



SimiVal, a multi-criteria map comparison tool for land-change model projections



Andrew V. Bradley^{a,*}, Isabel M.D. Rosa^a, Robert G. Pontius Jr.^b, Sadia E. Ahmed^c, Miguel B. Araújo^a, Daniel G. Brown^d, Amintas Brandão Jr.^e, Gilberto Câmara^f, Tiago G.S. Carnerio^{g,h}, Andrew J. Hartleyⁱ, Matthew J. Smith^c, Robert M. Ewers^a

^a Department of Life Sciences, Imperial College London, Silwood Park, Ascot, SL5 7PY, UK

^b Graduate School of Geography, Clark University, Worcester, MA, USA

^c Computational Science Laboratory, Microsoft Research, 21 Station Road, Cambridge, CB1 2FB, UK

^d School of Natural Resources and Environment, University of Michigan, MI, USA

^e Imazon, Travessa Dom Romualdo de Seixas, 1698, Edifício Zion Business, 11º andar, Belem, Brazil

^f Image Processing Division, National Institute for Space Research, São José dos Campos, São Paulo, Brazil

^g Department of Computer Science, Federal University of Ouro Preto, Minas Gerais, Brazil

^h Environmental Change Institute, Oxford University Centre for the Environment, South Parks Road, Oxford, OX1 3QY, UK

ⁱ Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK

ARTICLE INFO

Article history:

Received 14 October 2015

Received in revised form

14 April 2016

Accepted 15 April 2016

Keywords:

Land-cover modelling

Validation

Landscape metrics

Land-cover change

Model similarity

Quantity allocation

ABSTRACT

The multiple uses of land-cover models have led to validation with choice metrics or an *ad hoc* choice of the validation metrics available. To address this, we have identified the major dimensions of land-cover maps that ought to be evaluated and devised a Similarity Validation (SimiVal) tool. SimiVal uses a linear regression to test a modelled projection against benchmark cases of, perfect, observed and systematic-bias, calculated by rescaling the metrics from a random case relative to the observed, perfect case. The most informative regression coefficients, p-value and slope, are plot on a ternary graph of 'similarity space' whose extremes are the three benchmark cases. SimiVal is tested on projections of two deliberately contrasting land-cover models to show the similarity between intra- and inter-model parameterisations. We find metrics of landscape structure are important in distinguishing between different projections of the same model. Predictive and exploratory models can benefit from the tool.

Crown Copyright © 2016 Published by Elsevier Ltd. All rights reserved.

Software

Release version SimiVal.v1.1. (09/10/15).
Availability and cost Free software written in R, with demonstration files and a user guide, available at: <http://www.lufllondon.co.uk/lulcc-models.html>.
System requirements R (programming language) www.r-project.org, which runs on most desktop platforms under Unix, Windows and Mac operating systems, see www.r-project.org for detailed hardware requirements.

Developer

Andrew Bradley, Silwood Park,
Imperial College, Ascot, SL5 7PY. UK.

1. Introduction

Land-Cover Change Models (LCCMs) can be used to support science and decision-making for many domains, such as environmental change and sustainability (National Research Council, 2014). LCCMs are a popular topic for this journal, fresh articles address, the conversion of natural habitats (Ralha et al., 2013; Soares-Filho et al., 2013; Tayyebi et al., 2014), agricultural change (Celio et al., 2014; Tayyebi et al., 2014; Versteegen et al., 2014; Olmedo et al., 2015), urban change (Fuglsang et al., 2013; Pijanowski et al., 2014; Tayyebi et al., 2014; Liao et al., 2016), improved LCCM methodologies (Haase et al., 2012; Magliocca et al., 2015; Liao et al., 2016; Versteegen et al., 2016) and, inter-model

* Corresponding author.

E-mail address: a.bradley@imperial.ac.uk (A.V. Bradley).

comparison (Mas et al., 2014). These examples, and similar LCCMs, produce a land-cover change map, and usually involve evaluation of the properties which signify change within these maps. However, after the following review of modelling motivation and practices, it is clear these methods are inconsistent and we identify the need for, and provide, a more comprehensive and dependable assessment.

The major goals of LCCMs are to project land-cover change within or beyond a historical period, and understand how processes contribute to land-cover change, according to the influences of socioeconomic and biophysical drivers input into the model (Verburg et al., 2003, 2004). The influence of these drivers can often change, i.e. they are often non-stationary, which may affect the predictive accuracy of any model both within and beyond the historical period (Verstegen et al., 2016). Hence multiple models, or many intra-model, projections are often made, to predict, or project a range of possible outcomes. Furthermore, on a typical land-cover scenario the outcome may vary between different models because of the way individual LCCMs carry out similar procedures (Mas et al., 2014). In these cases, recent studies have allowed for model variability using multiple models (Pérez-Vega et al., 2012; Olmedo et al., 2015), and using multiple calibrations of the same model (Soares-Filho et al., 2013). Following evaluation, how similar these model projections are to reality or a speculated outcome, and what land-cover changes they determine, can give an insight into which land-cover change processes are significant (e.g. Celio et al., 2014). The operator is then better informed about how to interpret, use or improve a LCCM (Pontius and Millones, 2011).

Quantification of this similarity between projections and the observed map can be accomplished with validation, a statistical measure representing the goodness of fit between the projected and observed maps over the same time period (Costanza, 1989; Pontius and Schneider, 2001; Pontius, 2002). There are many dimensions of land-cover maps that can be validated, each with their own metric or metrics (Section 2). However, modellers are free to choose, and may well favour, which dimension to validate depending on the function of a LCCM and the discipline of a researcher (Verburg et al., 2004). Moreover, recent literature illustrates how similarity between maps is reported in different ways. For example, metrics may be summarised as bar charts (Tayyebi et al., 2014), multiple metrics may be summarised as cumulative bar graphs (Ralha et al., 2013; Olmedo et al., 2015), and validation scores from intra-model projections have been visualised on radar plots (Soares-Filho et al., 2013). This lack of consistency in evaluation, reporting and appraisal of model outputs can be problematic for retrospective inter-model comparisons (Rosa et al., 2014). Some basic, consistent reporting would broaden the application of LCCMs into the ever increasing environment of inter-disciplinary research. This paper presents a validation tool that goes some way to address these issues with land-cover change validation. In doing so we respond to a call for improved validation methods (Brown et al., 2013), a challenging prospect since LCCMs have multiple goals.

2. Validation methods

Spatially explicit LCCMs produce land-cover maps that show spatial patterns in the landscape, the location of change, and the quantity of change between different land-cover types (unless quantity is explicitly specified in the LCCM itself). Increasingly modellers have turned their attention to the evaluation of model products, particularly if the goals of the modelling are predictive in nature. This takes place with some form of ‘pattern validation’ (Brown et al., 2013) of the model projection against an observed reference. The observed reference is usually a land-cover map derived from aerial photographs (e.g. López and Sierra, 2010) or

satellite data sources, such as Landsat (e.g. Messina and Walsh, 2001), or MODIS (e.g. Etter et al., 2006), but they can also include other features such as fire hotspots (e.g. Silvestrini et al., 2011). In contrast to predictive modelling, researchers may project with an explanatory model to design policies and address specific societal issues (Filatova et al., 2013). Validation of output patterns from an explanatory model is generally considered to be of limited value because, the interest of the investigation is not to replicate actual outcomes, but to explain which parameterisations lead to particular spatial outcomes (Deadman et al., 2004). In these cases, attempts are made to validate the processes in the model (e.g. Haase et al., 2010; Sun and Müller, 2013). While reproduction of realistic output patterns may contribute to *process validation*, we acknowledge that additional information about the specific rules and mechanisms, temporal dynamics of model output, and sensitivity of model outputs to changes in parameters or processes are also needed. Here we focus on the procedures for ‘pattern validation’ of map-based outputs from LCCMs.

In the last few decades, researchers have concentrated on the development of pattern validation methods (see Pontius et al., 2011) and a suite of procedures have appeared, each with their own advantages and limitations (Table 1). These validation methods help score the accuracy of model calibration and output with respect to key dimensions of the land-cover maps. The key dimensions are identified as: differences in location, quantity, landscape structure, change probability and optimal model resolution. However, some metrics such as the popular Kappa statistics (Pontius, 2000), have now been described to have limited value to validation (Pontius and Millones, 2011).

Validation rarely covers all of these key dimensions, and it is only possible to evaluate LCCM performance with the dimension that a researcher has chosen to appraise their model with. Model to model this appears *ad hoc*, as no consistent validation procedure is ever applied, thus it is difficult to evaluate and compare the strengths and weaknesses of different models (Rosa et al., 2014). Presently, researchers may select one of the existing validation methods, use several together, and sometimes combine those methods or use their own validation scheme. As a result, the LCCM literature is replete with different combinations of approaches to validating land-cover maps: K_{fuzzy} and Receiver Operating Characteristic (ROC) (Lapola et al., 2011); the Kappa indices (Michalski et al., 2008); Kappa indices and Area Under Curve (AUC) (Sangermano et al., 2012); error matrix, ROC and Chi squared (López and Sierra, 2010; Pérez-Vega et al., 2012); ROC (Pontius and Pacheco, 2004; Tayyebi et al., 2014); ROC and expert knowledge (Wassenaar et al., 2007); ROC and conversion probabilities (Vance and Iovanna, 2008); multiple resolution filtering and landscape metrics (Soares-Filho et al., 2002); landscape metrics (Verstegen et al., 2014); fuzzy similarity (Almeida et al., 2008; Soares-Filho et al., 2013); ROC, fuzzy map and time series analysis (Silvestrini et al., 2011); AUC, distance-based metrics and pixel-by-pixel comparisons (Rosa et al., 2013); producers accuracy and visual inspection (Walker et al., 2004); cross tabulation (Etter et al., 2006; Messina and Walsh, 2001; Geoghegan et al., 2004; Walsh et al., 2008; Ralha et al., 2013; Celio et al., 2014; Olmedo et al., 2015); and comparing a time series of land-cover maps to a statistical envelope of several model projections (Evans et al., 2001). Spatial and quantity allocation are the key dimensions usually covered by these combinations, whilst configuration of landscape structure is less frequently considered and should be included (Mas et al., 2014). This may be because modellers consider that the correct spatial allocation will logically correct the problems with landscape structure. Conversely, the reverse may also be true in that correct simulation of landscape structure may help spatial allocation, particularly if a model is built to simulate land-cover change via

Table 1

A selection of land-cover model validation procedures, indicating the validated dimension, quantitative measure, and limitations.

Validation and dimension	Quantification	Comments
Visual inspection General appraisal of model projection in many dimensions	Qualitative	Intuitively identifies problem areas that a computer cannot but it is subjective and the operator may choose only to scrutinise one dimension of the output.
Receiver Operating Characteristic (Pontius and Schneider, 2001) Characterises relationship between predicted probability of events (e.g. transitions) and observed discrete phenomenon	0–1 (0–100%)	Often incorrectly used in validation of final land-cover projections but underexploited and open to more detailed interpretation in the validation of calibration data (Pontius and Parmentier, 2014).
Cross Tabulation Matrices (Congalton and Green, 1999) Correct transitions and quantity of transitions. Methods can be two (Pontius and Cheuk, 2006) and three dimensional (Pontius and Millones, 2011)	True positives, true negatives, false positives (commission error), false negatives (omission error)	It becomes more complex to explain and validate land-cover processes with increasing numbers of land-covers. No structural information about the landscape.
Quantity allocation kFuzzy maps (and multiple resolution filtering) (Costanza, 1989; Hagen, 2003; Almeida et al., 2008)	% and window resolution	Intra- and inter-model comparison may not be possible as the best resolution may vary for each parameterisation. Does not evaluate landscape structure.
Spatial allocation Kappa statistics (Pontius, 2000) (i) <i>Kappa</i> – quantity (ii) <i>Kstandard</i> – quantity (iii) <i>Kno</i> – quantity (iv) <i>Klocation</i> – location (v) <i>Kquantity</i> – quantity	1 to –1 % % %	Criticised by creator as having limited use (Pontius and Millones, 2011) because of: 1. Ambiguity over the numerator/denominator determining value of ratio; 2. Does not separate two components of disagreement, quantity and allocation; 3 The <i>Kstandard</i> is no different from proportion correct, and; 4. Kappa compares to randomness not a null or naïve model.
Three way map comparison statistics (Pontius et al., 2008) (i) <i>Quantity disagreement</i> Quantity allocation (ii) <i>Near and far location disagreement</i> Spatial allocation (iii) <i>Users and producers accuracy</i> Quantity allocation (iv) <i>Figure of merit</i> Overall score	% % % and score >1 % and score >1	These metrics concentrate on spatial and quantity allocation and do not evaluate landscape structure.
Landscape metrics (McGarigal and Marks, 1995) Landscape structure and configuration	Depends on the metric	With no reference envelope as to what a bad value is, it is difficult to say how ‘out’ the value of the landscape metric is in comparison to the observed value.

landscape structural changes, e.g. the expander and patch function in DINAMICA (Soares-Filho et al., 2002). Awkwardly, metrics of landscape structure do not lend themselves well to validation as it is uncertain how different the projected landscape metric values have to be from the observed landscape metric values to signal a poorly replicated landscape structure. A similar issue arises when quantifying the degree of similarity in spatial allocation.

To reduce the temptation for researchers to use a selection of metrics that favours a particular model, a procedure that avoids *ad hoc* selection of comparison metrics and covers many different dimensions of land-cover maps, is required. It must be a generic approach that evaluates the key dimensions of quantity allocation, location accuracy, and overall spatial configuration between maps. The procedure must also overcome the difficulty of metrics with boundless ranges that cannot describe the magnitude of difference between observed and predicted projections, such as landscape metrics. As there are a number of goals in land-change modelling, such a validation method would need to be flexible in that the method can provide:

- a comparison, for modellers who wish to validate how well the model parameterisation correctly replicates the observed land-cover patterns, allowing them to evaluate the similarity of projections beyond a historical period, assuming stationary processes. This may be when the goal is the validation of a single model run;
- a consistent assessment, for several model runs, so a modeller will be better informed on how to interpret or reject different model parameterisations based on their strengths

and weaknesses, and thereby help inform a subsequent round of model improvement. This may be when the goal is for an intra- or inter-model comparison;

- a relative measure, to show the range of land-cover possibilities, compared against either reality or a speculative target land-cover map, to understand how different parameterisations (or processes) of the same model project land-cover change. This may be when exploratory modelling is the goal.

To address these needs we have developed and present the Similarity Validation (SimiVal) tool. The output presents an intuitively interpretable graphic so a modeller can quickly assess how dissimilar the projections are from reality and, understand how different parameterisations of models influence the key dimensions of land-cover projections. To demonstrate the capabilities of SimiVal, we deliberately use two structurally different LCCMs to provide degrees of dissimilarity from the observed land-cover and between each model projection. We do not focus on the meaning of the outcomes of the model projections as the aim of the paper is to demonstrate the versatility of the validation tool.

3. The similarity validation tool (SimiVal)

3.1. SimiVal design

SimiVal requires three maps: (1) the modelled land-cover projections (p) as a categorical map at the projected end time point (p^{t2}), and generated from a land-change model calibrated between

two observed (o) time points (o^{t0} and o^{t1}); and (2, 3) the observed change as two categorical maps for the beginning and end time points of the projection period (o^{t1} and o^{t2}). The tool then creates a fourth random (r) change categorical map at the second time point (r^{t2}), that is constrained to duplicate the same amount of land-cover change and land-cover transitions that occurred in each category between o^{t1} and o^{t2} . For each of these categorical maps, SimiVal computes a series of component metrics representing different dimensions of land-cover change, classified as non-spatial (quantity) and spatial (structural) metrics. Metric scores from o^{t2} represent the perfect metric scores for which the modeller may aim. Metric scores from r^{t2} represent the random metric scores from which the modeller may want to avoid. Metric scores at p^{t2} represent the metric scores that the modeller wishes to evaluate relative to the other two. SimiVal compares these component metrics with a linear regression. However, since: (i) many of these metrics are of different orders of magnitude, the metrics are first log transformed as some metrics exert undue leverage on the summary regression statistics while others exert very little and; (ii) because there is no reference value to determine how wide of the mark a projection is in a regression between the observed and projected metrics, the random metric scores are scaled relative to the observed scores, providing benchmark metric values for the perfect and random cases. These benchmark conditions are calculated as the percentage difference between the perfect metric scores and the random metric scores and then standardised by setting the perfect metric scores to zero. This procedure also provides a 'bad' score for a boundless metric to compare to the observed. The perfect case is, the regression between the standardised observed and standardised random metrics, a flat line, and the random case is, a regression of the standardised random metrics against each other, a line at 45° (Fig. 1).

When the modelled metrics, p^{t2} , are standardised and plot against the standardised random metric scores on the same graph, a modeller can now draw rigorous conclusions of how close or far away a modelled projection is from either case.

SimiVal assumes that all modelled projections are similar to the coefficients of the random case until proven otherwise. From the regression statistics the test of how similar a modelled projection is to a random model can be found with the p-value and the coefficients of the simple linear regression line (Eq (1)).

$$y = mx + c \quad (1)$$

where m = slope and c = intercept. If we suppose that all models are random until proven otherwise, then a perfect fit to the metrics of a random model has a significant relationship between the diff % predicted and diff % random values, p-value = 0, with a slope $m = 1$, and intercept $c = 0$. By contrast, if the model projection conforms to the null hypothesis it is perfect and therefore is not a random model, there would be no significant relationship between the diff % predicted and diff % random values, the p-value would approach 1, and the regression would have a slope $m = 0$, and intercept $c = 0$. In reality, the model projection that is tested will have values between these coefficients, providing a measure of how similar the modelled projection is to the random or perfect case. The intercept c , may also be ± 0 . There can also be cases when the regression has a p-value = 0 but with a slope $m = -1$ and an intercept $c = 0$. Although the p-value suggests a random model, the slope coefficient predicts an opposite polarity to the random model metrics, a condition that does not conform to the random model or the perfect model. In these cases SimiVal is indicating a model projection that incorporates a systematic, non-random bias, and might suggest that the land-cover change in the modelled projection is structured and non-random, i.e. the model does not produce a

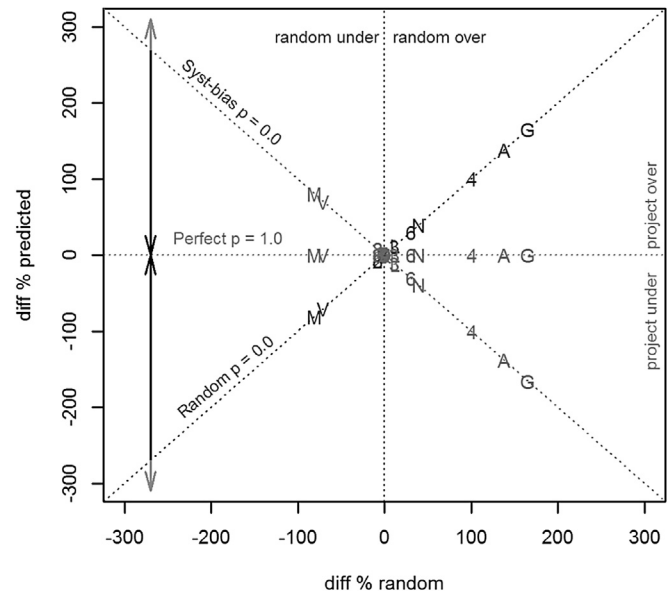


Fig. 1. The benchmark cases of random, perfect and systematic-bias of land-cover projections. Axes are: the percentage difference between the random metrics (or the modelled projection metrics) and the observed metrics (diff % predicted) versus the percentage difference between the random model metrics and the observed metrics (diff % random). Coefficients for model projections will depend how similar they are to, a completely random model (slope = 1, $p = 0$), the observed 'perfect' model, (slope = 0, $p = 1$), or whether the model contains a systematic-bias, (slope = -1, $p = 0$). Regression lines that rotate towards the observed regression (dark arrows) show increasing similarity to the observed projection. Regression lines that rotate away from the observed regression with a slope steeper than -1, are worse than a systematic-bias or with a slope steeper than 1, are worse than random (light arrows). Metric symbols are: N – number of patches, A – average-area-to-edge-length ratio, V – variance-average-area-to-edge-length ratio, M – Moran's I, G – Geary's C, L – distance allocation metric, 1, 5 & 9 persistence Px:x, and 2, 3, 4, 6, 7 & 8 the transitions Tx:y. (see Section 3.2.1 for full details on Px:x and Tx:y).

random map but cannot fully replicate the processes that led to the observed map patterns. To represent the case of systematic-bias, a naïve model, where change is simulated adjacent to an obvious predetermining factor (Pontius et al., 2007), could be used. However, a naïve model may return many manifestations of patterns when limited to a fixed allocation of change, so we assume the case of a perfect systematic-bias is a regression based on metric scores with opposite behaviour to the random metric (Fig. 1).

Using the three benchmark cases, SimiVal is able to evaluate, quantify, and visualise a modelled projection and help determine if that model is random, perfect or contains a systematic-bias. The regression also allows a deeper investigation into the structural performance of models, as it is possible to see whether individual metrics are being over-, or under-estimated as indicated by their residuals from the linear regression. SimiVal also makes a direct visual comparison between different modelled projections for an inter- and intra-model comparison by plotting each modelled projection in a ternary graph representing what we refer to as 'similarity space'. At each extreme, or apex, of the 'similarity space', is the benchmark score for the perfect, random and systematic-bias cases (Fig. 2). The acceptable level of similarity to the perfect, observed case is at the user's discretion depending on their specific land-cover modelling goals.

3.2. SimiVal component metrics

The component metrics in SimiVal have been selected so land-change modellers can identify how the structural features and

parameterisations within their land-cover change models influence their different land-cover projections. These metrics are firstly calculated for the observed land-cover map at time 2 (o^{t2}), projected map at time 2 (p^{t2}), and random map at time 2 (r^{t2}). The component metrics represent non-spatial and spatial differences between the o^{t1} , p^{t2} and r^{t2} , land-cover maps. Non-spatial properties are defined here as quantity of change metrics (total quantity for each land-cover, quantity of transition between land-covers and the persistence of each land-cover). Spatial metrics are defined here as spatial autocorrelation of the overall change, landscape structure, and distance allocation.

3.2.1. Non-spatial (quantity) metrics

These metrics are computed to evaluate how well the model reproduced how much of each land-cover change transitions (between all pairs of classes) actually happened. They are calculated by comparing the observed number of land-cover transitions between o^{t1} and o^{t2} with the projected number of transitions between o^{t1} to p^{t2} . SimiVal categorises, calculates and records the number of pixels for all the possible pairwise transitions, including persistence (Table 2 and Eq. (2) and Eq. (3)), and then sums each transition to give the total transition between the two time periods (Eq. (4))

$$x^{t1} \cap x^{t2} = Px:x^{1-n} \quad (2)$$

$$x^{t1} \cap y^{t2} = Tx:y^{1-n} \quad (3)$$

$Px:x$ = total units of persistence of land-cover x following the union of the categorical land-cover maps from $t1$ and $t2$ for each transition 1 to n and, $Tx:y$ = the total units of transition from land-cover x to land-cover y following the union of the categorical land-cover maps from $t1$ and $t2$ for each transition 1 to n .

$$\sum (Tx:y^1, Tx:y^2 \dots Tx:y^n) = \text{totaltrans}^{t1-t2} \quad (4)$$

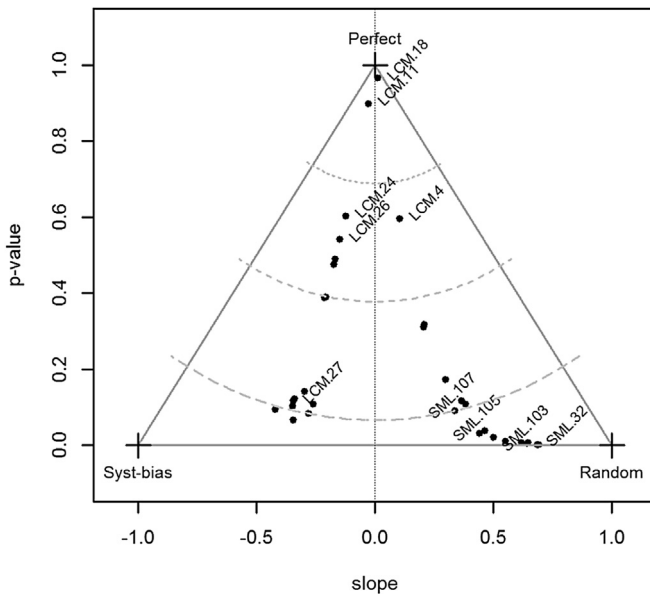


Fig. 2. Similarity space. Cross-hairs mark the benchmark apices for the perfect, random and systematic-bias cases. Model predictions that fall inside the ternary graph show how close that projection is to one of the three extreme cases, while any that might fall outside the ternary graph are either worse than random or are extreme cases with a systematic-bias. Dashed lines represent user-defined acceptable thresholds of similarity to the perfect case. Dots represent values for models used in the worked example, described in Section 4 are, all LCM projections (LCM1-30), SML realisation 32, and the SML cumulative probability maps (SML101-107).

Table 2

Transitions (T) and persistence (P) between three land-covers.

	Class	$t2$		
		1	2	3
$t1$	1	P1:1	T1:2	T1:3
	2	T2:1	P2:2	T2:3
	3	T3:1	T3:2	P3:3

where, $\text{totaltrans}^{t1-t2}$ is the total sum of all possible transitions, $Tx:y$ 1 to n , between $t1$ and $t2$.

3.2.2. Spatial metrics

Spatial metrics are calculated from the categorical maps of total observed and total projected land-cover change and quantify the ability of land-cover models to replicate the spatial structure of overall change in the landscape.

3.2.2.1. Spatial autocorrelation. The degree of clustering, randomness and dispersion in the map of transitions is quantified using Moran's I (Gittleman and Kot, 1990) and Geary's C (Cliff and Ord, 1973) indices. A Moran's Index value near +1.0 indicates clustering while an index value near -1.0 indicates dispersion. Geary's C ranges are generally inverse to Moran's I beginning with a value of 0 for perfect spatial autocorrelation and with a value of 1 indicating complete absence of spatial autocorrelation. Both indices are used as Geary's C is more sensitive to localised clustering of data, whereas Moran's I has better global sensitivity. These metrics indicate how well the model has predicted patchiness and dispersion for all the combined land-cover transitions.

3.2.2.2. Landscape structure. Landscape metrics provide statistics describing patterns in the landscape. Many landscape metrics are correlated and provide redundant information (Riitters et al., 1995), so for a concise account of landscape structure SimiVal uses: number-of-patches (nop), average-perimeter-to-area-ratio (paAve), and the variance-in-perimeter-to-area-ratio across all patches (paVar) with units in metres. These landscape metrics were quantified for the observed, random and all projected model runs to determine if the maps had similar numbers of isolated land-cover patches (nop), and if the individual patches have similar shape or compactness (paAve and paVar). Higher values indicate less compact shapes. The R algorithm we used to define a patch includes adjacent corners (Chang et al., 2004).

3.2.2.3. Distance allocation of change pixels. The SimiVal spatial allocation metric, uses the Wilcoxon rank sum test statistic, W (Crawley, 2005), to measure and compare how near or far instances of change pixels in the projected model and the random model are from the observed change pixels. This is done by generating a W score for a perfect match (observed compared to observed) and a W score for the difference from observed to each model projection and to the random projection. A model projection similar to the observed will have a W score close to the perfect match. A random model is unlikely to be the same as the observed and will have a much larger W score, whereas a projected model ought to have a W score that falls between the perfect match and the random model W scores, assuming the model projection is not worse than random. The final W score is reported as a percentage of the perfect W score (Eq (5)).

$$(W^{proj}/W^{perfect}) \times 100 \quad (5)$$

4. Data, and land-cover change models

4.1. Data description

We used an area of township boundaries that intersect, or are inside, the watershed of the Plumb Island Ecosystems (PIE) ecological research site, Massachusetts, USA (Fig. 3). In these townships, housing is replacing forest and knock on effects of the land-cover change are thought to be contributing to the depletion of aquifers and increases of algal blooms in the coastal bays (Kirkby et al., 2000).

Land-cover data at 30 m resolution for 1971, 1985 and 1999 was used. There are 21 land-cover classes in these data sets, which were simplified into the land-cover classes 'forest', 'built' and 'other' for the modelling. The 'other' class consists of variety of relatively small categories such as mines, dedicated open land, pasture, cropland, wetlands and water bodies. Two land-cover-change models, with contrasting structures, were used to demonstrate SimiVal: the IDRISI Land Change Modeller (Clark Labs, 2009), referred herein as LCM, and StocModLCC (Rosa et al., 2013), referred herein as SML. Additional explanatory variables for the models were either calculated from the land-cover maps or downloaded (<http://pie-iter.ecosystems.mbl.edu/>). LCM used elevation data, geology, distance-to-1971-built-land and the 1971 land-cover map as predictor variables, whereas SML used distance to major settlements, wetland areas, protected areas, farmland, distance-to-roads, national landscapes, flood areas, geology and planning zones. Each model was calibrated between the years 1971 (t_0) and 1985 (t_1). We applied the land-cover models assuming that land-cover-change processes were stationary, and projected forward another 14 years to 1999 (t_2). SimiVal calculated the metrics, regression and plot the coefficients in the 'similarity space' described in section 3.

4.2. Model 1: Land Change Modeller (LCM)

LCM bases its projection on transition sub-models that use the calculated transition potentials between the calibration dates for each land-cover type and chosen explanatory variables. Three sub-models were defined: (i) built-gain-from-forest; (ii) other-gain-

from-forest and; (iii) built-gain-from-other, using combinations of elevation, 1971 land-cover, surface geology and distance to-1971-built-land as explanatory variables. Transition maps that indicate the actual change in land-cover during the calibration period were used as the input data to calculate the transition potentials of each of the explanatory variables. The transition models then provided transition evidence maps, which were used as the basis for the transition matrix calculated by Markov Chain prediction. The transition matrix supplied the proportion of change for each land-cover transition, which remained the same for each projection. In this instance, two transition models were used: the multiple layer perceptron neural network model, and a similarity weighted method. The model thus provided two different internal structures, with three transitions and 15 combinations of explanatory variables to produce a total of 30 land-cover projections, labelled as LCM.1 to LCM.30. For this exercise, the categorical map, or hard classification, calculated from a land allocation model built into the IDRISI software, was used. This model will tend to select locations where the most influential explanatory variables coincide with locations where transition potentials are highest, and we may therefore expect land-cover change to occur as small contiguous patches, which may be isolated or adjacent to the existing 'built' class. Considering these model characteristics, we would anticipate that the SimiVal metrics of the modelled projections may show a high spatial autocorrelation, consistency in compactness, and a similar allocation distance to the observed. Different transition evidence maps will vary the location and size of these patches between each projection.

4.3. Model 2: StocModLCC (SML)

SML determines that land-cover change occurs when the transition probability of a pixel is high enough for the land-cover transition. This probability is calculated from the combined individual posterior probability distributions for each of the explanatory variables, and reflects the land-cover condition of surrounding neighbourhood pixels as well as their proximity to a combination of other explanatory variables. The posterior probability for each parameter is drawn from a likelihood distribution that has been

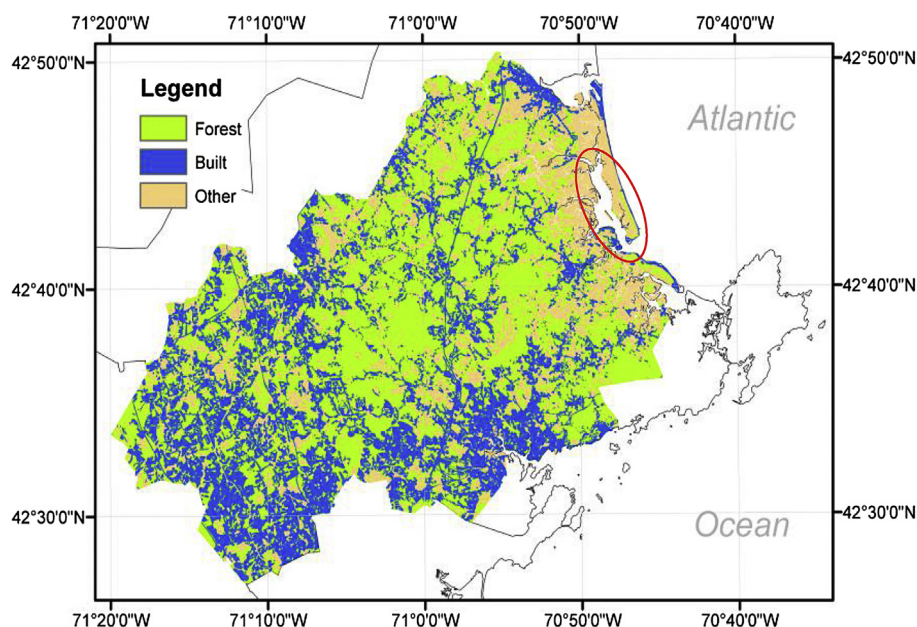


Fig. 3. Land-cover within the township boundaries that intersect the watersheds draining into the Plumb Island Sound (ringed), Massachusetts, USA.

calculated during the calibration with Markov Chain Monte Carlo sampling, which themselves have been validated for goodness of fit with a 50:50 split between calibration and test data. Three transitions were calibrated using the following explanatory variables: (i) forest-to-built-or-other, which used distance to roads, protected areas, farms, national landscape areas, flood areas, planning zones and geology plus the internal variable, 'proportion of forest neighbours'; (ii) built-to-other-or forest, which used distance to roads, protected areas, farms, geology and the internal variable 'proportion of built neighbours', and; (iii) other-to-forest-or-built which used distance to roads, protected areas, farms, national landscape areas, flooded areas, planning zones and the internal variable 'proportion of other class neighbours'. The built-back-to-forest transition did not occur in the calibration time period and, therefore could not be projected, leaving the following transitions: forest-to-other-or-built, other-to-forest-or-built (competition, multinomial regression) and, built-back-to-other (no competition, binomial regression). The predictions of the three models were simultaneously combined to produce the final land-cover-change projection. To allow for uncertainty in the calibration of this model there were 100 iterations, or realisations, i.e., 100 equally probable versions, labelled SML.1-SML.100, produced by a stochastic process, normally combined to create a probability change map (Rosa et al., 2013). However, for SimiVal a hard classification land-cover map is required, meaning the probability maps emerging from SML needed to be categorised. For this study, it was sufficient to cap the highest probability pixels with a threshold calculated from the average number of change pixels of the 100 SML realisations, giving a map showing pixels of change or no change. This land-cover map is referred to as the cumulative probability map. To evaluate the influence of the number of realisations on the final cumulative probability map, thresholds were calculated for 4×25 (SML.1-SML.25, SML.26-SML.50, SML.51-SML.75, SML.76-SML.100), 2×50 (SML.1-SML.50, SML.51-SML.100), and 1×100 (SML.1-SML.100) realisations to create an additional seven cumulative probability maps labelled SML.101-SML.107. SML selects individual pixels for change based on a weighted probability 'coin flip'. Thus contiguity in land-cover change is less likely to occur than in LCM and change pixels could be close to or far away from the observed change. Considering the model characteristics, we would anticipate that the SimiVal metrics are likely to show, low values of spatial auto correlation, variation in compactness of patches, large differences in quantity allocation, and a wide range of variation in distance allocation against the observed map.

5. Results

5.1. Modelled spatial patterns

The differences between the observed land-cover change and three examples of contrasting modelled projections are shown as subsets of the full model area (Fig. 4). Between 1985 (o^{t1}) and 1999 (o^{t2}), the observed land-cover class transitions (print/online colour version) were mainly the appearance of small parcels and expansion of forest-to-built (medium grey to dark grey/green to blue) areas (Fig. 4a and b). The projection 1985 (o^{t1}) to SML.107 (p^{t2}) illustrates transitions of both forest-to-built (medium grey to dark grey/green to blue) and instances of other-to-forest (light grey to dark grey/beige to blue), and transitions were distributed predominantly as individual pixels or small groups of pixels rather than as small parcels or expanding parcels (Fig. 4c). The projection from 1985 (o^{t1}) to LCM.18 (p^{t2}) and to LCM.24 (p^{t2}) also projected forest-to-built (medium grey to dark grey/green to blue), but did not replicate the 1999 (o^{t2}) patterns, allocating different quantities in the same

and different places to 1999 (o^{t2}). The projection of change in LCM.18 (p^{t2}) had a more parcel like appearance of built (dark grey/blue) areas (Fig. 4d) whereas LCM.24 (p^{t2}) tended to systematically expand existing built (dark grey/blue) areas (Fig. 4e). At this particular location projection LCM.18 (p^{t2}) showed a slight propensity to project more land-cover transitions between other-to-built (light grey to dark grey/beige to blue), than did LCM.24 (p^{t2}).

5.2. Metric summary

A summary of the metrics for the observed, random, LCM and SML model projections are shown in Table 3. For the non-spatial metrics (quantity, Tx:y, Px:x), the random model used the same total quantity of change as the observed change, but over allocated the change pixels across the transitions for the least likely transitions (e.g. increases in T1:3 (built-to-other), T2:1 (built-to-forest) and T2:3 (forest-to-other)), but under allocated change to the more likely transition (e.g. increases in T1:2 (forest-to-built)). The mean scores of the two models showed that both LCM and SML projected a lower quantity of change pixels than was observed.

As expected for the dominant transitions during each model calibration, change pixels were allocated to T1:2 (forest-to-built) and T3:2 (other-to-built), but the models could not project T2:1 (built-to-forest). The remaining transitions either had low numbers of pixels or no pixels at all. The predicted quantity allocation between the two models also differed (Table 3). For LCM there was no variation around the mean, and the mean values were the same in each individual projection because the quantity of change for each transition was fixed by the transition matrix. By contrast, SML had variation around the mean quantity allocation because, rather than impose a fixed change quota, a change pixel was determined by 'coin flip'. The transitions in the cumulative probability map, LCM.107, slightly underestimated the mean SML projection because some of the high probability values of the combined realisations were excluded during the probability capping.

For the spatial metrics scores (nop, paAve, paVar, M, G and alloc) there was a wide variation between the observed and random models. This was primarily due to larger contiguous change patches occurring in the observed maps, whereas the random projected allocated change to many isolated pixels and small patches. As anticipated, the LCM projections predicted reasonably high levels of spatial autocorrelation (M & G), whereas SML predicted much lower levels. There were fewer patches (nop) predicted in LCM than SML, and the compactness of the patches in LCM was closer to the observed than to the random, whereas the compactness of patches in the SML projections was more similar to the random than to the observed. There was more variability in compactness from the LCM than SML projections. The variability occurred because the SML patches were consistently small groups of pixels that did not form large contiguous areas of change, whereas the range of patch sizes in the LCM projections was much larger. The allocation distance was poorer in the LCM projections than the SML projections, probably because the more widely distributed selection of change pixels in SML ensures those pixels can end up being anywhere from quite close to quite far from the observed change, whereas with the LCM projections contiguous clusters of many pixels may collectively be far away meaning that each component pixel in the change patch has a similar distance to the observed change. Overall, the mean spatial metric values for LCM were closer to the observed than to the random model, whereas the opposite was true for the SML realisations. However, the SML cumulative probability map (SML.107) was less like the random model than the individual 100 realisations, demonstrating that the cumulative probability map encouraged clustering of pixels with the effect of improving the overall spatial predictions of SML.

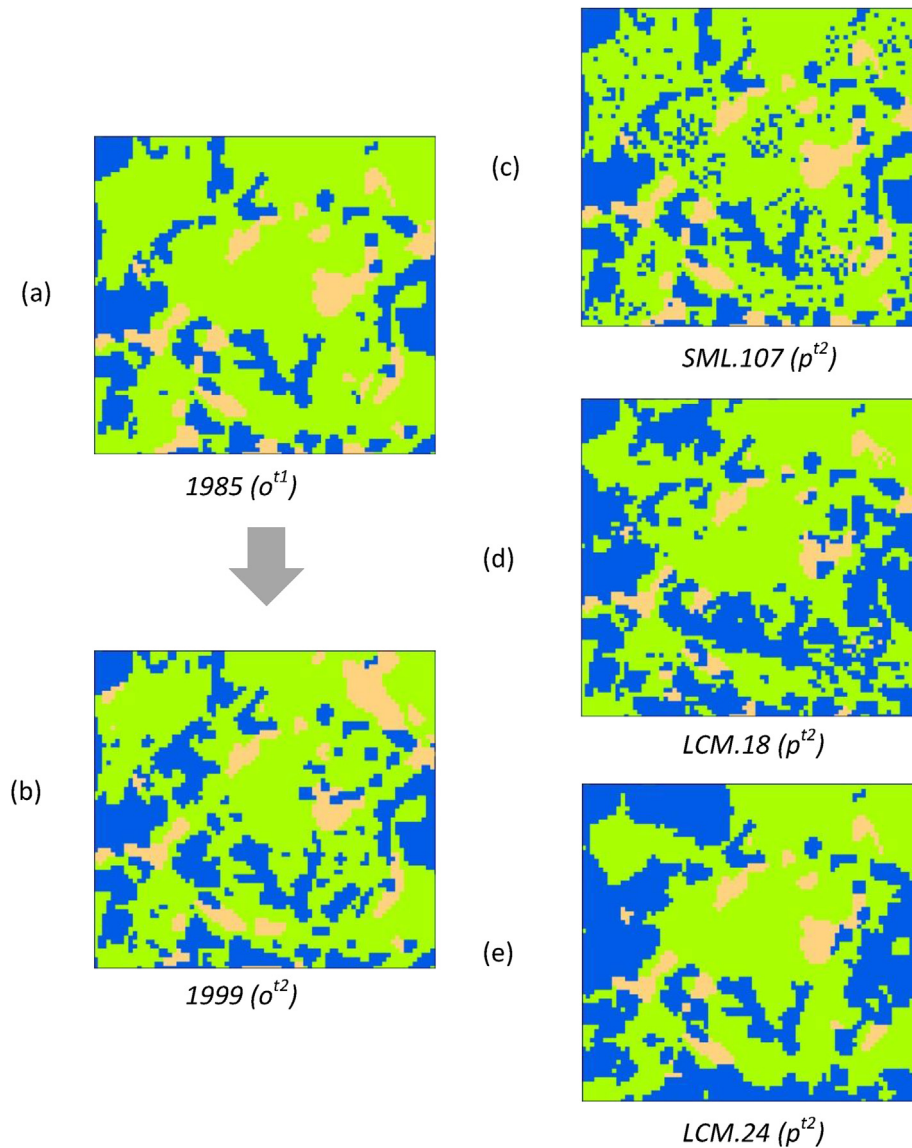


Fig. 4. Land-cover maps for, (a) 1985, (b) 1999, (c) SML cumulative probability map 107, (d) LCM projection 18, (e) LCM projection 24. The arrow between (a) and (b) signifies observed change. Land-cover classes are (print/online colour version), forest (medium grey/green), built (dark grey/blue), other (light grey/beige). Pixels are 30 m by 30 m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.3. SimiVal statistics

Typical examples of the SimiVal regression illustrate the relative similarity of the validation metrics for the LCM and SML projections to the perfect, random, and systematic-bias cases (Fig. 5). The metrics of one SML realisation, SML.32, scored very close to the random case metrics with a slope 0.68, intercept -14.95 and p -value <0.01 (Fig. 5a). This was largely due to the near-random selection of individual pixels that did not replicate the observed patterns of change well. However, combining all the 100 realisations into the cumulative probability map (SML.107) improved this situation, changing the regression coefficients to a moderate slope 0.34, intercept -11.48 and p -value 0.09 (Fig. 5b). For LCM.18 many of the metrics had values similar to those of the observed, perfect case with a shallow regression slope of 0.01, intercept of -76.79 and p -value of 0.97 (Fig. 5c), reflecting the ability of the LCM model to create patches of contiguous pixels that were more similar to the

observed map. By contrast some metrics in LCM.24 exhibited a tendency to move towards the systematic-bias, resulting in a negative slope -0.12 , intercept -1.90 and p -value of 0.60 (Fig. 5d). In this case the LCM model was concentrating change in a few places rather than across the landscape. The remaining projections showed variations of the themes in Fig. 5 and are summarised in Table S1.

The plot between all the p -value and slope scores provides an inter- and intra-comparison of all the modelled projections together (Fig. 2.). The ternary plot provides an instant visual evaluation of the relative similarity of all these projections to the perfect, random, and systematic-bias cases. This is much simpler to interpret than the table of values (Table S1) and gives a collective summary of the individual regressions (e.g. Fig. 5). Model projections with the shallowest gradients, LCM runs 4, 11, 18, 24 and 26 had the highest p -values and were most similar to the observed, perfect model. All other projections had greater positive or negative

Table 3

Summary of the metrics for the observed ($n = 1$), random ($n = 1$), LCM ($n = 30$), and SML ($n = 100$) projections and for the selected individual projections LCM.18, LCM.24 and SML.107. Quantity – total number of pixels of change of all land-cover transitions, T – refers to transition and P – refers to persistence between t^1 and t^2 . nop – number of patches, paAve – average area to edge length ratio, paVar – variance average area to edge length ratio, M – Morans' I, G – Geary's C, Alloc – distance allocation metric. Units are in pixels.

Metric	Observed	Random	LCM (mean)	LCM (SD)	LCM 18	LCM 24	SML (mean)	SML (SD)	SML 107
Quantity	95,361	95,361	72,605	0	72,605	72,605	58,750	327	49,993
P1:1	489,425	503,063	494,172	0	494,172	494,172	506,723	266	514,833
T1:2	47,260	20,663	41,521	0	41,521	41,521	36,754	269	28,710
T1:3	7,497	20,456	8,489	0	8,489	8,489	704.5	56	639
T2:1	121	15,507	0	0	0	0	0	0	0
P2:2	409,261	379,981	411,035	0	411,035	411,035	410,754	21	410,857
T2:3	1,653	15,547	0	0	0	0	281	21	178
T3:1	13,935	11,544	0	0	0	0	22	11	1
T3:2	24,895	11,644	22,595	0	22,595	22,595	20,988	204	20,465
P3:3	266,094	281,736	282,329	0	282,329	282,329	283,914	204	284,458
nop	2,917	68,806	3,294	2,436	3,774	2,882	34,843	166	10,823
paAv	0.05	0.13	0.08	0.02	0.09	0.09	0.13	0.83×10^{-4}	0.12
paVar	5.31×10^{-4}	1.56×10^{-4}	9.18×10^{-4}	3.87×10^{-4}	13.44×10^{-4}	12.99×10^{-4}	2.23×10^{-4}	0.02×10^{-4}	4.70×10^{-4}
M	0.73	0.11	0.63	0.20	0.72	0.77	0.18	8.35×10^{-4}	0.54
G	0.30	0.99	0.41	0.23	0.32	0.25	0.92	9.37×10^{-4}	0.52
Alloc	100	183	151	14	158	154	141	<1	129

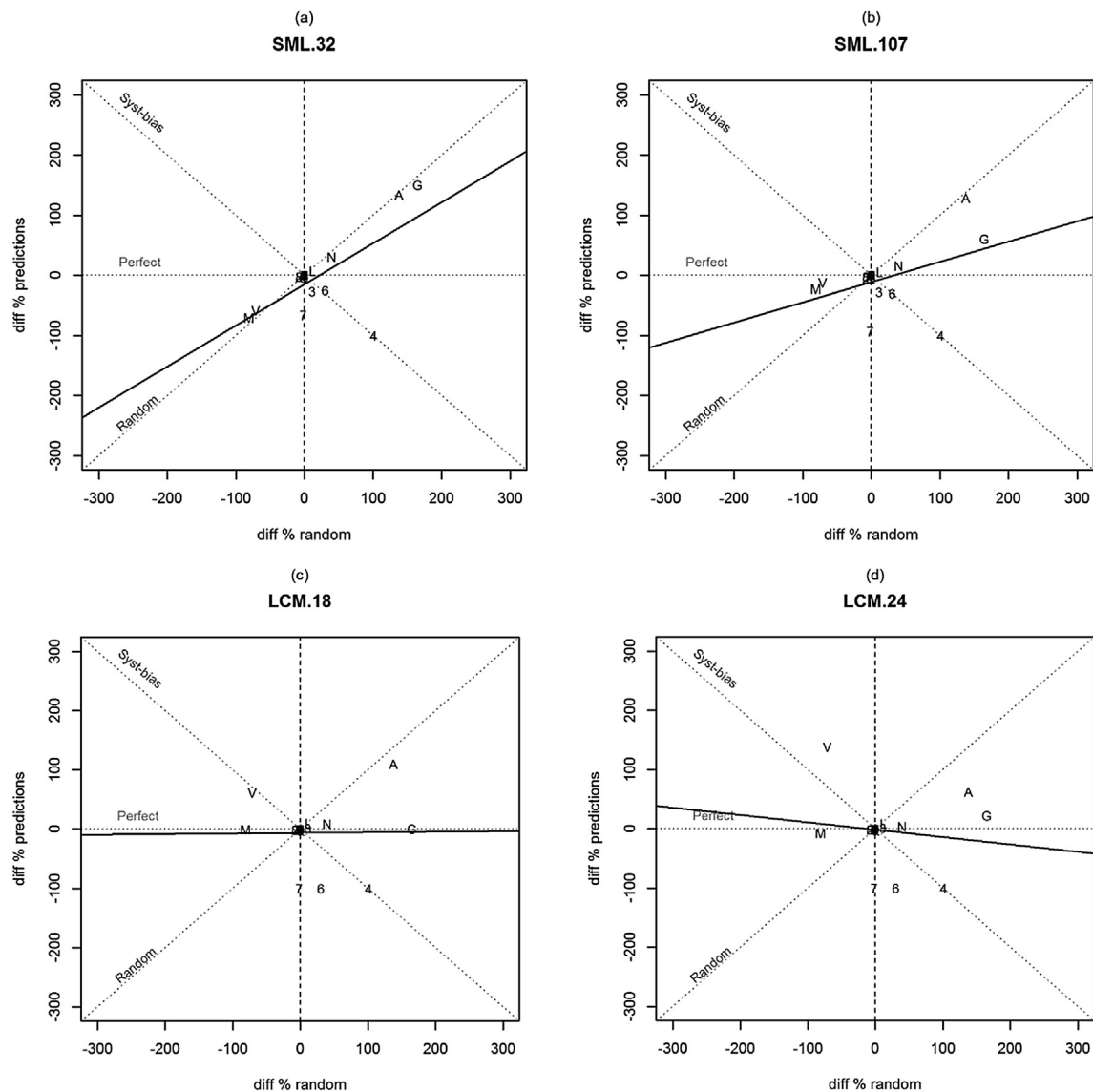


Fig. 5. Linear regression of SimiVal metrics for four model projections (a) SML.32 (b) SML.107, (c) LCM.18, (d) LCM.24. Metric symbols are described in Fig. 2.

slopes, in particular a group of LCM projections with negative gradients less than -0.2 and thereby exhibiting systematic-bias, whereas the SML projections tended to have positive slopes greater than 0.3, and thereby exhibiting a tendency to the random case. Accordingly the p -values decrease to the base of the ternary plot as the modelled projections become less similar to the observed, perfect case.

6. Discussion

To validate land-cover maps from LCCMs we have built a tool in R code that includes a set of metrics covering the major dimensions of land-cover change maps: quantity allocation; distance allocation and structural characteristics. When these metrics are combined and analysed collectively the similarity between different model projections can be visualised. The visualisation provides an instant recognition of how similar or dissimilar models can be to the three cases of a perfect, random, or systematic-bias. This is useful for: (i) validation of a single model projection; (ii) validation of multiple projections from for an intra- or inter-model comparison and; (iii) identifying how different model parameterisations influence the component metrics. This approach encourages consistent validation to evaluate the outcomes of one or many projections avoiding *ad hoc* choices of metrics and arbitrary decisions that can highlight specific objectives and model features.

The use of 'similarity space' (Fig. 2), avoids the first-past-the-post vision that validation tends to invoke, and instead provides a measure of similarity for each projection depending on how close the projection is to each of three benchmark apices. A perfect model will sit close to the perfect apex, calculated from what was actually observed to occur, and indicates a parameterisation of land-cover processes that leads to the same outcome as reality. A projection that plots towards the centre of similarity space is becoming more dissimilar to the perfect, and correspondingly exhibiting greater similarity to the random or systematic-bias cases. Additionally, when the projection is positioned towards either the random or systematic-bias apices, the modelled run can be judged as being parameterized in such a way that the land-cover change processes represented in the model are not realistic and that the resultant outcome is not accurately reflecting reality. Exactly where the boundary of realism sits in 'similarity space' may well be a decision for the individual analyst who ultimately judges the realism of the modelled land-cover projections and, how good a model needs to be before it is acceptable for a particular purpose. By drawing threshold lines across 'similarity space', the user can add certainty to this subjective task, but the location of the thresholds will likely be user-specific (Fig. 2).

The wide range of dimensions that are included in SimiVal is an advance on existing validation procedures. In addition to the metrics which are highly informative with respect to allocation (e.g. Costanza, 1989; Hagen, 2003; Almeida et al., 2008; Pontius et al., 2008, 2011), we paid particular attention to evaluation of land-cover maps including metrics of spatial correlation and landscape structure, and therefore provide a more diverse pool of metrics. Other authors have used metrics of landscape structure (e.g. Soares-Filho et al., 2002; Dale et al., 1994), and this added dimension is particularly useful to models that specifically drive change *via* landscape structure (e.g. DINAMICA Soares-Filho et al., 2002), as opposed driving change *via* quantity and probability (e.g. LCM and SML). Alarming, we found that when using quantity metrics alone it was difficult to distinguish between different projections from the same model (Table 3), and in fact was impossible for LCM. This represents a critical problem, as each LCM projection had a different modelled spatial pattern of change (Fig. 4). Thus the addition of spatial metrics is useful in differentiating and

understanding these projections. For example, although LCM.18 (p^{t2}) did not spatially allocate change in the right places, the model was better at predicting the structural features of land-cover change than LCM.24 (p^{t2}).

Although SimiVal is not the ultimate answer to validation, it is a step towards consistency. The method avoids inadvertently imbalanced assessments and contributes to the need for improved land-cover model validation (Brown et al., 2013), reducing the difficulties identified during inter-model comparison (Rosa et al., 2014). In our illustration of SimiVal we found models varied greatly in their ability to recreate observed change, and that models made erroneous predictions in very different ways. Using the model positions in 'similarity space' we can explicitly quantify on: (i) parameterisations that produced acceptable land-cover maps and; (ii) parameterisations that produced inappropriate maps. This approach could be used for explicit model comparison, although results need to be interpreted with caution. For example, we found that some parameterisations of LCM performed exceptionally well in predicting observed patterns of change, whereas SML projections did not generally perform so well. However, it should be noted that SML was explicitly designed to model deforestation at large spatial scales (Rosa et al., 2013), as opposed to simulating secluded urban parcels over relatively small spatial scales as we used in this exercise. A strict comparison of model performance in this single, limited case might tell us what modelling approach is best suited to projecting land-cover change in this type of situation, but tells us nothing about the relative strengths and weaknesses of the two approaches in other situations. Perhaps, then, these models were unfairly compared in this exercise, but our explicit objective here was to evaluate the versatility of SimiVal and not to make generalised statements about the relative performance of different modelling platforms.

Nonetheless, our worked example of applying SimiVal does demonstrate how the tool can be used to assist in the selection of the best suited parameterisation for exploratory studies and future projection in a given situation. Of course, one has to assume land-cover change processes are stationary and, that the credibility of the final landscape projection and the representation of the land-cover change processes have been considered to be appropriate. If these are reasonable assumptions, the position of each parameterisation in similarity space is useful information for trying to improve models, or during exploratory modelling when an analyst is trying to understand parameterisations that lead to specific land-cover outcomes. Using similarity space in this way is therefore useful for understanding and improving predictive models where validation has previously been considered to have limited value (Deadman et al., 2004). With our exercise, for example, we were able to use visual inspection of Fig. 2 to identify projections with systematic-bias, and on further examination of those parameterisations we found that they represented malfunctioning models that produced improbable landscape structures (e.g. one model predicted stripes of change across the landscape). Such information, along with more detailed examination of the SimiVal regression output (Fig. 5) helps guide the next round of model improvements, an approach suggested by Pontius and Millones (2011). If an analyst was trying to exactly replicate the observed land-cover map, in depth analysis of the regression plots will give clues to where a particular model parameterisation requires attention. As a specific example, projection LCM.24 had problems with allocation in the T3:1, T2:3 and T2:1 transitions, and difficulty replicating the structural dimensions, paAve and paVar. Such information could tell an analyst two things that will help improve future parameterisations: (i) that there was a problem because the model was not calibrated with these transitions; and (ii) that the model does not re-create landscape structure well and the

simulation of patch size and clustering requires attention.

The applications of SimiVal reach into many disciplines and the information that each discipline seeks is often quite different. For example, an economist or policy analyst may be interested in different outcomes to infer how a particular parameterisation influenced specific policies or market mechanisms, whereas an ecologist or spatial analyst may be interested in landscape structure for conservation studies or development planning. Whatever the purpose of the LCCM, using a similarity tool that provides a comprehensive validation, can rigorously guide a user's choice of model and their interpretation of land-cover projections.

7. Conclusion

Different parameterisations of the same land-cover model, and different land-cover modelling approaches can produce a wide range of land-change predictions from the same scenario and data. Using a limited validation that focusses solely on specific dimensions of change, such as quantity allocation, it is not always possible to distinguish how similar land-cover change projections are to other land-change projections or to the actual, observed land-cover change. More dimensions need to be considered, and in particular metrics that consider landscape structure. Moreover, it is also difficult to estimate how dissimilar particular model projections are unless they are compared to values from a known benchmark condition. These benchmark conditions require the rescaling of all metrics relative to the observed values, and the subsequent regression plots provide coefficients to test the similarity of a modelled projection. The most informative coefficients of the regression are the p-value and slope coefficient, which when plot simultaneously in a ternary graph of 'similarity space', indicate how modelled projections vary with different model parameterisations, relative to the cases of perfect, random, and systematic-bias. This approach, delivered as a program SimiVal, provides a consistent method for land-cover map validation, improving on the selective procedures widely practiced at present.

Acknowledgements

NERC Biodiversity & Ecosystem Service Sustainability (BESS) for funding the Tansley working group for land-cover modelling, grant NERC/PR100027. The Plum Island Ecosystems - Long Term Ecological Research site supplied the data via the United States National Science Foundation grant OEC-1238212. Mukesh Subedee (graduate student at Clark University), who ran and archived the LCM projections. AVB, IMDR and RME were supported by ERC grant 281986. This paper represents a contribution to Imperial College's Initiative in Grand Challenges in Ecosystems and the Environment.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.envsoft.2016.04.016>.

References

- Almeida, C.M., Gleriani, J.M., Castejon, E.F., Soares-Filho, B.S., 2008. Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *Int. J. Geogr. Inf. Sci.* 22 (9), 943–963.
- Brown, D.G., Verburg, P.H., Pontius Jr., R.G., Lange, M.D., 2013. Opportunities to improve impact, integration and evaluation of land change models. *Curr. Opin. Environ. Sustain.* 5, 452–457.
- Celio, E., Koellner, T., Grêt-Regamey, A., 2014. Modeling land use decisions with Bayesian networks: Spatially explicit analysis of driving forces on land use change. *Environ. Model. Softw.* 52, 222–233.
- Chang, F., Chen, C.-J., Lu, C.-J., 2004. A linear-time component-labeling algorithm using contour tracing technique. *Comput. Vis. Image Underst.* 93, 206–220.
- Clark Labs, 2009. The Land Change Modeler for Ecological Sustainability. IDRISI Focus Paper. Clark University, Worcester, M. <http://www.clarklabs.org/applications/uplaad/Land-Change-Modeler-IDRISI-Focus-Paper-pdf>.
- Cliff, A.D., Ord, J.K., 1973. *Spatial Autocorrelation*. Pion, London, p. 178.
- Congalton, R., Green, K., 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC/Lewis Press, Boca Raton, FL, p. 137.
- Costanza, R., 1989. Model goodness of fit: A multiple resolution procedure. *Ecol. Model.* 47, 199–215.
- Crawley, M.J., 2005. *Statistics. An Introduction Using R*. John Wiley and Sons, Chichester, England, p. 327.
- Dale, V.H., O'Neill, R.V., Southworth, F., Pedlowski, M., 1994. Modeling Effects of Land Management in the Brazilian Amazonian Settlement of Rondônia. *Con. Bio.* 8, 196–206.
- Deadman, P., Robinson, D., Moran, E., Brondizio, E., 2004. Colonist household decisionmaking and land-use change in the Amazon Rainforest: an agent-based simulation. *Environ. Plan. B Plan. Des.* 31, 693–709.
- Etter, A., McAlpine, C., Wilson, K., Phinn, S., Possingham, H., 2006. Regional patterns of agricultural land use and deforestation in Colombia. *Agric. Ecosyst. Environ.* 114, 369–386.
- Evans, T.P., Manire, A., de Castro, F., Brondizio, E., McCracken, S., 2001. A dynamic model of household decision-making and parcel level landcover change in the eastern Amazon. *Ecol. Model.* 143, 95–113.
- Filatova, T., Verburg, P.H., Parker, D.C., Stannard, C.A., 2013. Spatial agent-based models for socio-ecological systems: Challenges and prospects. *Environ. Model. Softw.* 45, 1–7.
- Fuglsang, M., Münier, B., Hansen, H.S., 2013. Modelling land-use effects of future urbanization using cellular automata: An Eastern Danish case. *Environ. Model. Softw.* 50, 1–11.
- Geoghegan, J., Schneider, L.C., Vance, C., 2004. Temporal dynamics and spatial scales: Modeling deforestation in the southern Yucatán peninsular region. *Geojournal* 61 (4), 353–363.
- Gittleman, J.L., Kot, M., 1990. Adaptation: Statistics and a null model for estimating phylogenetic effects. *Syst. Zool.* 39, 227–241.
- Haase, D., Lautenbach, S., Seppelt, R., 2010. Modeling and simulating residential mobility in a shrinking city using an agent-based approach. *Environ. Model. Softw.* 25, 1225–1240.
- Haase, D., Haase, A., Kabisch, N., Kabisch, S., Rink, D., 2012. Actors and factors in land-use simulation: The challenge of urban shrinkage. *Environ. Model. Softw.* 35, 92–103.
- Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps. *Int. J. Geogr. Inf. Sci.* 17 (3), 235–249. <http://dx.doi.org/10.1080/13658810210157822>.
- Kirkby, R., Claessens, L., Hopkinson Jr., C.S., Rastetter, E., Vallino, J.J., 2000. Modelling the effects of Land-Use Change on Nitrogen Biogeochemistry in the Ipswich Watershed, Massachusetts. *Biol. Bull.* 199, 218–219.
- Lapola, D.M., Schaldach, R., Alcamo, J., Bondeau, A., Msangi, S., Priess, J.A., Silvestrini, R., Soares-Filho, B.S., 2011. Impacts of climate change and the end of deforestation on land use in the Brazilian Legal Amazon. *Earth Interact.* 15, 1–29.
- Liao, J., Tang, L., Shao, G., Su, X., Chen, D., Xu, T., 2016. Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. *Environ. Model. Softw.* 75, 163–175.
- López, S., Sierra, R., 2010. Agricultural change in the Pastaza River Basin: A spatially explicit model of native Amazonian cultivation. *Appl. Geogr.* 30, 355–369.
- Magliocca, N.R., van Vliet, J., Brown, C., Evans, T.P., Houet, T., Messerli, P., Messina, J.P., Nicholas, K.A., Ornetsmüller, C., Sagebiel, J., Schweizer, V., Verburg, P.H., Yu, Q., 2015. From meta-studies to modeling: Using synthesis knowledge to build broadly applicable process-based land change models. *Environ. Model. Softw.* 72, 10–20.
- Mas, J.-F., Kolb, M., Paegelow, M., Olmedo, M.T.C., Houet, T., 2014. Inductive pattern-based land use/cover change models: a comparison of four software packages. *Environ. Model. Softw.* 51, 94–111.
- McGarigal, K., Marks, B.J., 1995. *Fragstats. Spatial Analysis Program for Quantifying Landscape Structure*. U.S. Department of Agriculture Forest Service.
- Messina, J.P., Walsh, S.J., 2001. 2.5D Morphogenesis: modeling landuse and land-cover dynamics in the Ecuadorian Amazon. *Plant Ecol.* 156, 75–88.
- Michalski, F., Peres, C.A., Lake, I.R., 2008. Deforestation dynamics in a fragmented region of southern Amazonia: evaluation and future scenarios. *Environ. Conserv.* 35 (2), 93–103.
- National Research Council, 2014. *Advancing Land Change Modeling. Opportunities and Research Requirements*. The National Academies Press, Washington, D.C., p. 142.
- Olmedo, M.T.C., Pontius Jr., R.G., Paegelow, M., Mas, J.-M., 2015. Comparison of simulation models in terms of quantity and allocation of land change. *Environ. Model. Softw.* 69, 214–221.
- Pérez-Vega, A., Mas, J.-F., Ligmann-Zielinska, A., 2012. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environ. Model. Softw.* 29, 11–23.
- Pijanowski, B.C., Tayyebi, A., Doucette, J., Pekin, B.K., Braun, D., Plourde, J., 2014. A big data urban growth simulation at a national scale: Configuring the GIS and neural network based Land Transformation Model to run in a High Performance Computing (HPC) environment. *Environ. Model. Softw.* 51, 250–268.
- Pontius, R.G., 2000. Quantification error versus location error in comparison to categorical maps. *Photogramm. Eng. Remote Sens.* 66 (8), 1011–1016.

- Pontius, R.G., 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogramm. Eng. Remote Sens.* 68 (10), 1041–1049.
- Pontius, R.G., Cheuk, M.L., 2006. A generalised cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *Int. J. Geogr. Inf. Sci.* 20 (1), 1–30.
- Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* 32, 4407–4429.
- Pontius, R.G., Pacheco, P., 2004. Calibration and validation of a model of forest disturbance in the Western Ghats, India 1920–1990. *Geojournal* 61, 325–334.
- Pontius, R.G., Parmentier, B., 2014. Recommendations for using the relative operating characteristic (ROC). *Landsc. Ecol.* 29, 367–382.
- Pontius, R.G., Schneider, L.C., 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agric. Ecosyst. Environ.* 85, 239–248.
- Pontius, R.G., Walker, R., Yao-Kumah, R., Arima, E., Aldrich, S., Caldas, M., Vergara, D., 2007. Accuracy Assessment for a Simulation Model of Amazonian Deforestation. *Ann. Assoc. Am. Geogr.* 97 (4), 677–695.
- Pontius, R.G., Boersma, W., Castella, J.-C., Clarke, K., de Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstien, N., Kok, K., Koomen, E., Lippit, C.D., McConnell, W., Sood, A.M., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T.N., Veldkamp, A.T., Verburg, P.H., 2008. Comparing the input, output, and validation maps for several models of land change. *Ann. Region. Sci.* 42, 11–37.
- Pontius Jr., R.G., Peethambaram, S., Castella, J.-M., 2011. Comparison of three maps at multiple resolutions: A case study of land change simulation in Cho Don district, Vietnam. *Ann. Assoc. Am. Geogr.* 101 (1), 45–62.
- Ralha, C.G., Abreu, C.G., Coelho, C.G.C., Zaghetto, A., Macchiavello, B., Machado, R.B., 2013. A multi-agent model system for land-use change simulation. *Environ. Model. Softw.* 42, 30–46.
- Riitters, K.H., O'Neill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B., Jackson, B.L., 1995. A factor analysis of landscape pattern and structure metrics. *Landsc. Ecol.* 10, 23–39.
- Rosa, I., Purves, D., Souza Jr., C., Ewers, R.M., 2013. Predictive modelling of contagious deforestation in the Brazilian Amazon. *PLOS One* 8 (10), e77231. <http://dx.doi.org/10.1371/journal.pone.0077231>.
- Rosa, I.M.D., Ahmed, S.E., Ewers, R.M., 2014. The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. *Glob. Change Biol.* 20 (6), 1707–1722.
- Sangermano, F., Toledano, J., Eastman, R.J., 2012. Land cover change in the Bolivian Amazon and its implications for REDD+ and endemic biodiversity. *Landsc. Ecol.* 27, 571–584. <http://dx.doi.org/10.1007/s10980-012-9710-y>.
- Silvestrini, R.A., Soares-Filho, B.S., Nepstad, D., Coe, M., Rodrigues, H., Assunção, R., 2011. Simulating fire regimes in the Amazon in response to climate change and deforestation. *Ecol. Appl.* 21, 1573–1590.
- Soares-Filho, B.S., Cerqueira, G.C., Pennachin, C.L., 2002. DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecol. Model.* 154, 217–235.
- Soares-Filho, B., Rodrigues, H., Follador, M., 2013. A hybrid analytical-heuristic method for calibrating land-use change models. *Environ. Model. Softw.* 43, 80–87.
- Sun, Z., Müller, D., 2013. A framework for modelling payments for ecosystem services with agent-based models, Bayesian belief networks and opinion dynamics models. *Environ. Model. Softw.* 45, 15–28.
- Tayyebi, A., Pijanowski, B.C., Linderman, M., Gratton, C., 2014. Comparing three global parametric and local non-parametric models to simulate land use change in diverse areas of the world. *Environ. Model. Softw.* 59, 202–221.
- Vance, C., Iovanna, R., 2008. Cities and Satellites: Spatial Effects on Unobserved Heterogeneity in Modelling Urban Growth. *Transp. Res. Rec.* 27, 87–96.
- Verburg, P.H., Groot, W.T., Veldkamp, A.J., 2003. Methodology for multi-scale land-use change modelling: concepts and challenges. In: Dolman, A.J., et al. (Eds.), *Global Environmental Change and Land Use*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 17–51.
- Verburg, P.H., Schot, P.P., Dijst, M.J., Veldkamp, A., 2004. Land use change modelling: current practice and research priorities. *Geojournal* 61, 309–324.
- Verstegen, J.A., Karssenbergh, D., van der Hilst, F., Faaij, A.P.C., 2014. Identifying a land use change cellular automaton by Bayesian data assimilation. *Environ. Model. Softw.* 53, 121–136.
- Verstegen, J.A., Karssenbergh, D., van der Hilst, F., Faaij, A.P.C., 2016. Detecting systemic change in a land use system by Bayesian data assimilation. *Environ. Model. Softw.* 75, 424–438.
- Walker, R., Drzyzga, S.A., Li, Y., Qi, J., Caldas, M., Arima, E., Vergara, D., 2004. A behavioral model of landscape change in the Amazon Basin: the colonist case. *Ecol. Appl.* 14, S299–S312.
- Walsh, S.J., Messina, J.P., Mena, C.F., Malanson, G.P., Page, P.H., 2008. Complexity theory, spatial simulation models, and land use dynamics in the Northern Ecuadorian Amazon. *Geoforum* 39, 867–878.
- Wassenaar, T., Gerber, P., Verburg, P.H., Rosales, M., Ibrahim, M., Steinfeld, H., 2007. Projecting land use changes in the Neotropics: the geography of pasture expansion into forest. *Glob. Environ. Change – Hum. Policy Dimens.* 17, 86–104.