

2017

## Linking Climate Change and Socio-economic Impact for Long-term Urban Growth in Three Mega-cities

Qi Lu

*University of Central Florida*



Part of the [Environmental Engineering Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Lu, Qi, "Linking Climate Change and Socio-economic Impact for Long-term Urban Growth in Three Mega-cities" (2017). *Electronic Theses and Dissertations, 2004-2019*. 5538.

<https://stars.library.ucf.edu/etd/5538>



LINKING CLIMATE CHANGE AND SOCIO-ECONOMIC IMPACT  
FOR LONG-TERM URBAN GROWTH IN THREE MEGACITIES

by

QI LU

B.A. University of Illinois at Chicago, 2015

A thesis submitted in partial fulfillment of the requirements  
for the degree of Master of Science  
in the Department of Civil, Environmental, and Construction Engineering  
in the College of Engineering and Computer Science  
at the University of Central Florida  
Orlando, Florida

Summer Term  
2017

Major Professor: Ni-Bin Chang

© 2017 Qi Lu

## **ABSTRACT**

Urbanization has become a global trend under the impact of population growth, socio-economic development, and globalization. However, the interactions between climate change and urban growth in the context of economic geography are unclear due to missing links in between the recent planning megacities. This study aims to conduct a multi-temporal change analysis of land use and land cover in New York City, City of London, and Beijing using a cellular automata-based Markov chain model collaborating with fuzzy set theory and multi-criteria evaluation to predict the city's future land use changes for 2030 and 2050 under the background of climate change.

To determine future natural forcing impacts on land use in these megacities, the study highlighted the need for integrating spatiotemporal modeling analyses, such as Statistical Downscale Modeling (SDSM) driven by climate change, and geospatial intelligence techniques, such as remote sensing and geographical information system, in support of urban growth assessment. These SDSM findings along with current land use policies and socio-economic impact were included as either factors or constraints in a cellular automata-based Markov Chain model to simulate and predict land use changes in megacities for 2030 and 2050. Urban expansion is expected in these megacities given the assumption of stationarity in urban growth process, although climate change impacts the land use changes and management. More land use protection should be addressed in order to alleviate the impact of climate change.

## **ACKNOWLEDGMENTS**

The research is in partnership with University of Exeter in the U.K. and Tsinghua University in China and supported by the Global Innovation Initiative (British Council GII206), funded by the British Council and the Department for Business, Innovation and Skills. The author wishes to thank her Thesis Committees, Dr. Ni-Bin Chang, Dr. Martin Wanielista and Dr. Kelly Kibler. The author also wish to thank her co-workers for their assistance in this work including Ms. Sanaz Imen and Mr. Justin Joyce.

## TABLE OF CONTENTS

LIST OF FIGURES	ix
LIST OF TABLES	xii
CHAPTER 1 INTRODUCTION	1
1.1. Urbanization and Urban Growth Model (UGM).....	1
1.2. Population Growth and World Urbanization Trends .....	3
1.3. The Growth of Megacities.....	4
1.4. Modeling Urban Growth .....	6
1.5. Cellular-Automata (CA) Modeling .....	7
1.6. Essential Features of Megacities and Applications of CA-based Models .....	11
1.7. Perspective and Challenges .....	13
1.8. Limitations .....	14
1.9. References .....	15
CHAPTER 2 LINKING SOCIO-ECONOMIC DEVELOPMENT, SEA LEVEL RISE, AND CLIMATE CHANGE IMPACTS ON URBAN GROWTH IN NEW YORK CITY WITH A FUZZY CELLULAR AUTOMATA-BASED MARKOV CHAIN MODEL	22
2.1. Introduction .....	22
2.2. Study Area.....	25
2.3. Software and Data Collection .....	26
2.3.1. Statistical Downscaling Model (SDSM).....	26

2.3.2.	CA-based MC Model and Data Collection .....	27
2.4.	Methodology .....	30
2.4.1.	Statistical Downscaling Model (SDSM).....	30
2.4.2.	CA-MC Model - Land Use Change Prediction.....	31
2.5.	Results and discussion.....	44
2.5.1.	Statistical Downscaling Model (SDSM).....	44
2.5.2.	Results of the Fuzzy CA – based MC Model .....	47
2.6.	Conclusion.....	57
2.7.	References .....	58
CHAPTER 3 EXPLORING THE POTENTIAL CLIMATE CHANGE IMPACT ON URBAN GROWTH IN LONDON BY A CELLULAR AUTOMATA-BASED MARKOV CHAIN MODEL .....		67
3.1.	Introduction .....	67
3.2.	Study area.....	70
3.3.	Data and Software .....	71
3.4.	Methodology .....	76
3.4.1.	Flood impact concern & SDSM.....	76
3.4.2.	CA-based MC model .....	80
3.5.	Results .....	90
3.5.1.	SDSM calibration and validation.....	90

3.5.2.	SDSM future projections (2030 & 2050).....	94
3.5.3.	TerrSet Transition probability matrices .....	95
3.5.4.	TerrSet suitability maps .....	96
3.5.5.	TerrSet validation and prediction.....	100
3.6.	Discussion .....	102
3.7.	Conclusion.....	104
3.8.	References .....	105
CHAPTER 4 PREDICTING LONG-TERM LAND USE CHANGES IN BEIJING (CHINA) WITH CONSIDERATION OF CLIMATE CHANGE, SOCIO-ECONOMIC DEVELOPMENT, AND POLICY-ORIENTED FACTORS		110
4.1.	Introduction .....	110
4.2.	Study Area.....	116
4.3.	Applied Software and Data Collection.....	117
4.4.	Methodology .....	119
4.4.1.	Statistical Downscaling Model (SDSM).....	119
4.4.2.	CA-MC Model – LULC Change Prediction .....	125
4.5.	Results .....	133
4.5.1.	SDSM Calibration, Validation, and Prediction.....	133
4.5.2.	CA-based MC model .....	137
4.6.	Discussion .....	147



4.7. Conclusion.....	150
4.8. References .....	151
CHAPTER 5 CONCLUSIONS	154
APPENDIX: ACCEPTED MATERIAL FOR PUBLICATION	155

## LIST OF FIGURES

<b>Figure 1-1</b> Percentage of population living in urban areas in different regions of the world (1990-2050) – Source: (Data source: United Nations, 2014).....	4
<b>Figure 1-2.</b> Average annual rate of population change in percent during 2010-2015 in the world’s mega-cities (Data source: United Nations, 2014).....	6
<b>Figure 1-3.</b> Levels for modelling the urban growth (Cheng, 2003).....	7
<b>Figure 2-1.</b> Study framework (*CA: cellular automata; MCE-WLC: multi-criteria evaluation-weighted linear combination) .....	33
<b>Figure 2-2.</b> Predicted SLR impact area in 2050 in NYC (NOAA, 2011) .....	36
<b>Figure 2-3.</b> High risk flooding area in NYC (FEMA 2016) .....	37
<b>Figure 2-4.</b> Monthly Mean Rainfall (Observed vs. SDSM) for each station from Jan 1990-Jan 2000: (a) JFK Airport, (c) Central Park and (e) LaGuardia Airport and monthly variance (observed vs. SDSM) for (b) JFK Airport, (d) Central Park and (f) LaGuardia Airport.....	46
<b>Figure 2-5.</b> Suitability maps for each land use class with continuous suitability scaled from 0 to 255.....	52
<b>Figure 2-6.</b> Predicted land use map in 2030 (a) and 2050 (b) with difference circled .....	54
<b>Figure 2-7</b> Urban changes under urban growth and SLR simultaneously in (a) 2030 and (b) 2050 .....	56
<b>Figure 3-1.</b> London land use in 2000 with re-classified land use classes .....	74
<b>Figure 3-2.</b> 2.5° latitude x 3.75° longitude grid system used by Hadley Centre’s coupled ocean–atmosphere GCMs and NCEP. (Source: Wilby et al., 2002).....	78

<b>Figure 3-3.</b> Study framework; *CA: cellular automata; MCE-WLC: multi-criteria evaluation-weighted linear combination; MC: Markov Chain .....	81
<b>Figure 3-4.</b> Recorded flood area from 1706 to 2015 in London .....	85
<b>Figure 3-5.</b> Land use classes' interaction relationship associated with their factors .....	86
<b>Figure 3-6.</b> Observed vs. SDSM during validation period (1976-1990) for Longford Station ...	92
<b>Figure 3-7.</b> Observed vs. SDSM during validation period (1976-1990) for Woodford Station ..	92
<b>Figure 3-8.</b> Observed vs. SDSM during validation period (1976-1990) for Hogsmill Valley Station with month of June as an outlier (a) and without month of June as an outlier (b) .....	93
<b>Figure 3-9.</b> Historical Rainfall (a) and SDSM Projected Rainfall (b) and (c) for Woodford Station for the years 2030 and 2050 using HADCM3 A2 scenario .....	94
<b>Figure 3-10.</b> Suitability maps for water, urban, open space, forest, and agriculture (0 indicates no suitability, and 1 indicates the highest suitability).....	99
<b>Figure 3-11.</b> Predicted LULC maps in 2030 and 2050 with difference in comparison highlighted by red squares .....	100
<b>Figure 4-1.</b> Geographical location of Beijing in China.....	117
<b>Figure 4-2.</b> Collected LULC data (a) 2000, (b) 2005, and (c) 2010 .....	118
<b>Figure 4-3.</b> Annual rainfall trends for period of 1990-2015 for (a) North station, (b) South station, (c) East station and (d) West station .....	121
<b>Figure 4-4.</b> Flowchart of the integrated urban growth model; *CA: cellular automata; MCE-WLC: multi-criteria evaluation-weighted linear combination.....	126
<b>Figure 4-5.</b> Geographical locations of Beijing, Tongzhou district, Tianjin, and Xiong'an connected by the road systems .....	127
<b>Figure 4-6.</b> Flood records location in Beijing from 2013 to 2017 (source: NASA, 2017).....	128

<b>Figure 4-7.</b> Validation results for period of January 2005-January 2010 for (a) North station, (b) South station, (c) East station and (d) West station .....	135
<b>Figure 4-8.</b> Projected annual rainfall for year 2030 at (a) North station, (b) South station, (c) East station and (d) West station.....	136
<b>Figure 4-9.</b> Projected annual rainfall for year 2050 at (a) North station, (b) South station, (c) East station and (d) West station.....	137
<b>Figure 4-10.</b> Observed maps (a) 2000 and (b) 2010 vs. scenario 1 projections (c) 2030 and (d) 2050.....	141
<b>Figure 4-11.</b> Observed maps (a) 2000 and (b) 2010 vs. scenario 2 projections (c) 2030 and (d) 2050.....	142
<b>Figure 4-12.</b> Observed maps (a) 2000 and (b) 2010 vs. scenario 3 projections (c) 2030 and (d) 2050.....	143
<b>Figure 4-13.</b> Observed maps (a) 2000 and (b) 2010 vs. scenario 4 projections (c) 2030 and (d) 2050.....	144

## LIST OF TABLES

<b>Table 1-1.</b> Types of CAs and the definition of each type (Krishna et al., 2012) .....	9
<b>Table 1-2.</b> List of previous CA-based models applied in mega-cities .....	11
<b>Table 2-1</b> Rainfall Stations with Continuous Period (30+yrs) of Record in NYC .....	27
<b>Table 2-2.</b> LULC data summary in percentage (%) .....	28
<b>Table 2-3.</b> Data sources and description .....	29
<b>Table 2-4.</b> Constraints and factors and their related sources or policies .....	38
<b>Table 2-5.</b> Calibration Results for Daily Rainfall (1961-1981) with observed predictors.....	45
<b>Table 2-6.</b> Transition probability matrix in MC calculated using land use maps 2001 and 2006 .....	48
<b>Table 2-7.</b> Extracted weights based on AHP of urban land use class .....	50
<b>Table 2-8.</b> Calibration results based on the observed and predicted land use maps in 2006 .....	53
<b>Table 2-9.</b> Land use data comparison among 2011, 2030, and 2050 in percentage .....	55
<b>Table 3-1.</b> Applied dataset and source .....	71
<b>Table 3-2.</b> Reclassified land use class area and difference comparison (unit: hectares) .....	75
<b>Table 3-3.</b> Rainfall Stations in London with Continuous Period of Daily Rainfall Record .....	77
<b>Table 3-4.</b> Factors and constraints applied in the model.....	87
<b>Table 3-5.</b> Correlation ( $R^2$ ) between observed rainfall and SDSM predictions at Longford, Woodford and Hogsmill stations for the period 1961-1975 .....	91
<b>Table 3-6.</b> Transition probability matrix between 2000 and 2006 for predicting LULC map in 2012 .....	95
<b>Table 3-7.</b> Factors and weights for urban suitability map ( $CR = 0.03 < 0.1$ ) .....	97
<b>Table 3-8.</b> AUC values for all land use classes.....	99

<b>Table 3-9.</b> Land use area comparison between 2012, 2030, and 2050 .....	101
<b>Table 3-10.</b> The scenario comparison between the inclusion or exclusion of flood impact for land use change prediction in 2030 and 2050.....	103
<b>Table 4-1.</b> Summary of the previous studies related to the simulation and prediction of LULC change in Beijing area.....	111
<b>Table 4-2.</b> Collected data for LULC changes prediction .....	119
<b>Table 4-3.</b> Locations of Daily Beijing Rainfall Data .....	121
<b>Table 4-4.</b> GCM climate predictors used for each station.....	123
<b>Table 4-5.</b> Applied driving factors and constraints .....	129
<b>Table 4-6.</b> Monthly Calibration Results for Daily Rainfall (1990-2000) with daily GCM predictors .....	134
<b>Table 4-7.</b> Transition probability matrix for predicting LULC map in 2010.....	138
<b>Table 4-8.</b> Transition probability matrix for predicting LULC map in 2030.....	138
<b>Table 4-9.</b> Transition probability matrix for predicting LULC map in 2050.....	139
<b>Table 4-10.</b> Comparison of validation results between scenarios 1, 2, 3, and 4 .....	140
<b>Table 4-11.</b> Predicted results of Scenario 1.....	145
<b>Table 4-12.</b> Predicted results of Scenario 2.....	145
<b>Table 4-13.</b> Predicted results of Scenario 3.....	146
<b>Table 4-14.</b> Predicted results of Scenario 4.....	146
<b>Table 4-15.</b> Comparison between previous studies and the current study for Beijing metropolitan area .....	148

# CHAPTER 1 INTRODUCTION

## 1.1. Urbanization and Urban Growth Model (UGM)

The world is undergoing the largest wave of *urban growth* in human history. Around 50% of the world's population currently lives in an urban environment, and more than 70% of world's population is expected to live in cities by 2050 (United Nation, 2014). According to the UN-Habitat Global Report, the world's population is expected to grow by more than 2 billion people by 2050, resulting in a total global population of 9.2 billion (PBL, 2014), which may further accelerate urban growth. Urban metabolism is intimately tied to a few networked infrastructure systems, including drinking water, wastewater, stormwater, waste management, power, transportation, and telecommunication. A positive consequence of urbanization may be stronger economic growth due to the higher concentration of activities leading to economies of scale. Continuous population growth and migration during the urbanization process will in turn contribute to higher demands on water, energy, and environmental resources and speed up economic development. The interdependence among those infrastructure systems to meet the unprecedented demand of urban population complicates the managerial framework and policy-making in the context of urban growth.

The term of megacity was defined by the United Nations as a metropolitan area with at least 10 million residents (United Nation, 2014). The phenomenon of rapid urbanization, especially in megacities, was recognized in the 20<sup>th</sup> century, when the United Nations Conference on Environment and Development (UNCED) suggested working towards finding a solution to environmental problems caused by urbanization and economic growth (Lavalle et al., 2001; International Federation of Surveyors, 2010). The rapid growth of mega-cities has resulted in high

urban densities, informal development, unsustainable land-use, commuting problems, food, water, and energy insecurity, lack of basic services, poor natural hazards management, increase crime and social inequalities, environmental degradation, climate change, and inefficient administration (International Federation of Surveyors, 2010).

The global process of urbanization incorporates many components that vary with time. Hence, planning sustainable urban development requires an understanding of urban dynamics. Advances in technology and developments can provide information about the trends of historical growth patterns, which can be used in an urban growth model (UGM) to provide understanding of probable future developments. Among different types of UGMs, the ability of a cellular automata (CA)-based model in dealing with extreme complex content of urban dynamic elements makes it a popular tool among urban modelers (Batty, 2005). Its ability is not limited to the spatial land-use variation, but also includes intensive human activities such as population density and immigration (Croes et al., 2012). Hence, this type of CA-model is able to incorporate both spatial and temporal dimensions of urban processes (Sante et al., 2010). CA-based models also have the ability to generate new urban growth patterns and land-use changes for the future. This makes CA-based models suitable for simulating and predicting the intertwined urban growth processes, specifically for future megacities.

However, a disadvantage with CA-based models is that neighborhood selection and transition rule identification had to be determined by human experts. In recent years, several attempts have been made in literatures to automate the identification of CA-based models. For instance, studies focused on using evolutionary algorithms such as a genetic algorithm (GA) to automate CA design (Yang and Billings, 2000; Chavoya and Duthen, 2006; Maeda and Sakama, 2007). Since the

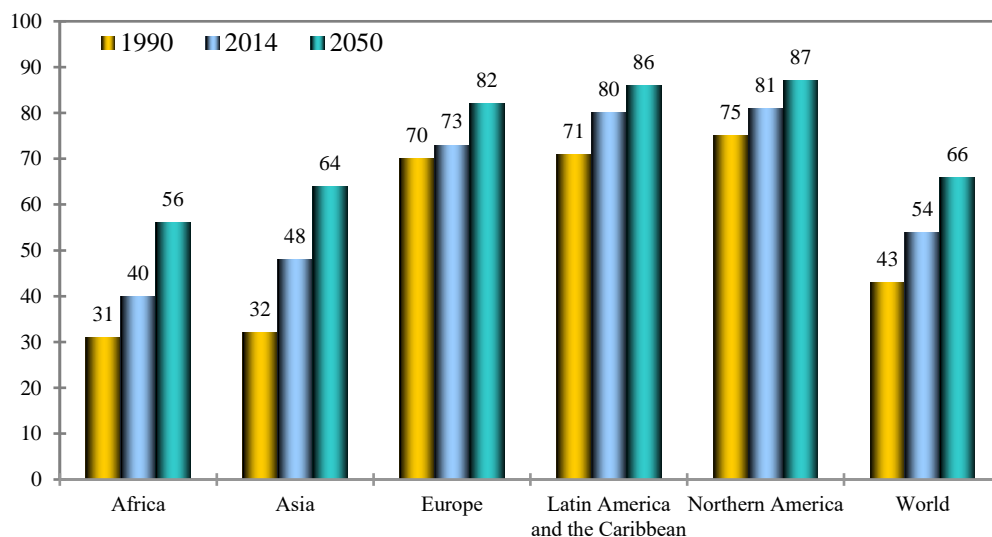


application of a GA is time-consuming in real-world applications, Adamatzky (1994) proposed a method to extract rules faster from observed data. However, the drawback of the proposed method is that it may extract redundant cells. Zhao and Billings (2006) developed a new algorithm based on mutual information to detect the initial temporal and spatial range in the identification of CA. Later, Sun et al. (2011) developed a new algorithm that was faster and more space efficient than the method proposed by Zhao and Billing (2006). More recently, Placzek (2014) applied data mining techniques based on rough set theory to select neighborhood and determine updated rules for CA. Despite the literatures, behavior of a CA-based model at different scales requires more studies (Pan et al., 2010). Pan et al. (2010) showed that the variation of scale in a CA-based model affects the model's behavior. On one hand, small cell size and neighborhood lead to incorrect expression of land-use transition. On the other hand, increasing in the neighborhood size with ring shape decreases the precision of simulation (Pan et al., 2010). Consequently, the initial step in modeling urban growth in megacities is knowing the transition rule and the scale of each existing CA-based model to select the suitable structure to achieve the highest precision in simulation. For this purpose, the main objective of this paper is to provide comprehensive information on the existing CA-based models and compare these models from a wealth of aspects to help users in selecting a suitable one with appropriate software support.

## 1.2. Population Growth and World Urbanization Trends

In 1950, one-third of the world's population resided in urban areas. Since 2007, the global urban population has exceeded the global rural population and about half of the world's population resided in urban areas in 2014 (United Nations, 2014). By 2050, it is projected that the distribution of global rural-urban population will be opposite of that in the mid-twentieth century and two-

thirds of the world's population will be living in urban regions (United Nation, 2014). In 2014, Northern America, Latin America and the Caribbean, and Europe were the most urbanized regions. At the same time, Africa and Asia were less urbanized (United Nations, 2014). As shown in Figure 1-1, although it is projected that Africa and Asia have been urbanizing faster than other parts of the world by the year 2050, it is expected that these two regions will be less urbanized than other regions of the world by that time (United Nations, 2014). It has also been projected that urban growth will occur in smaller cities and towns rather than megacities by the year 2025 (Cohen, 2004).



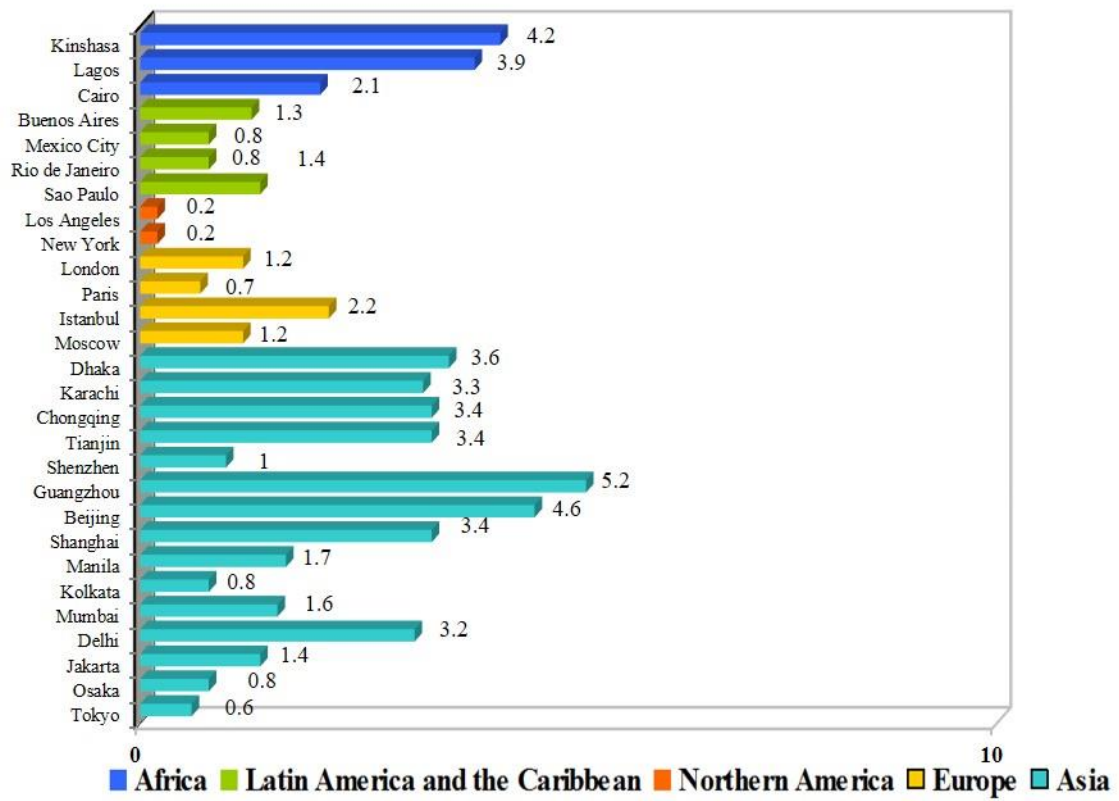
**Figure 1-1** Percentage of population living in urban areas in different regions of the world (1990-2050) – Source: (Data source: United Nations, 2014)

### 1.3. The Growth of Megacities

According to the 2014 world urbanization prospects report of the United Nation's department of economic and social affairs, there are 28 megacities which are the home of 453 million people

(Figure 1-2). The top 10 largest megacities in the world are Tokyo, Delhi, Shanghai, Mexico City, São Paulo, Mumbai, Osaka, Beijing, New York, and Cairo. As shown in Figure 2, most of the megacities are located in Asia. Six megacities belong to China and it is projected that China will add one more megacity by 2030 (United Nations, 2014). In addition to three current megacities in India, four more megacities are projected in this country by 2030 (United Nations, 2014). As shown in Figure 1-2, there are seven other megacities found in Asia, outside of China and India.

In Africa, there are only three megacities, and it has been projected that three more megacities will be added by 2030 (United Nations, 2014). As shown in Figure 2, Asia and Africa had the highest average annual rate of population growth during 2010-2015, and in the Latin America and the Caribbean, there are only four megacities. The total number of megacities in Latin America and the Caribbean is expected to increase to 6 by 2030 (United Nations, 2014). In Northern America, the current number of megacities is equal to two, and it is expected that the population of New York will decline by 2030 (United Nations, 2014). Planning the rapid growth of megacities has been a grand challenge in urban modeling worldwide. Finding the driving forces behind urban growth is of great importance to deal with this challenge.

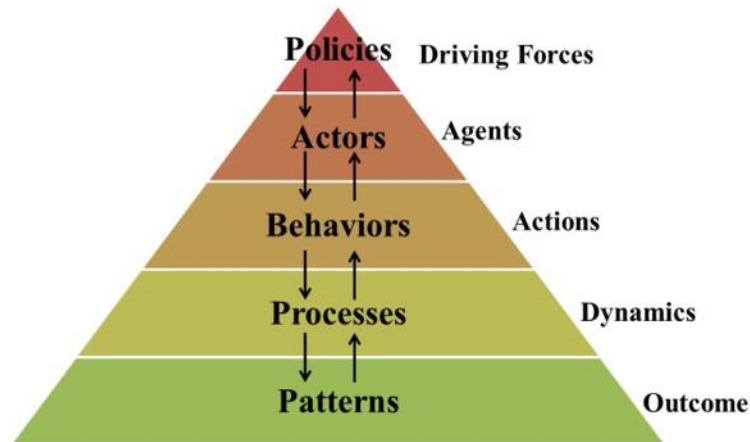


**Figure 1-2.** Average annual rate of population change in percent during 2010-2015 in the world's megacities (Data source: United Nations, 2014)

#### 1.4. Modeling Urban Growth

The aim of modeling urban growth is to support the planning of urban development and create sustainable urban growth management (Cheng, 2003). However, urban development encompasses various physical and socioeconomic factors, and it is considered as a complex system because of the unknown amount of aforementioned factors and their complex interactions (Li, 2014). Sources of complexity in urban growth can be categorized as: (1) spatial complexity, (2) temporal complexity, and (3) decision making complexity (Cheng et al., 2003). Managing complexity requires a systematic perspective. A systematic perspective to understand urban growth consists

of five levels: policy, actor, behavior, process, and pattern (Figure 1-3). Policy refers to the level confirmed to be the most effective factor and driving force of urban growth on the macro scale.



**Figure 1-3.** Levels for modelling the urban growth (Cheng, 2003)

Actor refers to the agent of behaviors, and behavior represents the decision making of actors. Processes demonstrate the sequence of changes in space (i.e. spatial process) and time (i.e. temporal process). Pattern is the observable outcome and it can be viewed as: (1) urban growth system which includes development units, and (2) planned urban system, developed non-urban system, and developed urban system (Cheng, 2003). To model the complexity of urban growth, different models including CA-based modeling, agent-based modeling, artificial neural network, fractal modeling, linear-logistic modeling, and decision tree modeling have been developed. A brief description of each aforementioned model is presented.

### 1.5. Cellular-Automata (CA) Modeling

The theory of CA was introduced by John von Neumann in the late 1940s through his work on developing an abstract model of self-reproduction in biology. In 1951, he applied Stanislaw

Ulam's suggestion to simplify his model and develop a 2D CA. CA-based models have attracted the attention of researchers because of their ability to model and visualize the complex processes that are spatially distributed from simple rules (Takeyama and Couclelis, 1997). The elements of a basic CA include lattice, cell states, neighborhood, transition rules, and time. Lattice is described as the space on which the automaton exists (Torrens, 2000). The two dimensional lattice is generally applied to model urban growth. In addition to a regular fashion lattice, some irregular lattices have been applied for urban modeling as well (Jiao, 2003). A CA includes a set of similar elements called cells. Each cell occupies a node of a regular, discrete, infinite spatial network (Bandini et al., 2001). Each cell represents a state from a finite set. In discrete time steps, the automaton evolves by changing the states of all its cells based on a local rule at each time step. In urban CA, states can be: (a) binary values (e.g. urban, non-urban), (b) qualitative values (e.g. different land uses), (c) quantitative values (e.g. population density), and (d) a vector of several attributes (Sante et al., 2010). Another element is the neighborhood around the automata (Torrens, 2000). The neighborhood is defined as the radius of affected area (Batty et al., 1999). Radius means the number of cells left and right of the central cell. There are several types of neighborhood in 2D CA such as: Von Neumann neighborhood, Moore Neighborhood, Margolus neighborhood, unaligned rectangular neighborhood, Hexagonal neighborhood, and small unaligned hexagonal neighborhood. Among these types, Moore, Van Neumann, and extended Moore-Van Neumann are frequently used in urban CA models. The engine of changes in a CA is transition rules (Jiao, 2003). It defines the behavior of the automaton (Torrens, 2000). The local transition rule depends on a small number of adjacent cells and the cell itself. Therefore, the state of each cell is updated at each time step based on its current state and the state of its neighboring cells in the preceding step according to a set of transition rules (Krishna et al., 2012). Lastly, time is the temporal space in

which automata exists (Torrens, 2000). Time in CA is discrete and can be different from one CA model to another. In the case of a longer time gap, the CA evolution will be more discrete (Jiao, 2003). Different types of CAs are summarized in Table 1-1.

**Table 1-1.** Types of CAs and the definition of each type (Krishna et al., 2012)

<b>Type</b>	<b>Definition</b>	<b>Reference</b>
Linear CA	All cells of CA have linear rules that include only XOR <sup>1</sup> (Exclusive-OR) logic	Nandi et al. 1994
Complement CA	All cells of CA have complement rule that includes only XNOR <sup>2</sup> logic.	Nandi et al. 1994
Additive CA	Cells of CA have rules that are combination of XOR and XNOR logic	Nandi et al. 1994
Uniform CA	All cells of CA have the same rule	Nandi et al. 1994
Hybrid CA	All cells of CA have different rules	Anghelescu et al. 2007
Null Boundary CA	Both left and right neighbor of the leftmost and rightmost terminal cell are connected to logic 0	Kundu et al. 2008
Periodic Boundary CA	The rightmost cell as the left neighbor of leftmost cell and similarly, the leftmost cell is considered as the right neighbor of rightmost cell	Anghelescu et al. 2007
Reversible CA	CA always return to its initial state	Seredynski et al. 2005

Type	Definition	Reference
Non-Linear CA	Non-linear CA uses all possible logic available in CA	Das et al. 2006
Fuzzy CA	CA employed with Fuzzy logic.	Maji et al. 2004

<sup>1</sup>XOR: The output is "true" if either, but not both, of the inputs are "true"

<sup>2</sup>XNOR: The output is true, if the inputs are the same, and “False” is the inputs are different

CAs have several advantages that make them a suitable candidate for modeling urban growth such as: (1) capability of dealing with spatial phenomena, (2) highly decentralized, (3) affinity with geographic information system and remote sensing, (4) capability of handling fine-scale dynamics with computational efficiency, (5) equal attention to space, time, and system attributes, (6) flexibility to allow multiple timescales to be represented in the simulation, (7) infusion of complexity theory, (8) simplicity, (9) linking macro- to micro-approaches, and (10) visualization (Torrens, 2000).

CAs have some disadvantages. Incorporating all internal and external rules in CAs models leads to complicated transition rules and reduces the generalization potential of CA model. In addition, incorporating large numbers of factors in CA model increases the calibration and simulation time (Wang, 2012). It is also unable to reveal the feedback of system and social economic influence on decision-making (Shanthi and Rajan, 2012).



### 1.6. Essential Features of Megacities and Applications of CA-based Models

The first step in selecting a suitable CA-based model to explore urban growth in megacities is having knowledge of the essential features of each megacity. Some CA-based models such as DUEM, SLEUTH, SimLand, and SIMLUCIA concentrate on processes and patterns of urban growth are not able to examine people's movement or economic processes within the city (Jankovic et al., 2005). Thus, CA-based models such as CAST and ISGM have been developed to bridge this gap. A summary of UGM applied to various megacities is listed in Table 1-2.

**Table 1-2.** List of previous CA-based models applied in megacities

<b>CA-Model</b>	<b>Megacity</b>	<b>Reference</b>
DUEM	Buffalo, NY	Xie, 1996
MOLAND	Lagos	Barredo et al., 2004
Type of CA-Based Model	Los Angeles  (Santa Monica Mountains National Recreation Area)	Syphard et al., 2005
DINAMICA	Dhaka	Ahmed and Bramley, 2015
DINAMICA	Sao Paulo	Almeida et al., 2005  Compos et al., 2015

<b>CA-Model</b>	<b>Megacity</b>	<b>Reference</b>
A CA-based Land-use Model	Tokyo	Arai and Akiyama, 2004
UES  UED	Beijing	He et al., 2006  He et al., 2008
An Urban CA model  Ant Colony Optimization-CA  Kernel-based non-linear CA  A Linguistic CA	Guangzhou	Li et al., 2008  Liu et al., 2008a,  Liu et al., 2008b  Wu, 1996
GIS-based CA Modeling  CA-Support Vector Machines	Shenzhen	Sui and Zeng, 2001  Yang et al., 2008
Population Surface Modeling  and CA	London  (Southeast England)	Wu and Martin, 2002

In China, economic growth caused the fast expansion of the urban population. This expansion, especially in Beijing and Shanghai, led to considerable increases in energy consumption and air pollution (Chan and Yao, 2008). To deal with the effects of urban expansion and dynamic processes with feedback in a system, there is a need to develop a CA-based model, such as UES,

which can represent the effects of macro-scale political, economic, and cultural driving forces. Another challenge in some megacities is the existing infrastructure. For instance, half of Mumbai commuters spend at least 1-2 hours traveling to work. Sao-Paolo also suffers from the lack of a well-developed transportation system. Among the mentioned CA-based models SLEUTH, METRONAMICA, CAST, and FCAUGM model consider the transportation system. Another issue is the effect of sea level rise on coast-hugging megacities such as Manila, Lagos, Buenos Aires, Rio de Janeiro, Los Angeles, New York, Istanbul, and so on. SIMLUCIA can assess the impacts of climate change on the territory of coastal countries and islands. The current CA-based models have been applied to megacities, reveals there is an acute need to design a next-generation UGM in response to future challenges and meeting the task of sustainable development.

### 1.7. Perspective and Challenges

Urban growth has significant impacts on society with political, ecological, and environmental implications constrained by infrastructural conditions. Accelerated urban growth combined with large number of inhabitants, especially in megacities, increases the risk of unsustainable development. This paper discusses the trends of population growth and urbanization in general to pin down the relevant issues of fast growth in megacities around the world. Given that urban planning and policy making can limit the impacts of accelerated urban growth, there is a need to employ a wealth of UGMs to simulate future urban growth and assess the effects of relevant policies. During the past 40 years, several UGMs such as CA-Based models, agent-based models, ANN-based models, fractal-based models, linear-logistic regression-based models, and decision tree models have been developed for many purposes. Among these models, CA-based models have received wide attention because of their ability to provide end users with a better understanding of

urban dynamics with respect to spatial and temporal dimensions of urban processes and patterns. Although several CA-based models have been developed and applied since the 1980s, these models are different based on their varying types of transition rules. In addition, these models can be integrated with different types of models and consider different urban development influence factors. In addition, the radius of affected area and neighborhood selection methods are different from one model to another in the sense that end users can be very selective when achieving different objectives from each CA-based model. Objectives of these models can be categorized as descriptive, perspective, and predictive. Since the 2000s, CA-based models have been integrated with other models driven by selective advantages of the integrated modeling process. Calibration is an important step in CA-based models as it helps to achieve a reliable simulation. Although scholars proposed several methods to calibrate CA-based models, only a few methods have been proposed to validate CA-based models. Prior to selecting a suitable CA-based model for a megacity, it is necessary to have a broad knowledge about the main challenges and perspectives related to the types of driving forces in that megacity. There is a need to design a next-generation UGM to deal with the intertwined challenges in fast growing mega-cities.

### 1.8. Limitations

Rapid urbanization has resulted in a growing number of megacities with varying levels of urban metabolism. Modeling urban growth requires knowledge of urban dynamics associated with different driving factors and ability to simulate the observed patterns to assess the effects of the urban planning policies. Recently, cellular automata (CA)-based urban growth models (UGMs) have become a popular tool for modeling such complex dynamics of urban systems. In the application of a CA-based UGM, the adopted scale and identified transition rules have significant

effects on the simulation precision. In order to analyze the urban growth pattern under the background of climate change, three megacities including New York City in the U.S., London in the U.K., and Beijing in China were selected as study area to predict land use land cover changes in 2030 and 2050. Different scenarios were considered and analyzed for comparison reason.

### 1.9. References

- Adamatzky A., 1994, Identification of cellular automata, In Encyclopedia of complexity and system science, Meyers R.A., pp. 4739-4751, ISBN: 978-0-387-75888-6
- Ahmed S., Bramley G., 2015, How will Dhaka grow spatially in future? Modeling its urban growth with a near-future planning scenario perspective, International Journal of Sustainable Built Environment, 4, pp. 359-377
- Almeida C.M.D., Monteiro A.M.V., Camara G., Soares-Filho B.S., Cerqueira G.C., Pennachin C.L., Batty M., 2005, GIS and remote sensing as tools for the simulation of urban land-use change, International Journal of Remote Sensing, 26(9), pp. 759-774
- Angelescu, P., Ionita, S., Safron, E., 2007, Block Encryption Using Hybrid Additive Cellular Automata. In: 7th International Conference on Hybrid Intelligent Systems. IEEE
- Arai T., Akiyama T., 2004, Empirical analysis for estimating land-use transition potential functions-case in the Tokyo metropolitan region, Computers, Environment and Urban System, 28(1-2), pp. 65-84.
- Bandini S., Mauri G., Serra R., 2001, Cellular automata: from a theoretical parallel computational model to its application to complex systems, Parallel Computing, vol. 27, pp. 539-553.

- Barredo J.I., Demicheli L., Lavalle C., Kasanko M., McCromick N., 2004, Modeling future urban scenarios in developing countries: an application case study in Lagos, Nigeria”, *Environ. Plann. Design.* 31, pp. 65-84.
- Batty, M. 2005, *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals.* MIT Press, Cambridge. The MIT Press, Cambridge, USA.
- Chan C.K., Yao, X., 2008, Air pollution in mega-cities in China, *Atmospheric Environment*, 42, 1-42.
- Chavoya A., Duthen Y., 2006, Using a genetic algorithm to evolve cellular automata for 2D/3D computational development, In: *Genetic and Evolutionary Computation Conf.*, 00. 231-232.
- Cheng J., Masser I., Ottens H., 2003, Understanding urban growth systems: theories and methods, proceeding of the 8<sup>th</sup> international conference on computer techniques for urban planning and management, CD-Rom Proceedings, Sendai, Japan.
- Cohen B., 2004, Urban growth in developing countries: a review of current trends and a caution regarding existing forecasts, *World Development*, 32(1), pp. 23-51, doi: 10.1016/j.worlddev.2003.040008
- Compos P.B.R., Queiroz Filho A.P., Pedro A.A., 2015, Cellular automata approach for analysis of land use cover in the vicinity of the Unified Center of Education of Paz, Sao Paulo, between 2002 and 2010, 27<sup>th</sup> International Cartographic Conference, Aug. 23-28, Rio. de Janeiro, Brazil.
- Crols, T., White, R., Uljee, I., Engelen, G., Canters, F., Peolmans, L., 2012, Development of an Activity-based Cellular Automata Land-use Model: the case of Flanders, Belgium. 2012 International Congress on Environmental Modelling and Software Managing Resources of

- a Limited Planet: Pathways and Visions under Uncertainty, Sixth Biennial Meeting, Leipzig, Germany. vol. F1, pp. 2000-2007
- Das, S., 2006, Theory and Applications of Nonlinear Cellular Automata In: VLSI Design. Ph.D Thesis, B.E. College.
- He C., N. Okada, Q. Zhang, P. Shi, and J. Li, 2008. Modelling dynamic urban expansion process incorporating a potential model with cellular automata. *Landscape and Urban Planning* 86(1): 79-91
- He C., Okada N., Zhang Q., shi P., Zhang J., 2006. Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China, *Applied Geography*, 26: 323-345
- International Federation of Surveyors (FIG), 2010, Rapid urbanization and mega-cities: the need for spatial information management, ISBN: 978-87-90907-78-5
- Jankovic L., Hopwood L., Alwan Z., 2005. CAST-city analysis simulation tool: an integrated model of land-use, population, transport, and economic. In proceedings of the 9<sup>th</sup> computers in urban planning and urban management conference, London.
- Jiao J., 2003, Transition rule elicitation for urban cellular automata models, case study: Wuhan, China, International Institute for Geo-information Science and Earth Observation, Enschede, the Netherlands, 121 p.
- Krishna P.V., Babu M.R., Ariwa E. (Eds.), 2012, Global trends in computing and communication systems, 4<sup>th</sup> international conference, ObCom 2011, Part I, CCIS 269, pp. 753-762, Springer-Verlag Berlin Heidelberg, ISSN: 1865-0929.

- Kundu, A., Pal, A.R., Sarkar, T., Banarjee, M., Guha, S.K., Mukhopadhyay, D., 2008, Comparative Study on Null Boundary and Periodic Boundary Neighbourhood Multiple Attractor Cellular Automata for Classification. IEEE 2008.
- Lavalle C., Demicheli L., Turchini M., Casals-Carrasco P., Neiderhuber M., 2001, Monitoring megacities: the MURBANDY/MOLAND approach, *Development in Practice*, 11(2 &3), pp. 350-357.
- Li C., 2014, Monitoring and analysis of urban growth process using remote sensing, GIS, and cellular automata modeling a case study of Xuzhou City, China, A doctorate dissertation, TU Dortmund University.
- Li X., Yang Q., Liu X., 2008, Discovering and evaluating urban signature for simulating compact development using cellular automata, *Landsc. Urban Plann.*, 86, 177-186.
- Liu X., Yang Q., Liu X., 2008a, Discovering and evaluating urban signature for simulating compact development using cellular automata, *Landsc. Urban Plann.*, 86, 177-186.
- Liu X., Li X., Shi X., Wu S., Liu T., 2008b, simulating complex urban development using Kernel-based non-linear cellular automata, *Ecol. Model.*, 211, pp. 169-181.
- Maeda K.I., Sakama C., 2007, Identifying cellular automata rules, *Journal of Cellular Automata*, 2(1-20), pp. 1-20.
- Maji, P., Pal Chaudhuri, P.: FMACA, 2004, A Fuzzy Cellular Automata Based Pattern Classifier. In: Lee, Y., Li, J., Whang, K.-Y., Lee, D. (eds.) DASFAA 2004. LNCS, vol. 2973, pp. 494–505. Springer, Heidelberg.
- Nandi, S., Kar, B.K., Pal Chaudhuri, P., 1994 Theory and Application of Cellular Automata in Cryptography. *IEEE Transaction on Computers* 43.



- Pan Y., Roth A., Yu Z., Doluschitz R., 2010, The impact of variation in scale on the behavior of a cellular automata used for land-use modeling, *Computers, Environment and Urban Systems*, 34, pp. 400-408.
- PBL, 2104, *Towards a World of Cities in 2050: An Outlook on Water-related Challenges Background Report* submitted to the UN-Habitat Global Report, Ligtoet, W. and Hilderink, H. (editors), PBL Netherlands Environment Assessment Agency.
- Placzek B., 2014, Neighborhood selection and rules identification for cellular automata: a rough sets approach, Chapter in: *Parallel Processing and Applied Mathematics*, 8385, Lecture notes in computer science, ISBN: 978-642-55194-9.
- Sante I., Garcia A.M., Miranda D., Crecente R., 2010, Cellular automata models for the simulation of real-world urban processes: a review and analysis, *Landscape and Urban Planning*, vol. 96, pp. 108-122.
- Seredynski, M., Bouvry, P., 2005, *Block Cipher Based on Reversible Cellular Automata*. New Generation Computer.
- Shanthi M., Rajan E.G., 2012, Agent based cellular automata: a novel approach for modeling spatiotemporal growth process, *International Journal of Application or Innovation in Engineering and Management*, vol.1, no. 3, pp. 56-61.
- Sui D.Z., Zeng H., 2001. Modeling the dynamics of landscape structure in Asia's emerging Desakota regions: a case study in Shenzhen, *Landsc. Urban Plann.*, 53, pp. 37-52.
- Sun X., Rosin P.L., Martin R.R., 2011. Fast rule identification and neighborhood selection for cellular automata, *IEEE trans. On Systems, Man, and Cybernetics, Part B: Cybernetics*, 41(3), pp. 749-760.

- Syphard A., Clarke K.C., Franklin J., 2005, Using a cellular automaton model to forecast the effects of urban growth on habitat pattern in Southern California, *Ecological Complexity*, 2(2), pp. 185-203
- Takeyama M., Couclelis H., 1997, Map dynamics: integrating cellular automata and GIS through Geo-Algebra, *Int. J., Geographical Information Science*, vol. 11, no. 1, pp. 73-91.
- Torrens P.M., 2000, How cellular models of urban systems work?, WP-28, Center for advanced spatial analysis (CASA), University College London. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=278E8F2236C17B255D6387FEA792FB52?doi=10.1.1.15.2030&rep=rep1&type=pdf>.
- United Nations, Department of Economic and Social Affairs, Population Division. 2014, World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352).
- Wang F., 2012, A cellular automata model to simulate land-use changes at fine spatial resolution, Department of Geomatics Engineering, University of Calgary, Available at: <http://www.geomatics.ucalgary.ca/graduatetheses>
- Wu F., Martin D., 2002, Urban expansion simulation of Southeast England using population surface modeling and cellular automata, *Environment and Planning A.*, 34, pp. 1855-1876.
- Wu F., 1996, A linguistic cellular automata simulation approach for sustainable land development in a fast-growing region, *Comput. Environ. Urban Syst.*, 20(6), pp. 367-387.
- Xie Y., 1996, A generalized model for cellular urban dynamics, *Geographical Analysis*, vol. 28, no. 4, pp. 350-373.
- Yang Q., Li X., Shi X., 2008, “Cellular automata for simulating land-use changes based on support vector machine”, *Computers and Geosciences*, vol. 34, no. 6, pp. 592-602.

Yang Y., Billings S.A., 2000, Extracting Boolean Rules from CA patterns, IEEE Transactions of Systems, Man, and Cybernetics- Part B: Cybernetics, 30(4), pp. 573- 581

Zhao Y., Billings S., 2007, The identification of cellular automata, Journal of Cellular Automata, 2(1), pp. 47-65.

## **CHAPTER 2 LINKING SOCIO-ECONOMIC DEVELOPMENT, SEA LEVEL RISE, AND CLIMATE CHANGE IMPACTS ON URBAN GROWTH IN NEW YORK CITY WITH A FUZZY CELLULAR AUTOMATA-BASED MARKOV CHAIN MODEL**

This chapter has been accepted for publication as follows:

Lu, Q., Joyce, J., Imen, S., Chang, N.B., 2017. Linking Socioeconomic Development, Sea Level Rise, and Climate Change Impacts on Urban Growth in New York City with a Fuzzy Cellular Automata-based Markov Chain Model. *Environment and Planning B: Urban Analytics and City Science*. Accepted for publication.

### **2.1. Introduction**

Coastal regions, particularly within megacities, can encounter additional challenges since these regions are easily impacted by accelerated sea level rise (SLR), long-term rainfall variability, and other climate change influences such as extreme weather events (Nicholls, 1995; Roland, 2010). SLR can have a profound impact on the land use change and land cover types because of the risk of coastal flooding (Murali and Kumar, 2015). SLR was a global concern throughout the 20<sup>th</sup> century and could accelerate through the 21<sup>st</sup> century due to global warming (Nicholls and Cazenave, 2010).

Dynamism and growth are two common characteristics of most urban areas (Barredo et al., 2004). Modeling these two in each megacity depends on the stochasticity of an urban system and its major urban growth factors. Urban growth simulations to characterize land use changes can help understand past growth patterns, as well as offer insights into future land use development under certain assumptions (Moghadam and Helbich, 2013). Advances in technology and

socioeconomic developments can provide information about the trends of historical growth patterns, which can be used in an urban growth model (UGM) to provide understanding of probable future developments. Different UGMs, including the cellular automata (CA)-based model (Batty, 2007), the agent-based model (ABM) (Railsback and Grimm, 2011), the spatial statistics model (Xie et al., 2005) (e.g., Markov chain (MC) analysis, multiple regression analysis, principal components analysis, linear regression, and logistic regression), artificial neural networks (ANNs)-based model (Guan et al., 2005), and the fractal-based model (Cheng et al., 2003) have been developed and applied since the 1980s. As with most mathematical models, however, there exist gaps or drawbacks. For instance, ANNs have been used widely for urban process simulations, but they do not provide direct clues regarding the contribution of each of the driving forces (Guan et al., 2005). As for fractal-based modeling, the temporal dimension is not included and the spatial results are difficult to interpret since the same value of fractal dimension may represent different forms and structures (Cheng et al., 2003). The ABM is based on our understanding of how individual behavior leads to the global outcome, and random definitions of initial conditions and interaction rules can cause variable results (Couclelis, 2002).

Among different types of UGMs, the ability of a cellular automata (CA)-based model to deal with the extremely complex content of urban dynamic elements makes it a popular tool among urban modelers (Batty, 2007; Arsanjani et al., 2013). CA models are based on a bottom-up approach in which complex systems are driven by the interactions of local conditions and surrounding subsystems (Liu et al., 2008). CA models consider the influence of neighborhood cells and the interactive relationship within each cell while the predefined transition rules enable CA models to determine the change of each cell (Batty et al., 1999). Yet the process of the CA-based

model is able to resolve spatial phenomena and considers multiple timescales with respect to a cascade impact (Torrens, 2000a, b). Earlier studies showed the success of CA-based models in simulating land use and land cover (LULC) changes worldwide (Batty, 1995; Batty, 2005), and currently CA-based models are integrated with different techniques and applications such as logistic regression, support vector machines, statistical models, and so on (Shafizadeh-Moghadam et al., 2017b; Tayyebi et al., 2016). CA models, however, are unable to account for the actual amount of change over varying temporal scales (Moghadam and Helbich, 2013). MC models can bridge this gap by retrieving the quantity of change over varying temporal scales under uncertainty of the urban growth process (Guan et al., 2011). These MC models may describe a stochastic process that the probability of one state is to change to another over differing time scales (Arsanjani et al., 2013). One limitation of these MC models is the inability to account for impacts of neighborhood cells over the spatial dimension (Arsanjani et al., 2013). Hence, the integration of the CA and MC models is able to incorporate both spatial and temporal dimensions of urban processes (Sante et al., 2010). In these integrated CA and MC models, factors and constraints within a CA-MC model are normally considered for decision analysis of possible land use change pathways under uncertainty via an integrated fuzzy membership function and multi-criteria evaluation. In this regard, the weighted linear combination (MCE-WLC) approach to entail the behavior-oriented transition rules in such a CA-MC model became popular (Wu, 1998). Therefore, the integration of CA and MC models may further provide an opportunity to simulate spatial and temporal land-use variations and predict future land use patterns (Memarian et al., 2012; Agarwal, 2002).

A potential impact on land use change that is typically not considered is the effect of SLR and long-term climate change. While SLR and climate change impacts on land use and land cover have typically focused on agronomics (Dinar & Mendelsohn, 2011), just as important is the impact of SLR and climate change, such as future rainfall modeling, on land changes in urban areas. Although urban development increases flood risk in cities due to increasing urban runoff (Huong and Pathirana, 2013), the effect of climate change combined with potential SLR on coastal flooding and urban development has not yet been well studied.

The aim of this chapter is thus to demonstrate how an integrated modeling system consisting of the CA, MC, fuzzy membership, and MCE-WLC, can be designed, calibrated, validated, and applied to predict the future land use change of New York City (NYC) for 2030 and 2050 under traditional and emerging impacts (i.e., rainfall runoff variations and SLR) simultaneously given the assumptions of stationarity in the urban growth process. The following scientific questions are explored in this study: (1) How will rainfall vary in 2030 and 2050 with the impact of climate change in NYC? (2) How can the land use change in 2030 and 2050 under known land use policies as well as uncertain SLR and long-term rainfall variability impacts be modeled for decision analysis of future land use patterns? and (3) How significantly will SLR and climate change affect land use changes in NYC in 2030 and 2050?

## 2.2. Study Area

NYC is located on the east coast of the United States, geographically adjacent to the Atlantic Ocean. The city has five boroughs: Manhattan, Bronx, Brooklyn, Queens, and Staten Island. According to the United States Census Bureau, the total area is 1,214 square kilometers in which

the land area is 789 square kilometers and the water area is 425 square kilometers. The total population of NYC was 8,491,079 in 2014 while the population density was 10,756 per square kilometer (United States Census Bureau, 2015). NYC ranks first in population density in the United States. Due to NYC being highly urbanized, most of the stormwater in NYC flows over impervious surfaces and subsequently the storm-sewer networks. These impervious surfaces cover approximately 72% of 305 square miles (790 square kilometers) of NYC's land area and generate a significant amount of stormwater (City of New York, 2016). In addition, according to the New York State government, there was at least 30.45 cm (1 foot) of SLR in New York State since 1900, primarily attributed to the warming ocean water. The New York State government has predicted a SLR ranging from 0.381 meters (15 inches) to 1.905 meters (75 inches) by 2100 (New York State Department of Environmental Conservation, 2015).

### 2.3. Software and Data Collection

#### 2.3.1. Statistical Downscaling Model (SDSM)

In order to determine the potential impact of future rainfall on flooding in NYC, a SDSM was utilized to project future rainfall for 2030 and 2050 (Wilby and Dawson, 2007). A SDSM is a statistical downscaling approach, which is based on multiple linear regression techniques, between large-scale climate and local climate such as daily rain gauge records. A SDSM also allows for incorporating future climate predictors from global climate models (GCMs). For this study, daily observed rainfall was sought as an input to the SDSM based on the outputs from a GCM (Wilby and Dawson, 2007). Daily rainfall data were obtained from the NOAA Climate Data Online database. Stations with the longest continuous period (30+yrs) of daily rainfall record were chosen and are summarized in Table 2-1.



**Table 2-1** Rainfall Stations with Continuous Period (30+yrs) of Record in NYC

Station Name	Station ID	Period of Record	Lat/Long (Decimal Deg)
JFK Airport	COOP:305803	1948-Present	40.63 N / 73.76 W
NY City Central Park	COOP 305801	1876-Present	40.78 N / 73.96 W
LaGuardia AP	COOP:305811	1944-Present	40.77 / 73.88 W

### 2.3.2. CA-based MC Model and Data Collection

In this study, the integration of a GIS and TerrSet software was applied to realize the urban growth process and possible future land use patterns. LULC data were collected for 2001, 2006, and 2011 for characterizing the pathways of multi-temporal land use change from a historical perspective and predicting future land use change in 2030 and 2050. These LULC data were available from the Multi-Resolution Land Characteristics Consortium (MRLC), which provides national LULC data. These national LULC data are produced using a decision-tree classification of Landsat 5 Thematic mapper imagery acquired from the United States Geological Survey Earth Resources Observation and Science Center Landsat archive, where they have been radiometrically and geometrically calibrated (MRLC, 2015).

The LULC data include fifteen categories: open water, developed open space, developed low intensity, developed medium intensity, developed high intensity, barren land (rock/sand/clay), deciduous forest, evergreen forest, mixed forest, shrub/scrub, grassland/herbaceous, pasture/hay,

cultivated crops, woody wetlands, and emergent herbaceous wetlands. According to the USGS Land Cover Institute (2016), the aforementioned fifteen land-use categories can be consolidated into eight categories including water, open space, urban, barren, forest, rangeland, agriculture, and wetland to meet the goal of this study. The percentages of each land use category in 2001, 2006, and 2011 are summarized in Table 2.2.

**Table 2-2.** LULC data summary in percentage (%)

Land use classes	2001	2006	2011	2001-2006	2006-2011	2001-2011
Water	1.02%	1.02%	1.01%	0.00%	-0.01%	-0.01%
Open space	6.68%	6.27%	6.16%	-0.41%	-0.11%	-0.52%
Urban	81.09%	81.58%	82.11%	0.49%	0.53%	1.02%
Barren	0.95%	0.96%	0.98%	0.01%	0.02%	0.03%
Forest	4.13%	4.08%	3.87%	-0.05%	-0.20%	-0.26%
Rangeland	1.82%	1.82%	1.75%	0.00%	-0.07%	-0.07%
Agriculture	0.13%	0.13%	0.11%	0.00%	-0.02%	-0.02%
Wetlands	4.18%	4.14%	4.00%	-0.04%	-0.14%	-0.18%
Total	100.00%	100.00%	100.00%			

Other data required for this study, which include road data, protected areas, a digital elevation model, SLR, and flood hazard data, were obtained from the sources shown in Table 2.3. These layers were re-projected and resampled to 30 meters by 30 meters. Therefore, all the processed layers have the same numbers of rows and columns and the same pixel size.

**Table 2-3.** Data sources and description

Dataset	Source	Date	Resolution
LULC	MRLC	2001, 2006, 2011	30m×30m
Roads	New York City Open Data	2016	N/A
The protected area	NYPAD	2013	N/A
DEM	New York City Open Data	2010	1 arc-second
SLR	National Oceanic Atmospheric Administration (NOAA)	2011	N/A
Flood hazard	FEMA	2016	N/A

## 2.4. Methodology

### 2.4.1. Statistical Downscaling Model (SDSM)

#### 2.4.1.1. *Climate Predictors*

Climate predictors are important factors in a SDSM since a particular set or combination of climate predictors can impact the results when the predicted values are compared statistically to our field observations. The selection of the best predictors is an iterative process (Wilby et al, 2002; Hessami et al., 2008; Teutschbein et al., 2011) which removes the least significant predictors until all remaining predictors are statistically significant. Such an iterative process leads to the generation of a clear relationship between predictor variables and predictands such as rainfall. The SDSM relies on both observed-derived and GCM-derived atmospheric predictor variables. Observed predictor variables such as those from the National Center for Environmental Prediction (NCEP) (Kalnay et al., 1996) are used for SDSM calibration and validation while GCM-derived predictor variables are utilized within the “Scenario Generator” component of the SDSM for future projections (Wilby and Dawson, 2007).

The GCM-derived atmospheric predictor variables were obtained from the Hadley Centre Coupled Model, version 3(HADCM3) GCM A2 scenario of the Intergovernmental Panel on Climate Change Fourth Assessment Report (Solomon, 2007), representing the worst-case scenario. All atmospheric predictor variables were re-gridded to a standard coordinate system ( $2.5^{\circ}$  latitude  $\times$   $3.75^{\circ}$  longitude) used in HADCM3 over the period 1961 to 2099 and normalized with respect to the base period (1961 to 1990) averages. The predictor variables of concern include mean sea level pressure, surface airflow strength, surface vorticity, 500hPa airflow strength,

500hPa meridional velocity, 500 hPa surface vorticity, 850 hPa airflow strength, 500hPa meridional velocity, 500 hPa surface vorticity, 850 hPa geopotential height and surface specific humidity.

#### *2.4.1.2. Calibration, Validation, and Future Projection*

Given 41 years of NCEP daily observed-predictor variables (1961-2001), the first 20 years (1961–1981) were used for calibrating the SDSM model, while 1990-2000 predictor variables were used to validate the SDSM model. Before calibration and validation, SDSM requires daily historical data to be close to a normal distribution. Daily rainfall is typically positively-skewed statistically; thus a transformation of daily rainfall data was required to obtain a near-normal distribution using a log transformation of the observed daily rainfall data.

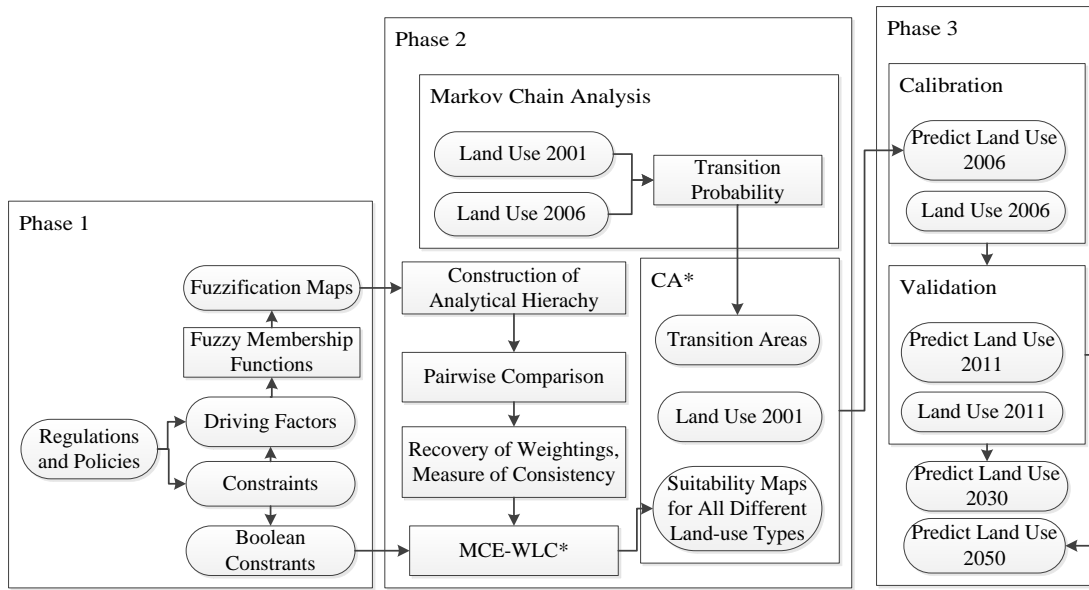
Future projections of rainfall were conducted using the scenario generator in the SDSM, which produces ensembles of future daily weather series for future climate using the calibrated SDSM model and GCM-derived atmospheric predictor variables obtained from the HADCM3 GCM A2 scenario. Annual mean rainfall projections from a 25-member SDSM ensemble were considered to provide a range of possible projections of annual mean rainfall values for 2030 and 2050; this range was compared to each year of the observed historical record (1961-2015) to retrieve and assess the historical vs. future trend. These results will become part of the factors and constraints in our UGM modeling analysis.

#### *2.4.2. CA-MC Model - Land Use Change Prediction*

An integrated hybrid UGM consisting of the MC, CA, fuzzy membership, and MCE-WLC was implemented to predict NYC's projected land use changes for 2030 and 2050. The

methodological framework is shown in Figure 2-1. This framework is designed to represent the multi-level urban growth dynamics with the aid of stochastic and linguistic terms accounting for both spatial, temporal, and social uncertainties. The suggested framework incorporates a set of fuzzy sets (i.e., linguistic terms) and MC (i.e., stochastic terms), each of which focuses on different aspects of the urban growth dynamics, while the systems' structure and function provide a wealth of connection points and a workflow that is closer to the conceptualization of the phenomenon of urban dynamics.

In detail, the proposed framework (Figure 2-1) includes three phases to conduct the integrated hybrid UGM analysis. In a broader sense, city regulations can limit the development of lands in the future, while financial and wildlife considerations are also important to developers and environmentalists (Eastman, 2015). Hence, the main objective of Phase I is to determine the driving factors and constraints based on the NYC land use policies and regulations discussed in the previous section. Next, in Phase II, an integrated hybrid model consisting of CA, MC, fuzzy membership, and MCE-WLC is applied to generate suitability maps for each land use class and to predict the land use in 2006 and 2011 for calibration and validation, respectively, in the next phase. Lastly, Phase III is designed for calibration and validation of the developed UGM for NYC and for prediction of LULC in 2030 and 2050.



**Figure 2-1.** Study framework (\*CA: cellular automata; MCE-WLC: multi-criteria evaluation-weighted linear combination)

#### 2.4.2.1. Phase I: Criteria Selection and Preprocessing

##### 2.4.2.1.1. Factors and Constraints

Based on the land use policies and regulations, criteria can be defined as factors or constraints in such a decision analysis. Constraints limit our decision analysis to a specific geographic location and are always Boolean images with values of 0 and 1. In contrast, factors indicate the degree of suitability of pixels located in different geographic regions. Therefore, a factor is continuous in character and is applied to measure the suitability ranging from value 0 to 1. In this study, all factors are considered in terms of fuzzy decision rules in which degrees of unsuitability and

suitability of pixels are defined as continuous values between 0 and 1 via the use of fuzzy membership functions.

#### 2.4.2.1.2. Land use policies related to factors and constraints

Policies and regulations related to land use dominate the fuzzy decision rules to capture the land use variation behavior in the model (Wu, 1998). Human-activity related factors are more essential to land use variations in the city, reflected by corresponding land use policies and regulations. Therefore, collecting and analyzing land use policies and converting to corresponding constraints and factors are the first steps in the modeling analysis.

NYC has comprehensive policies and laws regulating the utilization of land. According to the New York Protected Areas Database (NYPAD, 2013), protected lands including conservation lands, open space, and natural areas are prevented from urban development. Freshwater wetlands with an area larger than 0.057 square kilometers (12.4 acres) are also under protection from being developed based on the NYC Wetlands Strategy (NYC Department of Park and Recreation, 2012). In this case, freshwater wetlands do have the probability to be converted to other land use types if the area is smaller than 0.057 square kilometers (12.4 acres). In addition, the NYC Wetlands Strategy states that the buffer area of 46 meters (150 feet) adjacent to wetlands cannot be developed, but urban development is allowed greater than 91 meters (300 feet) from wetlands (NYC Department of Park and Recreation, 2012).

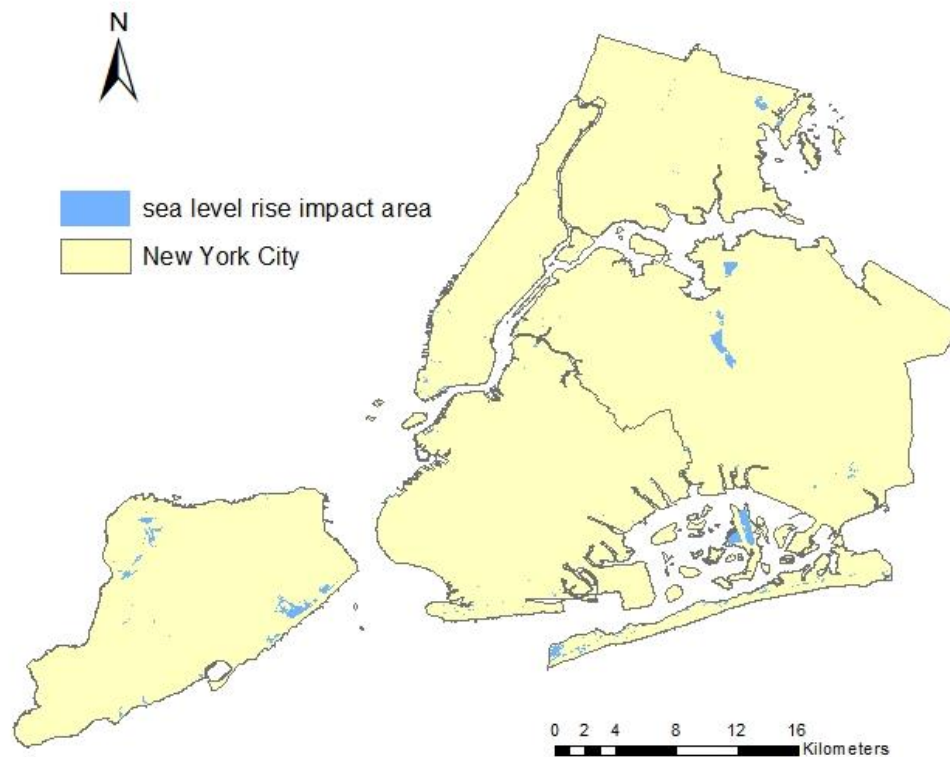
Socioeconomic factors are reflected as economic development, driven by urban expansion, transportation, and infrastructure (Xiao et al., 2006). Road networks and transportation are commonly considered factors of urban sprawl as they connect urban and rural areas and provide



linear branch development (Bhatta, 2010). The closer an area is to roads, the higher the suitability to promote urban and economic growth. Additionally, population growth is commonly considered a driving factor of environmental and land use changes (e.g., deforestation) (Meyer and Turner, 1992). According to the NYC Department of Planning (2016), NYC has the highest population density in the USA. Residential areas with different intensities are present as developed areas based on the definitions of the land use classes from the USGS Land Cover Institute (2016). Moreover, slope is another factor for urbanization, especially for spontaneous neighborhood growth (Silva and Clarke, 2002). In addition, all the LULC classes should be driving factors in order to achieve spontaneous growth. The further away from the original LULC class, the higher the probability that it will be changed to other LULC classes.

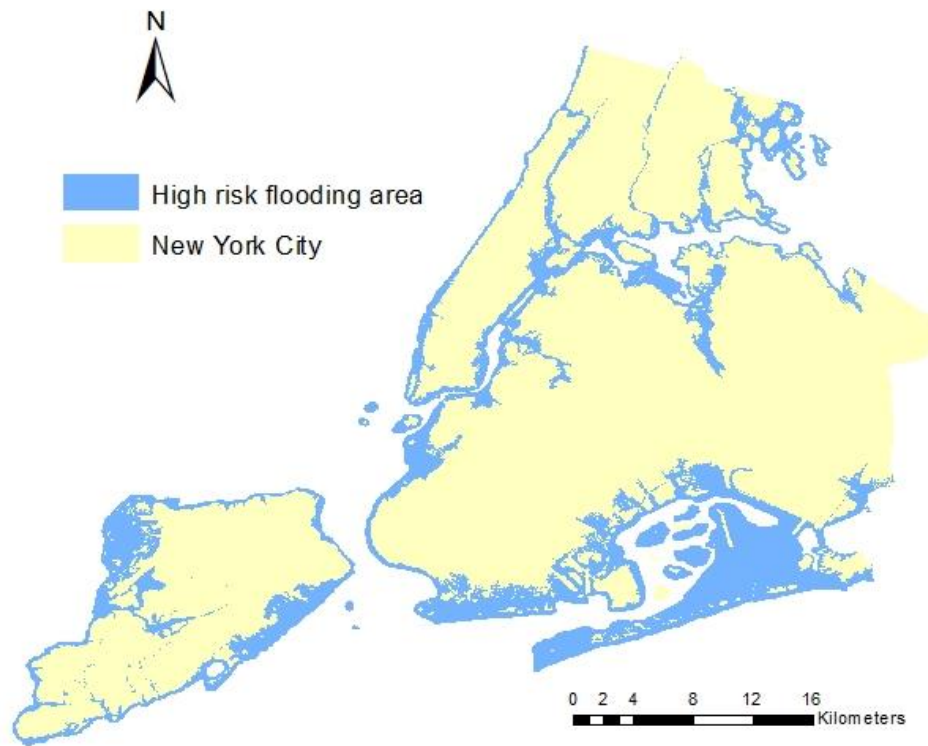
#### 2.4.2.1.3. SLR and the impact of flooding

In this study, the SLR data are acquired from NOAA which includes predicted SLR in future scenarios. The SLR impacted area is confirmed by comparing the most frequently flooded area predicted in 2050, which is extracted as an independent layer and input as a factor in the model (Figure 2-2). Compared to other policy-oriented factors in terms of land use changes, the factor of SLR is more related to climate change and has more impacts on urban, agriculture and rangeland where the agglomerations of human activities are concentrated. Other land use classes including water, barren, forest, and wetlands involve relatively less human activity. Thus, the factor of SLR is assigned a lower weight compared to other factors in the model. However, the impact of SLR may increase if climate change continues, resulting in accelerated SLR. The model can be trained and applied based on real-world changes to policies and variations in relevant environmental issues.



**Figure 2-2.** Predicted SLR impact area in 2050 in NYC (NOAA, 2011)

Data on the impact of flooding were acquired from FEMA, which were available up to 2016, including the impact of Hurricane Sandy. The flood hazard data reflect flood hazard and risk from low to high. The high risk flooding area was extracted as a spatial data layer, representing a factor associated with the impact of flooding in the model (Figure 2-3). The constraints and factors used with their source or related policies are summarized in Table 2-4.



**Figure 2-3.** High risk flooding area in NYC (FEMA 2016)

**Table 2-4.** Constraints and factors and their related sources or policies

	Name	Sources or Policies
Constraints	Protected area	NYPAD, 2013
	Wetlands larger than 0.05 km <sup>2</sup> (12.4 acres)	NYC Department of Park and Recreation, 2012
	Actual roads	
Factors	Wetlands smaller than 0.05 Km <sup>2</sup> (12.4 acres) with buffer larger than 45.72 m (150 ft)	NYC Department of Park and Recreation, 2012
	Slope less than 15%	
	Distance to roads	
	Distance to high flood risk area	FEMA, 2016
	Distance to SLR	New York State Department of Environmental Conservation, 2015
	Distance to each LULC class	

#### 2.4.2.1.4. Fuzzification maps for factors and Boolean maps for constraints

After selecting the factors and constraints for this study, a GIS-based *fuzzy* logic approach was applied to process the data for all factors. Thematic fuzzy systems provide useful intermediate results in regard to urban growth potentials, making it easier to conduct further decision analysis based on simple comprehensible rules. These corresponding fuzzy variables consist of three ordered fuzzy symbols (low, average, high) and delineate the suitability of each area to urbanization based on the input factors (Zadeh, 1968). The use of fuzzy CA to reflect spatial uncertainty leads to some flexibility for environmental planning and management (Liu and Phinn, 2003). For each driving factor, the “Distance module” in TerrSet is used to calculate the distance from a feature (e.g., distance to the road network, distance to water bodies, distance to the slope, and distance to SLR impact zone). Then the maps produced in the previous step with distance are standardized using the fuzzy module. The types of fuzzy membership functions need to be selected among three available in TerrSet including linear, sigmoidal and J-shaped. Control point values also need to be defined in the fuzzy module, which are the minimum and maximum distances calculated from distance images.

In contrast to factors, constraints are presented as Boolean images with only 0 and 1 values. “1” indicates a pixel will be developed in the future, while 0 indicates that the pixel will remain in its current condition. If a given area has value of 0 for a constraint, then the area can never be suitable for development, regardless of the factors.

#### 2.4.2.2. *Phase II: Integrated Hybrid Model*

##### 2.4.2.2.1. Analytical Hierarchy Process (AHP), MCE-WLC, and Suitability Maps

The modeling of land use suitability is an analytical process that describes the fitness of a certain unit of land for a specified use based on multiple criteria (Saaty, 1990). However, the process of determining the weight associated with each criterion is a decision analysis issue. The AHP applies pairwise comparison to determine the weights of fuzzy potential maps (Saaty, 1990; Carr and Zwick, 2007). Expert knowledge and qualitative interviews can help assign weights via the AHP. Factors are weighted based on their impacts on the land use changes and the interaction relationship between each factor. After each round of pairwise comparison, the consistency ratio (CR) is measured to test the reliability of judgment (Saaty, 1991). In the case that the CR is not acceptable (i.e., greater than 0.1), the assigned weights need to be reproduced through an additional AHP analysis. It should be noted that considering different driving factors or assigning different weights to fuzzy potential maps result in different suitability maps.

In this study, suitability maps for all different land use classes can be generated based on the MCE-WLC approach after inputting the Boolean maps for constraints and fuzzy maps for factors with their corresponding weights. Suitability maps are the final results of the MCE-WLC decision analysis. Suitability maps with continuous values are used to determine the best place or site which will change to a specific land use class. Suitability maps are normally scaled by a color scheme. A value 0 indicates no suitability; and a value of 255 is the highest suitability.

#### 2.4.2.2.2. MC

An MC model can be applied to analyze the actual quantity of variations between land use categories temporally and non-spatially (Muller and Middleton, 1994). It describes the likelihood of a random process triggering stochastic variables (e.g., forest) and changing to another state (e.g., urban) within a certain time period, involving transition probabilities (López et al., 2001). In general, the MC process is performed to estimate the transition matrix between the two past and documented dates (date 1 and date 2) and to estimate probabilities of change for the third date (date 3) to be predicted. Hence, a transition probability matrix is produced to express the possibility that a cell of a given land use category will change into any other category, and a transition area matrix contains the total area (in cells) expected to change in the next target year.

In this study, the matrix of transition probabilities, which is a square *matrix* used to describe the *transitions* of a MC, is built based on the assumption of stationarity and observation of temporal changes between 2001 and 2006. These transition probabilities lead to the formalization of a transition probability matrix for future prediction of land use changes in 2011, 2030 and 2050. Mathematically, the element in the transition probability matrix can be denoted as  $P_{ij}$  in which  $i$  and  $j$  represent two different land use categories. Thus, the transition probability can be described as follows: in time  $t_0$ , the land use type is  $i$  that will change to land use type  $j$  in time  $t_1$ . The prediction of the transition probability for each element (cell) is based on a specific predetermined time interval, which can be demonstrated in the matrix  $P$  shown below:

$$(V_i \times P_{ij}) = (V_1, V_2, V_3 \dots V_n) \times \begin{pmatrix} P_{11}, & P_{12}, & P_{13}, & \dots & P_{1m} \\ P_{22}, & P_{22}, & P_{23}, & \dots & P_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{n1}, & P_{n2}, & P_{n3}, & \dots & P_{nm} \end{pmatrix} \quad (2-1)$$

where  $V_i \times P_{ij}$  = Proportion of land cover of the second date;  $P_{ij}$  = Matrix of probability of land cover transition;  $V_i$  = Proportion of land cover of the first date (Vector);  $i$  = type of land cover in the first date;  $j$  = type of land cover in the second date. For example,  $P_{11}$  = the probability that a land cover 1 will change into land cover 1 from the first date to the second date;  $P_{12}$  = the probability that a land cover 1 will change into land cover 2 from the first date to the second date and so on;  $m$  = the number of land cover types in the study area (i.e., the system has  $m$  states).

As the projection moves forward on a yearly basis, the corresponding MC model is carried out from 2006 for 5 years for validation, 24 years for prediction, and 44 years for prediction in 2011, 2030, and 2050, respectively. The output of this modeling process includes a transition probability matrix and a transition area matrix for the projection of land use in each of the three target years. The projected transition area matrix for 2011, 2030, and 2050 is thus produced by multiplication of each column in the corresponding transition probability matrix by the number of cells of the corresponding land use in the later image in the year 2006 (Eastman, 2015).

#### 2.4.2.2.3. CA

CA-based UGMs are spatially dynamic models that can be applied to study LULC changes (White and Engelen, 1993). In contrast to MC models, CA models have the capability to process spatial interaction relationships to simulate dynamic land use changes in complex urban systems (Neumann and Burks, 1966). A CA-based MC model estimates the spatial distribution of LULC at a later date (date 3) which is 2011, 2030, and 2050 in this study. Using the output from equation (1) produced by the MC analysis, the predicting model needs to apply a contiguity filter to grow



the LULC cell by cell from date 2, which is 2006 in this study, to a later time (date 3). This filter in the CA-based MC model develops a spatially explicit contiguity-weighting factor to change the state of a cell based on its neighboring cells. Thus, a CA-based MC model exhibits the unique integrative ability to analyze spatiotemporal land use changes based on the transition probability of each neighboring cell over the study area (Subedi et al., 2013). Both the spatial contiguity and the probable spatial transitions that indicate the specific state of each land pixel (i.e., cell) at a specified time are handled with inherent linkages in the proposed fuzzy CA-based MC model with uncertainty, which is a strength of this study.

To train the fuzzy CA-based MC model, the required input files include a base map (LULC in 2001 in this case), a Markov transition area matrix (output of the MC process), and a group of suitability maps (one for each land use class). In our case, a 5×5 contiguity spatially explicit filter is applied to decrease the suitability of the cells far from the existing cells (Houet and Hubert-Moy, 2006). By multiplying each suitability map with the Boolean mask filter for each constraint and considering the fuzziness of each factor, preferences can be assigned to the contiguous suitable pixels. This produces the predicted land use maps for 2011, 2030 and 2050 based on the MCE-WLC approach.

#### *2.4.2.3. Phase III: Model Calibration, Validation, and Prediction*

The observed and predicted land use maps in 2006 were compared to test the accuracy of the model calibration of the fuzzy CA-based MC models. Similarly, the observed land use map for 2011 was used for validation purposes. The Kappa statistic was used to measure the accuracy of the model results between the predicted and observed land use maps (Bishop et al., 1977). If the

difference is within an acceptable range, then the model is valid and ready for predicting land use maps for 2030 and 2050.

## 2.5. Results and discussion

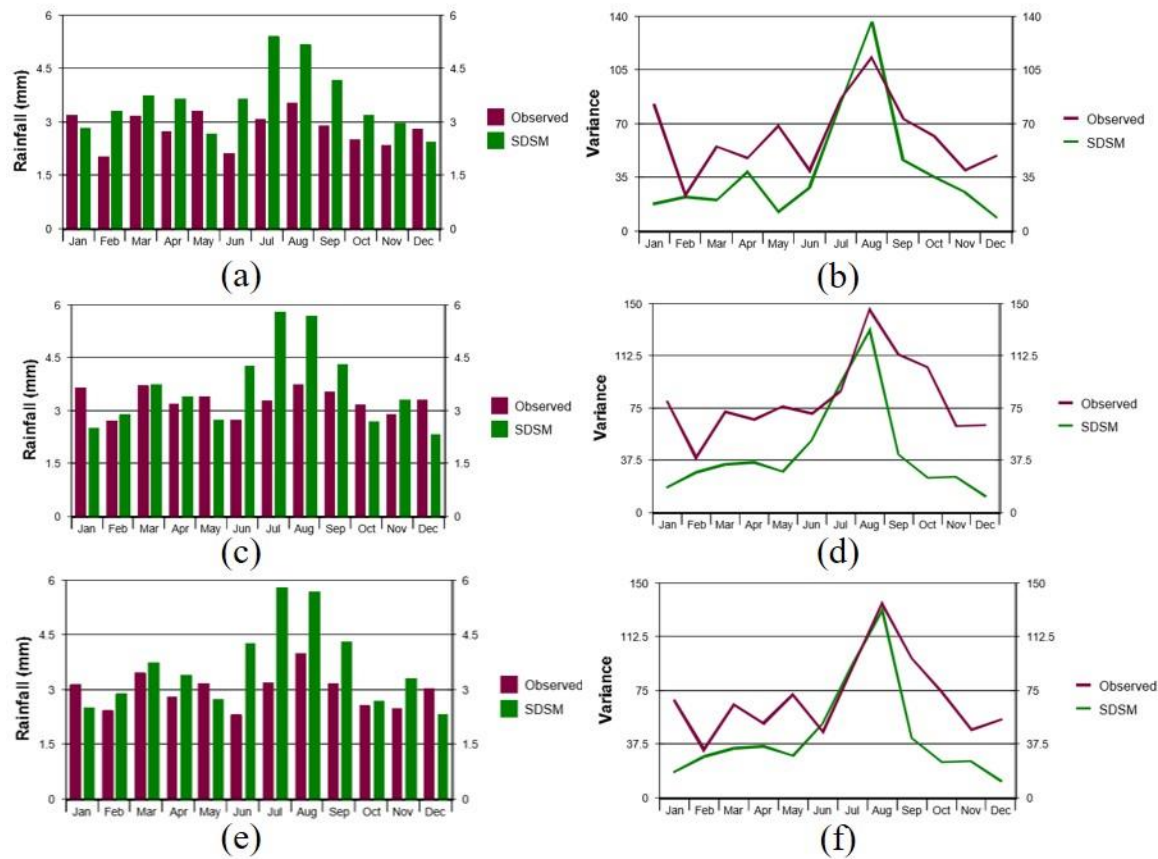
### 2.5.1. Statistical Downscaling Model (SDSM)

#### 2.5.1.1. *Calibration & Validation*

Results of the SDSM calibration for all rainfall stations (JFK Airport, Central Park and LaGuardia Airport) are provided in Table 2-5 based upon a monthly analysis for the period 1961-1981. As reflected in the literature, such as Wilby et al. (2002) and Gulacha and Mulungu (2016),  $R^2$  values for calibration of rainfall were low (Table 2-5) due to the difficulty in downscaling a highly variable predictand such as daily rainfall. Rainfall is modeled as a conditional process in a SDSM since there is no direct link between climate predictors and rainfall. An intermediate variable, such as the probability of wet-day occurrence, is provided in the conditional process. For the validation period, from January 1990 to January 2000, the goal for SDSM was to capture both mean monthly rainfall and monthly variance as shown in Figure 2-4, respectively. There is a general agreement in the time series patterns between the observed and predicted values.

**Table 2-5.** Calibration Results for Daily Rainfall (1961-1981) with observed predictors

	JFK Airport Location	Central Park Location	LaGuardia Airport Location
Month	R <sup>2</sup>	R <sup>2</sup>	R <sup>2</sup>
January	0.334	0.309	0.355
February	0.299	0.288	0.423
March	0.226	0.241	0.238
April	0.329	0.337	0.425
May	0.275	0.284	0.287
June	0.081	0.208	0.198
July	0.153	0.238	0.196
August	0.237	0.216	0.236
September	0.317	0.190	0.191
October	0.382	0.330	0.358
November	0.382	0.374	0.321
December	0.464	0.430	0.470



**Figure 2-4.** Monthly Mean Rainfall (Observed vs. SDSM) for each station from Jan 1990-Jan 2000: (a) JFK Airport, (c) Central Park and (e) LaGuardia Airport and monthly variance (observed vs. SDSM) for (b) JFK Airport, (d) Central Park and (f) LaGuardia Airport

#### 2.5.1.2. *SDSM Future Projections (2030 & 2050)*

In determining the future impacts of rainfall over NYC, projected annual rainfall means for 2030 and 2050 were compared with historical annual rainfall means (1961-2015) for all stations considered. Overall, the projected annual mean rainfall will increase during the years 2030 and

2050 across all stations according to the projections from the HADCM3 A2 scenario. The 2030 and 2050 average annual means of rainfall values at the JFK airport station for the 25-member ensemble are 4.3 mm and 7.4 mm, respectively. The same values for Central Park and LaGuardia airport are 4.9 mm and 8.2 mm, and 4.6 mm and 7.1 mm, respectively. All three results support the idea that future rainfall impacts needs to be considered as an important factor for land use changes in NYC, in addition to SLR.

## 2.5.2. Results of the Fuzzy CA – based MC Model

### 2.5.2.1. *Markov Probability Matrixes*

The Markov transition probability matrix and transition area matrix are created by using the Markov modeler in TerrSet, representing the probabilities of each land use class changing to another land use class from 2001 to 2006 (Table 2-6). From the probability matrix, it is clear that open space has a very high probability of changing to urban area, which is 0.0650 (6.5%) compared to other land use classes. Other land use classes including forest, rangeland, agriculture, and wetlands, have probabilities of 0.0067 (6.7%), 0.0015(1.5%), 0.0027 (2.7%), and 0.0069 (6.9%), respectively, of changing to urban area. However, the probabilities of water and barren changing to urban areas are negligible (0.0008 (0.8%) and 0.0002 (0.2%)) based on the MC analysis.

**Table 2-6.** Transition probability matrix in MC calculated using land use maps 2001 and 2006

	open					rangelan	agricultu	wetland
	water	space	urban	barren	forest	d	re	s
	0.994		0.000	0.002	0.000			
water	8	0.0003	8	0	0	0.0020	0.0000	0.0000
	0.000		0.065	0.000	0.000			
open space	0	0.9350	0	0	0	0.0000	0.0000	0.0000
	0.000		1.000	0.000	0.000			
urban	0	0.0000	0	0	0	0.0000	0.0000	0.0000
	0.000		0.000	0.998	0.000			
barren	1	0.0002	2	8	0	0.0004	0.0000	0.0002
	0.000		0.006	0.002	0.985			
forest	0	0.0046	7	6	2	0.0008	0.0000	0.0002
	0.000		0.001	0.000	0.001			
rangeland	0	0.0008	5	3	7	0.9955	0.0000	0.0003
	0.000		0.002	0.000	0.000			
agriculture	0	0.0036	7	0	0	0.0000	0.9937	0.0000
	0.000		0.006	0.000	0.000			
wetlands	1	0.0019	9	3	0	0.0000	0.0000	0.9908

#### 2.5.2.2. *Suitability Maps*

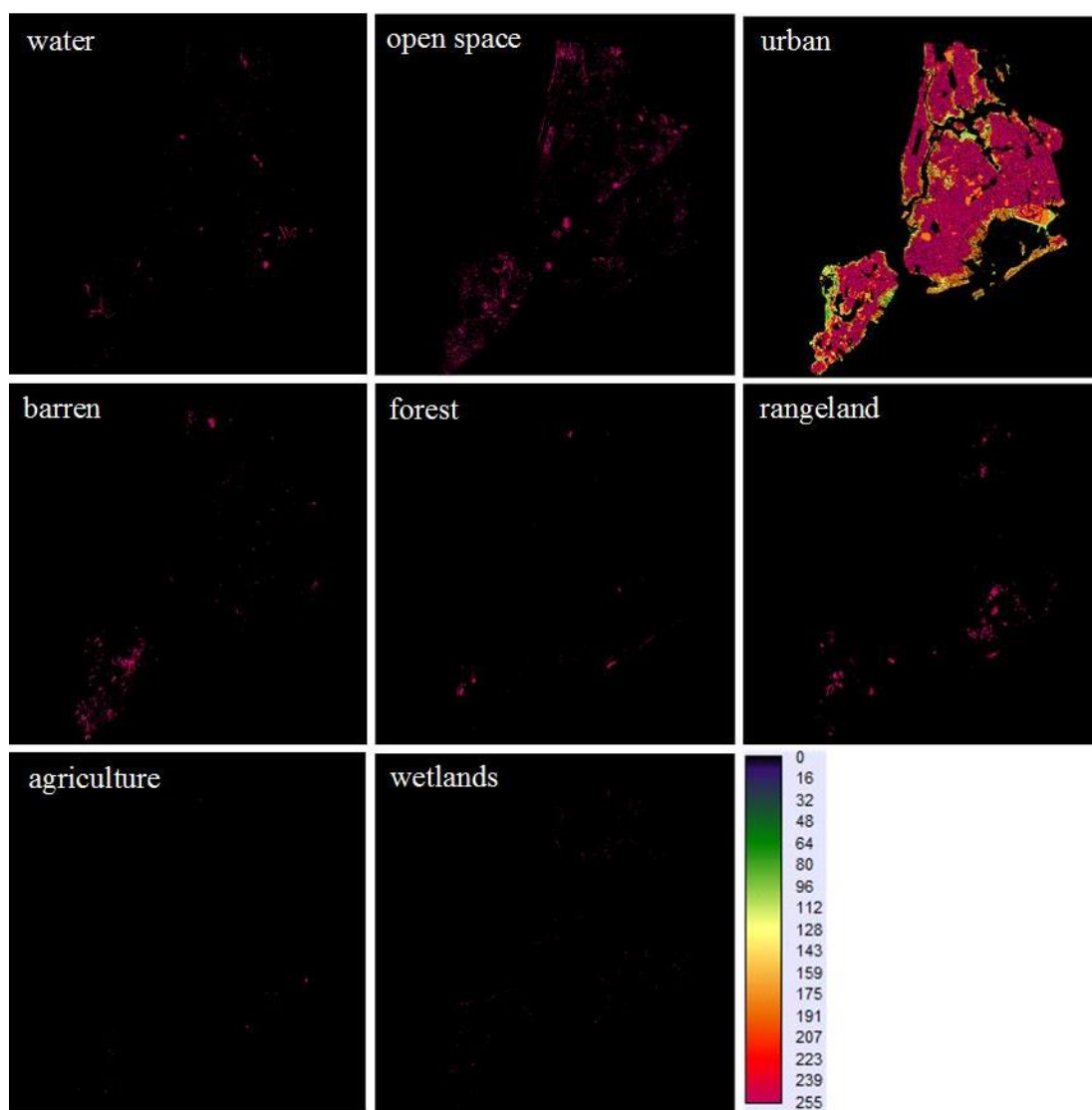
The weight of each factor for possible urban land use changes is assigned based on the AHP outputs and listed in Table 2-7. Their corresponding functions and control points are also included. All constraints and factors with their weights were the inputs to the MCE-WLC, which generated the suitability maps for each land use class.

**Table 2-7.** Extracted weights based on AHP of urban land use class

Factors		Functions	Control points	Weights
Distance from roads	J-shape decreasing		0-50m highest suitability	0.2978
			50-7km decreasing suitability	
			>7 km lowest suitability (equal)	
Distance to wetlands	Sigmoidal increasing		0-46m no suitability	0.0993
			46-91m increasing suitability	
			>91m equal suitability	
Slope	J-shape decreasing		0% highest suitability	0.0367
			0-15% decreasing suitability	
			>15% no suitability	
Distance to high flood risk area	Sigmoidal increasing		0-200m increasing suitability	0.2831
			>200 highest suitability	
Distance to SLR area	Sigmoidal increasing		0-100m increasing suitability	0.2831
			>100 highest suitability	



The suitability map for each land use class is the input to the UGM model. According to the modeling output from Terrset, the suitability map for each land use class reveals the holistic potential for urban growth. As shown by the continuous suitability scaled from 0 to 255 as indicated in the legend of Figure 2-5, the larger the number in the scaled color scheme, the higher the suitability of being the designated land use class such as water, open space, urban, barren, forest, and rangeland if it is not the case at present. Since the suitability of areas adequate for agriculture and wetland is too small, the corresponding suitable areas in Figure 2-5 are not visible for more interpretation.



**Figure 2-5.** Suitability maps for each land use class with continuous suitability scaled from 0 to

255

### 2.5.2.3. Calibration, Validation, and Prediction

The observed and predicted 2006 land use maps are in agreement (Table 2-8). Moreover, the validation result between the projected and observed 2011 land use maps is assessed using the Kappa statistic. Since all the Kappa statistics ( $K_{\text{standard}} = 0.9519$ ,  $K_{\text{no}} = 0.9823$ , and  $K_{\text{location}}$

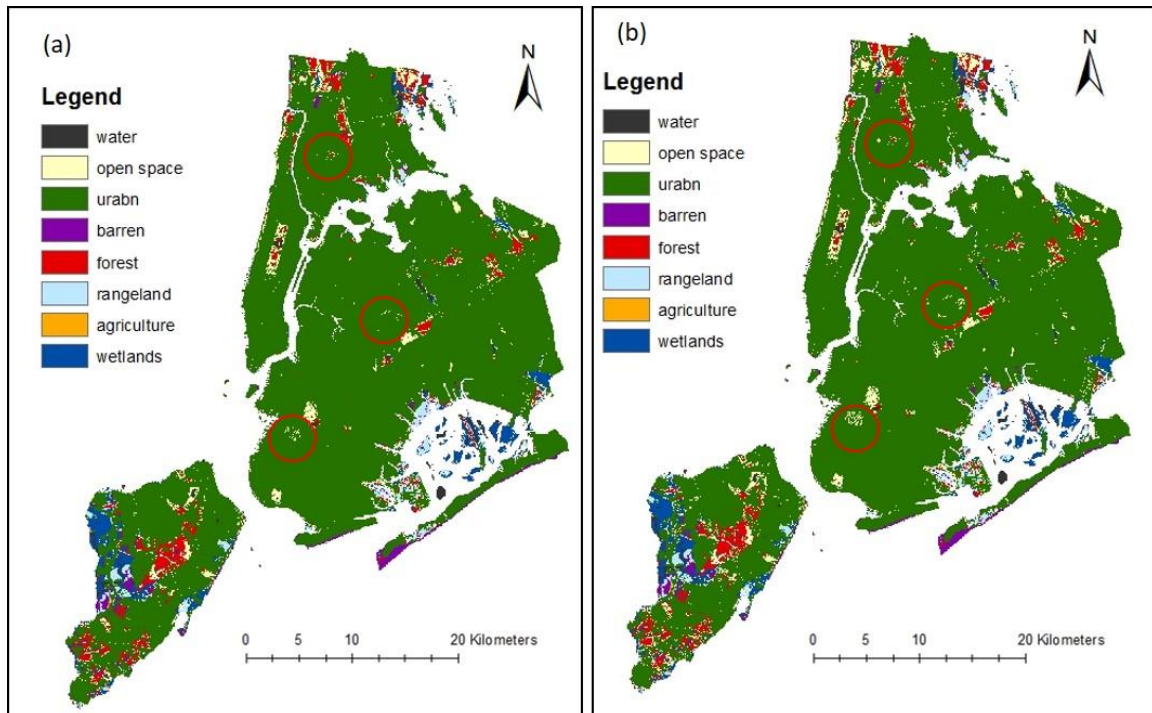
= 0.9692) indicate an outstanding model performance, the fuzzy CA-based MC model is therefore credible and the modeling output applied for land use change prediction is therefore considered valid.

**Table 2-8.** Calibration results based on the observed and predicted land use maps in 2006

	observed 2006	predicted 2006	difference in %
water	1.02%	1.02%	0.00%
open space	6.27%	6.59%	0.32%
urban	81.58%	81.18%	-0.40%
barren	0.96%	0.95%	-0.01%
forest	4.08%	4.13%	0.05%
rangeland	1.82%	1.82%	0.00%
agriculture	0.13%	0.13%	0.00%
wetlands	4.14%	4.17%	0.03%

The predicted results are shown in Figure 2-6, and quantitative analysis is available in Table 2-9. By 2030 and 2050, open space is expected to decrease by 1.51% and 2.51%, respectively, contributing to the growth of the urban area. Besides the urbanization impact, open space is also affected by the impact of flooding. By 2030 and 2050, urban area is still expected to expand by

about 1.36% and 2.63%, respectively, although NYC has been highly developed. There is a very slight change for other land use classes including water, rangeland, and agriculture. Wetlands will slightly decrease because of urbanization in 2050. In general, urban area will keep expanding within a limited space in the next few decades.

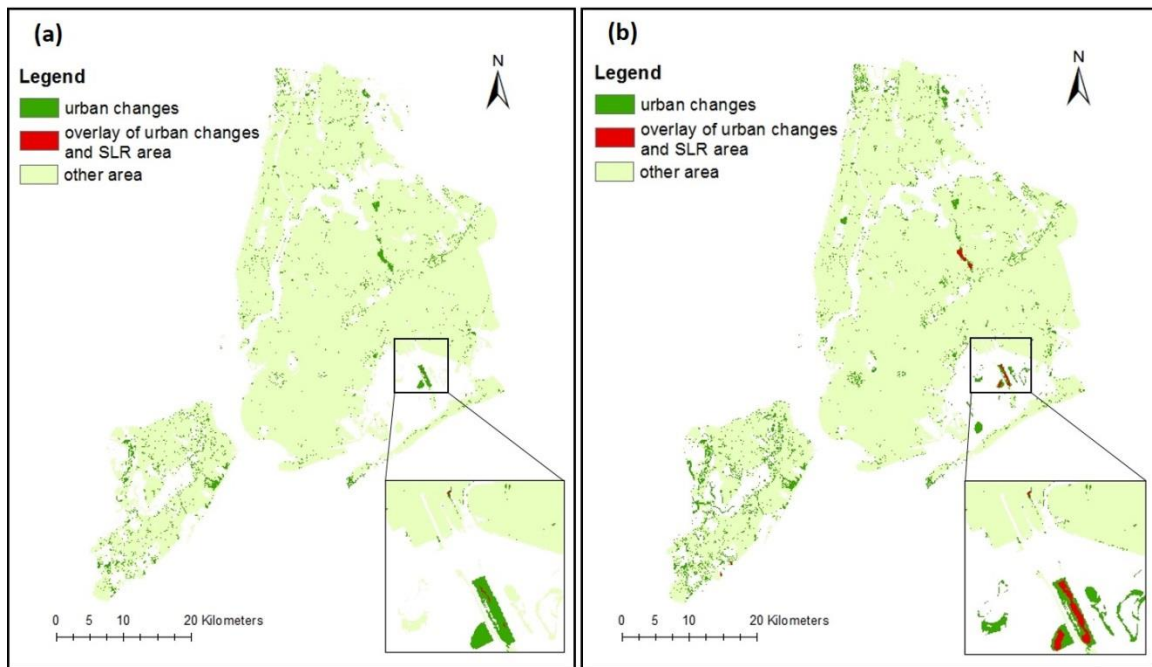


**Figure 2-6.** Predicted land use map in 2030 (a) and 2050 (b) with difference circled

**Table 2-9.** Land use data comparison among 2011, 2030, and 2050 in percentage

land use classes	2011	2030	2050	2011-2030	2011-2050
water	1.01%	1.01%	1.00%	0.00%	-0.01%
open space	6.16%	4.65%	3.65%	-1.51%	-2.51%
urban	82.11%	83.47%	84.74%	1.36%	2.63%
barren	0.98%	1.03%	1.07%	0.05%	0.09%
forest	3.87%	3.91%	3.75%	0.04%	-0.13%
rangeland	1.75%	1.80%	1.79%	0.05%	0.04%
agriculture	0.11%	0.13%	0.12%	0.02%	0.01%
wetlands	4.00%	4.00%	3.88%	0.00%	-0.12%

The overlay maps of urban changes and SLR areas were generated for comparison in 2030 and 2050 (Figure 2-7). Urban changes involve two scenarios: (1) urban area is expected to change to other LULC classes; and (2) other LULC classes are expected to change to urban. Within this urban change detection, the SLR area was overlain on top, indicating the possible areas to be affected. There is a trend that the area affected by SLR could be expanding in the future, giving the assumption that urban growth will still be achieved under the model prediction limited by prescribed climate change scenarios. The major urban changes are along the seashore. More land use protection should be addressed around these conflicting areas to alleviate the impact of SLR especially near the airport region next to Jamaica Bay.



**Figure 2-7** Urban changes under urban growth and SLR simultaneously in (a) 2030 and (b) 2050

Additionally, two different scenarios – the inclusion and exclusion of SLR and the impact of flooding - are compared in this study. The predicted results indicate that there is no obvious difference between these two scenarios in terms of LULC changes. It is also noticeable that the weights associated with SLR and the impact of flooding are both 0.28, which are significant compared to policy-oriented and socioeconomic factors. However, if the climate change situation keeps deteriorating, we might have to increase the weights associated with SLR and the impact of flooding in future modeling scenarios, resulting in significant influences on land use changes. Based on the SDSM prediction results, the mean rainfall will increase by 2030 and 2050. More green space and low impact development may be required to minimize the climate change impact.

Corresponding to this demand, the NYC Green Infrastructure (GI) plan, which emerged in 2010 aims to build more green space, promote low impact development, and decrease the impact of flooding (Bloomberg, 2010).

## 2.6. Conclusion

Land use changes have profound linkages to the constraints and driving factors of urban growth patterns in NYC. The proposed fuzzy CA-based MC model is integrated with an-AHP-based decision analysis framework - MCE-WLC - for modeling spatiotemporal land use changes in NYC under the impacts of socioeconomic development, SLR, and climate change simultaneously for 2030 and 2050. The UGM has been developed, calibrated, validated, and applied for urban growth assessment in this study to explore these profound linkages over the planning horizon. Such a UGM analysis has proved its capability in simulating and predicting urban land use changes. By using the SDSM scenarios to predict future rainfall-runoff variations given the relevant projections of SLR, the results indicate that open space is expected to decrease by 1.51% and 2.51% and urban area is expected to expand by about 1.36% and 2.63% in 2030 and 2050, respectively. Although the mean rainfall will increase in the future based on the SDSM results and SLR is occurring, it is indicative that SLR and the impact of flooding would not affect land use patterns in NYC significantly over the planning horizon.

However, the limitation of the study is that the weights of SLR and the impact of flooding might have to increase in the future, and future weights that are inherently driven by constraints and factors in the analysis of AHP could vary depending on external environmental impacts. Under the impacts of climate change and SLR, coastal megacities like NYC have been continuously

threatened by the acceleration of SLR and the impact of flooding. With the increasing risk of SLR and the impact of flooding in the next decades, finding adaptation strategies to respond to such natural impacts and other socioeconomic influences on land use changes is in an urgent need for NYC as well as other coastal megacities. As such, the GI plan should be promoted, and more green space and low impact development plans are required to minimize the rain-induced flooding hazard under climate change while also incorporating natural defenses along the coastline, such as the coastal wetland restoration, to reduce the accelerating impact of SLR. The NYC government has already implemented some plans since 2010, such as the NYC GI Plan, to provide a more sustainable approach for urban growth.

The proposed model is transferrable and the outcome of this study may further contribute to the advanced assessment of coastal inundation and marsh migration in advanced numerical simulation analyses.

## 2.7. References

- Agarwal, C. (2002) A review and assessment of land-use change models dynamics of space, time, and human choice. Newtown Square, Pa.: U.S. Dept. of Agriculture, Forest Service, Northeastern Research Station.
- Arsanjani, J.J., Helbich, M., Kainz, W., and Boloorani, A.D. (2013) Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion, International Journal of Applied Earth Observation and Geoinformation, 21, 265-275.



- Barredo, J.I., Demicheli, L., Lavalle, C., Kasanko, M., and McCormick, N. (2004) Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria, *Environment and Planning B: Planning and Design*, 32, 65-84.
- Batty, M. (2005) *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. MIT Press, Cambridge. The MIT Press, Cambridge, USA.
- Batty, M., Xie, Y., and Sun, Z. (1999) Modeling urban dynamics through GIS-based cellular automata, *Computers, environment and urban systems*, 23, 205–233.
- Batty, M. (2007) *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press.
- Bhatta, B. (2010) Causes and consequences of urban growth and sprawl, *Analysis of Urban Growth and Sprawl from Remote Sensing Data*. Springer –Verlag Berlin Heidelberg, pp. 17-36.
- Bishop, Y.M., Fienberg, S.E., Holland, P.W., Light, R.J., Mosteller, F. (1977) Book Review: *Discrete Multivariate Analysis: Theory and Practice*. *Applied Psychological Measurement* 1, 297-306.
- Bloomberg, M.R. (2010) NYC Green Infrastructure Plan. NYC Environmental Protection, 1-141.  
Retreieve from [http://www.nyc.gov/html/dep/pdf/green\\_infrastructure/NYCGreenInfrastructurePlan\\_LowRes.pdf](http://www.nyc.gov/html/dep/pdf/green_infrastructure/NYCGreenInfrastructurePlan_LowRes.pdf)
- Carr, H. M. and Zwick, P. D. (2007) *Smart Land-Use Analysis - The LUCIS Model*, ESRI Press, Redwood, CA.

- Cheng, J, Masser, I., and Ottens, H. (2003) Understanding urban growth systems: theories and methods, proceeding of the 8<sup>th</sup> international conference on computer techniques for urban planning and management, Sendai, Japan.
- Couclelis, H. (2002) Modeling frameworks, paradigms, and approaches. Geographic information systems and environmental modelling, Prentice Hall, London.
- City of New York (2016) NYC Environmental Protection: Stormwater. Accessed December 2016. Retrieved from <http://www.nyc.gov/html/dep/html/stormwater/index.shtml>
- Dinar, A., and Mendelsohn, R. O., (2011). Handbook on climate change and agriculture (ed.). Northampton, MA: Edward Elgar Publishing
- Eastman, J.R. (2015) TerrSet tutorial, Clark Labs. Retrieved from [www.clarklabs.org](http://www.clarklabs.org)
- FEMA (2016) Mayor De Blasio and FEMA Announce Plan to Revise NYC's Flood Maps Accessed December 2016. Retrieved from <https://www.fema.gov/newsrelease/2016//17/mayor-de-blasio-and-fema-announce-plan-revise-nycs-flood-maps>
- Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., Hokao, K. (2011) Modeling urban land use change by the integration of cellular automaton and Markov model. Ecological Modelling 222, 3761-3772.
- Guan, Q., Wang, L., Clarke, K.C. (2005) An artificial-neural-network-based, constrained CA model for simulating urban growth. Cartography and Geographic Information Science 32, 369-380.
- Gulacha, M.M. and Mulungu, D.M.M., 2016: Generation of climate change scenarios for precipitation and temperature at local scales using SDSM in Wami-Ruvu River Basin Tanzania,

- Hessami, M., Gachon, P., Ouarda, T. B.M.J., and St-Hilaire, A., 2008: Automated regression-based statistical downscaling tool, *Environmental Modelling & Software*, 23(6): 813-834.
- Houet, T., Hubert-Moy, L. (2006) Modelling and projecting land-use and land-cover changes with a cellular automaton in considering landscape trajectories: an improvement for simulation of plausible future states. *EARSeL eProceedings* 5, 63-76.
- Huong, H.T.L. and Pathirana, A. (2013) Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam, *Hydrology and Earth System Sciences*, 17, 379-394.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D. (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*. 77(3) 35.
- Liu, X., Li, X., Liu, L., He, J., Ai, B. (2008) A bottom-up approach to discover transition rules of cellular automata using ant intelligence. *International Journal of Geographical Information Science* 22, 1247-1269.
- Liu, Y., Phinn, S.R. (2003) Modelling urban development with cellular automata incorporating fuzzy-set approaches. *Computers, Environment and Urban Systems* 27, 637-658.
- López, E., Boccoa, G., Mendozaa, M., and Duhau, E. (2001) Predicting land-cover and land-use change in the urban fringe. A case in Morelia city, Mexico, *Landscape and Urban Planning*, 55, 271–285.

- Mani Murali, R., and Dinesh Kumar, P.K. (2015) Implications of sea level rise scenarios on land use /land cover classes of the coastal zones of Cochin, India, *Journal of Environmental Management*, 148, 124-133.
- Memarian, H., Balasundram, S., Talib, J., Sung, C., Sood, A., and Abbaspour, K. (2012) Validation of CA-Markov for Simulation of Land Use and Cover Change in the Langat Basin, Malaysia, *Journal of Geographic Information System*, 4, 542-554.+
- Meyer, W.B., Turner, B.L. (1992) Human population growth and global land-use/cover change. *Annual review of ecology and systematics* 23, 39-61.
- 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.
- Muller, M.R., Middleton, J. (1994) A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. *Landscape Ecology* 9, 151-157.
- Multi-Resolution Land Characteristics Consortium (MRLC) (2015) National Land Cover Database. Retrieved from <http://www.mrlc.gov/>
- Neumann, J. von and Burks, A. W. (1966) *Theory of Self-Reproducing Automata*, University of Illinois Press, Champaign, Ill, USA.
- New York City Department of Park and Recreation (2012) *New York City Wetlands Strategy*, New York, NY.
- New York City Department of Planning (2016) *New York City Population*. Retrieved from <https://www1.nyc.gov/site/planning/data-maps/nyc-population/population-facts.page>

New York Protected Area Database (NYPAD) (2013) Overview of NYPAD Data. Retrieved from <http://www.nypad.org/>

New York State Department of Environmental Conservation (2015) DEC Proposes New Sea-Level Rise Projection Regulations Continues National Leadership in Efforts to Improve Coastal Resiliency. Retrieved from <http://www.dec.ny.gov/press/103925.html>

Nicholls, R. J. (1995) Coastal megacities and climate change, *GeoJournal*, 37(3), 369-379.

Nicholls, R.J., Cazenave, A. (2010) Sea-level rise and its impact on coastal zones, *Science*, 328, 1517-1520.

NOAA (2011) NOAA's 1981-2010 Climate Normals. Retrived from <http://www.ncdc.noaa.gov/oa/climate/normals/usnormals.html>

Murali, R. M., and Kumar, P. D. (2015) Implications of sea level rise scenarios on land use/land cover classes of the coastal zones of Cochin, India. *Journal of environmental management*, 148, 124-133.

Railsback, S.F., Grimm, V. (2011) Agent-based and individual-based modeling: a practical introduction. Princeton university press, U.S. New Jersey.

Roland, J. F. (2010) Cities at risk: Asia's coastal cities in an age of climate change, *Asia- Pacific Issues*, 96, 1-12.

Saaty, T.L. (1990) How to make a decision: the analytic hierarchy process. *European Journal of Operational Research* 48, 9-26.

Saaty, T.L. (1991) Some mathematical concepts of the analytic hierarchy process. *Behaviormetrika* 18, 1-9.

- Sante, I., Garcia, A.M., Miranda, D., and Crecente, R. (2010) Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning*. 96, 108-122.
- Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., and Taleai, M. (2017) Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. *Computers, Environment, and Urban Systems* 64, 297-308.
- Silva, E.A., Clarke, K.C. (2002) Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment and Urban Systems* 26, 525-552.
- Solomon, S., 2007. *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*. Cambridge University Press.
- Subedi, P., Subedi, K., and Thapa, B. (2013) Application of a hybrid cellular automaton – Markov (CA-Markov) model in land-use change prediction: A case study of Saddle Creek Drainage Basin, Florida, *Applied Ecology and Environmental Sciences*, 1(6), 126-132.
- Tayyebi, A., Tayyebi, A.H., Arsanjani, J.J., Shafizadeh-Moghadam, H., and Omrani, H. (2016) FSAUA: A framework for sensitivity analysis and uncertainty assessment in historical and forecasted land use maps. *Environmental Modeling & Software* 84, 70-84.
- The USGS Land Cover Institute (2016) NLCD 92 Land Cover Class Definitions. Retrieved from <https://landcover.usgs.gov/classes.php>
- Teutschbein, C., Wetterhall F. and Seibert J., 2011: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dynamics*, 37(9): 2087–2105.
- Torrens, P.M. (2000a) How cellular models of urban systems work (1. Theory). Centre for

Advanced Spatial Analysis Working, WP-28, University College London

Torrens, P.M., (2000b). How cellular models of urban systems work (1. Theory).

United States Census Bureau (2015) Annual estimates of the resident population: April 1, 2010-2014 population estimates. Retrieved from [www.census.gov](http://www.census.gov)

Wilby, R. L., & Dawson, C. W. (2007) SDSM 4.2-A decision support tool for the assessment of regional climate change impacts. User Manual. London, UK.

Wilby, R.L., Dawson, C.W., Barrow, E.M. (2002) SDSM—a decision support tool for the assessment of regional climate change impacts. *Environmental Modeling and Software*. 17(2): 145-157.

White, R., Engelen, G. (1993) Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment and planning A* 25, 1175-1199.

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. (1998) Simulating urban encroachment on rural land with fuzzy-logic-controlled cellular automata in a geographical information system. *Journal of Environmental Management* 53, 293-308.

Xiao, J., Shen, Y., Ge, J., Tateishi, R., Tang, C., Liang, Y., Huang, Z. (2006) Evaluating urban expansion and land use change in Shijiazhuang, China, by using GIS and remote sensing. *Landscape and urban planning* 75, 69-80.

Xie, C., Huang, B., Claramunt, C., Chandramouli, C. (2005) Spatial logistic regression and GIS to

model rural-urban land conversion, Proceedings of PROCESSUS Second International Colloquium on the Behavioural Foundations of Integrated Land-use and Transportation Models: frameworks, models and applications. University of Toronto, pp. 12-15.

Zadeh, L.A. (1968) Probability measures of fuzzy events. Journal of mathematical analysis and applications 23, 421-427.



## **CHAPTER 3 EXPLORING THE POTENTIAL CLIMATE CHANGE IMPACT ON URBAN GROWTH IN LONDON BY A CELLULAR AUTOMATA-BASED MARKOV CHAIN MODEL**

### **3.1. Introduction**

As one of the megacities in Europe with the highest population density, London has its unique geographic location having the interactive relationship with its networked cities and corridors achieving synchronous growth. Land use changes are directly impacted by most economic, social, and environmental activities, reflecting the urban development and growth (Litman, 1995). However, in parallel with its population increase and economic development as the financial center of Europe, the area of forest and agriculture had decreased significantly between 2000 and 2006, contributing to the growth of urban and open space. As flooding has been threatening the economic development of the city since the 1700s (Greater London Authority, 2002), the impacts of climate changes need to be addressed in order to analyze the driving factors of land use changes in future scenarios. However, the interactions between climate change and urban growth in the context of economic geography are unclear due to miss links in between within the recent urban planning in London by more deliberative way.

Economic geography is considered as a terminology in terms of location, distribution, and spatial organization involving economic activities worldwide (Clark et al., 2003). The combination of the fundamental network structure and the large size of systems consists of the complex networks that could be a striking approach for growth dynamics modeling (Andersson et al., 2006). Yet incomputable networks and connections are extremely difficult to be counted or identified within the systems. Analyzing some schematic structure or simple systems to represent a dynamic growth on a large-scale region remains difficult to achieve. Therefore, a comprehensive and

systematic approach based on different levels of factors and networks is required in order to study the complex evolving system (Barabási and Albert, 1999).

Macro-level modeling analysis involves significant portion of internal dynamics of the objects, relatives, and implications (Andersson et al., 2006). Such dynamic models can predict potential changes with global parameters, which differs with conceptual models designed simply to address some fundamental questions. As one of the representative models in urban planning regime, grid-based cellular automata (CA) models can simulate the multiplicative growth at micro level from more detailed local activities of simple systems. Furthermore, CA models can be agglomerated in varying scales to analyze the spatial variation of network as well as to predict potential urban evolutionary changes in a long period of time. The history of CA dates back the 1940s (Von Neumann and Burks, 1966) and was further developed by numerous studies (Batty, 2007; Clarke et al., 1997; Takeyama and Couclelis, 1997; Torrens, 2000; Xie, 1996). As extensions of CA, some approaches such as artificial neural networks (Hopfield, 1988; Yegnanarayana, 2009), logistic regression (Hu and Lo, 2007; Peduzzi et al., 1996), and Markov chain models (MC) (Guan et al., 2011; Kamusoko et al., 2009) are more related to time-series prediction, which are general methodologies rather than only applications for simulating urban evolutionary processes.

This study relies on rainfall prediction, with the aid of the Statistical DownScaling Model (SDSM) (Wilby et al., 2002), which provides the statistical foundation for future flooding potential in within the urban space of London while considering major socioeconomic policies related to land use management. In this context, a CA-based MC model was applied for a spatiotemporal assessment taking advantage of the spatial of CA models and temporal coverage of MC models simultaneously (Memarian et al., 2012). In other words, CA models can spatially simulate land

use and land cover (LULC) changes in large scale based on the grid-based principle, whereas MC models are able to quantify the actual amount of changes over time periods dynamically (Moghadam and Helbich, 2013). This can be achieved due to the fact that MC models use the mathematical theory to calculate the probability of changes during two certain time periods (Arsanjani et al., 2013). Therefore, CA-based MC models are ideally integrated as a modeling system to simulate LULC changes spatially based on the calculated likelihood of LULC changes over specified temporal scales mathematically. The proposed model that addresses spatiotemporal complexity specifically can thus consider long-term precipitation variability and climate change impact as a driving force to search for better urban planning strategies

The objective of this study is therefore to simulate and predict LULC changes for 2030 and 2050 in London by using an integrated CA-based MC model regarding the interactions between climate change and urban growth from local networks and large scale dynamics to increase the insight for future scenarios of LULC changes. Rainfall forecasting conducted by SDSM provides the statistical foundation of climate change and flood impact in 2030 and 2050. Two scenarios are developed to explore the difference between inclusion and exclusion of climate and flood impacts. The applied software including geographic information system (GIS) and IDRISI integrated in TerrSet can collaborate and realize the goal of simulation and prediction of land use changes. With this model settings, the study may lead to analyze and answer the following three scientific questions: (1) How will rainfall change in 2030 and 2050 and what are its potential impacts on flooding within the urban space of London? (2) How would the LULC patterns of London change under the impact of climate change (i.e., rainfall variability) in 2030 and 2050 as a whole? and (3)

What are the differences between the inclusion and exclusion of flood impact regarding to LULC changes in 2030 and 2050?

### 3.2. Study area

London, the capital of the UK located in the southeast England, is the largest metropolitan area in the country, as well as one of the largest urban zones in Europe. The city covers about or 1572 square kilometers (607 square miles) with more than 8.67 million populations in 2015. London is vulnerable to flooding from various sources, including storm surge and fluvial flood from the Thames River and surrounding tributaries as a result of heavy rainfall and from the overwhelmed drainage systems (Greater London Authority, 2002). Flooding can also be exacerbated by increase in urbanized regions and lack of green space to retain runoff.

Following World War I, green space was seen as an alternative to shift rapid urban densification in the city's center toward with more focus on suburbanization. However, the need for suburbanization threatened available green-space as policy makers sought to develop these areas in near the mid-20<sup>th</sup> century (Rotenberg, 2008). Toward the late 20<sup>th</sup> century, policy makers were seeking new solutions to balance urban growth while protecting as much of the green space surrounding London (Urban Task Force, 2002). Recent practices on the policy level for the 21<sup>st</sup> century have sought to not only protect green space surrounding the center of London but also to incorporate more green space within the city as part of sustainable drainage systems (SuDS), which can be utilized to address concerns over flood risk in London (Greater London Authority, 2015).

### 3.3. Data and Software

The data required in this study involves land use, land cover, road, Green Belt land, Central Activities zone boundary, areas for intensification points, opportunity area points, strategic industrial location points, London brownfield sites, digital elevation data, parks and gardens, and buildings (Table 3-1). The Greater London Authority established a new center of commercial and residential community in the Old Oak and Park Royal Development Corporation (OPDC) in 2015 (OPDC, 2017). In this case, the OPDC and the central activities zone can be integrated as one layer for the following data process. All the data require pre-processing and appropriate conversions between shapefile and raster file formats to ensure no extra error is introduced during the process.

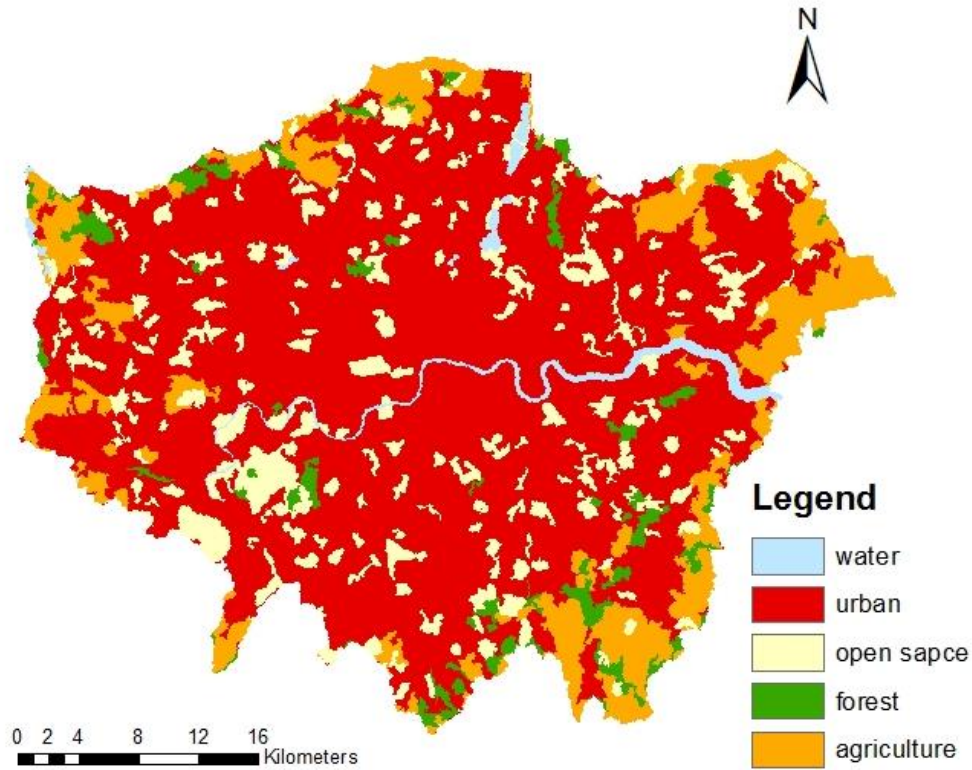
**Table 3-1.** Applied dataset and source

<b>Data</b>	<b>Description</b>	<b>Date</b>	<b>Source</b>
Land use land cover	100x100m resolution	2000,2006, 2012	Corine Land Cover
Road	Existing road	2016	Ordnance Survey Open Data
Central Activities Zone Boundary	For strategic finance, specialist retail, tourist and cultural uses and activities, residential, and other local functions	2009	London Datastore

<b>Data</b>	<b>Description</b>	<b>Date</b>	<b>Source</b>
Green Belt land	—	2011	London Datastore
Areas for intensification points	For residential, employment and other uses	2009	London Datastore
Opportunity area points	For new employment and housing	2009	London Datastore
Strategic industrial location points	For general business, industrial, warehousing, waste management, and transport sectors	2009	London Datastore
OPDC	New centre and community for development	2016	London Datastore
London brownfield sites	For redevelopment and reuse	2010	London Datastore
Digital elevation data	90x90m resolution	2010	USGS SRTM 1 Arc-Second Global
Parks and gardens	Historic parks and gardens	2016	Historic England
Recorded flood area	Records from 1706 to 2015	1706-2015	UK Environment Agency

<b>Data</b>	<b>Description</b>	<b>Date</b>	<b>Source</b>
Buildings	Existing buildings	2016	Historic England

The original LULC data (2000, 2006, and 2012) include 19 to 20 land use classes which is excessively precise but not easy for distinguishing especially when land use changes. According to the data source (Corine Land cover, 2017), the data can be classified by 3 approaches: coarse classification (4 classes), moderate classification (15 classes), and sophisticated classification (44 classes). The original attribute table of the land use data presents the land use class as the sophisticated classification approach, but they are not consistent for land use data in 2000, 2006, and 2012. For example, land use data in 2000 does not include intertidal flats, but 2006 and 2012 land use data have this land use class. To better understand the land use changes and unify the land use classes, the initial land use classes are reclassified by combining the first and second approaches which are finally displayed as five classes including water, urban, open space, forest, and agriculture (Figure 3-1). The reclassified water includes water courses, water bodies, and estuaries. Similarly, urban includes continuous urban fabric, discontinuous urban fabric, industrial or commercial units, road and rail networks and associated land, port areas, airports, mineral extraction sites, dump sites, and construction sited. Green urban areas, and sport and leisure facilities are reclassified as open space. Forest includes broad-leaved forest, coniferous forest, and mixed forest. Agricultural indicates non-irrigated arable land, pastures, complex cultivation patterns, fruit trees and berry plantations, land principally occupied by agriculture with significant areas of natural vegetation.



**Figure 3-1.** London land use in 2000 with re-classified land use classes (source: CLC, 2000)

The reorganized data sets (Table 3-2) show that the area of forest and agriculture decreased significantly between 2000 and 2006, contributing to the growth of open space. However, the difference of each land use class between 2006 and 2012 is not significant. The reason could be human activity-oriented policies and planning strategies changed after 2006. The analysis of the land use data provides basic understanding of the land use changes and statistic foundation of the model.



**Table 3-2.** Reclassified land use class area and difference comparison (unit: hectares)

land use class	2000	2006	2012	2000-2006 difference	2006-2012 difference
Water	2736	2750	2750	0.51%	0.00%
Urban	109807	110670	110943	0.79%	0.25%
open space	19580	22540	22526	15.12%	-0.06%
Forest	5972	3224	3215	-45.01%	-0.28%
agriculture	21501	20412	20162	-5.06%	-1.22%

In this study, the software of GIS and TerrSet are mainly used. GIS provides numerous functions to visualize and analyze data and has been used in comprehensive disciplines. Data pre-processing and format unification are achieved by using GIS. TerrSet, developed by Clark Labs at Clark University, is an integrated geospatial software with the ability to incorporate the IDRISI GIS analysis for monitoring and modeling purpose (Eastman, 2015). The submodule IDRISI integrated in TerrSet, for digital geospatial information analysis and display, is also adopted to quantify and predict land use changes.

### 3.4. Methodology

#### 3.4.1. Flood impact concern & SDSM

To determine the effect of possible changes in climate would have on rainfall, statistical modeling techniques were sought to analyze historical rainfall within the London city limits and combine data with climate modeling to project rainfall for 2030 and 2050. Increases (decreases) in annual rainfall can reveal that more flooding (less flooding) can occur. For this reason, this study aims to determine the annual rainfall for the years 2030 and 2050. The SDSM useful in this regard and was applied in this study to determine statistical relationships, based on multiple linear regression techniques, between large-scale climate and local climate. These relationships were developed using observed weather data and, thus these relationships can be used to obtain downscaled local information for some future time period using Global Climate Model (GCM) derived atmospheric predictors assuming that they remain valid in the future.

For this study, daily observed rainfall was a required input into SDSM. Daily rainfall data was obtained from the Meteorological Office of the UK (Met Office) via the Met Office Integrated Data Archive System (MIDAS). Stations with the longest, continuous period of daily rainfall record were chosen for at least a 20-yr period or longer and are summarized in Table 3-3. The period of 1961-1990 was a limiting factor for SDSM calibration and validation since two out of the three rainfall stations had a continuous record for this period in question.

**Table 3-3.** Rainfall Stations in London with Continuous Period of Daily Rainfall Record

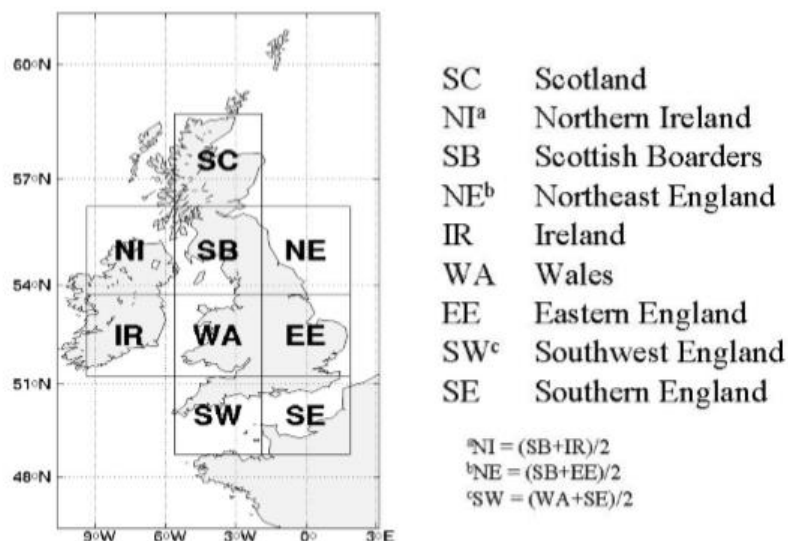
(20+ yrs.)

MIDAS Station Name	Lat/Long (Decimal Deg)	Continuous Period of Record
Longford-Perry Oaks Station  (SRC* ID: 5652)	51.47 N/ 359.5 W	1961-2000
Woodford Station  (SRC* ID: 5533)	51.59 N/ 0.00 E	1961-1990
Hogsmill Valley  (SRC* ID: 6570)	51.40 N/ 359.7 W	1961-1990

*3.4.1.1. SDSM predictor variables*

Selection of the best predictors is an iterative process, which removes the least significant predictors until the remaining predictors are statistically significant, such that a clear relationship between predictor variables and the “predictand” or observed variable of concern could be determined. SDSM relies on both observed-derived and GCM-derived atmospheric predictor variables. Global observed predictor variables such as those from National Center for Environmental Prediction (NCEP) (Kalnay et al., 1996) are used for SDSM calibration and validation while GCM-derived predictor variables are utilized within the “Scenario Generator”

component of SDSM for future projection (Wilby et al., 2002). GCM-derived atmospheric predictor variables were obtained from the Hadley Centre Coupled Model, version 3 (HADCM3) GCM A2 scenario of the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC, 2007). HADCM3 was chosen due to its source from one of the leading centers on climate change research in the UK, the Hadley Centre. All atmospheric predictor variables were re-gridded to a standard co-ordinate system ( $2.5^\circ$  latitude  $\times$   $3.75^\circ$  longitude) used in HADCM3 (Figure 3-2) covering a period from 1961 to 2099 [<http://www.cccsn.ec.gc.ca/?page=pred-hadcm3>] and normalized with respect to base period (1961 to 1990) averages. Normalization is required since bias or error can occur from large scale predictions not performing well at simulating the climate of a particular location such as large differences between observed and larger scale conditions.



**Figure 3-2.**  $2.5^\circ$  latitude  $\times$   $3.75^\circ$  longitude grid system used by Hadley Centre's coupled ocean-atmosphere GCMs and NCEP. (Source: Wilby et al., 2002)

#### *3.4.1.2. SDSM-Calibration & Validation*

Given 41 years of global NCEP daily observed-predictor variables (1961-2001), the first 14 years (1961–1975) are used for calibrating the SDSM model, while the years 1976-1990 of predictor variables are used to validate the model. Before calibration and validation, SDSM requires that the data is normally distributed. Daily observed rainfall over a period of record typically exhibits positively skewed behavior since the occurrence of small rainfall events is higher than occurrences of large rainfall events such as a high intensity rain event. As a result, a transformation of the data was required to obtain a normal distribution. Near-normal distribution was achieved using the log transformation of the rainfall data for both stations. The following predictors were used for SDSM model calibration: Surface airflow strength, Surface zonal velocity, Surface meridional velocity, Surface vorticity, Surface wind direction, Surface divergence, 500 hPa airflow strength, 500 hPa zonal velocity, 500 hPa meridional velocity, 500 hPa vorticity and 500 hPa geopotential height.

#### *3.4.1.3. SDSM – Future projections (2030 & 2050)*

Future projections of rainfall were conducted using the scenario generator in SDSM. The scenario generator produces ensembles of synthetic daily weather series for future climate using the calibrated SDSM model and GCM-derived future atmospheric predictor variables. GCM-derived future atmospheric predictor variables can only be used to project future climate since the previously used NCEP observed predictor variables are not valid for future predictions: The following predictors, from the HADCM3 GCM A2 scenario, were used to project future climate: Surface airflow strength, Surface zonal velocity, Surface meridional velocity, Surface vorticity,

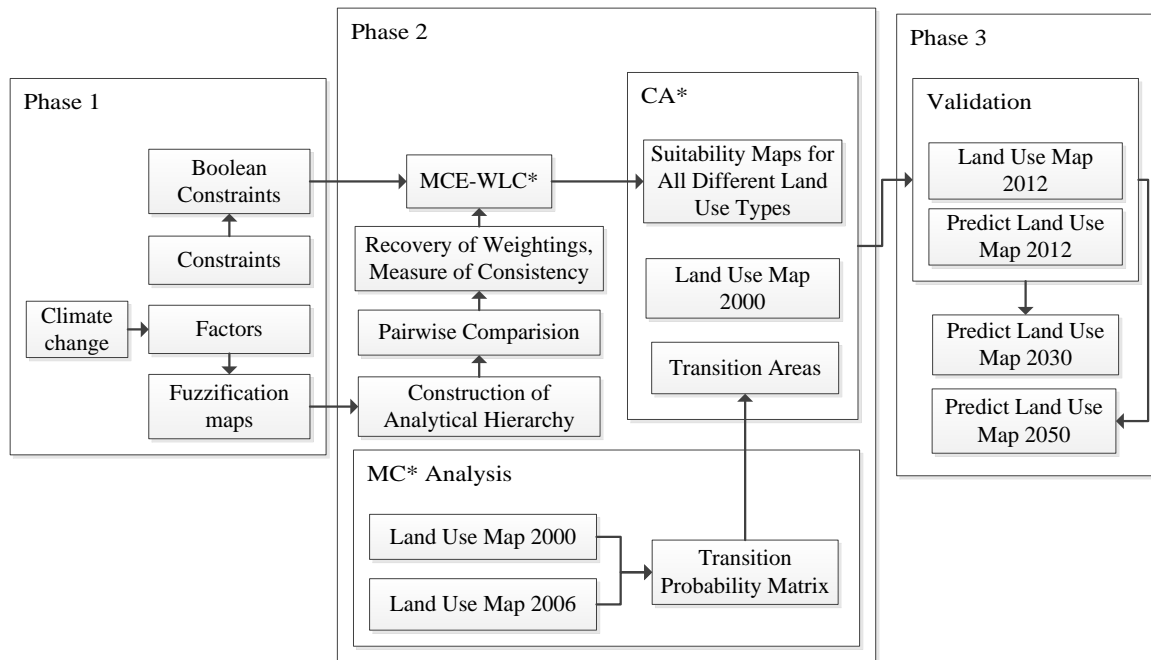
Surface wind direction, Surface divergence, 500 hPa airflow strength, 500 hPa zonal velocity, 500 hPa meridional velocity, 500 hPa vorticity and 500 hPa geopotential height. After future predictors are established, annual rainfall projections over a 25-member SDSM ensemble were sought to provide a range of possible annual rainfall values for the years 2030 and 2050. From these possible range of values, each year of the historical record (1961-1990) can be compared to the 20-member ensemble projections for the year 2030 and 2050 to determine if projections show an increase or decrease in annual rainfall.

#### 3.4.2. CA-based MC model

An integrated CA-based MC model is applied in this study to predict LULC changes for the city of London in the target year 2030 and 2050 with the consideration of climate change (i.e. flood impact). MC models are able to quantify the actual changes between land use classes, but not spatially (Moghadam and Helbich, 2013). By incorporating with MC model, a CA-based MC model can quantify and predict the spatiotemporal changes based on the transition probabilities for each local pixel (Subedi et al., 2013). The CA-based MC model is applied to simulate the stochastic process with the assumption of stationary probability distribution over time.

In general, the model structure and flowchart can be summarized in Figure 3-3. In phase 1, the model starts with setting up factors and constraints for each land use class, followed by applying fuzzy function for each factor, and assigning Boolean values (0 or 1) for constraints. In phase 2, construction of analytical hierarchy is used for further process of fuzzification maps by using pairwise comparison and assigning weights for each factor. The consistency ratio (CR) is used to measure the consistency of the weights. The Weighted factors and Boolean constraints are imported to multi-criteria evaluation-weighted linear combination (MCE-WLC) process to generate

suitability maps for all different land use types. In the meantime, the transition probability matrix is built by using MC analysis based on the land use maps in 2000 and 2006. The transition areas file is also an outcome of MC analysis and generated from Markov transition matrix with the quantity analysis of expected LULC change in the next time period. CA model can be run by inputting the collection of the suitability maps for all different land use types, land use map 2000, and transition areas file. A 5 by 5 continuity filter needs to be applied in the CA model setting in order to impact the neighborhood cell changes. The projected land use map 2012 is the outcome of the CA model, as well as the input of validation in phase 3 that compares the level of agreement between the observed and the projected land use maps of 2012 based on the Kappa statistics. If the validation result is valid, then the model is prepared for predicting land use map in 2030 and 2050.



**Figure 3-3.** Study framework; \*CA: cellular automata; MCE-WLC: multi-criteria evaluation-weighted linear combination; MC: Markov Chain

### 3.4.2.1. MC

MC models can analyze the likelihood of a random process from one stage to another via the use of transition probability matrices (López et al., 2001) such that they have been widely applied to quantify the temporal variations between land use classes (Moghadam and Helbich, 2013). In this study, a probability matrix based on the likelihood of the land use changes between 2000 and 2006 was used for predicting LULC map in 2012, 2030, and 2050. Assuming a constant rate of transition and setting up 2000 as base map, the corresponding time period to project forward from 2006 should be 6 years, 24 years, and 44 years for 2012, 2030, and 2050, respectively. The mathematic formation of MC is shown as follow:

$$\sum_{j=1}^m P_{ij} = 1 \quad i = 1, 2 \dots m \quad (3-1)$$

where  $P_{ij}$  represents the probability that land use changes from type  $i$  in time  $t_0$  to type  $j$  in time  $t_1$ . The matrix is built based on the combination of the probabilities that different land use types change to other land use types on a certain time interval. The structure of the probability matrix  $P$  can be present as follow (Adhikari and Southworth, 2012):

$$(V_i \times P_{ij}) = (V_1, V_2, V_3 \dots V_n) \times \begin{pmatrix} P_{11}, P_{12}, P_{13} \dots P_{1m} \\ P_{21}, P_{22}, P_{23} \dots P_{2m} \\ P_{m1}, P_{m2}, P_{m3} \dots P_{mm} \end{pmatrix} \quad (3-2)$$

where  $V_i \times P_{ij}$  describes the proportion of the second land use reference year;  $P_{ij}$  is the transition probability of land use;  $V_i$  is the proportion of the first land use reference year (vector);  $i$  is the land use type in the first land use reference year;  $j$  is the land use type in the second land use reference year;  $P_{11}$  is the probability that land use type 1 in the first land use reference year changes to land use type 1 in the second land use reference year;  $P_{12}$  is the probability that land



use type 1 in the first land use reference year changes to land use type 2 in the second land use reference year, and so on;  $m$  is the number of land use types classified in the study.

In the CA-based MC model trained in TerrSet, the first outputs are the transition probability matrices. Area-based tables are constructed to present the probabilities for each land use in a convenient and straightforward way. The transition matrices indicate the probability of each land use change to other land use types during a certain time period. Since the 2000 land use map is identified as the base map, the transition probability matrix is generated based on the time difference between 2000 and 2006, and the 6 years interval is set up to project forward from 2006. Year intervals need to be updated for predicting LULC maps in 2030 and 2050 additionally.

#### 3.4.2.2. *CA-based MC Model*

CA models can predict LULC changes based on two time intervals (Yang et al., 2012) such that they have been adopted to simulate the spatiotemporal variations of LULC in urban growth modeling (Neumann and Burks, 1966). CA models are dynamic and discrete models in terms of time, space, and state (Balzter et al., 1998). Four parameters including lattice  $L$ , state space  $Q$ , neighborhood template  $\delta$ , and local transition function  $f$  are applied to describe CA (Adamatzky, 1994):

$$A = [L, Q, \delta, f] \quad (3-3)$$

The state of each cell is defined based on  $Q$  as well as cells are connected geometrically in different ways. The variation of all the cell states simultaneously indicates discrete time-steps, regulated by the neighborhood transition rules. With the collaboration of MC, cells are able to pass

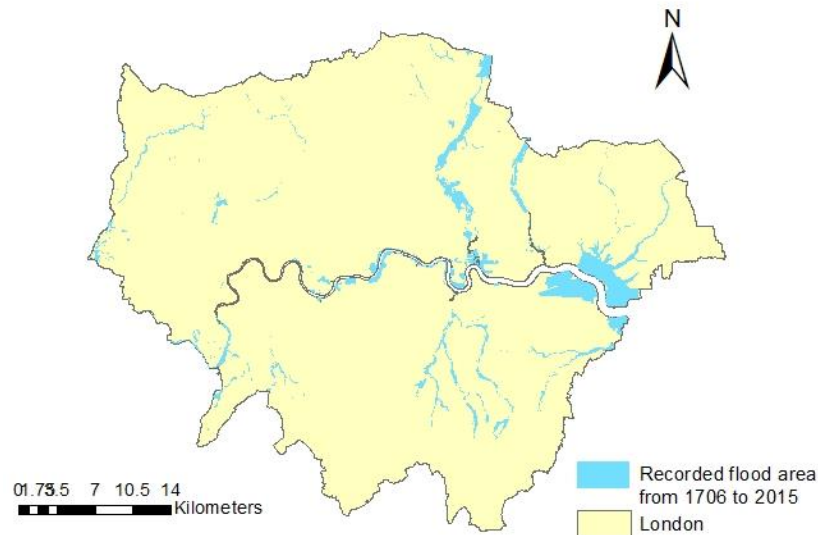
separately through the MC under the control of the neighborhood transition rules with a certain probability in the CA-based MC model. Spatio-temporal analysis is conducted in this model to describe the transition probabilities of each cell continuously.

#### *3.4.2.3. Factors and Constraints*

Factors and constraints play important roles in the model since they will impact the formation of suitability maps for each land use class. Generally, constraints limit the specific geographic location since they are Boolean images comprising only values of 0 and 1. Value 0 indicates that the location does not meet the developing criteria and there is no suitability of land use changes; while value 1 indicates that the location meets the development criteria and has the potential to change to another land use type. Factors or driving factors also indicate the suitability of the location to be developed, but present as continuous values from 0 to 1.

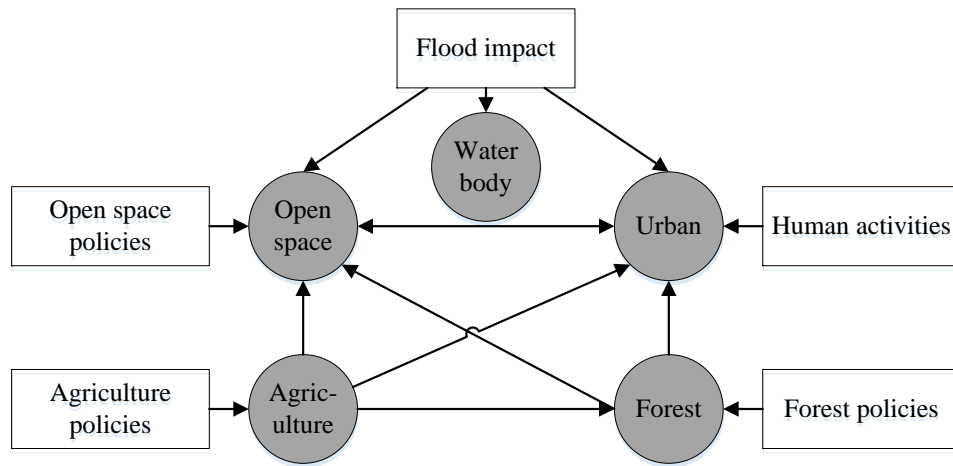
To set up factors for each land use class, it is important to collect and analyze policies and regulations in terms of land use before modeling. Since the major land use changes happened among open space, forest, and agriculture, policies related to these land use types are the key to analyze the changes. Specially, open space plays an important role in flood storage responding to climate change and flood impact (Mayor of London, 2009). In this case, the change of open space depends on the dominant national policy rather than human activities or local policy variations. For another important land use class – forest, especially urban forest, also has the ability to reduce the rate and volume of stormwater runoff and eliminate flooding impact (Dwyer et al., 1992). Large portions of the forest and open space are located in the Green Belt which is defined as a mix-used region and prevent urban sprawl and expansion (Gant et al., 2011). In this regard, forest

will be a significant important factor contributing to the growth of open space, which should be reflected in the model.



**Figure 3-4.** Recorded flood area from 1706 to 2015 in London

In recent years, there has been concern on how climate variability will affect flooding along the Thames River and surrounding tributaries in the future compared to the present (Greater London Authority, 2002) (Figure 3-4). Flood impact area has been considered as a factor of open space that supports natural ecosystem processes and offers direct and indirect benefits for flood control such as absorbing and retaining rainfall (Costanza et al., 1997). In this study, the flood impact is considered as an important factor for both urban and open space since these two land use classes consist of the majority of the total area and they are more sensitive to flood impact. From the holistic perspective, land use variations within each land use class associated with their factors are visualized in Figure 3-5. The overall factors and constraints involved in this study are summarized in Table 3-4.



**Figure 3-5.** Land use classes' interaction relationship associated with their factors

**Table 3-4.** Factors and constraints applied in the model

<b>Factors</b>	<b>Constraints</b>
Distance from roads	Existing buildings
Distance to central activities zone	Water bodies
Slope	The Green Belt
Distance to industrial location	
Distance to intensification points	
Distance to opportunity area	
Distance to London brownfield sites	
Distance to parks and gardens	
Distance to buildings	
Distance to flood area	

#### 3.4.2.4. *Boolean values for constraints and fuzzy membership for factors*

The constraints are assigned as a Boolean value such that the fuzzy membership function only needs to be applied to the corresponding factors. Each land use class needs to be identified with

the corresponding factors and constraints that will become the compositions in the fuzzy membership function. The fuzzy membership function is also incorporated as a necessity to measure the suitability of land use variations, which the potential growth is classified as low, medium, or high for each land use class. Environmental planning and management may impact the fuzzy logic, resulting in a certain degree of flexibility of fuzzy CA. The distance among features (e.g. distance to road, distance to slope, distance to buildings, etc.) are calculated to create buffer zones for building the fuzzification maps. Three types of functions, including linear, sigmoidal, and J-shaped, are determined based on the character of each feature to set up the fuzzy membership. Increasing or decreasing functions also need to be identified, as well as the control point values which are the minimum and maximum distance values calculated previously.

#### 3.4.2.5. Analytic hierarchy process (AHP), MCE-WLC, suitability maps, and CA filters

Given that constraints have been already assigned with values of 0 and 1, each factor defined by fuzzy membership functions is assigned a weight according to an analytical hierarchy process (AHP). Factors are weighted based on their impact on each land use class and interactive relationship between factors. After a weight is assigned to each factor, the consistency ratio (CR) is computed to measure the reliability of the weights of factors assigned to each land use class. The CR is calculated based on producing a new set of ratings according to the computed weights. Ratings are generated by deriving one weight in the comparison with another weight. When the CR converges to 0, it indicates that the assigned weights are ideal. Weights should be retrieved again by AHP if the CR is greater than 0. Herein we accept a maximum CR of 0.1. In this study,

AHP is generated based on scenario analysis using the judgement to adjust the input. Weights are assigned based on socio-economic traditions and policies, and assessed by RC.

After assigning weights to each land use class, MCE-WLC is used to generate the suitability maps for each land use class. There are three types of MCE including WLC, Boolean intersection, and ordered weighted average. In this study, MCE-WLC is used because it offers much more flexibility than the Boolean intersection and allows the criteria to be differentially weighted with full tradeoff capacity (Eastman, 2015).

Suitability maps are used to describe the suitability of land use changes of a certain land use class for each pixel. Value 0 represents no suitability, and 255 indicates the highest suitability of land use changes. All the constraints with Boolean values and factors with continuous values from 0 to 1 need to be incorporated in a MCE-WLC function, generating the suitability map for each land use class. Collecting and grouping all the suitability maps as one file is required in CA-MC model. Before running the model, CA filters need to be identified by combining the suitability maps with MC. In this study, a standard 5×5 Boolean mask filter is used to analyze the neighborhood definition that the suitability weight of the pixels will decrease far from the existing areas and allocate preference to the neighboring suitable areas (Houet and Hubert-Moy, 2006).

In order to assess the accuracy of the suitability maps with the fuzzy membership functions and MCE-WLC approaches, the relative operating characteristic (ROC) is used in this study. The output of the ROC is a text file reporting the area under the curve (AUC), which has been widely used for measuring the accuracy and discriminatory capacity of classification and distribution models (Beck and Shultz, 1986; Greiner et al., 2000; Jiménez-Valverde, 2012). According to Swets (1988), the ROC is defined by rating of likelihood with the binary procedure under different

criteria. Usually five or more category scales with continuous quantity are set up ranging from “very unlikely” to “very likely”. AUC values represent the degree of accuracy between perfect accurate ( $AUC=1$ ), high accurate ( $0.9 < AUC < 1$ ), moderate accuracy ( $0.7 < AUC < 0.9$ ), less accuracy ( $0.5 < AUC < 0.7$ ), and random locations ( $AUC=0.5$ ) (Swets, 1988).

### 3.5. Results

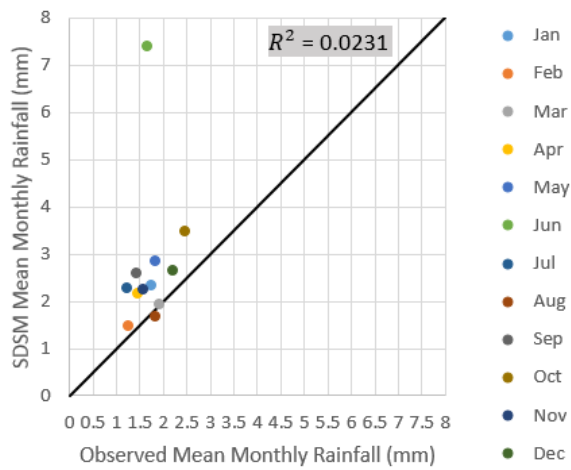
#### 3.5.1. SDSM calibration and validation

Results of calibration for the Longford, Woodford and Hogsmill Valley stations are provided in Table 3-5 on a monthly analysis for the period of 1961-1975. For SDSM calibration results, the R-square value represents the goodness of fit of predictor variables in explaining occurrence of observed rainfall on a monthly basis for each station. Calibration for rainfall was low due to difficulty in large scale predictors in explaining occurrence of local daily rainfall, which has been reflected in literature for certain study regions (Nguyen et al., 2005; Yang et al., 2012). Bias correction and variance inflation parameters in SDSM help improve model after initial calibration from predictors alone.

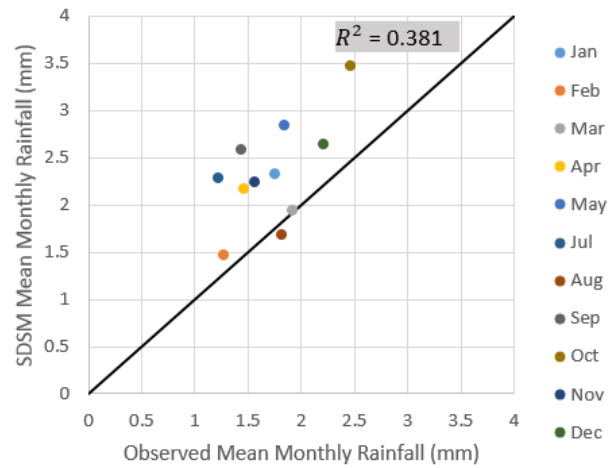


**Table 3-5.** Correlation ( $R^2$ ) between observed rainfall and SDSM predictions at Longford, Woodford and Hogsmill stations for the period 1961-1975

<b>Month</b>	<b>Longford</b>	<b>Woodford</b>	<b>Hogsmill Valley</b>
January	0.249	0.195	0.248
February	0.234	0.217	0.328
March	0.312	0.275	0.277
April	0.208	0.166	0.285
May	0.107	0.122	0.146
June	0.204	0.269	0.444
July	0.224	0.286	0.361
August	0.114	0.129	0.186
September	0.326	0.279	0.446
October	0.162	0.177	0.207
November	0.250	0.255	0.292
December	0.233	0.171	0.168

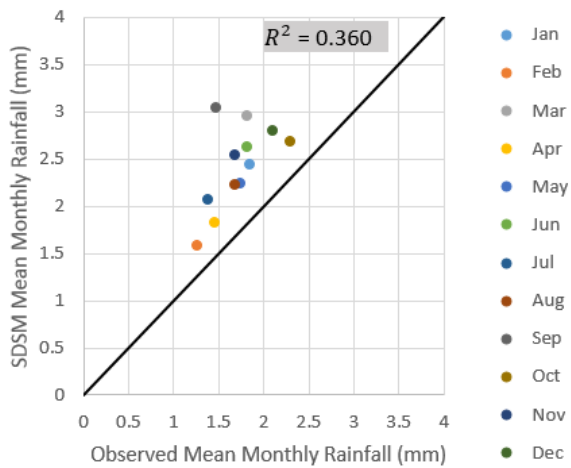


(a)

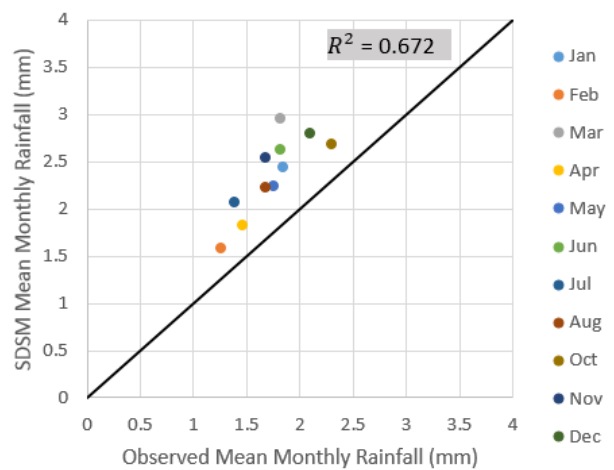


(b)

**Figure 3-6.** Observed vs. SDSM during the validation period (1976-1990) for Longford Station with month of June as an outlier (a) and without month of June as an outlier (b)

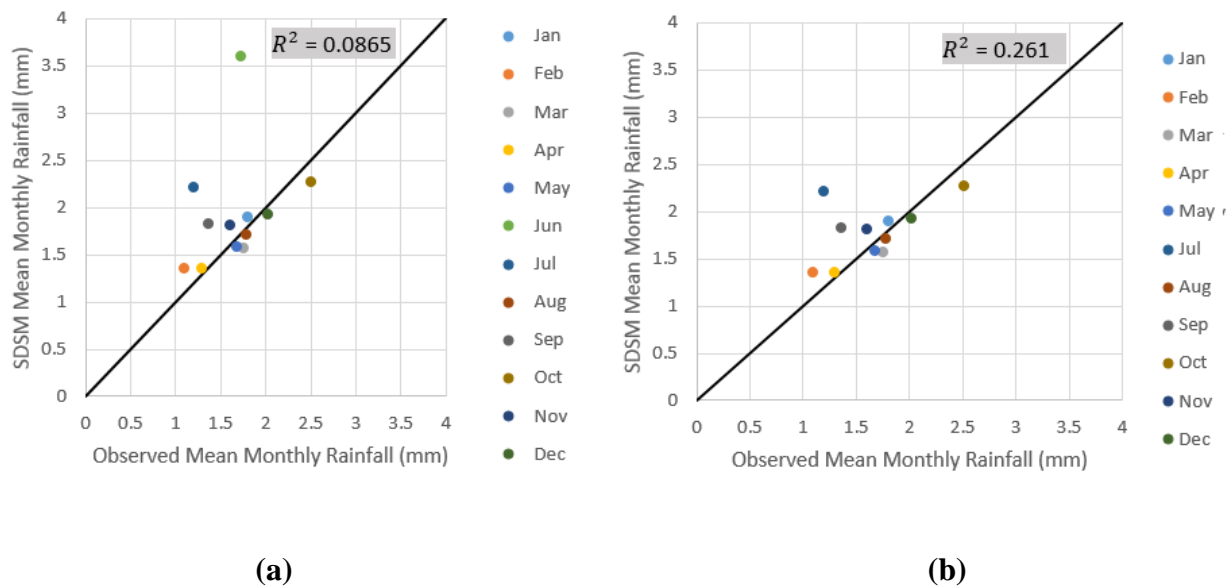


(a)



(b)

**Figure 3-7.** Observed vs. SDSM during the validation period (1976-1990) for Woodford Station with month of September as an outlier (a) and without month of September as an outlier (b)

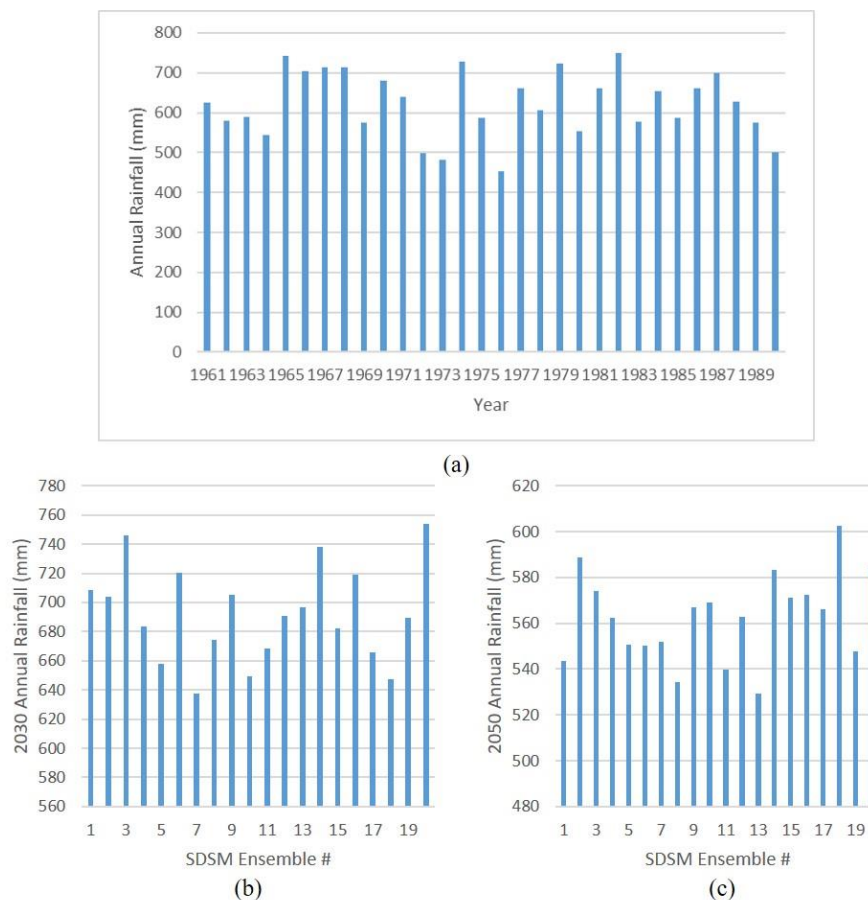


**Figure 3-8.** Observed vs. SDSM during the validation period (1976-1990) for Hogsmill Valley Station with month of June as an outlier (a) and without month of June as an outlier (b)

Results of validation for Longford, Woodford and Hogsmill Valley stations are provided in Figures 3-6, 3-7 and 3-8 respectively, on a monthly basis. Overall, SDSM slightly overestimated the monthly mean rainfall during the validation period (1976-1990) for the Longford station (Figure 3-6) with the exception of the month of February and March. A significant “bias” is evident for the month of June due to model bias for summer rainfall. For the Woodford Station (Figure 3-7), which is slightly northeast of the Longford station, SDSM also slightly overestimated the monthly mean rainfall overall. As opposed to the Longford station however, “bias” is not as evident for the month of June during the validation period. Greater bias is evident for the month of September at the Woodford station. Similar to the Longford station, the Hogsmill Valley station showed significant bias for the month of June (Figure 3-8).

### 3.5.2. SDSM future projections (2030 & 2050)

Due to the poor fit in SDSM model results for the Longford and Hogsmill Valley stations for the validation period, the Woodford Station was used to define the model for future rainfall projections. Annual rainfall projections for the year 2030 and 2050 are presented in Figure 3-9(b) and 3-9(c) respectively. Projected annual rainfall for the year 2030 increases, as compared to historical trends in Figure 3-9(a), for five out of 20-members ensemble under the HADCM3 A2 scenario. Projected annual rainfall for the year 2050 decreases overall as compared to historical trends in Figure 3-9(a).



**Figure 3-9.** Historical Rainfall (a) and SDSM Projected Rainfall (b) and (c) for Woodford Station for the years 2030 and 2050 using HADCM3 A2 scenario

Based on the SDSM prediction results, it can be concluded that the annual rainfall will increase for the year 2030 with respect to baseline years 1961-1990. Irrespective of frequency and duration of rain events, there is 25% chance (from a 20-member ensemble) that the area surrounding the Woodford location will experience greater rainfall on an annual basis in the year 2030 only as compared to historical record resulting in potential for greater chance of flooding with increased urban growth. This analysis needs to be reflected in the CA-MC model to consider the climate change impact in the future land use scenario.

### 3.5.3. TerrSet Transition probability matrices

The transition probability matrix between 2000 and 2006 for predicting LULC map in 2012 is presented in Table 3-6. Based on the matrix, open space, forest, and agriculture contribute most to the growth of urban with probabilities 0.0414, 0.0330, and 0.0417 respectively. Water body has relatively smaller probability to change to urban which is 0.0095.

**Table 3-6.** Transition probability matrix between 2000 and 2006 for predicting LULC map in 2012

	water	urban	open space	forest	agriculture
water	0.9653	0.0095	0.0241	0.0000	0.0011
urban	0.0003	0.9903	0.0078	0.0001	0.0014

open space	0.0006	0.0414	0.9328	0.0061	0.0192
forest	0.0017	0.0330	0.3135	0.4585	0.1934
agriculture	0.0023	0.0417	0.0690	0.0163	0.8707

#### 3.5.4. TerrSet suitability maps

Each land use class needs to set up its constraints and factors based on related policies and transition matrices. Among the land use types, urban is mostly impacted by human activities and policies. Thus, the constraints and factors are more complex and interactive. The factor weights with their functions and control points assigned for urban are shown in Table 3-7.

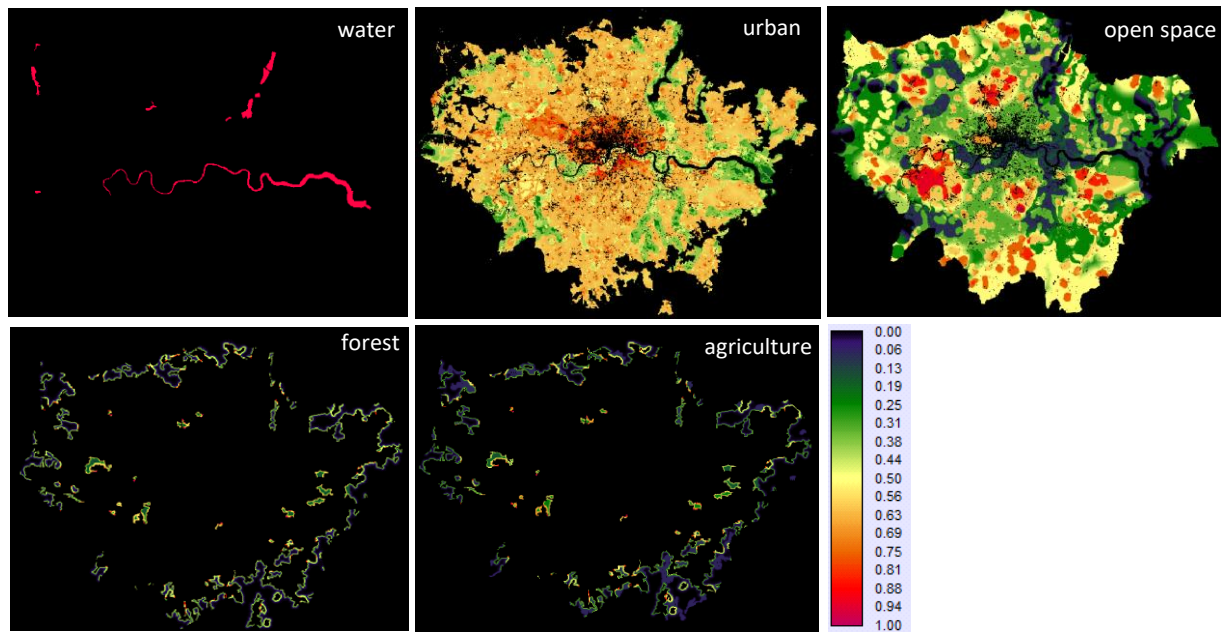
**Table 3-7.** Factors and weights for urban suitability map (CR = 0.03 < 0.1)

<b>Factors</b>	<b>Functions</b>	<b>Control points</b>	<b>Weights</b>
Distance from roads	J-shape	0-50m highest suitability	0.1531
	decreasing	50-1km decreasing suitability	
		>1 km lowest suitability (equal)	
Distance to central activities zone	J-shape	0-1km highest suitability	0.1004
	decreasing	1km- 5km decreasing suitability	
		>2 km lowest suitability (equal)	
Slope	J-shape	0% highest suitability	0.016
	decreasing	0-15% decreasing suitability	
		>15% no suitability	
Distance to industrial location	J-shape	0-500m highest suitability	0.0606
	decreasing	500m- 5km decreasing suitability	
		>2 km lowest suitability (equal)	
Distance to intensification	J-shape	0-200m highest suitability	0.051
	decreasing	200m- 2km decreasing suitability	
		>2 km lowest suitability (equal)	
Distance to opportunity area	J-shape	0-200m highest suitability	0.051
	decreasing	200m- 2km decreasing suitability	
		>2 km lowest suitability (equal)	
Distance to London brownfield sites	J-shape	0-100m highest suitability	0.112
	decreasing	100m- 1km decreasing suitability	
		>1 km lowest suitability (equal)	
Distance to parks and gardens	J-shape	0-200m highest suitability	0.0678
	decreasing	200m- 1km decreasing suitability	
		>1 km lowest suitability (equal)	
		0% highest suitability	

Distance to	Linear	0 – 5km decreasing suitability	0.194
		>5km no suitability	
Distance to flood area	J-shape increasing	0-500m lowest suitability	0.194
		500-1km increasing suitability	
		>1km highest suitability (equal)	

After assigning factor weights for all the land use types, the suitability maps can be generated by using MCE-WLC. The suitability maps for each land use type are presented in Figure 3-10. The value ranges from 0 to 1 which indicates the lowest suitability to the highest suitability. It should be noted that water area is almost constant which is consistent with the transition matrix data as well as the observed land use map in 2006. The highest suitability for urban is still at the center of the city since it is the agglomeration of human activities. Open space basically involves parks and gardens, but the boundary between open space and forest is obscure since they may interconvert if the definitions of these terminology are vague or capricious. Forest and agriculture are essentially located in the Green Belt or nearby the Green Belt. Some of them are also adjacent to open space, corresponding to the suitability maps for open space as well.





**Figure 3-10.** Suitability maps for water, urban, open space, forest, and agriculture (0 indicates no suitability, and 1 indicates the highest suitability)

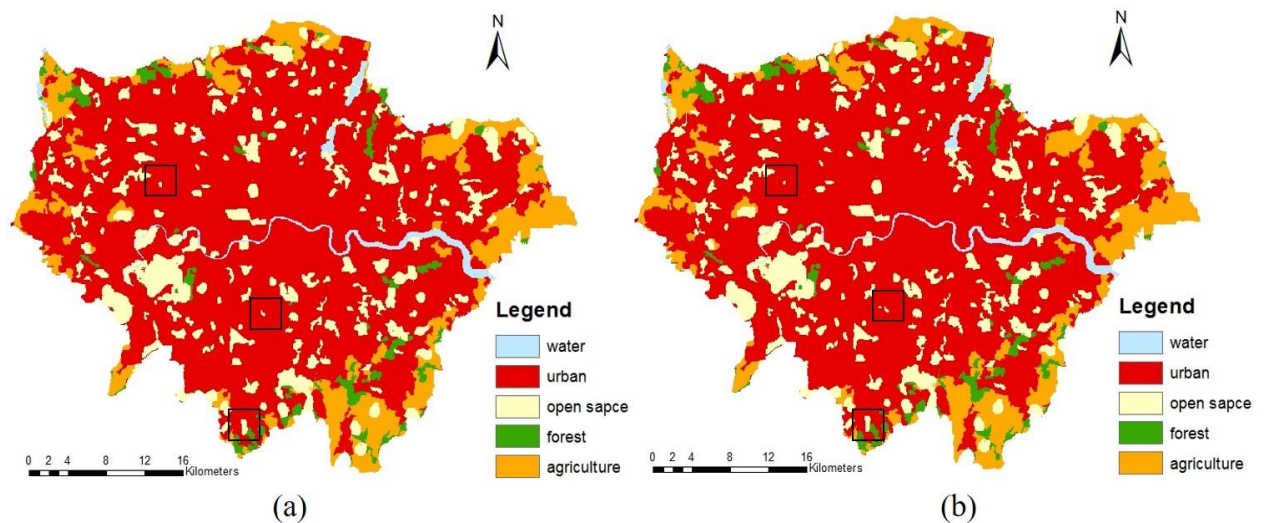
To measure the accuracy of these suitability maps for each land use class, AUC was used and summarized in Table 3-8. Since all the AUC values are well above 0.9, the model performance of the suitability maps present high accuracy (Swets, 1988). These high values of AUC can approve the reliability of the previous assumptions regarding to AHP and MCE-WLC. Thus, the model is ready for prediction LULC map in 2012, which would be used for validation purpose in the next step.

**Table 3-8.** AUC values for all land use classes

land use class	AUC
water	0.9988
urban	0.9299
open space	0.9471
forest	0.9677
agriculture	0.9799

### 3.5.5. TerrSet validation and prediction

By using the VALIDATE module in TerrSet, the location agreement between the observed and projected LULC layers can be evaluated. The overall Kappa statistics ( $K_{\text{standard}} = 0.8545$ ,  $K_{\text{no}} = 0.9123$ ,  $K_{\text{location}} = 0.8715$ ) for the validation result are well above 80 percent, indicating the good agreement between the observed and projected LULC layers. After implementing validation of the model, the model is ready for predicting land use changes in 2030 and 2050. From Figure 11, it is not easy to distinguish the land use change differences, but the land use area data can indicate the changes and differences from 2012 to 2050 comparatively (Table 3-9).



**Figure 3-11.** Predicted LULC maps in (a) 2030 and (b) 2050 with major difference in comparison highlighted by black squares

**Table 3-9.** Land use area comparison between 2012, 2030, and 2050

land use classes	observed	predicted		differenc e	differenc e	differenc e	differenc e
	2012	2030	2050	2012- 2030	2012- 2050	2012- 2030	2012- 2050
	unit: hectares					in %	
water	2750	2762	2770	12	20	0.44%	0.73%
urban	110943	112616	113073	1673	2130	1.51%	1.92%
open space	22526	24154	25204	1628	2678	7.23%	11.89%
forest	3215	3672	3489	457	274	14.21%	8.52%
agriculture	20162	16392	15060	-3770	-5102	-18.70%	-25.31%

### 3.6. Discussion

Based on the prediction results, water body will expand slightly from 2012 to 2030 and 2050. Urban area is expected to expand about 16.73 and 21.30 km<sup>2</sup> (in percentage of 1.51 and 1.92 compared to 2012) in 2030 and 2050 respectively. Meanwhile, open space is also expected to increase significantly which is 16.82 and 26.78 km<sup>2</sup> (in percentage of 7.23 and 11.89 compared to 2012) in 2030 and 2050 respectively. The increasing open space will reduce the flood impact as well as alleviate other climate change impact. The open space protection policies and the food impact driving factor play important role in protecting and supporting the expansion of open space. Although forest is also under the policy protection, the impact of climate change and urbanization still obstruct the development of forest. After the significant reduction of forest from 2000 to 2006, the policy changed to protect forest from being developed. In 2030, forest will increase slightly, but decrease later because of climate change. Agriculture is expected to decrease about 37.70 and 51.02 km<sup>2</sup> (in percentage of 18 and 25 compared to 2012) in 2030 and 2050 respectively. The reduction of agriculture area could be the result driven by policies and climate variability simultaneously, converting to the growth of urban and open space.

Comparing the scenarios of flood impact factor inclusion and exclusion for both urban and open space, is required to analyze the flood impact in terms of land use changes for long term period. The comparison between inclusion and exclusion of flood impact scenarios can be made possible side by side (Table 3-10). When flood impact to urban and open space is excluded, there will be more urban area and less open space. Other changes such as water and agriculture are very slight and negligible since flood impact was not considered as driving factor for these land use classes. However, the use of SuDS to mitigate the flood impact was not considered in this scenario.

Inclusion of flood impact factor leads to decreasing of urban area and increasing of open space. Therefore, it can be concluded that the closer to the flood impact zone, the higher the suitability to be converted to green space from its current land use class.

**Table 3-10.** The scenario comparison between the inclusion or exclusion of flood impact for land use change prediction in 2030 and 2050

land use classes	2030		2050	
	flood inclusion	flood exclusion	flood inclusion	flood exclusion
	unit: square kilometers			
water	27.62	27.75	27.70	27.82
urban	1126.16	1126.25	1130.73	1131.21
open space	241.54	241.40	252.04	251.91
forest	36.72	36.72	34.89	34.89
agriculture	163.92	163.84	150.60	150.13

The differences between the inclusion and exclusion of flood impact are not significant, which may be results by the small weight assigned to the flood impact. Since the socio-economic and policy-oriented driving factors are more important and contribute more to urban growth, the

associated weights assigned to these factors are higher than the flood impact factor. However, if the situation of climate change keeps deteriorating in the future resulting in more rainfall and higher flood risk, the weight of flood impact may increase leading to more dramatic land use changes when flood impact is included. This has to be paid attention by all urban planners and decision makers regarding to the impact of climate change and flood risk.

### 3.7. Conclusion

London has been suffering from flood impact since 1700s which threatens the development of the city in the future, especially under the projected climate variability. There is an urgent need to simulate and predict land use changes in terms of flood impact as well as other human activities with varying constraints and factors. CA-based MC model is used in this study to simulate and predict land use changes for 2030 and 2050 in London, collaborating with the rainfall prediction via SDSM in order to provide statistical foundation for future potential flooding area. Results of SDSM show that annual rainfall will increase over London for the year 2030 however will decrease in 2050 overall according to the UK-based GCM, HADCM3 A2 scenario.

The results of CA-MC model indicate that urban and open space are expected to increase more than 16 and 20 km<sup>2</sup> (in percentage of 1.51 and 1.92 compared to 2012) in 2030 and 2050 respectively in London. Forest will slightly increase in 2030, but decrease after that due to climate change. Agriculture are expected to decrease significantly due to urbanization and climate change. The total amount of water body will slightly increase resulting by more rainfall is expected in 2030. Two scenarios comparison between inclusion and exclusion of flood impact factor was developed to measure the climate change and flood impact in London. Inclusion of potential flood

impact from future rainfall results in higher suitability of the development of open space, but less suitability for urban area. More support of developing open space from policy and planning views should be addressed to respond to the rigorous climate change and flood impact issues in the future.

### 3.8. References

Adamatzky, A.I., 1994. Identification of cellular automata. CRC Press.

Adhikari, S., Southworth, J., 2012. Simulating forest cover changes of Bannerghatta National Park based on a CA-Markov model: a remote sensing approach. Remote Sensing 4, 3215-3243.

Andersson, C., Frenken, K., Hellervik, A., 2006. A complex network approach to urban growth. Environment and Planning A 38, 1941-1964.

Arsanjani, J.J., Helbich, M., Kainz, W., Boloorani, A.D., 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. International Journal of Applied Earth Observation and Geoinformation 21, 265-275.

Balster, H., Braun, P.W., Köhler, W., 1998. Cellular automata models for vegetation dynamics. Ecological modelling 107, 113-125.

Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. science 286, 509-512.

Batty, M., 2007. Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals. The MIT press.

Beck, J.R., Shultz, E.K., 1986. The use of relative operating characteristic (ROC) curves in test performance evaluation. Archives of pathology & laboratory medicine 110, 13-20.

Clark, G.L., Gertler, M.S., Feldman, M.P., 2003. The Oxford handbook of economic geography.

Oxford University Press.

Clarke, K.C., Hoppen, S., 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, 247-261.

Costanza, R., d'Arge, R., De Groot, R., Faber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., 1997. The value of the world's ecosystem services and natural capital.

Dwyer, J.F., McPherson, E.G., Schroeder, H.W., Rowntree, R.A., 1992. Assessing the benefits and costs of the urban forest. *Journal of Arboriculture* 18, 227-227.

Eastman, J., 2015. *TerrSet manual*. Clark Labs. Clark University. Worcester, MA, USA.

Gant, R.L., Robinson, G.M., Fazal, S., 2011. Land-use change in the 'edgelands': Policies and pressures in London's rural-urban fringe. *Land Use Policy* 28, 266-279.

Greater London Authority, 2002. *Flooding in London: A London Assembly Scrutiny Report*. London, UK.

Greater London Authority, 2015. *London Sustainable Drainage Action Plan*. Greater London Authority.

Greiner, M., Pfeiffer, D., Smith, R., 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventive veterinary medicine* 45, 23-41.

Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., Hokao, K., 2011. Modeling urban land use change by the integration of cellular automaton and Markov model. *Ecological Modelling* 222, 3761-3772.

Hopfield, J.J., 1988. Artificial neural networks. *IEEE Circuits and Devices Magazine* 4, 3-10.



- Houet, T., Hubert-Moy, L., 2006. Modelling and projecting land-use and land-cover changes with a cellular automaton in considering landscape trajectories: an improvement for simulation of plausible future states. *EARSeL eProceedings* 5, 63-76.
- Hu, Z., Lo, C., 2007. Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems* 31, 667-688.
- Jiménez-Valverde, A., 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography* 21, 498-507.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society* 77, 437-471.
- Kamusoko, C., Aniya, M., Adi, B., Manjoro, M., 2009. Rural sustainability under threat in Zimbabwe—simulation of future land use/cover changes in the Bindura district based on the Markov-cellular automata model. *Applied Geography* 29, 435-447.
- Litman, T., 1995. Land use impact costs of transportation. *World Transport Policy and Practice* 1, 9-16.
- López, G.G.I., Hermanns, H., Katoen, J.-P., 2001. Beyond memoryless distributions: Model checking semi-Markov chains, *Process Algebra and Probabilistic Methods. Performance Modelling and Verification*. Springer, pp. 57-70.
- Mayor of London, 2009. Open space strategies: Best practice guidance. Retrieved from <https://www.designcouncil.org.uk/sites/default/files/asset/document/open-space-strategies.pdf>.
- Memarian, H., Balasundram, S.K., Talib, J.B., Sung, C.T.B., Sood, A.M., Abbaspour, K., 2012.

- Validation of CA-Markov for simulation of land use and cover change in the Langat Basin, Malaysia.
- Moghadam, H.S., 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.
- Nguyen, V., Nguyen, T., Gachon, P., Lee, J., Lam, K., 2005. An evaluation of statistical downscaling methods for simulating daily precipitation and extreme temperature series at a local site, *Environmental hydraulics and sustainable water management*. Volume 1, environmental hydraulics: Volume 2, sustainable water management in the Asia-Pacific region. *Proceedings of the 4th International Symposium on Environmental Hydraulics and the 14th Congress of Asia and Pacific Division, International Association of Hydraulic Engineering and Research, Hong Kong, 15-18 December 2004*. Taylor & Francis, pp. 1911-1916.
- OPDC, 2017. Old oak and park royal development corporation. Mayor of London, retrieved from <https://www.london.gov.uk/about-us/organisations-we-work/old-oak-and-park-royal-development-corporation-opdc>.
- Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R., 1996. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology* 49, 1373-1379.
- Rotenberg, R., 2008. *The European City and Green Space: London, Stockholm, Helsinki and St. Petersburg, 1850–2000*—Edited by Peter Clark. *International Journal of Urban and Regional Research* 32, 523-525.
- Subedi, P., Subedi, K., Thapa, B., 2013. Application of a hybrid cellular automaton–Markov

- (CA-Markov) Model in land-use change prediction: a case study of saddle creek drainage Basin, Florida. *Applied Ecology and Environmental Sciences* 1, 126-132.
- Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science* 240, 1285-1293.
- Takeyama, M., Couclelis, H., 1997. Map dynamics: integrating cellular automata and GIS through Geo-Algebra. *International Journal of Geographical Information Science* 11, 73-91.
- Torrens, P.M., 2000. How cellular models of urban systems work (1. Theory).
- Urban Task Force, 2002. Towards an urban renaissance: final report of the urban task force chaired by Lord Rogers of Riverside, The Department of the Environment. Transport, and Regions (DoE), London.
- Von Neumann, J., Burks, A.W., 1966. Theory of self-reproducing automata. *IEEE Transactions on Neural Networks* 5, 3-14.
- Wilby, R.L., Dawson, C.W., Barrow, E.M., 2002. SDSM—a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software* 17, 145-157.
- Xie, Y., 1996. A generalized model for cellular urban dynamics. *Geographical Analysis* 28, 350-373.
- Yang, T., Li, H., Wang, W., Xu, C.Y., Yu, Z., 2012. Statistical downscaling of extreme daily precipitation, evaporation, and temperature and construction of future scenarios. *Hydrological Processes* 26, 3510-3523.
- Yegnanarayana, B., 2009. Artificial neural networks. PHI Learning Pvt. Ltd.

## **CHAPTER 4 PREDICTING LONG-TERM LAND USE CHANGES IN BEIJING (CHINA) WITH CONSIDERATION OF CLIMATE CHANGE, SOCIO-ECONOMIC DEVELOPMENT, AND POLICY-ORIENTED FACTORS**

### **4.1. Introduction**

Changes of land-use and land-cover (LULC) are part of the surface earth processes driven by the interaction among natural and socio-economic driving factors and constraints at different scales. Yet, the dynamic and complexity of urban growth makes us difficult to distinguish and separate the function and impact of each driving factor or constraint especially for multiscale urban growth simulations facing heterogeneous environments in megacities. In fact, LULC changes should reflect not only the accelerating speed of urbanization facing globalization, population growth/migration and economic development but also regional planning, such as the corridor effect, and climate change. Comprehensive understanding of the urban growth pattern and characteristics of each driving force (i.e., factors and constraints) locally and regionally is important for simulation and prediction in LULC changes (Liu, 2013).

As one of the megacities with the highest population density and an unprecedented pace of urbanization, Beijing in China has demonstrated its potential to keep expansion in the next few decades. There were a suite of studies related to LULC changes in Beijing area focusing on typical socio-economic driving factors such as population density, gross domestic product (GDP), distance to urban center, distance to road and railway, and so on. These previous studies indicate the plausible urban growth trend with different scenarios of consideration (Table 4-1). They attempted to simulate and predict LULC changes around the urban area with the consideration of

most popular socio-economic driving factors under different scenario assumptions. However, few combined the traditional urban growth driving factors with the concerns of the potential impact of the possible climate change and the new corridor effects together for the future.

**Table 4-1.** Summary of the previous studies related to the simulation and prediction of LULC change in Beijing area

Num ber	Study area	Model	Simula- tion year	Predic- tion year	Result	Source
1	Beijing	binary auto- logistic model	2000	2010	Urban development took place in east, south and northeast part of Beijing in the last decade. Mostly in areas beyond the fourth ring and the majority was between the fifth ring and sixth ring.	Deng and Srinivasan, 2016
2	Beijing	CLUE- s and Markov -based CA model	1985, 2000, 2010	2020	The geographical environment limits the expansion of urban built-up land, but the conversion of cultivated land to built-up land in mountainous areas will be more prevalent by 2020	Han et al., 2015

Num ber	Study area	Model	Simula- tion year	Predic- tion year	Result	Source
3	Beijing	CLUE- s CA model	1995, 2000	2015	The key location susceptible to land-use changes were found to be located at the central urban Beijing and the surrounding regions inclding Yanqing County, Changping District and Fangshan District	Hu et al, 2013
4	Beijing	urban expansi on dynami c (UED) CA model	1991, 1997, 2000, 2004	2015	The dominant pattern of urban development was the limitless expansion to form a "big pancake" by encroaching cultivated land and green space in the urban-rural fringe.	He et al., 2008
5	Beijing	urban expansi on scenari o (UES) CA model	1991 - 2004	2004, 2020	A dilemma of urban expansion versus limited water resource and environment deterioration exists	He et al., 2006

Num ber	Study area	Model	Simula- tion year	Predic- tion year	Result	Source
6	Beijing	Markov chains and regressi on analyse s model	1986- 2001	2021	Most of the urban growth and loss of agriculture land occurred in inner and outer suburbs	Wu et al., 2006
7	Haidian district, Beijing	CLUE- s CA model	1991- 2001		The neighborhood factor analyse is important to simulate urban land-use changes	Duan et al., 2004
8	Beijing	remote sensing	1975, 1984, 1991, 1997		The large-scale land use/cover change took place in Beijing area in the process of rapid urbanization from 1975 to 1997 with major characteristics of urban expansion at the expanse of encroaching cultivated land in the plain area and of obvious structural readjustment occurring in other land use/cover types	He et al., 2001

Different than other megacities in developed or developing countries, the major driving factors impacting LULC changes in China are more linked to national policies and regulations. For instance, Shenzhen once was a small village located in the south coastal region of China. After being assigned as a special economic free trade zone in the early 1980's by the central government, Shenzhen has developed rapidly as a modern metropolis that bridges Hong Kong and the mainland China, geographically and economically. The GDP of Shenzhen ranked top 5 in Chinese cities in 2013 (China Knowledge, 2014) with population of 10.8 million in 2016 (World Population Review, 2016). The Chinese central government is keen to dominate the economic development and population growth with varying policies that are more important to govern the land use management and affect LULC changes. In April 2017, a new economic free trade zone, named Xiong'an new zone, located in Hebei province of China (i.e., about 100 km southwest of Beijing), was established. Xiong'an is considered as another Shenzhen in north China, which is expected to achieve rapid economic growth in the next decades. China aims to build Xiong'an as a new corridor city between Beijing and the southeast coast of China geographically and economically forming a golden triangle among Shenzhen, Shanghai-Pudong, and Xiong'an itself although another corridor city of Tianjin has provided an economic driving force in concert with Beijing's economic booming since the 1980s. However, the *effects* of urbanization and *climate change* are converging in dangerous ways triggering more flooding impact that requires further attention.

Cellular automata (CA) model is one of the most popular models applied to simulate and predict the LULC changes (Batty, 2007; Arsanjani et al., 2013). The CA model has the ability to cooperate with other approaches for temporal analyses such as the Markov Chain (MC) model (Wu et al., 2006), the Conversion of land Use and its Effects at Small regional extent (CLUE-S)



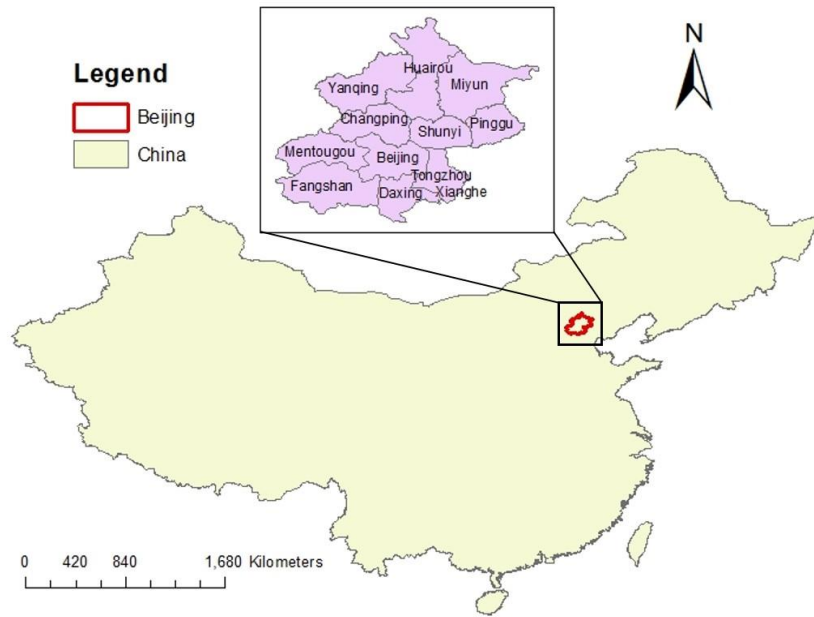
model (Duan et al., 2004), and so on. Among these models, MC is a mathematical approach that can calculate the probability of each step of LULC change from one state to the other within a given horizon (Guan et al., 2011). The temporal characteristic of a MC model can be in concert with the spatial nature of a CA model to collectively simulate and predict spatiotemporal interactions during the process of LULC changes, in which the multi-criteria evaluation – weighted linear combination (MCE-WLC) can make a step wise decision of the change of LULC. In addition, the coupled CA and MC model can work with fuzzy membership functions to reflect the plausible uncertainties embedded in such step wise decision analyses.

In this study, an integrated model of CA, MC, and MCE-WLC was developed and applied to simulate and predict LULC changes for 2030 and 2050 in Beijing. In order to consider the climate change and flooding impact, a statistical downscaling model (SDSM) was used to predict future rainfall and flood potential, providing the fundamental climate change background (Wilby et al., 2002). In order to analyze the impact of the new economic free trade zone in Xiong'an and climate change simultaneously, four different scenarios were included for exploring the differential effect of these concerns systematically. Scenario 1 is designed to address the traditional pathway considering only prescribed socio-economic driving factors. Scenario 2 considers the impact of the new economic free trade zone in Xiong'an and the new corridor with Tianjin. Scenario 3 simulates the recent policy that the central government will be relocated to Tongzhou district at southeast suburban area of Beijing to reduce the overdevelopment pressure at downtown of Beijing. Hence, Scenarios 2 and 3 are simply expanded cases with a new initiative on the top of Scenario 1, respectively. Scenario 4 examines the potential impact of climate change and flood risk, adding more insight on the previous three scenarios.

The science questions to be answered in this study include: (1) What will be the possible LULC change of Beijing in 2030 and 2050 in association with the four prescribed scenarios? (2) How will the government policies in regard to the new economic free trade zone in Xiong'an and government relocation to Tongzhou district affect the LULC changes of Beijing in 2030 and 2050? and; (3) Will climate change and flooding potential alter the pattern LULC changes of Beijing in 2030 and 2050 significantly?

#### 4.2. Study Area

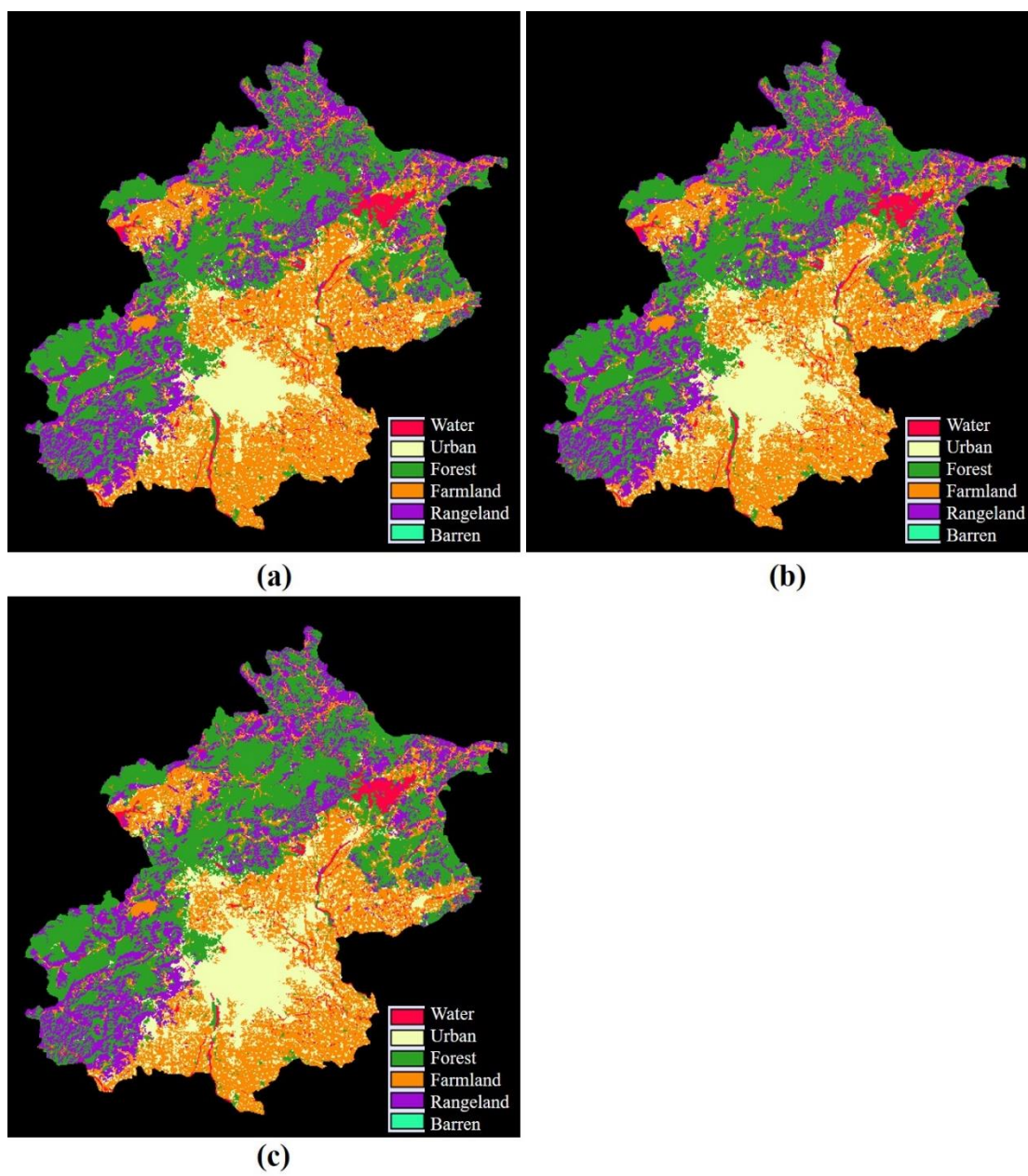
Beijing is the capital city of China and is one of the fastest-growing megacities in the world covering total area of 16,410 square kilometers. There are 16 districts including the central part of the city (Xicheng and Dongcheng), the suburbs (Haidian, Chaoyang, Fengtai, and Shijingshan), the peri-urban (Tongzhou, Shunyi, Fangshan, Daxing, Changping, Mentougou, Huairou, Pinggu, Minyu, and Yanqing). The city of Beijing in Figure 4-1 includes the urban area with following six districts Xicheng, Dongcheng, Haidian, Chaoyang, Fengtai, and Shijingshan. Since China started the economic reform in 1979 and became a member of the World Trade Organization (WTO) in 2001, Beijing has experienced dramatic economic growth and rapid urbanization. After the successful hosting the Olympic Games in 2008, the local infrastructure and transportation systems were prompted to a new level (Owen, 2005). The total population exceeded 21 million as of 2015 (Beijing Economic and Social Development Statistical Bulletin, 2015) and the annual growth rate of GDP has kept higher than 6.9% since the 1990s (World Bank, 2015).



**Figure 4-1.** Geographical location of Beijing in China

#### 4.3. Applied Software and Data Collection

SDSM was applied in this study to simulate and predict rainfall and flooding potential in Beijing for the predetermined target years. In order to predict LULC changes in 2030 and 2050, ArcGIS and TerrSet were applied to simulate the urban growth processes. The collected data for simulating the LULC changes includes three-year's LULC data (2000, 2005, and 2010) (Figure 4-2), road system, railway system, digital elevation model (DEM), population census, GDP, and flood records (Table 4-2). All the data needs to be pre-processed to 30×30 m raster data prepared for simulation in TerrSet to identify possible LULC changes. Whereas the DEM data were collected from the United States Geological Survey (USGS), the flood data were collected from the National Aeronautics and Space Administration (NASA).



**Figure 4-2.** Collected LULC data (a) 2000, (b) 2005, and (c) 2010

**Table 4-2.** Collected data for LULC changes prediction

Data	Description	Resolution	Source
LULC	2000, 2005, 2010	30x30 m	resdc.cn
Road			Digital Chart of the World
Railway			Beijing City Lab
DEM			USGS
Population			resdc.cn
GDP			resdc.cn
Flood records	2013-2017		NASA

#### 4.4. Methodology

##### 4.4.1. Statistical Downscaling Model (SDSM)

###### *4.4.1.1. Beijing Rainfall and Flood Potential*

The climate of Beijing is characterized by heavy rainfall during the summer months (i.e., wet season) and low to moderate rainfall during the winter to spring months (i.e., dry season). Within several decades, Beijing has seen rapid urban growth, and subsequent changes in

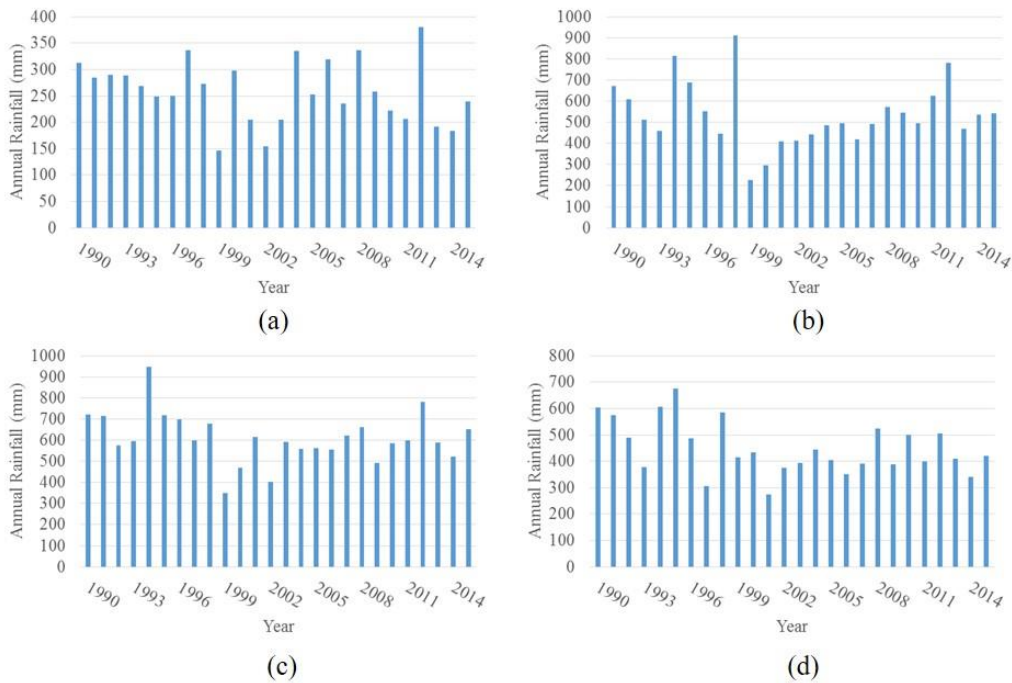
atmospheric circulation patterns and rainfall variations have been apparent (Li et al. 2005; Miao et al., 2011). With greater imperviousness due to rapid urbanization, more urban runoff is anticipated. This could present challenges to drainage systems that are more vulnerable with age or incapacities to handle increasing urban runoff from the increases in surrounding urban space. In addition, increased urban space reduces evaporation and particularly when linking urban space and rainfall patterns in Beijing (Zhang et al., 2009). Zhang et al. (2009) found that urbanization resulted in weaker summer convective rainfall patterns; however, other studies showed that the opposite was true with urbanization processes resulted in increased summer rainfall over and downwind of major cities. It will be helpful to assess the trends of existing rainfall patterns and make projections using SDSM models for 2030 and 2050.

#### *4.4.1.2. Statistical Downscaling of Beijing Rainfall*

In order to determine the potential impact of future rainfall within Beijing city limits for the year 2030 and 2050, the SDSM is potentially useful (Wilby et al., 2002). SDSM is designed by a statistical downscaling approach which applies statistical relationships, based on multiple linear regression techniques, between large-scale climate, such as humidity and temperature from global climate models (GCMs) and local climate such as daily rain gauge records. Daily records are required for SDSM and daily rainfall gauge data were sought for the purposes of this study. Due to the nature of Beijing as a megacity, this study looked at several stations across the Beijing city limits in 4 key directions: North, South, East and West (Table 4-3) with respect to center of Beijing. Analyzing rainfall data across various locations can assist in determining the relative differences in rainfall different areas of Beijing could face, as apparent in Figure 4-3.

**Table 4-3.** Locations of Daily Beijing Rainfall Data

Relative Location of Station w.r.t center of Beijing	Latitude (N)	Longitude (E)
North	40.09	116.39
South	39.56	116.37
East	40.09	117.06
West	39.58	115.41



**Figure 4-3.** Annual rainfall trends for period of 1990-2015 for (a) North station, (b) South station, (c) East station and (d) West station

#### *4.4.1.3. DSM Climate Predictor Variables*

Climate predictors are important factors in SDSM analysis since a particular set or combination of climate predictors can impact calibration, validation, and ultimately future climate projections. The best combination of predictors is an iterative process (Wilby et al, 2002; Hessami et al., 2008; Teutschbein et al., 2011) which is a process that removes the least significant predictors until statistically significant predictors are remaining. SDSM relies on both observed-derived and GCM-derived atmospheric predictor variables. For this study, GCM-derived atmospheric predictors are used for both calibration, validation, and future projection.

GCM-derived atmospheric predictor variables were obtained from the Hadley Centre Coupled Model, version 3 (HADCM3) GCM B2 scenario of the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC, 2007) and were obtained from the Government of Canada website which provides required SDSM predictors. The B2 scenario describes on a global level, local and regional solutions to economic, social, and environmental sustainability with continued population growth, economic development, and a moderate increasing trend in greenhouse gas emissions (IPCC, 2007). All atmospheric predictor variables were re-gridded to a standard coordinate system ( $2.5^{\circ}$  latitude  $\times$   $3.75^{\circ}$  longitude) used in HADCM3 over a period from 1961 to 2099 and normalized with respect to base period (1961 to 1990) averages. The grid centered at  $40.0^{\circ}$  N latitude and  $116.25^{\circ}$  E longitude was chosen for extracting climate predictors for the study area. The predictor variables of concern include mean sea level pressure, 500hPa airflow strength, 500hPa zonal velocity, 500 hPa surface vorticity, 500hPa meridional velocity, 500 hPa geopotential height, 500 hPa wind direction, 500 hPa divergence, relative surface



humidity, surface specific humidity and mean surface temperature. Table 4-4 provides a summary of the climate variables used for each station.

**Table 4-4.** GCM climate predictors used for each station

<b>Predictors</b>	<b>North</b>	<b>South</b>	<b>East</b>	<b>West</b>
Mean sea level pressure	Yes	Yes	Yes	Yes
500hPa airflow strength	Yes	Yes	Yes	Yes
500hPa zonal velocity	Yes	Yes	Yes	Yes
500 hPa surface vorticity	Yes	Yes	Yes	Yes
500hPa meridional velocity	Yes	Yes	Yes	Yes
500 hPa geopotential height	Yes	Yes	Yes	Yes
500 hPa wind direction	No	No	Yes	Yes

<b>Predictors</b>	<b>North</b>	<b>South</b>	<b>East</b>	<b>West</b>
500 hPa divergence	No	No	Yes	Yes
Relative surface humidity	Yes	No	No	Yes
Surface specific humidity	No	No	Yes	Yes
Mean surface temperature	Yes	Yes	Yes	Yes

Note: hPa = hecto-Pascals (denoting pressure)

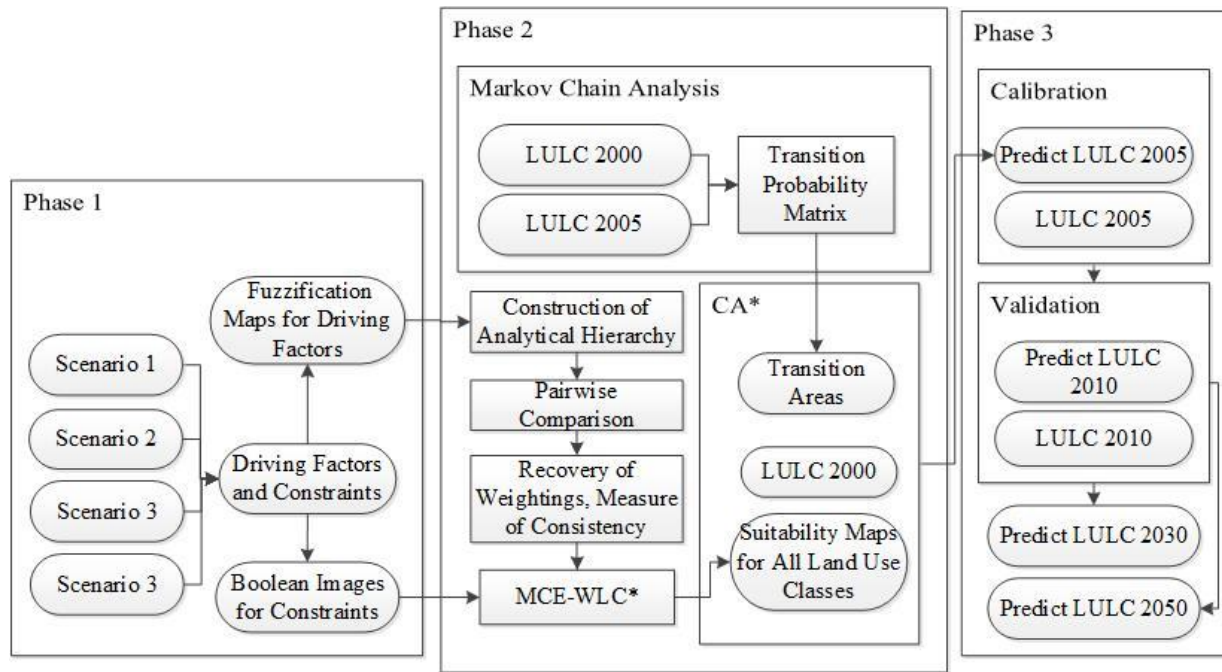
#### *4.4.1.4. SDSM Calibration, Validation, and Future Projection*

A ten-year period spanning from January 1990 until January 2000 are used for calibrating the SDSM model across all four stations, while the period of January 2005 until January 2010 was used to validate the SDSM model across all four stations. Future projections of rainfall were conducted using the weather generator in SDSM. The weather generator was used to produce ensembles of future daily weather series for future climate using the calibrated SDSM model and GCM-derived atmospheric predictor variables obtained from the HADCM3 GCM B2 scenario. Annual rainfall projections from a 20-member SDSM ensemble were considered to provide a range of possible projections of annual mean rainfall values for years 2030 and 2050. The range of possible projections of annual rainfall values from the 20-member ensemble for 2030 and 2050

are compared to each year of the observed historical record (1990-2015) to retrieve and assess the historical vs. future trend.

#### 4.4.2. CA-MC Model – LULC Change Prediction

This study conducted an integrated model analysis consisting of CA and MC to address the spatiotemporal relationship in urban growth with a decision-making module (MCE-WLC). It is designed to simulate and predict LULC changes in 2030 and 2050 in Beijing under uncertainty depicted by using fuzzy membership functions. The flowchart of the integrated modeling analysis is summarized in Figure 4-4. There are three phases in the integrated modeling analysis and they include: 1) determination of driving factors and constraints for four scenarios in Phase 1, 2) production of suitability maps for all land use types and performing MC analysis in Phase 2, and 3) calibration, validation, and prediction LULC changes in 2030 and 2050 in Phase 3. They are described below.

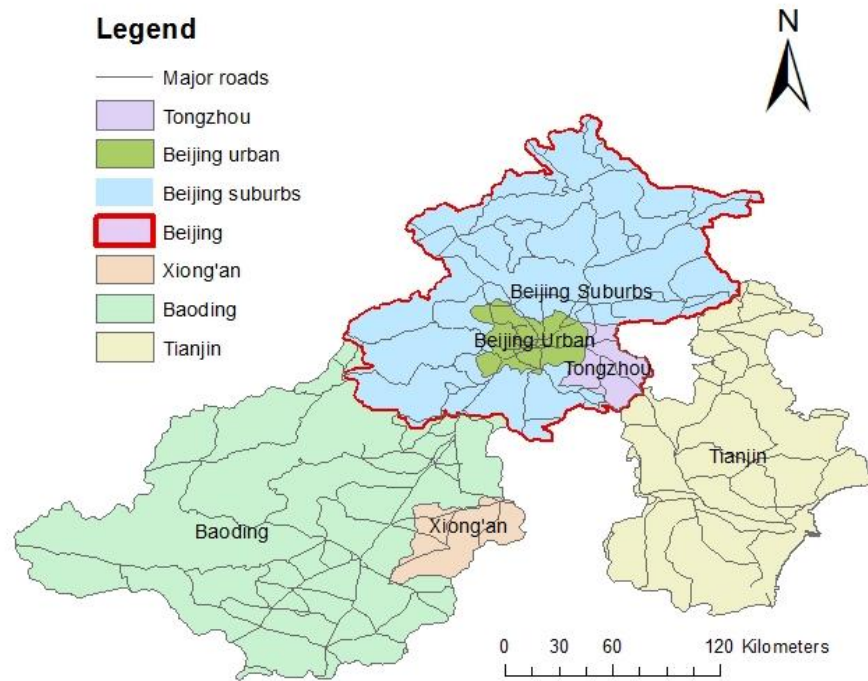


**Figure 4-4.** Flowchart of the integrated urban growth model; \*CA: cellular automata; MCE-WLC: multi-criteria evaluation-weighted linear combination.

#### 4.4.2.1. Phase 1: Factors and Constraints

The LULC changes depend on the land transfer rules related to driving factors and constraints that push and shape the changes of land use collectively. The starting point of this integrated model analysis is to select driving factors and constraints for Beijing based on the type of urban growth and the socio-economic background. The corridor city of Tianjin has provided economic driving force and natural resources for Beijing since the 1980s. The economic booming of Beijing has close tie with the corridor effect from Tianjin. After Xiong'an was announced as a new economic free trade zone in the southwest of Beijing, the newly formed corridor among Beijing, Tianjin, and Xiong'an is expected to become a new economic development hub and promote the next economic booming regionally.

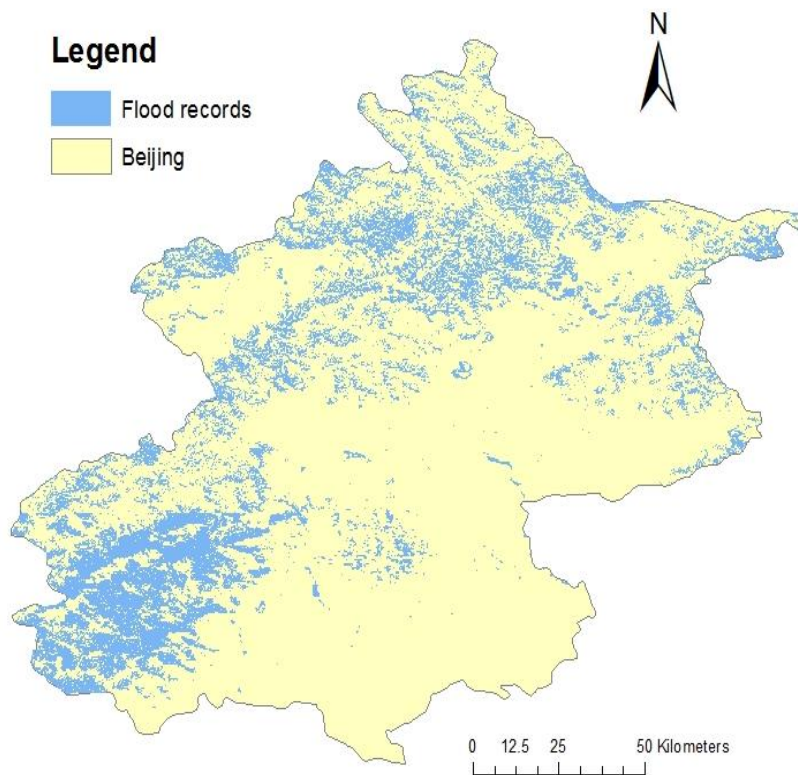
In addition, the proposal suggesting the central government relocation from the central part of the city to the peri-urban Tongzhou district to avoid the overdevelopment issue. The peri-urban Tongzhou district has more potential and space to be developed. Thus, Tongzhou district is expected to become a new political and business center of Beijing and promote economic development around this area. The geographical locations of Beijing, Tongzhou district, Xiong'an, and Tianjin as well as the road system for connection are shown in Figure 4-5.



**Figure 4-5.** Geographical locations of Beijing, Tongzhou district, Tianjin, and Xiong'an connected by the road systems

Based on the socio-economic and political background of Beijing, the applied driving factors for this study can be identified. The driving factors include distance to road, distance to urban center, distance to network cities (Tianjin), distance to new special economic zone Xiong'an,

distance to proposed new political center Tongzhou district, distance to slope, distance to population density, distance to GDP, and distance to flood records. The map of flood records is presented in Figure 4-6 (NASA, 2017). Specially, different driving factors may result in differing LULC changes for each land use class since the corresponding weight and associated driving factors may vary. In addition, constraints shape the area that are not suitable to be developed, such as actual roads and rivers. All these driving factors and constraints play important roles impacting the LULC changes and they are summarized for four scenarios in Table 4-5.



**Figure 4-6.** Flood records location in Beijing from 2013 to 2017 (source: NASA, 2017)

**Table 4-5.** Applied driving factors and constraints

	Description
Driving factors	Distance to road
	Distance to urban center
	Distance to urban
	Distance to Xiong'an
	Distance to Tongzhou district
	Distance to Tianjin
	Distance to slope
	Distance to population density
	Distance to GDP
	Distance to flood records
Constraints	Actual roads
	Rivers

All the driving factors need to be pre-processed to fit the fuzzification maps by using fuzzy membership functions in order to determine the urban growth potential in the decision-making

process (i.e., the multicriteria evaluation process). Fuzzy membership function was used to reflect the degrees of special uncertainty as continuous values from 0 to 1. Distance to each factor was calculated which generated distance images, followed by the creation of a fuzzy membership function to each factor. There are three types of fuzzy membership functions selected in this study and they are liner, sigmoidal, and J-shaped. In addition, control points were identified to determine the maximum and minimum distance from the previous distance images. On the other hand, constraints were pre-processed as Boolean functions containing values of either 0 or 1 for dichotomous choices. Value 0 indicates that the cell does not meet the required condition. In contrast, value 1 indicates that the cell meets the development requirement and is allowed to experience change of its land use class. Actual roads and rivers are considered as non-suitability to any possible development (i.e., changes of land use class) in this study.

#### *4.4.2.2. Phase 2: MCE, MC and CA*

MCE was applied to simulate the land development potential based on the trade-off of multiple development factors (Wu, 1998). MCE helps consider the preference of land development in the model by an analytical hierarchy process (AHP) to obtain behavior-oriented transition rules. According to Figure 4, AHP was applied in Phase 2 to measure the pairwise comparison and determine the weights for each factor to be used in MCE. The weight of each factor is assigned based on the impact on land use types as well as the interactive relationship between each other, with respect to some prior planning knowledge and experience gained from the interview survey. The consistency ratio (CR) was applied to calculate the reliability of the weights to measure whether the assignment for the weights is consistent throughout the interview survey. The closer



to the value 0, the better the weights that the CR may prove. If the CR is higher than 0.1, such an AHP needs to be re-processed and new weights are to be re-assigned.

Because weighted linear combination (WLC) can allow possible trade-offs with each other among prescribed criteria and has more flexibility than other two functions (Boolean intersection and ordered weighted average) in the MCE process (Eastman, 2015), MCE-WLC was selected to generate the suitability map for each land use type. The weight associated with each factor for each land use type was fed into the formulated MCE-WLC, collaborating with these Boolean images associated with constraints to generate such a suitability map for each land use type. Suitability maps consisting of continuous value from 0 to 255 indicate the suitability of a certain type of land use change at each cell in the CA model after the spatiotemporal trade-offs via the MCE-WLC scheme. Thus, a group of suitability maps were finally produced after applying MCE-WLC for each land use type.

MC that is the key model to support temporal trade-offs in the context of MCE-WLC is a mathematical approach that can calculate the probability over alternative changes of prescribed LULC types between two consecutive time periods spatially (Muller and Middleton, 1994). It has a close relationship to dynamic distribution consisting of two primary components: a transition probability matrix and a transition area matrix. These two matrices indicate the probability and area of each LULC type changing to other LULC types. The mathematic formation of a transition probability matrix can be expressed as follows:

$$(V_i \times P_{ij}) = (V_1, V_2, V_3 \dots V_n) \times \begin{pmatrix} P_{11}, P_{12}, P_{13} \dots P_{1m} \\ P_{21}, P_{22}, P_{23} \dots P_{2m} \\ P_{31}, P_{32}, P_{33} \dots P_{3m} \end{pmatrix} \quad (4-1)$$

where  $i$  is the LULC type in the first time period;  $j$  is the LULC type in the second time period;  $V_i$  is the fraction of a LULC type in the first time period;  $P_{ij}$  is the probability of LULC type  $i$  shifting to LULC type  $j$ ;  $V_i \times P_{ij}$  is the fraction of LULC type in the second time period;  $P_{11}$  is the probability that LULC type 1 will change into LULC type 1 from the first time period to the second time period;  $P_{12}$  = the probability that LULC type 1 will change into LULC type 2 from the first time period to the second time period and so on;  $m$  is the number of the LULC types in the study area.

CA is the one of the most popular urban growth models to simulate dynamic LULC changes spatially (White and Engelen, 1993; Batty, 2007). As mentioned above, the CA-based MC model has the unique ability to quantify the temporal and spatial LULC variations simultaneously (Subedi et al., 2013). Transition rules of each factor are defined by such an integrated approach and reflected by the variations in those predicted results.

#### 4.4.2.3. Phase 3: Calibration, Validation, and Prediction

Model calibration was performed by comparing the observed and predicted LULC maps in 2005, and validation was conducted by comparing the observed and predicted LULC maps in 2010. A set of indexes related to Kappa statistics was applied to measure the accuracy between the observed and predicted LULC maps in all the predetermined scenarios. After the completion of calibration and validation, the model is ready to predict LULC changes in 2030 and 2050.

## 4.5. Results

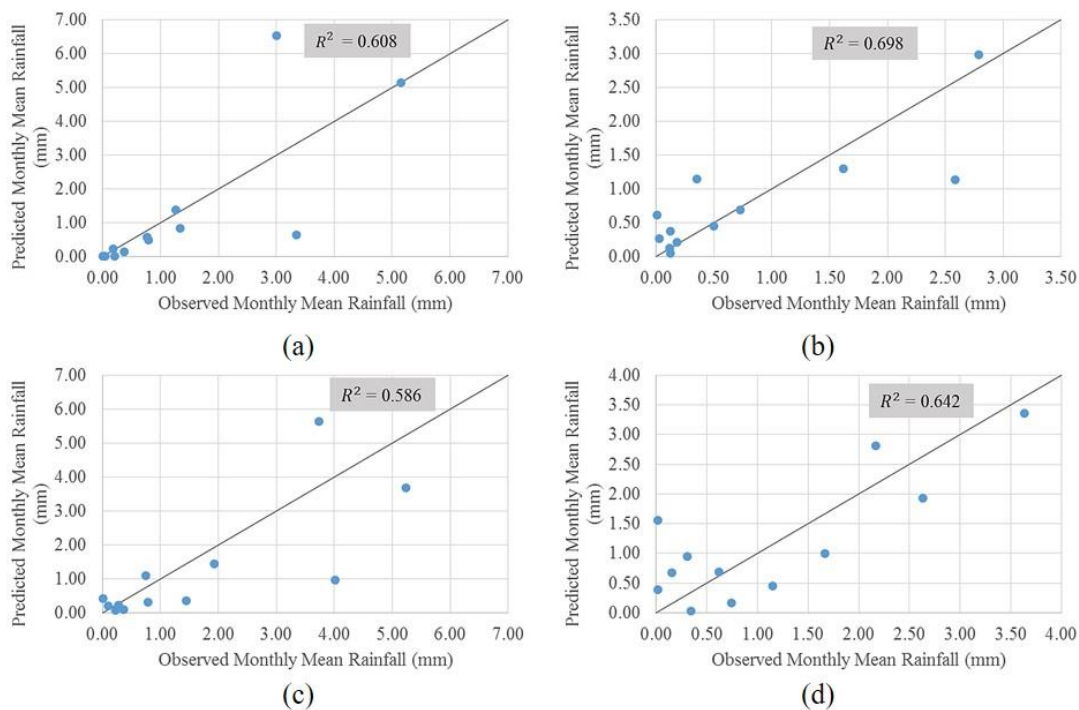
### 4.5.1. SDSM Calibration, Validation, and Prediction

Results of SDSM calibration for all rainfall stations are provided in Table 4-6 based upon monthly analysis for the period 1990-2000. Reflective of literature, such as Wilby et al. (2002) and Gulacha and Mulungu (2016),  $R^2$  values for calibration can be particularly low for rainfall for particular regions and time periods.  $R^2$  values for rainfall across all stations were low during spring and summer months in particular. This is due to the highly variability of daily rainfall. Calibration values for the winter months are particularly good except for the month of January for the North station and the month of February for the West station. For the fall to winter months (October, November and December), the North station calibration values are overall better than the other stations. Validation for each station performed relatively well except for the East station (Figure 4-7). Variance inflation and bias correction were applied for all stations to obtain best-case validation. Overall,  $R^2$  values were within the range of near 0.60 to near 0.70. The south station provided the best  $R^2$  value from all the stations.

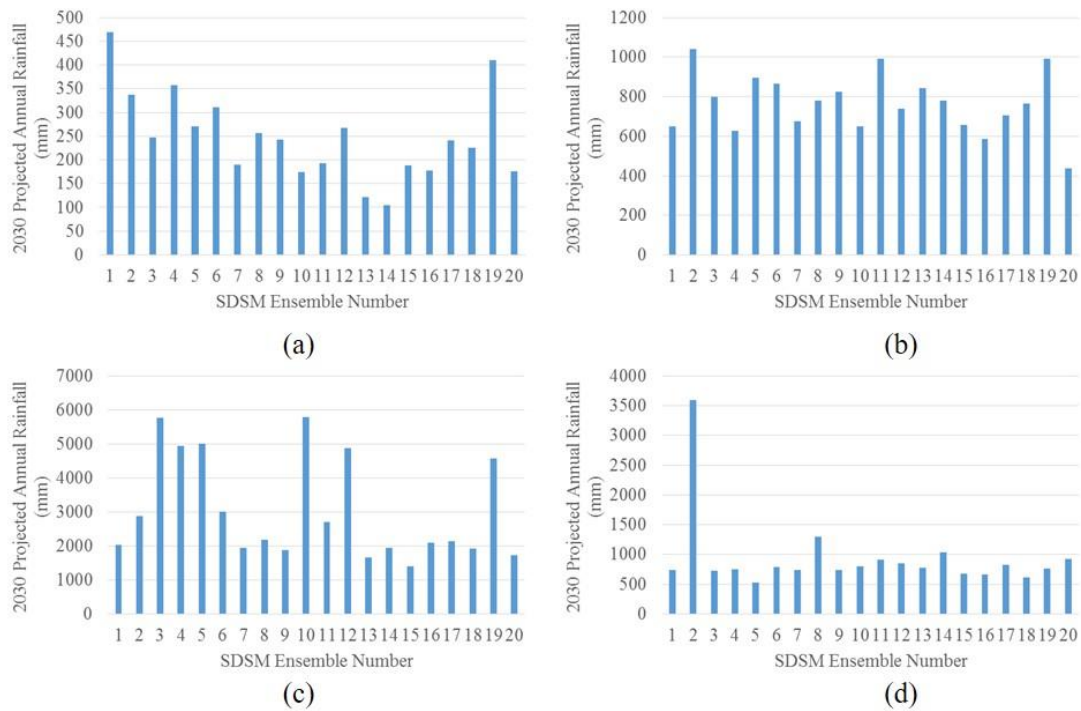
**Table 4-6.** Monthly Calibration Results for Daily Rainfall (1990-2000) with daily GCM predictors

	North	South	East	West
Month	R <sup>2</sup>	R <sup>2</sup>	R <sup>2</sup>	R <sup>2</sup>
January	0.253	1.00	1.00	1.00
February	0.876	0.810	0.573	0.237
March	0.208	0.316	0.382	0.28
April	0.0419	0.099	0.210	0.171
May	0.199	0.275	0.157	0.242
June	0.234	0.127	0.122	0.0767
July	0.09	0.121	0.0875	0.0770
August	0.187	0.151	0.141	0.173
September	0.07	0.0623	0.120	0.0663
October	0.637	0.0876	0.403	0.399
November	0.462	0.357	0.175	0.484
December	0.992	0.769	0.910	0.979

The south (Figure 4-8b) and east stations (Figure 4-8c) show significant model bias in projected annual rainfall with the majority of SDSM model ensembles reporting significant annual rainfall that physically may not be possible. Ensemble #2 of the west station (Figure 4-8d) also has significant bias compared with the other SDSM ensembles of Figure 8d. Model projections of 2030 annual rainfall show increased annual rainfall for the south, east and west stations overall compared to 1990-2015 period on an annual basis. The opposite is true for the north station for which the 2030 projected annual rainfall is at or less than the rainfall on an annual basis for the 1990-2015 period.

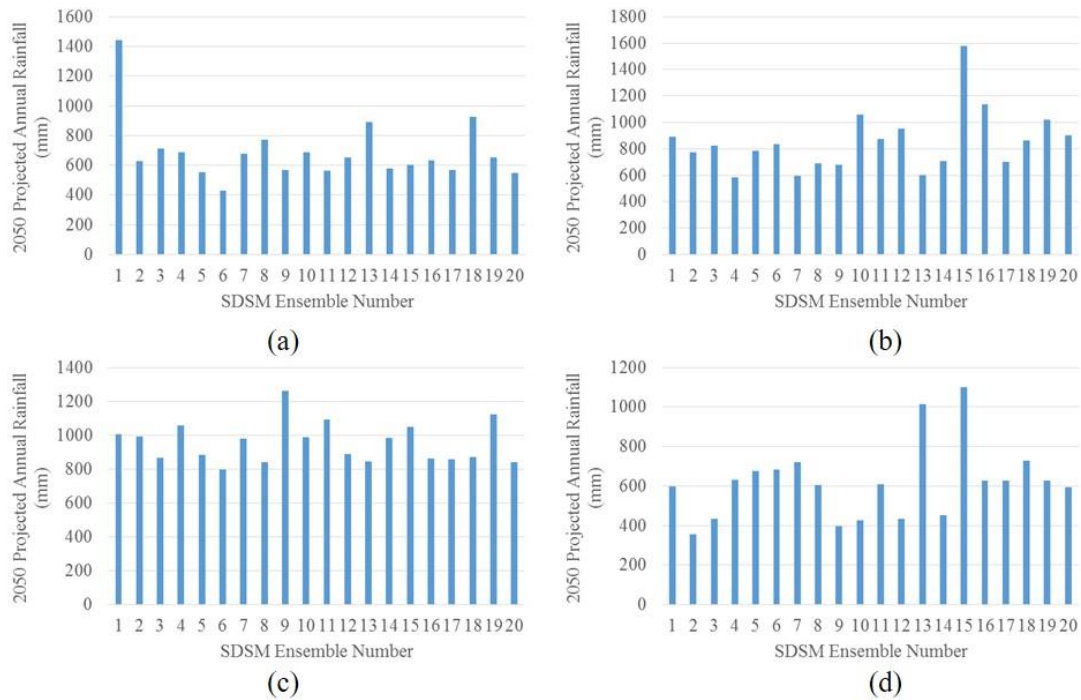


**Figure 4-7.** Validation results for period of January 2005-January 2010 for (a) North station, (b) South station, (c) East station and (d) West station



**Figure 4-8.** Projected annual rainfall for year 2030 at (a) North station, (b) South station, (c) East station and (d) West station

North, south and east stations (Figures 4-9a-c) show projected increases in annual rainfall for the year 2050 as compared with annual rainfall reported for the 1990-2015 period. The west station is the exception with projected annual rainfall in most cases less than majority of annual rainfall during the 1990-2015 period. Changes in model bias are also apparent for the year 2050 with the model reporting much lower annual rainfall for the south, east and west stations as compared to the year 2030.



**Figure 4-9.** Projected annual rainfall for year 2050 at (a) North station, (b) South station, (c) East station and (d) West station

#### 4.5.2. CA-based MC model

##### 4.5.2.1. MC Transition Probability Matrices

The generated MC transition probability matrices are presented as follows (Tables 4-7, 4-8, and 4-9). Table 4-7 is used to predict LULC map in 2010 since the project forward years are 5 years after 2005. Similarly, Tables 4-8 and 4-9 represent the matrices used for predicting LULC maps in 2030 and 2050 and the corresponding projection forward years are 25 and 45 years, respectively.

**Table 4-7.** Transition probability matrix for predicting LULC map in 2010

	water	urban	forest	farmland	rangeland	barren
water	0.9191	0.0285	0.0038	0.0339	0.0147	0.0000
urban	0.0009	0.9854	0.0019	0.0111	0.0008	0.0000
forest	0.0007	0.0047	0.9852	0.0025	0.0069	0.0000
farmland	0.0020	0.0815	0.0046	0.9086	0.0034	0.0000
rangeland	0.0008	0.0021	0.0138	0.0036	0.9798	0.0000
barren	0.0000	0.0000	0.0251	0.0008	0.0313	0.9428

**Table 4-8.** Transition probability matrix for predicting LULC map in 2030

	water	urban	forest	farmland	rangeland	barren
water	0.6568	0.1409	0.0192	0.1218	0.0613	0.0000
urban	0.0037	0.9374	0.0093	0.0450	0.0045	0.0000
forest	0.0030	0.0243	0.9292	0.0109	0.0326	0.0000
farmland	0.0075	0.3303	0.0203	0.6272	0.0148	0.0000
rangeland	0.0033	0.0132	0.0644	0.0150	0.9041	0.0000
barren	0.0004	0.0024	0.1123	0.0044	0.1355	0.7450



**Table 4-9.** Transition probability matrix for predicting LULC map in 2050

	water	urban	forest	farmland	rangeland	barren
water	0.4705	0.2429	0.0344	0.1593	0.0928	0.0000
urban	0.0059	0.9018	0.0169	0.0670	0.0085	0.0000
forest	0.0048	0.0446	0.8781	0.0175	0.0550	0.0000
farmland	0.0105	0.4889	0.0329	0.4443	0.0234	0.0000
rangeland	0.0051	0.0275	0.1084	0.0230	0.8359	0.0001
barren	0.0011	0.0081	0.1818	0.0084	0.2117	0.5887

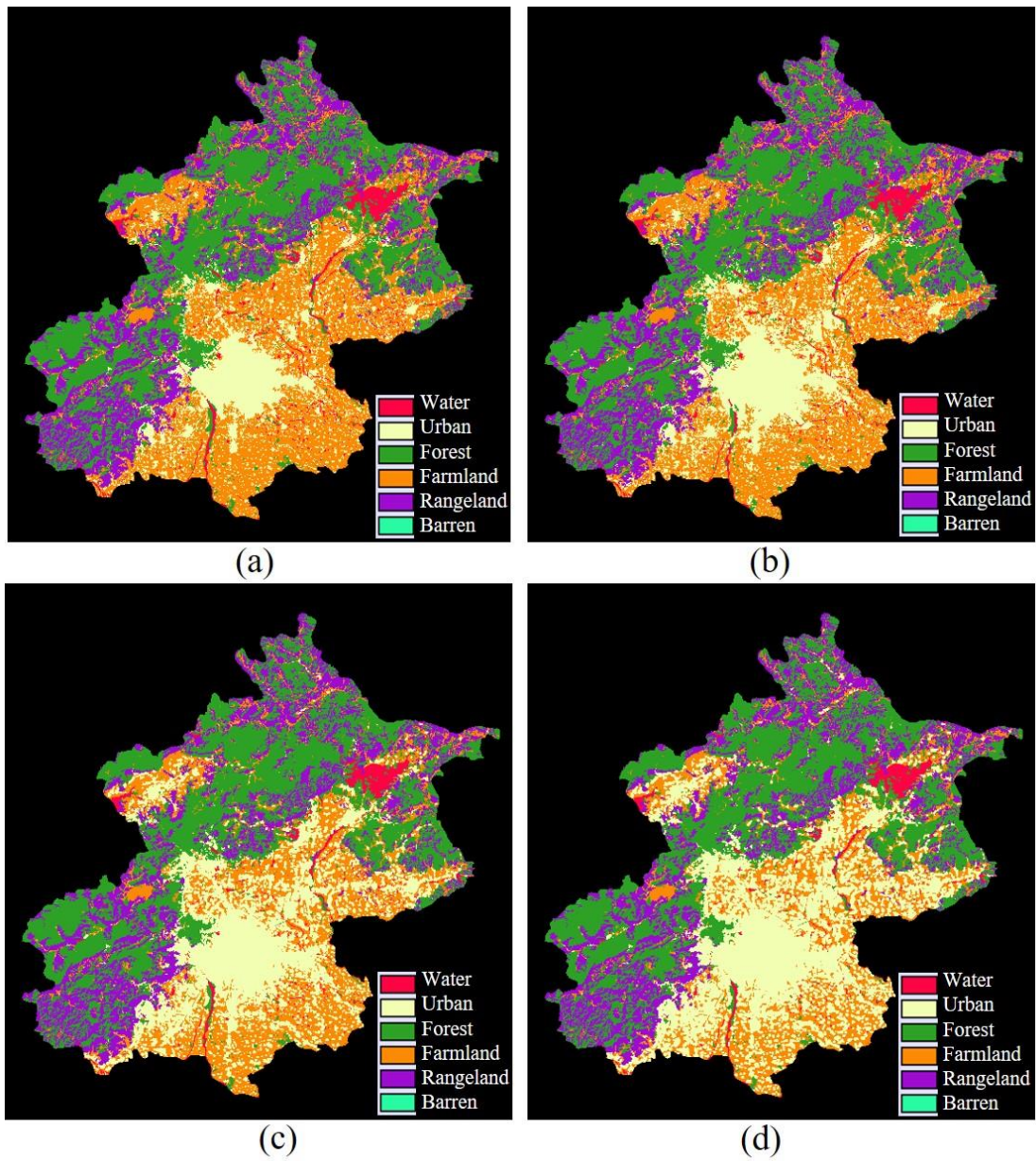
#### *4.5.2.2. Prediction LULC Maps*

Four scenarios were discussed herein. The comparison of validation results for these four scenarios is summarized in Table 4-10. Since the indexes related to Kappa statistics are all well above 80%, the urban growth model is trained reliably to predict possible changes of LULC in 2030 and 2050.

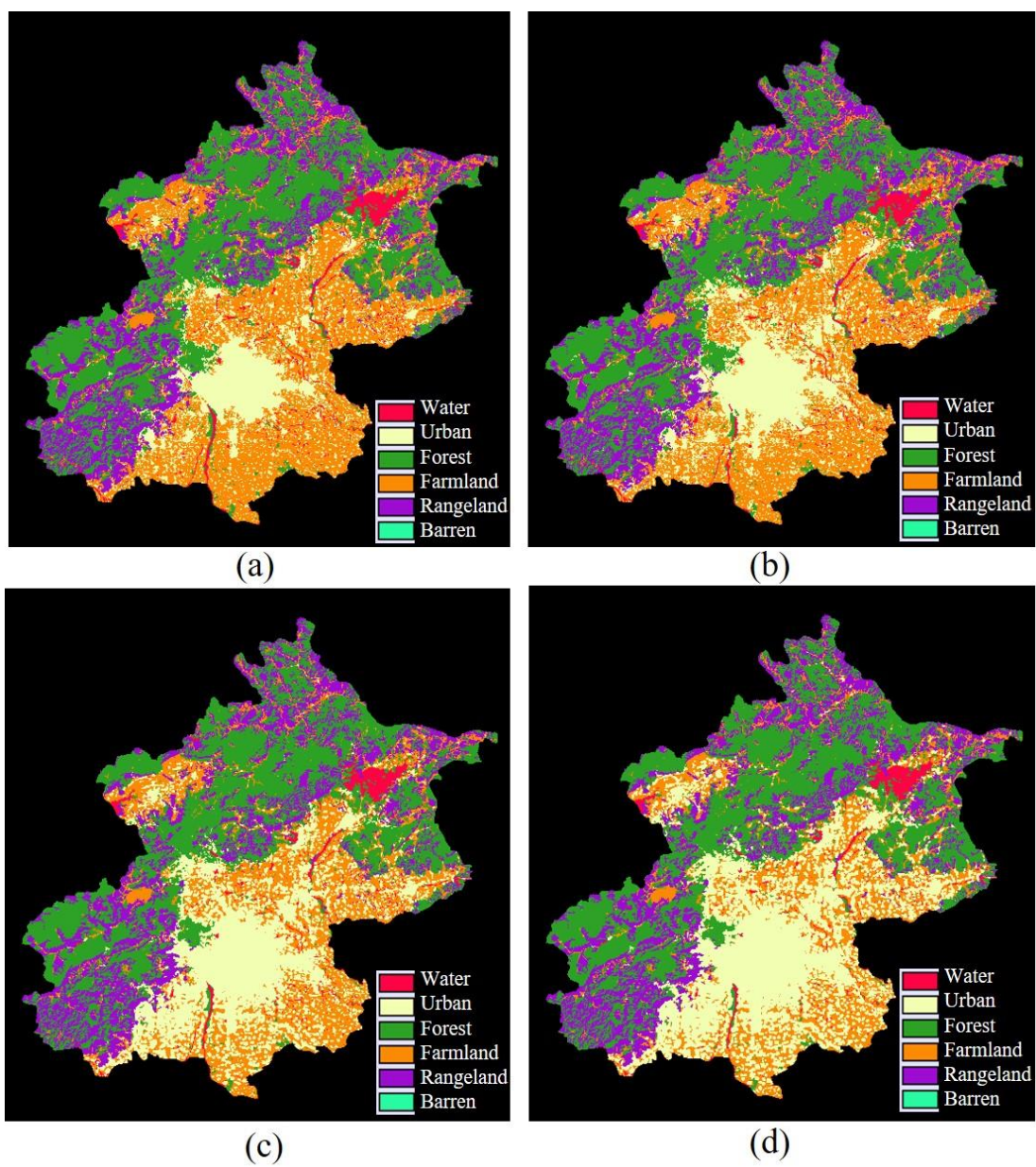
**Table 4-10.** Comparison of validation results between scenarios 1, 2, 3, and 4

Kappa	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Kstandard	0.9275	0.9282	0.9268	0.9253
Kno	0.9351	0.9357	0.9334	0.9331
Klocation	0.9419	0.9425	0.9410	0.9396

The prediction results are presented in Figures 4-10, 4-11, 4-12, and 4-13 for each scenario and comparison is presented in Tables 4-11, 4-12, 4-13, and 4-14. Compared to the observed LULC map in 2010, the projected 2030 and 2050 maps indicate the urban expansion in the southeast, southwest, and north of Beijing, especially around the railway and road systems. In the northwest (Yanqing district), urbanization will also occur with the conversion from farmland. In general, urban area is expected to increase more than 6% and 11% in 2030 and 2050, respectively, compared to the situation of 2010. Overall, farmland will decrease significantly due to the fast urbanization. Forest, rangeland, and water body are expected to decrease slightly in the future, resulting from the urban expansion and climate change.

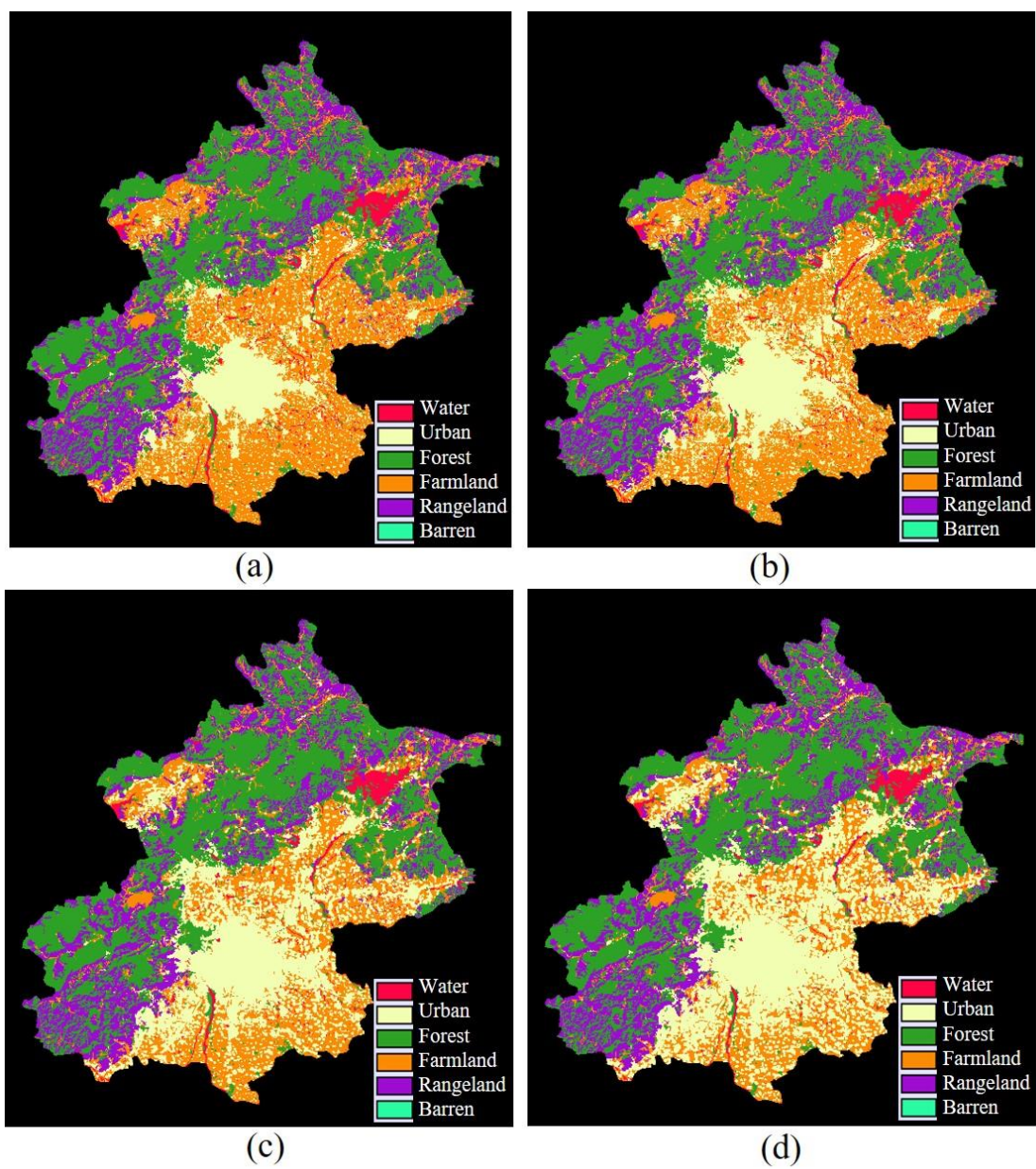


**Figure 4-10.** Observed maps (a) 2000 and (b) 2010 vs. scenario 1 projections (c) 2030 and (d) 2050

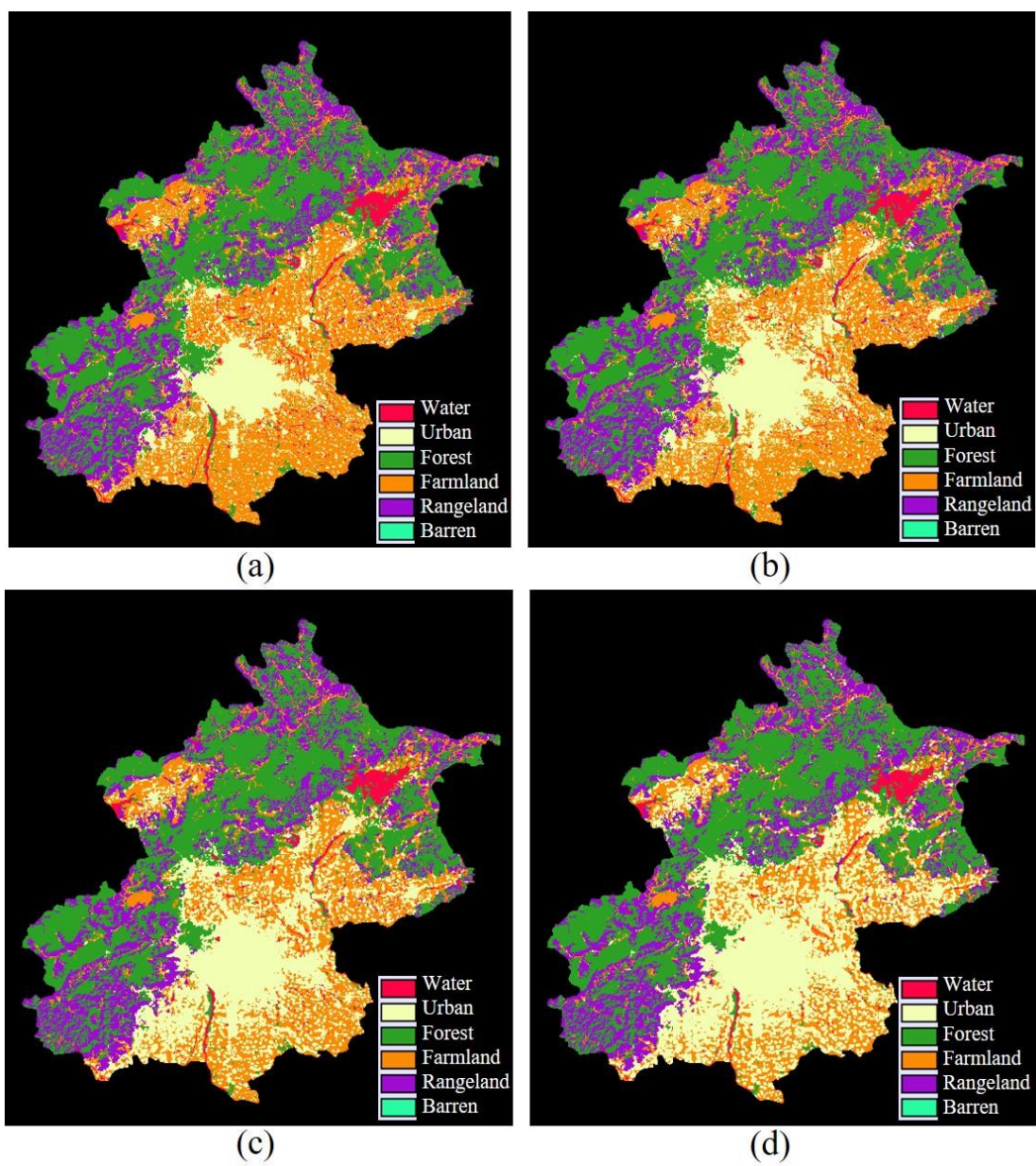


**Figure 4-11.** Observed maps (a) 2000 and (b) 2010 vs. scenario 2 projections (c) 2030 and (d) 2050





**Figure 4-12.** Observed maps (a) 2000 and (b) 2010 vs. scenario 3 projections (c) 2030 and (d) 2050



**Figure 4-13.** Observed maps (a) 2000 and (b) 2010 vs. scenario 4 projections (c) 2030 and (d) 2050

**Table 4-11.** Predicted results of Scenario 1

land use	Scenario 1					
	2030		2010-2030	2050		2010-2050
	unit:km2	in %	in %	unit: km2	in %	in %
water	409.5720	2.52%	-0.38%	349.8219	2.16%	-0.75%
urban	3806.7030	23.45%	6.18%	4631.1570	28.53%	11.26%
forest	5391.8217	33.22%	-0.21%	5341.1976	32.90%	-0.52%
farmland	3488.6169	21.49%	-5.55%	2802.6180	17.27%	-9.78%
rangeland	3134.4687	19.31%	-0.04%	3106.3995	19.14%	-0.21%
barren	1.1583	0.01%	0.00%	1.1466	0.01%	0.00%

**Table 4-12.** Predicted results of Scenario 2

land use	Scenario 2					
	2030		2010-2030	2050		2010-2050
	unit:km2	in %	in %	unit: km2	in %	in %
water	409.6737	2.52%	-0.38%	349.8840	2.16%	-0.75%
urban	3806.7030	23.45%	6.18%	4631.1570	28.53%	11.26%
forest	5391.7200	33.22%	-0.21%	5341.1427	32.90%	-0.52%
farmland	3488.6169	21.49%	-5.55%	2802.6180	17.27%	-9.78%
rangeland	3134.4687	19.31%	-0.04%	3106.4454	19.14%	-0.21%
barren	1.1583	0.01%	0.00%	1.0935	0.01%	0.00%

**Table 4-13.** Predicted results of Scenario 3

land use	Scenario 3					
	2030		2010-2030	2050		2010-2050
	unit:km2	in %	in %	unit: km2	in %	in %
water	409.5891	2.52%	-0.38%	349.8201	2.16%	-0.75%
urban	3806.7030	23.45%	6.18%	4631.1570	28.53%	11.26%
forest	5391.8046	33.22%	-0.21%	5341.1999	32.90%	-0.52%
farmland	3488.6169	21.49%	-5.55%	2802.6180	17.27%	-9.78%
rangeland	3134.4687	19.31%	-0.04%	3106.3995	19.14%	-0.21%
barren	1.1583	0.01%	0.00%	1.1466	0.01%	0.00%

**Table 4-14.** Predicted results of Scenario 4

land use	Scenario 4					
	2030		2010-2030	2050		2010-2050
	unit:km2	in %	in %	unit: km2	in %	in %
water	409.8321	2.52%	-0.38%	349.8840	2.16%	-0.75%
urban	3806.7030	23.45%	6.18%	4631.1570	28.53%	11.26%
forest	5391.5616	33.21%	-0.21%	5341.1427	32.90%	-0.52%
farmland	3488.6169	21.49%	-5.55%	2802.6180	17.27%	-9.78%
rangeland	3134.4687	19.31%	-0.04%	3106.4454	19.14%	-0.21%
barren	1.1583	0.01%	0.00%	1.0935	0.01%	0.00%



Based on the comparison among these four scenarios, the major land use changes are attributed to changes of water and forest. With the proposal of the new economic free trade zone (Xiong'an) (scenario 2), water body is expected to increase and forest will decrease due to the formation of new corridor relative to scenario 1. More forest land is expected to contribute to the urban growth in the south and southwest of Beijing. For scenario 3, the area of water body will be more than that in scenario 1 but less than that in scenario 2. There will be more forest land since the government will be relocated to the southeast and far away from the forest area in the west and north. When considering the impact of climate change in scenario 4, there will be more water body since rainfall is expected to increase in the future. Forest area will decrease holistically because of the impact of climate change and urbanization. Scenario 4 indicates the worst case that more area will face the flood risk while less forest area can be available to relieve the impact of climate change.

#### 4.6. Discussion

Compared to the previous research of urban growth modeling for Beijing, the prediction results of the current study conform to the previous findings in general that indicates the possible locations and expansion directions of urban growth within the traditional city boundary in the future, although the prediction years were different (Table 4-15). Yet, this study not only considered the traditional socio-economic driving factors but also included the newest policy of establishing a new economic free trade zone (Xiong'an), which may significantly impact the economic and urban growth of Beijing regionally in the next decades. In addition, climate change and flood impact were also considered to link the global climate impact to socio-economic and policy-oriented factors. Such comprehensive consideration of current work could provide new ideas to urban growth modeling and make contribution to urban planning and management in the future.

**Table 4-15.** Comparison between previous studies and the current study for Beijing metropolitan area

Categories	Previous studies	The current study
Model type	CA based models	CA-based MC model
Driving factors	Socio-economic	Socio-economic, new policy, corridor effect, climate change
Prediction year	2004-2021	2030 and 2050
Prediction results	<ul style="list-style-type: none"> <li>• Urban expansion in rural areas under different scenarios.</li> <li>• Geographical environment limits urban growth.</li> <li>• LULC variations among each LULC types</li> </ul>	<ul style="list-style-type: none"> <li>• Urban expansion in rural areas due to government relocation.</li> <li>• Scenario comparison indicates the impact of the new policy</li> <li>• Climate change and flooding potential.</li> </ul>
Categories	Previous studies	The current study
Model type	CA based models	CA-based MC model
Driving factors	Socio-economic	Socio-economic, new policy, corridor effect, climate change
Prediction year	2004-2021	2030 and 2050

Categories	Previous studies	The current study
Prediction results	Urban expansion in rural areas under different scenarios. Geographical environment limits urban growth. LULC variations among each LULC types	Urban expansion in rural areas. Scenario comparison indicates the impact of the new policy and climate change.

However, there are some limitations in this study as well. The difference of predicted urban growth among the four scenarios is not phenomenal. Due to the availability of flood data from the NASA, the flood record that only contains a 4-year data set from 2013 to 2017 can only permit partial impact on the long term urban growth prediction. The longer the time period of our data, the more comprehensive and reliable the model will be. Another limitation of this study is that model can only reflect driving factors by setting distance to each factor with their fuzzy membership functions within a given geographical boundary. It is difficult to realize the impacts from external regions remotely such as the corridor effect and reflect those impacts in the model. The formation of the new corridor (Beijing-Tianjin, and Beijing-Xiong'an) was not being reflected completely in this integrated modeling analysis. Corridor effect and regional planning must be quantified and reflected by using a regionally planning model to be coupled with the current model to address the scaling effect. Therefore, to improve the current urban growth model, future work may focus on obtaining a more sophisticated and comprehensive dataset and enriching the form of reflecting

different driving factors. Regional planning can be coupled into the current model to promote the modeling efficacy over geographical boundaries.

#### 4.7. Conclusion

Urbanization has become a global trend making tremendous LULC changes. As one of the fastest growing megacities, Beijing has experienced huge land development in the past three decades. In order to analyze the LULC changes in Beijing, an integrated model of CA, MC, and MCE-WLC was applied to simulate and predict the possible LULC changes in 2030 and 2050 under uncertainty in this study. The driving factors not only included the traditional socio-economic factors, but also involved the newest policy of a new economic free trade zone at Xiong'an in 2017. In addition, SDSM was used in this study to predict rainfall variations in the future, which helps support and analyze the impact of climate change on the ultimate urban growth.

Four scenarios were included in this study to examine the impacts of socio-economic, policy-oriented, and climate change as three major driving factors. Compared to the baseline in 2010, the predicted results clearly indicate that urban expansion is expected to increase more than 6% and 11% in 2030 and 2050, respectively. The directions of urban growth are in the southeast, southwest, and north area of the urban area in Beijing, especially around the transportation systems. Farmland that will be largely converted for urbanization will make great contribution to the urban growth and economic development. Forest is expected to decrease due to the projected urbanization (scenarios 2 and 3) and climate change (scenario 4) and affect the area of water body in different scenarios. Based on the scenario 4, climate change will threaten the development of the city and bring more rainfall in the future. More research should be conducted in this context to

respond to climate change and other environmental issues such as urban sustainability in a food-water-energy nexus.

#### 4.8. References

- Arsanjani, J.J., Helbich, M., Kainz, W., Boloorani, A.D., 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation* 21, 265-275.
- Batty, M., 2007. *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press.
- Eastman, J., 2015. *TerrSet manual*. Clark Labs. Clark University. Worcester, MA, USA.
- Beijing Economic and Social Development Statistical Bulletin, 2015 (in Chinese). Beijing Bureau of Statistics. Retrieved June 10, 2017.
- Guan, Q., Wang, L., Clarke, K.C., 2005. An artificial-neural-network-based, constrained CA model for simulating urban growth. *Cartography and Geographic Information Science* 32, 369-380.
- Wu F., 1998, SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules, *Int. J. Geographical Information Science*, vol. 12, no. 1, pp. 63-82.
- China Knowledge, 2013. Shenzhen (Guangdong) City information. Retrieved from <http://www.chinaknowledge.com/CityInfo/City.aspx?Region=Coastal&City=Shenzhen>
- World Bank, 2015. World Bank national accounts data and OECD National Accounts data files. Retrieved from <http://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=CN>

World Population Review, 2016. Shenzhen Population. Retrieved from

<http://worldpopulationreview.com/world-cities/shenzhen-population/>

Gulacha, M.M. and Mulungu, D.M.M., 2016. Generation of climate change scenarios for precipitation and temperature at local scales using SDSM in Wami-Ruvu River Basin Tanzania, Physics and Chemistry of the Earth, Parts A/B/C, 1-11.

Hessami, M., Gachon, P., Ouarda, T. B.M.J., and St-Hilaire, A., 2008. Automated regression-based statistical downscaling tool, Environmental Modelling & Software, 23(6): 813-834.

Intergovernmental Panel on Climate Change, Working Group I, 2007. Climate change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC. Cambridge University Press, Cambridge, UK.

Li, Q., Zhang, C., and Miao, S., 2005. The distribution characteristics of rainfall and the effect of land use in Beijing area, J. Desert Res., 25 (Suppl.), 60–65.

Liu, W.D., 2013. Thinking of economic geography. Sciences Press: Beijing, China. 165-165.

Miao, S., Chen, F., Li, Q. and Fan, S., 2011. Impacts of urban processes and urbanization on summer precipitation: a case study of heavy rainfall in Beijing on 1 August 2006, Journal of Applied Meteorology and Climatology 50 (4), 806-825.

Muller, M.R., Middleton, J., 1994. A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. Landscape Ecology 9, 151-157.

Owen, J.G., 2005. Estimating the cost and benefit of hosting Olympic Games: What can Beijing expect from its 2008 Games? The Industrial Geographer; Terre Haute 3.1 (Fall 2005):1-18.

Subedi, P., Subedi, K., and Thapa, B. (2013) Application of a Hybrid Cellular Automaton – Markov (CA-Markov) Model in Land-Use Change Prediction: A Case Study of Saddle

- Creek Drainage Basin, Florida, *Applied Ecology and Environmental Sciences*, 1(6), 126-132.
- Teutschbein, C., Wetterhall F. and Seibert J., 2011: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dynamics*, 37(9): 2087–2105.
- White, R., Engelen, G., 1993. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment and planning A* 25, 1175-1199.
- Wilby, R.L, Dawson, C.W., Barrow, E.M., 2002: SDSM—a decision support tool for the assessment of regional climate change impacts, *Environ. Model Software.*, 17 (2): 145–157.
- Zhang, C.L., Chen, F., Miao, S.G., Li, Q. C., Xia, X.A. and Xuan, C.Y., 2009: Impacts of urban expansion and future green planting on summer precipitation in the Beijing metropolitan area. *J. Geophys. Res.*, 114, D02116

## **CHAPTER 5 CONCLUSIONS**

Based on the simulation and prediction results of each megacity, the urban growth types can be identified. Since cities have different urban growth patterns with their unique driving factors, land use changes and management are various. With the development of urbanization and globalization, the three study megacities are expected to achieve urban growth in different levels. Although climate change and flood impact do not significantly influence the land use changes and urban development now, the situation may keep deteriorating in the future and impact urban growth, which needs to be paid attention by all urban planners and decision makers.

Future work may involve creating and obtaining a more sophisticated and comprehensive dataset and enriching the form of reflecting different driving factors. Regional planning can be coupled into the current model to promote the modeling efficacy over geographical boundaries. In order to respond to climate change and flood impact, more research should be conducted focusing on flood impact assessment especially in megacities. Appropriate measurement such as green infrastructure to and low impact development should be done to alleviate the impact of climate change.



## **APPENDIX: ACCEPTED MATERIAL FOR PUBLICATION**

Chapter 2 of this thesis incorporate material that has been accepted for publication as follows:

### **Chapter 2**

Lu, Q., Joyce, J., Imen, S., Chang, N.B., 2017. Linking Socioeconomic Development, Sea Level Rise, and Climate Change Impacts on Urban Growth in New York City with a Fuzzy Cellular Automata-based Markov Chain Model. *Environment and Planning B: Urban Analytics and City Science*. Accepted for publication.