Barreira González, P., Aguilera-Benavente, F. & Gómez-Delgado, M. 2015, "Partical validation of cellular automata based model simulation of urban growth: an approach to assessing factor influence using spatial methods", *Enviromental Modelling and Software,* vol. 69, pp. 77-89.

Available at http://dx.doi.org/10.1016/j.envsoft.2015.03.008

(Article begins on next page)

Title:

Partial validation of cellular automata based model simulations of urban growth: an approach to assessing factor influence using spatial methods.

Authors names and affiliations:
M.Sc. Pablo Barreira González* [a]
Ph.D. Francisco Aguilera Benavente [b]
Ph.D. Montserrat Gómez Delgado [c]
[a,b,c] Departamento de Geología, Geografía y Medio Ambiente, Universidad de Alcalá
Facultad de Filosofía y Letras, c/Colegios nº2, 28801, Alcalá de Henares, Spain.

*Corresponding author:
Tel. +34918856384,
Fax +34918854439

E-mail addresses:
pablo.barreira@edu.uah.es [a]; f.aguilera@uah.es [b] ; montserrat.gomez@uah.es [c].

Abstract:
Cellular Automata (CA) based models have a high aptitude to reproduce the characteristics of urban processes and are useful to explore future scenarios. However, validation of their results poses a major challenge due to the absence of real future data with which to compare them. A partial validation applied to a CA-based model for the Madrid Region (Spain) is presented as a proposal for determining the influence of given factors on the results and testing their spatial variability. Several simulations of the model were computed by different combinations of factors, and results were compared using flexible map comparison methods in order to study spatial pattern matches and similarities between them. Main and total effects of these factors were calculated for each method, by applying a simplified Global Sensitivity Analysis approach. Frequency maps showing the most frequent cells with changed land use in the results were generated.

Keywords:
Cellular automata models, urban growth, partial validation, map comparison, frequency maps.

Highlights:

- A set of runs were computed for all possible combinations of main input factors.
- Generated simulations were compared using spatial metrics and buffer zones.
- The main and total effects of each parameter were computed for each method employed, applying a simplified GSA.
- A stability (frequency) map was made for the different simulations generated to detect the more robust areas.
- This simple methodology is an attempt to improve modeller-planner communication and understanding.

1. INTRODUCTION

Land use change processes, and particularly the expansion of urban areas, are among the most significant changes currently taking place in Europe (European Environment Agency, 2006; Kasanko et al., 2006), the U.S.A. (Buyantuyev and Wu, 2009; Wu et al., 2011) and other less developed countries all over the world (Barredo et al., 2004). These phenomena have substantial environmental and regional implications that require meticulous assessment (Berling-Wolff and Wu, 2004; Grimm et al., 2008).

In this context, the usefulness of urban simulation tools is evident: besides helping us to better understand the processes and driving forces of urban growth (Cheng and Masser, 2003), urban simulation models can reproduce past land use dynamics and simulate the possible evolution of these processes in the future (Hansen, 2010; Paegelow and Camacho-Olmedo, 2008; Sante et al., 2010). They can therefore be useful for assessing changes in cities and metropolitan growth patterns (Aguilera et al, 2010; Aguilera et al., 2011; Berling-Wolff and Wu, 2004; Li et al., 2008) and the possible effects of such growth on the landscape and region where they occur (Aguilera and Talavera, 2009; Forman, 1995; Mitsova et al., 2011)

Among the several types of simulation models, Cellular Automata (CA) has shown a high capacity for reproducing the main characteristics of urban expansion processes (Benenson and Torrens, 2004; Batty, 2007), such as complexity or self-similarity (Itami, 1994; Frankhauser, 1998; Torrens, 2000). Previous theoretical work carried out by White and Engelen, (1993), Couclelis (1997), White et al (1997), Benenson and Torrens (2004), Couclelis et al (2005), Batty (2007) among others, and also existing applications (Barredo et al, 2003; Petrov et al, 2009; Hewitt et al, 2014) have shown the potential of CA-based models to simulate urban growth and land use dynamics in different contexts. The *NASZ* model scheme, a CA-based model framework proposed by White et al. (1997), which includes four parameters (*neighbourhood*, *accessibility*, *suitability* and *zoning status)*, has become widely popular when simulating urban growth, leading to a considerable proliferation of simulation tools and studies based on this approach (see Sante et al., 2010 and Triantakonstantis and Mountrakis, 2012 for a detailed review).

Urban growth simulation using CA-based models follows three main steps (Figure 1). Once a model is implemented, a calibration process is necessary in order to adjust its internal parameters. Calibration is performed at a past period of time (t_0 - t_1) using historical land use data, and in essence the model has to reproduce a real dynamic within that period and compare its prediction with real data at time t_1. When the comparison results are optimal, it is assumed that the model can reproduce past dynamics that continue to the present (demonstrated through a validation process using real data), enabling subsequent trend simulations whereby present dynamics are projected into the future. Hence, calibration and validation process are closely related since they compare simulation results with real data (Engelen and White, 2008). Many of the studies developed, have therefore adopted a predictive approach (Paegelow and Camacho-Olmedo, 2008), embodied in a methodology similar to that shown in Figure 1.

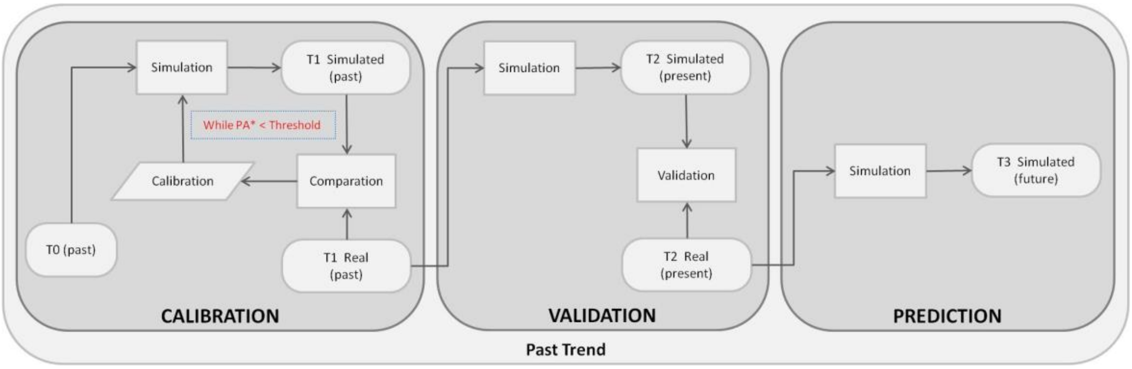

Figure 1. Simulation process with CA-based models. Validation is performed before simulation. *PA= Percentage of Agreement.

However, other studies (Berdoulay, 2009) have highlighted the problem, or at least the complexity, of adopting a purely predictive approach to simulations of urban dynamics based on a positivist conception of the urbanization process. Within this approach, urban CA models are developed (among other reasons) to improve urban planning. Therefore, simulating only past trend dynamics into the future may not be enough to inform planners about the possible futures that may happen. In addition to this, urban processes are subject to a significant degree of uncertainty related to political and administrative decision-making processes, unforeseen economic circumstances or emergence of new influential factors (Barredo et al., 2003).

Thus, we believe that the aim of these models should be to simulate several possible future situations under different circumstances. This alternative approach would employ urban growth simulation models in conjunction with future scenarios, not only to simulate past trends into the future (Figure 1), but also to explore different development possibilities (Figure 2), such as external exploratory scenarios (Borjeson et al., 2006).

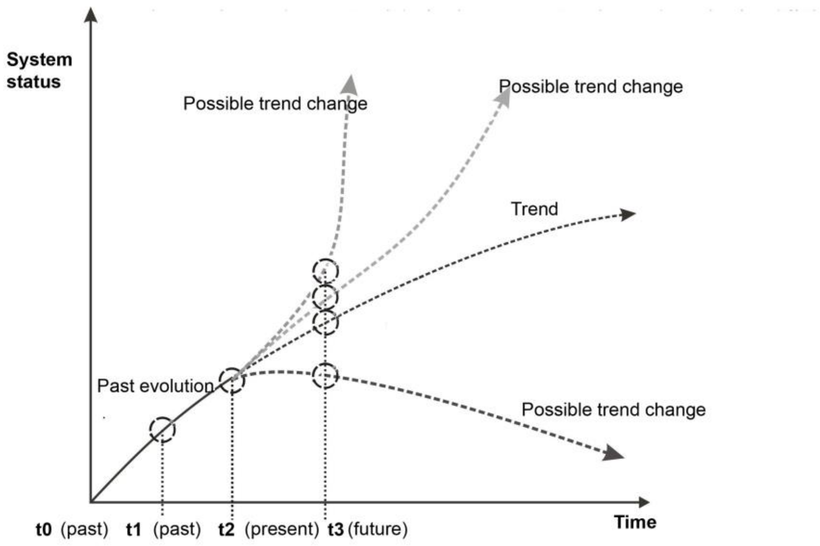

Figure 2. Possible changes on urban dynamics trends may make it necessary to use scenarios: evolution of time is shown in the axis of abscissa and the state of the different variables that describe the system, i.e. urban growth quantity, are shown in the axis of ordinates.

Hence, after the model calibration process and comparison with real data to confirm that the model reproduces past dynamics, the model may be used to simulate different future scenarios. These scenarios may require changes to the model in order to simulate different possible futures. Therefore, the results of future simulations should be subjected to a second validation process (Figure 3). The main problem at this stage (future time) lies in the fact that there are no real data to compare the model's simulation with. Therefore, the concept of validation in this context should be adapted and made flexible, since only partial validation is possible (Paegelow and Camacho-Olmedo, 2008).

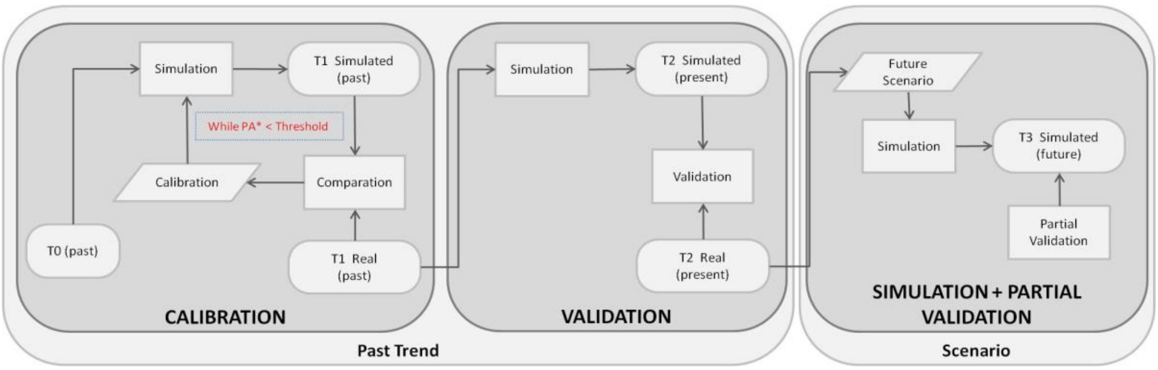

Figure 3. Validation and simulation approach to explore future scenarios. *PA= Percentage of Agreement.

There is no agreement within the scientific community about which set of procedures should comprise a validation process (see tables 1 to 6 in Santé et al., 2010), even less in relation to a partial validation. However, stability and robustness determination of the results through a Sensitivity Analysis (SA) is considered one of the most important processes. In the field of spatially explicit models, we can find many examples in the literature that consider variations in the input variables and analyse quantitatively the differences in the outputs (e.g. Crosetto et al., 2001; Crosetto and Tarantola, 2001; Store and Kangas, 2001; Gómez-Delgado and Tarantola, 2006; Lilburne and Tarantola, 2009; Ligmann-Zielinska and Sun, 2010, Chu-Agor et al., 2011, Plata-Rocha et al., 2012). In recent years, there is an increasing effort to develop SA based methods that obtain spatially-explicit representation of the results of the SA (e.g. Brown et al, 2005; Chen et al., 2010, Chen et al., 2013, Ligmann-Zielinska and Jankowski, 2014, Marrel et al., 2011). These can quantify the influence that each parameter implemented in the model has on the results generated. Such SA-based methods have been widely applied to spatial models such as urban growth or land use change models (based on Multicriteria Evaluation, Agent-based models, etc). With regard to the spatial CA-based models, these methods have traditionally been applied only in order to determine the influence of a single parameter, such as neighbourhood or randomness (Garcia et al., 2011; Kocabas and Dragicevic, 2006; Ménard and Marceau, 2005; Pan et al., 2010), or a given characteristic of the data used, such as spatial or temporal resolution (Dietzel and Clarke, 2004; Jantz and Goetz, 2005; Liu and Andersson, 2004; Samat, 2006; Yeh and Li, 2006).

However, bearing in mind the CA-based models' complexity and peculiarity, it may be of interest to conduct a partial validation assessing the influence of each of the model parameters on the simulation's results. This partial validation may be helpful to (i) determine whether each factor is driving the results as they were modelled to do (in other words, whether we are building a relevant model), (ii) to determine the spatial variation of the results due to each factor, and (iii) to determine the most robust spatial results in the area studied.

Within this context, the aim of the present study is to propose a general SA-based methodology for partial validation that can be applied to CA-based models. Our proposal is based on a simplification of global sensitivity analysis (GSA) approach (Saltelli et al., 2000). It consists of switching on/off the parameters of a CA-based model to generate different simulations. Comparing these simulations with the simulation generated by the baseline model (the model including the four parameters, or "true" model) we will be able to assess how the presence of the inputs parameters impacts the robustness of the baseline simulation and which of the model parameters needs to be improved in order to get closer to the real world representation.

To test this methodology we have used a CA-based model developed to simulate urban growth in one of the most urbanized regions in Spain (the urban region of Madrid) (Gómez-Delgado and Rodríguez Espinosa, 2012). It includes four parameters (*neighbourhood*, *accessibility*, *suitability* and *stochasticity*) that have been widely considered to be relevant in the development of similar models (e.g. Barredo et al, 2003; 2004; Petrov et al, 2009; Aguilera et al, 2010; Hewitt et al, 2014). The comparative study of the simulations generated was conducted by means of different map comparison methods including spatial metrics and distance buffers analysis. In addition, the main and total effects of the parameters implemented were calculated, applying a simplified Global Sensitivity Analysis approach. Finally, a location analysis of the areas that presented the greatest robustness was carried out with the aid of a map showing the cells which appeared most frequently in the simulations.

A detailed description of the methodology employed is given in section 2, together with a brief description of the model being validated, whilst the results obtained are presented in section 3. A discussion follows in section 4 and conclusions are given in section 5.

2. MATERIALS AND METHODS

2.1. Simulation model analysed and study area.

In recent decades, the Madrid region (Figure 4) has become one of the most dynamic areas of urban growth in the Iberian Peninsula and Europe (Plata-Rocha et al., 2010), as a result of strong residential growth spurred by the housing bubble (Burriel, 2011). This dynamism has also led to changes in the patterns of urban growth, with important consequences for the urban structure (OSE, 2006). Furthermore, the Madrid region contains a large amount of protected land, including some areas, such as the Sierra de Guadarrama, which has recently been incorporated into the National Parks network. This represents a significant constraint on urban growth, although the natural attractions of such areas often lead to an increased interest in developing the surrounding sites (Plata-Rocha et al., 2011).

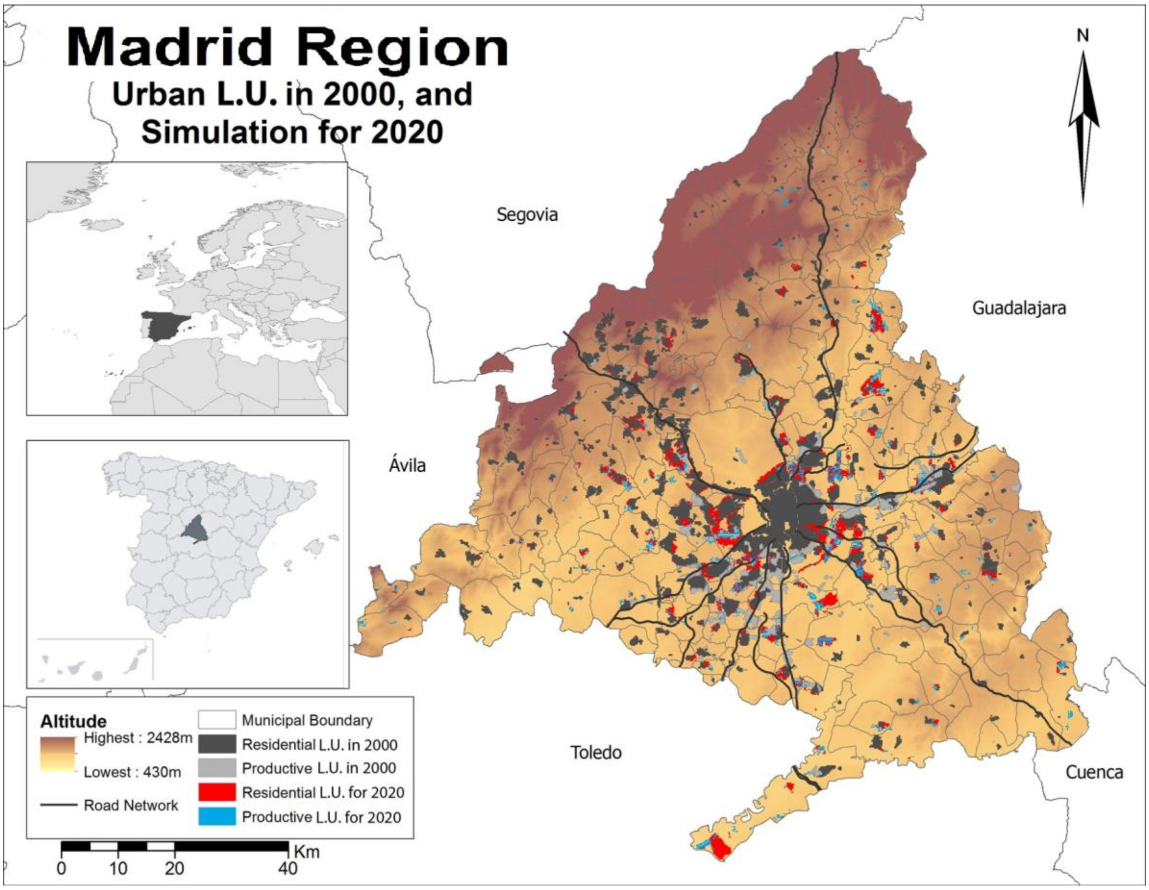

Figure 4. Study area. Land uses state in 2000 is represented and it is also shown the model's simulation for 2020.

Consequently, the Madrid region was considered an interesting place to study urban dynamics, and has already been used as a study area in the SIMURBAN project [1], which simulated several future scenarios (Plata-Rocha et al., 2011).

The proposed methodology was applied to a modified version of a model designed to simulate different urban growth scenarios in the Urban Agglomeration of Granada (Aguilera et al.2010; Aguilera et al., 2011), adapted to the specific characteristics of the Madrid region. This model was developed and implemented with IDRISI Andes software (Eastman, 2006). It is based on theoretical premises proposed by White et al (1997), which are the basis for other simulation models (Barredo et al., 2003; Petrov et al., 2009; Hewitt et al., 2014). The model uses a 200 m cell resolution and can simulate the location of new urban sites (the quantity of simulated urban growth should be established by the user) for three specific uses: productive land uses (industrial and commercial areas) and high and low density residential land uses. To simulate urban growth the model employs an urban land-use map (figure 4) as the main input. For each cell in the study area, the model estimates a transition potential for each of three simulated uses (Pj), obtained through a combination of four factors: neighbourhood, accessibility, suitability and a random

[1] http://www.geogra.uah.es/simurban1/

parameter, which were all transformed on a scale from 0 to 1. The transition potential (P) for each urban use (j) was finally obtained by combining these parameters with Equation 1:

$$P_j = N_j \times S_j \times A_j \times V_j$$

Eq [1]: Transition potential (P_j) equation where Neighbourhood (N_j), Suitability (S_j), Accessibility (A_j) and Randomness (V_j), are combined.

Neighbourhood (N) is an inherent factor in CA models. This parameter estimates the probability of change for each cell, depending on existing neighbouring urban land uses and the distance to the central cell. Each cell in the neighbourhood window exerts an effect of attraction or repulsion over the central cell for different urban uses. This effect depends on the type of urban land use existent in each of the neighbouring cells.

Accessibility (A) refers to ease of access to the road network and to certain other elements in the area (e.g. shopping and business centres) in each of the cells. Areas contiguous to the road network, shopping centres, etc. have a value of 1, and those farther away have a value of 0. For the rest, a linear transformation function is used.

Suitability (S) represents the capacity of each part of the study area to accommodate urban use, considering several variables such as slope, elevation, population growth or distance to water.

Stochasticity (R): also known as randomness. The objective of this parameter is to incorporate a random component typical of urban spatial processes (Batty and Xie, 1997). It is obtained with equation 2:

$$R = 1 + (-\ln(rand))^{\alpha}$$

Eq [2]: Randomness calculation (R), where *rand* is a random number between 0 and 1, and α is a parameter that permits an adjustment of the degree of perturbation (dispersion).

The value of α (0.3*)* was computed, on the basis of the radial dimension (Barredo et al, 2003), and validated to better simulate the urban growth patterns identified for the Madrid region.

Once the model was calibrated and validated, it was used to simulate urban growth from 2000 to 2020 under one specific scenario which takes into account a balance between social, economic, and environmental aspects (named "Innovation and Sustainability"). Compared with the growth recorded in the region of Madrid for the period 1987-2000 (OSE, 2006), this scenario proposes a reduction in growth of 17% (for the simulated period, 2000-2020) for residential land use and 25% for productive sector land use, implying a total growth of 508 ha/year for productive land use, equal to 127 cells/year, and 916 ha/year for residential land use (229 cells/year), divided between high and low density residential land use.

Based on these growth quantities and using the method described above, the model locates the different growths for each type of land use, generating a spatial simulation of urban growth for 2020 (starting in 2000 with annual iterations) (Figure 4).

2.2. Methodological proposal for partial validation.

The proposed methodology is based on a simplification of Global Sensitivity Analysis (GSA) methods (Saltelli et al., 2000) that have been already discussed and applied in spatial models (see e.g. Lilburne and Tarantola, 2009; Marrel et al. 2011; Ligmann-Zielinska and Jankowski, 2014). This method is easy to implement and computationally cheap.

The difference introduced in our proposal with respect to a strict GSA is that no probabilistic variations are introduced in input factors. Instead, we perform a progressive elimination of the original factors (see Table 1), with a full factorial design, and compute the model output for all possible combinations of the input factors (S1 to S16). These factors are treated as triggers since they are only switched on or off, i.e. each factor is characterized by a discrete distribution of only two values [0, 1] and the 16 simulations cover the full factorial design. By comparing the spatial results obtained from the different simulations it is possible to determine how they differ, to assess their robustness and to analyse with ease the consequent possible spatial implications for the studied area. In addition, it is possible to determine whether each factor is steering the results in the model in a likely direction, in other words, whether the model is correctly built.

FACTORS / SIMULATIONS	Neighbourhood (N)	Accessibility (A)	Suitability (S)	Stochasticity (R)
S1	X	X	X	
S2	X	X		
S3	X		X	
S4		X	X	
S5	X			
S6		X		
S7			X	
S8 (baseline model)	X	X	X	X
S9				X
S10	X			X
S11		X		X
S12			X	X
S13		X	X	X
S14	X		X	X
S15	X	X		X
S16				

Table 1: List of the model simulations computed and the factors which are included in each of them. S16 is a constant since it does not include any parameter.

To evaluate the results of the method, it was necessary to compare the simulation results among themselves, obtained from the different simulations of the model (S1 to S16). This assessment was performed by comparing the simulated maps obtained using different map comparison methods. These techniques have been substantially improved in recent decades and they encompass simple methods such as a visual comparison of the results (Barredo et al., 2003; White and Engelen, 1997), statistical indices to determine how many cells are the same in two maps (Pontius, 2002), fuzzy comparison metrics (Hagen, 2003), fractal dimensions and even the use of metrics related to or designed for other purposes (Aguilera et al., 2011; Aguilera et al., 2014).

Early methods such as cross tabulation or the Kappa index were based on cell-to-cell (pixel-to-pixel) comparison and provided information about the spatial agreement of the results. However, this form of comparison was excessively rigid since it compared each cell on one map with the same cell on the other but did not consider neighbouring cells. Moreover, the suitability of using these types of indices to compare maps has been widely debated (Pontius and Millones, 2011). Nevertheless, some authors have maintained that they provide more than merely spatial information on the maps comparison (Hagen, 2006), arguing that although the results generated by two models may not present exact spatial agreement (cell-to-cell), this does not imply that the results are far from similar.

Furthermore, a cell-to-cell map comparison is even less effective when analysing patterns and shapes of urban growth, where it is more appropriate to study how shape of urban fragments (patches formed by contiguous cells showing the same use) varies between one map and another (Li et al., 2008). Mas et al. (2012) further reaffirmed this stance in the case of models that simulate land use changes in various scenarios, where spatial agreement may not be as important as exploring the various patterns generated.

Therefore, the following methods were used to compare results and conduct the partial validation: the first method was the application of distance **buffers** (Tveite and Langaas, 1999), which, unlike visual comparison or the Kappa, can capture qualitative differences or similarities between two maps index (Power et al., 2001) . The second method involved the application of **spatial metrics** (Herold et al., 2003) to the maps obtained and their comparison, providing information on the shape, dispersion and fragmentation of land uses, enabling a comparison of results in terms of urban land use patterns (Mas et al., 2012). These two methods partially comply with the spatial (buffers) and aggregate (spatial metrics) validation proposed by Brown et al. (2005).

2.2.1. Evaluation of the results using distance buffers.

Distance buffers can be used to generate zones of influence around newly created cells in the reference simulation (S8) results and to determine whether the cells obtained from other simulation results are located within that area. Thus, the percentage of simulated land use located within and outside the area of influence can be calculated. If the percentages of agreement between the distance buffer and the simulation being compared with are high, this would indicate that both simulations generate similar results, even though the results are not in exactly the same cell as in the reference simulation.

This technique is based, to a certain extent, on the use of the epsilon band to verify the accuracy of linear elements mapping. This band has traditionally been used as an uncertainty zone within which there is a degree of probability that the real line will be located (Veregin, 1999). Based on the delimitation of this band, some authors have developed processes to measure the accuracy of the position of linear elements, through the generation of corridors (Goodchild and Hunter, 1997). Thus, Tveite and Langaas (1999) proposed a method called BOS (Buffer-Overlay-Statistics), whereby a corridor is established around two linear databases (the one evaluated and the most accurate one) and they are then compared using mathematical calculations of area, perimeter, etc, in addition to using different sized corridors. Similarly, Mas et al. (2012) reported an interesting approach comparing the results of two land use/cover change models with the actual changes observed through the generation of longer tolerance distances.

Given the foregoing and in accordance with the proposal made by Tveite and Langaas (1999), the process was repeated with differently sized buffers, although the distance of these must fit well, since if the buffer is very small the comparison will be similar to a cell-to-cell one and would not therefore provide any additional information. Conversely, if the buffer is too large, the comparison will lose relevance.

Taking this and the spatial characteristics of the study area into account, we generated distance buffers of 1000m and 2000m from the new urban areas simulated for each of the uses in the S1 results (which was used as the reference for comparison). In this simulation, the areas generated by the buffer formed part of the simulated area so that a comparison with other simulations would reveal whether the results were in approximate spatial agreement with those generated by the reference simulation.

Lastly, it should be noted that the use of map comparison methods to evaluate the degree of fitness or similarity between two maps raises a recurrent question: Should the total area of land use be used in the comparison, namely the initial and simulated area simultaneously, as is usually the case when calibrating this type of model? Or on the contrary, should only the simulated area be used in the comparison? In this latter case, utilization of the total area of land use would significantly increase the values of the numerical results, reducing the differences between each comparison, since the stable area on each map is much greater than the simulated area. Therefore, and in order to clarify the differences between each simulation resulting from the factors applied in each case, we chose to use solely the simulated area for this method.

2.2.2. Evaluation of the results using spatial metrics.

Spatial metrics (Herold et al., 2003) have been employed on multiple occasions to determine urban growth patterns (Aguilera et al., 2011; Buyantuyev and Wu, 2010; Dietzel et al., 2005; Li et al., 2008; Wu et al., 2011), on the basis of shapes and distribution of urban patches. It has been proven that these metrics can be applied to map comparison due to their capacity to quantify and subsequently compare the spatial and morphological characteristics of different thematic maps (Herold et al., 2005), and to compare urban shapes (Aguilera et al., 2011; Berling-Wolff and Wu, 2004; Schwarz, 2010). However, the selection of the metrics to employ in order to assess the spatial characteristics of urban growth (whether real or simulated) remains one of the fundamental questions when applying spatial metrics to the study of urban processes. There is a general lack of consensus on the set of metrics to use, although some authors have explored the most suitable metrics for the study in question (Schwarz, 2010).

In accordance with several studies (Aguilera et al., 2011; Buyantuyev et al., 2010; Dietzel et al., 2005; Hepinstall-Cymerman et al., 2013; Herold et al., 2005; Kaza, 2013; Seto and Fragkias, 2005; Wu et al., 2011), and with the main spatial features that would be useful for characterising urban settlement patterns and, therefore, for comparing results, it was decided to employ metrics that quantify: 1) the degree of growth fragmentation, for which the number of patches metric (NP) is selected (Table 2); 2) the shape of the patches with the Mean Shape index (SHAPE_MN); and 3) the relative distance between patches with the Mean Euclidean Nearest Neighbour metric (ENN_MN). Meanwhile, other, frequently used metrics in the literature, such as AREA_MN (Mean patch size) or PD (Patch Density), were not employed. It should be noted that since the simulated area remained constant in the different simulations (only the spatial location of the growth varied), the NP metric was sufficient to assess fragmentation, without the need for AREA_MN. The metrics cited above were calculated and applied using the FRAGSTATS software package (for a detailed description of the metrics see McGarigal et al., 2002).

The three selected metrics were calculated for the results of the different simulations, enabling a comparison of the percentages of degree of similarity between each simulation and the reference simulation (S8), using the three spatial characteristics as measured by the metrics. It should be noted that in this case, the results obtained were calculated for the entire urban area, in other words, for the area existing at baseline together with the simulated area, since we considered that the growth patterns generated by the simulation should be analysed in conjunction with pre-existing areas rather than in isolation. Table 2 provides a description of the three metrics used.

Spatial Metric	Nomenclature	Characteristic determined	Description
Number of Patches	NP	Fragmentation	Number of fragments of each land use. Higher results imply greater land use fragmentation. Lower results imply land use aggregation.
Mean Shape Index	SHAPE_MN	Shape	Mean relation between the perimeter of a land use patch and the perimeter that the simplest patch (circle) would have with the same area. As the index value raise, the patches will be more stretched.
Mean Euclidean distance	ENN_MN	Dispersion	Mean Euclidean distance to the closest patch with the same land use. The higher the value, the higher the dispersion of the patches.

Table 2: Description of the spatial metrics used.

2.3. Main and Total Effects.

The results deriving from the buffers and the spatial metrics provide *per se* a great deal of information about how each factor may drive the model towards outputs of one kind or another. But we are now in a position to exploit the outputs further with a view to determining quantitatively each factor's contribution through variance decomposition. On the basis of Sobol's (2001) methodology, which is used in global variance based SA (Saltelli et al., 2000), the main and total effects can be calculated for each factor and method. The main effects show what proportion of variance in the results may be accounted for by each of the model's factors, while total effects show what proportion may be accounted for by a single factor and its interactions with the rest of the factors.

To be more precise, the model's output Y, in this case for a spatial metrics ((buffer or spatial metrics), taken as the percentage of agreement between two simulations may be expressed as a function of the parameters implied in the model:

$$Y = f(X1, \dots, Xj, \dots, Xn)$$

Eq [3]: Output Y for a model, expressed as a function of n factors.

Variance of Y may be decomposed as shown in equation 4, where V_i represents the component of variance depending on a single factor, V_{ij} the component of variance due to the interaction of two factors, etc.

$$V(Y) = \sum_{i=1}^{n} V_i + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} V_{ij} + \cdots + V_{ij\ldots n}$$

Eq [4]: Decomposition of variance V for output Y according to additive (V_i's) and interaction terms.

The *main* and *total effects* may then be calculated using the values for partial variances. Thus, we may define η_i[2](Eq[5]) as the first order or main effect sensitivity index for factor i, or, to put it another way, the proportion of variance in the model's output Y due to factor i in itself. Similarly, η_{ij}, η_{ijk} and η_{ijkl} will be second, third and fourth order sensitivity indices respectively, determining the proportion of variance due to interactions among factors.

$$\sum_{i=1}^{n} \eta_i + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \eta_{ij} + \cdots + \eta_{i,j\ldots n} \;=\; 1\,; \quad where \quad \eta_i = \frac{V_i}{V(Y)}$$

Eq [5]: Sensitivity index calculation for a model with n factors.

Finally, each factor's total contribution may be determined by calculating the total indices or total effects. Thus, for a given factor i, the total index η_{Ti} is the sum of each of sensitivity indices in which that factor is present:

$$\eta_{Ti=} \eta_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \eta_{ij} + \cdots + \eta_{i,j\ldots n}$$

Eq [6]: Calculation of sensitivity indices for a given factor i in a model of n factors.

These indices allow us to quantify the influence of each of the four factors which compose the model on the location of the simulations' outputs (distance buffers) or on their spatial pattern (spatial metrics).

In this case, the computation of sensitivity indices is strongly simplified by the fact that input factors have discrete distributions with only two values [0,1]. So, considering 4 factors, we can perform a full factorial design of all possible 2^4=16 combinations of all these 4 factors. Then, sensitivity indices are easily computed applying the definitions for first order indices:

[2] Main and total effects are usually named S and S_T in the literature. In order to avoid confusions, we have decided to use η and η_T since S has previously been used in the text to name each model simulation (S1-S16).

$V_i = V(E(Y|X_i))$

and interactions:

$V_{ij} = V(E(Y|X_i, X_j))$

where the conditioning operator is applied for the values 0 and 1 in turn for each individual factor or combinations of factors, i.e. X_i= 0, 1 for main effects; X_i, X_j = [0 0] [0 1] [1 0] [1 1] for second order interactions and so on. This easily provides the full variance decomposition.

2.4. Identification of the most frequent areas simulated by the model.

Frequency maps were used to validate the results of the simulations, which reflect the number of times a site or location appears in the results of a model. All that is necessary is to sum together all the simulations results to obtain the frequency of occurrence as a result of urban development in each cell.

This method has been used by some authors as part of an uncertainty analysis, as a tool to determine the propagation of errors in a model (Canters et al., 2002; Gómez-Delgado and Bosque-Sendra, 2004). Others have applied it in the same context to determine the extent to which areas are vulnerable (Lowry et al., 1995). This method is similar to the *stability map* proposed by Brown et al. (2005) which identifies the variant and invariant regions through a binary map, but by obtaining the frequency of occurrence for each cell.

Meanwhile, the proposed methodology generates many results which are sometimes difficult to interpret, although there are always patterns or areas that are frequently repeated in each model. Thus, regardless of the factors implemented in the model, these are recurring areas each time the model was run and they seem to present the requisite characteristics for development.

As pointed above, this indicates the advisability of identifying the areas that appear most frequently in the results obtained from the various derived models, since they indicate the more robust sites generated by each simulation. Determination of the spatial location of these sites could lead to significant improvements in urban planning, reducing the risk and uncertainty associated with decision-making in this field (Gómez-Delgado and Bosque-Sendra, 2004).

3. RESULTS

3.1. The simulations generated.

The simulation generated by the baseline model, S8, includes all factors and was used as the reference for all comparisons. The following figures show some of the results generated by the various simulations, specifically the most representative ones in terms of results (S1, S2, S3, S4 and S8).

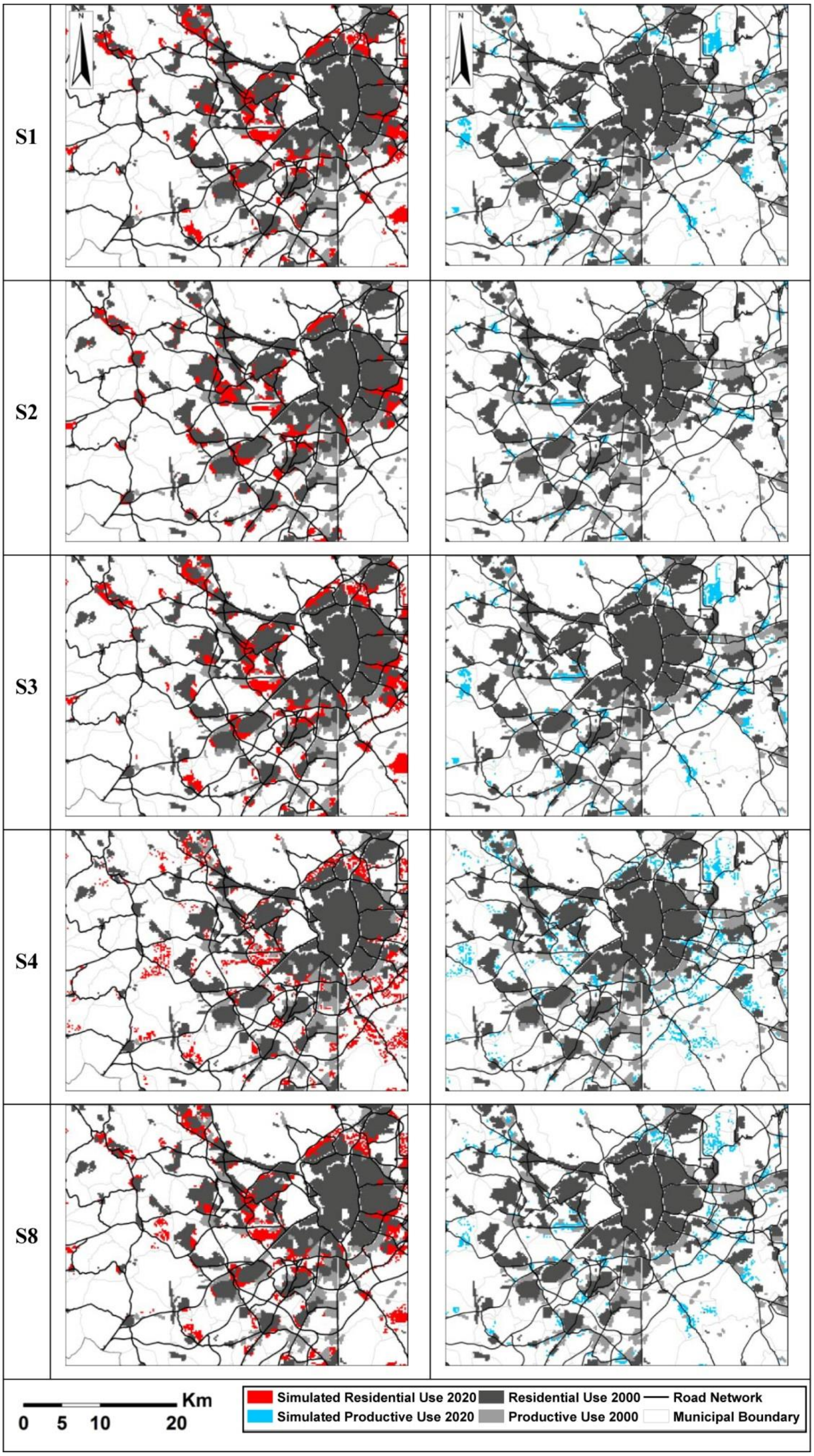
S1
S2
S3
S4
S8
N
Km
0 5 10 20
Simulated Residential Use 2020
Residential Use 2000
Road Network
Simulated Productive Use 2020
Productive Use 2000
Municipal Boundary

Figure 5. Results of the most representative simulations, showing first the results for residential land use and secondly the results for productive land use for S1(N,A,S), S2(N,A), S3(N,S), S4(A,S) and S8(N,A,S,R – baseline model) simulations.

3.2. Comparison of results using distance buffers.

Table 3 shows the results for each of the simulations and simulated urban land uses. Firstly (a), the percentage of cell-to-cell agreement is given for the results of the baseline simulation (S8, including all factors) and the simulation being compared (PA), which is simply the percentage of cells that are spatially coincident in both maps. In the following columns, the cells located within the distance buffers generated at 1000m and 2000m, respectively, are added to this percentage. Secondly (b), the main effects and the total effects of each parameter are given for the three comparison methods.

(a) PA* WITH REFERENCE SIMULATION S8

SIMULATIONS	Residential Land Use			Productive Land Use		
	PA*	PA 1000m	PA 2000m	PA	PA 1000m	PA 2000m
S1 (N,A,S)	81.14	99.89	99.98	74.92	99.65	99.84
S2 (N,A)	37.27	72.36	85.41	40.08	79.09	85.04
S3 (N,S)	80.85	99.76	99.96	74.06	99.53	99.76
S4 (A,S)	40.79	75.46	85.90	20.28	73.74	88.11
S5 (N)	36.97	72.10	85.26	40.08	79.37	85.20
S6 (A)	5.90	40.57	57.62	22.01	55.12	71.30
S7 (S)	41.27	74.76	85.11	21.26	75.31	88.74
S9 (R)	6.92	28.93	48.69	3.07	31.02	56.65
S10 (N,R)	37.53	72.10	84.89	39.21	77.91	84.33
S11 (A,R)	7.40	37.77	55.00	11.89	46.50	65.87
S12 (S,R)	40.96	75.00	85.57	20.31	73.31	87.40
S13 (A,S,R)	41.22	75.55	85.41	21.06	76.61	89.76

S14 (N,S,R)	80.50	98.95	99.52	70.98	98.74	99.29
S15 (N,A,R)	36.92	70.68	83.69	39.72	77.24	83.94
S16 ( )	0.92	10.55	24.72	1.42	19.25	42.68

(b) MAIN (η) AND TOTAL (η_T) FACTOR EFFECTS

EFFECTS	Residential Land Use			Productive Land Use		
	PA	PA 1000m	PA 2000m	PA	PA 1000m	PA 2000m
η_S	0.529	0.504	0.420	0.214	0.388	0.494
η_{TS}	0.553	0.558	0.550	0.317	0.462	0.603
η_N	0.436	0.415	0.399	0.651	0.493	0.337
η_{TN}	0.459	0.471	0.533	0.751	0.568	0.449
η_A	0.003	0.009	0.014	0.018	0.021	0.025
η_{TA}	0.015	0.046	0.081	0.055	0.089	0.114
η_R	0.003	0.001	0.003	0.001	0.000	0.001
η_{TR}	0.015	0.016	0.042	0.030	0.013	0.029
$\sum \eta_i$	0.972	0.929	0.837	0.884	0.902	0.856

*PA=Percentage-of-Agreement

Table 3: First block (a) shows the agreement between each simulation S and reference S8 for cell-to-cell comparison, 1,000m buffer and 2,000m buffer. For each simulation S, the factors included are given in parenthesis. Second block (b) shows the values for the main and total effects of each of the parameters implemented in the model for each method and the sum of all main effects.

The S1 simulation yielded the best cell-to-cell agreement values for residential and productive land use of around 81% and 75%, respectively, whilst an almost complete match was obtained in the results for the first distance buffer of 1,000 meters. Since S1 excludes randomness, it could be said that this factor induced a dispersion of 1 in every 6 cells for residential land use and 1 in every 4 for productive land use, all in the immediately surrounding area (1,000m). The best agreement of S1 is followed by S3 where the absence of the accessibility factor suggests that this parameter has limited influence, barely altering the simulated results. The results for S14 are equivalent to those for S3. In this case, accessibility was the only parameter switched off and the differences between S14 and S1 were very small (less than 1%). These results were corroborated by analyzing the main effects values of both accessibility and stochasticity parameters (η_A and η_R), and the PA values of S6 and S9. Also, their effect in combination with the rest of factors (η_{TA} and η_{TR}) was less than expected. Comparison of PA values from simulations including and excluding these factors led to the same conclusions (i.e. S1 with S3 or S14, S2 with S5 or S15, S10 with S15, S4 with S12 or S13).

Meanwhile, the S15 simulation presented a low PA value for residential and productive land use, with spatial matches remaining below 80% until the second distance buffer. Therefore, this simulation, which did not include suitability, reveals that this factor has a strong influence on the simulated results of both uses. The result is analogous to S2, where randomness is also excluded. The S13 simulation, which excludes neighbourhood, also presented few cell-to-cell matches. This phenomenon is especially pronounced in the case of productive land use, which presents a lower agreement, indicating that the neighbourhood factor has a stronger influence on the location of the new industrial zones simulated. The result is analogous to S4, where randomness is also excluded. Main effects and total effects for suitability and neighbourhood confirm their significant influence on the location of the results in each simulation model. Nevertheless, there are some differences between simulated land uses, as is the case of productive land use, where the effect of neighbourhood is three times greater than that of suitability. This may be explained by the fact that productive land use has a greater attraction effect than residential land use when locating new productive areas. The same conclusions can be drawn if S3 is compared with S5 or S7, or if S14 is compared with S10 or S12.

3.3. Comparison of the results using spatial metrics.

Spatial metrics values applied to the different models are given in figure 6 and the values of main and total effect for each parameter of each spatial metric are given in table 4.

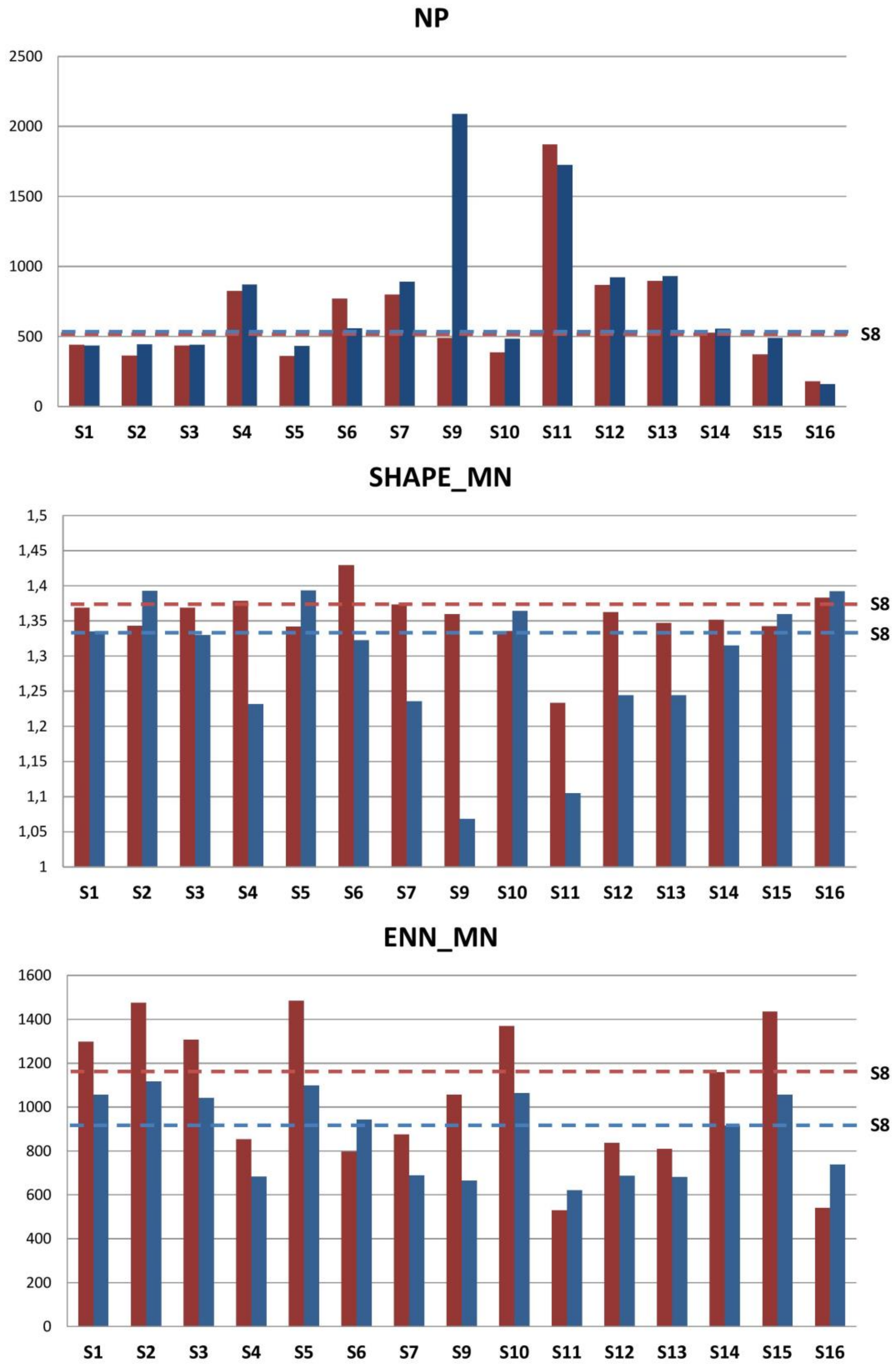

Figure 6: Spatial metrics (NP, SHAPE_MN and ENN_MN) comparison for residential land use (in red) and productive land use (in blue). Values obtained for the reference simulation (S8) are shown as blue lines (productive land use) and red lines (residential land use).

INDICES	Residential Land Use			Productive Land Use		
	NP	SHAPE_MN	ENN_MN	NP	SHAPE_MN	ENN_MN
η_S	0.007	0.062	0.094	0.011	0.007	0.048
η_N	0.293	0.005	0.001	0.306	0.419	0.793
η_A	0.106	0.011	0.071	0.000	0.000	0.004
η_R	0.080	0.213	0.071	0.199	0.157	0.069
η_{NA}	0.112	0.036	0.068	0.000	0.001	0.002
η_{NS}	0.003	0.004	0.042	0.017	0.031	0.003
η_{NR}	0.050	0.150	0.102	0.135	0.085	0.000
η_{AS}	0.099	0.017	0.075	0.000	0.001	0.003
η_{AR}	0.015	0.076	0.058	0.009	0.006	0.010
η_{SR}	0.035	0.074	0.096	0.136	0.162	0.006
η_{NAS}	0.097	0.009	0.070	0.000	0.000	0.004
η_{NAR}	0.019	0.125	0.054	0.008	0.004	0.006
η_{NSR}	0.051	0.092	0.086	0.158	0.116	0.037
η_{ASR}	0.017	0.072	0.058	0.010	0.003	0.009
η_{TS}	0.325	0.383	0.576	0.343	0.327	0.117
η_{TN}	0.641	0.475	0.477	0.635	0.663	0.852
η_{TA}	0.482	0.399	0.508	0.038	0.023	0.043
η_{TR}	0.283	0.855	0.579	0.666	0.540	0.143
$\sum \eta_i$	0.485	0.292	0.236	0.516	0.583	0.913

Table 4: First order η_i, second order η_{ij}, third order η_{ijk} and total η_T sensitivity indices of Neighbourhood (N), Accessibility (A), Suitability (S) and Randomness (R) factors for spatial metrics NP, SHAPE_MN and ENN_MN. Sum of first order indices is given in the last row.

The NP values show the extent to which the land use analysed is fragmented or aggregated. S1, S3 and S14 results are close to those obtained for reference S8. This suggests that the inclusion or exclusion of accessibility does not alter land use fragmentation. On the other hand, S11, S12, S13 and S9 (this last for productive land use only) yield higher values in comparison to S8, which suggests that neighbourhood, included in none of them, has a strong influence on the aggregation pattern. The values for S4, S6 and S7 confirm this statement. Including stochasticity slightly increases fragmentation if pairs of models are compared (i.e. S1 with S8, S3 with S15, S6 with S11). It is also noticeable that comparison of S2, S5, S10 and S15 with S8 suggests that suitability has no significant influence on the aggregation pattern. These results are corroborated by the main and total effects values in table 4, where neighbourhood yields the highest value as having greatest effect on aggregation (NP). Nevertheless, total effects for all factors are at least twice as large as the main effect for the same factor. This means that factor interactions are more influential than each factor on itself in the simulated patterns.

In relation to shape (SHAPE_MN) (Fig. 6b), the values are generally very similar, although S1, S3 and S14 present again the most similarity with S8. This finding indicates that the accessibility factor in S3 and S14, and the randomness excluded in S1 do not substantially change the elongation of the simulated patches. S4, S7, S9 and S11 are the simulations that present the most difference in the case of productive land use, with a considerable reduction in SHAPE_MN. The elimination of neighbourhood in those simulations prevents the development of productive land use along roads and thus less elongation is produced. Furthermore, the higher fragmentation appearing in the results of these simulations, with new, small, almost circular patches also meant that SHAPE_MN values are lower. The S2, S5 and S10 simulations present similar tendencies to the reference S8, with slightly more complex shapes (higher values of SHAPE_MN than S8) in the case of productive land use, which are directly related to the elimination of suitability. As the values for all simulations differed less than those for NP values, the main effects are more evenly distributed. In the case of residential land use, stochasticity accounts for more than 21% of the total variance (0.213 value) for SHAPE_MN index, whereas in the case of productive land use neighbourhood takes the principal role, although stochasticity also yields a high value. The high values of η_{TR} suggest that stochasticity in combination with the other parameters exerts a greater than expected effect on the shape of land use patterns. This can be explained by the fact that the stochastic parameter introduces dispersed patterns into the results: even if there is only one isolated pixel, its shape is close to a circle's and therefore determines the values of SHAPE_MN.

Finally, patch dispersion (ENN_MN) (Figure 6c) differentiates the distribution of simulated urban areas according to their mutual proximity. Again, the spatial distribution for S1, S3 and S14 simulations is most similar to S8 for both uses. ENN_MN values decrease when neighbourhood is excluded from simulations (S4, S6, S12, S13) but increase with the exclusion of suitability (S2, S5, S10, S15). The exclusion of stochasticity increases the values of ENN_MN, which is to be expected as its inclusion generates more patches (NP values higher for S10, S11, S12, S13, S14, S15 than for S5, S6, S7, S4, S3, S2 respectively): more patches imply less distance between them. In this case we found dissimilar values between both land uses in terms of main and total effect. For residential use each factor seems to have more or less the same influence on the spatial distributions of land use (less than 10% difference between them). Neighbourhood has the strongest main effect for productive land use, at no point increasing beyond 6% in the total effect. This means that the spatial distribution of productive land use can practically be accounted for in terms of neighbourhood alone.

Overall, we can affirm that there are more changes to the spatial pattern of productive than residential land use, possibly explained by the different implementation of the neighbourhood factor and the inherent characteristics of productive land use.

3.4. Mapping of the most frequent areas.

Figure 7 shows the locations of new growth generated by the 15 simulations from the model discussed here, based on the number of times these cells appear in the results. In the case of residential land use, a strong trend is observed towards locating new sites in the vicinity of the Madrid metropolitan area, in addition to a tendency to fill in the spaces between various urban centres in close proximity, and this generates the highest values (appearing between 8 and 15 times in the simulations).

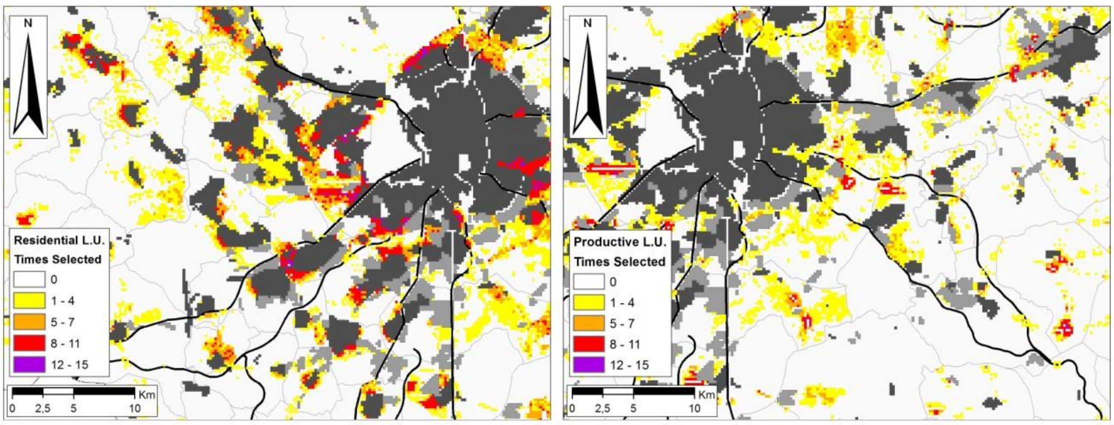

Figure 7: Stability maps for residential and productive land uses showing the number of times that one cell is selected through the different simulations of the model (15; S16 is not included since it is a constant).

On the other hand, productive land use presents a different trend, with less growth for 2020 than residential land use. Few cells appear in all or almost all simulations, and those that do, are mainly located in areas close to other industrial and/or commercial centres already in existence or nearby or along the main transportation routes. Furthermore, as mentioned earlier in the comparison of spatial metrics and more specifically in the case of SHAPE_MN, we found a greater discrepancy in productive land use when compared with the reference simulation. This is reflected in the disparate result shown Figures 7a and 7b.

4. DISCUSSION

4.1. Methods selected to assess spatial differences.

Distance buffer assessment helped us to detect that the random factor may have exerted an excessively strong influence, determining nearly 17% of residential land use cells and 25% of productive ones. Distance buffers also helped us to detect that neighbourhood exerted its greatest influence on the results of productive land use, whilst suitability was the factor that most determined the location of the results as a whole. This was demonstrated in the comparison between the S6 or S11 simulations (in which suitability and neighbourhood were eliminated) and the reference simulation, S8. Further demonstration was provided by computing main and total effects for each factor within the percentage of agreement and distance buffers, when neighbourhood and suitability proved their key role in locating urban growth in the simulations performed here.

The use of spatial metrics provided a more explicit assessment than other methods of the stability of the results in terms of the spatial configuration of the simulated growth. This question is especially important for spatial models used to simulate urban growth given the considerable environmental and land use implications of different urban growth patterns. The use of spatial metrics has also shown that first neighbourhood, then suitability, exert the strongest influence in terms of preserving spatial patterns in the simulations. In addition, we have shown that when randomness is incorporated into the simulations, it leads to the fragmentation (increased number of patches) and decreases the dispersion (proximity between patches) of the results, whereas suitability increases the dispersion of the results for residential land use and neighbourhood generates the degree of aggregation required for new uses to be situated in the vicinity of existing ones. These latter two factors are thus reaffirmed as the two cornerstones of the model. Although the results of main effects for these metrics were lower than expected, we believe that in this case this is due to the iteraction or total effect of the factors for fragmentation, shape and relative distance between patches. In this case, total effect reflects the expected sensitivity behaviour.

Using mapping to determine the cells most repeatedly selected in each simulation (similar to the variant-invariant map proposed by Brown et al., 2005) enabled the most frequent areas to be identified, thus ensuring that land use planning of these areas is more robust and reliable. We believe this map may help planners become aware of the areas with a greater predisposition for transformation into urban areas and of those which may not be appropriate for development.

Lastly, all these methods allowed us to test the operational validity of the model. We can conclude that the current structure, or at least the modelling of some factors, is not suited to the purpose of the scenario simulated. Accessibility had poor influence in the results, so it might be necessary to re-implement this factor into the model. This may be related to the fact that the calculation of suitability includes some spatial distances to concrete elements in the environment. The method proposed here enabled us to identify this problem, which should be recalculated in future research.

4.2. Limitations of the study.

The study presented here is subject to some limitations. A complete SA conducted on CA models should probably deal with all associated sources of uncertainty. In CA-based models, each parameter (according to the NASZ model scheme) and also input data are sources of uncertainty. For example, suitability may depend on several factors (height, slope, soil type, etc); consequently, how it is evaluated and then mapped is subject to some uncertainty. The neighbourhood parameter is also affected by uncertainty, clear sources of which are the size of the neighbourhood window and the calibration of the attraction-repulsion values between land uses (Kocabas and Dragicevic, 2006; Ménard and Marceau, 2005). Other considerations such as cell size (Dietzel and Clarke, 2004; Samat, 2006), future scenario design, etc, are sources of uncertainty too and should therefore be part of a complete SA.

Nevertheless, we believe the proposed method can act as a framework for either assessing the influence of the different factors as presented here or developing a more complete SA, through the in-depth analysis of existing sources of uncertainty (e.g. cell size, calibration values, neighbourhood size).

Another limitation is related to the scenario simulated. In this proposal we have simulated only one scenario, carrying out 16 simulations which combine the different factors of the model (S1 to S16) for it. An analysis of the influence of the factors for several scenarios could help us to present a more complete assessment. To be more precise, mapping the more frequent areas for different scenarios could be a powerful tool in the urban planning process. This map may make it easier to identify the cells that change to urban land use in several scenarios and which therefore require specially consideration in future urban planning.

However, we conducted the assessment for one scenario only in order to focus sharply on the influence of the factors of the NASZ model and to ensure that the results were more readable and clear. We believe that if we make our models and their related validation process easier to understand, planners will be more inclined to include them in the decision-making process.

Although different map comparison methods were employed, other map comparison methods are available, such as those associated with fuzzy set theory (Hagen, 2003), most of which are included in the Map Comparison Kit software package (Visser and de Nijs, 2006). Comparisons performed using these kinds of methods take location and category fuzziness into account. In our case, we could not tackle category fuzziness since we studied each category or land use in isolation.

Furthermore, since we are not evaluating the model as a whole, but rather in parts, in order to understand the influence of each factor on the results, we considered it more appropriate to use comparative methods which facilitate the identification of differences between maps and provide clearly differentiated information.

5. CONCLUSIONS

This paper presents a method of validating future scenarios simulated with a CA-based model. This methodology could serve as a framework for validating future urban growth derived from CA models (or other kind of land use change models). It could also give a first impression of how factors implemented in the model influence results in terms of spatial location. Assessment of the factor's influence was conducted by applying a simple method based on the GSA approach, where it is assumed that each of the factors implemented will present a different influence on the results generated.

The results obtained are relatively stable and enable a clearer identification of each of the characteristics on account of which each factor contributed to the model, although those concerning the accessibility factor present minor differences since this type of model is constructed on the assumption that this factor determines the location of new urban areas. It therefore appears necessary to review the manner in which this factor is modelled and implemented in order to assume its proper function within the model and also to continue exploring factor stability through other map comparison methods.

Hence, the proposed method provides a clear picture of how each individual factor, and the set of factors in conjunction, influences or determines the location, shape, distribution and fragmentation of the patches generated in simulations of future urban growth. Nevertheless, other future scenarios should be explored using this method since the stability of only one future scenario was examined in this study.

ACKNOWLEDGEMENTS

This research was performed within the context of SIMURBAN2 project (Geosimulation and environmental planning on metropolitan spatial decision making. Implementation to intermediate scales) (CSO2012-38158-C02-01), funded by the Spanish Ministry of Economy and Competitiveness. The Universidad de Alcalá supports the first author within the “Ayudas para la Formación de Personal Investigador (FPI) 2012” framework. The authors would also like to thank the anonymous reviewers for their very helpful comments.

REFERENCES

Aguilera, F., Talavera, R., 2009. Valoración de escenarios futuros a través de la conectividad del paisaje; Future scenario assessment through landscape connectivity. Observatorio Medioambiental(12) 17-36.

Aguilera, F., Valenzuela, L.M., Bosque, J., 2010. Simulación de escenarios futuros en la aglomeración urbana de granada a través de modelos basados en autómatas celulares. Boletin De La Asociacion De Geografos Espanoles(54) 271.

Aguilera, F., Valenzuela, L.M., Botequilha-Leitao, A., 2011. Landscape metrics in the analysis of urban land use patterns: A case study in a Spanish metropolitan area. Landsc. and Urban Plan. 99(3-4) 226-238.

Aguilera, F., Botequilha-Leitão, A., Díaz-Varela, E., 2014. Detecting multi-scale urban growth patterns and processes in the Algarve region (Southern Portugal). Appl. Geogr. 53 234-245.

Barredo, J.I., Demichelli, L., Lavalle, C., Kasanko, M., McCormick, N., 2004. Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. Environ. and Plan. B: Plan. & Des. 31(1), 65-84.

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

Batty, M., 2007. Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals. The MIT press.

Batty, M., Xie, Y., 1997. Possible urban automata. Environ. and Plan. B: Plan. & Des. 24, 175-192.

Benenson, I., Torrens, P.M., 2004. Geosimulation : automata-based modelling of urban phenomena. John Wiley & Sons, Hoboken, NJ.

Berdoulay, V., 2009. The history of geography and the challenge of strategic foresight. Boletin De La Asociacion De Geografos Espanoles (51), 9-23.

Berling-Wolff, S., Wu, J., 2004. Modeling urban landscape dynamics: a case study in Phoenix, USA. Urban Ecosyst. 7(3), 215-240.

Börjeson, L., Höjer, M., Dreborg, K.-H., Ekvall, T., Finnveden, G., 2006. Scenario types and techniques: towards a user's guide. Futures 38(7), 723-739.

Brown, D.G., Page, S., Riolo, R., Zellner, M., Rand, W., 2005. Path dependence and the validation of agent-based spatial models of land use. Int. J. of Geogr. Inf. Sci. 19(2), 153-174.

Burriel, E.L., 2011. Subversion of land-use plans and the housing bubble in Spain. Urban Res. & Pract. 4(3), 232-249.

Buyantuyev, A., Wu, J., 2009. Urbanization alters spatiotemporal patterns of ecosystem primary production: A case study of the Phoenix metropolitan region, USA. J. of Arid Environ. 73(4), 512-520.

Buyantuyev, A., Wu, J., 2010. Urban heat islands and landscape heterogeneity: linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. Landsc. Ecol. 25(1), 17-33.

Buyantuyev, A., Wu, J.G., Gries, C., 2010. Multiscale analysis of the urbanization pattern of the Phoenix metropolitan landscape of USA: Time, space and thematic resolution. Landsc. and Urban Plan. 94(3-4), 206-217.

Canters, F., De Genst, W., Dufourmont, H., 2002. Assessing effects of input uncertainty in structural landscape classification. Int. J. of Geogr. Inf. Sci. 16(2), 129-149.

Couclelis, H., 1997. From cellular automata to urban models: new principles for model development and implementation. Environ. and Plan. B: Plan. & Des. 24, 165-174.

Couclelis, H., 2005. " Where has the future gone?" Rethinking the role of integrated land-use models in spatial planning. Environ. and Plan. A: Environ. & Plan. 37(8), 1353-1371.

Chen, Y., Yu, J., Khan, S., 2010. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. Environ. Model. & Softw. 25(12), 1582-1591.

Chen, Y., Yu, J., Khan, S., 2013. The spatial framework for weight sensitivity analysis in AHP-based multi-criteria decision making. Environ. Model. & Softw. 48, 129-140.

Cheng, J., Masser, I., 2003. Urban growth pattern modeling: a case study of Wuhan city, PR China. Lands. and Urban Plan. 62(4), 199-217.

Chu-Agor, M., Muñoz-Carpena, R., Kiker, G., Emanuelsson, A., Linkov, I., 2011. Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. Environ. Model. & Softw. 26(5) 593-604.

Crosetto, M., Moreno Ruiz, J.A., Crippa, B., 2001. Uncertainty propagation in models driven by remotely sensed data. Remote Sens. of Environ. 76(3) 373-385.

Crosetto, M., Tarantola, S., 2001. Uncertainty and sensitivity analysis: tools for GIS-based model implementation. Int. J. of Geogr. Inf. Sci.15(5) 415-437.

Dietzel, C., Clarke, K.C., 2004. Spatial Differences in Multi-Resolution Urban Automata Modeling. Trans. in GIS 8(4), 479-492.

Dietzel, C., Herold, M., Hemphill, J.J., Clarke, K.C., 2005. Spatio-temporal dynamics in California's central valley: Empirical links to urban theory. Int. J. of Geogr. Inf. Sci. 19(2), 175-195.

Eastman, R., 2006. IDRISI Andes version 15. Clark Labs, Clark University, Worcester, USA. URL: http://www.clarklabs.org.

Engelen, G., White, R., 2008. Validating and calibrating integrated cellular automata based models of land use change, The Dynamics of Complex Urban Systems. Springer, pp. 185-211.

European Environment Agency., 2006. Land accounts for Europe 1990-2000 : towards integrated land and ecosystem accounting. European Environment Agency, Copenhagen.

Forman, R.T., 1995. Land mosaics: the ecology of landscapes and regions. Cambridge university press.

Frankhauser, P., 1998. Fractal geometry of urban patterns and their morphogenesis. Discret. Dyn. in Nat. and Soc. 2(2), 127-145.

Garcia, A.M., Sante, I., Crecente, R., Miranda, D., 2011. An analysis of the effect of the stochastic component of urban cellular automata models. Comput. Environ. and Urban Syst. 35(4), 289-296.

Gómez, M., Bosque, J., 2004. Sensitivity analysis in multicriteria spatial decision-making: A review. Hum. and Ecol. Risk Assess. 10(6), 1173-1187.

Gómez-Delgado, M., Rodríguez-Espinosa, V.M., 2012. Análisis de la dinámica urbana y simulación de escenarios de desarrollo futuro con tecnologías de la información geográfica. RA-MA EDITORIAL, p. 352.

Gómez-Delgado, M., Tarantola, S., 2006. GLOBAL sensitivity analysis, GIS and multi-criteria evaluation for a sustainable planning of a hazardous waste disposal site in Spain. Int. J. of Geogr. Inf. Sci. 20(4) 449-466.

Goodchild, M.F., Hunter, G.J., 1997. A simple positional accuracy measure for linear features. Int. J. of Geogr. Inf. Sci. Science 11(3), 299-306.

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., Briggs, J.M., 2008. Global change and the ecology of cities. Sci. 319 (5864), 756-760.

Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps. Int. J. of Geogr. Inf. Sci. 17(3), 235-249.

Hagen, A., 2006. Map comparison methods that simultaneously address overlap and structure. J. of Geogr.l Syst. 8(2), 165-185.

Hansen, H.S., 2010. Modelling the future coastal zone urban development as implied by the IPCC SRES and assessing the impact from sea level rise. Landsc. and Urban Plan. 98(3), 141-149.

Hepinstall-Cymerman, J., Coe, S., Hutyra, L.R., 2013. Urban growth patterns and growth management boundaries in the Central Puget Sound, Washington, 1986-2007. Urban Ecosyst. 16(1), 109-129.

Herold, M., Couclelis, H., Clarke, K.C., 2005. The role of spatial metrics in the analysis and modeling of urban land use change. Comput., Environ. and Urban Syst. 29(4), 369-399.

Herold, M., Goldstein, N.C., Clarke, K.C., 2003. The spatiotemporal form of urban growth: measurement, analysis and modelling. Remote Sens. of Environ. 86(3), 286-302.

Hewitt, R., Van Delden, H., Escobar, F., 2014. Participatory land use modelling, pathways to an integrated approach. Environ. Model. & Soft. 52, 149-165.

Itami, R.M., 1994. Simulating spatial dynamics: cellular automata theory. Landsc. and Urban Plan. 30, 27-47.

Jantz, C.A., Goetz, S.J., 2005. Analysis of scale dependencies in an urban land-use-change model. Int. J. of Geogr. Inf. Sci. 19(2), 217-241.

Kasanko, M., Barredo Cano, J.I., Lavalle, C., McCormick, N., Demichelli, L., Sagris, V., Brezger, A., 2006. Are European cities becoming dispersed? A comparative analysis of 15 european urban areas. Landsc. and Urban Plan. 77(1-2), 111-130.

Kaza, N., 2013. The changing urban landscape of the continental United States. Landsc. and Urban Plan. 110, 74-86.

Kocabas, V., Dragicevic, S., 2006. Assessing cellular automata model behaviour using a sensitivity analysis approach. Comput., Environ.t and Urban Syst. 30(6), 921-953.

Li, X., Yang, Q.S., Liu, X.P., 2008. Discovering and evaluating urban signatures for simulating compact development using cellular automata. Landsc. and Urban Plan. 86(2), 177-186.

Ligmann-Zielinska, A., Jankowski, P., 2014. Spatially-explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation. Environ. Model. & Soft. 57 235-247.

Ligmann-Zielinska, A., Sun, L., 2010. Applying time-dependent variance-based global sensitivity analysis to represent the dynamics of an agent-based model of land use change. Int. J. of Geogr. Inf. Sci. 24(12) 1829-1850

Lilburne, L., Tarantola, S., 2009. Sensitivity analysis of spatial models. Int. J. of Geogr. Inf. Sci. 23(2), 151-168.

Liu, X.H., Andersson, C., 2004. Assessing the impact of temporal dynamics on land-use change modeling. Comput., Environ. and Urban Syst. 28(1-2), 107-124.

Lowry, J.H., Miller, H.J., Hepner, G.F., 1995. A Gis-Based Sensitivity Analysis of Community Vulnerability to Hazardous Contaminants on the Mexico/Us Border. Photogramm. Engineer. and Remote Sens. 61(11), 1347-1359.

Marrel, A., Iooss, B., Jullien, M., Laurent, B., Volkova, E., 2011. Global sensitivity analysis for models with spatially dependent outputs. Environmetrics 22(3) 383-397.

Mas, J.F., Perez-Vega, A., Clarke, K.C., 2012. Assessing simulated land use/cover maps using similarity and fragmentation indices. Ecol. Complex. 11, 38-45.

McGarigal, K., Cushman, S., Neel, M., Ene, E., 2002. FRAGSTATS: spatial pattern analysis program for categorical maps. Computer software program produced by the authors at the University of Massachusetts, Amherst.

Ménard, A., Marceau, D.J., 2005. Exploration of spatial scale sensitivity in geographic cellular automata. Environ. and Plan. B: Plan. and Des. 32(5), 693-714.

Mitsova, D., Shuster, W., Wang, X.H., 2011. A cellular automata model of land cover change to integrate urban growth with open space conservation. Landsc. and Urban Plan. 99(2), 141-153.

OSE, O.d.l.S.e.E., 2006. Cambios de ocupación del suelo en España. Implicaciones para la sostenibilidad. Estudio realizado a partir del proyecto CORINE LAND COVER. Mundi-Prensa Libros, Madrid.

Paegelow, M., Olmedo, M.T.C., 2008. Modelling Environmental Dynamics. Springer.

Pan, Y., Roth, A., Yu, Z., Doluschitz, R., 2010. The impact of variation in scale on the behavior of a cellular automata used for land use change modeling. Comput., Environ. and Urban Syst. 34(5), 400-408.

Petrov, L.O., Lavalle, C., Kasanko, M., 2009. Urban land use scenarios for a tourist region in Europe: Applying the MOLAND model to Algarve, Portugal. Landsc. and Urban Plan. 92(1), 10-23.

Plata-Rocha, W., Gomez-Delgado, M., Bosque-Sendra, J., 2011. Simulating urban growth scenarios using GIS and multicriteria analysis techniques: a case study of the Madrid region, Spain. Environ. and Plan. B: Plan. & Des. 38(6), 1012-1031.

Plata, W., Gómez, M., Bosque, J., 2010. Desarrollo de modelos de crecimiento urbano óptimo para la Comunidad de Madrid aplicando Métodos de Evaluación Multicriterio y Sistemas de Información Geográfica. GeoFocus 10, 103-134.

Plata-Rocha, W., Gómez-Delgado, M., Bosque-Sendra, J., 2012. Proposal for the Introduction of the Spatial Perspective in the Application of Global Sensitivity Analysis. J. of Geogr. Inf. Syst. 4, 503-513.

Pontius Jr, R.G., 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogramm. Engineer. and Remote Sens. 68(10), 1041-1050.

Pontius Jr, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. of Remote Sens. 32(15), 4407-4429.

Power, C., Simms, A., White, R., 2001. Hierarchical fuzzy pattern matching for the regional comparison of land use maps. Int. J. of Geogr. Inf. Sci. 15(1), 77-100.

Saltelli, A., Chan, K., Scott, E.M., 2000. Sensitivity analysis. Wiley New York.

Samat, N., 2006. Characterizing the scale sensitivity of the cellular automata simulated urban growth: a case study of Seberan Perai Region, Penang State, Malaysia. Comput., Environ. and Urban Syst. 30(6), 905-920.

Sante, I., Garcia, A.M., Miranda, D., Crecente, R., 2010. Cellular automata models for the simulation of real-world urban processes: A review and analysis. Landsc. and Urban Plan. 96(2), 108-122.

Schwarz, N., 2010. Urban form revisited-Selecting indicators for characterising European cities. Landsc. and Urban Plan. 96(1), 29-47.

Seto, K.C., Fragkias, M., 2005. Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. Landsc. Ecology 20(7), 871-888.

Sobol', I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math. and computers in Simul. 55(1-3) 271-280.

Store, R., Kangas, J., 2001. Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modelling. Landsc. and Urban Plan. 55(2) 79-93.

Torrens, P., 2000. How cellular models of urban systems work, CASA Working Serie Papers. Centre for Advanced Spatial Analysis (University College London).

Triantakonstantis, D., Mountrakis, G., 2012. Urban Growth Prediction: A Review of Computational Models and Human Perceptions. J. of Geogr. Inf. Syst. 4(6), 555-587.

Tveite, H., Langaas, S., 1999. An accuracy assessment method for geographical line data sets based on buffering. Int. J. of Geogr. Inf. Sci. 13(1), 27-47.

van Vliet, J., Bregt, A.K., Hagen-Zanker, A., 2011. Revisiting Kappa to account for change in the accuracy assessment of land-use change models. Ecol. Model. 222(8), 1367-1375.

Veregin, H., 1999. Data quality parameters, In: Longley, P.A., Goodchild, M.F., Mcaguire, D.J., Rhind, D.W. (Eds.), Geogr. Inf. Syst., pp. 177-189.

Visser, H., de Nijs, T., 2006. The Map Comparison Kit. Environ. Model. & Soft. 21(3), 346-358.

White, R., Engelen, G., 1993. Cellular-Automata and Fractal Urban Form - a Cellular Modeling Approach to the Evolution of Urban Land-Use Patterns. Environ. and Plan. A 25(8), 1175-1199.

White, R., Engelen, G., 1997. Cellular automata as the basis of integrated dynamic regional modelling. Environ. and Plan. B-Plan. & Des. 24 (2), 235-246.

White, R., Engelen, G., Uljee, I., 1997. The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environ. and Plan. B: Plan. & Des. 24(3), 323-343.

Wu, J., Jenerette, G.D., Buyantuyev, A., Redman, C.L., 2011. Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States. Ecol.l Complex. 8(1), 1-8.

Yeh, A.G.-O., Li, X., 2006. Errors and uncertainties in urban cellular automata. Computers, Environ. and Urban Syst. 30, 10-28.