Relationships between Cellular Automata based land use models parameters and spatial metrics: Enhancing understanding in a calibration context

Charles P. Newland, Holger R. Maier, Jeffrey P. Newman, Hedwig Van Delden, Aaron C. Zecchin

Abstract

This research determined the ability of different metrics to capture behavior of land use simulation outputs driven by adjustments to neighborhood rules, the defining component of transition potential based Cellular Automata land use models. Following a series of tests, the metric clumpiness, when used to evaluate the class housing low density, exhibited the most ideal behavior defined to capture adjustments to neighborhood rules.

C. P. Newland (Corresponding author) • H. R. Maier • J. P. Newman • A. C. Zecchin
Department of Civil, Environmental and Mining Engineering, University of Adelaide, South Australia
Email: charles.p.newland@adelaide.edu.au

H.R. Maier
Email: holger.maier@adelaide.edu.au

J.P. Newman
Email: jeffrey.newman@adelaide.edu.au

A.C. Zecchin
Email: aaron.zecchin@adelaide.edu.au

H. Van Delden
Research Institute of Knowledge Systems
Maastricht, the Netherlands
Email: hvdelden@riks.nl

1. Introduction

Land use, both urban and rural, evolves from a series of incremental land use changes over time. These changes are the result of a dynamic anthropogenic system where people make land use decisions in response to a number of factors including: the vicinity to other land uses; access to resources and services; and the suitability of land for certain activities. The analysis of these factors facilitates the development of more effective planning policies for the assessment of the impacts of future scenarios. Of the tools available for such assessment, Cellular Automata (CA) based land use models are popular because of their ability to reproduce complex dynamics from a simple rule structure, because they incorporate socio-economic and biophysical processes that drive land use change.

Numerous Land Use Cellular Automata (LUCA) models have been developed for scenario assessment. As the application of LUCA models has grown, spatial modelling frameworks have been developed for generic use that are applied to different case studies as part of decisions support systems (Van Delden et al., 2011). One class of LUCA model that has been applied extensively as part of decision support systems is transition-potential based constrained CA land use models (Van Delden et al., 2011), which use a number of different factors to calculate the likelihood of land use at a particular location in the future.

The development of widely used spatial modelling frameworks has mitigated the need to design case-specific models, providing major time saving advantages and better testing and verification of model concepts and mechanisms. As such, the emphasis is now on calibration, the adaptation of the existing frameworks to particular case studies (Hewitt et al., 2014). The common current calibration procedure for transition potential based LUCA models is manual, which is time and knowledge intensive. Given the benefits these models provide, there has been a recent push to automate the calibration process, specifically the parameter tuning stage.

The objective of parameter tuning is to generate model simulations that replicate observed data as accurately and realistically as possible through the manual adjustment of model parameters. Methods for the assessment of the quality of the parameter tuning process range from visual interpretation performed by experts (Barredo et al., 2004) to methods based entirely on analysis by spatial metrics, making the process more objective and repeatable (Hagen-Zanker, 2009). Metrics are generally classified as measuring either locational accuracy or landscape structure (Van Vliet, 2013). However, as there are a large amount of metrics available to characterize all elements of simulation results, particularly for landscape structure, current manual parameter tuning methods tend to make use of both visual interpretation and metrics (Van Delden et al., 2012).

The use of spatial metrics has facilitated the development of automated parameter tuning methods using optimization. This approach uses an automated procedure that changes LUCA model parameters such that the difference between spatial metrics calculated for both simulation derived land use maps and actual land use maps is minimized, called a fitness function. However, as highlighted in Table 1.1, previous studies that have introduced automated procedures for calibrating LUCA models have all used different metrics. In addition, the metrics that have been used in automated procedures are not consistent with those suggested for a manual calibration procedure. While it is known that calibration success is dependent on the quality of metrics used to assess model parameterizations, to date previous studies have not investigated the performance of calibration with respect to metric choice.

Table 1.1 Different metric combinations used to assess model performance

Reference	Calibration Mode	Metrics
(García et al., 2013)	Automatic, Genetic Algorithm	• Global Index [a] • Number of Patches [b] • Mean Patch Area [b] • Edge Density [b]
(Li et al., 2013)	Automatic, Genetic Algorithm	• Percentage of Landscape [b] • Largest Patch Index [b] • Landscape Division [b]
(Van Delden et al., 2012)	Manual	• Kappa Simulation [a] • Clumpiness Index [b] • Fractal Dimension [b] • Rank Size Distribution [b] • Enrichment Factor [b] • Visual inspection/interpretation [a, b]

a. Measure of locational accuracy
b. Measure of landscape structure

There is clear space to investigate the performance of calibration with respect to metric choice, incorporating the present knowledge of visual interpretation and metrics used as part of manual calibration processes to enhance automated calibration. One approach is formally defining the application of different metrics to capture what experts interpret intuitively during manual parameter tuning. Based on visual interpretation or a spatial metrics, an expert knows intuitively how to improve the simulated pattern and adjusts parameters accordingly. However, there is limited formal definition of how parameters should be adjusted based on spatial metrics. As highlighted by Table 1.1 there is little consistency between different approaches.

Ideally the optimal parameters obtained using automatic calibration procedures with metrics as fitness functions should generate results that are realistic, and thus represent processes correctly by the parameter values obtained. Presently this is better achieved using manual calibration methods. Thus, the impact the selection of different metrics has on the optimal parameters obtained as part of automated calibration processes must be understood. A first step in this process is the development of an understanding of the relationships between certain model parameters and metrics. Thus, the main objective of this paper is to determine the ability of different metrics to capture behavior of land use model simulation outputs driven by adjustments to different model parameters. Ideally, these links will enhance understanding when posing parameter tuning as an optimization problem, allowing for improvement on previous automation attempts.

2. Methodology

In order to investigate the relationship between optimal parameters and different metrics, a sensitivity analysis approach was followed. The sensitivity of the difference between simulated and observed metrics was calculated, using a manually calibrated model as a baseline with a set of parameter values that were considered optimal, which were systematically perturbed as part of the testing regime. This method had three components:

- A land use model;
- A number of metrics to investigate; and
- A series of tests that perturbed different parameters within the land use model.

2.1 Cellular Automata Land Use Model and Case Study Application

This study used the constrained CA, transition potential based land use allocation model Metronamica (Van Delden and Hurkens, 2011) that has three types of land use classes: passive, which only change as a result of other land use dynamics; active, which are actively modelled based on exogenous demands; and static, which do not change but influence dynamics. As a constrained LUCA model, the demand for each land use class is defined exogenously. For each time step, representing one year, land use classes are allocated based on locations with the highest potential. Potential is calculated for each land use class for each cell based on transition rules shown in Equation 2.1:

$$\mathrm{P}_{k,i} = v \times \mathrm{A}_{k,i} \times \mathrm{S}_{k,i} \times \mathrm{Z}_{k,i} \times \mathrm{N}_{k,i} \quad (2.1)$$

Where $\mathrm{P}_{k,i}$ is the potential for land use k in cell i, v is a scalable random perturbation term, $\mathrm{A}_{k,i}$ is the accessibility, the effect of the nearness and importance of different types of transport networks and infrastructure, for land use k in cell i, $\mathrm{S}_{k,i}$ is the suitability, the effect of a location's physical properties, for land use k in cell i, $\mathrm{Z}_{k,i}$ is the zoning status, the influence of policy and restrictions, for land use k in cell i, and $\mathrm{N}_{k,i}$ is the neighborhood rule, defining the interactions of land use classes, for land use k in cell i,

Neighborhood rules are functions that define the interactions between cells based on their distance to the location of interest. Existing land use patterns influences future land use patterns in three ways: through the inertia of land uses in a location; through the ease of conversion from one land use to another; and the attraction or repulsion effects exerted by land uses situated in the neighborhood of a location (Van Vliet et al., 2013b). This is reflected in neighborhood rules having two distinct regions as shown in Figure 2.1: Inertia/conversion, the effects that land uses exert at the point of interest, and attraction/repulsion, the effect exerted at a cell distance of greater than zero.

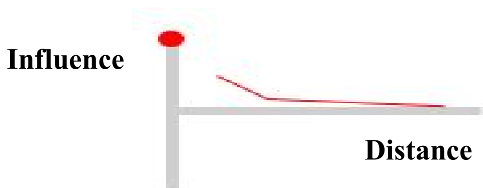

Fig. 2.1 The dissociation of neighborhood rules into different components, the point of inertia/conversion (dot) and the attraction/repulsion relationship (line)

Multiple neighborhood rules are required for a Metronamica application, equal to the product of the total number of land use classes and the number of function land use classes. Rules were classified into two groups as shown in Table 2.1. First, rules were classified as persistence if they defined the interaction of a land use with itself. These neighborhood rules have an inertia point at a cell distance of zero. Second, rules were classified as change if they defined the interaction between different classes of land use. These neighborhood rules have a conversion point at a cell distance of zero. By this classification scheme there were 10 persistence and 81 change neighborhood rules.

Table 2.1 A simplified representation of the classification scheme used to group different neighborhood effects. Persistence rules define land use class self-interaction, change rules define inter-class interaction

Land use Class	Residential	Industrial	Agricultural
Residential	*Persistence*	Change	Change
Industrial	Change	*Persistence*	Change
Agricultural	Change	Change	*Persistence*

Neighborhood rules are considered the defining element of a LUCA model (Van Vliet et al., 2013b). By expressing the influence exerted on land use dynamics by both the land use in a location and the land uses in neighboring locations, neighborhood rules have the greatest influence over model outputs. Thus, this research focused exclusively on the parameters of the neighborhood rules.

A Metronamica application for the Randstad area, a conurbation of the four largest cities in the Netherlands and the surrounding area, was used as a case study. The application was manually calibrated prior to testing. Simulations were run from a period of 2000 to 2030. The land use for the year 2000 and for a simulation to the year 2030 is shown in Figure 2.2, with the corresponding land use classes and categories shown in Table 2.2. There were a total of 16 land use classes, 10 of which were functions. Thus the application had 160 possible neighborhood functions that could be defined. 91 were manually calibrated prior to testing (detailed previously). Ten result maps were generated using the calibrated parameters to serve as synthetic data.

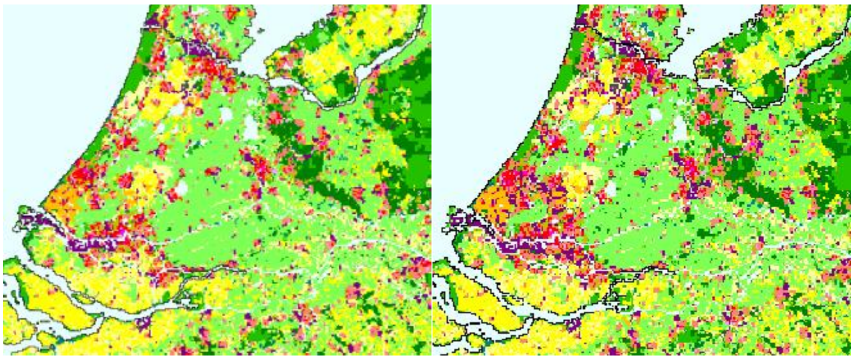

Fig. 2.2 The Randstad region investigated, simulated from 2000 (left) to 2030 (right)

Table 2.2 The Randstad region land use class categorization and color code

Class	Category	Color
Other agriculture	Vacant	
Pastures	Vacant	
Arable land	Vacant	
Greenhouses	Function	
Housing low density	Function	
Housing high density	Function	
Industry	Function	
Services	Function	
Socio cultural uses	Function	
Forest	Function	
Extensive grasslands	Function	
Nature	Function	
Recreation areas	Function	
Airport	Feature	
Fresh water	Feature	
Marine water	Feature	

2.2 Metric Classification

Due to the inherent uncertainty and complexity of land use change processes, it is not appropriate to validate a land use model exclusively on the ability to reproduce historic land use changes. This results in over-calibration at the cost of realism (Kok et al., 2001). Instead, a more comprehensive validation approach is required where land use models are evaluated on the ability to be accurate and realistic (Hagen-Zanker and

Marten, 2008). Within this scope, metrics are classified as quantifying either locational accuracy or landscape structure, alternately termed predictive and process accuracy respectively (Brown et al., 2005).

Predictive accuracy assesses whether land use changes are allocated correctly (Van Vliet et al., 2013a). Process accuracy assesses how realistically a pattern has been simulated. Two types of process accuracy metrics are currently used. First are landscape metrics derived from landscape ecology (McGarigal, 2014) that are sub-categorized as measuring one of four components of a landscape: area/edge/density, shape, contagion/interspersion and diversity (Peng et al., 2010). Second are metrics derived from complexity science such as Zipf's law (Gabaix, 1999) and residential rank-cluster-size distribution (White, 2006).

For this research landscape metrics derived from landscape ecology were tested. Eight different metrics were selected, based on the subcategorization scheme used in landscape ecology to determine if this distinction was meaningful in a land use modelling context. Table 2.3 summarizes the different metrics used, providing the sub-category and a brief description. The calculation of the metrics was performed using the Map Comparison Kit (http://mck.riks.nl/). The formula for each metric can be found in the reference given.

Table 2.3 Summary of tested metrics

Metric	Sub-category	Description
Patch size [a]	Area/Edge/Density	The average size of patches for the entire landscape
Edge density [c]	Area/Edge/Density	The average length of edge segments in the landscape relative to the total landscape area
Shape index [b]	Shape	A measure of the geometric complexity of the patches composing the landscape
Fractal dimension [b]	Shape	A ratio of complexity comparing how detail in a pattern changes relative to the scale it is measured
Interspersion and juxtaposition of edges [b]	Contagion/Interspersion	The extent to which patch types are interspersed within the landscape
Clumpiness [c]	Contagion/Interspersion	The proportional deviation of the proportion of like adjacencies involving the

		corresponding class from that expected under a spatially random distribution
Shannon's diversity index [c]	Diversity	The proportional abundance of each patch type relative to the abundance of that patch type in the landscape
Simpson's diversity index [c]	Diversity	The proportional abundance of each patch type relative to the abundance of that patch type in the landscape

a. Leitão et al., 2006
b. Hagen-Zanker, 2006
c. McGarigal, 2014

Certain spatial metrics can only be calculated for land use classes, and not aggregated for the entire landscape. Of the metrics selected, clumpiness required selection of such a class. For the research performed, the class housing low density was selected for this purpose. This selection was based on the assumption that urban regions were the major investigative purpose of the case study used, and this class was the best proxy for urban regions. However, this definition was arbitrary and is not necessarily transferable, as different application of the model will not always have the specified class or be constructed around investigating urban regions.

2.3 Testing Procedure

To determine if an identifiable relationship existed between the neighborhood rules of a transition potential based, constrained CA model and the landscape ecology metrics selected for testing (or the general subcategory) a procedure was designed using a one-at-a-time sensitivity analysis approach. The procedure varied components of the neighborhood rules, and measured the difference between observed and simulated metric values, to explore the relationship between parameter values and metrics.

Tests were designed using the different categorizations and components of the neighborhood rules that could be manipulated. The different tests, listed in Table 2.4, were designed to alter various components of the neighborhood rules as follows: Persistence, inertia point; change, conversion point; persistence, tail; conversion, tail. Using a one-at-a-time sensitivity analysis approach parameter values were varied about percentage intervals, from 20% to 300% in 20% intervals, of the manually obtained calibrated values. The manually obtained (optimum) parameters were used to generate optimum metric values against which the difference between perturbed metrics was calculated. Therefore, the minimum difference was expected at the 100% interval.

Table 2.4 Tests performed for evaluation of different components of neighborhood rules

Test	Name	Explanation
1	Persistence Tail pt. 2	Percentage variation of the point of the tail at a cell distance of 2 for the 10 persistence functions
2	Persistence Tail pt. 1	Percentage variation of the point on the tail at a cell distance of 1 for the 10 persistence functions
3	Persistence Tail	Percentage variation of all points on the tail (cell distance 1-8) for the 10 persistence functions
4	Persistence Inertia	Percentage variation of the inertia point (cell distance 0) for the 10 persistence functions
5	Persistence All	Percentage variation of all points for the 10 persistence functions
6	Change Tail	Percentage variation of all points on the tail (cell distance 1-8) for the 81 change functions
7	Change Conversion	Percentage variation of the conversion point (cell distance 0) for the 81 change functions
8	Change All	Percentage variation of all point for the 81 change functions

Testing was performed by inputting the values for each neighborhood rule based on the test being conducted into an Excel spreadsheet. A PHP script was used to extract the inputs from the spreadsheet and input these as neighborhood rules that were run through the command line version of Metronamica. The stochastic nature of the model (random component from transition potential equation) meant multiple runs for the same set of parameters were required. Thus each iteration of the neighborhood rule intervals was run ten times. A second PHP script was used to generate a log file to store result maps that were used for evaluation. The MCK was used to calculate metric values for each result map. Post processing compared the results by calculating the difference for the ten synthetic data maps and ten simulated maps. Thus, 100 difference measurements were calculated per interval per test.

3 Results

Determining the ability of different metrics to capture behavior of land use model simulation outputs driven by adjustments to different model parameters required the definition of trends in the fitness function (difference between simulation and data metric values), referred to hence forth as ideal behavior, as optimal. It was expected that the difference between metrics calculated for data and for simulated outputs would be minimized for all metrics for the manually calibrated parameter combination (100% of the calibrated value). Therefore further definition was required to distinguish better performing metrics. For behavior to be considered ideal, two further trends in the difference between metrics calculated for simulations and data were expected in model results, shown in Figure 3.1 (a). The first was that a steep, linear gradient towards the optimum parameter combination, to make the search for the optimum solution faster, unlike Figure 3.1 (b). Second, behavior was considered ideal if there was only one sign change at the optimum parameter combination, which implied that only a single optimum solution existed, unlike Figure 3.1 (c).

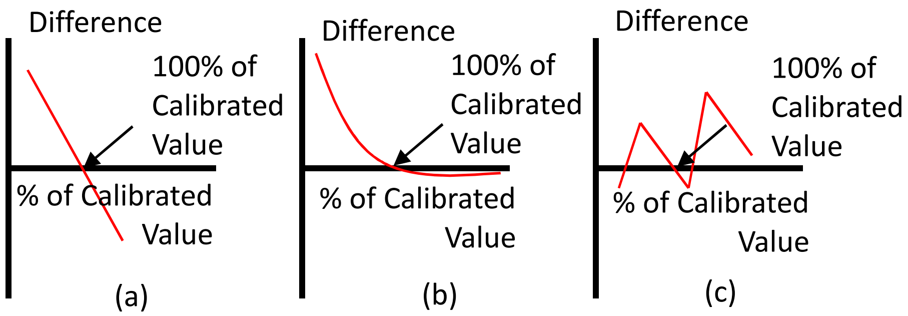

Fig. 3.1 Idealized representations of different trends expected as results. (a) Exhibits ideal behavior because of a steep gradient towards a single optimum solution. (b) Does not exhibit ideal behavior because the trend is not a steep, linear gradient. (c) Does not exhibit ideal behavior because there are multiple sign changes.

To summarize, when calculating the difference between observed and simulated data, ideal behavior was exhibited by a difference between simulated and observed metric values that:

- Had a steep, linear gradient to the calibrated value; and
- Had a single sign change at the calibrated value.

With ideal behavior defined results were interpreted to distinguish which landscape structure metrics were better suited to identifying parameter changes in neighborhood functions. Results were compiled graphically for each metric used, sorted by test.

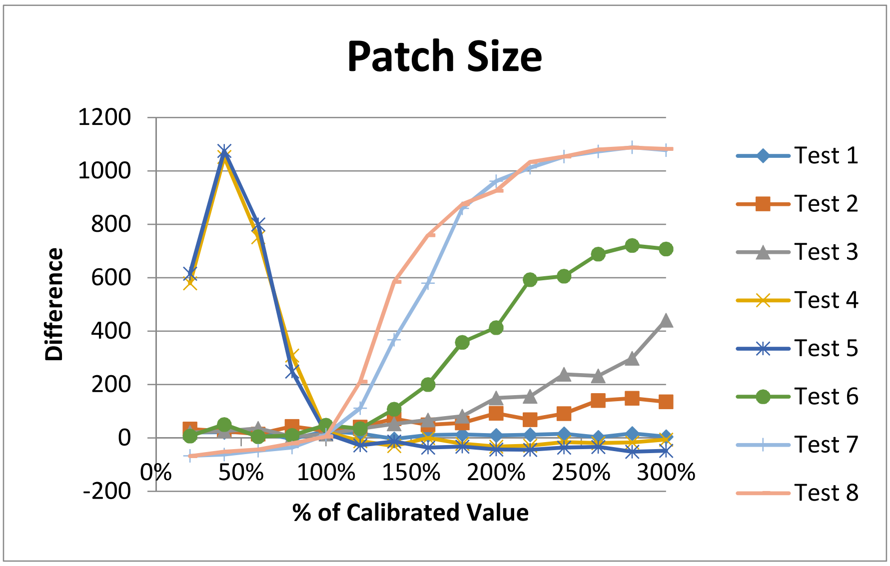

Fig.3 2. The average difference of the patch size measured for simulated and observed data for variations of neighborhood rule components

As shown in Figure 3.2, the trends between the differences in patch size exhibited non-ideal behavior of the gradient towards the calibrated value. Also, there was an unexpected increase in the difference for tests four and five despite parameters closer to the optimal parameter set. Also for tests two, three and six the difference had no sign change, whilst tests one had more than one.

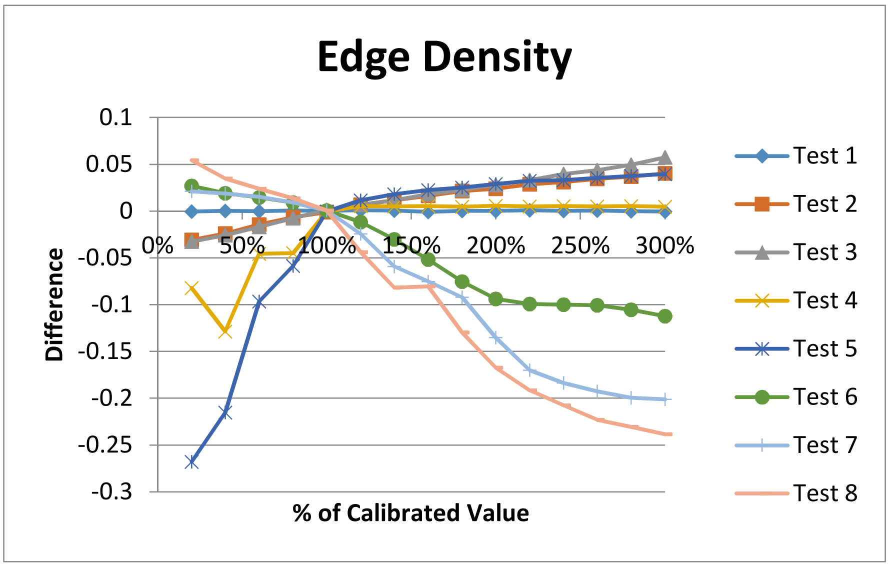

Fig. 3.3 The average difference of the edge density measure for simulated and observed data for variations of neighborhood rule components

As shown in Figure 3.3, the difference in edge density only exhibited ideal gradient trends for test two and three. Although other tests exhibited steep trends, they were not smooth, for example the local minimum in test four. All tests resulted in a single sign change except for test one.

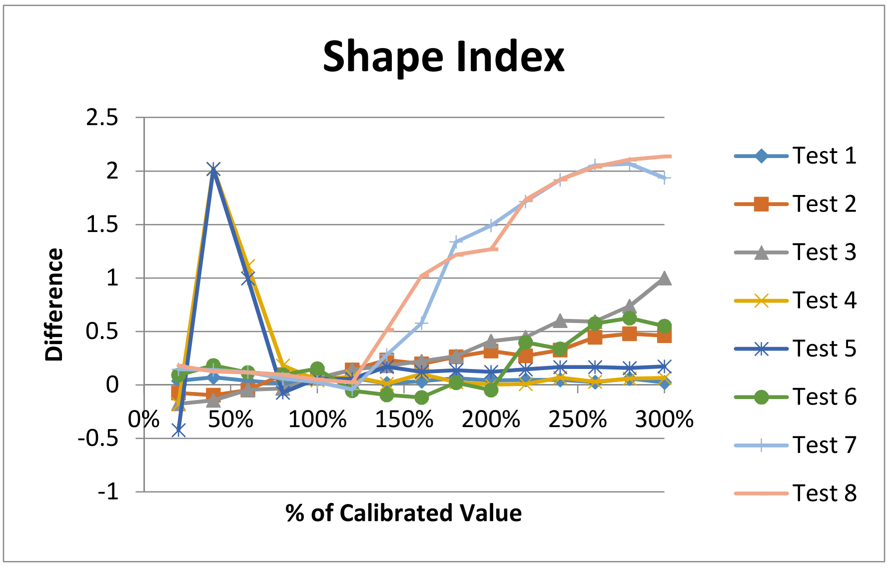

Fig. 3.4 The average difference of the shape index measured for simulated and observed data for variations of neighborhood rule components

The trends in the differences for shape index, shown in Figure 3.4, did not exhibit ideal behavior. The gradients were not ideal for any of the tests performed and only tests two, three and four had a single sign change.

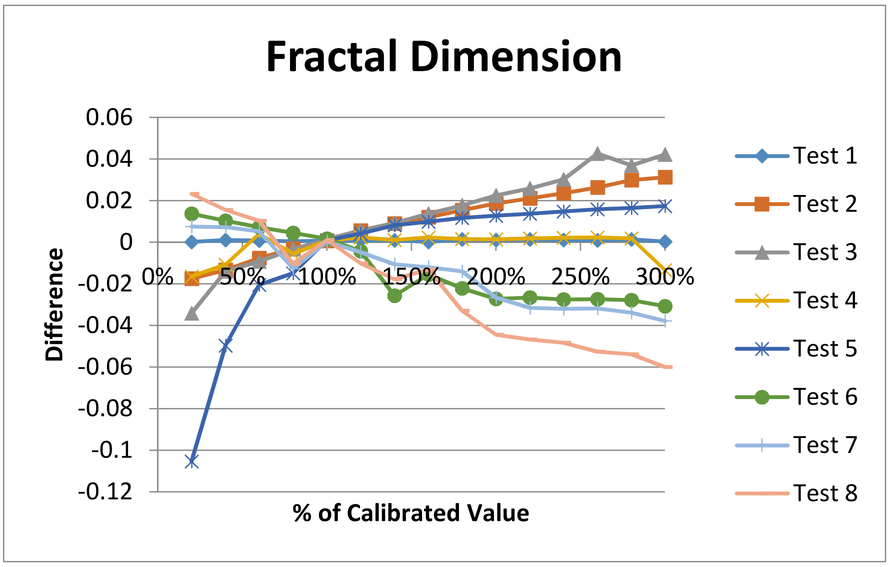

Fig. 3.5 The average difference of the fractal dimension measured for simulated and observed data for variations of neighborhood rule components

As Figure 3.5 shows, gradient trends in the difference for fractal dimension were only ideal for test two. Other tests exhibited inconsistent gradients, and test one had no sign change, and tests four, seven and eight each had greater than one sign change.

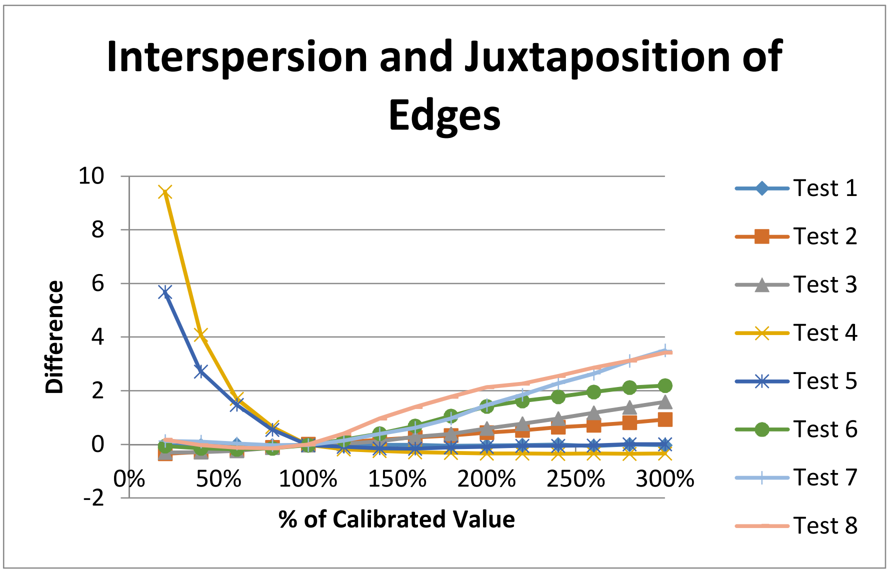

Fig. 3.6 The average difference of the interspersion and juxtaposition of edges measured for simulated and observed data for variations of neighborhood rule components

The behavior of interspersion and juxtaposition of edges, shown in Figure 3.6, was not ideal. Only tests two and three resulted in a gradient that was considered ideal for generating optimal parameters. Tests two, three, four and six resulted in difference trends with a single sign change.

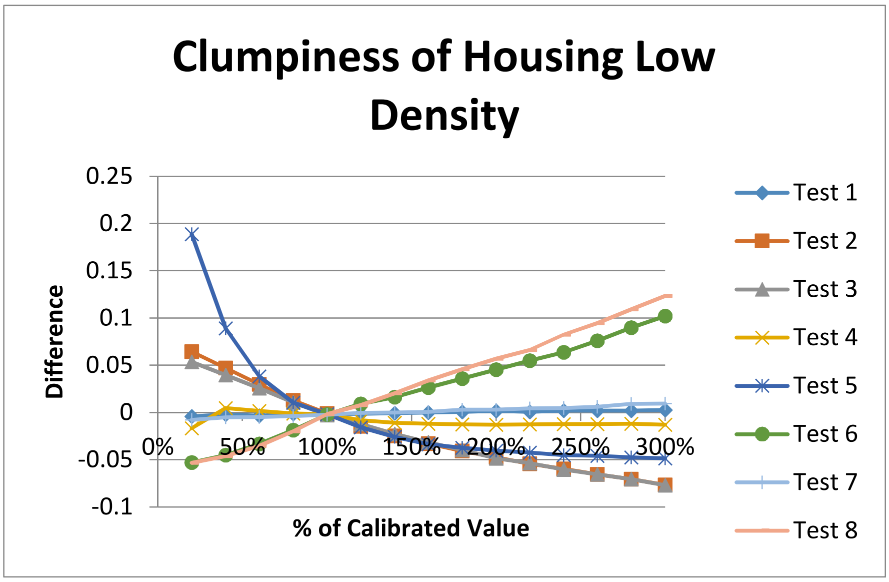

Fig. 3.7 The average difference of the clumpiness of the class housing low density measured for simulated and observed data for variations of neighborhood rule components

The behavior of the difference in the clumpiness of the class housing low density for a majority of the tests was ideal. Only test four and five exhibited non-ideal gradient behavior, and the only test resulting in greater than one sign change was test four.

As previously stated it was assumed the class housing low density was the appropriate clumpiness class to measure. This was an important assumption. A class had to be selected for the clumpiness metric to be calculated, because the metric cannot be calculated for the entire landscape (unlike all other metrics used). As stated previously, the selection of the class housing low density was based on the assumption that urban regions were the major investigative purpose of the case study used, and this class was the best proxy for an urban region. This definition was arbitrary and not necessarily transferable. Therefore, to determine how dependent the performance of the metric was to the class being measured, clumpiness was calculated for different functional land use classes using data from test 4 (adjustments to the inertia points of persistence neighborhood rules). The results are presented in Figure 3.8.

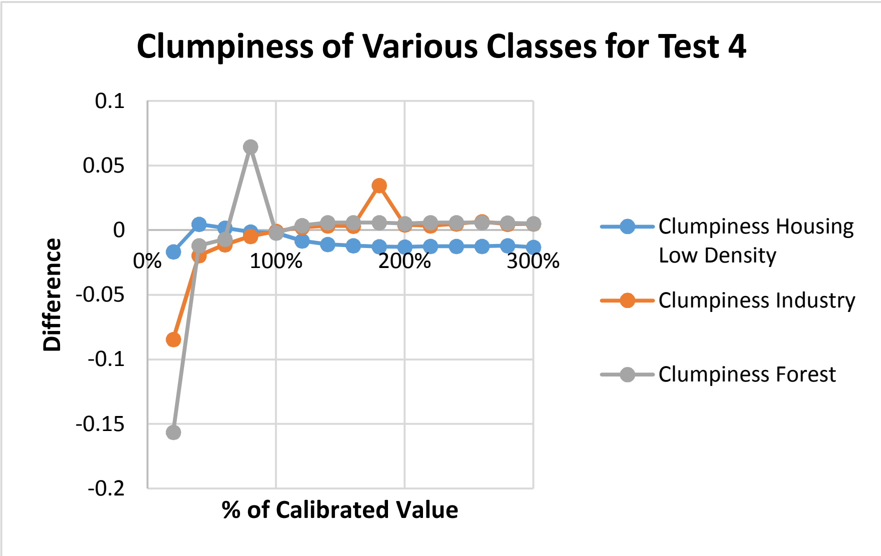

Fig. 3.8 The average difference of the clumpiness of three different land use classes measured for simulated and observed data for variations of neighborhood rule components

The class industry was selected as another proxy for urban areas. Forest was chosen for comparison if environmental space was being studied. The results in Figure 3.8 show trends that still reach optimum solutions for the calibrated value at 100%. However, what was notable for the clumpiness measures of the alternative classes used was the presence of random spikes that totally distorted the gradient of the difference trend. This reinforced that using the class housing low density was appropriate.

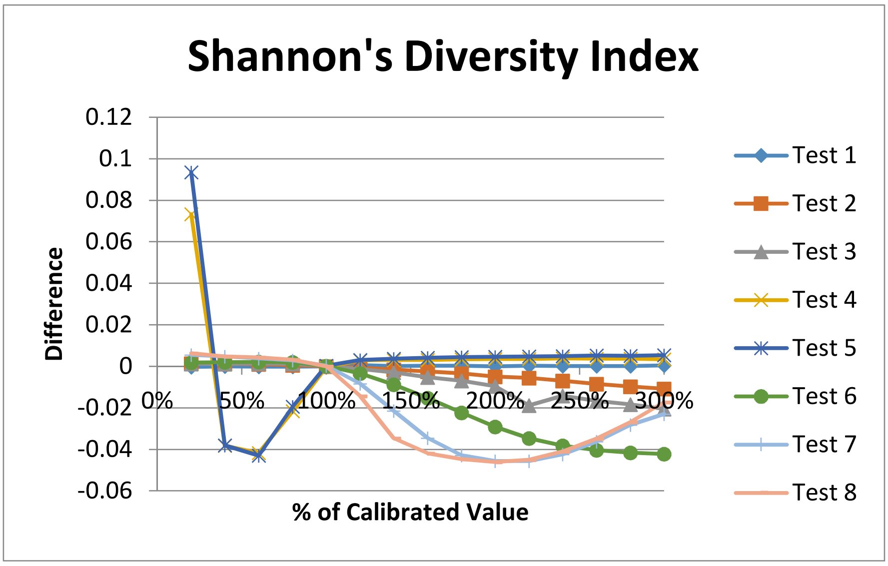

Fig. 3.9 The average difference of Shannon's diversity index measured for simulated and observed data for variations of neighborhood rule components

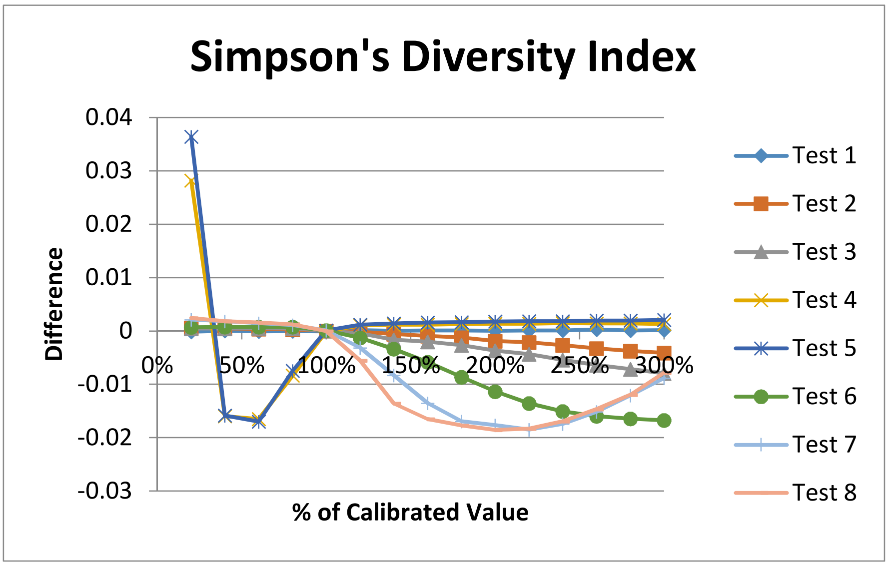

Fig. 3.10 The average difference of Simpson's diversity index measured for simulated and observed data for variations of neighborhood rule components

As shown in Figures 3.9 and 3.10 the diversity indices of both Shannon and Simpson behaved almost identically (when adjusted for scale). None of the trends in differences exhibited an ideal gradient. Tests one, four, and five had non-ideal sign change trends for both indices, with greater than one sign change.

As highlighted in Table 3.1 the metrics clumpiness of housing low density was comparatively the most ideal metric, based on the percentage of tests that resulted in trends with ideal behavior. This was due to the fitness function trends for most clumpiness tests adhering to the definition of an ideal gradient.

Table 3.1. Summary of results obtained from testing based on ideal behavior

	Percentage of tests with ideal fitness function behavior			
Metric	**Gradient**	**Sign change**	**Both**	**Rank**
Patch Size	0.0 %	50.0 %	0.0 %	5th
Edge Density	37.5 %	87.5 %		2nd
Shape Index	0.0 %	37.5 %	0.0 %	5th
Fractal Dimension	12.5 %	50.0 %	12.5 %	4th
Interspersion and Juxtaposition of Edges	25.0 %	50.0 %	25.0 %	3rd
Clumpiness of housing low density	75.0 %	87.5 %	75.0 %	1st
Shannon's diversity index	0.0 %	62.5 %	0.0 %	5th
Simpson's diversity index	0.0 %	62.5 %	0.0 %	5th

4. Conclusions and Recommendations

Previous attempts to automate the calibration of LUCA models have selected metrics based on the assumption that any metric can be used in a fitness function. Whilst it is true that all metric will exhibit optimum values for a fully calibrated model, this research was designed to more thoroughly investigate the relationship between metrics and parameters to make the process of selection more fully informed. The specific research objective was to determine if certain metrics derived from landscape ecology were more sensitive to variations in neighborhood rules, considered the defining element of a LUCA model (Van Vliet, 2013), and thus exhibited ideal behavior for use as part of automated calibration. Ideal behavior was defined based on graphical interpretation as:

- Having a smooth, steep gradient towards the calibrated value; and
- Having a single sign change at the calibrated value

From interpreting the results it was concluded that clumpiness of the class housing low density was the metric which exhibited the most ideal behavior for use as a fitness function. Based on the results, edge density was the next most appropriate, but exhibited less ideal gradient trends, implying that solutions would take longer to find.

Further research is required on additional case studies, with the use of historical periods to allow for comparison with real data, to verify the conclusions of hits research. Additionally, only landscape metrics have been analyzed. Further research will also focus on predictive accuracy metrics and process accuracy metrics derived from complexity science. Finally, the role of additional processes used as part of calculating transition potential must be assessed.

References

BARREDO, J. I., DEMICHELI, L., LAVALLE, C., KASANKO, M. & MCCORMICK, N. 2004. Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. *Environment and Planning B: Planning and Design,* 32**,** 65-84.

BROWN, D. G., PAGE, S., RIOLO, R., ZELLNER, M. & RAND, W. 2005. Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science*, 19:2**,** 153-174.

GABAIX, X. 1999. Zipf's law for cities: An Explanation. *The Quarterly Journal of Economics,* 114**,** 739-767.

GARCÍA, A. M., SANTÉ, I., BOULLÓN, M. & CRECENTE, R. 2013. Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm. Application to a small urban settlement of NW Spain. *International Journal of Geographical Information Science,* 27**,** 1593-1611.

HAGEN-ZANKER, A. 2006. Map comparison methods that simultaneously address overlap and structure. *Journal of Geographical Systems,* 8**,** 165-185.

HAGEN-ZANKER, A. 2009. An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. *International Journal of Geographical Information Science,* 23**,** 61-73.

HAGEN-ZANKER, A. & MARTEN, P. Map comparison methods for comprehensive assessment of geosimulation models. *In:* GERVASI, O., MURGANTE, B., LAGANA, A., TANIAR, D., MUN, Y. & GAVRILOVA, M., eds. Lecture Notes in Computer Science, Proceedings of the 2008 International Conference of Computational Science and Its Application ICCSA, 2008 Berlin, Germany. Springer, 194-209.

HEWITT, R., VAN DELDEN, H. & ESCOBAR, F. 2014. Participatory land use modelling, pathways to an integrated approach. *Environmental Modelling and Software,* 52**,** 149-165.

KOK, K., FARROW, A., VELDKAMP, A. & VERBURG, P. H. 2001. A method and application of multi-scale validation in spatial land use models. *Agriculture, Ecosystems and Environment,* 85**,** 223-238.

LEITÃO, A. B., MILLER, J., AHERN, J. & MCGARIGAL, K. 2006. *Measuring Landscapes,* Washington, DC, Island Press.

LI, X., LIN, J., CHEN, Y., LIU, X. & AI, B. 2013. Calibrating cellular automata based on landscape metrics by using genetic algorithms. *International Journal of Geographical Information Science,* 27**,** 594-613.

MCGARIGAL, K. 2014. *FRAGSTATS Help,* University of Massachusetts, Amherst.

PENG, J., WANG, Y., ZHANG, Y., WU, J., LI, W. & LI, Y. 2010. Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. *Ecological Indicators,* 10**,** 217-223.

PONTIUS JR., R. G., BOERSMA, W., CASTELLA, J.-C., CLARKE, K., DE NIJS, T., DIETZEL, C., DUAN, Z., FOTSING, E., GOLDSTEIN, N., KOK, K., KOOMEN, E., LIPPITT, 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. *Annals of Regional Science,* 42**,** 11-37.

VAN DELDEN, H., DÍAZ-PACHECO, J., SHI, Y. & VAN VLIET, J. Calibration of cellular automata based land use models: lessons learnt from practical experience. CAMUSS, 2012 Oporto, Portugal.

VAN DELDEN, H. & HURKENS, J. 2011. A generic Integrated Spatial Decision Support System for urban and regional planning. *MODSIM.* Perth, Australia.

VAN DELDEN, H., VAN VLIET, J., RUTLEDGE, D. T. & KIRKBY, M. J. 2011. Comparison of scale and scaling issues in integrated land-use models for policy support. *Agriculture, Ecosystems and Environment,* 142**,** 18-28.

VAN VLIET, J. 2013. *Calibration and validation of land-use models.* PhD, Wageningen.

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 Modelling,* 222**,** 1367-1375.

VAN VLIET, J., HAGEN-ZANKER, A., HURKENS, J. & VAN DELDEN, H. 2013a. A fuzzy set approach to asses the predictive accuracy of land use simulations. *Ecological Modelling,* 261-262**,** 32-42.

VAN VLIET, J., NAUS, N., VAN LAMMEREN, R. J. A., BREGT, A. K., HURKENS, J. & VAN DELDEN, H. 2013b. Measuring the

neighbourhood effect to calibrate land use models. *Computers, Environment and Urban Systems,* 41**,** 55-64.
WHITE, R. 2006. Pattern based map comparisons. *Journal of Geographic Systems,* 8**,** 145-164.