the local weight of the z_1-th criterion for δ_{23}; z_1 is a criterion and q_{23} indicates the neighbourhood of δ_{23}. n represents the total number of criterion maps, which is four here. w_{z_1} is the global importance weight of criterion z_1; $r_{z_1}^{q_{23}}$ is the local range of criterion z_1 values in neighbourhood q_{23}. r_{z_1} is the global range of criterion z_1 values. $min_{q_{23}}\, c_{z_1}^{q_{23}}$ and $max_{q_{23}}\, c_{z_1}^{q_{23}}$ are the minimum and maximum values of criterion z_1 in neighbourhood q_{23}. $min\, c_{z_1}$ and $max\, c_{z_1}$ are the global minimum and maximum values of criterion z_1, respectively; and $min_{q_{23}}\, c_{z_n}^{q_{23}}$ and $max_{q_{23}}\, c_{z_n}^{q_{23}}$ are the minimum and maximum values of criterion z_n in neighbourhood q_{23}, respectively.

According to Table 4.2, the global weight of proximity to shopping centres for the high-income single agent is 0.21. If it is assumed that z_1 is proximity to shopping centres, the local importance weight of proximity to shopping centres for δ_{23} would be (see Figures 4.12 and 4.13):

$$w_{z_1}^{q_{23}} = \frac{.21 * \frac{(3.6-1)}{(4.2-0)}}{\left[.21 * \frac{(3.6-1)}{(4.2-0)} + \left(.05 * \frac{(5.4-3)}{(6.4-0)}\right) + \left(.35 * \frac{(67-22)}{(67-0)}\right) + \left(.39 * \frac{(183-140)}{(200-87)}\right)\right]} = .24 \quad (4.27)$$

In the same way, the importance weight of proximity to education centres, the residential intensity index, and housing price for δ_{23} would respectively be: 0.04, 0.44, and 0.28.

Next, the combination process is performed using the local OWA function (see Equation 4.25). The overall suitability of δ_{23} is obtained as follows:

$$E_{\delta_{23}} = LOWA_{\delta_{23}} = \left[\sum_{n=1}^{4} \left(\frac{u_{z_n}^{q_{23}} \lambda_{z_n}}{\sum_{n=1}^{4} u_{z_n}^{q_{23}} \lambda_{z_n}}\right) b_{\delta_{23} z_n}^{q_{23}}\right] \prod LAR_{\delta_{23}} \quad (4.28)$$

where $E_{\delta_{23}}$ is the overall suitability of δ_{23} for residential development. $u_{z_n}^{q_{23}}$ is the local importance weight of criterion z_n after reordering. λ_{z_n} is the order weight associated with criterion z_n. $b_{\delta_{23} z_n}^{q_{23}}$ is the standardized value of δ_{23} in the z_n-th criterion according to its neighbourhood q_{23} after reordering. $LAR_{\delta_{23}}$ is the suitability of δ_{23} for residential development based on the local authority rules (feasible or infeasible).

Figure 4.19 highlights the values of δ_{23} among all criterion maps and the constraint map (see Figures 4.14, 4.16, and 4.17). The criterion values of δ_{23} in order are:

$$b_{\delta_{23} z_1}^{q_{23}} = .93, \quad b_{\delta_{23} z_2}^{q_{23}} = .76, \quad b_{\delta_{23} z_3}^{q_{23}} = .54, \quad b_{\delta_{23} z_4}^{q_{23}} = .53 \quad (4.29)$$

where $b_{\delta_{23} z_1}^{q_{23}}$ is the highest standardized criterion value for δ_{23} and the value of proximity to shopping centres ($b_{\delta_{23} z_3}^{q_{23}}$) is the third highest value and its associated weight ($u_{z_3}^{q_{23}}$) is 0.24. By assuming that all criteria need to be used equally in the suitability evaluation process, all order weights, λ_{z_n}, would be equal to $\frac{1}{4}$. Moreover, according to Figure 4.14, the constraint value of δ_{23} is 1. Consequently, Equation 4.28 can be calculated as:

$$E_{\delta_{23}} = \left[\frac{.28 * .25}{\Sigma} * .93 + \frac{.44 * .25}{\Sigma} * .76 + \frac{.24 * .25}{\Sigma} * .54 + \frac{.04 * .25}{\Sigma} * .53\right] \prod LAR_{\delta_{23}} = .75 \quad (4.30)$$

$$\sum_{n=1}^{4} u_{z_n}^{q_{23}} \lambda_{z_n} = (.28 * .25) + (.44 * .25) + (.24 * .25) + (.04 * .25) = .25 \quad (4.31)$$

Thus, the overall suitability of δ_{23} is equal to 0.75.

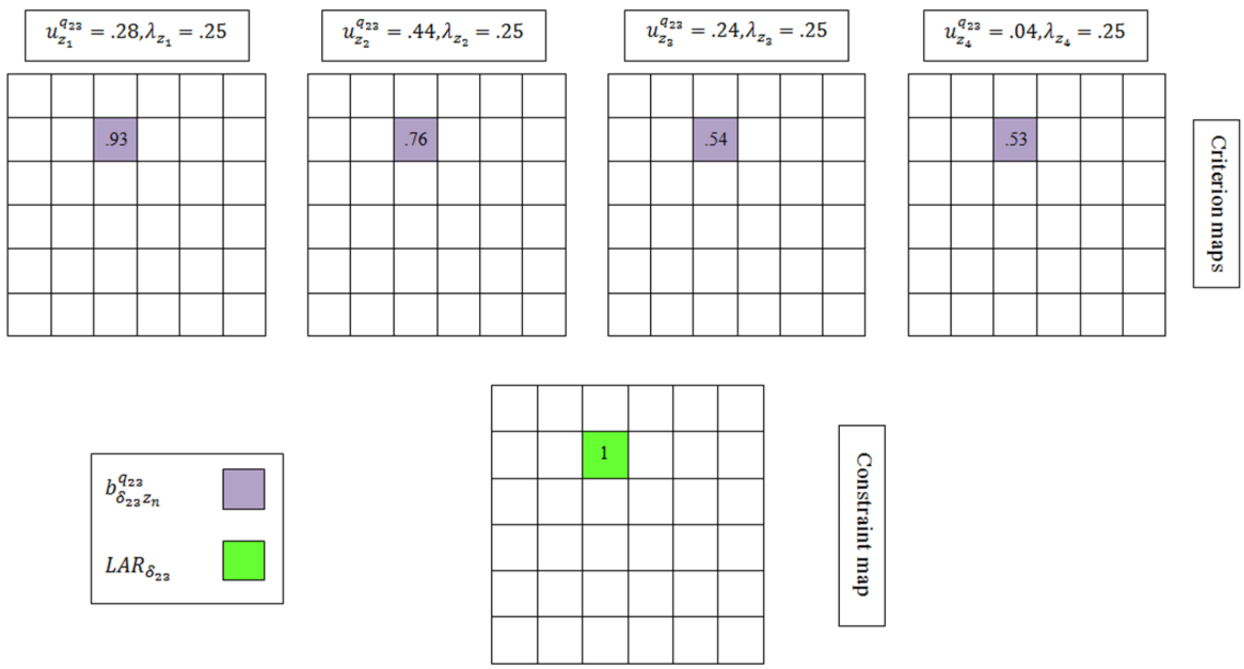


Figure 4.19: Criterion and constraint values for cell δ_{23}

Overall suitability values for other cells are calculated in the same way. Figures 4.20a and 4.20b show the overall suitability for residential development for high-income singles and low-income couples with children, respectively.

(a)

.79	0	0	0	.69	0
0	0	.75	0	0	0
0	0	0	0	0	0
0	0	.71	0	0	0
0	0	.62	0	.46	0
0	0	0	0	0	0

(b)

.65	0	0	0	.56	0
0	0	.7	0	0	0
0	0	0	0	0	0
0	0	.71	0	0	0
0	0	.56	0	.43	0
0	0	0	0	0	0

Figure 4.20: Overall suitability maps for (a) high-income single; (b) low-income couple with children

The next step in the geosimulation-multicriteria modelling process involves ordering the potential alternatives based on their overall suitability scores and selecting the most suitable cells (see Figure 4.10). According to Figure 4.20, δ_{11} is the most suitable cell for residential development from a high-income single perspective and δ_{43} is the most desirable cell for a low-income couple with children. Therefore, these two cells will be converted to the residential type to meet the demand for residential areas. Since δ_{11} and δ_{43} are not available for the next time step, the map of the restricted areas must be updated by the local authority agent by adding δ_{11} and δ_{43} to infeasible areas for residential development (Figure 4.21).

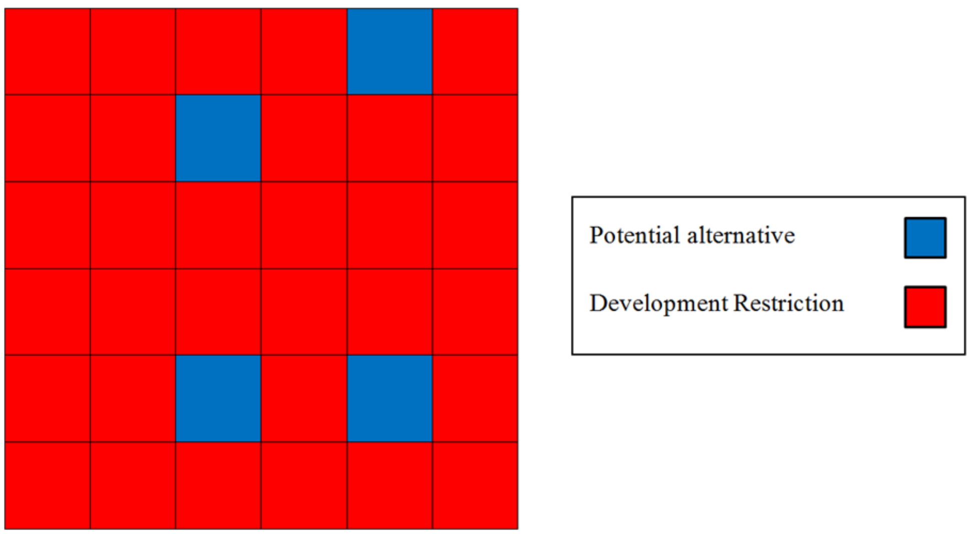


Figure 4.21: Feasible and infeasible cells according to the local authorities' rules after updating

Since the demand has been met, the model execution stops and a map is generated. Figure 4.22 shows the final structure of the landscape (land-use pattern), which can be served as input data in multi-stage geosimulation-multicriteria models.

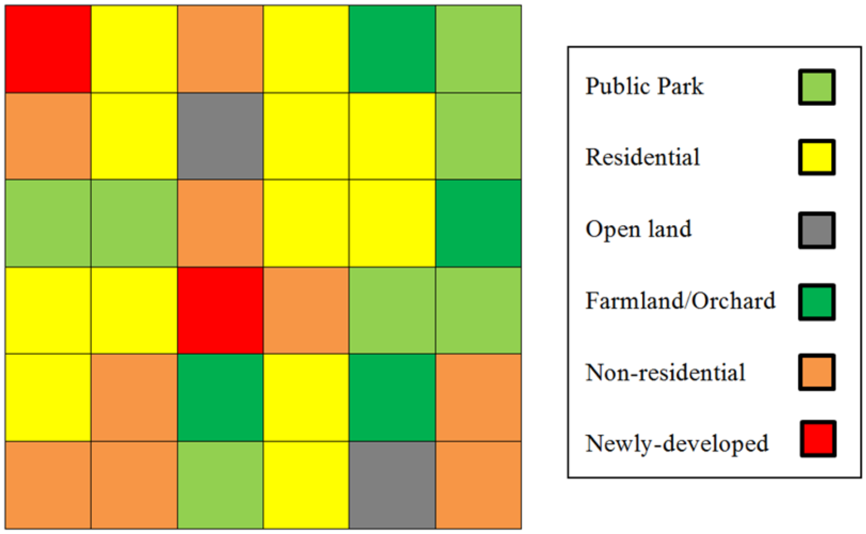


Figure 4.22: New spatial pattern of land-uses

Chapter 5

5 Analysis

5.1 Introduction

In this Chapter, the procedure of acquiring and preparing data for simulating residential land development in Tehran between 1996 and 2006 is explained. The data will be used in Chapter 6 to test the framework/model that was developed in Chapter 4, and compare the results of local and global modelling. Two types of input data are required for operationalizing the geosimulation-multicriteria procedure (see Chapter 4): geographical data (satellite images and criterion maps) and preferential data (that is, preferences of experts/agents regarding evaluation criteria). Details about the geographical data can be found in Appendix D.

5.2 Geographical data: Landscape representation

Landscape is represented by a grid of cells using satellite images. Two satellite images of the study area have been acquired from the U.S. Geological Survey (2016): (i) Landsat TM (Thematic Mapper) image at 30 meter spatial resolution with seven bands for 1996, and (ii) Landsat ETM (Enhanced Thematic Mapper) image at 30 meter spatial resolution with seven bands for 2006. The study area includes the administrative boundaries of Tehran as of 2006 (see Section 3.2.1). The boundary of the study area was delineated on the images. Also, the atmospheric effects and geometric errors were eliminated. Then, satellite images were classified using the maximum likelihood approach (ERDAS, 1999; Erbek, Özkan, and Taberner, 2004). Three classes were defined in the study area: public park/farmland/orchard, built-up, and open land. Figure 5.1 shows the derived images.

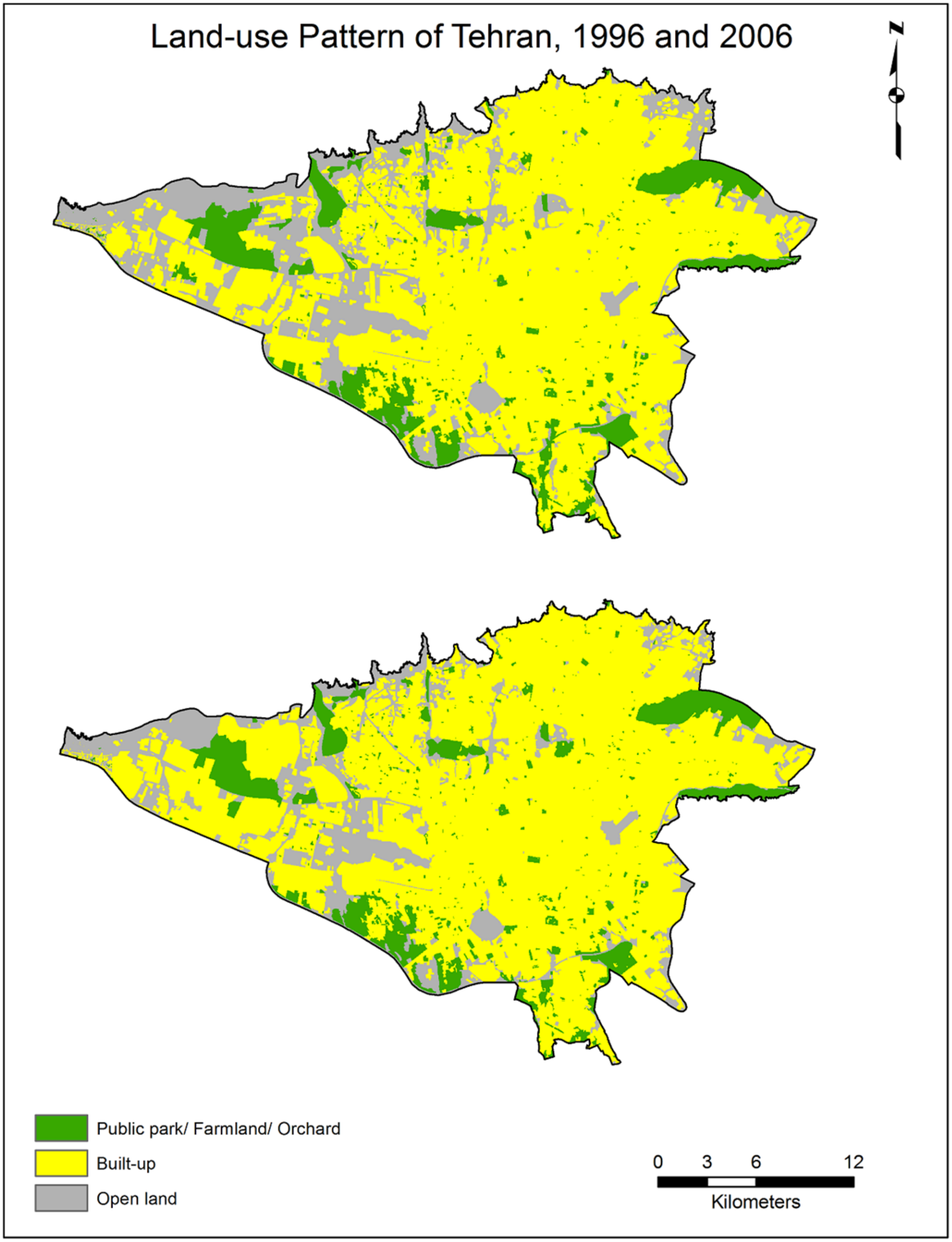


Figure 5.1: Land-use pattern of Tehran: 1996 (top) and 2006 (bottom)

5.2.1 Accuracy assessment

A set of ground truth points (reference points) were randomly collected to evaluate the accuracy of the derived images. Ground truth data were collected using a set of high-resolution WorldView-2 images (provided by DigitalGlobe Foundation, 2017), and aerial photos from the study area (provided by National Cartographic Center, 2015). Eighty well-distributed random points were identified for each land-use type as the reference points. The confusion matrix was produced for the 240 points to compare the result of classification to the reference data. A confusion matrix is a cross-tabulation matrix that compares reference data and classification outputs. The confusion matrices for the accuracy assessment of 1996 and 2006 land-use maps are shown in Tables 5.1 and 5.2. The overall accuracy of the classification for 1996 and 2006 are 87.08% and 89.17%, respectively. The overall accuracy is the ratio of correctly classified points to total number of points. As suggested by Anderson et al. (1976) any overall accuracy higher than 85% is considered acceptable. The Kappa indices for the results of classifications are 87.16% and 96.4 for 1996 and 2006, respectively. The index is a measure of agreement between two images (Congalton and Mead, 1983). Specifically, it compares the accuracy of a classified image to the accuracy expected to be obtained if the image was classified randomly. The z_{score} associated with the Kappa coefficients of 1996 and 2006 are 17.5 and 17.82, respectively. The null hypothesis states that the observed agreement between two images is insignificant. Based on the observed z_{score} the null hypothesis is rejected for both images and therefore the Kappa coefficients are statistically significant at $p = 0.05$.

Table 5.1: Classification accuracy for the Landsat image of 1996

Classified	Reference			
	PFO*	Built-up	Open land	Total
PFO*	69	3	5	77
Built-up	7	75	10	92
Open land	4	2	65	71
Total	80	80	80	240

* PFO = Public park/Farmland/Orchard

Table 5.2: Classification accuracy for the Landsat image of 2006

Classified	Reference			
	PFO*	Built-up	Open land	Total
PFO*	74	0	2	76
Built-up	5	79	17	101
Open land	1	1	61	63
Total	80	80	80	240

* PFO = Public park/Farmland/Orchard

5.2.2 Land-use change statistics

The satellite image of 1996 was employed as the initial state of the land-use pattern and it was intended to model the land-use pattern of 2006 in terms of residential development. Since there is one class for all built-up areas in the derived images, existing land-use maps of 1996 and 2006 were used to make distinction between residential and non-residential areas. Also, it is required to differentiate between farmland/orchard and public park to find out how much farmland/orchard areas were sacrificed for residential development. Therefore, the two images were reclassified into five categories: public park, residential area, open land, farmland/orchard, and non-residential area. Table 5.3 shows the area and percentage of each land-use type. It is found that 24,546 cells have been converted into residential type between 1996 and 2006. Since the spatial resolution of the Landsat images is 30 meters, the area covered by each cell is 900 square meters. Therefore, the residential areas increased by 22.09 square km or 2,209 ha in a ten-year interval.

Table 5.3: Area (ha) and percentage of each land-use types in 1996 and 2006

Land-use type / Year	Public park	Residential area	Open land	Farmland/ Orchard	Non-residential area	Total
1996	4,154 (6.5%)	31,342 (49.3%)	13,071 (20.6%)	2,892 (4.6%)	12,098 (19%)	63,557 (100%)
2006	4,530 (7.1%)	33,551 (52.8%)	9,863 (15.5%)	2,283 (3.6%)	13,330 (21%)	63,557 (100%)

5.3 Preferential data

This section explains how the set of evaluation criteria for residential development was selected based on the experts' opinions. It also describes the procedure for eliciting criterion weights and value functions based on experts' judgments. There are a number of approaches that can be used to elicit the expert's preferences regarding evaluation criteria to be used in geosimulation-multicriteria modelling (Keeney, 1992; Hobbs and Meier, 2012). Examining relevant literature (e.g., Mendoza and Prabhu, 2000; Belton and Stewart, 2002), and surveying opinions using methods such as questionnaires, interviews, focus groups, and the Delphi technique are the most often used methods in GIS-based multicriteria analysis applications (e.g., Fitzsimons et al., 2012; Wang et al., 2013; Comino et al., 2014). This study employs a combination of two approaches: the review of relevant literature (see Chapter 2 and Section 4.4.2) and the focus group technique (Morgan, 1997; Bryman, 2016). In the context of multicriteria analysis, a focus group approach is a form of qualitative research in which participants are asked about their preferences, opinions, and beliefs regarding the decision/evaluation problem and related concepts such as criterion weighting and value function.

Six experts familiar with the study area were selected and asked to participate in the process of identifying a set of evaluation criteria and their preferences with respect to criterion importance and value functions (see Appendices B1, B2 and B3). A meeting of the participants took place on the 17th of June, 2015 in Tehran. The meeting was organized in a workshop/focus group format. I acted as the focus group facilitator. I gave a workshop on the case study and the geosimulation-multicriteria procedure (see Sections 4.4 and 5.2) and assisted the groups through the various stages, eliciting relevant expertise and judgments from the participants. The group was guided through the relevant stages of the geosimulation-multicriteria modelling with appropriate displays of the procedures and results for all to see. The purpose of the focus group meeting was to acquire information and identify preferences required for implementing the geosimulation-multicriteria model. Specifically, the meeting aimed at obtaining information about three elements of the multicriteria procedure: selecting criteria, criteria weighting and value functions (or criteria standardization).

5.3.1 Selecting criteria

The process of selecting criteria (objectives and attributes) involved a four-step procedure (see Section B1 of Appendix B). First, a list of potential criteria was created using a review of relevant literature (see Chapter 2). The set of criteria identified by reviewing geosimulation-multicriteria case studies in Iran is given in Appendix B (see Table B1.1). Second, each participant of the focus group was asked to specify a list of criteria. The lists of criteria suggested by participants individually were then combined. Third, following the group discussion about the criteria identified by the literature review and the combined list of criteria, a set of nine evaluation criteria and four constraints for use in this study was selected (see Table B1.1 and Section 4.4). Fourth, each criterion was associated with an agent's objective. The objectives are to maximize: (i) profit, (ii) accessibility, and (iii) neighbourhood quality, and define (iv) restrictive areas for development. Once the criteria and objectives had been identified, they were organized into a hierarchical structure, which decompose the overall goal (the land suitability for residential development) into the objectives of the three groups of agents and associated criteria (attributes) (see Figure 4.9).

5.3.2 Estimating criterion weights

Different classes of households have different preferences concerning a suitable location for residential area and make their evaluation based on their preferences (see Section 4.4.2.1.2). For example, a couple with children put more emphasis on the proximity to education centres compared to pensioners; working-age households are more sensitive to accessibility to workplaces than retired people (Waddell, 2000). These types of preferences are reflected in different importance weights assigned to the evaluation criteria (attributes) (see Section 4.4.2.1.2). The vector of seven attribute weights associated with each household type and also the weights of two attributes associated with real estate developer agents were obtained based on the experts' judgments. The participants were asked to assign weights of relative importance to the criteria using a two-step procedure (see Section B2 of Appendix B). The procedure involved: (i) ranking the criteria based on their importance by taking into account the range of criterion values (that is, the difference between the best and worst criterion values), and (ii) allocating points among the criteria (rating criteria), with more points to be given to more important criteria (Belton and Stewart, 2002; Hobbs and Meier, 2012). The participants were divided into two groups and then they were asked to rank the nine criteria and then allocate points to the criteria. The individual group's weights were then reported to a plenary session, any significant differences between groups' weightings were discussed, and each group was given the opportunity to revise its weights. It proved possible in the plenary discussion to reach a consensus weighting for each criterion: all participants were content to accept an average of the individual group's weightings, as amended following the plenary discussion, where there remained any difference in those weightings. Table 5.4 summarizes the weights of attributes associated with household and developer agents provided by the experts.

Table 5.4: The weights of attributes

		Household agent							Developer agent	
Household structure	Attribute	Proximity to education centres (w_{kEC})	Proximity to major workplaces (w_{kMW})	Proximity to shopping centres (w_{kSC})	Proximity to major roads (w_{kMR})	Proximity to public transit (w_{kPT})	Green space index (w_{kGS})	Residential intensity index (w_{kRI})	Cost of land acquisition (w_{CL})	Housing price (w_{HP})
Young singles	Low-income	0.00	0.15	0.07	0.06	0.11	0.04	0.05	0.29	0.23
	Medium-income	0.01	0.13	0.06	0.06	0.08	0.08	0.06	0.29	0.23
	High-income	0.01	0.10	0.06	0.08	0.02	0.12	0.09	0.29	0.23
Old couples	Low-income	0.00	0.00	0.16	0.05	0.1	0.12	0.05	0.29	0.23
	Medium-income	0.00	0.00	0.12	0.07	0.08	0.16	0.05	0.29	0.23
	High-income	0.00	0.00	0.1	0.08	0.03	0.19	0.08	0.29	0.23
Couples with children	Low-income	0.10	0.04	0.07	0.04	0.11	0.06	0.06	0.29	0.23
	Medium-income	0.08	0.05	0.05	0.03	0.06	0.13	0.08	0.29	0.23
	High-income	0.08	0.03	0.06	0.07	0.01	0.15	0.08	0.29	0.23
Couples without children	Low-income	0.00	0.12	0.07	0.06	0.13	0.05	0.05	0.29	0.23
	Medium-income	0.00	0.10	0.08	0.08	0.07	0.1	0.05	0.29	0.23
	High-income	0.00	0.07	0.09	0.09	0.01	0.14	0.08	0.29	0.23

The results indicate that participants felt that high-income households care more about neighbourhood quality while low-income households emphasize the importance of accessibility criteria (see Table 5.4). The low-income households of working-age are more willing to live closer to their workplaces with a good access to public transit. Having large amounts of green spaces in the vicinity of a location is of great importance to old couples and high-income households. Except for couples with children who put considerable emphasis on proximity to education centres, other households do not attach any importance to this criterion.

5.3.3 Assessing value functions

A single-criterion value function expresses the relative value of outcomes within the range of criterion (attribute) values (Beinat, 1997; Belton and Stewart, 2002; Malczewski and Rinner, 2015). Given the range of each criterion, a value function was developed to specify the relationship between changes along the range of criterion scores and its value defined on a scale of 0–1. The criterion value functions were obtained using the bisection method (Keeney, 1992; Hobbs and Meier, 2012). The bisection method aided participants to express their opinions about the shape of the criterion value functions (see Section B3 of Appendix B). The experts were divided into two groups and then they were asked to identify a shape of the value function for each criterion. Any significant differences between the value functions generated by the groups were discussed at a plenary session to reach a consensus on the shapes of the value functions. The results of the experts' judgments regarding the shape of the value functions are depicted in Figure 5.2.

According to the experts' judgments, six of the value functions have convex shapes, two have concave shapes, and one has a linear form. The results show that households are very sensitive to distance to shopping centres, workplaces and public transit; as distance of a location to shopping centres, workplaces, or public transit stations increases from zero slightly, the value of the location with respect to these three criteria drops significantly. Figure 5.2 shows that for the five criteria associated with accessibility there can be found a point beyond which people feel indifferent about the distance to the facility. For example, people express no preferences for a location that is 6 km away from their workplace to a location that is 8 km away. The same interpretation applies to the cost of land acquisition criterion. When the price of land is very low, a small change decreases the standardized value substantially, while in high values even a big change has a very little effect on the outcome. For the two attributes associated with the neighbourhood quality, the shapes of the functions are very close to a linear function. Indeed, for residential intensity index, the value function is a linear function. Moreover, the shape of the value function for housing price is monotonically increasing and concave.

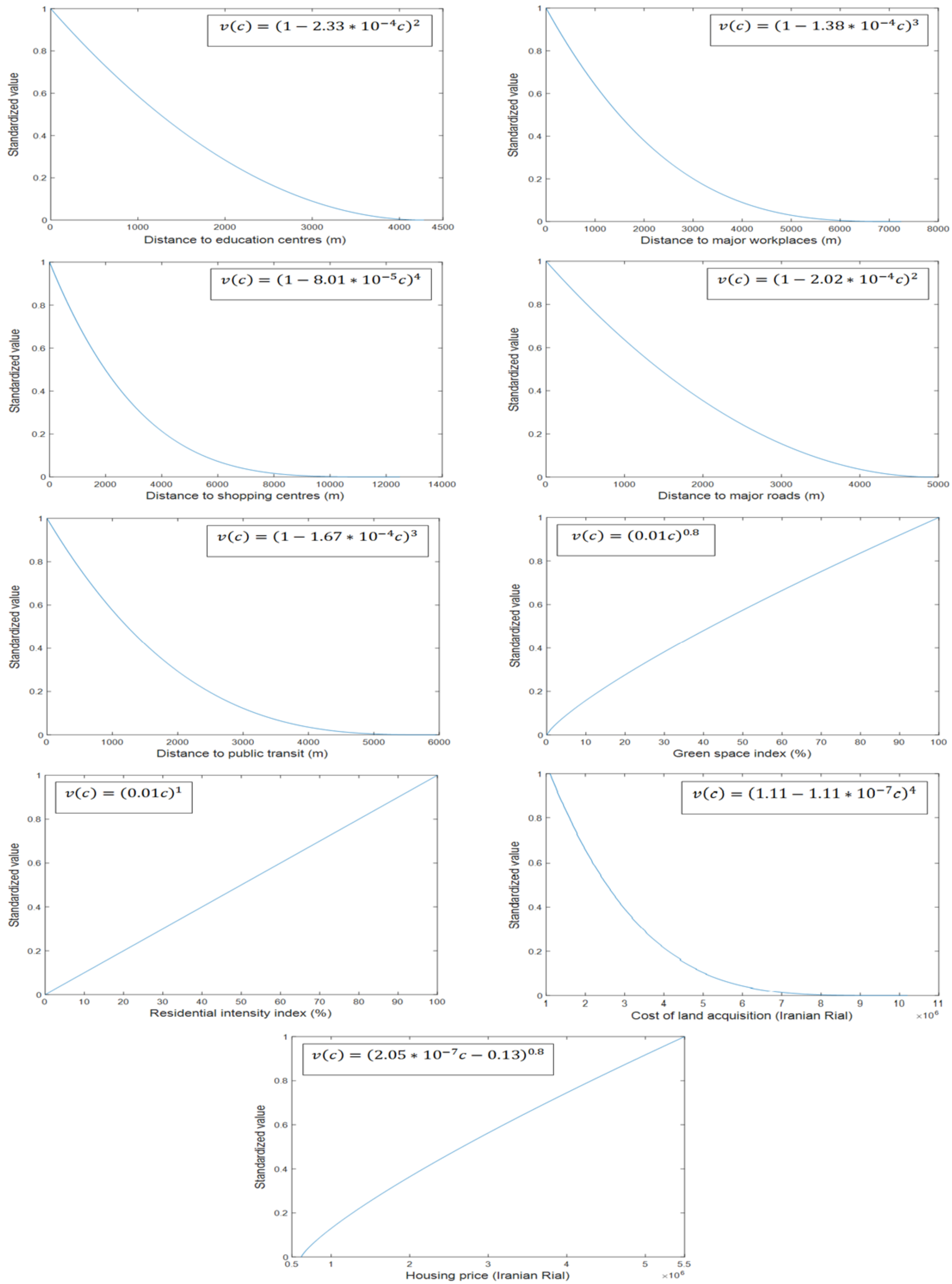


Figure 5.2: The criterion value functions

5.4 Data about agents

5.4.1 Household agents

To initialize household agents, the number of agents in each household type must be determined. The total number of households in Tehran as of 2006 was 2,245,601 (Ranji et al., 2013). Table 5.5 shows how these 2,245,601 households were distributed among the four household structures defined in Section 4.4.2.1.2.

Table 5.5: Classification of households by household structure in 2006 (Source: Ranji et al., 2013)

Household structure		Number of households	Percentage
Old couples		461,167	20.5
Young singles		88,537	4.0
Couples	With children	1,396,984	62.2
	Without children	298,913	13.3
Total		2,245,601	100

By assuming that the three classes of income have been distributed evenly among the four household structures, the percentages and the number of households in each class are obtained (see Table 5.6).

Table 5.6: Classification of households by household structure and income in 2006

Structure \ Income	Low	Middle	High	Total
Old couples	184,139 (8.2%)	138,104 (6.15%)	138,104 (6.15%)	460,347 (20.5 %)
Young singles	35,930 (1.6%)	26,947 (1.2%)	26,947 (1.2%)	89,824 (4%)
Couples with children	558,706 (24.88%)	419,029 (18.66%)	419,029 (18.66%)	1,396,764 (62.2%)
Couples without children	119,466 (5.32%)	89,600 (3.99%)	89,600 (3.99%)	298,666 (13.3%)
Total	898,241 (40%)	673,680 (30%)	673,680 (30%)	2,245,601 (100%)

Classified images showed an increase of 24,546 cells in the number of cells that belongs to the residential type from 1996 and 2006 (see Section 5.2.2). Since it is assumed that each cell is selected by one household agent, the initial number of household agents in the model is equal to the number of cells that have been converted to the residential type. The number of each type of household agents that needs to be initialized at the beginning of the modelling process is calculated through multiplying the total number of agents by the percentage of each household type. Table 5.7 shows the initial number of each type of household agents.

Table 5.7: The initial number of each type of household agents

Structure \ Income	Low	Middle	High	Total
Old couples	2,012	1,510	1,510	5,032
Young singles	393	295	295	983
Couples with children	6,107	4,580	4,580	15,267
Couples without children	1,306	979	979	3,264
Total	9,818	7,364	7,364	24,546

5.4.1.1 Household agent behaviour

Each household agent examines any parcel of land among seven criteria to find the most suitable location for residential development (see Section 4.4.2.1.3). The seven criteria include: proximity to education centres, proximity to major workplaces, proximity to shopping centres, proximity to major roads, proximity to public transit, green space index, and residential intensity index. Figure 5.3 shows the spatial distribution of education centres (including public schools and universities), major workplaces (major industrial and commercial areas), shopping centres (large stores), and major roads for 1996 and subway stations for 2006. The spatial layers for 1996 were used because they represent the state of the landscape at the beginning of the model execution. However, the first reliable GIS data for subway stations has been generated in 2006. By looking at the distribution of the education and shopping centres and subway stations, one can find a relatively low concentration of facilities in the western part of the city. This can be attributed to the lower population and residential areas in the western section of the city before 2006 (see Section 3.3.2 and Figure 5.4). Moreover, most major workplaces are agglomerated on the south-western and western part of Tehran. Figure 5.4 demonstrates public parks and residential areas for 1996. It can be seen that most major parks are located on the peripheral districts of Tehran. There is a high concentration of residential areas in the central districts of the city, while peripheral districts, especially on the western parts, are less developed in terms of residential areas.

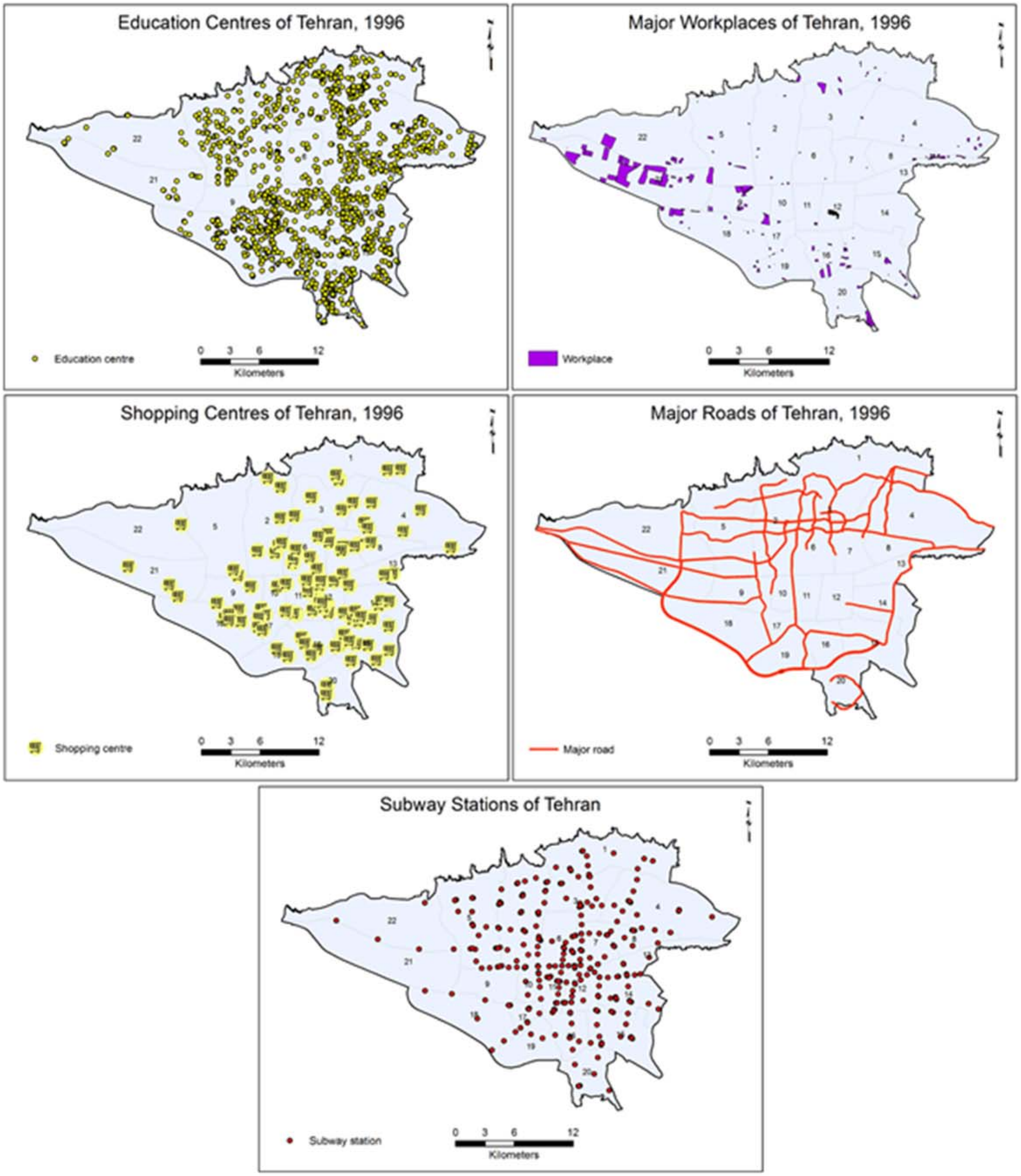
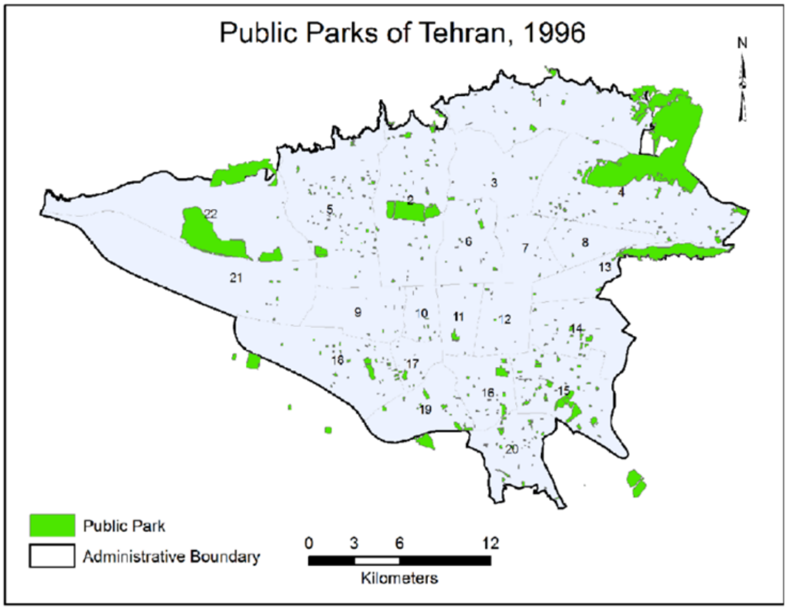


Figure 5.3: Distribution of facilities and amenities in Tehran (Data source: Iranian National Cartographic Center)



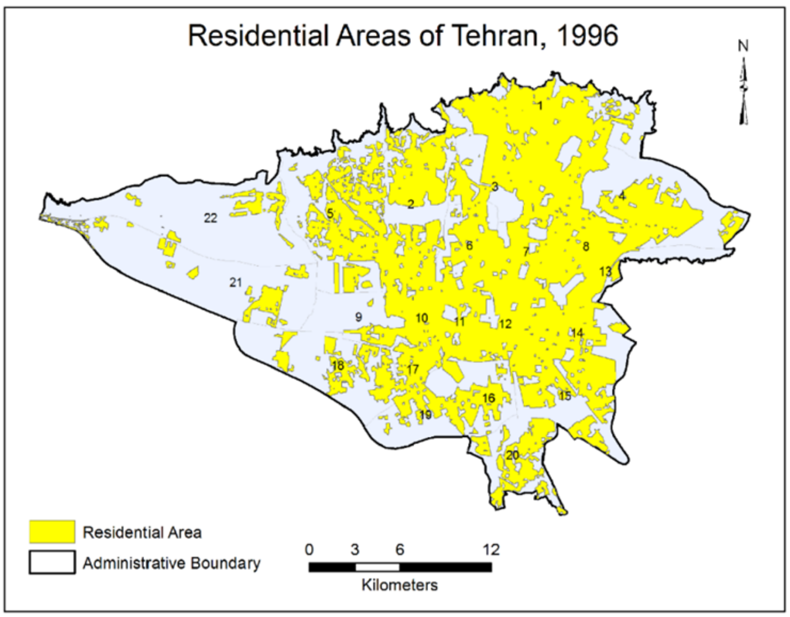


Figure 5.4: Public parks (top); residential areas (bottom) in Tehran, 1996 (Data source: Iranian National Cartographic Center)

5.4.2 Real estate developer agent

As discussed in Section 4.4.2.2, real estate developer agents try to maximize their profits. Therefore, they evaluate the suitability of each cell for development using two criteria: housing price and cost of land acquisition. Figure 5.5 shows spatial patterns of housing price and cost of land acquisition in Tehran. The patterns indicate that the most expensive houses and lands can be found in the northern part of the city. As one moves from the northern to the southern districts, the housing price and cost of land decrease gradually. The reason behind this pattern is related to differences between northern and southern parts in terms of physical and social conditions (see Sections 3.2.2, 3.2.3, and 3.4). Also, the west section of the city is characterized with low land cost and low-to-medium housing price. The southernmost part of Tehran is characterized by the lowest housing prices and land costs.

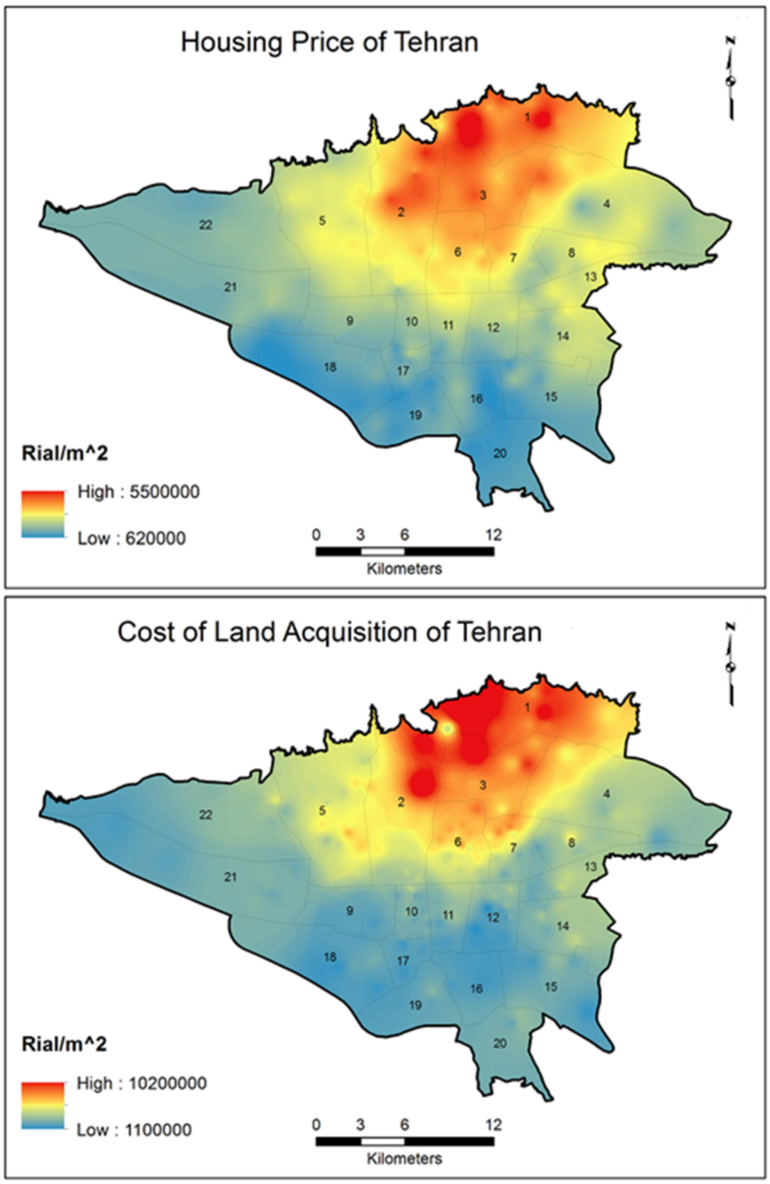


Figure 5.5: Housing price (top); cost of land acquisition (bottom) in Tehran, 1996 (Data source: Iranian Ministry of Roads and Urban Development)

5.4.3 Local authority agent

As explained in Section 4.4.2.3, developer agents must follow the rules established by local authority agents. Developing residential areas in restrictive lands is not permitted by local authorities. Figure 5.6 demonstrates the restrictive areas at the beginning of the geosimulation-multicriteria model execution. Restrictive areas are updated after each model run. The initial map was generated by aggregating the four map layers associated with four development constraints (see Section 4.4.2.3). Most feasible lands for development are located in the westernmost districts of the city (i.e., Districts 21 and 22). Also, there is ample opportunity for development in some south-western and southern districts of Tehran.

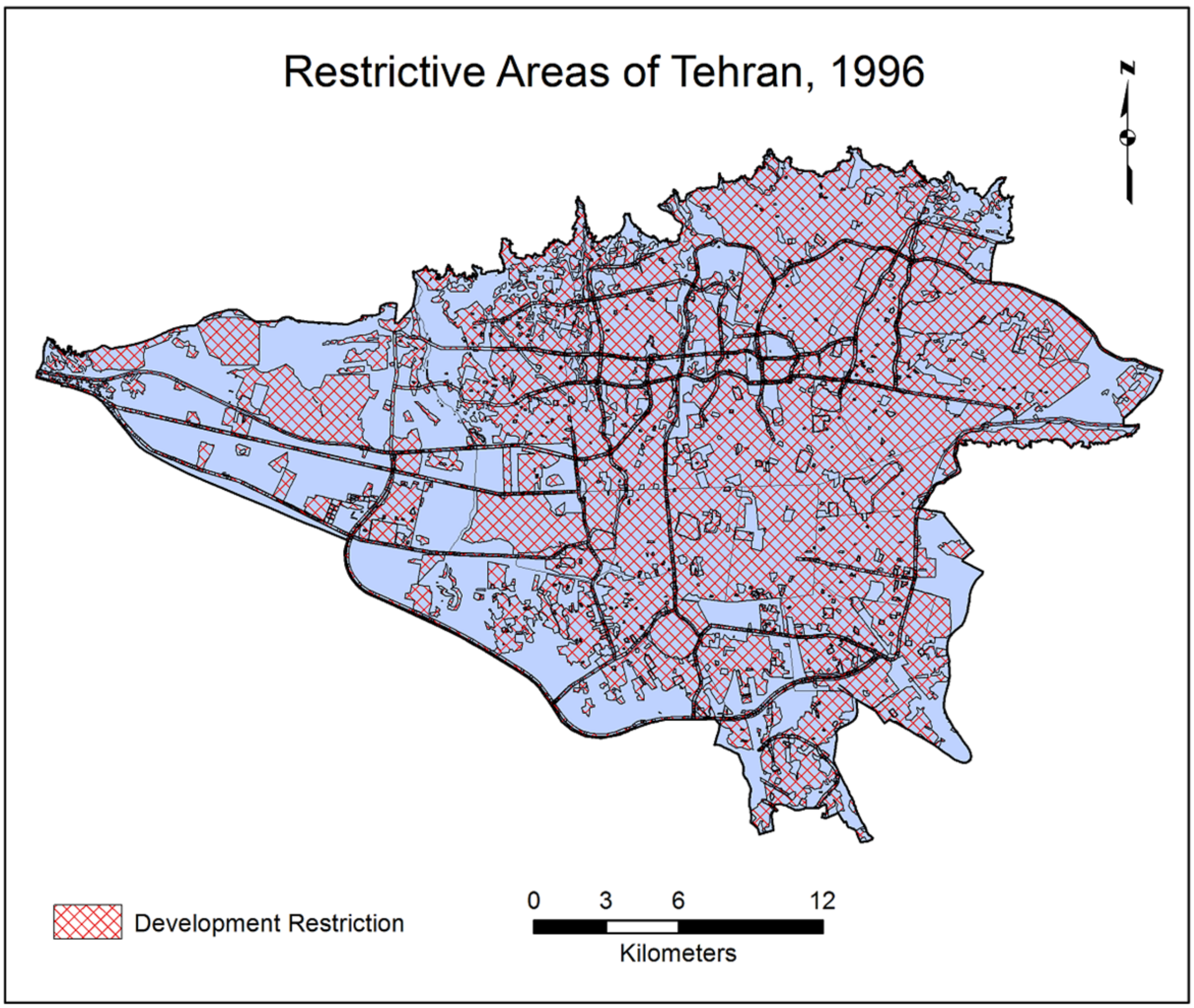


Figure 5.6: Restrictive areas for residential development in Tehran, 1996

Chapter 6

6 Results and discussion

6.1 Introduction

This chapter focuses on the application of the geosimulation-multicriteria modelling framework (see Chapters 4 and 5). Forty-two scenarios of residential development in Tehran are defined and then the results obtained by the scenarios are examined and evaluated. The output of each scenario is compared to the results of other scenarios and to the actual land-use pattern in the city in 2006. The analysis centres on comparing the results generated by different scenarios in terms of the two components of the geosimulation-multicriteria model: the linguistic quantifiers (or associated order weights) and neighbourhood size (or order of contiguity) used for local multicriteria modelling. A series of hypotheses is put forward to analyse how the linguistic quantifiers and size of neighbourhood affect the results of global and local geosimulation-multicriteria models.

6.2 Results of geosimulation-multicriteria modelling

6.2.1 Defining scenarios

The spatial pattern of residential development was simulated using 42 scenarios: 35 local scenarios and 7 global scenarios. The former were created based on different definitions of neighbourhood and sets of order weights (or associated linguistic quantifies). The global scenarios were generated based on the seven sets of order weights. The linguistic quantifier method (see Section 4.3.1.1.2.1) was used for calculating the order weights for both models (see Equation 4.13). The advantage of this approach is that the order weights are generated according to linguistic statements (see Section 4.3.1.1.2.1).

The definition of neighbourhood size and type lies at the core of local multicriteria analysis (see Section 4.3.1.1.2.3). Indeed, the output of the local multicriteria function and, therefore, the result of the geosimulation-multicriteria procedures depend on the neighbourhood structure (e.g., Eldrandaly, 2013; Liu et al., 2014; Malczewski and Liu, 2014; Cabrera-Barona et al., 2015). There is no theoretical justification regarding the best neighbourhood structure in urban models (Liu, 2008). Ligmann-Zielinska and Sun (2010) and Yu et al. (2011) used Moore neighbourhood (first-order Queen contiguity). On the other hand, some studies employed larger neighbourhood sizes. White and Engelen (1994) and White, Engelen, and Uljee (1997) used 113 cells surrounding a target cell as its neighbours. According to Liu (2008), most applications in urban studies apply larger neighbourhoods than studies in natural sciences. This study used the extended Moore neighbourhood with five different sizes. The contiguity order, range, and size of the neighbourhoods (see Section 4.2.1.2) and the area they cover are given in Table 6.1. The 5×5 Moore neighbourhood was selected as the smallest neighbourhood size. It is the most often used neighbourhood size in geosimulation studies (e.g., Myint and Wang, 2006; Mitsova, Shuster, and Wang, 2011; Zhang et al., 2011; Moghadam and Helbich, 2013; Mokadi, Mitsova, and Wang, 2013; Terra, dos Santos, and Costa, 2014; Bozkaya et al., 2015; Nourqolipour et al., 2015; Hyandye and Martz, 2017; Sakieh, Salmanmahiny, and Mirkarimi, 2017). The largest neighbourhood has a range of 960 meters that is very close to 1 kilometer suggested by Liu (2008) as the very large neighbourhood size. The five contiguity orders are the sequence of powers of 2; that is, 2^1, 2^2, 2^3, 2^4, and 2^5.

Table 6.1: Neighbourhood and linguistic quantifier definition

Neighbourhood definition						
Definition	Very Small	Small	Medium	Large	Very Large	Study area
Contiguity order	2	4	8	16	32	Global (778)
Range (meters)	60	120	240	480	960	N/A
Window size	5×5	9×9	17×17	33×33	65×65	N/A
Area (ha)	2.25	7.29	26.01	110.25	40.01	63557.19

Linguistic quantifier definition							
Quantifier	At least one	Few	Some	Half	Many	Most	All
α	0.0001	0.1	0.5	1	2	10	1000

6.2.2 Land suitability analysis

This section examines a selection of maps to illustrate the differences between the results of the global and local OWA models in terms of the main components of the land suitability model: criterion values (value functions), criterion weights and overall evaluation scores (see Equation 4.25).

6.2.2.1 Global and local value functions

Value functions are used to standardize the criterion values to [0, 1] interval. Figures 6.1 and 6.2 show the global and local standardized criterion maps that were created based on the global and local value function models, respectively (see Equations 4.6 and 4.15, Section 5.3.3, and Figure 5.2). When comparing the spatial patterns generated by the global and local models, it is clear that the local model generates more extreme criterion values. This is due to the fact that high and low criterion values appear in each neighbourhood in the local modelling. The extreme values in the local models are more isolated relative to the global models. In the global models, the high and low criterion values tend to cluster around global extreme values. This pattern is, in particular, exemplified by the cost of land acquisition and housing price criteria. While high values for the two criteria can be spotted in a few parts of Tehran based on the results of the global models, there are a large number of local high values generated by the local models. Moreover, the global spatial patterns of these two criteria are characterized by a high degree of aggregation, while the spatial patterns generated by the local models are relatively dispersed. These kinds of differences between the spatial patterns generated by the global and local value functions can also be identified in other criterion maps; however, it is not as evident as in the case of the cost of land acquisition and housing price criteria.

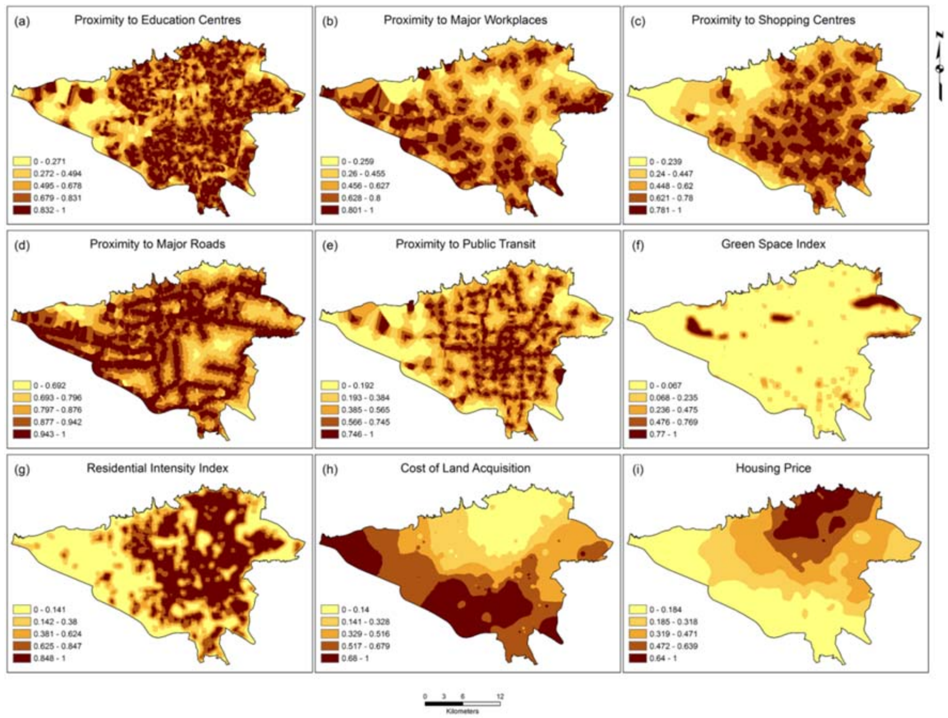


Figure 6.1: The standardized criterion maps generated by the global value functions

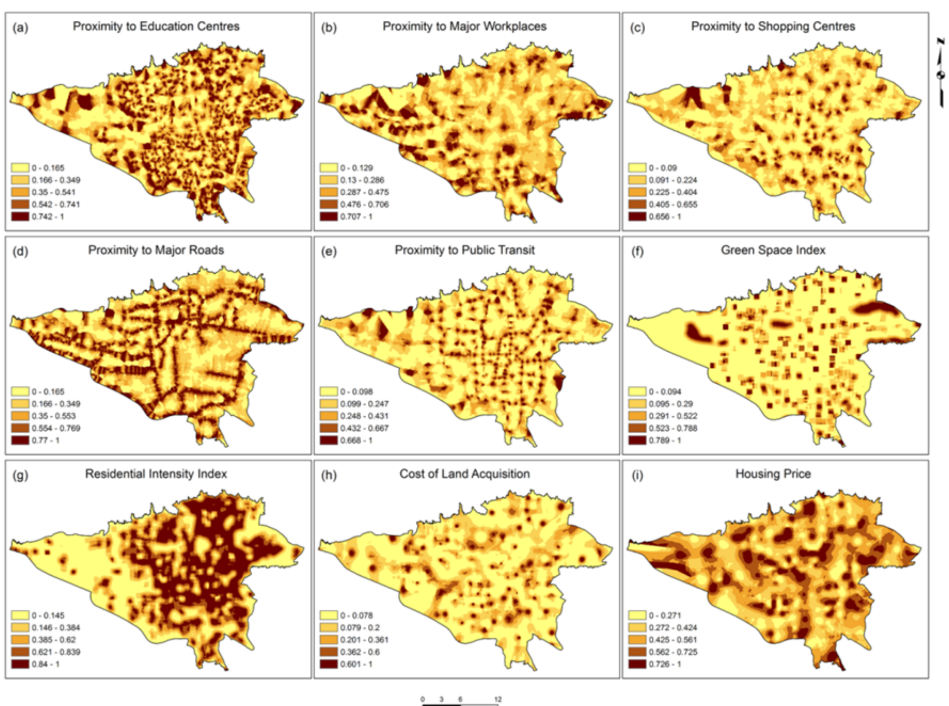


Figure 6.2: The standardized criterion maps generated by the local value function with contiguity order of 32

6.2.2.1.1 Value functions and different neighbourhood sizes

Figure 6.3 shows how the pattern of standardized criterion maps changes with increasing neighbourhood size. The maps were created for cost of land acquisition because it is easier to see the dissimilarities between different models. As can be seen, the spatial pattern of standardized values gets smoother by increasing the neighbourhood size. In the global model, high values are clustered in the western and southern parts and low values are concentrated in the northern parts of the city. However, extreme values are dispersed all over the study area in the local models. The model with contiguity order of 2 has the most dispersed pattern with a large number of isolated high values. By increasing the size of the neighbourhood, the pattern of high values becomes more aggregated.

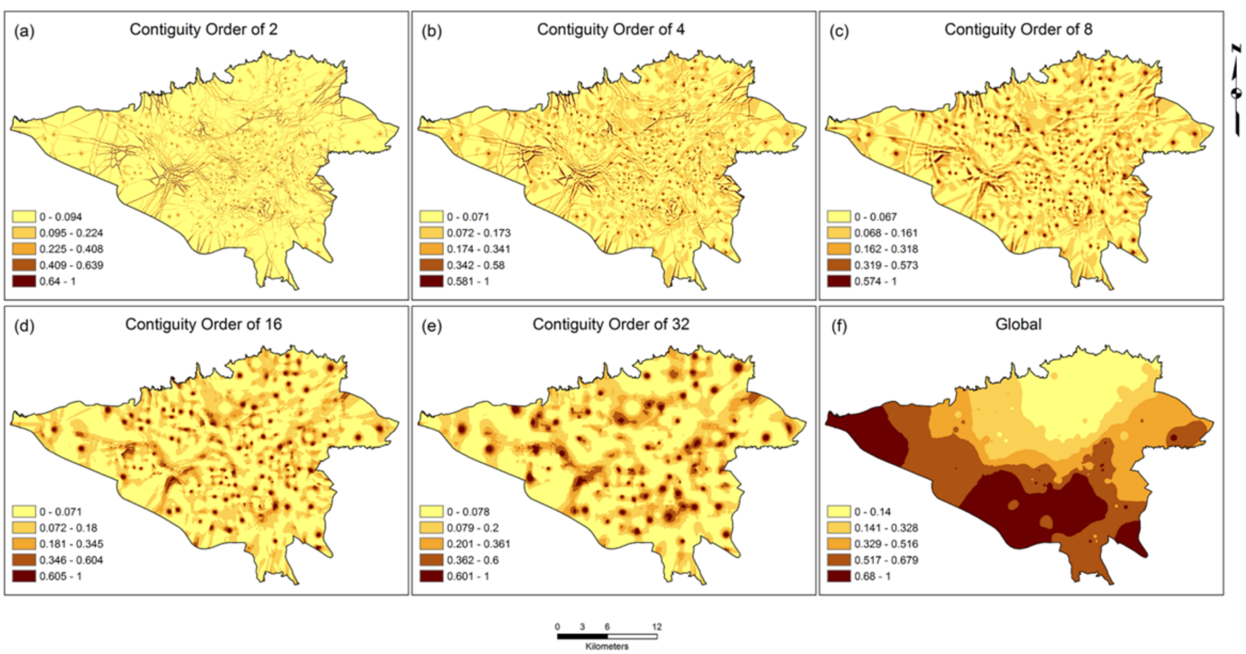


Figure 6.3: Standardized criterion maps created for cot of land acquisition using different neighbourhood sizes

6.2.2.2 Global and local criterion weights

The preferences of different types of agents with respect to the evaluation criteria, which are reflected in criterion weights, are constant over the study area in the global geosimulation-multicriteria modelling; that is, each evaluation criterion is assigned a single weight of relative importance in the global model (see Section 5.3.2 and Table 5.4). The agents' preferences (criterion weights) change from one location to another in local geosimulation-multicriteria modelling. Figure 6.4 shows the spatial pattern of preferences for low-income couples with children and real estate developer agents for the contiguity order of 32. Low-income couples with children were selected because it has the most number of initial agents in the geosimulation procedure (see Table 5.7). Seven of the maps (Figures 6.4a to 6.4g) show the spatial pattern of criterion weights for low-income couples with children and the other two (Figures 6.4h and 6.4i) are related to the real estate developer agent. Since the distribution of facilities and events are not uniform across the spatial space, the local ranges of values vary over the study area and the criterion weights change based on the local range (see Equation 4.17). Examining the spatial pattern of local weights reveals that the accessibility criteria, including proximity to education centres, major workplaces, shopping centres, major roads, and public transit (see Section 4.4.2.1.3.1) are relatively more important in the western parts of the city. This is due to the lack of urban facilities and poor transportation networks in those parts (see Section 3.3.2 and Figure 5.3). This also causes the criteria of proximity to education centres and public transit, which are globally very important for low-income couples with children (see Table 5.4), to have relatively very high importance in the western sections of the city and lower importance in the central parts. The weights of proximity to education centres and public transit in western Tehran range roughly from 0.12 to 0.3 and 0.12 to 0.27, respectively. On the other hand, low-income couples with children put less emphasis on five criteria related to accessibility in the central part of the study area. This can be attributed to the abundance of urban facilities and good transportation networks in the core area of Tehran (see Figure 5.3). With respect to two criteria related to the neighbourhood quality, i.e., green space index and residential intensity index, the central parts of the city that are more developed receive relatively less importance as compared to the peripheral districts that are less developed. The two criteria that are associated with the real estate developer agent, i.e., cost of land acquisition and housing price (see Section 4.4.2.2), have relatively lower importance in western Tehran and higher importance in central parts of the city.

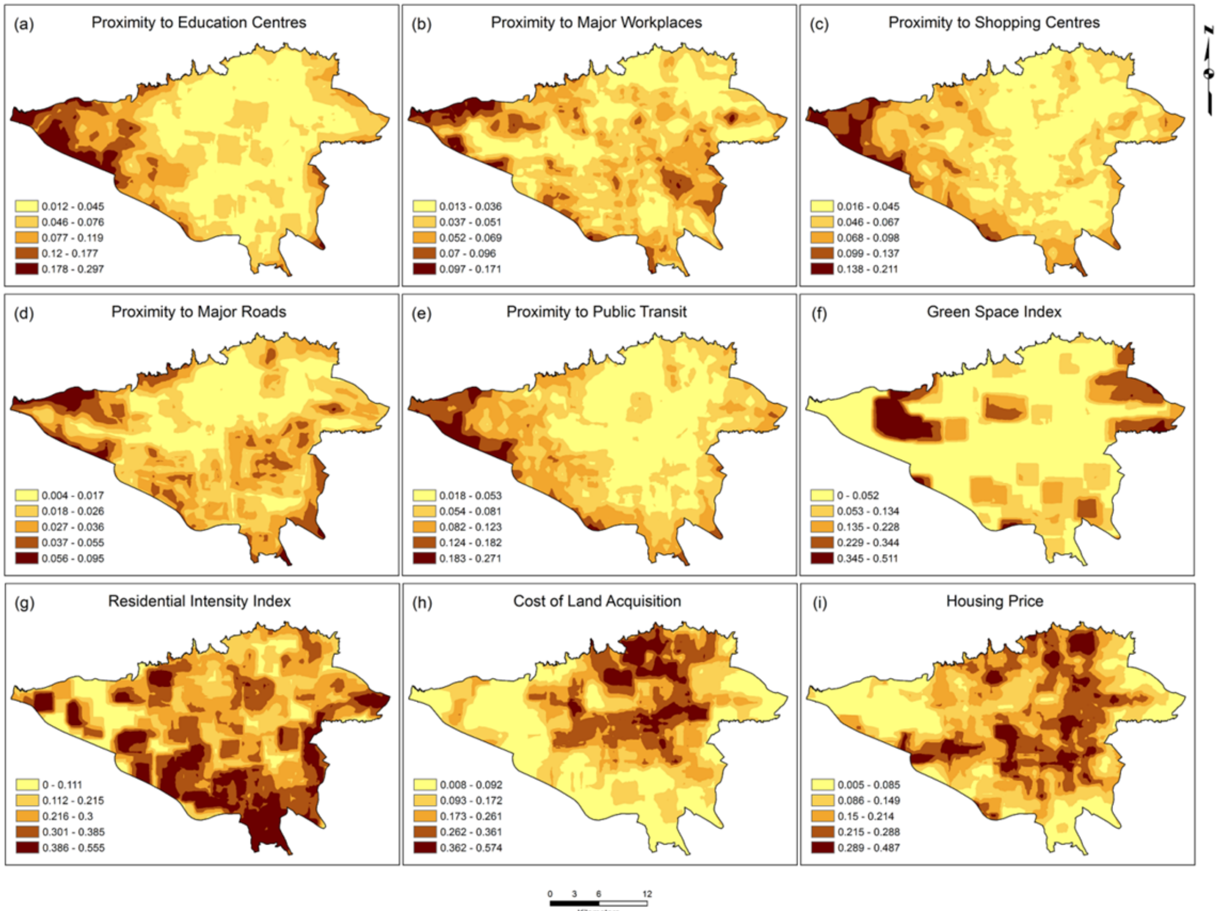


Figure 6.4: Spatial patterns of local preferences for the criteria associated with low-income couples with children (a to g) and the real estate developer agent (h and i)

6.2.2.2.1 Local criterion weights and different types of household agents

To examine how difference in household agents' preferences affects the pattern of local weights, the spatial pattern of preferences for high-income old couples generated using the local model with the contiguity order of 32 (Figure 6.5). The spatial pattern of preferences for two criteria, i.e., proximity to education centres and proximity to major workplaces, are completely different from the ones for low-income couples with children. The reason of that is related to the global weights of these two criteria (see Table 5.4). The global importance weights associated with these two criteria is zero for high-income old couples. Apart from these two criteria, Figure 6.5 shows that the general patterns of local criterion weights for high-income old couples is very similar to the ones for low-income couples with children. For instance, for proximity to shopping centres, major roads, and public transit, the highest values can be identified in the western parts of the city. This is due to the fact that the range of values for each criterion within in each neighbourhood is independent of the type of households (see Equation 4.17). However, the ranges of local weights are quite different because of the difference in the global importance of the criteria. For example, Table 5.4 shows that the global importance weights for proximity to major roads are 0.04 and 0.08 for low-income couples with children and high-income old couples, respectively. As a result, the range of local weights for proximity to major roads is roughly 0.04 to 0.1 for low-income couples with children and 0.1 to 0.3 for high-income old couples in the western parts of the study area. To examine if the same patterns exist for other types of households, a map of local weight for proximity to shopping centres is created for all twelve types of households (Figure 6.6). Proximity to shopping centres was selected because it is non-zero for all household types. Figure 6.6 confirms that the spatial patterns of local weights are very similar irrespective of the type of household. However, the range of the local weights heavily depends on the global weights.

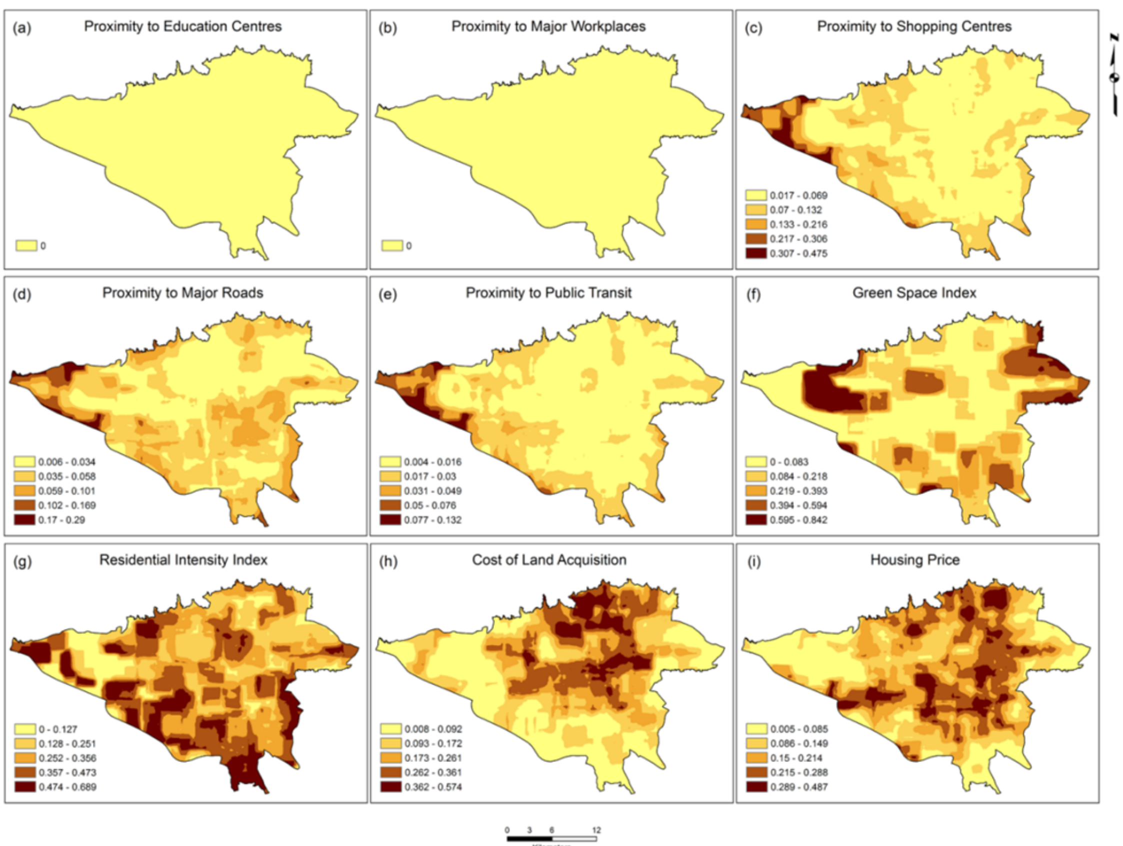


Figure 6.5: Spatial patterns of local preferences for the criteria associated with high-income old couples (a to g) and the real estate developer agent (h and i)

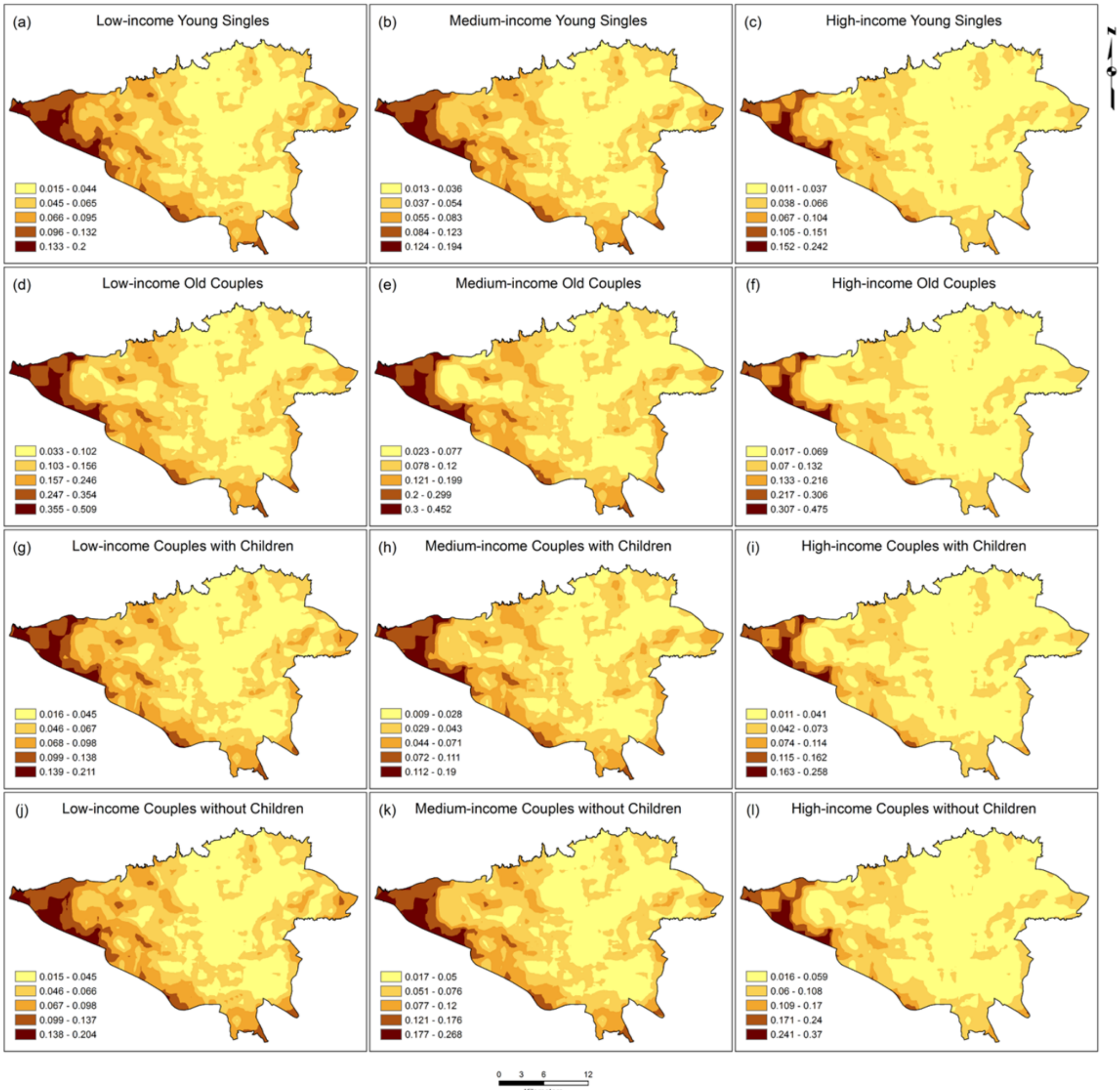


Figure 6.6: Spatial patterns of local preferences with respect to proximity to shopping centres for all types of household agents

6.2.2.2.2 Local criterion weights and different neighbourhood sizes

To examine how difference in neighbourhood sizes affects the pattern of local weights, the spatial pattern of preferences for proximity to shopping centres for low-income couples with children was generated by the local model with five contiguity orders (see Table 6.1). Figure 6.7 shows the result of different contiguity orders. As it can be seen, the spatial pattern of local preferences gets smoother by increasing the size of the neighbourhood. This came as no surprise because sudden changes happen more often in smaller neighbourhood sizes. Moreover, the size of the neighbourhood has little to do with the range of the local weights, as it was expected.

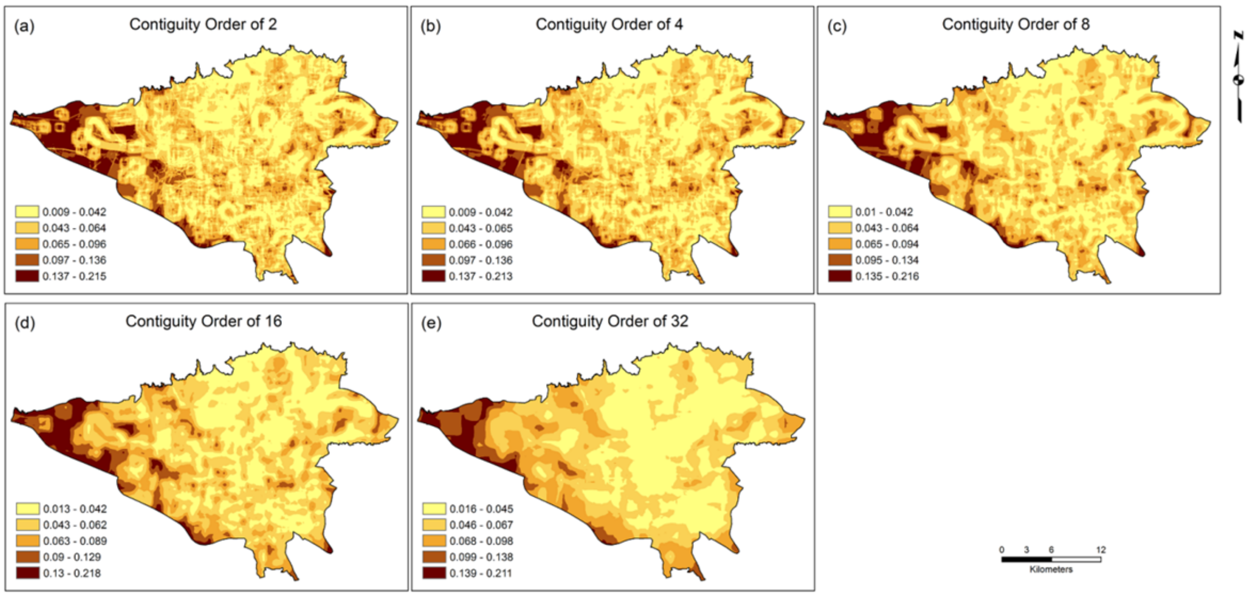


Figure 6.7: Spatial patterns of local preferences with respect to proximity to shopping centres for different neighbourhood sizes

6.2.2.3 Global and local evaluation scores

Figures 6.8 and 6.9 show the suitability of each cell for development from the perspective of low-income couples with children and the real estate developer using global and local models, respectively. The order of the contiguity is 32 for the local models and the already developed cells are assigned zeros. Suitability maps were generated by combining contributory criteria and their associated importance weights (see Section 4.4.3 and Equation 4.25) for each linguistic quantifier. Associated with each linguistic quantifier, there is a corresponding α parameter (see Table 6.1). The value of α increases as one moves from the “At least one” to “All” quantifier. As was expected, the range of evaluation scores decreases with an increase in the value of α, because the suitability of each cell is evaluated by emphasizing the negative aspects of it. The negative aspects of each cell are those criteria in which the cell performs worse. Accordingly, by moving gradually from “At least one” to “All” quantifier, higher order weights are assigned to lower criterion values at a given cell and vice versa. When the “All” quantifier is applied in the model, all criteria need to be satisfied by an acceptable alternative (see Section 4.3.1.1.2.1). This embodies the extremely pessimistic situation (the worst-case scenario). In this situation, the suitability of cells is evaluated on the basis of lowest criterion values. Therefore, the evaluation scores for the “All” quantifier are relatively lower than other quantifiers in both global and local models. Moreover, the “All” quantifier has the highest number of zeros as compared to other quantifiers. On the other hand, the “At least one” quantifier represents the extremely optimistic scenario. In this situation, cells are assessed based on the highest criterion values. Accordingly, the “At least one” quantifier has, relatively, the highest evaluation scores.

Comparing the spatial patterns of evaluation scores generated by the global and local models, one can make the following observations: (i) since the order of the contiguity is 32, which is relatively large, the patterns are very similar in some cases; e.g., the “At least one” scenario; (ii) the results of both global and local modelling indicate that the north-western and south-western sections of the city have relatively higher suitability values compared to other parts; and (iii) the spatial patterns of suitability scores are more dispersed in the local models as compared to the global ones (this can be attributed to the tendency of low values and high values to cluster around absolute minimum and maximum values in the global modelling).

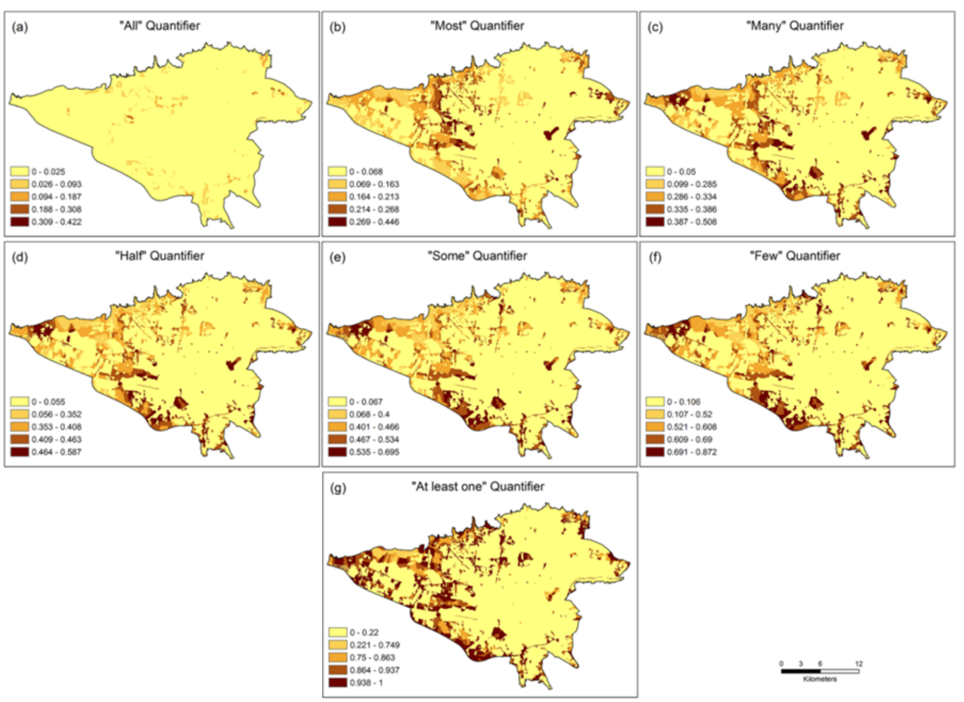


Figure 6.8: Suitability maps generated by the global models among different linguistic quantifiers

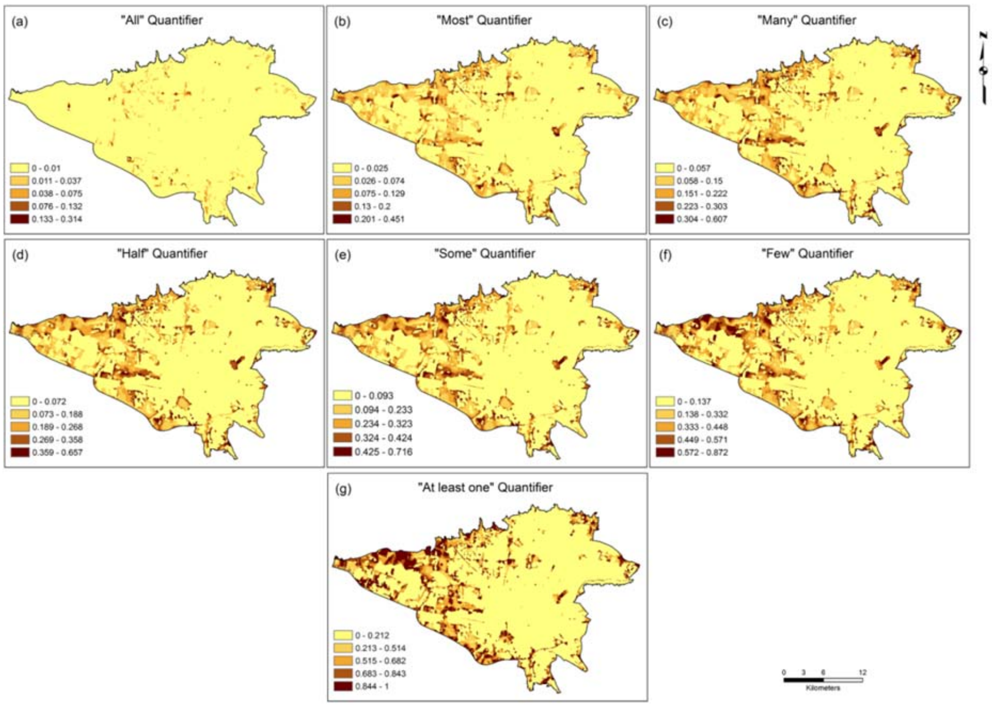


Figure 6.9: Suitability maps generated by the local models with contiguity order of 32 among different linguistic quantifiers

6.2.2.3.1 Local evaluation scores and different neighbourhood sizes

Figure 6.10 shows how the pattern of evaluation scores changes with increasing the neighbourhood size. The suitability maps were generated for low-income couples with children and the linguistic quantifier “Half”. The linguistic quantifier “Half” was selected because it represents the weighted linear combination function (see Section 4.3.1.1.2.1). An interesting finding is that the range of suitability scores decreases with increasing the neighbourhood size. The largest range of suitability is associated with the local model with the contiguity order of 2, which is roughly from 0 to 0.93; while the global model has the smallest range, which is roughly from 0 to 0.59. This can be attributed to the fact that in the local modelling, values are standardized with respect to the neighbourhood within which they are located (see Sections 4.3.1.1.2.3 and 6.2.2.1, and Equation 4.15). Therefore, as the size of the neighbourhood decreases, there are more local extreme values in each criterion map after standardization. Having more extreme values in the criterion maps, results in having greater range of evaluation scores after combination process.

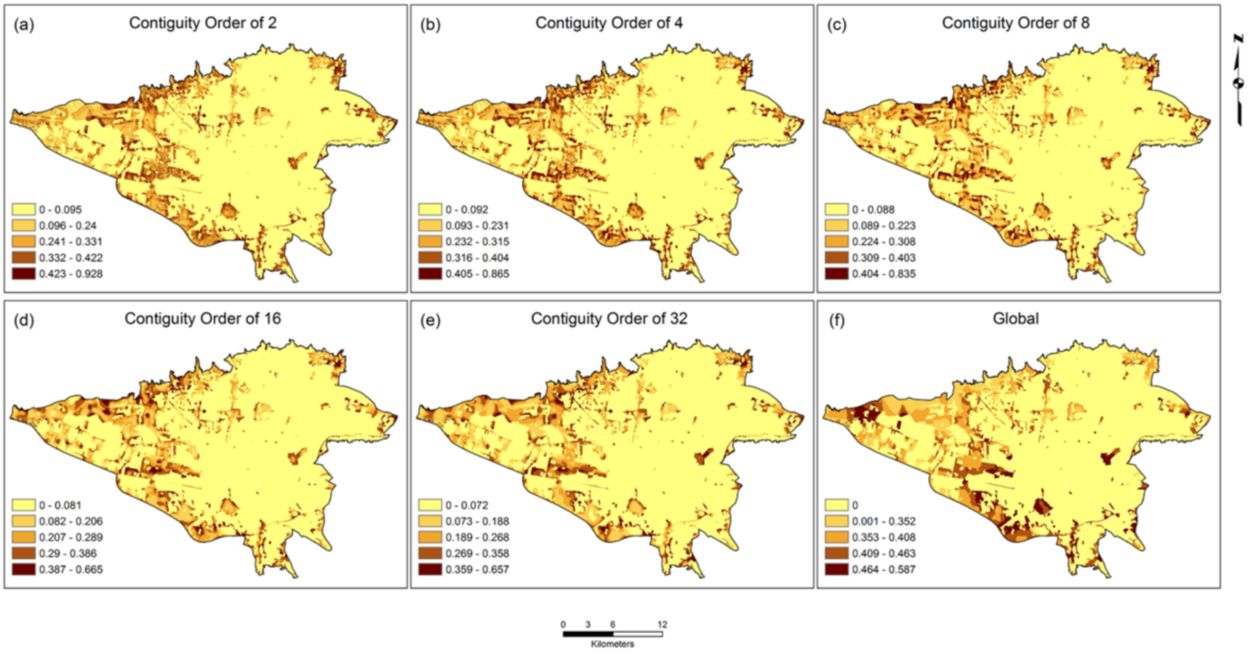


Figure 6.10: Suitability maps for low-income couples with children using the "Half" linguistic quantifier and different neighbourhood sizes

6.2.3 Evaluating and comparing global and local models

The evaluation of the results of the global and local geosimulation-multicriteria models can be done by analysing cross-tabulation matrices and morphological/spatial characteristics of land-use patterns (Li and Liu, 2007). Both approaches are used in this study. The cross-tabulation matrix (confusion matrix) provides the base for determining accuracy assessment metrics (see Section 6.2.3.1). Moreover, several measures are considered for evaluating morphological/spatial properties of land-use patterns (see Section 6.2.3.2). Since the aim is to simulate residential development, the focus is on the residential land-use type while assessing the performance of the 42 scenarios.

6.2.3.1 Accuracy assessment metrics

To evaluate the result of each scenario, a cross-tabulation matrix of the simulation output and reference data for 2006 was constructed. However, instead of five categories, two categories of land-uses were considered: undeveloped and residential. The undeveloped category includes two types of lands: open lands and farmland/orchards. The main diagonal of the cross-tabulation matrix contains the number of correctly classified (simulated) cells and the counter diagonal includes the number of cells classified incorrectly. To describe the cross-tabulation matrix, the following indices were employed in this research: overall accuracy, error of commission, error of omission, and two Kappa indices (Congalton and Green, 1999; Liang, 2004). Overall accuracy shows the percentage of the cells correctly converted to the residential type to the total number of observations in the matrix. However, this index is not very useful here, because the cells in the undeveloped category were not classified by the simulation process, and therefore, there is an overestimation in the simulation accuracy. A more useful approach is to compute the error of the residential category. There are two indices that describe the error of a single category in the confusion matrix: error of commission and error of omission. Error of commission happens when undeveloped cells in the actual image are incorrectly included in the residential category in simulation results. Error of omission occurs when residential cells in the actual image are left out of the residential category in simulation outputs. Kappa index (κ) is the measure of agreement between simulated and actual (observed) land-use patterns. The index is calculated as follows (Congalton, 1981; Congalton, Oderwald, and Mead, 1983):

$$\kappa = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \times x_{+i})} \tag{6.1}$$

where *r* is the number of rows in a cross-tabulation matrix; x_{ii} indicates the value of element ii in the matrix; x_{i+} and x_{+i} are the marginal totals for row *i* and column *i*, respectively; *N* is the total number of observations in the matrix.

Kappa index considers both residential and undeveloped categories in calculating the agreement between two spatial patterns. However, the present study is primarily concerned with the agreement between two maps of residential land-use types. Therefore, the following equation is applied to measure the agreement of a single class (element) i between two maps (Congalton and Green, 1999; Paine and Kiser, 2003):

$$\kappa_i = \frac{N(x_{ii}) - x_{i+} \times x_{+i}}{N(x_{i+}) - x_{i+} \times x_{+i}} \tag{6.2}$$

Therefore, the Kappa index is also calculated for the residential category (Kappa$_r$ or κ_r) to gain better insights into the accuracy of the simulation outputs. Figures C1-C7 and Table C1 summarize the results generated by the geosimulation-multicriteria models/scenarios (see Appendix C).

There are linear relations between κ_r and κ, and between κ_r and overall accuracy (see Table 6.2). The κ_r index is characterized by an inverse linear relation with the commission and omission errors. Thus, a scenario with a relatively higher κ_r is described by high values of overall accuracy and κ and relatively low values of the commission and omission errors. The κ_r index will be used for analysing the outputs. Given the relations between accuracy measures, similar conclusions can be reached by analysing other measures of accuracy.

Table 6.2 Relations between κ_r and measures of accuracy: the overall accuracy, error of commission, error of omission, and κ (Data sources: Appendix C, Table C1)

Relation between κ_r and:	Linear function
Overall accuracy (*OA*)	$\kappa_r = 3.360 \times OA - 236.352$
Commission error (*CE*)	$\kappa_r = -1.224 \times CE + 100.127$
Omission error (*OE*)	$\kappa_r = -1.211 \times OE + 98.928$
Kappa index (κ)	$\kappa_r = 0.996 \times \kappa - 0.092$

6.2.3.1.1 The Kappa index-based comparisons

Table 6.3 shows the values of κ_r for the geosimulation-multicriteria models (or scenarios). The results indicate that, in general, the local models perform better than the global methods. Specifically, the “Some” quantifier scenario with the contiguity order of 32 is characterized by the highest value of κ_r (53.37%). The worst outcome in terms of accuracy is related to the global model with the “At least one” quantifier (44.66%). The two extreme scenarios (i.e., “All” and “At least one”) result in the most and least accurate outputs, respectively, for both global and local models (see Figure 6.11). Therefore, scenarios with the “All” quantifier tend to be characterized by the lowest allocation disagreement, and the “At least one” models generate highest allocation disagreement.

Table 6.3: The results of geosimulation-multicriteria procedures: the values of κ_r index (Data sources: Appendix C, Table C1)

Linguistic quantifier		All	Most	Many	Half	Some	Few	At least one
Contiguity order	2	52.53	50.01	50.13	50.8	50.86	50.52	48.71
	4	52.58	50.22	50.19	50.31	50.67	50.64	48.98
	8	52.55	50.25	50.27	50.33	50.85	50.75	48.80
	16	52.71	50.48	50.45	50.86	50.99	50.95	48.55
	32	52.96	49.24	49.47	52.98	53.37	53.01	48.52
	Global	53.09	49.51	46.27	53.04	50.65	46.5	44.66

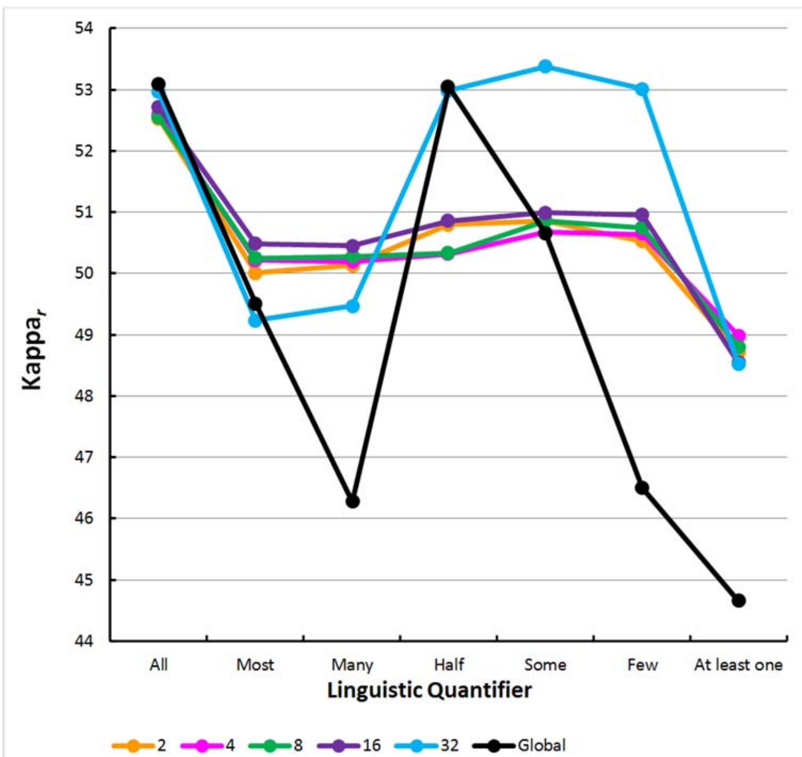


Figure 6.11: The κ_r index of the global and local models

The values of κ and κ_r seem to be relatively low. However, some facts about these two indices need to be considered. First, the results show that all indices are statistically significant since the z_{score} associated with each index is higher than 1.96 and, therefore, the null hypothesis is rejected at p = 0.05. The null hypothesis states that the value of Kappa index can be achieved by a randomly generated pattern. Second, as suggested by Landis and Koch (1977), Kappa values can be classified in three groups: values smaller than 40% show weak agreement; values between 40% and 80% represent moderate agreement; and values higher than 80% show strong agreement. Accordingly, all Kappa indices, in this research, show moderate agreement between the simulated patterns and actual pattern of residential development. Third, in many simulation studies the κ value is calculated based on the agreement between all cells in the simulated and actual land-use patterns (e.g., Henríquez, Azócar, and Romero, 2006; Akın, Sunar, and Berberoğlu, 2015; Gong et al., 2015). This approach results in an overestimation in the values of κ and κ_r, because the initial state of the landscape (base map) should not be regarded in the indices' calculation if the accuracy of the modelling process is examined. Fourth, the number of observations in the cross-tabulation matrix affects the result of the simulation and also the values of κ and κ_r. In this research, the number of observations in the cross-tabulation matrix is large, and therefore, a small allocation disagreement has a large impact on the values of the two indices. In some studies when the number of observations in the matrix is large, any output with the Kappa index higher than 40% is considered acceptable (e.g., Park et al., 2011).

6.2.3.1.1.1 Global versus local models and linguistic quantifiers

Several studies suggest that the results generated by the GIS-based OWA modelling depend on the linguistic quantifiers (e.g., Rinner and Malczewski, 2002; Malczewski, 2006b; Eldrandaly, 2013; Malczewski and Liu, 2014; Cabrera-Barona et al., 2015). To answer the research question about the significance of differences between the global and local models a set of hypotheses is tested.

Hypothesis 1A: *There is no difference between the value of global* κ_r (κ_{rg}) *and the mean value of* κ_r *for local models* (κ_{rl}). This hypothesis is tested for each of the linguistic quantifies: "All", "Most", "Many", "Half", "Some", "Few", and "At least one" (see Table 6.4). The hypothesis is examined using a single sample *t*-test for comparing means (Rogerson, 2015). The test compares the mean of a single sample of scores to a known or hypothetical population mean or a single score; that is, for a given linguistic quantifier the κ_r index obtained by the global model (κ_{rg}) is compared with the mean value of κ_r generated by the local models (κ_{rl}). Table 6.4 indicates that there are significant differences between the global and local models for all linguistic quantifiers but one (the difference is marginally insignificant for the “Some” models). These results confirm findings of previous studies that compare the global and local GIS-based OWA methods (Liu, 2013; Malczewski and Liu, 2014; Cabrera-Barona et al., 2015). For example, Malczewski and Liu (2014) show that the OWA models generate considerably different results depending on the linguistic quantifiers or the sets of order weights. They also demonstrate that there are significant differences between the local and global OWA models’ outcomes for the same quantifiers; e.g., the global “All’ model generates results considerably different from the local “All” model (see also Liu et al., 2014).

Table 6.4: The results of the *t*-test of the difference between the κ_r value obtained by the global model and the mean κ_r value of the local models

Linguistic quantifier	*t*-statistic	*p*-value
All	-5.306	0.003*
Most	2.483	0.034*
Many	22.956	0.000*
Half	-4.013	0.008*
Some	1.374	0.121
Few	10.064	0.001*
At least one	47.974	0.000*

Note: *significant at $p < 0.1$

6.2.3.1.1.2 Global versus local models and neighbourhood sizes

The results of multicriteria analysis depend on the spatial scale at which the analysis is performed. Given a study area (that is, the geographic/operational scale), any change in the neighbourhood size affects the results generated by local models (Malczewski and Rinner, 2015). Specifically, evaluating parcels of land for residential development may result in different overall suitability scores depending on the size of the neighbourhood (e.g., Can, 1992; Lopez Ridaura et al., 2005). In turn, this can influence the accuracy of the results of geosimulation-multicriteria procedures measured by the κ_r index. A set of hypotheses is analysed to verify the significance of the difference between the global and local models with respect to the neighbourhood size parameter.

Hypothesis 1B: *There is no difference between the mean value of* κ_{rg} (*the global model*) *and the mean value of* κ_{rl} (*the local model*). This hypothesis is tested for each of the five contiguity orders (neighbourhood sizes): 2, 4, 8, 16 and 32 (see Table 6.5) using a two samples *t*-test for comparing means; H0: Mean(κ_{rl}) = Mean(κ_{rg}), and Ha: Mean(κ_{rl}) ≠ Mean(κ_{rg}) (Rogerson, 2015); that is, for a given contiguity order the mean value of κ_{rg} for the global-linguistic quantifier models is compared with the mean value of κ_{rl} for the local-linguistic quantifier models.

Table 6.5: The results of the *t*-test for the difference between the mean values of global κ_{rg} and local κ_{rl}

Contiguity order	*t*-statistic	*p*-value
2	1.042	0.159
4	1.052	0.157
8	1.071	0.153
16	1.186	0.130
32	1.491	0.081*

Note: * significant at $p < 0.1$

The results show that the mean value of κ_{rl} for the local models with any of the five neighbourhood definitions is higher than the mean value of κ_{rg}; however, there are insignificant differences between the global and each of the local models except for the model with the contiguity order of 32. The statistics indicate that the accuracy of the results increases with contiguity order. This fact shows the importance of choosing an appropriate neighbourhood size for the local multicriteria analysis (see Malczewski and Rinner, 2015). The size of the neighbourhood can increase until a local becomes a global multicriteria model (McHenry and Rinner, 2016). The optimum contiguity order can be 32 or any other number between 32 and the largest possible contiguity order. Further examination is needed to see if the accuracy of the outputs increases with the neighbourhood size. Section 6.2.4 explains what will happen if the order of the contiguity is 64.

6.2.3.2 Morphological/spatial metrics

Analysing the cross-tabulation matrix gives insufficient information about the pattern accuracy, because it does not reflect the morphological/spatial properties. To analyse the accuracy of spatial pattern, three indices were employed in this study: mean parcel size (*MPS*), aggregation index (*AI*), and average nearest neighbour (*ANN*) (see Ligmann-Zielinska, 2009; Hosseinali, Alesheikh, and Nourian, 2015; Dezhkam et al. 2017). In the following definitions, a parcel of land can be assumed as either a single isolated cell or a set of connected cells that were selected by a model to be converted to residential areas. The mean parcel size is the ratio of the area of the newly developed residential parcels to the total number of parcels (McGarigal et al., 2002):

$$MPS = \frac{\sum_{p=1}^{n} A_p}{n} \tag{6.3}$$

where A_p represents the area of the *p*-th parcel and *n* is the total number of parcels. The *MPS* index can be any values greater than zero. The larger parcels of land in a spatial pattern are, the higher the *MPS* index would be. The weakness of the *MPS* index is related to the fact that it does not consider the shape of the land parcels in the calculations. For example, if the area and number of parcels are identical for two patterns, then the *MPS* indices will be equal as well irrespective of the shape of the parcels.

The aggregation index eliminates the weakness of the *MPS* index by considering conditions of the neighbourhood for each cell (not necessarily parcel) in the calculations. The *AI* index for two patterns with the same area and number of parcels but different parcels' shapes would be different. The aggregation index is the number of similar adjacencies in a class (cells selected for residential development) divided by the highest possible number of similar adjacencies (McGarigal et al., 2002):

$$AI = \frac{g_\mu}{\max\limits_{\mu} g_\mu} 100 \tag{6.4}$$

where g_{μ} is the observed number of similar adjacencies in class μ; and $\max_{\mu} g_{\mu}$ is the maximum possible number of similar adjacencies in class μ. Aggregation index ranges from 0 to 100. When the selected cells are maximally dispersed, *AI* would be equal to 0. As the selected cells become more aggregated, the *AI* index increases and reaches 100 when the pattern is completely aggregated into a single square parcel. Although the *AI* index gives information about the shape of the parcels, it provides no prospect of how far these parcels of land are located with respect to each other.

Average nearest neighbour index (*ANN*) shows how individual land parcels, including isolated cells, are positioned within the landscape. The *MPS* and *AI* indices do not provide any information about the distance between parcels of land. The distance between developed areas can be used to detect urban sprawl or uncontrolled growth. This index is calculated based on the Euclidean distance between parcels of land as follows (Mitchel, 2005):

$$ANN = \frac{\overline{D}_O}{\overline{D}_E} \tag{6.5}$$

where $\overline{D}_O$ is the observed mean between each parcel and its nearest neighbour in a given pattern, and $\overline{D}_E$ is the expected mean between each parcel and its nearest neighbour in a random pattern. These two variables are calculated as below:

$$\overline{D}_O = \frac{\sum_{p=1}^{n} d_p}{n} \tag{6.6}$$

$$\overline{D}_E = \frac{0.5}{\sqrt{n/A}} \tag{6.7}$$

where n is the total number of parcels; and d_p is the Euclidean distance between parcel p and its nearest neighbouring parcel that belongs to the same category. The Euclidean distance is calculated from the geometric centre of parcels. Therefore, the size and shape of the parcels have little to do with this measure. As the spatial pattern of land parcels becomes more compact, the *ANN* index decreases. By compact pattern it means the selected parcels and individual cells are located very close to each other over the landscape. For *ANN* less than 1, selected parcels are located closer to each other than what would be by a random pattern, and for values higher than 1, the pattern is less compact than the random pattern.

6.2.3.2.1 The *MPS*-based comparisons

Table 6.6 shows the mean parcel size of the selected cells for residential development. The results indicate that the global "Most" scenario is characterized by the lowest value of *MPS* (1.432 hectares); and the "At least one" model with the contiguity order of 32 generated the largest average parcel size (2.989 hectares). The *MPS* values decrease from "All" to "Most" scenarios; then, they tend to slightly increase from the "Most" scenarios up to the "Few" scenarios; and the "At least one" models generate the largest size of parcels, on average (see also Figure 6.12). This is a 'hockey stick" curve representing the relation between *MPS* and linguistic quantifiers (or the α parameter of the OWA model). One can argue that the *MPS* depends on the spatial extent (size) of the most suitable land for residential development within a given study area (e.g., Rutledge, 2003). The extent of the most suitable area, in turn, depends on the linguistic quantifiers: the mean size of the parcels has the highest values at the two extreme scenarios, i.e., "All" and "At least one" scenarios; for other quantifiers, the mean size increases as one moves from "Most" to "Few" quantifiers; that is, it increases along with increasing the value of α, or the average area that could be recommended for residential development gets larger and larger (see Jiang and Eastman, 2000, Malczewski 2006b; Cabrera-Barona et al., 2015).

Table 6.6: The results of the geosimulation-multicriteria models: the mean parcel size (*MPS*) for residential development (in ha) (Data sources: Appendix C, Table C1)

Linguistic quantifier		All	Most	Many	Half	Some	Few	At least one
Contiguity order	2	2.003	1.731	1.780	1.830	1.962	1.969	2.608
	4	1.877	1.656	1.665	1.686	1.701	1.714	2.617
	8	1.891	1.650	1.734	1.844	1.879	1.880	2.714
	16	2.010	1.739	1.856	1.872	1.901	1.926	2.646
	32	2.029	1.770	1.850	1.858	1.883	1.929	2.989
	Global	1.783	1.432	1.512	1.513	1.624	1.752	2.569

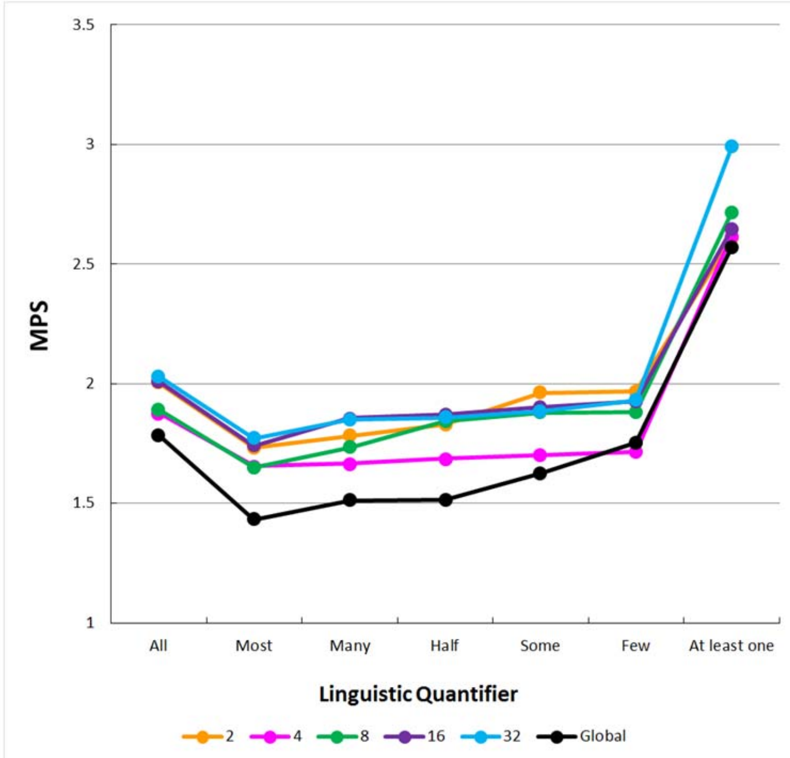


Figure 6.12: The mean parcel size (*MPS*) index of the global and local models

6.2.3.2.1.1 Global versus local models and linguistic quantifiers

Hypothesis 2A: *There is no difference between the value of MPS for the global model* (MPS_g) *and the mean value of MPS for the local models* (MPS_l). This hypothesis is verified for each of the linguistic quantifies: “All”, “Most”, “Many”, “Half”, “Some”, “Few”, and “At least one” (see Table 6.7). It is tested using a single sample *t*-test (see Section 6.2.3.1.1.1). The results provide evidence for significant differences between the global and local models. Indeed, the mean value of MPS_l is significantly greater than MPS_g for all linguistic quantifiers. This can be attributed to the strength of local multicriteria models in identifying more locally extreme high values in the study area (see Section 6.2.2.3.1). When the demand for new residential areas is high, having more extreme high values can be beneficiary to some extent. New residential areas will be formed around those extreme high values, but the shape of the parcels may not be ideally aggregated. One can argue that irrespective of the shape of the residential areas, local multicriteria models generate larger parcels of land as compared to a global model for a given linguistic quantifier.

Table 6.7: The results of the *t*-test for the difference between the global MPS_g value and the mean value of MPS_l for local models

Linguistic quantifier	*t*-statistic	*p*-value
All	5.559	0.003*
Most	11.613	0.000*
Many	7.354	0.001*
Half	9.041	0.001*
Some	5.525	0.003*
Few	2.945	0.021*
At least one	2.053	0.055*

Note: * significant at $p < 0.1$

6.2.3.2.1.2 Global versus local models and neighbourhood sizes

Hypothesis 2B: *There is no difference between the mean value of* MPS_g (*the global models*) *and the mean value of* MPS_l (*the local models*). This hypothesis is tested for each of the five contiguity orders: 2, 4, 8, 16 and 32 (see Table 6.8) using the two sample *t*-tests for comparing means (see Section 6.2.3.1.1.2); that is, for a given contiguity order, the mean value of MPS_g for the global-linguistic quantifier models is compared with the mean value of MPS_l for the local-linguistic quantifier models. The results show that there is an insignificant difference between the two models for three contiguity orders. It indicates that the local multicriteria models with three smaller neighbourhood sizes produce parcels of land that are almost as large as the ones created by the global models. However, in two cases (the contiguity orders of 16 and 32), the mean size of land parcels are significantly higher than the mean value of MPS_g. The results can be contrary to expectations, as one would anticipate that the larger neighbourhood sizes would produce closer results to the global models. However, the results show that the mean size of the parcels increases gradually with the contiguity order until the order of 32, and after that it drops. The results of *MPS* for contiguity order of 64 can be found in Appendix C (see Table C1). Given to the fact that each cell covers 900 square meters, having larger land parcels for residential development can be preferable. In this case, real estate developers can develop big residential areas in the selected land parcels. Considering the *MPS* index of the actual image of 2006 also confirms that having larger parcels for future development is desirable. Although having larger land parcels is preferable for residential areas, the shape of the parcels is of significant importance as well. The *MPS* index does not provide any information about the shape of the land parcels.

Table 6.8: The results of the two sample *t*-test for the difference between the mean values of global MPS_g and local MPS_l

Contiguity order	*t*-statistic	*p*-value
2	1.319	0.106
4	0.530	0.303
8	1.015	0.165
16	1.362	0.099*
32	1.397	0.094*

Note: * significant at $p < 0.1$

6.2.3.2.2 The *AI*-based comparisons

Table 6.9 summarizes the results of geosimulation-multicriteria modelling in terms of the aggregation index, which measures the degree of aggregation of land parcels for residential development. The values of *AI* range from 60.24 (the local scenario with the "Most" quantifier and contiguity order of 2) to 77.55 (the global scenario with the "At least one" quantifier). In general, the "All" and "At least one" models generate more aggregated patterns than the models in between these two scenarios. Figure 6.13 shows that there is a U- or W- shaped relation between the *AI* values and the α parameter (the linguistic quantifies). The most spatially aggregated patterns are obtained using two extreme quantifier scenarios; that is, the "At least one" models (the smallest value of α) and the "All" models (the largest value of α). The remaining scenarios are characterized by the *AI* values that are considerably lower than those for the two extreme scenarios. It is important to note that the "All" and "At least one" scenarios represent a non-compensatory modelling approach, while the remaining scenarios are compensatory (allowing for a trade-off between evaluation criteria) (see Jiang and Eastman, 2000). This finding confirms the results of other studies on the relations between the values of α

and the spatial patterns of land suitability (e.g., Jiang and Eastman, 2000; Malczewski 2006b; Cabrera-Barona et al., 2015).

Table 6.9: The results of the geosimulation-multicriteria models: the aggregation index (*AI*) (Data sources: Appendix C, Table C1)

Linguistic quantifier		All	Most	Many	Half	Some	Few	At least one
Contiguity order	2	66.07	60.24	60.62	60.79	60.65	60.34	73.42
	4	65.83	60.37	60.98	61.29	61.24	61.03	73.45
	8	65.27	60.83	61.06	61.67	61.03	61.36	73.88
	16	66.98	62.40	64.31	63.65	63.80	62.69	75.11
	32	68.92	64.54	64.49	64.60	64.31	62.88	76.88
	Global	71.89	64.62	64.68	66.75	64.87	62.87	77.55

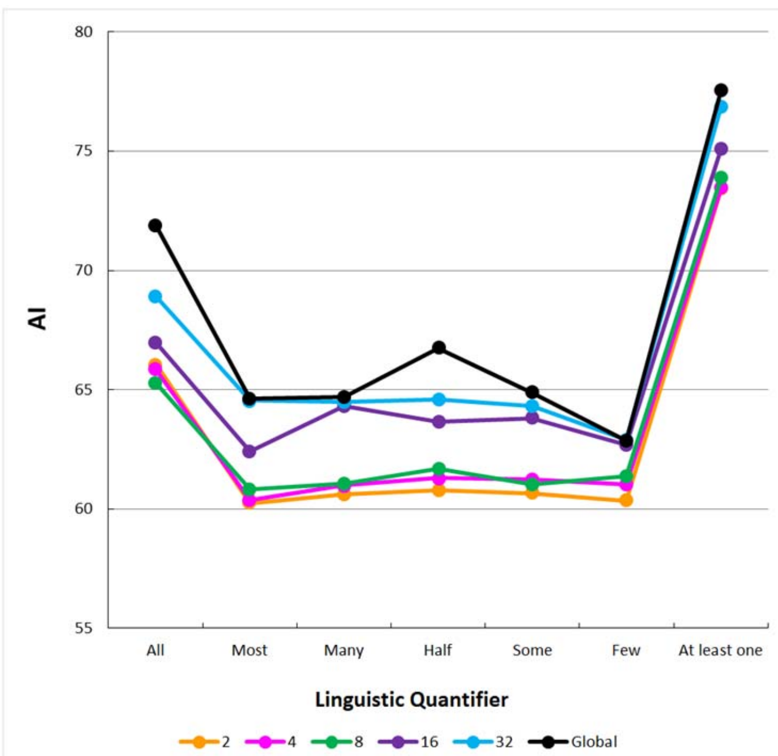


Figure 6.13: The aggregation index (*AI*) of the global and local models

6.2.3.2.2.1 Global versus local models and linguistic quantifiers

Hypothesis 3A: *There is no difference between the value of AI for the global model* (AI_g) *and the mean value of AI for the local model* (AI_l). This hypothesis is tested for each of the linguistic quantifiers: “All”, “Most”, “Many”, “Half”, “Some”, “Few”, and “At least one” (see Table 6.10). It is tested using the single sample *t*-test (see Section 6.2.3.1.1.1). The results provide strong evidence for significant differences between the global and local models. It confirms the findings of previous studies on land suitability analysis with local and global multicriteria models (e.g., Liu, 2013; Malczewski and Liu, 2014), according to which, for a given α value, global multicriteria models generate spatially more aggregated patterns than local multicriteria models. This is due to the fact that in global multicriteria models, each cell is evaluated with respect to all other cells in the area. Therefore, more suitable cells for residential development tend to cluster around global maximum values. However, the spatial pattern created by local multicriteria models would be more disaggregated since each cell is evaluated with respect to its neighbouring cells. Accordingly, the global model generates more aggregated land parcels as compared to the local models, for a given linguistic quantifier.

Table 6.10: The results of the *t*-test for the difference between the global AI_g value and the mean value of AI_l for the local models

Linguistic quantifier	*t*-statistic	*p*-value
All	-8.256	0.001*
Most	-3.622	0.011*
Many	-2.763	0.026*
Half	-5.931	0.002*
Some	-3.482	0.013*
Few	-2.475	0.035*
At least one	-4.557	0.005*

Note: * significant at $p < 0.1$

6.2.3.2.2.2 Global versus local models and neighbourhood sizes

Hypothesis 3B: *There is no difference between the mean value of AI_g (the global models) and the mean value of AI_l (the local models).* This hypothesis is tested for each of the five contiguity orders: 2, 4, 8, 16 and 32 (see Table 6.11) using the two sample *t*-test for comparing means (see Section 6.2.3.1.1.2); that is, for a given contiguity order, the mean value of AI_g for the global-linguistic quantifier models is compared with the mean value of AI_l for the local-linguistic quantifier models. The results show that there are significant differences between the local models with small contiguity orders (i.e., 2, 4, and 8) and the global model; and there are insignificant differences between the local models with large contiguity orders (i.e., 16 and 32) and the global model. It confirms the findings of previous studies, which state that as the size of the neighbourhood increases gradually, a local multicriteria model exhibits more similar behaviour to a global multicriteria model in terms of clustering pattern (see Carter and Rinner, 2014; McHenry and Rinner, 2016). This is due to the fact that local multicriteria models highlight local extremes as opposed to global models (see Section 6.2.2.3). According to Mahiny and Clarke (2012), having larger and more aggregated parcels of land is usually preferred in land-use planning. Therefore, higher values of *AI* are desirable for residential development. Examining the actual value of *AI* for 2006 also confirms that higher values of *AI* will result in more realistic simulated patterns (see Section 6.2.4).

Table 6.11: The results of the *t*-test for the difference between the mean values of global AI_g and local AI_l

Contiguity order	*t*-statistic	*p*-value
2	-1.625	0.065*
4	-1.547	0.074*
8	-1.495	0.081*
16	-0.784	0.225
32	-0.349	0.367

Note: * significant at $p < 0.1$

6.2.3.2.3 The *ANN*-based comparisons

Table 6.12 contains the values of the average nearest neighbour index for the model outputs. The results indicate that the local model with "Many" quantifier and the contiguity order of 32 generates the most 'compact' pattern (the lowest value of *ANN* index); and the "At least one" model with contiguity order of 2 creates the least 'compact' pattern (the highest value of *ANN* index). The "At least one" scenarios are characterized with the most dispersed pattern irrespective of the neighbourhood size. For the remaining scenarios, the values of *ANN* are very close and no distinctive pattern can be found (see Figure 6.14). The higher values of *ANN* can be an indication of urban sprawl in the spatial structure of residential areas. Herold, Goldstein, and Clarke (2003) claimed that having a large distance between individual urban areas is not desirable. The distance between the most suitable areas for residential development, in turn, depends on the linguistic quantifiers in some cases: the average distance between suitable areas reaches the maximum at the "At least one" scenarios; the average distance between individual residential areas looks very similar for other linguistic quantifiers.

Table 6.12: The results of the geosimulation-multicriteria models: the average nearest neighbour index (*ANN*) (Data sources: Appendix C, Table C1)

Linguistic quantifier		All	Most	Many	Half	Some	Few	At least one
Contiguity order	2	0.315	0.319	0.318	0.314	0.315	0.313	0.380
	4	0.318	0.315	0.310	0.321	0.318	0.317	0.356
	8	0.314	0.318	0.317	0.315	0.319	0.324	0.364
	16	0.307	0.313	0.310	0.312	0.311	0.315	0.370
	32	0.308	0.310	0.305	0.306	0.310	0.309	0.366
	Global	0.316	0.318	0.318	0.320	0.322	0.323	0.374

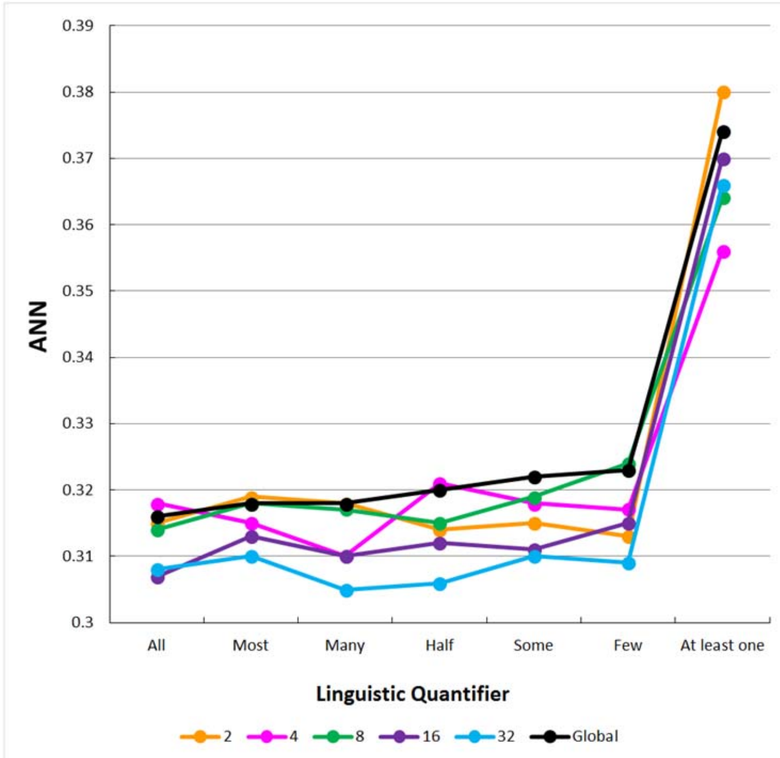


Figure 6.14: The average nearest neighbour (*ANN*) index of the global and local models

6.2.3.2.3.1 Global versus local models and linguistic quantifiers

Hypothesis 4A: *There is no difference between the value of ANN for the global model* (ANN_g) *and the mean value of ANN for the local models* (ANN_l). This hypothesis is tested for each of the linguistic quantifies: “All”, “Most”, “Many”, “Half”, “Some”, “Few”, and “At least one” (see Table 6.13) using the single sample *t*-test (see Section 6.2.3.1.1.1). The results show that the value of ANN_g is significantly greater than the mean value of ANN_l for all linguistic quantifiers. This means that the average distance between residential areas is significantly larger when a global multicriteria model is applied for a given α value. This index depends on the structure of the urban area. Since the central section of the city has already been developed, most development should take place in peripheral districts. When a global model is applied, global high-suitable cells (global extreme values) are selected for new residential development. These highly-suitable cells tend to cluster around absolute maximum values (see Section 6.2.2.3). The results indicate that global high-suitable cells are located relatively far from each other as compared to local high-suitable cells. Accordingly, local models generate a pattern that is significantly more compact (less sprawl) than global models for a given linguistic quantifier.

Table 6.13: The results of the *t*-test for the difference between the global ANN_g value and the mean value of ANN_l for local models

Linguistic quantifier	*t*-statistic	*p*-value
All	-1.705	0.082*
Most	-1.826	0.071*
Many	-2.470	0.035*
Half	-2.644	0.029*
Some	-4.098	0.008*
Few	-2.982	0.021*
At least one	-1.731	0.080*

Note: * significant at $p < 0.1$

6.2.3.2.3.2 Global versus local models and neighbourhood sizes

Hypothesis 4B: *There is no difference between the mean value of* ANN_g (*the global model*) *and the mean value of* ANN_l (*the local model*). This hypothesis is tested for each of the five contiguity orders: 2, 4, 8, 16 and 32 (see Table 6.14) using the two sample *t*-test for comparing means (see Section 6.2.3.1.1.2); that is, for a given contiguity order, the mean value of ANN_g for the global-linguistic quantifier models is compared to the mean value of ANN_l for the local-linguistic quantifier models. Having a large open space between individual residential areas can be taken as an indication of undesirable urban sprawl (see Herold, Goldstein, and Clarke, 2003). The results indicate that although the average distance between individual urban areas is larger when global multicriteria models are used, there is insignificant difference between local and global methods with respect to the *ANN* index. At first look, it seems that this finding is contrary to the results of the *AI* index. However, the focus of *AI* index is on the level of aggregation of cells in a single parcel of residential areas, while the focus of *ANN* is on how far each residential parcel is located from its closest residential parcel on the landscape (spatial distribution of parcels) (see Section 6.2.3.2).

Table 6.14: The results of the *t*-test for the difference between the mean values of global ANN_g and local ANN_l

Contiguity order	*t*-statistic	*p*-value
2	-0.201	0.422
4	-0.528	0.304
8	-0.277	0.394
16	-0.658	0.262
32	-0.962	0.178

6.2.4 Comparing scenarios and actual patterns

Table 6.15 gives the values of *MPS*, *AI*, and *ANN* for the best and worst models as well as the observed values based on the actual pattern of residential development. Higher values for *MPS* and *AI* indices are desirable (see Sections 6.2.3.2.1 and 6.2.3.2.2), while lower values for *ANN* index is preferable (see Section 6.2.3.2.3). Comparing the results based on the mean size of parcels of land confirms that the actual pattern has larger residential parcels on average. Comparing the values of *AI* index for the actual pattern and the model outputs reveals that the actual pattern is more aggregated than simulated patterns. Interpreting the results with respect to the *ANN* index is more complicated. According to the literature, having large open spaces between residential areas is not desirable (see Section 6.2.3.2.3); if the models are compared based on this concept, then the model with the lowest *ANN* value is the best model; however, the value of *ANN* for the worst model (the highest *ANN* value) is even considerably lower than what was observed from the actual pattern. This discrepancy can be a result of government planning to leave more open lands between residential areas, or it can be the result of uncontrolled development between 1996 and 2006 (see Section 3.3.1).

Table 6.15: Evaluation metrics: the best models and observed residential development

Metric	Scenario (model)	Best model	Value of metrics	Worst model	Value of metrics	Observed residential development 1996-2006
Mean parcel size (*MPS*)	Linguistic quantifier	At least one	2.989	Most	1.432	6.555
	Contiguity order	32		Global		
Aggregation index (*AI*)	Linguistic quantifier	At least one	77.55	Most	60.240	87.510
	Contiguity order	Global		2		
Average nearest neighbour (*ANN*)	Linguistic quantifier	Many	0.305	At least one	0.380	0.471
	Contiguity order	32		2		

To summarize, Tables 6.4, 6.7, 6.10, and 6.13 suggest that local models can produce a result with less allocation disagreement (higher κ_r index) and more desirable morphological/spatial properties for a given linguistic quantifier (except for the aggregation). By considering the neighbourhood size, the local models with the contiguity order of 32 produced the most accurate and desirable results except for aggregation property in which the global models performed insignificantly better (see Tables 6.5, 6.8, 6.11, and 6.14). Another point that deserves attention is that global models can be seen as a special case of local models when the contiguity order (neighbourhood size) is so large that it covers the whole study area. In the case of Tehran, if the contiguity order of 778 is applied, local models will be reduced to global ones; it is the largest contiguity order needed to cover all of the study area and it is operationalized when the central cell is examined. Accordingly, the geosimulation-multicriteria model was executed for six contiguity orders (i.e., 2, 4, 8, 16, 32, and 778). The contiguity order of 32 performed better than all other scenarios. However, to get a better approximation about the best neighbourhood size, the procedure was executed for the contiguity order of 64 with 7 linguistic quantifiers (64 was chosen because it is the next number in the sequence of powers of 2 after 32). The results of the analysis can be found in Appendix C (see Table C1). The results indicate that the scenarios with the contiguity order of 64 generate less accurate results with less desirable morphological/spatial characteristics, except for the *AI* index that is insignificantly greater. Consequently, the scenarios with the contiguity order of 32 generate the ‘best’ simulation outputs.

Chapter 7

7 Conclusions

7.1 Summary

There were three main objectives of this research. First, a framework/model for simulating residential land development in the City of Tehran was developed. The framework integrated local multicriteria models into geosimulation procedures. Specifically, the local form of the ordered weighted averaging (OWA) model was used as a method for modelling agents' behaviours (preferences) in the geosimulation procedure. Second, the framework was tested in the context of residential land development in the City of Tehran between 1996 and 2006. The focus of the empirical research was on identifying the spatial patterns of land suitability for residential development by taking into account the preferences of three groups of actors (agents): households, developers, and local authorities. Third, a comparative analysis of the results of the geosimulation-multicriteria models was performed. Forty-two scenarios (global and local geosimulation-multicriteria models) of residential development in Tehran were defined and then the results obtained by the scenarios were evaluated and examined. The output of each geosimulation-multicriteria model was compared to the results of other models and to the actual pattern of land-use in the city. The analysis focused on comparing the results of the local and global geosimulation-multicriteria models with respect to the linguistic quantifiers and the neighbourhood sizes employed for the local multicriteria modelling.

Two types of measures were used in the comparative analysis. First, five accuracy (cross-tabulation matrix) measurements (i.e., overall accuracy, error of commission, error of omission, κ index, and κ_r index) were employed by focusing on the results obtained using the κ_r index. Second, three spatial metrics (i.e., mean parcel size (*MPS*), aggregation index (*AI*), and average nearest neighbour (*ANN*)) were used to compare the morphological properties of the residential land-use patterns. The results showed that, in general, the local geosimulation-multicriteria models performed better than the global methods with respect to the cross-tabulation matrix measurements. The difference between the two models was significant in several cases. The local geosimulation-multicriteria model with the contiguity order of 32 produced the most accurate results (smallest allocation disagreement). When the results were compared using morphological/spatial metrics, the local model with the contiguity order of 32 generated the most desirable results in terms of *MPS*. Moreover, the results showed that there is a significant difference between the local and global models for small neighbourhood sizes with respect to the *AI* index. Furthermore, if the models were compared based on the *ANN* index, no significant differences can be identified between the local and global forms of the geosimulation-multicriteria models. By juxtaposing the outputs of the scenarios with the actual residential pattern of 2006, it was concluded that the local multicriteria analysis with the contiguity order of 32 generated the closest pattern to the real-world situation.

7.2 Implications

The results of this research make a substantial contribution to Geographic Information Science and spatial analysis by developing a new approach to the geosimulation-multicriteria analysis. Although many studies applied multicriteria methods to examine land-use/cover changes and urban development, there has been no research dealing with the integration of local multicriteria modelling and geosimulation procedures. Furthermore, there is a very limited volume of empirical research about the differences between the local and global multicriteria analysis. This study represents a unique effort to 'localize' the conventional, global OWA method and to demonstrate the differences between the global and local methods empirically. Although this research focuses on applying geosimulation-multicriteria methods to analyse residential land development, the proposed framework/model is generic enough to accommodate a wide range of decision/evaluation situations in urban and regional planning.

Urban planners and local authorities can derive substantial benefits from the results of geosimulation-multicriteria modelling. The municipality of Tehran plays a key role in the future land-use pattern by enforcing comprehensive land-use plans, approval processes for development applications, zoning policies, and designing public facilities and transportation networks. A significant loss of farmland/orchard in Tehran over the last three decades shows that government measures have been insufficient to counter the environmental impacts of land-use changes. Low percentages of open lands and farmlands/orchards in Tehran cause serious concerns about the environmental conditions of the city in the near future if the current trend of land-use changes continues. As population growth puts pressure on land resources, preparing a judicious land-use plan by the municipality is becoming increasingly crucial. In order to make a good plan for the future and minimize negative impacts on the environment, the trajectory of past land-use changes needs to be tracked. The geosimulation-multicriteria modelling can help urban planners and decision-makers to examine how location decisions of different agents (interest groups) contributed to the existing land-use pattern. The approach can also provide urban planners and decision-makers with a decision support tool for representing the future outcomes of different scenarios. Based on the result of scenarios, one can establish some policies and regulations to control future residential growth.

7.3 Limitations and outlooks

The research focused on descriptive geosimulation-multicriteria modelling; that is, the framework was used to simulate past residential developments. However, the proposed approach can be applied as a predictive tool to forecast the future structure of urban areas by using the most recent and accurate land-use image as the base map. Geosimulation-multicriteria modelling can also be extended to serve as a prescriptive tool to provide users with advice on what action should be taken to 'optimize' land-use pattern.

Since the geosimulation-multicriteria modelling focused on the two-dimensional development of the study area, the vertical growth is ignored in the modelling process. In the future, the vertical structure of the residential areas can be considered as well. For example, an undeveloped parcel of land that is highly suitable for residential development from the perspective of different type of agents is more likely to be converted to a high-rise building.

There are also some limitations and possibilities for extending the geosimulation-multicriteria procedure with respect to the structure and behaviour of the agents participating in the process of residential land development. The limitations are related to the assumptions behind the geosimulation-multicriteria model, including: (i) household structure was assumed to be the same across the modelling process, (ii) agents had complete information about the residential land suitability/site selection problem, (iii) one agent represented all real estate developers operating within the study area, which implies that there is no competition among real estate developers, and (iv) preferences of different groups of agents remained the same over a given time period. By relaxing these assumptions one can extend the geosimulation-multicriteria model to improve the accuracy of the results and gain new insights into the process of residential land development.

References

Abbasi-Shavazi, M. J., McDonald, P., and Hosseini-Chavoshi, M. (2009). *National and Provincial Level Fertility Trends in Iran, 1972–2006*. Berlin: Springer.

Akın, A., Sunar, F., and Berberoğlu, S. (2015). Urban change analysis and future growth of Istanbul. *Environmental Monitoring and Assessment*, 187(8), 1-15.

Allen, P. M. (1997). Cities and regions as evolutionary, complex systems. *Geographical Systems*, 4, 103-130.

Alonso, W. (1964). *Location and Land Use: Towards a General Theory of Land Rent*. Cambridge: Harvard University Press.

Anderson, J. R., Hardy, E. E., Roach, J. T., and Witmer, R. E. (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. Washington: US Government Printing Office.

Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1), 3-25.

Anselin, L., and Getis, A. (2010). Spatial statistical analysis and geographic information systems. In *Perspectives on Spatial Data Analysis* (pp. 35-47). Berlin: Springer.

Arthur, W. B., Durlauf, S. N., and Lane, D. A. (1997). *The Economy as an Evolving Complex System II*. Reading, MA: Addison-Wesley.

Barasa, B., Majaliwa, J. G. M., Lwasa, S., Obando, J., and Bamutaze, Y. (2011). Magnitude and transition potential of land-use/cover changes in the trans-boundary river Sio catchment using remote sensing and GIS. *Annals of GIS*, 17(1), 73-80.

Basu, N., and Pryor, R. J. (1997). Growing a market economy. *Sandia Report SAND97-2093 UC-905. Sandia National Laboratories*. Albuquerque, NM.

Batty, M., Couclelis, H., and Eichen, M. (1997). Urban systems as cellular automata. *Environment and Planning B: Planning and Design*, 24(2), 159-164.

Beinat, E. (1997). *Value Functions for Environmental Management*. Dordrecht: Kluwer Academic Publishers.

Belton, V., and Stewart, T. J. (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*. Boston: Kluwer Academic Publishers.

Ben-Arieh, D. (2005). Sensitivity of multi-criteria decision making to linguistic quantifiers and aggregation means. *Computers and Industrial Engineering*, 48, 289-309.

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

Benenson, I. (1999). Modeling population dynamics in the city: from a regional to a multi-agent approach. *Discrete Dynamics in Nature and Society*, 3(2-3), 149-170.

Benenson, I., Omer, I., and Hatna, E. (2002). Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel Aviv. *Environment and Planning B*, 29(4), 491-512.

Benenson, I., and Torrens, P. M. (2004a). *Geosimulation: Automata-based Modeling of Urban Phenomena*. New York: John Wiley and Sons.

Benenson, I., and Torrens, P. M. (2004b). Geosimulation: object-based modeling of urban phenomena. *Computers, Environment and Urban Systems*, 28(1-2), 1-8.

Benenson, I., Aronovich, S., and Noam, S. (2005). Let's talk objects: generic methodology for urban high-resolution simulation. *Computers, Environment and Urban Systems*, 29(4), 425-453.

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

Bertaud, A. (2003). *Tehran spatial structure: constraints and opportunities for future development*. Tehran: Ministry of Housing and Urban Development.

Bone, C., Dragicevic, S., and White, R. (2011). Modeling-in-the-middle: bridging the gap between agent-based modeling and multi-objective decision-making for land use change. *International Journal of Geographical Information Science*, 25(5), 717-737.

Boroushaki, S., and Malczewski, J. (2008). Implementing an extension of the analytical hierarchy process using ordered weighted averaging operators with fuzzy quantifiers in ArcGIS. *Computers and Geosciences*, 34(4), 399-410.

Bozkaya, A. G., Balcik, F. B., Goksel, C., and Esbah, H. (2015). Forecasting land-cover growth using remotely sensed data: a case study of the Igneada protection area in Turkey. *Environmental Monitoring and Assessment*, 187(3), 59-77.

Bryman, A. (2016). *Social Research Methods*. Oxford: Oxford University Press.

Burke, R. (2003). *Getting to Know ArcObjects: Programming ArcGIS with VBA*. Redlands, CA: ESRI Press.

Burrough, P. A. (1986). *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford: Clarendon Press.

Cabrera-Barona, P., Murphy, T., Kienberger, S., and Blaschke, T. (2015). A multi-criteria spatial deprivation index to support health inequality analyses. *International Journal of Health Geographics*, 14, 1-14.

Can, A. (1992). Residential quality assessment: alternative approaches using GIS. *The Annals of Regional Science*, 23(1), 97–110.

Carter, B., and Rinner, C. (2014). Locally weighted linear combination in a vector geographic information system. *Journal of Geographical Systems*, 16(3), 343-361.

Cecchini, A., and Rizzi, P. (2001). The reasons why cellular automata are a useful tool in the working-kit for the new millennium urban planner in governing the territory. *In CUPUM 2001 Proceedings*, Honolulu.

Chen, Y., Li, X., Liu, X., and Liu, Y. (2010). An agent-based model for optimal land allocation (AgentLA) with a contiguity constraint. *International Journal of Geographical Information Science*, 24(8), 1269-1288.

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(2), 167-194.

Chowdhury, P. R., and Maithani, S. (2014). Modelling urban growth in the Indo-Gangetic plain using nighttime OLS data and cellular automata. *International Journal of Applied Earth Observation and Geoinformation*, 33, 155-165.

Clark, W. W. A., and Dieleman, F. M. (1996). *Households and Housing: Choice and Outcomes in the Housing Market*. Rutgers University, NJ: CUPR Press.

Clarke, K. C., Hoppen, S., and Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design*, 24(2), 247-261.

Clarke, K. C., Olsen, G., and Brass, J. A. (1993). Refining a cellular automaton model of wildfire propagation and extinction. In *Proceedings of the Second International Conference on the Integration of Geographic Information Systems and Environmental Modeling*. Breckenridge, CO.

Clarke, K. C., Riggan, P. J., and Brass, J. A. (1995). A cellular automaton model of wildfire propagation and extinction. *Photogrammetric Engineering and Remote Sensing*, 60, 1355-1367.

Clifford, N. (2008). Models in geography revisited. *Geoforum*, 39(2), 675-686.

Collier, N., Howe, T., and North, M. (2003). Onward and upward: the transition to Repast 2.0. In *Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference*, Pittsburgh, PA.

Comino, E., Bottero, M., Pomarico, S., and Rosso, M. (2014). Exploring the environmental value of ecosystem services for a river basin through a spatial multicriteria analysis. *Land Use Policy*, 36, 381-395.

Congalton, R. G. (1981). *The use of discrete multivariate analysis for the assessment of Landsat classification accuracy*. Master dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Congalton, R. G., and Green, K. (1999). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, FL: Lewis Publishers.

Congalton, R. G., and Mead, R. A. (1983). A quantitative method to test for consistency and correctness in photointerpretation. *Photogrammetric Engineering and Remote Sensing*, 49(1), 69-74.

Congalton, R. G., Oderwald, R. G., and Mead, R. A. (1983). Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing*, 49(12), 1671-1678.

Costanza, R., Sklar, F. H., and White, M. L. (1990). Modeling coastal landscape dynamics. *BioScience*, 91-107.

Cox, W. (2015). Demographia world urban areas. *11th Annual Edition ed. St. Louis: demographia. Retrieved from http://www.demographia.com/db-worldua.pdf*. Date of access, Oct 2015.

de Noronha Vaz, E., Nijkamp, P., Painho, M., and Caetano, M. (2012). A multi-scenario forecast of urban change: a study on urban growth in the Algarve. *Landscape and Urban Planning*, 104(2), 201-211.

Deadman, P., and Gimblett, R. H. (1994). A role for goal-oriented autonomous agents in modeling people-environment interactions in forest recreation. *Mathematical and Computer Modelling*, 20(8), 121-133.

Dezhkam, S., Jabbarian Amiri, B., Darvishsefat, A. A., and Sakieh, Y. (2017). Performance evaluation of land change simulation models using landscape metrics. *Geocarto International*, 32(6), 655-677.

DigitalGlob Founadation. (2017). *Home Page*. Retrieved from http://www.digitalglobefoundation.org.

Dyer, R. F., and Forman, E. H. (1992). Group decision support with the analytic hierarchy process. *Decision Support Systems*, 8(2), 99-124.

Eastman, J. R. Kyem P. A. K., Toledano, J., and Jin, W. (1993). *GIS and Decision Making*. Geneva: UNITAR European Office.

Eastman, J. (1997). *Idrisi for Windows, Version 2.0: tutorial exercises, Clark Labs for Cartographic Technology and Geographic Analysis*. Clark University, Worcester, MA.

Einhorn, H. J., and Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*, 13(2), 171-192.

Eldrandaly, K. A. (2013). Exploring multi-criteria decision strategies in GIS with linguistic quantifiers: an extension of the analytical network process using ordered weighted averaging operators. *International Journal of Geographical Information Science*, 27(12), 2455–2482.

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Generative Social Science: Studies in Agent-Based Computational Modeling*, 4(5), 4-46.

Erbek, F. S., Özkan, C., and Taberner, M. (2004). Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *International Journal of Remote Sensing*, 25(9), 1733-1748.

ERDAS. (1999). *ERDAS Field Guide*. Atlanta: ERDAS Inc.

Fagiolo, G., Windrum, P., and Moneta, A. (2006). *Empirical validation of agent-based models: A critical survey* (No. 2006/14). LEM Working Paper Series, Laboratory of Economics and Managements, Sant'Anna School of Advanced Studies, Italy.

Fischer, G. W. (1995). Range sensitivity of attribute weights in multiattribute value models. *Organizational Behavior and Human Decision Processes*, 62(3), 252-266.

Fitzsimons, J., Pearson, C. J., Lawson, C., and Hill, M. J. (2012). Evaluation of land-use planning in greenbelts based on intrinsic characteristics and stakeholder values. *Landscape and Urban Planning*, 106(1), 23-34.

Forrester, J. W. (1969). *Urban Dynamics*. Cambridge: MIT Press.

Fotheringham, A. S., and P. Rogerson. (1994). *Spatial Analysis and GIS*. London: Taylor and Francis.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2000). *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Publication Sage.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2003). *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. New York: John Wiley and Sons.

Frazier, C., and Kockelman, K. (2005). Spatial econometric models for panel data: incorporating spatial and temporal data. *Transportation Research Record: Journal of the Transportation Research Board*, (1902), 80-90.

Fullér, R., and Majlender, P. (2003). On obtaining minimal variability OWA operator weights. *Fuzzy Sets and Systems*, 136(2), 203-215.

Gärling, T., and Friman, M. (2002). A psychological conceptualization of residential choice and satisfaction. In *Residential Environments: Choice, Satisfaction, and Behavior* (pp. 55-80). London: Bergin and Garvey.

Gaube, V., and Remesch, A. (2013). Impact of urban planning on household's residential decisions: an agent-based simulation model for Vienna. *Environmental Modelling and Software*, 45, 92-103.

Ghavami, S. M., and Taleai, M. (2016). Towards a conceptual multi-agent-based framework to simulate the spatial group decision-making process. *Journal of Geographical Systems*, 1-24.

Gilbert, N. (2008). *Agent-based Models (Quantitative Applications in the Social Sciences).* London: Sage Publications.

Gong, W., Yuan, L., Fan, W., and Stott, P. (2015). Analysis and simulation of land use spatial pattern in Harbin prefecture based on trajectories and cellular automata—Markov modelling. *International Journal of Applied Earth Observation and Geoinformation*, 34, 207-216.

Habibi, N. (2010). The impact of sanctions on Iran-GCC economic relations. *Middle East Brief*, 45, 1.

Habibpour, G. K., and Ghaffary, G. (2011). A study on the causes of rising marriage age among girls. *Women in Development and Politics* (in Persian), 9(1), 7-34.

Hansen, H. S. (2010). Modelling the future coastal zone urban development as implied by the IPCC SRES and assessing the impact from sea level rise. *Landscape and Urban Planning*, 98(3), 141-149.

Hansen, H. S. (2012). Empirically derived neighbourhood rules for urban land-use modelling. *Environment and Planning B: Planning and Design*, 39(2), 213-228.

Henríquez, C., Azócar, G., and Romero, H. (2006). Monitoring and modeling the urban growth of two mid-sized Chilean cities. *Habitat International*, 30(4), 945-964.

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(3), 286-302.

Hiebeler, D. (1994). The swarm simulation system and individual-based modeling. *Advanced Technology for Natural Resource Management*, 20-26.

Hinshaw, M., and Allott, K. (1972). Environmental preferences of future housing consumers. *Journal of the American Institute of Planners*, 38(2), 102-107.

Hobbs, B.F., Meier, P., 2012. *Energy Decision and the Environment: A Guide to the Use of Multicriteria Methods*. New York: Springer.

Hosseinali, F., Alesheikh, A. A., and Nourian, F. (2013). Agent-based modeling of urban land-use development, case study: simulating future scenarios of Qazvin city. *Cities*, 31, 105-113.

Hosseinali, F., Alesheikh, A. A., and Nourian, F. (2015). Assessing urban land-use development: developing an agent-based model. *KSCE Journal of Civil Engineering*, 19(1), 285-295.

Huu Phe, H., and Wakely, P. (2000). Status, quality and the other trade-off: Towards a new theory of urban residential location. *Urban studies*, 37(1), 7-35.

Hyandye, C., and Martz, L. W. (2017). A Markovian and cellular automata land-use change predictive model of the Usangu Catchment. *International Journal of Remote Sensing*, 38(1), 64-81.

Iranian Ministry of Roads and Urban Development. (2017). Retrieved from: http://hmi.mrud.ir/sabaa.

Irwin, E. G., Jayaprakash, C., and Munroe, D. K. (2009). Towards a comprehensive framework for modeling urban spatial dynamics. *Landscape Ecology*, 24(9), 1223-1236.

Itami, R., and Gimblett, H. (2001). Intelligent recreation agents in a virtual GIS world. *Complexity International*, 8, 1-14.

Jiang, H., Eastman, J. R. (2000). Application of fuzzy measures in multi-criteria evaluation in GIS. *International Journal of Geographical Information Science*, 14, 173–184.

Jiao, J., and Boerboom, L. (2006). Transition rule elicitation methods for urban cellular automata models. *Innovations in Design and Decision Support Systems in Architecture and Urban Planning*, 53-68.

Jokar Arsanjani, J. (2012). *Dynamic Land-use/cover Change Simulation: Geosimulation and Multi Agent-based Modelling*. Springer Theses. Berlin: Springer.

Jokar Arsanjani, J., Helbich, M., and de Noronha Vaz, E. (2013). Spatiotemporal simulation of urban growth patterns using agent-based modeling: the case of Tehran. *Cities*, 32, 33-42.

Keeney, and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley and Sons.

Keeney, R. L. (1992). *Value-focused Thinking: A Path to Creative Decision-making*. Cambridge: Harvard University Press.

Keshtkar, H., and Voigt, W. (2016). A spatiotemporal analysis of landscape change using an integrated Markov chain and cellular automata models. *Modeling Earth Systems and Environment*, 2(1), 10.

Kohler, T. A., and Gummerman, G. J. (2000). *Dynamics of Human and Primate Societies: Agent-based Modeling of Social and Spatial Processes*. Oxford: Oxford University Press.

Lambin, E. F. (1997). Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography*, 21(3), 375-393.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., and George, P. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, 11(4), 261-269.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Lansing, J. B., and Morgan, J. N. (1955). Consumer finances over the life cycle. *Consumer Behavior*, 2(4), 36-50.

Lau, K. H., and Kam, B. H. (2005). A cellular automata model for urban land-use simulation. *Environment and Planning B: Planning and Design*, 32(2), 247-263.

Lee Jr, D. B. (1973). Requiem for large-scale models. *Journal of the American Institute of Planners*, 39(3), 163-178.

Leverett, F., and Leverett, H. M. (2013). *Going to Tehran: Why the United States Must Come to Terms with the Islamic Republic of Iran*. New York: Metropolitan Books.

Li, X., and Yeh, A. G. O. (2001). Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment and Planning A*, 33(8), 1445-1462.

Li, X., and Yeh, A. G. O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), 323-343.

Li, X., and Liu, X. (2007). Defining agents' behaviors to simulate complex residential development using multicriteria evaluation. *Journal of Environmental Management*, 85(4), 1063-1075.

Li, X., Chen, Y., Liu, X., Li, D., and He, J. (2011). Concepts, methodologies, and tools of an integrated geographical simulation and optimization system. *International Journal of Geographical Information Science*, 25(4), 633-655.

Li, W., Wu, C., and Zang, S. (2014). Modeling urban land use conversion of Daqing City, China: a comparative analysis of "top-down" and "bottom-up" approaches. *Stochastic Environmental Research and Risk Assessment*, 28(4), 817-828.

Li, C., and Zhao, J. (2017). Assessment of future urban growth impact on landscape pattern using cellular automata model: a case study of Xuzhou city, China. *Journal of Environmental Engineering and Landscape Management*, 25(1), 23-38.

Liang, S. (2004). *Quantitative Remote Sensing of Land Surfaces*. New York: John Wiley and Sons.

Ligmann-Zielinska, A. (2009). The impact of risk-taking attitudes on a land use pattern: an agent based model of residential development. *Journal of Land Use Science*, 4(4), 215 — 232.

Ligmann-Zielinska, A., and Jankowski, P. (2010). Exploring normative scenarios of land use development decisions with an agent-based simulation laboratory. *Computers, Environment and Urban Systems*, 34(5), 409-423.

Ligmann-Zielinska, A., and Sun, L. (2010). Applying time-dependent variance-based global sensitivity analysis to represent the dynamics of an agent-based model of land use change. *International Journal of Geographical Information Science*, 24(12), 1829-1850.

Ligtenberg, A., Bregt, A. K., and Van Lammeren, R. (2001). Multi-actor-based land use modelling: spatial planning using agents. *Landscape and Urban Planning*, 56(1), 21-33.

Ligtenberg, A., Wachowicz, M., Bregt, A. K., Beulens, A., and Kettenis, D. L. (2004). A design and application of a multi-agent system for simulation of multi-actor spatial planning. *Journal of Environmental Management*, 72(1), 43-55.

Liu, Y. (2008). *Modelling Urban Development with Geographical Information Systems and Cellular Automata*. New York: CRC Press.

Liu, X. (2013). *GIS-based Local Ordered Weighted Averaging: A Case Study in London, Ontario*. Master dissertation, Department of Geography, The University of Western Ontario, London, ON.

Liu, Y., Lv, X., Qin, X., Guo, H., Yu, Y., Wang, J., and Mao, G. (2007). An integrated GIS-based analysis system for land-use management of lake areas in urban fringe. *Landscape and Urban Planning*, 82(4), 233-246.

Liu, R., Zhang, K., Zhang, Z., and Borthwick, A. G. (2014). Land-use suitability analysis for urban development in Beijing. *Journal of Environmental Management*, 145, 170-179.

Lloyd, C. D. (2010). *Local Models for Spatial Analysis*. New York: CRC Press.

Loibl, W., and Toetzer, T. (2003). Modeling growth and densification processes in suburban regions—simulation of landscape transition with spatial agents. *Environmental Modelling and Software*, 18(6), 553-563.

Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W. (2001). *Geographic Information System and Science*. New York: John Wiley and Sons.

Lopez Ridaura, S., van Keulen, H., van Ittersum, M.K., Leffelaar, P.A. (2005). Multiscale methodological framework to derive criteria and indicators for sustainability evaluation of peasant natural resource management systems. *Environment, Development and Sustainability*, 7(1), 51–69.

Lotfi, V., Stewart, T. J., and Zionts, S. (1992). An aspiration-level interactive model for multiple criteria decision making. *Computers and Operations Research*, 19(7), 671-681.

Lowry, I. S. (1964). *A Model of Metropolis*. Santa Monica, CA: Rand Corporation.

Madanipour, A. (2006). Urban planning and development in Tehran. *Cities*, 23(6), 433-438.

Mahdavi, M. S., Kaldi, A., and Jamand, B. (2016). Examining the social factors behind young people reluctance to marriage in urban areas (case study: Tehran). *Urban Studies* (in Persian), 6(19), 33-60.

Mahiny, A. S., and Clarke, K. C. (2012). Guiding SLEUTH land-use/land-cover change modeling using multicriteria evaluation: towards dynamic sustainable land-use planning. *Environment and Planning B: Planning and Design*, 39(5), 925-944.

Makropoulos, C., Argyrou, E., Memon, F., and Butler, D. (2007). A suitability evaluation tool for siting wastewater treatment facilities in new urban developments. *Urban Water Journal*, 4(2), 61-78.

Makropoulos, C., and Butler, D. (2006). Spatial ordered weighted averaging: incorporating spatially variable attitude towards risk in spatial multi-criteria decision-making. *Environmental Modelling and Software*, 21(1), 69-84.

Malczewski, J. (1999). *GIS and Multicriteria Decision Analysis*. New York: John Wiley and Sons.

Malczewski, J. (2000). On the use of weighted linear combination method in GIS: common and best practice approaches. *Transactions in GIS*, 4(1), 5-22.

Malczewski, J. (2004). GIS-based land-use suitability analysis: a critical overview. *Progress in Planning*, 62(1), 3-65.

Malczewski, J. (2006a). GIS-based multicriteria decision analysis: a survey of the literature. *International Journal of Geographical Information Science*, 20(7), 703-726.

Malczewski, J. (2006b). Ordered weighted averaging with fuzzy quantifiers: GIS-based multicriteria evaluation for land-use suitability analysis. *International Journal of Applied Earth Observation and Geoinformation*, 8(4), 270-277.

Malczewski, J. (2011). Local weighted linear combination. *Transactions in GIS*, 15(4), 439-455.

Malczewski, J., and Liu, X. (2014). Local ordered weighted averaging in GIS-based multicriteria analysis. *Annals of GIS*, 20(2), 117-129.

Malczewski, J., and Rinner, C. (2005). Exploring multicriteria decision strategies in GIS with linguistic quantifiers: a case study of residential quality evaluation. *Journal of Geographical Systems*, 7(2), 249-268.

Malczewski, J., and Rinner, C. (2015). *Multicriteria Decision Analysis in Geographic Information Science*. Berlin: Springer.

Manganelli, B., Di Palma, F., Amato, F., Nolè, G., and Murgante, B. (2016). The effects of socio-economic variables in urban growth simulations. *Procedia-Social and Behavioral Sciences*, 223, 371-378.

Manson, S. M. (2005). Agent-based modeling and genetic programming for modeling land change in the Southern Yucatan Peninsular Region of Mexico. *Agriculture, Ecosystems and Environment*, 111(1), 47-62.

McGarigal, K., Cushman, S. A., Neel, M. C., and Ene, E. (2002*). FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps*. Retrieved from www.umas.edu/landeco/research/fragstats/fragstats.html.

McHenry, M., and C., Rinner. (2016). The impact of multi-criteria decision analysis parameters on an urban deprivation index. *Cartographica*, 51(4), 179–197.

Mendoza, G. A., and Prabhu, R. (2000). Development of a methodology for selecting criteria and indicators of sustainable forest management: a case study on participatory assessment. *Environmental Management*, 26(6), 659-673.

Meng, Y., Malczewski, J., and Boroushaki, S. (2011). A GIS-based multicriteria decision analysis approach for mapping accessibility patterns of housing development sites: a case study in Canmore, Alberta. *Journal of Geographic Information System*, 3(01), 50.

Mitchel, A. (2005). *The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics*. Redlands, CA: ESRI Press.

Mitsova, D., Shuster, W., and Wang, X. (2011). A cellular automata model of land cover change to integrate urban growth with open space conservation. *Landscape and Urban Planning*, 99(2), 141-153.

Moghadam, H. S., and Helbich, M. (2013). Spatiotemporal urbanization processes in the megacity of Mumbai, India: a Markov chains-cellular automata urban growth model. *Applied Geography*, 40, 140-149.

Mokadi, E., Mitsova, D., and Wang, X. (2013). Projecting the impacts of a proposed streetcar system on the urban core land redevelopment: the case of Cincinnati, Ohio. *Cities*, 35, 136-146.

Molin, E. J. E. (1999). *Conjoint Modeling Approaches for Residential Group Preferences*. PhD dissertation, Faculty of Architecture, Eindhoven University of Technology, Eindhoven, the Netherlands.

Morgan, D.L. (1997). *Focus Groups as Qualitative Research*. London: Sage Publications.

Moulin, B., Chaker, W., Perron, J., Pelletier, P., Hogan, J., and Gbei, E. (2003). MAGS project: multi-agent geosimulation and crowd simulation. In *International Conference on Spatial Information Theory* (pp. 151-168). Berlin: Springer.

Munda, G. (1998). *Multicriteria Evaluation in a Fuzzy Environment: Theory and Applications in Ecological Economics*. Berlin: Springer.

Myint, S. W., and Wang, L. (2006). Multicriteria decision approach for land use land cover change using Markov chain analysis and a cellular automata approach. *Canadian Journal of Remote Sensing*, 32(6), 390-404.

National Cartographic Center. (2015). *Home Page*. Retrieved from http://www.ncc.org.ir.

Von Neumann, J., and Burks, A. W. (1966). Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, 5(1), 3-14.

Nourqolipour, R., Shariff, A. R. B. M., Ahmad, N. B., Balasundram, S. K., Sood, A. M., Buyong, T., and Amiri, F. (2015). Multi-objective-based modeling for land use change analysis in the South West of Selangor, Malaysia. *Environmental Earth Sciences*, 74(5), 4133-4143.

Nourqolipour, R., Shariff, A. R. B. M., Balasundram, S. K., Ahmad, N. B., Sood, A. M., Buyong, T., and Amiri, F. (2015). A GIS-based model to analyze the spatial and temporal development of oil palm land use in Kuala Langat district, Malaysia. *Environmental Earth Sciences*, 73(4), 1687-1700.

Nourqolipour, R., Shariff, A. R. B. M., Balasundram, S. K., Ahmad, N. B., Sood, A. M., and Buyong, T. (2016). Predicting the effects of urban development on land transition and spatial patterns of land use in Western Peninsular Malaysia. *Applied Spatial Analysis and Policy*, 9(1), 1-19.

O'Hagan, M. (1988). Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. In *Proceedings Twenty-Second Annual IEEE Asilomar Conference on Signals, Systems and Computers* (Vol. 2, pp. 681-689). Pacific Grove, CA: IEEE.

O'sullivan, D., and Torrens, P. M. (2001). Cellular models of urban systems. In *Theory and Practical Issues on Cellular Automata* (pp. 108-116). London: Springer.

Paine, D. P., and Kiser, J. D. (2003). *Aerial photography and Image Interpretation*. New York: John Wiley and Sons.

Park, S., Jeon, S., and Choi, C. (2012). Mapping urban growth probability in South Korea: comparison of frequency ratio, analytic hierarchy process, and logistic regression models and use of the environmental conservation value assessment. *Landscape and Ecological Engineering*, 8(1), 17-31.

Park, S., Jeon, S., Kim, S., and Choi, C. (2011). Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea. *Landscape and Urban Planning*, 99(2), 104-114.

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

Pimpler, E. (2015). *Programming ArcGIS with Python Cookbook*. Birmingham: Packt Publishing Ltd.

Polhill, J. G., Parker, D., Brown, D., and Grimm, V. (2008). Using the ODD protocol for describing three agent-based social simulation models of land-use change. *Journal of Artificial Societies and Social Simulation*, 11(2), 3.

Pontius, G. R., and Malanson, J. (2005). Comparison of the structure and accuracy of two land change models. *International Journal of Geographical Information Science*, 19(2), 243-265.

Pooyandeh, M., and Marceau, D. J. (2013). A spatial web/agent-based model to support stakeholders' negotiation regarding land development. *Journal of Environmental Management*, 129, 309-323.

Rafiee, R., Mahiny, A. S., Khorasani, N., Darvishsefat, A. A., and Danekar, A. (2009). Simulating urban growth in Mashad city, Iran through the SLEUTH model (UGM). *Cities*, 26(1), 19-26.

Ranji, H., Erfani, J., Haghgoo, N., and Ghayoomi, Z. (2013). *Examining and Comparing the Properties of Population and Housing in 22 Districts of Tehran Based on 2011 and 2006 Census Data*. Retrieved from http://web2.srtc.ac.ir/Files/Proposal-

DataSend/Proposal_View/Article_2015-04-27_11.01.36_Manategh%2022%20gane%20Tehran%201390-Abstract.pdf.

Rinner, C., and Malczewski, J. (2002). Web-enabled spatial decision analysis using Ordered Weighted Averaging (OWA). *Journal of Geographical Systems*, 4(4), 385-403.

Rogerson, P. A. (2015). *Statistical Methods for Geography: A Student's Guide*. Los Angeles: SAGE Publications.

Rovai, M., Andreoli, M., Gorelli, S., and Jussila, H. (2016). A DSS model for the governance of sustainable rural landscape: a first application to the cultural landscape of Orcia Valley (Tuscany, Italy), *Land Use Policy,* 56, 217–237.

Roy, B. (1996). *Multicriteria Methodology for Decision Analysis*. Dordrecht: Kluwer Academic Publishers.

Rutledge, D. T. (2003). Landscape indices as measures of the effects of fragmentation: can pattern reflect process?. In *DOC Science International Series 98 Department of Conservation*. Wellington.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.

Sabri, S., Ludin, A. N. M. M., and Ho, C. S. (2012). Conceptual design for an integrated geosimulation and analytic network process (ANP) in gentrification appraisal. *Applied Spatial Analysis and Policy*, 5(3), 253-271.

Sakieh, Y., Amiri, B. J., Danekar, A., Feghhi, J., and Dezhkam, S. (2015). Scenario-based evaluation of urban development sustainability: an integrative modeling approach to compromise between urbanization suitability index and landscape pattern. *Environment, Development and Sustainability*, 17(6), 1343-1365.

Sakieh, Y., Salmanmahiny, A., and Mirkarimi, S. H. (2017). Tailoring a non-path-dependent model for environmental risk management and polycentric urban land-use planning. *Environmental Monitoring and Assessment*, 189(2), 91.

Sayer, R. (1979). Understanding urban models versus understanding cities. *Environment and planning A*, 11(8), 853-862.

Silva, E. A., and Clarke, K. C. (2002). Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment and Urban Systems*, 26(6), 525-552.

Simon, H. A. (1957). *Models of Man; Social and Rational*. New York: John Wiley and Sons.

Singh, S. K., Mustak, S., Srivastava, P. K., Szabó, S., and Islam, T. (2015). Predicting spatial and decadal LULC changes through cellular automata Markov chain models using earth observation datasets and geo-information. *Environmental Processes*, 2(1), 61-78.

Sklar, F. H., Costanza, R., and Day, J. W. (1985). Dynamic spatial simulation modeling of coastal wetland habitat succession. *Ecological Modelling*, 29(1), 261-281.

Statistical Center of Iran. (2017a). *Home Page*. Retrieved from http://www.amar.org.ir/.

Statistical Center of Iran. (2017b). *Selected Results of the 2016 National Population and Housing Census* . Retrieved from https://www.amar.org.ir/Portals/1/census/2016/Iran_Census_2016_Selected_Results.pdf.

Statistical Center of Iran. (2017c). *Selected Results of the 2006 National Population and Housing Census for the city of Tehran* . Retrieved from https://www.amar.org.ir/Portals/0/Files/fulltext/1388/tehran-manategh-85.pdf.

Stevens, D., and Dragićević, S. (2007). A GIS-based irregular cellular automata model of land-use change. *Environment and Planning B: Planning and Design*, 34(4), 708-724.

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*, 53(1), 37-52.

Sun, Y., Tong, S. T., Fang, M., and Yang, Y. J. (2013). Exploring the effects of population growth on future land use change in the Las Vegas Wash watershed: an integrated approach of geospatial modeling and analytics. *Environment, Development and Sustainability*, 15(6), 1495-1515.

Surabuddin Mondal, M., Sharma, N., Kappas, M., and Garg, P. (2013). Modeling of spatio-temporal dynamics of land use and land cover in a part of Brahmaputra River basin using geoinformatic techniques. *Geocarto International*, 28(7), 632-656.

Tehran Municipality. (2017a). *Home Page*. Retrieved from http://www.atlas.tehran.ir.

Tehran Municipality. (2017b). *Climate and Air Pollution*. Retrieved from http://atlas.tehran.ir/Default.aspx?tabid=242.

Tehran Municipality. (2017c). *Population Increase in Tehran Districts*. Retrieved from http://atlas.tehran.ir/Default.aspx?tabid=264.

Tehran Municipality. (2017d). *Population Density*. Retrieved from http://atlas.tehran.ir/Default.aspx?tabid=269.

Tehran Municipality. (2017e). *Quality of Life in Tehran*. Retrieved from http://atlas.tehran.ir/Default.aspx?tabid=331.

Terra, T. N., dos Santos, R. F., and Costa, D. C. (2014). Land use changes in protected areas and their future: the legal effectiveness of landscape protection. *Land Use Policy*, 38, 378-387.

Tian, G., Ouyang, Y., Quan, Q., and Wu, J. (2011). Simulating spatiotemporal dynamics of urbanization with multi-agent systems—a case study of the Phoenix metropolitan region, USA. *Ecological Modelling*, 222(5), 1129-1138.

Torrens, P. M., and O'Sullivan, D. (2001). *Cellular Automata and Urban Simulation: Where Do We Go From Here?*. London: SAGE Publications.

Torrens, P. M. (2003). Cellular automata and multi-agent systems as planning support tools. In *Planning Support Systems in Practice* (pp. 205-222). Berlin: Springer.

Torrens, P. (2006). Geosimulation and its application to urban growth modeling. *Complex Artificial Environments*, 119-136.

U.S. Geological Survey. (2016). *Home Page*. Retrieved from https://www.usgs.gov.

Van Dyke Parunak, H., Brueckner, S. A., Matthews, R., and Sauter, J. (2006). Swarming methods for geospatial reasoning. *International Journal of Geographical Information Science*, 20(9), 945-964.

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

Von Neumann, J., and Burks, A. W. (1966). Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, 5(1), 3-14.

Waddell, P. (2000). A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim. *Environment and Planning B: Planning and Design*, 27(2), 247-263.

Wang, H., Shen, Q., Tang, B. S., and Skitmore, M. (2013). An integrated approach to supporting land-use decisions in site redevelopment for urban renewal in Hong Kong. *Habitat International*, 38, 70-80.

Wang, Y. M., and Parkan, C. (2005). A minimax disparity approach for obtaining OWA operator weights. *Information Sciences*, 175(1), 20-29.

White, R., and Engelen, G. (1994). Cellular Dynamics and GIS: modelling spatial complexity. *Geographical Systems*, 1, 237-253.

White, R., Engelen, G., and Uljee, I. (1997). The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. *Environment and Planning B: Planning and Design*, 24(3), 323-343.

Wilensky, U. (1999). *NetLogo*. Retrieved from: http://ccl.northwestern.edu/netlogo.

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature*, 311(5985), 419-424.

Wooldridge, M., and Jennings, N. R. (1995). Intelligent agents: theory and practice. *Knowledge Engineering Review*, 10(02), 115-152.

Wu, F. (1998). SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science*, 12(1), 63-82.

Wu, F. (2000). A parameterised urban cellular model combining spontaneous and self-organising growth. *Geocomputation: Innovation in GIS*, 7, 73-85.

Wu, F. (2005). Introduction-urban Simulation. In P. M. Atkinson, G. M. Foody, S. E. Darby, and F. Wu (Ed.), *GeoDynamics* (Chapter 15). Boca Raton, FL: CRC Press.

Wu, F., and Webster, C. J. (1998). Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B*, 25, 103-126.

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1), 183-190.

Yager, R. R. (1996). Quantifier guided aggregation using OWA operators. *International Journal of Intelligent Systems*, 11(1), 49-73.

Yu, J., Chen, Y., Wu, J., and Khan, S. (2011). Cellular automata-based spatial multi-criteria land suitability simulation for irrigated agriculture. *International Journal of Geographical Information Science*, 25(1), 131-148.

Zadeh, L. A. (1983). A computational approach to fuzzy quantifiers in natural languages. *Computers and Mathematics with Applications*, 9(1), 149-184.

Zandbergen, P. A. (2013). *Python Scripting for ArcGIS*. Redlands, CA: Esri Press.

Zhang, Q., Ban, Y., Liu, J., and Hu, Y. (2011). Simulation and analysis of urban growth scenarios for the Greater Shanghai Area, China. *Computers, Environment and Urban Systems*, 35(2), 126-139.

Zhang, H., Jin, X., Wang, L., Zhou, Y., and Shu, B. (2015). Multi-agent based modeling of spatiotemporal dynamical urban growth in developing countries: simulating future scenarios of Lianyungang city, China. *Stochastic Environmental Research and Risk Assessment*, 29(1), 63-78.

Zhou, B., and Kockelman, K. M. (2008). Neighborhood impacts on land use change: a multinomial logit model of spatial relationships. *The Annals of Regional Science*, 42(2), 321-340.

Appendix A

Table A1: Review of geosimulation-multicriteria studies

No	Authors	Year	Title	Journal	Study area (country)	Objective
1	Akın, A., Sunar, F., and Berberoğlu, S.	2015	Urban change analysis and future growth of Istanbul	*Environmental Monitoring and Assessment*	Istanbul (Turkey)	Prediction of future land-use changes
2	Bone, C., Dragicevic, S., and White, R.	2011	Modeling-in-the-middle: bridging the gap between agent-based modeling and multi-objective decision-making for land use change	*International Journal of Geographical Information Science*	Chilliwack (Canada)	Integrating agent-based modelling and MODM to simulate land-use change
3	Bozkaya, A. G., Balcik, F. B., Goksel, C., and Esbah, H.	2015	Forecasting land-cover growth using remotely sensed data: a case study of the Igneada protection area in Turkey	*Environmental Monitoring and Assessment*	Igneada (Turkey)	Forecasting land cover changes to protect natural areas
4	Cheng, J., and Masser, I.	2004	Understanding spatial and temporal processes of urban growth: cellular automata modelling	*Environment and Planning B: Planning and Design*	Wuhan (China)	Prediction of urban growth
5	Chowdhury, P. R., and Maithani, S.	2014	Modelling urban growth in the Indo-Gangetic plain using nighttime OLS data and cellular automata	*International Journal of Applied Earth Observation and Geoinformation*	Indo Gangetic plain (India)	Modelling urban growth using inexpensive data
6	de Noronha Vaz, E., Nijkamp, P., Painho, M., and Caetano, M.	2012	A multi-scenario forecast of urban change: A study on urban growth in the Algarve	*Landscape and Urban Planning*	Algarve region (Portugal)	Modelling urban growth to evaluate loss of ecosystems
7	Ghavami, S. M., and Taleai, M.	2016	Towards a conceptual multi-agent-based framework to simulate the spatial group decision-making process	*Journal of Geographical Systems*	Zanjan (Iran)	Simulating the approval process for an urban land-use master plan

No	Method	No of criteria/ objectives	MCDA/MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
1	CA/Markov	5	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	5	Actual
2	ABM/heuristic optimization	3	MODA	Goal programming (single-objective linear programming)	Vector	Land-use/cover maps	Not applicable	Households, commercial enterprises	7	Scenario-based
3	CA/Markov	8	MCDA	Fuzzy linear function	Raster	Landsat	30m	Land-use/cover cells	8	Actual
4	CA/Logistic function	4	MCDA	WLC	Raster	SPOT PAN/XS	100m	Land-use/cover cells	2	Actual
5	CA/Markov	8	MCDA	WLC	Raster	DMSP satellite imagery	1km	Land-use/cover cells	2	Actual
6	CA/Markov	5	MCDA	WLC	Raster	Land-use/cover maps	100m	Land-use/cover cells	5	Scenario-based
7	MAS	4	MCDA	AHP	Raster	Landsat	30m	Land-use/cover cells	5	Actual

No	Authors	Year	Title	Journal	Study area (country)	Objective
8	Gong, W., Yuan, L., Fan, W., and Stott, P.	2015	Analysis and simulation of land use spatial pattern in Harbin prefecture based on trajectories and cellular automata—Markov	*International Journal of Applied Earth Observation and*	Harbin (China)	Simulating the pattern of land se change
9	Hansen, H. S.	2010	Modelling the future coastal zone urban development as implied by the IPCC SRES and assessing the impact from sea	*Landscape and Urban Planning*	Aalborg (Denmark)	Simulating the future urban growth on the coastal zone
10	Hansen, H. S.	2012	Empirically derived neighbourhood rules for urban land-use modelling	*Environment and Planning B: Planning and Design*	Aalborg (Denmark)	Improving CA-based land-use change through a new method for defining the interactions between adjacent land parcels
11	Henríquez, C., Azócar, G., and Romero, H.	2006	Monitoring and modeling the urban growth of two mid-sized Chilean cities	*Habitat International*	Chillán and Los Ángeles (Chile)	Analyzing land-use/cover changes
12	Hosseinali, F., Alesheikh, A. A., and Nourian, F.	2013	Agent-based modeling of urban land-use development, case study: Simulating future scenarios of Qazvin city	*Cities*	Qazvin (Iran)	Modelling urban development by simulating engaging agent behaviours
13	Hosseinali, F., Alesheikh, A. A., and Nourian, F.	2015	Assessing urban land-use development: Developing an agent-based model	*KSCE Journal of Civil Engineering*	Qazvin (Iran)	Modelling urban sprawl by taking risk attitude of land developers into consideration
14	Hyandye, C., and Martz, L. W.	2017	A Markovian and cellular automata land-use change predictive model of the Usangu Catchment	*International Journal of Remote Sensing*	Usangu Catchment (Tanzania)	Prediction of future land-use/cover changes

No	Method	No of criteria/ objectives	MCDA/MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
8	CA/Markov	9	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	5	Actual
9	Constrained CA	2	MCDA	WLC	Raster	Land-use/cover maps	100m	Land-use/cover cells	7	Scenario-based
10	CA	2	MCDA	WLC	Raster	Land-use/cover maps	100m	Land-use/cover cells	8	Actual
11	CA/Markov	3	MCDA	WLC	Raster	Land-use/cover maps	10m	Land-use/cover cells	2	Actual
12	MAS	3	MCDA	WLC	Raster	Land-use/cover maps	100m	Young person, high-income developers, rich people, low-income people, moderate to low-income people	2	Actual
13	MAS	3	MCDA	WLC	Raster	Land-use/cover maps	100m	Young person, high-income developers, rich people, low-income people, moderate to low-income people	5	Scenario-based
14	CA/Markov	8	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	5	Actual

No	Authors	Year	Title	Journal	Study area (country)	Objective
15	Jokar Arsanjani, J., Helbich, M., and de Noronha Vaz, E.	2013	Spatiotemporal simulation of urban growth patterns using agent-based modeling: The case of Tehran	*Cities*	Tehran (Iran)	Simulating urban growth
16	Keshtkar, H., and Voigt, W.	2016	A spatiotemporal analysis of landscape change using an integrated Markov chain and cellular automata models	*Modeling Earth Systems and Environment*	Central Germany	Predicting future land-cover changes
17	Lau, K. H., and Kam, B. H.	2005	A cellular automata model for urban land-use simulation	*Environment and Planning B: Planning and Design*	Melbourne (Australia)	Developing a new framework to simulate land-use changes
18	Li, X., and Liu, X.	2007	Defining agents' behaviors to simulate complex residential development using multicriteria evaluation	*Journal of Environmental Management*	Guangzhou (China)	Modelling residential development by simulating the decision behaviour of interest groups
19	Li, C., and Zhao, J.	2017	Assessment of future urban growth impact on landscape pattern using cellular automata model: a case study of Xuzhou	*Landscape and Urban Planning*	Xuzhou (China)	Evaluating the consequences of future urban growth
20	Ligmann-Zielinska, A.	2009	The impact of risk-taking attitudes on a land use pattern: an agent-based model of residential development	*Journal of Land Use Science*	Hypothetical	Evaluating the spatial consequences of using different utility functions that reflect people's risk attitude
21	Ligmann-Zielinska, A., and Sun, L.	2010	Applying time-dependent variance-based global sensitivity analysis to represent the dynamics of an agent-based model of land use change	*International Journal of Geographical Information Science*	Hypothetical	Examining the uncertainty of ABM for land-use change

No	Method	No of criteria/ objectives	MCDA/MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
15	MAS	17	MCDA	WLC	Raster	Landsat	30m	Resident, government, developer	2	Actual
16	CA/Markov	4	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	5	Scenario-based
17	CA	4	MCDA	WLC	Raster	Land-use/cover maps	1km	Land-use/cover cells	10	Pilot application
18	MAS	12 (for resident agent)	MCDA	WLC	Raster	Landsat	100m	low-income without children, middle-income without children, high-income without children, low-income with children, middle-income with children, high-income with children, government, developers	5	Actual
19	CA	6	MCDA	WLC	Raster	Landsat	100m	Land-use/cover cells	4	Scenario-based
20	ABM	2	MCDA	Ideal point method	Raster	Not applicable	Not specified	Developers	2	Pilot application
21	ABM	3	MCDA	Ideal point method	Raster	Not applicable	Not specified	Developers	2	Pilot application

No	Authors	Year	Title	Journal	Study area (country)	Objective
22	Ligmann-Zielinska, A., and Jankowski, P.	2010	Exploring normative scenarios of land use development decisions with an agent-based simulation laboratory	*Computers, Environment and Urban Systems*	Washington (USA)	Examining if other arrangements of land-use structure is possible from property developers and planning agencies point of view to mitigate the negative impacts of current suburban sprawl
23	Ligtenberg, A., Bregt, A. K., and Van Lammeren, R.	2001	Multi-actor-based land use modelling: spatial planning using agents	*Landscape and Urban Planning*	Nijmegen (The Netherlands)	Developing a new framework for spatial planning
24	Ligtenberg, A., Wachowicz, M., Bregt, A. K., Beulens, A., and Kettenis, D. L.	2004	A design and application of a multi-agent system for simulation of multi-actor spatial planning	*Journal of Environmental Management*	Land van Maas en Waal (The Netherlands)	Simulating individuals decision making behaviours to examine spatial scenarios for spatial planning process
25	Liu, Y., Lv, X., Qin, X., Guo, H., Yu, Y., Wang, J., and Mao, G.	2007	An integrated GIS-based analysis system for land-use management of lake areas in urban fringe	*Landscape and Urban Planning*	Wuhan (China)	Land-use allocation based on the suitability of land parcels for the conservation of aquatic ecosystems
26	Liu, R., Zhang, K., Zhang, Z., and Borthwick, A. G.	2014	Land-use suitability analysis for urban development in Beijing	*Journal of Environmental Management*	Beijing (China)	Generating suitability map for urban development
27	Loibl, W., and Toetzer, T.	2003	Modeling growth and densification processes in suburban regions—simulation of landscape transition with spatial agents	*Environmental Modelling and Software*	the Greater Vienna Region (Austria)	Modelling suburban development
28	Mahiny, A. S., and Clarke, K. C.	2012	Guiding SLEUTH land-use/land-cover change modeling using multicriteria evaluation: towards dynamic sustainable land-use planning	*Environment and Planning B: Planning and Design*	Gorgan Township (Iran)	Upgrading the SLEUTH model to simulate land-use changes

No	Method	No of criteria/ objectives	MCDA/ MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
22	ABM/ MOLA	3 objectives	MCDA	Ideal point method	Raster	Land-use/cover maps	127m	Developers	3	Scenario-based
23	MAS	Not specified	MCDA	Weighted summation	Raster	Land-use/cover maps	100m	Municipality of Nimegen, the new rich, nature and	11	Scenario-based
24	MAS	Not specified	Not specified	Not specified	Raster	Land-use/cover maps	Not specified	Regional authorities, farmers' organization, environmentalists' organization	20	Scenario-based
25	CA	11	MCDA	AHP	Raster	Land-use/cover maps	100m and 200m	Land-use/cover cells	9	Scenario-based
26	CA	10	MCDA	OWA / Ideal point method	Raster	Not applicable	100m	Land-use/cover cells	Not specified	Actual
27	MAS/CA	8	MCDA	WLC	Raster	Corona, IRCS-1C, Landsat	100m	High-income and highly educated households, moderate to high-income households, moderate-income and highly educated younger households, low-income households, weekend-home seekers, enterprise founders/owners	4	Actual
28	CA	15	MCDA	Fuzzy linear function	Raster	Landsat	30m	Land-use/cover cells	9	Actual

No	Authors	Year	Title	Journal	Study area (country)	Objective
29	Manganelli, B., Di Palma, F., Amato, F., Nolè, G., and Murgante, B.	2016	The effects of socio-economic variables in urban growth simulations	*Procedia-Social and Behavioral Sciences*	Vulture-Alto Bradano (Italy)	Comparing the results of two urban growth simulations, one with considering socio-economic variables and the other without socio-economic variables
30	Manson, S. M.	2005	Agent-based modeling and genetic programming for modeling land change in the Southern Yucatan Peninsular Region of Mexico	*Agriculture, Ecosystems and Environment*	Southern Yucatán Peninsular Region (Mexico)	Developing a new method for modelling land-use/cover changes
31	Mitsova, D., Shuster, W., and Wang, X.	2011	A cellular automata model of land cover change to integrate urban growth with open space conservation	*Landscape and Urban Planning*	27 counties in Ohio, Indiana, Kentucky (USA)	Developing a new model of urban development by considering environmentally sensitive areas into the modelling process
32	Moghadam, H. S., and Helbich, M.	2013	Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model	*Applied Geography*	Mumbai (India)	Simulating past urban lan-use changes and predicting future pattern
33	Mokadi, E., Mitsova, D., and Wang, X.	2013	Projecting the impacts of a proposed streetcar system on the urban core land redevelopment: The case of Cincinnati, Ohio	*Cities*	Cincinnati (USA)	The effect of the new streetcar project on the urban core land redevelopment
34	Myint, S. W., and Wang, L.	2006	Multicriteria decision approach for land use land cover change using Markov chain analysis and a cellular automata approach	*Canadian Journal of Remote Sensing*	Norman (USA)	Identifying land-use/cover change
35	Nourqolipour, R., Shariff, A. R. B. M., Ahmad, N. B., Balasundram, S. K., Sood, A. M., Buyong, T., and Amiri, F.	2015	Multi-objective-based modeling for land use change analysis in the South West of Selangor, Malaysia	*Environmental Earth Sciences*	Selangor (Malaysia)	Simulating land-use change

No	Method	No of criteria/ objectives	MCDA/ MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
29	CA	5	MCDA	AHP	Raster	Landsat	30m	Land-use/cover cells	8	Scenario-based
30	MAS/CA	Not specified	MCDA	WLC	Raster	Satellite imagery	28.5 m^2	Households, institutions	7	Actual
31	CA/Markov	Not specified	MCDA	Fuzzy linear function	Raster	Land-use/cover maps	Not specified	Land-use/cover cells	5	Scenario-based
32	CA/Markov	5	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	5	Actual
33	CA/Markov	7	MCDA	WLC	Raster	Land-use/cover maps	Not specified	Land-use/cover cells	4	Scenario-based
34	CA/Markov	7	MCDA	WLC	Raster	Landsat	30m, 79m	Land-use/cover cells	7	Actual
35	CA/Markov	3 objectives	MCDA/MODA	WLC	Raster	Land-use/cover maps	20m	Land-use/cover cells	5	Actual

No	Authors	Year	Title	Journal	Study area (country)	Objective
36	Nourqolipour, R., Shariff, A. R. B. M., Balasundram, S. K., Ahmad, N. B., Sood, A. M., Buyong, T., and Amiri, F.	2015	A GIS-based model to analyze the spatial and temporal development of oil palm land use in Kuala Langat district, Malaysia	*Environmental Earth Sciences*	Kuala Langat district (Malaysia)	Simulating the expansion of oil palm land-use
37	Nourqolipour, R., Shariff, A. R. B. M., Balasundram, S. K., Ahmad, N. B., Sood, A. M., and Buyong, T.	2016	Predicting the effects of urban development on land transition and spatial patterns of land use in western Peninsular Malaysia	*Applied Spatial Analysis and Policy*	Town of Banting and the adjacent townships (Malaysia)	Quantifying the effect on urban growth on the dynamic of land-use
38	Park, S., Jeon, S., Kim, S., and Choi, C.	2011	Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea	*Landscape and Urban Planning*	South Korea	Comparing the result of urban land-use change prediction using four different methods to generate the suitability index. Frequency ratio, AHP, logistic regression, artificial neural network
39	Park, S., Jeon, S., and Choi, C.	2012	Mapping urban growth probability in South Korea: comparison of frequency ratio, analytic hierarchy process, and logistic regression models and use of the environmental conservation value assessment	*Landscape and Ecological Engineering*	South Korea	Comparing three different models to simulate urban growth patterns. Frequency ratio, AHP, logistic regression
40	Pontius, G. R., and Malanson, J.	2005	Comparison of the structure and accuracy of two land change models	*International Journal of Geographical Information Science*	The town of Worcester and the nine adjacent towns in central Massachusetts (USA)	Comparing the predictive power of two land change models
41	Pooyandeh, M., and Marceau, D. J.	2013	A spatial web/agent-based model to support stakeholders' negotiation regarding land development	*Journal of Environmental Management*	Elbow river watershed (Canada)	Developing a spatial web/ABM system to support the negotiation process of stakeholders for land development
42	Sabri, S., Ludin, A. N. M. M., and Ho, C. S.	2012	Conceptual design for an integrated geosimulation and analytic network process (ANP) in gentrification appraisal	*Applied Spatial Analysis and Policy*	Not applicable	Developing a conceptual framework for modelling the process of gentrification

No	Method	No of criteria/ objectives	MCDA/ MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
36	CA/Markov	9	MCDA	WLC	Raster	Land-use/cover maps	20m	Land-use/cover cells	6	Actual
37	CA/Markov	9	MCDA	WLC	Raster	Landsat/Spot	20m	Land-use/cover cells	5	Actual
38	CA and frequency ratio/AHP/logistic regression/ANN	9	MCDA	AHP	Raster	Land-use/cover maps	30m	Land-use/cover cells	7	Actual
39	CA and frequency ratio/AHP/logistic regression9	9	MCDA	AHP	Raster	Land-use/cover maps	30m	Land-use/cover cells	7	Actual
40	CA/Markov	2	MCDA	Not specified	Raster	Land-use/cover maps	30m	Land-use/cover cells	2	Actual
41	ABM	Different for each type of agent	MCDA	Fuzzy AHP	Not specified	Land-use/cover maps	Not specified	Stakeholders	Not specified	Scenario-based
42	MAS/CA	2 objectives	MCDA	ANP	Raster	Not specified	Not specified	Residential	4	Not specified

No	Authors	Year	Title	Journal	Study area (country)	Objective
43	Sakieh, Y., Amiri, B. J., Danekar, A., Feghhi, J., and Dezhkam, S.	2015	Scenario-based evaluation of urban development sustainability: an integrative modeling approach to compromise between urbanization suitability index and landscape pattern	*Environment, Development and Sustainability*	Karaj (Iran)	Examining the relationship between land suitability and land-use patterns
44	Sakieh, Y., Salmanmahiny, A., and Mirkarimi, S. H.	2017	Tailoring a non-path-dependent model for environmental risk management and polycentric urban land-use planning	*Environmental Monitoring and Assessment*	Gorgan and Ali-Abad (Iran)	Developing a framework for environmental risk management and urban land-use allocation
45	Singh, S. K., Mustak, S., Srivastava, P. K., Szabó, S., and Islam, T.	2015	Predicting spatial and decadal LULC changes through cellular automata Markov chain models using earth observation datasets and geo-information	*Environmental Processes*	Allahabad (India)	Simulating spatial and temporal land-use/cover changes
46	Sun, Y., Tong, S. T., Fang, M., and Yang, Y. J.	2013	Exploring the effects of population growth on future land use change in the Las Vegas Wash watershed: an integrated approach of geospatial modeling and analytics	*Environment, Development and Sustainability*	Las Vegas Wash watershed (USA)	Modelling future land-use pattern to mitigate environmental side effects of urban growth
47	Surabuddin Mondal, M., Sharma, N., Kappas, M., and Garg, P.	2013	Modeling of spatio-temporal dynamics of land use and land cover in a part of Brahmaputra River basin using geoinformatic techniques	*Geocarto International*	Brahmaputra River basin (India)	Spatiotemporal modelling to monitor and predict land-use/cover changes
48	Terra, T. N., dos Santos, R. F., and Costa, D. C.	2014	Land use changes in protected areas and their future: The legal effectiveness of landscape protection	*Land Use Policy*	Southern Sao Paulo State (Brazil)	Simulating land-use changes to examine the impact of restrictive legal instrument to preserve protected areas
49	Wu, F.	1998	SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules	*International Journal of Geographical Information Science*	Guangzhou (China)	Developing a new model and system to simulate land-use changes

No	Method	No of criteria/ objectives	MCDA/ MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
43	CA	9	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	Not specified	Scenario-based
44	CA/Markov	27	MCDA	WLC	Raster	Land-use/cover maps	30m	Land-use/cover cells	8	Scenario-based
45	CA/Markov	7	MCDA	WLC	Raster	Landsat	30m	Land-use/cover cells	7	Actual
46	CA/Markov	2	MCDA	WLC	Raster	Land-use/cover maps	30m	Land-use/cover cells	5	Scenario-based
47	CA/Markov	Not specified	MCDA/ MOLA	Not specified	Raster	Landsat/IRS	23.5m	Land-use/cover cells	14	Actual
48	CA/Markov	4	MCDA	WLC	Raster	digitized aerial photographs/ World View	3m (squared unit)/ 900m² (hexagonal unit)	Land-use/cover cells	7	Actual
49	CA	5	MCDA	WLC	Raster	Landsat	200m	Land-use/cover cells	5	Pilot application\ Scenario-based

No	Authors	Year	Title	Journal	Study area (country)	Objective
50	Wu, F., and Webster, C. J.	1998	Simulation of land development through the integration of cellular automata and multicriteria evaluation	*Environment and Planning B*	Guangzhou (China)	Applying MCDA to define transition rules of CA-based models
51	Yu, J., Chen, Y., Wu, J., and Khan, S.	2011	Cellular automata-based spatial multi-criteria land suitability simulation for irrigated agriculture	*International Journal of Geographical Information Science*	Macintyre Brook (Australia)	Developing a new framework for land suitability analysis
52	Zhang, Q., Ban, Y., Liu, J., and Hu, Y.	2011	Simulation and analysis of urban growth scenarios for the Greater Shanghai Area, China	*Computers, Environment and Urban Systems*	The Greater Shanghai Area (China)	Modelling urban growth
53	Zhang, H., Jin, X., Wang, L., Zhou, Y., and Shu, B.	2015	Multi-agent based modeling of spatiotemporal dynamical urban growth in developing countries: simulating future scenarios of Lianyungang city, China	*Stochastic Environmental Research and Risk Assessment*	Lianyungang (China)	Spatial and temporal simulation of urban growth

No	Method	No of criteria/ objectives	MCDA/ MODA	MCA method	Data model	Base map	Spatial resolution	Type of agents	No of land-use types	Type of implementation
50	CA	6	MCDA	AHP	Raster	Landsat	200m	Land-use/cover cells	3	Scenario-based
51	CA	7	MCDA	WLC	Raster	Land-use/cover maps	100m	Land-use/cover cells	9	Actual
52	CA/Markov	Total of 7 (different number for each type of land-use)	MCDA/MODA	WLC	Raster	Landsat, China-Brazil Earth Resource Satellite image	30m	Land-use/cover cells	6	Scenario-based
53	MAS/CA	8	MCDA	WLC	Raster	Land-use/cover maps	Not specified	Resident agents, farmer agents, industrial enterprise agents, environmentalist agents	9	Scenario-based

Appendix B

To collect the preferential information for this research, ten experts familiar with the study area were contacted. Six of them agreed to collaborate. The names and contact information of the participants are available from the author (email: hhossei7@uwo.ca).

B1: Selecting criteria

B1.1. *Please list up to five criteria that you consider relevant for evaluating a parcel of land in terms of its suitability for residential development in Tehran*.

1. ………………………………………………………………………………
2. ………………………………………………………………………………
3. ………………………………………………………………………………
4. ………………………………………………………………………………
5. ………………………………………………………………………………

B1.2. *The participants are presented with the list of criteria identified by review of literature about geosimulation-multicriteria modelling of urban growth in Iran*.

B1.3. *The two lists of criteria obtained in B1.1 and B1.2 are compared and discussed using the focus group format to select a final list of criteria.*

B1.4. *The criteria are classified according to underlying objectives of agents*: *households, real estate developers and local authorities*.

The results of this procedure are shown in Table B1.1.

Table B1.1: Criteria for evaluating the land suitability for residential development in Tehran

Step B1.1	Step B1.2	Step B1.3	Step B1.4			
Criteria identified by experts individually	**Criteria identified in the literature review**	**Criteria identified by experts collectively**	**Criteria of household agent**	**Criteria of developer agent**	**Criteria (constraints) of local authority**	**Objectives**
Cost of land acquisition	Air quality	Cost of land acquisition		x		Maximize profit
Distance to airports	Construction expenses	Distance to airports			x	Constraint
Conservation areas	Distance to CBD	Conservation areas			x	Constraint
Distance wasteland	Distance to industrial sites/areas	Distance to military zones			x	Constraint
Distance to military zones	Distance to nearby cities	Housing price		x		Maximize profit
Elevation	Distance to protected/conservation areas	Proximity to education centres	x			Maximize accessibility
Green space index	Easting coordinates	Green space index	x			Maximize neighbourhood quality
Housing price	Elevation	Proximity to major roads	x			Maximize accessibility
Land-use/cover	Household income	Proximity to public transit	x			Maximize accessibility
Population density	Land-use/cover	Proximity to shopping centres	x			Maximize accessibility
Population structure by age	Northing coordinates	Proximity to major workplaces	x			Maximize accessibility
Profit on investment	Open lands	Residential intensity index	x			Maximize neighbourhood quality
Proximity to education centres	Percentage of young population	Slope gradient			x	Constraint
Proximity to major roads	Population density					
Proximity to public transit	Proximity to building blocks					
Proximity to shopping centres	Proximity to CBD					
Proximity to major workplaces	Proximity to interchange					
Residential	Proximity to parks					

intensity index						
Slope gradient	Proximity to power lines					
Safety	Proximity to residential areas					
	Proximity to rivers and water bodies					
	Proximity to roads					
	Proximity to single buildings					
	Proximity to town edges					
	Underground water depth					
	Seismic hazards/distance to geological faults					
	Slope gradient					
	Type of geological structure					

B2: Eliciting criterion weights

The ranking exercise is a technique in which criteria are ranked from the most important to the least important. Ranking is a commonly used method to prioritize criteria in GIS-based multicriteria analyses and often combined with the point allocation method where points are allocated over criteria to reflect their relative importance.

B2.1. Ranking

Imagine the starting point is at the worst level for each criterion. Identify which criterion you would like to improve first to its best level (then assign rank 1 to that criterion); identify which criterion you would like to improve second to its best level (then assign rank 2 to that criterion); etc. Table B2.1 contains the list of criteria and the range of values for each criterion. The experts were asked to write the ranks in the third column of the table.

Table B2.1: Ranking criteria

Criterion	Range of values	Rank
Proximity to education centres	Best: 0 m Worst: 4,294 m	
Proximity to major workplaces	Best: 0 m Worst: 7,240 m	
Proximity to shopping centres	Best: 0 m Worst: 12,480 m	
Proximity to major roads	Best: 0 m Worst: 4,957 m	
Proximity to public transit	Best: 0 m Worst: 5,975 m	
Green space index	Best: 100 % Worst: 0 %	
Residential intensity index	Best: 100 % Worst: 0 %	
Housing price	Best: 5,500,000 Rial/m^2 Worst: 620,000 Rial/m^2	
Cost of land acquisition	Best: 1,100,000 Rial/m^2 Worst: 10,200,000 Rial/m^2	

B2. Allocating points

Give the first-rank criterion 100 points; and then, allocate points to other criteria relative to the range of the most important criterion. Table B2.2 contains the list of criteria and the range of values for each criterion. The experts were asked to allocate a point to each criterion from 0 to 100.

Table B2.2: Allocating points to criteria

Criterion	Range of values	Points
Proximity to education centres	Best: 0 m Worst: 4,294 m	
Proximity to major workplaces	Best: 0 m Worst: 7,240 m	
Proximity to shopping centres	Best: 0 m Worst: 12,480 m	
Proximity to major roads	Best: 0 m Worst: 4,957 m	
Proximity to public transit	Best: 0 m Worst: 5,975 m	
Green space index	Best: 100 % Worst: 0 %	
Residential intensity index	Best: 100 % Worst: 0 %	
Housing price	Best: 5,500,000 Rial/m^2 Worst: 620,000 Rial/m^2	
Cost of land acquisition	Best: 1,100,000 Rial/m^2 Worst: 10,200,000 Rial/m^2	

B3: Eliciting value functions

A value function transforms the raw criterion scores into a scaled value ranging from 0 (the worst criterion outcome) to 1 (the best criterion outcome). A value function standardizes incommensurate criterion. The following procedure is applied for identifying the shape of the value function for each of the nine criteria.

B3.1. Identify the worst (c_{worst}) and best (c_{best}) scores for a given criterion (see Table B3.1)

Table B3.1: Worst and best criterion values

Criteria	Worst criterion value (c_{worst})	Best criterion value (c_{best})
Proximity to education centres	4294	0
Proximity to major workplaces	7240	0
Proximity to shopping centres	12480	0
Proximity to major roads	4957	0
Proximity to public transit	5975	0
Green space index	0	100
Residential intensity index	0	100
Housing price	620,000	5,500,000
Cost of land acquisition	10,200,000	1,100,000

B3.2. Set $v(c_{worst}) = 0$, $v(c_{best}) = 1$ (see Figure B3.1)

B3.3. Identify the 'bisection point' m_1 for which moving from c_{worst} to m_1 is just as valuable as moving from m_1 to c_{best}. The relative value of m_1 must be 0.5. You now have 3 points on the value function curve.

B3.4. To get more points, identify the bisection point m_2 between c_{worst} and m_1. It has relative value of 0.25; then the bisection point m_3 between m_1 and c_{best}, which has value of 0.75.

B3.5. Given the five points on the curve, a continuous value function is estimated (see Figure B3.2)

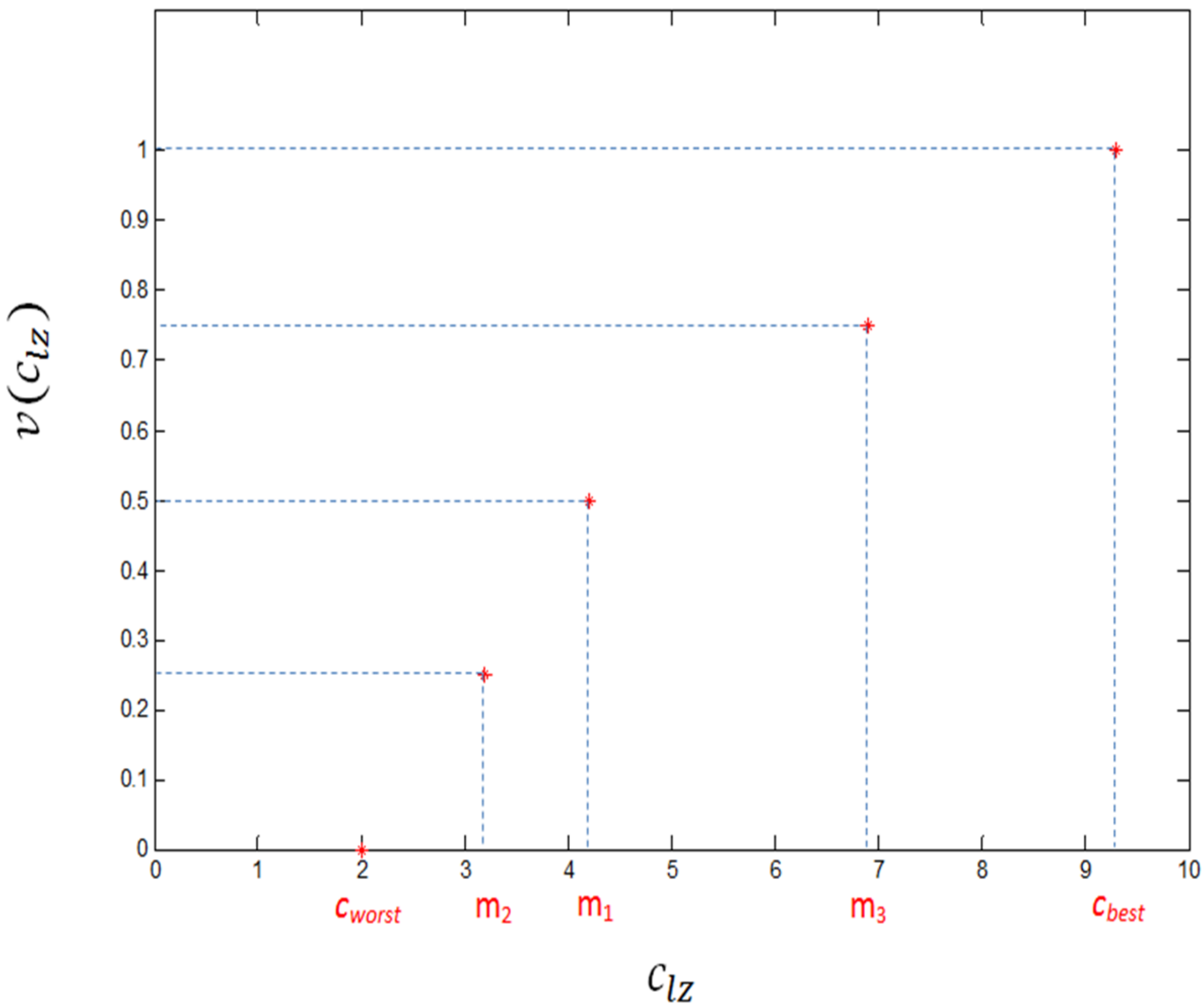


Figure B3.1: Finding the bisection points

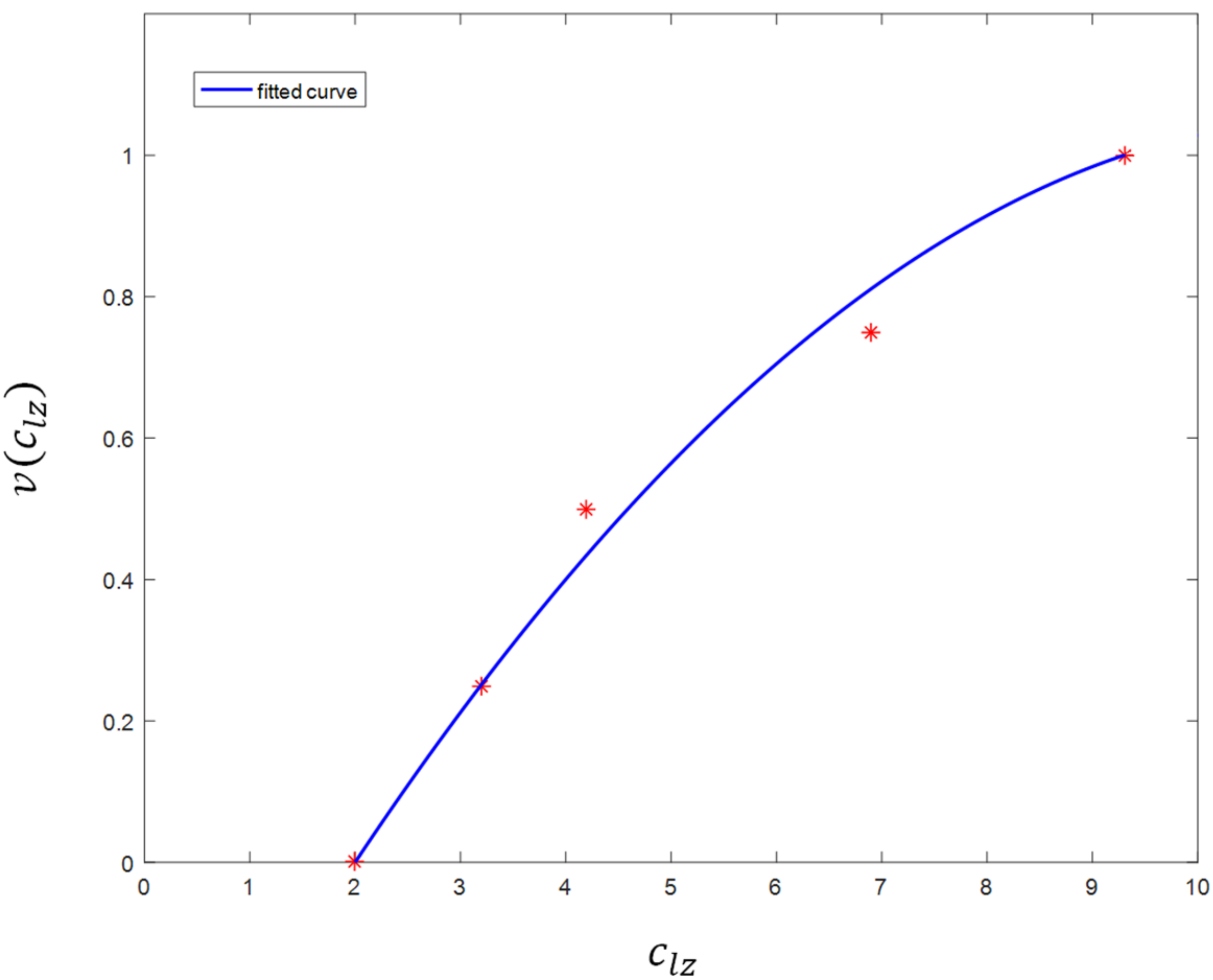


Figure B3.2: The estimated value function fitted to the bisection points

Appendix C

Table C1: The results of the evaluation of outputs

Order	Linguistic quantifier			Reference data			Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
				Residential	Undeveloped	Total								
2	All	Simulation	Residential	15,002	9,544	24,546								
			Undeveloped	9,319	100,585	109,904	85.97	38.88	38.32	52.83	52.53	2.003	66.07	0.315
			Total	24,321	110,129	134,450								
	Most		Residential	14,496	10,050	24,546								
			Undeveloped	9,825	100,079	109,904	85.22	40.94	40.4	50.3	50.01	1.731	60.24	0.319
			Total	24,321	110,129	134,450								
	Many		Residential	14,519	10,027	24,546								
			Undeveloped	9,802	100,102	109,904	85.25	40.85	40.3	50.41	50.13	1.780	60.62	0.318
			Total	24,321	110,129	134,450								
	Half		Residential	14,653	9,893	24,546								
			Undeveloped	9,668	100,236	109,904	85.45	40.3	39.75	51.08	50.8	1.830	60.79	0.314
			Total	24,321	110,129	134,450								
	Some		Residential	14,667	9,879	24,546								
			Undeveloped	9,654	100,250	109,904	85.47	40.25	39.69	51.15	50.86	1.962	60.65	0.315
			Total	24,321	110,129	134,450								
	Few		Residential	14,598	9,948	24,546								
			Undeveloped	9,723	100,181	109,904	85.37	40.53	39.98	50.81	50.52	1.969	60.34	0.313
			Total	24,321	110,129	134,450								
	At least one		Residential	14,234	10,312	24,546								
			Undeveloped	10,087	99,817	109,904	84.83	42.01	41.47	48.99	48.71	2.608	73.42	0.380
			Total	24,321	110,129	134,450								

Order	Linguistic quantifier	Simulation	Reference data: Residential	Reference data: Undeveloped	Reference data: Total	Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
4	All	Residential	15,012	9,534	24,546	85.99	38.84	38.28	52.88	52.58	1.877	65.83	0.318
		Undeveloped	9,309	100,595	109,904								
		Total	24,321	110,129	134,450								
	Most	Residential	14,537	10,009	24,546	85.28	40.78	40.23	50.5	50.22	1.656	60.37	0.315
		Undeveloped	9,784	100,120	109,904								
		Total	24,321	110,129	134,450								
	Many	Residential	14,531	10,015	24,546	85.27	40.8	40.25	50.47	50.19	1.665	60.98	0.310
		Undeveloped	9,790	100,114	109,904								
		Total	24,321	110,129	134,450								
	Half	Residential	14,555	9,991	24,546	85.31	40.7	40.15	50.59	50.31	1.686	61.29	0.321
		Undeveloped	9,766	100,138	109,904								
		Total	24,321	110,129	134,450								
	Some	Residential	14,628	9,918	24,546	85.41	40.41	39.85	50.96	50.67	1.701	61.24	0.318
		Undeveloped	9,693	100,211	109,904								
		Total	24,321	110,129	134,450								
	Few	Residential	14,621	9,925	24,546	85.4	40.43	39.88	50.92	50.64	1.714	61.03	0.317
		Undeveloped	9,700	100,204	109,904								
		Total	24,321	110,129	134,450								
	At least one	Residential	14,289	10,257	24,546	84.91	41.79	41.25	49.26	48.98	2.617	73.45	0.356
		Undeveloped	10,032	99,872	109,904								
		Total	24,321	110,129	134,450								

Order	Linguistic quantifier		Reference data: Residential	Reference data: Undeveloped	Reference data: Total	Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
∞	All	Simulation: Residential	15,006	9,540	24,546	85.98	38.87	38.3	52.85	52.55	1.891	65.27	0.314
		Undeveloped	9,315	100,589	109,904								
		Total	24,321	110,129	134,450								
	Most	Residential	14,544	10,002	24,546	85.29	40.75	40.2	50.54	50.25	1.650	60.83	0.318
		Undeveloped	9,777	100,127	109,904								
		Total	24,321	110,129	134,450								
	Many	Residential	14,547	9,999	24,546	85.29	40.74	40.19	50.55	50.27	1.734	61.06	0.317
		Undeveloped	9,774	100,130	109,904								
		Total	24,321	110,129	134,450								
	Half	Residential	14,559	9,987	24,546	85.31	40.69	40.14	50.61	50.33	1.844	61.67	0.315
		Undeveloped	9,762	100,142	109,904								
		Total	24,321	110,129	134,450								
	Some	Residential	14,664	9,882	24,546	85.47	40.26	39.71	51.14	50.85	1.879	61.03	0.319
		Undeveloped	9,657	100,247	109,904								
		Total	24,321	110,129	134,450								
	Few	Residential	14,643	9,903	24,546	85.44	40.34	39.79	51.03	50.75	1.880	61.36	0.324
		Undeveloped	9,678	100,226	109,904								
		Total	24,321	110,129	134,450								
	At least one	Residential	14,252	10,294	24,546	84.85	41.94	41.4	49.08	48.8	2.714	73.88	0.364
		Undeveloped	10,069	99,835	109,904								
		Total	24,321	110,129	134,450								

Order	Linguistic quantifier			Reference data			Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
				Residential	Undeveloped	Total								
16	All	Simulation	Residential	15,038	9,508	24,546	86.02	38.74	38.17	53.01	52.71	2.010	66.98	0.307
			Undeveloped	9,283	100,621	109,904								
			Total	24,321	110,129	134,450								
	Most		Residential	14,589	9,957	24,546	85.36	40.56	40.01	50.76	50.48	1.739	62.40	0.313
			Undeveloped	9,732	100,172	109,904								
			Total	24,321	110,129	134,450								
	Many		Residential	14,584	9,962	24,546	85.35	40.59	40.04	50.74	50.45	1.856	64.31	0.310
			Undeveloped	9,737	100,167	109,904								
			Total	24,321	110,129	134,450								
	Half		Residential	14,667	9,879	24,546	85.47	40.25	39.69	51.15	50.86	1.872	63.65	0.312
			Undeveloped	9,654	100,250	109,904								
			Total	24,321	110,129	134,450								
	Some		Residential	14,692	9,854	24,546	85.51	40.15	39.59	51.28	50.99	1.901	63.80	0.311
			Undeveloped	9,629	100,275	109,904								
			Total	24,321	110,129	134,450								
	Few		Residential	14,685	9,861	24,546	85.5	40.17	39.62	51.24	50.95	1.926	62.69	0.315
			Undeveloped	9,636	100,268	109,904								
			Total	24,321	110,129	134,450								
	At least one		Residential	14,202	10,344	24,546	84.78	42.14	41.61	48.83	48.55	2.646	75.11	0.370
			Undeveloped	10,119	99,785	109,904								
			Total	24,321	110,129	134,450								

Order	Linguistic quantifier			Reference data			Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
				Residential	Undeveloped	Total								
32	All	Simulation	Residential	15,088	9,458	24,546	86.1	38.53	37.96	53.26	52.96	2.029	68.92	0.308
			Undeveloped	9,233	100,671	109,904								
			Total	24,321	110,129	134,450								
	Most		Residential	14,340	10,206	24,546	84.99	41.58	41.02	49.52	49.24	1.770	64.54	0.310
			Undeveloped	9,981	99,923	109,904								
			Total	24,321	110,129	134,450								
	Many		Residential	14,386	10,160	24,546	85.05	41.39	40.85	49.75	49.47	1.850	64.49	0.305
			Undeveloped	9,935	99,969	109,904								
			Total	24,321	110,129	134,450								
	Half		Residential	15,092	9,454	24,546	86.1	38.52	37.95	53.28	52.98	1.858	64.60	0.306
			Undeveloped	9,229	100,675	109,904								
			Total	24,321	110,129	134,450								
	Some		Residential	15,171	9,375	24,546	86.22	38.19	37.62	53.67	53.37	1.883	64.31	0.310
			Undeveloped	9,150	100,754	109,904								
			Total	24,321	110,129	134,450								
	Few		Residential	15,099	9,447	24,546	86.11	38.49	37.92	53.31	53.01	1.929	62.88	0.309
			Undeveloped	9,222	100,682	109,904								
			Total	24,321	110,129	134,450								
	At least one		Residential	14,195	10,351	24,546	84.77	42.17	41.63	48.79	48.52	2.989	76.88	0.366
			Undeveloped	10,126	99,778	109,904								
			Total	24,321	110,129	134,450								

Order	Linguistic quantifier	Simulation		Reference data			Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
				Residential	Undeveloped	Total								
64	All	Simulation	Residential	14,787	9,759	24,546	85.65	39.76	39.2	51.75	51.46	2.124	68.82	0.320
			Undeveloped	9,534	100,370	109,904								
			Total	24,321	110,129	134,450								
	Most		Residential	14,271	10,275	24,546	84.88	41.86	41.32	49.17	48.9	1.725	64.60	0.315
			Undeveloped	10,050	99,854	109,904								
			Total	24,321	110,129	134,450								
	Many		Residential	14,767	9,779	24,546	85.62	39.84	39.28	51.65	51.36	1.688	64.51	0.312
			Undeveloped	9,554	100,350	109,904								
			Total	24,321	110,129	134,450								
	Half		Residential	14,243	10,303	24,546	84.84	41.97	41.44	49.03	48.76	1.790	64.98	0.317
			Undeveloped	10,078	99,826	109,904								
			Total	24,321	110,129	134,450								
	Some		Residential	14,270	10,276	24,546	84.88	41.86	41.33	49.17	48.89	1.891	64.85	0.318
			Undeveloped	10,051	99,853	109,904								
			Total	24,321	110,129	134,450								
	Few		Residential	14,460	10,086	24,546	85.16	41.09	40.55	50.12	49.84	1.905	62.68	0.318
			Undeveloped	9,861	100,043	109,904								
			Total	24,321	110,129	134,450								
	At least one		Residential	14,159	10,387	24,546	84.72	42.32	41.78	48.61	48.34	2.671	76.79	0.359
			Undeveloped	10,162	99,742	109,904								
			Total	24,321	110,129	134,450								

Order	Linguistic quantifier			Reference data: Residential	Reference data: Undeveloped	Reference data: Total	Overall accuracy	Error of Commission	Error of omission	κ	κ_r	*MPS*	*AI*	*ANN*
Global	All	Simulation	Residential	15,114	9,432	24,546	86.14	38.43	37.86	53.39	53.09	1.783	71.89	0.316
			Undeveloped	9,207	100,697	109,904								
			Total	24,321	110,129	134,450								
	Most		Residential	14,339	10,207	24,546	84.98	41.58	41.04	58.96	49.51	1.432	64.62	0.318
			Undeveloped	9,982	99,922	109,904								
			Total	24,321	110,129	134,450								
	Many		Residential	13,744	10,802	24,546	84.1	44.01	43.49	46.53	46.27	1.512	64.68	0.318
			Undeveloped	10,577	99,327	109,904								
			Total	24,321	110,129	134,450								
	Half		Residential	15,105	9,441	24,546	86.12	38.46	37.89	53.34	53.04	1.513	66.75	0.320
			Undeveloped	9,216	100,688	109,904								
			Total	24,321	110,129	134,450								
	Some		Residential	14,624	9,922	24,546	85.41	40.42	39.87	50.94	50.65	1.624	64.87	0.322
			Undeveloped	9,697	100,207	109,904								
			Total	24,321	110,129	134,450								
	Few		Residential	13,790	10,756	24,546	84.17	43.82	43.3	46.76	46.5	1.752	62.87	0.323
			Undeveloped	10,531	99,373	109,904								
			Total	24,321	110,129	134,450								
	At least one		Residential	13,420	11,126	24,546	83.62	45.33	44.82	44.91	44.66	2.569	77.55	0.374
			Undeveloped	10,901	99,003	109,904								
			Total	24,321	110,129	134,450								

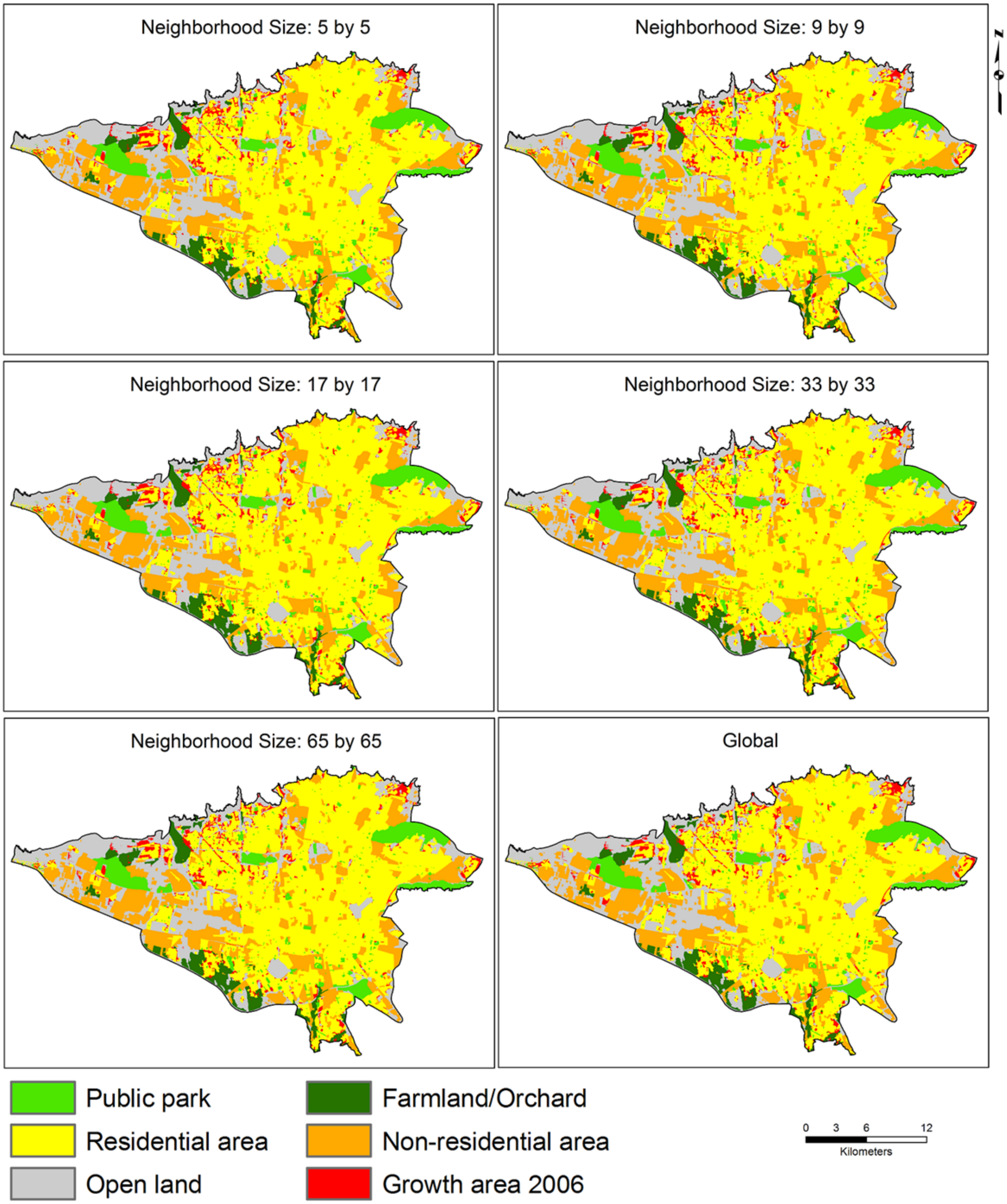


Figure C1: Land-use patterns generated by the "All" quantifier scenarios

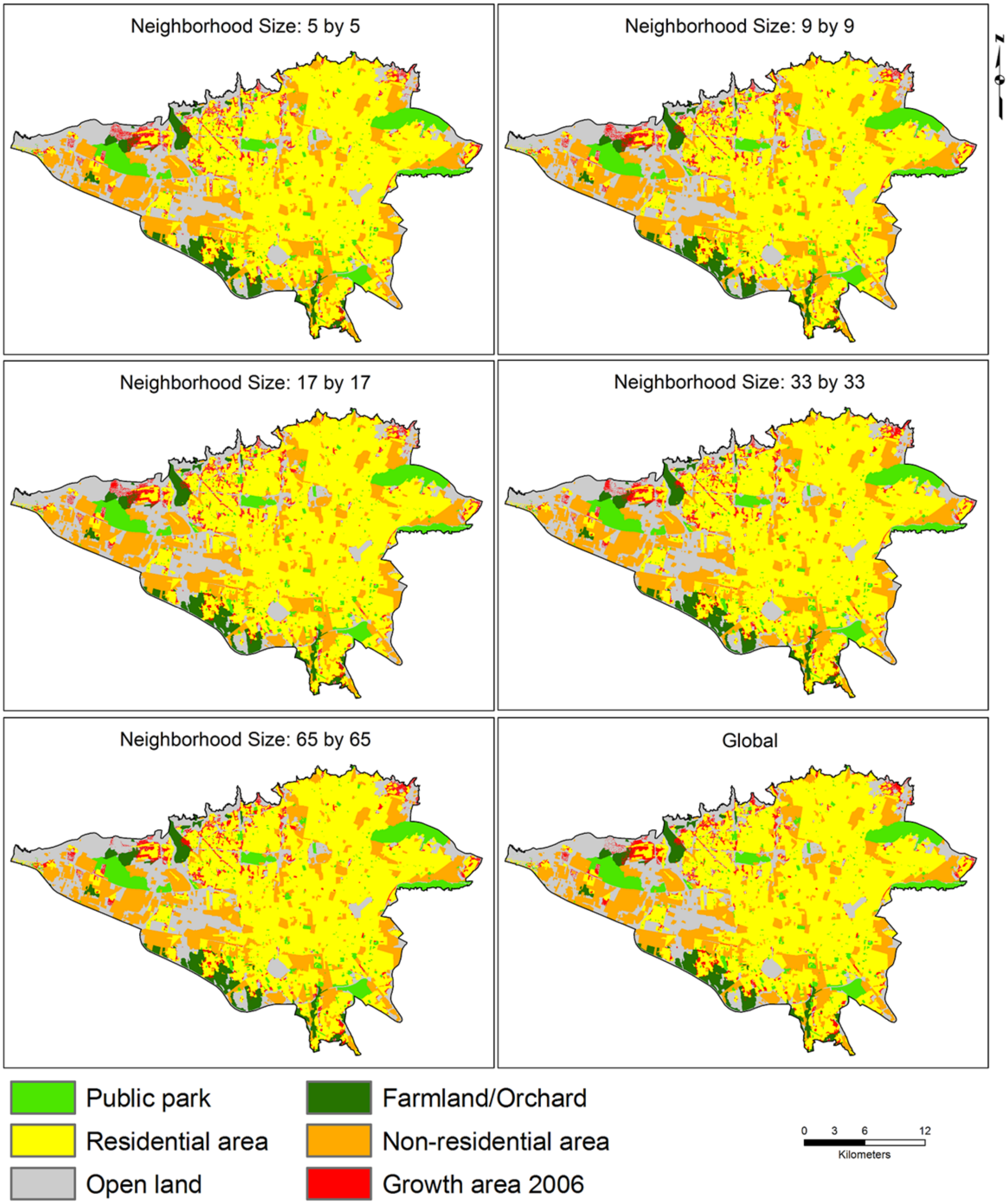


Figure C2: Land-use patterns generated by the "Most" quantifier scenarios

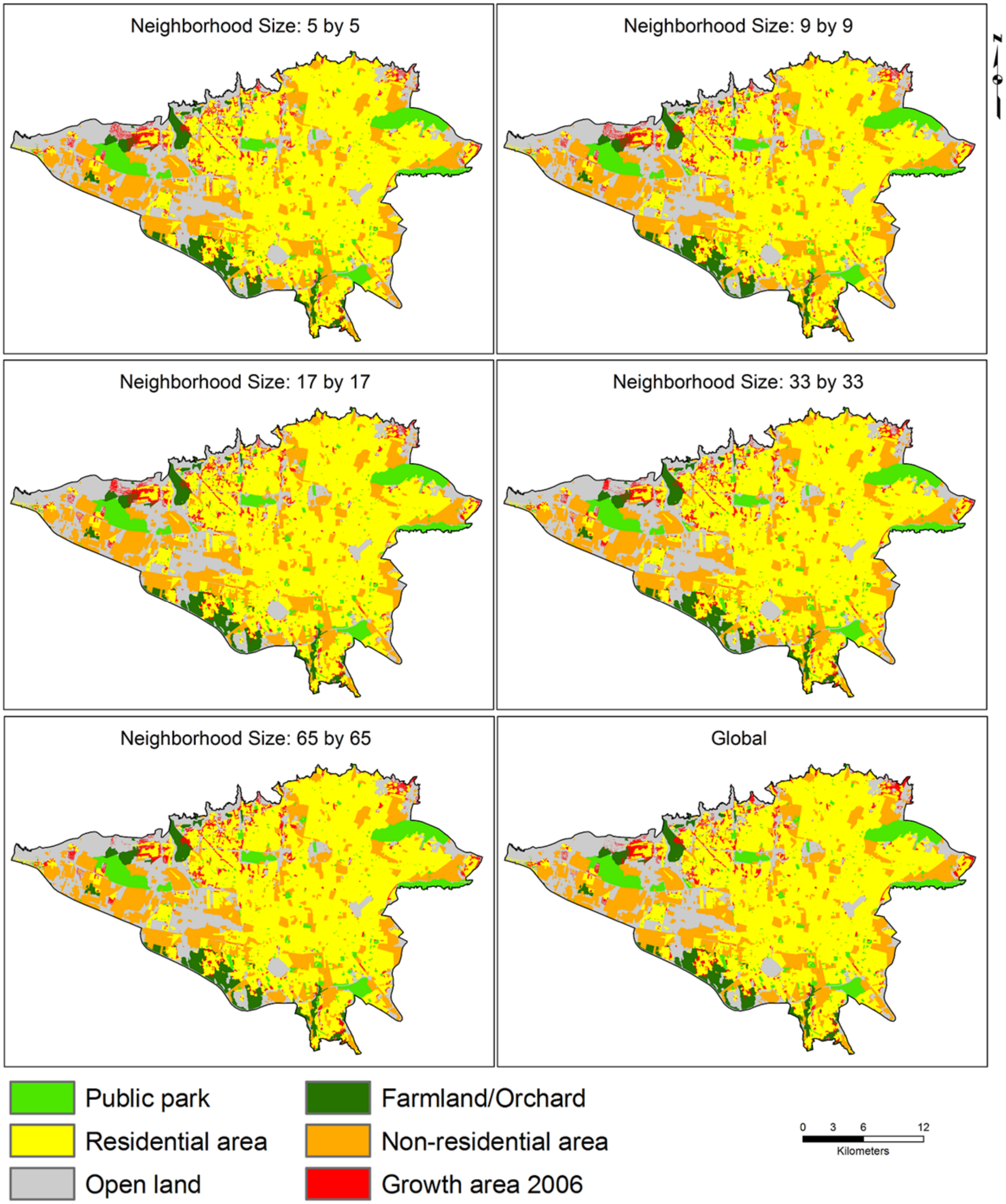


Figure C3: Land-use patterns generated by the “Many” quantifier scenarios

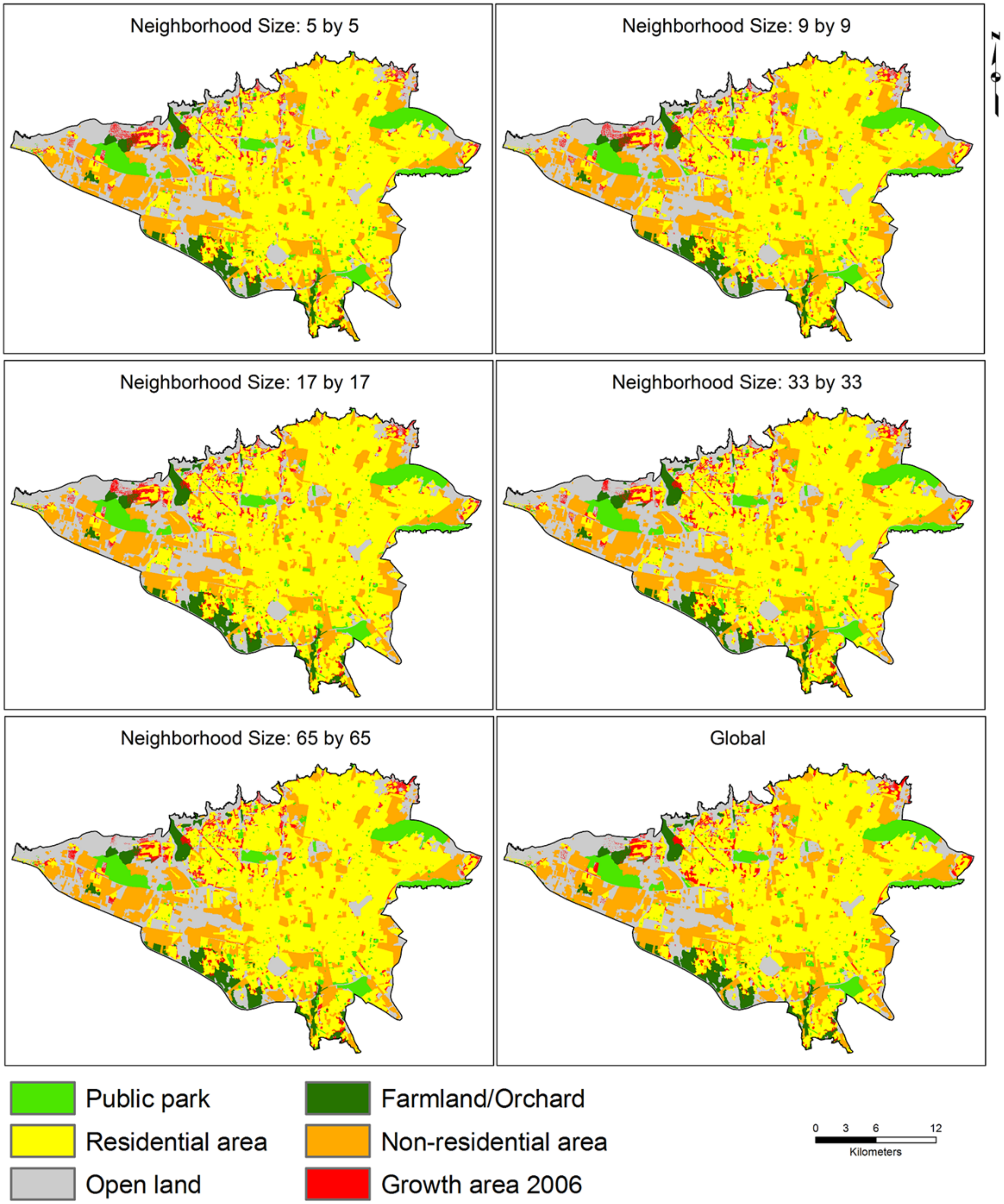


Figure C4: Land-use patterns generated by the "Half" quantifier scenarios

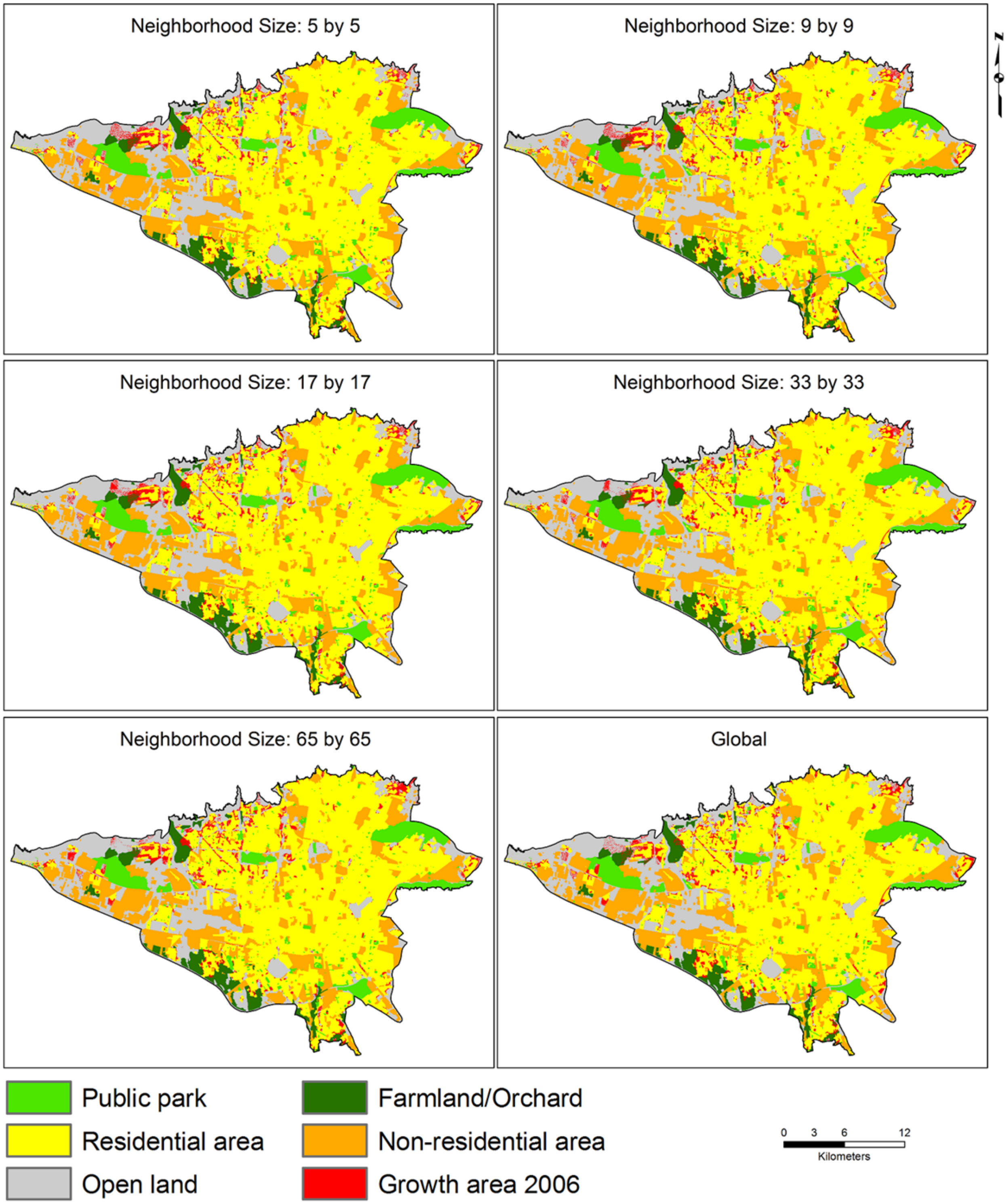


Figure C5: Land-use patterns generated by the "Some" quantifier scenarios

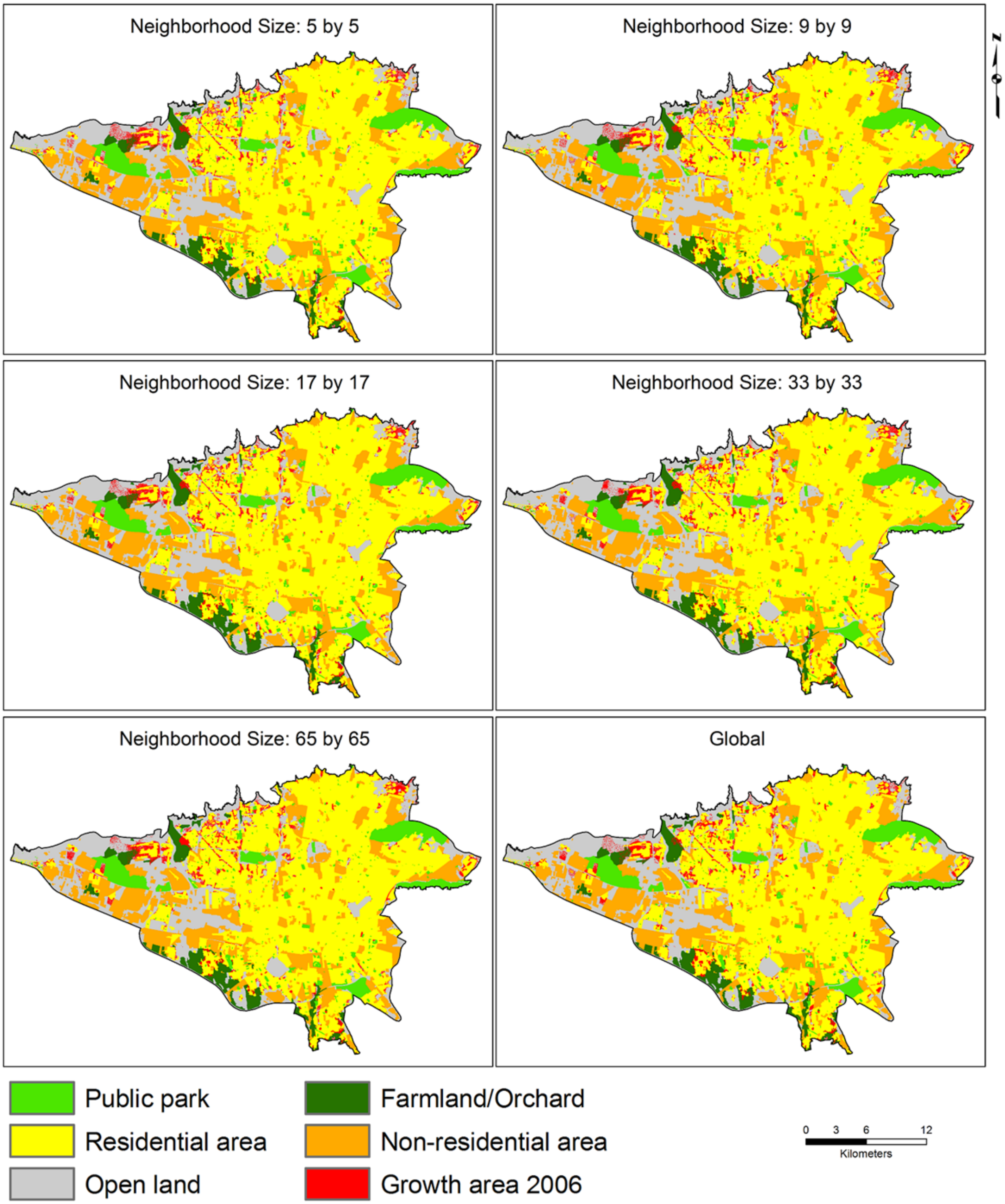


Figure C6: Land-use patterns generated by the "Few" quantifier scenarios

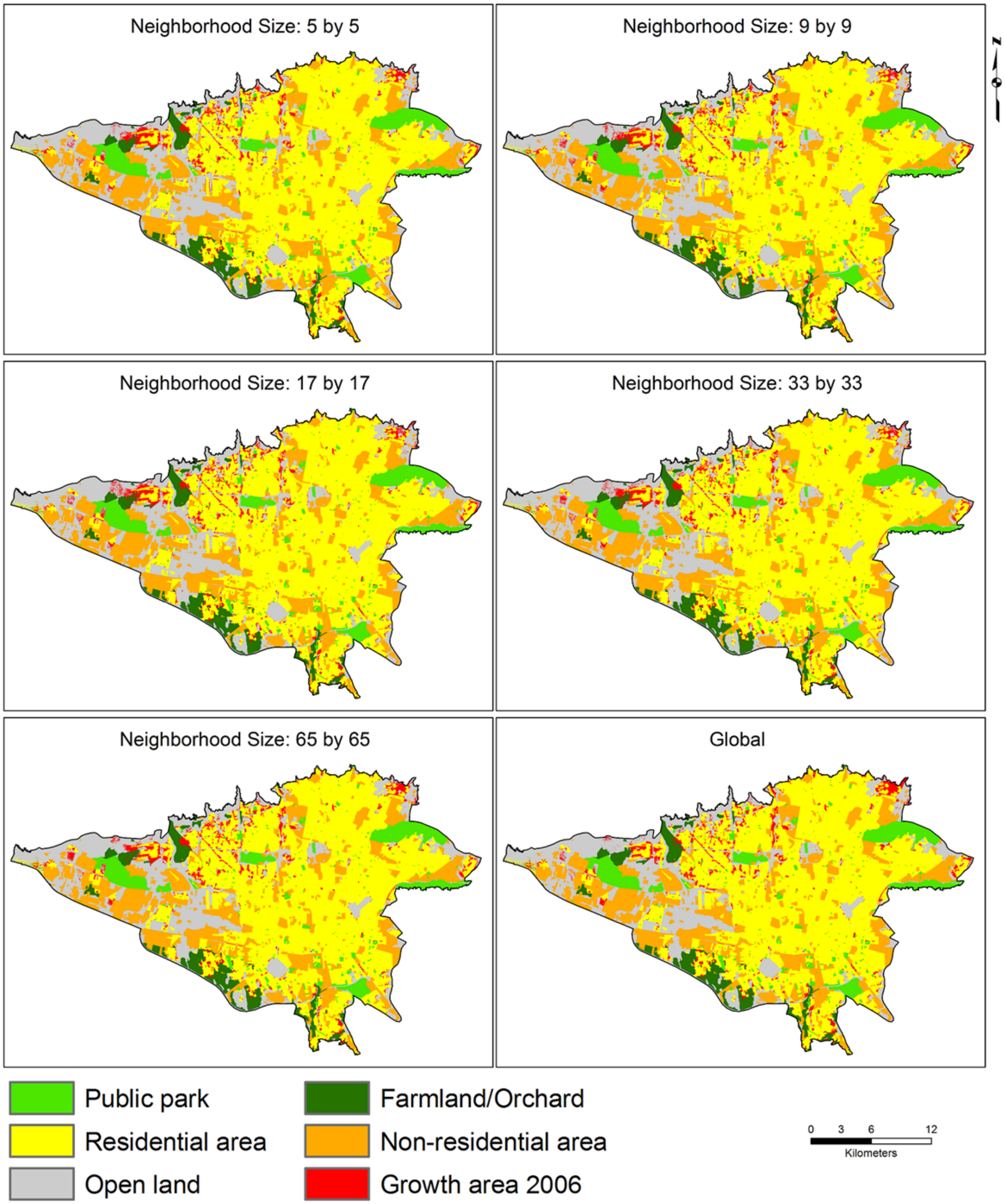


Figure C7: Land-use patterns generated by the "At least one" quantifier scenarios

Appendix D: Metadata

Data	Producer	Date published	Data type	Coordinate system	Resolution
Landsat image	USGS	1996	Raster	WGS_1984_UTM_Zone_39N	30m
Landsat image	USGS	2006	Raster	WGS_1984_UTM_Zone_39N	30m
WorldView-2 image	DigitalGlobe Foundation	2006	Raster	WGS_1984_UTM_Zone_39N	0.5m
DEM	Iranian National Cartographic Center	2006	Raster	WGS_1984_UTM_Zone_39N	30m
Education centre	Iranian National Cartographic Center	1996	Vector/Point	WGS_1984_UTM_Zone_39N	N/A
Commercial area	Iranian National Cartographic Center	1996	Vector/Polygon	WGS_1984_UTM_Zone_39N	N/A
Industrial area	Iranian National Cartographic Center	1996	Vector/Polygon	WGS_1984_UTM_Zone_39N	N/A
Major road	Iranian National Cartographic Center	1996	Vector/Polyline	WGS_1984_UTM_Zone_39N	N/A
Road	Iranian National Cartographic Center	1996	Vector/Polyline	WGS_1984_UTM_Zone_39N	N/A
Shopping centre	Iranian National Cartographic Center	1996	Vector/Point	WGS_1984_UTM_Zone_39N	N/A
Subway station	Iranian National Cartographic Center	2006	Vector/Point	WGS_1984_UTM_Zone_39N	N/A

Public park	Iranian National Cartographic Center	1996	Vector/Polygon	WGS_1984_UTM_Zone_39N	N/A
Land-use	Tehran Municipality	1996	Vector/Polygon	WGS_1984_UTM_Zone_39N	N/A
Land-use	Tehran Municipality	2006	Vector/Polygon	WGS_1984_UTM_Zone_39N	N/A
Housing price	Iranian Ministry of Roads and Urban Development	1996	Vector/Point	WGS_1984_UTM_Zone_39N	N/A
Land cost	Iranian Ministry of Roads and Urban Development	1996	Vector/Point	WGS_1984_UTM_Zone_39N	N/A

Appendix E: Curriculum Vitae

Name: Hossein Hosseini

EDUCATION

M.Sc. in Geomatics Engineering – Geographic Information System (GIS) (Aug. 2011)

K. N. Toosi University of Technology, Tehran, Iran

Dissertation: Developing the Ordered Weighted Averaging (OWA) operator in site selection

GPA:18.55/20

B.Sc. in Geomatics Engineering - Surveying (Sept. 2008)

K. N. Toosi University of Technology, Tehran, Iran

GPA:15.33/20

RESEARCH PROJECTS

PhD Project, July 2014 – Present

Advanced Spatial Analysis Course, Jan 2014 – Apr 2014

Master Project, Sep 2009 – Sep 2011

Seminar Course, July 2009 – Oct 2009

Summer Internship, Sazeh Karan Gabric Co., Tehran, Iran, July 2008 – Oct 2008

TEACHING EXPERIENCE

Teaching and Research Assistant, University of Western Ontario, London, ON, Canada, Sep 2013 – present

- Teaching Assistant for the following courses:
  - Spatial Statistics (three times)
  - Quantitative Analysis in Geography (two times)
  - Introduction to Geographic Information Science (for graduate students)
  - Geographic Information Science I
  - Location Theory and Analysis

- Private tutoring classes, June 2009 – Sep 2012
  - Taught C# to university students
  - Taught MS SQL Server to university students
  - Taught JAVA to university students
  - Taught ESRI ArcGIS to university students
  - Taught MATLAB to university students

COMPUTER SKILLS

- **Programming Languages:** Python, C#.NET, VB.NET, JAVA, C++, C, VB, Pascal, Fortran, ASP.NET, PHP, JavaScript, T-SQL, CSS
- **Markup Languages:** HTML5, XML, GML
- **Software:** ESRI ArcGIS Desktop, GeoDA, AutoCAD, SPSS, ENVI, MapInfo, Google Earth, MATLAB, Microsoft Office, Microsoft Visual Studio 2008-2010-2012, Microsoft SQL Server 2008-2012-2014, MySQL, ECLIPSE, IntelliJ IDEA, NetBeans, ZEND Studio, Adobe Dreamweaver, Notepad++, Aptana Studio
- **Libraries:** jQuery
- **Technology:** ArcObjects, ArcSDE, Mobile Programming

OTHER ACTIVITIES

Credentials:

- Certificate in C#.Net Programming Language
- Certificate in ASP.Net Programming Language
- Certificate in Microsoft SQL Server
- Certificate in Network Fundamentals
- Certificate in JAVA
- Certificate in Web Design
- Certificate in JAVASCRIPT