

2014

## Application of a Spatially Explicit, Agent-Based Land Use Conversion Model to Assess Water Quality Outcomes under Buffer Policies

Nazia N. Arbab

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

---

### Recommended Citation

Arbab, Nazia N., "Application of a Spatially Explicit, Agent-Based Land Use Conversion Model to Assess Water Quality Outcomes under Buffer Policies" (2014). *Graduate Theses, Dissertations, and Problem Reports*. 5118.

<https://researchrepository.wvu.edu/etd/5118>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact [researchrepository@mail.wvu.edu](mailto:researchrepository@mail.wvu.edu).

**Application of a Spatially Explicit, Agent-Based Land Use Conversion Model  
to Assess Water Quality Outcomes under Buffer Policies**

**Nazia N. Arbab**

**Dissertation submitted to the  
Davis College of Agriculture, Natural Resources and Design  
at West Virginia University  
in Partial Fulfillment of the Requirements for the Degree of**

**Doctor of Philosophy  
in  
Natural Resource Economics**

**Alan R. Collins, Ph.D., Chair  
Jamison Conley, Ph.D.  
Michael P. Strager, Ph.D.  
Gregory A. Elmes, Ph.D.  
Jerald J. Fletcher, Ph.D.  
Timothy T. Phipps, Ph.D.**

**Agricultural and Resource Economics Program  
Division of Resource Management  
West Virginia University**

**Morgantown, West Virginia  
2014**

**Keywords: Spatially Explicit Land Use Conversion, Agent-Based Model, Spatial Logistic  
Regression, Monte Carlo Simulations, Water Quality, ArcSWAT**

**Copyright 2014 Nazia N. Arbab**

# ABSTRACT

## **Application of a Spatially Explicit, Agent-Based Land Use Conversion Model to Assess Water Quality Outcomes under Buffer Policies**

Nazia N. Arbab

Land use changes within watersheds have spatially explicit dynamics and involve decision making by individuals. The role of the spatial dimension of human behavior and its impact on land use change has been analyzed using agent-based modelling approaches. Agent-based land use change has received a significant theoretical attention; however, these models lack empirical implementation and testing due to the lack of spatial modelling tools and data that can capture human land use dynamics.

This research presents a methodology for projecting land use conversions through the implementation of a spatially explicit agent-based simulation model in the Opequon Creek watershed of Berkeley County, West Virginia. Empirical estimates for factors that influence the land use conversion probability are captured using a spatial logistic regression model. Then, agent-based probabilistic land use conversion (APLUC) model is programmed on Python language within a geographic information system (GIS) to explore the impacts of policies on land use conversion decisions using estimates from actual land use change from 2001-2011. A series of model runs are executed under buffer policy scenarios. Three policy scenarios are developed: (1) a scenario where there is no policy implemented, (2) a scenario where 50 ft buffer zones are applied to all streams, and (3) a scenario where 50 ft buffers are applied only on critical source areas (CSAs) watersheds. The land use patterns project in APLUC model are driven by individual land conversion decisions over 50 model runs of 10 iterations each under each policy scenario. The APLUC model is validated at sub-basin level and outcomes are analyzed to identify the influence of various land use policies on land use patterns. The results show that a 50 ft buffer policy everywhere in watershed, greatly reduced the residential land use conversions. Spatial patterns generated under a 50 ft buffer policy in CSAs only showed that future projected land use changes occurred close to major highways. In the baseline policy, most conversions occurred near existing residential land use and urban centers. Results from the APLUC model also suggests that forest is serving as “distant amenity” for residential land conversion.

Finally, the impacts of these three policies on water quality are estimated using an ArcSWAT model, a graphical user interface for SWAT (Soil and Water Assessment Tool). This model indicates that the 50 ft buffer policy in CSAs is most effective among the three policies in reducing the pollutant loads. This study suggests that carefully designed policies, which discourage residential land use conversion in CSAs, result in less pollutant loads by shifting the location of residential conversion to less critical areas where agricultural land is dominant in the watershed.

## ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my advisor, Dr. Alan Collins, for his constant encouragement, direction and patience throughout the entire process of my dissertation work. Without his outstanding guidance and unconditional help, this dissertation would not have been possible.

I would like to thank Dr. Jamison Conley for the excellent guidance and expertise regarding the model implementation, and python programming. His useful inputs on empirically strengthening my research were invaluable. I gratefully acknowledge the useful suggestions and advice of Dr. Mike Strager on watershed quality modeling and spatial data. My appreciation also goes to Dr. Gregory Elmes for providing his great insight on the theoretical framework of the model. Special thanks to Dr. Jerald Fletcher and Dr. Timothy Phipps for their useful suggestions, support and guidance.

I also would like to thank Sharon Waltman and Dr. James Thompson from the Geospatial Research Unit, WVU for their support and help in many ways over the years. Furthermore, I am also thankful to administrative staff including Lisa, Ellen, Melanie, Ben, Mark, and Dan for their constant help and care.

I have also enjoyed the great fellowship of my friends and fellow graduate students who offered encouragement, support, and help.

My friend Marc provided a thorough edit for which I am very grateful.

I am forever thankful to my dear friend Andy for his friendship and great humor. His memory will be eternal.

Most importantly, I wish to thank my father and siblings for providing extraordinary support, love, strength and encouragement in pursuing my goals.

*To the memory of my beloved Mother, Rasheeda Tabassum who taught me to pursue greater understanding and knowledge.*

# Contents

<b>1. Introduction</b> .....	1
<b>1.1 Background</b> .....	1
<b>1.2 Problem Statement</b> .....	3
<b>1.3 Study Area</b> .....	7
<b>1.4 Objectives</b> .....	10
<b>1.5 Research Questions</b> .....	11
<b>Chapter 2</b> .....	13
<b>2. Literature Review</b> .....	13
<b>2.1 Approaches to Modeling Land-Use/Cover Change in Economic Framework</b> .....	13
<b>2.2 Spatially Explicit Models of Land Use Change</b> .....	15
<b>2.3 Cellular Automata (CA) Land Use Change Models</b> .....	17
<b>2.4 Agent-based Models (ABM) of Land Use Change</b> .....	18
<b>2.5 Coupled ABM Models of Land Use</b> .....	21
<b>Chapter 3</b> .....	25
<b>3. Theoretical Framework</b> .....	25
<b>Chapter 4</b> .....	32
<b>4. Conceptual Framework</b> .....	32
<b>4.1 A Spatial Logistic Regression (SLR) Model</b> .....	34
<b>4.2 Agent-Based Probabilistic Land Use Conversion (APLUC) model</b> .....	39
<b>4.3 ArcSWAT Model</b> .....	48
<b>4.4 Policy Determination</b> .....	53
<b>4.5 Baseline Policy</b> .....	54
<b>4.6 Implementation of 50 foot Riparian Buffer Zone Policy on every Stream in the Watershed</b> .....	55
<b>4.7 Implementation of 50 foot Riparian Buffer Zone Policy in CSAs</b> .....	56
<b>Chapter 5</b> .....	59
<b>5. Model Validation</b> .....	59
<b>CHAPTER 6</b> .....	69
<b>6. Data Description</b> .....	69
<b>6.1 General Data</b> .....	69
<b>6.2 Spatial Logistic Regression Data</b> .....	71

<b>6.3 Data for ArcSWAT</b> .....	73
<b>Chapter 7</b> .....	75
<b>7 Results</b> .....	75
<b>7.1 Spatial Logistic Model Analysis and Results</b> .....	75
<b>7.2 Agent-based Probabilistic Land Use Conversion Model Results</b> .....	85
<b>7.3 ArcSWAT Model Results</b> .....	95
<b>Chapter 8</b> .....	106
<b>8. Conclusions, Limitations and Future works</b> .....	106
<b>References</b> .....	114
<b>APPENDICES</b> .....	144
<b>APPENDIX I: Spatial data</b> .....	145
<b>APPENDIX II: Variation in Conversions across 50 Model Runs</b> .....	158
<b>APPENDIX III: Monthly Mean Sediment, Mean TP and Mean TN Loads for 42 Sub-basins</b> <b>Estimated by ArcSWAT over 10 years.</b> .....	162
<b>APPENDIX IV: Land Use, Soil, and Slope from ArcSWAT</b> .....	170
<b>APPENDIX V: Requirements by Berkeley County Planning Commission, 2009</b> .....	183
<b>APPENDIX VI: Code for the APLUC in Python using ArcPy (ESRI, 2014) (Baseline scenario)</b> .....	184

## LIST OF TABLES

Table 1. Variables and definitions of pollutant loads in ArcSWAT. ....	52
Table 2. Comparison among actual and projected residential conversions at parcel level. ....	65
Table 3. Spearman correlation analysis between actual pollutant loading and projected pollutant loadings for 2011. ....	68
Table 4. Land uses/land cover classes and their description, (EPA, 2007; USGS, 2014). ....	70
Table 5. Descriptive statistics of parcel size for the Opequon Creek watershed. ....	71
Table 6. Explanatory variables descriptions and mean values. ....	73
Table 7. Spatial logistic regression model results. ....	76
Table 8. Classification of cases & odds ratio using 0.5 threshold. ....	78
Table 9. Classification of cases & odds ratio using 0.11 threshold. ....	78
Table 10. The Likelihood ratio statistics and chi-square significance for each predictor. ....	81
Table 11. Marginal effects, change in odds and percentage change in the odds. ....	83
Table 12. Descriptive Statistics of parcel conversions over 50 model runs in each policy scenario. ....	87
Table 13. Percentage of each land use type in watershed area under each policy scenario. ....	96
Table 14. Average pollutant loads per month for the Opequon Creek watershed projected over a 10 year period. ....	97



## LIST OF FIGURES

Figure 1. Location map of the Opequon Creek within Berkeley County, WV .....	7
Figure 2. Conceptual framework for this research.....	33
Figure 3. Logistic function.....	35
Figure 4. Decision algorithm of agent-based probabilistic land use conversion model .....	48
Figure 5. ArcSWAT model procedure for the Opequon Creek Watershed .....	50
Figure 6. Additional actual residential parcels in 2011.....	60
Figure 7. Projected 2011 land use conversions.....	62
Figure 8. Correctly projected land use conversions .....	62
Figure 9. Potential growth zone .....	63
Figure 10. Projected land use conversions within growth zone.....	63
Figure 11. Residential Land Use Projection .....	65
Figure 12. Sediment yield in each sub-basin for actual and projected 2011 land use data.....	66
Figure 13. TN yield in each sub-basin for actual and projected 2011 land use data.....	66
Figure 14. TP yield in each sub-basin for actual and projected 2011 land use data .....	67
Figure 15. Predicted land use conversion global probability.....	84
Figure 16. Number of parcels converted from each developable land use type .....	87
Figure 17. Area of land converted from each developable land use type .....	88
Figure 18. Land use conversions under baseline policy .....	90
Figure 19. Land use conversions under 50 ft agricultural buffer scenario.....	92
Figure 20. Land use conversions under 50 ft forest buffer scenario.....	92
Figure 21. Land use conversions under 50 ft agricultural buffers in CSAs .....	93
Figure 22. Land use conversions under 50 ft forest buffers in CSAs .....	94
Figure 23. Land use distribution by area [ha] in each scenario .....	96
Figure 24. Dominant land use within each sub-basin delineated by ArcSWAT under each scenario .....	98
Figure 25. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with actual 2011 land use data.....	100
Figure 26. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with baseline policy scenario .....	100
Figure 27. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT in 50 ft agricultural buffer everywhere.....	101
Figure 28. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT in 50 ft forest buffer everywhere .....	101
Figure 29. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with 50 ft agricultural buffer in CSAs .....	102
Figure 30. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with 50 ft forest buffer in CSAs .....	102

# Chapter 1

## 1. Introduction

### 1.1 Background

Land use change within watersheds is coupled to socio-ecological systems in which human decisions to convert land from one type to another impact natural systems. Entities in natural resources system are interconnected and influenced by biophysical and socio-economic conditions, spatial juxtaposition of neighboring land uses and location specific characteristics (Parker & Tatianaa, 2008). The interconnectedness of socio-ecological factors in such systems results in complexity occurring within both spatial and temporal dynamics (Lambin et al., 2003; Rammel et al., 2007). Changes at the macro level such as infrastructure, neighboring features, and policy interventions, affect the land use decisions of the individuals that produce micro-scale localized changes (Bell and Irwin, 2002). Aggregation of these micro-scale changes result into emergent spatial patterns of large scale land use changes, which impact biophysical processes. Individuals as intrinsic entities in such non-linear, interactive and transformative land use systems can tell us much about the spatial expression of human relationships with natural systems.

Most land use change in the U.S. has been a result of agricultural or forest land being converted into residential development due to changes in socio-economic factors such as population and income growth (Polyakov and Zhang, 2008; Alig, 2010). Rural areas closer to metropolitan and urban centers, having less strict zoning criteria and lower property values are more likely to have a rapid land use change (Goetz et al., 2003; White et al., 2009; Qian, 2010, HUD, 2012). In the context of potential consequences of land use change, the significance of residential development on hydrological systems have been observed in previous studies (Tong and Chen, 2002; Coutu and Vega, 2007).

One of the major consequences of residential development in watersheds is increased impervious surfaces due to additional paved areas such as streets, parking lots, curbs, sidewalks and driveways. These impervious surfaces alter the characterization of morphological features of watersheds and result in impaired streams (Weng, 2001; Corbett et al., 1997). Impervious surfaces lack infiltration, which facilitates the transportation of non-point source pollution, elevates water runoff, and soil loss during rain events. In several watersheds of the U.S., rapid urbanization has been matched by stream quality degradation due to decreased permeability (Bhaduri et al., 2001; Schueler, 2009; Mejia et al., 2014).

The most common non-point pollutants in watersheds causing disruptions in biophysical functions of hydrological systems have been identified as sediment, nitrogen, and phosphorus (Carpenter et al., 1998; Niraula et al., 2013). For example, research has shown that excessive phosphorus and nitrogen are entering the Chesapeake Bay estuary from its surrounding tributaries (Kaushal et al., 2011; Duan et al., 2012). Increased non-point source pollution in the Chesapeake Bay's network of streams and rivers has been influenced by the high percentage of developed land due to urbanization (Jantz et al., 2005; Dauer et al., 2000).

As urban development including high residential density areas increases, watershed hydrology changes due to changing runoff, peak flow and infiltration of sediments (Carlson, 2004; Coutu and Vega, 2007). Watershed hydrology spatially varies with changing land uses (Anbumozhi et al., 2005; Tu, 2011; Niraula et al., 2013; Sun et al., 2014). Several studies suggest that water quality is more sensitive to land uses near streams when compared to land uses over an entire watershed area (Osborne and Wiley, 1988; Hunsaker and Levine, 1995; Johnson et al., 1997). Ecological studies have examined the effectiveness of vegetative riparian buffers in filtering

pollutants entering from nearby lands (Johnson et al., 1997; Broadmeadow and Nisbet, 2004; Jontos, 2004).

Often coupled land use water quality studies focus on percentage or proportions of each land use type such as urban, forest, agriculture and wetlands from the past or current land use data within watershed (Sliva and Williams 2001; Jung et al., 2008; Lee et al., 2009; Li et al., 2012) . Therefore, the land parcels as choice making unit is ignored at watershed scale. A spatially-explicit model of future land use patterns can provide useful information to watershed management. A simulation model using spatially explicit data in discrete land units is a relatively new modeling technique to predict urban development. Clarke and Gaydos (1998) have developed a cellular automaton model to predict urban growth with simulation modeling framework using geographic information system (GIS) data. The simulation modeling technique provides a methodology to project land use conversions of each location and helps to forecast future human impact on hydrological processes.

This research examines the complexity of land use change and water quality by analyzing factors that influence the parcel level land use change and then modeling projected land use patterns under riparian buffer policies. Finally, land use patterns are linked with a hydrological watershed model to assess the effects of land use conversions on watershed quality.

## **1.2 Problem Statement**

In the ecological literature, relationships between land use change and drivers of land use change have been explored to estimate the impacts of these changes on natural systems using the indicators of human impacts (Gergel et al., 2002; Allan, 2004). These human indicators were imposed exogenously such as estimating the urban land use area within a catchment (Miller and Plantinga, 1999; Sliva and Williams, 2001; Gergel et al., 2002). However, land use changes due to

urbanization are inherently related to land use conversion decisions made by land owners, developers, planners, and policy makers. These decisions influence the state of the land at a specific location.

One of the challenges in land use research is that macro patterns can be quantified but the decisions cannot be observed or measured directly. The characterization of decision making in land use models requires location specific information as well as the responses of individuals to policies. A typical land use decision taking place at the parcel level involves the probability of land conversion from one discrete use to another as a function of neighboring land use ( e.g., agriculture and forest) and location characteristics ( e.g., distance to the urban center) (Bockstael, 1996; Bockstael and Bell, 1998).

Analyzing the causes and consequences of land use conversion decisions at the parcel level provides spatially explicit details of land use change driven by factors such as proximity and spatial spillover effects at disaggregate scale (Irwin and Geoghegan, 2001; Sidharthan and Bhat, 2012). Knowledge about probability of land use conversion from non-developed to residential provides key insights to environmental management and designing land use planning and policy. This micro-scale information can be useful in informing policy makers of projected land use patterns under different policy scenarios.

Among spatially-explicit land use models, the class of cellular automata (CA) models provides a simulation framework for modeling land use conversion decision making where landscape is divided among equal-sized cells. Typically, the CA models predict the land use patterns, which is driven by transitional probabilities, and use cell or pixel in raster based grid as a unit of analysis. However, in reality, the decision to change land from one use to an alternative use are made within land parcel boundaries. The land use conversion decision at the cell level

treats the cells as independent entities (observations) within the same parcel boundary and results in biased interactions among entities (Irwin, 2010).

An agent-based model (ABM) framework in GIS (Geographical Information Systems) has the capability to model spatially explicit land use decision making at the actual decision making unit such as plot or land parcels (Najlis and North, 2004; Johnston, 2013). An ABM is a simulation modelling framework where, various behaviors of individuals or agents are programmed through their interactions and choice principles (Gimblett et al., 2002; O'Sullivan et al., 2012). Parker et al., (2003) suggested that multi-agent systems of land use change are useful in analyzing the feedback of agents to their environment within complex systems. To assess the impact of human behavior on natural resource systems, the land use patterns generated from a spatially explicit, land use ABM can be used as an input into models of water quality, ecological systems, or environmental economics (Ng et al., 2011; Guzy, 2008; Heckbert et al., 2010).

Within spatially explicit ABM of land use, neighboring land uses and their impact on land parcel conversion provides an implicit interdependency and interaction among agents (Irwin, 2010; Irwin and Geoghegan, 2001). Previous studies have investigated these neighboring impacts in land use conversion using Cellular Automata (CA) and ABMs (Parker and Meretsky, 2004; Polhill et al., 2008). Due to the unavailability of micro scale data (property parcels) in these models, the parameter measures that influence the value of land parcels reflected only relative neighboring land use impacts which were based on researcher generated estimates (Parker and Meretsky, 2004; Polhill et al., 2008).

In the real world, the neighboring land use and site specific characteristics are important to residential land use conversion. One way to improve estimates of neighboring impacts is through an empirical parameterization of spatial externalities using spatial data obtained from a study area.

With this empirical approach, spatial externalities represent quantified impacts of neighboring land use and location features on land use conversion decisions. Incorporating empirically estimated local scale influences in agent-based models for projecting the land use conversion, requires both a spatially explicit (GIS environment) agent-based and an empirical modelling framework that link both quantified neighboring externalities and location specific influences to parcel based land use conversion decisions.

This research seeks to contribute to our understanding of the processes of land use change and water quality outcome effects of these land uses changes using a linked modelling approach. First, spatial logistic regression analysis is used to estimate the spatial externalities (neighboring impacts and proximity factors) influencing residential land use conversions. These empirically based estimates are then implemented into an ABM for land use conversions. This ABM is designed to simulate land-use decisions of agricultural and forest land owners to convert land from non-develop land to residentially developed land as a dichotomous choice. Lastly, the results from the ABM are linked with a water quality model to provide key understandings of land use processes and watershed hydrology under various policy scenarios. These land use patterns are applied to a GIS extension of the Soil and Water Assessment Tool (ArcSWAT) to project the impact of land use conversion on watershed hydrology in terms of the amount of nonpoint pollutant loadings from sub-watersheds.

Modeling allows for simulations of land use policy scenarios. Three types of policy scenarios are applied to examine the impact of residential land use conversions on water quality. These policy scenarios include: (1) a baseline policy with no restriction on land use conversion, which is the current land use policy without any additional regulations of protected areas; (2) establishing a protection zone of a riparian buffer of 50 foot placed on each side of all the streams

in the study area; and (3) establishment of a 50 foot buffer zone on each side of streams only with high priority sub-watersheds which are identified as critical source areas in watersheds. The buffer zone policies are aimed to mitigate the water quality impacts of land use conversion.

### 1.3 Study Area

The study area comprises the Opequon Creek watershed area located in Berkeley County, West Virginia (WV) (Figure 1). Opequon Creek starts near Winchester Virginia, flows to the north into West Virginia, and drains into the Potomac River. In West Virginia, the majority of Opequon Creek falls into Berkeley County, WV, and covers a drainage area of 38,100 hectares.

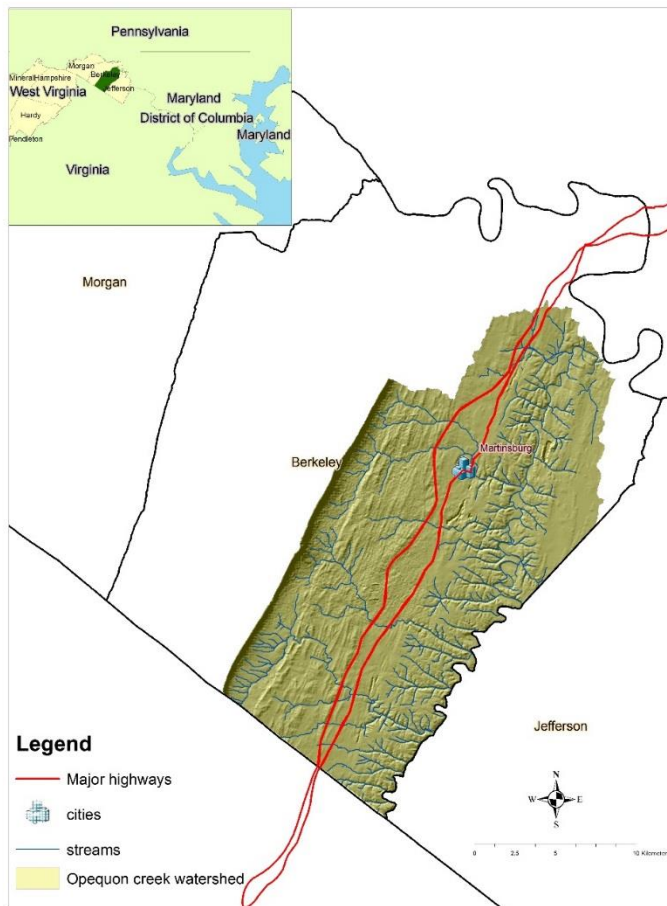


Figure 1. Location map of the Opequon Creek within Berkeley County, WV



Berkeley County is transitioning from a rural area into suburban communities for the Washington DC metropolitan area (Stocks, 2010). This county has the second largest population growth of all counties in West Virginia (Berkeley County Development Authority, 2014). From 2000 to 2010, there was a 37.2 % increase in Berkeley County's population (Berkeley County Development Authority, 2014; Christiadi, 2011). From 2010 to 2012, the proportion of the net in-migration into the city of Martinsburg, the largest city in Berkeley County was approximately 66 % of population growth (HUD, 2012). Recent population projection estimates of West Virginia show an expected growth rate of 1.3 % per year for Berkeley County between 2010 and 2030 (BBER, WVU, 2014). The increase in net-in-migration has been attributed to affordable housing and proximity to Washington, D.C and Hagerstown, Maryland (HUD, 2012). This geographic location places Berkeley County in a desirable housing market. The Opequon Creek Project Team reported a high percentage of building permits in Berkeley County during the year 2004 (The Opequon Creek Project Team, 2014). This high percentage is the result of increased population and minimum requirements of the subdivision by County government. Another report shows that between 2004 and 2006, the County approved several subdivisions that result into 6,985 lots and developed 3,653 acres of land (Goodspeed, 2007). Since 2000, the number of lots receiving final approval from the County has increased (Goodspeed, 2007).

A major concern of this residential growth is its impact on the Opequon watershed, which is already showing an increased level of phosphorus and nitrogen due to extensive farming (VT CTMDLWS, 2006; Karigomba, 2009). In terms of its location specific significance, Opequon Creek is a tributary of the Potomac River. Among the Potomac tributaries in West Virginia, Opequon Creek has the highest priority for restoration due to its elevated nutrients and sediment level (WVDEP, 2008; Water Resources and TMDL Center, 2008). According to the Opequon

Creek project team report (2009), WVDEP identified Opequon Creek and 13 of its tributaries in West Virginia as “impaired.” The criteria for the impairment in the entire Opequon Creek, WV is based on the total maximum load (TMDL) of nutrients, fecal coliform bacteria and sediments (Opequon Creek IP Steering Committee, 2006; WVDEP, 2008). These pollutants likely originate from both point sources such as treatment plants and non-point sources such as developed and agricultural land. Most of the bank erosion in Opequon Creek is the result of the increased urbanization and impervious surfaces (Water Resources and TMDL Center, 2008).

Recent efforts of local government and the EPA has diverted towards the problem of “non-point source” pollution due to storm water run-off from surfaces such as roofs, parking lots, streets, and sidewalks. Reduced riparian cover and increased impervious surface area from urban growth and development cause stream channel erosion and stream flow instability in Opequon Creek area (WVDEP, 2008). The general growth management plan for Berkeley County divides the County into zoning districts and suggests density requirements for each district (Berkeley County Planning Commission, 2006). The Berkeley County Planning Commission (2006) is looking at the possibility of creating a buffer zone along the streams in critical sub-watersheds within the Opequon Creek watershed.

The Opequon Creek watershed as sub-region of Berkeley County serves as a good case study area because of the existing non-point source pollution problem and the heterogeneous exurban landscape that includes agricultural and forested areas. In order to implement a policy within Berkeley County, the factors that need to be considered should not only be related to indicators of residential development but to what causes the development. The analysis of human impacts on watershed requires the identifications of stressors at each sub-basin or hydrological response unit.

## **1.4 Objectives**

The primary goal of this study is to demonstrate how modelling of individual decision-making processes of land use conversion can provide a better understanding of emerging land use patterns under various land use policies, and to assess the impacts of these policies on surface water quality. Accordingly, the objective of this research is not only to project land use patterns, but to provide an explanation of the land use patterns through location specific factors and spatial externalities. To achieve this goal, a model of land use conversion is created with empirical parameterization from actual land use conversion data obtained from the study area. The choice of land use conversion model is based on the recognition that land use conversion decisions are probabilistic discrete choices (Bockstael, 1996). The discrete choice model offers forward looking patterns of land use under different policy regimes. The resulting land use patterns are then used as a database for an assessment of water quality impacts to Opequon Creek.

The following goals are met in order to achieve the primary objective of this study.

1. Calculate empirical parameters for spatial externalities of land use conversion for inclusion into an agent-based model.
2. Build an agent-based, probabilistic land use model in a spatially explicit environment using property parcels as the unit of analysis.
3. Use an agent-based model to project patterns of residential land use changes within Berkeley County, WV by simulating multiple agents' decisions to develop land parcels.
4. Use spatially determined residential land use changes to simulate impacts on the transport of sediment nitrogen and phosphorus in the Opequon Creek watershed of Berkeley County, WV by linking an agent-based model of land use conversion with the ArcSWAT model.

5. Analyze and compare the outcomes and relative importance of different land use policy scenarios on land use patterns and surface water quality outcomes.

## **1.5 Research Questions**

The key research questions to be addressed in this study are:

- (i) What spatial factors cause changes in spatial pattern of land use in the Opequon Creek Watershed during the period of a decade?
- (ii) Do empirically driven estimates explain the local level probabilities of residential land use conversions and if so, how effectively does it address the watershed level land use processes and patterns?
- (iii) How do the linkages of the three models respond to conditions reflected in the policies?
- (iv) How much does location matter for environmental policy such as water quality improvement?

## **1.6 Organization of this Study**

The research is organized into eight chapters. The current chapter provides the relevant background and the problems with existing approaches of modeling the impacts of land use change. This chapter discussed the selection of the study area and land use change and water quality issues and indicates the research objectives and research questions. In Chapter 2, selective studies are reviewed for their methodological and empirical frameworks. This chapter covers the strengths and shortcomings of different modeling approaches. Chapter 3 emphasizes theoretical frameworks. Chapter 4 presents a detailed overall model structure for this research. Discussions include how the land use conversions are modelled within an ABM framework using empirically driven parameters and the linkages between projected land use patterns and the water quality

model. This chapter also discusses riparian buffer policy scenarios suggested for the study area. Chapter 5 describes the validation approaches utilized for the ABM in this study. Chapter 6 provides a description of the data. The results of each model will be discussed in Chapter 7. Finally, Chapter 8 summarizes the findings, contributions, limitations of the current approach, and future research directions.

# Chapter 2

## 2. Literature Review

Land use change has been investigated from both the economic and socio-ecological disciplines (Irwin and Geoghegan, 2001; Veldkamp et al., 2011). In general, economic models of land use have focused on the value of land as influenced by location specific characteristics and the price of land as derived from the equilibrium of demand and supply (Palmquist, 1989; Madison, 2000; Buurman et al., 2001). Socio-ecological studies have utilized approaches that emphasize stressors of land use change on ecological systems (Turner, 1989; Alberti and Waddell, 2000; Lambin and Meyfroidt, 2010). This chapter discusses major theoretical and empirical studies from economic and socio-ecological fields and their respective techniques to model and predict land use change.

### 2.1 Approaches to Modeling Land-Use/Cover Change in Economic Framework

Von Thünen is (1826) best-known for his classic work “Der Isolierte Staat” and was among the first inquirers of land use patterns with relation to location. Von Thünen introduced the concept of economic land rent as a function of distance from the market within the context of an agricultural economy. His location theory of land use focuses on the significance of an urban center with surrounding patterns of agricultural land. Transportation cost is a primary factor influencing economic rent of land and is defined in terms of distance from the central business district (CBD). Thus, the competition among various agricultural lands is due to the proximity to the CBD. The representation of classic Von Thünen theory was based on the following assumptions (Weinschenk et al., 1969):

- a) There is only one central business district (CBD) ,
- b) The CBD is an isolated center, with no trade and the economic system is similar to a closed economy,

- c) Land is spatially uniform in terms of productivity and exists upon a featureless plain (non- heterogeneous landscape), and
- d) Transportation cost is a linear function of distance.

Although the Von Thünen approach is only applicable under strict assumptions, the model places great emphasis on spatial location variation in determining economic rent. The simplified concept of economic rent is central to subsequent land use models in urban economics.

Newer versions of the Von Thünen model are monocentric bid-rent models (monocentric model) of urban economics (Alonso, 1964; Mills, 1967; Muth, 1969). Based upon Von Thünen's framework, Alonso (1964) formalized monocentric bid-rent theory, in which individual utilities are a function of accessibility to an urban center. The model suggests that individuals are willing to pay the land bid-rent that maximized their utility based upon accessibility to the urban center.

Further extensions of monocentric land use theory were developed by incorporating the transportation infrastructure in defining population density at each residential location around the CBD (Mills, 1967; Muth, 1969). Location choices in these models are based upon the relationships between time travel and urban density. These locational choice models provide a great insights into how the spatial arrangement of economic infrastructure influences the choices and utility functions of individuals, but they ignore the heterogeneity in land use which are location specific and not restricted to accessibility. In this regard, the hedonic modeling approach offers a framework that accounts for the importance of spatial amenities and externalities from the site specific attributes and surrounding land use in defining the property parcel value (Bockstael, 1996; Geoghegan et al., 1997; Bastian et al., 2002; Irwin, 2002).

Bockstael (1996) combined the hedonic approach with land use decision making of developers in predicting the probability of land use conversion from non-developed to residential

land use in the Patuxent Watershed. The characterization of decision making was based upon land unit value. This approach was used to estimate the value of the land parcel based upon two major features: (1) proximity to infrastructure such as distance to population centers, roads and metropolitan area and (2) surrounding land use patterns. Bockstael's findings suggests that economic centers, open space, and pasture land have positive influence on surrounding lands.

Geoghegan et al. (1997) emphasized an application of spatially explicit data in an empirical hedonic model to explain spatial heterogeneity when estimating residential land value within a 30 mile radius area of Washington DC. This study found that along with parcel specific features, the neighboring land use, distance to urban center and natural amenities are significant in determining the value that individuals place on land use. They also found that land value varies according to spatial diversity and fragmentation across different scales such as rural, urban, and suburban regions. Their findings suggest that open space is a positive amenity at a smaller scale, but acts as a negative amenity at a larger scale.

The concept of hedonic price is based upon the equilibrium outcomes of demand and supply, which result in one dimensional measure of distance from location (Bell and Irwin, 2002). In this regard, hedonic models can be combined with micro simulation models such as UrbanSim (Waddell, 2002) in order to provide spatially diverse information by assigning choice behavior of land allocation to intrinsic entities in the system (Waddell, 2002; Waddell et al., 2003).

## **2.2 Spatially Explicit Models of Land Use Change**

Often socio-ecological models use simulation techniques to understand the behavior of entities through the decision rules to predict land use changes (Verburg et al., 2002; Arsanjani, 2013). Land use change analysis in these models often is based upon spatially explicit data (Mertens and Lambin, 1997; Nowak and Walton, 2005; Radeloff et al., 2012; Sohl et al., 2014). Drivers of land use change within a socio-ecological framework are based upon proximity measures, biophysical



variables such as soil, slope, and land productivity, and socio-economic data such as population size, density, and economic conditions (Veldkamp and Fresco, 1996; Vionov et al., 1999).

Selection of data is crucial in defining land use processes at different scales. On an aggregate level, determining factors of land use to allocate land in different uses can be measured through aggregate data (Alig and Healy, 1987; Nagubadi and Zhang, 2005). Parcel level or disaggregated data is well suited for explaining land use conversion decisions among alternative uses at a choice making unit (Carrión-Flores and Irwin, 2004; Irwin and Bockstael, 2002). Models that utilize individual level data base to explain land use conversion are generally estimated under a discrete choice framework (Bockstael, 1996; Polyakov and Zhang, 2008). A discrete choice model can empirically estimate the probability of conversion of any particular unit, as a function of driving factors of land use change (Geoghegan et al., 2001; Polyakov and Zhang, 2008).

The prediction of land use change is often defined in socio-ecological research within a deterministic or probabilistic frameworks. A deterministic modeling framework is derived through the use of defined transition rules at each discrete location to investigate the evolving spatial composition of a landscape (White and Engelen, 1993; Balzter et al., 1998; Ozah et al., 2012). This results in emerging complex behavior from simple empirically quantified local rules (Manson, 2001). On the other hand, probabilistic models introduce uncertainty and probability of land use change at each location using micro-simulation techniques (Almeida et al., 2003; Batty, 2012).

Among micro-simulation land use models, both CA and ABM have the capability to generate land use patterns by incorporating spatial variations and interactions among entities in the system (Clarke and Gaydos, 1998; Gimblett, 2002; Irwin and Bockstael, 2002; Heppenstall and Crooks, 2012).

## 2.3 Cellular Automata (CA) Land Use Change Models

Cellular automata (CA) includes mathematical simulation models which are based upon a system behavior of landscape from transition rules (Clarke and Gaydos, 1998; White and Engelen, 2000; Feng & Liu, 2013). The landscape in CA model typically consists of a matrix of grid cells. A change in the discrete state of each cell results from the previous state and the state of neighboring cells according to some set of deterministic or probabilistic choice rules. In a probabilistic CA model, Markov models are often integrated with CA models to generate the probability of land use change from defined variables (Balzter et al., 1998; Li and Reynolds, 1997). The application of CA in land use modeling efficiently captures the scale effects of land use change projections as discrete events (per month, year, or conversion event) using a spatially interactive environment. Spatial heterogeneity is characterized by dividing the landscape into a lattice of cells based upon several attributes of the actual landscape. This functionality allows location specific factors, and neighboring land uses to influence the transition in cell state using empirical techniques (Clarke and Gaydos, 1998; White and Engelen, 2000). Empirical techniques to estimate probabilities of land conversion have included spatially explicit econometric equations such as a logit model (Wu, 2002) and a Bayesian functional form (Almeida et al., 2003).

In CA models, the decision making of land use conversion is exogenously expressed through a random number of agents in the lattice cells landscape and defining the transition rules as a surrogate for decision making. However, agents in these models are not essentially belong to the observed location resulting in biased spatial relationships between neighboring locations (Parker et al., 2003). A limitation of this approach from the perspective of an economic land use modeling is that quantification of decisions occurs at individual cells or simplification of decision units into a tessellated landscape, instead of actual decision making individuals or decision making

units. Data on the actual unit of decision making such as boundaries of property parcels instead of boundaries between equal sized square cells is important in distinguishing the policy impacts among various land owners. CA models are good representations of the biological and ecological phenomena of land use systems by reproducing the rules, but to understand behavior in a land use process requires exploration of methods that represents decision making at actual choice making level (Parker et al., 2003). In this regard, ABM offers the flexibility to represent agents at their location and can be combined or defined in a CA modeling framework.

## **2.4 Agent-based Models (ABM) of Land Use Change**

The use of agent-based modeling in land use systems is appropriate due to the fact that land itself has many attributes such as slope, soil type, productivity, property parcel size, accessibility, and buildable structure. Due to its multiple attributes, individual units of land can itself be regarded as agents (Le et al., 2010).

Several studies analyzed emerging land use patterns as an outcome of land use systems. These systems have represented markets, social integration and segregation, or policy regimes (Berger, 2001; Parker and Meretsky, 2004; Mathews et al., 2007). The aggregate land use patterns depend upon interdependencies between agents with landscape and other agents plus agent characteristics (Parker et al., 2003). Interactions in an ABM may include spatially dependent influences, networking, information transmission, and/or social structure (Parker, 2000; Irwin and Bockstael, 2002; Epstein, 2006; Ma and Nakamori, 2009).

The implementation of ABM to explain changes in land use are relatively new within the socio-ecological modeling literature (Parker, 2000; Berger, 2001; Irwin and Bockstael, 2002). Anderies et al. (2004), Le et al. (2010), Manson (2005), and Parker et al. (2003) offer examples of ABM applied to land use systems. One well-known study using ABM for agricultural communities

was based on the fall of the Anasazi civilization around 1300 CE in Arizona (Dean et al., 2000). The evidence from this Anasazi model shows that settlement changes of the Anasazi people were related to the availability of water resources (Dean et al., 2000).

Additionally, several studies have used an ABM approach for land use, land cover processes (Millington et al., 2008, Parker et al., 2003). Some of these models focus on feedback effects between the decision-making of agents and the physical landscape (Parker et al., 2003). Parker et al. (2003) assesses multi-agent system models to explain land-use/cover change. This study provides examples of agent-based/LUCC models that are suitable for analyzing complex systems where decentralized decisions are made at the local scale (Parker et al., 2003).

Parker (2000) examined distant-dependent spatial externalities defined as edge effect externalities. The results from this study illustrates that zoning and buffer zoning are important for efficient allocation of land uses. Irwin and Bocksteal (2002) examined negative externalities resulting into negative spatial interactions or a “repelling” effect among residential properties which leads to fragmented patterns of development in the rural-urban boundary area in Maryland. Manson (2005) utilized a genetic programming approach for land use change in the Southern Yucatán Peninsular Region (SYPR) of Mexico. This application used a multi-criteria evaluation method. Key socio-economic and political drivers such as market institutions and land tenures, as well as ecological factors, especially characteristics of land use land cover, history, and fragmentation were found to be important for decision making about land use change.

Le et al. (2010) developed a Land Use Dynamic Simulator (LUDAS) within a multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. They assessed policies of forest protection zoning, agricultural extension, and agrochemical subsidies. The results suggest that policy implementation that focuses on forest protection in sensitive

watershed areas results in agricultural production in less critical zones. Thus, policies that help in forest restoration provide economic incentives and opportunities to generate future income for communities.

Brown and Robinson (2006) utilized an ABM to examine the process of urbanized residential development. Heterogeneity of agents was characterized using the choices of location for residential development from survey results of southeastern Michigan residents. Residential agents picked residential locations on the basis of attributes such as proximity to urban economic and employment centers, aesthetic landscape, and agent's similarity to their neighboring agents. The model showed that heterogeneity in preferences influence land use patterns such as sprawl. The weakness of the model was a lack of validation as model results were not compared with landscape changes.

Millington et al. (2008) examined wildfire risk using spatially explicit land use agent-based modeling for Mediterranean agricultural land uses. The results suggest that size of the parcel impacts the land management. The study strongly recommended the spatially explicit agent based model to explain the impacts of land use changes. A spatially explicit model allows investigation of the spatial variation of wildfire from one location to another. The model also observed spatial variation of wildfire under various scenarios of land tenure. The model did not rely on a predictive approach due to data difficulties. Therefore they used a heuristic model structure.

Bithell and Brasington (2009) developed a preliminary model to simulate land-use change using a coupled modeling environment linking ABM of subsistence farming with forest dynamics and hydrological processes. This linked modeling system estimated distributed soil moisture and basin scale water fluxes at catchment scale and found that soil moisture is connected to the topographic features of landscape. They found the variation in soil moisture with different land

covers, which impacts the pattern of storm runoff. The shortcomings of the model was a lack of spatial and temporal variation in rainfall distribution across the catchment scale, which is important for farming. In addition heterogeneity of soil properties was not incorporated into the model.

Le (2008) presented a theoretical model of multi-agents for a spatio-temporal simulation of a coupled human landscape system in rural forest boundaries using a forest growth model. The model examined self-organization among agents and the landscape environment in a real-time feedback pixel based landscape. The heuristically ruled based decisions within a utility maximization framework is defined in spatial multi-nominal logistic functions. The model was develop to test which scenario in land use decisions can improve the livelihood and decrease the negative effects of land use change.

## **2.5 Coupled ABM Models of Land Use**

In recent studies, ABM have been coupled with hydrological models to quantify the impacts of land use change on water quality (Tsai et al., 2013; Bithell and Brasington, 2009; Ng et al., 2011; Millington et al., 2008).

Tsai et al. (2013) formulated a land use transition agent-based model (LTABM) to simulate land use decision making of farmers buying and selling lands. In this model, the resulting land use change was estimated for the Missisquoi Watershed in Vermont. A hypothetical probability distribution of financial stress was used to represent the financial conditions of farmers, which played an important role in their decision making. Financial stress criteria were determined from interview survey responses by farmers and classified as major stress, moderate stress, or feel good financially conditions. The results of the model show that farmers with major financial stress are likely to abandon 25%, 75%, and 90% of the agriculture, barren land, and grass lands respectively. Farmers with moderate financial stress are assumed not to change land into another use. When

farmers feel good financially, the results show that 25%, 90%, and 90% of the forest, barren and grass lands are cultivated by these farmers respectively. As an initial modeling effort, this study did not incorporate heterogeneity of farmers in terms of their interaction with the neighboring land use.

Ng et al. (2011) developed an ABM in combination with SWAT (Soil and Water Assessment Tool) to identify factors impacting farmers' decisions on adopting Best Management Practices (BMPs) and to estimate corresponding nitrate load reductions in the Salt Creek Watershed, Illinois. Fifty semi-hypothetical farmers were characterized in ABM using deterministic and stochastic models for modeling the farmers' optimizing decisions on crop and BMP, which were linked to a hydrologic-agronomic model of the watershed. Stream quality was evaluated by nitrate load output in the SWAT model. Farmers' decisions were modeled as a function of risk aversion, future market prices, productivity, and weather conditions. These decision impacts the cultivation of crops, which influence stream quality. Farmer perceptions were represented by a probability distribution and updated using Bayesian inference.

The findings suggest that driving factors of farmers' decisions to practice BMPs are crop prices, production costs, and yields. The model predicted that farmers who interact with their neighbors, quickly adapted expectations, and were less risk averse were the farmers who were found more open to practice BMPs. However, hypothetical farmers were used as agents, therefore the results of their model may not accurately verify farmers' decision of adopting BMPs. The model incorporated the dynamic behavior of farmers, but the choice of crop and best management practices were based upon the hypothetical agent-based landscape not the observed landscape. The spatially explicit landscape at catchment scale can provide farmers behavior as an attribute to land use practices along with the identification of vulnerable locations due to farm practices.

Reeves and Zellner (2010) simulated the potential interactions between dynamic urban pattern and groundwater health in a prototype application for Monroe County, Michigan by linking an ABM with MODFLOW, a ground water flow model. Results of the linked simulation model show the potential effect of decisions of residents and golf courses under zoning scenarios on groundwater levels. The location features for residential preferences included municipal water, soil quality, density, sewer coverage, roads, and proximity factors representing the distances to recreational areas, school, or business centers. In the model, Monroe County was represented by a grid of cells. Agents randomly select cells and develop those cells that provide the highest utility functions. The coefficient values for locational and proximity factors for each agent were assigned using a probability distribution function based on results from the personal surveys. Zoning was imposed by the development threshold. Therefore, the cells having high density development were recognized as unavailable cells.

The model generated land use patterns, which linked to groundwater flow. However, the model used a regional level groundwater model instead of capturing the local scale ground water flow and stressors. Another shortcoming of this model is the use of cell based observation instead of individually owned land parcels. The cells do not represent the actual boundaries between neighbors. Such parcel boundaries would also provide information on impacts of zoning across large and small property parcels.

Although understanding local scale spatial processes are challenging for socio-ecological research due to data limitations and difficulty in anticipating future change within a computational modeling approach. This can be overcome with the use of observed data in modeling. The present section has reviewed selected studies to provide background about the agent based model for land



use change and the capability of these model to couple with ecological and hydrological modeling frameworks. Each study has its own limitations and strengths.

# Chapter 3

## 3. Theoretical Framework

Spatial patterns of land use conversions are the result of various underlying socioeconomic and biophysical processes. Theories of land use change typically examine the processes according to their respective individual discipline, thus lacking the complex interrelationships of several underlying driving forces that define land use change. These underlying processes cannot be explained by one unifying theory due to the interaction of multi-scale socio-ecological factors that explain the driving forces of land use change (Veldkamp et al., 2001; Verburg et al., 2002; Braimoh, 2004). In order to anticipate future land use change, there is a need for a theoretical model that not only allows us to project land use patterns using explanatory factors, but also represents the underlying choice behavior that generate these patterns. In this chapter such a modelling approach is presented. The derived modeling framework will be implemented for projecting future land use conversions in the Opequon Creek watershed, Berkeley County.

Following bid-rent theory (Alonso, 1964; Mills, 1967; Muth, 1969), the present modeling approach is based upon the assumption that each land parcel is allocated to the use that maximizes the utility of its owner. The necessary condition for land use change/conversion from undeveloped agricultural or forest land into residential land use depends upon the land value for residential use being higher than agricultural or forest use. This necessary condition is suggested due to the identification of urbanization as the major factor of farmland land conversion into developed land (Osion and Osion, 1999; Koontz, 2001; Rosenberger et al., 2002). Studies found that factors such as population, urban growth, employment, and proximity to urban centers cause land use changes in farmland land by influencing the value of land (Fleming, 1989; Nzaku and Bukenya, 2005).

Taking into account this necessary condition in the model, this research grounded in two strands of theoretical approaches to explaining residential land value:

- (1) Land value as a function of distance to economic locations, and
- (2) Land value as a function of surrounding, location specific attributes.

Theoretically, land use conversions are driven by decisions of land owners about whether to convert the land due to the difference in the relative rent between potential uses of land (Polyakov and Zhang, 2008; Lambin and Meyfroidt, 2010). Similar to the modeling structure suggested by Polyakov and Zhang (2008), the basic assumption of the model is that land use change is the outcome of property owner's choice to allocate an undeveloped property parcel into a residential parcel. The decision to convert land depends upon the net present value of the land parcel. Other than through land prices, the net present value is not directly observable but can be characterized in observable attributes, which influence conversion decisions. These observable attributes are the location specific features and neighboring land use patterns. Therefore, a land use conversion decision based upon the utility of land parcel conversion conditional on the initial state can be expressed as:

$$(3.1) \quad U_{is+1} > U_{is}$$

Where  $U_{is}$  is the utility from parcel  $i$  in state  $s$ . The property owner would convert the land if the utility of the converted state is higher than the current state. Following Polyakov and Zhang (2008), the utility of land use conversion can be formulated as:

$$(3.2) \quad U_{is} = V_{is+} + \varphi_{is}$$

Where  $V_{is} = V(A_{is} + N_{is} + M_{is})$  the conversion utility of parcel  $i$  in state  $s$  is based upon  $A_{is} + N_{is} + M_{is}$  attributes. Where  $A_{is}$  is distance to economic locations,  $N_{is}$  is distance to the neighboring land uses,  $M_{is}$  is the distance to the amenities and  $\varphi_{is}$  is all other factors that affect

the utility but that are not included in utility function. Error function  $\varphi_{it}$  is distributed randomly in the model. In the present modeling approach, spatial interactions and dependencies are embedded in neighboring externalities. Neighborhood externalities are estimated influence of each land use on surrounding parcel value.

The utility of conversion from non-developed to residential parcel is assumed to be defined as a linear function of spatial externalities (observable attributes) across all property parcel owners in state  $s$ .

$$(3.3) \quad V_{is} = \delta_{0s} + \beta_i A_{is} + \alpha_i N_{is} + \Omega_i M_{is}$$

Where  $\delta_{0s}$  is a constant parameter in each state,  $\beta_i$  is a vector of coefficients representing the influences of distance from economic location on parcel  $i$ 's value, and  $\alpha_i$  is a coefficient vector that represents the neighboring spatial externalities on parcel  $i$  and  $\Omega_i$  is the coefficient estimates of effect of amenities. Distances to each land use type in the model are regarded as a surrogate for neighboring spatial externalities, distance to forest and streams are regarded as spatial externalities due to amenities (Roe et al., 2004; Irwin and Bockstael, 2004; Poudyal et al., 2008).

Land use decisions often involve uncertainty and expectations about the future, which vary from one land owner to another (Bockstael, 1996). Due to stochastic elements within spatial land use dynamics of simulation, probabilistic discrete choice is often recommended in the complexity literature (Heppenstall and Crooks, 2012). However, such an approach is not common, due to data limitations and problems with computational modeling structure for estimation of explanatory variables (Bayoh, 2006).

The probabilistic approach employed in the present modeling framework is based on Monte Carlo simulations. Monte Carlo simulations have been widely used in systems where uncertainty exists in the problem, such as risks, costs and failures need to be predicted

(Papadopoulos and Yeung, 2001; Glasserman 2003; Breuer et al., 2006). Monte Carlo simulations have been used in computational models of CA as an experimental probabilistic method for the spatial distribution of probabilities (Clarke and Gaydos, 1998). Since the sample of parcels is already taken from secondary data, this model uses Monte Carlo simulations to estimate the probability of conversion for each parcel agent by using the law of large numbers. With the law of large numbers for simulations, the average estimated conversion probability was assumed approximate the true value of conversion probability (McLeish, 2003). By assuming that each parcel taking on either a 0 or 1 value at the end of each simulation, the probability of conversion in n model runs is calculated using Monte Carlo technique to forecast the spatial patterns as:

$$(3.4) P_i(\text{conversion}) = \frac{\sum_{n=1}^N (C_k)}{n}$$

Where  $P_i$  the probability of conversion for parcel  $i$ ,  $n$  is the number of model runs,  $C$  is the Boolean conversion in each model run with  $k$  is 1 or 0, where 1 indicates conversion and 0 represents no conversion. Within the context of land use modeling approach, Monte Carlo method generate patterns of likely residential developments. Monte Carlo simulations are suitable for complex systems such as Opequon Creek watershed to assess uncertainty and predictability of the model for each policy scenario and to analyze variability in the model results. With a Monte Carlo method, peculiarities in land use patterns and sensitivity of the model to the policy parameters can be observed.

The behavior of individuals within a linear framework may lead to incorrect predictions regarding the distribution of land use conversion decisions over a landscape. For example, a land use change measured over 10 years may have higher land use conversion within the first few years or the last few years. Therefore, this theoretical framework provides a strong basis for conceptualizing probabilistic land use conversion of individuals as a decision making variable in

non-linear land use processes. In particular, for each individual the characterization of land use conversion decisions is based upon the probability of land use conversion from non-developed parcel to residential parcel in n number of times.

Probability of conversion is a function of parcel based observable spatial externalities, which include location specific factors of land use conversion ( $A_{is}$ ), neighboring features ( $N_{is}$ ), neighboring amenities ( $M_{is}$ ) and random portion,  $\varphi_{is}$  which is unobservable (Bockstael, 1996; Heppenstall and Crooks, 2012).

$$(3.5) \text{ prob}(\text{develop}) = \text{prob}(A_{is} + N_{is} + M_{is} + \varphi_{is})$$

The empirical structure of the logistic regression function describes the functional relationship between the land use conversion and the a set of explanatory variables that influence conversion probability, which makes it best suited for modeling land use change (Xie et al., 2005). This approach accounts the behavioral aspect of economic decision making of conversion in defining the underlying processes (Irwin and Geoghegan, 2001).

Logistic regression can be used to explain the probability of land use conversion (conversion=1, non-conversion = 0) of each location given a set of explanatory variables. Following Wooldridge (2002),

$$(3.6) P(y = 1|x) = P(y = 1|x_1, x_2, x_3, x_4, x_5, x_6)$$

This approach explicitly connect conversion observations to residential land value changes. As is explained further in Chapter 4, the explanatory variables utilized in this research to explain parcel land value (rent) into residential land use are: (1) distance to urban centers; (2) distance to major highways; (3) distance to streams; (4) distances to agricultural land use; (5) distances to forest land use; and (6) distance to residential land use.

Distance to urban center is a surrogate for proximity to economic activity centers, schools, shopping centers, railway station, and public services (Kitamura et al., 1997). Distances to the roads and urban centers are conceptualized upon the intuition of Von Thünen and the bid-rent theory of urban economics where distance to the urban center and roads explains the land rent and transportation cost respectively under the relaxed assumptions of spatial variation in the landscape (Von Thünen , 1826; Alonso, 1964; Mills, 1967; Muth, 1969). Distance to a city is defined as the major factor in monocentric bid-rent theory (Alonso 1964). As the distance from the city center increases, accessibility decreases which results in higher transportation costs. Distance to roads can be regarded as a proxy for accessibility of metropolitan and urban areas, workplace, shopping, and schools (Serneels and Lambin, 2001).

The framework that spatial features and spatial structure of land patterns surrounding parcels impact the parcel value is taken from the economic theory of hedonic property values (Wu, 2006). In particular, the theoretical basis of distances to streams and forests are regarded as a relative measures for aesthetic amenities, where closeness to streams and forests determines the value of parcel-level characteristics. Since streams and forest count as proxy variables for amenities, therefore, they are expected to positively influence the probabilities of land use conversions (Irwin et al., 2014). The distance to agricultural and residential land use accounts for neighboring externalities.

Within bid-rent theory, the distance from the economic locations is conceptualized by concentric rings of development density around the CBD, which does not realistically model the spatial variations in land use patterns. In this regard, empirical hedonic property value models typically include property characteristic variables to explain changes in prices for residential property and account for local scale variations. These characteristics can be property specific such

as structural features as well as surrounding characteristics such as open space, forest, and streams. These factors affect the value of a parcel in its use and enter into the conversion decisions.

The model utilized in this research includes only economic locations and surrounding, location specific attributes. The theoretical basis of the hedonic model defines how amenities or locational characteristics impact decision making and how this decision making is embedded into the spatial land use pattern. Using a similar approach to Geoghegan et al. (1997) the hedonic model is conceptualized as follows:

$$(3.7) \quad R = \alpha + A\beta + N\alpha + M\Omega + \varepsilon$$

Where  $R$  is a vector of parcel value;  $A$  is a vector of distance based variables (such as distances from economic location);  $\beta$  is a parameter vector for  $A$ ,  $N$  is a vector of neighborhood characteristics (such as land uses);  $\alpha$  is associated parameter vector;  $M$  is a vector of amenities and  $\Omega$  is associated parameter, and  $\varepsilon$  is a vector of random error terms.

The distance based characteristics can be altered by spatially restricted policies such as restricting residential development within a certain distance from the stream, which influences the parcel value. Therefore, land use policies in influencing land use conversion are important policy parameters in this research.



# Chapter 4

## 4. Conceptual Framework

This research is divided into three interconnected models to illustrate the concept of driving factors of land use change, land conversion decisions, and linking these factors to water quality indicators. The three models are: (1) a spatial logistic regression model, (2) an agent-based probabilistic land use conversion (APLUC) model, and (3) an ArcSWAT surface water quality model. To examine the impact of land use conversions in the Opequon Creek watershed, a linked agent-based probabilistic land use model and surface water quality model is employed under three buffer policy scenarios. These models and their connections are represented in Figure 2.

The methodology used in this study provides a hybrid approach of spatial logistic regression to calibrate an APLUC model, which not only quantifies land use change, but provides coefficients for explanatory variables of land use change from real observations of land use conversions. The implementation of spatial logistic regression allows empirically derived parameters for spatial externalities for the APLUC model.

Using the parameters from the spatial logistic regression model, land use conversion by parcel agents in the APLUC model is determined by three criteria: (1) neighboring externalities in terms of distance from each land use type, (2) distance from location features such as roads, streams, and urban centers, and (3) land use conversion probability. The APLUC model simulates the decisions to convert developable land into residentially developed land, given land parcel attributes, and generates the land use patterns at a disaggregated level.

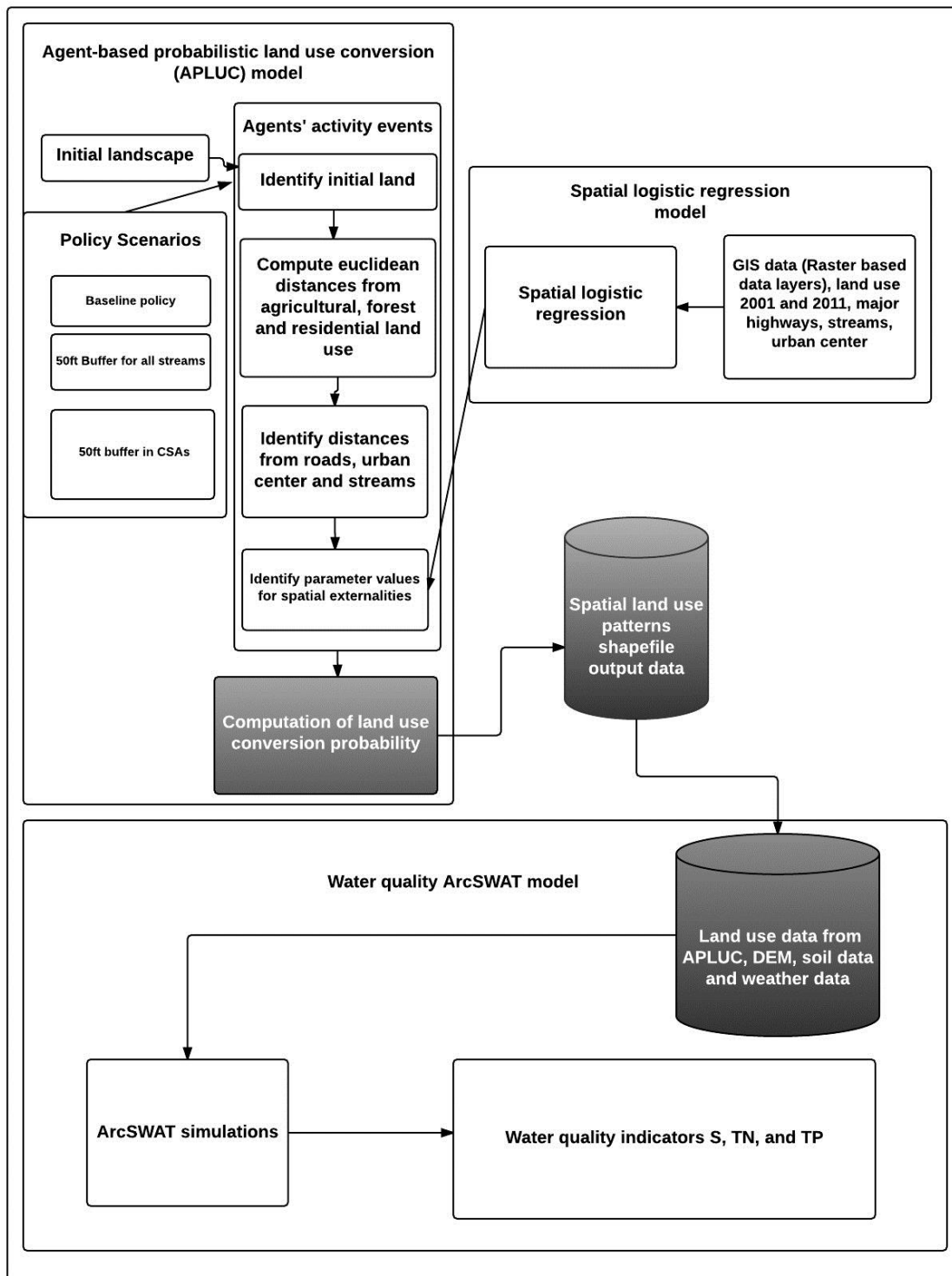


Figure 2. Conceptual framework for this research.

The ArcSWAT model is used to simulate the impact of spatially explicit conversion decisions from the APLUC model on the movement of pollutants across the Opequon Creek watershed. A linked APLUC/ArcSWAT model is used to provide the impact of residential development on the hydrology of the watershed. The impact of land use change is analyzed by three variables: total phosphorus (TP), total nitrogen (TN), and sediments.

#### **4.1 A Spatial Logistic Regression (SLR) Model**

In order to predict the macro scale land use conversion probabilities, driving factors of land use conversion were estimated to examine the probability of land use conversion during the period 2001-2011. To examine the change in spatial residential land use patterns, a spatial logistic regression analysis was developed in IDRISI Selva Software of Clark Labs to estimate the influence of driving factors on spatial land use trends in the Opequon Creek watershed. Logistic regression offers the functionality to incorporate binary dependent variables as a presence or absence of occurrence and suitability for discrete, categorical, or continuous explanatory variables (Atkinson and Massari 1998; Lee, 2005).

The empirically estimated relationship between the conversions of residential development and the driving factors can be expressed as the following logistic functional form:

$$(4.1) \quad P(Y = 1|x) = \frac{\exp(\Sigma \beta X)}{1 + \exp(\Sigma \beta X)}$$

Where  $P(Y = 1|X)$  is the predicted probability value of the binary or dichotomous dependent variable  $Y$  and where  $Y=1$  means if a cell in raster map changes from a non-residential land use in 2001 to residential land use in 2011 and  $Y=0$ , otherwise. The probability of  $Y$  having the value of 1 is assumed to follow a logistic curve as defined in equation (4.1) (Wooldridge, 2002).  $X$  is the full set of explanatory variables and  $\beta$  are linear combination of estimated parameters. Equation

4.1 is the cumulative distribution function for a standard logistic random variable shown in Figure 3.

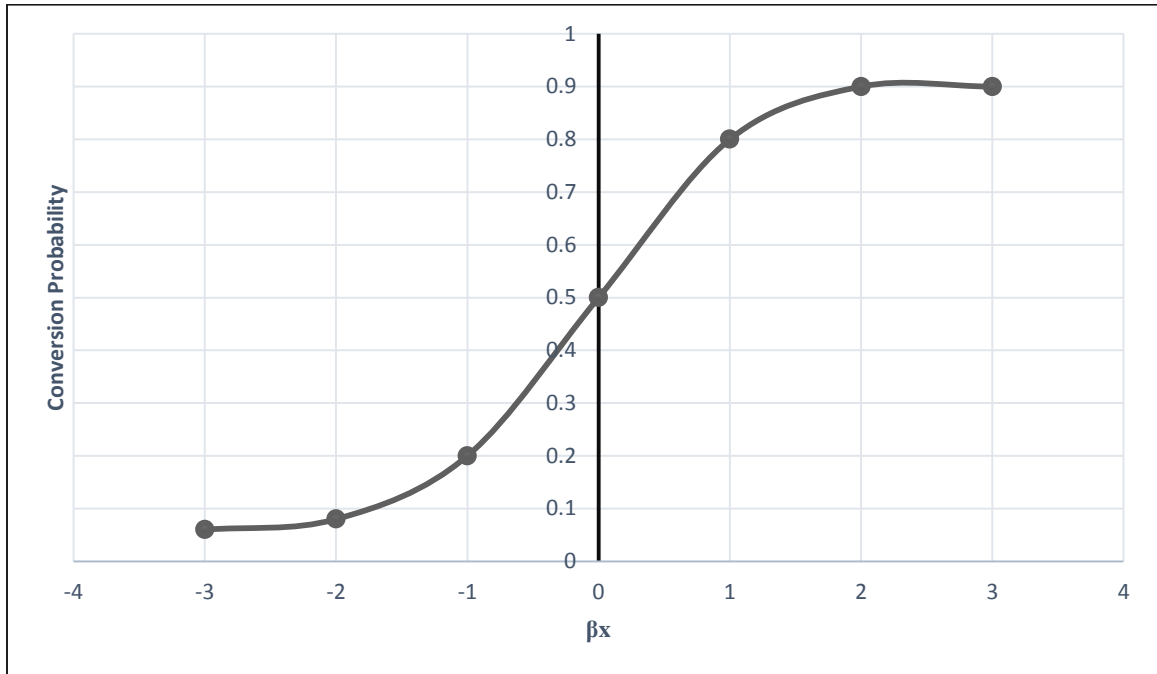


Figure 3. Logistic function

Logistic function takes into account the linear probability in a set of parameters, by having the range of probability between zero and one. The following linear logit transformation on both sides of equation 4.1 was used to estimate the coefficients (Menard, 1995):

$$(4.2) \quad Y = \text{logit}(p) = \ln\left(\frac{p_k}{(1-p_k)}\right) = \beta_0 + \beta_1 x_{1k} + \beta_2 x_{2k} + \beta_3 x_{3k} + \beta_4 x_{4k} + \beta_5 x_{5k} + \beta_6 x_{6k}$$

Where Y is the probability that the dependent variable (Y) is 1,  $p_k$  is the predicted probability of dependent variable of non-residential land use (agricultural and forest) conversion to residential land use.  $\beta_0$  is the intercept, and  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \text{ and } \beta_6$  are coefficients for distance to the existing agriculture, distance to the existing forests, distance to the existing residential areas, distance to streams, distance to major highways, and distance to urban center respectively. These

coefficients measure the influence of each independent variable ( $x_1, x_2, x_3, x_4, x_5,$  and  $x_6$ ) on the variations in probability of land use conversion from non-residential land use to residential land use (Y).

The sign of the parameter indicates the influence of each explanatory variable to the conversion probability. A negative sign for a parameter shows that a decrease in distance would increase the probability of conversion. Similarly, the positive sign of parameter shows that as the distance from the cell to land use or location feature increases, the probability of conversion would increase.

Additionally, the nonlinear nature of logistic function makes it difficult to interpret the direct increase or decrease of continuous variables  $x_i$  on the land use conversion probability (Zhang and Nagubadi, 2005). Several research studies suggested the use of marginal effect of land use conversion at the mean value of each explanatory variable for interpretation of model results (Zhang and Nagubadi, 2005; Poudyal et al., 2008). Therefore, marginal effects of continuous variables on the land use conversion probability was calculated in the model as the following partial derivative form (Wooldridge, 2002):

$$(4.3) \quad \frac{\partial p(x)}{\partial x_j} = g(\beta_o + x\beta)\beta_j$$

Where  $x_j$  is the continuous explanatory variable,  $p(x)$  is probability response,  $g$  is the probability density function and defined as (Wooldridge, 2002):

$$(4.4) \quad g(z) \equiv \frac{dG}{dz}(z)$$

Where  $G$  is the strictly increasing cumulative density function (cdf) of continuous random variable, therefore  $g(z) > 0$  for all  $z$ . This means, partial effects of each explanatory variable on probability of conversion has the same sign as  $B_j$  (Wooldridge, 2002).

Since the marginal change is the change in the land use conversion probability from non-residential to residential land per unit change of explanatory variable, it is interpreted as per meter change in land use conversion due to per meter change in distance dependent explanatory variables. The marginal change is calculated in the model using log odds by one unit increase (per meter increase in probability) as below (IDRE, 2014):

$$(4.5) \quad L_1 = \log\left(\frac{p}{1-p}\right) = \text{logit}(p) = \beta_0 + \beta_n \overline{x_{nk}}$$

Where  $\beta_0$  is the intercept and  $\overline{x_{nk}}$  is mean of explanatory variable  $n$ .

The one unit increase in mean is

$$(4.6) \quad L_2 = \log\left(\frac{p}{1-p}\right) = \text{logit}(p) = \beta_0 + (\beta_n(\overline{x_{nk}} + 1))$$

The marginal effect obtained by taking the difference between equation (4.5) and (4.6)

$$(4.7) \quad [L_1 - L_2]$$

Another interpretation of the model can be provided by estimating the odds value as the exponential of the parameter estimate values (Hosmer and Lemeshow, 1989; IDRE, 2014):

$$(4.8) \quad \exp[L_1 - L_2]$$

This provides the percentage change in land use conversion due to corresponding explanatory variable holding other explanatory variables constant.

## **Model Evaluation**

The SLR model was validated using the quantitative measurement of Receiver Operator Characteristic (ROC) statistics, which is suggested as a reliable approach for model validation by several studies (Pontius and Schneider 2001; Dendoncker et al., 2007; Arsanjani et al., 2013). The ROC predicts the location of conversion by comparing the actual change between 2001 and 2011 in a Boolean map and the suitability (fitted) change between 2001 and 2011, which is defined as excellent statistics to measure the goodness of fit of a logistic regression (Swets, 1986; Pontius and Schneider, 2000; Verburg et al., 2002; Pijanowski et al., 2009; Tayyebi et al., 2010; Clark Labs, 2014). The ROC varies between 0 and 1. Within the ROC range, 1 shows a perfect fit and 0.5 shows a random fit. The larger ROC values show the better association between explanatory variables and dependent variable (Clark Labs, 2012).

A Chi-square distribution of likelihood ratio statistics tests the null hypothesis that all variables measuring distances from economic locations, surrounding land uses or amenities have no impact on whether a cell will be converted from non-residential to residential land use. Due to the non-linear functional form of logistic model, OLS is not applicable (Wooldrige, 2002). Therefore, maximum likelihood estimation (MLE) is used. MLE estimator is good for the distribution of  $y$  given  $x$  which includes heteroskedasticity in  $\text{VAR}(y|x)$  and is consistent, asymptotically normal, and asymptotically efficient (Wooldrige, 2002). The MLE provides the predictive power to the model in terms of MLE of model parameters.

Spatial externalities within the APLUC model are deduced from coefficient estimates in the spatial logistic regression. Further, these empirically estimated parameters showing per meter spatial externalities are implemented in the APLUC model to model the land use conversion decision making of agents.

## **Sampling approach**

The explanatory variables in the samples were tested for multicollinearity. Because cells (pixels) are used as sampling units in raster based spatial logistic models, a very large number of samples are generated. This results in similarity among cells (pixels) and causes autocorrelation in these models (Lo and Watson, 1998; Clark Labs, 2012). The autocorrelation results in inefficient estimates and type I errors (Clark and Granato, 2005). To quantitatively minimize the spatial patterns of errors (autocorrelation) between the connectivity of cells (pixels) in the sample, a pixel thinning method is the suggested method for raster-based data (Lo and Watson, 1998; Clark Labs, 2012). Using the pixel thinning method in IDRISI Selva software, raster data for all explanatory variables and dependent variable are contracted by the contraction factor of 10<sup>th</sup> lag. Due to the small scale study area, with 10<sup>th</sup> lag, every 10<sup>th</sup> cell (pixel) is selected, which provide a wider spatial distance between each cell to minimize the effects of spatial autocorrelation and reduces the number of cells in the sample.

## **4.2 Agent-Based Probabilistic Land Use Conversion (APLUC) model**

The land use patterns are conditional to the sequential land use conversions resulting in variation in neighboring land uses. This sequential process can best be defined by updating the local probability of land use conversion to explore land use patterns resulting from micro-scale dynamics (Wu, 2002). Spatial logistic regression provides empirically valid estimates of positive or negative effects on land use conversion probabilities as measured by geographic location and nearby featuring attributes. Theoretically, direct implementation of explanatory variables in the SLR to model local probability in each time period (event/step) can be achieved. But due to data limitations, it is not technically feasible to model local probabilities on an incremental (roughly year to year) basis (Wu, 2002). The process of simulating land use conversion in the APLUC



model not only allows the discrete conversion of each parcel but also to explore land use change in increments instead of just the start and terminating points.

### **APLUC Model Environment**

The APLUC is based upon a geographic environment where each decision unit or property parcel has the following fundamental characteristics:

- Parcel boundaries,
- Spatial externalities estimates,
- Distance estimates for neighboring features, and
- Land category.

### **Landscape**

The landscape of the model comprises Berkeley County within the Opequon Creek watershed area, consisting of approximately 58.51 square miles. The model employs a GIS environment composed of actual property parcels in Berkeley County (see Appendix I). The parcels are assigned with land use conversion behavior under three land use policy scenarios, which are described later in this chapter.

### **Agents**

The parcel agents are developable parcels representing as land owners' choice making units. The developable parcels are forest and agricultural properties, and have no restriction on the density of development. Each parcel is assumed to act independently by being owned and controlled by a single owner, therefore each property parcel is characterized as an agent. In reality, multiple parcels are owned by a single owner, but the same property owner may convert the property parcel in one location, without converting a property parcel owned at a different location. The APLUC approach reasonably captures the link between location of property parcel and property owner.

The APLUC model is built upon parcels' probability of conversion, which reflect the conversion behavior of property owners.

### **Assumptions**

The key assumptions underlying the land use model dynamics are:

- Agents know their property parcel, location, land use conversion probability value of all other parcel agents, and distances from each land use,
- Agents do not foresee the effects of land use decisions of their neighboring parcel for more than one event period,
- The action of land use conversion is regarded as the assumption that agricultural and forest land owners either sell to a residential developer or convert into a residentially developed area. Both actions are pre-assumed as a conversion event of developable land into residentially developed land in the model,
- The parcels that are residentially developed by the agents are assumed to remain as residentially developed parcels in every iteration hereafter. Once the parcel is residentially developed, it is not available to the pool of undeveloped parcels, and
- Agents are not assumed to exhibit optimizing behavior on an inter-temporal basis.

### **Decision Variables**

In the APLUC model, an agent's behavior is formulated based on empirical rules of land use conversion. This method represents agents' decisions to convert land using a probabilistic approach. A similar approach has been used in studies where agents are characterized within a bounded rationality framework (Benenson and Torrens, 2004; Valbuena et al., 2010). Bounded rationality implies that agents have limited information (Simon, 1955).

The agents' conversion decisions vary with the spatial distances from each neighboring land use over a period of 10 iterations (to rough approximate a 10 ten year time period), where

each iteration is assumed as a conversion event possibility. Additionally, the modelling factor that influences decision making is path dependency in which the initial conversions influences future conversions within each model run.

In classic CA land use models, transition probabilities of land use do not change over time in Markov processes and therefore result in equilibrium states (Almeida et al., 2005; Ozah et al., 2012). In the APLUC model, these probabilities are not non-stationary due to the dependence of local probabilities on neighboring land use parcels.

By modifying the Beneson and Torrens (2004) approach in the APLUC model, each parcel agent's probability *Prob* of conversion from developable *i* to residential state *j* in each iteration is modeled as:

$$(4.9) \text{ Prob}_a(S_i \rightarrow S_j) = S(N(a))$$

Where *N* (*a*) represents parcel agent *a*'s neighbors and *S* represents state of parcel *a*. Decision rules and initial conditions such as distance to streams, roads, and urban centers of each property parcels do not change over the course of operation for the APLUC model. The decision rules that are implemented involve investigating the spatial features of that parcels' neighborhood, and then conversion of the property parcel into residential land use based upon the probabilities influenced by driving factors (Dietzel and Clarke, 2004).

The model employs a Monte Carlo process (Hagerstrand, 1965; Wu, 2002) to generate the result of a stochastic APLUC model. Due to uncertainty, the probability function is used to condition the residential conversions using a random number generator (Heppenstall and Crooks, 2012). For undeveloped parcels, the conversion decision is based upon the comparison between a random number and the probability value generated from equation (4.10) for each parcel. The

random number generator has a random distribution that is uniform between 0 and 1. The agents adopted the following rule of land use conversion in each iteration

(4.10) if  $\text{rand}(\varphi_i) < P_{it}$  then  $A_{it+1} = k$

And  $\text{rand}(\varphi_i) \sim N(\sigma, \mu^2)$

Where  $k$  represents the land use class of residential development.  $P$  is the probability of conversion to residential development for each parcel  $i$ ,  $A$  is the conversion event and  $t$  is iteration. Agents first assess the probability of conversion by comparing it with a random number. If the value of probability is higher than the random number, the agent converts the parcel into a residentially developed parcel. If not, then the parcel remains in its current undeveloped state. This means that if the uncertainty factor is higher than the probability of conversion based upon the favorable surrounding features for conversion, the agent would not convert the developable parcel into a residential parcel. The random number generator incorporates a stochastic element into the APLUC model, which mimics uncertainty and allow randomness in the model.

Land use conversions change the land use allocation of parcels in each iteration. The spatial distribution of changes in land use conversion from undeveloped to developed parcels impacts decisions of neighboring parcels to convert in subsequent iterations due to neighborhood externality impacts. Changes in distances from each land use helps in investigating the path-dependence and stochastic processes on residential land use patterns (Magliocca et al., 2011).

The APLUC model employs spatial logistic regression for the calibration to project the residential land use conversions in each policy scenario. The purpose of calibration is to extract the coefficient or parameter values for the APLUC model from the observed land use pattern at each iteration  $t$  and subsequent iterations  $t+1$ . The empirical structure of agents' land use conversion probability is generated through parameters from the multiple logistic regression. The

probability of a property parcel to be converted into a residential parcel by each parcel agent is computed as:

$$(4.11) P_i = \frac{\exp(\sum \beta X)}{1 + \exp(\sum \beta X)}$$

Where  $P_i$  is the probability of conversion for each parcel agent;  $X$  is the set of explanatory variables and  $\beta$  is linear combination of spatial externalities of each explanatory variable. The spatial externalities parameters are measured in terms of meters in raster based environment and assigned to each parcel.

### **Process Overview and Scheduling**

The model was created using the Python 2.7 programming language with integration of ArcGIS 10.2 to reflect spatial dynamics, using a “bottom up” approach. The model proceeds in discrete event steps and generated a series of projected residential conversion and non-conversion data sets. A total of 10 iteration steps are included in each model run. The number of iterations steps is based upon how data were generated for the spatial logistic regression. The raster from spatial logistic regression consists of an aggregation of 10 years of land use change from 2001 to 2011. A single iteration represents the duration of a single time period as counted in the SLR model (Fragkias and Seto, 2007). All land use conversions generated synchronously at the end of each iteration.

Parcel agents incorporate the estimated SLR coefficients and in each iteration, they calculate a new set of explanatory variables for neighboring land uses. New explanatory variables are created in each iteration due to changes in spatial patterns of land use parcel data. Thus, as the parcel landscape changes, explanatory variables are recalculated by each parcel agent.

The landscape is initialized as the actual land use vector layer for the year 2011. Agents start their activity by identifying whether the land parcel is developable or not. Then they compute the mean Euclidean distances from agricultural, forest, and residential lands. Once the distances

are calculated, they identify the parameter values for spatial externalities and identify the distances from roads, urban center, and streams.

Having assessed land uses in its type, neighboring land uses, and features distances, agents incorporate this information into their computation of probabilities. The initial or global probability of land use change from the SLR model is used at zero step event for each parcel agent. This probability raster is converted into polygon data where each polygon probability value is based upon the majority cells (each cell representing corresponding probability value) from rasterized data. After this zero event, agents compute the probability of residential land use conversion as defined in equation 4.12 and update the probability using local spatial patterns for each parcel/polygon throughout the simulation iteration. Within each event or step, the agents are making residential land use conversion decisions based upon a generated probability value ranging from  $0 \leq P_{it} \leq 1$ , where  $i$  represents each parcel agent in each event  $t$ .

The agents make their conversion decisions on each developable parcel based on constant information feedback of distances and spatial externalities in each model run in a continuously iterative fashion (Liu et al., 2013). The conversion decision is not only influenced by the neighboring land use conversion but by the assigned coefficient values (parameters), which exhibit the influence of each proximity factor (spatial externalities) on the probability of land use conversion.

The probability of conversion is further transformed into stepwise probability for evaluations. For 10 iterations it is calculated as the following;

$$(4.12) \quad P_i = 1 - \left[ \frac{\exp(\Sigma \beta X)}{1 + \exp(\Sigma \beta X)} \right]^{\frac{1}{10}}$$

$$= P_i = 1 - P^{.10}$$

The stepwise probability ensures the final conversion probability for all 10 steps will match the SLR. The interaction among agents is not explicitly modeled but occurs implicitly as defined by the neighboring land use effects on land use conversion probability. The sign of each parameter from SLR shows the type of interaction among each parcel agent.

To account for the probabilistic nature of conversions, the Monte Carlo process in the APLUC model is repeated for 50 model runs. Each model run generated a different sets of land use conversion sites. Employing, several set of runs helped in testing the fluctuation among model runs. Due to the defined empirical structure of local probability, the model shows fluctuations but at consistent rate. Therefore, the choice of 50 model runs for each policy scenario was found adequate for testing the path dependency and stochastic processes in spatially explicit land use simulation.

Projections of land use conversions for 50 model runs allowed for mapping of probabilities within each developable parcels. The probability of each parcel within fifty model runs, where each model has 10 iterations are calculated as:

$$(4.13) P_i = \sum_{n=1}^{50} C_k / 50$$

Where  $P_i$  the probability of conversion for parcel  $i$ ,  $n$  the number of model runs,  $C$  the Boolean conversion in each model run with  $k$  is 1 or 0, where 1 indicates conversion and 0 represents no conversion. Once the Monte Carlo probabilities are mapped for each parcel, the threshold probability representing future residential land use conversion rate is assigned to generate projected residential land use conversion data (Fragkias and Seto, 2007). Based upon several urban studies, thresholds for probability cut-off points that range between 0.50 to 1 have been selected (Zeeb and Burns, 1998; Sohn and Park, 2008; Fragkias and Seto, 2007). Logically, the projected

probability is interpreted as parcels which have at least 50% likelihood or  $\geq 0.50$  probability of land use conversion is residentially converted parcel or 1 in the model, while projected parcel with  $< 0.50$  probability are assumed as no conversion or takes the zero value. Intuitively, it shows the likelihood of parcel conversion from non-developed to residential land, if conversion decision is made for 500 times (10 iterations over 50 model runs).

The projected Monte-Carlo spatial land use patterns were used as data input for the ArcSWAT model. The flow chart of the decision algorithm in the model is shown in Figure 4. A data generator step and a land use conversion step are performed in each iteration. For each policy option, the model will be used to determine the total number of parcels converted to residential development using a Monte Carlo process. The number of iterations and model runs are also input parameters of the model.

First, the APLUC model was run with the base case scenario, which represents no policy in place. In this scenario, an actual property parcel vector map is initialized with no regulations on residential land use conversions. This means all forest and agricultural parcels are spatially unrestricted. The second model is run with a 50 foot buffer spatially restricted policy. In this policy, the GIS environment has spatially restricted parcel agents from development on 50 foot buffer zones on both sides of all of the streams in the Opequon Creek watershed. The third and last model is run for a 50 foot buffer policy only for streams within high priority sub-basins in the Opequon Creek watershed, where development within buffers in high priority sub-basins is restricted. Identification of high priority sub-basins is based upon Karigomba (2009). For the policy scenarios in the model, developable areas, and non-developable areas are assumed to be designated by the local government as policy regimes.



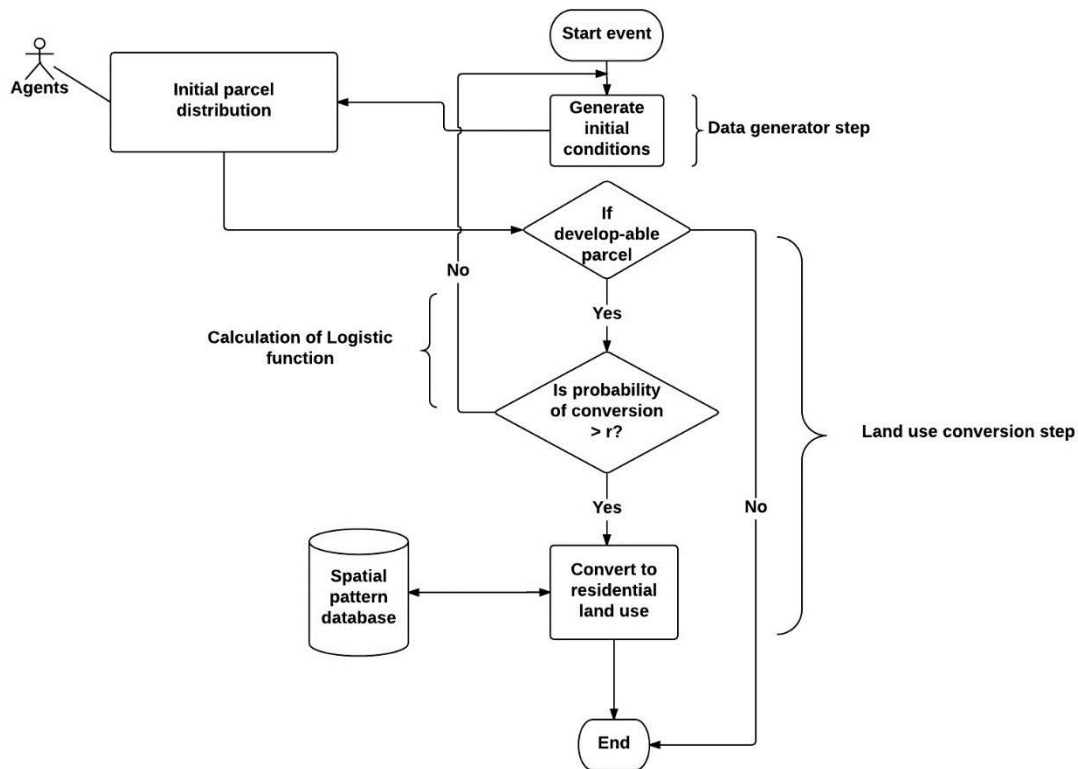


Figure 4. Decision algorithm of agent-based probabilistic land use conversion model

### 4.3 ArcSWAT Model

ArcSWAT, an ArcGIS extension of Soil and Water Assessment tool (SWAT), is used in this study to estimate water quality outcomes stemming from land use conversion changes. SWAT, developed by the USDA Agricultural Research Service, is a continuous, distributed parameter, daily time step model to assess the effects of land management practices on the hydrology, nutrients, sediments and non-point pollutant transport in watersheds under various slope, soils, and land use in a continuous-time (Arnold et al., 1998; Park et al., 2011). SWAT is used for various water quality assessment projects by federal agencies such as EPA, NOAA, and NRCS. Several

studies have recognized its strength in input data availability and predictions using processes based methods for complex watershed systems (Baumgart, 2005; Neitsch et al., 2005; Park et al., 2011; Getachew and Melesse, 2013; Niraula, 2013). The SWAT model has the functionality to divide a watershed into sub-basins and then further divides each sub-basin into hydrological response units (HRUs). This model represents a realistic projection given specific biophysical features such as land use, soil, topography, hydrology, climate, and policy effects at sub-watershed area (Baumgart, 2005).

Following a modeling structure in Getachew and Melesse, (2013), the land use hydrological model comprises steps in ArcSWAT as shown in figure 5. The ArcSWAT model is set up using data on the Opequon Creek Watershed terrain (30 meter resolution digital elevation model (DEM)), land use, soil type, and local meteorological conditions. A DEM of 30 meters was the input to delineate the watershed sub-basins using topography, such as overland slope and slope length (in meters) to analyze the drainage patterns of the landscape and define the area of the sub-basin in the watershed. ArcSWAT delineated the physical characteristics of the Opequon Creek such as size, boundaries, and stream network based upon the digital elevation model (DEM), and divided the watershed into 42 hydrologically and spatially connected sub-basins.

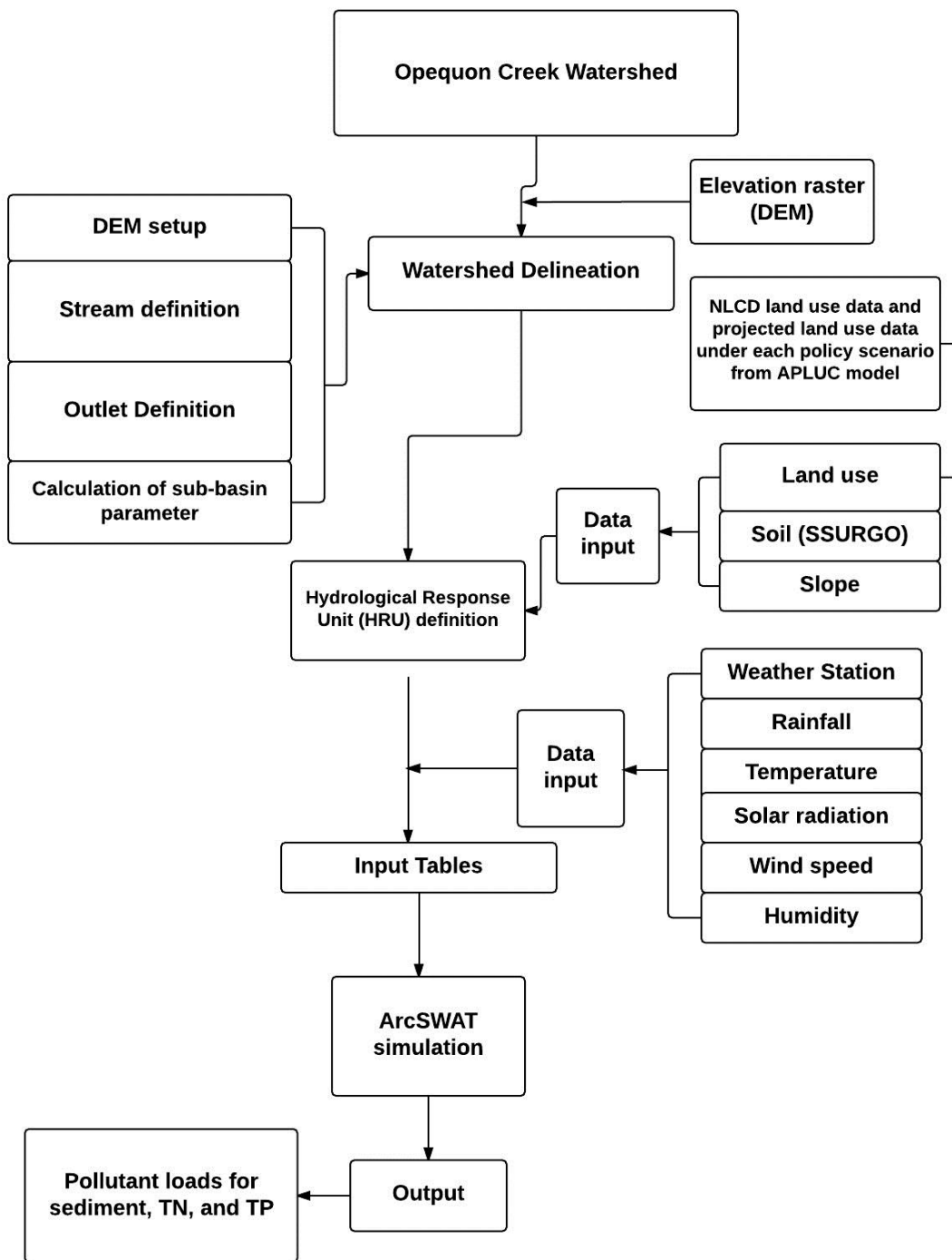


Figure 5. ArcSWAT model procedure for the Opequon Creek Watershed

Using ArcSWAT, the Opequon Creek watershed was partitioned into sub-basins using sub-basin outlet locations. This division allows spatial reference of each sub-basins to one another. Land use classes are matched with the SWAT code for each type of land cover/ land use generated by APLUC data. The SSURGO soil data layer was linked with the soil database. The land use/land cover data from the APLUC, the SSURGO soil data, and the slope class layers were overlaid to derive unique HRUs or sub-basins. For the distribution of HRU's, dominant land use, soils, and slope were used. HRUs are defined as an area that has a unique combination of land, soil type, and slope characteristics. ArcSWAT provides the utility of readily available input data on weather and has the functionality to implement the spatial land patterns data from the APLUC model as land use input to assess water quality responses under each policy scenario.

In ArcSWAT, the hydrology at each HRU is predicted using the water balance equation (Neitsch et al., 2005; Getachew and Melesse, 2013) :

$$(4.14) \quad SW_T = SW_0 + \sum_{i=1}^t (R_{day} - Q_{sruf} - E_a - W_{seep} - Q_{gw})$$

where  $SW_t$  is the final soil water content (mm water),  $SW_0$  is the initial soil water content (mm water),  $t$  is the time (days),  $R_{day}$  is the amount of precipitation on day  $i$  (mm water),  $Q_{surf}$  is the amount of water in the various zone from the soil (mm) on day  $i$  (mm water),  $E_a$  is the amount of evapotranspiration on day  $i$  (mm water),  $W_{seep}$  is the amount of return flow on day  $i$  (mm water) and  $Q_{gw}$  is the amount of ground flow (mm).

For this research, ArcSWAT simulations were run annual over a 10 years' worth of data. The choice of a 10 year time period for simulation was based upon the corresponding training data time frame in the SLR model. ArcSWAT quantified the water quality impacts of land use policy scenarios as captured into the land use hydrological data base at a sub-basin scale. Pollutant loading data were the outcome of the model. The identification criteria is the pollutant load releasing out of the watershed rather than the final pollutant load flowing into the watershed. Having this

consideration, total phosphorus, total nitrogen and sediment were selected as pollutant loadings from ArcSWAT output defined in Table 1.

Table 1. Variables and definitions of pollutant loads in ArcSWAT.

Variable	Definition
TOT_P	Total phosphorus transported with water out of reach during time step (kg P).
TOT_N	Total nitrogen transported with water out of reach during time step (kg N).
SED_OUT	Sediment transported with water out of reach during time step (metric tons).

ArcSWAT calculates the sediment yield in each sub basin using the (MUSLE) Modified Universal Soil Loss Equation (Williams, 1995; Neitsch et al., 2005):

$$(4.15) \quad Sed = 11.8. (Q_{surf} \cdot q_{peak} \cdot area_{hru})^{0.56} \cdot K_{USLE} \cdot C_{USLE} \cdot P_{USLE} \cdot LS_{USLE} \cdot CFRG$$

Where  $Sed$  is the sediment yield on a given day (metric tons),  $Q_{surf}$  is the surface runoff volume ( $mm H_2O/ha$ ),  $q_{peak}$  is the peak runoff rate ( $m^3/S$ ),  $area_{hru}$  is the area of the HRU (ha),  $K_{USLE}$  is the USLE (Universal Soil Loss Equation) soil erodibility factor ( $0.013 \text{ metric ton m}^2 \text{ hr}/(\text{m}^3\text{-metric ton cm})$ ),  $C_{USLE}$  is the USLE cover and management factor,  $P_{USLE}$  is the USLE support practice factor,  $LS_{USLE}$  is the USLE topographic factor, and  $CFRG$  is the coarse fragment factor.

ArcSWAT calculates the amount of sediment released out of the watershed on a given day as a function of final concentration. The sediment Outflow in ArcSWAT is calculated as (Neitsch et al., 2005):

$$(4.16) \quad sed_{flowout} = conc_{sed,f} \cdot V_{flowout}$$

Where  $sed_{flowout}$  is the amount of sediment released out of the water with outflow (metric tons),  $conc_{sed,f}$  is the final sediment concentration ( $Mg/m^3$ ), and  $V_{flowout}$  is the volume of outflow from the impoundment ( $m^3$  of water).

In ArcSWAT, a regression model estimates loadings such as total nitrogen and total phosphorus as a function of impervious area, land use, and rainfall. The general equation developed in ArcSWAT to predict loadings in watersheds is (Neitsch et al., 2005)

$$(4.17) Y = \frac{\beta_0 \cdot (R_{day}/25.4)^{\beta_1} \cdot (DA \cdot imp_{tot}/2.59)^{\beta_2} \cdot (imp_{tot} \cdot 100 + 1)^{\beta_3} \cdot \beta_4}{2.205}$$

Where  $Y$  is the total constituent load (kg),  $R_{day}$  is precipitation on a given day ( $mm H_2O$ ),  $DA$  is the HRU drainage area ( $km^2$ ),  $imp_{tot}$  is the function of the total area that is impervious, and the  $\beta$  variables are regression coefficients. The conversion factors to implement metric units in equations were used: 25.4 mm/inch, 2.59  $km^2/ mi^2$ , and 2.205 lb/kg. ArcSWAT assigns the annual precipitation to each sub-basin by aggregating the monthly precipitation from the weather generator data (Neitsch et al., 2005).

Once the sediment, total nitrogen, and total phosphorus are calculated under each policy scenario, the relative performance of each land use policy scenarios were compared. The ArcSWAT model was run for three policy scenarios as discussed further.

#### 4.4 Policy Determination

The APLUC model provides the capacity to be used as a policy tool for assessing different policies at a watershed scale. One effective way to prevent pollution from residential development is the adaptation or enforcement of a zoning ordinance that prevents residential development from taking place in sensitive locations or, at a minimum, requires specific measures to mitigate the impacts of development. These measures in general include low impact development or riparian buffer policies. In West Virginia, strict land use zoning at the county level may result in political repercussions.

One possible policy to assess is low impact development. Size requirements for developed and open areas are required for low impact development to protect watersheds. Also storm water

best-management practices, such as municipal separate storm sewer systems (MS4) permits must be required from builders for residential land use development (West Virginia WIP Development Team, 2012). Current modeling structure and data limitations do not allow for the construction of parcels that require information on infrastructure cost per acreage or sequential structure for different types of residential development. Therefore low impact development is not feasible in the model.

It also is possible to improve the water quality by creating spatial buffers by drawing a boundary line of a specified horizontal distance from the bank of the streams. In addition, a spatial buffer can be linked to the ArcSWAT model. In this regard, three spatial policy scenarios are considered in this research. These buffer areas are a 50 foot buffer zone for all the streams in the Opequon Creek area (see Appendix I) and a 50 foot buffer zone in critical source areas (see Appendix I). Two types of land use were implemented within buffer zones to observe the effectiveness of the type of buffer. These land use types are agricultural and forest. In the APLUC model, the buffer zone areas are set as no development zones. ArcGIS 10.2 was used to prepare spatially restricted buffers for the APLUC model.

#### **4.5 Baseline Policy**

The baseline scenario involves no additional regulation or spatial restriction in land use conversion. The importance of this policy is to simulate the water quality impacts of land use conversion under the existing regulatory framework in Berkeley County. The spatially restricted policy scenarios will be compared with this policy of no additional buffer zone protection with an assessment of water quality impacts.

#### **4.6 Implementation of 50 foot Riparian Buffer Zone Policy on every Stream in the Watershed**

Riparian buffer zones can prevent adverse impacts to water quality of streams from impervious surface runoff by filtering the nutrients and sediments loadings (Dosskey et al., 2010; Goetz et al., 2003). The use of buffer zones for stream restoration is well documented in watershed management and catchment scale assessment studies (Osborne and Kovacic, 1993; Xiang, 1996; Silva and Williams, 2001). Vegetative buffers along urban streams naturally store and retain nutrients of nitrogen and phosphorus, and have the ability to alter the biochemical properties such as mineralization and denitrification (Groffman and Crawford, 2003). Small watersheds in an urbanized settings are particularly vulnerable to the sedimentation, total phosphorus (TP) and soluble reactive phosphorus (SRP) due to the higher proportion of water-sediment contact zones (Duan et al., 2012).

The Source Water Assessment and Protection (SWAP) Program of Berkeley County, West Virginia recommended riparian buffers to protect the quantity and quality of water from the impacts of impervious surfaces (Berkeley County, 2004). In addition, the Opequon Creek Project Team established buffer streams as a priority in order to raise the community awareness of nonpoint source pollution (Water Resources and TMDL Center, 2008).

An important characteristic of any buffer zone policy is the influence of buffer width on water quality. The Section 402.5.5 of Stream Buffers by Berkeley County, Subdivision Ordinance, sets the minimum of thirty-five feet (35') width on each side of a stream as a buffer with vegetative land cover in the design requirement (Berkeley County Planning Commission, 2009) (See Appendix V). West Virginia Interagency Review Team (IRT) in their WV Stream and Wetland Valuation Metric (SWVM) Development recommended an extended buffer zone width incentive



of inner buffer 0-100' (or 0-50' on each side) and an outer buffer of 101-300' ( or 51-150' on each side) (Hatten et al., 2011). One of the high priority practices identified by the Opequon Implementation Plan Development Steering Committee is a policy of permanent easements on riparian buffers (Opequon Implementation Plan Development Steering Committee, 2006).

In this analysis, a linear 50 foot buffer zone on each side of all streams was delineated as a buffer policy in the Opequon Creek watershed of Berkeley County, WV. Based upon the recommendation of IRT, this research focused on buffer zones surrounding the streams as a spatially restrictive policy for non-point source pollution. Studies found that land use near streams influence hydrological response more than land use in other areas (Johnson et al., 1997; Huang et al., 2013). Using buffer analysis in ArcGIS, a 50 foot buffer on each side of streams was drawn parallel to the stream bank. The streams were delineated by the ArcSWAT model. Buffers were shown as non-developable areas in APLUC model. These areas were later classified in the ArcSWAT model as riparian buffer zones.

The implementation of a buffer policy was regarded as a useful tool for a watershed protection plan of the Opequon Creek watershed. Buffer zoning is assumed as a constraint to residential land use conversion in Opequon Creek, therefore fewer residential land use conversions are expected to result from the buffer policy. The riparian buffer with agricultural and forest land use type in the buffer zone would enhance effectiveness of these buffers due to the associated land use management for water quality. Additionally, the agricultural and distance based buffers are restricted areas but have spatial externalities that influence the parcels' value.

#### **4.7 Implementation of 50 foot Riparian Buffer Zone Policy in CSAs**

Despite the small streams channel size, sub-basins in a watershed system play a key role in nutrient and pollutant loadings (Peterson et al., 2001). Pinoke et al. (2000) examined large amounts of storm flow and nutrient yields in the Chesapeake Bay Watershed and found that that they are tied

to small areas in watersheds. These smaller areas of concern within a watershed are identified as critical source areas (CSAs). Critical source areas are defined as the areas that show highest loading of total phosphorus (TP), total nitrogen (TN), and sediments (S) (Niraula et al., 2013). For water quality management, it is important to identify and assess the impacts of these critical source areas and set them as high priority for land use watershed based policies (Pinoke et al., 2000).

Common approaches to identifying CSAs are index methods, such as the phosphorus index (PI) approach (Srinivasan and McDowell, 2007), the topographic index (TI) approach (Bevin and Kirkby, 1979), and the universal soil loss equation (USLE) factor map (Sivertun and Prange, 2003). However, most of these index approaches do not quantify the nutrient runoff and transformation processes on pollutant loads (Srinivasan et al., 2005). In this research, CSAs are those sub-watersheds identified by the Karigomba (2009) study. In this study, the sub-basins are identified as high priority by both the Watershed Characterization and Modeling System (WCMS) nutrient levels and public participation prioritization method (Karigomba, 2009; Strager et al., 2010). WCMS, developed by the West Virginia University Natural Resources Analysis Center (NRAC), estimates the pollution concentration based upon hydrologically connected digital elevation model, which accounts for flow path and drainage area (NRAC, 2007; Karigomba, 2009). This approach helps in identifying the high priority sub-basins with consideration of hydrological pathways that influence the pollution responses.

Another effective method is a public participation prioritization method utilized by Karigomba (2009). In this prioritization approach, participation from the general public living in the Opequon Creek watershed was used to identify, rank, and prioritize sub-basins within the watershed which show high concentration of pollution and need immediate reductions in pollutant loads. Utilizing the results from Karigomba (2009), three sub-basins: Mill, Tuscarora and Middle

Creeks are identified as high priority sub-basins within the Opequon Creek watershed. The locations of these sub-basins are shown in Appendix I.

## Chapter 5

### 5. Model Validation

Validation is important for assessing and evaluating land use change models (Clarke & Gaydos, 1998; Fragkias and Seto, 2007). The criteria for spatially explicit land use change model validation needs a level of agreement between the projected land use change and actual land use change (Pontius et al., 2004). While calibration seeks to maximize the spatial relationship between the model behavior and historic land use conversion data at specific locations, the results from several studies suggest that, due to the stochastic component in the model and uncertainty resulting from the data, model and several other factors impacting future land use changes, precise land use conversion projections are not always feasible in validation (Pontius and Neeti, 2010; Memerian et al., 2012).

Having these considerations, the validation method employed in this research utilizes Monte Carlo simulations of the observed historical data in order to provide evidence that explanatory variables used in APLUC are suitable to project residential growth in the Opequon Creek watershed. The criterion for validity consists of spatial and statistical validity. Validity was assessed by the projected probabilities to account for the residential land use conversion at the parcel level unit. The model employs 2001 land use data as an initial condition and projected the residential land use conversion for 2011 by employing the APLUC model functional form. All the explanatory variables in validation are calculated by taking the distance measure from each land use, streams, urban center, and roads in the model. The probability of conversion is calibrated using estimated coefficient for each explanatory variable from the SLR model. The model compares the projected outcome with the actual residential land use conversion during 2001-2011. This method not only provides the validation of the overall APLUC model with the spatial

relationships defined in the conceptual model but also validates the parameter values that best project the model outcome into the future.

The results from the actual land use change between 2001 and 2011 shows that 4,748 parcels are converted into residential parcels, which comprise 1,114.85 hectares of land in the Opequon Creek watershed (Table 2 and Figure 6).

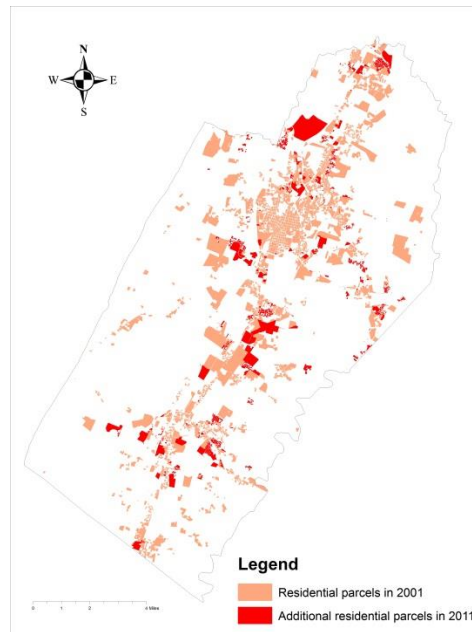


Figure 6. Additional actual residential parcels in 2011.

The residential land use conversions were observed for 10 iterations in the APLUC model to match the 10 year time period between 2001 and 2011. The spatial robustness of the model was assessed through the projections using a Monte Carlo probability derived from 50 model runs. Parcels that had between a 0.50 to 1.00 probabilities of land use conversion over the 50 runs were assumed to convert to residential land use. This probability threshold represents that property parcels that had at least a 50% likelihood of conversion would be converted into residential parcels. The model projected that 2,394 parcels were potentially converted into residential parcels, comprising an area of 1,373.26 hectares (Table 2 and Figure 7).

The spatial and statistical validity of the model was based upon two criteria. First, at least 21.17% of parcels should be precisely measured at the same location where the actual residential development has occurred. This 21.17% threshold criteria was based upon the SLR percent correctly predicted (PCP) cells where the number of fitted conversion matches the number of observed conversions. Second, location at the sub-basin level is an important spatial feature for conversion projection due to the fact that the APLUC model outcome is used in water quality assessment. Therefore, the APLUC model needs to be spatially validated at the sub-basin watershed level compared to actual 2011 land use change data.

The statistical validity was assessed at the parcel level using the percent correctly predicted (PCP) parcels. This method was suggested for a discrete choice models where the calibrated coefficients are utilized for two time periods (Fragkias and Seto, 2007; Pontius et al., 2004). The model accurately replicated 722 parcels (156.67 hectares), which is 15.20% similar spatially located parcels as observed in actual 2011 land use data (Table 2 and Figure 8). Projection accuracy for PCP was measured by spatially intersecting simulated residential land use parcels with actual land use parcel conversions between 2001 and 2011 in ArcGIS (Figure 8).

The correctly projected conversions are lower than what spatial logistic regression predicted in terms of global probability. This difference is expected due to the stochastic nature of the APLUC model. Secondly, the SLR model provides the projected global probability where the threshold probability in each cell would result in conversion, while the APLUC model generates local probability projections of what parcels agents would convert when making a conversion choice based on 500 choices (10 iterations over 50 model runs).

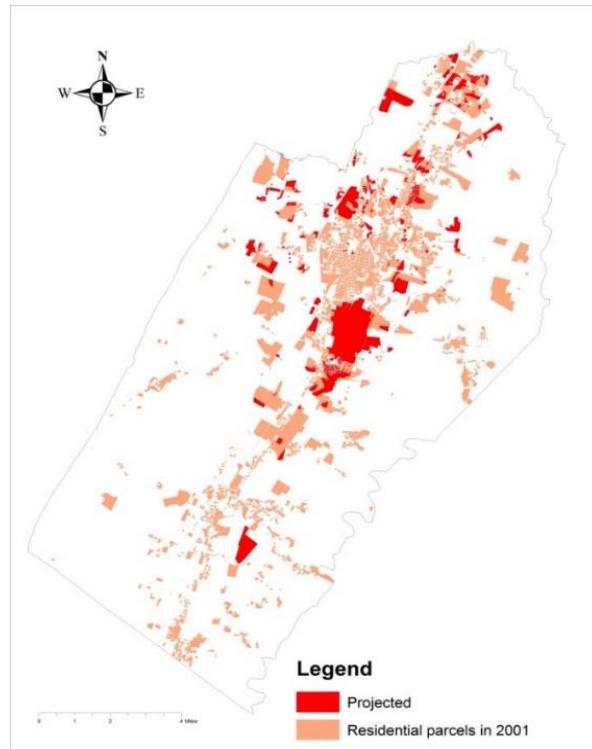


Figure 7. Projected 2011 land use conversions

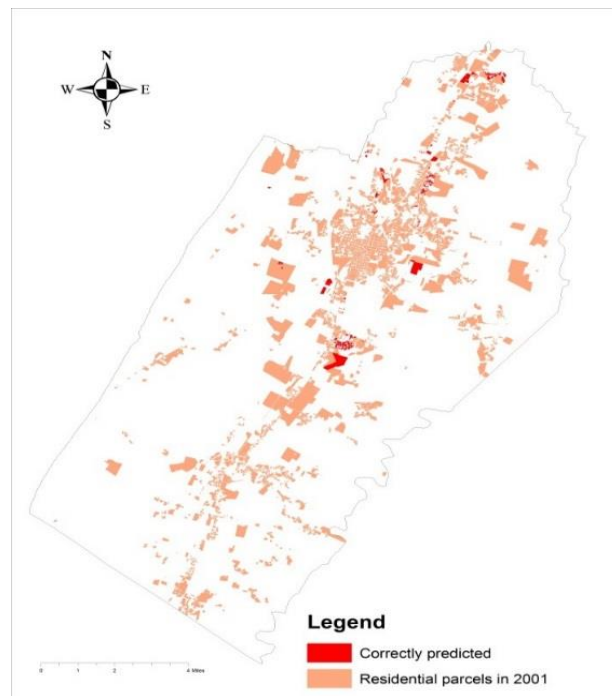


Figure 8. Correctly projected land use conversions

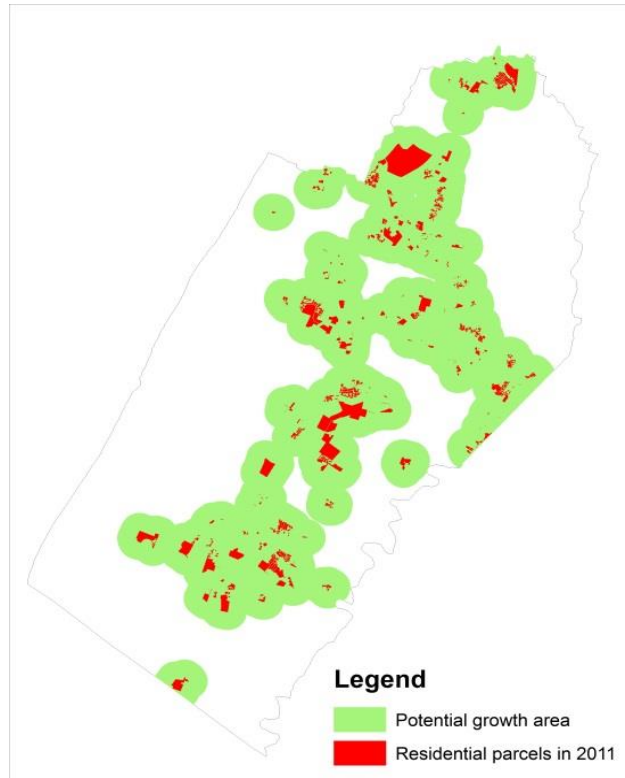


Figure 9. Potential growth zone

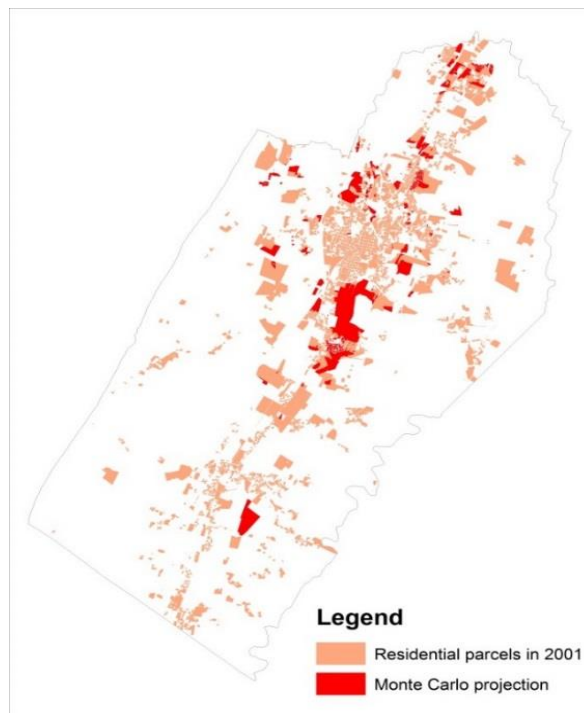


Figure 10. Projected land use conversions within growth zone



Additionally, to investigate how close the projected land use patterns were to observed patterns, one mile buffers from the 2001-2011 residential growth areas were set as a potential standardized geographical area for urban growth zones (Weiler, 2007) (Figure 9). Within the potential growth zones, it was expected that the validation model would closely match the spatial pattern of actual land use conversion occurring from 2001 to 2011. The projected converted parcels within potential growth area are 2,157, comprising 973.11 hectares, which is 87.29% of actual residential land use in terms of area (Figure 10 and Table 2).

Differences between the actual land use conversion in 2011 and projected land use conversion in 2011 model outputs varied in both the magnitude and spatial component of residential land use conversion. However, the historical distribution of residential land use is relatively closer to the observed residential properties in 2011 because the distance based explanatory variables are embedded in the land use distribution, which reflects local characteristics and is predicted by the model. This means that the APLUC model projects land use conversion at those locations which are nearby existing residential areas through the value of spatial externalities.

Using the Monte Carlo projections, the PCP threshold from the SLR and 0.5 mile distance from the residential areas, a cross- tabular comparison was created between the 2011 projected and actual residential land use conversion and compares the results from the projected data sets with the observed dataset. The results suggest that much of the anticipated land use conversion was close but not quite accurate (Table 2 and Figure 10).

Table 2. Comparison among actual and projected residential conversions at parcel level.

Area/ Number of parcels	Actual 2011 additional residential growth	Overall projection	Percent correctly projected (PCP) projection	Projection within growth area
Area ( hectares)	1,114.85	1,373.26	156.67	973.11
# of parcels	4,748	2,394	722	2,157

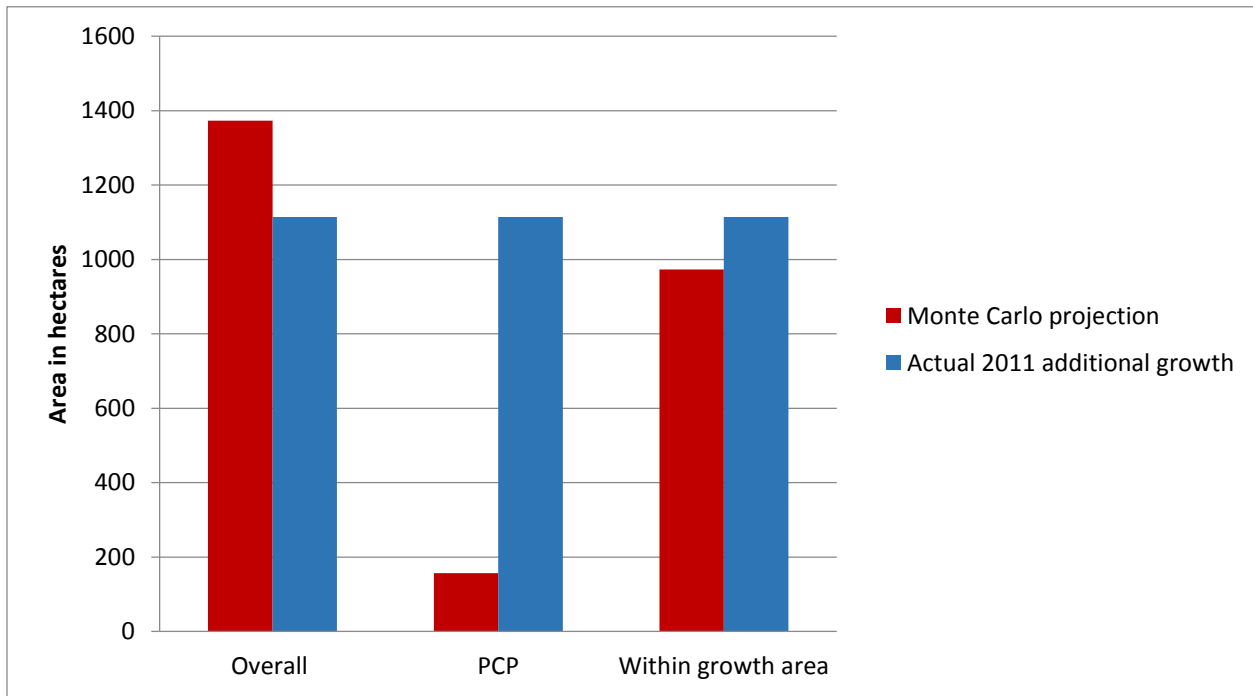


Figure 11. Residential Land Use Projection

The importance of these land use patterns is whether these patterns are statistically and spatially validated at watershed scale. If such a relationship exists, then these pattern can be linked with watershed quality outcomes. In addition to the parcel level validation, sub-basin level validation between projected and actual land use data is examined. The simulated 2011 and actual 2011 land use data are both utilized in ArcSWAT to estimate pollutant loadings by sub-basins. For comparison, each sub-basin was rank ordered by monthly average load over three different pollutants – sediment, TP, and TN. Ranking was on a per hectare basis and in tons for sediments and kg for TP and TN (Figure 12, 13 and 14).

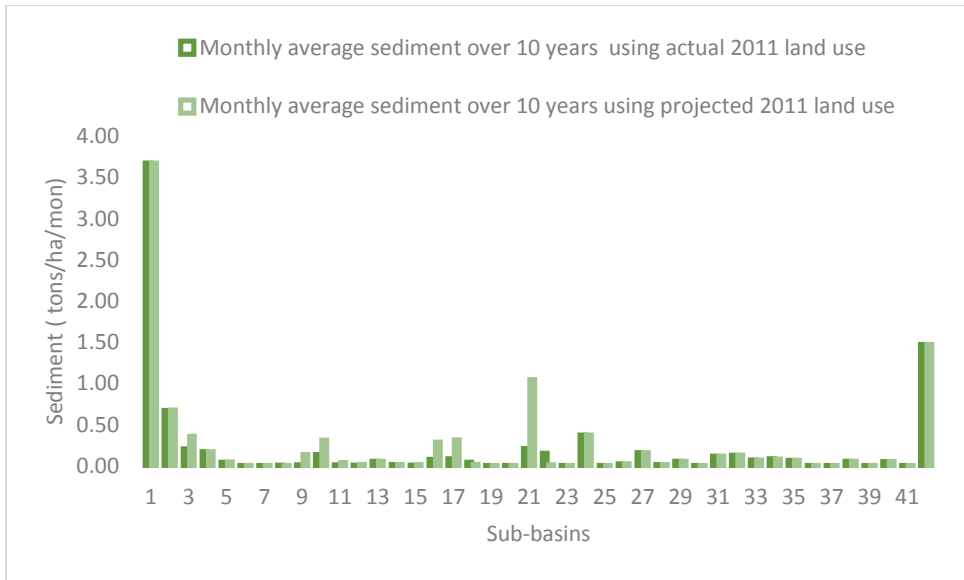


Figure 12. Sediment yield in each sub-basin for actual and projected 2011 land use data

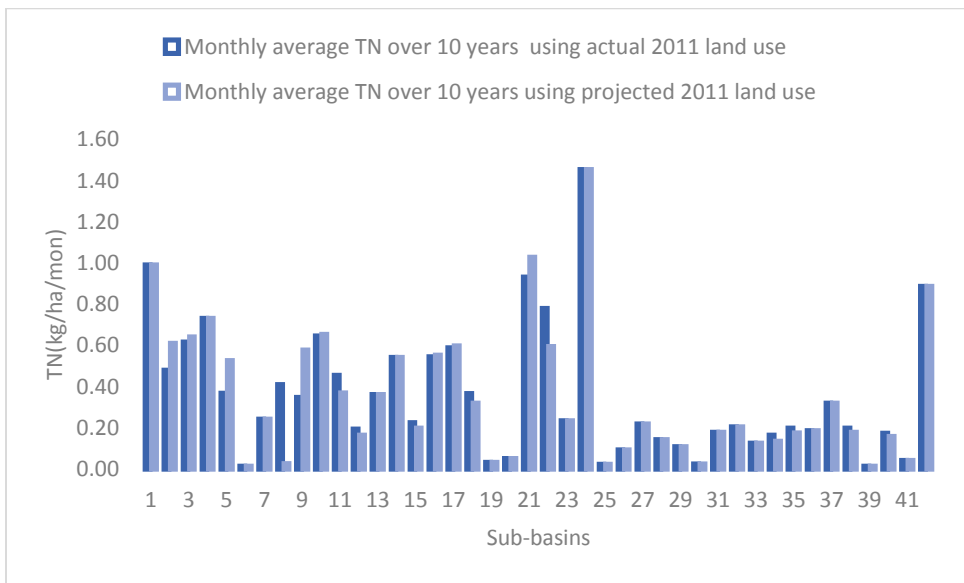


Figure 13. TN yield in each sub-basin for actual and projected 2011 land use data

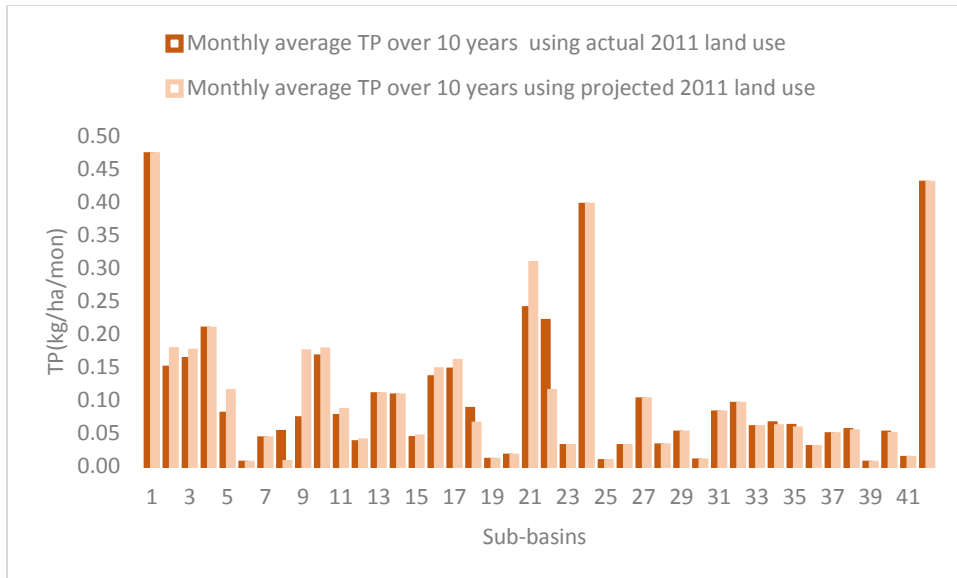


Figure 14. TP yield in each sub-basin for actual and projected 2011 land use data

Figures 12, 13 and 14 show that the overall trend of pollutant loadings for projected land use data is consistent with the actual land use data. However, there are some fluctuations in the cases of monthly average sediment for sub-basins 9 and 21, TP for sub-basins 9, 21 and 22 and in TN for sub-basins 5, 8, 9, 21 and 22 (Figure 12, 13 and 14). For statistical validation, a Spearman Rank correlation coefficient was calculated for each pollutant loading (sediment, TN, and TP) comparing actual land use in 2011 versus projected land use in 2011. The ranking was done in ascending order from highest to lowest pollutant yield.

The Spearman Rank correlation coefficient is defined as:

$$(5.1) \quad \rho = 1 - \frac{6 \sum d_i^2}{42(42^2 - 1)}$$

Where  $d_i^2$  shows the difference between two ranks. The rank coefficient between monthly average loading over 10 years for each sub-basin using actual land use 2011 data and the monthly average loading for sediment, TN, and TP over 10 years for each sub-basin using projected land

use 2011 is 0.60, 0.47, and 0.51 respectively with 42 degrees of freedom (the number of sub-basins) (Table 3).

Table 3. Spearman correlation analysis between actual pollutant loading and projected pollutant loadings for 2011.

<b>Pollutant type</b>	<b>Rank Coefficient (<math>\rho</math>)</b>	<b>Sig.(One-Tailed)</b>
Sediment	0.60	0.000**
Total Nitrogen (TN)	0.47	0.000**
Total Phosphorus (TP)	0.51	0.000**

\*\*P <0.01

The one-tailed value of P at 0.01 significance level is 0.000, 0.000, and 0.000 for sediment, TN, and TP respectively (Table 3). These significance tests show that associations exist between the two land use datasets and are statistically significant for all pollutant types, which provides statistical validity of projected 2011 land use data at a sub-basin watershed level.

# CHAPTER 6

## 6. Data Description

In this chapter, data and variables for general use and all three models (spatial logistic regression (SLR), agent-based probabilistic land use conversion (APLUC), and ArcSWAT) are discussed. Definitions, units, data sources, and preparation of data are described. All spatial data are projected into NAD UTM 1983 Zone 17N.

### 6.1 General Data

#### Land Use

The National Land Cover Database (NLCD) for the years 2001 and 2011 were derived from the Multi-Resolution Land Characteristics Consortium (MLRC) (Homer et al., 2007; Jin et al., 2013). These databases were created by MLRC using Landsat satellite data with a spatial resolution of 30 meters (USGS, 2014). The most recent land use/land cover data available on the MLRC website was for the year 2011. These land use/land cover data were used for all three models. Fifteen land cover classes in land use data for 2001 and 2011 were found for the Opequon Creek watershed of Berkeley County, WV (see Appendix I). The selection of NLCD data for 2001 and 2011 was based upon the pixel to pixel comparison due to similar classification of land use/ land cover.

The land use datasets for 2001 and 2011 were further reclassified into seven land use/ land cover classes using a common scale for the SLR model and for the APLUC model (see Appendix I). The seven land use classes and their definitions utilized in the dataset are shown in Table 4.

The definitions are based upon the MLRC descriptions for each land use/ land cover class (USGS, 2014). The rows and columns of the reclassified 2001 and 2011 datasets were aligned using 30 meter spatial resolution in the geo-processing environment.

Table 4. Land uses/land cover classes and their description, (EPA, 2007; USGS, 2014).

<b>Land uses/land cover</b>	<b>Description</b>
1. Open water	Land comprised of open water with 25% or less of vegetation or soil areas.
2. Open space	Consists less than 20 % of impervious surfaces for total cover and mainly include, large-lot-single family housing units, parks, recreational, and aesthetic areas.
3. Developed, residential	Areas characterized by 20% or greater of constructed materials (e.g. asphalt, concrete, buildings, etc.). Land areas commonly include single, multi-family housing units, and apartment complexes and are characterized by 20% or more impervious surfaces.
4. Barren	Areas of characterized 15% or less vegetation of total cover and commonly include bedrock, strip mines, gravel, and scarps.
5. Forest	Forest areas include mixed forest, deciduous forest, evergreen forest, shrub/scrub, and grassland/herbaceous vegetation.
6. Agriculture	Vegetation for food, feed, or fiber. In general, Berkeley County is dominated by pasture/hay and some areas of cultivated crops.
7. Wetlands	Include woody wetlands and emergent herbaceous wetlands.

Raster based land use data sets for 2001 and 2011 were used in the SLR model. In raster based data, the proportion of each land use type is represented by cell count. The land use data were converted into vector based (polygon data) for use in the APLUC model. Each parcel has a land use category defined by zonal statistics of ArcGIS 10.2. In zonal statistics, zones were defined by property parcels. These zones were based upon the single output value of land use data (the value raster) representing the most common cells within the parcel zone (ESRI, 2011).

## Property Parcel Data

Property parcel level data contained actual property parcels were obtained from the Berkeley County Assessor's office for the year 2011. Property parcel feature data for Berkeley County were extracted for Opequon Creek watershed (see Appendix I). The descriptive statistics in Table 5 show the variation in the size of property parcels in the study area. In Berkeley County, the majority of large size single family housing, farms, and forest properties are located in rural areas, and the small size single and multi-family housing, commercial, and industrial units are in the Martinsburg area. The minimum size of parcel was based upon Berkeley County Commission requirement for minimum residential density (Berkeley County Planning Commission, 2009).

Table 5. Descriptive statistics of parcel size for the Opequon Creek watershed.

Statistics	Value (acres)
Minimum	0.17
Maximum	838.15
Mean	1.66
Standard deviation	11.04

## 6.2 Spatial Logistic Regression Data

### Urban Center

In this research, urban centers were characterized as having three features: (1) those areas with the highest average population density per square mile, (2) a train station with a rail line connected to Baltimore-Washington, D.C. metropolitan area, and (3) being located within the city of Martinsburg. The population density by census tract data provides a demographic basis for the urban fringe (Pozzi and Small, 2005). Population density data for urban centers were collected from U.S. Census Bureau (2000). Data for the year 2001 were not available, therefore, population density data by Census 2000 Tracts was used as a base year for the 2001 data in the SLR and APLUC models. The average population density per square mile for the year 2000 was 825 persons



per square mile in the Opequon Creek watershed. To draw the demographically driven boundary of the urban center, the three highest population tracts with unique six digit codes were selected: 971500, 971600, and 971700. These tracts had population densities of 2,846, 2,705 and 3,340 persons per square mile, respectively. The total area of this urban center was found to be 4.51 square miles. Martinsburg was identified as major urban activity center within the Berkeley County in the model.

Another important feature for Berkeley County was accessibility of the Baltimore-Washington metropolitan area through public transportation. The Martinsburg train station provides service for the Maryland Area Regional Commuter (MARC) train that connects Martinsburg, WV to Harford County, Maryland; Baltimore City; Washington D.C.; Brunswick, Maryland and Frederick, Maryland (DOT, Maryland Transit Administration, 2014). The final layer of the urban center was created from the centroid of Martinsburg and centroid of the train station using a point feature data layer and extracted within the demographic boundary of the urban center (see Appendix I).

### **Major highways**

Major highways as road features were collected from the U.S. Department of Transportation (1997). In general, road features typically remain constant over long periods of time. Data for 1997 was used for both the baseline year 2001 in the SLR model and for the baseline year 2011 in the APLUC model. Major highways Interstate-81 and U.S.11 were selected for the study area (see Appendix I).

### **Streams**

Data on streams were delineated through the ArcSWAT Watershed delineation based on digital elevation model (DEM) raster for the Opequon Creek watershed. The elevation was in meters having 30 x 30 cell size. ArcSWAT draws the location of the stream network based upon the flow

direction and accumulation using DEM grid (see Appendix I). The minimum and maximum, and ArcSWAT defined sub-watershed drainage areas were 107, 21,327, and 426.54 hectares, respectively.

Six explanatory variables were used in the SLR model. These variables are distance to residential land use, distance to agricultural land use, distance to forest land use, distance to urban center, distance to streams, and distance to major highways (see Appendix I). All variables are raster based layers in the SLR model. The layers for distance to residential land use, agricultural, forest, streams, urban centers, and major highways were calculated using the cost distance tool in ArcGIS. Description and mean values for each explanatory variable is shown in Table 6.

Table 6. Explanatory variables descriptions and mean values.

<b>Variable</b>	<b>Description</b>	<b>Mean ( meters)</b>
agdist	Distance from the cell to the nearest agricultural cell	45.04
foresdist	Distance from the cell to the nearest forest cell	73.24
residist	Distance from the cell to the nearest residential cell	483.32
urbandist	Distance from the cell to the nearest urban center	7848.16
highwaydist	Distance from the cell to the nearest major highway	2985.70
streamdist	Distance from the cell to the nearest stream	473.22

### **6.3 Data for ArcSWAT**

#### **Digital elevation model (DEM)**

The digital elevation model (DEM) data for elevation was obtained from the National Elevation Dataset (NED) by the U.S. Geological Survey to define flow direction, flow accumulation, and

watershed delineation. This DEM dataset was in a raster format with spatial resolution of 30 meters. 30 meter resolution were found consistent with land use dataset from NLCD.

### **Land use**

Three land use data sets were utilized in ArcSWAT: (1) NLCD data for year 2011, (2) projected 2011 land use data from the APLUC model for model validation; and (3) simulated land use data under each policy scenario obtained from the APLUC model.

### **Soil**

The Soil Survey Geographic (SSURGO) database was used in ArcSWAT in defining the hydrological response units. SSURGO data provides smaller polygons (soil map units) and higher resolution with fine details (see Appendix I). Each soil map unit represents a soil type in the Opequon Creek. The detailed SSURGO database is well suited for a small scale study area (Mednick et al., 2008). Due to having correct soil texture details, simulations generated using SSURGO database improve the prediction in small scale watersheds (Anderson et al., 2006).

### **Slope**

In ArcSWAT multiple slope classes in percentage were used to define the hydrological response units. The slope is broken down into five classes to represent the variation in topography of the Opequon Creek watershed (see Appendix I). The majority of the Opequon Creek area is having a relatively flatter terrain (range between 0-19 percentages of slope) (see Appendix I).

### **Weather Data**

Weather data were obtained from monthly weather database from first order stations in ArcSWAT, which has weather data from the year 1960 to 2010. The data on rainfall, temperature, relative humidity, solar radiation and wind speed were simulated using the weather generator function in ArcSWAT for each sub-basin.

# Chapter 7

## 7 Results

This chapter analyses and interprets the empirical results of the three linked models utilized in this research: spatial logistic regression (SLR), agent-based probabilistic land use conversion (APLUC), and ArcSWAT (ArcGIS extension of Soil and Water Assessment Tool). The spatial logistic regression (SLR) model provides the empirically estimated spatial externalities to be used in the APLUC model. The application of the spatially explicit land use conversion model helps to explore the spatial location and patterns of residential land use in Berkeley County at the watershed scale via a computational modeling approach. The APLUC projects the spatial pattern of land use conversion under each policy scenario and summarizes the findings of the simulations. The impacts of each policy scenario on the surface water quality in the Opequon Creek watershed are assessed with the ArcSWAT model.

### 7.1 Spatial Logistic Model Analysis and Results

The overall model statistics from logistic regression analysis are summarized in Table 7. By using a pixel thinning method in the IDRISI software, negative impacts of spatial interdependence are reduced. The result is 3,468 sampled observations are used in SLR, of which 2.45% (85 cells) are converted from non-residential to residentially developed cells between 2001 and 2011. Statistical significance for the overall model in logistic regression is tested by a chi-square.

Table 7. Spatial logistic regression model results.

Statistics	Value
Number of total observations	3,468
Number / percentage of 0s in sampled area	3,383 (97.55%)
Number/ percentage of 1s in sampled area	85 (2.45%)
Chi-square (6)	123.82 (p-value 0.00001)
-2logL0	798.37
-2log(likelihood)	674.55

The chi-square is also known as a likelihood ratio statistic, which is comparable to the F statistics in linear regression analysis. The SLR model uses maximum likelihood estimation (MLE) to find the best fitting set of parameters (coefficients). The maximum likelihood function is defined as (Clark Labs, 2014):

$$(7.1) \quad L = \prod_{i=1}^N \mu_i^{y_i} * (1 - \mu_i)^{(1-y_i)}$$

Where L is the likelihood,  $\mu_i$  is the predicted value of the dependent variable for sample i, and  $y_i$  is the observed value of dependent variable for sample i. The predicted value of the dependent variable for sample i is calculated as:

$$(7.2) \quad \mu_i = \frac{\exp(\sum_{k=0}^6 b_k x_{ik})}{(1 + \exp(\sum_{k=0}^6 b_k x_{ik}))}$$

Where  $k$  is the number of observable variables, and  $b_k$  is the coefficient of each variable ( $x$ ). To maximize the equation 7.2, the solution to the following simultaneous nonlinear equations are solved by IDRISI Selva software using the Newton-Raphson algorithm (Clark Labs, 2012):

$$(7.3) \quad \sum_{i=1}^N (y_i - u_i) * x_{ij} = 0$$

Where  $x_{ij}$  is the observed value of the independent variable  $j$  for sample  $i$ .

The goodness of fit test in SLR was based upon the likelihood ratio principle, where the ratio is defined through the following statistics (7.4) and (7.5) (Clark Labs, 2012):

$$(7.4) \quad -2\log(L_0)$$

Where  $L_0$  represents the value of the likelihood function if all coefficients other than the intercept term are 0:

$$(7.5) \quad -2\log(\text{Likelihood})$$

Where Likelihood represents the value of the likelihood function for the full model as fitted. Based upon equations (7.4) and (7.5), Chi Square ( $k$ ) are computed as:

$$(7.6) \quad \text{ChiSquare} (k) = -2(\log(\text{likelihood}) - \log(L_0))$$

With a chi-square value of 123.82, the null hypothesis that distances from economic location, amenities, or surrounding land uses have no impact on residential land use conversion is rejected (p-value is 0.00001).

Initially, cells were classified using a SLR predicted probability threshold of 0.5 where cell probabilities of less than 0.5 were classified as not converted (0) and cell probabilities greater than 0.5 were classified as converted (1). The two-by-two contingency table is shown in Table 8.

Table 8. Classification of cases & odds ratio using 0.5 threshold.

<b>Observed</b>	<b>Fitted 0</b>	<b>Fitted 1</b>	<b>Percent correct</b>
0 (No conversion)	3381 ( $f_{11}$ )	2 ( $f_{12}$ )	99.9409
1 (Residential conversion)	83 ( $f_{21}$ )	2 ( $f_{22}$ )	2.3529

Based upon this threshold, the odds ratio is calculated as (Clark Labs, 2012):

$$(7.7) \text{ Odds Ratio} = \frac{(f_{11} * f_{22})}{(f_{12} * f_{21})}$$

The resulting odds ratio is 40.73 with a 0.50 threshold probability. Instead of using 0.50 threshold for conversions, SLR employs a new threshold of 0.11, which determines that land use conversion occurs at the cell (pixel) level when the probability is 0.11 or above. This means that if  $x\beta_i \geq 0.11$ , the cell will convert into residential land use and if  $x\beta_i \leq 0.11$ , the cell does not undergo any change. The resulting outcome provides the percent correct land use conversion and non-conversions (Table 9). The resulting value of the adjusted odds ratio is 13.29.

Table 9. Classification of cases & odds ratio using 0.11 threshold.

<b>Observed</b>	<b>Fitted 0</b>	<b>Fitted 1</b>	<b>Percent correct</b>
0 (No conversion)	3316 ( $f_{11}$ )	67 ( $f_{12}$ )	98.0195
1 (Residential conversion)	67 ( $f_{21}$ )	18 ( $f_{22}$ )	21.1765

By assigning a 0.11 cutting threshold for predicted probability, true positive and false positive are calculated as (Clark Labs, 2012):

$$\text{True Positive} = \frac{f_{22}}{(f_{21} + f_{22})}$$

$$\text{False Positive} = \frac{f_{12}}{(f_{11} + f_{12})}$$

The value of true positive is 46.15% and false positive is 1.98%. ROC is performed by comparing the fitted cells that are converted and actual cells that are converted during 2001-2011. ROC represents the model's ability to predict the probability of conversion at various locations in the study area (Tayyebi et al., 2010). The resultant ROC for the SLR model shows a higher value of 0.80 with 100 thresholds (Table 10).

There are 85 cells that are converted into residential land use and are represented by 1s in the model. The fitted 1s are 18 and fitted 0s are 6. Therefore there are 21.17% correctly predicted cells compared to the actual land use change. Essentially, this indicates that 21.17% percent are correctly predicted in terms of exact location by fitting the residential land use conversion with the number of observed residential land use in the dependent variable. Since most of the land within 10 years did not change. Thus, the probability of no conversion is highest. Therefore, 98.01% of the cells are correctly predicted for non-conversions, which shows that the model is good in predicting non-conversions.

The logistic regression model results showing parameter estimates are shown in Table 10. In general, the signs of the coefficients show residential land use conversion trends that are consistent with the study area. Positive coefficient signs for explanatory variables indicate that as the distance from the cell of explanatory variable to an economic location, amenities, or



neighboring land use gets larger, the impact on land use conversion probability gets higher. Conversely, negative coefficient signs show the land use conversion probability increases as the distance decreases between a cell and economic locations, amenities, or land uses.

The results in Table 10 indicate that the closer a non-residential cell is to surrounding residential land, highways, or urban center the higher the probability of conversion. The negative coefficient for urban center is consistent with Von Thünen and bid-rent theory. Positive coefficient signs are estimated for distance to forest land use and distance to agricultural land use. These signs are consistent with expected spatial influences of these land use types because in most of these areas, residential land use conversion is limited by the availability of the public water and the public sewer systems (Berkeley County Planning Commission, 2006).

The Z statistics of coefficient estimates were not included due to the unavailability of standard errors in IDRISI Selva software. In general, Z statistics for raster based SLR are not very common in land use research (Cohen et al., 2005; Gorsevski et al., 2006; Polyakov and Zhang, 2008; Arsanjani et al., 2013). Alternatively, likelihood ratio tests provide a measure of the significance for individual predictors by deviance statistics. The deviance statistics are generally known as negative two log likelihood (-2LL) (Cohen et al., 2003). Each explanatory variable is judged by comparing the deviance with a full model with the deviance in the model without a predictor. The difference between these two deviance values is Chi-square test for goodness of fit and calculated as follows:

$$(7.8) \quad \begin{aligned} \chi^2 &= D_k - D_{k-1} \\ &= -2LL_k - (-2LL_{k-1}) \end{aligned}$$

Where  $D_k$  is the deviance for the model containing all k variable (explanatory variables) and  $D_{k-1}$  is the deviance for the model with one fewer explanatory variable. This would help in testing the

significance of the predictor that is not included in  $D_{k-1}$ . To test the significance, Chi-square is assessed using five degrees of freedom (number of predictors minus one).

Table 10. The Likelihood ratio statistics and chi-square significance for each predictor.

Explanatory Variables	Coefficient values	$-2LL_k$	$-2LL_{k-1}$	$\chi^2 = -2LL_k - (-2LL_{k-1})$	$P >  \chi^2$
agdist	0.0034	674.550	686.737	12.186	0.032*
foresdist	0.0040	674.550	710.936	36.385	<0.000**
residdist	-0.0027	674.550	701.628	27.077	<0.000**
streamdist	-0.0004	674.550	676.299	1.748	0.883
highwaydist	-0.0001	674.550	678.538	3.987	0.551
urbandist	-0.0001	674.550	690.628	16.077	0.007**
Constant	-2.2730	-	-	-	-
ROC	0.8056	-	-	-	-

ROC=1 indicates a perfect fit; and ROC=0.5 indicates a random fit. \* shows significant at 0.05 and \*\* shows significant at 0.01.

The results are reported in Table 10. Except distance to highway and distance to streams, all other predictors are found to be statistically significant. Based upon the study area, streams and highways are recognized as important factors in land use conversion decisions. Therefore, these variables are not omitted on basis of the significance test results.

The SLR function represents coefficient values based upon the raster based data derived for the study area. These coefficients can be regarded as weights to produce the global probability of change (Shirzadi et al., 2012). Interpretation of these coefficients in terms of their relationship with the dependent variable (Boolean map of land use conversion 2001-2011) is not feasible in a logistic regression functional form due to non-linearity. Therefore, in order to show the

relationship between the explanatory variable and the dependent variable, marginal effects are analyzed as shown in equation 4.3 of chapter 4.

The marginal effects calculated from the logistic regression estimated results are reported in Table 11. These results represent the per meter change in land use conversion probability for each explanatory variables. The marginal effect represents a one meter closer distance. Therefore, in terms of interpretation, each explanatory variable is linear with its influence on conversion probability. The marginal effects further were multiplied by a conversion factor of 1000 to represent marginal effect per kilometer. The negative relationship between the distance from forest and agriculture and land use conversion probability indicate that the cells that are located farther from the agriculture and forest cells have higher probability of land use conversion. While cells located closer to the residential, urban center, highway, and streams have higher land use conversion probability and thus the marginal effects of these variables are positive.

In addition to the marginal effects, the change in odds value show the percentage increase and decrease in the odds of cells being converted into residential land use (Table 11). Since the values are in meters, compared to a watershed area which is kilometers in area, the value of the change in odds is very small for all variables. The results show relatively high percentage decrease in the odds of being residential land use due to close proximity to the agricultural and forest land cells.

Table 11. Marginal effects, change in odds and percentage change in the odds.

<b>Explanatory Variables</b>	<b>Marginal effects (meters)</b>	<b>Marginal effects (km)</b>	<b>Change in odds</b>	<b>% Change in the odds</b>
agdist	0.0034	-3.3632	0.996642449	0.34% (decrease)
foresdist	0.0040	-3.9604	0.996047432	0.4%(decrease)
residdist	-0.0027	2.66703	1.00267059	0.2% (increase)
streamdist	-0.0004	0.44498	1.000445079	0.04% (increase)
highwaydist	-0.0001	0.12974	1.000129748	0.01% ( increase)
urbandist	-0.0001	0.10346	1.000103465	0.01% (increase)

The SLR model predicts the conversion probability of each cell to be converted into residential land use, from which a probability map was generated (Figure 15). The highest probability of residential land use conversion is 0.99 and the lowest is 0.00 (Figure 15). The maps shows the highest probability is in and around the urban center (Martinsburg area). The red area in the map shows more than 0.60 probability of conversion. Therefore for each single developable cell located in red area has more than 60% chance of conversion.

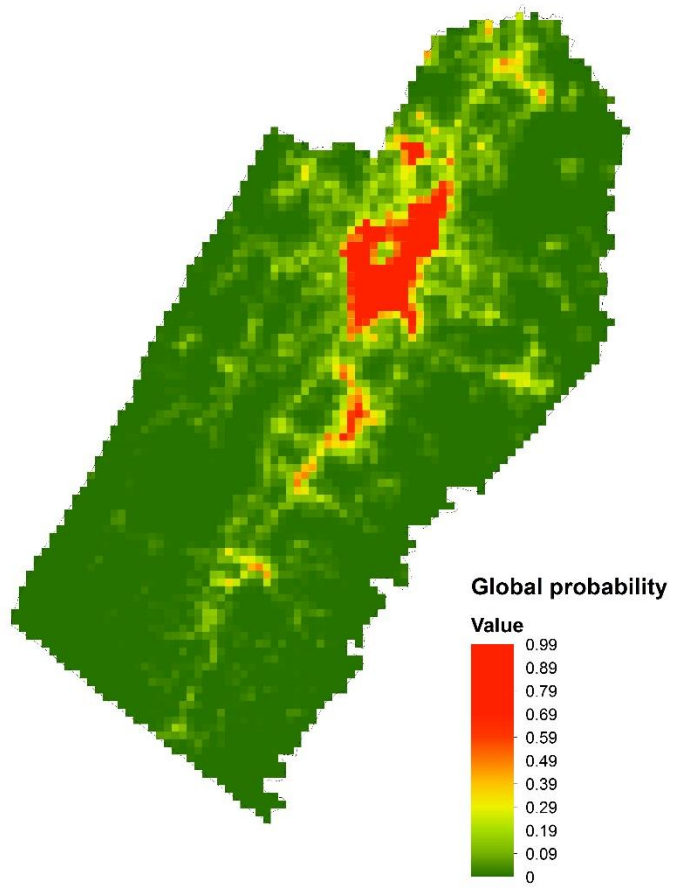


Figure 15. Predicted land use conversion global probability.

## 7.2 Agent-based Probabilistic Land Use Conversion Model Results

The parameter values were calibrated through spatial logistic regression and incorporated into agents' land use conversion decisions.

The  $\beta X$  function is calculated as:

$$(7.9) \beta X = -2.2730 + 0.003363 * agric_{dist} + 0.003960 * forest_{dist} - 0.002667 * resid_{dist} - 0.00044 * stream_{dist} - 0.000129 * highway_{dist} - 0.000103 * urban_{dist}$$

The local probabilities were stochastically run through the Monte Carlo method. Land use change and changes in the probability due to land use conversions was recorded using a vector map. The model projections were a set of 50 Monte Carlo simulations of land use conversion showing projected developed and non-developed land use patterns over 50 model runs, each run consisting of 10 iterations.

The consistency and stability across these runs were considered important and observed within model iterations and across model runs. The overall results from each policy scenario suggests that there is not much variation between models runs (See Appendix II). Model calibration also impacted the parameter setting of the model. Thus, the results show deterministic behavior within stochastic settings. Further, Monte Carlo probabilities are mapped for each property parcel. For model testing, the spatial land use patterns were observed in each iteration. It was observed in each model run that the number of conversions is higher in earlier iterations compare to later iterations.

The parcel level probabilities from the Monte Carlo simulations were divided into two classes where a probability less than 0.50 was considered as non-conversion and parcels with a probability equal or more than 0.50 are considered as residentially converted parcels. The 0.50 threshold of probability has been suggested by various research studies to match the urban growth

rates for the calibration period (Brooks et al., 1999; Fragkias and Seto, 2007). Compared to the probability observed by SLR cutting threshold, this 0.5 threshold is a relatively strict criteria.

The various sizes of land parcels and neighboring interactions with surrounding land uses incorporate spatial heterogeneity into the APLUC model. Differences among parcel agents are due to differing land use types surrounding each parcel based on its location. Resulting differences in conversion probabilities were then assigned to land use conversion decisions by comparing the probability value with a random number in each iteration to incorporate uncertainty into the APLUC model. This accounts for stochastic element within the model.

To illustrate the capability of APLUC to simulate alternative scenarios, three major policy scenarios were tested to project anticipated future residential land use development for the Opequon Creek watershed. These policies included: a baseline policy, which incorporated no additional regulation on residential land use conversions, a policy with 50 foot buffers along all streams in the Opequon Creek watershed, and a 50 foot buffer policy only within critical source areas sub-basins identified by Karigomba (2009). Two types of buffers are employed in the model: agricultural land use buffer and forest land use buffer. Thus, a total of five policy scenarios were tested. The results show a considerable variation in land use conversions and patterns for each of the policy scenarios.

Conversion decision spatially linked with adaptation of buffer policies facilitates the comparisons and assessments of the potential impacts of policies on watershed changes. For example, with these results it is possible to compute the potential number of conversions in a certain policy, as well as the potential changes in allocation of land use in each sub-basin due to these policies.

Table 12. Descriptive Statistics of parcel conversions over 50 model runs in each policy scenario.

Policy Scenario	Mean # of converted residential parcels	Minimum ( # of converted residential parcels)	Maximum ( # of converted residential parcels)
Baseline Policy	1,527	1,486	1,562
50 ft buffers around all streams (agricultural)	399	358	430
50 ft buffers around all streams (forest)	442	422	465
50 ft buffers in CSAs (agricultural)	1,277	1,250	1,306
50 ft buffers in CSAs (forest)	1,441	1,421	1,473

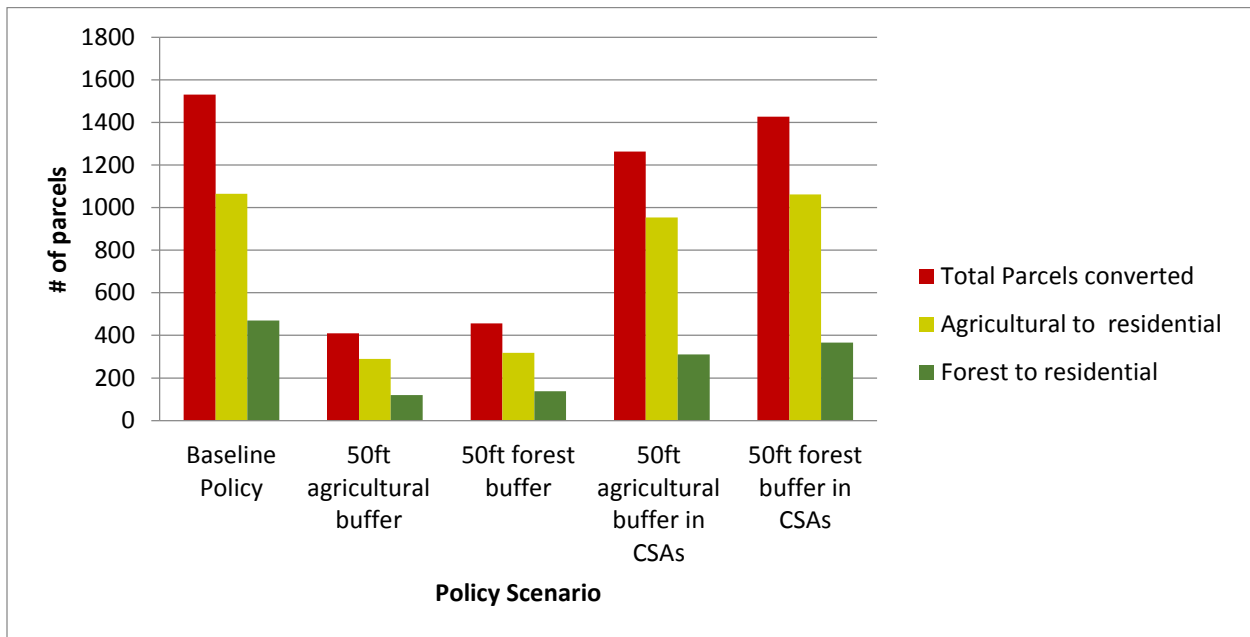


Figure 16. Number of parcels converted from each developable land use type



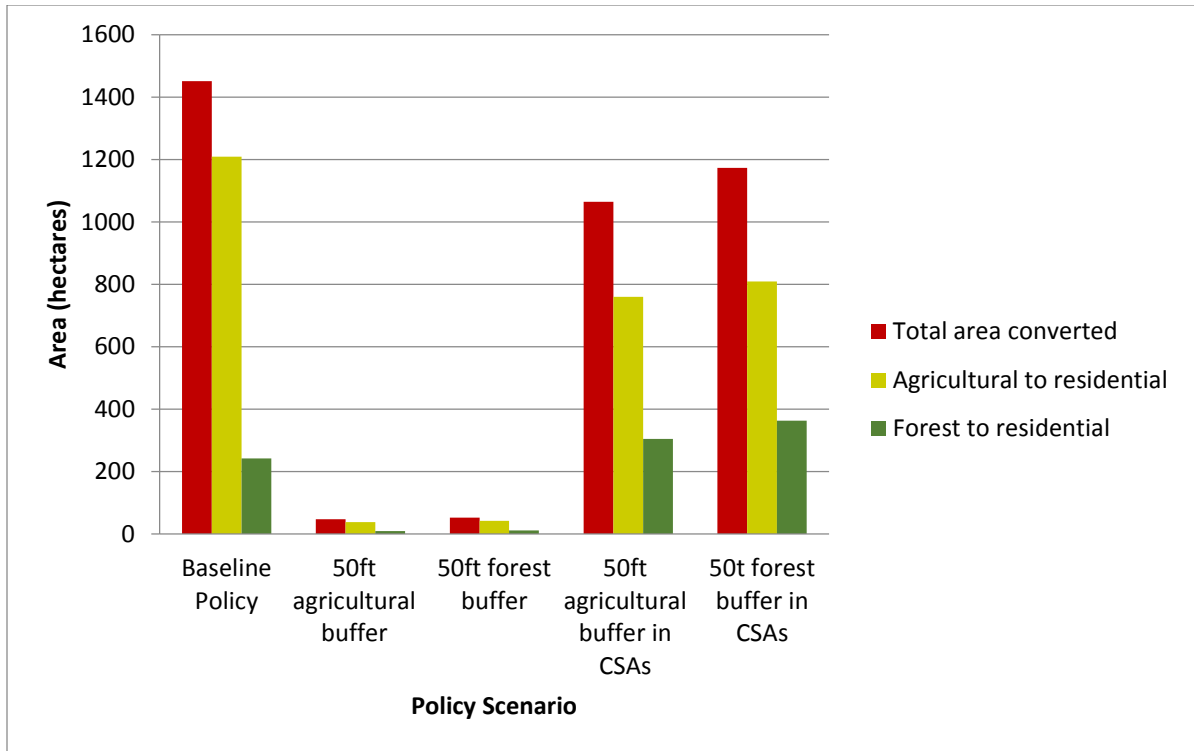


Figure 17. Area of land converted from each developable land use type

### Baseline Policy Scenario

In the baseline policy scenario, there are no additional density restrictions from city or local governments on residential development. This scenario approximates the current policy in Berkeley County, WV. Results show the highest number of land use conversion in the Opequon Creek watershed results from baseline policy scenario compared to all buffer policies (Table 12 & Figures 16-18). Under the baseline policy, converted residential parcel during 50 models runs ranged between 1,486 to 1,562 property parcels (Table 12). The Monte Carlo projections under this policy scenario shows that, on average, a total of 1,531 parcels are converted, which consisted of 1451.26 hectares (or 4.08% of total land) (see Appendix II). Most of the residential conversions took place on larger, agricultural property parcels, thereby causing a decrease in agricultural area of 12.88% from the initial state (9,379.92 hectares) (see Appendix II). Overall, 83.29% residential land use change (in terms of area) occurred from agricultural land consisting of 1,064 previously

agricultural parcels with total area of 1,208.81 hectares (see Appendix II). This could be caused by a high percentage of agricultural land use (developable parcels) being closer to residential parcels. Following SLR results, parcels closer to existing residential properties had higher impacts on residential conversion probability than parcels further away. Most of the predicted conversions captured in the model are close to the existing residential properties, which is the most predictable type of residential growth with current calibration from SLR (Figure 18). Conservatively, these potential converted parcels by Monte Carlo simulations are regarded as potential residential growth areas. Results also show that majority developable parcels have a clear tendency to convert into residentially developed areas closer to the urban center (Martinsburg area) in the baseline policy scenario.

Additionally, land use conversions from forest to residential land use are taking a small proportion of the total residential land use conversions (Figure 16 and 17). Only 16.70 % of land area converted into residential land from forest land use (see Appendix II). Most of the forest parcels are located in close proximity to agricultural parcels in the study area, which makes forest parcels less likely to convert. But residential parcels are more likely to convert when forests are at a distance from the properties. These findings suggest that forest can serve as “distant amenities” and influence property parcels conversion at distances.

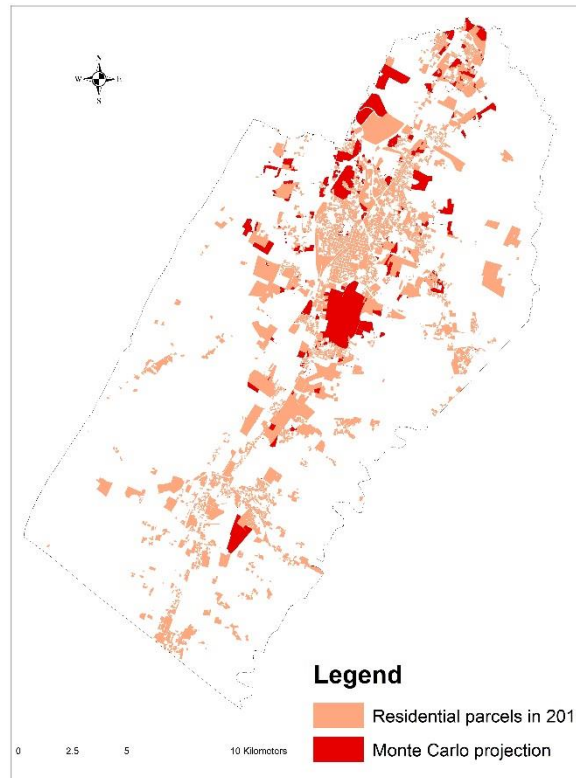


Figure 18. Land use conversions under baseline policy

### **50 Foot Buffer Policy**

A 50 foot riparian buffer policy scenario developed in the APLUC model that incorporates residential density restriction in a buffer zone to protect water quality. The 50 foot riparian buffer is a spatially restricted area around all streams on each side. The adoption of the riparian buffer policy restricted agents from developing residential land in buffer zones and resulted in less residential conversion projections.

This policy was analyzed with agricultural and forest land use types as buffer zones. The application of these two types of buffers was not only based upon the high percentage of these land uses in the area but also captures the influence of these land uses on conversion probability. In the SLR model, the coefficient value is positive for both the distance from forest and agricultural land

use. These distance based spatial externalities result in a decrease in probability of conversion of many properties that are closer to agricultural and forest land use. Therefore, not only do agricultural and forest buffers constrain the development of residential land, but additionally act as spatial externalities which reduce nearby conversions (Figure 19 and 20).

Results from conversion trends across 50 model runs show that the average conversion rate was 399 parcels under a 50 foot agricultural buffer within 50 model runs (Table 12). However, there was a slight increase (10.77%) in the average rate of parcel conversions with forest buffers compare to agricultural buffers (Table 12). As fewer parcels converted, the Monte Carlo based projected total residential land use area drastically dropped to 96.76% and 96.36% less total area compared to baseline policy in agricultural and forest type buffers respectively (see Appendix II).

Buffer policies also showed differences between parcel agent types by property size (Figure 19 and 20). The spatial distribution of parcel agent by size was not homogeneous throughout the watershed, and therefore, the adoption of the policy also was not equally distributed among small and large parcel agents. Since large parcels occupy most of the area, most of the streams are spatially located across these parcels. Specifically, it is likely that many large parcel agents were required to adopt a buffer policy and resulting in a decrease of these large parcels conversions. The change converted land use percentages between agricultural and forest type buffer was due to the fact that the value of the forest parameter in SLR is slightly higher than agricultural, which results in relatively higher parcels conversion under a 50 foot forest type riparian buffer (Figure 20 and Appendix II). Intuitively, the results suggest that forest can serve as a distance amenity to developable parcel properties that would slightly increase the land use conversion rate, which still remained far below the baseline policy's conversion rate.

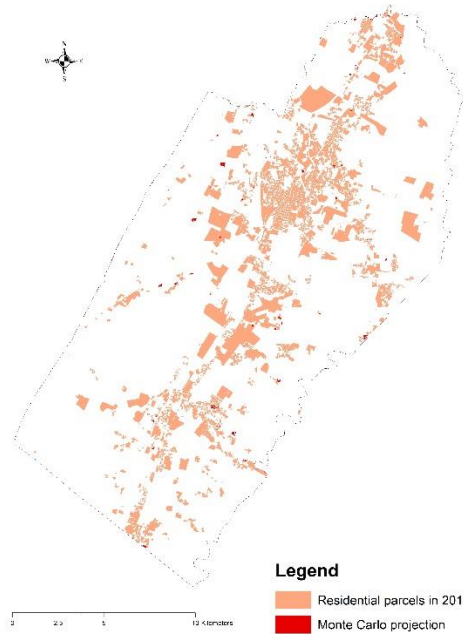


Figure 19. Land use conversions under 50 ft agricultural buffer scenario

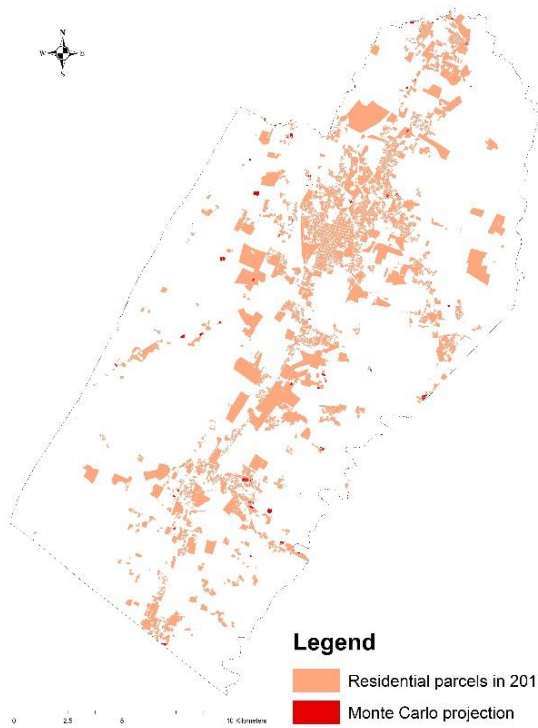


Figure 20. Land use conversions under 50 ft forest buffer scenario

## 50 Foot Buffer in Critical Source Areas (CSAs)

Under a 50 foot riparian buffer in high priority sub-basins in Opequon Creek watershed scenario, there was very little decline in land use conversion compared to the baseline policy. Under this policy, 1,064.39 hectares of land converted into residential land use with agricultural buffers and 1,172.72 hectares of land under forest buffers (see Appendix II).

Similar to the previous results of a 50 foot buffer on all streams, more land was converted into residential land use from agricultural land compared to forest land in each type of buffers (Figure 21 and 22). However, the 10.17% higher area converted when forest buffers were implemented compared to agricultural buffers (see Appendix II).

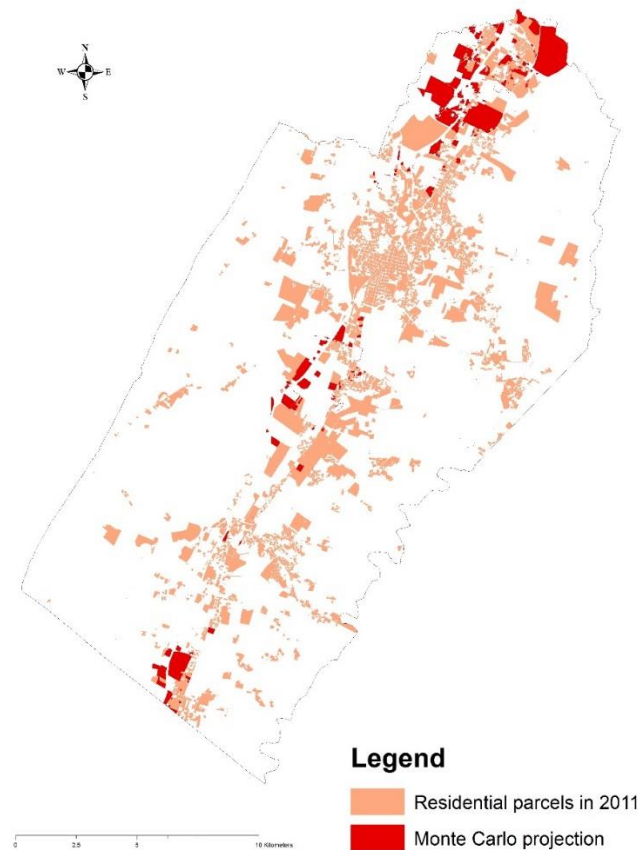


Figure 21. Land use conversions under 50 ft agricultural buffers in CSAs

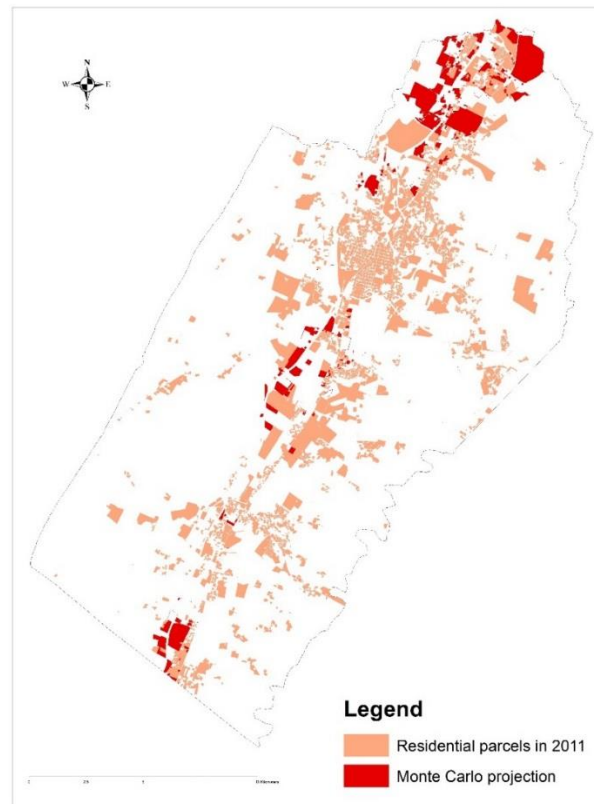


Figure 22. Land use conversions under 50 ft forest buffers in CSAs

With a buffer policy only within CSA sub-basins watersheds, most conversions are occurring along the major highways. Specifically, bigger parcels are converted and located in the northern part of the Opequon Creek watershed. This result implied that the presence of highways is more influential on residential land use conversion under this policy. Lastly, these projected land use conversions were used as data inputs for future land use patterns within the Opequon Creek watershed to link with the ArcSWAT model for investigating the resultant water quality outcomes.

### **7.3 ArcSWAT Model Results**

Land use data layers generated from the APLUC model for each policy scenario along with actual land use from 2011 were implemented in ArcSWAT. From these land use data, pollutant loads were calculated for each sub-basin in the watershed over a 10 year simulation period (see Appendix III). Land use classifications within ArcSWAT generated four types of land uses: residential, open space, forest, and agriculture (Figure 23 and Appendix IV). ArcSWAT has its own classification system to define land use types for each sub-basin. In ArcSWAT, residential land use is defined as high density residential land while open space includes areas that have mixed land use including major highways, low density residential areas along with the combination of forest and agricultural related vegetation. In this research, pasture was assigned to an agricultural land use class due to its being the observed dominant land use in Berkeley County. Deciduous forest was selected for forest class in ArcSWAT due to its observed high percentage of land use in Berkeley County.

The details of the Opequon Creek watershed in hectares and the percentage of each land use type is shown in Figure 23 and Table 13. The dominant land use in each sub-basin watershed was calculated in ArcSWAT using actual 2011 land use data and land use data generated by the APLUC model (Figure 24). The highest residential land use was found in the baseline scenario and with the 50 foot buffer in CSAs only (either agricultural and forest). To eliminate the effect of seasonal differences in flows and loadings, average of monthly loadings over a 10 year period were used for each sub-basin. Table 14 shows the average pollutant per month within 10 year simulation in Opequon Creek watershed.



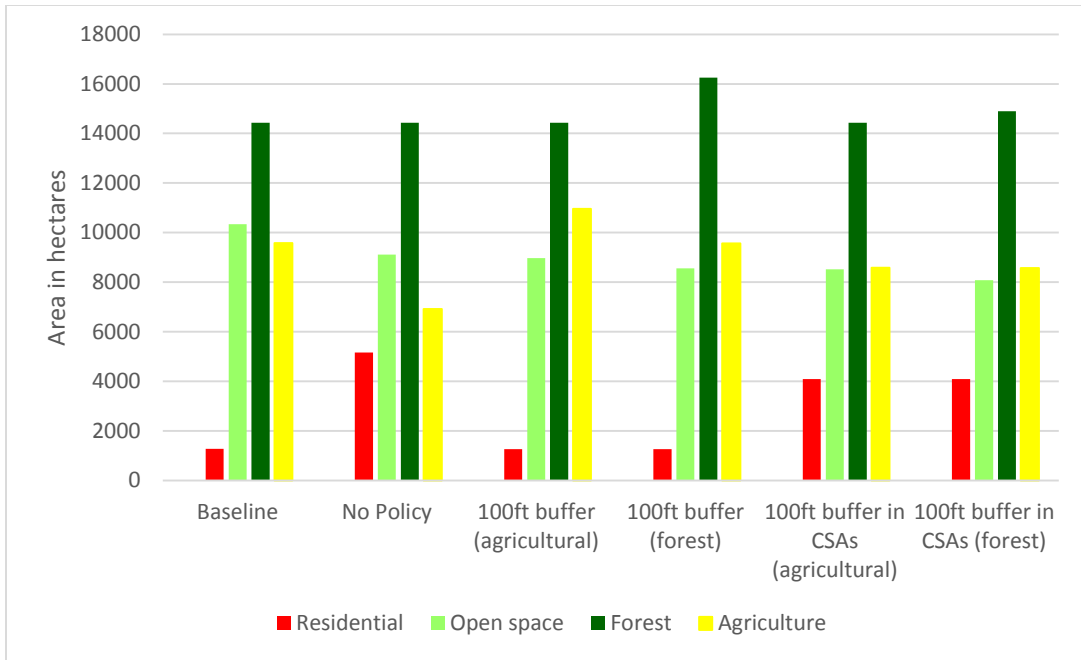


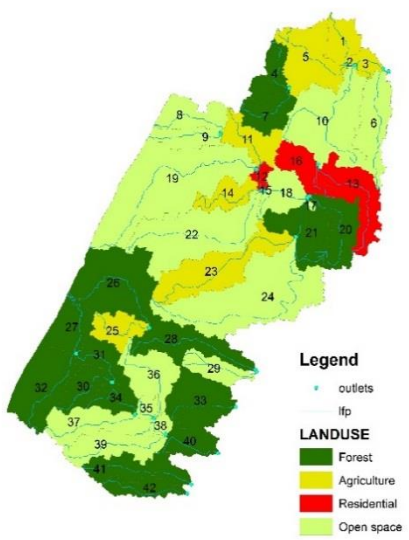
Figure 23. Land use distribution by area [ha] in each scenario

Table 13. Percentage of each land use type in watershed area under each policy scenario.

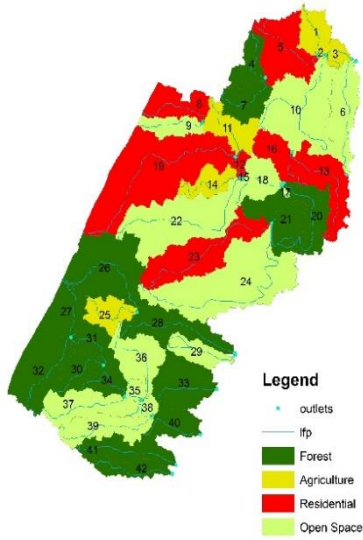
Scenario	Residential (%)	Open space (%)	Forest (%)	Agriculture (%)
Actual 2011	3.58	29.03	40.51	26.88
Baseline	14.50	25.59	40.51	19.40
50 ft buffer around all streams (agricultural)	3.55	25.18	40.51	30.76
50 ft buffer around all streams (forest)	3.55	24.02	45.60	26.83
50 ft buffer in CSA (agricultural)	11.49	23.91	40.51	24.09
50 ft buffer in CSA (forest)	11.49	22.67	41.79	24.05

Table 14. Average pollutant loads per month for the Opequon Creek watershed projected over a 10 year period.

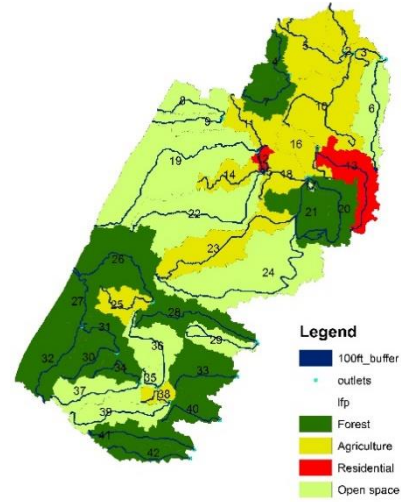
<b>Scenario</b>	<b>Sediment (tons/ha)</b>	<b>Total Nitrogen (kg/ha)</b>	<b>Total Phosphorous (kg/ha)</b>
Actual 2011	8.30	14.91	4.23
Baseline	7.88	14.91	4.01
50 ft buffer (agricultural) for all streams	8.60	15.39	4.50
50 ft buffer (forest) for all streams	8.28	13.86	4.06
50 ft buffer (agricultural) in CSAs	3.75	14.95	3.82
50 ft buffer (forest) in CSAs	3.75	14.86	3.80



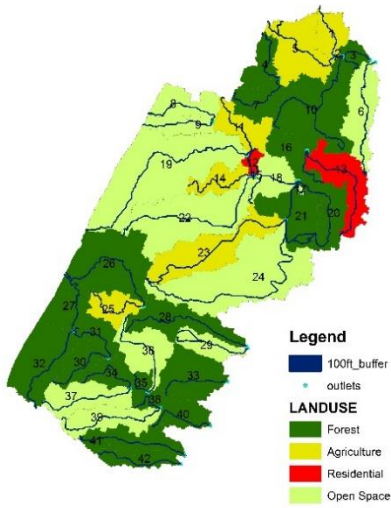
(a) 2011 land use



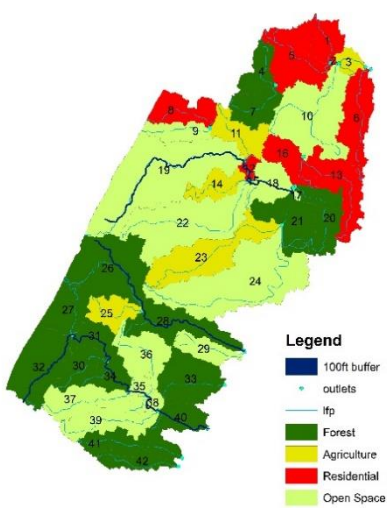
(b) baseline



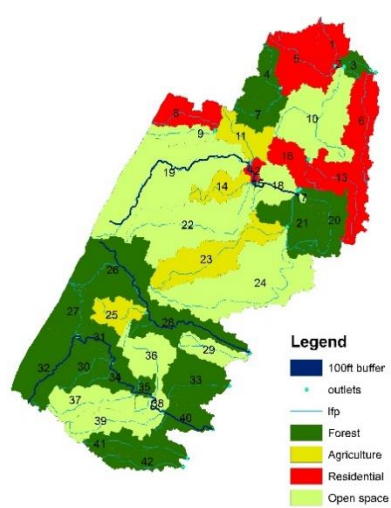
(c) 50 ft agricultural buffer (everywhere)



(d) 50 ft forest buffer (everywhere)



(e) 50 ft agricultural buffer in CSAs



(e) 50 ft forest buffer in CSAs

Figure 24. Dominant land use within each sub-basin delineated by ArcSWAT under each scenario

By computing Spearman Rank Correlation Coefficients, statistical relationships were assessed between residential land use conversion and pollutant loadings. Compared variables were: (1) the percentage of residential land use under the baseline scenario for each sub-basin, and (2) the baseline policy pollutant load by sub-basin minus the pollutant load for actual 2011 land use. Rank correlation coefficients were computed for the pollutants sediment, total nitrogen, and total phosphorus. The null hypothesis for each coefficient was:

$$H_0: \rho_S = 0$$

The alternative hypothesis was directional such that a positive correlation between the two variables was assumed:

$$H_1: \rho_S > 0$$

The rank correlation coefficients are reported in Table 15. The null hypotheses were rejected for the pollutants of sediment and TP, but not for TN. Total phosphorus in general has a relationship between biomass and suspended solid loads, therefore, reduction in sediment leads to a reduction in total phosphorus (Neitsch, 2005). It is evident from the Spearman Rank Correlation Coefficients that pollutant loads for sediment and TP are correlated with residential land use at a sub-basin watershed level. Therefore, decreasing of these pollutant loads are important considering residential land use policies to protect water quality.

Table 15. Spearman rank correlation analysis by pollutant type.

<b>Pollutant Type</b>	<b>Rank Coefficient (<math>\rho</math>)</b>	<b>Sig.(One-Tailed)</b>
Sediment	0.43	0.002**
Total Nitrogen (TN)	0.21	0.091
Total Phosphorus (TP)	0.47	0.000**

\*\* One tail test, P <0.01. Degree of freedom = 42

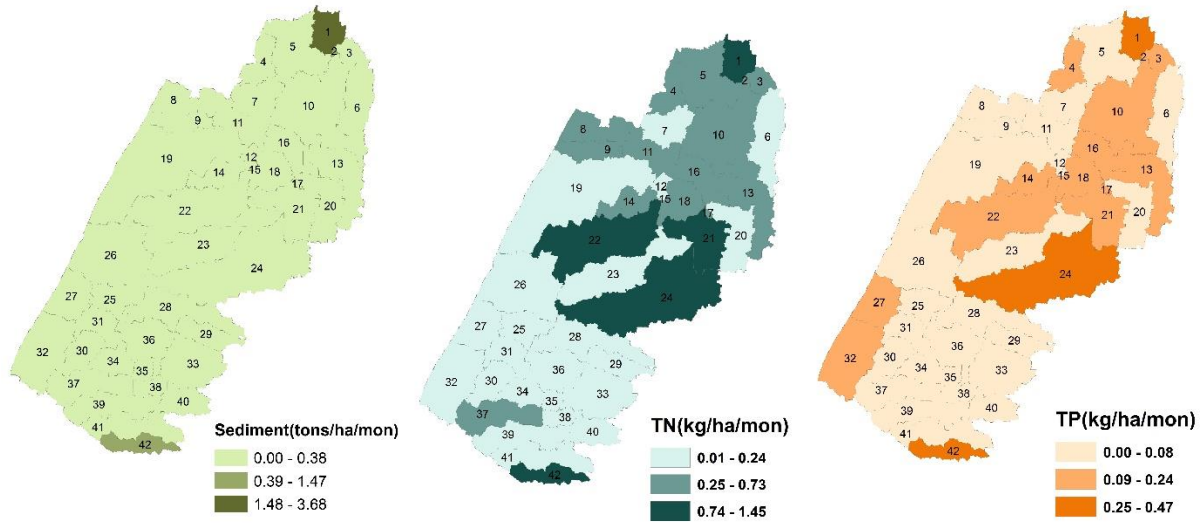


Figure 25. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with actual 2011 land use data

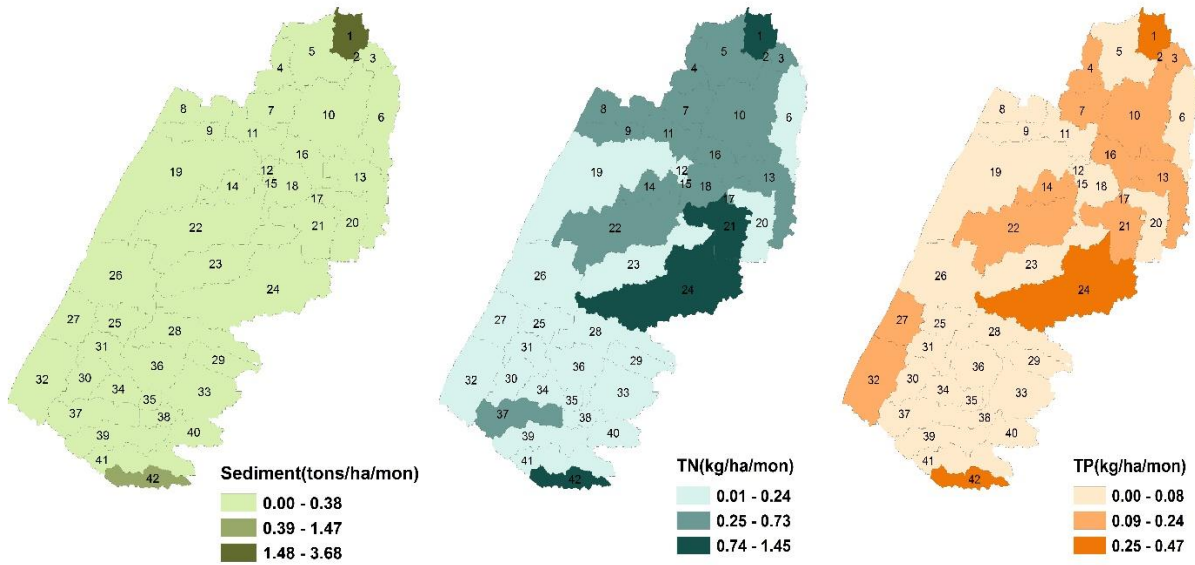


Figure 26. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with baseline policy scenario

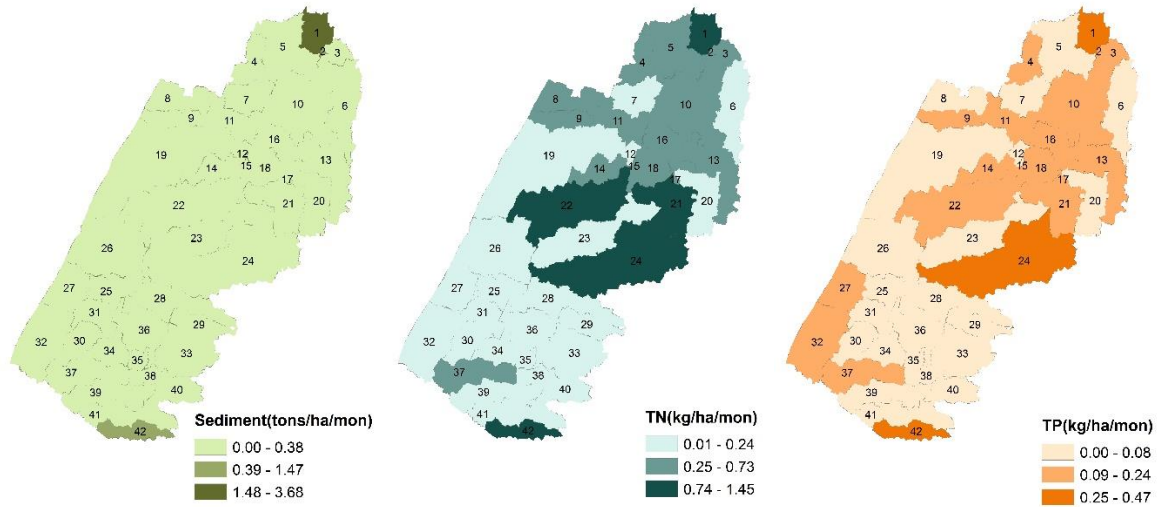


Figure 27. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT in 50 ft agricultural buffer everywhere

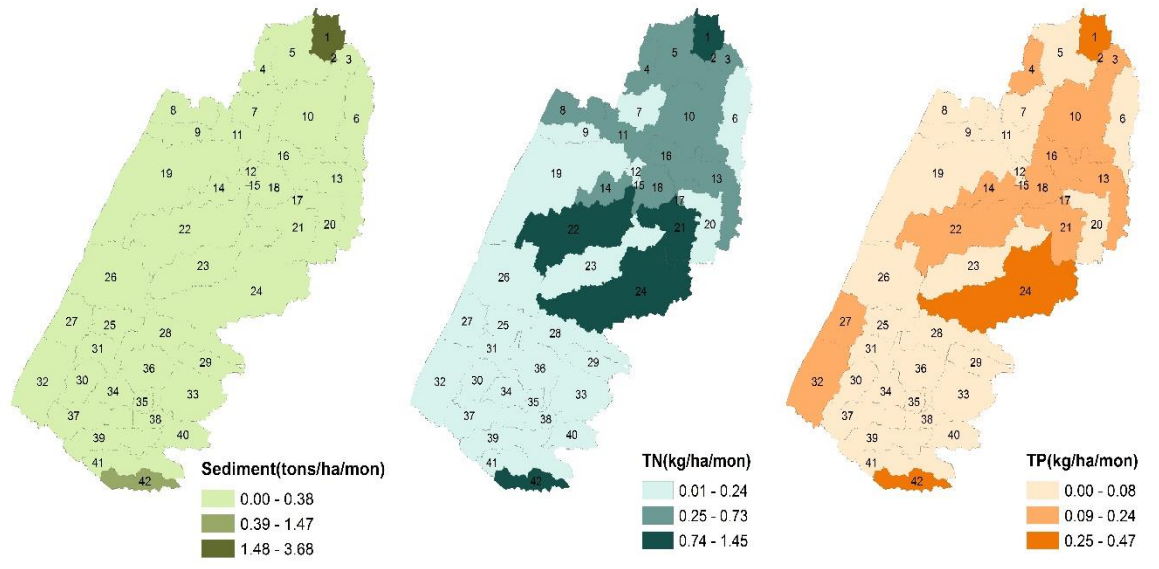


Figure 28. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT in 50 ft forest buffer everywhere

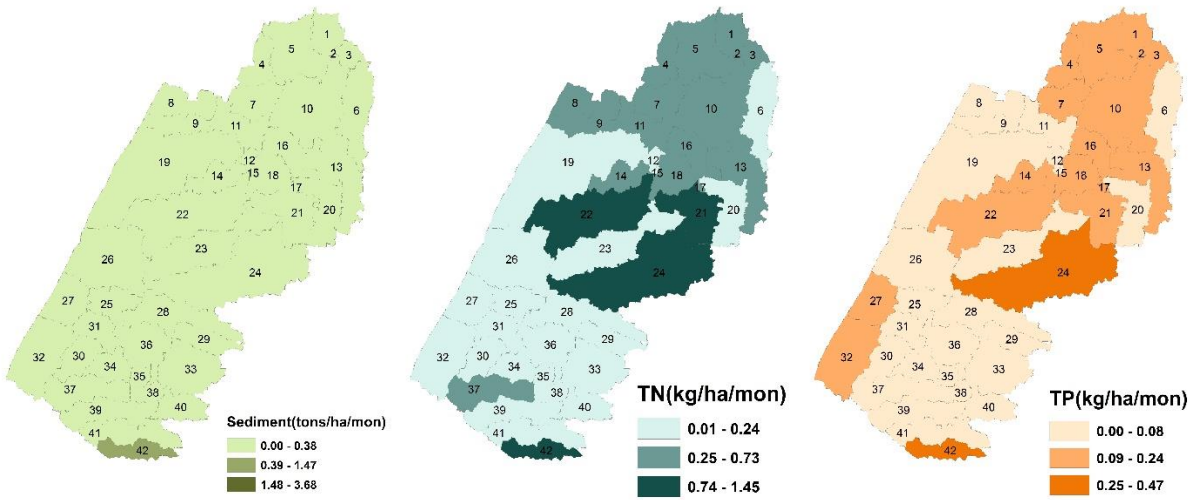


Figure 29. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with 50 ft agricultural buffer in CSAs

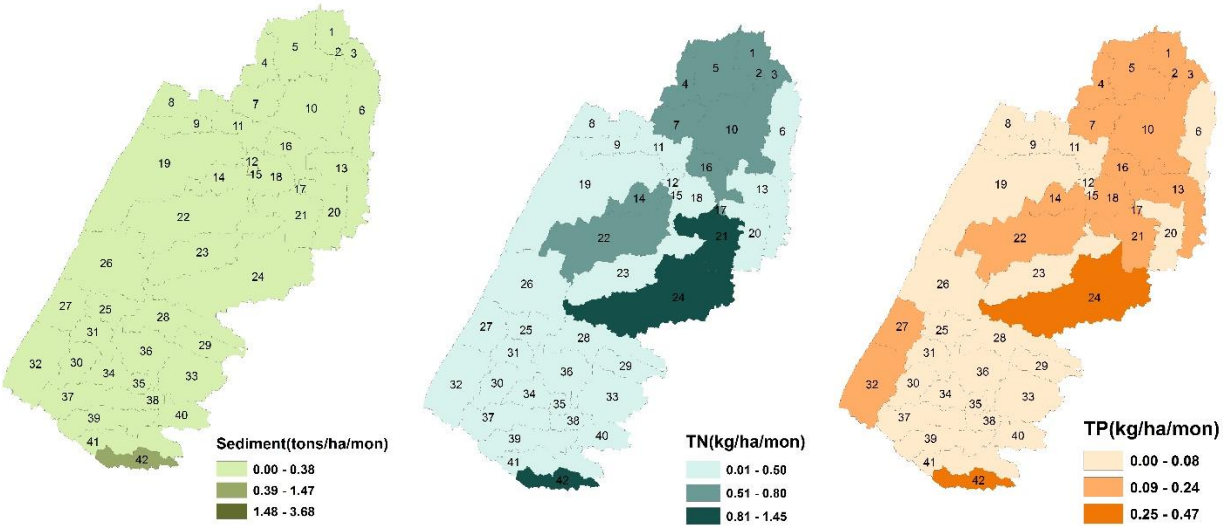


Figure 30. Sediment, TP and TN yields from each sub-basin as estimated by ArcSWAT with 50 ft forest buffer in CSAs

## **Water quality outcomes under the Baseline Policy**

Under the baseline policy scenario projections from the APLUC model, percentage land cover in the Opequon Creek watershed was estimated to be 14.50 % residential, 40.51% forest, 19.40% agricultural land, and 25.59% open space (Table 13). Average monthly sediment per hectare yield over 10 years by sub-basin ranged from a high 3.68 tons/ha/month to a low of 0.0009 tons/ha/month based on the ArcSWAT results (Figure 26). Compared to actual land use 2011, there is a 5.06% (Table 14) decrease in average sediment load per month per hectare under the baseline policy, however, there is a 75.31% increase in residential land use projected under this policy (Table 13). The average TN and TP yields for the entire watershed were found to be 14.91 kg/ha/month and 4.01 kg/ha/month respectively (Table 14). High proportion of TN and TP were found in sub-basins that are closer to sub-basins with residential or in agricultural land use as dominant land use (Figure 26). Most notable is sub-basin 1 where agricultural land is the dominant land use and there are high loadings of all pollutants.

## **Water quality outcomes with 50 foot buffers for all streams**

Land use patterns under this buffer policy showed 3.55% residential land use with either agricultural or forest buffers (Table 13). Under this policy, there was a 75.52% decrease in residential land use when compared to the baseline policy (Table 13). Despite this reduced residential land use, there were not substantial decreases in any of the three pollutant loads watershed wide or at the sub-basin level (Table 14 and Figure 27 and 28).

Compared to the baseline policy scenario there is a 3.22% increase in monthly average/ha TN loadings over entire watershed with agricultural buffers (Table 14). However with forest buffers monthly average/ha TN loadings declined by 7.04% compared to the baseline policy (Table



14). Since the Opequon Creek watershed is having relatively high percentage of agricultural land use (which is largely pasture), the ArcSWAT results indicate that this land use contributes largely to nitrogen loads. Thus, implementing a 50 foot agricultural buffer along all streams may not pinpoint locations that are sensitive to hydrological responses. With agricultural buffers, agricultural land increased in the watershed, consequently increasing monthly average/ha TN pollutant loads (Table 14). Forest riparian buffers are comparatively a slightly better policy than agricultural buffers due to a slight decrease in monthly average/ha TN loads.

### **Water quality outcomes with 50 foot buffers only in CSA sub-basins**

This policy showed much lower sediment and TP loadings in the Opequon Creek watershed than either of the other policies. Under this policy, there are 56.39% and 54.71% monthly average/ha decreases in sediment when compared to everywhere agricultural buffer and forest buffer, respectively (Table 14). The monthly average/ha sediment decrease compared to baseline is about 52.41% (Table 14). There is 15.11% monthly average/ha decrease in TP with 50 ft agricultural buffer in CSAs compare to 50 ft agricultural buffer everywhere and 6.40% monthly average/ha decrease in TP with 50 ft forest buffer in CSAs compare to 50 ft forest buffer everywhere (Table 14).

Specifically, loadings from sub-basin 1 were investigated further. This sub-basin is located in the northern part of the watershed near the junction of Opequon Creek with the Potomac River. In this sub-basin, the mean sediment loading was 0.03 tons/ha compared to 3.68 tons/ha for baseline or buffer everywhere policy (see Appendix III and Figures 26-30). This 3.68 tons/ha is the highest sediment load and estimated in sub-basin 1. In baseline and buffer everywhere policy, sub-basin 1 was dominant in agriculture (pasture) land and only has 60.235 hectares of residential land use. While with buffer policy in CSAs only, this sub-basin has 267.125 and 284.348 hectares

of residential land use with agricultural buffers and forest buffers, respectively. These increases in residential land made this land use type the dominant land use in sub-basin 1 (Figure 24). The result was a dramatic decrease in mean sediment loading. In addition, mean TP loadings declined along with monthly sediment loading in sub-basin 1 from 0.47 to 0.10 kg/ha (See Appendix III). This 0.47 kg/ha is the highest monthly mean loadings found in sub-basin 1 under baseline and 50 ft buffer (around all streams) policies (Figures 26-28).

The results suggest that location of residential land use is more important in the Opequon Creek watershed than the conservative spatial restrictive policy. Agriculture is causing higher pollutant loads in the watershed. Thus, drastically decreasing the residential land use would not solve pollutant loading problems. Instead implementing buffers in high priority watersheds would result in effective water quality outcomes. The buffers shift the location of conversion in sub-basins that are dominant in agricultural land use and reducing the sedimentation from agricultural sources. In this way residential land conversion in sub-basins that are generating agricultural related pollutant loads are improving water quality.

## Chapter 8

### **8. Conclusions, Limitations and Future works**

The overall modeling effort was based upon linking of three models: a spatial logistic regression model (SLR), an agent-based probabilistic land use conversion (APLUC) model, and an ArcSWAT (GIS extension of the Soil and Water Assessment Tool) model. Three land use policies were evaluated with these models: a baseline policy, a 50 foot buffer for every stream in entire watershed, and buffers only in critical source areas (CSAs). This study provides a detailed and comprehensive analysis of projected land use change for Opequon Creek watershed, based on prior observations of land use change. The SLR model provides information on how distance based explanatory variables impact land use conversions over a 10 year period (2001-2011). These explanatory variables were then used as parameters within the APLUC model to project land use conversions under each policy. Finally, the ArcSWAT model relates the spatial relationships between projected residential land use conversion and water quality in terms of sediment, total nitrogen (TN), and total phosphorus (TP) loadings. In this way, these models provide useful descriptive details of spatial policies' impact on land use conversion and water quality.

Based upon the research results, below is a summary of responses to the four questions presented in Chapter 1.

- (i) What spatial factors cause changes in spatial pattern of land use in the Opequon Creek Watershed during the period of a decade?
- (ii) Do empirically driven estimates explain the local level probabilities of residential land use conversions and if so, how effectively does it address the watershed level land use processes and patterns?

- (iii) How do the linkages of the three models respond to conditions reflected in the policies?
- (iv) How much does location matter for environmental policy such as water quality improvement?

**(i) What spatial factors cause changes in spatial pattern of land use?**

To demonstrate some important spatial factors that influence residential land use conversion, a SLR model was run for the period between 2001 and 2011 using land cover data in the Opequon Creek watershed. Coefficient values for distances to urban centers and neighboring land uses of forest, agriculture and residential land were found to have statistically significant influences on residential land use conversion. The distance to urban center variable represented the Von Thünen monocentric Alonso and bid-rent theory. The variable representing distances to neighboring land uses were explained by the Hedonic theory of land valuation.

The positive coefficient signs for distances to agricultural and forest land are important to note as they mean that close proximity to these land uses in suburban areas create a negative spatial externalities for residential land use conversion. These findings illustrate a strong agglomeration effect between residential land conversion and distant amenities. The distance to existing residential land use had a negative coefficient and had a positive marginal effect on residential land use conversion. Overall, the SLR model provided empirically valid parameters to project land use change probabilities within the APLUC model.

**(ii) Do empirically driven estimates explain the local level probabilities and address the watershed level land use processes and patterns?**

The theoretical basis of the empirical model established the linkage between spatial logistic model and bid-rent, Von Thünen and Hedonic models. The calibration results from the spatial logistic

regression affirm the hedonic and bid-rent model. However, APLUC model contributes to the understanding of the land use change processes at decision making units (property parcel) by accounting stochastic component using empirically observed spatial externalities. The pattern of residential development results from stochastic observations over 50 model runs for 10 iterations each in each policy scenario. The probabilities are based upon local interactions of parcels through their proximity influence, and account for these influences in iterative manner (10 iterations). This iterative functionality allows the model to adapt to changing local conditions, based upon estimated parameters from the SLR model. In this manner, the APLUC model incorporates uncertainty in land use conversion projections due to stochastic elements within the model. However, empirically estimated parameters which do not change over iterations in the APLUC model make this probabilistic model of land use conversion behave in a somewhat deterministic manner.

The use of micro level parcel based data on land use conversion (as opposed to percentages of land use types) offers a spatially explicit view of the impacts from residential land use conversions on water quality improvements. The APLUC model was validated at sub-basin watershed level. Spearman Rank Correlation Coefficients were calculated to compare pollutant loadings between actual land use in 2011 and projected land use in 2011 using data on monthly average loadings for sediment, TN, and TP over 10 years by sub-basins. The APLUC model was validated with statistically significant ranking coefficients for each pollutant loading.

Residential land use conversion patterns were taken from the APLUC model results and put into the ArcSWAT model at a sub-basin level. This link of APLUC with ArcSWAT allowed ArcSWAT to simulate the impact of residential conversion decisions on water quality. The conversion decisions generated residential land use patterns that were assessed at a sub-basin level.

The pollutant load results from these linked models indicated that effective policies discouraged residential land use in CSAs and moved conversion to agricultural dominated sub-basins.

**(iii) How do the linkages of the three models respond to conditions reflected in the policies?**

The APLUC model results showed that each buffer policy has different impacts in terms of location and residential land use conversions. Under the baseline policy, the urban center was the driving factor of residential land use conversion. Thus, most of the projected residential conversion occurred close to the urban center under this policy scenario. For residential land use conversions under the 50 ft buffer policy on all streams, much lower residential conversion took place than under the baseline policy. The resulting residential patterns showed a sparse, small conversions to residential land due to buffer zones limiting larger parcel conversions. The buffers in CSAs policy resulted in a slight decrease in the quantity of residential land use conversions and shifted residential development to sub-basins located in the northern part of the Opequon Creek watershed.

This study addresses water quality changes within the Opequon Creek watershed as the result of changing land use patterns under different buffer policies. By linking the three models, observed residential land use conversion decisions were connected to projected conversions, which were then tied to water quality changes. By estimating individual sub-basin level loadings, spatial variation in water quality impacts was observed. The use of Spearman Rank Correlation Coefficients provided quantitative evidence that residential land use conversion was linked with sediment and TP loadings across the sub-basins.

The ArcSWAT results showed that location is crucial in small scale watershed policies. The impact of restrictive buffer zones in CSAs was not effective in decreasing the quantity of

residential land use conversion, but was effective in terms of location of residential conversions and in reducing pollutant loadings for sediment and TP. When restrictive buffers were placed on all streams, pollutant loadings did not decrease (except for a small reduction in TN with forest buffers). These water quality result occurred despite large decreases in residential conversions projected under this buffer policy. Thus, buffer restrictions in CSAs were more effective in reducing pollutant loads compared to a policy with buffers on all streams.

One reason for such a weak hydrological response of buffers on all streams is due to the fact that most of the land use in the Opequon Creek watershed was pasture-based agriculture or open space. The results show that when buffers were agricultural land, this expanded agricultural land use as the dominant land use in sub-basin watersheds. The ArcSWAT model showed the highest TN loadings under this buffer policy. Forest buffers were relatively more effective in achieving water quality improvements than agricultural buffers.

**(iv) How much does location matter for environmental policy such as water quality improvement?**

This research shows that an efficient allocation (in terms of reduced pollutant loadings) of residential land use conversion in the Opequon creek watershed results in a movement of conversion to the northern part of the watershed, especially in sub-basin one. Sub-basin one showed a high level of residential land use conversion from what was formerly an agricultural dominated sub-basin under the 50 foot buffer in CSAs policy. The results were a dramatic reduction in sediment loading and a small reduction in TP loading. Therefore, sub-basin level considerations in policy is informative for water quality improvement.

A spatially explicit APLUC framework to empirically model land use change provides knowledge about impacts of land use conversions and where the impacts occurred. Using this

spatially explicit data, residential land use conversions at a parcel level were related to sub-basin level dominant land use in the ArcSWAT model. This linkage of residential land use conversions with water quality modeling measured outcomes in pollutant loads. These results demonstrated that scale dependent APLUC model with disaggregated units of observation can be effective in understanding how changes in one part of the watershed affect the water quality outcomes in surrounding sub-basins. Understanding these linkages is the key to showing how effective land use policies can be for water quality improvement. Thus, the focus of land use policies should be on location of residential land use conversions instead of watershed wide land use conversion restrictions.

### **Methodological Issues and Future Development**

There are numerous dimensions in which the model adopted in this study can be further improved. The present modeling effort is addressing a small scale watershed to forecast short term projections. The calibration results should not be extended for long term projections or for larger scale watershed, since the underlying driving factors for land use can change over longer times and over broader scales. Having primary as well as secondary source data sources such as community surveys, historical data, and parcel based socio-economic data with more frequent calibrations can be employed to make the projections more reliable.

This study calibrated a land use conversion model for only one watershed. It is uncertain to what extent the results can be applied to other study areas and whether the results are showing general trends or trends specific to the Opequon Creek area. Since Opequon Creek is located in the Chesapeake system, similar modeling exercises for other watersheds in the Chesapeake Bay are suggested. Second, what are the value of spatial externalities if other modeling structures are applied for calibration such as Bayesian probability instead of logit transformed probability? Does



information on perceptions of property owners change the projections? Also in terms of the choice of explanatory variables, the current modeling structure has been limited by the lack of information on markets and other equally important factors. Inclusion of information on dynamic markets, socioeconomic and demographic factors of land owner agents would improve the validation and model structure.

In terms of water quality, this research included only three pollutants and examined homogenous sized buffers to represent the potential water quality effects under buffer policies. It would be interesting to simulate the APLUC data for other pollutants under varying sizes of buffers. It is also important to include sensitive water bodies, downstream, and water heads to link within the whole watershed for understanding of hydrological pathways. This information is not only critical for evaluation of water quality but also to determining buffer strip effectiveness.

Additionally, the land use change prediction can be assessed by proximate causes of land use change, however the projections of future land use change requires the understanding of the forces that may be difficult to observe. These forces often involve cultural values, market conditions, and policy changes which are difficult to forecast. Also, as with any model forecast, there is a degree of uncertainty associated with a model to anticipate changes between 2001 and 2011.

To improve the modeling framework, future work should include parcel based socioeconomic information in assessing land use conversions. Further, broader land use classes for conversions and different initial conditions would be implemented to understand the interspersion of different land use types and resulting water quality indicators. The study would also consider other land use specific practices such as sustainable management and best management practices in conjunction with riparian management. The APLUC model would be integrated with other

ecological models such as impact on biodiversity and ecosystem services to provide multiple assessments of the impact of land use change. This would make the APLUC model a valuable policy tool to anticipate future land use systems. Current and further extension of this research would help in answering questions like, would the present policies change if the future consequences of these policies are anticipated? Can policy makers evaluate the urban growth control policies by spatially explicit ABM modeling? How can effective decisions be made knowing predicted land use patterns?

## References

- Alberti, M., & Waddell, P. (2000). An integrated urban development and ecological simulation model. *Integrated Assessment*, 1(3), 215-227.
- Alig, Ralph J. (2010). Economic modeling of effects of climate change on forest sector and mitigation options: a compendium of briefing papers. USDA
- Alig, R. J., & Healy, R. G. (1987). Urban and built-up land area changes in the United States: An empirical investigation of determinants. *Land Economics*, 63(3), 215-226.
- Allan, D. J. (2004). Landscapes and riverscapes: The influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution and Systematics*, 35, 257-284.
- Alonso, W. (1964). *Location and land use*. Cambridge, MA. Harvard University Press.
- Almeida, C.M., Batty, M., Monteiro, A.M.V., Câmara, G., Soares-Filho, B. S., Cerqueira, G. C., Pennachin, C. L. (2003). Stochastic cellular automata modeling of urban land use dynamics: Empirical development and estimation. *Computers, Environment and Urban Systems*, 27(5), 481-509.
- Almeida, C.M., Monteiro, A.M.V., Câmara, G., Soares-Filho, B. S., Cerqueira, G. C., Pennachin, C. L, Batty, M. (2005). GIS and Remote Sensing as Tools for the Simulation of Urban Land Use Change. *International Journal of Remote Sensing*. 26(4):759-774.

- Anbumozhi, V., Radhakrishnan, J., & Yamaji, E. (2005). Impact of riparian buffer zones on water quality and associated management considerations. *Ecological Engineering*, 24(5), 517-523.
- Anderies, J. M., Janssen, M. A., & Ostrom, E. (2004). A framework to analyze the robustness of social-ecological systems from an institutional perspective. *Ecology and Society*, 9(1), 18.
- Anderson, R. M., Koren, V. I., & Reed, S. M. (2006). Using SSURGO data to improve Sacramento model a priori parameter estimates. *Journal of Hydrology*, 320(1-2), 103-116.
- Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large-area hydrologic modeling and assessment: Part I. model development. *Journal of American Water Resources Association*, 34(1), 73-89.
- Arsanjani, J.J., Helbich, M., Kainz, W., & Darvishi, A, B. (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion - the case of Tehran. *International Journal of Applied Earth Observation and Geoinformation*, 21(0), 265-275.
- Balzter, H., Braun, P. W., & Köhler, W. (1998). Cellular automata models for vegetation dynamics. *Ecological Modelling*, 107(2-3), 113-125
- Bastian, C. T., McLeod, D. M., Germino, M. J., Reiners, W. A., & Blasko, B. J. (2002). Environmental amenities and agricultural land values: A hedonic model using geographic information systems data. *Ecological Economics*, 40(3), 337-349.
- Batty, M. (2012). A generic framework for computational spatial modeling. In A. J. Heppenstall, A. T. Crooks, L. M. See & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 19-50) Springer, New York, NY.

Baumgart, P. D. (2005). *Lower green bay and lower fox tributary modeling report*. Prepared for Oneida Tribe of Indians of Wisconsin and Green Bay Remedial Action Plan Science and Technical Advisory Committee. Retrieved from

[http://www.uwgb.edu/watershed/reports/related\\_reports/load-allocation/lowerfox\\_tss-p\\_load-allocation.pdf](http://www.uwgb.edu/watershed/reports/related_reports/load-allocation/lowerfox_tss-p_load-allocation.pdf)

Bayoh, I, Irwin, E. G., & Haab, T. (2006). Determinants of residential location choice: How important are local public goods in attracting homeowners to central city locations? *Journal of Regional Science*, 46, No. 1, 2006, pp. 97–120

Bell, K. P., & Irwin, E. G. (2002). Spatially explicit micro-level modelling of land use change at the rural–urban interface. *Agricultural Economics*, 27(3), 217-232.

Benenson I, & Torrens PM. (2004). *Geosimulation: Automatabased modeling of urban phenomena*. West Sussex: Wiley.

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

Berkeley County Assessor's office (2014). Retrieved from <http://www.theassessor.org/>

Berkeley County Development Authority. (2014)

Retrieved from <http://www.developmentauthority.com/>

Berkeley County Planning Commission. (2009). *Subdivision ordinance (2009)*. Draft ordinance with county commission approved changes. Retrieved from [http://www.berkeleycountycomm.org/docs/draft\\_subreg0409.pdf](http://www.berkeleycountycomm.org/docs/draft_subreg0409.pdf)

Berkeley County Planning Commission. (2006). Berkeley County, comprehensive plan update. Retrieved from <http://www.berkeleycountycomm.org/docs/2006BCCCompPlan.pdf>

Berkeley County, W. V. (2004). *Source water assessment & protection (SWAP) project*. Retrieved from <http://www.berkeleywater.org/pdfs/SWAP.pdf>

Bevin, K. J., & M. J. Kirkby. (1979). A physically based variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24, 43-69.

Bhaduri, B., Minner, M., & Tatalovich, S. H., J. (2001). Long-term hydrologic impact of urbanization: A tale of two models *Journal of Water Resources Planning and Management*, 127, 13-19.

Bithell, M., & Brasington, J. (2009). Coupling agent-based models of subsistence farming with individual-based forest models and dynamic models of water distribution. *Environmental Modelling & Software*, 24(2), 173-190.

Bockstael, N. E., & K.P. Bell. (1998). Land use patterns and water quality: The effect of differential land management controls. In R. Just, & S. Netanyahu (Eds.), *International water*

*and resource economics consortium, conflict and cooperation on trans-boundary water resources* (pp. 169-191) Kluwer Academic Publishers, Norwell, MA.

Bockstael, N. E. (1996). Modeling economics and ecology: The importance of a spatial perspective. *American Journal of Agricultural Economics*, 78(5), 1168-1180.

Braimoh, A. K. (2004). Modeling land-use change in the Volta basin of Ghana. *Ecology and development series no 14* (pp. 175) Cuvillier Publishers, Gottingen, Germany.

Breuer, L., Huisman, J. A., & Frede, H. G. (2006). Monte Carlo assessment of uncertainty in the simulated hydrological response to land use change. *Environmental Modeling & Assessment*, 11(3), 209-218.

Broadmeadow, S., & Nisbet, T. R. (2004). The effects of riparian forest management on the freshwater environment: A literature review of best management practice. *Hydrology and Earth System Sciences*, 8(3), 286-305.

Brown, D. G., & Robinson, D. T. (2006). Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl *Ecology and Society*, 11(1), 46.

BBER. (2014). *Population trends in West Virginia through 2030*. Morgantown, WV: Bureau of Business and Economic Research, College of Business and Economics, West Virginia University.

Buurman, J., Rietveld, P., & Scholten, H. (2001). The land market: A spatial economic perspective Stillwell, J & Scholten, H. (Eds.), *Land use simulation for Europe*, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 65–83

- Carlson, T. N. (2004). Analysis and prediction of surface runoff in an urbanizing watershed using satellite imagery1. *JAWRA Journal of the American Water Resources Association*, 40(4), 1087-1098.
- Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, 8(3), 559-568.
- Carrión-Flores, C., & Irwin, E. G. (2004). Determinants of residential land-use conversion and sprawl at the rural-urban fringe. *American Journal of Agricultural Economics*, 86(4), 889-904.
- Christiadi. (2011). *2000-2010 demographic changes in West Virginia*. (No. CENSUS 2010 BRIEF). Morgantown, WV: Bureau of Business and Economic Research, College of Business and Economics, West Virginia University.
- Clark Labs. (2012). IDRISI Selva, Clark University  
<http://www.clarklabs.org/>
- Clarke, H., & Granato, J. (2005). Autocorrelation. In K. Kampf-Leonard (Ed.), *Encyclopedia of social measurement* () San Diego: Academic Press.
- 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.



- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple Regression/Correlation analysis for the behavioral sciences*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Cohen, M. J., Shepherd, K. D., & Walsh, M. G. (2005). Empirical reformulation of the universal soil loss equation for erosion risk assessment in a tropical watershed. *Geoderma*, 124(3-4), 235-252.
- Corbett, C. W., Wahl, M., Porter, D. E., Edwards, D., & Moise, C. (1997). Nonpoint source runoff modeling. A comparison of a forested watershed and an urban watershed on the south carolina coast. *Journal of Experimental Marine Biology and Ecology*, 213(1), 133-149.
- Coutu, G. W., & Vega, C. (2007). Impacts of land use changes on runoff generation in the east branch of the Brandywine creek watershed using a GIS-based hydrologic model. *Middle States Geographer*, 40, 142-149.
- Dauer, D. M., Weisberg, S. B., & Ranasinghe, J. A. (2000). Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. *Estuaries*, 23(1), 80-96.
- Dean, J. S., Gummerman, G. J., Epstein, J. M., Axtell, R. L., Swedlund, A. C., Parker, M. T., & McCarroll, S. (2000). Understanding Anasazi culture change through agent-based modeling. In T. A. Kohler, & G. J. Gummerman (Eds.), *In dynamics in human and primate societies* (pp. 179-205). Oxford, UK: New York: Oxford University Press.

Department of Transportation (DOT) Maryland Transit Administration. (2014)

Retrieved from: <http://www.mdot.maryland.gov/>

Dietzel, C., & Clarke, K. (2004). Replication of spatio-temporal land use patterns at three levels of aggregation by an urban cellular automata. *Computer Science*, 3304 (2004), pp. 523–532

Dosskey, M. G., Vidon, P., Gurwick, N. P., Allan, C. J., Duval, T. P., & Lowrance, R. (2010). The role of riparian vegetation in protecting and improving chemical water quality in Streams. *JAWRA Journal of the American Water Resources Association*, 46(2), 261-277.

Duan, S., Kaushal, S. S., Groffman, P. M., Band, L. E., & Belt, K. T. (2012). Phosphorus export across an urban to rural gradient in the Chesapeake Bay watershed. *Journal of Geophysical Research: Biogeosciences*, 117(G1), - G01025.

Epstein, J. M. (2006). *Generative social science: Studies in agent-based computational modeling*. Princeton, NJ: Princeton Univ. Press

ESRI 2011. (2011). *ArcGIS desktop: Release 10*. Redlands, CA: Environmental Systems Research Institute.

ESRI. (2014). ArcGIS Help 10.2. Retrieved from:

<http://resources.arcgis.com/en/help/main/10.2/index.html>

Feng, Y., & Liu, Y. (2013). A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing. *International Journal of Geographical Information Science*, 27(3), 449-466.

- Fleming, K. (1989). Farming in the Shadow of the City. *Farm Management*. (1989) Yearbook of Agriculture.
- Fragkias, M., Seto, K.C. (2007). Modeling urban growth in data-sparse environments: A new approach. *Planning and Design*, 34(5), 858-883.
- Geoghegan, J., Wainger, L & Bockstael, N. (1997). Spatial landscape indices in a Hedonic framework: An ecological economics analysis using GIS. *Ecological Economics* 23: 251–264.
- Geoghegan, J., Villar, S. C., Klepeis, P., Mendoza, P. M., Ogneva-Himmelberger, Y., Chowdhury, R. R., & Vance, C. (2001). Modeling tropical deforestation in the Southern Yucatán peninsular region: Comparing survey and satellite data. *Agriculture, Ecosystems and Environment*, 85, 25-46.
- Gorsevski P.V., Gessler P.E., Foltz R.B., & Elliot W.J. (2006) Spatial prediction of landslide hazard using logistic regression and ROC analysis. *Transactions in GIS* 10(3):395–415
- Gergel, S. E., Turner, M. G., Miller, J. R., Melack, J. M., & Stanelly, E. H. (2002). Landscape indicators of human impacts to riverine systems. *Aquatic Sciences*, 64(2), 118-128.
- Getachew, H. E., & Melesse, A. M. (2013). The impact of land use change on the hydrology of the Angereb Watershed, Ethiopia. *International Journal of Water Sciences*, 1(4), 1-7.
- Gimblett, H., Roberts C., Daniel T., Ratliff M., Meitner M., Cherry S., Stallman D., Bogle R., Allred R., Kilbourne, D., & Bieri, J. (2002): "An intelligent agent-based model for simulating and evaluating river trip scenarios along the Colorado River in Grand Canyon National Park".

In: Gimblett, H (ed.), *Integrating geographic information systems and agent-based modelling techniques for simulating social and ecological processes*. Oxford University Press, pp. 245-275.

Gimblett, R. H. (2002). *Integrating geographic information systems and agent-based modeling techniques for stimulating social and ecological processes*. USA: Oxford University Press.

Glasserman, P. (2003). Monte Carlo methods in financial engineering. *Stochastic modelling and applied probability*. 53, 596. Springer.

Goetz, S.J., Smith, A.J., Jantz, C., Wright, R.K., Prince, S.D., Mazzacato, M.E., & Melchior, B. (2003). Monitoring and predicting urban land use change - applications of multi-resolution multi-temporal satellite data. *IGARSS 2003: IEEE International Geoscience and Remote Sensing Symposium, I-Vii*, 1567-1569.

Goetz, S.J., Wright, R.K., Smith, A.J., Zinecker, E. & Schaub, E. (2003). IKONOS imagery for resource management: tree cover, impervious surfaces, and riparian buffer analyses in the mid-Atlantic region. *Remote Sensing of Environment*, 88, pp. 195–208.

Goodspeed, R. (2007). Leapfrog' sprawl in West Virginia.

Retrieved from <http://goodspeedupdate.com/2007/2104>

Gorsevski, P. V., Gessler, P. E., Foltz, R. B., & Elliot, W. J. (2006). Spatial prediction of landslide hazard using logistic regression and ROC analysis. *Transactions in GIS*, 10(3), 395-415.

- Groffman PM, & Crawford MK. (2003). Denitrification potential in urban riparian zones. *Journal of Environmental Quality*, 32(3), 1144-1149.
- Guzy, M. R., Smith, C. L., Bolte, J. P., Hulse, D. W., & Stanley, V. G. (2008). Policy research using agent-based modeling to assess future impacts of urban expansion into farmlands and forests. *Ecology and Society*, 13(1), 37.
- Hagerstrand, T., (1965). A Monte Carlo approach to diffusion. *Archive of European Sociology*, VI, 43-67.
- Hatten, M., Lapp, J., Bennett, D., & Stottlemeyer, D. (2011). WV stream and wetland valuation (SWVM) metric development. *Appalachian Stream Mitigation Workshop*, Lexington, KY.
- Heckbert, S., Baynes, T., & Reeson, A. (2010). Agent-based modelling in ecological economics. *Annals of the New York Academy of Sciences*, 1185, 39-53.
- Heppenstall, A. J., & Crooks, A. T. (2012). In Batty M., See L. M. (Eds.), *Agent-based models of geographical systems* Springer.
- Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., and Wickham, J. (2007). Completion of the 2001 national land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 73(4), 337-341.
- Hosmer, D. W., & Lemeshow S. (1989). *Applied logistic regression* New York: John Wiley & Sons, Inc.

- Huang, J.L., Li, Q.S., Glimore, R., Pontius Jr., Klemas, V., & Hong, H. (2013). Detecting the dynamic linkage between landscape characteristics and water quality in a subtropical coastal watershed, southeast china. *Environmental Management*, 51, 32-44.
- Hunsaker, C. T., & Levine, D. A. (1995). Hierarchical approaches to the study of water quality in rivers. *Bioscience*, 45(3, Ecology of Large Rivers), 193-203.
- IDRE (Institute for Digital Research and Education). (2014). Retrieved from [http://www.ats.ucla.edu/stat/mult\\_pkg/faq/general/odds\\_ratio.htm](http://www.ats.ucla.edu/stat/mult_pkg/faq/general/odds_ratio.htm)
- Irwin, E. G. (2002). The effects of open space on residential property values. *Land Economics*, 78(4), 465-480.
- Irwin, E. G. (2010). New directions for urban economic models of land use change: Incorporating spatial dynamics and heterogeneity. *Journal of Regional Science*, 50(1), 65-91.
- Irwin, E. G., & Bockstael, N. E. (2002). Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography*, 2(1), 31-54.
- 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 & Environment*, 85(1-3), 7-24.

- Irwin, E. G., Wilner Jeanty, P., & Partridge, M. D. (2014). Amenity values versus land constraints: The spatial effects of natural landscape features on housing values. *Land Economics*, 90, 61-78.
- Jantz, P., Goetz, S., & Jantz, C. (2005). Urbanization and the loss of resource lands in the Chesapeake Bay watershed. *Environmental Management*, 36(6), 808-825.
- Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the national land cover database to circa 2011. *Remote Sensing of Environment*, 132(0), 159-175.
- Johnson, L., Richards, C., Host, G., & Arthur, J. (1997). Landscape influences on water chemistry in Midwestern stream ecosystems. *Freshwater Biology*, 37(1), 193-208.
- Johnston, K. M. (2013). *Agent analyst: Agent-based modeling in ArcGIS*. Redlands, California: ESRI Press.
- Jontos, R. (2004). Vegetative buffers for water quality protection: An introduction and guidance document. *Connecticut Association of Wetland Scientists White Paper on Vegetative Buffers., Draft version 1.0. 22pp.*
- Jung, K. W., Lee, S. W., Hwang, H. S., & Jang, J. H. (2008). The effects of spatial variability of land use on stream water quality in a costal watershed. . *Paddy Water Environ*, 6, 275-284.
- Karigomba, W. (2009). *A spatial optimization approach to watershed water quality management: A case of the Opequon watershed*. (Unpublished PhD). West Virginia University,

- Kaushal, S. S., Groffman, P. M., Band, L. E., Elliott, E. M., Shields, C. A., & Kendall, C. (2011). Tracking nonpoint source nitrogen pollution in human-impacted watersheds. *Environmental Science & Technology*, *45*, 8225-8232.
- Kitamura, R., Mokhtarian, P. L., & Laidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco bay area. *Transportation*, *24*, 125-158.
- Koontz, T. M. (2001). Money Talks—But to whom? Financial versus nonmonetary motivations in land use decisions. *Society and Natural Resources*, *14*, 51-65.
- Lambin, E. F., Geist, H. J., & Lepers, E. (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, *28*, 205-241.
- Lambin, E. F., & Meyfroidt, P. (2010). Land use transitions: Socio-ecological feedback versus socio-economic change. *Land use Policy*, *27*(2), 108-118.
- Le, Q. B., Park, S. J., & Vlek, P. L. G. (2010). Land use dynamic simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system: 2. scenario-based application for impact assessment of land-use policies. *Ecological Informatics*, *5*(3), 203-221.
- Le, Q. B., Park, S. J., Vlek, P. L. G., & Cremers, A. B. (2008). Land-use dynamic simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system. I. structure and theoretical specification. *Ecological Informatics*, *3*(2), 135-153



- Lee, S. (2005). Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing*, 26(7), 1477-1491.
- Lee, S., Hwang, S., Lee, S., Hwang, H., & Sung, H. (2009). Landscape ecological approach to the relationships of land use patterns in watersheds to water quality characteristics. *Landscape and Urban Planning*, 92(2), 80-89.
- Li, Y. L., Liu, K., Li, L., & Xu, Z. X. (2012). Relationship of land use/cover on water quality in the Liao river basin, china. *Procedia Environmental Sciences*, 13(0), 1484-1493.
- Liu Y., Kong X., Liu Y., & Chen Y. (2013). Simulating the conversion of rural settlements to town land based on multi-agent systems and cellular automata. *Plos One*, 8(11), e79300.
- Li, H., Reynolds, J.F. (1997). Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: a combined Markov and cellular automaton approach. In: Quattrochi, D.A., Goodchild, M.F. (Eds.), *Scale in Remote Sensing and GIS*. Lewis Publishers, Boca Raton, Florida, pp. 211–230
- Lo, C. P., & Watson, L. J. (1998). The influence of geographic sampling methods on vegetation map accuracy evaluation in a swampy environment. *Photogrammetric Engineering & Remote Sensing*, 64(12), 1189-1200.
- Ma, T., & Nakamori, Y. (2009). Modeling technological change in energy systems – from optimization to agent-based modeling. *Energy*, 34(7), 873-879.

- Madison, D. (2000). A hedonic analysis of agricultural land prices in England and Wales. University College London and University of East Anglia. *European Review of Agricultural Economics*, 27: 510-532.
- Magliocca, N., Safirova, E., McConnell, V., & Walls, M. (2011). An economic agent-based model of coupled housing and land markets (CHALMS). *Computers, Environment and Urban Systems*, 35(3), 183-191.
- Manson, S. M. (2005). Agent-based modeling and genetic programming for modeling land change in the Southern Yucatán' n Peninsular Region of Mexico. *Agriculture, Ecosystems and Environment*, 111, 47-62.
- Manson, S. M. (2001). Simplifying complexity: A review of complexity theory. *Geoforum*, 32(3), 405-414.
- Matthews, R. B., Gilbert, N., Roach, A., Polhill, G., & Gotts, N. M. (2007). Agent-based land-use models: A review of applications. *Landscape Ecology*, 22(10), 1447-1459.
- McLeish, D. L. (2005). *Monte Carlo simulation and finance*. Wiley.
- Mednick A.C., Sullivan J., & Watermolen D.J. (2008). Comparing the use of STATSGO and SSURGO soils data in water quality modeling: A literature review. *Research/Management Findings*, (60)
- Mejia, A., Daly, E., Rossel, F., Jovanovic, T., & Gironas, J. (2014). A stochastic model of stream flow for urbanized basins. *Water Resources Research*, 50

- Memarian, H., Balasundram, S. K., Talib, J. B., Sung, C. T. B., Sood, A. M., & Abbaspour, K. (2012). Validation of CA-markov for simulation of land use and cover change in the Langat basin, Malaysia. *Journal of Geographic Information System, 4*, 542-554.
- Menard, S. (1995). Applied logistic regression analysis. *Sage University Paper Series on Quantitative Applications in Social Sciences, 106*, 98.
- Mertens, B., & Lambin, E. F. (1997). Spatial modelling of deforestation in southern Cameroon: Spatial disaggregation of diverse deforestation processes. *Applied Geography, 17*(2), 143-162.
- Miller, D. J., & Plantinga, A. J. (1999). Modeling land use decisions with aggregate data. *American Journal of Agricultural Economics, 81*(1), 180-194.
- Millington, James, Romero-Calcerrada, Raúl, Wainwright, John and Perry, George. (2008). An agent-based model of Mediterranean agricultural land-Use/Cover change for examining wildfire risk. *Journal of Artificial Societies and Social Simulation, 11*(4), 4.
- Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review, 57*(2), 197-210.
- Muth, R. F. (1969). *Cities and housing*. Chicago: University of Chicago Press.
- Nagubadi, R. V., & Zhang, D. (2005). Determinants of timberland use by ownership and forest type in Alabama and Georgia. *Journal of Agricultural and Applied Economics, 37*(1), 173-186.

- Najlis, R., & North, M. J. (2004). *Repast for GIS. Agent 2004: Social Dynamics: Interaction, Reflexivity and Emergence*, Chicago, IL.
- Natural Resource Analysis Center (NRAC). (2007). *Watershed Characterization Modeling System for ArcGIS 9.2*. Release 1.0. Natural Resource Analysis Center, West Virginia University, Morgantown, WV.
- Neitsch, S. L., Arnold, J. G., Kiniry, J. R., & Williams, J. R. (2005). *Soil and water assessment tool - theoretical documentation – version 2005*. Grassland, Soil and Water Research Laboratory, Agricultural Research Service and Blackland Research Center, Texas Agricultural Experiment Station, Temple, Texas. Retrieved from <http://swat.tamu.edu/media/1292/swat2005theory.pdf>
- Ng, T. L., Eheart, J. W., Cai, X., & Braden, J. B. (2011). An agent-based model of farmer decision-making and water quality impacts at the watershed scale under markets for carbon allowances and a second-generation biofuel crop. *Water Resources Research*, 47(9), - W09519.
- Niraula, R., Kalin, L., Srivastava, P., & Anderson, C. J. (2013). Identifying critical source areas of nonpoint source pollution with SWAT and GWLF. *Ecological Modelling*, 268(0), 123-133.
- Nowak, D. J., & J. T. Walton. (2005). Projected urban growth (2000–2050) and its estimated impact on the U.S. forest resource. *Journal of Forestry*, 103(8), 383-389.
- Nzaku, K., & Bukenya, J. O. (2005). Economic analysis of agricultural land conversion in the southeast. *American Agricultural Economics Annual Meeting*, July 24-27, 2005. Providence, Rhode Island.

Olson, R. K., & A. H. Olson. (1999). Farmland loss in America. In R. K. Olson, & T. A. Lyson (Eds.), *Under the blade: The conversion of agricultural landscapes* (pp. 15-52) Westview Press, Boulder, CO.

Opequon Creek IP Steering Committee. (2006). *Opequon creek watershed TMDL implementation plan*. Retrieved from

<http://www.deq.virginia.gov/Portals/0/DEQ/Water/TMDL/ImplementationPlans/opeqip.pdf>

Opequon Creek Project Team. (2009).

Retrieved from <http://www.opequoncreek.org/Publications/Brochure-09.PDF>

Opequon Creek Project Team. (2014). Retrieved from <http://www.opequoncreek.org/>

Osborne, L. L., & Kovacic, D. A. (1993). Riparian vegetated buffer strips in water-quality restoration and stream management. *Freshwater Biology* 29:243–258

Osborne L. L., & Wiley M. J. (1988). Empirical relationships between land use/cover and stream water quality in an agricultural watershed. *Journal of Environmental Management*, 26, 9-27.

O’Sullivan, D., Millington, J., Perry, G., & Wainwright, J. (2012). Agent-based models – because they’re worth it? In Heppenstall, A. J., Crooks, A.T., See, L.M., Batty, M. (Ed.), *Agent-based models of geographical systems*. Springer.

Ozah, A. P., Dami, A., & Adesina, F. A. (2012). A deterministic cellular automata model for simulating rural land use dynamics: A case study of Lake Chad basin. *Journal of Earth Science & Engineering*, 2(1), 22.

- Palmquist, R. B. (1989). Land as a differentiated factor of production: A hedonic model and its implications for welfare measurement. *Land Economics*, 65(1), 23-28.
- Papadopoulos, C. E., & Yeung, H. (2001). Uncertainty estimation and Monte Carlo simulation method. *Flow Measurement and Instrumentation*, 12(4), 291-298.
- Park, J. Y., Park, M. J., Joh, H. K., Shin, H. J., Kwon, H. J., Srinivasan, R., & Kim, S. J. (2011). Assessment of MIROC3.2 HiRes climate and CLUE-s land use change impacts on watershed hydrology using SWAT. *Transactions of the ASABE*, 54(5), 1713-1724.
- Parker, D. (2000). *Edge-effect externalities: Theoretical and empirical implications of spatial heterogeneity*. Ph.D. Thesis, University of California, Davis.
- 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. *Annals of the Association of American Geographers*, (94), 314-337.
- Parker, D. C., & Meretsky, V. (2004). Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. *Agriculture, Ecosystems and Environment*, 101, 233-250.
- 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.

- Parker, D. C., & Tatiana, F. (2008). A conceptual design for a bilateral agent-based land market with heterogeneous economic agents. *Computers, Environment and Urban Systems*, 32(6), 454-463.
- Peterson, B.J., Wollheim, W.M., Mulholland, P.J., Webster, J.R., Meyer, J.L., Tank, J.L., Martí, E., Bowden, W.B., Valett, H.M., Hershey, A.E., McDowell, W.H., Dodds, W.K., Hamilton, S.K., Gregory, S., & Morrall, D.D. (2001). Control of nitrogen export from watersheds by headwater streams. *Science*, 292 (5514), 86-90.
- Pionke, H. B., Gburek, W. J., & Sharpley, A. N. (2000). Critical source area controls on water quality in an agricultural watershed located in the Chesapeake basin. *Ecological Engineering*, 14(4), 325-335.
- Polhill, J. Gary., Parker, Dawn., Brown, Daniel., & Grimm, Volker. (2008). Using the ODD protocol for describing three agent-based social simulation models of land-use change. *Journal of Artificial Societies and Social Simulation*, 11(2), 3.
- Polyakov, M., & Zhang, D. (2008). Population growth and land use dynamics along Urban–Rural gradient. *Journal of Agricultural and Applied Economics*, 40(2), 649-666.
- Pontius Jr, R. G., Huffaker, D., & Denman, K. (2004). Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling*, 179(4), 445-461.
- Pontius Jr, R. G., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment*, 85(1–3), 239-248.

- Pontius Jr, R. G., Neeti, N. (2010). Uncertainty in the difference between maps of future land change scenarios. *Land use and Ecosystems*, 5, 39-50.
- Pontius, R. G., & Schneider, L. C. (2000). Land-use change model validation by a ROC method. *Agriculture, Ecosystems and Environment*, 85, 269-280.
- Poudyal, N. C., Cho, S., Strickland, J. D., & Hodges, D. G. (2008). Socio-demographic and market forces of forest land use change on the northern Cumberland Plateau, Tennessee. *International Journal of Ecological Economics & Statistics*, 10(W08), 53-62.
- Pozzi, F., & Small, C. (2005). Analysis of urban land cover and population density in the United States. *Photogrammetric Engineering and Remote Sensing*, 71, 719-726.
- Qian, Z. (2010). Without zoning: Urban development and land use controls in Houston. *Cities*, 27(1), 31-41.
- Radeloff, V.C., Nelson, E., Plantinga, A .J., Lewis, D.J., Helmers, D., Lawler, J.J., Withey, J.C., Beaudry, F., Martinuzzi, S., Bustic, V., Lonsdorf, E., White, D., & Polasky, S. (2012). Economic-based projections of future land use in the conterminous United States under alternative policy scenarios. *Ecological Applications*, 22(3), 1036-1049.
- Rammel, C., Stagl, S., & Wilfing, H. (2007). Managing complex adaptive systems - A co-evolutionary perspective on natural resource management. *Ecological Economics*, 63(1), 9-21.
- Reeves, H. W., & Zellner, M. L. (2010). Linking MODFLOW with an agent-based land-use model to support decision making. *Ground Water*, 48(5), 649-660.



- Roe, B., Irwin, E. G., & Marrow-Jones, H. A. (2004). The effects of farmland, farmland preservation, and other neighborhood amenities on housing values and residential growth. *Land Economics*, 80(1), 55-75.
- Rosenberger, R. S., Gebremedhin, T. G., & Hailu, Y. G. (2002). *An economic analysis of urbanization of agricultural land in West Virginia*. (Research Paper No. 8). West Virginia Regional Research Institute.
- Schueler, T., Fraley-McNeal, L., & Capiella, K. (2009). Is impervious cover still important? Review of recent research. *Journal of Hydrologic Engineering*, 14(4), 309-315.
- Serneels, S., & Lambin, E. F. (2001). Proximate causes of land-use change in Narok district, Kenya: A spatial statistical model. *Agriculture, Ecosystems and Environment*, 85, 65-81.
- Shirzadi, A., Saro, L., Joo, O. H., & Chapi, K. (2012). A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavatabad case study, Kurdistan, Iran. *Natural Hazards*, 64(2), 1639-1656.
- Sidharthan, R., & Bhat, C. R. (2012). Incorporating spatial dynamics and temporal dependency in land use change models. *Geographical Analysis*, 44(4), 321-349.
- Simon HA. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69 (1), 99-118.
- Sivertun, Å., & Prange, L. (2003). Non-point source critical area analysis in the gisselö watershed using GIS. *Environmental Modelling & Software*, 18, 887-898.

- Sliva, L., & Williams, D. (2001). Buffer zone versus whole catchment approaches to studying land use impact on river water quality. *Water Research*, 35(14), 3462-3472.
- Sohl, T.L., Sayler, K.L., Bouchard, M. A., Reker, R.R., Friesz, A.M., Bennett, S.L., Sleeter, B.M., Sleeter, R.R., Wilson, T.S., Soulard, C., Knuppe M., & Hofwegen T. V. (2014) Spatially explicit modeling of 1992–2100 land cover and forest stand age for the conterminous United States. *Ecological Applications*, 24(5), pp. 1015–1036
- Sohn, K. T., & Park, S. M. (2008). Guidance on the choice of threshold for binary forecast modeling. *Advances in Atmospheric Sciences*, 25(1), 83-88.
- Srinivasan, M. S., Gérard-Marchant, P., Veith, T. L., Gburek, W. J., & Steenhuis, T. S. (2005). Watershed scale modeling of critical source areas of runoff generation and phosphorus transport. *JAWRA Journal of the American Water Resources Association*, 41(2), 361-377.
- Srinivasana, M. S., & McDowellb, R. W. (2007). Hydrological approaches to the delineation of critical-source areas of runoff. *New Zealand Journal of Agricultural Research*, 50(2)
- Stocks, L. J. (2010). *A spatio-temporal analysis of land use and land cover change and sinkhole development in Opequon creek watershed, West Virginia: 1984-2009*. (Doctor of Philosophy) Kent State University.
- Strager, M. P., Fletcher, J. J., Strager, J. M., Yuill, C. B., Eli, R. N., Todd Petty, J., & Lamont, S. J. (2010). Watershed analysis with GIS: The watershed characterization and modeling system software application. *Computers & Geosciences*, 36(7), 970-976.

- Sun, Y., Guo, Q., Liu, J., & Wang, R. (2014). Scale effects on spatially varying relationships between urban landscape patterns and water quality. *Environmental Management*, 54, 272-287.
- Swets, J. A. (1986). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293.
- Tayyebi, A., Delavar, M. R., Yazdanpanah, M. J., Pijanowski, B. C., Saeedi, S., & Tayyebi, A. H. (2010). A spatial logistic regression model for simulating land use patterns, a case study of the shiraz metropolitan area of Iran In Chuvieco, E., Li, J., Yang, X. (Ed.), *Advances in earth observation of global change*, Springer Press.
- Tong, S. T. Y., & Chen, W. (2002). Modeling the relationship between land use and surface water quality. *Journal of Environmental Management*, 66(4), 377-393.
- Tsai, Y., Zia, A., Koliba, C., Guilbert, J., Bucini, G., & Beckage, B. (2013). Impacts of land managers' decisions on landuse transition within Missisquoi watershed Vermont: An application of agent-based modeling system. Paper presented at the *Systems Conference (SysCon), 2013 IEEE International*, 824-829.
- Tu, J. (2011). Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Applied Geography*, 31(1), 376-392.
- Turner, M. G. (1989). Landscape ecology: The effect of pattern on ecology. *Annual Review of Ecology and Systematics*, 20, 171-197.

- U.S. Census Bureau. (2000). Population density by census tract, 2000. Retrieved from <http://www.census.gov/>
- HUD (U.S. Department of Housing and Urban Development). (2012). *Comprehensive housing market analysis: Hagerstown-Martinsburg, Maryland-West Virginia*. (Research No. U.S. Department of Housing and Urban Development). Retrieved from [http://www.huduser.org/publications/pdf/HagerstownMD\\_comp\\_12.pdf](http://www.huduser.org/publications/pdf/HagerstownMD_comp_12.pdf)
- U.S. Department of Transportation (1997): Bureau of Transportation Statistics.  
<http://www.rita.dot.gov/bts/home>
- U.S. Geological Survey. (2014). National land cover database (NLCD). Retrieved from <http://www.mrlc.gov/index.php>
- United States Environmental Protection Agency. (2007). Multi-resolution land characteristics consortium (MLRC). Retrieved from <http://www.epa.gov/mrlc/definitions.html>
- Valbuena D., Verburg P., Bregt A.K., & Ligtenberg A. (2010). An agent-based approach to model land-use change at a regional scale. *Landscape Ecology*, 25, pp. 185–199
- Veldkamp A., Verburg P.H., Kok K., De koning G.H.J., Priess J., & Bergsma, A.R. (2001). The need for scale sensitive approaches in spatially explicit land use change modeling. *Environmental Modeling and Assessment*, 6, 111-121.
- Veldkamp, T., Polman, N., Reinhard, S., & Slingerland, M. (2011). From scaling to governance of the land system: Bridging ecological and economic perspectives. *Ecology and Society*, 16(1), 1.

- Veldkamp, A., & Fresco, L. O. (1996). CLUE: A conceptual model to study the conversion of land use and its effects. *Ecological Modelling*, 85(2-3), 253-270.
- Verburg P.H., Soepboer W., Veldkamp A., Limpiada R., Espaldon V., & Mastura S.S.A. (2002). Modeling the spatial dynamics of regional land use: The CLUE-S model. *Environmental Management*, 30, 391-405.
- Voinov, A., Costanza, R., Wainger, L., Boumans, R., Villa, F., Maxwell, T., & Voinov, H. (1999). Patuxent landscape model: Integrated ecological economic modeling of a watershed. *Environmental Modelling & Software*, 14(5), 473-491.
- Von Thünen, J. H. (1826). In P.G. Hall e. (Ed.), *Die isolierte staat in beziehung auf landwirtschaft und nationalökonomie*. (Wartenberg C M in 1966 Trans.). Pergamon Press, New York.
- VT CTMDLWS. (2006). *Opequon creek watershed TMDL implementation plan*. Virginia tech center for TMDL and watershed studies. Retrieved from [http://www.tmdl.bse.vt.edu/uploads/File/pub\\_db\\_files/Opequon%20Creek%20TMDL%20IP%2007-12-06.pdf](http://www.tmdl.bse.vt.edu/uploads/File/pub_db_files/Opequon%20Creek%20TMDL%20IP%2007-12-06.pdf)
- Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association*, 68(3), 297-314.
- Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M., & Ulfarsson, G. (2003). Microsimulation of urban development and location choices: Design and implementation of UrbanSim. *Networks and Spatial Economics*, 3(1), 43-67.

Water Resources and TMDL Center. (2008). Total maximum daily loads for selected streams in the Potomac direct drains watershed, West Virginia. Prepared for: West Virginia Department of Environmental Protection Division of Water and Waste Management Watershed Branch, TMDL section.

Weiler, S. (2007). Economics and contemporary land use policy: Development and conservation at the rural-urban fringe ? edited by robert J.J and stephen K. S. *Growth and Change*, 38(1), 163-166.

Weinschenck, G., Henrichsmeyer, W., & Aldinger, F. (1969). The Theory of Spatial Equilibrium and Optimal Location in Agriculture: A Survey. *Review of Marketing and Agricultural Economics*. 37:3-70

Weng, Q. (2001). Modeling urban growth effects on surface runoff with the integration of remote sensing and GIS. *Environmental Management*, 28(6), 737-748.

West Virginia Department of Environmental Protection, Canaan Valley Institute, & the Opequon Creek Project Team, Inc. (2008). *Watershed based plan for mill creek "A tributary of Opequon creek, in the Potomac direct drains watershed" Berkeley County, WV*. Retrieved from [http://www.dep.wv.gov/WWE/Programs/nonptsources/WBP/Documents/WBP/MillCreekOpequon\\_WBP.pdf](http://www.dep.wv.gov/WWE/Programs/nonptsources/WBP/Documents/WBP/MillCreekOpequon_WBP.pdf)

West Virginia WIP Development Team. (2012). *West Virginia's Chesapeake Bay TMDL final phase II watershed implementation plan*. Chesapeake Bay Program. Retrieved from

- White, R., & Engelen, G. (2000). High-resolution integrated modeling of spatial dynamics of urban and regional systems. *Computers, Environment, and Urban Systems* 24:383–400.
- White R., & Engelen, G. (1993). Cellular automata and fractal urban form: A cellular modelling approach to the evolution of urban land-use patterns. *Environment and Planning A*, 25(8), 1175-1199.
- White, E. M., Morzillo, A. T., & Alig, R. J. (2009). Past and projected rural land conversion in the US at state, regional, and national levels. *Landscape and Urban Planning*, 89(1–2), 37-48.
- Williams, J.R. (1995). The EPIC Model. pp 909-1000 In Singh, V. P. *Computer models of watershed hydrology*, Water Resources Publications, Highlands Ranch, CO.
- Wooldridge, J. (2002). *Introductory econometrics* South-Western, New York.
- 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, J. (2006). Environmental amenities, urban sprawl, and community characteristics. *Journal of Environmental Economics and Management*, 52(2), 527-547.
- Xiang, W. (1996). GIS-based riparian buffer analysis: Injecting geographic information into landscape planning. *Landscape and Urban Planning*, 34(1), 1-10.
- Xie, C., Huang, B., Claramunt, C., & Chandramouli, M. (2005). Spatial logistic regression and GIS to model rural-urban land conversion. *Proceedings of PROCESSUS Second International*

*Colloquium on the Behavioural Foundations of Integrated Land-use and Transportation Models: Frameworks, Models and Applications*, University of Toronto, Canada.

Zeeb, C. N., & Burns, P. J. (1998). *A comparison of failure probability estimates by Monte Carlo sampling and Latin hypercube sampling*. (Technical Report). Sandia National Laboratories.

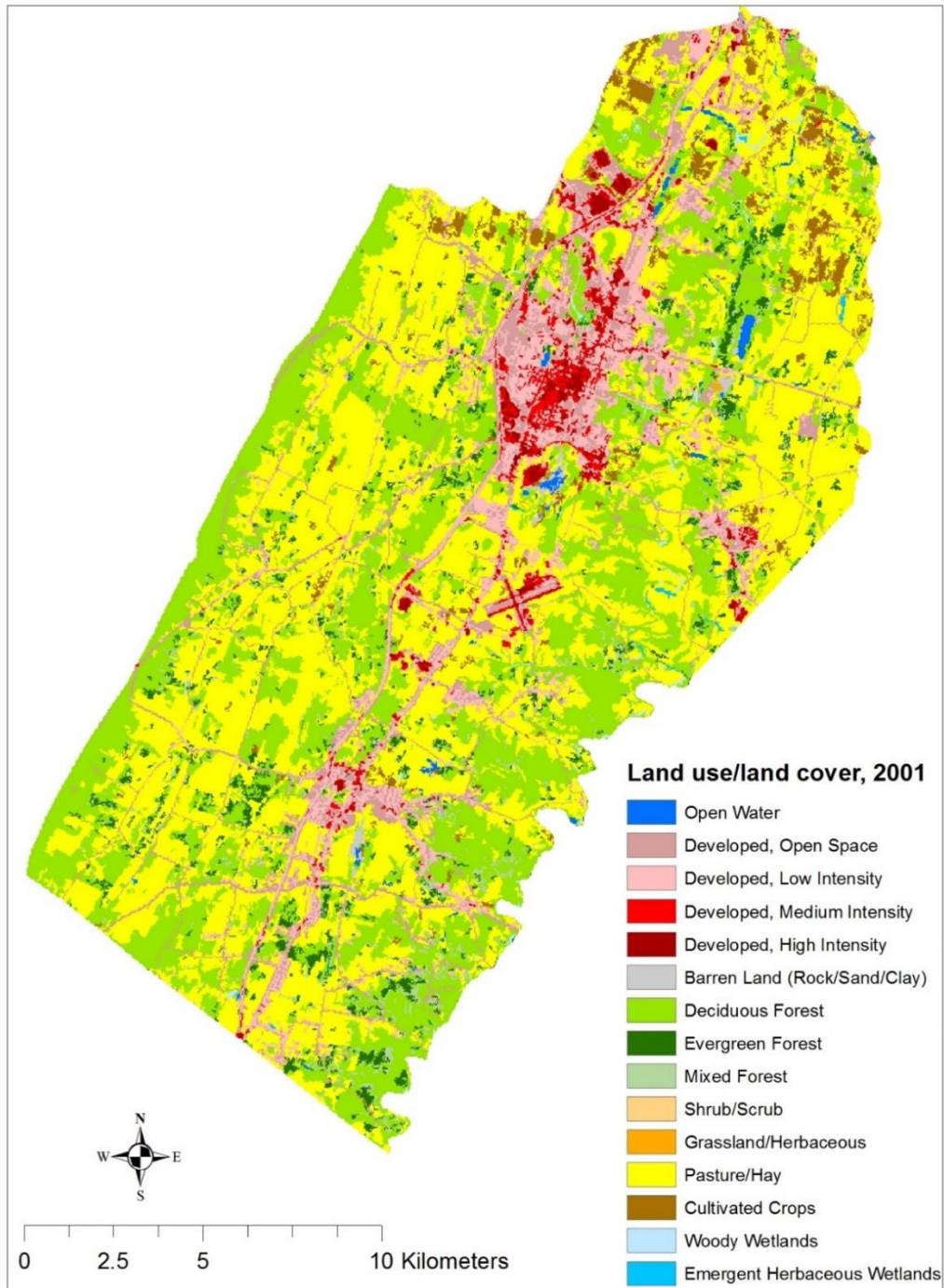
Retrieved from <https://www.palisade.com/downloads/pdf/MCvsLHsampling.pdf>

Zhang, D. T., & Nagubadi, R. V. (2005). The influence of urbanization on timberland use by forest type in the southern United States. *Forest Policy and Economics*, 7, 721-731.

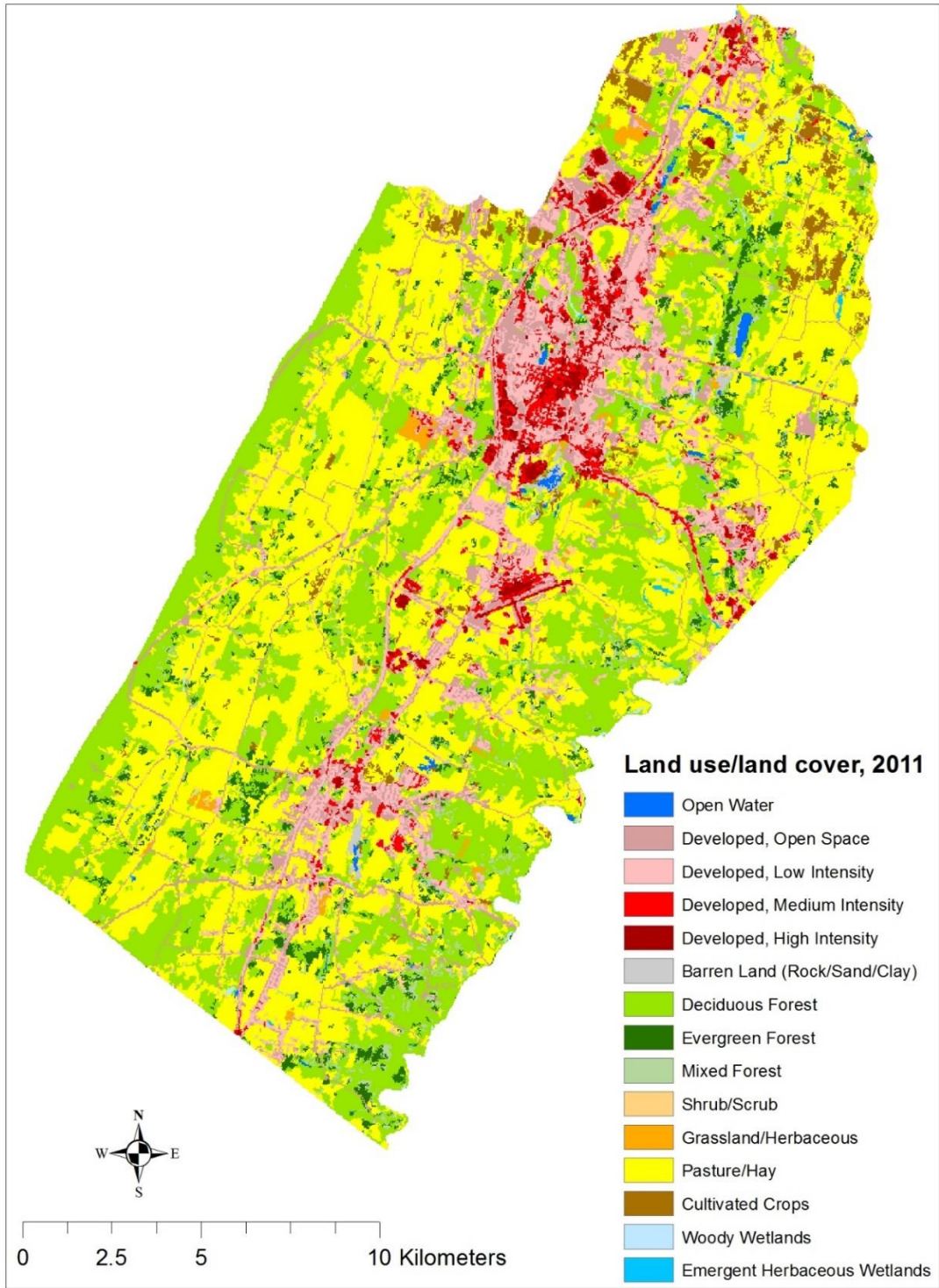


## **APPENDICES**

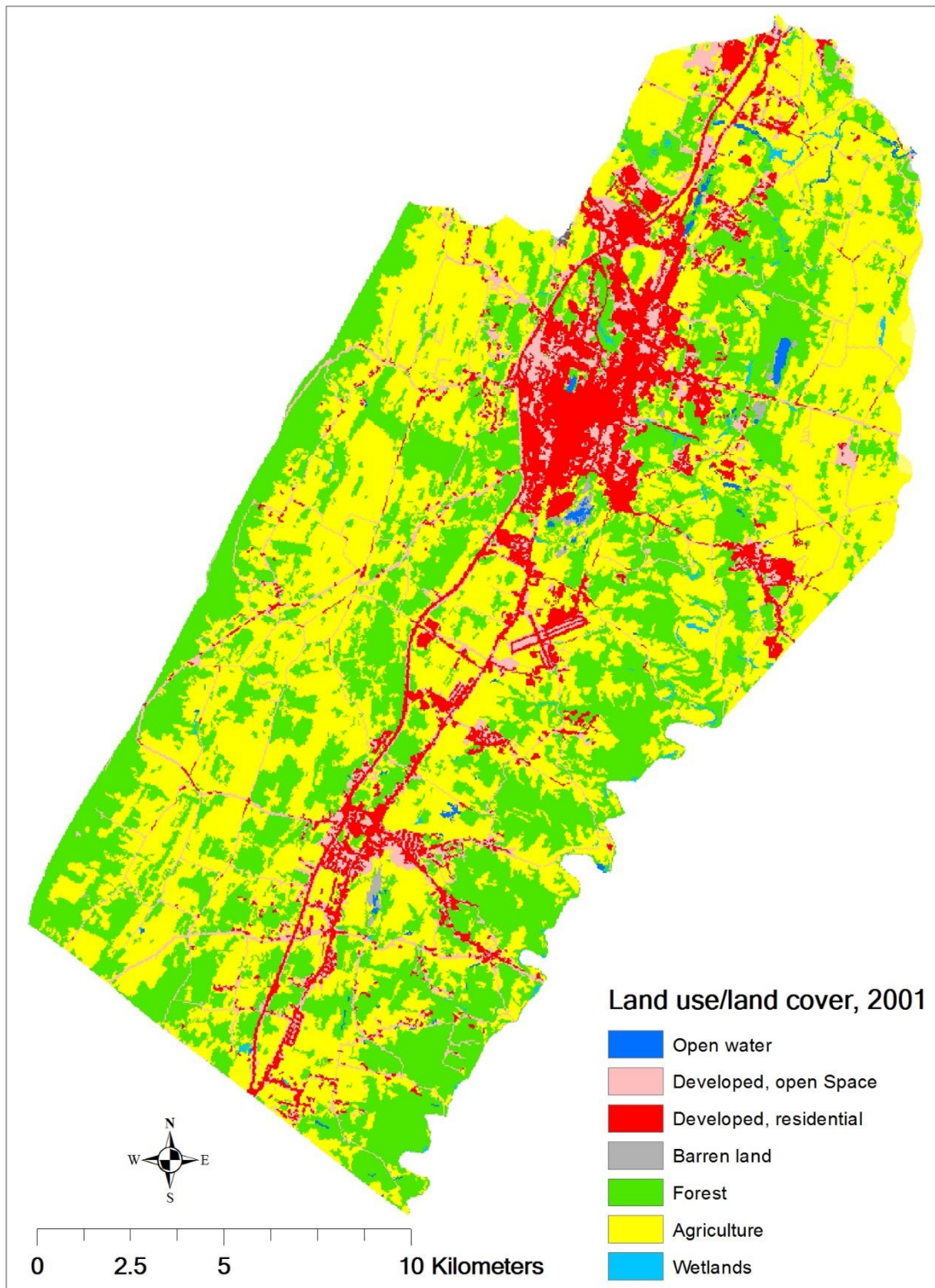
## APPENDIX I: Spatial data



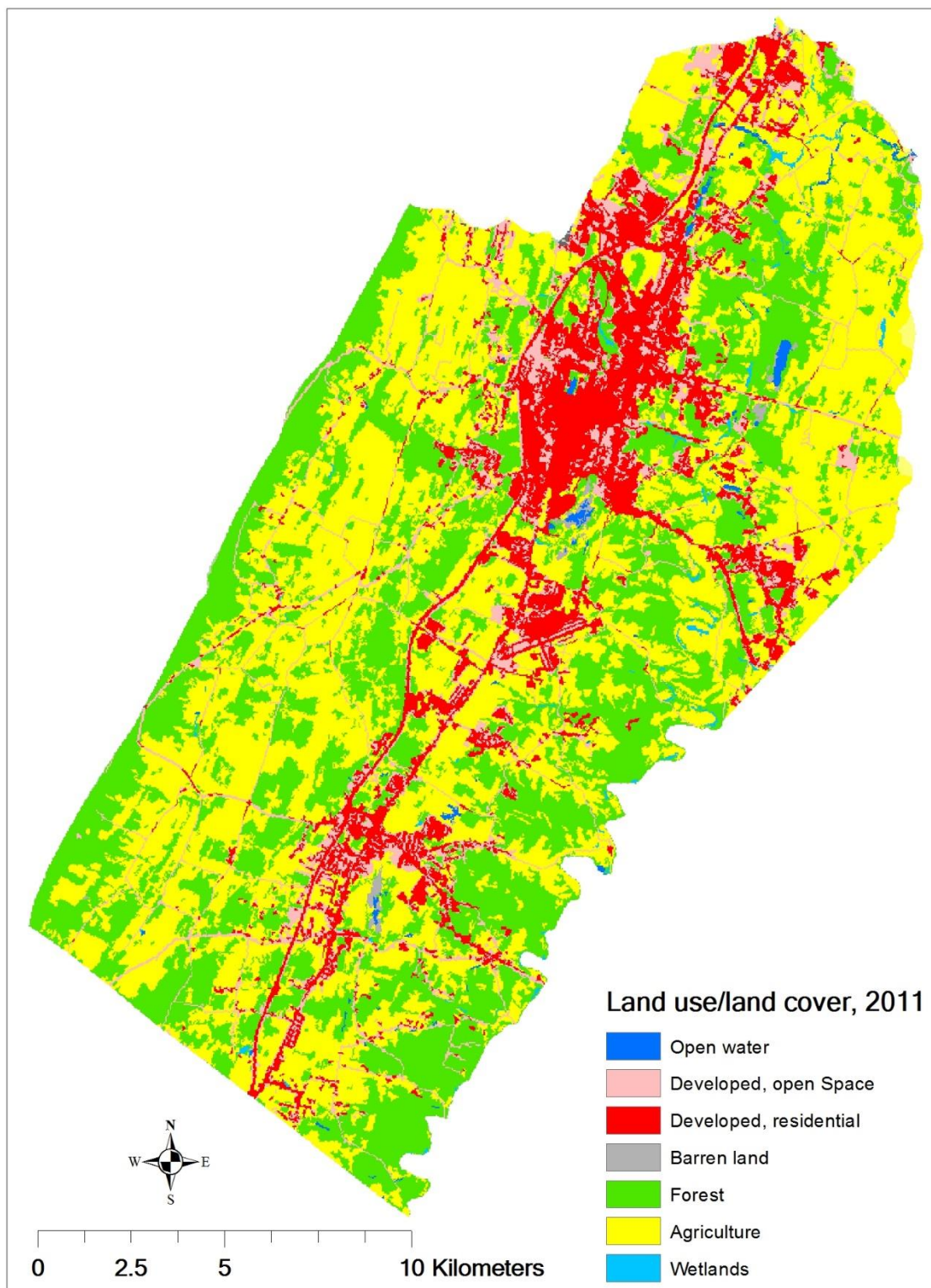
NLCD land use/land cover for the Opequon Creek watershed, WV, 2001.



NLCD land use/ land cover for the Opequon Creek watershed, WV, 2011.



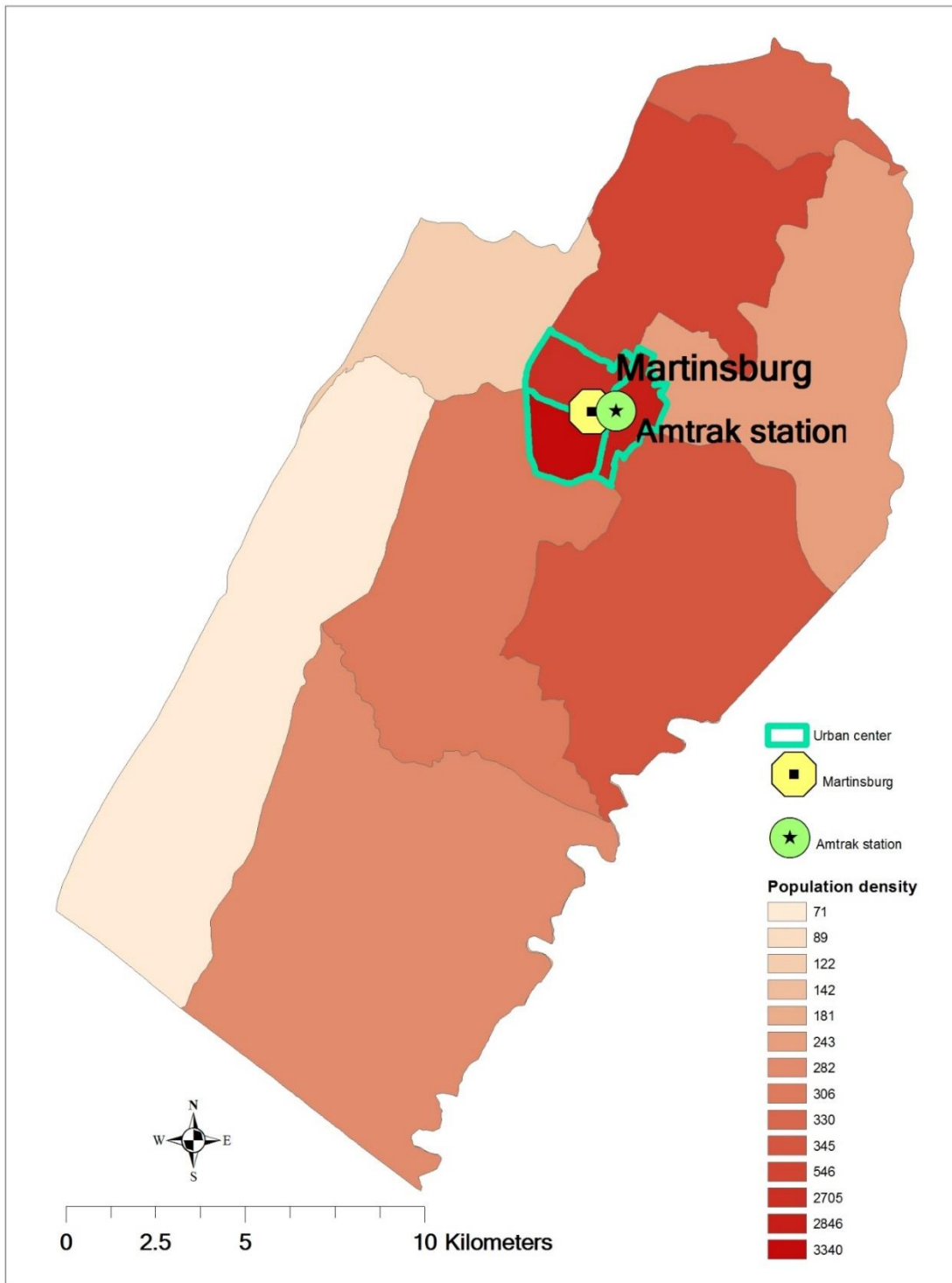
Reclassified land use/land cover for the Opequon Creek Watershed, WV, 2001.



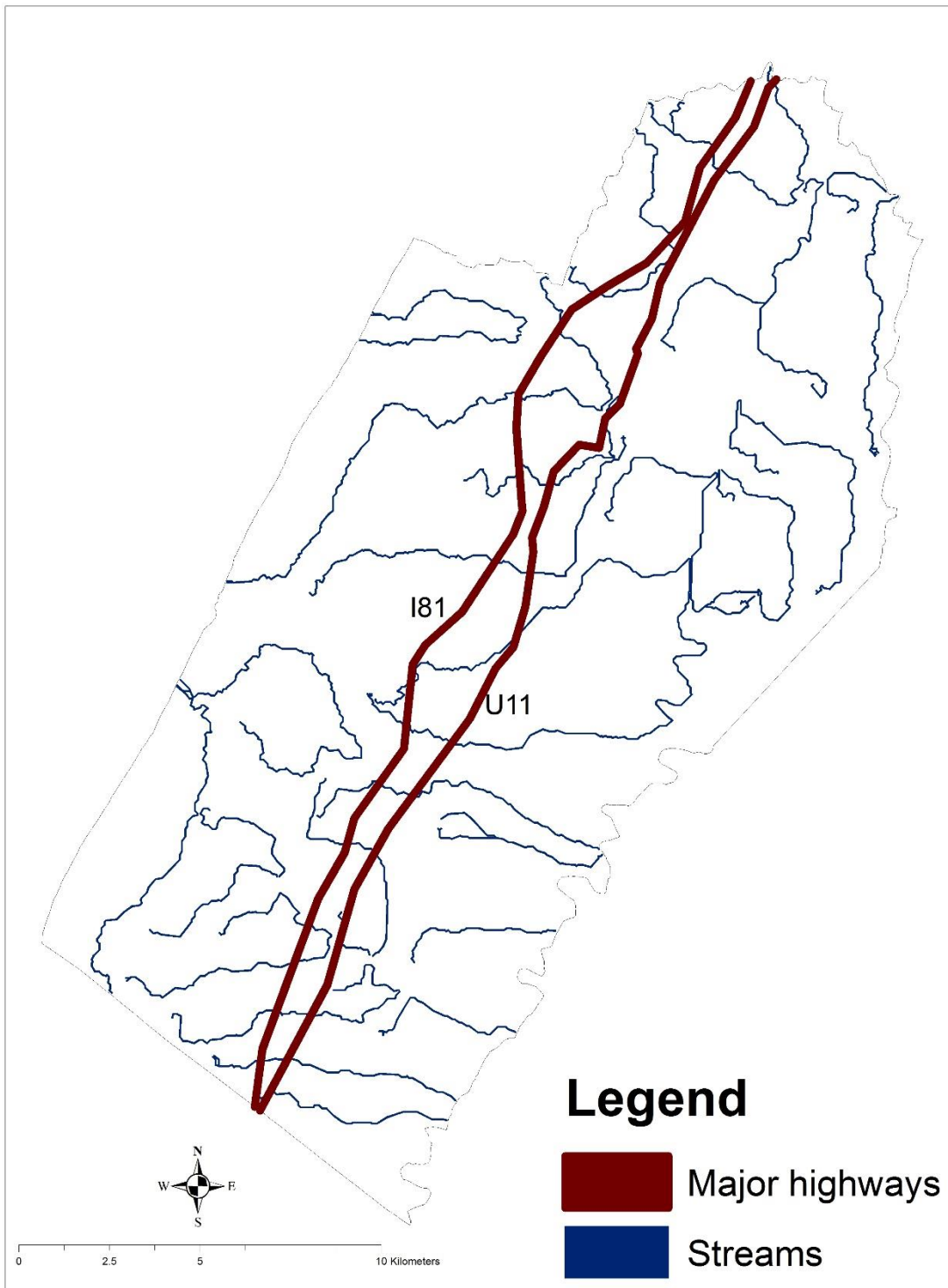
Reclassified land use/land cover for the Opequon Creek watershed, WV, 2011.



Property parcels in the Opequon Creek watershed in 2011.

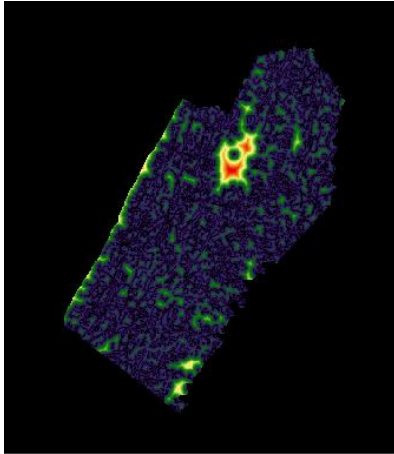


Urban center designation in the Opequon Creek watershed, 2000.

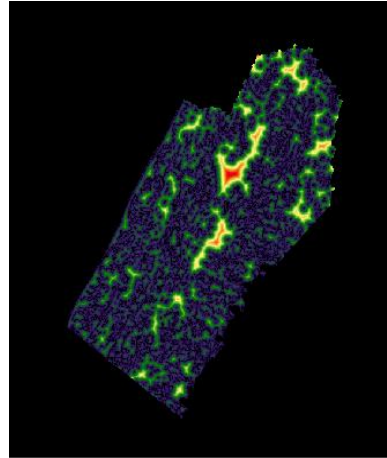


Major highways and streams in the Opequon Creek watershed.

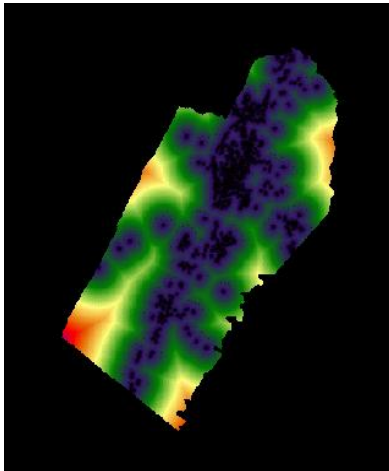




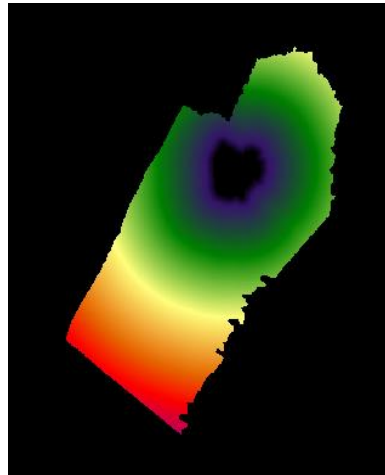
Distance from agricultural land (meters)



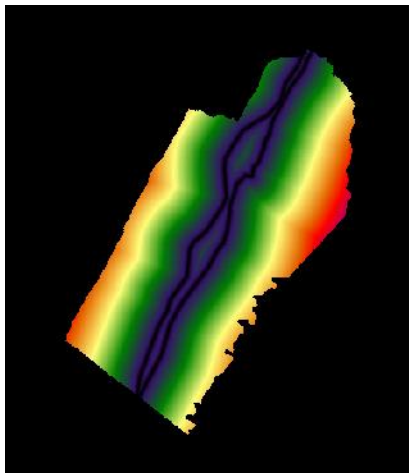
Distance from forest land (meters)



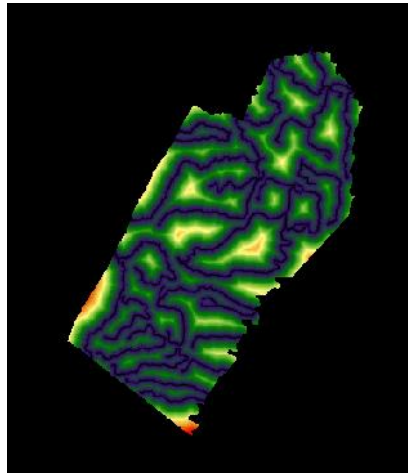
Distance from residential land use (meters)



Distance from urban center (meters)

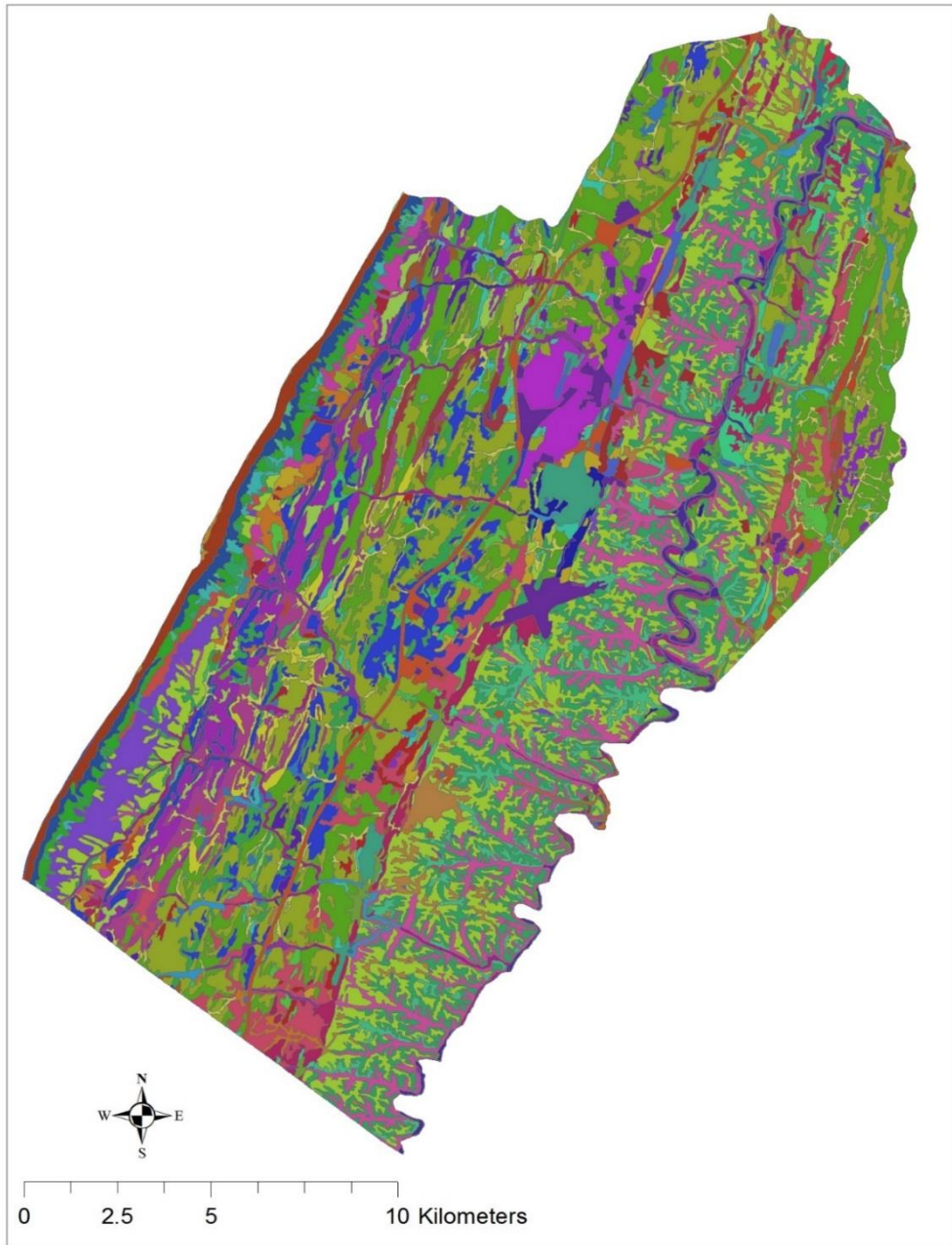


Distance from roads (meters)

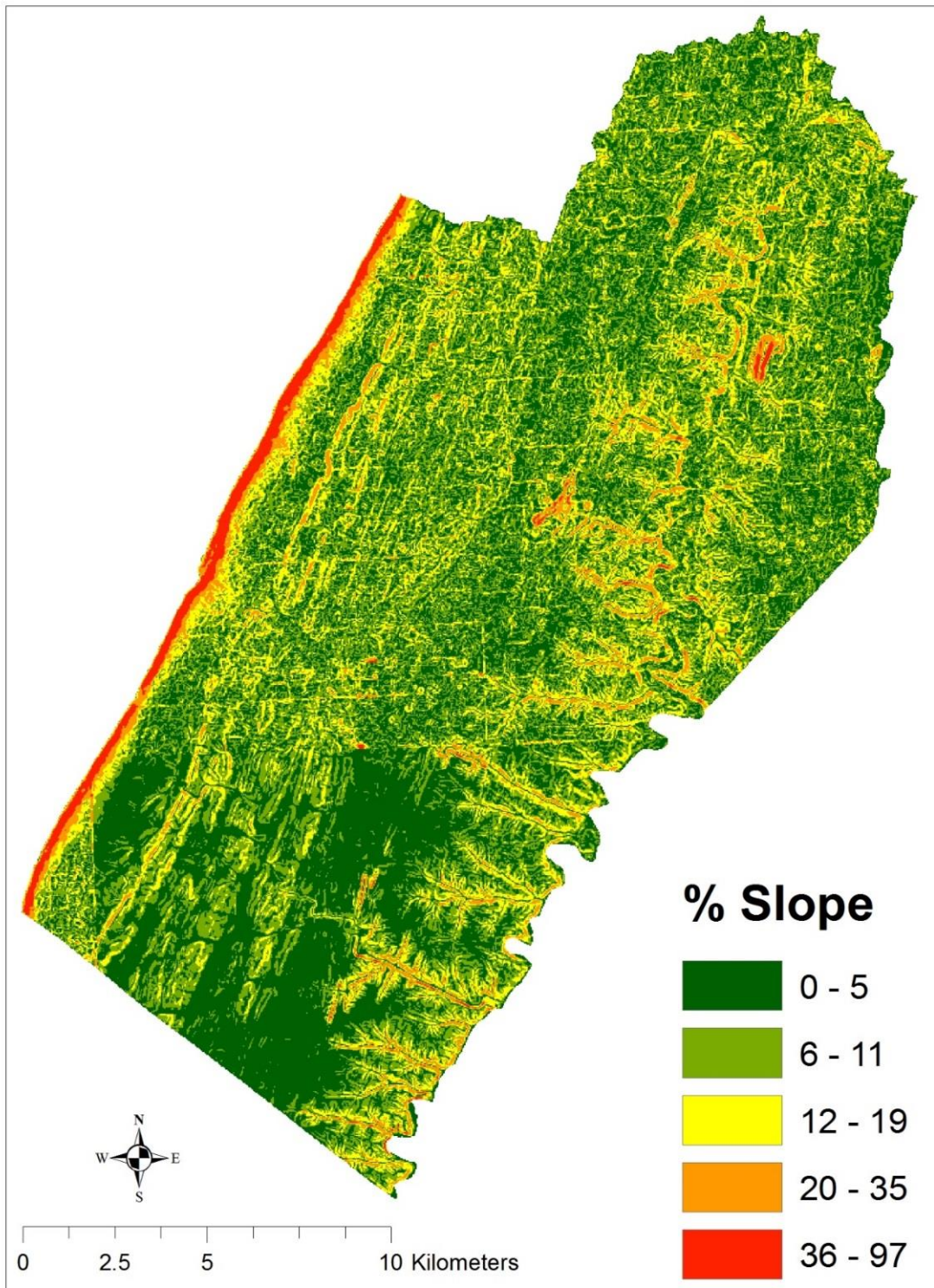


Distance from streams (meters)

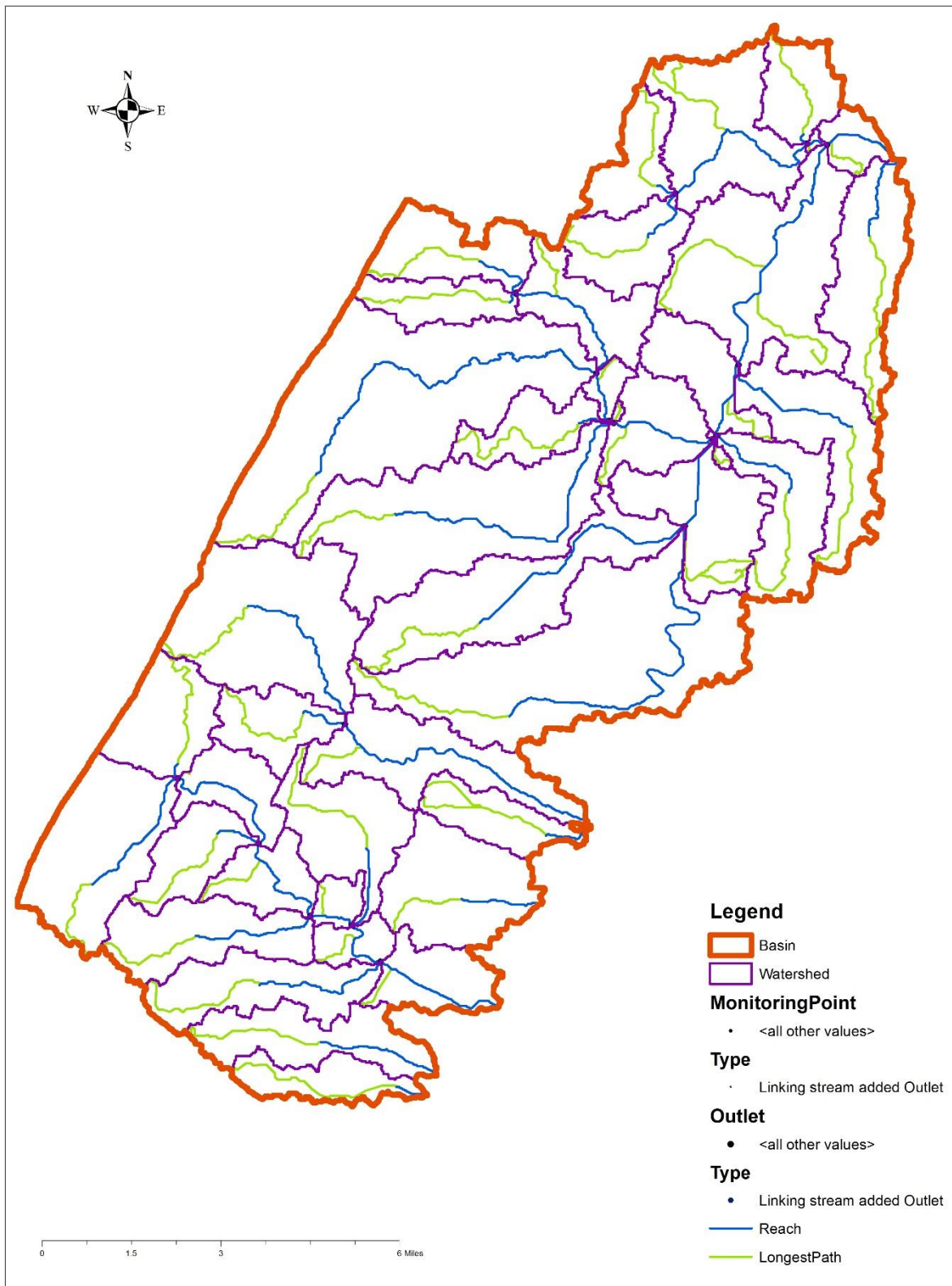
Visual demonstration of the proximity variables within the Opequon Creek watershed that were used in spatial logistic regression model.



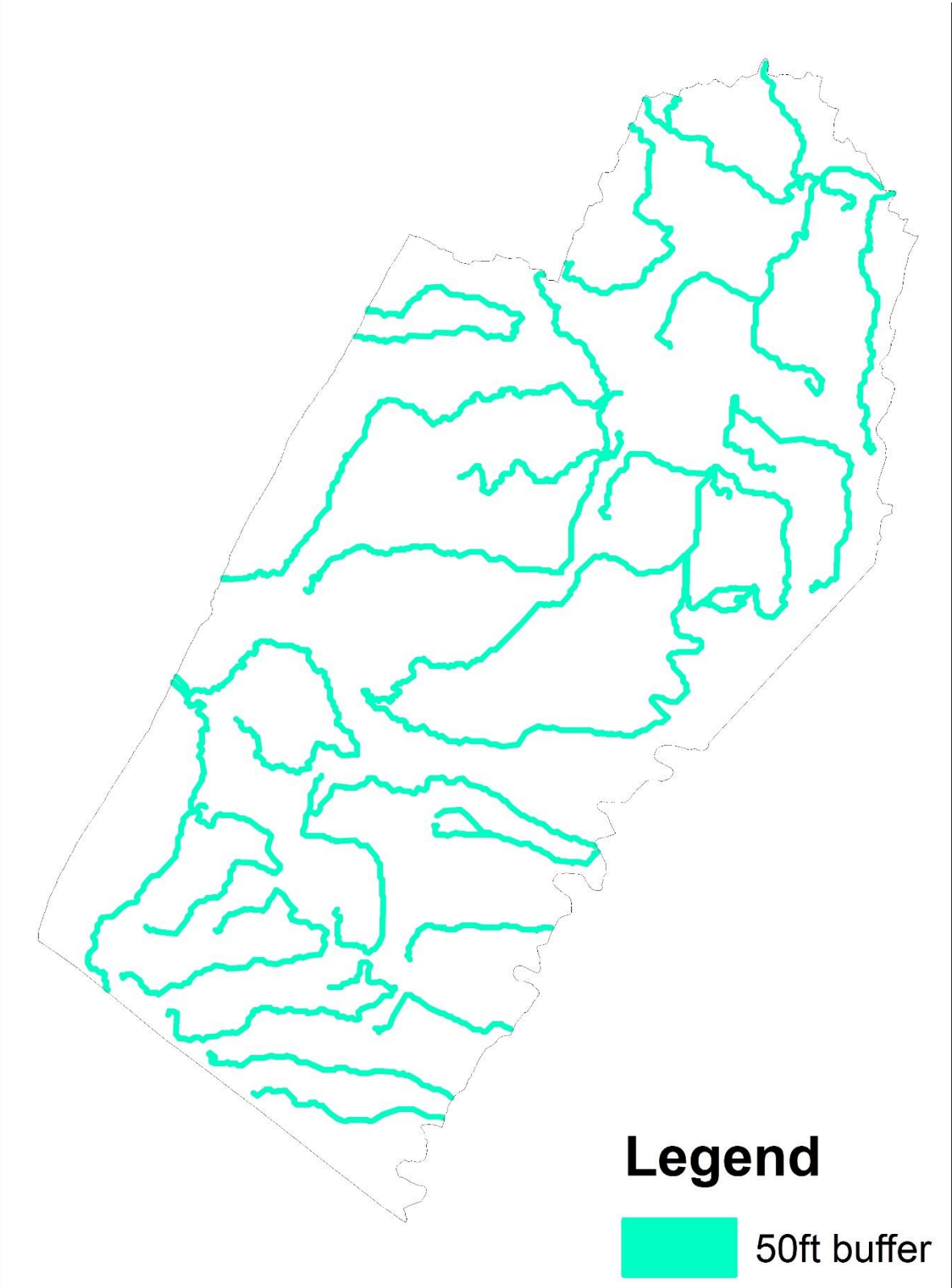
Soil survey geographic (SSURGO) database in the Opequon Creek watershed.



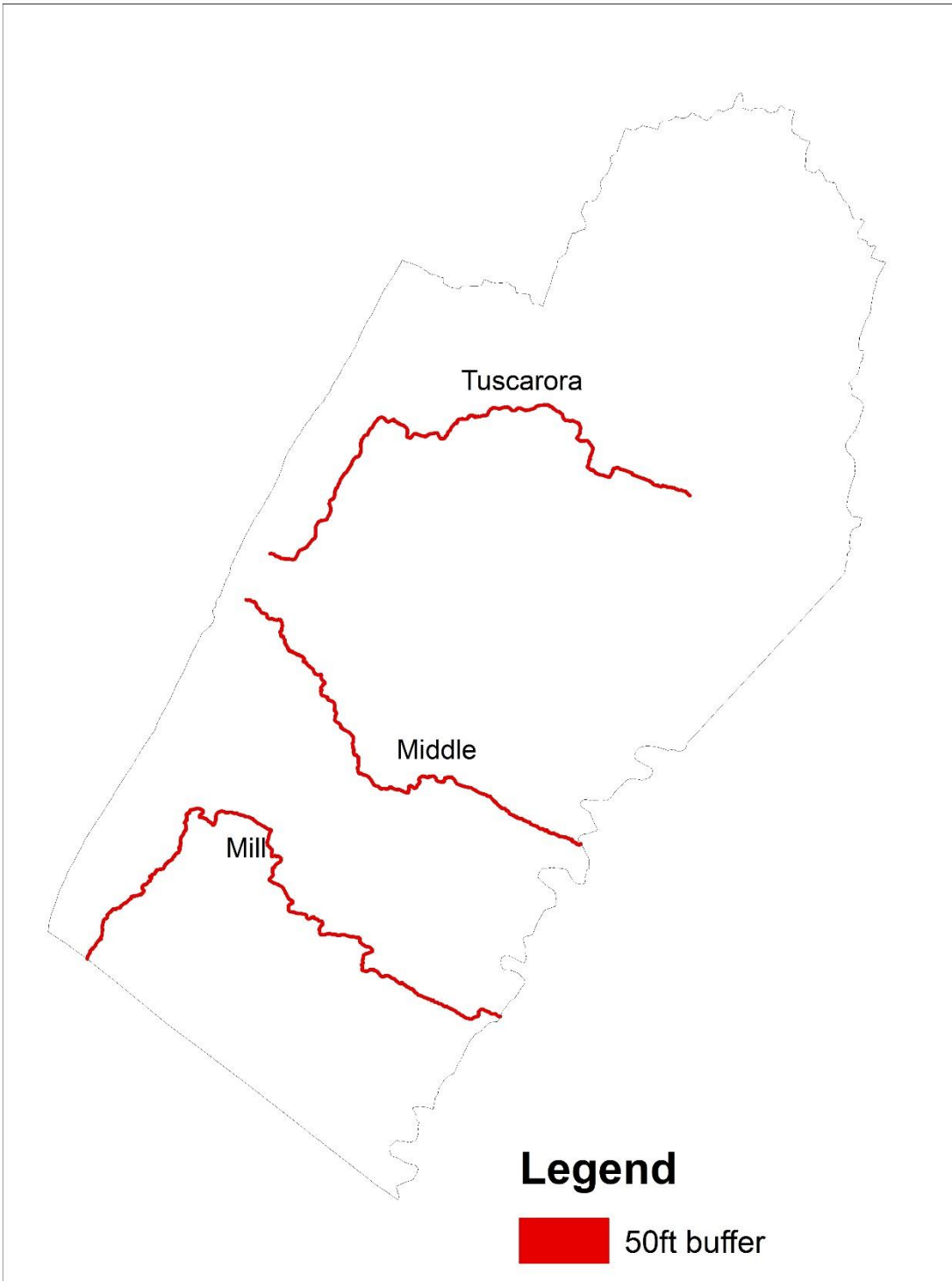
Visual demonstration of slope classes expressed as percentage in the ArcSWAT model



Watershed delineation of Opequon Creek watershed in ArcSWAT model



50 foot Riparian Buffer Policy (all streams)

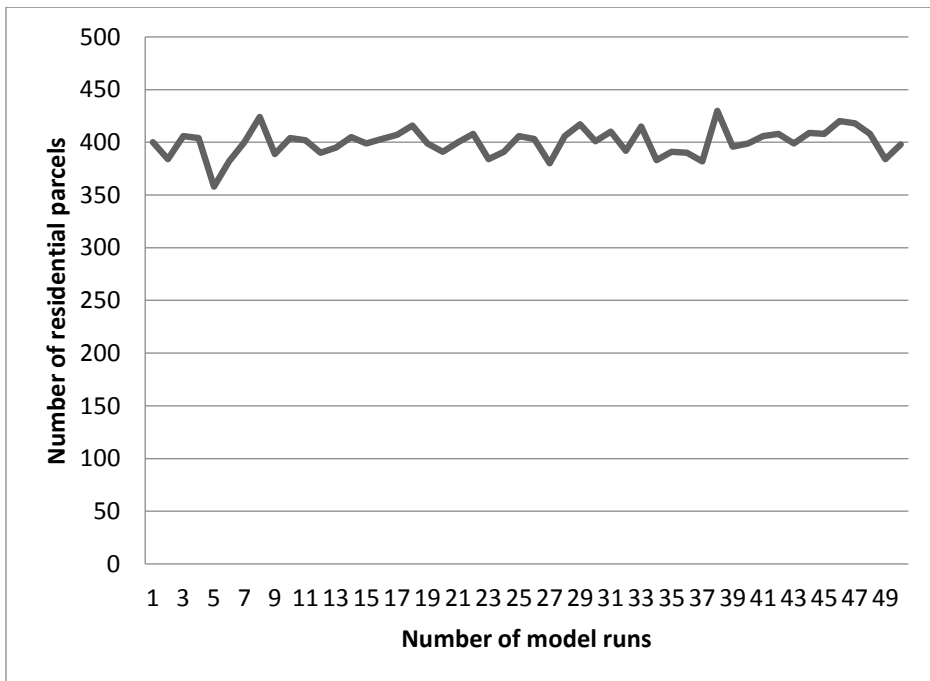


50 foot Riparian Buffer Policy in CSAs

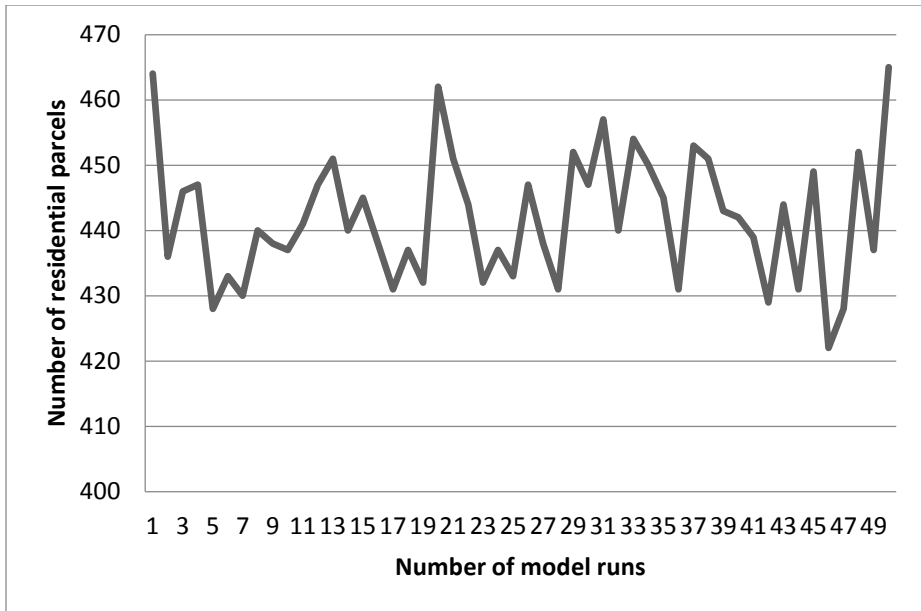
## APPENDIX II: Variation in Conversions across 50 Model Runs



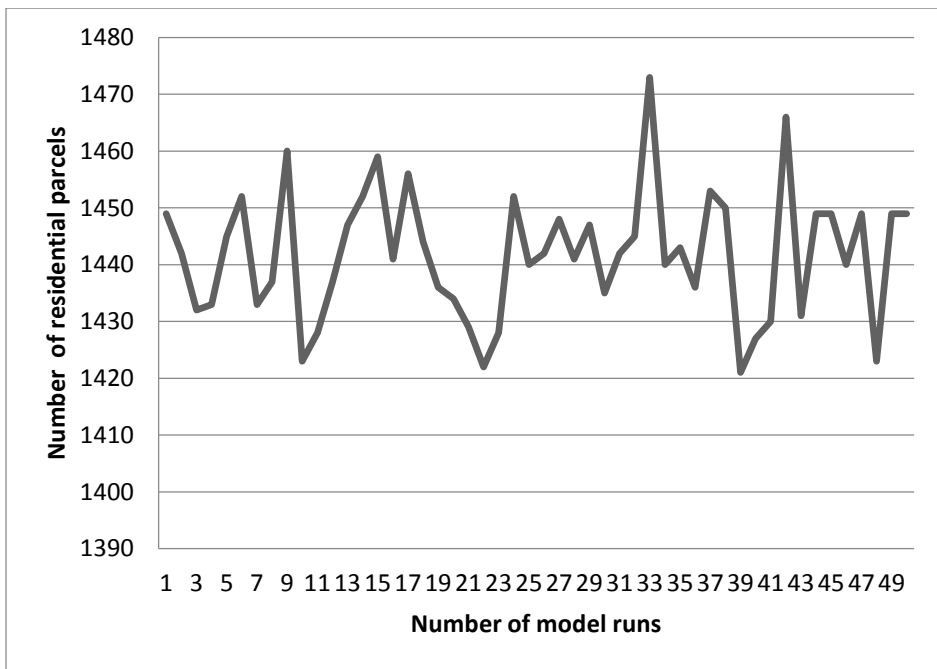
Stochastic variation in conversions within 50 model runs under baseline policy scenario



Stochastic variation in conversions within 50 model runs under 50 ft agricultural buffer (everywhere)

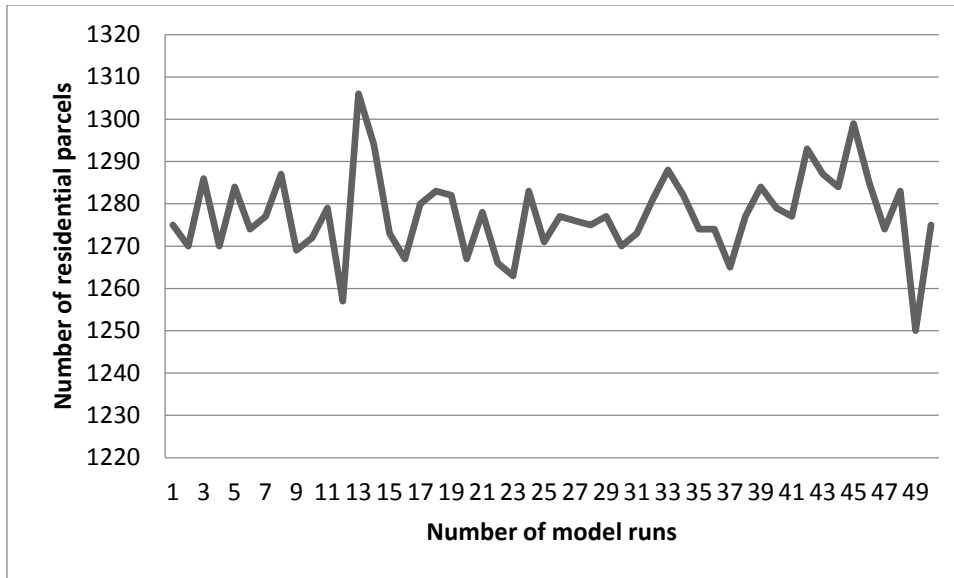


Stochastic variation in conversions within 50 model runs under 50 ft forest buffer (everywhere)



Stochastic variation in conversions within 50 model runs under 50 ft forest buffer in CSAs





Stochastic variation in conversions within 50 model runs under 50 ft agricultural buffer in CSAs

Land use in 2011 (Initial condition)

# of parcels	Residential Parcels	Agricultural Parcels	Forest Parcels
Total Parcels	20,852	8,957	7,311
Area in hectares	5,104.131	9,379.923	1,1532.452

Total area at first iterations = 35553.364 hectares

Number of parcels converted from each developable land use type

# of parcels	Baseline Policy	50 ft Agricultural Buffer (everywhere)	50 ft Forest Buffer (everywhere)	50 ft Agricultural Buffer in CSAs	50 ft Forest Buffer in CSAs
Total Parcels converted	1531	409	456	1263	1426
Agricultural to residential	1064	290	318	953	1062
Forest to residential	469	119	138	310	366

Area of land converted from each developable land use type

<b>Area(hectares)</b>	<b>Baseline Policy</b>	<b>50 ft Agricultural Buffer</b>	<b>50 ft Forest Buffer</b>	<b>50 ft Agricultural Buffer in CSAs</b>	<b>50 ft Forest Buffer in CSAs</b>
Total area converted	1451.26	46.89	52.74	1064.39	1172.72
Agricultural to residential	1208.81	37.58	41.74	759.87	809.39
Forest to residential	242.45	9.32	11.00	304.52	363.33

**APPENDIX III: Monthly Mean Sediment, Mean TP and Mean TN Loads for 42 Sub-basins Estimated by ArcSWAT over 10 years.**

**For land use 2011**

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
1	3.68	0.99	0.47
2	0.67	0.48	0.15
3	0.21	0.62	0.16
4	0.17	0.73	0.21
5	0.04	0.37	0.08
6	0.00	0.01	0.00
7	0.00	0.24	0.04
8	0.01	0.41	0.05
9	0.01	0.35	0.07
10	0.14	0.65	0.17
11	0.01	0.46	0.07
12	0.01	0.19	0.04
13	0.06	0.36	0.11
14	0.02	0.54	0.11
15	0.01	0.22	0.04
16	0.08	0.54	0.13
17	0.09	0.59	0.15
18	0.04	0.37	0.09
19	0.00	0.03	0.01
20	0.01	0.05	0.01
21	0.21	0.93	0.24
22	0.15	0.78	0.22
23	0.00	0.23	0.03
24	0.38	1.45	0.40
25	0.00	0.02	0.01
26	0.02	0.09	0.03
27	0.16	0.22	0.10
28	0.01	0.14	0.03
29	0.06	0.11	0.05
30	0.00	0.02	0.01
31	0.12	0.18	0.08
32	0.13	0.20	0.09
33	0.07	0.13	0.06
34	0.08	0.16	0.06

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
35	0.07	0.20	0.06
36	0.00	0.19	0.03
37	0.00	0.32	0.05
38	0.05	0.20	0.05
39	0.00	0.01	0.00
40	0.05	0.17	0.05
41	0.00	0.04	0.01
42	1.47	0.89	0.43

### **For Baseline Policy**

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
1	3.68	0.99	0.47
2	0.65	0.52	0.14
3	0.19	0.61	0.15
4	0.02	0.60	0.11
5	0.01	0.42	0.07
6	0.00	0.01	0.00
7	0.02	0.54	0.10
8	0.01	0.41	0.05
9	0.01	0.35	0.07
10	0.13	0.63	0.15
11	0.01	0.46	0.07
12	0.01	0.19	0.04
13	0.06	0.36	0.11
14	0.02	0.54	0.11
15	0.01	0.22	0.04
16	0.06	0.53	0.12
17	0.07	0.57	0.13
18	0.01	0.34	0.06
19	0.00	0.03	0.01
20	0.01	0.05	0.01
21	0.21	0.93	0.24
22	0.01	0.59	0.11
23	0.00	0.23	0.03
24	0.38	1.45	0.40
25	0.00	0.02	0.01
26	0.02	0.09	0.03
27	0.16	0.22	0.10

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
28	0.01	0.14	0.03
29	0.06	0.11	0.05
30	0.00	0.02	0.01
31	0.12	0.18	0.08
32	0.13	0.20	0.09
33	0.07	0.13	0.06
34	0.08	0.16	0.06
35	0.07	0.20	0.06
36	0.00	0.19	0.03
37	0.00	0.32	0.05
38	0.05	0.20	0.05
39	0.00	0.01	0.00
40	0.05	0.17	0.05
41	0.00	0.04	0.01
42	1.47	0.89	0.43

**For 50 ft buffer around all streams (agricultural buffer)**

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
1	3.68	0.99	0.47
2	0.67	0.48	0.15
3	0.21	0.62	0.16
4	0.17	0.73	0.21
5	0.04	0.37	0.08
6	0.00	0.01	0.00
7	0.00	0.24	0.04
8	0.01	0.41	0.05
9	0.13	0.58	0.17
10	0.14	0.65	0.17
11	0.04	0.51	0.10
12	0.02	0.21	0.04
13	0.06	0.36	0.11
14	0.02	0.54	0.11
15	0.02	0.25	0.05
16	0.08	0.55	0.14
17	0.09	0.60	0.15
18	0.05	0.38	0.09

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
19	0.00	0.03	0.01
20	0.01	0.05	0.01
21	0.21	0.93	0.24
22	0.15	0.78	0.22
23	0.00	0.23	0.03
24	0.38	1.45	0.40
25	0.00	0.02	0.01
26	0.02	0.09	0.03
27	0.16	0.22	0.10
28	0.01	0.14	0.03
29	0.06	0.11	0.05
30	0.00	0.02	0.01
31	0.12	0.18	0.08
32	0.13	0.20	0.09
33	0.07	0.13	0.06
34	0.08	0.16	0.06
35	0.08	0.21	0.07
36	0.00	0.19	0.03
37	0.07	0.38	0.11
38	0.06	0.21	0.06
39	0.00	0.01	0.00
40	0.06	0.18	0.06
41	0.00	0.04	0.01
42	1.47	0.89	0.43

**For 50 ft buffer around all streams (forest buffer)**

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
1	3.68	0.99	0.47
2	0.67	0.47	0.15
3	0.20	0.61	0.16
4	0.17	0.73	0.21
5	0.04	0.37	0.08
6	0.00	0.01	0.00
7	0.00	0.24	0.04
8	0.01	0.41	0.05
9	0.00	0.03	0.01
10	0.13	0.64	0.16
11	0.01	0.38	0.06

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
12	0.01	0.17	0.03
13	0.06	0.36	0.11
14	0.02	0.54	0.11
15	0.01	0.20	0.04
16	0.08	0.54	0.13
17	0.08	0.58	0.14
18	0.04	0.35	0.08
19	0.00	0.03	0.01
20	0.01	0.05	0.01
21	0.21	0.93	0.24
22	0.15	0.78	0.22
23	0.00	0.23	0.03
24	0.38	1.45	0.40
25	0.00	0.02	0.01
26	0.02	0.09	0.03
27	0.16	0.22	0.10
28	0.01	0.14	0.03
29	0.06	0.11	0.05
30	0.00	0.02	0.01
31	0.12	0.18	0.08
32	0.13	0.20	0.09
33	0.07	0.13	0.06
34	0.08	0.13	0.06
35	0.07	0.12	0.05
36	0.00	0.19	0.03
37	0.00	0.01	0.00
38	0.05	0.13	0.04
39	0.00	0.01	0.00
40	0.05	0.12	0.04
41	0.00	0.04	0.01
42	1.47	0.89	0.43

**For 50 ft buffer in CSAs streams (agricultural buffer)**

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
1	0.03	0.55	0.10
2	0.02	0.57	0.11
3	0.12	0.63	0.16
4	0.02	0.60	0.11
5	0.02	0.58	0.11
6	0.00	0.01	0.00
7	0.02	0.54	0.10
8	0.01	0.41	0.05
9	0.01	0.35	0.07
10	0.14	0.65	0.17
11	0.01	0.46	0.07
12	0.01	0.19	0.04
13	0.06	0.36	0.11
14	0.02	0.54	0.11
15	0.01	0.22	0.04
16	0.08	0.54	0.13
17	0.09	0.59	0.15
18	0.04	0.37	0.09
19	0.00	0.03	0.01
20	0.01	0.05	0.01
21	0.21	0.93	0.24
22	0.15	0.78	0.22
23	0.00	0.23	0.03
24	0.38	1.45	0.40
25	0.00	0.02	0.01
26	0.02	0.09	0.03
27	0.16	0.22	0.10
28	0.01	0.14	0.03
29	0.06	0.11	0.05
30	0.00	0.02	0.01
31	0.12	0.18	0.08
32	0.13	0.20	0.09
33	0.07	0.13	0.06
34	0.08	0.16	0.06
35	0.07	0.20	0.06
36	0.00	0.19	0.03
37	0.00	0.32	0.05
38	0.05	0.20	0.05



<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
39	0.00	0.01	0.00
40	0.05	0.17	0.05
41	0.00	0.04	0.01
42	1.47	0.89	0.43

**For 50 ft buffer in CSAs streams (forest buffer)**

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
1	0.03	0.55	0.10
2	0.02	0.57	0.11
3	0.12	0.63	0.16
4	0.02	0.60	0.11
5	0.02	0.58	0.11
6	0.00	0.01	0.00
7	0.02	0.54	0.10
8	0.01	0.41	0.05
9	0.01	0.35	0.07
10	0.14	0.65	0.17
11	0.01	0.46	0.07
12	0.01	0.19	0.04
13	0.06	0.36	0.11
14	0.02	0.54	0.11
15	0.01	0.22	0.04
16	0.08	0.54	0.13
17	0.09	0.59	0.15
18	0.04	0.37	0.09
19	0.00	0.03	0.01
20	0.01	0.05	0.01
21	0.21	0.93	0.24
22	0.15	0.78	0.22
23	0.00	0.23	0.03
24	0.38	1.45	0.40
25	0.00	0.02	0.01
26	0.02	0.09	0.03
27	0.16	0.22	0.10
28	0.01	0.14	0.03
29	0.06	0.11	0.05
30	0.00	0.02	0.01
31	0.12	0.18	0.08

<b>Sub-basin</b>	<b>Mean Sediment (tons/ha)</b>	<b>Mean TN (kg/ha)</b>	<b>Mean TP (kg/ha)</b>
32	0.13	0.20	0.09
33	0.07	0.13	0.06
34	0.08	0.13	0.06
35	0.07	0.18	0.06
36	0.00	0.19	0.03
37	0.00	0.32	0.05
38	0.05	0.18	0.05
39	0.00	0.01	0.00
40	0.05	0.16	0.05
41	0.00	0.04	0.01
42	1.47	0.89	0.43

## APPENDIX IV: Land Use, Soil, and Slope from ArcSWAT

Watershed level land use, soil and slope report from ArcSWAT with land use 2011 data

	Area [ha]	Area[acres]	
<b>Watershed</b>	35,635.41	88,056.88	
	Area [ha]	Area[acres]	%Wat.Area
<b>LAND USE:</b>			
Agriculture	9577.98	23667.67	26.88
Forest	14434.47	35668.30	40.51
Open Space	10345.50	25564.25	29.03
Residential	1277.46	3156.67	3.58
<b>SOILS (Mukey):</b>			
516224	4015.80	9923.24	11.27
516261	16.56	40.92	0.05
516293	6310.44	15593.41	17.71
516252	8487.63	20973.36	23.82
516244	8144.64	20125.81	22.86
516290	608.49	1503.61	1.71
516219	27.36	67.61	0.08
516282	1654.38	4088.06	4.64
516249	1528.83	3777.82	4.29
516279	2273.76	5618.57	6.38
516223	2038.86	5038.13	5.72
516236	528.66	1306.35	1.48
<b>SLOPE:</b>			
0-5	25638.30	63353.52	71.95
10-20	13.86	34.25	0.04
0-5	9983.25	24669.11	28.01

Sub-basin level land use, soil and slope report from ArcSWAT with land use 2011 data

<b>Sub-basin(HRU)</b>	<b>Landuse/Soil(Mukey)/Slope</b>	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area %Sub.Area</b>
1	Agriculture/516224/5-10	491.22	1213.83	1.38
2	Agriculture/516261/5-10	16.56	40.92	0.05
3	Forest/516293/5-10	231.75	572.67	0.65
4	Agriculture/516252/5-10	501.57	1239.40	1.41
5	Open Space/516252/5-10	1198.71	2962.07	3.36
6	Forest/516244/5-10	1001.34	2474.36	2.81
7	Open Space/516244/5-10	625.77	1546.31	1.76
8	Open Space/516244/5-10	683.73	1689.53	1.92
9	Open Space/516244/5-10	458.64	1133.32	1.29
10	Agriculture/516293/5-10	1968.21	4863.55	5.52
11	Residential/516252/5-10	655.11	1618.81	1.84
12	Residential/516290/5-10	122.40	302.46	0.34
13	Agriculture/516244/5-10	1052.46	2600.68	2.95
14	Residential/516290/5-10	486.09	1201.15	1.36
15	Residential/516293/10-20	13.86	34.25	0.04
16	Open Space/516293/5-10	569.97	1408.42	1.60
17	Open Space/516219/5-10	27.36	67.61	0.08
18	Open Space/516293/5-10	601.65	1486.71	1.69
19	Forest/516244/5-10	3217.50	7950.60	9.03
20	Forest/516282/5-10	826.65	2042.69	2.32
21	Open Space/516224/5-10	962.37	2378.06	2.70
22	Agriculture/516252/5-10	2161.80	5341.92	6.07
23	Open Space/516249/5-10	1528.83	3777.82	4.29
24	Agriculture/516293/5-10	2925.00	7227.82	8.21
25	Forest/516252/5-10	525.33	1298.12	1.47
26	Forest/516279/5-10	1792.44	4429.21	5.03
27	Forest/516223/0-5	708.03	1749.58	1.99
28	Open Space/516252/0-5	1095.12	2706.10	3.07
29	Forest/516224/0-5	495.00	1223.17	1.39
30	Forest/516236/0-5	528.66	1306.35	1.48
31	Forest/516279/5-10	481.32	1189.37	1.35
32	Forest/516223/0-5	1330.83	3288.55	3.73
33	Forest/516224/0-5	1081.53	2672.51	3.03
34	Open Space/516252/0-5	441.99	1092.18	1.24
35	Open Space/516252/0-5	161.01	397.86	0.45
36	Open Space/516244/0-5	873.45	2158.34	2.45
37	Open Space/516252/0-5	885.15	2187.25	2.48
38	Open Space/516244/0-5	231.75	572.67	0.65

Sub-basin(HRU)	Landuse/Soil(Mukey)/Slope	Area [ha]	Area[acres]	%Wat.Area %Sub.Area
39	Forest/516252/0-5	861.84	2129.65	2.42
40	Forest/516224/5-10	524.52	1296.12	1.47
41	Forest/516282/0-5	827.73	2045.36	2.32
42	Agriculture/516224/0-5	461.16	1139.55	1.29

Watershed level land use, soil and slope report from ArcSWAT under baseline

	Area [ha]	Area[acres]	
<b>Watershed</b>	35635.41	88056.88	
	Area [ha]	Area[acres]	%Wat.Area
<b>LAND USE:</b>			
Agriculture	6914.61	17086.35	19.40
Forest	14434.47	35668.30	40.51
Open Space	9118.08	22531.23	25.59
Residential	5168.25	12771.00	14.50
<b>SOILS (Mukey):</b>			
516224	4015.80	9923.24	11.27
516261	16.56	40.92	0.05
516293	6310.44	15593.41	17.71
516252	8487.63	20973.36	23.82
516244	8144.64	20125.81	22.86
516290	608.49	1503.61	1.71
516219	27.36	67.61	0.08
516282	1654.38	4088.06	4.64
516249	1528.83	3777.82	4.29
516279	2273.76	5618.57	6.38
516223	2038.86	5038.13	5.72
516236	528.66	1306.35	1.48
<b>SLOPE:</b>			
10-5	25638.30	63353.52	71.95
10-20	13.86	34.25	0.04
0-5	9983.25	24669.11	28.01

Sub-basin level land use, soil and slope report from ArcSWAT under baseline

<b>Sub-basin(HRU)</b>	<b>Landuse/soil(Mukey)/slope</b>	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area %Sub.Area</b>
1	Agriculture/516224/5-10	491.22	1213.83	1.38
2	Agriculture/516261/5-10	16.56	40.92	0.05
3	Forest/516293/5-10	231.75	572.67	0.65
4	Residential/516252/5-10	501.57	1239.40	1.41
5	Open Space/516252/5-10	1198.71	2962.07	3.36
6	Forest/516244/5-10	1001.34	2474.36	2.81
7	Residential/516244/5-10	625.77	1546.31	1.76
8	Open Space/516244/5-10	683.73	1689.53	1.92
9	Open Space/516244/5-10	458.64	1133.32	1.29
10	Agriculture/516293/5-10	1968.21	4863.55	5.52
11	Residential/516252/5-10	655.11	1618.81	1.84
12	Residential/516290/5-10	122.40	302.46	0.34
13	Agriculture/516244/5-10	1052.46	2600.68	2.95
14	Residential/516290/5-10	486.09	1201.15	1.36
15	Residential/516293/10-20	13.86	34.25	0.04
16	Open Space/516293/5-10	569.97	1408.42	1.60
17	Open Space/516219/5-10	27.36	67.61	0.08
18	Residential/516293/5-10	601.65	1486.71	1.69
19	Forest/516244/5-10	3217.50	7950.60	9.03
20	Forest/516282/5-10	826.65	2042.69	2.32
21	Open Space/516224/5-10	962.37	2378.06	2.70
22	Residential/516252/5-10	2161.80	5341.92	6.07
23	Open Space/516249/5-10	1528.83	3777.82	4.29
24	Agriculture/516293/5-10	2925.00	7227.82	8.21
25	Forest/516252/5-10	525.33	1298.12	1.47
26	Forest/516279/5-10	1792.44	4429.21	5.03
27	Forest/516223/0-5	708.03	1749.58	1.99
28	Open Space/516252/0-5	1095.12	2706.10	3.07
29	Forest/516224/0-5	495.00	1223.17	1.39
30	Forest/516236/0-5	528.66	1306.35	1.48
31	Forest/516279/5-10	481.32	1189.37	1.35
32	Forest/516223/0-5	1330.83	3288.55	3.73
33	Forest/516224/0-5	1081.53	2672.51	3.03
34	Open Space/516252/0-5	441.99	1092.18	1.24
35	Open Space/516252/0-5	161.01	397.86	0.45
36	Open Space/516244/0-5	873.45	2158.34	2.45
37	Open Space/516252/0-5	885.15	2187.25	2.48
38	Open Space/516244/0-5	231.75	572.67	0.65

Sub-basin(HRU)	Landuse/soil(Mukey)/slope	Area [ha]	Area[acres]	%Wat.Area %Sub.Area
39	Forest/516252/0-5	861.84	2129.65	2.42
40	Forest/516224/5-10	524.52	1296.12	1.47
41	Forest/516282/0-5	827.73	2045.36	2.32
42	Agriculture/516224/0-5	461.16	1139.55	1.29

Watershed level land use, soil and slope report from ArcSWAT under agricultural buffer (everywhere)

	Area [ha]	Area[acres]	
<b>Watershed</b>	35635.41	88056.88	
	Area [ha]	Area[acres]	% Wat.Area
<b>LAND USE:</b>			
Agriculture	10962.99	27090.10	30.76
Forest	14434.47	35668.30	40.51
Open Space	8974.35	22176.07	25.18
Residential	1263.60	3122.42	3.55
<b>SOILS (Mukey):</b>			
516224	4015.80	9923.24	11.27
516261	16.56	40.92	0.05
516293	6310.44	15593.41	17.71
516252	8487.63	20973.36	23.82
516244	8144.64	20125.81	22.86
516290	608.49	1503.61	1.71
516219	27.36	67.61	0.08
516282	1654.38	4088.06	4.64
516249	1528.83	3777.82	4.29
516279	2273.76	5618.57	6.38
516223	2038.86	5038.13	5.72
516236	528.66	1306.35	1.48
<b>SLOPE:</b>			
0-5	25638.30	63353.52	71.95
10-20	13.86	34.25	0.04
0-5	9983.25	24669.11	28.01

Sub-basin level land use, soil and slope report from ArcSWAT under agricultural buffer (everywhere)

<b>Sub-basin(HRU)</b>	<b>Land use/soil(Mukey)/slope</b>	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area %Sub.Area</b>
1	Agriculture/516224/5-10	491.22	1213.83	1.38
2	Agriculture/516261/5-10	16.56	40.92	0.05
3	Forest/516293/5-10	231.75	572.67	0.65
4	Agriculture/516252/5-10	501.57	1239.40	1.41
5	Open Space/516252/5-10	1198.71	2962.07	3.36
6	Forest/516244/5-10	1001.34	2474.36	2.81
7	Open Space/516244/5-10	625.77	1546.31	1.76
8	Open Space/516244/5-10	683.73	1689.53	1.92
9	Agriculture/516244/5-10	458.64	1133.32	1.29
10	Agriculture/516293/5-10	1968.21	4863.55	5.52
11	Residential/516252/5-10	655.11	1618.81	1.84
12	Residential/516290/5-10	122.40	302.46	0.34
13	Agriculture/516244/5-10	1052.46	2600.68	2.95
14	Residential/516290/5-10	486.09	1201.15	1.36
15	Agriculture/516293/10-20	13.86	34.25	0.04
16	Open Space/516293/5-10	569.97	1408.42	1.60
17	Agriculture/516219/5-10	27.36	67.61	0.08
18	Open Space/516293/5-10	601.65	1486.71	1.69
19	Forest/516244/5-10	3217.50	7950.60	9.03
20	Forest/516282/5-10	826.65	2042.69	2.32
21	Open Space/516224/5-10	962.37	2378.06	2.70
22	Agriculture/516252/5-10	2161.80	5341.92	6.07
23	Open Space/516249/5-10	1528.83	3777.82	4.29
24	Agriculture/516293/5-10	2925.00	7227.82	8.21
25	Forest/516252/5-10	525.33	1298.12	1.47
26	Forest/516279/5-10	1792.44	4429.21	5.03
27	Forest/516223/0-5	708.03	1749.58	1.99
28	Open Space/516252/0-5	1095.12	2706.10	3.07
29	Forest/516224/0-5	495.00	1223.17	1.39
30	Forest/516236/0-5	528.66	1306.35	1.48
31	Forest/516279/5-10	481.32	1189.37	1.35
32	Forest/516223/0-5	1330.83	3288.55	3.73
33	Forest/516224/0-5	1081.53	2672.51	3.03
34	Open Space/516252/0-5	441.99	1092.18	1.24
35	Open Space/516252/0-5	161.01	397.86	0.45
36	Open Space/516244/0-5	873.45	2158.34	2.45
37	Agriculture/516252/0-5	885.15	2187.25	2.48
38	Open Space/516244/0-5	231.75	572.67	0.65



Sub-basin(HRU)	Land use/soil(Mukey)/slope	Area [ha]	Area[acres]	%Wat.Area %Sub.Area
39	Forest/516252/0-5	861.84	2129.65	2.42
40	Forest/516224/5-10	524.52	1296.12	1.47
41	Forest/516282/0-5	827.73	2045.36	2.32
42	Agriculture/516224/0-5	461.16	1139.55	1.29

Watershed level land use, soil and slope report from ArcSWAT under forest buffer (everywhere)

	Area [ha]	Area[acres]	
<b>Watershed</b>	35635.41	88056.88	
	Area [ha]	Area[acres]	% Wat.Area
<b>LAND USE:</b>			
Agriculture	9561.42	23626.75	26.83
Forest	16250.67	40156.22	45.60
Open Space	8559.72	21151.50	24.02
Residential	1263.6	3122.42	3.55
<b>SOILS (Mukey):</b>			
516224	4015.8	9923.24	11.27
516261	16.56	40.92	0.05
516293	6310.44	15593.41	17.71
516252	8487.63	20973.36	23.82
516244	8144.64	20125.81	22.86
516290	608.49	1503.61	1.71
516219	27.36	67.61	0.08
516282	1654.38	4088.06	4.64
516249	1528.83	3777.82	4.29
516279	2273.76	5618.57	6.38
516223	2038.86	5038.13	5.72
516236	528.66	1306.35	1.48
<b>SLOPE:</b>			
0-5	25638.3	63353.52	71.95
10-20	13.86	34.25	0.04
0-5	9983.25	24669.11	28.01

Sub-basin level land use, soil and slope report from ArcSWAT under forest buffer (everywhere)

<b>Sub-basin(HRU)</b>	<b>Land use/soil(Mukey)/slope</b>	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area %Sub.Area</b>
1	Agriculture/516224/5-10	491.22	1213.83	1.38
2	Forest/516261/5-10	16.56	40.92	0.05
3	Forest/516293/5-10	231.75	572.67	0.65
4	Agriculture/516252/5-10	501.57	1239.40	1.41
5	Open Space/516252/5-10	1198.71	2962.07	3.36
6	Forest/516244/5-10	1001.34	2474.36	2.81
7	Open Space/516244/5-10	625.77	1546.31	1.76
8	Open Space/516244/5-10	683.73	1689.53	1.92
9	Forest/516244/5-10	458.64	1133.32	1.29
10	Agriculture/516293/5-10	1968.21	4863.55	5.52
11	Residential/516252/5-10	655.11	1618.81	1.84
12	Residential/516290/5-10	122.4	302.46	0.34
13	Agriculture/516244/5-10	1052.46	2600.68	2.95
14	Residential/516290/5-10	486.09	1201.15	1.36
15	Forest/516293/10-20	13.86	34.25	0.04
16	Open Space/516293/5-10	569.97	1408.42	1.6
17	Open Space/516219/5-10	27.36	67.61	0.08
18	Open Space/516293/5-10	601.65	1486.71	1.69
19	Forest/516244/5-10	3217.5	7950.60	9.03
20	Forest/516282/5-10	826.65	2042.69	2.32
21	Open Space/516224/5-10	962.37	2378.06	2.7
22	Agriculture/516252/5-10	2161.8	5341.92	6.07
23	Open Space/516249/5-10	1528.83	3777.82	4.29
24	Agriculture/516293/5-10	2925	7227.82	8.21
25	Forest/516252/5-10	525.33	1298.12	1.47
26	Forest/516279/5-10	1792.44	4429.21	5.03
27	Forest/516223/0-5	708.03	1749.58	1.99
28	Open Space/516252/0-5	1095.12	2706.10	3.07
29	Forest/516224/0-5	495	1223.17	1.39
30	Forest/516236/0-5	528.66	1306.35	1.48
31	Forest/516279/5-10	481.32	1189.37	1.35
32	Forest/516223/0-5	1330.83	3288.55	3.73
33	Forest/516224/0-5	1081.53	2672.51	3.03
34	Forest/516252/0-5	441.99	1092.18	1.24
35	Open Space/516252/0-5	161.01	397.86	0.45
36	Open Space/516244/0-5	873.45	2158.34	2.45
37	Forest/516252/0-5	885.15	2187.25	2.48
38	Open Space/516244/0-5	231.75	572.67	0.65

Sub-basin(HRU)	Land use/soil(Mukey)/slope	Area [ha]	Area[acres]	%Wat.Area %Sub.Area
39	Forest/516252/0-5	861.84	2129.65	2.42
40	Forest/516224/5-10	524.52	1296.12	1.47
41	Forest/516282/0-5	827.73	2045.36	2.32
42	Agriculture/516224/0-5	461.16	1139.55	1.29

Watershed level land use, soil and slope report from ArcSWAT under agricultural buffer in CSAs

	Area [ha]	Area[acres]	
<b>Watershed</b>	35635.41	88056.88	
	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area</b>
<b>LAND USE:</b>			
Agriculture	8585.19	21214.43	24.09
Forest	14434.47	35668.30	40.51
Open Space	8521.02	21055.87	23.91
Residential	4094.73	10118.28	11.49
<b>SOILS (Mukey):</b>			
516224.00	4015.80	9923.24	11.27
516261.00	16.56	40.92	0.05
516293.00	6310.44	15593.41	17.71
516252.00	8487.63	20973.36	23.82
516244.00	8144.64	20125.81	22.86
516290.00	608.49	1503.61	1.71
516219.00	27.36	67.61	0.08
516282.00	1654.38	4088.06	4.64
516249.00	1528.83	3777.82	4.29
516279.00	2273.76	5618.57	6.38
516223.00	2038.86	5038.13	5.72
516236.00	528.66	1306.35	1.48
<b>SLOPE:</b>			
5-10	25638.30	63353.52	71.95
10-20	13.86	34.25	0.04
0-5	9983.25	24669.11	28.01

Sub-basin level land use, soil and slope report from ArcSWAT under agricultural buffer in CSAs

Sub-basin(HRU)	Landuse/soil(Mukey)/slope	Area [ha]	Area[acres]	%Wat.Area %Sub.Area
1	Residential/516224/5-10	491.22	1213.83	1.38
2	Agriculture/516261/5-10	16.56	40.92	0.05
3	Forest/516293/5-10	231.75	572.67	0.65
4	Residential/516252/5-10	501.57	1239.40	1.41
5	Residential/516252/5-10	1198.71	2962.07	3.36
6	Forest/516244/5-10	1001.34	2474.36	2.81
7	Residential/516244/5-10	625.77	1546.31	1.76
8	Open Space/516244/5-10	683.73	1689.53	1.92
9	Open Space/516244/5-10	458.64	1133.32	1.29
10	Agriculture/516293/5-10	1968.21	4863.55	5.52
11	Residential/516252/5-10	655.11	1618.81	1.84
12	Residential/516290/5-10	122.40	302.46	0.34
13	Agriculture/516244/5-10	1052.46	2600.68	2.95
14	Residential/516290/5-10	486.09	1201.15	1.36
15	Residential/516293/10-20	13.86	34.25	0.04
16	Open Space/516293/5-10	569.97	1408.42	1.60
17	Open Space/516219/5-10	27.36	67.61	0.08
18	Open Space/516293/5-10	601.65	1486.71	1.69
19	Forest/516244/5-10	3217.50	7950.60	9.03
20	Forest/516282/5-10	826.65	2042.69	2.32
21	Open Space/516224/5-10	962.37	2378.06	2.70
22	Agriculture/516252/5-10	2161.80	5341.92	6.07
23	Open Space/516249/5-10	1528.83	3777.82	4.29
24	Agriculture/516293/5-10	2925.00	7227.82	8.21
25	Forest/516252/5-10	525.33	1298.12	1.47
26	Forest/516279/5-10	1792.44	4429.21	5.03
27	Forest/516223/0-5	708.03	1749.58	1.99
28	Open Space/516252/0-5	1095.12	2706.10	3.07
29	Forest/516224/0-5	495.00	1223.17	1.39
30	Forest/516236/0-5	528.66	1306.35	1.48
31	Forest/516279/5-10	481.32	1189.37	1.35
32	Forest/516223/0-5	1330.83	3288.55	3.73
33	Forest/516224/0-5	1081.53	2672.51	3.03
34	Open Space/516252/0-5	441.99	1092.18	1.24
35	Open Space/516252/0-5	161.01	397.86	0.45
36	Open Space/516244/0-5	873.45	2158.34	2.45
37	Open Space/516252/0-5	885.15	2187.25	2.48
38	Open Space/516244/0-5	231.75	572.67	0.65

Sub-basin(HRU)	Landuse/soil(Mukey)/slope	Area [ha]	Area[acres]	%Wat.Area %Sub.Area
39	Forest/516252/0-5	861.84	2129.65	2.42
40	Forest/516224/5-10	524.52	1296.12	1.47
41	Forest/516282/0-5	827.73	2045.36	2.32
42	Agriculture/516224/0-5	461.16	1139.55	1.29

Watershed level land use, soil and slope report from ArcSWAT under forest buffer in CSAs

	Area [ha]	Area[acres]	
<b>Watershed</b>	35635.41	88056.88	
	Area [ha]	Area[acres]	%Wat.Area
<b>LAND USE:</b>			
Agriculture	8568.63	21173.51	24.05
Forest	14893.02	36801.40	41.79
Open Space	8079.03	19963.69	22.67
Residential	4094.73	10118.28	11.49
<b>SOILS (Mukey):</b>			
516224.00	4015.80	9923.24	11.27
516261.00	16.56	40.92	0.05
516293.00	6310.44	15593.41	17.71
516252.00	8487.63	20973.36	23.82
516244.00	8144.64	20125.81	22.86
516290.00	608.49	1503.61	1.71
516219.00	27.36	67.61	0.08
516282.00	1654.38	4088.06	4.64
516249.00	1528.83	3777.82	4.29
516279.00	2273.76	5618.57	6.38
516223.00	2038.86	5038.13	5.72
516236.00	528.66	1306.35	1.48
<b>SLOPE:</b>			
5-10	25638.30	63353.52	71.95
10-20	13.86	34.25	0.04
0-5	9983.25	24669.11	28.01

Sub-basin level land use, soil and slope report from ArcSWAT under forest buffer in CSAs

<b>Sub-basin(HRU)</b>	<b>Landuse/soil(Mukey)/slope</b>	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area %Sub.Area</b>
1	Residential/516224/5-10	491.22	1213.83	1.38
2	Forest/516261/5-10	16.56	40.92	0.05
3	Forest/516293/5-10	231.75	572.67	0.65
4	Residential/516252/5-10	501.57	1239.40	1.41
5	Residential/516252/5-10	1198.71	2962.07	3.36
6	Forest/516244/5-10	1001.34	2474.36	2.81
7	Residential/516244/5-10	625.77	1546.31	1.76
8	Open Space/516244/5-10	683.73	1689.53	1.92
9	Open Space/516244/5-10	458.64	1133.32	1.29
10	Agriculture/516293/5-10	1968.21	4863.55	5.52
11	Residential/516252/5-10	655.11	1618.81	1.84
12	Residential/516290/5-10	122.40	302.46	0.34
13	Agriculture/516244/5-10	1052.46	2600.68	2.95
14	Residential/516290/5-10	486.09	1201.15	1.36
15	Residential/516293/10-20	13.86	34.25	0.04
16	Open Space/516293/5-10	569.97	1408.42	1.60
17	Open Space/516219/5-10	27.36	67.61	0.08
18	Open Space/516293/5-10	601.65	1486.71	1.69
19	Forest/516244/5-10	3217.50	7950.60	9.03
20	Forest/516282/5-10	826.65	2042.69	2.32
21	Open Space/516224/5-10	962.37	2378.06	2.70
22	Agriculture/516252/5-10	2161.80	5341.92	6.07
23	Open Space/516249/5-10	1528.83	3777.82	4.29
24	Agriculture/516293/5-10	2925.00	7227.82	8.21
25	Forest/516252/5-10	525.33	1298.12	1.47
26	Forest/516279/5-10	1792.44	4429.21	5.03
27	Forest/516223/0-5	708.03	1749.58	1.99
28	Open Space/516252/0-5	1095.12	2706.10	3.07
29	Forest/516224/0-5	495.00	1223.17	1.39
30	Forest/516236/0-5	528.66	1306.35	1.48
31	Forest/516279/5-10	481.32	1189.37	1.35
32	Forest/516223/0-5	1330.83	3288.55	3.73
33	Forest/516224/0-5	1081.53	2672.51	3.03
34	Forest/516252/0-5	441.99	1092.18	1.24
35	Open Space/516252/0-5	161.01	397.86	0.45
36	Open Space/516244/0-5	873.45	2158.34	2.45
37	Open Space/516252/0-5	885.15	2187.25	2.48
38	Open Space/516244/0-5	231.75	572.67	0.65

<b>Sub-basin(HRU)</b>	<b>Landuse/soil(Mukey)/slope</b>	<b>Area [ha]</b>	<b>Area[acres]</b>	<b>%Wat.Area %Sub.Area</b>
39	Forest/516252/0-5	861.84	2129.65	2.42
40	Forest/516224/5-10	524.52	1296.12	1.47
41	Forest/516282/0-5	827.73	2045.36	2.32
42	Agriculture/516224/0-5	461.16	1139.55	1.29

## **APPENDIX V: Requirements by Berkeley County Planning Commission, 2009**

### **Section 402.5.5 Stream Buffers**

Where an application contains any portion of a perennial or intermittent stream, a buffer shall be established and shown on the plan. The width of the buffer shall be a minimum of thirty-five feet (35'), measured from and perpendicular to the top of the stream bank. The buffer shall be expanded to include any floodplain determined by the Floodplain Ordinance, any field verified non-tidal wetland areas and/or any area of severe slope as defined in this Ordinance.

Within the stream buffer vegetative ground cover shall be maintained at all times. The U.S. National Resources Conservation Service (NRCS) may recommend planting species and methods when no ground cover exists in the buffer or additional planting to improve existing ground cover. No permanently affixed building shall be permitted within the stream buffer except those designed to improve water quality in the stream or structures such as fences designed to limit access to the stream. No septic system shall be constructed within the buffer nor shall any septic reserve area be established within the buffer.



## APPENDIX VI: Code for the APLUC in Python using ArcPy (ESRI, 2014) (Baseline scenario)

```
import arcpy
from arcpy import env
from arcpy.sa import *
import random
import math

env.workspace = "C:/data"

# Set the extent environment.
arcpy.env.outputCoordinateSystem = arcpy.SpatialReference("NAD 1983 UTM Zone 17N")
arcpy.env.extent = "C:/data/dem"
arcpy.env.snapRaster = "C:/data/dem"
arcpy.env.mask = "C:/data/dem"
arcpy.env.cellSize = "C:/data/dem"
arcpy.env.overwriteOutput = True

for run in range(100):
    inFeatures = "parcel.shp"
    outLocation = "C:/data"
    outFeatureClass = "resid_zonal.shp"
    # Execute FeatureClassToFeatureClass
    arcpy.FeatureClassToFeatureClass_conversion(inFeatures, outLocation, outFeatureClass)

    for iteration in range(10):
        random.seed()
        print "starting iteration " + str(iteration)

#-----

# Set local variables(distance measurement from agricultural land use)

in_features = "resid_zonal.shp"
where_clause = "'landuse' = 6"
out_feature_class = "C:/data/agParcels.shp"
# Execute Select
arcpy.Select_analysis(in_features, out_feature_class, where_clause)
outDirectionRaster = "C:/data/eucdirect"

# Set local variables
inSourceData = "agParcels.shp"
maxDistance = 25000
cellSize = 30
outDirectionRaster = "C:/data/eucdirect"
# Check out the ArcGIS Spatial Analyst extension license
arcpy.CheckOutExtension("Spatial")
```

```

# Execute EucDistance
outEucDistance = EucDistance(inSourceData, maxDistance, cellSize, outDirectionRaster)

# Save the output
outEucDistance.save("C:/data/eucdist")

#Set local variables
inZoneData = "resid_zonal.shp"
zoneField = "FID"
inValueRaster = "eucdist"

# Execute ZonalStatistics
outZonalStatistics = ZonalStatistics(inZoneData, zoneField, inValueRaster, "MEAN")

# Save the output
outZonalStatistics.save("C:/data/zonestatout")

outTimes = Raster("zonestatout")
outTimes.save("C:/data/int_agri")
inRaster = "int_agri"

# Execute Int
outInt = Int(inRaster)

# Save the output
outInt.save("C:/data/outint")
inRaster = "outint"
outPolygons = "c:/data/int_agri1.shp"
field = "VALUE"

# Execute RasterToPolygon
arcpy.RasterToPolygon_conversion(inRaster, outPolygons, "NO_SIMPLIFY", field)
# Save the output
arcpy.SpatialJoin_analysis("resid_zonal.shp", "int_agri1.shp", "agric_zonal")

#-----

# Set local variables (distance measurement from forest land use)
in_features = "resid_zonal.shp"
where_clause = "'landuse' = 5"
out_feature_class = "C:/data/forParcels.shp"
# Execute Select
arcpy.Select_analysis(in_features, out_feature_class, where_clause)
outDirectionRaster = "C:/data/eucdirect1"

```

```

# Set local variables
inSourceData = "forParcels.shp"
maxDistance = 25000
cellSize = 30
outDirectionRaster = "C:/data/eucdirect1"
# Check out the ArcGIS Spatial Analyst extension license
arcpy.CheckOutExtension("Spatial")

# Execute EucDistance
outEucDistance = EucDistance(inSourceData, maxDistance, cellSize, outDirectionRaster)

# Save the output
outEucDistance.save("C:/data/eucdist1")

#Set local variables
inZoneData = "resid_zonal.shp"
zoneField = "FID"
inValueRaster = "eucdist1"

# Execute ZonalStatistics
outZonalStatistics = ZonalStatistics(inZoneData, zoneField, inValueRaster, "MEAN")

# Save the output
outZonalStatistics.save("C:/data/zonestatout1")

outTimes = Raster("zonestatout1")
outTimes.save("C:/data/int_for")
inRaster = "int_for"

# Execute Int
outInt = Int(inRaster)

# Save the output
outInt.save("C:/data/outint1")
inRaster = "outint1"
outPolygons = "c:/data/int_for1.shp"
field = "VALUE"

# Execute RasterToPolygon
arcpy.RasterToPolygon_conversion(inRaster, outPolygons, "NO_SIMPLIFY", field)
# Save the output
arcpy.SpatialJoin_analysis("agric_zonal.shp", "int_for1.shp", "for_zonal")

#-----

# Set local variables (distance measurement from residential land use)

in_features = "resid_zonal.shp"

```

```

where_clause = "'landuse' = 3'
out_feature_class = "C:/data/residParcels.shp"
# Execute Select
arcpy.Select_analysis(in_features, out_feature_class, where_clause)
outDirectionRaster = "C:/data/eucdirect2"

# Set local variables
inSourceData = "residParcels.shp"
maxDistance = 25000
cellSize = 30
outDirectionRaster = "C:/data/eucdirect2"
# Check out the ArcGIS Spatial Analyst extension license
arcpy.CheckOutExtension("Spatial")

# Execute EucDistance
outEucDistance = EucDistance(inSourceData, maxDistance, cellSize, outDirectionRaster)

# Save the output
outEucDistance.save("C:/data/eucdist2")

#Set local variables
inZoneData = "resid_zonal.shp"
zoneField = "FID"
inValueRaster = "eucdist2"

# Execute ZonalStatistics
outZonalStatistics = ZonalStatistics(inZoneData, zoneField, inValueRaster, "MEAN")

# Save the output
outZonalStatistics.save("C:/data/zonestatout2")

outTimes = Raster("zonestatout2")
outTimes.save("C:/data/int_resid")
inRaster = "int_resid"

# Execute Int
outInt = Int(inRaster)

# Save the output
outInt.save("C:/data/outint2")
inRaster = "outint2"
outPolygons = "c:/data/int_res.shp"
field = "VALUE"

# Execute RasterToPolygon
arcpy.RasterToPolygon_conversion(inRaster, outPolygons, "NO_SIMPLIFY", field)
# Save the output
arcpy.SpatialJoin_analysis("for_zonal.shp", "int_res.shp", "resid_zonal")

```

```

#-----
# Implementation of coefficients from SLR model

beta0 = -2.27300013 # intercept
beta1 = -0.00044498 #distance from streams
beta2 = -0.00012974 #distance from highways
beta3 = -0.00010346 # distance from urban center
beta4 = 0.00336320 #distance from agricultural lands
beta5 = 0.00396040 # distance from forest lands
beta6 = -0.00266703 # distance from residential lands

# To observe each iteration and resulting land use outcome
inSourceData = ("C:/data/resid_zonal.shp")
inFeatures = "resid_zonal.shp"
fieldName = "luR" + str(run) + "I" + str(iteration)
fieldPrecision = 0
fieldAlias = "landUse" + str(iteration)

# Add fields
print "the field name is " + fieldName
arcpy.AddField_management(inSourceData, fieldName, "SHORT", fieldPrecision, "",
"",fieldAlias, "NULLABLE")

cursor = arcpy.UpdateCursor(inSourceData)

convertTotal = 0
for row in cursor:

    if (row.getValue("landUse") == 5 or row.getValue("landUse") == 6):
        # use distances and logistic regression equation to get conversionProb
        x1 = row.getValue("dis_strea")
        x2 = row.getValue("dis_roads")
        x3 = row.getValue("dis_urban")
        x4 = row.getValue("GRIDCODE") # distance from agric
        x5 = row.getValue("GRIDCODE_1") # distance from forest
        x6 = row.getValue("GRIDCODE_2") # distance from resid

        exponent = beta0+beta1*x1+beta2*x2+beta3*x3+beta4*x4+beta5*x5+beta6*x6
        if exponent > 100:
            exponent = 100
        conversionProb = (math.exp(exponent))/(1+math.exp(exponent))
        # stepwise conversion probability
        stepwiseConversionProb = 1-math.pow(1-conversionProb,0.1)

        row.setValue("prob", conversionProb)

```

```

cursor.updateRow(row)

#if convert this parcel

if random.random() < stepwiseConversionProb:
    convertTotal = convertTotal + 1
    row.setValue("landUse", 3)
    cursor.updateRow(row)

row.setValue(fieldName, row.getValue("landuse"))

cursor.updateRow(row)

# another way to check the conversion in each iteration
print "I converted " + str(convertTotal) + " parcels in this step."
inFeatures = "resid_zonal.shp"
outLocation = "C:/data"
outFeatureClass = "model" + str(run) + ".shp"
# Execute FeatureClassToFeatureClass
arcpy.FeatureClassToFeatureClass_conversion(inFeatures, outLocation, outFeatureClass)

#delete the GRIDCODE, GRIDCODE_1, and GRIDCODE_2 fields
arcpy.DeleteField_management(inFeatures, ["GRIDCODE", "GRIDCODE_1", "GRIDCODE_2"])

del row
del cursor

```