RESEARCH ARTICLE

Spatially-Explicit Sensitivity Analysis of an Agent-Based Model of Land Use Change

Arika Ligmann-Zielinska*

Department of Geography and Environmental Science and Policy Program, Michigan State University, East Lansing, USA

Reprint

Published in *International Journal of Geographical Information Science*, Special Issue on Land change modelling: moving beyond projections, 27 (9), 1764-1781

*Corresponding author: Email: ligmannz@msu.edu

Abstract
The complexity of land use and land cover change (LULC) models is often attributed to spatial heterogeneity of the phenomena they try to emulate. The associated outcome uncertainty stems from a combination of model unknowns. Contrarily to the widely shared consensus on the importance of evaluating outcome uncertainty, little attention has been given to the role a well-structured spatially-explicit sensitivity analysis (SSA) of LULC models can play in corroborating model results. In this paper, I propose a methodology for SSA that employs sensitivity indices (SI), which decompose outcome uncertainty and allocate it to various combinations of inputs. Using an agent-based model of residential development, I explore the utility of the methodology in explaining the uncertainty of simulated land use change. Model sensitivity is analyzed using two approaches. The first is spatially inexplicit in that it applies SI to scalar outputs, where outcome land use maps are lumped into spatial statistics. The second approach, which is spatially explicit, employs the maps directly in SI calculations. It generates sensitivity maps that allow for identifying regions of factor influence, that is, areas where a particular input contributes most to the clusters of residential development uncertainty. I demonstrate that these two approaches are complementary, but, at the same time, can lead to different decisions regarding input factor prioritization.

1 Introduction

Land use and land cover change (LULC) is one of the most compelling manifestations of global anthropogenic impacts on natural environment (Liu et al., 2007, Gutman et al., 2004, Turner et al., 2007). Although the direct and indirect drivers of LULC have already been identified and described (Rindfuss et al., 2008, Verburg et al., 2004), there is still limited understanding of the linkages among these drivers and the causality chains, which jointly constitute LULC processes (Verburg, 2006, Claessens et al., 2009). To better understand these trajectories, many complex spatially-explicit LULC models have been developed and employed (Agarwal et al., 2002, Milne et al., 2009, Verburg et al., 2006). Many scholars argue, however, that the models are of limited value if their epistemological and ontological uncertainties are not explicitly accounted for and communicated (Turner et al., 2007, Messina et al., 2008, Ascough Ii et al., 2008, Pontius and Neeti, 2010, Visser et al., 2006, Warmink et al., 2010). While the quantifiable probabilistic uncertainties are frequently acknowledged in the reported research, they oftentimes are inadequately addressed, partially due to the lack of methods and tools that could decompose the uncertainty estimates and distribute them among model variables and functional relationships. This procedure of model outcome decomposition is of particular value if we want to understand the relative influence of the interconnected dynamic drivers on LULC.

This paper proposes one such procedure of LULC model outcome decomposition, in which we employ variance-based global sensitivity analysis (Saltelli et al., 2008, Lilburne and Tarantola, 2009) to spatially heterogeneous outputs. In particular, I present a framework that quantifies input sensitivity maps that augment model output analysis. I demonstrate the utility of this framework using an agent-based model of residential development – ABM. The reported ABM is a representation of a complex LULC system, and it incorporates heterogeneous land actors that convert undeveloped land to urban land use. The consequences of their individual decisions feed back into the simulated landscape. These environmental changes affect future agent decisions constituting nonlinear and path dependent phenomena.

I define a comprehensive uncertainty and sensitivity analysis (U-SA) as a systematic quantitative and qualitative approach to estimating model outcome variation and apportioning it to different model components including parameters, variables, and functions (Saltelli et al., 2008, Lilburne and Tarantola, 2009, Makler-Pick et al., 2011). I argue that uncertainty analysis and sensitivity analysis should be applied in tandem. First, uncertainty analysis is employed to quantify outcome variability given model input uncertainties and to generate a distribution of values of the dependent variable like, for example, a selected land use fragmentation statistic (FS). Sensitivity analysis (SA) is then applied to the quantified output variability to investigate the influence of inputs on model results. For instance, the variance of FS can be partitioned into sensitivity measures that represent a fractional contribution of the inputs to the uncertainty of FS values.

The number of studies that employ comprehensive U-SA in environmental modeling increased rapidly over the last decade. Examples include exploring ecosystem vulnerability to climate change (Chu-Agor et al., 2011), parameterizing watershed and hydrological models (Makler-Pick et al., 2011, Yang, 2011), calibrating soil and crop models (Varella et al., 2010), evaluating species interaction and community stability (Hosack et al., 2009), and simulating aquatic ecosystems (Melbourne-Thomas et al., 2011, Estrada and Diaz, 2010). The utilization of U-SA in geographical modeling has also been steadily growing. For instance, spatial U-SA has been applied in multicriteria evaluation and spatial decision support (Feick and Hall, 2004, Gomez-Delgado and Tarantola, 2006, Chen et al., 2010), GIS-based flood forecasting (Crosetto and Tarantola, 2001), conservation reserve delineation (Pantus et al., 2008, Rae et al., 2007, Humphries et al., 2010), managing invasive species (Roura-Pascual et al., 2010), coastal oil spills (Li et al., 2000), groundwater dynamics (Dixon, 2005, Lilburne and Tarantola, 2009), and agent-based modeling of land use change (Ligmann-Zielinska and Sun, 2010, Ligmann-Zielinska and Jankowski, 2010) to name just a few. While these efforts represent considerable advancements in studying uncertainties of spatial models, they either apply a rudimentary one-parameter-at-a-time sensitivity analysis (Chen et al., 2010, Rae et al., 2007), or use an aspatial representation of model results, in which maps are aggregated to one scalar measure like a best composite score among spatial alternatives (Gomez-Delgado and Tarantola, 2006), total nitrate concentration at a site (Lilburne and Tarantola, 2009), or a selected fragmentation statistic of land use patterns (Ligmann-Zielinska and Jankowski, 2010).

In this article, the focus of LULC modeling is directed to spatial representations of comprehensive SA. The report on the proposed framework proceeds in five sections. Following this introduction, a particular approach to SA, called variance-based global sensitivity analysis, is discussed. What follows is a description of the spatial SA framework. The next section outlines the ABM and the computational experiments. Subsequently the analysis of results is presented, first by looking at the uncertainty of the simulated LULC patterns, and then model factor sensitivities. Sensitivity maps of LULC patterns are further contrasted with selected scalar measures (spatial statistics) of development patterns. The final section summarizes the findings and concludes the article.

2 Variance-Based Global Sensitivity Analysis

SA is frequently perceived as an optional and onerous step in modeling, which can be omitted without a substantial loss of information. However, SA offers many benefits that could improve the LULC modeling. Not only does SA improve model validity by recognizing its critical components, but it also provides means of model simplification (factor reduction), which is especially valuable when simulating complex LULC systems. A common approach to SA involves modifying the value of one input (while keeping all the other factors constant) and observing the effects of this change on model result. This popular SA method is often referred to as one-parameter-at-a-time (OAT) approach. Unfortunately, OAT has many serious limitations that diminish its analytical value (Saltelli and Annoni, 2010). In the context of LULC modeling, three are especially worth mentioning. First, the choice of which parameter, where, and by what amount to change is problematic. Even if the modeler has extensive knowledge of the system under study, the magnitude of direct and indirect drivers may be hard to determine. Second, OAT assumes a linear relationship between inputs and outputs. For example, if a LULC modeler increases the value of the *road density* input a few times and observes an increase in output *residential land compactness*, she may conclude that, in her model, the compactness is positively correlated with road density. However, the relationship may be more complicated if *road density* is analyzed in combination with *noise* or *air pollution from transportation*. It is therefore crucial to incorporate factor interaction effects in U-SA of LULC models. Third, OAT is qualitative in nature – it does not compute any fractional contribution of a particular input factor to outcome uncertainty.

Global SA (GSA) has been proposed as an alternative to OAT (Saltelli et al., 2000). Unlike OAT, GSA is based upon perturbations of the entire parameter space, where input factors are examined both individually and in combinations (Campolongo et al., 2000). Different methods of performing GSA in spatial modeling have been developed. Lilburne and Tarantola (2009) categorized the methods based on their model dependence, computational efficiency, and algorithmic complexity. For example, regression-based approaches assume that the underlying model behaves linearly and, consequently, are of limited use for studying dynamic causality (Millington et al., 2007, Mac Nally, 1996). Other common methods, like the importance measures based on Latin Hypercube Sampling (Helton and Davis, 2003, Xu et al., 2005), require well-defined stratification of inputs, a very large number of model executions, and, oftentimes, do not estimate higher-order effects that account for factor interactions (Lilburne and Tarantola, 2009, Saltelli et al., 2004, Helton and Davis, 2003). Here, I propose to employ a model independent variance-based GSA, which obviates the assumptions of model linearity and offers an acceptable compromise in computational efficiency. Variance-based GSA breaks down the total variance (V) of model output, caused by the changes in (z) model inputs, and apportions it to individual factor k (V_k) and k's combinations with other factors $i, j, \ldots, z$ ($V_{kij\ldots z}$) with an increasing level of dimensionality (Saisana et al., 2005, Lilburne and Tarantola, 2009, Crosetto and Tarantola, 2001). For example, if an ABM has the following three input factors a = resident developer's preference for a selected landscape feature (e.g. distance to roads), b = developer's attitude to risk-taking, c = number of parcels known to the developer agent, then the variance of the output (like the ranking of parcels to develop)[1] can be decomposed as follows:

[1] represented using a scalar e.g. an average shift in ranks (Saisana et al., 2005)

$$V = V_a + V_b + V_c + V_{ab} + V_{ac} + V_{bc} + V_{abc} \qquad (1)$$

For example, V_a is the variance of ranking due to the agent's preference alone, and V_{ab} is share in the total variance of parcel ranking caused by the interaction between agent's preference and its risk-taking attitude. This variance is then applied to compute two sensitivity measures: first order (S_k) and total-effect (ST_k) sensitivity index for every factor *k*. For instance for factor *a* defined above:

$$S_a = \frac{V_a}{V} \qquad (2)$$

$$ST_a = \frac{V - VC}{V} = S_a + S_{ab} + S_{ac} + S_{abc}; VC = V_{bc}\{E_a(Y \mid b, c)\} \qquad (3)$$

Where *Y* represents the ranking of parcels to develop (the dependent variable), *V* is *Y*'s unconditional (total) variance, V_a stands for the variance of *Y* due to factor *a* alone, S_a captures the single impact of agent's preference on the ranking of parcels, and ST_a embodies the overall contribution of agent's preference (including its interactions with all other factors) to the variability in the ranking of parcels. *VC* (Saisana et al., 2005) is the conditional variance which, in our example, is the total contribution to the variance in ranking due to the other two factors (*b* and *c*) i.e. the attitude and knowledge of available land.

3 Framework for Spatially-explicit Sensitivity Analysis

Given that LULC models produce LULC maps, it is legitimate to assume that the results of U-SA should also be presented as maps. Here, I propose to define a *spatially-explicit* SA as a method of model evaluation in which outcome uncertainties and factor sensitivities are computed and mapped for every spatial entity within the area of interest (Figure 1).

Insert figure1about here

In addition, we can define a *spatially-inexplicit* SA as an approach in which scalar outputs (like land use fragmentation statistics), are substituted for sensitivity maps. As an example, consider a model of deforestation in the Amazon. In the spatially-inexplicit SA we could employ *acreage of deforested land* as the scalar model result used to evaluate outcome uncertainties and factor sensitivities. Conversely, we could employ spatially-explicit SA by using the LULC maps directly. In this case, we would generate deforestation frequency (probability) maps and map which factors (and where) contribute to the locational uncertainty of the potential deforestation regions (*SA-maps*).

3.1 Generating SA-maps

In this section I describe the procedure of spatially-explicit SA in LULC modeling. To demonstrate the concept, I assume that an ABM of residential development is employed.

Given a set of output binary maps (undeveloped and developed land) from *N* number of LULC model realizations, we can calculate a surface depicting the overall frequency (proportion) of development ($F_d map$):

$$F_d map : F_{dl} = \frac{d_l}{N}, \forall l \qquad (4)$$

Where, l is location (a raster cell, a vector point etc.) and d is the number of times the location was developed out of the total number of model executions N. The $F_d map$ serves as an input to S_k and ST_k calculation, which is performed independently for each location. Based on predefined thresholds, we then subdivide the F_d *map* into three distinct regions: [a] areas where development is very unlikely, [b] areas where development is almost certain, and [c] areas of development uncertainty (Brown et al., 2005). As the next step, we select locations l for S_k and ST_k evaluation using the following conditions:

$$(F_{dl} < \theta_{max}) \text{ AND } (F_{dl} > \theta_{min}) \text{ AND } (S_{kl} > \varsigma) \qquad (5)$$

$$\theta_{max} < 1; \ \theta_{min} > 0; \varsigma > 0$$

Where S_{kl} is the first order sensitivity index for factor k at location l, and θ_{max}, θ_{min} are development frequency thresholds. Note that the location selection conditions are applied to ST_{kl} in the same way as they are defined for S_{kl} in formula (5).

Therefore, for every location l, we check if its development frequency F_{dl} is larger than a user-defined threshold value θ_{min} and smaller than a user-defined threshold value θ_{max}. The θ_{min} is larger than zero and θ_{max} is smaller than one, so that the selected locations are characterized by spatial uncertainty[2]. Such sites comprise areas similar to the variant region (VR) as defined by Brown and colleagues (2005), that is, a region where locations are sometimes developed and sometimes not depending on the simulation run. Only when the location l passes both threshold values, we calculate the indices (i.e. S_{kl} and ST_{kl}) for every factor k for location l. When drawing the resultant sensitivity maps, we also want to avoid visualizing values of sensitivity indices close to zero, which obscures the exploratory analysis. Hence, a third condition for *SA-map*

[2] In extreme cases i.e. when $F_{dl} = 0$ or $F_{dl} = 1$, the total variance of development allocation at a given site is $V_l(Y) = 0$ and, as per formulas (2) and (3), we cannot compute sensitivity indices.

rendering is introduced, i.e. $S_{kl} > \varsigma$ ($ST_{kl} > \varsigma$), where ς is a threshold for a minimum sensitivity index value. The latter term means that we map only those locations that have sufficiently high sensitivities for a particular factor. Consequently, given the conditions in formula (5), the number of S_{kl} and ST_{kl} indices to display and analyze will be usually much smaller than the total number of locations (cells, points etc.) considered by the model and the spatial sensitivity analysis will be confined to relatively high uncertainty areas.

Below I provide an example to demonstrate the procedure. Suppose that a LULC produced a development frequency map ($F_d map$) as depicted in Figure 2(a). Assuming $\theta_{min} = .25$, $\theta_{max} = .85$, we partition the $F_d map$ into two regions (Figure 2(b)): approximately certain (depicted with *E*) and relatively uncertain (depicted with a fraction value F_{dl}). We then calculate sensitivities for all *non-E* locations (Figure 2(c)). Finally assuming $\varsigma = .05$, that is, excluding locations with $S_{kl} <= .05$ for factor *k*, we obtain an *SA-map* as shown in Figure 2(d).

Insert figure2 about here

Note that the *SA-maps* (separate for the first order and the total effect indices) serve as synthesized U-SA visualization. The reclassified development frequency map is first used to delineate the VR. Then, the sensitivity hot and cold spots are identified within that region. We have to acknowledge, however, that in this approach the heterogeneity of spatial uncertainty is lost due to the arbitrariness of θ_{min} and θ_{max} values. Thus, uncertainty is confined only to mapping out the variant regions. As a consequence, the proposed U-SA framework assumes that the non-variant regions are insensitive to the uncertain model inputs.

In summary, the spatial SA of LULC models uses the simulated land use maps directly as input to variance decomposition. This approach to spatial SA is different from the one described by Lilburne and Tarantola (2009), in which they use scalar outputs in place of maps. It is also distinct from the pioneering procedure described by Marrel et al. (2011), which, to the author's knowledge, is the only reported attempt to compute and analyze spatially-explicit sensitivities. The method proposed by Marrel and colleagues (2011) employs Gaussian process (Gp) metamodeling to emulate a model of groundwater contamination with radioactive waste. Instead of using the contamination diffusion model directly, they employ the Gp surrogate to calculate *SA-maps*. The Gp emulator has one important advantage over the exact-model SA, namely, computational efficiency (Marrel et al., 2011, Saltelli et al., 2008). In particular, a much smaller number of model executions is needed to converge to an acceptable solution. At the same time, it is important to recognize two notable shortcomings of metamodeling approaches. Firstly, since a metamodel is a mathematical approximation of the underlying more complex model, there is always a risk of inaccurate representation (Saltelli et al., 2008). Secondly, especially in the context of ABM, metamodeling may be envisioned as a less transparent simulation approach, that is, an efficient but difficult to interpret substitute.

4 Agent-Based Model and Simulation Setup

The following two sections outline the ABM of residential development and the computational procedure used to demonstrate the spatial SA framework.

Figure 3 depicts the diagram of the model. A given agent (*g*), who represents a residential developer, starts from selecting a random parcel *p* (a vector polygon) out of a pool of developable sites. This selection triggers a landscape retrieval procedure, which reads values of the underlying landscape characteristics (represented as raster layers) like land value of p (LV_p), landscape natural (scenic) beauty of p (NB_p), and distance to roads called accessibility of parcel p (AC_p). These characteristics serve as criteria used in g's decision making about land conversion from undeveloped to residential. It is further assumed that the decision making is affected by the density of development in the surrounding environment of p (D_p). Hence, a second step of landscape retrieval calculates the density of development in the adjacent neighborhood:

$$D_p = (\sum_n dev_n) / s; n \in N_p \tag{6}$$

Where N_p is the set of parcels that neighbor parcel p, n is its neighboring parcel, s is the number of parcels in N_p, and dev_n = 1 when n is developed and zero otherwise.

Insert figure3 about here

Agent's decision making is driven by its normalized weights (preferences) for the three selected landscape characteristics (WLV_g, WNB_g, and WAC_g, respectively). These weights are used to compute the utility (score) of a selected parcel p (S_{pg}), which is distinct for every agent g. The utility is calculated using the weighted summation aggregation function. Agent g is also equipped with a preference for development density called a neighborhood effect (E_g), scaled to a range of [-1,1]. When E_g is positive, the agent wants to settle close to other agents. When E_g is negative, the agent avoids other agents in the neighborhood. The rationale behind a different treatment of the three preferences and the E_g is as follows. While the retrieved landscape attribute values are only indirectly affected by the surrounding values (through spatial autocorrelation), the E_g represents a spatially explicit proximal agent-agent interaction. In other words, WLV_g, WNB_g, and WAC_g, are aspatial weights used in utility calculation (i.e. they are uniform over the entire study area) , whereas E_g contributes to a spatial weight (Feick and Hall, 2004) referred to as WE_{pg} (neighborhood effect weight of agent g for parcel p) independently applied to the initial parcel utility:

$$WE_{pg} = D_p * E_g \quad (7)$$

$$U_{pg} = S_{pg} + S_{pg} * WE_{pg} \quad (8)$$

Hence, the initial p's utility for g i.e. S_{pg} is post-hoc adjusted with the neighborhood effect weight to obtain the final utility – U_{pg} (Rinner and Heppleston, 2006). U_{pg} is further compared to the score of the best parcel so far evaluated (this score is initially set to zero). If, for agent g, the utility of p is higher than the best parcel utility, its best parcel is set to p.

The procedure of utility calculation is repeated for a sample of locations. Once the parcels are evaluated, the best scoring parcel is developed by the agent (Figure 3). This development feeds back to the landscape layers. Two feedbacks are employed as described in Ligmann-Zielinska and Sun (2010): land value increase (*LVI*) and natural beauty decrease (*NBD*). Both feedbacks can be applied to a variable size neighborhood *N*.

4.1 Computational Experiments

To control for factor distribution and hence facilitate the interpretation of results, I employed hypothetical data. Figure 4 shows the three landscape characteristics used in experimentation.

Insert figure4 about here

An additional discrete object (vector) layer was generated by partitioning the extent into 3600 equal-size square polygons. This partitioning of space introduces the decision-making layer – a parcel layer – which is separate from the landscape layers (continuous rasters) in figure 4. Therefore, the agents do not make decisions on changing the land use of individual pixels but rather, as is the case in real world residential decision-making, decide on the development of the whole parcel lot.

Nine heterogeneous developer agents are generated during ABM initialization. Each agent develops one parcel per time step by randomly sampling 6% of the parcels (216 sites). The agent has no preference for any particular parcel. Instead, they select the top scoring parcel from the 216 sites included in the sample. The model runs for 40 time steps. As a result, at the end of the simulation, 10% of the landscape becomes developed, which amounts to 360 parcels at the end of model execution.

The ABM is equipped with seven input parameters described in the previous section: three weights per agent (WLV_g, WNB_g, WAC_g), three agent-environment feedback mechanisms (*N*, *LVI*, *NBD*), and agent-agent interaction (E_g). The weights are further combined for all agents to form factor groups (Saltelli et al., 2004) i.e. *WLV*, *WNB*, *WAC*, and *E*, whereas the three feedbacks are single model-level factors. A factor group is a set of a particular type of factors (e.g. land value weights) assembled for the whole collection of agents. Hence, each factor group in the ABM is composed of nine single factors. Factor grouping allows for parameter reduction and, consequently, decreases the number of model executions. In the following sections, both single factors and factor groups used in the ABM are referred to as ‘factors’.

The probability density functions for input factors are summarized in Table 1. I used SimLab open source software (http://simlab.jrc.ec.europa.eu/) to generate samples and calculate sensitivity indices. I also employed the *extended FAST on groups* sampling method to generate factor samples and produce a total of 19275 model realizations. This sampling method puts together subsets of factors and applies the 'extended Fourier Amplitude Sensitivity Test' (FAST) method of sensitivity analysis, in which a multidimensional integral over all the uncertain model inputs is transformed to a one-dimensional integral used in outcome decomposition (Saltelli et al., 2004, Gomez-Delgado and Tarantola, 2006).

The ABM was implemented in Python programming language (http://www.python.org/) and is available from the model library on the OpenABM website (http://www.openabm.org/).

Insert table1 about here

4.2 U-SA Setup for the Demonstration Study

Individual spatial realizations of development patterns, that is, the final binary maps with developed and undeveloped locations, were first summarized by calculating the F_d *map* depicted in Figure 5. Given the gradient of change in the map, I assumed that the VR would be delineated with θ_{min} = 20% and θ_{max} = 90%. All sites (parcel centroids) that are bounded by these thresholds are grey in Figure 5. To compare the F_d *map* with an aggregate measure of development I employed the *directional distribution* spatial statistics, which is also called a standard deviational ellipse (de Smith et al., 2009). In particular, I used two measures of the directional distribution: ellipse area (DDA), and ellipse eccentricity (DDE). The former serves as a proxy for land development dispersion, whereas the latter quantifies development elongation, and, hence, unevenness of spatial distribution.

5 Results and Discussion

The DDA and DDE statistics are summarized in Figure 5-top and examples of the ellipses are mapped Figure 6. Observe that both statistics exhibit a substantial amount of variability (DDE: C_V=.23 and DDA: C_V=.18). A simple visual evaluation of DDE and DDA with the three landscape layers (Figures 4 and 6), suggests that DDA is the smallest when the majority of agents prefer *LV* over the other two attributes, whereas the DDE is the largest (most elongated) when a high importance is assigned to *AC* and *NB*, so that an area of a large marginal decrease in *NB* is selected. Moreover, DDA is the largest and DDE is the smallest when (for the majority of agents) *NB* plays the most important role in deciding where to develop. However, as described in the next section, these relationships prove to be more complex when sensitivity analysis is employed. In sum, Figure 5 summarizes the uncertainty associated with the development using two scalars (DDA and DDE), and one F_d *map* with the variant region sketched out.

Insert figure5 about here

Insert figure6 about here

5.1 Spatially-inexplicit and Spatially-explicit Sensitivity Analysis

Published research demonstrates that the S_k and ST_k indices are good composite indicators for a plethora of model performance characteristics (Saltelli et al., 2004, Tarantola et al., 2002). Below I suggest questions that could be addressed with spatial SA using the ABM of LULC as an example. The interpretation of SA results is based on the pie charts shown in Figure 7 and maps shown in Figure 8.

Insert figure7 about here

Insert figure8 about here

The major goal of SA is to identify those components of a particular model that affect its outcome uncertainty. In this context, we can therefore pose the following questions: *Which model factors influence outcome uncertainty the most? Which of them are spatially heterogeneous?*

Figure 7 depicts variance decomposition of the spatially-indirect (aggregate) representations of the LULC patterns. Both DDE and DDA show similar relative sensitivities. When analyzed singly (Figure 7 top), *E, WLV, WNB,* and *WAC* affect the dispersion of the ellipse statistics the most. The pronounced impact of *E* is not surprising since this factor can have a very different influence on development when set to a negative versus positive value. The negative value forces the agents to locate away from each other, while the positive value brings agents together, leading to a more compact development. The influence of weights on the uncertainty of DDE and DDA is also understandable. As seen in figure 4, all three layers are spatially distinct, providing quite different land use configurations for extreme preference values. The *WNB* is especially influential since the *NB* layer is the only one that exhibits patchiness leading to diverging land configurations (the other two layers are spatially symmetric). What is most intriguing, however, is the fact that the first order indices account only for a third of the unconditional output variance (38% for DDE and 31% for DDA, respectively). The remainder of the variance must be attributed to factor interactions.

The spatial distribution of the first order indices is rather uninteresting. Most of the factors exhibit a nearly random pattern. An exception is a map of the interaction effects (Figure 8-top), which is characterized by a considerable spatial heterogeneity with a few distinct clusters. I could therefore conclude that, when the factors of this model are taken singly, an aggregate representation of patterning (like the *directional distribution* statistics) is sufficient to evaluate factor's individual influence. In this case, mapping the S_k indices is unnecessary.

Some factors are characterized by nonlinear behavior, depending on the way the results are summarized. The *N* factor has a particularly interesting nature of sensitivity (Figures 6-bottom and 7-bottom). Its role in the aggregate pattern measures (DDE and DDA) is rather minimal. When its total effect is considered, it contributes only 7 to 10% to the variance of the ellipse measures. However, when mapped for individual locations, it proves to be more influential. Not only is it higher (i.e. explains from 10 to 20% of the total variance mapped for each location in VR) but it has, in general, higher values when compared to the other maps of ST_k, especially to the map of *E*.

Given that the majority of the output variance cannot be explained by individual factors alone, we should concentrate on analyzing the ST_k indices. The following questions can be asked: *Which combinations of factors influence the uncertainty of development locations, suggesting that LULC change is driven by a myriad of interconnected forces? Are there any regions where factors interact with each other?*

As mentioned above, over 60% of outcome uncertainty, measured using the directional distribution, can be attributed to factor interactivity (the grey fractions of the pies in Figure 7).

We can therefore infer that the model is highly interactive in terms of its variables. This factor interactivity can be better assessed when mapped against the VR (Figure 8-top). Factor interactions for all sites in the VR contribute 46-61% to the total variance, which further confirms that the model is functionally complex. Moreover, we can observe that factor interactions concentrate (cluster) in locations with large differences among the three landscape layers (Figure 4).

Based on the formulation of the three feedback factors (*LVI*, *NBD*, and *N*) we can hypothesize that their influence on the variability of land use configuration will be more pronounced when the interactions are investigated (Figure 7 bottom). Indeed, for both DDE and DDA, all three feedbacks contribute to variance in around 10% each (the ST_k pies), suggesting that a substantial part of this ABM behavior is fairly complex. Consequently, it would be incorrect to evaluate this model by applying OAT SA, which is so common in land use research. Moreover, if we finished the analysis at the first order level (for example by employing regression in place of variance decomposition) we would erroneously assume that *LVI*, *NBD*, and *N* have a negligible impact on output variance. Only if we employ ST_k calculations, can we determine the influence of the higher-order (coupling) factors. Finally, out of all maps of ST_k, the spatial maps for feedbacks are the least spatially dependent. Overall, they contribute to development uncertainty, but their influence is spread roughly evenly over the whole area.

We can also analyze spatial distributions of sensitivities of the selected factors one-at-a-time, focusing on locations where LULC is the most uncertain due to a given factor. We would then consider the following questions*: Where are the sensitivity hot (cold) spots of factor k located? How do they compare to hot (cold) spots of the other factors?*

To demonstrate this issue, we will focus on the sensitivity of *WNB* (Figure 8, bottom left). The spatial distribution of the total effect index *WNB* map follows the distribution of its underlying landscape layer (Figure 4, middle). Specifically, values of high total effect sensitivity are encountered in locations that are characterized by a steep gradient in *NB*, that is, areas where a rapid change in *NB* value is observed, signifying edge effects. Furthermore, when we compare the total effect SA maps of *WNB* and *WLV* (Figure 8) we can observe that areas of high sensitivity to *WNB* are also characterized by low sensitivity to *WLV* (e.g. hot spots of *WLV* are cold spots of *WNB*). This is not surprising, given that the ST_k is a fraction of the total variance.

Out of the seven factors, the *E* requires more explanation. As seen in Figure 8, a cluster of high values of *E* total effect sensitivity coincides with an area of the highest cumulative value of all three landscape attributes. This result suggests that, within the most competitive locations, where the utility of the developable parcel is high regardless of preferences, and where feedbacks (*LVI*, *NBD*, and *N*) only strengthen this relationship, it is the dual (push/pull) nature of *E* that drives agent's decision to develop. This interesting observation may corroborate the conventional belief that LULC ABMs are complex models driven by interrelated and nonlinear forces.

5.2 Why use spatially-explicit SA?

Saltelli et al. (2004) list four major settings that justify the use of SA: factor prioritization, factor fixing, variance reduction, and factor mapping. Here, I restate these rationales in the context of LULC.

Factor prioritization (FP) focuses on establishing which factors and in which locations generate most of the spatial output uncertainty. In the ABM example, the neighborhood effect proved to be the most influential in driving land use change. Its impact clustered around sites with high overall utility. Since FP provides guidance for data acquisition, with the spatially-explicit SA in the hypothetical problem, we could concentrate our data collection efforts around the area of the highest overall potential for development. In particular, we would survey the homebuyers in this area to find out if they seek solitude or want to live close to other residents. *Factor fixing and model simplification* is employed when, as a result of sensitivity analysis, we reduce the number of factors by fixing the non-influential ones to their best-fit values. In the context of LULC, if the sensitivities of a factor *k* are low and spatially independent we can, in principle, substitute *k*'s value with a constant. *Variance reduction* can be achieved when, after identifying input factors that drive model's spatial uncertainty the most, we focus on better estimation of these factors in the preferred locations. Finally, in *factor mapping,* we utilize spatial SA as a tool for scenario generation and policy exploration, by concentrating on factors that are primarily responsible for producing desirable landscape configurations (Lempert et al., 2003).

6 Summary and Conclusions

In this paper, I presented a spatially-explicit method of sensitivity analysis that aims at model's outcome uncertainty decomposition. I used sensitivity maps of input factors as a means of model outcome evaluation. In the context of land use and land cover change, mapping the sensitivities of model input factors allows for identifying those model variables that truly contribute to output uncertainly in the areas of interest. Indeed, sensitivity maps provide a way to visualize the influence of a particular model factor on the simulated LULC maps. Without them, we would not be able to isolate the effects of a particular LULC driver within the area of interest.

The proposed SA framework was demonstrated using an agent-based model of residential development. The *SA-maps* were contrasted with SA-pies of spatially-inexplicit aggregate outcome pattern statistics. This comparative analysis showed that the maps and pies are complementary. The pies show factor sensitivities relative to each other but give little information on their spatial distributions. The maps do not provide a transparent view of the relative 'between factor' sensitivities but allow for spatial evaluation of the 'within factor' variability of sensitivities. Specifically, the SSA helps to uncover which inputs contribute to the formation of specific clusters of high development uncertainty. Based on the ABM example, I argue that complex LULC models exhibit nonlinear and spatially heterogeneous behavior. Therefore, modelers should be judicious when replacing outcome maps of such models with simpler scalar representations.

Acknowledgements

The author would like to acknowledge the valuable comments of an anonymous reviewer and Dr. Robert Gilmore Pontius Jr, co-editor of this special issue, on the previous versions of the manuscript. Thank you.

Table 1 Probability Density Functions (PDFs) for seven factors used in the computational experiment (U – uniform, DU – discrete uniform, * – factor group). Values for all factors are independently drawn from their respective distributions. N_F is measured in raster cells. All other values are unitless.

Factor	**PDF**
Land value weight * (*WLV*)	U[0.0, 1.0]
Natural beauty weight * (*WNB*)	U[0.0, 1.0]
Accessibility weight * (*WAC*)	U[0.0, 1.0]
Neighborhood effect *(*E*)	U[-1.0, 1.0]
Natural beauty decrease (*NBD*)	U[0.01, 0.1]
Land value increase (*LVI*)	U[0.01, 0.1]
Size of neighborhood affected by feedbacks (*N*)	DU{1,2,3,4}

Figure Captions

Figure 1 Spatially-explicit uncertainty and sensitivity analysis.

Figure 2 Example of deriving a sensitivity map **(***Fd* – development frequency; *E* – excluded location; S_k – First order sensitivity index for factor *k***).**

Figure 3 Agent-based model used in this study.

Figure 4 Spatial distributions of landscape characteristics at the beginning of the simulation. Darker shade indicates higher (and hence more preferred) locations. Raster maps are composed of 121x121 cells.

Figure 5 Uncertainty analysis visualization: top – spatially-inexplicit, where development is summarized using area and eccentricity of the directional distribution statistics, bottom – spatially-explicit, where development is summarized using a frequency development map and a variant region (enclosed by a line).

Figure 6 Sample ABM output realizations. The dots symbolize the centers of developed parcels. *Eccentricity* and *Area* refer to the standard deviational ellipses of the development depicted in the maps.

Figure 7 Sensitivity pies of spatially aggregated outcome maps.

Figure 8 Sensitivity maps of LULC outcome maps.

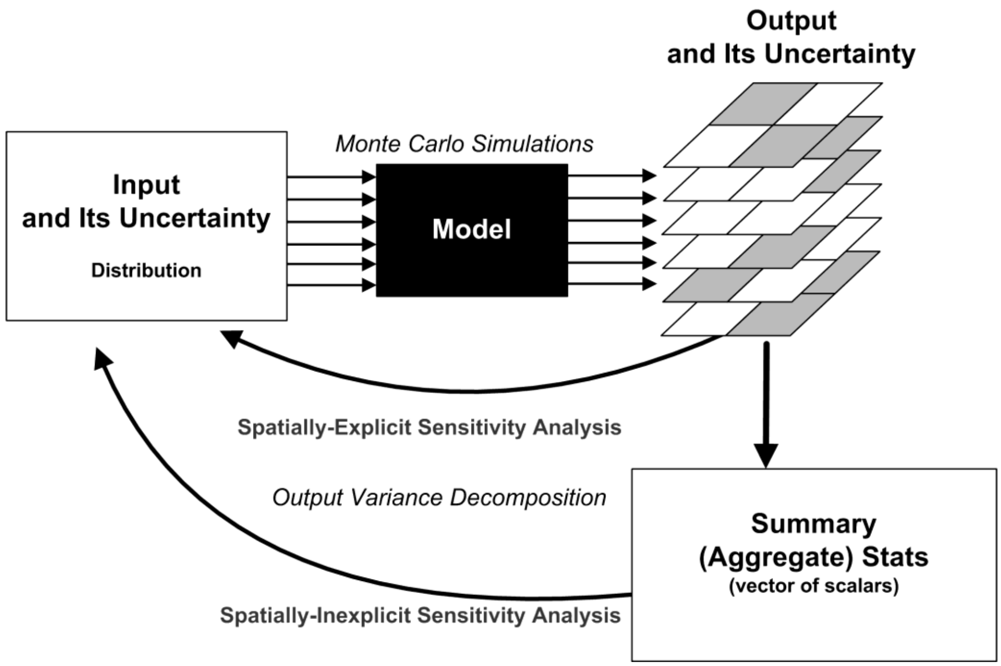


Figure 1

0	0	.005	.07
.3	.4	.7	.8
.5	.8	1	1
.9	.99	.97	1

(a) F_d map

E	E	E	E
.3	.4	.7	.8
.5	.8	E	E
E	E	E	E

(b) F_d map selection for sensitivity calculation

E	E	E	E
0	.03	.5	.1
.67	.55	E	E
E	E	E	E

(c) S_k map

E	E	E	.E
E	E	.5	.1
.67	.55	E	E
E	E	E	E

(d) Final (drawn) S_k map

Figure 2

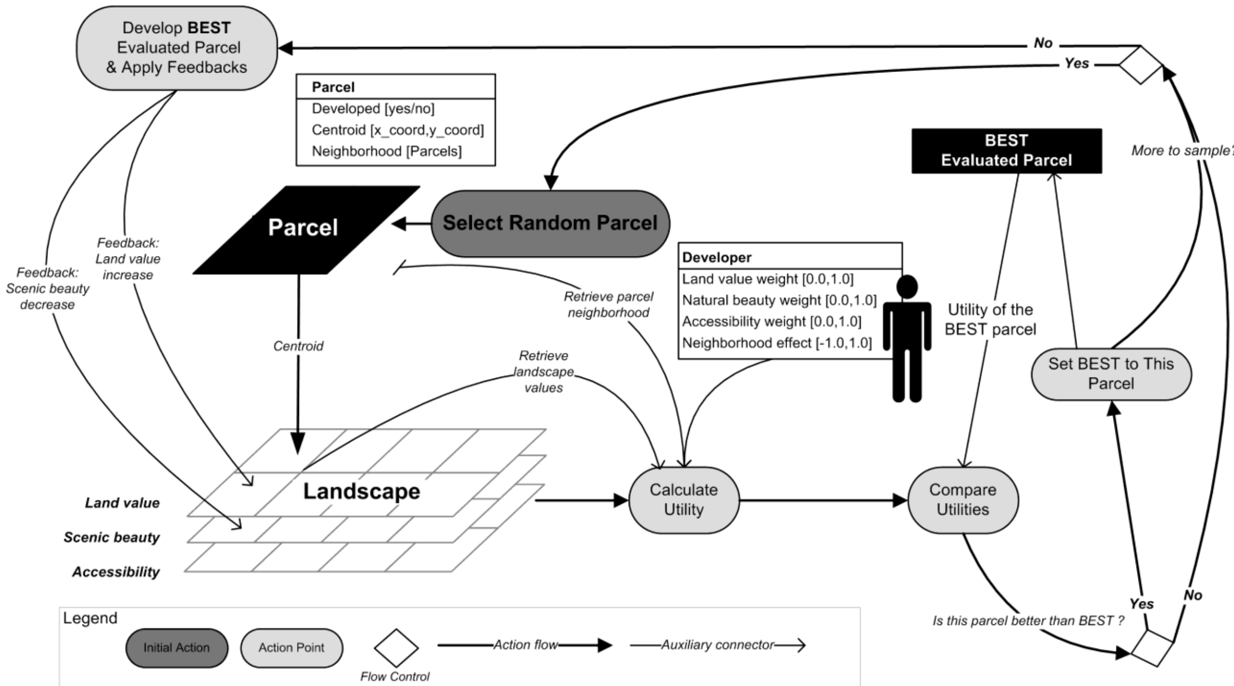


Figure 3

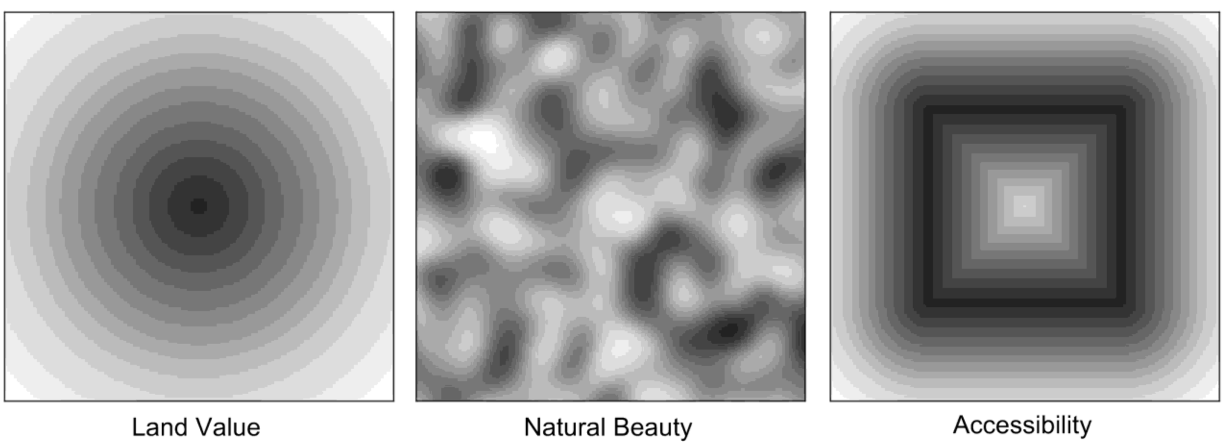


Figure 4

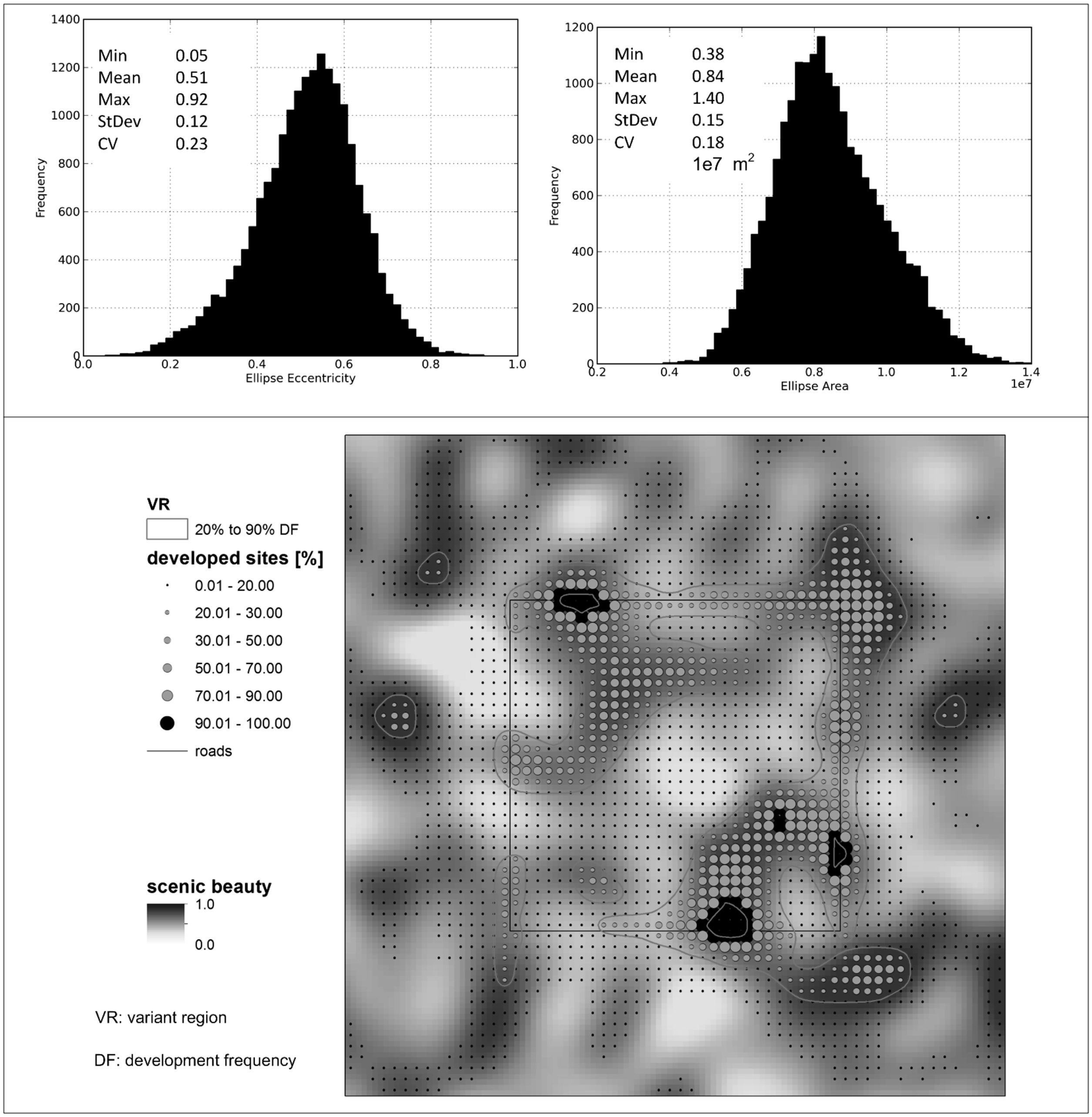

Min 0.05
Mean 0.51
Max 0.92
StDev 0.12
CV 0.23
Frequency
Ellipse Eccentricity
Min 0.38
Mean 0.84
Max 1.40
StDev 0.15
CV 0.18
1e7 m2
Frequency
Ellipse Area
1e7
VR
20% to 90% DF
developed sites [%]
0.01 - 20.00
20.01 - 30.00
30.01 - 50.00
50.01 - 70.00
70.01 - 90.00
90.01 - 100.00
roads
scenic beauty
1.0
0.0
VR: variant region
DF: development frequency

Figure 5

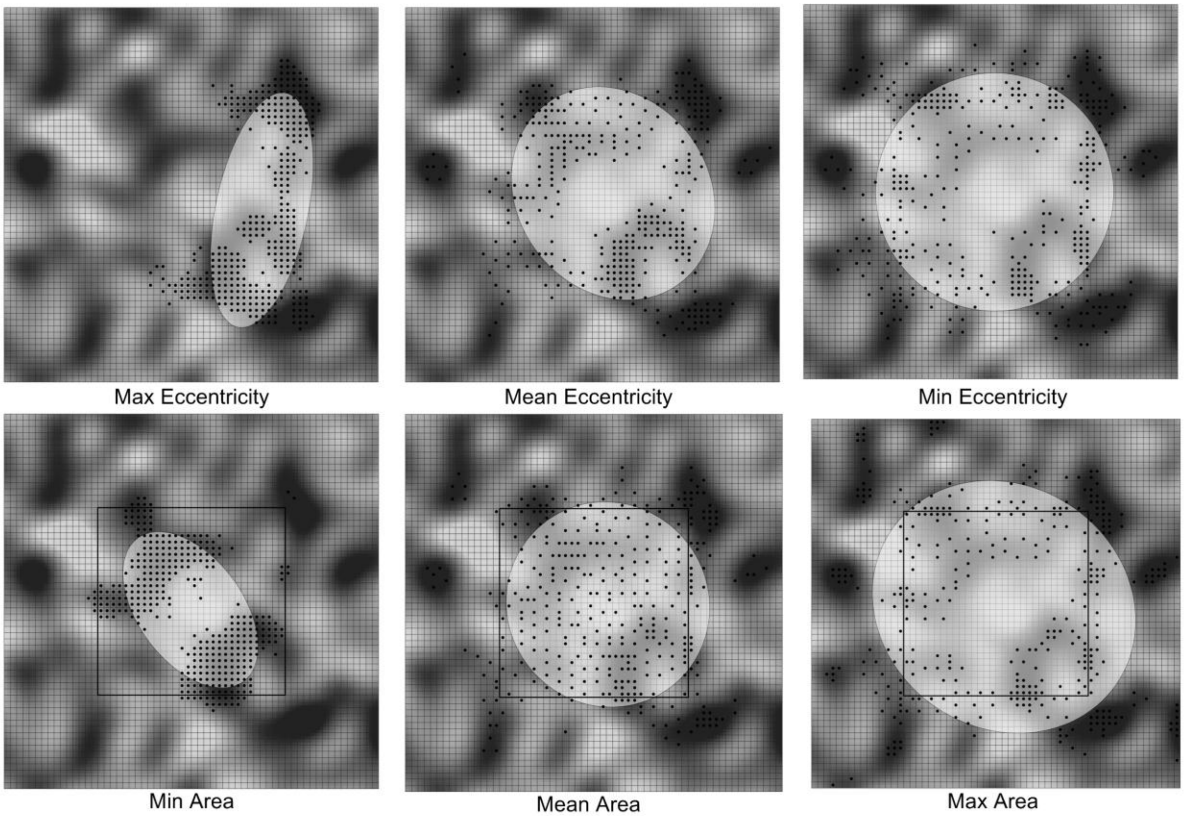

Max Eccentricity
Mean Eccentricity
Min Eccentricity
Min Area
Mean Area
Max Area

Figure 6

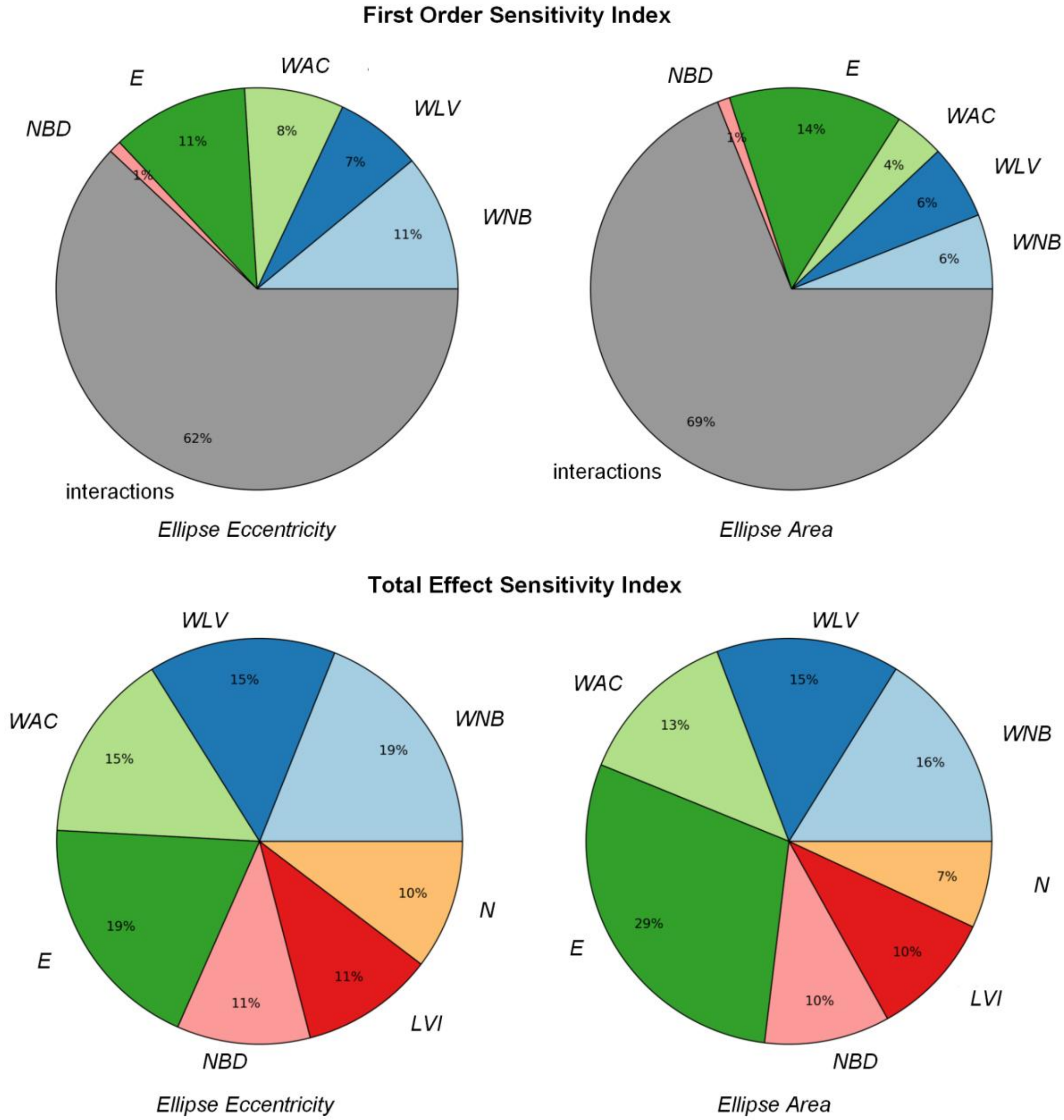
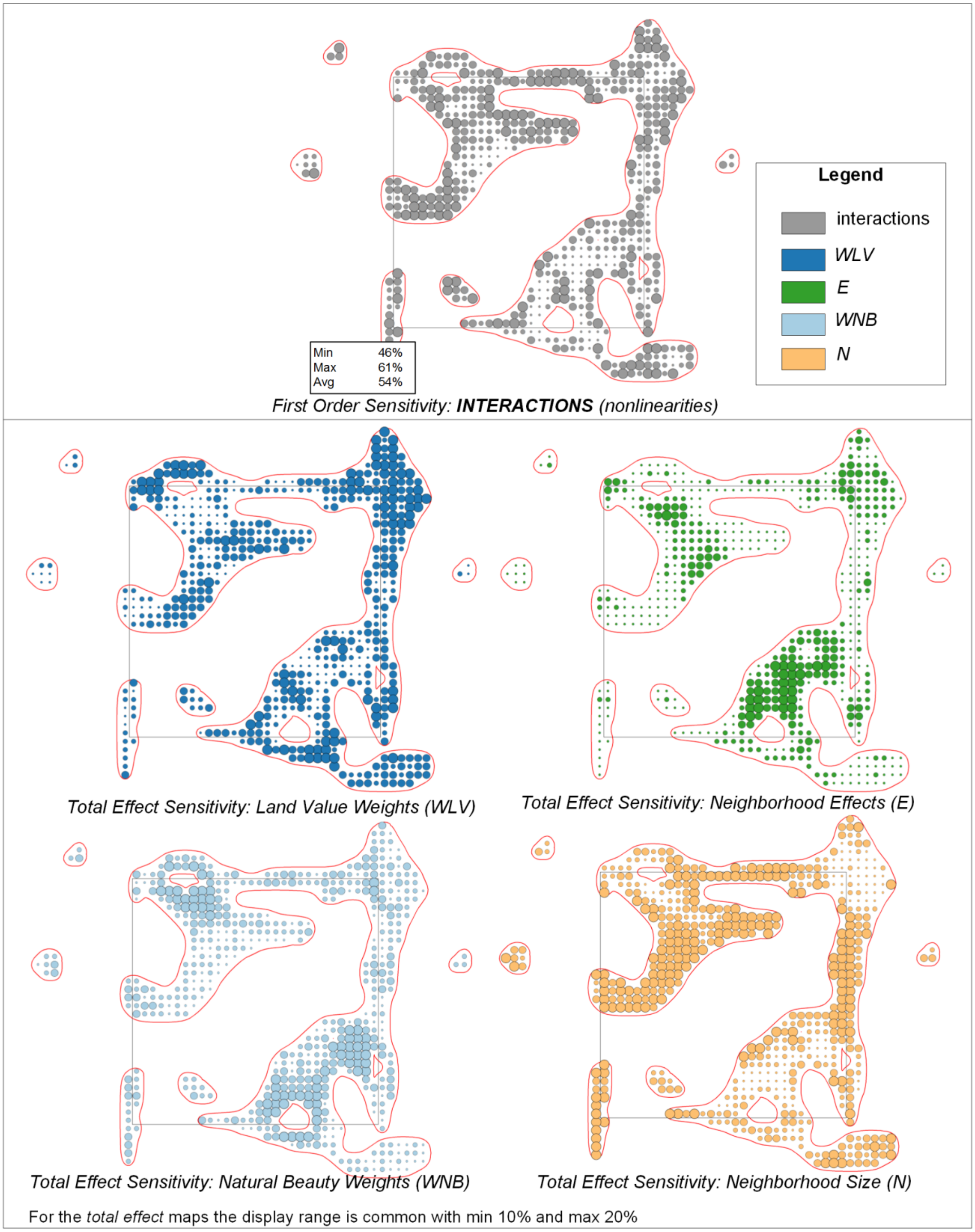

First Order Sensitivity Index
E
WAC
WLV
NBD
WNB
11%
8%
7%
1%
11%
62%
interactions
Ellipse Eccentricity
NBD
E
WAC
WLV
WNB
1%
14%
4%
6%
6%
69%
interactions
Ellipse Area
Total Effect Sensitivity Index
WLV
WAC
WNB
15%
15%
19%
19%
10%
N
E
11%
11%
LVI
NBD
Ellipse Eccentricity
WLV
WAC
15%
13%
WNB
16%
7%
N
E
29%
10%
LVI
10%
NBD
Ellipse Area

Figure 7



Legend
interactions
WLV
E
WNB
N
Min 46%
Max 61%
Avg 54%
First Order Sensitivity: INTERACTIONS (nonlinearities)
Total Effect Sensitivity: Land Value Weights (WLV)
Total Effect Sensitivity: Neighborhood Effects (E)
Total Effect Sensitivity: Natural Beauty Weights (WNB)
Total Effect Sensitivity: Neighborhood Size (N)
For the total effect maps the display range is common with min 10% and max 20%

Figure 8

References

AGARWAL, C., GREEN, G. M., GROVE, J. M., EVANS, T. P. & SCHWEIK, C. M. 2002. A review and assessment of land-use change models: dynamics of space, time, and human choice.

ASCOUGH II, J. C., MAIER, H. R., RAVALICO, J. K. & STRUDLEY, M. W. 2008. Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling,* 219, 383-399.

BROWN, D. G., PAGE, S., RIOLO, R., ZELLNER, M. & RAND, W. 2005. Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science,* 19, 153 - 174.

CAMPOLONGO, F., SALTELLI, A., SORENSEN, T. & TARANTOLA, S. 2000. Hitchhiker's Guide to Sensitivity Analysis. *In:* SALTELLI, A., CHAN, K. & SCOTT, E. M. (eds.) *Sensitivity Analysis.* Chichester, England: Wiley-Interscience.

CHEN, Y., YU, J. & KHAN, S. 2010. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. *Environmental Modelling & Software,* 25, 1582-1591.

CHU-AGOR, M. L., MUÑOZ-CARPENA, R., KIKER, G., EMANUELSSON, A. & LINKOV, I. 2011. Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. *Environmental Modelling & Software,* 26, 593-604.

CLAESSENS, L., SCHOORL, J. M., VERBURG, P. H., GERAEDTS, L. & VELDKAMP, A. 2009. Modelling interactions and feedback mechanisms between land use change and landscape processes. *Agriculture, Ecosystems & Environment,* 129, 157-170.

CROSETTO, M. & TARANTOLA, S. 2001. Uncertainty and sensitivity analysis: tools for GIS-based model implementation. *International Journal of Geographical Information Science,* 15, 415 - 437.

DE SMITH, M. J., GOODCHILD, M. F. & LONGLEY, P. A. 2009. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools,* Leicester, Troubador Publishing.

DIXON, B. 2005. Applicability of neuro-fuzzy techniques in predicting ground-water vulnerability: a GIS-based sensitivity analysis. *Journal of Hydrology,* 309, 17-38.

ESTRADA, V. & DIAZ, M. S. 2010. Global sensitivity analysis in the development of first principle-based eutrophication models. *Environmental Modelling & Software,* 25, 1539-1551.

FEICK, R. D. & HALL, G. B. 2004. A method for examining the spatial dimension of multi-criteria weight sensitivity. *International Journal of Geographical Information Science,* 18, 815-840.

GOMEZ-DELGADO, M. & TARANTOLA, S. 2006. GLOBAL sensitivity analysis, GIS and multi-criteria evaluation for a sustainable planning of a hazardous waste disposal site in Spain. *International Journal of Geographical Information Science,* 20, 449-466.

HELTON, J. C. & DAVIS, F. J. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety,* 81, 23-69.

HOSACK, G. R., LI, H. W. & ROSSIGNOL, P. A. 2009. Sensitivity of system stability to model structure. *Ecological Modelling,* 220, 1054-1062.

HUMPHRIES, H., BOURGERON, P. & REYNOLDS, K. 2010. Sensitivity Analysis of Land Unit Suitability for Conservation Using a Knowledge-Based System. *Environmental Management,* 46, 225-236.

LEMPERT, R., POPPER, S. & BANKES, S. 2003. Shaping the Next One Hundred Years New Methods for Quantitative, Long-Term Policy Analysis. Santa Monica, CA: RAND.

LI, Y., BRIMICOMBE, A. J. & RALPHS, M. P. 2000. Spatial data quality and sensitivity analysis in GIS and environmental modelling: the case of coastal oil spills. *Computers, Environment and Urban Systems,* 24, 95-108.

LIGMANN-ZIELINSKA, A. & JANKOWSKI, P. 2010. Exploring Normative Scenarios of Land Use Development Decisions with an Agent-Based Simulation Laboratory. *Computers, Environment and Urban Systems,* 34, 409-423.

LIGMANN-ZIELINSKA, A. & SUN, L. 2010. Applying Time Dependent Variance-Based Global Sensitivity Analysis to Represent the Dynamics of an Agent-Based Model of Land Use Change. *International Journal of Geographical Information Science,* 24**,** 1829-1850.

LILBURNE, L. & TARANTOLA, S. 2009. Sensitivity analysis of spatial models. *International Journal of Geographical Information Science,* 23**,** 151-168.

MAC NALLY, R. 1996. Hierarchical partitioning as an interpretative tool in multivariate inference. *Australian Journal of Ecology,* 21**,** 224-228.

MAKLER-PICK, V., GAL, G., GORFINE, M., HIPSEY, M. R. & CARMEL, Y. 2011. Sensitivity analysis for complex ecological models - A new approach. *Environmental Modelling & Software,* 26**,** 124-134.

MARREL, A., IOOSS, B., JULLIEN, M., LAURENT, B. & VOLKOVA, E. 2011. Global sensitivity analysis for models with spatially dependent outputs. *Environmetrics,* 22**,** 383-397.

MELBOURNE-THOMAS, J., JOHNSON, C. R. & FULTON, E. A. 2011. Characterizing sensitivity and uncertainty in a multiscale model of a complex coral reef system. *Ecological Modelling,* 222**,** 3320-3334.

MESSINA, J. P., EVANS, T. P., MANSON, S. M., SHORTRIDGE, A. M., DEADMAN, P. J. & VERBURG, P. H. 2008. Complex systems models and the management of error and uncertainty. *Journal of Land Use Science,* 3**,** 11 - 25.

MILLINGTON, J., PERRY, G. & ROMERO-CALCERRADA, R. 2007. Regression Techniques for Examining Land Use/Cover Change: A Case Study of a Mediterranean Landscape. *Ecosystems,* 10**,** 562-578.

MILNE, E., ASPINALL, R. & VELDKAMP, T. 2009. Integrated modelling of natural and social systems in land change science. *Landscape Ecology,* 24**,** 1145-1147.

PANTUS, F. J., ELLIS, N., POSSINGHAM, H. P. & VENABLES, W. 2008. Sensitivity analysis of spatially aggregated responses: A gradient-based method. *International Journal of Geographical Information Science,* 22**,** 575 - 599.

PONTIUS, R. & NEETI, N. 2010. Uncertainty in the difference between maps of future land change scenarios. *Sustainability Science,* 5**,** 39-50.

RAE, C., ROTHLEY, K. & DRAGICEVIC, S. 2007. Implications of error and uncertainty for an environmental planning scenario: A sensitivity analysis of GIS-based variables in a reserve design exercise. *Landscape and Urban Planning,* 79**,** 210-217.

RINDFUSS, R. R., ENTWISLE, B., WALSH, S. J., AN, L., BADENOCH, N., BROWN, D. G., DEADMAN, P., EVANS, T. P., FOX, J., GEOGHEGAN, J., GUTMANN, M., KELLY, M., LINDERMAN, M., LIU, J., MALANSON, G. P., MENA, C. F., MESSINA, J. P., MORAN, E. F., PARKER, D. C., PARTON, W., PRASARTKUL, P., ROBINSON, D. T., SAWANGDEE, Y., VANWEY, L. K. & VERBURG, P. H. 2008. Land use change: complexity and comparisons. *Journal of Land Use Science,* 3**,** 1 - 10.

RINNER, C. & HEPPLESTON, A. 2006. The spatial dimensions of multi-criteria evaluation - Case study of a home buyer's spatial decision support system. *In:* RAUBAL, M., MILLER, H. J., FRANK, A. U. & GOODCHILD, M. F. (eds.) *Geographic Information Science Proceedings 4th International Conference, GIScience 2006, Münster, Germany, September 20-23, 2006.* Berlin/Heidelberg: Springer.

ROURA-PASCUAL, N., KRUG, R. M., RICHARDSON, D. M. & HUI, C. 2010. Spatially-explicit sensitivity analysis for conservation management: exploring the influence of decisions in invasive alien plant management. *Diversity and Distributions,* 16**,** 426-438.

SAISANA, M., SALTELLI, A. & TARANTOLA, S. 2005. Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 168**,** 307-323.

SALTELLI, A. & ANNONI, P. 2010. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software,* 25**,** 1508-1517.

SALTELLI, A., CHAN, K. & SCOTT, E. M. 2000. *Sensitivity Analysis,* Chichester, England, Wiley-Interscience.

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.

SALTELLI, A., TARANTOLA, S., CAMPOLONGO, F. & RATTO, M. 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models,* Chichester, England, Wiley.

TARANTOLA, S., GIGLIOLI, N., JESINGHAUS, J. & SALTELLI, A. 2002. Can global sensitivity analysis steer the implementation of models for environmental assessments and decision-making? *Stochastic Environmental Research and Risk Assessment,* 16**,** 63-76.

TURNER, B. L., LAMBIN, E. F. & REENBERG, A. 2007. The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences,* 104**,** 20666-20671.

VARELLA, H., GUÉRIF, M. & BUIS, S. 2010. Global sensitivity analysis measures the quality of parameter estimation: The case of soil parameters and a crop model. *Environmental Modelling & Software,* 25**,** 310-319.

VERBURG, P. H. 2006. Simulating feedbacks in land use and land cover change models. *Landscape Ecology,* 21**,** 1171-1183.

VERBURG, P. H., KOK, K., PONTIUS, J. R. G. & VELDKAMP, A. 2006. Modeling Land-Use and Land-Cover Change. *In:* LAMBIN, E. F. & GEIST, H. J. (eds.) *Land-Use and Land-Cover Change: Local Processes and Global Impacts.* Berlin: Springer.

VERBURG, P. H., VAN ECK, J. R. R., DE NIJS, T. C. M., DIJST, M. J. & SCHOT, P. 2004. Determinants of land-use change patterns in the Netherlands. *Environment and Planning B: Planning and Design,* 31**,** 125-150.

VISSER, H., PETERSEN, A. C., BEUSEN, A. H. W., HEUBERGER, P. S. C. & JANSSEN, P. H. M. 2006. Guidance for uncertainty assessment and communication: checklist for uncertainty in spatial information and visualising spatial uncertainty. *In:* JANSSEN, P. & VAN DER SLUIJS, J. (eds.) *RIVM/MNP Guidance on Uncertainty Assessment and Communication.* Netherlands Environmental Assessment Agency (RIVM/MNP).

WARMINK, J. J., JANSSEN, J. A. E. B., BOOIJ, M. J. & KROL, M. S. 2010. Identification and classification of uncertainties in the application of environmental models. *Environmental Modelling & Software,* 25**,** 1518-1527.

XU, C., HE, H. S., HU, Y., CHANG, Y., LI, X. & BU, R. 2005. Latin hypercube sampling and geostatistical modeling of spatial uncertainty in a spatially explicit forest landscape model simulation. *Ecological Modelling,* 185**,** 255-269.

YANG, J. 2011. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. *Environmental Modelling & Software,* 26**,** 444-457.