

## Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use” that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

**Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation**

Printing note: If you do not wish to print this page, then select “Pages from: first page # to: last page #” on the print dialog screen



The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

## **ABSTRACT**

### **PROJECTING LAND USE CHANGES USING PARCEL-LEVEL DATA: MODEL DEVELOPMENT AND APPLICATION TO HUNTERDON COUNTY, NEW JERSEY**

**Florencio Ballesteros, Jr.**

#### **Abstract**

This dissertation is to develop a parcel-based spatial land use change prediction model by coupling various machine learning and interpretation algorithms such as cellular automata (CA) and decision tree (DT). CA is a collection of cells that evolves through a number of discrete time steps according to a set of transition rules based on the state of each cell and the characteristics of its neighboring cells. DT is a data mining and machine learning tool that extracts the patterns of decision process from observed cell behaviors and their affecting factors. In this dissertation, CA is used to predict the future land use status of cadastral parcels based on a set of transition rules derived from a set of identified land use change driving factors using DT. Although CA and DT have been applied separately in various land use change models in the literature, no studies attempted to integrate them. This DT-based CA model developed in this dissertation represents the first kind of such integration in land use change modeling. The coupled model would be able to handle a large set of driving factors and also avoid subjective bias when deriving the transition rules. The coupled model uses the cadastral parcel as a unit of analysis, which has practical policy implications because the responses of land use changes to various policy usually take place at the parcel level. Since parcel varies by their sizes and shapes, its use as a unit of analysis does make it difficult to apply CA, which initially designed to handle regular grid cells. This dissertation improves the treatment of the irregular cell in CA-based land use change models in literature by defining a cell's neighborhood as a fixed distance buffer along the parcel boundary.

The DT-based CA model was developed and validated in Hunterdon County, New Jersey. The data on historical land uses and various land use change driving factors for Hunterdon County were collected and processed using a Geographic Information

System (GIS). Specifically, the county land uses in 1986, 1995 and 2002 were overlaid with a parcel map to create parcel-based land use maps. The single land use in each parcel is based on a classification scheme developed thorough literature review and empirical testing in the study area. The possible land use status considered for each parcel is agriculture, barren land, forest, urban, water or wetlands following the land use/land cover classification by the New Jersey Department of Environment Protection. The identified driving factors for the future status of the parcel includes the present land use type, the number of soil restrictions to urban development, and the size of the parcel, the amount of wetlands within the parcel, the distribution of land uses in the neighborhood of the parcel, the distances to the nearest streams, urban centers and major roads. A set of transition rules illustrating the land use change processes during the period 1986-1995 were developed using a TD software J48 Classifier. The derived transition rules were applied to the 1995 land use data in a CA model Agent Analyst/RePast (**R**ecursive **P**orous **A**gent **S**imulation **T**oolkit) to predict the spatial land use pattern in 2004, which were then validated by the actual land use map in 2002. The DT-based CA model had an overall accuracy of 84.46 percent in terms of the number of parcels and of 80.92 percent in terms of the total acreage in predicting land use changes. The model shows much higher capacity in predicting the quantitative changes than the locational changes in land use. The validated model was applied to simulate the 2011 land use patterns in Hunterdon County based on its actual land uses in 2002 under both “business as usual” and policy scenarios. The simulation results shows that successfully implementing current land use policies such as down zoning, open space and farmland preservation would prevent the total of 7,053 acres (741 acres of wetlands, 3,034 acres of agricultural lands, 250 acres of barren land, and 3,028 acres of forest) from future urban development in Hunterdon County during the period 2002-2011. The neighborhood of a parcel was defined by a 475-foot buffer along the parcel boundary in the study. The results of sensitivity analyses using two additional neighborhoods (237- and 712-foot buffers) indicate the insignificant impacts of the neighborhood size on the model outputs in this application.

**PROJECTING LAND USE CHANGES USING PARCEL-LEVEL DATA:  
MODEL DEVELOPMENT AND APPLICATION TO HUNTERDON COUNTY,  
NEW JERSEY**

**by  
Florencio C. Ballesteros, Jr.**

**A Dissertation  
Submitted to the Faculty of  
New Jersey Institute of Technology  
And  
Rutgers, The State University of New Jersey  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Environmental Science**

**Department of Chemistry and Environmental Science**

**May 2008**

Copyright © 2008 by Florencio C. Ballesteros, Jr.

**ALL RIGHTS RESERVED**

APPROVAL PAGE

PROJECTING LAND USE CHANGES USING PARCEL-LEVEL DATA:  
MODEL DEVELOPMENT AND APPLICATION TO HUNTERDON COUNTY,  
NEW JERSEY

Florencio C. Ballesteros, Jr.

4/11/2008

---

Dr. Zeyuan Qiu, Dissertation Advisor  
Assistant Professor of Chemistry and Environmental Science, NJIT

Date

4/11/08

---

Dr. Maurie Cohen, Committee Member  
Associate Professor of Chemistry and Environmental Science, NJIT

Date

April 11, 2008

---

Dr. Daniel J. Watts, Committee Member  
Executive Director, York Center for Environmental Engineering and Science, NJIT

Date

4/11/08

---

Dr. Richard G Lathrop, Jr., Committee Member  
Professor of Ecology, Evolution, & Natural Resources, Rutgers University

Date

11 April 2008

---

Dr. Daniel Van Abs, Committee Member  
Senior Director of Planning and Science, New Jersey Highlands Council, Chester, NJ

Date

## **BIOGRAPHICAL SKETCH**

**Author:** Florencio C. Ballesteros, Jr.

**Degree:** Doctor of Philosophy

**Date:** May 2008

### **Undergraduate and Graduate Education:**

- Doctor of Philosophy in Environmental Science,  
New Jersey Institute of Technology, Newark, NJ, 2008
- Master of Science in Environmental Engineering,  
University of Illinois, Champaign, IL, 1998
- Bachelor of Science in Chemical Engineering,  
Saint Louis University, Baguio City, Philippines, 1982

**Major:** Environmental Science

## ACKNOWLEDGMENT

Successfully defending a Ph.D. dissertation is a highlight in one's academic career. It is also something accomplished with the aid of numerous people.

Foremost among them was Dr. Zeyuan Qiu, my advisor. In guiding me to the completion of my dissertation he was an exemplar of patience and encouragement, hallmarks of a great teacher. Dr. Qiu was not only an advisor but a mentor as well. I will never be able to thank him sufficiently.

My profound thanks go to Dr. Daniel Watts, who made it possible for me to come to the New Jersey Institute of Technology to pursue my Ph.D. degree through partial support in the form of teaching and research assistantships. I will always remember his eager readiness to help.

I was privileged and delighted to discuss the results of my research with Dr. Richard Lathrop and Dr. Daniel Van Abs. In spite of their demanding workloads, they generously shared time with me to offer insights and advice throughout my research. I also thank Dr. Maurie Cohen for his many insights on improving my manuscript, and in particular for staying up late for my dissertation defense.

I gratefully acknowledge the assistance of John Tyrawsky of the New Jersey Department of Environmental Protection, Bureau of GIS who helped me interpret land use data from their website. Many thanks also to Patricia Leidner, GIS Coordinator of the Hunterdon County Office of GIS who provided me with land use and parcel maps of the County. As well, I am grateful to Jen Zhang and Robert O'Neil of the New Jersey Water Supply Authority for the references on build out analysis.

I am also grateful to the New Jersey Water Resources Research Institute, which provided a grant to partially fund my dissertation. I acknowledge as well the Faculty and Staff of the Department of Chemistry and Environmental Science. They were supportive throughout my work.

A special recognition is given to my lab mates, Sherry Kuo and Noor Dhaliwal. Their encouragement was always valuable and will always be appreciated.

My deepest gratitude I reserve for my family. I felt always strengthened by their prayers. Their optimism and cheerfulness helped me to remain positive and joyful. Finally, I thank the many friends – too numerous to mention here – for moral support they provided. I remember you all with deep gratitude.

## TABLE OF CONTENTS

Chapter	Page
1 INTRODUCTION.....	1
1.1 Definitions: Land Use, Land Cover, Land Use/Cover Changes.....	6
1.1.1 Land Cover and Land Use.....	7
1.1.2 Land Use Change and Land Cover Change.....	10
1.1.3 Land Use Modeling.....	11
1.2 Research Priorities in Land Use Change Modeling.....	14
1.3 Research Objective .....	15
1.4 Research Contributions.....	17
1.5 Assumptions.....	19
1.6 Modeling Procedure.....	20
1.6.1 Preparing Parcel-based Maps.....	22
1.6.2 Generating Transition Rules Using Decision Tree.....	23
1.6.3 Predicting Land Use Change Through Cellular Automata.....	23
1.6.4 Model Validation.....	24
1.7 Expected Output.....	24
1.7.1 Parcel-based Maps.....	25
1.7.2 Workable Model for Land Use Change in Hunterdon County, New Jersey.....	25
1.7.3 Protocol for Map Comparison.....	25
1.7.4 Land Use Impacts of Various Land Use Policy Scenarios.....	26

**TABLE OF CONTENTS**  
**(Continued)**

<b>Chapter</b>	<b>Page</b>
1.8 Dissertation Structure.....	26
<b>2 LITERATURE REVIEW.....</b>	<b>28</b>
2.1 Elements of Cellular Automata (CA).....	28
2.2 CA Applications in Land Use Change Modeling.....	31
2.2.1 Spatiality.....	32
2.2.2 Decentralization.....	32
2.2.3 Affinity with Techniques of Spatial Analysis Notably GIS.....	33
2.2.4 Simplicity and Ability to Integrate Micro and Macro Approaches.....	33
2.2.5 Examples of CA Land Use Change Models.....	33
2.3 Cell Shape and Neighborhood Configuration.....	35
2.3.1 Using Irregular Cell Shapes.....	35
2.3.2 Defining an Appropriate Neighborhood .....	38
2.4 Driving Factors of Land Use Change.....	42
2.4.1 Neighborhood Factors.....	42
2.4.2 Accessibility Factors.....	43
2.4.3 Suitability Factors.....	43
2.4.4 Policy Controls and Socio-economic Factors.....	44
2.5 Generating Transition Rules.....	46
2.5.1 Regression Method.....	46
2.5.2 Artificial Neural Network-based Method.....	47

**TABLE OF CONTENTS**  
**(Continued)**

<b>Chapter</b>	<b>Page</b>
2.5.3 Multiple Criteria Evaluation-Analytical Hierarchy Process Method .....	48
2.5.4 Fuzzy Logic.....	49
2.6 Decision Tree.....	51
2.6.1 Generating the Decision Tree.....	51
2.6.2 From Decision Tree to Transition Rule.....	54
2.6.3 Advantages of Using Decision Tree.....	55
2.7 Accuracy Assessment.....	56
2.8 Summary.....	61
3 METHODOLOGY.....	63
3.1 Developing a Parcel-Based Land Use Map.....	63
3.2 Neighborhood Configurations.....	68
3.3 Model Components.....	72
3.3.1 Decision Tree Module: Generating the Transition Rule.....	73
3.3.2 Cellular Automata Module: Predicting Future Land Use/Cover Change...	78
3.4 Model Integration.....	80
3.4.1 The Decision Tree Algorithm.....	81
3.4.2 Configuring Agent Analyst Prior to Program Execution.....	82
3.4.3 Agent Analyst Program Execution.....	84
3.4.4 Using Decision Trees in Agent Analyst .....	92
3.5 Model Validation.....	94

**TABLE OF CONTENTS**  
**(Continued)**

<b>Chapter</b>	<b>Page</b>
3.6 Sensitivity Analysis.....	96
3.7 Summary.....	97
<b>4 STUDY AREA PROFILE AND DATA PROCESSING.....</b>	<b>98</b>
4.1 Hunterdon County, New Jersey.....	98
4.1.1 Residential Development.....	101
4.1.2 Commercial and Industrial Development.....	101
4.1.3 Residential Development-related Policies.....	103
4.2 Land Use Policies.....	104
4.2.1 Large-lot Zoning.....	104
4.2.2 Open Space Zoning.....	105
4.2.3 Farmland Preservation.....	106
4.2.3.1 Purchase of Development Rights (PDR).....	108
4.2.3.2 County and Municipal Planning Incentive Grants (PIG).....	108
4.2.3.3 Municipality-Approved Farmland Preservation Program (MAFPP)	109
4.2.3.4 Fee Simple .....	109
4.2.3.5 Direct Easement and Emergency Easement Purchase.....	110
4.2.4 Open Space Development.....	110
4.2.5 Highlands Water Protection and Planning Act.....	111
4.3 Parcel-based Land Uses.....	112
4.4 Understanding Urban Land Conversions.....	118

**TABLE OF CONTENTS**  
**(Continued)**

<b>Chapter</b>	<b>Page</b>
4.5 Driving Factors of Land Use Change.....	122
4.5.1 Soil.....	123
4.5.2 Slope.....	127
4.5.3 Proximity to Transportation Networks and Urban Centers.....	129
4.5.4 Distance to Streams.....	133
4.6 Summary.....	134
5 RESULTS AND DISCUSSION.....	136
5.1 Transition Rules.....	137
5.2 Land Use Change Model Validation.....	141
5.2.1 Land Use Prediction for 2004.....	141
5.2.2 Validation.....	144
5.2.2.1 Overall Prediction Accuracy.....	144
5.2.2.2 Kappa Index and Its Variants .....	148
5.3 Simulation of Land Use Change.....	156
5.3.1 Baseline (Business as Usual) Scenario.....	156
5.3.1 Simulated Policy Scenario.....	159
5.4 Sensitivity Analysis.....	166
5.5 Summary.....	167
6 SUMMARY AND CONCLUSION.....	169
6.1 Summary of Approach.....	169

**TABLE OF CONTENTS**  
**(Continued)**

<b>Chapter</b>	<b>Page</b>
6.3 Summary of Findings.....	171
6.2.1 Reliability of the Decision Tree-based CA Model.....	171
6.2.2 Effect of Driving Factors.....	173
6.2.3 Suitability of the Parcel as Unit of Analysis.....	174
6.2.4 Land Use Outcome of Policy Implementation .....	174
6.3 Contributions of this Research.....	175
6.4 Directions for Further Research.....	177
6.5 Conclusions.....	179
APPENDIX A SCRIPT FOR CALCULATING NEIGHBORHOOD LAND USE DISTRIBUTION.....	182
APPENDIX B DRIVERS OF LAND USE CHANGE .....	184
APPENDIX C J48 PRUNED DECISION TREE.....	186
REFERENCES .....	204

## LIST OF TABLES

<b>Table</b>	<b>Page</b>
3.1 An Example of Input Training Data for Decision Tree .....	73
3.2 Calculation of Kappa and its Variants .....	95
4.1 Comparison of the Total Areas of the Re-classified Land Uses in Parcels to the Original Land Uses in NJDEP Land Use/Cover Maps in 1986, 1995 and 2002 .....	113
4.2 The Flow of Land Use Changes in Acres in Hunterdon County during the 1986 - 1995 and 1995 - 2002 Periods .....	118
4.3 Driving Factors for Modeling Land Use Change .....	123
5.1 The Predicted Land Use Distribution for 2004 in Terms of the Number of Parcels .	142
5.2 The Land Use Pattern for 2004 in Terms of the Total Acreage .....	142
5.3 The Basic Statistics of the Converted Parcel Size by Land Uses .....	144
5.4 Confusion Matrix for Evaluating Accuracy in Terms of Number of Parcels and Parcel Area .....	147
5.5 Modified Confusion Matrix Expressed in Proportionality Values .....	150
5.6 Estimation of Kappa and Its Variants .....	151
5.7 The Predicted Land Use Change Pattern During 2002-2011 in Terms of the Number of Parcels in the Baseline Scenario .....	157
5.8 The Predicted Land Use Change Patterns During 2002-2011 in Terms of the Total Acreage in the Baseline Scenario .....	157
5.9 The Predicted Land Use Change Pattern Under the Policy Scenario in Terms of the Number of Parcels.....	160
5.10 The Predicted Land Use Change Pattern During the 2002-2011 Period in the Policy Scenario in Terms of Total Acreage .....	160
5.11 Differential Land Use Matrix Between Baseline and Policy Scenarios .....	161

**LIST OF TABLES**  
**(Continued)**

<b>Table</b>	<b>Page</b>
5.12 The Impacts of Two Neighborhood Sizes on the CA Modeling Accuracy and the Resulting Kappa Index .....	167

## LIST OF FIGURES

Figure	Page
1.1 Exurban area in New Jersey .....	2
1.2 A flowchart of the modeling procedure .....	2
2.1 One-dimensional cellular automata .....	29
2.2 Two-dimensional cellular automata .....	31
2.3 Construction of a Voronoi (Thiessen) Diagram .....	37
2.4 Neighborhood configurations in formal cellular automata .....	40
2.5 Neighborhood configurations for irregular polygons used in CA modeling .....	41
2.6 Decision tree with 27 terminal nodes .....	53
2.7 Confusion matrix for two maps with common measures of classification .....	59
3.1 Land use classification criteria .....	67
3.2 Neighborhood configurations with a buffer from the centroid (a) and the edge (b) of a parcel .....	69
3.3 Neighborhood configurations with buffer from the centroid for a small (a) and big (b) parcel .....	70
3.4 Neighborhood configuration as defined in this study .....	71
3.5 A sample J48 decision tree output .....	75
3.6 Rule Set from J48 Decision Tree .....	78
3.7 Integration steps .....	81
3.8 Retrieving the model .....	85
3.9 Opening the model dialogue box .....	86
3.10 Editing model steps .....	87

**LIST OF FIGURES  
(Continued)**

<b>Figure</b>	<b>Page</b>
3.11 Actions editor dialogue box (a) .....	88
3.12 Actions editor dialogue box (b) .....	89
3.13 Actions editor dialogue box (c) .....	90
3.14 Compiling the model .....	91
3.15 Model run .....	91
3.16 Commence simulation dialogue box .....	92
3.17 Step definition in Agent Analyst for transition rule .....	92
4.1 Hunterdon County Map .....	99
4.2 Hunterdon County landscape change 1984-2001 .....	106
4.3 Hunterdon County land use 2002 .....	114
4.4 Comparison of the total areas of the re-classified land uses among three years in Hunterdon County .....	116
4.5 Transformation from urban land to barren land: transitional areas .....	119
4.6 Transformation from urban land to forest: old field .....	120
4.7 Transformation from urban: rural residential to agriculture: other agriculture .....	121
4.8 Transformation from urban: rural residential to wetland .....	122
4.9 Soil suitability for development in Hunterdon County .....	125
4.10 Location of new urban areas relative to soil suitability .....	127
4.11 Urbanization along roads and highways .....	130
4.12 Distance of new urban areas relative to nearest local road .....	131

**LIST OF FIGURES  
(Continued)**

<b>Figure</b>	<b>Page</b>
4.13 Distance of new urban areas to nearest urban center .....	132
4.14 The distribution of the distance of the urbanized parcel during the 1995-2002 period to the nearest streams .....	134
5.1 Model timeline schematic .....	137
5.2 Spatial distribution of newly urbanized areas during the 1995-2004 period versus the soil suitability for development .....	154
5.3 Spatial distribution of misclassified land parcels in the 2004 predicted land use pattern .....	155
5.4 Projected new urban areas for 2011 in the baseline scenario .....	158
5.5 Comparative baseline and policy scenario predictions .....	163
5.6 Location of newly urbanized areas in 2011 compared to suitability of soils for development .....	165

# CHAPTER 1

## INTRODUCTION

The landscape in the United States has been experiencing rampant changes over the last several decades. Based on data assembled by the Natural Resource Conservation Service (2003), the nation's cropland comprised 368 million acres in 2003 which is about 12 percent lower than the 420 million acres in 1982. The percentage of total non-cultivated cropland increased by 11 percent from 44 million acres in 1982 to 58 million acres in 2003. In 2003, developed land was estimated at 108.1 million acres, up from 98.3 million acres in 1997. An underlying cause of these changes is the growth of urban areas where the demand for urbanized land is chiefly supplied at the expense of forest and agricultural land.

City centers have been historically densely populated because workers and their families tended to live in multi-family dwellings near their areas of employment. As the economy grew, a financially empowered middle class began to buy single-family residences away from the city center (Jackson, 1985). Initially limited to the surrounding communities of urban centers, this early stage of "suburbanization" steadily spread, aided by increased car ownership (Muller, 1986), and the passing of the Federal Interstate Highway Act of 1950 that created multi-lane highways. Aided by these and other factors such as zoning ordinances and environmental conditions, suburbs continued to expand resulting in "leapfrog" development, or growth that occurs around lakes or areas designated as preserved open space (Cutsinger and Galster, 2006; Wu and Plantinga, 2003). Simultaneously, cities in the United States became "deindustrialized" as

legislators sought to alleviate pollution in the urban areas. Increasing environmental awareness accelerated the deindustrialization trend (Bjelland, 2004, Kowala-Stamm, 2006). Deindustrialization also led to the suburbanization of employment as industries and corporate offices relocated to the suburbs (Gong and Wheeler, 2002). As the suburbs began to exhibit the same negative urban aspects of the older cities (Yang and Jarkowsky, 2006; Oh, 2005; Gordon and Richardson, 2004), development was moved farther out from the urban core. Exurbanization emerged as a new type of development, i.e. residential developments in large lots (typically up to 20 acres) in rural settings (Davis, et al., 1994; Nelson, 1992). Figure 1.1 gives a glimpse of the growing number of residences in New Jersey with on-site horses. Together with suburbanization, exurbanization makes up what is called urban sprawl, development that blurs the traditional rural and urban interfaces and complicates land use planning and management.



**Figure 1.1** Exurban area in New Jersey. (Source: Downloaded from <http://www.sellhorseproperties.com/ForSale/equestrianproperties/6957595.jpg> November 2, 2007)

These urbanizing developments came at the expense of other land cover types such as forest and agricultural lands. Depending on government regulations in the

locality, these changes manifested themselves differently in various states. In some parts of the United States, agricultural land became the ready source of land conversions. In other locations, marginal or non-agricultural land supplied the demand for land used for urban development (Stansfield, 1998; Matlack, 1997; Abler et al., 1976).

As a part of the dynamic New York metropolitan area, northern and central New Jersey has been experiencing some of the country's most dramatic changes in land use and development to accommodate economic and population growth. According to land cover data for 2001<sup>1</sup>, approximately 1.48 million acres in New Jersey equivalent to 31 percent of the state's total is developed. This number represents a 23 percent increase from 1.2 million acres in 1984. The development of previously barren lands characterized by thin soil, sand, or rocks and generally low vegetation grew by 21,000 acres during the same period. Much of this increase came from cultivated grassland that lost approximately 157,000 acres, forest land contributed approximately 77,000 acres, and wetlands contributed a further 63,000 acres.<sup>2</sup>

The amount of land annually converted to development now stands at 18,000 acres (or 28 square miles), which is roughly two times the size of Jersey City (NJ Facts, 2003). At this rate, it is expected that New Jersey will be completely built out in 40 years (Hasse and Lathrop, 2001).

Land use changes do not only bring about significant social and economic changes, but they also have profound impacts on human health and the natural

---

<sup>1</sup> Classification and analysis of Landsat satellite image data. Grant F. Walton Center for Remote Sensing & Spatial Analysis, Rutgers, The State University of New Jersey. Downloaded January 31, 2008, from [http://www.crssa.rutgers.edu/projects/lc/download/reportsdata00\\_01/landuselancov\\_web.pdf](http://www.crssa.rutgers.edu/projects/lc/download/reportsdata00_01/landuselancov_web.pdf).

<sup>2</sup> Classification and analysis of Landsat satellite image data. Grant F. Walton Center for Remote Sensing & Spatial Analysis, Rutgers, The State University of New Jersey. Downloaded January 31, 2008, from [http://www.crssa.rutgers.edu/projects/lc/download/reportsdata00\\_01/landuselancov\\_web.pdf](http://www.crssa.rutgers.edu/projects/lc/download/reportsdata00_01/landuselancov_web.pdf).

environment. In particular, urban sprawl has tremendous health impacts primarily because of people's dependence on the automobile to travel across greater distances (CDC, 2002). As urban sprawl is associated with intensive driving activities, and driving contributes to air quality deterioration, increase in morbidity and mortality occurs. Furthermore, sprawl also promotes less walking and a sedentary lifestyle thereby increasing risks of cardiovascular disease and stroke. In addition to its effects on health, the lack of physical activity is also blamed for obesity and other diseases associated with being overweight (Mokdad et al., 2001; Lee et al., 1999; Kampert et al., 1996).

The environmental consequences of land use changes also include eutrophication and acidification of lakes and other water bodies, land degradation and desertification, decrease in water quantity and deterioration of water quality, groundwater pollution, marine and coastal pollution, and biodiversity loss (Briassoulis, 1994; Brouwer et al. 1991; Blaikie and Brookfield, 1987; Jongman, 1995; Laws, 1983; Ortolano, 1984; Noble, 1999; Stephenson, 1994; Buttle and Xu, 1988; Higgins, 2007; Gaston et al., 2003). Eutrophication is the process by which a body of water becomes rich in dissolved nutrients from fertilizers or sewage, thereby encouraging the growth and decomposition of oxygen-depleting plant life and resulting in harm to other organisms. It is a natural process that nutrients and sediment are continually transported into river systems. However, urbanization and other land use changes accelerate this process by accelerating erosion and runoff. Likewise, acidification of lakes is caused by acid rain of which the precursors are sulfides in the atmosphere. With increased urbanization comes the need for more power generating facilities. Since power generation typically involves fossil fuel combustion, sulfur emissions to the atmosphere increases causing acidification of lakes.

Moreover, to supply the demand for urban land, forest lands are cleared and impervious surfaces are built. When this happens, rainfall is not effectively absorbed into the soil and the groundwater reserve. In its place, rainfall is diverted as storm water and flows directly into streams and rivers (Harder et al., 2007; Niehoff, Fritsch, and Bronstert, 2002; Galle, Ehrmann, and Peugeot, 1999; Stephenson, 1994). If left unmitigated, a reduction in groundwater discharge would have devastating effects since about a third of all communities in the United States depend on groundwater as a source of drinking water (USEPA 2002). Storm water from parking lots, highways, and other impervious surfaces carries contaminants such as oil, grease, and other chemicals into streams as non-point pollution. Hence, both water quantity and quality are affected by land use changes. Biodiversity loss is an outcome of forest clearing that destroys the habitat for many species of flora and fauna. Higher biodiversity or species richness requires a healthy balance of environmental conditions especially in climate and land-use patterns (Higgins, 2007).

The New Jersey Department of Environmental Protection, (NJDEP, 2003) identified 178 environmental and human health stressors in the state. The conversion of agricultural and forest land to developed land ranks first in the list. The next four threats on the list are indoor and outdoor air pollution, loss of biodiversity and introduction of invasive species, deterioration of water quality, and decline in drinking water supply. It is to be noted that all these concerns are related to land use changes.

Because of the profound impacts of land use changes on health, environment, economy, and society, substantial research has been conducted to predict how land uses and the interfaces between urban and rural (and/or between high intensive agriculture and

non-commercial farms) will shift and change over time so that these impacts can be anticipated and appropriately addressed from a policy perspective. Since such changes result from the interplay of complex socioeconomic and biophysical processes, they are virtually impossible to duplicate through experiments or analysis by empirical observation (Verburg and Veldkamp, 2005; Walker, 2003; Pittman, 2003; Baker, 1989). One approach to solving this dilemma is through land use change modeling (Chen et al., 2002; Briassoulis 2000; Baker, 1989). Appropriately calibrated computer models can provide the most systematic and accurate way to predict future land use changes. Aside from forecasting, models can be used to explore land use system response to policy interventions through “what-if” scenarios. Predictions of land use changes based on scenario analysis are frequently used in the context of policy making on local, and regional issues as well as at a larger scale – that of global climate change (Salmun and Molod, 2006; Reid et al., 2005; Ramankutty et al., 2005; Kalnay, 2003; Pielke et al., 1999; Riebsame et al., 1994).

This dissertation reviews the land use change modeling efforts that have been carried in the past and aims to develop a fine-resolution spatial land use change prediction model that uses various machine learning and interpretation algorithms. The model was developed, calibrated and tested for Hunterdon County located in central New Jersey.

### **1.1 Definitions: Land Use, Land Cover, Land Use/Cover Changes**

Before moving to specific discussion of land use change modeling, it is necessary to define several key terms employed in this dissertation such as *land use*, *land cover*, and

*land use change and land cover change*. These definitions vary according to the particular application, scale, and context that is being studied. This section defines these terms in the context of this research.

### **1.1.1 Land Cover and Land Use**

*Land cover* is often referred to as the biophysical state of the earth's surface and its subsurface in its natural (vegetation) or artificial constructions, and *land use* to how land cover is manipulated in a particular human activity (Turner et al., 1995; Skole, 1994; Burley, 1961). For example, land cover refers to the vegetation that covers a parcel of land such as grass, trees, or water. Meyer & Turner (1994) included quantity and type of vegetation, water and earth materials in their definition of land cover. Land use would refer to land classification according to commercial purposes such as stores, office buildings, and apartments; industrial purposes such as factories and assembly plants; or agricultural purposes such as cropland, pasture, rangeland, and forest. Cihlar and Jansen (2001) caution that these terms are not synonymous and do not have a one-to-one relationship, and thus should not be used interchangeably.

In small-scale land use change studies, these distinctions are of primary significance. As an illustration, two land parcels may have similar land cover, but have different land uses. An industrial plant engaged in assembling electronic parts may appear similar from the outside to an office building with a warehouse in an adjacent lot. In this case, the assembly plant would be designated as industrial use while the warehouse is designated as commercial use. Another case would be forest land cover that can also be used as urban land if it is used as hunting grounds for recreational

purposes. Conversely, two land parcels may have a single land use classification, but have multiple types of land cover. For example, a golf course and an office condominium are both commercial land uses. The former would have a land cover of grass while the latter would be considered developed.

However, some studies use the term land use and land cover interchangeably because the purposes for which lands are used have associated types of land cover (Anderson et al., 1976). Land use maps are usually deduced from information taken from satellite images. The interpreter uses patterns, tones, shapes, and textures – information about land cover to deduce or derive information about land use activities. Some studies do not distinguish land use from land cover particularly in large-scale models where subtle differences in land use/land cover are averaged out.

Nonetheless, information gleaned from land use or land cover alone does not always adequately meet the requirements of particular users. To cite an example, in order to evaluate the vulnerability of forest land to effects of invasive species it is necessary to supplement forest cover information with information on how much of the area is used for activities such as hunting for recreational purposes (land use information). To accommodate the requirements of the majority of users, NJDEP (as well as many other agencies) tends to integrate both land use and land cover when evaluating landscape changes. NJDEP eventually followed a land use/land cover classification scheme when it released its land use maps for public use.

The land use/land cover classification system used in New Jersey is a modified version of the 1976 Anderson Classification System. The modifications consist of a few deviations from Anderson's general level I classifications for some special New Jersey

classes identifying disturbed wetlands. The Anderson Classification System has several levels with level 1 as the most general having 8 classes. These are urban or built-up land, agricultural land, rangeland, forest land, water, wetlands, barren land, and managed wetlands. Of these eight classes, rangeland is not found in New Jersey. Consequently only seven categories are used in the NJDEP mapping system of 1986. Managed wetlands were dropped from the 1995 mapping system making a total of six land use types in the current system. The descriptions of each class discussed below are based on the NJDEP Modified Anderson System of 2002.

Urban land is the designation assigned to areas where the landscape has been altered by human activities. This class includes areas used as residential, commercial, industrial, transportation, industrial and commercial complexes, mixed urban, recreational, and other urban or built-up lands. The last category includes cemeteries, and undeveloped land within urban areas among others.

Agricultural land includes acreage primarily used for the production of food and fiber including the structures used in this production. Included in this class are cropland and pastureland, orchards, vineyards, nurseries and horticultural areas, and confined feeding operations.

The forest land class pertains to any land with woody vegetation excluding wetlands. It includes sub-classifications such as deciduous, coniferous, mixed deciduous-coniferous, and brush land.

Another classification is water. It includes streams and canals, natural lakes, artificial lakes and reservoirs, and bays and estuaries. Closely related to the water class are wetlands. According to Anderson et al. (1976), these are “areas that are inundated or

saturated by surface or ground waters at a frequency and duration sufficient to support vegetation adapted for life in saturated soil conditions.” The wetlands category includes naturally vegetated swamps, marshes, bogs, and savannas.

Lands in the barren land category are characterized by thin soil, sand, or rocks which generally have low vegetation. Included in this category are beaches, bare exposed rocks, sites used in extractive mining, and landfills.

The last category is Managed Wetlands. Included in this type of wetlands are landscaped or maintained areas that appear saturated with water, but that are unable to support vegetation typically found in wetlands. Examples are swales and managed lawns in business parks. Used until 1986, this classification was removed in 1995. Thereafter, managed wetlands were subsumed under Wetlands in the level 1 category.

### **1.1.2 Land Use Change and Land Cover Change**

The changes in land cover and land use encompass conversion and modification (Turner et al. 1995). With regard to land cover change, conversion involves change of land cover types while modification involves an alteration in physical property or function, but that does not necessarily involve a change from one type of land cover to another. A change in the dominant type of vegetation in a wetland area is an example of modification. By the same token, conversion and modification can also occur under circumstances of land use change. Conversion in land use is exemplified by changes from one land use to another while modification occurs in a much more subtle way. For example, modification in agricultural land use could be the intensification of farming. In urban lands, modification could be situations such as changes from low-income to high-income

residential areas, or changes of suburban forests from their natural state to recreation uses.

To define change in land use and in land cover, consideration should be made on the spatial and temporal scale at which the analysis is conducted. This means that a change in areal extent of land use can be detected only when there is a sufficient spatial level of detail (i.e., the map resolution should be detailed enough to see subtle changes in topography). Along with spatial detail, the duration for which the observation is made should be long enough to be able to detect protracted changes in the landscape. These two considerations influence biophysical, social and other processes that shape land use change. At fine scales, social variables tend to be significant while biophysical variables have more significance on larger scales (McConnell, 2002). Knowledge of what scale these processes and variables are discernible can enhance the accuracy of analysis.

This research aims to model the quantitative changes in areal extent of a given type of land use or land cover and their spatial locations. It is important to note that this analysis is concerned with more than just numbers as changes can occur in various subtle forms. Each of these changes needs to be addressed using appropriate land management measures. Given a clear understanding of such basic concepts, decision makers should be able to discern the appropriate measures to mitigate any profound impacts related to the changes in land use and land cover.

### **1.1.3 Land Use Change Modeling**

A model is “an idealized and structured representation of the real” (Johnston et al., 1994:273) or “an experimental design based on a theory” (Harris 1966:258). Land use

change models developed generally fall into two categories: those that predict total land use change for a region or state and those that predict land use for specific parcels or grid cells (CARA, 2006). The former group predicts how much new development will occur in a geographic area but does not forecast where exactly these changes will occur. These models are commonly used to estimate how much development will occur based on projections of population and population growth rates (Alig and Platinga, 2004; Stephenne and Lambin, 2004), or to analyze habitat loss and other environmental impacts arising from increased demand for urban land (Du et al, 2006). Some groups (CARA, 2006) subdivide models that predict land use change for specific locations into three broad categories: build-out analyses, cellular automata models and, agent-based models. A build-out analysis projects maximum allowable future development in a specific area based on zoning regulations (Conway and Lathrop, 2005a, 2005b; Underwood and Davis, 2005; Stony Brook-Millstone Watershed Association, 2003). A build-out analysis is an important tool for government planners as it provides a way to evaluate potential outcomes of existing zoning regulations. It can also be used to project the resulting land use patterns of any proposed changes in the regulations. It should be noted that build-out analysis only projects what could happen under a certain regulatory regime. It would not predict when or whether the complete build out would occur (Montgomery Planning Commission, 1996). Such predictions could be made by the other two types of models – cellular automata and agent-based models. A cellular automata is used to simulate the spatial pattern of land use change using a set of transition rules. The region or area being modeled is subdivided into cells. The model user defines a set of rules that describes the probability that a cell will transition to another land use. The set of rules called transition

rules are based on cell characteristics such as topography, proximity to existing transportation networks and urban centers, and characteristics of its neighboring cells (Torrens, 2000). The set of rules is then used to provide quantitative and spatial predictions of future land use changes.

Agent-based models complement cellular automata models. Unlike cellular automata models where the focus is on the interaction among neighboring cells, agent-based models focus on interactions among decision makers such as land owners, institutions, and other organizations (Parker, et al., 2003). Agents refer to all stakeholders in the land management process. A landowner's decision to sell his property, for example, can potentially precipitate decisions by other similarly situated landowners or groups of landowners to sell their properties as well. Agent-based models simulate interactions among such groups of decision makers that contribute to changes in land use in a given area. Given the ability of agent-based models to represent human decision processes, more and more applications of this technique have appeared in the literature. However, precisely because of these features, agent based models have their own important limitations. Owing to their complexity, techniques for model calibration, verification, and validation are not yet well developed thereby making communication and appreciation of results difficult (Parker et al., 2003). A better understanding of the output is needed before these procedures are in place. Agent based models also require voluminous data and intensive computing (Koomen, Rietveld, and de Nijs, 2008). For these reasons, the cellular model is deemed to be a more suitable modeling technique for this research.

## 1.2 Research Priorities in Land Use Change Modeling

After reviewing the theory, rationale, and implementation of various aspects of the past land use models including (1) level of analysis; (2) cross-scale dynamics; (3) driving forces; (4) spatial interaction and neighborhood effects; (5) temporal dynamics; and (6) level of integration Verburg, Schot, et al. (2004) identified four research priorities for future land use models..

First, a future modeling approach should better address the multi-scale characteristics of land use systems including scale dependency, scale-up effects, and interactions of processes operating at different scales. Briassoulis (2000) voiced a similar appeal and argued that the impacts of land use change should not only be confined to one scale or one level, but should also be dispersed at different levels and scales.

Second, the future modeling approach should be an integration of multiple disciplines in methodology, evaluation of the impacts caused by land use changes and analysis of urban-rural interaction. The Land Use and Land Cover Change (LUCC) project of the International Geosphere-Biosphere Program (IGBP) calls for the integration of different disciplines involved in land use studies in an effort to develop models that approximate reality more accurately (Lambin et al, 2000). Geoghegan, et al. (1998) called for land use change models “socializing the pixel and pixelizing the social” to integrate the social sciences with geographic and ecological models.

Third, the future modeling approach should pay explicit attention to temporal dynamics of land use change such as the interaction between spatial and temporal dimensions, influence of non-linear pathways of change, feedbacks, and time-lags. For instance, incremental impacts of land use change at lower levels of the spatial and

temporal scales lead to cumulative impacts after some period of time at the regional scale. Examples of these impacts include desertification and salinization: Niang, et al. (2008); Orlovsky, et al. (2006); Versace, et al. (2008).

Finally, the researchers should develop new methods to assess and quantify neighborhood effects in land use change models. To a large extent, it can be said that land use conversions result from the occurrence of land uses in the neighborhood (Verburg, de Nijs, et al., 2004; Torrens, 2000). For example, the methods to quantify neighborhood effects in cellular automata models are frequently based on expert knowledge. Although this method can be satisfactory in many situations, it is however prone to error resulting from subjectivity or bias from the experts. A more sophisticated way to quantify neighborhood effects is needed.

### **1.3 Research Objective**

The objective of this dissertation is to develop a fine-resolution spatial land use change prediction model that uses machine learning and interpretation algorithms built upon the previous land use change modeling efforts. Specifically, this research will expand previous Cellular Automata (CA)-based land use change modeling using cadastral parcels as a unit of analysis. The model will be developed and validated using historical land use data in Hunterdon County, New Jersey. The validated model will be used to predict future land use changes under different land use policy scenarios. Although the model is developed using Hunterdon County as its study area, the method employed in this research can be adapted for land use planning in other counties of the state and at regional level where the parcel and land use maps are readily available.

CA models are considered to be an improvement over large urban land use models (Torrens, 2000) and they have been widely used because of several outstanding features. CA models are especially equipped to handle complexities in spatial data that characterize urban land use systems (White, Engelen and Uljee, 1997). Simplicity is an inherent feature of CA systems as shown in their straightforward transition rules. This model type handles dynamism of urban systems with simple rules that facilitate computational efficiency. With regard to the complexity of urban systems, Torrens (2000) contends that complexity is allowed to emerge from within a CA model. CA models can mimic the self-organizing nature of complex systems since the interactions between various actors, residents, planners, and developers at the local level are simulated by the interaction of different cells in the CA lattice (O'Sullivan and Torrens, 2001). Lastly, CA models can integrate various scales in the modeling process as shown by the macro and micro sub models in the Environment Explorer Model (Engelen, White, and de Nijs, 2003).

Although square-shaped grids or cells are traditionally used as the unit of analysis in most CA models because of their convenience, cadastral parcels are the most ideal spatial unit of analysis in land use change modeling (Landis and Zhang, 1998a, 1998b; Allen and Lu, 2003; Irwin and Geoghegan, 2001). Stakeholders' behaviors such as purchasing, selling, and developing land are made and observed at the parcel level. It is also at the parcel level that most land use policies such as zoning are crafted and implemented. Parcels contain socio-economic information through the municipal tax assessment database – information which can be used in modeling studies. Additionally, land use change predictions at the parcel level can be easily aggregated at the regional

scale to economic, ecological, and environmental impacts by linking them to certain regional models. However, disadvantages are not entirely lacking. Parcel data can pose difficulties in modeling work. Aside from their highly variable area, parcel lots may include multiple land covers and development constraints. The methodology herein tries to address these difficulties.

#### **1.4 Research Contributions**

CA land use change models have been used to model urban growth both at both regional and smaller municipal scales (Clarke and Gaydos, 1998, Engelen, White, Uljee and Drazan, 1995; Landis, 1994; Cheng and Masser, 2004; Li and Yeh, 2000; White et al., 1997). All of these models have used regular square-shaped raster grids as their unit of analysis except Stevens et al. (2007) who used irregular shaped parcel as a unit of analysis to evaluate urban development in Saskatoon, Saskatchewan, Canada. There are tremendous challenges when using individual parcels as the unit of analysis. Stevens et al. (2007) addressed many research issues, but some challenges remain. Chapter 2 presents details of those challenges and discusses how they were treated in this research. There are other limitations in the model of Stevens et al. (2007). First, the others focused on parcels representing only two classification types in urban development: urban and non-urban uses. Such a procedure is overly restrictive and limits the applicability of their model. Second, they assigned weights to various influencing factors to derive land use conversion rules which are dependent on expert knowledge and as such prone to subjectivity bias.

This research extends the parcel-based CA model developed by Stevens et al. (2007) to predict land use changes in three areas. First, it uses the more realistic and finer land use types instead of a simple, dichotomous, urban/non-urban classification scheme. Second, it addresses a greater range of complications in defining driving factors and neighborhood effects when using parcels as the unit of analysis. Third, it applies an innovative data-mining scheme to elicit land use transition rules in CA. Specifically, as discussed in Chapter 2, a decision tree approach was used in this research. Since this method uses a machine learning and interpretation algorithm, it avoids the subjectivity bias that usually becomes manifest when expert knowledge is used. Li and Yeh (2004) applied the decision tree in their CA land use change model, but used artificial areal units (e.g., pixel) instead of parcels as the unit of analysis and therefore did not address the modeling issues that arise when using irregular parcels as unit of analysis. The machine learning and interpretation approach for deriving transition rules does not require extensive quantitative skills and would be better appreciated by non-technical users such as stakeholders and land use change decision makers. The decision tree approach provides a better and easier way to take into account all relevant driving factors of land use change to develop a more accurate transition rule for use in cellular automata. Together with the use of the irregular parcel as unit of analysis, this research can contribute to the development of CA-based land use change models.

This research also addresses several research priorities in land use change modeling spelled out by Verburg et al. (2004b). The straightforward manner in which transition rules are developed can easily be followed by users from other disciplines other than the natural sciences, thereby providing a vehicle to which they can comfortably

manipulate and couple with their own models. Through the application of a data-mining technique (i.e., decision tree in an irregular CA model), this research should be able to accurately quantify the effects of neighborhood and other driving factors of land use change.

### **1.5 Assumptions**

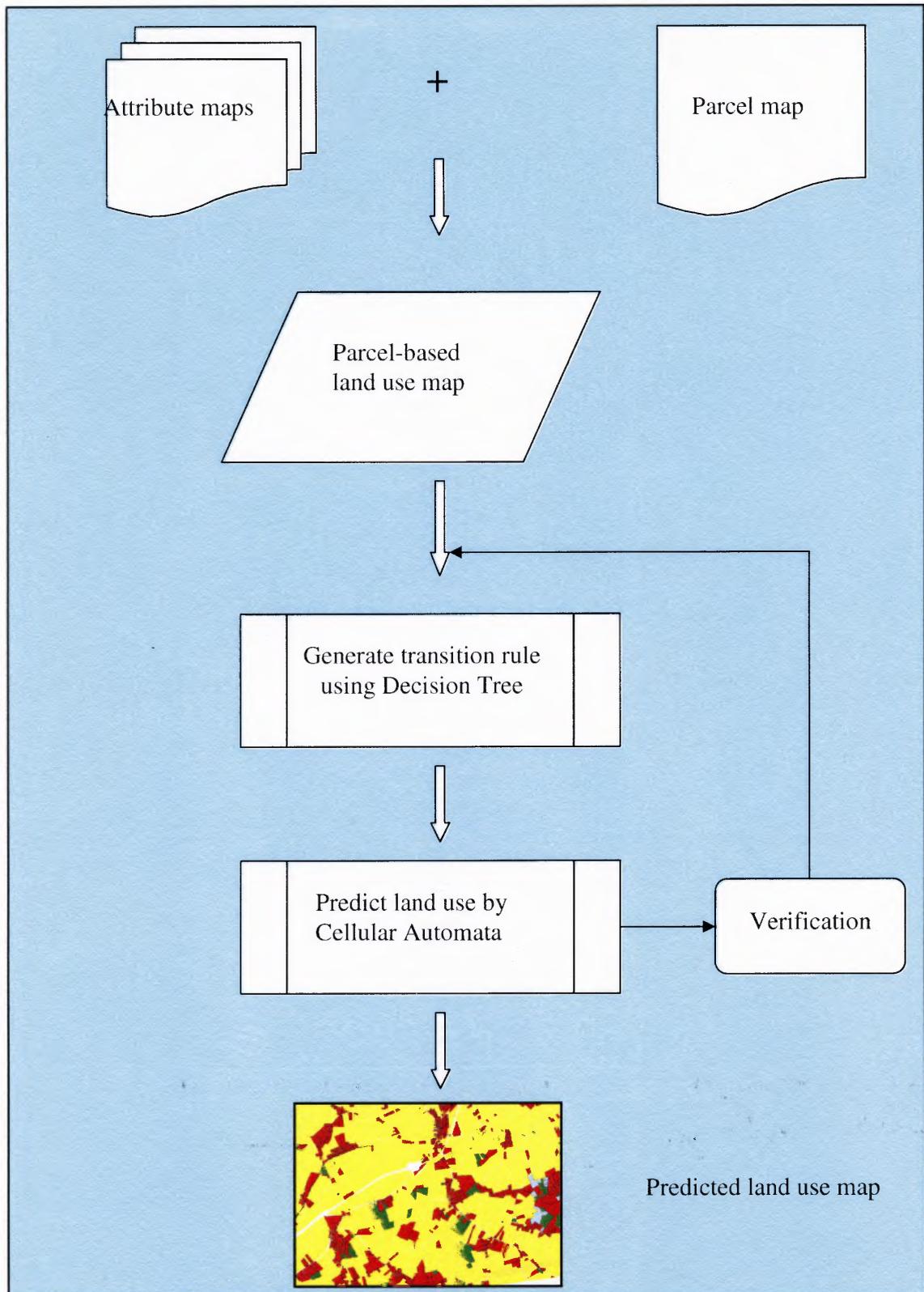
In developing the model framework, the following assumptions are made:

- i. Transition “rules” are stable over time. It is assumed that the factors driving land use and land cover change during the period remain constant and their behavior can thus be used to extrapolate into “what-if” scenarios.
- ii. Land cover (i.e., land classification) is an effective surrogate measure for land use. Since this research is primarily focused on modeling and quantifying land use change, changes in land cover are assumed as quantitatively equivalent to changes in land use. A corollary to this is that land use change are likely to cause land cover change. As Turner and Meyer (1994:5) stated, “A single land use may correspond fairly well to a single land cover: pastoralism to unimproved grassland, for example.”
- iii. Land use change occurs reasonably independent of spatial and economic processes outside the study area. The driving factors included in the modeling process are the relevant driving factors that significantly contribute to land use change in Hunterdon County. The contributions of supra-local factors are assumed to be relatively minimal.

iv. It is recognized that parcels can be subdivided and their boundaries change over time. Although the consideration of this feature contributes to a more sophisticated model, it greatly complicates the modeling process. For this reason this research assumes that the boundaries of the parcels remain the same over all the modeling periods.

### **1.1 1.6 Modeling Procedure**

Figure 1.2 presents the flowchart of the procedure to develop the land use change model. In general, all thematic maps regarding the driving factors of land use change must be converted to parcel-based maps. Using these factors, a decision tree was used to derive the transition rules which were subsequently used in the CA model to predict land use change. The predicted results were mapped in Geographical Information System (GIS). A brief discussion of each step is presented below and Chapter 3 gives more detailed consideration to some of the procedures.



**Figure 1.2** A flowchart of the modeling procedure.

### **1.6.1 Preparing Parcel-based Maps**

The database requirements for this research consist of parcel and land use maps containing spatial data on driving factors of land use change. Parcel maps were obtained from the Office of Geographical Information of Hunterdon County, New Jersey.

Land use maps for 1986, 1995 and 2002 were obtained from the NJDEP Bureau of GIS. The maps for 1986 and 1995 show the land use on a county basis while the 2002 map uses watershed management area thereby making it necessary to geo process the 2002 land use map to the same boundaries for the study area.

Spatial data on driving factors were obtained from digital elevation models (DEM), land use/cover (1986, 1995, and 2002), soil maps, maps of protected areas, census data, and Tiger/Line files for location of urban centers and roads, streets and major highways. All these data sets were available from the NJDEP Bureau of GIS in digital-data form. These maps were then overlaid with a parcel map of the county to generate parcel-based data.

The process of overlaying land use with parcel maps produced parcels with multiple land use types since the boundaries of the LULC patches did not coincide with the parcel boundaries. In order to assign a single land use type to each parcel, classification criteria were developed and this procedure is explained in Chapter 3. Overlaying the parcel with the soil map also resulted in multiple soil types for each parcel. Assigning a single soil type for each parcel was based on the predominant soil type in the parcel. Such preliminary processing was needed before the data were used in model calibration. After all driving factors were derived and expressed on a parcel basis, it was possible to begin building the decision tree.

### **1.6.2 Generating Transition Rules Using Decision Tree**

Simply stated, a decision tree represents a collection of rules that are derived by extracting patterns in a dataset of driving factors of land use change leading to a target value – the land use type in parcels.

Land use type and spatial data on the driving factors taken from the previously prepared parcel-based maps for 1986 and 1995 were used as training data for generating the decision tree. This process used the J48 WEKA Version of the C4.5 algorithm to generate the decision tree. The software can be downloaded from <http://www.cs.waikato.ac.nz/ml/weka>. Once constructed, the decision tree is converted into a java file, compiled, and then used as transition rules in cellular automata. The procedure for using the program is explained in more detail in Chapter 3.

### **1.6.3 Predicting Land Use Change Through Cellular Automata**

Agent Analyst, a software program that serves as an interface between GIS and a cellular automata program was used for this step. The modeling and forecasting process in cellular automata was simplified by allowing predictor variables or driving factors in digital format as input and likewise projecting the results in the same form. Using the 1995 dataset of input variables and the decision tree generated in the previous step as transition rule, cellular automata predicts land use change for 2004 and subsequently for 2013. From this, a predicted land use map is then displayed in GIS. In a subsequent step, a 2002 dataset is used for validation of the model.

#### **1.6.4 Model Validation**

A validation procedure evaluated whether the model was able to achieve its purpose. The procedure entailed an assessment of the accuracy of the model in simulating land use change. The literature is replete with ways to undertake accuracy assessments as there is still no standard method that is universally accepted. In keeping with the objective of developing a model that lends itself to easy interpretation, simple yet adequate accuracy metrics were used. To this end, the following metrics were employed:

- a) Three variants of Kappa (standard Kappa, Kappa location, and Kappa quantity) to indicate location and quantity predicting capability;
- b) Errors of omission and commission sometimes referred to as User's and Producer's accuracy respectively; and
- c) Overall accuracy computed as the ratio of correct classification to the total number of classifications.

#### **1.7 Expected Output**

This research aims to produce a more sophisticated modeling system that addresses some practical difficulties in extant land use change models while at the same time demonstrating practical applicability. Hunterdon County was used as a case study to assess how accurately the model would have predicted recent changes using historical land use data. The model then predicted future land use change in the County. In doing so, the advantages of the enhanced model are illustrated. The output of this work includes the following:

### **1.7.1 Parcel-based Maps**

Thematic maps of the various driving factors used in the model were overlaid with a parcel map of Hunterdon County. After preliminary processing of the resulting overlay map, a parcel-based map of the various driving factors was produced. This map provides land use change including past and predicted land use, soil type, slope, urban centers, transportation network, and stream locations all of which are referenced to the parcel. The data include both spatial and tabular information.

### **1.7.2 Workable Model for Land Use Change in Hunterdon County, New Jersey**

Another output of this study is a streamlined integration of several processes employed during the implementation and calibration of the model. The components of the working model include: a) land use type identification algorithm: Overlaying a parcel map with the existing land use map of the county produced multiple land use types for every parcel. This algorithm assigns a single land use type for each parcel using the criteria discussed in Chapter 3; b) decision tree: Using driving factors of land use change as inputs, this program generates the transition rule used in cellular automata; c) cellular automata: This program uses a transition rule generated by the decision tree program and driving factors of land use change corresponding to the current period to forecast land use change.

### **1.7.3 Protocol for Map Comparison**

To enhance its usefulness, the model's predictive capacity is assessed for accuracy. In essence, the procedure entails a parcel-by-parcel comparison between a map of the predicted results and a reference map. Accuracy is measured in terms of the model's ability to predict the correct land use type, as well as to assess the correct location for the

predicted parcel. Documentation of this procedure can be used as a protocol for map-comparison exercises.

#### **1.7.4 Land Use Impacts of Various Land Use Policy Scenarios**

The model as applied to the study area has capabilities in predicting changes in land use, including the facility for evaluating the effects of alternative policies. Efforts by state government to preserve the rural and agricultural character of New Jersey were considered in this research. These initiatives have included open space and farmland preservation, as well as down-zoning policies. Likewise, the model lends itself to including other policy controls in the simulation. For example, provisions for the protection of wetlands using buffers can be included as an additional transition rule. Lastly, the model can also be applied to other similar areas provided that all the necessary inputs used in this study are available or can be derived.

### **1.8 Dissertation Structure**

The remaining parts of the dissertation are organized as follows. Chapter Two summarizes past and current research on land use change modeling with specific emphasis on cellular automata-based models. Chapter Three presents in detail the formal model and methodology in detail used in the dissertation. Chapter Four is devoted to the case study and a detailed description of the study area. This chapter discusses the various policies on land use, a historical background of land development, and the present state of land use and land cover distributions in the area. Chapter Five presents the modeling

results with regard to the case study. Lastly, Chapter Six summarizes the findings of the research, its contributions, and outlines future research directions.

## **CHAPTER 2**

### **LITERATURE REVIEW**

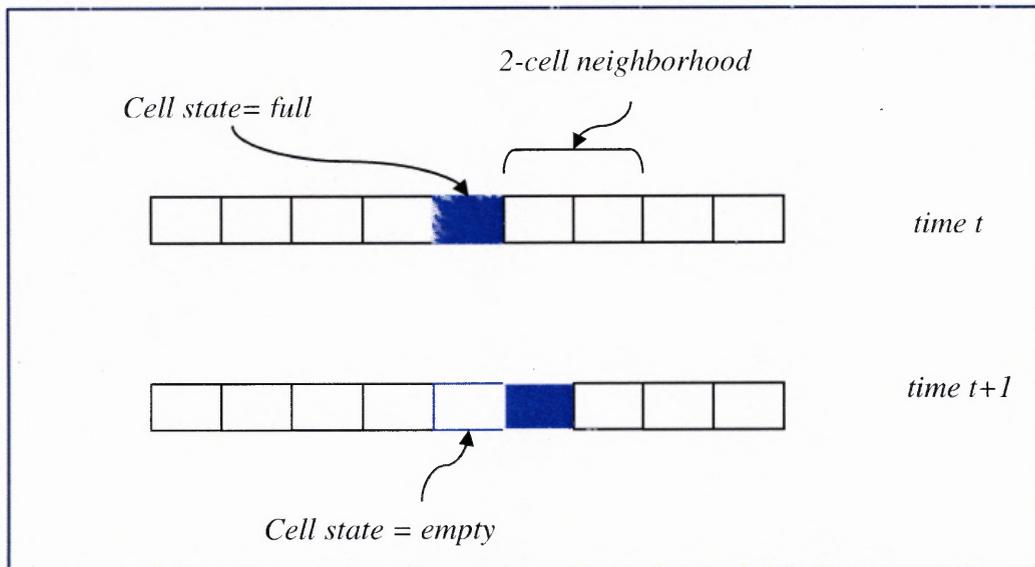
This chapter traces the development of cellular automata (CA) as a tool for modeling land use change. It describes how the formal CA of Ulam and Wolfram in the 1930s was modified for urban land use change modeling. The chapter then discusses a particular element of a CA model, namely the elicitation of its transition rules, and finally the remaining issues in CA modeling that this research intends to address.

#### **2.1 Elements of Cellular Automata (CA)**

The cellular automata model traces its origins to the work of the British mathematician Alan Turing in the 1930s. He conceived a Universal Turing Machine as a hypothetical automaton, which if given a suitable initial program and capacity for self evolution, would be able to perform any computation assigned to it. The machine reads information (commonly binary) received and decides a definite action to take based on a given set of rules. Levy (1992) defines an automaton as “a machine that processes information, proceeding logically, inexorably performing its next action after applying data received from outside itself in light of instructions programmed within itself (p. 15).”

CA has five principal elements --- namely cell state, lattice, neighborhood, time and transition rule. Lattice refers to the space or cell in which the CA exists and evolves over time. The lattice of a formal CA can be one-dimensional as in Figure 2.1 or two dimensional with the latter being more commonly used in land use studies. Although the

cells are squares in Figure 2.1, they may theoretically be in any geometric shape (O'Sullivan, 2001; White and Engelen, 2000).



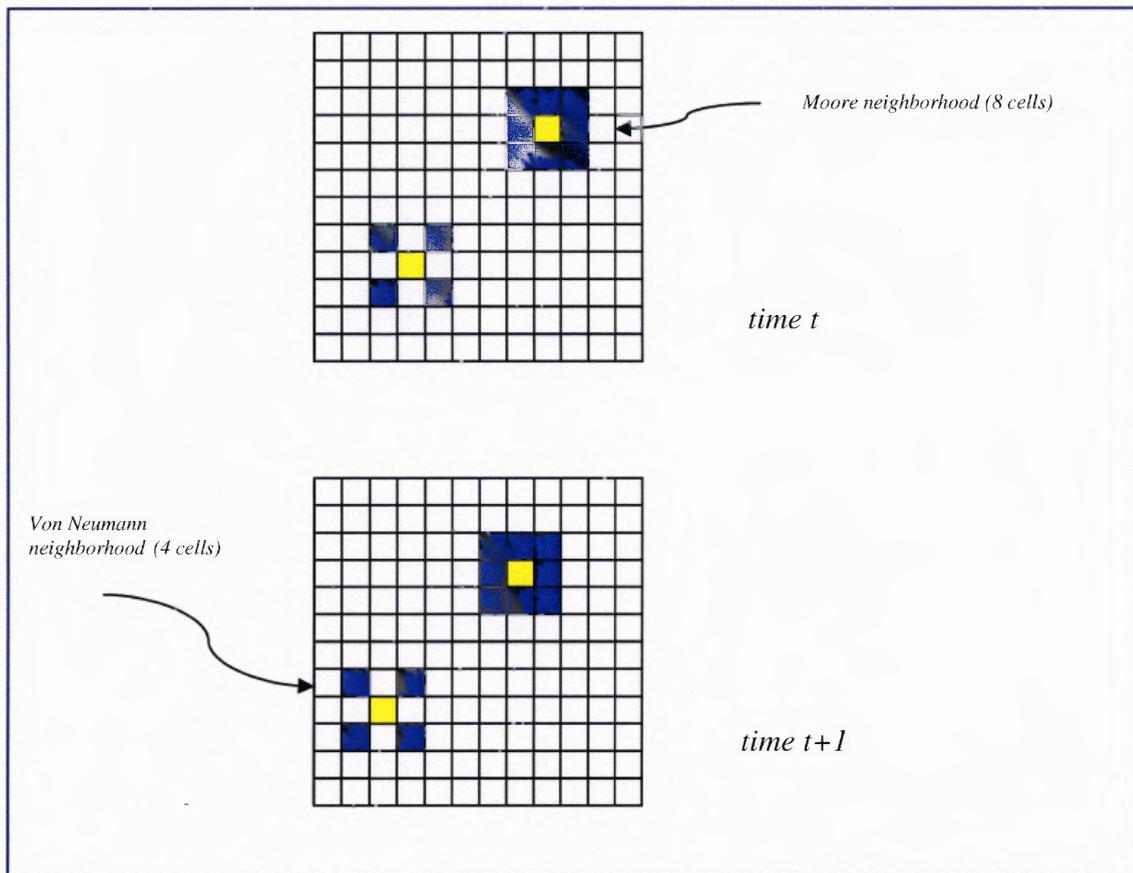
**Figure 2.1** One-dimensional cellular automata.  
Adapted from: Torrens (2000)

The neighborhood comprises the localized region of a CA lattice. The neighborhood templates for a two-dimensional CA are the Moore neighborhood and the von Neumann neighborhood as shown in Figure 2.2. In a two dimensional lattice, a Moore neighborhood is depicted as square shaped composed of the cells adjacent to a central cell. A 3 x 3 Moore neighborhood would thus be composed of eight cells while a 5 x 5 neighborhood size would include 24 cells. In contrast, a von Neumann neighborhood includes only the cells that are orthogonally situated around a central cell. Time is the duration of iterative steps for which a cell evolves. During each step, all cells are updated in a synchronous fashion.

Transition rules specify how cells change. They generally consist of IF, THEN, and ELSE statements that are evaluated using input from neighborhood of the cell. Boerboom and Jiao (2003) defined a transition rule as:

$$(2.1) \quad TP_{T+1} = f(S_T, NF),$$

where  $TP_{T+1}$  is the transition potential of cell at time  $T + 1$ ,  $S_T$  is state of cell at time  $T$ , and  $NF$  is a set of neighborhood factors. Note that the transition rule implies that change takes place purely as a function of what happens in a cell's neighborhood. The rule is applied uniformly on all cells thereby assuming a homogenous lattice. In reality however, these two assumptions (single influencing factor and lattice homogeneity) do not hold since each cell is associated with a set of different driving factors other than those arising from its immediate neighborhood (Ward, Murray and Phinn, 2000; de Almeida et al., 2003).



**Figure 2.2** Two-dimensional cellular automata.  
Adapted from: Torrens (2000)

## 2.2 CA Applications in Land Use Change Modeling

The introduction of CA systems in geographical applications was catalyzed by Tobler (1979) who first recognized the advantages of CA in solving geographical problems. In Tobler's cellular space model (1979), he postulated that the state of the cell is determined by the states of the cell's neighbors according to a uniform location-independent rule (White and Engelen, 1993). Following Tobler's pioneering work, other researchers applied CA to urban land use change modeling (White and Engelen, 1993; Batty and Xie, 1994; Clarke and Gaydos, 1998; Wu and Webster, 1998). CA models were considered to be an improvement over other large urban land use models (Torrens, 2000) and became

widely used because of several outstanding features such as: spatiality, decentralization, affinity with GIS, and simplicity and ability to integrate micro and macro approaches (Torrens, 2000; Parker, 2003; Batty, Xie and Sun, 1999).

### **2.2.1 Spatiality**

Many traditional urban land use models because of their large area of application (e.g. regions) tend to abstract over spatial detail (Torres, 2000). However, CA-based models are especially equipped to handle complexities in spatial data that characterize urban land use systems (White, Engelen and Uljee, 1997). These models make use of relationships among neighboring cells to determine a given cell's reaction or behavior. This process requires treatment of space in absolute terms as in GIS while other models treat space in a comparative manner (Torrens, 2000; Couclelis, 1997). Examples of the latter include the spatial interaction and gravity models (Batten and Boyce, 1986; Haynes and Fotheringham, 1984; Batty, 1976), the econometric models (Kitamura et al., 1997; Morita et al., 1997) and the location allocation models (Yeh and Chow, 1996; Duerden, 1996).

### **2.2.2 Decentralization**

The gradual evolution of the mono-centric city into multi-urban centers indicated a moving away from the central control theory in modeling (Resnick, 1994). As this development became widespread, CA models became more popular since that multiple growth centers can be modeled as individual spatial entities whose growth is governed by local transition rules in CA. A bottom-up approach entailed development at the local level that then shaped developments at the regional level.

### **2.2.3 Affinity with Techniques of Spatial Analysis Notably GIS**

CA is commonly regarded as having a “natural affinity” with raster GIS data (Takeyama and Couclelis, 1997). For example, the themes in GIS are similar to the state-spaces in CA. Moreover, CA organizes space in grids while GIS uses polygons or grids. The similarity in features between CA and GIS enable modelers to couple these two systems to facilitate input/output data processing and presentation.

### **2.2.4 Simplicity and Ability to Integrate Micro and Macro Approaches**

Simplicity is an inherent feature of CA systems, as shown in its straightforward transition rules. However, as to the complexity of urban systems, Torrens (2000) contends that complexity is allowed to emerge from within a CA model. CA can model local processes that interact at different scales and then build up into larger regional or global scales. Analogously, complex entities such as cities and other large urban systems exhibit fractal dimensions which are also mimicked in CA modeling systems (Torrens, 2000).

### **2.2.5 Examples of CA Land Use Change Models**

RIKS, an acronym for **R**esearch **I**nstitute for **K**nowledge **S**ystems of the University of Maastricht, is one of the earliest CA models. The model was developed as a decision support tool for the management of river basins, coastal zones, cities and regions (Engelen et al., 1995). RIKS is a package consisting of several modules; one of which is for land use change. Designed to be used in different geographical areas, it has both micro and macro components. The macro component forecasts land use demand from socio-economic indicators and existing policies relating to land use. Macro component

estimates are subsequently used by a cellular automata model in the micro component to allocate land among different land use types.

Another noticeable example is SLEUTH which is an acronym for an urban land use change prediction model taken from its input requirements such as: **S**lope, **L**and use map, **E**xcluded areas from urbanization, **U**rban areas, **T**ransportation map, and **H**ill shading area (Clarke et al., 1997). The behavior of each individual cell is dictated by rules described by the state of its neighborhood cells. These rules govern various urban growth trajectories such as diffusion, breeding, and spreading of small urban areas. It also takes into consideration slope resistance to urban settlement, and road gravity which refers to growth resulting from proximity to road systems.

As discussed later, many CA-based land use change models have been developed by introducing new methods to elicit the transition rules in the models. Wu (1996) is among the first who introduced the use of fuzzy logic in CA models to simulate different scenarios resulting from the implementation of different urban development policies. Li and Yeh (2002) and Pijanowski et al. (2002) used artificial neural networks to elicit or “learn” patterns from driving factors of land use change and calculate a conversion probability for a given cell based on those factors.

It should be also realized that the land use change process is stochastic in nature. The CA models discussed above are deterministic. Another direction in CA-based land use modeling is to incorporate the stochastic land use change process. The way to CA-based models being stochastic is to develop the stochastic transition rules. Examples of these models are Guan, et al. (2005); de Kok, et al. (2001); and Ward, et al. (2000).

However, this dissertation only concerns the deterministic CA-based land use model for simplicity.

### **2.3 Cell Shape and Neighborhood Configuration**

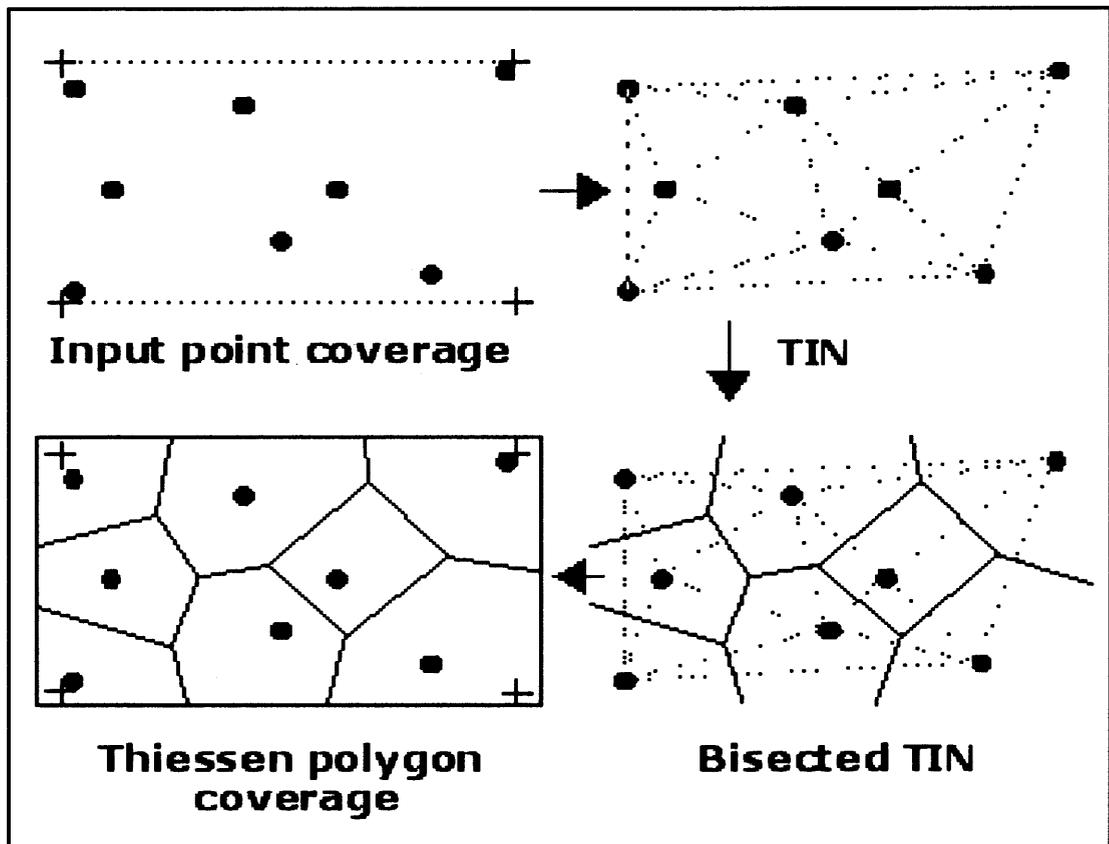
While CA models are seen as having an extensive application in urban land use change modeling, some peculiarities of this research present certain challenges. The first one is on using a cell size and shape that is different from the conventional raster grid in common land use change CA models. Related to this issue is the use of a non-regular neighborhood configuration, a departure from the usual von Neumann and Moore configuration used in standard models. In the same manner as the cell, the size and shape of the neighborhood result to different land use change path leading to different conclusions (Liu, 2007; Kocabas and Dragicevic, 2006). The following sections discuss the issues as reported in the literature.

#### **2.3.1 Using Irregular Cell Shapes**

One of the simplifying assumptions of the cellular automata method as applied to urban land use change modeling is that classical CA models represent space using a raster grid of cells. Such a scheme is not an accurate representation of irregular sizes and shapes that characterizes urban land where it is often given in parcels. To apply a regular grid system to approximate an urban parcel unit results with inaccuracies since raster boundaries usually do not coincide with parcel boundaries. Some CA models that used regular grids were developed by Wu (2003), Yeh and Li (2001), White and Engelen (2000), and Batty and Xie (1999). Couclelis (1985) suggested the use of irregular grids

in CA modeling as early as 1985. Nonetheless, there have been few instances where this approach was used. The lukewarm reception may be due to the apparent difficulty of encoding a complicated neighborhood due to its irregular shape and size (Batty, 2007). Another reason is the ease of use afforded by the compatibility of regular grids with raster data in GIS. Notwithstanding these difficulties, several studies (Shi and Pang, 2000; Pang and Shi, 2002; Flache and Hegselmann, 2001; O'Sullivan 2001; 2002) tried to use irregular grid approaches. They use the Voronoi polygons and Delaunay triangulation. Voronoi polygons partition space into cells where each cell consists of points closer to one particular object than to any other (Okabe, 2000). If a line is drawn between any two points whose Voronoi domains touch, a set of triangles is obtained called Delaunay triangulation. The procedure is described in the following paragraph using Figure 2.3 as reference.

To construct a Voronoi diagram using vector data in GIS, the centroid of each polygon is first created. The centroids, represented as points, are used to form a triangulated irregular network (TIN) that meets the Delaunay criterion. The perpendicular bisectors for each triangle edge are then generated to form the edge of a Voronoi polygon. The locations at which the bisectors intersect determine the locations of the Voronoi polygon vertices.



**Figure 2.3** Construction of a Voronoi (Thiessen) Diagram.

Source: ESRI, ARCGIS 9.2 Desktop Manual, 2006

Shi and Pang used Voronoi polygons to model the effect of the presence of urban centers, transport facilities, and other amenities on population movements. In another study, Flache and Hegselmann (2001) compared model results using a regular grid structure with irregular grids. They discovered that using irregular grids gives better insights on model results than a regular grid structure does. Stevens, Dragicevic, and Rothley (2007) pointed out, however, that although Voronoi polygons are irregular in shape and size, they do not represent useful units such as cadastral land parcels. Approximating parcel areas with these units leads to computational errors since their boundaries do not necessarily coincide. On the other hand, O'Sullivan's model of urban gentrification (urban housing renewal induced by displacement of lower income groups by more

affluent groups) generated a graph structure to model the neighborhood of irregularly shaped parcels using a Delaunay triangulation approach. However, this approach is limited when used in simulating urban growth (Kocabas and Dragicevic, 2006; Stevens et al., 2007).

### **2.3.2 Defining an Appropriate Neighborhood**

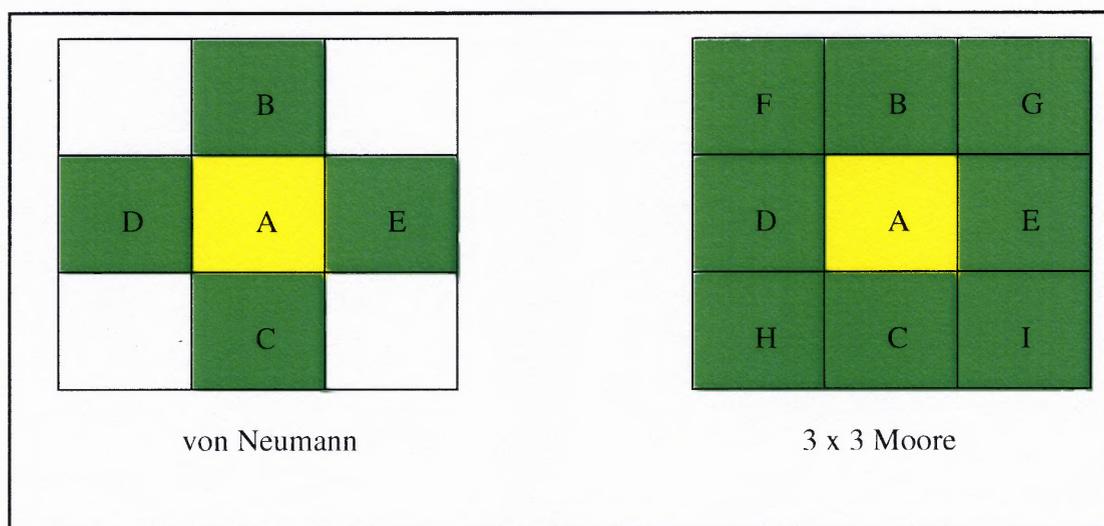
The neighborhood system of a cell in formal CA systems includes the cell itself and a group of eight cells forming a square around the cell (Moore neighborhood), or four adjacent cells (von Neumann neighborhood) as shown in Figure 2.2. While the formal CA model restricts interaction within the area defined by this neighborhood, land use change processes transcend the immediate neighborhood of a cell in urban systems. Social interactions, for example, can involve people in one's neighborhood and those beyond municipal borders as well (Torrens and O'Sullivan, 2000). White and Engelen (1997) extended the standard neighborhood to a region of eight cell radius around a central cell for a total of 196 cells. Verburg (2004) experimented with neighborhood sizes to quantify neighborhood effects.

Lou et al. (2000) asserted that CA theory regards the neighbors of a cell as an aggregation of all cells that are capable of affecting the central cell's next state. Accordingly, the authors expanded the coverage of neighborhood which considers not only spatial adjacency (Moore or von Neumann configuration) but also attribute correlation. Attribute correlation takes account of those non-adjacent cells that affect the state of the cell. Lou et al. (2000) grouped the spatial relationships between two cells into three categories, namely spatial adjacency, spatial neighborhood and complicated

separation. Spatial adjacency occurs when the distance between two cells is zero or less than a pre-specified numerical value. Spatial neighborhood occurs when the distance between two cells is larger than zero or a pre-specified numerical value. Complicated separation includes all the other relations between two cells exclusive of spatial adjacency and neighborhood.

Limiting the neighborhood space to von Neumann or Moore configurations or even their extensions (White and Engelen, 1993) is incompatible with the nature of human relationships that influence the land use decision-making processes (Flache and Hegselmann, 2001). Flache and Hegselmann (2001) argued that regular grids impose restrictions on transitivity of neighborhood relationships and preclude variation in the number of neighbor cells in given locations. Transitivity in relationships refers to a situation wherein the first element is in relation to the third element if the first element is related to a second element which in turn is related to a third element. However, with a von Neumann neighborhood, it is not possible to have transitive relationships between neighbors. To illustrate, referring to the von Neumann configuration in Figure 2.4, B is a neighbor of A and A is a neighbor to C, but B is not a neighbor of C. To put in another way: The neighbors of a center cell (here: A) are not neighbors among themselves (Hegselmann, 2008). The situation is different in a Moore neighborhood. B is a neighbor of A and A is a neighbor of C. As shown in the figure, B can also be a neighbor of C. However, while transitivity is made possible in a Moore neighborhood, it precludes variation in the number of transitive relationships a cell can have. In other words, owing to uniformity in size and shape, all Moore neighborhoods in a lattice will have the same number of transitive relationships. In real situations, the number of cells in a given

neighborhood affects or influences decisions made by individuals. In the Schelling's neighborhood segregation model (Schelling, 1971), actors strive to migrate to locations or neighborhoods having a minimum proportion of neighbors in their own category. Likewise, the number of neighbors may also affect opinion formation dynamics. The number of neighbors sharing a given set of opinions can have a bearing on how a single decision maker forms his/her opinion (Flache and Hegselmann, 2001).

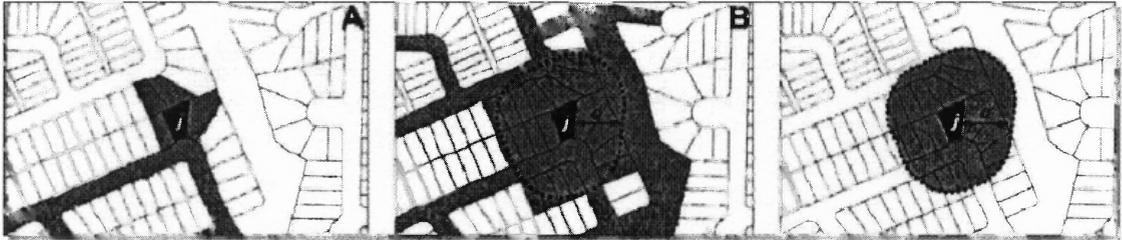


**Figure 2.4** Neighborhood configurations in formal cellular automata.

The challenges posed by defining an appropriate neighborhood configuration for irregular remain. The application of Voronoi diagrams is limited since they do not accurately represent useful land parcel units. Likewise, the generalizing process in Delauney triangulation where each parcel is reduced to a single point disregards the variety in sizes and shapes of land parcels (Stevens et al., 2007). Consequently, defining the neighborhood in this manner is not appropriate for urban land use change modeling.

Stevens et al. (2007) proposed an urban growth model that addresses the issues on irregular spatial structure. In the iCity model of Stevens et al. (2007), three neighborhood

configurations are listed; each one is used for a different function. Depicted in Figure 2.5, these are: adjacency neighborhood, distance neighborhood, and clipped distance neighborhood.



**Figure 2.5** Neighborhood configurations for irregular polygons used in CA modeling. (Adopted from Stevens, 2005)

In Figure 2.5, the gray polygons around polygon  $j$  represent the neighborhood. Three types of neighborhood configurations are shown: (A) the adjacency neighborhood includes all polygons having a common edge with the central polygon  $j$ ; (B) distance neighborhood includes polygons that fall completely or partially within a distance  $d$  map units of polygon  $j$ ; and (C) clipped distance neighborhood includes all polygons that fall completely and portions of the polygons that are partially within distance  $d$  map units of polygon  $j$ . Each neighborhood configuration is used for different functions. Stevens (2005) used these configurations to estimate attractiveness scores for residential land parcels and transition rules for commercial and industrial parcels. For example, a residential land parcel's attractiveness is enhanced if it is adjacent to a park (adjacency neighborhood). Distance neighborhood is used to determine whether a commercial or industrial parcel is within a given distance from a residential land parcel  $j$ . Conversely, a clipped distance neighborhood is used to determine the number of residents within a given distance from a commercial or industrial land parcel  $j$ .

## **2.4 Driving Factors of Land Use Change**

Lu (2002) suggested three criteria in selecting an appropriate set of driving factors for modeling land use change. First, the factors should include all physical, economic, demographic and social factors that affect all types of land use change. Second, they must have spatial attributes. Third, they must reflect the properties and characteristics of the parcel. Elsewhere, Boerboom and Jiao (2003) grouped various driving forces into five categories namely: neighborhood, accessibility, suitability, policy, and socio-economic factors.

### **2.4.1 Neighborhood Factors**

Neighborhood factors refer to the interaction between neighboring cells or between land use types. Values for these effects can either be positive reflecting attraction or negative for repulsion (Landis and Zhang, 1998; Engelen et al., 1999). Depending on the state of its neighbor, a residential cell may stay at its current state or be transformed to another state. For instance, a residential cell may locate adjacent to other cells that are residential (positive influence) and away from commercial and industrial areas, and highways (negative influence). Li and Yeh (2002) represented the neighborhood factors by the quantifying the amount of cropland, orchards, construction sites, built-up areas, forests, and water in a cell's surrounding area. Other models represented neighborhood factors by the percentage of developed cells in the neighborhood of a given cell (Wu and Webster, 1998; Li and Yeh, 2001).

### **2.4.2 Accessibility Factors**

These accessibility factors reflect von Thunen's theory of land use that the demand for sites, measured by land prices and densities, are greatest near major city centers because of minimum transportation costs (Walker, 2004). The accessibility effects on land use are supported by the observation that a city draws more trade from nearby towns than it does from more distant ones (Reilly, 1931). This implies that the intensity of interaction between two entities is dependent on distance. When extending this observation to urbanizing effects, there is a greater likelihood of urban development in areas closer to major transportation routes and urban centers (McMillen, 1989; Clarke et al., 1997; Arai & Akiyama, 2004). Such rule constitutes an essential element of the early models based on Newton's gravity principles. Several studies include transport/location accessibility variables such as distance to the major city, closest sub-urban areas, and closest road in their land use conversion models (Li and Yeh, 2002; Landis and Zhang, 1998). Tyrell et al. (2004) identified the distance to urbanized areas as one of the top drivers of fragmentation of forested lands – a consequence of urbanization thus indirectly land use change. Similarly, Alig et al. (2005) found that land quality represented by attributes such as soil fertility or the distance of urban plots to amenities are significant determinants of forest fragmentation.

### **2.4.3 Suitability Factors**

The suitability factors represent the conditions of the natural characteristic of the cell, such as slope, terrain, and soil type for different land uses. Taillefumier and Piegay (2003) and Bender et al. (2005) assert the importance of slope and elevation in

determining land use in mountain areas because of their impacts on development costs and aesthetic value of a property. Soil type was used in several models to indicate suitability for various land uses (Li and Yeh, 2002; Wu and Yeh, 1997; Voogd 1983). NJWSA (2002) discussed the required soil properties to support six community development applications in the Raritan River Basin including: septic tank absorption, foundations for dwellings with and without basements, local streets and roads, foundations for small commercial buildings, and lawns, landscaping and golf fairways. For example, soil properties for septic tank applications should have good percolation properties to accommodate on-site septic systems for waste treatment and disposal. A severely limited soil is indicated by slow or no percolation lowering the capacity of septic systems to function properly (Quisenberry, Brown, Smith, and Hallahan, 2006; Jantrania, 2004; Middle, 1996). Soil property requirement for foundations for dwellings with and without basements as well as foundations for small commercial establishments are designed to support loads without movement (Nelson and Miller, 1997). Local roads and streets carry automobile and light traffic all year and as such require soil that can easily be excavated and graded. For lawns and landscaping, the requirements are devised to support the growth of turf and ornamental plants. Lastly, since golf fairways are subjected to heavy foot traffic, the soil should have properties that support plant growth at the same time enable traffic mobility once growth is established (USDA, 2006).

#### **2.4.4 Policy Controls and Socio-economic Factors**

Seto and Kaufmann (2003) evaluated urban land use change in the Pearl River Delta in China and listed several socio-economic factors for urbanization and other land use

changes. These factors include direct investment, agricultural land rent, industrial land rent, urban and rural population and wage rates. Direct investment, foreign or domestic, can trigger a demand for more housing units to accommodate an increase in the labor force (Li and Yeh, 2002). Agricultural and industrial land rents represent the land value or the productivity arising from using land. A land with a high value, for example, will more likely remain in its original land use than one with a low value. According to Von Thunen's regional land use model, the relative cost of transporting different agricultural commodities to the central market determines the agricultural land use around a city. In other words, the most productive activities will tend to be located close to the city while less productive ones will locate farther from the city. Population dynamics play a significant role in land conversion as well since urban congestion leads to the expansion of the urban fringe (Entwisle et al., 2005; La Gro and Degloria, 1992).

Land use policies are incorporated into land use change models as independent variables in regression models (Overmars et al. 2007; Braimoh and Onishi, 2007; Fang, Gertner and Anderson, 2007; Silber and Wytrzens, 2006). Another approach to incorporating land use policies is the use of spatial constraints to represent those policies. For example, an open space preservation policy can be modeled by imposing non-development restrictions on those designated open space areas. Using a logistic regression model, Conway and Lathrop (2005b) compared various policy scenarios including down zoning, cluster development, wetland buffers and open space protection in order to evaluate their effectiveness in attaining set ecological goals in Barnegat Bay and Mullica River Watersheds in Southern New Jersey. Down zoning was represented by constructing buffers around urban cells taking into account prescribed house density, e.g.

>1.3 ha in unsewered areas (NJDEP 2001). Cluster development was modeled by relocating newly urbanized dispersed cells near existing urban cells (Conway and Lathrop, 2005b). Protection of wetlands and open space was achieved by designating these areas as non-developable land (Irwin and Bockstael, 2004). Other policies on land use management include preservation of farmlands (Duke, 2004; Nickerson and Lynch, 2001; Tulloch et al., 2003). These policies are implemented using several mechanisms which include the purchase of development rights and the transfer of development rights. Under these schemes, a landowner voluntarily agrees to relinquish development rights over his/her land in exchange for monetary compensation. The landowner may still sell the land with the understanding that the development restriction will apply to the new owners of the property. Through this arrangement, permanent preservation of farmlands is achieved.

Appendix B lists the various drivers of land use change used in several studies reported in the literature.

## **2.5 Generating Transition Rules**

In a CA model, a transition rule defines how driving factors determine the evolution of a cell from one state to another. As discussed previously, CA-based land use change models vary by how the transition rules are derived. In general, there are four methods to generate transition rules.

### **2.5.1 Regression Method**

The transition rules can be specified as a regression function as follows:

$$(2.8) \quad \text{Transition potential (T.P.)} = \sum_i^n \alpha + \beta_i X_i + \varepsilon$$

where  $\alpha$ ,  $\beta$  are the regression coefficients,  $\varepsilon$  is the error term, and  $x$  are the independent variables or the transition rule parameters that represent the driving factors. These driving factors can be any of the factors enumerated in section 2.4. Wu (2000) used neighborhood factors; Arai and Akiyama (2004) added accessibility factors while Sui and Zeng (2001) included neighborhood, accessibility and suitability factors. The transition potential is the probability whether a cell evolves from one land use type to another. After sample data are derived for change in land use type and driving factors, a multiple logistic regression procedure is usually performed to estimate the coefficients of the independent factors and the estimated regression model is subsequently used to estimate the transition potentials of all the cells in the study area. A threshold value such as 0.5 can be used to determine whether a particular cell remain in its original land use type or evolve into a land use type such as developed land.

### **2.5.2 Artificial Neural Network–based Method**

Statistical models present difficulties in the selection of a complete and appropriate independent variables, the classical “identification problem” in regression analysis since the method is dependent on the modeler’s experience and professional background. An alternate method that addresses this limitation is an artificial neural network (ANN) (Liu et al., 2008; Guan, Wang, and Clarke, 2005; Aitkenhead, Mustard, and McDonald, 2004; Li and Yeh, 2002). The artificial neural network approach is appropriate for large datasets and where the relationships between dependent and independent variables are not readily apparent (Razi and Athappily, 2005).

In a neural network model, simple nodes (called variously "neurons" or processing units) are connected together to form a network of nodes, similar to the vast connection of neurons in the brain, hence the term "neural networks" (Li and Yeh, 2002). Nodes in a single or several hidden layers can process several variables or driving factors in the input layer. The system adapts/changes its structure based on external or internal information that flows through the network, thereby finding patterns in the data linked to a given output (Pijanowski et al., 2002). The ability to process large datasets makes neural network-based modeling useful in land use change studies.

ANNs are not too dependent on expert knowledge to determine relevant influencing factors in developing transition rules. Furthermore, ANNs can be particularly useful in situations involving a non-linear complex relationship between driving factors and land use change (Dai et al., 2005; Pal and Mather, 2002). However, ANNs have drawbacks because of their being a black box. Consequently, their usefulness is not much appreciated by users with limited analytical background (Razi and Athappily, 2005; Li & Yeh, 2004).

### **2.5.3 Multiple Criteria Evaluation – Analytical Hierarchy Process Method**

Wu and Webster (1998) developed transition rules using multiple criteria that reflected people's preferences in development of urban areas. The driving factors in the transition rule were selected based on their knowledge of the area. The logistic regression was used to narrow down the list of factors for inclusion in the analysis. Weights were assigned to each of these factors based on decision makers' opinion on what and how urban development should proceed. The opinions of all stakeholders were also considered in estimating weights for these factors. It can be seen that the performance of the model

depended on a large extent on the knowledge of all decision makers on urban development. However, this procedure suffers from uncertainties due to subjectivity on the part of the decision makers.

#### **2.5.4 Fuzzy Logic**

Fuzzy set theory deals with ambiguities of real-world phenomena which cannot be defined strictly as “completely true (having a degree of truth 1.0)” or “completely false (having a degree of truth 0.0)” (Zadeh, 1972). It is applied as a method of reasoning (Fuzzy logic) to interpret vague concepts, such as blackness or crowdedness. For example, it might be said that an area is crowded, with a degree of truth of 0.8. Fuzzy logic is also employed in geographical studies to represent landscape in a flux.

Prior to its application to CA modeling, fuzzy logic was used extensively in other GIS applications. Fuzzy logic was used in land suitability analysis work done by Wang et al. (1990), and Davidson et al. (1994). Kollias and Voliotis (1991), and Wang (1994) used fuzzy logic in GIS query functions. Altman (1994) and Banai (1993) used it to represent geographic features and spatial relationships in spatial analysis.

Wu (1996) introduced the use of fuzzy logic in defining the transition rule for a CA. Whereas the definition of a CA set of transition rules in other models are expressed as crisp rules (Dragicevic, 2004) e.g. Boolean representation either 1 or 0, fuzzy sets are more akin to uncertain knowledge of system behavior. Taking the case of Conway's game of life, the rule dictates that a cell dies if surrounded by three or more live cells because of overcrowding. In fuzzy CA, rules can be stated to allow for degrees of overcrowding, e.g., crowded, very crowded, not so crowded, etc. (Wu, 1996). The fuzzy

set theory developed by Zadeh (1972) enabled rules to process “soft” data as contrasted to “hard” (definite) data in conventional CA. Transition rules are formulated using several criteria. An example is carrying criteria used by Wu (1996) as shown in Equation (2.9). Wu (1996) defined carrying capacity (land availability) as the overall proportion of undeveloped land to land area:

$$(2.9) \quad Capacity = \begin{cases} 0 & \text{if } \sum count(all) - count(water) = 0 \\ \frac{count(cultivated) + count(wood)}{\sum count(all) - count(water)} & \text{otherwise} \end{cases}$$

To illustrate the above criteria, a pixel has a zero carrying capacity when the count of all pixels (of all land uses) in its neighborhood is equal to the count of water pixels. This is a situation where there is no room for urban development since the pixel is surrounded by water. In the case of an inland pixel, land availability for urban development is computed as the ratio of undeveloped land (total of cultivated and wood land) to the total pixels representing the total land area. Three other criteria are used such as agglomeration, accessibility (to transportation and urban centers), and uniformity (to discourage scattered development). These criteria are evaluated for each land conversion/transitions and assigned a membership value. These are conversion to urban land, preservation of cultivated land, and preservation of woodland. Once all individual criteria have been assigned membership values, the composite membership value is calculated. The transition with the highest membership value is then used to determine the new state of the cell.

At this point, it is worth stressing that the threshold values used in the membership functions were formulated based on expert knowledge. As such, this method

is susceptible to uncertainties caused by subjectivity like the other methods discussed above.

A final method of eliciting transition rules is through the use of decision trees. Like ANNs and fuzzy sets, it is adept at handling a large set of parameters. This method is further explained in the next section.

## **2.6 Decision Tree**

In geographical studies, the decision tree (called data mining in other studies) was initially used as a classification tool (Moore et al., 1991; Meyer, et al., 2001; Speybroeck et al., 2004, Wu, Silvan-Cardenas, and Wang, 2007). Basically, a decision tree's structure entails a series of yes/no questions in which the sequence of the questions that are asked depends on the answers given in the previous question. Applied to land use/cover classification, the specific questions assume values equivalent to land attributes, the sequence of which eventually determines the appropriate land use or land cover classification (Aalders and Aitkenhead, 2005). Aside from its application for mapping purposes, decision trees have been used to predict land use changes (McDonald and Urban, 2006; Liu, et al., 2007). Its mechanism of classification and its application in generating transition rules for cellular automata models are explained below.

### **2.6.1 Generating the Decision Tree**

One of the earliest decision tree learning programs is the ID3 developed by J. Ross Quinlan in 1975. Later versions developed include C4.5, CART, IB1, IB2, MPIL1, and MPIL2. Of these seven algorithms, C4.5 as developed by Quinlan (1991) remains the

most popularly used for data mining. In addition, Quinlan (1999; 1996) stressed that C4.5, unlike many statistical approaches, does not depend on assumptions about the distribution of variables or on the independence of the variables themselves.

Initially, the program selects a case at random from a training dataset,  $S$ , and then classifies it to a particular land use class. The entropy (or information) from this procedure is calculated using equation 2.10 (Quinlan, 1996b) taking the case of two land use types namely urban and non-urban:

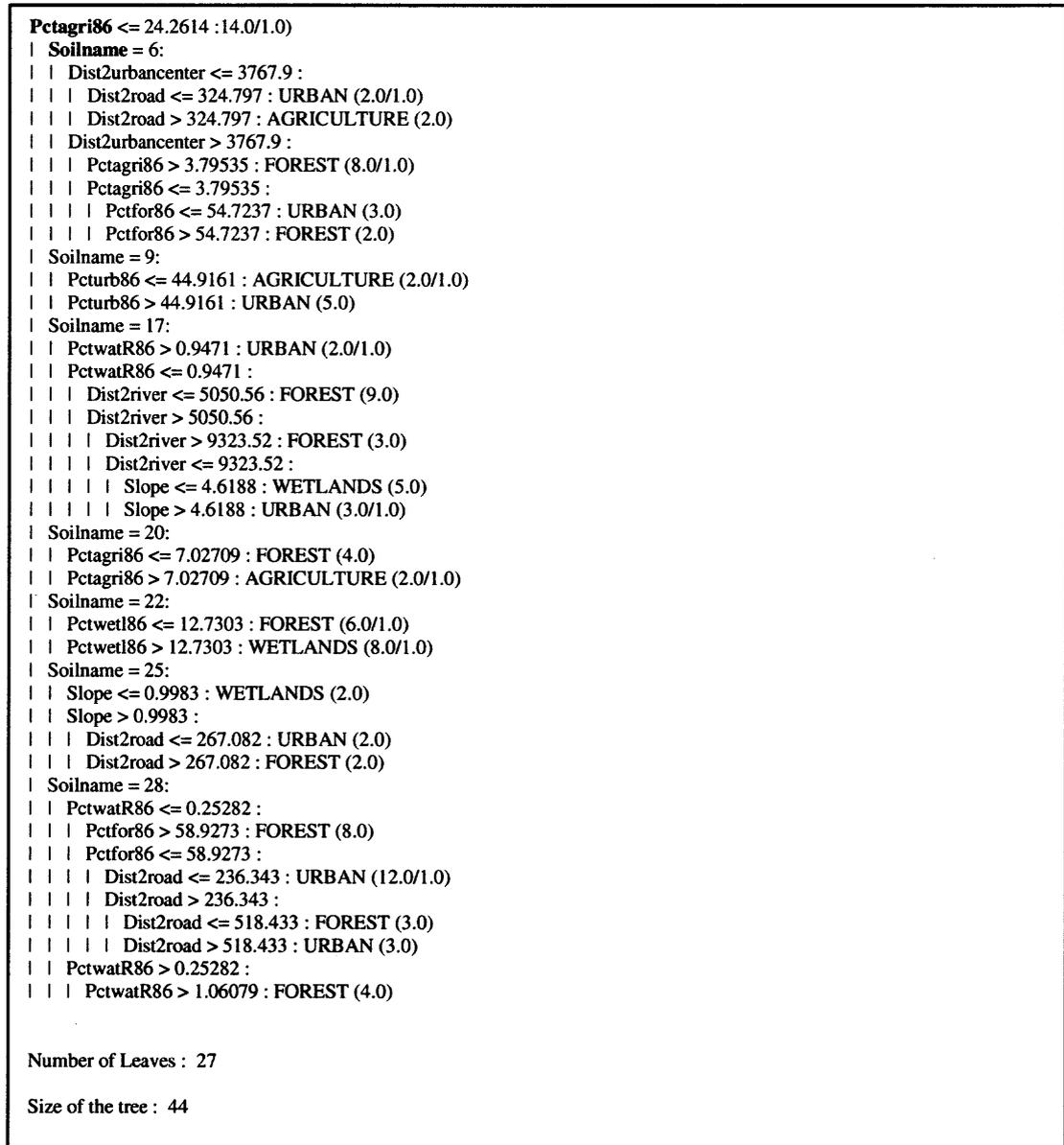
$$(2.10) \text{ entropy}(S) = - \left( \left( \frac{\text{freq}(\text{urban})}{|S|} \cdot \log_2 \frac{\text{freq}(\text{urban})}{|S|} \right) + \left( \frac{\text{freq}(\text{non-urban})}{|S|} \cdot \log_2 \frac{\text{freq}(\text{non-urban})}{|S|} \right) \right)$$

where  $\text{freq}(\text{urban})$  is the number of instances in  $S$  belonging to class **urban**, and  $|S|$  is the total observations in  $S$ . The entropy of a dataset is a measure of the non-uniformity of the data i.e., the higher the entropy, the more disordered the data are. Elsewhere, entropy is called information since the higher the entropy, the more information is required to describe completely the data (Goodman and Smyth, 1988). Let  $S$  be partitioned into  $n$  outcomes for a test attribute, say **soil type**. The entropy from this set can be written as:

$$(2.11) \text{ entropy gain}(S, \text{urban}) = \sum_{i=1}^n \frac{|\text{freq}(\text{soiltype} = i)|}{|S|} \text{entropy}(S)$$

Information gained by splitting dataset  $S$  using attribute **soil type** is calculated by subtracting equation (2.11) from equation (2.10). This concept of entropy gain is what drives the decision tree generation process. Recursive splitting of dataset  $S$  is governed by maximum entropy gain applied at each node until each leaf node contains only observations from a single class indicated by a zero entropy gain. For this reason, entropy

gain is used to define a preferred sequence of attributes that most rapidly converges to a leaf node. The result of the above procedure is a sequence whose structure is called a decision tree. A portion of a decision tree is revealed in Figure 2.6 showing 24 terminal nodes.



**Figure 2.6** Decision tree with 27 terminal nodes.

The sample decision tree shown above was obtained using the variables percent land use in the neighborhood of the parcel, distance to road, distance to urban center, soil name, and distance to river. It has a total of 27 terminal nodes. It indicates that the percent of agricultural land in the neighborhood and soil name which split the dataset at the highest nodes are the most important classification variables.

### 2.6.2 From Decision Tree to Transition Rule

The form of the decision tree shown in Figure 2.6 can sometimes be too large and complicated to use directly. It is thus necessary to transform or present the decision tree in another form such as a transition rule. Each path in the decision tree structure from the root to the terminal node or leaf can be translated into a transition rule. To illustrate, the path from top (root) to the first two terminal nodes (leaf) in Figure 2.6 can be represented as two transition rules:

Rule 1:

(percent agriculture  $\leq$  26.34) and (soilname is 6) and (distance to urban center  $\leq$  3767.9) and distance to road  $\leq$  324.7) ----- > URBAN

Rule 2:

(percent agriculture  $\leq$  26.34) and (soilname is 6) and (distance to urban center  $\leq$  3767.9) and distance to road  $>$  324.7) ----- > AGRICULTURE

Other transition rules can be extracted from the tree by following the next sequence all the way to a terminal node. Since there are a total of 27 terminal nodes, the decision tree will yield an equivalent number of transition rules.

### 2.6.3 Advantages of Using Decision Tree

In the traditional methods of eliciting transition rules, expert knowledge is required in selecting the appropriate factors to include in the transition rule. As a consequence, such methods are prone to subjectivity error since it is subject to the influence of individual knowledge and the preference of whoever is making the decisions. Moreover, the process becomes more complicated when the functional relationship between independent and dependent variables is unknown. There is also a tendency to include more explanatory variables in the regression equation as a good way of avoiding bias. Although omitted variables and the bias they cause are avoided using this strategy, the irrelevant variables that will inevitably be present can cause high variances (Kennedy, 1998).

Creating transition rules through the decision tree method is more helpful than using statistical regressions in the following cases where: (1) there is a large number of attributes to project land use change (Speybroeck et al., 2004; Pal and Mather, 2003, 2002); (2) there is the existence of non-linear relationships between variables in the data (Razi and Athappilly, 2005); and (3) the underlying relationship between dependent and independent variables is not known (Pal and Mather, 2002). Although decision trees share these advantages with ANN approach, the latter is not intuitive to policy makers and land use planners because of its black box nature. Also, in many instances, the decision tree can achieve comparable if not better results than ANNs (Pal and Mather, 2003; Ghaffari et al., 2006).

Compared to other machine learning algorithms such as ANNs, a decision tree has more advantages. Decision tree programs can automatically perform feature selections from the raw dataset (Perner et al., 2001). During the feature selection, only the

most relevant attributes are chosen from the whole set of attributes for the construction of decision rules in the nodes. A number of studies show that neural classifiers depend on a range of user defined factors that consequently limit their use whereas a decision tree requires a smaller number of these factors (Pal and Mather 2002). Decision tree program performs better with smaller datasets even with the presence of a large number of irrelevant attributes (Brown, Corruble, and Pittard 1993). Data preparation for a decision tree is minimal (Perner et al., 2001; Pal and Mather, 2002). Decision trees can handle a broad range of data types including nominal and categorical data. Neural methods can only be used with numerical variables while other models can only work with nominal variables (Quinlan 1996a;1996b). Decision tree is a white box model. A decision tree is more easily interpreted compared to a neural network where the derivation of the results is not explained or readily available. Input variables are processed in hidden layers within the network that do not easily lend to interpretation by the end-user, e.g., decision maker (Breiman et al., 1984; Quinlan, 1996b; Li and Yeh, 2004). Decision tree processing entails shorter processing/computing time (Pal and Mather, 2002; Perner et al., 2001; Razi and Athappily, 2005).

## **2.7 Accuracy Assessment**

The complex nature of human-environment interactions leading to land use change presents many challenges to the modeler and researcher. Ranging from the conceptual to the empirical, these challenges provided the impetus in the past to develop models that properly inform policy and practice. A careful consideration of a model's prediction accuracy is necessary for modeling to continue this role in land use planning. To this

end, accuracy assessment should not merely quantify but also provide a framework by which it is interpreted in a meaningful manner.

The value of a land use map depends to a large extent on the accuracy as referenced to a truth map or reality map. Congalton (1991) traced the development of accuracy assessment studies in four stages that started as a simple visual inspection of the map. This early method entailed considerable errors as it is fraught with subjectivity by the one making the assessment. A second stage sought to overcome subjectivity by including a quantity aspect to the assessment. Such method involved comparing the areal extent of each predicted class to the areal extent in a reference map. Note that this method considers only aggregate area of a predicted class. It does not give information as to location. It may give a correct proportion of each class relative to a reference map but it may not be in the correct location. Not only quantity but also quality is reflected in methods belonging to the third stage. It sought to remedy the non-site specific drawback of the previous method by computing the percentage accuracy for a given class in a specific location. The percentage accuracy is computed for the cases that were correctly assigned in that particular class. This process is done repeatedly for the entire study area to come up with an overall accuracy. The last stage is a further improvement as it gives more information thus presenting accuracy in more meaningful form. A common feature of this stage is the use of a confusion matrix (sometimes referred to as contingency table or error matrix). A confusion matrix is simply a cross tabulation of the predicted class against the same class in a reference map. The diagonal element of the matrix represents correct classifications. Referring to the confusion matrix in Figure 2.7, overall accuracy is calculated as the sum of the items in the diagonals divided by the sum of all row and

column entries. Although overall accuracy is useful, it does not give much information about the accuracy of the individual land use types. Some assessment studies supplement it with an evaluation of user's and producer's accuracy sometimes referred to as errors or commission and errors of omission respectively. User's accuracy indicates the probability that a sample from a classified (predicted) data actually represents that category in the reference map (Story and Congalton, 1986). It is computed by dividing the number of correctly classified cases by the total number of cases that were classified as belonging to that land use type. Producer's accuracy, on the other hand expresses the probability that a reference sample has been classified correctly. This is computed by dividing the number of cases that have been classified correctly by the total number of reference cases in that land use type.

Another criticism of overall accuracy metric is that it does not take into account correctly classified cases as a result of pure chance from those that come from the model itself. As such, it can happen that the overall accuracy is an overestimate (Pontius, 2000; Congalton, 1991).

		Reality				
Simulated		A	B	C	D	$\Sigma$
A		$n_{AA}$	$n_{AB}$	$n_{AC}$	$n_{AD}$	$n_{A+}$
B		$n_{BA}$	$n_{BB}$	$n_{BC}$	$n_{BD}$	$n_{B+}$
C		$n_{CA}$	$n_{CB}$	$n_{CC}$	$n_{CD}$	$n_{C+}$
D		$n_{DA}$	$n_{DB}$	$n_{DC}$	$n_{DD}$	$n_{D+}$
$\Sigma$		$n_{+A}$	$n_{+B}$	$n_{+C}$	$n_{+D}$	$n$

Percentage correct =  $\frac{\sum_{k=1}^q n_{kk}}{n} \times 100$

Kappa coefficient =  $\frac{n \sum_{k=1}^q n_{kk} - \sum_{k=1}^q n_{k+} n_{+k}}{n^2 - \sum_{k=1}^q n_{k+} n_{+k}}$

User's accuracy =  $\frac{n_{ii}}{n_{i+}}$

Producer's accuracy =  $\frac{n_{ii}}{n_{+i}}$

**Figure 2.7** Confusion matrix for two maps with common measures of classification accuracy. Adopted from Foody (2002)

A more accurate metric that compensates for this feature is Cohen's Kappa coefficient. Kappa has a maximum value of 1 indicating perfect agreement and a minimum of zero indicating maximum uncertainty. Kappa is evaluated using the following formula (Foody, 2002):

$$(2.6) \quad \hat{\kappa} = \frac{P_{correct} - P_{chance}}{1 - P_{chance}}$$

where  $P_{chance}$  is calculated as:

$$(2.7) \quad P_{chance} = \sum_{i=1}^n P_{row(i)} P_{column(i)}$$

where  $P_{row(i)}$  is the proportion of total entries in row  $i$ , and  $P_{column(i)}$  is the proportion of total entries in column  $i$ . Referring to the confusion matrix shown in Figure 2.7, the Kappa index is thus calculated as:

$$(2.8) \quad \hat{\kappa} = \frac{n \sum_{k=1}^q n_{kk} - \sum_{k=1}^q n_{k+} n_{+k}}{n^2 - \sum_{k=1}^q n_{k+} n_{+k}}$$

While it has been suggested that Kappa be adopted as the accuracy standard (Smits et al., 1999), some quarters argue that although Kappa compensates for chance agreement, such capability is not unique to this metric and that the Kappa coefficient alone does not suffice as indicator of accuracy assessment (Foody, 2002; Pontius, 2000; Stehman and Czaplewski, 1998). Moreover, Pontius (2000) contends that Cohen's Kappa has several limitations including insensitivity to location and quantification errors. Cohen's Kappa does not penalize for large errors or reward close predictions. Consequently, Pontius (2000) suggested separately considering quantity (i.e., proportion of area allocated to each land use type) and location. Accordingly, Pontius (2000) derived three variants of Cohen's Kappa namely: Kappa location, Kappa quantity and Kappa<sub>na</sub>. Kappa location as the name suggests is an accuracy measure indicating success in predicting location. Similarly, Kappa quantity indicates ability to successfully predict quantity while Kappa<sub>na</sub> indicates no ability to correctly specify location and quantity. Elsewhere, Hagen (2003) and Hagen-Zanker et al. (2005) introduced a kappa variant for use in cellular models involving raster data. Hagen (2003) developed the Fuzzy Kappa statistic using fuzzy set theory. Fuzzy set theory was used to assess map representations and map comparisons in studies made by Metternicht (1999), Lewis and Brown (2001), and Power et al. (2001). A

positive aspect of Fuzzy Kappa is its ability to mimic the way humans compare maps (Visser and de Nijs, 2006; Fritz and See, 2005).

Foody (2002) in his review of the status of land cover classification accuracy noted that despite calls for a single standard for accuracy assessment, the call has remain largely unheeded. Langford et al. (2006) report that techniques for assessing classification accuracy are still generally specified in terms of percentage of misclassified pixels. Other metrics have also been used including entropy (Yeh and Li 2001), and Moran's I (Wu, 2002; Li and Yeh, 2004).

Given several options to use in accuracy assessment, Foody (2002) suggested using a general format of reporting accuracy that would consist of at least one metric (Muller, 1998; Stehman, 1997), minimum level of overall accuracy of at least 85% with no single class accuracy below 70% (Thomlinson et al, 1999), and a confusion matrix (Stehman, 1997).

## **2.8 Summary**

The literature survey presented above described the emergence of CA in land use and land cover change modeling. CA addressed some of the limitations of large urban models that caused the failure of some to develop fully to the operational stage. Its popularity can be traced to several advantages including as its ability to handle spatiality better, a decentralized approach, simplicity, ability to integrate micro and macro modeling approaches, and affinity to GIS.

In order to apply the technique in an urban context, the basic formalisms of the CA were relaxed. Different modifications were made in the four elements of CA

including lattice, cell state, transition rule and time. Focusing on the transition rule element, this literature survey presented current ways by which they (transition rules) are developed. The challenges in eliciting suitable transition rules were discussed including those that a decision tree is most adept at resolving. Being in a more explicit format than the mathematical equations used in the other methods, decision trees are more intuitive thus more useful to users of the model. Defining what constitutes the cell's neighborhood was another focus of these efforts. The survey concludes with current techniques by which prediction accuracy is assessed.

## **CHAPTER 3**

### **METHODOLOGY**

This chapter describes the methodology used to develop a parcel-based prediction model for land use change using machine learning and interpretation algorithms. It is divided into six sections. The first section describes the method used to prepare parcel-based land use for the model. The second section describes the neighborhood configuration from which driving factors are derived. The third section elaborates on the two components of the model: decision trees and cellular automata sub-models. The fourth section deals with the implementation tools to run the model. These tools are J48 Classifier to implement the decision tree and Agent Analyst for cellular automata. The remaining two sections discuss model validation and sensitivity analysis.

#### **3.1 Developing a Parcel-Based Land Use Map**

This research used the parcel as a basic unit of analysis in its land use change prediction model with its application to Hunterdon County in New Jersey. Accordingly, a single land use class should be assigned to each parcel as do the rest of driving factors for land use change. The other alternative would be to proceed with the modeling process using a multiple land use class for the parcels. However like the raster data this scheme also requires re-aggregation to larger units that are useful in the planning process, i.e. re-aggregation of model prediction of each sub-parcel. As pointed out in the literature, this can lead to the so called Modifiable Areal Unit Problem caused when boundaries of re-aggregated units do not exactly coincide with the larger unit (Openshaw, 1983).

Although the parcel is the proper unit of analysis for land use change modeling, it is not a simple task to assign a single land use class to each parcel in a study area. The local land assessor's office may have records on the intended use of a parcel, but the actual land use might be different. The land use classification used for taxation purpose might not be consistent with the need to understand the land use change process. In addition, records on the historical land use changes for all parcels in a study area might be enormous. It is difficult to trace all the land use changes in all parcel in a larger study area based on land assessment data.

Alternatively, land use change modelers tend to develop land uses in parcels from land use maps derived from satellite images and aerial photography through image identification and interpretation (Erol and Akdeniz, 2005; Aplin and Atkinson, 2001; Aplin, Atkinson, and Curran, 1999).

Most land use maps are prepared using the conventional per-pixel classification method which uses the spectral signature of each pixel to allocate it to a land use class. Although the per-pixel classification method has superior capabilities, it has its limitations and disadvantages. Spectral information can contain noise data resulting from atmospheric effects (Smith and Fuller, 2001) and natural variations within land cover types (Hutchinson, 1982). Pixels in low resolution maps that cross land cover boundaries will have multiple reflectance data (Fuller, Wyatt, and Barr, 1998; Crapper, 1984). Because of such disadvantages, several authors affirm that per-parcel field classification is the more suitable and accurate approach (Erol and Akdeniz, 2005; Aplin and Atkinson, 2001; Berberoglu, Lloyd, Atkinson, and Curran, 2000; Tso and Mather, 1999). Alternatively, the per-parcel method has the added advantage of allowing the use of

certain parcel attributes as classification criteria (Wu, Silvan-Cardenas, and Wang, 2007). Other researchers have combined these two approaches. Based on spectral, textural and contextual properties, land use type is classified by each pixel (Munoz-Villers and Lopez-Blanco, 2008; Erol and Akdeniz, 2005; De Wit and Clevers, 2004; Dean and Smith, 2003; Fuller, Smith, Sanderson, and Thompson, 2002). The majority class of pixels within a parcel is the designated class for that field or parcel.

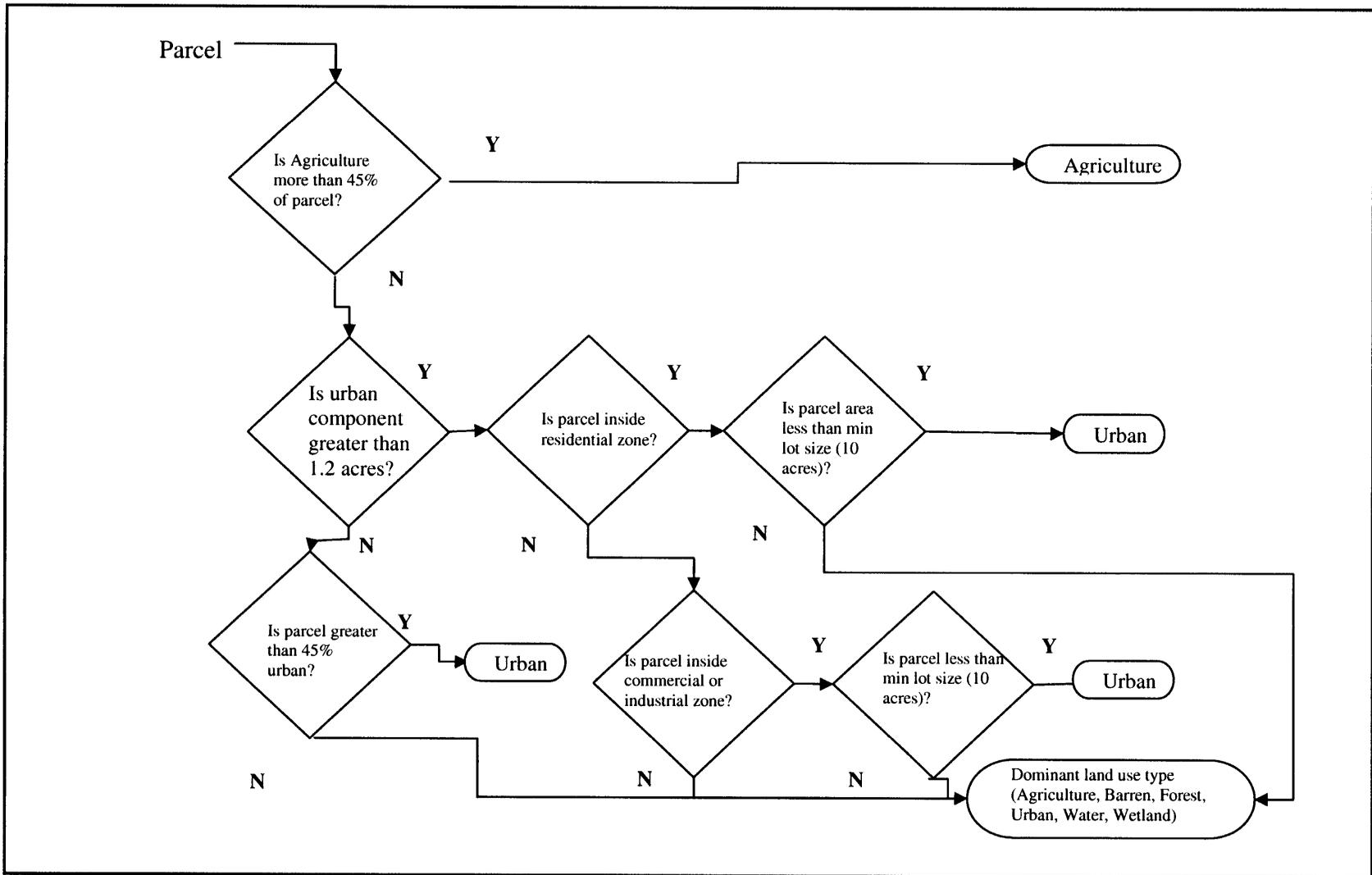
Wu et al. (2007) used tax parcel boundaries in preparing a parcel land use map. In generating a detailed urban land use map, they used geometrical, textural and contextual information as parameters to characterize different types of urban use. A total of 12 field attributes were used to characterize 9 urban land use types. The attributes used included building count, building shape, proximity to road, parcel size, maximum building area, and others. The classification method of Wu et al. (2007) is remarkable for the fact that it emphasizes the use of tax parcels instead of remote sensing images, thereby reducing costs.

This research develops a parcel-based land use map from the historical land use maps derived from aerial photography and remote sensing data. The historical land use maps for 1986, 1995, and 2002 were developed and are maintained by NJDEP and the parcel boundary in the study area is maintained and provided by Hunterdon County Division of GIS. The land use classification scheme shown in Figure 3.1 is used to assign a single land use/land cover type to each parcel. Each parcel is initially tested for any agricultural area present. A 45 percent threshold value was used as a filter to classify parcels as agricultural land. This threshold is based on the percentage of agricultural land in all parcels in 2002, which has a mean of 19.19 percent with a standard deviation of

25.85 percent. To take into account large agricultural parcels, the mean plus a standard deviation was used as the threshold value.

To classify parcels as urban the first consideration is that any agricultural portion contained within the parcel should be less than 45 percent of the parcel area. Furthermore, the urban component should be greater than 1.2 acres, located in a residential zone, and have an area that is less than the minimum lot size. On the other hand, if a parcel has less than 1.2 acres devoted to urban use and the urban portion represents more than 45 percent of the total parcel area, it is also classified as urban. A threshold value of 1.2 acres is employed to represent the house footprint that includes the area occupied by house, driveway, patio, pool, etc. (NJWSA, 2003).

Should a parcel fail these filters, it is subjected to two further tests before it is finally rejected as non-urban land. The parcel should be situated in a commercial or industrial zone and be less than the zone-specified minimum lot size of 10 acres. The latter specification suggests that parcels that are less than the minimum lot size will not undergo transformation to another land use since the size of the parcel rules out subdividing it to accommodate other land use types. Failing any of these last two filters, the parcel is a candidate for the other land use types (e.g. wetland, barren, forest or water). The final classification is assigned to the dominant land use in the parcel.

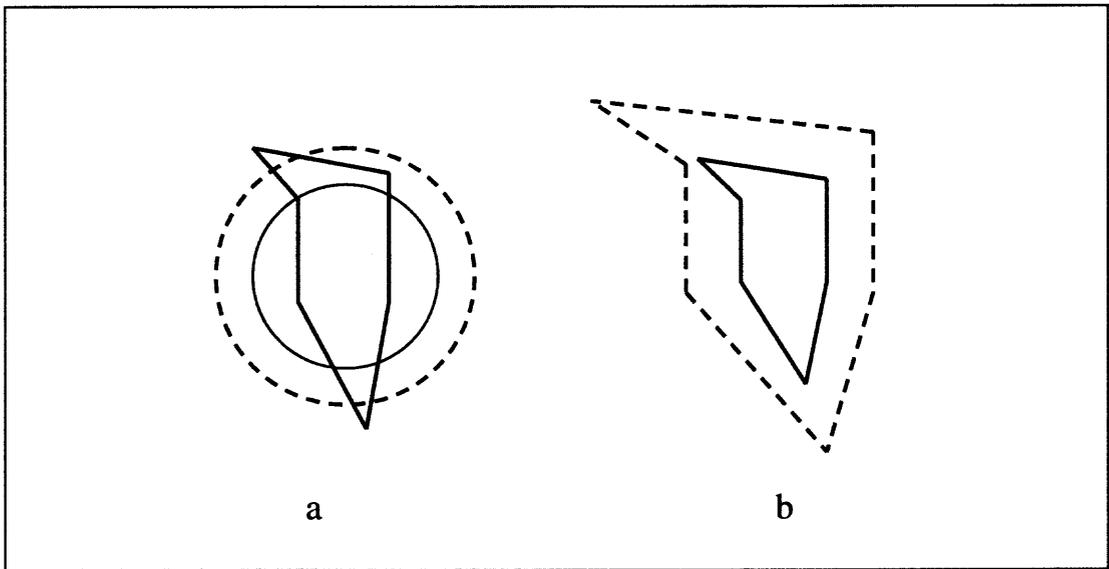


**Figure 3.1** Land use classification criteria.

### 3.2 Neighborhood Configurations

Another dataset to be specified comprises land use distribution in the neighborhood of each parcel. The usual neighborhood configurations used in CA models include the von Neumann or the Moore patterns that assume a lattice composed of regularly shaped cells or grids. Since irregular shaped parcels are used in this research, an alternative neighborhood configuration is needed. Irregular shapes based on a Voronoi diagram were used in CA models of Shiyuan and Deren (2004), Shi and Pang (2000), and O'Sullivan (2002). However, these shapes do not correspond to real geographical objects such as cadastral parcels (Stevens, 2005; Moreno and Marceau, 2007). In his parcel-based cellular automata model on urban sprawl (iCity model), Stevens (2005) used three neighborhood patterns to evaluate site attractiveness for residential development. These configurations are depicted in Figure 2.5 as simple adjacency (Pattern A), distance neighborhood (Pattern B), and clipped distance neighborhood (Pattern C). Stevens (2005) hypothesized that in urban sprawl modeling, a major consideration for selecting areas for residential development is the site attractiveness or desirability. Attractiveness scores were estimated based on a site's proximity to parks and unattractiveness in the case of proximity to industrial or commercial areas. For adjacency to parks, Pattern A was used. Pattern B was used to determine whether commercial land is within a distance "d" from residential parcel  $j$ . Pattern C, the third configuration was used to determine the number of residences within distance "d" from commercial parcel  $j$ . This research modified the neighborhood configuration Pattern C to estimate areal distributions of the different land use types in a parcel's immediate neighborhood.

Instead of depicting a parcel's neighborhood as a circular buffer with a fixed radius from its centroid as shown in pattern C, this research used a buffer drawn around the edge of each parcel as presented in Figure 3.2. This approach is adopted primarily to minimize land use identification errors. When using a concentric circular buffer (represented in dashed lines in Figure 3.2a), portions or even a whole neighborhood parcel may be excluded, thereby leaving some areas in a parcel's neighborhood unaccounted for. On the other hand, by using the edge of a given parcel as the basis for creating a buffer as in Figure 3.2b, such omission errors may be avoided.

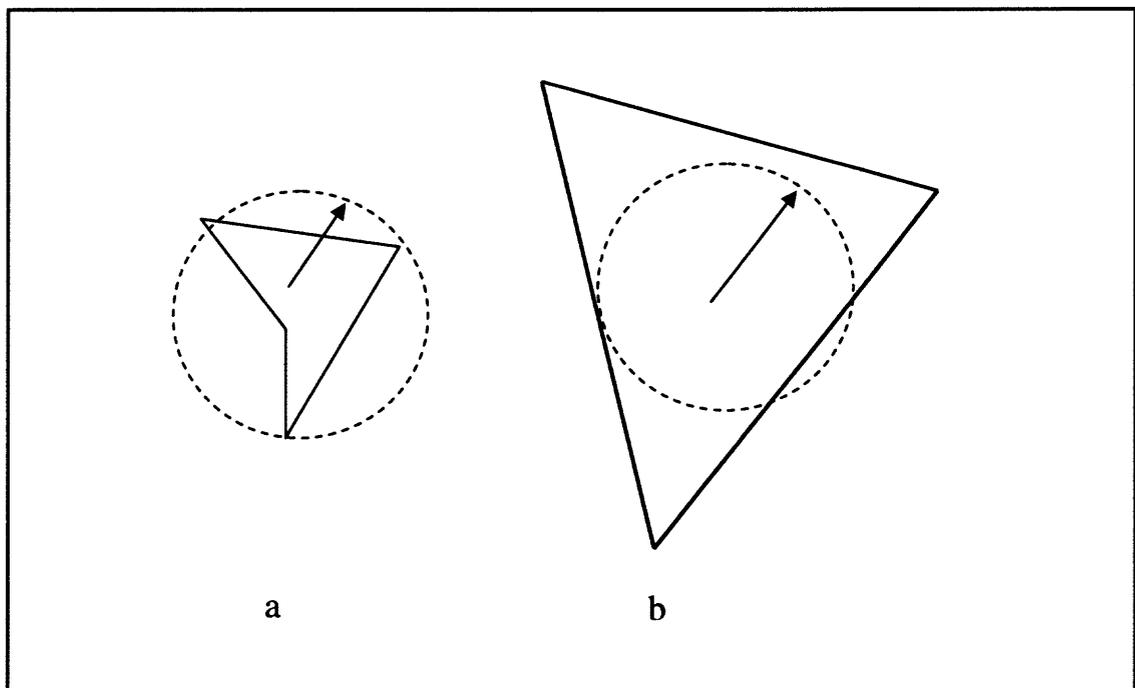


**Figure 3.2** Neighborhood configurations with a buffer from the centroid (a) and the edge (b) of a parcel.

Concern about such omission errors has also been shared by economists who study how land values are affected by the surrounding land use distribution. In estimating the acreage of different land use types around a given parcel, Geoghegan et al. (1997), Geoghegan (2002), and Irwin and Bockstael (2001) made use of concentric circular buffers. However, in their studies on the use of parcel-level data in developing a hedonic

land pricing model, Lynch, Bucholtz and Malcolm (2002) argue that such an approach may produce biased and inconsistent parameter estimates. The neighborhood configuration with a buffer from the edge of the parcel might help improve their hedonic pricing model.

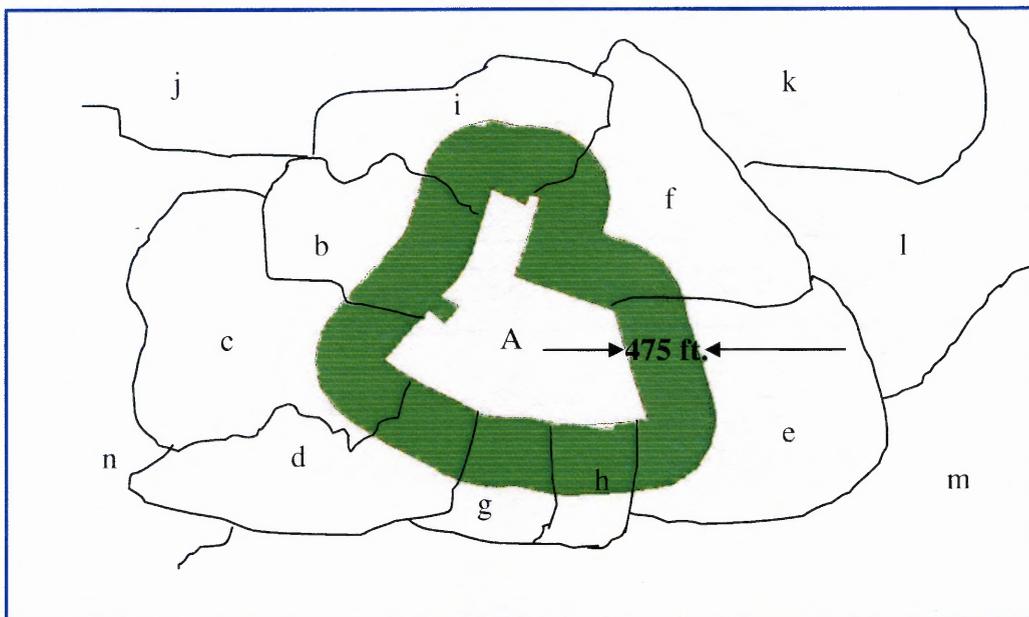
The neighborhood configuration with the buffer from the edge of each parcel ensures that a neighborhood is created outside the parcel as shown in Figure 3.2b. Such property is extremely important especially when the parcel itself is large. An example of such a situation becomes apparent in a parcel shown in Figure 3.3b. It appears that a buffer drawn from the center of a large parcel does not correctly represent a neighborhood area at all.



**Figure 3.3** Neighborhood configurations with buffer from the centroid for a small (a) and big (b) parcel.

The neighborhood configuration to be used in this research is shown in Figure 3.4. An external buffer around the edge of a parcel defines the neighborhood of the

parcel. For Parcel A, the neighborhood includes the portions of Parcels b, c, d, e, f, g, h, and i within the external buffer. A distance of 475 feet was used initially as the buffer thickness. The use of 475 feet is based on the average size of all parcels in Hunterdon County. Assuming that all the lots are squares, the side length of an average lot is estimated at 475 feet. The percentages of different land use types within the buffer are calculated by overlaying the buffer with the land use maps. These numbers are used as important parts of the driving factors in the land use change model to determine the future land use of the parcel.



**Figure 3.4** Neighborhood configuration as defined in this study.

A script was written in ArcView 3.x Avenue Scripting Language to define the neighborhoods of the parcels in the study area and to calculate the percentages of land use types within each neighborhood (See Appendix A). Since land uses in the neighborhood of a parcel have important impacts on the future land use of the parcel, a sensitivity analysis was also conducted by varying the size of the buffer that defines the neighborhood as discussed later in this chapter.

### 3.3 Model Components

The proposed land use change prediction model has two modules. The first module consists of a machine learning algorithm using a decision tree to generate the transition rule and the other module is a cellular automata to predict future land use change. A set of transition rules is expressed as a combination of driving factors leading to a future land use class. The driving factors as discussed previously are: present land use type, percentages of land use types in the neighborhood, distance to nearest urban center, distance to major roads, distance to streams, soil suitability, slope, parcel area, and wetland area inside of a parcel if any. A decision tree takes in all possible values of these driving factors and examines all their possible combinations or permutations. It then compares this combination of attributes to the land use designation of the parcel possessing that combination of attributes in order to elicit patterns in the training data. This stage constitutes the learning phase of the land use change model. Subsequently, the generated pattern of driving factors is arranged into a tree structure and is received as transition rule input by the cellular automata sub-model in the next stage. A parcel land use map with accompanying spatial data is keyed in as additional input to the cellular automata sub-model. The cellular automata algorithm then examines the spatial data of each parcel within the study area and applies the transition rule to decide the next state of the parcel. Policy variables are also processed by the model in the form of rules that override transition rules in certain instances. A policy on farmland preservation, for example, is recorded by the model and is used to designate parcels as non-developable

regardless of the land use type that a transition rule may decree. The two modules are thoroughly discussed below.

### 3.3.1 Decision Tree Module: Generating the Transition Rule

The decision tree model is used to generate the transition rules for the land use change model. Specifically, the parcels that change in land use type between 1986 and 1995 and their corresponding attributes were noted and were used as the training data in the decision tree model. For illustrative purposes, a hypothetical set of input data is shown in Table 3.1. The case represents the parcel number. The current land use is the initial land use state of the parcel at time  $t$  (i.e., 1986). The percent of agricultural land, percent of urban land and percent of forest land are the percentages of agriculture, urban and forest uses in the surrounding neighborhood of the parcel, respectively, as defined above. Note that the distance to road, soil suitability expressed as restrictions, slope, including the surrounding land use distribution, are the parcel's attributes for time  $t$  (i.e., 1986). The target attribute is the land use type at time  $t+1$  (i.e., 1995).

**Table 3.1 An Example of Input Training Data for Decision Tree**

Case	Current Land Use (1986)	Pct Agr	Pct Urban	Pct Forest	Distance to road	Soil res.	slope	Parcel Area	Wetland Area in parcel	Target land use type (1995)
1	Agriculture	24	50	26	5	1	0.5	67655	0	Urban
2	Forest	20	31	49	10	3	1	7845	76	Forest
3	Forest	62	10	28	1	5	1	9087	567	Agriculture
4	Forest	10	85	5	2	6	1	3639	0	Urban
5	Forest	13	17	70	2	5	3	3453	343	Forest
.										
.										
.										
.										

N

The decision-tree algorithm used J48, a WEKA implementation of the latest public release (Version 8) of C4.5, a standard decision tree algorithm that is widely used for practical machine learning. J48 was developed by the Machine Learning Group of the University of Waikato, New Zealand. WEKA which contains the J48 algorithm, is a publicly available machine-learning package consisting of a collection of algorithms for data-mining tasks. Written in Java, the software includes a uniform interface to a number of standard machine-learning techniques. A chief function of J48 is classification that can be applied to a dataset specified by the user. The dataset that J48 handles is essentially a description of a list of data instances having a set of attributes. The output of this operation is a textual output specifying the classification decision tree. J48 also includes a classify/graph function. This is similar to the classify function, but with added graphical representation of the decision tree created by J48. The J48 algorithm operates by recursively splitting training data input based on feature values or attributes to produce a tree that preferably generates just one branch. The first attribute to be chosen is designated as the root of the tree. The examples are then split among branches based on their attribute values. If the values are continuous, then each branch takes a certain range of values. A new attribute feature is then chosen, and the process is repeated for the remaining examples. The process stops when the classification of a branch is pure (i.e., it contains only examples in a certain class). As to what attribute to use for a given split, the choice is based on the attribute having the largest value for information gain as discussed in Chapter 2.

The J48 decision tree software learns from these input training data as presented in Table 3.1 and builds a decision tree as shown in Figure 3.5.

```

J48 pruned tree
-----
CURRLANDTYPE = URBAN
| GISACRES <= 2.32482
| | PCTBARN <= 32.366
| | | PCTURB <= 50.782
| | | | PCTFOR <= 32.492
| | | | | WETLANDAREA <= 1333.9717
| | | | | | SLOPE <= 8.2215
| | | | | | | PCTFOR <= 15.789: URBAN (809.0/37.0)
| | | | | | | | PCTFOR > 15.789
| | | | | | | | | SOILTYPE = 0: URBAN (50.0/4.0)
| | | | | | | | | SOILTYPE = 1: URBAN (326.0/15.0)
| | | | | | | | | SOILTYPE = 2: URBAN (73.0/3.0)
| | | | | | | | | SOILTYPE = 3: URBAN (75.0/4.0)
| | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | SOILTYPE = 5
| | | | | | | | | | PCTWATR <= 36.025: URBAN (76.0/4.0)
| | | | | | | | | | PCTWATR > 36.025: FOREST (2.0)
| | | | | | | | | | SOILTYPE = 6: URBAN (115.0/10.0)
| | | | | | | | | | SLOPE > 8.2215
| | | | | | | | | | | GISACRES <= 2.13293: URBAN (87.0/10.0)
| | | | | | | | | | | GISACRES > 2.13293: FOREST (4.0)
| | | | | | | | | | | WETLANDAREA > 1333.9717: URBAN (249.0/31.0)
| | | | | | | | | | | PCTFOR > 32.492
| | | | | | | | | | | | WETLANDAREA <= 10256.44142
| | | | | | | | | | | | | PCTURB <= 13.21
| | | | | | | | | | | | | | PCTBARN <= 3.557
| | | | | | | | | | | | | | | PCTWETL <= 0.08: URBAN (31.0/16.0)
| | | | | | | | | | | | | | | PCTWETL > 0.08
| | | | | | | | | | | | | | | | SOILTYPE = 0: URBAN (0.0)
| | | | | | | | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | | | | | | | | PCTWETL <= 13.579: URBAN (14.0)
| | | | | | | | | | | | | | | | | PCTWETL > 13.579: FOREST (3.0/1.0)
| | | | | | | | | | | | | | | | | | SOILTYPE = 2: URBAN (3.0)
| | | | | | | | | | | | | | | | | | SOILTYPE = 3: URBAN (6.0/1.0)
| | | | | | | | | | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | | | | | | | | | SOILTYPE = 5: URBAN (2.0)
| | | | | | | | | | | | | | | | | | SOILTYPE = 6: URBAN (14.0/3.0)
| | | | | | | | | | | | | | | | | | PCTBARN > 3.557: URBAN (2.0/1.0)
| | | | | | | | | | | | | | | | | | PCTURB > 13.21
| | | | | | | | | | | | | | | | | | | GISACRES <= 1.77651
| | | | | | | | | | | | | | | | | | | | DIST2STREAM <= 73.93447: URBAN (84.0/20.0)
| | | | | | | | | | | | | | | | | | | | DIST2STREAM > 73.93447
| | | | | | | | | | | | | | | | | | | | | SOILTYPE = 0: URBAN (93.0/11.0)
| | | | | | | | | | | | | | | | | | | | | SOILTYPE = 1: URBAN (425.0/22.0)
| | | | | | | | | | | | | | | | | | | | | SOILTYPE = 2: URBAN (131.0/6.0)
| | | | | | | | | | | | | | | | | | | | | SOILTYPE = 3: URBAN (293.0/29.0)
| | | | | | | | | | | | | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | | | | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | | | | | | | | | | | | | PCTWATR <= 31.044: URBAN (118.0/6.0)
| | | | | | | | | | | | | | | | | | | | | | PCTWATR > 31.044: FOREST (2.0)

```

Figure 3.5 A sample J48 decision tree output.

To explain how the decision tree structure is represented above, consider the following portion of the output:

```

CURRLANDTYPE = URBAN
| GISACRES <= 2.32482
| | PCTBARN <= 32.366
| | | PCTURB <= 50.782
| | | | PCTFOR <= 32.492
| | | | | WETLANDAREA <= 1333.9717
| | | | | | SLOPE <= 8.2215
| | | | | | | PCTFOR <= 15.789: URBAN (809.0/37.0)
| | | | | | | PCTFOR > 15.789
| | | | | | | | SOILTYPE = 0: URBAN (50.0/4.0)
| | | | | | | | SOILTYPE = 1: URBAN (326.0/15.0)
| | | | | | | | SOILTYPE = 2: URBAN (73.0/3.0)
| | | | | | | | SOILTYPE = 3: URBAN (75.0/4.0)
| | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | SOILTYPE = 5
| | | | | | | | | PCTWATR <= 36.025: URBAN (76.0/4.0)
| | | | | | | | | PCTWATR > 36.025: FOREST (2.0)
| | | | | | | | | SOILTYPE = 6: URBAN (115.0/10.0)
| | | | | | | | | SLOPE > 8.2215
| | | | | | | | | GISACRES <= 2.13293: URBAN (87.0/10.0)
| | | | | | | | | GISACRES > 2.13293: FOREST (4.0)
| | | | | | | | | WETLANDAREA > 1333.9717: URBAN (249.0/31.0)

```

Each line or row represents a node in the tree. The lines that follow and start with a “|”, are child nodes of the first line. A node (or line) with one or more “|” characters before the rule is a child node of the node having one “|” less than said node. For example, nodes 4 and 5 are child nodes of node 3 (or line 3).

The next part of the line is a decision rule that evaluates the expression. If the expression is true for a given instance, either classify it if the rule is followed by a colon and in which case a class designation is given, otherwise continue to the next node in the tree (i.e., the first child node immediately following the node that was just evaluated). If the expression is instead false, continue to the “sister” node (the node having the same number of “|” characters before it in the same parent node) of the node that was just evaluated. Consider node 7 as an example. If upon evaluation of this node, the expression

is found to be true, proceed to its child node (node 8). In node 8, a classification is made – urban. If in the case when node 7 is false, proceed to its sister node --- node 19.

Note that either one or two numbers in parenthesis follow a node that generates a classification. In the node below,

| Pctfor86 > 54.7237 : FOREST (2.0),

the first number in the parenthesis, in this case 2, indicates the number of correctly classified instances in the training set made by this node. The second number represents the number of instances incorrectly classified by the node. A missing second number usually is taken as a zero.

The next step is to convert the newly built tree (Figure 3.5) into a rule-set format for subsequent use in cellular automata. Each path from the root of the tree down to the terminal node where a classification is shown can be translated into a transition rule. As an example, a portion of the decision tree depicted in Figure 3.5 will yield the following transition rules:

**Rule 1:**

(land type = urban) and (Parcel area is less than or equal to 2.32 acres) and (percent barren in neighborhood is less than or equal to 32.37) and (percent urban in neighborhood is less than or equal to 50.78) and (percent forest in neighborhood is less than or equal to 32.49) and (wetland portion of parcel is less than or equal to 1333.97 sq. ft.) and (slope is less than or equal to 8.22) and (percent forest is less than or equal to 15.79) then predict as URBAN.

**Rule 2:**

(land type = urban) and (Parcel area is less than or equal to 2.32 acres) and (percent barren in neighborhood is less than or equal to 32.37) and (percent urban in neighborhood is less than or equal to 50.78) and (percent forest in neighborhood is less than or equal to 32.49) and (wetland portion of parcel is greater than 1333.97 sq. ft.) then predict as URBAN.

**Rule 3:**

(land type = urban) and (Parcel area is less than or equal to 2.32 acres) and (percent barren in neighborhood is less than or equal to 32.37) and (percent urban in neighborhood is less than or equal to 50.78) and (percent forest in neighborhood is less than or equal to 32.49) and (wetland portion of parcel is less than or equal to 1333.97 sq. ft.) and (slope is less than or equal to 8.22) and (percent forest is greater than 15.79) and (soil is type no. 5) and (percent water in neighborhood is greater than 36.03) then predict as FOREST.

**Figure 3.6** Rule Set from J48 Decision Tree.

It can be observed that according to these rules, current land use type, parcel area, and percent urban in the neighboring parcels are the significant attributes needed to make a prediction of land use type.

### **3.3.2 Cellular Automata Module: Predicting Future Land Use/Cover Change**

In a subsequent step, the transition rule derived using the decision tree algorithm J48 is used by a cellular automata model to evaluate how parcels are converted from their current land uses to their future land uses. Agent Analyst/RePast is the computer software chosen to implement the CA model. Spelled out as **Recursive Porous Agent**

Simulation Toolkit, RePast is developed by a group of researchers at the University of Chicago and Argonne National Laboratory (Sallach, et al., 2006). The RePast system, including the source code, is available directly from the web (<http://repast.sourceforge.net>). This research uses Agent Analyst, an extension of ArcGIS that integrates and extends the functionalities of RePast modeling and simulation environment with GIS. Through Agent Analyst, modelers can create, edit, and run RePast models from within the ArcGIS 9 geo-processing framework. This graphical user interface allows the modeler to create agents<sup>1</sup>, schedule simulations, establish mappings to ArcGIS layers, and specify the behavior and interactions of the agents. Aside from having the power and flexibility of ArcGIS, Agent Analyst/RePast has two outstanding features relevant to this study. First, the model has provisions to allow the modification of agent properties, agent behavioral equations, and model properties during run time. Second, it has libraries for genetic algorithms and, neural networks, including the ability to handle irregular grids or vector data as a model component.

Utilization of the transition rules in RePast is implemented as the *step* function of the Agent Analyst model. First, RePast reads from an array list of all spatial data represented in the transition rule and stores the attributes into a *featureList*. Subsequently, an instance of LTCMain is created and assigned to the variable named *ltc*. The *classify* method of the *ltc* variable is then invoked with *featureList* as an input parameter. In this manner, the program feeds into the decision tree the attributes placed inside *featureList*. After processing, the decision tree will then make a prediction (and thus return a value)

---

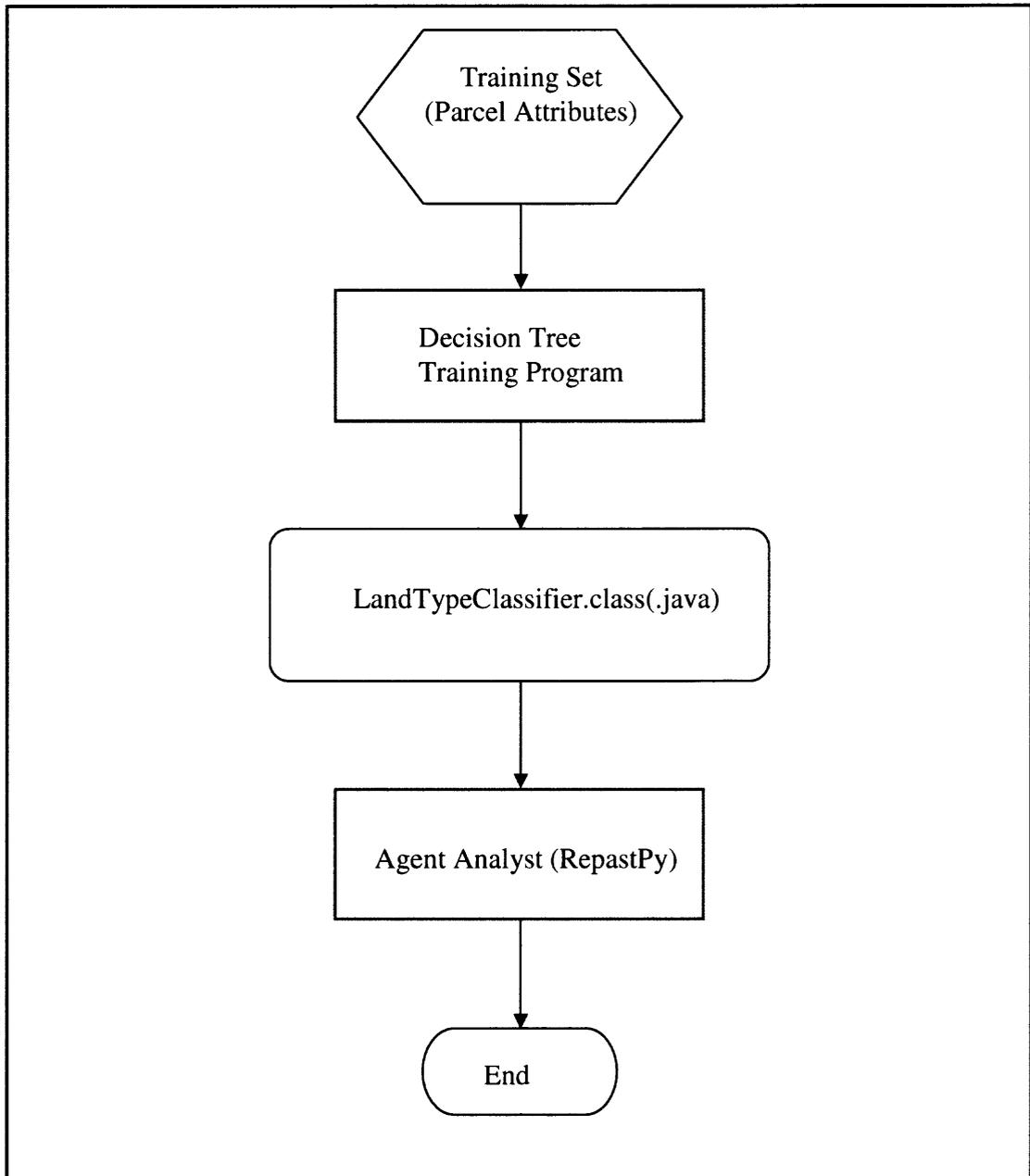
<sup>1</sup> An individual, household or institution that takes specific actions according to its own decision rules which drive land use/cover change (McConnel, 2001).

on the next land usage based on the attribute values stored in *featureList*. The value returned by the method is then stored as the new land use type.

The transition rules embedded in the CA model can also be modified to include government regulatory policies on land use changes. Examples of these policies include regulations or limitations on the conversion of agricultural land to developed land, the steering of urban developments to sites where the soils are considered low value for agriculture, or the designation of zoning laws. Following the calibration of the CA land use change model, the parameter estimates are used to simulate future growth patterns arising from the implementation of planned land use policies. Simulating the impact of policy implementation on land use changes requires additional transition rules as input to the cellular sub-model. Several land use related policies from the New Jersey State Planning Commission's State Development and Redevelopment Plan of 2001 can be translated into transition rules. Examples would be a down-zoning policy converted to a transition rule that disallows conversion to urban land for parcels less than 1.3 hectares. The preservation of farmland and open space can be represented by specifying transition rules designating these areas as non-developable.

### **3.4 Model Integration**

The decision tree and cellular automata modules are implemented using J48 WEKA version and Agent Analyst/Repast, respectively. Figure 3.7 depicts the steps for integrating these modules and the details of each step are also described below in the following sub-sections.



**Figure 3.7** Integration steps.

### 3.4.1 The Decision Tree Algorithm

The original (Java) program created is a (C4.5) decision tree learner/trainer based on the implementation in the WEKA software. The trainer program uses the C4.5 decision tree

implementation that is provided by the WEKA library. What a trainer program does is take a properly formatted .csv file as a training set and build a decision tree via the C4.5 algorithm. The output of the program is a .java file (named as LandTypeClassifier.java in this study) that was then compiled by the trainer program to produce the actual program (LandTypeClassifier.class). This program is the decision tree based on the training set implemented as a Java program that will be used later in the Agent Analyst model.

### 3.4.2 Configuring Agent Analyst Prior to Program Execution

Before describing the steps in calibrating the Agent Analyst model, it is necessary to define the terms used in describing the procedures.

- *Java class file.* A Java class file is a file produced when a Java source code is compiled. It consists of bytecodes representing the program translated into a language that the Java virtual machine can understand and thus use to execute a Java-based program.
- *JAR file.* JAR stands for Java Archive. It is a special type of compressed file (similar to .zip or .rar) that consists of compiled Java class files that can be called or used by other Java programs.
- *Classpath of a Java program.* To put it simply, a classpath of a Java program is the absolute or relative path that the Java program will search for other Java class files and libraries (JAR files) that it will use or call when it executes the program.

Once the Java program representing the decision tree has been created, it is ready for use in the Agent Analyst Model. The Java program is compiled into a .class file, LandTypeClassifier.class file that is then placed into a JAR file (named as LTP.jar in this study). The Jar file is subsequently used by the Agent Analyst model. Included in the JAR file is another Java class file named LTCMain.class that represents an intermediary or interface program to handle the processing of the input values into the decision tree (i.e. LandTypeClassifier.class). This is the class file that is used later in the input of

attributes for classification in the Agent Analyst model. Once the JAR file has been created, it is then placed inside the library subfolder (lib) within the Agent Analyst folder.

The next step is to modify the classpath of the Agent Analyst so that it will then be able to recognize the JAR file that has been created (and ultimately the LandTypeClassifier.class) and use it in the simulations. To do this, the contents of two files inside the Agent Analyst folder are modified, namely the agent\_analyst.bat and agent\_analyst\_run.bat.

Below is how the agent\_analyst.bat file looks when opened using a text editor:

```
rem @echo off
set LIB=.lib
%AGENT_ANALYST_VOL%
cd "%AGENT_ANALYST%"
.jre\bin\java-Xmx512M-classpath./repastpy.jar;
%LIB%/repastj.jar;%LIB%/arcobjects.jar;%LIB%/jintegra.jar;%LIB%/openmap.jar;%LIB%/geotools_
repast.jar;./lib/velocity-dep-1.4.jar;%LIB%\jakarta-
poi.jar;%LIB%\mediaplayer.jar;%LIB%\jmf.jar;%LIB%\plot.jar;%LIB%\asm.jar;%LIB%\junit.jar;%LI
B%\colt.jar;%LIB%\trove.jar;%LIB%\jh.jar;%LIB%\commons-logging.jar;%LIB%\commons-
collections.jar;%LIB%\multiplayer.jar;%LIB%\ProActive.jar;%LIB%\joone.jar;%LIB%\beanowl.jar;
%LIB%\violinstrings-1.0.2.jar;%LIB%\jode-1.1.2-pre1.jar uchiago.src.simbuilder.Main -ArcGIS %1
%2 %3 %4 %5 %6 %7 %8
pause
```

Below is how agent\_analyst.bat file looks after modification (modified sections shown in boldface and underlined)

```
rem @echo off
set LIB=.lib
%AGENT_ANALYST_VOL%
cd "%AGENT_ANALYST%"
.jre\bin\java-Xmx512M-classpath/repastpy.jar;%LIB%/repastj.jar;
%LIB%/LTP.jar;%LIB%/arcobjects.jar;%LIB%/jintegra.jar;%LIB%/openmap.jar;%LIB%/geotools_
repast.jar;./lib/velocity-dep-1.4.jar;%LIB%\jakarta-
poi.jar;%LIB%\mediaplayer.jar;%LIB%\jmf.jar;%LIB%\plot.jar;%LIB%\asm.jar;%LIB%\junit.jar;%LI
B%\colt.jar;%LIB%\trove.jar;%LIB%\jh.jar;%LIB%\commons-logging.jar;%LIB%\commons-
collections.jar;%LIB%\multiplayer.jar;%LIB%\ProActive.jar;%LIB%\joone.jar;%LIB%\beanbowl.jar;
%LIB%\violinstrings-1.0.2.jar;%LIB%\jode-1.1.2-pre1.jar uchicago.src.simbuilder.Main -ArcGIS %1
%2 %3 %4 %5 %6 %7 %8
pause
```

Similarly, the modified agent\_analyst\_run.bat will appear as:

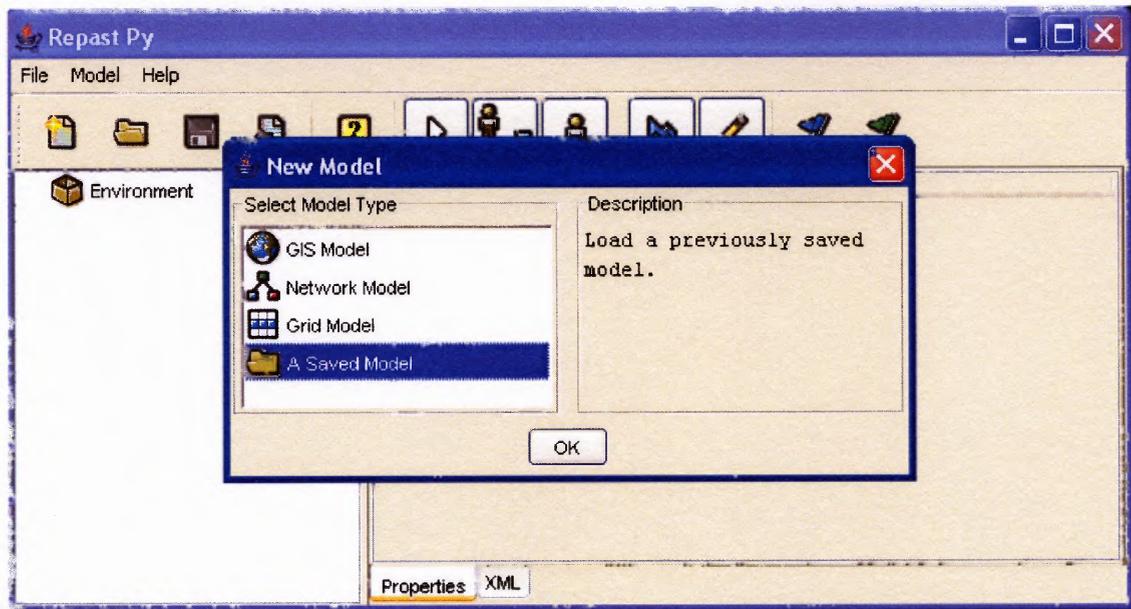
```
rem @echo off
set LIB=.lib
%AGENT_ANALYST_VOL%
cd "%AGENT_ANALYST%"
start.jre\bin\javaw-Xmx512M-classpath
./repastpy.jar; %LIB%/LTP.jar;%LIB%/repastj.jar;%LIB%/arcobjects.jar;%LIB%/jintegra.jar;%LIB
%/openmap.jar;%LIB%/geotools_repast.jar;./lib/velocity-dep-1.4.jar;%LIB%\jakarta-
poi.jar;%LIB%\mediaplayer.jar;%LIB%\jmf.jar;%LIB%\plot.jar;%LIB%\asm.jar;%LIB%\junit.jar;%LI
B%\colt.jar;%LIB%\trove.jar;%LIB%\jh.jar;%LIB%\commons-logging.jar;%LIB%\commons-
collections.jar;%LIB%\multiplayer.jar;%LIB%\ProActive.jar;%LIB%\joone.jar;%LIB%\beanbowl.jar;
%LIB%\violinstrings-1.0.2.jar;%LIB%\jode-1.1.2-pre1.jar uchicago.src.simbuilder.Main -ArcGIS -run
%1 %2 %3 %4 %5 %6 %7 %8
```

Having made the necessary modifications, the decision tree program can now be used by the Agent Analyst model.

### 3.4.3 Agent Analyst Program Execution

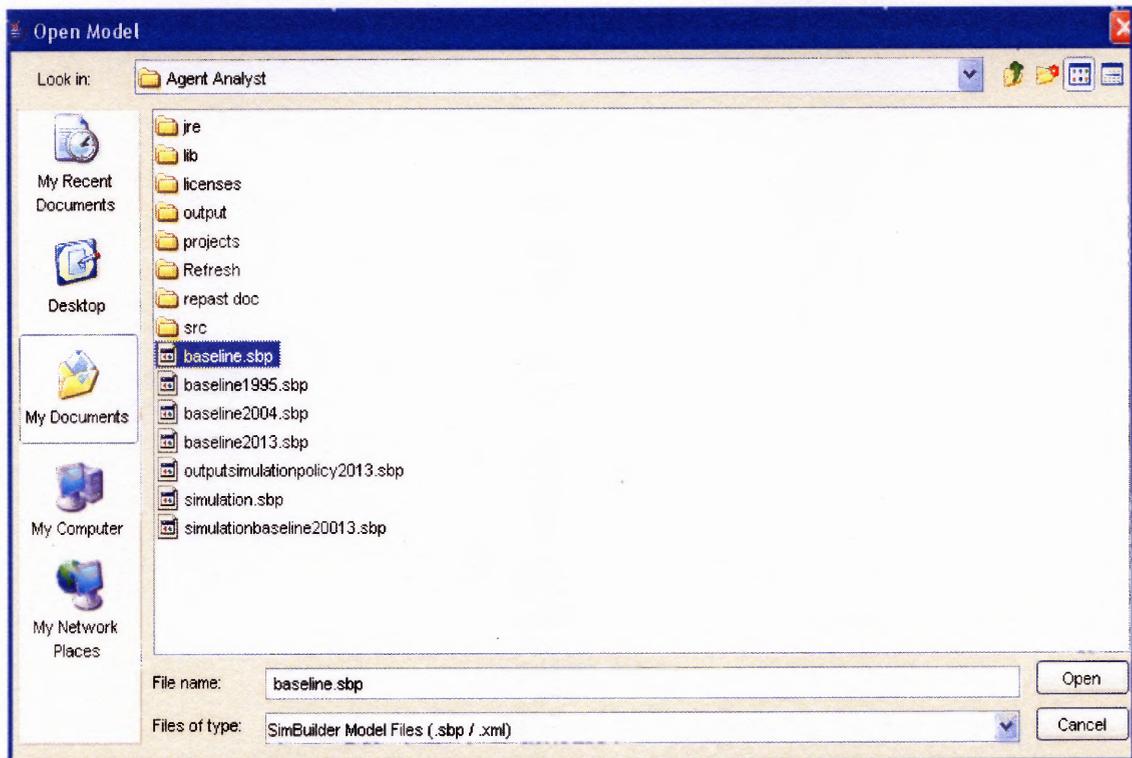
For clarity, program execution is illustrated in a step-wise procedure format below.

- 1) Open Agent Analyst. When prompted to choose a Model Type, select “A Saved Model” then click on “OK.”



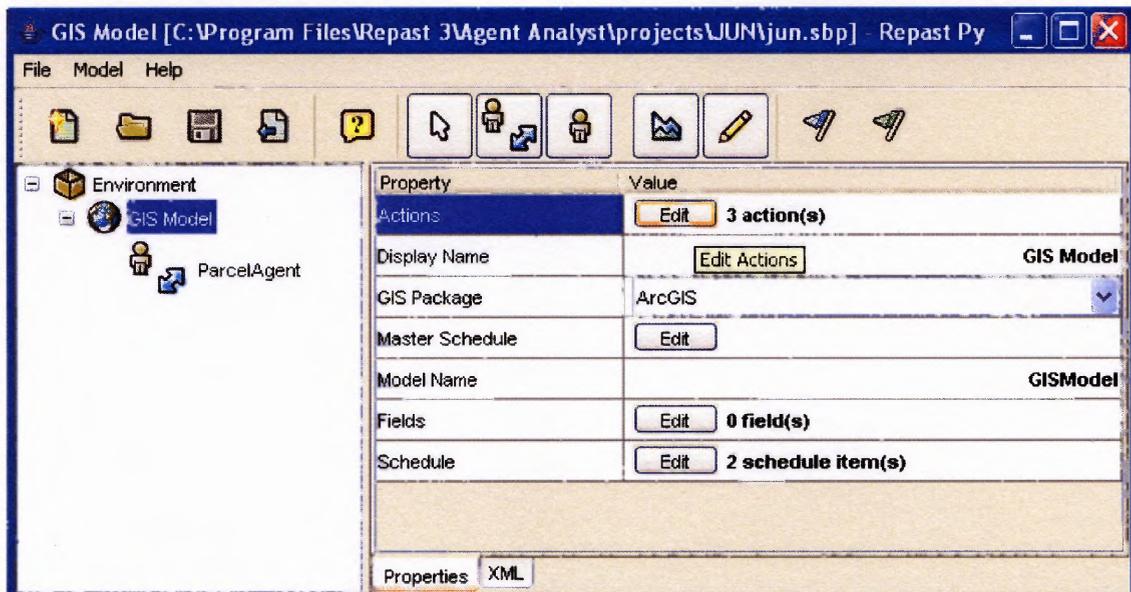
**Figure 3.8** Retrieving the model.

- 2) When prompted, load the agent analyst file (filename.sbp) that was previously copied into new folder created inside the “projects” folder within the {AGENT\_ANALYST\_FOLDER}.



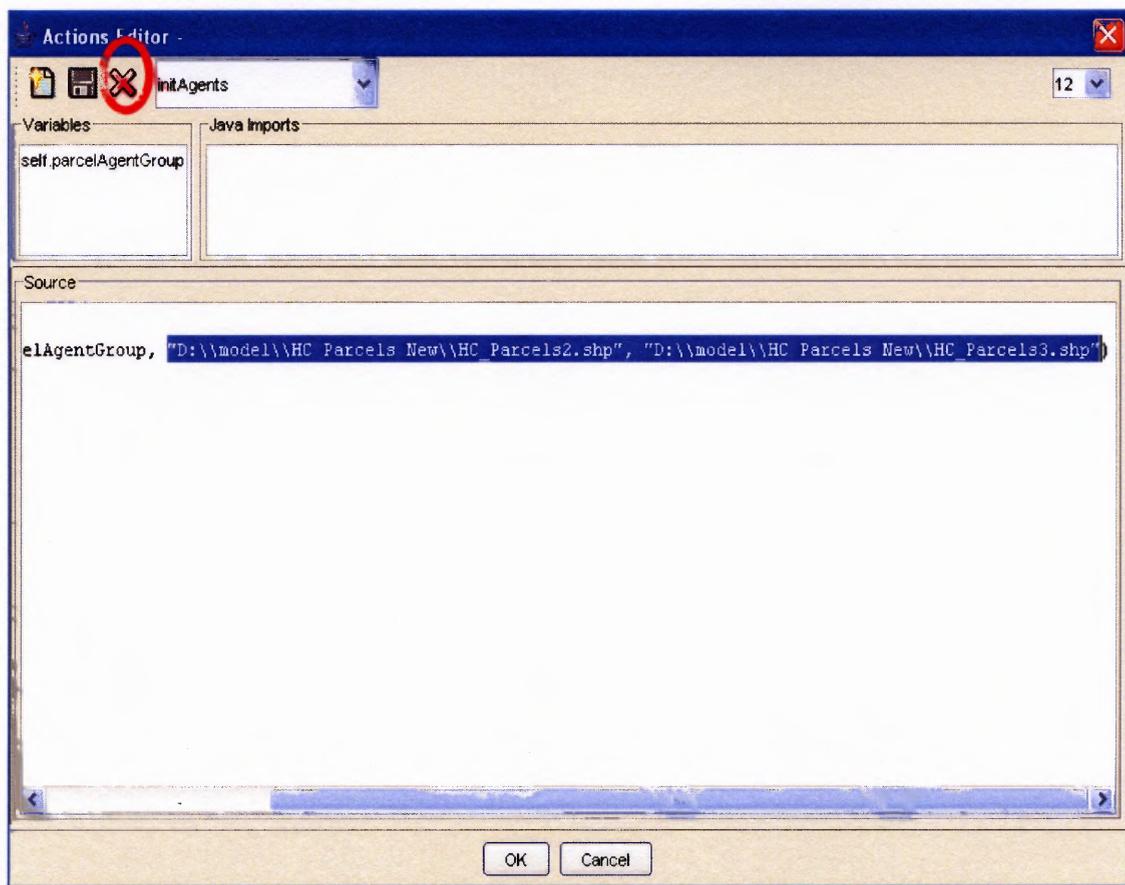
**Figure 3.9** Opening the model dialogue box.

- 3) After the model has been loaded, click on the “GIS Model” icon on the project pane (left portion of the dialogue box, Figure 3.10). Click on the “Edit” button in the “Actions” property on the right menu.



**Figure 3.10** Editing model steps.

- 4) A dialogue box opens (Figure 3.11). In the illustration below, a code for the `initAgents` action is shown. The highlighted part shows the source and destination shapefiles from which the parcel data will be derived and written to respectively. Make the necessary changes to reflect the location of the input shapefiles to be used by the program.

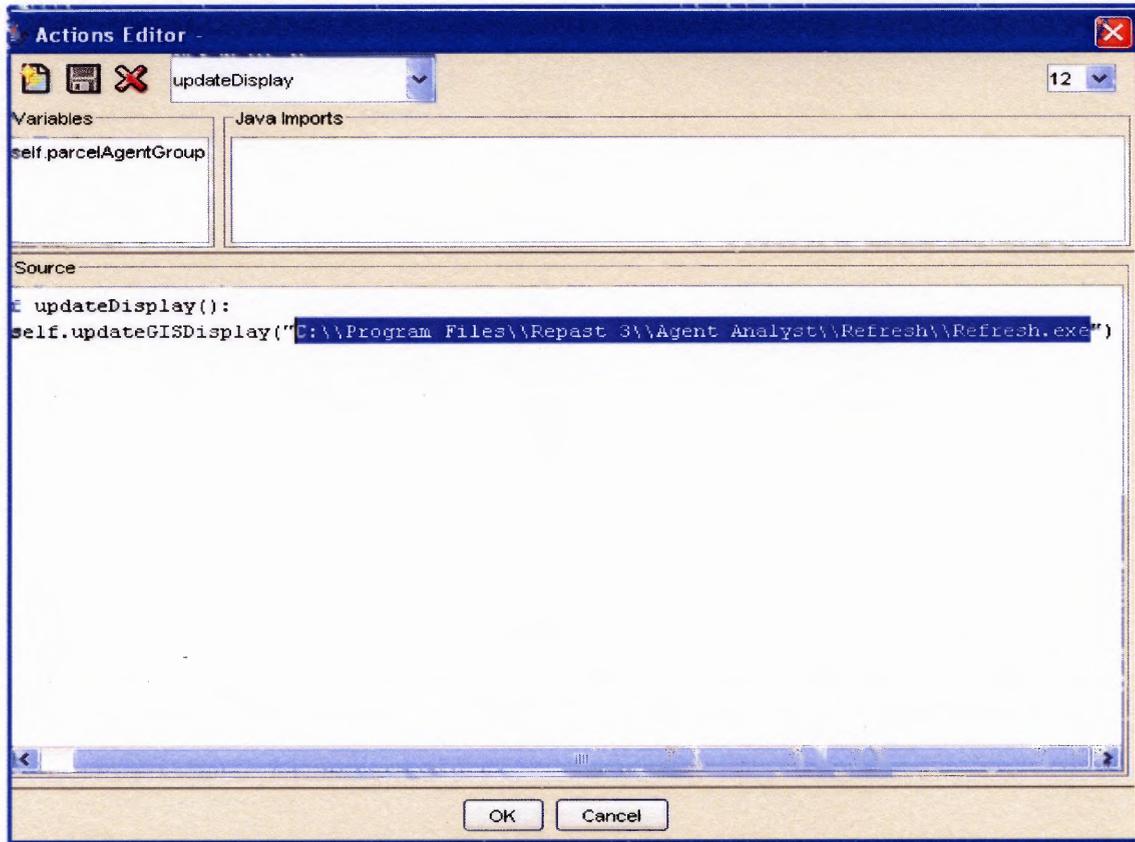


**Figure 3.11** Actions editor dialogue box (a).

- 5) Click on the dropdown box that appears in the upper corner of the new window and choose “updateDisplay.” A code will be shown as presented in Figure 3.12 representing the updateDisplay action. The highlighted part in Figure 3.12 represents the location of the executable file that will be used to update the display in ArcMap. This file (i.e., executable file) is placed inside the “Refresh” folder inside {AGENT\_ANALYST\_FOLDER}. Necessary changes should be made on the “C:\\Program Files\\Repast 3\\Agent Analyst” to reflect the actual folder where the Agent Analyst is located in the user’s computer. For example, if the Agent Analyst software is located in D:\\Agent Analyst, the resulting value that should replace the highlighted part of the image below will

be “D:\\Agent Analyst\\Refresh\\Refresh.exe” (note the double backslash).

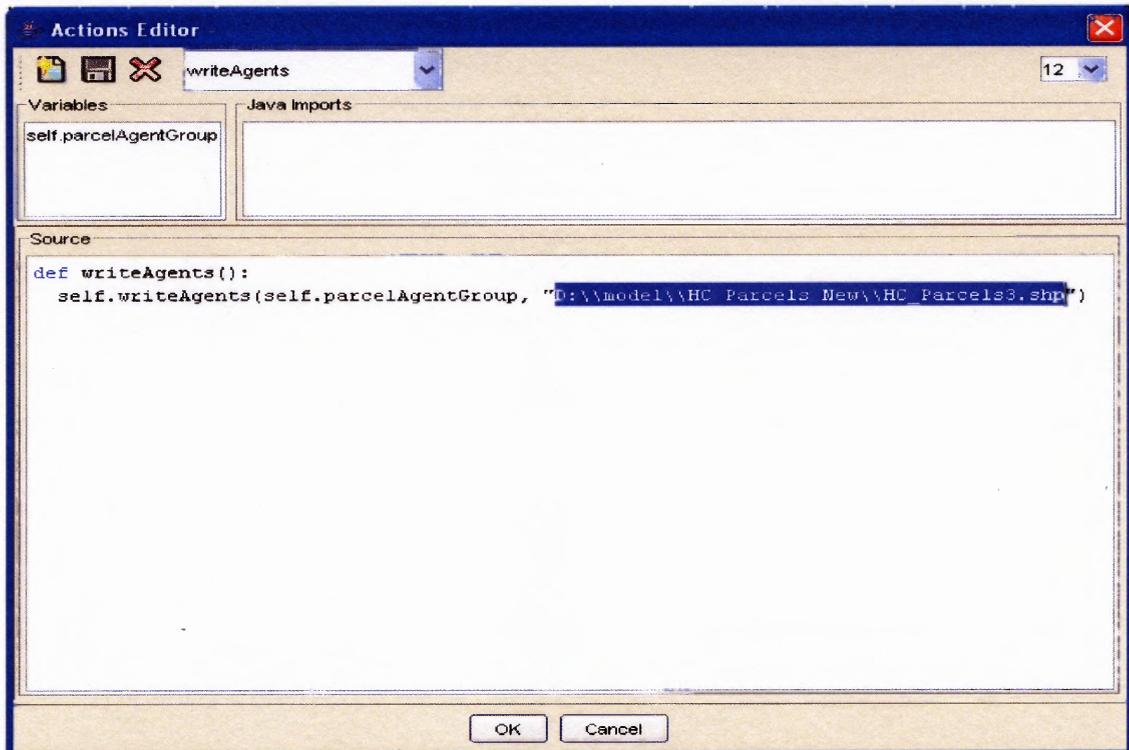
Save the changes by clicking the Disk icon (encircled in red) in Figure 3.11.



**Figure 3.12** Actions editor dialogue box (b).

- 6) The next step is to click on the dropdown box again and choose “writeAgents”. The code in the bottom pane of Figure 3.13 is the code needed to write to the shapefile the updates that will be performed by the simulation later on. As with the previous item, make the necessary changes accordingly and save the changes. Go back to the Agent Analyst main menu by clicking “OK.”
- 7) This step invokes the decision tree output to be used as the transition function in Agent Analyst. Click on the parcel agent icon on the left pane of the

dialogue box. Click the “Edit” button in the “Actions” property on the right menu. Enter the appropriate source code (explained in section 3.4.4 below).



**Figure 3.13** Actions editor dialogue box (c).

- 8) In the Agent Analyst main menu, compile the program by clicking the blue flag on the upper right-hand side of the window. The “Compile” button translates the changes made earlier into a computer readable and executable program.

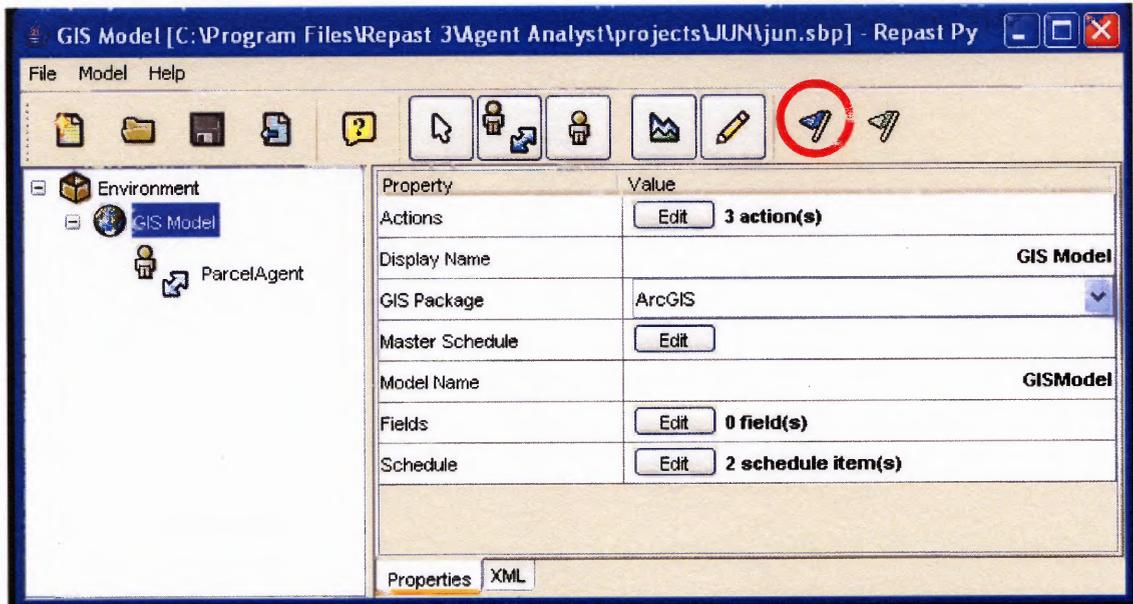


Figure 3.14 Compiling the model.

- 9) Assuming no errors or exceptions were reported, open ArcMap and display the shapefile.
- 10) Once the map has been displayed, click on the green flag of the Agent Analyst Window to proceed with the simulation.

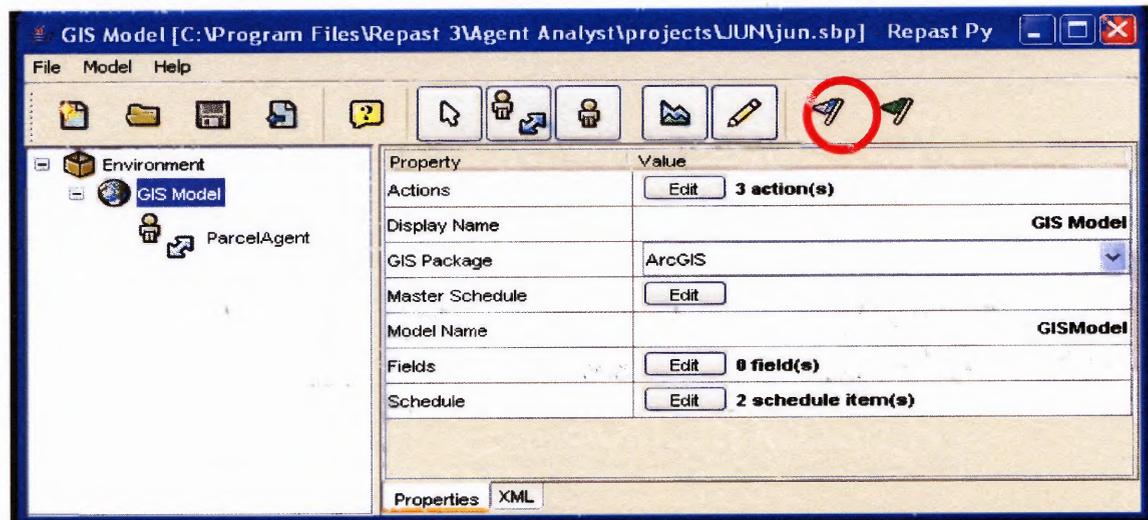


Figure 3.15 Model run.

- 11) Two windows will appear that represent the simulation console. Click on the play button to commence simulation.



**Figure 3.16** Commence simulation dialogue box.

### 3.4.4 Using Decision Trees in Agent Analyst

In step 7 of section 3.4.3 above, the output from the decision tree output is used as the transition function of the cellular automata model. A transition rule is implemented as the step function of the Agent Analyst model. The following script for the step function in the Agent Analyst model is used to define the transition function referred to in step 7 above.

```
def step():
    featureList = ArrayList()
    featureList.add(self.SOIL_SUITE)
    featureList.add(Double(self.GIS_ACRES))
    featureList.add(Double(self.PCTAGRI))
    featureList.add(Double(self.PCTBARN))
    featureList.add(Double(self.PCTFOR))
    featureList.add(Double(self.PCTURB))
    featureList.add(Double(self.PCTWTR))
    featureList.add(Double(self.PCTWET))
    featureList.add(Double(self.SLOPE_MEAN))
    featureList.add(Double(self.URBAN_NEAR))
    featureList.add(Double(self.HWDIST))
    featureList.add(Double(self.STREAM_NEA))
    featureList.add(self.LANDTYPE)
    ltc = LTCMain()
    self.LANDTYPE = ltc.classify(featureList)
```

**Figure 3.17** Step definition in Agent Analyst for transition rule.

The featureList variable here is implemented as an ArrayList (a Java object representing an array with dynamic memory allocation). This will serve as the storage for

the data that are required for feeding into the DT program later. The Agent Analyst then stores the necessary attributes one by one (using the add function) into featureList. Subsequently, an instance of LTCMain is created and is assigned to the variable named ltc. The classify method of the ltc variable is then invoked, with featureList as an input parameter. What the program will do (after the processing of the input as discussed in the earlier part) is to feed into the decision tree the attributes placed inside featureList. After processing, the decision tree will then make a prediction (and thus return a value) on the next land usage based on the attribute values stored in featureList. The value returned by the method is then stored in the LANDTYPE variable of the Agent Analyst model.

Besides describing the usage of the decision tree in Agent Analyst, there is a need to explain the significance of using a separate Java class LTCMain file instead of directly using LandTypeClassifier as the transition function in Agent Analyst. The reason behind this approach is that in the creation of LandTypeClassifier it is not possible to directly intervene or revise the main code of the Java program since this is done by the WEKA library. Hence, any modifications on the behavior of the program should be done outside LandTypeClassifier, and in this case, through a separate program (LTCMain) which calls LandTypeClassifier. Two modifications are needed. One is the processing of the attribute values of the input instance to make sure that the values are correctly formatted and thus can be properly processed by the decision tree. The second is the addition of rules that may not have been derived via the training of the decision tree (e.g. down zoning: if a parcel area is less than 1.3 hectares, then next land use cannot be “urban”).

### 3.5 Model Validation

Validation is a standard procedure for evaluating a model's reliability in achieving the objectives for which it was developed. In the case of simulation models, this process essentially involves an assessment of the conformity of the model's output to reality.

The literature survey in Chapter 2 listed various ways of assessing model performance. This discussion further argued that there is still no standard measure for accuracy that is universally accepted within the land use research community. Some have argued that no single metric should be adopted since different models have been developed for different objectives. Thus a single metric that provides information on a certain aspect can be appropriate for a model from a certain perspective, but for another model developed for a different objective, the same metric may not be suitable at all. To foster a consensus to make models more useful to users, Foody (2002) suggested using more than one model metric by providing multiple accuracy descriptions. Following this indication, the metrics listed below will be used:

- a) Three variants of Kappa (standard Kappa, Kappa location, and Kappa quantity) to indicate location and quantity predicting capability,
- b) Errors of omission and commission sometimes referred to as User's and Producer's accuracy respectively, and
- c) Overall accuracy computed as the ratio of correct classification to the total number of classifications.

The formula for Kappa and its variants are given in Table 3.2. The significance of the terms in the table and the procedures for computing errors of omission and commission were explained in Section 2.8.

**Table 3.2. Calculation of Kappa and its Variants<sup>2</sup>**

<b>Kappa</b>	<b>Formula</b>
Standard Kappa, <b>K<sub>std</sub></b>	$\frac{P_o - MQNL}{1 - MQNL}$
Kappa for location specification, <b>K<sub>loc</sub></b>	$\frac{P_o - MQNL}{MQPL - MQNL}$
Kappa for quantity specification, <b>K<sub>quantity</sub></b>	$\frac{P_o - NQML}{PQML - NQML}$
Proportion of correctly classified parcels, <b>P<sub>o</sub></b>	Sum of diagonals in confusion matrix
No ability to specify quantity, perfect ability to specify location, <b>NQPL</b>	$\sum_{j=1}^j \min\left(\left(\frac{1}{j}\right), R_j\right)$
No ability to specify quantity, medium ability to specify location, <b>NQML</b>	$\left(\frac{1}{j}\right) + K_{loc} \left( NQPL - \left(\frac{1}{j}\right) \right)$
Medium ability to specify quantity, no ability to specify location, <b>MQNL</b>	$\sum_{j=1}^j (S_j R_j)$
No ability to specify quantity, no ability to specify location, <b>NQNL</b>	$\frac{1}{j}$
Medium ability to specify quantity, perfect ability to specify location, <b>MQPL</b>	$\sum_{j=1}^j \min(S_j, R_j)$
Perfect ability to specify quantity, no ability to specify location, <b>PQNL</b>	$\sum_{j=1}^j (R_j^2)$
Perfect ability to specify quantity, medium ability to specify location, <b>PQML</b>	$PQNL + K_{loc} (1 - PQNL)$

<sup>2</sup> Adopted from Pontius (2000). S and R in the formulas denote quantities in the simulation and reality rows and columns respectively.

### 3.6 Sensitivity Analysis

Developing the correct transition rule is a key component in CA modeling (Torrens, 2000, Childress et al., 1996). Directly affecting the outcome of transition rules are the neighborhood size and shape. When evaluating a CA model's robustness and reliability, altering the transition rule by varying the neighborhood size and shape is frequently undertaken (Kocabas and Dragicevic, 2007; Dietzel and Clarke, 2004; Chen and Mynett, 2003; Jenerette and Wu, 2001; White, Engelen, and Uljee, 1997). In these studies, the neighborhood consisted of uniform cell sizes represented in a Von Neumann, Moore, or circular neighborhood configurations. Since irregular cell shapes and size cannot conform to von Neumann or Moore neighborhood arrangements, the focus will be on varying neighborhood size to investigate its effect on the conversion of the central parcel and hence model output.

The sensitivity analysis was performed using two neighborhood sizes, namely 237 and 712 feet measured from the parcels' edge as the buffer distance. These distances represent 0.5 and 1.5 times the buffer distance in the baseline scenario. As stated previously, 475 feet is the edge length of an average square parcel in the county.

By varying the neighborhood size, proportions of land use types around each parcel were changed. These values were then used as input to the model and the results were compared with a reference map to assess its deviation or sensitivity to changing input. To measure sensitivity, a Kappa index was used. In the same way as Kappa was used to measure agreement between predicted and a reference maps, Kappa was used to indicate the relative correspondence of the output resulting from varying model inputs to the reference map. On one hand, similar Kappa values between model results from

different input indicate that the model results are likewise similar and consequently not sensitive to such perturbations. On the other hand, varying Kappa values indicates a model that is sensitive to the changed input.

### **3.7 Summary**

The foregoing sections have outlined the approach for developing a land use change model that relies on realistic geographical entities such as the parcel as the unit of analysis. A cellular automata model is proposed that employs transition rules derived from parcel-level data. At the same time, the methodology specified for a procedure for generating a transition rule that is simple, intuitive, and yet free from subjective errors to which other methods are susceptible. A knowledge-discovery technique, namely the decision tree, fulfills these requirements.

## CHAPTER 4

### STUDY AREA PROFILE AND DATA PROCESSING

Hunterdon County is chosen as the study area for this study because of its readily available spatial data for land use change modeling. Hunterdon is one of the few counties in the State of New Jersey with the digitized parcel maps. The spatial data on selected driving factors of land use change can be easily obtained from the County's Division of GIS. This chapter summarizes the land use changes and their driving factors in Hunterdon County, the research study area. It begins with the processing of spatial information by overlaying maps since all driving factors needed to be aggregated at and derived from the parcel. This chapter also briefly explains current trends of land use including anthropogenic activities that shaped landscape patterns in the study area. The remaining sections discuss characteristic driving factors for land use change in Hunterdon County.

#### 4.1 Hunterdon County, New Jersey

Hunterdon County is one of 21 counties of New Jersey. Bounded by Warren County to the northwest, Morris and Somerset Counties to the east, and Mercer County to the south (Fig. 4.1), Hunterdon encompasses 438 square miles of the western portion of the State. It ranks eighth among New Jersey's counties in terms of land area and has 26 municipalities classified into townships, boroughs, and towns. The County is traversed from east to west by the I-78 interstate highway designed to carry traffic between regions of the state and to serve as a corridor between Port Newark/Liberty Airport and points westward.

Accessibility between municipalities and adjoining counties is provided by a network of county and municipal roads that includes Routes 12, 31, 202, and 517.

Hunterdon County is still considered to be a mostly rural and country suburban county. Rural areas defined by a predominance of undeveloped land are mostly found in the southern sections of the County and in Lebanon and Franklin Townships in the north. Country suburban areas characterized by more development than rural areas, predominate in the central part of the county that includes the greater Flemington area (HCPB, 2005c).

Hunterdon County is home to approximately 129,000 people as of 2004 (New Jersey Department of Labor and Workforce Development (NJDLWD), 2006). Decade censuses from 1950 to 2000 show a population growth rate ranging from 13 to 29 percent (US Census Bureau, 2005). According to NJDLWD (2006), the population in Hunterdon County grew by 87 percent between 1970 and 2004 making it the third fastest growing county in New Jersey. By 2020, the population is estimated to be 160,932 (HCPB, 2005a). Hunterdon County also experienced considerable economic growth owing to its proximity to high growth areas in the state with firms like Exxon, Foster Wheeler, and Merck establishing their corporate offices there during the 1980s and 1990s. Growth has slowed down and continued expansion will likely continue in coming decades though at a slower pace (HCPB, 2005b). Given these trends, demand for land development is expected to increase. The following sections explain how these trends have shaped recent development patterns in the county and discuss past and current policies that directly or indirectly relate to land use management.

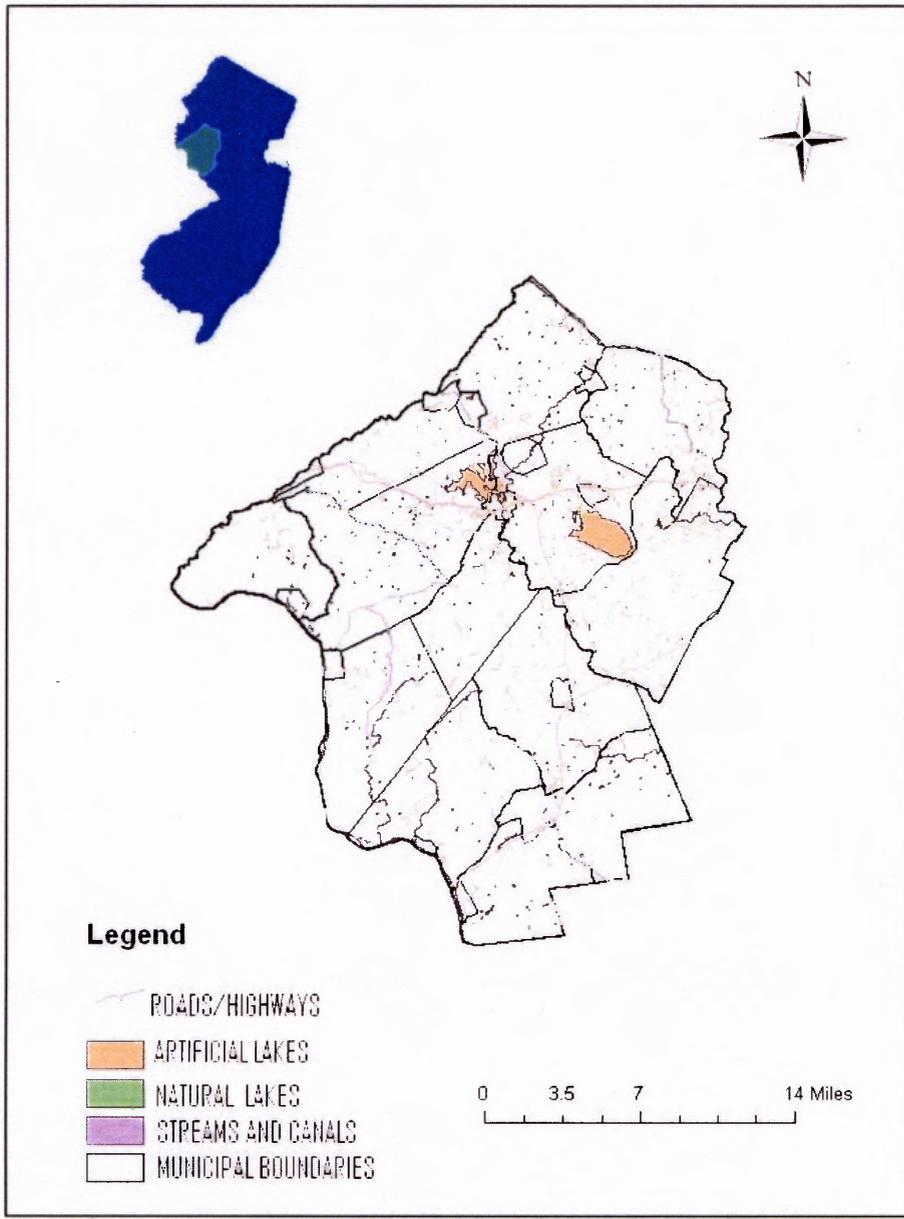


Figure 4.1 Hunterdon County Map

#### **4.1.1 Residential Development**

Municipalities in Hunterdon County experienced high density residential developments during the 1970s and 1980s. Single-family homes on small lots, townhouses, apartments, and condominiums were built in Raritan Township, Union Township, Glen Gardner, Lambertville, Clinton Township, Clinton Town, and Readington Township. This phase of activity was followed by a period of high density “inclusionary developments”, that entailed the construction of housing units with a significant portion dedicated to low- and medium-income bracket families and the rest for market-rate units (HCPB, 2005a).

At present, newly developed residential areas are mostly comprised of single-family houses on large lots located in country suburban and rural areas. Many of these developments are in areas not yet covered by sewer services. Thus, to prevent groundwater contamination, some municipalities have legislated zoning laws that specify the minimum lot size to 1.5 acres or more to accommodate individual septic systems. In order to balance the need for more housing while at the same time preserving open space, some municipalities have allowed the construction of wastewater treatment facilities, thereby making construction on smaller lots possible. Reduced lot sizes in turn have made it possible to leave more large contiguous areas available for preservation as open space or farmland.

#### **4.1.2 Commercial and Industrial Development**

Commercial and industrial development has centered on pharmaceuticals, computer technology, biotechnology, and medical research. These were some of the leading

employers that located their headquarters in New Jersey in the 1990s (HCPB, 2005b). Initially, many of these were based in Somerset and Union Counties located to the east of Hunterdon. The land use effect to Hunterdon County was essentially an increase in residential development as the employees of these firms settled in the County. As I-78 extended westward, industrial development eventually spilled into Hunterdon County. However, recent reports project that these trends are on the decline (Hughes and Seneca, 2006). Whereas New Jersey accounted for 5.2 percent of the nation's total high-technology employment base in 1990, the state's share dropped to 4.0 percent in 2005 (Hughes and Seneca, 2006). This pattern is reflected in every single high technology sector. It was observed that since the 2001 economic recession, high-end investments by US corporations have mostly been directed outside the State. Although US economic statistics showed an increase in the number of non-farm employment of 7.2 percent from 2000-2005 (US Census Bureau, 2005), the increase occurred mostly in below-average-paying job sectors. The Hughes and Seneca (2006) report stated that during the 2000–2005 period, New Jersey lost 117,600 high-paying advanced services and manufacturing jobs. In contrast, the private sector experienced growth in the below-average-paying jobs such as education and health services (+60,800 jobs), leisure and hospitality services (+35,900 jobs), and other services (+16,500 jobs). Total additional employment in these three sectors was 113,200 jobs. The report affirmed that although the high employment afforded by a seemingly prosperous state economy obscures the consequence of high paying job losses, the long-term picture, if the trend continues, is that there will be fewer high technology companies in the State. Such a statewide pattern can have a significant impact on Hunterdon County being nearby to a number of high-tech industries in the

eastern boundary on I-287 and in the south in the Princeton/Lawrenceville area along Route 1 (HCPB, 2007a). As the high-paying jobs in these sectors are gradually replaced by low-paying jobs, per-capita income can be expected to decrease. The projected population increase in the county will be most likely be employment in low paying jobs. One of the impacts of these trends will be on the housing types that will be needed to accommodate new residents from this income class. This income group may stimulate the demand for higher density housing in the county.

#### **4.1.3 Residential Development-related Policies**

State efforts to build the economy shaped land use changes in the county. The impacts are seen in the character and growth of new residential developments in the county.

Hunterdon is one of six counties situated in an extensive growth area known as "the wealth belt". Where 35 percent of new jobs created from 1970 to 1997 (Hughes and Seneca, 1999), the "wealth belt area is characterized by high property values, high population, plenty of jobs and high personal income." High income workers bought large houses with a large area typical in many places in Hunterdon County.

Another factor that brought about the construction of new residential units in the county is the "ratables chase" policy of the State. Essentially, the "ratables chase" policy was designed to encourage the influx of investors to balance the tax structure of the county. Because higher taxes are levied on these "ratables" or investors, they compensate for the relatively low tax rate on residential entities. This encourages local governments to permit the construction of new residential units since funds are available to finance their public service requirements (e.g. sewer, solid waste collection, etc). The

construction boom during the 1980s, when thousands of new homes and commercial buildings were built, was partly attributed to the notion that “ratables” lower property taxes (HCPB, 2007a).

The Fair Housing Act of 1985 creating the Council on Affordable Housing (COAH) also affected the characteristic features of new residential developments in the county. The COAH adopted an “inclusionary zoning policy” requiring developers to set aside a percentage of rental or for-sale units in housing developments for low- and moderate-income residents. The rest of the residential units can then be sold at market value to offset losses. As an incentive, developers are entitled to build more units, or receive other cost benefits, from the government. The goal of this policy is to create, mixed-income communities but often at higher densities than are prevalent in the municipality.

## **4.2 Land Use Policies**

Land management in Hunterdon County is carried out using a number of land use related policies. The Hunterdon County Planning Board is the office tasked with preparing a master plan identifying physical development in the county. The Board oversees the implementation of the County's Traditional Farmland Preservation Program and the 8-year Municipally Approved Farmland Preservation Program. These policies are explained in the following sections.

### **4.2.1 Large-lot Zoning**

One factor affecting land use development patterns in many parts of Hunterdon County is large lot zoning. Relatively inexpensive land and large lot homes have drawn residents

increasingly further from traditional urban centers. Coupled with these developments have been zoning requirements aimed at maintaining the appearance of a rural and agricultural character in several of the county's the communities. Such zoning ordinances require every house to be located on at least two acres and often five or more. In its efforts to preserve its rural landscape character, Hunterdon County municipalities have frequently made use of this zoning technique. Large-lot zoning produces one type of housing: large, single-family homes. However, there are social, economic and environmental implications of this type of zoning. From an agricultural standpoint, a loss of farmland results when more land is consumed per dwelling. This counteracts the County's farmland and open space preservation efforts. Though such drawback can be addressed by directing large lot zoning to areas that are less agriculturally productive. This scheme also leads to dependence on the automobile for transportation needs.

There are several reasons why large lot zoning is a preferential mode of development for both local governments and developers. Since large-lot subdivisions are relatively easier to develop compared to other housing schemes, municipalities that are under pressure to provide accommodations to a fast growing population find this system an appealing alternative. Large lot subdivisions also do not normally require long negotiations with professional planners and attorneys. Developers favor large-lot subdivisions because their attractive big houses on sizeable lots appeal to a wealthy class of prospective new residents.

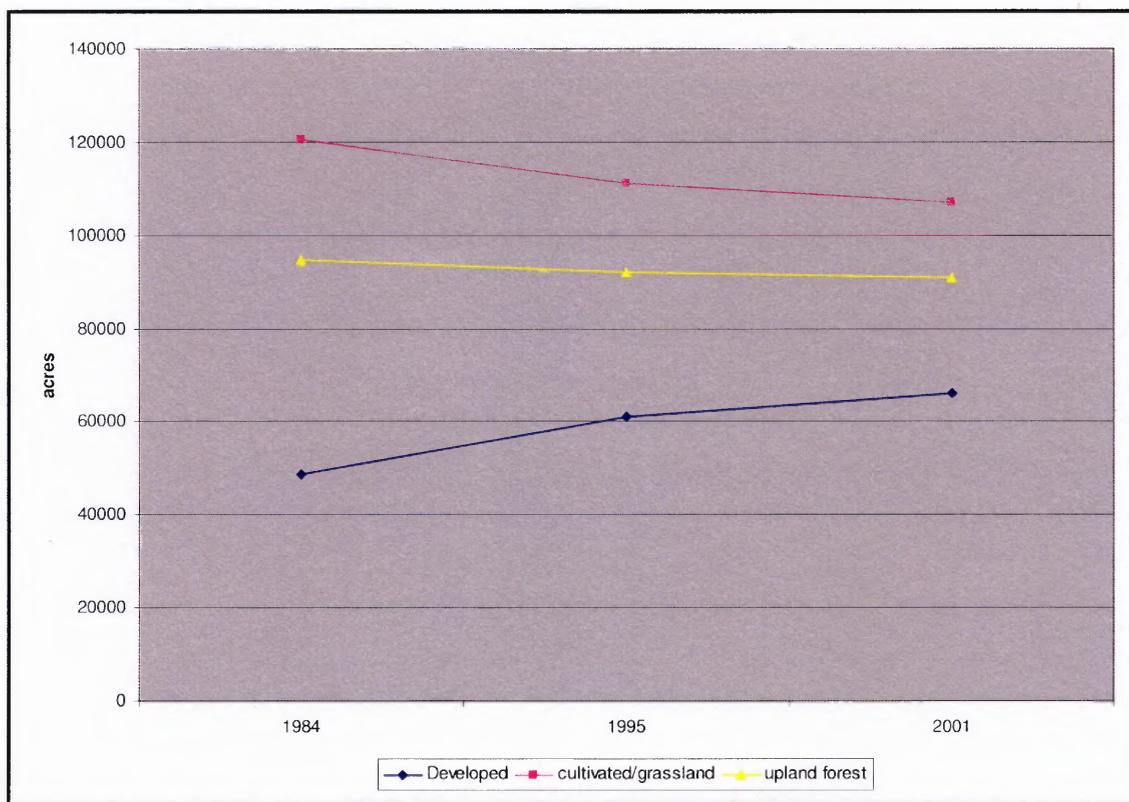
#### **4.2.2 Open Space Zoning**

Similar to "conventional zoning," open space zoning also regulates the overall amount of development in a municipality. The key difference is that open space zoning requires new

construction to be located only on a portion of the parcel. The remaining open space is permanently protected under conservation-easement arrangements between the land owner and the municipality. Open space zoning is also known as “clustering.” Clustering is employed to achieve the state’s goal to preserve as much farmland as possible. Data show that although Hunterdon ranks second statewide in its number of acres of preserved farmland, the acreage is only 18.6 percent of farmland assessed properties in the county and 7.2 percent of the total county area (HCPB, 2007a). While Hunterdon seeks to increase its coverage, lack of sufficient funds to finance farmland preservation programs hampers the realization of this goal.

#### **4.2.3 Farmland Preservation**

Hunterdon County data show that considerable farmland and forest areas are being lost to development (Figure 4.2). From 1984 to 2001, the county’s developed land increased by 17,613 acres. During the same period cultivated/grassland and upland forests decreased by 13,630 and 3,731 acres respectively (Rutgers CRSSA, 2004).



**Figure 4.2** Hunterdon County landscape change 1984-2001.

Source: Rutgers CRSSA Land Cover Classification of Landsat Imagery

Anticipating continued demand for more urban land and preventing further loss of its remaining quality farms and open space to development, Hunterdon County actively participates in State-initiated land preservation programs. This strategy employs a combination of five purchase programs that include purchase of development rights, municipally-approved farmland preservation program, fee simple, direct easement and emergency easement purchase, and planning incentive grant. The New Jersey State Agriculture Development Committee as of December 31, 2007 puts the amount of preserved farmlands in Hunterdon County at 23,299 acres for a total cost of \$181,383,758 (NJDA, 2007). Of this amount, 70 percent represents cost shared by the State.

#### **4.2.3.1 Purchase of Development Rights (PDR)**

Instituted in 1983 through the State Agriculture Retention and Development Act, the PDR (also known as County Easement Purchase Program) has enrolled 128 farms with a total area of 13,678 acres as of August 24, 2007 (HCPB, 2007b). The program encourages landowners to sell the development rights of their farm in exchange for a permanent restriction on its conversion. To be eligible, the farm should be located in an Agricultural Development Area (ADA), should have a minimum of 40 acres, and be predominantly (at least 50 percent) tillable farmland. ADAs are designated by the County Agriculture Development Board as locations where agricultural operations are likely to continue in the future. The agricultural district requirement is meant to encourage the preservation of aggregated parcels as opposed to isolated chunks of farmland within the county. A group of parcels with a minimum total acreage of 250 that has been enrolled, or will be enrolled once the application to the farmland preservation program has been processed, can be classified as an agricultural district. Starting in 2008, however, the PDR program will be replaced by the County Planning Incentive Grant (PIG) program.

#### **4.2.3.2 County and Municipal Planning Incentive Grants (PIG)**

Administered by the State Agricultural Development Committee, the Planning Incentive Grant (PIG) program aims to preserve large tracts of farmland. Essentially, the program is a support fund available to municipal or county governments to purchase multiple parcels situated in a contiguous manner. The financial package features innovative financial strategies, including installment payment programs, option agreements, donations, and bargain sales. Approval of an application by SADC is made on the basis

of contiguity and location, (e.g. located in Agricultural Development Areas). Both county and municipal PIGs are operated in the same manner. However, county PIGs are designed to enable counties to target farm lands of special interest that are not considered for preservation by a municipality. The county PIG program took effect on July 2, 2007. Latest records available show that as of August 24, 2007, 64 farms have received municipal PIGs for a total of 3,348 acres of preserved farmland.

#### **4.2.3.3 Municipally-Approved Farmland Preservation Program (MAFPP)**

Complementing the PDR program of the County at the municipal level is the Municipally Approved Farmland Preservation Program (MAFPP). This program offers incentives for landowners who agree to use their farms exclusively for agriculture for eight years. Incentives include grants for soil and water conservation projects, exemptions from the law on eminent domain, and protection from energy and water restrictions. Unlike the PDR, the MAFPP has no minimum requirement for lot size. This scheme encourages more participation initiating the growth of new agricultural districts. County records show that there are twenty farms enrolled in this program. The total area preserved under this program amounts to 811 acres (Hunterdon County Agricultural Development Board 2006a).

#### **4.2.3.4 Fee Simple**

As of August 24, 2007, eleven farms with a combined area of 1,547 acres were transferred to the county under the fee simple scheme (HCPB, 2007b). As a permanent land acquisition program, fee simple entails the sale of the entire farm to either the county or state. The relevant government body then elects to deed-restrict the farm for

agricultural use only. Although the county or state can eventually opt to sell the land through auction, the deed-restriction clause will continue to apply to the new owners.

#### **4.2.3.4 Direct Easement and Emergency Easement Purchase**

Development easement refers to the land owner's right to develop a piece of land for non-agricultural purposes. Unlike fee simple, purchase of development easement is made directly by the State rather than the county or municipality. The State Agricultural Development Committee purchases development rights from a landowner with a stipulation that development is limited to those uses permitted by the easement restrictions. To date, 45 farms with an aggregate of 3,455 acres in the county have been preserved under this scheme (HCPB, 2007b).

#### **4.2.4 Open Space Development**

In November 1999, the county established the Open Space, Recreation, Farmland and Historic Preservation Trust Fund from real estate tax levies to purchase land for recreation, conservation, general open space, and farmland preservation purposes. Subsequently, an open space tax was approved in a county referendum that took effect in 2000. This tax levies three cents for every \$100 of assessed property value to finance the acquisition of lands for preservation. The first phase expired in 2004, but it was reauthorized for another five years until December 31, 2009 (HCDPR, 2008). To date, \$37 million has been collected and used to preserve 6,258 acres of farmland, 2,939 acres of county parklands and 5,622 acres of municipal and non-profit land acquisitions (HCBP, 2007a). Each municipality was allocated an average of \$2 million to finance land

acquisitions. In 2006, Hunterdon County allocated close to \$2 million to local municipalities for further land acquisition (HCPB, 2007c).

#### **4.2.5 Highlands Water Protection and Planning Act**

The New Jersey Highlands is part of a bigger National Highlands Region that straddles the boundaries New York, Connecticut, Pennsylvania, and New Jersey. New Jersey's portion covers an area of 800,000 acres stretching from Ringwood in the northeast to Philipsburg in the southeast (Phelps and Hoppe, 2002; NJDEP, 2005). Of this total area, around half, or 398,000 acres, is considered to be of exceptional natural resource value (NJDEP, 2005). The area includes portions of Bergen, Hunterdon, Morris, Passaic, Somerset, Sussex, and Warren Counties (New Jersey Assembly, 2004). Recognizing that suburban sprawl development leads to increases in demand for water supply and increases pollution and runoff, the State passed the Highlands Water Protection and Planning Act of 2004 (New Jersey Assembly, 2004). The Highlands Act regulates the development of some 389,000 acres designated as a preservation area. Some 13 townships, towns, and boroughs in Hunterdon County fall partly or completely within the Highlands Preservation Area. The Act established the Highlands Water Protection and Planning Council and together with NJDEP this body was entrusted with the preparation of a regional master plan for the preservation area and the planning area, which is the remaining portion of the 88 municipalities affected by the Act. The master plan includes measures to promote compatible residential, commercial, industrial development, (and redevelopment) so as to avoid piecemeal growth or fragmentation of the surrounding areas. Agriculture, which currently comprises dairy farms and row crops, is also

regulated under the Act so as to maintain agricultural production while at the same time to address preservation goals of the Highlands area (Barringer et al., 2007).

### **4.3 Parcel-based Land Uses**

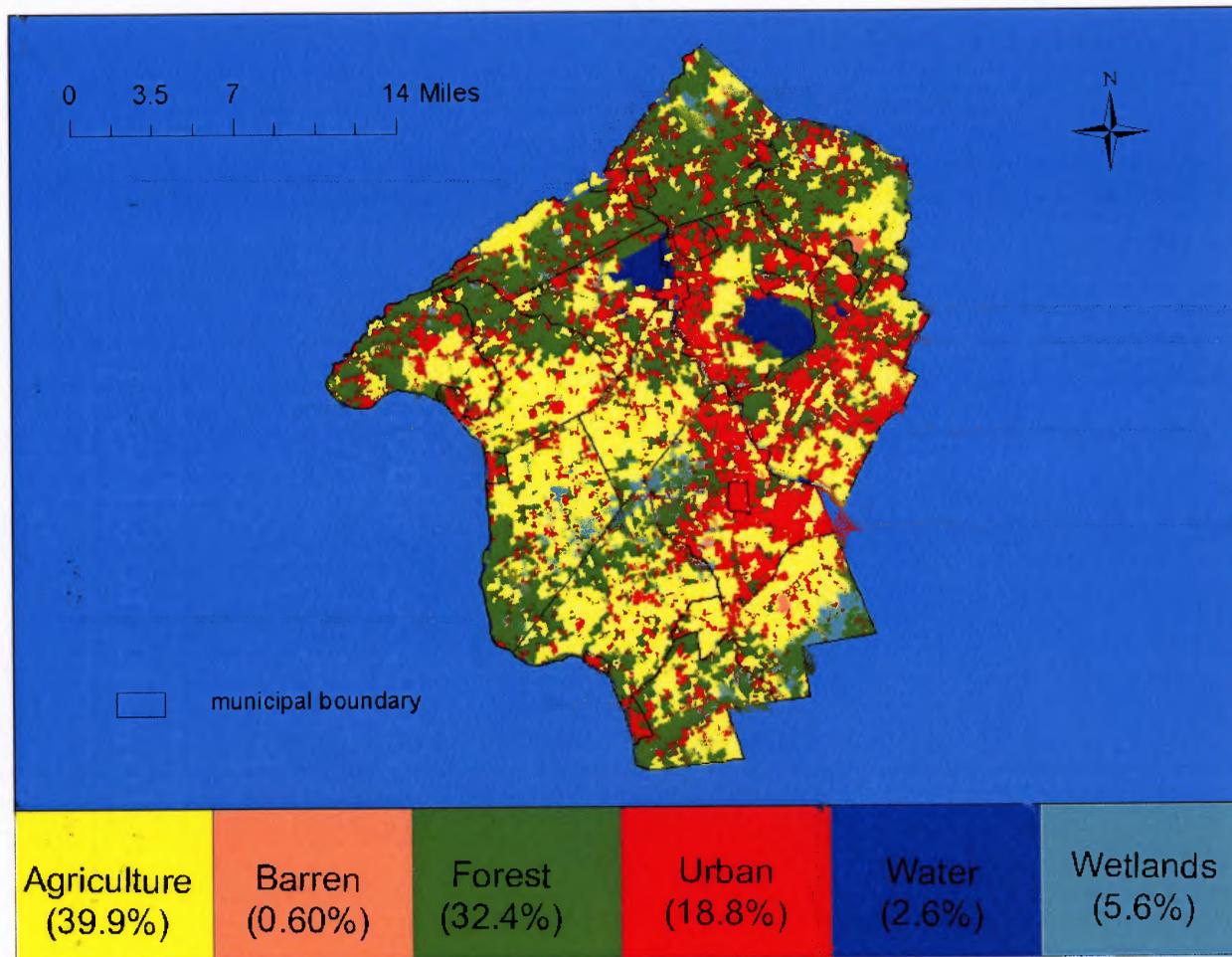
As discussed in Chapter 3, the first and foremost task for implementing the CA-based land use change modeling framework is to assign a single land use type to each parcel based on the land use/land cover maps in 1986, 1995 and 2002 maintained by NJDEP. The classification scheme presented in Figure 3.1 was applied to assign a single land use type to each land parcel in the county in those three years. Table 4.1 summarizes the classification results. The second column represents the total land areas of land uses compiled from the NJDEP land use/land cover maps. The third column is the total land areas of the re-classified land uses in all parcels in the County. The fourth column indicates their differences and the fifth column the percentage changes. Although a 100 percent accuracy of the classification scheme is desirable, achieving this level was not possible given the complex data requirements of the process. Its effect on model accuracy is thereby recognized.

It seems that the classification scheme used in this study consistently over allocates land to agriculture parcels by approximately 26 percent in all three years as shown in Table 4.1. At the same time, the model under allocates the land areas to forest, urban use, and wetlands parcels by 13 percent, 21 percent, and 38 percent, respectively. A possible explanation of the discrepancy can be found in the way specific land use categories are represented in the NJDEP map, particularly that of wetlands. For example, by virtue of the NJ Freshwater Wetlands Protection Act, some areas exhibiting at least

some wetland functions are classified generally as wetlands although some do not support typical vegetation as in those found in urban areas (Hasse and Lathrop, 2001). Many of these atypical wetlands also exist in forest and agricultural settings and are consequently showed interspersed within these areas. It should be recalled that the classification criteria in this study uses agriculture as a primary filter in assigning a land use class to each parcel. An agriculturally dominant parcel with some of its area comprised of wetlands (e.g., agriculture wetlands) will be classified as agriculture. As a consequence, land areas are added to agriculture at the expense of wetlands. Similar situations can be found in urban (e.g. disturbed wetlands) or forestlands (forested wetlands). By the same token, urban land and forest lands situated within a predominantly agricultural parcel will be erroneously classified as agricultural land.

**Table 4.1 Comparison of the Total Areas of the Re-classified Land Uses in Parcels to the Original Land Uses in NJDEP Land Use/Cover Maps in 1986, 1995 and 2002**

2002	NJDEP Land Use/Cover Map (acres)	Re-classified (this study) (acres)	Difference (acres)	Percent
Agriculture	82,089	107,935	25,846	31.49
Barren	1,528	1,700	172	11.23
Forest	101,996	87,515	-14,481	-14.20
Urban	62,992	50,861	-12,131	-19.26
Water	6,774	6,964	191	2.82
Wetlands	24,390	15,200	-9,190	-37.68
<b>1995</b>				
Agriculture	89,749	115,056	25,307	28.20
Barren	1,117	774	-343	-30.68
Forest	100,566	87,369	-13,198	-13.12
Urban	57,038	44,144	-12,895	-22.61
Water	6,108	6,722	613	10.04
Wetlands	25,190	16,111	-9,078	-36.04
<b>1986</b>				
Agriculture	105,083	123,872	18,789	17.88
Barren	1,920	1,558	-362	-18.86
Forest	97,213	86,321	-10,892	-11.20
Urban	49,930	39,877	-10,053	-20.13
Water	6,039	6,634	594	9.84
Wetlands	19,918	11,914	-8,004	-40.19

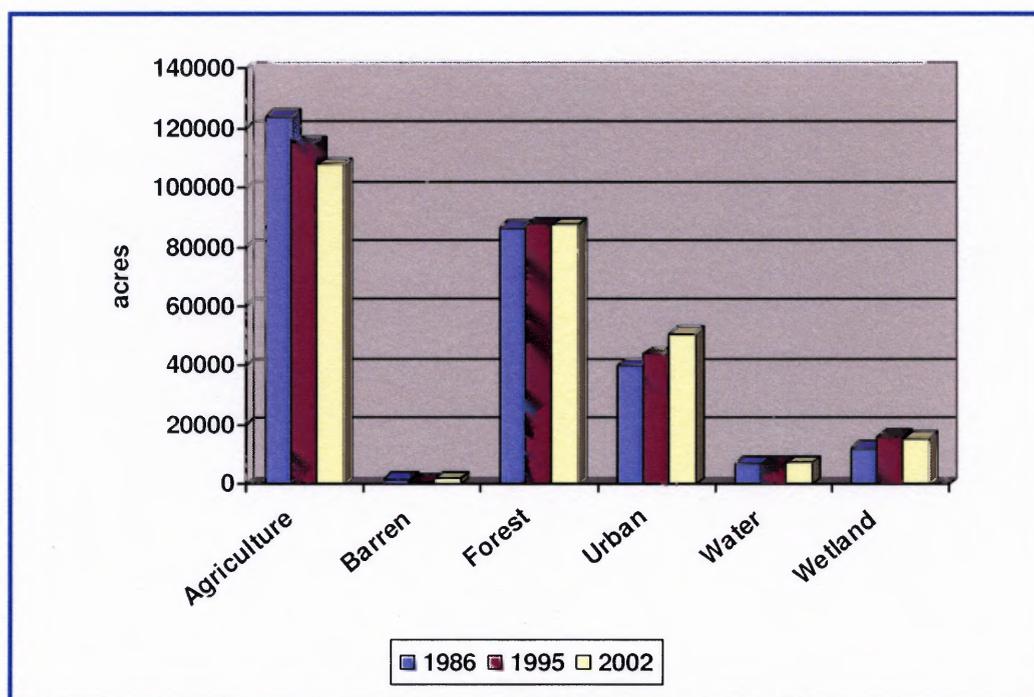


**Figure 4.3** Hunterdon County land use 2002.

Sources: Overlay Map generated from Hunterdon County Parcel Map, Division of GIS & NJDEP Land Use Land Cover Map 2002

Figure 4.3 presents the spatial distribution of the parcel-based land use classification in Hunterdon County in 2002. Forest lands make up the major portions of Holland, Bethlehem, Tewksbury, and Lebanon Townships. Forest lands are present in some portions of Alexandria in the northwest and Kingwood and West Amwell Townships in the south. Residential and other urban lands are spread in Readington, Flemington, Raritan Townships, and much of Clinton Town and Lebanon Borough. Like Califon Borough, High Bridge Borough has a mix of forest and urban lands with some small areas classified as water. Clinton Township is mostly urban and agricultural. Vast acreage of agricultural lands is found in most of Franklin, Kingwood, Delaware, East Amwell, and Alexandria Townships. The remaining wetlands of the county are mostly found in East Amwell and Raritan Townships.

Figure 4.4 compares the total areas of the re-classified land uses in three years. Agriculture is the predominant land use and occupies some 40 percent of the county area. There was a continuous declines in agricultural acreage in the county from 1986 to 2002. Forest lands stay steady and constitute the second major category, accounting for about a third of the county's total area. The third largest category is urban land occupying 20 percent of the total county area and this category shows a steady increase from 1986 to 2002. Other land use types occupy relatively small portions of the county including wetlands, water and barren land in a decreasing order in terms of acreage. While the acreages in barren land and water, are relatively unchanging, the acreage in wetlands shows an increasing pattern. As discussed earlier, the increases in wetlands are primarily caused by the NJDEP classification rules.



**Figure 4.4** Comparison of the total areas of the re-classified land uses among three years in Hunterdon County.

From Figure 4.4, one could easily conclude that the increase in urban land is coming from the loss in agricultural lands because other land use categories stay relatively steady over those periods. This though is not necessarily true because Figure 4.4 only presents the net changes in each land use category. Table 4.2 presents the flow of land use changes during the 1986-1995 (the upper panel) and 1995-2002 (the lower panel) periods based on the re-classified parcel-based land uses. As shown in the lower panel of Table 4.2, the total amount of non-urban land converted to urban land was 10,442 during the 1995-2002 period. Agriculture is the largest source of conversion to urban land followed closely by forest lands. From 1995 to 2002, a total of 4,995 acres of agricultural land was converted to urban use while 4,467 acres of urban land were converted from forest lands. These two land use types combined constitute 90 percent of the land converted into urban lands during that period. Meanwhile, there were also

considerable losses of urban land to agriculture and forest, which are 1,165 and 2,295 acres during that period, respectively. It is not intuitive to understand the conversion process from urban land to non-urban. Detailed discussion on the process will be presented in the next Section.

It is noteworthy that although agricultural and forest lands are the primary supply to the new urban lands, the total amount of forest land in the county has remained relatively constant between 1995 and 2002. The loss of the forest land has apparently been offset to some extent by the gains from agriculture to forest lands. The total amount of agricultural land converted to forest land is about 5,399 acres during that period. As a consequence, loss of agricultural lands is accelerated.

The information presented above can be used to estimate the amount of the remaining land available for development in the County. As shown in Table 4.1, agricultural, forest, and barren land added up to 197,148 acres in 2002. Among them, 4,392 acres of lands are not suitable for development because of their steep slope of more than 15 percent<sup>1</sup> and 31,000 acres are protected from development.<sup>2</sup> Therefore the amount of land available for development in Hunterdon County is about 161,756 acres. Using an average annual urban growth of 1,600 acres,<sup>3</sup> and assuming a steady rate of growth into the future, the county is estimated to be fully built up in about 100 years.

---

<sup>1</sup> Calculated from DEM Raster file of Hunterdon County provided by the NJ DEP Bureau of GIS

<sup>2</sup> Taken from Hunterdon County Open Space GIS Coverage. Retrieved April 17, 2008, from [http://gis.co.hunterdon.nj.us/Hunterdon/HC\\_GIS\\_MAP\\_Download.htm](http://gis.co.hunterdon.nj.us/Hunterdon/HC_GIS_MAP_Download.htm).

<sup>3</sup> Average urban loss is estimated from the period 1986 to 2002 from Table 4.2

**Table 4.2 The Flow of Land Use Changes in Acres in Hunterdon County during the 1986-1995 and 1995-2002 Periods**

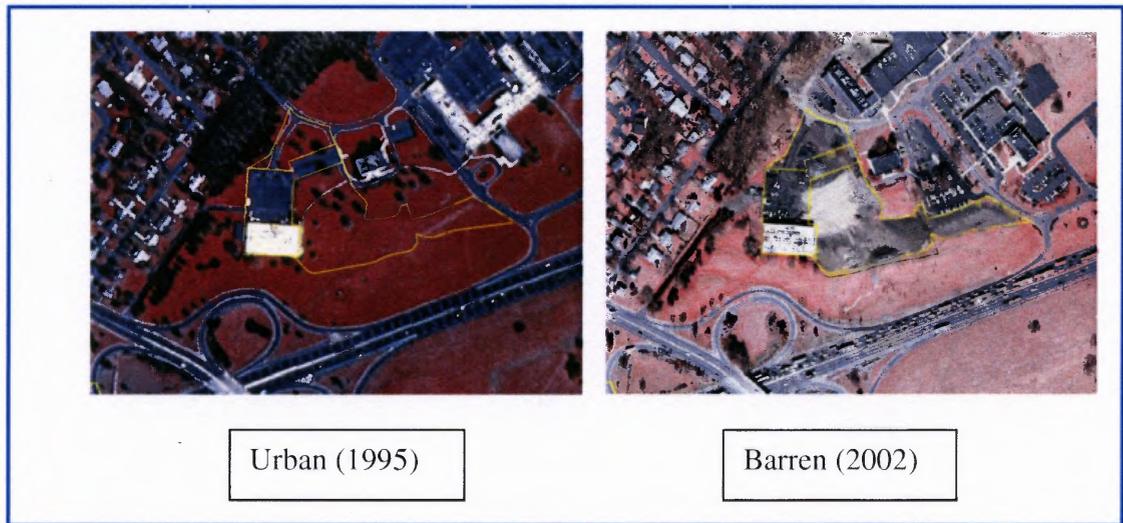
1995							
1986	Total	Agriculture	Barren	Forest	Urban	Water	Wetlands
Agriculture	123,872.18	106,713.04	340.73	6,229.62	7,370.29	10.62	3,207.88
Barren	1,557.57	1.59	291.41	489.29	729.08	9.72	36.49
Forest	86,321.36	5,142.60	96.18	75,083.66	4,516.19	51.14	1,431.58
Urban	39,876.91	3,062.38	45.55	5,107.51	31,356.35	16.28	288.84
Water	6,633.58			2.95	3.27	6,627.36	
Wetlands	11,913.66	135.95	0.59	455.6	168.42	6.48	11,146.61
<b>Total</b>	<b>270,175.26</b>	<b>115,055.56</b>	<b>774.45</b>	<b>87,368.63</b>	<b>44,143.60</b>	<b>6,721.59</b>	<b>16,111.41</b>
2002							
1995	Total	Agriculture	Barren	Forest	Urban	Water	Wetlands
Agriculture	115,055.56	103,315.36	965.71	5,399.59	4,995.50	61.67	317.72
Barren	774.45	7.85	270.36	114.36	381.89		
Forest	87,368.63	2,870.26	378.9	78,912.89	4,467.41	98.48	640.69
Urban	44,143.60	1,165.24	80.19	2,294.89	40,419.78	7.32	176.17
Water	6,721.59			43.79	24.92	6,643.72	9.17
Wetlands	16,111.41	576.05	4.55	749.6	571.61	153.09	14,056.52
<b>Total</b>	<b>270,175.26</b>	<b>107,934.76</b>	<b>1,699.72</b>	<b>87,515.12</b>	<b>50,861.11</b>	<b>6,964.29</b>	<b>15,200.26</b>

#### 4.4 Understanding Urban Land Conversions

As shown in Table 4.2 and discussed above, there has been a considerable loss of urban land to different non-urban land. While the conversion from non-urban to urban land is easily perceived, the reverse process appears counterintuitive. However, as observed from the data gathered for this research, such transformation does indeed happen. This section describes these processes involving urban use conversion. It should be recalled that the urban land use category encompasses a wide range of specific land use types which vary greatly in the amount of impervious cover or development that is associated with them (NJDEP GIS Metadata for 2002 Land Use Map).

The urban to barren or urban to forest cases usually happen to parcels that were undergoing a re-development process. Examples would be abandoned industrial sites

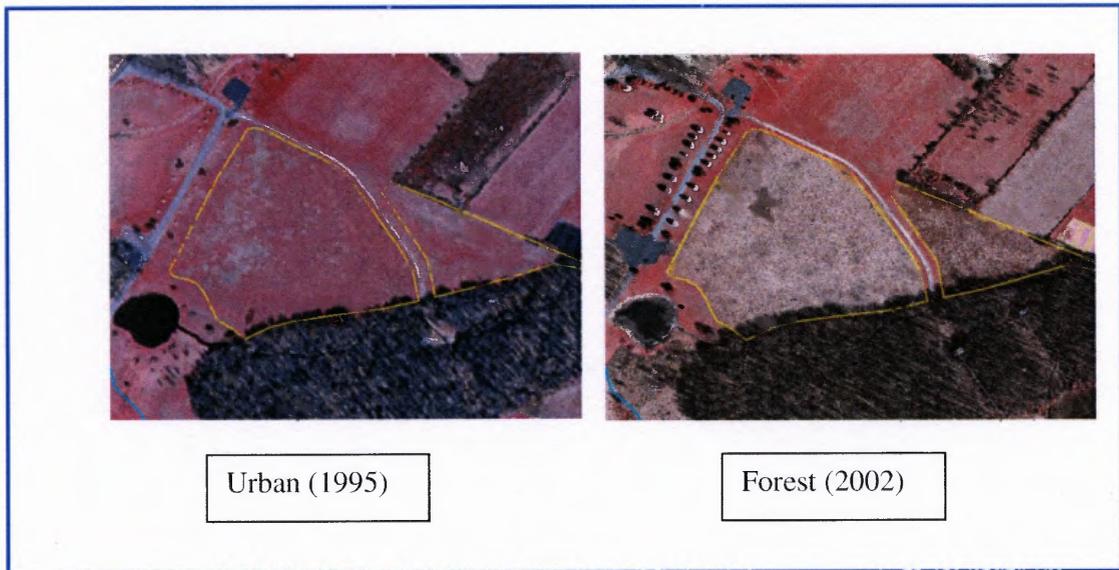
awaiting re-development. In the interim, buildings may be removed and land areas graded, but until any new development is completed, these areas would be classified under the Barren Land category, called Transitional Areas. Another case would be an industrial facility that is not being demolished, but is being altered. The portion of the site that is being altered would be called Barren Land as shown in Figure 4.5.



**Figure 4.5** Transformation from urban land to barren land: transitional areas.

Source: NJDEP Bureau of GIS

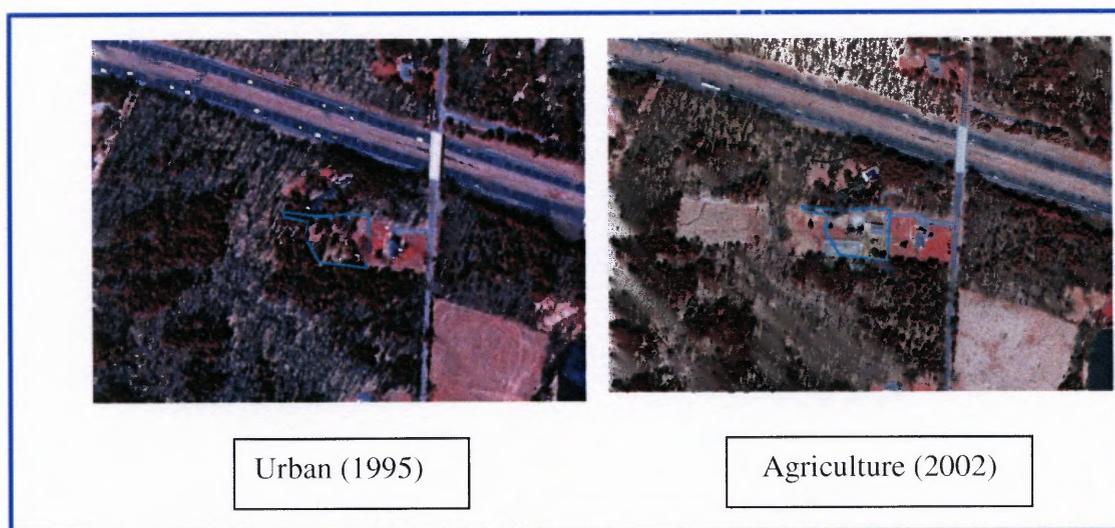
If these types of parcels remain this way with no further intervention for a while, or if lawn areas are no longer maintained, grasses and shrubs will grow. At this stage, such areas would then be classed as Old Fields or Scrub/Shrub areas, both of which fall under the general category of Forest. An example is presented in Figure 4.6.



**Figure 4.6** Transformation from urban land to forest: old field.

Source: NJDEP Bureau of GIS

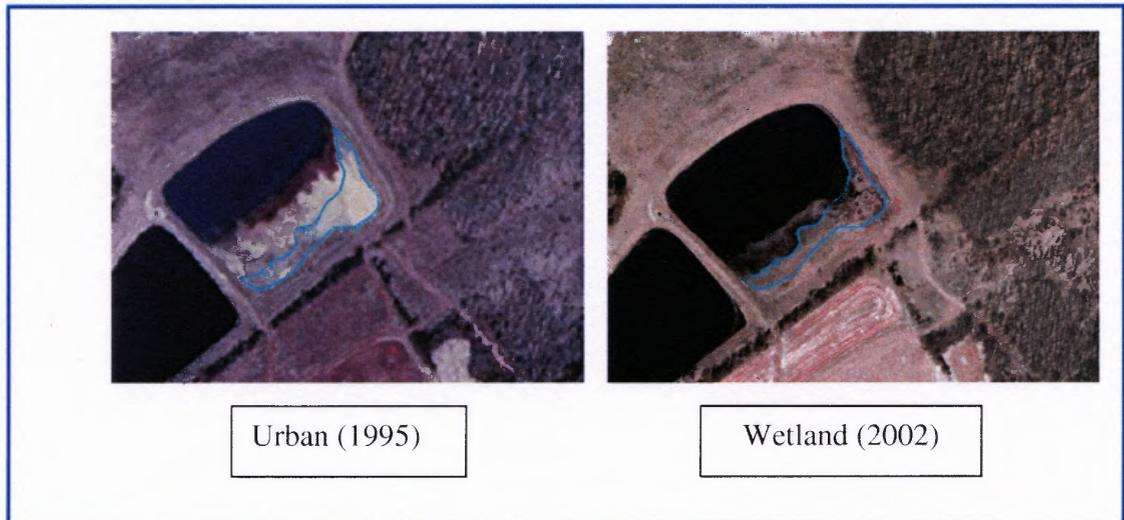
Urban land can also be changed to cropland or to other agricultural categories. These categories could be a horse farm or outbuildings (e.g., hothouses, barns, storage structures) associated with agricultural operations. These kinds of areas fall under the general agriculture category, but are in the specific category of Other Agriculture such as orchards, vineyards, nurseries, or horticultural areas. Figure 4.7 is an illustration of this situation (Urban: Rural Residential to Agriculture: Other Agriculture). The area highlighted in Figure 4.7 includes portions of a rural residence that has been converted from lawn and trees to a site with a small barn and some horse or other animal corrals. These types of parcels do not have active cropland associated with them, but are nevertheless mapped as Agricultural land (Tyrawski, 2007).



**Figure 4.7** Transformation from urban: rural residential to agriculture: other agriculture.  
Source: NJDEP Bureau of GIS

The Urban to Wetland cases generally involve small areas that are being converted as part of wetland mitigation measures. Wetlands that are part of mitigation measures are those that have hydric soils (i.e., evidence of soil saturation on the imagery), but because of artificial alterations do not have typical wetlands vegetation. Examples would be farm fields that are being cultivated, but are considered altered wetlands. Such areas may be under some state wetland jurisdiction and so are mapped under the general Wetlands category. These types of areas are not generally mapped as wetlands but are included as wetlands in the state datasets because of the regulatory implications (NJDEP/DLUR 2007; Tyrawski, 2007). In other cases, the urban to wetland change arises from an activity that results in a small expansion of the wetland area. This situation is exemplified in Figure 4.8 where a pond has been expanded. The expanded area evolved from an apparent upland managed urban lawn to an apparent wetland area, with obvious saturation and even a new small drainage channel created. As with all other land use

categories, these wetland types can be identified by the more detailed land use code and land use label associated with any of the polygons in the NJDEP Land Use Maps.



**Figure 4.8** Transformation from urban: rural residential to wetland.  
Source: NJDEP Bureau of GIS

#### 4.5 Driving Factors of Land Use Changes

The last section describes data processing on the driving factors used in the land use change model. Based on the criteria suggested by Lu (2002), a total of 14 variables were selected as appropriate for this study. Table 4.3 lists these factors and the method used to obtain them. Driving factors used in other studies are given in Appendix B. Spatial data on soil suitability, slope, and distances to urban centers, roads, and streams are discussed below. The other factors were discussed previously.

**Table 4.3 Driving Factors for Modeling Land Use Change**

<b>Driving factors</b>	<b>Method</b>
Percent agriculture	Buffer analysis in Arc GIS
Percent urban land	Buffer analysis in Arc GIS
Percent forest	Buffer analysis in Arc GIS
Percent wetland	Buffer analysis in Arc GIS
Percent barren	Buffer analysis in Arc GIS
Percent urban	Buffer analysis in Arc GIS
Distance to urban center	Near distance function in Arc GIS
Distance to streets and roads	Near distance function in Arc GIS
Distance to streams	Near distance function in Arc GIS
Slope	Spatial Analyst Extension
Soil type	Soil Map from NJ SSURGO
Land use type	Land use map from NJDEP
Wetland area inside parcel	Land use map from NJDEP
Parcel size	Parcel map

#### 4.5.1 Soil

Soil is expressed in terms of soil suitability for development and is dictated by soil type. Each parcel is given a soil type designation corresponding to the dominant soil type in the parcel. The soil attribute was derived from the Soil Survey Geographic (SSURGO) data available from the Natural Resources Conservation Service (NRCS) of the US Department of Agriculture (USDA). Following the procedure in evaluating soil suitability in the Raritan River Basin, the soil attribute assumes values 0 to 6 indicating the number of limitations for 6 development applications in the county.

Although the USDA-NRCS soil survey lists 26 community development applications when evaluating soil suitability, six of these applications were considered for rating the suitability of the soil in the study area. The New Jersey Water Supply Authority (NJWSA), (2003) rated the soils in the Raritan River Basin according to the following six development applications: septic tank absorption fields, foundations for

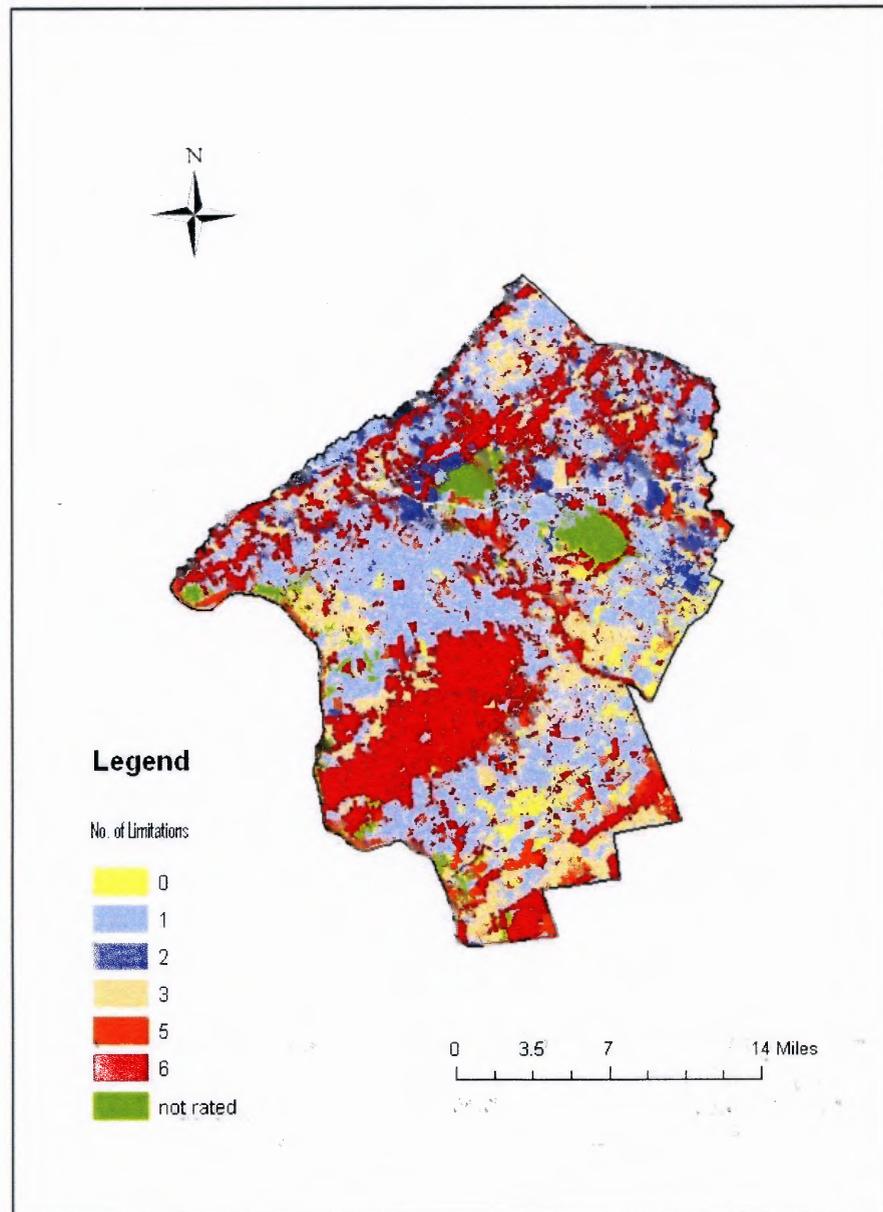
dwellings with and without basements, local streets and roads, foundations for small commercial buildings, and lawns, landscaping, and golf fairways. Soils were designated as “most suitable,” “moderately suitable,” and “least suitable” for development based on the number of soil limitations to those applications. A rating of “most suitable” indicates that the soil is not severely limited in any of the six development applications stated above. On the other hand, “moderate” suitability indicates limitation in one up to three in six of the development applications, while four to six indicates “least” suitability for development.

Using the same parameters, data from the NRCS Soil Survey were downloaded to create a soil suitability map for Hunterdon County. The suitability map in Figure 4.8 depicts the number of development applications that limit the use of the soils in the county. Only 3.2 percent (8,566 acres) of the total land area of Hunterdon County can be used in all six development applications cited above. More than half (59.3 percent), representing 160,266 acres, have one to three limitations in development applications while 33.0 percent (89,233 acres) is deemed least suited for development as these soils are limited in at least four categories. The remaining 4.5 percent was not rated as this acreage was comprised of either rough broken land consisting of shale or bodies of water.

In terms of the number of parcels, the percentages of the parcels that have limitations in 1, 2, 3, 5, and 6 development applications are 43, 9, 13, 9 and 20 percent, respectively. Only 8,566 parcels (or 3.5 percent) of the total parcels are suited for all six development applications and the rest of the parcels are not rated.

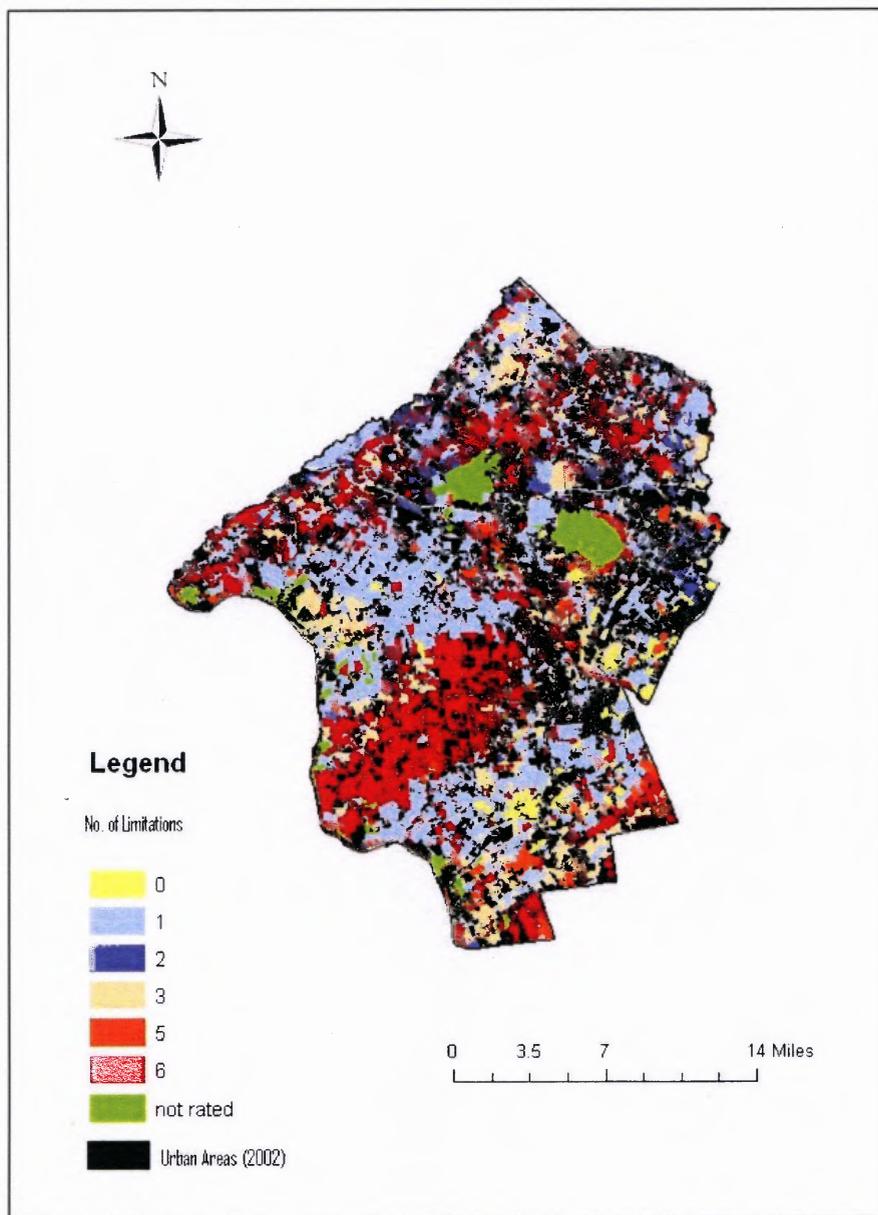
Rather than employing the rating scheme used in the NJWSA (2003) study, this research used the number of restrictions to development as a surrogate for the soil

suitability driving factor. These values range from 0 to 6 to indicate the number of development restrictions. The use of actual numbers instead of descriptions gives more information and therefore accuracy, in the modeling process.



**Figure 4.9** Soil suitability for development in Hunterdon County.

Bucks silt loam series (BucB and BucC2) are most suited for development and consist of approximately 7,175 acres. The largest “least suitable” soils in terms of acreage are the Parker cobbly loam (ParEe) and chalfont silt loam (CheB), which comprise approximately 15,900 and 12,832 acres of the county, respectively. Of the 120 soil types in the county, only 20 percent are suited for development applications such as dwellings with or without basements and commercial buildings. All 120 soils are severely limited in certain applications including septic tank absorption and for the construction of local roads and streets. Considering that much of the area is unsewered, it is necessary to employ soil treatment methods to enable adequate septic tank operation. Figure 4.10 compares the soil suitability map to newly urbanized lands in 2002. It seems that much of the new development occurred in areas not amenable for many types of development applications. A large concentration of urban areas is located in those areas that have the least to moderately suited soils, such as the townships of Raritan and Clinton, the City of Lambertville, and the boroughs of Califon, Flemington, Glen Gardner, Hampton, High Bridge, Lebanon and Milford.



**Figure 4.10** Location of urban areas relative to soil suitability.

#### 4.5.2 Slope

The average slope was estimated for each parcel from a Digital Elevation Model (DEM ) file provided by the United States Geological Survey. This is a data file that contains the elevation of the terrain over a specified area. The NJDEP Bureau of GIS data download

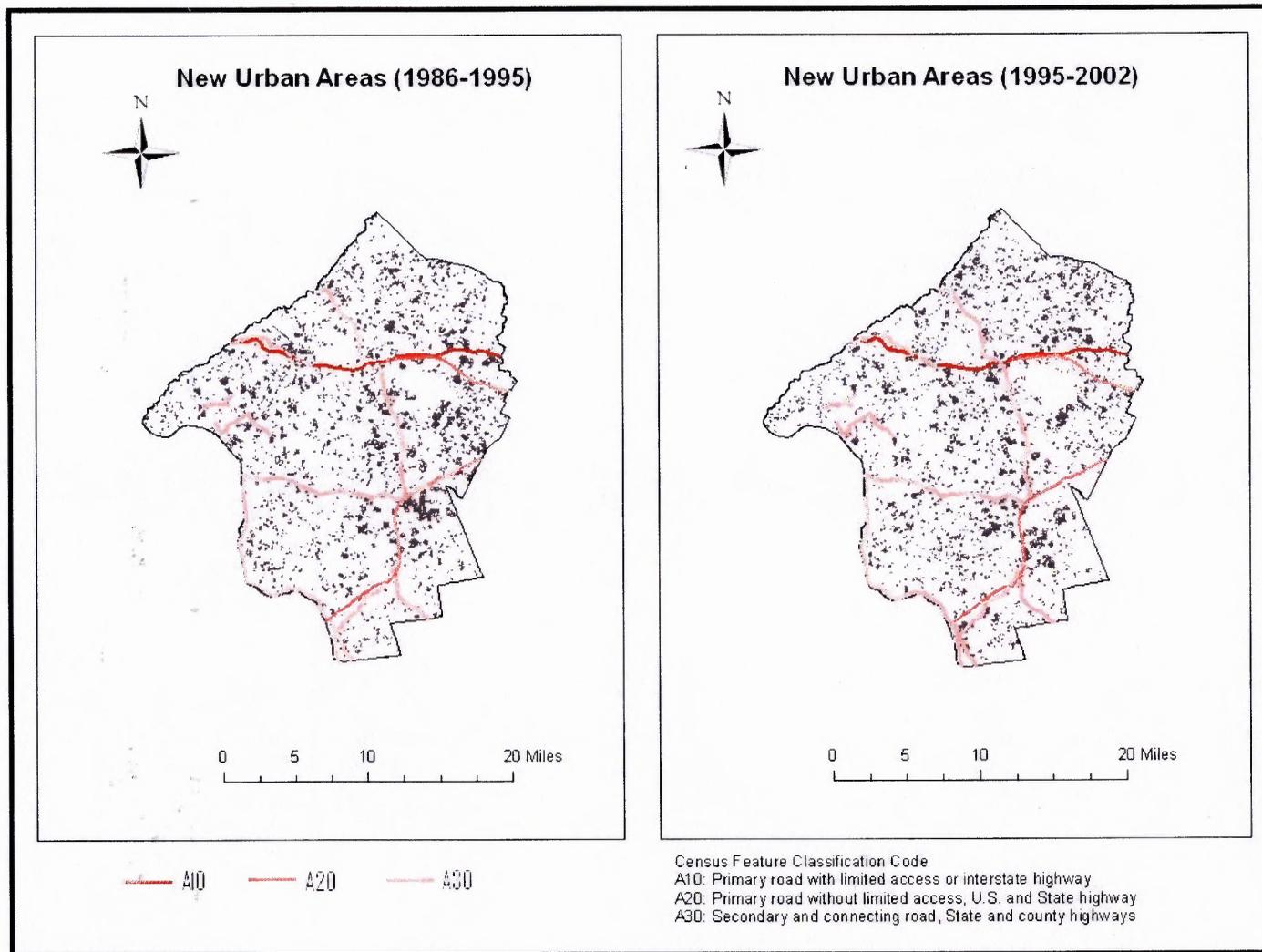
site did not have a ready DEM file for Hunterdon County. Instead, four DEM files for the four Watershed Management Areas to which Hunterdon County is situated were fused and then clipped using the county-boundary shapefile. The local statistics function in ArcGIS was then used to calculate the average parcel slope. The downloaded DEM file was in raster digital data format using a 10-meter digital elevation grid.

Slope consideration in land use planning is based on avoiding damage to the environment, structures, and utilities. Different land uses have varying slope requirements. Among several land types, urban land, playgrounds, industrial sites, and high speed highways have the most stringent slope requirement ranging from 2-3 percent (Marsh, 1992). Safety considerations account for such requirements. Slope requirements for cropland are generally assigned to a maximum of 10 degrees as activities in these areas require operating tractors and other farm machineries.

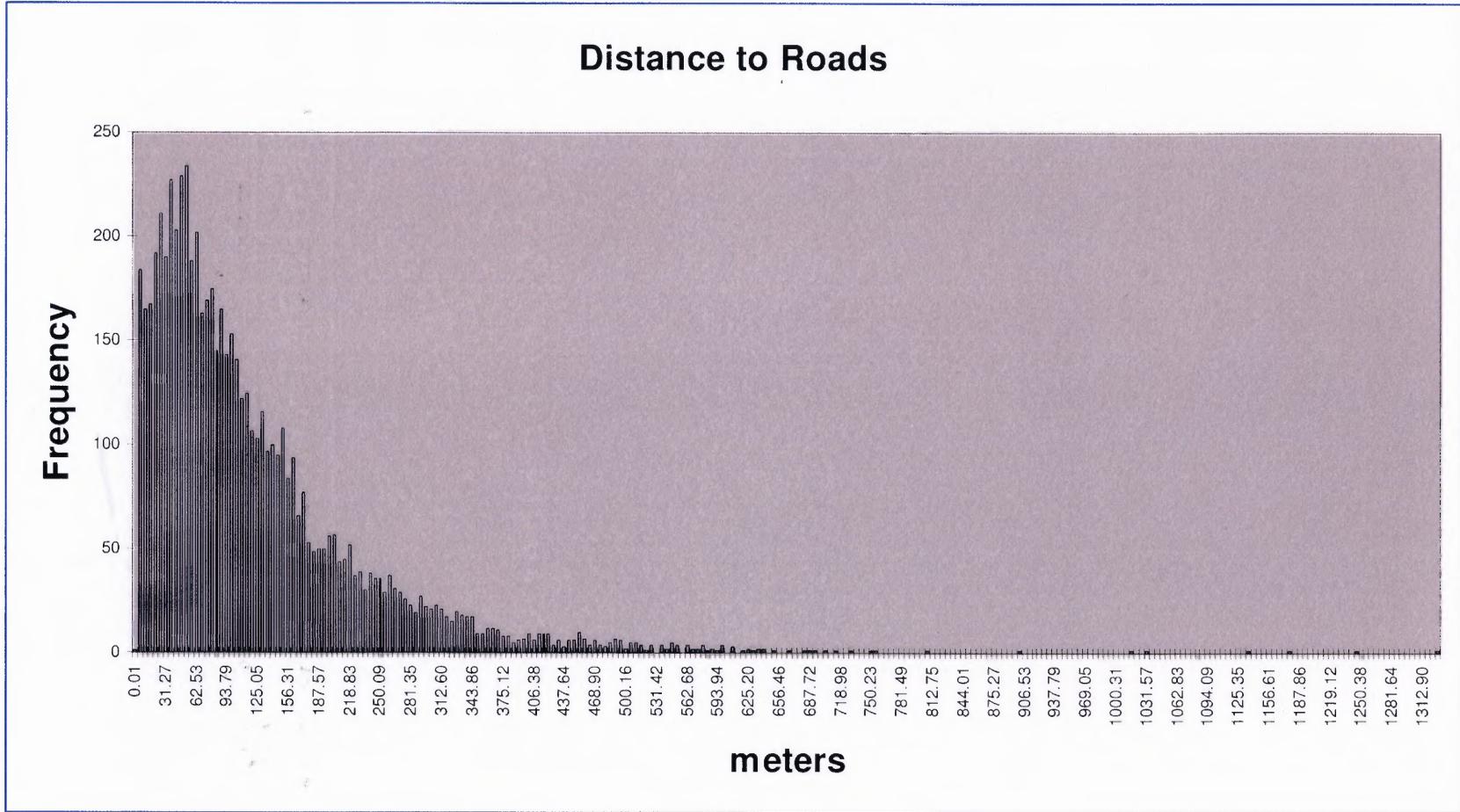
Slope values for Hunterdon County are relatively uniform between urban and non-urban land use types. Of the 40,174 parcels classified as urban in 2002, there were 812 parcels that had slopes greater than the 15 percent threshold and are not suitable for development. These parcels represent a total area of 2,454 acres, around six percent of the total urban area. These parcels are primarily distributed in Lambertville City and Tewksbury and Lebanon Townships. On a countywide scale, the average slope for all urban parcels is 4.40 percent. The average slope value for non-urban land use types is roughly the same at 4.38 percent.

### **4.5.3 Proximity to Transportation Networks and Urban Centers**

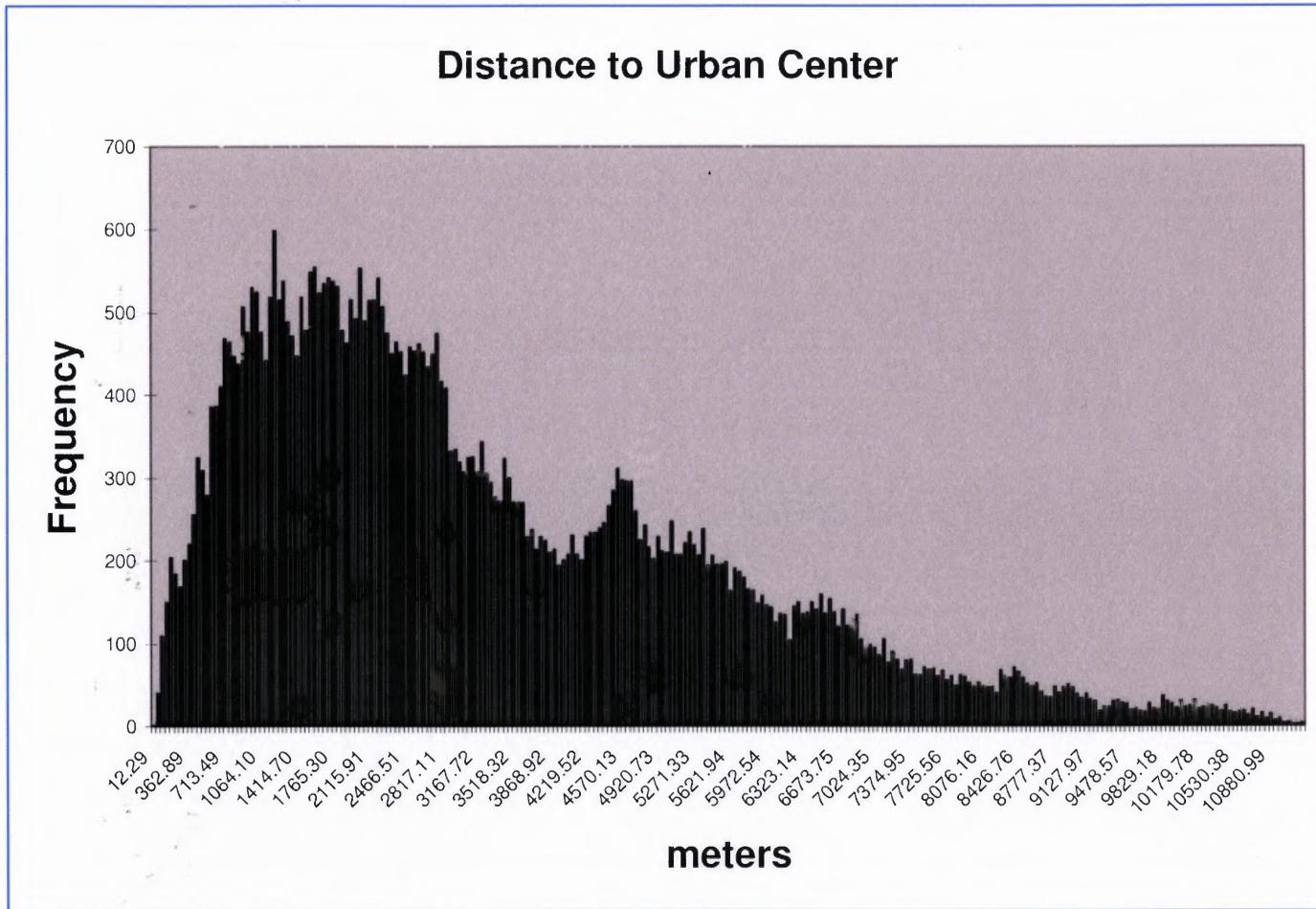
Site selection is heavily influenced by accessibility. The modern phenomenon of urban sprawl is the result of changing transportation patterns and accessibility (Ozmen-Ertekin, 2007; Lucy and Philips, 1997; Lewis and Maund, 1976). Figure 4.11 presents the spatial distributions of the major road network versus the urbanized parcels during the 1986-1995 (left panel) and 1995-2002 (right panel) periods. The newly-urbanized areas are more or less evenly distributed throughout the county. The impacts of road networks on the urbanization are not apparent. One reason for this could be that highways are not too distant from each other. A prospective real estate developer can choose to locate a project anywhere in the county and still be close to any highway. Another reason is the presence of an extensive network of access roads providing easy driving to the major highways. To verify whether local roads will give a better indication of urban location trends, a frequency distribution was generated showing the distance of newly urbanized areas in 2002 and distance to local roads. Figure 4.12 shows that a proportional relationship exists between urban location and distance to local roads. There is a greater concentration of urban areas near the local roads. By the same token, Figure 4.13 shows that urban centers tend to promote urbanization in its vicinity. Urban locators are concentrated at a distance of 1 to 2.5 km from the nearest urban center. Thereafter, the number of urban locators tapers off with distance.



**Figure 4.11** Urbanization along roads and highways.



**Figure 4.12** Distance of new urban areas relative to nearest local road.



**Figure 4.13** Distance of new urban areas to nearest urban center.

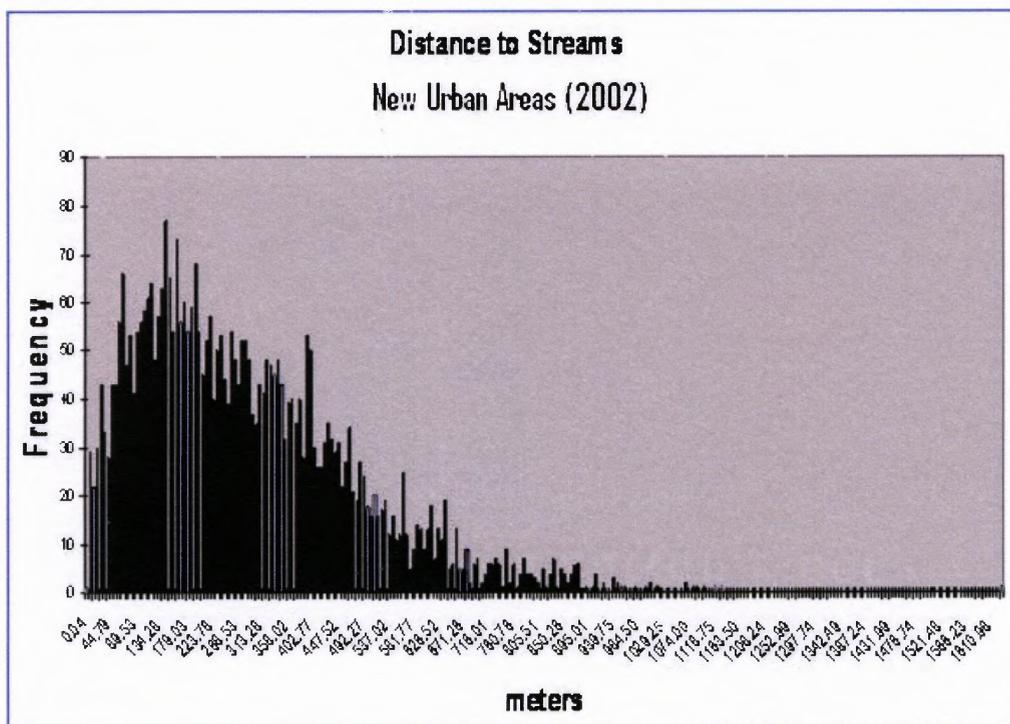
The distance to the nearest urban center and major highway was estimated using the Euclidean Distance function in ArcGIS. This tool measures the straight-line distance from the centroid of each parcel to the closest source which can be the objects of interest, such as wells, roads, or schools. As to the location of urban centers, population density from census data was used. The census tract having the highest population density in a municipality was then the designated location of its urban center.

#### **4.5.4 Distance to Streams**

The last attribute affecting land use is distance to streams. People place a high premium on living near aquatic ecosystems because of their aesthetic or recreational features (Qiu et al., 2006). For instance, Wilson and Carpenter (1999) have demonstrated in their hedonic pricing studies that land values are higher in places near lakes and streams. Yet Walsh et al. (2003) claimed that few modeling studies include lakes and streams as predictors of land use/cover distribution. Walsh attributes the trend to lack of studies that quantified the effect of aquatic ecosystems on land use distribution. Another reason is the changing attitudes towards these ecosystems (i.e., wetlands whose role in maintaining surface water quality have not been appreciated recently).

Figure 4.14 presents the distribution of the distances of the 3,942 newly urbanized parcels to the nearest streams during the 1995-2002 period. The number of parcels (the frequency indicated by the vertical axis) increases as their distances to the nearest streams increases and then decreases gradually as the distance continues to increase. The average distance of an urbanized parcel during the 1995-2002 period to the nearest stream was 282 meters. The average distance of the 971 urbanized parcels during the 1986-1995 to

the nearest streams was 263 meters. The increased distance between those two periods may reflect recent regulations and practices protecting the stream corridors and riparian areas in the interest of maintaining water quality.



**Figure 4.14** The distribution of the distance of the urbanized parcel during the 1995-2002 period to the nearest streams.

#### 4.6 Summary

The foregoing discussion provided a snapshot of the current state of land uses in Hunterdon County. It was followed by an enumeration and discussion of the driving factors of land use changes – the parcel attributes that will be used in calibrating the land use change model presented in the next chapter.

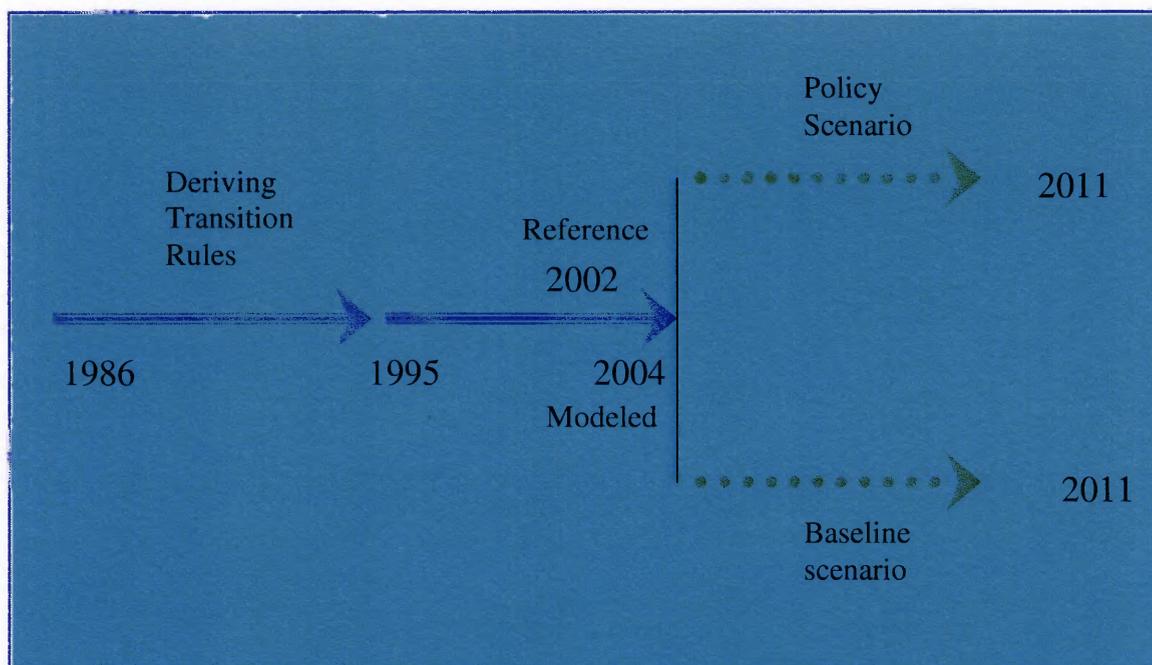
The driving factors were quantified in order to provide a benchmark from which the magnitude of land use change can be calculated. To evaluate the effectiveness of land

use policies, estimates of land use conversions resulting from implementation of these policies are necessary. A better understanding of these driving factors and how they interact can provide some useful insights on the magnitude of their effects in a changing landscape - information that will aid in model assessment.

## **CHAPTER 5**

### **RESULTS AND DISCUSSION**

Chapter 5 reports and discusses the land use change modeling results for Hunterdon County, New Jersey. As presented in Figure 5.1, the transition rule will be derived using 1986 and 1995 parcel based land use changes and their driving factors. The derived transition rules will be used in the CA module to predict land use changes from 1995 to 2004 using the parcel based land use data in 1995 since the time interval in which the transition rules were derived was nine years (1986 to 1995). The model was validated by comparing the predicted land use pattern in 2004 to the parcel-based land use data for 2002 as discussed in Chapters 3 and 4 because of the data limitation. By using the 2002 land use data as a reference to validate the prediction for 2004, it is assumed that changes within this two year period would not be considerable. The prediction accuracy was evaluated using various Kappa statistical values as discussed in Chapters 2 and 3. The validated model was used to predict the land use changes from 2002 to 2011 based on the land use data in 2002. Two scenarios were modeled. A baseline scenario is a “business as usual” situation wherein policy interventions by the government are not included in the modeling process. The policy scenario incorporates government policy interventions such as down-zoning and the preservation of farmlands and open space by modifying the validated decision tree. The chapter ends with a sensitivity analysis undertaken for the model to test its behavior to varying inputs.



**Figure 5.1** Model timeline schematic

### 5.1 Transition Rules

As discussed in Chapter 3, the transition rules used by the CA module of the land use change model were generated by the decision tree module using the J48 software. Appendix C presents the results of the transition rules from the J48 Decision Tree program. The first split in the decision tree is current land use type. Among the six land use types, urban land is considered first. Within the urban land use designation, parcel area, percent barren and forest land in the neighborhood of the parcel significantly determines the future land use status of those parcels. Urban parcels less than 2.32 acres tend to remain urban. Some exceptions exist for parcels with steep slopes. Steep slopes tend to discourage development since construction costs would be prohibitive. These parcels tend to be converted to forest lands. Another exception in the small urban parcels category is that the parcels with significant amount of water and wetland in their

neighborhoods tend to be converted to forest lands as well. Although large urban parcels (>2.32 acres) mostly remain in urban use, some of them could be converted to forest and agricultural uses. These parcels tend to be converted if they have a big portion of wetlands within them, three and more soil restrictions to urban development, lower percentage of urban and barren land, and higher percentage of water, agricultural, forest, wetlands in the neighborhood, and are farther away from highways and urban centers with steeper slope.

Some agricultural parcels can be converted into urban, forest, barren and wetlands depending primarily on their neighborhood land use distribution and parcel size. Agricultural parcels tend to be converted into urban uses when there is high percentage of urban land in their neighborhood and the parcel size is small. Even with a low percentage of urban land in their neighborhood, agricultural parcels between 1 and 5 acres tend to be converted to urban use provided the slope is not steep. These characteristics are apparently consistent with restrictions related to construction and zoning laws. The conversion from agriculture to forest could occur to those parcels with a high percentage of urban land in their neighborhood when there are severely restricted soils for development and a high percentage of forest (at least 21 percent), agriculture, and barren lands already existing in the neighborhood for those parcels. Agricultural parcels with a small percentage of urban land in their neighborhood can be converted into forest lands because of steep slopes. High erosion potential of steep lands may hamper farming leading farmers to abandon the area which eventually initiates plant growth. The high percentage of forest in the neighborhood of those parcels tends to increase the likelihood of converting into forest. Agricultural parcels with a significant amount (at least 11

percent) of barren land in their neighborhood have the potential of becoming newly barren land as well. The neighborhood characteristics of these parcels might also include the low percentage of forest lands and the smaller wetland area within them. The conversion to wetlands usually occurs to the large agricultural parcels that have significant wetlands.

Forest parcels with a high percentage of urban land and a low percentage of barren land in their neighborhood are usually converted to urban use. This type of conversion tends to occur in the case of small forest parcels. Forest parcels could be converted to agriculture or wetlands when there is a significant presence of agricultural land already in their neighborhood or wetlands within these parcels, respectively.

The actual amount of wetlands within a wetland parcel usually determines its future status. The wetland parcel with a large amount of wetlands (>49,000 square feet) within it always remain as wetlands. However, wetland parcels with smaller amounts of wetlands have higher likelihood of converting into urban in a high urban neighborhood or they become barren lands if the parcel has three restrictions to urban development.

Barren land parcels can be developed into urban lands or remain as barren. When their neighborhoods have a relatively small percentage of barren land ( $PCTBAR < 0.052$  percent), the barren land parcels are always converted into urban. If there is a high percentage of barren land in their neighborhoods, the smaller barren lands ( $\leq 15$  acres) are likely to be developed into urban land, and the large barren land parcels might stay barren. Finally, the land parcels classified as water, or artificial and natural lakes usually do not change land use type.

In summary, the presence of urban and barren land in neighboring parcels and the size of the parcel are two important factors affecting the future status of the parcel besides its current land use type.

In theory, the accuracy of the decision tree model, or any data-mining algorithms for that matter, depends heavily on the preprocessing done to the data set used to build the model. For (continuous) numeric attributes (e.g. parcel area, percent urban, percent agriculture), very minimal difficulty is seen as the algorithm used to build the decision tree has been empirically proven to be effective in handling numeric attributes (Quinlan, 1996b). The accuracy of the decision tree depends on how (discrete) nominal attributes (e.g., soil type, parcel-based land uses) are preprocessed.

The impact of nominal attribute preprocessing was actually evident when soil types (having more than 40 nominal values) were initially used to generate the decision tree. The resulting decision tree posted a poor accuracy of 65 – 70 percent on re-substitution (i.e., testing the model using the training dataset) alone. Having too many values on a nominal attribute causes high entropy (i.e., a high degree of disorderliness for that attribute). According to information theory as discussed in Section 2.6.1, an attribute with higher entropy generally requires more splitting before a terminal node is obtained. For this reason, the soil limitations for development were used to characterize the impacts of soils in developing the transition rules from the decision tree. When the number of values that characterizes the impacts of soils was effectively reduced to 6 (i.e. 6 soil restrictions), the accuracy was increased to 81.4 percent in re-substitution, and ~85 percent for the test set data using the bootstrap method in evaluating the decision tree model.

The accuracy of the decision tree is also affected by how single land use can be correctly assigned to a parcel based on the land use map derived from aerial imagery as discussed in Chapter 4. Although great efforts have been given to develop a reasonable classification scheme, limitations do exist. Some land use classes were overestimated, while others were underestimated. Small errors from the outset of a multi-step process can affect the accuracy of the final classification. Nevertheless, these errors were not too significant to affect the performance of the model as will be shown in the next section.

## **5.2 Land Use Change Model Validation**

### **5.2.1 Land Use Prediction for 2004**

The decision tree-based transition rule was generated using 1986 and 1995 parcel-based land use maps. Since the time interval for this duration which the transition rules were derived was nine years (1986 to 1995), the CA model predicted the land use pattern for 2004 using the 1995 land use data.

The predicted land use distribution for 2004 in terms of the number of parcels is presented in Table 5.1. As shown in Table 5.1, urban parcels increase to 39,386 in 2004 from 33,866 in 1995 while the number of other land use parcels decrease (except water remains the same). Of those parcels converted to urban use, 3,306 (or 32 percent of total parcels) were previously forest land. The second largest contributor to urban land is agriculture with 1,434 parcels that represent 24 percent of the total number of agricultural parcels in 1995. Almost half (48 percent) of the wetlands parcels in 1995 were converted to urban use. About 80 barren land parcels were converted into urban during the same

period. There were also conversions from agriculture (24 parcels) and wetlands (35 parcels) to forest and from urban to barren (38 parcels) and forest (1 parcel).

**Table 5.1 The Predicted Land Use Distribution for 2004 in Terms of the Number of Parcels**

Predicted 2004 Land use (no. of parcels)							
1995	Agriculture	Barren	Forest	Urban	Water	Wetlands	Total
Agriculture	4,463		24	1,434			5,921
Barren		318		80			398
Forest			6,872	3,306			10,178
Urban		38	1	33,827			33,866
Water					142		142
Wetlands			35	739		763	1,537
<b>Total</b>	<b>4,463</b>	<b>356</b>	<b>6,932</b>	<b>39,386</b>	<b>142</b>	<b>763</b>	<b>52,042</b>

Table 5.2 presents the predicted land use patterns for 2004 in terms of total acreage. Forest lost the most use amounting to 24,735 acres followed by agriculture with 6,024 acres. Wetlands contributed a further 5,131 acres to urban development in that period. The loss of wetlands is quite significant as this amount represents almost a third of the county's wetlands in 1995. The forest loss was partially offset with the reforestation of 355 acres from wetlands, agriculture, and urban parcels.

**Table 5.2 The Land Use Pattern for 2004 in Terms of the Total Acreage**

Predicted 2004 Land use (acres)							
1995	Agriculture	Barren	Forest	Urban	Water	Wetlands	Total
Agriculture	108,962		69	6,024			115,056
Barren		616		159			774
Forest			62,634	24,735			87,369
Urban		19	2	44,123			44,144
Water					6,722		6,722
Wetlands			284	5,131		10,696	16,111
<b>Total</b>	<b>108,962</b>	<b>634</b>	<b>62,989</b>	<b>80,172</b>	<b>6,722</b>	<b>10,696</b>	<b>270,172</b>

Table 5.3 compares size characteristics of parcels predicted to change with actual changed parcels. During the 1995-2004 period, among the parcels predicted to convert to

urban uses, the average sizes for agriculture, barren, forest and wetlands parcels are 4.21, 1.98, 7.48 and 6.94 acres, respectively. The average sizes of the parcels converted to urban uses are very similar except for forest parcels, whose average size decreases by half to 3.54 acres. Among the reforested parcels, the average size for agricultural, urban and wetland parcels are 2.88, 2.14 and 8.12 acres, respectively, during the 1995-2004 period.

Table 5.3 compares the sizes of predicted and actual parcels converting to other uses. The model underpredicts the sizes of those reforested parcels. The actual reforested parcels are generally larger than the predicted parcels with greater standard deviation for both agricultural and wetland parcels. The actual parcels converted to urban use are smaller than those converted parcels predicted by the model. However, the differences between the average size of predicted and actual converting parcels are much smaller and ranges from 0.87 acres for agricultural parcels to 4.74 acres for forest parcels. As for the urban parcels converted to barren lands, the sizes of the actually changed parcels are larger than those predicted by the model. The differences between the actual and predicted changes in term of parcels are hard to be explained fully, which offers the additional opportunity for research and evaluation.

**Table 5.3 The Basic Statistics of the Converted Parcel Size by Land Uses**

<b>Parcel Sizes (acres)</b>										
<b>Change</b>	<b>1995-2004 Predicted Conversions</b>				<b>1995-2002 Actual Conversions</b>				<b>Difference</b>	
	<b>Mean</b>	<b>Std Dev</b>	<b>Max</b>	<b>Min</b>	<b>Mean</b>	<b>Std Dev</b>	<b>Max</b>	<b>Min</b>	<b>Mean</b>	<b>Std Dev</b>
<b>To Forest</b>										
<b>Agriculture</b>	2.88	1.65	5.65	0.36	14.25	27.64	266.19	0.01	-11.37	-25.99
<b>Wetland</b>	8.12	8.08	36.7	1.89	12.49	20.92	147.94	0.66	-4.37	-12.84
<b>To Urban</b>										
<b>Agriculture</b>	4.2	8.73	202.6	0.01	3.33	9.1	238.08	0.05	0.87	-0.37
<b>Barren</b>	1.98	1.63	8.99	0.05	1.03	1.58	15.5	0.02	0.95	0.05
<b>Forest</b>	7.48	19.57	395.78	0.002	2.74	11.01	395.78	0.01	4.74	8.56
<b>Wetland</b>	6.94	15.35	249.62	0.02	2.71	2.52	18.17	0.03	4.23	12.83
<b>To Barren</b>										
<b>Urban</b>	0.49	0.58	1.99	0.02	3.82	0.08	3.92	15.56	-3.33	0.5

## 5.2.2 Validation

Because the actual land use data in 2004 are not available, the model was validated by comparing the predicted land use pattern in 2004 to a reference dataset for 2002. The latter was based on an actual 2002 LULC map from NJDEP. By using the 2002 land use as a reference for comparison, it is assumed that changes within the two-year period are not considerable. An assessment of model prediction capability entails estimating several complementary accuracy metrics. Some of the metrics used in literature are discussed in this section.

### 5.2.2.1 Overall Prediction Accuracy

Table 5.3 presents the contingency table in terms of the number of parcels (the upper panel) and of the area of parcels (the lower panel) using the predicted land use distribution for 2004 and the reference map of 2002 actual land use. As shown in the

upper panel of the table, the overall prediction accuracy in terms of the number of parcels, computed as the sum of the agreements in the diagonals divided by the total number of parcels, is 84.46 percent. The values of the prediction accuracy obtained in this study are similar to the values reported in the literature. Li and Yeh (2004) reported an overall accuracy of 82 percent using a decision tree based-CA for predicting land use change in an urbanizing city in Southern China. Similarly, Lu (2002) using a multinomial logistic model with the parcel as the unit of analysis achieved an overall success rate of 80.76 percent in terms of number of parcels predicted.

Supplementing the information regarding overall accuracy are two types of prediction errors. Errors of omission are shown in the row totals, while errors of commission are shown in the column totals. Omission errors are cases of certain land use categories that were omitted or excluded and misclassified into other categories. Conversely, commission errors are cases that truly belong to one category, but have been mistakenly included in another category. Theoretically, the measures should be equal. Table 5.3 shows that the error of omission and the error of commission are 15.54 percent in terms of the total parcels in the County. Omission error varies across land use categories. Barren land has the largest error of omission at 91.7 percent, followed by wetland at 51.67 percent and water at 44.15 percent. On the other hand, taken all together, all three land use categories account for only 3.23 percent of the total parcels in the county. The major land use categories of urban land and agriculture were the least misclassified with omission error of 6.25 percent and 26.90 percent respectively. Urban lands account for 70.18 percent of the county's total number of parcels while agriculture comprises 8.33 percent. Just as in the error of omission, the error of commission also

varies by land use types. The barren land parcels have the highest error of commission of 95.79 percent. The agricultural parcels have the second highest error of commission of 28.95 percent and water is misclassified at 26.06 percent in terms of the error of commission. The errors of commission for forest and wetlands parcels are almost the same, which is about 16 percent. The urban parcels have the least error of commission at 13.06 percent.

The same method is also used to evaluate the model accuracy in terms of the total acreage. As shown in the lower panel of Table 5.3, the overall prediction accuracy is 80.92 percent. The overall error of omission and/or error of commission is 19.1 percent. Although the error of omission for the barren land is 84.49 percent, it is comprised of the smallest land area in the County at roughly 0.6 percent. On the other hand, agricultural land constituting the largest portion of the County has a low error of omission of 7.89 percent. In terms of acreage, the error of commission also varies by land use types. Errors of commission for barren, agricultural, wetland, and water were 58.45 percent, 8.76 percent, 9.54 percent, and 1.16 percent, respectively. The error of commission for urban land use is 44.78 percent in terms of acreage, which appears much higher than in terms of the number of parcels.

Table 5.4 also shows that agricultural and urban lands are consistently under-predicted, while water and wetlands are over-predicted in terms of both parcel number and acreage.

**Table 5.4 Confusion Matrix for Evaluating Accuracy in terms of Number of Parcels and Parcel Area**

Simulated 2004 Land Use Class (the Number of Parcels)									
2002 Reference Land Use Class	Agriculture	Barren	Forest	Urban	Water	Wetlands	Total Parcels	Misclassified	Error of omission %
Agriculture	3,171		98	1,057		12	4,338	1,167	26.90
Barren	97	15	15	52		2	181	166	91.71
Forest	270	9	5,787	3,385	13	32	9,496	3,709	39.06
Urban	912	332	955	34,241	17	68	36,525	2,284	6.25
Water	4		25	40	105	14	188	83	44.15
Wetland	9		52	611	7	635	1,314	679	51.67
Total Parcels	4,463	356	6,932	39,386	142	763	52,042		
Misclassified	1,292	341	1,145	5,145	37	128		8,088	
Error of commission, %	28.95	95.79	16.52	13.06	26.06	16.78			15.54
Accuracy, %	71.05	4.21	83.48	86.94	73.94	83.22			84.46
Simulated 2004 Land Use Class (the Acreage of Parcels)									
2002 Reference Land Use Class	Agriculture	Barren	Forest	Urban	Water	Wetland	Total Area	Misclassified	Error of omission, %
Agriculture	99,418		1,209	6,986		322	107,935	8,516	7.89
Barren	918	264	323	191		5	1,700	1,436	84.49
Forest	4,795	113	58,344	23,830	44	389	87,515	29,171	33.33
Urban	3,457	258	2,640	44,268	25	210	50,858	6,590	12.96
Water	61		81	84	6,644	95	6,964	321	4.60
Wetland	313		392	4,810	9	9,676	15,200	5,525	36.35
Total Area	108,962	634	62,989	80,169	6,722	10,696	270,172		
Misclassified	9,544	371	4,645	35,901	78	1,020		51,559	
Error of commission, %	8.76	58.45	7.37	44.78	1.16	9.54			19.08
Accuracy, %	91.24	41.55	92.63	55.22	98.84	90.46			80.92

### 5.2.2.2 Kappa Index and Its Variants

The overall prediction accuracy is usually recognized as an overestimate since the method used to calculate it does not account for agreements that would have occurred by chance (Lillesand and Kiefer, 2000). Another way to assess prediction errors is to make use of the Cohen's Kappa index, a measure that takes into account overestimates of the computed percentage of correct values caused by agreements made by pure chance (Pontius and Cheuk, 2006; Foody, 1992; Monserud and Leemans, 1992; Cohen, 1960). This index ranges between 0 and 1 and is interpreted as the proportionate reduction in error achieved by the model being evaluated as compared with the error of a completely random prediction model. Thus a Kappa index of 0.8 would mean that the model is avoiding 80 percent of the errors that a totally random model would generate (Lillesand and Kiefer, 2000).

Equation 2.8 in Chapter Two is used to estimate the Kappa index in terms of the number of parcel numbers. There is no interest in calculating the Kappa index in terms of the acreage since the CA-based land use change model focused on changes of the land use status of individual parcels. In order to calculate the Kappa index using Equation 2.8, the numbers of parcels in the upper panel of Table 5.3 was converted into proportional measures by dividing those numbers by the total number of parcels, (i.e., 52,042). Table 5.5 presents the resulting proportional measures. Using Equation 2.8, the calculated Kappa index is 0.644. This value indicates that the model is able to avoid 64 percent of the errors that a completely random model would generate. The Kappa values ranging from 1 – 0.8 indicate strong agreement, 0.79 – 0.40 moderate agreement, and 0.39 - 0 as

poor agreement (Congalton and Green, 1999; Landis and Kock, 1977). Therefore, the model prediction accuracy is “moderate” based on the calculated Kappa index of 0.644.

As discussed in chapter 2, Cohen’s Kappa index also has limitations in assessing prediction accuracy. As Pontius (2000) pointed out, Cohen’s Kappa index does not give information about location and quantification errors. On the other hand, quantification error occurs when the number of parcels or cells for a given land use type in the predicted map is different from the reference map. Location error occurs when the predicted land use type of a given parcel is different from that in the reference map.

The standard Kappa index,  $K_{\text{standard}}$ , as discussed above does not penalize for large errors nor reward close predictions and is thus an inadequate measure of accuracy (Pontius 2000). As a supplementary measure, two other variants of the Kappa index are presented namely:  $K_{\text{location}}$  for Kappa location, and  $K_{\text{quantity}}$  for Kappa quantity. Table 5.6 lists the formulas for calculating related parameters, the estimated parameter values, and the estimated values for  $K_{\text{location}}$  and  $K_{\text{quantity}}$ . The parameters used in these equations were discussed in Section 2.8 and Section 3.5.

**Table 5.5 Modified Confusion Matrix Expressed in Proportionality Values**

<b>Simulated 2004 Land Use Class (S)</b>							
<b>2002 Reference Land Use Class (R)</b>	<b>Agriculture</b>	<b>Barren</b>	<b>Forest</b>	<b>Urban</b>	<b>Water</b>	<b>Wetland</b>	<b>Total (<math>\Sigma R_j</math>)</b>
<b>Agriculture</b>	0.060931555	0	0.001883094	0.020310518	0	0.000230583	0.083355751
<b>Barren</b>	0.001863879	0.000288229	0.000288229	0.000999193	0	3.84305E-05	0.00347796
<b>Forest</b>	0.005188117	0.000172937	0.111198647	0.065043619	0.000249798	0.000614888	0.182468007
<b>Urban</b>	0.017524307	0.006379463	0.018350563	0.657949349	0.000326659	0.001306637	0.701836978
<b>Water</b>	7.6861E-05	0	0.000480381	0.00076861	0.002017601	0.000269013	0.003612467
<b>Wetlands</b>	0.000172937	0	0.000999193	0.011740517	0.00134507	0.012201683	0.025248837
<b>Total (<math>\Sigma S_j</math>)</b>	0.085757657	0.006840629	0.133200108	0.756811806	0.002728565	0.014661235	1

**Table 5.6 Estimation of Kappa and its Variants<sup>1</sup>**

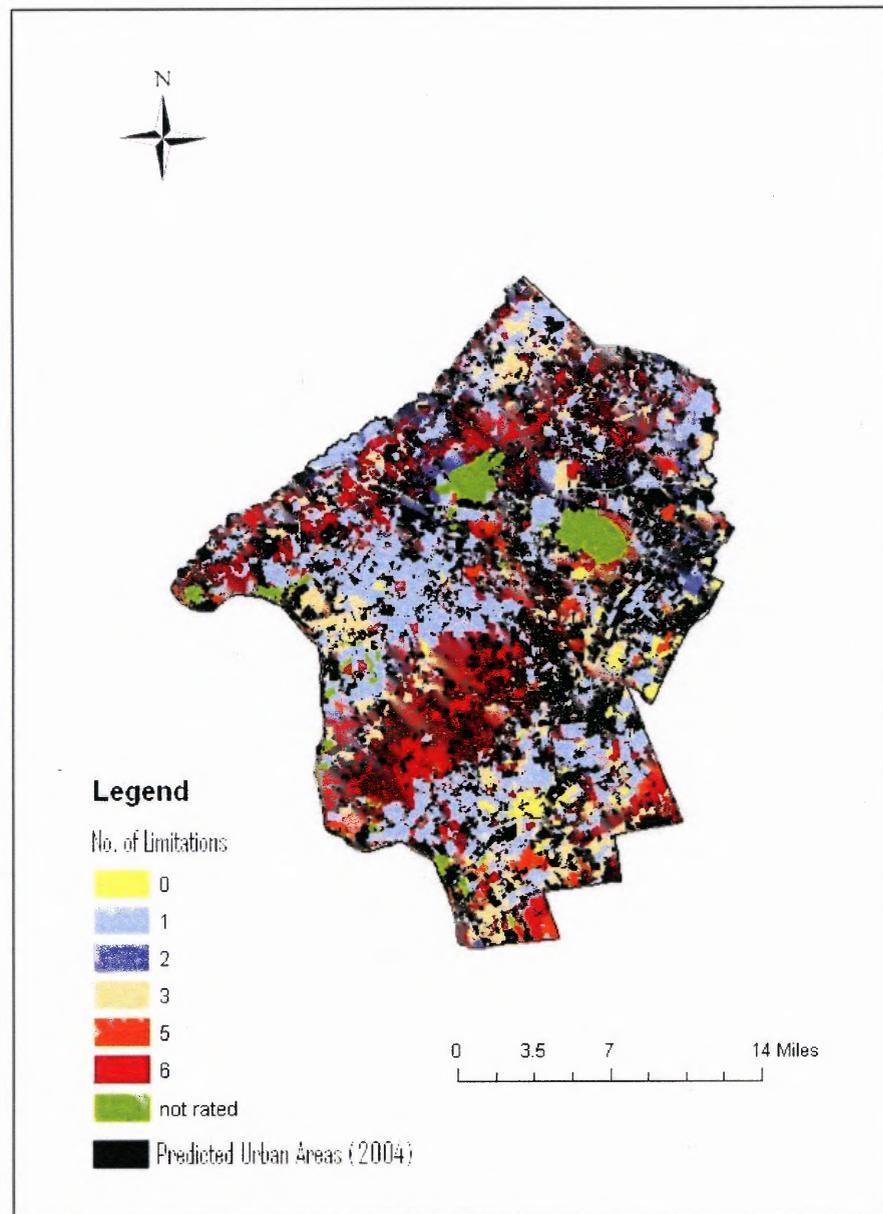
Constants	Estimate	Formula
Po	0.84459	Sum of diagonals in confusion matrix
NQPL	0.44903	$\sum_{j=1}^j \min\left(\left(\frac{1}{j}\right), R_j\right)$
NQML	0.37798	$\left(\frac{1}{j}\right) + K_{location} \left(NQPL - \left(\frac{1}{j}\right)\right)$
MQNL	0.56301	$\sum_{j=1}^j (S_j R_j)$
NQNL	0.16667	$\frac{1}{j}$
MQPL	0.93926	$\sum_{j=1}^j \min(S_j, R_j)$
PQNL	0.53348	$\sum_{j=1}^j (R_j^2)$
PQML	0.88261	$PQNL + K_{location} (1 - PQNL)$
K <sub>standard</sub>	0.644	$\frac{P_o - MQNL}{1 - MQNL}$
K <sub>location</sub>	0.748	$\frac{P_o - MQNL}{MQPL - MQNL}$
K <sub>quantity</sub>	0.925	$\frac{P_o - NQML}{PQML - NQML}$

<sup>1</sup> Adopted from Pontius (2000)

$K_{\text{location}}$  indicates the model's ability to specify location correctly and is equal to 0.748.  $K_{\text{quantity}}$  indicates the model's ability to specify quantity correctly and is equal to 0.925. As stated previously, one of the criticisms of Kappa index is that it does not penalize for large errors nor does it reward estimates having small errors (i.e., close to the true value) (Pontius, 2000; 2005). Put in another way, all errors regardless of their magnitude are all treated equally in calculating Kappa index. The occurrence of either location or quantification error will always subtract or decrease  $K_{\text{standard}}$ . In the case of the current model, location and quantification errors have apparently contributed to a  $K_{\text{standard}} = 0.644$  – only a “moderate” accuracy (Congalton and Green, 1999; Landis and Kock, 1977). However, the model has good quantification ability as indicated by  $K_{\text{quantity}} = 0.925$ . It can be said that a moderate accuracy (based on  $K_{\text{standard}}$ ) is not a complete assessment of the model since it has good ability to predict quantity.

The calculated parameters presented in Table 5.6 can be used to guide potential corrective efforts for improving model prediction accuracy. For instance, PQML - which indicates high ability to specify quantity but medium for location is 88.26 percent. The model achieved an overall correct prediction of 84.45 percent. By focusing on improving the model's ability to specify quantity, the maximum improvement in the overall accuracy would only be 3.8 percent, (i.e, the different between those two measurements). In the same manner, the increase in accuracy can be as high as 9.47 percent by improving the model's ability to specify location because an MQPL is 93.9 percent. Therefore to maximize the percentage of parcels classified correctly, efforts should focus on improving  $K_{\text{location}}$  (i.e., correct specification of location).

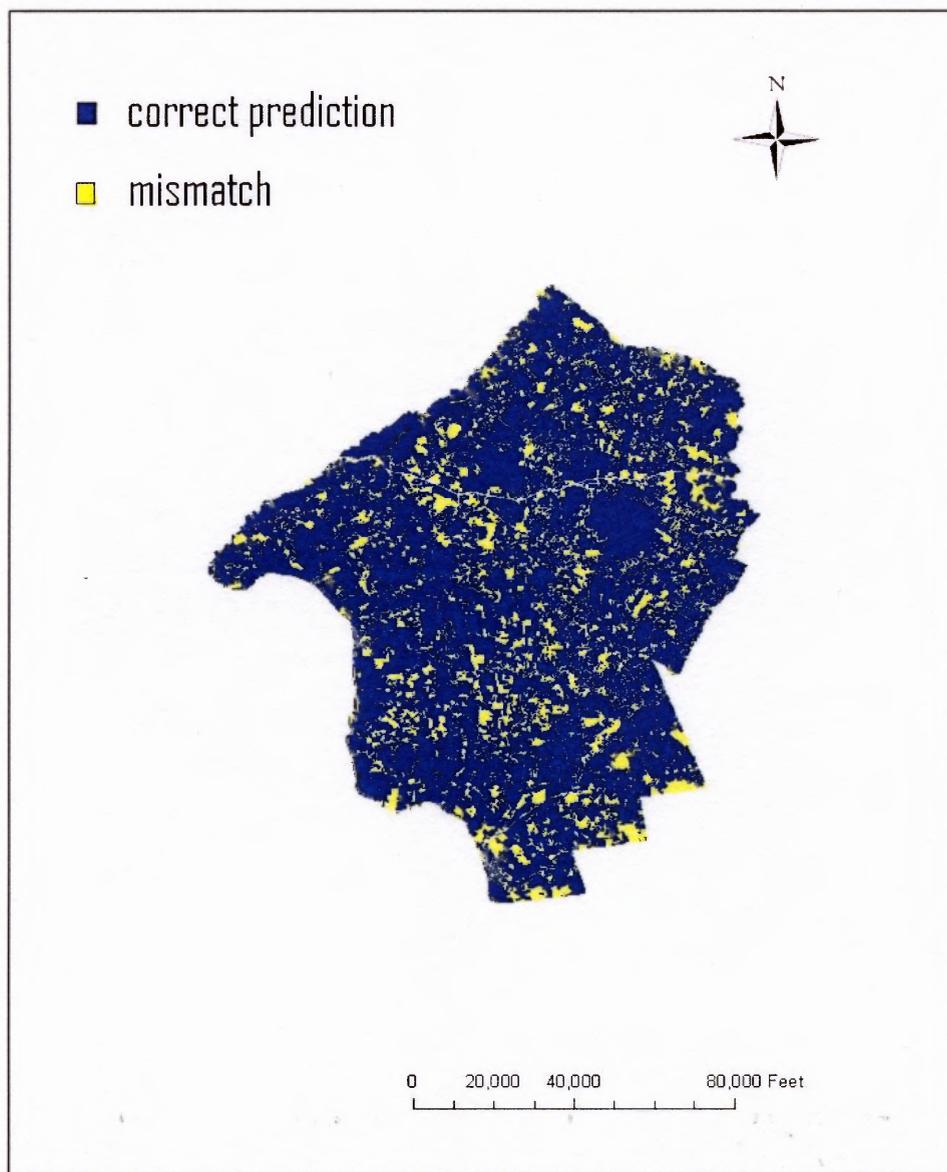
A further validation of the model results is shown in Figure 5.2 which depicts the spatial location of new urban areas versus soil restrictions to urban development as in Section 4.5.1. A cursory survey shows that the model did not predict much urbanization occurring in areas with the largest number of restrictions (highlighted in red). Conversely, growth of new urban areas is seen spread out in places having at most one restriction (blue areas). This observation supports the assumption that least suitable soils require additional treatment thereby increasing development costs. Among other requirements, for land parcels to be eligible for development they should have soils that have good absorption capacities for proper septic tank operation, be able to support load for building foundations, and be relatively level. Absence of any of these requirements increases the development cost.



**Figure 5.2** Spatial distribution of newly urbanized areas during the 1995-2004 period relative to soil suitability (limitations) for development

Finally, Figure 5.3 shows the spatial distribution of parcels that were predicted correctly (shown in green) and incorrectly (in yellow) when comparing the 2004 predicted land use pattern to the 2002 reference level. Inspection of the map shows that the mismatches are

somewhat randomly distributed throughout the county which suggests that systematic errors in the modeling process were somewhat minimal.



**Figure 5.3** Spatial Distribution of misclassified land parcels in the 2004 predicted land use pattern.

### **5.3 Simulation of Land use Change**

As shown in Figure 5.1, two future land use change scenarios are predicted using the calibrated model. They are a baseline scenario and a policy scenario. The baseline scenario assumes “business as usual” and the calibrated transition rules are used to predict the future land use change in 2011 based on the land use pattern in 2002. The policy scenario incorporates government policy interventions such as down-zoning and preservation of farmland and open space by modifying the calibrated transition rules. The policy scenario might more realistically represent future land use changes than the baseline scenario. These policies were adopted and became effective after 2000 in New Jersey (Lathrop, 2002; Richardson, 2007) and the resulting significant changes in land use would have been expected to occur for some years thereafter.

#### **5.3.1 Baseline (Business as Usual) Scenario**

Table 5.7 presents the land use change pattern in terms of the number of parcels from 2002 to 2011 in the baseline scenario. There are 2,361 non-urban parcels converted to urban uses which constitute 95 percent of all converting parcels during the 2002-2011 period. Urbanization came primarily at the expense of agriculture and forest land, each contributing 1,027 parcels and 896 parcels respectively. During the projection period, wetland suffers an additional loss of 119 parcels, among which were 99 converted to urban and 20 to forest land. The total number of forest parcels continues to decline even when 74 parcels become reforested. Among the reforested parcels, 53 come from

agriculture and 20 come from wetlands. There are no other land use parcels converted into agriculture and wetlands.

**Table 5.7 The Predicted Land Use Change Pattern during 2002-2011 in Terms of the Number of Parcels in the Baseline Scenario**

2011 (Number of Parcels)							
2002	Agriculture	Barren	Forest	Urban	Water	Wetlands	Total
Agriculture	3,366	17	53	1,027			4,463
Barren		17		339			356
Forest		2	6,034	896			6,932
Urban		31	1	39,354			39,386
Water					142		142
Wetlands			20	99		644	763
Total	3,366	67	6,108	41,715	142	644	52,042

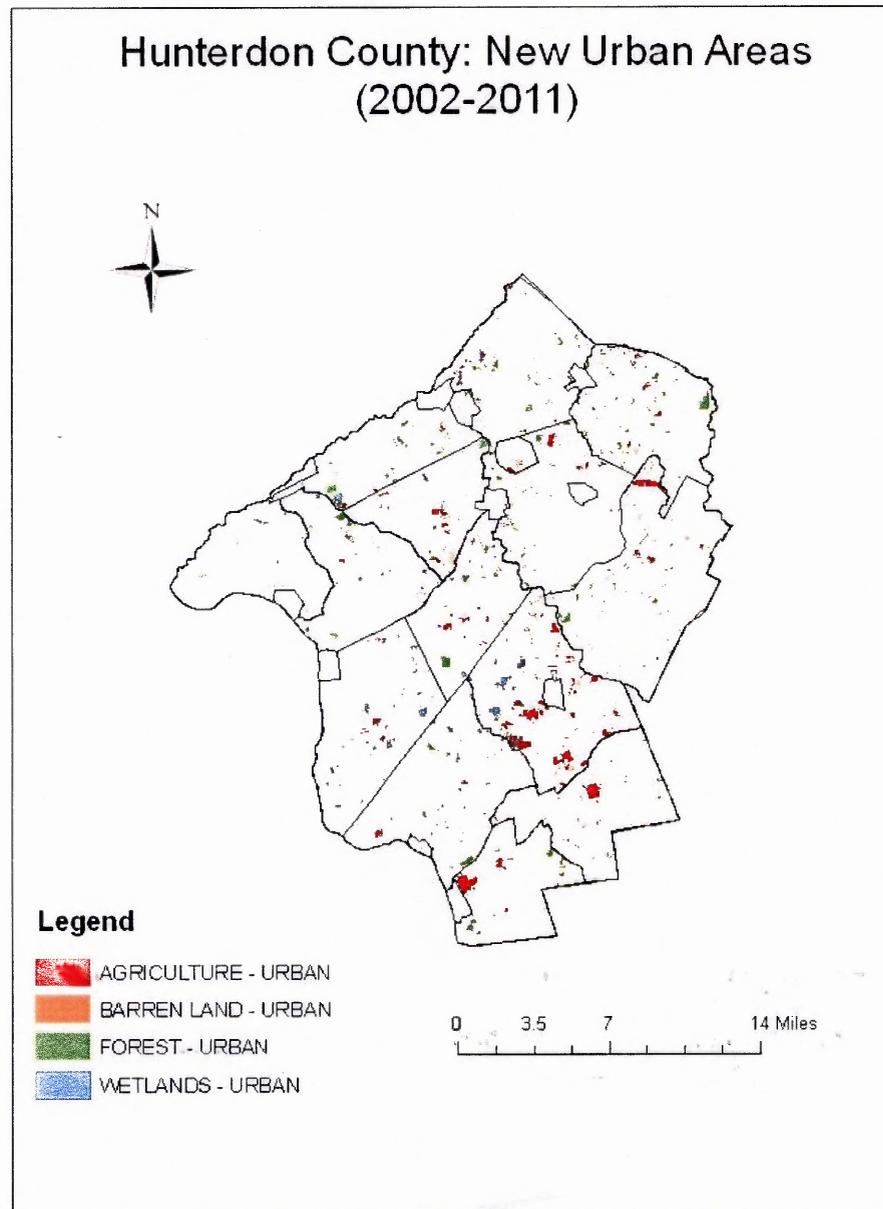
Table 5.8 presents the predicted land use change pattern in terms of total acreage occurring for the 2002-2011 period under the baseline scenario. As the table shows, agriculture, forest, and wetlands contribute 4,204, 3,179, and 749 acres to urban development during this period, respectively. There are also 351 acres of reforested land, among which are 171 acres of previously agricultural land and 178 acres of wetlands.

**Table 5.8 The Predicted Land Use Change Patterns during 2002-2011 in Terms of the Total Acreage in the Baseline Scenario**

2011							
2002	Agriculture	Barren	Forest	Urban	Water	Wetland	Total
Agriculture	104,560	27	171	4,204			108,962
Barren		286		349			634
Forest		5	59,805	3,179			62,989
Urban		17	2	80,153			80,172
Water					6,722		6,722
Wetland			178	749		9,769	10,696
Total	104,560	335	60,156	88,634	6,722	9,769	270,172

Figure 5.4 shows new urban areas built from 2002 - 2011 under the baseline scenario. A marked decrease in urbanization is noted compared to the trends in the previous period (depicted in Figure 5.2). Conversion of agricultural, forest, and wetlands decreased by a

range of 30 to 87 percent indicating that continuing with current land management practices without adopting new restrictions to development, a slowing down of urban spread can still be achieved. Incidentally, although not apparent from Figure 5.4, barren land converting to urban land actually increased by 120 percent. This implies that these abandoned areas are being redeveloped.



**Figure 5.4** Projected new urban areas for 2002-2011 in the baseline scenario.

### 5.3.2 Simulated Policy Scenario

Three land use policies were considered in the policy scenario, namely down-zoning and preservation of farmlands and preservation of open spaces. Down-zoning is reflected in the Water Quality Management Planning Rules implemented in 2001 by NJDEP (2001a). Under the down-zoning rule, the lot size for residential development in non-sewered areas will have to be greater than 1.3 hectares. Conway and Lathrop (2005) reported that the rule was rescinded because of some implementation difficulties. However, it is likely that the rule will be reinstated as a result of public pressure to protect water quality (Springer, 2002; McKeon, 2001).

Down-zoning and the two other land preservation policies were implemented in the model by adopting a spatial constraints approach (Swenson and Franklin 2000; Schneider and Pontius 2001). This approach is used in cellular automata by implementing a set of transition rules wherein parcels that are classified as preserved farmland, or open space. As regards down-zoning, parcels in non-sewered areas that do not conform to development requirements (i.e., greater than 1.3 hectares) are likewise not allowed to convert to urban lands. These transition rules will override all other factors in the land use change modeling process when simulating the policy scenario.

Table 5.9 presents the predicted land use change pattern during the 2002-2011 period under the policy scenario in terms of the number of parcels. Successfully implementing these land use policies stated previously tends to slow down the process of urbanization. Under this policy scenario, only 54 agricultural parcels and 26 forest parcels are converted to urban use. Recall the agricultural and forest land parcels

converted to urban land would be 1,027 and 896 parcels, respectively, in the baseline scenario.

**Table 5.9 The Predicted Land Use Change Pattern Under the Policy Scenario in Terms of the Number of Parcels**

		2011						
2002	Agriculture	Barren	Forest	Urban	Water	Wetland	Total	
<b>Agriculture</b>	4,339	17	53	54			4,463	
<b>Barren</b>		349		7			356	
<b>Forest</b>		2	6,904	26			6,932	
<b>Urban</b>		31	1	39,354			39,386	
<b>Water</b>					142		142	
<b>Wetland</b>			20	2		741	763	
<b>Total Parcels</b>	4,339	399	6,978	39,443	142	741	52,042	

Table 5.10 presents the predicted land use change pattern during this period in terms of total acreage. Compared to the baseline scenario as discussed previously, the policy scenario protects a total of 7,053 acres of non-urban lands from urban development. The protected non-urban lands include 741 acres of wetlands, 3,034 acres of agricultural lands, 250 acres of barren, and 3,028 acres of forest.

**Table 5.10 The Predicted Land Use Change Pattern during 2002-2011 in the Policy Scenario in Terms of Total Acreage**

		2011						
2002	Agriculture	Barren	Forest	Urban	Water	Wetland	Total	
<b>Agriculture</b>	107,594	27	171	1,170			108,962	
<b>Barren</b>		536		99			634	
<b>Forest</b>		5	62,833	151			62,989	
<b>Urban</b>		17	2	80,153			80,172	
<b>Water</b>					6,722		6,722	
<b>Wetland</b>			178	8		10,510	10,696	
<b>Total Area (Acres)</b>	107,594	585	63,184	81,581	6,722	10,510	270,172	

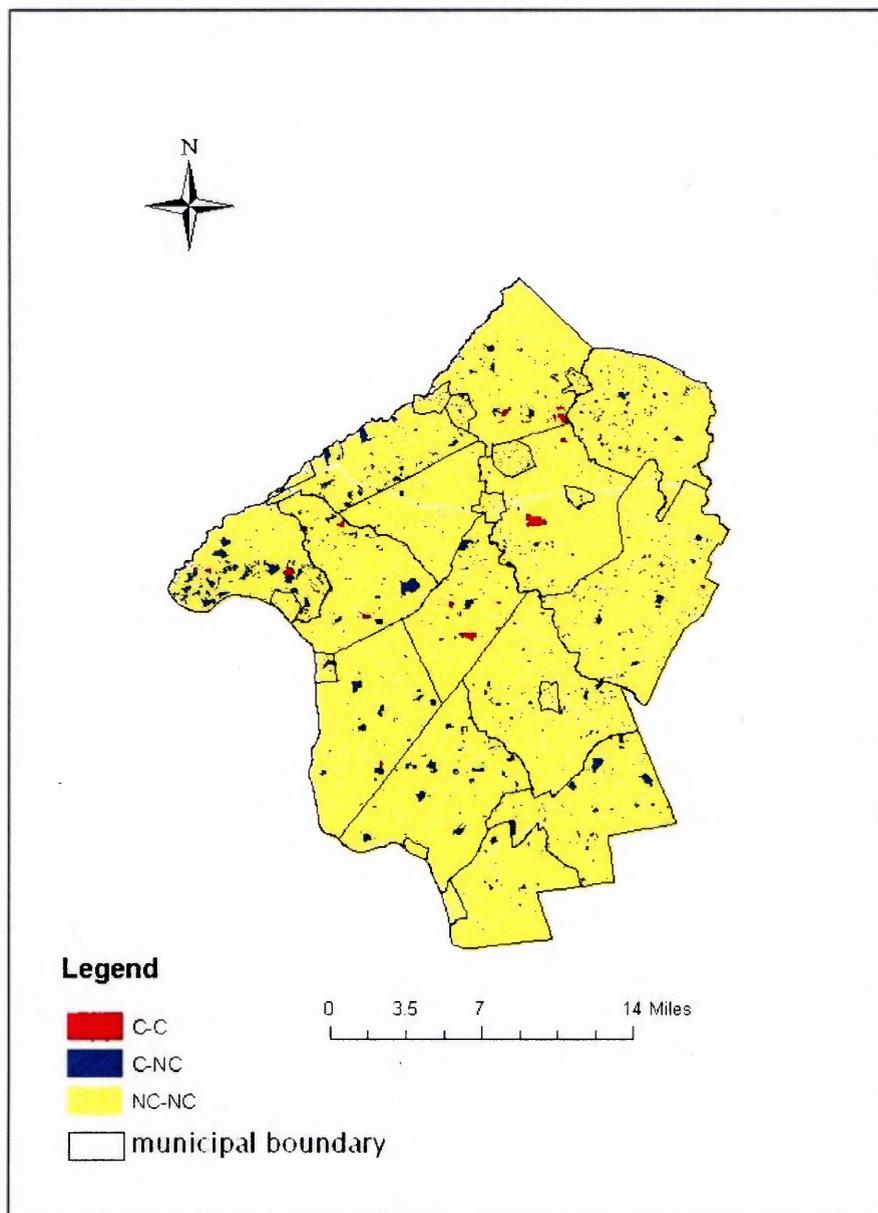
Table 5.11 compares the baseline and policy scenarios in terms of the numbers of the converted and on-converted parcels by land use categories. A total of 1,097 agricultural parcels were converted to other uses in the baseline scenario, whereas only 124 parcels

were converted to urban (54 parcels), forest (53 parcels) and barren land (17 parcels) in the policy scenario. As for forest parcels, the baseline scenario shows 898 parcels converted to other land use types whereas only 28 parcels are converted in the policy scenario.

**Table 5.11 Differential Land Use Matrix Between Baseline and Policy Scenarios**

Agricultural Parcels				
Baseline	Policy Scenario			
	No. of Parcels	Change	No Change	
	Change	124	Barren = 17 Forest = 53 Urban = 54	973
	No Change	0		3,366
Forest Parcels				
Baseline	Policy Scenario			
	No. of Parcels	Change	No Change	
	Change	28	Barren = 2 Urban = 26	870
	No Change	0		6,034
Wetland Parcels				
Baseline	Policy Scenario			
	No. of Parcels	Change	No Change	
	Change	22	Forest = 20 Urban = 2	97
	No Change	0		645
Urban Parcels				
Baseline	Policy Scenario			
	No. of Parcels	Change	No Change	
	Change	32	Barren = 31 Forest = 1	0
	No Change	0		39,372
Barren Parcels				
Baseline	Policy Scenario			
	No. of Parcels	Change	No Change	
	Change	7	Urban = 7	332
	No Change	0		18

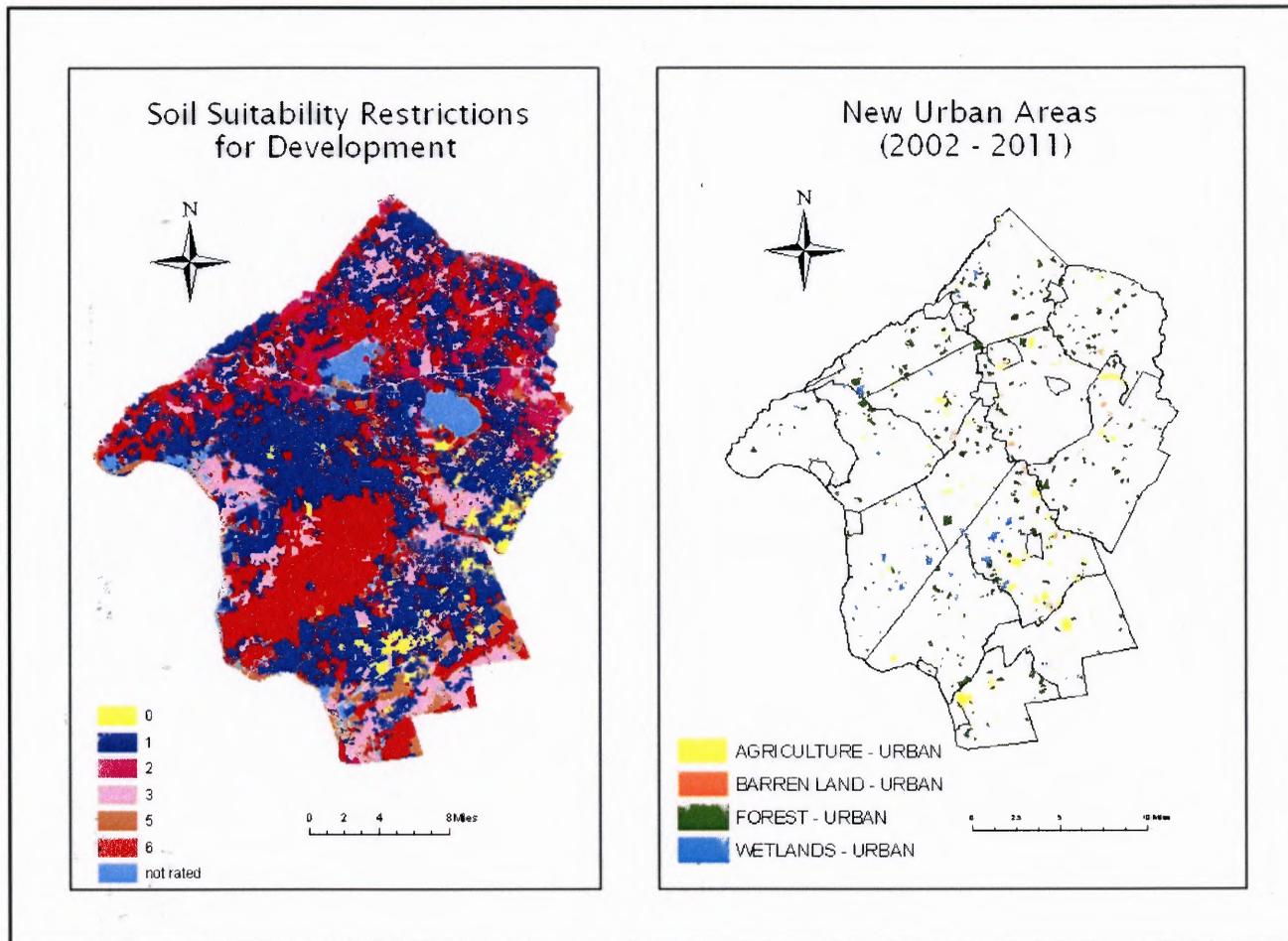
Figure 5.5 presents the spatial distributions of the converted and not converted parcels predicted by both scenarios. There are converted and not converted parcels under two scenarios, which results four possibilities: (a) converted parcels predicted by both scenarios; (b) not converted parcels predicted by both scenarios; (c) converted parcels predicted by the baseline scenario but not converted parcels predicted by the policy scenarios and (d) not converted parcels predicted by the baseline scenario but converted parcels predicted by the policy scenarios. As indicated by Table 5.11, the category (d) does not exist, which leaves the categories. It is not surprising that the majority of parcels are in the category (b), which is shown in yellow in Figure 5.10. The category (c) represents the large preserved parcels by the policy scenario, which are mostly located on the western and southern portion of the county. Holland and Bethlehem Townships in the west have the largest share of preserved lands. Alexandria, Delaware, East Amwell and Kingwood Townships also have large preserved lands but to a lesser degree. In terms of number of parcels, a considerable quantity of preserved parcels is found in Tewksbury and Readington Townships as well as High Bridge, Glen Gardner and Hampton Boroughs. The parcels that belong to the category (a), i.e. the parcels that are converted in both scenarios, are primarily located in Lebanon and Franklin Townships. As indicated by Table 5.11, they are mostly agricultural and forest parcels that are converted to urban use.



**Figure 5.5** Comparative baseline and policy scenario predictions.

C-C: Change predicted by baseline and policy scenarios, C-NC: Change predicted by baseline but not in policy scenario, NC-NC: No changes predicted by both baseline and policy scenarios,

Figure 5.6 presents the spatial distribution of the newly urbanized areas during the prediction period under the policy scenario versus the spatial distribution of soil restrictions to development. As shown in Figure 5.6 most urban expansion would occur in the areas that have fewer restrictions to development. In this scenario, 39.37, 12.5, and 15.9 percent of new urban parcels have one, two and three development restrictions, respectively. NJWSA (2003) considers the land with three development restrictions to be moderate for urban development in a build-out analysis in the Raritan River Basin.



**Figure 5.6** Location of newly urbanized areas in 2011 compared to suitability of soils for development.

#### 5.4 Sensitivity Analysis

Recall that the transition rules are based on the neighborhood size of 475 feet further from the boundary of a parcel. The sensitivity analysis evaluates whether the different neighborhood sizes improve the modeling accuracy. A change in the neighborhood size directly affects the values of the driving factors hence the transition rule and model outcomes. Two neighborhood scenarios considered are 237 feet and 712 feet away from the boundary of a parcel. Table 5.12 presents the contingency table and Kappa indices when comparing the predicted 2004 land use pattern to the 2002 reference land use pattern for each scenario.

There is not much difference in the prediction results of future land use changes using the two neighborhood sizes. A survey of the three land use classes with large areas – agriculture, forest, and urban land indicates that the differences are negligible. The model using the 237-foot neighborhood size predicts 3,217 agricultural parcels in 2004 while the model with the 712-foot buffer neighborhood gave 3,142 agricultural parcels. The difference is only 75 parcels, equivalent to 2.3 percent of total agricultural parcels. Predictions for forest and urban lands differed by 24 (0.41 percent) and 313 (0.92 percent) parcels, respectively.

The Kappa index is 0.64 when comparing the predicted land use pattern using a neighborhood buffer of 237 feet to the 2002 reference land use pattern. The index is 0.65 when comparing the model results using a buffer of 712 feet with the reference map. The relatively small discrepancy in the two values of the Kappa index confirms previous observations indicating that there is no marked difference in model outcomes between the two neighborhood sizes in the application.

Caution is given about generalizing this result to other studies and areas. Some studies showed that the CA model was sensitive to changes in model elements such as the neighborhood size. The land uses in a neighborhood are always significant driving factors for future land use change. Extensive investigation on the impacts of the neighborhood size should be conducted in every new application.

**Table 5.12 The Impacts of Two Neighborhood Sizes on the CA Modeling Accuracy and the Resulting Kappa Index**

Simulated Parcel Count (Neighborhood buffer = 237 feet)							
Actual Parcel Count	Agriculture	Barren	Forest	Urban	Water	Wetlands	Total
<b>Agriculture</b>	3,217		104	1,006		11	4,338
Barren	97	15	17	50		2	181
<b>Forest</b>	297	9	5,785	3,359	13	33	9,496
Urban	1,013	376	982	34,077	17	60	36,525
<b>Water</b>	4		25	39	105	15	188
Wetlands	9		44	644	7	610	1,314
<b>Total</b>	4,637	400	6,957	39,175	142	731	52,042
Overall Accuracy							84.18
<b>Kappa Standard</b>							<b>0.64</b>
Simulated Parcel Count (Neighborhood buffer = 712 feet)							
Actual Parcel Count	Agriculture	Barren	Forest	Urban	Water	Wetlands	Total
<b>Agriculture</b>	3,142		91	1,094		11	4,338
Barren	88	13	15	63		2	181
<b>Forest</b>	256	8	5,809	3,376	13	34	9,496
Urban	804	297	944	34,390	17	73	36,525
<b>Water</b>	3		28	38	105	14	188
Wetlands	10		47	590	7	660	1,314
<b>Total</b>	4,303	318	6,934	39,551	142	794	52,042
Overall Accuracy							84.78
<b>Kappa Standard</b>							<b>0.65</b>

## 5.5 Summary

The straightforward interpretation of transition rules generated through the decision tree method affords better appreciation of the hierarchy of driving factors according to their significance in effecting land use change. It was shown that the two most important

attributes that influence a parcel's behavior are the presence of urban and barren land in neighboring parcels, and parcel area. The other attributes have varying effects on the changes occurring in different land use types.

The transition rules were then considered as input to a cellular automata to model and to predict land use change in Hunterdon County. The results show that new urban development is spread more or less evenly throughout the county. Urban development was made mostly at the expense of forest and agricultural lands. Forest loss was the major urban land contributor during the period 1995 to 2002. Although contributing to a small share of the total demand for urban land, wetlands lost 32 percent of their original land area to urbanization during the same period. Without government intervention, the downward trend of remaining forest, agricultural, and wetland areas will continue. With down-zoning, and farmland and open space preservation policies in place, the loss to urbanization is reduced by some 2,272 non-urban parcels equivalent to 7,053 acres. To give credence to these figures, the prediction accuracy of the model was assessed using several indices suggested in the literature. The model gives high overall accuracy and moderate Kappa indices. Accordingly, the model can be considered as having reasonable predictive capabilities.

## CHAPTER 6

### SUMMARY AND CONCLUSION

This dissertation has sought to develop decision tree based CA model to predict future land use changes with parcel-level data at regional level. It focuses on model development and its application to Hunterdon County in New Jersey. The modeling approach is summarized and the model's findings are discussed in the first two sections of this chapter, followed by contributions of the research to land use-change modeling, possible areas for further investigation, and conclusions.

#### 6.1 Summary of Approach

This research focused largely on developing a decision tree based CA model that would use the parcel as its modeling unit. The model extends the classical raster-based CA model by defining the modeling space as a collection of geographic objects of irregular shape that are spatially represented by the parcel and defining the transition rules using a knowledge discovery algorithm decision tree.

Since parcel is the basic unit where a transition rule is built upon in the CA model, a major portion of this research revolved around preparing the parcel data and generating suitable transition rules. It is a challenging task to prepare the parcel-level data on historical land uses and the driving factors for land use changes required by the model since the boundaries of land use polygons and some driving factors do not correspond to the boundary of the parcels. Intersecting the parcels with land use polygons usually results in multiple land use classes within a parcel. Consequently, an empirical

classification scheme in the study area was created to assign a single land use class to each parcel.

The neighborhood of each parcel was defined as an external buffer along the boundary of the parcel, which is an improvement over the ways of defining the neighborhood in the existing CA-based land use change models that also use parcel as a unit of analysis. A parcel's future status of land use is usually significantly influenced by the land use conditions in those parcels that are partially or completely located within its neighborhood. The land use distribution in this neighborhood is generated by overlapping the neighborhood with the actual land use maps and makes up an set of important driving factors in the input dataset. Other driving factors such as soil suitability, slope, wetland area, and proximity to urban centers, streets and roads, and streams are similarly derived through spatial analysis in GIS.

Historical land uses of parcels and a dataset of driving factors referenced to the parcels are imported into a knowledge-discovery process decision tree to extract or learn patterns how driving factors affect the land use changes in the study area. The learned pattern constitutes the transition rules in a CA model to predict the future land use changes using the current land use and driving factors. This decision tree approach has superior accuracy compared to other transition rule elicitation methods. It is not dependent on expert knowledge in identifying relevant transition-influencing factors. As such, it is free from the subjective biases that are usually associated with other methods. Finally, an assessment approach is developed to evaluate the modeling results. The assessment entails a cell-to-cell comparison approach and uses several accuracy metrics. By itemizing sources of agreement and disagreement between a predicted and a reference

map, possible useful alternatives to improve the model are identified to set the future agenda in parcel-based CA modeling.

## **6.2 Summary of Findings**

The findings of this research are summarized in four parts: (1) reliability of the decision tree-based CA model, (2) effects of driving factors, (3) suitability of the parcel as unit of analysis, and (4) land use outcome of policy implementation. Details of these findings are explained below.

### **6.2.1 Reliability of the Decision Tree-based CA model**

The accuracy measurement of the prediction results show that a decision tree-based cellular automata model can accurately predict land use change using the driving factors listed in this study. Using actual changes during the period 1995-2002 as a reference, an overall accuracy of 84.46 percent in terms of the number of parcels was achieved for the prediction period 1995-2004. In terms of the total area of changed land use, the overall accuracy is 80.92 percent. These numbers are close to the 85 percent requirement suggested in the literature (Foody, 2002). These estimates of accuracy were complemented with other accuracy indicators to provide a more informative assessment. Kappa statistics that indicate the model's ability to correctly predict location and quantity were calculated. They were moderate and high with the overall accuracies of 74.8 and 92.5 percent respectively. Besides the overall accuracy, the prediction accuracy is also calculated in terms of the number of parcels for each land use class. The prediction accuracies ranged from 71.05 to 86.94 percent for all the categories, except barren land.

In terms of the total acreage of the changed land uses, the calculated accuracies range from 55.22 percent to 98.84 percent, with lowest prediction success for urban land.

In addition to these accuracy measurements, the following are also of particular significance:

- 1) The area-based accuracy for all land use classes, with the exception of the urban class, are higher than the accuracy based on the total number of parcels. This observation may be useful in cases where the model is used to evaluate the magnitude of the environmental impacts of land use change. Although this observation still needs further validation, the finding suggests that the prediction results of the CA-based modeling approach are more reliable in terms of the total amounts of changed land uses than the total amount of changed parcels.

- 2) Most errors of omission and commission are associated with barren land, suggesting that the model has difficulty differentiating barren land from other land use classes.

- 3) In terms of both the total number of parcels and land area, the agriculture losses and the increases in urban areas are consistently over-predicted, while the losses in water and wetlands are under-predicted. Such results may be related to the inaccuracy in assigning single land use change to each parcel as described earlier. This finding suggests that overall prediction accuracy may significantly be improved by developing more precise classification scheme that classify a smaller number of predicted urban and agriculture parcels compared to the scheme used in the present study.

### **6.2.2 Effect of Driving Factors**

For parcels that are currently urban, agriculture, forest, and barren lands, the presence of barren and urban land in the neighborhood of the parcels significantly affects their future land use status. Thereafter, the other attributes acting singly or together have varying effects on the changes occurring in different land use types.

In the case of urban lands, their future status is influenced primarily by parcel size and secondarily by land use in their neighboring parcels. Parcels that are smaller than 2.32 acres will likely remain as urban. The future status of a larger urban parcel is subject to the area of wetlands within the parcel and the presence of forest area around the parcel.

In predicting future status of agricultural parcels, the presence of urban land in its neighborhood primarily increases its likelihood of converting to urban lands. Parcels whose neighborhood is more than 43.37 percent urban are likely to transform into urban land. For parcels with less urban land in their neighborhoods, the wetland area within the parcel and the percentage of forest in its neighborhood determine whether they will retain their status as agricultural land or not.

As to forest parcels, the presence of barren and urban lands in their neighborhood is the significant determinant of their future changes. Forest parcels with a high percentage of barren land in their neighborhood are most likely to be converted to barren lands even they are located close to streams or rivers. The presence of high percentage of urban in the neighborhood triggers the conversion to urban lands from forest. This process is to some extent slowed down by the presence of agricultural lands in the neighborhood.

The decision-tree structure has shown that wetlands can remain as wetlands or can be converted into urban lands depending on the amount of wetland area within the parcel. The wetland parcels with large size wetlands (>49,000 square feet) always remain as wetlands. Small wetlands may stay as wetlands but at times can be converted into to urban uses or barren lands.

Barren land can either remain as barren or can be converted to urban land. Interestingly, the presence of barren land in the neighborhood always triggers the conversion of barren land parcels to urban uses.

### **6.2.3 Suitability of the Parcel as Unit of Analysis**

The results of the research indicate that the parcel is an appropriate analysis unit for modeling land use change at the county level. As expected, small spatial units enhance modeling practice by increasing spatial precision. As noted in previous chapters, although spatial representation in raster maps is smaller, it does not represent useful geographical entities related to political, economic and policy changes. Parcels are the smallest legal units of land and are thus more precise in representing land use than other units such as census blocks or tracts. By achieving a high accuracy in predicting land use change, the model was able to demonstrate the feasibility and effectiveness of using parcel-level data in land use change modeling resulting from the political and economic processes and policy changes.

### **6.2.4 Land Use Outcome of Policy Implementation**

The model has shown considerable land preserved from urbanization when down-zoning, farmland and open space preservation policies are in place. In forecasting the results, the

model presupposed that all three policies are fully implemented throughout the simulation period. Full compliance is also assumed. It is recognized that any of these policies may be modified because of political considerations or by changing priorities of the part of stakeholders. If such situations arise, some of the model's parameters should be modified accordingly to accommodate these changes. For one, the stakeholder's priorities can be treated as additional transition rule in the model that can be programmed to override other transition rules. The same applies to political considerations.

In simulating the effects of policy implementation, all three policies were simultaneously considered and universally applied to the whole study area. A particular policy may be more applicable to one area while another policy is better implemented in another area. To have a better assessment of their effectiveness, the policies should be modeled separately and comparison be made to assess their applicability in a given situation.

### **6.3 Contributions of this Research**

This dissertation has demonstrated the benefits of coupling decision tree and cellular automata in modeling land use change. In this way, the strengths of both methods are added and their weaknesses are lessened. A decision tree can elicit land use patterns from a large set of driving factors. This procedure is mainly data driven, which avoids the subjective bias and increase the flexibility of deriving transition rules. In this dissertation, aside from the transition rule generated by a decision tree, policies on farmland and open space preservation and down-zoning were entered as additional transition rules in cellular automata. Through this method, it is possible to assess and to anticipate the impact of

planned policy initiatives. In this sense, the method in this dissertation is an improvement over the previous work of using decision tree in CA-based land use models. For example, Li (2003) relied on the purely data-driven transition rule to predict the future land use change.

From an application perspective, the decision tree-based CA model has significant potential as a major planning tool. The acquisition of knowledge about the complex nature of land use systems will help to identify additional driving factors. This model offers the facility to test the significance of these factors in a straightforward manner. Land managers, planners, and policymakers at the municipal and state levels can modify the model to adapt it to their own requirements. This is especially true when there is a need to investigate and predict the effects of future policies before they are implemented. All that is required is that the data are available.

With continued computerization of cadastral data and rapid advancement of GIS capabilities, more and more parcel maps will become available within the state and across the nation. Consequently, there will be more intensive use of parcel-based land use maps in land use planning and modeling. The usefulness of an explicit and easily understood land use change model will become more apparent as the stakeholders involved in land use planning increase in number and diversity. Many of these individuals may not have a technical background. It is hoped that the model will contribute to a better appreciation of the land use-change modeling process.

#### **6.4 Directions for Further Research**

This dissertation work opens possibilities for future work. Three areas are proposed that are aimed at improving model accuracy. One proposal focuses on an alternative classification scheme for assigning land use classes to each parcel. A second proposal is focused on defining an appropriate neighborhood configuration while the third seeks to enhance the model through the inclusion of more driving factors and other influencing factors of land use change.

An assumption made in this research is that each parcel falls only into one land use class. During the investigation process however, it was observed that large parcels had several land use classes. Barren, forest, wetlands were interspersed within large tracts of agricultural lands. Yet because of the classification method used, the parcel is classified as agriculture, thus resulting in an underestimation of the minor land use classes present in the parcel. To reduce this type of error, large parcels can be split into smaller ones based on land-cover boundaries. Advanced GIS techniques should be able to accomplish this task. A rule of thumb for this type of situation is “the smaller the parcels the smaller the errors will be.” Future work of this kind may build on the classification criteria used which showed a bias for agricultural lands over wetlands. It is assumed that a significant number of wetland parcels were subsumed as agricultural lands and classified as such. It is anticipated that reducing such discrepancies will enhance the modeling process.

Another possibility for further investigation is in the area of neighborhood configuration. The external buffer adopted as a parcel’s neighborhood in this study was defined as the area enclosed by a 475 feet distance measured from the boundary of the

parcel. The sensitivity analysis showed that 213 foot and 725 foot buffer sizes did not result in significant changes in prediction results, possibly because there was not much variability in land use distributions within these neighborhood sizes. Future work in this area can focus on investigating other buffer sizes and can determine which neighborhood size will result in improved predictions. It is possible that the 475 foot external buffer used in this research is not the optimal neighborhood size.

To a certain extent, the model can be enhanced by the inclusion of additional relevant driving factors. While the model development stage tried to include a good number of these factors, more can still be included. In as much as the model was designed to handle large datasets, adding more driving factors should be a straightforward process. Possible additional driving factors can include land use intensity as represented by degree of imperviousness, proximity to airports, wastewater treatment plants and other utilities. Moreover, the influence of political factors and the role of specific stakeholders in shaping the outcome of land use decisions have not been extensively dealt with in this dissertation. One possible alternative is to treat this aspect as an additional transition rule similar to how down-zoning and land preservation policies were processed in the model. Another consideration is the issue of outside economic influences that were assumed as exogenous to the model. To address this issue, outside factors can be inserted in the model structure as a sub-module. This sub-module which is essentially macro in scope consist of regional and global factors that will then be coupled with the present land use-change model in the same way as how constrained cellular automata are developed.

A final consideration is to explore the inclusion of stochastic features in generating transition rules to further enhance the model for urban growth applications. Stochasticity is incorporated in the transition rules using constraints. Constraints can be based on land quality or allowable amount of land consumption. Such models would be useful in sustainability studies particularly in land use management.

### **6.5 Conclusions**

A myriad of complex and dynamic processes affects land use and land-cover change. The need to understand the intricacies of this phenomenon is urgent. The potential large impact on the environment, life, and property of uncontrolled land use is documented and has stimulated research across all disciplines and even geographical boundaries. This is evident in the efforts of the International Geosphere–Biosphere Programme (IBGP) and the International Human Dimensions Programme on Global Environmental Change (IHDP) to adopt land use/cover as one of its chief research areas. The first step to understand a complex phenomenon such as land use change and being able to deal with it practically (such as policy making and land management for sustainable land use) is to simplify the processes and relationships to manageable and understandable dimensions. This task is carried out through the use of models. Land use-change models have been extensively used in various applications through the years. Because models can describe the spatial and temporal relationships between the driving factors and the resulting patterns of land uses and their changes, they have been used to provide decision support in various decision and policy-making contexts. Frequently, models have been used to predict future land use patterns under various scenarios of biophysical and socioeconomic

change. The motivation for this effort is that models of land use change can be used to assess the impact of past or future anthropogenic activities in the environmental and/or the socioeconomic spheres. These assessments can focus on qualitative and/or quantitative changes of land use caused by planned changes in one or more of its determinants as in policy making, or on the environmental and socio-economic impacts of changes in land use itself. Some examples of effects of changes in land use are land degradation, desertification, health and safety hazards.

This research contributed in developing simple models for land use planners. The results of the study as applied to Hunterdon County indicate that the model can provide fairly good predictions for land use change. The spatial accuracy for five land use classes was higher on an area basis (98.84 percent) than on the basis of correctly predicted parcels (87.94 percent). This outcome suggests that the decision tree-based CA model can be useful not only in land use planning but also in evaluating the magnitude of environmental impacts of land use. With the exception of barren land, the model is particularly good at predicting all land use classes giving success rates ranging from 71.05 percent to 87.94 percent in terms of the number of converted parcels and 55.22 percent to 98.84 percent in terms of the number of converted area, respectively. By achieving such success rates, the model was able to support the hypothesis that the parcel is a suitable spatial unit in land use change modeling.

The model may have limitations as expected from any endeavor to represent complex realities in simplified forms. Limitations and new challenges will always be present as technology advances and the tools for better mapping techniques become more available. At the same time, people are facing new and competing pressures that make

them change how they use and manage land --- from demographic change to climate change. If we are to avoid the unfortunate consequences of unregulated land use, we can not recoil from our duty to address the seemingly formidable challenges that lay ahead. To this end, there should be a continuous search for ways to improve land-use prediction. One option to consider is to come up with a more comprehensive land use-change theory that fuses both social and natural science perspectives. Taken separately, these theories do not consider all the relevant factors; in the few instances that they do, there is frequently inadequate consideration of the importance of each factor owing to biases in the investigator's training. For example, the urban economic models are almost exclusively concerned with the economic determinants of land use change while the sociological models with the social factors of land use change. An integrated approach would be able to provide better explanation of the effects of a variety of driving factors on patterns of land use change.

Many issues about spatial scale remain. Driving factors and processes operate at various spatial levels within the same or different time scales. For example, the influence of climatic factors on land use change involves a time frame of centuries at regional scales, while policy changes are discernible at time frames of years at lower spatial levels. There are no theories yet that account for this spatio-temporal complexity (Briassoulis 2001). As long as we aspire for a better quality of life, we will always need development. However, urban development need not necessarily be at the expense of the environment. For this to be so, we need to have better informed systems of planning and managing land use. Much of this depends on how far we can uncomplicate the complex reality of land use change.

## APPENDIX A

### SCRIPT FOR CALCULATING NEIGHBORHOOD LAND USE DISTRIBUTION

(Written in ArcView 3.x Avenue scripting language)

```
theProject = av.GetProject
theView = av.GetActiveDoc

theParcelTheme = theView.FindTheme("Parcels.shp")
theParcelFtab = theParcelTheme.GetFtab
fldParcelShape = theParcelFtab.FindField("Shape")
fldParcelID = theParcelFtab.FindField("Objectid")

theLUTheme = theView.FindTheme("Landuse1995.shp")
theLUFtab = theLUTheme.GetFtab
fldLUShape = theLUFtab.FindField("Shape")
fldLUTypeInitial = theLUFtab.FindField("LUType")

aDirName = "C:\Temp".AsFileName
aFileName = aDirName.MakeTmp( "parcel", "dbf" )
aOutputVtab = VTab.MakeNew (aFileName, dBASE)
aOutputTable = Table.Make(aOutputVtab)
aOutputVtab.StartEditingWithRecovery
fldParcelArea = Field.Make ("AreaPar", #FIELD_DECIMAL, 16, 0)
fldDonutArea = Field.Make ("AreaDon", #FIELD_DECIMAL, 16, 0)
aNewParcelIDField = Field.Make ("ParcelID", #FIELD_DECIMAL, 16, 0)
aOutputVtab.BeginTransaction
aOutputVtab.AddFields({ aNewParcelIDField, fldParcelArea, fldDonutArea })
aOutputVtab.EndTransaction

theLUTypeVtab = theProject.FindDoc("landusetypes.dbf").GetVtab
fldLUType = theLUTypeVtab.FindField("Lutype")
lstLUType = {}
lstLUTypeFields = {}

for each i in theLUTypeVtab
  strLUType = theLUTypeVtab.ReturnValue(fldLUType, i)
  aNewLUField = Field.Make (strLUType, #FIELD_DECIMAL, 16, 0)
  aNewPCTLUField = Field.Make ("Pct"+strLUType, #FIELD_DECIMAL, 16, 3)
  aOutputVtab.BeginTransaction
  aOutputVtab.AddFields({ aNewLUField, aNewPCTLUField })
  aOutputVtab.EndTransaction
end
```

```

for each i in theParcelFtab.GetSelection
  aParcel = theParcelFtab.ReturnValue(fldParcelShape, i)
  aParcelID = theParcelFtab.ReturnValue(fldParcelID, i)
  aParcelBuffered = aParcel.ReturnBuffered (475)
  aParcelDonut = aParcelBuffered.ReturnMerged (aParcel)
  'aGraphicShape = GraphicShape.Make (aParcelDonut)
  'aGraphicShape.SetDisplay( theView.GetDisplay)
  'aGraphicShape.Draw
  'aGraphicShape.Invalidate
  'aGraphicShape.IsVisible
  aOutputVtab.BeginTransaction
  aNewRec = aOutputVtab.AddRecord
  aOutputVtab.SetValue(aNewParcelIDField,aNewRec,aParcelID)
  aOutputVtab.SetValue(fldParcelArea,aNewRec,aParcel.ReturnArea)
  aOutputVtab.SetValue(fldDonutArea,aNewRec,aParcelDonut.ReturnArea)
  aOutputVtab.EndTransaction
  theLUFtab.SelectByPolygon (aParcelDonut, #VTAB_SELTYPE_NEW)
  for each j in theLUFtab.GetSelection
    aLUShape = theLUFtab.ReturnValue(fldLUShape, j)
    aLUType = theLUFtab.ReturnValue(fldLUTypeInitial, j)
    aLUIntersection = aLUShape.ReturnIntersection (aParcelDonut)
    'msgbox.info(aLUIntersection.ReturnArea.AsString ++ aLUType, "")
    aPrevArea = aOutputVtab.ReturnValue(aOutputVtab.FindField(aLUType),aNewRec)
    aOutputVtab.BeginTransaction
    aOutputVtab.SetValue(aOutputVtab.FindField(aLUType), aNewRec,
aPrevArea+aLUIntersection.ReturnArea)
    aOutputVtab.SetValue(aOutputVtab.FindField("Pct"+aLUType), aNewRec,
(aPrevArea+aLUIntersection.ReturnArea)/aParcelDonut.ReturnArea*100)
    aOutputVtab.EndTransaction
  end
end

aOutputVtab.StopEditingWithRecovery(TRUE)

```

## APPENDIX B

### DRIVERS OF LAND USE CHANGE

<b>Author</b>	<b>Variables Used</b>	<b>Modeling Method</b>	<b>Unit of Analysis</b>
Stevens et al. (2007)	distance to parks	Cellular Automata	parcel
	distance to commercial areas		
	distance to light industrial land		
	distance to heavy industrial land		
	adjacency to existing developed and undeveloped land		
Li and Yeh (2004)	urban conversion	Cellular automata	parcel
	distance to city proper		
	distance to town centers		
	distance to roads		
	distance to expressways		
	distance to railways		
	number of developed cells in the neighborhood		
	current land use		
agricultural suitability			
	slope		
Allen & Lu, 2001	distance to tourist attraction features	Logistic regression	parcel
	distance to roads		
	distance to sewer line		
	distance to central business district		
	distance to nearest neighborhood		
	Elevation		
	Slope		
	parcel size		
	parcel ownership		
	Drainage		
	Policy constraints: Protected land		
Policy constraint: Residential zone			
Policy constraint: Commercial zone			
Policy constraint: Subdivision			

Author	Variables Used	Modeling Method	Unit of Analysis
	Policy constraint: Urban boundary Population density		
	Housing unit density Housing unit value		
	Initial land use		
Waddell, 2000	current development in neighbor policy constraint: zoning regulations land and improvement values distance to highways distance to existing development regional accessibility to population	Logistic regression	parcel
Moreno and Marceau, 2006	area of parcel distance to adjacent polygon transformation probability to neighbor land use type based on area	Cellular automata	parcel

## APPENDIX C

### J48 PRUNED DECISION TREE

```
CURRLANDTYPE = URBAN
| GISACRES <= 2.32482
| | PCTBARN <= 32.366
| | | PCTURB <= 50.782
| | | | PCTFOR <= 32.492
| | | | | WETLANDAREA <= 1333.9717
| | | | | | SLOPE <= 8.2215
| | | | | | | PCTFOR <= 15.789: URBAN (809.0/37.0)
| | | | | | | PCTFOR > 15.789
| | | | | | | | SOILTYPE = 0: URBAN (50.0/4.0)
| | | | | | | | SOILTYPE = 1: URBAN (326.0/15.0)
| | | | | | | | SOILTYPE = 2: URBAN (73.0/3.0)
| | | | | | | | SOILTYPE = 3: URBAN (75.0/4.0)
| | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | SOILTYPE = 5
| | | | | | | | PCTWATR <= 36.025: URBAN (76.0/4.0)
| | | | | | | | PCTWATR > 36.025: FOREST (2.0)
| | | | | | | | SOILTYPE = 6: URBAN (115.0/10.0)
| | | | | | | SLOPE > 8.2215
| | | | | | | | GISACRES <= 2.13293: URBAN (87.0/10.0)
| | | | | | | | GISACRES > 2.13293: FOREST (4.0)
| | | | | | | | WETLANDAREA > 1333.9717: URBAN (249.0/31.0)
| | | | | | | | PCTFOR > 32.492
| | | | | | | | WETLANDAREA <= 10256.44142
| | | | | | | | PCTURB <= 13.21
| | | | | | | | PCTBARN <= 3.557
| | | | | | | | PCTWETL <= 0.08: URBAN (31.0/16.0)
| | | | | | | | PCTWETL > 0.08
| | | | | | | | | SOILTYPE = 0: URBAN (0.0)
| | | | | | | | | SOILTYPE = 1
| | | | | | | | | PCTWETL <= 13.579: URBAN (14.0)
| | | | | | | | | PCTWETL > 13.579: FOREST (3.0/1.0)
| | | | | | | | | SOILTYPE = 2: URBAN (3.0)
| | | | | | | | | SOILTYPE = 3: URBAN (6.0/1.0)
| | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | SOILTYPE = 5: URBAN (2.0)
| | | | | | | | | SOILTYPE = 6: URBAN (14.0/3.0)
| | | | | | | | | PCTBARN > 3.557: URBAN (2.0/1.0)
| | | | | | | | | PCTURB > 13.21
| | | | | | | | | GISACRES <= 1.77651
| | | | | | | | | DIST2STREAM <= 73.93447: URBAN (84.0/20.0)
| | | | | | | | | DIST2STREAM > 73.93447
| | | | | | | | | | SOILTYPE = 0: URBAN (93.0/11.0)
| | | | | | | | | | SOILTYPE = 1: URBAN (425.0/22.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (131.0/6.0)
| | | | | | | | | | SOILTYPE = 3: URBAN (293.0/29.0)
| | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | PCTWATR <= 31.044: URBAN (118.0/6.0)
```

```

| | | | | | | | | | PCTWATR > 31.044: FOREST (2.0)
| | | | | | | | | | SOILTYPE = 6
| | | | | | | | | | SLOPE <= 19: URBAN (405.0/30.0)
| | | | | | | | | | SLOPE > 19: FOREST (8.0/3.0)
| | | | | | | | | | GISACRES > 1.77651
| | | | | | | | | | PCTBARN <= 0.418
| | | | | | | | | | SOILTYPE = 0: URBAN (12.0/4.0)
| | | | | | | | | | SOILTYPE = 1: URBAN (133.0/24.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (23.0/2.0)
| | | | | | | | | | SOILTYPE = 3: URBAN (54.0/16.0)
| | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 5: URBAN (24.0/10.0)
| | | | | | | | | | SOILTYPE = 6
| | | | | | | | | | WETLANDAREA <= 3048.37106
| | | | | | | | | | PCTWETL <= 4.444
| | | | | | | | | | PCTWETL <= 2.278: URBAN (46.0/9.0)
| | | | | | | | | | PCTWETL > 2.278: FOREST (2.0)
| | | | | | | | | | PCTWETL > 4.444: URBAN (17.0)
| | | | | | | | | | WETLANDAREA > 3048.37106: URBAN (2.0/1.0)
| | | | | | | | | | PCTBARN > 0.418: URBAN (21.0/4.0)
| | | | | | | | | | WETLANDAREA > 10256.44142: URBAN (56.0/6.0)
| | | | | PCTURB > 50.782
| | | | | DIST2STREAM <= 105.66225
| | | | | PCTWATR <= 3.726
| | | | | PCTFOR <= 11.589: URBAN (128.0/4.0)
| | | | | PCTFOR > 11.589
| | | | | WETLANDAREA <= 9247.16022: URBAN (197.0/31.0)
| | | | | WETLANDAREA > 9247.16022
| | | | | SOILTYPE = 0: FOREST (2.0/1.0)
| | | | | SOILTYPE = 1: URBAN (8.0/1.0)
| | | | | SOILTYPE = 2: URBAN (0.0)
| | | | | SOILTYPE = 3: URBAN (1.0)
| | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | SOILTYPE = 5: URBAN (1.0)
| | | | | SOILTYPE = 6: URBAN (5.0/3.0)
| | | | | PCTWATR > 3.726: URBAN (109.0)
| | | | | DIST2STREAM > 105.66225
| | | | | PCTFOR <= 9.826: URBAN (6006.0/16.0)
| | | | | PCTFOR > 9.826
| | | | | SLOPE <= 17
| | | | | GISACRES <= 0.06614: URBAN (807.0/2.0)
| | | | | GISACRES > 0.06614
| | | | | PCTAGRI <= 9.745
| | | | | GISACRES <= 2.09973
| | | | | WETLANDAREA <= 5473.23491
| | | | | SOILTYPE = 0: URBAN (161.0/5.0)
| | | | | SOILTYPE = 1: URBAN (1556.0/38.0)
| | | | | SOILTYPE = 2
| | | | | PCTBARN <= 0.012: URBAN (514.0/15.0)
| | | | | PCTBARN > 0.012
| | | | | SLOPE <= 8.28114: URBAN (6.0)
| | | | | SLOPE > 8.28114: FOREST (3.0/1.0)
| | | | | SOILTYPE = 3: URBAN (511.0/12.0)
| | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | SOILTYPE = 5: URBAN (443.0/23.0)
| | | | | SOILTYPE = 6: URBAN (641.0/26.0)

```

```

| | | | | | | | | | WETLANDAREA > 5473.23491: URBAN (67.0/1.0)
| | | | | | | | | | GISACRES > 2.09973: URBAN (95.0/16.0)
| | | | | | | | | | PCTAGRI > 9.745: URBAN (766.0/20.0)
| | | | | | | | | | SLOPE > 17
| | | | | | | | | | DIST2HW <= 113: URBAN (21.0)
| | | | | | | | | | DIST2HW > 113
| | | | | | | | | | SOILTYPE = 0: FOREST (4.0/1.0)
| | | | | | | | | | SOILTYPE = 1: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 2: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 3: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 6
| | | | | | | | | | PCTWATR <= 0.078: URBAN (3.0)
| | | | | | | | | | PCTWATR > 0.078
| | | | | | | | | | PCTFOR <= 33.197: FOREST (6.0)
| | | | | | | | | | PCTFOR > 33.197: URBAN (2.0)
| | | | | | | | | | PCTBARN > 32.366: BARREN LAND (25.0/2.0)
| | | | | | | | | | GISACRES > 2.32482
| | | | | | | | | | WETLANDAREA <= 34804.11083
| | | | | | | | | | PCTFOR <= 45.031
| | | | | | | | | | GISACRES <= 3.81563: URBAN (1211.0/205.0)
| | | | | | | | | | GISACRES > 3.81563
| | | | | | | | | | PCTAGRI <= 33.897
| | | | | | | | | | PCTFOR <= 18.192: URBAN (140.0/16.0)
| | | | | | | | | | PCTFOR > 18.192
| | | | | | | | | | PCTBARN <= 1.698
| | | | | | | | | | PCTURB <= 62.839
| | | | | | | | | | SOILTYPE = 0: URBAN (21.0/7.0)
| | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | PCTBARN <= 0.055
| | | | | | | | | | DIST2HW <= 246: URBAN (37.0/10.0)
| | | | | | | | | | DIST2HW > 246
| | | | | | | | | | PCTWETL <= 0.765: URBAN (31.0/9.0)
| | | | | | | | | | PCTWETL > 0.765
| | | | | | | | | | DIST2HW <= 440
| | | | | | | | | | PCTWATR <= 1.2
| | | | | | | | | | DIST2HW <= 400
| | | | | | | | | | PCTURB <= 42.698: FOREST (10.0/4.0)
| | | | | | | | | | PCTURB > 42.698: URBAN (13.0/2.0)
| | | | | | | | | | DIST2HW > 400: AGRICULTURE (6.0/2.0)
| | | | | | | | | | PCTWATR > 1.2: URBAN (3.0)
| | | | | | | | | | DIST2HW > 440: FOREST (12.0/4.0)
| | | | | | | | | | PCTBARN > 0.055: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (21.0/9.0)
| | | | | | | | | | SOILTYPE = 3: URBAN (35.0/18.0)
| | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | DIST2UC <= 17690.63559: URBAN (13.0/2.0)
| | | | | | | | | | DIST2UC > 17690.63559: FOREST (5.0/1.0)
| | | | | | | | | | SOILTYPE = 6
| | | | | | | | | | DIST2UC <= 8250.50094
| | | | | | | | | | PCTWETL <= 1.342: URBAN (4.0)
| | | | | | | | | | PCTWETL > 1.342: FOREST (13.0/5.0)
| | | | | | | | | | DIST2UC > 8250.50094: URBAN (36.0/17.0)
| | | | | | | | | | PCTURB > 62.839: URBAN (40.0/4.0)

```

| | | | | | PCTBARN > 1.698: URBAN (16.0)  
 | | | | | PCTAGRI > 33.897  
 | | | | | | SLOPE <= 5.29694  
 | | | | | | | SOILTYPE = 0: URBAN (14.0/7.0)  
 | | | | | | | SOILTYPE = 1: URBAN (123.0/62.0)  
 | | | | | | | SOILTYPE = 2  
 | | | | | | | GISACRES <= 7.69757  
 | | | | | | | | DIST2HW <= 247: URBAN (2.0)  
 | | | | | | | | DIST2HW > 247: AGRICULTURE (7.0)  
 | | | | | | | | GISACRES > 7.69757: URBAN (4.0)  
 | | | | | | | | SOILTYPE = 3: URBAN (23.0/12.0)  
 | | | | | | | | SOILTYPE = 4: URBAN (0.0)  
 | | | | | | | | SOILTYPE = 5: URBAN (7.0/4.0)  
 | | | | | | | | SOILTYPE = 6: URBAN (25.0/10.0)  
 | | | | | | | SLOPE > 5.29694: URBAN (39.0/16.0)  
 | | | | PCTFOR > 45.031  
 | | | | | GISACRES <= 4.80418  
 | | | | | | PCTWATR <= 0.026  
 | | | | | | | SOILTYPE = 0: URBAN (21.0/10.0)  
 | | | | | | | SOILTYPE = 1  
 | | | | | | | | PCTURB <= 32.332: FOREST (87.0/43.0)  
 | | | | | | | | PCTURB > 32.332: URBAN (85.0/24.0)  
 | | | | | | | | SOILTYPE = 2: URBAN (26.0/11.0)  
 | | | | | | | | SOILTYPE = 3  
 | | | | | | | | PCTAGRI <= 22.463  
 | | | | | | | | SLOPE <= 6.61882  
 | | | | | | | | | GISACRES <= 3.02281  
 | | | | | | | | | | DIST2UC <= 17280.02338: URBAN (20.0)  
 | | | | | | | | | | DIST2UC > 17280.02338  
 | | | | | | | | | | PCTWETL <= 3.432: FOREST (4.0/1.0)  
 | | | | | | | | | | PCTWETL > 3.432: URBAN (5.0)  
 | | | | | | | | | | GISACRES > 3.02281  
 | | | | | | | | | | PCTWETL <= 0.521: URBAN (10.0/2.0)  
 | | | | | | | | | | PCTWETL > 0.521: FOREST (14.0/4.0)  
 | | | | | | | | | | SLOPE > 6.61882: FOREST (31.0/11.0)  
 | | | | | | | | | | PCTAGRI > 22.463: URBAN (3.0/1.0)  
 | | | | | | | | | | SOILTYPE = 4: URBAN (0.0)  
 | | | | | | | | | | SOILTYPE = 5: URBAN (15.0/6.0)  
 | | | | | | | | | | SOILTYPE = 6  
 | | | | | | | | | | SLOPE <= 6.85472: URBAN (57.0/20.0)  
 | | | | | | | | | | SLOPE > 6.85472: FOREST (92.0/33.0)  
 | | | | | | | | | | PCTWATR > 0.026: URBAN (41.0/7.0)  
 | | | | | | | | | | GISACRES > 4.80418  
 | | | | | | | | | | PCTBARN <= 0.064  
 | | | | | | | | | | SLOPE <= 9.08501  
 | | | | | | | | | | PCTFOR <= 47.001: URBAN (22.0/4.0)  
 | | | | | | | | | | PCTFOR > 47.001  
 | | | | | | | | | | | SOILTYPE = 0: URBAN (4.0/2.0)  
 | | | | | | | | | | | SOILTYPE = 1  
 | | | | | | | | | | | | PCTWATR <= 0.411  
 | | | | | | | | | | | | DIST2HW <= 505  
 | | | | | | | | | | | | DIST2UC <= 14240.19296  
 | | | | | | | | | | | | DIST2STREAM <= 1320.60388  
 | | | | | | | | | | | | | PCTAGRI <= 13.957: FOREST (9.0)  
 | | | | | | | | | | | | | PCTAGRI > 13.957: URBAN (3.0/1.0)  
 | | | | | | | | | | | | | DIST2STREAM > 1320.60388: URBAN (16.0/4.0)

```

| | | | | | | | | | DIST2UC > 14240.19296
| | | | | | | | | | DIST2UC <= 15222.04438: URBAN (2.0/1.0)
| | | | | | | | | | DIST2UC > 15222.04438
| | | | | | | | | | PCTAGRI <= 13.017: FOREST (4.0)
| | | | | | | | | | PCTAGRI > 13.017: AGRICULTURE (2.0/1.0)
| | | | | | | | | | DIST2HW > 505: URBAN (8.0/3.0)
| | | | | | | | | | PCTWATR > 0.411: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (14.0/8.0)
| | | | | | | | | | SOILTYPE = 3
| | | | | | | | | | WETLANDAREA <= 187.51091
| | | | | | | | | | PCTWATR <= 0.141
| | | | | | | | | | SLOPE <= 3.59657: URBAN (5.0/2.0)
| | | | | | | | | | SLOPE > 3.59657: FOREST (14.0/2.0)
| | | | | | | | | | PCTWATR > 0.141: FOREST (2.0)
| | | | | | | | | | WETLANDAREA > 187.51091: URBAN (3.0/1.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | WETLANDAREA <= 5618.80592: URBAN (6.0/2.0)
| | | | | | | | | | WETLANDAREA > 5618.80592: FOREST (2.0)
| | | | | | | | | | SOILTYPE = 6: FOREST (39.0/20.0)
| | | | | | | | | | SLOPE > 9.08501
| | | | | | | | | | SOILTYPE = 0: FOREST (2.0)
| | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | PCTFOR <= 69.704: FOREST (4.0)
| | | | | | | | | | PCTFOR > 69.704: URBAN (2.0/1.0)
| | | | | | | | | | SOILTYPE = 2
| | | | | | | | | | SLOPE <= 9.8715: FOREST (2.0)
| | | | | | | | | | SLOPE > 9.8715: URBAN (3.0/1.0)
| | | | | | | | | | SOILTYPE = 3: FOREST (9.0/2.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5: FOREST (1.0)
| | | | | | | | | | SOILTYPE = 6: FOREST (14.0/4.0)
| | | | | | | | | | PCTBARN > 0.064: URBAN (5.0/2.0)
| | | | | | | | | | WETLANDAREA > 34804.11083
| | | | | | | | | | PCTBARN <= 0.023: URBAN (271.0/110.0)
| | | | | | | | | | PCTBARN > 0.023
| | | | | | | | | | SOILTYPE = 0: URBAN (1.0)
| | | | | | | | | | SOILTYPE = 1: URBAN (4.0/2.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (1.0)
| | | | | | | | | | SOILTYPE = 3: AGRICULTURE (1.0)
| | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 5: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 6: URBAN (5.0)
CURRLANDTYPE = AGRICULTURE
| | | | | | | | | | PCTURB <= 43.37
| | | | | | | | | | PCTBARN <= 10.976
| | | | | | | | | | WETLANDAREA <= 45332.69624
| | | | | | | | | | PCTFOR <= 49.336
| | | | | | | | | | PCTAGRI <= 39.848
| | | | | | | | | | GISACRES <= 5.87164
| | | | | | | | | | PCTURB <= 13.078: AGRICULTURE (32.0/15.0)
| | | | | | | | | | PCTURB > 13.078
| | | | | | | | | | SLOPE <= 9.56742
| | | | | | | | | | PCTAGRI <= 22.169: URBAN (117.0/43.0)
| | | | | | | | | | PCTAGRI > 22.169
| | | | | | | | | | GISACRES <= 3.7568

```





```

| | | | | | | | | | SOILTYPE = 6
| | | | | | | | | | PCTWATR <= 4.137
| | | | | | | | | | WETLANDAREA <= 3074.62293: AGRICULTURE (28.0/5.0)
| | | | | | | | | | WETLANDAREA > 3074.62293
| | | | | | | | | | GISACRES <= 4.32097: URBAN (2.0)
| | | | | | | | | | GISACRES > 4.32097: AGRICULTURE (2.0/1.0)
| | | | | | | | | | PCTWATR > 4.137: AGRICULTURE (2.0/1.0)
| | | | | | | | | | WETLANDAREA > 41534.98293: FOREST (3.0/1.0)
| | | | | | | | | | GISACRES > 7.20898: AGRICULTURE (533.0/20.0)
| | | | | PCTFOR > 49.336
| | | | | GISACRES <= 5.66532
| | | | | PCTAGRI <= 29.763
| | | | | GISACRES <= 1.01464: FOREST (34.0)
| | | | | GISACRES > 1.01464
| | | | | PCTAGRI <= 3.693
| | | | | PCTFOR <= 80.102
| | | | | SOILTYPE = 0: URBAN (1.0)
| | | | | SOILTYPE = 1: URBAN (30.0/11.0)
| | | | | SOILTYPE = 2: URBAN (8.0/4.0)
| | | | | SOILTYPE = 3: FOREST (6.0/2.0)
| | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | SOILTYPE = 5: URBAN (0.0)
| | | | | SOILTYPE = 6: FOREST (15.0/6.0)
| | | | | PCTFOR > 80.102: FOREST (8.0)
| | | | | PCTAGRI > 3.693
| | | | | PCTWATR <= 0.656
| | | | | WETLANDAREA <= 3074.62293
| | | | | PCTFOR <= 64.128
| | | | | PCTWETL <= 0.056
| | | | | DIST2HW <= 510: AGRICULTURE (9.0/1.0)
| | | | | DIST2HW > 510: FOREST (5.0/2.0)
| | | | | PCTWETL > 0.056: URBAN (30.0/15.0)
| | | | | PCTFOR > 64.128: FOREST (27.0/8.0)
| | | | | WETLANDAREA > 3074.62293
| | | | | SOILTYPE = 0: FOREST (0.0)
| | | | | SOILTYPE = 1
| | | | | PCTWETL <= 10.54: URBAN (2.0/1.0)
| | | | | PCTWETL > 10.54: FOREST (4.0)
| | | | | SOILTYPE = 2: FOREST (0.0)
| | | | | SOILTYPE = 3: FOREST (0.0)
| | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | SOILTYPE = 5: FOREST (1.0)
| | | | | SOILTYPE = 6: FOREST (3.0)
| | | | | PCTWATR > 0.656: AGRICULTURE (3.0)
| | | | | PCTAGRI > 29.763
| | | | | SOILTYPE = 0: FOREST (1.0)
| | | | | SOILTYPE = 1: AGRICULTURE (37.0/9.0)
| | | | | SOILTYPE = 2: AGRICULTURE (4.0/2.0)
| | | | | SOILTYPE = 3: FOREST (7.0/2.0)
| | | | | SOILTYPE = 4: AGRICULTURE (0.0)
| | | | | SOILTYPE = 5: AGRICULTURE (6.0/1.0)
| | | | | SOILTYPE = 6: FOREST (3.0)
| | | | | GISACRES > 5.66532
| | | | | PCTAGRI <= 11.461
| | | | | DIST2UC <= 10792.3702
| | | | | SOILTYPE = 0: FOREST (1.0)

```

```

| | | | | | | | SOILTYPE = 1
| | | | | | | | WETLANDAREA <= 23531.62946: FOREST (10.0)
| | | | | | | | WETLANDAREA > 23531.62946: AGRICULTURE (2.0/1.0)
| | | | | | | | SOILTYPE = 2: FOREST (1.0)
| | | | | | | | SOILTYPE = 3: AGRICULTURE (2.0)
| | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | SOILTYPE = 5: FOREST (1.0)
| | | | | | | | SOILTYPE = 6: FOREST (5.0)
| | | | | | | | DIST2UC > 10792.3702: AGRICULTURE (23.0/6.0)
| | | | | | | | PCTAGRI > 11.461: AGRICULTURE (114.0/18.0)
| | | | WETLANDAREA > 45332.69624
| | | | PCTWETL <= 40.975
| | | | GISACRES <= 4.43808
| | | | | WETLANDAREA <= 50652.74355: AGRICULTURE (4.0)
| | | | | WETLANDAREA > 50652.74355: WETLANDS (29.0/9.0)
| | | | GISACRES > 4.43808
| | | | | PCTWETL <= 14.838: AGRICULTURE (616.0/53.0)
| | | | | PCTWETL > 14.838
| | | | | WETLANDAREA <= 231161.824
| | | | | GISACRES <= 8.06483
| | | | | PCTAGRI <= 5.324: FOREST (2.0)
| | | | | PCTAGRI > 5.324
| | | | | | WETLANDAREA <= 81412.09324: AGRICULTURE (20.0/1.0)
| | | | | | WETLANDAREA > 81412.09324
| | | | | | | | DIST2HW <= 482
| | | | | | | | SOILTYPE = 0: WETLANDS (0.0)
| | | | | | | | SOILTYPE = 1: URBAN (2.0)
| | | | | | | | SOILTYPE = 2: WETLANDS (1.0)
| | | | | | | | SOILTYPE = 3: URBAN (2.0)
| | | | | | | | SOILTYPE = 4: WETLANDS (0.0)
| | | | | | | | SOILTYPE = 5: WETLANDS (0.0)
| | | | | | | | SOILTYPE = 6: WETLANDS (4.0)
| | | | | | | | DIST2HW > 482: AGRICULTURE (5.0/1.0)
| | | | | | | | GISACRES > 8.06483: AGRICULTURE (129.0/10.0)
| | | | | | | | WETLANDAREA > 231161.824
| | | | | | | | GISACRES <= 37.00253
| | | | | | | | DIST2STREAM <= 2129.95969
| | | | | | | | WETLANDAREA <= 472545.3402
| | | | | | | | GISACRES <= 18.28326
| | | | | | | | DIST2UC <= 15487.87676: WETLANDS (8.0)
| | | | | | | | DIST2UC > 15487.87676: AGRICULTURE (3.0/1.0)
| | | | | | | | GISACRES > 18.28326: AGRICULTURE (18.0/3.0)
| | | | | | | | WETLANDAREA > 472545.3402: WETLANDS (15.0)
| | | | | | | | DIST2STREAM > 2129.95969: FOREST (2.0)
| | | | | | | | GISACRES > 37.00253
| | | | | | | | WETLANDAREA <= 1068559.04: AGRICULTURE (73.0/4.0)
| | | | | | | | WETLANDAREA > 1068559.04
| | | | | | | | GISACRES <= 70.41792: WETLANDS (11.0/4.0)
| | | | | | | | GISACRES > 70.41792: AGRICULTURE (12.0)
| | | | | PCTWETL > 40.975: WETLANDS (61.0/19.0)
| | | PCTBARN > 10.976
| | | PCTFOR <= 40.804
| | | PCTAGRI <= 13.199
| | | | PCTBARN <= 35.17: URBAN (21.0/10.0)
| | | | PCTBARN > 35.17: BARREN LAND (48.0/6.0)
| | | | PCTAGRI > 13.199

```

```

| | | | | SOILTYPE = 0: BARREN LAND (4.0/2.0)
| | | | | SOILTYPE = 1
| | | | | | PCTBARN <= 47.551
| | | | | | | DIST2STREAM <= 331.18571: BARREN LAND (7.0/3.0)
| | | | | | | DIST2STREAM > 331.18571
| | | | | | | GISACRES <= 1.96382
| | | | | | | | PCTAGRI <= 48.96
| | | | | | | | | PCTAGRI <= 24.322: AGRICULTURE (6.0/2.0)
| | | | | | | | | PCTAGRI > 24.322
| | | | | | | | | | PCTWETL <= 17.189
| | | | | | | | | | | DIST2HW <= 237: AGRICULTURE (5.0/2.0)
| | | | | | | | | | | DIST2HW > 237: BARREN LAND (8.0)
| | | | | | | | | | | PCTWETL > 17.189: AGRICULTURE (2.0/1.0)
| | | | | | | | | | | PCTAGRI > 48.96: AGRICULTURE (15.0/1.0)
| | | | | | | | | | | GISACRES > 1.96382
| | | | | | | | | | | GISACRES <= 2.9787: URBAN (3.0/1.0)
| | | | | | | | | | | GISACRES > 2.9787
| | | | | | | | | | | | PCTWATR <= 1.188
| | | | | | | | | | | | | PCTFOR <= 19.431: FOREST (3.0/1.0)
| | | | | | | | | | | | | PCTFOR > 19.431: AGRICULTURE (3.0)
| | | | | | | | | | | | | PCTWATR > 1.188: AGRICULTURE (2.0)
| | | | | | | | | | | | | PCTBARN > 47.551: BARREN LAND (8.0)
| | | | | | | | | | | | | SOILTYPE = 2: AGRICULTURE (1.0)
| | | | | | | | | | | | | SOILTYPE = 3
| | | | | | | | | | | | | | WETLANDAREA <= 7089.1844
| | | | | | | | | | | | | | GISACRES <= 2.08217: BARREN LAND (3.0/1.0)
| | | | | | | | | | | | | | GISACRES > 2.08217: URBAN (3.0)
| | | | | | | | | | | | | | WETLANDAREA > 7089.1844: AGRICULTURE (2.0/1.0)
| | | | | | | | | | | | | | SOILTYPE = 4: AGRICULTURE (0.0)
| | | | | | | | | | | | | | SOILTYPE = 5: AGRICULTURE (2.0)
| | | | | | | | | | | | | | SOILTYPE = 6: AGRICULTURE (1.0)
| | | | | | | | | | | | | | PCTFOR > 40.804
| | | | | | | | | | | | | | DIST2UC <= 2134.43375: URBAN (2.0/1.0)
| | | | | | | | | | | | | | DIST2UC > 2134.43375: FOREST (35.0/10.0)
| | | | | | | | | | | | | | PCTURB > 43.37
| | | | | | | | | | | | | | GISACRES <= 4.95473
| | | | | | | | | | | | | | | WETLANDAREA <= 26963.33125
| | | | | | | | | | | | | | | PCTBARN <= 0.748
| | | | | | | | | | | | | | | PCTAGRI <= 19.074
| | | | | | | | | | | | | | | GISACRES <= 3.10139: URBAN (953.0/36.0)
| | | | | | | | | | | | | | | GISACRES > 3.10139
| | | | | | | | | | | | | | | PCTAGRI <= 1.236: URBAN (31.0/1.0)
| | | | | | | | | | | | | | | PCTAGRI > 1.236
| | | | | | | | | | | | | | | SOILTYPE = 0: URBAN (8.0/3.0)
| | | | | | | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | | | | | | PCTFOR <= 38.25
| | | | | | | | | | | | | | | | DIST2STREAM <= 745.47823: URBAN (10.0/1.0)
| | | | | | | | | | | | | | | | DIST2STREAM > 745.47823
| | | | | | | | | | | | | | | | SLOPE <= 5.28094
| | | | | | | | | | | | | | | | PCTWETL <= 1.715: URBAN (5.0)
| | | | | | | | | | | | | | | | PCTWETL > 1.715: AGRICULTURE (3.0/1.0)
| | | | | | | | | | | | | | | | SLOPE > 5.28094: AGRICULTURE (3.0)
| | | | | | | | | | | | | | | | PCTFOR > 38.25: FOREST (4.0/1.0)
| | | | | | | | | | | | | | | | SOILTYPE = 2: URBAN (8.0/3.0)
| | | | | | | | | | | | | | | | SOILTYPE = 3: URBAN (7.0/1.0)
| | | | | | | | | | | | | | | | SOILTYPE = 4: URBAN (0.0)

```

```

| | | | | | | | SOILTYPE = 5: URBAN (3.0/1.0)
| | | | | | | | SOILTYPE = 6: URBAN (3.0)
| | | | | | | | PCTAGRI > 19.074
| | | | | | | | GISACRES <= 2.39638: URBAN (278.0/31.0)
| | | | | | | | GISACRES > 2.39638
| | | | | | | | DIST2HW <= 342
| | | | | | | | SOILTYPE = 0: URBAN (9.0/3.0)
| | | | | | | | SOILTYPE = 1: AGRICULTURE (51.0/21.0)
| | | | | | | | SOILTYPE = 2: URBAN (8.0/4.0)
| | | | | | | | SOILTYPE = 3: URBAN (5.0/1.0)
| | | | | | | | SOILTYPE = 4: AGRICULTURE (0.0)
| | | | | | | | SOILTYPE = 5: AGRICULTURE (2.0)
| | | | | | | | SOILTYPE = 6
| | | | | | | | GISACRES <= 3.29758: AGRICULTURE (2.0)
| | | | | | | | GISACRES > 3.29758: URBAN (3.0/1.0)
| | | | | | | | DIST2HW > 342
| | | | | | | | PCTWATR <= 0.111: URBAN (30.0/5.0)
| | | | | | | | PCTWATR > 0.111: AGRICULTURE (2.0)
| | | | | | | | PCTBARN > 0.748
| | | | | | | | PCTFOR <= 21.374
| | | | | | | | SOILTYPE = 0: URBAN (18.0/3.0)
| | | | | | | | SOILTYPE = 1
| | | | | | | | DIST2STREAM <= 1125.81821: URBAN (103.0/11.0)
| | | | | | | | DIST2STREAM > 1125.81821
| | | | | | | | PCTBARN <= 28.339: URBAN (17.0/5.0)
| | | | | | | | PCTBARN > 28.339: BARREN LAND (3.0/1.0)
| | | | | | | | SOILTYPE = 2: URBAN (31.0/6.0)
| | | | | | | | SOILTYPE = 3: URBAN (72.0/7.0)
| | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | SOILTYPE = 5: URBAN (3.0/1.0)
| | | | | | | | SOILTYPE = 6: URBAN (16.0/1.0)
| | | | | | | | PCTFOR > 21.374
| | | | | | | | SOILTYPE = 0
| | | | | | | | PCTAGRI <= 2.468: URBAN (3.0)
| | | | | | | | PCTAGRI > 2.468: FOREST (2.0)
| | | | | | | | SOILTYPE = 1: URBAN (15.0/6.0)
| | | | | | | | SOILTYPE = 2: URBAN (2.0/1.0)
| | | | | | | | SOILTYPE = 3: URBAN (15.0/3.0)
| | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | SOILTYPE = 5: URBAN (0.0)
| | | | | | | | SOILTYPE = 6: FOREST (2.0/1.0)
| | | | | | | | WETLANDAREA > 26963.33125
| | | | | | | | PCTBARN <= 8.786
| | | | | | | | PCTFOR <= 0.059: WETLANDS (5.0)
| | | | | | | | PCTFOR > 0.059: URBAN (29.0/11.0)
| | | | | | | | PCTBARN > 8.786: FOREST (2.0/1.0)
| | | | | | | | GISACRES > 4.95473
| | | | | | | | DIST2HW <= 72: URBAN (16.0/7.0)
| | | | | | | | DIST2HW > 72
| | | | | | | | WETLANDAREA <= 241111.1366
| | | | | | | | PCTBARN <= 1.992
| | | | | | | | PCTAGRI <= 18.705
| | | | | | | | GISACRES <= 11.28423
| | | | | | | | PCTFOR <= 6.686: URBAN (7.0/1.0)
| | | | | | | | PCTFOR > 6.686: AGRICULTURE (44.0/23.0)
| | | | | | | | GISACRES > 11.28423: AGRICULTURE (40.0/8.0)

```



```

| | | | | | | | | | GISACRES > 2.49904: FOREST (257.0/58.0)
| | | | | | | | | | PCTBARN > 0.772
| | | | | | | | | | SOILTYPE = 0: FOREST (3.0/1.0)
| | | | | | | | | | SOILTYPE = 1: FOREST (18.0/2.0)
| | | | | | | | | | SOILTYPE = 2: FOREST (8.0/1.0)
| | | | | | | | | | SOILTYPE = 3: FOREST (6.0/1.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | PCTWETL <= 3.126: FOREST (9.0/1.0)
| | | | | | | | | | PCTWETL > 3.126: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 6: FOREST (17.0/1.0)
| | | | | | | | | | SLOPE > 9.65234
| | | | | | | | | | DIST2HW <= 262
| | | | | | | | | | SOILTYPE = 0
| | | | | | | | | | PCTFOR <= 69.638
| | | | | | | | | | GISACRES <= 0.06997: URBAN (3.0)
| | | | | | | | | | GISACRES > 0.06997
| | | | | | | | | | DIST2STREAM <= 454.61784
| | | | | | | | | | PCTWATR <= 0.165: FOREST (9.0/2.0)
| | | | | | | | | | PCTWATR > 0.165: URBAN (4.0)
| | | | | | | | | | DIST2STREAM > 454.61784: FOREST (10.0)
| | | | | | | | | | PCTFOR > 69.638: FOREST (11.0)
| | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | PCTAGRI <= 9.986: FOREST (9.0/2.0)
| | | | | | | | | | PCTAGRI > 9.986: URBAN (3.0)
| | | | | | | | | | SOILTYPE = 2: FOREST (9.0/4.0)
| | | | | | | | | | SOILTYPE = 3: FOREST (14.0/3.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5: FOREST (9.0/3.0)
| | | | | | | | | | SOILTYPE = 6: FOREST (139.0/23.0)
| | | | | | | | | | DIST2HW > 262: FOREST (140.0/7.0)
| | | | | | | | | | PCTAGRI > 21.207
| | | | | | | | | | SOILTYPE = 0: FOREST (6.0/1.0)
| | | | | | | | | | SOILTYPE = 1: FOREST (40.0/6.0)
| | | | | | | | | | SOILTYPE = 2: FOREST (10.0/5.0)
| | | | | | | | | | SOILTYPE = 3: FOREST (19.0/4.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | PCTWATR <= 0.163
| | | | | | | | | | PCTWETL <= 4.753: FOREST (4.0)
| | | | | | | | | | PCTWETL > 4.753: URBAN (2.0/1.0)
| | | | | | | | | | PCTWATR > 0.163: FOREST (2.0)
| | | | | | | | | | SOILTYPE = 6: FOREST (34.0/6.0)
| | | | | GISACRES > 3.84451
| | | | | PCTFOR <= 49.99
| | | | | GISACRES <= 9.54958: FOREST (422.0/67.0)
| | | | | GISACRES > 9.54958
| | | | | SOILTYPE = 0: FOREST (13.0/1.0)
| | | | | SOILTYPE = 1: FOREST (59.0/13.0)
| | | | | SOILTYPE = 2
| | | | | PCTBARN <= 0.104
| | | | | PCTWATR <= 0.002
| | | | | PCTAGRI <= 28.432
| | | | | PCTURB <= 28.635: AGRICULTURE (3.0)
| | | | | PCTURB > 28.635: FOREST (2.0)
| | | | | PCTAGRI > 28.432: FOREST (6.0)

```

```

| | | | | | | | | PCTWATR > 0.002: FOREST (2.0)
| | | | | | | | | PCTBARN > 0.104: FOREST (2.0)
| | | | | | | | | SOILTYPE = 3: FOREST (22.0/5.0)
| | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | SOILTYPE = 5: FOREST (10.0/1.0)
| | | | | | | | | SOILTYPE = 6: FOREST (62.0/7.0)
| | | | | | | | | PCTFOR > 49.99: FOREST (1551.0/64.0)
| | | WETLANDAREA > 34713.31772
| | | | GISACRES <= 11.51917
| | | | | WETLANDAREA <= 120646.5863
| | | | | | GISACRES <= 4.93076: FOREST (88.0/28.0)
| | | | | | GISACRES > 4.93076
| | | | | | | SOILTYPE = 0: FOREST (9.0)
| | | | | | | SOILTYPE = 1: FOREST (46.0/2.0)
| | | | | | | SOILTYPE = 2: FOREST (9.0/3.0)
| | | | | | | SOILTYPE = 3: FOREST (35.0/2.0)
| | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | SOILTYPE = 5: FOREST (12.0/1.0)
| | | | | | | SOILTYPE = 6
| | | | | | | | DIST2STREAM <= 416.32843: FOREST (48.0/10.0)
| | | | | | | | DIST2STREAM > 416.32843
| | | | | | | | | DIST2STREAM <= 983.2542
| | | | | | | | | | WETLANDAREA <= 85204.21494: FOREST (7.0)
| | | | | | | | | | WETLANDAREA > 85204.21494: WETLANDS (4.0/1.0)
| | | | | | | | | | DIST2STREAM > 983.2542: FOREST (26.0)
| | | | | | | | | | WETLANDAREA > 120646.5863: FOREST (29.0/13.0)
| | | | | GISACRES > 11.51917
| | | | | | WETLANDAREA <= 200228.0363
| | | | | | | PCTURB <= 29.356: FOREST (212.0/21.0)
| | | | | | | PCTURB > 29.356
| | | | | | | | SOILTYPE = 0: URBAN (2.0/1.0)
| | | | | | | | SOILTYPE = 1: FOREST (14.0/3.0)
| | | | | | | | SOILTYPE = 2
| | | | | | | | | SLOPE <= 3.01881: URBAN (2.0/1.0)
| | | | | | | | | SLOPE > 3.01881: FOREST (4.0)
| | | | | | | | | SOILTYPE = 3: FOREST (6.0)
| | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | SOILTYPE = 5: FOREST (5.0/2.0)
| | | | | | | | | SOILTYPE = 6: FOREST (12.0/1.0)
| | | | | | | | | WETLANDAREA > 200228.0363: FOREST (197.0/28.0)
| | | PCTURB > 44.314
| | | | GISACRES <= 1.98787
| | | | | SLOPE <= 9.99945
| | | | | | GISACRES <= 0.03029: URBAN (263.0/2.0)
| | | | | | GISACRES > 0.03029
| | | | | | | PCTFOR <= 27.562
| | | | | | | | PCTAGRI <= 0.299: URBAN (247.0/21.0)
| | | | | | | | PCTAGRI > 0.299
| | | | | | | | | PCTWATR <= 0.177: URBAN (136.0/30.0)
| | | | | | | | | PCTWATR > 0.177
| | | | | | | | | PCTBARN <= 1.484
| | | | | | | | | | DIST2UC <= 8345.46455: URBAN (11.0/4.0)
| | | | | | | | | | DIST2UC > 8345.46455: FOREST (7.0)
| | | | | | | | | | PCTBARN > 1.484: FOREST (2.0/1.0)
| | | | | | | | | | PCTFOR > 27.562
| | | | | | | | | | DIST2STREAM <= 392.62976

```

```

| | | | | | | | | | SOILTYPE = 0: FOREST (6.0)
| | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | WETLANDAREA <= 3193.56649: FOREST (12.0/2.0)
| | | | | | | | | | WETLANDAREA > 3193.56649: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (3.0/1.0)
| | | | | | | | | | SOILTYPE = 3: FOREST (6.0/2.0)
| | | | | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | PCTWETL <= 8.861
| | | | | | | | | | PCTAGRI <= 1.9: URBAN (4.0/1.0)
| | | | | | | | | | PCTAGRI > 1.9: FOREST (6.0)
| | | | | | | | | | PCTWETL > 8.861: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 6: FOREST (34.0/12.0)
| | | | | | | | | | DIST2STREAM > 392.62976
| | | | | | | | | | DIST2HW <= 149
| | | | | | | | | | WETLANDAREA <= 854.88664
| | | | | | | | | | SOILTYPE = 0: FOREST (1.0)
| | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | PCTFOR <= 42.334
| | | | | | | | | | PCTAGRI <= 1.756: URBAN (19.0/3.0)
| | | | | | | | | | PCTAGRI > 1.756
| | | | | | | | | | PCTURB <= 44.835: URBAN (2.0)
| | | | | | | | | | PCTURB > 44.835
| | | | | | | | | | PCTFOR <= 30.988: URBAN (2.0/1.0)
| | | | | | | | | | PCTFOR > 30.988: FOREST (3.0)
| | | | | | | | | | PCTFOR > 42.334: FOREST (7.0/1.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (11.0/3.0)
| | | | | | | | | | SOILTYPE = 3
| | | | | | | | | | DIST2STREAM <= 1103.6516: FOREST (8.0/1.0)
| | | | | | | | | | DIST2STREAM > 1103.6516: URBAN (4.0)
| | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 5
| | | | | | | | | | PCTFOR <= 34.343: URBAN (5.0)
| | | | | | | | | | PCTFOR > 34.343: FOREST (7.0)
| | | | | | | | | | SOILTYPE = 6
| | | | | | | | | | PCTURB <= 57.005: URBAN (12.0/4.0)
| | | | | | | | | | PCTURB > 57.005: FOREST (9.0/1.0)
| | | | | | | | | | WETLANDAREA > 854.88664: FOREST (2.0)
| | | | | | | | | | DIST2HW > 149
| | | | | | | | | | PCTBARN <= 0.102
| | | | | | | | | | SOILTYPE = 0: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 1: URBAN (46.0/12.0)
| | | | | | | | | | SOILTYPE = 2
| | | | | | | | | | GISACRES <= 0.21044: FOREST (4.0)
| | | | | | | | | | GISACRES > 0.21044: URBAN (21.0/5.0)
| | | | | | | | | | SOILTYPE = 3: URBAN (6.0/1.0)
| | | | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | | | SOILTYPE = 5: URBAN (7.0/1.0)
| | | | | | | | | | SOILTYPE = 6: URBAN (29.0/8.0)
| | | | | | | | | | PCTBARN > 0.102: URBAN (14.0)
| | | | | | | | | | SLOPE > 9.99945
| | | | | | | | | | PCTBARN <= 5.091
| | | | | | | | | | SOILTYPE = 0: FOREST (15.0)
| | | | | | | | | | SOILTYPE = 1: URBAN (6.0/2.0)
| | | | | | | | | | SOILTYPE = 2: URBAN (2.0)
| | | | | | | | | | SOILTYPE = 3: URBAN (6.0/2.0)

```

```

| | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | SOILTYPE = 5: URBAN (2.0/1.0)
| | | | | SOILTYPE = 6
| | | | | PCTFOR <= 50.7
| | | | | | PCTAGRI <= 2.348
| | | | | | | DIST2STREAM <= 1512.66386: FOREST (35.0/2.0)
| | | | | | | DIST2STREAM > 1512.66386: URBAN (2.0)
| | | | | | | PCTAGRI > 2.348: URBAN (2.0/1.0)
| | | | | | | PCTFOR > 50.7: URBAN (3.0)
| | | | | | | PCTBARN > 5.091: URBAN (8.0/4.0)
| | | | | GISACRES > 1.98787
| | | | | GISACRES <= 6.67422
| | | | | SLOPE <= 10
| | | | | | PCTURB <= 67.711
| | | | | | PCTFOR <= 21.646
| | | | | | | PCTBARN <= 1.937: FOREST (47.0/21.0)
| | | | | | | PCTBARN > 1.937: URBAN (5.0/1.0)
| | | | | | | PCTFOR > 21.646
| | | | | | | WETLANDAREA <= 29346.49604
| | | | | | | SOILTYPE = 0: FOREST (14.0/2.0)
| | | | | | | SOILTYPE = 1: FOREST (114.0/38.0)
| | | | | | | SOILTYPE = 2: FOREST (34.0/15.0)
| | | | | | | SOILTYPE = 3: FOREST (50.0/16.0)
| | | | | | | SOILTYPE = 4: FOREST (0.0)
| | | | | | | SOILTYPE = 5
| | | | | | | | PCTFOR <= 38.019: FOREST (10.0)
| | | | | | | | PCTFOR > 38.019: URBAN (9.0/3.0)
| | | | | | | | SOILTYPE = 6
| | | | | | | | GISACRES <= 2.20401: URBAN (9.0/2.0)
| | | | | | | | GISACRES > 2.20401: FOREST (38.0/8.0)
| | | | | | | | WETLANDAREA > 29346.49604: FOREST (21.0/9.0)
| | | | | | | PCTURB > 67.711
| | | | | | | SOILTYPE = 0: FOREST (6.0/2.0)
| | | | | | | SOILTYPE = 1
| | | | | | | | PCTBARN <= 0.082
| | | | | | | | DIST2UC <= 3081.02952: FOREST (7.0/1.0)
| | | | | | | | DIST2UC > 3081.02952: URBAN (38.0/9.0)
| | | | | | | | PCTBARN > 0.082: FOREST (2.0)
| | | | | | | | SOILTYPE = 2
| | | | | | | | WETLANDAREA <= 23497.06122
| | | | | | | | DIST2UC <= 8700.36724
| | | | | | | | GISACRES <= 2.41072: URBAN (2.0)
| | | | | | | | GISACRES > 2.41072: FOREST (8.0/1.0)
| | | | | | | | DIST2UC > 8700.36724: URBAN (11.0/1.0)
| | | | | | | | WETLANDAREA > 23497.06122: FOREST (2.0)
| | | | | | | SOILTYPE = 3
| | | | | | | | DIST2UC <= 3941.38572: URBAN (5.0)
| | | | | | | | DIST2UC > 3941.38572: FOREST (11.0/3.0)
| | | | | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | | | | SOILTYPE = 5: URBAN (5.0/2.0)
| | | | | | | | SOILTYPE = 6
| | | | | | | | GISACRES <= 4.36368
| | | | | | | | DIST2UC <= 2917.38998: FOREST (2.0)
| | | | | | | | DIST2UC > 2917.38998: URBAN (9.0/1.0)
| | | | | | | | GISACRES > 4.36368: FOREST (3.0)
| | | | | | | SLOPE > 10: FOREST (39.0/3.0)

```

```

| | | | GISACRES > 6.67422
| | | | | DIST2HW <= 58: FOREST (6.0/4.0)
| | | | | DIST2HW > 58
| | | | | | PCTWATR <= 4.353
| | | | | | | PCTAGRI <= 14.183: FOREST (84.0/10.0)
| | | | | | | PCTAGRI > 14.183
| | | | | | | | PCTWATR <= 1.028: FOREST (33.0/8.0)
| | | | | | | | PCTWATR > 1.028: URBAN (2.0/1.0)
| | | | | | | | PCTWATR > 4.353: FOREST (5.0/2.0)
| | PCTBARN > 26.718
| | | PCTURB <= 40.548
| | | | DIST2STREAM <= 628.00571: FOREST (6.0)
| | | | DIST2STREAM > 628.00571: BARREN LAND (63.0/4.0)
| | | PCTURB > 40.548
| | | | PCTBARN <= 43.133: URBAN (34.0/8.0)
| | | | PCTBARN > 43.133: BARREN LAND (9.0)
CURRLANDTYPE = WETLANDS
| | WETLANDAREA <= 73.41142
| | | WETLANDAREA <= 17.78172
| | | | SOILTYPE = 0: URBAN (0.0)
| | | | SOILTYPE = 1: URBAN (18.0)
| | | | SOILTYPE = 2: URBAN (0.0)
| | | | SOILTYPE = 3
| | | | | PCTFOR <= 21.292
| | | | | | SLOPE <= 4.81566: URBAN (19.0)
| | | | | | SLOPE > 4.81566: BARREN LAND (2.0)
| | | | | | PCTFOR > 21.292: BARREN LAND (3.0)
| | | | | SOILTYPE = 4: URBAN (0.0)
| | | | | SOILTYPE = 5: URBAN (0.0)
| | | | | SOILTYPE = 6: URBAN (4.0)
| | | WETLANDAREA > 17.78172: URBAN (2.0/1.0)
| | WETLANDAREA > 73.41142
| | | WETLANDAREA <= 49473.47602
| | | | GISACRES <= 1.6309
| | | | | PCTWATR <= 1.901
| | | | | | SOILTYPE = 0: WETLANDS (5.0)
| | | | | | SOILTYPE = 1
| | | | | | | GISACRES <= 0.72945: WETLANDS (20.0/1.0)
| | | | | | | GISACRES > 0.72945
| | | | | | | | WETLANDAREA <= 20994.34131: URBAN (5.0/1.0)
| | | | | | | | WETLANDAREA > 20994.34131: WETLANDS (9.0)
| | | | | | | | SOILTYPE = 2: WETLANDS (7.0)
| | | | | | | | SOILTYPE = 3: WETLANDS (30.0/2.0)
| | | | | | | | SOILTYPE = 4: WETLANDS (0.0)
| | | | | | | | SOILTYPE = 5: WETLANDS (12.0/3.0)
| | | | | | | | SOILTYPE = 6: WETLANDS (51.0/8.0)
| | | | | | | | PCTWATR > 1.901: WETLANDS (55.0)
| | | | GISACRES > 1.6309
| | | | | PCTWATR <= 3.273
| | | | | | PCTURB <= 18.944: WETLANDS (14.0/1.0)
| | | | | | PCTURB > 18.944
| | | | | | | PCTWATR <= 1.094
| | | | | | | | WETLANDAREA <= 20557.29429: URBAN (9.0/1.0)
| | | | | | | | WETLANDAREA > 20557.29429
| | | | | | | | | DIST2STREAM <= 2633.73643
| | | | | | | | | DIST2UC <= 8609.08467

```

```

| | | | | | | | | | GISACRES <= 1.81574: WETLANDS (3.0/1.0)
| | | | | | | | | | GISACRES > 1.81574: URBAN (6.0)
| | | | | | | | | | DIST2UC > 8609.08467
| | | | | | | | | | GISACRES <= 3.52393
| | | | | | | | | | SOILTYPE = 0: WETLANDS (0.0)
| | | | | | | | | | SOILTYPE = 1
| | | | | | | | | | GISACRES <= 1.88495: WETLANDS (4.0)
| | | | | | | | | | GISACRES > 1.88495: URBAN (2.0/1.0)
| | | | | | | | | | SOILTYPE = 2: WETLANDS (0.0)
| | | | | | | | | | SOILTYPE = 3
| | | | | | | | | | PCTFOR <= 10.765: URBAN (2.0)
| | | | | | | | | | PCTFOR > 10.765: WETLANDS (4.0)
| | | | | | | | | | SOILTYPE = 4: WETLANDS (0.0)
| | | | | | | | | | SOILTYPE = 5: WETLANDS (0.0)
| | | | | | | | | | SOILTYPE = 6: WETLANDS (8.0)
| | | | | | | | | | GISACRES > 3.52393: URBAN (3.0)
| | | | | | | | | | DIST2STREAM > 2633.73643: URBAN (4.0/1.0)
| | | | | | | | | | PCTWATR > 1.094: FOREST (4.0/2.0)
| | | | | | | | | | PCTWATR > 3.273: WETLANDS (7.0/1.0)
| | WETLANDAREA > 49473.47602: WETLANDS (575.0/24.0)
CURRLANDTYPE = BARREN LAND
| PCTBARN <= 0.052: URBAN (1331.0/13.0)
| PCTBARN > 0.052
| | GISACRES <= 5.94512
| | | PCTBARN <= 15.109: URBAN (17.0/2.0)
| | | PCTBARN > 15.109: BARREN LAND (3.0/1.0)
| | GISACRES > 5.94512: BARREN LAND (14.0/4.0)
CURRLANDTYPE = WATER: WATER (88.0/7.0)

```

Number of Leaves : 602

Size of the tree : 974

## REFERENCES

- Aalders, I. H., & Aitkenhead, M. J. (2006). Agricultural census data and land use modelling. *Computers, Environment and Urban Systems*, 30(6), 799-814.
- Aitkenhead, M. J., Mustard, M. J., & McDonald, A. J. S. (2004). Using neural networks to predict spatial structure in ecological systems. *Ecological Modelling*, 179(3), 393-403.
- Alig, R. J., & Plantinga, A. J. (2004). Future forestland area: Impacts from population growth and other factors that affect land values. *Journal of Forestry*, 102(8), 19-24.
- Allen, J., & Lu, K. (2003). Modeling and prediction of future urban growth in the Charleston region of South Carolina: A GIS-based integrated approach. *Ecology and Society*, 8(2).
- Anderson, J.R., Hardy, E.E., Roach, J.T. and Witmer, R.E. 1976. A land use and land cover classification system for use with remote sensor data. Geographical Survey Professional Paper 964, U.S. Geological survey, Washington, D.C.
- Aitkenhead, M. J., Mustard, M. J., & McDonald, A. J. S. (2004). Using neural networks to predict spatial structure in ecological systems. *Ecological Modelling*, 179(3), 393-403.
- Altman, D. (1994). Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal of Geographical Information Systems*, 8(3), 271-289.
- Aplin, P., Atkinson, P. M., & Curran, P. J. (1999). Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the united kingdom. *Remote Sensing of Environment*, 68(3), 206-216.
- Aplin, P., & Atkinson, P. M. (2001). Sub-pixel land cover mapping for per-field classification. *International Journal of Remote Sensing*, 22(14), 2853-2858.
- Arai, T., & Akiyama, T. (2004). Empirical analysis for estimating land use transition potential functions - case in the Tokyo Metropolitan Region. *Computers, Environment and Urban Systems*, 28(1-2), 65-84.
- Baker, W. L. (1989). A review of models of landscape change. *Landscape Ecology*, 2(2), 111-133.

- Banai, R. (1993). Fuzziness in geographical information systems: Contributions from the analytic hierarchy process. *International Journal of Geographical Information Systems*, 7(4), 315-329.
- Barringer, J. L., Bonin, J. L., Deluca, M. J., Romagna, T., Cenno, K., Alebus, M., et al. (2007). Sources and temporal dynamics of arsenic in a new jersey watershed, USA. *Science of the Total Environment*, 379(1), 56-74.
- Batten, D.F. and D.E. Boyce. 1986. "Spatial Interaction, Transportation and Interregional Commodity Flow Models." In *Handbook of Regional and Urban Economics*. Vol. 1. ed. P. Nijkamp, 357-406. Amsterdam: North-Holland.
- Batten, D. F., & Boyce, D. E. (1986). Spatial interaction, transportation, and interregional commodity flow models. *Handbook of regional and urban economics*. Vol. 1: regional economics, , 357-406.
- Batty, M. 1976. *Urban Modeling: Algorithms, Calibrations, Predictions*. Cambridge: Cambridge University Press.
- Batty, M. Personal Communication, February 19, 2007
- Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems*, 23(3), 205-233.
- Bell, K. P., & Bockstael, N. E. (2000). Applying the generalized-moments estimation approach to spatial problems involving micro level data. *Review of Economics and Statistics*, 82(1), 72-82.
- Berberoglu, S., Lloyd, C. D., Atkinson, P. M., & Curran, P. J. (2000). The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. *Computers and Geosciences*, 26(4), 385-396.
- Bjelland, M. D. (2004). Brownfield sites in Minneapolis-St. Paul: The interwoven geographies of industrial disinvestment and environmental contamination. *Urban Geography*, 25(7), 631-657.
- Braimoh, A. K., & Onishi, T. (2007). Spatial determinants of urban land use change in Lagos, Nigeria. *Land Use Policy*, 24(2), 502-515.
- Briassoulis, H., Analysis of Land Use Change: Theoretical and Modeling Approaches, The Web Book of Regional Science. Retrieved September 8, 2007, from <http://www.rri.wvu.edu/WebBook/Briassoulis/contents.htm>
- Briassoulis, H. (1994). Pollution prevention for sustainable development: The land-use question. *INT.J.SUSTAINABLE DEV.WORLD ECOL.*, 1(2), 110-120.

- Brown, D. E., Corruble, V., & Pittard, C. L. (1993). A comparison of decision tree classifiers with back propagation neural networks for multimodal classification problems. *Pattern Recognition*, 26(6), 953-961.
- Burley, T. M. 1961. Land use or land utilization. *Professional Geographer* 13(6):18-20.
- Buttle, J. M., & Xu, F. (1988). Snowmelt runoff in suburban environments, (Ontario, Canada). *Nordic Hydrology*, 19(1), 19-40.
- CARA, 2006 Land Use Primer. Consortium for Atlantic Regional Assessment. Retrieved November 7, 2007, from <http://www.cara.psu.edu/land/lu-primer/luprimer09.asp>
- CDC. (2002). Centers for Disease Control, Urban Sprawl and Public Health, Public Health Reports, May-June 2002, Vol. 117, Retrieved November 6, 2007, from <http://www.cdc.gov/healthyplaces/articles/Urban%20Sprawl%20and%20Public%20Health%20-%20PHR.pdf>
- Chen, J., Gong, P., He, C., Luo, W., Tamura, M., & Shi, P. (2002). Assessment of the urban development plan of Beijing by using a CA-based urban growth model. *Photogrammetric Engineering and Remote Sensing*, 68(10), 1063-1071.
- Chen, Q., & Mynett, A. E. (2003). Effects of cell size and configuration in cellular automata based prey-predator modelling. *Simulation Modelling Practice and Theory*, 11(7-8), 609-625.
- Childress, W. M., Rykiel Jr., E. J., Forsythe, W., Li, B. -, & Wu, H. -. (1996). Transition rule complexity in grid-based automata models. *Landscape Ecology*, 11(5), 257-266.
- Cihlar, J., & Jansen, L. J. M. (2001). From land cover to land use: A methodology for efficient land use mapping over large areas. *Professional Geographer*, 53(2), 275-289.
- Clarke, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco bay area. *Environment and Planning B: Planning and Design*, 24(2), 247-261.
- Clarke, K. C., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: Long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12(7), 699-714.
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1), 35-46.

- Conway, T. M., & Lathrop Jr., R. G. (2005a). Modeling the ecological consequences of land-use policies in an urbanizing region. *Environmental Management*, 35(3), 278-291.
- Conway, T. M., & Lathrop, R. G. (2005b). Alternative land use regulations and environmental impacts: Assessing future land use in an urbanizing watershed. *Landscape and Urban Planning*, 71(1), 1-15.
- Couclelis, H. (1985). Cellular worlds: A framework for modeling micro-macro dynamics. *Environment & Planning A*, 17(5), 585-596.
- Dai, E., Wu, S., Shi, W., Cheung, C. -, & Shaker, A. (2005). Modeling change-pattern-value dynamics on land use: An integrated GIS and artificial neural networks approach. *Environmental Management*, 36(4), 576-591.
- Davidson, D. A., Theocharopoulos, S. P., & Bloksma, R. J. (1994). A land evaluation project in Greece using GIS and based on Boolean and fuzzy set methodologies. *International Journal of Geographical Information Systems*, 8(4), 369-384.
- Davis, J. S., Nelson, A. C., & Dueker, K. J. (1994). The new 'burbs': The exurbs and their implications for planning policy. *Journal - American Planning Association*, 60(1), 45-59.
- de Almeida, C. M., Batty, M., Monteiro, A. M. V., Câmara, G., Soares-Filho, B. S., & Cerqueira, G. C. et al. (2003). Stochastic cellular automata modeling of urban land use dynamics: Empirical development and estimation. *Computers, Environment and Urban Systems*, 27(5), 481-509.
- De Kok, J. -, Engelen, G., White, R., & Wind, H. G. (2001). Modeling land-use change in a decision-support system for coastal-zone management. *Environmental Modeling and Assessment*, 6(2), 123-132.
- de Wit, A. J. W., & Clevers, J. G. P. W. (2004). Efficiency and accuracy of per-field classification for operational crop mapping. *International Journal of Remote Sensing*, 25(20), 4091-4112.
- Dean, A. M., & Smith, G. M. (2003). An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing*, 24(14), 2905-2920.
- Dietzel, C., & Clarke, K. C. (2004). Spatial differences in multi-resolution urban automata modeling. *Transactions in GIS*, 8(4), 479-492.
- Dragicevic, S. (2004). Coupling fuzzy sets theory and GIS-based cellular automata for land-use change modeling. 1 203-207.

- Du, B., Zhang, K., Song, G., & Wen, Z. (2006). Methodology for an urban ecological footprint to evaluate sustainable development in China. *International Journal of Sustainable Development and World Ecology*, 13(4), 245-254.
- Duerden, F. (1996). Land allocation in comprehensive land claim agreements: The case of the Yukon land claim. *Applied Geography*, 16(4), 279-288.
- Duke, J. M. (2004). Participation in agricultural land preservation programs: Parcel quality and a complex policy environment. *Agricultural and Resource Economics Review*, 33(1), 34-49.
- Engelen, G., White, R., Uljee, I., & Drazan, P. (1995). Using cellular automata for integrated modelling of socio-environmental systems. *Environmental Monitoring and Assessment*, 34(2), 203-214.
- Engelen, G., White, R., & de Nijs, T. (2003). Environment explorer: Spatial support system for the integrated assessment of socio-economic and environmental policies in the Netherlands. *Integrated Assessment*, 4(2), 97-105.
- Entwisle, B., Walsh, S. J., Rindfuss, R. R., & VanWey, L. K. (2005). Population and upland crop production in Nang Rong, Thailand. *Population and Environment*, 26(6), 449-470.
- Erol, H., & Akdeniz, F. (2005). A per-field classification method based on mixture distribution models and an application to landsat thematic mapper data. *International Journal of Remote Sensing*, 26(6), 1229-1244.
- Fang, S., Gertner, G. Z., & Anderson, A. B. (2007). Prediction of multinomial probability of land use change using a bisection decomposition and logistic regression. *Landscape Ecology*, 22(3), 419-430.
- Fangju Wang. (1994). Towards a natural language user interface: An approach of fuzzy query. *International Journal of Geographical Information Systems*, 8(2), 143-162.
- Flache, A, and Hegselmann, R. "Do Irregular Grids make a Difference? Relaxing the Spatial Regularity Assumption in Cellular Models of Social Dynamics. *Journal of Artificial Societies and Social Simulation* vol. 4, no. 4. Retrieved December 8, 2007, <http://jasss.soc.surrey.ac.uk/4/4/6.html>
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185-201.
- Fritz, S., & See, L. (2005). Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science*, 19(7), 787-807.

- Fuller, R. M., Smith, G. M., Sanderson, J. M., Hill, R. A., & Thomson, A. G. (2002). The UK land cover map 2000: Construction of a parcel-based vector map from satellite images. *Cartographic Journal*, 39(1), 15-25.
- Fuller, R. M., Wyatt, B. K., & Barr, C. J. (1998). Countryside survey from ground and space: Different perspectives, complementary results. *Journal of Environmental Management*, 54(2), 101-126.
- Galle, S., Ehrmann, M., & Peugeot, C. (1999). Water balance in a banded vegetation pattern. A case study of tiger bush in Western Niger. *Catena*, 37(1-2), 197-216.
- Gaston, K. J., Blackburn, T. M., & Klein Goldewijk, K. (2003). Habitat conversion and global avian biodiversity loss. *Proceedings of the Royal Society of London - Biological Sciences*, 270(1521), 1293-1300.
- Geoghegan, J. (2002). The value of open spaces in residential land use. *Land Use Policy*, 19(1), 91-98.
- Geoghegan, J., Pritchard, L., Ogneva-Himmelberger, Y., Chowdhury, R., Sanderson, S., and Turner II, B. 1998. " 'Socializing the Pixel' and 'Pixelizing the Social' in Land-Use and Land-Cover Change." In *People and Pixels: Linking Remote Sensing and Social Science*. D. Liverman, E. Moran, R. Rindfuss, and P. Stern, (eds.) Committee on the Human Dimensions of Global Environmental Change, National Research Council. Washington, D.C. National Academy of Science Press, pp. 51-69.
- Geoghegan, J., Wainger, L. A., & Bockstael, N. E. (1997). Spatial landscape indices in a hedonic framework: An ecological economics analysis using GIS. *Ecological Economics*, 23(3), 251-264.
- Lynch, L., Bucholtz, S., and Malcolm, S., (2002). Parcel Level GIS Data and Measurement Error: Are we producing biased and inconsistent estimates? Retrieved December 8, 2007, from [http://www.aere.org/meetings/0306workshop\\_Lynch.pdf](http://www.aere.org/meetings/0306workshop_Lynch.pdf)
- Ghaffari, A., Priestnall, G., & Clarke, M. L. (2006). Artificial neural networks and decision tree classifier performance on medium resolution ASTER data to detect gully networks in southern Italy.
- Gong, H., & Wheeler, J. O. (2002). The location and suburbanization of business and professional services in the Atlanta Metropolitan Area. *Growth and Change*, 33(3), 341-369.
- Goodman, R. M., & Smyth, P. (1988). Decision tree design from a communication theory standpoint. *IEEE Transactions on Information Theory*, 34(5 pt 1), 979-994.
- Gordon, P., & Richardson, H. W. (2004). Exit and voice in U.S. settlement change. *Review of Austrian Economics*, 17(2-3), 187-202.

- Guan, Q., Wang, L., & Clarke, K. C. (2005). An artificial-neural-network-based, constrained CA model for simulating urban growth. *Cartography and Geographic Information Science*, 32(4), 369-380.
- Guan, Q., Wang, L., & Clarke, K. C. (2005). An artificial-neural-network-based, constrained CA model for simulating urban growth. *Cartography and Geographic Information Science*, 32(4), 369-380.
- Hagen, A. (2003). Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science*, 17(3), 235-249.
- Hagen-Zanker, A., Straatman, B., & Uljee, I. (2005). Further developments of a fuzzy set map comparison approach. *International Journal of Geographical Information Science*, 19(7), 769-785.
- Harder, S. V., Amatya, D. M., Callahan, T. J., Trettin, C. C., & Hakkila, J. (2007). Hydrology and water budget for a forested Atlantic coastal plain watershed, South Carolina. *Journal of the American Water Resources Association*, 43(3), 563-575.
- Harris, B. 1966. "The Use of Theory in the Simulation of Urban Phenomena." *Journal of the American Institute of Planners* 32(5): p.258.
- Hasse, J., and Lathrop, R.G., (2001). Measuring Urban Growth in New Jersey, Retrieved April 25, 2008, from <http://crssa.rutgers.edu/projects/lc/urbangrowth/index.html>
- Haynes, K.E. and A.S. Fotheringham. 1984. *Gravity and Spatial Interaction Models*. Beverly Hills: Sage.
- Haynes, K. E., & Fotheringham, A. S. (1991). The impact of space on the application of discrete choice models. *Review of Regional Studies*, 20(2), 39-49.
- HCDPR. (2008). Hunterdon County Department of Parks and Recreation. Retrieved March 19, 2008, from <http://www.co.hunterdon.nj.us/depts/parks/parks.htm>
- HCPB. (2005a). Hunterdon County Planning Board, People and Housing Profile. Retrieved March 11, 2008, from <http://www.co.hunterdon.nj.us/pdf/hcpb/demographics.pdf>
- HCPB. (2005b). Hunterdon County Planning Board, Economic Development Profile. Retrieved March 11, 2008, from <http://www.co.hunterdon.nj.us/pdf/hcpb/econdev.pdf>
- HCPB. (2005c). Hunterdon County Planning Board, Land Development Profile. Retrieved March 11, 2008, from <http://www.co.hunterdon.nj.us/pdf/hcpb/landdev.pdf>

- HCPB. (2007a). Hunterdon County Planning Board. 2007 Growth Management Plan. Retrieved March 19, 2008, from <http://www.co.hunterdon.nj.us/pdf/hcpb/2007GrowthManagmentPlan.pdf>
- HCPB. (2007b). Hunterdon County Planning Board. Hunterdon County Agriculture Development Board. Hunterdon County Comprehensive Farmland Preservation Plan. Retrieved March 19, 2008, from <http://www.co.hunterdon.nj.us/pdf/cadb/farmlandplan07.pdf>
- HCPB. (2007c). Hunterdon County Planning Board. Hunterdon County Open Space Trust Fund Program. Retrieved March 19, 2008, from <http://www.co.hunterdon.nj.us/openspac.htm>
- Hegselmann, R. Personal communication. February 19, 2008.
- Higgins, P. A. T. (2007). Biodiversity loss under existing land use and climate change: An illustration using northern South America. *Global Ecology and Biogeography*, 16(2), 197-204.
- Hughes, J.W., & Seneca, J.J. (1999). The Emerging Wealth Belt: New Jersey's New Millennium Geography. Edward J. Bloustein School of Planning and Public Policy. Retrieved December 8, 2007, from <http://www.heldrich.rutgers.edu/uploadedFiles/Publications/WealthBelt.pdf>
- Hughes, J.W., & Seneca, J.J. (2006) New Jersey's New Economy: Growth Challenges, Rutgers Regional Report, Issue Paper No. 25, July 2006, Edward J. Bloustein School of Planning and Public Policy. Retrieved December 8, 2007, from [http://www.policy.rutgers.edu/news/reports/RRR/RRR\\_July\\_2006.pdf](http://www.policy.rutgers.edu/news/reports/RRR/RRR_July_2006.pdf)
- Hutchinson, C. F. (1982). Techniques for combining landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering & Remote Sensing*, 48(1), 123-130.
- Irwin, E. G., & Bockstael, N. E. (2001). The problem of identifying land use spillovers: Measuring the effects of open space on residential property values. *American Journal of Agricultural Economics*, 83(3), 698-704.
- Irwin, E. G., & Bockstael, N. E. (2004). Land use externalities, open space preservation, and urban sprawl. *Regional Science and Urban Economics*, 34(6), 705-725.
- Irwin, E. G., & Geoghegan, J. (2001). Theory, data, methods: Developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and Environment*, 85(1-3), 7-23.

- Jantrania, A. R. (2004). Evaluating field performance of a pre-engineered land based effluent dispersal system under wide range of soil and site conditions found in Virginia. 482-489.
- Jenerette, G. D., & Wu, J. (2001). Analysis and simulation of land-use change in the central Arizona - phoenix region, USA. *Landscape Ecology*, 16(7), 611-626.
- Johnston, R.J., D. Gregory and D.M. Smith. 1994. *The Dictionary of Human Geography*, 3<sup>rd</sup> edition. Oxford: Blackwell.p.273.
- Jongman, R. H. (1995). Nature conservation planning in Europe: Developing ecological networks. *Landscape and Urban Planning*, 32(3), 169-183.
- Kalnay, E., & Cai, M. (2003). Impact of urbanization and land-use change on climate. *Nature*, 423(6939), 528-531.
- Kampert, J. B., Blair, S. N., Barlow, C. E., & Kohl III, H. W. (1996). Physical activity, physical fitness, and all-cause and cancer mortality: A prospective study of men and women. *Annals of Epidemiology*, 6(5), 452-457.
- Kitamura, R., Chen, C., & Pendyala, R. M. (1997). *Generation of synthetic daily activity-travel patterns*. Transportation Research Record (1607), 154-162.
- Kitamura, T., M. Kagatsume, S. Hoshino, H. Morita. 1997. *A Theoretical Consideration on the Land Use Change Model for the Japan Case Study Area*. Interim Report IR-97-064. Laxenburg: International Institute for Applied Systems Analysis.
- Kocabas, V., & Dragicevic, S. (2006). Assessing cellular automata model behaviour using a sensitivity analysis approach. *Computers, Environment and Urban Systems*, 30(6), 921-953.
- Kollias, V. J., & Voliotis, A. (1991). Fuzzy reasoning in the development of geographical information systems. FRIS : A prototype soil information system with fuzzy retrieval capabilities. *International Journal of Geographical Information Systems*, 5(2), 209-223.
- Koomen, E., Rietveld, P., & Nijs, T. (2008). Modelling land-use change for spatial planning support. *Annals of Regional Science*, 42(1), 1-10.
- LaGro Jr., J. A., & DeGloria, S. D. (1992). Land use dynamics within an urbanizing non-metropolitan county in New York state (USA). *Landscape Ecology*, 7(4), 275-289.
- Lambin, Eric F and Helmut, J. (Eds.) 2006. Land-Use and Land-Cover Change. Local processes and Global Impacts series: Global Change-The IGBP Series. Springer-Verlag.

- Landis, J. D. (1994). Future tense. *Planning (APA)*, 60(2), 22-25.
- Landis, J., & Zhang, M. (1998a). The second generation of the California urban futures model. part 1: Model logic and theory. *Environment and Planning B: Planning and Design*, 25(5), 657-666.
- Landis, J., & Zhang, M. (1998b). The second generation of the California urban futures model. part 2: Specification and calibration results of the land-use change submodel. *Environment and Planning B: Planning and Design*, 25(6), 795-824.
- Langford, W. T., Gergel, S. E., Dietterich, T. G., & Cohen, W. (2006). Map misclassification can cause large errors in landscape pattern indices: Examples from habitat fragmentation. *Ecosystems*, 9(3), 474-488.
- Lathrop, R. G., Tulloch, D. L., & Hatfield, C. (2007). Consequences of land use change in the new york-new jersey highlands, USA: Landscape indicators of forest and watershed integrity. *Landscape and Urban Planning*, 79(2), 150-159.
- Lee, I., Sesso, H. D., & Paffenbarger Jr., R. S. (1999). Physical activity and risk of lung cancer. *International Journal of Epidemiology*, 28(4), 620-625.
- Levy, S. *Artificial Life: The Quest for a New Creation*, Random House, NY, 1992
- Lewis, H. G., & Brown, M. (2001). A generalized confusion matrix for assessing area estimates from remotely sensed data. *International Journal of Remote Sensing*, 22(16), 3223-3235.
- Li, X., & Yeh, A. G. -. (2000). Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science*, 14(2), 131-152.
- Li, X., & Yeh, A. G. -. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), 323-343.
- Li, X., & Yeh, A. G. -. (2004). Data mining of cellular automata's transition rules. *International Journal of Geographical Information Science*, 18(8), 723-744.
- Li, X., & Yeh, A. G. -. (2005). Knowledge discovery for geographical cellular automata. *Science in China, Series D: Earth Sciences*, 48(10), 1758-1767.
- Liu, Y. (2007). The impact of neighbourhood size on the accuracy of cellular automata-based urban modelling. *Applied GIS*, 3(3).

- Liu, X. P., Li, X., Yeh, A. G. -, He, J. Q., & Tao, J. (2007). Discovery of transition rules for geographical cellular automata by using ant colony optimization. *Science in China, Series D: Earth Sciences*, 50(10), 1578-1588.
- Liu, X., Li, X., Shi, X., Wu, S., & Liu, T. (2008). Simulating complex urban development using kernel-based non-linear cellular automata. *Ecological Modelling*, 211(1-2), 169-181.
- Liu, Y. (2007). The impact of neighbourhood size on the accuracy of cellular automata-based urban modelling. *Applied GIS*, 3(3).
- Matlack, G. R. (1997). Land use and forest habitat distribution in the hinterland of a large city. *Journal of Biogeography*, 24(3), 297-307.
- McConnell, W. (2002). Meeting in the middle: The challenge of meso-level integration. *Land Use Policy*, 19(1), 99-101.
- McConnell, W. (2001). Meeting the Challenge of Complexity. *Proceedings of a Special Workshop on Land-use/Land-Cover Change*. Retrieved April 17, 2008, from <http://www.csiss.org/events/other/agent-based/additional/proceedings.pdf>
- McDonald, R. I., & Urban, D. L. (2006). Spatially varying rules of landscape change: Lessons from a case study. *Landscape and Urban Planning*, 74(1), 7-20.
- Metternicht, G. (1999). Change detection assessment using fuzzy sets and remotely sensed data: An application of topographic map revision. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(4), 221-233.
- Meyer, D., Wylie, B., Robbins, C., Vierling, L., and Kozak, P. Decision tree classification of multirate AVIRIS data for mapping woodland encroachment into the great plains. Retrieved July 24, 2007, from [ftp://popo.jpl.nasa.gov/pub/docs/workshops/01\\_docs/2001Meyer\\_web.pdf](ftp://popo.jpl.nasa.gov/pub/docs/workshops/01_docs/2001Meyer_web.pdf)
- Middle, G. (1996). Environmental requirements for the disposal of effluent from wastewater disposal systems. *Desalination*, 106(1-3), 323-329.
- Montgomery Planning Commission, 1996 Shaping Future Development: The role of current zoning. A Build out study of Central and Western Montgomery County. Retrieved February 9, 2008, from <http://www.epa.gov/greenkit/pdfs/futgrow0.pdf>
- Moore, D. M., Lees, B. G., & Davey, S. M. (1991). A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Environmental Management*, 15(1), 59-71.
- Morita, H., S. Hoshino, M. Kagatsume, and K. Mizuno. 1997. *An Application of the Land Use Change Model for the Japan Case Study Area*. Interim Report IR-97-065,

November. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Moreno, N. and D.J. Marceau, 2006. A vector-based cellular automata model to allow changes of polygon shape. *Proceedings of The 2006 SCS International Conference on Modeling and Simulation - Methodology, Tools, Software Applications*, R. Huntsinger, H. Vakilzadian and T. Ören, eds, July 31 - August 2, 2006, Calgary, Canada, pp. 85-92. Retrieved December 8, 2007, from [http://www.ucalgary.ca/gcl/articles/Moreno\\_2006\\_Modeling.pdf](http://www.ucalgary.ca/gcl/articles/Moreno_2006_Modeling.pdf)

Muller, S. V., Walker, D. A., Nelson, F. E., Auerbach, N. A., Bockheim, J. G., Guyer, S. et al. (1998). Accuracy assessment of a land-cover map of the Kuparuk river basin, Alaska: Considerations for remote regions. *Photogrammetric Engineering and Remote Sensing*, 64(6), 619-628.

Muñoz-Villers, L. E., & López-Blanco, J. (2008). Land use/cover changes using landsat TM/ETM images in a tropical and biodiverse mountainous area of central-eastern Mexico. *International Journal of Remote Sensing*, 29(1), 71-93.

Nelson, A. C. (1992). Characterizing exurbia. *Journal of Planning Literature*, 6(4), 350-368.

Nelson, J., and Miller, D.J.:1997, *Expansive Soils: Problems and Practice in Foundation and Pavement Engineering*, Wiley-IEEE, Hoboken, NJ.

NJDEP 2003. New Jersey Department of Environmental Protection. New Jersey Comparative Risk Project, New Jersey Department of Environmental Protection, March 2003.

NJDLD 2006. New Jersey Department of Labor and Workforce Development. Northern Regional Community Fact Book. Hunterdon County Edition. Retrieved April 25, 2008, from <http://www.wnjp.state.nj.us/OneStopCareerCenter/LaborMarketInformation/lmi12/hunfct.pdf>

NJWSA 2003. Watershed Model for Watersheds of the Spruce Run Reservoir. New Jersey Water Supply Authority; Retrieved December 8, 2007, from <http://www.njwsa.org/WPU/SRI/Publications/WatershedModel/Section2DataDevelopment.pdf>

Niang, A. J., Ozer, A., & Ozer, P. (2008). Fifty years of landscape evolution in southwestern Mauritania by means of aerial photos. *Journal of Arid Environments*, 72(2), 97-107.

Nickerson, C. J., & Lynch, L. (2001). The effect of farmland preservation programs on farmland prices. *American Journal of Agricultural Economics*, 83(2), 341-351.

- Niehoff, D., Fritsch, U., & Bronstert, A. (2002). Land-use impacts on storm-runoff generation: Scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany. *Journal of Hydrology*, 267(1-2), 80-93.
- NJDA. (2007). New Jersey Department of Agriculture. New Jersey Farmland Preservation Program. Summary of Preserved Farmland. Retrieved March 19, 2008, from <http://www.nj.gov/agriculture/sadc/farmpreserve/progress/stats/preservedsummary.pdf>
- NJDCA. (2007). New Jersey Department of Community Affairs. Mission Statement. Council on Affordable Housing. Retrieved December 8, 2007, from <http://www.state.nj.us/dca/coah/>
- NJDEP. (2003). New Jersey Department of Environmental Protection. Final Report of the New Jersey Comparative Risk Project. Retrieved December 8, 2007, from <http://www.state.nj.us/dep/dsr/njcrp/>
- NJDEP. (2005). New Jersey Department of Environmental Protection Guidelines for the Highlands Protection and Planning Act. Retrieved March 11, 2008, from [http://www.nj.gov/dep/highlands/faq\\_info.htm](http://www.nj.gov/dep/highlands/faq_info.htm)
- Noble, C. (1999). Lifeline for a landscape. *American Forests*, 105(1), X7-X8.
- Oh, J. (2005). Social disorganizations and crime rates in U.S. central cities: Toward an explanation of urban economic change. *Social Science Journal*, 42(4), 569-582.
- Openshaw, S. (1983). The modifiable areal unit problem. *CATMOG (Concepts & Techniques in Modern Geography)*, 38
- O'Sullivan, D. (2001). Exploring spatial process dynamics using irregular cellular automaton models. *Geographical Analysis*, 33(1), 1-18.
- O'Sullivan, D. (2002). Toward micro-scale spatial modeling of gentrification. *Journal of Geographical Systems*, 4(3), 251-274.
- O'Sullivan, D., and Torrens, P. M. (2000) "Cellular Models of Urban Systems". *CASA Working Paper 22*. University College London, Centre for Advanced Spatial Analysis. Retrieved September 8, 2007, from [http://www.casa.ucl.ac.uk/working\\_papers/paper22.pdf](http://www.casa.ucl.ac.uk/working_papers/paper22.pdf)
- Orlovsky, L., Kaplan, S., Orlovsky, N., Blumberg, D., & Mamedov, E. (2006). Monitoring land use and land cover changes in Turkmenistan using remote sensing. 99 463-471.

- Overmars, K. P., Verburg, P. H., & Veldkamp, T. (. ). (2007). Comparison of a deductive and an inductive approach to specify land suitability in a spatially explicit land use model. *Land Use Policy*, 24(3), 584-599.
- Pal, M., & Mather, P. M. (2002). A comparison of decision tree and back propagation neural network classifiers for land use classification. *I* 503-505.
- Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86(4), 554-565.
- Pang, M. Y. C., & Shi, W. (2002). Development of a process-based model for dynamic interaction in spatio-temporal GIS. *GeoInformatica*, 6(4), 323-344.
- Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers*, 93(2), 314-337.
- Perner, P., Zscherpel, U., & Jacobsen, C. (2001). A comparison between neural networks and decision trees based on data from industrial radiographic testing. *Pattern Recognition Letters*, 22(1), 47-54.
- Phelps and Hoppe, 2002 Phelps, M., Hoppe, M., 2002. New York–New Jersey Highlands Regional Study: 2002 Update. NA-TP-02-03, USDA Forest Service, Northeastern Area State and Private Forestry, Newtown, PA.
- Pielke Sr., R. A., Walko, R. L., Steyaert, L. T., Vidale, P. L., Liston, G. E., & Lyons, W. A. et al. (1999). The influence of anthropogenic landscape changes on weather in south Florida. *Monthly Weather Review*, 127(7), 1663-1672.
- Pijanowski, B. C., Brown, D. G., Shellito, B. A., & Manik, G. A. (2002). Using neural networks and GIS to forecast land use changes: A land transformation model. *Computers, Environment and Urban Systems*, 26(6), 553-575.
- Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, 23(5), 479-510.
- Pontius Jr., R. G. (2000). Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing*, 66(8), 1011-1016.
- Power, C., Simms, A., & White, R. (2001). Hierarchical fuzzy pattern matching for the regional comparison of land use maps. *International Journal of Geographical Information Science*, 15(1), 77-100.

- Quinlan, J. R. (1987). Simplifying decision trees. *International Journal of Man-Machine Studies*, 27(3), 221-234.
- Quinlan, J. R. (1996a). Learning decision tree classifiers. *ACM Computing Surveys*, 28(1), 71-72.
- Quinlan, J. R. (1996b). Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Research*, 4, 77-90.
- Quinlan, J. R. (1999). Simplifying decision trees. *International Journal of Human Computer Studies*, 51(2), 497-510.
- Quisenberry, V., Brown, P., Smith, B., & Hallahan, D. F. (2006). In-situ liquid storage capacity measurement of subsurface wastewater absorption system products. *Journal of Environmental Health*, 69(4), 9-15.
- Ramankutty, N., Foley, J. A., Norman, J., & McSweeney, K. (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography*, 11(5), 377-392.
- Razi, M. A., & Athappilly, K. (2005). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Systems with Applications*, 29(1), 65-74.
- Reid, R. S., Thornton, P. K., McCrabb, G. J., Kruska, R. L., Atieno, F., & Jones, P. G. (2004). Is it possible to mitigate greenhouse gas emissions in pastoral ecosystems of the tropics? *Environment, Development and Sustainability*, 6(1-2), 91-109.
- Resnick, M. (1994). 'Changing the centralized mind'. *Technology Review* (July): 32-40.
- Riebsame, W. E., Meyer, W. B., & Turner II, B. L. (1994). Modeling land use and cover as part of global environmental change. *Climatic Change*, 28(1-2), 45-64.
- Rutgers CRSSA, 2004. Rutgers University Center for Remote Sensing and Spatial Analysis. New Jersey Land Cover Mapping: 1972, 1984, 1995, 2001. Retrieved April 25, 2008, from <http://crssa.rutgers.edu/projects/lc/>
- Salmun, H., & Molod, A. (2006). Progress in modeling the impact of land cover change on the global climate. *Progress in Physical Geography*, 30(6), 737-749.
- Seto, K. C., & Kaufmann, R. K. (2003). Modeling the drivers of urban land use change in the pearl river delta, china: Integrating remote sensing with socioeconomic data. *Land Economics*, 79(1), 106-121.
- Silber, R., & Wytrzens, H. K. (2006). *Modelling the probability of land abandonment at parcel level*.

- Shi, W., & Pang, M. Y. C. (2000). Development of voronoi-based cellular automata-an integrated dynamic model for geographical information systems. *International Journal of Geographical Information Science*, 14(5), 455-474.
- Shiyuan, H., and Deren. L. Vector Cellular Automata Based Geographical Entity in Proceedings of 12<sup>th</sup> International Conference on Geoinformatics – geospatial information Research: Bridging the Pacific and Atlantic, University of Gavle, Sweden, June 7-9, 2004. Retrieved December 8, 2007, from <http://www.hig.se/~bjg/geoinformatics/files/p249.pdf>
- Skole, D. L. (1994). Data on global land-cover change: Acquisition, assessment, and analysis. *Changes in land use and land cover: a global perspective*, 437-471.
- Smith, G. M., & Fuller, R. M. (2001). An integrated approach to land cover classification: An example in the island of jersey. *International Journal of Remote Sensing*, 22(16), 3123-3142.
- Speybroeck, N., Berkvens, D., Mfoukou-Ntsakala, A., Aerts, M., Hens, N., & Van Huylenbroeck, G. et al. (2004). Classification trees versus multinomial models in the analysis of urban farming systems in central Africa. *Agricultural Systems*, 80(2), 133-149.
- Soil Survey Geographic Database (SSURGO) (2004)* USDA-NRCS, USDA, Washington, DC. Retrieved December 8, 2007, from <http://www.ncgc.nrcs.usda.gov/products/datasets/ssurgo>
- Official soil series descriptions data access (2003)* U.S. Department of Agriculture-Natural Resources Conservation Service USDA-NRCS, Retrieved December 8, 2007, from <http://ortho.ftw.nrcs.usda.gov/osd/osd.html>
- Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1), 77-89.
- Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sensing of Environment*, 64(3), 331-344.
- Stephenne, N., & Lambin, E. F. (2004). Scenarios of land-use change in sudano-sahelian countries of Africa to better understand driving forces. *GeoJournal*, 61(4), 365-379.
- Stephenson, D. (1994). Comparison of the water balance for an undeveloped and a suburban catchment. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 39(4), 295-307.

- Stevens, D. 2005, Integration of cellular automata approach and geographic information systems for high-resolution modeling of urban growth. Unpublished master thesis. Simon Fraser University, Burnaby, BC, Canada.
- Stevens, D., Dragicevic, S., & Rothley, K. (2007). iCity: A GIS-CA modelling tool for urban planning and decision making. *Environmental Modelling and Software*, 22(6), 761-773.
- Stony Brook-Millstone Watershed Association, 2003 Build out Analysis of West Amwell Township, Hunterdon County, New Jersey. Retrieved February 9, 2008, from <http://www.giscenter.org/resources/library/buildout.westamwell.pdf>
- Story, M., & Congalton, R. G. (1986). Accuracy assessment: A user's perspective. *Photogrammetric Engineering & Remote Sensing*, 52(3), 397-399.
- Sui, D. Z., & Zeng, H. (2001). Modeling the dynamics of landscape structure in Asia's emerging Desakota regions: A case study in Shenzhen. *Landscape and Urban Planning*, 53(1-4), 37-52.
- Takeyama, M., & Couclelis, H. (1997). Map dynamics: Integrating cellular automata and GIS through geo-algebra. *International Journal of Geographical Information Science*, 11(1), 73-91.
- Thomlinson, J. R., Bolstad, P. V., & Cohen, W. B. (1999). Coordinating methodologies for scaling landcover classifications from site-specific to global: Steps toward validating global map products. *Remote Sensing of Environment*, 70(1), 16-28.
- Tobler, W. R. (1979). Cellular geography. *Philosophy in geography*, 379-386.
- Torrens, P. M. (2000) "How cellular models of urban systems work". *CASA Working Paper 28*. University College London, Centre for Advanced Spatial Analysis. Retrieved September 8, 2007, from [http://www.casa.ucl.ac.uk/working\\_papers/paper28.pdf](http://www.casa.ucl.ac.uk/working_papers/paper28.pdf)
- Tso, B. C. K., & Mather, P. M. (1999). Classification of multisource remote sensing imagery using a genetic algorithm and markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3 I), 1255-1260.
- Tulloch, D. L., Myers, J. R., Hasse, J. E., Parks, P. J., & Lathrop, R. G. (2003). Integrating GIS into farmland preservation policy and decision making. *Landscape and Urban Planning*, 63(1), 33-48.
- Turner, B.L. II and B.L. Meyer. 1994. "Global Land Use and Land Cover Change: An Overview." In *Changes in Land Use and Land Cover: A Global Perspective*, eds. W.B. Meyer and B.L. Turner II, 3-10. Cambridge: Cambridge University Press.

- Turner II, B. L., Skole, D., Sanderson, S., Fischer, G., Fresco, L., & Leemans, R. (1995). Land-use and land-cover change: Science/research plan.
- Tyrawski, John. M., NJDEP, Personal Communication, January 8, 2008.
- Underwood, K. L., & Davis, C. L. (2005). GIS-based planning tools to mitigate channel enlargement and water quality impacts resulting from impervious surfaces. Proceedings of the 2005 Watershed Management Conference - Managing Watersheds for Human and Natural Impacts: Engineering, Ecological, and Economic Challenges, p. 1133.
- US Census Bureau. (2005). American Community Survey. Retrieved March 12, 2008, from <http://www.epodunk.com/top10/countyPop/coPop31.html>
- Verburg, P. H., de Nijs, T. C. M., van Eck, J. R., Visser, H., & de Jong, K. (2004a). A method to analyse neighbourhood characteristics of land use patterns. *Computers, Environment and Urban Systems*, 28(6), 667-690.
- Verburg, P. H., Schot, P. P., Dijst, M. J., & Veldkamp, A. (2004b). Land use change modelling: Current practice and research priorities. *GeoJournal*, 61(4), 309-324.
- Verburg, P. H., & Veldkamp, A. (2005). Introduction to the special issue on spatial modeling to explore land use dynamics. *International Journal of Geographical Information Science*, 19(2), 99-102.
- Versace, V. L., Ierodiaconou, D., Stagnitti, F., & Hamilton, A. J. (2008). Appraisal of random and systematic land cover transitions for regional water balance and revegetation strategies. *Agriculture, Ecosystems and Environment*, 123(4), 328-336.
- Walker, R. (2003). Evaluating the performance of spatially explicit models. *Photogrammetric Engineering and Remote Sensing*, 69(11), 1271-1278.
- Wang, F., Hall, G. B., & Subaryono. (1990). Fuzzy information representation and processing in conventional GIS software: Database design and application. *International Journal of Geographical Information Systems*, 4(3), 261-283.
- Ward, D. P., Murray, A. T., & Phinn, S. R. (2000). A stochastically constrained cellular model of urban growth. *Computers, Environment and Urban Systems*, 24(6), 539-558.
- White, R., & Engelen, G. (1993). Cellular automata and fractal urban form: A cellular modelling approach to the evolution of urban land-use patterns. *Environment & Planning A*, 25(8), 1175-1199.

- White, R., & Engelen, G. (2000). High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems*, 24(5), 383-400.
- White, R., Engelen, G., & Uljee, I. (1997). The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. *Environment and Planning B: Planning and Design*, 24(3), 323-343.
- Wu, F. (1996). A linguistic cellular automata simulation approach for sustainable land development in a fast growing region. *Computers, Environment and Urban Systems*, 20(6), 367-387.
- Wu, F. (2002). Calibration of stochastic cellular automata: The application to rural-urban land conversions. *International Journal of Geographical Information Science*, 16(8), 795-818.
- Wu, F. (2003). Simulating temporal fluctuations of real estate development in a cellular automata city. *Transactions in GIS*, 7(2), 193-210.
- Wu, F., & Webster, C. J. (1998). Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design*, 25(1), 103-126.
- Wu, S., Silván-Cárdenas, J., & Wang, L. (2007). Per-field urban land use classification based on tax parcel boundaries. *International Journal of Remote Sensing*, 28(12), 2777-2801.
- Yang, R., & Jargowsky, P. A. (2006). Suburban development and economic segregation in the 1990s. *Journal of Urban Affairs*, 28(3), 253-273.
- Yeh, A. G., & Chow, M. H. (1996). An integrated GIS and location-allocation approach to public facilities planning - an example of open space planning. *Computers, Environment and Urban Systems*, 20(4-5), 339-350.
- Yeh, A. G., & Li, X. (1998). Sustainable land development model for rapid growth areas using GIS. *International Journal of Geographical Information Science*, 12(2), 169-189.
- Yeh, A. G. -, & Li, X. (2001). Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. *Photogrammetric Engineering and Remote Sensing*, 67(1), 83-90.
- Zadeh, L. A. (1972). Fuzzy-Set-Theoretic interpretation of linguistic hedges. *J Cybern*, 2(3), 4-34.

Zhao, J., Chen, H., & Xu, H. (2007). Simulation of land use evolution based on cellular automata and artificial neural network. *Tongji Daxue Xuebao/Journal of Tongji University*, 35(8), 1128-1132.