A Markov Chain Monte Carlo Cellular Automata Model to Simulate Urban Growth

Ahmed Mustafa[1,2,*], Gen Nishida[2], Ismaïl Saadi[1], Mario Cools[1], Jacques Teller[1]
[1]Local Environment Management & Analysis, ArGEnCo, University of Liège, Belgium
[2]Computer science department, Purdue University, USA
*Email: a.mustafa@ulg.ac.be, a-mustafa@purdue.edu

Abstract— **This paper investigates the potential of a cellular automata (CA) model based on logistic regression (logit) and Markov Chain Monte Carlo (MCMC) to simulate the dynamics of urban growth. The model assesses urbanization likelihood based on (i) a set of urban development driving forces (calibrated based on logit) and (ii) the land-use of neighboring cells (calibrated based on MCMC). An innovative feature of this CA model is the incorporation of MCMC to automatically calibrate the CA neighborhood transition rules. The MCMC based CA model is applied to Wallonia region (Belgium) to simulate urban growth from 1990 to 2000 using Corine Land Cover data (CLC). The outcome of logit model is evaluated by the relative operating characteristic (ROC). The simulated map of 2000 is then validated against 2000 actual map based on cell-to-cell location agreement. The model outcomes are realistic and relatively accurate confirming the effectiveness of the proposed MCMC-CA approach.**

Keywords- cellular automata; Markov chain Monte Carlo; logistic regression.

I. INTRODUCTION

Among the various urbanization modelling approaches, the cellular automata approach has gained popularity for urban modelling. Since the pioneering work of Tobler [1], there has been considerable interest in modifying standard CA models to make them more suitable for urban modelling [2]–[4]. Key challenges in CA are calibrating the transition rules. Early methods for CA calibration were based on trial and error [5] and/or a visual test, to determine the model's parameters. Recently, a variety of automated methods based on statistics [6], machine learning [7], artificial neural networks [8] and optimization algorithms [9] have begun to be widely employed. This paper contributes to such automated calibration methods by using MCMC to calibrate CA neighborhood transition rules.

In this paper, CA model is employed to simulate urban growth based on urbanization probability of a cell according to a number of driving forces of urban growth and state of the cell and its neighbors. Logit method is used to calibrate the driving forces parameters whereas MCMC is used to calibrate neighborhood rules.

This paper is structured as follows. Section II presents the study area. Section III describes the CA model. Section IV gives and discusses the results. Finally, Section V presents our conclusions.

II. STUDY AREA

The study area is located in southern Belgium (Wallonia region). It accounts for 55% of the territory of Belgium with a total area of 16,844 km². The main metropolitan areas are Charleroi, Liège, Mons, and Namur (Fig. 1). They are all characterized by a historical city-center, around which the urban development expanded.

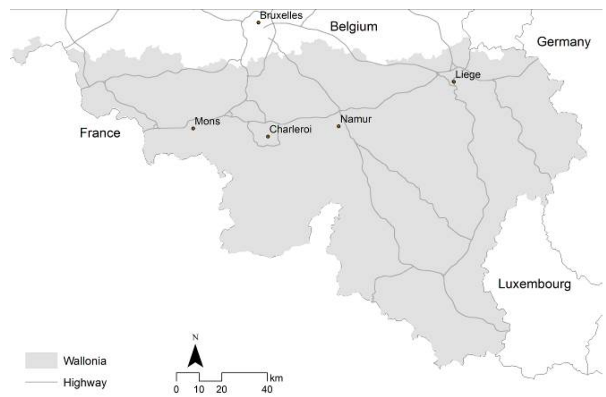

Figure 1. Study area.

The total population in 2010 was 3,498,384 inhabitants that makes up a third of Belgium population.

III. METHODS

The initial state of the simulation starts from land-use in 1990 and proceeds to simulate an urban growth of 2000. The analysis of land-use change is based on the CLC with resolution of 100×100m for the years 1990 and 2000. The 44 classes of CLC datasets have been reclassified into 7 classes (urban lands, arable lands, grasslands, forests, wetlands, water bodies and others). The quantity of change was constrained to the actual quantity of new urban cells in 1990-2000 divided evenly by 10 (the number of years).

The quantity of change is spatially allocated based on two decision rules. The first rule set concerns the main urban growth driving forces, using logit. The second decision rule deals with the neighborhood interactions, using MCMC. The input dependent variable (Y) for logit model is a binary map of the actual non-urban/urban changes within 1990-2000. The independent variables (Xn) are distance to roads, distance to major cities, slope, access to jobs, and zoning. All Xn do not take into account neighboring regions or states.

 ISBN: 978-1-61208-539-5

All Xn are standardized and show a very low degree of multicollinearity (variance inflation factors ranging from 1.01 to 2.76). Logit is calibrated using a random sample of 50,000 cells in order to minimize spatial autocorrelation. MCMC is used to calibrate neighborhood interaction which is arranged in five square distances from the cell. We use the most popular MCMC formulation based on the Metropolis-Hastings algorithm that yields a sequence of samples whose stationary distribution eventually converges to a specified probability density function. The objective function of the MCMC is the maximization of cell-to-cell (CTC) location agreement. The sample with higher score is sampled more than the sample with lower score. In this manner, the algorithm smartly samples from a parameter space, and the global optimal solution can be obtained in a relatively small number of runs.

IV. RESULTS AND DISCUSSIONS

Table I lists logit calibration of driving forces. These coefficients reveal that the location of a new urban development is strongly correlated with the zoning status. Distances to different road classes and cities also play an important role in explaining urban development at a specific location but far less than the zoning status.

TABLE I. COEFFICIENT VALUES OF THE DRIVING FACTORS.

Driving factor	*Coefficient*
Intercept (constant)	-0.9816
Slope	0.0002
Dist to cities	-0.1982
Dist to highway	-0.1962
Dist to major roads	-0.2292
Dist to secondry roads	-0.3185
Dist to local roads	-0.5677
Access to jobs	0.0004
Zoning	2.6809

The weights calibrated by MCMC that defines the neighborhood interactions are illustrated in Fig. 2. The calibration shows that the impact of existing urban lands on new urban development is extremely significant, whereas other land-uses have far less effect than urban land in the immediate neighborhood of the cell. The neighborhood effect is strongest in the immediate neighborhood of the cell, decreases and becomes neutral at a distance of around 5 cells.

The ROC value of the probability map, generated by logit, is 0.78. The cell-to-cell location agreement is 32.75%.

V. CONCLUSION

This paper presents a CA model based on MCMC. The MCMC allows to automate the calibration of the model without losing flexibility and analysis capability. The model is calibrated based on the observed urban growth in 1990–2000 and used to simulate 2010 urban growth in Wallonia.

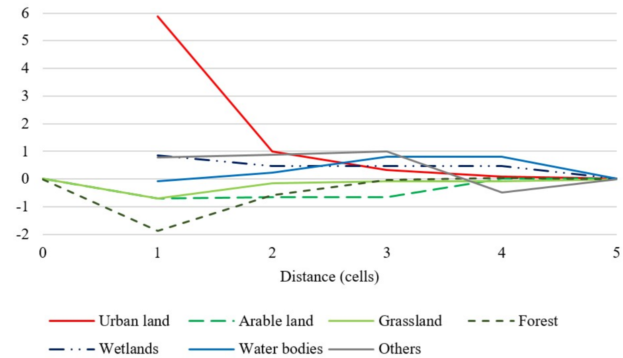

Figure 2. Weighting parameters (Y axis) that represent the interaction between an urban cell and other land-uses

The cell-to-cell location agreement, which measures for the new urban cells in the 1990-2000, is similar to the numbers reported for the best performing CA urban models. The results confirm that MCMC is a method with great potential for urban CA calibration.

ACKNOWLEDGMENT

The research was funded through the ARC grant for Concerted Research Actions, financed by the Wallonia-Brussels Federation.

REFERENCES

[1] W. R. Tobler, “Cellular Geography,” in *Philosophy in Geography*, S. Gale and G. Olsson, Eds. Springer Netherlands, 1979, pp. 379–386.

[2] K. C. Clarke and L. J. Gaydos, “Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore,” *Int. J. Geogr. Inf. Sci.*, vol. 12, no. 7, pp. 699–714, 1998.

[3] D. Guan, et al., “Modeling urban land use change by the integration of cellular automaton and Markov model,” *Ecol. Model.*, vol. 222, no. 20–22, pp. 3761–3772, 2011.

[4] A. Mustafa, M. Cools, I. Saadi, and J. Teller, “Urban Development as a Continuum: A Multinomial Logistic Regression Approach,” in *Computational Science and Its Applications -- ICCSA 2015*, O. Gervasi, B. Murgante, S. Misra, M. L. Gavrilova, A. M. A. C. Rocha, C. Torre, D. Taniar, and B. O. Apduhan, Eds. Springer International Publishing, 2015, pp. 729–744.

[5] R. White and G. Engelen, “Cellular Automata as the Basis of Integrated Dynamic Regional Modelling,” *Environ. Plan. B Plan. Des.*, vol. 24, no. 2, pp. 235–246, 1997.

[6] A. M. García, I. Santé, M. Boullón, and R. Crecente, “Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm. Application to a small urban settlement of NW Spain,” *Int. J. Geogr. Inf. Sci.*, vol. 27, no. 8, pp. 1593–1611, 2013.

[7] A. Rienow and R. Goetzke, “Supporting SLEUTH – Enhancing a cellular automaton with support vector machines for urban growth modeling,” *Comput. Environ. Urban Syst.*, vol. 49, pp. 66–81, 2015.

[8] S. Berberoğlu, A. Akın, and K. C. Clarke, “Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach,” *Landsc. Urban Plan.*, vol. 153, pp. 11–27, 2016.

[9] K. Al-Ahmadi, L. See, A. Heppenstall, and J. Hogg, “Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia,” *Ecol. Complex.*, vol. 6, no. 2, pp. 80–101, 2009.

 ISBN: 978-1-61208-539-5