Computers, Environment and Urban Systems 77 (2019) 101366

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Validating land change models based on configuration disagreement

Brian R. Pickard[a,*], Ross K. Meentemeyer[b]

[a] *Tetra Tech, Raleigh, NC, USA*
[b] *Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA*

ARTICLE INFO

Keywords:
Model validation
Accuracy assessment
FUTURES
Landscape heterogeneity
Simulation

ABSTRACT

Land change models are increasingly being employed to predict future landscapes and influence policy and decision-making. To ensure the highest model accuracy, validation methods have become commonplace following a land change simulation. The most common validation method employed uses quantity and allocation disagreement. However, these current measures may not account for differences in the configurations of land change, placing them in potential conflict with the principals of heterogeneity and spatial patterning of landscape ecology. We develop a new metric, termed configuration disagreement, designed to focus on the size, shape, and complexity of land change simulations. Using this metric, we demonstrate the value of including errors of configuration disagreement – in addition to quantity and allocation error – in the assessment of land change models. Four computational experiments of land change that vary only in spatial pattern are developed using the FUTURES land change model. For each experiment, configuration disagreement and the traditional validation metrics are computed simultaneously. Results indicate that models validated only with consideration of quantity and allocation error may misrepresent, or not fully account for, spatial patterns of landscape change. The research objective will ultimately guide which component, or components, of model disagreement are most critical for consideration. Yet, our work reveals why it may be more helpful to validate simulations in terms of configuration accuracy. Specifically, if a study requires accurately modeling the spatial patterns and arrangements of land cover. Configuration disagreement could add critical information with respect to a model's simulated changes in size, shape, and spatial arrangements, and possibly enhance ecologically meaningful land change science.

1. Introduction

Models of land use and land change can be powerful tools for simulating future patterns of landscape change and guiding decision making (Eigenbrod et al., 2011; Nedkov & Burkhard, 2012; Nelson et al., 2009; Pickard, Gray, & Meentemeyer, 2017; Renard, Rhemtulla, & Bennett, 2015; Tayyebi, Pijanowski, & Pekin, 2015). Computational in nature, they enable experiments at varying spatial scales to investigate how land cover can change under differing conditions and the ecological implications of such alterations (van Vliet et al., 2016). Land change models can incorporate social, environmental, institutional, and economic processes, thereby creating a wide variety of methodological approaches. Land change models use statistical correlations (inductive), explicitly describe processes (deductive), infer underlying processes from observed patterns of land change (patter-based), or simulate individual decision makers (agent-based) (Mas, Kolb, Paegelow, Camacho-Olmedo, & Houet, 2014). Examples include CLUE (Verburg & Overmars, 2009), SLEUTH (Clark, Hoppen, & Gaydos, 1997), Metronamica (White, Engelen, & Uljee, 1997), FUTURES (Meentemeyer et al., 2013), IDRISI's suite of tools (LCM, GEOMOD, CA_MARKOV), and DINAMICA, with new models constantly being developed. Given their variety and prevalence, accurate, ecologically-relevant simulations are needed to advance meaningful land change science and to provide a better basis for decision making and policy formulation that relies on the use of such models.

Consensus regarding model accuracy assessments focuses on quantifying two specific types of accuracy: *quantity* and *allocation* (Chen, Li, & Ai, 2014; Pontius Jr & Millones, 2011). Quantity disagreement is the difference between observed and simulated maps attributable to the difference in proportions of map categories (e.g., simulating too much or too little change). Allocation disagreement is the differences between observed and simulated maps attributed to differences in matching spatial allocation of categories (e.g., simulating change at a location where no change was observed). While these methods provide useful information of model accuracy, they do not explicitly address the degree to which simulations match the spatial arrangement of observed

* Corresponding author.
E-mail address: brian.pickard@tetratech.com (B.R. Pickard).

https://doi.org/10.1016/j.compenvurbsys.2019.101366
Received 3 December 2018; Received in revised form 8 July 2019; Accepted 9 July 2019
Available online 18 July 2019

landscape patterns. Landscape ecology focuses on the ecological effects of spatial patterning (Forman, 1995; Turner, 1989) and particularly important in a wide range of landscape scale science (e.g. habitat connectivity, surface water flow, urban planning). Therefore, in the context of validating land change models employed in ecology-based work, greater emphasis on the spatial arrangement of patches is needed, with more focus on configuration and less on correctness of single pixels. For example, a simulation that places the same number of change pixels into a single large patch can have identical allocation disagreement compared to a simulation that has several fragmented patches (Pickard, Van Berkel, Petrasova, & Meentemeyer, 2017). This implies that current validation methods may be misleading land change scientists as to how accurate their models really are.

Pontius Jr et al. (2018) show that the quantity of simulated change can confound the accuracy results. Therefore, to understand model accuracy, it is important to simulate the correct number of pixels that change. Pontius Jr et al. (2018) demonstrated that models that predict less change than observed may have higher accuracy than those that predict more change. When the exact number of simulated change pixels is achieved (i.e. no quantity error) a model's behavior can be evaluated with respect to allocation (or configuration). However, no research to date has modeled multiple simulations with zero quantity error, thereby varying only in spatial arrangement of simulated change in order to assess configuration disagreement.

Using a simple example (Fig. 1), we demonstrate the added context that configuration accuracy might play in the validation of a land change model. Consider a forested landscape experiencing urbanization, with each map representing a simulated example of where new development is expected to occur. Each map consists of thirty-six pixels, with each pixel being classified as forest (green), existing urban (dark gray), or projected new development (white). We also provide a reference map which illustrates the "observed" development from which the accuracy of each simulation is evaluated (Fig. 1, box with dotted box surrounding it). Although each of the maps has the same number of forest pixels simulated as changing from forest to developed (i.e., four pixels in the white category), they differ in the specific configurations of where changes occur. Using the same number of change pixels in each scenario eliminates quantity error in these computational experiments. Based on the current validation method of allocation disagreement (Pontius Jr & Millones, 2011), each map is equally accurate. However, most readers will clearly note varying accuracy in how well the simulations replicate the observed configuration. The ecological implications of these configurations are likely to vary considerably depending on which configuration is considered. This inaccuracy highlights a critical methodological gap in the assessment of a model's overall accuracy.

In this paper we propose a metric for assessing accuracy of land change models, called configuration disagreement, designed to be complimentary to the existing validation methods of quantity and allocation. We demonstrate its need using a spatially-explicit land change model (FUTURES; Meentemeyer et al., 2013) capable of simulating variation in spatial configuration. Using a rapidly urbanizing region, Charlotte, North Carolina and the surrounding nine counties, we present a case study analyzing computation experiments of varying configurations of urbanization. As with the simplified examples in Fig. 1, each computational experiment differs only in the spatial arrangement of simulated change pixels the number of change pixels are held constant across all experiments. By calculating both allocation and configuration disagreement simultaneously for each experiment, we demonstrate the need for considering configuration to more holistically understand model accuracy. Our results illustrate that allocation disagreement does not indicate configuration and can lead to spurious results if the purpose of the model relies on correctly predicting observed configurations. Our proposed metric of configuration disagreement likely provides a better metric for specific instances concerned with configuration.

2. Methods

2.1. Computation experiments

Four computational experiments of differing configurations of urban development were designed and compared to observed urban development. Using the observational dataset, the number of pixels that were converted from undeveloped to developed were identified annually and used as inputs for each configuration experiment. By creating multiple experiments depicting different configurations of newly assigned development pixels, and ensuring the same number of observed change pixels were allocated in each experiment, we can directly compare validation methods. Allocation and configuration disagreement were calculated for each experiment and compared to the observational dataset to identify which computational experiment had

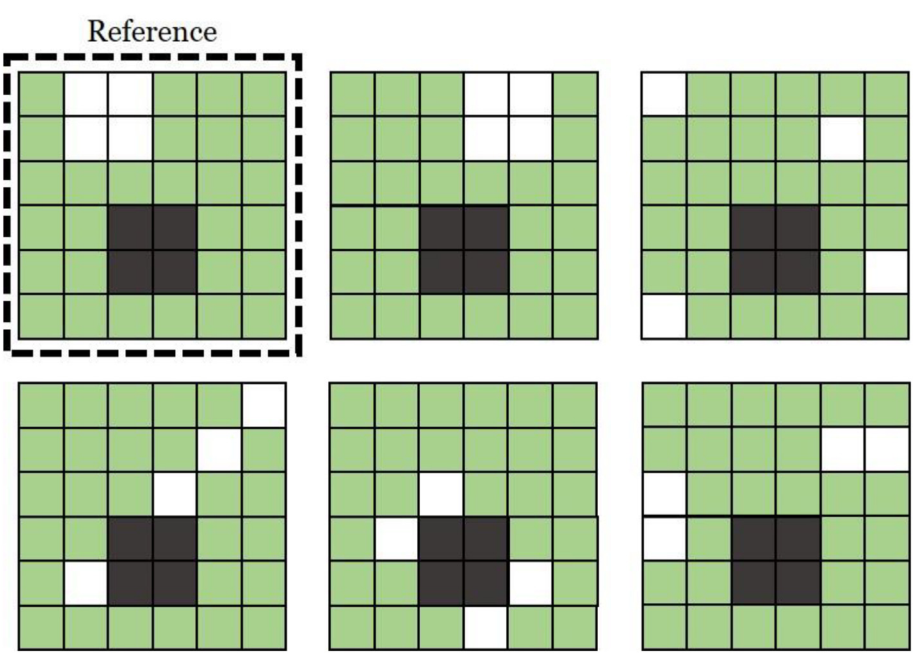

Fig. 1. Reference (dotted box) and comparison maps showing multiple configurations with identical quantity and allocation disagreement values. The white boxes represent simulated change pixels. Each of the five configurations have zero quantity disagreement and 22% allocation disagreement. Using current validation methods (Pontius Jr & Millones, 2011) each map is considered identical in accuracy compared to the reference map.

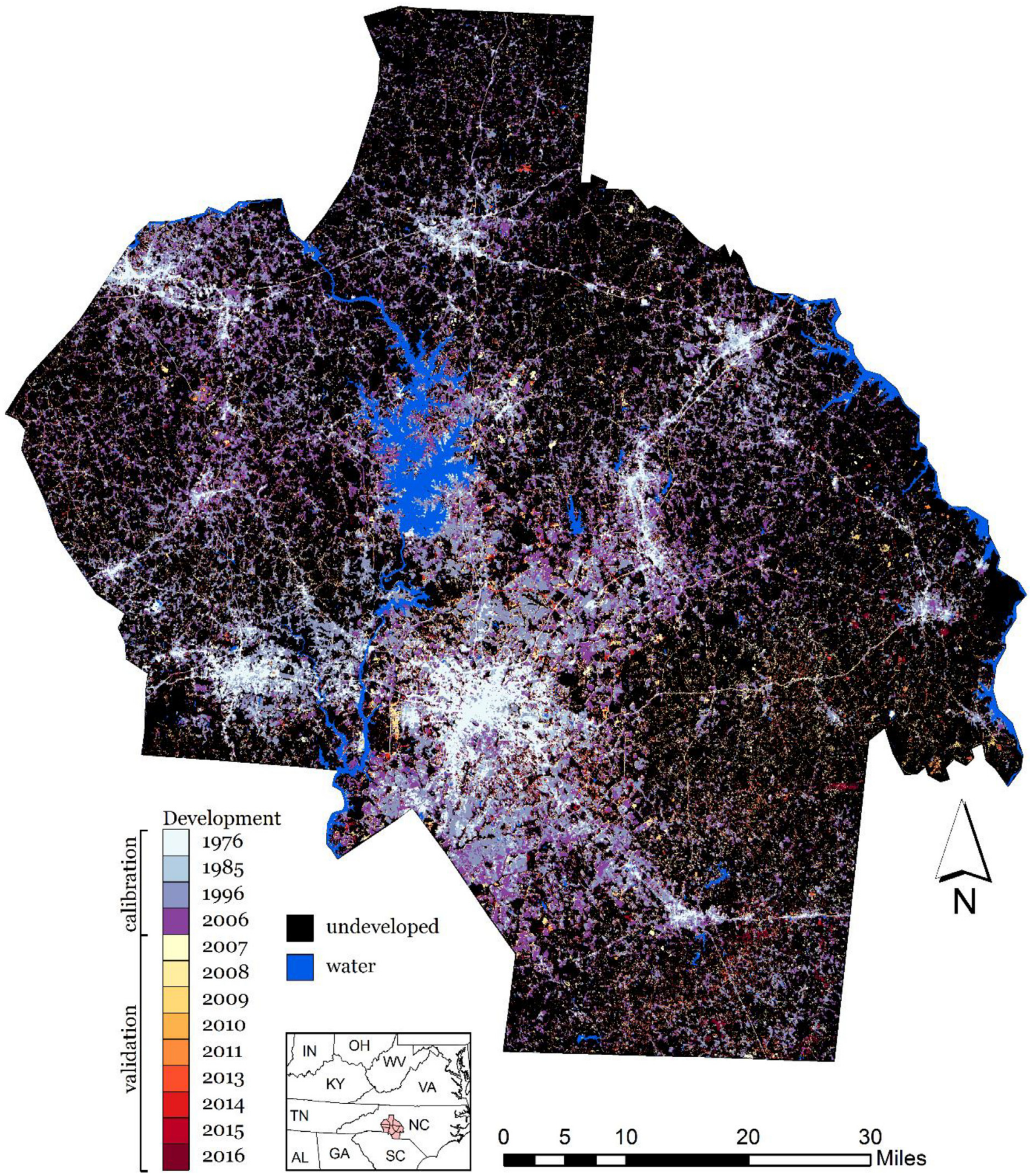

Fig. 2. Nine county study extent depicting developed lands observed during calibration (1976, 1985, 1996, and 2006) and validation (2007–2016) phases.

the lowest levels of disagreement. A total of nine years (2006–2016, excluding 2012) were used to validate the modeled results to account for the stochastic, human-ecological interactions that shape the processes driving land change (Olmedo, Pontius Jr, Paegelox, & Mas, 2015). Therefore, validation results can be evaluated while accounting for a land change models' difficulty in capturing these complex interactions (Kolb, Mas, & Galicia, 2013; Olmedo et al., 2015; Perez-Vega, Mas, & Lingmann-Zielinska, 2012).

2.2. Study extent

The study extent is located within the Piedmont physiographic province of Central North Carolina (Fig. 2), and is a key urban center within the "Charlanta" megaregion (Florida, Gulden, & Mellander, 2008). We selected nine counties within and surrounding North Carolina's most urbanized portion of the state, intersecting three rapidly expanding metropolitan areas that are expected to double in population by 2030 (North Carolina State Demographic Office, 2017). Development, thus far, has manifested as a pattern of urban sprawl typical of much of the United States in the late 20th and early 21st century (Terando et al., 2014).

2.3. Input data

We used satellite imagery from Landsat MultiSpectral Scanner (MSS; Landsat 4), Thematic Mapper (TM; Landsat 5), and Operational

Table 1
Accuracy results based on Olofsson et al. (2013) of Landsat imagery classification using Vegetation-Impervious Surface-Soil (VIS) method.

	Undeveloped	Developed	Total	W_i	User's	Producer's	Overall
2016 – (OLI)							
Undeveloped	227	23	250	0.647	0.908	0.998	0.939
Developed	1	249	250	0.354	0.996	0.856	
Total	228	272	500	1.000			
2015 – (OLI)							
Undeveloped	225	25	250	0.650	0.900	0.993	0.931
Developed	3	247	250	0.349	0.988	0.841	
Total	228	272	500	1.000			
2014 – (OLI)							
Undeveloped	206	44	250	0.657	0.824	0.997	0.883
Developed	1	249	250	0.343	0.996	0.747	
Total	207	293	500	1.000			
2013 – (OLI)							
Undeveloped	214	36	250	0.664	0.856	0.998	0.903
Developed	1	249	250	0.335	0.996	0.777	
Total	215	285	500	1.000			
2011 – (TM)							
Undeveloped	214	36	250	0.676	0.856	0.996	0.900
Developed	2	248	250	0.324	0.992	0.768	
Total	216	284	500	1.000			
2010 – (TM)							
Undeveloped	216	34	250	0.682	0.864	1.000	0.907
Developed	0	250	250	0.318	1.000	0.774	
Total	216	284	500	1.000			
2009 – (TM)							
Undeveloped	215	35	250	0.685	0.860	0.998	0.903
Developed	1	249	250	0.315	0.996	0.761	
Total	216	284	500	1.000			
2008 – (TM)							
Undeveloped	221	29	250	0.701	0.884	0.998	0.918
Developed	1	249	250	0.299	0.996	0.786	
Total	222	278	500	1.000			
2007 – (TM)							
Undeveloped	219	31	250	0.715	0.876	0.995	0.908
Developed	3	247	250	0.285	0.988	0.761	
Total	222	278	500	1.000			
2006 – (TM)							
Undeveloped	225	225	250	0.732	0.900	0.982	0.915
Developed	11	11	250	0.268	0.956	0.778	
Total	236	236	500	1.000			
1996 – (TM)							
Undeveloped	237	237	250	0.828	0.948	0.988	0.947
Developed	14	14	250	0.172	0.944	0.791	
Total	251	251	500	1.000			
1985 – (TM)							
Undeveloped	242	242	250	0.947	0.968	0.986	0.956
Developed	63	63	250	0.053	0.748	0.565	
Total	305	305	500	1.000			
1976 – (MSS)							
Undeveloped	218	218	250	0.967	0.872	0.999	0.875
Developed	9	9	250	0.033	0.964	0.204	
Total	227	227	500	1.000			

Land Manager (OLI; Landsat 8), to classify the amount of development that occurred within the region from 1976 to 2016 (Table 1). Land cover data was divided into two datasets: 1) a calibration phase consisting of four time points (1976, 1985, 1996, 2006), and 2) a validation phase consisting of annual maps (2007–2016). Land cover for the year 2012 was omitted from the validation phase because a suitable set of images free of scan line corrector (SLC, Arvidson, Goward, Gasch, & Williams, 2006) errors with cloud cover < 20% was not available. Land change models require at a minimum two dates for calibration (e.g. Verburg & Overmars, 2009), yet recent trends in research suggest calibrating with multiple time points provides greater information regarding temporal trends in land change (e.g. Clark et al., 1997; Meentemeyer et al., 2013; Terando et al., 2014).

For each year, we produced fractional components representing the vegetation, impervious and soil components of each Landsat image pixel using vegetation-impervious surface-soil (VIS) classification and unconstrained linear spectral unmixing analysis (Ridd, 1995; Wu, 2004). We then selected training sites to represent pure end members for each VIS component using aerial orthophotography. We next classified each pixel within the map as developed or undeveloped by employing logistic regression comparing the fraction images generated by the spectral unmixing analysis and the interpretations of 1000 points in the current year's orthophotography. Sing only two classes reduces the likelihood of error when compared to multi-class image classification

(Pontius Jr & Maliza, 2004). Each final classified map was designed to allow for both gross gain and loss of urban, however in this case study there was little to no loss of urban in any specific year.

Several methods have been proposed to report error found in observed maps (Olofsson et al., 2014; Olofsson, Foody, Stehman, & Woodcock, 2013; Pontius Jr & Li, 2010; Pontius Jr & Lippitt, 2006), and images classified in this analysis followed the "good practice recommendations" of Olofsson et al. (2014). Each land cover dataset was evaluated against 500 stratified random sample points (250 each for developed and undeveloped) using concurrent high-resolution orthophotography. Samples for assessing classification error for each year were selected separately from the other samples. Error matrices were quantified for each year, reporting raw point counts, estimated area proportions, and user's, producers, and overall accuracy.

2.4. FUTURES land change model

The FUTure Urban-Regional Environmen Simulation model (FUTURES; Meentemeyer et al., 2013) is a fully open source, multilevel, spatio-temporal land change model that is run through a suite of modules in GRASS GIS. Here we focus on conversion of forest and farmland to impervious development. Processes of landscape change are represented through three sub-models: 1) specification of the quantity, or amount, of new development (DEMAND sub-model), 2) location of development (POTENTIAL sub-model) based on local site suitability factors, and 3) the spatial pattern of development simulated by a stochastic patch growing algorithm (PGA sub-model).

The DEMAND sub-model of FUTURES requires the user to manually specify the number of new development pixels to be allocated each year. Using the observed annual land cover datasets generated for 2007–2016, we identified the specific number of pixels developed each year and used these pixels as inputs for the FUTURES model. Specifying the exact same amount of new development eliminates the quantity disagreement from each computational experiment. To determine the locations of where new development will be sited, FUTURES requires the input of a site suitability surface (e.g., Dorning, Kock, Shoemaker, & Meentemeyer, 2013; Meentemeyer et al., 2013; Pickard, Gray, & Meentemeyer, 2017). We used a linear mixed-effects model to determine the relationship between natural lands converted to developed based on environmental conditions (Meentemeyer et al., 2013) to create a site suitability surface. We then selected a set of key indicators explaining locations of urban growth (Table 2) using principal component analysis. Model parameters were determined based on the binary response (land converted or not) for approximately 20,000 randomly sampled points. We accounted for variability among counties by assigning a random effect to the intercept and development pressure variable. These random effects account for unexplained development factors that likely influence development but were not included in the model, such as zoning constraints.

To simulate new development experiments, the FUTURES model uses the specified quantity of new development with the site suitability surface to project development based on an iterative, stochastic site selected process and a patch-based region growing algorithm designed to mimic distinct spatial structures. The PGA stochastically selects a location for development across the site suitability surface. An urban patch is successfully developed in the model if the chosen location survives a randomized (i.e., Monte Carlo) challenge. Locations that survive this challenge spread into discrete patches based on distributions of patch sizes and shapes derived from observed development patterns. The PGA stochastically selects a patch within the library and allocates it to the successful location. When the total number of new cells for a computational experiment year are allocated (determined by user specified quantity), the development pressure variable is updated, and the site suitability surface is recalculated. This process is repeated for each experiment until the final year of development occurs.

Computational experiments of development configuration can be explored in FUTURES using the PGA incentive parameter. This parameter applies a power transformation to the site suitability surface, changing the distribution and spatial configuration of new patches across the landscape (Fig. 3). By varying this parameter, we generated multiple computational experiments ranging from extreme sprawl (Experiment 1) to compact infilling (Experiment 4). In each of these, all model parameters and amount of new development are constant, with the only variation coming from the spatial arrangement of the new patches.

2.5. Quantity and allocation disagreement

We assessed current validation methods as prescribed by Pontius Jr and Millones (2011) for each of the FUTURES experiments. Quantity disagreement is the difference between observed and simulated maps attributable to the difference in proportions of categories (Pontius Jr & Millones, 2011), and is calculated as:

Table 2
Coefficients from the multilevel mixed effects model used to develop the site suitability surface within FUTURES (Meentemeyer et al., 2013).

Fixed effects	Estimate	SE	*P* value
Intercept[a]	−1.95	0.65	0.002
Distance to interchange	0.22	0.03	< 0.001
Distance to roads	−0.21	0.03	< 0.001
Slope	−0.03	0.01	< 0.001

Random effects

County	Intercept	Development pressure
Cabarrus	−4.3	0.11
Catawba	−4.01	0.10
Lincoln	−4.36	0.11
Rowan	−4.39	0.12
Iredell	−4.08	0.11
Stanley	−4.71	0.12
Gaston	−4.21	0.09
Mecklenburg	−4.14	0.09
Union	−4.38	0.11

[a] Varies by county.

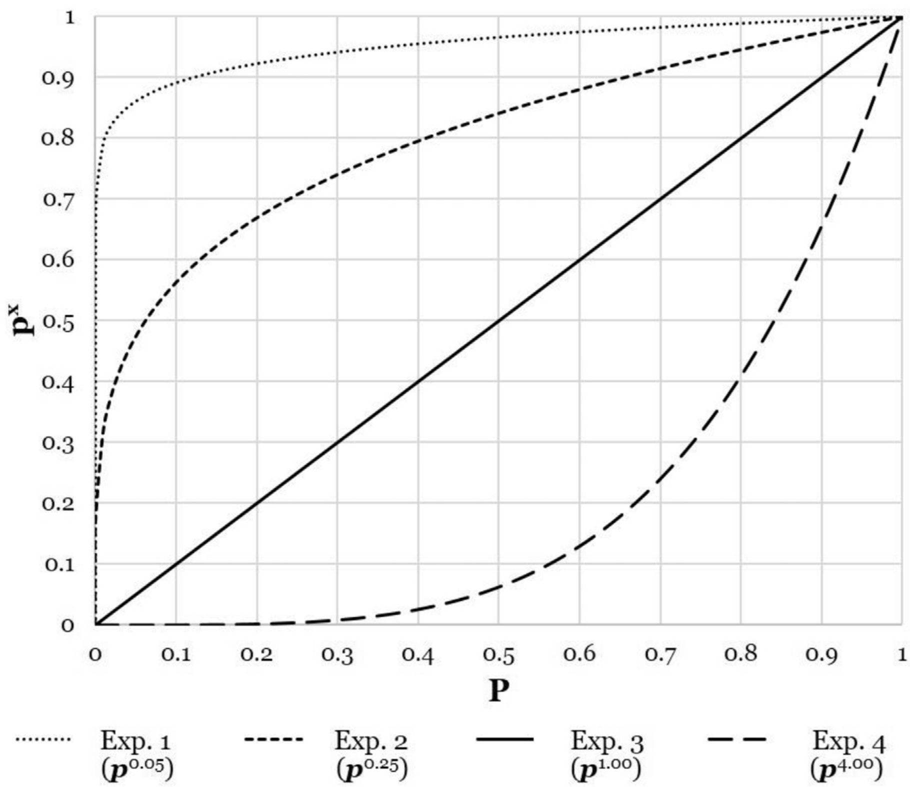

Fig. 3. Range of power functions (INCENTIVE parameter) for transforming the site suitability surface (P). Four scenarios were developed to test multiple configuration scenarios. Curves were adapted from Meentemeyer et al. (2013).

$$q_g = \left| \left(\sum_{i=1}^{J} p_{ig} \right) - \left(\sum_{j=1}^{J} p_{gj} \right) \right| \tag{1}$$

where, p_{ig} and p_{gj} represent the estimated proportion of class g in the simulated and reference maps, respectively. For the purposes of this analysis, we designed all the experiments such that quantity disagreement is equal to zero. Allocation disagreement is the difference between observed and simulated maps attributed to the differences in matching spatial allocation of categories (Pontius Jr & Millones, 2011), computed as:

$$a_g = 2\min\left[\left(\sum_{i=1}^{J} p_{ig} \right) - p_{gg}, \left(\sum_{j=1}^{J} p_{gj} \right) - p_{gg} \right] \tag{2}$$

where, the first argument within the minimum function is the omission of class g and the second argument is the commission of class g. To determine the total disagreement between the simulated and observed map, a user simply sums the quantity and allocation disagreements.

2.6. Configuration disagreement

In addition to current accuracy measures we applied a validation metric meant to be complimentary, termed configuration disagreement. Initial developments of configuration disagreement (Chen et al., 2014; Pickard, Van Berkel, et al., 2017) have relied on quantifying spatial metrics using the open-source FRAGSTATS package (McGarigal, Cushman, Neel, & Ene, 2002). Previously, configuration disagreement has been determined by comparing four metrics of the simulated development class with observed development class maps, yet the quantity of newly simulated pixels differed (Pickard, Gray, & Meentemeyer, 2017). Here, we improve upon this method by modifying the number and selection of specific FRAGSTATS metrics to be evaluated across the landscape and by eliminating quantity disagreement from these computational experiments. This allows for focusing only on how well a model matches the size, shape, complexity, aggregation, dispersion, patch variability, fragmentation, proximity to other classes, edge effects and perimeter-area ratios of observed new development. Seven class-level metrics were identified base don the results of Cushman, McGarigal, and Neel (2008) (Table 2). These seven metrics (Table 2) were incorporated into a final index value, configuration disagreement, calculated as:

$$GYRATE_{AM}, FRAC_{AM}, CORE0_{AM}, ENN_{AM}, NP = \frac{|Sim - Obs|}{Obs} x100 \tag{3}$$

$$ENN_{CV}, ECON_{AM} = |Sim - Obs| \tag{4}$$

$$C_{dis} = \frac{1}{7} \sum GYRATE_{AM}, FRAC_{AM}, CORE_{AM}, ENN_{AM}, ECON_{AM}, NP, ENN_{CV} \tag{5}$$

where, $GYRATE_{AM}$, $FRAC_{AM}$, $CORE_{AM}$, ENN_{AM} are area-weighted means of the patch parameters gyration, fractal dimension index, core area, and Euclidean nearest neighbor, respectively (see Section 2.7.1); NP is the number of patches (Section 2.7.3); ENN_{CV} and $ECON_{CV}$ are the coefficients of variation for Euclidean nearest neighbor and edge contrast; Sim and Obs are metric values for the simulated and observed classes, respectively (Table 3).

2.7. FRAGSTATS metrics for simulated and observed maps

2.7.1. Patch-based metrics

Within FRAGSTATS multiple metrics are calculated for both the observed and simulated maps. The following equations listed below are used to quantify specific FRAGSTATS metrics. Eqs. (6)–(10) describe how to calculate patch-based metrics to be used in Eq. (11) for calculating area weighted means (AM). Eq. (12) describes how the number of patches, or NP, is calculated. Lastly, Eqs. (13)–(15) detail how to

Table 3
Components of configuration, description and FRAGSTAT metrics used to calculate configuration disagreement. Specific FRAGSTAT metrics were selected based on the analysis of Cushman et al. (2008).

Component Name	Description	FRAGSTAT metric
Edge Contrast	Degree of contrast between the focal class and its neighborhood, where contrast represents the magnitude of difference between classes.	GYRATE_AM
Patch shape complexity	Shape complexity of patches, where shape is defined by perimeter-area relationships.	FRAC_AM
Aggregation	Degree of aggregation of cells in the class, where large, compact clusters of cells are considered to be aggregated.	CORE_AM
Nearest neighbor distance	Proximity of patches of the focal class, based on the area-weighted average distance between neighbors.	ENN_AM
Patch dispersion	Spatial dispersion of patches, reflecting whether patches tend to be uniformly distributed or over-dispersed.	ECON_AM
Large patch dominance	Degree of concentration of focal class area in a few, large patches with large core areas.	ENN_CV
Neighborhood similarity	Degree of isolation of patches from nearby patches of the same or similar class.	# of Patches (NP)

calculate the coefficient of variation for the Euclidean nearest neighbor metric (ENN).

The radius of gyration, or GYRATE, equals the mean distance (m) between each cell in the patch and the patch centroid. It is a measure of patch extent and is affected by both patch size and patch compaction. GYRATE is calculated as follows:

$$GYRATE = \sum_{r=1}^{z} \frac{h_{ijr}}{z} \tag{6}$$

where, h_{ijr} equals the distance (m) between cell ijr, located within patch ij, and the centroid of patch ij (the average location), based on the cell center-to-cell distance; z equals the number of cells in patch ij. The fractal dimension index (FRAC) accounts for shape complexity across a range of spatial scales and patch sizes. It is computed as:

$$FRAC = \frac{2\ln(0.25p_{ij})}{\ln(a_{ij})} \tag{7}$$

where, p_{ij} equals the perimeter of patch ij; a_{ij} equals the area of patch ij. Core area is defined as the area within a patch beyond a specified edge distance. CORE is computed as:

$$CORE = \frac{a_{ij}^{c}}{10{,}000} \tag{8}$$

where, a_{ij} equals the core area in meter of patch ij that is further than the specified c depth-of-edge distance from the patch perimeter. The edge contrast index (ECON) is a relative measure of the amount of contrast along a patch perimeter. It is computed as:

$$ECON = \frac{\sum_{k=1}^{m} (p_{ijk} \text{ x } d_{ik})}{p_{ij}} \text{ x } 100 \tag{9}$$

where, $p_{ijk}$is the length of edge of patch ij adjacent to patch type (class) k; d_{ik} is the dissimilarity (edge contrast weight) between patch types i an k; p_{ij} equals the length of perimeter of patch ij. The Euclidean nearest neighbor distance (ENN) has been used extensively to understand patch isolation. It is defined as the shortest straight-line distance between the focal patch and its nearest neighbor of the same class, it is computed as:

$$ENN = h_{ij} \tag{10}$$

where, h_{ij} is the distance in meters from patch ij to nearest neighboring patch of the same type (class), based on patch edge-to-edge distance, computed from cell center to cell center.

2.7.2. Area-weighted mean calculations

Prior to use in the configuration metric (Eq. (3)), patch metrics (Eqs. (4)–(8)) require an area-weighted mean calculated as:

$$GYRATE_{AM}, FRAC_{AM}, CORE_{AM}, ECON_{AM} ENN_{AM} = \sum_{j=1}^{n} \left[X_{ij} \left(\frac{a_{ij}}{\sum_{j=1}^{n} a_{ij}} \right) \right] \tag{11}$$

where, AM equals the sum, across all patches of the corresponding patch type n, of the corresponding metric value X_{ij} multiplied by the proportional abundance of the patch a_{ij} divided by the sum of the patch areas.

2.7.3. Number of patches

Number of patches is a simple metric that quantifies the extent of subdivision of fragmentation of the patch type. It equals the number of patches of a particular patch type and is computed as:

$$NP = n_i \tag{12}$$

where, n_i is the number of patches in the landscape of patch type (class) i.

2.7.4. Euclidean nearest neighbor of coefficient of variation

Euclidean nearest neighbor equals the distance in meters to the nearest neighboring patch of the same type, based on shortest edge-to-edge distance. The coefficient of variation is quantified as the standard deviation of ENN divided by the mean of ENN, multiplied by 100 to convert to a percentage. The following equations are used to compute ENN_{CV}:

$$ENN_{SD} = \sqrt{\frac{\sum_{j=1}^{n} \left[h_{ij} - \left(\frac{\sum_{j=1}^{n} h_{ij}}{n_i} \right) \right]}{n_i}} \tag{13}$$

$$ENN_{MN} = \frac{\sum_{j=1}^{n} h_{ij}}{n_i} \tag{14}$$

$$ENN_{CV} = \frac{ENN_{SD}}{ENN_{MN}} \text{ x } 100 \tag{15}$$

where, h_{ij} is the distance in meters from patch ij to nearest neighboring patch of the same type (class), based on patch edge-to-edge distance, computed from cell center to cell center; n_i is the number of patches in the landscape of patch type (class) i.

3. Results

No quantity disagreement was identified for each experiment, consistent with the design of this study. Each computational experiment had the exact same number of new development pixels distributed across the study area for each year. Fig. 4 provides a zoomed example for part of the ten-county study area showing the varying spatial configuration of each experiment. Therefore, we were able to confirm that each experiment differed only in the arrangement of pixels within the study area.

Allocation disagreement for each experiment is summarized in Fig. 5, demonstrating the amount of disagreement using the current validation methods. Allocation disagreement was the lowest for Experiment 4, with disagreement reaching 12.7% by 2016. Experiment 3 had allocation disagreement values reaching 13.3% after ten years. Experiments 1 and 2 had allocation disagreement values of 14.8 and 14.4% over the same time period. Using current validation methods, we

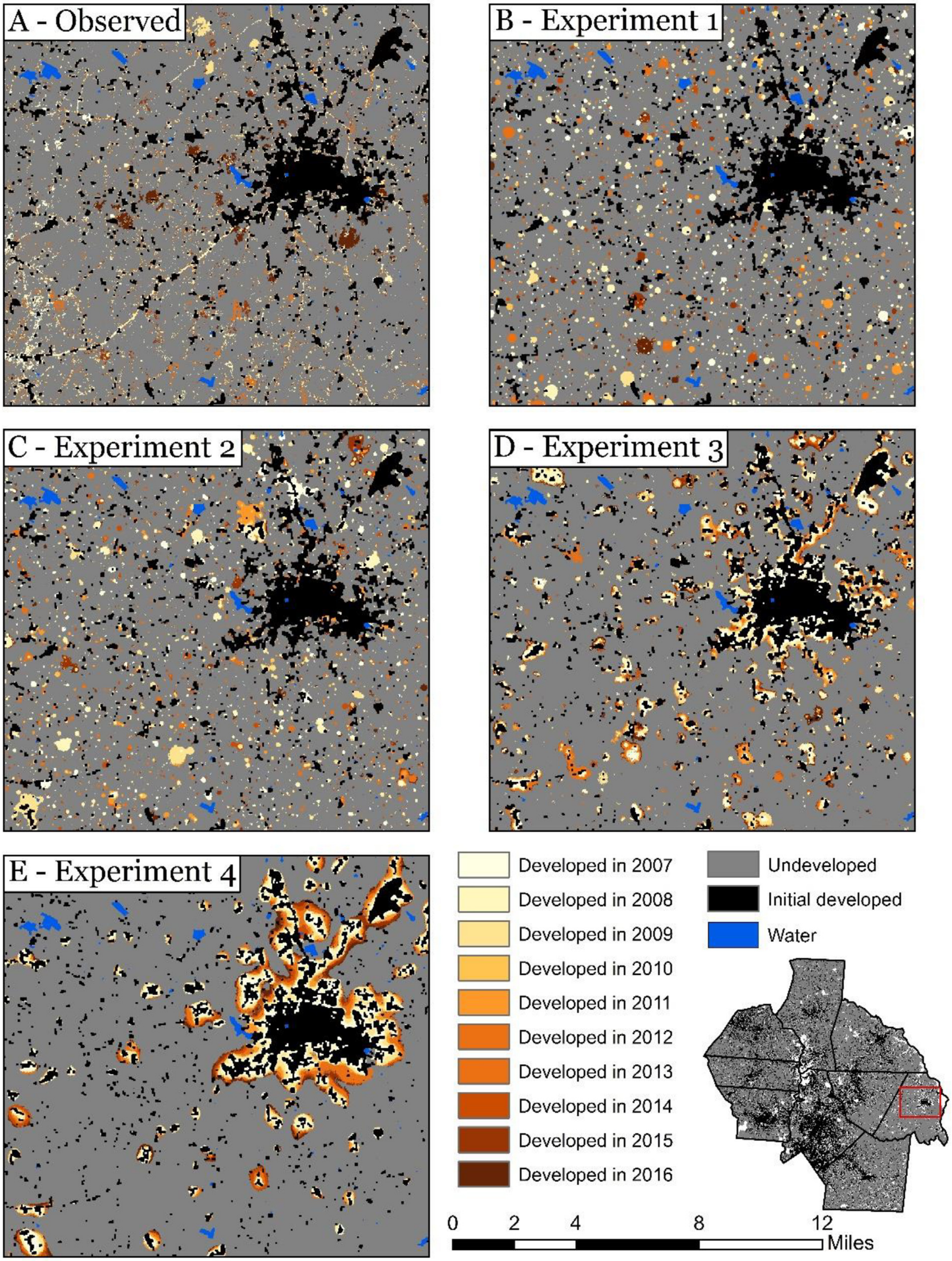

Fig. 4. An area demonstrating observed (A-Observed) and each configuration experiment (B-E) from 2006 to 2016. Visually, Experiment 4 is the poorest match of the patterns and shapes of new development of the four experiments, however it performs well with respect to the commonly applied accuracy measure of allocation.

would conclude that Experiment 4 had the lowest disagreement among the four experiments (Fig. 5). However, allocation differences among each experiment were not statistically significant. Finally, we examined configuration disagreement for each experiment and compared that result to each experiments allocation disagreement. No single experiment had the lowest disagreement (highest accuracy) for both configuration and allocation (Fig. 5). Configuration disagreement results also demonstrated greater inter-annual variability within the computational

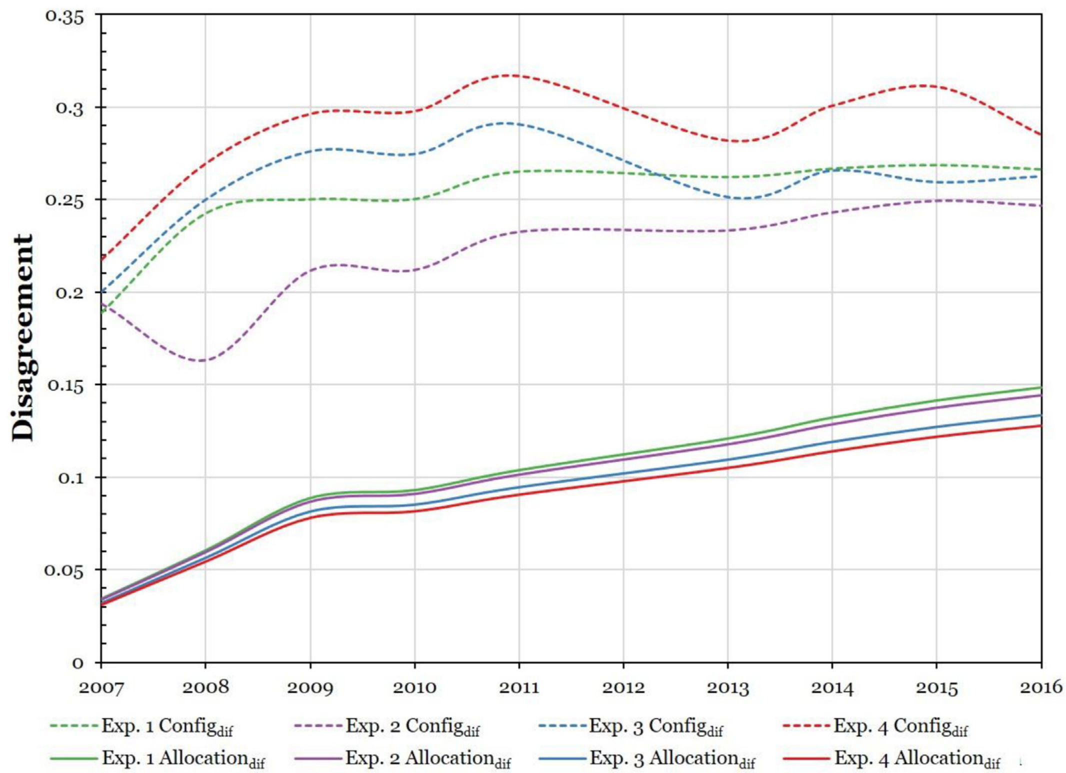

Fig. 5. Annual allocation and configuration disagreement from 2007 to 2016 for each experiment. Experiment 4 has the lowest allocation disagreement but the highest configuration disagreement. Comparatively, Experiment 2 has the second highest allocation disagreement but the lowest configuration disagreement.

experiments, indicating that at times some experiments exhibited less disagreement than others and vice-versa. We found that while Experiment 4 had the lowest allocation disagreement, it had the greatest configuration disagreement (Fig. 5). Intuitively this result makes sense, as visual inspection of Experiment 4 shows a distinctly different pattern of simulated new development than the observed dataset. Configuration disagreement for Experiment 2 appears to best simulate the observed configuration and patterns of new development with the lowest configuration disagreement for the entire ten-year validation period, in spite of its higher allocation disagreement.

4. Discussion

Environmental research, decision making, and policy formation rely on land change models, creating an imperative that they be ecologically meaningful, accurate, and repeatable (DeAngelis & Yurek, 2017). This work builds upon previous research to better understand land change model accuracy (Olmedo et al., 2015; Pickard, Van Berkel, et al., 2017; Pontius Jr & Millones, 2011; Pontius Jr & Parmentier, 2014; van Vliet, Bregt, & Hagen-Zanker, 2011) and introduces a methodology that may provide greater ecological context with respect to model accuracy. The metric we tested here, configuration disagreement, can provide clarity on the performance of land change simulations with respect to the size, shape, and spatial arrangement of simulated pixels. Our results indicate that current validation methods may misrepresent, or altogether fail to account for, the accuracy of a simulation in realistically mimicking the spatial arrangement of land change. We revealed that current validation methods (e.g. Pontius Jr & Millones, 2011) identified simulations with configurations that were dissimilar to observed patterns, yet having the highest allocation accuracy. Therefore, land change research that is motivated by spatial patterning will likely achieve more realistic results by focusing on configuration over allocation accuracy.

This work demonstrates inherent misrepresentations in the way that we determine land change simulation accuracy. By eliminating quantity disagreement and only manipulating the spatial arrangement of newly simulated pixels, we identified substantial differences in computational experiments that previously would be considered equally accurate. Landscape ecology focuses on broad spatial scales and the ecological effects of spatial patterning of ecosystems (Forman, 1995; Turner, 1989). Validation of land change simulations should, likewise, focus on a broad spatial scales and take into account spatial patterning (i.e., configuration), with less concern for how correct a single pixel location may be. The configuration of a landscape can impact many ecological processes and functions, species richness (Weibull, Ostman, & Granquist, 2003), fragmentation (Nagendra, Munroe, & Southworth, 2004) or edge effects (Villasenor, Driscoll, Escobar, Gibbons, & Lindenmayor, 2014), pollination (Kennedy et al., 2013), temperature and urban heat islands (Connors, Galletti, & Chow, 2013), water quality (Chaplin-Kramer et al., 2016), and ecosystem services (Eigenbrod et al., 2011; Pickard, Gray, & Meentemeyer, 2017), and should not be overlooked when evaluating land change model accuracy. With the increasing emphasis of coupling land change model outputs with ecological processes models (e.g, Pickard, Van Berkel, et al., 2017; Xie, Huang, He, & Zhao, 2018), configuration accuracy is critical for realistically assessing the ecological ramifications resulting from land change.

Assessment of configuration disagreement should be considered complimentary to current validation metrics (Pontius Jr & Millones, 2011; Pontius Jr & Santacruz, 2014), rather than a replacement. Understanding the quantity, allocation, and configuration disagreement of a land change simulation collectively provides a more comprehensive understanding of simulation accuracy then any single metric alone (Aquejdad, Houet, & Hubert, 2017; Pickard, Van Berkel, et al., 2017). While the relative importance of different components of model accuracy will depend on the study objective, limiting quantity disagreement should always be considered a key priority (Pontius Jr et al., 2018). Beyond quantity, land change science is potentially reaching a crossroads, where the next generation of models may better serve the research community by focusing on configuration, even at the cost of decreased allocation accuracy. Furthermore, with the decrease in time and resource costs to develop multiple validation maps, we suggest that, regardless of specific metric, future accuracy assessments be required at multiple time points. Understanding trends in model accuracy can be far more enlightening than any one assessment at a single time point. Our results are consistent with previous research showing decreasing trends in model accuracy as a simulation extends further from the initialization timepoint (Aguilera, Valenzuela, & Botequilha, 2011; Olmedo et al., 2015; Peterson, Bergen, & Brown, 2009). By assessing model accuracy at multiple time points, it is possible to weight the choice of extending a simulation further into the future compared to the observed trend in accuracy.

We have demonstrated several reasons why it might be more helpful to validate simulations in terms of configuration accuracy, yet the methods to measure configuration come with some limitations. First, configuration disagreement is an index, making it difficult to understand the differences between two configuration values without evaluating specific accuracy scores for each FRAGSTAT metric. For example, configuration disagreement in 2008 for simulations 1 and 4 was approximately 16 and 26%, respectively (Fig. 5). While the index values are not easily interpretable beyond their direct comparison (e.g., a map with 16% disagreement is a better representation of configuration compared to a map with 26%), they do identify differences in map accuracy that are likely useful in validating pattern and spatial arrangements. When combined with a robust sensitivity analysis of each individual FRAGSTAT metric, model accuracy can be improved by modifying specific components (e.g. number of patches, shape or size). Using Experiment 4 as an example, by evaluating each individual metric we could conclude that to improve the model's accuracy more patches of particular sizes would likely be needed. Configuration disagreement provides an overarching value to compare model accuracy, and when combined with individual FRAGSTAT metrics the researcher can now be better informed of how the model is performing. FRAGSTATS contains dozens of metrics that could be used to interpret configuration (Cushman et al., 2008), further research is needed that focuses on the selection of specific metrics for use in assessing landscape configuration.

This research evaluated configuration disagreement at the broadest spatial extent simulated for each Experiment. In this case configuration disagreement values provide an estimate of how well each experiment performed across the entire simulated landscape, however future research is needed to explore the relationship between scale and accuracy. Within each Experiment it is likely that some geographic areas perform differently in simulating new development compared to the overall configuration disagreement score. Approaches such as using a moving window or randomly selecting smaller geographical units could provide a different level of detail with respect to model accuracy. Ultimately, the choice of scale or boundary unit for use in quantifying configuration disagreement is at the researcher's discretion and should be chosen based on the specific research objectives. For example, research focusing on flooding within an urban neighborhood projected to further develop may have better results by evaluating configuration disagreement for only that neighborhood, rather than the entire county. The purpose of this research is to provide new methods for understanding model accuracy with respect to modeling patch shapes and sizes, and for the researcher to adapt this index to their specific needs.

Like previous model validation studies, our results include some limitations and assumptions. We classified Landsat imagery into categorical maps, distinguishing between developed and undeveloped. Discretizing Landsat imagery introduces error into the reference map. Accuracy assessments of the classified imagery can provide information regarding this error (Olofsson et al., 2013; Olofsson et al., 2014). However, final results of model accuracy tests must consider the underlying error associated with the reference maps. Landsat imagery is used often in validation studies, likely because it is free and has high temporal coverage across the United States. Whether 30-m spatial resolution provides high enough detail to adequately capture realistic patterns of new development has been widely debated (Cadenasso, Pickett, & Schwarz, 2007). Using finer scale resolution imagery may better capture the spatial heterogeneity of urban cities and suburbs and represents an emerging land change modeling research frontier. Lastly, we used a land change model that attempts to correlate environmental variables with observed changes in landscapes, yet other socioeconomic drivers exist in the determination of where new development is sited. Inclusion of socioeconomic or development planning related information may increase validation accuracy and further increase the realism of simulations.

5. Conclusion

As modeling capabilities continue to increase exponentially, it is critical that validation approaches simultaneously expand. Using four computational experiments we demonstrate the importance of including errors of configuration disagreement in addition to allocation and quantity disagreement in the assessment of land change models. Ultimately, the research objective will guide which component, or components, of model disagreement are most critical for consideration. Here, we demonstrate why it may be more helpful to validate simulations in terms of configuration accuracy if a study requires accurately modeling the spatial patterns and arrangements of land cover. When configuration disagreement is considered, ecologically meaningful land change science could be enhanced. Further study will be required to better understand the best methods to quantify spatial patterning and arrangement at the landscape scale. Configuration disagreement provides a common starting framework for comparing across land change models with respect to the form and arrangement of simulated land change.

Acknowledgements

This work was supported by the Center for Geospatial Analytics at North Carolina State University, and by Tetra Tech Inc. The authors would like to thank Robert Gilmore Pontius Jr. for his thorough, critical review of an earlier version of this manuscript.

Appendix

Appendix A Results of each FRAGSTATS metric for the observed and experiments 1–4.

Number of patches (NP)									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	36,369	66,023	74,470	74,793	79,982	86,239	84,975	86,869	87,129
Exp 1	26,046	34,065	42,111	43,281	46,098	50,393	53,234	55,445	57,119
Exp 2	23,734	29,557	35,478	36,335	38,571	41,575	43,596	44,960	46,051
Exp 3	17,431	17,824	18,032	18,231	18,659	18,985	19,270	19,587	19,478
Exp 4	15,821	15,307	14,690	14,690	14,452	14,044	13,775	13,554	13,343

GYRATE_AM									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	7499	8034	10,228	10,365	11,666	12,076	13,028	13,965	14,115
Exp 1	4790	4696	5163	5140	5116	5239	5309	5322	5320
Exp 2	4874	6717	6642	6649	6608	6694	6742	6743	6793
Exp 3	5312	5461	6018	6091	6150	7551	7599	8352	8325
Exp 4	5384	5610	6358	6393	6512	7914	7987	8041	9349
FRAC_AM									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	1.248	1.261	1.283	1.285	1.291	1.299	1.308	1.317	1.320
Exp 1	1.219	1.217	1.219	1.219	1.219	1.221	1.222	1.223	1.223
Exp 2	1.223	1.233	1.235	1.236	1.237	1.239	1.241	1.243	1.244
Exp 3	1.234	1.238	1.244	1.246	1.248	1.256	1.257	1.261	1.261
Exp 4	1.220	1.218	1.218	1.218	1.219	1.221	1.221	1.221	1.223
CORE_AM									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	29,504	34,304	44,342	45,424	51,223	54,929	62,384	66,653	67,979
Exp 1	16,875	16,814	20,374	20,375	20,350	21,750	22,380	22,649	22,890
Exp 2	17,426	29,381	29,835	30,147	30,170	31,821	32,733	33,072	33,911
Exp 3	20,472	21,859	25,813	26,356	26,739	37,798	38,542	42,423	42,521
Exp 4	21,063	23,386	29,069	29,361	30,176	41,780	42,934	43,862	51,263
ENN_AM									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	71.454	66.478	64.360	64.249	63.756	63.330	62.949	62.754	62.631
Exp 1	78.718	76.155	73.698	73.288	72.586	71.498	70.809	70.198	69.837
Exp 2	78.865	76.381	74.143	73.795	72.988	71.956	71.263	70.787	70.438
Exp 3	80.007	78.108	76.411	75.865	74.886	73.560	72.547	71.998	71.860
Exp 4	82.613	82.203	81.481	81.136	80.696	80.225	80.066	79.573	79.133
ENN_CV									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	0.289	0.297	0.296	0.296	0.293	0.291	0.291	0.290	0.290
Exp 1	0.168	0.196	0.216	0.219	0.225	0.237	0.242	0.247	0.251
Exp 2	0.167	0.193	0.211	0.214	0.220	0.229	0.234	0.237	0.241
Exp 3	0.140	0.161	0.176	0.181	0.191	0.201	0.208	0.213	0.214
Exp 4	0.085	0.085	0.086	0.089	0.093	0.093	0.094	0.096	0.093
ECON_CV									
Year	2007	2008	2009	2010	2011	2013	2014	2015	2016
Obs	94.144	95.793	95.769	95.768	95.798	95.860	95.195	95.963	96.037
Exp 1	94.901	94.998	95.077	95.093	95.152	95.227	95.289	95.285	95.329
Exp 2	94.954	95.099	95.243	95.273	95.334	95.456	95.499	95.541	95.603
Exp 3	95.028	95.066	95.058	95.103	95.161	95.220	95.249	95.279	95.304
Exp 4	94.456	94.179	93.632	93.609	93.537	93.252	93.090	92.991	93.076

References

Aguilera, F., Valenzuela, L. M., & Botequilha, A. (2011). Landscape metrics in the analysis of urban land use patterns: A case study in a Spanish metropolitan area. *Landscape and Urban Planning, 99*, 226–238.

Aquejdad, R., Houet, T., & Hubert, L. (2017). Spatial validation of land use change models using multiple assessment techniques: A case study of transition potential models. *Environmental Model Assessment, 22*, 591–606.

Arvidson, T., Goward, S., Gasch, J., & Williams, D. (2006). Landsat-7 long term acquisition plan: Development and validation. *Photogrammetric Engineering and Remote Sensing, 72*, 1137–1146.

Cadenasso, M. L., Pickett, S. T. A., & Schwarz, K. (2007). Spatial heterogeneity in urban ecosystems: Reconceptualizing land urban land cover and a framework for classification. *Frontiers in Ecology and the Environment, 5*, 80–88.

Chaplin-Kramer, R., Hamel, P., Sharp, R., Kowal, V., Wolny, S., et al. (2016). Landscape configuration is the primary driver of impacts on water quality associated with agricultural expansion. *Environmental Research Letters, 11*, 074012.

Chen, Y., Li, X., & Ai, B. (2014). Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automation with a patch-based simulation strategy. *International Journal of Geographic Information Science, 28*, 234–255.

Clark, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automation model of historical urbanization in the San Francisco Bay area. *Environmental Planning, 24*, 247–261.

Connors, J. P., Galletti, C. S., & Chow, W. T. L. (2013). Landscape configuration and urban heat island effects: Assessing the relationship between landscape characteristics and land surface temperature in Phoenix, Arizona. *Landscape Ecology, 28*, 271–283.

Cushman, S. A., McGarigal, K., & Neel, M. C. (2008). Parsimony in landscape metrics: Strength, universality, and consistency. *Ecological Indicators, 8*, 691–703.

DeAngelis, D. L., & Yurek, S. (2017). Spatially explicit modeling in ecology: A review. *Ecosystems, 20*, 284–300.

Dorning, M. A., Kock, J., Shoemaker, D. A., & Meentemeyer, R. K. (2013). Simulating urbanization scenarios reveals trade-offs between conservation planning strategies. *Landscape and Urban Planning, 136*, 28–39.

Eigenbrod, F., Bell, V. A., Davies, H. N., Heinemeyer, A., Armsowrth, P. R., & Gaston, K. J. (2011). The impact of projected increases in urbanization on ecosystem services. *Proceedings of the Royal Society B, 278*, 3201–3208.

Florida, R., Gulden, T., & Mellander, C. (2008). The rise of the mega-region. *Cambridge Journal of Regions, Economy, and Society, 1*, 459–476.

Forman, R. T. T. (1995). Some general principals of landscape and regional ecology. *Landscape Ecology, 10*, 133–142.

Kennedy, C. M., Lonsdorf, E., Neel, M. C., Williams, N. M., Ricketts, T. H., et al. (2013). A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. *Ecology Letters, 16*, 1–16.

Kolb, M., Mas, J. F., & Galicia, L. (2013). Evaluating drivers of land-use change and transition potential models in a complex landscape in Southern Mexico. *International Journal of Geographic Information Science, 27*, 1804–1827.

Mas, J. F., Kolb, M., Paegelow, M., Camacho-Olmedo, M. T., & Houet, T. (2014). Inductive pattern-based land use/cover change models: A comparison of four software packages. *Environmental Modelling and Software, 51*, 94–111.

McGarigal, K., Cushman, S. A., Neel, M. C., & Ene, E. (2002). *FRAGSTATS: Spatial pattern analysis program for categorical maps.* Amherst: University of Massachusetts.

Meentemeyer, R. K., Tang, W., Dorning, M. A., Vogler, J. B., Cunniffe, N. J., & Shoemaker, D. A. (2013). FUTURES: Multilevel simulations of emerging urban-rural landscape structure using a stochastic patch-growing algorithm. *Annals of the Association of American Geography, 103*, 785–807.

Nagendra, H., Munroe, D. K., & Southworth, J. (2004). From pattern to process: Landscape fragmentation and the analysis of land use/land cover change. *Agriculture, Ecosystems, and Environment, 101*, 111–115.

Nedkov, S., & Burkhard, B. (2012). Flood regulating ecosystem services-mapping supply and demand, in the Entropole municipality, Bulgaria. *Ecological Indicators, 21*, 67–79.

Nelson, E., Mendoza, G., Regetz, J., Polasky, S., Tallis, H., et al. (2009). Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers of Ecology and the Environment, 7*, 4–11.

North Carolina State Demographic Office (2017). State Demographics Branch. County/state population projections. Retrieved July 15, 2017 from: https://www.osbm.nc.gov/demog/county-projections.

Olmedo, M. T. C., Pontius, R., Jr., Paegelox, M., & Mas, J. F. (2015). Comparison of simulation models in terms of quantity and allocation of land change. *Environmental Modelling and Software, 69*, 214–221.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of the Environment, 148*, 42–57.
Olofsson, P., Foody, G. M., Stehman, S. V., & Woodcock, C. E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of the Environment, 139*, 122–131.
Perez-Vega, A., Mas, J. K., & Lingmann-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. *Environmental Modelling and Software, 29*, 11–23.
Peterson, L. K., Bergen, K. M., & Brown, D. G. (2009). Forested land-cover patterns and trends over changing forest management areas in the Siberian Baikal region. *Forest Ecology and Management, 257*, 911–922.
Pickard, B. R., Gray, J., & Meentemeyer, R. K. (2017). Comparing quantity, allocation and configuration accuracy of multiple land change models. *Land, 6*, 1–21.
Pickard, B. R., Van Berkel, D., Petrasova, A., & Meentemeyer, R. K. (2017). Forecasts of urbanization scenarios reveal trade-offs between landscape change and ecosystem services. *Landscape Ecology, 32*, 617–634.
Pontius, R. G., Jr., Castella, J. C., Nijs, T., Duan, Z., Fotsing, E., et al. (2018). Lessons and challenges in land change modeling derived from synthesis of cross-case comparisons. In M. Behnisch, & G. Meinel (Eds.). *Trends in spatial analysis and modelling.* Cham: Geotechnologies and the Environment. Springer.
Pontius, R. G., Jr., & Li, X. (2010). Land transition estimates from erroneous maps. *Journal of Land Use Science, 5*, 31–44.
Pontius, R. G., Jr., & Lippitt, C. D. (2006). Can error explain map differences over time? *Cartography and Geographic Information Science, 33*, 159–171.
Pontius, R. G., Jr., & Maliza, N. R. (2004). Effect of category aggregation on map comparison. Geographic information science. *Lecture Notes in Computer Science, 3234*, 251–268.
Pontius, R. G., Jr., & Millones, M. (2011). Death to kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing, 232*, 4407–4429.
Pontius, R. G., Jr., & Parmentier, B. (2014). Recommendations for using the relative operating characteristic (ROC). *Landscape Ecology, 29*, 367–382.
Pontius, R. G., Jr., & Santacruz, A. (2014). Quantity, exchange and shift components of differences in a square contingency table. *International Journal of Remote Sensing, 35*, 7543–7554.
Renard, D., Rhemtulla, J. M., & Bennett, E. M. (2015). Historical dynamics in ecosystem service bundles. *Proceedings of the National Academy of Sciences, 112*, 13411–13416.
Ridd, M. K. (1995). Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities. *International Journal of Remote Sensing, 16*, 2165–2185.
Tayyebi, A., Pijanowski, B. C., & Pekin, B. K. (2015). Land use legacies of the Ohio River Basin: Using a spatially explicit land use change model to assess past and future impacts on aquatic resources. *Applied Geography, 57*, 100–111.
Terando, A. J., Costanza, J., Belyea, C., Dunn, R. R., McKerrow, A., & Collazo, J. A. (2014). The Southern Megalopolis: Using the past to predict the future of urban sprawl in the Southeast US. *PLoS One, 9*, 1–8.
Turner, M. G. (1989). Landscape ecology: The effect of pattern on process. *Annual Review of Ecological Systems, 20*, 171–191.
Verburg, P. H., & Overmars, K. (2009). Combining top-down and botton-up dynamics in land use modeling: Exploring the future of abandoned farmlands in Europe with the DynaCLUE model. *Landscape Ecology, 24*, 1167–1181.
Villasenor, N. R., Driscoll, D. A., Escobar, M. A. H., Gibbons, P., & Lindenmayor, D. B. (2014). Urbanization impacts on mammals across urban-forest edges and a predictive model of edge effects. *PLoS One, 9*, e97036.
van Vliet, J., Bregt, A. K., Brown, D. G., Delden, H., Heckbert, S., & Verburg, P. H. (2016). A review of current calibration and validation practices in land-change modeling. *Environmental Modeling and Software, 82*, 174–182.
van Vliet, J., Bregt, A. K., & Hagen-Zanker, A. (2011). Revisiting Kappa to account for change in the accuracy assessment of land use change models. *Ecological Modeling, 222*, 1367–1375.
Weibull, A. C., Ostman, O., & Granquist, A. (2003). Species richness in agroecosystems: The effect of landscape, habitat and farm management. *Biodiversity and Conservation, 140*, 1335–1355.
White, R., Engelen, G., & Uljee, I. (1997). The use of constrained cellular automata for high-resolution modeling of urban land use dynamics. *Environment and Planning B, 24*, 323–343.
Wu, C. S. (2004). Normalized spectral mixture analysis for monitoring urban composition using ETM plus imagery. *Remote Sensing of the Environment, 93*, 480–492.
Xie, W., Huang, Q., He, C., & Zhao, X. (2018). Projecting the impacts of urban expansion on simultaneous losses of ecosystem services: A case study in Beijing, China. *Ecological Indicators, 84*, 183–193.