Environmental Modelling & Software 32 (2012) 1–15

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

An integrated monitoring/modeling framework for assessing human–nature interactions in urbanizing watersheds: Wappinger and Onondaga Creek watersheds, New York, USA

Bongghi Hong[a], Karin E. Limburg[b,*], Myrna H. Hall[b], Giorgos Mountrakis[b], Peter M. Groffman[c], Karla Hyde[b], Li Luo[b], Victoria R. Kelly[c], Seth J. Myers[b]

[a] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853 USA
[b] State University of New York, College of Environmental Science and Forestry, Syracuse, NY 13210 USA
[c] Cary Institute of Ecosystem Studies, Millbrook, NY 12545 USA

ARTICLE INFO

Article history:
Received 31 December 2010
Received in revised form
23 July 2011
Accepted 18 August 2011
Available online 10 September 2011

Keywords:
Urbanization
Impervious surface
Uncertainty
Remote sensing
Socio-economic factors
Water resources

ABSTRACT

In much of the world, rapidly expanding areas of impervious surfaces due to urbanization threaten water resources. Although tools for modeling and projecting land use change and water quantity and quality exist independently, to date it is rare to find an integrated, comprehensive modeling toolkit to readily assess the future course of urban sprawl, and the uncertainties of its impact on watershed ecosystem health. We have developed a combined socio-economic–ecological toolbox, running on the ArcGIS platform, to analyze subsequent impacts on streamflow and nutrient export using the spatial pattern of urbanization in response to anticipated socio-economic conditions and scenarios. We have applied our toolbox to two New York State catchment areas, Onondaga Creek watershed and Wappinger Creek, that have undergone rapid development in the last decades. Uncertainties in temporal trends of new housing permits, spatial distribution of development detection and development potential, and stream conditions were evaluated using three separate toolsets (ArcECON, ArcGEOMOD, and ArcGWLF, respectively). The toolbox capabilities are demonstrated through a year 2020 scenario prediction and analysis, where the aforementioned tools were explicitly linked to determine future housing development, spread of impervious areas, runoff generation, and stream nitrate flux. Higher economic growth was estimated to induce increased new housing permits and spread of impervious surface areas, leading to flashier streamflow as well as worsening stream condition, which was aggravated when only the forest lands were allowed to be developed.

© 2011 Elsevier Ltd. All rights reserved.

Software availability

Name: ArcECON, ArcGEOMOD, ArcGWLF;
Hardware requirements: Windows-compatible PC; Software requirements: ESRI ArcGIS and Microsoft Excel;
Program language: Visual Basic for Applications;
Program size: approximately 500 MB;
License type: free;
Availability information: http://www.esf.edu/cue/integratedmodeling/.

1. Introduction

Increasing urbanization continues to reduce fresh water availability, deteriorate water quality, and threaten the health of aquatic ecosystems (Postel, 2000). Indeed, it is as grave a threat to water resources as global warming and climatic change (Vörösmarty et al., 2000). The spread of impervious surface areas (ISAs) constitutes a major driver of water problems. Loss of permeability has been linked to alterations in hydrology, nutrients, sediment, and toxicant loading, as well as general stream degradation (Forman and Alexander, 1998; Paul and Meyer, 2001; Klein, 1979; Center for Watershed Protection, 2003), with attendant loss of ecosystem function and biodiversity. ISAs increase peak discharges from storm and snowmelt events, increasing the likelihood of downstream flooding as storm waters exceed stream channel capacities (Berry and Horton, 1974). Furthermore, pollutants concentrated on the land surface degrade the biological, chemical, and physical characteristics of lakes, streams, and estuaries receiving urban runoff.

In many parts of the U.S., the expansion of urban and suburban areas is currently the most important driver of land use change (Naiman and Turner, 2000). Thus, accurate prediction of future trends of urbanization is essential to the assessment of stream

* Corresponding author.
E-mail address: klimburg@esf.edu (K.E. Limburg).

1364-8152/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2011.08.006

ecosystem health, and is a need voiced at local to national scales. The pressure for land use change has been most acute around urban centers, via the process referred to as "urban sprawl," defined as "the spread of urban congestion into adjoining suburbs and rural sections in an irregular, unordered, and chaotic way" (Merriam-Webster Dictionary, 2002; Ewing, 1994). Because this expansion of urban areas often takes the form of low intensity development, with more roadways and driveways per household and person, it can in fact have higher impact on water quality than an equal amount of development in existing urbanized areas according to an EPA study (USEPA, 2007). In the "Rust Belt" cities of the Northeastern U.S. there are many abandoned properties that lie in the downtown areas waiting for development; for instance, there are 1600 in Syracuse, NY alone. In-city, in-fill development or revitalization on these properties would add no new impervious surface area, would conduct storm water to existing waste water treatment systems, and could incorporate new low impact development building technologies or green infrastructure.

The future course of urban sprawl, and its impact on ecosystem and watershed health, can be appropriately evaluated with the help of a combined socio-economic–ecological model that assesses what factors create demand for new land development in response to anticipated socio-economic events (Li et al., 2003). Jay Forrester, as early as the 1960s, modeled heuristically the attracting and repelling demographic and economic forces of urban dynamics that lead to urban sprawl (Forrester, 1970). However, the explicit, quantitative links among socio-economic systems, land use change, and ecosystem health have rarely been established. A barrier to such holism derives from the fact that those systems, with their own complexities and nonlinearities, have been studied by different academic disciplines, at different temporal and spatial scales (Nilsson et al., 2003; Veldkamp and Verburg, 2004).

Over the past decade, we have developed a conceptual framework (Limburg et al., 2005; Erickson et al., 2005) exploring the following three sets of research questions: (1) how much does human social and economic activity create demand for new land for development? (2) where in the landscape does land use change in response to the demand for new land? and (3) to what extent does land use change alter stream ecosystem condition? The pursuit of these research questions resulted in the development of three separate quantitative analyses: a social accounting matrix (SAM) model describing the socio-economic structure (Nowosielski and Erickson, 2007), a land use change model based on a binary logit regression (Polimeni, 2005; Polimeni and Erickson, 2007), and an ecosystem assessment relating land use to stream health (Stainbrook et al., 2006; Limburg and Stainbrook, 2006). Recently, we developed an integrated modeling framework, running on the MATLAB (http://www.mathworks.com/) platform, that links these analyses explicitly (Hong et al., 2009). The model takes an economic impact scenario from the user, runs three analyses consecutively, and reports estimated changes in the economy, land use, and stream ecosystem condition. Exploring more general use of this modeling framework, we have encountered two major weaknesses of our approach: (1) although some of these analyses were appropriate for describing the current conditions of the system of interest in detail, it was essentially a "snapshot" in time, rather than dynamic, and (2) some approaches were highly data-demanding, often requiring an extensive dataset that is not publically available and time-consuming to process.

In view of these considerations, our new interdisciplinary team has developed a spatially explicit, integrated ecological–economic assessment toolbox, running on the ArcGIS (http://www.esri.com/software/arcgis/) platform, a popular software package for geographic analysis. The toolbox is composed of three "building blocks" simulating social and economic structures (ArcECON), spatial pattern of urbanization (ArcGEOMOD), and watershed health indexed by stream water quality (ArcGWLF), with explicit links among them (Fig. 1). Each building block is composed of several toolsets that can be designed in a modular way so that alternative models can be readily substituted according to the needs of the user.

The problem can be viewed as three interrelated tasks: (1) socio-economic drivers provide a general expected land use change (e.g., 5% urbanization increase); (2) this general change metric is then used with a set of allocation rules to generate specific spatial footprints of expected change (e.g., urbanization maps); and lastly (3) change specifics are input to hydrological models to assess changes in water quality/quantity. For the first task, our toolbox provides a simple regression-based model using readily available socio-economic statistics to model the socio-economic drivers of land use change. For the second task, namely a detailed, spatially explicit description of existing land cover upon which change will be effected, we use modules from a geospatially explicit model, GEOMOD (Hall et al., 1995a,b; Pontius et al., 2001) to allocate new development following a rule-based approach. Third, we use modules from a hydrology and nutrient loading model GWLF (Generalized Watershed Loading Functions) (Haith and Shoemaker, 1987) to evaluate consequential alterations in water quantity and quality. Each of these tools includes capabilities of assessing uncertainties in its estimation. We note that the proposed modular approach supports integration of additional algorithms and/or data sources; a module can be replaced by another if desired by a user, assuming comparable output between the old and new modules.

To date there is no modeling suite of tools within any GIS system, but most particularly not within ArcGIS, that includes the capacity to evaluate economic drivers of land use change, rate and pattern of change, and water quality impacts induced by that change. Some individual tools that reside within a GIS system include the ArcHydro hydrological modeling suite within ArcGIS (Maidment, 2002). Conservation International has recently partnered with Clark Labs (IDRISI GIS software) to create The Land Change Modeler from which biodiversity and carbon impacts can be assessed. This toolset has been ported to ArcGIS and is available as an add on package in both systems (Eastman, 2011). It includes

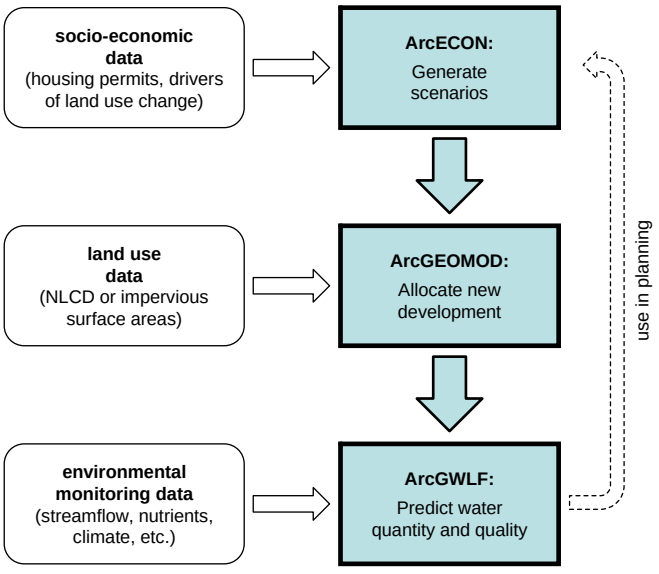

Fig. 1. Conceptual diagram of the key elements of our integrated assessment approach, including various tools mentioned in this paper.

a number of land use change modeling approaches including Markov chain analysis, cellular automata, logistical regression and multinomial logistical regression, GEOMOD, and artificial neural networks. It does not include capacity for hydrological simulation as a function of land use change.

In this study, we applied our toolbox to two New York State catchment areas, Onondaga Creek watershed in upstate New York and Wappinger Creek watershed in the Hudson River Valley (Fig. 2). Both watersheds are characterized by distinct rural–urban gradients and have undergone rapid development in the last decades. A summary of input data used in this study is given in Table 1. In the following sections we describe the modeling framework, tools, and data used for our analysis, demonstrate our findings, and discuss the significance and limitations of our approach.

2. Materials and methods

2.1. Toolbox description

Here, we describe three modules we have incorporated into our assessment framework. The first, ArcECON, is the socio-economic driver algorithm to create demand for new housing. ArcGEOMOD then allocates that demand spatially, and ArcGWLF computes hydrological consequences of land use change. A key input to the assessment framework is the satellite-derived remote sensing data on impervious surfaces, using a new method to estimate ISA (Section 2.3.2). The toolbox and application data are available at http://www.esf.edu/cue/ultraex.htm.

2.1.1. ArcECON

The first of three sets of tools (Fig. 1), the socio-economic toolbox called "ArcECON," statistically determines the factors, both biophysical and socio-economic, that create the demand for new development. Historic economic, social, and demographic information that coincides both spatially and temporally

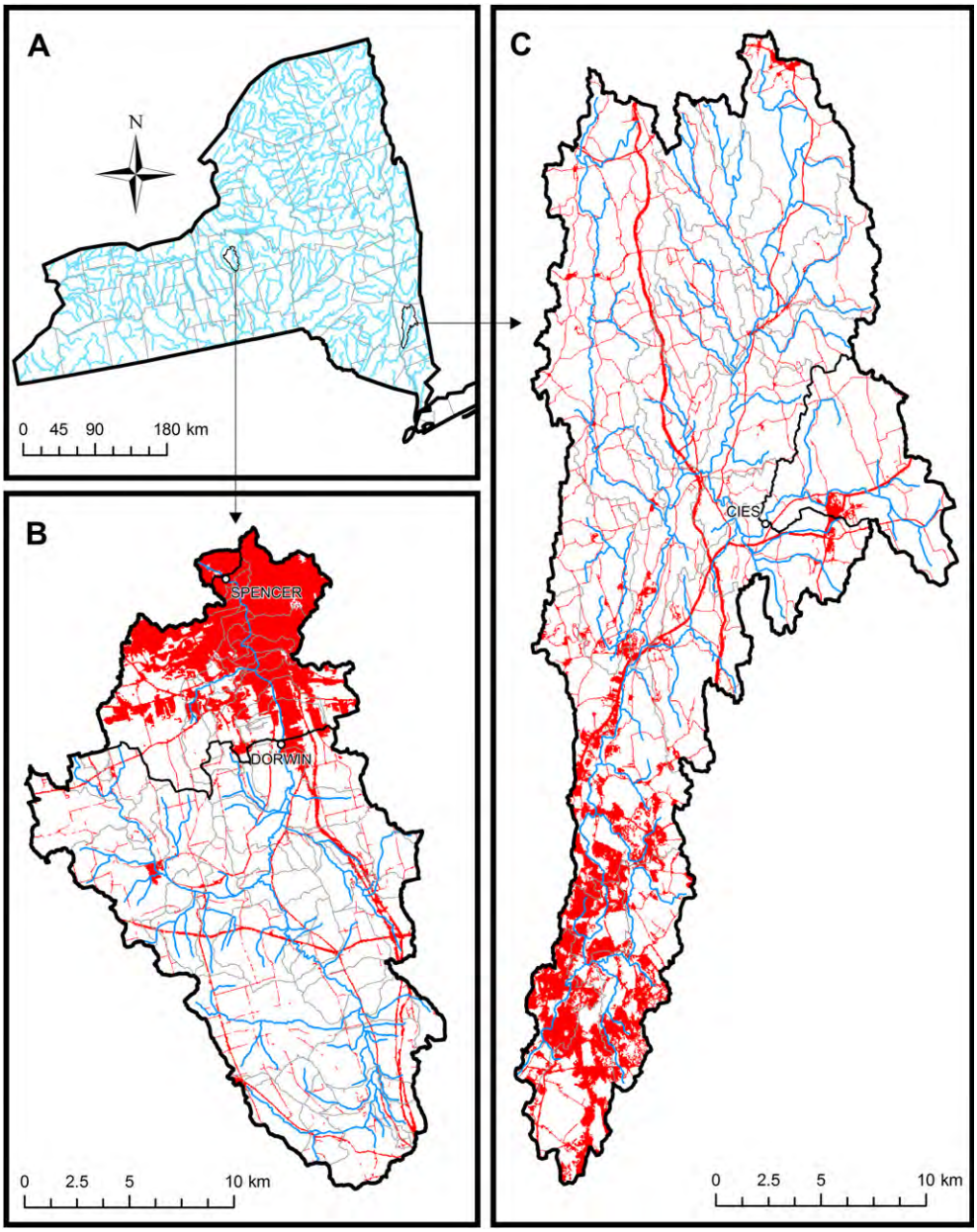

Fig. 2. (A) Map of New York State, (B) Onondaga Creek watershed in Onondaga County, NY, and (C) Wappinger Creek watershed in Dutchess County, NY. Red pixels are urban areas identified by 2001 NLCD (National Land Cover Data). Stream network is shown in blue. The urban region in northern Onondaga County is the City of Syracuse. "Spencer" and "Dorwin" are sites of gaging stations mentioned in the text. "CIES" in (C) is the Cary Institute. Gray lines in (A) indicate county boundaries; gray lines in (B) and (C) indicate boundaries of 72 and 20 subbasins in Onondaga and Wappinger Creek watershed, respectively, including the Spencer (B; upper part of map), Dorwin (B; lower part of map), and CIES (C; right-most part of map) subbasins shown in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Summary of input data used by integrated assessment toolbox.

Toolbox	Name	Source	Resolution	Purpose
ArcECON	Housing permits	SOCDS Building Permits Database[a]	Annual	Input to ArcECON (dependent variable)
	Socio-economic data	See Table 2 for detail	Annual	Identify drivers of new development
ArcGEOMOD	Impervious surface areas	See Section 2.3.2 for detail	30 m × 30 m raster	Input to ArcGEOMOD
	Accuracy maps	See Section 2.3.2 for detail	30 m × 30 m raster	Input to ArcGEOMOD (uncertainty analysis)
	Elevation	National Map Seamless Server[b]	30 m × 30 m raster	Identify drivers of new impervious surface areas
	Slope	Calculated from elevation	30 m × 30 m raster	Identify drivers of new impervious surface areas
	Distance to roads	Calculated from road network[b]	30m × 30 m raster	Identify drivers of new impervious surface areas
ArcGWLF	Subbasin maps	Delineated from elevation	30 m × 30 m raster	Locate subbasin areas
	Precipitation	NCDC[c]	Daily	Climatic input to ArcGWLF
	Temperature	NCDC[c]	Daily	Climatic input to ArcGWLF
	Land use (NLCD)	National Map Seamless Server[b]	30 m × 30 m raster	Assign nutrient runoff coefficients
	Streamflow monitoring data	See Section 2.3.3 for detail	15 min or daily	Calibrate and validate ArcGWLF
	Nutrient grab samples	See Section 2.3.3 for detail	Monthly	Calibrate and validate ArcGWLF

[a] Department of Housing and Urban Development (http://socds.huduser.org/permits/index.html).
[b] U.S. Geological Survey National Map Seamless Server (http://seamless.usgs.gov/index.php).
[c] National Climatic Data Center (http://lwf.ncdc.noaa.gov/oa/ncdc.html).

with new housing permits issued over time is analyzed using best subsets regression analysis. Upon opening the ArcECON tool, the user can select any set of explanatory variables to be regressed against any set of observed housing permits (e.g., single family, single and multiple family, total, rural, urban, etc.) in order to explore correlation with local or regional watershed building activity, and build a model for projection. Once the best model is identified, the user-specified scenarios describing future economic conditions, i.e., varying model inputs, are applied to obtain the estimated number of future housing permits.

We also developed a separate model and a data analysis procedure using the Statistical Analysis System (SAS, 2000) to test the ability of ArcECON to produce a defensible best-fit model. To test for collinearity of the independent variables we applied the SAS functions CORR and INSIGHT and reviewed Pearson's correlation coefficients (r) and scatter plots. Variables that were more than 0.3 correlated with each other were parsed into separate sub-models. Sub-models were then designed to capture all possible models with a high overall effect as measured by the coefficient of determination (R^2) using the SAS REG function with the SELECTION STEPWISE option (for step-wise regression). Mallows' C_p statistic was used as an alternate "goodness of fit" test (Rawlings et al., 1998). Only models with low VIF (Variance Inflation Factor) (less than 5) were retained for further evaluation, and tested for normality with the Shapiro–Wilks test. Based on the results of these tests using each of the dependent variables and a subset of the independent variables, in an iterative process we created a matrix of best fit models that we then compared to ArcECON output.

2.1.2. ArcGEOMOD

The second set of tools, called "ArcGEOMOD," is a raster-based land use-land cover change (LULCC) model that builds upon the original version (Hall et al., 1995a,b; Pontius et al., 2001) now available in the IDRISI GIS software package (Clark Labs, 2010). Through an iterative calibration/validation procedure, the program builds a model of a potential future change surface, called the "suitability" or change potential (CP) map where each grid cell is assigned a number indicating its relative rank in terms of likelihood of change compared to all other landscape cells. This GEOMOD-type, empirically-based, calibration process for determining vulnerability of landscape cells to future change will hereafter be referred to as the "histogram method" (see Pontius et al., 2001 for more in depth description of GEOMOD). The second method available in the toolbox for creating a CP surface is the logistic-regression approach described below.

For ArcGEOMOD, we developed an optimization algorithm that is not available in former versions of GEOMOD. The algorithm seeks to maximize the fit between a CP surface based on a combination of multiple-factor maps by assigning weights to each of them that reflect their relative ability in combination with the other important factors to accurately predict the actual landscape change that occurred in the validation time period. The fit is measured by the area under the Relative Operating Characteristic (ROC) curve (Pontius and Schneider, 2001). The area under the ROC curve measures the ability of a CP map to discriminate among developed and non-developed pixels by assigning a higher CP value to the former. The new algorithm (1) applies an equal weight to all CP maps by creating an arithmetic mean from all included individual factor-based CP maps such that e.g. elevation receives a weight of 0.33, slope 0.33 and distance to roads 0.33; (2) evaluates the fit of the resulting weighted composite CP map to the validation time period using the area under the curve; (3) slightly increases and decreases the weight associated with an individual CP map from its center location while keeping the weights assigned to the other CP maps stationary; (4) recalculates the area under the curve using the weight at the center location, the weight adjusted slightly upward, and the weight adjusted slightly downward. The weight with the best fit of the three is then used as the center for another round of incremental change in weight for that same individual CP map, until no further improvement is gained. At that point, the operation is repeated but using a smaller incremental increase and decrease in weights. This is continued until no increase in fit is found with an arbitrarily small change in weights; (5) steps 3 and 4 are performed in turn for each individual CP map to create a combination of weighting factors; (6) weights assigned to each CP map at the end of step 5 are used as new centers and steps 3 to 5 are repeated until no improvement in the weighted composite CP map is attained. The process of finding individual CP weights is (re)iterated in this way to be certain that shifts in weighting combinations do not create an opportunity within the search space to make further improvements in fit.

The weighted composite CP map giving the best fit during the iterative processing is used to project development into the future. ArcGEOMOD selects from this surface, those cells with the highest rank, or likelihood of development, up to the number projected by the ArcECON analysis. In other words, cells with the highest CP are designated as being developed until the number estimated to be developed by some future time (determined by ArcECON) has been met. This is a computational way of "developing the most desirable land first."

In addition to the histogram method described above, ArcGEOMOD also provides an option of using the regression method to assess change potential. Unlike the histogram method, the regression method treats potential explanatory factors as independent variables in a binary logit regression (single or multiple), and uses the coefficients of the regression equation to create the change potential maps. These individual potential maps then may be weighted and combined in the same way as described above. Detailed description of the regression method is given in Hong et al. (2009).

Finally, for each binary ISA map used for change detection, ArcGEOMOD provides the user an option of specifying the corresponding "accuracy" map (see Section 2.3.2 for detail) to estimate uncertainties in land use change. When only the binary ISA maps of two periods (before and after change) are provided by the user, ArcGEOMOD creates the change map (that is used to create the change potential map as described above) as a binary map containing either 0 (not changed) or 1 (changed). When the accuracy maps are provided with the binary ISA maps, the change map created by ArcGEOMOD contains continuous values between 0 and 1, representing the likelihood of each cell changed between two periods. As an example, if a non-ISA cell in the "before" period has an accuracy value of 0.9, and the same cell becomes the ISA cell in the "after" period with an accuracy value of 0.8, the cell is 72% likely to be changed (0.9×0.8). After the continuous change map is created, it is used for the creation of the change potential map in the same way as described above, except that it is treated as continuous instead of binary (e.g., for the regression approach, the dependent variable is treated as a continuous variable instead of binary).

2.1.3. ArcGWLF

The last of three sets of tools is the watershed assessment tool "ArcGWLF" estimating changes in water quantity and quality in response to the projected land use change determined by ArcGEOMOD. It is based on a simple, lumped-parameter model GWLF (Haith and Shoemaker, 1987), capturing basic hydrology and nutrient export from a watershed. The model uses daily temperature and precipitation data to simulate runoff generation, infiltration to surface soils, evapotranspiration, snowmelt, groundwater recharge, and streamflow discharge at a daily time step. The watershed can be divided into different land-use/soil-type categories as specified by the user. Runoff nutrient coefficients can be specified by the user for each category to calculate nutrient loading. The model and its variations (e.g., CSIM, Mörth et al., 2007) have been applied to a wide range of geographic scales in the U.S. (Howarth

et al., 1991; Swaney et al., 1996; Schneiderman et al., 2002; Lee et al., 2000, 2001) and Europe (Mörth et al., 2007; Smedberg et al., 2006).

In this study, Visual Basic code from the Microsoft Excel version of GWLF (GWLFXL; http://www.eeb.cornell.edu/biogeo/usgswri/GWLFXL/gwlfxl.doc) was obtained from its developer (D. P. Swaney, personal communication) and imported into ArcGIS so that the land use projection map from ArcGEOMOD can be directly linked. In addition to the daily climate data, ArcGWLF can read subbasin and land use maps as input (Table 1). For investigation of urban runoff generation, a new parameter was added, specifying the effective impervious fraction of ISA cells. This fraction of daily rainfall and snowmelt on the ISA cells becomes runoff, and the rest is added to the groundwater through infiltration. The groundwater pool was further divided into "shallow" and "deep" compartments, similar to the modification made in CSIM (Mörth et al., 2007), each compartment with its own recession coefficient ("fast" and "slow," respectively), for the purpose of improving groundwater discharge simulation.

More sophisticated and data intensive calculations, such as calibration and Bayesian uncertainty analysis, were performed using GWLFXL. The calibration feature of GWLFXL makes use of the Solver utility within Excel to find the parameter sets minimizing the sum of squared deviations between the observations and predictions (user-specified daily, monthly, or annual values of streamflow, nutrient loading, or sediment yield). Uncertainties in model parameters were evaluated by applying a Bayesian technique called GLUE (Generalized Likelihood Uncertainty Estimation, Beven and Binley, 1992; Viola et al., 2009). Various Bayesian parameter estimation methods (e.g., Hong et al., 2005) share some common features with the standard Monte-Carlo analyses, in that the selected model parameters are assumed to have a probabilistic distribution instead of a single value and, in turn, simulation output varies stochastically. The step unique to the Bayesian parameter estimation is comparing each set of these simulations with available observations, that can be any variable(s) of the user's choice at any time period and location as long as the simulation outputs to which they can be compared exist, and assigning "weights" (referred to as "likelihood" in Bayesian terminology) to each simulation. Simulations yielding estimates closer to the observations generally get higher weights. The procedure for assigning these weights has been applied through an equation referred to as Bayes' equation; the actual application may be performed using different techniques with varying complexity. Sample sizes and computational requirements are generally determined by the complexity of this step. After the combined weights of each simulation output are determined, the simulations are resampled according to these weights to obtain their updated stochastic distributions. The simulation outputs that better match the observations get higher weights and thus are more likely to be resampled, along with the corresponding input parameters. The degree of updating of each input parameter after the resampling gives a sense of sensitivity; more sensitive parameters to the observations tend to result in narrower ranges of values relative to the initially assumed distributions. Bayesian parameter estimation is an appropriate technique for linking a model simulating a system and the observations available for the system, enhancing in the process our understanding of the system and associated uncertainties. Any type of data at any time period and location can be incorporated into the analysis, and updates can be made repeatedly as new observations become available. The GLUE procedure involves (1) sampling model parameters from their initial ("prior") probability distributions, (2) performing Monte-Carlo simulations and generating prior output distributions, (3) calculating "goodness-of-fit" measure (generalized likelihood) for each simulation (e.g., squared deviations), and (4) resampling according to the likelihood values and generating the updated ("posterior") distributions of simulation output and model parameters. Detailed discussion of various Bayesian parameter estimation techniques, including GLUE, can be found in Hong et al. (2005) and Lamb et al. (1998). Hydrological and nutrient parameters that have been shown to have important controls over the amount and timing of streamflow and nutrient export (Haith and Shoemaker, 1987; Mörth et al., 2007) were chosen for calibration. The GLUE procedure was also applied to the same sets of parameters, with the prior uncertainty ranges obtained from literature review (Haith and Shoemaker, 1987; Mörth et al., 2007).

2.1.4. Linking models for future projection

All the tools in our integrated assessment toolbox are explicitly linked, i.e., output from one tool is directly used as input to another (Fig. 1). For example, output from ArcECON, the observed new housing permits and the projected numbers in response to anticipated socio-economic conditions, is read by ArcGEOMOD tools as input. ArcGEOMOD calculates the annual change rates of housing permits (observed in the county), compares them with the changes in the number of new ISA cells (observed in the watershed), and determines the projected numbers of new ISA cells equivalent to the projected housing permits. ArcGEOMOD then distributes the new ISA cells in each year of the scenario and generates a projected ISA map as output, which in turn is read by ArcGWLF as input to calculate the impervious areas in the simulated subbasins.

2.2. Study sites

Onondaga Creek watershed (325 km^2) lies in central New York, where the city of Syracuse occupies its lower (northern) reaches (Fig. 2). The watershed is the principal tributary to Onondaga Lake, designated as an EPA Superfund site in 1994 for its toxic sediments and surrounding contaminant-contributing industrial sites. The lower reaches of Onondaga Creek typify what has recently been called "urban stream syndrome" (Walsh et al., 2005), with highly degraded physical, chemical, and biological conditions. The watershed is also home to the Onondaga Indian Nation, which recently filed a unique land claim in Federal Court to restore the watershed, and the County of Onondaga is under Federal court order to restore water quality in Onondaga Lake. Land and water quality managements are complicated by different stakeholder goals, so exploring different scenarios in this watershed is a priority. In the years 1990–2005 Onondaga County did not experience a housing "boom" on the scale observed in much of the U.S. (Dugas, 2009), primarily because of flat or declining population numbers over this same period of time. Nonetheless, development of new homes, particularly in the rural areas surrounding the city of Syracuse continued, contributing to the phenomenon known as "Sprawl without Growth" (Pendall, 2003) and the "Shrinking Cities" syndrome (Schilling and Logan, 2008).

Wappinger Creek watershed (541 km^2) within Dutchess County, New York, has long been moderately agricultural (Swaney et al., 2006), but urban sprawl is an increasing concern (Erickson et al., 2005; Limburg et al., 2005). Land use intensity follows a development gradient from the rural (and largely unzoned) northeast toward the largest urban centers located in the southwestern part of the county along the Hudson River (Fig. 2). Since 1984, the Environmental Monitoring Program at the Cary Institute of Ecosystem Studies has collected data on streamflow, stream chemistry, air, precipitation, meteorological, solar radiation, and physical stream parameters at a site on the Wappinger Creek (Section 2.3.3). These data were incorporated into our monitoring/modeling framework to calibrate and test our watershed assessment tools.

2.3. Data collection and preparation

The datasets required to run the three modules are listed in Table 1 and described in more detail as follows.

2.3.1. Economic data

Number of housing permits is the dependent variable of choice, as it is commonly used as a marker of society's economic growth or stagnation, is correlated with construction of expanded service and market facilities, and therefore, accounts for most of the new ISA added to a watershed. From 1990 to 2005, on average, a total of 1146 and 894 housing permits per year were issued in Onondaga and Dutchess County, New York, corresponding to 0.55 and 0.42 permits/km^2/year, respectively. The independent factors evaluated for their influence on the economy are those generally assumed to create job opportunities, increase Gross Domestic Product (GDP) and per capita income, particularly in the middle to upper income brackets, and hence drive the number of new housing permits issued. For the simple regression-based model we gathered data as diverse as home values, mortgage loan rates, employment, earnings by sector, Federal and State Aid, crime statistics, school district tax rates and expenditures, the federal interest rate (the Fed), global and U.S. crude oil prices, and local and U.S. GDP (Table 2). While some data sets represent driving forces external to the county, most of this information is aggregated at the county or town or city level, not the watershed. To get around this frequently encountered problem when trying to assess ecological–economic interactions, i.e. the non-congruency of political and ecological boundaries, we worked with the basic assumption that the county-level and city-level dynamics explained the *rate* of housing development in the watershed. All independent variables were assessed for collinearity and final models eliminated redundancy. We explored many potentially significant determinants of new development, including school achievement and inner city crime rates, but report results that were significant in both watersheds. While the independent factors analyzed do not explicitly indicate *where* new development will take place (that is left to ArcGEOMOD, described in Section 2.1.2), they do explain the driving forces behind the *amount* of new development and ISA imposed on the landscape in the region of interest.

2.3.2. ISA and accuracy maps

The extraction of impervious surface areas for years 1991, 2001, and 2006 in Onondaga and Wappinger Creek watersheds was based on satellite imagery. The training of the image classifier made use of a supervised classification process. Reference data, classifying each pixel in the impervious or the non-impervious class, were obtained from high-resolution aerial photography of approximately 0.6 m ground pixel size. The 0.6 m binary map was up-sampled to a 30 m resolution using a minimum presence rule. The new 30 m binary layer was the reference data used to train a Landsat ETM + satellite image classifier. Looking further into the multi-image classification process, a scene from 2001 covering Central New York (CNY) acted as the base image, as that was the scene where the reference data were collected. The developed classifier was later propagated to the other scenes after normalizing images to the selected CNY 2001 image. As an example, Fig. 3A and B depicts two impervious surface classification products from the year 2001 for the CNY and Wappinger Creek, respectively.

A pixel-based accuracy metric is created to accompany every classification product as a natural byproduct of the classification methodology. Instead of using a single classification method, an ensemble of classifiers is implemented, each

Table 2
Description of socio-economic data and their sources.

Name	Description
Adjustable mortgage rate	HSH Associates Financial Publishers (ARM 1 YR)
Annual average domestic crude oil prices	Energy Information Administration
County unemployment rate	Local Area Unemployment Statistics, Bureau of Labor Statistics
Federal funds effective rate	Board of Governors of the Federal Reserve System
GDP (corrected to 2005)	National Economic Accounts, BEA U.S. Dept. of Commerce
Intergovernmental transfers	Local Government and School Accountability, NY State Office of the Comptroller
Labor force participation rate	(Seas) Civilian Labor Force Participation Rate, Bureau of Labor Statistics, Labor Force Statistics from the Current Population Survey
National unemployment rate	(Seas) Unemployment rate, Bureau of Labor Statistics, Labor Force Statistics from the current population survey
Net earnings by place of residence	Regional Economic Profiles Table CA30, BEA U.S. Dept. of Commerce
Per capita income	Regional Economic Profiles Table CA05, BEA U.S. Dept. of Commerce
Population density	Calculated from USA Counties populations estimates Table, U.S. Census Bureau
Revenue from NYS for health	Local Government and School Accountability, NY State Office of the Comptroller
Revenue from NYS for transportation	Local Government and School Accountability, NY State Office of the Comptroller
Revenue in federal aid	Local Government and School Accountability, NY State Office of the Comptroller
Revenue in NYS aid	Local Government and School Accountability, NY State Office of the Comptroller
Total full-time and part-time employment	Regional Economic Profiles Table CA30 BEA U.S. Dept. of Commerce
United States crude oil (price per barrel)	United States spot price FOB weighted by estimated import volume (dollars per barrel), Energy information administration
Wage and salary jobs	Regional Economic Profiles Table CA30 BEA U.S. Dept. of Commerce
World crude oil (price per barrel)	All countries spot price FOB weighted by estimated export volume (dollars per barrel), Energy Information Administration

addressing different portions of the classification. For example, one classifier may extract water from the scene, followed by another that tries to separate bright impervious surfaces from soil. This multi-step approach allows on-the-fly algorithmic complexity adjustment depending on classification difficulty through targeted classifier selection. Details of the methodology can be found in technical remote sensing journals (Mountrakis, 2008; Mountrakis et al., 2009; Luo and Mountrakis, 2010, in press; Mountrakis and Luo, 2011). Fig. 3C and D depicts the accompanied accuracy metrics corresponding to the impervious surface maps of Fig. 3A and B, respectively.

2.3.3. Environmental monitoring data

Several USGS streamflow monitoring sites exist within the Onondaga Creek watershed, including the Dorwin (USGS 04239000) and Spencer (USGS 04240010) sites along the mainstream of Onondaga Creek watershed (Fig. 2), with the upstream areas of 242 km^2 and 323 km^2, respectively. The areas of upstream Dorwin are relatively rural, whereas the downstream areas (i.e., areas between Dorwin and Spencer) cover the major portion of the urban area. Daily streamflow data were obtained from the USGS website (http://waterdata.usgs.gov/nwis/rt) for calibration and validation of ArcGWLF. Streamflow from the urban portion of the watershed was estimated by calculating the difference of the volume of streamflow between the Spencer and Dorwin sites.

For the Wappinger Creek watershed, streamflow and temperature gauging equipment are located on the East Branch of Wappinger Creek (delineated subcatchment in Fig. 2C) near Millbrook, NY, approximately 1.6 km downstream from a small sewage treatment plant (http://www.ecostudies.org/emp.html). The stream is a tributary to the main branch of Wappinger Creek, which flows into the Hudson River at Wappingers Falls. Streamflow, monitored at 15-min intervals (upstream areas of 50.5 km^2), is calculated from continuous measurements of stream height and rating curves developed on site. Stream (grab) samples (500 ml) have been collected at the end of every month at 2–4 replicate sites and analyzed for nitrate since 1983. Nitrate fluxes at the Cary Institute site were estimated using the USGS software Fluxmaster (Schwarz et al., 2006) from the observed streamflow aggregated to the daily level and the nitrate concentrations obtained from the grab samples. The estimated N fluxes were then used to calibrate and validate ArcGWLF.

3. Results and discussion

3.1. ArcECON

From 1990 to 2005 the total number of new single family housing permits in rapidly urbanizing Dutchess County reflected the economic growth evidenced in both rising local household income (Net Earnings by Place of Residence), and the national employment picture (National Labor Participation Rate). The factors supporting home building, but not necessarily economic growth, were more complicated in the already more urbanized County of Onondaga. There, household income (Net Earnings by Place of Residence) was also a significant contributor to the annual rate of housing construction (single family and duplex) but Adjustable Rate Mortgages and State Aid for Transportation also were significant and retained by the model. County Revenues from Interest Earnings (on its investments) was retained in both models and was negatively correlated with housing construction (the lower the interest rate, the greater the building activity). Although a number of variables in single factor analyses (Table 3) were highly correlated with numbers of single family building permits (e.g. population with $r = 0.74$ and per capita income with $r = 0.82$ in Dutchess County), the best explanatory multiple-factor regression model (Fig. 4A; $R^2 = 0.89$) was achieved using the number of Single Family and Duplex Building Permits in Onondaga County, excluding the City of Syracuse, as follows:

$$\begin{aligned}\text{New Housing Permits} = {} & -2915.05 + 63.52 \\ & \times (\text{Adjustable Rate Mortgage (1 year) Rate}) \\ & + 0.35 \\ & \times (\text{Net Earnings by Place of Residence}) \\ & - 1.91 \times 10^{-5} \\ & \times (\text{Revenue from Interest Earnings}) \\ & + 5.76 \times 10^{-5} \\ & \times (\text{Revenue from NYS for Transportation})\end{aligned} \quad (1)$$

whereas the best economic model for Dutchess County explaining the number of Single Family Building Permits county-wide (Fig. 4B; $R^2 = 0.79$) was:

$$\begin{aligned}\text{New Housing Permits} = {} & -1.74 \times 10^4 + 249.95 \\ & \times (\text{National Labor Participation Rate}) \\ & - 5.56 \times 10^{-5} \\ & \times (\text{Revenue from Interest Earnings}) \\ & + 2.47 \times 10^{-4} \\ & \times (\text{Net Earnings by Place of Residence})\end{aligned} \quad (2)$$

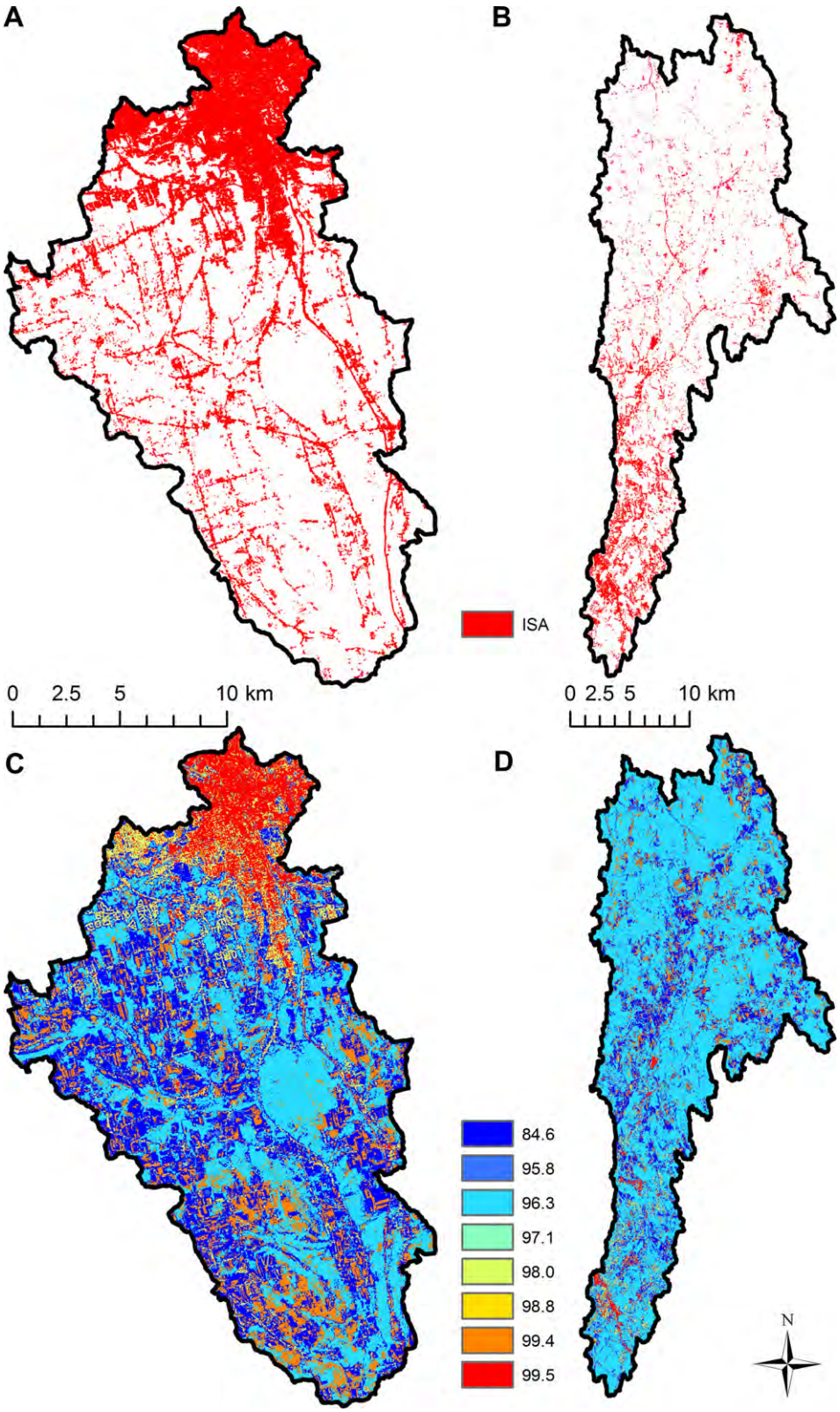

Fig. 3. ISA (A and B) and associated accuracy maps (C and D) for Onondaga (A and C) and Wappinger Creek (B and D) watersheds in 2001.

Table 3
Pearson's correlation coefficients (r) between 19 explanatory socio-economic variables, tested in Onondaga and Dutchess Counties, and the dependent variable (single family and duplex permits in Onondaga County outside the City of Syracuse, and single family housing permits county-wide in Dutchess County) that produced the best final model in each area.

Explanatory variable	Dependent variable	
	Onondaga County (single family and duplex permits outside city)	Dutchess County (single family housing permits)
Adjustable mortgage rate	0.26	−0.50
County unemployment rate	0.52	−0.60
Federal funds effective rate	−0.05	−0.20
GDP (corrected to 2005)	−0.04	0.84
Intergovernmental transfers	0.37	0.40
Labor force participation rate	−0.57	0.17
National unemployment rate	0.51	−0.52
Net earnings by place of residence	0.53	0.81
Per capita income	0.08	0.82
Population	0.17	0.74
Revenue from interest earnings	−0.51	0.10
Revenue from NYS for health	−0.10	0.42
Revenue from NYS for transportation	0.35	0.45
Revenue in federal aid	0.32	−0.03
Revenue in NYS aid	0.26	0.55
Total full-time and part-time employment	0.43	0.63
United States crude oil (price per barrel)	0.30	0.28
Wage and salary jobs	0.28	0.59
World crude oil (price per barrel)	0.31	0.27

ArcECON results (from best subsets regression analysis) matched those of the in-depth SAS analysis (from stepwise regression analysis).

3.2. ArcGEOMOD

Percent impervious area in Onondaga Creek watershed, determined from the ISA maps (Fig. 3), showed a steady increase from 21.7% in 1991 to 27.6% in 2001, and to 31.5% in 2006. The percent impervious area was much lower in Wappinger Creek watershed (4.36% in 1993 and 9.69% in 2001), but the rate of increase from the early 1990s through 2001 was similar (Onondaga Creek watershed = 0.59% per year, Wappinger Creek watershed = 0.66% per year) in both regions.

In both watersheds, a higher fraction of cells changed from non-ISA to ISA cells with decreasing elevation (Fig. 5A and B), slope (Fig. 5C and D), and distance to roads (Fig. 5E and F) between the early 1990s and 2001. Although the discrete histogram approach of ArcGEOMOD and the continuous binary logit regression approach showed generally similar results in both watersheds, there was an indication that the regression approach might have introduced a systematic bias in its estimation (for example, positive residuals at the low and high distance to roads, and negative residuals in the middle). The histogram approach, that stratifies the landscape into blocks of areas with similar properties and makes estimations of change potential in each block, may be a preferred choice where the regression approach introduces a systematic bias in the residuals of estimation. The steepest change (most important factor) was in the distance to roads.

These relationships were applied to create individual change potential maps for each driver (elevation, slope, and distance to roads). Using the third ISA map available in 2006 for the Onondaga Creek watershed, the relative importance of these potential maps were weighted and combined (Fig. 6A). When only a single land use change driver was considered, the highest ROC was obtained from the change potential map calculated from the distance to roads (Table 4). The ROC statistics generally increased as more drivers were combined, and the highest value was obtained when all three drivers were considered.

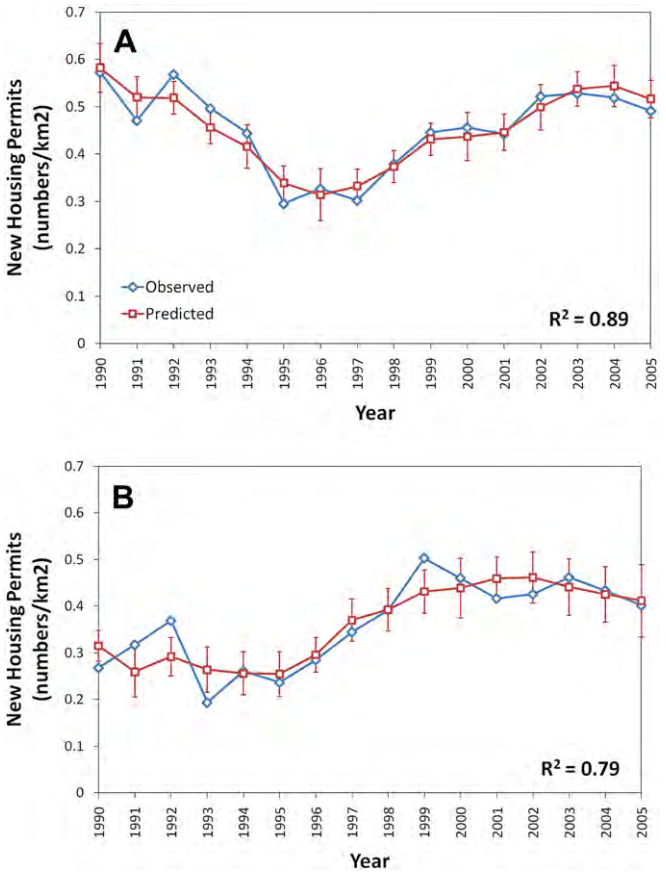

Fig. 4. Observed (blue) new housing permits in Onondaga (A) and Dutchess (B) Counties, NY, and predictions (red) from the best economic regression model. Error bars are 95% prediction interval. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Incorporating the accuracy maps (Fig. 3C, described in Section 2.3.2) in the change detection procedure resulted in expressing ISA change as a continuous variable between 0 (not changed) and 1 (changed), instead of as a binary variable. As a result, more gradual changes were introduced from the highest to the lowest change potential values when the individual change potential maps were created (i.e., the change from high to low potential values became "smoother" when the accuracy maps were added; distinguishing between high and low potential values became "more uncertain"). Consequently, a "fuzzier" map of combined change potentials (all drivers included) was obtained when the accuracy maps were incorporated (Fig. 6B).

Creating individual change potential maps using the simple binary regression method, and weighting and combining these maps afterwards, generally resulted in ROC statistics comparable to those from the GEOMOD histogram approach (Table 4). The multiple regression approach showed the lowest ROC statistics in most cases. Although the difference was relatively small between ROC statistics calculated with and without the accuracy maps, the highest value was obtained by applying the histogram approach with all the drivers (elevation, slope, and distance to roads) and the accuracy maps included (0.746).

3.3. ArcGWLF

Using available monitoring data (Section 2.3.3), ArcGWLF was calibrated during the 1997–2001 period and validated during the

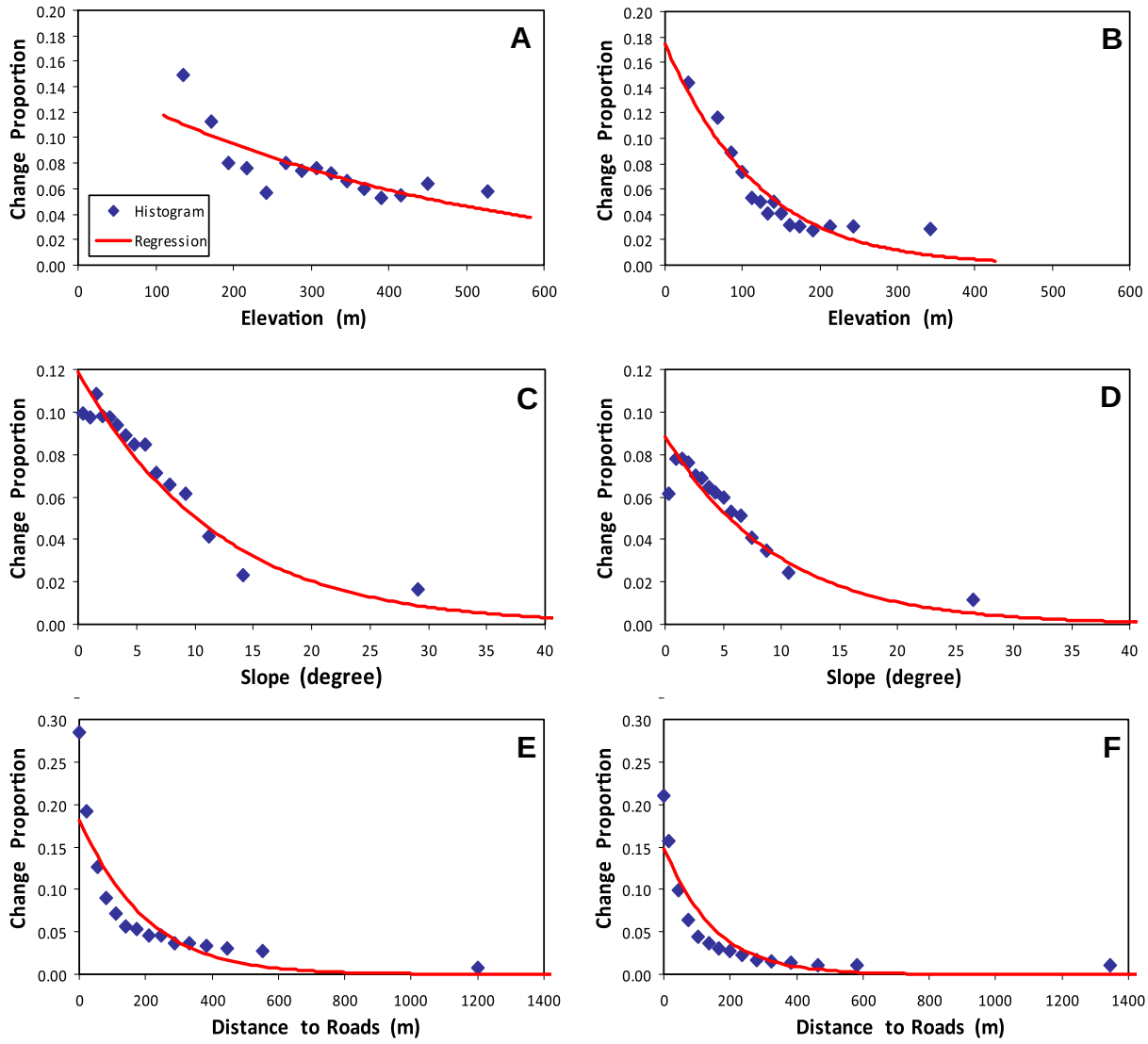

Fig. 5. Drivers (elevation (A and B), slope (C and D), and distance to roads (E and F)) of new impervious cells in Onondaga (A, C, and E; 1991–2001) and Wappinger (B, D, and F; 1993–2001) Creek watersheds estimated with discrete histogram (blue dots) and continuous binary logit regression (red line) approaches of ArcGEOMOD. The regression approach used each binary grid cell value in the study area (1 = changed from non-ISA to ISA cell; 0 = no change) as dependent variable; the histogram approach grouped the grid cells into 15 equal numbered blocks to calculate the change proportion in each group.

2002–2005 period, at the Dorwin and Spencer sites within the Onondaga Creek watershed (streamflow) and at the Cary Institute within the Wappinger Creek watershed (streamflow and nitrate export) (Fig. 7). *R*-square values ranged from 0.60 to 0.78 and from 0.59 to 0.64 during the calibration and validation periods, respectively.

Parameter sensitivity was determined based on the change in the standard deviation (i.e., variation) of parameter distribution from prior to posterior (Hong et al., 2005) obtained from the results of the Bayesian GLUE analysis (Table 5). Out of five hydrological parameters, the growing season evapotranspiration cover factor (controlling the amount of streamflow) and shallow groundwater recession coefficient (controlling the groundwater discharge rate, thus the shape of hydrograph) were identified to be most sensitive. The calibrated value and the average posterior distribution were very close for these sensitive parameters (Table 5). The deep groundwater recession coefficient was not sensitive, indicating that the groundwater pool of these systems can simply be approximated as a first-order decay process. Although relatively insensitive, the effective impervious fraction of ISA cells was estimated to range around 0.2–0.3. The two nutrient parameters, estimated based on the observed nitrate fluxes at the Cary Institute, were both found to be sensitive, with the agricultural runoff N coefficient of 7 mg/l (similar to the value reported in Haith and Shoemaker, 1987) and the impervious surface runoff N coefficient of 3–4 mg/l.

The observed proportion of volume (m^3) of streamflow from urban areas (estimated as net areas upstream of Spencer) of the Onondaga Creek watershed showed a distinct monthly pattern (Fig. 8), with lower values around 0.2–0.3 during the dormant season (November–April) and higher values of 0.3–0.45 during the growing season (May–October), peaking in August. The ArcGWLF simulation also reproduced this pattern, although there were some overestimations in the growing season, as well as in October. Evapotranspiration loss by the soil and plants during the growing season may reduce the streamflow contribution from the relatively rural areas (upstream of Dorwin), whereas swift runoff delivered from impervious surface in the urban portion of the watershed into combined storm/sewer overflows may be preventing evaporation losses.

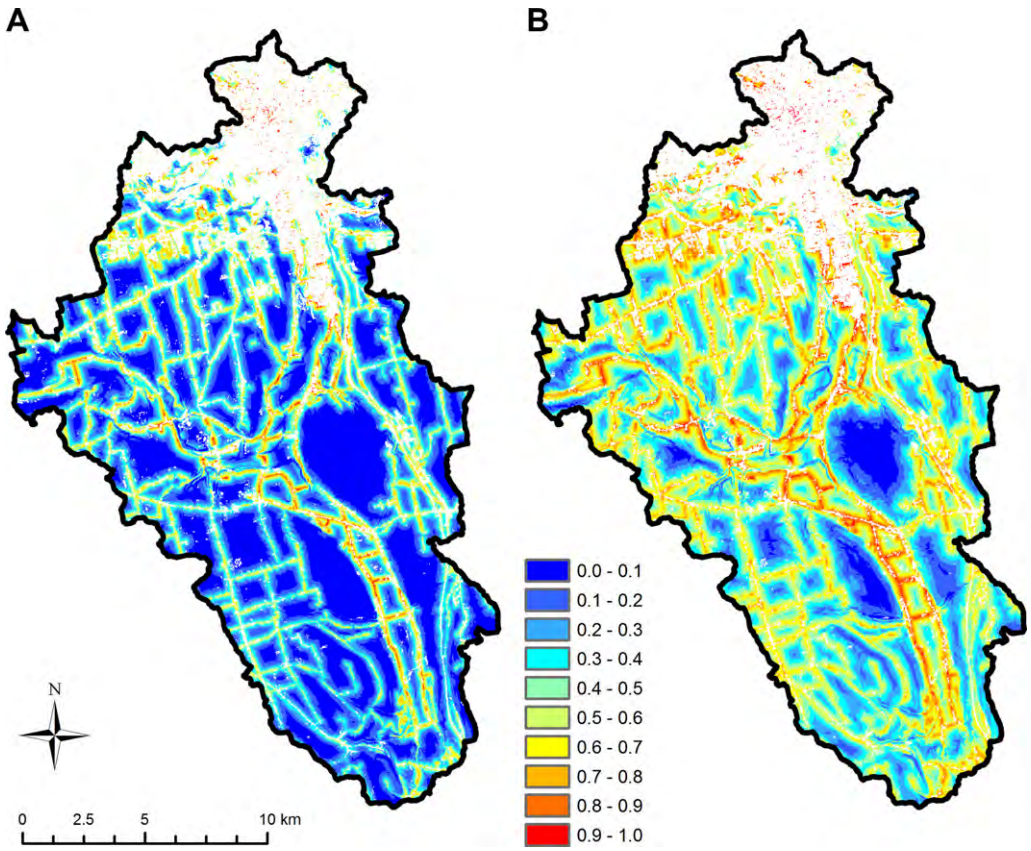

Fig. 6. Map of combined CP (change potentials), relative values scaled from 0 to 1 indicating least to most likely areas of new development, respectively, for the Onondaga Creek watershed calculated from drivers of new ISA cells (elevation, slope, and distance to roads), created from the binary ISA maps only (A) or both the ISA and associated accuracy maps (B).

3.4. Scenario analysis

As an example of scenario analysis, all the three toolsets illustrated above (ArcECON, ArcGEOMOD, and ArcGWLF, describing the economic, land use, and hydrologic conditions of the system, respectively) were explicitly linked by using output from one toolset as input to another. Onondaga Creek watershed was chosen, where three sets of ISA maps (1991, 2001, and 2006) were available to validate the change potential map. Based on the data availability (the 15-year period from 1991 to 2005), we generated projections up to 2020 (another 15 years from 2006). The climatic variables (daily precipitation and temperature) observed during the 1991–2005 period were repeated during the 2006–2020 period for the hydrologic projections.

Three different economic scenarios were used to drive future projections (Table 6): (1) the "base" scenario assuming that the future economic conditions, including Adjustable Rate Mortgage Rates, Net Earnings by Place of Residence, Revenue from Interest Earnings, and Revenue from NYS for Transportation (the best economic model variables for Onondaga County; See Eq. (1) in Section 3.1), will be repeated as observed over the last 15 years, resulting in the same number of new housing permits annually ($0.446\ km^{-2}\ yr^{-1}$), (2) a "low development" scenario assuming that the revenue from interest earnings (with a negative regression coefficient) will be 10 % higher than what had been observed over the previous 15 years, and the rest (with positive regression coefficients) set to 10 % lower, resulting in smaller number of new housing permits ($0.244\ km^{-2}\ yr^{-1}$), and (3) a "high development" scenario, which assumes setting these parameters to change by 10% in the opposite direction, resulting in greater number of new housing permits ($0.647\ km^{-2}\ yr^{-1}$).

Under these economic scenarios, the projected number of new ISA cells was computed each year assuming that the rate of new impervious cell generation is proportional to the rate of new

Table 4
ROC (relative operating characteristic) statistics for the Onondaga Creek watershed.

Variable	ISA only			ISA + accuracy maps		
	Histogram	Simple regression	Multipe regression	Histogram	Simple regression	Multiple regression
Elevation	0.627	0.646	–	0.635	0.646	–
Slope	0.582	0.586	–	0.576	0.586	–
Distance to roads	0.692	0.693	–	0.692	0.693	–
Elevation + slope	0.655	0.662	0.640	0.658	0.663	0.638
Elevation + distance	0.736	0.731	0.711	0.738	0.724	0.715
Slope + distance	0.713	0.709	0.695	0.711	0.702	0.688
Elevation + slope + distance	0.745	0.741	0.712	0.746	0.733	0.709

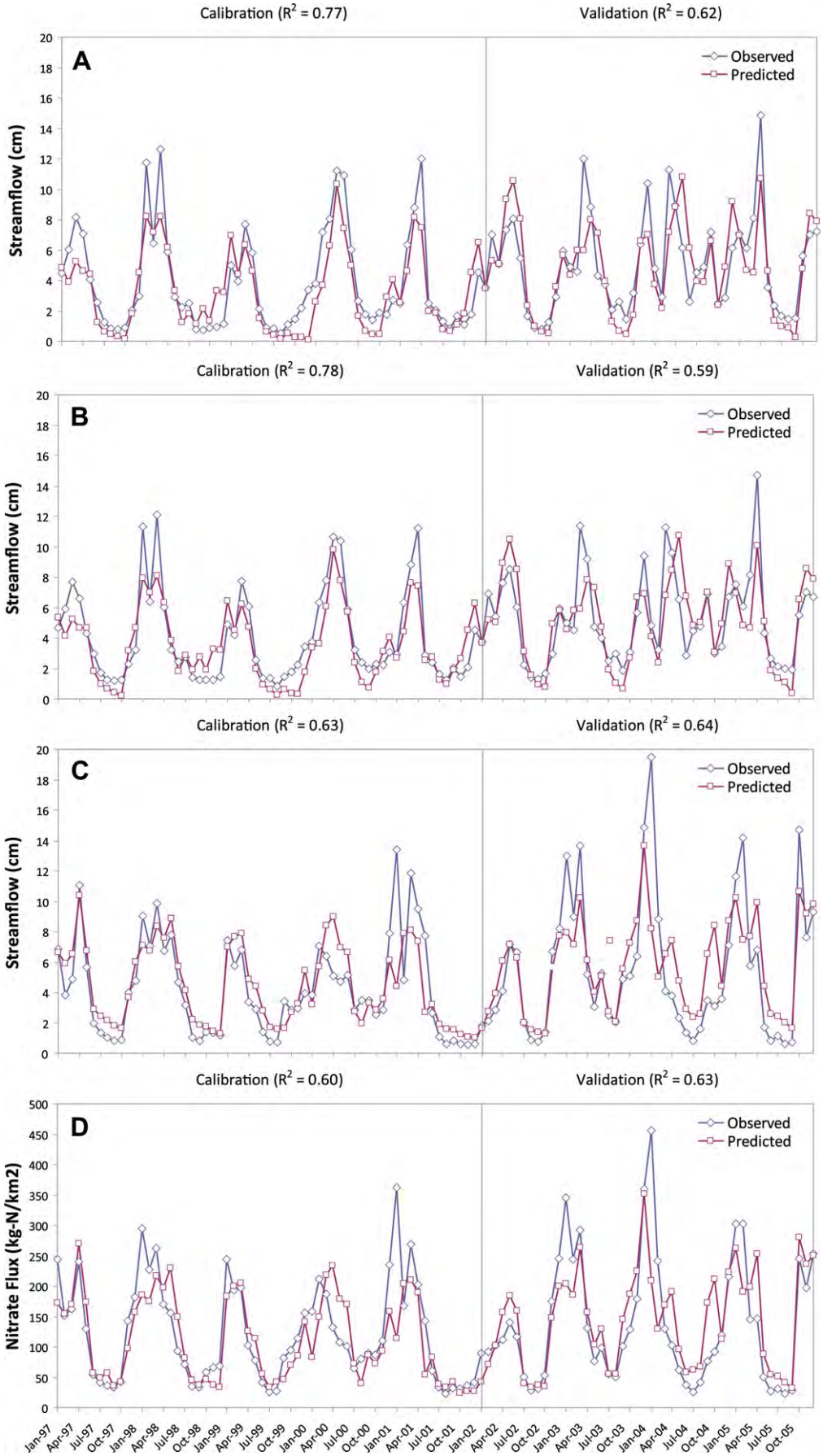

Fig. 7. Observed (blue) and simulated (red) streamflow at Dorwin (A) and Spencer (B) sites within Onondaga Creek watershed and at the Cary Institute within Wappinger Creek watershed (C), and nitrate export at the Cary Institute (D), calibrated during 1997–2001 periods and validated during 2002–2005 periods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
ArcGWLF parameters and simulation outputs during 1997–2001 periods calibrated and estimated using the Bayesian parameter estimation technique.

Variable	Prior Dist	Prior Avg.	Prior Std.	Posterior Avg.	Posterior Std.	Calibrated	Observed
Streamflow, Dorwin, Onondaga							
Growing season ET cover factorgrowing season ET cover factor	U(0,1)	0.478	0.285	0.732	0.152	0.771	0.126
Shallow groundwater recession coefficient	U(0,1)	0.495	0.288	0.145	0.069	0.134	
Deep groundwater recession coefficient	U(0,0.1)	0.050	0.029	0.053	0.028	0.034	
Shallow to deep groundwater flow fraction	U(0,0.1)	0.049	0.029	0.050	0.030	0.096	
Effective impervious fraction of ISA cells	U(0,0.5)	0.247	0.144	0.268	0.150	0.289	
Average daily streamflow (cm/day)		0.147	0.103	0.111	0.044	0.106	
Streamflow, Spencer, Onondaga							
Growing season ET cover factorgrowing season ET cover factor	U(0,1)	0.478	0.285	0.681	0.164	0.708	0.131
Shallow groundwater recession coefficient	U(0,1)	0.495	0.288	0.128	0.067	0.106	
Deep groundwater recession coefficient	U(0,0.1)	0.050	0.029	0.054	0.027	0.027	
Shallow to deep groundwater flow fraction	U(0,0.1)	0.049	0.029	0.050	0.030	0.104	
Effective impervious fraction of ISA cells	U(0,0.5)	0.247	0.144	0.256	0.147	0.178	
Average daily streamflow (cm/day)		0.147	0.101	0.119	0.048	0.115	
Streamflow, Cary Institute, Wappinger							
Growing season ET cover factorgrowing season ET cover factor	U(0,1)	0.478	0.285	0.718	0.196	0.960	0.139
Shallow groundwater recession coefficient	U(0,1)	0.495	0.288	0.103	0.069	0.055	
Deep groundwater recession coefficient	U(0,0.1)	0.050	0.029	0.049	0.028	0.004	
Shallow to deep groundwater flow fraction	U(0,0.1)	0.049	0.029	0.051	0.029	0.032	
Effective impervious fraction of ISA cells	U(0,0.5)	0.247	0.144	0.269	0.152	0.335	
Average daily streamflow (cm/day)		0.206	0.132	0.174	0.063	0.147	
Nitrate flux, Cary Institute, Wappinger							
Agricultural runoff N coefficient (mg/l)	U(0,10)	5.056	2.888	7.062	0.527	7.150	4.012
Impervious surface runoff N coefficient (mg/l)	U(0,10)	5.124	2.881	3.803	1.529	3.449	
Average daily nitrate flux (kg-N/km^2/day)		3.000	1.302	3.715	0.260	3.717	

housing permits as observed during the 1991–2005 period. These new ISA cells were distributed to undeveloped cells of the Onondaga Creek watershed annually, with the highest CP cells selected first, resulting in 2006–2020 average ISA densities of 0.383, 0.357, and 0.409 km^2 km^{-2} under the base, low, and high economic conditions, respectively. (The ISA densities in the year of 2020 were projected to be 0.435, 0.386, and 0.485 km^2 km^{-2}, respectively.)

In response to the increase in ISA, the 15-year average daily simulated runoff increased from 0.012 cm d^{-1} during 1991–2005 to 0.017 cm d^{-1} during 2006–2020 (0.016 and 0.019 cm d^{-1} under the low and high economic conditions, respectively). Nevertheless, there was virtually no change in the simulated streamflow (runoff + groundwater flow). With unchanging streamflow but higher runoff, the simulated stream becomes increasingly "flashy" with the spread of impervious areas.

The nutrient parameters estimated at the Cary Institute were applied to evaluate possible impacts of increasing ISA on the stream nitrate export. The daily nitrate flux showed an increase from 4.29 kg-N km^{-2} d^{-1} (1991–2005 average) to 4.41 kg-N km^{-2} d^{-1} (2006–2020 average), or 4.38 kg-N km^{-2} d^{-1} and 4.43 kg-N km^{-2} d^{-1} under the low and high economic scenarios. As an additional analysis, two land use change scenarios were tested under each economic scenario: "only forest lands allowed to be developed" and "only agricultural lands allowed to be developed," imposing restrictions on distributing new ISA cells to undeveloped lands. Higher nitrate export was predicted when only the forest lands were allowed to be developed, whereas it was slightly reduced when the new ISA cells were distributed only to the agricultural lands. These results indicate that which type of lands the new impervious areas are created from can be as important as the economic scenarios determining how much new impervious area is created.

3.5. Benefits, limitations, and ongoing and future work

As described above, the toolbox presented in this paper is an enhancement of its prototype (Hong et al., 2009), originally running on the MATLAB platform, that had a similar structural framework. This version implements all the features of its prototype, including the event-driven economic analysis based on the SAM input-output socio-economic model, regression-based assessment of land use change potential, and prediction of the spatial pattern of various watershed conditions based on percent land use (Hong et al., 2009). Overall, the new features described in this paper represent capabilities to analyze the dynamics of socio-economic–ecological systems (instead of "snapshots"), to make spatially and temporally detailed projections into the future, and to apply sophisticated uncertainty estimation techniques readily incorporating publicly available data. Specifically, our improvements include:

(1) Whereas the previous socio-economic tool based on the social accounting matrix could generate only event-based economic

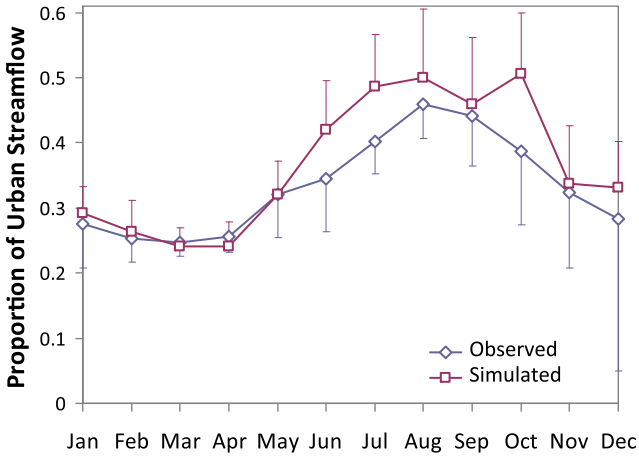

Fig. 8. Observed (blue) and simulated (red) fraction of streamflow from urban portion (net upstream areas of Spencer and below Dorwin) of the Onondaga Creek watershed during the 1997–2005 periods. Error bars indicate standard deviations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Simulated economic, land use, and hydrologic conditions in Onondaga Creek watershed in the time period 1991–2005 (annual average ± standard deviation) and projected changes to year 2020 under three economic scenarios.

Variable	1991–2005	2006–2020, Base	2006–2020, Low	2006–2020, High
New housing permits, Onondaga County (numbers/km^2/year)	0.446 ± 0.078	0.446 ± 0.078	0.244 ± 0.075	0.647 ± 0.081
Impervious surface, Onondaga watershed (km^2/km^2)	0.271 ± 0.035	0.383 ± 0.031	0.357 ± 0.016	0.409 ± 0.046
Average daily streamflow, Onondaga watershed (cm/day)	0.15 ± 0.038	0.15 ± 0.037	0.15 ± 0.037	0.15 ± 0.037
Average daily runoff, Onondaga watershed (cm/day)	0.012 ± 0.0025	0.017 ± 0.0030	0.016 ± 0.0026	0.019 ± 0.0036
Average daily nitrate flux, Onondaga watershed (kg-N/km^2/day)	4.29 ± 1.06	4.41 ± 1.08	4.38 ± 1.09	4.43 ± 1.09
Average daily nitrate flux, Onondaga watershed (kg-N/km^2/day), only forest lands allowed to be developed	4.29 ± 1.06	4.57 ± 1.13	4.46 ± 1.11	4.69 ± 1.16
Average daily nitrate flux, Onondaga watershed (kg-N/km^2/day), only agricultural lands allowed to be developed	4.29 ± 1.06	4.05 ± 1.03	4.18 ± 1.04	3.93 ± 1.02

impacts, ArcECON can perform temporal analyses of socio-economic trends, enabling future economic projection. Socio-economic events can still be evaluated and added to the projected trends.

(2) Incorporating the widely used GEOMOD land use change analysis algorithm, ArcGEOMOD allows users to engage in in-depth spatial analyses of drivers of land use change and urban development projection using various types of data as input. These include time-series of LULCC derived from e.g. National Land Cover Data (NLCD) or ISA maps versus potentially important biophysical constraints to the location of new development such as distance from roads, streams or topography, political factors like zoning, and socio-economic indicators of neighborhood attractiveness, like income, real property values, or school achievement scores.

(3) Whereas the previous version of the watershed assessment tool estimated only average expected stream water conditions based on the projected change in percent land use, ArcGWLF now provides estimations of daily streamflow and nutrient export, with newly added parameters designed to test hydrological impacts of impervious areas and evaluate their uncertainties through Bayesian parameter estimation.

(4) The toolbox allows rigorous integration of various datasets (e.g., socio-economic data, maps of ISA, USGS streamflow monitoring data, etc.) into the assessment framework, making it more feasible to parameterize and test relevant tools, and incorporate accuracy/uncertainty information.

(5) Running in ArcGIS, the toolbox provides a user-friendly interface reading map inputs and writing map outputs without the need for pre- or post-processing the intermediate spatial data. All the tools are explicitly linked, so that the output from one tool is used directly by another tool as input.

Linking models is still problematic, and information may be lost during transfer from one tool to another. For example, in this study ArcGEOMOD generated projections of ISA maps at the spatially detailed scale of 30 m × 30 m. However, ArcGWLF only takes the subbasin areas of different land use types as input and does not consider their spatial configuration. Thus the current approach does not allow for evaluation of the impact of, for example, forested buffers along streams (Kleppel et al., 2004) or impervious areas directly connected to streams by pipes or lined drains (Hatt et al., 2004). In addition, the daily time step at which ArcGWLF is operating limits our ability to assess the fine-scale impacts of impervious areas on the hydrograph. We are currently developing for addition to the toolbox modifications to the EPA Storm Water Management Model (SWMM) that routes storm water overland flow at high resolution (1 m) between different urban land covers to storm drains, projects the receiving conduit hydrograph and volume addition to the urban stream. It provides the ability to test the impacts of addition of green infrastructure technologies as alternatives to impervious surfaces (Sun et al., in preparation). We have found considerable reductions in runoff are possible through the implementation of both tree planting and green roofs. This addition will make the suite of tools developed to date even more powerful in terms of testing the impacts on stream integrity of two varying policies and patterns of development, such as sprawl versus urban in-fill.

Ongoing and future work can lead to more accurate and realistic projections of the future and also more generic and flexible application of the toolbox. For example, in this study we used the same daily climatic variables observed during the 1991–2005 periods for the projection of the 2006–2020 stream conditions (Section 3.4), to estimate the "net" impact of spreading ISA on streamflow and nutrient export (Table 6). Alternatively, the hydrological responses may also be driven by projected climatic conditions, for example based on the downscaled IPCC (Intergovernmental Panel on Climate Change) climate projections or synthetically generated weather conditioned on regional climate projections. Although we simulate growth here only as sprawl, we have in fact run the model under a "business as usual" sprawl scenario versus one of lower demand combined with a strategy of urban in-fill. In the latter we reduced the quantity of demand by 1500 homes, based on an assumed economic slow-down, and relocated new housing development to 1500 abandoned properties within the City of Syracuse. We did this to assess the impacts on urban canopy layer temperatures and the urban heat island regionally. This illustrates the capability of the toolbox for simulating a variety of demand and pattern projections, as well as its value in assessing different ecosystem impacts.

In terms of the data required to drive the modeling toolkit (Table 1), most of inputs are generic datasets publicly available on the Web, or can be replaced with generic datasets. For example, our analyses presented here used ISA and accuracy maps developed for the study areas (Fig. 3). Although a binary ISA map is a crude approximation of actual impervious surface distribution, the incorporation of an accuracy map associated with impervious surface detection compares favorably to such products as the NLCD that only provide summary accuracy metrics. Furthermore, given that the NLCD is updated in 5–10 year intervals, and the recent decision by the U.S. Government to make freely available past and future Landsat satellite imagery, our methodology can provide a cost-effective, on demand means to describe the spread of ISA, a much needed benefit for land use planners. Nevertheless, NLCD or other land use maps may still be used as input where the ISA maps are not available. The overall procedure of calibrating and validating the individual toolkits can be applied as described in this paper without having to reconfigure the model structure. In this study, we applied the modeling framework to two separate systems, one highly

Edited by Foxit Reader
Copyright(C) by Foxit Software Company,2005-2008
For Evaluation Only.

urbanized (Onondaga) and one rapidly urbanizing (Wappinger), demonstrating the toolbox's flexibility.

Our initial goal in this study was to investigate the overall impact of spreading impervious areas (as driven by economic growth) on various stream conditions, including water quantity, stream nutrient export, and sediment transport. To accomplish this goal, it was imperative to have an accurate estimation of water fluxes, as the accurate estimation of nutrient and sediment fluxes depends on it. Thus, much work was devoted to the evaluation of streamflow simulation (e.g., Fig. 8), although a fair amount of work was done with nutrient (nitrate) flux simulation and validation with available data. The hydrology model also produced simulations of other target variables including the sediment yield, but those results are not reported in this paper due to limited availability of validation datasets.

We suggest that a modular approach is an effective way to develop integrated watershed assessment tools. Although we have developed and applied specific modules, other users could implement their own preferred socio-economic drivers, water quality impact algorithms, or other algorithms describing impacts not included here (e.g., terrestrial ecological impacts, or quality of life impacts). Such flexibility has potential to generate a users' community of regional planning tools, based on biophysical and socio-economic data, for sustainable development. Our work presented in this paper has a goal similar to that of previous studies of integrated assessment (Voinov and Bousquet, 2010; Caminiti, 2004), aiming to inform catchment managers and stakeholders with objective quantitative analyses. Future developments toward this goal could include dynamic feedbacks among modules (Fig. 1), e.g., water quality explicitly affecting the decision of where to build, when such data become available. However, in spite of recent conceptual and modeling studies establishing links between the biophysical and socio-economic systems in the integrated assessment (e.g., Kragt et al., 2011; Jakeman and Letcher, 2003), such coupled socio-ecological datasets are currently rare.

While these tools may seem less needed now that new home construction and development in general have slowed across the U.S., it has not stopped altogether and will likely increase again. The amount of open land available to support human activities and the wastes emanating from those activities has diminished substantially over the last 25 years. Hence, every new acre of impervious surface will have more consequence per unit area now than it did in the 1970s. Modeling tools such as these, developed to help communities make smart choices between economic demands and ecological needs are, therefore, required more than ever before. We hope that by making them freely available at the SUNY ESF Center for Urban Environment webpage (http://www.esf.edu/cue/ultraex.htm), the ESRI webpage, and the Ecological Society of America Coupled Natural/Human Systems Modeling web page that groups will find them and use them to plan for locally more sustainable futures.

Acknowledgements

This work was supported by grants from the Syracuse Center of Excellence CARTI Program (which is supported by a grant from U.S. Environmental Protection Agency [Award No: X-83232501-0]), the National Science Foundation (awards GRS-0648393 and BSC-0948952) and by the National Aeronautics and Space Administration (awards NNX08AR11G, NNX09AK16G). We thank D. Nowak, J. Walton, and D. Matthews for providing data and sharing their expertise, D. P. Swaney for sharing the GWLFXL model code used in the hydrological analysis, and three anonymous reviewers for helpful comments on an earlier draft.

References

Berry, B.J.L., Horton, F.E., 1974. Urban Environmental Management: Planning for Pollution Control. Prentice-Hall, Inc., Englewood Cliffs, NJ.
Beven, K., Binley, A., 1992. The future of distributed models – model calibration and uncertainty prediction. Hydrological Processes 6, 279–298.
Caminiti, J.E., 2004. Catchment modelling – a resource manager's perspective. Environmental Modelling and Software 19 (11), 991–997.
Center for Watershed Protection, 2003. Impacts of impervious cover on aquatic ecosystems. In: Watershed Protection Research Monograph No. 1. Center for Watershed Protection, Ellicott City, MD.
Clark Labs, 2010. IDRISI Software. http://www.clarklabs.org/index.cfm.
Dugas, C., 2009. USA Today, 1/5/2009.
Eastman, R., 2011. http://www.clarklabs.org/products/Land-Change-Modeler-Overview.cfm.
Erickson, J.D., Limburg, K.E., Gowdy, J.M., Stainbrook, K.M., Nowosielski, A., Hermans, C., Polimeni, J., 2005. Anticipating change in the Hudson River Watershed: an ecological economic model for integrated scenario analysis. In: Bruins, R., Heberling, M. (Eds.), Economics and Ecological Risk Assessment: Applications to Watershed Management. CRC Press, Boca Raton, FL, pp. 341–370.
Ewing, R., 1994. Characteristics, causes, and effects of sprawl: a literature review. Environmental and Urban Issues, 1–15.
Forman, R.T.T., Alexander, L.E., 1998. Roads and their major ecological effects. Annual Review of Ecology and Systematics 29, 207–231.
Forrester, J.W., 1970. Urban Dynamics. M.I.T. Press, Cambridge, MA.
Haith, D.A., Shoemaker, L.L., 1987. Generalized watershed loading functions for stream flow nutrients. Water Resources Bulletin 23, 471–478.
Hall, C.A.S., Tian, H., Qi, Y., Pontius, G., Cornell, J., Uhlig, J., 1995a. Spatially explicit models of land use change and their application to the tropics. In: CDIACOak Ridge National Lab (Ed.), DOE Research Summary, No. 31.
Hall, C.A.S., Tian, H., Qi, Y., Pontius, G., Cornell, J., Uhlig, J., 1995b. Modeling spatial and temporal patterns of tropical land use change. Journal of Biogeography 22, 753–757.
Hatt, B.E., Fletcher, T.D., Walsh, C.J., Taylor, S.L., 2004. The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams. Environmental Management 34, 112–124.
Hong, B., Limburg, K.E., Erickson, J.D., Gowdy, J.M., Nowosielski, A.A., Polimeni, J.M., Stainbrook, K.M., 2009. Connecting the ecological-economic dots in human-dominated watersheds: models to link socio-economic activities on the landscape to stream ecosystem health. Landscape and Urban Planning 91 (2), 78–87. doi:10.1016/j.landurbplan.2008.11.012.
Hong, B., Strawderman, R.L., Swaney, D.P., Weinstein, D.A., 2005. Bayesian estimation of input parameters of a nitrogen cycle model applied to a forested reference watershed, Hubbard Brook Watershed Six. Water Resources Research 41, W03007. doi:10.1029/2004WR003551.
Howarth, R.W., Fruci, J.R., Sherman, D., 1991. Inputs of sediment and carbon to an estuarine ecosystem: influence of land use. Ecological Applications 1, 27–39.
Jakeman, A.J., Letcher, R.A., 2003. Integrated assessment and modelling: features, principles and examples for catchment management. Environmental Modelling and Software 18 (6), 491–501.
Klein, R.D., 1979. Urbanization and stream water quality impairment. Water Resources Bulletin 15, 948–963.
Kleppel, G.S., Madewell, S.A., Hazzard, S.E., 2004. Responses of emergent marsh wetlands in upstate New York to variations in urban typology. Ecology & Society 9 (1). http://www.ecologyandsociety.org/vol9/iss5/art1.
Kragt, M.E., Newham, L.T.H., Bennett, J., Jakeman, A.J., 2011. An integrated approach to linking economic valuation and catchment modeling. Environmental Modelling and Software 26 (1), 92–102.
Lamb, R., Beven, K.J., Myrabø, S., 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. Advances in Water Resources 22 (4), 305–317.
Lee, K.Y., Fisher, T.R., Jordan, T.E., Correll, D.L., Weller, D.E., 2000. Modeling the hydrochemistry of the Choptank River basin using GWLF and Arc/Info: 1. Model calibration and validation. Biogeochemistry 49, 143–173.
Lee, K.Y., Fisher, T.R., Rochelle-Newall, E.P., 2001. Modeling the hydrochemistry of the Choptank River basin using GWLF and Arc/Info: 2. Model validation and application. Biogeochemistry 56, 311–348.
Li, L., Sato, Y., Zhu, H., 2003. Simulating spatial urban expansion based on a physical process. Landscape and Urban Planning 64, 67–76.
Limburg, K.E., Stainbrook, K.M., Erickson, J.D., Gowdy, J.M., 2005. Urbanization consequences: case studies in the Hudson Valley. In: Brown, L.R., Gray, R.H., Hughes, R.M., Meador, M. (Eds.), The Effects of Urbanization on Stream Ecosystems. American Fisheries Society Symposium, vol. 47, pp. 23–37.
Limburg, K.E., Stainbrook, K.M., 2006. Assessing ecosystem health in Dutchess County, New York. In: Erickson, J.D., Gowdy, J.M. (Eds.), Frontiers in Environmental Valuation and Policy. Edward Elgar, Cheltenham, UK Chapter 6.
Luo, L., Mountrakis, G., 2010. Integrating intermediate inputs from partially classified images within a hybrid classification framework: an impervious surface estimation example. Remote Sensing of the Environment 114 (6), 1220–1229.
Luo, L., Mountrakis, G. A multi-process model of adaptable complexity for impervious surface detection. International Journal of Remote Sensing, in press.
Maidment, D.R., 2002. Arc Hydro: GIS for Water Resources. ESRI Press, Redlands, CA.

Merriam-Webster Dictionary, 2002. Merriam-Webster. On-line at: http://www.m-w.com/home.htm.
Mörth, C.-M., Humborg, C., Eriksson, H., Danielsson, Å., Rodriguez Medina, M., Löfgren, S., Swaney, D.P., Rahm, L., 2007. Modelling riverine nutrient transport to the Baltic Sea – a large scale approach. Ambio 36, 124–133.
Mountrakis, G., 2008. Next generation classifiers: focusing on integration frameworks. Photogrammetric Engineering and Remote Sensing 74 (10), 1178–1180.
Mountrakis, G., Watts, R., Luo, L., Wang, J., 2009. Developing collaborative classifiers using an expert-based model. Photogrammetric Engineering and Remote Sensing 75 (7), 831–844.
Mountrakis, G., Luo, L., 2011. Enhancing and replacing spectral information with intermediate structural inputs: a case study on impervious surface detection. Remote Sensing of Environment 115 (5), 1162–1170.
Naiman, R.J., Turner, M.G., 2000. A future perspective on North America's freshwater ecosystems. Ecological Applications 10 (4), 958–970.
Nilsson, C., Pizzuto, J.E., Moglen, G.E., Palmer, M.A., Stanley, E.H., Bockstael, N.E., Thompson, L.C., 2003. Ecological forecasting and the urbanization of stream ecosystems: challenges for economists, hydrologists, geomorphologists, and ecologists. Ecosystems 6, 659–674.
Nowosielski, A.A., Erickson, J.D., 2007. Regional economic modeling and spatial key sector analysis. In: Erickson, J.D., Messner, F., Ring, I. (Eds.), Ecological Economics of Sustainable Watershed Management: Advances in the Economics of Environmental Resources. Elsevier Science, Amsterdam, pp. 167–182.
Paul, M.J., Meyer, J.L., 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32, 333–365.
Pendall, R., October 2003. Sprawl Without Growth: the Upstate Paradox. Survey Series. The Brookings Institution Center on Urban and Metropolitan Policy, Washington.
Pontius Jr., R.G., Cornell, J., Hall, C., 2001. Modeling the spatial pattern of land-use change with GEOMOD2: application and validation for Costa Rica. Agriculture, Ecosystems and Environment 85 (1–3), 191–203.
Pontius Jr., R.G., Schneider, L., 2001. Land-use change model validation by a ROC method. Agriculture, Ecosystems and Environment 85, 239–248.
Polimeni, J.M., 2005. Simulating agricultural conversion to residential use in the Hudson River Valley: scenario analyses and case studies. Agriculture and Human Values 22, 377–393.
Polimeni, J.M., Erickson, J.D., 2007. Residential location theory, modeling, and scenario analysis of urban growth and planning. In: Erickson, J.D., Messner, F., Ring, I. (Eds.), Ecological Economics of Sustainable Watershed Management: Advances in the Economics of Environmental Resources. Elsevier Science, Amsterdam, pp. 183–210.
Postel, S.L., 2000. Entering an era of water scarcity: the challenges ahead. Ecological Applications 10 (4), 941–948.
Rawlings, J.O., Pantula, S.G., Dickey, D.A., 1998. Applied Regression Analysis: a Research Tool, second ed. Springer-Verlag, Inc., New York.
SAS Institute Inc, 2000. SAS/STAT User's Guide, Version 8. SAS Institute, Cary, NC. http://www.sas.com/.
Schilling, J., Logan, J., 2008. Greening the rust belt: a green infrastructure model for right sizing America's shrinking cities. Journal of the American Planning Association 74 (4), 451–466.
Schneiderman, E.M., Pierson, D.C., Lounsbury, D.G., Zion, M.S., 2002. Modeling the hydrochemistry of the Cannonsville watershed with generalized watershed loading functions (GWLF). Journal of the American Water Resources Association 38 (5), 1323–1347.
Schwarz, G.E., Hoos, A.B., Alexander, R.B., Smith, R.A., 2006. The sparrow surface water quality model: theory, application and user documentation. In: U.S. Geological Survey Techniques and Methods Book 6 Section B, Chapter 3.
Smedberg, E., Mörth, C.-M., Swaney, D.P., Humborg, C., 2006. Modelling hydrology and silicon-carbon interactions in taiga and tundra biomes from a landscape perspective - Implications for global warming feedbacks. Global Biogeochemical Cycles 20. doi:10.1029/2005GB002567.
Stainbrook, K.M., Limburg, K.E., Daniels, R.A., Schmidt, R.E., 2006. Long-term changes in ecosystem health of two Hudson Valley watersheds, New York, USA, 1936–2001. Hydrobiologia 571, 313–327.
Sun, N., Hong, B., Hall, M. Calibration of the U.S. EPA SWMM 5.0 model within a Bayesian generalized likelihood uncertainty estimation (GLUE) framework for a high resolution urban sewershed, in preparation.
Swaney, D.P., Limburg, K.E., Stainbrook, K.M., 2006. Some historical changes in the patterns of population and land use in the Hudson River watershed. American Fisheries Society Symposium 51, 75–112.
Swaney, D.P., Sherman, D., Howarth, R.W., 1996. Modeling water, sediment, and organic carbon discharges in the Hudson/Mohawk basin: coupling to terrestrial sources. Estuaries 19 (4), 833–847.
USEPA (Environmental Protection Agency), 2007. Protecting Water Resources with Higher-density Development. EPA Publication 231-R-06-001. Available: http://www.epa.gov/smartgrowth/pdf/protect_water_higher_density.pdf (accessed 30.03.07).
Veldkamp, A., Verburg, P.H., 2004. Modelling land use change and environmental impact. Journal of Environmental Management 72, 1–3.
Viola, F., Noto, L.V., Cannarozzo, M., La Loggia, G., 2009. Daily streamflow prediction with uncertainty in ephemeral catchments using the GLUE methodology. Physics and Chemistry of the Earth, Parts A/B/C 34, 701–706.
Voinov, A., Bousquet, F., 2010. Modelling with stakeholders. Environmental Modelling and Software 25 (11), 1268–1281.
Vörösmarty, C.J., Green, P., Salisbury, J., Lammers, R.B., 2000. Global water resources: vulnerability from climate change acid population growth. Science 289 (5477), 284–288.
Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., Morgan, R.P., 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24, 706–723.