ASSESSING SIMULATED LAND USE/COVER MAPS USING SIMILARITY AND FRAGMENTATION INDICES

Jean-François Mas, Researcher
Centro de Investigaciones en Geografía Ambiental
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO (UNAM)
Antigua Carretera a Pátzcuaro No. 8701
Col. Ex-Hacienda de San José de La Huerta
C.P. 58190 Morelia Michoacán MEXICO
jfmas@ciga.unam.mx

Azucena Pérez Vega, Researcher
Departamento de Ingeniería Civil - UNIVERSIDAD DE GUANAJUATO
Av. Juárez 77, Colonia Centro C.P. 3600 Guanajuato Gto MEXICO
azuperezvega@yahoo.com

Keith Clarke, Professor
Department of Geography – UNIVERSITY OF CALIFORNIA – SANTA BARBARA
Department of Geography
U.C. Santa Barbara
Santa Barbara CA 93106-4060 USA
kclarke@geog.ic.ucsb.edu

ABSTRACT

Land use/cover changes (LUCC) are significant to a range of issues central to the study of global environmental change. Over the last decades, a variety of models of LUCC have been developed to predict the location and patterns of land use/cover dynamics. The simulation procedures of most computational LUCC models can be sub-divided into three basic steps, 1) a non-spatial procedure which calculates the quantity of each transition, 2) a spatial procedure which allocates changes to the more likely locations and eventually replicates the patterns of the landscape and, 3) an evaluation procedure which compares a simulated land use/cover map with the true map of the same date. However, most of the evaluation techniques are focused in assessing the location of the simulated changes in comparison with the true changes and do not assess the ability of the model to simulate the landscape patterns (e.g. size, shape and distribution of patches). This study aims at evaluating simulated land use/cover maps obtained by two models (DINAMICA and Land Change Modeler). Simulated maps were evaluated using a fuzzy similarity index which takes into account the fuzziness of location within a cell neighborhood and fragmentation indices. Results show that more realistic simulated landscapes are often obtained at the expense of the location coincidence. When patterns of landscape are important (e.g. when considering fragmentation effects on biodiversity), it is important to incorporate indices that take into account the spatial patterns, and not merely location, during the model assessment procedure.

INTRODUCTION

Land use/cover change (LUCC), especially the conversion of forested areas into other uses, has been identified as a contributing factor to climate change, accounting for 33 percent of the increase in atmospheric CO_2 since 1850, and is a leading factor in the loss of biological diversity (de Sherbinin, 2002). Over the last decades, a range of "spatially explicit" computational models of LUCC have been developed and represent an important suite of techniques for the projection of alternative scenarios into the future, for conducting experiments that test our understanding of key processes, and for describing the latter in quantitative terms (Veldkamp and Lambin, 2001; Xiang and Clarke, 2003). Among these models, process-based models, closely related with geographic information systems, view land use and cover changes as transition process from one state to other states. Typical examples are models based on cellular automata and Markov process models, such as the two models used in the present study.

MATERIAL

For LUCC modeling, the programs DINAMICA EGO (version 1.4.0) and Land Change Modeler in IDRISI (version 16.05) were used. DINAMICA EGO is a cellular automata-based model which has been applied in a variety of studies, including modeling tropical deforestation from local to basin-wide scales (Soares-Filho et al., 2002, 2006, Cuevas and Mas, 2008; Texeira et al., 2009) and urban growth and dynamics (Almeida et al., 2003; Godoy and Soares-Filho, 2008). Land Change Modeler (available in IDRISI) provides tools for the assessment and projection of land cover change, and their implications for species habitat and biodiversity (Eastman, 2006; Václavík and Rogan, in press; Gontier et al., 2009). Statistical analysis and graphs were created using R (R Development Core Team, 2009).

Modeling was carried out using the data set supplied with the IDRISI tutorial, developed by the Conservation International's Center for Applied Biodiversity Science at the Museo Noel Kempff Mercado in Bolivia. The data set consists of land cover (LC) maps and ancillary information from a rapidly changing area in the Bolivian lowlands. The study area is about 200 km to the north/northwest of Santa Cruz de la Sierra. This is a region of rolling hills at the ecotone between the Amazonian forest and deciduous dryland tropical forest. It is not well suited to mechanized agriculture, but has economic value for both cattle and timber production (Eastman, 2009.)

The data used in the present study are the LC maps of 1986 and 1994 and several maps used as explanatory variables (maps of distance from urban areas, distance from roads, slope, distance from disturbance, elevation). As the LC maps were intended for ecosystem monitoring, they do not distinguish between settlements and agriculture, both included in the category of anthropogenic disturbance. This also includes secondary forest – once disturbed, land remains in that class. The vast majority of disturbed areas are used for pasture – either for dairy (primarily in the south east) or for beef production (Eastman, 2009).

METHODS

The present research has two-fold goals: 1) creation of modeled LC maps using the two programs; and 2) the assessment of these maps using two approaches: a) based on the spatial coincidence; and b) landscape metrics.

LUCC Modeling

LUCC are modeled empirically by using past change to develop a mathematical model; and GIS data layers influence the transition potential. First the model is calibrated using a map of LUCC obtained through the comparison of LC maps at two different dates (1986 and 1994 in the present case). A non spatial procedure calculates the quantity of each type of change and a spatial analysis allows the identification of more likely change locations using a set of explanatory variables (typically slope, elevation, distance to roads, distance to human settlements, or previous change). The result is a change susceptibility map for each transition. From these maps, two outputs can be obtained: 1) a map of susceptibility to change for the selected set of transitions and; 2) a LC map for a future date. The present study is focused on the elaboration and assessment of the LC map. In order to create the simulated LC map a cellular automata procedure replicates the pattern of the landscape.

Finally, an assessment procedure allows comparing the simulated map with the true map at the posterior date. In the present study, as we are interested in assessing landscape pattern simulation rather than predictive performance, we simulated a 1994 LC map from the model calibrated over the period 1986-94 (e.g. simulation and calibration period are the same). In the following paragraphs, the calibration and modeling procedures are described more precisely.

Quantity of change estimate. The changes were computed from a Markov matrix obtained through the comparison of LC maps of two previous dates, which is the standard procedure in DINAMICA and LCM, although both programs allow using a quantity of change provided by an external model.

Potential for change. Based upon the relationship between the different transitions and the explanatory variables, maps of change susceptibility are produced for each transition. In order to establish this relationship, DINAMICA and LCM use the weights-of-evidence method and an artificial neural network (ANN) respectively. In the present study, the DINAMICA model uses the map of probability elaborated by the LCM in order to obtain comparable results.

Reproduction of spatial patterns. DINAMICA and IDRISI use a cellular automata approach in order to obtain a proximity effect (areas which are close to existing areas of a certain class are more likely to change to this class) and eventually simulate landscape pattern.

In IDRISI, the process involves a 3x3 filter which traverses the image and reclassifies pixels to incorporate the effects of neighboring pixels on a current pixel value. There is no option to control the CA behavior.

DINAMICA uses two complementary transition functions: 1) the Expander; and 2) the Patcher. The first process is dedicated only to the expansion or contraction of previous patches of a certain class. The second process is designed to generate new patches through a seeding mechanism. The combination of DINAMICA's transition function presents numerous possibilities with respect to the generation of spatial patterns of change. The user can set parameters to control the size and shape of the simulated patches, such as mean patch size, patch size variance, and isometry. Increasing patch size leads to model with a less fragmented landscape, increasing patch size variance to a more diverse landscape, and setting isometry greater than one leads to the creation of more isometric patches. Additionally, a prune factor multiplies the expected number of cells, based on their spatial probability, that take part in the selection mechanism of new patch nuclei. Therefore, increasing the prune factor allows simulated changes to occur in less likely areas.

Model assessment. The evaluation of the LC prospective map was based on the comparison between the simulated and the observed (true) map. Paegelow and Camacho Olmedo (2005) pointed out that modeled LC maps can be very close to reality due to the persistence of LC over time. The comparison was therefore focused on change areas (both observed and simulated changes). This comparison was done using two approaches: a) the spatial coincidence between modeled and true change; and, b) the spatial pattern of modeled and true change patches.

In order to assess the spatial coincidence between simulated and true changes, we used the method implemented in DINAMICA. The fuzzy similarity test is based on the concept of fuzziness of location, in which a representation of a cell is influenced by the cell itself and by the cells in its neighborhood (Hagen 2003). Two-way comparison was conducted, applying the fuzziness to the simulated and the true maps of change in turn. As random maps tend to score higher, we picked up the minimum fit value from the two-way comparison.

In order to assess the spatial configuration of simulated and true changes, we calculated, for each transition, the amount of change with respect to the map of change probability. For this, a map of susceptibility categories was first obtained by reclassifying susceptibility maps into 10 categories and overlaid with the maps of changes. Additionally, some metrics used to characterize landscape such as the number and the size of the patches (mean and standard deviation) and total edge (mean and standard deviation) were computed.

RESULTS

LUCC modeling

During 1986-94, the main LUCC processes in the study area were the conversion of deciduous mature forest, savanna, amazonian mature forest and woodland savanna to anthropogenic disturbance (Table 1). Only these 4 principal transitions were modeled using as explanatory variables the distance from 1986 urban areas, the distance from roads, the slope, the distance from 1986 disturbance, the elevation and a map created by determining the relative frequency with which different LC categories occurred within the areas that transitioned from 1986 to 1994. The values thus express the likelihood of finding the LC at the pixel in question if this were an area that would transition.

Table 1. Principal Transitions during 1986-94

Acronym	From-to transition category	Area (ha)
Transition 1	Woodland savanna to anthropogenic disturbance	1,897
Transition 2	Amazonian mature forest to anthropogenic disturbance	3,200
Transition 3	Savanna to anthropogenic disturbance	11,378
Transition 4	Deciduous mature forest to anthropogenic disturbance	19,699

Figure 1 shows the map of change susceptibility obtained by the neural network for each one of the four transitions. These maps and the 1986-94 transition matrix were used by both models to build the 1994 simulated LC map. In the case of DINAMICA, various settings of prune factors, patch sizes and isometry was tested. Table 2 shows the parameters which gave the better results.

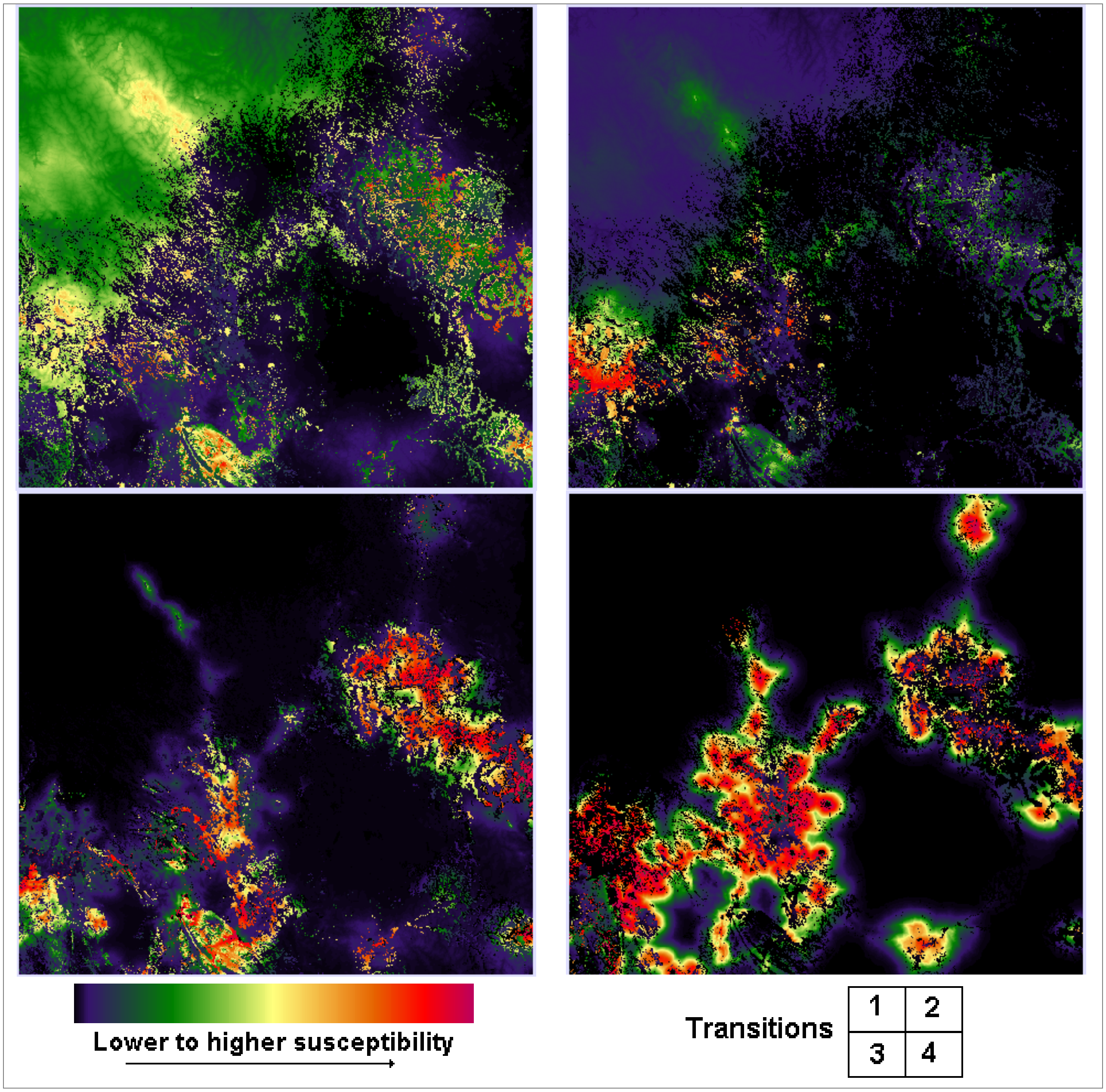

Figure 1. Maps of change susceptibility.

Table 2. DINAMICA Parcher and Expander parameters

Acronym	Expander Prune = 5 Mean / variance / isometry	Parcher Prune = 8 Mean / variance / isometry
Transition 1	2 / 40 / 2	2.5 / 40 / 2
Transition 2	4 / 60 / 2	2.5 / 50 / 2
Transition 3	4 / 90 / 2	3 / 70 / 2
Transition 4	9 / 60 / 2	10 / 60 / 2

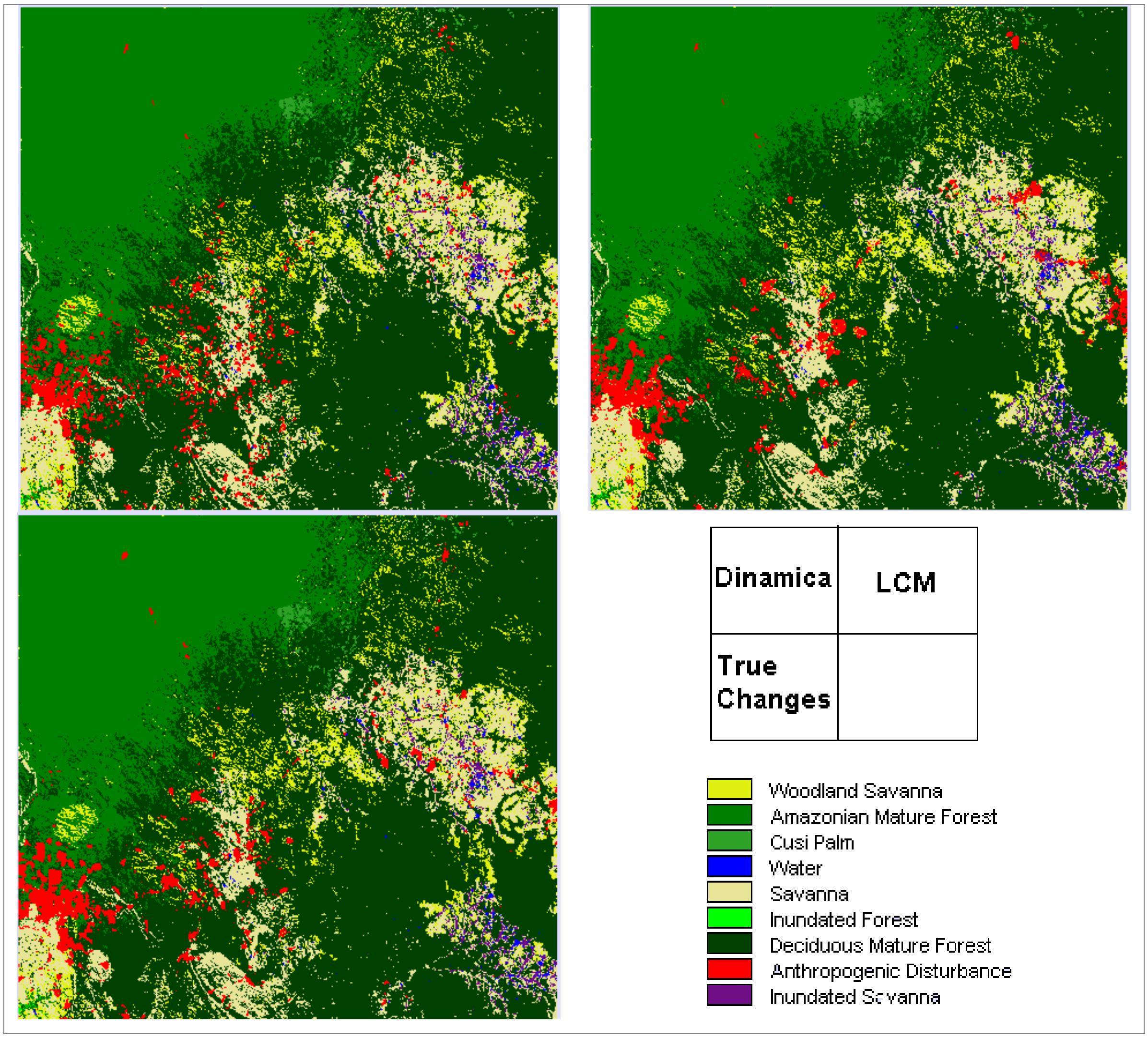

Figure 2. Modeled LC maps by DINAMICA and LCM.

Model Assessment

Figure 3 shows that without fuzzy tolerance (the case corresponding to a "hard" per pixel comparison) and with little tolerance (fuzzy tolerance distance < 1000 m) the coincidence between the true changes and the changes modeled by LCM is much higher than with DINAMICA. This result was expected as LCM tends to collocate simulated changes only in the areas with higher change susceptibility. A comparison between the LCM modeled map and a map obtained by simply thresholding the transition susceptibility maps to the correct quantity of change indicates that 91% of the change pixels coincide in both maps. Since DINAMICA makes an attempt to create patches and simulates change in less likely areas (if the prune factor value is set high), the coincidence between the true and simulated changes is likely to be lower. However, with higher fuzzy tolerance values, DINAMICA presents a higher score because it has some (fuzzy) coincidence of simulated patches located in less likely areas. This does not occur with LCM that restricts the simulated change to the more susceptible areas only. Therefore, DINAMICA presents a better coincidence "as a broad picture" whereas LCM exhibits a better coincidence on a per-pixel comparison or with little fuzzy tolerance.

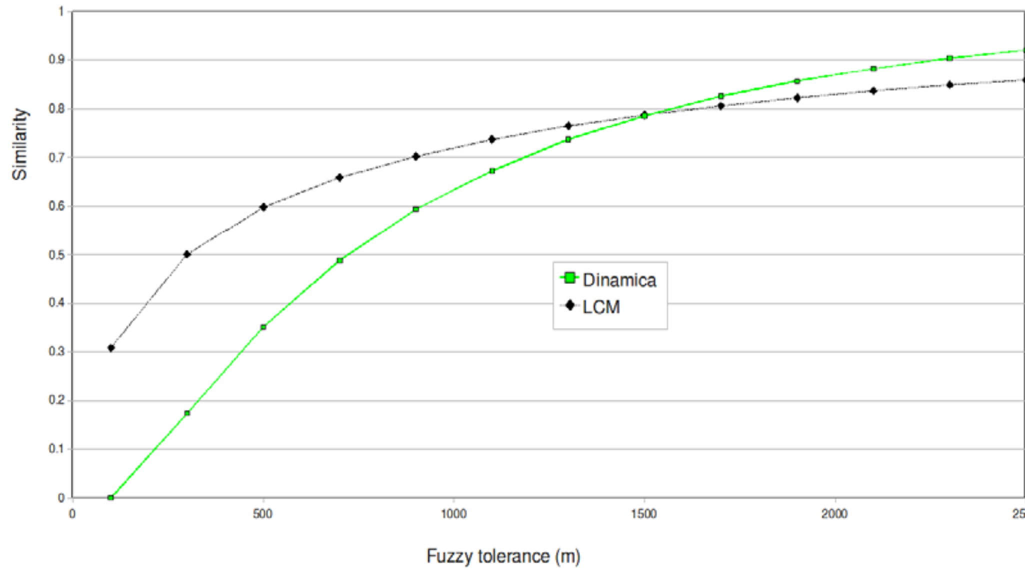

Figure 3. Fuzzy similarity as a function of distance tolerance.

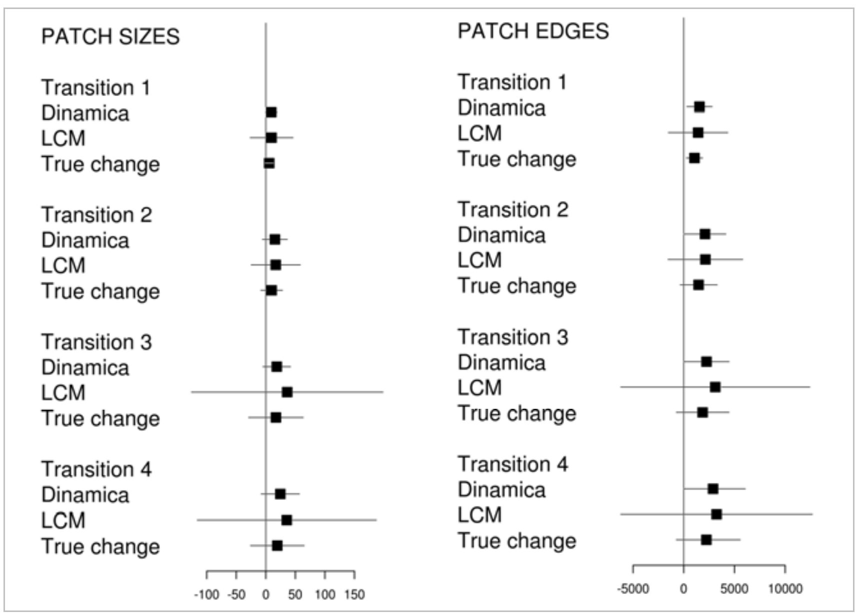

Figure 4. Means and standard deviations of patch sizes and patch edge lengths for each transition.

Table 3. Number of patches for each transitions in the true and modeled maps

Transition	True changes	LCM	DINAMICA
Transition 1	332	192	204
Transition 2	326	190	208
Transition 3	666	299	615
Transition 4	1017	547	804

Figure 4 shows that with DINAMICA it was possible to obtain for each transition simulated patches of change which present broadly the same size as true patches of change. In the case of LCM, the simulation produced some very large patches corresponding to the higher susceptibility areas, resulting in larger values for the mean and the standard deviation of patch size. A similar pattern can be observed for patch edge lengths. However, there are fewer patches on the true change map than on either of the simulated change maps. The map modeled with LCM has 47% fewer patches than the true map (Table 3).

Figure 5 shows that true changes do not occur only in more susceptible areas and that this tendency depends on the transition. Transition 1 occurred mainly in the more susceptible area whereas transition 2 is frequent even in areas with medium susceptibility. The changes simulated by LCM are limited to the areas with higher susceptibility. The setting of the prune factor allowed DINAMICA to generate a map with a distribution of change closer to the observed change.

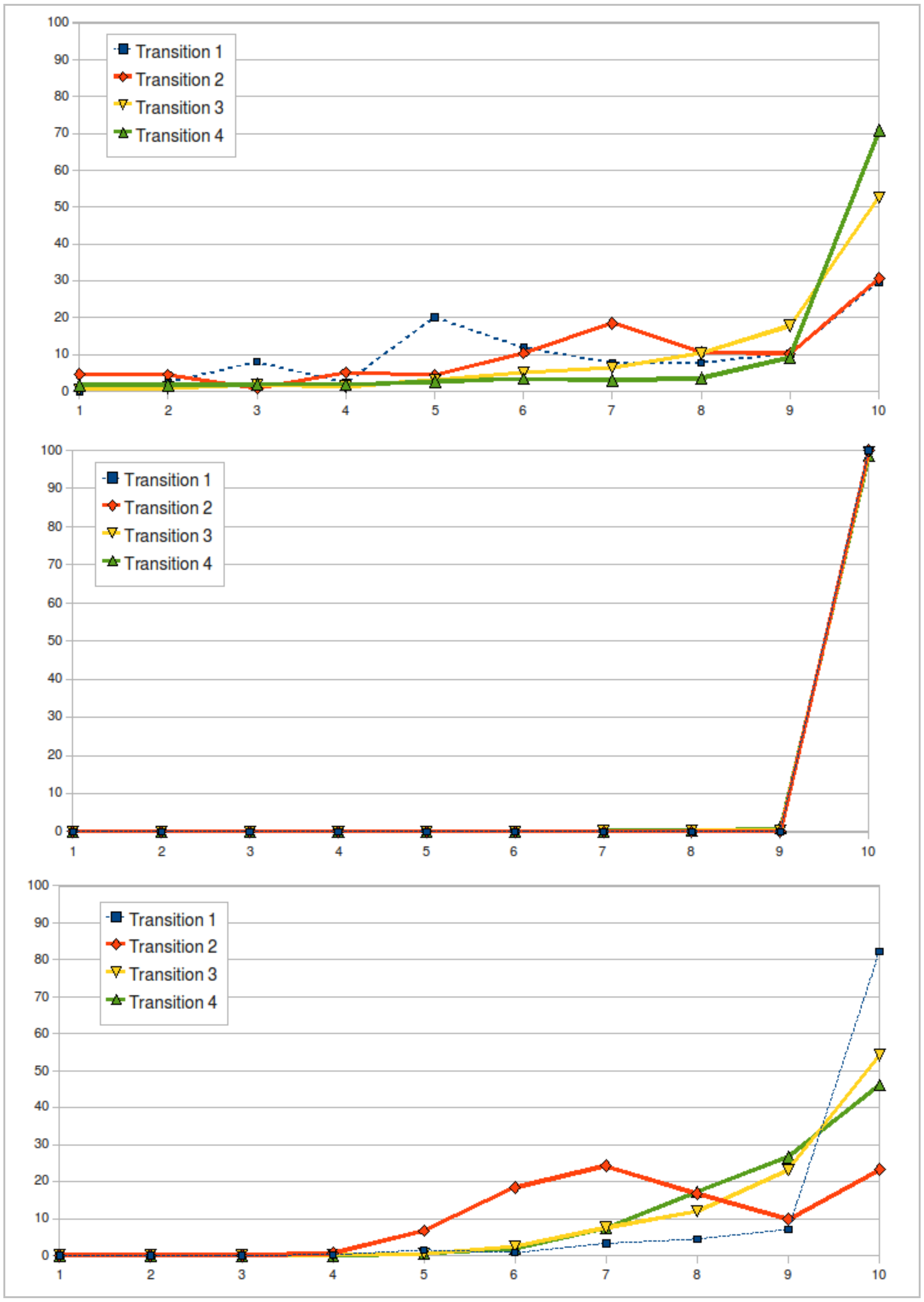

Figure 5. Distribution of change in categories of change susceptibility. Top figure is true changes, middle is change simulated by LCM and bottom simulated by DINAMICA. Categories 1 to 10 represent areas with increasing susceptibility to change.

DISCUSSION AND CONCLUSION

Due to the more sophisticated approach offered by DINAMICA, it was possible to generate more realistic prospective LC maps with respect to landscape pattern. However, the setting of the parameters which control the cellular automata is not a straight forward process because their action depends also on the different susceptibility maps and on the landscape pattern on the previous LC map. The model we developed could have been improved for example by using cellular automata for each transition, and allowing the setting of a prune factor for each transition.

The best manner of producing a prospective LC map, where simulated changes fit better with the true changes, is by thresholding the susceptibility map, because the majority of the changes occur in the more likely locations (with the exception of cases in which the pattern of changes during the calibration and the simulation periods is radically different). The realism of the landscape pattern in the prospective LC map is obtained at the expense of the accuracy of the locations of the change. This is particularly obvious when models simulate the occurrence of changes in unlikely areas. For example, SLEUTH's urban growth model recognizes the possibility of change that occurs without any spatial logic, allowing a small amount of change at randomly assigned locations. The prune factor in DINAMICA also allows the occurring of change in less likely areas, or if the prune factor is set to the maximum value, in random areas. Although such simulations do not generate LC maps with accurate spatial allocation, they can provide valuable information.

Depending on the objective of the modeling, different qualities of prospective maps can be critical. When modeling aims at identifying areas with more propensity to change, it seems logical to look for the best fit between modeled and true changes. However, in this case a fuzzy map that expresses the degree to which locations might potentially change in the future is probably a better option. When modeling aims at producing LC maps that can represent a possible future given a certain scenario, the accuracy of the spatial allocation of change is not necessarily a critical issue. For example, in the assessment of LUCC on biodiversity, it can be important to know that homogeneous forest areas will be perforated by small agriculture fields although the exact location of the fields remain unknown.

However, the common procedures of assessment of prospective LC map are based upon the spatial coincidence of the simulated map and a “true” observed map. Although fuzzy assessment as used in this study provides interesting information, they do not evaluate the landscape pattern.

Therefore, the modeling and the assessment procedures have to be adapted to the critical feature the model has to achieve. When landscape pattern is an important feature, the computing of landscape metrics can provide valuable insights to evaluate the model performance.

ACKNOWLEDGEMENTS

This research has been supported by the *Consejo Nacional de Ciencia y Tecnología* - CONACyT (Sabbatical and posdoctoral stays at the University of California - Santa Barbara of the first and second author respectively) and the *Dirección General de Asuntos del Personal Académico* (DGAPA) at the Universidad National Autónoma de México (additional support for the sabbatical stay).

REFERENCES

Almeida, C.M., M. Batty, M. Vieira Monteiro, G. Câmara, B.S. Soares-Filho, G. Coutinho Cerqueira and C. Lopes Pennachin, 2003. Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation, *Computers, Environment and Urban Systems*, 27:481-509.

Cuevas, G and J-.F., 2008. Land use scenarios: a communication tool with local Communities. In: *Modelling Environmental Dynamics*, Springer-Verlag. pp. 223-246.

de Sherbinin, A., 2002. *CIESIN Thematic Guide to Land Land-Use and Land Land-Cover Change* (LUCC). Center for International Earth Science Information Network (CIESIN) Columbia University Palisades, NY, USA. http://sedac.ciesin.columbia.edu/guides/lu/CIESIN_LUCC_TG.pdf.

Eastman, J.R. 2009. *IDRISI taiga Tutorial*. Accessed in IDRISI 16.05. Worcester, MA: Clark University: 333 p.

Eastman, J.R., M.E. Van Fossen and L.A. Solarzano, 2005. Transition Potential Modeling for Land Cover Change, in *GIS, Spatial Analysis and Modeling*, D. Maguire, M. Goodchild and M. Batty, eds., ESRI Press Redlands, California.

Godoy, M.M.G. and B.S. Soares-Filho, 2008. Modelling intra-urban dynamics in the Savassi neighbourhood, Belo Horizonte city, Brazil. In: *Modelling Environmental Dynamics*. Springer-Verlag, pp. 319-338.
Gontier, M., U. Mörtberg and B. Balfors, in press. Comparing GIS-based habitat models for applications in EIA and SEA, *Environmental Impact Assessment Review*.
Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps, *International Journal of Geographical Information Science*, 17(3): 235–249.
Paegelow, M., M.T. Camacho Olmedo, 2005. Possibilities and limits of prospective GIS land cover modelling - a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain). *International Journal of Geographical Information Science*, 19(6): 697-722.
R Development Core Team, 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria, URL http://www.R-project.org.
Soares-Filho, B.S., C.L. Pennachin and G. Cerqueira, 2002. DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling*, 154(3):217-235.
Soares-Filho, B.S., D. Nepstad, L. Curran, E. Voll, G. Cerqueira, R.A. García, C.A. Ramos, A. Mcdonald, P. Lefebvre and P. Schlesinger, 2006. Modelling conservation in the Amazon basin. *Nature*, 440:520-523.
Texeira, A. M., B.S. Soares Filho, S. Freitas and J.P.W. Metzger, 2009. Modeling Landscape dynamics in the Atlantic Rainforest domain: Implications for conservation. *Forest Ecology and Management*, 257:1219–1230.
Václavík, T. and J. Rogan, 2009. Identifying Trends in Land Use/Land Cover Changes in the Context of Post-Socialist Transformation in Central Europe: A Case Study of the Greater Olomouc Region, Czech Republic, *GIScience & Remote Sensing*, 46(1):54-76.
Veldkamp, A. and E. Lambin, 2001. Predicting land-use change. *Agriculture, Ecosystems and Environment*, 85:1-6.
Xiang, W-N and K.C. Clarke, 2003. The use of scenarios in land-use planning, *Environment and Planning B: Planning and Design*, 30(6) 885-909.