

# Masters Program in **Geospatial Technologies**



## **URBAN LAND COVER CHANGE DETECTION ANALYSIS AND MODELLING SPATIO-TEMPORAL GROWTH DYNAMICS USING REMOTE SENSING AND GIS TECHNIQUES**

**“A CASE STUDY OF DHAKA, BANGLADESH”**

**BAYES AHMED**

Dissertation submitted in partial fulfilment of the requirements  
for the Degree of *Master of Science in Geospatial Technologies*



ifgi  
Institute for Geoinformatics  
University of Münster

Supported by:



Erasmus Mundus

# URBAN LAND COVER CHANGE DETECTION ANALYSIS AND MODELLING SPATIO-TEMPORAL GROWTH DYNAMICS USING REMOTE SENSING AND GIS TECHNIQUES

“A CASE STUDY OF DHAKA, BANGLADESH”

Dissertation Supervised by

**Dr. Pedro Latorre Carmona**

Professor, Institute of New Imaging Technologies (INIT)

Universitat Jaume I (UJI), Castellón, Spain

Dissertation Co-Supervised by

**Dr. Mário Caetano**

Professor, Instituto Superior de Estatística e Gestão de Informação (ISEGI)

Universidade Nova de Lisboa (UNL), Lisbon, Portugal

**Dr. Edzer Pebesma**

Professor, Institute for Geoinformatics (ifgi)

Westfälische Wilhelms-Universität (WWU), Münster, Germany

**Dr. Nilanchal Patel**

Professor, Department of Remote Sensing

Birla Institute of Technology Mesra, Jharkhand, India

**February 2011**



ifgi  
Institute for Geoinformatics  
University of Münster

Supported by:



**Erasmus Mundus**

## **Candidate's Declaration**

This is to certify that this research work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources).

All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institutes.

It is hereby also declared that this dissertation or any part of it has not been submitted elsewhere for the award of any degree or diploma.

---

**BAYES AHMED**

Date: February 28, 2011



ifgi  
Institute for Geoinformatics  
University of Münster

Supported by:



**Erasmus Mundus**

## Dedication

I want to dedicate this research work to

My beloved Parents

*Md. Abdus Sattar*

and

*Mrs. Suariya Begum*



ifgi  
Institute for Geoinformatics  
University of Münster

Supported by:



Erasmus Mundus



## Abstract

Dhaka, the capital of Bangladesh, has undergone radical changes in its physical form, not only in its vast territorial expansion, but also through internal physical transformations over the last decades. In the process of urbanization, the physical characteristic of Dhaka is gradually changing as open spaces have been transformed into building areas, low land and water bodies into reclaimed builtup lands etc. This new urban fabric should be analyzed to understand the changes that have led to its creation.

The primary objective of this research is to predict and analyze the future urban growth of Dhaka City. Another objective is to quantify and investigate the characteristics of urban land cover changes (1989-2009) using the Landsat satellite images of 1989, 1999 and 2009. Dhaka City Corporation (DCC) and its surrounding impact areas have been selected as the study area. A fisher supervised classification method has been applied to prepare the base maps with five land cover classes. To observe the change detection, different spatial metrics have been used for quantitative analysis. Moreover, some post-classification change detection techniques have also been implemented. Then it is found that the 'builtup area' land cover type is increasing in high rate over the years. The major contributors to this change are 'fallow land' and 'water body' land cover types.

In the next stage, three different models have been implemented to simulate the land cover map of Dhaka city of 2009. These are named as 'Stochastic Markov (St\_Markov)' Model, 'Cellular Automata Markov (CA\_Markov)' Model and 'Multi Layer Perceptron Markov (MLP\_Markov)' Model. Then the best-fitted model has been selected based on various Kappa statistics values and also by implementing other model validation techniques. This is how the 'Multi Layer Perceptron Markov (MLP\_Markov)' Model has been qualified as the most suitable model for this research. Later, using the MLP\_Markov model, the land cover map of 2019 has been predicted. The MLP\_Markov model shows that 58% of the total study area will be converted into builtup area cover type in 2019.

The interpretation of depicting the future scenario in quantitative accounts, as demonstrated in this research, will be of great value to the urban planners and decision makers, for the future planning of modern Dhaka City.

**Key Words:** Remote Sensing, Land Cover, Markov Chain, Cellular Automata, Multi Layer Perceptron Neural Network, Change Detection, Supervised Classification, GIS.

## **Acknowledgement**

At the outset, all praises belong to the almighty ‘Allah’, the most merciful, the most beneficent to all the creatures and their dealings.

First of all, I would like to express my gratitude to the European Commission and Erasmus Mundus Consortium (Universitat Jaume I, Castellón, Spain; Westfälische Wilhelms-Universität, Münster, Germany and Universidade Nova de Lisboa, Portugal) for awarding me the Erasmus Mundus scholarship in Master of Science in Geospatial Technologies. It is a great opportunity for my lifetime experiences to study in the reputed universities of Europe.

It is a great pleasure to acknowledge my sincere and greatest gratitude to my dissertation supervisor, Dr. Pedro Latorre Carmona, Professor, Institute of New Imaging Technologies, University Jaume I, Spain; for his untiring effort, careful supervision, thoughtful suggestions and enduring guidance at every stage of this research. This thesis would not be in its current shape without his continuous exertion and support.

I am very grateful to my dissertation co-supervisors Dr. Mário Caetano, Dr. Edzer Pebesma and Dr. Nilanchal Patel; for accepting my thesis proposal at the very early stage and also for their valuable time and effort in contributing information and practical suggestions on numerous occasions.

My heartfelt thanks goes to Dr. Raquib Ahmed, Professor and Director, Institute of Environmental Science, University of Rajshahi, Bangladesh; for his initial encouragement for conducting this research and helping me developing the research proposal.

I also want to thank Dr. Filiberto Pla Bañón, Professor, Institute of New Imaging Technologies, Universitat Jaume I, Spain; for his time and some comments. I am pleased to extend my gratitude to Prof. Dr. Joaquín Huerta Guijarro, Dolores C. Apanewicz, Prof. Dr. Jorge Mateu and Dr. Christoph Brox; for their support and hospitality during my stay in Spain and Germany.

My thanks and best wishes also conveyed to my classmates and lovely friends, from all over the world, for sharing their knowledge and giving me inspirations during the last eighteen months in Europe. Special thanks goes to Dipu da, Anik vai, Shiuli vabi, Irene, Mauri, Carlos, Sherzod, Neba, Shahin vai, Diyan vai, Pathak, Freska and Pearl for their patronage and helping me coping with this new and challenging European environment.

Finally yet importantly, I want to express deep gratitude and indebtedness to my beloved parents for a life-long love and affection. I would also like to thank them for their continuous inspiration and encouragement, regarding the completion of this thesis and for their overall support throughout the years of my studies.

# Index of the Text

<b>Content</b>	<b>Page No</b>
<b>Abstract</b>	i
<b>Acknowledgement</b>	ii
<b>Index of the Text</b>	iii
<b>Index of Figures</b>	viii
<b>Index of Tables</b>	x
<b>Abbreviations and Acronyms</b>	xii
<b><u>Chapter 1: Introduction</u></b>	
1.1 Background of the Research	1
1.2 Statement of the Problem	1
1.3 Study Area Profile	5
1.4 Objectives of the Research	9
1.5 Research Hypotheses	9
1.5.1 Related to Objective 1	9
1.5.2 Related to Objective 2	9
1.6 Limitations of the Research	10
1.6.1 Collection of Satellite Images	10
1.6.2 Seasonal Variation	10
1.6.3 Collection of Reference Data	10
<b><u>Chapter 2: Theoretical Framework and Methodology</u></b>	
2.1 Basic Terminologies	11
2.1.1 Remote Sensing (RS)	11
2.1.2 Geographic Information System (GIS)	11
2.1.3 Land	12
2.2 Literature Review	12
2.2.1 Examples Related to Land Cover Change Detection	13
2.2.2 Examples Related to Future Land Cover Prediction	14
2.3 Methodology of the Research	14

2.3.1 Selection of Study Area	14
2.3.2 Problem Identification and Research Objectives	14
2.3.3 Data Collection	16
2.3.3.1 Satellite Images	16
2.3.3.2 Reference Data	18
2.3.3.3 Literature Review	18
2.3.4 Base Map Preparation and Accuracy Assessment	18
2.3.5 Change Detection Analysis	18
2.3.6 Model Calibration/ Simulation	18
2.3.7 Model Validation and Selection	19
2.3.8 Future Prediction	19
2.3.9 Directions for Future Planning	19
2.3.10 Report Writing	19
2.4 Tools Used for this Research	19

### **Chapter 3: Base Map Preparation and Accuracy Assessment**

3.1 Image Enhancement	20
3.2 Composite Generation	20
3.3 Image Classification	22
3.3.1 Training Site Development	22
3.3.2 Signature Development	24
3.3.3 Classification	24
3.3.4 Generalization	24
3.4 Accuracy Assessment	26
3.4.1 Assessment Procedure	27
3.4.2 Results and Discussion	30

### **Chapter 4: Land Cover Change Detection Analysis**

4.1 Change Detection	31
4.2 Terminologies	31
4.3 Analysis and Interpretation of the Change Detection Techniques	32
4.3.1 Number of Patches and Largest Patch Index	32
4.3.2 Edge Density and Mean Fractal Dimension Index	33
4.3.3 Mean Euclidean Nearest-Neighbour Distance and CPLAND	34

4.3.4 Change in Area	35
4.3.5 Gains and Losses by Category	35
4.3.6 Contributors to Net Change Experienced by Builtup Area	36
4.3.7 Transition to Builtup Area	36
4.3.8 Gains and Losses in Land Cover Types	38
4.4 Summary of Land Cover Change Detection Analysis	38

## **Chapter 5: Stochastic Markov Model**

5.1 Stochastic Process	40
5.2 Markov Chain	40
5.2.1 Markov Property	41
5.2.2 Transition Matrix for a Markov Chain	41
5.2.3 Example of Markov Chain	42
5.2.3.1 Weather Prediction	43
5.3 Stochastic Markov Model	43

## **Chapter 6: Cellular Automata Markov Model**

6.1 Cellular Automata	48
6.1.1 What are Cellular Automata?	48
6.1.2 The Elements of Cellular Automata	48
6.1.3 The Cell Space	49
6.1.4 The Cell States	50
6.1.5 The Cell Neighbourhood	50
6.1.6 The Transition Rules	51
6.1.7 The Temporal Space	52
6.1.8 Mathematical Notation of Cellular Automata	52
6.1.9 Running a Simulation	53
6.2 Cellular Automata Markov Model	54
6.3 How CA_Markov Model Works?	54
6.3.1 Suitability Maps for Land Cover Classes	56
6.3.2 Preparing Suitability Maps	57
6.3.4 Future Prediction	63

## **Chapter 7: Multi Layer Perceptron Markov Model**

7.1 Artificial Neural Network	64
7.2 Basic Concept of Artificial Neural Network (ANN)	64
7.2.1 Types of Artificial Neural Network	65
7.3 Multi Layer Perceptron (MLP)	65
7.3.1 Input Layer	66
7.3.2 Hidden Layer	66
7.3.3 Output Layer	66
7.3.4 The Feed-Forward Concept of MLP Neural Network	67
7.3.5 Number of Nodes	68
7.3.6 Number of Training Samples and Iterations	68
7.4 Multi Layer Perceptron Markov Modelling	69
7.4.1 Testing Potential Explanatory Power	72
7.4.2 Transition Potential Modelling	73
7.4.3 Future Prediction	75

## **Chapter 8: Model Validation and Selection**

8.1 Model Validation	76
8.1.1 Per Category Method	77
8.1.2 Fraction Correct	78
8.1.3 Fuzzy Sets	78
8.1.4 Fuzzy Kappa	79
8.2 Actual Base Maps vs. Simulated Maps	81
8.2.1 Base Map (2009) vs. St_Markov (2009)	81
8.2.1.1 Analysis of the Results of St_Markov	84
8.2.2 Base Map (2009) vs. CA_Markov (2009)	85
8.2.2.1 Analysis of the Results of CA_Markov	88
8.2.3 Base Map (2009) vs. MLP_Markov (2009)	89
8.3 Model Selection	91

## **Chapter 9: Future Prediction and Analysis**

9.1 Future Prediction	92
9.1.1 Creating Boolean Images (2009)	92
9.1.2 Creating Distance Images (2009)	93

9.1.3 Creating Land Cover Transition Image	94
9.1.4 Selecting Driving Variables	94
9.1.5 Testing Potential Explanatory Power of the Driving Variables	96
9.1.6 Transition Potential Modelling	96
9.1.7 Markov Chain Analysis	98
9.2 Analysis of the Predicted Map	99
9.3 Limitations of MLP_Markov Model	99

## **Chapter 10: Recommendations and Conclusions**

10.1 Answers to Research Questions	101
10.1.1 Answers Related to Objective 1	101
10.1.2 Answers Related to Objective 2	101
10.2 Recommendations for Future Work	102
10.3 Recommendations for Real World Plan Preparation	103
10.4 Epilogue	104
<b>References</b>	105

## **Appendices**

<b>Appendix A:</b> Photographs: Some Existing Urban Problems within Dhaka City	113
<b>Appendix B:</b> Reference Maps	114
<b>Appendix C:</b> Details of Accuracy Assessment	117
C 1 Basic Terminologies	117
C 1.1 Ground Truth	117
C 1.2 Overall Accuracy	117
C 1.3 Producer's Accuracy	117
C 1.4 User's Accuracy	117
C 1.5 Kappa	118
C 1.6 Kno	120
C 1.7 Klocation	121
C 1.8 Khisto	121
C 2 Accuracy Assessment Report	122
C 2.1 Accuracy Assessment Report (1989)	122
C 2.2 Accuracy Assessment Report (1999)	123
C 2.3 Accuracy Assessment Report (2009)	123

<b>Appendix D: Terminologies Used for Land Cover Change Detection Analysis</b>	125
D 1 Patch	125
D 2 Landscape	125
D 3 Core Area	125
D 4 Number of Patches (NP)	126
D 5 Edge Density (ED)	126
D 6 Largest Patch Index (LPI)	126
D 7 Core Area Percentage of Landscape (CPLAND)	127
D 8 Mean Euclidean Nearest-Neighbour Distance (ENN_MN)	127
D 9 Mean Fractal Dimension Index (FRAC_MN)	127
<b>Appendix E: Quantitative Values for Change Detection Analysis</b>	128
<b>Appendix F: Summary of Land Cover Classification Statistics (1989 - 2009)</b>	129
<b>Appendix G: Land Use Suitability</b>	130
G 1 Land Use Suitability Analysis	130
G 2 Different Methods for Land Use/Cover Suitability Analysis	130
G 2.1 Computer-Assisted Overlay Mapping Technique	130
G 2.2 Multi Criteria Decision Making Methods	130
G 2.3 Artificial Intelligence (AI) Methods	131
<b>Appendix H: Sigmoid Function and Cramer's V</b>	132
H 1 Sigmoid Function	132
H 2 Cramer's V	132

## **Index of Figures**

Figure 1.1: Changing Patterns of Dhaka City in Area and Population	3
Figure 1.2: Satellite Images Showing Urban Growth of Dhaka	6
Figure 1.3: Location of Dhaka City	7
Figure 1.4: Location of the Study Area within Greater Dhaka City	8
Figure 2.1: Flow Chart of Methodology	15
Figure 2.2: Dhaka City Corporation and its Surroundings	17
Figure 3.1: Composites Using Different Band Combinations	20
Figure 3.2: False Color Composite (RGB=4, 3 and 2) Maps	21
Figure 3.3: Land Cover Maps of Dhaka City	25
Figure 3.4: Random Points Cell Listing	27



Figure 3.5: Selected Stratified Random Points for Ground Truthing	29
Figure 4.1: Number of Patches and Largest Patch Index	32
Figure 4.2: Edge Density and Mean Fractal Dimension Index	33
Figure 4.3: ENN_MN and CPLAND	34
Figure 4.4: Land Cover Change in Area (Percentage)	35
Figure 4.5: Gains and Losses of Land Covers by Category	36
Figure 4.6: Contributors to Net Change Experienced by Builtup Area	37
Figure 4.7: Transition of Other Land Cover Types into Builtup Area	37
Figure 4.8: Gains and Losses in Land Cover Types (1989-2009)	39
Figure 5.1: Example of a Markov Chain	41
Figure 5.2: Stochastic Weather Model	43
Figure 5.3: Markovian Conditional Probability Images	45
Figure 5.4: Stochastic Markov Predicted Land Cover Map (2009)	47
Figure 5.5: Final St_Markov Predicted Land Cover Map (2009)	47
Figure 6.1: One-Dimensional Cellular Automata	49
Figure 6.2: Two-Dimensional Cellular Automata Grid	49
Figure 6.3: Examples of Variegated Cells	50
Figure 6.4: Two-Dimensional CA Neighbourhoods	51
Figure 6.5: The $3 \times 3$ Mean Contiguity Filter for CA_Markov Modelling	55
Figure 6.6: Fuzzy Linear Membership Function	58
Figure 6.7: Boolean Images of each Land Cover Type (1999)	59
Figure 6.8: Distance Images of each Land Cover Type (1999)	60
Figure 6.9: Suitability Images of each Land Cover Type	61
Figure 6.10: Aggregated Land Cover Suitability Map of Dhaka City (1999)	62
Figure 6.11: CA_Markov Projected Land Cover Map of Dhaka City (2009)	63
Figure 7.1: Artificial Neural Network Processing Element	64
Figure 7.2: A Multi Layer Perceptron Neural Network Model	65
Figure 7.3: Working Methodology of a Neuron in MLP Network	66
Figure 7.4: Transitions of Land Covers between 1989 and 1999	69
Figure 7.5: Transition from All to Builtup Area (1989-1999)	70
Figure 7.6: Distance Image of Transition from All to Builtup Area	71
Figure 7.7: Empirical Likelihood Image (1989-1999)	71
Figure 7.8: RMS Error Monitoring Curve	73
Figure 7.9: Transition Potential Maps	74

Figure 7.10: MLP_Markov Projected Land Cover Map of Dhaka City	75
Figure 8.1: Cell-by-cell Map Comparison	76
Figure 8.2: Cell-by-cell Map Comparison of Two Maps	77
Figure 8.3: Per Category Map Comparison of Two Maps	78
Figure 8.4: Fuzzy vs. Crisp Set Membership Functions	79
Figure 8.5: Difference between ‘Cell-by-cell’ and ‘Fuzzy’ Comparison	80
Figure 8.6: Levels of Agreement for Fuzzy Comparison Method	80
Figure 8.7: Maps for Model Validation	81
Figure 8.8: Levels of Agreement for Kappa	82
Figure 8.9: Per Category Comparison Method	83
Figure 8.10: Maps for Model Validation	85
Figure 8.11: Levels of Agreement for Kappa	85
Figure 8.12: Per Category Comparison Method	86
Figure 8.13: Fuzzy Kappa Result Map	87
Figure 8.14: Maps for Model Validation	89
Figure 8.15: Levels of Agreement for Kappa	89
Figure 8.16: Per Category Comparison Method	90
Figure 9.1: Boolean Images of each Land Cover Type (2009)	92
Figure 9.2: Distance Images of each Land Cover Type (2009)	93
Figure 9.3: Transition from All to Builtup Area (1999-2009)	94
Figure 9.4: Distance Image of Transition from All to Builtup Area	95
Figure 9.5: Empirical Likelihood Image (1999-2009)	95
Figure 9.6: RMS Error Monitoring Curve	97
Figure 9.7: Transition Potential Maps	97
Figure 9.8: MLP_Markov Projected Land Cover Map of Dhaka City	98
Figure 9.9: Change in Area (%) over the Years (1989-2019)	99
Figure 9.10: Gains and Losses of Land Cover Types (2009-2019)	100
Figure D 1: Classes of Landscape Pattern	125
Figure H 1: Sigmoid Function	132

### **Index of Tables**

Table 1.1: Growth of Dhaka City in Urban Agglomerations (1950-2025)	2
Table 1.2: Historical Growth of Dhaka City in Terms of Area	4
Table 2.1: Details of Landsat Satellite Images	16

Table 3.1: Properties of the Processed Raster Images for Analysis	23
Table 3.2: Details of the Land Cover Types	23
Table 3.3: Accuracy Assessment Cell Array	28
Table 5.1: Markov Probability of Changing among Land Cover Types	44
Table 5.2: Cells Expected to Transition to Different Classes	44
Table 5.3: Stochastic Random Process for Selecting Land Cover Type	46
Table 6.1: The Factor Weights Evaluated for the Suitability Map (1999)	62
Table 7.1: Cramer's V of the Driving Factors	72
Table 7.2: Running Statistics of MLP Neural Network	73
Table 7.3: Transition Probabilities Grid for Markov Chain	75
Table 8.1: Per Category Land Cover Change	82
Table 8.2: Per Category Kappa Statistics	84
Table 8.3: Fuzzy Kappa per Category	87
Table 8.4: Per Category Land Cover Change	88
Table 8.5: Per Category Kappa Statistics	88
Table 8.6: Per Category Kappa Statistics	91
Table 8.7: Overall Kappa Statistics and Fraction Correct	91
Table 9.1: Cramer's V of the Driving Factors	96
Table 9.2: Running Statistics of MLP Neural Network	96
Table 9.3: Transition Probabilities Grid for Markov Chain	98
Table C 1: An Example of Error Matrix	118
Table C 2: Contingency Table	119
Table C 3: Strength of Agreement for Kappa Statistic	120
Table C 4: Error Matrix (1989)	122
Table C 5: Accuracy Totals (1989)	122
Table C 6: Conditional Kappa for each Category (1989)	122
Table C 7: Error Matrix (1999)	123
Table C 8: Accuracy Totals (1999)	123
Table C 9: Conditional Kappa for each Category (1999)	123
Table C 10: Error Matrix (2009)	124
Table C 11: Accuracy Totals (2009)	124
Table C 12: Conditional Kappa for each Category (2009)	124

## Abbreviations and Acronyms

<b>AHP</b>	Analytical Hierarchy Analysis
<b>AI</b>	Artificial Intelligence
<b>ANN</b>	Artificial Neural Network
<b>BBS</b>	Bangladesh Bureau of Statistics
<b>BTM</b>	Bangladesh Transverse Mercator
<b>BUET</b>	Bangladesh University of Engineering and Technology
<b>CA</b>	Cellular Automata
<b>CA_Markov</b>	Cellular Automata Markov Model
<b>CPLAND</b>	Core Area Percentage of Landscape
<b>DAP</b>	Detail Area Plan
<b>DCC</b>	Dhaka City Corporation
<b>DMA</b>	Dhaka Metropolitan Area
<b>DMDP</b>	Dhaka Metropolitan Development Plan
<b>DSMA</b>	Dhaka Statistical Metropolitan Area
<b>ED</b>	Edge Density
<b>EIU</b>	Economist Intelligence Unit
<b>ENN_MN</b>	Mean Euclidean Nearest Neighbour Distance
<b>ETM+</b>	Enhanced Thematic Mapper Plus
<b>FAO</b>	Food and Agriculture Organization
<b>FRAC_MN</b>	Mean Fractal Dimension Index
<b>GIS</b>	Geographic Information System

<b>GoB</b>	Government of the People’s Republic of Bangladesh
<b>GPS</b>	Global Positioning System
<b>LPI</b>	Largest Patch Index
<b>MCDM</b>	Multi Criteria Decision Making
<b>MCE</b>	Multi Criteria Evaluation
<b>MLP</b>	Multi Layer Perceptron
<b>MLP_Markov</b>	Multi Layer Perceptron Markov Model
<b>MOLA</b>	Multi Objective Land Allocation
<b>NASA</b>	National Aeronautics and Space Administration
<b>NDVI</b>	Normalized Differential Vegetation Index
<b>NP</b>	Number of Patches
<b>OWA</b>	Ordered Weighted Averaging
<b>RAJUK</b>	<i>Rajdhani Unnayan Kartripakkh</i> [Bangla] (Capital City Planning and Development Authority)
<b>RS</b>	Remote Sensing
<b>SoB</b>	Survey of Bangladesh
<b>St_Markov</b>	Stochastic Markov Model
<b>TM</b>	Thematic Mapper
<b>UAP</b>	Urban Area Plan
<b>USGS</b>	United States Geological Survey
<b>UTM</b>	Universal Transverse Mercator
<b>WGS</b>	World Geodetic System
<b>WLC</b>	Weighed Linear Combination

# Chapter 1

## Introduction

---

### **1.1 Background of the Research**

Urbanization is one of the most evident human induced global changes worldwide. In the past 200 years, the world population has increased 6 times and the urban population has multiplied 100 times [1].

Like many other cities in the world Dhaka, the capital of Bangladesh, is also the outcome of spontaneous rapid growth without any prior or systematic planning. As the growth of population in Dhaka is taking place at an exceptionally rapid rate, it has become one of the most populous Mega Cities<sup>1</sup> in the world.

Dhaka City has undergone radical changes in its physical form, not only in its vast territorial expansion, but also through internal physical transformations over the last decades. These have created entirely new kinds of urban fabric.

In the process of urbanization, the physical characteristics of Dhaka City are gradually changing as plots and open spaces have been transformed into building areas, open squares into car parks, low land and water bodies into reclaimed built-up lands etc.

This new urban fabric is to be analyzed to understand the changes that have led to its creation. Therefore it is necessary to track the changes of modern Dhaka City which mainly includes the changes in the physical form of the city.

### **1.2 Statement of the Problem**

Dhaka is enriched with rich culture, history and heritage of about 400 years [2, 3]. Dhaka was worldwide famous for its fine Muslin (a kind of very smooth cloth), mosques and trade [4, 5]. But the scenario has been changed over the centuries.

<sup>1</sup> The term 'Mega City' is frequently used as a synonym for words such as super-city, giant city, conurbation and megalopolis. Mega Cities are defined as the cities with more than 10 million inhabitants [6].

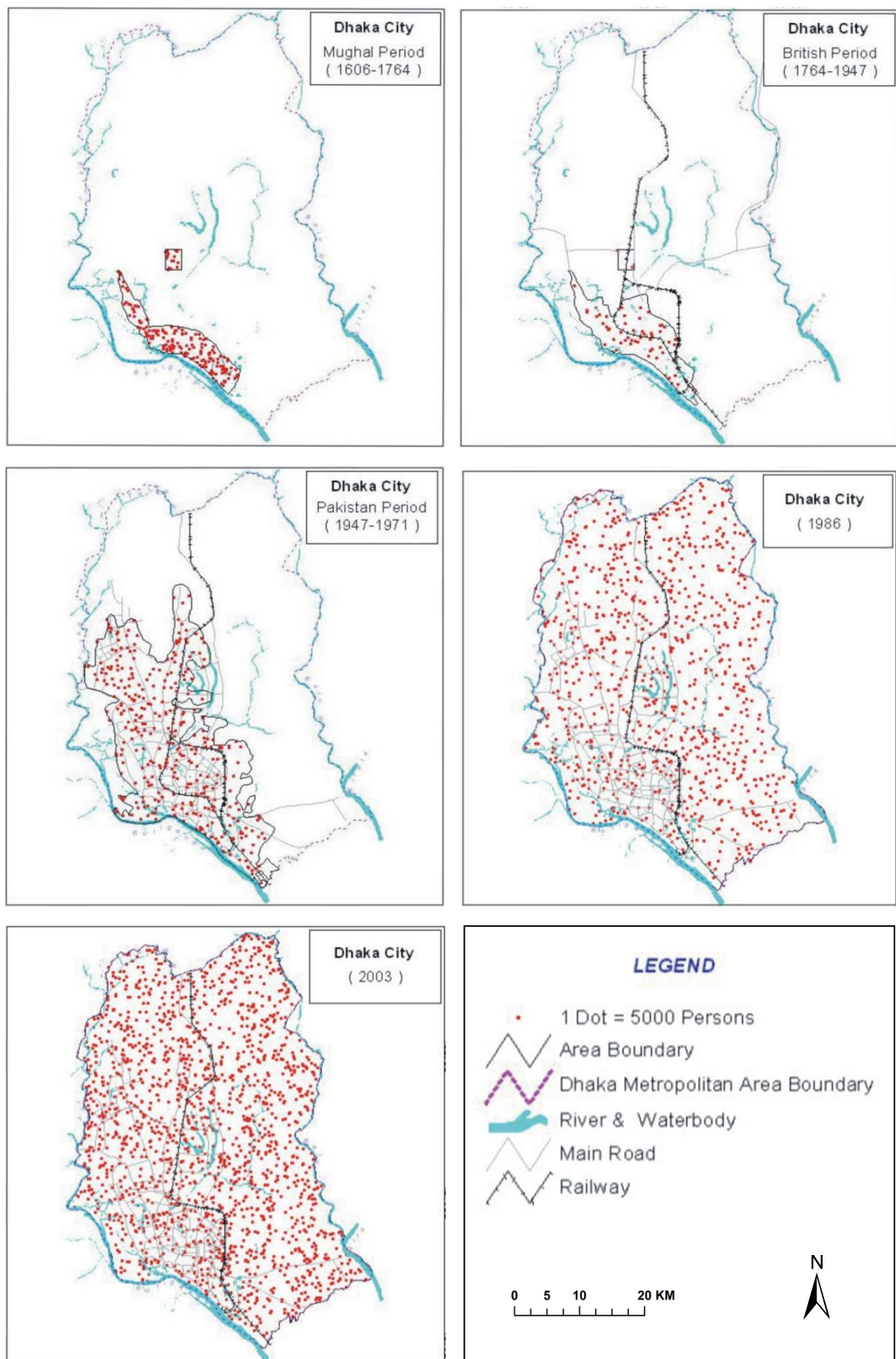
Dhaka is now attracting a huge amount of rural-urban migrants from all over the country due to well-paid job opportunities, better educational, health and other daily life facilities [7]. This kind of increasing and over population pressure is putting adverse impacts on Dhaka city like converting wetlands/natural vegetation/open space/bare soil to urban built-up areas [8, 9]. All these are creating numerous problems like unplanned urbanization, extensive urban poverty, water logging, growth of urban slums and squatters, traffic jam, environmental pollution and other socio-economic problems [7].

Dhaka has turned into one of the world's biggest megacities [1]. Dhaka is now one of the world's most populous, most densely populated and most urban agglomerated<sup>2</sup> cities [10, 11, and 12]. According to the World Bank Annual Report (2007), Dhaka is going to be world's third largest city by 2020. The overall scenario of physical growth of Dhaka City has been depicted in Table 1.1, Figure 1.1 and Table 1.2.

**Table 1.1: Growth of Dhaka City in Urban Agglomerations with 750,000 Inhabitants or More (1950-2025)**

Year	City Population (Thousands)	Urban Population Residing in Each Agglomeration (%)	Total Population Residing in Each Agglomeration (%)	Average Annual Rate of Change (%)
1950	336	18.0	0.8	3.94 (1950-1955)
1955	409	18.0	0.8	
1960	508	18.3	0.9	4.34 (1955-1960)
1965	821	21.7	1.3	9.60 (1960-1965)
1970	1374	26.2	2.0	10.30 (1965-1970)
1975	2221	28.6	2.8	9.61 (1970-1975)
1980	3266	24.3	3.6	7.71 (1975-1980)
1985	4660	25.9	4.5	7.11 (1980-1985)
1990	6621	28.9	5.7	7.02 (1985-1990)
1995	8332	30.0	6.5	4.60 (1990-1995)
2000	10285	31.0	7.3	4.21 (1995-2000)
2005	12555	31.9	8.2	3.99 (2000-2005)
2009	14251	31.9	8.8	3.08 (2005-2010)
2010	<b>14648</b>	31.7	<b>8.9</b>	2.53 (2010-2015)
2015	16623	30.8	9.5	2.38 (2015-2020)
2020	18721	29.8	10.1	2.24 (2020-2025)
2025	20936	28.7	10.7	

Source: Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, *World Population Prospects: The 2008 Revision and World Urbanization Prospects: The 2009 Revision*, <http://esa.un.org/wup2009/unup/>, retrieved on Wednesday, January 05, 2011.



**Figure 1.1: Changing Patterns of Dhaka City in Area and Population**

*Source: GIS Division, Bangladesh Centre for Advanced Studies, 2007*



<sup>2</sup> The term “Urban Agglomeration” refers to the population contained within the contours of a contiguous territory in-habited at urban density levels without regard to administrative boundaries. It usually incorporates the population in a city or town plus that in the suburban areas lying outside of but being adjacent to the city boundaries [13].

**Table 1.2: Historical Growth of Dhaka City in Terms of Area**

<b>Year</b>	<b>Area (sq.km)</b>
1600	1
1608	2
1700	40
1800	4.5
1867	10
1872	20
1881	20
1891	20
1901	20
1931	20
1941	25
1951	85
1961	125
1974	336
1981	510
1991	1353
2001	1530

*Source: Taylor, J., Sketch of the Topography and Statistics of Dacca (Calcutta: Military Orphan Press, 1840); Bangladesh Bureau of Statistics (BBS), Bangladesh National Population Census Report - 1974 (Dhaka: Ministry of Planning, 1977); Bangladesh Population Census 1991 Urban Area Report (Dhaka: Ministry of Planning, 1997); Population Census 2001 Preliminary Report (Dhaka: Ministry of Planning, 2001) and [10, 14, 15].*

If the trend of urbanization (in terms of population and area) of Dhaka is analyzed over time, then it is clear how spontaneously the city is growing (Table 1.1, 1.2 and Figure 1.1). The increase in population and area is degrading the living standard of Dhaka city.

Following all these adverse trends, Dhaka has clearly been announced as the second worst liveable city (following Harare, Zimbabwe) in the world according to a survey conducted by the Economist Intelligence Unit (EIU), affiliated with the UK-based weekly the Economist on February 2010 [16]. The EIU surveyed 140 cities of the world to create a global liveability report. For each city, the EIU considered over 30 qualitative and quantitative factors across five broad categories: stability; healthcare; culture and environment; education; and infrastructure [16].

All these are really a matter of special consideration. If this situation continues then Dhaka would soon become an urban slum with the least liveable situation for the city dwellers. Moreover, the land cover<sup>3</sup> pattern of Dhaka city is changing from vegetation (green) to builtup areas (brown) gradually (Figure 1.2). In this regard, it is much needed to track the land cover changes over-time and predict the future scenario of Dhaka city. This kind of research would be of great importance for the politicians, decision-makers and urban planners of Bangladesh to take early steps towards facing the worst situations about Dhaka city.

Some photographs depicting the existing urban life problems, due to this haphazard urban growth, of Dhaka city are presented in Appendix A.

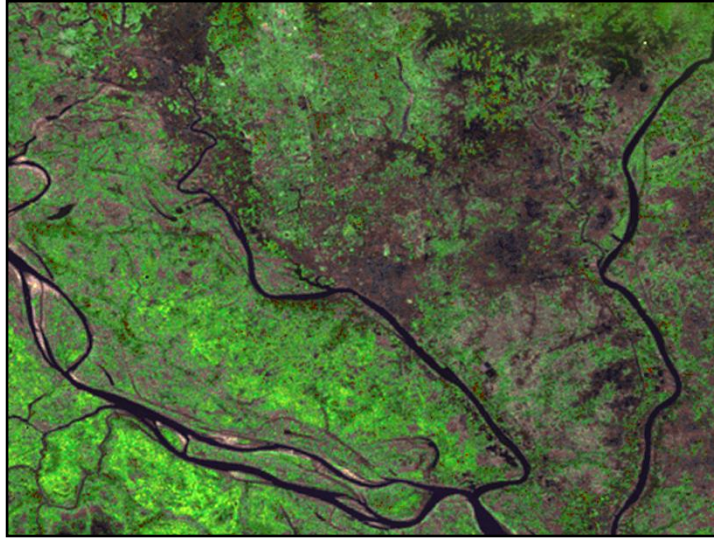
### **1.3 Study Area Profile**

There are 64 districts in Bangladesh. Dhaka City is located in Dhaka District that is surrounded by rivers. Dhaka is located in central Bangladesh at 23°43'0"N, 90°24'0"E, on the eastern banks of the Buriganga River [17]. Dhaka city area is under jurisdiction of different authorities that are known as Dhaka City Corporation (DCC), Dhaka Metropolitan Area (DMA), Dhaka Statistical Metropolitan Area (DSMA) and Dhaka Metropolitan Development Plan (DMDP) area. DCC comprises of 90 wards<sup>4</sup> (Figure 1.3).

The proposed study area for this research is Dhaka City Corporation (DCC) and its surrounding impact areas (Figure 1.4). The study area covers the oldest organic core part of Dhaka city (old Dhaka), the planned areas and even the unplanned new generation organic areas that are called 'Informal Settlements'. This selected study area almost covers the biggest urban agglomeration and is the central part of Bangladesh in terms of social and economic aspects [8, 10]. Therefore, this area has huge potentiality to face massive urban growth in near future based on the current trend of rapid urbanization.

<sup>3</sup> *'Land Cover' describes the physical state of the earth's surface and immediate subsurface in terms of the natural environment (such as vegetations, soils, surfaces and groundwater) and the man-made structures (e.g. buildings). While 'Land Use' relates to the manner in which the biophysical assets are used by humans. 'Land Use' is the human employment of a land-cover type [18, 19].*

<sup>4</sup> *'Ward' is the smallest electoral unit of urban areas in Bangladesh [10].*



**a.** Dhaka on 28-12-1972 taken by Landsat-1-MSS



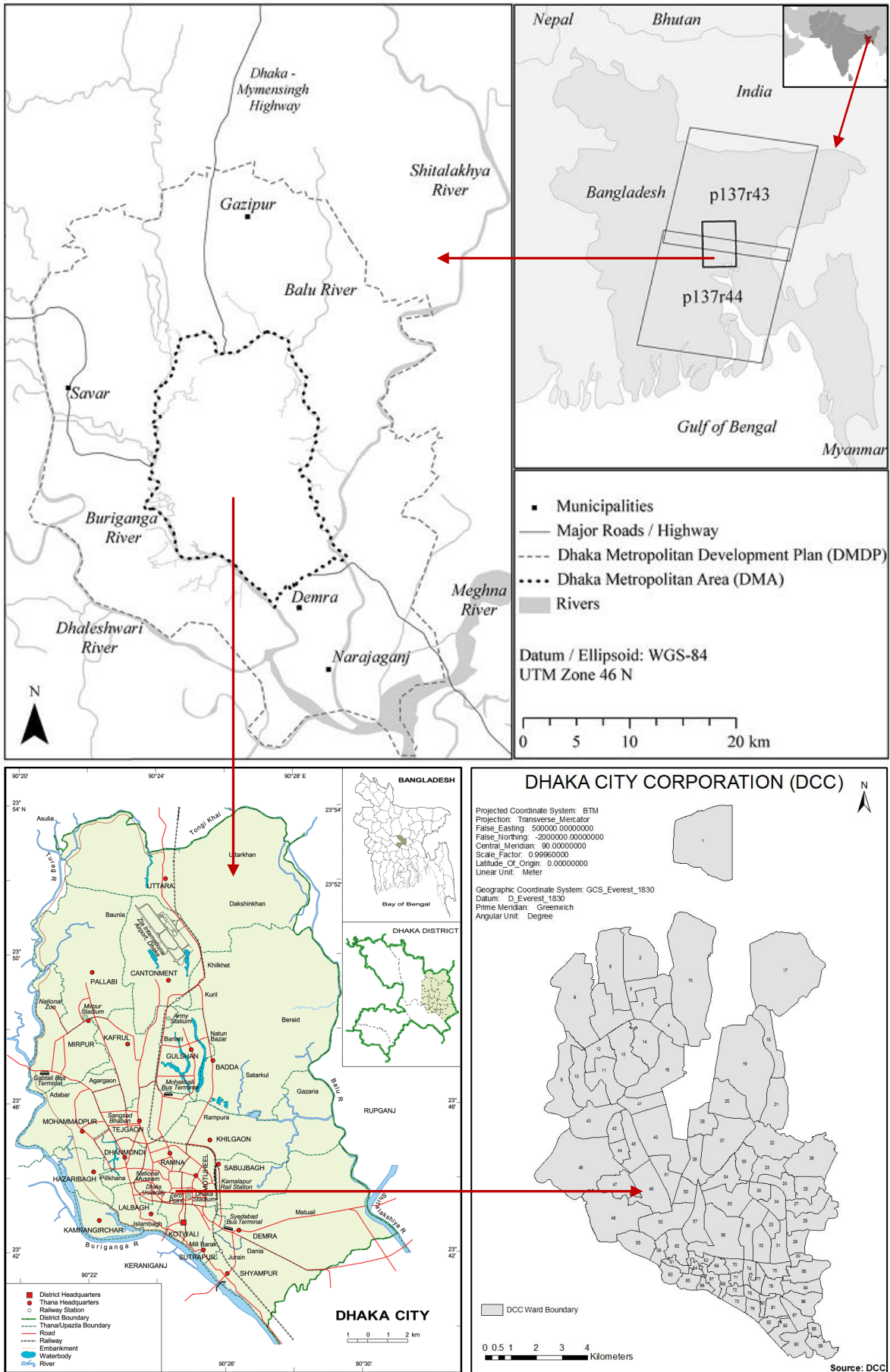
**b.** Dhaka on 13-02-1989 taken by Landsat-5-TM showing further growth since 1972



**c.** Dhaka on 29-01-2001 taken by Landsat-7-ETM+ showing further urban growth since 1989

**Figure 1.2: Satellite Images Showing Urban Growth of Dhaka (1972 to 2001)**

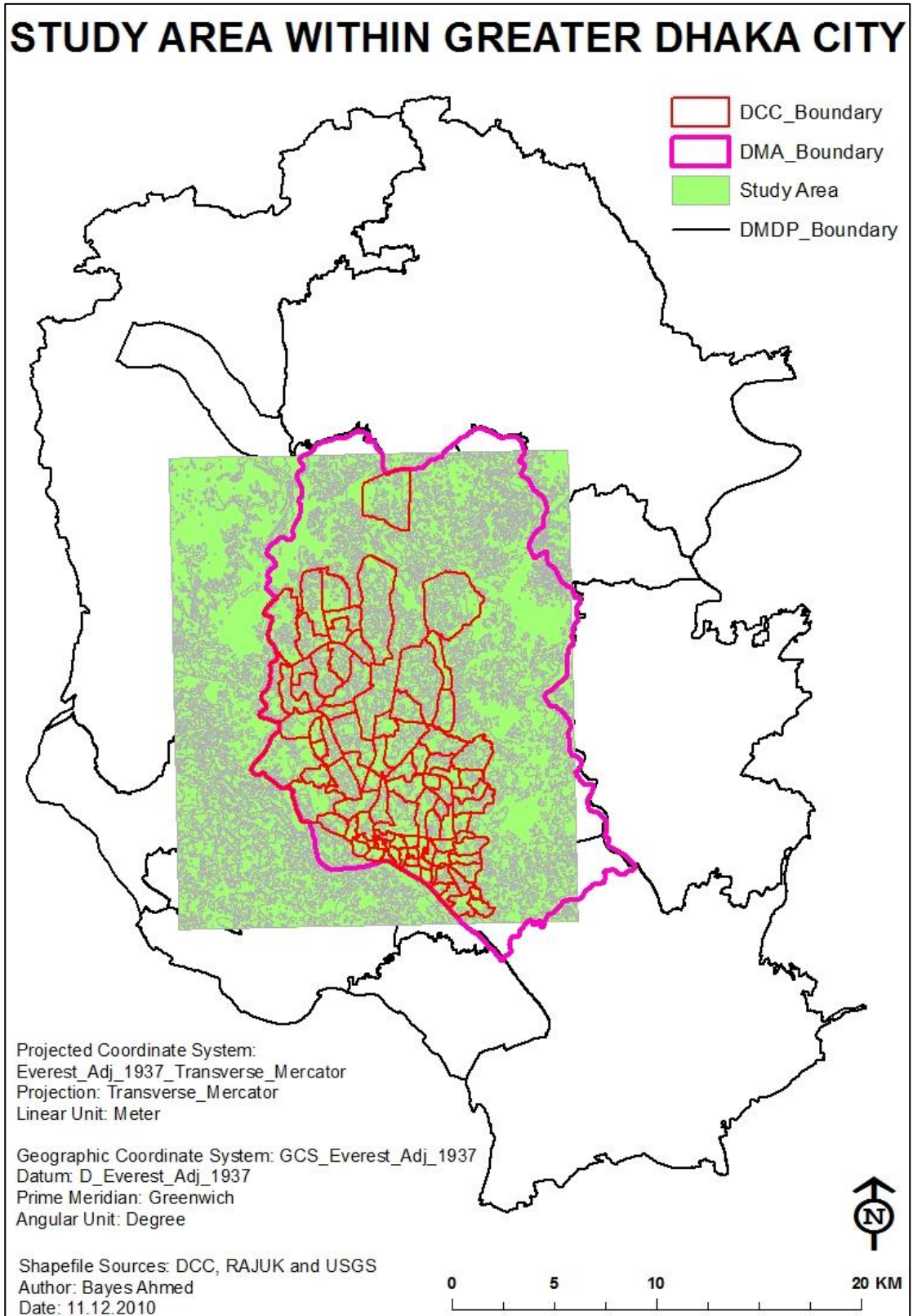
Source: NASA/Goddard Space Flight Center, Scientific Visualization Studio, 2001 (URL: <http://svs.gsfc.nasa.gov/goto?2323>)



**Figure 1.3: Location of Dhaka City**

Source: Top Map [20]; Bottom Left- Banglapedia, National Encyclopedia of Bangladesh, 2006; Bottom Right- DCC





**Figure 1.4: Location of the Study Area within Greater Dhaka City**

## **1.4 Objectives of the Research**

The general objective of this research is to map, detect, quantify, analyze and predict the land cover changes of Dhaka City over time. For this reason, various existing modelling techniques have been implemented with the help of some multi-spectral and multi-temporal remotely sensed and other geospatial datasets. The primary objective is to forecast the future urban land cover change of the selected study area within Dhaka city.

To address the above mentioned problems (section 1.2), this research has been conducted to achieve the following two broad objectives:

1. To quantify and investigate the characteristics of urban land cover changes (1989-2009) within the study area using satellite images.
2. To predict and analyze the future urban growth of Dhaka City.

## **1.5 Research Hypotheses**

The following research questions or hypotheses have been considered to fulfil the above mentioned research objectives:

### **1.5.1 Related to Objective 1**

1. Which datasets are available to conduct the whole research?
2. Which techniques are available for analyzing the change detection of land cover types of Dhaka city?
3. What is the general trend of land cover changes of the study area over time (1989-2009)?

### **1.5.2 Related to Objective 2**

1. Which methods are available to simulate the future urban growth of the study area?
2. Which method is suitable for forecasting the future land cover changes?
3. Are the satellite images, GIS and Remote Sensing tools and different methods/ techniques used for this research adequate or useful?
4. What will be the future scenario of Dhaka city based on the analysis and simulation of past trends?

## **1.6 Limitations of the Research**

### **1.6.1 Collection of Satellite Images**

To perform this type of Spatio-temporal analysis, it is important to select the satellite images of the same time interval. Again the spatial resolution of the images is important. For this research purpose, Landsat satellite images have been chosen that are only commercially available but can be found in free public-domain. Another reason for choosing these images is that the time interval is found equal 10 years of interval (1989, 1999, and 2009). The main problem of working with Landsat images is low resolution. The spatial resolution of Landsat Image is 30 meter [21]. IKONOS, QuickBird or other satellite images with higher resolution can be better option, but those images are commercial. Therefore due to limitations of resources, only free public-domain data have been used for this research.

### **1.6.2 Seasonal Variation**

Another important point, while selecting satellite images, is seasonal variation. Seasonal variation is an important aspect for tropical countries like Bangladesh. The change in vegetation, wet land, low land and water body land cover types are evident due to different seasons. Therefore, in an ideal situation, satellite images of the same season are selected for this kind of research. But there exist some sorts of seasonal variation for Landsat satellite images collected for this research. The images collected for 1999 (November) and 2009 (October) are from the same winter season. But the image of 1989 (February) is from another season, summer. This kind of variation creates problems while preparing base maps for analysis.

### **1.6.3 Collection of Reference Data**

The next limitation regarding this research is the collection of reference data or maps. The reference data are necessary for ground truthing purpose of the base maps (1989, 1999 and 2009) that have been prepared from the Landsat satellite images. But reference maps of the respective years (1989, 1999 and 2009) are not available. Therefore the base maps of Dhaka city of the years 1987, 1995 and 2001, collected from Survey of Bangladesh (SoB), have been used for referencing purpose. Google Earth images (2010) are used as reference data for ground truthing the base map of 2009.

## Chapter 2

### Theoretical Framework and Methodology

---

The detail descriptions of the methodology of the research along with the literature reviews have been stated in this chapter.

#### **2.1 Basic Terminologies**

Two basic terms have been used throughout this research. These are Remote Sensing (RS) and Geographic Information System (GIS). RS and GIS mean different things to different users. To some, it is a tool that allows the generation of custom maps. To some, it is kind of software that helps in analyzing different aspects of geographic data. To some, it helps in finding out oneself or knowing the current position of someone with the help of a detailed digital map. Though it is a hectic task to define RS and GIS in an easy manner, the general definitions of these two terms, along with the definition of Land, have been described in this section.

##### **2.1.1 Remote Sensing (RS)**

According to Lambin (2004), the definition of Remote Sensing (RS) is as follows: *“Remote sensing is the science and art of obtaining information about the Earth's surface through the analysis of data acquired by a device which is at a distance from the surface”* [22]. The broad definition of RS may include fields like vision, astronomy, medical imaging, Earth observation and so on [23]. But for this research purpose, RS only refers to the Earth Observation. In this kind of RS analysis, aerial photographs and images from Earth observing satellites are most commonly used [22].

##### **2.1.2 Geographic Information System (GIS)**

Geographic information means information about the earth. More specifically it refers to earth's surface information [24]. According to Michael F. Goodchild (2001): *“Geographic Information Systems (GISs) are defined as software systems and their relationships to other activities connected with geographic information”* [25].

Chrisman (1999) has defined GIS as: *“Organized activity by which people measure and represent geographic phenomena then transforms these representations into other forms while interacting with social structures”* [26].



According to Maguire (1991), some people believe GIS as the system of hardware and software while others believe GIS as applications or information processing [27]. In general, GIS support any operation on geographic information: acquisition, editing, manipulation, analysis, modelling, visualization, publication, and storage [28].

GIS has four basic subsystems: input, storage, analysis and output. All these things act like a perfect system. Finally these entities work in a body to get the final result.

RS and GIS techniques are comprehensively used for analytical purposes in numerous fields' like-urban planning, geography, earth science, transportation planning, environmental planning and disaster management and so on.

### **2.1.3 Land**

The Food and Agriculture Organization (FAO) defines land as:

*“Land is a delineable area of the earth's terrestrial surface, encompassing all attributes of the biosphere immediately above or below this surface including those of the near-surface climate the soil and terrain forms, the surface hydrology (including shallow lakes, rivers, marshes, and swamps), the near-surface sedimentary layers and associated groundwater reserve, the plant and animal populations, the human settlement pattern and physical results of past and present human activity (terracing, water storage or drainage structures, roads, buildings, etc.)”* [29].

## **2.2 Literature Review**

RS and GIS techniques are being widely used to assess natural resources and monitor environmental changes. It is possible to analyze land use/ land cover change dynamics using time series of remotely sensed data and linking it with socio-economic or bio-physical data using GIS. The incorporation of GIS and RS can help analyzing this kind of research in variety of ways like land cover mapping, detecting and monitoring land cover change over time, identifying land use attributes and land cover change hot spots etc [22].

With the advancement of technology, reduction in data cost, availability of historic spatio-temporal data and high resolution satellite images, GIS and RS techniques are now very useful for conducting researches like land cover change detection analysis and predicting the future scenario [30].

Urban area is a complex dynamic system. The growth of a city depends on numerous driving factors like social, economic, demographic, environmental, geographic, cultural and other phenomenon. Therefore, modelling this kind of highly dynamic urban areas is not an easy task. A significant number of urban growth and land use change models have been developed based on different theories. Examples of some popular models are ‘the von Thünen Model’ by Johann Heinrich von Thünen, ‘Concentric Zone Theory’ by E.W. Burgess, ‘Central Place Theory’ by Walter Christaller, ‘Sector Theory’ by Homer Hoyt, ‘Multiple Nuclei Theory’ by C.D. Harris and E.L. Ullman [1] etc.

Many researchers have conducted number of researches to detect the land use/ land cover change pattern over time and predict the future growth of urban areas. They have introduced and applied different new or existing techniques and methods to achieve the research objectives. Some examples are as follows:

### **2.2.1 Examples Related to Land Cover Change Detection**

1. Basak (2006) has classified some Landsat images of Dhaka Metropolitan Development Planning (DMDP) area using index-based expert classification process. The main objective is to identify the Spatio-temporal trends and dimension of urban form in DMDP area from 1989 to 2003 [8].
2. Griffiths *et al* (2010) have approached to map the urban growth of Dhaka megacity region (1990 to 2006) using multi-sensoral data. They have used a Support Vector Machine (SVM) classifier and post-classification comparison to reveal Spatio-temporal patterns of urban land-use and land-cover changes [20].
3. Dewan and Yamaguchi (2009) have tried to evaluate land cover changes and urban expansion in greater Dhaka, between 1975 and 2003 using satellite images and socio-economic data. A supervised classification algorithm and the post-classification change detection technique in GIS have been implemented by them. They have found the accuracy of the Landsat-derived land cover maps ranged from 85% to 90% [7].
4. Emch and Peterson (2006) have quantified mangrove forest cover change in the Sundarbans of south-west Bangladesh from 1989 to 2000 using Landsat Thematic Mapper (TM) satellite imagery. They have used three image processing techniques: Normalized Differential Vegetation Index (NDVI), maximum likelihood classification and sub-pixel classification [31].

### **2.2.2 Examples Related to Future Land Cover Prediction**

1. Kashem (2008) has implemented SLEUTH urban growth model to simulate the historical growth pattern of Dhaka Metropolitan Area. SLEUTH model incorporates Slope, Landuse, Exclusion layer (where growth cannot occur), Urban, Transportation and Hill-shade data layers. SLEUTH uses a modified Cellular Automata (CA) to model the spread of urbanization [1].
2. Lahti (2008) has predicted the urban growth of Sydney, Australia till 2106. The modelling has been performed using the CA model Metronamica, developed by the Research Institute for Knowledge Systems (RIKS) in the Netherlands [32].
3. Li and Yeh (2002) have introduced a new method, integrating Artificial Neural Networks (ANN) and CA (ANN\_CA model), for simulating the land-use map of 2005. They implemented the proposed model in a city called Dongguan in southern China using 1988 and 1993 TM satellite images [33].
4. Cabral and Zamyatin (2006) have implemented three land change models to forecast the urban dynamics in Sintra-Cascais municipalities of Portugal, for the year 2025. The models are CA Markov chain model (CA\_Markov), CA\_Advanced and Geomod. They have used image segmentation and texturing procedures to classify the Landsat images of 1989, 1994 and 2001 [34].

### **2.3 Methodology of the Research**

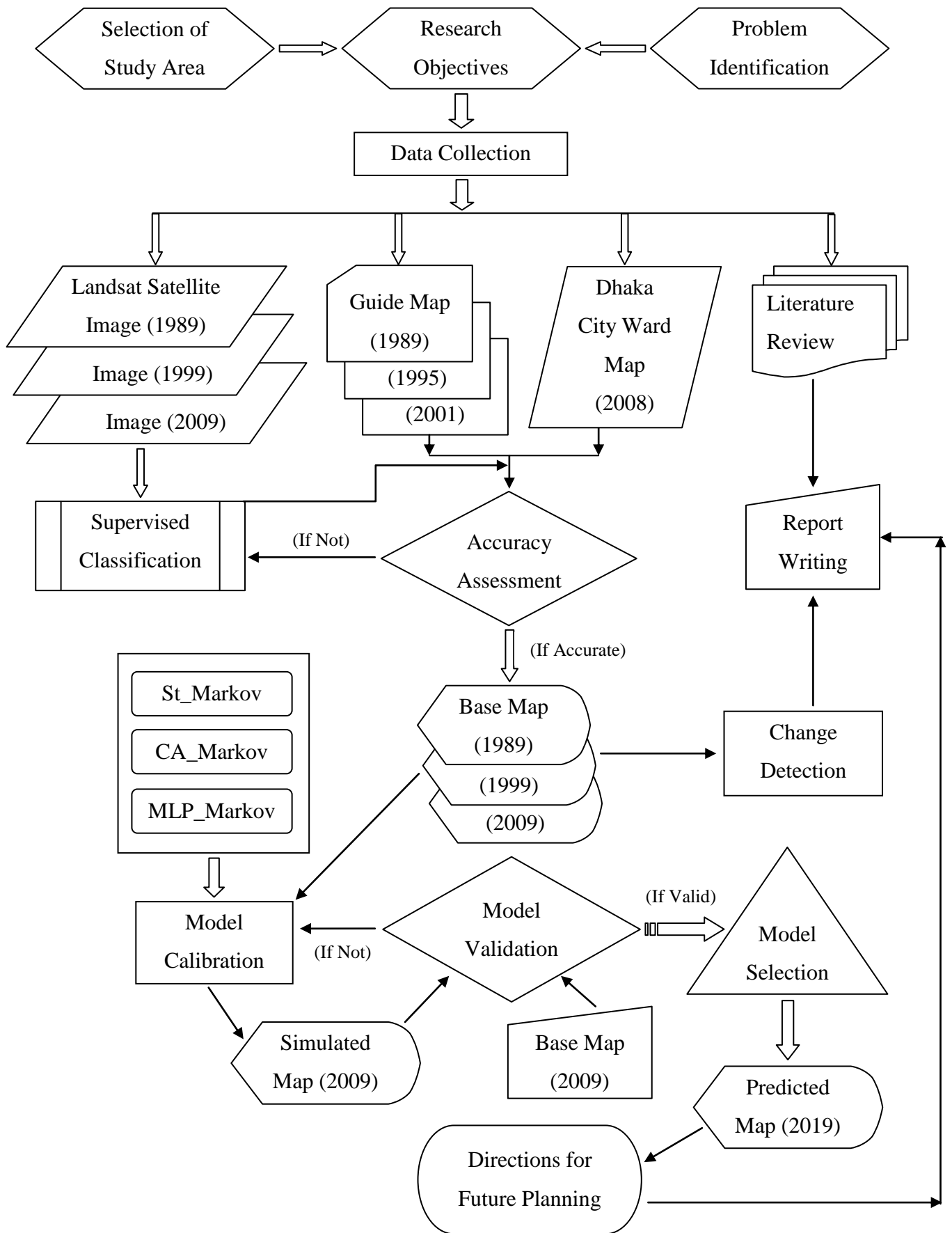
The steps followed to achieve the objectives and carry out the entire research have been described in this section. A flow chart depicting the main steps relating to carry out the research is also attached in Figure 2.1.

#### **2.3.1 Selection of Study Area**

Dhaka City Corporation (DCC) and its surrounding impact area have been selected as the study area for this research. A detail on this is described in Section 1.3 of Chapter 01.

#### **2.3.2 Problem Identification and Research Objectives**

To carry out this research in an effective way and to reach the goal, two broad objectives have been formulated. The basis for selecting the research objectives are based on the problems identified in Section 1.2 of Chapter 1.



**Figure 2.1: Flow Chart of Methodology**

### **2.3.3 Data Collection**

#### **2.3.3.1 Satellite Images**

This research is dependent on secondary data. To prepare the base maps for analysis purpose and applying the different methods to achieve the research objectives, Landsat satellite images (1989, 1999 and 2009) have been collected from the official website of U.S. Geological Survey (USGS). Table 2.1 shows the details of the Landsat satellite images used for analysis.

**Table 2.1: Details of Landsat Satellite Images**

<b>Respective Year</b>	<b>Date Acquired (Day/Month/Year)</b>	<b>Sensor</b>	<b>Quality (100% Cloud Free)</b>
<b>1989</b>	13/02/1989	Landsat 4-5 Thematic Mapper (TM)	7
<b>1999</b>	24/11/1999	Landsat 7 Enhanced Thematic Mapper Plus (ETM+)	9
<b>2009</b>	26/10/2009	Landsat 4-5 Thematic Mapper (TM)	9

*Source: U.S. Geological Survey, 2010 [21]*

Landsat Path 137 Row 44 covers the whole study area. Map Projection of the collected satellite images is Universal Transverse Mercator (UTM) within Zone 46 N– Datum World Geodetic System (WGS) 84 and the pixel size is 30 meters [21].

Figure 2.2 illustrates the location of DCC on the Landsat satellite images for different years. The surroundings of DCC have also been included within the study area for predicting the future land cover changes.

The Band Combination used, for the base Landsat satellite images (Figure 2.2), is 432 Red-Green-Blue (RGB). Map Projection used for DCC Boundary is Bangladesh Transverse Mercator (BTM) and datum is D\_Everest\_1830.



**Figure 2.2: Dhaka City Corporation and its Surroundings**

*Map Prepared by the Researcher; Source: USGS, 2010 and DCC, 2008*

### **2.3.3.2 Reference Data**

For the purpose of ground-truthing/ referencing, several base maps of Dhaka City (for the year of 1987, 1995 and 2001) have been collected from the Survey of Bangladesh (SoB). Again, for comparing the images some other reference satellite images (IRS image of 1996 and Landsat satellite image of 2003) have been collected from the Department of Urban and Regional Planning, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. The land use maps of all the wards of DCC (2008) have been collected from DCC and *Rajdhani Unnayan Kartripakkha* (RAJUK) or Capital City Planning and Development Authority. Google Earth is another option to get some ideas about the recent land cover pattern of Dhaka city. These reference data have been used for training site selection and preparing land cover maps.

The collected base maps are attached in Appendix B.

### **2.3.3.3 Literature Review**

To understand the concept and methods of urban land cover change pattern analysis and prediction, and collect some necessary information; many papers and documents have been collected. In this regard various journal papers, reports, conference papers and dissertations have been overviewed.

### **2.3.4 Base Map Preparation and Accuracy Assessment**

For image classification purpose, a fisher supervised classification method has been used. Then after achieving satisfactory accuracy results, the base maps have been finalized. Details of the procedures of base map preparation and accuracy assessment have been described in Chapter 3.

### **2.3.5 Change Detection Analysis**

The change detection of urban land cover types have been analyzed using different available techniques. Details can be found in Chapter 4.

### **2.3.6 Model Calibration/ Simulation**

For simulating the land cover map of Dhaka, three different methods have been implemented in this research. The detailed descriptions of the methods and the process of implementation have been discussed from Chapter 5 to Chapter 7.

### **2.3.7 Model Validation and Selection**

At first the classified base land cover images of Dhaka city of 1989 and 1999 have been used to predict the land cover map of 2009 using the selected three methods. Then map comparison techniques have been applied between the predicted and classified land cover map of 2009. This is known as model validation. Out of the three methods implemented, the most suitable model for this particular research has been chosen based on the kappa statistics. Model validation and selection techniques and the results are described in detail in Chapter 8.

### **2.3.8 Future Prediction**

The land cover map of 2019 of Dhaka city has been predicted using the method obtained from model selection process. The simulation process and the analysis of the predicted map (2019) have been elaborated in Chapter 09.

### **2.3.9 Directions for Future Planning**

A few recommendations have been devised for future planning of Dhaka city, based on the results found from this research. This can be found in Chapter 10.

### **2.3.10 Report Writing**

After getting all relevant data and information related to the concerned aspects, data analysis and the subsequent interpretations; all these are presented in this thesis report.

## **2.4 Tools Used for this Research**

Different software programs have been used for conducting this research. These are:

Image Classification and Model Simulation: ‘IDRISI 16: The Taiga Edition’.

Accuracy Assessment: ‘ERDAS IMAGINE 9.1’.

Change Detection Analysis: ‘Fragstats 3.3’.

Model Validation: ‘The Map Comparison Kit: version 3.2.0’.

Map Projections and Preparation: ‘ArcGIS 9.3.1 Desktop’.

Other Purposes: CorelDRAW Graphics Suite X5, Microsoft Word and Excel 2007.



## Chapter 3

### Base Map Preparation and Accuracy Assessment

---

This chapter deals with explaining the procedure of preparing the base maps for analysis and the accuracy assessment of those maps.

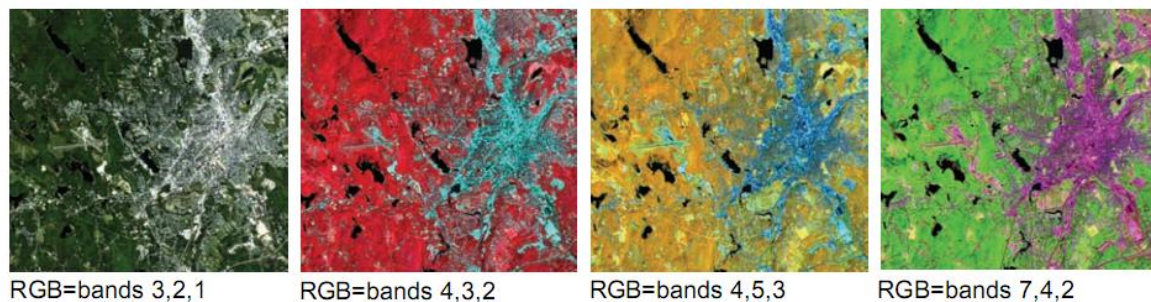
The collected Landsat satellite images (1989, 1999 and 2009) have been used for preparing the base maps for land cover change detection and future prediction. The basic steps for preparing the base maps are as follows:

#### **3.1 Image Enhancement**

Image enhancement is a kind of image modification that enables the capabilities of human vision to identify and select regions of interests [35]. Composite generation technique has been performed for this particular research.

#### **3.2 Composite Generation**

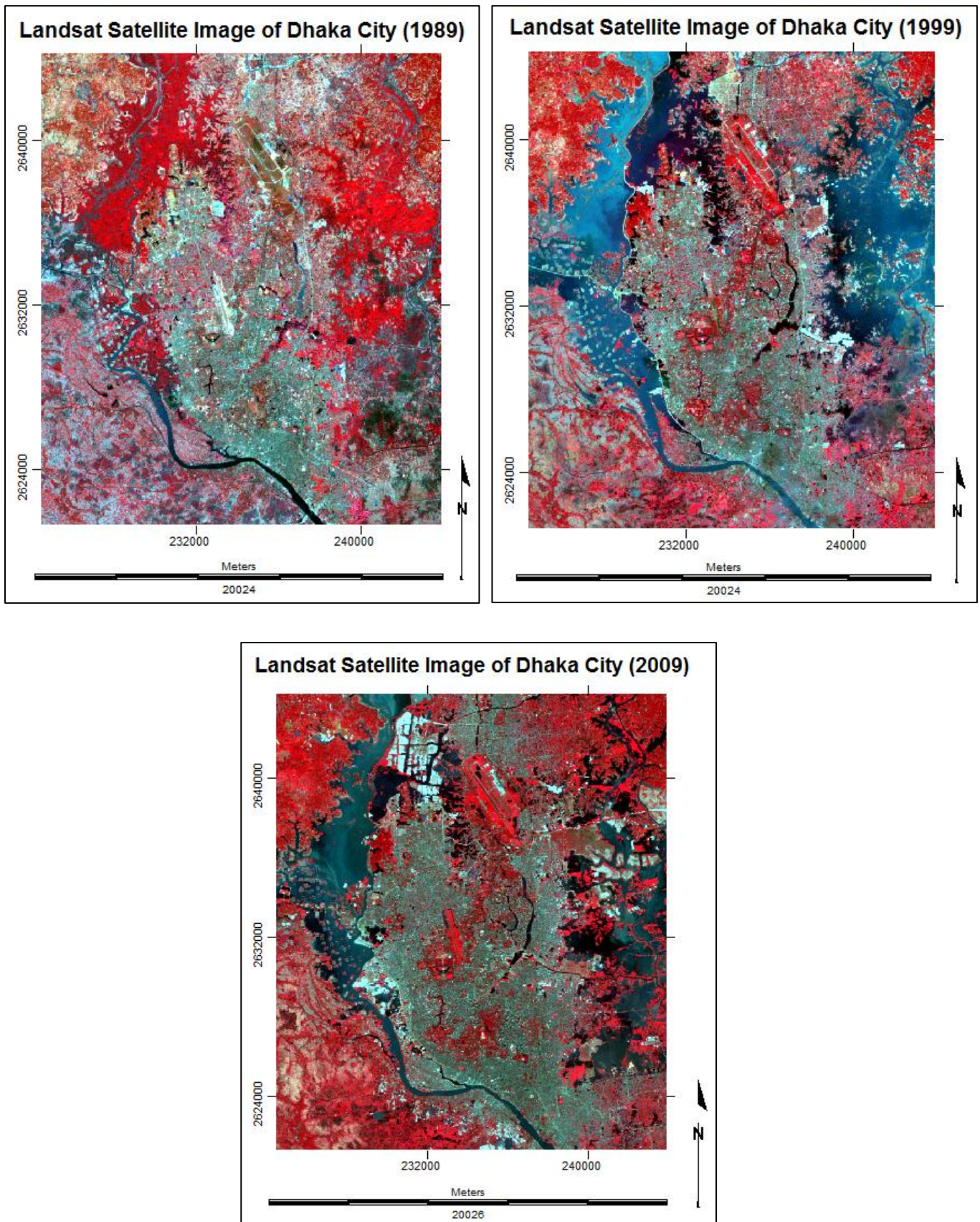
Landsat TM records 7 spectral bands. For visual purpose any 3 bands are combined that are acting a False Color Composite (FCC). Figure 3.1 shows several composites using different band combinations from the same TM images [36].



**Figure 3.1: Composites Using Different Band Combinations [36]**

Using the basic colours red, green and blue (RGB) it is possible to prepare different FCC images [37]. These FCC images are useful to distinguish between different cover types or ground objects like buildings, roads, and vegetation.

The FCC of RGB= bands 4, 3 and 2 has been chosen for this research. This combination normally makes urban areas appear blue, vegetation red, water bodies from dark blue to black, soils with no vegetation from white to brown [38]. Figure 3.2 shows the FCC (RGB= bands 4, 3 and 2) image of the study area for different time periods.



**Figure 3.2: False Color Composite (RGB=4, 3 and 2) Maps of the Study Area**

The study area is selected by choosing the same geographic position of all 7 bands for the same time period. This helps to maintain the same number of rows and column. The properties of each image are described in Table 3.1.

### **3.3 Image Classification**

Image classification refers to *grouping image pixels into categories or classes to produce a thematic representation* [39]. Image classification comprehends various operations that can be applied to photographic or image data. These include image restoration, image pre-processing, enhancement, compression, spatial filtering, and pattern recognition and so on [39]. There are two basic methods of image classification: supervised and unsupervised [37]. Supervised classification relies on the priori knowledge of the study area [39]. Therefore, for this research, a supervised classification method has been used.

Supervised classification can be defined as: *“A procedure for identifying spectrally similar areas on an image by identifying ‘training’ sites of known targets and then extrapolating those spectral signatures to other areas of unknown targets”* [39].

In case of supervised classification, the user develops statistical description for various known land cover types that is called signature development. Then a procedure is used to identify the similar pixels/signature for different land cover types for the whole image. The steps that are followed for this supervised classification are as follows:

#### **3.3.1 Training Site Development**

Training sites are the areas defined for each land cover type within the image. The chosen colour composite is used for digitizing polygons around each training site for similar land cover. Then a unique identifier is assigned to each known land cover type [36]. More than one training site for each type has been identified for making the land cover images. Five land cover types have been identified for this research (Table 3.2). The training sites developed for this research are based on the reference data and ancillary information collected from various sources as mentioned earlier [36]. For developing training sites properly, all the satellite images have been analyzed with respect to spectral and spatial profiles. This is performed to make sure that the digital numbers (DNs) of different land cover types are acceptable prior to final classification [36].

**Table 3.1: Properties of the Processed Raster Images for Analysis**

<b>Image Characteristic</b>	<b>Description</b>
File Format	Raster
File Type	Binary
Data Type	Byte
Columns	650
Rows	762
Reference System	UTM-46 North
Reference Unit	Meter
Datum	WGS 84
Minimum X	224445
Maximum X	243945
Minimum Y	2621445
Maximum Y	2644305
X Resolution	30
Y Resolution	30

**Table 3.2: Details of the Land Cover Types**

<b>Land Cover Type</b>	<b>Description</b>
<b>Builtup Area</b>	All residential, commercial and industrial areas, villages, settlements and transportation infrastructure.
<b>Water Body</b>	River, permanent open water, lakes, ponds, canals and reservoirs.
<b>Vegetation</b>	Trees, shrub lands and semi natural vegetation: deciduous, coniferous, and mixed forest, palms, orchard, herbs, climbers, gardens, inner-city recreational areas, parks and playgrounds, grassland and vegetable lands.
<b>Low Land</b>	Permanent and seasonal wetlands, low-lying areas, marshy land, rills and gully, swamps, mudflats, all cultivated areas including urban agriculture; crop fields and rice-paddies.
<b>Fallow Land</b>	Fallow land, earth and sand land in-fillings, construction sites, developed land, excavation sites, solid waste landfills, open space, bare and exposed soils.



### **3.3.2 Signature Development**

This is the stage of creating the spectral signature for each type of land cover. This is done by analyzing the pixels of the training sites [36]. When the digitization of training sites is finished, the statistical characterizations of each land cover class are needed. These are called signatures.

Signatures are developed incorporating the vector files containing training sites and the bands used for analysis. These signature files contain statistical information about the reflectance values of the pixels of the training sites for each land cover type [36].

### **3.3.3 Classification**

After developing signature files for all land cover classes the next step is to classify the images based on these signature files. This can be done by two ways: hard or soft classifiers [37]. These are kind of statistical techniques to classify the whole image pixel by pixel based on each known type particular signature [36].

In case of hard classifications, each pixel is assigned in a way that has the most similar signature for a particular land cover type. On the other hand, soft classifications take into consideration the degree of membership of the pixel in all classes [37].

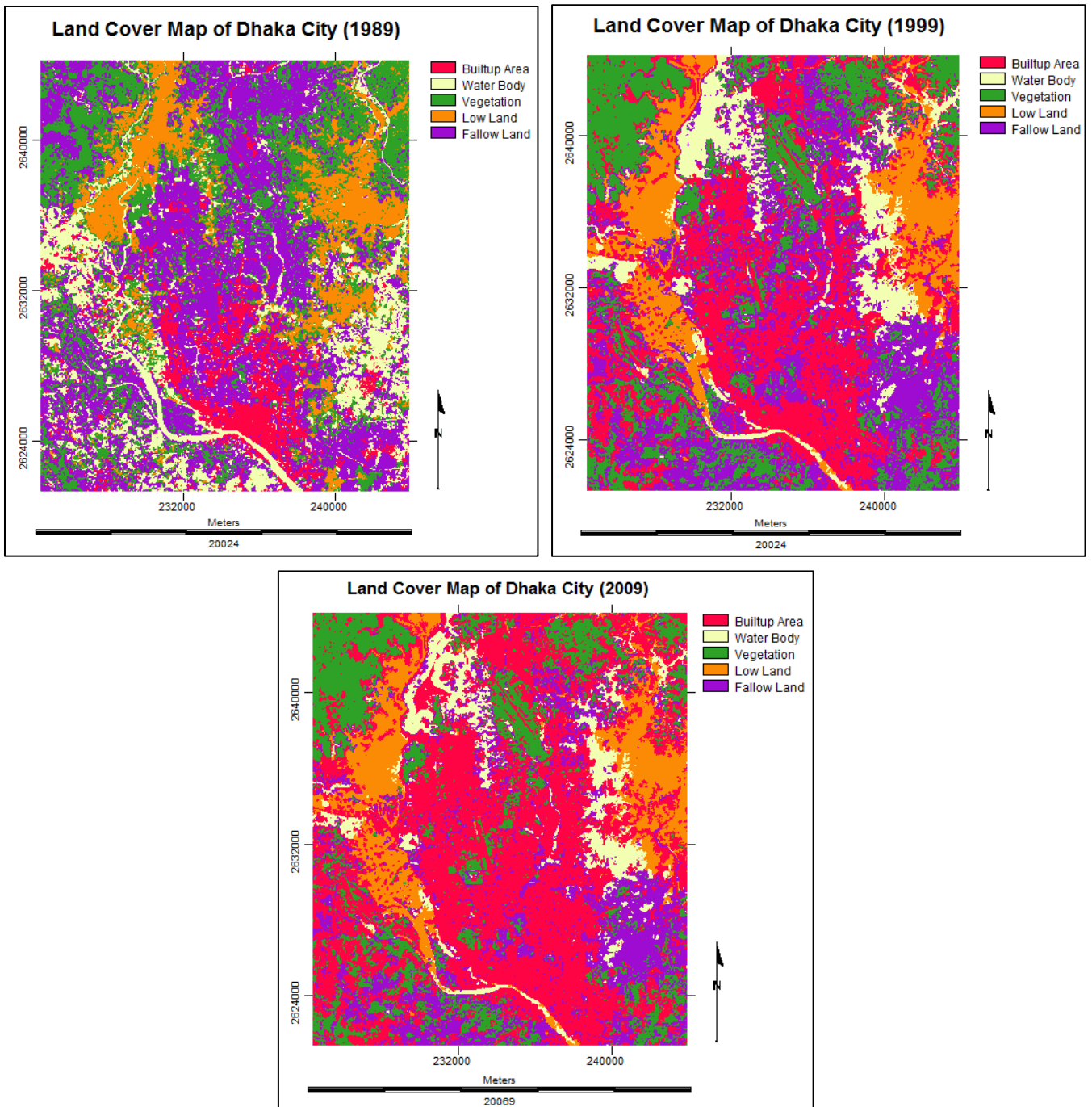
For this research, a hard classifier called 'Fisher Classifier' has been chosen. Fisher classifier uses the concept of the linear discrimination analysis [36]. Fisher Classifier performs well when there are very few areas of unknown classes and when the training sites are representative of their informational classes [37]. This is why fisher classifier is appropriate for this particular research, because most areas for the classes are known.

### **3.3.4 Generalization**

After image classification, sometimes many isolated pixels may be found [37]. These isolated pixels may belong to one or more classes that differ from surrounding pixels. Therefore it is necessary to generalize the image and remove the isolated pixels.

Filtering is the solution for this type of problem [36]. In RS analysis, filtering refers to *the removal of certain spectral or spatial frequencies to highlight features in the remaining image* [39]. Mode filters are good for filling gaps between polygons after a vector-to-raster conversion [37].

Therefore, a 3×3 mode filter has been applied to generalize the fisher classified land cover images. This post-processing operation replaces the isolated pixels to the most common neighbouring class. Finally the generalized image is reclassified to produce the final version of land cover maps for different years (Figure 3.3).



**Figure 3.3: Land Cover Maps of Dhaka City**

### **3.4 Accuracy Assessment**

The final stage of image classification process is accuracy assessment. Accuracy Assessment is a kind of process to compare the classification with ground truth or other data. It allows evaluating a classified image file [36].

It is not typical to ground truth each and every pixel of the classified image. Therefore some reference pixels are generated. Reference pixels are points on the classified image. Each point of reference pixels represents specific geographic coordinate of the image. These reference pixels are randomly selected [40].

The randomly selected points within the classified image list two sets of class. The first set of class values represents the actual land cover type. The second set of class values are known as reference values. These reference values are input by the researcher that is based on ground truth data [41].

The ground truthing of reference values are possible through field visit, comparing base maps, aerial photos, previously tested maps or other data [41].

The number of reference pixels is an important issue for accuracy assessment. It has been proved that more than 250 reference pixels ( $\pm 5\%$ ) are needed to estimate the mean accuracy of classification [40].

Therefore 250 reference pixels have been generated for each classification image for this research to perform accuracy assessment.

There is a common tendency while selecting the reference pixels by the analysts. The researchers often select the same pixels as reference pixels for testing the classification that had already been used for training samples [40]. This creates kind of biasness.

This is why the reference pixels are normally selected randomly to eliminate this kind of impartiality [40].

Therefore for this research, the reference pixels have been selected using stratified random distribution process.

In stratified random distribution process the number of points is stratified to the distribution of the thematic layer classes within the reference pixels [41].

### 3.4.1 Assessment Procedure

The collected base maps (Appendix B) have been used to find the land cover types of the reference points. Figure 3.4 and Table 3.3 (not real, just an example to show how the mechanism works within the overall procedure) show how the land cover types for the randomly selected points are chosen.

Here (Figure 3.4) the representation is like; 1= Builtup Area, 2= Water Body, 3= Vegetation, 4= Low Land and 5= Fallow Land.

<b><u>Cell 1</u></b> (5)	<b><u>Cell 2</u></b> (2)	<b><u>Cell 3</u></b> (4)	<b><u>Cell 4</u></b> (1)
5 5 5	1 2 2	5 5 4	1 1 3
5 <b>5</b> 5	2 <b>2</b> 1	4 <b>4</b> 4	3 <b>1</b> 3
5 5 5	2 2 2	4 5 4	1 1 1
<b><u>Cell 5</u></b> (1)	<b><u>Cell 6</u></b> (5)	<b><u>Cell 7</u></b> (3)	<b><u>Cell 8</u></b> (1)
4 4 4	2 5 5	1 1 3	1 1 1
1 <b>1</b> 1	2 <b>5</b> 5	3 <b>3</b> 3	5 <b>1</b> 1
1 1 1	2 5 5	4 3 3	5 5 1

**Figure 3.4: Random Points Cell Listing**

Figure 3.4 shows the process of selecting the stratified random points for different land cover types within a specific cell.

To select the random reference pixels, a 3×3 window has been selected for cell listing. Then the mid-value of each cell is generated as random point (Figure 3.4). The numbers in parenthesis are the chosen land cover type values for each cell.

This is how 250 random points showing class values have been selected from 250 cells.



**Table 3.3: Accuracy Assessment Cell Array**

Point	Name	'X' Co-ordinate	'Y' Co-ordinate	Class	Reference	Result
1	ID#1	232230.00	2643810.00	5	5	Correct
2	ID#2	230220.00	2628480.00	<b>2</b>	<b>1</b>	<b>Incorrect</b>
3	ID#3	232560.00	2630907.00	4	4	Correct
4	ID#4	225180.00	2868000.00	1	1	Correct
5	ID#5	225270.00	2634120.00	1	1	Correct
6	ID#6	239220.00	2643210.00	<b>5</b>	<b>4</b>	<b>Incorrect</b>
7	ID#7	243870.00	2637990.00	3	3	Correct
8	ID#8	238590.00	2633880.00	1	1	Correct

Table 3.3 illustrates the cell array for accuracy assessment. The right-most columns respectively represent the class values and reference values. The class values are generated from as per the rules mentioned in Figure 3.4. The reference values, cross-checked from base maps, are input from the researcher.

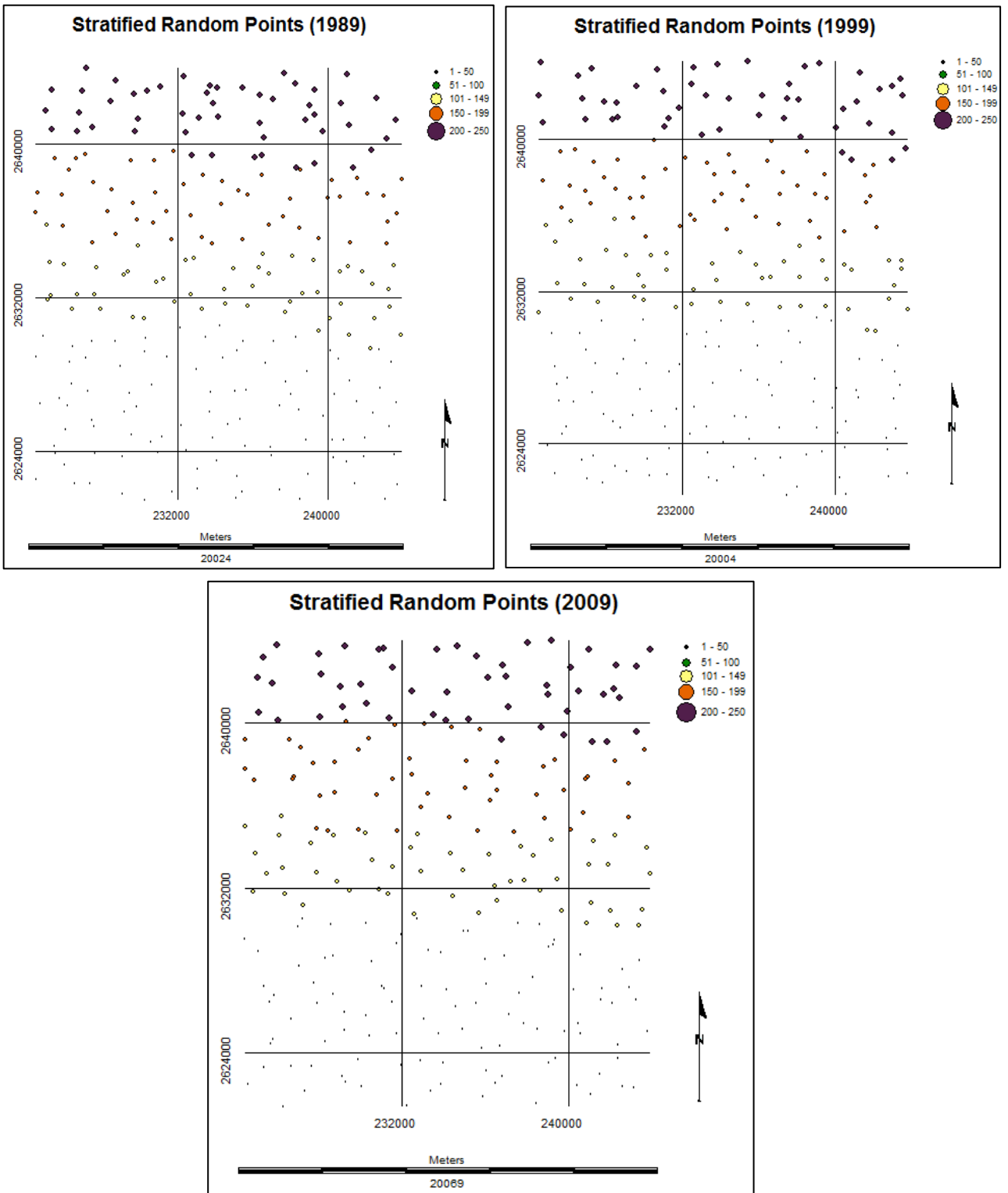
Moreover, this figure shows the point IDs' and the geographic coordinates (X and Y values) for each random point. At the end, 250 land cover class and reference values have been randomly selected for the base years of 1989, 1999 and 2009.

Figure 3.5 shows the randomly selected 250 points for each base year.

From the accuracy assessment cell array, three types of reports are generated [41]. These are:

- a) **Error Matrix:** it compares the reference points to the classified points in a  $c \times c$  matrix, where  $c$  is the number of land cover classes.
- b) **Accuracy Totals:** it calculates statistics of the percentages of accuracy that is based on the error matrix and
- c) **Kappa Coefficient.**

Details of the basic terms used for accuracy assessment have been explained in Appendix C. Moreover, the related tables of the accuracy reports for different time periods have been stated from Table C 4 to C 12 (Appendix C).



**Figure 3.5: Selected Stratified Random Points for Ground Truthing**

### **3.4.2 Results and Discussion**

The assessment of classification accuracy of the land cover maps (1989, 1999 and 2009) has been carried out using error matrices (Table C 4, C 7 and C 10).

User's accuracy measures the proportion of each land cover class which is correct. On the other hand, producer's accuracy measures the proportion of the land base which is correctly classified. Producer's and User's accuracy are also found consistently high ranging from 71%-100% (Table C 5, C 8 and C 11) for all the years. Therefore, higher values of producer's and user's accuracy indicate that the prepared base maps are quite good enough for further analysis.

Overall accuracy takes no account of source of error. Kappa coefficient is a measure of the proportional (or percentage) improvement by the classifier over a purely random assignment to classes. The overall accuracies for 1989, 1999 and 2009 are found 85.20%, 86.80% and 91.60% respectively, with Kappa statistics of 0.8054, 0.8294 and 0.8592 (Appendix C). The conditional kappa values for each category for different years are also found higher (Table C 6, C 9 and C 12).

It means few misclassifications have been observed in the classified land cover maps of Dhaka city. This may be because of the same spectral characteristics of some land cover types. For example, in case of 1989 base map, certain built-up areas were misclassified as fallow land. Again, in most cases it was really difficult to separate water bodies and low/cultivable lands categories. The reasons may be the seasonal variations of the satellite images for different years and the similar spectral properties of land covers in some cases.

Moreover, less image spectral resolution has directed to spectral mixing of different land cover types. This has caused spectral confusions among the cover types. It is also important to mention that the images of 1999 and 2009 represent winter season while the image of 1989 represents spring season. Therefore other seasonal images can be important evaluating the land cover change pattern of this kind of highly dynamic urban environment like Dhaka.

It is typical that most land cover classification images are 85% accurate [37]. Therefore, it can be stated that the classification accuracy achieved for this research is above satisfactory level.

## **Chapter 4**

### **Land Cover Change Detection Analysis**

---

The different techniques and spatial metrics that have been implemented for change detection analysis are described in this chapter.

#### **4.1 Change Detection**

Urban area is highly dynamic. Numerous factors put impact on the growth or changing pattern for a particular city. Therefore, understanding the urban dynamics is a complex task while planning for a planned and sustainable urban development [48]. To predict the future change of highly dynamic urban areas like Dhaka city, it is necessary to assess and monitor the urban land cover change on a regular basis. In remote sensing, ‘Change Detection’ is defined as the process of determining and monitoring the changes in the land cover types in different time periods. It provides the quantitative analysis of the spatial distribution in the area of interest [48]. Change detection is important because it helps the researcher to understand and monitor the land cover change pattern (e.g. urbanization, deforestation, agricultural land management) within the study area.

Change detection can be divided into two major groups: pre-classification and post-classification methods. A number of approaches like image differencing, image rationing, vegetation index differencing (NDVI), principal component analysis, artificial networks, fuzzy sets, object-oriented methods, image regression etc; have been applied in numerous studies to determine the spatial extent of land cover types [48]. Among these methods, post-classification is the most common and suitable method for detecting land cover change analysis [30]. The accuracy of change detection is dependent on the accuracy of each land cover type. Therefore, performing change detection analysis should be done after land cover classification [30]. Considering all these things post-classification method has method has been applied for this research.

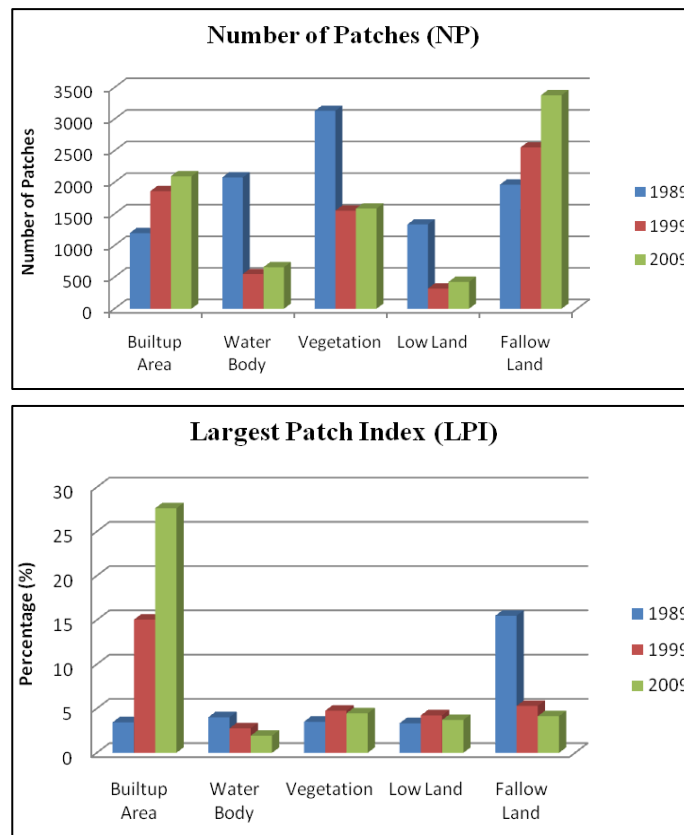
#### **4.2 Terminologies**

In this chapter, the change among different land cover types over the years has been described. To do this analysis, a number of spatial metrics have been used. The definitions and mathematical expressions of the spatial metrics used in this research are described in Appendix D. It helps to understand the inner meaning and the applicability.

### **4.3 Analysis and Interpretation of the Change Detection Techniques**

Class-level metrics are integrated over all the patches of a given type (class) [50]. For this research purpose, class-level metrics have been used. There are five broad categories of land cover types in this study. Therefore, the spatial metrics values for all these types have been generated for all the base years (1989, 1999 and 2009). The raw values in tabular format are attached in Appendix E. In this section, the interpretation of the generated values of selected spatial metrics and other techniques has been discussed.

#### **4.3.1 Number of Patches and Largest Patch Index**



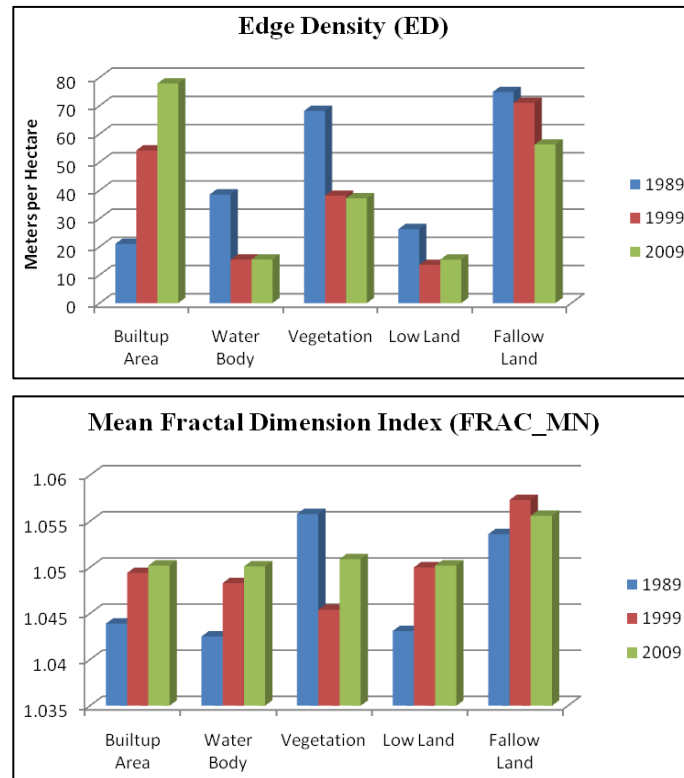
**Figure 4.1: Number of Patches and Largest Patch Index**

The numbers of patches, urban blocks in this case, of builtup area and fallow land types have increased over time. Again the decrease of water body and vegetation is prominent (Figure 4.1). This reveals the gradual development of urban infrastructures converting water bodies, vegetation and low lands.

LPI of builtup area type is increasing in a high rate and fallow land type is decreasing (Figure 4.1). This indicates the rapid growth of urbanization within the study area over time (1989-2009).

### **4.3.2 Edge Density and Mean Fractal Dimension Index**

Edge density (ED) of builtup area is increasing, while ED of fallow land, vegetation and water body cover types are decreasing (Figure 4.3). The total length of the edge of the land cover patches (urban patch) increases with an increase in the land use fragmentation and development of continuous urban features [48]. Therefore, this gradual increment in builtup area proves an increase in the total length of the edge of the urban patches.



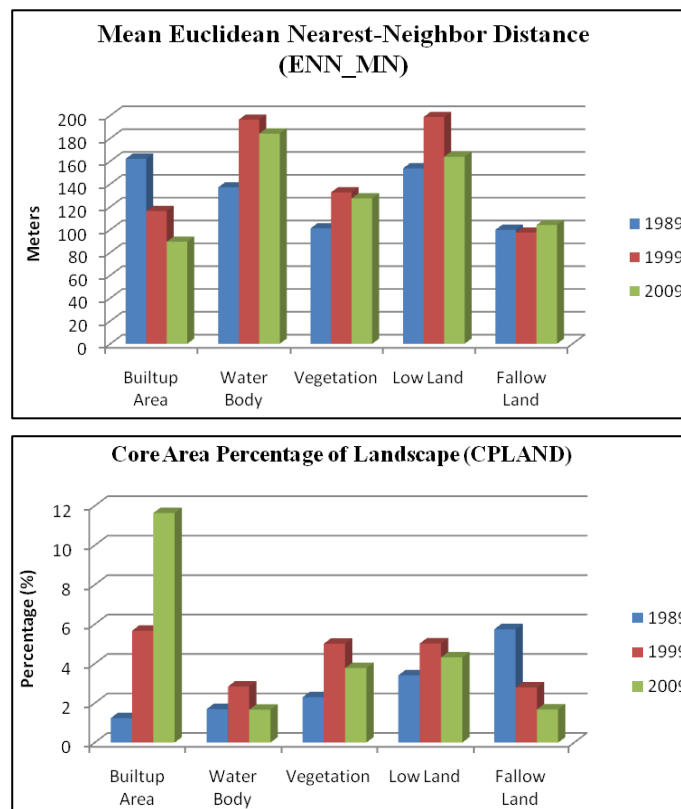
**Figure 4.2: Edge Density and Mean Fractal Dimension Index**

A FRAC\_MN value greater than 1 for a 2-dimensional patch indicates a departure from Euclidean geometry (i.e., an increase in shape complexity). FRAC\_MN approaches 1 for shapes with very simple perimeters such as squares, and approaches 2 for shapes with highly convoluted, plane-filling perimeters [50].

FRAC\_MN values of builtup area, low land and water body land cover types are increasing (Figure 4.2). All the values are showing just above 1 and less than two. This proves the shape of builtup cover type is changing from simple square shape to complex shapes over time. The more complex the shapes the more the urban growth is dispersed and unplanned. Therefore, the FRAC\_MN values indicate massive and haphazard growth of urban patches.

### **4.3.3 Mean Euclidean Nearest-Neighbour Distance and CPLAND**

Mean Euclidean Nearest-Neighbour Distance (ENN\_MN) approaches 0 as the distance to the nearest neighbour decreases. It is perhaps the simplest measure of patch context and has been used extensively to quantify patch isolation [50]. ENN\_MN of builtup area land cover type is decreasing over time (Figure 4.4). This proves that the urban patches of builtup area type are getting closer to each other. It means the urban areas are getting clumsy indicating absence of proper planning.

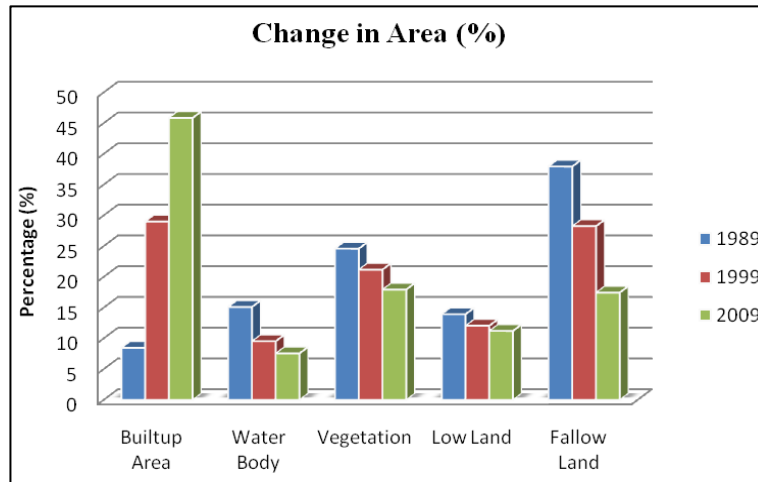


**Figure 4.3: ENN\_MN and CPLAND**

Core Area Percentage of Landscape (CPLAND) approaches 0 when core area of the corresponding patch type (class) becomes increasingly rare in the landscape, because of increasing smaller patches and/or more convoluted patch shapes. CPLAND approaches 100 when the entire landscape consists of a single patch type (i.e., when the entire image is comprised of a single patch) [50]. Here edge influence distance of 100 m has been specified for generating CPLAND values. In this case, CPLAND of builtup area type is increasing and fallow land type is decreasing (Figure 4.3). It means the core area of builtup area type is increasing over time and this situation is vice-versa for fallow land.

### **4.3.4 Change in Area**

There are three possible change detection analyses over time for this research. These are change between 1989 and 1999, 1999-2009 and 1989-2009. The summaries of land cover classification statistics (1989-1999) have been attached in Appendix F.



**Figure 4.4: Land Cover Change in Area (Percentage)**

It is clear that over the years (1989 to 2009) builtup area has increased in huge percentage (from 8.4% to 46%). It is also noteworthy that fallow has decreased in good rate (from 38% to 17%). Other land cover types have decreased in a very small amount (Figure 4.4).

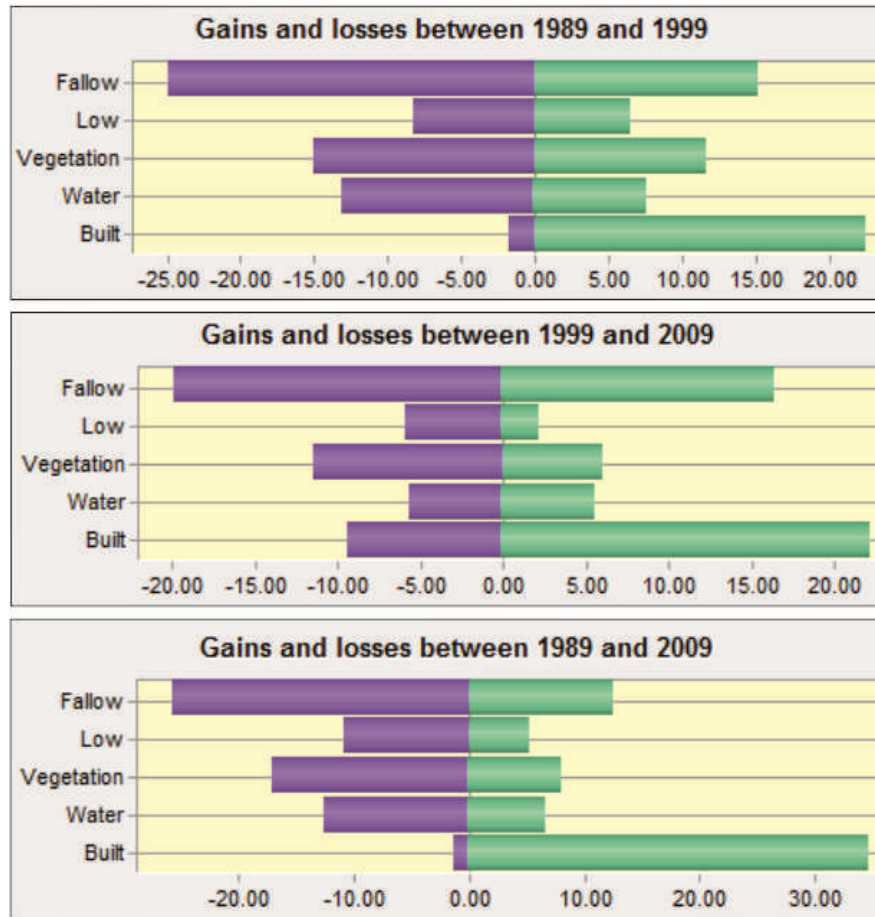
### **4.3.5 Gains and Losses by Category**

Figure 4.5 illustrates that builtup area has increased over the years while there is slight loss in this category. It means some parts of the previously existed builtup areas have converted to some other land cover classes, while vast new area has transformed into builtup area from other classes.

Gains in builtup area are evident in all three combinations. Again fallow land cover type is decreasing in large percentage in all the years. The changes (in terms of gains and losses) in other land cover types are almost the same or not influencing.

Therefore an abrupt increase in builtup area and decrease in fallow land cover type is quite clear from this kind of analysis.





**Figure 4.5: Gains and Losses of Land Covers by Category (Unit: % of Area)**

#### **4.3.6 Contributors to Net Change Experienced by Builtup Area**

Figure 4.6 illustrates which land cover type is contributing more to net change in built-up area. It is found that fallow land is contributing most converting towards builtup area followed by water body and vegetation.

#### **4.3.7 Transition to Builtup Area**

Figure 4.7 shows which areas from other land cover types have been converted to built-up areas. The dominance of fallow land (orange color) is clear here which justifies the explanation of Figure 4.6.

Overall it can be concluded that builtup area is increasing and the contributions are mainly from fallow land and water body.

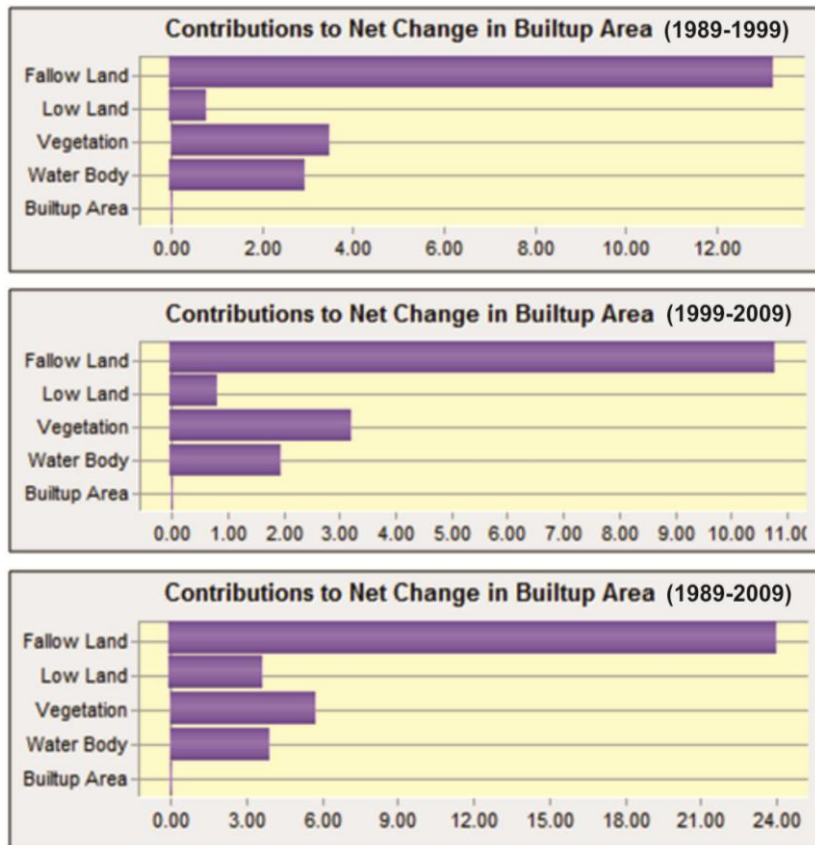


Figure 4.6: Contributors to Net Change Experienced by Builtup Area (Unit: % of Area)

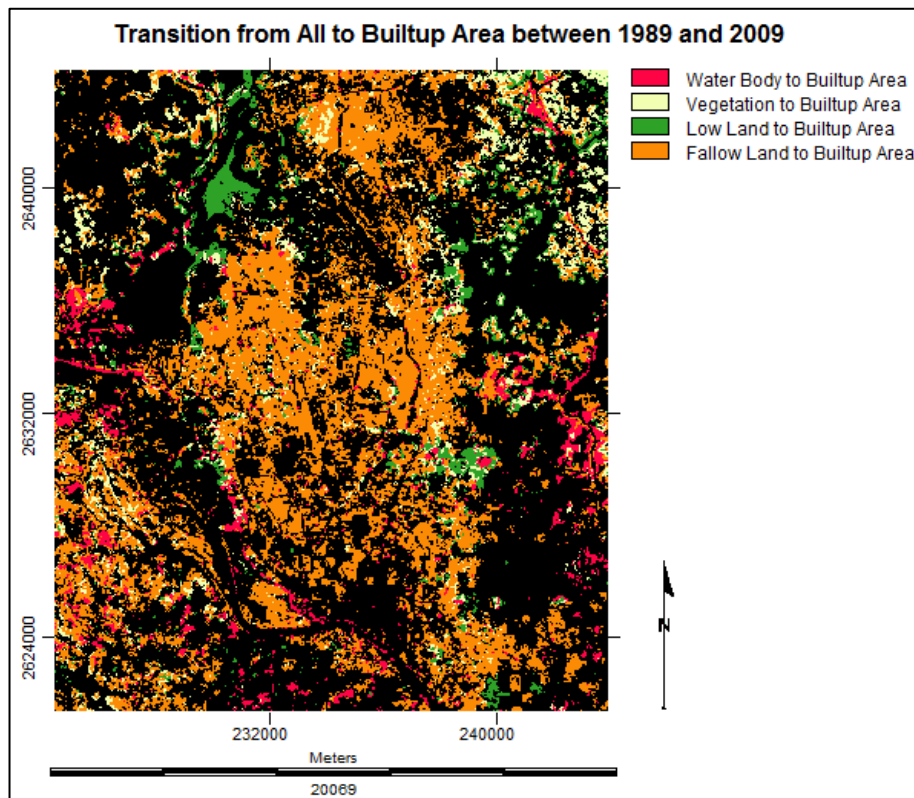


Figure 4.7: Transition of Other Land Cover Types into Builtup Area (1989-2009)

#### **4.3.8 Gains and Losses in Land Cover Types**

In case of builtup area, the core southern part of Dhaka city has remained the same. While the north-east and south-west parts have converted to builtup areas. The northern part of Dhaka city has gained water body followed by a massive decrease in the south-east and south-west parts (Figure 4.8).

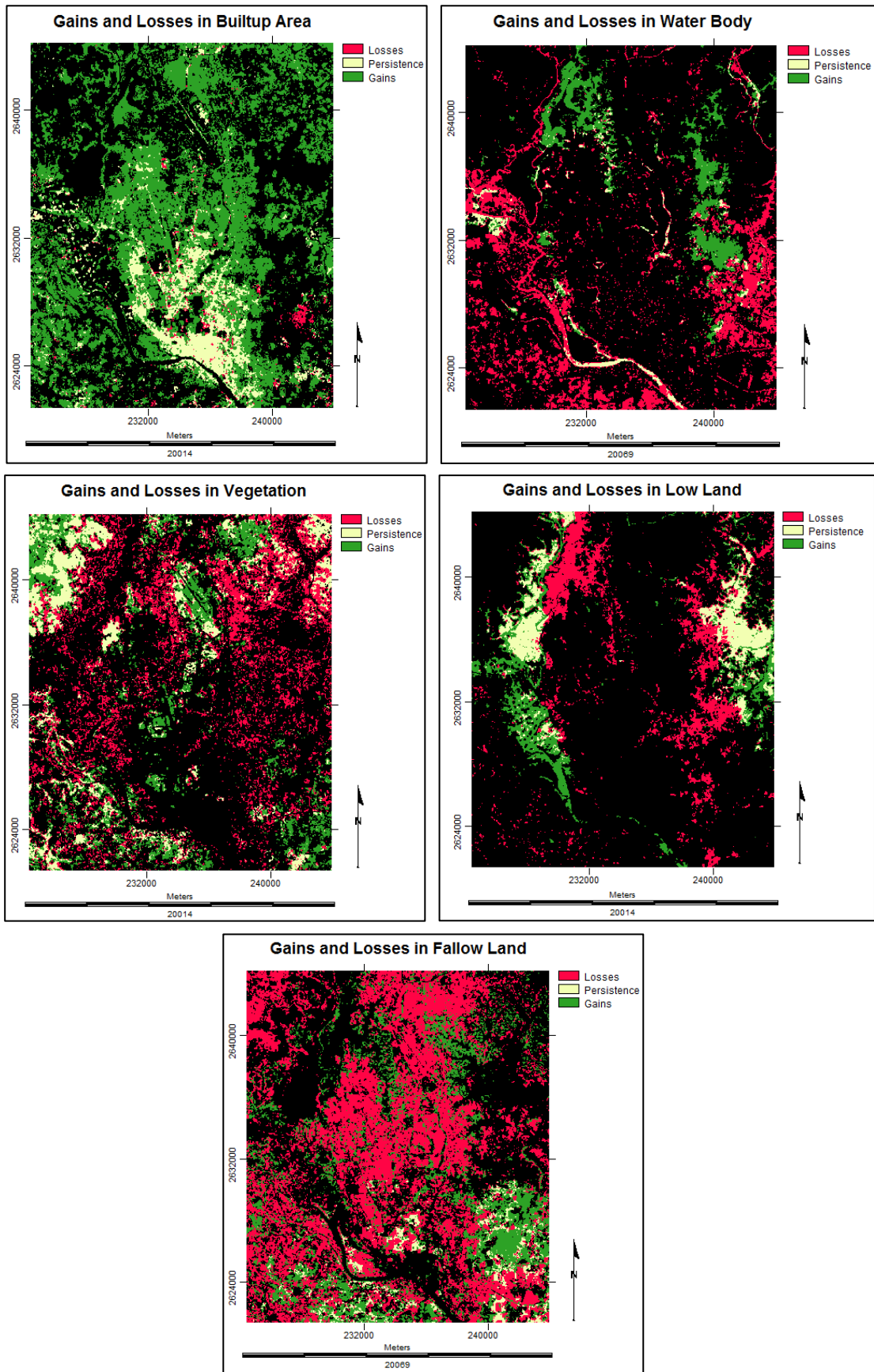
No particular pattern on gains or losses is found for vegetation. In cases of low land the changes are evident in eastern and western parts. Fallow land has decreased markedly and the losses are clear in north-western and mid parts (Figure 4.8).

#### **4.4 Summary of Land Cover Change Detection Analysis**

In summary, it can be stated that in any combinations of time periods (1989-1999 or 1999-2009 or 1989-2009) some similarities have been found. Those are as follows:

- a) The increase in builtup area is prominent.
- b) The net contribution to builtup area is mainly from fallow land followed by water body.
- c) The expansion of urban area is following the northern and western portions of Dhaka city.
- d) The shape of urban patches is converting from simple square shape to complex shapes.
- e) The growth of urbanization is haphazard indicating the absence of proper planning.

Therefore, fallow land and water body types are basically converting into builtup areas over the years. This is the general trend of land cover change pattern for Dhaka city from 1989-2009.



**Figure 4.8: Gains and Losses in Land Cover Types (1989-2009)**

## Chapter 5

### Stochastic Markov Model

---

The details of stochastic Markov modelling process have been described in this chapter.

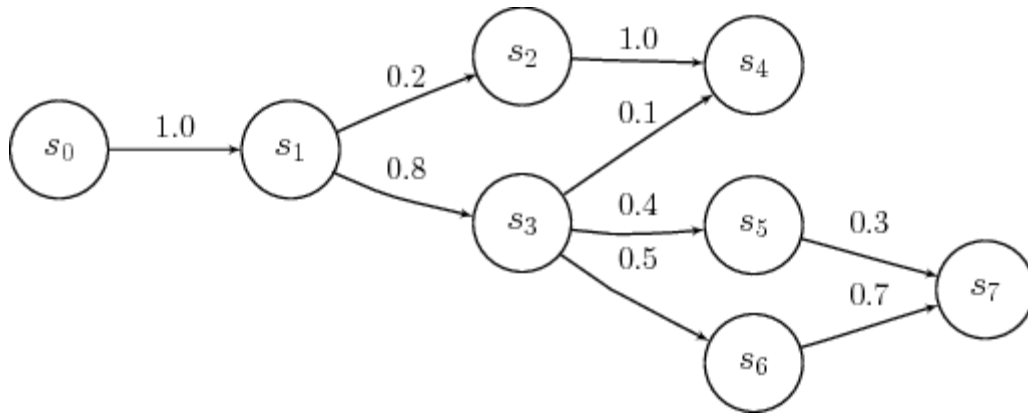
#### **5.1 Stochastic Process**

It is a statistical process involving a number of random variables depending on a variable parameter (e.g. time). The cases of stochastic process can be considered separately as discrete time and continuous time. A discrete time stochastic process  $X_d = \{X_n, n = 0, 1, 2, \dots\}$  is a countable collection of random variables indexed by non-negative integers. A continuous time stochastic process  $X_c = \{X_t, 0 \leq t < \infty\}$  is an uncountable collection of random variables indexed by non-negative real numbers. In general stochastic process is a process in which the transition from one state to another may include some measure of uncertainty [51].

**Example:** A stochastic or random process is a sequence of random variables ( $\xi_1, \xi_2, \xi_3, \dots$ ) based on the same sample space  $\Omega$ . Suppose, there are 4 telephone lines and at a given time (moment) 0, 1, 2, 3 or 4 of those lines can be busy. Now the task is to observe how many lines will be busy once in a minute. The outcome will be a random variable with  $\Omega_\xi = \{0, 1, 2, 3, 4\}$ . Let  $\xi_1$  be the number of busy line/lines at the first observation time,  $\xi_2$  be the number of busy line/lines at the second observation time etc. The sequence of the number of busy lines then forms a random process with discrete random variables and a discrete time parameter [52].

#### **5.2 Markov Chain**

Andrei Andreevich Markov (1856-1922), a Russian Mathematician, invented 'Markov Chain' concept in 1906 [53]. A Markov chain is a stochastic process (based on probabilities instead of certainties) with discrete state space and discrete or continuous parameter space with the Markov property [54]. In this random process the current state of a variable or system is independent of all the past states, except current/present state [53]. In simple, the state of a system  $s$  at time  $(t+1)$  depends only on the state of the system at time  $t$ , not on the previous states. Some examples of Markov processes are flow of traffic, behaviour of business or economy (stock/share price), progress of an epidemic etc. Figure 5.1 depicts an example of a Markov Chain [51].



**Figure 5.1: Example of a Markov Chain [51]**

The probability of being in state  $s_6$  at a certain time depends on state  $s_3$ , not on all other previous states ( $s_1$  or  $s_0$ ). This is a typical Markov Chain (Figure 5.1).

### 5.2.1 Markov Property

In a Markov chain the probability of the next state is only dependent upon the current state. This is called Markov property and stated as [52]:

$$P(\xi_{t+1} = x_{i_{t+1}} | \xi_1 = x_{i_1}, \dots, \xi_t = x_{i_t}) = P(\xi_{t+1} = x_{i_{t+1}} | \xi_t = x_{i_t})$$

The probability of a Markov chain  $\xi_1, \xi_2, \dots$  can be calculated as [52]:

$$\begin{aligned} P(\xi_1 = x_{i_1}, \dots, \xi_t = x_{i_t}) &= \\ &= P(\xi_1 = x_{i_1}) \cdot P(\xi_2 = x_{i_2} | \xi_1 = x_{i_1}) \cdot \\ &\dots \cdot P(\xi_t = x_{i_t} | \xi_{t-1} = x_{i_{t-1}}) \end{aligned}$$

The conditional probabilities:  $P(\xi_{t+1} = x_{i_{t+1}} | \xi_t = x_{i_t})$

These are called the ‘Transition Probabilities’ of the Markov chain.

### 5.2.2 Transition Matrix for a Markov Chain

Let’s consider a Markov chain with  $n$  states  $s_1, s_2, \dots, s_n$ . Let  $p_{ij}$  denote the transition probability from state  $s_i$  to state  $s_j$ , i.e.  $P(\xi_{t+1} = s_j | \xi_t = s_i)$

The transition matrix of this Markov process is then defined as [52]:

$$P = \begin{bmatrix} p_{11} & \dots & p_{1n} \\ \dots & \dots & \dots \\ p_{n1} & \dots & p_{nn} \end{bmatrix}, \quad p_{ij} \geq 0, \quad \sum_{j=1}^n p_{ij} = 1, \quad i = 1, \dots, n$$

Predictions of the future state probabilities can be calculated by solving the matrix equation [55]:

$$p(t) = p(t-1) \cdot P$$

With increasing time steps, a Markov chain may approach to a constant state probability vector, which is called limiting distribution [55]:

$$p(\infty) = \lim_{t \rightarrow \infty} p(t) = \lim_{t \rightarrow \infty} p(0) \cdot P^t$$

### **5.2.3 Example of Markov Chain**

Let's assume a simple weather model:

**1. Raining Today:**

⇒ 40% rain tomorrow

⇒ 60% no rain tomorrow

**2. Sunny Today:**

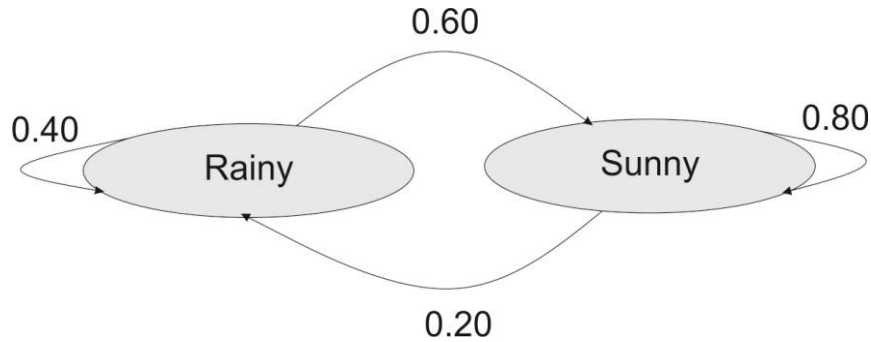
⇒ 20% rain tomorrow

⇒ 80% no rain tomorrow

Therefore, the transition matrix:

$$P = \begin{bmatrix} 0.40 & 0.60 \\ 0.20 & 0.80 \end{bmatrix}$$

This matrix represents a weather model. As this is a stochastic matrix, therefore the rows sum to 1. The matrix depicts, a rainy day is 40% likely to be followed by a rainy day, while a rainy day is 20% likely to be followed by a sunny day (Figure 5.2).



**Figure 5.2: Stochastic Weather Model**

### **5.2.3.1 Weather Prediction**

Suppose the weather on ‘Day 0’ is known to be ‘Rainy’. This represents a vector with 100% ‘Rainy’ and 0% ‘Sunny’ day.

$$p(0) = [1 \quad 0]$$

Now the weather on ‘Day 1’ can be predicted as follows:

$$p(1) = p(0). P = [1 \quad 0] \begin{bmatrix} 0.40 & 0.60 \\ 0.20 & 0.80 \end{bmatrix} = [0.40 \quad 0.60]$$

The weather on ‘Day 2’ can be predicted as:

$$p(2) = p(1). P = [0.40 \quad 0.60] \begin{bmatrix} 0.40 & 0.60 \\ 0.20 & 0.80 \end{bmatrix} = [0.28 \quad 0.72]$$

It means the weather on ‘Day 2’ is expected to be 28% rainy or 72% sunny. This is how the weather can be predicted using Markov chain analysis.

## **5.3 Stochastic Markov Model**

The first model that has been implemented is given the name as ‘Stochastic Markov Model (St\_Markov), because this model combines both the stochastic processes as well Markov chain analysis techniques. This kind of predictive land cover change modelling is suitable when the past trend of land-cover change pattern is known.

In general, in Markovian processes the future state of a system in time  $t_2$  can be modelled/predicted based on the immediate preceding state; time  $t_1$ . Therefore, the future state can be predicted not based on the past but rather the present. Here the past and future are independent [36].



Markov chain produces a transition matrix (Table 5.1), a transition areas matrix (Table 5.2) and a set of conditional probability images by analyzing two qualitative land cover images (Figure 5.3) from two different dates (1989 and 1999) [36].

**Table 5.1: Markov Probability of Changing among Land Cover Types**

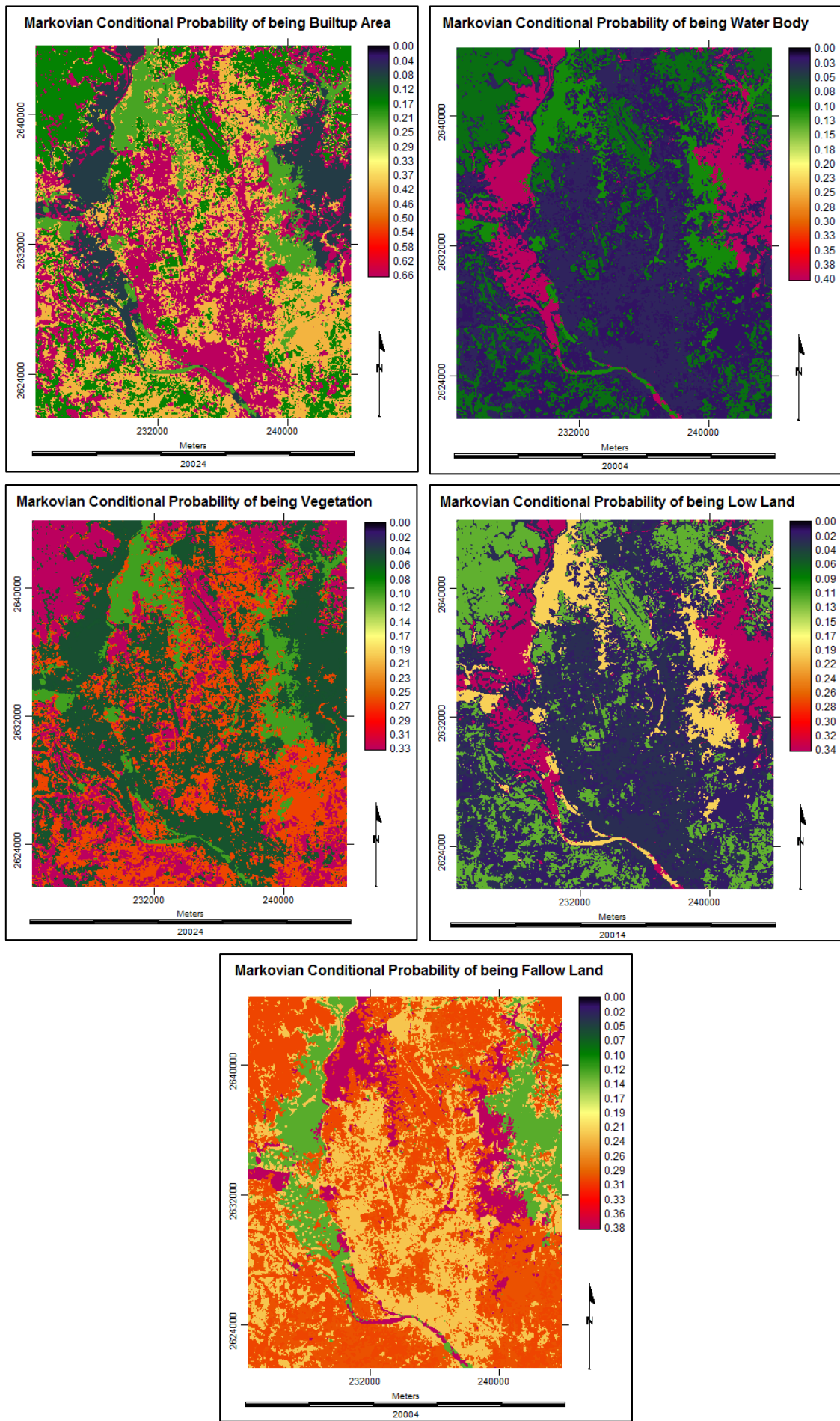
	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>Builtup Area</b>	0.6649	0.0268	0.0533	0.0298	0.2252
<b>Water Body</b>	0.2125	0.1074	0.1030	0.1969	0.3802
<b>Vegetation</b>	0.1675	0.0853	0.3304	0.1173	0.2995
<b>Low Land</b>	0.0766	0.4006	0.0514	0.3446	0.1267
<b>Fallow Land</b>	0.4126	0.0144	0.2603	0.0199	0.2928

**Table 5.2: Cells Expected to Transition to Different Classes**

	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>Builtup Area</b>	95476	3845	7658	4278	32332
<b>Water Body</b>	10034	5074	4865	9300	17953
<b>Vegetation</b>	17569	8945	34655	12302	31409
<b>Low Land</b>	4574	23914	3070	20572	7566
<b>Fallow Land</b>	57732	2009	36415	2789	40964

The matrix of transition probabilities (Table 5.1) shows the probability that each land cover category will change to other categories in 2009. Table 5.2 represents the number of cells/pixels (30 m × 30 m) that will be transformed over time from one land cover type to other types.

Markov Chain Analysis also produces related conditional probability images (Figure 5.3) with the help of transition probability matrices. These images are called conditional because the probability is conditional to the current state. These maps/images, which have been projected from the two previous land cover images, are useful for future prediction of land cover change. Each conditional probability image shows the possibility of transitioning to another land cover class.



**Figure 5.3: Markovian Conditional Probability Images**

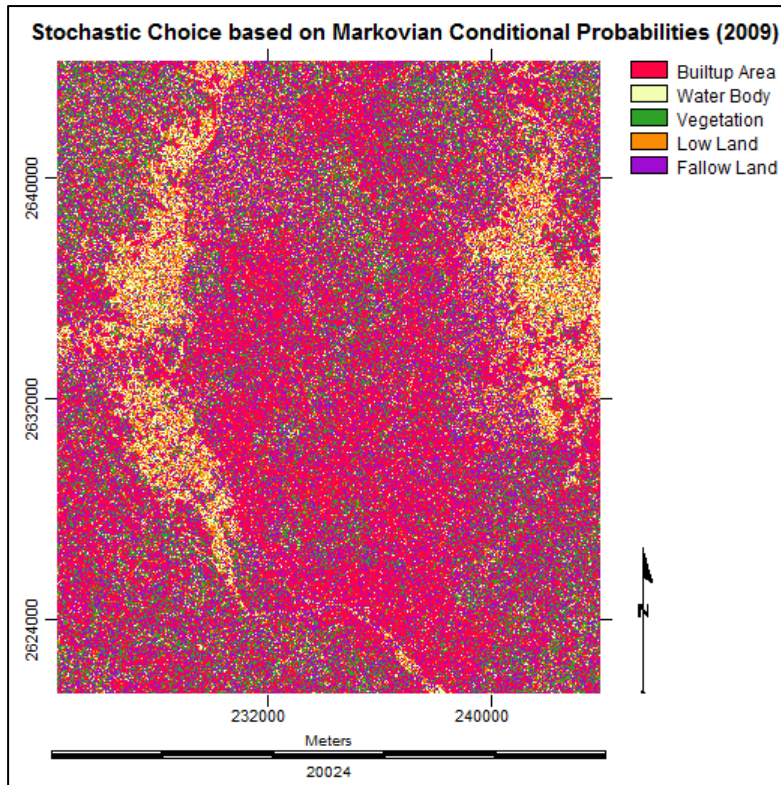
After analyzing Figure 5.3, it is clear that most areas will convert into builtup areas. The Markovian conditional probability of being builtup area ranges upto 0.66 which is the highest among all the land cover types. This probabilistic prediction is dependent upon the past trend of the last ten years (1989-1999). From change detection analysis, it is found that most areas are being converted to builtup areas whereas the Markovian conditional probability images are also showing the same trend.

The next step is to make one single land cover map for future prediction aggregating all the Markovian conditional probability images. The final predicted land cover map of 2009 will be based on the past ten year's land cover change pattern on the basis of Markov chain analysis. This prediction is performed by a stochastic choice decision model [37]. Stochastic Choice creates a stochastic land cover map by evaluating and aggregating the conditional probabilities in which each land cover can exist at each pixel location against a rectilinear random distribution of probabilities [36]. For each type of land cover, this decision model gives a random value from 0.0 to 1.0 for each pixel [37]. Then for each pixel this process continues cumulatively adding conditional probabilities, in the order of the Markovian conditional probability images. When the pixel surpasses the random value, that particular pixel is predicted to be the land cover type for the next time period [37]. Table 5.3 illustrates the stochastic mechanism.

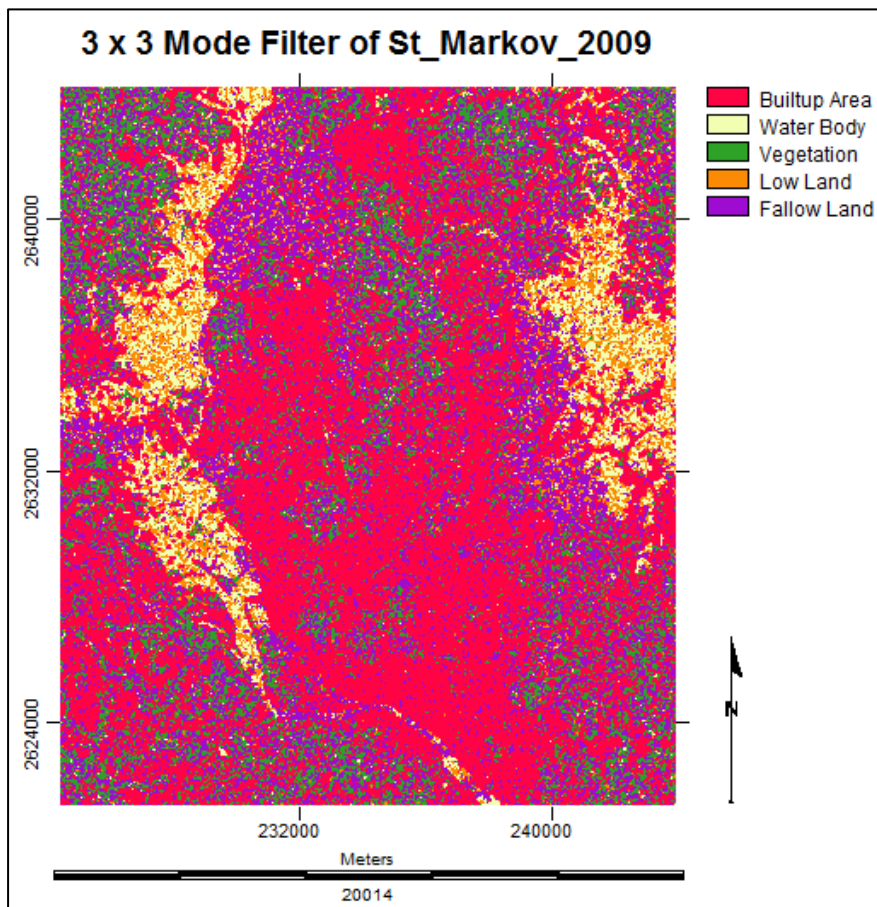
**Table 5.3: Stochastic Random Process for Selecting Land Cover Type**

Let's assume for a scenario, the random value = 0.76		
Land Cover Type	Conditional Probability	Cumulative Sum
Residential Area	0.27	0.27
Wet Land	0.36	0.27+0.36=0.63
<b><i>Open Space</i></b>	<b><i>0.17</i></b>	<b><i>0.63+0.17=0.80</i></b>
Industrial Area	0.06	0.80+0.06=0.86
Forests	0.14	0.86+0.14=1.0
The chosen land cover = <b>OPEN SPACE</b>		

In the case of Table 5.3, open space has been selected because the random value (0.76) for a particular pixel exceeds (0.80) in case of the land cover type; open space. This is how the whole process works. The Stochastic Markov (St\_Markov) predicted land cover map of 2009 is shown in Figure 5.4. At the end, a 3×3 Mode filter has been applied to generate a much clear St\_Markov predicted map of 2009 (Figure 5.5).



**Figure 5.4: Stochastic Markov Predicted Land Cover Map (2009)**



**Figure 5.5: Final St\_Markov Predicted Land Cover Map (2009)**



## Chapter 6

### Cellular Automata Markov Model

---

The details of cellular automata Markov modelling process with its theoretical frameworks have been described in this chapter.

#### **6.1 Cellular Automata**

The concept of Cellular Automata (CA) was first introduced by Johann Louis von Neumann in 1948. The purpose was modelling biological self-reproduction [56]. At the same time Stanislaw Marcin Ulam suggested the same idea by introducing the concept of ‘Cellular Spaces’ [57]. Therefore, both of them are considered the developers of this CA concept [58]. In the seventies, John Horton Conway made a popular application called ‘Game of Life’ based on CA concept. This was a kind of simulation game [57]. This concept was originated in computer science but now it is being used by other disciplines like Urban Planning, Geography, Mathematics, Physics, Natural Science, Transportation, and Biology etc. [59]. CA has also been used to better understand and predict urban growth [60].

#### **6.1.1 What are Cellular Automata?**

According to Torrens (2000), “*An automaton essentially comprises a finite state machine that exists in some form of tessellated cell-space*” [61]. Stephen Wolfram (1983) defined CA as follows: “*Cellular automata are simple mathematical idealizations of physical systems in which space and time is discrete, and physical quantities take on finite set of discrete values*” [56].

Cellular automata have been defined as simple dynamic spatial systems where the state of each cell in an array depends on the previous state of the cells within a neighbourhood, according to a set transition rules [62].

#### **6.1.2 The Elements of Cellular Automata**

The components that comprise an elementary cellular automaton are as follows [61]:

- a) The physical environment or the space represented by an array of cells, on which the automaton exists (its lattice)

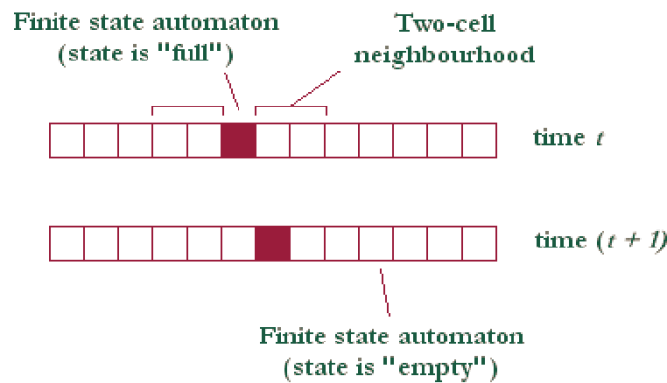
- b) The cell in which the automaton resides that contains its state(s)
- c) The neighbourhood around the automaton
- d) Transition rules that describe the behaviour of the automaton
- e) The temporal space in which the automaton exists

These elements have been described in brief to better understand the concept of CA:

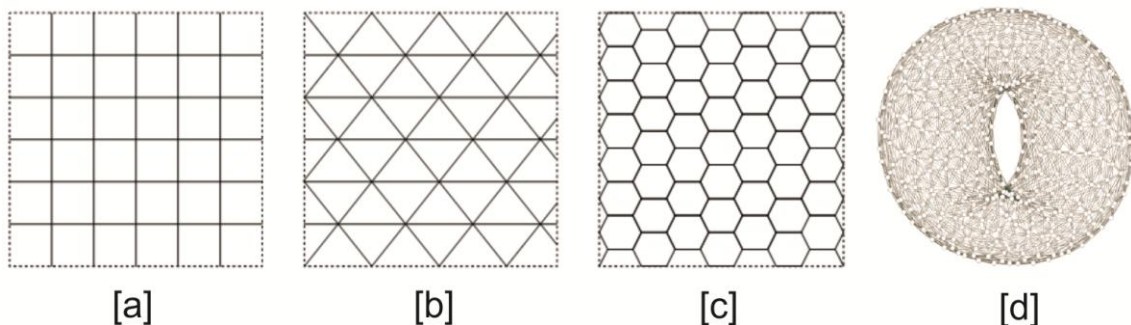
### 6.1.3 The Cell Space

The cell space (lattice) is composed of individual cells. In an elementary CA, the lattice is one-dimensional (Figure 6.1). This is a linear string of cells called an elementary cellular automaton [57]. But it can be  $n$ -dimensional. CA lattices can be of infinite proportions with any dimension.

In most cases, lattices are defined in a regular fashion: grid squares, triangles, hexagons, torus (ring-shaped surface) etc. (Figure 6.2). Figure 6.1 and Figure 6.2 have been adopted from [61] and [57].



**Figure 6.1: One-Dimensional Cellular Automata [61]**

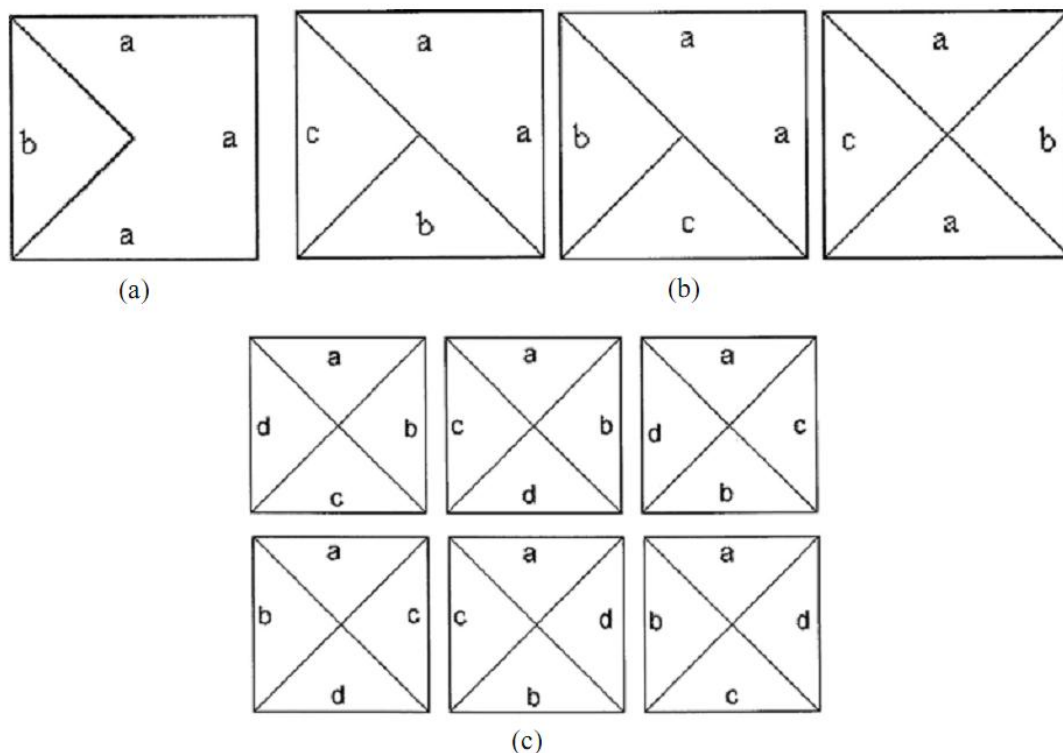


**Figure 6.2: Two-Dimensional Cellular Automata Grid: [a] Rectangular, [b] Triangular, [c] Hexagonal and [d] Torus [57]**

### **6.1.4 The Cell States**

Each cell in the lattice can exist in a number of distinct states, which define the occupancy of the cell. The cell can be empty or contain a specific ingredient (e.g. molecule, particle, organism, land use etc.) [58]. A cell's state is typically represented by an integer (binary state), but it can also have a continuous range of values [57].

Recently, the idea of variegated cell, in which each edge can have its own independent rules for interacting with each other, has been introduced. Examples of some types of variegated cells are shown in Figure 6.3 [59].

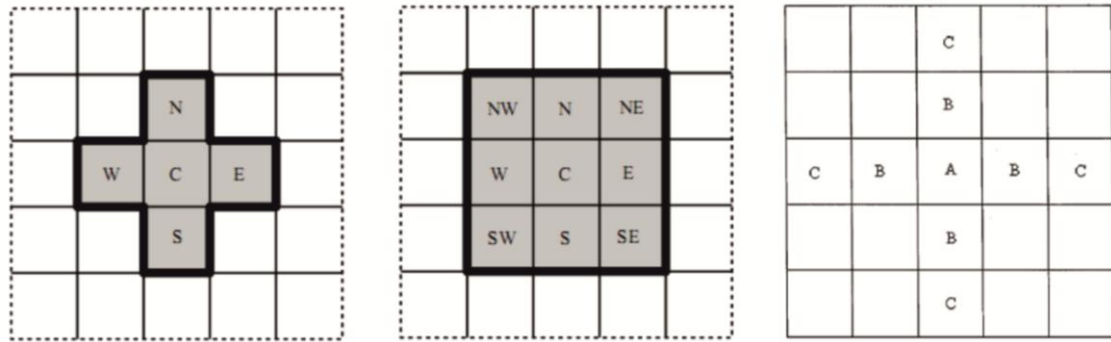


**Figure 6.3: Examples of Variegated Cells: (a) Two different types of edges with different rules, (b) Three different types of edges with different rules and (c) Four different types of edges with different rules [59]**

### **6.1.5 The Cell Neighbourhood**

The movements/ actions of an ingredient on the lattice are governed by some rules. These rules depend on the nature cells in close proximity to the ingredient. This proximity environment of a cell is called neighbourhood [58]. In one-dimensional cellular automata (1D lattice), each cell has two neighbours, while in a 2D rectangular lattice there are several possibilities [57]. The most common neighbourhood in 2D CA

is called ‘von Neumann Neighbourhood’ with a radius of 1 there are four adjacent cells (north, east, south and west) including the cell itself. One more popular example is ‘Moore Neighbourhood’ including the previous five cells as well as four north-east, south-east, south-west and north-west diagonal cells. Another useful example is ‘extended von Neumann Neighbourhood’ where the four C cells lying beyond the four B cells of von Neumann Neighbourhood. All the examples are shown in Figure 6.4 [58].



**Figure 6.4: Two-Dimensional CA Neighbourhoods: (Left) the von Neumann Neighbourhood, (Middle) the Moore Neighbourhood and (Right) the extended von Neumann Neighbourhood of cell A [58]**

### 6.1.6 The Transition Rules

A transition rule (also called function) acts upon a cell and its neighbourhood, such that the cell’s stage changes from one discrete time to another (i.e. the system’s iterations). The same rule is subsequently applied to all the cells in parallel [57]. These rules are the heart of CA that guides its dynamic evolution [59].

Transition rules influence the probability that an ingredient (e.g. land cover type) will transform to a different ingredient during iterations of the simulation [58]. For example suppose:

**Case 1:** If  $P_T(AB) = 1.0$ , then the transition  $A \rightarrow B$  is certain to occur.

**Case 2:** If  $P_T(AB) = 0.0$ , then the transition  $A \rightarrow B$  will never occur.

**Case 3:** If  $P_T(AB) = 0.5$ , then during each iteration, there will be a 50% change that the transition  $A \rightarrow B$  will occur.

Case 1 and Case 2 are called ‘Deterministic CA Model’, as there are no possibilities for different outcomes.



Case 3 is considered as ‘Stochastic (Probabilistic) CA Model’ since it allows different outcomes, like the ingredient might remain unchanged or it might transform to a different state.

### **6.1.7 The Temporal Space**

A CA evolves at sequence of discrete time steps. The temporal evolution of cells destroys the independence of initial cell states [61]. At each step, the cells are updated simultaneously based on transition rules.

### **6.1.8 Mathematical Notation of Cellular Automata**

A CA model represents discrete dynamic system consisting of four elements [57]:

$$CA = (\mathcal{L}, \Sigma, \mathcal{N}, \delta)$$

Where,  $\mathcal{L}$  = the discrete lattice or the physical environment

$\Sigma$  = the set of possible states, where each  $i^{th}$  cell of the lattice at time step  $t$  has a state  $\sigma_i(t) \in \Sigma$

$$N_s(r) = \{(i_1, i_2) / dist((i_1, i_2)) < r\}$$

$\mathcal{N}$  = the neighbourhood of a cell automaton, which is defined as all cells that fall within a radius  $r \in \mathcal{N}$  around the actual cell. It denotes the idea of neighbourhood template [63]:

$$N_s(r) = \left\{ \frac{i_1, i_2}{dist((i_1, i_2))} < r \right\}, \text{ where, } dist((i_1, i_2)) = \sqrt{(i_1^2 + i_2^2)}$$

$N_s(r)$  = the relative index of all neighbours of a particular cell

$\delta$  = the local transition rule, which is denoted as follows [57]:

$$\delta : \Sigma^{|\mathcal{N}|} \longrightarrow \Sigma : \bigcup_{j \in \mathcal{N}_i(t)} \sigma_j(t) \longmapsto \sigma_i(t+1)$$

It shows that the state of the  $i^{th}$  cell at the next time step  $t+1$  is computed by  $\delta$  based on the states of all the cells in its neighbourhood at the current time step  $t$ .

Here,  $N_i(t)$  = the associated neighbourhood with  $i^{th}$  cell at time  $t$

$|\mathcal{N}|$  = the number of cells in the neighbourhood

The local transition rule is given by a rule table where given the sizes of  $\Sigma$  and  $N$ , the total number of possible rules equals [57]:

$$|\Sigma^{\Sigma^N}|$$

Where each of the  $|\Sigma^N|$  possible configurations of a cell's neighbourhood is mapped to the number of possible states a cell can be in. Now considering the ordered set of all the states of all cells collectively at time step  $t$ , a CA's global configuration can be denoted as follows [57]:

$$\mathcal{C}(t) = \bigcup_{j \in \mathcal{L}} \sigma_j(t)$$

Now applying the local transition rule to all the cells in the CA's lattice, the next configuration of the CA can be computed by its induced global map [57]:

$$G : \Sigma^{\mathcal{L}} \longrightarrow \Sigma^{\mathcal{L}} : \mathcal{C}(t) \longmapsto \mathcal{C}(t + 1)$$

In brief, the standard CA can be generalized as follows [64]:

$$S^{t+1} = f(S^t, N)$$

Where,  $S$  = the set of all possible states of the cellular automata;  $N$  = a neighbourhood of all cells providing input values for the function  $f$  and  $f$  = a transition function that defines the change of the state from  $t$  to  $t+1$

### **6.1.9 Running a Simulation**

After defining the basic elementary parameter values, the next step is to consider the following conditions for the simulation [58]:

- a) The nature and number of starting items (e.g. land cover type)
- b) The configuration of the initial state of the system
- c) The number of runs to be carried-out for the simulation
- d) The length of the runs or the number of iterations it should include

The classical approach of a CA model has been described in this section. Over the years, the classical CA model has been modified (like CA\_Agent Model, CA\_Markov Model etc.) by researchers, to solve particular problems and overcome the shortcomings of the classical approach.

## **6.2 Cellular Automata Markov Model**

The second model that has been implemented to predict the future land cover map of Dhaka city is known as 'CA\_Markov Model'.

CA\_Markov combines the concepts of Markov Chain, Cellular Automata, Multi-Criteria Evaluation (MCE) and Multi-Objective Land Allocation (MOLA) [66].

CA\_Markov is useful for modelling the state of several categories of a cell based on a matrix of Markov transition areas; transitional suitability images and a user defined contiguity filter [34].

The concepts of MCE and MOLA have been described in Appendix G.

## **6.3 How CA Markov Model Works?**

Initially CA\_Markov takes as input the base land cover map from which the changes in the future will be projected. Other inputs are transition matrix, suitability maps and a contiguity filter.

In Markov model, the extent of each land cover category in time  $t_2$  is obtained by the degree of difference of each category in time  $t_1$  and  $t_0$ .

The output is a non-spatial transition matrix which is based on the assumption that transition probability or transition area matrices are constant over time. But this assumption is not true in all cases or in reality [34].

Before it starts the iterative process, it is necessary to input the total number of iterations that will be implemented for future prediction. In general, within iteration, each land cover class can lose some of its land to one or more other land cover classes.

The opposite can also happen as each land cover class can gain some lands from one or more other land cover classes. The estimating year is 2009 ( $t_2$ ) that is based on the transitional years 1999 ( $t_1$ ) and 1989 ( $t_0$ ).

In this regard, there exist two classes: host and claimant class [36]. Claimant class gains lands from the host class (host can be one or more classes). This selection of lands by the claimant class is chosen based on the suitability maps. There is a clear competition for specific land parcels that is performed using Multi-Objective Land Allocation (MOLA) [37].

Basically the loss or gain of land parcels is determined by the suitability maps, ranking and contiguity filter [34]. The suitability maps determine which pixels will change as per the highest suitability of each land cover type. The higher the suitability of a pixel, the possibility of the neighbouring pixels to change into that particular class is higher.

This kind of Markov model applies contiguity rule like a pixel near to an urban area is most likely to be changed into urban area [34]. In this research, a 3×3 mean contiguity filter has been used (Figure 6.5).

0	1	0
1	1	1
0	1	0

**Figure 6.5: The 3 × 3 Mean Contiguity Filter for CA\_Markov Modelling**

This kind of filtering is applied on suitability images for each land cover class. This is a defined neighbourhood. The suitability of a pixel is determined by the pixel values within this defined filtering kernel. The more pixels of the same category of land cover exist in the neighbourhood, the more the suitability value for that particular land cover type increases. Otherwise the pixel value remains the same [34].

This kind of CA contiguity filter rules out to change land cover randomly. It adds some spatial character to the model [36].

The mean contiguity filter together with a Boolean filter mask produces a resulting value of 1 when it is entirely within the existing class and gives the value of 0 when it is entirely outside the class [36]. This output is then multiplied by the class being considered from the original suitability map [67].

Therefore the purpose of this filter is to give less weight of suitability values to the pixels far away from the existing areas of land cover class. At the end, it gives priority to more contiguous suitable areas [67].

At this stage, it generates a new suitability map for input into MOLA [37]. MOLA maximizes the suitability of lands as per some assigned weights for a certain objective [66]. MOLA process solves the problem with land allocation.

MOLA allocates a particular cell for each land cover type as per the specific objective to the highest weighted suitability. This reduces the amount of area to be assigned for each class [67].

MOLA allocates  $1/n^{th}$  of the required land in the first run,  $2/n^{th}$  in second run and so on, until the full allocation of each land cover is obtained. Here the transition areas file determines the amount of land is allocated for each land cover class over the years [37].

Ranking is an important issue to determine which pixel is going to change to another category or remain the same. Here the ranking is based on a final probabilistic value of a land cover change. It means if a pixel has the highest final probabilistic value, it certainly denotes the highest ranked value to change to that land cover class [34]. In each step of iteration, a new land cover map is produced with overlaying all the MOLA resulting operations. This new map is used as the input for the next iteration [37].

### **6.3.1 Suitability Maps for Land Cover Classes**

There exist some transition rules for preparing suitability maps in land cover change modelling. These rules are based on some socio-economic and biophysical driving forces as well as spatial dependencies [67]. These driving forces differ for each land cover type.

For example: in case of built-up areas one driving force may be the distance from main road. In general, pixels closer to main roads are more suitable for urbanization. There may be many other factors for change detection. Even these factors are applicable for all land cover classes. In case of water body, the driving factors for preparing suitability maps are different.

The presence of some constraints is also possible. Like there may be some protected water bodies or forest areas within the areas of interest. These constraints are expressed in the form of Boolean (logical) maps. In this type of Boolean map, the values 0 represent the areas excluded from consideration and the values 1 represent the areas that will be included for preparing suitability maps [37].

Therefore it is needed to standardize these factors. For this research purpose, the factors or land cover constraints are standardized to a continuous scale of suitability from 0 (the least suitable) to 255 (the most suitable). The 0-255 range provides byte data type [37].

This kind of continuous scale helps to avoid the hard Boolean decision of labelling whether any particular pixel is absolutely suitable or not. This continuous scale for suitability is known as soft or ‘fuzzy’ concept [36]. It gives a certain value to all locations that represent the degree of suitability for each pixel.

These suitability maps are prepared based on Multi Criteria Evaluation (MCE) process where fuzzy set membership is used for standardizing criteria [36].

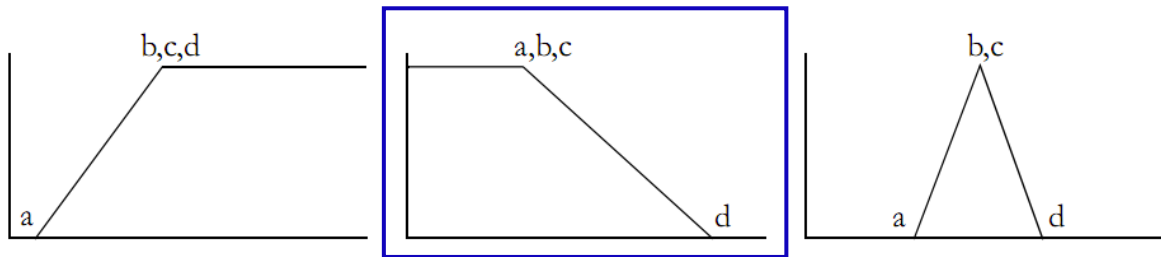
### **6.3.2 Preparing Suitability Maps**

Preparing a suitability map for each land cover type is difficult in terms of data and information availability. It is not possible to incorporate all types of factors or constraints that exist within the study area. Therefore a simple assumption has been assumed for fuzzy factor standardization.

The basic assumption for preparing suitability images: the pixel closer to an existing land cover type has higher suitability. It means a pixel completely within vegetation has the highest suitability value (255) and pixels far from existing vegetation pixels will have less suitability values. The farthest pixels from vegetation will show the lowest suitability values. Here the suitability decreases with distance. Therefore a simple linear distance decay function is appropriate for this basic assumption.

It serves the basic idea of contiguity. A pixel near to an urban area is more likely to convert into urban area in near future. Though this idea is not perfect always like in case of water body and vegetation the scenario can be different. In reality, a pixel near to forest is more likely to convert into builtup area rather in forest. But this can be considered to be the research limitation.

The fuzzy membership function used here is linear type. This kind of fuzzy set analysis requires the positions of 4 inflection points to control the shape of the curve. These inflection points are indicated in Figure 6.6 as ‘a’, ‘b’, ‘c’ and ‘d’ [36].

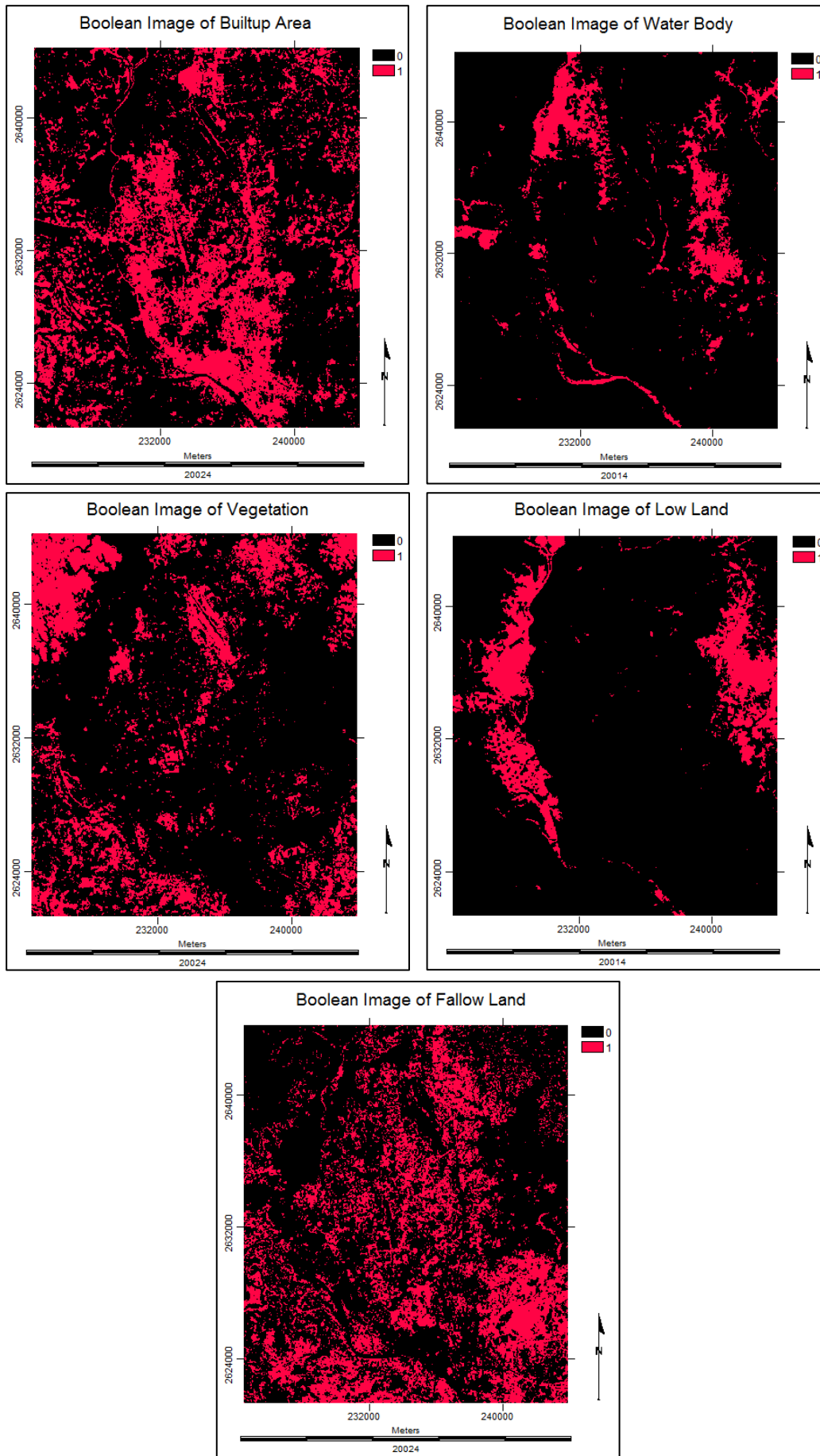


**Figure 6.6: Fuzzy Linear Membership Function**

This linear function can take different shapes. The second curve shows (Figure 6.6) the ‘Monotonically Decreasing’ function. This shape begins at 255 (inflection point ‘c’), then falls and stays at 0 (inflection point ‘d’) [36]. Monotonically Decreasing Linear function has been chosen for this research as there is no complex relationship in the basic assumption (the far from an existing land cover type, the lower the suitability value for a pixel). This completely depicts the scenario for the second curve. The first and third curves respectively show the ‘Monotonically Increasing’ and ‘Symmetric’ functions (Figure 6.6).

At the very beginning, the Boolean images (1999) for each land cover type have been prepared (Figure 6.7). For these Boolean images, the values 1 represent the areas of interest (the particular land cover type) and the values 0 represent the areas of no interest. Then the distance images (Figure 6.8) for each of these Boolean land cover images have been generated. These distance images are important to measure the values of suitability for the pixels of land cover classes. The distance images are produced using simple Euclidean distance function which measures the distance between each cell from the featured image [66]. The unit of measurement is ‘meter’ here.

The lowest and highest values obtained from the distance images have been used as the input for fuzzy set membership analysis. For this case, the lowest value is considered as the input value for the inflection point ‘c’ whether the highest value is used for the inflection point ‘d’ (the middle curve of Figure 6.6). This is how all the Boolean land cover images have been standardized to the same continuous suitability scale (0-255) using fuzzy set membership analysis process (Figure 6.9).



**Figure 6.7: Boolean Images of each Land Cover Type (1999)**



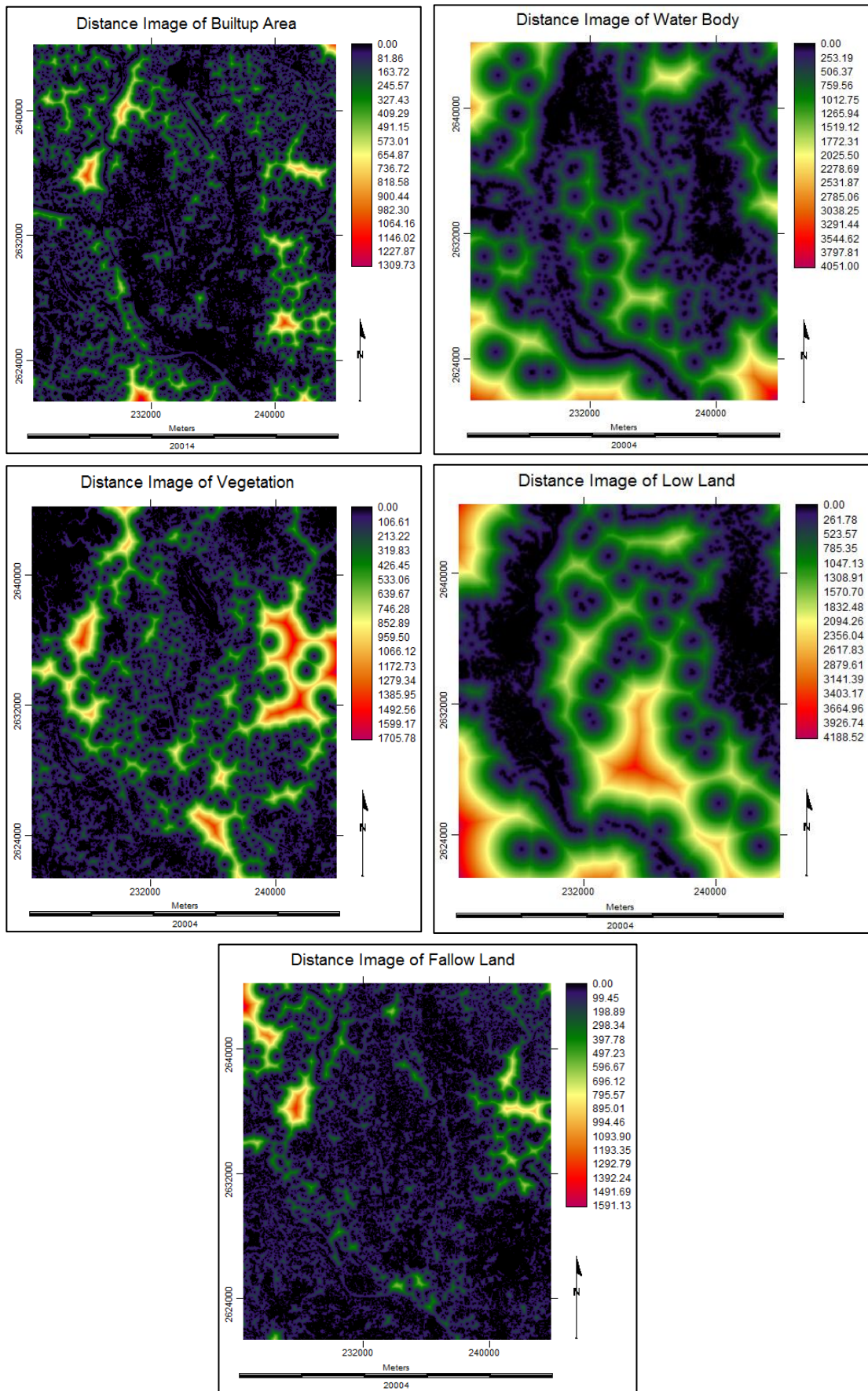
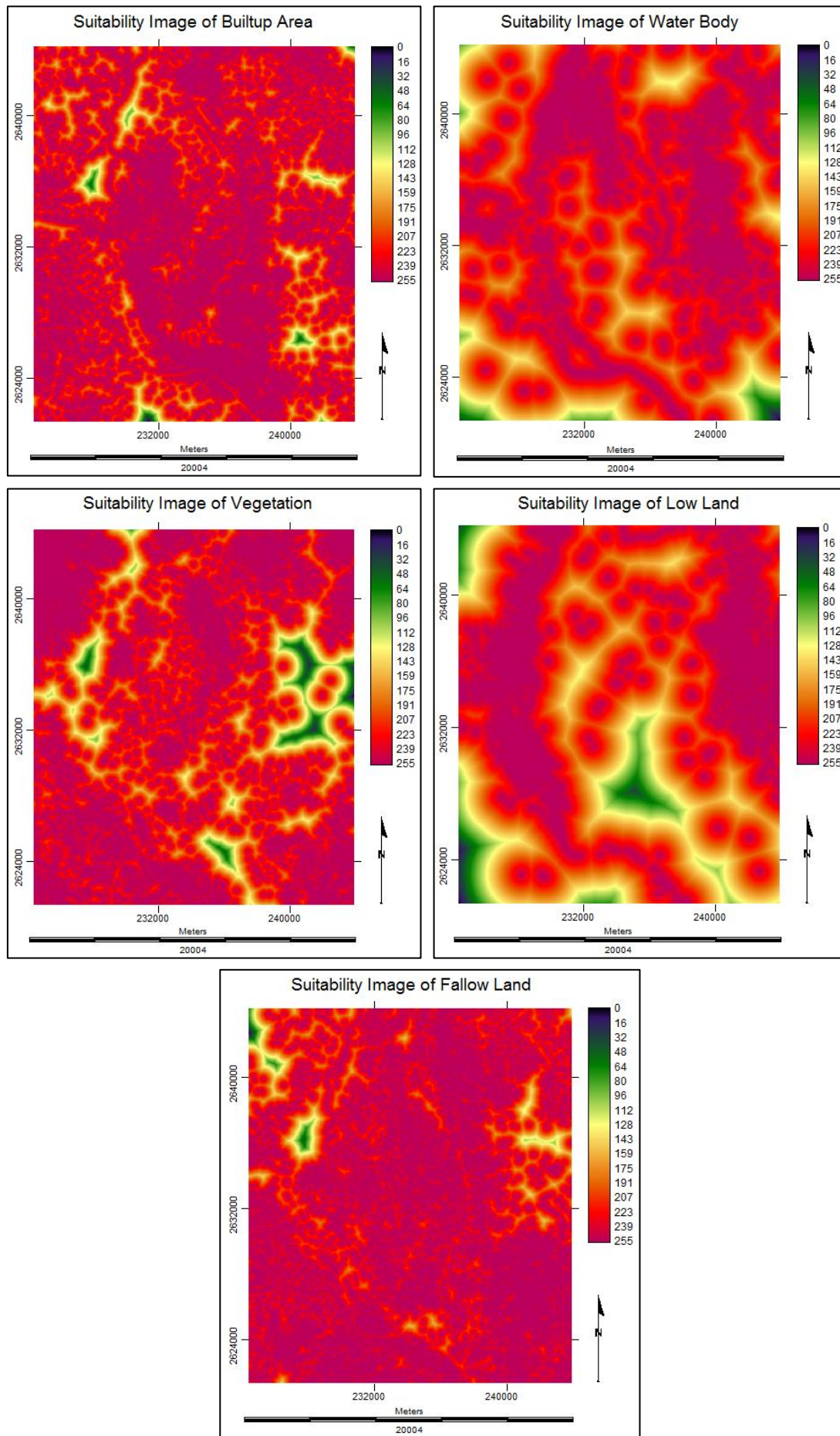


Figure 6.8: Distance Images of each Land Cover Type (1999)

[Unit: Meter]



**Figure 6.9: Suitability Images of each Land Cover Type**

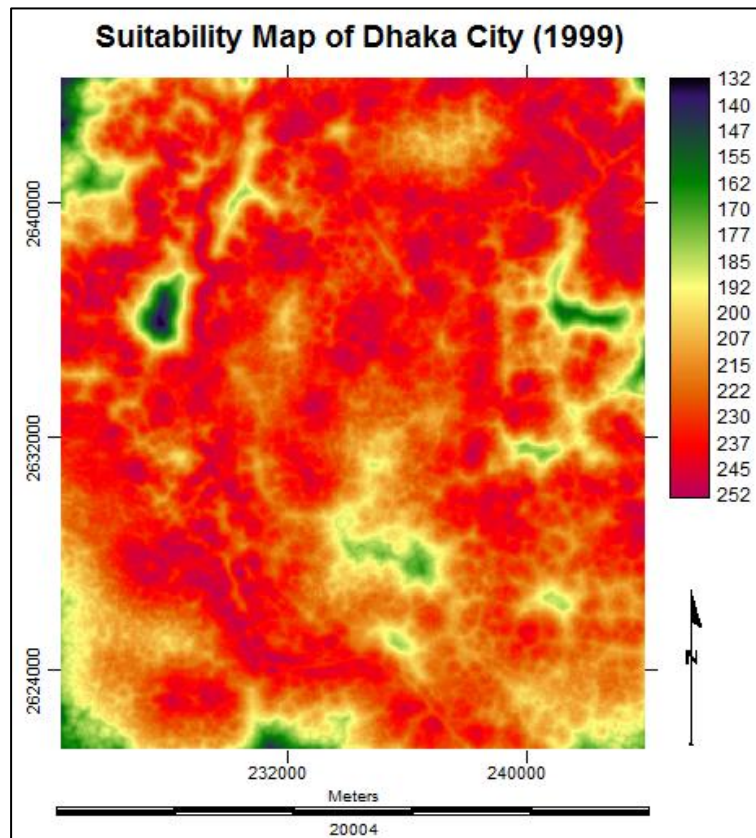


All the 5 images of Figure 6.9 are known as ‘Factor Images’. Now it’s the turn to assign weights for all the factor images. Factor weights are assigned to specify the relative importance of each factor in determining the aggregate output value [37].

**Table 6.1: The Factor Weights Evaluated for the Suitability Map (1999)**

Factor Name	Factor Weight
Suitability Image of Builtup Area	0.2000
Suitability Image of Water Body	0.2000
Suitability Image of Vegetation	0.2000
Suitability Image of Low Land	0.2000
Suitability Image of Fallow Land	0.2000
<b>Total Weight = 1.0000</b>	

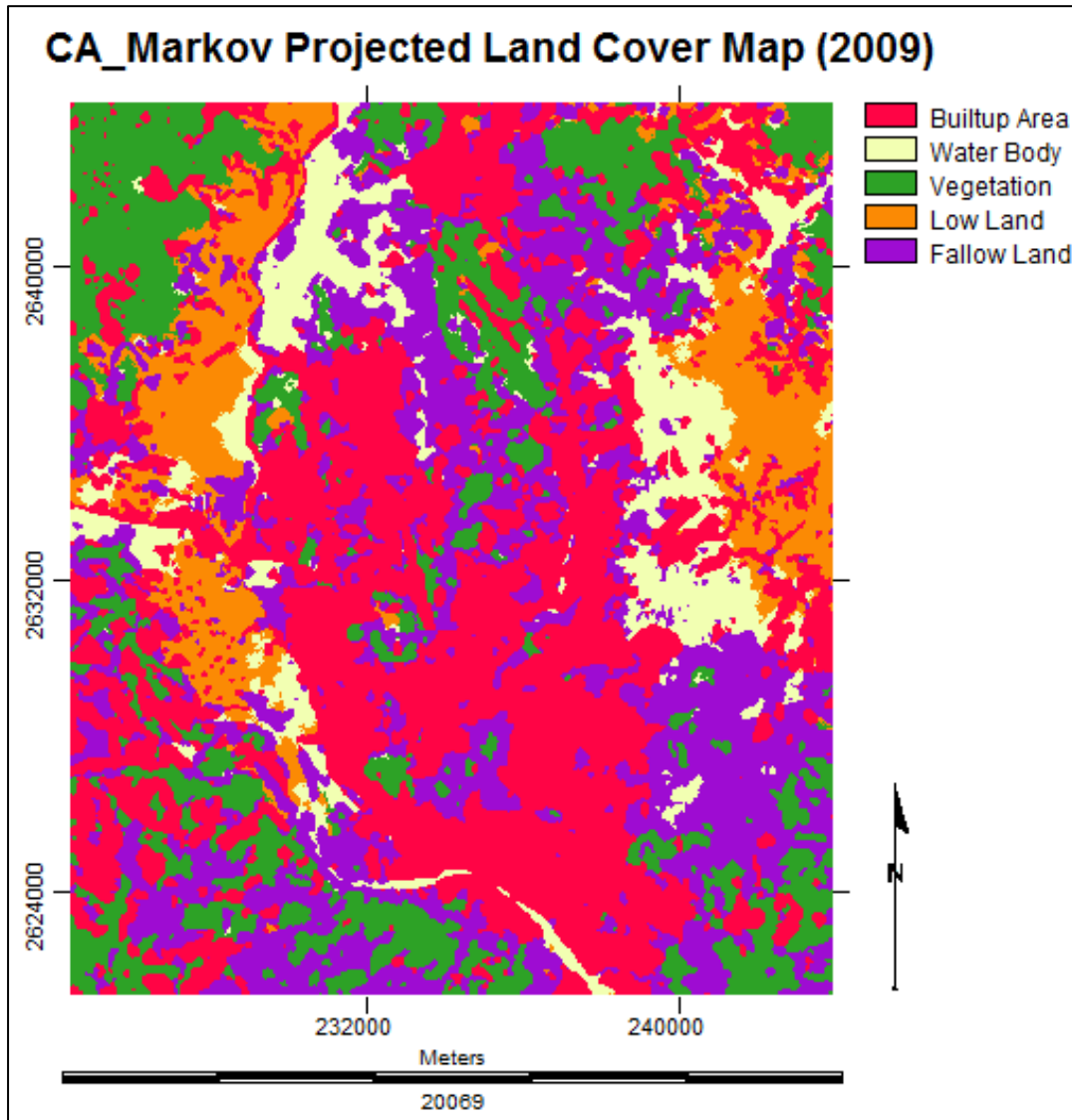
For this research purpose, equal weights have been assigned to all the factor images (Table 6.1). The reason behind this is that all the land cover types are of same importance for predicting the future. This is how; the aggregated multi-criteria resulted image has been created (Figure 6.10) using Weighed Linear Combination (WLC).



**Figure 6.10: Aggregated Land Cover Suitability Map of Dhaka City (1999)**

### **6.3.4 Future Prediction**

At the end, the Markov transition area matrix (Table 5.2), all the suitability images (Figure 6.9), the 3×3 CA contiguity filter and the base map of Dhaka city (1999) have been used to predict the land cover map of 2009. The CA\_Markov predicted final land cover image (2009) of Dhaka city is illustrated in Figure 6.11.



**Figure 6.11: CA\_Markov Projected Land Cover Map of Dhaka City (2009)**

## Chapter 7

### Multi Layer Perceptron Markov Model

The basic concept of multi layer perceptron neural network and the modelling technique using this concept has been depicted in this chapter.

#### 7.1 Artificial Neural Network

The term ‘Artificial Neural Network (ANN)’ has been inspired by human biological nervous system [68]. It is a kind of conversation from human neuron to artificial neurons. Artificial Neural Networks are models that try to simulate and parallel the decision making process of the human brain [69]. This is why the basic concept is termed as ‘Artificial Neural Network (ANN)’. However, in terms of human brain ANN is far simpler [70]. The definitions of some terminologies used for this chapter have been described in Appendix H.

#### 7.2 Basic Concept of Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is an interconnected group of nodes. In a typical ANN model, simple nodes (also known as ‘neurons’) are connected together to form a network of nodes. Some of these nodes are called input nodes; some are output nodes and in between there are hidden nodes.

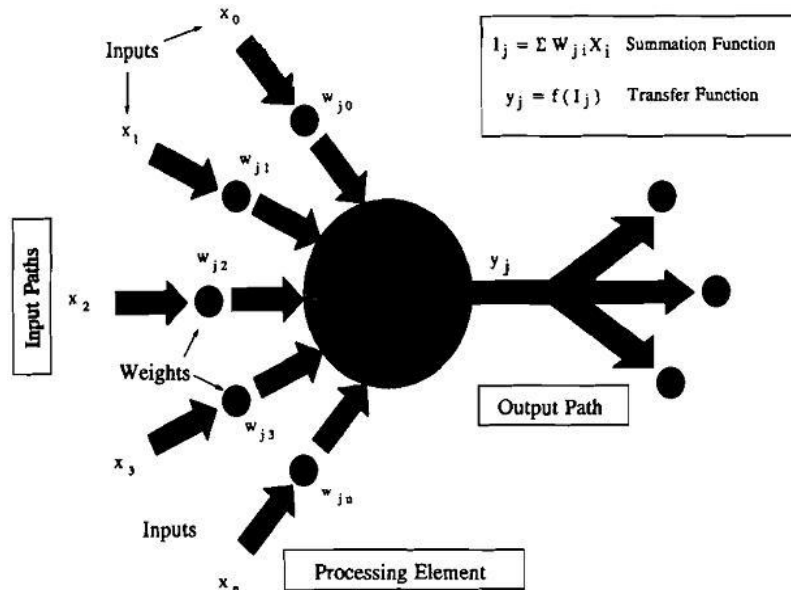


Figure 7.1: Artificial Neural Network Processing Element [69]

In a typical Artificial Neural Network, the processing element has many input paths like the dendrites of human brain (Figure 7.1). The information transferred through these paths by a variety of mathematical functions (e.g. simple summation). The input paths are weighted ( $W_{ij}$ ) like the synaptic strength of neural connections. The result of these combined inputs is known as Internal Activity (I) that is then used for the receiving processing element.

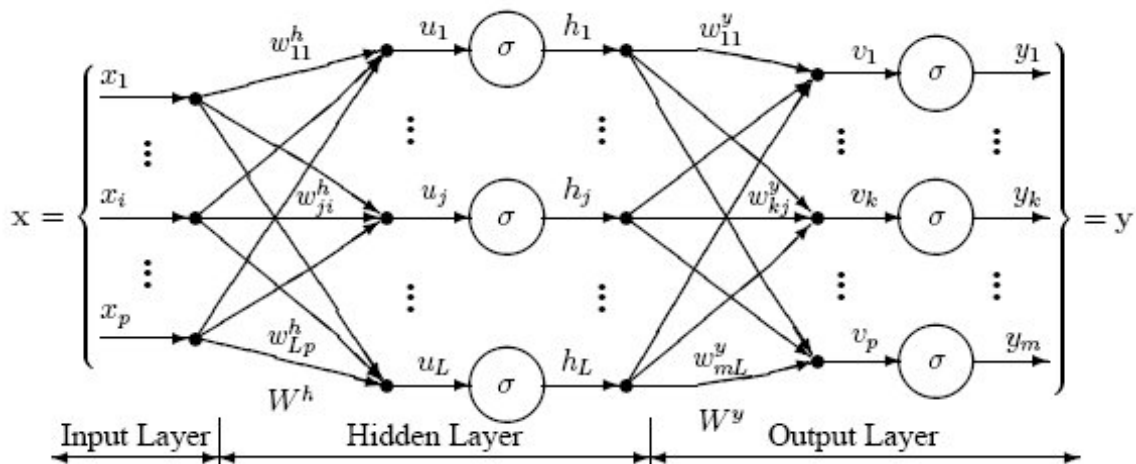
This internal activity is then modified by a transfer function (f) before being passed to other connected processing elements. This transfer function is commonly either a sigmoid or a hyperbolic tangent function [69].

### 7.2.1 Types of Artificial Neural Network

There are many types of ANN like Single Layer Perceptron, Multi Layer Perceptron, Recurrent Network, Hopfield Network, Spiking Neural Network and many more [70].

### 7.3 Multi Layer Perceptron (MLP)

Multi Layer Perceptron (MLP) is a feed-forward Neural Network with one or more layers between input and output layers. Feed-forward means that data flows in one direction from input to output layer (forward) [66].



**Figure 7.2: A Multi Layer Perceptron Neural Network Model [72]**

An example of MLP is illustrated in Figure 7.2 [72]. It consists of an input layer, one hidden layer and an output layer; all with three neurons. This kind of three layer network is also known as ‘Back Propagation Network’ [66].

### **7.3.1 Input Layer**

Here the nodes are the elements of a feature vector. This vector might be the wavebands of a data set or the texture of an image [73]. This can also be ancillary data like slope, elevation, soil type, zoning etc [69]. A vector of predictive variable values ( $x_1, \dots, x_p$ ) is presented in Figure 7.2. The  $x$  values are the inputs of the network.

### **7.3.2 Hidden Layer**

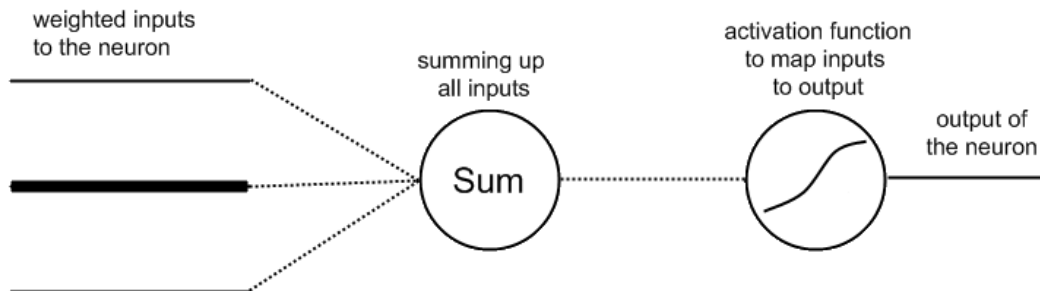
The second layer is an internal or hidden layer. It does not contain output units. One hidden layer can represent any Boolean function [73]. The activity of each hidden unit is determined by the activities of the input units and the weights of the connections between the input and the hidden units [68]. After arriving to the hidden layer, the value from each input neuron is multiplied by a weight ( $W_{ji}$ ). The resulting weighted values are added together producing a combined value  $u_j$ . The weighted sum ( $u_j$ ) is then modified with a transfer function ( $\sigma$ ) that outputs a value  $h_j$ . The outputs from the hidden layer are distributed to the output layer [72].

### **7.3.3 Output Layer**

The output layer presents the output data. For image classification, the number of nodes in the output layer equals to the classes in the classification [73].

After arriving to the output layer, the value from each hidden layer neuron is again multiplied by a weight ( $W_{kj}$ ). The resulting weighted values are added together producing a combined value  $v_j$ . The weighted sum ( $v_j$ ) is then modified with a transfer function ( $\sigma$ ) that outputs a value  $y_k$ . The  $y$  values are the outputs of the network [72].

Therefore, each neuron works as shown in Figure 7.3 [74].



**Figure 7.3: Working Methodology of a Neuron in MLP Network [74]**

### **7.3.4 The Feed-Forward Concept of Multi Layer Perceptron Neural Network**

MLP neural network uses the back propagation (BP) algorithm. The calculation is based on information from training sites [66]. Back propagation involves two major steps, forward and backward propagation. The input that a single node receives is weighted as:

$$net_j = \sum \omega_{ji} O_i$$

Where,  $w_{ij}$  = the weights between node i and node j;  $O_i$  = the output from the node i

The output from a given node j is computed as [73]:

$$O_i = f (net_j)$$

f = a non-linear sigmoid function that is applied to the weighted sum of inputs before the signal passes to the next layer

This is known as 'Forward Propagation'. Once it is finished, the activities of the output nodes are compared with their expected activities. In normal circumstances, the network output differs from the desired output (a set of training data, e.g. known classes). The difference is termed as the error in the network [73]. The error is then back-propagated through the network. Now the weights of the connections are corrected as follows [66]:

$$\Delta\omega_{ji}(n + 1) = \eta(\delta_j O_i) + \alpha\Delta\omega_{ji}(n)$$

$\eta$  = the learning rate parameter

$\delta_j$  = an index of the rate of change of the error

$\alpha$  = the momentum parameter

The process of the forward and backward propagation is repeated iteratively, until the errors of the network minimized or reaches an acceptable magnitude [73]. The purpose of training the network is to get proper weights both for the connection between the input and hidden layer, and between the hidden and the output layer for the classification of unknown pixels [66].

Several factors affect the capabilities of the neural network to generalize [73]. These include:



### **7.3.5 Number of Nodes**

In general, the larger the number of nodes in the hidden layer, the better the neural network represents the training data [73]. The number of hidden layer nodes is estimated by the following equation [66]:

$$N_h = INT (\sqrt{N_i * N_o})$$

Where,  $N_h$  = the number of hidden nodes

$N_i$  = the number of input nodes

$N_o$  = the number of output nodes

### **7.3.6 Number of Training Samples and Iterations**

The number of training sample also affects the training accuracy. Too few samples may not represent the pattern of each category while too many samples may cause overlap.

Again too many iterations can cause over training that may cause poor generalization of the network [66]. Over training can be prevented by early stopping of training [68].

The acceptable error rate is evaluated based on the Root Mean Square (RMS) Error [68]:

$$RMS = \frac{\sum(e_i)^2}{N} = \frac{\sum(t_i - a_i)^2}{N}$$

Where,  $N$  = the number of elements

$i$  = the index for elements

$e_i$  = the error of the  $i^{\text{th}}$  element

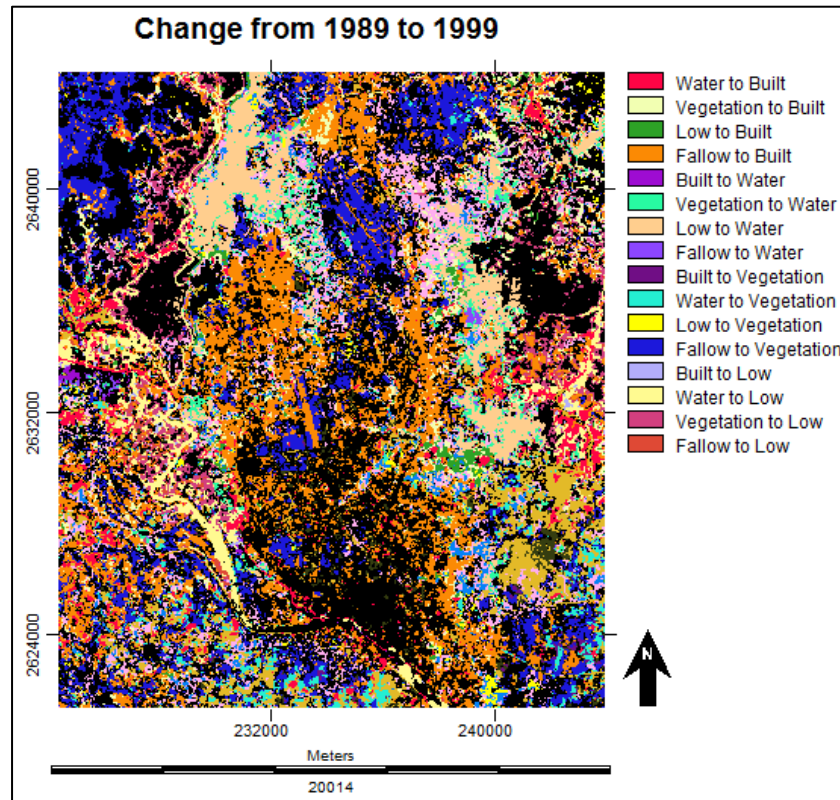
$t_i$  = the target value (measured) for  $i^{\text{th}}$  element

$a_i$  = the calculated value for the  $i^{\text{th}}$  element

Till now, the theories and mathematical notations behind multi layer perceptron neural network have been explained. In the next section, how all these techniques have been incorporated, to predict the land cover map (2009) of Dhaka city, is described in detail.

## **7.4 Multi Layer Perceptron Markov Modelling**

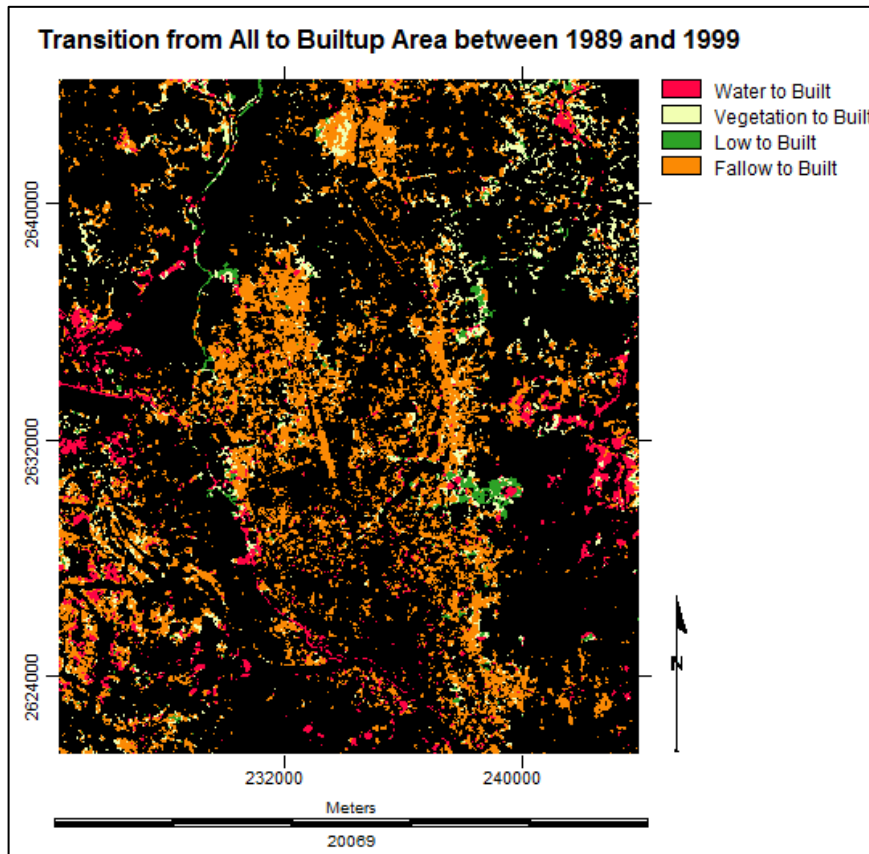
The change/ transition among the land covers is an important aspect for modelling the future land cover change. Preparing a potential map of land to go through transitions is important to predict the land cover change.



**Figure 7.4: Transitions of Land Covers between 1989 and 1999**

Figure 7.4 illustrates the change among land covers over the years. But considering all the transitions is not necessary. The basic concept of modelling with MLP neural network, for this research, is to consider the change in builtup area over the years. It is already proved that builtup area has been increasing at a high rate, while other land covers are decreasing. In general, it means other land cover types are primarily contributing to increase the builtup area.

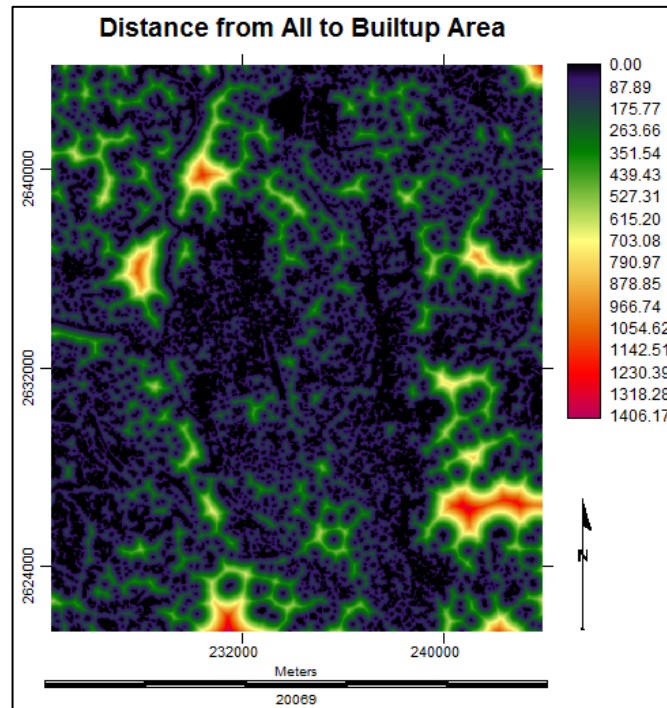
Therefore, only the transitions from ‘water body to builtup area’, ‘vegetation to builtup area’, ‘low land to builtup area’ and ‘fallow land to builtup area’ have been considered for model simulation. These four transitions have been termed as “All” here. Figure 7.5 exhibits the transition from all to builtup area.



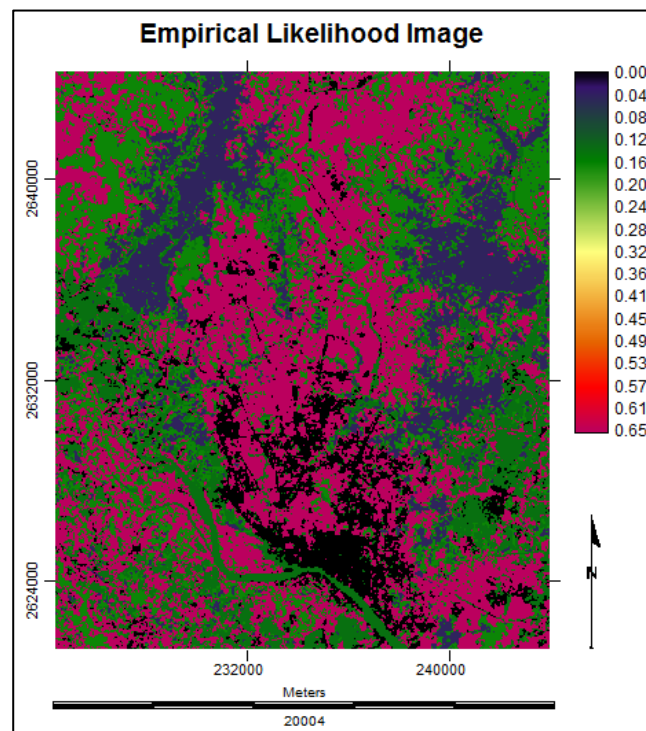
**Figure 7.5: Transition from All to Builtup Area (1989-1999)**

To predict the change, it is important to empirically model each of these four transitions. These maps are called transition potential maps. The great advantage of using MLP perceptron neural network is that it gives the opportunity to model several or even all the transitions at once [66]. As the driving force (change to builtup area) for all these transitions are the same here, therefore it is possible to create a common group of explanatory variables that can adequately model all the transitions.

At this stage, the issue of which variables affect the change to builtup area (1989-1999) has been considered. Therefore, distance from all to builtup area has been chosen (Figure 7.6). It is logical that new areas will be converted to builtup area where there are existing builtup areas. Other variables that can put impact on this change are also considered like images of distance from water body, vegetation, low land and fallow land (Figure 6.8). Another important aspect that might be useful for future land cover prediction is finding out the likelihood of all changes for transforming into builtup areas based on the base image of 1989 (Figure 7.7).



**Figure 7.6: Distance Image of Transition from All to Builtup Area (1989-1999)**



**Figure 7.7: Empirical Likelihood Image (1989-1999)**

This kind of empirical likelihood transformation is an effective means of incorporating categorical variables into the analysis (Figure 7.7). It has been produced by determining the relative frequency of different land cover types occurred within areas of transition (1989 to 1999). The numbers (legend) indicate the likelihood of changing into builtup area. The higher the value the likeliness of the pixel to change into builtup area is more.

### **7.4.1 Testing Potential Explanatory Power**

Now there are 6 driver variables for the model (Distance from all to builtup area, distance from water body, distance from vegetation, distance from low land, distance from fallow land and empirical likelihood image). It is important to test the potential explanatory power of each variable.

The driver variable test procedure is based on contingency table analysis [66]. The quantitative measures of the variables have been tested through Cramer's V. A high Cramer's V indicates that the potential explanatory value of the variable is good. It is suggested that the variables that have a Cramer's V of about 0.15 or higher are useful while those with values of 0.4 or higher are good [66]. A variable can be discarded if the Cramer's V is low.

**Table 7.1: Cramer's V of the Driving Factors**

<b>Driving Variable / Cover Class</b>	<b>Overall V</b>	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
Distance from All to Builtup Area	0.1143	0.0000	0.7848	0.2185	0.2886	0.2024
Distance from Water Body	0.1254	0.0000	0.2352	0.8467	0.2701	0.1512
Distance from Vegetation	0.1601	0.0000	0.3521	0.2545	1.0000	0.4408
Distance from Low Land	0.1303	0.0000	0.2169	0.2134	0.2320	0.8862
Distance from Fallow Land	0.1532	0.0000	0.4547	0.2226	0.3449	0.4399
Empirical Likelihood Image	0.2527	0.0000	0.4189	0.3926	0.3128	0.3999

Table 7.1 shows that the potential explanatory values of the driving variables are useful (Cramer's V > 0.15) and in most cases it is good (Cramer's V > 0.40). Though the overall Cramer's V is low but the individual class values are showing higher values. The overall Cramer's V is showing low values because the values for builtup area are 0. As the change has been assumed to be transition from all to builtup areas, therefore, the Cramer's V for builtup area is the lowest value 0. Considering this situation, the overall Cramer's V is quite acceptable for all the 6 driving variables chosen for MLP model.

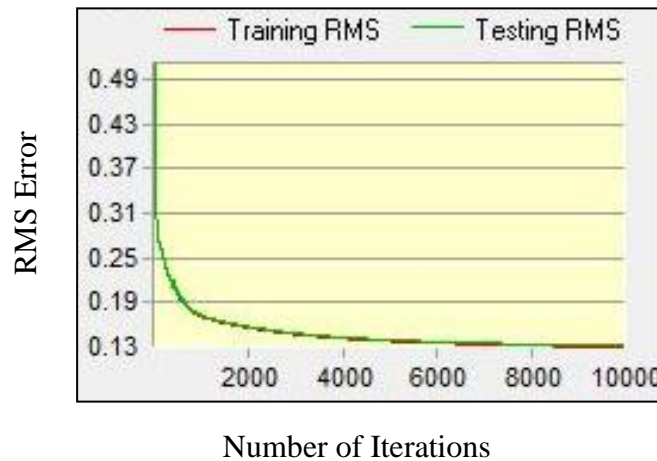
### 7.4.2 Transition Potential Modelling

After getting satisfactory Cramer's V values for all the driving variables, now the turn is to run MLP neural network model. For each principal transition particular weights have to be obtained. The MLP running statistics are shown in Table 7.2.

**Table 7.2: Running Statistics of MLP Neural Network**

Maximum Sample Size	4794
Iterations	10000
Training RMS	0.1323
Testing RMS	0.1341
Accuracy Rate	<b>91.36%</b>

The MLP running statistics gives a very high accuracy rate of 91.36% (Table 7.2). The minimum number of cells that transitioned from 1989 to 1999 is 4794. Therefore, the maximum sample size has been chosen as 4794. The RMS error curve has been found smooth and descent after running MLP neural network (Figure 7.8).



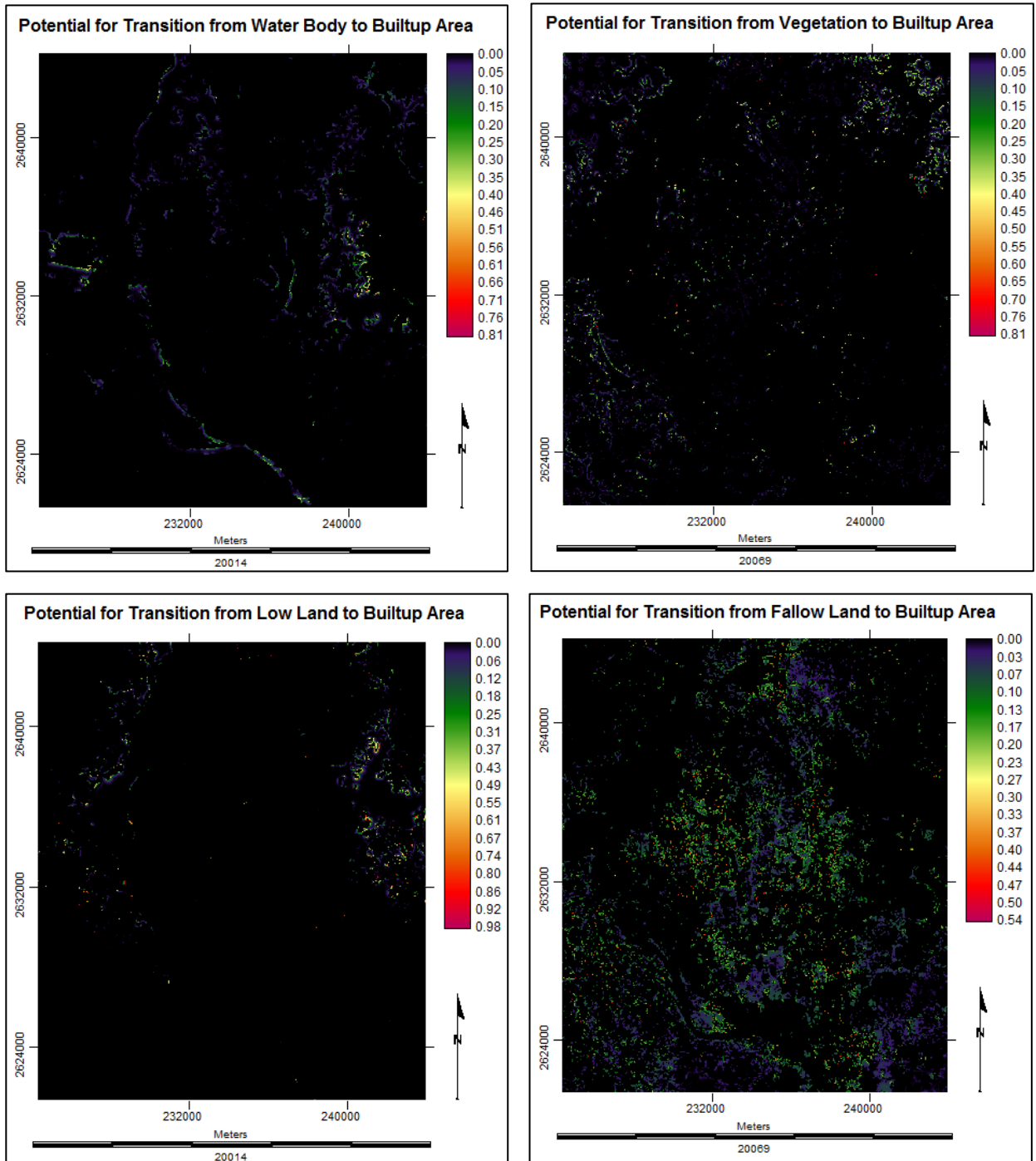
**Figure 7.8: RMS Error Monitoring Curve**

Both the training and testing RMS curves are showing progressive decrease in the RMS errors meaning substantial increase in accuracy rate (Figure 7.8). It means the training results are satisfactory. This is why; high accuracy rate (91.36%) has been achieved. This accuracy rate is based on the sampling specifications for the training and testing pixels per category [66].

Based on these running statistics four transition potential maps have been produced (Figure 7.9). These maps depict, for each location, the potential it has for each of the modelled transitions [37]. These are not the probability maps where the sum of values



for a particular pixel location will not be 1. The reason behind this is because the MLP neural network outputs are obtained by applying fuzzy set to the signals into values from 0-1 with activation function (sigmoid). Here the higher values represent a higher degree of membership for that corresponding land cover type [66].



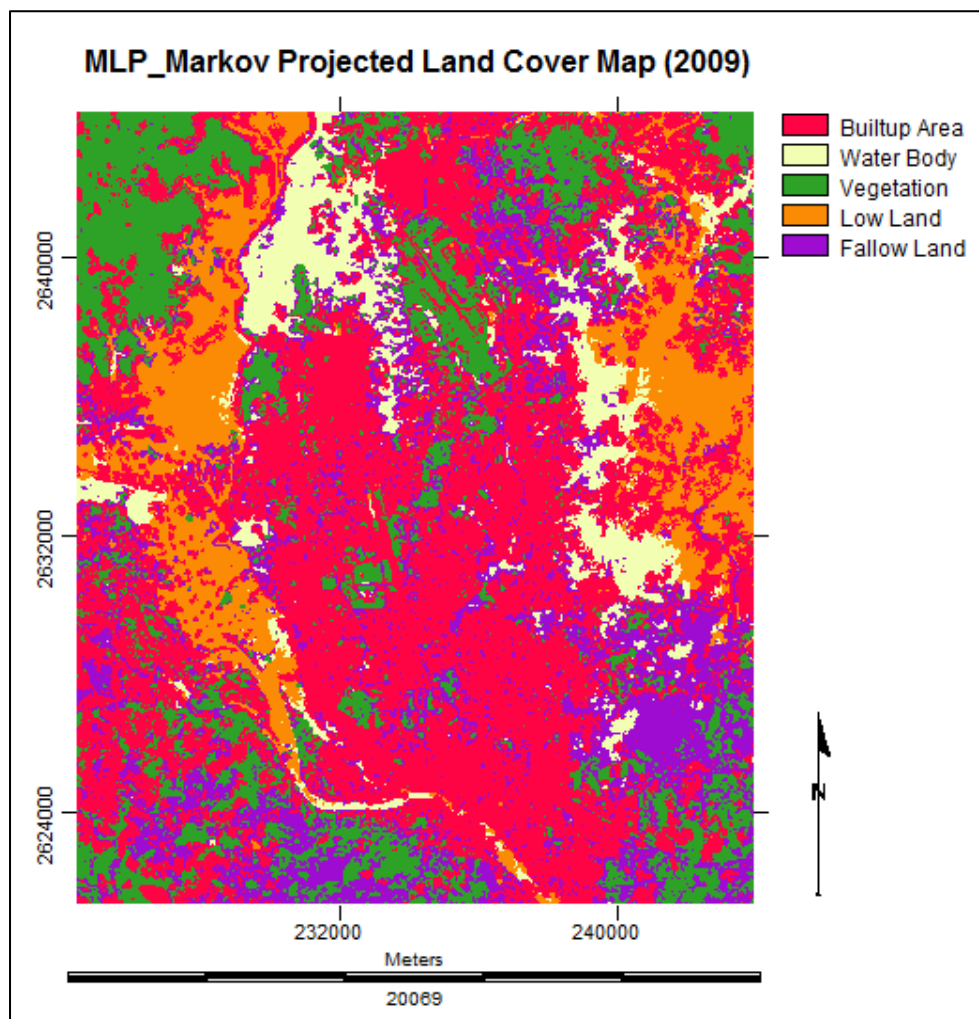
**Figure 7.9: Transition Potential Maps**

### 7.4.3 Future Prediction

Using this MLP neural network analysis it is possible to determine the weights of the transitions that will be included in the matrix of probabilities of Markov Chain for future prediction. The transition probabilities are shown in Table 7.3. Based on all these information from MLP neural network, the final land cover map of 2009 (Figure 7.10) has been simulated through Markov chain analysis. The whole procedure, for predicting the land cover map by this way, has been termed as ‘MLP\_Markov’ model.

**Table 7.3: Transition Probabilities Grid for Markov Chain**

	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>Builtup Area</b>	0.7823	0.0174	0.0347	0.0194	0.1463
<b>Water Body</b>	0.2079	0.1264	0.1008	0.1927	0.3721
<b>Vegetation</b>	0.1529	0.0779	0.3887	0.1071	0.2734
<b>Low Land</b>	0.0695	0.3634	0.0467	0.4054	0.1150
<b>Fallow Land</b>	0.3825	0.0133	0.2413	0.0185	0.3445



**Figure 7.10: MLP\_Markov Projected Land Cover Map of Dhaka City (2009)**



## Chapter 8

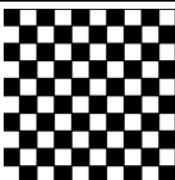
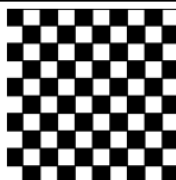
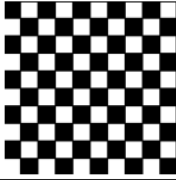
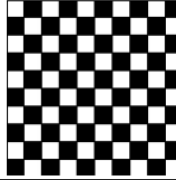
### Model Validation and Selection

---

This deals with the model validation techniques and selecting the best fitted model.

#### 8.1 Model Validation

Model validation refers to comparing the simulated and reference maps. Sometimes the simulated maps can give misleading results. In that case, it is necessary to validate the projected/simulated map with the base/reference map. Therefore, model validation is an important step in case of predictive change modelling [36]. Often it is termed as ‘Map Comparison’. Map comparison is vital in evaluating the analytical techniques for spatial data. It produces numerical expressions to compare two maps [75]. An example of map comparison is shown in Figure 8.1 [75].

Map 1	Map 2	Cell-by-cell comparison
		100 %
		0%

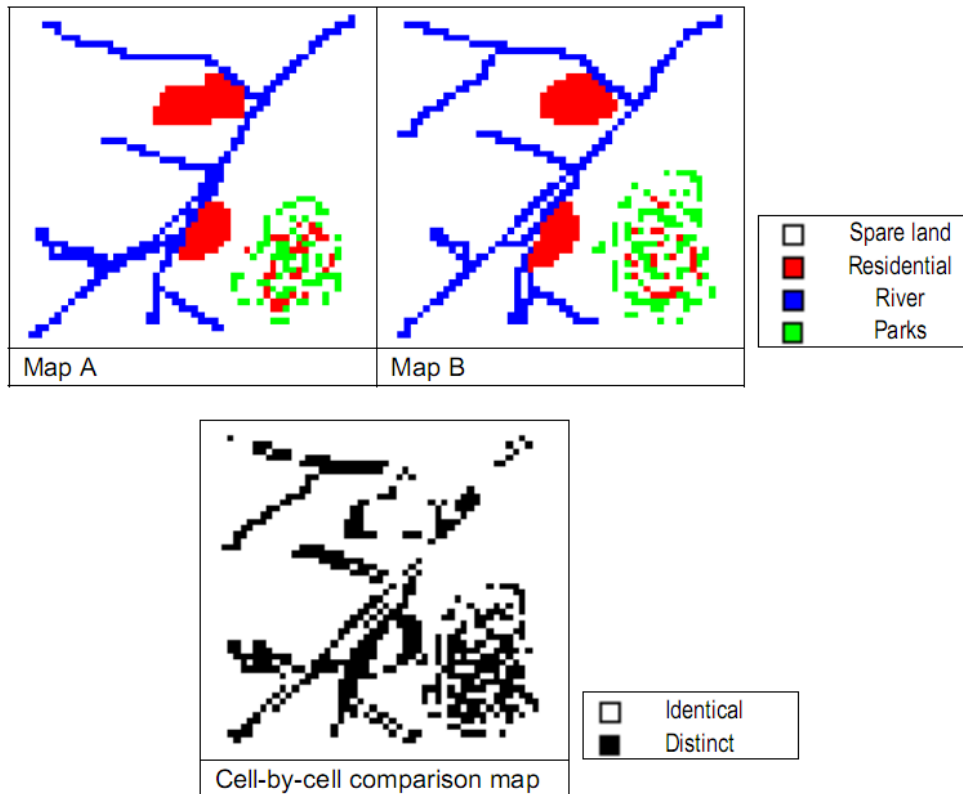
**Figure 8.1: Cell-by-cell Map Comparison (left and right maps are compared) [75]**

Here the cell in the upper left corner (Map 1) that is black; matches with the same cell position of Map 2 (Figure 8.1). This is applicable for all the cells between these two maps. Therefore the average observers will conclude the decision: ‘Map 1’ and ‘Map 2’ are similar (100% agreement) in the first case. The situation is opposite in the second case. In case of land cover change prediction, model validation is a challenge. Because there is no agreed criteria or method to assess the performance of the method applied for modelling purpose. The best way to validate a model is to compare the predicted map of time  $t_2$  with the reference map of time  $t_2$  [76].

The different methods that have been used to validate the model in this research have been described in the following sections.

### **8.1.1 Per Category Method**

The ‘Per Category Method’ performs cell-by-cell comparison for a particular user defined category [75]. In case of cell-by-cell comparison, the output is a binary map that denotes for each cell whether the maps are identical or distinct on that location [43].

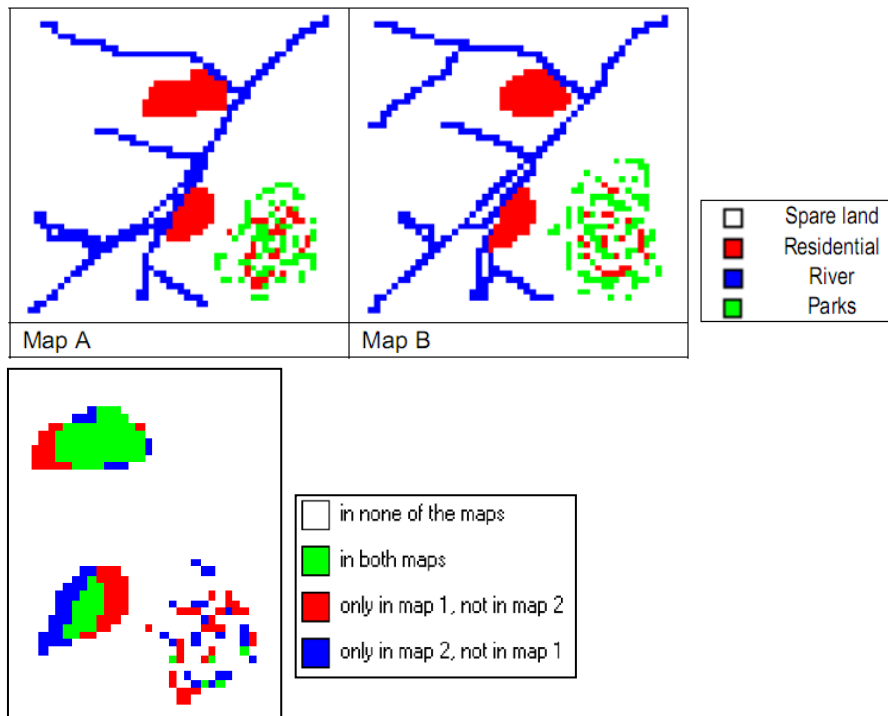


**Figure 8.2: Cell-by-cell Map Comparison of Two Maps [75]**

Figure 8.2 is a typical example of cell-by-cell comparison. An informative enhancement of this cell-by-cell comparison is ‘Per Category Method’. The result of ‘Per Category Method’ gives the output with the following legends [43]:

- Category present in both Maps
- Category present in neither of the Maps
- Category present in Map A not in Map B
- Category present in Map B not in Map A

Figure 8.3 illustrates an example of per category method [75].



**Figure 8.3: Per Category Map Comparison of Two Maps [75]**

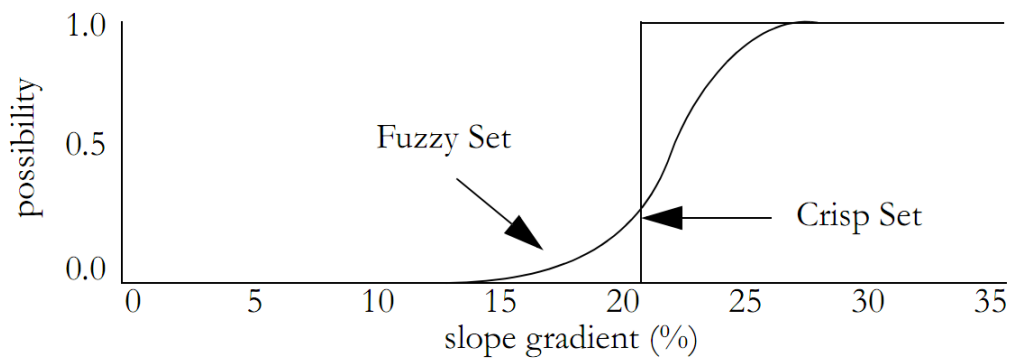
### **8.1.2 Fraction Correct**

Fraction Correct is calculated as the number of equal cells divided by the total number of cells [75]. It represents the percentage (%) of correct cells.

### **8.1.3 Fuzzy Sets**

“A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by membership (characteristic) function which assigns to each object a grade of membership ranking between zero and one” [77]. The ranges from 0.0 to 1.0, indicates a continuous increase from non-membership to complete membership function [36].

An example is given in Figure 8.4 to understand fuzzy sets [36]. The figure evaluates the steepness of a slope. It is assumed that a slope of 10% has a membership of 0.0, while a slope of 25% has a membership of 1.0. Therefore, in between 10% to 25%, the fuzzy membership of a slope gradually increases on a scale from 0.0 to 1.0. This figure also includes the classic ‘Crisp Set’ that has distinct boundaries. One example of crisp set membership function is ‘Boolean Constraints’.



**Figure 8.4: Fuzzy vs. Crisp Set Membership Functions [36]**

### **8.1.4 Fuzzy Kappa**

Fuzzy kappa map comparison shows the grades of similarity between pairs of cells.

The main difference with the cell-by-cell map comparison is that fuzzy kappa map comparison takes into account the neighbourhood of a cell [75]. Then it represents the cell values between 0.00 (fully distinct) to 1.0 (fully identical).

The statistic of fuzzy kappa is similar to Kappa. The main difference lies in the calculation of the expected similarity [78].

$$K_{Fuzzy} = \frac{P_o - P_e}{1 - P_e}$$

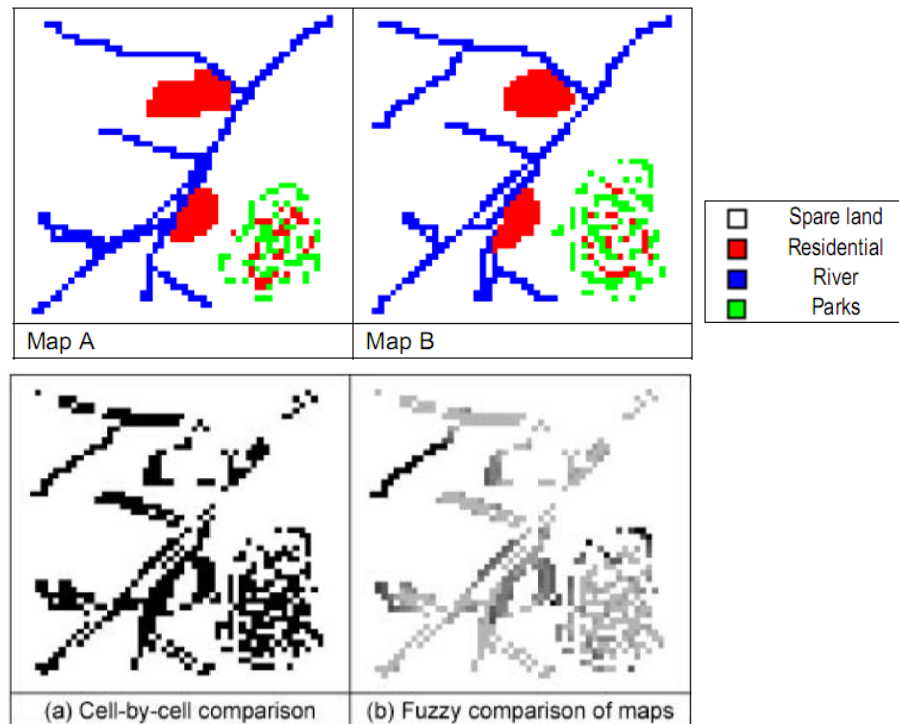
Where,  $K_{Fuzzy}$  = Fuzzy Kappa

$P_o$  = observed percentage of agreement (average similarity)

$P_e$  = expected similarity based upon given histograms

Figure 8.5 gives an example of fuzzy kappa [78].

[The detail description on Kappa Statistics has been explained in Appendix C].



**Figure 8.5: Difference between ‘Cell-by-cell’ and ‘Fuzzy’ Comparison of Maps [75]**

Figure 8.5 shows the difference between two methods. In cell-by-cell comparison method, the cells are in white when the categories are same in both maps while the cells are black where the categories differ. But in fuzzy comparison method, the lighter cells are more similar in both maps than the darker cells.

The main advantage of fuzzy set comparison is that the result gives more information and opportunity to analyze the situation in further details. For example, it is possible to distinguish between total agreement, medium similarity and low similarity (Figure 8.6) by creating a suitable range of classification of fuzzy set values [78].



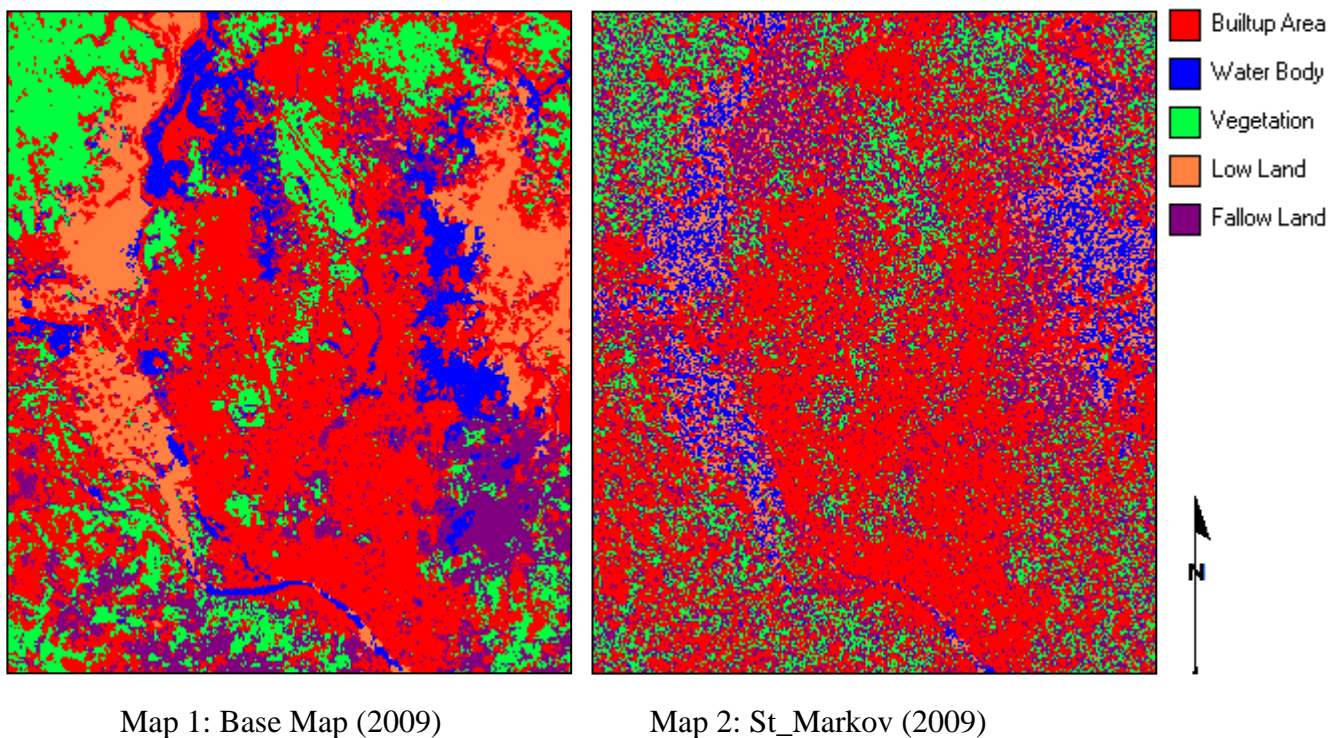
**Figure 8.6: Levels of Agreement for Fuzzy Comparison Method [78]**

## **8.2 Actual Base Maps vs. Simulated Maps**

In this section, the comparisons between the actual base map (2009) and the simulated maps (St\_Markov, CA\_Markov and MLP\_Markov) of year 2009 have been performed. This is known as model validation. The main objective of model validation is to find out whether the simulation is giving any abrupt results or not. This justifies the modelling output in terms of reality.

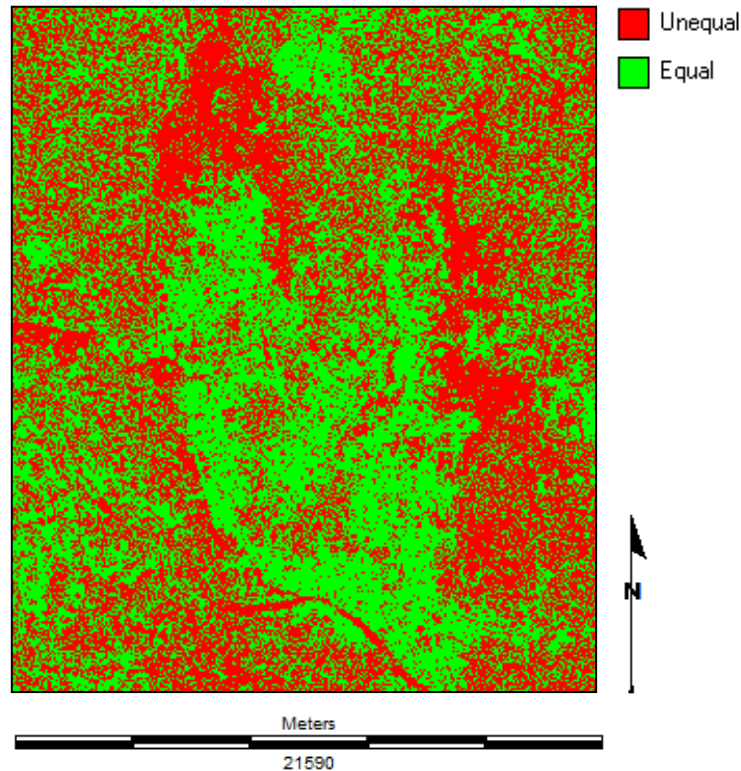
For validating the simulated maps, two different approaches have been adopted. The first one is visual approach. This approach helps to reveal the spatial patterns in a quick look. The visual approach is subjective. Another one is statistical approach. This approach is important because it explains the scenario in a quantitative way [76].

### **8.2.1 Base Map (2009) vs. St\_Markov (2009)**



**Figure 8.7: Maps for Model Validation**

Figure 8.7 shows the maps to be compared in this section for model validation purpose. Here map 1 represents the base map of 2009 and map 2 represents the St\_Markov simulated map of 2009.



**Figure 8.8: Levels of Agreement for Kappa**

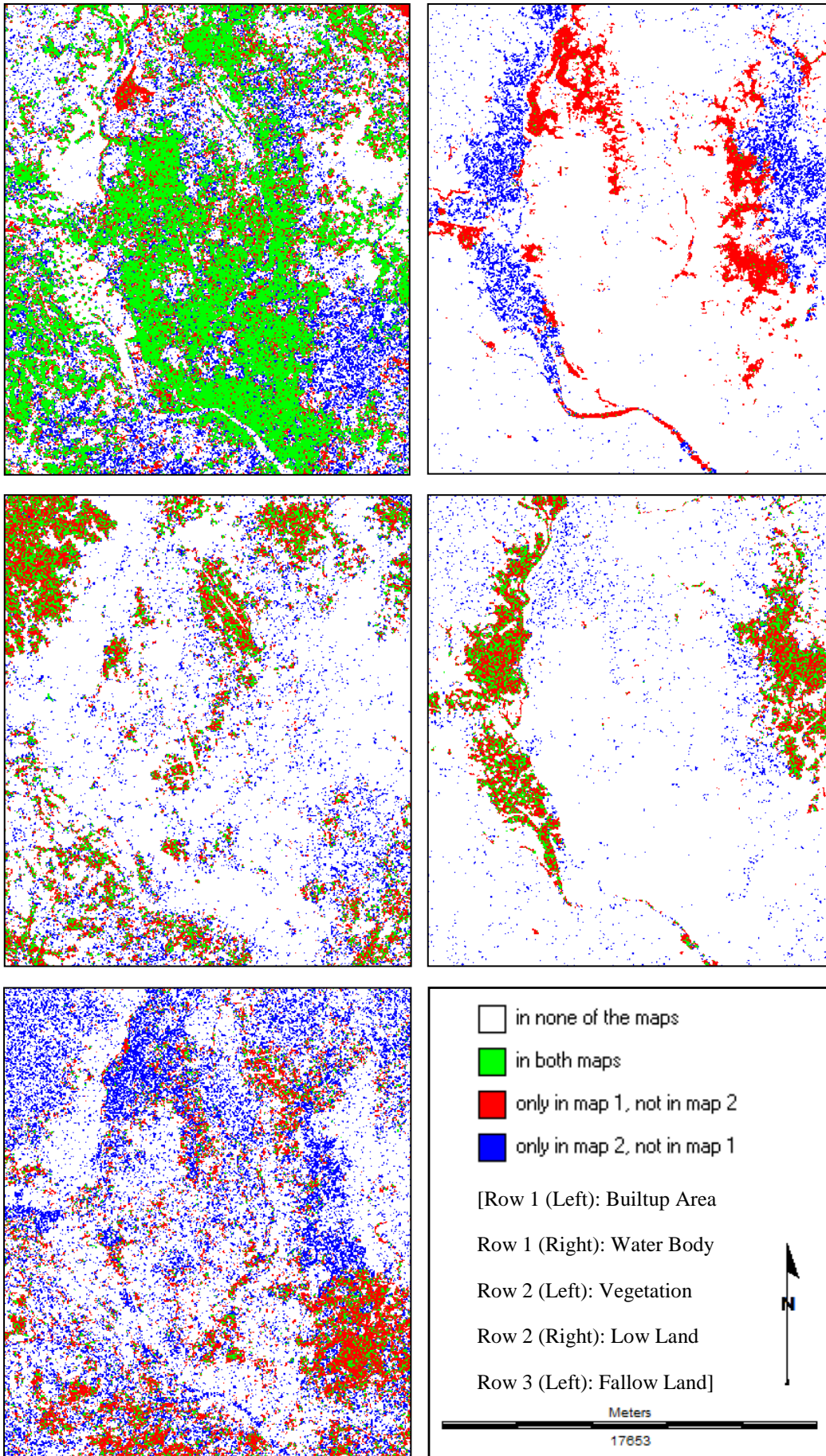
Figure 8.8 shows the cell-by-cell comparison of two maps for kappa statistic. It is clear that majority of the cells do not match each other. Therefore, the agreement is not good.

**Table 8.1: Per Category Land Cover Change**

Land Cover Type	Base Map (2009)		St_Markov (2009)		Change in Area (%)
	Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%	
<b>Builtup Area</b>	204.4008	45.853	210.3903	47.197	<b>+1.344</b>
<b>Water Body</b>	33.6645	7.552	30.0429	6.739	<b>-0.813</b>
<b>Vegetation</b>	79.9578	17.937	64.3068	14.426	<b>-3.511</b>
<b>Low Land</b>	49.9914	11.215	30.9870	6.952	<b>-4.263</b>
<b>Fallow Land</b>	77.7555	17.443	110.0430	24.686	<b>+7.243</b>
Total	<b>445.770</b>	<b>100</b>	<b>445.770</b>	<b>100</b>	<b>0</b>

Table 8.1 shows the per category change in area between the two maps. In the most ideal case, the change in area should be 0. The more the deviation, the less sophisticated is the model. In terms of area, it is clear that builtup area and water body remain almost the same. But fallow land has increased a lot, while vegetation and low land have decreased. It means vegetation, low and fallow land can show unexpected results in case of quantitative analysis.





**Figure 8.9: Per Category Comparison Method**



Figure 8.9 shows the method that performs cell-by-cell comparison for each land cover category. It is known as per category comparison. It gives information about the occurrences of the selected category in both maps [75]. The outputs are showing 4 legends indication 4 different state of comparison. The more there will be the amount of ‘both maps’, the better the simulation result.

In case of builtup area, the class ‘in both maps’ is higher. It means the agreement in cells for builtup area is quite good. On the other hand, in water body there is no sign of the class ‘in both maps’. This means there is no agreement in water body. In case of vegetation and low land the agreement among cells seems moderate. Fallow land shows high amount of ‘only in map 2, not in map 1’ and ‘only in map 1 and not in map 2’ classes. It means low degree of agreement exists in fallow land.

**Table 8.2: Per Category Kappa Statistics**

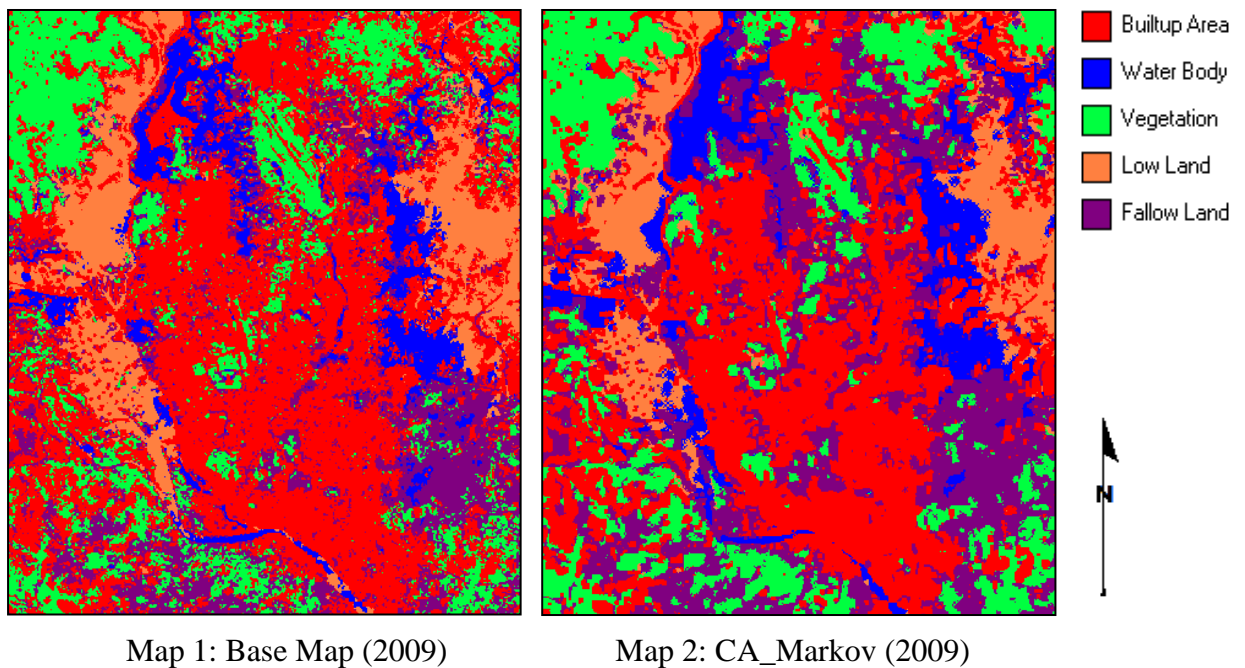
	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>Kappa</b>	0.47161	-0.03474	0.35805	0.43137	0.02969
<b>Klocation</b>	0.48470	-0.03700	0.41114	0.58036	0.03788
<b>Khisto</b>	0.97300	0.93879	0.87086	0.74328	0.78390

Table 8.2 demonstrates the per category kappa coefficients. Builtup area shows the highest kappa values. Water body is showing negative values, it means there is no effective agreement. Fallow land represents very low degree of agreement while vegetation and low land is moderate high.

#### **8.2.1.1 Analysis of the Results of St Markov**

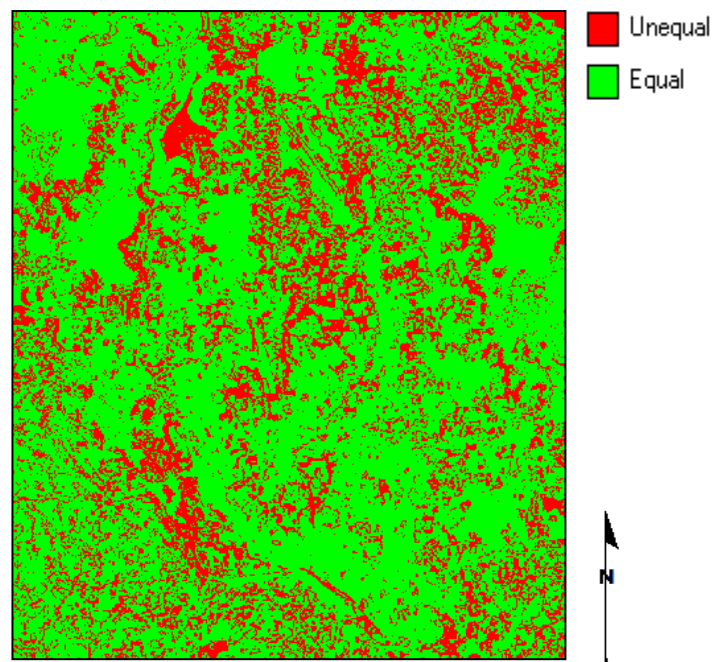
- a) The amount of areas (%) on a per category basis is almost accurate.
- b) There is no knowledge of spatial distribution of the occurrences within each category.
- c) This model lacks the knowledge of spatial dependency.

### 8.2.2 Base Map (2009) vs. CA Markov (2009)



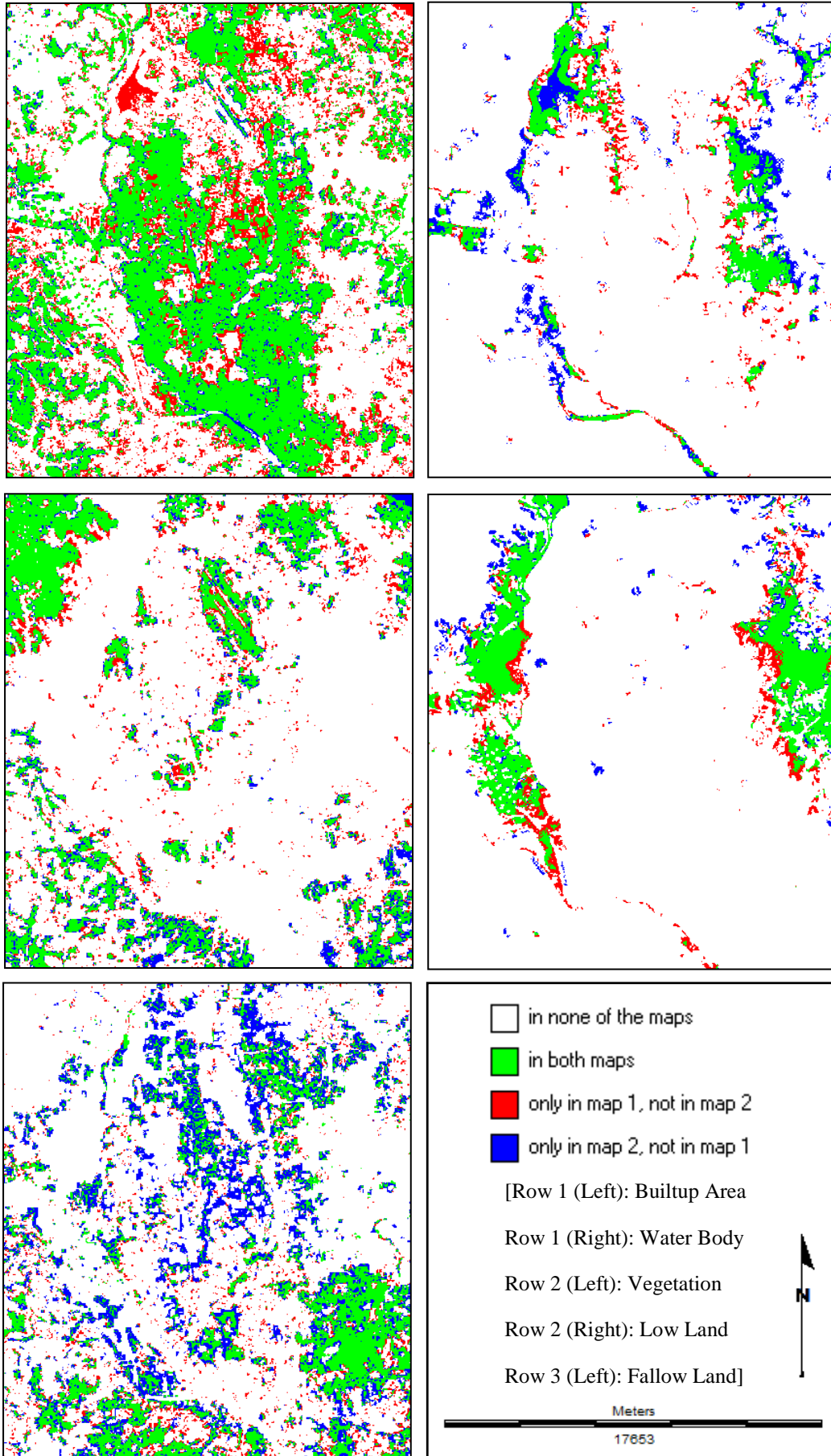
**Figure 8.10: Maps for Model Validation**

Figure 8.10 shows the maps to be compared. These are the base map of 2009 and the St\_Markov simulated map of 2009.



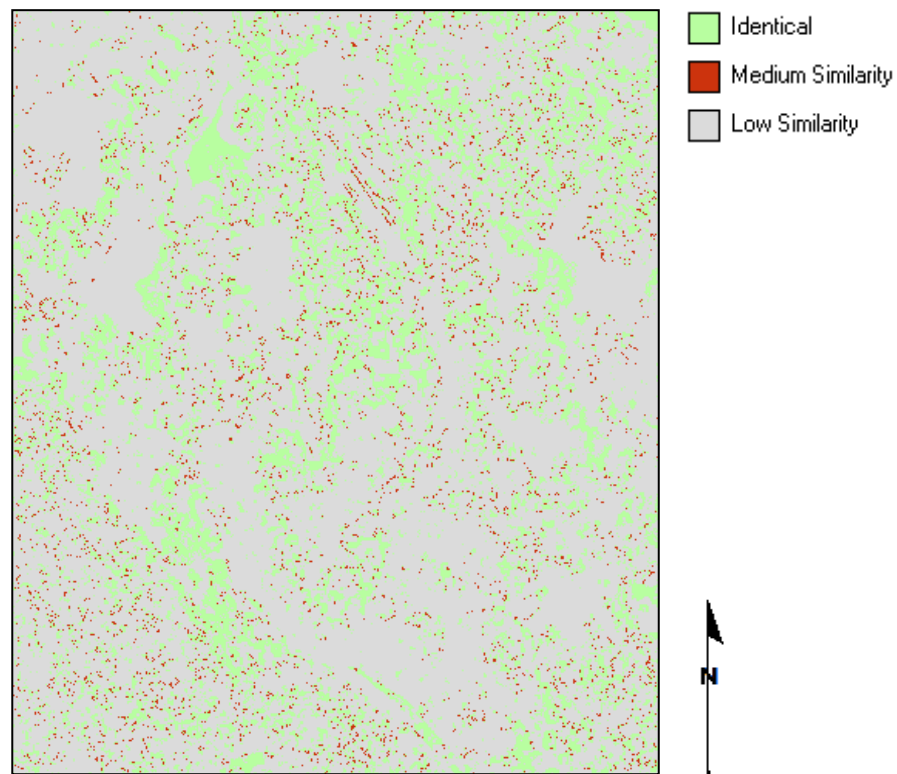
**Figure 8.11: Levels of Agreement for Kappa**

It is clear from Figure 8.11 that most of the cells match each other. Therefore, the levels of agreement among the cells are high.



**Figure 8.12: Per Category Comparison Method**

From Figure 8.12 it can be stated that builtup area is showing the dominance of ‘in both maps’ class. The second majority is ‘only in map 1, not in map 2’. It means builtup area has a moderately good agreement. Vegetation and low land are showing also showing high agreement. On the other hand, water body and fallow land having ‘only in map 2, not in map 1’ class in good amount. It means the agreement is lower than others.



**Figure 8.13: Fuzzy Kappa Result Map**

Figure 8.13 illustrates the fuzzy kappa classified comparison map. This figure has been generated using a 4×4 radius of neighbourhood with a linear decay function considering slope parameter value of 0.5. This gives more detail information about the grades of similarity between pairs of cells. It is clear that most cells have low similarity.

**Table 8.3: Fuzzy Kappa per Category**

<b>Fuzzy Kappa</b>	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
0.53851	0.52853	0.50481	0.59895	0.65579	0.30326

The fuzzy kappa for each category is showing moderate high values (Table 8.3). Fallow land category is showing the lowest value of all.

**Table 8.4: Per Category Land Cover Change**

Land Cover Type	Base Map (2009)		CA_Markov (2009)		Change in Area (%)
	Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%	
<b>Builtup Area</b>	204.4008	45.853	166.8402	37.427	<b>-8.426</b>
<b>Water Body</b>	33.6645	7.552	39.4074	8.841	<b>+1.289</b>
<b>Vegetation</b>	79.9578	17.937	78.0039	17.499	<b>-0.438</b>
<b>Low Land</b>	49.9914	11.215	44.3214	9.944	<b>-1.271</b>
<b>Fallow Land</b>	77.7555	17.443	117.1971	26.289	<b>+8.846</b>
Total	<b>445.770</b>	<b>100</b>	<b>445.770</b>	<b>100</b>	<b>0</b>

The CA\_Markov simulation shows that builtup area has decreased and fallow land has increased in good amount than in reality (Table 8.4). There is some dissimilarity. Rest land cover types are showing almost perfect match in terms of area.

**Table 8.5: Per Category Kappa Statistics**

	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>Kappa</b>	0.66079	0.57780	0.68680	0.72295	0.49241
<b>Klocation</b>	0.79816	0.63186	0.69728	0.77504	0.66185
<b>Khisto</b>	0.82789	0.91444	0.98497	0.93280	0.74400

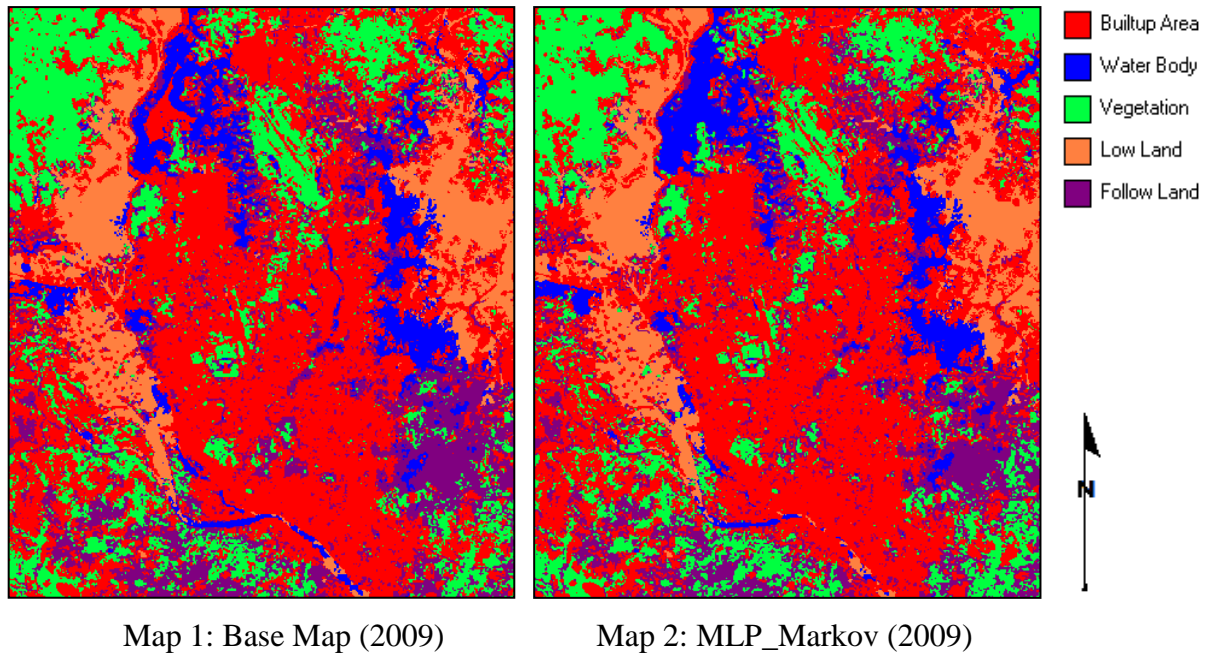
The kappa coefficients of per category are showing high values (Table 8.5). This proves that the agreement between the two maps is substantial. Only fallow land type is showing less value than the average of the other categories.

### **8.2.2.1 Analysis of the Results of CA\_Markov**

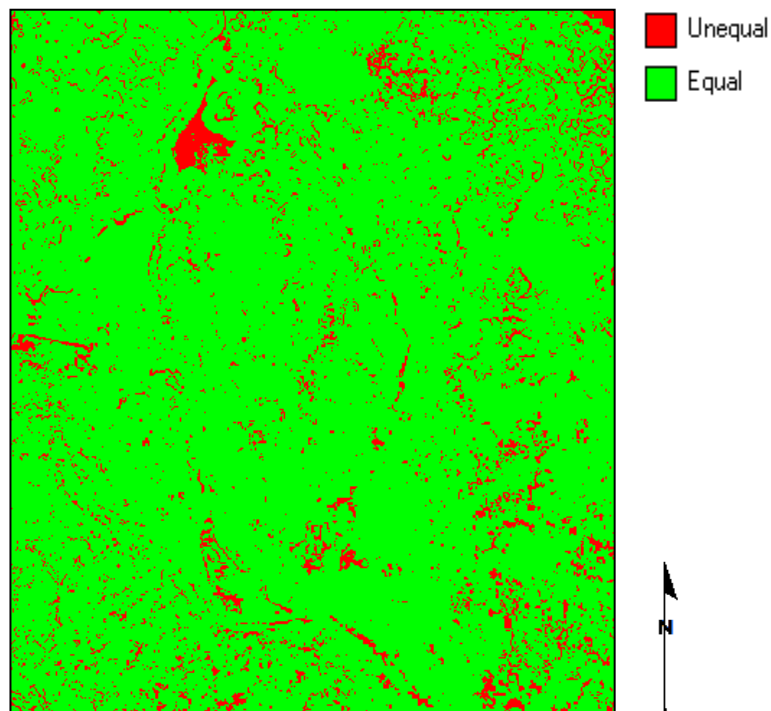
- a) The final output of CA\_Markov is quite acceptable based on kappa values.
- b) Though there exist some dissimilarity comparing to the original base map, but it can be solved by incorporating better suitability maps.
- c) No constraint and factor maps have been included while preparing the suitability maps for the study area. This is one drawback for this type of modelling.
- d) The suitability maps have been prepared taking into consideration only one basic assumption of spatial proximity. This is due to the lack of proper information/data.
- e) If these shortcomings can be minimized then the situation would be different.



### 8.2.3 Base Map (2009) vs. MLP Markov (2009)

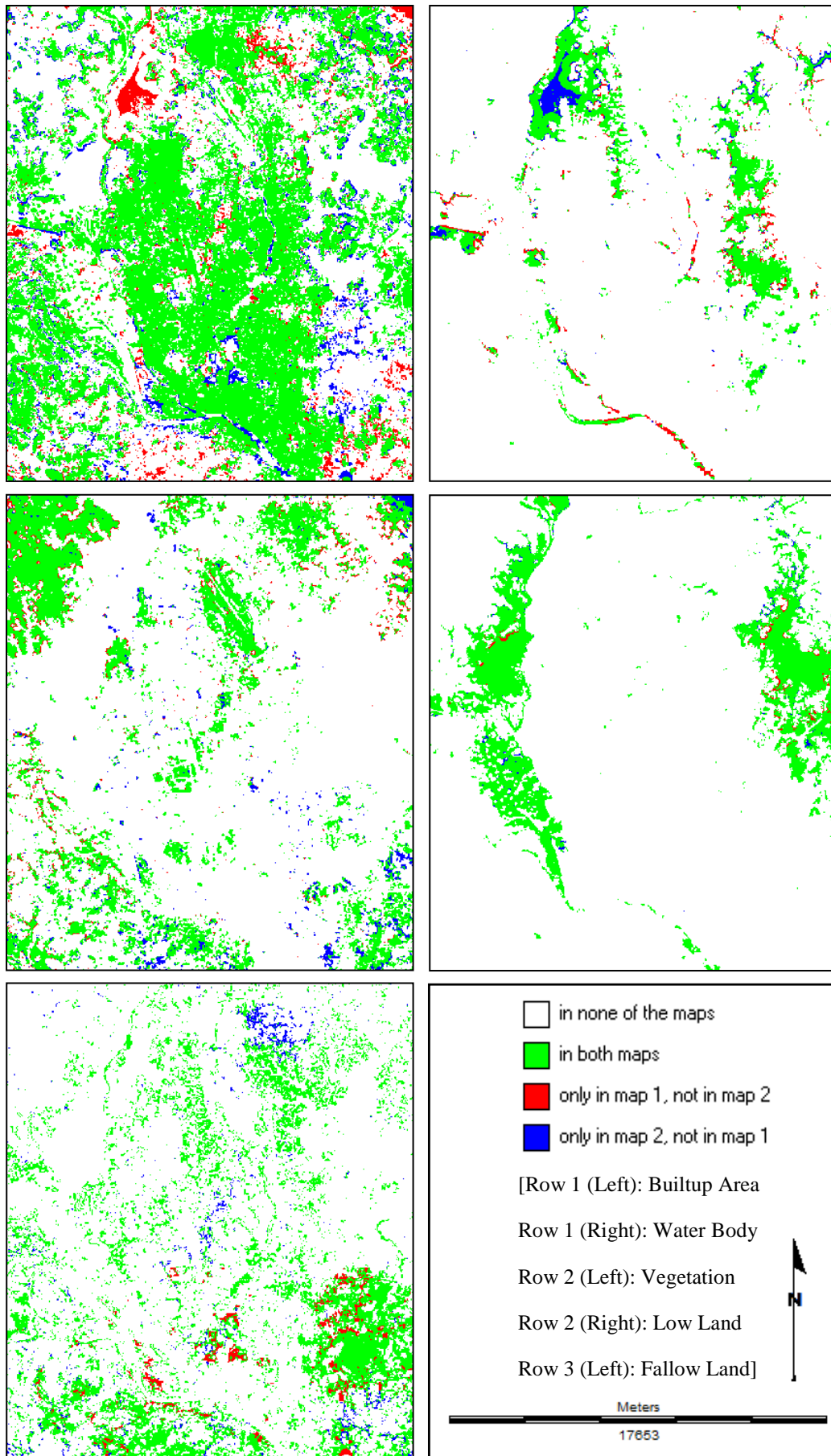


**Figure 8.14: Maps for Model Validation**



**Figure 8.15: Levels of Agreement for Kappa**

Figure 8.14 illustrates the maps to be compared for MLP\_Markov model validation. The MLP\_Markov simulated map is almost perfect to the actual base map, as there is very high level of agreement (Figure 8.15).



**Figure 8.16: Per Category Comparison Method**

All the land cover categories are showing almost perfect agreement of cells between the two maps (Figure 8.16). Some mismatching has been found particularly in builtup area and water body cover types. Still the level of agreement is almost perfect.

**Table 8.6: Per Category Kappa Statistics**

	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>Kappa</b>	0.83853	0.84933	0.90185	0.97450	0.91262
<b>Klocation</b>	0.83855	0.84936	0.90187	0.97451	0.91262

The various kappa coefficients of each category are showing higher values (Table 8.6). This also proves the level of agreement is almost perfect.

### **8.3 Model Selection**

Three different models have been implemented for this research to predict the future land cover change. Now the task is to select the most suitable model, out of these three, for this particular research. The selection has been performed based on the overall kappa statistics. The assumption is the higher the overall kappa values, the better the model.

**Table 8.7: Overall Kappa Statistics and Fraction Correct**

	<b>St_Markov (2009)</b>	<b>CA_Markov (2009)</b>	<b>MLP_Markov (2009)</b>
<b>Fraction Correct</b>	0.50275	0.72558	0.91982
<b>Kno</b>	0.37842	0.65704	0.89983
<b>Klocation</b>	0.33206	0.72616	0.88691
<b>Khisto</b>	0.87764	0.86177	0.93228
<b>Kappa</b>	<b>0.29143</b>	<b>0.62578</b>	<b>0.88689</b>

It is possible to validate the similarity between an observed/actual base map and a model map, by calculation the relative measures of various kappa statistics. After analyzing Table 8.7, it can be concluded that MLP\_Markov model is showing the highest values of kappa coefficients among the three.

Therefore, the MLP\_Markov model has been selected for predicting the land cover map of Dhaka City for the year of 2019.



# Chapter 9

## Future Prediction and Analysis

### 9.1 Future Prediction

MLP\_Markov model has been selected for simulating land cover map of Dhaka city for the year 2019 (Chapter 08). In this chapter, only the very basic steps of the prediction procedure are described, because the details of MLP\_Markov process have been explained in Chapter 07. The first step is to create the Boolean and respective distance images.

#### 9.1.1 Creating Boolean Images (2009)

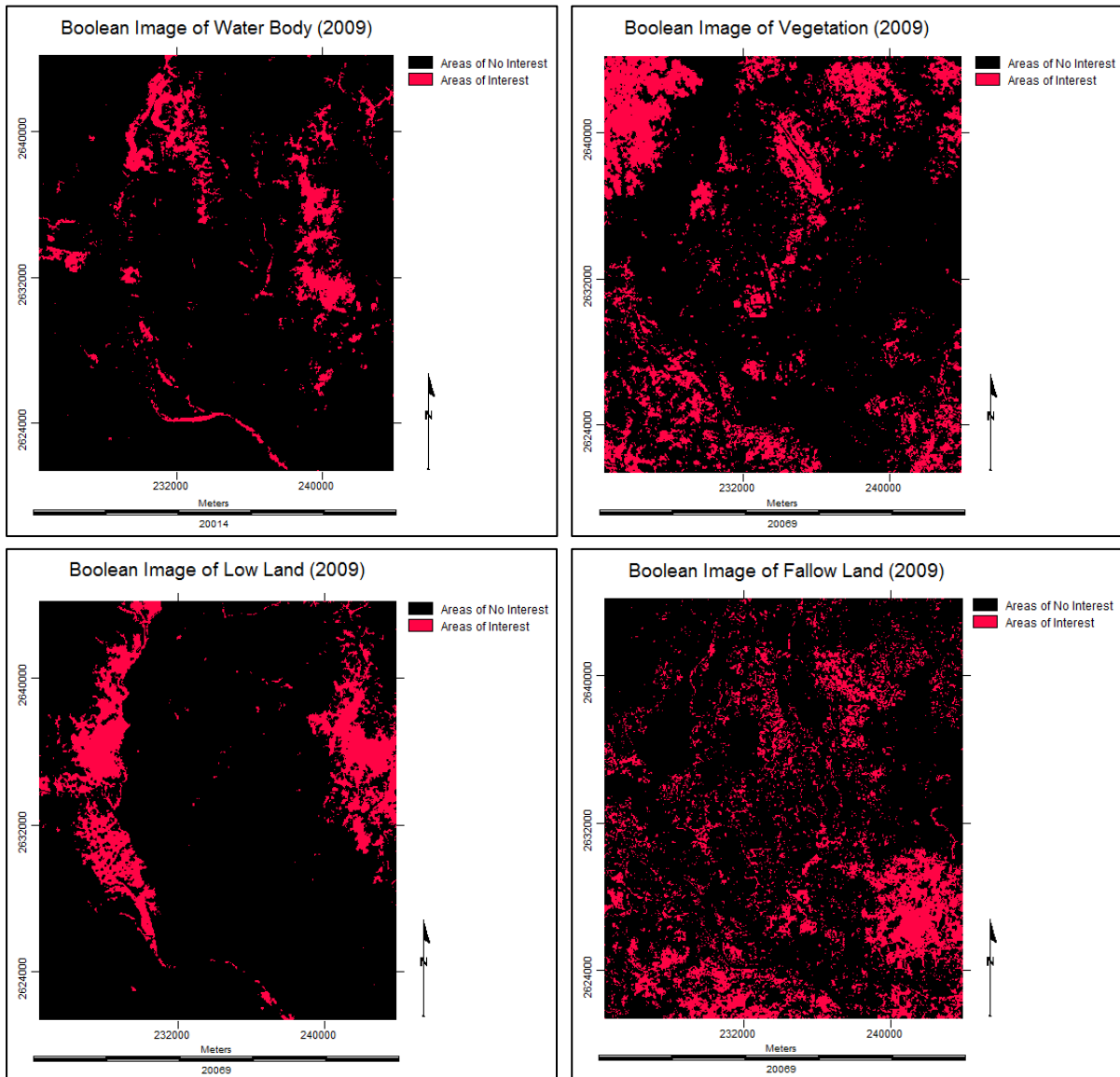
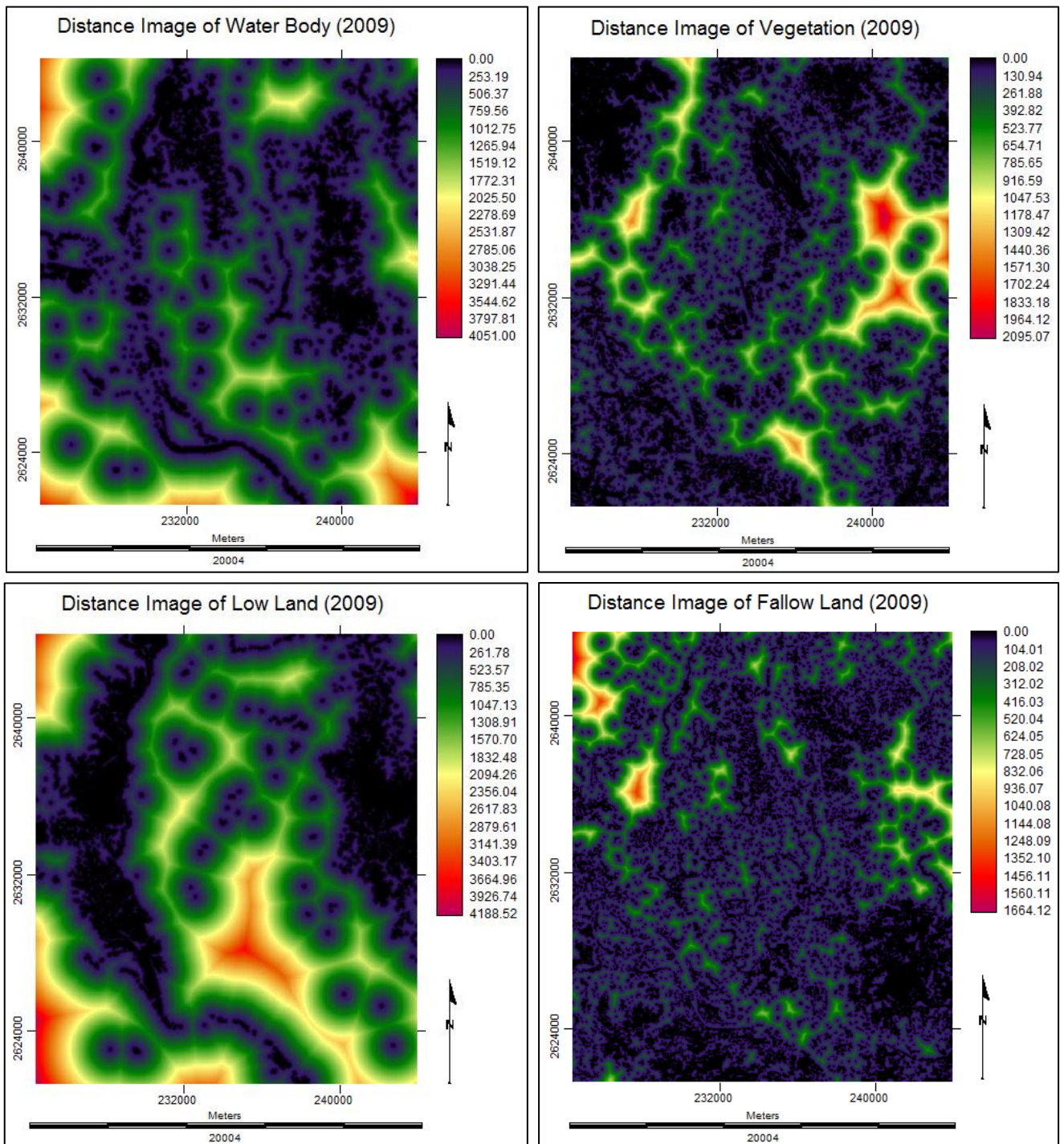


Figure 9.1: Boolean Images of each Land Cover Type (2009)

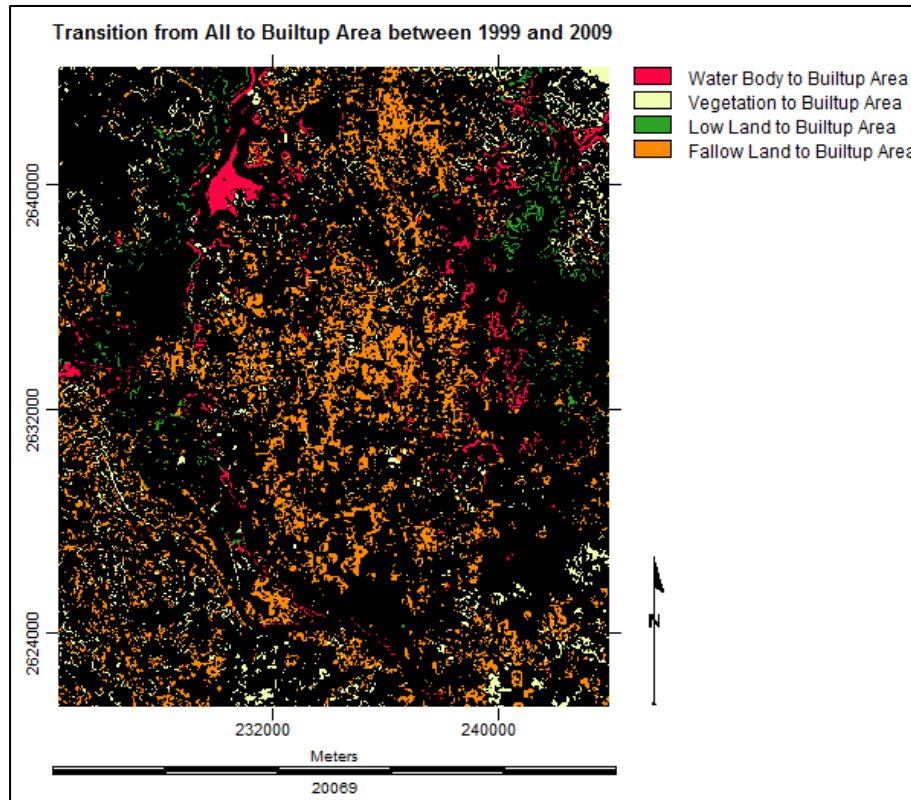
### 9.1.2 Creating Distance Images (2009)



**Figure 9.2: Distance Images of each Land Cover Type (2009)**  
**[Unit: Meter]**

### **9.1.3 Creating Land Cover Transition Image**

The basic concept of modelling with MLP neural network is to consider the change in builtup area over the years. Figure 9.3 depicts the transition from all land cover types to builtup area. This figure has been produced considering the transitions from all other land cover types to only builtup area. Other changes have been ignored.

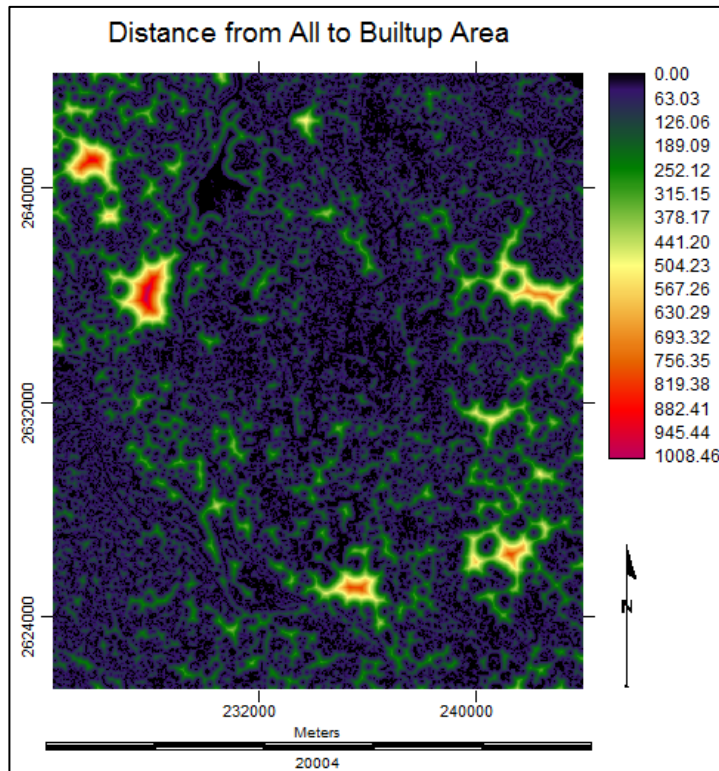


**Figure 9.3: Transition from All to Builtup Area (1999-2009)**

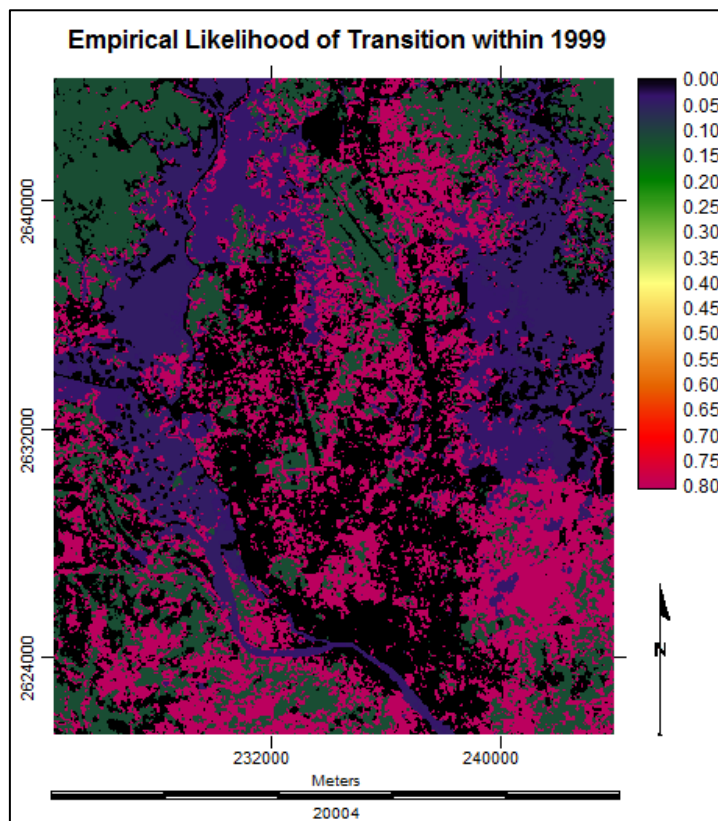
### **9.1.4 Selecting Driving Variables**

At this stage, the issue of which variables affect the change to builtup area (1999-2009) has been considered. Therefore, distance from all to builtup area has been chosen (Figure 9.4). Another important aspect is to find out the empirical likelihood of all changes for transforming into builtup areas based on the base image of the year 1999 (Figure 9.5). The highest value 0.80 is showing a high probability of converting other land cover types to builtup area. Therefore, 6 driver variables have been selected for the model. These are: distance from all to builtup area, distance from water body, distance from vegetation, distance from low land, distance from fallow land and empirical likelihood image.





**Figure 9.4: Distance Image of Transition from All to Builtup Area (1999-2009)**



**Figure 9.5: Empirical Likelihood Image (1999-2009)**

### **9.1.5 Testing Potential Explanatory Power of the Driving Variables**

The quantitative measures of the variables have been tested through Cramer's V. It is suggested that driving variables having Cramer's V of about 0.15 or higher are useful. Table 9.1 shows that the potential explanatory values of the driving variables are useful (Cramer's V > 0.15).

**Table 9.1: Cramer's V of the Driving Factors**

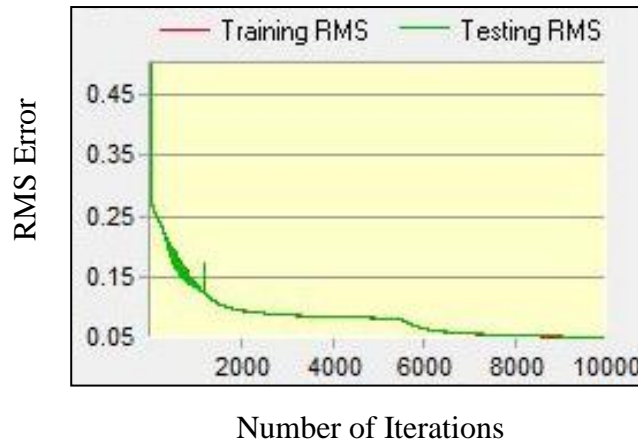
<b>Driving Variable / Cover Class</b>	<b>Overall V</b>	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
Distance from All to Builtup Area	0.0795	0.0000	0.5070	0.2987	0.2122	0.1442
Distance from Water Body	0.1231	0.0000	0.2615	0.8209	0.2835	0.1308
Distance from Vegetation	0.1566	0.0000	0.4600	0.2420	1.0000	0.4062
Distance from Low Land	0.1272	0.0000	0.2830	0.1896	0.2154	0.8701
Distance from Fallow Land	0.1511	0.0000	0.4851	0.1488	0.2528	0.3617
Empirical Likelihood Image	0.6470	0.0000	0.7254	0.8804	0.9020	0.9601

### **9.1.6 Transition Potential Modelling**

The MLP running statistics gives a very high accuracy rate of 99.29% (Table 9.2). The minimum number of cells that transitioned from 1989 to 1999 is 4794. The RMS error curve has also been found smooth and descent after running MLP neural network (Figure 9.6). It means the training result is satisfactory.

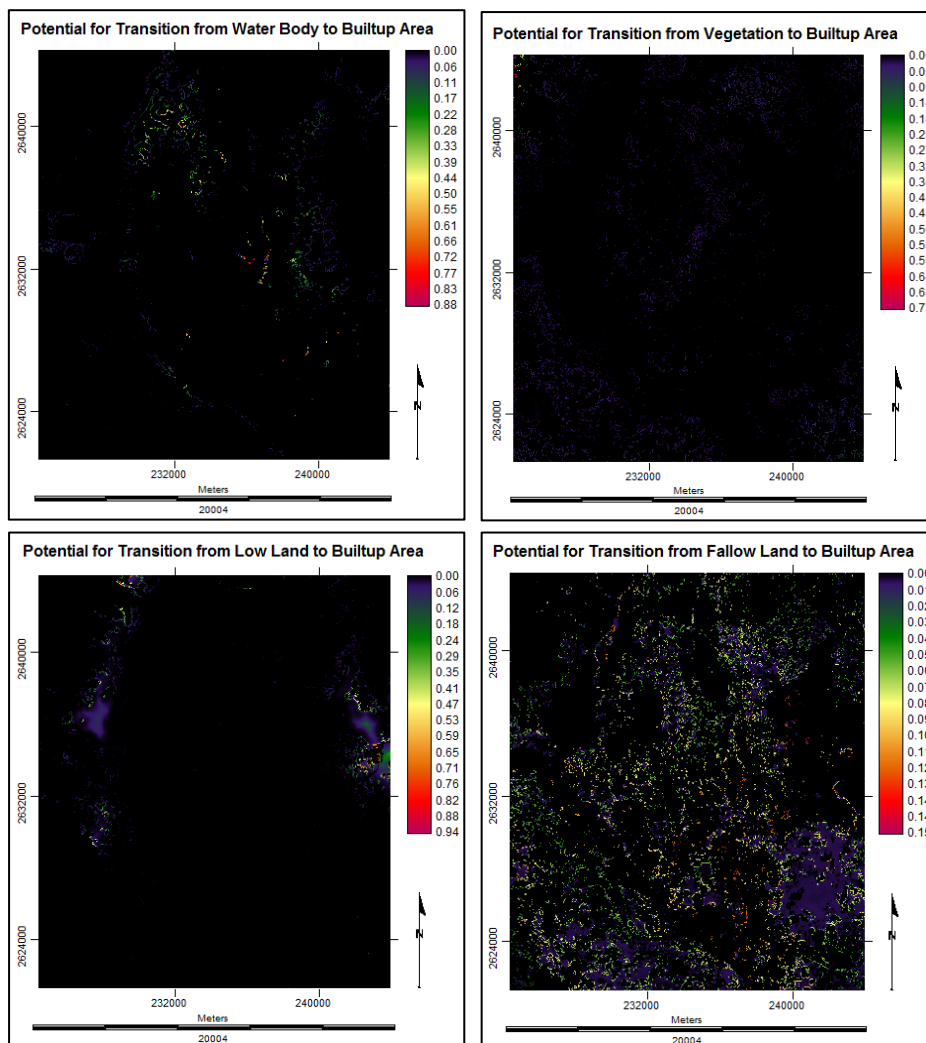
**Table 9.2: Running Statistics of MLP Neural Network**

Maximum Sample Size	4150
Iterations	10000
Training RMS	0.0533
Testing RMS	0.0541
Accuracy Rate	<b>99.29%</b>



**Figure 9.6: RMS Error Monitoring Curve**

Based on these running statistics four transition potential maps have been produced. The dominance of water body and fallow land to builtup area type is clear here (Figure 9.7).



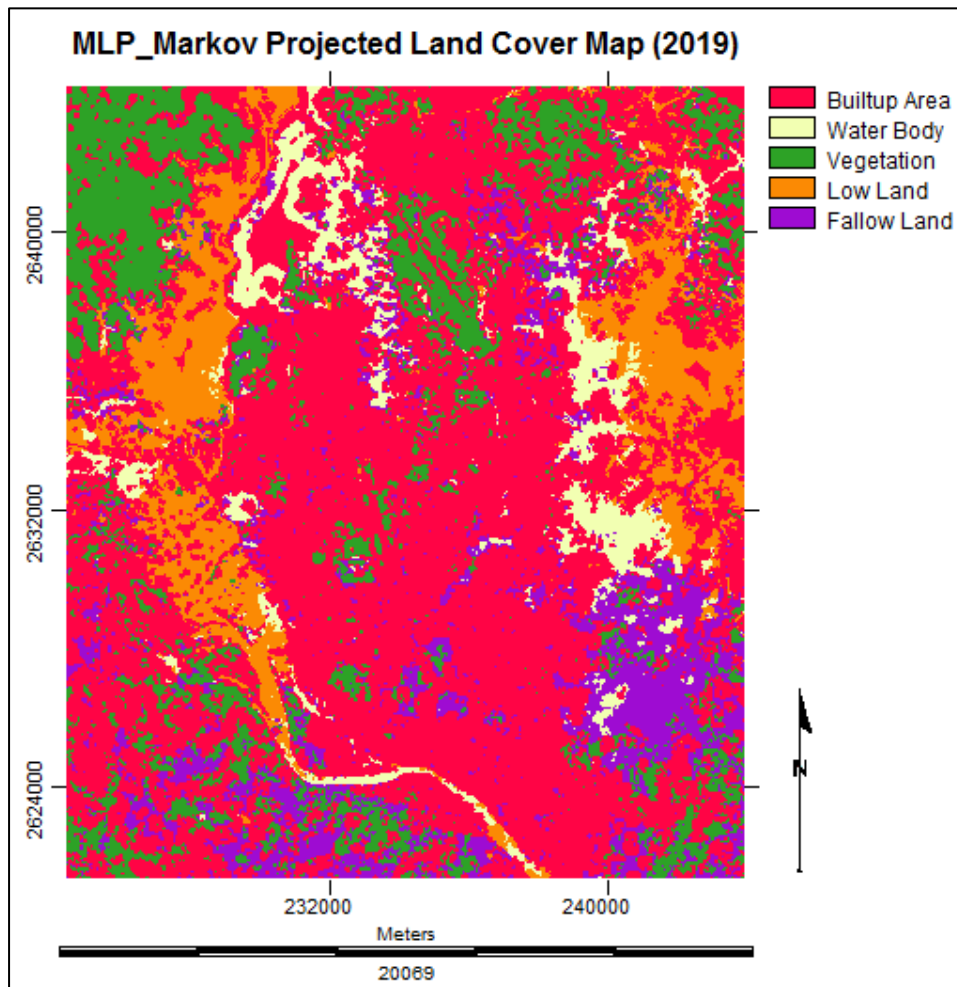
**Figure 9.7: Transition Potential Maps**

### 9.1.7 Markov Chain Analysis

Using this MLP neural network analysis it is possible to determine the weights of the transitions that will be included in the matrix of probabilities of Markov Chain for future prediction. The transition probabilities are shown in Table 9.3. Based on all these information from MLP neural network, the final predicted map of 2019 has been simulated through Markov chain analysis (Figure 9.8).

**Table 9.3: Transition Probabilities Grid for Markov Chain**

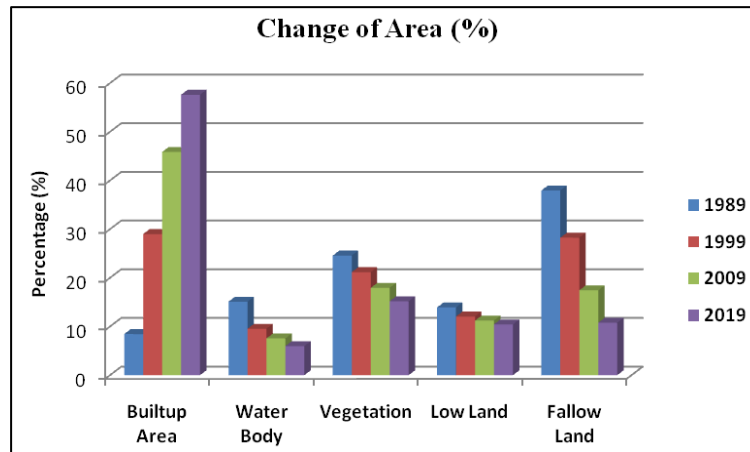
	Builtup Area	Water Body	Vegetation	Low Land	Fallow Land
Builtup Area	0.6782	0.0275	0.0785	0.0328	0.1830
Water Body	0.3601	0.1024	0.1579	0.1242	0.2553
Vegetation	0.3073	0.0853	0.2352	0.1081	0.2641
Low Land	0.2093	0.1997	0.1038	0.2429	0.2444
Fallow Land	0.4719	0.0384	0.1923	0.0497	0.2477



**Figure 9.8: MLP\_Markov Projected Land Cover Map of Dhaka City (2019)**

## **9.2 Analysis of the Predicted Map**

The predicted map of 2019 reveals that 58% of the total area will be occupied by builtup area cover type (Figure 9.9). On the other hand, water body (6%) and fallow land (11%) types are going to decrease in a notable way.



**Figure 9.9: Change in Area (%) over the Years (1989-2019)**

Gains in builtup area land cover type are prominent while most of the areas will be persistent (Figure 9.10). Slight loss in water body, low land and vegetation cover types will also be found. But fallow land type will decrease in good amount in near future.

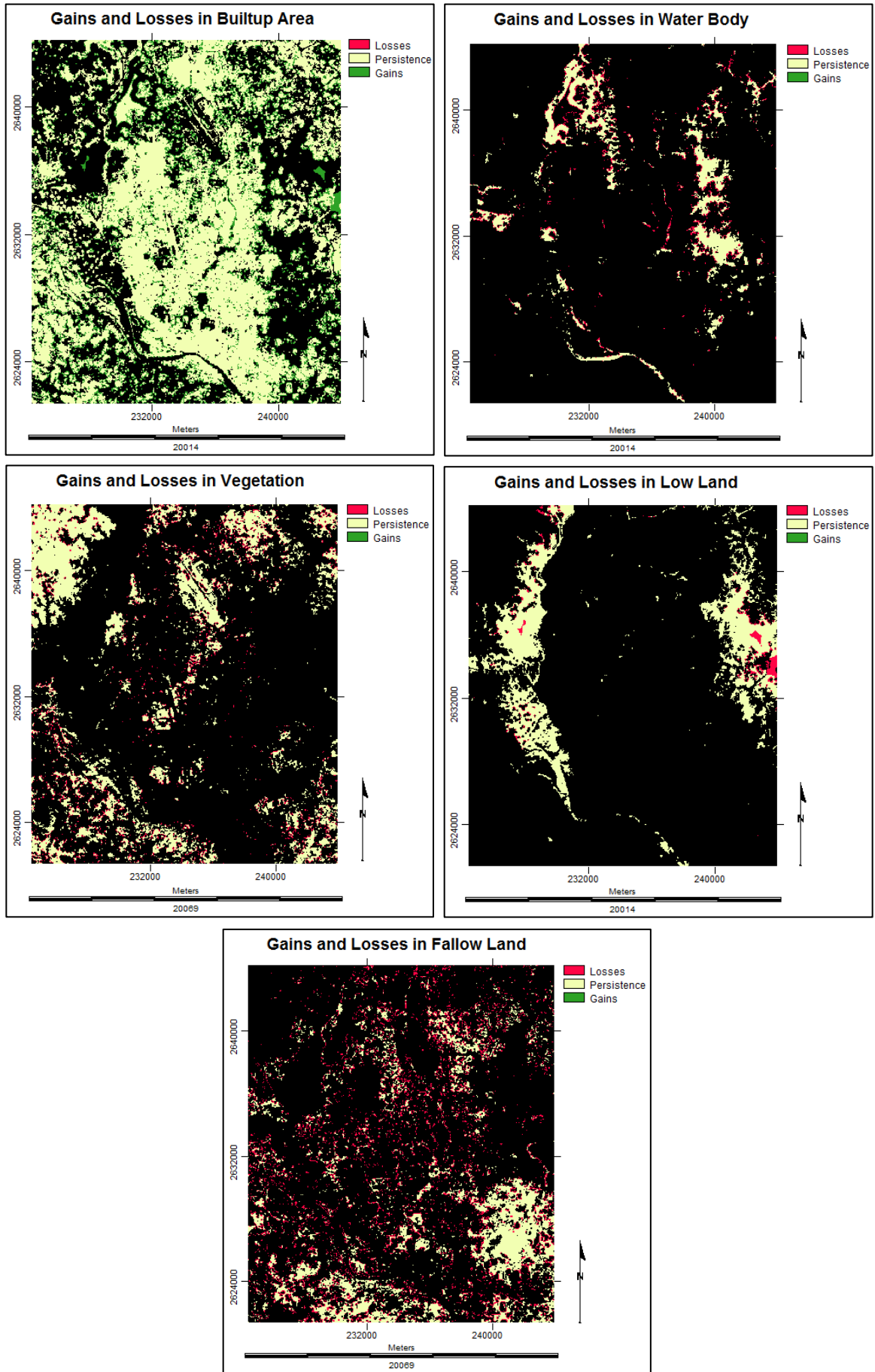
## **9.3 Limitations of MLP Markov Model**

Long term predictions are not possible by applying the MLP\_Markov model. As the urban growth rate is very high in case of Dhaka, it is giving errors in predicting land cover maps for long terms (2030 or 2050). It needs more horizontal space to depict the future change. As there is no provision for extending the study area in the horizontal dimension, long term prediction is not possible. It means in 2019, the study area will reach its threshold limit in builtup area type. Some major drawbacks may be as follows:

- a) Selecting few number of driving factors for model calibration.
- b) Not incorporating the basic road network and elevation layers.

Therefore, CA\_Markov model is better option for long term prediction for this research. Another drawback is that these kinds of land cover modelling techniques do not consider the vertical growth of the city. Like the rise in building heights can contribute to the urban expansion in another adverse way. This causes the increase in urban density. But all these three models, which have been implemented for the future prediction, only consider the horizontal expansion of urban areas.





**Figure 9.10: Gains and Losses of Land Cover Types (2009-2019)**

# Chapter 10

## Recommendations and Conclusions

---

### **10.1 Answers to Research Questions**

The purpose of this research is to achieve the objectives mentioned in Chapter 01 (Section 1.4). This research has also been conducted to fulfil the research hypotheses (Section 1.5). The brief answers to the research questions are as follows:

#### **10.1.1 Answers Related to Objective 1**

**Question 01:** Which datasets are available to conduct the whole research?

**Answer 01:** Landsat Satellite Images have been used for this purpose.

**Question 02:** Which techniques are available for analyzing the change detection of land cover types of Dhaka city?

**Answer 02:** To explain the change detection of Dhaka city, different spatial metrics have been used for quantitative analysis. Moreover, some post-classification change detection techniques have also been implemented.

**Question 03:** What is the general trend of land cover changes of the study area over time (1989-2009)?

**Answer 03:** Builtup area land cover type is increasing in an alarming rate over the years. While the major contributors are fallow land and water body land cover types.

#### **10.1.2 Answers Related to Objective 2**

**Question 01:** Which methods are available to simulate the future urban growth of the study area?

**Answer 01:** Three different methods have been implemented to simulate the land cover map of Dhaka city of 2009. The methods have been named as ‘Stochastic Markov (St\_Markov)’ Model, ‘Cellular Automata Markov (CA\_Markov)’ Model and ‘Multi Layer Perceptron Markov (MLP\_Markov)’ Model. Later, using the MLP\_Markov model, the land cover map of 2019 has been predicted.

**Question 02:** Which method is suitable for forecasting the future land cover changes?

**Answer 02:** The ‘Multi Layer Perceptron Markov (MLP\_Markov)’ model has been found most suitable among the three models for future prediction. The best-fitted model has been selected based on Kappa statistics values and also by implementing other map comparison or model validation techniques.

**Question 03:** Are the satellite images, GIS and Remote Sensing tools and different methods/ techniques used for this research adequate or useful?

**Answer 03:** The RS and GIS techniques or methods implemented, for achieving the goals this research, have been found quite useful. The low resolution of the Landsat satellite images (30 m) is a problem.

**Question 04:** What will be the future scenario of Dhaka city based on the analysis and simulation of past trends?

**Answer 04:** Based on the MLP\_Markov model simulation, it has been found that 58% of the total study area will be converted into builtup area cover type in 2019. This should be a matter of real concern, for the urban planners and city dwellers as well.

## **10.2 Recommendations for Future Work**

The following recommendations can be incorporated for similar research works in future:

- a) It has been proved that leaf-off season (dry season) images are better than leaf-on season satellite images for identifying builtup areas [79]. This means seasonal variation is an important aspect conducting this kind of research. High resolution satellite images with equal time interval is another important aspect.
- b) This kind of research needs extensive field visit for image classification and assessing the accuracies of the satellite images. Another point is the verification of the older images. For older images it is not possible to visit the field to find out the actual land cover types. These things can be improved by recent field visit to collect Global Positioning System (GPS) data for land cover verification/ground truthing purpose. To solve the problems with older satellite images, the historical base maps of similar years should be collected.

- c) Lack of data collection is another problem that can be improved while conduction similar future works. For example, in case of MLP\_Markov modelling, it is possible to incorporate basic road layer, elevation and other driving variables for future prediction. Again, for CA\_Markov modelling, identifying some constraints like environmental protection areas, flood retention ponds, preserved wetlands or water bodies etc. are important. Moreover, while creating factors images it is possible to develop some criteria like distance from existing roads or urban builtup areas. All these issues have been neglected in this research. The reasons behind this are lack of necessary data and the low resolution of Landsat satellite images. Working with Landsat satellite images, it is not possible to identify the detail urban pattern like classifying the image into more land cover types. This problem can be solved using object-oriented method rather using the traditional pixel-based classification method [30].
- d) Further studies can be conducted, to find out whether these kinds of model are suitable for future prediction, in areas where urban growth is not higher like Dhaka city. This would be an interesting analysis to discover which model is best-suited for particular research problems.
- e) Only simulating the future land cover map is not fruitful, until further steps are taken. Some more important research works can be recommended like- trying to compare the predicted map with the existing Master plans of the city. If there any conflicts exist, then it should be taken into special consideration.

### **10.3 Recommendations for Real World Plan Preparation**

Predicting the future of anything is not an easy task. On the other hand, urbanization is a complex and dynamic system. Numerous physical, social, economic and environmental factors put impact on shaping the urban form of a city. Considering all these driving and influencing factors, it is a difficult and hectic job to predict the future of any city or urban area. The future land cover map of Dhaka city (2019) has been generated, in this research, using MLP\_Markov modelling techniques.

Despite all kinds of limitations, it can be concluded that the accuracy results of the supervised classified images, Kappa co-efficient values for model validation and the overall predictive results have been found exceeding the agreeable levels.

The Dhaka Metropolitan Development Plan (DMDP) of RAJUK has already prepared a Structure Plan (1995-2015) and an Urban Area Plan (1995-2005). The Detailed Area Plan (DAP) is the third and last tier of Development Plan for Dhaka City [80]. DAP is a very vital part of the DMDP as far as spatial development and development controlled is concerned. The Government of the People's Republic of Bangladesh (GoB) is soon going to implement DAP. The GoB is also planning to extend the DMDP from 2015-2035 in phase-II [80]. Therefore, the concerned urban planning and development authorities can go for this kind of analytical research works, for controlling and monitoring the future urban growth of Dhaka city. The traditional plan making process, that involves extensive surveying, is huge expensive and time consuming. This kind of research to forecast the future is not only time and money saving, but also effective.

#### **10.4 Epilogue**

Dhaka, as the capital city of Bangladesh, encounters a host of problems perpetuating for decades. With the growing importance of the city and mounting up pressure of population, the problems are getting more and more precarious every year. Amid all the problems, uncontrolled urban growth/ development is perhaps the most irritating one particularly for the planning and development control authorities like DCC and RAJUK.

The unwanted urban growth originates not only due to ineffective development control but also due to failure of the equilibrium between demand and supply of buildable and liveable urban land. Therefore, regular monitoring of the plan implementation is necessary together with the urban development trend in new areas. An early measure in tackling problems can not only save huge public money but also reduce the miseries of the city dwellers. In this regard, performing this kind of change detection analysis to monitor the urban growth is highly recommended. Moreover, it is also important to predict the future scenario for controlling the unwanted urban sprawl or growth. This can be facilitated, by applying the RS and GIS based future prediction methods as explained in this research. This kind of research has a high potential to contribute towards the sustainable urban growth both at any local and regional level in the world.

The interpretation of depicting the future scenario in quantitative accounts, as demonstrated in this research, will be of great value to the urban planners and decision makers, for the future planning of the modern and much liveable Dhaka City.

## REFERENCES

---

- [1] Kashem, M.S.B., 2008, *Simulating Urban Growth Dynamics of Dhaka Metropolitan Area: A Cellular Automata Based Approach*. Master's Thesis, Department of Urban and Regional Planning, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh.
- [2] Ahmed, S.U. (Ed.s), 1991, *Dhaka: Past Present Future* (Dhaka: The Asiatic Society of Bangladesh).
- [3] Mamun, M., 1993, *Dhaka Smriti Bismritir Nagari* (Bengali) (Dhaka: Bangla Academy).
- [4] Taifoor, S.M., 1956, *Glimpses of Old Dhaka* (Dhaka: Pioneer).
- [5] Dani, A.H., 1962, *Dacca: A Record of its Changing Fortunes* (Dacca: Asiatic Press).
- [6] Hazel, G.P., e Miller, D., 2004, *Megacity Challenges: A Stakeholder Perspective*, A Research Project conducted by GlobeScan and MRC McLean Hazel Sponsored by Siemens, Siemens AG, Corporate Communications (CC), Wittelsbacherplatz 2, 80333 Munich, Germany.
- [7] Dewan, A. M., e Yamaguchi, Y., 2009, Landuse and Land Cover Change in Greater Dhaka, Bangladesh: Using Remote Sensing to Promote Sustainable Urbanization. *Journal of Applied Geography*, **29**, 390–401.
- [8] Basak, P., 2006, *Spatio-Temporal Treads and Dimensions of Urban Form in Central Bangladesh: A GIS and Remote Sensing Analysis*. Master's thesis, Department of Urban and Regional Planning, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh.
- [9] Khan, N. S., 2000, Temporal Mapping and Spatial Analysis of Land Transformation Due to Urbanization and its Impact on Surface Water System: A Case from Dhaka Metropolitan Area, Bangladesh. *International Archives of Photogrammetry and Remote Sensing*, **Vol. XXXIII, Part B7**, Amsterdam.

- [10] Ahmed, B., Hasan, R., e Ahmad, S., 2008, *A Case Study of the Morphological Change of Four Wards of Dhaka City over the Last 60 Years (1947-2007)*. Bachelor Thesis, Department of Urban and Regional Planning, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh.
- [11] Bangladesh Bureau of Statistics (BBS), 2009, *Statistical Pocket Book of Bangladesh 2008*, Bangladesh Bureau of Statistics, Planning Division, Ministry of Planning, Government of the People's Republic of Bangladesh (GoB), Dhaka.
- [12] Dhaka City Corporation, 2010, *History of Dhaka City Corporation >>Dhaka as the Capital of Bangladesh*, Nagar Bhaban, Fulbaria, Dhaka 1000, Bangladesh (URL: [www.dhakacity.org](http://www.dhakacity.org), retrieved on 06-01-2011).
- [13] Zlotnik, H., 2010, *World Urbanization Prospects: The 2009 Revision Population Database*, Population Division, United Nations, 2 United Nations Plaza, Rm. DC2-1950, New York, NY 10017, USA (URL: <http://esa.un.org/wup2009/wup/source/country.aspx>, retrieved on 05-01-2011).
- [14] Islam, N., e Khan, F.K., 1964, High Class Residential Areas in Dhaka City, 1608-1962. *Oriental Geographer*, **8:1**, pp 1-40.
- [15] Nilufar, F., 1997, *The Spatial and Social Structuring of Local Areas in Dhaka City- A Morphological Study of the Urban Grid with Reference to Neighbourhood Character within Naturally-Grown Areas*. PhD Dissertation, Unit for Advanced Architectural Studies, The Bartlett School of Graduate Studies, University College London, University of London, UK.
- [16] The Economist Intelligence Unit Limited, 2010, *Liveability Ranking and Overview (Sample)*, 26 Red Lion Square London WC1R 4HQ, United Kingdom (URL: <http://store.eiu.com/product/475217632-sample.html>, retrieved on 06-01-2011).
- [17] The World Factbook, 2010, *Country Profile: Bangladesh*, The Central Intelligence Agency, Office of Public Affairs, Washington, D.C. 20505, USA (URL: <https://www.cia.gov/library/publications/the-world-factbook/geos/bg.html>, retrieved on 06-01-2011).
- [18] Malczewski, J., 2004, GIS-Based Land-Use Suitability Analysis: A Critical Overview. *Progress in Planning*, **62**, 3–65.



- [19] Jansena, L.J.M., e Gregorio, A.D., 2003, Land-Use Data Collection Using the “Land Cover Classification System”: Results from a Case Study in Kenya. *Land Use Policy*, **20**, 131–148.
- [20] Griffiths, P., Hostert, P., Gruebner, O., e Linden, S.V.D., 2010, Mapping Mega City Growth with Multi-Sensor Data. *Remote Sensing of Environment*, **114**, 426–439.
- [21] U.S. Department of the Interior, 2010, *Landsat Product Information*, U.S. Geological Survey (URL: <http://landsat.usgs.gov>, retrieved on 09-01-2011).
- [22] Lambin, E.F., 2001, Remote Sensing and GIS Analysis. In *International Encyclopedia of the Social and Behavioral Sciences*, edited by N.J. Smelser e P.B. Baltes (Oxford: Pergamon), pp. 13150-13155.
- [23] Carmona, P.L., e Bañón, F.P., 2008, *Introduction to Image Processing Techniques for Remote Sensing* [CD-ROM] (e-Treballs d'informàtica i tecnologia: Universitat Jaume I, Castellón, Spain).
- [24] Goodchild, M.F., 2009, Geographic Information Systems and Science: Today and Tomorrow. *Procedia Earth and Planetary Science*, **1**, 1037–1043.
- [25] Goodchild, M. F., 2001, Geographical Information Systems. In *International Encyclopedia of the Social and Behavioral Sciences*, edited by N.J. Smelser e P.B. Baltes (Oxford: Pergamon), pp. 6175-6182.
- [26] Chrisman, N. R., 1999, What does ‘GIS’ mean? *Transactions in GIS*, **3(2)**, 175-86.
- [27] Maguire, D.J., 1991, An Overview and Definition of GIS. In *Geographic Information Systems: Principles and Applications*, edited by D. J. Maguire, M. F. Goodchild e D. W. Rhind (London: Longman), Vol 1, pp. 9-20.
- [28] Goodchild, M. F., 2009, GIScience and Systems, In *International Encyclopedia of Human Geography*, edited by R. Kitchin and M. Thrift, (New York: Springer), pp. 536-538.

- [29] Food and Agriculture Organization (FAO), 1995, Planning for Sustainable Use of Land Resources, FAO Land and Water Bulletin 2, Rome: Food and Agriculture Organization of the United Nations, (URL: <http://www.fao.org/docrep/v8047e/v8047e04.htm#land%20and%20land%20resources>, retrieved on 18-01-2011).
- [30] Das, T., 2009, *Land Use / Land Cover Change Detection: an Object Oriented Approach*, Münster, Germany. Thesis: Master of Science in Geospatial Technologies, Institute for Geoinformatics, University of Münster, Germany.
- [31] Emch, M. e Peterson, M., 2006, Mangrove Forest Cover Change in the Bangladesh Sundarbans from 1989-2000: A Remote Sensing Approach. *Geocarto International*, **21: 1**, 5-12.
- [32] Lahti, J., 2008, *Modelling Urban Growth Using Cellular Automata: A Case Study of Sydney, Australia*. Thesis: Master of Science in Geo-information Science and Earth Observation for Environmental Modelling and Management, International Institute for Geo-information Science and Earth Observation, Enschede, The Netherlands.
- [33] Li, X., e Yeh, A.G.O., 2002, Neural-Network-Based Cellular Automata for Simulating Multiple Land Use Changes Using GIS. *International Journal of Geographic Information Science*, **16:4**, 323-343.
- [34] Cabral1, P. e Zamyatin, A., 2006, Three Land Change Models for Urban Dynamics Analysis in Sintra-Cascais Area. *1st EARSeL Workshop of the SIG Urban Remote Sensing*, (Germany: Humboldt-Universität zu Berlin).
- [35] Billah, M., e Rahman, G. A., 2004, Land Cover Mapping of Khulna City Applying Remote Sensing Technique. In *Proceedings of the 12<sup>th</sup> International Conference on Geoinformatics* (Sweden: Geospatial Information Research: Bridging the Pacific and Atlantic, University of Gävle), pp. 707-714.
- [36] Eastman, J. R., 2009, IDRISI Taiga Guide to GIS and Image Processing (Manual Version 16.02) [Software] (Massachusetts, USA: Clark Labs, Clark University).
- [37] Eastman, J. R., 2009, IDRISI Taiga Tutorial (Manual Version 16.02) [Software] (Massachusetts, USA: Clark Labs, Clark University).

- [38] Geospatial Data Service Centre, 2008, Netherlands (URL: [http://gdsc.nlr.nl/gdsc/information/earth\\_observation/band\\_combinations](http://gdsc.nlr.nl/gdsc/information/earth_observation/band_combinations) , retrieved on 03-10-2010).
- [39] Canada Centre for Remote Sensing, 2010, Natural Resources Canada, 588 Booth Street, Ottawa, Ontario, K1A 0Y7, Canada (URL: [http://www.ccrs.nrcan.gc.ca/glossary/index\\_e.php](http://www.ccrs.nrcan.gc.ca/glossary/index_e.php) , retrieved on 15-11-2010).
- [40] Congalton, R., 1991, A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment*, **37**, 35-46.
- [41] The ERDAS IMAGINE 9.1 On-Line Help System, 2006, Leica Geosystems Geospatial Imaging, LLC [Software] (Worldwide Headquarters, 5051 Peachtree Corners Circle, Norcross, GA 30092-2500, USA).
- [42] National Oceanic and Atmospheric Administration (NOAA), 2008, NOAA Coastal Services Center, 2234 South Hobson Avenue, Charleston, SC 29405-2413, USA (URL: [http://www.csc.noaa.gov/crs/lca/faq\\_tech.html#q4](http://www.csc.noaa.gov/crs/lca/faq_tech.html#q4), retrieved on 18-11-2010).
- [43] Hagen, A., 2002, *Technical Report: Comparison of Maps Containing Nominal Data*. Research Institute for Knowledge Systems BV, Maastricht, The Netherlands.
- [44] Hagen, A., 2002, Multi-Method Assessment of Map Similarity. In: Ruiz, M., Gould, M., Ramon, J. (Eds.), *Proceedings of the Fifth AGILE Conference on Geographic Information Science* (Palma, Spain), pp. 171-182.
- [45] Simon, S., 2008, *What is a Kappa coefficient? (Cohen's Kappa)* (URL: <http://www.childrens-mercy.org/stats/definitions/kappa.htm>, retrieved on 12-11-2010).
- [46] Landis, J. R., e Koch, G. G., 1977, The Measurement of Observer Agreement for Categorical Data. *Biometrics*, **33**, 159-174.
- [47] Pontius, R. G., 2000, Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering & Remote Sensing*, **66 (8)**, 1011-1016.
- [48] Araya, Y.H., 2009, *Urban Land Use Change Analysis and Modelling: A Case Study of Setúbal and Sesimbra, Portugal*. Thesis: Master of Science in Geospatial Technologies, Institute for Geoinformatics, University of Münster, Germany.

- [49] Cabral, P., Geroyannis, H., Gilg, J.P., e Painho, M., 2005, Analysis and Modelling of Land-Use and Land-Cover Change in Sintra-Cascais Area. *In Proceedings of 8<sup>th</sup> Agile Conference* (Estoril: Portugal).
- [50] McGarigal, K., S. A. Cushman, M. C. Neel, e E. Ene, 2002, *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps*. Computer software program produced by the authors at the University of Massachusetts, Amherst, USA. (URL: <http://www.umass.edu/landeco/research/fragstats/fragstats.html>, retrieved on 08-01-2011).
- [51] Thesis Blog, 2010, Stochastic Processes and Markov Chains (URL: <http://research.sourcebyte.com/blog/?p=133>, retrieved on 19-12-2010).
- [52] Kuebler, S., 2009, *Stochastic Processes, Markov Chains, and Markov Models*. Lecture of Course: L645, Department of Linguistics, Indiana University, USA.
- [53] Basharin, G. P., Langville, A. N., e Naumov, V. A., 2004, The Life and Work of A. A. Markov. *Linear Algebra and its Applications*, **386**, 3–26.
- [54] Balzter, H., 2000, Markov Chain Models for Vegetation Dynamics. *Ecological Modelling*, **126**, 139–154.
- [55] Charoenjit, K., 2009, *Application of Markov-Cellular Automata and Social Models for Land Use Prediction in Central Petchaburi Watershed, Thailand*. Master Dissertation, Faculty of Graduate Studies, Mahidol University, Thailand.
- [56] Wolfram, S., 1983, Statistical Mechanics of Cellular Automata. *Reviews of Modern Physics*, **55** (3), 601-644.
- [57] Maerivoet, S., e Moor, B. D., 2005, Cellular Automata Models of Road Traffic. *Physics Reports*, **419** (1), 1-64.
- [58] Kier, L.B., Seybold, P.G., e Cheng, C.K., 2005, *Modeling Chemical Systems Using Cellular Automata* (Springer: The Netherlands).
- [59] Singh, A. K., 2003, *Modelling Land Use Land Cover Changes Using Cellular Automata in a Geo-Spatial Environment*. Master Dissertation, International Institution for Geo-Information and Earth Observation, Enschede, The Netherlands.

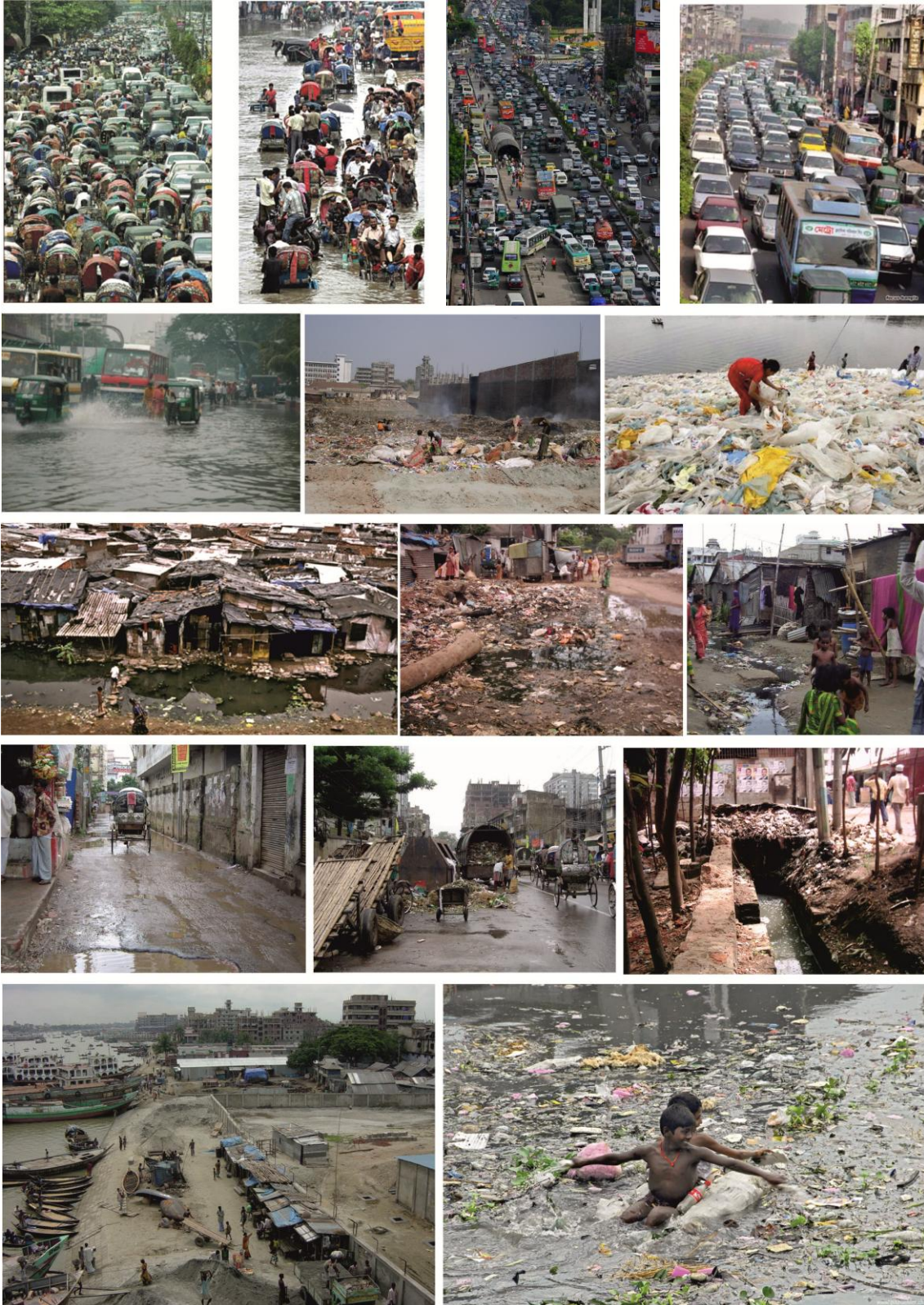
- [60] Santé, I., García, A. M., Miranda, D., e Crecente, R., 2010, Cellular Automata Models for the Simulation of Real-World Urban Processes: A Review and Analysis. *Landscape and Urban Planning*, **96**, 108–122.
- [61] Torrens, P.M., 2000, *How Cellular Models of Urban Systems Work*, Working Paper 28, Centre for Advanced Spatial Analysis (CASA), University College London, UK.
- [62] Barredo, J. I., Kasanko, M., McCormick, N., e Lavalle, C., 2003, Modelling Dynamic Spatial Processes: Simulation of Urban Future Scenarios through Cellular Automata. *Landscape and Urban Planning*, **64**, 145–160.
- [63] Soltani, A., 2004, *Towards Modeling Urban Growth with Using Cellular Automata (CA) and GIS*, Research Report, University of South Australia, Australia.
- [64] Li, X., e Yeh, A.G.O., 2000, Modelling Sustainable Urban Development by the Integration of Constrained Cellular Automata and GIS. *International Journal of Geographical Information Science*, **14 (2)**, 131-152.
- [65] Collins, M.G., Steiner, F.R., e Rushman, M.J., 2001, Land-Use Suitability Analysis in the United States: Historical Development and Promising Technological Achievements. *Environmental Management*, **28 (5)**, 611–621.
- [66] IDRISI Taiga Help System, 2009, Clark Labs, Clark University, 950 Main Street, Worcester, MA 01610-1477, USA (URL: [http://www.clarklabs.org/support/upload/IDRISI Taiga Help.zip](http://www.clarklabs.org/support/upload/IDRISI%20Taiga%20Help.zip), retrieved on 22-10-2010).
- [67] Houet, T. e Hubert-Moy, L., 2006, Modelling and projecting land-use and land-cover changes with a Cellular Automaton in considering landscape trajectories: An improvement for simulation of plausible future states [Electronic version]. *EARSeL eProceedings*, **5(1)**, 63-76.
- [68] Karul, C., e Soyupak, S., 2006, A Comparison between Neural Network Based and Multiple Regression Models for Chlorophyll-a Estimation. In Recknagel, F. (ed), *Ecological Informatics* (Berlin: Springer), pp. 309-323.

- [69] Civco, D. L., 1993, Artificial neural networks for land-cover classification and mapping. *International Journal of Geographical Information Science*, **7: 2**, 173-186.
- [70] WordIQ Encyclopedia, 2010, Neural Network – Definition (URL: [http://www.wordiq.com/definition/Neural\\_network](http://www.wordiq.com/definition/Neural_network), retrieved on 15-12-2010).
- [71] Cramér, H. (Ed.), 1999, *Mathematical Methods of Statistics* (USA: Princeton University Press).
- [72] DTREG [Software] Online Help, 2010, Multilayer Perceptron Neural Networks, (URL: <http://www.dtreg.com/mlfn.htm>, retrieved on 16-12-2010).
- [73] Atkinson, P. M., e Tatnall, A. R. L., 1997, Introduction Neural networks in remote sensing. *International Journal of Remote Sensing*, **18: 4**, 699-709.
- [74] Stergiou, C. e Siganos, D., 2007, *Neural Networks* (URL: [http://www.doc.ic.ac.uk/~nd/surprise\\_96/journal/vol4/cs11/report.html](http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html), retrieved on 15-12-2010).
- [75] Vliet, J. V., 2009, Map Comparison Kit 3 User Manual (Manual Version 3.2) [Software] (Maastricht, The Netherlands: Research Institute for Knowledge Systems BV).
- [76] Pontius, G. R., e Chen, H., 2006, *GEOMOD Modeling: Land-Use and Cover Change Modelling*. Clark University, USA.
- [77] Zadeh, L. A., 1965, Fuzzy Sets. *Information and Control*, **8**, 338-353.
- [78] Hagen, A., 2003, Fuzzy Set Approach to Assessing Similarity of Categorical Maps. *International Journal of Geographical Information Science*, **17: 3**, 235-249.
- [79] Yang, L., Huang, C., Homer, C.G., Wylie, B.K., e Coan, M.J., 2003, An Approach for Mapping Large-Area Impervious Surfaces: Synergistic Use of Landsat 7 ETM+ and High Spatial Resolution Imagery. *Canadian Journal of Remote Sensing*, **29:2**, pp. 230-240.
- [80] SHELTECH (Pvt.) Ltd. & DIAL Consultants Ltd., 2010, *Preparation of Detailed Area Plan (DAP) for DMDP, Location-10*. Final Plan Report, *Rajdhani Unnayan Kartripakkha* (RAJUK), Ministry of Housing and Public Works, Government of the People's Republic of Bangladesh.



## APPENDIX A

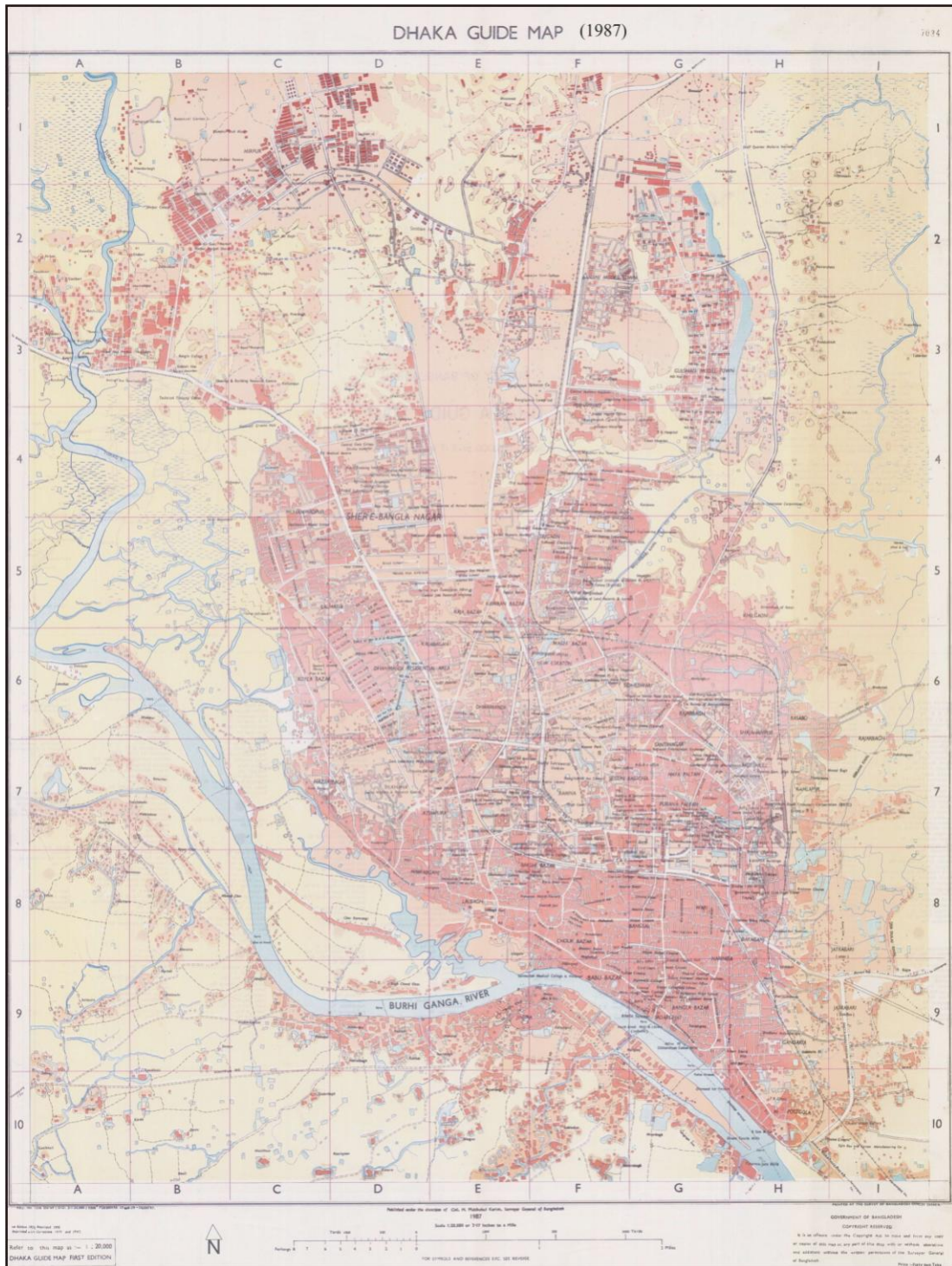
### Photographs: Some Existing Urban Problems within Dhaka City





# APPENDIX B

## Reference Maps

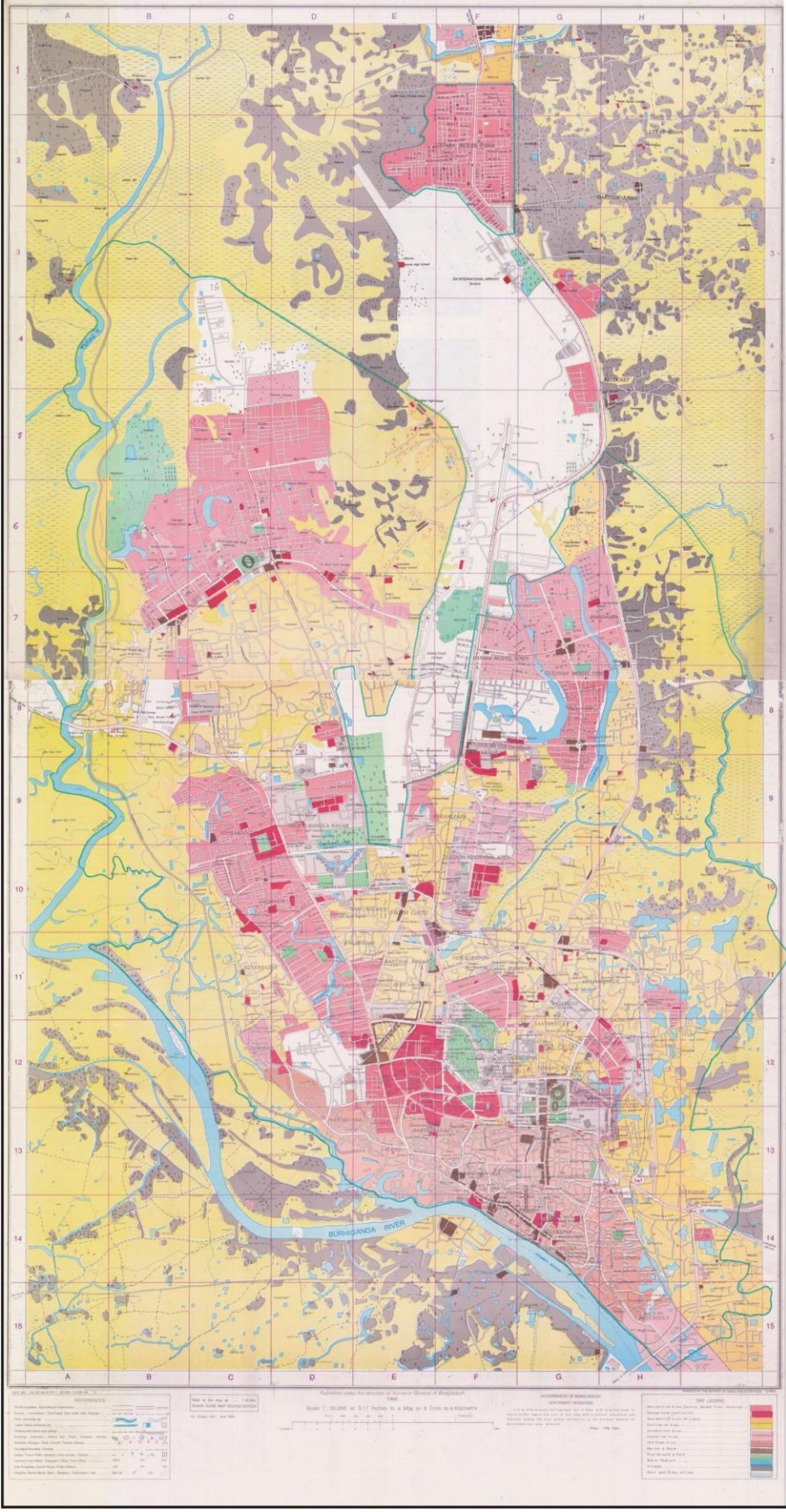


**Dhaka City Guide Map (1987)**



DHAKA GUIDE MAP (1995)

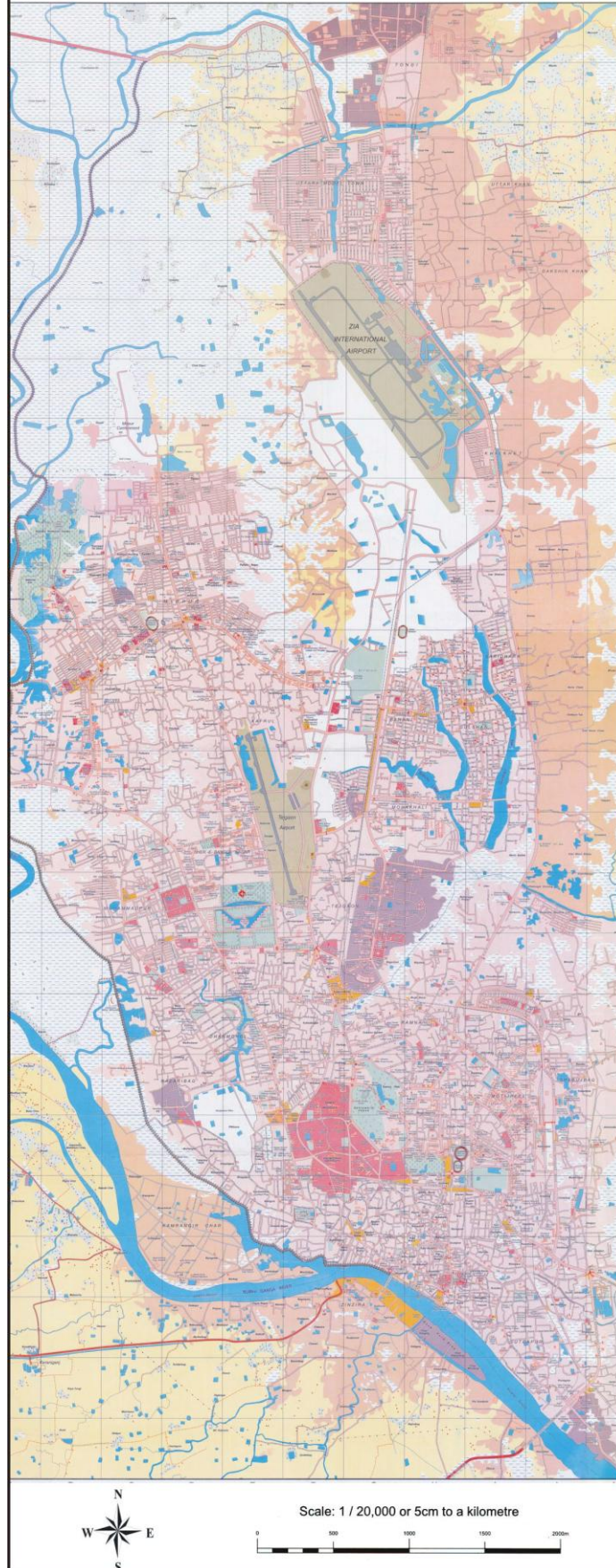
591



Dhaka City Guide Map (1995)



Dhaka City Guide Map (2001)



Dhaka City Guide Map (2001)

## APPENDIX C

---

### Details of Accuracy Assessment

#### C 1 Basic Terminologies

##### C 1.1 Ground Truth

Information acquired by field study for the purpose of calibration or verification of remotely sensed data [39].

##### C 1.2 Overall Accuracy

It represents the percentage of correctly classified pixels [39]. It is achieved by dividing the number of correct observations by the number of actual observations. It is represented by the following equation [42]:

$$O = \frac{\sum A}{\sum B} \times 100$$

Where, O = Overall Accuracy

A = the number of pixels assigned to the correct class

B = the number of pixels that actually belong to the class

##### C 1.3 Producer's Accuracy

It is a measure of how accurately the analyst classified the image data by category [42]. It shows what percentage of a particular ground class has been correctly classified. It is calculated by dividing the number of correct pixels for a class by the actual number of reference/ ground truth pixels for that class [39].

##### C 1.4 User's Accuracy

It is a measure of how well the classification performed in the field by category [42]. It tells what percentage of a class corresponds to the reference/ ground truth class. It is calculated by dividing the number of correct pixels for a class by the total pixels assigned to that class [39].

The details of calculations of overall accuracy, user's accuracy and producer's accuracy have been described in Table C 1. This is simply an example, not real data.

**Table C 1: An Example of Error Matrix**

Classified Data	Reference Data				
	A	B	C	D	Row Total
A	65	4	22	24	115
B	6	81	5	8	100
C	0	11	85	19	115
D	4	7	3	90	104
Column Total	75	103	115	141	434

**Land Cover Categories:**

Here, A = Arable Land, B = Bare Soil, C = Conifer Trees and D = Deciduous Trees.

$$\text{Overall Accuracy} = \frac{65+81+85+90}{434} = \frac{321}{434} = 74\%$$

**Producer's Accuracy**

$$A = \frac{65}{75} = 87\%$$

$$B = \frac{81}{103} = 79\%$$

$$C = \frac{85}{115} = 74\%$$

$$D = \frac{90}{141} = 64\%$$

**User's Accuracy**

$$A = \frac{65}{115} = 57\%$$

$$B = \frac{81}{100} = 81\%$$

$$C = \frac{85}{115} = 74\%$$

$$D = \frac{90}{104} = 87\%$$

**C 1.5 Kappa**

Kappa is a statistical measure of overall agreement between two maps (e.g. output of classified/simulated map and ground truth/reference map) or a matrix [39].

The Kappa statistic (K) is defined according to the following equation [43]:

$$K = \frac{P(A) - P(E)_{RL}}{1 - P(E)_{RL}}$$

Where, P(A) = Percentage of cells in the map that are identical; P(E)<sub>RL</sub> = Random Location (RL) conditional to the observed distribution in both maps.

The calculation of Kappa is based on the ‘Contingency Table’ or ‘Confusion Matrix’ [44]. It shows (Table C 2) how the distribution of categories in Map A differs from Map B [43].

**Table C 2: Contingency Table**

Map A categories	Map B categories				Total
	1	2	...	c	
1	$p_{11}$	$p_{12}$	...	$p_{1c}$	$p_{1.}$
2	$p_{21}$	$p_{22}$	...	$p_{2c}$	$p_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
c	$p_{c1}$	$p_{c2}$	...	$p_{cc}$	$p_{c.}$
Total	$p_{.1}$	$p_{.2}$	...	$p_{.c}$	1

Where,  $p_i$  = the proportion of c cells that is of category i in map A

$p_{.i}$  = the proportion of c cells that is of category i in map B

$p_{ij}$  = the proportion of cells that is of category i in map A and category j in map B

On the basis of the contingency table the following equations can be derived [44]:

$$P(A) = \sum_{i=1}^c p_{ii}$$

$$P(E)_{RL} = \sum_{i=1}^c p_{i.} p_{.i}$$

Kappa value ranges from -1 to +1. The maximum value +1 indicated, perfect agreement and the minimum value -1, indicates maximum disagreement.

Kappa value 0 indicates, the observed agreement matches the agreement expected by random arranging of all cells [43].

Kappa can be negative. This is a sign that there is no effective agreement between the two rates [45].

There are different qualitative classifications of degree of agreements for Kappa values [39]. One example is given in Table C 3 [46].

**Table C 3: Strength of Agreement for Kappa Statistic**

<b>Kappa Statistic</b>	<b>Strength of Agreement</b>
< 0	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect/Perfect

A number of variations of the Kappa statistic have been introduced by Pontius such as Kappa for no information (denoted as Kno) and Kappa for grid-cell level location (denoted as Klocation) [47].

### **C 1.6 Kno**

“Kno indicates the proportion classified correctly relative to the expected proportion classified correctly by a simulation with no ability to specify accurately quantity or location” [47].

It is a statistic similar to Kappa, but better capable of expressing similarity both in quantity and location [43].

$$K_{no} = \frac{P(A) - P(E)_{RC}}{1 - P(E)_{RC}}$$

Where,  $P(E)_{RC}$  = the fraction of expected agreement for the whole map, where RC stands for random category



### **C 1.7 Klocation**

“Klocation indicates the success due to the simulation’s ability to specify location divided by the maximum possible success due to a simulation’s ability to specify location perfectly” [47].

It depends strictly on the spatial distribution of the categories on the map. It indicates how well the grid cells are located on the landscape [47].

$$P(\max) = \sum_{i=1}^c \min(p_{i.}, p_{.i})$$
$$Klocation = \frac{P(A) - P(E)_{RL}}{P(\max) - P(E)_{RL}} = \frac{\sum_{i=1}^c (p_{ii} - p_{i.} p_{.i})}{\sum_{i=1}^c (\min(p_{i.}, p_{.i}) - p_{i.} p_{.i})}$$

Where,  $P(\max)$  = the maximum success rate

$P(E)_{RL}$  = the fraction of expected agreement for the whole map, where RL stands for random location

### **C 1.8 Khisto**

Khisto depends on the total number of cells taken in by each category [44].

It is a statistic that can be calculated directly from the histograms of two maps [43].

$$Khisto = \frac{P(\max) - P(E)_{RL}}{1 - P(E)_{RL}} = \frac{\sum_{i=1}^c \min(p_{i.}, p_{.i}) - \sum_{i=1}^c (p_{i.} p_{.i})}{1 - \sum_{i=1}^c (p_{i.} p_{.i})}$$

## C 2 Accuracy Assessment Report

Here 1= Builtup Area, 2= Water Body, 3= Vegetation, 4= Low Land and 5= Fallow Land. All these are applicable only for land cover class types.

### C 2.1 Accuracy Assessment Report (1989)

**Table C 4: Error Matrix (1989)**

Classified Data	Reference Data					Row Total
	1	2	3	4	5	
1	18	2	0	1	0	21
2	0	35	3	0	0	38
3	0	1	54	5	1	61
4	0	2	1	31	1	35
5	0	6	8	6	75	95
Column Total	18	46	66	43	77	<b>250</b>

**Table C 5: Accuracy Totals (1989)**

Class Name	Reference Totals	Classified Totals	Number Correct	Producer's Accuracy	User's Accuracy
1	18	21	18	100.00%	85.71%
2	46	38	35	76.09%	92.11%
3	66	61	54	81.82%	88.52%
4	43	35	31	72.09%	88.57%
5	77	95	75	97.40%	78.95%
<b>Totals</b>	<b>250</b>	<b>250</b>	<b>213</b>		

Overall Classification Accuracy = **85.20%**

**Table C 6: Conditional Kappa for each Category (1989)**

Class Name	Kappa
1	0.8461
2	0.9033
3	0.8441
4	0.8620
5	0.6958

Overall Kappa Statistics = **0.8054**.

**C 2.2 Accuracy Assessment Report (1999)**

**Table C 7: Error Matrix (1999)**

Classified Data	Reference Data					Row Total
	1	2	3	4	5	
1	60	1	4	1	6	72
2	1	20	0	0	3	24
3	1	3	47	1	1	53
4	0	1	0	29	0	30
5	0	3	5	2	61	71
Column Total	62	28	56	33	71	<b>250</b>

**Table C 8: Accuracy Totals (1999)**

Class Name	Reference Totals	Classified Totals	Number Correct	Producer's Accuracy	User's Accuracy
1	62	72	60	96.77%	83.33%
2	28	24	20	71.43%	83.33%
3	56	53	47	83.93%	88.68%
4	33	30	29	87.88%	96.67%
5	250	250	217	85.92%	85.92%
<b>Totals</b>	<b>250</b>	<b>250</b>	<b>217</b>		

Overall Classification Accuracy = **86.80%**

**Table C 9: Conditional Kappa for each Category (1999)**

Class Name	Kappa
1	0.7784
2	0.8123
3	0.8541
4	0.9616
5	0.8033

Overall Kappa Statistics = **0.8294**.

**C 2.3 Accuracy Assessment Report (2009)**

**Table C 10: Error Matrix (2009)**

Classified Data	Reference Data					Row Total
	1	2	3	4	5	
1	108	3	0	2	2	115
2	1	17	0	0	1	19
3	1	1	41	1	1	45
4	1	1	0	24	2	28
5	1	0	2	1	39	43
Column Total	112	22	43	28	45	<b>250</b>

**Table C 11: Accuracy Totals (2009)**

Class Name	Reference Totals	Classified Totals	Number Correct	Producer's Accuracy	User's Accuracy
1	112	115	108	96.43%	93.91%
2	22	19	17	77.27%	89.47%
3	43	45	41	95.35%	91.11%
4	28	28	24	85.71%	85.71%
5	45	43	39	86.67%	90.70%
<b>Totals</b>	<b>250</b>	<b>250</b>	<b>229</b>		

Overall Classification Accuracy = **91.60%**

**Table C 12: Conditional Kappa for each Category (2009)**

Class Name	Kappa
1	0.8897
2	0.8846
3	0.8926
4	0.8391
5	0.8866

Overall Kappa Statistics = **0.8821.**

## APPENDIX D

### Basic Terminologies Used for Land Cover Change Detection Analysis

#### D 1 Patch

The term patch defines scale-independent homogeneous regions in a landscape (e.g., grassland, forest, urban etc.) [49].

For this research purpose, 8-cell patch neighbourhood rule has been applied. The 8-cell rule considers all 8 adjacent cells, including the 4 orthogonal and 4 diagonal neighbours.

#### D 2 Landscape

It is defined by an interacting mosaic of patches relevant to the phenomenon under consideration (at any scale) [50]. Real landscapes contain complex spatial patterns in the distribution of resources that vary over time.

There are four basic types of landscape patterns (Figure D 1): spatial point patterns, linear network patterns, categorical or thematic map patterns and surface or continuous patterns [50]. This research deals with categorical map pattern.

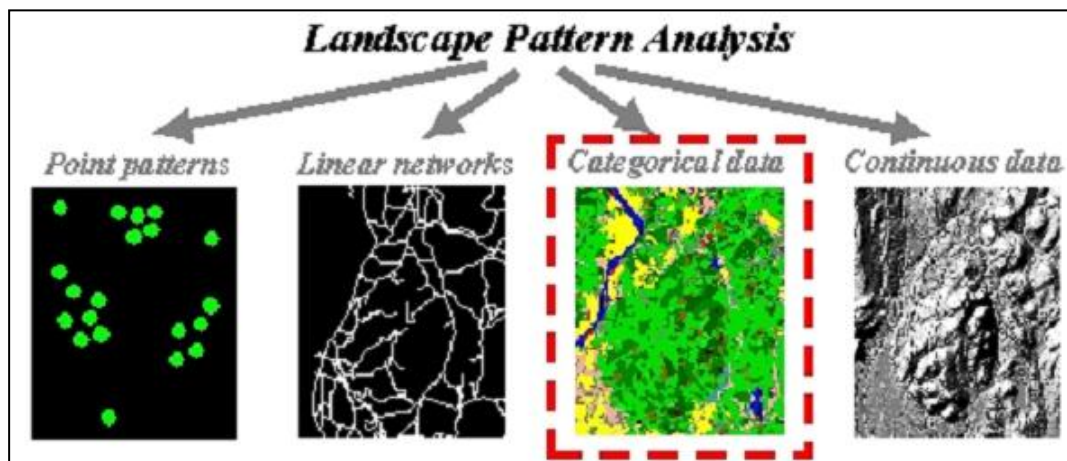


Figure D 1: Classes of Landscape Pattern [50]

#### D 3 Core Area

Core area is defined as the area within a patch beyond some specified depth-of-edge influence (i.e., edge distance) or buffer width [50]. Core area is affected by patch shape.

For example, given a 30 m cell size and a specified edge depth of 50 m, the mask will be rounded up to 2 cells (60 m) wide in the orthogonal directions. The non-orthogonal

directions will be rounded similarly, producing a near circular mask. Cells within the mask are eliminated from the ‘core’ of the patch. After all bounding cells are treated in this manner, the remaining cells not masked constitute the ‘core’ of the patch [50].

#### **D 4 Number of Patches (NP)**

NP equals the number of patches of the corresponding patch type (class). It is defined as follows [50]:

$$NP = n_i$$

Where,  $n_i$  = number of patches in the landscape of patch type (class) i

Number of patches of a particular class or land cover type is a simple measure of the extent of subdivision or fragmentation of the class.

#### **D 5 Edge Density (ED)**

ED equals the sum of the lengths (m) of all edge segments involving the corresponding patch type, divided by the total landscape area ( $m^2$ ), multiplied by 10,000 (to convert to hectares) [50].

$$ED = \frac{\sum_{k=1}^m e_{ik}}{A} (10,000)$$

Where,  $e_{ik}$  = total length (m) of edge in landscape involving patch type (class) i; includes landscape boundary and background segments involving patch type i

A = total landscape area ( $m^2$ ), it includes any internal background present

#### **D 6 Largest Patch Index (LPI)**

LPI equals the percentage of the landscape comprised by the largest patch. LPI equals the area ( $m^2$ ) of the largest patch of the corresponding patch type divided by total landscape area ( $m^2$ ), multiplied by 100 (to convert to a percentage). It is a simple measure of dominance. It is defined as follows [50]:

$$LPI = \frac{j=1^a \max(a_{ij})}{A} (100)$$

Where,  $a_{ij}$  = area ( $m^2$ ) of patch ij; A = total landscape area ( $m^2$ )

### **D 7 Core Area Percentage of Landscape (CPLAND)**

CPLAND equals the percentage the landscape comprised of core area of the corresponding patch type. CPLAND equals the sum of the core areas of each patch (m<sup>2</sup>) of the corresponding patch type, divided by total landscape area (m<sup>2</sup>), multiplied by 100 (to convert to a percentage). It is defined as follows [50]:

$$CPLAND = \frac{\sum_{j=1}^n a_{ij}^c}{A} (100)$$

Where,  $a_{ij}^c$  = core area (m<sup>2</sup>) of patch ij based on specified edge depths (m)

A = total landscape area (m<sup>2</sup>)

CPLAND facilitates comparison among landscape of varying size.

### **D 8 Mean Euclidean Nearest-Neighbour Distance (ENN MN)**

ENN equals the distance (m) to the nearest neighboring patch of the same type, based on shortest edge-to-edge distance. It is defined as follows [50]:

$$ENN_{MN} = h_{ij} * \frac{\sum_{i=1}^m \sum_{j=1}^n x_{ij}}{N}$$

Where,  $h_{ij}$  = distance (m) from patch ij to nearest neighboring patch of the same type (class), based on patch edge-to-edge distance, computed from cell center to cell center

MN = Mean; m = number of patch types; n = number of patches in the same class

$x_{ij}$  = urban patch ij and N = the total number of patches in the landscape

### **D 9 Mean Fractal Dimension Index (FRAC MN)**

FRAC equals 2 times the logarithm of patch perimeter (m) divided by the logarithm of patch area (m<sup>2</sup>); the perimeter is adjusted to correct for the raster bias in perimeter. It reflects shape complexity across a range of spatial scales (patch sizes). It is defined as follows [50]:

$$FRAC_{MN} = \frac{2 \ln(0.25 p_{ij})}{\ln a_{ij}} * \frac{\sum_{i=1}^m \sum_{j=1}^n x_{ij}}{N}$$

Where,  $p_{ij}$  = perimeter (m) of patch ij;  $a_{ij}$  = area (m<sup>2</sup>) of patch ij



## APPENDIX E

### Quantitative Values of Spatial Metrics for Change Detection Analysis

<b>1989</b>					
<b>Land Cover Type/ Spatial Metrics</b>	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>NP</b>	1198	2079	3132	1334	1964
<b>LPI</b>	3.4551	4.0222	3.5120	3.3454	15.4884
<b>ED</b>	21.0155	38.5187	68.1446	26.2272	74.8469
<b>CPLAND</b>	1.2407	1.6939	2.2958	3.4103	5.7454
<b>ENN_MN</b>	162.0899	137.0132	101.1428	153.7087	99.8811
<b>FRAC_MN</b>	1.0439	1.0425	1.0558	1.0431	1.0536

<b>1999</b>					
<b>Land Cover Type/ Spatial Metrics</b>	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>NP</b>	1861	549	1553	320	2552
<b>LPI</b>	15.0426	2.7725	4.7579	4.2360	5.2944
<b>ED</b>	54.1470	15.4297	38.0826	13.5453	71.0317
<b>CPLAND</b>	5.6707	2.8383	5.0149	5.0204	2.7939
<b>ENN_MN</b>	116.3086	196.2941	132.6763	198.6695	97.4143
<b>FRAC_MN</b>	1.0494	1.0483	1.0454	1.0500	1.0573

<b>2009</b>					
<b>Land Cover Type/ Spatial Metrics</b>	<b>Builtup Area</b>	<b>Water Body</b>	<b>Vegetation</b>	<b>Low Land</b>	<b>Fallow Land</b>
<b>NP</b>	2096	658	1586	425	3379
<b>LPI</b>	27.6267	1.9408	4.4650	3.7163	4.1399
<b>ED</b>	77.8558	15.4560	37.1452	15.4459	56.2319
<b>CPLAND</b>	11.6378	1.6657	3.7745	4.3154	1.6794
<b>ENN_MN</b>	89.4140	184.1595	127.4179	163.8193	104.0553
<b>FRAC_MN</b>	1.0502	1.0501	1.0509	1.0502	1.0556

## APPENDIX F

### Summary of Land Cover Classification Statistics (1989 - 2009)

Land Cover Type	1989		1999		Change in Area (%) [1989-1999]
	Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%	
<b>Builtup Area</b>	37.6569	8.447	129.2292	28.991	<b>+20.544</b>
<b>Water Body</b>	67.3218	15.102	42.5034	9.535	<b>-5.567</b>
<b>Vegetation</b>	109.5714	24.581	94.3929	21.175	<b>-3.406</b>
<b>Low Land</b>	62.0640	13.923	53.7264	12.052	<b>-1.871</b>
<b>Fallow Land</b>	169.1559	37.947	125.9181	28.247	<b>-9.70</b>
Total	<b>445.770</b>	<b>100</b>	<b>445.770</b>	<b>100</b>	<b>0</b>

Land Cover Type	1999		2009		Change in Area (%) [1999-2009]
	Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%	
<b>Builtup Area</b>	129.2292	28.991	204.4008	45.853	<b>+16.862</b>
<b>Water Body</b>	42.5034	9.535	33.6645	7.552	<b>-1.983</b>
<b>Vegetation</b>	94.3929	21.175	79.9578	17.937	<b>-3.238</b>
<b>Low Land</b>	53.7264	12.052	49.9914	11.215	<b>-0.837</b>
<b>Fallow Land</b>	125.9181	28.247	77.7555	17.443	<b>-10.804</b>
Total	<b>445.770</b>	<b>100</b>	<b>445.770</b>	<b>100</b>	<b>0</b>

Land Cover Type	1989		2009		Change in Area (%) [1989-2009]
	Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%	
<b>Builtup Area</b>	37.6569	8.447	204.4008	45.853	<b>+37.406</b>
<b>Water Body</b>	67.3218	15.102	33.6645	7.552	<b>-7.55</b>
<b>Vegetation</b>	109.5714	24.581	79.9578	17.937	<b>-6.644</b>
<b>Low Land</b>	62.0640	13.923	49.9914	11.215	<b>-2.708</b>
<b>Fallow Land</b>	169.1559	37.947	77.7555	17.443	<b>-20.504</b>
Total	<b>445.770</b>	<b>100</b>	<b>445.770</b>	<b>100</b>	<b>0</b>

## APPENDIX G

---

### Land Use Suitability

#### **G 1 Land Use Suitability Analysis**

According to Collins *et al* (2001), “*Land-use suitability analysis aims at identifying the most appropriate spatial pattern for future land uses according to specific requirements, preferences, or predictors of some activity*” [65].

In general, this is a GIS based method used to find out the best location for a particular land use in a given area.

#### **G 2 Different Methods for Land Use/Cover Suitability Analysis**

According to Malczewski (2004), there are three major groups of approaches of GIS based land use suitability analysis [18]:

- a) Computer-Assisted Overlay Mapping
- b) Multi Criteria Evaluation (MCE) Methods, and
- c) Artificial Intelligence (Soft Computing or Geo-computation) Methods

#### **G 2.1 Computer-Assisted Overlay Mapping Technique**

In the early years, there were manual techniques for map overlaying. Computer-Assisted Overlay Mapping Technique has upgraded from the traditional overlaying technique. This technique is useful for overlaying large data and maps. This kind of technique is applied in the form of Boolean operations and Weighed Linear Combination (WLC) within the GIS environment. This technique is popular because it is easy to understand and implement. The main drawback of this model is failing to choose appropriate assumptions/criteria for suitability analysis [18].

#### **G 2.2 Multi Criteria Decision Making Methods**

The limitation of the above technique can be solved using the Multi Criteria Decision Making (MCDM) methods. The MCDM procedures combine three things [18]:

- (i) The geographical data
- (ii) The decision maker’s preferences, and

(iii) Manipulation of the data and preferences as per some specific decision rules.

There are two types of decision rules:

**a) Multi-Objective Methods:** These methods are mathematical programming based models. It consists of two or more objective functions and a set of constraints [18]. To solve the problems related to multi-objective ‘Multi-Objective Land Allocation (MOLA)’ module has been developed [36]. To run the MOLA module the specified objectives, their relative weights, the ranked suitability maps for each and the areas that should be allocated to each; are required. The module then undertakes the iterative procedure of allocating the best ranked cells to each objective according to the areal goals and resolving the conflicts [36].

**b) Multi-Attribute Methods:** There have been many multi-attributes (sometimes simply known as Multi-criteria) methods used in GIS field. These methods are much popular and are now being widely used in different analysis and decision making processes. There are many procedures under these methods like-the Boolean operations, WLC, Ordered Weighted Averaging (OWA), and Analytical Hierarchy Analysis (AHP).

Among these, Weighed Linear Combination (WLC) is most popular. It is based on the weighted average method. The researcher/decision maker puts weight as per practical experience or importance for each criterion/attribute of the data layer [18].

The problem related to this method is there is standard format for selecting the criteria for a particular problem for a given area. The decision maker has to choose the criteria. Therefore, the result may vary for solving the same problem. Moreover, it is also a problem, trying to fix which method is suitable for a particular situation [18].

### **G 2.3 Artificial Intelligence (AI) Methods**

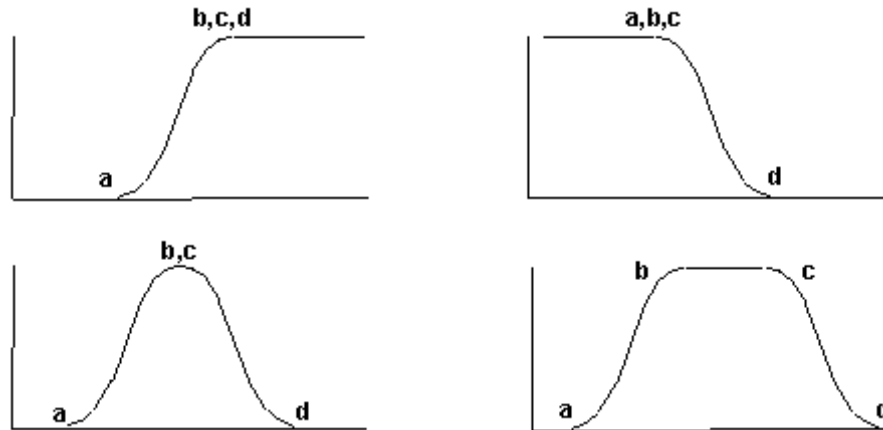
The problems related to MCDM methods could be minimized using AI methods. AI incorporates modern mathematical techniques to describe complex situations and decision making for a particular case. There are many approaches with the integration of GIS and AI techniques such as Fuzzy Logic Techniques, Artificial Neural Networks (ANN), Evolutionary (genetic) Algorithms and Cellular Automata (CA) [18] etc.

## APPENDIX H

---

### H 1 Sigmoid Function

The Sigmoid (S-shaped) function is commonly used in Fuzzy Set theory. It is a cosine function. It requires 4 co-ordinates along with the X-axis to construct the shape of the curve. These are indicated in Figure H 1 [66]:



**Figure H 1: Sigmoid Function [66]**

Here, a = membership rises above 0; b = membership becomes 1; c = membership falls below 1 and d = membership becomes 0.

### H 2 Cramer's V

Cramer's V is a statistic measuring the strength of association or dependency between two categorical variables. Cramer's V ranges from 0 to 1. The value close to 0 indicates little association between the variables while close to 1 shows strong association [66]. It is defined as [71]:

$$V = V(X, Y) = \sqrt{\frac{\chi^2}{n * \min(M - 1, N - 1)}}$$

Where, V = Cramer's V

X and Y are two categorical variables; X has M distinct classes and Y has N distinct classes

$\chi^2$  = Chi-squared statistic;  $\min(M-1, N-1)$  = the minimum of the two values.



# Masters Program in **Geospatial Technologies**

**URBAN LAND COVER CHANGE DETECTION ANALYSIS AND  
MODELLING SPATIO-TEMPORAL GROWTH DYNAMICS  
USING REMOTE SENSING AND GIS TECHNIQUES**

**“A CASE STUDY OF DHAKA, BANGLADESH”**

**BAYES AHMED**