



Jul 1st, 12:00 AM

Modeling Urban Growth and Land Use Change with Cellular Automata and Genetic Algorithms

Arlex Sanchez

Wilmer Barreto

Neiler Medina

Zoran Vojinovic

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>

Sanchez, Arlex; Barreto, Wilmer; Medina, Neiler; and Vojinovic, Zoran, "Modeling Urban Growth and Land Use Change with Cellular Automata and Genetic Algorithms" (2012). *International Congress on Environmental Modelling and Software*. 30.
<https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/30>

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Modeling Urban Growth and Land Use Change with Cellular Automata and Genetic Algorithms

Arlex Sanchez¹, Wilmer Barreto², Neiler Medina¹, Zoran Vojinovic¹

¹ Department of Hydroinformatics and Knowledge Management, UNESCO-IHE,
Delft, The Netherlands

² University Lisandro Alvarado, Venezuela
torre3@unesco-ihe.org

Abstract: It is a well accepted fact that urbanisation, climate change and population growth represent an enormous challenge for urban water managers. In this respect, computer models coupled with spatial mapping techniques have proved to be invaluable. The present paper demonstrates the use of a cellular automata model (Dinamica Ego) for modeling land use change process on a case study of Birmingham (UK). Two approaches were evaluated for this purpose. The results obtained show that the model based on optimization of parameters that deal with the process of expansion/contraction is capable of producing promising results. The analysis was carried out using the Corine dataset for the years 1990 and 2000 and the fuzzy similarity test was used as the objective function. To minimize computational demands the optimization loop was simulated with the NSGA II algorithm using the parallel computing approach.

Keywords: Urban growth, Land Use Change Modeling, Cellular Automata, Genetic Algorithms.

1. INTRODUCTION

Cities can be considered as complex systems considering their characteristics of emergence, self-similarity, self-organization and non-linear behaviour of land use changes with time [Batty and Langley 1994]. The use of computer modeling tools can help in the understanding of the above-mentioned characteristics and to gain knowledge about the patterns and mechanisms behind urban dynamics. Cellular automata (CA) models have been used increasingly in many applications and tested on several case studies.

Several authors have used CA models to predict land-use changes at the catchment scale. Their results have shown that the application of CA for land-use change modelling is feasible and that the outcome of the models is close to reality. The Moland framework, [Barredo et al. 2003] an example of such a model that has been successfully applied and calibrated in case studies in Europe. Several examples using cellular automata models like Dinamica EGO and SLEUTH among others have been applied in different cities around the world [see for example, Barredo et al. 2003, Engelen et al. 2007, Soares et al. 2011].

The present paper describes the integration between a cellular automata model (Dinamica Ego) with the Non dominated Sorting genetic algorithm NSGA II, developed by [Deb 2002] for the purpose of modeling land-use change in Birmingham (UK). The analysis was done using the Corine dataset for the years 1990 and 2000. To assess the spatial performance of the model the fuzzy similarity test proposed by [Hagen 2003] and incorporated in Dinamica Ego [Soares 2011] was used as the objective function. Four different sets of parameters were used

during the optimization process: 1.) the initial set of weights of evidence within a bound of ± 1.2 the initial value. 2.) The parameters for the modulate change matrix. 3.) The set of parameters for the Expander (i.e. mean patch size, patch size variance, isometry). 4.) The set of parameters for the Patcher (i.e. mean patch size, patch size variance, isometry).

2. METHODOLOGY

In this study, DINAMICA EGO software is used as a simulation platform for our urban dynamics model [Soares et al. 2011]. DINAMICA employs, as input, a set of maps. This includes the initial and final map of land use, also known as landscape maps. The model considers that a landscape could be viewed as a bi-dimensional array of land use types. The other inputs are two sets of ancillary maps: the static and dynamic variables, the latter named so because they are iteratively updated by the model. These two sets of variables control the spatial allocation of changes, since they are used to calculate the weights of evidence. The weights of evidence represents each variables influence on the spatial probability of transition a to b. The variables are combined by summing their weights of evidences [Goodacre et al. 1993, Bonham-Carter 1994, Soares-Filho et al. 2011] to produce a transition probability map, which depicts the most favourable areas for change [Soares-Filho et al. 2002, 2004, 2011]. The "weights of evidence" method is a Bayesian method, in which the effect of each spatial variable on a transition is calculated independently of a combined solution.

DINAMICA EGO uses as a local CA rule, a transition engine composed of two complementary transition functions, the Expander and the Patcher [Soares-Filho et al. 2002]. DINAMICA splits the cell selection mechanism into these two processes. The first process is dedicated only to the expansion or contraction of previous patches of a certain class, and it is called Expander. The second process is designed to generate or form new patches through a seeding mechanism, and it is called Patcher. For each transition of land use to another, the percentage of cell changes executed by the Expander function in relation to the Patcher must be defined. Within the software this process is handled by the Modulate Change matrix function. This matrix specifies the percentage of changes (number of cells per transition) that will be executed by the expander and the Patcher. These values need to be fine tune in the calibration process.

The size of new patches and expansion fringes are set according to a lognormal probability distribution. Therefore, it is necessary to specify the parameters of this distribution represented by the mean and variance of the patch sizes to be formed [Soares et al. 2011]. The size and variance of the patches and expansion varies from case to case because of the peculiarities of each urban area. Because the range over which parameter values can vary is large, the model requires calibration. The other parameter that requires calibration is the patch isometry index, which varies from 0 to 2. The patches assume a more isometric form as this index increases [Soares et al. 2011].

The validation of the outcome of the cellular automata model consists of a comparison between the model results and a reference map, in this case, the land use map at the simulation final time. There are several map comparison techniques to assess the spatial match between two maps. Nonetheless, there is no consensus about which technique yields the most appropriate validation. The fuzzy comparison method described by [Hagen 2003] was adapted to be used in Dinamica and in this study.

For this study, four different sets of parameters were used during the optimization process. The initial set of weight of evidence functions that shows the attraction/repulsion of each transition in respect to each variable were optimized

using a bound of ± 1.2 of the initial value. This is done according with the results of previous research done by the developers of DINAMICA EGO. In their study, different genetic algorithms were tested in conjunction with several combinations of adjusting bounds for the weight of evidences [Soares-Filho et al. 2011]. After this step, the best set of functions is then fixed and used to estimate the parameters that regulate the expansion/contraction of the clusters. Two approaches were used to calibrate this part of the model. The first consisted in changing the parameters of every object each at the time before moving to the next. The second consisted in changing all the parameters of the patcher, expander and the modulate change matrix at the same time.

To handle the optimization process the NSGA-II algorithm developed by [Deb 2002] was used. The NSGA-II has been tested and proved to develop good Pareto fronts and can manage several objective functions and constraints [i.e. Barreto et al. 2006, Muschalla 2006]. The DINAMICA EGO 1.6 developed by [Soares et al. 2011] is used as computational engine to simulate the land use changes.

To link the genetic algorithm with the computational engine a set of intermediate link routines were written. The first one runs reads the original model and gets the original values of the parameters that need to be adjusted and computes the original spatial correlation between the outcome of the model and the real map. The second routine directly links the NSGA II and DINAMICA by interpreting the randomly generated population of the GA for the variables and computing the value of the objective functions that are passed to the GAs for evaluation and generation of further populations. The stopping criteria for the optimization loop is set by the number of combinations (Population size times the number of generations) to be evaluated, Figure 1 shows the optimization loop.

The optimization loop to calibrate the parameters was run for 144 generation and 12 populations (for a total of 1728 evaluations). The objective function consisted of the fuzzy similarity fitness indicator

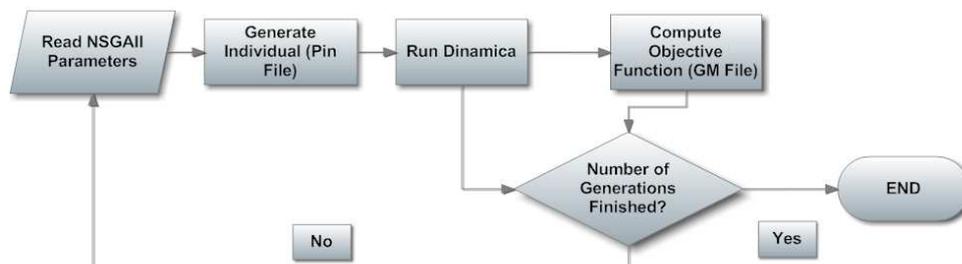


Figure 1. Optimization loop used for calibration.

3. CASE STUDY

The case of study city is Birmingham city located in the West Midlands county of England. The city total area is 268 km². Birmingham city is the most populous British city outside London, with a population of 1,028,700 in 2009 based on [Birmingham's City Council 2012]. The city experienced rising population as the growth rate of population increase from the year 2000 to year 2010 is 4.2 %.

3.1 Setup and Operating the Cellular Automata Model

The Dinamica model requires two set of maps for land use, for this particular case the land use maps corresponding to the Corine Dataset from the European Spatial

Agency were used. The maps correspond to the land-use classification for the years 1990 and 2000. The Corine land use classes were regrouped into 4 classes to be modeled.

To setup Dinamica Ego all the input maps have to be transformed to raster maps. It is very important that all of them are projected in the same coordinate system, so that they overlay exactly one on top of the other. The cell size has to be the same for all the spatial dataset, in this case the cell size is 100 by 100 meters. Figure 2 presents the spatial dataset (Variables or evidence used to explain the land use changes) that was collected and processed to set-up the model. The maps were processed using ArcGIS 9.3. These raster maps are static variables that do not change during the simulation.

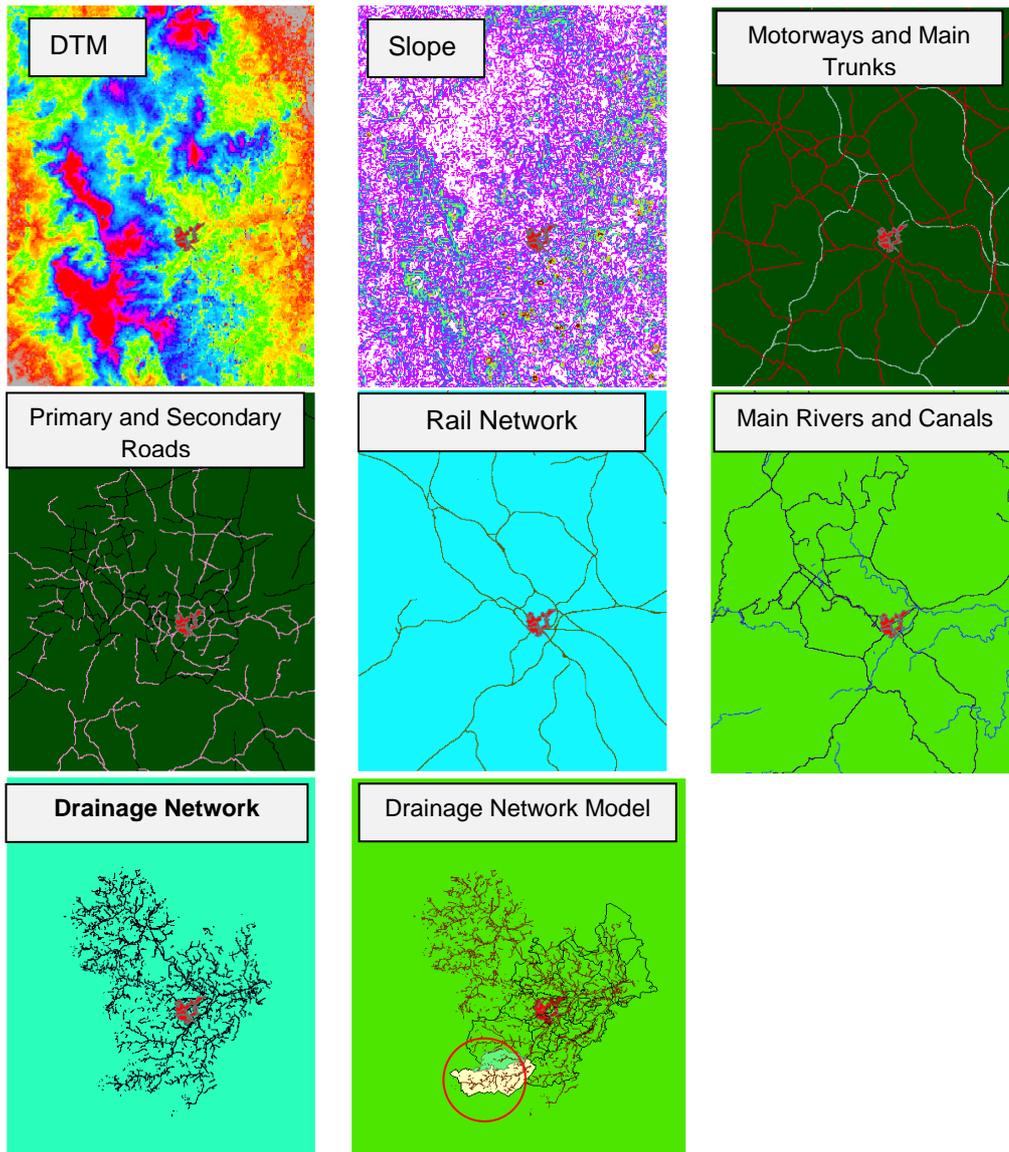


Figure 2. Input map data set for Dinamica Ego.

The model was run with initial land use map for the year 1990 and for time span of 10 years to simulate the land use for the year 2000. The simulated map is shown in figure 4 with comparison to the real land use map of the year 2000. To compare

the output of the simulation with the real land use map DINAMICA can calculate a similarity map to check the spatial correlation between the two.

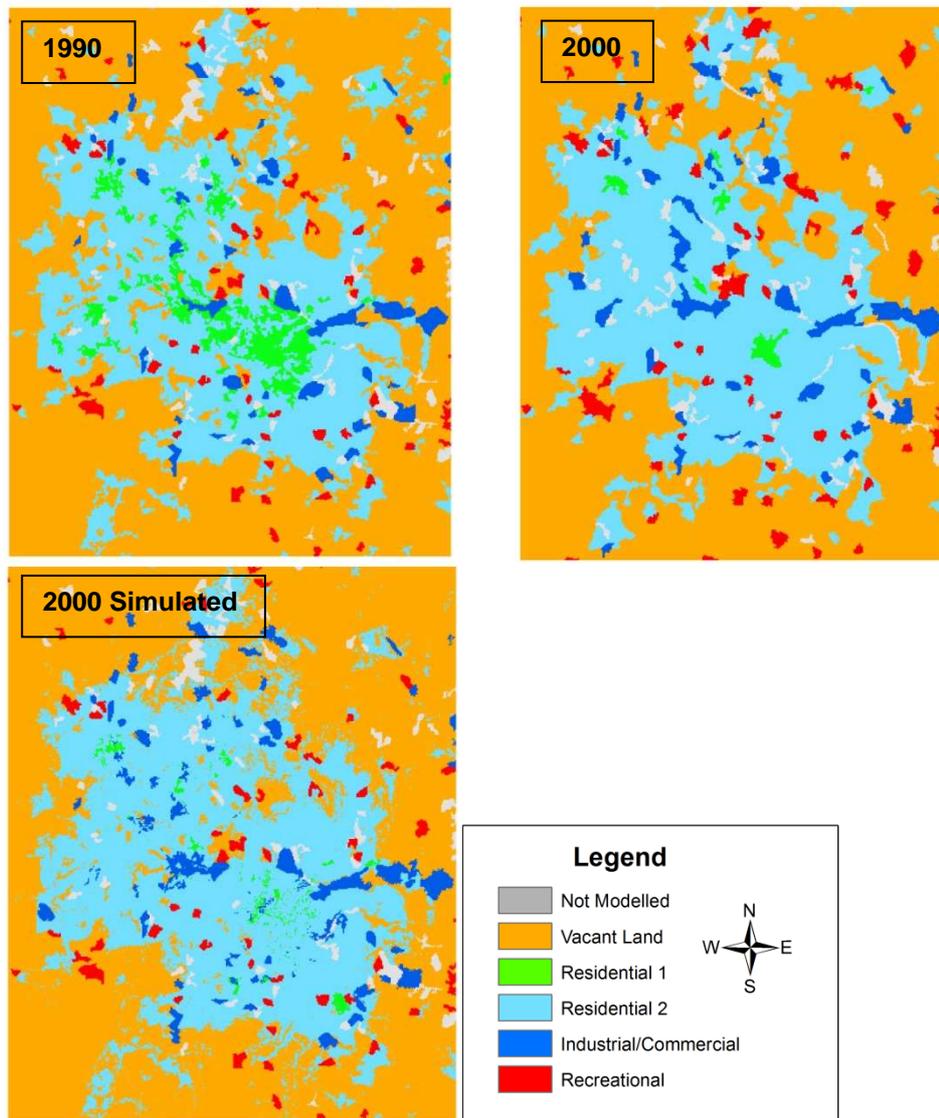


Figure 3. Actual land use for the years 1990 and 2000 and simulated land use for the year 2000.

For the calibration of the model the parameters that define the Expander, the Patcher and the Modulate change matrix that regulate the percentage of changes done by each of these two processes were linked with the NSGA II Genetic algorithm. A set of tests were conducted to evaluate the performance of optimization by using the option of DINAMICA to automatically detect the number of processors available in the machine or by passing the number of processors to be used as a fixed parameter. The outcome of this test was that is better to use the automatic detection option to detect the number of processors available in the computer that is included in Dinamica. For a single machine the number of simulations achieved was 8.5 per minute. The computer used for the runs had an Intel I5 processor, which has 4 processors in parallel. The total RAM was 4 gigabytes and the operating system corresponded to Windows 7.

The first step was to adjust the Weights of Evidence matrix that is initially computed by Dinamica. This consisted of 150 functions that describe the attraction/repulsion

effect of each variable per transition. The functions were adjusted within the bound of ± 1.2 of the initial value. The ranges used for the parameters of the Modulate change matrix, Expander and Patcher are presented in Table 1.

Table1 . Variables and intervals used for Calibration

Transitions		Modulate Change Matrix (%)	Expander			Patcher		
From	To		Mean Patch Size (Ha)	Patch Size Variance (Ha)	Patch Isometry	Mean Patch Size (Ha)	Patch Size Variance (Ha)	Patch Isometry
1	3	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
1	4	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
2	3	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
2	4	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
3	1	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
3	2	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
3	4	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
4	3	0 - 1	1 - 2000	2 - 3000	0 - 2	1 - 2000	2 - 3000	0 - 2
Number of Variables for Calibration		8	24			24		

The first run was to adjust the Weights of evidence matrix. This run alone improved the fitness of the model from 0.18 obtained initially to 0.41. The best Weight of evidence matrix was then fixed to perform the following experiments:1.) The modulate change matrix (8 Variables, 1728 simulations), 2.) The Expander (24 Variables, 1728 Simulations), 3.) The Patcher (24 Variables, 1728 Simulations). The results of the best simulation are presented in table 2, as well as the best fitness value.

Table2 . Values for each variable and best fitness value

Transitions		Modulate Change Matrix (%)	Expander			Patcher		
From	To		Mean Patch Size (Ha)	Patch Size Variance (Ha)	Patch Isometry	Mean Patch Size (Ha)	Patch Size Variance (Ha)	Patch Isometry
1	3	0.86	231.78	1856.25	0.26	984.98	1146.50	1.13
1	4	0.71	2.36	1037.73	0.68	1493.70	2802.61	1.26
2	3	0.90	1918.89	2544.72	0.52	1417.12	783.16	0.85
2	4	0.10	1863.19	1370.60	1.63	1443.31	1026.31	1.66
3	1	0.56	481.85	1879.25	0.66	1787.29	1796.29	1.60
3	2	0.22	15.12	523.95	0.76	1345.92	231.77	1.56
3	4	0.75	945.30	1860.94	1.24	805.57	2499.00	1.55
4	3	0.93	566.09	1985.54	1.70	1461.52	870.11	0.89
Fitness of the best simulation		0.46	0.45			0.46		

The best values found in the calibration of each individual function (i.e. Modulate change matrix, Expander and Patcher) were then used to build a single model with the best parameters. The value achieved of the fitness indicator was 0.42. This value of fitness is lower than the ones achieved by the calibration of each individual function. This can be explained because there is a direct relation between the amount of changes or cells that are allocated by the model with the Patcher and the Expander. Therefore, the combination of these parameters is important, rather than the adjustment of each individual sub-set of parameters per function. Another test was prepared for adjusting the parameters of the Modulate change matrix, the Patcher and the Expander at the same time. The results show similar values for the parameters of each function (Modulate change matrix, Expander, Patcher) and the value of the fitness was 0.45. Not big differences were achieved.

As presented in table 2, the automatic calibration process is promising because all the evaluations needed to improve the model performance are time-consuming to be done manually. Particularly, in this case the interest in developing the model is to simulate the internal changes that occur in the city. With the amount of land use classes to be modelled the case becomes more complex. The fuzzy similarity indicator to compare the simulated land use map and the real map showed a good correlation, more than 0.40. Even though, it is possible that all the variables that can explain the internal urban dynamics are not used. Moreover, it is possible that the necessary information and datasets needed to achieve higher correlation factor are probably not measured in reality. This approach is currently being extended to use the NSGAXp version of the NSGA II that runs in parallel developed by [Barreto et al. 2006]. Initial test showed that by using a cluster of 16 processors a number of 44 simulations per minute can be achieved. The idea is to increase the computational power to better assess the interval to select the set of parameters for the patch size, patch size variance and isometry used by the Patcher and the Expander. The interval of these parameters is quite big and there are not many values reported in the literature or other studies.

4. CONCLUSIONS

The present paper describes the preliminary results obtained from an ongoing research project which aims to develop a calibration approach for a land use change model. The analysis was carried out using the Corine dataset for the years 1990 and 2000 and the fuzzy similarity test was used to assess the spatial correlation and as the objective function in the calibration process. In order to minimize computational demands the optimization loop was simulated with the NSGA II algorithm using the parallel computing approach. The best option to make better use of the multiprocessor computer is to set the auto-detection number of processors included in Dinamica. Two approaches were evaluated to conduct the calibration of the model. The results obtained show that the two calibration approaches produce similar results. Nevertheless, the calibration approach that uses all the parameters that deal with the process of expansion/contraction in the optimization loop is recommended. The step by step approach shows that a good correlation is found by an individual set at the beginning of the process, in the next steps a lower correlation can be achieved or is hardly increased. The highest gain in correlation is achieved initially by adjusting the weight of evidence matrix that is at the heart of the land use transitions. The overall results demonstrate the possibility of using an integrated modeling approach as a planning tool to evaluate future scenarios in urban areas.

ACKNOWLEDGMENTS

This study has been carried out within the framework of the European research project SWITCH (Sustainable Urban Water Management Improves Tomorrow's

City's Health). SWITCH is supported by the European Commission under the 6th Framework Programme and contributes to the thematic priority area of "Global Change and Ecosystems" [1.1.6.3] Contract n°018530 -2.

REFERENCES

- Barredo J., Kasanko M., McCormick N., Lavalle C. Modeling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Journal of Landscape and urban planning*. Vol 64. pp 145-160. Elsevier Science. 2003.
- Barreto W. J., Vojinovic Z., Price R.K., Solomatine D.P. Approaches to multi-objective multi-tier optimization in urban drainage planning. 7th international conference on hydro-informatics. Nice, France. 2006.
- Birmingham's City Council. <http://www.birmingham.gov.uk>. Accessed via Web. 2012
- Bonham-Carter, G Geographic information systems for geoscientists: modeling with GIS. Pergamon, New York, 1994.
- Engelen G., C. Lavalle , J.I. Barredo, M. van der Meulen and R. White. The Moland modeling framework for urban and regional land-use dynamics. In: *Modeling Land Use change: Progress and applications*. Edited by Koomen Eric, Stillwell John, Bakema Aldrik, Scholten Henk J. Springer. The Netherlands. 2007.
- Goodacre CM., Bonham-Carter GF., Agterberg FP., Wright DF. A statistical analysis of spatial association of seismicity with drainage patterns and magnetic anomalies in western Quebec. *Tectonophysics*, v. 217, 205-305, 1993.
- Hagen A. Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science*, v. 17, 235–249, 2003.
- Muschalla D. Optimization of water resources systems using multi-objective evolution strategies. 7th international conference on hydro-informatics. Nice, France. 2006.
- Soares Filho BS., Pennachin CL., Cerqueira G. DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling*, v. 154, 217-235, 2002.
- Soares-Filho B., Alencar A., Nepstad D., Cerqueira G., Vera Diaz M., Rivero S., Solórzano L., Voll E. Simulating the Response of Land-Cover Changes to Road Paving and Governance Along a Major Amazon Highway: The Santarém-Cuiabá Corridor. *Global Change Biology*, v. 10, 745-764, 2004.
- Soares-Filho BS., Cerqueira G., Araujo W., Voll E. DINAMICA project. <http://www.csr.ufmg.br/dinamica> visited February, 2011.