Computers, Environment and Urban Systems 33 (2009) 44–54

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Implementation of a dynamic neighborhood in a land-use vector-based cellular automata model

Niandry Moreno [a,b], Fang Wang [a], Danielle J. Marceau [a,*]

[a] Geocomputing Laboratory, Department of Geomatics Engineering, University of Calgary, 2500 University Drive N.W, Calgary, AB, Canada T2N 1N4
[b] Centro de Simulación y Modelos (CESIMO), Universidad de Los Andes, Núcleo La Hechicera – Edif. B – 3er Nivel – Ala Norte, Mérida 5101, Venezuela

A R T I C L E I N F O

Keywords:
Vector-based cellular automata
Dynamic neighborhood
Land-use/land-cover changes

A B S T R A C T

While cellular automata (CA) models have been increasingly used over the last decades to simulate a wide range of spatial phenomena, recent studies have illustrated that they are sensitive to cell size and neighborhood configuration. In this paper, a new vector-based cellular automata (VecGCA) model is described to overcome the scale sensitivity of the raster-based CA models. VecGCA represents space as a collection of geographic objects of irregular shape and size corresponding to real-world entities. The neighborhood includes the whole geographic space; it is dynamic and specific to each geographic object. Two objects are neighbors if they are separated by objects whose states favor the land-use transition between them. The shape and area of the geographic objects change through time according to a transition function that incorporates the influence of the neighbors on the specific geographic object. The model was used to simulate land-use/land cover changes in two regions of different landscape complexity, in Quebec and Alberta, Canada. The results revealed that VecGCA produces realistic spatial patterns similar to reference land-use maps. The space definition removes the dependency of the model to cell size while the dynamic neighborhood removes the rigid, arbitrarily defined zone of influence around each geographic object.

1. Introduction

Cellular automata (CA) are dynamic models originally conceived by Ulam and Von Neumann in the 1940s to provide a formal framework for investigating the behavior of complex systems (Von Neumann & Burks, 1966). A basic CA consists of five components: (1) a grid space on which the model acts, (2) cell states in the grid space, (3) transition rules that determine the spatial dynamic process, (4) a neighborhood that influences the central cell, and (5) time steps. It has been demonstrated that CA can capture complex spatially distributed processes, as well as providing insights into a wide variety of local behaviors and global patterns (Wolfram, 1984). In addition, temporal and spatial complexities of many phenomena can be well modeled by properly defining transition rules in CA models. Therefore, during the last decade, CA models have been increasingly used for simulating various spatial phenomena including land-use and land-cover changes (Almeida et al., 2003; Li & Yeh, 2002; Ménard & Marceau, 2007; Wu, 2002), urban growth (Almeida, Gleriani, Castejon, & Soares-Filho, 2008; Batty, Xie, & Sun, 1999; Dietzel & Clarke, 2006; White, Engelen, & Uljee, 2000), fire propagation (Berjak & Hearne, 2002; Favier, Chave, Fabing, Schwartz, & Dubois, 2004; Yassemi, Dragicevic, & Schmidt, 2008), species competition (Arii & Parrott, 2006; Chen, Mynett, & Minns, 2002; Matsinos & Troumbis, 2002; Rietkerk, Dekker, Ruiter, & Koppel, 2004), and traffic flow (Sun & Wang, 2007; Wahle, Neubert, Esser, & Schreckenberg, 2001).

In these applications of CA models, geographic space is typically represented as a grid of regular cells and the neighborhood is defined as a collection of cells based on physical adjacency. Recent studies have demonstrated that such cell-based CA models are sensitive to the modifiable units used in the models; in other words, the modeling results may vary according to the cell size and the neighborhood configuration. Jenerette and Wu (2001) have used two different cell resolutions in their CA model to study urban expansion and have shown that it generates significantly different land-use patterns. Chen and Mynett (2003) have investigated the impact of cell size and neighborhood configuration in a raster-based CA prey–predator model and observed that they affect both the resulting spatial patterns and the system stability. Jantz and Goetz (2005) have examined the modeling results of a widely used CA based urban model, SLEUTH, in response to different cell sizes and have indicated that the cell size at which the land-use data are represented can impact the quantification of the land-use patterns and the ability of the model to replicate spatial patterns. Ménard and Marceau (2005), Kocabas and Dragicevic (2006) have

* Corresponding author. Tel.: +1 403 220 5314; fax: +1 403 284 1980.
E-mail address: dmarceau@ucalgary.ca (D.J. Marceau).

0198-9715/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2008.09.008

also demonstrated that their raster-based land-use and urban growth CA are sensitive to different cell resolutions and neighborhood configurations.

One approach to eliminate or minimize the scale sensitivity consists in representing space using an irregular tessellation rather than the traditional regular grid. In the initial work undertaken towards this objective, Voronoi diagrams have been used (Flache & Hegselmann, 2001; Shi & Pang, 2000). The basic idea of constructing a CA model utilizing a Voronoi diagram is to break down space into Voronoi polygons using spatial objects (such as points, lines and areas) as generators and to define the neighborhood of each spatial object as the objects that share a common Voronoi boundary with it. The change of state of each spatial object is based on the attributes of its neighbors. It has been shown that Voronoi CA are able to depict the spatial interactions between any irregular spatial objects and to simulate their dynamics. Another irregular space representation is the Delaunay triangle network used in an urban growth CA proposed by Semboloni (2000). The neighborhood of each triangle is defined by its adjacent triangles; new triangles are generated by additional nodes in the Delaunay triangle network. O'Sullivan (2001a,b) developed an irregular CA where space is represented as a planar graph composed of vertices and edges. Each vertex stands for an object and has one of a collection of states (such as a type of building), whose neighborhood is defined as a set of vertices linked to it by edges. This graph CA model simulates the changes of state of the vertices. All these studies have demonstrated the usability of the irregular tessellation in CA models. One limitation of these approaches however, is that the polygons are generated automatically and might not correspond to the real-world entities composing the landscape a user would perceive as meaningful. In addition, the neighborhood definition is rigid and limited since it relies only on topology (White & Engelen, 2000).

In the search for improving space representation in CA models, Benenson, Omer, and Hatna (2002) applied an entity-based approach to simulate urban residential dynamics, where infrastructure elements, such as land parcels and houses are directly described, while the neighborhood still remains defined by Voronoi polygons. Torrens and Benenson (2005) proposed the geographic automata system (GAS) that combines characteristics of both CA and multi-agent models, which incorporates irregular vector objects as automata to represent real-world entities such as roads, buildings and parks composing an urban system. This framework is considered more spatially realistic than the traditional raster-based CA models. Stevens, Dragicevic, and Rothley (2007) set up their CA model based on irregular cadastral land parcels; the neighborhood is defined by the adjacent parcels, parcels accessible from a road, and parcels within a buffer. Although this progress in space representation contributes to the development of more flexible and spatially realistic CA models, the geometry of the objects remains invariant, that is, the models do not allow irregular growth or decrease as indicated by the change of shape and size of the objects. This is an important limitation since such changes are prevalent in the real world.

Pursuing the work initiated by Torrens and Benenson (2005), Hammam, Moore, and Whigham (2007) recently introduced vector agents (VA) to overcome the limitation of the GAS model in which the geometry of the objects remains invariant through the time frame of the simulations. A vector agent is goal-oriented, adaptable, defined by a Euclidian geometry, and able to change its own geometry while interacting with other agents in its neighborhood using a set of rules. This approach allows real-world objects to be naturally defined and to explicitly control their geometry, including their location. It has been shown that the spatial patterns simulated by the vector agents are similar to land-use parcels and urban patches as they appear in the real world. However, at this point, the vector agents are predominately driven by geometry and the transition rules do not explicitly capture the driving factors responsible for the dynamic geometric changes.

In attempts to incorporate these two key characteristics: a vector representation of real-world entities and their geometrical transformation, a vector-based CA model (VecGCA) has been recently proposed (Moreno & Marceau, 2006; Moreno, Ménard, & Marceau, 2008). In VecGCA, space is represented as a collection of interconnected irregular geographic objects, corresponding to real-world entities, and the neighborhood is defined as an external buffer around each geographic object that corresponds to an influence area. The neighbors are all the objects located within this area. The geographic objects evolve through time according to a transition function that determines their change of shape and area, which depends on the area of the neighbors within the neighborhood and their influence on the specific geographic object. It has been shown that VecGCA produces more realistic spatial patterns than those generated by a raster-based CA model, while the space definition removes the dependency of the model to cell size. However, a sensitivity analysis has demonstrated that VecGCA remains sensitive to the neighborhood configuration and that this sensitivity depends on the landscape configuration of the study area (Marceau & Moreno, 2008; Moreno & Marceau, 2007).

Neighborhood configuration is a key component in a CA model since it represents the zone of influence determining the change of state of a cell in a raster-based CA model, or the state of a polygon in a vector-based CA model (Verburg, Nijs, Eck, Visser, & Jong, 2004). The neighborhood in conventional raster-based CA is usually defined as a set of geometrically nearest cells to a central cell and it is the same for all cells that compose the space. In vector-based CA, where space is represented as a variety of interconnected irregular polygons, the neighborhood is typically defined based on connectivity, adjacency, or distance to the central object (Torrens & Benenson, 2005). Various alternative neighborhoods have been investigated for vector-based CA, including Delaunay triangle links (Semboloni, 2000), planar graphs (O'Sullivan, 2001a,b) and Voronoi polygons (Flache & Hegselmann, 2001; Hu & Li, 2004; Shi & Pang, 2000). These neighborhoods are usually extended and there is no rule to identify the proper zone of influence, which is determined by a mixture of data availability, intuition, computing considerations and trial and error (Couclelis, 1985). This explains the sensitivity of CA models to neighborhood configuration.

To overcome or limit the problem of neighborhood sensitivity, various approaches have been proposed. White and Engelen (1993) have introduced an enlarged circular neighborhood which incorporates a weighting function depending on distance. The circular neighborhood treats all directions equally, so it overcomes the source of differences in the distance between the neighborhood and the central cell. However, the radius of the circle still affects the modeling results. In the geo-algebra proposed by Takeyama and Couclelis (1997), the neighborhood can be spatially variant or invariant. Verburg et al. (2004) have analyzed neighborhood characteristics of land-use patterns and shown that neighborhood interactions among land-use types differ in different parts of the study area; they have suggested parameterizing CA models using derived neighborhood characteristics. Stewart-Cox, Britton, and Mogie (2005) have used two kinds of neighborhood in their raster-based CA model, global and local, for simulating pollination and seed setting processes. The separation of the neighborhoods allows the differentiation of the two processes, which is more realistic than using only one uniform neighborhood. However, the configuration of the local neighborhood still influences the modeling results. Ménard and Marceau (2005) and Kocabas and Dragicevic (2006) have proposed to conduct a sensitivity analysis to the neighborhood configuration to gain knowledge before defining a specific neighborhood in a CA model.

A solution to the problem of neighborhood sensitivity is to define a separate neighborhood and transition rule for each cell or object, at each point of time, and for every likely set of external events (Couclelis, 1985). In this paper, we propose a dynamic neighborhood where the neighbors are defined during the modeling process and may vary for each object at each time step. In this new definition, two objects are neighbors if they are separated by 0, 1 or more objects whose states favor the transition from the state of one object to the other. This dynamic neighborhood has been implemented in VecGCA and tested to simulate land-use changes in two regions of different landscape complexity in southern Quebec and southern Alberta, Canada.

In the following section, the dynamic neighborhood is described and the update of the VecGCA model is presented. Section 3 describes the study areas and the methodology used for the definition of the land-use VecGCA models. Finally, results and conclusion are presented in Sections 4 and 5.

2. A dynamic neighborhood definition

The dynamic neighborhood proposed in this paper is a neighborhood that changes through time; there is no distance or fixed area that delineates it and it is specific to each geographic object. The neighborhood includes the whole geographic space, and the neighborhood relationships between two objects depend on the properties of each geographic object. Objects A and B are neighbors if they are adjacent or separated by other objects which states are favorable to the change of state from A to B. A n x m binary matrix describes if a state X is favorable to the transition from the state Y to Z, where n is the number of possible states of a geographic object and m is the number of possible transitions in the model. In this matrix, the 1 values indicate that a state X is favorable to a transition and the 0 values indicate the opposite. The number of intermediate objects between two objects A and B can be 0, 1 or any number. The only required condition is that the state of these intermediate objects (which can be in different states) is favorable to the transition from the state of A to the state of B.

For example, let's suppose a geographic space as defined in Fig. 1, composed of six geographic objects that represent patches of different land uses/land covers (U = undeveloped land, D = developed land, P = park, C = commercial land and W = water). Let's suppose that the possible transitions are U to D, U to C, U to P, P to D, and P to C. Let's suppose that the matrix M represents the favorable states to the transitions in the model.

$$M = \begin{array}{c} \\ U \to D \\ U \to C \\ U \to P \\ P \to D \\ P \to C \end{array} \begin{array}{c} \begin{array}{ccccc} U & D & C & P & W \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

This matrix indicates for example, that commercial land is favorable to the transition from undeveloped to developed land, from undeveloped land to commercial land and from park to developed land. Using this matrix, we can say that the neighbors of object A in Fig. 1 are the adjacent objects B, C, D and the non adjacent object E. E is neighbor of A in two cases: first, E and A are separated by B that is a commercial land and this state is favorable to the transition from undeveloped to developed land; second, E and A are separated by the objects C and D that are in the states water and park, respectively, and these states are favorable to the transition from undeveloped to developed land. However, we cannot say that A is neighbor of E, because the transition from developed to undeveloped land is not possible.

This matrix is obtained from the analysis of historical spatial data that reveals when a change of state of a geographical object has occurred due to the influence of its non adjacent neighbors and what are the states of these intermediate objects. The procedure to calculate this matrix is described in Section 3.

In the previous version of the VecGCA model (Marceau & Moreno, 2008; Moreno et al., 2008), each neighbor exerts an influence on the central object producing its change of shape if this influence is higher than a threshold value (λ) that represents the resistance of the geographic object to change state for the state of its neighbor. An exponential function was used to define the influence value, which varies between 0 and 1, where 0 indicates no influence and 1 the highest influence and it is constant on the whole surface of the central object. This function depends on a parameter α which is defined in function of the variables that control the influence value of a geographic object on another. The transition function quantifies the area of the central object that changes its state for the neighbor's state. A geometrical transfor-

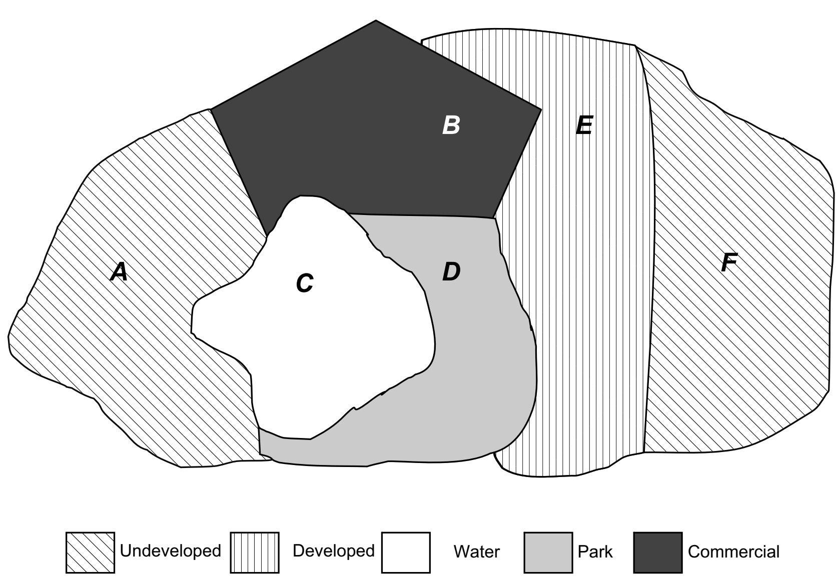

Fig. 1. Geographic space composed of six geographic objects; the geographic object A has four neighbors: the adjacent objects B, C and D, and the non adjacent object E.

mation procedure is performed to reduce this area from the region nearest to the corresponding neighbor.

With the new neighborhood definition, the influence value is variable on the surface of the central object. Fig. 2 presents the graph of the new proposed influence function. The influence value increases when the neighbor is closer to the central object; the maximum value (g_{max}) is obtained in the object's border and decreases inside the object. If g_{max} is higher than λ, then the geometrical transformation procedure is performed. The area that changes is not specified in the transition function. Instead, the transition function determines exactly the buffer size that is used in the geometrical transformation procedure to take a portion of the central object and add it to the corresponding neighbor. This buffer size is calculated from the point where the influence value inside the central object is equal to λ (corresponding to P1 in Fig. 2). Therefore, the transition function is derived from the influence function. Table 1 presents a comparison of the definition of VecGCA for the previous and new versions.

A combined function is proposed to define the influence function, where the influence outside the geographic object is represented as the exponential function used in the previous version of the model (Marceau & Moreno, 2008), and an inverse exponential function defines the influence value inside the geographic object (Eq. (1)). The parameter α is defined as a function that depends on the factors that determine the influence of a neighbor to the central object: the transition probability from the neighbor's state and the central object's state, the common border between the neighbor and the central object (if they are adjacent) or the common border between the neighbor and the object that connects it to the central object, and the minimum distance between the neighbor and the central object (Eq. (2))

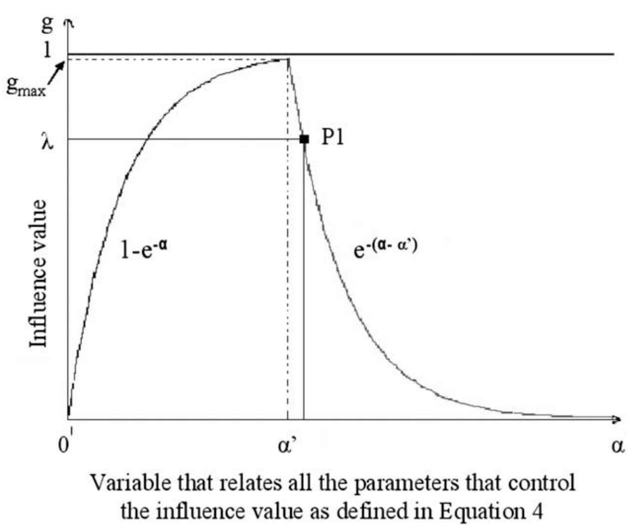

Fig. 2. Influence function used to implement the dynamic neighborhood.

$$g_{AB} = \begin{cases} 1 - e^{\alpha_{AB}} & \text{if } 0 \leqslant \alpha \leqslant \alpha' \\ e^{-(\alpha_{AB}-\alpha_{AB'})} & \text{if } \alpha > \alpha' \end{cases} \tag{1}$$

where g_{AB} is the influence of A on B, α_{AB} is defined in Eq. (2), and α'_{AB} is the value α_{AB} on the border.

$$\alpha_{AB} = p^{1/2}\left(\frac{a_A}{a_B} + \frac{cb}{b_B} + e^{-d_{min}}\right) \tag{2}$$

where p is the transition probability from the B's state to A's state, a_A is the A's area, a_B is the B's area, cb is the common border between A and B, b_B is the B's perimeter, and d_{min} is the minimum distance between A and B.

If A and B are not adjacent neighbors, cb is the common border between A and another object adjacent to A that was used in the analysis of the favorable states to the transition from A to B, and b_B is replaced by the perimeter of this other object.

When the neighbor is larger than the central object (a_A/a_B is higher than 1), the α_{AB} value is high and it produces a high influence value. Therefore, to control the effect of the neighbor's area, a normalization of a_A/a_B is introduced to regulate its value between 0 and 1, which limits the α value between 0 and 3 and does not produce the influence values close to 1 for all neighbors of large area. Without this normalization, the neighbor's area would control the influence function. To account for the normalization, Eq. (2) is rewritten as Eq. (3)

$$\alpha_{AB} = p^{1/2}\left(\frac{\frac{a_A}{a_B}}{\frac{a_{max}}{a_{min}}} + \frac{cb}{b_B} + e^{-d_{min}}\right), \quad 0 \leqslant \alpha \leqslant 3 \tag{3}$$

where a_{max} is the largest object's area within the whole geographic space, and a_{min} is the smallest object's area.

The transition function calculates the value of d_{min} for which the influence value is equal to the threshold value (λ) when α_{AB} is higher than α_{AB}', that is the influence inside the object. From Eq. (1), α_{AB} can be defined by Eq. (4), and by replacing (3) in (4)

$$\alpha_{AB} = \alpha'_{AB} - \mathrm{Ln}(\lambda) \tag{4}$$

$$p^{1/2}\left(\frac{\frac{a_A}{a_B}}{\frac{a_{max}}{a_{min}}} + \frac{cb}{b_B} + e^{-d_{min}}\right) = \alpha'_{AB} - \mathrm{Ln}(\lambda) \tag{5}$$

d_{min} can be found from Eq. (5) and defines the transition function (Eq. (6))

$$f_{AB} = -\mathrm{Ln}\left[\frac{1}{p^{1/2}}(\alpha'_{AB} - \mathrm{Ln}(\lambda)) - \frac{\frac{a_A}{a_B}}{\frac{a_{max}}{a_{min}}} - \frac{cb}{c_B}\right] \tag{6}$$

where, f_{AB} is the transition function that determines the size of the buffer that is built around A to take a portion of B.

To include stochasticity in the model that represents the influence of driving factors that are not included in the transition function, a random variable (β) (limited between 0 and 1) was introduced in Eqs. (6) and (7)

$$f_{AB} = -\beta^* \ln\left[\frac{1}{p^{1/2}}(\alpha'_{AB} - \mathrm{Ln}(\lambda)) - \frac{\frac{a_A}{a_B}}{\frac{a_{max}}{a_{min}}} - \frac{cb}{b_B}\right] \tag{7}$$

where β is a random variable limited between 0 and 1.

Table 1
Differences between the previous version of VecGCA and the new version proposed

Previous version of VecGCA. Buffer neighborhood	New version of VecGCA. Dynamic neighborhood
The neighborhood is defined as a buffer around each geographic object	The neighborhood is the whole geographic space
An object A is a neighbor of B if it is (partially or totally) within a neighborhood delineated by the user-defined buffer	An object A is a neighbor of B if they are adjacent or separated by other objects which states are favorable to the change of state from B to A
The influence of a neighbor on the central object is constant over the whole surface of the central object	The influence of a neighbor on the central object is variable over the surface of the central object
The transition function quantifies the area of an object that changes state for the state of its neighbor. The buffer size built around the neighbor to take a portion of the central object is automatically adjusted in the geometrical transformation procedure	The transition function calculates the buffer size that is used in the geometrical transformation procedure to take a portion of the central object and add it to the corresponding neighbor

3. Methodology

To test the new dynamic neighborhood and to determine if it overcomes the sensitivity to the neighborhood size that is present in the previous version of VecGCA, two land-use VecGCA models were built for two regions of varying landscape configuration in Canada. The results were compared with the ones obtained using the previous version of VecGCA. The implemented models are simple land-use change models based on transition probabilities to ensure that the obtained results are generated by the logic implemented in the conceptual VecGCA model and not by other external driving factors that would increase the complexity of the models.

3.1. Study areas

The first study area is the Maskoutains region, an agricultural region covering 1312 km^2, located in Southern Quebec, Canada. Data used for the study include two land-use maps originating from Landsat Thematic Mapper images acquired in 1999 and 2002, at 30 m spatial resolution (Soucy-Gonthier et al. 2003). The landscape is characterized by small forest patches within a large agriculture matrix. The second study area is a sub-region of the Elbow river watershed, located in Southwest Alberta, Canada, which covers approximately 731 km^2. This subregion comprises the land area drained by the Elbow River and its tributaries, excluding the portion corresponding to the Alberta's Rocky Mountains. Three land-use maps generated from Landsat Thematic Mapper images acquired in the summer of 1996, 2001 and 2006 were also available for that region. This landscape is more fragmented and composed of numerous polygons of smaller extent compared to the Maskoutains region.

3.2. The land-use VecGCA model

In VecGCA, space is represented as a collection of patches of different land uses, where each patch corresponds to a polygon of the vector land-use map of the study area. The influence function and the transition function are given in Eqs. (1) and (7), respectively.

The transition probabilities are obtained from the comparison of two land-use maps of different dates and are calculated as the area that changes from the state X at time t to the state Y at $t+1$ divided by the total area that changes from the state X to all other states at $t+1$ (including Y). The land-use maps of 1999 and 2002 were used for the Maskoutains region, while the maps corresponding to 1996 and 2001 were used for the Elbow river watershed. The transition probabilities for a temporal resolution of one year were calculated using the exponential method presented by Yeh and Li (2006).

The threshold value (λ) is used as a condition to execute the geometrical transformation procedure. It represents the resistance of the geographic object A to change its state for the state of its neighbor B. These threshold values were calculated using the method described in Moreno et al. (2008).

Finally, the binary matrix that describes if a state X is favorable to the transition from the state Y to Z is obtained from the comparison of two historical vector land-use maps. This procedure checks when an object A changes, in a portion or in the totality of its surface, from the state Y to the state Z without having adjacent objects in the Z state. The procedure is looking for the closest object in the state Z; the objects that separate the object A to this closest object are considered intermediate objects, and the states of these intermediate objects are considered favorable states to the transition from Y to Z. The result is a binary matrix where 1 for the state Y to the transition X → Z indicates that in the historical data there is at least one case where an object has changed from the state X to Z while being separated by an object in state Y; 0 indicates that this case has never been found in the historical data.

3.3. Model simulations

Five simulations on each study area were performed to compare the results of the new dynamic neighborhood with the results obtained using the previous version of VecGCA (for four different neighborhood sizes: 10 m, 30 m, 60 m and 120 m) and the reference land-use/land-cover maps of the two areas. The simulations were performed from 1999 to 2002 for the Maskoutains region and the results were compared with the 2002 land-use/land-cover map. For the Elbow river watershed, the simulations were performed from 1996 to 2006, and the results obtained were compared with the 2001 and 2006 land-use/land-cover maps. The temporal resolution of the simulations is one year. Using a dynamic neighborhood, an additional simulation for 1996–2016 was performed to assess the stability of the model over a longer simulation period.

Since a pseudo-random number generator was used to implement the random variable (β) in the transition function, five replicates of each simulation were performed for each model and the mean was calculated. VecGCA was implemented in Java and it uses two additional libraries: OpenMap library (OpenMap, 2005) for the handling and display of shape files, and JTS Topology Suite (JTS, 2004) for the handling of geometric objects (points, lines, polygons, polylines), buffer construction and geometric operations (intersection, difference, union, etc).

4. Results

When using the previous version of VecGCA in the Maskoutains region, the results reveal that for the neighborhood sizes of 10 m, 30 m and 60 m, the simulated proportion of forested and agricultural land for 2002 differs in less than 2% of the proportion calculated from the 2002 land-use/land-cover map. When using the neighborhood size of 120 m, the difference between the simulated land-use/land-cover proportions for 2002 and the calculated proportions from the 2002 map is higher and might exceed 6% (Table 2). A spatial overlay of the 2002 land-use map and the simulation results for 2002 reveals that when using a neighborhood size of 10 m, 98.72% of agricultural area generated by VecGCA coincides with the agricultural area present in the 2002 land-use map, in comparison with 93.21% when a neighborhood size of 120 m is used (Table 2). For the forested area, the obtained results are opposite; the coincident proportion is smallest when the neighborhood size is 10 m (Table 2). These results can be explained by the fact that when the neighborhood size increases, small forested patches disappear due to the high influence of the agricultural area producing a decrease of the forested area and a reduction of the number of forested patches.

The neighborhood size has a less pronounced impact on the simulation outcomes in the Elbow river watershed. The simulation outcomes were compared with the proportion of land use and spatial distribution of patches calculated from two land-use maps corresponding to 2001 and 2006. In both cases, the results obtained are similar; the proportions of agricultural and urban areas slightly increase with the neighborhood size while it is the opposite for the proportion of forested area (Table 3). When the 2001 simulation outcomes are compared with the 2001 reference land-use map, the largest difference (3.12% points) is found for the agricultural land-use class using a neighborhood size of 120 m. For the other classes, the simulated and calculated proportions differ by less than 3% points when using different neighborhood sizes. These re-

Table 2
Proportion of land uses for 2002 produced by VecGCA and percentage of coincidence of the simulation outcomes with the 2002 land-use map for the Maskoutains region using the neighborhood defined as a buffer

		Proportion of simulated land uses. Buffer neighborhood				2002 simulated land uses – 2002 reference land-use map	% Coincident with the ref. land-use map
		1999	2000	2001	2002		
Forest	Neigh. 10 m	16.57	16.46	16.40	16.34	1.51	79.49
	Neigh. 30 m	16.57	16.17	15.86	15.52	0.69	83.75
	Neigh. 60 m	16.57	16.02	15.31	14.20	−0.63	84.57
	Neigh. 120 m	16.57	12.86	9.89	8.45	−6.38	87.66
	Ref. land-use map	16.57	–	–	14.83	–	–
Agriculture	Neigh. 10 m	80.70	80.81	80.87	80.93	−1.52	98.72
	Neigh. 30 m	80.70	81.10	81.42	81.75	−0.70	97.12
	Neigh. 60 m	80.70	81.25	81.96	83.07	0.62	96.46
	Neigh. 120 m	80.70	84.41	87.38	88.82	6.37	93.21
	Ref. land-use map	80.70	–	–	82.45	–	–

Table 3
Proportion of land uses produced by VecGCA and percentage of coincidence of the simulation outcomes with the reference land-use maps for the Elbow river watershed using the neighborhood defined as a buffer

		Proportion of simulated land uses. Buffer neighborhood			Simulated land-use – ref. land-use map		% Coincident with the ref. land-use map	
		1996	2001	2006	2001	2006	2001	2006
Forest	Neigh. 10 m	45.09	44.93	44.83	−1.56	−0.18	93.78	95.58
	Neigh. 30 m	45.09	44.48	44.23	−2.01	−0.78	95.41	96.52
	Neigh. 60 m	45.09	44.09	43.84	−2.40	−1.17	96.04	97.32
	Neigh. 120 m	45.09	43.82	43.56	−2.67	−1.45	96.45	97.45
	Ref. land-use map	45.09	46.49	45.01	–	–	–	–
Agriculture	Neigh. 10 m	25.86	25.86	25.88	2.51	0.9	89.54	88.34
	Neigh. 30 m	25.86	26.02	26.10	2.67	1.12	88.11	87.28
	Neigh. 60 m	25.86	26.24	26.29	2.89	1.31	87.41	85.79
	Neigh. 120 m	25.86	26.47	26.53	3.12	1.55	86.98	82.63
	Ref. land-use map	25.86	23.35	24.98	–	–	–	–
Urban	Neigh. 10 m	4.83	4.99	5.06	−0.41	−1.12	94.75	95.41
	Neigh. 30 m	4.83	5.28	5.44	−0.12	−0.74	92.44	87.69
	Neigh. 60 m	4.83	5.44	5.64	0.04	−0.54	89.54	85.19
	Neigh. 120 m	4.83	5.48	5.69	0.08	−0.49	87.89	78.31
	Ref. land-use map	4.83	5.40	6.18	–		–	–

sults can be explained by the fact that in the previous version of VecGCA, the influence and the transition function are directly proportional to the neighbors' area within the neighborhood. The neighbors' area varies with the neighborhood size and the landscape configuration. In the Maskoutains region, the majority of the objects have only one neighbor; they are small forested patches located inside a large matrix of agricultural land (Fig. 3). The influence of the agriculture matrix on the small forested patches and the area that changes from forest to agriculture (calculated from the transition function) increase when the neighborhood size increases.

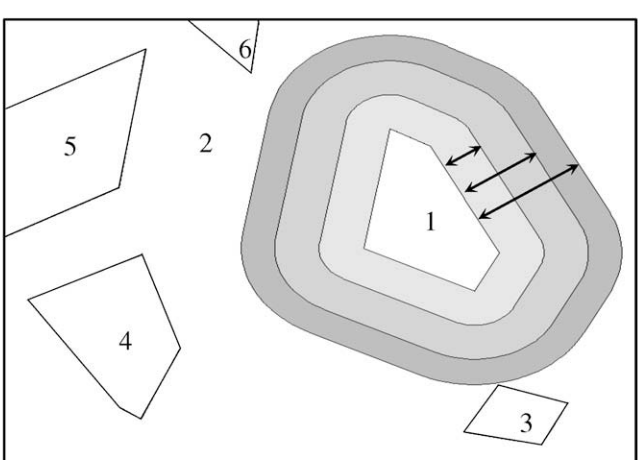

Fig. 3. Schematic representation of the landscape configuration of the Maskoutains region where the objects have only one neighbor for different neighborhood sizes.

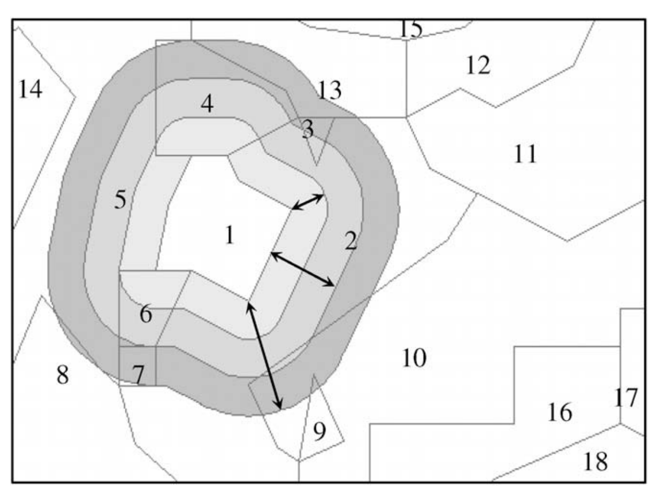

Fig. 4. Schematic representation of the Elbow river landscape configuration where the objects have several neighbors; when the neighborhood size increases the number of neighbors and the neighbors' area within the neighborhood also increase.

The landscape configuration for the Elbow river watershed is different. Each geographic object has several neighbors of different states where an increase of the neighborhood size can produce an increase of the neighbors' area within the neighborhood as well as an increase in the number of neighbors (Fig. 4). The increase of the neighbors' area is less important than the one produced in a landscape composed of a large matrix containing numerous small objects, such as in the Maskoutains region. Consequently, the increase of influence and the area to change from the object' state to the neighbor's state is often not significant when the neighborhood size increases. When the number of neighbors increases, the influence and the area to change for the neighbors present in the neighborhood do not increase significantly due to the fact that these neighbors are distant geographic objects separated by other objects. Therefore, in this landscape configuration the simulation outcomes are less sensitive to the neighborhood size.

Using the dynamic neighborhood, the results obtained for both regions reveal a good performance of the model independently of the landscape configuration. For the Maskoutains region, the proportion of forested and agricultural land differs in less than 1% point of the proportion calculated from the 2002 land-use map (Table 4). A spatial overlay of the simulation outcomes and the 2002 land-use map shows that the spatial distribution of patches produced by the dynamic neighborhood corresponds to 85.72 % and 98.41% of the spatial distribution of forested and agricultural patches, respectively, present in the 2002 land-use map (Table 4). Additionally, the Moran index calculated for the 2002 simulation outcomes (0.04) is very similar to the index calculated for the 2002 reference land-use map (0.05).

For the Elbow river watershed, the simulation results obtained are also very similar to the landscape configuration represented on the 2001 land-use map. The proportion of forested land for 2001 corresponds to 44.24% in comparison to 46.49% of forested

Table 4
Proportion of land uses for 2002 produced by VecGCA and percentage of coincidence of the simulation outcomes with the 2002 land-use map for the Maskoutains region using the dynamic neighborhood

	2002 Proportion of simulated land uses. Dynamic neighborhood	2002 Reference land-use map	2002 Simulated land uses – 2002 reference land-use map	% Coincident with the ref. land-use map
Forest	14.21	14.83	−0.62	85.72
Agriculture	83.06	82.45	0.61	98.41

Table 5
Proportion of land uses produced by VecGCA and percentage of coincidence of the simulation outcomes with the reference land-use maps for the Elbow river watershed using the dynamic neighborhood

	Proportion of simulated land uses. Dynamic neighborhood		Reference land-use map		Simulated land uses – ref. land-use map		% Coincident with the ref. land-use map	
	2001	2006	2001	2006	2001	2006	2001	2006
Forest	44.24	44.12	46.49	45.01	−2.25	−0.89	98.21	95.68
Agriculture	25.90	25.39	23.35	24.98	2.55	0.41	87.24	89.47
Urban	5.63	6.26	5.40	6.18	0.23	−0.08	87.98	85.48

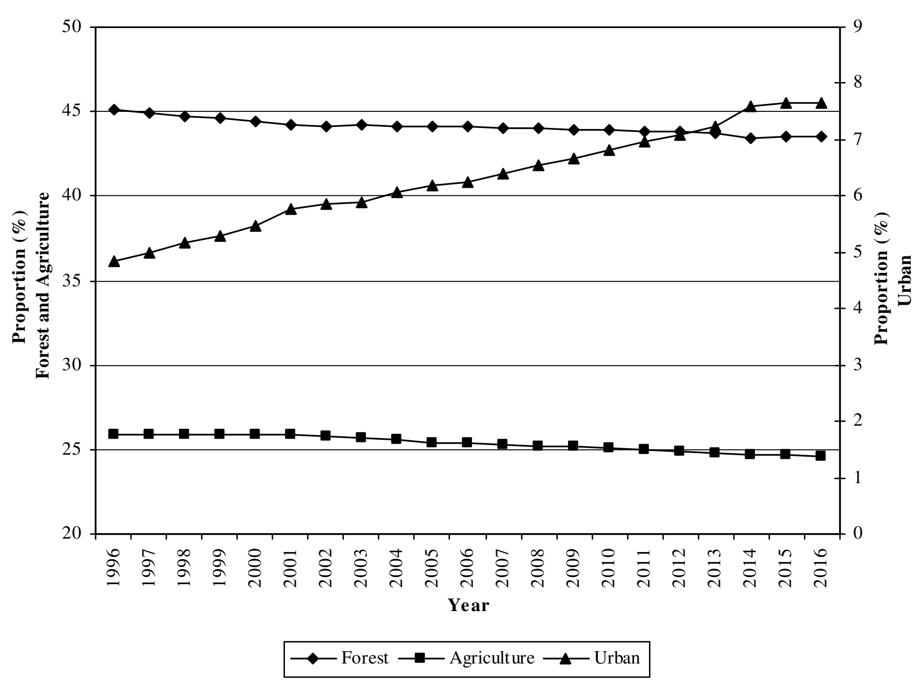

Fig. 5. Proportion of forest, agriculture and urban areas produced by VecGCA using a dynamic neighborhood, for the simulation period 1996–2016.

land calculated from the 2001 land-use map (Table 5). In general the proportion of land uses generated by VecGCA varies by less than 3% points when compared to the 2001 land-use map. The results reveal a high correspondence of the landscape generated by VecGCA using a dynamic neighborhood and the landscape present in the 2001 land-use map. A proportion of 98.21%, 87.24% and 87.98% of forested, agricultural and urban areas, respectively, corresponds to the forested, agricultural and urban patches present in

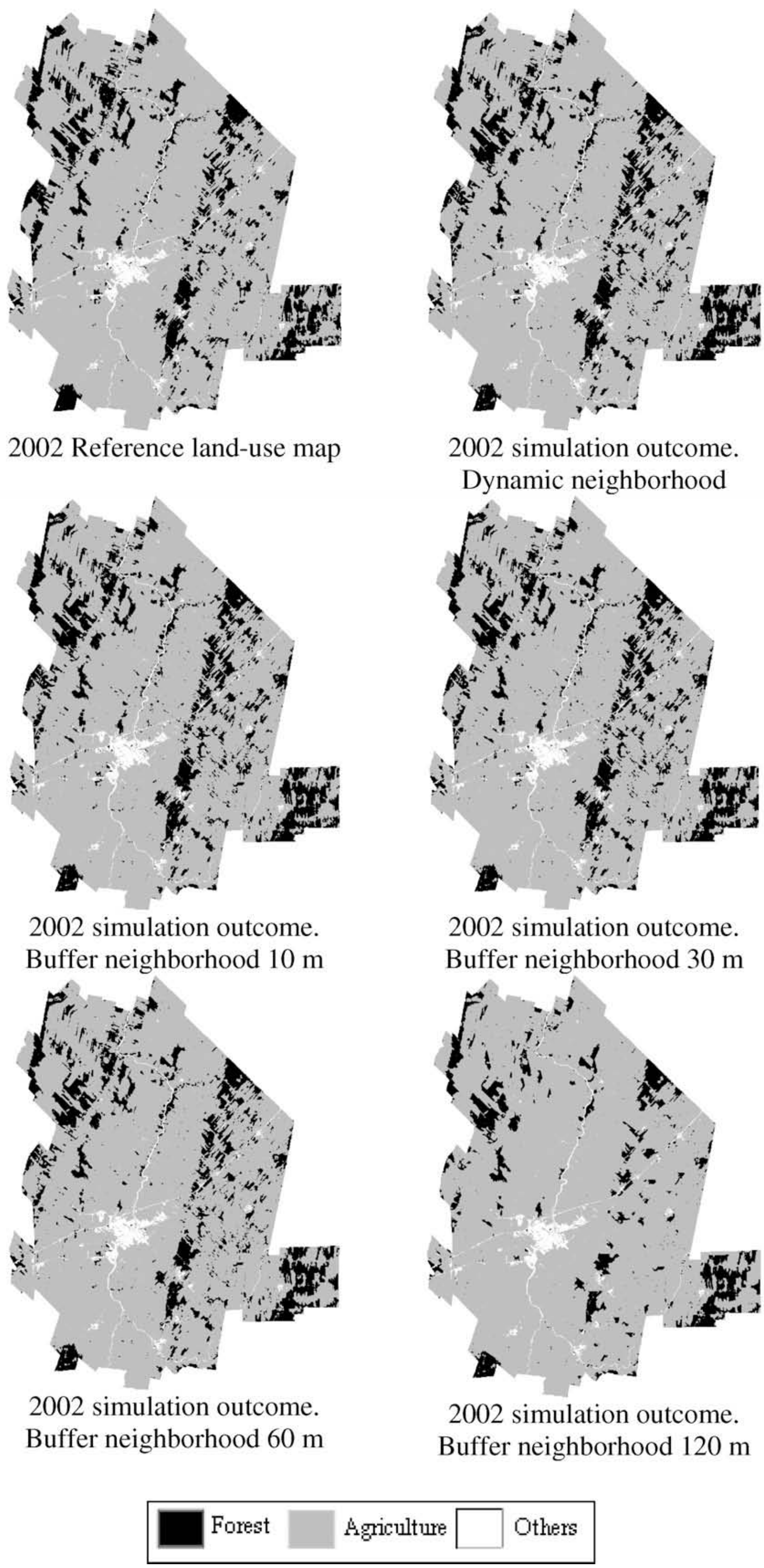

Fig. 6. Simulation outcomes for the Maskoutains region for 2002 using the buffer neighborhood with different neighborhood sizes and the dynamic neighborhood.

the 2001 land-use map (Table 5). Additionally, the spatial autocorrelation represented in the Moran index is very similar for the 2001 simulation outcomes and the 2001 reference land-use map, being 0.12 and 0.15, respectively. The same analysis performed using the 2006 land-use map (Table 5) leads to similar results which suggests that the model adequately captures the dynamics in the region. When the model was run over an additional 10 years (until 2016) the trends remain the same: the forest and agricultural areas slightly decrease while the urban area increases steadily until 2014 where it seems to reach a plateau (Fig. 5).

Fig. 6 presents the simulated maps obtained for the Maskoutains region. A visual comparison of these maps with the 2002 reference land-use map reveals the similitude between the spatial patterns generated by VecGCA using buffer sizes of 10 m, 30 m, 60 m and a dynamic neighborhood and the spatial patterns present in the study area, but a marked difference can be observed when a buffer neighborhood of 120 m is used. Similar results can be observed for the Elbow river watershed (Fig. 7). The spatial patterns generated by the 10 m buffer neighborhood and the dynamic neighborhood are similar to the spatial patterns present in the

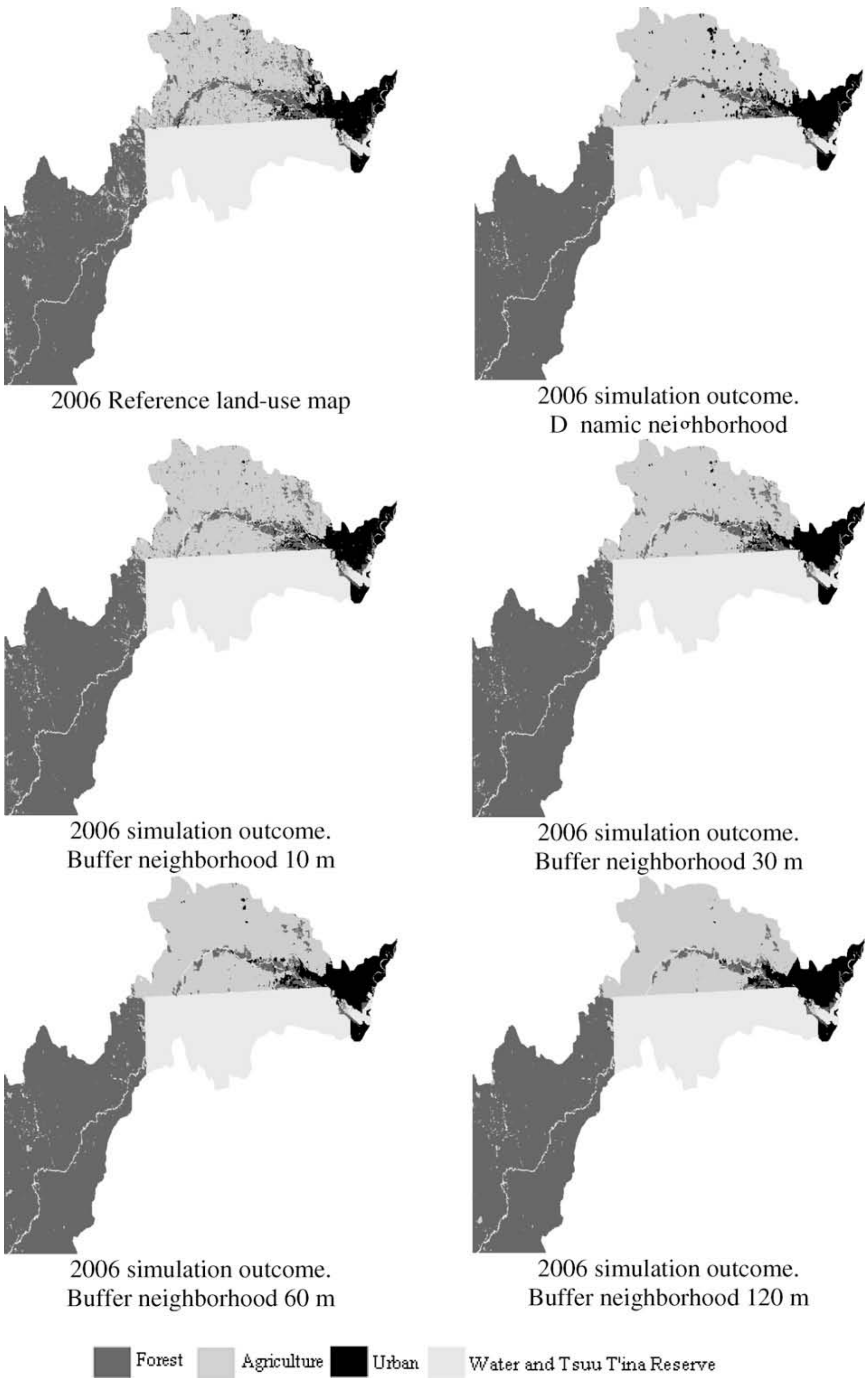

Fig. 7. Simulation outcomes for the Elbow river watershed for 2006 using the buffer neighborhood with different neighborhood sizes and the dynamic neighborhood.

study area, but a difference is noticeable when using a buffer neighborhood of 30 m and higher.

The results obtained with the implementation of a dynamic neighborhood in VecGCA demonstrate that the model can adequately represent the dynamics in both study areas. In addition, this neighborhood definition removes the neighborhood size sensitivity of VecGCA since the buffer of influence associated to each geographic object has been eliminated. The neighborhood's area is now the whole geographic space. Therefore a geographic object could be a neighbor of all other geographic objects that compose space. For example, in a geographic space composed of 100 geographic objects, a geographic object could have from zero to ninety-nine neighbors. A dynamic neighborhood allows the representation of all possible neighborhood sizes in a unique neighborhood configuration. The neighbors of a geographic space can be adjacent or separated by any distance where the limit is the extent of the geographic space.

VecGCA using a dynamic neighborhood is computationally intensive due to the number of geometric operations executed in the selection of the neighbors in addition to the geometric operations in the change of shape of objects. For example, for the Maskoutains region, the computation time was approximately 48 h (for three iterations from 1999 to 2002) for each replication. For the Elbow river watershed model, the computation time was approximately 60 h (for ten iterations from 1996 to 2006) for each replication. However, using VecGCA with a dynamic neighborhood eliminates the computation time associated to the sensitivity analysis that should be conducted to determine the best combination of cell size and neighborhood configuration in a raster-based CA model or the neighborhood size in the previous version of VecGCA, which is a considerable advantage.

5. Conclusion

Raster-based CA models are useful tools to simulate geographical phenomena; however previous studies on the impact of spatial resolution and neighborhood configuration on the simulation results of these models indicate that their arbitrary selection can generate outcomes that do not represent adequately the dynamics of the system under study. Vector-based CA models previously proposed can simulate spatial phenomena based on an irregular representation of space. However, these models still suffer some limitations, including a rigid and oversimplified definition of the objects and their neighborhood based on topology, and the lack of a dynamic representation of the geometry of the objects.

VecGCA is a new vector-based geographical cellular automata model presented as a solution to overcome the limitations of raster-based CA models and previous vector-based models by allowing the representation of space as a collection of real-world objects (polygons) with proper behavior that evolves through time. In this paper, a new dynamic neighborhood was proposed where the neighborhood relationships among objects are described semantically and the neighborhood is not associated to a fixed distance. An object A is neighbor of another object B if they are separated by objects which states are favorable to the change of state from A to B. The principal advantage of this new neighborhood definition is that it is independent of a fixed influence zone and it uses the whole geographic space to evaluate which geographic objects exert an influence on others to generate a geometric transformation or change of shape. In comparison with the traditional raster-based CA model, sensitivity analyses are not necessary to determine the most appropriate cell size and neighborhood configuration. The definition of a geographic object is no longer associated to an arbitrary partition of space; its geometric representation is an attribute that changes through time according to an influence function that defines its behaviour. Additionally, the dynamic neighborhood encompasses all possible neighborhood sizes in a unique neighborhood configuration; there are no fixed influence areas that limit the neighborhood relationships between objects. VecGCA is a computationally intensive model due to numerous geometric operations that must be performed when an object changes shape; a code optimization is under progress to address this issue. However, VecGCA eliminates the computation time associated to the sensitivity analysis to scale that should be conducted in a traditional raster-based CA, which is a considerable advantage.

The dynamic neighborhood implemented in VecGCA produces spatial patterns similar to the reference land-use maps for two study areas of different spatial complexity, which suggests that it is independent of the landscape configuration. These models were simple models based on transition probabilities and a stochastic factor to ensure that the results obtained were generated by the logic implemented in the conceptual VecGCA model and not by other external driving forces that would increase the complexity of the model. Ongoing researches aim at developing a more complex land-use change model that explicitly includes the significant driving factors that participate in the land-use dynamics of the study area.

The new version of VecGCA is a generic powerful tool to simulate land-use/land-cover changes or other spatio-temporal phenomena that implies geometric transformations of objects. Despite the increased computation time, the model can be easily implemented and generalized to different applications such as epidemic propagation, deforestation process, fire propagation, among others. With the inclusion of a dynamic neighborhood, VecGCA becomes independent of the cell size, the neighborhood configuration and the landscape configuration and ensures a more realistic representation of the geographic space and the evolution of the objects composing it.

Acknowledgments

The authors are very grateful to the three anonymous reviewers for their constructive comments and suggestions. This project was funded by two scholarships awarded to N. Moreno by the OAS (Organization of American States) and the University of Calgary, a scholarship awarded to F. Wang by the University of Calgary, and by a NSERC (Natural Sciences and Engineering Research Council) research grant awarded to D.J. Marceau. Some of the remote sensing dataset have been acquired through a collaborative project with the Calgary Regional Partnership funded by the Ministry of Municipal Affairs.

References

Almeida, C. M., Batty, M., Monteiro, A. M. V., Câmara, G., Soares-Filho, B. S., Cerqueira, G. C., et al. (2003). Stochastic cellular automata modeling of urban land use dynamics: Empirical development and estimation. *Computers, Environment and Urban Systems, 27*, 481–509.

Almeida, C. M., Gleriani, J. M., Castejon, E. F., & Soares-Filho, B. S. (2008). Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *International Journal of Geographical Information Science, 22*(9), 943–963.

Arii, K., & Parrott, L. (2006). Examining the colonization process of exotic species varying in competitive abilities using a cellular automaton model. *Ecological Modelling, 199*, 219–228.

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems, 23*, 205–233.

Benenson, I., Omer, I., & Hatna, E. (2002). Entity-based modeling of urban residential dynamics: The case of Yaffo, Tel Aviv. *Environment and Planning B: Planning and Design, 29*, 491–512.

Berjak, S. G., & Hearne, J. W. (2002). An improved cellular automata model for simulating fire in a spatially heterogeneous Savanna system. *Ecological Modelling, 148*(2), 133–151.

Chen, Q., & Mynett, A. E. (2003). Effects of cells size and configuration in cellular automata based prey–predator modeling. *Simulation Modelling Practice and Theory, 11*(7–8), 609–625.

Chen, Q., Mynett, A. E., & Minns, A. W. (2002). Application of cellular automata to modelling competitive growths of two underwater species *Chara aspera* and *Potamogeton pectinatus* in Lake Veluwe. *Ecological Modelling, 147*(3), 253–265.

Couclelis, H. (1985). Cellular worlds: A framework for modeling micro–macro dynamics. *Environment and Planning A, 17*(5), 585–596.

Dietzel, C., & Clarke, K. C. (2006). The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. *Computers, Environment and Urban Systems, 30*(1), 78–101.

Favier, C., Chave, J., Fabing, A., Schwartz, D., & Dubois, M. A. (2004). Modelling forest–savanna mosaic dynamics in man-influenced environments: Effects of fire, climate and soil heterogeneity. *Ecological Modelling, 171*, 85–102.

Flache, A., & Hegselmann, R. (2001). Do irregular grids make a difference? Relaxing the spatial regularity assumption in cellular models of social dynamics. *Journal of Artificial Societies and Social Simulation, 4*(4), 6.1–6.27.

Hammam, Y., Moore, A., & Whigham, P. (2007). The dynamic geometry of geographical vector agents. *Computers, Environment and Urban Systems, 31*(5), 502–519.

Hu, S., & Li, D. (2004). Vector cellular automata based geographical entity. In *Proceedings of the 12th international conference on geoinformatics – geospatial information research: Bridging the Pacific and Atlantic* (June 7–9, pp. 249–256). Sweden: University of Gävle.

Jantz, C. A., & Goetz, S. J. (2005). Analysis of scale dependencies in an urban land-use change model. *International Journal of Geographical Information Science, 19*(2), 217–241.

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

JTS (2004). *JTS Technology Suite*. Vivid Solutions, Victoria, British Columbia, Canada. <http://www.vividsolutions.com/jts/jtshome.htm>.

Kocabas, V., & Dragicevic, S. (2006). Assessing cellular automata model behavior using a sensitivity analysis approach. *Computers, Environment and Urban Systems, 30*(6), 921–953.

Li, X., & Yeh, A. G. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science, 16*(4), 323–343.

Marceau, D. J., & Moreno, N. (2008). An object-based cellular automata to mitigate scale dependency. In T. Blaschke, S. Lang, & G. J. Hay (Eds.), *Object-based image analysis* (pp. 43–73). Springer-Verlag.

Matsinos, Y. G., & Troumbis, A. Y. (2002). Modeling competition, dispersal and effects of disturbance in the dynamics of a grassland community using a cellular automaton model. *Ecological Modelling, 149*, 71–83.

Ménard, A., & Marceau, D. J. (2005). Exploration of spatial scale sensitivity in geographical cellular automata. *Environment and Planning B: Planning and Design, 32*, 693–714.

Ménard, A., & Marceau, D. J. (2007). Simulating the impact of forest management scenarios in an agricultural landscape of southern Quebec, Canada, using a geographic cellular automata. *Landscape and Urban Planning, 79*(3–4), 253–265.

Moreno, N., & Marceau, D. J. (2006). A vector-based cellular automata model to allow changes of polygon shape. In *Proceedings of the 2006 SCS international conference on modeling and simulation – methodology, tools, software applications* (July 31–August 2, pp. 85–92). Canada: Calgary.

Moreno, N., & Marceau, D. J. (2007). Performance assessment of a new vector-based geographic cellular automata model. In *Proceedings of the international conference on geocomputation* (September 3–5). Ireland: Maynooth.

Moreno, N., Ménard, A., & Marceau, D. J. (2008). VecGCA: A vector-based geographic cellular automata model allowing geometrical transformations of objects. *Environment and Planning B: Planning and Design, 35*, 647–665.

OpenMap (2005). *OpenMapTM. Open Systems Mapping Technology*. BBN Technologies. <http://openmap.bbn.com>.

O'Sullivan, D. (2001a). Graph-cellular automata: A generalised discrete urban and regional model. *Environment and Planning B: Planning and Design, 28*, 687–705.

O'Sullivan, D. (2001b). Exploring spatial process dynamics using irregular cellular automaton models. *Geographical Analysis, 33*(1), 1–18.

Rietkerk, M., Dekker, S. C., Ruiter, P. C., & Koppel, J. (2004). Self-organized patchiness and catastrophic shifts in ecosystems. *Science, 305*(5692), 1926–1929.

Semboloni, F. (2000). The growth of an urban cluster into a dynamic self-modifying spatial pattern. *Environment and Planning B: Planning and Design, 27*(4), 549–564.

Shi, W., & Pang, M. Y. C. (2000). Development of Voronoi-based cellular automata – an integrated dynamic model for geographical information systems. *International Journal of Geographical Information Science, 14*(5), 455–474.

Soucy-Gonthier, N., Marceau, D. J., Delage, M., Cogliastro, A., Domon, G., Bouchard, A. (2003). *Détection de l'évolution des superficies forestières en Montérégie entre juin 1999 et août 2002 à partir d'images satellitaires Landsat TM* [Survey of the evolution of forested areas in the Monteregie region, June 1999–August 2002, from Landsat-TM remote sensing images], report presented to the Agence Forestière de la Montérégie; copy available from the authors.

Stevens, D., Dragicevic, S., & Rothley, K. (2007). iCity: A GIS-CA modelling tool for urban planning and decision making. *Environmental Modelling and Software, 22*(6), 761–773.

Stewart-Cox, J., Britton, N., & Mogie, M. (2005). Pollen limitation or mate search need not induce an Allee effect. *Bulletin of Mathematical Biology*, *67*(5), 1049–1079.

Sun, T., & Wang, J. (2007). A traffic cellular automata model based on road network grids and its spatial and temporal resolution's influences on simulation+. *Simulation Modelling Practice and Theory, 15*, 864–878.

Takeyama, M., & Couclelis, H. (1997). Map dynamics: Integrating cellular automata and GIS through geoalgebra. *International Journal of Geographical Information Science, 11*(1), 73–91.

Torrens, P. M., & Benenson, I. (2005). Geographic automata systems. *International Journal of Geographical Information Science, 19*(4), 385–412.

Verburg, P. H., Nijs, T. C. M., Eck, J. R., Visser, H., & Jong, K. (2004). A method to analyse neighborhood characteristics of land use patterns. *Computers, Environment and Urban Systems, 28*(6), 667–690.

Von Neumann, J., & Burks, A. W. (1966). *Theory of self reproducing automata*. Urbana, Illinois: University of Illinois Press.

Wahle, J., Neubert, L., Esser, J., & Schreckenberg, M. (2001). A cellular automata traffic flow model for online simulation of traffic. *Parallel Computing, 27*(5), 719–735.

White, R., & Engelen, G. (1993). Cellular automata and fractal urban form: A cellular modelling approach to the evolution of urban land-use patterns. *Environment and Planning A, 25*(8), 1175–1199.

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

White, R., Engelen, G., & Uljee, I. (2000). Modelling land-use change with linked cellular automata and socio-economic models: A tool for exploring the impact of climate change on the island of St. Lucia. In M. J. Hill & R. J. Aspinell (Eds.), *Spatial information for land-use management* (pp. 189–204). Amsterdan: Gordon and Breach.

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature, 311*, 419–424.

Wu, F. (2002). Calibration of stochastic cellular automata: The application to rural–urban land conversions. *International Journal of Geographical Information Science, 16*(8), 795–818.

Yassemi, S., Dragicevic, S., & Schmidt, M. (2008). Design and implementation of an integrated GIS-based cellular automata model to characterize forest fire behaviour. *Ecological Modelling, 210*, 71–84.

Yeh, A. G.-O., & Li, X. (2006). Errors and uncertainties in urban cellular automata. *Computers, Environment and Urban Systems, 30*, 10–28.