Spatial Dynamic Modelling of Deforestation in the Amazon

Arimatéa C. Ximenes, Cláudia M. Almeida, Silvana Amaral,
Maria Isabel S. Escada and Ana Paula D. Aguiar
National Institute for Space Research,
General Coordination for Earth Observation,
Av. dos Astronautas, 1758, PO Box 515, São José dos Campos, SP
Brazil

1. Introduction

New GIS technologies have been employed to support public policies and actions towards environmental conservation, aiming to preserve biodiversity and mitigate the undesirable side-effects of human activities. The spatio-temporal simulation of systems dynamics is an example of such new technologies and helps scientists and decision-makers to understand the driving forces lying behind processes of change in environmental systems. In assessing how systems evolve, it is possible to figure out different scenarios, given by diverse socio-economical, political and environmental conditions (Soares-Filho et al., 2001), and hence, anticipate the occurrence of certain events, like land cover and land use change, including deforestation. According to Openshaw (2000), computer simulation models provide qualitative and quantitative information on complex natural phenomena. In this sense, spatial dynamic models may be defined as mathematical representations of real-world processes or phenomena, in which the state of a given place on the Earth surface changes in response to changes in its driving forces (Burrough, 1998).

Spatial dynamic models are commonly founded on the paradigm of cellular automata (CA). Wolfram (1983) defines CA as "[…] mathematical idealisations of physical systems in which space and time are discrete, and physical quantities take on a finite set of discrete values. A cellular automaton consists of a regular uniform lattice (or 'array'), usually infinite in extent, with a discrete variable at each site ('cell'). […] A cellular automaton evolves in discrete time steps, with the value of the variable at the site being affected by the values of variables at sites in its 'neighbourhood' on the previous time step. The neighbourhood of a site is typically taken to be the site itself and all immediately adjacent sites. The variables at each site are updated simultaneously ('synchronously'), based on the values of the variables in their neighbourhood at the preceding time step, and according to a definite set of 'local rules'." (Wolfram, 1983, p. 603).

This work applies a CA model – Dinamica EGO – to simulate deforestation processes in a region called *São Félix do Xingu*, located in east-central Amazon. EGO consists in an environment that embodies neighbourhood-based transition algorithms and spatial feedback approaches in a stochastic multi-step simulation framework. Biophysical variables

drove the simulation model of the present work, and statistical validation tests were then conducted for the generated simulations (from 1997 to 2000), by means of multiple resolution fitting methods. This modelling experiment demonstrated the suitability of the adopted model to simulate processes of forest conversion, unravelling the relationships between site attributes and deforestation in the area under analysis.

2. Study area

The region of *São Félix do Xingu* is regarded as one of the current three main occupation fronts in the Brazilian Amazon. Recent official data indicate that *São Félix do Xingu* was the Amazonian municipality owning the highest deforestation rates at the end of last decade and the beginning of this decade. According to Becker (2005), *São Félix do Xingu* and other occupation fronts represent the new inland Amazonian frontiers, namely mobile frontiers, which differ from the frontiers observed in the early stages of human occupation in this region in the 1970s with regard to three aspects: i) the prevailing migration is intra-regional, and mostly rural-urban; ii) the private capital plays a crucial role in such fronts, which present a great diversity of local actors, mainly loggers, cattle raisers, and grains producers, and iii) these fronts own greater accessibility and connectivity, relying on a denser (air, terrestrial, and fluvial) transportation network as compared to the one available in the 1970s. Recent data on deforestation show that *São Félix do Xingu* was also the Amazonian municipality presenting the greatest absolute values of deforested areas between the years 2000 and 2006. Out of the total deforested area assessed in the Brazilian Legal Amazon in the years 2005 (665,854 km^2) and 2006 (679,899km^2), *São Félix do Xingu* alone accounted for 13,626 km^2 (2.0%) and 14,496 km^2 (2.1%), respectively (Brazilian National Institute for Space Research or *Instituto Nacional de Pesquisas Espaciais* [INPE], 2006). Part of the occupation history in this region can be ascribed to natural rubber and mahogany exploitation, mining, cattle raising, and huge private and public rural settlements projects, among which cattle raising is the chief economic activity (Escada et al., 2007).

The study area comprises most of the *São Félix do Xingu* municipality and its surroundings, located in the State of *Pará* (PA), east-central Amazon, north of Brazil. The *Xingu* river, one of the major tributaries of the Amazon river, crosses *Pará* in the longitudinal direction (Fig. 1). The data used in this research are set in the Polyconic Projection System, Datum SAD-69, between longitude coordinates 52° 30' and 51° 00'W, and latitude coordinates 05° 52' and 07° 07'S. Besides the municipality of *São Félix do Xingu*, the study area also comprises the municipal seats of *Ourilândia do Norte* and *Tucumã*, and part of the municipalities of *Marabá*, *Parauapebas*, and *Água Azul do Norte*. Within *São Félix do Xingu*, the study area contains the following villages and districts: *Vila Taboca*, *Nereu*, *Tancredo Neves*, *Carapanã*, *Minerasul*, and *Ladeira Vermelha*. Part of the Indian reserves *Apyterewa*, *Kayapó*, and *Xinkrin do Cateté* are also included in the study area.

The portion of *São Félix do Xingu* embraced by the study area presents already a consolidated occupation. In the 1980s, this area sheltered pioneer fronts, marked by intense land occupation processes concentrated along the PA-279 road, which connects *Tucumã* to *São Félix do Xingu* (Shimink & Wood, 1992; Research Network for Environmental Modelling in the Amazon [GEOMA], 2004). This region is currently characterised by the presence of large farms, landed property concentration, and very often, illegal practices of land ownership (Escada et al., 2005).
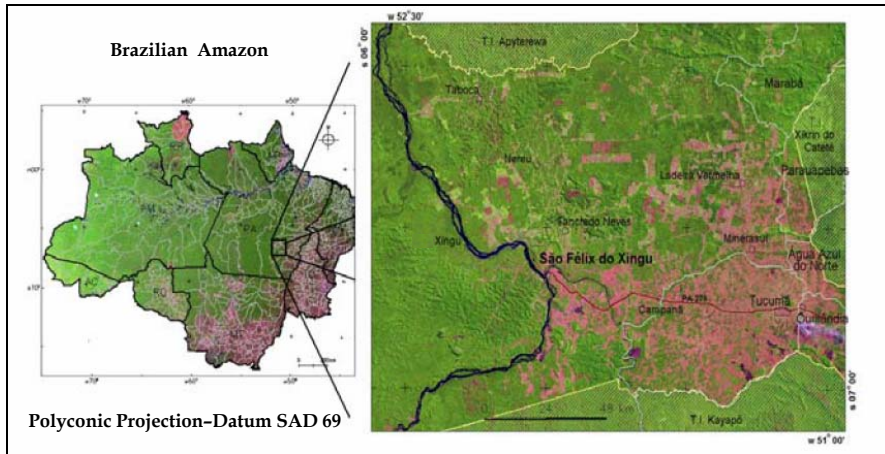
Fig. 1. Location of study area: *São Félix do Xingu* municipality and surroundings

Although this region has cattle raising as the leading economic activity, ore exploitation (mainly cassiterite and gold) has had an important role in its economy and regional spatial structure since the 1970s. Initially, ore was transported along the *Xingu* river and its tributaries, and also through small aircrafts (Santana, 2007). From the first half of the 1990s onwards, mining activities entered into decline, leaving behind them a dense network of roads connecting farms in the region, which considerably reduced the importance of the air and fluvial transport (Amaral et al., 2006; Escada et al., 2007).

3. Input data for the spatial dynamic model of deforestation

A platform for the spatio-temporal modelling of landscape dynamics – Dinamica EGO – was used to generate the simulations of deforestation in the analysed study area. Dinamica EGO consists in a cellular automata environment that embodies transition algorithms operating commonly through a Moore neighbourhood (3 x 3 window) as well as spatial feedback approaches in a stochastic multi-step simulation framework. Dinamica EGO is a free domain platform and was developed by the Centre for Remote Sensing of the Federal University of Minas Gerais – CSR-UFMG[1] (Soares-Filho et al., 2002, 2009; Rodrigues et al., 2007).

Real data on deforestation from 1997 to 2000 drove the simulation model, helping to assess the total amount of forest conversion into other land cover classes in the study area for the given period of analysis. A set of explaining variables related to deforestation together with internal parameters of Dinamica EGO were jointly combined to generate a simulation for the year 2000. In the following sections, methodological procedures employed at each stage of the modelling process (data acquisition, variables selection, exploratory analysis, calculation of transition probabilities, parameterisation, and accuracy assessment) will be dealt with in a thorough manner.

---

[1]The CA modelling platform Dinamica EGO is available for download at
http://www.csr.ufmg.br/dinamica/

3.1 Deforestation data

The original deforestation map, containing the land cover classes forest, grasslands, rivers, deforested areas until 1997, and deforested areas from 1997 to 2000, was acquired from the digital PRODES (Brazilian Deforestation Assessment Project) database (INPE, 2006) and is shown in Fig. 2. This map and other input data were pre-processed in the software IDRISI Kilimanjaro (Eastman, 2003).
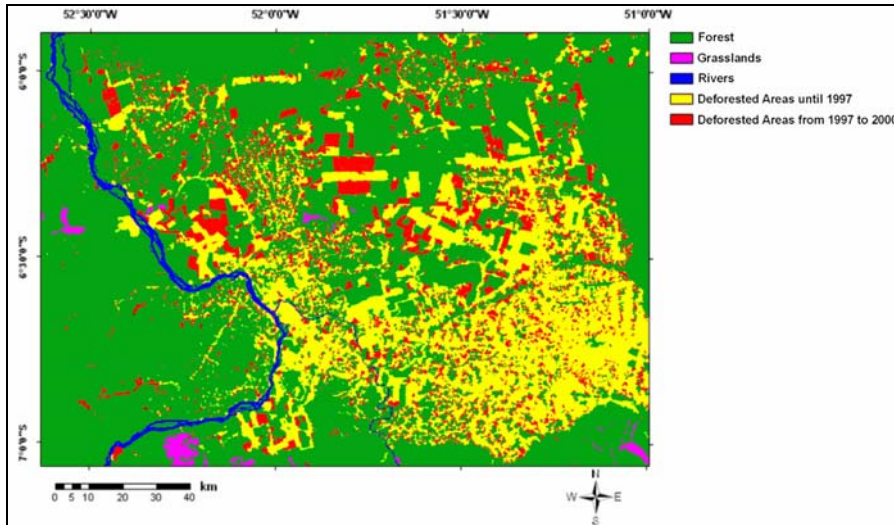

Fig. 2. PRODES deforestation map for the study area from 1997 to 2000

This PRODES deforestation map was reclassified, so as to generate the land cover maps for the initial time of simulation – year 1997 (Fig. 3), and for the final time of simulation – year 2000 (Fig. 4). Considering that PRODES methodology does not take into account deforestation over natural grasslands, this work similarly restricted itself to simulate deforestation only over forested areas. Due to generalisation procedures adopted in the PRODES maps, only the *Xingu* river and its major tributary, *Fresco* river, are visible, since minor streams were disregarded in face of the spatial resolution adopted in the maps (120 x 120 m).

3.2 Explaining variables

The proper choice of a set of explaining variables is critical for the success of a spatial dynamic model. The forested cells suitability to undergo deforestation precisely depends on the relations between such variables and the response variable (deforestation). In this experiment, the analysed variables were selected based upon similar previous studies (Alves, 2002; Laurance et al., 2002; Aguiar et al., 2007; Soares-Filho et al., 2006; Pereira et al., 2007; Brandão Júnior et al., 2007), which report the prevailing variables in deforestation processes. The factors which influence deforestation are manifold. However, the difficulties in data acquisition limit the input data actually employed in the modelling process. Six biophysical variables were evaluated: i) distance to paved roads; ii) distance to non-paved roads; iii) distance to urban centres; iv) distance to rivers; and v) distance to previously deforested areas in 1997.
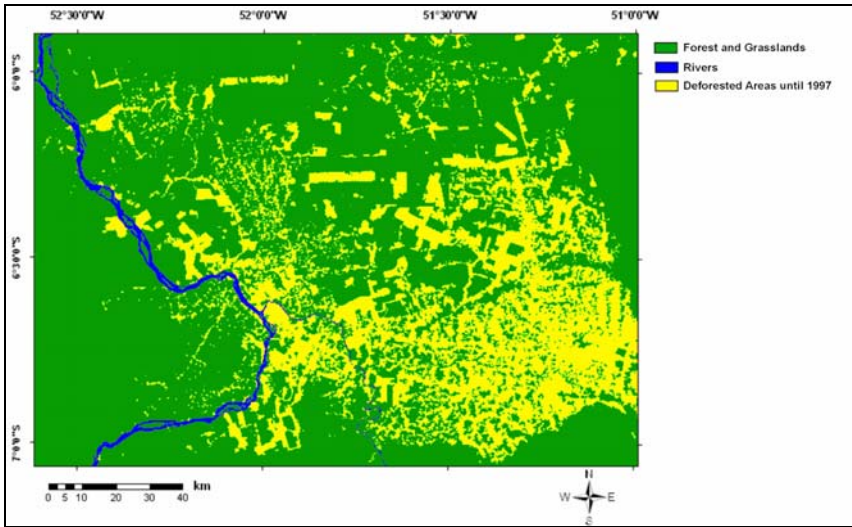
Fig. 3. Land cover map in 1997, derived from the PRODES deforestation map reclassification
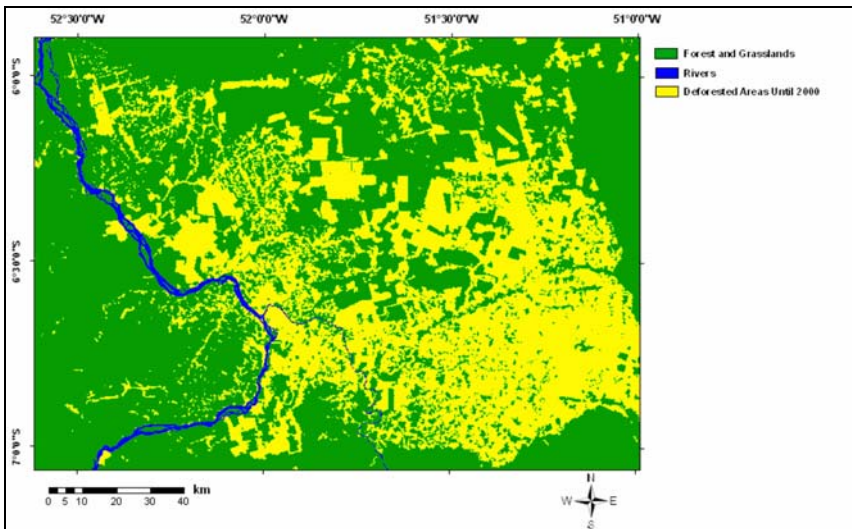

Fig. 4. Land cover map in 2000, derived from the PRODES deforestation map reclassification

By means of a visual analysis of the original deforestation map in 1997, it was possible to observe that both the distance to roads and to urban centres have a limited influence on deforestation processes. In 1997, the occupation in the study area was already well-established, reducing the availability of land for deforestation close to roads and urban settlements. Although the fluvial transport plays a minor role in the region, as previously exposed, it was observed the mushrooming of deforestation patches in the vicinities of the *Xingu* and *Fresco* rivers during the analysed period, which although occurred to a reduced

extent in farther areas, justified the inclusion of the variable 'distance to rivers' in the model. In CA models, two types of variables can be used: i) the static ones, which are kept constant throughout the model run, and ii) the dynamic ones, that suffer changes throughout the successive time steps, which are then continuously updated at each iteration. Both of them were built upon basis of the PRODES reclassified map in the year 1997 (Fig. 3). The static variable corresponds to the map of distance to rivers (Fig. 5), and the map of distance to previously deforested areas in 1997 (Fig. 6) entered the model as a dynamic variable. In order to categorise the grids of distances, i.e. generate optimal discrete ranges of distances, special automatic routines available in the Dinamica EGO were used, which are based on algorithms of lines generalisation (Agterberg & Bonham-Carter., 1990; Goodacre et al., 1993).
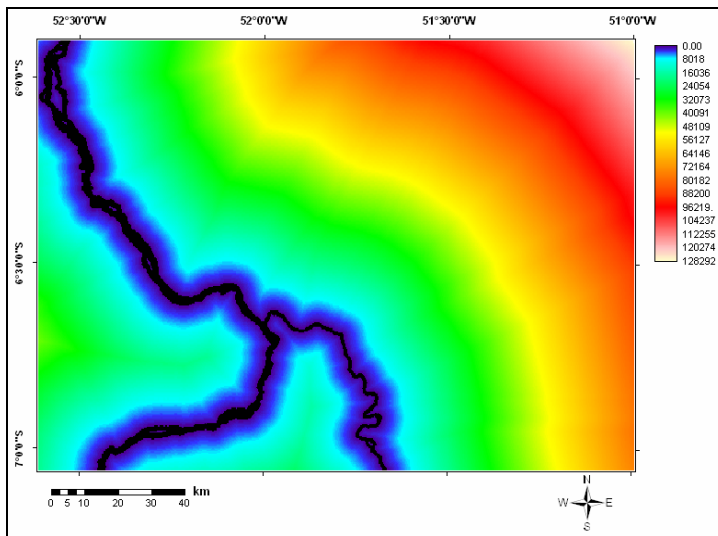

Fig. 5. Colour-sliced map of distance to rivers (static variable), defined in meters

4. The spatial dynamic modelling experiment

4.1 Exploratory analysis

The statistical method 'weights of evidence' was employed for the parameterisation of this modelling experiment. Such method is entirely based on Bayes´s theorem, also known as the conditional probability theorem, which assumes the independence of events. In this sense, the eventual existence of spatial dependence (or spatial association) between pairs of explaining variables must be initially verified. For this end, two statistical indices were used: the Cramer´s Coefficient (V) and the Joint Information Uncertainty (U) (Bonham-Carter, 1994). Both of them are based on the ratio of overlapping areas among the different categories (in this case, ranges of distance) belonging to two maps of explaining variables (or evidences). The Cramer´s Coefficient operates with absolute values of area, while the Joint Information Uncertainty deals with relative (percentage) values, and hence, tends to be more robust than the former index, for it avoids the risk of bias represented by absolute area values.

Bonham-Carter (1994) reports that values less than 0.5 either for V or U suggest less association rather than more. In this way, the threshold of 0.5 was adopted to decide whether a variable should remain in the model (V or U < 0.5) or be excluded from it (V or U ≥ 0.5). In this work, the Cramer's Coefficient and the Join Information Uncertainty presented low values (V = 0.21 and U = 0.0485), indicating that both variables could be simultaneously employed in the model.

4.2 Global transition rates

The global transition rates refer to the total amount of change in the simulation period, regardless of its spatial distribution, i.e. without taking into account spatial peculiarities at the local level, which are those related to biophysical and infrastructural characteristics of each cell in the study area.

In this modelling experiment and in other experiments where the initial and final land cover maps are available, the transition rates were calculated by a cross-tabulation operation between the land cover maps from 1997 and 2000, producing as output a transition matrix with land cover change rates observed during this period.
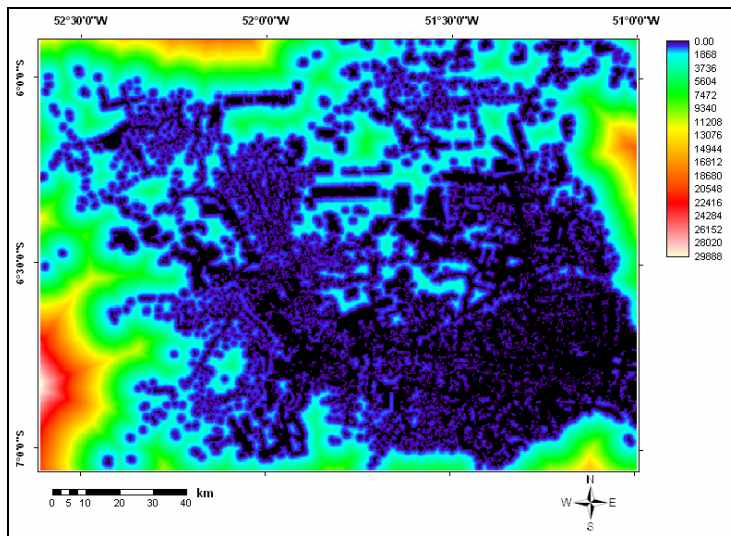

Fig. 6. Colour-sliced map of distance to previously deforested areas in 1997 (dynamic variable), defined in meters

4.3 Local transition probabilities

The local transition probabilities, different from the global transition rates, are calculated for each cell considering the natural and anthropic characteristics of the site. For estimating the land cover transition probabilities in each cell, represented by its coordinates x and y, an equation converting the *logit* formula into a conventional conditional probability was used. The *logit* corresponds to the natural logarithm of odds, which consists in the ratio of the probability of occurring a given land cover transition to its complementary probability, i.e.

the probability of not occurring the transition. This concept derives from the Bayesian weights of evidence method, from which the land cover transition probability can be obtained through algebraic manipulations of the *logit* formula, as follows (Bonham-Carter, 1994):

$$P(T_i^\alpha / V_i^1,...,V_i^{m_\alpha}) = O(T_i^\alpha) \cdot e^{\sum_{v=1}^{m_q} W_{1,v}^+} / 1 + \sum_{\alpha=1}^{\eta} O(T_i^\alpha) \cdot e^{\sum_{v=1}^{m_q} W_{1,v}^+} \qquad (1)$$

where P corresponds to the probability of transition in a cell; i corresponds to a notation of cells positioning in the study area, defined in terms of x,y coordinates; α represents a type of land cover transition, e.g. from a class c to a class k, within a total of η transitions (in this particular experiment, there is only one type: from forest to deforested areas); V_i^1 corresponds to the first variable observed in cell i, used to explain transition α; $V_i^{m_\alpha}$ corresponds to the m-th variable observed in cell i, used to explain transition α; O (T_i^α) represents the odds of transition T^α in the i-th cell, expressed by the ratio of the probability of occurrence of T_i^α over its complementary probability, i.e., P (T_i^α)/ P ($\overline{T}_i^\alpha$); and $W_{i,v}^+$ corresponds to the positive weight of evidence for the i-th cell regarding the v-th variable range, defined as:

$$W_{i,v}^+ = \log_e P(V_i^{m_\alpha} / T_i^\alpha) / P(V_i^{m_\alpha} / \overline{T}_i^\alpha) \qquad (2)$$

where P ($V_i^{m_\alpha}/ T_i^\alpha$) is the probability of occurrence of the m-th variable range observed in cell i, used to explain transition α, in face of the previous presence of transition T_i^α, given by the number of cells where both $V_i^{m_\alpha}$ and T_i^α are found divided by the total number of cells where T_i^α is found; and P($V_i^{m_\alpha}/ \overline{T}_i^\alpha$) is the probability of occurrence of the m-th variable range observed in cell i, used to explain transition α, in face of the previous absence of transition T_i^α, given by the number of cells where both $V_i^{m_\alpha}$ and $\overline{T}_i^\alpha$ are found, divided by the total number of cells where T_i^α is not found.

The W^+ values represent the attraction between a determined landscape transition (in this case, from forest to deforested areas) and a certain variable range. The higher the W^+ value is, the greater is the probability of a certain transition to take place. On the other hand, negative W^+ values indicate lower probability of a determined transition in the presence of the respective variable range. Using the W^+ values concerning the several distances ranges of the static and dynamic variables employed in the analysis of deforestation, the Dinamica EGO model calculates the cells transition probabilities according to Equation (1). The grid cells are assigned a value of probability and a probability map is then generated. In order to evaluate if the model is well calibrated, i.e. if the employed explaining variables are appropriate and if the categorisation of the numerical grids is optimal, this map must present the area with the highest transition probability values as close as possible to the areas that actually underwent deforestation processes.

4.4 Defining the dinamica EGO internal parameters

Dinamica EGO incorporates two empirical land cover allocation algorithms (or transition functions) called: expander and patcher. The expander function accounts for the expansion of previous patches of a certain land cover class. The patcher function, on its turn, is designed to generate new independent patches (seed cells), which are not physically

connected with previous patches of the same land cover class (Soares-Filho et al., 2002). In summary, the expander function performs transitions from a state i to a state j only in the adjacent vicinities of cells with state j. And the patcher function performs transitions from a state i to a state j only in the adjacent vicinities of cells with state other than j.

Based on a visual analysis of the PRODES deforestation map (Fig. 2) during the analysed period (1997-2000), it was not observed the occurrence of deforestation by means of diffusion processes, i.e. amidst the virgin forest, what is emulated by the patcher function. In this way, only the expander function was employed in the model, which simulates the formation of deforestation patches through the expansion of previously deforested areas, as exposed above.

The average size and variance of size (in hectares) of the new deforestation patches are also required as internal parameters by Dinamica EGO. The definition of these parameters is done heuristically. The ideal average size (μ) was set to 300 ha, and the variance (σ^2) to 500 ha. The model contains another heuristically determined parameter, which is the so-called 'patch isometry index (PII)'. This index represents a numerical value ranging from 0 to 2, which is multiplied by the probability values of the eight cells belonging to the Moore neighbourhood, before the application of the transition function. A high isometry index results in compact patches, while low values are reflected in more fragmented landscape patterns. In this modelling experiment, an isometry index of 1.5 was adopted, what represents a balance between compactness and fragmentation. This value generates results in compliance with the deforestation pattern observed in the study area, which presents a mixture of geometrically stable (compact) deforestation patches produced by capitalised farmers that use tractors for the forest clear-cut, as well as fragmented patches generated by small farmers, deprived of sophisticated means for the forest removal.

5. Validation

For assessing the accuracy of the CA simulation model performance, fuzzy similarity measures applied within a neighbourhood context were used. Several validation methods operating on a pixel vicinity basis have been proposed (Costanza, 1989; Pontius, 2002; Hagen, 2002, 2003; Hagen-Zanker et al. 2005; Hagen-Zanker, 2006), aimed at depicting the spatial patterns similarity between a simulated and a reference map, so as to relax the strictness of the pixel-by-pixel agreement. The fuzzy similarity method employed in this work is a variation of the fuzzy similarity metrics developed by Hagen (2002), and has been implemented in the DINAMICA model by the CSR team.

Hagen´s method is based on the concept of fuzziness of location, in which the representation of a cell is influenced by the cell itself and, to a lesser extent, by the cells in its neighbourhood. Not considering fuzziness of category, the fuzzy neighbourhood vector can represent the fuzziness of location. In the fuzzy similarity validation method, a crisp vector is associated to each cell in the map. This vector has as many positions as map categories (land cover classes), assuming 1 for a category = i, and 0 for categories other than i. Thus, the fuzzy neighbourhood vector (V_{nbhood}) for each cell is given as:

$$V_{nbhood} = \begin{bmatrix} \mu_{nbhood\ 1} \\ \mu_{nbhood\ 2} \\ \vdots \\ \mu_{nbhood\ C} \end{bmatrix} \qquad (3)$$

$$\mu_{nbhood\ i} = |\ \mu_{nbhood\ i,1} * m_1,\ \mu_{crisp\ i,2} * m_2,\ \ldots,\ \mu_{crisp\ i,N} * m_N\ |_{Max} \qquad (4)$$

where $\mu_{nbhood\ i}$ represents the membership for category i within a neighbourhood of N cells (usually $N=n^2$); $\mu_{crisp\ ij}$ is the membership of category i for neighbouring cell j, assuming, as in a crisp vector, 1 for i and 0 for categories other than i (i $\subset$ C); m_j is the distance-based membership of neighbouring cell j, where m accounts for a distance decay function, for instance, an exponential decay (m = $2^{-d/2}$), with *d* being the unitary distance measured in between two cells centroids). The selection of the most appropriate decay function and the size of the window depend on the vagueness of the data and the spatial error tolerance threshold (Hagen, 2003). As it is intended to assess the model spatial fit at different resolutions, besides the exponential decay, a constant function equal to 1 inside the neighbourhood window and to 0 outside can also be applied. Equation (5) sets the category membership for the central cell, assuming the highest contribution is found within a neighbourhood window n x n. Next, a similarity measure for a pair of maps can be obtained through a cell-by-cell fuzzy set intersection between their fuzzy and crisp vectors:

$$S(V_A, V_B) = [\ |\ \mu_{A,1}, \mu_{B,1}\ |_{Min},\ |\ \mu_{A,2}, \mu_{B,2}\ |_{Min},\ \ldots,\ |\ \mu_{A,i}, \mu_{B,i}\ |_{Min}\ ]_{Max} \qquad (5)$$

where V_A and V_B refer to the fuzzy neighbourhood vectors for maps A and B, and $\mu_{A,i}$ and $\mu_{B,i}$ are their neighbourhood memberships for categories i $\subset$ C in maps A and B, as in Equation (6). According to Hagen (2003), since the similarity measure S (V_A,V_B) tends to overestimate the spatial fit, the two-way similarity is instead applied:

$$S_{twoway}(A, B) = |\ S(V_{nbhoodA}, V_{crispB}),\ S(V_{crispA}, V_{nbhoodB})\ |_{Min} \qquad (6)$$

The overall similarity of a pair of maps can be calculated by averaging the two-way similarity values for all map cells. However, when comparing a simulated map to the reference map (real land cover in the final time of simulation), this calculation carries out an inertial similarity between them due to their areas that did not suffer any change. To avoid this problem, the CSR team introduced a modification into the overall two-way similarity method of Dinamica EGO, using two maps of differences, which present value 1 for the cells that underwent change, and 0 for those that did not change. In this way, each type of change is analysed separately using pair-wise comparisons involving maps of differences: (i) between the initial land cover map and a simulated one, and (ii) between the same initial land cover map and the reference one. This modification is able to tackle two matters. First, as it deals with only one type of change at a time, the overall two-way similarity measure can be applied to the entire map, regardless of the different number of cells per category. Second, the inherited similitude between the initial and simulated maps can be eliminated from this comparison by simply ignoring the null cells from the overall count. However, there are two ways of performing this function. One consists of counting only two-way similarity values for non-null cells in the first map of difference, and the other consists in doing the opposite. As a result, three measures of overall similarity are obtained, with the third representing the average of the two ways of counting. As random pattern maps tend to score higher due to chance depending on the manner in which the null cells are counted, it is advisable to pay close attention to the minimum overall similarity value. This method has proven to be the most comprehensive when compared to the aforementioned methods, as it yields fitness measures with the highest contrast for a series of synthetic patterns that depart from a perfect fit to a totally random pattern (Soares-Filho et al., 2009).

6. Results and discussion

Table 1 presents the transition matrix resulting from the cross-tabulation operation between the initial (1997) and the final (2000) land cover maps. It provides the percentages of forest conversion, what corresponds in the particular case of this work to the global transition rates from forest to deforested areas, calculated as 13.8%. The classes 'deforested areas' and 'rivers' did not suffer any change during the study period.

Land Cover Classes		2000		
		Forest/ Grasslands	Deforested Areas	Rivers
1997	Forest/Grasslands	0.861641846	**0.138358154**	0
	Deforested Areas	0	1	0
	Rivers	0	0	1

Table 1. Land cover transition matrix for the period from 1997 to 2000

According to what was previously explained in Section 4.3, the calculation of the local transition probabilities, i.e. the probabilities of land cover change for each cell, is based on the values of the positive weights of evidence (W^+). Tables 2 and 3 present the values of W^+ for each distance range of the dynamic variable 'distance to previously deforested areas in 1997' and to the static variable 'distance to rivers', respectively. Fig. 7a and 7b graphically present the behaviour of the W^+ values in relation to the successive distance ranges of these two explaining variables.

The curve of W^+ for the variable 'distance to previously deforested areas in 1997' (Fig. 7a) reveals the concentration of weights with the greatest values in the first distance ranges, what demonstrates the predominance of new deforestation patches in the surroundings of pioneer areas (Alves, 2002; Aguiar et al., 2007). Said in other words, the pattern of deforestation expansion in the study area during the analysed period mostly presents patches of large extensions, following a trend to take place in the vicinities of previously deforested areas, also in face of the reported landed property concentration in this region.
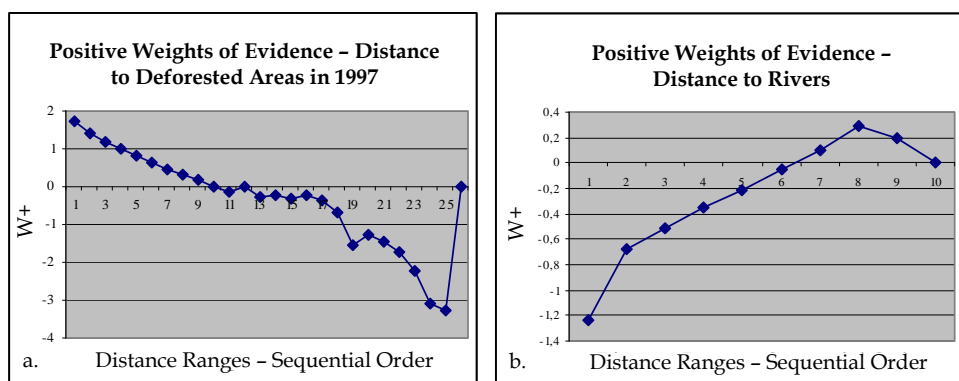

Fig. 7. Behaviour of the positive weights of evidence (W^+) in relation to the distance ranges for: a. distance to previously deforested areas in 1997, and b. distance to rivers

Values of the Positive Weights of Evidence		
Ranges Sequential Order	**Distance to Deforested Areas (m)**	**W+**
1	0 - 170	1.74594629
2	170 – 240	1.41006378
3	240 - 268	1.18230916
4	268 - 339	0.983065336
5	339 - 360	0.836426843
6	360 - 480	0.650478547
7	480 - 509	0.460166267
8	509 - 537	0.312746731
9	537 – 720	0.179722679
10	720 – 805	-0.0177924833
11	805 – 960	-0.143699617
12	960 – 975	-0.0110620882
13	975 – 988	-0.287776212
14	988 – 1,017	-0.221559314
15	1,017 – 1,045	-0.313774626
16	1,045 – 1,108	-0.247802748
17	1,108 – 1,137	-0.383389447
18	1,137 – 3,325	-0.686325958
19	3,325 – 3,335	-1.54902
20	3,335 – 3,340	-1.25484549
21	3,340 – 3,390	-1.4529907
22	3,390 – 3,757	-1.73329591
23	3,757 – 4,539	-2.2167735
24	4,539 – 4,720	-3.09234848
25	4,720 – 29,889	-3.28531919
26	29,889 – 2,147,483,647	0

Table 2. Values of the positive weights of evidence (W+) for the distance ranges of the dynamic variable 'distance to previously deforested areas in 1997'

Values of the Positive Weights of Evidence		
Ranges Sequential Order	**Distance to Rivers (m)**	**W+**
1	0 – 240	-1.23872974
2	240 – 360	-0.681253599
3	360 – 1,320	-0.515391121
4	1,320 – 1,440	-0.353605584
5	1,440 – 25,800	-0.212681384
6	25,800 – 25,920	-0.0495070727
7	25,920 – 29,400	0.104666667
8	29,400 – 29,760	0.286801266
9	29,760 – 32,880	0.193898811
10	32,880 – 2,147,483,647	0

Table 3. Values of the positive weights of evidence (W+) for the distance ranges of the static variable 'distance to rivers'

Regarding the variable 'distance to rivers' (Fig. 7b), it is possible to observe, however, that the distance ranges closest to rivers present the lowest W^+ values. This can be explained by the fact that, although deforestation processes to a reduced extent occur nearby rivers, these new deforestation patches account for a very limited share of the total deforested area in this region, what causes a decrease in the weights values referring to such distance ranges. As ranges are located ever further from rivers, their weights gradually start to assume increasing positive values. These farthest ranges actually correspond to the very bordering areas of well-established occupations, which are exactly those prone to experience deforestation processes. In brief, this variable acted as a fine tuning device for the variable 'distance to previously deforested areas in 1997'.

As previously stated in Section 4.3, Dinamica EGO generates maps of local probabilities based on the calculated W^+ values, assigning to each map cell a value of transition probability. Fig. 8 presents the map of local land cover change probabilities (from forest to deforested areas), generated in this experiment.
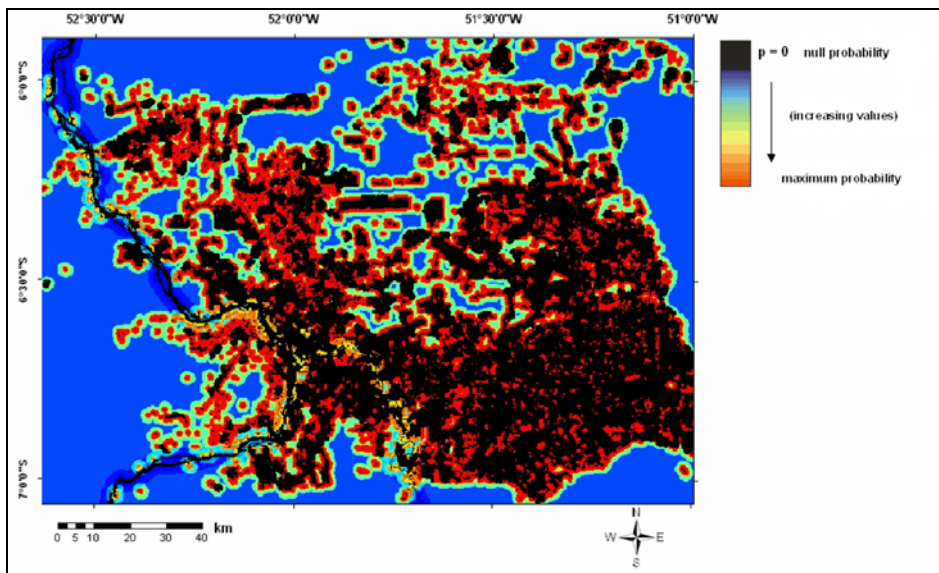

Fig. 8. Map of local land cover change probabilities (from forest to deforested areas) for the period from 1997 to 2000

The areas in black have null values of transition probability. The blueish and greenish cells own very low and low values of probability, while the areas in tones of yellow, orange, and red respectively present intermediate, upper intermediate and high values of transition probability. The areas with the highest probability values pretty much correspond to the areas where deforestation indeed occurred, as it can be observed in the PRODES deforestation map (Fig. 2).

The simulated deforestation map produced by the model (Fig. 9), considering both the static and the dynamic variables, presented high fuzzy similarity indices for multiple spatial resolutions, i.e. for multiple sampling window sizes (Table 4), in the case of the constant decay, what denotes the good performance of the model.

Comparing the simulated land cover map with the real land cover map in 2000 (Fig. 4), it is possible to notice that very regularly-shaped patches generated by real deforestation processes undertaken by capitalized farmers could not be reproduced in the simulation. This can be ascribed to the fact that the current transition functions available in Dinamica EGO still cannot cope with extremely rigid requirements regarding the patches geometry. Nevertheless, the spatial pattern of deforestation patches in the simulated map is very similar to the one found in the real scene, indicating the CA simulation model efficacy. It is worth mentioning that the aim of modelling is not to reproduce reality as close as possible, but solely to detect main spatial patterns and trends of land cover change. Spatial patterns refer to morphological aspects in the patches formation, i.e. whether they are geometrically stable or irregular, whether they are originated by expansion or diffusion processes, and so on, whilst trends concern the directions (or vectors) of deforestation propagation in space.
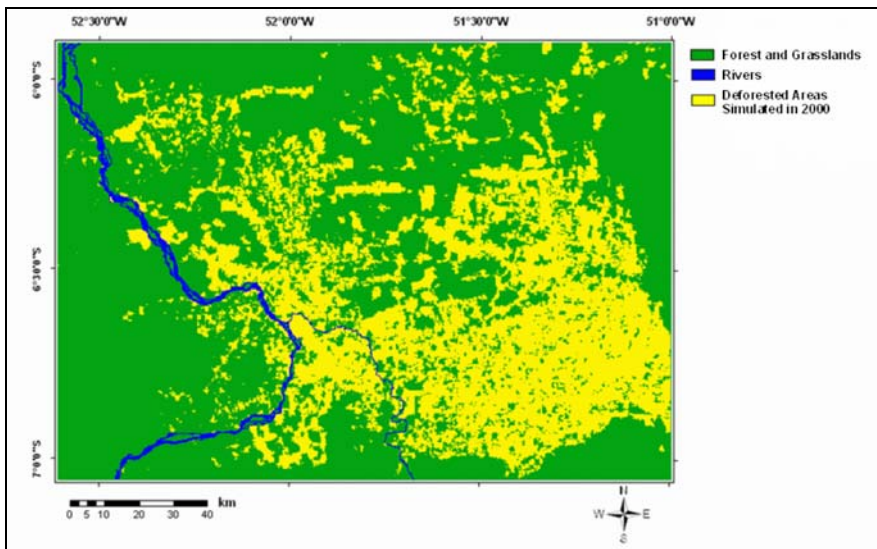

Fig. 9. Map of simulated deforestation for the year 2000, with patches average size set to 300 ha, and variance of size to 500 ha

Since the fuzzy similarity index (FSI) is a flexible method for CA model validation, in the sense that it does not operate on a pixel basis but rather on multiple levels of resolution, the values of this index in the cases where the constant decay is adopted tend to be slightly superior when compared to indices based on strict agreement, which are those derived from a direct pixel-by-pixel comparison between the real and the simulated scene. The agreement increases with the size of the sampling window until it reaches a resolution of around 11 x 11 or 13 x 13 pixels, when the similarity index stabilises (Fig. 10), what shows that the FSI is inefficient to assess the model fitness with rough spatial resolutions. It is worth reminding that the use of agreement indices based on multiple spatial resolutions to assess the performance of CA models can be justified by the fact that it is unfeasible not only to reproduce the configurations of changes in the past, but also and mainly to foresee future

land cover transitions with a very fine spatial accuracy, given the intrinsic randomness of land cover change processes.

CA Deforestation Model	Window Size (Pixels)	Parameters of Patches	FSI* (constant decay)
	3 x 3		0.877
	5 x 5	μ = 300 ha	0.891
	7 x 7	σ^2 = 500 ha	0.900
Simulation from 1997 to 2000	9 x 9	PII** = 1.5	0.903
	11 x 11		0.904
	13 x 13		0.905

*FSI = Fuzzy Similarity Index / PII** = Patch Isometry Index

Table 4. Result of the model validation using different sampling window sizes
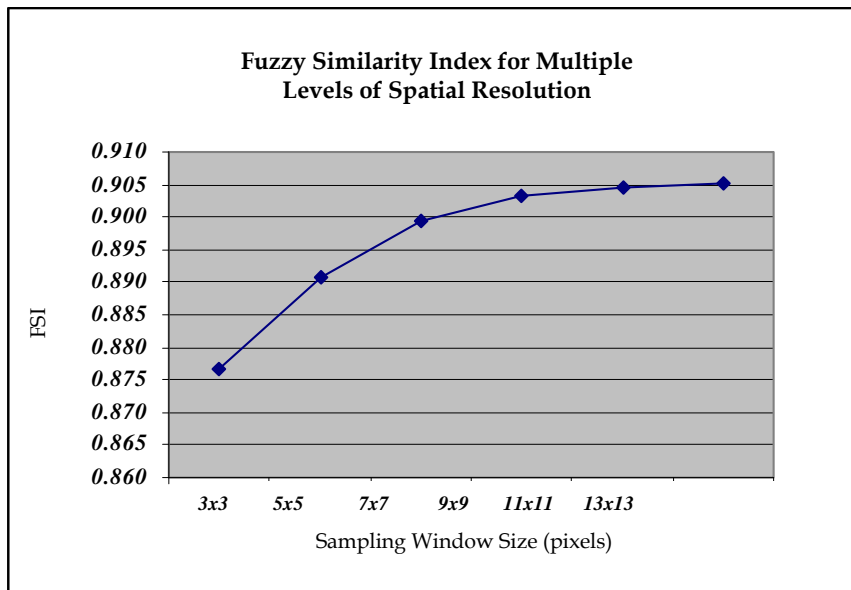

Fig. 10. Variation of the fuzzy similarity index (FSI) as a function of different sampling window sizes

7. Conclusion

This chapter presented an experiment on spatial dynamic modelling based on cellular automata, designed to simulate deforestation in the region of *São Félix do Xingu*, southeast of *Pará* State, north of Brazil, during the period from 1997 to 2000. The employed explaining variables (evidences) do not exhaustively represent the set of deforestation drivers in the area, but they correspond however to strategic inducing factors of such type of land cover change.

The dynamic variable 'distance to previously deforested areas in 1997' was decisive for simulating deforestation processes observed in the vicinities of forest clear-cut areas. In face of the historically reported landed property concentration in this region, the deforestation pattern in the analysed period mostly presents patches of large extensions, which are predominantly found in the surroundings of pioneer areas. The static variable 'distance to rivers', on the contrary, revealed the limited importance of the fluvial transport for deforestation and human occupation in this region. Nevertheless, considering that deforestation still takes place nearby rivers, this variable acted as an alternative for the fine tuning of the variable 'distance to previously deforested areas in 1997'. The obtained result demonstrates the suitability of the employed CA model for simulating deforestation processes in the study area for the time span from 1997 to 2000, what was confirmed by the high values of the fuzzy similarity index.

The Dinamica EGO platform proved to be appropriate for achieving the goals of this research experiment, also in view of one of its transition allocation algorithms – the expander function – which is able to reproduce the spatial pattern of deforestation expansion in the vicinities of previously occupied areas, as it is the case in *São Félix do Xingu*. Further advantages of Dinamica EGO concern its flexible structure that accepts different parameterization methods (Almeida et al, 2008) and varied sets of static and dynamic variables, which, when associated in a thoughtful manner, may meet the modelling requirements of the most diverse study areas, with peculiar land cover (or land use) change patterns. These properties allow the transferability of this and other models developed in this platform to other areas in the Amazon (Soares-Filho et al., 2009; Maeda et al., 2010), owning distinct occupation histories, at different consolidation stages, and involving specific local actors, with very particular characteristics.

As directions for future research, the authors intend to explore a wider scope of explaining variables, like indicators of the land ownership legal status, local unpaved roads as well as vectors of ongoing occupation fronts, besides operating with further stochastic parameterization methods. The authors as well intend to deal with longer temporal series, so as to use a more in-depth knowledge on the area occupation history, and hence, provide differentiated future simulation scenarios for *São Félix do Xingu* in the short- and medium-term. These forecasts would be based on plausible political, socio-economic, and infrastructural arrangements at the local and regional level.

Although models in general have continuously been the target of severe criticism, mainly in view of their reductionism and constraints to fully capture the reality inherent complexity (Briassoulis, 2000), it can be argued that they ought to exist, for they offer an incomparable way of abstracting patterns, order and main dynamic trends of real world processes (Batty, 1976).

Actually, urban models should be conceived, handled, applied and interpreted in a wise and critical way, so that modelers, practitioners, public and private decision-makers as well as citizens as a whole could take the best of what they can offer and sensibly acknowledge their limits (Almeida, 2003; Almeida et al., 2005). Spatial dynamic models, of which CA is one of the best representatives, are still the most promising means for rendering land cover change simulations communicable and transparent to politicians, planners, decision-makers as well as to the lay public in general.

8. Acknowledgement

9. References

Agterberg, F.P. & Bonham-Carter, G.F. (1990). Deriving weights of evidence from geo-science contour maps for the prediction of discrete events, *Proceedings of the XXII International Symposium AP-COM*, pp. 381–395, Berlin, Germany, September 17-21, 1990

Aguiar, A.P.D.; Câmara, G. & Escada, M.I.S. (2007). Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. *Ecological Modelling*, Vol. 209, No. 02-04, pp. 169–188, ISSN 0304-3800

Almeida, C.M. (2003). Spatial Dynamic Modelling as a Planning Tool: Simulation of Urban Land Use Change in Bauru and Piracicaba (SP), Brazil, In: *Biblioteca do INPE*, 15.09.2010, Available from http://mtc-m12.sid.inpe.br/col/sid.inpe.br/ jeferson/2003/12.18.07.36/doc/publicacao.pdf

Almeida, C.M.; Gleriani, J.M.; Castejon, E.F. & Soares-Filho, B.S. (2008). Using neural networks and cellular automata for modeling intra-urban land use dynamics. *International Journal of Geographical Information Science*, Vol. 22, No. 9, pp. 943–963, ISSN 1365-8816

Almeida, C.M.; Monteiro, A.M.V.; Câmara, G.; Soares-Filho, B.S.; Cerqueira, G.C.; Pennachin, C.L. & Batty, M. (2005). GIS and remote sensing as tools for the simulation of urban land use change. *International Journal of Remote Sensing*, Vol. 26, No. 04, pp. 759-774, ISSN 0143-1161

Alves, D.S. (2002). Space-time dynamics of deforestation in Brazilian Amazônia. *International Journal of Remote Sensing,* Vol. 23, No. 14, pp. 2903-2908, ISSN 0143-1161

Amaral, S.; Monteiro, A.M.V.; Câmara, G.; Escada, M.I.S. & Aguiar, A.P.D. (2006). Redes e conectividades na estruturação da frente de ocupação do Xingu-Iriri – Pará. *Geografia*, Vol. 31, No. 3, pp. 655-675, ISSN 0100-7912

Batty, M. (1976). *Urban modelling: algorithms, calibrations, predictions* (1.ed.), Cambridge University Press, ISBN 052-1134-36-6, Cambridge, UK

Becker, B.K. (2005). Geopolítica da Amazônia. *Revista de Estudos Avançados*, Vol.19, No. 53, pp. 71-86, ISSN 0103-4014

Bonham-Carter, G.F. (1994). *Geographic Information Systems for Geoscientists: Modelling with GIS*. Pergamon, ISBN 008-0418-67-8, Ontario, Canada

Brandão Júnior, A.O.; Souza Júnior, C.M.; Ribeiro, J.G.F. & Sales, M.H.R. (2007). Desmatamento e estradas não-oficiais da Amazônia, *Proceedings of the XIII Brazilian Symposium on Remote Sensing*, pp. 2357-2364, ISBN 978-851-7000-31-7, Florianópolis, Santa Catarina, Brazil, April 21-26, 2007

Briassoulis, H. (2000). Analysis of land use change: theoretical and modeling approaches, In: *University of Aegean, Greece*, 08.05.2010, Available from http://www.rri.wvu.edu/WebBook/Briassoulis/contents.htm

Burrough, P.A. (1998). Dynamic modelling and geocomputation, In: *Geocomputation: a primer*, Longley, P.A.; Brooks, S.M.; McDonnell, R. & MacMillan, B., pp. 165-192, John Wiley & Sons, ISBN 047-1985-76-7, Chichester, UK

Costanza, R. (1989). Model goodness of fit: a multiple resolution procedure. *Ecological Modelling*, Vol. 47, No. 03-04, pp. 199-215, ISSN0304-3800

Eastman, J.R. (2003). *IDRISI: The Kilimanjaro edition*, Clark University, Worcester, MA, USA

Escada, M.I.S; Amaral, S.; Monteiro, A.M.V.; Almeida, C.A.; Carriello, F. & Almeida, A. (2007). Padrões de mudança de uso e cobertura da terra na fronteira agropecuária de São Félix do Xingu, PA, *Proceedings of the I Symposium of the Research Network for Environmental Modelling in the Amazon [GEOMA]*, Petrópolis, RJ, Brazil, October 29-31, 2007

Escada, M.I.S.; Vieira, I.C.; Amaral, S.; Araújo, R.; Veiga, J.B.; Aguiar, A.P.D.; Oliveira, I.V.M.; Pereira, J.L.V.; Filho, A.C.; Fearnside, P.M.; Venturieri, A.; Carriello, F.; Thales, M.; Carneiro, T.S.G.; Monteiro, A.M.V. & Câmara, G. (2005). Processos de ocupação nas novas fronteiras da Amazônia (o interflúvio do Xingu/Iriri). *Estudos Avançados*, Vol. 19, No. 54, pp. 9-23, ISSN 0103-4014

Research Network for Environmental Modelling in the Amazon [GEOMA]. (2004). *Dinâmica territorial da frente de ocupação de São Félix do Xingu-Iriri: Subsídios para o desenho de políticas emergenciais de contenção do desmatamento*, Ministério da Ciência e Tecnologia, Brasília, Brazil

Goodacre, A.K.; Bonham-Carter, G.F.; Agterberg, F.P. & Wright, D.F. (1993). A statistical analysis of the spatial association of seismicity with drainage and magnetic anomalies in western Quebec. *Tectonophysics*, Vol. 217, No. 03-04 , pp. 285-305, ISSN 0040-1951

Hagen, A. (2002). Multi-method assessment of map similarity, *Proceedings of the 5th AGILE Conference on Geographic Information Science*, pp. 171–182, Palma de Mallorca, Spain, April 25-27, 2003

Hagen, A. (2003). Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science*, Vol. 17, No. 3, pp. 235-249, ISSN 1365-8816

Hagen-Zanker, A.; Straatman, B. & Uljee, I. (2005). Further developments of a fuzzy set map comparison approach. *International Journal of Geographical Information Science*, Vol. 19, No. 7, pp. 769-785, ISSN 1365-8816

Hagen-Zanker, A. (2006). Map comparison methods that simultaneously address overlap and structure. *Journal of Geographical Systems*, Vol. 8, No. 2, pp. 165-185, ISSN 1435-5930

Instituto Nacional de Pesquisas Espaciais [INPE]. (2006). PRODES, In: *OBT-INPE*, 28.11.2007, Available from http:// http://www.obt.inpe.br/ prodes/

Laurance, W.F.; Albernaz, A.K.M.; Schroth, G.; Fearnside, P.M.; Bergen, S.; Venticinque, E.M. & Costa, C. (2002). Predictors of deforestation in the Brazilian Amazon. *Journal of Biogeography*, Vol. 29, No. 05-06, pp. 737-748, ISSN 1365-2699

Maeda, E.E.; Almeida, C.M.; Ximenes, A.C.; Formaggio, A.R.; Shimabukuro, Y.E.; Pellikka, P. (2010). Dynamic modeling of forest conversion: Simulation of past and future scenarios of rural activities expansion in the fringes of the Xingu National Park, Brazilian Amazon. *International Journal of Applied Earth Observation and Geoinformation*, ISSN 0303-2434, In press, DOI 10.1016/j.jag.2010.09.008

Openshaw, S. (2000). Geocomputation, In: *Geocomputation,* Openshaw, S.; Abrahart, R. J., (Eds.), pp. 1-31, Taylor & Francis, ISBN 978-047-1985-76-1, New York, USA

Pereira, L.M.; Escada, M.I.S. & Rennó, C.D. (2007). Análise da evolução do desmatamento em áreas de pequenas, médias e grandes propriedades na região centro-norte de Rondônia, entre 1985 e 2000, *Proceedings of the XII Brazilian Symposium on Remote Sensing*, pp. 6905-6912, ISBN 851-7000-18-8, Goiânia, GO, Brazil, April 16-21, 2005

Pontius Jr., G. (2002). Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering & Remote Sensing*, Vol. 68, No. 10, pp. 1041-1049, ISSN 0099-1112

Rodrigues, H.O.; Soares-Filho, B.S. & Costa, W.L.S. (2007). Dinamica EGO, uma plataforma para modelagem de sistemas ambientais, *Proceedings of the XIII Brazilian Symposium on Remote Sensing*, pp. 3089-3096, ISBN 978-851-7000-31-7, Florianópolis, Santa Catarina, Brazil, April 21-26, 2007

Santana, L.F. (2007). *São Félix do Xingu e sua história: 1889 – 1997*, Prefeitura Municipal de São Félix do Xingu, São Félix do Xingu

Schimink, M. & Wood, C.H. (1992). *Contested Frontiers in Amazonia*, Columbia University Press, New York, USA

Soares-Filho, B.S.; Assunção, R.M. & Pantuzzo, A.E. (2001). Modeling the spatial transition probabilities of landscape dynamics in an Amazonian colonization frontier. *BioScience*, Vol. 51, No. 12, pp. 1059-1067, ISSN 0006-3568

Soares-Filho, B.S.; Cerqueira, G.C. & Pennachin, C.L. (2002). DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling*, Vol. 154, No. 03, pp. 217-235, ISSN 0304-3800

Soares-Filho, B.S.; Nepstad, D.C.; Curran, L.M.; Cerqueira, G.C.; Garcia, R.A.; Ramos, C.A.; Voll, E.; McDonald, A.; Lefebvre, P. & Schlesinger, P. (2006). Modelling conservation in the Amazon basin. *Nature*, Vol. 440, No. 7083, pp. 520-523, ISSN 0028-0836

Soares-Filho, B.S.; Rodrigues, H.O. & Costa, W.L.S. (2009). Modeling Environmental Dynamics with Dinamica EGO, In: *CSR-UFMG*, 23.06.2008, Available from http://www.csr.ufmg.br/dinamica/, ISBN 978-85- 885910119r-r0-2

Wolfram, S. (1983). Statistical mechanics of cellular automata. *Review of Modern Physics,* Vol. 55, pp. 601-643, ISSN 0034-6861