Chapter 36
A Short Presentation of the Land Change Modeler (LCM)

J.R. Eastman and J. Toledano

Abstract The Land Change Modeler is a land change projection tool for land planning. It uses historical land cover change to empirically model the relationship between land cover transitions and explanatory variables to map future scenarios of change.

1 Introduction

The Land Change Modeler (LCM) was developed (Eastman 2006) as an empirically parameterized land change projection tool to support a wide range of planning activities. Based on an analysis of historical land cover change, the system develops an empirical model of the relationship between land cover transitions and a set of explanatory variables. Mappings of future change are then based on this empirical relationship and a projection of quantity derived from a Markov Chain. The result is a business-as-usual (BAU) projection of change without subjective intervention. It is designed to support applications with strict BAU baseline needs such as REDD (Reducing Emissions from Deforestation and forest Degradation) climate mitigation projects.

2 Description of the Methods Implemented in the Model

At present, three separate empirical model development tools are provided in LCM: a Multi-Layer Perceptron neural network (MLP), Logistic Regression (LR) and SimWeight (SW). The MLP procedure is the default, and is the most mature and

J.R. Eastman (✉) · J. Toledano
Clark Labs, Clark University, Worcester, USA
e-mail: reastman@clarku.edu
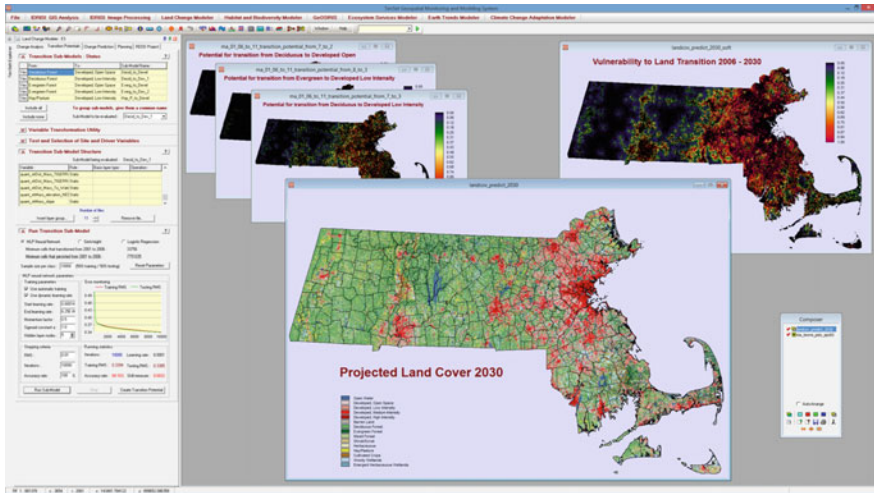
Fig. 1 In LCM, the process of land change modeling is organized into major stages embodied by tabs in the interface. The most important stage is that in which empirical models are developed relating historical changes to explanatory variables. In this example, the default Multi-Layer Perceptron Neural Network is used to develop transition potential maps (small maps, *upper left*)— empirically derived statements of the potential of land to undergo specific transitions. These are used in the subsequent change prediction tab to generate both future scenarios (large image, *center*) and maps of vulnerability to change (*upper left*)

primary focus of LCM (Fig. 1). MLP is also the only procedure that can model multiple transitions at the same time. Logistic Regression is provided primarily for pedagogic reasons while SimWeight is an experimental machine learning procedure based on a K-Nearest Neighbor variant (Sangermano et al. 2010). In each case, analysis of two land cover layers in the recent past is used to train and evaluate the model.

2.1 Training

For the Multi-Layer Perceptron, LCM examines each of the transitions over the historical period to determine the number of pixels that went through the transition being modeled (change pixels) and the number that were eligible to, but which did not (persistence pixels). The user is then required to specify the sample sizes to use for training the model. For MLP and SimWeight, equal-sized samples of change and persistence are required. However, the default sample sizes are very different— 10,000 for MLP and 1000 for SimWeight. The difference relates to how they are used—for iterative learning in the case of MLP and characterization for SimWeight. For Logistic Regression, the sample chosen is proportional to the relative number of change and persistence pixels for the individual transition being modeled. The user is able to indicate the sampling proportion (the default is 10%) and the method of spatial sampling—stratified random sampling (the default) or systematic.

2.2   Simulation

In LCM, the simulation proceeds in three stages. The first is the development of transition potentials—mappings of the readiness of land to go through each of the transitions under consideration. The second is the estimation of the expected quantity of change and the third is the spatial allocation of the estimated change based on the transition potentials.

Transition Potentials
A transition potential is a continuous value from 0–1 that expresses the relative potential of a pixel to transition from one state to another. The metric varies from one empirical modeling procedure to another but in the final stage of the simulation, only the relative value of the metric matters.

If MLP is used as the modeling procedure, the transition potential is the activation level of the output neurons which represents the posterior probability of transition under an assumption of equal probability of change/persistence. With Logistic Regression, the transition potential is the probability of change assuming an identical quantity to that which transitioned during the historical period. With SimWeight, the value is unitless, but monotonic with the posterior probability of transition.

Estimation of the Quantity of Change
In the second stage, a Markov Chain analysis is used to determine the quantity of change for the forecast date selected. A Markov Chain assumes that the rate of change (but not the quantity of change) remains constant over time. The calculation proceeds (Eastman 2014) by first computing a cross tabulation of transitions between the land cover maps for the two historical dates. From this, the basic transition probability matrix (X) is calculated. If the date being projected forward is an even multiple of the training period, then the new transition probability matrix is calculated through a simple powering of the base matrix (Kemeny and Snell 1976). For example, if the training period is from 2002 to 2011 (9 years) then the transition probability matrix for 2020 from 2011 (9 years forward) is X^1, for 2019 (18 years forward) is X^2, for 2047 (36 years forward) is X^4, and so on. However, if the projected time period is in between even multiples of the training period, then the power rule is used to generate 3 transition matrices that envelop the projection time period (if the 3 time periods are times A, B and C, the period to be interpolated will be between A and B). The three values at each cell in the transition probability matrix are then fed into a quadratic regression (thus there will be a separate regression for each transition probability matrix cell). Given that a quadratic regression ($Y = a + b_1X + b_2X^2$) has 3 unknowns and we have three data points, it yields a perfect fit. This equation is then used to interpolate the unknown transition probability. From these transition probabilities, the projected quantity of change is determined for each transition being modeled.

Spatial Allocation

Given a set of transition potential maps and the projected quantity of change for each transition, LCM then allocates change based on a greedy selection algorithm. The greedy selection is based on the simple assumption that the areas with the highest transition potential will always transition first. Because a single pixel may be selected for multiple transitions, a competitive strategy is used whereby it will be assigned to the transition with the highest marginal transition potential. This will lead to some transitions being allocated less than the expected quantity of change. Thus the procedure iterates through a process of selecting pixels with lower transition potentials until all transitions achieve their required quantity.

LCM recognizes that some explanatory variables may be based on land cover, and thus change as transitions progress. For example, a model focused on deforestation may use a variable of proximity to existing agriculture. As agriculture expands, this variable will constantly be changing. Such variables are termed *dynamic* as opposed to *static*. Thus the user has the ability to predict in stages with dynamic variables being automatically recomputed at each stage. The procedure chosen for re-computation can be as simple as a single distance calculation to a complex macro. For example, a user may have empirically determined the potential for transition based on the age of the forest post-agriculture, and thus may use the macro option to add time and then re-compute the transition potential at each step.

2.3   Validation

Validation is handled differently for each of the empirical modeling procedures. For MLP, half of the training data are reserved for validation. Validation is a critical component of the training process. At each stage in the training, learning is refined with one half of the data and the quality of the model is assessed by comparison with the other half. Accuracy and model skill over all transitions and persistences combined are dynamically reported during the training process. Model skill is reported as a Heidke Skill Score (Heidke 1926), also known as Kappa (Cohen 1960), which ranges from −1 to +1 with 0 indicating a skill no better than random allocation. At the end of the empirical modeling procedure, LCM provides a detailed accounting of accuracy and skill for each transition and each persistence category. It also provides a wealth of information about the contribution of each variable to the model including a backwards stepwise assessment that allows for a very easy determination of the most parsimonious model.

For SimWeight, again, 50% of the training data are reserved for validation. From these, a Peirce Skill Score (Joliffe and Stephenson 2003) is evaluated—a value similar in nature to a Heidke Skill Score in that it ranges from −1 to +1 with 0 representing the point where the hit rate and false alarm rate are equal. SimWeight also reports the relevance of variables by examining the variability of a variable within historical samples of transition relative to all areas (Sangermano et al. 2010).

For Logistic Regression, validation is handled by a Goodness of Fit measure which expresses the degree to which the fitted values of the modelled regression match the training data. Relevance of variables is assessed by the slope coefficients and t-score values, although use of standardized variables is recommended for this purpose.

2.4   REDD

An important application of this kind of empirically-modeled projection tool is the climate change mitigation strategy known as REDD—Reducing Emissions from Deforestation and forest Degradation. To meet the special needs of these programs, LCM provides a special set of tools for the development of REDD projects. Tools provide for the definition of the project and leakage areas, specification of carbon pools to be considered, method of calculation and carbon density in the evaluation of CO_2 emissions. Non-CO_2 emissions are also considered. Leakage, success and effectiveness rates are then specified for each of the reporting stages. In the end, 19 tables are produced following the BioCarbon Fund methodology. However, these are easily re-formatted into any of the prevailing approved methodologies.

3   Applications

Land Change Modeler has been evaluated and applied across many disciplines in varied geographic areas since its release in 2007. It has been evaluated against other land change methods (Eastman et al. 2005; Fuller et al. 2011; Mas et al. 2014; Paegelow and Camacho Olmedo 2008). LCM has been applied to forest monitoring and deforestation (Khoi and Murayama 2011; Valle Jr et al. 2012) and its impacts on biomass (Eckert et al. 2011; Fuller et al. 2011; Saha et al. 2013), biodiversity (Dean and Salim 2012; Uddin et al. 2015) and species habitat impact assessment (Fuller et al. 2012). LCM has even been employed to model post-socialist land change in Eastern Europe (Václavík and Rogan 2010). LCM is an accepted tool by the Verified Carbon Standard (VCS 2014) and is extensively used in REDD (Areendran et al. 2013; Centro de Conservación 2012; Kim and Newell 2015) and REDD+ project planning (Moore et al. 2011; Sangermano et al. 2012; Scheyvens et al. 2014; VCS 2014).

4   Final Considerations

LCM has now been in public use for more than a decade. It was commissioned by Conservation International and the conservation community still constitutes the largest body of users. Companion software components have been developed to

work closely with LCM including the Habitat and Biodiversity Modeler and the Ecosystem Services Modeler. Development of LCM continues with two new machine learning procedures (Weighted Normalized Likelihoods and a Support Vector Machine) currently in testing. Additionally, a cloud-based implementation is currently under development. Given the pace of anthropogenic land conversion, the ability to develop defensible and skillful land change models is of critical importance.

References

Areendran G, Raj K, Mazumdar S, Puri K, Shah B, Mukerjee R, Medhi K (2013) Modelling REDD+ baselines using mapping technologies: a pilot study from Balpakram-Baghmara Landscape (BBL) in Meghalaya, India. Int J Geoinf 9(1):61–71

Centro de Conservación, Investigación y Manejo de Areas Naturales (2012) Cordillera Azul National Park REDD Project, Peru

Cohen JA (1960) Coefficient of agreement for nominal scales. Educ Psychol Measur 20(1):37–46

Dean A, Salim A (2012) Technical background material, Kalimantan InVest and LCM modeling. Heart of Borneo: investing in nature for a green economy. WWF HoB Global Initiative

Eastman JR (2006) IDRISI 15.0: The Andes edn. Clark University, Worcester MA

Eastman JR (2014) TerrSet geospatial monitoring and modeling system. Clark University, Worcester, MA

Eastman JR, Solorzano L, Van Fossen M (2005) Transition potential modeling for land-cover change. In: Maguire DJ, Batty M, Goodchild MF (eds) GIS, spatial analysis and modeling. ESRI Press, Redlands, CA, pp 357–385

Eckert S, Ratsimba HR, Rakotondrasoa LO, Rajoelison LG, Ehrensperger A (2011) Deforestation and forest degradation monitoring and assessment of biomass and carbon stock of lowland rainforest in the Analanjirofo Region, Madagascar. For Ecol Manag 262(11):1996–2007

Fuller DO, Hardiono M, Meijaard E (2011) Deforestation projections for carbon-rich peat swamp forests of central Kalimantan, Indonesia. Environ Manag 48(3):436–447

Fuller DO, Parenti M, Hassan AN, Beier JC (2012) Linking land cover and species distribution models to project potential ranges of malaria vectors: an example using Anopheles arabiensis in Sudan and Upper Egypt. Malaria J 11(1):264

Guidolini JF, Batista Pedroso L, Neves Araújo MV (2013) Análise Temporal Do Uso e Ocupação Do Solo Na Microbacia Do Ribeirão Do Feijão, Município De São Carlos–SP, Entre Os Anos De 2005 e 2011. In: Anais XVI Simpósio Brasileiro De Sensoriamento Remoto–SBSR, Foz Do Iguaçu, PR, Brasil, 13 a 18 De Abril, 4503–4509. INPE 2013

Heidke P (1926) Berechnung des erfolges und der gute der windstarkvorhersagen im sturm-warnungsdienst. Geogr Ann 8:301–349

Houet T, Vacquié L, Sheeren D (2015) Evaluating the spatial uncertainty of future land abandonment in a mountainous valley (Vicdessos, Pyrenees–France): insights from model parameterization and experiments. J Mt Sci 12(5):1095–1112

Joliffe IT, Stephenson DB (2003) Forecast verification: a practitioners guide in atmospheric science. Wiley, Chichester, West Sussex, England

Kemeny JG, Snell JL (1976) Finite markov chains. Springer, New York

Khoi DD, Murayama Y (2011) Modeling deforestation using a neural network-Markov model. In: Murayama Y, Thapa RB (eds) Spatial analysis and modeling in geographical transformation process: GIS-based applications. Springer, New York, pp 169–192

Kim OS, Newell JP (2015) The 'Geographic Emission Benchmark' model: a baseline approach to measuring emissions associated with deforestation and forest degradation. J Land Use Sci 10 (4):466–489

Mas JF, Kolb M, Paegelow M, Camacho Olmedo MT, Houet T (2014) Modelling land use/cover changes: a comparison of conceptual approaches and softwares. Environ Model Softw 51:94–111

Moore C, Eickhoff G, Ferrand J, Khiewvongphachan X (2011) Technical Feasibility Assessment of the Nam Phui National Protected Area REDD+ Project in Lao PDR. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). http://theredddesk.org/sites/default/files/resources/pdf/2013/giz-clipad-technical-feasibiltiy-study-nam-phui-npa-lao-pdr-v1.4.pdf. Accessed 5 Dec 2016

Paegelow M, Camacho Olmedo MT (eds) (2008) Modelling environmental dynamics. Springer, Berlin

Saha GC, Paul SS, Li J, Hirshfield F, Sui J (2012) Investigation of land-use change and groundwater–surface water interaction in the Kiskatinaw River Watershed, Northeastern British Columbia, Parts of NTS. Geoscience BC Summary of Activities 2012. Geoscience BC, Report 2013-1:139–148

Sangermano F, Eastman JR, Zhu H (2010) Similarity weighted instance based learning for the generation of transition potentials in land change modeling. Trans GIS 14(5):569–580

Sangermano F, Toledano J, Eastman JR (2012) Land cover change in the Bolivian Amazon and its implications for REDD+ and endemic biodiversity. Landsc Ecol 27(4):571–584

Scheyvens H, Gilby S, Yamanoshita M, Fujisaki T, Kawasaki J (2014) Snapshots of REDD+ Project Designs. Institute for Global Environmental Strategies (IGES) https://pub.iges.or.jp/pub/redd-projects-snapshots-redd-project-designs. Accessed 5 Dec 2016

Uddin K, Chaudhary S, Chettri N, Kotru R, Murthy M, Chaudhary RP, Ning W, Shrestha SM, Gautam SK (2015) The changing land cover and fragmenting forest on the Roof of the World: A case study in Nepal's Kailash Sacred Landscape. Landsc Urban Plan 141:1–10

Valle RFD Jr, Evangelista Siqueira H, Ferreira Guidolini J, Abdala VL, Ferreira Machado M (2012) Diagnóstico de Mudanças e Peristência de Ocupação do Solo Entre 1978 e 2011 No IFTM-Campus Uberaba, Utilizando o 'Land Change Modeler (LCM)'. Enciclopédia Biosfera 8 (15):672–681

Václavík T, Rogan J (2010) 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. GISci Remote Sens 46(1):54–76

Verified Carbon Standard (2014) Carmen del Darien (CDD) REDD+ Project. 2014. 3rd ed. http://database.v-c-s.org/sites/v-c-s.org/files/CDD%20REDD+%20Project%20Description%20v5.9.pdf. Accessed 5 Dec 2016