URBAN GROWTH MODELING OF PHNOM PENH, CAMBODIA USING REMOTELY SENSED DATA AND LOGISTIC REGRESSION MODEL

Kompheak Mom

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Geoinformatics

Suranaree University of Technology

Academic Year 2015

การสร้างแบบจำลองการเจริญเติบโตของกรุงพนมเปญประเทศกัมพูชาจาก ข้อมูลการรับรู้จากระยะไกลและแบบจำลองการถดถอยโลจิสติก

นายกมแกะ มม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาภูมิสารสนเทศ

มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2558

URBAN GROWTH MODELING OF PHNOM PENH, CAMBODIA USING REMOTELY SENSED DATA AND LOGISTIC REGRESSION MODEL

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee

____
(Assoc. Prof. Dr. Songkot Dasananda)
Chairperson

____
(Assoc. Prof. Dr. Suwit Ongsomwang)
Member (Thesis Advisor)

____
(Asst. Prof. Dr. Sunya Sarapirome)
Member

____
(Dr. Pantip Piyatadsananon)
Member

____
(Prof. Dr. Sukit Limpijumnong)
Vice Rector for Academic Affairs
and Innovation

____
(Prof. Dr. Santi Maensiri)
Dean of Institute of Science

KOMPHEAK MOM : การสร้างแบบจำลองการเจริญเติบโตของกรุงพนมเปญประเทศกัมพูชาจากข้อมูลการรับรู้จากระยะไกลและแบบจำลองการถดถอยโลจิสติก (URBAN GROWTH MODELING OF PHNOM PENH, CAMBODIA USING REMOTELY SENSED DATA AND LOGISTIC REGRESSION MODEL) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.สุวิทย์ อ๋องสมหวัง, 141 หน้า.

กรุงพนมเปญกำลังเผชิญกับอัตราการเพิ่มขึ้นของประชากรอย่างรวดเร็ว ที่มีอัตราการขยายตัวของเมืองสูงในทวีปเอเชีย และประสบปัญหาเรื่องการวางผังเมืองที่ดี ส่งผลทำให้พื้นที่ทางธรรมชาติและพื้นที่เกษตรกรรมลดลง ตัวบ่งที่ชัดเจนในเรื่องนี้คือ กรุงพนมเปญมีการขยายตัวของเมืองอย่างไม่เป็นระเบียบ (sprawl) และการเติบโตในลักษณะไม่ยั่งยืน ฉะนั้น ความเข้าใจพลวัตเชิงพื้นที่และเชิงเวลาและปัจจัยขับเคลื่อนการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดิน จึงเป็นสิ่งจำเป็นเพื่อช่วยสนับสนุนนโยบายการวางผังเมืองเพื่อนำไปสู่การแก้ไขปัญหาดังกล่าวข้างต้น วัตถุประสงค์ของการศึกษาคือ (1) เพื่อประเมินสถานภาพและการเปลี่ยนแปลงของสิ่งปกคลุมดิน (2) เพื่อใช้แบบจำลองการถดถอยโลจิสติกค้นหาปัจจัยขับเคลื่อนการเจริญเติบโตของกรุงพนมเปญ และ (3) เพื่อคาดการณ์รูปแบบการเจริญเติบโตของกรุงพนมเปญในปี ค.ศ. 2030 วิธีการศึกษาประกอบด้วย 4 องค์ประกอบหลัก ได้แก่ (1) การรวบรวมข้อมูล (2) การเตรียมข้อมูล (3) การจำลองและตรวจสอบความสมเหตุสมผลของแบบจำลอง และ (4) การคาดการณ์การเจริญเติบโตของเมือง

ผลการศึกษา พบว่า พื้นที่เมืองและสิ่งปลูกสร้างเพิ่มขึ้นอย่างต่อเนื่องตั้งแต่ปี ค.ศ. 2002 ถึง 2015 ส่งผลทำให้พื้นที่เพาะปลูก พืชพรรณ และแหล่งน้ำลดลง ในขณะเดียวกัน พื้นที่เบ็ดเตล็ดที่ส่วนใหญ่เป็นที่ดินเปิดโล่งสำหรับการก่อสร้างและการพัฒนามีความผันผวน ในขณะที่รูปแบบของการเจริญเติบโตของเมืองขยายไปทางทิศใต้ เหนือ และตะวันตกของกรุงพนมเปญ ในระหว่างปี ค.ศ. 2002-2009 ซึ่งมีรูปแบบการเจริญเติบโตของเมืองทุกรูปแบบ ประกอบด้วย Infill growth Expansion growth Linear branch และ Isolated growth อย่างไรก็ตาม ในระหว่างปี ค.ศ. 2009-2015 รูปแบบการเจริญเติบโตของเมืองเกิดขึ้นในทุกทิศทางที่มีรูปแบบการเจริญเติบโตของเมืองประกอบด้วย Expansion growth Clustered branch และ Isolated growth

ผลลัพธ์ของการค้นหาปัจจัยอิทธิพลต่อการเจริญเติบโตของเมืองระหว่างปี ค.ศ. 2002-2015 2002-2009 และ 2009-2015 ได้บ่งชี้ว่า ในระหว่างปี ค.ศ. 2002-2015 ที่เป็นคาบเวลาระยะยาว ปัจจัยอิทธิพล 3 อันดับแรก ที่ส่งผลทางลบต่อการเจริญเติบโตของเมือง ประกอบด้วย ระยะห่างจากสนามบินกรุงพนมเปญ ถนนหลักและพื้นที่อุตสาหกรรม อย่างไรก็ตาม ในคาบเวลาระยะสั้นระหว่างปี ค.ศ. 2002-2009 และ 2009-2015 ปัจจัยอิทธิพลที่ส่งผลทางลบต่อการเจริญเติบโตของ

เมือง ได้แก่ ระยะห่างจากกลุ่มเมืองที่คงอยู่เป็นปัจจัยสำคัญลำดับแรกของ 2 คาบเวลา แต่ปัจจัยลำดับที่สองและสามจะแตกต่างกัน ในที่นี้ ระยะห่างจากพื้นที่อุตสาหกรรมและสนามบินกรุงพนมเปญเป็นปัจจัยอิทธิพลลำดับที่สองและสามซึ่งส่งผลทางลบในคาบเวลาระหว่างปี ค.ศ. 2002-2009 ในขณะที่ ระยะห่างจากศูนย์กลางย่านธุรกิจและพื้นเบ็ดเตล็ดเป็นปัจจัยอิทธิพลที่ส่งผลทางลบและทางบวกในคาบเวลาระหว่างปี ค.ศ. 2009-2015 ตามลำดับ ผลลัพธ์สุดท้าย จากการคาดการณ์การเจริญเติบโตของเมืองจากแบบจำลองโลจิสติกที่เหมาะสม พบว่า กรุงพนมเปญจะขยายตัวไปทางทิศเหนือ ใต้ และตะวันออกในปี ค.ศ. 2030 ในทุกประเภทสิ่งปกคลุมดิน

จากผลการศึกษาที่ได้รับ สามารถสรุปได้ว่า แบบจำลองโลจิสติกที่บูรณาการร่วมกับการข้อมูลการรับรู้จากระยะไกลและเทคนิคจากระบบสารสนเทศภูมิศาสตร์เป็นวิธีการที่มีประสิทธิภาพสำหรับใช้อธิบายอิทธิพลของปัจจัยขับเคลื่อนที่มีผลต่อการเจริญเติบโตของกรุงพนมเปญ วิธีการนี้สามารถให้สารสนเทศที่มีสำคัญเพื่อใช้เป็นแนวทางสำหรับนักวางแผนและผู้ตัดสินใจ

สาขาวิชาการรับรู้จากระยะไกล ลายมือชื่อนักศึกษา____________________________

ปีการศึกษา 2558 ลายมือชื่ออาจารย์ที่ปรึกษา______________________

KOMPHEAK MOM : URBAN GROWTH MODELING OF PHNOM PENH, CAMBODIA USING REMOTELY SENSED DATA AND LOGISTIC REGRESSION. THESIS ADVISOR : ASSOC. PROF. SUWIT ONGSOMWANG, Dr. rer. Nat. 141 PP.

URBAN GROWTH MODELING / LOGISTIC REGRESSION MODEL / LAND COVER / PHNOM PENH CITY / CAMBODIA

Phnom Penh City has faced rapid population growth with high urban expansion rate in Asia, and it has been encountered with poor urban planning that resulted in reducing natural area and agricultural land. This indicated that it recently goes toward sprawl and grow in unsustainable way. To find out this problem, an understanding of spatial and temporal dynamic and the driving forces of land use and land cover change should be conducted in order to enhance urban planning policy. The specific objectives of this study are (1) to assess land cover status and its change; (2) to employ logistic regression model (LR) to discover the driving factors of Phnom Penh urban growth; and (3) to predict the future urban growth pattern of Phnom Penh in 2030. Four main components of research methodology are here conducted included (1) data collection; (2) data preparation; (3) model simulation and validation; and (4) urban growth prediction.

Results showed that urban and built-up area continuously increased from 2002 to 2015 resulting a major decline of arable land, vegetation and water body, while miscellaneous land that majority was bare land for construction and development sites was shown as fluctuation. Meanwhile, pattern of urban growth expanded toward the

southern, northern and western direction of Phnom Penh during 2002-2009 with all type of growth including infill growth, expansion growth, linear branch and isolated growth. However, during 2009-2015, urban growth pattern occurred in all direction with expansion growth, clustered branch and isolated growth.

Result of finding the key influential factors for 2002-2015, 2002-2009 and 2009-2015 periods were indicated that during 2002-2015 as long term period, top three negative influence factors were distance to Phnom Penh airport, distance to road and distance to industrial area. However, the short term period results of 2002-2009 and 2009-2015 showed the negative influence of distance to existing urban cluster, which is the first ranking of both periods, but the second and third ranking were different. Herein, distance to industrial area and Phnom Penh airport were the second and third negative influential factors of 2002-2009 period, while distance to CBD and miscellaneous land were shown as negative and positive influential factors in 2009-2015 period. Lastly, the prediction from the optimum LR model revealed that Phnom Penh tends to expand in the north, south and east direction in 2030 on all types of land cover.

In conclusion, result of finding showed that LR model incorporated with remote sensing data and GIS technique are efficient approach to apply for understanding the influence of driving factors to the urban growth of Phnom Penh. It contributes the significant information as a guideline for planner and decision maker.

School of Remote Sensing Student's Signature ________________________

Academic Year 2015 Advisor's Signature________________________

ACKNOWLEDGEMENTS

This research project would not have been possible without the support of many people. I wish to express my deepest appreciation and gratitude to my advisor, Assoc. Prof. Dr. Suwit Ongsomwang who was extremely supportive and helpful for providing idea, discussion and technical assistance through the learning process of this master thesis.

I would like to express my gratitude to the committee members, Asst. Prof. Dr. Songkot Dasananda for comments that helpfully improve the research scheme, Asst. Prof. Dr. Sunya Sarapirome for guidance with some suggestion on correcting this thesis, and Dr. Pantip Piyatadsananon for sharing her precious time to discuss ideas on urban planning theory, as well as the process of developing thesis proposal.

Gratefulness are also due to the support from ASEA-UNINET Thailand On-Place Scholarship, the Thailand Higher Educational Office, and the Suranaree University of Technology without their support in allowance, accommodation and university fee, this study would not have been successful.

Special thanks also to Mr. Meang Sotha, the GIS consultant at Ministry of Land Management Urban Planning and Construction Cambodia and a lecture at Paññāsāstra University of Cambodia for his sharing related data in Phnom Penh.

I extend my thanks to all the staffs and students of School of Remote Sensing for their help and support during my study here, especially Mr. Krittawat Boonchoo, Mr. Athiwat Phinyoyang and Miss Salisa Saraisamrong.

Finally, I wish to express my love and gratitude to my beloved families for their understanding and encouragement through the duration of my study.

Kompheak MOM

CONTENTS

CONTENTS (Continued)

Page

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF TABLES (Continued)

Table **Page**

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure **Page**

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

CA	=	Cellular Automata
CBD	=	Central Business District
FA	=	Factor Analysis
GDP	=	Gross Domestic Product
GIS	=	Geographic Information System
LL_0	=	Log-Likelihood of the Model with Intercept Only
LL_m	=	Log-Likelihood of the Model
LR	=	Logistic Regression Model
LULC	=	Land Use Land Cover
NGD	=	National Geographic Department
NIS	=	National Institute of Statistics
OLS	=	Ordinary Least Squares
OR	=	Odds Ratio
PA	=	Producer’s Accuracy
PCA	=	Principle Component Analysis
PPSEZ	=	Phnom Penh Special Economic Zone
ROC	=	Relative Operating Characteristic
SE_b	=	Standard Error
TOL	=	Tolerence
UA	=	User’s Accuracy

LIST OF ABBREVIATIONS (Continued)

UNFPA	=	United Nations Population Fund
VIF	=	Variance Inflation Factor

CHAPTER I

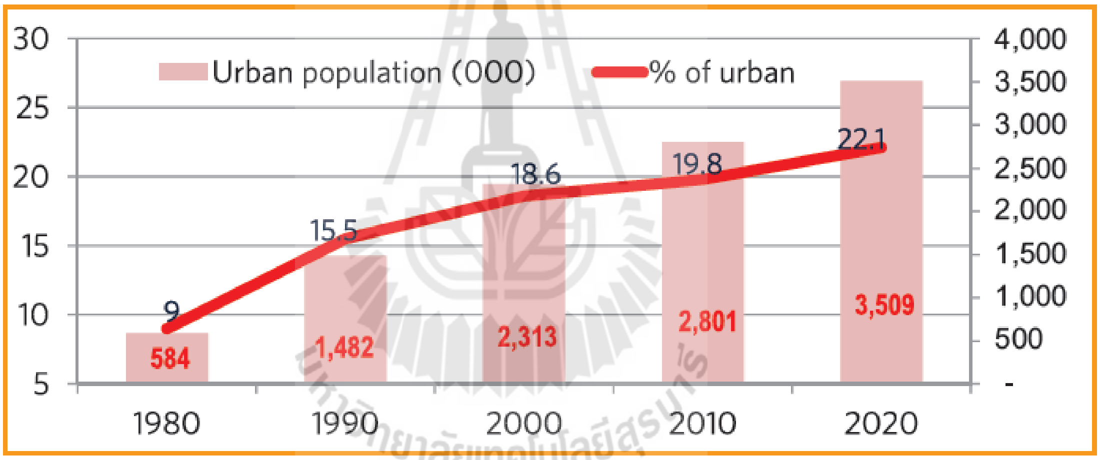
INTRODUCTION

1.1 Background and significance of the study

Global urbanization have become a major issue in urban land use planning, and as a result many researches in recent years have been conducted the study on urban growth modeling in order to achieve sustainable urban development (Achmad, Hasyim, Dahlan, and Aulia, 2015). There are many major problems related to the process of urbanization that results in reducing land use and land cover (LULC) to become impervious surface, and it is caused by the economic development and population growth (Shalaby and Tateishi, 2007). Bhatta (2010) stated that the impact of human population growth may result in negative consequence that generally lead to uncontrolled urban sprawl and serious environmental problems. This outcome of sprawl may result in losing precious land and resource, landscape alteration, environmental pollution, traffic congestion, infrastructure pressure, rising taxes, and neighborhood conflicts (Allen and Lu, 2003; Aguayo, Wiegand, Azocar, Wiegand, and Vega, 2007). Therefore, to reduce these negative impact of sprawl and to prevent the city to grow improperly, urban policy planning should be devised to control urban spatial expansion in sustainable way (Shu, Zhang, Li, Qu, and Chen, 2014; Achmad et al., 2015; Manu, Twumasi, Lu, and Coleman, 2015). In addition to this policy planning, driving factors behind urban spatial expansion may also consider as important, and it should be addressed and identified as a guideline for the policy maker to make a decision for sustainable urban growth in the future. Because it can help to improve the understanding of the spatial and temporal dynamic process of LULC change from the past to the future by analyzing the relationship between the driving factors and the historical urban growth pattern trend, so that prediction future urban pattern distribution could be identified and could be helpful to answer the influence of historical driving factors affecting the future urban growth pattern (Aguayo et al., 2007; Manu et al., 2015).

For the case of Cambodia which is still a developing country, but presently it has revealed as a fast urbanization country according to the United Nations Population Fund (UNFPA) Cambodia (2014). The report from National Institute of Statistics (NIS) (2013) also revealed that over the last two decades, urban population of Cambodia grew from 17.7% of total population in 1998 to 21.4% of total population in 2013. Annual urban growth rate of 2.2% between 1998 and 2008 increased to 3.7% between 2008 and 2013. These indicate the trend of increasing population is in large scale, and the trend will be expected to accelerate more in the next 5 and 15 years. As a result from UN Population Division revealed that urban population of Cambodia might reach 3.5 million or 22.1% of total population in 2020 (Figure 1.1), and population projection of UNFPA Cambodia (2014) also revealed that urban population will reach by 5.4 million in 2030 particularly in a major city of Phnom Penh. Meanwhile, beside the population growth, urban spatial expansion rate in Cambodia was found as the fastest rate about 4.3 percent per year in Asia after Lao PDR based on the report studied by World Bank Group (2015). This shows that Cambodia urbanization is facing an expansion in high rate, especially, in the Phnom

Penh capital city. Moreover, Phnom Penh has been encountered a problem with poor planning as well as the ineffective master plan and lack of official land use zoning, so that the city has been found many negative impacts (Doyle, 2012). For instance, the wide expansion of the urban sprawl of Phnom Penh into natural areas such as lakes surrounding the core existing urban area and agricultural lands in the suburban area. This shows that the ineffective master plan and lack of land use zoning accompanied with the fastest expansion rate in Phnom Penh may inevitably lead to haphazard sprawl that will give more environmental impact to the city in the future.



Source: UNFPA Cambodia (2014)

Figure 1.1 Cambodia urban trend and projection.

Concerning to the model of identifying the driving factors of urban growth, three types of them are rule-based simulation, system dynamic simulation and empirical estimation model (Hu and Lo, 2007). However, empirical estimation model has been proved as efficient model and has been widely used to model and interpret the relationship of driving factors that influence the city expansion around the world, especially logistic regression model (LR) which is one model among the empirical estimation model. Many researchers including Allen and Lu (2003), Aguayo et al. (2007), Hu and Lo (2007), Alsharif and Pradhan (2014), Shu et al. (2014), and Achmad et al. (2015) conducted the study using LR model, a statistical algorithm approach, incorporated with GIS and remote sensing techniques to identify and interpret the driving factors and predicting the future urban growth pattern. However, little attention has been paid to the modeling of future urban growth pattern in developing country that is facing high urban expansion rate, like Cambodia. Therefore, LR model is here chosen to identify the most influential factors that drive urban growth from the past to recent periods for predicting the future Phnom Penh urban growth pattern.

1.2 Research objectives

The specific objectives of the study are:

(1) To assess land cover status and its change in three different periods in 2002-2009, 2009-2015, and 2002-2015;

(2) To employ logistic regression model to discover the driving factors of Phnom Penh urban growth in three periods;

(3) To predict the future urban growth pattern of Phnom Penh in 2030 based on the integration of LR model and Markov Chain model.

1.3 Scope of the study

According to the three proposed objectives above, scope of the study can be briefly described as follows:

(1) Three Landsat images of 2002, 2009, and 2015 are classified land cover types, and their thematic maps are assessed accuracy by using topographic map of 2002, high resolution Google earth image of 2010 and 2015. Then land cover change evaluation of three different period of 2002-2009, 2009-2015, and 2002-2015 are implemented by using post comparison change detection algorithm.

(2) The influential factors behind land cover change due to the urban expansion are identified. The influential factors applied in LR model, usually called independent variable or predictor variables, are used as input variables while the output variable is resulted as dichotomous 1 and 0 that represents as urban and non-urban area in the study. The output map reflects the capability of growth along the continuous value of 1 as the highest one to 0 as the lowest one then validation with ROC is implemented by finding the goodness of fit between the simulation and actual map.

(3) After satisfactory validation, prediction of urban growth in 2030 is conducted using the optimum urban growth demand that is estimated by the future population growth rate, Markov Chain model and simple linear regression.

1.4 Limitation of the study

(1) Regarding the limitation of the available information in Cambodia, this study uses all available information and discovers influential factors which are appropriate to the nature of study area. However, land price that is the essential factor for urban expansion cannot be here applied due to unavailable data in 2009 and incomplete information in 2015.

(2) Since prediction ability of logistic regression cannot quantify amount of urban growth demand, the future urban growth area by year 2030 is here calculated based on change ratio between population growth and urbanized land. In addition, Markov Chain model and simple linear regression are also applied to estimate area of urban growth in 2030. These results are further applied to quantify of change for allocating urban distribution pattern and identified the best practice for urban growth study.

(3) In this study, the pattern of urban distribution is allocated only in horizontal direction.

1.5 Study area

Phnom Penh, Cambodia capital city and the largest urbanized area in Cambodia, is located on an embankment of the intersection between Tonle Sap and Mekong rivers, and it has been recognized as the main national economic and industrial center (Figure 1.2). The city is laid at latitude 11^{o} 33’ North and longitude 104^{o} 55’ East that covers total area of 678.5 sq. km. Phnom Penh provides home for inhabitant to settle down about 1.6 million of total population of 14.8 million (Cambodia Inter-Censal Population Survey (CIPS), 2013) and it is subdivided into nine administrative districts or Khan in Cambodian language. Recently, three new districts are created by splitting from the existing districts due to the increase of population density, so that Phnom Penh totally have 12 districts.

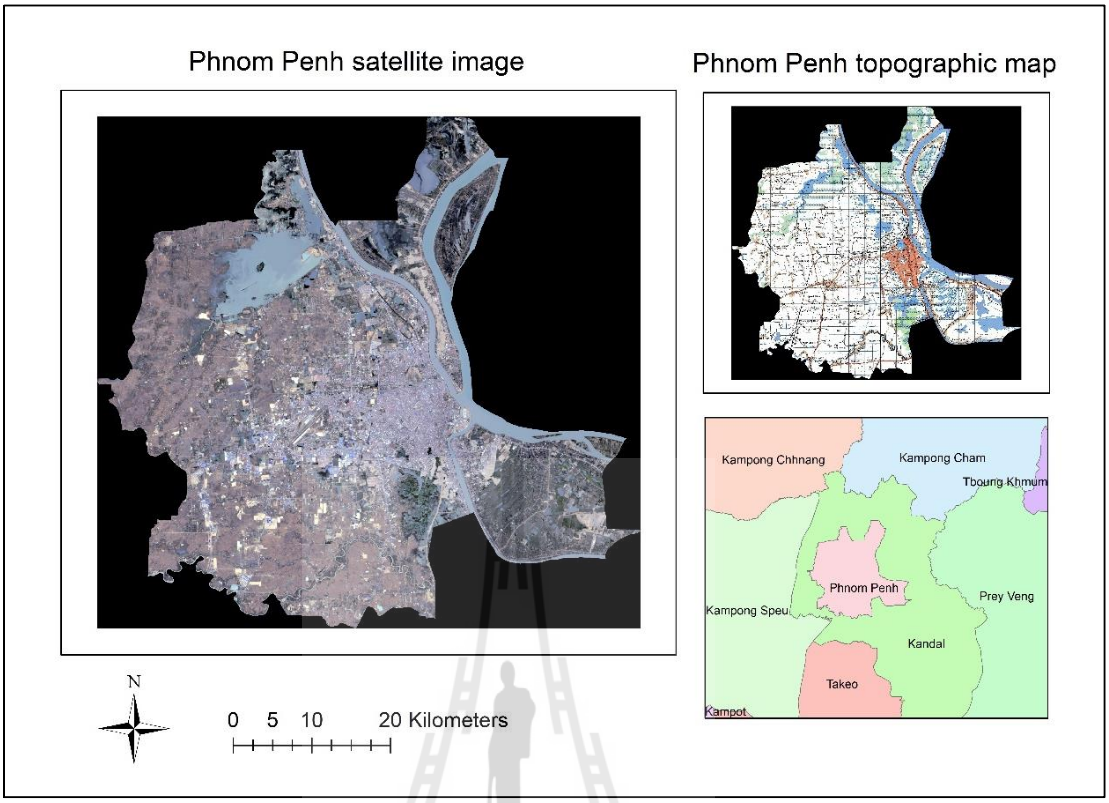


Figure 1.2 Map and location of Phnom Penh city, Cambodia.

1.6 Benefits of the study

The accomplishment of this research provides benefits as follows:

(1) The understanding of LULC status of Phnom Penh in 2002, 2009, and 2015 and its change in three periods of 2002-2009, 2009-2015, and 2002-2015;

(2) The identification of urban growth pattern of Phnom Penh in the study periods;

(3) The identification of influential factors for driving urban growth of Phnom Penh in short period of time (2002-2009 and 2009-2015) and long period of time (2002-2015);

(4) The achievement of the best practice for predicting urban growth of Phnom Penh in 2030 using allocation method based on an optimum logistics regression model;

(5) The obtained outputs are useful for governor, city planners, decision-makers and relevant researchers for their development activities, studies and planning that support sustainable urban growth and welfare of the community. In addition, the initial use of logistic regression model for urban growth prediction in Phnom Penh can be used as basic guideline and information to improve the future research in urban planning, management, and development in Phnom Penh study area.

CHAPTER II

BASIC CONCEPTS AND LITERATURE REVIEW

This chapter explains the concepts and theories, covering the three main topics related to the research: (1) basic concept of urban growth and sprawl, (2) urban growth model, (3) logistic regression model and urban growth studies.

2.1 Basic concepts of urban growth and sprawl

This section reviews and highlight some basic concepts of urban growth modelling including its cause and consequence of urban growth and sprawl, its influential factors, type of urban growth modeling, logistic regression model and its previous research study.

2.1.1 Cause of urban growth and sprawl

There are many factors that contribute toward urban growth. One of the most prominent is the rise of human population. As number of people in the urban area start growing over the capacity, the process of urbanization will continue to spread and expand step by step from the city center along the outskirt area. The result of growing is coming from the vast majority of rural immigration since the development of the cities grow very fast in term of economic, infrastructure and industrialization. Process of urbanization does not define only the cause of population growth, but it is also involved with the change in economic, social and political structure of a region. Thus, the effect of urbanization will impact to the environmental and social change in the urban environment and will also relate to global change issues. Moreover, through the rapid growth of urbanization, cities will encounter the problem of its infrastructure to provide the sustainable facilities and services such as energy, education, health care, transportation, sanitation and physical security (Bhatta, 2010). In addition to the migration to urban areas, it is often decomposed to two factors which are push and pull factors. Push factor refers to the circumstance of people moving out of the rural area whereas pull factor is the circumstance of people are attracted to live in the city. Generally, push factor occurs to the place where migrants confront an unfavorable condition in term of economic and security, for example, high unemployment and political persecution. Pull factor actually occurs when the outcome of push factor necessarily need to make a possible change of its condition, for example, flow of rural people to cities because of high opportunites of job employment and better living facilities.

Economic growth is also another factor for causing urban growth since the Gross domestic product (GDP) income of people are high and their great demand of living space is also big issue (Bhatta, 2010). So the rapid construction of new buildings and houses in the open space or suburb will be needed. This developed construction will lead to increase of highways, roads, bridges, and infrastructures that will cause sprawl along their physical factors.

Industrialization such as new economic zone, industrial and garment factories will consume a lot of land for development, so that some land use like agricutural field will be preferred and changed to built up area. For instance, the development of new economic zone will need factories, accomodations for labors, industrial parks, and parking lots (Bhatta, 2010).

Transportation includes roads and railways is used in many study as a factor for modeling urban sprawl (Yang and Lo, 2003). Roads and railways are important of modeling urban growth because of its usefulness for transportation facilities that will connect the urban city to its neighborhood. Construction plan of new road will be able to influence to the development of urban growth, real estate, location of preference house, land development and building density (Rui and Ban, 2011). Road width is also a main problem of causing sprawl because if the road width is too small, it is very difficult for vehicle to assess into the sites like high rise building and multistorey apartment located in the city center. Thus the narrow road may restrict the construction of high rise building resulting in waste of vertical space, so that the waste of vertical space will lead a demand of land and the growth will be in horizontal space (Bhatta, 2010).

Demand of more living space in many developing countries, residents of the core city lack sufficient living space. This encourages countryside development for more living space. People can buy more living space in the countryside than in the inner city since the cost of property is less in the countryside. However, consumption of more living space not always causes sprawl. Population density is a major concern in this issue. Cities in developing countries are three times denser than the cities in developed countries. Therefore, higher per capita consumption of built-up area (or living space) is desired in many instances. In such cases, higher per capita consumption of living space may indicate better and extended living facilities within the confine of compact urban growth. However, if the demand of more living space forces rapid low-density development in the countryside then it must be an indication of sprawl (Bhatta, 2010).

Due to physical geography, the sprawl is occurred because of unsuitable physical terrain such as rugged terrain, wetlands, mineral lands, or water bodies, etc. This often creates “leap-frog” development sprawl (Barnes, Morgan, Roberge, and Lowe, 2001).

One single family houses also cause the sprawl due to the consumption of large lot size. Most individual family houses consume large space rather than high rise buildings that waste alot of space area as non-developed (Bhatta, 2010).

Restrictive land-use policies in one political jurisdiction may lead development to “jump” to one that is favourably disposed toward development or is less able to prevent or control it (Barnes et al., 2001). Often dissimilarities in development regulations, land-use policies, and urban services among the neighbouring municipalities (or local governments) may cause discontinuous development (Bhatta, 2010).

Lack of consistent and well-experimented planning policies may also cause urban sprawl. A city may be planned with exclusive zoning policies; this means separation of residential, commercial, industrial, office, institutional, or other land uses. Completely separate zoning created isolated islands of each type of development. In most cases, the automobile had become a requirement for transportation between vast fields of residentially zoned housing and the separate commercial and office strips, creating issues of automobile dependency and more fossil fuel consumption and thereby pollution. A mixed land-use policy is preferred to fight against sprawl. Having a proper planning policy is not enough, rather its successful implementation and enforcement is more important. Unsuccessful enforcement of land-use plans is one of the reasons of sprawl in developing countries, since the enforcement is often corrupt and intermittent in these countries (Bhatta, 2010).

Credit and loan facility, low interest rate, etc. are also responsible for rapid urban growth in advance. In this sense, people can buy homes before achieving the financial capability. Therefore, the growth will occur in advance than actually supposed to be (Bhatta, 2010).

Residents of countryside are often former urbanites who desire the solitude and perceived amenities of country-living as rural retreats. Despite traffic congestion and long commutes to work, moving to the suburbs remains a goal for many city residents who perceive quality of life in the suburbs as better. Unless this perception changes and the conditions of urban life improve, sprawl development will continue as the flight from cities to suburbs continues (Barnes et al., 2001).

Urbanites purchase second homes in the countryside as future investments (Barnes et al., 2001). This encourages the developers for construction at the countryside in advance. These homes often left vacant but the government is forced to maintain urban facilities and services in a low-density area. Low interest rate and high housing demand make the countryside-housing investment more attractive.

2.1.2 Consequence of urban growth and sprawl

As urban populations grow more and more, people want privacy accomodation and the greenspace area for living. The demand of people will push the development of urban surface to expand further and end up as sprawl that may reduce the land cover and bring the impacts to the community if the consideration of new development and effective planning does not invovle. Thus, urban sprawl offers number of impacts as mentioned by Luther (2005). There are the ecological, social, and economic effects of urban sprawl.

Ecological effects of urban sprawl:

- Destruction of wildlife habitat;
- Introduction of non-native invasive plants and animals into natural areas;
- Increased human and pet exposure to diseases such as Rabies and Lyme disease;
- Increased risks of water pollution from oil and gasoline washing off paved surfaces and from pesticides, lawn fertilizers, and other chemicals;
- Increased potential for flooding and soil erosion due to impervious surfaces such as concrete or pavement;
- Decrease in groundwater for wells and irrigation caused by abundance of impervious surfaces; and
- Increased risk to life and property from wildfires.

Social and economic effects of urban sprawl:

- Increased community costs for maintaining roads, school bus routes, sewers, and other services needed when businesses and residences are spread out;
- Ongoing increases in property taxes to meet growing need for services, which may pressure rural landowners to sell to developers;

- Increased need for automobiles; increased noise, traffic, pollution; reduced potential for bicycling and walking;
- Isolation of the young, poor, and elderly who cannot drive or lack access to cars;
- Increased cost and difficulty of providing public transportation;
- Increased time needed for transportation and resultant reduction in time available to spend with family and friends or contributing to the community;
- Loss of agricultural and forestry jobs, and traditional land practices;
- Reduction of rural character or community sense of place;
- Increased ordinances that regulate logging, noise, or odors.

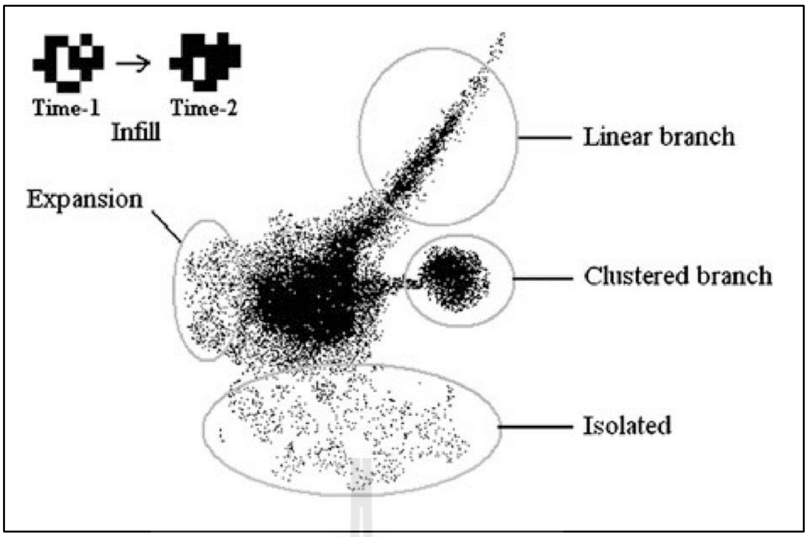
2.1.3 Type of urban growth pattern

According to Wilson, Hurd, Civco, Prisloe, and Amold (2003), urban growth refers to the transition change of landscape pattern from land cover to urban land use by the development of the residential area, commercial building, and industrial land use. Thus, studying on modeling urban growth must involve with the pattern of development rather than only study of the population density. Three categories of urban growth are infill, expansion and outlying, while outlying is separated into isolated, linear branch and clustered branch growth. In brief, to know the type of urban growth can help to find out how the existing developed areas grow and its relation to the factors of growing (See Figure 2.1).

Infill growth is characterized by non-developed pixels being converted to urban use and surrounded by at least 40% existing developed pixels. It can be defined as the development of a small tract of land mostly surrounded by urban land-cover.

Expansion growth is characterized by non-developed pixels being converted to developed and surrounded by no more than 40% existing developed pixels. This conversion represents and expansion of the existing urban patch.

Outlying growth is characterized by a change from non-developed to developed land cover occurring beyond existing developed area. Isolated growth is characterized by one or several non-developed pixels some distance from an existing developed area being developed. This class of growth is characteristic of a new house or similar construction surrounded by little or no developed land. Linear branch can be defined as an urban growth such as a new road, corridor, or a new linear development that is generally surrounded by non-developed land and is some distance from an existing developed land. A linear branh is different from isolated growth in that pixels that change to urban are connected in linear fashion. Clustered branch defines a new urban growth that is neither linear nor isolated, but instead of a cluster or a group. It is typical of a large, compact, and dense development.



Source: Bhatta (2010)

Figure 2.1 Urban growth type and pattern.

Identification of the three categories of urban growth, the expansion of its development will have a negative connotation through its growth characteristic that normally refer to sprawl. Planning, controlling and coordinating the outgrowth area is very important for preventing the sprawl pattern in order to avoid uncontroled diffusion and bad consequences of sprawl like inflated structure, loss of farmland and poor of infrastructure (Bhatta, 2010).

2.1.4 Influential factors of urban growth modeling

The significant impact of urban growth to the land cover change will consequently lead the future situation of a city become worse if the planning policy are not regulated appropriately to support urban expansion. The impact can cause urban sprawl that effects on social and economic, and it also abruptly causes the land cover change such as a major change of agricultural land and deforestation. In order to reduce the negative effects of urban expansion, it is necessary to understand the spatial and temporal dynamics pattern of land cover change process, especially, by finding a relationship of influential factors that drive the mechanisms of urban land development. According to the previous researches of Cheng and Masser (2003), Lui (2005), Braimoh and Onishi (2007), Hu and Lo (2007), Dewan and Yamaguchi, (2009), Thapa and Murayama (2010), Eyoh, Olayinka, Okwuashi, Isong, and Udoudo (2012), Duwal (2013), Shu et al. (2014), and Achmad et al. (2015), there are many driving factors used in their study that can be separately categorized in many types as shown in Table 2.1 with their descriptions as below:

- Bio-physical factor is a fundamental determinant factor of urban land expansion, involves with natural environment, includes topography, soils, rivers, wetlands and other ecological sensitive areas like agricultural land, grassland, and types of land served important ecological service functions (Dewan and Yamaguchi, 2009; Thapa and Murayama, 2010; and Shu et al., 2014).

- Economic factor refers to the economic activities that involve with job opportunity, business trade and industrialization. These economic activity also contribute a high attraction of rural immigrants that result will lead to a higher rate of urban expansion. In this context of economic factor, GDP and land value, is appropriately use as driving factor to identify the existing business center or industrialization for people or amount of rural immigrant working or living in that exact location of that area (Braimoh and Onishi, 2007; Dewan and Yamaguchi, 2009; and Duwal, 2013).

- Social factor including population density shows a high rate of urbanization due to a great demand of new settlement and open space area. It also concerns with race, per capita income, employment rate and housing rent (Hu and Lo, 2007; and Dewan and Yamaguchi, 2009).

- Accessibility and proximity including distance to major roads, railway, airport, seaport, urban, university, central business district (CBD), etc. plays very important role in illustrating future urban growth pattern since its potential benefit provides a good opportunity for economic development (Cheng and Masser, 2003; Hu and Lo, 2007; Eyoh et al., 2012; Shu et al., 2014; and Achmad et al., 2015).

- Planning and policy refers to land use zoning, land protection, regional economic planning and urban transportation planning is main driving factor that guides the direction of urban expansion in the future (Lui, Zhan, and Deng, 2005; Braimoh and Onishi, 2007; Thapa and Murayama, 2010; and Shu et al., 2014).

- Spatial interaction (number of urban cells within a neighborhood) can be used to explain the land use conversions, especially in the context of urban growth because the probability of new urban is highly occurred at the fringe of the existing urban area. For example, the existing ecological, cultural, water and transport networks are highly influence for new urban development (Hu and Lo, 2007; Duwal, 2013; and Alsharif et al., 2013). It is useful to include neighborhood interactions as a driving factor (Verburg, Nijs, Eck, Visser, and Jong, 2003).

Table 2.1 List of driving factors of urban growth.

Factor	Cheng and Masser (2003)	Hu and Lo (2007)	Braimoh and Onishi (2007)	Dewan and Yamaguchi (2009)	Shu et al., (2014)	Achmad et al. (2015)
Bio-physical factor	Distance to river	Slope percentage, Bare land, Crop land, Forest	Slope and elevation, distance from water	Slope and elevation	Ecological sensitivity, and distance to main river ways.	Distance to green open space area, historical area, river, Tsunami prone area
Economic factor		Per capita income, Employment rate, Poverty rate, median housing rent	Income potential	GDP		
Social factor		Population density	Population potential surface	Population		Population density
Accessibility and proximity	Distance to railway, roads, bridges and distance to socioeconomic center such as CBD, industrial center.	Distance to CBD, Distance to active economic centers, distance to nearest urban cluster	Distance to roads, commercial center, harbor and airport		Distance to road, town centers, exits of thruways, ports and docks	Distance to CBD and road
Planning and policy	Master plan (urban and non-urban)	Conservation area	Distance from protected forest and distance from public supplied piped water.		Prime cropland protection	
Neighborhood factor	Neighboring waters, developed area, and industrial area	Number of urban cell within 7 x 7 pixel window	Proportion of urban, forest, and industry		Proportion of urban land within 7x7 pixel window	

2.2 Urban growth models

Modeling urban growth tends to study about the land use system of its connection to the land cover change which are caused by human activities. Rapid land use change can be a distressful feeling to the city planners, geographers and relevant decision makers to figure out, so that the usefulness of spatial land use change is very helpful for them to understand the processes of urbanization in order to find out the possible development in the future under different scenarios (Vermeiren, Anton, Maarten, Eria, and Paul, 2012). Over the past decades, serveral type of modeling methods have been used to study in LULC changes in order to monitor land use changes, evaluate planning strategy and plan for new development (Yeh, 1999). Three types of them are rule-based simulation models, system dynamic simulation models and empirical estimation models that will be used to model land use change (Hu and Lo, 2007; and Duwal, 2013). Among three of them, only empirical simulation model has been used efficiently to study and understand the driving forces of urban growth.

According to Hu and Lo (2007), rule based simulation model likes cellular automata (CA), that shows a great capability to handle temporal dynamics, will attempt to analyze land use land cover pattern based on state, cell neighborhoods and transition rules. Since CA focuses on simulation of spatial pattern, it is difficult to interpret the change affected by social and economic factors.

System dynamic simulation model arranges and describes the complicated connections among each element in different levels, and also deals with dynamic processes with feedback in a system. This model can be used to couple to cellular automata for including spatial dynamic (He, Okada, Zhang, Shi, and Zhang, 2006);

however, its ability of identifying the factors is weak and inefficient (Guo, Liu, Huang, Fuller, Zou, and Yin, 2001; and Liu, Li, and Yeh, 2006)

On the other hand, empirical simulation model proves as an efficiency technique to figure out the drivers of causing land use changes and gives a good result of understanding of what causing the urbanization process. For example, logistic regression model uses statistics to find out the relationship between influential factors and land use change base on historical data (Hu and Lo, 2007), so that it is successful in interpreting socio-economic activities (Lopez, Bocco, Mendoza, and Duhau, 2001). Cheng and Masser (2003) defined the empirical model as spatial statistic model that consists of Markov Chain analysis, multiple regression analysis, principal component analysis (PCA), factor analysis (FA) and LR.

Since the majority of techniques of empirical simulation model are efficiently used to model the urban growth driving forces, these models are briefly described in Table 2.2 as follows:

1. Markov Chain analysis (Lopez et al., 2001): Markov Chain represents a dynamic system of special classes involving transition probabilities and areas, described as symmetric matrices (Luenberger, 1979; Logofet and Lesnaya, 2000). The transition probability matrix is a text file that records the probability that each land cover category will change to every other category (IDRISI, 2012).

2. Analytic hierarchy process (AHP) (Thapa and Murayama, 2010): It is best known as a comparative rating of various important factors that get involved by many decision makers. Its technique is based on pairwise comparison of elements (attributes or alternatives) and results in a comparison matrix in which the relative importance of each element or factor is determined numerically.

3. Multiple linear regression (Shu et al., 2014): The concept is similar to linear regression which is an approach of modelling the relationship between dependent variables and independent variables, but it has more than one independent varaible.

4. Principal component analysis (PCA) and factor analysis (FA): PCA is a statistical technique that linearly transforms an original set of variables into a substantially smaller set of uncorrelated new variables that represent most of the information of the original data set. Through this technique, it obtains a small set of uncorrelated variables, which is much easier to understand and uses in further analysis than a larger set of correlated variables (Lewis-Beck, 1994). Meanwhile, FA purpose is to identify and describe the latent factors within various variables. According to the review of many urban growth research papers conducted using PCA and FA, these techniques are not widely used to find out the important variables related to the factors that influence urban growth, and they are also difficult to reflect the spatial pattern information of urban growth. Thus PCA and FA are not comprehended the strength and weakness in this context.

5. Logistic regression (Hu and Lo, 2007; and Shu et al., 2014): Logistic regression analysis allows the finder to predict the value of dependent variable from the information of independent variables. It performs as dichotomous (0 and 1) while independent variables can be continuous and category.

Table 2.2 Strength and weakness of urban growth models under empirical simulation model.

Model	Strength	Weakness
Markov Chain (Lopez et al., 2001)	- Proves as useful to describe quantitatively the tendencies of LULC change by providing the transition area matrix.	- Prediction capability of LULC change is not strong.
Analytic hierarchy process (AHP) (Thapa and Murayama, 2010; and Shu et al., 2014)	- The evaluation is contributed by the participation like local residents, researchers and urban planners. Their expression may reflect their behavior interaction, their knowledge on environment based on their ideas and what they have observed in reality.	- The method is subjective since the model is based on expert's perception in order to find the relationship between factors and urban growth. - It does not provide the spatial pattern that reflects the urban growth.
Multiple linear regression (Menard, 2002; and Shu et al., 2014)	- Efficiently use to test the relationship of multiple influential factors and urban change area, as influential factors have equal contribution to urban growth.	- It is not applicable when dependent variables are categorical variables or variable with spatial attributes.
Logistic regression (Hu and Lo, 2007; Arsajani, Helbich, Kainz, and Boloorani, 2011; and Shu et al., 2014)	- LR focuses more on interpretation that provides an understanding between driving factors and urban growth. - It provides the probability of predicted map based on the influenced driving factors. It can bring the answer of urban growth pattern direction.	- It has weakness in quantification of change and temporal analysis. - Its ability of prediction is still limit that can not quantify amount of urban growth area in the future.

According to the description of strength and weakness of urban growth models in Table 2.2, LR seems to be more successful in urban growth modelling, since it brings the answer of interpreting the influential factors and their relationship with urban growth. Furthermore, outcome of probability map may give the direction information of future urban growth pattern despite of some limitations.

2.3 Logistic regression model and urban growth studies

In general, there are two types of logistic regression model that are multinomial logistic regression and binary logistic regression. Multinomial logistic regression model is used generally for the dependent variable that will be comprised more than two unordered categories. It is used for obtaining the set number of *m* categories of dependent variables as an example of land use type in land cover change analysis (Hao, Zhang, Li, Yao, Huang, and Meng, 2014). Binary logistic regression is a non-linear statistical method of regression analysis for binary depedent variables. Its output is measured with a dichotomous variable that means the two possible outcomes contain only data coded of 1 (Urban) and 0 (Non-urban). The goal of binary logistic regression model using in urban growth modeling is to describe the relationship between the two possible outputs of dependent variable (urban and non-urban) and a set of independent variables (predictor variables or driving factors of urban growth). This logistic regression will generate the coefficient of the predicted variables that describe the significance of each variable, the positive and negative influences, and the probability of the presence of characteristic of dependent variable (Arsanjani et al., 2013). Compared with the multinomial logistic model, binary logistic model is extremely flexible and easily applicable (Hosmer and Lemeshow, 1989).

2.4 Literature review

Binary LR model has been used to study for modeling urban growth since last decade and still contributes its work importantly at present. Obviously, many researchers have published about urban growth study from various study sites that can be here summarized as a literature review in this study.

Achmad et al. (2015) conducted a study on modeling of urban growth in Tsunami-prone city, Banda Aceh, Indonesia using LR based on LULC maps of 2005 and 2009. The aim of the study is to provide a vital information about its land cover change patterns and the driving factors of its rapid urbanization, since urban development of Banda Aceh surged up very fast after hitting by Tsunami in 2004. Land cover change detection was used to examine the process and the quantity of change of urban growth by classifying each image using supervised classification. Two SPOT images were used to classify four land categories, namely built-up area, vegetation, water body and wet land. According to LR model, the input of dependent variable must be dichotomous, so that each land use class of both image were reclassed as built-up and non built-up (vegetation and wet land) areas. The author selected seven independent variables which were two socio-economic (population density and distance to CBD and five biophysical (distances to green open space, historical area, river, highway and coastal area) as urban growth driving factors. The result showed that the city had changed drastically during 2005-2009 in particular, built-up areas that increased of 90.8% (1,016 ha). Socio-economic factors showed a positive influence of urban growth whereas the biophysical factors showed the negative influence. Socio-economic factors, namely population density and distance to CBD have a positive influence which mean the probability of growth is in the distance to CBD because Banda Aceh is not a large city based on its size and population; therefore, it's easy to access to each area. The distance to Tsunami prone area showed the positive effect which means the further the city is from the coast, the greater chance for the urban to grow. Furthermore, in term of biophysical factors that gave a negative effect such as green open space and distance to historical area reflecting the demand of people to live near green open space are high due to its good environment, recreation and beauty. Historical area is very important since it has been established as a heritage city. Distance to road and river gave a higher chance of urban growth that mean area closer to road and river have a likelihood of urban growth. The study also predicted the future growth of the city in 2029 by using the combination of Markov transition area matrix based on LULC data between 2005 and 2009 with the LR probability map to find the future growth map. Author gave a conclusion that the influence of the factors driving the growth of the city is important in drafting the regulation that involves spatial planning of area for future development. He also suggested to consider disaster mitigation to make public policies in order to achieve sustainable city.

Alsharif et al. (2013) applied the LR model with remote sensing data to study urban sprawl of Tripoli metropolitan area, Libya. In this research, the author aimed to predict the growth of Tripoli urban expansion in different scenarios. Because Tripoli, which is the political center of the state, the commercial center, industrial and the service capital of Libya, had a rapid haphazard growth over the past decade. The political unrest and corruption of the city may have effected the urban planning that lead to massive urban sprawl. Logistic regression with the eleven factors were used to investigate the influences of each factor to sprawl. The eleven factors are the distance to main active economic centers, the distance to a CBD, the distance to the nearest urbanized area, the distance to educational area, the distance to roads, and the distance to urbanized areas, easting and northing coordinates, slope, restricted area, and population density. All of the each factor were computed in the calibration phase using data from 1984 to 2002 and 2002 to 2010 were used in the validation by relative operating characteristic (ROC) method. After validation, the probability map was generated to estimate three scenarios of urban pattern for 2020 and 2025 by using the urbanization pattern index (Campbell et al., 2007). Result revealed that the population density factor has the highest odds ratio, which means the population increase will result in a spread pattern of urban expansion. The ROC result of 0.86 shows the high accuracy of probability assignment that could be used for prediction of urban expansion. The spatial pattern prediction resulted from growth ratio assumption of 5, 6, and 7 in the formular of Campbell et al. (2007) were allocated the estimated size of the future urban area to the urban probability map of LR model. The allocation starts from the highest probability cells to the lowest until the total area was equal to the estimated future area. Result of the three scenarios were concluded that the Tripoli situation might be worsen in the future due to the increase of land consumption so that effective planning and strict regulation would be implemented for sustainable urban development.

Shu et al. (2014) also used LR model to analyze the driving forces of urban land spatial expansion in the three port towns of Taicang city, China during 1989-2008. Unlike from the previous researches that used only socio-economic and biophysical factors, this research applied more driving forces of urban land spatial expansion that include natural eco-environment, accessibility, socio-economic, neighborhood factors, relevant planning and policies. However, the author denied to use socio-economic factors like population and GDP in his study since their spatial resolution could not achieve the requirements of LR and based on Li, Zhou, and Ouyang (2013) the missing of them would not effect the model performance. The factors of natural eco-environment are slope, distance to river bank, to Yangtze river, to reservoir shoreline, to fault zone, and to ecologically sensitive areas and land use status. Factors of land control policies consist of prime cropland which is the arable land that was protected according to the demand of agricultural product for social economy; therefore, its parameter was set as 0 and 1 for non-cropland and cropland. Accessibility factors consist of distance to town centers, to main road, to ports, to main river ways, and to exits of thruway. Neighborhood is defined as a proportion of urban land area within at 7x7 pixels window for each pixel grid with 30x30 m resolution. Result after applied the LR model were ranked in order to find and interpret the most influential factor of each port town in different period of 1989-1995, 1995-2001 and 2001-2008 with all ROC of each period are more than 0.8. Result revealed that the driving factors exist during different peroid in different regions and relative factors of importance also varied with time and space. Among the four groups of factor, the finding showed accessibility was a great dominant. Author suggested to formulate the policies to guide reasonable expansion of urban land.

Eyoh et al. (2012) conducted a research for modeling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using LR and GIS. The research used three date images which were 1984 and 2000 for calibration model and 2005 for validation model using Kappa statistic. After validation, the satisfactory result obtained from 2005 was used to predict urban expansion for 2030 based on the 1984 and 2000 calibration. This research used only 10 accessibility factors that included distance to water, to less dense urban, to dense urban, to major roads, to railway, to Lagos Island, to Lagos international airport, to Lagos seaport, to University of Lagos, and to Lagos state university. Result showed that all ten variables were significant at 95% based on yields p-value <0.05. This study showed a substantial agreement between the simulated map and reference map in 2000 that was proved by the Kappa statistic of 0.764 so that the calibration of 1984 and 2000 was used again to predict for 2005 map. The 2005 simulated map with the 2005 reference data were estimated by Kappa statistic that result of 0.6998. Then a prediction of 2030 was derived upon the satisfactory result of 2005 prediction. Result was concluded that distance to urban center had the highest effect that indicates urban growth to occur close to the nearest urban.

Vermeiren et al. (2012) conducted a research of urban growth of Kampala, a capital city of Uganda that experienced a fast growth rates of 5.6%. This study analyzed the pattern based on the historical Landsat images of 1989, 1995, 2003, and 2010 and to develop the future pattern of urban growth in 2030 based on the three alternative scenarios: a business as usual, restrictive, and stimulative scenarios. LR model was developed for calibration with the predicting variables included distance to road, distance to city center, topographic setting, slope gradient and built up potential. In this study, road was digitized from Google Earth, slope was derived from a 30 m resolution of ASTER DEM, so that topographic factor was obatained as well, built up potential is a relative measured of accessibility to the existing built up zones and was calculated by the gravity-based formula of Hansen (1959). Land cover change map was studied by a supervised classification and overlay method. The result revealed the correlation between built-up area and the distance to the main roads in all time periods. In addition, this research was also described the finding of quantity of total land during 1989 to 2010 and how it increases and changes of amount of built-up area until 2010 in term of its situation with each predictor variable. As in this case, result revealed that the total built-up area increased from 71 sq. km to 386 sq. km between

1989 and 2010 in Kampala captial city. Herewith 62 sq. km represented 87% of built-up area in 1989 and 255 sq. km respresented 66% of built up area in 2010 were situated within 500 m of the main road. Moreover, result of 35 sq. km built-up area in 1989 was situated within 5 km of the city center whereas 80 sq. km of built-up area in 2010 decreased from 49% to 21%. The calculated ROC value at 82% resulted in reliable predictions of the future urban expansion. The author used three scenarios to generate urban expansion of 2020 and 2030 by assessing the necessary area for future built-up area and selecting the pixels with the highest probability value of the derived LR model until the required area is met. The first scenario is the business scenario is based on extrapolation of the observed exponential population growth while assuming constant popultion density. The second scenario is restrictive scenario that have same amount of urban growth area but zoning laws prohibit new built-up area in wetlands and in a predefined list of open zones in the city center to assure the environmental urban life quality. The third scenario is stimulative scenario using the same constraint as the restrictive scenario but promotes development through the attractive center for new built-up area and high rise zone in the CBD. The result was concluded that the alternative policy options result in contrasting future urban sprawl patterns with a significant impact on the local quality of life.

Hu and Lo (2007) conducted a LR to model urban growth in Atlanta metropolitan area of Georgia where the region facing the risen of premier commercial, industrial, and transportation urban area of the southeastern United States and the rate of fast growing of human population. As a result, some tremendous land cover had changed and the urbanization had consumed vast acreages of forest and agricultural land to the city proper that pused the rural urban fringe farther and farther away from the original Atlanta urban core. The aim of modeling was to discover the relationship between urban growth and social, econometric and biophysical factors by using the historical LULC data in 1987 and 1997 from Landsat TM images. Authors used a historical review of urban growth in Atlanta to select the social predictors variable included population, race, income, employment, housing, as a continuous data and the biophysical predictors like proximity of major road, economic activity, existing land use status, and land conservation. The prediction of future urban growth was taken by the coefficient value of the predictor variables to estimate the logistic regression equation, so that the output probability image can be produced. Model was validated by the ROC method, a statistical analysis that compares the result of simulation image with the actual image to ensure the agreement of two maps being urbanized. The result revealed that two groups of factor were found to affect urban growth. The first group with odd ratios < 1 included population density, distances to nearest urban clusters, activity centers and roads, and high/low density urban uses and another group with odds ratios > 1 consisted of distance to the CBD, number of urban cell within a 7x7 window, bare land, crop/grass land, forest and UTM northing coordinate. ROC value of 0.85 indicated the probability map was valid. Authors also concluded that LR is very important model and result in deep understanding of the driving force of urban growth and the information of urban spatial pattern.

CHAPTER III

RESEARCH METHODOLOGY

Urban growth modeling using LR is the main important task to identify an associated influence of urban growth with the biophysical, social, and economic and its spatial interaction driving factors. Moreover, it is more helpful for not only use to define the influenced drivers, but also generates the urban expansion probability map that is able to indicate the direction of growth from the highest probability to the lowest probability in the future through their most influenced drivers. Thus, the research methodology framework using LR is developed and is displayed in Figure 3.1.

The main research consists of five components, namely: (1) data collection and description, (2) data preparation, (3) urban growth simulation by logistic regression model (4) urban growth validation, and (5) urban growth prediction. The details of five components of the research methodology are separately described in the following sections.

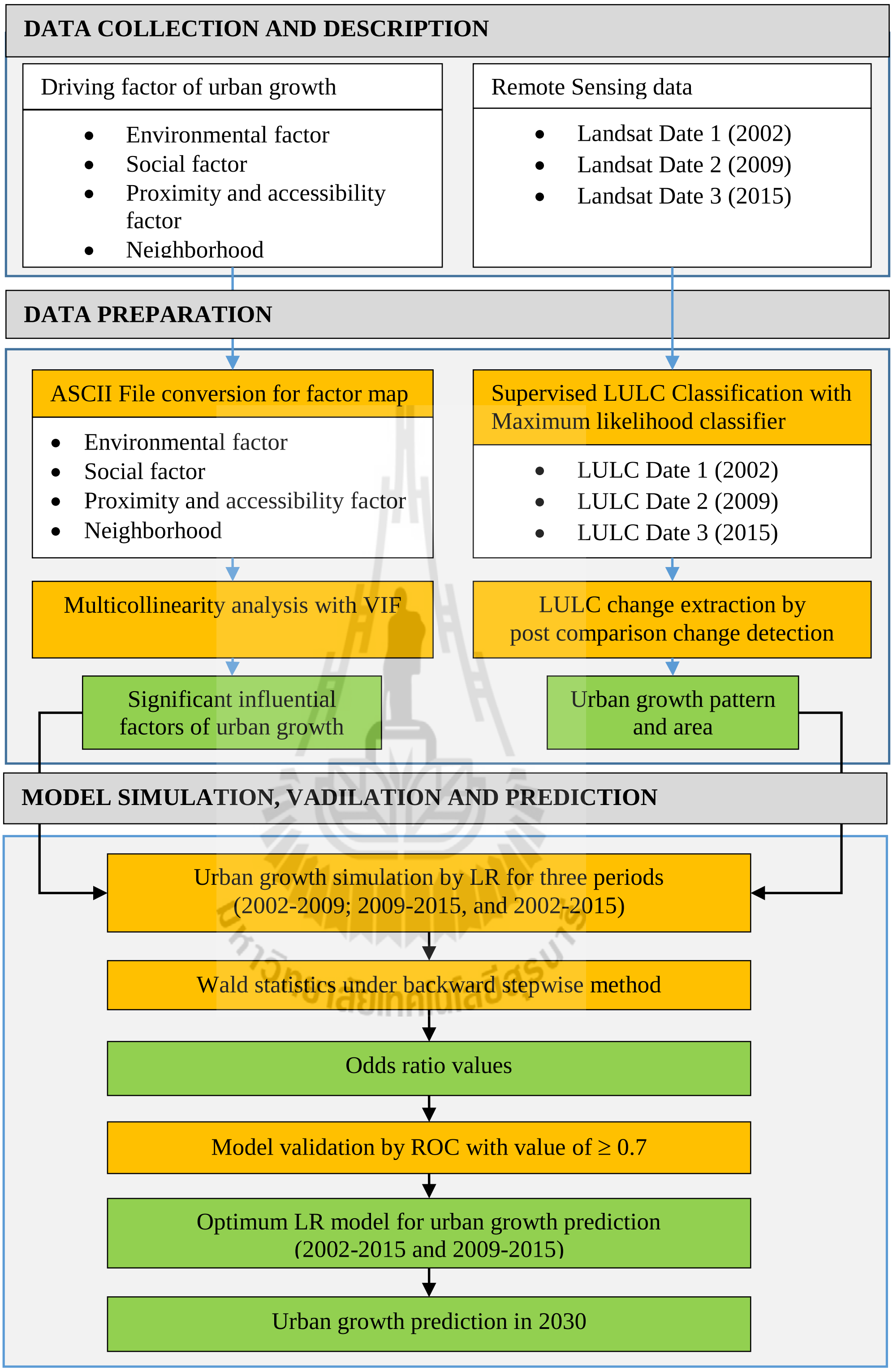


Figure 3.1 Schematic workflow of research methodology framework.

3.1 Data collection and description

Two major tasks of this step are about the compilation of available data used for modeling urban growth. There are two types of data include remotely sensed data (Table 3.1) and influential factors (Table 3.2) are here collected. The collected images and maps are presented in Appendix A. The main reason of selecting three remotely sensed data is related to availability of influential factors data. For year 2002, population data and topographic map that consists of land use and public service information are available. Meanwhile, the recent remote sensing data in 2015 was selected in order to update land cover and public service data from the existing websites and application like openstreetmap.org and Google Earth. For year 2009, it is chosen by selecting the equal interval time between 2002 and 2015 and since the annual population growth rate in Phnom Penh starts rising larger in 2008 as mentioned earlier in Chapter I. So the research can be conducted and applied urban growth modeling and to find out the result in three different periods: 2002-2009, 2009-2015, and 2002-2015.

Table 3.1 Basic information of remotely sensed data.

No	Path	Row	Resolution	Source	Acquisition date
1 Landsat 7 ETM+	126	52	30m	USGS	January 3rd, 2002
2 Landsat 5 TM	126	52	30m	USGS	January 14th, 2009
3 Landsat 8 OLI	126	52	30m	USGS	January 15th, 2015
4 Google earth image			1.6m	DigitalGlobe	Febuaury 3rd, 2003
5 Google earth image			1.6m	DigitalGlobe	January 7th, 2010
6 Google earth image			1.5m	Astrium	October 31st, 2015

Table 3.2 Influential factors of urban growth dataset.

No	Data	Year	Scale	Source	Description
1	Administrative boundary	2011	1:125,000	Economic census of Cambodia in 2011, NIS (2013)	Phnom Penh boundary, District and Commune.
2	Road network	2002	1:100,000	Cambodia National Geographic Department (NGD, 2002)	National main road and minor road from topographic map.
		2010 and 2015		Google earth/ high resolution images	Update main road and minor road
3	Water bodies	2002	1:100,000	Cambodia (NGD, 2002)	Water bodies consist of Tonle Sap and Mekong rivers, lakes, and streams.
		2009 and 2015		Landsat 5 and 8 images	Update water bodies in the study area
4	Public service	2002	1:100,000	Cambodia (NGD, 2002)	Public school, health center, and airport.
		2015		Openstreetmap.org	Update public school, health center, and airport.
5	Land use and land cover	2002	1:100,000	Cambodia (NGD, 2002)	Urban and built up area, vegetation, water bodies and miscellaneous land
		2009 and 2015		Landsat 5 and 8 images	Update land cover types
6	Population data	1998 and 2008		General population census 1998 and 2008, (NIS 2002; NIS, 2009)	Population density map
7	Industrial area	2002		White Book of Development and Management of Phnom Penh, 2020 (BAU, 2007)	Location of industrial area in 2002
		2009 and 2015		Google earth image interpretation	Update location of industrial area in 2009 and 2015

3.2 Data preparation

Two major activities of data preparation are urban growth influential factors extraction and LULC classification.

3.2.1 Urban growth driving factor extraction

The collected urban growth influential factor as maps or statistics data are firstly converted to ASCII files and they then used to analyze multicollinearity under IBM SPSS software. In principle, multicollinearity analysis is very important for logistic regression modeling in order to prevent the correlation of independent variables. Multicollinearity happens when the correlation among the predictor variables have perfect linear relationship, therefore the estimation of the model coefficients can not be possibly computed (Odzemir, 2011). In some case, if multicollinearity between two variables are very high, but not perfectly correlated, the model regression coefficients become more sensitive to individual predictor that can cause result of model coefficient to appear insignificant (Rogerson, 2010). According to Mernard (2002), tolerance (TOL) and variance inflation factor (VIF) are widely used to detect multicollinearity. He recommended that if TOL value is less than 0.2, it seems to occur multicollinearity and becomes more serious when TOL value is smaller than 0.1. TOL is the amount of variance in an independent variable and is not explained by the other independent variable (Rogerson, 2010). The TOL value can be calculated using the following equation:

$$TOL = 1- r^2 \qquad (3.1),$$

Meanwhile, VIF value is a reverse of TOL value calculated by 1/TOL. In the case that TOL equals or less than 0.1 will lead a reciprocal VIF value results to be equal or over

10, so that multicollinearity will occur (Rogerson, 2010). The VIF value can be calculated using following equation:

$$VIF = \frac{1}{1 - r^2} \tag{3.2}$$

where r^2 is associated with the regression of the independent variable on all other independent variables. In order to avoid multicollinearity, each variable that appears high value of *VIF* equal or more than 10 should be excluded. The process must repeat again and again until the value of *VIF* smaller than 10 which is acceptable.

3.2.2 Land cover supervised classification

In this study, maximum likelihood classifier is one of the most common method of image classification that have been widely used in supervised classification algorithm (Jensen, 2005) is applied for land cover extraction. Theoretically, decision rule of maximum likelihood is based on probability that a pixel belonging to each of predefined set of *m* classes is calculated, and the pixel is then assigned to the class for which the probability is the highest (Jensen, 2005) as:

$$p(X \mid w_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |V_i|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(X - M_i)^T V_i^{-1}(X - M_i)\right] \tag{3.3}$$

where: $p(X \mid w_i)$ is the probability density function of class w_i

$|V_i|$ is the determinant of the covariance matrix

$|V_i^{-1}|$ is the inverse of the covariance matrix

$(X - M_i)$ is the transpose of the vector $(X - M_i)$

The mean vectors (M_i) and covariance matrix (V_i) for each class are estimated from the training data.

The land cover classes of 2002, 2009, and 2015 are separately classified correspond to remotely sensed data. In this study land cover classification scheme was modified based on land cover classification system in topographic map of Cambodia NGD and it consists of (1) urban and built-up area, (2) arable land (3) vegetation, (4) water body and (5) miscellaneous land (see description in Table 3.3 and Figure 3.2). In addition, each classified land cover map of the recent year and the historical year is evaluated its accuracy (overall accuracy, producer's accuracy, user's accuracy, and Kappa hat coefficient) using the existing high spatial resolution of Google map of year 2010 and 2015. For land cover map accuracy of year 2002, it is verified by Cambodia NGD topographic map, scale 1:100,000 (see in Appendix A).

Table 3.3 Land cover classification scheme.

Class	Land cover classes	Description
1	Urban and built-up areas	It consists of village, town, residential, commercial, and industrial area, communication and utility center and transportation infrastructure.
2	Arable land	It consists of paddy field, abandoned paddy field, and open land.
3	Vegetation	It consists of crop, orchard, grassland, shrub land, flooded grassland, flooded shrub land, gardens, parks, garden village, and forest.
4	Water body	It consists of river, stream, pond and lake.
5	Miscellaneous land	It consists of bare land, marsh and swamp, sand bank, excavation site, and land fills

(a) **(b)**

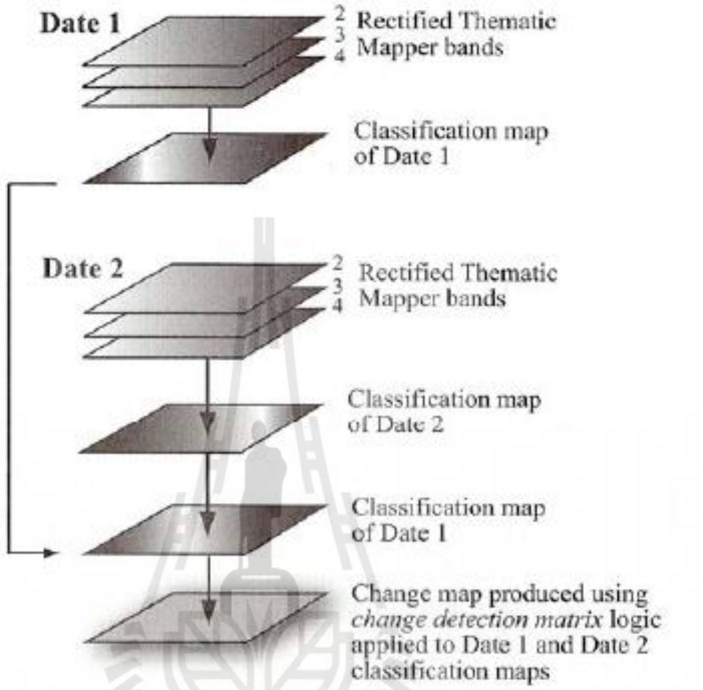
(c) **(d)**

Source: Google street view**.**

Figure 3.2 Land cover characteristic of Phnom Penh: (a) urban and built-up areas, (b) vegetation, (c) water body, and (d) arable land and miscellaneous land.

3.2.3 Land cover change detection

Three periods of land cover change: 2002-2009, 2009-2015, and 2002-2015 based on the classified land cover data from three different dates are extracted using post classification comparison change detection algorithm (Figure 3.3). This method provides a simple operation and easy to understand but its process relies heavily on the quantitative change of pixel of two comparison images using a change detection matrix (Jensen, 2005). It also provides from-to change map of each class information, so that the observation shall be convenient to notify the quantity of change from one class to other classes. The results of land cover change detection are further used to extract urban growth area, its pattern and to reveal the ratio of urbanized land change.



Source: Jensen (2005)

Figure 3.3 Algorithm of post classification comparison change detection.

3.2.4 Input data for logistic regression model

For running the simulation of LR model, input model data are prepared in ESRI ArcGIS 10.2 for three different dates of 2002, 2009, and 2015. Influential factor data of 2002 is extracted from topographic map that contains information of land cover class, public service location such as pubilc school and health center location, main and minor road, railway and airport. For data in 2009 and 2015, some variables remain the same and some vary through the time according to the city development. Therefore, some influential factors of 2002 are improved and brought new information up to date for data in 2009 and 2015. However, this research has limitation of some data availibility of 2009 like location of school and health center, that might affect to the modeling when the influential factors of 2009 are applied. The summary of each influential factor characteristic and its limitation is described as below.

Since land cover classification map's resolution size is 30 m, all factor maps are also produced the same resolution size of 30 m and then each factor map which is continous data are normalized into the range of 0-1 in order to get each spatial value distributed within the same scale.

Environmental factor consists of land cover classes like vegetation, arable land and miscellaneous land of each date image are used to determine the probability of transition of land cover change to urban as suggested by the previous study of Hu and Lo (2007) and Shu et al. (2014). Each land cover class is extracted from the Landsat image of 2002, 2009, and 2015 after classification. Distance to river, as review from literatures, is used in this research because of its major important for supplement to the people living in the village in the past, so that urban might grow along the river or extend further from the river. It was made by Euclidean distance tool in ESRI ArcGIS 10.2 and it remains constant for all years.

Population density (people per sq. km) is done by considering the amount of people per district area into district vector file. Then the vector file is rasterized by using the value of density estimation in each district vector file. Population census data is available for two dates only, 1998 and 2008 because Cambodia start conducting the census survey in 1998 and will redo every 10 years, so that the next census survey will be done in 2018. The missing population data at district level in 2002 and 2015 are estimated based on the derived annual growth rate of Phnom Penh between 1998 and 2013.

Proximity and accessibility factors are the distance of each pixel of image to the location of target point, line or polygon of each influential factor. There are 9 factors considering for proximity and accessibility: CBD, health center, public school, main road, minor road, railway, industiral area, industrial river port and airport. CBD of 2002 is determined by choosing the districts that meet the criteria of spectacular buildings, historical buildings, the Royal Palace, business areas with some commercial buildings, group of department stores, and expensive real estates. Due to the development, CBD of Phnom Penh in 2002 is expanded by including the nearby communes and it is used for influential factor of 2009 and 2015. Public school and health center are digitized as a point location from topographic map of 2002 while main and minor roads and railway are digitized as line. Furthermore, data of 2015 are extracted from openstreetmap website for updating the current road, public school and health center location. Main and minor roads are improved by digitizing additional road from Google historical image of 2010, Google current image of 2015 and the openstreetmap website. Distance to Phnom Penh airport and river port are digitized as polygon from Google earth image and it remains constant for all years. Similary, industrial areas in the outskirt of Phnom Penh, the vast majority are garment factories, are digitized by interpreting from the Google earth image of 2003, 2010, and 2015 by using the reference information from the White Book of Development and Management of Phnom Penh in 2020. On the other hand, to solve the limitation of data of health center and public school in 2009, the interpretation of the Google image in 2015 and the available information from openstreetmap.org are used by considering that if the health center and the university are established between 2002-2009, they will be assigned to use as a factor for 2009 and if they are established between 2009-2015, they will be assigned to use as a factor for 2015. In order to know the information of their establishment, each current university and health center are assessed through their websites or profile information to find out the date of their establishment.

Neighborhood factor is done by using mean focal statistic tool in ESRI ArcGIS 10.2 to produce the probability map of number of urban cell within 7x7 window cell size in 2002, 2009, and 2015. It is assumed that urban and built up areas will be high preferable developed in a place where there are existing nearby built up area located. There are many types of window cell size which can be used to calculate the probability map within their nighborhood such as 3x3, 5x5, 7x7, and 9x9 (Verburg et al., 2003). In this research, 7x7 window cell size is selected the same as applying in the previous studies of Hu and Lo (2007) and Shu et al., (2014) who also extracted LULC data from Landsat data with 30 m resolution. Also, distance to existing urban cluster provides more information about the effect of land cover transition from the nearest exisiting urbanized area is here applied by Euclidean distance tool in ESRI ArcGIS 10.2.

For economic factor that have been reviewed from literature, like GDP and land value, are not be taken in to account in this research since data are unavailable. Census data of 1998 and 2008 did not provide the GDP value per district in Phnom Penh. Similarly, the land value information of 2009 could not be assessed.

The summary of input influential factors used in LR model is presented in Table 3.4.

Table 3.4 Input influential factors used in logistic regression model.

Type of variable	Factor	Nature of variable	
Bio-physical factor	1: vegetation, 0: not vegetation	Dichotomous	(a)
	1: miscellaneous land, 0: not miscellaneous land	Dichotomous	(a)
	1: arable land, 0: not arable land	Dichotomous	(a)
	Distance to river	Continuous	(b)
Social factor	Population density (people/sq. km)	Continuous	(a)
Proximity and accessibility factor	Distance to CBD	Continuous	(c)
	Distance to health center		(a)
	Distance to public school		(a)
	Distance to main road		(c)
	Distance to minor road		(a)
	Distance to railway line		(b)
	Distance to Phnom Penh airport		(b)
	Distance to industrial area		(a)
	Distance to industrial river port		(b)
Neighborhood factor	Number of urban cells within a 7x7 cell window	Continuous	(a)
	Distance to existing urban cluster		(a)

Note a These independent variables are different in all dates.

b These independent variables remain the same for all dates.

c These independent variables are different only in 2002 and 2009.

3.3 Urban growth simulation by logistic regression model

The derived significant influential factor of urban growth as explanatory variable (x) and urban growth area as dependent variable (y) are used as input data of LR model. Herewith, a logistic regression model is used to associate the urban growth with influential factors and to generate an urban growth probability map (Hu and Lo, 2007; Alsharif et al., 2013; Duwal, 2013; and Shu et al., 2014). The general form of logistic regression can be described as follows:

$$y = a + b_1 x_1 + b_2 x_2 + ... + b_m x_m \quad (3.4)$$

Where x_1, x_2, ... x_m are explanatory variables, y is a linear combination function of the explanatory variables representing a linear relationship Equation 3.4. The parameter b_1, b_2, …b_m are the regression coefficients to be estimated and its contribution is to play as explanatory variable (Cheng and Masser, 2003). Because Equation 3.4 is not constant for value to stay on 0 and 1. Thus, the probability value related to x is made in nonlinear way by using logistic curve which has the equation as below (Rogerson, 2010):

$$P = \frac{e^y}{1+e^y} = \frac{e^{a+b_1x_1+b_2x_2+...+b_mx_m}}{1+e^{a+b_1x_1+b_2x_2+...+b_mx_m}} \quad (3.5)$$

The P means the probability of occurrence of a new unit. For example, the transition of cell from non-urban to urban or rural to urban. In logistic regression, the probability value can be a non-linear function of the explanatory variables as represented in Equation 3.5 (Cheng and Masser, 2003). Since the probability are not linearly related to x, the predicted probability is transformed to be linear related to x which is called as logistic transformation by executing the logarithm of P/(1-P) that is known as odds:

$$y = \log_e\left(\frac{P}{1-P}\right) = a + b_1x_1 + b_2x_2 + ... + b_mx_m = \log(P) \qquad (3.6)$$

Thus, the linear regression equation is the natural log-odds of the probability of being urban divided by the probability of being non-urban. Log-odds increase or decrease linearly, as *x* increases or decreases (Rogerson, 2010). The sign positive (+) and negative (-) of b_1, b_2,...b_m also determine the effect of the log-odds from explanatory variable whether log-odds are increasing or decreasing the probability of change for every one unit increase in *x*. For example, the distance to road has a negative influence on urban growth, so areas that are closer to roads reflect a likelihood to become urban.

3.3.1 Goodness of fit of logistic regression model

To assess the goodness of fit of logistic regression which is used to interpret the predictive ability and fitting degree of the model, pseudo R^2 was report from the finding of LR model. In this case, logistic regression does not have an equivalent to the R^2 value as there is found in Ordinary Least Squares (OLS) regression. It is used by other methods which is reported by SPSS logistic model for calculating the explained variation. Two of R^2 in LR model are Cox and Snell R^2 and Nagelkerke R^2 that generally have lower values than the OLS R^2, but they are interpreted in the same manner (Al-Ahmadi, Al-Ahmadi, and Al-Amri, 2013). However, Cox and Snell R^2 cannot reach the maximum value of 1, it should be better to report Nagelkerke R^2 which is used to modify the R^2 value, and if the pseudo R^2 value are between 0.2 and 0.4, it is considered as a good fit (Clark and Hosking, 1986, Hu and Lo, 2007 and Shu et al., 2014). It has formula as below:

$$f_{(x)} = \frac{1 - \exp\left(\frac{-2(LL_m - LL_0)}{N}\right)}{1 - \exp\left(\frac{2LL_0}{N}\right)} \tag{3.7}$$

Where LL_m and LL_0 are the log-likelihood of the model and intercept (the model without any explaining variable), respectively, and N is the sample size.

3.3.2 Wald statistic

Another parameter using in LR model is Wald statistic which is used to assess the contribution of individual logistic regression coefficients for each predictor whether it is significance to the outcome or not. The null hyphothesis is tested that the true value of the slope (the estimated coefficient) is equal to zero, so if the null hyphothesis is rejected that means the coefficent is different from zero, then it gives some evidence that the variable (x) is significance to understand dependent varaiable (y) (Rogerson, 2010). The Wald statistic is the ratio of the square of the regression coefficient to the square of the standard error of the coefficient and is asymptotically distributed as a chi-square distribution with one degree of freedom; (Menard, 2002). The formula is given as below (IBM SPSS, 2012; Menard, 2002):

$$Wald = \frac{b^2}{(SE_b)^2} \tag{3.8}$$

where b is a coefficient of independent variable

SE_b is a standard error that measures predictive accuracy

Wald statistic is interpreted that if the coefficient (b) is more than twice (approximately) its corresponding standard error (SE_b), it may be regarded as significantly different from zero (Rogerson, 2010). So, in other word, if the Wald statistic value is bigger than 4, the independent variables (*x*) are significant and influence to the model outcome of dependent variable (*y*) because the Wald value bigger than 4 gives a level of significance (P-value) of less than $\alpha=0.05$ (chi-square distribution table). Thus, in order to select the significance of independent variables, backward stepwise method is applied based on P-value obtained on 0.05 level of confidence. By applying backward stepwise, all independent variables are included in the equation in order to run the calculation. Then the largest significance variable that has P-value greater than the confidence interval $\alpha=0.05$ are proved as insignificant and are removed from the model. After removing, the calculation is repeated again until the remaining variables have P-value lower than the value of confidence interval (IBM SPSS, 2012).

3.3.3 Odds ratio

Generally, logistic regression has two parameters that provide a usefulness of interpretation include coefficient (*b*) and odds ratio (*OR*). The concept of using coefficient (*b*) is convenient for testing the usefulness of predictors, while odds ratios can be used to interpret as much easier than coefficient (*b*) (IBM SPSS, 2012). Hu and Lo (2007) used *OR* to interpret the effect of estimated independent variables on the probability of urban growth. *OR* is a ratio-change in the odds of the event of interest (odds of the outcome) by increasing a one-unit change in the predictor *x* and it is simply given by the equation:

$$OR = \frac{odds(x+1)}{odds(x)} = \frac{\dfrac{P_{(x+1)}}{1-P_{(x+1)}}}{\dfrac{P_{(x)}}{1-P_{(x+1)}}} = \frac{e^{b_0+b_1(x+1)}}{e^{b_0+b_1(x)}} = e^{b_1} \tag{3.9}$$

- If $b = 0$, the odds and probability are the same at all x levels (e^b=1)
- If $b > 0$, the odds and probability increase as x increases (e^b>1)
- If $b < 0$, the odds and probability decrease as x increases (e^b<1)

3.4 Urban growth validation

In order to evaluate the performance result of the LR model and to test the model accuracy for interpretation of output probability image, the ROC, a statistical techniques, is highly recommend for utilization due to its reliable method (Pontius and Schneider, 2001; Hu and Lo, 2007; Arsanjani et al., 2013). The ROC technique validates the simulated urban map in 2015 with the reference (classified) urban map in 2015 by doing the comparison between them and measuring the relationship of agreement of the pair of images that were urbanized. The area under the ROC curve varies from 0.5 to 1 in which 0.5 is a random assignment of probability and 1 is a perfect assignment of probability (Hu and Lo, 2007). A ROC value of 1 indicates a perfect spatial agreement between actual urban growth map and the predicted probability map. A ROC value of 0.5 indicates an agreement expected by chance (Hu and Lo, 2007; and Alsharif et al., 2013).

According to Pontius and Schneider (2001) and Hu and Lo (2007), the comparison of the simulation map and reference map are implemented by slicing the threshold levels which is the percentage of cells in the probability image to be reclassed as 1 in the preparation for comparison with the reference image. Each threshold values are firstly sliced as equal interval starting from 0-5%, 5-10% until 95-100%, so that 20 threshold intervals in total are obtained. Then the highest probability starting from 95-100% was reclassed as 1 and the remaining 95% downward were reclassed as 0. So the slice of threshold (95-100%) was compared with the referenced image. After that the ROC continues its process to the following 19 thresholds to compare with reference image. The main reason for 20 threshold slicing is to get the curve more smoother, the more threshold used, the smoother the curve will be (IDRISI, 2012). Furthermore, Pontius and Schneider (2001) also mentioned that the increase number of threshold, the higher the accuracy of the estimated the ROC. Each slice threshold of the predicted probability map was compared with the reference map using a two by two contingency table (Table 3.5). Where:

(A) is the number of true positive cells that predicted as urban growth and agree with the referenced image,

(B) is the number of false positive cells that predicted as urban growth but disagree with the referenced image,

(C) is the number of false negative cells that predicted as non-urban growth expansion but disagree with the referenced image, and

(D) is the amount of true negative cells that predicted as non-urban growth and agree with the referenced image.

In this study, the acceptance ROC value for the simulated urban growth validation of three periods: 2002-2009, 2009-2015, and 2002-2015 with reference data of urban area from 2009, 2015, and 2015, respectively should be equal or more than 0.7 because it indicates that independent variables have strong interpretability on dependent variable (Pontius and Schneider, 2001; and Shu et al., 2014).

Table 3.5 Contingency table of the slice image of predicted urban growth and reference image.

Predicted probability map	Actual map		Total
	Urban growth	Non urban growth	
Urban growth	A	B	A+B
Non urban growth	C	D	C+D
Total	A+C	B+D	A+B+C+D

Every contingency table creates one data point (X, Y) where X is rate of false positive and Y is rate of true positive, respectively. These data points are connected to create ROC curve from which the ROC statistic is computed for X and Y as below:

$$\text{False positive (\%)} = \frac{B}{B+D} \tag{3.10}$$

$$\text{True positive (\%)} = \frac{A}{A+C} \tag{3.11}$$

Figure 3.4 shows an example of ROC curve from Hu and Lo (2007).

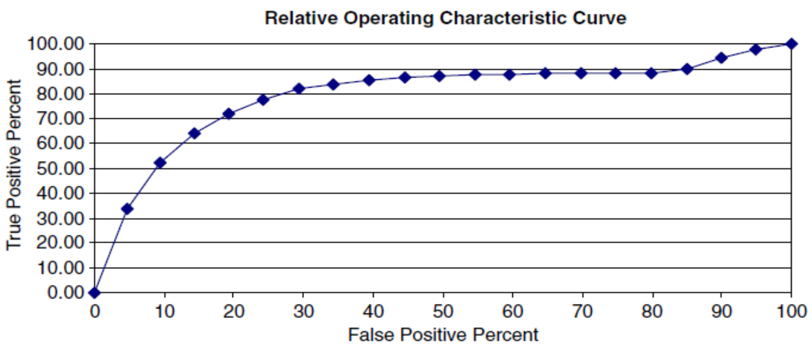


Figure 3.4 Example of ROC curve (Hu and Lo, 2007).

3.5 Urban growth prediction

The weakness points of LR model are quantification of change and temporal analysis, as it provides only the probability area and direction of the growth. Therefore, measuring change demand for urban growth is required in order to provide the pattern of urban growth in the future. According to Arsanjani et al. (2013), there are two methods to estimate of change demand for the future such as Markov Chain model and statistical extrapolation (by population prediction). Moreover, the work of Campbell et al. (2007) and Alsarif et al. (2013) proposed the change demand by using the population growth and land development rates into the equation, but it is impossible for making an assumption based only on amount of population growth due to the case of inconsistent of land development policy, for example, the building of multi-story apartment on the plot of land is more profitable than the building of one single family house (Arsanjani et al., 2013). According to many possible ways as described, three methods include (1) future population growth rate; (2) Markov Chain model and (3) simple linear regression of estimating change demand were here examined to identify as an optimal method for urban growth demand estimation. After that, the optimum predicted demand area is further used to allocate to the derived future urbanization probability map based on optimum LR model in order to generate urban growth pattern of Phnom Penh in 2030. In practice, allocation process start from the highest to the lowest cell until the total growth demand area are equal to the estimated future urbanized area in 2030.

3.5.1 Urban growth change demand by future population growth rate

According to the work conducted by Campbell et al. (2007) and Alsharif et al. (2013), one can propose the estimated quantity of land demand for future urban expansion by using the future developed land area and growth ratio while logistic regression probability output map was used to indicate the knowing direction of urban expansion. Therefore, to generate map of future urban distribution patterns, the future urban area should be estimated according to urban intensity of new development which is a growth ratio, existing area of urbanized land, expected and existing population, as shown in Equation 3.12.

$$A_{future} = A_{existing}\left(1 + R\frac{P_{future} - P_{existing}}{P_{existing}}\right) \qquad (3.12)$$

Where: A_{future} future area of urbanized land

$A_{exisitng}$ existing area of urbanized land

R (growth ratio), $\frac{\text{ratio of change in urbanized land (\%)}}{\text{ratio of population growth (\%)}}$

P_{future} the expected population in the future

$P_{existing}$ the existing population

A growth ratio is designated in the form of 1:1 indicates the value of proportion of population growth with the change of developed land. For instance, if 10% increase of population, there would be an increase of 10% of developed land. Campbell et al. (2007) used growth ratio of 5:1 to apply in his work since this ratio had been found and was used for future growth modeling conducted by the STI and SC Coastal Conservation League. Growth ratio of 5:1 means that if we assume the current developed land is 1,000 acres and the population is forecasted to increase by 10%, it prescribes the area will be increase 5 times or 50%. Therefore, the developed land will increase by 500 acres that totally is 1,500 acres. In his work, 576,336 acres of the latest year were given with overall growth ratio of 5:1, so that the predicted land by the year 2030 is 1,523,667 acres. Alsharif et al. (2013) adapted this work by putting a growth rate data of 1.41% per year that obtained from the General Authority of Information whereas ratio of urbanized area growth was calculated based on a situation of urban change between two maps of his study in 2002 and 2010 which equaled to 8.57% per year. Thus, the growth ratio was about 6%. He suggested to apply an assumption of three scenarios of growth ratio to predict future urban expansion due to the instability of economic, social, and political condition that might affect the change of urbanized land. Moreover, urban planning policy were unclear that need to estimate the different perspective of urban growth. So the first scenario is calculated by growth ratio of 6, second scenario is 7 and the third is decreased to 5.

3.5.2 Urban growth change demand by Markov Chain model

The Markov Chain model is able to predict the next status of change according to the previous status; therefore, a transition probability matrix and a transition area matrix were produced by this model indicating the amount of change between existing categories (Arsajani et al., 2013). The transition probability matrix records the probability that each land cover category will change to every other category, while the transition areas matrix records the number of pixels that are expected to change from each land cover type to each other land cover type over the specified number of time units (IDRISI, 2012).

3.5.3 Urban growth change demand by simple linear regression

For this method, this study use the simple linear regression to extrapolate urban growth area of 2030 based on urban area of 2002, 2009, and 2015 by Trend analysis under MS Excel spreadsheet.

CHAPTER IV

RESULTS AND DISCUSSIONS

Major results and findings are here separately described according to objectives of the study that has five parts include (1) development of land cover and its change during 2002 to 2015; (2) development of urban growth pattern; (3) urban growth simulation and influential factor identification; (4) urban growth model validation and optimum LR model for urbanization prediction; (5) urban growth pattern prediction of Phnom Penh in 2030.

4.1 Development of land cover and its change during 2002 to 2015

Results of the first part describe the status of land cover in three dates and their changes which were derived by maximum likelihood classification using multi temporal satellite images in 2002, 2009, and 2015 of Landsat 7, 5, and 8, respectively in order to further explain quantification of spatial temporal pattern of urban growth.

4.1.1 Land cover classification and accuracy assessment

The derived land cover information of 2002, 2009, and 2015, which were classified using supervised classification method by maximum likelihood clasifier, are presented as land cover distribution map in Figures 4.1 to 4.3, and area and percentage of them are comparatively summarized in Table 4.1.

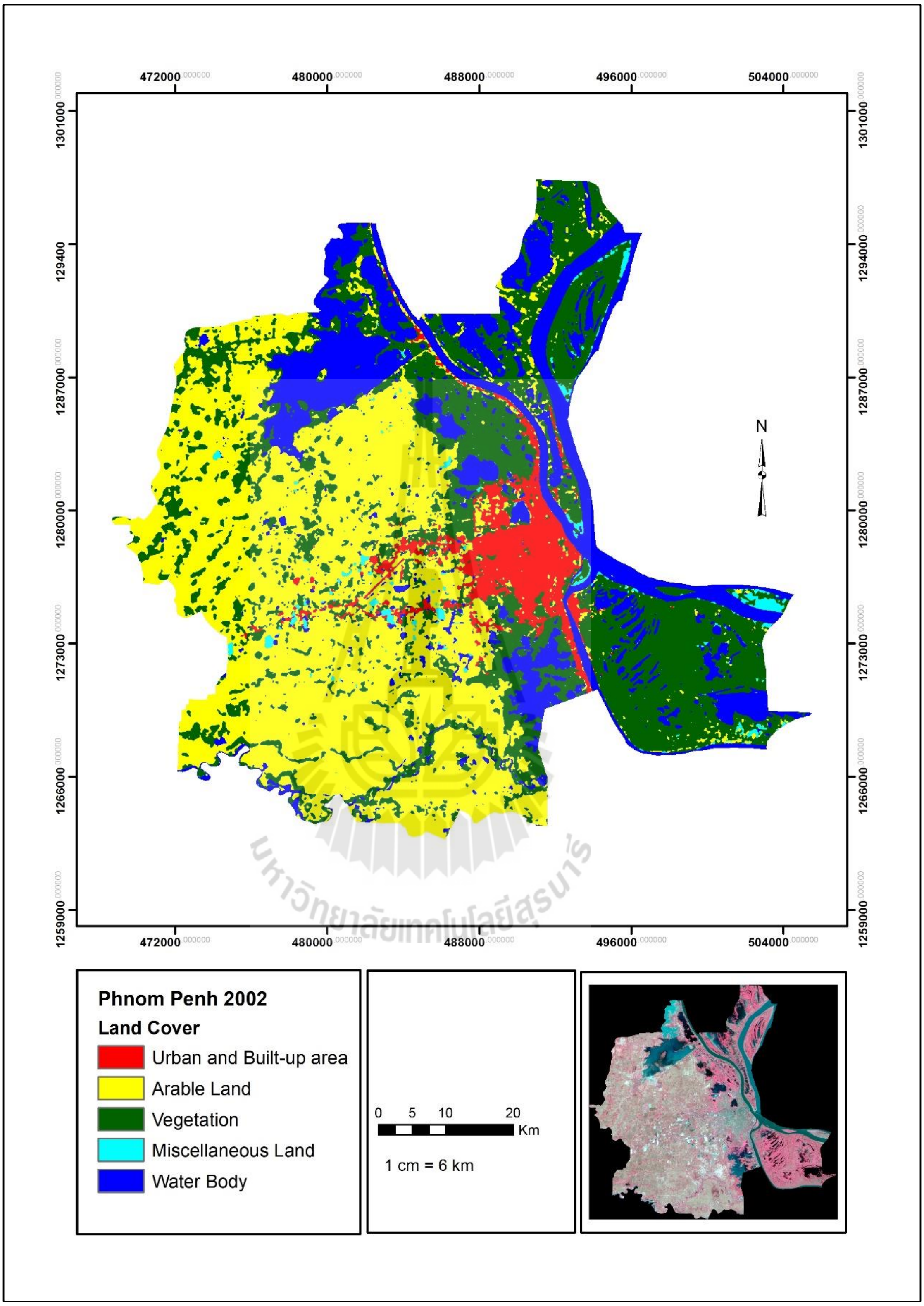


Figure 4.1 Distribution of land cover pattern in 2002.

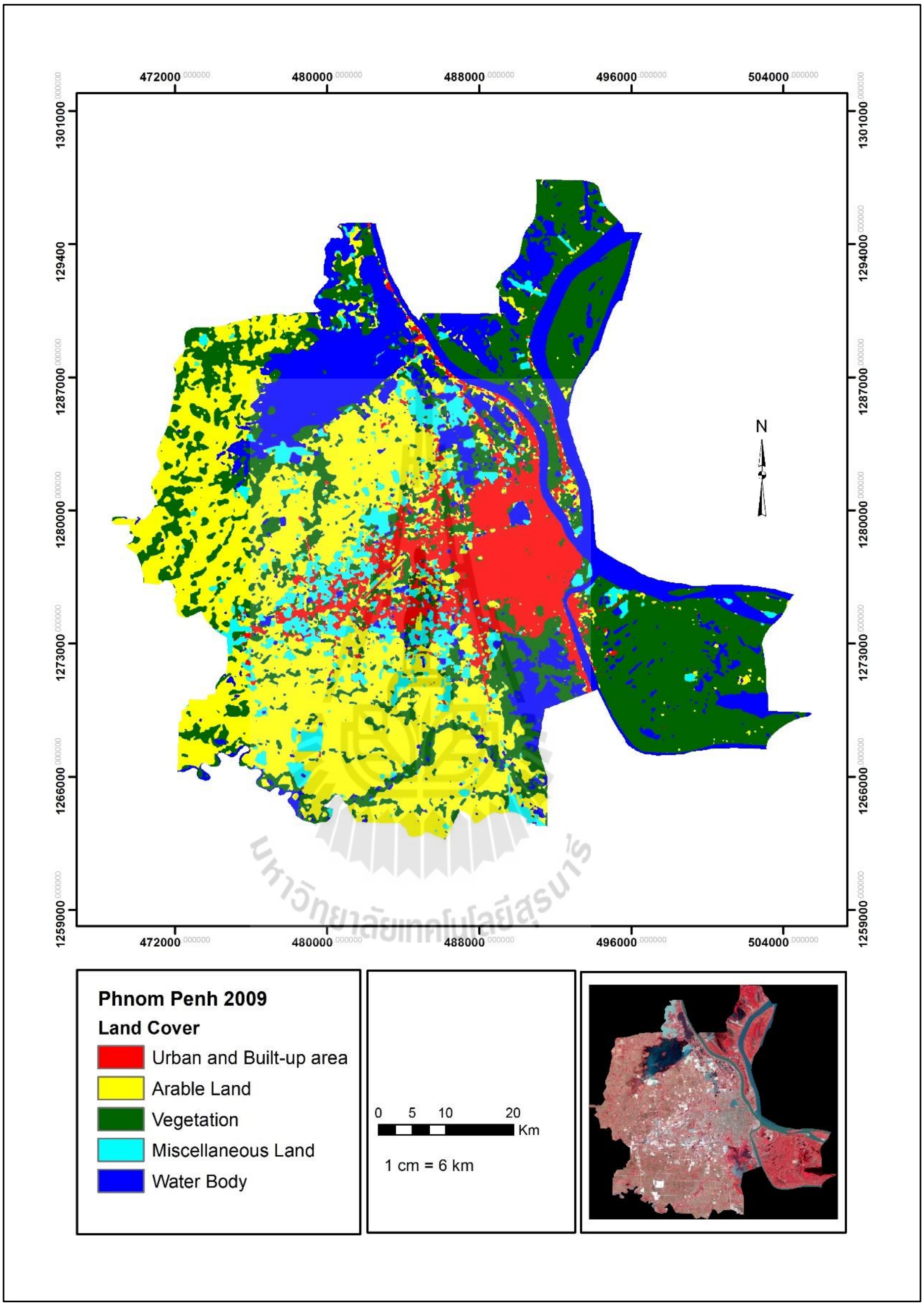


Figure 4.2 Distribution of land cover pattern in 2009.

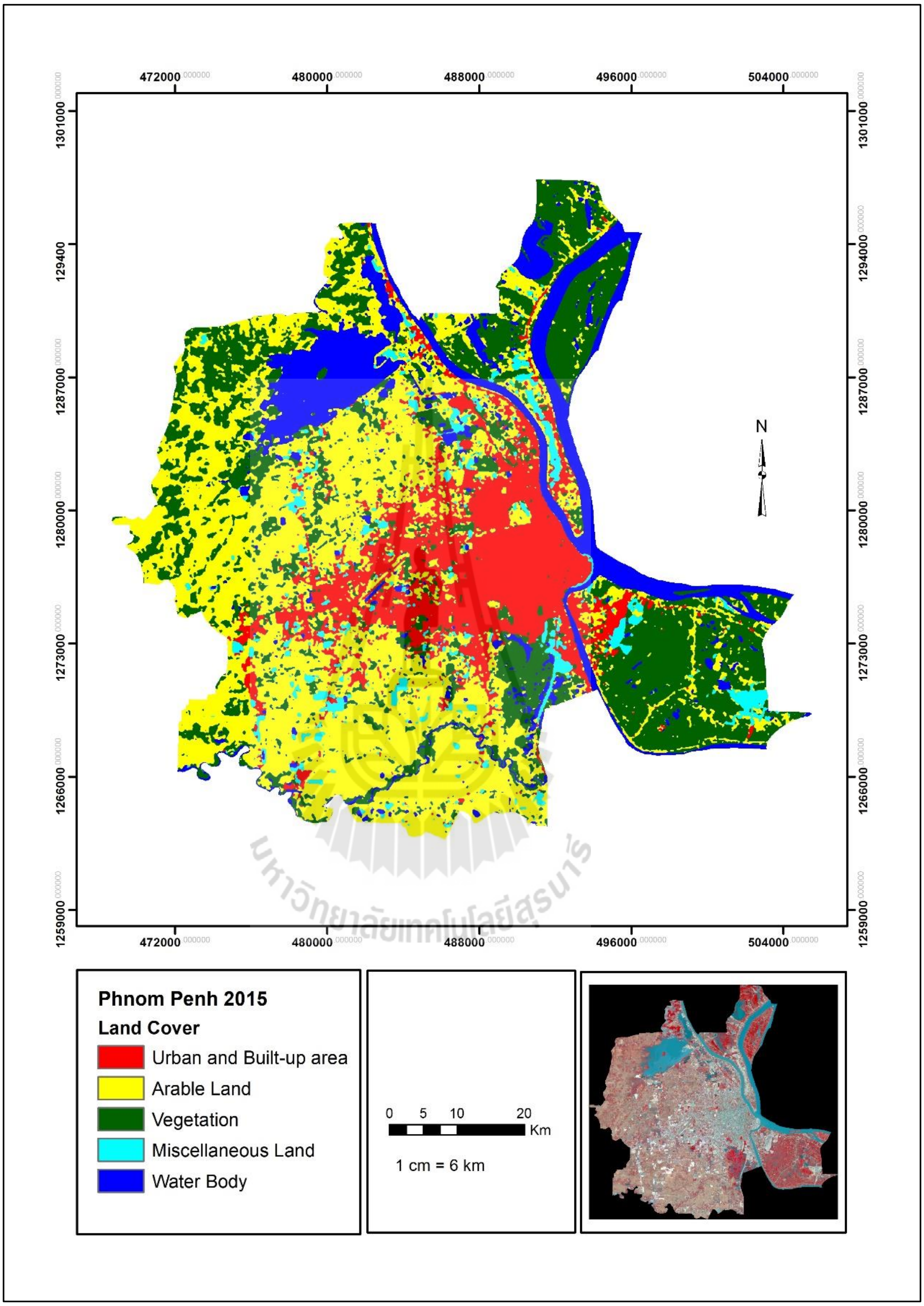


Figure 4.3 Distribution of land cover pattern in 2015.

Table 4.1 Area and percentage of land cover classes during 2002-2015.

Land cover classes	**2002**		**2009**		**2015**	
	sq. km	**%**	**sq. km**	**%**	**sq. km**	**%**
Urban and built up area	38.13	6	70.31	10	108.73	16
Arable land	292.39	43	225.88	33	275.92	40
Vegetation	215.28	31	213.13	31	186.56	27
Water body	132.57	19	132.18	19	88.54	13
Miscellaneous land	8.29	1	45.16	7	26.91	4
Total area	686.66	100	686.66	100	686.66	100

In term of land cover distribution, land cover classification in 2002 reveals that arable land (majority is paddy field) is predominant land cover class that consists of 292.39 sq. km, about almost half area of total of Phnom Penh land followed by vegetation, water body, urban and built up and miscellaneous land (Table 4.1 and Figure 4.1). The image indicates clearly that the western part of Phnom Penh from the river are arable land and the eastern part are vegetation (crop land and grass land). The core zone of urban in Phnom Penh is situated only at the confluence of the three rivers (Mekong, Tonle Sap, Bassac rivers) about 38.13 sq. km while the Phnom Penh airport is situated on the suburb area at the western part surrounded by arable land and some vegetation. However, 7 years later in 2009, urban area increased about two times of the urban area in 2002 that expanded along the main road to the airport and some were expanded to the north and the south, but for the eastern part of Phnom Penh, there were few because of some of the land there were wet land and easily flooded during the rainy season (Table 4.1 and Figure 4.2). This double expansion surged the urban up from 38.13 sq. km to 70.31 sq. km that most of the increase was consumed largely about 50% of arable land about 32.18 sq. km and small amount of vegetation about 2.15 sq. km. This shows abruptly decreased of arable land from 292.39 sq. km in 2002 to 225.88 sq. km in 2009 while another 50% consumption of arable land was converted to miscellaneous land that raised from 8.29 sq. km to 45.16 sq. km because of the development of urban like residential building and industrialization along the main road in the suburb area around the airport, so that in 2009 land cover classification shows most part of arable land were remarkably converted into bare land for construction site and development activity. This proves the high effect contribution of the construction sector to the public welfare of Phnom Penh started very high in 2009 by developing new construction project for industrialization that can offer more job and attract more rural immigrant. As a result, urban in 2015 was increased largely from the miscellaneous land in 2009 about 38.42 sq. km (Table 4.1 and Figure 4.3), however, during 2009 to 2015, miscellaneous land (bare land) still increased more and remain about 26.91 sq. km that indicates the development of urban expansion was in progress by transforming some arable land, vegetation and water body.

As shown in Table 4.1 and Figure 4.4, vegetation decreased from 213.13 to 186.56 sq. km and water body remain stable from 2002 to 2009 about 132 sq. km, but it decreased extremely to 88.54 sq. km due to the land infilling that convert water body to miscellaneous land. Arable land is showed slightly increase due to seasonal effect of the Landsat in 2015 which illustrated the northern part of Phnom Penh is a floating rice field. Due to the water flowing out to the large lake, the floating rice field became dry land and give spectral characteristic value as arable land, so that it made a different between Landsat in 2002 and 2009 which were water body because of that time floating rice field was flooded by rainy season water and remained stability until rainy season was over. Moreover, the increased of arable land in 2015 is also caused by the lake infilling for the development of satellite cities, and some area in the low land area of eastern part Phnom Penh.

In short, land cover categories in 2002, 2009, and 2015 as shown in Figure 4.4 illustrated that the land cover change shows the remarkably increased of urban and built up area continously while arable land and miscellaneous land are fluctuated and vegetation and water body are shown continously decrease.

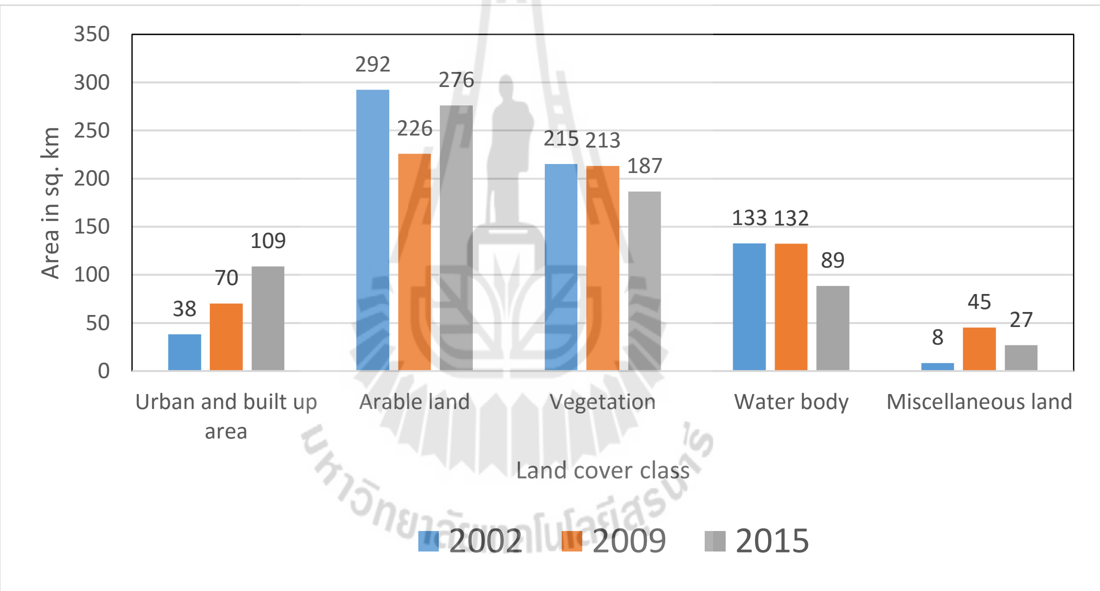


Figure 4.4 Comparison of each land cover type value in three different peroids.

Furthermore, thematic maps of land cover in 2002, 2009, and 2015 were further accessed its accuracy with 203 stratified random sample points with corresponding reference data include topographic map and historical google earth images. With the help of the high resolution images, ground verification was done for unclear areas for every classified image, especially the low radiometric sensitivity areas of Landsat images. Finally, the misclassified and doubtful areas of the three images were corrected using recode option in ERDAS Imagine in order to refine the classified data for the final land cover map as shown in Figues 4.1 to 4.3. Accuracy assessment result of three classified image is summarized in Table 4.2 include overall accuracy, producer's accuracy (PA), user's accuracy (UA) and kappa hat coefficient. Detail of error matrix for accuracy assesment of three dates is shown in Appendix B.

Table 4.2 Landsat accuracy classification report.

Land cover class	**2002**		**2009**		**2015**	
	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)
Urban and built up area	73.33	100.00	82.61	90.48	96.88	96.88
Arable land	92.86	90.70	87.50	94.03	93.51	87.80
Vegetation	90.63	90.63	91.94	90.48	87.93	92.73
Water body	100.00	97.44	100.00	89.74	92.86	100.0
Miscellaneous land	100.00	66.67	100.00	84.62	87.50	87.50
Overall accuracy	**92.16%**		**91.18%**		**92.12%**	
Kappa hat coefficient	**88.60%**		**88.11%**		**89.13%**	

As results, it reveals that the extraction of water body has the highest PA in all images for almost 100% and it is followed by miscellaneous land, arable land, vegetation and urban and built up area. Miscellaneous land also give high PA except in 2015 that 12.5% are misclassified to the water body. Result shows urban and built up areas in 2015 give a high PA about 97% whereas the omission error in 2002 and 2009 are found about 26% and 17% in the classification respectively. This is probably because of the conflict pixel value between urban and built up areas and bare land under miscellaneous land in 2009 and the problem of mixed pixels between urban and built up areas with vegetation and arable land in 2002 and 2009 which is difficult to classify by using medium spatial resolution as Landsat images. Meanwhile, vegetaion and arable land from three dates obtain PA over 85% which reflects a good result of classification and the minimal error is caused by phenological change on the land cover between them. Some arable land which are bare soil or paddy field are occupied by shrub and grassland that reflect spectral characteristic values similar to vegetation. In addition, results indicate that the high value of UA for all land cover classes are over 90% except for miscellaneous land that have UA lower than 90% for all years because of over classification of miscellaneous land by conversion of urban and built-up areas, arable land and water body pixels to miscellaneous land.

In summary, the overall accuracy of all classified images are greater than 90% which is demonstrated as a good result for further use to study urban growth. In this study, the obtained Kappa hat coefficient values of the classified land cover maps were 88.6%, 88.1%, and 89.1% for 2002, 2009, and 2015 respectively. Landis, and Koch (1977) states that Kappa hat coefficient values > 80% represents strong agreement between classification and referenced points, so that it is acceptable for performing as quantitative analysis.

4.1.2 Land cover change in 2002, 2009 and 2015

To understand land disparity and development progress in different land categories in three different periods (2002-2009 and 2009-2015 as short term period and 2002-2015 as long term), change detection matrix based on post comparision change detection algorithm are prepared and explained the result as below.

According the change matrix of land cover change between 2002 and 2009 (Table 4.3), urban and built up area had an increased area of 32.18 sq. km or 4.69% of the study area, or 4.6 sq. km per annum. The largest increased area came from arable land about 16.41 sq. km followed by vegetation about 13.7 sq. km and some minority land change of miscellaneous and water body. Figure 4.5 shows the pattern of urban growth and land cover change in 2002-2009 that illustrates the major transformation was arable land along the main road to the airport and some areas near industrial area. The transformation of vegetation is also presented near the main road, the lakes at the northern and southern part of Phnom Penh and in the neighborhood of existing urban and built-up area. At the same time, miscellaneous land had a massive increased area in amount of 36.87 sq. km or 5.37% of the study area with annual increase of 5.27 sq. km. The increase came from arable land which is the significance change about 27.33 sq. km, vegetation about 10.12 sq. km and the smallest change is water body about 5.8 sq. km. During this period, the transition of miscellaneous land was very impressive and remarkable, since the amount of growth have a lot more than amount of urban about 5 sq. km. The reason behinds this massive change is because of the development of six large urban areas which simply call “satellite city” that were approved by the Royal Government of Cambodia in the past ten years. The constructions were started in the purpose to connect those cities to the core existing urban of Phnom Penh resulting the decline in arable land, vegetation and some big lakes in the northern part of the core urban zone of Phnom Penh. Recently, 4 of those satellite cities are under construction: the first one is Diamond Island which is the small island in the east of Phnom Penh, and it was seen some major change of vegetation to bare land, the second is Grand Phnom Penh located at the upper part of

Phnom Penh which the major land cover change was vegetation and lake, the third is Camko City located next to the core existing urban in the north and it was filled by sand bar to convert lake into bareland, and the last one is Boeung Kak lake that just were being filled by sand in only some part. For some miscellaneous land pattern in the west of Phnom Penh, most of the tranformation came from arable land along the national road number 4 that linked Phnom Penh to the coastal province of Cambodia and because it is on the suburb area around Phnom Penh airport, many dry ports were located there and the construction site of industrialization like garment factories, inudstrial park, Phnom Penh Special Economic Zone (PPSEZ) and residential building were started to grow and consume a huge amount of arable land, so that it would extend the city to continue west.

For decreased land cover type, arable land lost an area of 66.5 sq. km or 9.69% of the study area or 9.5 sq. km per annum. It was changed to urban and built up area, miscellaneous land, water body, and vegetation due to seasonal effect of paddy field, wet land area and some parts of arable land that was became open land dominated by shrub and grass. Vegetation and water body decreased area of 2.15 sq. km and 0.39 sq. km respectively. Most of vegetation were converted to urban and built-up area, arable land and miscellaneous land, and they were converted to water body due to some parts of vegetation are situated in low land area that covered by flooded water during rainy season.

Table 4.3 Land cover change matrix between 2002 and 2009.

Land cover in 2002	Land cover in 2009 (Unit: sq. km)					
	U	**A**	**V**	**M**	**W**	**Total**
Urban and built-up area (U)	38.13	0.00	0.00	0.00	0.00	38.13
Arable land (A)	16.41	203.13	37.00	27.33	8.52	292.39
Vegetation (V)	13.70	17.72	151.54	10.12	22.20	215.28
Miscellaneous land (M)	0.69	1.14	3.60	1.90	0.95	8.29
Water body (W)	1.38	3.89	20.99	5.80	100.51	132.57
Total	**70.31**	**225.88**	**213.13**	**45.16**	**132.18**	**686.66**
Area of change (sq. km)	**32.18**	**-66.50**	**-2.15**	**36.87**	**-0.39**	
Percentage of study area (%)	**4.69**	**-9.69**	**-0.31**	**5.37**	**-0.06**	
Area per annum (sq. km)	**4.60**	**-9.50**	**-0.31**	**5.27**	**-0.06**	

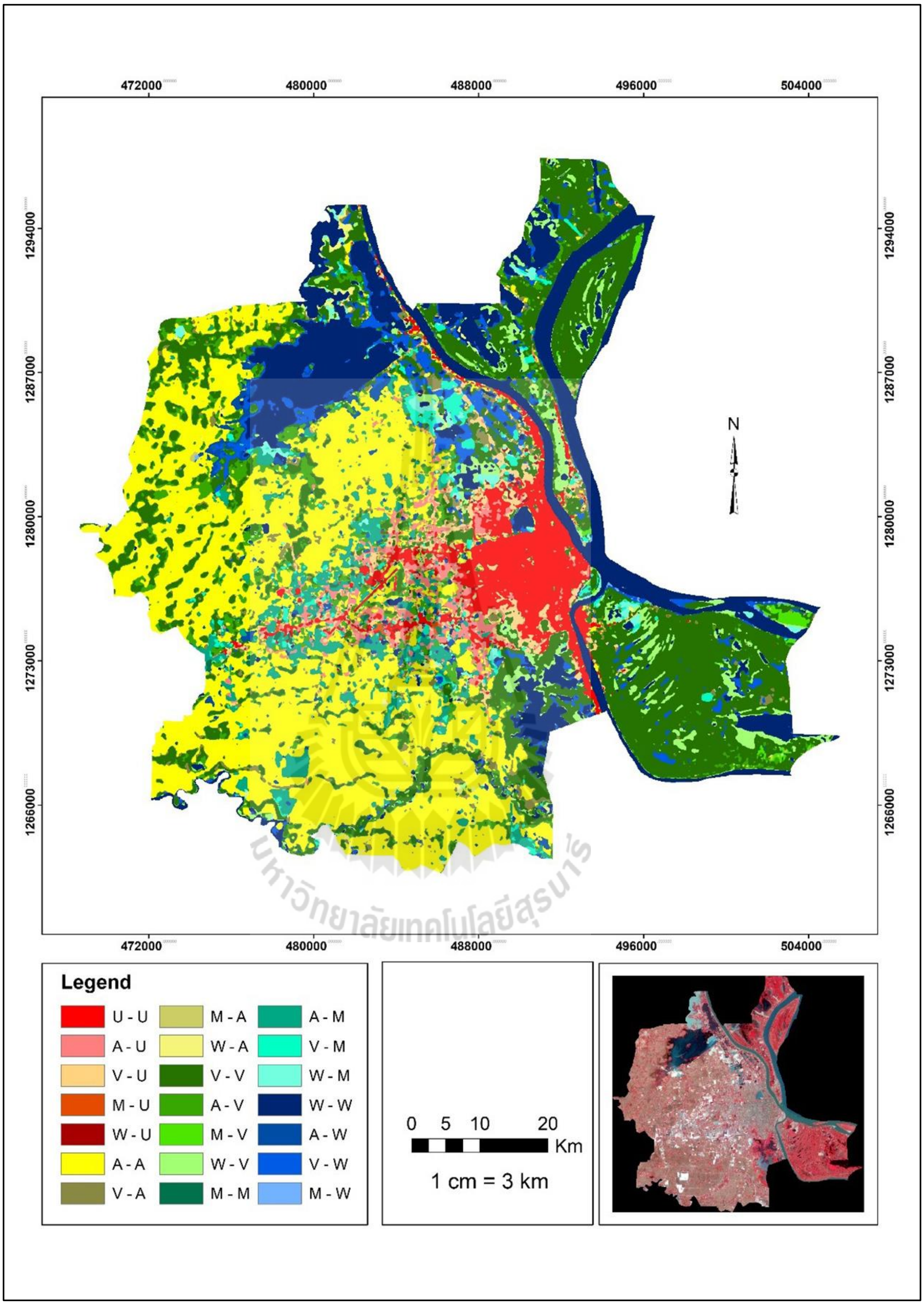


Figure 4.5 Land cover change between 2002-2009 period.

Refer to land cover change matrix during 2009-2015 as shown in Table 4.4, the increase of land cover type were urban and built up area and arable land about 38.42 sq. km and 50.04 sq. km respectively. Most of the transformation to urban came from miscellaneous land about 14.22 sq. km, vegetation about 12.36 sq. km, arable land about 7.75 sq. km and water body about 4.10 sq. km, and the transformation to arable land came from vegetation about 42.45 sq. km, miscellaneous land about 19.06 sq. km, water body about 23.01 sq. km. The result reveals that 14.22 sq. km of miscellaneous land presented in 2009 was converted to urban and built-up area in 2015, and the rest were coverted to arable land, vegetation and water body. This means that the construction of urban and built-up area are still in progress, especially the area of big project like satellite cities and industrial area at the north and west suburban Phnom Penh, so that as time moved forward for 6 years, most of the remaining bare land (miscellaneous land) in 2009 were became like free open land that had some shrub growing on that area, and it reflects the spectral characteristic value same like arable land around 19.06 sq. km. The conversion of miscellaneous land to vegetation and water body gave just only a small amount of number about 3.67 sq. km and 1.93 sq. km, and it is caused by the sesaonal effect, however some bare land of the Grand Phnom Penh satellite city had been finished the construction of golf course that indicated the amount of appearance of vegetation. Vegetation also gave a significance change to urban and built-up area about 12.36 sq. km that most of the change situated in the east of Phnom Penh on the wet land area and on the Diamond satellite city island. Moreover, about 42.45 sq. km of vegetation were transformed to arable land, as shown in Figure 4.6 showed that some of the change came from the expansion of road width and the development of new paved and earthen road. Beside that vegetation nearby urban and built-up area were occupied by land developer and they were kept them as vacant land. This causes the signature of vegetation likes bare soil. Majority of lakes were transformed to residential areas about 4.10 sq. km, and some arable land at PPSEZ in the west and at industrial area in the south were converted to urban and built-up area as well about 7.75 sq. km. Area of 23.01 sq. km of water body was converted to arable land, mostly on the big lakes like Beung Kak lake which is under construction satellite city in the middle of Phnom Penh, and the wet land area in the north east of Phnom Penh which is the future satellite city namely Garden City of Ly Yong Phat's group along National Road number 6. In fact, it should be bare land (miscellaneous land) since the lakes and wet land were filled up by sand, but because of the remaining areas from construction were kept freely, so that the spectral value looks similar to bare soil and paddy field that are in arable land class.

Vegetation, miscellaneous land and water body were the decreased land cover types. Vegetation in 2015 was decreased 26.57 sq. km or 3.87% of the study area. Even though it was a decreased class, its remaining area was 144.18 sq. km and it was increased from the conversion of arable land about 19.34 sq. km, especially at the north west of Phnom Penh and another increased area was from water body about 19.37 sq. km that most of the conversion was situated on the wetland of Beung Tom Pun lake at the south of Phnom Penh where morning glory were cultivated on that lake. The change of miscellaneous land about 3.67 sq. km to vegetation have been described above in the previous paragraph. It was changed due to the golf course construction of Grand Phnom Penh city. Miscellaneous land in 2015 also decreased about 18.24 sq. km or 2.66% of the study area, but the increase of miscellaneous land came from vegetation about 8.35 sq. km and water body about 6.15 sq. km. The change mostly located in the wetland area at the east of Phnom Penh and the right part of Beung Tom Pun lake at the south of Phnom Penh and it was changed due to the development of another satellite city which is called “Chroy Chongva City” and “ING City”. Arable land about 6.14 sq. km also converted to miscellaneous land as it was seen at the western part of PPSEZ and industrial area. Water body in 2015 showed the most declination about 43.65 sq. km or 6.36% of the study area to the other land cover types through the development of urbanization of Phnom Penh. The increase of water body is not so noticeable which is from arable land about 1.27 sq. km, vegetation about 5.8 sq. km and miscellaneous land about 1.93 sq. km.

Table 4.4 Land cover change matrix between 2009 and 2015.

Land cover in 2009	**Land cover in 2015** (Unit: sq. km)					
	U	**A**	**V**	**M**	**W**	**Total**
Urban and built-up area (U)	70.31	0.00	0.00	0.00	0.00	70.31
Arable land (A)	7.75	191.39	19.34	6.14	1.27	225.88
Vegetation (V)	12.36	42.45	144.18	8.35	5.80	213.13
Miscellaneous land (M)	14.22	19.06	3.67	6.28	1.93	45.16
Water body (W)	4.10	23.01	19.37	6.15	79.54	132.18
Total	**108.73**	**275.92**	**186.56**	**26.91**	**88.54**	**686.66**
Area of change (sq. km)	**38.42**	**50.04**	**-26.57**	**-18.24**	**-43.65**	
Percentage of study area (%)	**5.60**	**7.29**	**-3.87**	**-2.66**	**-6.36**	
Area per annum (sq. km)	**6.40**	**8.34**	**-4.43**	**-3.04**	**-7.27**	

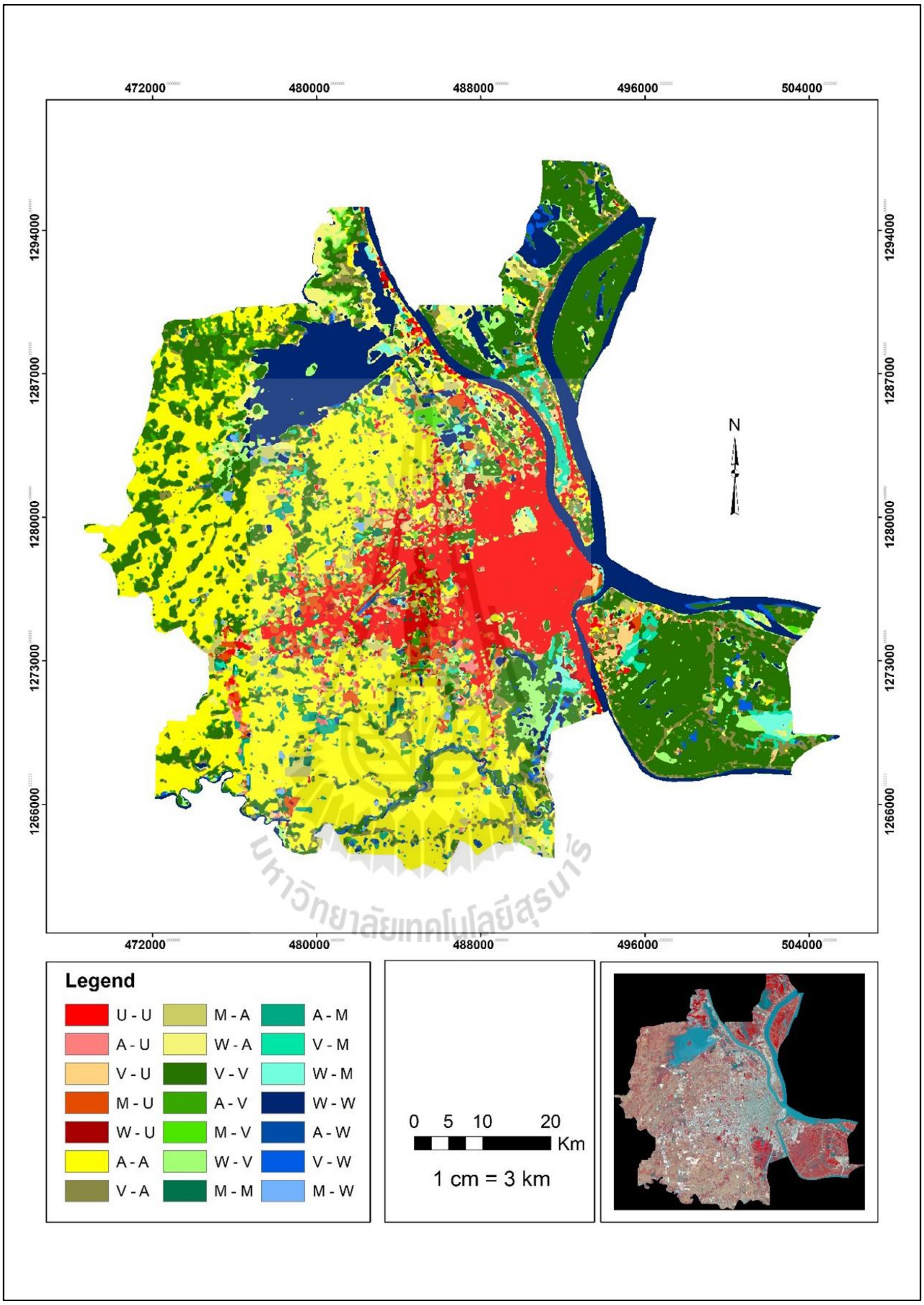


Figure 4.6 Land cover change between 2009 and 2015 period.

For long term peroid (2002-2015) of land cover change matrix as shown in Table 4.5 and land cover change map in Figure 4.7, urban and built up area was increased 70.6 sq. km or 10.28% of the study area or 5.43 sq. km per annum. Figure 4.7 shows the pattern of land cover change to urban and built-up area that most of the conversion was arable land about 34.96 sq. km. It started the change from the existing urban outer edge to the west of Phnom Penh, because the conversion of urban was occurred along the main road that linked Phnom Penh downtown to the airport in the suburb area. Therefore, through the development, many arable land around the airport were converted and continued westward and caused Phnom Penh airport eventually find itself in the middle of urban area. Morover, it was seen that the development project of PPSEZ and some arable land along the ring road and national road Number 3 also converted to urban and built-up area that caused an isolated growth pattern from the core urban area. Vegetation about 27.77 sq. km, situated in the north, south and east of Phnom Penh, was also changed into urban and built-up area, and most of conversion are residential areas and satellite cities. Miscellaneous land and water body were changed into urban and built-up area with a total change of 1.85 sq. km and 6.03 sq. km respectively. Furthermore, miscellaneous land was increased 18.62 sq. km or 2.71% of the study area with 1.43 sq. km per year. The increase reflected the development of urbanization and industrialization of Phnom Penh, since 10.85 sq. km of arable land was converted into bare land which is the construction site of PPSEZ and some industrial site located at westward Phnom Penh. Vegetation about 6.86 sq. km and water body about 8.76 sq. km indicated the significance change to miscellaneous land due to the development of Chroy Chonva City and Ing City.

The decreased of land cover type in this period was arable land about 16.47 sq. km or 2.4% of the study area and 1.27 sq. km per year. Vegetation gave a significance change of 48.62 sq. km where the change can be seen at the northward Phnom Penh. By that time in 2002, the location of change was a low wetland area that had occupied by aquaculture plants over the lake and flooded water, so that according to the development and urbanization, it was filled by land and was kept as open land for sales and construction that reflected the spectral characteristic as arable land. Moreover, due to some paved and earthen road construction and expansion, more vegetation were cleared that gave pattern of change as line in the eastward Phnom Penh. For water body, about 18.57 sq. km was changed to arable land by filling the sand to the lake and to the wetland area due to the development of Beung Kak lake and Garden satellite City in the middle and northward of Phnom Penh respectively. Some of water body area at the north of big lake were a location of floating rice field and due to the seasonal change, the land was dried so then it was become arable land. Vegetation decreased 28.72 sq. km or 4.18% per year and 2.21 sq. km per year. However, 36.79 sq. km of arable land were changed to vegetation. this probably came from seasonal effect while 24.63 sq. km of water body also changed to vegetation, especially it was found in the Boueng Tom Pun lake in the southward Phnom Penh. Water body shows a remarkably decrease about 44.04 sq. km or 6.41% and 3.39 sq. km per year. Most of the change was come from the development of construction projects such satellite cities, and residential buildings.

Table 4.5 Land cover change matrix between 2002 and 2015.

Land cover in 2002	**Land cover in 2015** (Units: sq. km)					
	U	**A**	**V**	**M**	**W**	**Total**
Urban and built-up area (U)	38.13	0.00	0.00	0.00	0.00	38.13
Arable land (A)	34.96	206.32	36.79	10.85	3.47	292.39
Vegetation (V)	27.77	48.62	122.31	6.86	9.71	215.28
Miscellaneous land (M)	1.85	2.42	2.82	0.44	0.77	8.29
Water body (W)	6.03	18.57	24.63	8.76	74.59	132.57
Total	**108.73**	**275.92**	**186.56**	**26.91**	**88.54**	**686.66**
Area of change (sq. km)	**70.60**	**-16.47**	**-28.72**	**18.62**	**-44.04**	
Percentage of study area (%)	**10.28**	**-2.40**	**-4.18**	**2.71**	**-6.41**	
Area per annum (sq. km)	**5.43**	**-1.27**	**-2.21**	**1.43**	**-3.39**	

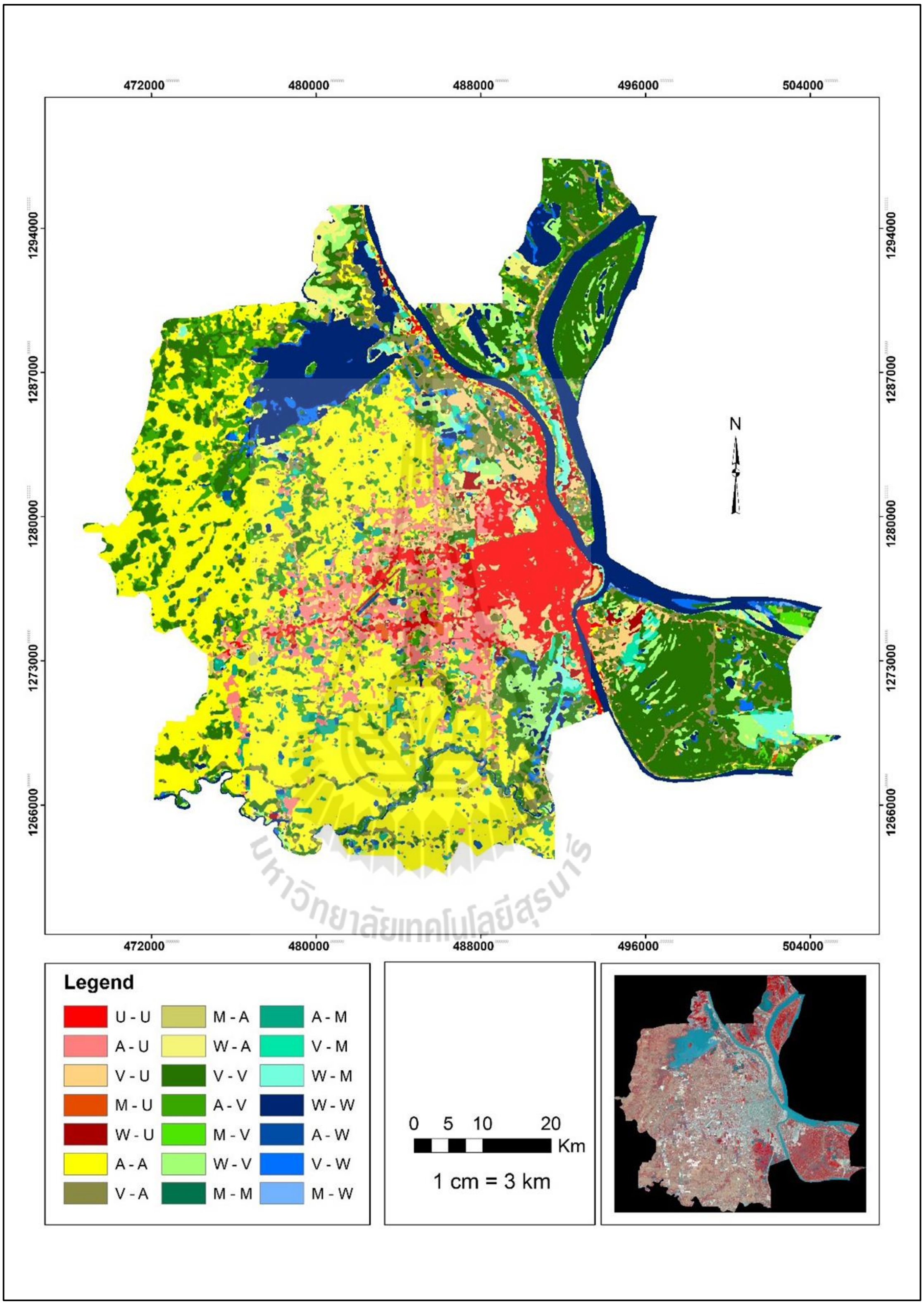


Figure 4.7 Land cover change between 2002 and 2015 period.

4.2 Development of urban growth pattern

Urban growth pattern occurred between 2002-2009 and 2009-2015 was here extracted using overlay analysis. From the literature review, there are three main types of spatial pattern of urban growth: (1) infill growth, (2) expansion growth and (3) outlying growth (linear branch, clusterd branch and isolated).

In the 2002-2009, it was found that Phnom Penh urban growth pattern illustrated in all type of growth such as infill growth, expansion growth, linear branch and isolated growth (Figure 4.8). Most of infill growth areas was inside the core existing urban and some parts at the northward Phnom Penh in Tuol Kok district where most of conversion came from vacant land which is arable land. Expansion growth was found some dispersed areas at the existing urban outer edge or boundary and some areas next to the airport. Moreover, linear branch was depicted as a signifcance pattern of Phnom Penh since it was seen as a majority of growth that were found as linear fashion along the main road to the airport, and also along the national road Number 4 to Kompong Speu province. Isolated growth was also found by some scattering areas with some far distance from the existing urban, since most of this growth pattern were industrial sites and the development of PPSEZ that surrounded by little developed land.

In 2009-2015, Figure 4.9 reveals that the major pattern of urban growth was depicted in three types: (1) expansion, (2) clustered branch, and (3) isolated growth. For expansion growth pattern, it was found in whole Diamond Satellite City Island and some areas around industrial site areas. Furthermore, it should be noticed that some isolated growth during 2002-2009 were become clustered branch due to some industrial sites had finished some part of their constructions like PPSEZ and some industrial factories. Isolated growth was also found on the northern part of ring road where most of the growth were factories.

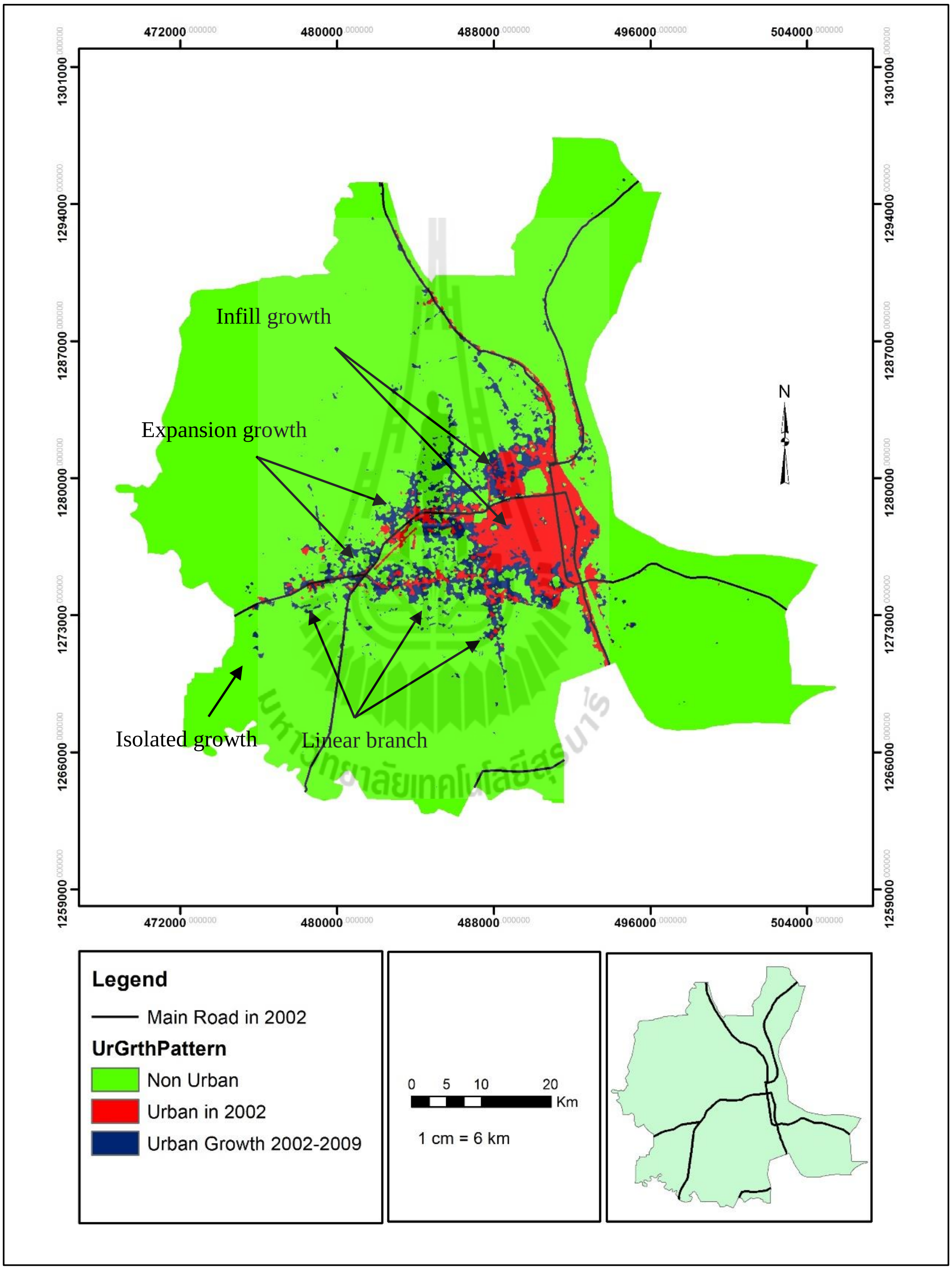


Figure 4.8 Spatial pattern of urban growth from 2002 to 2009.

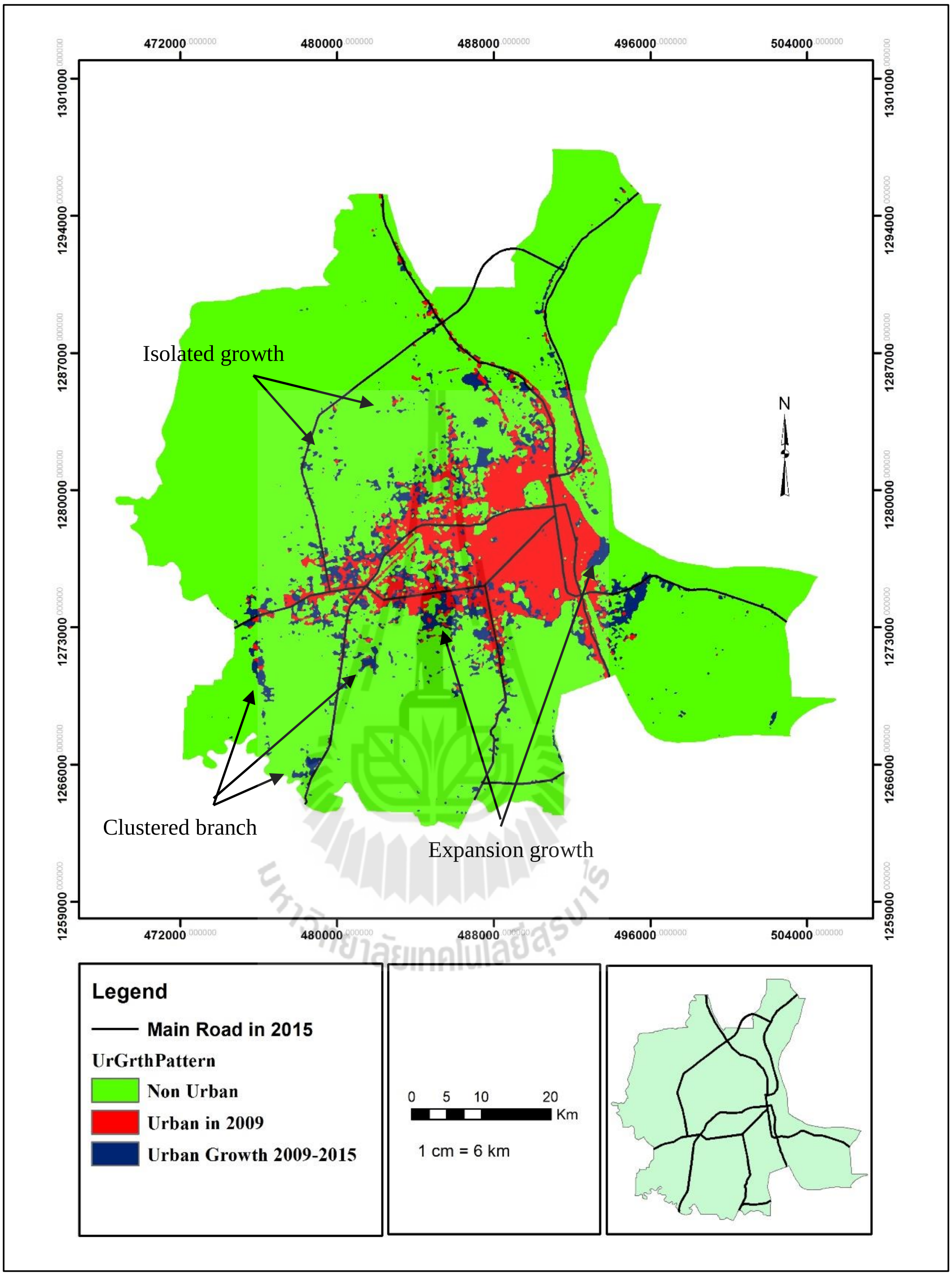


Figure 4.9 Spatial pattern of urban growth from 2009 to 2015.

4.3 Urban growth simulation and influential factor identification

Under this section, input data and multicollinearity test are prepared to input into logistic regression model for three periods: 2002-2015, 2002-2009, and 2009-2015. Appendix C provides an example of input LR model in SPSS statistic software.

4.3.1 Input data for logistic regression model

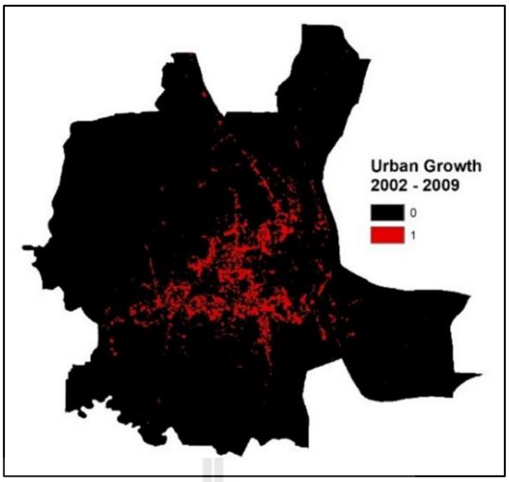
As describe in section 3.2.4, there are 16 factors that are included into the LR simulation model for three different peroids. Table 4.6 summarizes the list of independent factors with abbreviation for using in the model. First, all 16 variables were entered in the regression model in order to test their collinearity values and to find out only the significance variables. Figure 4.10 shows dependent variable maps and Figure 4.11, 4.12, and 4.13 are a factor maps of 2002, 2009, and 2015 that were used as input for finding the relationship between independent and dependent variables. Only 12 factors maps for 2009 and 2015 are shown in Figure 4.12 and 4.13 because other factors are same as factors of 2002.

Table 4.6 Input independent and dependent variables used in logistic regression model.

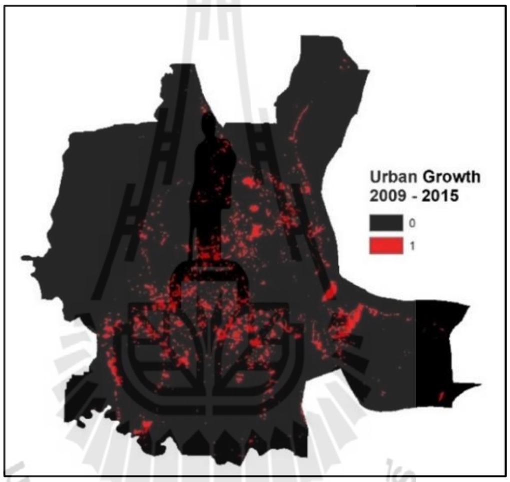
Factor	Abbreviation	Meaning	Data type
Dependent	Y	1 = Urban growth 0 = Non-urban growth	Dichotomous
Independent	X		
Bio-physical factor	VEGET	1:vegetation, 0:not vegetation	Dichotomous
	MISCELL	1:miscellaneous land, 0:not miscellaneous land	Dichotomous

Table 4.6 Input independent and dependent variables used in logistic regression model (Continued).

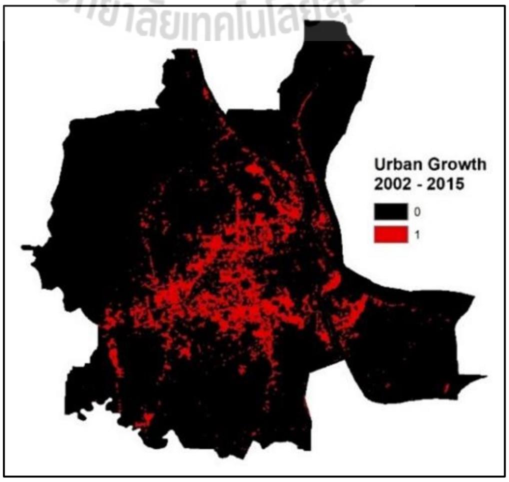
Bio-physical factor	ARABLE	1:arable land, 0:not arable land	Dichotomous
	DistRiver	Distance to river	Continuous
Social factor	DenPop	Population density (people/sq.km)	Continuous
Proximity and accessibility factor	DistCBD	Distance to CBD	Continuous
	DistHealth	Distance to health center	
	DistSchool	Distance to public school	
	DistMroad	Distance to main road	
	DistRoad	Distance to minor road	
	DistRailLine	Distance to railway line	
	DistAirport	Distance to Phnom Penh airport	
	DistIndustry	Distance to industrial area	
	DistIndPort	Distance to industrial river port	
Neighborhood factor	Num_Urban	Number of urban cells within a 7x7 cell window	Continuous
	DistUrban	Distance to existing urban cluster	



(a)



(b)



(c)

Figure 4.10 Dependent variable maps of (a) 2002-2009, (b) 2009-2015 and (c) 2002-2015.

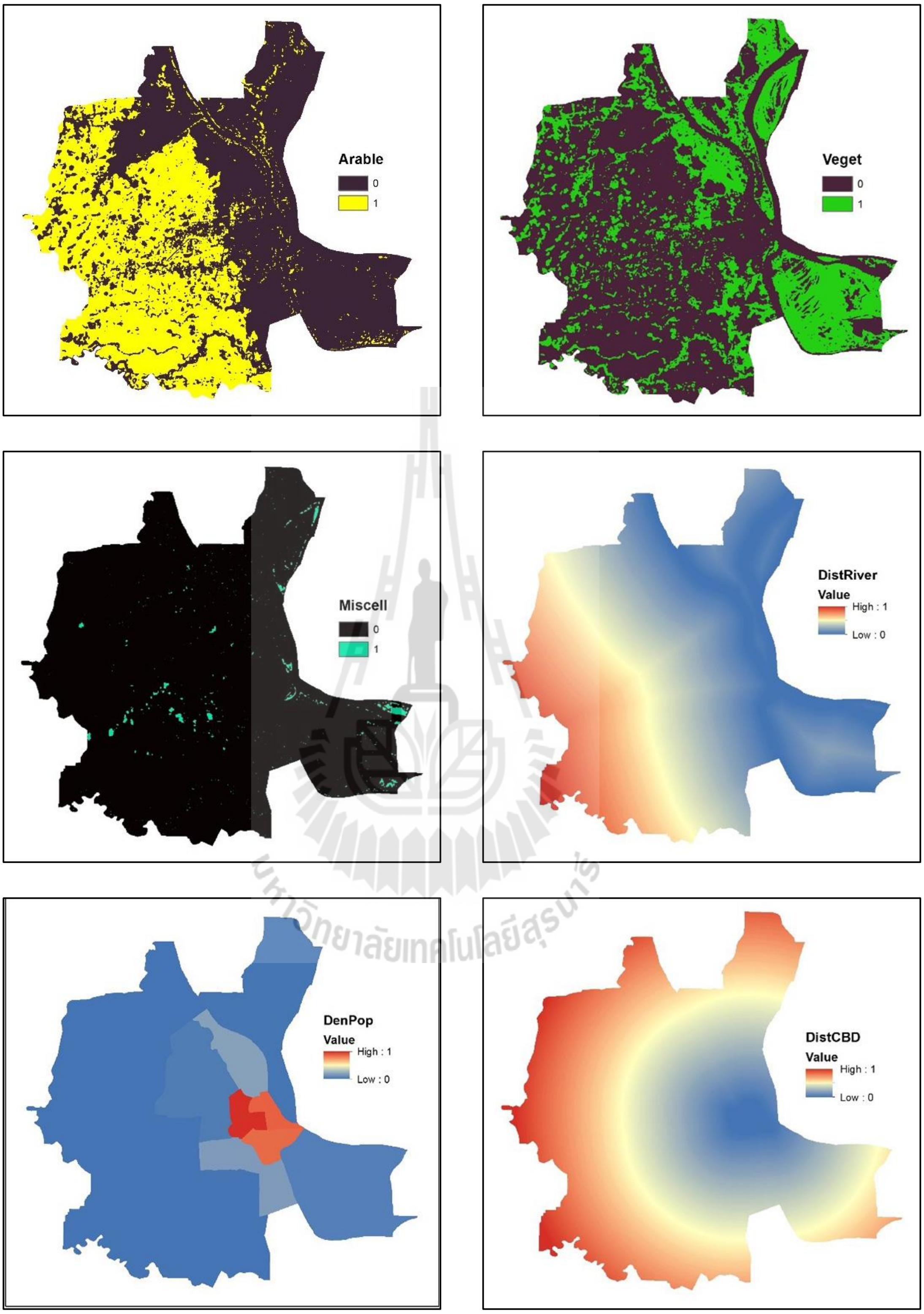


Figure 4.11 Factor maps of urban growth in 2002.

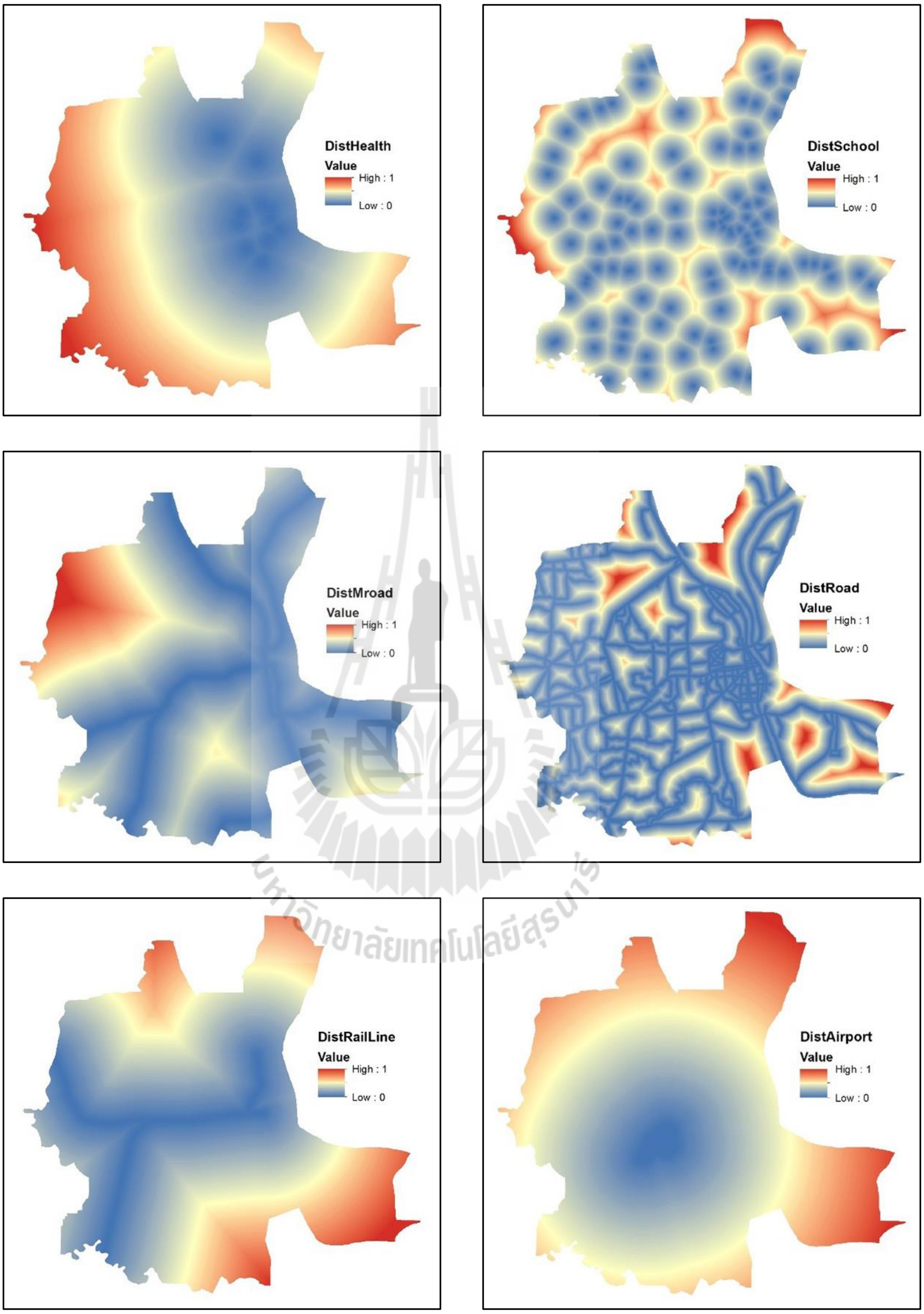


Figure 4.11 Factor maps of urban growth in 2002 (Continued).

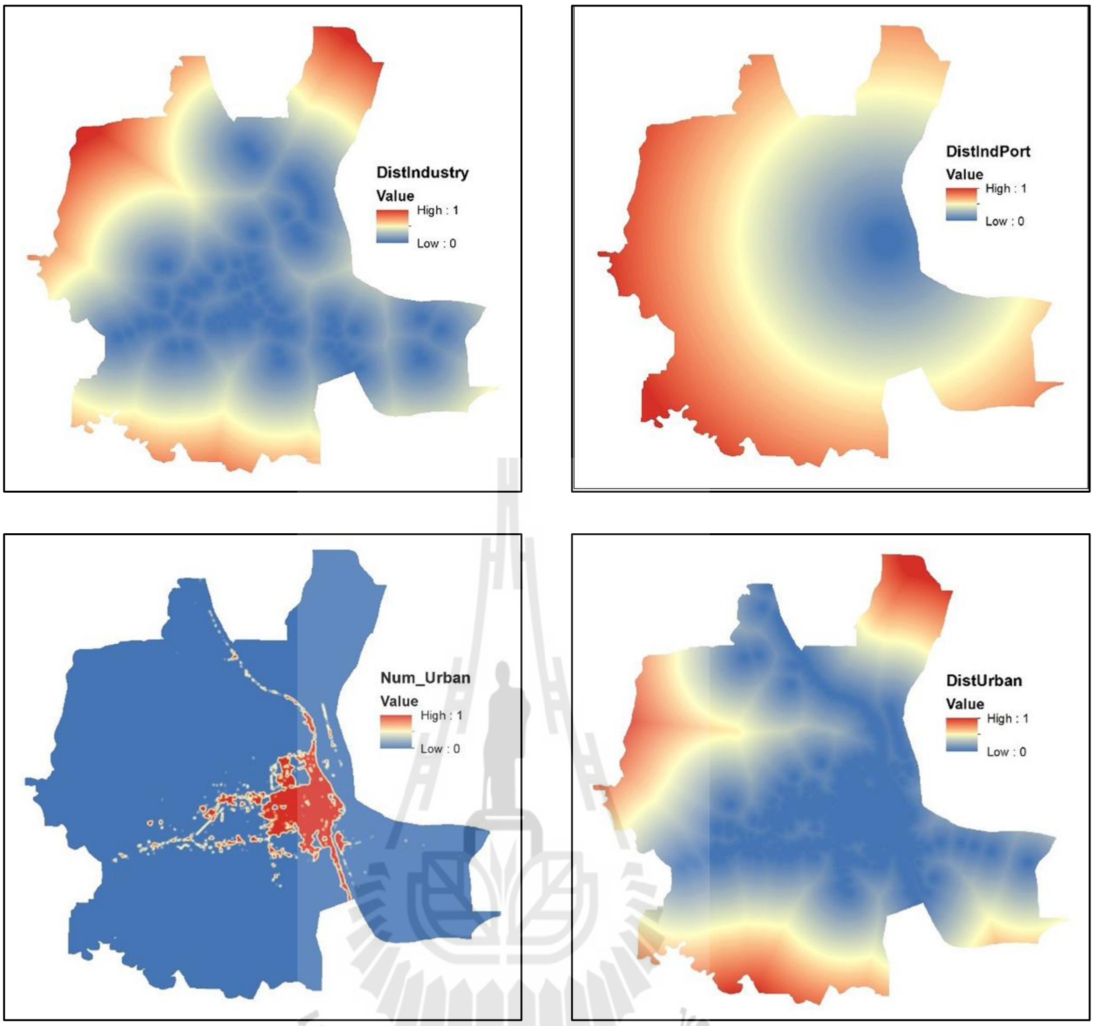


Figure 4.11 Factor maps of urban growth in 2002 (Continued).

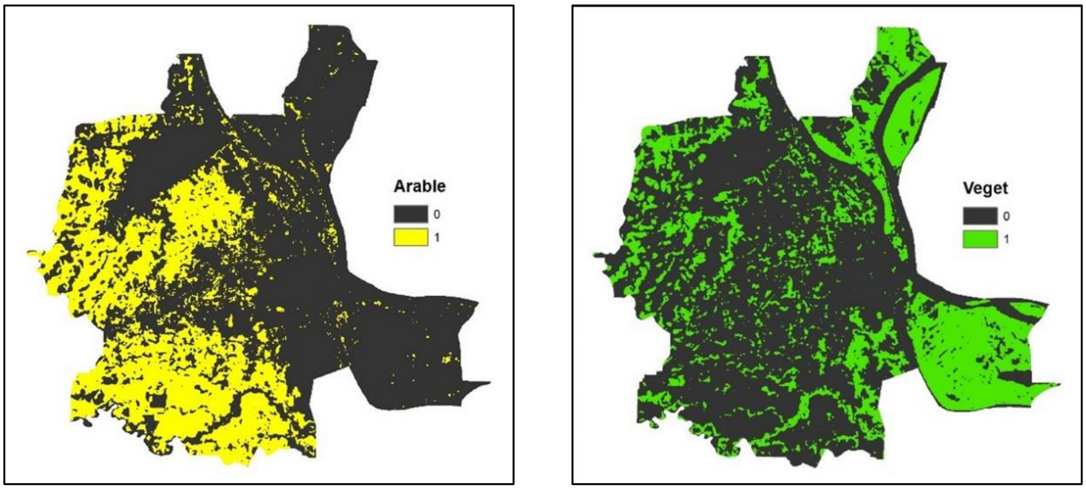


Figure 4.12 Factor map of urban growth in 2009.

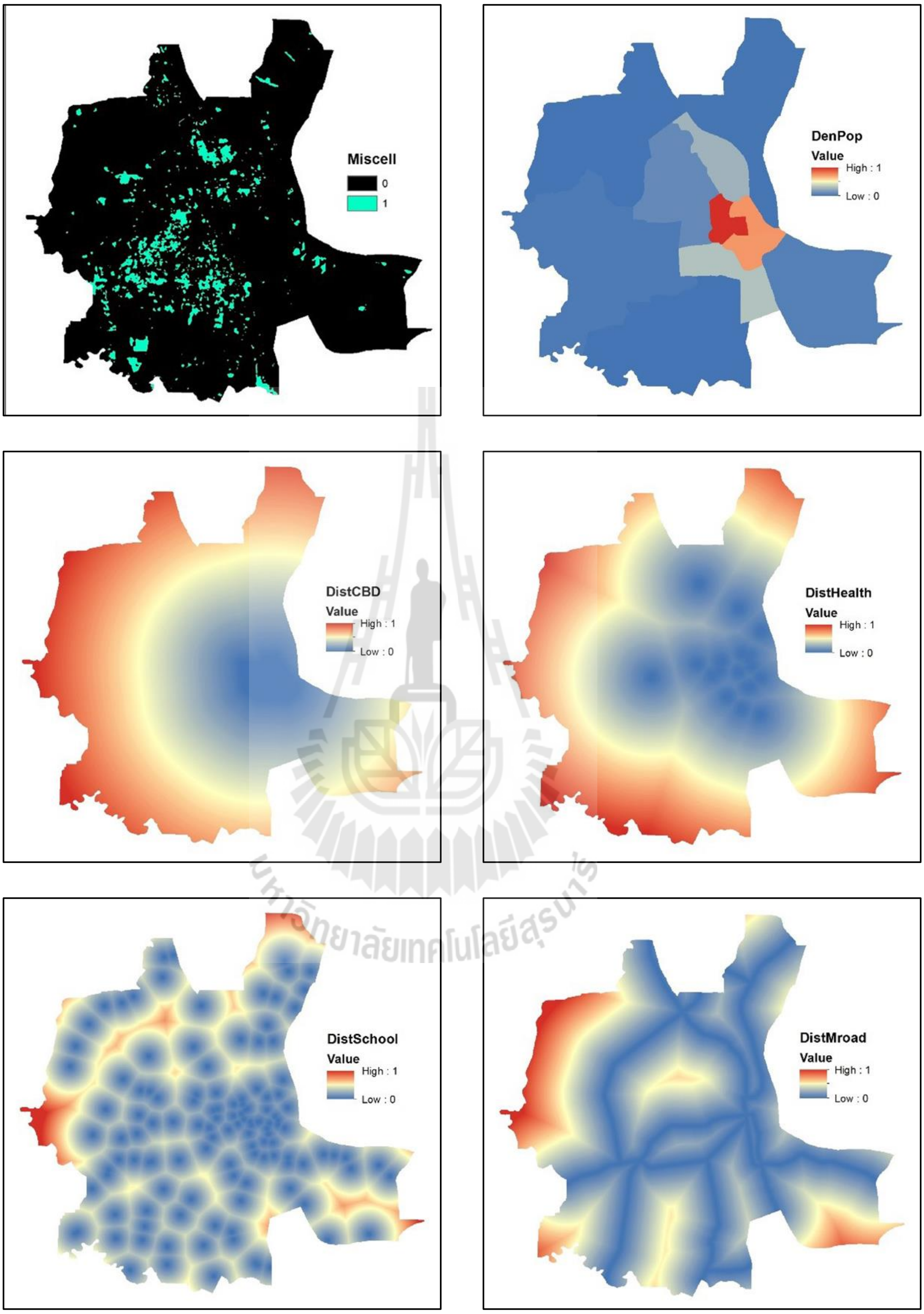


Figure 4.12 Factor map of urban growth in 2009 (Continued).

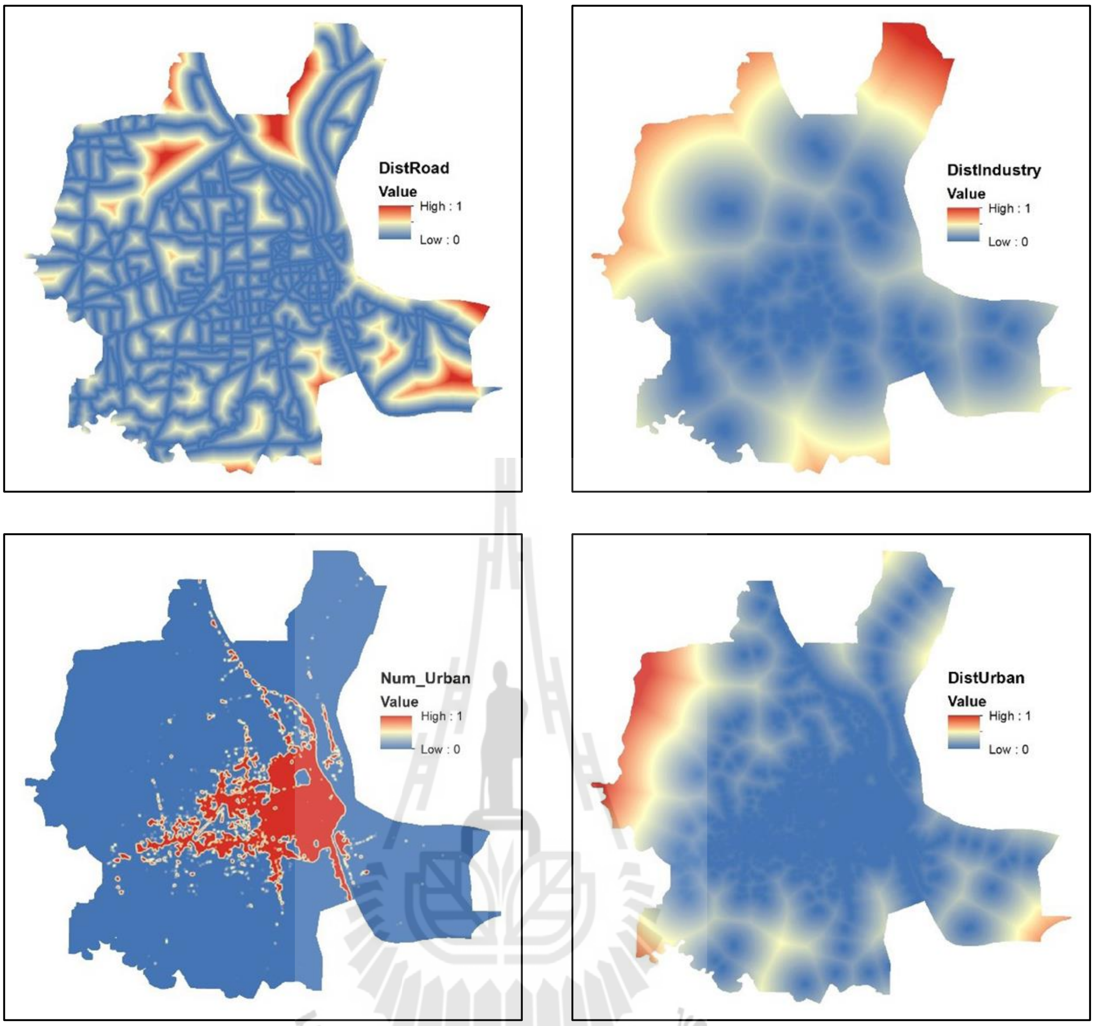


Figure 4.12 Factor map of urban growth in 2009 (Continued).

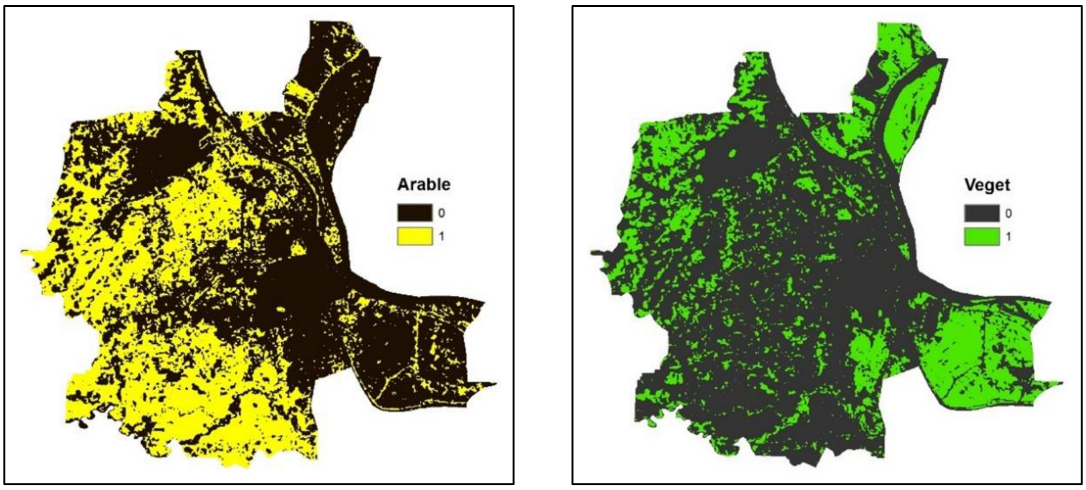


Figure 4.13 Factor maps of urban growth in 2015.

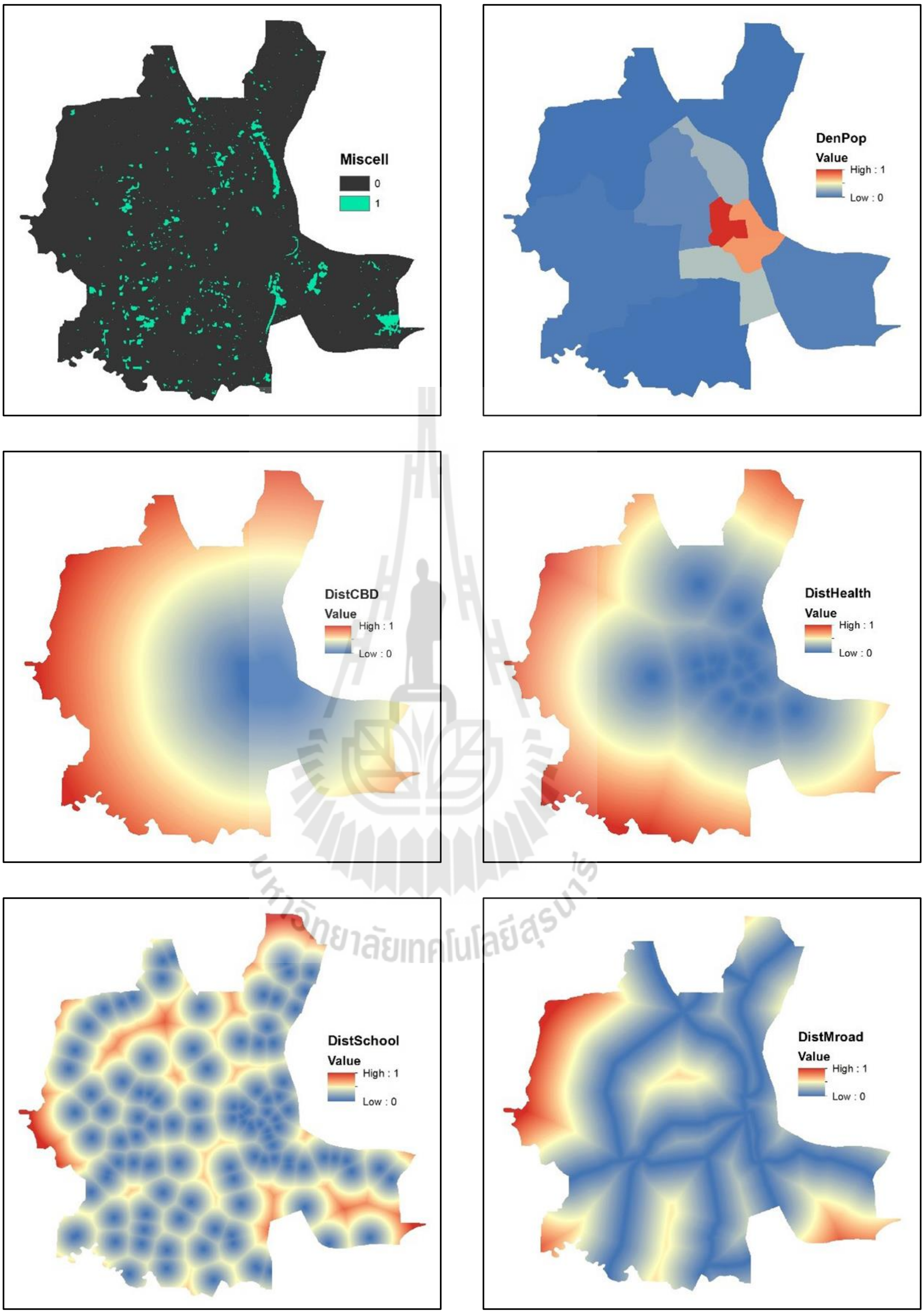


Figure 4.13 Factor maps of urban growth in 2015 (Continued).

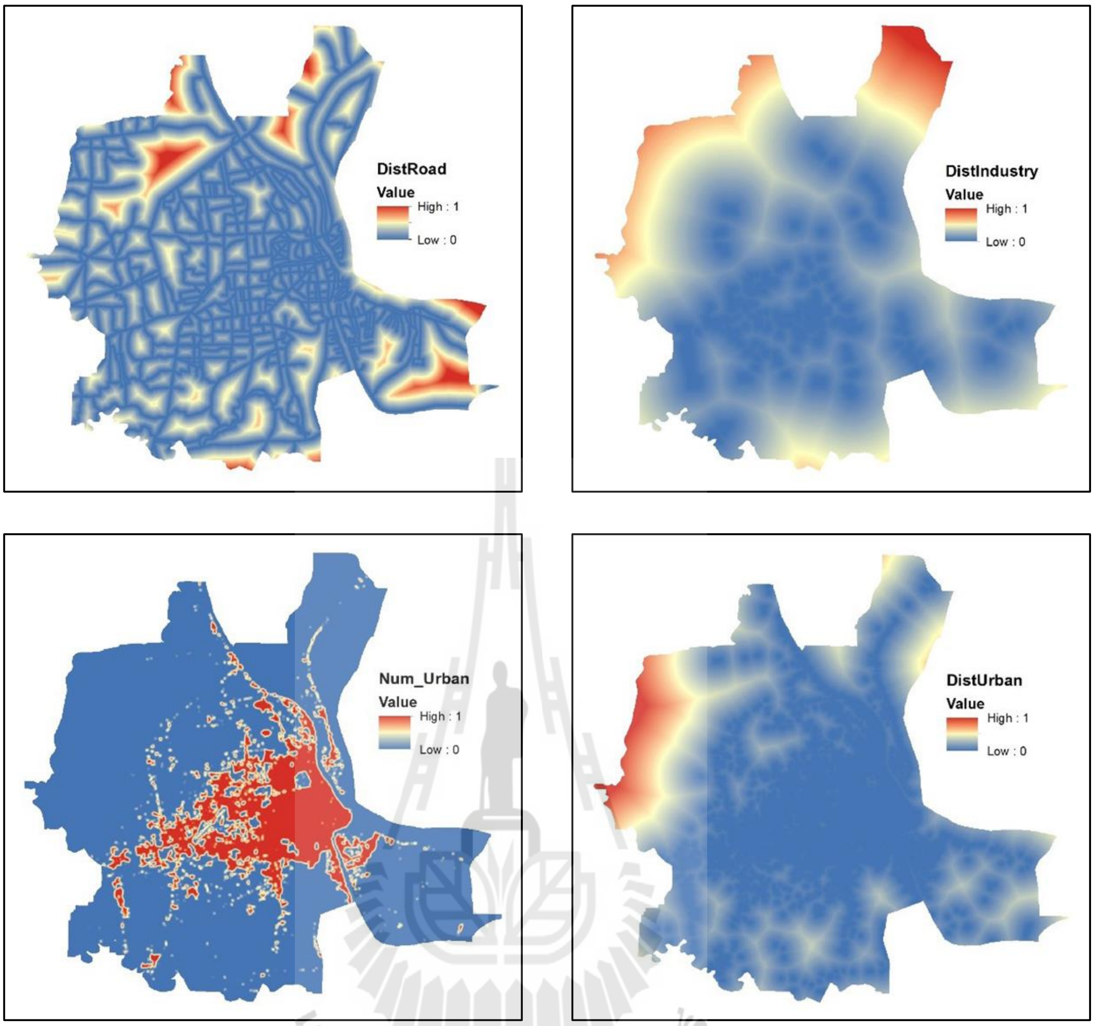


Figure 4.13 Factor maps of urban growth in 2015 (Continued).

4.3.2 Multicollinearity effects

Multicollinearity among the variables may cause result of LR model become poor and bias, so that before performing the LR anlysis, it is required to check the correlation among each independent variable by observing the result of TOL or VIF (Section 3.2.1). If one variable has VIF > 10, it should be removed from the analysis as it is indicated that the variable has a highly correlated which can be predicted from the other variables. Therefore, VIF test was checked for 2002-2009, and 2009-2015. The reason for unchecking the 2002-2015, because it was already checked by the factor maps of 2002 which is also performed in 2002-2009. The result of checking showed that DistRiverPort and DistRiver are highly correlated with VIF of 76.3 and 46.4 for 2002 and 58.8 and 38.8 for 2009 respectively (Table 4.7 and 4.9). After the observation, the top two highest VIF were removed from the checking since value are ridiculously high. The test of VIF was repeated again, so then Table 4.8 and 4.10 gave both VIF test among all 14 varaible with no idenfication of high VIF value. Therefore, all the 14 remaining variables were safely used for further analysis of LR model in the next step.

Table 4.7 Result of multicollinearity test of independent variables in 2002.

Model 2002 - 2015		**Collinearity statistics**	
N	**Variables**	**Tolerance**	**VIF**
1	Arable_2002	0.327	3.058
2	Veget_2002	0.493	2.030
3	Miscell_2002	0.935	1.069
4	DistRiver	0.022	46.399 (Remove)
5	DenPop_2002	0.433	2.311
6	DistCBD_2002	0.028	35.313
7	DistHealth_2002	0.035	28.381
8	DistSchool_2002	0.678	1.474
9	DistAirport	0.068	14.695
10	DistIndustry_2002	0.13	7.705
11	DistRiverPort	0.013	76.299 (Remove)
12	DistMroad_2002	0.598	1.673
13	DistRoad_2002	0.667	1.499
14	DistRailway	0.121	8.245
15	DistUrban_2002	0.172	5.802
16	Num_Urban_2002	0.428	2.336

Table 4.8 New result of multicollinearity test of independent variables in 2002.

	Model 2002 – 2015 new	Collinearity Statistics	
N	**Variables**	**Tolerance**	**VIF**
1	Arable_2002	0.328	3.047
2	Veget_2002	0.495	2.020
3	Miscell_2002	0.938	1.067
4	DenPop_2002	0.444	2.253
5	DistCBD_2002	0.102	9.830
6	DistHealth_2002	0.143	7.005
7	DistSchool_2002	0.733	1.364
8	DistAirport	0.266	3.758
9	DistIndustry_2002	0.131	7.620
10	DistMroad_2002	0.622	1.609
11	DistRoad_2002	0.679	1.473
12	DistRailway	0.448	2.233
13	DistUrban_2002	0.174	5.745
14	Num_Urban_2002	0.437	2.290

Table 4.9 Result of multicollinearity test of independent variables in 2009-2015.

	Model 2009 - 2015	Collinearity Statistics	
N	**Variables**	**Tolerance**	**VIF**
1	Arable_2009	0.373	2.678
2	Veget_2009	0.492	2.034
3	Miscell_2009	0.715	1.399
4	DistRiver	0.026	38.804 (Remove)
5	DenPop_2009	0.485	2.061
6	DistCBD_2009	0.039	25.880
7	DistHealth_2009	0.064	15.521
8	DistSchool_2009	0.58	1.724

Table 4.9 Result of multicollinearity test of independent variables in 2009-2015 (Continued).

9	DistAirport	0.088	11.351
10	DistIndustry_2009	0.294	3.398
11	DistRiverPort	0.017	58.851 (Remove)
12	DistMroad_2009	0.278	3.599
13	DistRoad_2009	0.697	1.435
14	DistRailway	0.087	11.431
15	DistUrban_2009	0.173	5.792
16	Num_Urban_2009	0.414	2.414

Table 4.10 New result of multicollinearity test of independent variables in 2009-2015.

Model 2009 – 2015 new		**Collinearity Statistics**	
N	**Variables**	**Tolerance**	**VIF**
1	Arable_2009	0.387	2.586
2	Veget_2009	0.515	1.943
3	Miscell_2009	0.72	1.388
4	DenPop_2009	0.497	2.013
5	DistCBD_2015	0.149	6.704
6	DistHealth_2009	0.159	6.305
7	DistSchool_2009	0.672	1.487
8	DistAirport	0.242	4.126
9	DistIndustry_2009	0.367	2.726
10	DistMroad_2015	0.279	3.587
11	DistRoad_2009	0.701	1.427
12	DistRailway	0.315	3.179
13	DistUrban_2009	0.183	5.471
14	Num_Urban_2009	0.418	2.392

4.3.3 Logistic regression model for 2002-2015 period

The model of 2002-2015 was obtained in the first step of backward stepwise method with all 14 variables. According to the Wald statistic test for significance of single predictor in Table 4.11, all 14 independent variables were significance, since the Wald statistics (Sig.) were smaller than assigned level of confidence (0.05). That means all 14 independent varaibles were useful and relevant for predicting the output. From the model output of SPSS, the report of LL_0 was 504848.106 and LL_m was 322511.506, thus the pseudo Nagelkerke R^2 value was 0.439. This indicates that the 2002-2015 model performs well, and the 14 independent variables could efficiently interpret the relationship of urban growth pattern of Phnom Penh. The equation of LR model is shown as below:

$$
\begin{aligned}
Y_{(2002\text{-}2015)} = {} & 0.92 + 1.656\ \text{Arable\_2002} + 1.690\ \text{Veget\_2002} \\
& + 1.984\ \text{Miscell\_2002} + 1.123\ \text{DenPop\_2002} \\
& - 3.351\ \text{DistCBD\_2002} - 0.448\ \text{DistHealth\_2002} \\
& - 0.439\ \text{DistSchool\_2002} - 4.155\ \text{DistAirport} \\
& - 3.624\ \text{DistIndustry\_2002} - 2.012\ \text{DistMroad\_2002} \\
& - 3.965\ \text{DistRoad\_2002} + 0.662\ \text{DistRailway} \\
& - 2.268\ \text{DistUrban\_2002} - 2.129\ \text{Num\_Urban\_2002} \qquad (4.1)
\end{aligned}
$$

The result of LR model indicated that all varibles have different degree of influence on the probability of urban growth. It was distinguished by observing the value of the coefficient (*b*) where the positive and negative values determined the phenomena of how strong or weak of the relationship between urban growth and the nature value of independent variables. Moreover, *OR* also results from each coefficient value, since the formula is e^b, and it is used to assist for comparing the probability of occurrence between two events by increasing a one unit change in predictor *x*, so that the effect of occurrence can be easily interpreted. Based on Table 4.11, it was seen that arable land, vegetation, miscellaneous land, population density and distance to railway have positive correlation with urban growth, while the remaining nine variables: distance to CBD, health, school, airport, industry, main road, road, and urban, and number of urban within 7x7 window cells have a negative correlation with urban growth. This means that the higher the distance from the negative correlation variables, the lower the probability of urban growth to occur, or in other words urban growth tends to occur in closed proximity of those variables, except number of urban within 7x7 window cells because its probability value is opposite the value of Euclidean distance which is the farther the distance from the variables, the higher the probability value, while number of urban within 7x7 window cells itself has the higher value at the proximity of urban.

The probability of urban growth in arable land is almost the same probability of urban growth in areas covered with vegetation, and the probability of urban growth of miscellaneous land is larger than arable land and vegetation. This can be seen by the *OR* values of 5.241 and 5.419 for arable and vegetation and 7.274 for miscellaneous land. The result seems true since the change detection matrix of 2002-2015 in Table 4.5 also identified the same percentage of change between arable land and vegetation approximately 12% and miscellaneous land about 20% of their whole areas. As their *OR* are also high, this can be notice that the development of urbanization are required very high on the land in Phnom Penh, mostly on the miscellaneous land which is bare land. Population density was indicated that urban growth tends to occur in an area of high population density with *OR* of 3.075 which is more than 1, reflecting the occurrence of urban will be increased 3.075 times with the unit increment in population density. This strong positive of population density seems reasonable, as pattern of Phnom Penh urban growth of 2002-2009 and 2009-2015 periods shown in Figure 4.8 and 4.9. They illustrated that most of the infill growth were appeared clearly on high population density district like some vacant land in downtown area and Toul Kork distric in the northern CBD, where most of urban at that time were one single family house mixed with arable land in the whole district. Therefore, urban were grew by converting the arable land and some areas in downtown to become congested urban. Moreover, the last positive influential factor is distance to railway line, and it showed that urban area tends to grow further away from railway line, as it gives *OR* of 1.938 which mean that if urban development is 1 km further away from the railway line then the probability of growth is increased about 1.938 larger than urban growth closer to the railway line. This is quite reasonable because after obervation the pattern of urban growth, there were no amount of growth laying along the railway line to the north and to the south of Phnom Penh. Even though it has some parts of railway line across the urban, urban tends to grow in opposite direction and leave some open space which is the buffer from the railway line. However, in reality some urban poor settle in small cottage along railway line.

For other distance factors concerning with the negative correlation with urban growth, the model shows that during period of 2002-2015, urban growth is significantly influenced by the distance to airport that indicates *OR* of 0.016 or 1/62.5. This means that the *OR* of urban growth in an area 1 km closer to the airport is 62.5 times as large as the *OR* of urban growth in an area 1 km further away from the airport. This result obviously demonstrates that the trend of urban pattern growth are continued westward to the airport which is located in the middle of Phnom Penh. Similary, result also reveals that urban growth has been controlled by road and main road accessibility. The *OR* for distance to road is 0.019 or 1/52.63 and for distance to main road is 0.134 or 1/7.42 which indicates that an area 1 km closer to road and main road is estimated as 52.63 and 7.42 times as large as the *OR* of urban growth in an area 1 km further away from road and main road respectively. This has contributed to the spatial pattern of the linear branch development, especially along the main road to the airport and some minor road outside the core existing urban area. Furthermore, distance to CBD and distance to industrial area also give a similar influence to the expansion of urban with *OR* of 0.035 or 1/28.57 and 0.027 or 1/37 respectively. CBD is located in the core urban area where some vacant land are available for development, so that the vacant land closer to CBD have high probability to transform for urban use, and it also demonstrates the occurrence of infilling growth in the CBD. Moreover, the reason why CBD is so influenced to urban growth because of the development of satellite city on the Diamond Island that seems closed to the CBD area, so that urban growth was occurred in the whole island by transforming its vegetation land. Around industrial zone in the suburban Phnom Penh and in the southern part of airport were developed very fast, since the economics of Phnom Penh and labor work are depended on the industry sector which the majority are garment factories. Many urban house facilities, accommodations, shop, restaurant, park and residential building for labors were built around the garment factories along the Veng Sreng road for rent and for sales, and they were become larger than the industry itself. This has demonstrated to the pulling factor of the urbanization in Phnom Penh, since it attracted more rural migration for job employment and quality of life. Distance to health center and school do not give much high impact to the cause of urban growth, since the *OR* are 0.639 or 1/1.56 and 0.645 or 1/1.55. Their degrees of influence are the lowest among all, because school, according to the topographic map, are located in almost every village in Phnom Penh. All the villages in the rural area are identified as garden villages, that means they are mixed with trees and some vegetation. Since the presence of school scatters in every location, there are some schools that closer and some are farther to the urban growth sample points. So that it does not give much impact to the urban growth. Distance to health center, referred to the national hospital at that time in 2002, were not so many, and more than half of them were situated in the downtown area and the remaining were located in the northern, thus the influence is effected mostly on the northern Phnom Penh.

Urban areas tend to grow closer to the nearest urban cluster that is shown in the model that distance to urban cluster has a coefficient of -2.268 with *OR* of 0.103 or 1/9.7. This means that probability of urban growth in an area 1 km close to the urban cluster is estimated as 9.7 times larger than the probability of urban growth in an area 1 km further away. Conversly, number of urban within a 7x7 cell window also gives a negative influence correlation with a coefficient value of -2.129 and *OR* of 0.119 or 1/8.4, but the nature of impact is different and is interpreted that urban is more likely to be growth further away from the neighboring urban fringe by increasing the 1 unit of urban cell further away within its neighborhood. That means the probability of growth is highly occurred outside the urban fringe about 8.4 times larger than clustering around the existing urban areas. The result seems reasonable, because of the large scale of urban expansion of Phnom Penh during 2002-2015 were huge, which means the expansion spreaded over the existing neighborhood cells, and as a result, the spatial pattern of growth can be seen that there were many outlying growth (scattering growth) more than infilling and expansion growth that occurred around the existing urban in Figure 4.8 and 4.9. The output probability map of LR model during 2002-2015 is presented in Figure 4.14.

Table 4.11 Variables in the equation of LR model for 2002-2015 period.

N	Variables	b	SE_b	Wald	df	Sig.	Exp(b)
1	Arable_2002	1.656	0.019	7,569.181	1	0.000	5.241
2	Veget_2002	1.690	0.018	9,323.729	1	0.000	5.419
3	Miscell_2002	1.984	0.036	3,107.967	1	0.000	7.274
4	DenPop_2002	1.123	0.055	415.490	1	0.000	3.075
5	DistCBD_2002	-3.351	0.065	2,666.872	1	0.000	0.035
6	DistHealth_2002	-0.448	0.059	57.646	1	0.000	0.639
7	DistSchool_2002	-0.439	0.044	98.855	1	0.000	0.645
8	DistAirport	-4.155	0.044	8,857.583	1	0.000	0.016
9	DistIndustry_2002	-3.624	0.091	1,600.025	1	0.000	0.027
10	DistMroad_2002	-2.012	0.050	1,602.164	1	0.000	0.134
11	DistRoad_2002	-3.965	0.052	5,817.872	1	0.000	0.019
12	DistRailway	0.662	0.039	289.141	1	0.000	1.938
13	DistUrban_2002	-2.268	0.090	634.178	1	0.000	0.103
14	Num_Urban_2002	-2.129	0.032	4,412.369	1	0.000	0.119
	Constant	0.920	0.026	1,260.478	1	0.000	2.510

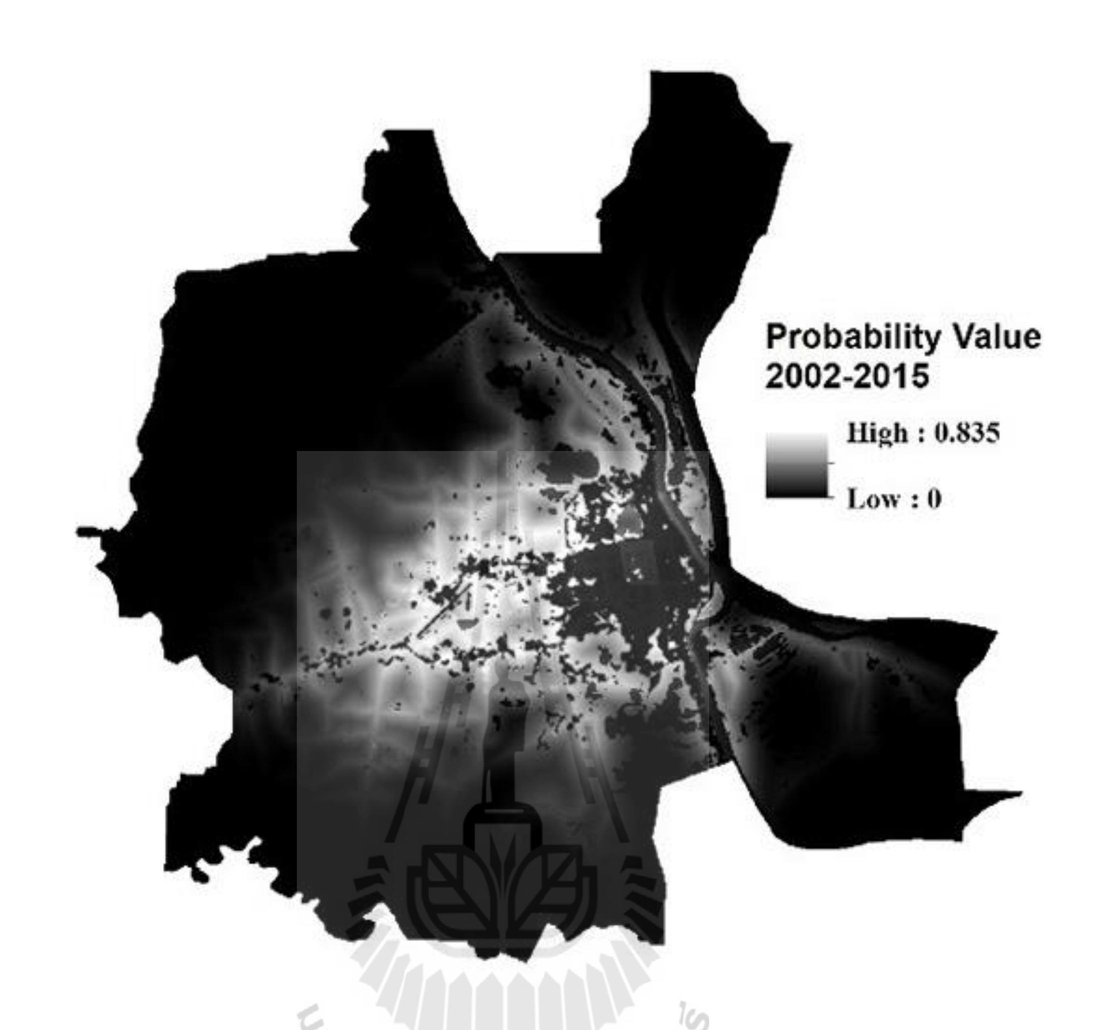


Figure 4.14 Urban growth probability map of Phnom Penh during 2002 to 2015.

4.3.4 Logistic regression model for 2002-2009 period

The short term model of 2002-2009 period was also obtained in the first step of backward stepwise method with all 14 variables. The Wald statistic test for significance of single predictor in Table 4.12 shows that all 14 independent variables were significance, since the Wald statistics (Sig.) were smaller than assigned level of confidence (0.05). That means all 14 independent varaibles were useful and relevant for predicting the outcome. From the model output of SPSS, the report of LL_0 was 288,420.469 and LL_m was 157,905.458, thus the pseudo Nagelkerke R^2 value was

0.5. This indicates that the 2002-2009 model performs well, and the 14 independent variables could efficiently interpret the relationship of urban growth in Phnom Penh. The equation of LR model is shown as below:

$$
\begin{aligned}
Y_{(2002\text{-}2009)} = & -1.50 + 3.725\ \text{Arable\_2002} + 3.801\ \text{Veget\_2002} \\
& + 3.465\ \text{Miscell\_2002} + 1.378\ \text{DenPop\_2002} \\
& - 1.468\ \text{DistCBD\_2002} - 3.175\ \text{DistHealth\_2002} \\
& - 1.686\ \text{DistSchool\_2002} - 4.829\ \text{DistAirport} \\
& - 6.320\ \text{DistIndustry\_2002} - 0.502\ \text{DistMroad\_2002} \\
& - 4.506\ \text{DistRoad\_2002} + 0.333\ \text{DistRailway} \\
& - 10.565\ \text{DistUrban\_2002} + 0.324\ \text{Num\_Urban\_2002} \qquad (4.2)
\end{aligned}
$$

According to Table 4.12, the model shows that there are 7 independent variables that show a positive correlation relationship with urban growth include arable land, vegetation, miscellaneous land, population density, distance to main road, distance to railway line and number of urban within a 7x7 cell window. All three land cover types show the significance role of controlling the urban growth area. The probability of growth is very high for vegetation with *OR* of 44.761, arable land with odd ratio of 41.462 and miscellaneous land with *OR* of 32. This means that the possibility of transformation to urban for vegetation at that time is larger than arable land and miscellaneous land, as the spatial pattern of growth also identified that most of vegetation below and upper the existing urban and arable land on the westward Phnom Penh were transformed to urban more remarkable than miscellaneous land. Meanwhile, population density and number of urban within a 7x7 cell window give a moderate influenced value with *OR* of 3.96 and 1.38 respectively, as observation in the spatial pattern of urban growth, almost half part of urban growth occurred inside the core existing urban where the high population density was presented, while another half were sprawl linearly to the center of Phnom Penh where airport was situated in the low population density area. This clarifies that urban in Phnom Penh at that time had faced haphazard growth at the suburban area outside the downtown, where the population density was still low. Furthermore, the model also proves that the growth is most likely to occur in the neighborhood cells of existing urban about 1.38 times as large as it occurs further way from the neighborhood cells. This can be seen in infill growth pattern and some linear branch pattern along the main road where urban growth was occurred in adjacent existing urban cells (see Figure 4.8). Distance to main road and railway has also the moderate influenced on the probability of urban growth during 2002-2009. The further the area to main road and railway line the higher is the probability of urban growth occurrence with the *OR* of 1.65 and 1.4 respectively. This reason is acceptable since the main road in 2002 at that time were used as a transport road for moving people out from the city to the countryside, therefore only 6 main roads were assigned in this model, and majority of urban growth during 2002-2009 seems not to occur along the main road and railway line.

All other distance factors such as distance to CBD, health center, school, airport, industrial area, road and urban cluster have negative correlation to urban growth. Urban growth tends to increase at area close to the existing urban cluster with a coefficient value of -10.565 and the lowest *OR* of 0.000026 or 1/38461.5 which means that the probability of urban growth of the region closer to urban cluster is equal to 38,461.5 times the *OR* of the area further away by 1 km from the nearest existing urban cluster. Therefore, the higher the distance from urban cluster, the lowest possibility of urban growth occurrence. This can be noticed that during 2002-2009, urban expansion in Phnom Penh was highly influenced in the neighborhood area of urban fringe that reflected the expansion pattern. Distance to industrial zone and airport have coefficient equal to -6.32 and -4.829 and *OR* of 0.002 or 1/500 and 0.008 or 1/125 respectively. This result indicates that the *OR* of urban growth in the area near the industrial area and airport are 500 and 125 times as great as the odds of urban growth in an area of 1 km further away from industrial area and airport. This short term period, industry and airport play a significance role after the distance to urban cluster for the occurrence of urban growth. In addition, distance to health center in this period plays as important role for urban growth, unlike the influence in long term period of 2002-2015, and it has the *OR* of 0.042 or 1/23.8. Thus, the possibility of urban growth in a region near the health center is about 24 times as big as the possibility of another area which is farther away from health center by 1 km. Similarly, distance to road still has a correlation of significant influence to the urban growth, as the *OR* is 0.011 or 1/90.9 that its possibility caused urban growth to take place in linear branch pattern which can be found in Figure 4.8. The last two factors are distance to CBD and school which the lowest negative influence among distance factors. They have *OR* of 0.23 or 1/4.37 and 0.185 or 1/5.4 respectively. CBD in this period does not have high impact of urban growth compare to the period of 2002-2015, because most of urban growth were linear branch and expansion growth that situated far away from CBD on the urban fringe and outside the core urban, while the influence of 4.37 times of urban to grow closer to CBD was come from the infill growth. In this period, CBD has lower influence than the school which has odds of occurrence about 5.4 for an area close to school. The output probability map of LR model during 2002-2009 period is presented in Figure 4.15.

Table 4.12 Variables in the equation of LR model for 2002-2009 period.

N	Variables	B	SE_b	Wald	df	Sig.	Exp(b)
1	Arable_2002	3.725	0.042	8,030.984	1	0.000	41.462
2	Veget_2002	3.801	0.042	8,358.283	1	0.000	44.761
3	Miscell_2002	3.465	0.060	3,340.280	1	0.000	31.987
4	DenPop_2002	1.378	0.073	358.709	1	0.000	3.968
5	DistCBD_2002	-1.468	0.109	180.162	1	0.000	0.230
6	DistHealth_2002	-3.175	0.107	873.835	1	0.000	0.042
7	DistSchool_2002	-1.686	0.068	616.011	1	0.000	0.185
8	DistAirport	-4.829	0.070	4,698.223	1	0.000	0.008
9	DistIndustry_2002	-6.320	0.170	1,389.508	1	0.000	0.002
10	DistMroad_2002	0.502	0.084	35.880	1	0.000	1.652
11	DistRoad_2002	-4.506	0.091	2,471.607	1	0.000	0.011
12	DistRailway	0.333	0.068	23.827	1	0.000	1.395
13	DistUrban_2002	-10.565	0.232	2,069.838	1	0.000	0.000026
14	Num_Urban_2002	0.324	0.049	43.008	1	0.000	1.383
	Constant	-1.500	0.054	764.741	1	0.000	0.223

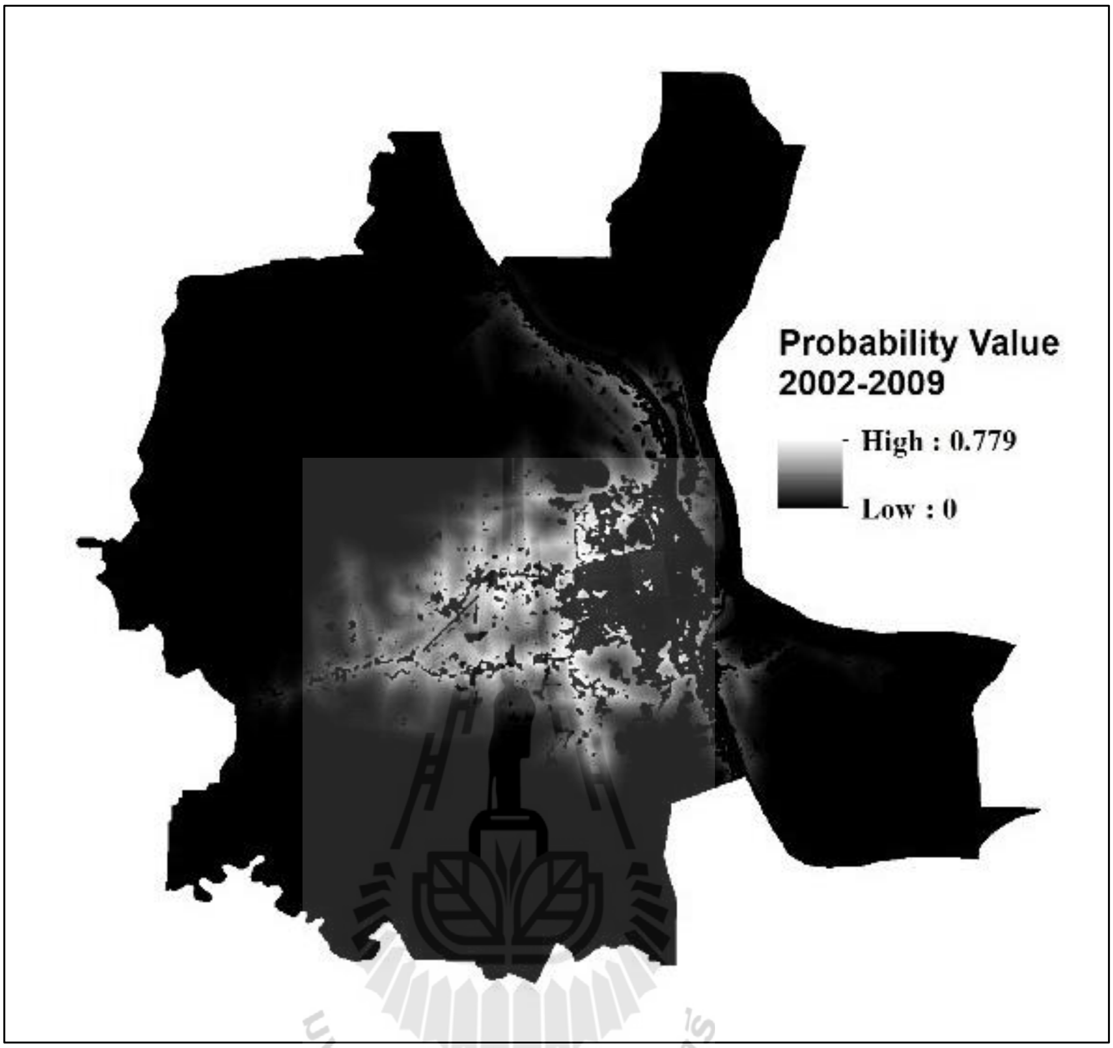


Figure 4.15 Urban growth probability map of Phnom Penh during 2002 to 2009.

4.3.5 Logistic regression model for 2009-2015 period

The next short term model of 2009-2015 was obtained in the second step of backward stepwise method with 13 variables that were proved as significance and relevant for predicting the relationship between independent variables and urban growth. The Wald statistic test for significance of single predictor in Table 4.13 shows that 13 independent variables were significance, while population density was not significance, since the Wald statistics (Sig.) was bigger than assigned level of confidence (0.05). That means it was removed as insignificance and the model was ran again with only qualified 13 independent variables, and the result is shown in Table 4.14 where all 13 variables were significance with value smaller than 0.05. From the model output of SPSS, the report of LL_0 was 328793.419 and LL_m was 222673.319, thus the Pseudo Nagelkerke R^2 value was 0.37. This indicates that the LR model of 2009-2015 period performs well, and the 13 independent variables could efficiently interpret the relationship of urban growth in Phnom Penh. The equation of LR model is shown as below:

$$
\begin{aligned}
Y_{(2009\text{-}2015)} = & -1.499 + 1.321\ \text{Arable\_2009} + 1.834\ \text{Veget\_2009} \\
& + 2.979\ \text{Miscell\_2009} - 2.973\ \text{DistCBD\_2015} \\
& + 1.943\ \text{DistHealth\_2009} - 0.322\ \text{DistSchool\_2009} \\
& - 0.159\ \text{DistAirport} - 2.782\ \text{DistIndustry\_2009} \\
& - 0.478\ \text{DistMroad\_2009} - 1.095\ \text{DistRoad\_2009} \\
& - 1.304\ \text{DistRailway} - 20.095\ \text{DistUrban\_2009} \\
& - 1.099\ \text{Num\_Urban\_2009} \qquad (4\text{-}3)
\end{aligned}
$$

According to Table 4.14, there are 4 factors that have positive correlation with urban growth include arable land, vegetation, miscellaneous land and distance to health center. For the group of land cover type, their *OR* are 3.75 for arable land, 6.26 for vegetation and 19.7 for miscellaneous land. The result seems true according to the change detection in Table 4.4, since percentages of change from miscellaneous land to urban and built-up area is very high about 30% of entire miscellaneous land area followed by vegetation about 5% of entire vegetation area and the last one is arable land about 3% of entire arable land area during period of 2009-2015. Therefore, among land cover type, miscellaneous land, the majority are bare land, is the most influential factor for possibility of urban occurrence during 2009-2015, because it was transformed for the preparation of the construction site for buildings. However, it should be noticed that in 2002-2015 period, miscellaneous land has less significance of influenced than 2009-2015 which means that the construction sector of Phnom Penh was boomed from 2009 like the industrialization of PPSEZ, the residential building, and the development of satellite cities. Moreover, distance to health in this period shows the positive influence, unlike in 2002-2009 period which is negative influence. Health center in 2009, refered to national hospital and private clinic and hospital, were increased and located inside the downtown, while only two or three were located at the surburban area. From this reason, urban tends to occur outside the downtown which is farther from the health center and gives the *OR* of 7 which means possibility of urban growth in an area 1 km further away from health center is 7 times larger than urban growth in an area near the health center.

There are nine factors showing the negative correlation to the urban growth include distance to CBD, school, airport, industrial area, main road, road, railway and existing urban cluster and number of urban within a 7x7 cell window. Result reveals that urban growth process are more likely to occur in areas close the urban clusters, as the distance to existing urban cluster has a coefficient of -20.1, which means that increasing the distance from the urban areas will decreases the urban growth probability. The *OR* of occurrence of urban growth in an area farther away from the urban cluster is almost zero, that reflects the *OR* of urban growth in an area close to urban cluster is extremely great. Similarly, distance to CBD is one of the major influence of urban growth in Phnom Penh, since the pattern of growth during 2009-2015 showed that the Diamond Island that close to CBD was developed as a satellite city where major business building and residential building were proceeded and finished the construction, so that the vacant land that distance close to CBD, have high possibility of urban growth occurrence with *OR* of 0.051 or 1/20. Distance to road and main road do not have much influenced to the urban growth, since the *OR* are 0.335 or 1/3 and 0.62 or 1/1.6 which reflects the influence of area near the road and main road have moderate probability of urban growth respectively. This explaination may be acceptable, since the spatial pattern of growth in Figure 4.9 also agreed that there are few pattern of linear branch along the road during 2009-2015 compared to the pattern that are detected with isolated and expansion growth and clustered branch. Distance to industrial area is still a significant function contributed to urban growth, as in the model output shows that the *OR* is 0.062 or 1/16.12. Thus, the possibility of urban growth in a region near the industrial area is still high about 16.12 times the possibility of urban growth in an area further away by 1 km. This demonstrates the increase of industrialization in Phnom Penh, especially the industrial park, urban housing and facilities around those areas. Moreover, distance to railway line and number of urban within a 7x7 cell window have *OR* of 0.272 or 1/3.67 and 0.333 or 1/3 respectively. This means urban tends to occur near the railway line about 3.67 times as large as it occur further away from railway line, because the extension of growth is continued from the 2002-2009 peroid that pushed urban to expand across the railway line, as a result, half part of railway line was found itself in the middle of built up area. Number of urban within a 7x7 cell window has a negative coefficient of -1.1 which reflects that most of urban are not likely to develop near the neighborhood area of existing urban, but its possibility is high about 3 times of occurence further away from the neighborhood area. This is reasonable, as the spatial pattern of urban growth in Figure 14 is illustrated that most of the pattern growth are isolated growth and clustered branch, excepted for some part of expansion and infill growth that can be found in some part of the neighborhood existing urban area. Distance to school and airport have the lowest influence with *OR* of 0.73 or 1/1.37 and 0.85 or 1/1.17 respectively. School in this period also have the same negative coefficient but has lower significance than in 2002-2009 period. For airport, it does not influence particularly the attraction of urban growth for this peroid, since the existing urban of 2009 is situated around the airport already, so that airport seem to be in the middle of urban and built-up area. The output probability map of LR model during 2009-2015 is presented in Figure 4.16.

Table 4.13 Variables in the equation of LR model for 2009-2015 period.

N	Variables	b	SE_b	Wald	df	Sig.	Exp(b)
1	Arable_2009	1.319	0.025	2,806.478	1	0.000	3.738
2	Veget_2009	1.831	0.023	6,272.413	1	0.000	6.240
3	Miscell_2009	2.977	0.024	14,987.897	1	0.000	19.638
4	DenPop_2009	-0.112	0.071	2.453	1	0.117	0.894
5	DistCBD_2015	-3.006	0.064	2,194.059	1	0.000	0.049
6	DistHealth_2009	1.957	0.067	854.035	1	0.000	7.079
7	DistSchool_2009	-0.314	0.069	20.856	1	0.000	0.731
8	DistAirport	-0.148	0.048	9.346	1	0.002	0.863
9	DistIndustry_2009	-2.761	0.082	1,147.581	1	0.000	0.063
10	DistMroad_2009	-0.481	0.062	59.465	1	0.000	0.618
11	DistRoad_2009	-1.091	0.083	174.753	1	0.000	0.336
12	DistRailway	-1.324	0.051	684.951	1	0.000	0.266
13	DistUrban_2009	-20.098	0.223	8,152.369	1	0.000	0.000
14	Num_Urban_2009	-1.085	0.033	1,064.275	1	0.000	0.338
	Constant	-1.487	0.033	2,002.405	1	0.000	0.226

Table 4.14 New variables in the equation of LR model for 2009-2015 period.

N	Variables	b	SE_b	Wald	df	Sig.	Exp(b)
1	Arable_2009	1.321	0.025	2,823.380	1	0.000	3.747
2	Veget_2009	1.834	0.023	6,335.013	1	0.000	6.260
3	Miscell_2009	2.979	0.024	15,008.184	1	0.000	19.668
4	DistCBD_2009	-2.973	0.061	2,408.053	1	0.000	0.051
5	DistHealth_2009	1.943	0.066	857.692	1	0.000	6.982
6	DistSchool_2009	-0.322	0.068	22.118	1	0.000	0.725
7	DistAirport	-0.159	0.048	11.040	1	0.001	0.853
8	DistIndustry_2009	-2.782	0.081	1,193.391	1	0.000	0.062
9	DistMroad_2009	-0.478	0.062	58.669	1	0.000	0.620
10	DistRoad_2009	-1.095	0.082	176.184	1	0.000	0.335
11	DistRailway	-1.304	0.049	711.739	1	0.000	0.272
12	DistUrban_2009	-20.095	0.223	8,146.131	1	0.000	0.000
13	Num_Urban_2009	-1.099	0.032	1,180.699	1	0.000	0.333
	Constant	-1.499	0.032	2,154.482	1	0.000	0.223

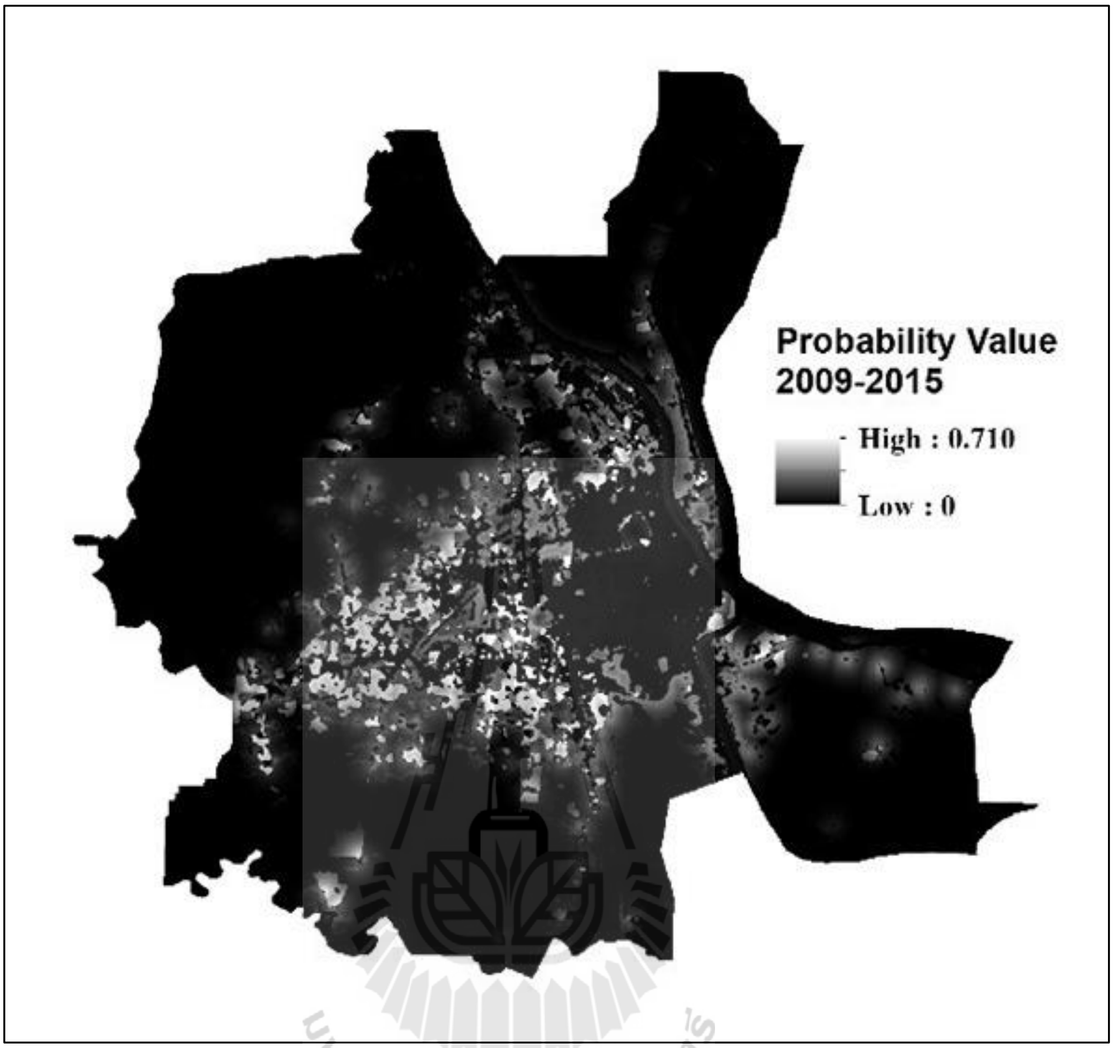


Figure 4.16 Urban growth probability map of Phnom Penh during 2009 to 2015.

In summary, all the important of three LR models are presented in Table 4.15 for the top three highest important factors for each period. The relative influence of influential factors varied along with urban and built up area pattern and periods. No influential factors that are stable impact to the urban growth in all periods, as a result showed that DistAirport and DistIndustry_2002 influenced only on the period of 2002-2009 and long term period of 2002-2015. This because of the large and linear expansion of urban growth along the road from the core existing urban to the suburban area where airport and industrial area situated. However, the influence was change in short period of 2009-2015 that urban growth occurred mostly around the neighborhood area of existing urban, and the development of Diamond Island that located closely to the CBD area, while the transformation of miscellaneous land to urban and built-up area indicated the increasing of construction site for urban development in Phnom Penh.

Table 4.15 Important of influence of factors on urban growth in three different periods.

Top three influential factors	**Period of logistic regression model analysis**		
	2002-2015	**2002-2009**	**2009-2015**
First rank	DistAirport	DistUrban_2002	DistUrban_2009
Second rank	DistRoad_2002	DistIndustry_2002	DistCBD_2009
Third rank	DistIndustry_2002	DistAirport	Miscell_2009

4.4 Urban growth model validation and optimum LR model for urbanization prediction

After getting the result of modeling, all three urban growth models have to be checked for their accuracy of calibration before using for predicting the probability value in the future. There are many ways to evaluate the performance of LR model including contingency table and ROC curve by comparing the actual map of each urban growth with the probability map of each model. Before comparing with actual map, probability map of each model was allocated amount of increasing cells, which are the amount of urban growth of each model period, from the highest probability cell to the lowest probability cells until the allocation process met the number of urban growth cells of each period (2002-2009, 2009-2015 and 2002-2015). So then urban distribution maps based on the model of each period were produced and were used to compare with the actual urban growth map (See Figure 4.10), as a result, contingency tables were made for evaluation of overall accuracy and percentages of correct prediction of urban area, in order to find out how strong each LR model can fit the location between modeling and actual map. Another method is ROC curved that were produced by the SPSS software by comparing the number of probability value of LR model with the binary number of dependent variable which is from urban growth actual map.

4.4.1 Urban growth model validation for 2002-2015 period

According to Table 4.16, the number of urban growth area between 2002 and 2015 is about 70.6 sq. km or 78,437 pixels, with overall accuracy of 91% of urban growth occurrence. This high number of overall accuracy come from the success fit of non-urban about 95% due to the fact that the study area consists of high number of non-urban cells that predicted correctly by the model, but not quite high for the success fit of urban which value is only 58%. This can be explained that the model of logistic regression with 14 influential factors can explain almost 60% of the phenomena of urban growth pattern in Phnom Penh, and it is suggested to fulfill another 40% by including unidentified factors that are not presented in this model, in order to improve the calibration capacity. As it was seen in contingency table and Figure 4.17, 33,015 cells of under prediction were allocated and replaced by over prediction around the region of urbanized area where probability value is higher than the region of isolated and linear growth. This means that most of the under prediction cells are isolated and linear growth, while over prediction cells are the cells laid closely to the infill and expansion growth around the neighborhood existing urban cells, as a result, the allocation of cell would go to the high probability cell value first in descending order from high to low probability value until total of allocation cell equated area of urban growth that leave some part of scattering growth blank (under prediction), since most of cells are already allocated into the infill cells. In conclusion, it is indicated that LR model may be not appropriate for growth scattering types prediction.

At the same time, ROC result reveals that the LR model for 2002-2015 has area under curve of ROC value of 0.903 (Figure 4.18). Pontius and Schneider (2001) states that ROC value of 1 indicates a perfect spatial agreement between actual urban growth map and the predicted probability map and ROC value of 0.5 indicates an agreement expected by chance. The results revealed that the validated calibration for urban growth of year 2002-2015 show good performance and can be accepted for urban growth prediction.

Table 4.16 Contingency table and accuracy for urban growth model of 2002-2015.

Unit in Pixels		**Reality**	
Model	**Urban**	**Non-Urban**	**Total**
Urban	45,422	33,021	78,443
Non-Urban	33,015	648,896	681,911
Total	78,437	681,917	760,354

Correct prediction = 694,318

Wrong prediction = 66,036

Overall accuracy = 0.91

Correct prediction for non-urban = 0.95

Correct prediction for urban = 0.58

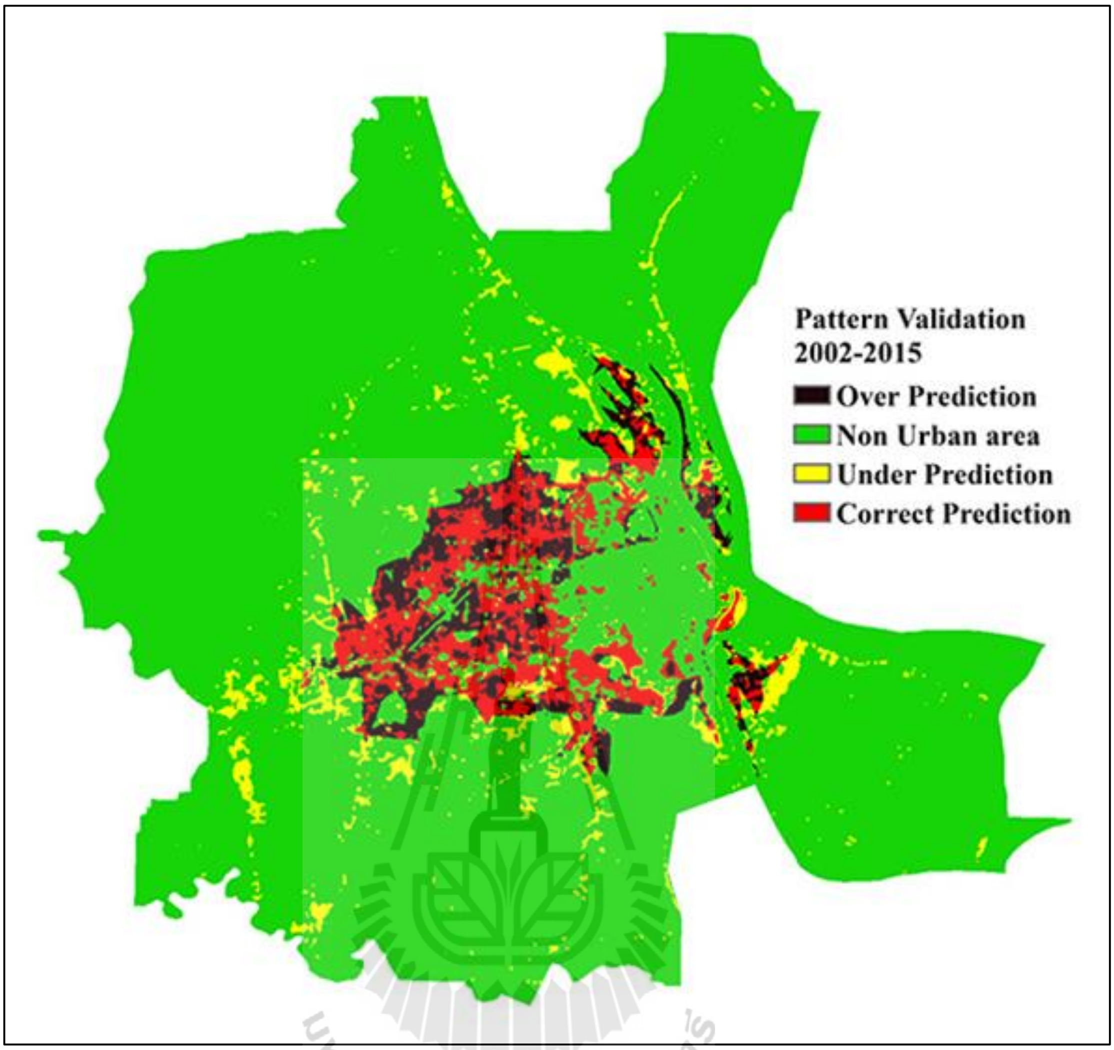


Figure 4.17 Validation pattern of LR model for 2002-2015 period.

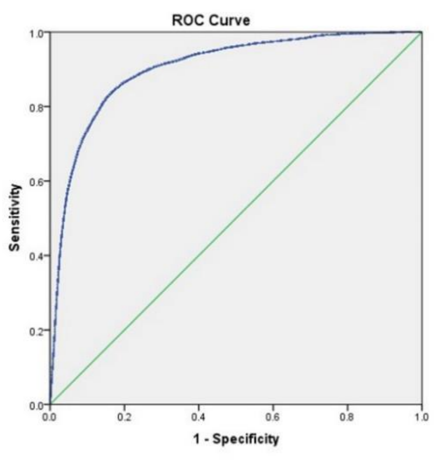


Figure 4.18 False positive (Specificity) and true positive (Sensitivity) of LR model 2002-2015.

4.4.2 Urban growth model validation for 2002-2009 period

According to Table 4.17, urban growth area between 2002 and 2009 is about 32.18 sq. km or 35,751 pixels, with overall accuracy of 96% of urban growth occurrence. This high number of overall accuracy come from the success fit of non-urban about 97% due to the fact that the study area consists of high number of non-urban cells that predicted correctly by the model, but not quite high for the success fit of urban which value is only 0.54. This can be explained that the model of logistic regression with 14 influential factors can explain about 54% of the phenomena of urban growth pattern in Phnom Penh during 2002-2009, and it is suggested to fulfill another 46% by including unidentified factors that are not presented in this model, in order to improve the calibration capacity. However, as observed from the contingency table and Figure 4.19, 16,473 cells of urban in reality was predicted as non-urban in model which demonstrated as under prediction, while 16,479 cells of themselves were replaced as over prediction which is non-urban in reality, and is predicted as urban in the model. This means that the allocation of cells go into the over prediction location where it has higher probability value, instead of allocation in the under prediction location. In short, during this short term period, LR model also indicated inappropriated estimation for urban scattering growth, where it is unable to have high probability value when compared to infill growth around the existing urban.

Meanwhile, ROC result reveals that the LR model for 2002-2009 has area under curve of ROC value of 0.948 (Figure 4.20). The results revealed that the validated calibration for urban growth of year 2002-2009 show good performance and can be further applied for urban growth prediction.

Table 4.17 Contingency table and accuracy for urban growth model of 2002-2009.

Unit in pixel		**Reality**	
Model	**Urban**	**Non-Urban**	**Total**
Urban	19,278	16,479	35,757
Non-Urban	16,473	708,124	724,597
Total	35,751	724,603	760,354

Correct prediction = 727,402

Wrong prediction = 32,952

Overall accuracy = 0.96

Correct prediction for non-urban = 0.97

Correct prediction for urban = 0.54

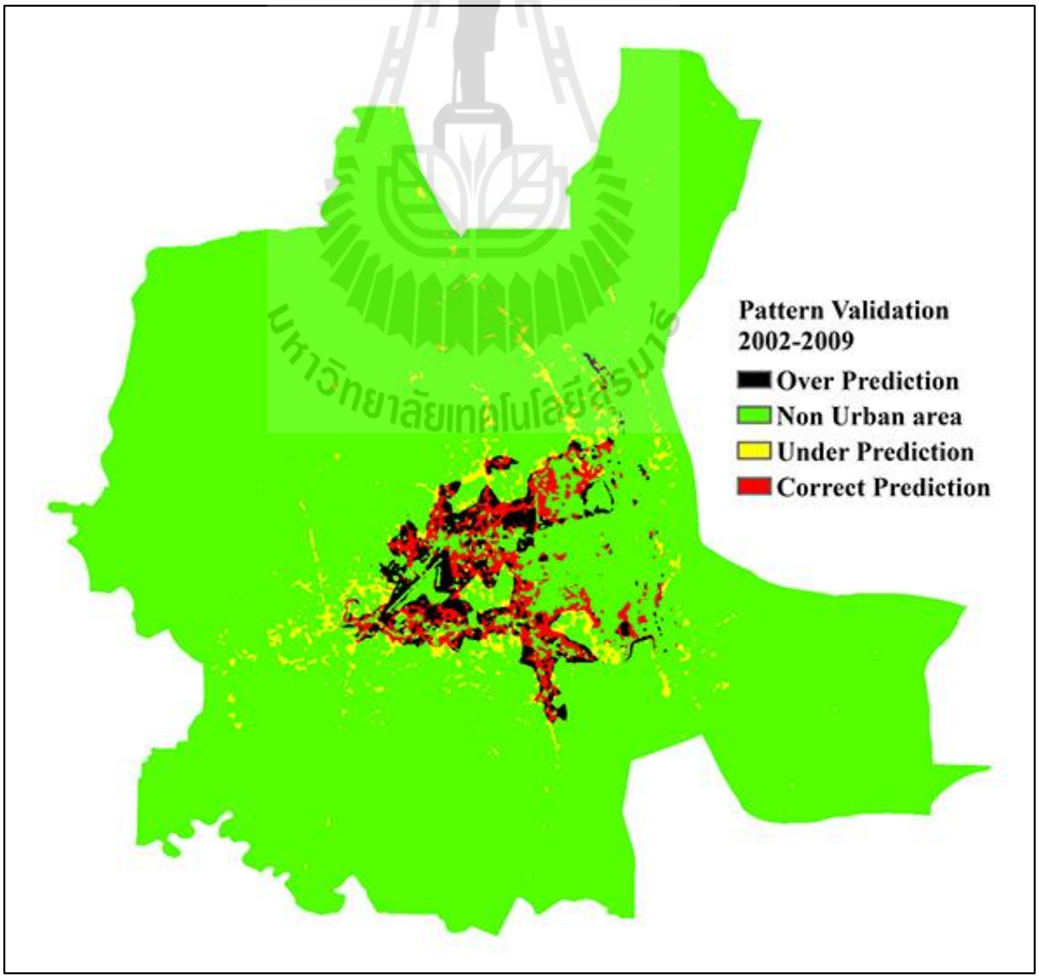


Figure 4.19 Validation pattern of LR model for 2002-2009 period.

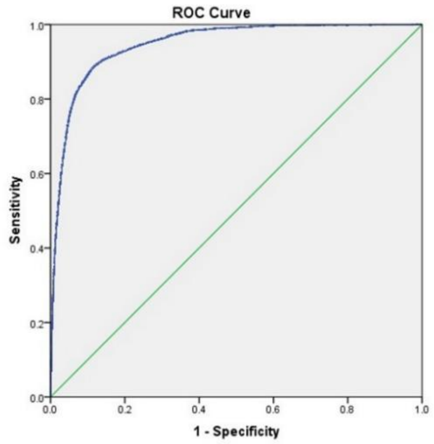


Figure 4.20 False positive (Specificity) and true positive (Sensitivity) of LR model 2002-2009.

4.4.3 Urban growth model validation for 2009-2015 period

Refer to Table 4.18, urban growth area between 2009 and 2015 is about 38.41 sq. km or 42,681 pixels, with overall accuracy of 94% of urban growth occurrence. This high number of overall accuracy come from the success fit of non-urban about 96% due to the fact that the study area consists of high number of non-urban cells that predicted correctly by the model, but low result for the success fit of urban which value is only 0.44. This can be explained that the model of logistic regression with 13 influential factors can explain about 44% of the phenomena of urban growth pattern in Phnom Penh during 2009-2015, and it is suggested to fulfill another 56% by including unidentified factors that are not presented in this model, in order to improve the calibration capacity. The reason is the same as the two models above, as allocation process did not fit the right location correctly (Figure 4.21). However, ROC result reveals that the LR model for 2009-2015 provides high ROC value of 0. 901 (Figure 4.22). The results can be accepted to apply for urban growth prediction.

Table 4.18 Contingency table and accuracy for urban growth model of 2009-2015.

Unit in pixel		**Reality**	
Model	**Urban**	**Non-Urban**	**Total**
Urban	18,626	24,071	42,697
Non-Urban	24,055	693,602	717,657
Total	42,681	717,673	760,354

Correct prediction = 712,228

Wrong prediction = 48,126

Overall accuracy = 0.94

Correct prediction for non-urban = 0.96

Correct prediction for urban = 0.44

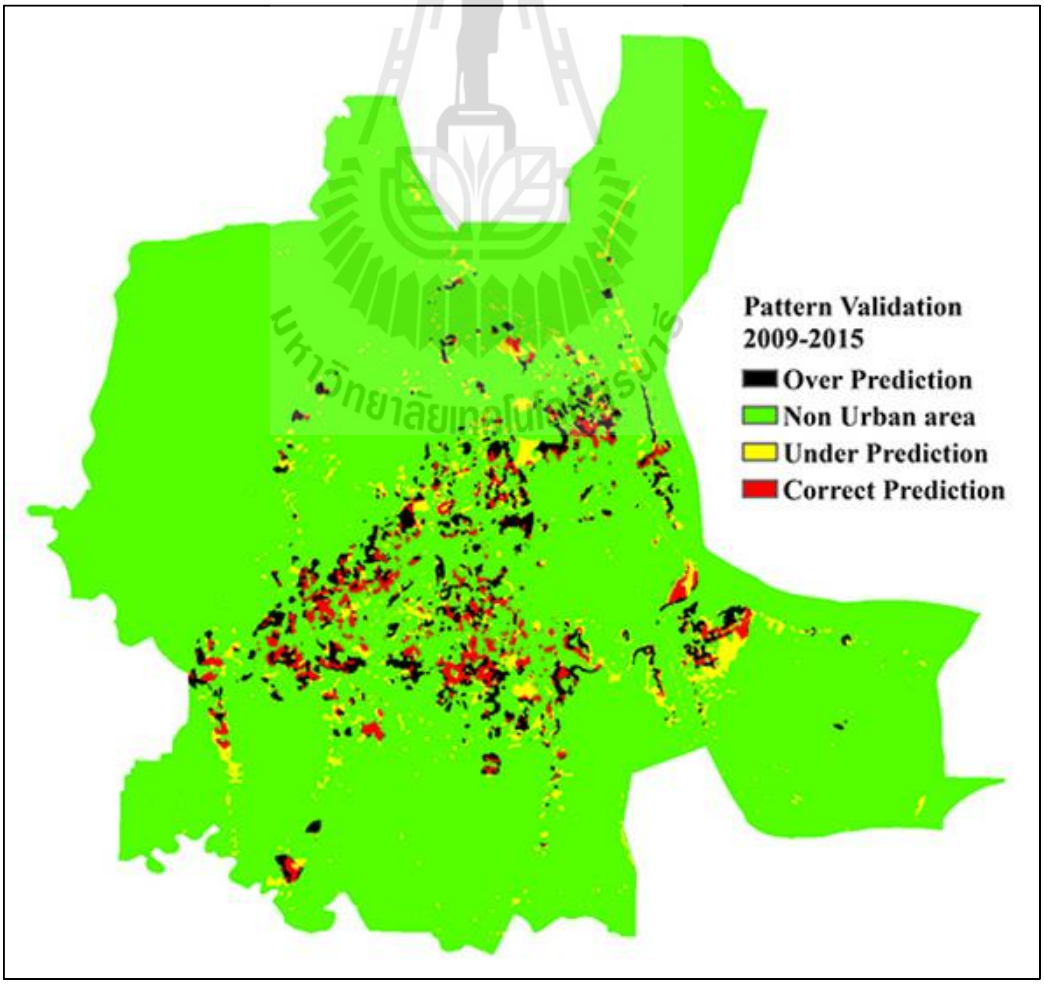


Figure 4.21 Validation pattern of LR model for 2009-2015 period.

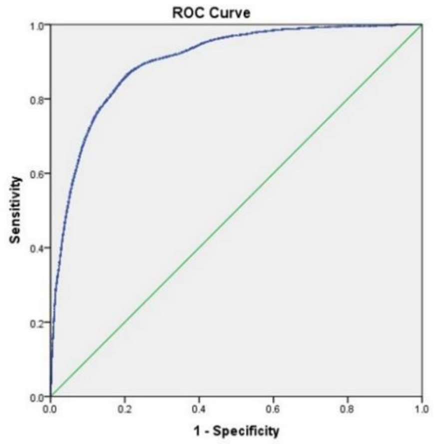


Figure 4.22 False positive (Specificity) and true positive (Sensitivity) of LR model 2009-2015.

4.4.4 Optimum LR model for urbanization prediction

According to the spaital and statistical pattern validation for urban growth in three periods as report earlier and presented in Figure 4.17, 4.19, and 4.21 with contingency Table 4.16, 4.17, and 4.18, the validation result is made a decision that LR model of 2002-2015 have conducted the satisfied calibration with reliable reason for Phnom Penh urban growth pattern in 2030 prediction. Because it can provide high ROC value and high percentages of overall accuracy with moderate satisfactory result of fitting urban allocation (Table 4.19). Moreover, it is a long term period model, the latest factor maps of 2015 could be updated according to city development, for example existing urban cluster of 2015 will replace the existing urban cluster of 2002 and many other factors like environmental, social and proximity and accessibility factors, so that the results would be more reliable and accurate compared to 2002-2009 model. This finding is also agreed with the previous work of Nduwayezu (2015) who also got good calibration of LR model in long term period.

The entire pattern of validation map of urban growth model of 2002-2015 period is shown in Figure 4.23. Appendix D provides the histogram of probability maps for three LR models.

Table 4.19 Comparison of LR model efficiency from three periods.

Criteria	LR model of 2002-2015	LR model of 2002-2009	LR model of 2009-2015
Length period	13 years	7 years	7 years
Pseudo R^2	0.439	0.50	0.37
Number of significance of influential factor involvement	14	14	13
ROC value	0.903	0.948	0.901
Allocation fit for urban pixels correction	58%	54%	44%

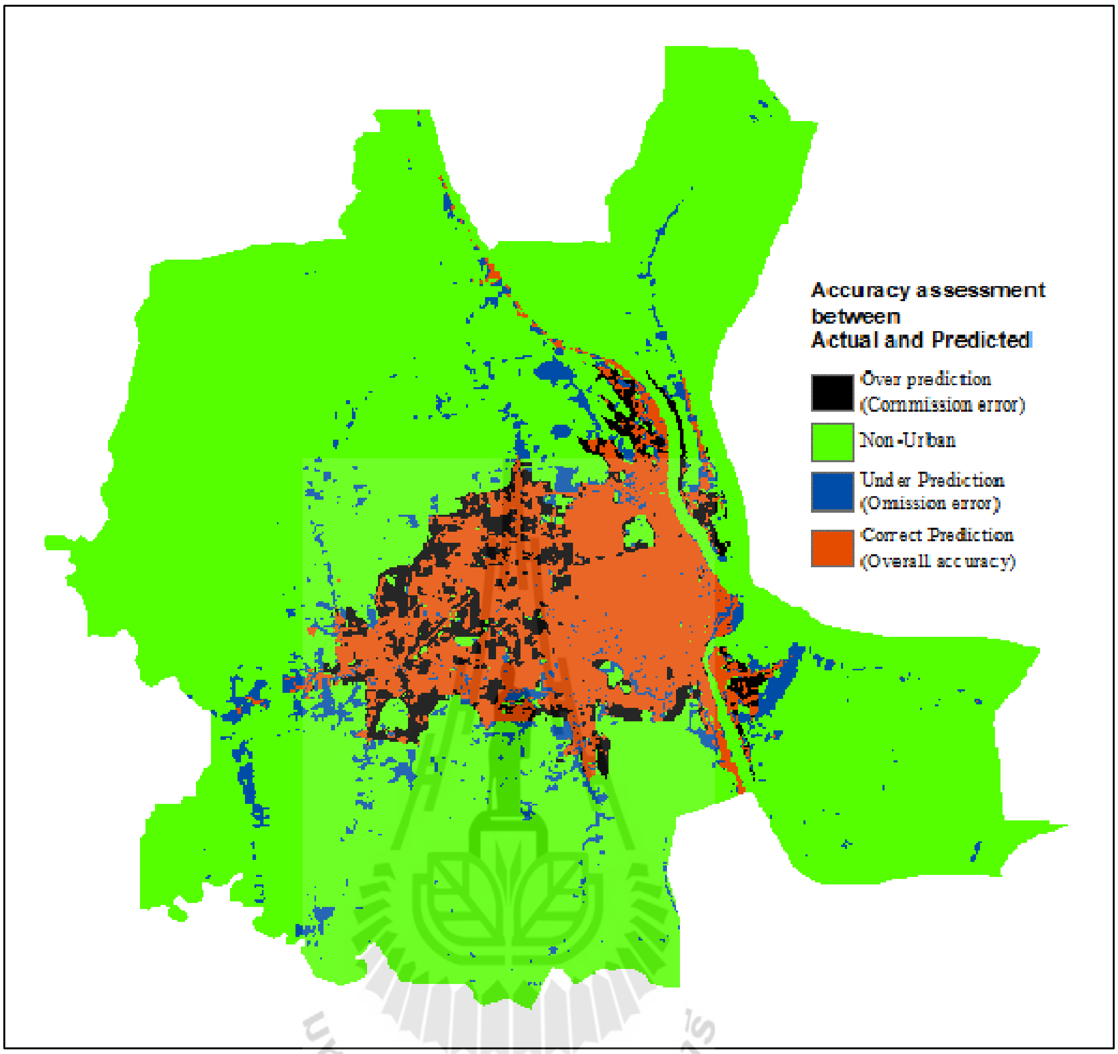


Figure 4.23 Entire pattern of validation map of urban growth model of 2002-2015 period.

4.5 Urban growth pattern of Phnom Penh in 2030

There are three methods include (1) future population growth rate, (2) Markov Chain model, and (3) simple linear regression had been used to identify as an optimal method for urban growth change demand estimation for urban growth pattern of Phnom Penh in 2030 prediction.

4.5.1 Urban growth change demand by future population growth rate

A summary of the urban and population growth of Phnom Penh during 2002-2015 which was used to estimate growth change demand area by future population growth rate with Equation 3.12 is presented in Table 4.20. All of the value in Table 4.20 were input into the Equation 3-12 in order to calculate urban growth change demand area as a result shown in Table 4.21.

The population of the city increased from 1,275,737 in 2002 to 1,835,090 in 2015. This represents an increase of 43.85% at an annual growth rate of 3.37%. Meanwhile, urban and built up areas of Phnom Penh increased from 38.12 sq. km in 2002 to 108.73 sq. km in 2015 according to the result from land cover change detection. The net increase of the urbanized areas in this period was 70.6 sq. km which represents an increase in urban size of 185.17% at an annual expansion rate of 14.24%. As a result, growth ratio (*R*) was obtained for 4.22, then the coefficient for multiplying with the existing urban area ($A_{existing}$) was 2.43. This means that urban area in 2030 is about 2.43 times the existing urban area in 2015, so that urban area in 2030 is extensive large about 262.78 sq. km or urban growth demand for 2030 allocation is 154.05 sq. km. The estimated urban growth change demand area is inconvincible and extensively high. If this area is allocated on the probability urban growth pattern of LR model, predicted urban will spread out in all direction of Phnom Penh.

Table 4.20 Summary of predicted future urban growth area by future population growth.

Population growth	**Value**
Population in 2002 (in person) [1]	1,275,737
Population in 2015 (in person) [2]	1,835,090
Net growth	559,353
Change (%)	43.85%
Annual growth rate (%)	3.37%
Urban growth rate	
Urban area in 2002 (sq. km)	38.1285
Urban area in 2015 (sq. km)	108.7308
Net growth	70.6023
Change (%)	185.1694%
Annual growth rate (%)	14.24 (%)

Sources: 1. Based on exponential annual growth rate between 1998 and 2008

2. Based on Population Projections for Cambodia between 2008-2030 which was studied by Statistics Japan corporated with JICA Cambodia (Statistics Japan, 2009).

Table 4.21 Urban growth change demand in 2030.

Urban growth change demand in 2030	**Value**
Growth Ratio (R)	4.22
Population in future (P_{2030})	2450717
$(P_{2030} – P_{2015})/P_{2015}$	0.34
$(1+R((P_{2030} – P_{2015})/P_{2015})$	2.43
Urban area in 2015 ($A_{existing}$)	108.7308 sq. km (120812 cells)
Urban area in 2030 (A_{future})	262.7792 sq. km (291977 cells)
Urban growth area demand in 2030	154.0484 sq. km (171165 cells)

4.5.2 Urban growth change demand by Markov Chain model

Markov Chain model was operated under IDRISI Selva Edition to calculate the transition probability matrix and transition area matrix of each land cover class based on the classified map of 2002 and 2015. However, this study used the quantity of change for urban area only, since scope of work dealt only with urban and built-up area. Refer to Table 4.22, Markov Chain model was estimated the probability transition matrix of 2030 based on 13 years of transition area, so that result of transition area matrix in 2030 was obtained in Table 4.23 with amount of urban area 185.83 sq. km or about 206,483 cells. This means that the increasing urban area during 2015 to 2030 is 77.1 sq. km (185.83-108.73 sq. km) or 85,667 cells.

Table 4.22 Markov probability value of 2030 based on transition matrix of 2002-2015.

Land cover class	**U**	**A**	**V**	**W**	**M**
Urban and built-up area (U)	1	0	0	0	0
Arable land (A)	0.1391	0.6538	0.1551	0.0151	0.0369
Vegetation (V)	0.1499	0.2726	0.4894	0.0548	0.0332
Water body (W)	0.0512	0.1721	0.2332	0.4802	0.0632
Miscellaneous Land (M)	0.2416	0.3126	0.3208	0.0904	0.0346

Table 4.23 Transition area matrix by Markov Chain model for 2030 in number of cells.

Land cover class	U	A	V	W	M
Urban and built-up area (U)	120,802	0	0	0	0
Arable land (A)	42,499	199,711	47,393	4,610	11,271
Vegetation (V)	30,982	356,345	101,160	11,333	6,870
Water body (W)	4,985	16,743	22,696	46,731	6,155
Miscellaneous land (M)	7,215	9,337	9,582	2,701	1,032
Total	206,483				

4.5.3 Urban growth change demand by simple linear regression

Simple linear regression was used to extrapolate urban growth area of 2030 based on urban area of 2002, 2009, and 2015 by Trend analysis of MS Excel spreadsheet as shown in Table 4.24. As a consequence, result showed with a perfect fit regression line with R^2 value of 0.99 (Figure 4.24) that means amount of urban change in three periods were proportional to temporal change. So that urban area of Phnom Penh in 2030 will be 187.76 sq. km and the increasing urban area during 2015 to 2030 is about 79 sq. km which is quite similar to the result from Markov Chain model.

Table 4.24 Estimate the future urban change demand by trend linear line.

Year	Number of year	Urban area (sq. km)
2002	1	38.1285
2009	8	70.3089
2015	14	108.7308
2030	29	187.76

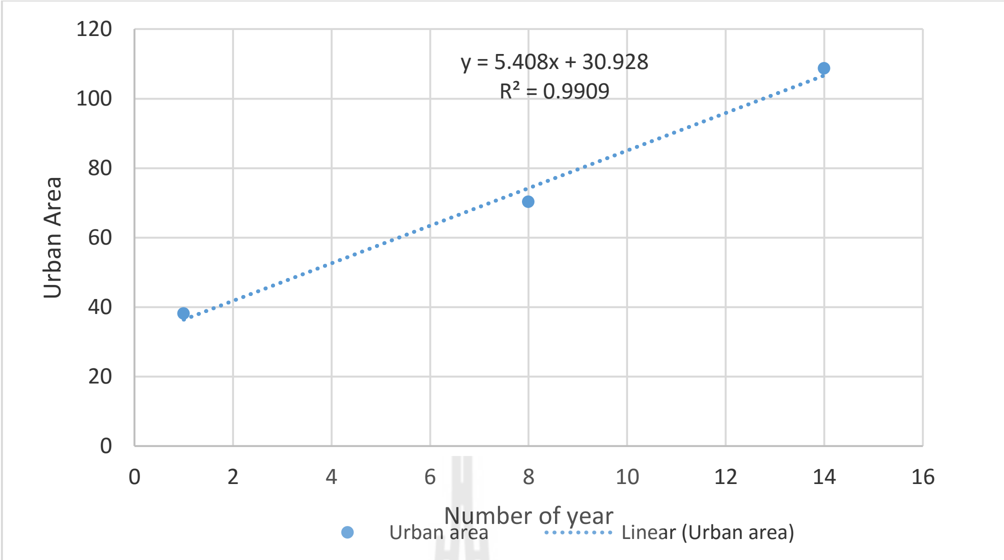


Figure 4.24 Linear line graph of equation and R^2 value.

4.5.4 Prediction of urbanization probability map

The probability of urbanization map was predicted by updating the new factor maps which is the latest information of independent variables in 2015 in Figure 4.13 into the equation of optimum LR model of 2002-2015 period. Finally, future probability image of urban growth was createad under Raster Calculator tool of ESRI ArcMap 10.2 as shown in Figure 4.25. Appendix D provides the histogram of future urbanization probability map.

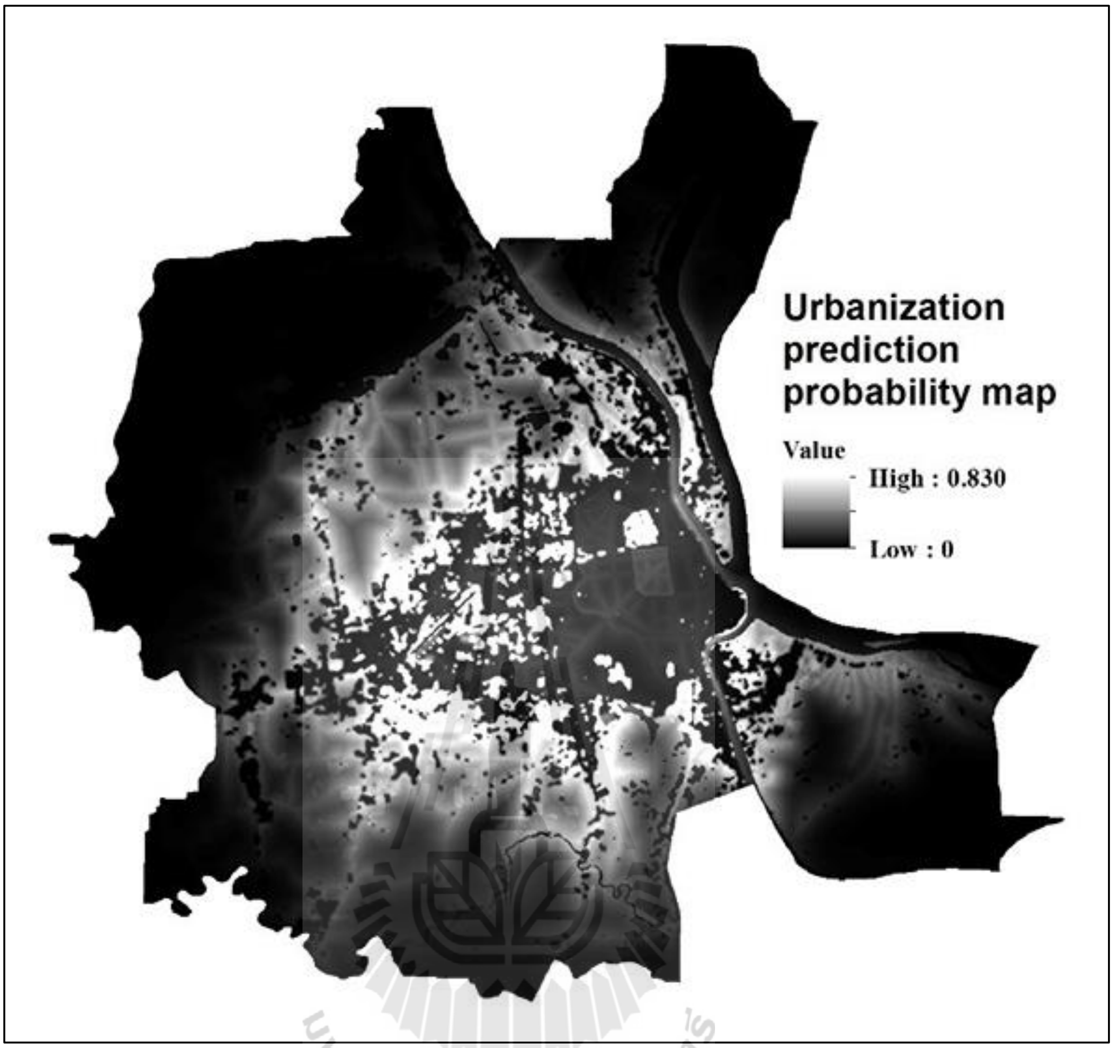


Figure 4.25 Future urbanization probability map of Phnom Penh in 2030.

According to result of the change demand area estimation from the three methods above, Markov Chain model and simple linear regression methods provided a reasonable result in term of urbanization trend that is quite logical and acceptable if they compared with future population growth rate method that estimated extremely high area, about 2.4 times of existing urban and built-up area of 2015. Furthermore, to consider about the policy development of urbanization as well as the urban vertical growth direction like the development of high rise building, condominium, and area of congested urban area, it is better to use urban growth change demand area from Markov Chain model and simple linear regression methods rather than future population growth rate method, because land policy and developer might build more apartments than one single family house due to urban planning restriction, social development and the growth of economic (Alsharif et al., 2014). Herein, Markov Chain model and simple linear regression methods can provide similar result. However, Markov Chain model applied both transition probability matrix and transition area matrix of all land cover class to predict urban and built-up area while simple linear regression applied only area of urban in three dates without considering transitional change among land cover class to estimate the future change demand area. So that, the derived future change demand area using Markov Chain model was chosen to allocate the urban growth pattern of Phnom Penh in 2030 based on future urbanization probability map based on LR model of 2002-2015 period in Figure 4.25. Finally, the result of urban growth pattern prediction of Phnom Penh in 2030 is presented in Figure 4.26.

It illustrates that the pattern of growth is occurred in the north, south and east direction around the existing urban cluster, and some are expected to occur along the major roads and ring roads. Furthermore, majority of the miscellaneous land that are the construction sites for development of satellite cities in 2015 are forecasted to finish their constructions and will become urban and built-up area in 2030.

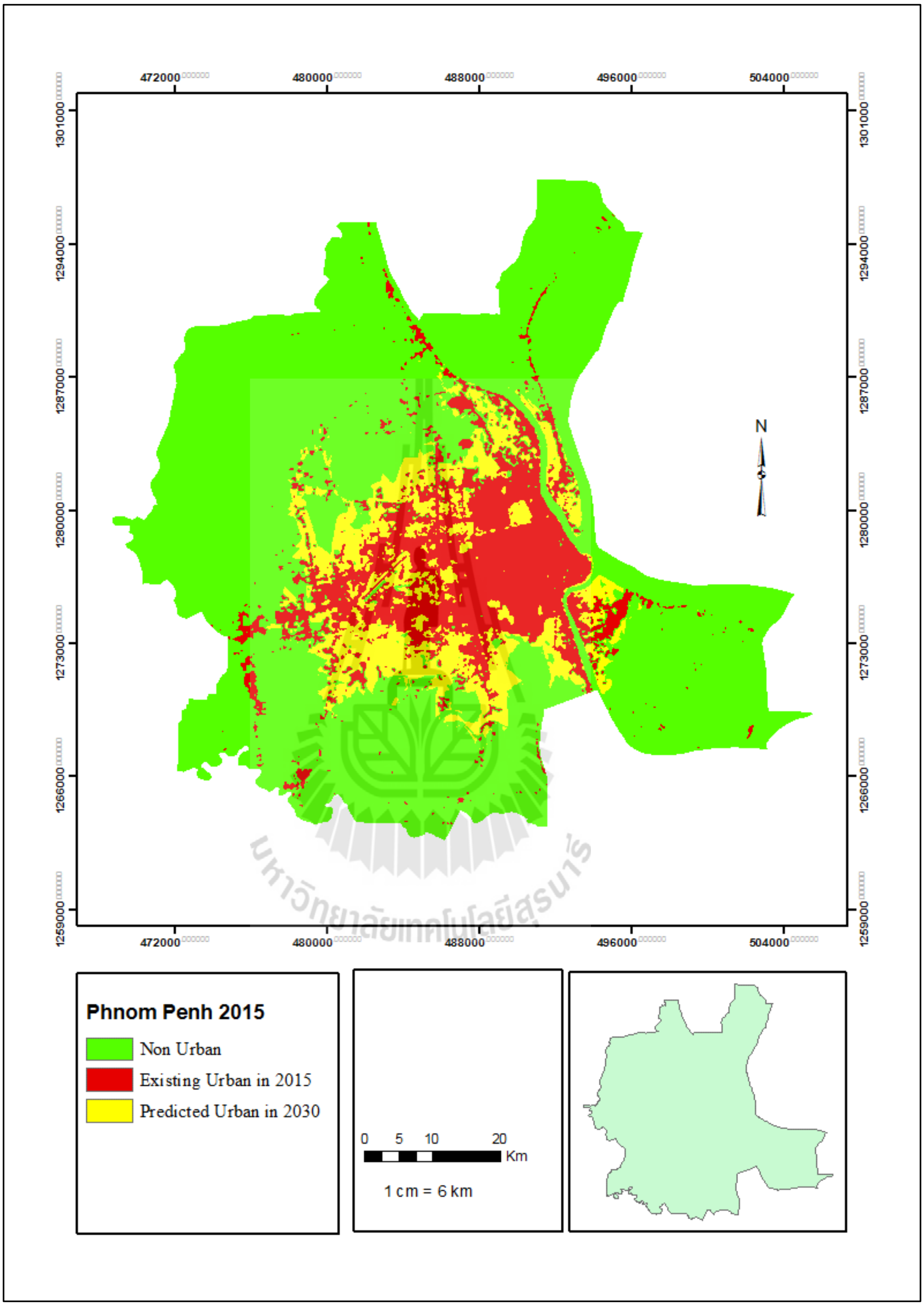


Figure 4.26 Urban growth pattern prediction of Phnom Penh in 2030 with the existing urban areas in 2015.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Under this chapter, three main results of the study according to the objectives included (1) to assess Phnom Penh land cover status and its change in three different periods, (2) to discover the driving factors of Phnom Penh urban growth in three periods and to identify optimum LR model for future urban growth pattern prediction, and (3) to predict future urban growth pattern of Phnom Penh in 2030 are here separately concluded and some recommendations are suggested for future research and development.

5.1 Conclusions

5.1.1 Land cover assessment and its change

The classification of derived land cover map of 2002, 2009, and 2015 were done by supervised classification method in order to quantify land cover status and its change including urban and built-up area, arable land, vegetation, water body and miscellaneous land. As results, it is found that arable land for rice production is the most predominant land cover type in Phnom Penh in three years. Since Cambodia economy is relied on agricultural sector, this result can be an evidence to be proved in term of spatial pattern from the past until recent years is agricultural land.

According to land cover change matrix of 2002-2009, 2009-2015, and 2002-2015 periods, it reveals that arable land and vegetation are the key land cover classes for urbanization process in three periods. The analysis of spatial temporal land cover change also reveal that urban and built-up area continuously increase from 38.13 sq. km in 2002 to 70.31 sq. km in 2009 and to 108.73 sq. km in 2015 and resulting a major decline of arable land, vegetation and water body. However, change for miscellaneous land that is bare land for construction and development sites is shown as fluctuation. This reason demonstrates the fast development of urbanization that reflects the increase of construction sector, the growth of economic and industrialization and the demand of housing and satellite cities for people living in Phnom Penh starting in 2009.

In addition, pattern of urban growth demonstrates that urban areas expand only toward the southern, northern and western direction of Phnom Penh during 2002-2009 with all type of growth including infill growth, expansion growth, linear branch and isolated growth, while eastern part of Phnom Penh does not show much of growth pattern due to that area mostly wet land. However, the expansion of urban growth during 2009-2015 occurs in all direction, and types of urban growth are changed to be expansion growth, clustered branch and isolated growth whereas infill growth during 2002-2009 has already been developed as urban, so that during 2009-2015, there are less developed area to be converted as urban as infill growth.

5.1.2 Optimum LR model and keys influential factors of urban growth in Phnom Penh

Model of logistic regression for two short term periods and one long term period of 2002-2009, 2009-2015, and 2002-2015 were here employed with 16 initial influential variables included environmental, social, proximity and neighborhood factors. The finding reveals that only 14 variables from 16 variables under LR model of three periods are qualified the multicollinearity test, except the distance to industrial port and river. However, only 13 variables from 14 variables are significance, except population density under LR model of 2009-2015 period based on Wald statistic test with P-Value > 0.005 while 14 variables are significance under LR model of 2002-2009 and 2002-2015 periods with P-Value < 0.005. Moreover, Nagelkerke Pseudo R^2 for measuring the degree of goodness of LR model of three periods with value greater than 0.2 indicates the goodness of regression line between sample points of independent variables and dependent variable. The final LR model of 2002-2009, 2009-2015, and 2002-2015 periods provide a good result of overall accuracy of 0.96, 0.94 and 0.95 and the success fit for urban allocation with the value of 0.54, 0.44 and 0.58, respectively. These result show that LR model of 2002-2015 period can provide the best LR calibration equation and was chosen an optimum LR model to predict the urban growth pattern of Phnom Penh in 2030.

Furthermore, it is found that the relative influence of influential factors varies according to urban and built up area pattern and time. No influential factors that permanently impact on the urban growth of Phnom Penh in all periods. During 2002-2015 as long term period, top three negative influential factors are distance to Phnom Penh airport, distance to road and distance to industrial area. However, during the short term period of 2002-2009 and 2009-2015 results show negative influence of distance to the existing urban cluster, which is the first ranking of both periods, but the second and third ranking are different. Herein, distance to industrial area and Phnom Penh airport are the second and third negative influential factors of 2002-2009 period, while distance to CBD and miscellaneous land are shown as negative and positive influential factor in 2009-2015 period.

During the short term period between 2002 and 2009, urban development in Phnom Penh city tends to grow in the neighborhood area resulting the pattern of urban growth in infilling growth inside the core existing urban and expansion growth type around the urban fringe area. This result is possible for urban development in developing country that vacant land in downtown is still remaining and urban likely grow on that land to become more crowded. However, Phnom Penh airport, road and industrial area are dominant influential factors when they consider in term of long period, because Phnom Penh airport and industrial area were located in suburb area which are linked to the core existing urban in 2002 by road and main road. Therefore, when urban expands due to the rise of population and the development of industrialization and satellite cities, the attraction of growth is took place along the main road and continued westward to area of airport and industrial area. The influence of Phnom Penh airport and industrial area still play certain role in 2002-2009 period, but CBD and miscellaneous land play more important role in 2009-2015 because of expansion of satellite city on Diamond Island near the CBD, and miscellaneous land become significance influence indicated by most of construction sites were prepared for urban development in Phnom Penh. On contrary Phnom Penh airport and industrial area swiftly lost their significance due to their surrounded areas had already been transformed to urban in 2009, so that it seems to be under controlled by urban cluster instead of Phnom Penh airport and industrial area during 2009-2015.

5.1.3 Future urban growth pattern in Phnom Penh City

Since logistic regression line can fit the sample points between binary variable and predictor variables, the urban growth pattern affected by its influential factors in the past can be used for the future prediction. However, the influential factors from the past may not be ensure to be certain for future influence, so that this study modeled the logistic regression with one long time and two short time periods and found that influential factors of urban growth are different. This finding may be helpful as an information or a guideline for decision makers to consider those important factors in different time periods, in order to use for predicting the future urban growth pattern.

In this study, LR model of 2002-2015 period as an optimum model based on its high accuracy and ROC value was applied to predict urban growth pattern of Phnom Penh in 2030. It reveals that Phnom Penh tends to expand in the north, south and east direction. This expansion can be influenced by the nearest existing urban of 2015, and the influence from road. Moreover, some patterns of linear growth are also detected along the major ring road at the north and some major roads at the southern of Phnom Penh to the Kompong Speu province. Most of the satellite cities which are bare land in 2015 like Bueng Kak Lake in the center of Phnom Penh, Grand Phnom Penh and Camko City in the northern Phnom Penh, Ing City in the southern Phnom Penh and Chroy Chongvar City in the eastern of Phnom Penh are expected to become urban and built-up area in 2030.

5.2 Recommendations

In this study, LR model was implemented to model urban growth by incorporating remotely sensed data with GIS techniques, in order to answer the key question of every objective in term of spatial pattern. However, this study may have limitation that need to be enhanced for future study as in the following recommendation.

(1) Although, this study try to review the influential factors to input them to the LR model as much as possible, the improvement of this model could be made in order to enhance more the accuracy of success fit for validation of urban growth by including the missing variables like economic variable (land value, per capital income, and poverty rate).

(2) This study have fulfilled the finding and understanding the relative influence between physical, social, environmental, proximity and neighborhood factors and urban growth pattern of Phnom Penh during 2002-2015 using the empirical simulation model (LR model), and it was applied to predict the future urban growth pattern in 2030 based on relative influence of driving factors. However, further studies should be added more by applying different urban growth models include rule-based simulation (CA) and system dynamic simulation models. Each model has their own advantages and disadvantages, so that different modeling techniques may describe a different view of result that can draw new conclusions and discussions for further study in urban planning.

REFERENCES

REFERENCES

Achmad, A., Hasyim, S., Dahlan, B., and Aulia, D. (2015). Modeling of urban growth in tsunami-prone city using logistic regression: Analysis of Banda Aceh, Indonesia. **Applied geography**. 62: 237-246.

Aguayo, M. I., Wiegand, T., Azocar, G. D., Wiegand, K., and Vega, C. E. (2007) Revealing the driving forces of mid-cities urban growth patterns using spatial modeling: a case study of Los Ángeles, Chile. **Ecology and Society**. 12(1): 13.

Al-Ahmadi, K., Al-Ahmadi, S., and Al-Amri, A. (2013). Exploring the association between the occurrence of earthquakes and the geologic-tectonic variables in the Red Sea using logistic regression and GIS. **Arab J Geosci**.

Allen, J., and Lu, K. (2003). Modeling and Prediction of Future Urban Growth in the Charleston Region of South Carolina: a GIS-based Integrated Approach. **Conservation Ecology**. 8(2): 2.

Alsharif, A. A. A., and Pradhan, B. (2014). Urban Sprawl Analysis of Tripoli Metropolitan City (Libya) Using Remote Sensing Data and Multivariate Logistic Regression Model. **J Indian Soc Remote Sens**. 42(1): 149-163.

Arsanjani, J. J., Helbich, M., Kainz, W., and Boloorani, A. D. (2013). Integration of logistic regression, Markov chain and cellular automata. **International of applied earth observation and geoinformation**. 21: 265-275.

Barnes, K. B., Morgan, J. M., III, Roberge, M. C., and Lowe, S. (2001). Sprawl development: Its patterns, consequences, and measurement. **A white paper, Towson university**.

BAU. (2007). **LIVRE BLANC DU DEVELOPPEMENT ET DE L'AMENAGEMENT DE PHNOM PENH.** CAMBODGE: Municipalité de Phnom Penh, Bureau des Affaires urbaines.

Bhatta, B. (2010). Analysis of Urban Growth and Sprawl from Remote Sensing Data, **Springer-Verlag Berlin Heidelberg**.

Braimoh, K. A., and Onishi, T. (2007). Spatial determinants of urban land use change in Lagos, Nigeria. **Land Use Policy**. 24: 502-515.

Campbell, C. E., Allen, J., and Lu, K. S. (2007). Modeling growth and predicting future developed land in the upstate of South Carolina. **Proceedings of the 2008 South Carolina Water Resources Conference**.

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

CIPS. (2013). Cambodia Inter-Censal Population Survey 2013 Final Report. **National Institute of Statistics, Ministry of Planning**. p. 19.

Clark, W. A., and Hosking, P. L. (1986). Statistical methods for geographers. **New York, Wiley**.

Dewan, A. M., and Yamaguchi, Y. (2009). Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. **Applied geography**. 29: 390-401

Doyle, S. E. (2012). City of Water: Architecture, Urbanism and the Floods of Phnom Penh. **Fulbright Research Fellow, Cambodia**.

Duwal, S. (2013). **Modeling urban growth in Kathmandu Valley**. Master Thesis. Faculty of Geo-Information Science and Earth Observation of the University of Twente, Enschede, the Netherlands.

Eyoh, A., Olayinka, D. N., Okwuashi, P. N. O., Isong, M., and Udoudo, D. (2012). Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic regression and GIS. **International Journal of Applied Science and Technology**. 2(5).

Guo, H. C., Liu, L., Huang, G. H., Fuller, G. A., Zou, R., and Yin, Y. Y. (2001). A system dynamics approach for regional environmental planning and management: A study for the Lake Erhai Basin. **Journal of Environmental Management**. 61: 93-111.

Hansen, W. G. (1959). How accessibility shapes land use. **Journal of the American Institute of Planners**. 25(2): 73-76.

Hao, C., Zhang, J., Li, H., Yao, F., Huang, H., and Meng, W. (2014). Integration of Multinomial-Logistic and Markov-Chain Models to Derive Land-Use Change Dynamics models to simulate urban expansion. **Journal of Urban Planning and Development**.

He, C. Okada, N., Zhang, Q., Shi, P., and Zhang, J. (2006). Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China. **Applied Geography.** 26: 323-345.

Hosmer, J. D. W., and Lemeshow, S. (1989). **Applied logistic regression**. New York: Wiley.

Hu, Z., and Lo, C. P. (2007). Modeling urban growth in Atlanta using logistic regression. **Computers, Environment and Urban Systems.** 31: 667-688.

IBM-SPSS. (2012). **IBM-SPSS File Help** (Version 21).

IDRISI. (2012). **IDRISI File Help** (Selva Edition). Clark Labs, Clark University.

Jensen, J. (2005). **Introductory digital image processing** (3rd edition). Prentic Hall Series in geographic information system.

Landis, J., and Koch, G. (1977). **The Measurement of Observer Agreement for Categorical Data. Biometrics**. 33: 159-174. (Quoted by Jensen, 2005).

Lewis-Beck, M. S. (1994). **Factor Analysis and Related Techniques**. SAGE Inc, Jurong, Singapore.

Li, X., Zhou, W., and Ouyang, Z. (2013). Forty years of urban expansion in Beijing: What is the relative importance of physical, socioeconomic, and neighborhood factors? **Applied geography**. 38: 1-10.

Liu, X., Li, X., and Yeh, A. G. O. (2006). Multi-agent systems for simulating spatial decision behaviors and land-use dynamics. **Science in China Series D: Earth Sciences**. 1184-1194.

Logofet, D. O., and Lesnaya, E. V. (2000). The mathematics of Markov models: what Markov chains can really predict in forest successions. **Ecological Modeling.** 126: 285-298.

Lopez, E., Bocco, G., Mendoza, 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.

Luenberger, D. G. (1979). **Introduction to Dynamic Systems. Theory, Models, and Applications**. Wiley, New York.

Lui, J., Zhan, J., and Deng, X. (2005). Spatio-temporal Patterns and Driving Forces of Urban Land Expansion in China during the Economic Reform Era. AMBIO: **A**

Journal of the Human Environment. 34(6): 450-455. Royal Swedish Academy of Sciences.

Luther, T. (2005). “The Effects of Urban Sprawl on Forest Fragmentation and Parcelization.” **Private Forest**.

Manu, A., Twumasi, Y. A., Lu, K. S., and Coleman, T. L. (2015). Predicting urban growth of a developing country city using a statistical modeling approach. **International journal of geomatics and geosciences**. 5(4).

Mernard, S. W. (2002). **Applied logistic regression analysis** (2nd ed.). Thousand Oaks, CA: SAGE Publications.

NGD. (2002). **Cambodia national topographic map.** Phnom Penh, Cambodia: National geographic department, Ministry of land management, urban planning and construction.

NIS. (2002). **General Population Census of Cambodia 1998 Final Census Result.** (2nd ed.). Phnom Penh, Cambodia: National Institute of Statistics (NIS), Ministry of Planning.

NIS (2009). **General Population Census of Cambodia 2008 National Report on Final Census Results.** Phnom Penh, Cambodia: National Institute of Statistics (NIS), Ministry of Planning.

NIS. (2013). **Cambodia Inter-Censal Population Survey 2013 Final Report**. Phnom Penh, Cambodia: National Institute of Statistics (NIS), Ministry of Planning.

NIS (2013). **Economic census of Cambodia in 2011 Provincial Report 12 Phnom Penh Municipality.** Phnom Penh, Cambodia: National Institute of Statistics (NIS), Ministry of Planning.

Nduwayezu, G. (2015). **Modeling urban growth in Kigali City Rwanda.** Master Thesis. Faculty of Geo-Information Science and Earth Observation of the University of Twente, Enschede, the Netherlands.

Odzemir, A. (2011). Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). **Journal of hydrology.** 405: 123-136.

Pontius, R. G. Jr., and Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. **Agriculture, Ecosystems and Environment.** 85: 239-248.

Statistics Japan. (2009). **POPULATION PROJECTIONS FOR CAMBODIA, 2008-2030.** Phnom Penh, Cambodia. [Online]. Available: http://www.stat.go.jp/info/meetings/cambodia/pdf/rp12_ch10.pdf.

Rogerson, P. (2010). **Statistical methods for geography.** London; Thousand Oaks, Calif. : SAGE Publications.

Rui, Y., and Ban, Y. (2011). Urban growth modeling with road network expansion and land use development. **Advances in Cartography and GIScience**.

Shalaby, A., and Tateishi, R. (2007). Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. **Applied geography**. 27: 28-41.

Shu, B., Zhang, H., Li, Y., Qu, Y., and Chen, L. (2014). Spatiotemporal variation analysis of driving forces of urban land spatial expansion using logistic regression: A case study of port towns in Taicang City, China. **Habitat international.** 43: 181-190.

Thapa, B. R., and Murayama, Y. (2010). Drivers of urban growth in the Kathmandu valley, Nepal: Examining the efficacy of the analytic hierarchy process. **Applied geography.** 30: 70-83.

UNFPA Cambodia. (2014). **Urbanization and its linkage to socio-economic and environmental issues**. Cambodia: United Nations Population Fund (UNFPA).

Verburg, P. H., Nijs, T. C. M., Eck, J. R., Visser, H., and Jong, K. (2003). A method to analyse neighbourhood characteristics of land use patterns. **Computers, Environment and Urban Systems.** 28: 667-690.

Vermeiren, K., Anton, V. R., Maarten, L., Eria, S., and Paul, M. (2012). Urban growth of Kampala, Uganda: Pattern analysis and scenario development. **Landscape and urban planning.** 106: 199-206.

Wilson, E. H., Hurd, J. D., Civco, D. L., Prisloe, M. P., and Amold, C. (2003). Development of a geospatial model to quantify, describe and map urban growth. **Remote Sensing of Environment.** 86: 275-285.

World Bank Group. (2015). **East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth**. Washington, DC 20433.

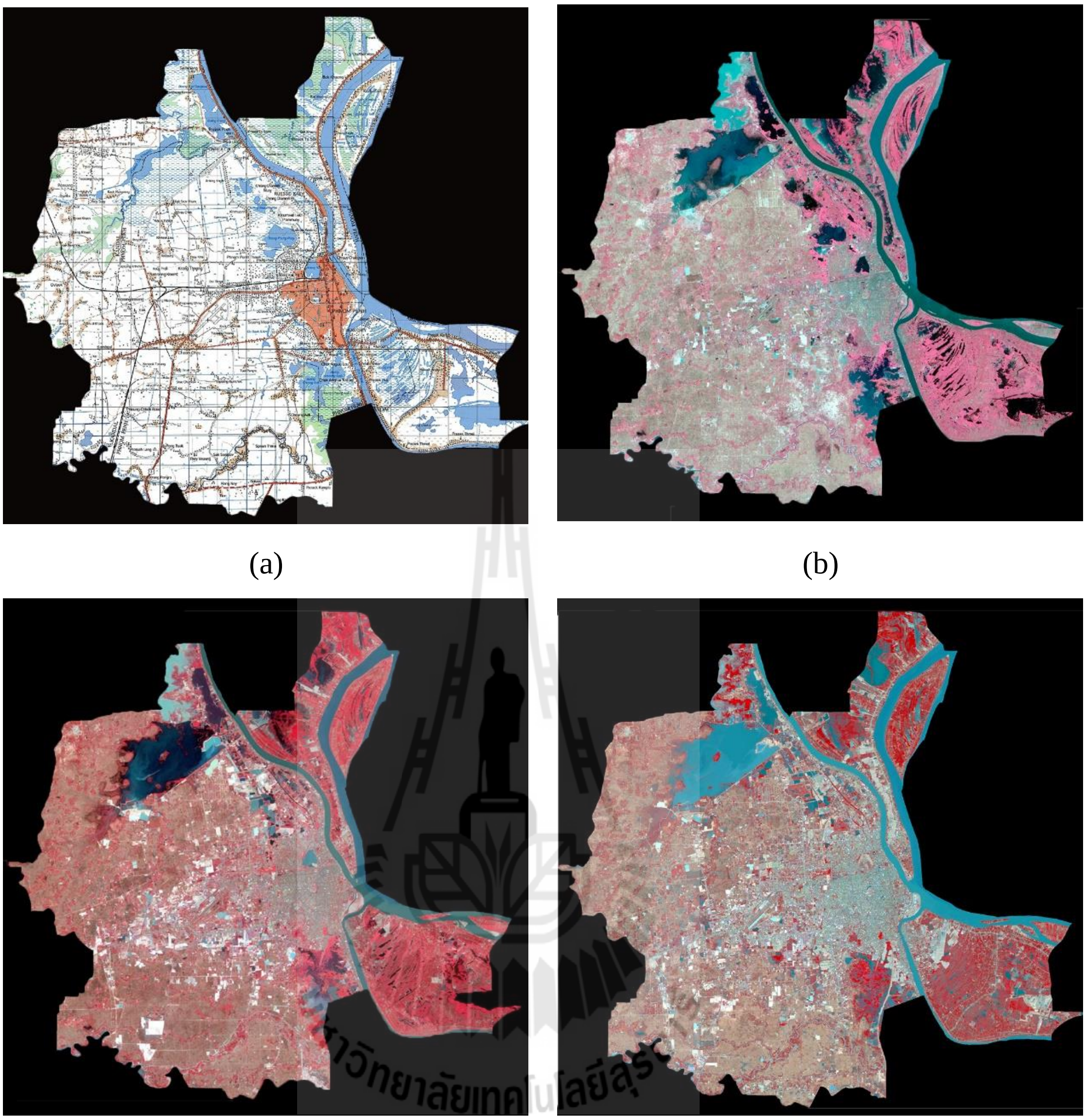
Yang, X., and Lo, C. P. (2003). Modelling urban growth and landscape changes in the Atlanta metropolitan area. **International journal of geographical information science**: 17(5): 463-488.

Yeh, A. G-O. (1999). Urban planning and GIS. **Geographical Information Systems: Principles and Technical Issues**, 2nd edition, New York, John Wiley & Sons. 877-888.

APPENDICES

APPENDIX A

DATA COLLECTION: TOPOGRAPHIC MAP, LANDSAT IMAGES FOR 2002, 2009 AND 2015, PUBLIC SERVICE DATA AND PHNOM PENH MASTER PLAN IN 2035



(a) (b)

(c) (d)

Figure A.1 Topographic map and Landsat images of Phnom Penh: (a) Topographic map in 2002, (b) Landsat 7 image, (c) Landsat 5 image, and (d) Landsat 8 image.

2003

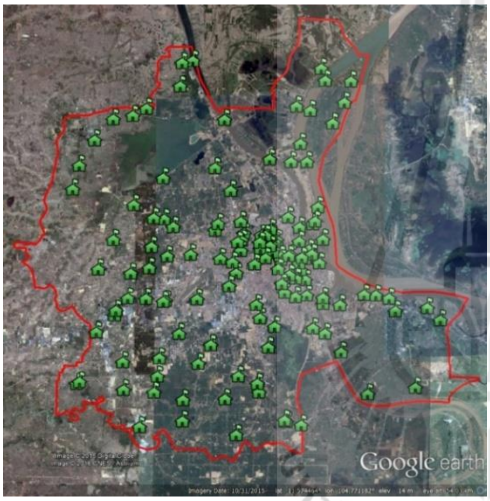
2010

2015

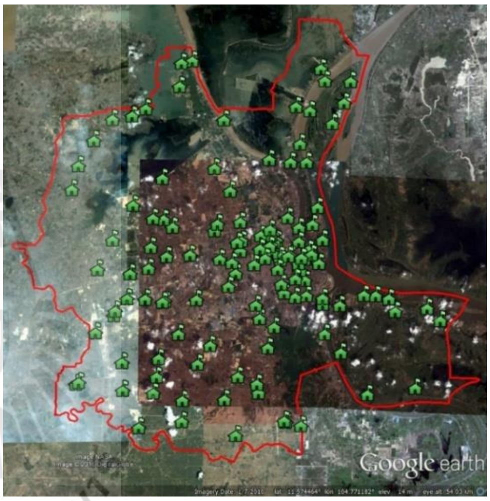
Figure A.2 Google earth high resolution images.

Main road and road in 2015

Road in 2009

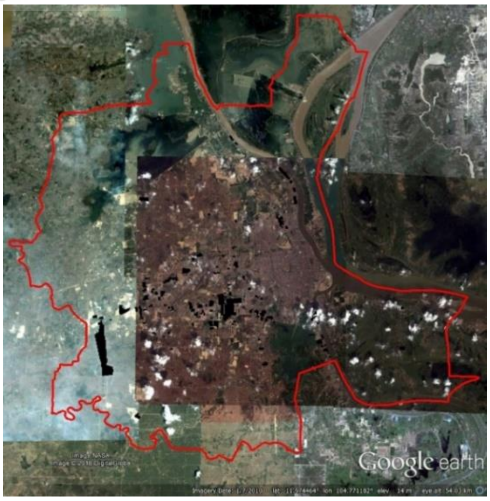


School in 2015



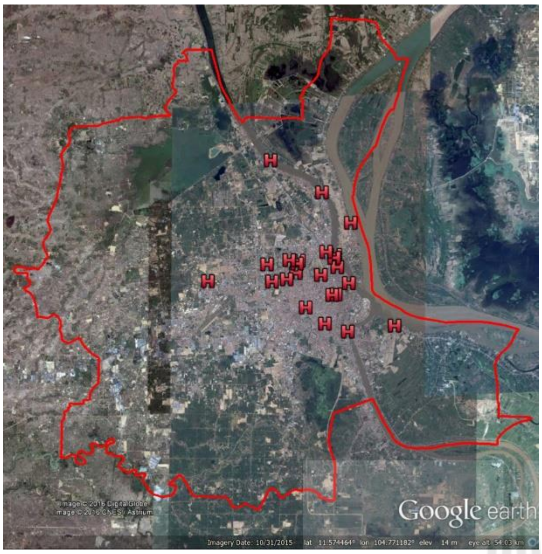
School in 2009

Industry in 2015

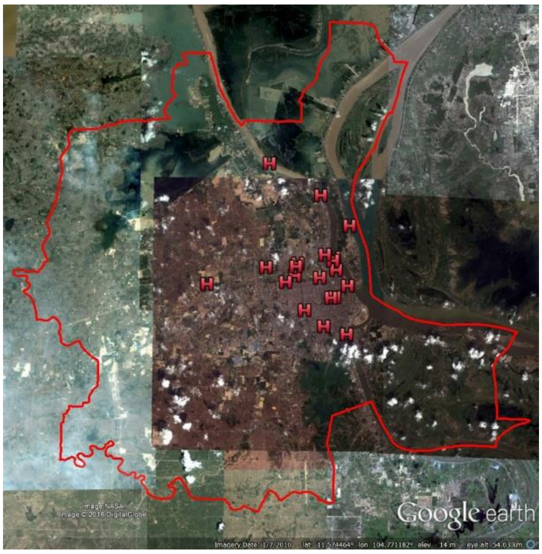


Industry in 2009

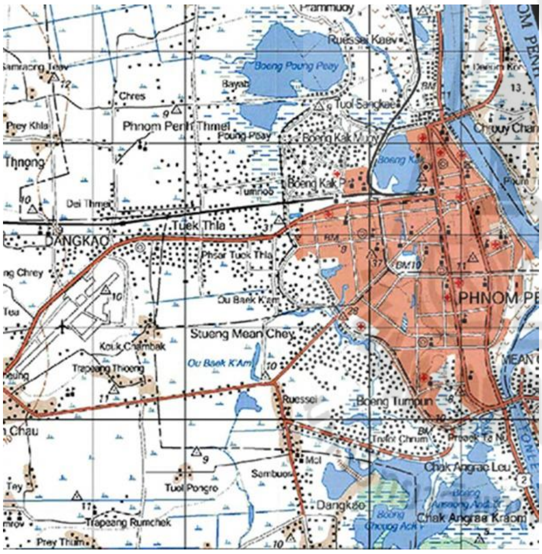
Figure A.3 Public service data.



Health Center in 2015



Health Center in 2009



Public Service data in 2002

Industry in 2002

Figure A.3 Public service data (Continued).

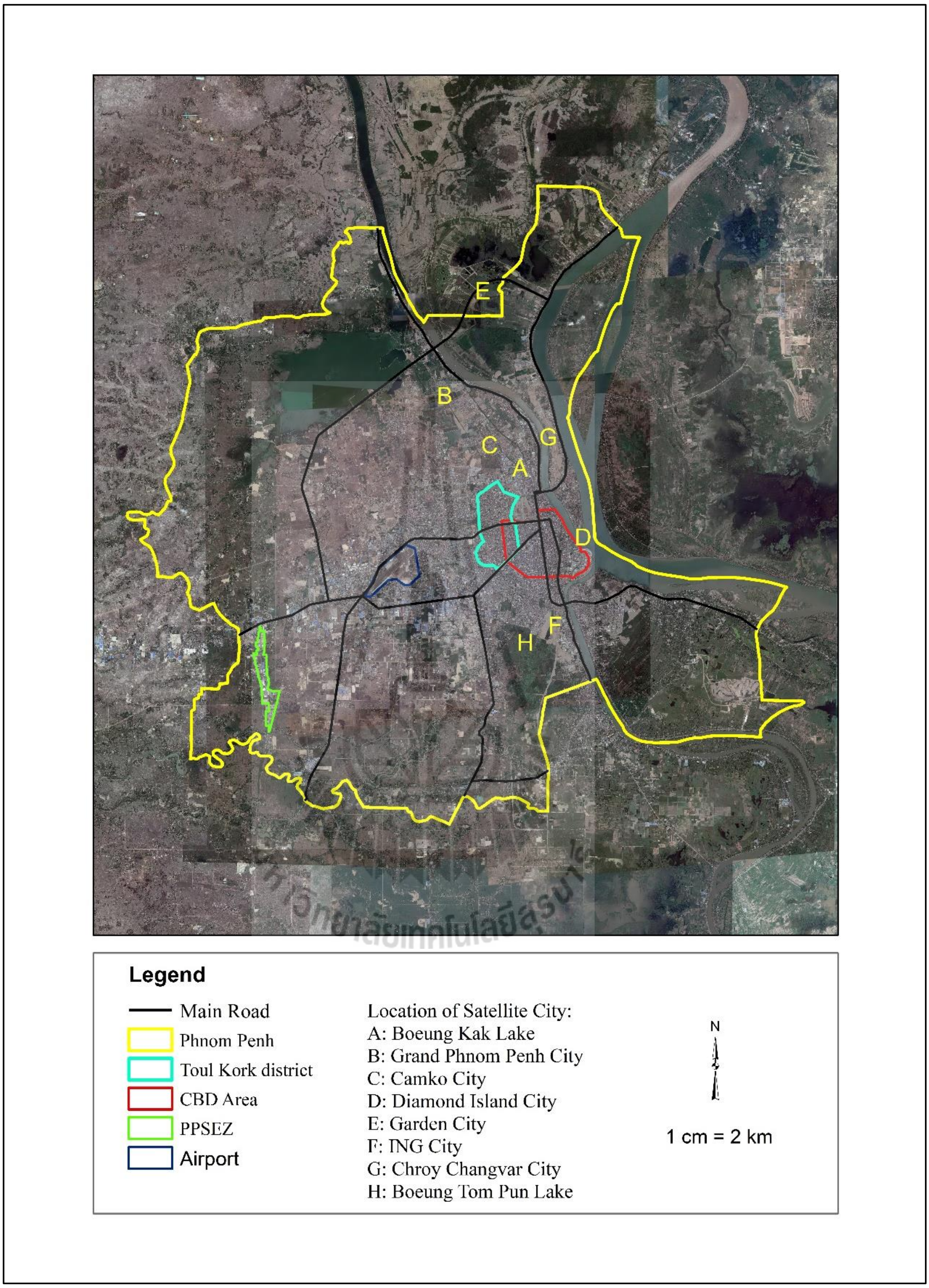


Figure A.4 Location of satellite cities in Phnom Penh.

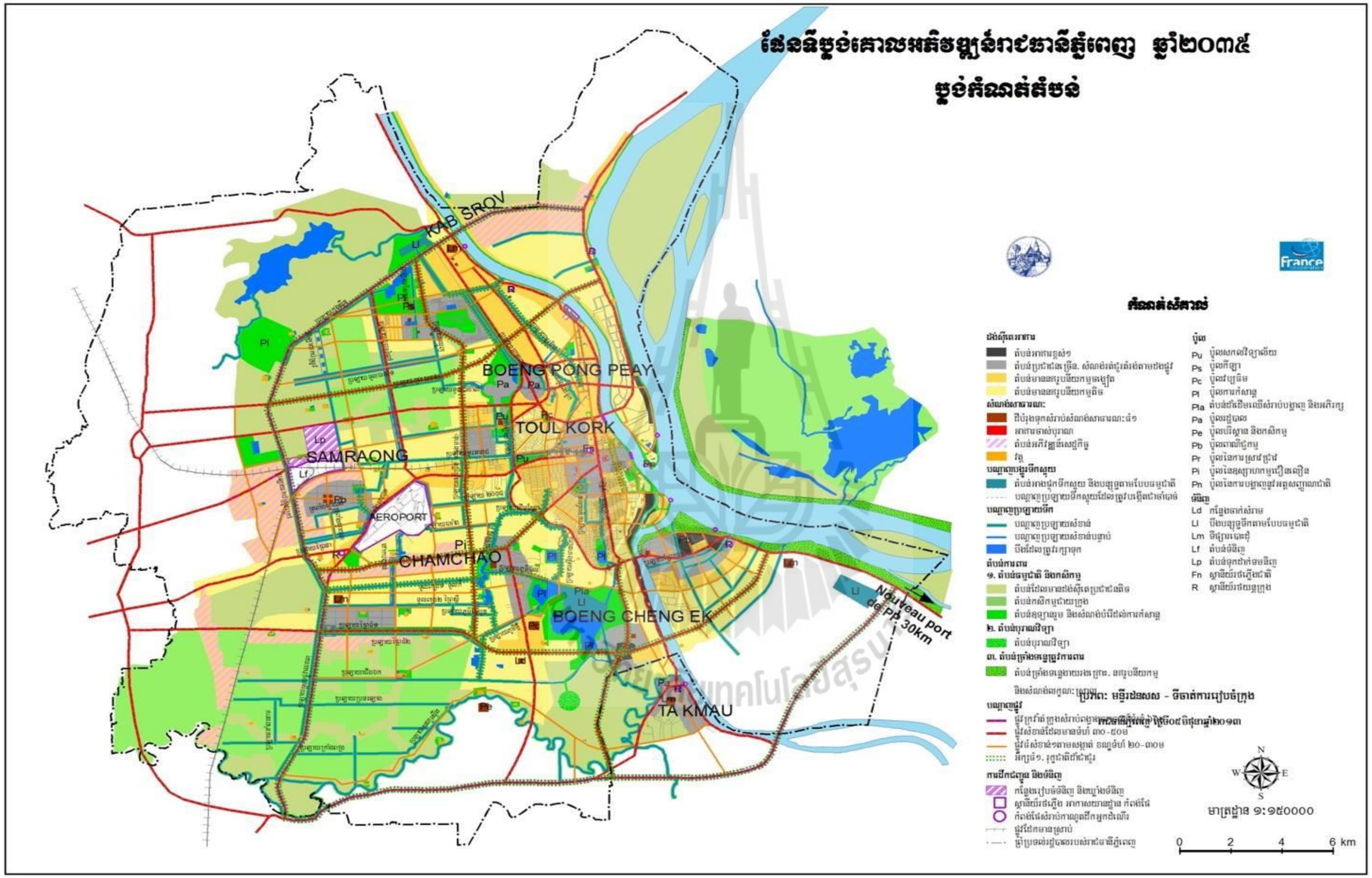


Figure A.5 Phnom Penh master plan in 2035.

APPENDIX B

ERROR MATRIX TABLE BETWEEN REFERENCE AND CLASSIFIED DATA

Table B1 Error matrix between reference and classified data of Landsat 8.

	Reference Data L8					
Classified Data	**U**	**A**	**V**	**M**	**W**	**Row Total**
Urban and built-up area (U)	31	1	0	0	0	32
Arable land (A)	1	72	7	1	1	82
Vegetation (V)	0	4	51	0	0	55
Miscellaneous land (M)	0	0	0	7	1	8
Water body (W)	0	0	0	0	26	26
Column Total	32	77	58	8	28	203

Table B2 Error matrix between reference and classified data of Landsat 7.

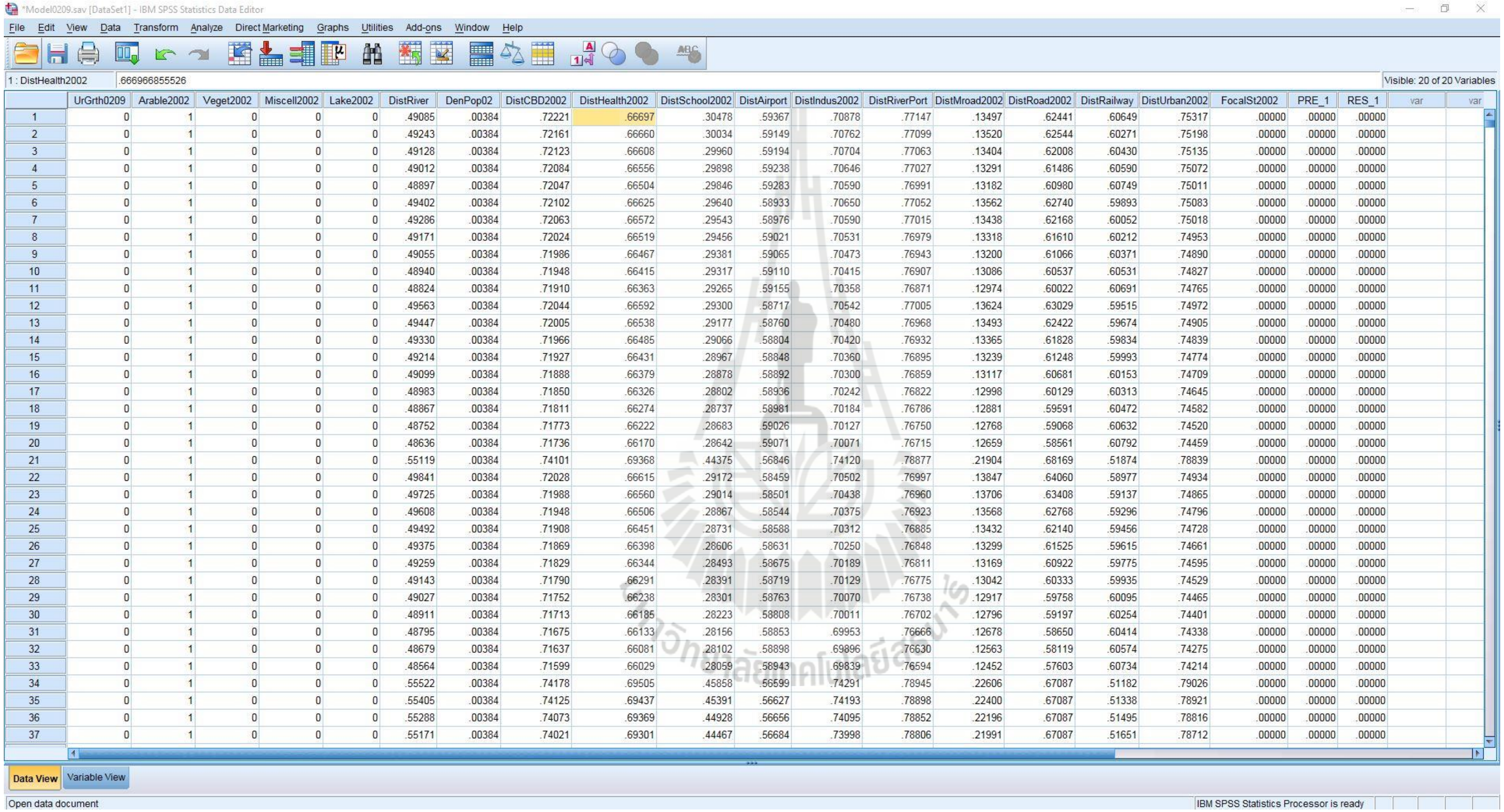
	Reference Data L7					
Classified Data	**U**	**A**	**V**	**M**	**W**	**Row Total**
Urban and built-up area (U)	11	0	0	0	0	11
Arable land (A)	3	78	5	0	0	86
Vegetation (V)	1	5	58	0	0	64
Miscellaneous land (M)	0	1	0	2	0	3
Water body (W)	0	0	1	0	38	39
Column Total	15	84	64	2	38	203

Table B3 Error matrix between reference and classified data of Landsat 5.

	Reference Data L5					
Classified Data	**U**	**A**	**V**	**M**	**W**	**Row Total**
Urban and built-up area (U)	19	2	0	0	0	21
Arable land (A)	3	63	1	0	0	67
Vegetation (V)	0	6	57	0	0	63
Miscellaneous land (M)	1	1	0	11	0	13
Water body (W)	0	0	4	0	35	39
Column Total	23	72	62	11	35	203

APPENDIX C

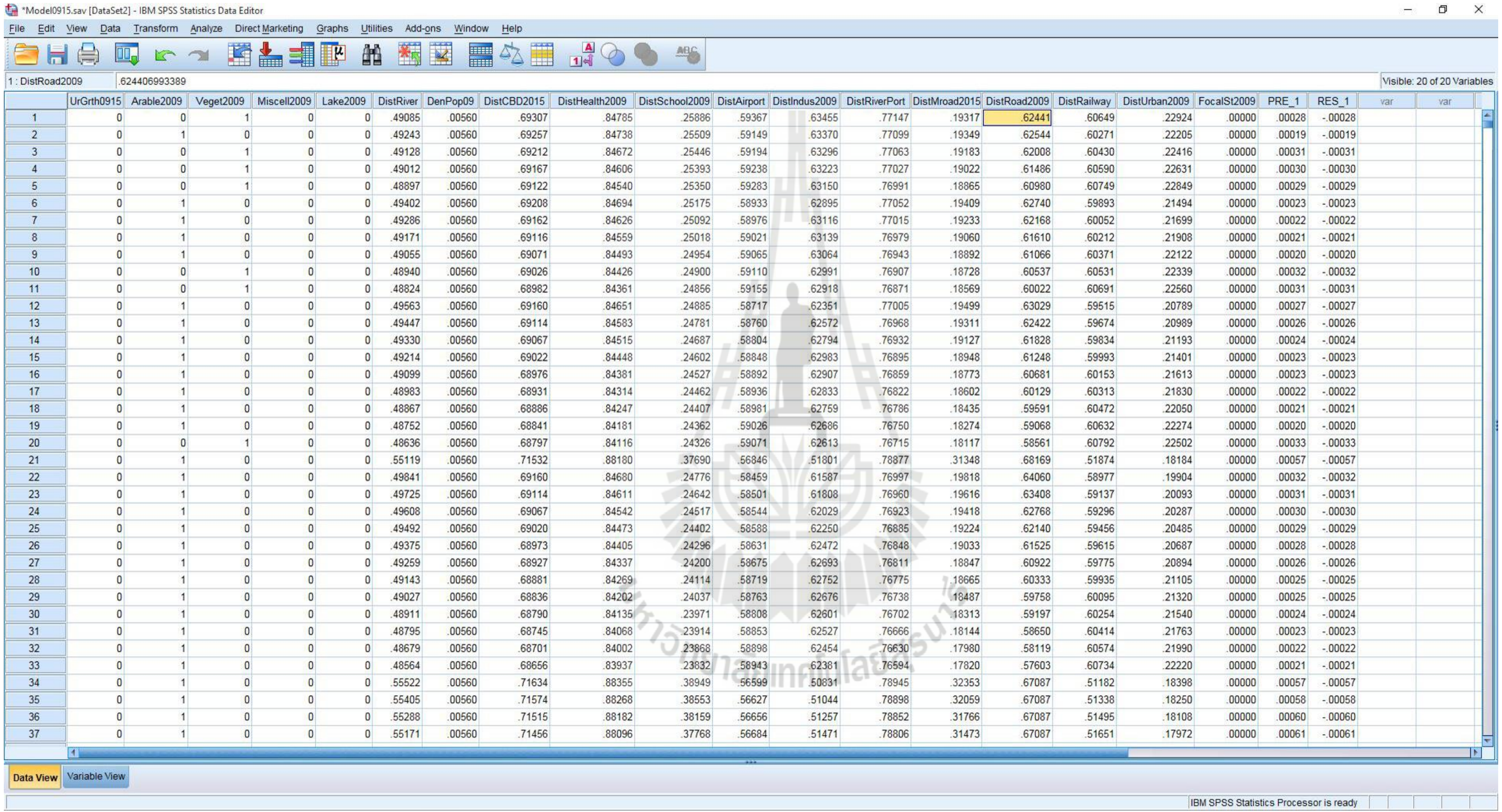
EXAMPLE OF INFLUENTIAL FACTOR DATA IN ASCII FILES



	UrGrth0209	Arable2002	Veget2002	Miscell2002	Lake2002	DistRiver	DenPop02	DistCBD2002	DistHealth2002	DistSchool2002
1	0	1	0	0	0	.49085	.00384	.72221	.66697	.30478
2	0	1	0	0	0	.49243	.00384	.72161	.66660	.30034
3	0	1	0	0	0	.49128	.00384	.72123	.66608	.29960
4	0	1	0	0	0	.49012	.00384	.72084	.66556	.29898
5	0	1	0	0	0	.48897	.00384	.72047	.66504	.29846
6	0	1	0	0	0	.49402	.00384	.72102	.66625	.29640
7	0	1	0	0	0	.49286	.00384	.72063	.66572	.29543
8	0	1	0	0	0	.49171	.00384	.72024	.66519	.29456
9	0	1	0	0	0	.49055	.00384	.71986	.66467	.29381
10	0	1	0	0	0	.48940	.00384	.71948	.66415	.29317
11	0	1	0	0	0	.48824	.00384	.71910	.66363	.29265
12	0	1	0	0	0	.49563	.00384	.72044	.66592	.29300
13	0	1	0	0	0	.49447	.00384	.72005	.66538	.29177
14	0	1	0	0	0	.49330	.00384	.71966	.66485	.29066
15	0	1	0	0	0	.49214	.00384	.71927	.66431	.28967
16	0	1	0	0	0	.49099	.00384	.71888	.66379	.28878
17	0	1	0	0	0	.48983	.00384	.71850	.66326	.28802
18	0	1	0	0	0	.48867	.00384	.71811	.66274	.28737
19	0	1	0	0	0	.48752	.00384	.71773	.66222	.28683
20	0	1	0	0	0	.48636	.00384	.71736	.66170	.28642
21	0	1	0	0	0	.55119	.00384	.74101	.69368	.44375
22	0	1	0	0	0	.49841	.00384	.72028	.66615	.29172
23	0	1	0	0	0	.49725	.00384	.71988	.66560	.29014
24	0	1	0	0	0	.49608	.00384	.71948	.66506	.28867
25	0	1	0	0	0	.49492	.00384	.71908	.66451	.28731
26	0	1	0	0	0	.49375	.00384	.71869	.66398	.28606
27	0	1	0	0	0	.49259	.00384	.71829	.66344	.28493
28	0	1	0	0	0	.49143	.00384	.71790	.66291	.28391
29	0	1	0	0	0	.49027	.00384	.71752	.66238	.28301
30	0	1	0	0	0	.48911	.00384	.71713	.66185	.28223
31	0	1	0	0	0	.48795	.00384	.71675	.66133	.28156
32	0	1	0	0	0	.48679	.00384	.71637	.66081	.28102
33	0	1	0	0	0	.48564	.00384	.71599	.66029	.28059
34	0	1	0	0	0	.55522	.00384	.74178	.69505	.45858
35	0	1	0	0	0	.55405	.00384	.74125	.69437	.45391
36	0	1	0	0	0	.55288	.00384	.74073	.69369	.44928
37	0	1	0	0	0	.55171	.00384	.74021	.69301	.44467

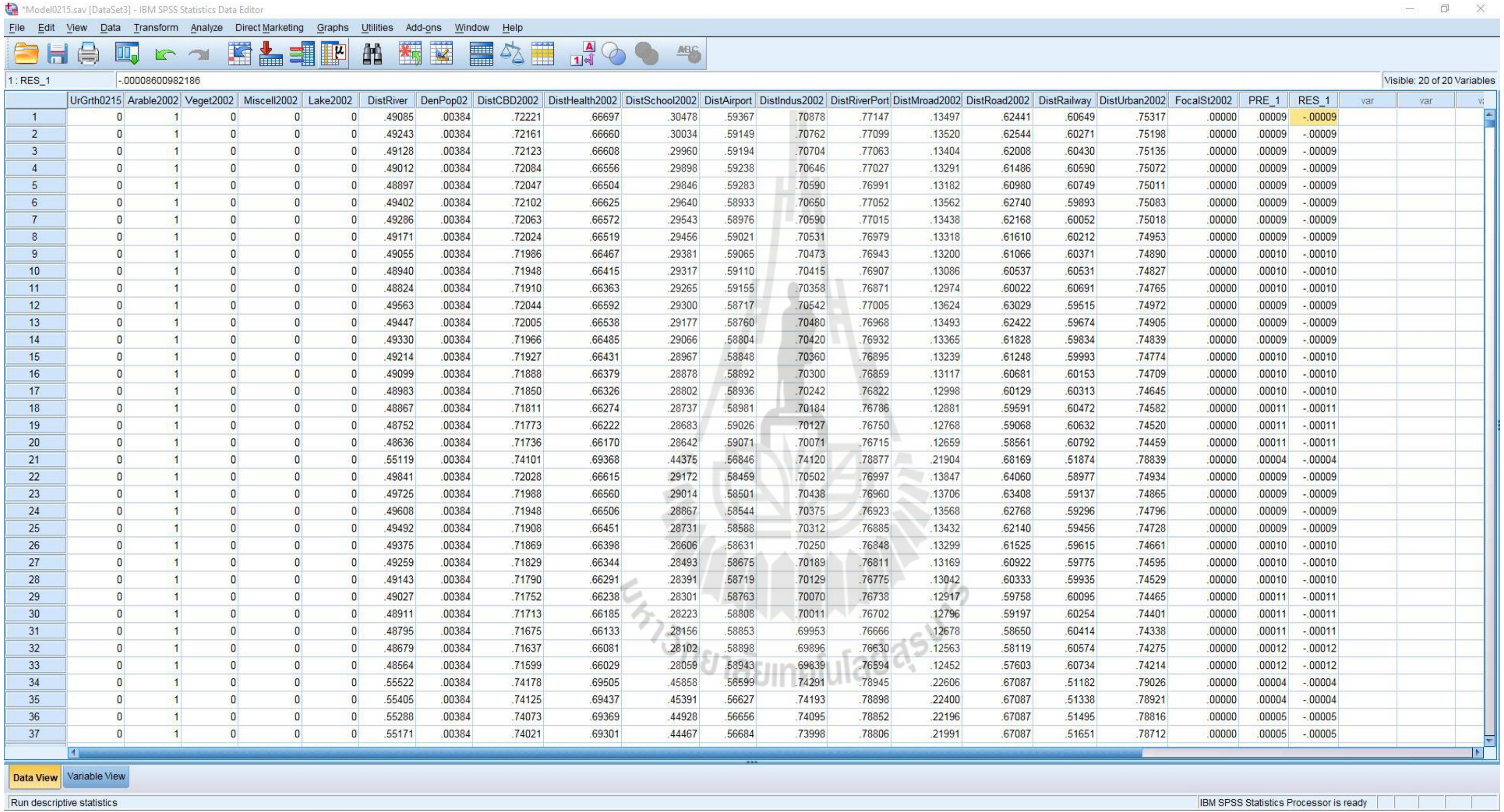
	DistAirport	DistIndus2002	DistRiverPort	DistMroad2002	DistRoad2002	DistRailway	DistUrban2002	FocalSt2002	PRE_1	RES_1	var	var
1	.59367	.70878	.77147	.13497	.62441	.60649	.75317	.00000	.00000	.00000		
2	.59149	.70762	.77099	.13520	.62544	.60271	.75198	.00000	.00000	.00000		
3	.59194	.70704	.77063	.13404	.62008	.60430	.75135	.00000	.00000	.00000		
4	.59238	.70646	.77027	.13291	.61486	.60590	.75072	.00000	.00000	.00000		
5	.59283	.70590	.76991	.13182	.60980	.60749	.75011	.00000	.00000	.00000		
6	.58933	.70650	.77052	.13562	.62740	.59893	.75083	.00000	.00000	.00000		
7	.58976	.70590	.77015	.13438	.62168	.60052	.75018	.00000	.00000	.00000		
8	.59021	.70531	.76979	.13318	.61610	.60212	.74953	.00000	.00000	.00000		
9	.59065	.70473	.76943	.13200	.61066	.60371	.74890	.00000	.00000	.00000		
10	.59110	.70415	.76907	.13086	.60537	.60531	.74827	.00000	.00000	.00000		
11	.59155	.70358	.76871	.12974	.60022	.60691	.74765	.00000	.00000	.00000		
12	.58717	.70542	.77005	.13624	.63029	.59515	.74972	.00000	.00000	.00000		
13	.58760	.70480	.76968	.13493	.62422	.59674	.74905	.00000	.00000	.00000		
14	.58804	.70420	.76932	.13365	.61828	.59834	.74839	.00000	.00000	.00000		
15	.58848	.70360	.76895	.13239	.61248	.59993	.74774	.00000	.00000	.00000		
16	.58892	.70300	.76859	.13117	.60681	.60153	.74709	.00000	.00000	.00000		
17	.58936	.70242	.76822	.12998	.60129	.60313	.74645	.00000	.00000	.00000		
18	.58981	.70184	.76786	.12881	.59591	.60472	.74582	.00000	.00000	.00000		
19	.59026	.70127	.76750	.12768	.59068	.60632	.74520	.00000	.00000	.00000		
20	.59071	.70071	.76715	.12659	.58561	.60792	.74459	.00000	.00000	.00000		
21	.56846	.74120	.78877	.21904	.68169	.51874	.78839	.00000	.00000	.00000		
22	.58459	.70502	.76997	.13847	.64060	.58977	.74934	.00000	.00000	.00000		
23	.58501	.70438	.76960	.13706	.63408	.59137	.74865	.00000	.00000	.00000		
24	.58544	.70375	.76923	.13568	.62768	.59296	.74796	.00000	.00000	.00000		
25	.58588	.70312	.76885	.13432	.62140	.59456	.74728	.00000	.00000	.00000		
26	.58631	.70250	.76848	.13299	.61525	.59615	.74661	.00000	.00000	.00000		
27	.58675	.70189	.76811	.13169	.60922	.59775	.74595	.00000	.00000	.00000		
28	.58719	.70129	.76775	.13042	.60333	.59935	.74529	.00000	.00000	.00000		
29	.58763	.70070	.76738	.12917	.59758	.60095	.74465	.00000	.00000	.00000		
30	.58808	.70011	.76702	.12796	.59197	.60254	.74401	.00000	.00000	.00000		
31	.58853	.69953	.76666	.12678	.58650	.60414	.74338	.00000	.00000	.00000		
32	.58898	.69896	.76630	.12563	.58119	.60574	.74275	.00000	.00000	.00000		
33	.58943	.69839	.76594	.12452	.57603	.60734	.74214	.00000	.00000	.00000		
34	.56599	.74291	.78945	.22606	.67087	.51182	.79026	.00000	.00000	.00000		
35	.56627	.74193	.78898	.22400	.67087	.51338	.78921	.00000	.00000	.00000		
36	.56656	.74095	.78852	.22196	.67087	.51495	.78816	.00000	.00000	.00000		
37	.56684	.73998	.78806	.21991	.67087	.51651	.78712	.00000	.00000	.00000		

Figure C.1 SPSS for LR model of 2002-2009.



	UrGrth0915	Arable2009	Veget2009	Miscell2009	Lake2009	DistRiver	DenPop09	DistCBD2015	DistHealth2009	DistSchool2009	DistAirport	DistIndus2009	DistRiverPort	DistMroad2015	DistRoad2009	DistRailway	DistUrban2009	FocalSt2009	PRE_1	RES_1	var	var
1	0	0	1	0	0	.49085	.00560	.69307	.84785	.25886	.59367	.63455	.77147	.19317	.62441	.60649	.22924	.00000	.00028	-.00028		
2	0	1	0	0	0	.49243	.00560	.69257	.84738	.25509	.59149	.63370	.77099	.19349	.62544	.60271	.22205	.00000	.00019	-.00019		
3	0	0	1	0	0	.49128	.00560	.69212	.84672	.25446	.59194	.63296	.77063	.19183	.62008	.60430	.22416	.00000	.00031	-.00031		
4	0	0	1	0	0	.49012	.00560	.69167	.84606	.25393	.59238	.63223	.77027	.19022	.61486	.60590	.22631	.00000	.00030	-.00030		
5	0	0	1	0	0	.48897	.00560	.69122	.84540	.25350	.59283	.63150	.76991	.18865	.60980	.60749	.22849	.00000	.00029	-.00029		
6	0	1	0	0	0	.49402	.00560	.69208	.84694	.25175	.58933	.62895	.77052	.19409	.62740	.59893	.21494	.00000	.00023	-.00023		
7	0	1	0	0	0	.49286	.00560	.69162	.84626	.25092	.58976	.63116	.77015	.19233	.62168	.60052	.21699	.00000	.00022	-.00022		
8	0	1	0	0	0	.49171	.00560	.69116	.84559	.25018	.59021	.63139	.76979	.19060	.61610	.60212	.21908	.00000	.00021	-.00021		
9	0	1	0	0	0	.49055	.00560	.69071	.84493	.24954	.59065	.63064	.76943	.18892	.61066	.60371	.22122	.00000	.00020	-.00020		
10	0	0	1	0	0	.48940	.00560	.69026	.84426	.24900	.59110	.62991	.76907	.18728	.60537	.60531	.22339	.00000	.00032	-.00032		
11	0	0	1	0	0	.48824	.00560	.68982	.84361	.24856	.59155	.62918	.76871	.18569	.60022	.60691	.22560	.00000	.00031	-.00031		
12	0	1	0	0	0	.49563	.00560	.69160	.84651	.24885	.58717	.62351	.77005	.19499	.63029	.59515	.20789	.00000	.00027	-.00027		
13	0	1	0	0	0	.49447	.00560	.69114	.84583	.24781	.58760	.62572	.76968	.19311	.62422	.59674	.20989	.00000	.00026	-.00026		
14	0	1	0	0	0	.49330	.00560	.69067	.84515	.24687	.58804	.62794	.76932	.19127	.61828	.59834	.21193	.00000	.00024	-.00024		
15	0	1	0	0	0	.49214	.00560	.69022	.84448	.24602	.58848	.62983	.76895	.18948	.61248	.59993	.21401	.00000	.00023	-.00023		
16	0	1	0	0	0	.49099	.00560	.68976	.84381	.24527	.58892	.62907	.76859	.18773	.60681	.60153	.21613	.00000	.00023	-.00023		
17	0	1	0	0	0	.48983	.00560	.68931	.84314	.24462	.58936	.62833	.76822	.18602	.60129	.60313	.21830	.00000	.00022	-.00022		
18	0	1	0	0	0	.48867	.00560	.68886	.84247	.24407	.58981	.62759	.76786	.18435	.59591	.60472	.22050	.00000	.00021	-.00021		
19	0	1	0	0	0	.48752	.00560	.68841	.84181	.24362	.59026	.62686	.76750	.18274	.59068	.60632	.22274	.00000	.00020	-.00020		
20	0	0	1	0	0	.48636	.00560	.68797	.84116	.24326	.59071	.62613	.76715	.18117	.58561	.60792	.22502	.00000	.00033	-.00033		
21	0	1	0	0	0	.55119	.00560	.71532	.88180	.37690	.56846	.51801	.78877	.31348	.68169	.51874	.18184	.00000	.00057	-.00057		
22	0	1	0	0	0	.49841	.00560	.69160	.84680	.24776	.58459	.61587	.76997	.19818	.64060	.58977	.19904	.00000	.00032	-.00032		
23	0	1	0	0	0	.49725	.00560	.69114	.84611	.24642	.58501	.61808	.76960	.19616	.63408	.59137	.20093	.00000	.00031	-.00031		
24	0	1	0	0	0	.49608	.00560	.69067	.84542	.24517	.58544	.62029	.76923	.19418	.62768	.59296	.20287	.00000	.00030	-.00030		
25	0	1	0	0	0	.49492	.00560	.69020	.84473	.24402	.58588	.62250	.76885	.19224	.62140	.59456	.20485	.00000	.00029	-.00029		
26	0	1	0	0	0	.49375	.00560	.68973	.84405	.24296	.58631	.62472	.76848	.19033	.61525	.59615	.20687	.00000	.00028	-.00028		
27	0	1	0	0	0	.49259	.00560	.68927	.84337	.24200	.58675	.62693	.76811	.18847	.60922	.59775	.20894	.00000	.00026	-.00026		
28	0	1	0	0	0	.49143	.00560	.68881	.84269	.24114	.58719	.62752	.76775	.18665	.60333	.59935	.21105	.00000	.00025	-.00025		
29	0	1	0	0	0	.49027	.00560	.68836	.84202	.24037	.58763	.62676	.76738	.18487	.59758	.60095	.21320	.00000	.00025	-.00025		
30	0	1	0	0	0	.48911	.00560	.68790	.84135	.23971	.58808	.62601	.76702	.18313	.59197	.60254	.21540	.00000	.00024	-.00024		
31	0	1	0	0	0	.48795	.00560	.68745	.84068	.23914	.58853	.62527	.76666	.18144	.58650	.60414	.21763	.00000	.00023	-.00023		
32	0	1	0	0	0	.48679	.00560	.68701	.84002	.23868	.58898	.62454	.76630	.17980	.58119	.60574	.21990	.00000	.00022	-.00022		
33	0	1	0	0	0	.48564	.00560	.68656	.83937	.23832	.58943	.62381	.76594	.17820	.57603	.60734	.22220	.00000	.00021	-.00021		
34	0	1	0	0	0	.55522	.00560	.71634	.88355	.38949	.56599	.50831	.78945	.32353	.67087	.51182	.18398	.00000	.00057	-.00057		
35	0	1	0	0	0	.55405	.00560	.71574	.88268	.38553	.56627	.51044	.78898	.32059	.67087	.51338	.18250	.00000	.00058	-.00058		
36	0	1	0	0	0	.55288	.00560	.71515	.88182	.38159	.56656	.51257	.78852	.31766	.67087	.51495	.18108	.00000	.00060	-.00060		
37	0	1	0	0	0	.55171	.00560	.71456	.88096	.37768	.56684	.51471	.78806	.31473	.67087	.51651	.17972	.00000	.00061	-.00061		

Figure C.2 SPSS for LR model of 2009-2015.


*Model0215.sav [DataSet3] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1 : RES_1 -.00008600982186 Visible: 20 of 20 Variables

	UrGrth0215	Arable2002	Veget2002	Miscell2002	Lake2002	DistRiver	DenPop02	DistCBD2002	DistHealth2002	DistSchool2002	DistAirport	DistIndus2002	DistRiverPort	DistMroad2002	DistRoad2002	DistRailway	DistUrban2002	FocalSt2002	PRE_1	RES_1
1	0	1	0	0	0	.49085	.00384	.72221	.66697	.30478	.59367	.70878	.77147	.13497	.62441	.60649	.75317	.00000	.00009	-.00009
2	0	1	0	0	0	.49243	.00384	.72161	.66660	.30034	.59149	.70762	.77099	.13520	.62544	.60271	.75198	.00000	.00009	-.00009
3	0	1	0	0	0	.49128	.00384	.72123	.66608	.29960	.59194	.70704	.77063	.13404	.62008	.60430	.75135	.00000	.00009	-.00009
4	0	1	0	0	0	.49012	.00384	.72084	.66556	.29898	.59238	.70646	.77027	.13291	.61486	.60590	.75072	.00000	.00009	-.00009
5	0	1	0	0	0	.48897	.00384	.72047	.66504	.29846	.59283	.70590	.76991	.13182	.60980	.60749	.75011	.00000	.00009	-.00009
6	0	1	0	0	0	.49402	.00384	.72102	.66625	.29640	.58933	.70650	.77052	.13562	.62740	.59893	.75083	.00000	.00009	-.00009
7	0	1	0	0	0	.49286	.00384	.72063	.66572	.29543	.58976	.70590	.77015	.13438	.62168	.60052	.75018	.00000	.00009	-.00009
8	0	1	0	0	0	.49171	.00384	.72024	.66519	.29456	.59021	.70531	.76979	.13318	.61610	.60212	.74953	.00000	.00009	-.00009
9	0	1	0	0	0	.49055	.00384	.71986	.66467	.29381	.59065	.70473	.76943	.13200	.61066	.60371	.74890	.00000	.00010	-.00010
10	0	1	0	0	0	.48940	.00384	.71948	.66415	.29317	.59110	.70415	.76907	.13086	.60537	.60531	.74827	.00000	.00010	-.00010
11	0	1	0	0	0	.48824	.00384	.71910	.66363	.29265	.59155	.70358	.76871	.12974	.60022	.60691	.74765	.00000	.00010	-.00010
12	0	1	0	0	0	.49563	.00384	.72044	.66592	.29300	.58717	.70542	.77005	.13624	.63029	.59515	.74972	.00000	.00009	-.00009
13	0	1	0	0	0	.49447	.00384	.72005	.66538	.29177	.58760	.70480	.76968	.13493	.62422	.59674	.74905	.00000	.00009	-.00009
14	0	1	0	0	0	.49330	.00384	.71966	.66485	.29066	.58804	.70420	.76932	.13365	.61828	.59834	.74839	.00000	.00009	-.00009
15	0	1	0	0	0	.49214	.00384	.71927	.66431	.28967	.58848	.70360	.76895	.13239	.61248	.59993	.74774	.00000	.00010	-.00010
16	0	1	0	0	0	.49099	.00384	.71888	.66379	.28878	.58892	.70300	.76859	.13117	.60681	.60153	.74709	.00000	.00010	-.00010
17	0	1	0	0	0	.48983	.00384	.71850	.66326	.28802	.58936	.70242	.76822	.12998	.60129	.60313	.74645	.00000	.00010	-.00010
18	0	1	0	0	0	.48867	.00384	.71811	.66274	.28737	.58981	.70184	.76786	.12881	.59591	.60472	.74582	.00000	.00011	-.00011
19	0	1	0	0	0	.48752	.00384	.71773	.66222	.28683	.59026	.70127	.76750	.12768	.59068	.60632	.74520	.00000	.00011	-.00011
20	0	1	0	0	0	.48636	.00384	.71736	.66170	.28642	.59071	.70071	.76715	.12659	.58561	.60792	.74459	.00000	.00011	-.00011
21	0	1	0	0	0	.55119	.00384	.74101	.69368	.44375	.56846	.74120	.78877	.21904	.68169	.51874	.78839	.00000	.00004	-.00004
22	0	1	0	0	0	.49841	.00384	.72028	.66615	.29172	.58459	.70502	.76997	.13847	.64060	.58977	.74934	.00000	.00009	-.00009
23	0	1	0	0	0	.49725	.00384	.71988	.66560	.29014	.58501	.70438	.76960	.13706	.63408	.59137	.74865	.00000	.00009	-.00009
24	0	1	0	0	0	.49608	.00384	.71948	.66506	.28867	.58544	.70375	.76923	.13568	.62768	.59296	.74796	.00000	.00009	-.00009
25	0	1	0	0	0	.49492	.00384	.71908	.66451	.28731	.58588	.70312	.76885	.13432	.62140	.59456	.74728	.00000	.00009	-.00009
26	0	1	0	0	0	.49375	.00384	.71869	.66398	.28606	.58631	.70250	.76848	.13299	.61525	.59615	.74661	.00000	.00010	-.00010
27	0	1	0	0	0	.49259	.00384	.71829	.66344	.28493	.58675	.70189	.76811	.13169	.60922	.59775	.74595	.00000	.00010	-.00010
28	0	1	0	0	0	.49143	.00384	.71790	.66291	.28391	.58719	.70129	.76775	.13042	.60333	.59935	.74529	.00000	.00010	-.00010
29	0	1	0	0	0	.49027	.00384	.71752	.66238	.28301	.58763	.70070	.76738	.12917	.59758	.60095	.74465	.00000	.00011	-.00011
30	0	1	0	0	0	.48911	.00384	.71713	.66185	.28223	.58808	.70011	.76702	.12796	.59197	.60254	.74401	.00000	.00011	-.00011
31	0	1	0	0	0	.48795	.00384	.71675	.66133	.28156	.58853	.69953	.76666	.12678	.58650	.60414	.74338	.00000	.00011	-.00011
32	0	1	0	0	0	.48679	.00384	.71637	.66081	.28102	.58898	.69896	.76630	.12563	.58119	.60574	.74275	.00000	.00012	-.00012
33	0	1	0	0	0	.48564	.00384	.71599	.66029	.28059	.58943	.69839	.76594	.12452	.57603	.60734	.74214	.00000	.00012	-.00012
34	0	1	0	0	0	.55522	.00384	.74178	.69505	.45858	.56599	.74291	.78945	.22606	.67087	.51182	.79026	.00000	.00004	-.00004
35	0	1	0	0	0	.55405	.00384	.74125	.69437	.45391	.56627	.74193	.78898	.22400	.67087	.51338	.78921	.00000	.00004	-.00004
36	0	1	0	0	0	.55288	.00384	.74073	.69369	.44928	.56656	.74095	.78852	.22196	.67087	.51495	.78816	.00000	.00005	-.00005
37	0	1	0	0	0	.55171	.00384	.74021	.69301	.44467	.56684	.73998	.78806	.21991	.67087	.51651	.78712	.00000	.00005	-.00005

Data View Variable View

Run descriptive statistics IBM SPSS Statistics Processor is ready

Figure C.3 SPSS for LR model of 2002-2015.

APPENDIX D

HISTOGRAM OF PROBABILITY OF URBAN GROWTH MAP OF THREE PERIODS

Table D-1 Histogram of probability value for urban growth of 2002-2015 period.

Lower Limit	Upper Limit	Frequency	Area (sq.km)
0.00	0.12	562985	506.6865
0.12	0.13	10222	9.1998
0.13	0.14	8709	7.8381
0.14	0.15	7885	7.0965
0.15	0.16	7183	6.4647
0.16	0.17	6556	5.9004
0.17	0.18	6018	5.4162
0.18	0.19	5446	4.9014
0.19	0.2	5218	4.6962
0.2	0.21	4815	4.3335
0.21	0.22	4642	4.1778
0.22	0.23	4373	3.9357
0.23	0.24	4133	3.7197
0.24	0.25	3903	3.5127
0.25	0.26	3748	3.3732
0.26	0.27	3470	3.123
0.27	0.28	3412	3.0708
0.28	0.29	3347	3.0123
0.29	0.3	3334	3.0006
0.3	0.31	3217	2.8953
0.31	0.32	3049	2.7441
0.32	0.33	2886	2.5974
0.33	0.34	2866	2.5794
0.34	0.35	2887	2.5983
0.35	0.36	2843	2.5587
0.36	0.37	3023	2.7207
0.37	0.38	3032	2.7288
0.38	0.39	3005	2.7045
0.39	0.4	3155	2.8395

Table D-1 Histogram of probability value for urban growth of 2002-2015 period (Continued).

0.4	0.41	3066	2.7594
0.41	0.42	3100	2.79
0.42	0.43	3012	2.7108
0.43	0.44	2846	2.5614
0.44	0.45	2816	2.5344
0.45	0.46	2610	2.349
0.46	0.47	2593	2.3337
0.47	0.48	2502	2.2518
0.48	0.49	2397	2.1573
0.49	0.5	2285	2.0565
0.5	0.51	2318	2.0862
0.51	0.52	2293	2.0637
0.52	0.53	2202	1.9818
0.53	0.54	2207	1.9863
0.54	0.55	2186	1.9674
0.55	0.56	2228	2.0052
0.56	0.57	2149	1.9341
0.57	0.58	2074	1.8666
0.58	0.59	2049	1.8441
0.59	0.6	1905	1.7145
0.6	0.61	1875	1.6875
0.61	0.62	1685	1.5165
0.62	0.63	1584	1.4256
0.63	0.64	1432	1.2888
0.64	0.65	1456	1.3104
0.65	0.66	1386	1.2474
0.66	0.67	1467	1.3203
0.67	0.68	1465	1.3185

Table D-1 Histogram of probability value for urban growth of 2002-2015 period (Continued).

0.68	0.69	1419	1.2771
0.69	0.7	1407	1.2663
0.7	0.71	1438	1.2942
0.71	0.72	1211	1.0899
0.72	0.73	1007	0.9063
0.73	0.74	813	0.7317
0.74	0.75	733	0.6597
0.75	0.76	559	0.5031
0.76	0.77	443	0.3987
0.77	0.78	274	0.2466
0.78	0.79	205	0.1845
0.79	0.8	117	0.1053
0.8	0.81	97	0.0873
0.81	0.82	62	0.0558
0.82	0.83	14	0.0126
0.83	0.84	4	0.0036
0.84	0.85	1	0.0009

Table D-2 Histogram of probability value for urban growth of 2002-2009 period.

Lower Limit	Upper Limit	Frequency	Area (sq.km)
0.00	0.12	666788	600.1092
0.12	0.13	3432	3.0888
0.13	0.14	3413	3.0717
0.14	0.15	3472	3.1248
0.15	0.16	3037	2.7333
0.16	0.17	3077	2.7693
0.17	0.18	2937	2.6433

Table D-2 Histogram of probability value for urban growth of 2002-2009 period (Continued).

0.18	0.19	3165	2.8485
0.19	0.2	2969	2.6721
0.2	0.21	2591	2.3319
0.21	0.22	2351	2.1159
0.22	0.23	2457	2.2113
0.23	0.24	2301	2.0709
0.24	0.25	2213	1.9917
0.25	0.26	2070	1.863
0.26	0.27	2003	1.8027
0.27	0.28	1912	1.7208
0.28	0.29	1985	1.7865
0.29	0.3	1903	1.7127
0.3	0.31	1841	1.6569
0.31	0.32	1830	1.647
0.32	0.33	1960	1.764
0.33	0.34	1898	1.7082
0.34	0.35	1829	1.6461
0.35	0.36	1607	1.4463
0.36	0.37	1625	1.4625
0.37	0.38	1593	1.4337
0.38	0.39	1587	1.4283
0.39	0.4	1762	1.5858
0.4	0.41	1736	1.5624
0.41	0.42	1629	1.4661
0.42	0.43	1528	1.3752
0.43	0.44	1447	1.3023
0.44	0.45	1315	1.1835
0.45	0.46	1239	1.1151

Table D-2 Histogram of probability value for urban growth of 2002-2009 period (Continued).

0.46	0.47	1169	1.0521
0.47	0.48	1128	1.0152
0.48	0.49	1037	0.9333
0.49	0.5	1091	0.9819
0.5	0.51	984	0.8856
0.51	0.52	991	0.8919
0.52	0.53	993	0.8937
0.53	0.54	935	0.8415
0.54	0.55	906	0.8154
0.55	0.56	888	0.7992
0.56	0.57	795	0.7155
0.57	0.58	792	0.7128
0.58	0.59	801	0.7209
0.59	0.6	766	0.6894
0.6	0.61	762	0.6858
0.61	0.62	794	0.7146
0.62	0.63	713	0.6417
0.63	0.64	631	0.5679
0.64	0.65	581	0.5229
0.65	0.66	541	0.4869
0.66	0.67	555	0.4995
0.67	0.68	421	0.3789
0.68	0.69	377	0.3393
0.69	0.7	324	0.2916
0.7	0.71	280	0.252
0.71	0.72	213	0.1917
0.72	0.73	147	0.1323
0.73	0.74	103	0.0927

Table D-2 Histogram of probability value for urban growth of 2002-2009 period (Continued).

0.74	0.75	48	0.0432
0.75	0.76	43	0.0387
0.76	0.77	23	0.0207
0.77	0.78	15	0.0135
0.78	0.79	5	0.0045

Table D-3 Histogram of probability value for urban growth of 2009-2015.

Lower Limit	Upper Limit	Frequency	Area (sq.km)
0.00	0.12	643067	578.7603
0.12	0.13	6792	6.1128
0.13	0.14	6662	5.9958
0.14	0.15	6562	5.9058
0.15	0.16	6372	5.7348
0.16	0.17	5950	5.355
0.17	0.18	5484	4.9356
0.18	0.19	5187	4.6683
0.19	0.2	4597	4.1373
0.2	0.21	3929	3.5361
0.21	0.22	3778	3.4002
0.22	0.23	3707	3.3363
0.23	0.24	3376	3.0384
0.24	0.25	3323	2.9907
0.25	0.26	3071	2.7639
0.26	0.27	2957	2.6613
0.27	0.28	2898	2.6082
0.28	0.29	2963	2.6667
0.29	0.3	2633	2.3697

Table D-3 Histogram of probability value for urban growth of 2009-2015 period (Continued).

0.3	0.31	2266	2.0394
0.31	0.32	2043	1.8387
0.32	0.33	1845	1.6605
0.33	0.34	1546	1.3914
0.34	0.35	1322	1.1898
0.35	0.36	1102	0.9918
0.36	0.37	1062	0.9558
0.37	0.38	1013	0.9117
0.38	0.39	929	0.8361
0.39	0.4	820	0.738
0.4	0.41	879	0.7911
0.41	0.42	942	0.8478
0.42	0.43	1066	0.9594
0.43	0.44	1125	1.0125
0.44	0.45	1195	1.0755
0.45	0.46	1338	1.2042
0.46	0.47	1359	1.2231
0.47	0.48	1525	1.3725
0.48	0.49	1488	1.3392
0.49	0.5	1295	1.1655
0.5	0.51	1168	1.0512
0.51	0.52	1010	0.909
0.52	0.53	888	0.7992
0.53	0.54	862	0.7758
0.54	0.55	763	0.6867
0.55	0.56	772	0.6948
0.56	0.57	765	0.6885
0.57	0.58	716	0.6444

Table D-3 Histogram of probability value for urban growth of 2009-2015 period (Continued).

0.58	0.59	722	0.6498
0.59	0.6	713	0.6417
0.6	0.61	588	0.5292
0.61	0.62	444	0.3996
0.62	0.63	380	0.342
0.63	0.64	282	0.2538
0.64	0.65	254	0.2286
0.65	0.66	194	0.1746
0.66	0.67	137	0.1233
0.67	0.68	94	0.0846
0.68	0.69	75	0.0675
0.69	0.7	44	0.0396
0.7	0.71	11	0.0099
0.71	0.72	3	0.0027
0.72	0.73	1	0.0009

Table D-4 Histogram of probability value of future urbanization prediction map.

Lower Limit	**Upper Limit**	**Frequency**	**Area (sq.km)**
0.00	0.12	552079	496.8711
0.12	0.13	9273	8.3457
0.13	0.14	8554	7.6986
0.14	0.15	8356	7.5204
0.15	0.16	8284	7.4556
0.16	0.17	7951	7.1559
0.17	0.18	8092	7.2828
0.18	0.19	7876	7.0884
0.19	0.2	7387	6.6483

Table D-4 Histogram of probability value of future urbanization prediction map (Continued).

0.2	0.21	7459	6.7131
0.21	0.22	7067	6.3603
0.22	0.23	6768	6.0912
0.23	0.24	6551	5.8959
0.24	0.25	6403	5.7627
0.25	0.26	6074	5.4666
0.26	0.27	5631	5.0679
0.27	0.28	5239	4.7151
0.28	0.29	4935	4.4415
0.29	0.3	4567	4.1103
0.3	0.31	4441	3.9969
0.31	0.32	4262	3.8358
0.32	0.33	4007	3.6063
0.33	0.34	4056	3.6504
0.34	0.35	3778	3.4002
0.35	0.36	3579	3.2211
0.36	0.37	3373	3.0357
0.37	0.38	3176	2.8584
0.38	0.39	3028	2.7252
0.39	0.4	2887	2.5983
0.4	0.41	2570	2.313
0.41	0.42	2533	2.2797
0.42	0.43	2311	2.0799
0.43	0.44	2390	2.151
0.44	0.45	2260	2.034
0.45	0.46	2228	2.0052
0.46	0.47	2195	1.9755
0.47	0.48	2123	1.9107

Table D-4 Histogram of probability value of future urbanization prediction map (Continued).

0.48	0.49	1948	1.7532
0.49	0.5	1856	1.6704
0.5	0.51	1731	1.5579
0.51	0.52	1742	1.5678
0.52	0.53	1645	1.4805
0.53	0.54	1542	1.3878
0.54	0.55	1502	1.3518
0.55	0.56	1316	1.1844
0.56	0.57	1270	1.143
0.57	0.58	1091	0.9819
0.58	0.59	988	0.8892
0.59	0.6	925	0.8325
0.6	0.61	791	0.7119
0.61	0.62	792	0.7128
0.62	0.63	700	0.63
0.63	0.64	717	0.6453
0.64	0.65	657	0.5913
0.65	0.66	604	0.5436
0.66	0.67	563	0.5067
0.67	0.68	534	0.4806
0.68	0.69	515	0.4635
0.69	0.7	506	0.4554
0.7	0.71	476	0.4284
0.71	0.72	436	0.3924
0.72	0.73	418	0.3762
0.73	0.74	353	0.3177
0.74	0.75	303	0.2727
0.75	0.76	210	0.189

Table D-4 Histogram of probability value of future urbanization prediction map (Continued).

0.76	0.77	153	0.1377
0.77	0.78	112	0.1008
0.78	0.79	103	0.0927
0.79	0.8	63	0.0567
0.8	0.81	29	0.0261
0.81	0.82	10	0.009
0.82	0.83	10	0.009

CURRICULUM VITAE

Name: Kompheak Mom

Date of Birth: 18 February, 1991

Place of Birth: Phnom Penh, Cambodia.

Education:

2008 - 2013 Bachelor of Engineering, Major in Civil-Engineering, Norton University, Phnom Penh, Cambodia.

Grants and Fellowships:

2014 - 2016 The ASEAN European Academic University Network (ASEA-UNINET) Scholarship for Master of Science at School of Remote Sensing, Institute of Science, Suranaree University of Technology, Nakhon Rachasima Province, Thailand.

Position and Place of Work:

Graduated Student, School of Remote Sensing, Suranaree University of Technology, Nakhon Ratchasima, Thailand.