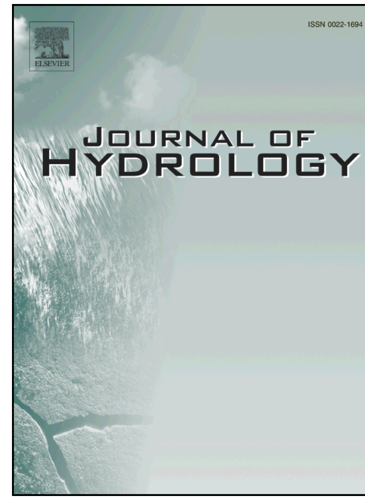
Accepted Manuscript



Research papers

Water Security Assessment of Current and Future Scenarios through an Integrated Modeling Framework in the Neshanic River Watershed.

Subhasis Giri, Nazia N. Arbab, Richard G. Lathrop

PII: S0022-1694(18)30373-1
DOI: https://doi.org/10.1016/j.jhydrol.2018.05.046
Reference: HYDROL 22823

To appear in: *Journal of Hydrology*

Received Date: 17 November 2017
Revised Date: 18 May 2018
Accepted Date: 20 May 2018

Please cite this article as: Giri, S., Arbab, N.N., Lathrop, R.G., Water Security Assessment of Current and Future Scenarios through an Integrated Modeling Framework in the Neshanic River Watershed., *Journal of Hydrology* (2018), doi: https://doi.org/10.1016/j.jhydrol.2018.05.046

Title Page

Water Security Assessment of Current and Future Scenarios through an Integrated Modeling Framework in the Neshanic River Watershed.

Subhasis Giri, Nazia N. Arbab, Richard G. Lathrop

Postdoctoral Research Associates and Professor, respectively in the Department of Ecology, Evolution, and Natural Resources, School of Environmental and Biological sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ-08901, USA.

Corresponding Author: Subhasis Giri, Postdoctoral Research Associate, Department of Ecology, Evolution, and Natural Resources, School of Environmental and Biological sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ-08901, USA, Email: Subhasis.giri@rutger.edu, Phone:+1(848)932-5573, Fax:+1(732) 932-2587

Water Security Assessment of Current and Future Scenarios through an Integrated Modeling Framework in the Neshanic River Watershed.

ABSTRACT

Water security assessment based on the concept of blue versus green water is becoming widely accepted globally. Blue water is the combination of surface runoff and deep aquifer recharge while green water is the summation of evapotranspiration and soil water content (i.e. mediated by plants). Due to the tight coupling between land use and the partitioning of blue and green water within a watershed, an integrated geospatial modeling framework that links land use and watershed hydrological processes is needed to predict the consequences of future land use change on blue versus green water security. By loosely coupling an agent based probabilistic land use change model with a hydrologic model, we investigated the consequences of present trends of urban growth to identify potential future hotspots of hydrological change across a watershed in central New Jersey undergoing suburbanization. The agent based probabilistic model, working at the scale of land ownership parcels, predicted that future urban development would be the result of forest, rather than farm land conversion. Using existing zoning maps, the housing unit density and human population of the urban growth areas was estimated. A consequence of the loss of forest land and increasing impervious surface leading was higher blue water but lower green water. While no severe blue water scarcity was observed, an increasing green water scarcity was found in some study area sub-basins. Such information will aid watershed managers' and policymakers' effort in sustainably managing water resources under changing land use and climate.

KEY WORDS

Water security, water scarcity, water withdrawal, blue water, green water flow, green water storage, soil and water assessment tool, agent based modeling, hydrological hotspot, zoning

1. Introduction

Availability of freshwater is critical to provide food security as well as safe drinking water to a rapidly growing population (Schuol et al., 2008). Adequate amounts of freshwater are also vital for supporting aquatic biodiversity (Harrison et al., 2016). However, freshwater is often limited in quantity and is vulnerable to human activities such as urbanization, agriculture, reservoir formation, irrigation, and inter-basin transfer (Vorosmarty et al., 2010). Ensuring the availability of freshwater is complicated by changing precipitation patterns, longer dry spells, higher evapotranspiration, and increased frequency of extreme events such as droughts and floods (Bogardi et al., 2012). Water scarcity, defined as the ratio of water demand over water availability, provides an indicator of the water stress within a system. The present water scarcity that is prevalent in many parts of the globe is expected to increase in the coming years due to population growth and climate change.

Water security is referred as "*an acceptable level of water related risks to humans and ecosystems, coupled with the availability of water of sufficient quantity and quality to support livelihoods, national security, human health, and ecosystem services*" (Bakker, 2012). Water security for humans addresses multi-faceted requirements starting from domestic water to maintenance of ecosystems and biodiversity. Water security assessment based on the concept of blue versus green water is becoming widely accepted globally (Veettil and Mishra, 2016; Rodrigues et al., 2014; Schuol et al., 2008). Blue water flows over or under the soil surface and is generally stored in rivers, lakes, aquifers, reservoirs, and wetlands. Green water derives from precipitation and is stored in the upper layers of the soil or vegetation and returns to atmosphere through evapotranspiration (Rodrigues et al., 2014; Zang et al., 2012). Green water can be further divided into green water flow and green water storage. Green water storage is the amount of soil moisture whereas green water flow is the sum of actual evapotranspiration (Schuol et al., 2008). In other words, blue water is the combination of surface runoff and deep aquifer recharge while green water is the summation of evapotranspiration and soil water content (i.e. mediated by plants) (Faramarzi et al., 2008).

The differentiation of blue versus green water in the assessment of water security reflects the growing recognition for the explicit consideration of the water required for biomass production such as food and timber. The water requirement for biomass production surpasses all other water dependent processes (Falkenmark and Rockstorm, 2006). Both blue and green water are directly or indirectly related to human consumption. For example, water withdrawal from a river for drinking purposes is the direct use of blue water for human consumption while use of soil water for crop production represents use of green water for human consumption. Consequently, management of both blue and green water is helpful in identifying the geographic hotspots in the watershed with limited freshwater availability. To help balance competing demands such as maintaining environmental flows, satisfying water requirements for agriculture, and water withdrawal for drinking and other purposes, water resource managers often rely on spatially distributed dynamic process models. Many approaches such as Crop-Wat decision support tool (Allen et al., 1998), data generalization (Shiklomanov and Rodda, 2003), general circulation models (Oki and Kanae, 2006), WASMOD-M (a conceptual water budgeting model) (Widen-Nilsson et al., 2007), LPJmL model (global vegetation and water balance model) (Rost et al., 2008), are available in the public domain to quantify blue and green water. In particular, the Soil and Water Assessment Tool (SWAT), due to its process based structure and spatially distributed implementation, has found widespread application in the evaluation of the spatiotemporal variation of blue and green water in various river basins across the globe (Faramarzi et al., 2008; Schuol et al., 2008; Zang et al., 2012; Rodrigues et al., 2014; Zuo et al., 2015; Veettil and Mishra, 2016). SWAT has the ability to model hydrology, plant growth related processes, and incorporate different agricultural management practices which are essential from water balance prospective (Giri et al., 2014). It can also be used to model nutrient runoff and water quality impacts from both agricultural and urban land uses (Qiu and Wang, 2014; Panagopoulos et al., 2015; Giri 2016 a).

The availability of both blue and green water are directly or indirectly affected by how humans are using the land. The latter half of the 20th into the first decade of the 21st Century, exhibited dramatic population growth and suburbanization in exurban areas across the United States adjacent to, but outside of, traditionally defined metropolitan areas (Qian, 2010; White et al., 2009; Alig, 2010). In these non-metro counties as well as the outer counties within metro regions, forest and agricultural lands were converted into residential and commercial land uses. While this rapid growth cooled during the recession, recent annual population estimates (2015-2016) suggest that this period may simply be a short-term interruption in suburbanization with outlying counties in metro areas once again growing faster than central urban counties (US Economic Research Service, 2017). Stated differently, while suburban sprawl and rural land conversion may have slowed, it isn't dead yet. More importantly, the legacy of past land use change continues to have implications for blue and green water resources in these exurban/suburban areas. Urban development leads to increasing impervious surface as a result of increase in streets, parking lots, sidewalk, driveways, and rooftop. Consequently, the watershed surface runoff is increased by increasing curve number (CN), streamflow, flood volume, and having shorter time of concentration. In contrast, groundwater recharge is decreased due to lesser infiltration as a result of increasing impervious surface. The change in both surface runoff and infiltration affects both blue and green water components. Therefore, understanding the consequences of past urbanization and the potential future land use conversion on blue and green water is requisite for better water resources management.

To analyze past land use conversion and to project the pace and pattern of future change, land use researchers and planners have turned to a variety of modelling approaches including spatial econometric, cellular automata and agent-based models (Bockstael, 1996, Parker and Meretsky, 2004; Polhill et al., 2008). An agent-based model (ABM) is a simulation modeling framework where various behaviors of individuals or agents are programmed through their interactions and choice principles (Gimblett et al., 2002; O'Sullivan et al., 2012). ABMs are used to predict land use pattern by incorporating spatial variations and interactions among different components or agents of the system (Heppenstall et al., 2012) as well as the inherent properties of the land (i.e., slope, soil type, productivity, property parcel size, accessibility, and buildable structures). ABMs can incorporate spatial context as well as the stochasticity and uncertainty in generating land use patterns to help mimic the complexity of coupled land use-watershed systems. As many land use decisions are made at the level of the individual ownership parcel, property parcels are suitably regarded as agents (Le et al., 2010). Applying an

ABM framework within geographical information system (GIS) greatly facilitates the modeling of spatially explicit land use decision making at property parcel or plot scale (Najlis and North, 2004; Johnston, 2013) and linking land use policies to potential land use patterns (d' Aquino et al., 2015). Additionally, ABM can also provide means to link changing land use to changes in water quality at a watershed scale (Arbab et al., 2016).

Due to the tight coupling between land use and the partitioning of blue and green water within a watershed, an integrated geospatial modeling framework that links land use and watershed hydrological processes is needed to predict the consequences of future land use change on water security. Such a framework would also enable a more thorough investigation of the spatio-temporal variation of blue and green water. By incorporating both current and future land uses in SWAT model along with ABM modeling framework, this study addresses following key questions: i) how does the spatial distribution of blue and green water vary across a Central New Jersey watershed under current and projected land use, and ii) how does the water security situation potentially change under future land use and population?

2. Methodology

2.1. Study Area

The Raritan Basin located in Central New Jersey (Fig.1) has increased population and urban land by 25.6 percent and 35.8 percent, respectively since 1990 to 2010 (Giri et al., 2016 b). The Neshanic River Watershed (NRW) is the headwater to Raritan Basin (hydrologic unit code is HUC 02030105030). While the NRW is currently less developed, regional development and population growth trends suggest that this area will be subject to increasing development pressures in the coming years. This increasing population and development trend combined with sporadic drought warning category based on water supply by the New Jersey Department of Environmental Protection (NJDEP) prompted us to select NRW as study area for water security analysis.

Out of total NRW watershed area (142 km^2), 39 percent consists of agricultural lands, 32 percent contains forest, 28 percent comprises of urban lands, and remainder are wetlands, barren land, and water. The elevation of the NRW ranges from 20 m to 201m above mean sea level. The NRW belongs to Piedmont physiographic region where the topography is primarily low rolling plains divided by higher ridges underlain by folded and faulted sedimentary and igneous rocks. The soils are well drained silty soil where major crops including hay, pasture, corn, and soybean are observed. Minor crops grown in the watershed are winter wheat, rye, and oats. The climate in the watershed varies from hot and humid summer to cold winter representing humid climate.

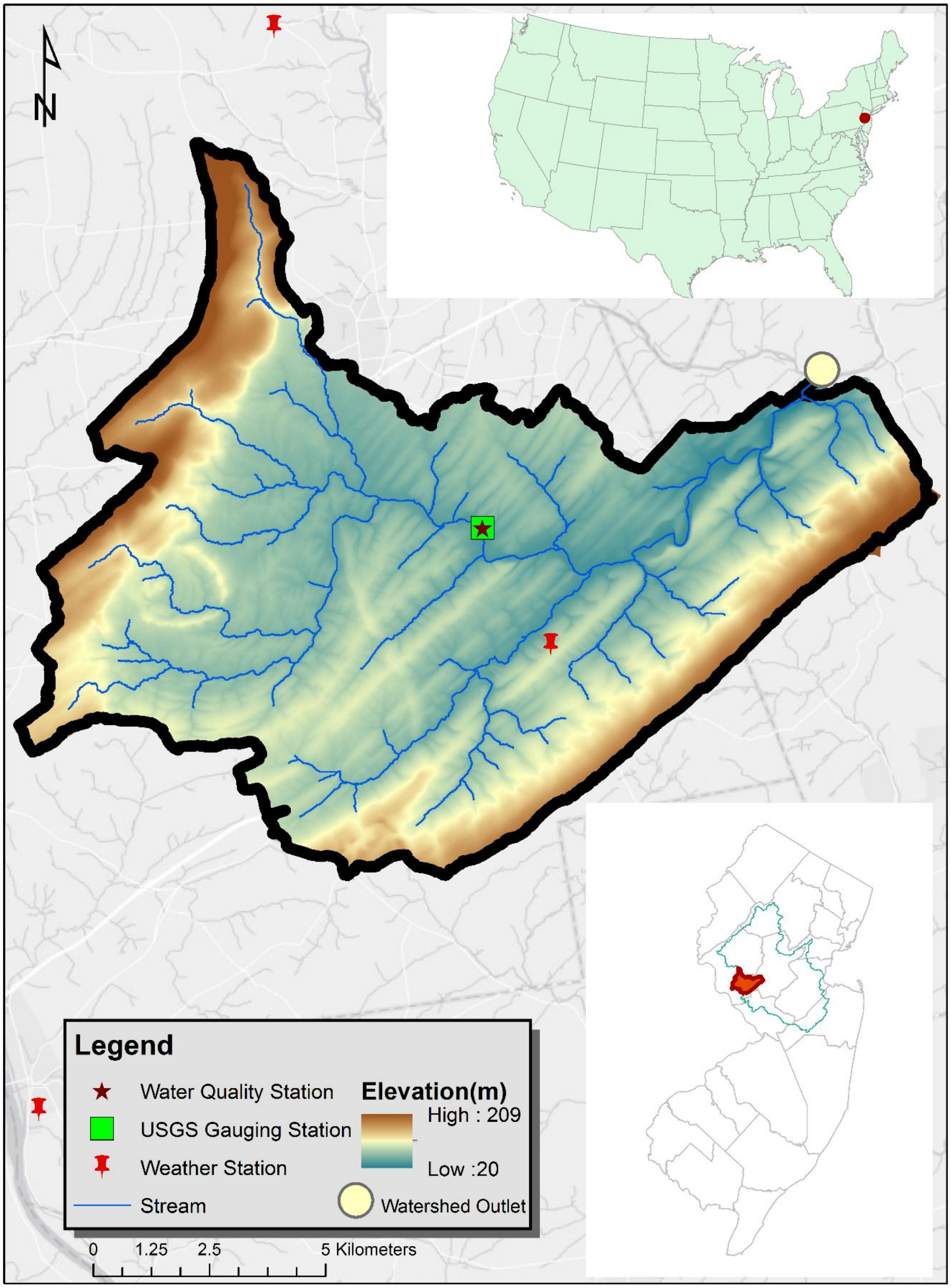


Fig. 1. Study area location. Red point on the top and bottom maps represent the study location with reference to United States and New Jersey, respectively.

2.2. Integrated Modeling Framework

The integrated modeling framework for this water security study consists of three primary components: 1) Hydrological modeling framework, 2) Agent-based modeling framework, and 3) Spatial data analysis framework (Fig. 2). The following section describes in detail of each modeling framework, data requirement, data source, and procedure.

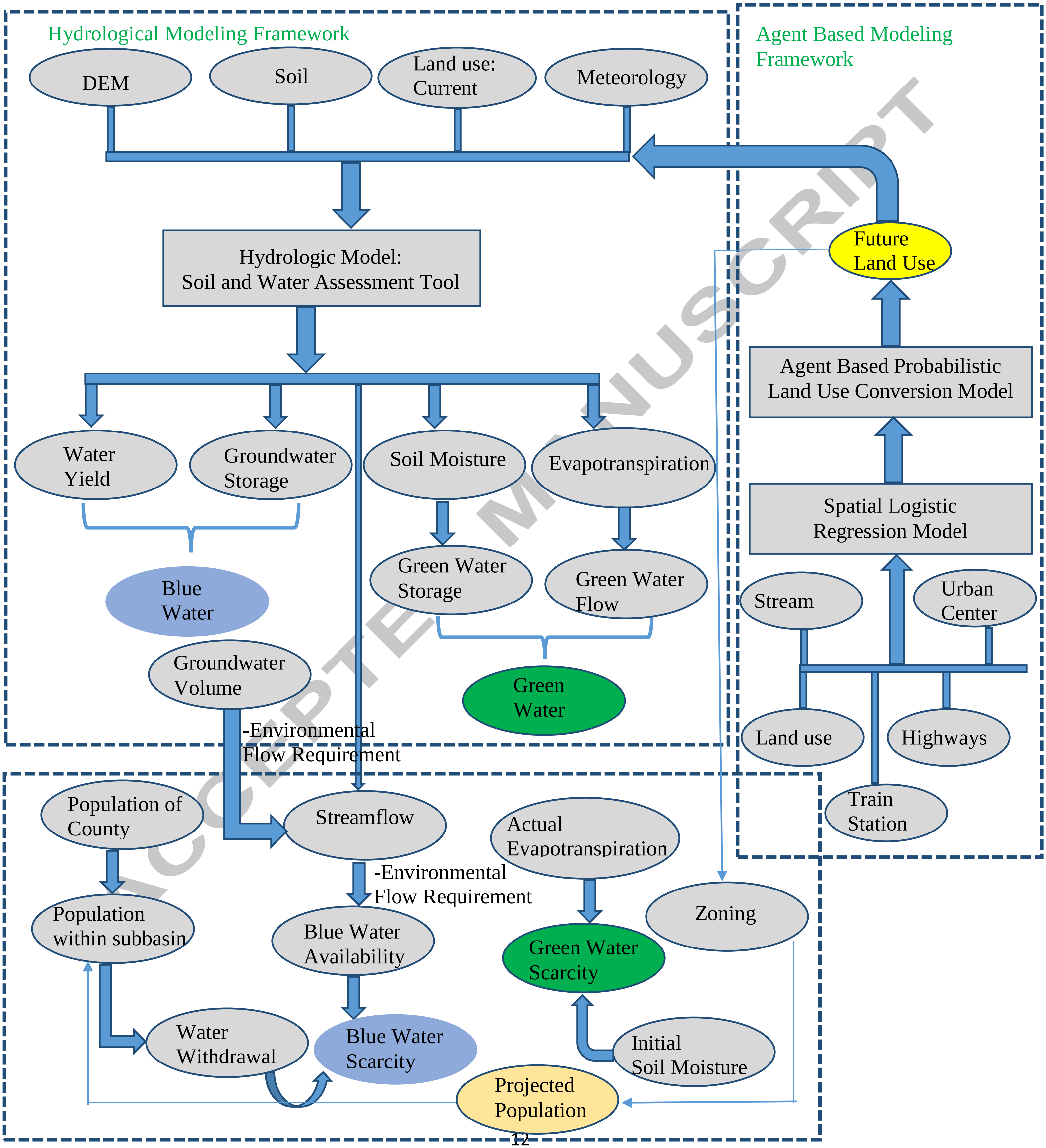
2.2.1. Hydrologic Modeling Framework

The SWAT model was used to simulate the hydrologic processes in the study area and calculated different components of blue and green water. SWAT is a physically based spatially distributed watershed scale model that simulates continuous daily time step streamflow, sediment, nutrients, chemicals, and pathogens transport in a watershed after incorporating different input data such as digital elevation model, land use land cover, soil, and meteorological data (Arnold et al., 1998; Neitsch et al., 2004).

SWAT divides the whole watershed into smaller units known as subbasin and it is based on the topography so that flow within each subbasin can direct to a single point called subbasin outlet. Subbasins are further partitioned into smaller unit known as hydrologic response unit (HRU) which has unique land use, soil, slope, and management. The major components of SWAT model are hydrology, soil, plant growth, weather, nutrients, pesticides, and land management practices (Gassman et al., 2007). Simulated processes include runoff, infiltration, evaporation, plant uptake, lateral flow, and percolation to shallow and deep aquifer. In this study, ArcSWAT version 2012.10_3.18 was used in ArcGIS 10.3 environment and NRW was divided into 115 subbasins and 9,291 HRUs.

2.2.1.1. Data Sources

SWAT model requires different types of data such as physiographic data, streamflow, and meteorological data for hydrology simulation in the watershed. A detail description of all the data and their sources used in the SWAT model can be found in Table 1. The land use of the study area was represented by NJDEP 2012 land use land cover data (NJDEP, 2012). The NJDEP land use data was categorized into six broader categories such as urban, wetlands, agricultural lands, water, and barren land while it has 51 subcategories based on modified Anderson Classification System (NJDEP, 2012). The 16 subcategories of urban land were assigned to the SWAT urban land category based on the percent of impervious surface. To identify specific cropping pattern within agricultural land, 2015 Crop Data Layer from USDA National Agricultural Statistics Service was used (NASS, 2015). Both land uses were combined in ArcGIS environment to process the data.


Hydrological Modeling Framework
DEM
Soil
Land use: Current
Meteorology
Hydrologic Model: Soil and Water Assessment Tool
Water Yield
Groundwater Storage
Soil Moisture
Evapotranspiration
Blue Water
Green Water Storage
Green Water Flow
Groundwater Volume
Green Water
-Environmental Flow Requirement
Agent Based Modeling Framework
Future Land Use
Agent Based Probabilistic Land Use Conversion Model
Spatial Logistic Regression Model
Stream
Urban Center
Land use
Highways
Train Station
Population of County
Streamflow
Actual Evapotranspiration
-Environmental Flow Requirement
Population within subbasin
Blue Water Availability
Zoning
Green Water Scarcity
Water Withdrawal
Blue Water Scarcity
Initial Soil Moisture
Projected Population

Spatial Data Analysis Framework

Fig. 2. Schematic of integrated modeling framework for water security analysis in the Neshanic River Watershed.

Table 1
A detail description of SWAT model input data for water security study in the Neshanic River Watershed.

Data Type	Description	Resolution	Source
Topography(raster)	Digital elevation model	3 m × 3m	New Jersey Department of Environmental Protection(NJDEP)
Land use or land cover(polygon)	2012-NJDEP land use	Converted to 30 m × 30 m	New Jersey Department of Environmental Protection(NJDEP)
Land use land cover(raster)	The Crop Data Layer	30 m × 30 m	United State Department of Agriculture (USDA)
Soil	Soil Survey geographic Database (SSURGO)	-	USDA- Natural Resources Conservation Services (NRCS)
Meteorological Data	Precipitation, maximum and minimum air temperature	Daily(mm)	National Climatic Data Center(NCDC)
Streamflow	Discharge	Daily(m^3/sec)	United States Geological Survey(USGS)
Management Practices	Tillage, fertilization, harvesting of different crops	-	Consultation with NRCS staff and local farmers

Meteorological data consisted of precipitation, temperature, solar radiation, relative humidity, and wind speed. Daily precipitation and temperature data were obtained from the National Climatic Data Center for 16 year time period (2000-2015) for three weather stations. One station was inside the study area while other two stations were located closer to the NRW edge (Fig.1). The remaining meteorological data were estimated using SWAT weather generator. Daily streamflow data for USGS gauging station 01398000 was obtained and converted into monthly basis to calibrate and validate the streamflow in the watershed (Fig.1).

To process the hydrology correctly, different management operations for different crops such as corn, soybean, winter wheat, hay, pasture, rye, and oats were developed. The management operation includes date and types of tillage equipment used, amount and date of fertilization, planting, and harvesting. These information were collected after series of interview with NRCS personnel and local stakeholders in the study area. In SWAT model, urban land has two parts: 1) impervious and 2) pervious part. In the pervious part, lawn management consists of mowing and fertilizer application were incorporated and these information was obtained from landscape professionals and residents inside the study area.

2.2.1.2. Sensitivity Analysis and Model Calibration

Sensitivity analysis performed by two prior studies (Giri et al., 2016a ; Qiu and Wang, 2014) was conducted to identify the most influential parameters with respect to model output. Calibration was performed using manual calibration helper in SWAT model using the sensitive parameters as well as understanding of the hydrologic characteristics of the study watershed. The calibration parameters used for streamflow were baseflow alpha factor, deep aquifer percolation fraction, surface runoff lag coefficient, hydraulic conductivity in the main channel, groundwater delay time, groundwater revap coefficient, soil evaporation compensation factor, curve number, available water capacity of the soil layer, maximum canopy storage, manning's n for overland flow, manning's n for main channel, threshold depth of shallow aquifer, threshold depth for revap, snowmelt base temperature, and snowpack temperature lag factor. These parameters were manually adjusted within the lower and upper bounds. The calibrated value of these parameters are presented in supplementary material (Table S1).

Validation was performed followed by model calibration to establish model reliability. A two years (2002 and 2003) warm up period was used to initialize the model parameters. Model calibration was performed on a monthly time step between 2004 to 2009 and validation was conducted between 2010 to 2014. Three model evaluation parameters Nash-Sutcliff efficiency (NSE), Percent bias (PBIAS), and Root mean square error (RMSE)-observations standard ratio (RSR) were used to evaluate model prediction. The NSE compares the observed and predicted data in terms of residual variance to measured data variance (Nash and Sutcliffe, 1970). PBIAS measures the quantitative difference of simulated data in terms of smaller or larger compared to observed value (Gupta et al., 1999) where RSR is the ratio between RMSE and standard deviation of measured data (Moriasi et al., 2007). The detail about NSE, PBIAS, and RSR can be found from Giri et al.(2012) and Moriasi et al.(2007).

2.2.2. Spatial Data Analysis Framework

Spatial data analysis framework was used to calculate both blue and green water scarcity in the study area. All the processes were performed in ArcGIS environment. The following section describes the calculation of blue and green water scarcity associated with other components.

2.2.2.1 Analysis of blue and green water

Blue water was calculated as the summation of water yield and groundwater storage (Fig. 2). Water yield is the total amount of water leaving the HRU and entering the main channel as streamflow and SWAT output file WYLD was used. Groundwater was calculated as the difference between amount of water entering to both shallow and deep aquifer, and the groundwater discharge into stream (Veettil and Mishra, 2016; Rodrigues et al., 2014). The SWAT output file used for water entering to both shallow and deep aquifer, and the groundwater discharge into stream are GW_RCHG and GW_Q, respectively. Green water was estimated as the summation of soil moisture and evapotranspiration (Fig.2) (Veettil and Mishra, 2016; Rodrigues et al., 2014; Schuol et al., 2008).

2.2.2.2. Blue Water Scarcity

Blue Water Scarcity is the ratio of the blue water required or withdrawn over the amount available (Veettil and Mishra, 2016; Rodrigues et al., 2014):

$$Blue\ water\ scarcity = \frac{Blue\ water\ requirement\ or\ withdrawal}{Blue\ water\ availability} \qquad (1)$$

A blue water scarcity values of less than one indicates no water scarcity. Blue water scarcity values of greater than one can be further categorized into mild or moderate or severe water scarcity based on the particular circumstances of the watershed.

Blue water withdrawal for this study was obtained from the USGS in the form of county level total water (surface and groundwater) withdrawal (USGS, 2017). The most recent (2010) total water withdrawal as well as total population for Hunterdon County (the county in which our study basin is located). The total water withdrawal was the summation of surface and groundwater withdrawal for different sectors such as public supply, domestic use, industrial purpose, irrigation, livestock, aquaculture, mining, and thermoelectric. Out of total water withdrawal, surface water contributed 86 percent while groundwater consisted of 14 percent. Among different sectors of water withdrawal, public supply and domestic use were the highest recipient of water from surface water and groundwater, respectively. Crop irrigation does not represent a significant use in this watershed. Total water withdrawal per person was achieved by dividing total water withdrawal to total population. To obtain the total water withdrawal within each subbasin, the most recent (2010) population data was obtained in the form of census tracts from the United States Census Bureau (USCB, 2017). The census tract was intersected by the study area boundary, however, the census tract boundaries did not exactly match with the subbasin boundaries. Therefore, percent of different census tract contributing to each subbasin was calculated to estimate the total population within each subbasin. The total water withdrawal within each subbasin was estimated by multiplying the total population within each subbasin with total water withdrawal per person.

Blue water availability for withdrawal should include environmental flow requirement. Blue water availability in stream is calculated based on the following equation (Veettil and Mishra, 2016; Rodrigues et al., 2014; Hoekstra et al., 2011):

$$Blue\ Water\ availability_{Stream} = Streamflow - Environmental\ flow\ requirment \qquad (2)$$

Where environmental flow requirement refers to minimum amount of streamflow required to maintain the downstream ecological integrity of most rivers or streams and is equal to 0.8 times of streamflow. Richter et al. (2012) developed this 80 percent stream flow protection standards to maintain ecological integrity of downstream aquatic ecosystems. This benchmark is also consistent with other studies such as by Brizaga et al. (2002) in Queensland, Australia as well as Carlisle et al. (2010) in the United States. A similar environmental flow requirement concept was also applied to groundwater to protect the potential impact of over exploitation of groundwater to many aquatic and riparian ecosystems (Gleeson and Richter, 2017). The blue water availability in groundwater was calculated based on the following equation:

$$Blue\ Water\ availability_{groundwater} = Groundwater - Environmental\ flow\ requirement \qquad (3)$$

Where environmental flow requirement is equal to 0.8 times of groundwater. The groundwater was calculated as the summation of area of each HRU within subbasin times the difference between amount of water entering to both shallow and deep aquifer, and the groundwater discharge into stream. The blue water availability was the summation of blue water in stream and groundwater calculated using equation 2 and 3. Annual streamflow and groundwater of 2010 predicted by SWAT model for each subbasin was used for blue water availability calculation. Streamflow and groundwater of 2010 was used for blue water availability calculation to maintain consistency with available total water withdrawal and population data.

2.2.2.3. Green Water Scarcity

Green water scarcity is an indicator which evaluates water security in a region. It is calculated using the following equation (Veettil and Mishra, 2016; Hoekstra et al., 2011):

$$Green\ water\ scarcity = \frac{Green\ water\ requirement\ or\ withdrawal}{Green\ water\ availability} \quad (4)$$

The green water scarcity values can be interpreted in a similar way as described for blue water scarcity. Green water requirement or withdrawal is the actual evapotranspiration which was obtained from the HRU output of the simulated SWAT model. Green water availability is the initial soil water content available for plant and soil evapotranspiration and it is calculated as the difference between soil moisture at the root zone and wilting point (Rodrigues et al., 2014). Initial soil water content was also obtained from HRU out of the simulated SWAT model and used for the green water scarcity calculation.

Temporal variation of precipitation, blue water, green water flow, and green water storage were generated using R-platform from monthly SWAT output during 2000 to 2015.

2.2.3. Agent based Modeling Framework

The ABM is based on two modules (spatial logistics regression (SLR) and agent-based probabilistic model (ABPM)) to illustrate the driving factors of land use change and land use conversions.

2.2.3.1. Spatial Logistic Regression Model

ABMs have incorporated values for neighboring land use impacts and these relative values are based on researcher postulated estimates (Parker and Meretsky, 2004; Brown et al., 2005; Polhill et al., 2008). One way to improve these ABM parameter estimates of neighborhood impacts is through an empirical parameterization using spatial land use change data. This empirical approach has the capability to quantify spatial externalities of neighborhood land uses and locational features using observed land use conversions (Arbab, 2014; Arbab et al., 2016). The SLR model is often used to provide spatially explicit empirical estimates of these neighborhood and proximate economic externalities (Arbab et al., 2016). The empirical structure of the logistic regression function describes a functional relationship between land use conversion and a set of explanatory factors that influence conversion probability for non-residential land (Xie et al., 2005). This empirical structure is used in characterizing the land use conversion decision for each property parcel in ABM. The SLR coefficients are derived from raster based data and these coefficients can be regarded as weights to produce the probability of conversion to developed land use (Shirzadi et al., 2012). The future land use patterns generated in ABM are conditional to the sequential land use conversions resulting in variation in neighboring land uses. This sequential process can best be defined by updating the probability of land use conversion influencing by proximity effect of neighboring features (Wu, 2002). The agents in ABM can be passive and social agents (Walsh et al., 2013). In land use conversion modeling, property parcels have been defined as passive agents and the type of each property parcel provides the further classification of these agents. Such as parcels that are available for residential conversions can be choice making unit for developers in ABM. Such system in this integrated framework provides the impact of land use conversion stemming from the local level conversion choice making of individuals.

Both SLR and ABM models account conversion in one way, from agricultural and forest properties to residential properties. A necessary condition in the ABM model for land use conversion from agricultural or forest land into residential land use is that, land values for residential use is higher than values for agricultural or forest use. This condition is set due to urbanization being the most important factor influencing farmland conversion (Olson and Olson, 1999; Rosenberger et al., 2002). Having this necessary condition, our research is grounded in two strands of theoretical approaches to explain residential land value: (1) land value as a function of proximity to economic locations, and (2) land value as a function of surrounding, location specific attributes.

To predict land use conversion probabilities from 2012 to 2022, a SLR model was used to calculate the probability of land use conversion from 2002 to 2012. A SLR model is developed in TerrSet software of Clarks Labs (TerrSet, 2016). The SLR model implementation in TerrSet requires spatial data in raster format. Therefore, raster cell is considered as a unit of observation in SLR.

The empirically estimated relationship between the non-residential land use conversion and the factors influencing the probability of conversion is expressed as the following functional form:

$$P\left(Y=1 \mid x\right)=\frac{exp\left(\sum \beta_i X_i\right)}{1+exp\left(\sum \beta_i X_i\right)} \tag{5}$$

where:

$P(Y = 1|X)$ = the predicted probability value of the binary or dichotomous dependent variable Y and Y=1 means a cell in raster map changes from a non-residential land use in 2002 to residential land use in 2012, otherwise Y=0.

X_i = independent variables including Xo for the constant term.

β_i = Coefficients of variables (parameters)

The logistic function takes into account the linear probability in a set of parameters, by having the range of probability between zero and one. The following linear logit transformation on both sides of equation (5) was used to estimate the coefficients (Menard, 1995):

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

where Y is the probability that the dependent variable is 1, p_k is the predicted probability of dependent variable of agricultural and forest land use conversion into residential land use. β_0 is the intercept, and β_1, β_2, β_3, β_4, β_5, and β_6 are coefficients for variables which reflect distance to the nearest: agricultural land use (x_1), forest (x_2), residential land use (x_3), stream (x_4), major highway (x_5), and distance to train station (x_6), respectively.

The sign of the parameter indicates the direction of the influence of each explanatory variable on conversion probability. A negative coefficient sign shows that an increase in distance between a cell and the neighborhood externality would decrease the probability of conversion. Conversely, a positive sign shows that as the distance of the cell to a land use or location feature increases, the probability of conversion would increase.

Following bid-rent theory (Alonso, 1964; Mills, 1967; Muth, 1969), the present ABPM modeling approach is based upon the assumption that each land parcel is allocated to the use that maximizes the utility of its owner. In the model, distance to urban center represents proximity to economic activity centers, schools, shopping centers, railway station, and public services (Kitamura et al., 1997). Distances to the major highways and urban center are conceptualized using Von Thünen and the bid-rent theory of urban economics where distance to the urban center and roads explains the land rent and transportation cost, respectively under the relaxed assumptions of spatial variation in the landscape (Von Thünen , 1826; Alonso, 1964; Mills, 1967; Muth, 1969). Distance to urban center is the major factor accounted in monocentric bid-rent theory (Alonso, 1964). As the distance from the urban center increases, accessibility decreases which results in higher transportation costs. Distance to roads represents accessibility to metropolitan and urban areas, workplace, shopping, and schools (Serneels and Lambin, 2001).

Spatial features surrounding property parcel impact the parcel value which is followed from the economic theory of hedonic property values (Wu, 2006). Distance to streams and forest are regarded as a relative measures for aesthetic amenities where closeness to streams and forests determines the value of parcel-level neighboring characteristics. Since streams and forest represent amenities, therefore, they are expected to positively influence the probabilities of land use conversions (Irwin et al., 2014). The distance to agricultural and residential land use accounts for neighboring externalities.

2.2.3.2. SLR Model Evaluation

The SLR model was validated using the quantitative measurement of Receiver Operator Characteristic (ROC) statistic which has been used as a reliable approach for model validation (Pontius and Schneider, 2001; Arsanjani et al., 2013). The ROC predicts the location of conversion by comparing the actual change between 2002 and 2012 in a Boolean raster data and the suitability (fitted) change between 2002 and 2012. Several studies considered ROC as reliable statistic to measure the goodness of fit for a logistic regression (Swets, 1986; Pontius and Schneider, 2001; Verburg et al., 2002; Tayyebi et al., 2010). The ROC varies between 0 and 1, where 1 shows a perfect fit and 0.5 shows a random fit. Overall, larger ROC values show a better association between explanatory variables and the dependent variable.

A Chi-square distribution of likelihood ratio statistics was used to test the null hypothesis that all variables in the model measuring distances from surrounding land uses or amenities have no impact on whether a cell will be converted from non-residential to residential land use. Since ordinary least square (OLS) is not applicable due to the nonlinear functional form of logistic model (Wooldridge, 2002), therefore, maximum likelihood estimator (MLE) was used. MLE estimator is suitable for the distribution of Y given X. This estimator includes heteroscedasticity in VAR(Y/X) and is consistent, asymptotically normal, and asymptotically efficient (Wooldridge, 2002).

2.2.3.3. Agent-based Probabilistic Model (ABPM)

ABPM is programmed in a GIS environment using python programming. Property parcels are assigned with land use conversion rule based upon empirical parameters from SLR model. Agents in ABPM are defined as the developable parcel including forest and agricultural non-protected properties representing land owners' choice-making units or passive parcel agents. Based upon the majority of rural residential properties in 2012, the assumption in the model implies that developable parcels (i.e., upland forest and agricultural land) can only convert into rural residential parcels. Each agent parcel is autonomous by being owned and controlled by a single owner. In reality, a single owner may have multiple properties, but the same property owner may convert the property parcel in one location, without converting a property parcel owned at a different location. In ABPM, an agent's decision rule is formulated based on empirical rules of land use conversion. A similar framework where agents are characterized within a bounded rationality framework has been used in studies such as Benenson and Torrens (2004) and Valbuena et al. (2010).

In ABPM, neighboring externalities are used as proxy for implicit spatial interactions and interdependencies. Neighborhood externalities were the estimated influence of each land use on surrounding parcels. The agents' conversion decisions is conditional upon the spatial distances from each neighboring land use over a period of 10 iterations (a 10 year time period), where each iteration is assumed to be a land use conversion event possibility. Initial conversions influences future conversions within each model run due to path dependent influence in decision making. With modification of Benenson and Torrens (2004) approach, parcel agent's probability (Prob) of conversion from developable state m to residential state r in each iteration is programmed similar to Arbab et al. (2016):

$$\mathrm{Pr}\,ob_i(S_m \to S_r) = S(N(i)) \tag{7}$$

Where N (i) represents parcel agent i's neighbors and S represents state of parcel i. Monte Carlo process (Hagerstrand, 1965; Wu, 2002) was employed to generate the results of a stochastic ABPM simulation model. To incorporate uncertainty and stochasticity in the ABPM, a probability function was used to condition the residential conversions utilizing a random number generator (Batty, 2012).

The conversion decision of agent in converting parcel into residential use was based upon a comparison between the random number generated and the probability value computed from equation (5) for each parcel as below (equation 8). In equation 8, the random number generator rand (αi) has a random distribution that is uniform between 0 and 1.

$$if\ \ \mathrm{rand}(\alpha_i) < P_{it} then\ \mathrm{A}_{it+1} = r \tag{8}$$

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 assess the probability of conversion by comparing it with a random number. Agents use *If* and *then* rule in parcel conversion. The agent converts the parcel into a residentially developed parcel if the probability is higher than the random number. If not, then the parcel remains in its current undeveloped state. This shows that probability which is driven by neighboring lands is indicator of parcels' conversion.

The model runs in discrete event steps and generates spatial output of projected residential converted parcels and non-converted parcels. A total of 10 iteration steps were included in each model run within ABPM. The number of iterations steps was matched with training land use data set used in the SLR model. The raster data in SLR consisted of 10 years of land use change from 2002 to 2012.

The landscape was initialized as the actual land use in shape file for the year 2012 in Arcpy program of python. First, agents identify whether the land parcel is developable or not. Then, agents compute the mean Euclidean distances from nearest agricultural, forest, and residential lands from their property parcel. After, the distances are calculated, agents identify the coefficient values for these spatial externalities and identify the distances from roads, train station, and streams. The coefficients of spatial externalities are imported in ABM from SLR model results. Parcel agents use the estimated SLR coefficients in their conversion decision and for each iteration, they calculate new values for the spatial externalities of neighboring lands due to changes in updated spatial patterns of land use parcel data in each iteration. Thus, as the parcel landscape changes, explanatory variables are recalculated by each parcel agent in the ABPM. Agents incorporate assessed land uses in its type, neighboring land uses, and features distances into conversion of land use in each iteration. The conversion decision due to uncertainty is probabilistic in ABPM.

Another factor in the residential land use conversions is generated probability value ranging from $0 \le Pit \le 1$, where i represents each parcel agent in each event t. The agents make their conversion decisions based on constant information feedback of distances and quantified spatial externalities in each model run in a continuously iterative fashion (Liu et al., 2013; Arbab et al., 2016). The conversion decision in ABPM 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 converted into stepwise probability. For 10 iterations the relationship between Pi (computed with SLR equation in ABPM) and Pa (a probability of conversion for each iteration) is formulated as:

$$P_i = 1 - [1 - P_a]^{10} \tag{9}$$

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

Pa is the average probability per iteration over the ten iterations and its value changes with each iteration due to changes in land use conversions. The stepwise probability ensures the final conversion probability for all 10 steps will match the 10 steps (10 year time period) of conversion probability from the SLR. The interaction among agents is implicitly defined by how changes in the neighboring land use affects land use conversion decisions in terms of probability. The sign of coefficients from SLR represent the type (negative or positive) of spatial externality for each parcel agent.

To account for the probabilistic nature of conversions, a Monte Carlo process of the ABPM model was used by simulating 100 model runs. Each model run generated a different set of spatial land use conversion distribution. Due to the defined empirical structure of local probability, model results showed fluctuations at consistent rate with 100 model runs. Therefore, the choice of 100 model runs is judged to be adequate for testing the path dependency and stochastic processes in ABPM.

Projections of land use conversions for 100 model runs are mapped and probabilities are assigned to all developable parcels. The probability of each parcel within 100 model runs, where each model has 10 iterations is calculated as:

$$P_j = \sum_{x=1}^{100} \frac{C_x}{100} \quad (11)$$

Where P_j is the probability of conversion for parcel j, x is the number of model runs, C is the Boolean conversion in each model run results in either one or zero, where one indicates conversion and zero represents no conversion. These probabilities are the Monte Carlo probabilities. Once the Monte Carlo probabilities are mapped, the threshold for probability is set to generate projected residential land use conversion data (Fragkias and Seto, 2007). Based upon several 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). This shows that the parcels which have at least a 0.50 probability of land use conversion are assumed as residentially developed parcels (value =1), while projected parcels with <0.50 probability are assumed as not converted parcels (value =0). A data generator step and a land use conversion step are performed in each iteration in ABPM. The projected spatial land use patterns from ABPM are further used as land use data layer for the SWAT model. The SWAT model was simulated with projected land use scenario for spatial analysis of blue and green water using previously calibrated model parameters.

2.2.3.3 Population Projection Using Zoning

In USA, the spatial land development is regulated through municipal zoning which depicts the acceptable types of future development. Consequently, the spatial map of zoning for Hunterdon and Somerset counties were collected from North Jersey Transportation Planning Authority (NJTPA, 2018). This zoning layer was used to estimate the available housing units that can be developed inside the study area. Data on housing units and population was extracted from U.S. Census data for the year 2010. We found 1 housing unit per acre and 3 person per housing unit based upon current zoning and population data, respectively. The zoning map was overlaid with projected 2022 new residential land use. Based on the housing unit density for different zoning categories, numbers of housing units (HU) were estimated. From per HU mean population, we estimated the additional population for the year 2022. The new population for each subbasin, which was the summation of the 2010 population of each subbasin and new projected population based on combination of zoning and projected residential land from ABPM model, was used to estimate the water scarcity for the year 2022.

3. Results and Discussions

3.1. SWAT Model Calibration and Validation

Model evaluation parameters (NSE, PBIAS, and RSR) are presented for monthly streamflow calibration and validation in the NRW for the water security analysis (Table 2). The overall modeling period was from 2004 to 2014 while the calibration period was between 2004 to 2009, and the validation period was between 2010 to 2014. Overall, the model is performing “Good” for streamflow due to 0.69(NSE), 5.60(PBIAS), and 0.56(RSR) based on the guideline developed by Moriasi et al. (2007).

The streamflow hydrographs of observed and simulated data is presented during calibration and validation period (Fig.3). In most cases, the observed streamflow is slightly higher than the simulated streamflow which suggests that SWAT model underestimates streamflow. These results may be due to higher percentage of urban area in the watershed that increases the impervious surface making the watershed flashier. A similar result was also observed by Giri et al. (2016 a) and Qiu and Wang (2014) in the same watershed.

Table 2
Model evaluation parameters for streamflow in the Neshanic River Watershed.

Constituent	Evaluation parameters	Overall period (2004 to 2014)	Calibration period (2004 to 2009)	Validation period (2010 to 2014)
Streamflow	NSE	0.69	0.66	0.71
	PBIAS	5.60	10.85	-1.46
	RSR	0.56	0.58	0.53

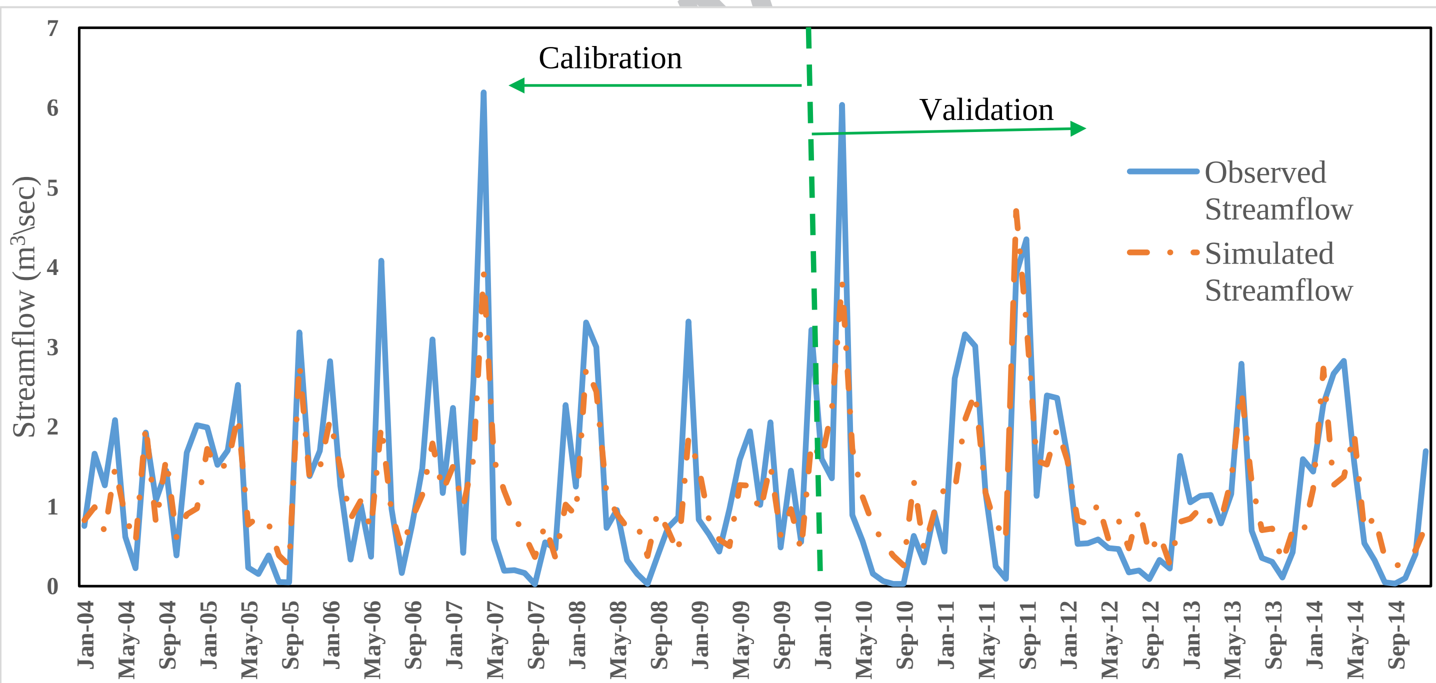


Fig. 3. Simulated and observed monthly streamflow at United States Geological Survey gauging station 01398000 in the Neshanic River Watershed.

3.2. Future Land use Conversion based on Agent-based Probabilistic Model

3.2.1. Spatial Logistic Regression

Model statistics for the SLR analysis are summarized in Table 3. There were 569,212 cells (sample observations) in raster map used in SLR, out of which 12,474 cells were converted from non-residential to residentially developed cells between 2002 and 2012. Statistical significance for the SLR model was found with a *p* value substantially less than 0.01. ROC was performed by comparing the fitted cells projected to convert with actual cells that did convert during 2002-2012. ROC represents the model's ability to predict the conversion probability at various locations in the study area (Tayyebi et al., 2010). The resultant ROC for the SLR model shows a high value of 0.7135 (Table 4).

The SLR coefficient are based upon the raster data and they can be regarded as weights to produce the global probability of change (Shirzadi et al., 2012). As the results in Table 4 indicate, the closer a developable parcel is to surrounding agricultural and forest land, the higher the probability of conversion. Positive coefficient signs were estimated for distances to residential parcels, streams, highways, and train stations which suggest that further the parcel is from these location features, higher the probability of conversion.

Table 3

Spatial logistic regression model results in Neshanic River Watershed.

Statistics	Value
Number of total observations	569212
Number or percentage of 0s in sampled area	556738
Number or percentage of 1s in sampled area	12474
Chi-square (6)	3215.70 (p-value 0.00001)
-2logL0	63120.82
-2log(likelihood)	59905.12

Table 4

Coefficient values for each explanatory variable in the Spatial Logistic Regression Model.

Explanatory Variables	Definition	Coefficient values
Agdist	Distance in meters to the nearest agriculture land	-0.0001

Foresdist	Distance in meters to the nearest forest land	-0.0303
Residdist	Distance in meters to the nearest residential land	0.0028
Streamdist	Distance in meters to the nearest stream	0.0062
Highwaydist	Distance in meters to the nearest highway	0.0044
Traindist	Distance in meters to the nearest train station	0.0017
Constant	-	-4.0889
ROC	-	0.7135

Note: ROC=1 indicates a perfect fit; and ROC=0.5 indicates a random fit.

3.2.2. Agent-based Probabilistic Model Results

The Monte Carlo projection in darker red color shows the area that are potential residential area during 2022 (Fig. 4). The land use conversion projection from ABM shows that most projected land converted from forest compared to agricultural land which may be due to forest being dominant land in the area. The total developable land in the study area consisted of 2,465 ha forest and 3,007 ha of agricultural land. Out of total 5,472 ha of developable land, approximately, 1,766 ha were converted into residential area. From the total newly developed residential area, 727 ha converted from agricultural land and the remaining 1,039 ha converted from forest (Table 5). Most of the predicted conversions captured in the model are close to the existing forest properties. This is the most predictable type of leapfrog residential growth pattern with current calibration from SLR model.

Table 5

Land converted from each developable land use type into residential area in the Neshanic River Watershed.

Land use converted	Residential land (in Hectare (ha))
From total developable land	1,766
From agriculture	727
From forest	1,039

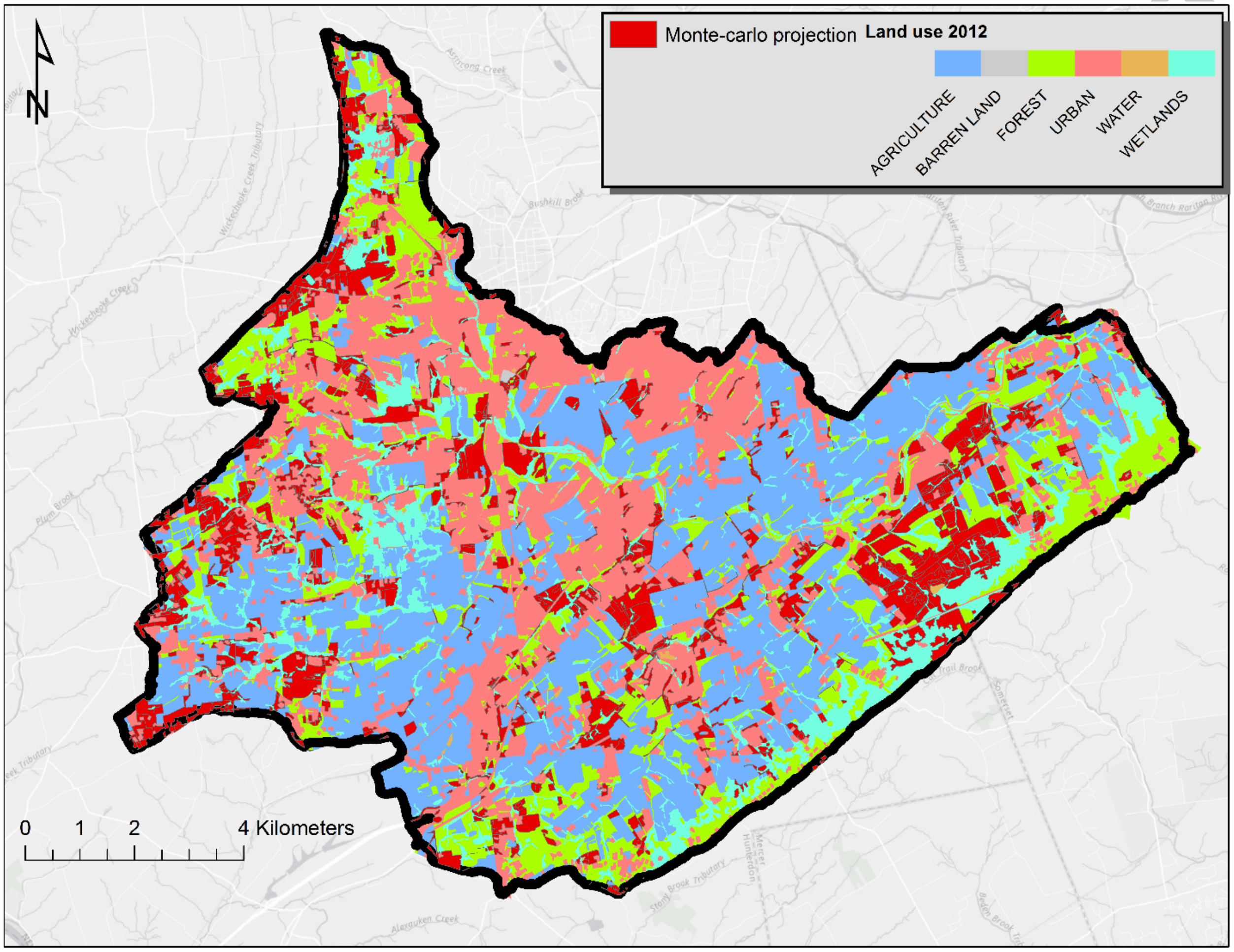


Fig. 4. Land use conversions based on Agent Based Probabilistic Model in the Neshanic River Watershed where darker red color represents the Monte Carlo projection of land converted to urban land uses between 2012 and 2022.

3.3. Water Security Analysis Based on Current Land Use and Population

Spatial variation of blue and green water in NRW is calculated based on the 16 years (2000 to 2015) of annual SWAT model output. The spatial distribution of blue water, green water flow, and green water storage follow expected patterns in relation to precipitation and land use (Figure 5 and Figure 6). Overall, blue water in the watershed varies between 35 to 48 percent of precipitation, green water flow ranges from 60 to 65 percent of precipitation, and green water storage varies between 8 to 17 percent of precipitation. The temporal variation of blue water, green water flow, and green water storage was primarily dependent on the evaporation. More detail description on temporal variation of blue water, green water flow, and green water storage is provided in the supplimentary material.

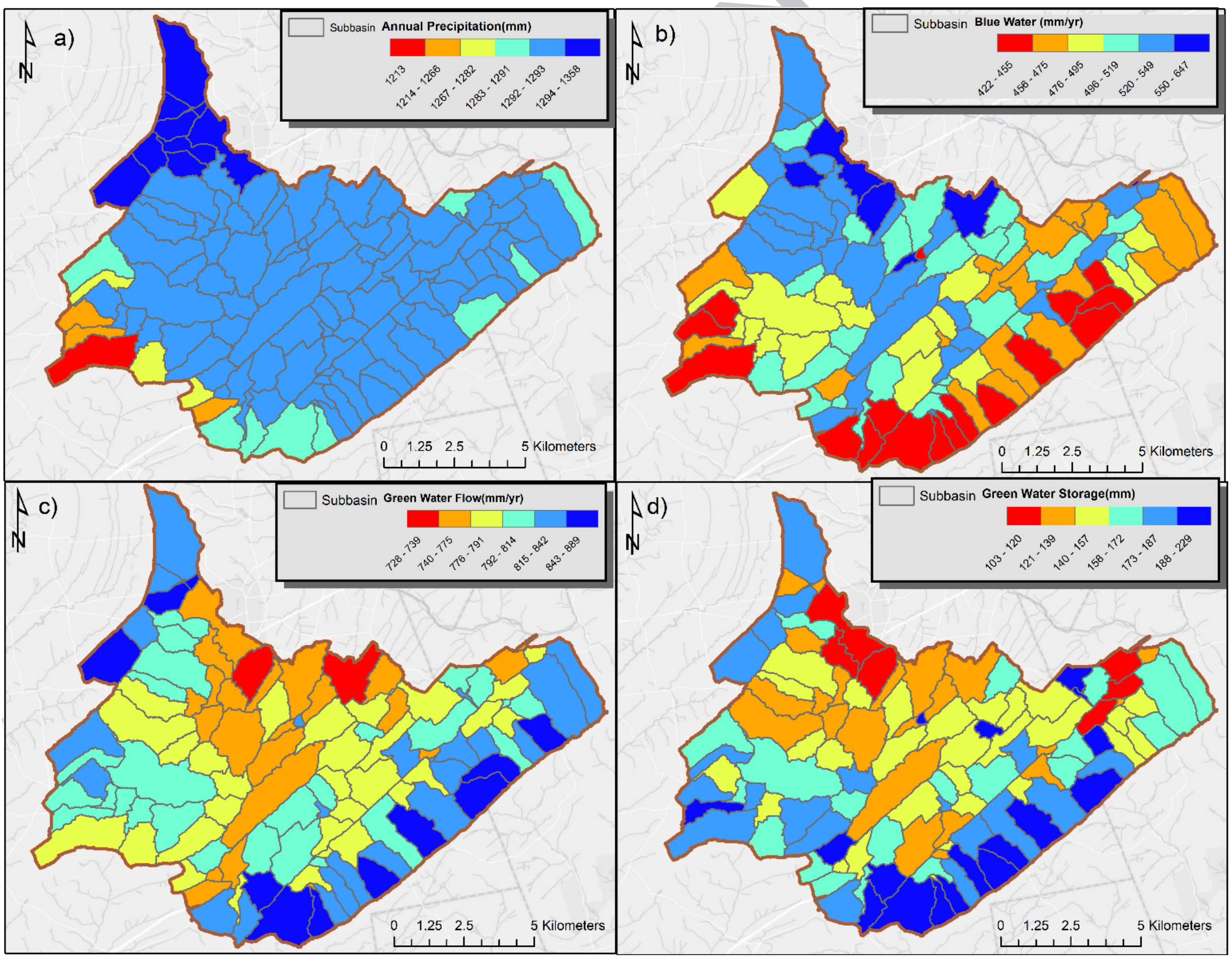


Fig. 5. Spatial distribution of (a) annual precipitation, (b) blue water, (c) green water flow, and (d) green water storage in the Neshanic River Watershed during 2000 to 2015.

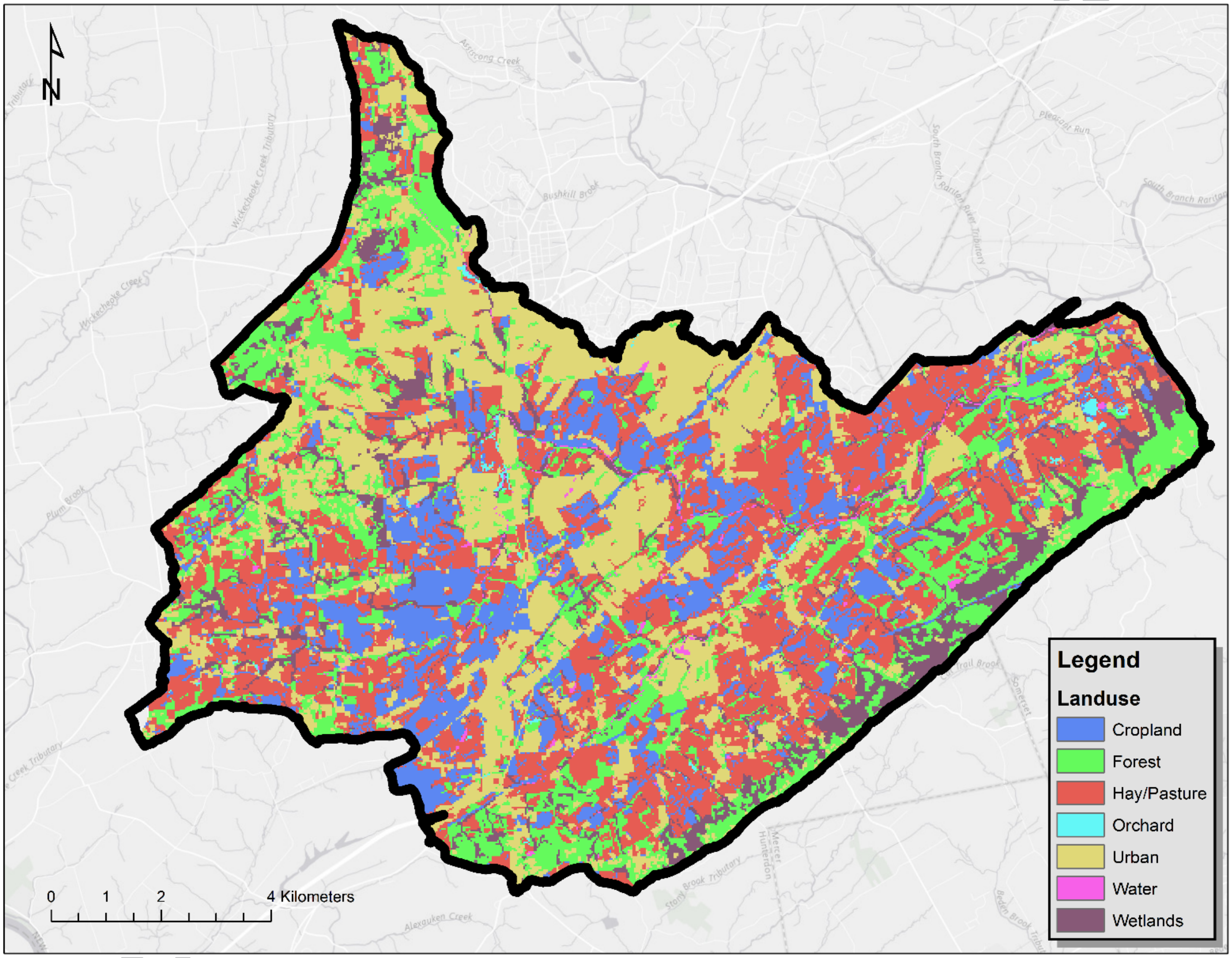


Fig. 6. Current land use land cover used in the water security analysis in the Neshanic River Watershed.

A water scarcity score of less than one indicates no water scarcity; scores greater than one were classed as mild or moderate or severe water scarcity based on the circumstances of the watershed. The spatial distribution of blue water scarcity varies from 0 to 5 (Fig. 7). The maximum blue water scarcity was predicted to occur on the north-east edge of the watershed along with some central subbasins associated with the most urbanized areas. The lower blue water scarcity was found in the central subbasins which are dominated by crop lands and pasture or hay (Fig. 6). The spatial distribution of green water scarcity varies from 3 to 15 (Fig. 8). The maximum green water scarcity was associated with areas of urban land use and higher intensity agricultural cropping. In urban areas, the rate of infiltration is less than other land uses due to presence of higher percentage of impervious cover leading to lower initial soil moisture and consequently, higher green water scarcity. The evapotranspiration in the crop land, hay or pasture land is likely to be higher leading to higher green water scarcity. The lowest green water scarcity is observed along the southeast edge of the watershed associated with the highest green water storage (Fig. 5 d).

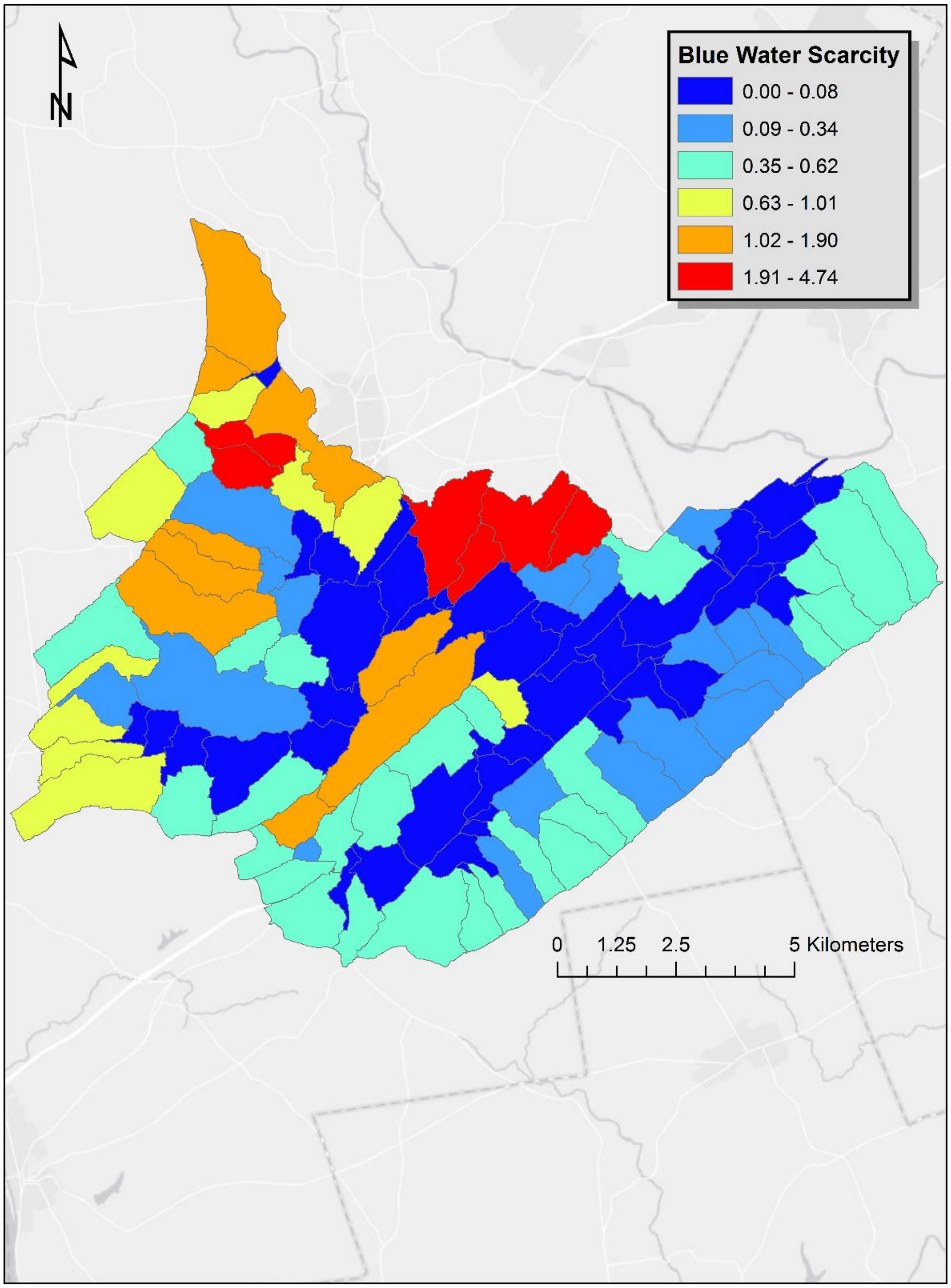


Fig. 7. Spatial distribution of blue water scarcity based on current land use and population in the Neshanic River Watershed.

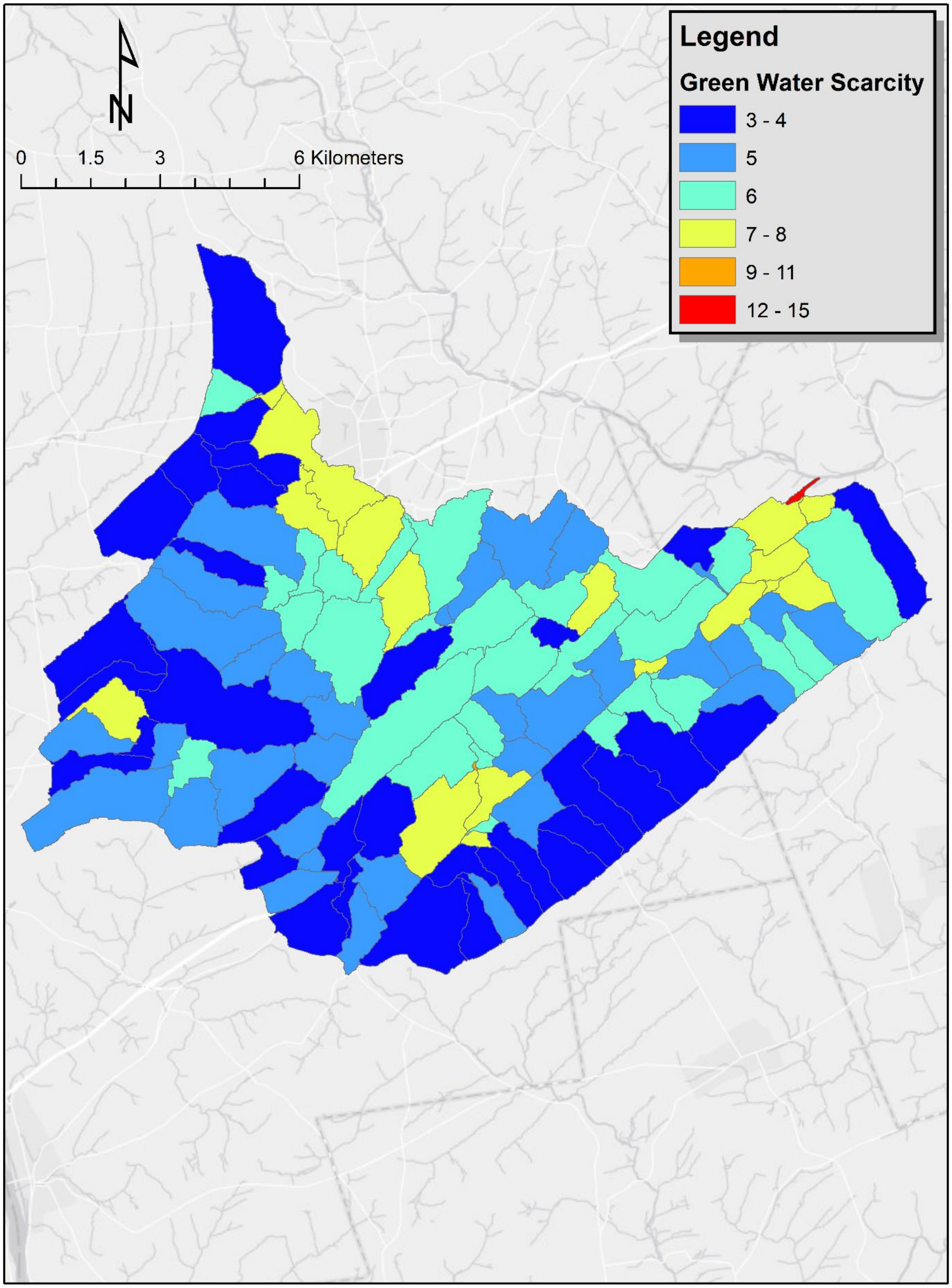


Fig. 8. Spatial distribution of green water scarcity based on current land use and population in the Neshanic River Watershed.

3.4. Water Security Analysis Based on Projected Land Use and Population

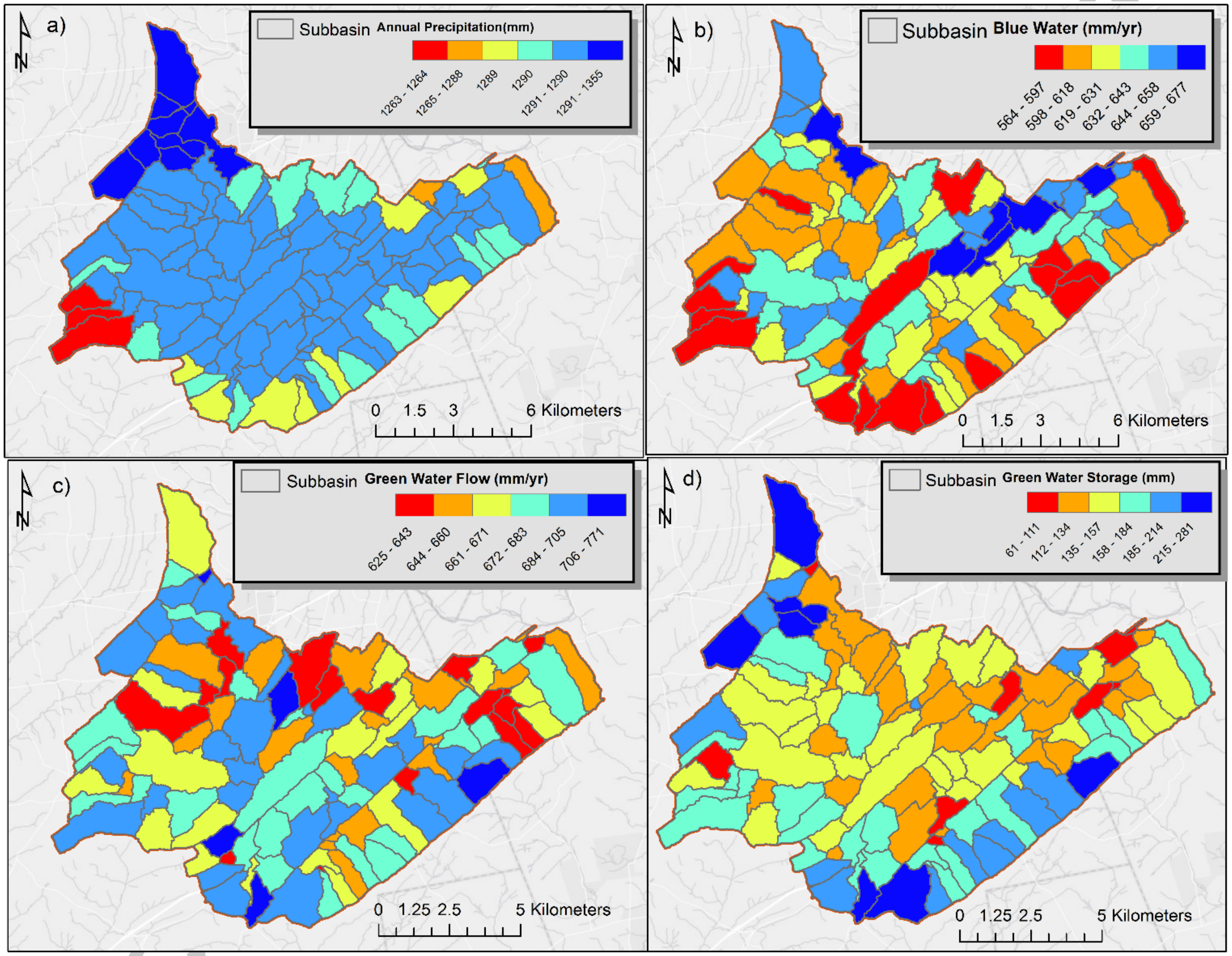


Fig. 9. Spatial distribution of (a) annual precipitation, (b) blue water, (c) green water flow, and (d) green water storage in the Neshanic River Watershed for 2022 land use conversion.

Based on the agent-based modelled land use as of 2022, the spatial distribution of blue water, green water flow, and green water storage was modelled (Fig.9). Overall, an increase in blue water was estimated for 2022 land use scenario compared to 2012 land use which may be due to increase in urban area (residential rea) leading to increased impervious surface and greater surface runoff. A decrease in green water flow and green water storage was predicted for the 2022 land use scenario compared to the 2012 land use. This may be due to lesser soil moisture availability due to lesser infiltration as a result of urbanization.

Even though increasing blue water was estimated in the projected land use and population scenario, blue water scarcity was predicted to increase slightly over the current situation (Fig. 10). This result may be due to increasing water demand in the residential areas due to a projected increase in population by nearly 100 percent. The projected population pattern follows the spatial distribution of land use conversion from the ABPM resulting into nearly doubling of population in the study area. The spatial distribution of increasing population is presented in the supplementary material (Fig. S5). Green water scarcity ranges from 2 to 10 (Fig. 11). Though green water flow and green water storage are expected to decrease somewhat under future conditions, increasing green water scarcity was observed in only a small number of study area sub-basins. However, if the increase in urban land trend continues, that may create problems for agricultural activities such as growing corn, soybean, and wheat in the watershed due to decreased green water storage.

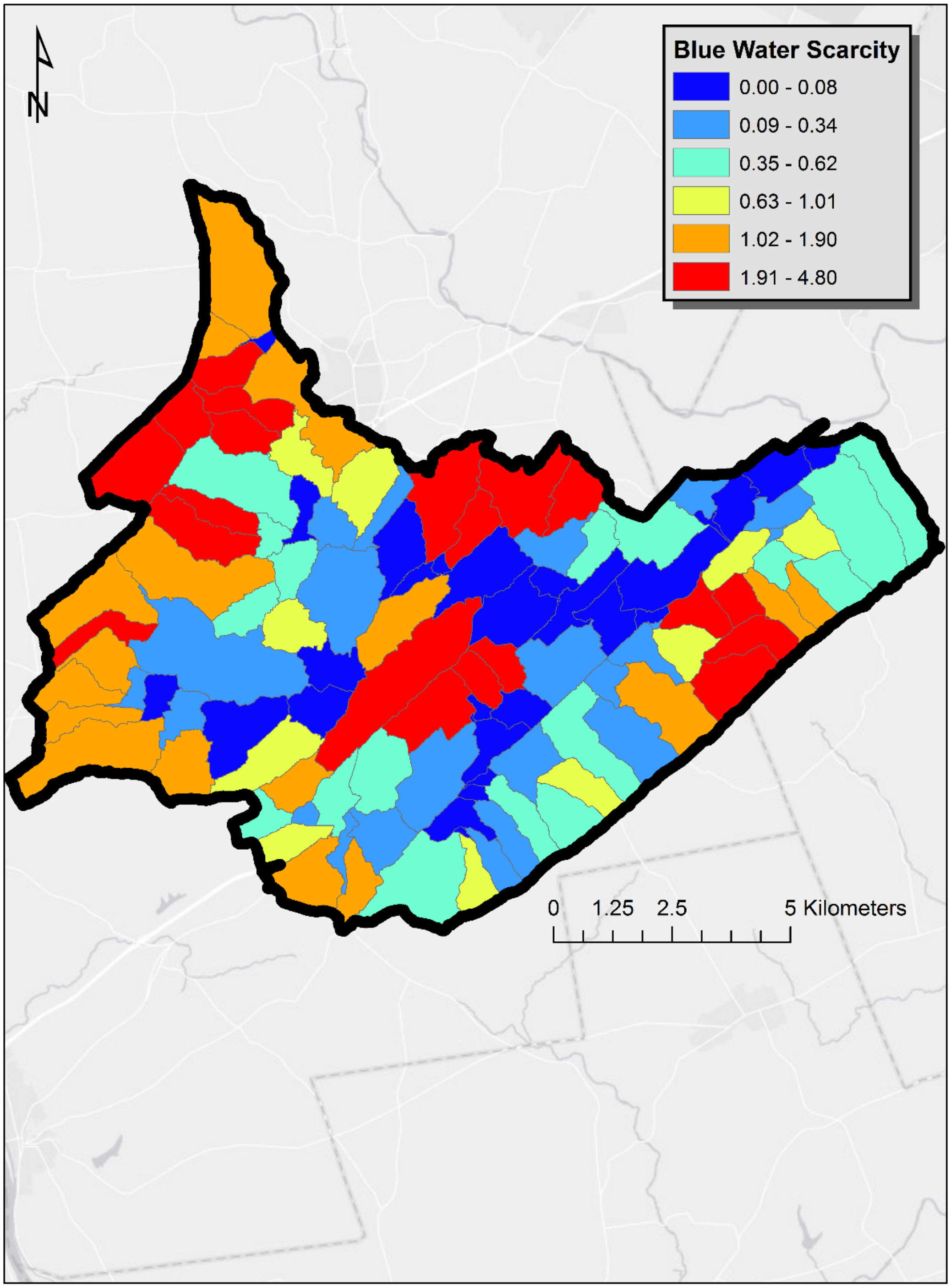


Fig. 10. Spatial distribution of blue water scarcity based on projected land use and population in the Neshanic River Watershed.

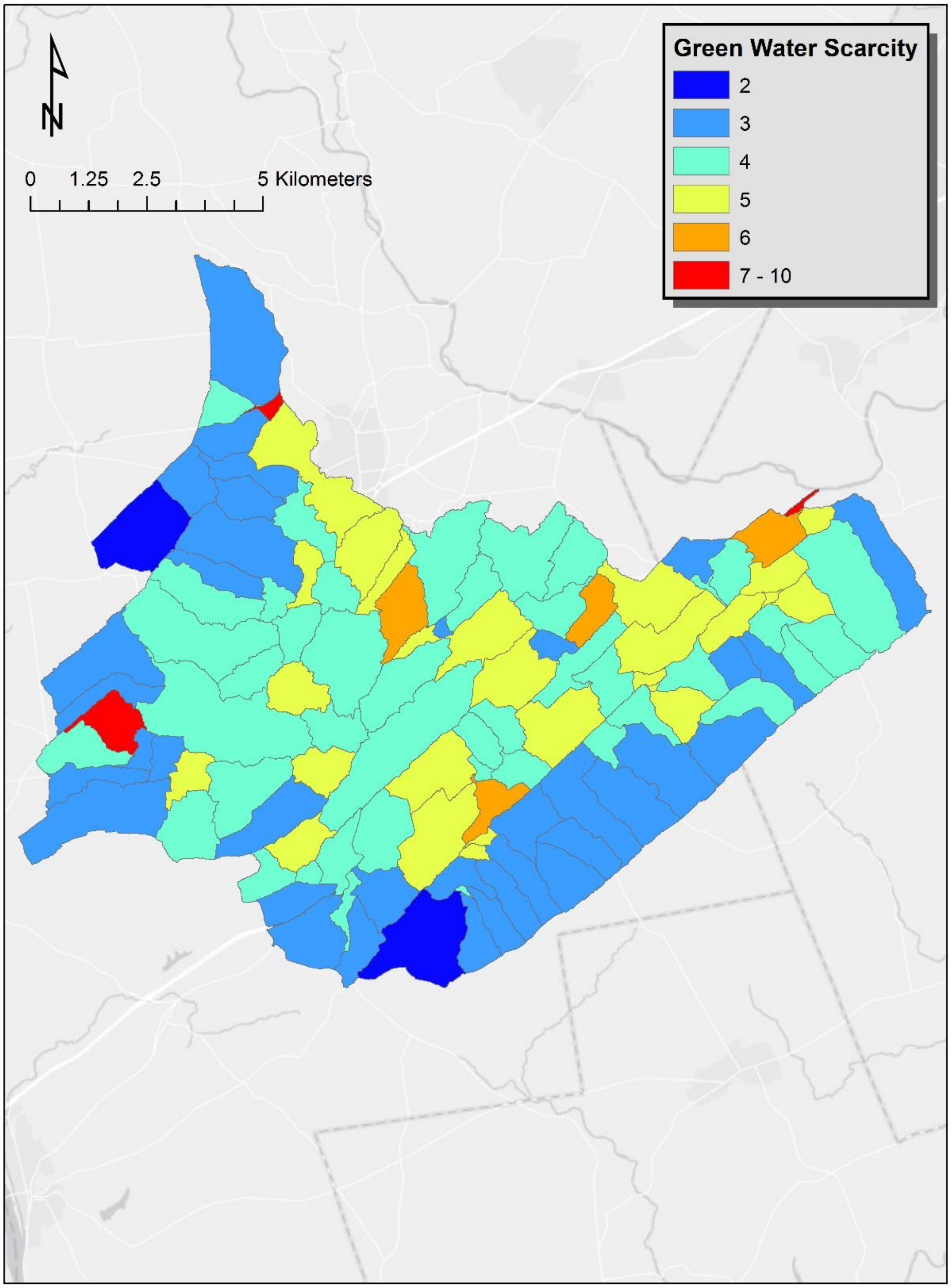


Fig. 11. Spatial distribution of green water scarcity based on projected land use and population in the Neshanic River Watershed.

4. Conclusions

The integrated modeling framework we have developed, consisting of a physically based spatially distributed hydrological model (SWAT) loosely coupled with an agent based probabilistic land use conversion model, permitted an analysis of the how the factors affecting blue versus green water security vary both spatially and temporally. In our study Central New Jersey watershed, the subbasins with the highest blue water scarcity were associated with higher amounts of urban land and higher water demand by residents. Green water scarcity was associated with urban land use and agricultural-dominated areas of lowest initial soil moisture and higher evapotranspiration. The agent based probabilistic model, working at the scale of land ownership parcels, predicted that future urban development would be the result of forest, rather than farm land conversion. A consequence of the loss of forest land and increasing impervious surface were lower infiltration rates leading to higher blue water but lower green water. However, a slight increase in blue water scarcity is expected compared to current scenario due to increase in urban land as a result of nearly 100percent increase in population in the study area. Though green water flow and green water storage are expected to decrease somewhat under future conditions, increasing green water scarcity was observed in only a small number of study area sub-basins. The integrated modeling framework we have employed here could be improved by i) a better representation of interbasin transfers by incorporating the actual source of water used by individual parcels (i.e., imported surface water vs. on-site well water) and the export via sewage treatment plants outside the basin, ii) incorporation of actual soil moisture and evapotranspiration data for calibration and validation of green water components, and iii) use of more than one land use conversion threshold to capture more possible land transformation scenarios.

Freshwater is widely regarded as critical to provide food security as well as safe drinking water to human society. However, the question remains: with the increasing population growth, climate change, and land cover transformation whether the limited amount of freshwater is sustainable to all parts of the globe? If not, then, how do we identify the locations where the availability of freshwater is inadequate considering the hydrological, climatic condition, ecological, and human consumption factors, simultaneously for current as well as future land use development. The integrated modeling framework we have developed informs policymakers and watershed managers a better understanding of how (interaction or change of different physical parameters in the watershed), where (hotspots), and when (which time period in year) water scarcity may occur across the landscape. The framework also allows changing future conditions, whether due to climate change or land use conversion, to be more readily analyzed. Framing this placed-based decision-making in the context of blue versus green water concepts shifts the focus of water resources planners and managers from traditional runoff management more towards a more rainwater management (Falkenmark and Rockstrom, 2006). Such a shift in focus will bring the critical role of land use activities on managing the various components of green water, as well as blue water, storage and flow.

Acknowledgements

This research was supported by the Johnson Family Chair in Water Resources & Watershed Ecology and Sustainable Raritan River Initiative at Rutgers University. Authors would like to thank Mr. John A. Bognar for providing parcel and zoning data for the study area and Mr. James L. Trimble for his suggestion during population calculation for each subbasin within the watershed. Special thanks to the reviewers for their valuable comments to improve our manuscript.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements, FAO Irrig. Drain., paper 56, FAO, Rome 300, D05109.

Alig, R.J., 2010. Economic modeling of effects of climate change on forest sector and mitigation options: a compendium of briefing papers, United States Department of Agriculture, PNW-GTR-833, 2010. Available at : https://www.hydrol-earth-syst-sci.net/21/3001/2017/hess-21-3001-2017.pdf. (accessed 1/8/2017).

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

Arbab, 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, Morgantown, WV.

Arbab, N., Collins, A. R., Conley, J.,2016. Projections of watershed pollutant loads using a spatially explicit, agent-based land use conversion model: A case study of Berkeley County, West Virginia. Appl. Spat. Anal. Polic. DOI 10.1007/s12061-016-9197-z.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: Model development 1. J.Am. Water Resour. As. 34, 73-89.

Arsanjani, J., Helbich, M., Kainz, W., Darvishi, A.B., 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion - the case of Tehran. Int. J. Applied Earth Obs. 21, 265-275.

Bakker, K., 2012. Water secuirity : research challenges and opportunities, Science 337, 914-915.

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.

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

Bogardi, J.J., Dudgeon, D., Lawford, R., Flinkerbusch, E., Meyn, A., Pahl-Wostl, C., Vielhauer, K., Vorosmarty, C., 2012. Water security for a planet under pressure: interconnected chanllenges of a changing world call for sustainable solutions. Curr. Opin. Env. Sust.4, 35-43.

Brizga, S.O., Arthington, A.H., Choy, S.C., Kennard, M.J., Mackay, S.J., Pusey, B.J., Werren, G.L., 2002. Benchmarking, a 'top-down' methodology for assessing environmental flows in Australian rivers. In Proceedings of International Conference on Environmental Flows for Rivers. Cape Town: Southern Waters Consulting.

Brown, D. G., Page, S. E., Riolo, R., Zellner, M., Rand, W., 2005. Path dependence and the validation of agent-based spatial models of land use. Int. J. Geogr. Inf. Sci. 19, 153–174.

Bockstael, N. E., 1996. Modeling economics and ecology: The importance of a spatial perspective. Am.J. Agr. Econ. 78(5), 1168-1180.

Carlisle, D.M., Wolock, D.M., Meador, M.R., 2010. Alteration of streamflow magnitudes and potential ecological consequences: a multi-regional assessment. Front. Ecol. Environ. DOI: 10.1890/100053.

Clark Labs., 2016. TerrSet, Clark University. Available at: http://www.clarklabs.org/. (accessed 1/8/2017).

d' Aquino, P., August, P., Balmann, A., Berger, T., Bousquet, F., Brondízio, E., 2015. Agent-based models of land-use and land-cover change. Paper presented at the International Workshop. Available at: http://www.globallandproject.org/Documents/LUCC_No_6.pdf. (accessed 1/8/2017).

Falkenmark, M., Rockstrom, J., 2006. The new blue and green water paradigm : breaking new ground for water resources planning and management. J. Water Resour. Plann. Manage.132,129-132.

Faramarzi, M., Abbaspour, K.C., Schulin, R., Yang, H., 2008. Modelling blue and green water resources availability in Iran, Hydrol. Process. 23, 486-501.

Fragkias, M., Seto, K. C., 2007. Modeling urban growth in data-sparse environments: a new approach. Environ. Plann. B. 34, 858‑883.

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. Trans.ASABE 50, 1211-1250.

Gimblett, H., Roberts, C., Daniel, T., Ratli, M., Meitner, M., Cherry, S., Stallman, D., Bogle, R., Allred, R., Kilbourne, D., Bieri, J., 2002. An intelligent agent-based model for simulating and evaluating river trip scenarios along the Colorado River in Grand Canyon National Park. In: Gimblett, H (ed.), Integrating Geographic Information Systems and Agent-

Based Modelling Techniques for Simulating Social and Ecological Processes. Oxford University Press, pp. 245-275.

Giri, S., Nejadhashemi, A.P., Woznicki, S., 2012. Evaluation of targeting methods for implementation of best management practices in the Saginaw River Watershed. J. Environ. Manage. 103, 24-40.

Giri, S., Nejadhashemi, A.P., Woznicki, S., Zhang, Z., 2014. Analysis of best management practice effectiveness and spatiotemporal variability based on different targeting strategies. Hydrol. Process. 28, 431–445.

Giri, S., Qiu, Z., Prato, T., Luo, B., 2016 a. An integrated approach for targeting critical source areas to control nonpoint source pollution in watersheds, Water Resour.Manage.30,5087-5100.

Giri, S., Krasnuk, D., lathrop, R. G., Malone, S. J., Herb, J., 2016 b. State of the Raritan Report, Volume 1, Sustainable Raritan River Initiative, Rutgers University, 2016. Available at: http://raritan.rutgers.edu/wp-content/uploads/2017/01/SOR-Final-2017-01-30.pdf. (accessed 1/8/2017).

Gleeson, T., Richter, B., 2017. How much groundwater can we pump and protect environmental flows through time? Presumptive standards for conjunctive management of aquifers and rivers. River Res. Appl. 34, 83-92.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J. Hydrol. Eng. 4,135-143.

Hagerstrand, T., 1965. A Monte Carlo approach to diffusion VI: Archive of European Sociology.

Harrison, I.J., Green, P.A., Farrell, T. A., Juffe-Bignoli, D., Saenz, L., Vorosmarty, C.J., 2016. Protected areas and freshwater provisioning: a global assessment of freshwater provision, threats, and management strategies to support human water security. Aquatic Conserv. Mar. Freshw. Ecosyst. 26, 103-120.

Heppenstall, A. J., Crooks, A. T., See, L. M., Betty, M., 2012. Agent-based models of geographical systems, Springer, DOI 10. 1007/978-90-481-8927-4_1.

Hoekstra, A.Y., Mekonnen, M.M., Aldaya, M.M., Chapagain, A.K., 2011. The water footprint assessment manual: Setting the Global Standard. Routledge.

Irwin, E. G., Jeanty, P. W., Partridge, M. D., 2014. Amenity values versus land constraints: The spatial effects of natural landscape features on housing values. Land Econ. 90, 61-78.

Johnston, K. M., 2013. Agent analyst: Agent-based modeling in ArcGIS. Redlands, CA: ESRI Press.

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.

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. Ecol. Inform. 5(3), 203-221.

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), 79300.

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.

Menard, S., 1995. Applied logistic regression analysis. Sage University Paper Series on Quantitative Applications in Social Sciences, 106, 98.

Moriasi, D.N., Arnold, J.G., VanLiew, M.W., Binger, R.L., Harmel, R.D., Veith, V., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50,885–900.

Mills, E. S., 1967. An aggregative model of resource allocation in a metropolitan area. Am. Econ. Rev. 57(2), 197–210.

Muth, R. F., 1969. Cities and housing. Chicago: University of Chicago Press.

Najlis, R., North, M.J., 2004. Repast for GIS. Agent 2004: Social Dynamics: Interaction, Reflexivity and Emergence. Chicago, IL.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: part 1. A discussion of principles, J. Hydrol. 10, 282-290.

National Agricultural Statistics Service(NASS): CropScape- Cropland Data Layer, 2015. Available at: https://nassgeodata.gmu.edu/CropScape/ (accessed 4/8/2017).

New Jersey Department of Environmental Protection (NJDEP): 2012. Land use land cover classification systems : NJDEP modified Anderson Classification Systems, NJDEP, Trenton, NJ.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2004. Soil and Water Assessment Tool Input/Output File Documentation Version 2005. Soil and Water Research Laboratory Agriculture Research Services & Black Land Research Center Texas Agricultural Experiment Station, Grassland.

North Jersey Transportation Planning Authority (NJTPA) 2018. Defining the vision and shaping the future. Available at: http://www.njtpa.org/data-maps/demographics. Newark, NJ. (accessed 3/15/2018).

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) Westview Press, Boulder, CO.

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. Science 313, 1068–1072.

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

Panagopoulos, Y., Gassman, P.W.,Jha, M.K., Kling, C.L., Cambell, T., Srinivasan, R., White, M., Arnold, J.G., 2015. A refined regional modeling approach for the corn belt-experiences and recommendations for large scale integrated modeling. J. Hydrol. 524, 348-366.

Parker, D.C., Meretsky, V., 2004. Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. Agr. Ecosyst. Environ. 101, 233-250.

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. J. Artif. Soc. S. 11, 3.

Pontius, R.G., Schneider L.C., 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agr. Ecosyst. Environ. 85, 239-248.

Qian, Z., 2010. Without zoning: urban development and land use controls in Houston. Cities 27, 31‑41.

Qiu, Z., Wang, L., 2014. Hydrological and water quality assessment in a suburban watershed with mixed land uses using the SWAT model. J. Hydrol. Eng. 19,816–827.

Richter, B.D., Davis, M.M., Apse, C., Konrad, C., 2012. A presumptive standard for environmental flow protection. River Res. Applic.28, 1312-1321.

Rodrigues, D.B.B., Gupta, H.V., Mendiondo, E.M., 2014. A blue/green water based accounting framework for assessment of water security. Water Resour. Res. 50, 7187-7205.

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, Regional Research Institute, West Virginia University, Morgantown, WV.

Rost, S., Gerten, D., Bondeau, A., Lucht, W., Rohwer, J., Schaphoff.,2008. Agricultural green and blue water consumption and its influence on the global water system. Water Resour. Res. 44, W09405, doi:10.1029/2007WR006331.

Schuol, J., Abbaspour, K.C., Yang, H., Srinivasan, R., Zehnder. J.B., 2008. Modeling blue and green water availability in Africa. Water Resour. Res. 44, W07406,doi:10.1029/2007WR006609.

Serneels, S., Lambin, E. F., 2001. Proximate causes of land-use change in Narok district, Kenya: A spatial statistical model. Agr. Ecosyst. Environ. 85, 65-81.

Shiklomanov, I. A., Rodda, J., 2003. World Water Resources at the Beginning of the 21st Century, Cambridge Univ. Press, New York.

Shirzadi, A., Saro, L., Joo, O.H.,Chapi K., 2012. A GIS-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavatabad case study, Kurdistan, Iran. Nat. Hazards. 64, 1639-1656.

Sohn, K. T., Park, S. M.,2008. Guidance on the choice of threshold for binary forecast modeling. Adv. Atmos. Sci. 25(1), 83–88.

Swets, J.A., 1986. Measuring the accuracy of diagnostic systems. Science 240, 1285-1293.

Tayyebi, A., Delavar, M.R., Yazdanpanah, M.J., Pijanowski, B.C., Saeedi, S., Tayyebi, A.H., 2010. A spatial logistic regression model for simulating land use patterns, a case study of the Shiraz metropolitan area of Iran. In Chuvieco, E., Li, J., Yang, X. (Ed.), Advances in earth observation of global change. Springer Press.

TerrSet (2017). TerrSet Geospatial Monitoring and Modeling Software, Clark University. Available at: https://clarklabs.org/terrset/. (accessed 3/15/2018).

United States Census Bureau (USCB): 2017. Download 2010 Census Tract relationship files. Available at: https://www.census.gov/geo/maps-data/data/tract_rel_download.html. (accessed 7/7/2017).

United States Economic Research Service. 2017. Shifting Geography of Population Change. Available at: https://www.ers.usda.gov/topics/rural-economy-population/population-migration/shifting-geography-of-population-change/. (accessed 7/7/2017).

United States Geological Survey (USGS): 2017. Estimated use of water in the United States County level data for 2010, .Available at: https://water.usgs.gov/watuse/data/2010/index.html. (accessed 7/7/2017).

Valbuena, D., Verburg, P., Bregt, A. K., Ligtenberg, A., 2010. An agent-based approach to model land-use change at a regional scale. Landscape Ecol. 25, 185‑199.

Verburg P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., Mastura, S.S.A., 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. Environ. Manage. 30, 391-405.

Veettil, A. V., Mishra, A.K., 2016. Water security assessment using blue and green water footprint concepts, J. Hydrol. 542, 589-602.

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.

Vorosmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity, Nature 467, 555-561.

Walsh SJ, Malanson GP, Entwisle B, et al., 2013. Design of an Agent-Based Model to Examine Population-Environment Interactions in Nang Rong District, Thailand. Appl. Geogr.(Sevenoaks,England).39:10.1016/j.apgeog.2012.12.010. doi:10.1016/j.apgeog.2012.12.010.

Wide´n-Nilsson, E., Halldin, S., Xu, C.Y., 2007. Global water-balance modelling with WASMOD-M: Parameter estimation and regionalization. J. Hydrol. 340, 105– 118.

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 Plan. 89, 37–48.

Wooldridge, J., 2002. Introductory econometrics. New York: South-Western.

Wu, F., 2002. Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int. J. Geogr. Inf. Sci. 16, 795–818.

Wu, J.,2006. Environmental amenities, urban sprawl, and community characteristics. J. of Environ. Econ. Manag. 52(2), 527-547.

Xie, C., Huang, B., Claramunt, C., Chandramouli, M., 2005. Spatial logistic regression and GIS to model rural-urban land conversion. Proceedings of PROCESSUS Second International

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

Zang, C.F., Liu, J., Velde, M. Van der., Kraxner, F., 2012. Assessment of spatial and temporal patterns of green and blue water flows under natural conditions in inland river basins in Northwest China, Hydrol. Earth Syst.Sci.16, 2859-2870.

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.

Zuo, D., Xu, Z., Peng, D., Song, J., Cheng, L., Wei, S., Abbaspour, K. C., Yang, H., 2015. Simulating spatiotemporal variability of blue and green water resources availability with uncertainity analysis, Hydrol. Process. 29, 1942-1955.

Highlights

- An integrated modeling framework is developed to better understand water scarcity using blue and green water concept
- Future land use transformation is projected using an Agent-based probabilistic model (ABPM)
- Hydrology in the watershed is estimated using Soil and Water Assessment Tool (SWAT)
- Future population projection is estimated based on zoning and projected land use
- Geospatial framework is developed to identify potential hotspots in the watershed