CrossMark

Projections of Watershed Pollutant Loads Using a Spatially Explicit, Agent-Based Land Use Conversion Model: A Case Study of Berkeley County, West Virginia

Nazia N. Arbab[1] · **Alan R. Collins**[2] ·
Jamison F. Conley[3]

Abstract This research presents a methodology to make projections of land use conversions in Berkeley County, West Virginia and then utilizes these projections to estimate water quality impacts on the Opequon Creek in Berkeley County. Empirical estimates for factors that influence the land use conversion probability are captured using parameters from a spatial logistic regression (SLR) model. Then, an agent-based, probabilistic land use conversion (APLUC) model is used to explore the impacts of policies on land use conversion decisions using estimates from actual land use change from 2001 to 2011 in SLR model. Three policy scenarios are developed: (1) no policy implementation, (2) a 15.24 m (50 ft) buffer zone policy of no development applied to all streams, and (3) 15.24 m buffer policy applied only on critical source area (CSA) watersheds. The projected land use patterns in the APLUC model are driven by individual land conversion decisions over 50 model runs of 10 iterations each under each policy scenario. The results

✉ Nazia N. Arbab
nazia.arbab@rutgers.edu

Alan R. Collins
acollins@mail.wvu.edu

Jamison F. Conley
Jamison.Conley@mail.wvu.edu

[1] Center for Resilient Landscapes, Department of Ecology, Evolution, and Natural Resources, School of Environmental and Biological Sciences, Rutgers, The State University of New Jersey, Room 109, 80 Nichol Avenue, New Brunswick, NJ 08901, USA

[2] Agricultural and Resource Economics Program, West Virginia University, Agricultural Sciences Building. Room 2026, Morgantown, WV 26506, USA

[3] Department of Geology and Geography, West Virginia University, 347 Brooks Hall, 98 Beechurst Ave, Morgantown, WV 26506, USA

show that with no policy scenario, most conversions occurred near existing residential land use and urban centers. Residential land use conversions are greatly reduced in a 15.24 m buffer policy around all streams in watershed. Spatial patterns generated under a 15.24 m buffer policy in CSAs only showed that future projected land use changes occurred close to major highways and shifted the residential development to the northern part of the Opequon Creek. 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 15.24 m 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 conversions in CSAs, result in less pollutant loads by shifting the location of residential conversions to less critical areas where agricultural land is dominant in the watershed.

Keywords Spatially explicit land use · Agent-based · Watershed pollutants · ArcSWAT

Introduction

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 which have 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 (Corbett et al. 1997; Weng 2001). 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 et al. 2009; Mejıa et al. 2014).

The most common non-point pollutants in watersheds causing disruptions in bio-physical 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). 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).

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 parcel as a choice-making unit is ignored at watershed scale. One of the challenges in land use research is that macro patterns can be quantified but the land use conversion decisions are difficult to observe or measure

directly. A typical land use decision occurring at the parcel level involves the 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 economic locations) (Bockstael 1996; Bockstael and Bell 1998).

The prediction of land use change is often defined in socio-ecological research within a deterministic or probabilistic framework. 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 cellular automata (CA) and agent-based models (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).

Among spatially-explicit land use models, CA models provide a simulation framework for modeling land use conversion decision-making where landscape is divided among equal-sized cells. Typically, CA models predict the land use patterns, which are driven by transitional probabilities, and use cell or pixel in raster based grid as a unit of analysis. 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).

A limitation of this approach from the perspective of economic land use modeling is that quantification of decisions occurs at individual cells simplifying decision units into a tessellated landscape instead of actual decision-making units. Additionally, in CA models, the decision-making of land use conversion is exogenously expressed through a random number of agents in the cellular lattice landscape and defining the transition rules as a surrogate for decision-making. Therefore, agents in these models do not belong to observed locations which result in biased spatial relationships between neighboring locations (Parker et al. 2003). 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. In this regard, ABM offers the flexibility to represent agents at their location and can be combined or defined in a CA modeling framework. 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, property parcel size and accessibility. Due to its multiple attributes, individual unit of land can itself be regarded as agent (Le et al. 2010).

In this regard, 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 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.

The specific objectives of this study are to: (1) use an agent-based, probabilistic land use conversion (APLUC) model to project patterns of residential land use changes within Berkeley County, West Virginia by simulating multiple agents' decisions to convert land parcels under alternative policy scenarios; (2) 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 a water quality model; and (3) analyze and compare the outcomes of different land use policy scenarios on land use patterns and surface water quality outcomes.

Following this introduction, the paper is organized as follows: the next section describes the study area. Section three presents a detailed overall model structure for this research. Validation, data and model results are discussed in section four, five and six respectively. The final section summarizes the findings, contributions, limitations of the current approach and future research directions.

Study Area

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

Berkeley County has the second highest population growth rate of all counties in West Virginia (Berkeley County Development Authority 2014). Recent population projection estimates show an expected growth rate of 1.3 % per year for Berkeley County between 2010 and 2030 (BBER 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 population growth has led to increased residential development in Berkeley County. From 2004 to 2006, the County approved several subdivisions that resulted in development of about 1500 ha 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).

The Opequon Creek watershed is selected as the case study area because of the existing non-point source pollution problem and the heterogeneous exurban landscape that includes agricultural and forested areas. 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). Reduced riparian cover and increased impervious surface area from urban growth and development cause stream
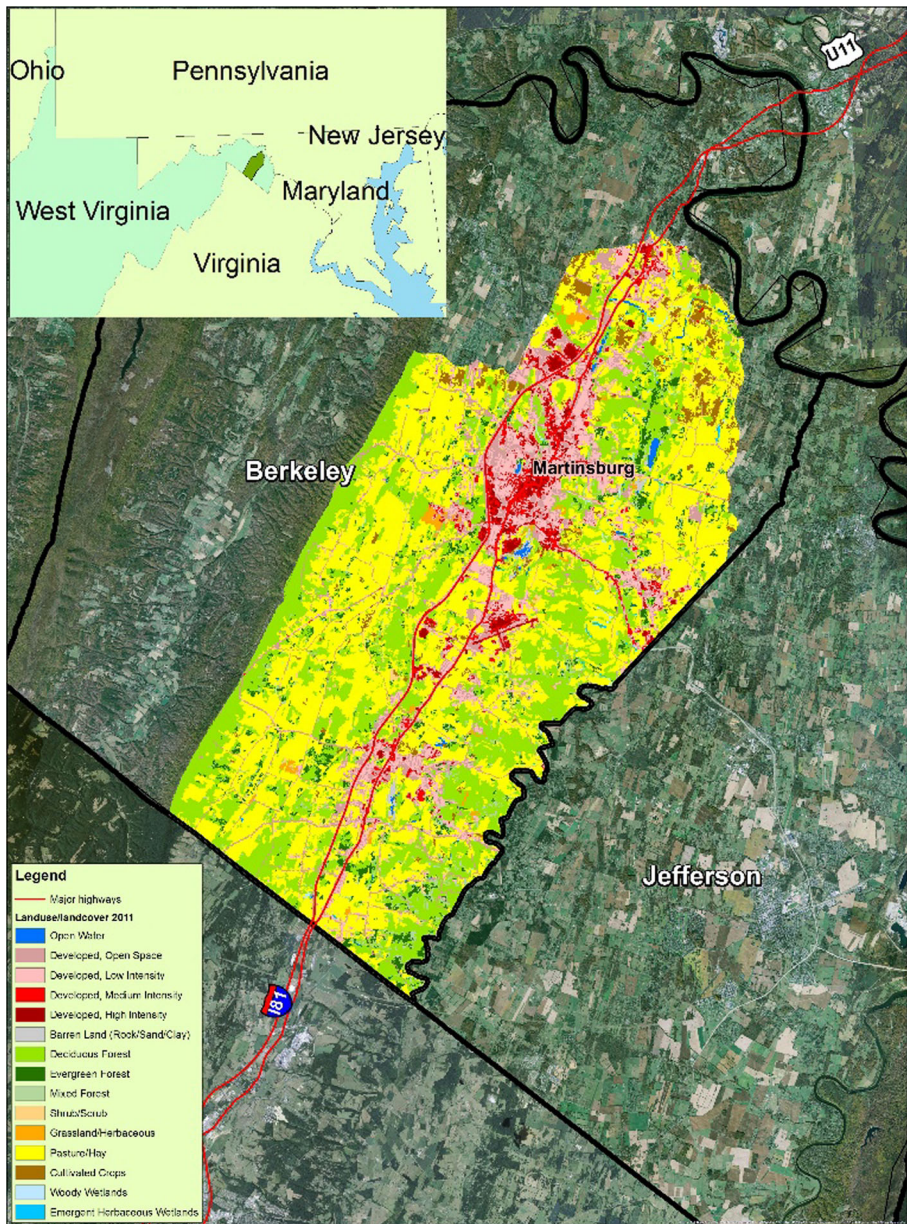
Fig. 1 Location map of the Opequon Creek within Berkeley County, WV. Source: Land use land cover 2011, USGS (2014)

channel erosion and stream flow instability in Opequon Creek area (WVDEP 2008). 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.

Data Description

Land Use

The National Land Cover Database (NLCD) for the years 2001 and 2011 are derived from the Multi-Resolution Land Characteristics Consortium (MLRC) (Homer et al. 2007; Jin et al. 2013). These databases are created by MLRC using Landsat satellite data with a spatial resolution of 30 m (USGS 2014). The most recent land use/land cover data available on the MLRC website is for the year 2011. These land use/land cover data are used for all three models. Fifteen land cover classes in land use data for 2001 and 2011 are found for the Opequon Creek watershed of Berkeley County, WV (Fig. 1). The selection of NLCD data for 2001 and 2011 is based upon the pixel to pixel comparison due to similar classification of land use/ land cover.

The land use datasets for 2001 and 2011 are further reclassified into seven land use/land cover classes using a common scale for the SLR model and for the APLUC model. The seven land use classes and their definitions utilized in the dataset are shown in Table 1 and Fig. 2. 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 are aligned using 30 m spatial resolution in the geo-processing environment.

Raster based land use data sets for 2001 and 2011 are 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 are 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 are defined by property parcels. These zones are based upon the single output value of land use data (the value raster) representing the most common

Table 1 Land uses/land cover classes and their description, (EPA 2007; USGS 2014)

| Land uses/land cover	Description
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 characterized by 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.
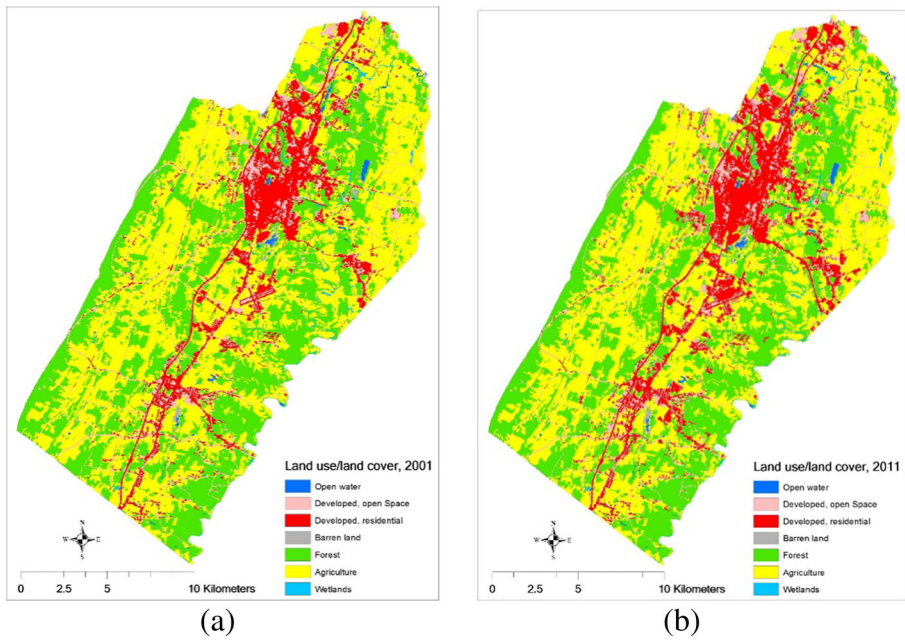
Fig. 2 Reclassified land use/land cover for the study area, 2001 (**a**) and 2011 (**b**)

cells within the parcel zone (ESRI 2014). In APLUC model, the developable parcels consist only agricultural and forest land use.

Property Parcel Data

Actual property parcels are obtained from the Berkeley County Assessor's office for the year 2011. Property parcel feature data for Berkeley County are extracted for Opequon Creek watershed. The descriptive statistics in Table 2 show the variation in the size of irregular 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 is based upon Berkeley County Commission's requirement for minimum residential density (Berkeley County Planning Commission 2009).

Table 2 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

Urban Center

In this research, urban centers are characterized by 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 is collected from U.S. Census Bureau (2000). Data for the year 2001 is not available, therefore, population density data by Census 2000 tracts is 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 is 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 are selected: 971,500, 971,600, and 971,700. These tracts had population densities of 2846, 2705 and 3340 persons per square mile, respectively. The total area of this urban center is found to be 4.51 mile^2. Martinsburg is identified as major urban activity center within the Berkeley County in the model.

Another important feature for Berkeley County is 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 is created from the centroid of Martinsburg, the train station and the demographic boundary of the urban center.

Major Highways

Major highways as road features are collected from the U.S. Department of Transportation (1997). In general, road features typically remain constant over long periods of time. Data for 1997 is 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 are selected for the study area.

Streams

Data on streams are delineated through the ArcSWAT (an ArcGIS extension of Soil and Water Assessment tool (SWAT)) watershed delineation based on digital elevation model (DEM) raster for the Opequon Creek watershed. The elevation is in meters having 30×30 cell size. ArcSWAT draws the location of the stream network based upon the flow direction and accumulation using DEM grid. The minimum and maximum, and ArcSWAT defined sub-watershed drainage areas are 107, 21,327, and 426.54 ha, respectively.

Methods

This research is based on three interconnected models to illustrate the concept of driving factors of land use change, land conversion decisions, and linking land use conversions to

water quality indicators (Fig. 3). This methodology provides a hybrid approach of a spatial logistic regression (SLR) model to calibrate APLUC model. SLR provides coefficients for explanatory variables of land use change from observed land use conversions. The APLUC model simulates the decisions to convert developable land into residentially developed land, given land parcel surrounding attributes, and generates land use patterns at a disaggregated scale. By linking the APLUC and ArcSWAT models, the impact of residential development on surface water quality in Opequon Creek is assessed.
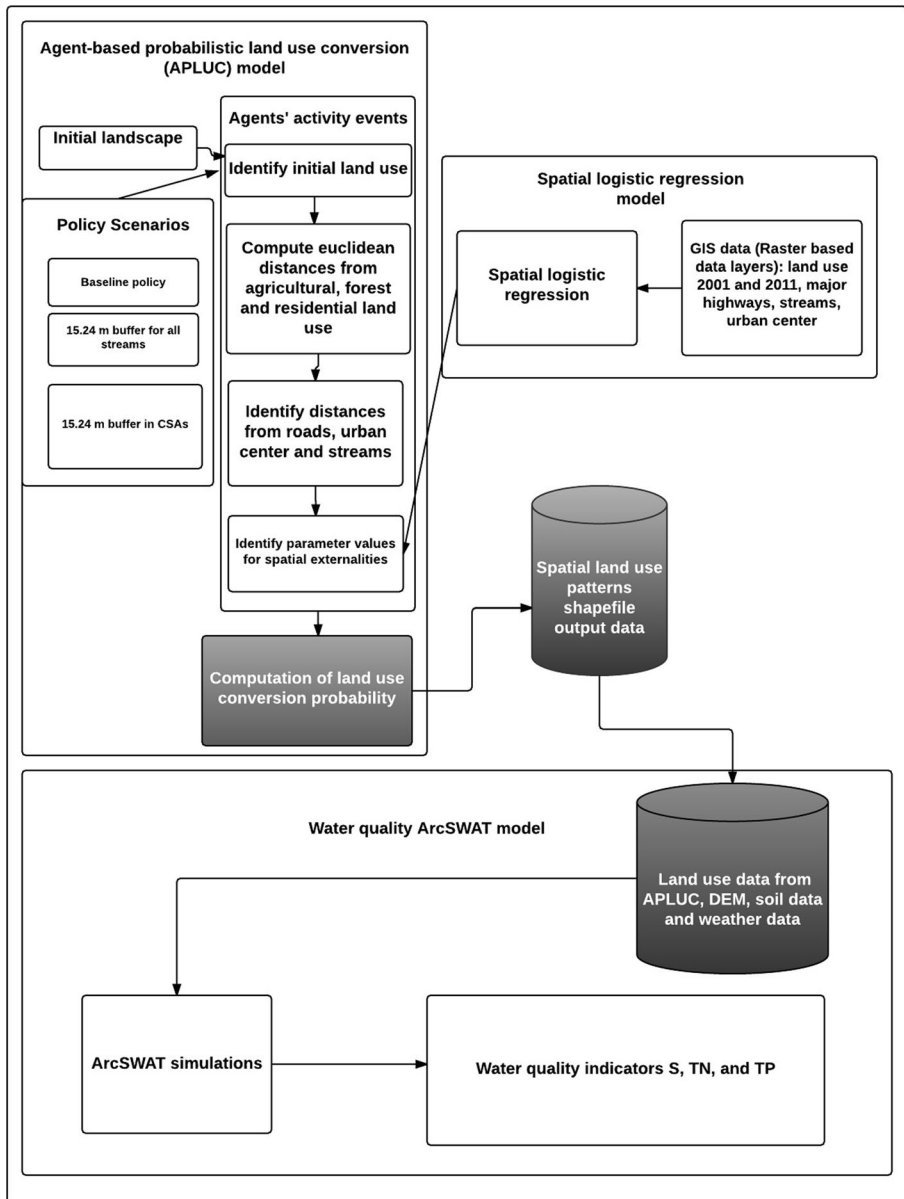

Fig. 3 Conceptual framework for this research

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:

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

Where P(Y = 1|X) is the predicted probability value of the binary or dichotomous dependent variable Y, 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. This logistic function has 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 (1) is used to estimate the β coefficients (Menard 1995):

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

Y is the probability that the dependent variable (Y) is 1, p_k is the predicted probability of the kth parcel of agricultural or forest land use conversion to residential land. β_0 is the intercept, and β_1, β_2, β_3, β_4, β_5, and β_6 are coefficients for distance to the existing agriculture (x_1), distance to the existing forests (x_2), distance to the existing residential areas (x_3), distance to streams (x_4), distance to major highways (x_5), and distance to urban center (x_6), respectively. These coefficients measure the influence of each independent variable on the variations in probability of land use conversion from non-residential land use to residential land use (Y). 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). Distance to the roads and urban centers are used to conceptualize the land rent and transportation cost, respectively, under the relaxed assumptions of spatial variation in the landscape based on the intuition of Von Thünen (1826) and the bid-rent theory of urban economics (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 centers, and schools (Serneels and Lambin 2001).

Spatial externalities within the APLUC model are deduced from coefficient estimates in the SLR based upon results from Arbab (2014). 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.

APLUC Model

This model is designed to simulate land-use decisions of agricultural and forest land owners to convert undeveloped land to residential land. A necessary condition in this model is higher residential land value compared to agricultural or forest land. This necessary condition is suggested based on high rates of conversion from farmland to developed land (Olson and Olson 1999; Koontz 2001; Rosenberger et al. 2002).

Model Environment

The model employs a geographical information systems (GIS) environment composed of actual property parcels in Opequon Creek watershed area within Berkeley County. Parcels are assigned with land use conversion rule under three land use policy scenarios, which are described later in this section.

The parcel agents are developable parcels representing 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. In the APLUC model, an agent's rule 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). The following key assumptions underlying the land use model dynamics are:

- Following logic from Polhill et al. 2008, agents predict future land use conversion in one way. They form probabilistic land use decision-making. Agents know their property parcel, location, land use conversion probability value of all other parcel agents, and distances from each land use,
- The model is spatial and not temporal in its scope. The events of conversion are discrete, therefore agents do not foresee the effects of land use decisions of their neighboring parcel for more than one event period,
- It is out of the scope to add selling and buying framework in the model, 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,
- Agents are not assumed to exhibit optimizing behavior on an inter-temporal basis, and
- Due to data limitations, all the convertible parcels are recognized as agricultural or forest land use parcels based upon majority of land use type in each parcel. Accurate information on proportion/percentage of each land use type in each parcel is not available. Additionally, division of parcels required additional zoning and subdivision ordinance policies, which is out of the scope of current study. Therefore, parcels are not split, but are consistent throughout the study time period.

In the present modeling approach, spatial interactions and dependencies are embedded in neighboring externalities. Neighborhood externalities are the estimated influence of each land use on surrounding parcels. 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).

The agents' conversion decisions vary with the spatial distances from each neighboring land use over a period of 10 iterations (to roughly approximate a 10 year time period), where each iteration is assumed to be a conversion event possibility. Additionally, the modeling factor that influences decision-making is path dependency in which the initial conversions influences future conversions within each model run. By modifying an approach from Benenson and Torrens (2004), each parcel agent's probability (Prob) of conversion from developable state m to residential state r in each iteration is modeled as:

$$\text{Prob}_i(S_m \rightarrow S_r) = S(N(i)) \tag{3}$$

Where N (i) represents parcel agent i's neighbors and S represents state of parcel i. Decision rules and initial conditions such as distance to streams, roads, and urban centers for each property parcels do not change over the course of operation for the APLUC model. The model employs a Monte Carlo process (Hagerstrand 1965; Wu 2002) to generate the results of a stochastic APLUC model. Due to uncertainty, a probability function is used to condition the residential conversions utilizing a random number generator (Batty 2012).

For undeveloped parcels, the conversion decision is based upon a comparison between the random number generated and the probability value computed from equation (1) for each parcel. The random number generator *rand* (φ_i) has a random distribution that is uniform between 0 and 1. The agents adopted the following rule of land use conversion in each iteration.

$$\text{if} \quad \text{rand}(\varphi_i) < P_{it} \text{ then } A_{it+1} = r \tag{4}$$

Where r 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. The random number generator incorporates a stochastic element into the APLUC model, which mimics uncertainty in the model.

SLR is used to calibrate projections of residential land use conversions under each policy scenario. The purpose of this calibration is to extract the coefficient values from the SLR into the APLUC model using observed land use pattern at each iteration t and subsequent iterations t + 1. The empirical structure of agent's land use conversion probability is generated through parameters from the SLR. The SLR coefficients are measured in terms of meters in raster based environment and assigned to each parcel.

Process Overview and Scheduling

The APLUC model is created using 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 generates 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 data used in the SLR model. The raster from SLR consists of an aggregation of 10 years of land use change from 2001 to 2011. A single iteration represents the duration of one time period (a year) as counted in the SLR model (Fragkias and Seto 2007). All land use conversions generated synchronously at the end of each iteration.

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 agents compute the mean Euclidean distances from agricultural, forest, and residential lands. Once the distances are calculated, agents identify the coefficient values for these spatial externalities and identify the distances from roads, urban center, and streams. Parcel agents incorporate the estimated SLR coefficients and for each iteration, they calculate new values for the spatial externalities 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.

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 generated by SLR model 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 Eq. (1) and update the probability using local spatial patterns for each parcel/polygon throughout the simulation iterations. 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, 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 the formulation of the relationship between P_i (computed with SLR equation) and P_a (a probability of conversion for each iteration) is:

$$P_i = 1 - [1 - P_a]^{10}$$
$$P_a = 1 - [1 - P_i]^{\wedge 0.1}$$

(5)

P_a is defined as the average probability per iteration over the ten iterations as this value changes with each iteration due to land use changes.

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

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.

To account for the probabilistic nature of conversions, a Monte Carlo process of the APLUC model is used by repeating 50 model runs. Each model run generated a different set of land use conversion sites. Employing 50 runs allowed examination of fluctuations among model runs. Due to the defined empirical structure of local probability, model results showed fluctuations at consistent rate. Therefore, the choice of 50 model runs for each policy scenario is judged to be adequate for testing the path dependency and stochastic processes in this 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 is calculated as:

$$P_j = \sum_{x=1}^{50} C_x \Big/ 50$$

(6)

Where P_j the probability of conversion for parcel j, x the number of model runs, C the Boolean conversion in each model run results in either one or zero, where one indicates conversion and zero 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 have ranged between 0.50 and 1.00 (Zeeb and Burns 1998; Louis and Raines 2003; Sohn and Park 2008; Fragkias and Seto 2007). Logically, the projected probability is interpreted as parcels which have at least a 0.50 probability of land use conversion are residentially developed parcels (value =1), while projected parcels with <0.50 probability are assumed as no conversion (value =0).

A data generator step and a land use conversion step are performed in each iteration. For each policy option, the model is used to determine the land use conversion of the total number of parcels converted to residential development using a Monte Carlo process. The projected Monte-Carlo spatial land use patterns are used as data input for the ArcSWAT model.

ArcSWAT Model

ArcSWAT is used in this study to estimate water quality outcomes stemming from land use conversion changes. The ArcSWAT model is set up using data on the Opequon Creek Watershed terrain (30 m resolution (DEM)), land use, soil type, and local meteorological conditions. A DEM of 30 m is 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.

Using ArcSWAT, the Opequon Creek watershed is 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 Soil survey geographic (SSURGO) soil data layer is linked with the soil database. The land use/land cover data from the APLUC, the SSURGO soil data, and the slope class layers are overlaid to derive unique hydrological response units (HRUs) or sub-basins. For the distribution of HRU's, dominant land use, soils, and slope are used. HRUs are defined as an area that has a unique combination of land, soil type, and slope characteristics (Neitsch et al. 2005). 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.

Land use data layers generated from the APLUC model for each policy scenario are implemented in ArcSWAT. From these land use data, pollutant loads are calculated for each sub-basin in the watershed over a 10 year simulation period. Land use classifications within ArcSWAT generated four types of land uses: residential, open space, forest, and agriculture. 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. Pasture is assigned to an agricultural

land use class due to its being the observed dominant land use in Berkeley County. Deciduous forest is selected for forest class in ArcSWAT due to its observed high percentage of land use in Berkeley County. For this research, ArcSWAT simulations are run annually over 10 years' worth of data. The choice of a 10 year time period for simulation is 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 database at a sub-basin scale. Pollutant loading data are the outcome of the model. The pollutant criteria is defined as the pollutant load releasing out of the watershed. Having this consideration, total phosphorus (TP), total nitrogen (TN), and sediment are selected as pollutant loadings from ArcSWAT output. With each pollutant load calculated per hectare under each policy scenario, the relative performance of each land use policy scenarios are compared for the three policy scenarios.

Policy Determination

The APLUC model provides the capacity to be used as a policy tool for assessing different policies at a watershed scale. In the APLUC model, the buffer zone areas are set as no development zones. ArcGIS 10.2 is used to prepare spatially restricted buffers for the APLUC model.

No Policy (Baseline Policy)

The no policy scenario involves no additional regulation or spatial restrictions in land use conversion. The importance of this policy is to simulate the water quality impacts of land use conversion under the existing land use regulatory framework in Berkeley County.

Implementation of 15.24 m Stream Buffer Policy

Riparian (forest) 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). 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). 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 0–3 (or 0–15.24 m on each side) (Hatten et al. 2011). In this analysis, a linear 15.24 m buffer zone on each side of all streams in the Opequon Creek watershed is delineated using buffer analysis in ArcGIS. The streams are delineated by the ArcSWAT model.

Implementation of 15.24 m Stream Buffer Policy in CSAs only

Sub-basins in a watershed system play a key role in nutrient and pollutant loadings (Peterson et al. 2001). Pionke 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). CSAs are defined as the areas that show the highest loading of total phosphorus (TP), total nitrogen (TN), and sediments (Niraula et al. 2013). For water quality management, it is important to identify and assess the impacts of these CSAs and set them as high priority for land use policies to protect water quality (Pionke et al. 2000). 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 the hydrologically corrected 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 (CSAs) within the Opequon Creek watershed.

Model Validation

To quantitatively minimize the spatial patterns of errors (autocorrelation) between the connectivity of cells (pixels) in the SLR model, a pixel thinning method is used. In this method raster data for all explanatory variables and dependent variable are contracted by the contraction factor of the 10th lag. With the 10th lag, every 10th cell (pixel) is selected that provides a wider spatial distance between each cell to minimize the effects of spatial autocorrelation and to reduce the number of cells in the sample. Due to the small study area and the pixel thinning method, the model does not allow use of a subset of 2001–2011 data for validation. Initially the data was divided among two subsets: 2001–2006 to parameterize the SLR model and 2006–2011 land data to be used as an independent source for validation. Using five years of land data, the model produces an extremely small sample size which does not generate any meaningful results. The sample size generated with the use of 2001–2006

data decreased the number of converted cells to 5 compared to 85 cells in 2001–2011 data. Therefore, the validation of the model is based upon 2001–2011 land use data. The purpose of using parameters based on 2001–2011 land use data is to project land use changes using agent-based simulation.

Calibration tends 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, uncertainty in the input data, and unanticipated future land use changes precise land use conversion projections are not always feasible in validation (Pontius and Neeti 2010; Memarian 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. Land use data from 2001 are employed as an initial condition and projected the residential land use conversion for 2011 using the APLUC model. 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.

Actual land use change between 2001 and 2011 shows that 4748 parcels are converted into residential parcels, which comprise 1114.85 ha of land in the Opequon Creek watershed. To match this 10 year period, residential land use conversions are observed for 10 iterations in the APLUC model. The spatial robustness of the model is 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 are 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 2394 parcels are converted into residential parcels, comprising an area of 1373.26 ha. The model accurately replicated 722 parcels (156.67 ha), which are 15.20 % similar spatially located parcels as observed in actual 2011 land use data (Table 3).

Table 3 Comparison among actual and projected residential conversions at parcel level

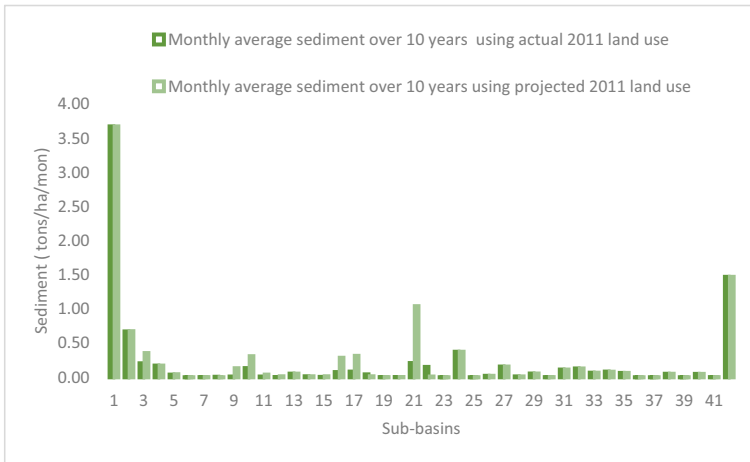|	Area (ha)	Number of parcels
Actual residential parcels	1114.85	4748
Projected residential parcels	1373.26	2394
Correct Projections	156.67	722

Fig. 4 Sediment yield in each sub-basin for actual and projected 2011 land use data

The data generated by ABM simulation is further used in the ArcSWAT model to estimate pollutant loads at sub-basin scale. In the present modeling approach, validation of the SLR based projection within ABM are examined from the sub-watershed level comparison in terms of the ArcSWAT model results. In this comparison, pollutant load outcomes are compared between projected residential conversion versus actual residential conversion between 2001 and 2011.

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 is rank ordered by monthly average load over three different pollutants: sediment, TP, and TN. Ranking is on a per hectare basis and in tons for sediments and kg for TN and TP (Figs. 4, 5 and 6).

Figures 4, 5 and 6 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 3, 9, 10, 16, 17, 21 and
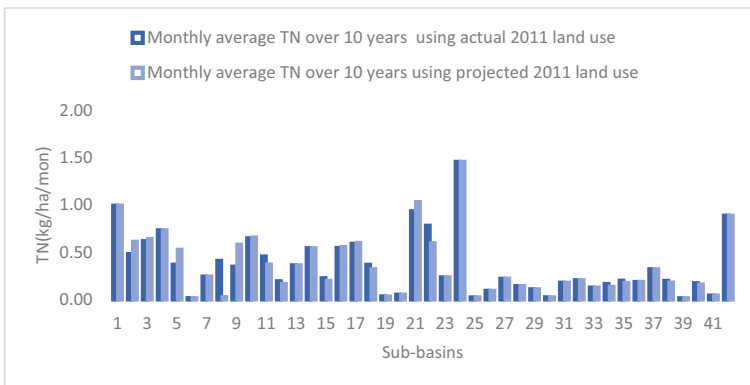

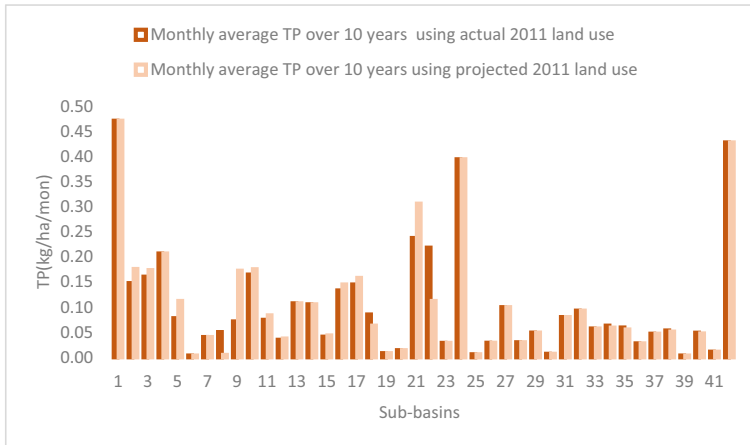Fig. 5 TN yield in each sub-basin for actual and projected 2011 land use data

Fig. 6 TP yield in each sub-basin for actual and projected 2011 land use data

22, TN for sub-basins 2, 5, 8, 9, 21 and 22 and TP for sub-basins 5, 8, 9, 21 and 22 (Figs. 4, 5 and 6). For statistical validation, a Spearman rank correlation coefficient is calculated for each pollutant loading (sediment, TN, and TP) comparing actual land use in 2011 versus projected land use in 2011. The ranking is done in descending order from highest to lowest pollutant yield.

The Spearman rank correlation coefficient is defined as:

$$\rho = 1 - \frac{6\sum d_i^2}{42\left(42^2 - 1\right)} \tag{7}$$

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 2011 actual land use data and the monthly average loading for sediment, TN, and TP over 10 years for each sub-basin using 2011 projected land use data are 0.60, 0.47, and 0.51 respectively with 42 degrees of freedom (the number of sub-basins) (Table 4).

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 4). These significance tests show that associations exist between the two land use datasets and are statistically significant for all pollutant types, which provide statistical validity of projected 2011 land use data at a sub-basin watershed level.

Table 4 Spearman correlation analysis between actual pollutant loadings and projected pollutant loadings for 2011

Pollutant type	Rank Coefficient (ρ)	Sig.(One-Tailed)
Sediment	0.60	0.000**
Total Nitrogen (TN)	0.47	0.000**
Total Phosphorus (TP)	0.51	0.000**

**P < 0.01

Results

This section analyses and interprets the empirical results of the three linked models utilized in this research: SLR, APLUC, and ArcSWAT model.

SLR and APLUC Models

The overall model statistics from logistic regression analysis are summarized in Table 5. By using a pixel thinning method in the IDRISI software, negative impacts of spatial interdependence are reduced. The result is 3468 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.

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

Initially, cells are classified using a SLR predicted probability threshold of 0.5 where cell probabilities of less than 0.5 are classified as not converted (0) and cell probabilities greater than or equal to 0.5 are classified as converted (1). Based upon this threshold, the odds ratio is calculated as (Clark Labs 2012):

$$\text{Odds Ratio} = \frac{(f_{11}*f_{22})}{(f_{12}*f_{21})} \tag{8}$$

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 converts into residential land use and if $x\beta_i \leq 0.11$, the cell does not change. The resulting outcome provides the percent correct land use conversions and non-conversions. The resulting value of the adjusted odds ratio is 13.29.

By assigning a 0.11 cutting threshold for predicted probability, true positive and false positive are calculated. The value of true positive is 46.15 % and false positive is 1.98 %. Receiver operator characteristic (ROC) is performed by comparing the fitted cells that are converted and actual cells that are converted during 2001–2011. ROC

Table 5 Spatial logistic regression model results

Statistics	Value
Number of total observations	3468
Number / percentage of 0 s in sampled area	3383 (97.55 %)
Number/ percentage of 1 s in sampled area	85 (2.45 %)
Chi-square (6)	123.82 (p-value 0.00001)
−2logL0	798.37
−2log(likelihood)	674.55

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.

There are 85 cells that are converted into residential land use and are represented by 1 s in the model. The fitted 1 s are 18 and fitted 0 s are 67. 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.

The parameter estimates from logistic regression model are shown in Table 6. 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 6 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 (1826) 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

Table 6 The Likelihood ratio statistics and chi-square significance for each predictor

Explanatory Variables	Coefficient values	$-2LL_k$	$-2LL_{k-1}$	$\chi2=-2LL_k - (-2LL_{k-1})$	$P > |\chi2|$
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

**shows significant at 0.01

availability of the public water and the public sewer systems (Berkeley County Planning Commission 2006).

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 difference between a full model and a model with one less predictor. The difference between these two values is Chi-square test for goodness of fit and calculated as follows:

$$\chi 2 = D_k - D_{k-1}$$
$$= -2LL_k - (-2LL_{k-1})$$

(9)

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 helps 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).

The results are reported in Table 6. Except distance to streams and distance to highways, 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 the basis of the significance test results.

The SLR function represents coefficient values based upon the raster data derived for the study area. These coefficients can be regarded as weights to produce the global probability of change (Shirzadi et al. 2012). The βx function is calculated as:[1]

$$\beta X = -2.2730 + 0.00336 \cdot \text{agric}_{\text{dist}} * + 0.00396 \cdot \text{forest}_{\text{dist}} ** - 0.00266 \cdot \text{resid}_{\text{dist}} ** \quad (10)$$

$$-0.0004 \cdot \text{stream}_{\text{dist}} - 0.00012 \cdot \text{highway}_{\text{dist}} - 0.00010 \cdot \text{urban}_{\text{dist}} **$$

Further, Monte Carlo probabilities are mapped for each property parcel. For model testing, the spatial land use patterns are observed in each iteration. It is observed in each model run that the number of conversions is higher in earlier iterations compared to later iterations.

The parcel level probabilities from the Monte Carlo simulations are divided into two classes where a probability less than 0.50 is considered as non-conversion and parcels with a probability equal to or more than 0.50 are considered as residentially developed parcels.

The various sizes of land parcels and influence of surrounding land uses incorporate spatial heterogeneity into the APLUC model. Differences among parcel agents are due to different land use types surrounding each parcel based on its location. Resulting differences in conversion probabilities were then assigned to land use conversion

---

[1] * shows significant at 0.05 and ** shows significant at 0.01.

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.

Policy Scenarios

To illustrate the capability of APLUC to simulate alternative scenarios, three policy scenarios are examined to project anticipated future residential land use conversions for the Opequon Creek watershed. Conversion decisions that are spatially linked with adaptation of buffer policies facilitate 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 under a certain policy, as well as the potential changes in allocation of land use in each sub-basin due to these policies.

No Policy

Under the no policy scenario, results show the highest number of land use conversions compared to two buffer policies (Table 7). Converted residential parcels ranged between 1486 to 1562 property parcels over the 50 model runs. The Monte Carlo projections under this policy scenario show that, on average, a total of 1531 parcels are converted, which consisted of 1451.26 ha (or 4.08 % of total land). Overall, 83.29 % residential land use change (in terms of area) occurred from agricultural land consisting of 1064 previously agricultural parcels with total area of 1208.81 ha.

This could be caused by a high percentage of agricultural land use (developable parcels) being closer to residential parcels. 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. Conservatively, these potential converted parcels by Monte Carlo simulations are regarded as potential residential growth areas. Results also show that the majority of developable parcels have a clear tendency to convert into residentially developed areas closer to the urban center (Martinsburg area) in the no policy scenario (Fig. 7). Additionally, land use

Table 7 Number of parcels converted from each developable land use type under each policy scenario

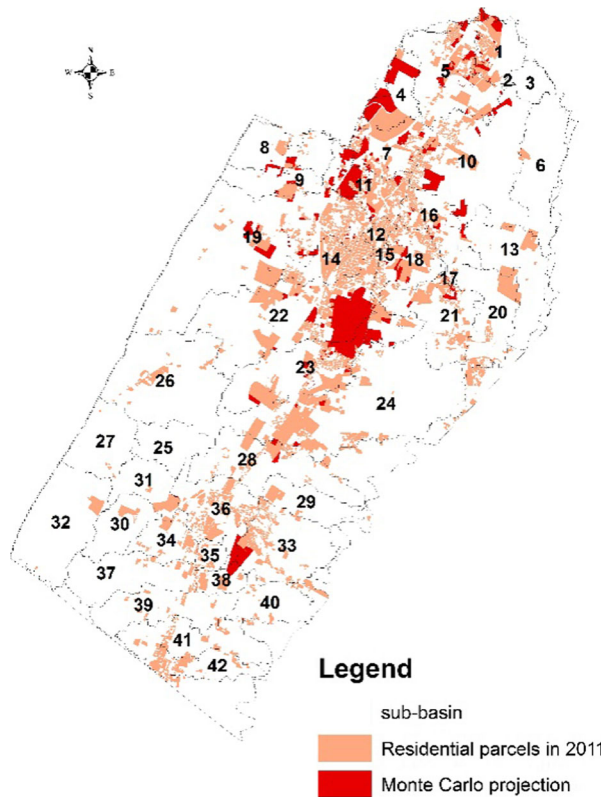| Parcels	No Policy	15.24 m Buffer (All Streams)	15.24 m Buffer (CSAs Only)
Total parcels converted	1531	456	1426
Agricultural to residential	1064	318	1062
Forest to residential	469	138	366

Fig. 7 Land use conversions under no policy

conversions from forest to residential land use are only a small proportion of the total residential land use conversions (16.70 % of land). 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.

15.24 m Stream Buffer Policy

The adoption of the riparian buffer policy restricted agents from developing residential land in buffer zones and resulted in the fewest number of residential conversion projections of any policy (Table 7). In the SLR model, the coefficient value is positive for the distance from forest land use. This distance based spatial externality results in a decrease in probability of conversion of many properties that are closer to forest land use. Therefore, not only do forest buffers constrain the development of residential land, but additionally act as spatial externalities which reduce nearby conversions. This buffer policy also resulted in majority of small sized parcel conversions than the other two
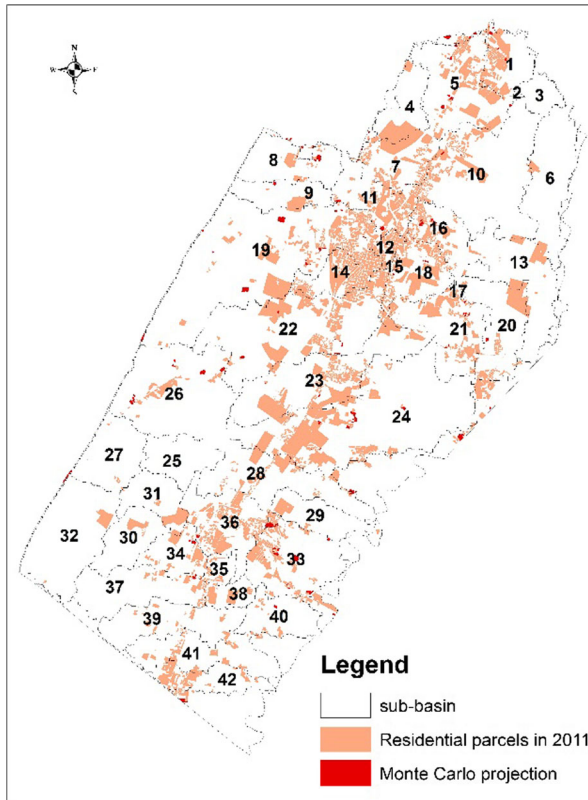
Fig. 8 Land use conversions under 15.24 m forest buffer scenario

policies (Fig. 8). The spatial distribution of parcel agent by size is not homogeneous throughout the watershed, and therefore, the adoption of the policy is 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 are required to adopt a buffer policy resulting in a decrease of these large parcel conversions.

15.24 m Buffer in CSAs Policy

Under a 15.24 m riparian buffer for streams in high priority sub-basins, there is very little decline in land use conversion compared to the no policy (Table 7). Under this policy, 1172.72 ha of land converted into residential land use. Most conversions under this policy are occurring along the major highways. Specifically, bigger parcels are converted and located in the northern part of the
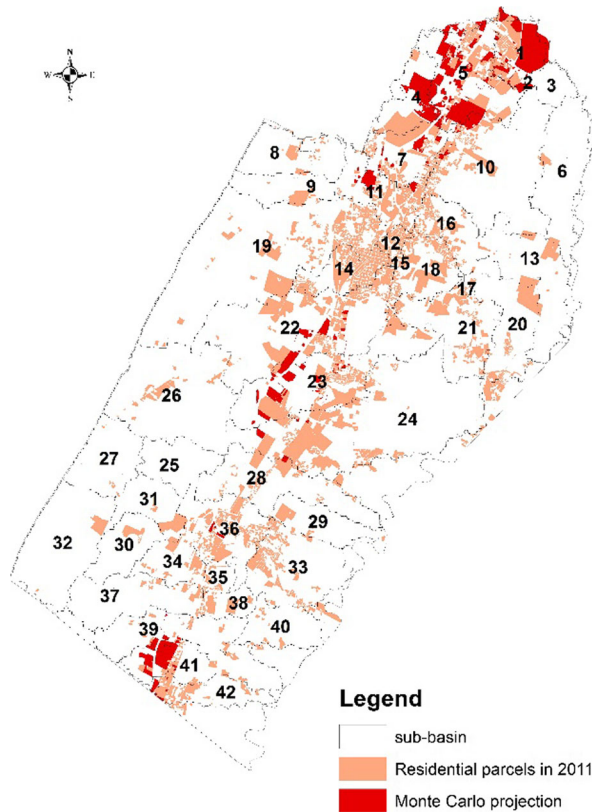
Fig. 9 Land use conversions under 15.24 m forest buffers in CSAs

Opequon Creek watershed (Fig. 9). This result implied that the presence of highways is more influential on residential land use conversion under this policy.

ArcSWAT Model Results

The null hypotheses of no relationship between residential land use and pollutant loading are rejected for the pollutants of sediment and TP, but not for TN (Table 8). In general, TP has a relationship with suspended solid loads, therefore, reduction in sediment leads to a reduction in TP (Neitsch et al. 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 (Table 8). Therefore, reductions of these pollutant loads are important considering residential land use policies to protect water quality.

Table 8 Spearman rank correlation analysis by pollutant type

Pollutant Type	Rank Coefficient (ρ)	Sig.(One-Tailed)
Sediment	0.43	0.002**
Total Nitrogen	0.21	0.091
Total Phosphorus	0.47	0.000**

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

The results of ArcSWAT are analyzed for entire watershed (Table 8) and for each sub-basin (Fig. 10).

Water Quality Outcomes under no Policy

To eliminate the effect of seasonal differences in flows and loadings, average monthly loadings over a 10 year period for the entire watershed are reported in Table 9. The average TN loading for the entire watershed is the largest of the three policy scenarios. High proportions of TN and TP are found in sub-basins with residential or agricultural land use as dominant land use (Fig. 10). Most notable is sub-basin 1 where agricultural land is the dominant land use and there are high loadings of all three pollutants.

Water Quality Outcomes with 15.24 m Stream Buffers around all Streams

Under this policy, there is a 75.52 % decrease in residential land use conversion when compared to the no policy scenario. Despite this reduced residential land use, there are not substantial decreases in any of the three pollutant loads watershed wide or at the sub-basin level (Table 9). Buffers on all streams have the lowest TN loadings of any policy by about 7 %. Since the Opequon Creek watershed has a relatively high percentage of agricultural land use (which is largely pasture), the ArcSWAT results indicate that this land use contributes largely to nitrogen loads (Fig. 10). Implementation of buffers causes slight decrease in monthly average/ha TN loads.

Water Quality Outcomes with 15.24 m Stream Buffers in CSA Sub-Basins

This policy showed much lower sediment loadings than either of the other policies. Under this policy, the sediment loadings decreased by over 50 % compared to both other policy scenarios (Table 9). TP loadings under this policy also are slightly lower than the other two policies.

Specifically, loadings from sub-basin 1 are 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 is 0.03 tons/ha/month under this policy compared to 3.68 tons/ha/month which is the highest of any sub-basin in both the no buffer and buffer everywhere policies (Figs. 10). In no policy and buffer everywhere policy, sub-basin 1 is dominant in agriculture (pasture) land and only has 60.235 ha of residential land use. While with buffer policy in CSAs only (Fig. 10), sub-
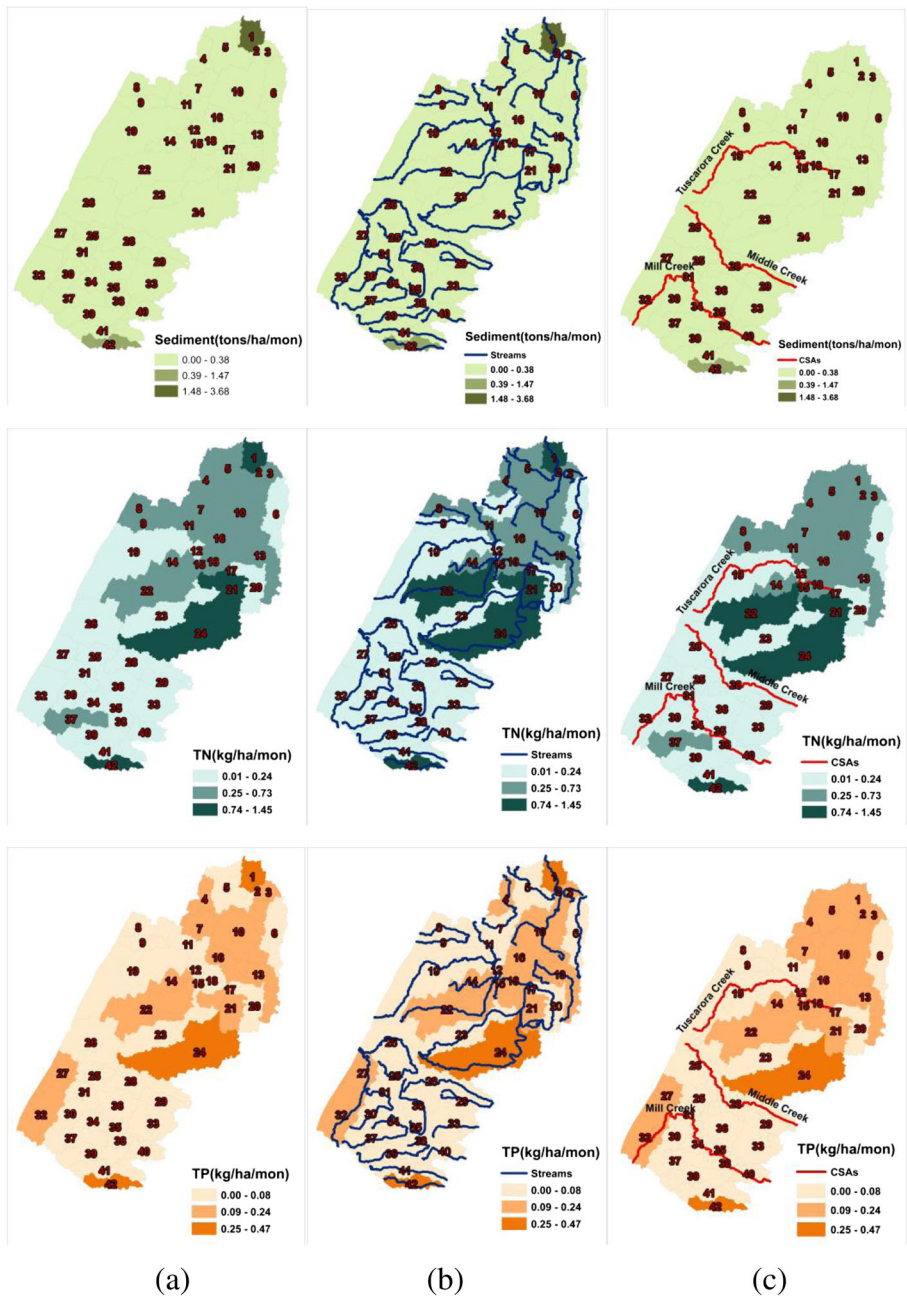
Fig. 10 Sediment, TN and TP yields from each sub-basin as estimated by ArcSWAT with **a** no policy, **b** 15.24 m stream buffers everywhere and **c** 15.24 m stream buffers in CSAs

basin 1 has 284.348 ha of residential land use. This increase in residential land made this land use type the dominant land use in the sub-basin. The result is a dramatic decrease in mean sediment loading. In addition, results show that mean TP loading

Table 9 Average pollutant loads per month for the entire Opequon Creek watershed projected over a 10 year period

Scenario	Sediment (tons/ha/month)	Total Nitrogen (kg/ha/month)	Total Phosphorous (kg/ha/month)
No Policy	7.88	14.91	4.01
15.24 m buffer (riparian) for all streams	8.28	13.86	4.06
15.24 m buffer (riparian) in CSAs	3.75	14.86	3.80

declined along with sediment loading in sub basin 1 from 0.47 kg/ha/month to 0.10 kg/ha/month, although not as dramatically.

Conclusions, Limitations and Future Works

This research provides a detailed and comprehensive analysis of projected land use change for the West Virginia portion of the Opequon Creek watershed, based on prior observations of land use change. SLR model provides estimates of how distance based explanatory variables impact land use conversions using data on observed conversions over a 10 year period (2001–2011) within the watershed. 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 negative spatial externalities for residential land use conversion. The distance to existing residential land use has a negative coefficient, and thus, proximity has a positive marginal effect on residential land use conversion. Using the SLR model results, empirical parameters to project land use change probabilities are provided for the APLUC model.

The APLUC model results show that each buffer policy has different impacts in terms of location and residential land use conversions. Under no policy, urban center and residential land use are the driving factors of residential land use conversions. Thus, most of the projected residential conversions occurred close to the urban center or residential properties under this policy scenario. For residential land use conversions under the 15.24 m buffer policy around all streams in the watershed, a much lower rate of residential conversion took place compared to no policy. The resulting residential patterns showed sparse, small conversions to residential land due to buffer zones limiting larger parcel conversions. Stream buffers only in the CSAs policy resulted in a slight decrease in the quantity of residential land use conversions when compared to the no policy result and shifted residential development to sub-basins located in the northern part of the Opequon Creek watershed.

The water quality results from ArcSWAT suggest that location of residential land use is more important in the Opequon Creek watershed than restricting land use conversion. Under ArcSWAT, agriculture land use causes higher pollutant loads in the watershed. Thus, drastically decreasing the residential land use conversion does not solve pollutant loading problems. Instead, implementing buffers in high priority watersheds results in more effective water quality outcomes. The buffers shift the

location of conversion to sub-basins that are dominant in agricultural land use, thereby reducing sedimentation from agricultural sources. In this way, residential land conversion in these sub-basins are reducing agricultural related pollutant loads and improving water quality.

There are numerous dimensions in which the model applied 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 data sources such as community surveys, historical data, and parcel based socio-economic data with more frequent calibrations can be employed to improve the projections.

This study calibrated a land use conversion model for only one watershed. There is an uncertainty in the implementation of the results to other study areas and if the results show general or specific land use patterns to the Opequon Creek area. Since Opequon Creek is located in the Chesapeake Bay system, similar modeling adaptations for other watersheds in the Chesapeake Bay are suggested. Secondly, what are the value of spatial externalities if other model structures are applied for calibration such as Bayesian probability instead of logit transformed probability? Does information on perceptions of property owners change the projections? Additionally, the current modeling framework is limited by the lack of information on real estate markets and other equally important factors. Inclusion of information on market dynamics, 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 direct causes of land use change, however the projections of future land use change require an understanding of the forces that may be difficult to observe. These forces may involve cultural values, market conditions, and policy changes which are difficult to forecast. Also, 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 socio-economic information in assessing land use conversions. Further, broader land use classes for conversions and different initial conditions could be implemented to understand the interspersion of different land use types and resulting water quality indicators. Such a study should also consider other land use specific practices such as sustainable management and best management practices in conjunction with riparian management. The APLUC model could be integrated with other ecological models to assess impacts on biodiversity and ecosystem services in order to provide multiple assessments of the impacts from land use change. This would make the APLUC model a valuable policy tool to anticipate future land use systems.

References

Alig, Ralph J. (2010). Economic modeling of effects of climate change on forest sector and mitigation options: a compendium of briefing papers. USDA

Almeida, C. M., Batty, M., Monteiro, A. M. V., Câmara, G., Soares-Filho, B. S., Cerqueira, G. C., et al. (2003). Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation. *Computers, Environment and Urban Systems, 27*(5), 481–509.

Alonso, W. (1964). *Location and land use*. Cambridge: MA. Harvard University Press.

Arbab, Nazia N. (2014). Application of a Spatially Explicit, Agent-Based Land Use Conversion Model to Assess Water Quality Outcomes under Buffer Policies. (PhD dissertation). West Virginia University.

Atkinson, P. M., & Massari, R. (1998). Generalised linear modelling of susceptibility to land sliding in the central Apennines, Italy. *Computers & Geosciences, 24*, 373–385.

Balzter, H., Braun, P. W., & Köhler, W. (1998). Cellular automata models for vegetation dynamics. *Ecological Modelling, 107*(2–3), 113–125.

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). New York, NY: Springer.

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.

Benenson, I., & Torrens, P. (2004). *Geosimulation: Automata-based modeling of urban phenomena*. West Sussex: Wiley.

Berkeley County Development Authority. (2014) "Facts and Figures" Retrieved from http://www.developmentauthority.com/

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

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

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.

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

Bockstael, N. E., & Bell, K. P. (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). Norwell, MA: Kluwer Academic Publishers.

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.

Clark Labs. (2012). IDRISI Selva, Clark University http://www.clarklabs.org/

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.

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.

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.

DOT (Department of Transportation) Maryland Transit Administration. (2014) Retrieved from: http://www.mdot.maryland.gov/

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.

EPA (United States Environmental Protection Agency). (2007). Multi-resolution land characteristics consortium (MLRC). Retrieved from http://www.epa.gov/mrlc/definitions.html

ESRI. (2014). ArcGIS Help 10.2. Retrieved from: http://resources.arcgis.com/en/help/main/10.2/index.html

Fragkias, M., & Seto, K. C. (2007). Modeling urban growth in data-sparse environments: a new approach. *Planning and Design, 34*(5), 858–883.

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

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*, 195–208.

Goodspeed, R. (2007). Leapfrog' sprawl in West Virginia. Retrieved from http://goodspeedupdate.com/2007/2104

Hagerstrand, T. (1965). *A Monte Carlo approach to diffusion* (pp. 43–67). VI: Archive of European Sociology.

Hatten, M., Lapp, J., Bennett, D., & Stottlemyer, D. (2011). *WV stream and wetland valuation (SWVM) metric development*. Lexington, KY: Appalachian Stream Mitigation Workshop.

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., et al. (2007). Completion of the 2001 national land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing, 73*(4), 337–341.

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

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.

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.

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.

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 and Water Environment, 6*, 275–284.

Karigomba, W. (2009). *A spatial optimization approach to watershed water quality management*: *A case of the Opequon watershed.* (PhD dissertation). 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.

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.

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.

Louis, J. S., & Raines, L. G. (2003). Genetic Algorithm Calibration of Probabilistic Cellular Automata for Modeling Mining Permit Activity, Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence, p.515, November 03–05, 2003

Manson, S. M. (2001). Simplifying complexity: a review of complexity theory. *Geoforum, 32*(3), 405–414.

Mejıa, 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.

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.

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

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.

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

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.

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.

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.

Peterson, B. J., Wollheim, W. M., Mulholland, P. J., Webster, J. R., Meyer, J. L., Tank, J. L., et al. (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. G., Parker, D., Brown, D., & Grimm, V. (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., & Neeti, N. (2010). Uncertainty in the difference between maps of future land change scenarios. *Land use and Ecosystems, 5*, 39–50.

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.

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., & Cappiella, 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: Salavat Abad case study, Kurdistan, Iran. *Natural Hazards, 64*(2), 1639–1656.

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.

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.

Strager, M. P., Fletcher, J. J., Strager, J. M., Yuill, C. B., Eli, R. N., Todd Petty, J., et al. (2010). Watershed analysis with GIS: the watershed characterization and modeling system software application. *Computers & Geosciences, 36*(7), 970–976.

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.

U.S. Census Bureau. (2000). Population density by census tract, 2000. Retrieved from http://www.census.gov/

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

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*, 185–199.

Von Thünen, J. H. (1826). In P.G. Hall e. (Ed.), *Die isolierte staat in beziehung auf landwirtshaft 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

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.

WVDEP (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/nonptsource/WBP/Documents/WBP/MillCreekOpequon_WBP.pdf

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

White, R., & Engelen, G. (1993). Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment & Planning A, 25*(8), 1175–1199.

White, 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.

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.

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.