**RESEARCH ARTICLE**

# Applying Time Dependent Variance-Based Global Sensitivity Analysis to Represent the Dynamics of an Agent-Based Model of Land Use Change

Arika Ligmann-Zielinska[a*] and Libo Sun[b]

[a]*Department of Geography and Environmental Science and Policy Program, Michigan State University, East Lansing, USA;*

[b]*Department of Statistics and Probability, Michigan State University, East Lansing, USA*

Reprint

Published in 2010 in IJGIS, 24(12), 1829-50

*Corresponding author: Email: ligmannz@msu.edu

**Abstract**

The growing body of knowledge on modeling land use systems points to epistemic uncertainty as one of the challenging obstacles in development and application of agent-based models (ABM). To decrease outcome uncertainty, sensitivity analysis (SA) is performed as part of model verification and validation. Oftentimes, however, it is inadequately addressed, partly due to the lack of tools and techniques that focus on an explicit evaluation of ABM dynamics. The nonlinear processes, inherent in such models, necessitate longitudinal SA with time path investigation of input-output relationships of endogenous variables. In response to the outlined deficiencies, the reported study investigates the potential of time dependent global sensitivity analysis (time-GSA) in examining the dynamics of outcome uncertainty of a simple ABM of land use change. Specifically, we apply first and total order sensitivity indices to decompose variance of output landscape fragmentation, apportioned to model inputs for multiple time steps and multiple realizations of the ABM. We focus the analysis on selected complex systems characteristics including preference uncertainty, path dependence, access to information, and magnitude of interactions and feedbacks. We conclude that the factor sensitivity measures vary significantly during model execution. Consequently, a static snapshot of ABM sensitivity, taken at the end of the simulation, is inadequate when deciding on factor prioritization and reduction. Assuming that ABM dynamics is a result of factor interaction, we observe a distinct time lag of nonlinearity, which unfolds after the formation of the seeds of development. Therefore, we argue for further application of time-GSA in ABM as one of the visual quantitative techniques contributing to evaluation of ABM nonlinearity.

## 1 Introduction

The complexity of land use systems (LUS) originates from interplay of many elements and elemental drivers (Claessens et al., 2009) that form a network of reciprocal relationships resulting in nonlinearities, path dependence, and feedbacks across scale, time, and space (Parker et al., 2008, Verburg, 2006, Liu et al., 2007, Bennett and McGinnis, 2008). LUS complexity imposes a number of challenges to long term policy making and spatial knowledge discovery. At the same time we recognize the pressing issues of unsustainable growth, which is a consequence of poorly managed land. To address these challenges, we often employ agent based models (ABMs), in which LUS actors like humans, organizations and institutions are modeled using individual entities that interact with each other and change their common land use through direct and indirect impact (Bousquet and Le Page, 2004, Matthews et al., 2007, Parker et al., 2003).

ABMs are inevitably prone to epistemic sources of uncertainty reflecting our poor knowledge of the processes and data describing the system (Helton and Burmaster, 1996, Saltelli et al., 1999). This uncertainty imposes one of the major obstacles in developing fully functional models. As Verburg (2006) and others point out, feedbacks in land use change systems cause strong model dependence on initial conditions, which calls for novel techniques of sensitivity analysis (SA), tailored to exploration of many possible land use trajectories (Brown et al., 2005, Manson, 2001, Oreskes et al., 1994).

The application of SA to ABM verification and validation is not new (Castella et al., 2005, An et al., 2005, Brown et al., 2008, Brown and Robinson, 2006, Li and Liu, 2007, Burke et al., 2006, Becu et al., 2003, Mosler and Martens, 2008, Schluter and Pahl-Wostl, 2007, Topping et al., 2010). Oftentimes, however, it is inadequately addressed, partly due to the lack of tools and techniques that focus on an explicit evaluation of ABM dynamics (Parker et al., 2003, Richiardi et al., 2006). Apart from explaining empirical regularities, complex systems models should allow for identifying the most important mechanisms of change that are hard to observe in the field (Verburg et al., 2006, Brugnach, 2005, Brugnach et al., 2008, Irwin et al., 2009). Consequently, the nonlinear processes, inherent in LUS models, necessitate longitudinal SA with time path investigation of input-output relationships of endogenous variables (Parker et al., 2003).

In response to the outlined deficiencies in SA as practiced in land use ABMs, we propose to use time dependent global sensitivity analysis (*time*-GSA), in the form of time series of sensitivity indices. Our approach borrows from research on GSA undertaken by Saltelli and colleagues (Saltelli et al., 2000, Lilburne and Tarantola, 2009, Gomez-Delgado and Tarantola, 2006, Crosetto and Tarantola, 2001, Saltelli et al., 1999). Traditionally, SA has been performed by evaluating the changes in final model results due to input perturbations. Contrary to this common approach, *time*-GSA opens possibilities to address a new kind of questions: What are the dynamics of model sensitivity to input factor variability? To what extent does the model behave nonlinearly? Are there any regions in the time series of sensitivities where a particular input variable dominates the others? Which input parameters contribute towards model stability? What combinations of thresholds make the model switch from one regime to another? As a consequence, the major advantage of *time*-GSA, as presented in this study, is the possibility of examining parameter sensitivity throughout model execution as opposed to ‘final time step’ testing.

In the following sections, we report on variance-based *time*-GSA of a simple agent-based model of residential land use. We focus the analysis on selected complex systems characteristics including: [a] preference variability related to three landscape attributes, [b] spatial path dependence (Brown et al., 2005, Manson, 2001), [c] magnitude of interactions defined by agent population size, [d] access to information, represented by the size of the sample of developable locations known to agents, and [e] the magnitude of agent-environment feedbacks.

The remainder of the manuscript is structured as follows. Section two summarizes the foundations of variance-based GSA. Section three outlines the ABM model used in this study. What follows is a description of spatial datasets and experiments designed to compare and contrast the impact of input perturbations on the dynamics of the generated land use pattern. In the final sections we summarize the results of *time*-GSA and conclude the paper by discussing future research challenges.

## 2 Variance-based global sensitivity analysis

Global sensitivity analysis studies the variability of model outputs due to a broad range of simultaneous perturbations in the whole set of uncertain input factors, which are examined independently and in combinations (Campolongo et al., 2000). GSA derives from numerical predictive modeling and experimental science and has been used in conjunction with equation-based models in the aspatial context (Saisana et al., 2005, Saltelli et al., 2000, Varella et al., 2010, Ziehn and Tomlin, 2009) as well as GIS-based applications (Lilburne and Tarantola, 2009, Crosetto et al., 2000, Gomez-Delgado and Tarantola, 2006, Tarantola et al., 2002).

Various GSA approaches have been proposed and evaluated (Campolongo et al., 2000). Here we utilize variance-based GSA, which obviates the assumptions of linearity inherent in regression-based approaches that are ill-suited for dynamic complex system modeling (Manson, 2007). Variance-based GSA decomposes the variance ($V$) of model output apportioned to changes in $k$ model inputs that are represented singly ($V_i$), and in combinations with an increasing level of dimensionality (Saisana et al., 2005, Gomez-Delgado and Tarantola, 2006):

$$V = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + \ldots + V_{12\ldots k} \quad (1)$$

For example, *Vij* represents the sensitivity of model output *Y* to the interaction between inputs *Xi* and *Xj*, $V_{ijm}$ is the share in the overall sensitivity of the model explained by the interaction among $X_i$, $X_j$, and $X_m$, and so forth. Such defined variance is further used to calculate first order (Si) and total-effect (STi) indices for every input parameter *i* (where *i*=1 to *k*):

$$Si = \frac{V_i}{V} = \frac{V_{X_i}[E_{X_{-i}}(Y \mid X_i)]}{V(Y)} \quad (2)$$

$$STi = \frac{V(Y) - V_{X_{-i}}[E_{X_i}(Y \mid X_{-i})]}{V(Y)} = S_i + S_{ij} + S_{im} + S_{ijm} + \ldots + S_{ij\ldots k} \quad (3)$$

Si quantifies a fractional contribution of an uncertain factor *i* to *V* taken independently from the other *k*-1 factors. STi is used to find the overall contribution of a given input *i* including its interaction with other inputs (Homma and Saltelli, 1996). $V_{X_{-i}}[E_{X_i}(Y \mid X_{-i})]$ stands for the total contribution of conditional variance of Y due to all *non-i* inputs so that STi includes first order and higher order terms that involve factor *i* (Saisana et al., 2005).

To calculate Si and STi we will use extended Sobol estimation procedure described in Saltelli (2002) and Lilburne and Tarantola (2009), which is available in SimLab (http://simlab.jrc.ec.europa.eu/). Since we focus on time variant GSA, we will compute normalized Si and STi for every time step of model execution, effectively creating a time series of output variance sensitivity. The interpretation of the pairs of indices (Si, STi) is summarized in Table 1.

*[Insert Table 1 about here]*

### *2.1 Experimental procedure*

To explore the land use ABM with *time*-GSA based on Sobol experimental design, the following steps are performed for any given computational experiment (Saisana et al., 2005, Lilburne and Tarantola, 2009):

1) Randomly generate *n* Monte Carlo (MC) input samples based on predefined probability density (or mass) functions using a quasi-random design
2) Execute the model *n* times for *t* time steps
3) Calculate fragmentation statistics of output land use patterns
4) Select a non-correlated subset of fragmentation statistics
5) Use the selected fragmentation statistics together with the input samples to calculate time series of Si and STi

Note that we will analyze the dynamics of pattern formation using aggregate measures in the form of landscape pattern metrics.

## 3 Agent-based model of residential land use change

The ABM presented in this research (Figure 1) is a simplified implementation of a real estate development process as outlined in Barrett and Blair (1988), focusing on site suitability assessment and investment decisions. The ABM is a discrete time stochastic model of decentralized decision making, composed of a cellular space (landscape), and developer agents, who are the major driving force of land use change. Three landscape attribute maps are utilized in agent decision process: *land value* representing the economic characteristics of the local property market, *scenic (natural) beauty* symbolizing the natural amenities of the area under consideration, and *accessibility* to already developed locations measured as Euclidean distance (Brown et al., 2005, Ligmann-Zielinska, 2009). The agents are equipped with preferences for the three landscape attributes. They make their decisions using an ideal point aggregation function (outlined below), followed by ordered choice heuristics (Benenson and Torrens 2004). In case of conflict over a piece of land, a simple rank-based bidding is employed to assign the parcel to one of the competing agents.

*[Insert Figure 1 about here]*

***3.1 Agent decision process***

Each agent enters the landscape at the beginning of the simulation, draws a sample of developable locations, and evaluates them using an ideal point (IP) decision rule (Hwang and Yoon, 1981, Malczewski, 1999):

1) For every standardized landscape attribute map, an agent finds the best and the worst values within the lattice, which are respectively called an ideal and a nadir.
2) The agent calculates gains and losses for the attributes:

$$A\_gain_{xy} = A_{xy} - A^{\wedge}$$

$$A\_loss_{xy} = A_{xy} - A^{*}$$

   Where
   $A\_gain_{xy}$ is a gain for attribute $A$ at location $xy$
   $A^{\wedge}$ is the nadir within the whole attribute $A$ layer
   $A\_loss_{xy}$ is a loss for attribute $A$ at location $xy$
   $A^{*}$ is the ideal within the whole attribute $A$ layer
   $A_{xy}$ is the value of attribute $A$ at location $xy$
3) The agent calculates the separations from ideals and nadirs:

$$S^{*}_{dxy} = \left( \sum_{A} \left( w_{dA} * A\_loss_{xy} \right)^{2} \right)^{0.5}$$

$$S^{\wedge}_{dxy} = \left( \sum_{A} \left( w_{dA} * A\_gain_{xy} \right)^{2} \right)^{0.5}$$

   Where
   $w_{dA}$ is developer agent's $d$ preference (weight) for attribute $A$
   $S^{*}_{dxy}$ is the separation from ideal for agent $d$ at location $xy$
   $S^{\wedge}_{dxy}$ is the separation from nadir for agent $d$ at location $xy$
4) Finally, agent $d$ calculates the utility $\left(U_{dxy}\right)$ of location $xy$:

$$U_{dxy} = \frac{S^{\wedge}_{dxy}}{S^{\wedge}_{dxy} + S^{*}_{dxy}} \qquad (4)$$

Applying IP for site utility calculation allows for an explicit representation of reference frames (ideals and nadirs) in decision making, which is consistent with the theory of choice psychology (Tversky and Kahneman, 1981). The IP decision rule belongs to the category of multi-attribute utility techniques and, unlike weighted linear combination, does not require the evaluation attributes to be independent.

Based on the utilities, the agent compiles an investment set by ordering the sampled sites from the best one to the worst one. Starting from the highest scoring site, the agent picks the desired (demanded) number of sites and creates a pending investment plan. When a site is selected by more than one developer agent, a rank-based bidding rule is invoked, in which the site of conflict is assigned to an agent who ranked the site the highest. When this is the case, the other competing agents lose the site from their investment plans. To compensate for their missing demands, these agents revisit their investment sets and choose additional sites to update their investment plans. This process continues until all conflicts are resolved and all demand is allocated to sites, which are then converted to developed land.

### *3.2 Agent-environment feedback*

Agents' decisions cause a number of reciprocal changes in the environment (Verburg, 2006). The ABM presented in this study emulates such feedback using a CA-based approach (Manson, 2001). First and foremost, the land use undergoes change due to the conversion of an undeveloped area to a developed one. Moreover, two decision attributes, namely, land value and scenic beauty, are updated as a result of property development. Specifically, the development has a positive impact on land value and a negative impact on scenic beauty. The explicit feedbacks are represented as a ratio of change in the nearest neighborhood of every developed cell.

The land value is increased using the following formula:

$$V_n(t+1) = V_n(t) + I_{lv} * V_p; \forall n \in N_p \quad (5)$$

Where *V* is the land value attribute, *p* is a developed parcel, and *n* is the neighboring cell of *p*. Therefore, with every time step passed, every *n*, located in the neighborhood of *p*, increases its land value by a fraction $I_{lv}$ of $p$'s land value. In this study, we assume that $N_p$ is a 3x3 Moore neighborhood, and $I_{lv}$ is in the range [0.0, 1.0]. If the updated land value exceeds 1.0, we set it to 1.0, which is the maximum possible score for all of the standardized attribute maps. Increasing $I_{lv}$ enhances the positive effect that the development has on the land value of adjacent locations. Observe that the new value is directly influenced by the development. Moreover, for $n$ with more than one adjacent $p$, its land value increases by $I_{lv}$ of each adjacent $p$, symbolizing a cumulative effect of neighborhood development.

Similarly to land value, scenic beauty is also updated within the neighborhood of $s$:

$$B_n(t+1) = B_n(t) - D_{sb} * B_p; \forall n \in N_p \tag{6}$$

Where $B$ is the scenic beauty attribute. Therefore, with every time step passed, every $n$, located in the neighborhood of $p$, decreases its scenic beauty score by a fraction $D_{sb}$ of $p$'s scenic beauty. We assume that $D_{sb}$ is in the range [0.0, 1.0]. If the updated scenic beauty falls below 0.0, we set it to 0.0, which is the minimum possible score for all of the standardized layers of this model. Obviously, increasing $D_{sb}$ intensifies the negative impact of development on scenic beauty of adjacent locations. Every $n$ with more than one adjacent $p$ decreases its beauty by a fraction $D_{sb}$ of each bordering $p$, symbolizing a cumulative effect of the whole neighboring development.

We assume for simplicity that the feedback coefficients are constant over time. In reality, it would be more appropriate to introduce attenuating feedbacks, represented as exponential decay functions, in which $I_{lv}$ and $D_{sb}$ decrease with time that passed since the development took place.

## 4 Landscape and input parameter values

The purpose of this study is to illustrate the utility of *time*-GSA using eleven computational experiments. To gain full control over experimentation, we employ a highly stylized environment composed of 14641 cells (121*121) and initialized with two landscape attribute maps depicted in Figure 2.

*[Insert Figure 2 about here]*

Preferred locations are represented by darker shades. Consequently, the land value layer scores highest in the center of the map, whereas the scenic beauty is composed of hot spot and cold spot value clusters. To reduce exogenous path dependence, the input land use map is set to undeveloped for every cell in the lattice. The model is executed for 70 time steps, which is a compromise between the computational cost of simulations and the minimal length of model execution that provides a representative time series of land use dynamics. Finally, the ABM generates constant development of 40 cells per time step, which amounts to 2800 developed cells (19.12%) at the end of model execution.

### *4.1 Input factor probability distributions*

Our major objective is to assess land use pattern sensitivity due to input parameter variation and sample size variation (Richiardi et al., 2006). In all experiments, also referred to as scenarios, we use seven independent input factors ($k$=7): number of developer agents ($\boldsymbol{A_{num}}$), preference (weight) for land value ($\boldsymbol{W_{lv}}$), preference for scenic beauty ($\boldsymbol{W_{sb}}$), preference for accessibility ($\boldsymbol{W_{a}}$), the fraction of developable land known to agents – sample size ($\boldsymbol{S_{s}}$), neighborhood decrease in scenic beauty ($\boldsymbol{D_{sb}}$), and neighborhood increase in land value ($\boldsymbol{I_{lv}}$). Table 2 lists the probability density functions (PDFs) reflecting the variable input factors for every scenario. It should be stressed that the PDFs are critical to defining model behavior, since the variances introduced into the PDFs reflect the magnitudes of variability associated with a particular input.

*[Insert Table 2 about here]*

We start from establishing a base case scenario, which represents a point of departure for the consecutive experiments. For all continuous factors we assume a normal distribution. In the majority of cases, we set the standard deviation of a factor to half of the mean value to impose a moderate fluctuation around the mean. Below we describe the PDFs in more detail.

The base-case preference allocation for landscape attributes is set to a third of unity and reflects equal preferences for all three characteristics. When one preference dominates the others, we increase the average value of the weight leaving the standard deviation unchanged. Preliminary experimentation with $D_{sb}$ and $I_{lv}$ showed that these two factors can have a huge impact on model behavior. In particular, when $D_{sb} > 0.1$ or $I_{lv} > 0.1$ the model immediately converges to equilibrium calculated as landscape pattern metrics. Therefore, in the consecutive experiments, we established 0.1 as the boundary condition for the normal PDFs of $D_{sb}$ and $I_{lv}$. Population size (number of agents) is defined using the discrete non-uniform distribution presented in Table 2. The following constant is assumed for every time step: $A_{num}$*demand_per_agent* = 40. In other words, when the number of agents increases, the demand for land per agent proportionally decreases, so that the developed area per time step remains constant. Furthermore, we assume that, for a particular ABM execution, all agents are homogeneous in preferences. Thus the variability of preferences is studied among alternate development histories (model runs) rather than among agents. While we acknowledge that such an assumption is a considerable simplification of a real land use system, we needed to reduce model dimensionality in order to trace the causality of the patterning.

To determine the necessary number of MC runs for the Sobol sample input generation procedure, we use the following formula (SimLab http://simlab.jrc.ec.europa.eu/): $(2k + 2)*2^{(4+j)}$, where $j>=0$ (Sobol, 2001). Due to the fact that our model is computationally demanding, taking over 10 min on Dell Precision T5400 Intel Xeon 3.16GHz processor, we decided to set $j = 1$, which amounts to 512 ABM executions per scenario.

The following operational questions were used as the basis for experimental design (Table 2):

Experiments 2 and 3: *What is the impact of variable preferences for landscape characteristics on outcome development pattern?*

Experiments 4 and 5: *What are the implications of agent-agent feedbacks on land use clustering?*

Experiments 6 and 7: *How does an increase (decrease) in agent-landscape feedback intensity influence the spatial outcome? Will the changes in feedback intensity result in a gradual modification of the pattern or will they cause an abrupt change?*

Experiments 8 and 9: *Assuming a constant amount of development, to what extent does the number of homogeneous agents influence output variability?*

Experiments 10 and 11: *How does agent's access to opportunities (information) shape land use morphology?*

Observe that we not only compare the distribution of outcome patterns, referred to as uncertainty analysis (UA), but we also analyze the influence of input factors on output variance.

Through *time*-GSA we try to infer which factors play the major role in obtaining a particular output distribution.

## 5 Experiments and results

The primary results of ABM simulations are binary land use maps of developed and undeveloped locations for every time step of each model execution. Figure 3 shows selected development frequency maps, calculated using the mean development per cell among all model executions of a given experiment. Observe a considerable variability in the spatial distribution of land development among scenarios.

*[Insert Figure 3 about here]*

Using two-dimensional array of output values, where each cell in the landscape is a separate outcome variable, may be computationally intractable (Lilburne and Tarantola, 2009). Therefore, in place of spatial layers, we use fragmentation statistics as the representative summaries of patterns.

### *5.1 Selecting fragmentation statistics for time-GSA analysis*

To summarize the patterns, we calculated the following fragmentation statistics (McGarigal and Marks, 1995): Largest Patch Index (LPI), Patch Density (PD), Edge Density (ED), Euclidean Nearest Neighbor Distance (ENN), and Aggregation Index (AI). From the computed spatial measures, ENN was dropped from post-processing analysis due to null values for some maps, in which the development clustered into one big patch (hence no neighbors). From the remaining statistics, we selected AI, PD, and LPI based on the linear correlation coefficient ($0.1<|r|<0.8$ for all pairs of the selected statistics). LPI is measured as the percentage of the landscape occupied by the largest patch of developed land, and is therefore dependent on the absolute area of development in any given time step. Therefore, we further normalized the LPI values so that they reflect the fraction of the developed area occupied by the largest patch, rather than the fraction of the total area under investigation. In this respect, we obtained a measure of clustering relative to the current development, and diminished the impact of linear constant growth. We call this measure Largest Patch of Developed area Index (LPDI).

### *5.2 Sensitivity estimation errors*

The accuracy of Si and STi depends on the number of model executions (Crosetto and Tarantola, 2001). Values of indices below zero usually indicate approximation errors due to a small MC sample. When these errors are minor (e.g. 10-15%) we can assume that the factors they represent are unimportant and their Si or STi can be reset to zero (Saltelli et al., 2008). Table 3 summarizes the highest errors obtained from *k* input factors for all eleven experiments. The errors were measured as values of Si or STi that fall below zero, normalized to 100%. Notice that, for experiments 2 and 7, the fragmentation statistics considerably exceed the 15% threshold. We consequently decided to repeat these two experiments using a larger MC sample of 1024 runs. These extended simulations improved Si and STi values for AI and LPDI. The PD statistics, however, maintained low accuracy (Table 3). To confirm that sensitivity estimation errors can noticeably affect the quality of *time*-GSA results, we compared the time series of normalized Si values for PD calculated for experiment 7 (Figure 4). Observe that both plots differ in the relative importance of factors over time. Therefore we decided to exclude PD from further analysis.

*[Insert Table 3 about here]*

*[Insert Figure 4 about here]*

### *5.3 Uncertainty plots of land use fragmentation statistics*

To establish output PDFs, we summarized AI and LPDI using ($\mu$, 1$\sigma$) pairs for every time step of model execution, rendered as time series of output uncertainty (Figure 5).

*[Insert Figure 5 about here]*

The AI statistic demonstrates a fairly consistent behavior with a general trend of a short-time change in value (increase or drop, depending on experiment), followed by a long-time period of relative stability with $\mu > 40\%$. Not surprisingly, the highest AI correlates with high importance assigned to land value either directly ($W_{lv}$, experiment 2) or indirectly (high $I_{lv}$ and low $D_{sb}$, experiment 7). In both cases the average AI scores are high due to the concentric spatial distribution of this layer (Figure 2) which has a tendency to build one big cluster. Observe that high $I_{lv}$ and low $D_{sb}$ (experiment 7) result in the lowest variance of AI (smallest $\sigma$) suggesting that this configuration of feedbacks ensures a high level of aggregation every time.

The behavior of LPDI differs from AI. First of all, it is much more variable with its 1$\sigma$-band of ~30% as opposed to AI where 1$\sigma$-band = 20% or less. Secondly, the time of convergence to equilibrium for LPDI is longer than for AI. For example, experiments 1 and 2 exhibit the convergence period of ~30 time steps for LPDI and ~10 for AI. The *Natural Beauty Weight* experiment (3) is the least variable. Its 1$\sigma$-band is the narrowest owing to the overwhelming impact of the polycentric nature of the natural beauty layer (Figure 2).

Moreover, all but one experiments exhibit an interesting behavior related to the relative size of the largest developed patch. After 3-5 time steps the model reaches the maximum LPDI value, then drops and maintains an almost steady value throughout the rest of the simulation. For instance, in experiment 11 LPDI starts around ~50%, then quickly goes up to ~60% and, after $t$=4, falls and stabilizes around 40%.

Finally, we should point to experiments 2, 3, and 7 that exhibit distinct LPDI time series. Both the *Land Value Weight* experiment and the *Natural Beauty Weight* experiment are directly dependent on the preference assigned to their respective attribute maps, with high LPDI for the concentric land value layer and low LPDI for the dispersed and polycentric clumps of high scenic beauty. Similarly to AI, experiment 7 demonstrates higher LPDI values due to the dominating impact of $I_{lv}$ relative to $D_{sb}$, which strengthens the influence of the land value layer on the resulting development.

### *5.4 Significance of different input uncertainties on landscape fragmentation*

To test the significance of different conceptions of input factor PDFs, we performed one-way analysis of variance for AI and LPDI among all experiments for selected time steps (one, three, five, ten and every $10^{th}$ afterwards). Table 4 shows selected ANOVA results for $t$=3, $t$=20, and $t$=70.

*[Insert Table 4 about here]*

For LPDI, the p-values are less than .0001 in both cases, indicating significant differences among the experiments. Moreover, based on the time series of differences (Figure 6), the largest difference at each tick decreases rapidly from over 25% at $t$=1 to ~17% at $t$=3, followed by an increase to 30% at step 20 and then a slight decline. After step 5 however, its variation is small. We can therefore conclude that after the initial short period of time, the largest difference of AI for all eleven experiments becomes stable.

*[Insert Figure 6 about here]*

Similarly to AI, LPDI p-values score less than .0001 in ANOVA, suggesting a significant impact of various PDF definitions of input factors on the relative size of the largest patch of developed area (Table 4). More importantly, based on the time series in Figure 6, the largest difference at each time step has an increasing trend with the peak at the end of the simulation. In summary, Figure 6 confirms the distinctive character of time step 30 as a point of gradual regime change, after which most of the experiments converge to equilibrium (Figure 5).

### *5.5 Time dependent global sensitivity analysis*

Figures 7 and 8 illustrate the *time*-GSA plots with sensitivity indices rendered cumulatively. Below we present the major observations of the dynamics of variance decomposition (refer to Table 1 for guidance).

*[Insert Figure 7 about here]*

*[Insert Figure 8 about here]*

With high Si we look for the important input factors that, if fixed independently, would substantially reduce the variance of AI and LPDI. The plots suggest a considerable variability in Si among scenarios, which is quite surprising given the relatively stable uncertainty plots in Figure 5. In the vast majority of cases preference for scenic beauty ($W_{sb}$) is the most influential input factor for both fragmentation statistics, followed by $W_{lv}$ and $D_{sb}$. The significance of $W_{sb}$ on AI and LPDI is not surprising given the fact that the natural beauty layer drives the magnitude of patchiness in the area. In experiments 7, 2, 3, and 5, $S_S$ becomes a moderately influential factor, especially at the beginning of the simulation. These four factors constitute the leverage points of clustered landscape configuration. The role of the other factors in fragmentation statistics variability is almost negligible, especially given the small values of STi.

Similarly to the uncertainty plots in Figure 5, scenarios 2, 3, and 7 diverge the most from the base case for both Si and STi of AI. Consider the most influential factors for experiment 7 and 3. Clearly, in the former experiment, $S_S$ is the factor that, taken singly, causes the majority of AI variance, as opposed to $D_{sb}$ in experiment 3. By investigating STi, we can observe that $D_{sb}$ and $S_S$ are still the most sensitive factors for these experiments, suggesting strong non-linear relationships between these variables (Table 1).

Using the sum of first-order sensitivity indices, we can assess to what extent the ABM behaves in an additive manner. The portion of output variance that cannot be explained by individual factors is drawn in white in Figures 7a and 8a. The width of this band reflects the fraction of interactions that affect output variance. Interestingly, factors interactivity differs between AI and LPDI, indicating that we should be cautious in selecting spatial metrics, since one fragmentation statistic is not enough to thoroughly describe pattern dynamics.

Given the PDFs used in the experiments, the ABM is more additive than expected and the interactions among its parameters are not substantial. For AI, the sum of first-order effects exceeds 70% for most of the scenarios, occasionally dropping to 60%. The major trend in the intensity of interactions among the inputs is that the model starts as an additive one and becomes more nonlinear later. Based on the total-effect indices, we can also observe that factor interactivity introduces some level of stability in the relative significance among inputs (STi of AI in experiments 1, 4, and 6). Moreover, the importance of factor interactivity becomes more evident for scenarios that are impacted by the availability of highly scoring opportunities located in the neighborhood of already developed sites (scenario 1, 4, and 11).

Comparing the uncertainly plots of AI with its sensitivity graphs leads to an interesting correlation. Whenever the value of AI stabilizes, the input factors become more interdependent. This behavior is representative of a dynamic equilibrium at the macro-level (Richiardi, Leombruni et al. 2006), in which disparate factor combinations contribute towards ABM outcome stability during different time periods.

When considering the LPDI statistic, the sum of its first order indices changes rather haphazardly. Observe that, apart from experiments 2, 3, and 7, the variance of LPDI seems to be shaped by the individual factors rather than their combinations, since the sum of first-order indices maintains 100% value most of the time. The interaction among inputs ($W_{lv}$, $W_{sb}$, and $D_{sb}$) is more evident in experiments with divergent importance assigned to the attribute maps (scenarios 2, 3, 7).

Furthermore, observe that the factor which defines a particular scenario is not necessarily the most significant one in shaping the behavior uncertainty of this scenario. For example, experiment 7 has a high value assigned to $I_{lv}$, and yet the sensitivity for AI of the tested uncertainty in this factor is negligible. In fact, the most significant factor here is $S_S$, suggesting that the high value of AI in this scenario (Figure 5) is mostly dictated by agents' access to information about development opportunities. This behavior can be easily explained. If the average value of a particular input is high, such factor has in consequence an overwhelming impact on model outcome dynamics to the point that the ABM becomes insensitive to the variability associated with this factor. Therefore, the influence of this factor on model outcomes is substantial but steady. This has a profound consequence on the input definition of a given experiment. Strictly speaking, a factor that controls a particular scenario can be essentially set to a relatively high fixed value yielding a similar model behavior as the one outlined above.

Given the *time*-GSA plots of most AIs and some LPDIs, we can observe that the model behaves more nonlinearly after $t$=30. We hypothesize that the small interaction among factors in the first time steps can be explained by spatial path dependence. The initial decisions of agents are based on randomly drawn samples of developable locations. Hence, the very first developed sites produce the seeds of growth that later glues to the existing patches. Other significant input variables are less influential. Not surprisingly then, the first few time steps are dominated by variations in $S_S$. The converse also supports the existence of spatial path dependence. With high $S_S$ in experiment 10, agents become more consistent in their choices of initial developments, reducing the role of $S_S$ variability in outcome sensitivity.

Is there anything that can be learned from the homogeneous population? Based on the negligible impact of the magnitude of $A_{num}$, we postulate that ABMs with multiple homogeneous agents can be substituted with the system dynamics approach employing one 'averaged' actor. Moreover, significant changes in population size (scenarios 8 and 9) increase factor interactivity. We hypothesize that this trend would be more pronounced with heterogeneous agents.

## 6 Concluding observations

The aim of this study was to investigate the potential of time dependent variance-based global sensitivity analysis in examining the dynamics of outcome uncertainty of a simple agent-based model of land use change. To achieve this objective we experimented with different configurations of uncertainties of a fixed number of input variables. We summarized output uncertainties with descriptive statistics and compared them with time series of first-order and total-effect sensitivity indices.

We have found that *time*-GSA allows for exploration of model behavior which would be hard to observe using other approaches. The sole analysis of outcome uncertainty, exemplified in probability density functions of selected fragmentation statistics, does not reveal the dynamics of model sensitivity to various input factors. We would be unable to estimate how uncertainties in different factors contribute individually and in interactions towards the stability of ABM outcome fragmentation over time. Based on the presented research, we argue that it is difficult to assess the nonlinearity of ABM without a thorough sensitivity analysis. ABM has been proclaimed to be complex and nonlinear, but little is reported to prove that the internal mechanisms of such models behave truly nonlinearly. As Phillips (2003) and Manson (2007) point out, not all complexity is nonlinear, and not all nonlinearity is complex. Going one step further, we propose different types of ABM nonlinearity: from functional (model definition), through factorial (input and output distribution), to behavioral (the nature and magnitude of interactions during model execution). As demonstrated here, *time*-GSA is particularly suitable for the latter case.

We conclude that the factor sensitivity measures vary significantly during model execution. Assuming that ABM dynamics is a result of factor interaction, we observe a distinct time lag of nonlinearity, which unfolds after the formation of the seeds of development. Consequently, a static snapshot of ABM sensitivity, taken at the end of the simulation, is inadequate when deciding on factor prioritization and reduction.

What should be done in the future? Considering that land use ABM is particularly suitable for studying the impact of human behavior on landscape configurations, a concurrent research project is undertaken to examine the effects of agent heterogeneity on land use dynamics. Finally, there is a clear need for better approximation methods of sensitivity indices.

As shown in section 5.2, the Monte Carlo technique applied here requires a large number of model executions. For empirical ABMs, which require longer computation time, large samples of simulations become prohibitive.

**Acknowledgements** The authors would like to acknowledge the constructive feedback provided by three anonymous reviewers on the previous version of the 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**,** 54-79.

BARRETT, G. V. & BLAIR, J. P. 1988. *How to Conduct and Analyze Real Estate Market and Feasibility Studies,* New York, Van Nostrand Reinhold.

BECU, N., PEREZ, P., WALKER, A., BARRETEAU, O. & LE PAGE, C. 2003. Agent based simulation of a small catchment water management in northern Thailand description of the CATCHSCAPE model. *Ecological Modelling,* 170**,** 319-331.

BENENSON, I. & TORRENS , P. 2004. *Geosimulation: Automata-based modeling of urban phenomena,* Chichester, England, Wiley.

BENNETT, D. & MCGINNIS, D. 2008. Coupled and complex: Human-environment interaction in the Greater Yellowstone Ecosystem, USA. *Geoforum,* 39**,** 833-845.

BOUSQUET, F. & LE PAGE, C. 2004. Multi-agent simulations and ecosystem management: a review. *Ecological Modelling,* 176**,** 313-332.

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.

BROWN, D. G. & ROBINSON, D. T. 2006. Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. *Ecology and Society,* 11.

BROWN, D. G., ROBINSON, D. T., AN, L., NASSAUER, J. I., ZELLNER, M., RAND, W., RIOLO, R., PAGE, S. E., LOW, B. & WANG, Z. F. 2008. Exurbia from the bottom-up: Confronting empirical challenges to characterizing a complex system. *Geoforum,* 39**,** 805-818.

BRUGNACH, M. 2005. Process level sensitivity analysis for complex ecological models. *Ecological Modelling,* 187**,** 99-120.

BRUGNACH, M., DEWULF, A., PAHL-WOSTL, C. & TAILLIEU, T. 2008. Toward a Relational Concept of Uncertainty: about Knowing Too Little, Knowing Too Differently, and Accepting Not to Know. *Ecology and Society,* 13.

BURKE, D. S., EPSTEIN, J. M., CUMMINGS, D. A. T., PARKER, J. I., CLINE, K. C., SINGA, R. M. & CHAKRAVARTY, S. 2006. Individual-based computational modeling of smallpox epidemic control strategies. *Academic Emergency Medicine,* 13**,** 1142-1149.

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.

CASTELLA, J. C., TRUNG, T. N. & BOISSAU, S. 2005. Participatory simulation of land-use changes in the northern mountains of Vietnam: the combined use of an agent-based model, a role-playing game, and a geographic information system. *Ecology and Society,* 10.

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.

CROSETTO, M., TARANTOLA, S. & SALTELLI, A. 2000. Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture Ecosystems & Environment,* 81**,** 71-79.

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. & BURMASTER, D. E. 1996. Guest editorial: Treatment of aleatory and epistemic uncertainty in performance assessments for complex systems. *Reliability Engineering & System Safety,* 54**,** 91-94.

HOMMA, T. & SALTELLI, A. 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety,* 52**,** 1-17.

HWANG, C. L. & YOON, K. 1981. *Multiple attribute decision making methods and applications: A state of the art survey,* Berlin, Springer-Verlag.

IRWIN, E., JAYAPRAKASH, C. & MUNROE, D. 2009. Towards a comprehensive framework for modeling urban spatial dynamics. *Landscape Ecology,* 24**,** 1223-1236.

LI, X. & LIU, X. 2007. Defining agents' behaviors to simulate complex residential development using multicriteria evaluation. *Journal of Environmental Management,* 85**,** 1063-1075.

LIGMANN-ZIELINSKA, A. 2009. The impact of risk-taking attitudes on a land use pattern: an agent-based model of residential development. *Journal of Land Use Science,* 4**,** 215 - 232.

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

LIU, J. G., DIETZ, T., CARPENTER, S. R., ALBERTI, M., FOLKE, C., MORAN, E., PELL, A. N., DEADMAN, P., KRATZ, T., LUBCHENCO, J., OSTROM, E., OUYANG, Z., PROVENCHER, W., REDMAN, C. L., SCHNEIDER, S. H. & TAYLOR, W. W. 2007. Complexity of coupled human and natural systems. *Science,* 317**,** 1513-1516.

MALCZEWSKI, J. 1999. *GIS and MULTICRITERIA DECISION ANALYSIS,* New York, John Wiley & Sons, Inc.

MANSON, S. M. 2001. Simplifying complexity: a review of complexity theory. *Geoforum,* 32**,** 405-414.

MANSON, S. M. 2007. Challenges in evaluating models of geographic complexity. *Environment and Planning B-Planning & Design,* 34**,** 245-260.

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**,** 1447-1459.

MCGARIGAL, K. & MARKS, B., J, 1995. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. *General Technical Report.* Portland, OR: USDA Forest Service, Pacific Northwest Research Station.

MOSLER, H.-J. & MARTENS, T. 2008. Designing environmental campaigns by using agent-based simulations: Strategies for changing environmental attitudes. *Journal of Environmental Management,* 88**,** 805-816.

ORESKES, N., SHRADERFRECHETTE, K. & BELITZ, K. 1994. VERIFICATION, VALIDATION, AND CONFIRMATION OF NUMERICAL-MODELS IN THE EARTH-SCIENCES. *Science,* 263**,** 641-646.

PARKER, D. C., ENTWISLE, B., RINDFUSS, R. R., VANWEY, L. K., MANSON, S. M., MORAN, E., AN, L., DEADMAN, P., EVANS, T. P., LINDERMAN, M., MUSSAVI RIZI, S. M. & MALANSON, G. 2008. Case studies, cross-site comparisons, and the challenge of generalization:

comparing agent-based models of land-use change in frontier regions. *Journal of Land Use Science,* 3**,** 41 - 72.
PARKER, D. C., MANSON, S. M., JANSSEN, M. A., HOFFMANN, M. J. & DEADMAN, P. 2003. Multi-agent systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers,* 93**,** 314-337.
PHILLIPS, J., D. 2003. Sources of nonlinearity and complexity in geomorphic systems. *Progress in Physical Geography,* 27**,** 1.
RICHIARDI, M., LEOMBRUNI, R., SAAM, N. J. & SONNESSA, M. 2006. A Common Protocol for Agent-Based Social Simulation. *Journal of Artificial Societies and Social Simulation,* 9**,** 15.
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. 2002. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications,* 145**,** 280-297.
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. & CHAN, K. 1999. A role for sensitivity analysis in presenting the results from MCDA studies to decision makers. *Journal of Multi-Criteria Decision Analysis,* 8**,** 139-145.
SCHLUTER, M. & PAHL-WOSTL, C. 2007. Mechanisms of resilience in common-pool resource management systems: an agent-based model of water use in a river basin. *Ecology and Society,* 12.
SOBOL, I. M. Year. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *In,* 2001. Elsevier Science Bv, 271-280.
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.
TOPPING, C. J., HØYE, T. T. & OLESEN, C. R. 2010. Opening the black box--Development, testing and documentation of a mechanistically rich agent-based model. *Ecological Modelling,* 221**,** 245-255.
TVERSKY, A. & KAHNEMAN, D. 1981. THE FRAMING OF DECISIONS AND THE PSYCHOLOGY OF CHOICE. *Science,* 211**,** 453-458.
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.
ZIEHN, T. & TOMLIN, A. S. 2009. GUI-HDMR - A software tool for global sensitivity analysis of complex models. *Environmental Modelling & Software,* 24**,** 775-785.

**Table 1** Si and STi interpretation

| **Measure** | **Interpretation** |
|---|---|
| Relatively high *Si* | A factor that is singly influential on the variability of model output |
| Sum of *Si* over all inputs | Percent of output variability due to the inputs taken independently; the remainder (to 100%) is the fraction of output variance due to the interactions among inputs |
| *STi* – *Si* relatively high | Input *i* is highly involved in interactions with other factors, all inputs with high difference are involved in interactions among each other<br>Note that input *i* can be singly insignificant (low *Si*), but influential when involved in interactions with other factors (high *STi*), *i* would therefore influence the output variance more through interactions than individually |
| Relatively low value of *STi* | An insignificant factor |

**Table 2** Probability density functions used for seven input factors in eleven ABM experiments: N – normal distribution probability density function (*mean, std*); M – probability mass (discrete) function defined with three values: *value_low* having p=0.25; *mean* having p=0.5; *value_high* having p=0.25

| **Id** | Experiment | agent num | value weight | beauty weight | access weight | sample size | decrease beauty (feedback) | increase value (feedback) |
|---|---|---|---|---|---|---|---|---|
| **1** | Base | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | N(.33, .165) | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **2** | Land value preference dominates | M(5, 8, 10) | **N(.66, .165)** | N(.33, .165) | N(.33, .165) | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **3** | Natural beauty preference dominates | M(5, 8, 10) | N(.33, .165) | **N(.66, .165)** | N(.33, .165) | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **4** | Low accessibility | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | **N(.165, .0825)** | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **5** | High accessibility | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | **N(.66, .165)** | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **6** | Feedbacks high-high | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | N(.33, .165) | N(.165, .0825) | **N(.08, .02)** | **N(.08, .02)** |
| **7** | Feedbacks low-high | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | N(.33, .165) | N(.165, .0825) | **N(.005, .002)** | **N(.08, .02)** |
| **8** | Agent population large | **M(40, 20, 10)** | N(.33, .165) | N(.33, .165) | N(.33, .165) | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **9** | Agent population small | **M(5, 4, 2)** | N(.33, .165) | N(.33, .165) | N(.33, .165) | N(.165, .0825) | N(.05, .02) | N(.05, .02) |
| **10** | Sample size large | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | N(.33, .165) | **N(.66, .165)** | N(.05, .02) | N(.05, .02) |
| **11** | Sample size small | M(5, 8, 10) | N(.33, .165) | N(.33, .165) | N(.33, .165) | **N(.0825, .04125)** | N(.05, .02) | N(.05, .02) |

**Table 3** The maximum errors in sensitivity index estimates (defined as % deviation below zero) for the Monte Carlo sample size of 512 runs. The highlighted experiments (2 and 7) have their errors calculated for 1024 runs; their previous errors (for 512 runs) are provided in parentheses.

| | AI | | LPDI | | PD | |
|---|---|---|---|---|---|---|
| Experiment Id | Si | STi | Si | STi | Si | STi |
| 1 | 2.8 | 0.0 | 6.1 | 1.4 | 3.4 | 2.7 |
| **2** | **7.6 (15.2)** | **1.3 (1.6)** | **14.5 (28.6)** | **0.7 (3.1)** | **3.9 (15.3)** | **18.3 (34.0)** |
| 3 | 7.1 | 1.2 | 5.8 | 4.1 | 5.1 | 4.4 |
| 4 | 1.7 | 0.0 | 4.3 | 2.3 | 5.4 | 0.5 |
| 5 | 4.4 | 0.5 | 5.0 | 0.1 | 3.8 | 4.7 |
| 6 | 4.4 | 1.1 | 5.2 | 2.4 | 3.4 | 1.4 |
| **7** | **5.7 (6.5)** | **5.9 (12.6)** | **3.0 (7.4)** | **1.6 (2.3)** | **14.6 (16.6)** | **39.5 (83.0)** |
| 8 | 3.2 | 1.4 | 6.1 | 1.0 | 3.9 | 7.1 |
| 9 | 9.6 | 0.5 | 9.1 | 1.9 | 7.6 | 16.8 |
| 10 | 1.8 | 0.7 | 6.6 | 2.4 | 5.9 | 6.0 |
| 11 | 8.3 | 0.8 | 11.8 | 2.1 | 5.9 | 3.2 |

**Table 4** ANOVA results for AI and LPDI for selected time steps

| **AI: time step 3** | | | | | |
|---|---|---|---|---|---|
| | Df | SS | MSE | F | p-value |
| Experiments | 10 | 614161 | 61416 | 194.07 | <0.0001 |
| Residuals | 5621 | 1778802 | 316 | | |
| **AI: time step 20** | | | | | |
| | Df | SS | MSE | F | p-value |
| Experiments | 10 | 966141 | 96614 | 239.12 | <0.0001 |
| Residuals | 5621 | 2271143 | 404 | | |
| **LPDI: time step 3** | | | | | |
| | Df | SS | MSE | F | p-value |
| Experiments | 10 | 504086 | 50409 | 68.41 | <0.0001 |
| Residuals | 5621 | 4141926 | 737 | | |
| **LPDI: time step 70** | | | | | |
| | Df | SS | MSE | F | p-value |
| Experiments | 10 | 1486257 | 148626 | 128.53 | <0.0001 |
| Residuals | 5621 | 6499626 | 1156 | | |

## Figures

Figure 1 Conceptual diagram of the agent-based model of residential development.

Figure 2 Input landscape attribute maps. Darker shades represent higher values.

Figure 3 Mean land development for t=10 and t=70 calculated among all model executions for selected scenarios, darker colour indicates higher frequency of development

Figure 4 Time series of the first order sensitivity index of PD fragmentation statistics calculated for Experiment 7 (*Feedbacks: Low Beauty, High Value*).

Figure 5 Time dependent uncertainty plots of [a] Aggregation Index, and [b] Largest Patch of Developed area Index.

Figure 6 Largest differences (%) among experiments plotted over time for AI and LPDI stats.

Figure 7 Time-GSA plots for Aggregation Index: [a] Si, [b] STi.

Figure 8 Time-GSA plots for Largest Patch of Developed area Index: [a] Si, [b] STi.

---

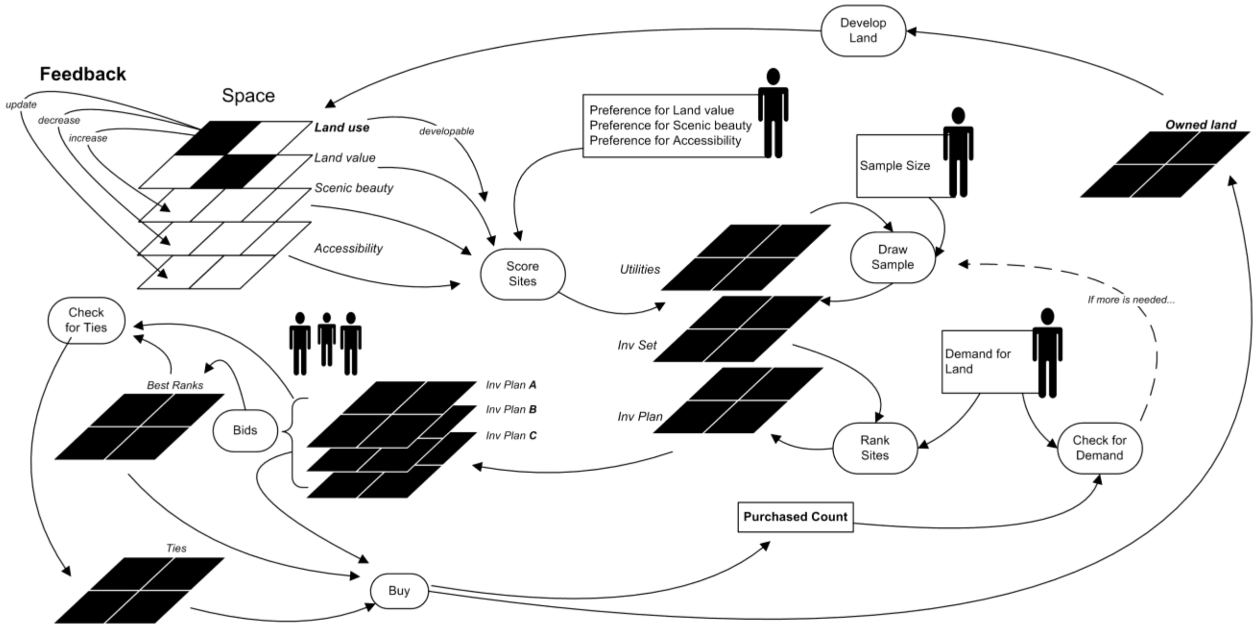



Figure 1

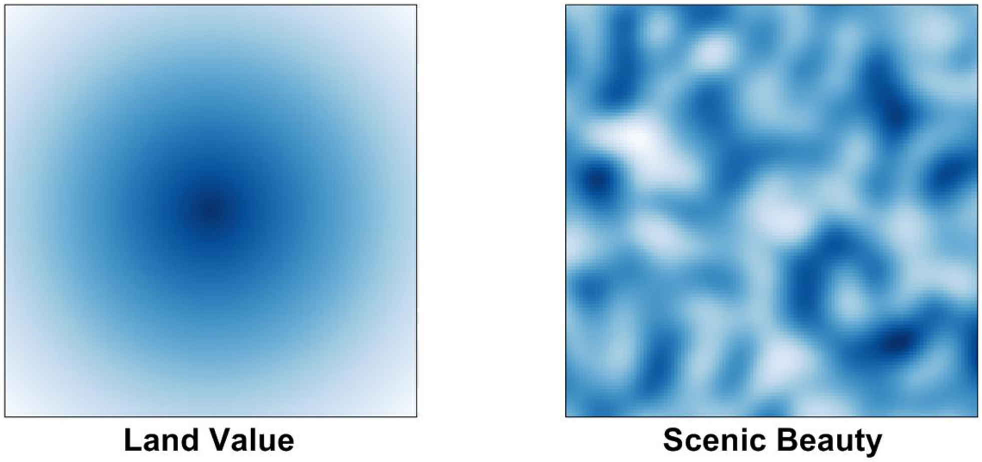



Figure 2

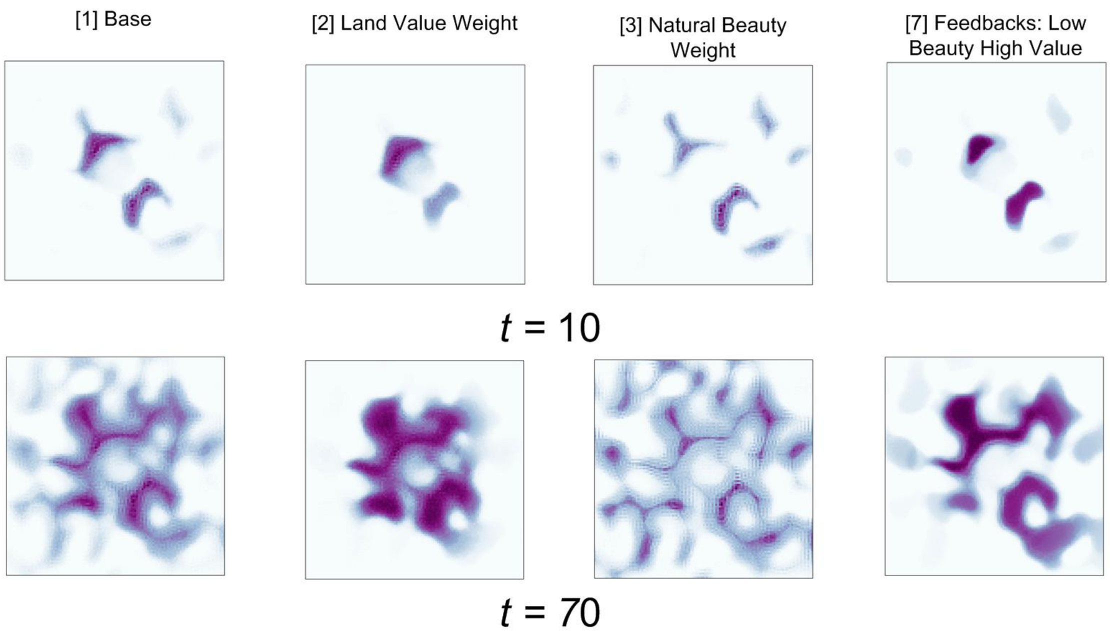
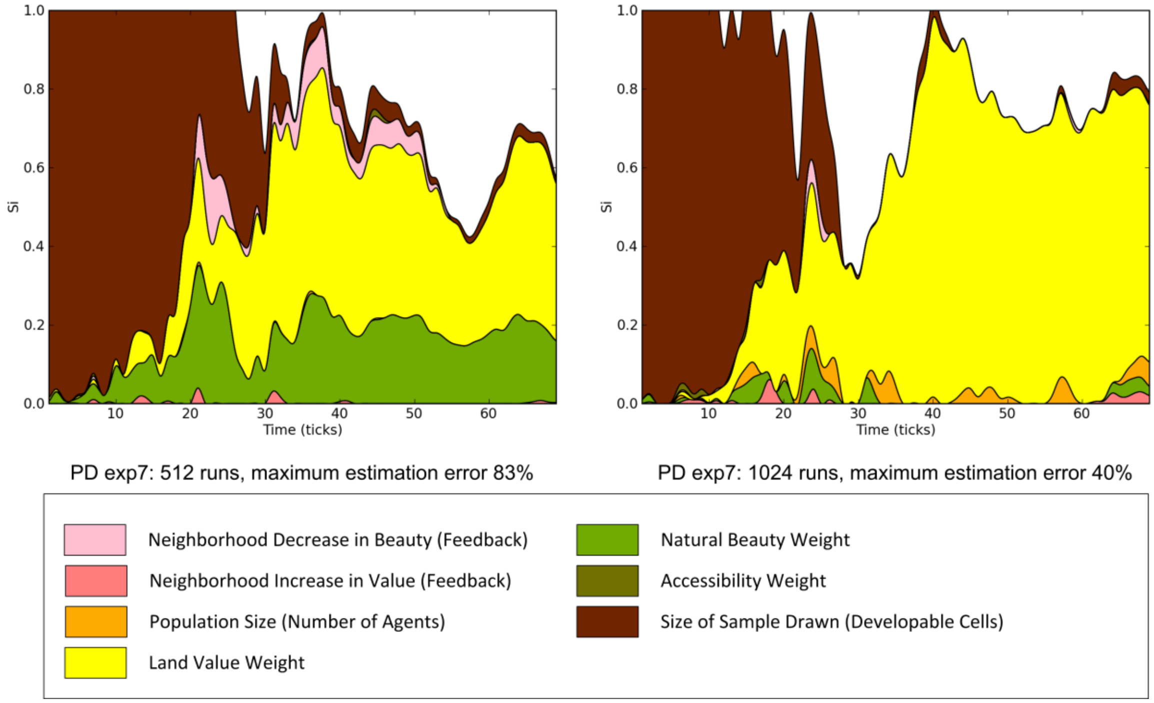
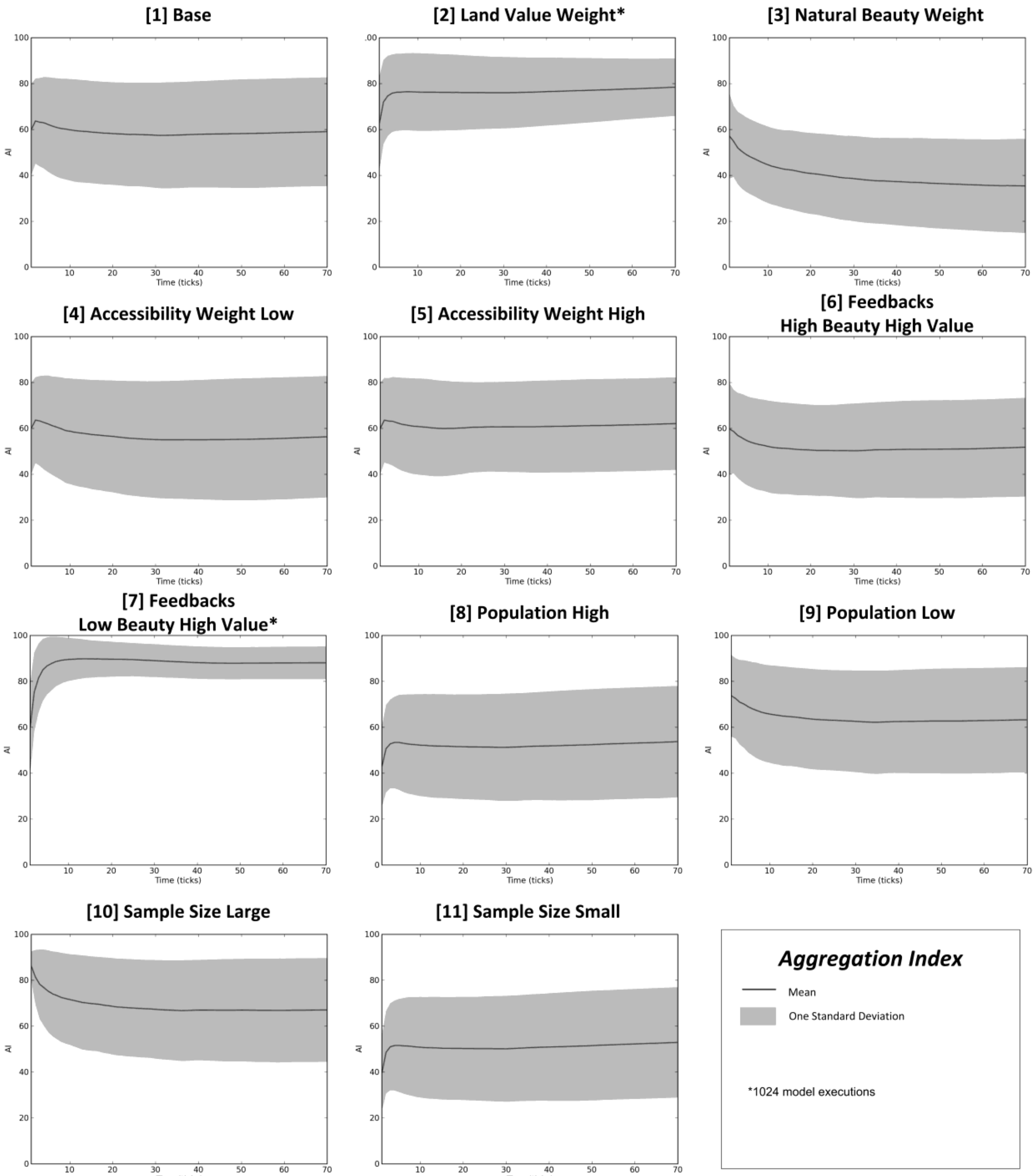



Figure 3



Si
Time (ticks)
PD exp7: 512 runs, maximum estimation error 83%
PD exp7: 1024 runs, maximum estimation error 40%
Neighborhood Decrease in Beauty (Feedback)
Neighborhood Increase in Value (Feedback)
Population Size (Number of Agents)
Land Value Weight
Natural Beauty Weight
Accessibility Weight
Size of Sample Drawn (Developable Cells)


Figure 4



[1] Base
[2] Land Value Weight*
[3] Natural Beauty Weight
[4] Accessibility Weight Low
[5] Accessibility Weight High
[6] Feedbacks
High Beauty High Value
[7] Feedbacks
Low Beauty High Value*
[8] Population High
[9] Population Low
[10] Sample Size Large
[11] Sample Size Small
AI
Time (ticks)
Aggregation Index
Mean
One Standard Deviation
*1024 model executions


Figure 5 [a]

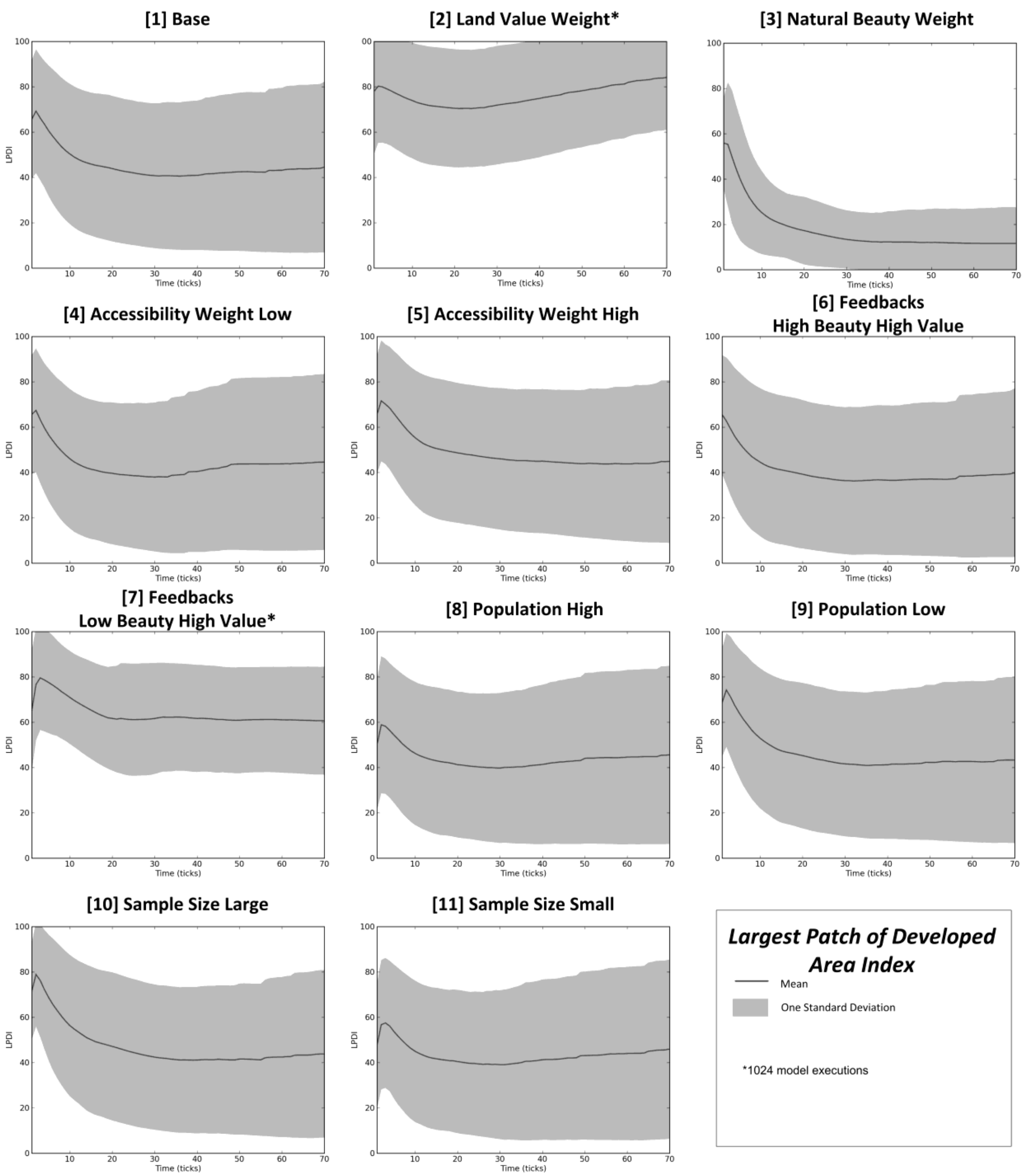

[1] Base
[2] Land Value Weight*
[3] Natural Beauty Weight
[4] Accessibility Weight Low
[5] Accessibility Weight High
[6] Feedbacks
High Beauty High Value
[7] Feedbacks
Low Beauty High Value*
[8] Population High
[9] Population Low
[10] Sample Size Large
[11] Sample Size Small
LPDI
Time (ticks)
Largest Patch of Developed Area Index
Mean
One Standard Deviation
*1024 model executions

Figure 5 [b]

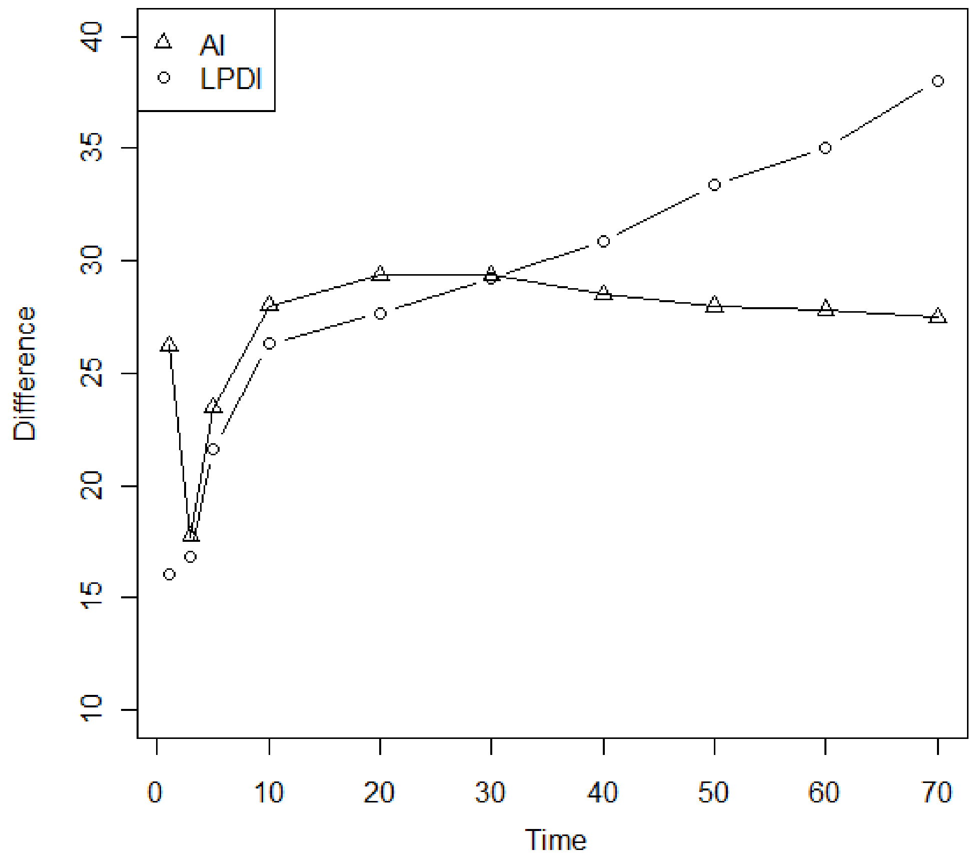



Figure 6

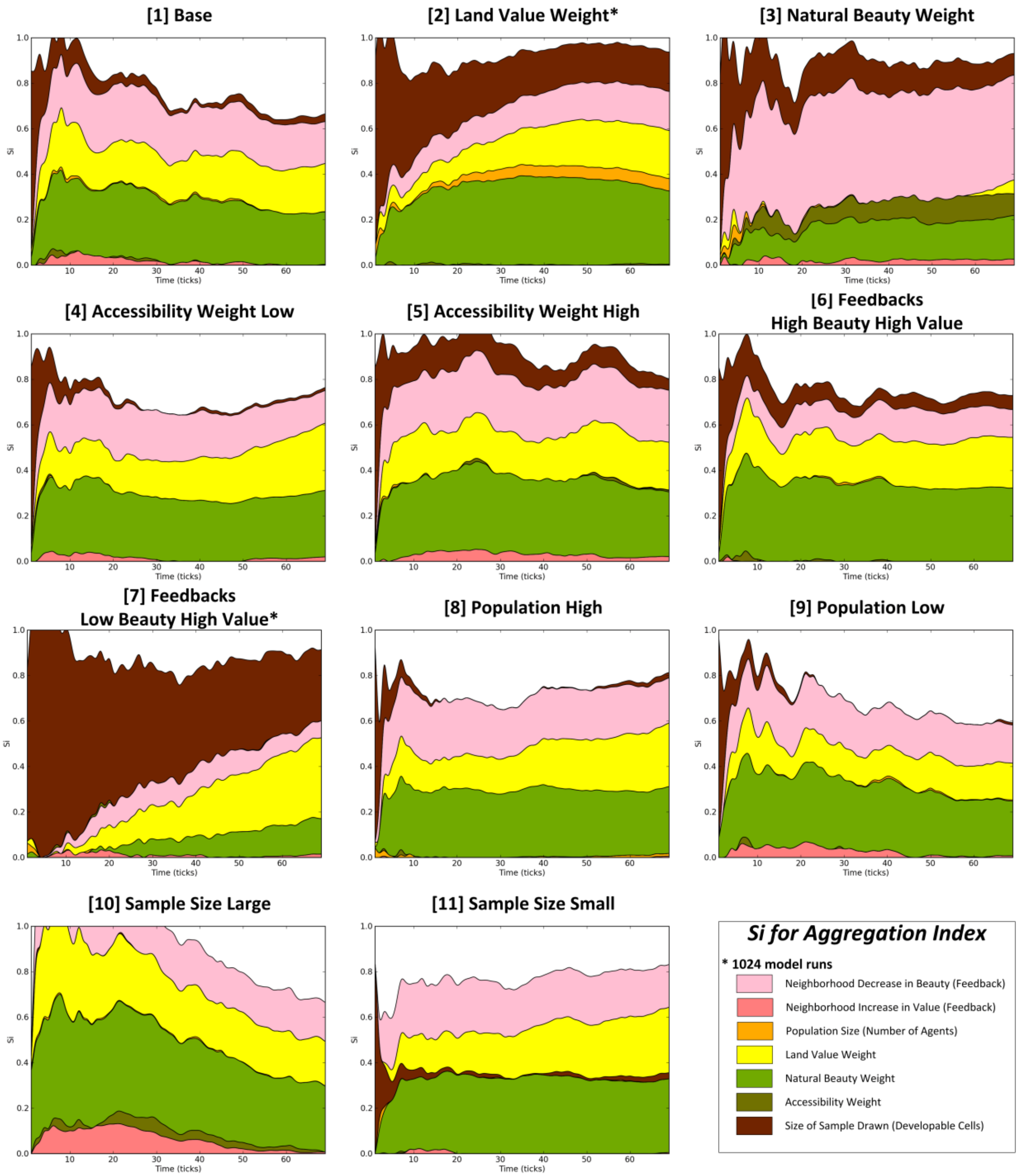

[1] Base
[2] Land Value Weight*
[3] Natural Beauty Weight
[4] Accessibility Weight Low
[5] Accessibility Weight High
[6] Feedbacks
High Beauty High Value
[7] Feedbacks
Low Beauty High Value*
[8] Population High
[9] Population Low
[10] Sample Size Large
[11] Sample Size Small
Time (ticks)
Si
Si for Aggregation Index
* 1024 model runs
Neighborhood Decrease in Beauty (Feedback)
Neighborhood Increase in Value (Feedback)
Population Size (Number of Agents)
Land Value Weight
Natural Beauty Weight
Accessibility Weight
Size of Sample Drawn (Developable Cells)

Figure 7 [a]

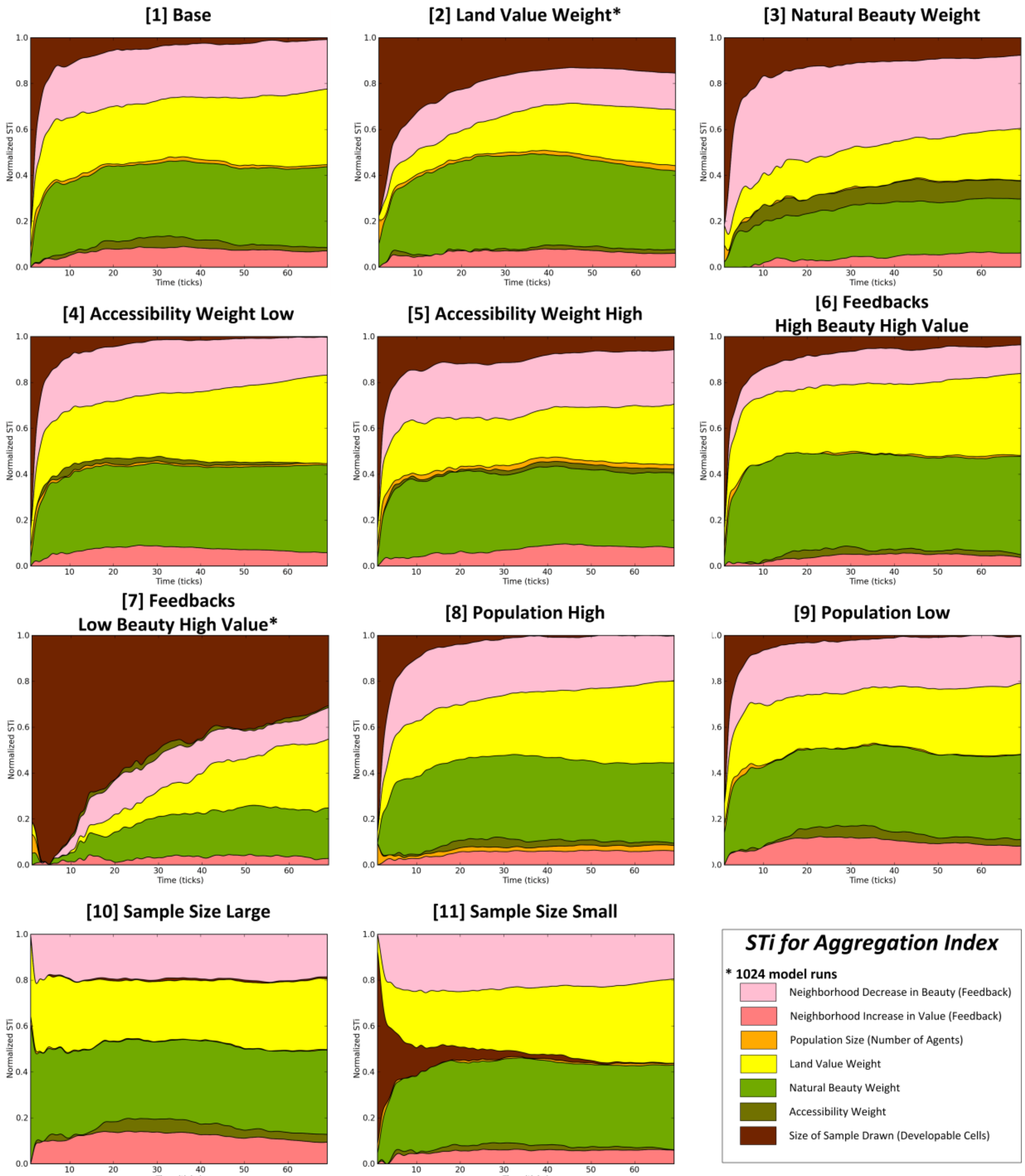

[1] Base
[2] Land Value Weight*
[3] Natural Beauty Weight
[4] Accessibility Weight Low
[5] Accessibility Weight High
[6] Feedbacks
High Beauty High Value
[7] Feedbacks
Low Beauty High Value*
[8] Population High
[9] Population Low
[10] Sample Size Large
[11] Sample Size Small
Normalized STi
Time (ticks)
STi for Aggregation Index
* 1024 model runs
Neighborhood Decrease in Beauty (Feedback)
Neighborhood Increase in Value (Feedback)
Population Size (Number of Agents)
Land Value Weight
Natural Beauty Weight
Accessibility Weight
Size of Sample Drawn (Developable Cells)

Figure 7 [b]

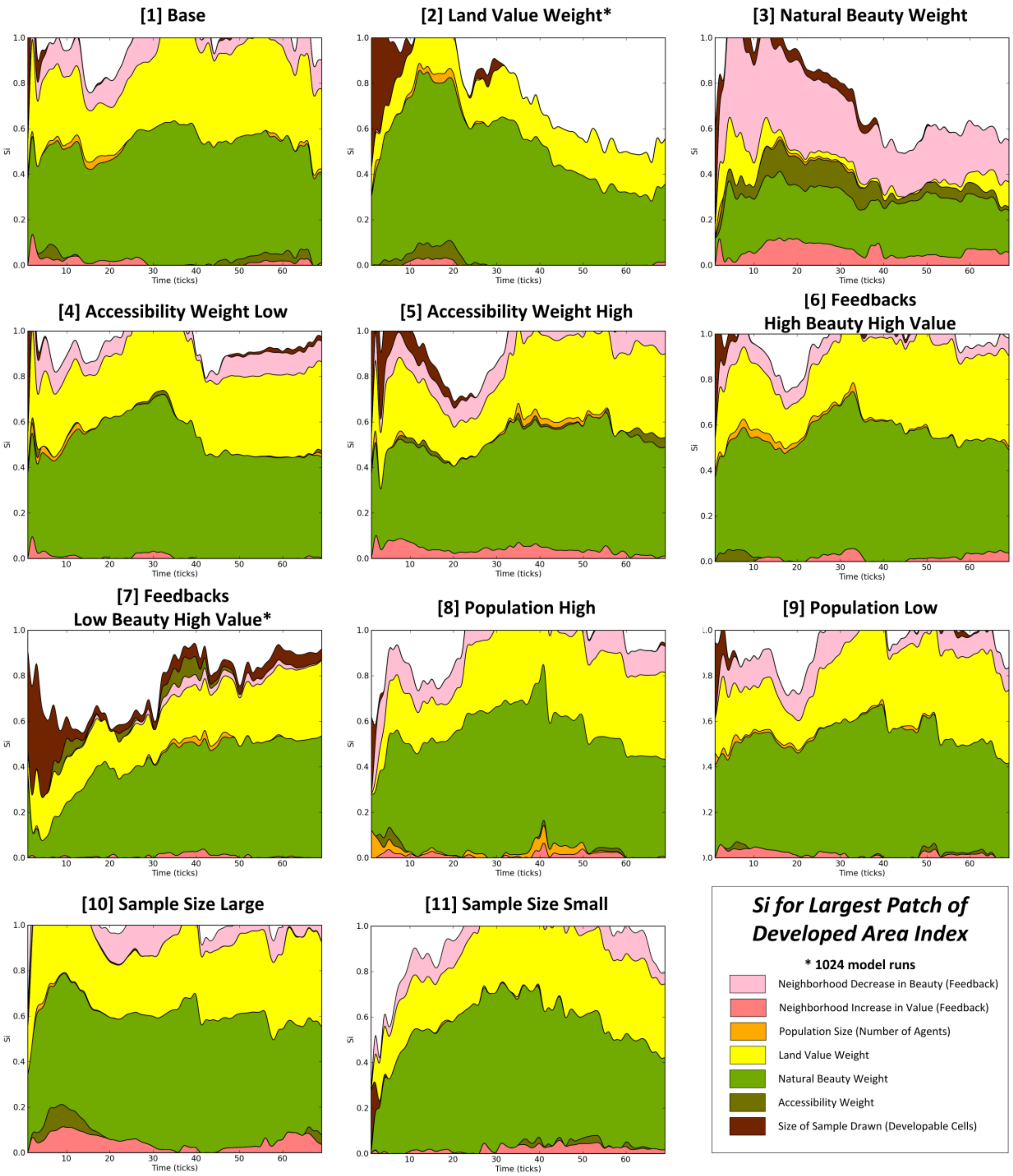

[1] Base
[2] Land Value Weight*
[3] Natural Beauty Weight
[4] Accessibility Weight Low
[5] Accessibility Weight High
[6] Feedbacks
High Beauty High Value
[7] Feedbacks
Low Beauty High Value*
[8] Population High
[9] Population Low
[10] Sample Size Large
[11] Sample Size Small
Si
Time (ticks)
Si for Largest Patch of Developed Area Index
* 1024 model runs
Neighborhood Decrease in Beauty (Feedback)
Neighborhood Increase in Value (Feedback)
Population Size (Number of Agents)
Land Value Weight
Natural Beauty Weight
Accessibility Weight
Size of Sample Drawn (Developable Cells)

Figure 8 [a]

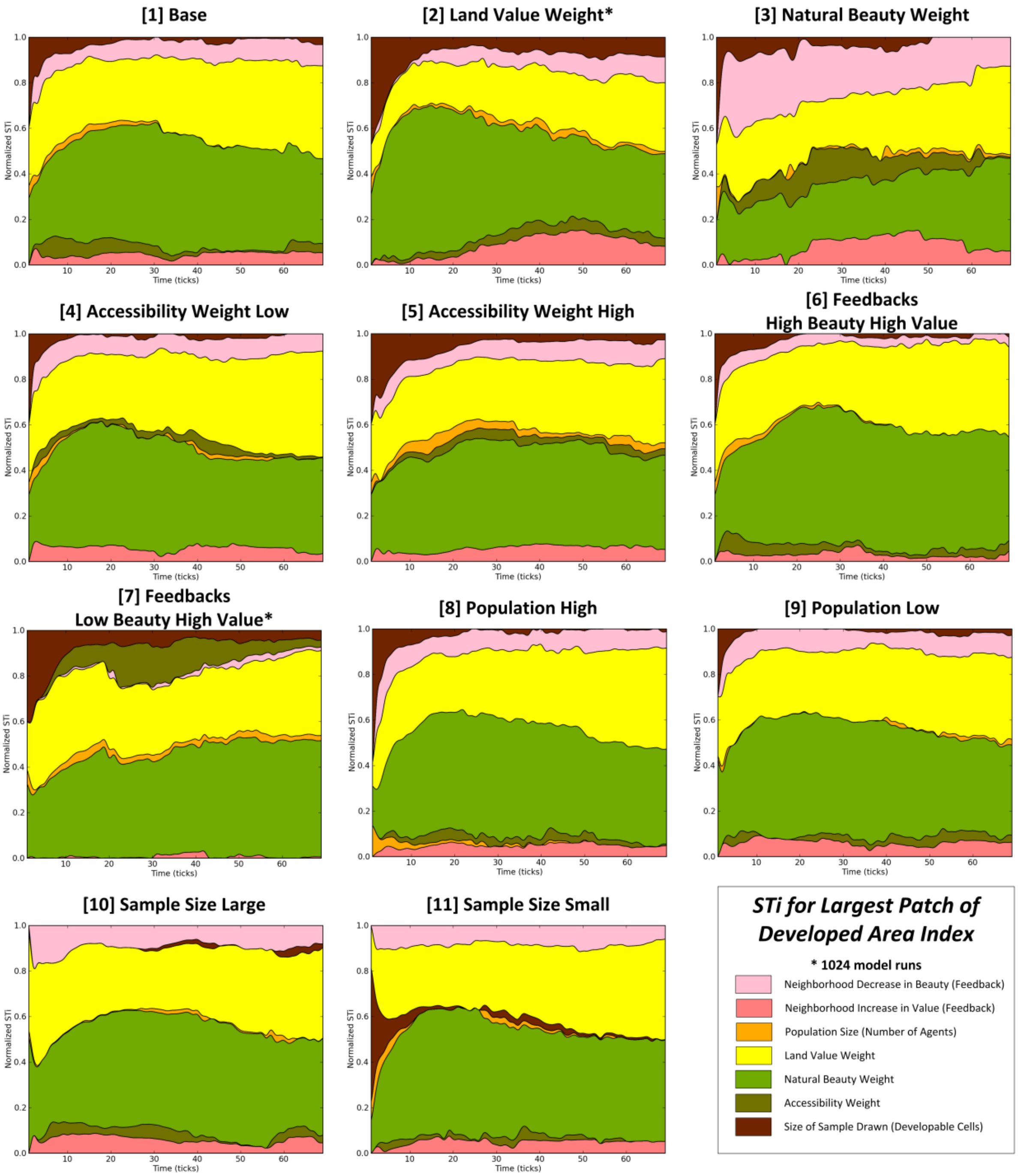

[1] Base
[2] Land Value Weight*
[3] Natural Beauty Weight
[4] Accessibility Weight Low
[5] Accessibility Weight High
[6] Feedbacks
High Beauty High Value
[7] Feedbacks
Low Beauty High Value*
[8] Population High
[9] Population Low
[10] Sample Size Large
[11] Sample Size Small
Normalized STI
Time (ticks)
STi for Largest Patch of Developed Area Index
* 1024 model runs
Neighborhood Decrease in Beauty (Feedback)
Neighborhood Increase in Value (Feedback)
Population Size (Number of Agents)
Land Value Weight
Natural Beauty Weight
Accessibility Weight
Size of Sample Drawn (Developable Cells)

Figure 8 [b]