SENSITIVITY ANALYSIS OF A GIS-BASED CELLULAR AUTOMATA MODEL

V. Kocabas [a, *], S. Dragicevic [a]

[a] Simon Fraser University, Department of Geography, Spatial Analysis Modeling Laboratory
8888 University Dr., Burnaby, BC, Canada, V5A1S6 – verdak@sfu.ca, suzanad@sfu.ca

KEY WORDS: Sensitivity Analysis, Cellular Automata, Modeling, Urban Growth, Spatial Metrics.

ABSTRACT:

Urban growth is dynamic and complex spatial process that has severe environmental and social impacts. High population growth results in the transformations of forested areas or high quality agricultural lands into urban land-use. The study and modeling of the urban growth and land-use change processes are complex due to the difficulties to represent the interactions between the physical and human environments. One of the models increasingly applied to urban research is based on cellular automata (CA) theory. Since models are approximations of the real world, they contain inherent errors due to the digital data input and are sensitive on model parameters and model misspecification thereby generating uncertainties in the results. The objective of this study is to explore these uncertainties through the sensitivity analysis (SA) of a GIS-based CA urban growth model. The impacts of changing CA neighbourhood size and type on the model outcome were addressed. The cross-classification, KAPPA statistic and spatial metrics were used as measures of sensitivity analysis in order to understand the CA model behavior and its limitations. The results from this research can provide better insights for improving the capabilities of current CA models to create more realistic output scenarios.

1. INTRODUCTION

Urban growth modeling can assist urban and regional planners to foresee impacts of their actions and policies (Wegener, 1994). To date, various urban growth models are developed due to the simple vs. complex, aspatial vs. spatial views of urban phenomena. For more then a decade, research focus has been on models using cellular automata (CA) theory as the approach which is capable to address the spatial complexity of the urban change process (Allen, 1997). CA models are receiving more attention due to the capability for handling spatial and temporal dimensions, using bottom-up approach, relying on geospatial data and capacity to couple with raster-based geographic information system (GIS) as well as with other approaches such as agent-based or multi-criteria evaluation (Batty, 1998; Wu, 1998; O'Sullivan and Torrens, 2000).

The main advantages of CA are in their simplicity, easy integration with raster GIS, and adaptability to various urban growth situations. CA models can generate complex patterns through the use of simple rules (Wolfram, 2002). In particular, it is possible to realistically represent spatial complexity and dynamics of urban growth change by choosing the configurations of basic elements of CA models such as cell states, cell size, neighborhood size and type, transition rules and temporal increments (Torrens, 2000; White and Engelen, 2000; Yeh and Li, 2001). In most of CA urban growth models the effects of varying different basic elements of CA are not yet fully addressed in the research literature. This study introduces an approach to sensitivity analysis of CA model in order to analyze the model responses and behavior with respect to the change of its elements more particularly neighborhood size and type. This study provides the assessment of CA model sensitivity versus its elements' variations, the consistency of model outcomes and the locational differences at specific land use types.

2. SENSITIVITY ANALYSIS

Sensitivity Analysis (SA) addresses the relationship of information flowing in and out of the model and deals with the sources of variation influencing the model outputs (Saltelli et al., 2000). It measures the change in the model output relative to a change in one or more of the input parameter values. In modeling practices, SA is a prerequisite since it determines the reliability of the model through assessing the uncertainties in the simulation results. It can also be considered as a resource optimization process since data gathering is the most important and expensive part of GIS. SA can be used also to test sub-models of the actual model and to determine the dependency of model outcomes on input parameters (Crosetto et al., 2000).

In current CA applications, SA is often used to help understand the behaviour of a model as well as the coherence between a model and the real world. The most common approach is based on variations of basic spatial and temporal CA elements, which represent input parameters in order to assess the outcome differences (White et al., 1997; Barredo et al., 2003). However, some recent studies have addressed the issue of errors and uncertainties related to CA models (Yeh and Li, 2003) and provided the analysis of CA model behavior with respect to changing model components (Clarke, 2003). The appropriate choice of transition rules were considered as the key component of the CA model (Childress et al., 1996). White et al. (1997) changed the transition rules in the CA model and compared the model results and simulation outputs visually and cell-by-cell with the actual land use. This approach can be regarded in current CA literature as a calibration procedure since model and simulation outputs were compared with the real data.

Variations of different cell size and cellular configurations were explored by Chen and Mynett (2003) and Jenerette and Wu (2001). Moreover, Liu and Andersson (2004) examined the effects of temporal dynamics on the behavior of a CA-based

* Corresponding author.

urban growth model.

2.1 SA method for CA modeling

In this study, univariate sensitivity analysis was performed in which the parameters i.e. the basic CA elements were assumed independent. This is mainly because of the fact that complicated mathematical operations are difficult to derive for the analysis of uncertainties and errors in CA simulation testing as stated by Yeh and Li (2003). Furthermore, KAPPA statistic can be used (Congalton and Mead, 1983) for the analysis of simulation outcomes since CA operates in raster environment. However, KAPPA statistic has some disadvantages since it does not quantify the patterns of map land-use classes, thus even a small difference of classes between two maps is shown as an inconsistency (White et al. 1997; Barredo et al. 2003; Straatman et al. 2004). In order to overcome these shortcomings this study proposes an integrated SA method that employs qualitative and quantitative approaches of cross-classification map, KAPPA index with coincidence matrices and spatial metrics. The effective SA was accomplished by varying the CA basic elements one at the time, more specifically neighborhood size and type. Subsequently, the model outputs i.e. the simulated maps were compared with each other in terms of the combination of different approaches to provide the quantitative and qualitative measures and achieve the high accuracy in map comparisons.

Under the heading of qualitative part of the approach, the visual comparison of the outcomes was used. The advantage of this method is in its capability to detect the uncertainties and inconsistencies while comparing the simulation results. White et al. (1997) and Mandelbrot (1983) pointed out that visual similarity is an important factor for comparison of complex fractal forms. Therefore, the cross-classification map, which is the result of multiple GIS overlay analysis showing all combinations, was produced to enable visual comparison and the analysis of visual similarities and differences between model outputs. The cross-classification map depicts the locations of the combinations of the map land-use classes of urban growth model output for the two maps that were being compared. The advantage of a cross-classification map is reflected in its capability to easily determine the locational differences of map land-use classes between the two maps.

Quantitative part of the developed approach relies on the following two categories: a coincidence matrix with KAPPA index and spatial metrics. The KAPPA index is computed with the coincidence matrix to compare the results of changing CA element values. It was introduced by Cohen (1960) and adapted for accuracy assessment in the remote sensing applications by Congalton and Mead (1983). It ranges from 0 to 1, and when it approaches 1 it indicates that the two maps are similar. In this study, the overall KAPPA index (Lillesand and Kiefer, 1994) was calculated in order to analyze the degree of similarity between outputs when varying different CA element configurations.

Landscape indices or metrics are quantitative indices that can measure the structure and pattern of a landscape (McGarigal and Marks, 1994; O'Neill et al., 1988). Their origins can be found in the information measures theory and fractal geometry (Mandelbrot, 1983). Recent studies use the term spatial metrics for the analysis of the urban phenomena. It has been shown that spatial metrics have significant advantages when applied in the analysis of heterogeneous urban areas (Parker and Meretsky 2004; Alberti and Waddell 2000; Barnsley and Bair 1997; Herold et al. 2002). In addition, Herold et al. (2003) stated that spatial metrics can be utilized for the accuracy assessment of CA model simulations. In turn, they have applied spatial metrics to test the accuracy of SLEUTH CA model.

In this study, fractal dimension (FD) and class area spatial metrics were employed to analyze the CA model output results. Fractal dimension (FD) illustrates the complexity and the fragmentation of a land-use class patch by a perimeter-area proportion. The derived version of FD, which is called Area Weighted Mean Patch Fractal Dimension (AWMPFD), is used in this study since it eliminates the overestimation of smaller land-use class patches (Milne, 1991). In addition to FD, class area metric compares the change in area of each class by varying radius from the city center. Class area is a measure used to calculate the surface of overall change for each class type. In order to analyze results of spatial metrics, radius zones were created to divide study area into subareas with the increasing radius of 10 km from the city center, and FD value is calculated for each sub-area. To compute metrics, PATCH ANALYST 3.1, an extension of ESRI ArcView GIS software, was used to facilitate the spatial analysis of landscape patches calculations and modeling of attributes associated with patches (Elkie et al., 1999).

3. SIMULATION FRAMEWORK

3.1 Study area and CA model

San Diego region, USA, was chosen for the study area. The digital map data for the study area was obtained from San Diego's Regional Planning Agency public Internet site (SANDAG, 2004). The vector map was rasterized to appropriate spatial resolutions in order to apply the CA model. The raster map represents the digital image of the city classified in nine land use classes, which are: housing area, commercial area, public area, industrial areas, recreational areas, water areas, transportation network, agricultural areas, and vacant land.

Since it is important to analyze different urban land use classes that change over time, housing and commercial areas were chosen as the most dynamic. Land uses such as agricultural areas and public areas, which would impose more constraints on urban growth pattern, were classified as other land use classes. Nonetheless, these areas are considered to be changed to housing throughout the simulation. Water areas and roads represent fixed land use types, which are assumed not to grow or change the location over time. Therefore, the simulation was configured to ensure that urban area can grow in any direction without limitations except for roads and water areas, in which urban growth is assumed to be impossible.

The constrained CA-based simulation model developed by White et al. (1997) was selected due to the fact that it has been widely adapted to simulation of different real cities growth. Simulations based on the CA model were performed by Cellular Automata extension of ESRI ArcView GIS software (Heuegger, 2002). The GIS-based CA simulations of urban growth of San Diego region were used to obtain different model outcomes when varying different CA elements more specifically neighborhood size and type.

3.2 Varying the CA elements

In CA models, the transition of the cell is based on the neighbourhood adopted in the CA simulation, since it affects the cell state change. Finding the areas of influence on the state of the cell is important for realistically modeling the urban land use change. Therefore in this study the variations of neighborhood size and type were examined to measure the sensitivity of these CA elements.

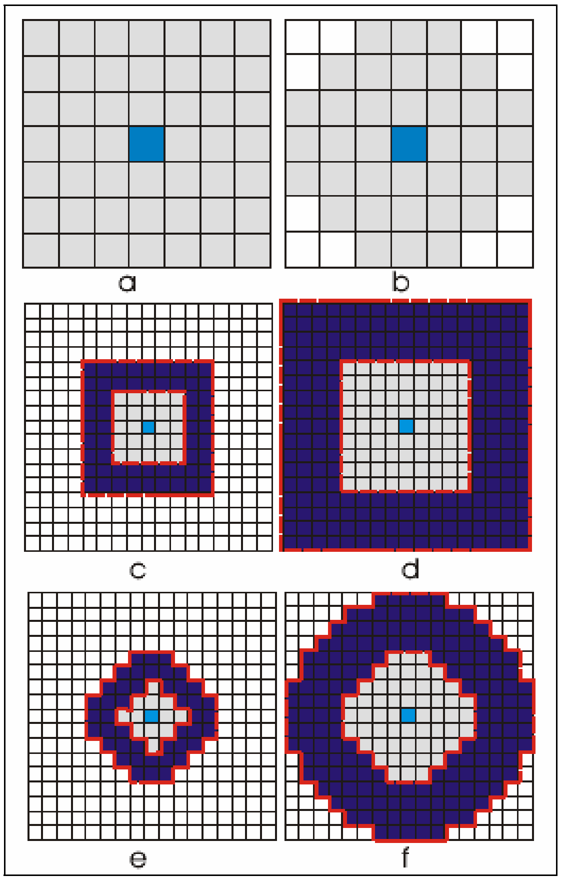

Figure 1. Size and type of neighbourhood a. rectangular neighborhood; b. circular neighbourhood; c. small rectangular neighborhood (inner: 2 cells, outer: 4 cells); d. large rectangular neighborhood (inner: 4 cells, outer: 8 cells); e. small circular neighborhood (inner: 2 cells, outer: 4 cells); f. large circular neighborhood (inner: 4 cells, outer: 8 cells)

3.2.1 Neighborhood type

The commonly used neighborhoods in the case of two-dimensional raster-based CA models are Von Neumann, rectangular – Moore, and circular neighbourhood. In this study the rectangular and circular neighborhood types (Figure 1 a, b). were chosen in order to evaluate the effects of neighbourhood type on simulation results. The impacts of these neighborhood sizes on the CA model outcome were addressed for 50m, 100m, 250m, and 500m spatial resolutions.

3.2.2 Neighborhood size

In the literature, both smaller and larger neighborhood sizes have been applied to the models of urban growth (Figure 1) (Clarke and Gaydos, 1998, Wu 1998, White and Engelen, 1993 and 1997). However, no particular validation on what is the appropriate neighborhood size (e.g. four or six cells in radius) has been made in these urban model applications.

Furthermore, due to the bifractal characteristics of the cities, as stated by White et al. (1997), White and Engelen (1993), and Batty and Longley (1994), an urban area can be divided into two zones. Inner zone corresponds to the area where the urban growth is considered finished or slow dynamics of transformations are expected. Otherwise, growth is considered still ongoing or faster in the outer zone (White et al, 1997; Barredo et al., 2003). In consideration of this fractal structure, in this study two zones are defined in a neighborhood: inner and outer on which development of the cell depends. Both the rectangular (Figure 1 c, d) and circular (Figure 1 e, f) neighborhood types were defined based on these two zones. Two different sizes were specified to represent small and large neighborhood size. The size of the small neighborhood is 2 cells and of the large is 4 cells surrounding the central cell. The outer size is also defined to contain 4 and 8 cells for small and large neighborhood sizes, respectively. Therefore, small and large neighborhood effects were compared with each other. The impacts of these neighborhood sizes on the model outcome were addressed for 50m, 100m, 250m, and 500m spatial resolutions.

4. RESULTS AND DISCUSSION

GIS-based urban CA simulations for San Diego Bay area were produced for temporal interval of 10 years and by using one-year time increment. The results were obtained by varying the neighbourhood size and type for different spatial resolutions and the outcomes were compared using the integrated approaches of sensitivity analysis.

The simulation results of different neighbourhood types were produced. It was observed from the results that the increase of the cell size cause a decrease in KAPPA - from 0.93 for 50m to 0.66 for 500m. This indicates that variation of neighborhood type does affect the discordance of the obtained map outputs. Figure 2 represents the cross-tabulation map of circular and rectangular neighbourhood simulations for 250m cell size. The cross-tabulation map indicates the emergence of commercial land-use areas when both rectangular and circular neighborhoods are used. The class area graph (Figure 3a) depicts the smaller discordances in generated surfaces for commercial land-use type but bigger discordance for housing land-use type. The graph of FD (Figure 3b) does not suggest any major variation.

The pattern of discordance in class area spatial metric graph was detected to have a tendency to increase in the simulation results at 500m cell size. However, no significant discordances were observed from the class area spatial metric graphs for 50m and 100m cell sizes. For all other spatial resolutions, the similar values of the FD were obtained.

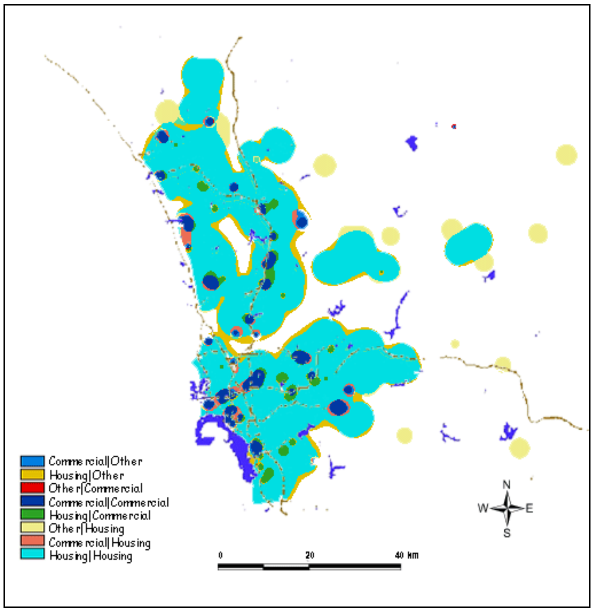

Figure 2. Cross-classification map of circular | rectangular neighbourhood type for 250m resolution

When simulation results of large and small neighbourhoods were assembled, the differences between them were not identified at the spatial resolutions of 50m and 100m respectively. The decrease of KAPPA index from 0.82 to 0.76 for both spatial resolutions was obtained. This indicates that model outcomes i.e. land-use maps obtained are quite similar when varying two neighborhood sizes. In addition, the graphs of FD and class area measures do not depict discordances. From the visual inspection of the cross-classification maps of 50m and 100m resolutions, new housing land-use areas were detected when simulations are performed for the small neighborhood size.

The simulations results point out that with the increase of a cell size KAPPA index decreases from 0.82 to 0.56 for 50m to 500m cell sizes respectively. When the CA model was performed on 250 m cell size, visual inspection of the cross-classification map (Figure 4) showed that large neighborhood size produces bigger changes in commercial land-use class areas than those produced for small neighborhood size. The class area graph (Figure 5a) indicates discordance in both land-use classes when varying neighborhood size. Discordance starts already in the radius of 10km for both commercial and housing land-use. Fractal dimension graph values (Figure 5b) for 250m cell size reveals similar and were stable for all spatial resolutions. For 500m spatial resolution, class area metrics graph demonstrated more discordance in commercial then in housing land-use type.

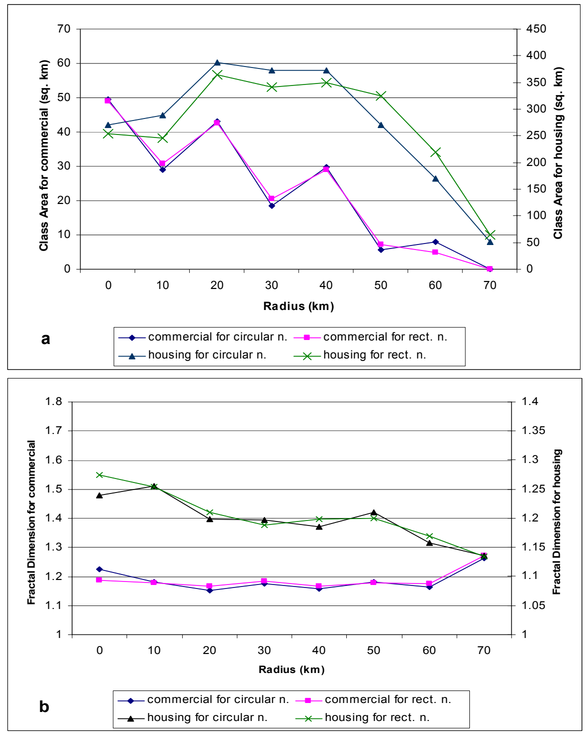

Figure 3. Spatial metrics plots for circular and rectangular neighbourhood type for 250m resolution a. Class area; b. Fractal dimension

5. CONCLUSION

The results of this study indicate that there are impacts of changing CA elements on urban growth modelling especially with respect to changing of neighborhood size and type. The proposed approach represents an exploratory method of sensitivity analysis that can contribute to finding of the appropriate neighborhood size and type for a CA model. The results reported in this study indicate that KAPPA statistic does change for different CA elements when varying spatial resolutions. However, CA model responses are different depending on the spatial metrics approach for neighborhood size and type, and indicate that the discordance in generated land-use classes is related to increase of the spatial resolution. It is worthwhile the efforts to expand this study to include a larger number of different CA element configurations and spatial resolutions as well as spatial metrics.

The selection of proper configurations of CA elements in urban CA models is important, since they can generate different model outputs. Therefore, the SA in CA modeling is a mandatory process to obtain better and more realistic modeling output scenarios. It is vital to understand the limitations of the CA model results pending on the impact of the variation of used CA elements in order to make proper decisions in the land use management process.

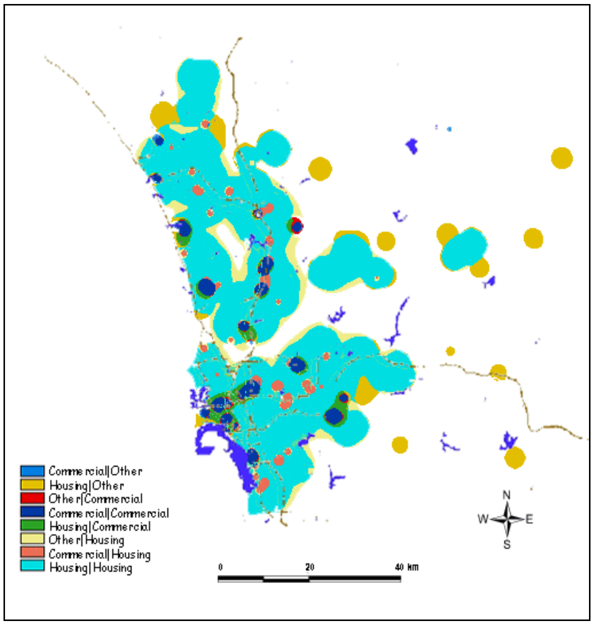

Figure 4. Cross-classification map of large | small neighbourhood size for 250m resolution

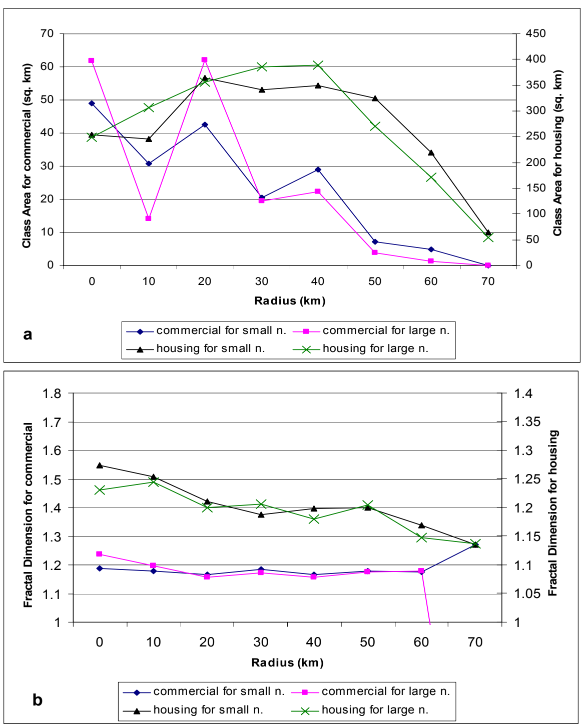

Figure 5. Spatial metrics plots for small and large neighbourhood size for 250m resolution a. Class area; b. Fractal dimension

6. REFERENCES

Alberti, M., Waddell, P., 2000. An integrated urban development and ecological simulation model. *Integrated assessment* 1(3), pp. 215-227.

Allen, P. M., 1997. Cities and regions as evolutionary, complex systems. *Geographical Systems*, 4, pp. 103-130

Barnsley, M. J., Barr, S. L., 1997. Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system. *Computers, Environment and Urban Systems* 21(3-4), pp. 209-225.

Barredo, J. I., Kasanko, M., McCormick, N., Lavalle, C., 2003. Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning* 64(3), pp. 145-160.

Batty, M., Longley, P., 1994. *Fractal cities : a geometry of form and function*. Academic Press: London ; San Diego

Batty, M., 1998. Urban evaluation on the desktop: simulation with the use of extended cellular automata. *Environment and Planning A* 30, pp. 1943-1967.

Chen, Q., Mynett, A. E., 2003. Effects of cell size and configuration in cellular automata based prey-predator modelling. *Simulation Modelling Practice and Theory* 11(7-8), pp. 609-625.

Childress, W. M., Rykiel, E. J., Forsythe, W., Li, B. L., Wu, H. I., 1996. Transition rule complexity in grid-based automata models. *Landscape Ecology*, 11(5), pp. 257-266.

Clarke, K, 2003. The limits of simplicity: toward geocomputational honesty in urban modeling, *Proceedings of the 7th International Conference on Geocomputation*, Southampton, UK

Clarke, K., Gaydos, L.J. 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science* 12(7), pp. 699-714.

Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, pp. 37-46.

Congalton R. G., and Mead, R. A., 1983. A Quantitative method to test for consistency and correctness in photo interpretation. *Photogrammetric Engineering and Remote Sensing* 49, pp. 69 - 74.

Crosetto, M., Tarantola, S., Saltelli, A., 2000. Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture, Ecosystems & Environment* 81(1), pp. 71-79.

Elkie, P. C., Rempel, R. S., Carr, A. P., 1999. Patch Analyst User's Manual: A tool for quantifying landscape structure. http://www2.uah.es/jvl/pa_manual.pdf (accessed 22 Mar. 2004)

Herold, M., Goldstein, N. C., Clarke, K. C., 2003. The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment* 86(3), pp. 286-302.

Herold, M., Scepan, J., Clarke, K.C., 2002. The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. *Environment and Planning A* 34(8), pp. 1443-1458.

Heuegger, M., 2002. The CA extension. <http://srf.tuwien.ac.at/lva/students/ca/> (accessed Mar 22, 2004)

Jenerette, G. D., Wu, J. G., 2001. Analysis and simulation of land-use change in the central Arizona-Phoenix region, USA. *Landscape Ecology* 16(7), pp. 611-626.

Lillesand, T., M., Kiefer, R.W., 1994. *Remote sensing and image interpretation*. Wiley \& Sons, New York

Liu, X., Andersson, C., 2004. Assessing the impact of temporal dynamics on land-use change modeling. *Computers, Environment and Urban Systems* 28(1-2), pp. 107-124.

Mandelbrot, B., 1983. *The Fractal Geometry of Nature*. New York, Freeman.

McGarigal, L. and Marks, B.J., 1994. FRAGSTATS manual: spatial pattern analysis program for quantifying landscape structure. <http://ftp.fsl.orst.edu/pub/fragstats.2.0> (accessed Mar 22, 2004)

Milne, B.T., 1991. Lessons from applying fractal models to landscape patterns*, In Quantitative methods in landscape ecology: the analysis and interpretation of landscape heterogeneity*. Springer-Verlag, New York, pp. 199-235.

O'Neill, R.V. et al., 1988. Indices of landscape pattern. *Landscape ecology*, 1, pp. 153-162.

O'Sullivan, D., Torrens, P. M., 2000. Cellular models of urban systems, *4th Int. Conference on Cellular Automata for Research and Industry*: Karlsruhe, Germany

Parker, D. C., Meretsky, V., 2004. Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. *Agriculture Ecosystems & Environment* 101(2-3), pp. 233-250.

Saltelli, A., Chan, K., and Scott, E. M., 2000. *Sensitivity Analysis*. John Wiley and Sons.

SANDAG, 2004. San Diego's Regional Planning Agency. Maps and GIS: GIS downloads. http://www.sandag.org/index.asp?subclassid=100&fuseaction=home.subclasshome (accessed 22 Mar. 2004)

Straatman, B., White, R., Engelen, G., 2004. Towards an automatic calibration procedure for constrained cellular automata. *Computers, Environment and Urban Systems* 28(1-2), pp. 149-170.

Torrens, P. M., 2000. Cities, cells, and complexity: developing a research agenda for urban geocomputation. In R. J. Abrahart and B.H.Carlisle (Ed.), *5th International Conference on GeoComputation*. "GeoComputation CD-ROM". University of Greenwich, United Kingdom

Wegener, M., 1994. Operational urban models: state of the art. *Journal of American Planning A*., 60, pp. 17-29.

White, R., Engelen, G., Uljee, I., 1997. The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. *Environment and Planning B-Planning & Design* 24(3), pp. 323-343.

White, R., Engelen, G., 1993. Cellular automata and fractal urban form: a cellular modelling approach to the evaluation of urban land-use patterns. *Environment and Planning A* 25, pp. 1175-1199.

White, R., Engelen, G., 1997. Cellular automata as the basis of integrated dynamic regional modelling. *Environment and Planning B-Planning & Design* 24(2), pp. 235-246.

White, R., Engelen, G., 2000. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems* 24(5), pp. 383-400.

Wolfram, S., 2002. *A New Kind of Science*. USA: Wolfram Media Inc, Champaign, IL.

Wu, F., 1998. An experiment on the generic polycentricity of urban growth in a cellular automatic city. *Environment and Planning B-Planning & Design* 25(5), pp. 731-752.

Yeh, A. G. O., Li, X., 2001. A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. *Environment and Planning B-Planning & Design* 28(5), pp. 733-753.

Yeh, A. G. O., Li, X., 2003. Error propagation and model uncertainties of cellular automata in urban simulation with GIS. *7th International Conference on GeoComputation*, University of Southampton, United Kingdom.

7. ACKNOWLEDGEMENTS

This study was supported by Natural Sciences Engineering Research Council (NSERC) of Canada Discovery Grant. Authors are grateful for the conference Travel Grant awarded to Verda Kocabas by the Department of Geography at Simon Fraser University.