Cellular automaton model for land-use change simulation

Andrés Manuel GARCÍA*, Inés SANTÉ[a], Rafael CRECENTE[a], Marcos BOULLÓN[a], Marcos SUÁREZ[a]

[a]Land Laboratory, Department of Agricultural and Forestry Engineering, University of Santiago de Compostela.

Escola Politécnica Superior, Campus Universitario s/n, 27002 Lugo, Spain

*Corresponding author: E-mail: andresmanuel.garcia@usc.es. Tel.: +34982252231 ext. 23292; fax +34982285926.
E-mail addresses: ines.sante@usc.es (Inés Santé), rafael.crecente@usc.es (R. Crecente), marcos.boullon@usc.es (M. Boullón), marcos.suarez@usc.es (M. Suárez)

Abstract:

Urban cellular automata have been widely used to simulate urban growth, due to their capability of reproducing complex emergent dynamics using simple rules which are based on the proximity of neighbouring land-uses.

One of the models that had a greatest impact in this domain is the model proposed by White and Engelen, because it is able to reproduce the dynamics of several land-uses using an extended neighbourhood, where the influence of the proximity of various land uses can be considered. This allows a greater flexibility and modelling power, using transition rules which are more consistent with the orthodox cellular automata structure.

The main shortcoming of this model is that it has a lot of coefficients to calibrate and therefore most of the models inspired on it are adjusted by trial and error or by expert knowledge. These methods reduce the accuracy of the results and make the calibration more cumbersome.

In this paper it is proposed a simplification of this kind of models, aimed at reducing the number of parameter so as to use an agent based model to automate the calibration and improve the model results. This will be achieved using logistic regressions to produce accessibility and aptitude maps and using linear functions to represent the influence of each neighbouring land use on the central cell.

Introduction

Cellular automata (CA) have been widely used to simulate complex urban growth processes, since CA can reproduce this kind of dynamics from simple rules. This characteristic provides these models with a great capacity of analysis of the simulated processes. In adition, as urban CA usually operate on a two-dimensional Euclidean space consisting of an array of square cells, they can be easily integrated in a GIS environment and incorporate raster data derived from aereal photographs or satellite images. At the same time, the GIS allows to visualize the results of the model and, therefore, to facilitate their interpretation.

CA were developed in the forties but did not reach a great development as urban growth simulation tools until the nineties. Most of these models were developed for research purposes, such as testing urban theories or studying the dynamics of real cities. In most of the cases, these models did not transcend beyond the scientific community due to: the lack of flexibility and adaptability to the simulation of different dynamics, the complexity of the transition rules which hinders the analysis of simulated processes, the lack of graphic interfaces to make models more user friendly and the lack of reliable calibration and validation methods.

Most of the CA models designed to simulate urban growth processes were applied to big cities such as: Bufalo (Xie, 1996), Cincinnati (White et al. 1997), San Francisco and the metropolitan area of Baltimore/Washinton (Clarke and Gaydos, 1998), Dublin (Barredo et al. 2003), Lagos (Barredo et al. 2004), Tokio (Arai e Akiyama, 2004) or San Diego (Kocabas and Dragicevic, 2006). In this kind of cities urban growth is high and fast, so urban processes are easily identified. In this context, information is usually enough to characterize the simulated dynamics. On the other side, the evolution of small urban areas characterized by low and slow growth processes has been rarely simulated. In this type of areas more complex models are required to better capture the urban dynamics and reproduce accurately the urban growth processes (García et al., in press). As aforeseen, the calibration of this type of models is more difficult because of their complexity and the high number of parameters used to calibrate them.

In this paper, an urban CA model that keeps the flexibility and potential of the model of White et al. (1997) but uses a lower number of calibration parameters is presented. Logistic regressions and a simplification of neighbourhood distance decay curves were

used to reduce the number of calibration parameters and a genetic algorithm was designed to improve the calibration of the urban CA. This model has been evaluated by applying it to the simulation of urban growth in Ribadeo, a small village of NW Spain.

Methodology

Selection of the base model

Formal CA rules present a series of shortcomings by the time of simulating urban growth processes. Couclelis (1985) proposed a set of relaxations of the formal rules, so that CA models could better represent urban growth dynamics. For example: formal CA usually consider two kinds of neighbourhood; the Moore neighbourhood (eigh cells adjacent to the central cell) and the von Newman neighbourhood (four cells adjacent to the sides or corners of the central cell). In real cities, the influence of neighbouring land-uses usually reaches higher distances that depend on the kind of considered land-uses. Another example is the uniform space considered in formal CA. Actually, the space presents a series of characteristics such as topographic irregularities or planning restrictions, among other factors, that condition the location of certain land-uses. There are also restrictions that are introduced into CA models to account for the temporal scale in which simulated processes take place. Nevertheless, if too many relaxations are considered to obtain more accurate results, models may get too complex and their transition rules may be quite far from those of formal CA (Santé et al. 2010). Furthermore, complex models are difficult to calibrate and the analysis of their results is quite cumbersome.

In the case study presented in this paper, another difficulty arises from the fact that Ribadeo is a small village, characterized by slow and low urban growth processes, which makes more difficult to identify clear growth patterns and the dynamics that generate these patterns. This leads to the need of making more complex simulations that can reflect with sufficient level of detail the particularities of the simulated areas (García et al. 2009). Consequently, in the study area is complicated to find a balance between the model simplicity and its capability to produce accurate simulations.

There have been too many models developed so far, which have been used both at a theoretical level to study urban dynamics and at a practical level to simulate urban growth in real cities. Among the latter models, which were developed to simulate urban

growth in real cities, the model of White et al. (1997) has been selected for several reasons. First, this model allows simulating the dynamics of several types of land-uses, avoiding errors derived from grouping several land-uses in the same category (García et al., in press). Considering the interactions between the different land-uses in the study area allows a greater detail in the simulation of land-use change dynamics.

Another advantage is the flexibility of the neighbourhood used in this model. This neighbourhood considers the influence of the proximity of neighbouring land-uses on the transition probability of the central cell. Hence, several kinds of iteractions can be considered and the model may be better adapted to the particular characteristics of each area.

Last, the transition rules of the model of White are quite close to the formal CA rules, thus the simplicity of CA models is conserved and the results of the model of White are easily analyzed.

The model of White et al. (1997) considers two kinds of land-uses; fixed land uses, which influence the dynamics of other land-uses but do not take part in the dynamics, and the active land-uses, which influence and participate in the dynamics of other land-uses. The transition potential of a cell from land-use h to each active land-use j is calculated according to the equation 1. Then, cells transition to the land-use for which they have the highest potential in each iteraction.

(ec. 1) $P_{hj} = vs_j \left(+ N_j \right) + H_j$

For the calculation of the neighbourhood effect (equation 2), N_j, the model uses a circular neighbourhood with a radius of 6 cells, where each cell is weighted with the coefficient m_{kd}, whose value depends on the land-use k of the cell and on its distance to the central cell, d, and I_{id} is equal to 1 if cell i at the distance d has the land-use k and equal to 0 otherwise. The flexibility of this type of neighbourhood is due to its ability to model the interactions among the different land-uses.

(ec. 2) $N_j = \sum_{k,i,d} m_{kd} I_{id}$

In addition to the neighbourhood effect, the model includes a suitability coefficient, s_j, of the cell for the land-use j, an inertia parameter, H_j, that models the resistence of land-use h to change to land-use j and a stochastic variable, v, calculated with equation3:

(ec. 3) $v = 1 + \left[-\ln(rand) \right]^{\alpha}$

where *rand* is a random number between 0 and 1 and α is a coefficient that controls the degree of randomness.

The flexibility and simplicity of the model of White et al. (1997) has enabled it to serve as a basis to inspire many other models (Engelen et al., 1999; White e Engelen, 2000; Barredo et al. 2003; Yüzer, 2004).

Design of the model

The main shortcoming of the model of White is the high number of parameters, and the consequent difficulty of calibration. In most of the cases, a trial and error calibration method is used, which is usually time-consuming and compromises the accuracy of the simulation. Consequently, automatic calibration methods that simplify the calibration tasks and make models more user-friendly are needed. For this reason, the model of White et al. (1997) has been modified, with the aim of reducing the number of parameters and improving the calibration process.

In the proposed model, land-uses were classified in the same way as in the model of White:

- Fixed land-uses; they influence the dynamics of other land-uses but they not change. In the case study, water bodies, roads, institutional land-uses, green areas and railway tracks were chosen as fixed land-uses.

- Active land uses; they take part in the dynamics and transition to other land-uses. In the case study, commercial, industrial and residential land-uses were chosen as active land-uses.

Agricultural and forest land uses were considered as fixed land-uses because they influence the simulated dynamics but do not participate in them, except as a land

reserve for urban uses. Thus, cells with these land-uses may transition to urban land uses, but forest and agricultural land dynamics are not simulated.

Equation 1 is used to calculate the transition potential for each land-use. Suitability is calculated by using a logistic regression to reduce the number of calibration coefficients. To this end, a logistic regression was carried out by using as independent variables those shown in Table 1 with the aim of identifying the variables that explain better the transition probability for each active land-use. Among these variables, those with a Pr(>|z|) value lower than 0.001 were chosen and used to calculate the suitability maps for each active land-use with a second logistic regression (fig. 1, 2 e 3).

The inertia coefficient reflects the difficulty for a specific land-use to change to other land-use. This coefficient must be calibrated. In the case study, the number of transitions from urban land-uses to non urban land-uses or to a different urban land use is very low. For this reason, the calibration of the inertia coefficient for urban land-uses is avoided by hindering the transition of urban land-uses to other uses. This is achieved by increasing the coefficient m_{kd} of the central cell when the land-use of this cell is the land-use for which the transition potential is being calculated. This way, the transition potential of an urban cell for its current land-use will be very high, so the cell will not transition.

The introduction of the aforementioned modifications in equation 1 gave rise to equation 5:

(Ec. 5) $P_{hj} = R_j * v * s_j^{\beta} * (+ N_j)$

The suitability coefficient, s_j, is calculated by means of a logistic regression, so it takes values between 0 and 1. Nevertheless, the neighbourhood effect, N_j, can take values much higher. For this reason the index β is introduced to scale the suitability factor. The stochastic variable, v, is already scaled with the coefficient α of equation 3. In addition, a series of restrictions are considered (R_j) to take into account the excluded areas for land-use j due to planning restrictions and to the presence of elements that may hinder the land-use change, such as coastal cliffs, cemeteries, churches, closed landfills, the training field of an agriculture school or an aerodrome. R_j takes the value 0 if the land-use j is excluded and 1 if it is allowed.

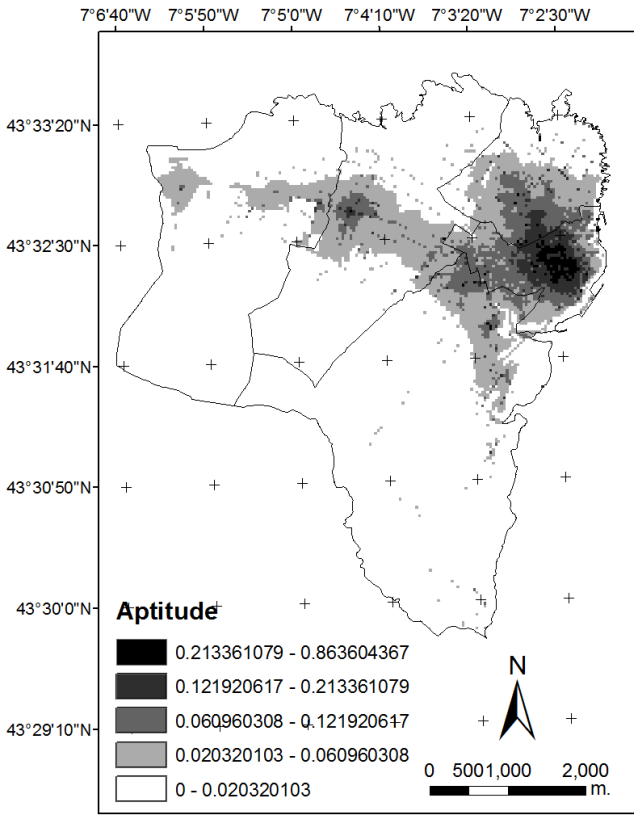
Table 1 Results of the logistic regression used to identify the variables that better explain the transition potential for each active land-use.

	Commercial			Industrial			Residential		
	Coefficients	Pr(>|z|)		Coefficients	Pr(>|z|)		Coefficients	Pr(>|z|)	
Area of parcels	-1.88E-05	0.05		-8.05E-06	0.009	**	-1.39E-05	2.78E-12	***
Distance to water bodies.	0.003	0.0002	***	0.001	2.32E-09	***	0.0008	7.27E-13	***
Distance to de Ribadeo down town	-7.28E-05	0.7		-0.0002	0.04		-0.0004	7.21E-23	***
Distancia to municipal roads	0.004	0.02		-0.0007	0.1		-0.002	1.26E-11	***
Distance to main roads	-0.02	2.14E-07	***	-0.002	3.19E-11	***	-0.002	2.18E-18	***
Distance to secondary roads	-0.0009	0.4		-0.002	1.33E-06	***	-0.002	5.01E-18	***
Distance to railway tracks	0.00282593	0.04		0.001	0.0004	***	0.0001	0.4	
Shape index of parcels	0.2	0.09		0.1	0.0009	***	0.2	8.65E-11	***
Slopes	0.01	0.7		-0.006	0.7		-0.04	1.28E-05	***

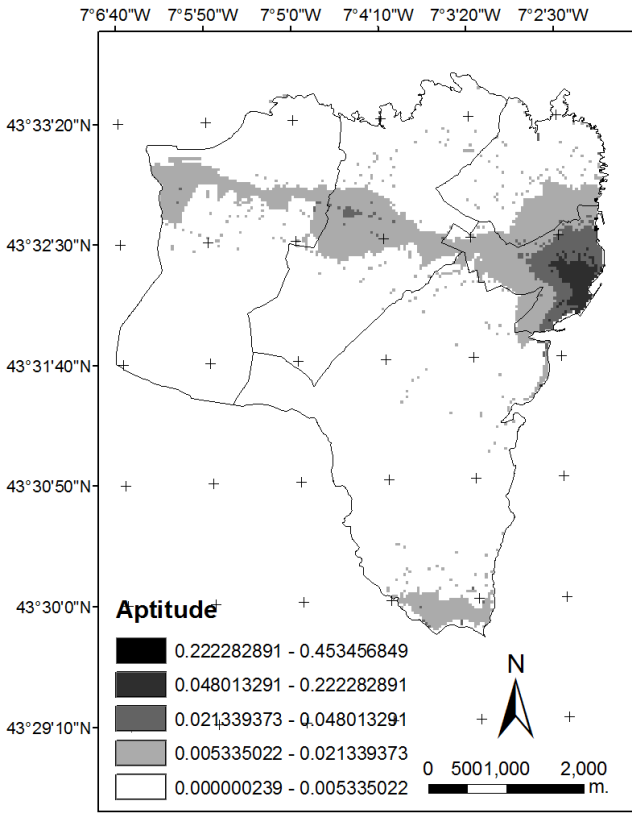
*** Pr(>|z|) < 0.001.

Figure 1 Suitability maps of active land-uses.

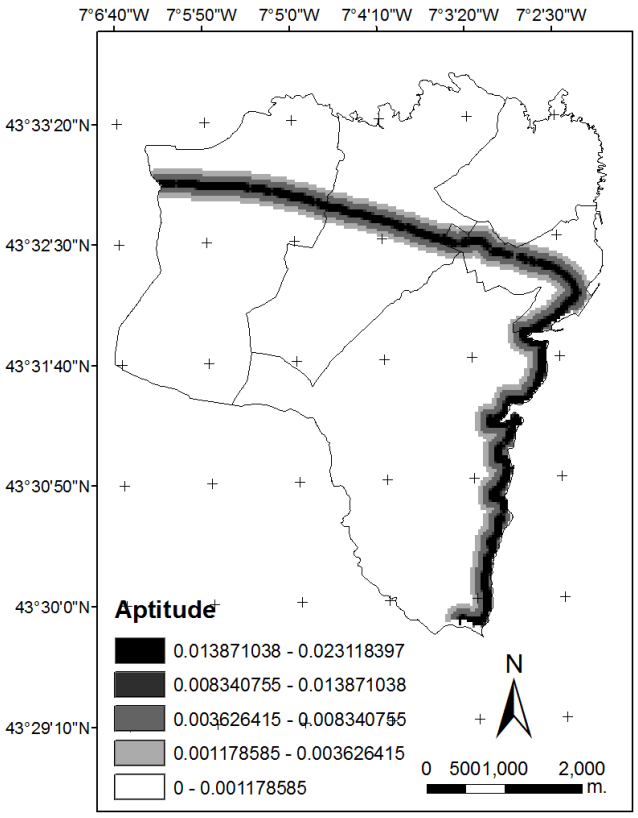
a) Residential



b)Industrial



c) Commercial



In spite of having reduced the number of calibration coefficients by using logistic regressions, the model is still complicated to calibrate due to the high number of parameters required to calculate the neighbourhood effect. In the proposed model a 3 cell radius neighbourhood is used. The influence of the neighbouring land-uses on the transition potential of the central cell could be calculated by using a function of the distance. However, this function could adopt several shapes (Xie, 1996). Bearing this in mind, this mathematical function was simplified by using two lines (a+bx and c+dx), this way, several distance decay influence functions can be represented in a simplied way, using only 4 parameters (a,b,c e d) for each couple of land-uses (fig. 2).

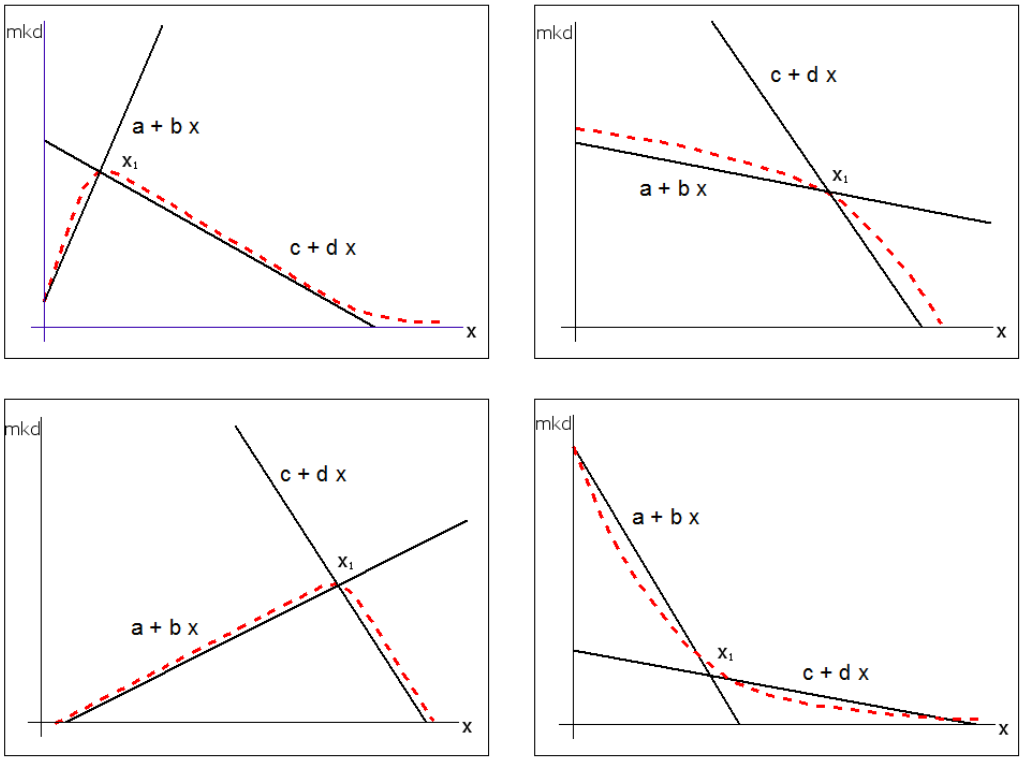


Figure 2. Examples of the simplification of distance decay functions that determine the influence of a land-use h on a land-use j for a distance (x), using two lines (a+bx and c+dx).

Consequently, a coefficient m_{kd} for each land-use k and distance d to the central cell in the neighbourhood is not needed. In the original model, with a 3 cell radius circular neighbourhood, it would be necessary to calibrate 7 coefficents for each land-use (figure 3). In the proposed model these coefficients can be calculated with the equations of the two lines defined by four parameters (a, b, c, d). To this end, the cross point between the

two lines is calculated with the equation 6, so, for values of x between 0 and x_1, the line a+bx is used, and for values greater than x_1 the line c+dx is used.

(Ec. 6) $x_1 = \frac{a-c}{d-b}$

			6			
	5	4	3	4	5	
	4	2	1	2	4	
6	3	1	0	1	3	6
	4	2	1	2	4	
	5	4	3	4	5	
			6			

Figure 3. Cells of a 3 cell radius circular neighbourhood, numbered from lowest to highest according to the distance to the central cell.

Calibration

In spite of the reduction in the number of parameters, this number remains high, therefore the model is still complicated to calibrate. In these cases heuristic methods, such as simulated annealing or genetic algorithms (GA), are usually used to improve the calibration process.

Genetic algorithms have been used in several studies to calibrate CA models (Shan et al. 2008, Li et al. 2007, D'Ambrosio et al. 2006, Goldstein , 2003). Al-Ahmadi et al. (2008) have proved that GA are more efficient than simulated annealing by the time of calibrating CA. GA were developed by Holland (1975) and reproduce the mechanisms of genetic evolution of species to find a solution. A GA algorithm is initiated by creating an initial random population of possible solutions. Each individual is composed by a cromosme in which the alleles are equivalent to the coefficients to be calibrated. Then, a fitness function is used to calculate the level of goodness of the solution represented by each individual. The best individuals of the initial population are selected to cross-breed and produce a generation that will be evaluated again to select the parents of the following generaton. In each generation mutations are introduced. These mutations vary randomly the values of some alleli so that the algorithm does not reach a local optimum. After several generations a near optimal solution is produced.

The GA consists of the following steps: initialización (the initial population of random individuals is generated), evaluation, selection (the best individuals are selected), cross-breeding (the selected individuals are cross-breeded to get the next generation) and mutation (mutations are introduced). There are many methods to implement each step (Goldstein 2003), and there is not an optimal method for each case; thus it is necessary to find the most suitable method for each specific case.

In the case study, a population of 1500 individuals was created with 114 alleli for each individual (108 alleli correspond to the parameters a, b, c and d for each pair of land-uses, 3 alleli correspond to the α coefficient of the stochastic variable and 3 alleli correspond to the β coefficient of the suitability factor). These alleli were generated randomly. The α coefficients were forced to vary between the interval (0,10] and the β coefficients were forced to vary between the interval (0,3]. The parameters b and d, which determine the slope of the lines, were assigned by generating a random angle and calculating its tangent. The parameters a and b were generated with a range of values within the interval [-100,100], so that the values of the neighbourhood were within a range similar to those proposed in the model of White et al. (1997).

Once the initial population was generated, parents were evaluated and selected. The tournament method was used, where two individuals are randomly chosen and that with the highest validation index is selected to be one of the parents. When two parents are obtained, they are cross-breeded to produce two sons. In the cross-breeding process two random cross points are stablished in the genes, in this points a recombination of the parents' alleli is produced. In the process of tournament every individual in a generation was forced to compete for being a parent at least once. The individual of the parent population with the best validation index survives in the next generation. Once the descendants are obtained, a mutation tax of 0.008% was applied to all individuals with the exception of the survivor of the parent generation.

Evaluation

validation evaluation function is needed to validate the results produced by each individual and identify the individuals that produce better simulations, so as to select the individuals that are going to produce the next generation.

The visual comparison of the results with the real data is the most common method to validate the results of an urban CA model. However, this is a qualitative method that do not allow to estimate cuantitatively the degree of correspondence between real and simulated data. Furthermore, visual comparison is complicated to use at an operational level in a GA. This is the reason why a series of mathematical indices were defined to calculate numerically the accuracy of the results of each individual. The most used indices are those that calculate the cell by cell coincidence between simulated and real maps. Nevertheless, a simulated map may have a very poor cell by cell coincidence and, at the same time, accurately reproduce growth patterns. Therefore, indices that calculate the degree of coincidence between the real and simulated growth patterns are needed.

Indices that measure cell by cell coincidence have a shortcoming due to the fact that these indices do not take into account that, if a simulated cell does not match the real one but is located near it, the result is better than if it was located in a more distant position. This is the reason why the index proposed by Pontius (2002) was selected for this study. To calculate this index (equation 7) the real map R and the simulated map S are ran over with windows that have several resolutions g (1 cell of side, 2, 3, 4, … n cells of side) in each window, and the number of cells n of each land-use j in the real map Rn, j and in the simulated map Sn, j is calculated. The minimum value for each land-use is chosen between the values for the two maps and the minimum values for all uses in the window are additioned. Then, each window is weighted by a coefficient Wn that is equal to the number of cells that the window covers and, finally, all values of each window are additioned and divided by the total number of cells of the map, $\sum_{n=1}^{Ng} Wn$.

(ec. 7) $$Pg = \frac{\sum_{n=1}^{Ng}\left[Wn \sum_{j}^{J} MIN(Rn, j, Sn, j) \right]}{\sum_{n=1}^{Ng} Wn}$$

Equation 7 is used to calculate an index for each window resolution Pg. This index has the value 1 if the number of cells of every land use in the real map and in the simulated map match and 0 if this number does not match for any land-use. In the fitness function used for the evaluation of the results the global Pontius index P (equation 8) was used.

This index is the result of a weighted addition of the indexes for all the window resolutions, where the indices of lower window resolutions were weighted more than those of higher window resolutions. The weighting coefficient Vg for each resolution g was obtained by an exponential curve (equation 8) which was scaled with the coefficient b, which was given a value of -1.2.

(ec. 8) $$P = \frac{\sum_{g}^{G} \exp^{b \times g} \times Pg}{\sum_{g}^{G} \exp^{b \times g}}$$

In addition to the previously described index, which measures the cell by cell coincidence, several spatial metrics have been used to compare the spatial pattern of the urban growth in the real map and the simulated map. The number of patches (NP), the mean patch area (AREA_MN) and the area weighted fractal index (FRAC_AM) were chosen. Other indices were not considered because they are normally correlated with the aforementioned ones (such as the largest patch index (LPI) which is correlated with NP and AREA_MN or the area weighted mean shape index (SHAPE_AM) which is correlated with FRAC_AM).

In the evaluation process, only the spatial metrics for active land uses were considered. The values of the spatial metrics calculated for the simulated maps, were substracted to those calculated for real maps. The inverse of the addition of the absolute values of these substractions was used as an evaluation index. This way, the closer to 1 the index is the more similar to reality the simulations will be (the differences between real and simulated values will be closer to 0), and the closer to 0 the index is, the farther the simulations will be from reality.

So that all indices contribute in the same way to the final evaluation index, the result of substracting real and simulated values was normalized between 0 and 1. This normalization was carried out by dividing each result by the maximum value that the result of the substraction of the real and simulated value could reach. In spite of this normalization, indices may vary between different ranges of values. Therefore, the indices were weighted by the coefficients shown in Table 2 in order to make them to vary within the same range.

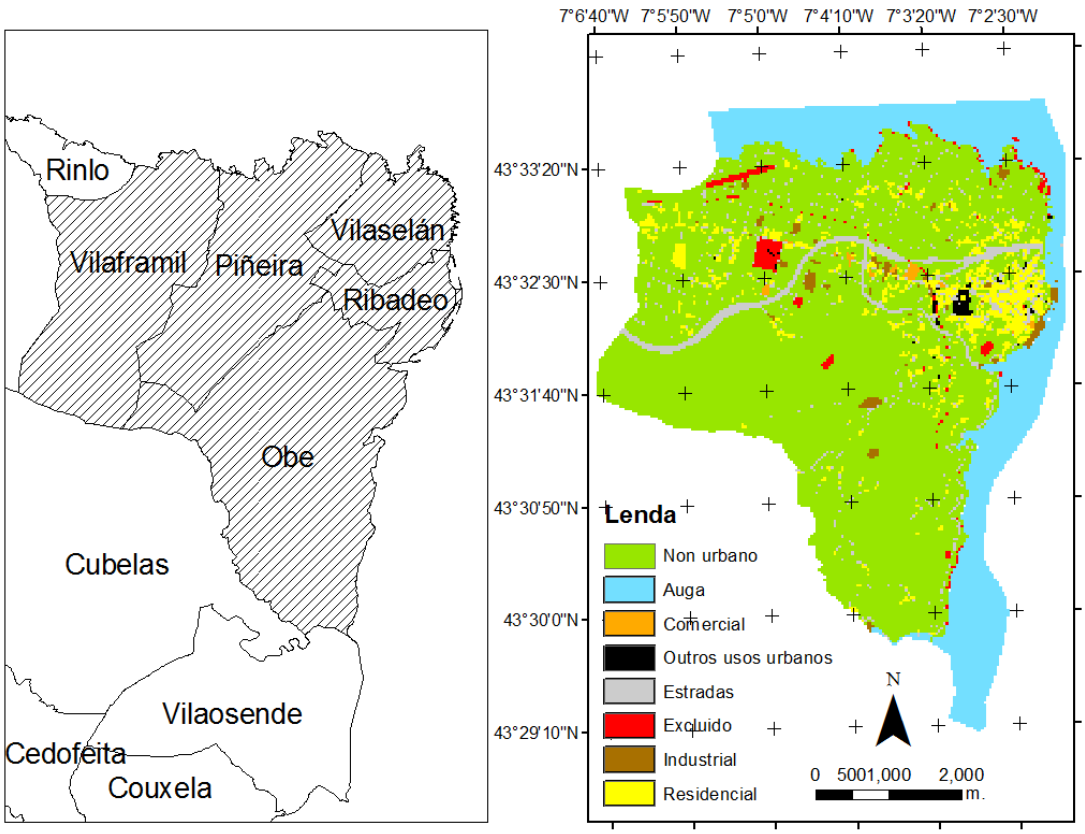
Table 2 Weighting coefficients for the evaluation indices.

ÍNDEX	WEIGHTING COEFFICIENTS
PONTIUS	1/3000000
NP	1/5000
AREA_MN	1/1.1
FRAC_AM	1/10

Results

The study area comprises the parishes (submunicipal administrative division in the region of Galicia) of Ribadeo, Vilaselán, Piñeira, Vilaframil and Obe (figure 4). For the model calibration two land-use maps of the years 1978 and 1995 with a resolution of 35 x 35 m were used. The genetic algorithm was run for 24 generations; until the maximum validation index did not increase during several generations.

Figure 4 Map of the study area and land-use map of the year 2007.



The coefficients obtained with the genetic algorithm were used to simulate the land-use change between 1995 and 2007 with the urban CA model. A land-use map of the year

2007 was used to validate the results. The amount of growth of each active land-use in each iteration of the model was calculated by dividing the real growth in the simulated period between the number of iterations of the model (1 iteraction for each year). The cells transitioned to the land-use with the highest transition potential, so that the simulated growth was equal to that estimated in each iteration for that land-use. If the cells with the highest transition potential for a certain land-use were not enough to achieve the real growth of that land-use, those cells, whose second highest transition potential corresponded to that land-use, would transition to that land-use, and so on. The simulation results were evaluated using the index proposed in Pontius (2002), spatial metrics and visual comparison.

The value of the index described in Pontius (2002) for the real and simulated map is 0.961. This result is very close to 1, therefore most of the simulated cells were located close to the real ones. This indicates that the simulation was quite good.

In spite of obtaining good values for this index, these results do not guarantee that simulated growth patterns are close to real ones. This is the reason why spatial metrics were used to check the match between simulated and real growth patterns (Table 4).

Table 4. Comparison of the spatial metrics of the simulated and real land-use patterns.

		NP	LPI	AREA_MN	FRAC_AM	ENN_AM
Simulated 2007	Industrial	56	0.11	0.63	1.1	148.23
Real 2007	Industrial	45	0.09	0.77	1.06	276.24
Simulated 2007	Residential	272	0.44	0.65	1.13	89.53
Real 2007	Residential	224	0.98	0.78	1.16	88.79
Simulated 2007	Commercial	61	0.01	0.17	1.02	106.89
Real 2007	Commercial	13	0.06	0.74	1.05	252.12

The values of the spatial metrics shown in Table 4 proved that simulated patterns of residential land-use are quite close to the real ones. The number of patches is a bit higher and the largest patch index a bit lower; however, the mean patch area, the FRAC_AM and the ENN_AM are quite close to the real one.

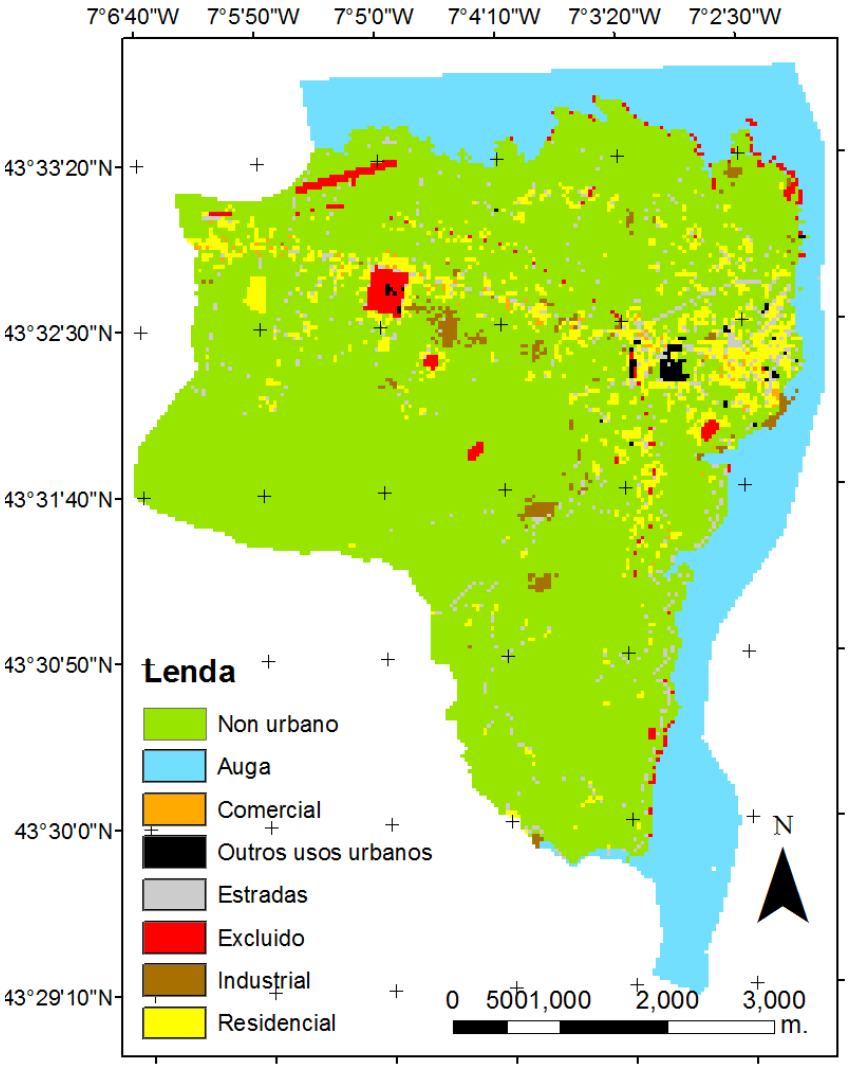
The patches of industrial land-use are closer to one another in the simulated map than in the real map. The model tended to group growth around existing patches. Most of the industrial land-use growth was produced by the growth of the port of Ribadeo and by the construction of industrial warehouses along the main road. This gave rise to a

disperse growth in large patches. There is not any factor that clearly influences the location of the disperse patches of industrial land-use, with the exception of the growth of the port and of a wide buffer along both sides of the main road. Hence, the spatial pattern of industrial land-use is very path dependent. According to Brown et al. (2005), this kind of growth is very influenced by estocasticity and it is very difficult to predict.

The growth patterns of the commercial land-use do not match well the real one. This is because the growth in the calibration period was too scarce and thus it was difficult to identify the factors that influence it.

In the map which was simulated by the proposed model for the year 2007 (Figure 5) it can be observed that the growth patterns are quite close to the real ones. The growth of the urban core of Ribadeo was slightly underestimated, whereas the growth around the area where the main road converges with an important secondary road was overestimated. As both infrastructures attract residential growth, a large patch of residential land-use was created in the area. The industrial land use, in spite of being located along the main road, tended to be concentrated around an existing industrial patch in the center of the map.

Figure 5 map of the year 2007 simulated with the proposed model.



In summary, the growth that took place on the north of Ribadeo urban core and along the main road was simulated correctly. The simulated patterns of small urban patches located at the south of the town and along the main road are also close to the real ones.

Conclusions

Results show that the proposed model has been able to simulate quite accurately urban growth processes in Ribadeo. The modifications of the base model allowed to simplify it, keeping its analysis capacity. This simplification allowed to use an heuristic method to improve the calibration tasks.

The simulation of the urban growth in the study area with the proposed model proved that genetic algorithms are a good calibration tool for urban CA models.

Future research lines should address the development of better validation methods which can capture the complexity of simulated growth patterns and assess the degree of coincidence with reality. In addition to genetic algorithms, there are other methods that can be used to calibrate urban CA models. In fact, some authors have used data mining techniques to find CA trainsiton rules. This technique should be explored in greater deep to find new calibration methods that can capture the simulated dynamics more accurately and produce better results.

Bibliography:

Al-Ahmadi, K.; See, L.; Heppenstall, H.; Hogg, J. (2008) Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. *Ecological Complexity,* 6(2), 80-101.

Arai, T.; Akiyama, T. (2004). Empirical analysis for estimating land use transition potential functions – case in the Tokyo metropolitan region. *Computers, Environment and Urban Systems*, 28(1-2), 65-84.

Barredo, J. I.; Demicheli, L.; Lavalle, C.; Kasanko, M.; McCormick, N. (2004). Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. *Environment and Planning B: Planning and Design,* 31(1), 65-84.

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

Brown, D. G., S. Page, R. Riolo, M. Zellner, and W. Rand, 2005. Path dependence and the validation of agent-based spatial models of land use. Int. J. Geogr. Inf. Sci. 19, 153-174.

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

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

D'Amborsion, D.; Spataro, W.; Iovine, G.(2006) Parallel genetic algorithms for optimising cellular automata models of natural complex phenomena: An application to debris flows. *Computers & Geosciences,* 32(7), 861-875.

Engelen, G., Geertman, S., Smits, P., Wessels, C., (1999) Dynamic GIS and strategic physical planning support: a practical application to the IJmond/Zuid-Kennemerland region. En: Stillwell, S., Geertman, S., Openshaw, S. (Eds.), Geographical Information and Planning. Springer-Verlag, Berlin.

García, A.M.; Santé, I., Crecente, R.; López, F.j.(2008) Analysis of the factors that condition the expansion of a coastal village trough cellular automata simulation. *Proceedings from the 12th International Congress in Project Engineering,* Zaragoza, July, 9-11, 2008.

García, A.M.; Santé, I., Boullón, M., Crecente, R., Miranda, D. A comparative analysis of cellular automata models for simulation of small urban areas in Galicia, NW Spain. *Computers, Environment and Urban Systems,* (in press).

Goldstein, N.C. (2003) Brains vs. Brawn-Comparative strategies for the calibration of a cellular automata-based urban growth model. 7th International Conference of Geocomputation, University of SouthHampton, UK.

Holland, J.H.; (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, vii, 183 pp.

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

Lavalle, C.; Barredo, J. I.; McCormick, N.; Engelen, G.; White, R.; Uljee, I.; (2004) The MOLAND Model for Urban and Regional Growth Forecast. A Tool for the Definition of Sustainable Development Paths, Technical Report EUR 21480/EN, Italy, Institute for Environmment and Sustainablitiy, Joint Research Centre, European Commission.

López, E.; Bocco, G.; Mendoza, M.; Duhau, E. (2001) Predicting land-cover and land-use change in the urban fringe: A case in Morelia city, Mexico. *Landscape and Urban Planning,* 55(4), 271-285.

Li, X.; Yang, Q.S.; Liu, X.P. (2007) Genetic Algorithms for determining the parameters of cellular automata in urban simulation. *Science in China, Series D: Earth Science*, 50(12), 1857-1866.

Myint, S. W.; Wang, L.(2006) Multicriteria decision approach for land use land cover change using Markov chain analysis and a cellular automata approach. *Canadian Journal of Remote Sensing*, 32(6), 390-404.

Pontius, R.G. (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing,* 68(10), pp.1041-1049.

Santé, I.; García, A.M.; Miranda, D.; Crecente, R. (2010) Cellular automata models for the simulation of real-world urban processes: A review and annalysis. *Landscape and Urban Planning*, 96(2), 108-122.

Shan, J.; Alkheder, S.; Wang, J. (2008) Genetic algorithms for the calibration of cellular automata urban growth modeling. *Photogrametric Engineering and Remote Sensing*, 74 (10, 1267-1277.

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

White, R.(1997) Cities and cellular automata. *Discrete Dynamics y Nature and Society*,2, 111-125.

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

Xie, Y. (1996). A generalizad model for cellular urban dynamics. *Geographical Analysis*, 28, 350-373.

Yüzer, M.A. (2004) Growth estimations in settlement planning using a land use cellular automata model (LUCAM). *European Planning Studies*, 12(4), 551-561.