iEMSs 2016 Conference

Environmental modelling and software for supporting a sustainable future

Proceedings | Volume 5 | Pages 1275-1335

8th International Congress on Environmental Modeling and Software (iEMSs)

July 10–14, 2016
Toulouse, France

Proceedings of the 8th International Congress on Environmental Modelling and Software (iEMSs)
July 10-14, 2016, Toulouse, FRANCE.

How to cite the full proceedings:

Sauvage, S., Sánchez-Pérez, J.M., Rizzoli, A.E. (Eds.), 2016. Proceedings of the 8th International Congress on Environmental Modelling and Software, July 10-14, Toulouse, FRANCE. ISBN: 978-88-9035-745-9

How to cite an individual paper:

Author, A., Author, B., Author, C…, 2016. This is the title of your paper. In: Sauvage, S., Sánchez-Pérez, J.M., Rizzoli, A.E. (Eds.), 2016. Proceedings of the 8th International Congress on Environmental Modelling and Software, July 10-14, Toulouse, FRANCE. ISBN: 978-88-9035-745-9

Peer Review:

Each paper has been peer reviewed by at least two independent reviewers with possible outcomes of reject, revise, and accept.

International Environmental Modelling and Software Society (iEMSs)
8th International Congress on Environmental Modelling and Software
Toulouse, France, Sabine Sauvage, José-Miguel Sánchez-Pérez, Andrea Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2016-proceedings

Combining Geographically Weighted and pattern-based models to simulate deforestation processes

Jean-François Mas
Centro de Investigaciones en Geografía Ambiental, Universidad Nacional Autónoma de México - jfmas@ciga.unam.mx

Abstract: In this study, two approaches were compared to carry out spatially explicit deforestation model using data driven pattern-based models. In the first approach, model was trained globally: A single matrix of change and set of weights of evidence was obtained taking into account the entire study area. Therefore, the relationship between change potential and expected amount of change and the drivers of change was established for the entire study area and used to simulated deforestation process. In the second approach, sub-regions which present different patterns of deforestation were first identified using a Geographically Weighted Regression model. Then model was trained and deforestation was simulated independently for each one of the region. Performance of both approaches was assessed through the comparison between simulated and true deforestation using a fuzzy coincidence index. The coincidence obtained by the region-based model was sightly superior to the global model. The coupling between spatial data mining techniques as Geographically Weighted Regression models can contribute to help understanding the land changes as complex processes involving both social and natural systems and increasingly develop models which take into account the processes of change and not only the patterns.

Keywords: *Geographically Weighted Regression; deforestation drivers; Mexico, Land change models*.

1 INTRODUCTION

Spatially explicit land use / cover change (LUCC) models aim at simulating the patterns of change on the landscape (Paegelow et al., 2013). Modelling is often based on a inductive pattern-based approach: LUCC is modelled empirically using past LUCC spatial distribution and rate to develop a mathematical model that estimates the change potential as a function of a set of explanatory spatial variables and the expected amount of change (Paegelow and Olmedo, 2005; Mas et al., 2014).

However, the relationship between change potential and expected amount of change, in one hand, and the drivers of change, on the other hand, is often established for the entire study area. On other words, the model is based on the assumption that change processes are stationary: the same drivers have the same effects on LUCC over the whole area. In the case of large or heterogeneous regions, different patterns of LUCC can be expected, that is LUCC and its drivers are usually location-dependent. The use of local exploratory models such as Geographically Weighted Regression (GWR) models, enables examining spatially varying relationships between a phenomenon and related factors. This study aims at evaluating the use of GWR models, to examine spatially varying relationships between the rate of deforestation, and related drivers and split the study area into subregions which changes are independently modelled.

2 STUDY AREA

The State of Michoacán (Figure 1), is located in west-central Mexico and encompasses about 60,000 km^2. It is one of the most diverse State of Mexico, a megadiverse country, and presents different types of tropical and temperate forests. It is also leader in agricultural production and presents important processes of land use/cover change (Bocco et al., 2001; Mas et al., submitted).

Figure 1 - Study area location

3 MATERIALS

3.1 Spatial database

In order to elaborate the spatial database, we used the following data:

- Land use / cover (LUC) maps of the state of Michoacán dated 2004, 2007 and 2014 elaborated at scale 1/100,000. The maps were elaborated using a procedure of semi-automatic interdependent classification (Mas and González, 2015), which enables the monitoring of LUCC (http://www.ciga.unam.mx/wrappers/proyectoActual/monitoreo/).
- Maps of ancillary data (digital elevation model, slope, roads maps, human settlements).
- Socio-economic data from the National Institute of Statistics and Geography (INEGI for its Spanish acronym) organized by settlements (Population census for 2005 and 2010).

3.2 Software programs

The R packages GWmodel (Lu et al., 2014; Gollini et al., 2015) and spgwr (Bivand et al., 2015) were used to carry out the spatial analysis and fit the GWR models. For the elaboration of the deforestation model, I used the freeware Dinamica EGO (hereafter DINAMICA, http://csr.ufmg.br/dinamica/). DINAMICA has been applied to a variety of studies, in particular models of tropical deforestation (Soares-Filho et al., 2002, 2006, 2013; Cuevas and Mas, 2008). We chose it due to its flexibility, its computing efficiency and the availability of tools designed for LUCC modelling such as Markov projection, automata cellular, regionalization of the models among others (Mas et al., 2014). Graphics were elaborated with R (R Development Core Team, 2014) and QGIS (QGIS Development Team, 2016).

4 METHODS

4.1 Geographically Weighted models

Geographically Weighted (GW) models, as GW statistics and Geographically Weighted Regression (GWR), are local spatial statistical techniques for exploring spatial non-stationarity that have been applied in the social, health and environmental sciences (Fotheringham et al., 2002). In GWR, local modelling of spatial relationships is done by fitting regression models using local data weighted by distance. For this, regression parameters are estimated using a weighting function that assigns larger weights to closer locations. Different from the usual global regression model, which produces a single regression equation by summarizing the overall relationships among the explanatory and dependent variables for the entire study area, GWR takes into account the spatial variation in the relationships among variables. This method enables the user elaborating maps that present the spatial distribution of the regression parameters estimates and exploring spatial non-stationarity. Fotheringham et al., (2002) provide with a full description of GWR.

In the present study, a random sample of points was used to collect information from the map of deforestation (period 2004-2007) and from the maps of drivers. The sample was stratified in order to have a sufficient number of observations in deforested areas. As the dependent variable was binary (deforestation / forest permanence), a logistic local regression was fitted. As explanatory variables, we selected only a reduced number of variables in order to avoid collinearity between explanatory variables and to have a simple model easier to interpret. These variables were elevation, slope, distance to agriculture, distance to roads, distance to rural settlements, population density and social margination.

First, we carried out the computing of local correlation between deforestation and the explanatory variables. As a following step, we carried out a GWR using the variables with higher correlation or variation of correlation over space. The values of the GWR parameters associated to each explanatory variable and the constant (y-intercept) were standardized and classified using a clustering algorithm. This enables identifying subregions where the same variables have a similar effect on the rate and the spatial distribution of deforestation. The different regions present different relationships between drivers and rate of deforestation and therefore different patterns of change.

4.1 Deforestation model

The deforestation model was based on a inductive pattern-based approach: deforestation was modelled empirically using past (2004-2007) deforestation spatial distribution and rate to estimate the change potential as a function of a set of explanatory spatial variables and the expected amount of change during the following period (2007-2014). In the case of the region-based model, the sub-regions identified in the previous step were used to split the deforestation model. Each subregion was trained separately. Transition probability maps were elaborated using the weight of evidence method, a Bayesian method of conditional probability. For each region, a Markov matrix was also fitted. In order to evaluate the sub-region based model, a global model was produced using the same methods and data but based on the entire area.

Then deforestation was simulated from 2007 to 2014 using the 2007 as initial LUC map and both, global and region-based models. The matrices of Markov were used to calculate expected annual deforested area and the set of weights of evidence to compute the probability of change.

In order to assess the models we used a fuzzy similarity index to compare the outputs of the models and the true changes during 2007-2014. The fuzzy similarity index is based on the concept of fuzziness of location: the coincidence between simulated and true deforestation areas is not restricted to a cell-by-cell overlay but includes the cells in a neighbourhood (Hagen, 2003).

5 RESULTS

5.1 Geographically Weighted models

LUC maps classification scheme was simplified to seven categories: 1) irrigated agriculture, 2) rain-fed agriculture, 3) permanent agriculture (orchards), 4) pasture, 5) temperate forest, 6) tropical forest and, 7) other categories. Two maps of change were produced, a map of LUCC taking into account all the transitions from the two forest categories to agriculture (3 categories) or pasture land and a simplified binary map of deforestation.

GW local coefficient of correlation between between deforestation and each explanatory variable was calculated using a systematic grid of points as analysis points. Figure 2 shows the local coefficient of correlation between deforestation and four explanatory variables (elevation, slope, margination and distance to rural settlements). It can be observed that the slope, the margination index and the distance to rural settlements exhibit a negative local coefficient of correlation over almost the entire study area, indicating that deforestation tend to be associated with areas with moderate slope, near to rural settlements and less marginalized. The strength of this relation varies over space, In some area the coefficient of correlation is positive. Elevation has a contrasting relationship with deforestation, with both positive and negative coefficient of correlation depending on the region. A GWR was carried out and the regression points were classified using a clustering algorithm to group sites with similar relationship between deforestation with explanatory variables. Figure 3 shows the four regions identified using the deforestation patterns observed during 2004-2007. These regions presents important similarity with economical and physiographic existing maps.

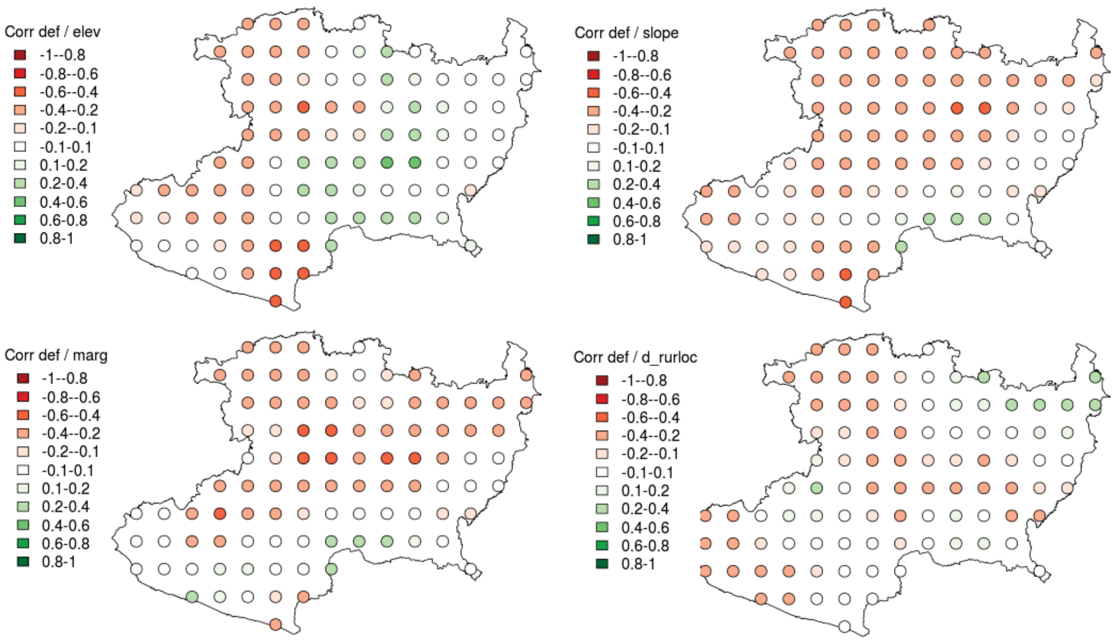

Figure 2 - GW local coefficient of correlation between deforestation and elevation, slope, margination and distance to rural settlements)

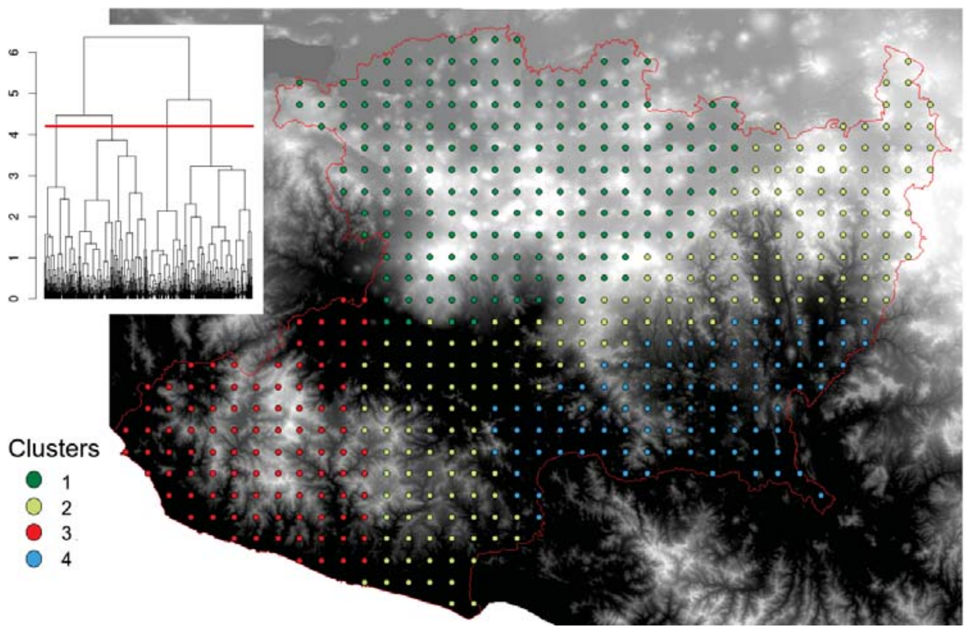

Figure 3 - Dendrogram used to produce the four clusters. Background is the DEM

5.2 Models of deforestation

Two models of deforestation were elaborated with the same explanatory variables. The global model used only one Markov matrix and one set of weights of evidence for the entire study area. The region-based model uses a matrix and a set of weights of evidence for each one of the four regions identified by the GWR model. The simulated LC maps for 2014 were compared to the true LC 2014 map. As shown in figure 4, the fuzzy coincidence is sightly larger for the region-based model in comparison with the global one.

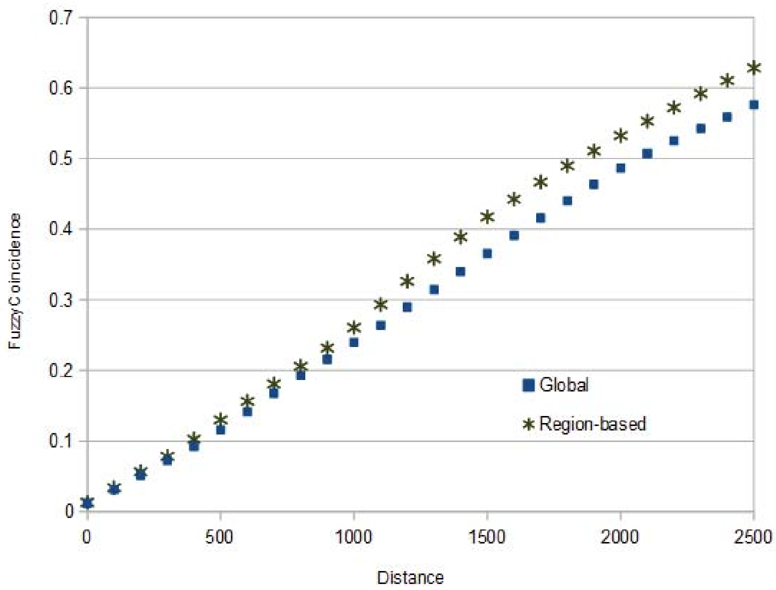

Figure 4 - Fuzzy coincidence between simulated and observed deforestation as a function of tolerance distance

5 DISCUSSION AND CONCLUDING REMARKS

Both models were able to identify the regions more likely to be deforested. The coincidence with little or no fuzzy tolerance was null, which means that the models were not able to predict the exact location of cleared areas (which seems impossible) but rather to identify roughly the areas of change. The coincidence between simulated and true change was approximately 60% with tolerance of 2500 m, which seems satisfactory taking into account that there is relatively little change (in comparison with forest area).

Both models overestimated largely the quantity of change: they simulated more than 40,000 ha of deforestation when less than 20,000 really occurred. This is due to the fact that they are based on the assumption of temporal stationarity. The matrix of Markov projected the simulated area of each transition based on the rate of change observed between 2004 and 2007. As there was a considerable diminution of the rates of deforestation during the second period, the projection overestimated the simulated areas of deforestation.

The performance of the region-based model was slightly superior to the global model. However in this exercise, I did not take full advantage of the local analysis because all the procedures were data driven: The Matrix matrix and the weights of evidence were calculated automatically and were not edited using expert knowledge (Mas et al., 2014). In fact, the GWR analysis enables users to identify sub-regions with different patterns of change, help understanding the complex social context of land change (e.g. role of the different drivers) and give insights to elaborate models increasingly process-based and able to link processes in social and natural systems (NRC, 2014). The use of local spatial data analysis techniques, as GWR, produces material which can facilitate the discussion between experts from different fields and creates opportunities to improve the LUCC models by evolving from data-driven to more knowledge-based models.

ACKNOWLEDGEMENTS

This study was supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT) and the Secretaría de Educación Pública through the project ¿Puede la modelacin espacial ayudarnos a entender los procesos de cambio de cobertura/uso del suelo y de degradacin ambiental? - Fondos SEP-CONACyT 178816.

REFERENCES

Bivand, R., Yu, D., Nakaya, T., García-López, M.A., 2005. SPGWR: Geographically weighted regression. Available online at: https://cran.r-project.org/ (Accessed Abril 15, 2016).

Bocco, G., Mendoza, M., Masera, O.R.. 2001. La dinámica del cambio del uso del suelo en Michoacán: Una propuesta metodológica para el estudio de los procesos de deforestación. Investigaciones geográficas, (44),18-36.

Cuevas, G., Mas, J.F., 2008. Land use scenarios: a communication tool. In: Paegelow and Camacho-Olmedo (eds), Modelling Environmental Dynamics, Springer Berlin Heidelberg, Berlin, 223-246.

Fotheringham, S.A., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression: the analysis of spatially varying relationships. John Wiley & Sons.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., Harris, P., 2015. GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. Journal of Statistical Software, 63(17), 1-50.

Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps. International Journal of Geographic Information Science, 17 (3), 235-249.

INEGI, 2005. Conteo de población y vivienda 2005. Indicadores del censo de Población y vivienda. Ed. INEGI, México.

INEGI, 2010. Censo de Población y Vivienda 2010, Ed. INEGI, México.

Lu, B., Harris, P., Charlton, M., Brunsdon, C., 2014. The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models, Geo-spatial Information Science, 17:2, 85-101.

Mas, J.F., González, R., 2015. Change detection and land use / land cover database updating using image segmentation, GIS analysis and visual interpretation. ISPRS Geospatial Week 2015, La Grande Motte, France, 28th September - 2nd October 2015. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-3/W3, 61-65.

Mas, J.-F., B. Soares-Filho, Rodrigues, H. 2015. Calibrating cellular automata of land use/cover change models using a genetic algorithmInt. ISPRS Geospatial Week 2015, La Grande Motte, France, 28th September - 2nd October 2015. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-3/W3, 67-70.

Mas, J.F., Kolb, M., Paegelow, M., Camacho Olmedo, M.T., Houet, T., 2014. Inductive pattern-based land use/cover change models: A comparison of four software packages. Environmental Modelling & Software 51, pp. 94-111.

Mas, J.F., González, R., Lemoine Rodríguez, R., López Sánchez, J., Piña Garduño. A., Herrera Flores, E, submitted. Evaluación de las tasas de deforestación en Michoacán (2004–2014). Revista Chapingo Serie Ciencias Forestales y del Ambiente.

NRC (National Research Council), 2014. Advancing Land Change Modeling: Opportunities and Research Requirements. Committee on Needs and Research Requirements for Land Change Modeling; Geographical Sciences Committee; Board on Earth Sciences and Resources; Division on Earth and Life Studies; National Academies Press, 152 p.

Paegelow, M., Camacho Olmedo, M.T., Mas, J.F., Houet, T., Pontius Jr., R.G., 2013. Land change modelling: moving beyond projections. International Journal of Geographical Information Science 27:1691-1695.

Paegelow, M., Camacho Olmedo, M.T., 2005. Possibilities and limits of prospective GIS land cover modellinga compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain). International Journal of Geographical Information Science 19(6), pp. 697-722.

QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org (Accessed 9 April 2016).

R Development Core Team, 2014. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org (Accessed 9 April 2016).

Soares-Filho, B.S., Coutinho Cerqueira, G., Lopes Pennachin, C., 2002. DINAMICA - A stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecological Modelling 154, 217-235.

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 440, 520–523.

Soares-Filho, B., Rodrigues, H., Follador, M., 2013. A hybrid analytical-heuristic method for calibrating land-use change models. Environmental Modelling and Software 43, 80-87.