International Congress on Environmental Modelling and Software

7th International Congress on Environmental Modelling and Software - San Diego, California, USA - June 2014

Jun 16th, 9:00 AM - 10:20 AM

Multiscale Spatial Sensitivity Analysis for Agent-Based Modelling of Coupled Landscape and Aquatic Systems

Arika Ligmann-Zielinska
Michigan State University, ligmannz@msu.edu

Wei Liu
Michigan State University, liuwei11@msu.edu

Daniel B. Kramer
Michigan State University, dbk@msu.edu

Kendra Spence Cheruvelil
Michigan State University, ksc@msu.edu

Patricia A. Soranno
Michigan State University, soranno@anr.msu.edu

Presenter/Author Information
Arika Ligmann-Zielinska, Wei Liu, Daniel B. Kramer, Kendra Spence Cheruvelil, Patricia A. Soranno, Piotr Jankowski, and Seda Salap

International Environmental Modelling and Software Society (iEMSs)
7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,
Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2014-proceedings

Multiscale Spatial Sensitivity Analysis for Agent-Based Modelling of Coupled Landscape and Aquatic Systems

Arika Ligmann-Zielinska[a], Wei Liu[a], Daniel B. Kramer[b], Kendra Spence Cheruvelil[c,d], Patricia A. Soranno[c], Piotr Jankowski[e], and Seda Salap[e]

[a] *Department of Geography, 673 Auditorium Rd, Michigan State University, East Lansing, MI, U.S. (ligmannz@msu.edu, liuwei11@msu.edu)*

[b] *James Madison College, 370 North Case Hall, 842 Chestnut Road, Michigan State University, East Lansing, MI, U.S. (dbk@msu.edu)*

[c] *Department of Fisheries and Wildlife, 480 Wilson Road, Michigan State University, East Lansing, MI, U.S. (ksc@msu.edu, soranno@anr.msu.edu)*

[d] *Lyman Briggs College, 919 E. Shaw Lane, Michigan State University, East Lansing, MI, U.S. (ksc@msu.edu)*

[e] *Department of Geography, 5500 Campanile Dr, San Diego State University, San Diego, CA, U.S. (pjankows@mail.sdsu.edu, salap@rohan.sdsu.edu)*

Abstract: Models of coupled landscape and aquatic systems (CLAS) are prone to input uncertainties that vary over space. To address this challenge, we employ a comprehensive model evaluation that: [1] quantifies the variability of model results (uncertainty analysis), and [2] decomposes this variability based on the relative contribution of inputs to identify major drivers in the model (sensitivity analysis). Our study simulates how agricultural land conversion from active to fallow lands reduces nutrient loading to lakes. We employ an agent-based model of farmer decision making coupled with a spatially-explicit biophysical lake model. A number of model inputs are uncertain including: variables reflecting farmer decision making, maps that represent the environmental benefits of land conservation, and variables that drive nutrient concentrations in CLAS. To be useful for policy analysis, the model requires simplification. To this end, we employ variance-based sensitivity analysis. We run the model multiple times to generate a distribution of lake total phosphorus concentration (TP) and evaluate the variability of TP using two spatial scales. First, the sensitivity analysis is run at a regional scale, at which the TP values from all lakes are lumped into a scalar calculated for the entire study area (aggregate analysis). Second, the sensitivity analysis is run at a lake scale focusing on TP values for individual lakes (fine-scale analysis). The aggregate analysis identifies the most critical components affecting the overall uncertainty of regional TP. The fine-scale analysis identifies the most crucial components affecting uncertainty of TP in individual lakes. A comparison of results from both scales provides useful insights for model simplification.

***Keywords**:* Sensitivity Analysis; Spatial Modelling; Agent-Based Modelling; Land Use Change; Lakes.

1 INTRODUCTION

In 1985, the United States initiated a new land Conservation Reserve Program (CRP) aimed at protecting terrestrial and aquatic landscapes from adverse effects of agricultural production (Lambert *et al.*, 2006; USDA FSA, 2012). This program was established to prevent soil erosion and degradation of water and air quality, as well as facilitate biodiversity conservation (USDA FSA, 2012). CRP participation is voluntary. Farmers make an offer to United States Department of Agriculture (USDA) Farm Service Agency (FSA) and, if enrolled, they convert their farmland to fallow for ten to fifteen years by planting grass and trees in place of agricultural production. In exchange, they receive rental payments for the enrolled land. It is estimated that, in the Great Lakes Region alone, from 500 to 600 thousand acres per year were enrolled in CRP between 2006 and 2010, intercepting annually about 1360 metric tons of phosphorus, which would have otherwise entered the waterways (USDA FSA, 2010). Notwithstanding these apparent environmental benefits of the program, it is unclear whether the existing policies that guide both farmers' and FSA's decision making result in the most efficient return on federal investment (Kramer *et al.*, 2013), or whether they have a positive effect on the quality and health of water bodies. Accurate modelling of land use change from agriculture to CRP-enrolled fallow lands, and the subsequent changes in nutrient loading to lakes, can help evaluate the effectiveness of CRP for improving lake water quality.

The landscape-CRP-lake system serves as a good example demonstrating the complex influence of land use change on lake nutrient concentrations. Its complexity and, by extension, uncertainty, results from a myriad of intertwined social, economic, and ecological conditions that affect the decision of land conservation. The uncertainty of the system is further exacerbated by the limited information on the biophysical processes affecting nutrient loading from the watershed to lakes, and the resulting lake nutrient concentrations. To study the landscape-CRP-lake system, we developed an integrated simulation platform composed of an agent-based model (ABM) of CRP participation and a lake model (LM) of total phosphorus concentration (TP). The ABM simulates CRP enrolment and produces maps of land use change in a given watershed. These land use maps serve as inputs to the LM that calculates TP loading to each lake.

Because we use an ABM, in which each agent is represented by a set of variables, the resulting number of model input parameters (referred to as *factors*) can amount to hundreds or even thousands. These factors influence the size (area) and location of CRP-enrolled land and, as a consequence, the variability of TP in lakes. To build a transparent landscape-CRP-lake model that will be useful for policy analysis, we first need to reduce its dimensionality by eliminating factors that have a negligible effect on the simulated lake TP. Given the nonlinear nature of the model, we employ variance-based sensitivity analysis (VBSA) (Saltelli *et al.*, 2010), in which the variability of the simulated TP is decomposed to quantify the influence of factors on model results. However an important question must be answered: *At what scale(s) should we evaluate lake TP? Should we evaluate individual watersheds, regions, or both?*

Due to the spatial heterogeneity inherent in modelling watershed phosphorus cycling and retention and therefore loading to lakes (Zhang *et al.*, 2012), we hypothesize that the scale of model evaluation will affect its later simplification (Ligmann-Zielinska, 2013). Consequently, we employ variance based sensitivity analysis at two different scales. We first use an aggregate approach, in which variance decomposition is applied to TP calculated for the entire region. We then move to a fine-scale approach, in which TP variance decomposition is performed for each individual lake (and its watershed). The former results in one pair of sensitivity indices per factor, whereas the latter produces a separate pair per factor per lake. In order to make prudent decisions on model simplification, we compare the results of both analyses and identify the non-influential factors common to both scales.

2 METHODS

The ABM is the fundamental component of the integrated model. ABM is a relatively recent approach to modelling complex human-environmental models (An *et al.*, 2005; Matthews *et al.*, 2007). The major premise of ABM is to identify the critical actors in the system i.e., people and institutions that make decisions affected by and affecting the environment. In this research, these actors comprise farmers and the FSA. Below we describe the models in more detail (Figure 1, left).

2.1 Coupled Agent-Based and Lake Model

The process of CRP enrolment is based on well-defined federal regulations (USDA FSA, 2012). A farmer, represented in the model by a farmer agent (FA), starts by evaluating their willingness to enrol in the program. To parameterize FA's enrolment decision, we use individual-level data obtained from the Agricultural Resource Management Survey (USDA ERS, 2012). Based on logistic regression, we identified three primary factors affecting FA's decision: *farmer's retirement* (yes/no), *total value of production* on the farm (in U.S. dollars), and *land tenure* (fraction of land owned by the farmer). Since we do not have data to identify the functional form of FA's choice behaviour, we use a collection of Ordered Weighted Averaging (*OWA*) aggregation functions that represent a continuum from risk averse to risk taking decision making (Yager, 1988).

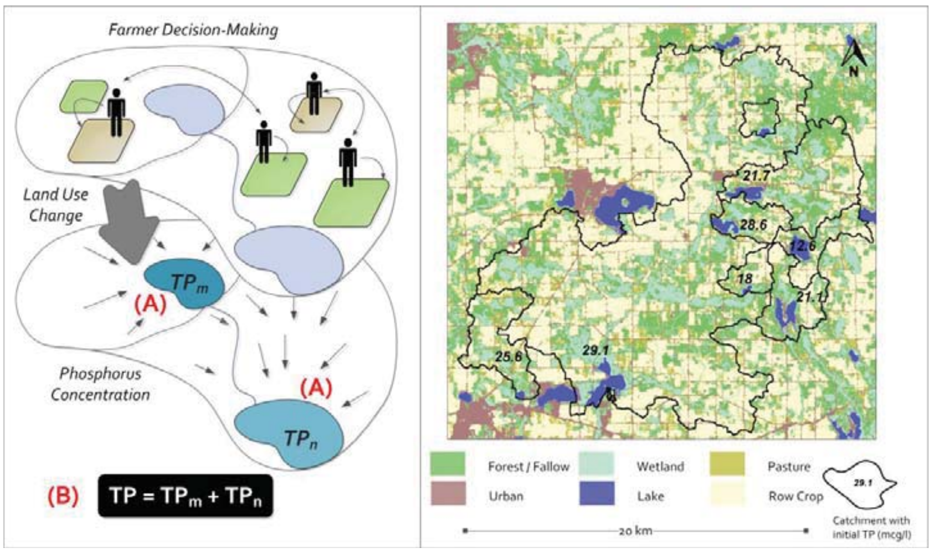

Figure 1. The coupled agent-based and lake model (left), and the study area (right). Refer to the text for symbol definition.

When FA's willingness to participate exceeds an empirically derived threshold, the agent selects a *portion of its farmland* and builds an offer by calculating an expected annual payment based on soil rental rates (USDA FSA, 2012). The location of land to enrol is based on three environmental characteristics: distance to water bodies, slope, and distance to forest. These characteristics are applied to rank the potential enrolment locations using *prioritization* specific to each agent (e.g. slope first, followed by distance to forest, followed by distance to water). To increase the competitiveness of their offer, the FA will apply a *bid* to their expected annual payment according to rules established by the USDA (USDA FSA, 2011). All offers are passed on to the FSA agent who selects a subset of offers based on the available budget empirically derived from U.S. Agricultural Census database (USDA, 2013), the discounted annual payments, and the spatially-heterogeneous *environmental benefit index* (EBI). Finally, the FSA announces the selected signup offers. This signup process is repeated annually. Once the offer is accepted, the land in contract changes to fallow. At the end of the AB simulation (after ten time steps representing a ten-year enrolment period), a new land use map is generated showing the distribution of CRP land.

The land use change map is then sent to the biophysical lake phosphorus concentration model (Figure 1, left), developed in our previous research (Kramer *et al.*, 2013; Zhang *et al.*, 2012). The ABM output map is the only factor that varies from run to run in the TP model. Other inputs include annual precipitation, surface water flow direction, lake area, and lake depth. In addition, the model uses a number of coefficients that account for land use-specific phosphorus attenuation along the flow path. We expect that, because row crops and pasture lands are important sources of phosphorus, the area of CRP participation and the heterogeneity of its spatial distribution can produce variable TP per lake.

2.2 Case Study

We selected adjacent watersheds of seven inland lakes in the Cass County in Michigan, U.S. The outlines of the watersheds are shown in Figure 1, right. They are labelled with their initial TP values, that is, the TP values approximated from the lake model using the input land use map.

2.3 Variance-Based Sensitivity Analysis and Experimental Design

Variance-based sensitivity analysis allows us to analyze the effect of different factors on TP variability and determine the best way to reduce the dimensionality of the model without sacrificing its exploratory power. Because we deal with a very large number of similar factors, that is, factors of the same type for multiple agents (e.g., *farmer's retirement*, *OWA*, and *prioritization* are separately defined for each FA), we first assembled factors of the same type into factor groups. Factor grouping limits the number of model runs by effectively reducing the number of parameters in the model, and allows for treating similar factors as one single factor (Saltelli *et al.*, 2008) without changing the model structure. All model factors are listed in section 2.1 in italics. Except for EBI, all factors are formulated as factor groups. In sum, the integrated model includes one model-level factor (pertaining to the environment) and seven factor groups (pertaining to the agents). For simplicity, both the individual factors and the factor groups are called *factors* in the following sections.

We selected variance-based sensitivity analysis (VBSA) as a technique for evaluating the uncertainty of the integrated AB-LM model. This model does not meet the criteria of additivity or linearity required by other SA approaches, and VBSA accounts for the interaction effects among factors (Saltelli *et al.*, 2008). In the reported experiments, we calculated and decomposed the variance of TP per lake ((A) in Figure 1 centre-left), and then the variance of TP summarized over the region ((B) in Figure 1 bottom-left). With the VBSA results, we computed two sensitivity indices for each factor (*f*): a first order index (S) that calculates the individual contribution of *f* to TP variance, and a total effects index (ST) that represents the contribution of *f* together with other factors to the variance of TP (i.e. accounts for *f*'s interactions). We used well-established formulas for calculating the (S,ST) pairs provided in the SA literature (Saltelli *et al.*, 2008; Saltelli *et al.*, 2010).

The probability density functions used to select the factor values are listed in the Appendix. We ran the integrated model using Monte Carlo simulations (N=4608). Factor samples were produced using the quasi-random Sobol' experimental design (Sobol', 1993). The model was implemented in the Python programming language (http://www.python.org/) and the sensitivity indices were calculated with the SimLab software (http://ipsc.jrc.ec.europa.eu/?id=756).

3 RESULTS

3.1 Sensitivity Analysis of TP at the Regional Scale

Figure 2 shows the results of SA for aggregated TP (i.e. calculated for the entire region). The sizes of the pie chart sectors are proportional to the values of sensitivity indices. The simulated decrease in TP ranges from 0.8 to 14.6 mcg/l, with $\mu = 3.4$ and $\sigma = 1.7$. Assuming the best case scenario, in which all agricultural land is converted to fallow, the results of the computational experiments demonstrate a decrease in TP of 0% to 30% as compared to this full land conversion ($\mu = 4.2$, $\sigma = 4.5$). These low TP decrease values can be attributed to the fact that only 2.9% of the agricultural land in our study area was annually enrolled in CRP during 2000-2010.

We assume that only factors with both S<0.06 and ST<0.06 can be simplified because we do not want to significantly reduce TP variance (it should only be done by improving model input data). Consequently, we will fix only those factors that have overall negligible impact on TP variability. Based on the aggregate results, we can conclude that three factors can be substituted with their means: *value of production*, *farmer retirement* and, to a lesser extent, *prioritization* of land characteristics. Our results also justify the selection of VBSA for sensitivity analysis, as 41% of TP variability can be attributed to factor interactions that other SA approaches cannot account for.

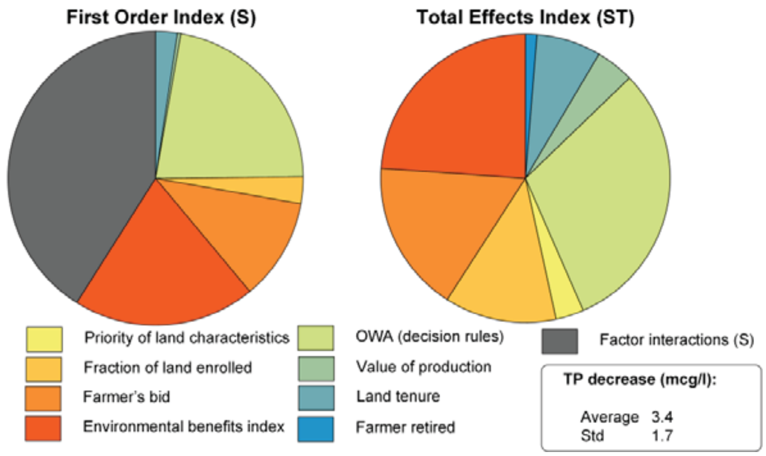

Figure 2 Sensitivity analysis results at the regional scale.

There exists a noticeable spatial dependence of TP reduction and its derivative (S, ST) indices (Figure 3). Thus, the regional results provide incomplete information to simplify our spatially-explicit model and we must also include the individual lake scale in our analysis.

3.2 Sensitivity Analysis of TP at the Lake Scale

Two observations can be made from the results of SA at the lake scale (Figure 3). First, the average decrease in TP is dependent on the value of the initial TP combined with the absolute area of agricultural land in the watershed. In particular, three lakes with the most agricultural land and the highest initial TP (Christiana, Coberts, and Paradise) end up with the highest average TP decreases of ~1 mcg/l (Figures 1 and 3). Second, watersheds with a relatively high interaction effects among factors (from 0.6 for Paradise Lake to 0.75 for Chain Lake) are characterized by higher variability in TP decrease ($\sigma > 1$).

Table 1 Factors with (S,ST) index values below 0.06. VP - value of production, FR - farmer retired, P - prioritization of land characteristics, LT- land tenure, LF - fraction of land enrolled in CRP, EBI - environmental benefits index layer, OWA - ordered weighted averaging decision rule, B - farmer's bid. Refer to the Appendix for details on factors.

	S	ST
REGION	VP,FR,P,LT,LF	FR,VP,P
Lakes		
Coberts	VP,FR,P,LT,LF,EBI,OWA	FR
Christiana	VP,FR,P,LT,LF,B	FR, P
Shavenhead	VP,FR,P,LT,LF,B,OWA	FR, P
Chain	VP,FR,P,LT,LF,EBI, OWA	FR
Paradise	VP,FR,P,LT,LF	FR,VP,P
Donnell	VP,FR,P,LT,LF,B	FR
Birch	VP,FR,P,LT,LF,B,EBI	FR,P

3.3 Comparison of Sensitivity Analysis of TP at Two Scales

At the regional scale, we identified three factors that could be simplified: *value of production*, *farmer retirement* and *prioritization* of land characteristics. When we compare the SA results for the entire region to the individual lakes, we find some interesting discrepancies both in terms of model simplification and factor prioritization. For most of the watersheds (excluding Paradise), the simplification of the model, based on the result of SA at the lake scale, would be different from its regional variant (Table 1). For example, at the individual lake scale, the *value of production* has ST~0.15 for Coberts, Chain, and Birch, whereas the *prioritization* of land characteristics has ST~0.09 for Coberts, Chain, and Donnell. Consequently, we can obtain a simpler equivalent model by fixing only *farmer retirement* to constant because it is the only factor with (S,ST) below 0.06 for all lakes and for the entire region.

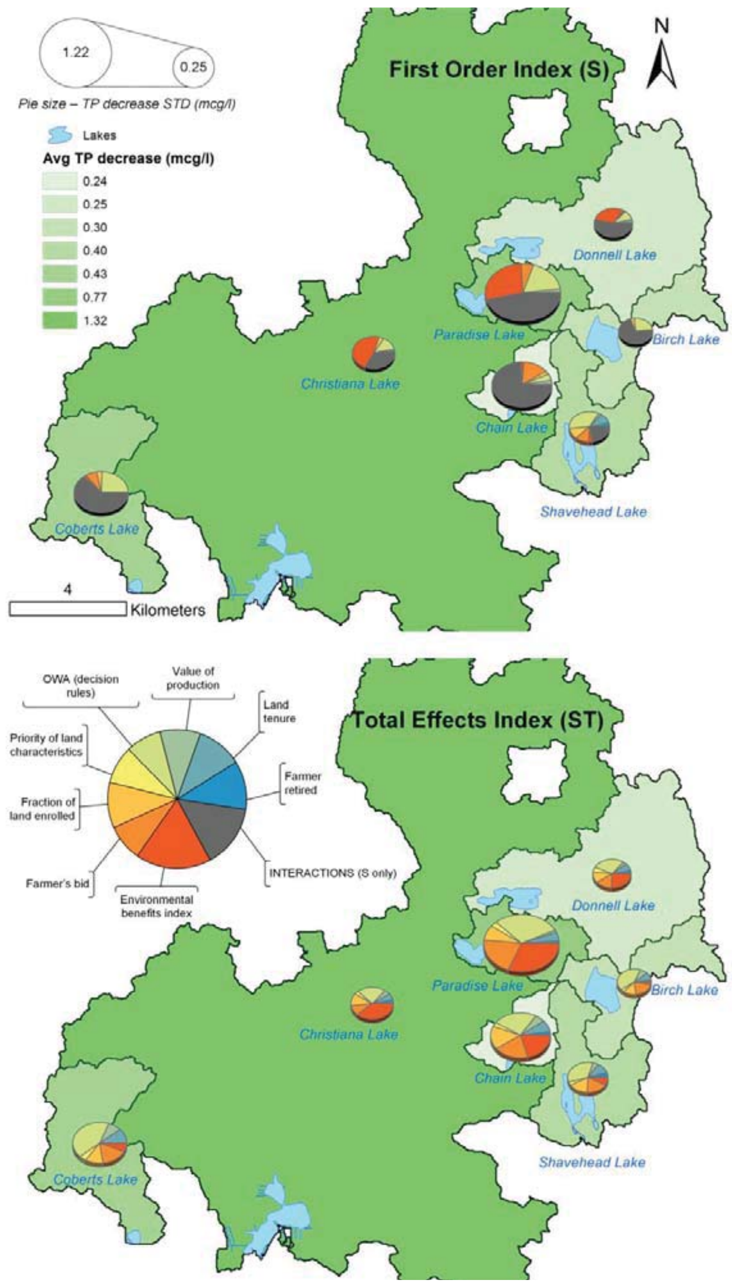

Figure 3 Sensitivity analysis results at the lake (watershed) scale.

When considering factor prioritization, we focus on factors that need more attention due to their considerable influence on outcome variability (i.e. factors with the highest (S,ST) values). The regional analysis suggests three such factors: *EBI, OWA,* and *bid*. The results at the watershed scale lead to slightly different conclusions. For example, the role of *EBI* in shaping the variance of TP for three lakes (Birch, Coberts, and Shavenhead) is much less pronounced. For these lakes, we should rather focus on *LT* and *LF* in addition to *OWA* and *bid*.

4 CONCLUSIONS AND RECOMMENDATIONS

We examined the role of scale in guiding a simplification of an integrated agent-based and biophysical model. We applied variance-based sensitivity analysis to model results measured at two different spatial scales: a lake in its watershed and a region. We demonstrated how model simplification choices, which can be achieved through factor fixing (Saltelli *et al.*, 2008), differ among spatial scales. The regional analysis produces fast and straightforward model simplification because it is based only on one set of indices (S,ST). However, analysis at the lake scale shows that the variance of the results and the subsequent sensitivity indices are spatially heterogeneous. Thus, a finer-scale analysis provides a more detailed and comprehensive picture about the uncertainty of the model, at the expense of brevity and transparency. As part of future research, we will extend our simulation to a wider geographic area that includes lakes with a larger range of TP and perform a thorough spatial statistical analysis to investigate whether general guidelines on variance-based multiscale model simplification can be proposed.

ACKNOWLEDGMENTS

Financial support for this work was provided by the Center for Water Sciences and the Environmental Science and Policy Program at Michigan State University, as well as the National Science Foundation Geography and Spatial Sciences Program Grant No. BCS 1263477. Any opinion, findings, conclusions, and recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation. We also appreciate the feedback provided by three anonymous reviewers on the previous version of this manuscript.

REFERENCES

An, L., Linderman, M., Qi, J., Shortridge, A. &Liu, J. (2005). Exploring complexity in a human-environment system: An agent-based spatial model for multidisciplinary and multiscale integration. *Annals of the Association of American Geographers* 95(1): 54-79.

Kramer, D., Zhang, T., Cheruvelil, K., Ligmann-Zielinska, A. &Soranno, P. (2013). A Multi-objective, Return on Investment Analysis for Freshwater Conservation Planning. *Ecosystems* 16(5): 823-837.

Lambert, D. M., Sullivan, P., Claassen, R. &Foreman, L. (2006).Conservation-Compatible Practices and Programs: Who Participates? In *USDA Economic Research Report*, 48: United States Department of Agriculture.

Ligmann-Zielinska, A. (2013). Spatially-explicit sensitivity analysis of an agent-based model of land use change. *International Journal of Geographical Information Science* 27(9): 1764-1781.

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G. &Gotts, N. M. (2007). Agent-based land-use models: a review of applications. *Landscape Ecology* 22(10): 1447-1459.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. &Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* 181(2): 259-270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. &Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer.* Chichester, England: Wiley-Interscience.

Sobol', I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment* 1: 407-414.

USDA (2013).United States Census of Agriculture, http://www.agcensus.usda.gov/.

USDA ERS (2012).Economic Research Service, http://www.ers.usda.gov/.

USDA FSA (2010).Conservation Reserve Program Annual Summary, http://www.fsa.usda.gov/. United States Department of Agriculture Farm Service Agency.

USDA FSA (2011).Conservation Reserve Program Sign-up 41 Environmental Benefits Index (EBI) Fact Sheet. United States Department of Agriculture Farm Service Agency.

USDA FSA (2012).Conservation Reserve Program Overview. United States Department of Agriculture Farm Service Agency, http://www.fsa.usda.gov/.

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics* 18: 183-190.

Zhang, T., Soranno, P., Cheruvelil, K., Kramer, D., Bremigan, M. &Ligmann-Zielinska, A. (2012). Evaluating the effects of upstream lakes and wetlands on lake phosphorus concentrations using a spatially-explicit model. *Landscape Ecology* 27(7): 1015-1030.

Appendix

Probability distributions for factors used in simulations. U: uniform distribution, D: discrete distribution (value, probability).

Factor Name (Symbol)	Factor Description	Probability Density Function
Farmer Retirement (FR)	Primary operator retired from farming (0: retired, 1: working).	D = { (0, .06), (1, .94) }
Value of production (VP)	Total value of production on a farm.	D = { (0,0), (.2, .06), (.4, .06), (.6,.11), (.8,.15),(1,.62) }
Land tenure (LT)	Ratio of owned to operated acres.	D = { (0,.04), (.2, .14), (.4, .18), (.6,.14), (.8,.15),(1,.35) }
Priority (P)	Prioritization of land characteristics used in ranking potential CRP locations.	D = { 6 combinations with equal probability }
OWA	Farmer agent decision rule based on ordered weighted averaging (Yager, 1988).	D = { 17 combinations with equal probability }
Land fraction (LF)	Fraction of parcel to set aside for conservation.	U = (0, 1]
Bid (B)	Voluntary reduction by the farmer of the offer value below the maximum payment rate.	D = { 0% to 16% of reduction with increments of 1, with equal probability of selection}
EBI	Environmental benefits index.	D = { 6 spatial layers representing different scenarios of EBI distribution in the study area with equal probability of selection }