CELLULAR AUTOMATA MODEL BASED ON MACHINE LEARNING METHODS FOR SIMULATING LAND USE CHANGE

Omar Charif

UTC & CEPS/INSTEAD
rue du Dr Schweitzer
60200 Compiègne, FRANCE

Hichem Omrani

CEPS/INSTEAD
3 Rue de la Fonte
L-4364 Esch-Sur-Alzette, Luxembourg

Reine-Maria Basse

CEPS/INSTEAD
3 Rue de la Fonte
L-4364 Esch-Sur-Alzette, Luxembourg

Philippe Trigano

Univeristy of Technology of Compiegne(UTC)
rue du Dr Schweitzer
60200 Compiègne, FRANCE

ABSTRACT

This paper presents an approach combining machine learning (ML), cross-validation methods and cellular automata (CA) model for simulating land use changes in Luxembourg and the areas adjacent to its borders. Throughout this article, we emphasize the interest in using ML methods as a base of CA model transition rule. The proposed approach shows promising results for prediction of land use changes over time. We validate the various models using cross-validation technique and Receiver Operating Characteristic (ROC) curve analysis, and compare the results with those obtained using a standard logit model. The application described in this paper highlights the interest of integrating ML methods in CA based model for land use dynamic simulation.

1 INTRODUCTION

In the last decades, land demands in Europe specifically in an attractive region like Luxembourg became an important issue for stakeholders. The population of Luxembourg keeps growing and the impact of this growth is visible especially when analyzing the land use changes over time. Land use is shaped by spatial interaction which make their features difficult to analyze. The spatial interactions induce neighborhood effects and so introduce the emergence of new patterns that need to be simulated in order to understand the hidden process of the land properties. Thus, the main question is how to simulate land process and consequently how to simulate the land use patterns?

The land use patterns of Luxembourg and areas adjacent to its borders can be simulated based on CA concept using several variables (e.g., slope, amount of built-up area surrounding a given cell, etc.). A built-up areas (also called artificial surfaces) contain all urban and industrial areas. These variables will be described in Section 5. The simulation of land use changes allows us to understand the spatial interaction as well as the land use patterns that emerged from the neighborhood effects on a cell and its assigned land use function.

The rest of this paper is organized in five sections. first, we show the motivations behind this work. Then, in the next section, we give a short overview of the CA model, its main components and its relevance to land use changes analysis. Section 4 describes the methodological aspects. Section 5 focuses on the application, its results and a comparison with various models. Section 6 draws conclusions and discussions.

2 MOTIVATION

The land use change is a very complex phenomenon due to: first, the complex relations, actions, interactions between different land use classes and second the variety and the complexity of the factors that influence the land use (e.g., economical conditions, transportation system). For the last three decade, to deal with this complexity, new dynamic and spatially explicit model based on cellular automata philosophy (Von Neumann 1951) emerged as a simulation tool for land use changes (Couclelis 1985).

Using conventional statistical methods in spatial land use analysis (like logistic regression: logit), is problematic because these methods assume the data to be statistically independent and identically distributed (iid). But spatial land use data have the tendency to be dependent, a phenomenon known as spatial autocorrelation (Xie et al. 2005). However, a number of papers recently published (Hu and Lo 2007) used logit model even with non mutually exclusive alternatives (e.g. the transition of urban cell to non urban cell is not possible). For this reason, we choose in the application two mutually exclusive states (built-up and not built-up land). In this paper, we simulate land use changes using ML models as a core of CA transition rule for Luxembourg and its cross-border region. The use of ML models is mainly motivated by their universal approximation capabilities and their high performance in a wide range of scientific fields.

3 RELATED WORK

The concept of CA was first introduced in the 1940's when Von Neumann (1951) and Ulam (1952) were investigating "complexity field" to analyze the behaviors of complex systems. Stephen Wolfram followed the same concept and introduced CA as a model of complexity (Wolfram 1984b; Wolfram 1984a). Later in the last century, CA have been applied in different fields like physics, mathematics, computer science, biology, philosophy and geography. From the end of 70s to the end of 90s, CA were presented as one of the most relevant tools for understanding complex systems (Tobler 1979; Couclelis 1985) particularly land use patterns (White and Engelen 1993; Batty, Xie, and Sun 1999; Torrens and Benenson 2005). Therefore, sensitivity of CA to the spatial configuration (e.g., land use features, transport networks, physical constraints) and neighborhood relationship (Ménard and Marceau 2005) has been extensively investigated by researchers in order to simulate urban growth and future land use requirements (Batty et al. 1999; Barredo et al. 2003; Torrens and Benenson 2005). Simple CA has five fundamental properties: (1) a regular discrete lattice of cells; (2) the evolution of each cell takes place in discrete time points (time step); (3) each cell is in one of a finite set of states that are exhaustive (no other states are possible) and exclusive (a cell cannot be in more than one state at any one time); (4) the future state of the cells is determined by the current states of the cell itself and the cells in the neighborhood following transition rules (these rules are identical for all cells in the lattice); (5) the neighborhood relation influences the studied cell. These properties allow CA to simulate complex systems (Wolfram 1984b; Couclelis 1985). Indeed, land use is among the most complex system that needs to be understood for better land planning (Lambin, Rounsevell, and Geist 2000).

The transition rule is one of the most important components of a CA model that needs to be well determined to provide good simulation in the context of CA based land use model (Couclelis 1985). Therefore a variety of methods was employed in the literature as transition rules for predicting land use, among them: standard logistic regression (logit) (Müller and Zeller 2002), a combination of linear and geometric formulations (Lau and Kam 2005) and rule-based model (Betel and Flocchini 2011). Recently, new artificial intelligence methods have been applied to simulate land use changes (Szuster, Chen, and Borger 2011). The two methods recently applied are artificial neural network (ANN) (Lin et al. 2005) and support vector machine (SVM) (Yang, Li, and Shi 2008). They both showed competitive performance (Szuster et al. 2011; Omrani et al. 2012). In the present paper, We apply these two methods and compare their results to those from a logit model. Table 1 presents a classification of methods used as transition rules in CA modeling.

Table 1: Methods of CA transition rules to predict land use dynamics in recent urban model applications.

Modelling method	Transition rule	References
Mathematic /	Logit	(Xie et al. 2005)
Econometric	MNL	(Müller and Zeller 2002)
– – –	LGF	(Lau and Kam 2005)
Conditional /	Basic	(Fulong 2002)
If-then	Fuzzy	(Betel and Flocchini 2011)
Mathematic /	ANN	(Lin et al. 2005)
AI	SVM	(Yang, Li, and Shi 2008)

Notes: Logit: logistic regression, MNL: multinomial logit, LGF: linear and geometric formulations, ANN: artificial neural network, SVM: support vector machine and AI: artificial intelligence.

4 METHODOLOGY

The land use modeling is a non linear regression problem which can be tackled using several approaches and methods. Here we use four different ones (one statistical and three other belonging to ML family) to define the transition rule of the CA model for predicting land use changes. Beside these transition rules, the CA model used in this work has the following properties: 1) Regular rectangular grid with a resolution of 100×100 meters, this grid presents the study area: Luxembourg and its bordering regions, 2) Time step of 10 years is chosen due to the availability of data as we possess the land use for two years 1990 and 2000, 3) Cell states are built-up and non built-up and 4) Classical Moore neighborhood with a window of 3×3 is used as the neighborhood that influences the studied cell. Here after, we briefly describe all the methods used as transition rule for the CA model.

4.1 Logit Model

Logit model is used to predict the probability in a binomial regression. It allows, after coefficients optimization, to model the link between a set of explanatory variables and a dichotomous dependent variable. Suppose that the dependent variable Y is binary, that is, it can have only two possible outcomes which we will denote as 1 and 0. Let $Y = \beta X + \varepsilon$ where the normalized error term ε follows a logistic distribution with mean 0 and variance $\frac{\pi^2}{3}$. X is a vector of explanatory variables which influence the outcome Y. The outcome probabilities can be defined as follows:

$$Pr(Y = 1|X) = \frac{\exp^{\beta X}}{1 + \exp^{\beta X}} \tag{1}$$

The parameters β are typically estimated by maximizing the likelihood function:

$$L = \prod_{i=1}^{n} \left[\frac{\exp^{\beta X}}{1 + \exp^{\beta X}} \right]^{y_i} \left[1 - \frac{\exp^{\beta X}}{1 + \exp^{\beta X}} \right]^{1-y_i} \tag{2}$$

4.2 Support Vector Machine (SVM)

SVM also called as maximum margin classifier is relatively new supervised ML algorithm, it was introduced with all its features by Vladimir Vapnik and his co-workers in 1995 (Cortes and Vapnik 1995). This method presents an extension to non linear models of Vapnik (1963) for pattern recognition using generalized portrait method. The concept behind the SVM algorithm is to find and to orient a separating hyperplane in such a way to be as far as possible from the nearest members of both classes which are called the support vectors. Thus, SVM tries to find a hyperplane where the margin (see Fig. 1) is the largest possible. This method assumes that new unseen input will often be close to the training patterns. The bigger is the margin between classes, the bigger the chances that the new input will be successfully classified and thus the better will be the ability to generalize.
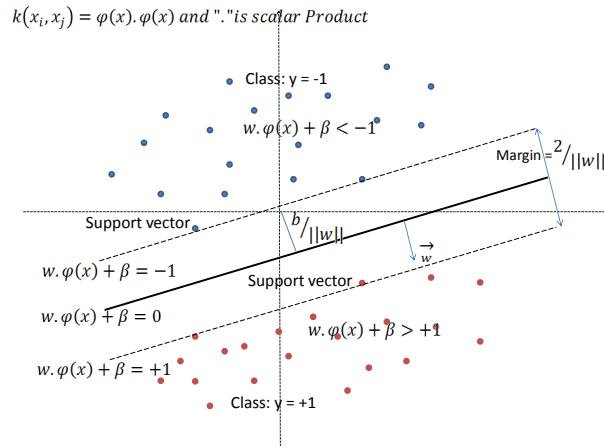
$$k(x_i, x_j) = \varphi(x).\varphi(x) \text{ and "." is scalar Product}$$

Class: y = -1

$w.\varphi(x) + \beta < -1$

Margin $= \frac{2}{||w||}$

Support vector

$\frac{b}{||w||}$

$w.\varphi(x) + \beta = -1$

Support vector

$\vec{w}$

$w.\varphi(x) + \beta = 0$

$w.\varphi(x) + \beta > +1$

$w.\varphi(x) + \beta = +1$

Class: y = +1

Figure 1: Dichotomous Classification problem using SVM.

The separating hyperplane is given by: $w \bullet \phi(x) + b = 0$, where w is perpendicular vector to the hyperplane, b is a constant such that $\frac{b}{||w||}$ is the perpendicular distance from the hyperplane to the origin of the space and $\phi(X)$ is the a projection, mapping function applied on the input vector x. This function will be used to define the kernel of SVM model given by: $K(x_i, x_j) = \phi(x_i) \bullet \phi(x_j)$. The kernel function is mainly used to project and remap the input vector on generally a higher dimension space in which this non linearly separable problem is transformed into linearly separable one in this new "feature" space. Thus, in case, of linearly separable problem, a linear kernel is used: $K(x_i, x_j) = x_i \bullet x_j$ with $\phi(x) = x$.

To successfully implement an SVM, it is enough to optimize w and b to maximize the margin given by: $\frac{2}{||w||}$. This is done using Lagrange multiplier and quadratic programming. Finally, in case of a binomial classification where the output is 1 or -1, each new input X is classified by evaluating $y = sgn(w \bullet \phi(x) + b)$. In this paper, we are dealing with non linearly separable problem thus we needed a non linear kernel function and we chose the radial basis function (RBF) as a Kernel function which is a real function that depend only on the distance from its argument to the origin or to another point in the space, the kernel we used is given by:

$$k(x_i, x_j) = \exp^{\left(\frac{x_i - x_j}{2\sigma^2}\right)} \tag{3}$$

4.3 Multi Layer Perceptron Artificial Neural Network (ANN-MLP)

We used a conventional Multi Layer Perceptron neural network (denoted as ANN-MLP) with one hidden layer trained using a supervised learning method, the standard "Backpropagation" learning algorithm. The hidden layer is a set of simple non linear hidden neurons. According to Heaton (2005), neural networks without a hidden layer will not be able to model any non linear function. The hidden layers hold the ability of neural network models to mimic non linear functions. Let $y = (y_1, y_2)$ be the output (e.g., built-up and not built-up land) which is expressed as a function of the input $x = (x_1, x_2, \ldots, x_q)$ (see Fig. 2) as follows (1):

$$y_k = \Psi\left(\sum_{j=1}^{p} v_{jk} \Phi\left(\sum_{i=1}^{q} \omega_{ij} x_i + \omega_{0j}\right) + v_{0k}\right) \tag{4}$$

where the ω_{ij} and v_{jk} are weights assigned to the connections between the input layer and the hidden layer, and between the hidden layer and the output layer, respectively, ω_{0j} and v_{0k} are biases (or threshold values in the activation of a unit). Φ is the activation function of the hidden layer and it is applied on the weighted sum of the input data. While Ψ is the activation function applied, by each output unit, to the weighted sum of the activations of the hidden layer. This expression can be generalized to networks with several
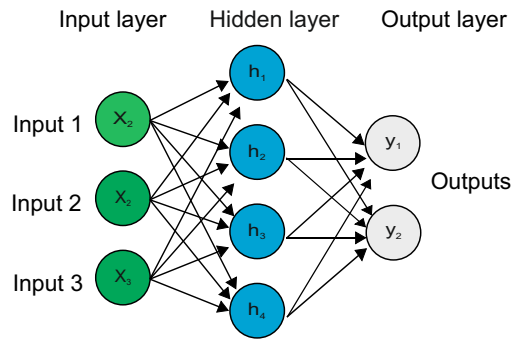
Figure 2: Architecture of ANN topology.

hidden layers. The output of the ANN will not be exactly equal to 0 or 1. When Ψ is Softmax function, the output will be the probabilities of belonging to a class. These values are in the range (0,1). Figure 2 shows the ANN architecture with an input layer, one hidden layer with h units and one output layer.

4.4 Ann-rbf

An ANN-RBF (Lowe and Broomhead 1988) is a neural network model that has the structure of MLP. But, in contradictory to ANN-MLP, the ANN-RBF formulates the problem in an interpolation way not in stochastic approximation way. An ANN-RBF neural network model assumes that the predicted value of an input x is likely to be equal to the output of other patterns (training patterns) that have close values to x. An ANN-RBF model is generally composed of one input, one or more hidden and one output layers. The hidden Layers contain a set of neurons, each neuron apply a non linear function (called Basis function). Neurons of the output layer apply a linear function. The output of an input x is given by:

$$y = \sum_{i=1}^{N} w_i \phi(\|x - c_i\|) \tag{5}$$

where N is the number of hidden neurons, w_i and c_i are respectively the optimized weight and the center of hidden neurons number i and $\phi(x)$ is the basis function. the two set of values w and c are optimized during the training procedure.

4.5 Methods Assessment and Comparison

We assess and compare methods by their performances in predicting the land use change. We used the Corine Land Cover data set of 1990 with other explanatory variables (see Section 5.2) to predict the land use change for the year 2000. We then compared the prediction with the Corine Land Cover 2000 in term of success rate. We define the success rate as follows: Suppose we have a set of observations which we are trying to classify into two classes A and B. The success rate of the class A is the number (nb) of the observations correctly classified (i.e. predicted as A and is actually A) expressed as a percentage of the total number of observations in class A(see equation 6).

$$\text{success rate(A)} = \frac{\text{nb correctly classified As A}}{\text{nb Total As A}} \tag{6}$$

Over-fitting is a well studied problem in neural networks and other complex methods (Lawrence, Giles, and Tsoi 1998). A model with a very good fit for a learning sample may yield a poor prediction on a test sample. The problem of over-fitting is addressed by fitting the model on a learning sample and evaluating the success rates on the complementary test sample. The model was validated by cross-validation technique (*N* replications) and Receiver Operating Characteristic (ROC) curve analysis.

4.5.1 Cross-validation

We want to avoid the vagaries of choosing a particular partitioning (sub-samples), and therefore we evaluate the success rates on a large number of partitions by cross-validation. We split the dataset S into training sub-sample L (60%) and testing sub-sample T (40%) by a random process. We fit the model to L and evaluate the success rate on T. We replicate this process of partitioning, fitting the model and evaluating the success rate, N times obtaining N sets of the success rates. Their average is the overall success rate.

4.5.2 ROC Curve Analysis

A summary of the predictions is obtained by the confusion matrix (classification table). The element in cell $(k;h)$, is the number of observations for which observed and predicted classes are respectively k and h. The numbers along the major diagonal of this matrix present the correct predictions, and the numbers off this diagonal represent the errors in prediction committed. The true positive rate (denoted by "tp rate", also called hit rate and recall) and the false positive rate (noted by "fp rate", also called false alarm rate) are estimated as:

$$\text{tp rate} = \frac{\text{Positive correctly classified}}{\text{Total positives}} \tag{7}$$

$$\text{fp rate} = \frac{\text{Positive incorrectly classified}}{\text{Total negatives}} \tag{8}$$

where total positives and total negatives are the numbers of observations observed respectively as positive and negative.
Additional metrics are:

$$\text{sensitivity} = \text{tp rate} \tag{9}$$

$$\text{specificity} = \frac{\text{Negative correctly classified}}{\text{Total negatives}} = 1 - \text{fp rate} \tag{10}$$

ROC graphs are two-dimensional graphs in which "sensitivity" is plotted on the Y axis and "1 - specificity" is plotted on the X axis. A ROC graph depicts relative trade-offs between (sensitivity) and (1 - specificity). Several points in ROC space are important to note. The lower left point (0;0) represents the strategy of never issuing a positive classification; such a classifier commits no false positive errors but also gains no true positives. The opposite strategy, of unconditionally issuing positive classifications, is represented by the upper right point (1;1) whereas the point (0;1) represents perfect classification.

5 APPLICATION

5.1 Study Area

Despite its small area ($2,586 \, km^2$) and population (around $500,000$ habitants), Luxembourg benefits from a strategic geographic location and appears as an important contributor in the decision making in the European Union (Fig. 3). Indeed, Luxembourg is one of the most attractive metropolitan areas in Europe (Sohn, Reitel, and Walther 2009). Its attraction is due to the socio-economic development of the society and specifically to the strong economic sectors (financial and industrial) that have been developing since the end of 1970's.

Surrounded by three countries (France, Germany and Belgium), Luxembourg attracted a lot of labor force to maintain its economic dynamism. This explains the importance of residential and daily mobility in Luxembourg and its bordering land (study area) (Gerber et al. 2008; Omrani et al. 2010). This paper assumes the hypothesis that the current situation of Luxembourg reflects a dynamic territory which will impact the land demands and consequently the land use changes in the near future. Thus, the objective of this paper is to investigate the way of possible land use changes of the study area between 2010 and 2050.
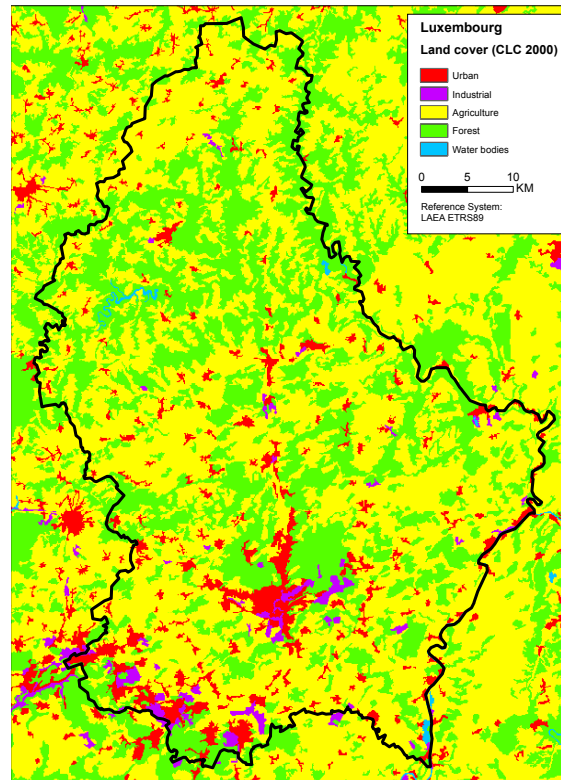
Figure 3: Land use/land cover map of study area, red signify built-up areas.

5.2 Main Inputs

Our simulation of urban changes uses multiple data sources: Corine Land Cover in two periods: 1990 and 2000 (Bossard et al. 2000), Digital Elevation Model (DEM) to show the physical constraint of the study area (Fig. 3). The source of the processed elevation data is the global SRTM dataset (Shuttle Radar Topography Mission) (Farr et al. 2007; Bodis 2009). The elevation model is used to calculate the slope presented in Fig. 4. The slope (or grade) is the amount of inclination of the topographic land form to the horizontal and it is given in percentage by: $slope = 100 \cdot \tan(\alpha)$ where α is the angle of inclination. The mentioned data sources were used to construct a dataset with unique patterns. This dataset was split randomly as shown in Section 4.5.1 in into training (60%) and testing datasets (40%), to be used in the ML models, as described in Table 2.

The inputs used in this application are: the slope value, the amount of built-up area in the neighborhood (Moore neighborhood configuration)and the state of each studied cell (Table 3). Table 4 shows the land use statistics in the observed periods (1990 and 2000). The ArcGIS software and Matlab tool were used respectively for spatial data processing and for simulation.

Table 2: Summary of distribution of land use classes.

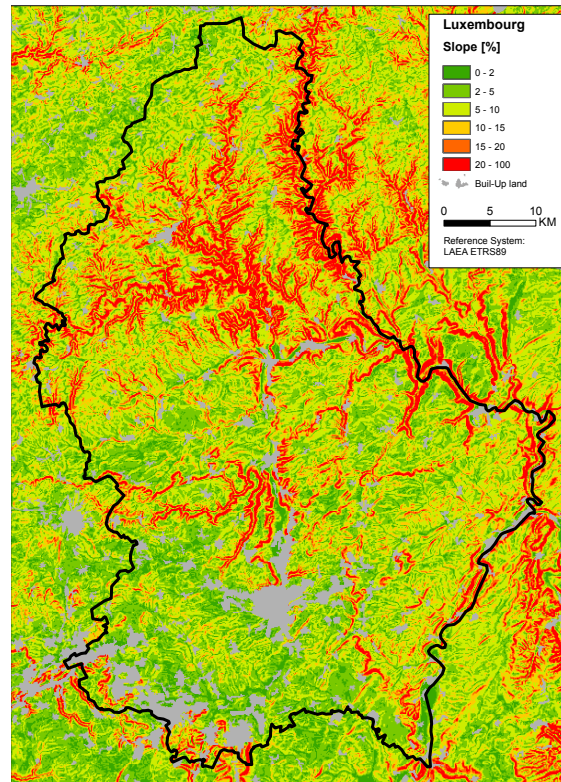| Land use	Total dataset		Training dataset		Test dataset	
	Number	Per (%)	Number	Per (%)	Number	Per (%)
Not built-up	3290	75.84	1956	75.17	1334	76.84
Built-up	1048	24.16	646	24.83	402	23.16
Sum	4338	100	2602	100	1736	100

Figure 4: Physical aspects of study area: slope with highlighted urban zone.

Table 3: Explanatory variables and description.

Category	Variable	Description	Number of modality
Spatial	Neighbours	Amount of built-up cells	1
		in the 3×3 neighborhood	
Physical	Slope	Slope value of cell	1
Physical	State	State of cell	1
		(0: built-up cell, 1: not built-up cell)	

Table 4: Land cover statistics in observed periods 1990 and 2000.

Year	Built-up land (number of cells)	Not built-up land (number of cells)	Proportion of built-up land (in %)
1990	28.251	492.654	5,42
2000	29.858	491.047	5,73

Note: Configuration of the raster grid: squared grid with a resolution of 100×100 meters (cell size of 1ha) and Moore neighborhood.

5.3 Results: Simulation of Land Use Changes

We present, in this work, CA based model to simulate the land use changes where three different ML models were used to define the transition rule. The different characteristics of these models were fixed by experiments. For example, the number of neurons in the hidden layer is found by changing it and observing the resulted rates of correct prediction (see Fig. 5).

The results presented in Table 5 show that the ANN-MLP outperformed the logit, ANN-RBF and the SVM models in terms of overall prediction accuracy. The ANN-MLP improved the overall prediction
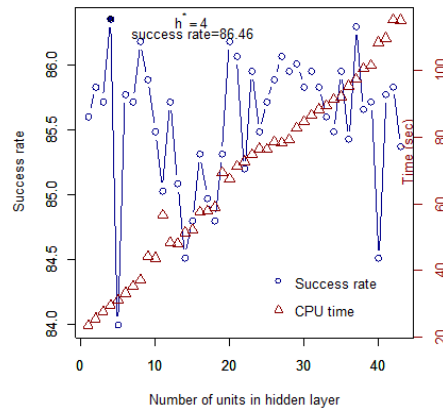
Figure 5: Success rate for number of units in the hidden layer, with optimal number highlighted ($h^\star = 4$; success rate=85.54 and computation time = 25sec).

accuracy (performance) nearly by 0.43% (85.54% vs. 85.11%) over the logistic model. Although the three models are very effective in predicting the not built-up land, with a success rate above 89%, ANN-RBF has a greater ability to predict not built up land and improved the success rate by 2.39%, 2.39%, and 4.1% compared respectively to logit, SVM and ANN-MLP. The ANN-RBF model tends to underestimate the built-up lands by misclassifying more built-up cells into not built-up cells. This may be caused by the domination of not built-up lands in the study area (75% of cases, Table 2). The ANN-MLP, logit and SVM have demonstrated a greater ability than ANN-RBF in predicting the less frequent class (built-up class) and they improved the prediction accuracy respectively by 25.08%, 17.96%, and 17.2%. The ANN-MLP has the highest the success rates (72.68%)in predicting built-up land (less frequent class) and its overall prediction accuracy is also high (85.54%).

Table 5: Success rate of the Logit, ANN and SVM models.

	Not built-up	Built-up	Overall	Performance
	(%)	(%)	(%)	Vs. Logit
Logit	91.37	65.56	85.11	0.00
SVM-RBF ($\gamma^\star = 1$)	91.49	64.80	85.03	-0.08
ANN-RBF ($\gamma^\star = 1, cost^\star = 1$)	93.76	47.60	82.14	-2.97
ANN-MLP ($h^\star = 4$)	89.66	72.68	85.54	$+0.43$

Note: An asterisk * is used to highlight the optimal value used in the appropriate model.

The work done in this paper is part of a project which has the objective to simulate the workers mobility for the future. Thus, even-though the preliminary performance of the model was not high enough to be able to simulate the land use for the next 100 years, it was important for us to have a view of the prospect of the land use changes. Further improvement will be done, in the near future, to improve the performance of the proposed model and to make more plausible the simulation up to the next 100 years (see perspective in Section 6). Figure 6 shows land use changes of the study area in 2000, 2050 and 2100 using ANN-MLP model. The evolution of built-up cells and not built-up cells from 1990 to 2100 are shown in Fig. 7. It can be observed in Fig. 6 that the south west part of the study area is more dynamic than the northern part. This situation can be explained by the existence of a higher density of urban cells (see Fig. 3) and the suitable relief in this area (see Fig. 4).

6 CONCLUSION, DISCUSSION AND PERSPECTIVES

Modeling land use changes has recently attracted a lot of attention. This attention led to the development of a wide range of methods and applications. In this paper, we used a CA based model in which four
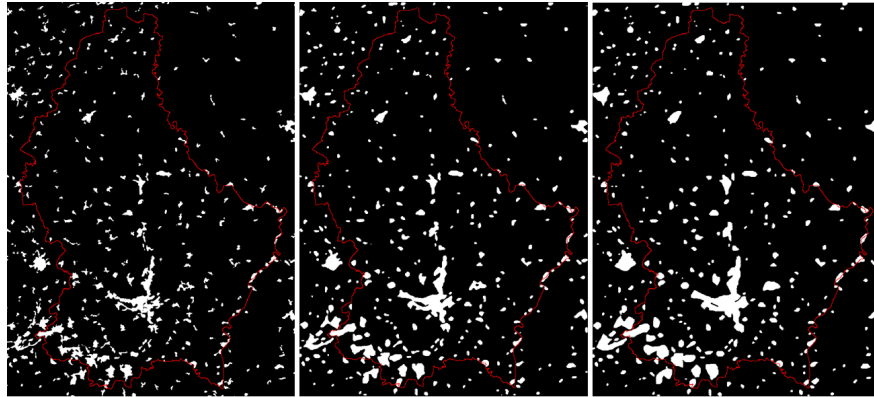
Figure 6: Land use changes of the study area in 2000, 2050 and 2100 using ANN-MLP. Not Built-up area, built-up area and administrative limit are displayed respectively in black, white and red colour.
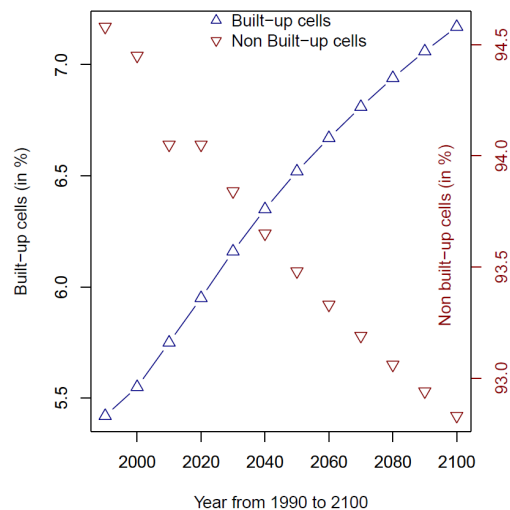

Figure 7: Evolution of Built-up and not built-up cells from 1990 to 2100 using ANN-MLP.

methods for non linear classification were tested as a transition rule. The performance of the different methods is influenced by the structure of the data. This is due to the separability of classes and also to the limited spatial representation of some classes (e.g. in the studied area: not built-up and built-up present respectively 75.84% and 24.16% of the total area). The results show that the ANN-MLP is slightly better then the other studied alternatives (ANN-RBF, SVM and logit) in terms of overall performance. This was mainly due to the good performance of ANN-MLP in predicting the less presented classes (built-up class) in our training data set. All applied methods were relatively equally capable of predicting the not built-up areas (success rate 90%). This can be explained by the existence of exhaustive set of patterns covering most of the cases that could lead to correctly classify not-built up area. The added value of this paper is the use of ANN in CA and cross-validation technique for extended validation. Our next challenge is to integrate various cell states (urban, industrial, agriculture, forest, water) and additional inputs (e.g., transport system, Luxembourg urban planning) into the model. In addition, following the results of this work, we believe that a decision fusion of some of ML models can improve the simulation results. Thus, We plan to combine several of ML models to define the transition rule for the CA based model for simulating the land use changes.

7 ACKNOWLEDGMENT

REFERENCES

Barredo, J., C. Lavalle, L. Demicheli, M. Kasanko, and N. McCormick. 2003. "Sustainable urban and regional planning: The MOLAND activities on urban scenario modelling and forecast. European Commision, Joint Reseach Centre". *Institute for Environment and Sustainability. Office for Official Publications of the European Communities, Luxembourg.*

Batty, M., Y. Xie, and Z. Sun. 1999. "Modeling urban dynamics through GIS-based cellular automata". *Computers, environment and urban systems* 23 (3): 205–233.

Betel, H., and P. Flocchini. 2011. "On the relationship between fuzzy and Boolean cellular automata". *Theoretical Computer Science* 412 (8–10): 703–713.

Bodis, K. 2009. "Development of a data set for continental hydrological modelling.". *Publications Office of the European Union.*

Bossard, M., J. Feranec, J. Otahel, and E. E. Agency. 2000. "CORINE land cover technical guide: addendum 2000".

Cortes, C., and V. Vapnik. 1995. "Support-vector networks". *Machine learning* 20 (3): 273–297.

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

Farr, T., P. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth et al. 2007. "The shuttle radar topography mission". *REVIEWS OF GEOPHYSICS-RICHMOND VIRGINIA THEN WASHINGTON-* 45 (2).

Fulong, W. 2002. "Calibration of stochastic cellular automata: the application to rural-urban land conversions". *International Journal of Geographical Information Science* 16 (8): 795–818.

Gerber, P., S. Carpentier, S. Petit, and I. Pigeron-Piroth. 2008. "Mobilités quotidienne et résidentielle au Luxembourg : un aperu à travers l'outil MobilluxWeb". *Population & Territoire*:13–16.

Heaton, J. 2005. *Introduction to neural networks with Java.* Heaton Research Inc.

Hu, Z., and C. Lo. 2007. "Modeling urban growth in Atlanta using logistic regression". *Computers, Environment and Urban Systems* 31 (6): 667–688.

Lambin, E., M. Rounsevell, and H. Geist. 2000. "Are agricultural land-use models able to predict changes in land-use intensity?". *Agriculture, Ecosystems & Environment* 82 (1-3): 321–331.

Lau, K., and B. Kam. 2005. "A cellular automata model for urban land-use simulation". *Environment and Planning B: Planning and Design* 32 (2): 247–263.

Lawrence, S., C. Giles, and A. Tsoi. 1998. "What size neural network gives optimal generalization? Convergence properties of backpropagation".

Lin, H., K. Lu, M. Espey, and J. Allen. 2005. "Modeling Urban Sprawl and Land Use Change in a Coastal Area–A Neural Network Approach". Technical report.

Lowe, D., and D. Broomhead. 1988. "Multivariable functional interpolation and adaptive networks". *Complex systems* 2:321–355.

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

Müller, D., and M. Zeller. 2002. "Land use dynamics in the central highlands of Vietnam: a spatial model combining village survey data with satellite imagery interpretation". *Agricultural Economics* 27 (3): 333–354.

Omrani, H., O. Charif, P. Gerber, K. Bódis, and R. M. Basse. 2012. "Simulation of land use changes using cellular automata and artificial neural network". Technical report, CEPS/INSTEAD working paper.

Omrani, H., O. Charif, O. Klein, P. Gerber, and P. Trigano. 2010. "An Approach for spatial and temporal data analysis: Application for mobility modeling of workers in Luxembourg and its bordering areas". Technical report.

Sohn, C., B. Reitel, and O. Walther. 2009. "Cross-border metropolitan integration in Europe: the case of Luxembourg, Basel, and Geneva". *Environment and Planning C: Government & Policy* 27 (5): 922–939.

Szuster, B., Q. Chen, and M. Borger. 2011. "A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones". *Applied Geography* 31 (2): 525–532.

Tobler, W. 1979. "CELLULAR GEOGRAPHY". *Philosophy geography*:349–386.

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

Ulam, S. 1952. "Random processes and transformations". Technical report.

Vapnik, V. 1963. "Pattern recognition using generalized portrait method". *Automation and Remote Control* 24:774–780.

Von Neumann, J. 1951. "The general and logical theory of automata". *Cerebral mechanisms in behavior*:1–41.

White, R., and G. Engelen. 1993. "Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns". *Environment and planning A* 25:1175–1175.

Wolfram, S. 1984a. "Cellular automata as models of complexity". *Nature* 311 (5985): 419–424.

Wolfram, S. 1984b. "Universality and complexity in cellular automata". *Physica D: Nonlinear Phenomena* 10 (1): 1–35.

Xie, C., B. Huang, C. Claramunt, and C. Chandramouli. 2005. "Spatial logistic regression and GIS to model rural urban land conversion". Technical report.

Yang, Q., X. Li, and X. Shi. 2008. "Cellular automata for simulating land use changes based on support vector machines". *Computers & geosciences* 34 (6): 592–602.

AUTHOR BIOGRAPHIES

OMAR CHARIF is a PhD fellow at the University of Technology of Compiegne, France and CEPS/INSTEAD, Luxembourg. He received a MSc in Intelligent Transport Systems from UTC, France. His research interest lie in using artificial intelligence techniques in specific ANN, CA and Multiagent systems for modeling the human mobility. His email address is omar.charif@ceps.lu.

HICHEM OMRANI is a researcher in GEODE Department of (Geography, development and spatial analysis) of CEPS/INSTEAD research institute. He earned a PhD and a Master degree in Information Technology from UTC (University of Technology of Compiegne-UTC, France) in 2007 and 2004 respectively. He received the engineering and master degrees in computer science from the ENSI (National School in computer science) and UTC, in 2002 and 2004, respectively. He is interested in cellular automata and multi-agent systems and their applications in mobility modeling. He is the author of several international peer-reviewed journal and IEEE/ACM conferences proceeding papers. His email address is hichem.omrani@ceps.lu.

REINE MARIA BASSE is a post doctoral fellow in geography at GEODE, CEPS/INSTEAD. She received her doctorate in geography and urban planing form the University of Nice-Sophia Antipolis, France. Her research interests are in Cellular Automata, Urban simulation and Land cover changes. Her email address is reine-maria.basse@ceps.lu.

PHILIPPE TRIGANO is a professor at the University of Technology of Compiegne, France. His research interest lie broadly in Human Machine Interaction(HMI), e-learning and software engineering. He is the author of several peer-reviewed journal and conference papers. His email is philippe.trigano@hds.utc.fr.