Improving Business-as-Usual Scenarios in Land Change Modelling by Extending the Calibration Period and Integrating Demographic Data

Romain Mejean, Martin Paegelow, Mehdi Saqalli, Doryan Kaced

Improving Business-as-Usual Scenarios in Land Change Modelling by Extending the Calibration Period and Integrating Demographic Data

Romain Mejean, Martin Paegelow, Mehdi Saqalli, Doryan Kaced

Abstract Land use and land cover change (LUCC) models are increasingly being used to anticipate the future of territories, particularly through the prospective scenario method. In the case of so-called trend or Business-as-Usual (BAU) scenarios, the aim is to observe the current dynamics and to extend them into the future. However, as they are implemented as baseline simulation in most current software packages, BAU scenarios are calibrated from a training period built from only two dates. We argue that this limits the quantitative estimation of future change intensity, and we illustrate it from a simple model of deforestation in Northern Ecuadorian Amazon using the Land Change Modeler (LCM) software package. This paper proposes a contribution to improve BAU scenarios calibration by mainly two enhancements: taking into account a longer calibration period for estimating change quantities and the integration of thematic data in change probabilities matrices. We thus demonstrate the need to exceed the linear construction of BAU scenarios as well as the need to integrate thematic and particularly socio-demographic data into the estimation of future quantities of change. The spatial aspects of our quantitative adjustments are discussed and tend to show that improvements in the quantitative aspects should not be dissociated from an improvement in the spatial allocation of changes, which may lead to a decrease in the predictive accuracy of the simulations.

1 Introduction

Over the past several decades, geographers developed a large spectrum of models to study land systems through land change, also called land use and land cover change (LUCC), whose socio-environmental impacts have been demonstrated (Chhabra et al. 2006; Mahmood et al. 2014; Oliver and Morecroft 2014). Among them, pattern-based models (PBM) of LUCC (Camacho Olmedo et al. 2018) are spatially explicit models allowed by knowledge of the drivers of change (Lambin et al. 2001; Carr 2003) and by change analysis methods (Mas 1999; Lambin et al. 2001; Comber et al. 2016). The purpose of PBM is to anticipate future changes in order to guide the present action e.g. in terms of public policy, by using prospective scenarios (Houet and Gourmelon 2014). Thus, the prospective scenario technique can be used both to observe the continuation of past and current trends in the future and to project alternative pathways (Veldkamp and Lambin 2001; Escobar et al. 2018). We will focus here on the first approach, called "business as usual" (BAU) scenarios which is a path-dependent approach (Houet et al. 2016) consisting therefore in extending the trend observed in the past over time. BAU scenarios are frequently found in the literature on LUCC modelling (Escobar et al. 2018) as well as in many cases of PBM application, in particular because PBM software packages includes BAU scenarios as baseline simulation (Mas et al. 2014).

According to Mas et al. (2014), the modelling process implemented in these PBM software packages can be divided into five steps : quantity of change estimate, change potential evaluation, spatial allocation of change, reproduction of temporal and spatial patterns and model evaluation. Although there is extensive literature on improving the spatial allocation of changes or model evaluation (Pontius and Millones 2011; Maestripieri and Paegelow 2013), there is little work on improving the quantitative estimation of change intensity. Indeed, most of the time, applying a BAU scenario means defining a single calibration period, between two training dates, according to the available data (Mas et al. 2018). The model uses notably this calibration period to estimate future change quantities, using generally one-order Markov chains (Camacho Olmedo and Mas 2018). Indeed, present-time software only allow the use of only two training dates (e.g. Land Change Modeler, CA_Markov, Dinamica EGO, Metronamica, ApoLUS, LucSim) and it has been shown that the choice of training period is not insignificant and that the simulation results obtained are different according to the training dates that have been chosen (Paegelow et al. 2014; Paegelow 2018).

The spatial expansion of the agricultural frontier in Northern Ecuadorian Amazon (NEA) can be observed over time from historical remote sensing images: settlement patterns and the forest clearing they induce are identifiable by their familiar fish-bone patterns, spread alongside the roads (Baynard et al. 2013). In the NEA, Mena et al. (2006) calculated an annual deforestation rate of 2.49% between 1986 and 1996 and of 1.78% between 1996 and 2002, i.e. a slowing of deforestation over time. We argue that, in a path-dependent approach like that of the BAU scenarios, such a slowdown in the rate of deforestation (i.e. in quantities of change) could not be deduced from purely spatial and linear assumptions, e. g. from only two training dates, but rather requires taking into account a longer period of time and the consideration of thematic data.

Based on a simple model of deforestation dynamics in the NEA using the Land Change Modeler (LCM) software (Eastman and Toledano 2018), we propose here a contribution to improve BAU scenarios and more specifically the quantitative estimation of change intensity by mainly two enhancements. First, this contribution tries to exceed the linearity of BAU scenarios resulting from taking into account only two training dates. Then, authors introduce available, socio-economic, especially demographic, driver data directly to make more realistic classic Markov matrices. Both approaches are implemented by adjusting Markovian transition probabilities.

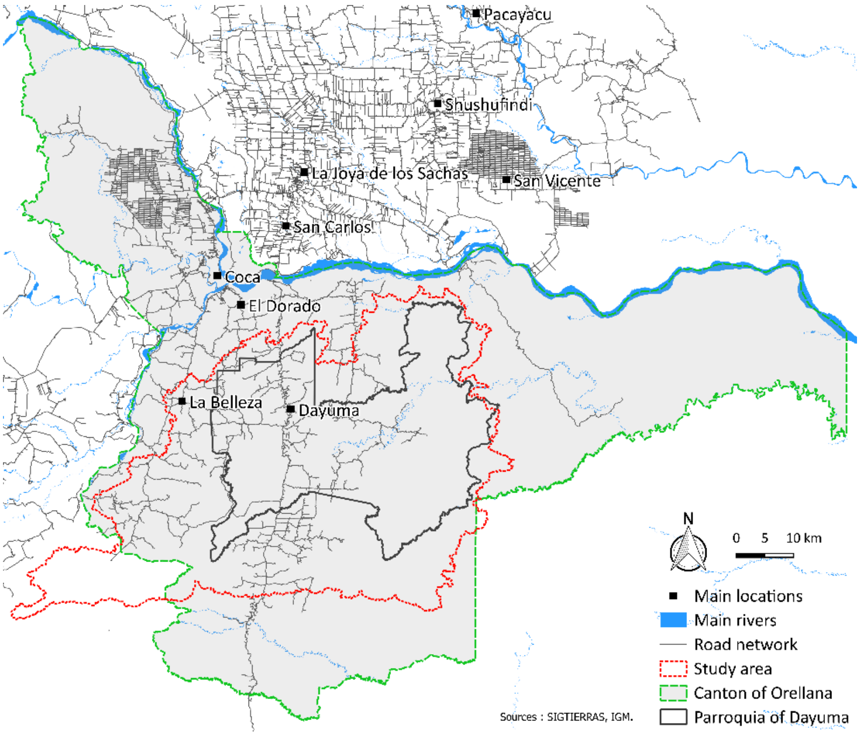
2 Materials and Methods

Context and study area

Northern Ecuadorian Amazon (NEA) is a region located in the western part of the Amazon basin, in the eastern part of Ecuador's national territory called "*Oriente*". This region is characterized by significant endemism to such a degree that it is known as one of the world's biodiversity hotspots (Orme et al. 2005). However, since the discovery of oil fields in the late 1960's, this territory is undergoing significant deforestation coupled with a fast population growth, due to free land accessibility, a high fertility rate and to a continue in-migration (Bilsborrow et al. 2004). Indeed, the road infrastructures built for oil extraction have enabled an agricultural colonization, mainly by small farmers from Andean and Coastal regions of Ecuador (Hiraoka and Yamamoto 1980; Bromley 1981; Brown et al. 1992). This agricultural colonization was spontaneous but also supported by the Ecuadorian authorities through two land reforms (Wasserstrom and Southgate 2013) and logistic support (Juteau-Martineau et al. 2014).

We will focus here on an area composed of a set of sub-watersheds, altogether surrounding and including the *parroquia of* Dayuma, inherited from another modelling approach dealing with environmental contamination (Houssou 2016). This area (Fig. 1), is located south of the city of Coca and Río Napo, in NEA. We have developed land cover classification for this study area using the following procedure detailed below.

Fig. 1: Study area, in NEA.



Data and image processing

The land cover data used in the modelling process were obtained by using relatively simple image processing, coupling supervised segmentation (Paegelow and Camacho Olmedo 2010) and classification based on *maximum likelihood algorithm*. Then, authors calculated the annual deforestation rates. In order to minimize classifications errors, we have chosen to classify land use into four major categories: water, forested areas, deforested areas (merging of the classes "bare soil", "crops" and "pastures") and urban areas. Tables 1, 2 and 3 indicate image characteristics, areas of each of the land cover classes and estimates of the annual rate of deforestation we have derived from it.

Table 1: Image characteristics

Satellite sensor	Path/row	Data acquired	Spatial resolution (m)
LANDSAT 5 - TM	9/60 and 9/61	25 september, 1998	30 m

LANDSAT 7 – ETM+	9/60 and 9/61	12 september, 2002	30m
LANDSAT 8 – OLI TIRS	9/60 and 9/61	2 september, 2013	30m
SENTINEL-2		1 september, 2017	10m decreased to 30m by generalization*
SENTINEL-2		8 february, 2018	10m decreased to 30m by generalization*

* Reduction in the number of columns and rows while decreasing cell resolution by a pixel thinning algorithm.

Table 2: Area per class (ha) on classification

Area per class (ha)	1998	2002	2013	2017
Water	53.46	0	5.67	0
Forested areas	259703.73	253322.01	240156.63	237400.83
Deforested areas	26889.39	33391.35	46470.15	49218.48
Urban areas	95.94	29.16	110.07	123.21

Table 3: Annual deforestation rate (%/year)

1998-2002	2002-2013	2013-2017
-0.61	-0.47	-0.29

Despite some problems in detecting water surfaces, we observe a trend similar to that observed by Mena et al. (2006) at a different period and further north, in an earlier colonized territory: a slowing down of the annual rate of deforestation. Indeed, according to our data, while 0.61% of forests disappeared each year between 1998 and 2002 in our study area, only 0.29% was disappearing each year over the 2013-2017 period.

Implemented pattern-based model

The software package chosen for the pattern-based modelling process, called Land Change Modeler (LCM), is integrated into TerrSet (Eastman 2014) and is used to develop prospective models of LUCC, based on observations of past changes, statistical and machine learning methods to calibrate functions describing the relationship between change and drivers of change (Mas et al. 2018).

Although many authors have focused on analysing the functioning of LCM (Mas et al. 2014; Eastman and Toledano 2018), it is necessary to recall here some essential points about it as a brief overview: to estimate future change quantities, LCM uses Markov chains from a calibration period purposely defined by two training dates in order to determinate matrices of future transition probability between land use classes. LCM allows the use of an external transition probability matrix. Then, in terms of spatial allocation of changes, LCM allows the user to choose between three different methods to determinate the location of future changes, based on the relationships between driver variables loaded into the model and changes that occurred during the training period: (i) a multi-layer perceptron (MLP) neural network (Mas 2004), (ii) Similarity-Weighted Instance-based Machine Learning (Sim-Weight) and (iii) Logistic Regression. The simulation results can be expressed in two forms:

a) a soft simulation, i.e. a map of projected potential for transitions, mapping the places most prone to change. It can then be validated by means of a Receiver Operating Characteristic (ROC) analysis (Pontius and Schneider 2001; Mas et al. 2013)
b) a hard simulation, i.e. a qualitative map of projected LUCC, which can be validated by pixel-by-pixel validation techniques (Chen and Pontius 2010).

LCM is also able to integrate dynamic drivers into the modelling process (recalculated at each time step of the simulation), such as land use or road network as well as incentives or constraints for change e. g. the presence of protected areas that reduce deforestation. Finally, especially for us, BAU-type trend scenarios are included in LCM as baseline simulations.

The model we developed with LCM was trained over the period 2002-2013 and we used it to do projections for the year 2017. For simplification purposes, the only transition considered by the model is the transition from forested to deforested areas. Tables 4 and 5 below show respectively the Markovian matrix of transition probabilities calculated by LCM and the spatial driver variables used by the MLP (method we have chosen to spatially allocate future changes).

Table 4: Markovian matrix of transition probabilities.

	Water	Forested areas	Deforested areas	Urban areas
Water	0	0.3333	0.3333	0.3333
Forested areas	0	0.9657	0.0343	0
Deforested areas	0	0.0882	0.9106	0.0012
Urban areas	0	0	0.2143	0.7857

Reading: from row to column.

Table 5: Spatial driver variables implemented in the model.

Driver variable	Source	Cramer V
Viviendas with one to two rooms	INEC, 2010	0.2777
Viviendas with three to five rooms	INEC, 2010	0.2307
Viviendas with six or more rooms	INEC, 2010	0.2271
People aged 0-14 years old	INEC, 2010	0.2543
People aged 65 years old and more	INEC, 2010	0.2419
People with 6 or more children	INEC, 2010	0.2272
Viviendas connected to the electricity network	INEC, 2010	0.237
Viviendas with WC facilities	INEC, 2010	0.2912
People born in the Sierra	INEC, 2010	0.2969
People born in the Oriente	INEC, 2010	0.259
Population density	INEC, 2010	0.2179
ED to deforested areas in 2002	Own data	0.4606
ED to roads	SIGTIERRAS	0.4606
ED to oil fields	PRAS 2016	0.2373

Drivers that are not Euclidean distances (ED) to relevant features like roads, oil fields or already deforested areas are socio-economic drivers selected from the last population census (Instituto Nacional de Estadística y Censos, INEC, 2010). More specifically, these are maps obtained by spatial interpolation (TIN method, Floriani and Magillo 2009) of census detail file data processed with REDATAM (De Grande 2016), from *localidades dispersas* and *manzanas*, which are census basic point units of the census. First, we selected the drivers to be included in the model based on our readings on deforestation processes in NEA or other South American contexts. So, we selected the variables we assumed to refer to: household size (Morin 2015), position of household in the lifecycle (Perz and Walker 2002), good living conditions (pull effect, Mena et al. 2006) and province of origin (push effect, to identify settlers). In a second step, among these drivers, we arbitrarily selected those with a Cramer V calculated with the class "deforested spaces" greater than 0.2.

LCM provides elements of model skill from the training process, based on analyses of a set of validation pixels: at each iteration, the MLP generates predicted class membership for each of the validation pixels and reports an overall accuracy rate and a skill score. According to the TerrSet documentation, the skill score represents the difference between the calculated accuracy using the validation data and expected accuracy if one were to randomly guess at the class memberships of the validation pixels. We obtained a training accuracy rate of nearby 80% and a skill measure of nearby 0.6. The only dynamic driver is ED to deforested areas and, still for simplification purposes, we have not used any constraint or incentive, although we have cadastral division and that our study area is crossed by the Yasuni National Park, in the east.

3 A classic Markovian BAU scenario and its adjustments

First, we performed a classic BAU trend scenario, as implemented by default in LCM, where future change quantities are estimated using Markov chains, based on two training dates. For this reason, we have named it "Markovian BAU". In a second step, we proposed two consecutive adjustments to the Markovian BAU: an adjustment to exceed its linearity, taking into account a larger time period (BAU-a), and an adjustment to integrate demographic data (BAU-b).

These two adjustments are made by corrections to the Markov transition matrix (Table 4) and are intended to improve quantitative estimation of change intensity in trend scenarios as part of path-dependent pattern-based modelling approaches. In a third step, we considered the spatial aspects of these adjustments.

Markovian BAU

The classic Markovian BAU scenario uses the Markovian matrix calculated by LCM (Table 4) based on the training period we have defined (2002-2013) to determinate future quantities of change in simulations. Under this scenario, the hard simulation produced by LCM overestimates deforestation quantities: as shown in Table 6, nearly 54700 ha of deforestation are estimated by the simulation in 2017, compared to almost 49200 ha on the classification (Table 2), that is about 11,1% overestimation.

Table 6: Markovian BAU scenario simulated areas (ha).

Land cover	2017 by classification	2017 by Markovian BAU predicted area	Model deviations (%)
Water	0	5.67	
Forested areas	237400.83	231847.28	-2.3
Deforested areas	49218.48	54690.28	+11.1
Urban areas	123.21	110.03	-10.7

We assume that this overestimation is mainly due to the non-inclusion of the observed trend of increasing deceleration in the rate of deforestation, as LCM only estimated the quantities of changes from only two training dates. We try to correct this below, by modifying the matrix to take into account the deceleration trend.

BAU-a

The first adjustment we made is therefore to take into account a longer period of time for model calibration. We assume indeed that an observation of the dynamics prior to those of the strict training period (2002-2017) would allow us to better integrate the slowing of the rate of deforestation and thus limit the overestimation of deforestation quantities by the model, that we have previously observed. We have therefore, in concrete terms, changed the original transition matrix to better integrate this deceleration, by multiplying the transition probabilities that interested us by one factor: the ratio between the annual deforestation rates for the periods 1998-2002 and 2002-2013 (Table 3). This ratio, about 0.77, was therefore used to weight the transition probability from forested to deforested area, in bold in Table 7. We then adjusted the cell of persistence of the forested areas class accordingly, in such a way that the sum of the row equals 1 (difference between 1 and the new transition probability). This new modified matrix (Table 7) has been implemented in LCM and has led to new simulations, the results of which in terms of area by class are presented in Table 8.

Table 7 : Modified matrix of Markovian transition probabilities : the BAU-a scenario.

	Water	Forested areas	Deforested areas	Urban areas
Water	0	0.3333	0.3333	0.3333
Forested areas	0	**0.9736**	**0.0264**	0
Deforested areas	0	0.0882	0.9106	0.0012
Urban areas	0	0	0.2143	0.7857

Table 8: BAU-a scenario simulated areas (ha).

Land cover	2017 by classification (ha)	2017 by BAU-a predicted area (ha)	Model deviations (%)
Water	0	5.67	
Forested areas	237400.83	233743.90	-1.5
Deforested areas	49218.48	52793.67	+7.3
Urban areas	123.21	110.03	-10.7

Under this new scenario, we observe this time a lower overestimation of deforestation quantities by the model: LCM overestimates only 7.3%.

BAU-b

Our second proposal to adjust the BAU trend scenario is to integrate population growth dynamics into the transition probability matrix, to make it more realistic. Indeed, population growth is often considered as a major driver of deforestation in the world, in Latin America and especially in NEA (Preston 1996; Armenteras et al. 2017; Jarrín-V. et al. 2017).

Therefore, using the available demographic data from the population censuses (INEC), we calculated a new ratio to reweight the transition matrix. On our study area, the only demographic data available at a fixed spatial scale over time, allowing the calculation of a population growth rate, were those at the cantonal level, and we focused on the canton of Orellana, which includes the Dayuma *parroquia* and most of our study area (Fig. 1). These data (Table 9) indicate that population growth slowed between 1990-2001 (10.32%) and 2001-2010 (8.14%), a trend effectively similar to that of the deforestation rate over a comparable period. We thus calculated the ratio between the annual population growth rates for the two periods (1990-2001 and 2001-2010). This ratio, of 0.79, was used to reweight the transition probability from forested to deforested area, this time in the BAU-a transition matrix (Table 7), in the same way as we did previously (i.e. by weighting the transition from forested to deforested areas by this new factor and then recalculating the other elements of the row). As before, the new matrix (Table 10), resulting from the calculation, was used in LCM to generate new simulations. The results in terms of quantities are presented in Table 11, in comparison with the surfaces of the classification.

Table 9: Demographic data from the population census (INEC) and calculation of the ratio between annual growth rates.

Population of Orellana Canton		
1990	2001	2010
19674	42010	72795
Growth rates (%)		
1990-2001	2001-2010	Ratio
10.32	8.14	0.79

Table 10: Modified matrix of Markovian transition probabilities: the BAU-b scenario.

	Water	Forested areas	Deforested areas	Urban areas
Water	0	0.3333	0.3333	0.3333
Forested areas	0	**0.9792**	**0.0208**	0
Deforested areas	0	0.0882	0.9106	0.0012
Urban areas	0	0	0.2143	0.7857

Table 11: BAU-b simulated areas (ha).

Land cover	2017 by classification (ha)	2017 by BAU-b predicted area (ha)	Model deviations (%)
Water	0	5.67	
Forested areas	237400.83	235088.43	-1
Deforested areas	49218.48	51449.13	+4.5
Urban areas	123.21	110.03	-10.7

As we can observe, after this second adjustment of the transition probability matrix, the overestimation by LCM is only 4.5% compared to the classification (Table 11). Our successive adjustments have therefore reduced the overestimation of quantities by more than half: whereas the classic Markovian BAU scenario simulated about 11.1% more deforestation than observed while the adjusted BAU-b scenario, the most advanced, generates only 4.5% of deviations to the model.

It seems that the adjustments have improved the quantitative estimation of change intensity, by exceeding the linearity of Markovian BAU scenarios based on only two dates and weighting the change probability matrix with demographic data. It is now a matter for us to briefly analyse the spatial effects of these adjustments.

Spatial aspects

In order to consider the spatial aspects of our successive quantitative adjustments, that led to the simulation results of the BAU-b scenario presented before, we use here the method developed by Chen and Pontius (2010). Based on the observation that the Kappa indices are ineffective for accuracy assessment (Pontius and Millones 2011) on the one hand, and the need for statistical assessment on the other (Pontius et al. 2004), a part of this method consists in categorizing pixels into four categories in order to identify omission and commission errors (Pontius 2000): *(i)* correct due to observed persistence predicted as persistence (null successes), *(ii)* error due to observed persistence predicted as change (false alarms), *(iii)* correct due to observed change predicted as change (hits) and *(iv)* error due to observed change predicted as persistence (misses).

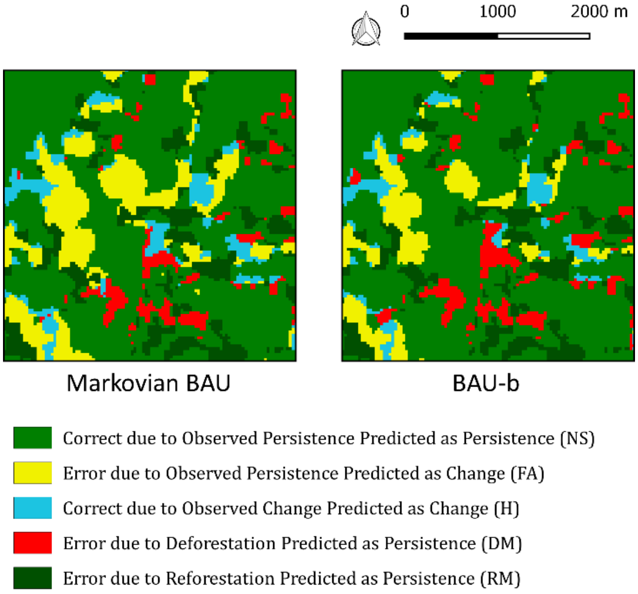
Table 12 shows the proportion of each of these categories in the Markovian BAU and in the BAU-b scenarios, calculated as a percentage of our study area. We have chosen to indicate separately errors due to reforestation (transition from deforested to forested area), which, as we recall, is not a process taken into account by the model. Figure 2 shows a portion of the territory simulated by the Markovian BAU scenario (left) and by BAU-b scenario (right), qualified according to this categorization of prediction successes and errors.

Table 12: Overall prediction successes and error across the entire study area for Markovian BAU and BAU-b scenarios (%).

	NS	FA	H	DM	RM
Markovian BAU	89.10	1.32	0.89	4.38	4.31
BAU-b	89.40	1.03	0.71	4.56	4.31

NS = null successes, FA = false alarms, H = hits, DM = deforestation misses, RM = reforestation misses.

Fig. 2: Accuracy components based on observed land cover (2013, 2017) and 2017 predicted land cover maps from the Markovian BAU and the BAU-b.



This analysis of the spatial aspects of simulation successes and errors demonstrate that quantitative adjustments to the probability matrices of change are not devoid of spatial consequences. Indeed, as shown in Table 12, hits, which refer to deforestation that occurred between 2013 and 2017 and correctly predicted by LCM, represent nearby 0.9% of the study area before adjustments compared to almost 0.7% after adjustments (BAU-b). This is also visible on the map (Fig. 2) showing a representative detail of the study area, where the hits appear in blue: they are more numerous on the left extract assessing the Markovian BAU than on the right extract assessing the BAU-b. Inversely, errors due to deforestation predicted as persistence (DM) are increasing after adjustments: they represented nearly 4.4% of the study area before adjustments compared to nearly 4.6% after adjustments (in red on map extracts).

Then, if we relate the area of hits in both cases (Markovian BAU and BAU-b) to the deforestation area simulated by the two scenarios, we also find that hits are decreasing after successive adjustments. Under the Markovian BAU scenario, about 3200 ha of hits are observed among the 54690 ha of simulated deforestation, or approximately 5.9%. For BAU-b, about 2045 ha of hits are observed for 51450 ha of simulated deforestation, or almost 4%: the proportion of hits in simulated deforestation decreases as a result of adjustments.

4 Discussion

This paper show that BAU scenarios as implemented in LCM, i.e. based on a training period established from only two dates, may be insufficient to provide a correct quantitative estimation of change intensity. Indeed, using a classic Markovian BAU scenario, the simple LCM model used here was not able to accurately reproduce the observed trend, i.e. a slowing down in the deforestation rate, since it overestimates quantities of deforestation. The work presented here explore two ways of improving change intensity prediction in BAU scenarios in land change modelling: extend the trend observation period and use thematic data to more accurately predict future quantities of change. These assumptions are applied by successive weightings of the model's transition probability matrix.

First of all, the results show that such adjustments of the probability transition matrix can improve path-dependant modelling approaches: they led to a lower overestimation of deforestation quantities in the simulations. These results therefore highlight the value of incorporating a longer time period and the benefits of taking socio-demographic data into account during the calibration step, to exceed the linearity in the construction of change quantities prediction and to make them more realistic. This last approach is in line with the idea of "socializing" pixels, which appeared in the 1990s (Martin and Bracken 1993; National Research Council 1998).

These results also imply that BAU scenarios would benefit from being better designed: as these are so-called "trend" scenarios, because they are path-dependent approaches, we believe that it would be more efficient to build them on the basis of a more in-depth understanding of the trends, i.e. beyond the only two training dates allowed by the current software packages. In this sense, the integration of higher-order Markov chains (Ching et al. 2013) into LUCC modelling tools could be a potential path to consider, because the successive adjustments of the Markovian matrices proposed here cannot constitute a robust methodology applicable to all cases and all types of thematic data.

However, these adjustments emphasize the need to integrate socio-economic data at each step of the LUCC modelling process, in one way or another and therefore not only at the spatial allocation of changes step as is currently the case in software packages. We believe that socio-economic data must be used to estimate future quantities of LUCC, as they are used to predict spatial allocation. This is obviously valid for all LUCC modelling approaches, above and beyond BAU scenarios. Because land use systems are characterized by multiple, non-linear and complex interactions between societies and environment, at different temporal and spatial scales (Geist et al. 2006), and cannot therefore be limited to the use of purely spatial or physical drivers, whatever the stage of the modelling process. In addition, it is likely that the coming decades will be characterized by the multiplication of accessible socio-economic data as well as those of big data. The latter represent a major challenge for many scientific disciplines and geography and geomatics are no exception (Kitchin 2013). Lastly, it is interesting to note that population projections studies have become increasingly numerous and accessible in recent years, including in Global South countries such as Ecuador, where they are produced and published by INEC. These data can be useful in the development of prospective BAU scenarios, especially when they are themselves trend-based.

Finally, the results show that in LCM, an improvement in the estimation of future quantities of change can lead to a decrease in the proportion of hits in predicted changes, i.e. changes that occurred and were correctly predicted. The improvement in the prediction of change quantities has indeed led to a decrease in the quality of the results in terms of spatial allocation in the case of the study presented here. This is simply because to spatially allocate changes, LCM selects the pixels with the highest change potential on the transition potential map calculated by the MLP neural network based on the relationship between the changes that occurred during the training period and drivers. But a reduction in quantities simply results in a smaller selection of pixels and therefore a smaller simulated change area. However, if the software simulates fewer changes, it is less likely to hit the target: the changes that occurred. Thus, this result suggests that quantitative improvements must be accompanied by progress in the spatial allocation of changes, in particular the reproduction of realistic patterns of change by models e.g. as Dinamica EGO software allows it better through its mechanism of expander/patcher (Soares-Filho et al. 2002; Rodrigues and Soares-Filho 2018). Process-based LUCC modelling approaches like agent-based models are also an interesting approach in this field (Parker et al. 2003; Matthews et al. 2007) and their coupling to pattern-based approaches is still an important scientific issue (Castella and Verburg 2007).

Although this approach allowed us to obtain results which contribute to a brief reflection on LUCC modelling practices and especially on the calibration stage of trend/BAU scenarios, it has however several limitations. These limitations are due both to the data used, to the choices made during the construction of the model in LCM and to the method by which we adjusted matrices.

Initially, these limitations concern the accuracy of the remotely sensed data as model input data. Indeed, it has been demonstrated that uncertainty is present at each step of the construction of land cover data and that is therefore significantly present in LUCC models (Garcia Alvarez 2018). Besides, this uncertainty is at the root of a numerous and ongoing work on improving satellite image classification techniques which show that there is always a scope for improvement (Lu and Weng 2007; Tso and Mather 2009). Nonetheless, it is important to remember that while supervised classification methods offer many advantages, including time savings, practical systematization and greater objectivity, their accuracy is often lower than manual classification by photo-interpretation. In brief, the results must be balanced according to the confidence we can place in land cover data, especially since they were not validated by field surveys.

Then, regarding the limitations inherent in the construction of the model, some choices made to simplify the model in order to improve understanding can be discussed. In particular, the non-inclusion of transitions from forested to deforested areas (reforestation process) can reduce model accuracy as much as the non-use of more dynamic data updated at each iteration like cadastral data or dynamic road modelling. However, the main purpose of this paper is to propose an improvement of the quantitative estimation of change intensity in trend scenarios, that is why the emphasis has been placed mainly on the quantitative aspects, to the detriment of certain details, which may nevertheless usually be essential for the development of a complete LUCC model.

Finally, the last limitation that we can highlight is the mismatch of spatial and temporal scales when the transition probability matrix has been modified. Indeed, for the second adjustment (the BAU-b scenario), we used cantonal demographic data for a model applied to a lower spatial level. Another bias lies in the fact that these cantonal data include several cities, characterized by specific demographic dynamics, while our territory is essentially rural. In addition, the time scale of the population censuses used to weight the matrix does not exactly match that of our classifications.

5 Conclusion

Based on a simple LUCC model developed with LCM, this work highlights the need to extend the trend observation period and to include thematic data in the calibration step of path-dependent pattern-based modelling approaches, to improve the quantitative estimation of change intensity. Indeed, the successive adjustments to the original Markov matrix of transition probabilities have minimized the model's overestimation of deforestation.

A quick spatial analysis of the results also recalls that improving the quantitative estimation of changes cannot be done independently of progress in the spatial allocation of changes.

References

Armenteras D, Espelta JM, Rodríguez N, Retana J (2017) Deforestation dynamics and drivers in different forest types in Latin America: Three decades of studies (1980–2010). Global Environmental Change 46:139–147 . doi: 10.1016/j.gloenvcha.2017.09.002

Baynard CW, Ellis JM, Davis H (2013) Roads, petroleum and accessibility: the case of eastern Ecuador. GeoJournal 78:675–695 . doi: 10.1007/s10708-012-9459-5

Bilsborrow RE, Barbieri AF, Pan W (2004) Changes in population and land use over time in the Ecuadorian Amazon. Acta Amazonica 34:635–647 . doi: 10.1590/S0044-59672004000400015

Bromley R (1981) The colonization of humid tropical areas in Ecuador. Singapore Journal of Tropical Geography 2:15–26 . doi: 10.1111/j.1467-9493.1981.tb00114.x

Brown LA, Sierra R, Southgate D, Labao L (1992) Complementary Perspectives as a Means of Understanding Regional Change: Frontier Settlement in the Ecuador Amazon. Environment and Planning A 24:939–961 . doi: 10.1068/a240939

Camacho Olmedo MT, Mas JF (2018) Markov Chain. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 441–445

Camacho Olmedo MT, Paegelow M, Mas JF, Escobar F (2018) Geomatic Approaches for Modeling Land Change Scenarios. An Introduction. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 1–8

Carr DL (2003) Proximate Population Factors and Deforestation in Tropical Agricultural Frontiers. Population and Environment 25:585–612 . doi: 10.1023/B:POEN.0000039066.05666.8d

Castella J-C, Verburg PH (2007) Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. Ecological Modelling 202:410–420 . doi: 10.1016/j.ecolmodel.2006.11.011

Chen H, Pontius RG (2010) Diagnostic tools to evaluate a spatial land change projection along a gradient of an explanatory variable. Landscape Ecology 25:1319–1331 . doi: 10.1007/s10980-010-9519-5

Chhabra A, Geist H, Houghton RA, Haberl H, Braimoh AK, Vlek PLG, Patz J, Xu J, Ramankutty N, Coomes O, Lambin EF (2006) Multiple Impacts of Land-Use/Cover Change. In: Lambin EF, Geist H (eds) Land-Use and Land-Cover Change. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 71–116

Ching W-K, Huang X, Ng MK, Siu T-K (2013) Higher-Order Markov Chains. In: Markov Chains. Springer US, Boston, MA, pp 141–176

Comber A, Balzter H, Cole B, Fisher P, Johnson S, Ogutu B (2016) Methods to Quantify Regional Differences in Land Cover Change. Remote Sensing 8:176 . doi: 10.3390/rs8030176

De Grande P (2016) El formato Redatam / The Redatam format. Estudios Demográficos y Urbanos 31:811 . doi: 10.24201/edu.v31i3.15

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

Eastman JR, Toledano J (2018) A Short Presentation of the Land Change Modeler (LCM). In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 499–505

Escobar F, van Delden H, Hewitt R (2018) LUCC Scenarios. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 81–97

Floriani LD, Magillo P (2009) Triangulated Irregular Network. In: Liu L, Özsu TT (eds) Encyclopedia of Database Systems. Springer US, Boston, MA, pp 3178–3179

Garcia Alvarez D (2018) Aproximación al Estudio de la Incertidumbre en la Modelización del Cambio de Usos y Coberturas del Suelo (LUCC). Universidad de Granada

Geist H, McConnell W, Lambin EF, Moran E, Alves D, Rudel T (2006) Causes and Trajectories of Land-Use/Cover Change. In: Lambin EF, Geist H (eds) Land-Use and Land-Cover Change. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 41–70

Hiraoka M, Yamamoto S (1980) Agricultural Development in the Upper Amazon of Ecuador. Geographical Review 70:423 . doi: 10.2307/214077

Houet T, Aguejdad R, Doukari O, Battaia G, Clarke K (2016) Description and validation of a "non path-dependent" model for projecting contrasting urban growth futures. Cybergeo. doi: 10.4000/cybergeo.27397

Houet T, Gourmelon F (2014) La géoprospective – Apport de la dimension spatiale aux démarches prospectives. Cybergeo. doi: 10.4000/cybergeo.26194
Houssou L (2016) Simulation sociale à base d'agents du comportement microéconomique des ménages en Amazonie équatorienne, face aux contaminations pétrolières, aux dynamiques économiques et aux politiques publiques. MSc thesis, Université nationale du Vietnam, Institut francophone international, Hanoi, Vietnam
Jarrín-V. PS, Tapia Carrillo L, Zamora G (2017) Demografía y transformación territorial: medio siglo de cambio en la región amazónica de Ecuador/ Demography and territorial transformation: half a century of change in the Amazonian Region of Ecuador. Eutopía, Revista de Desarrollo Económico Territorial 81 . doi: 10.17141/eutopia.12.2017.2913
Juteau-Martineau G, Becerra S, Maurice L (2014) Ambiente, petróleo y vulnerabilidad política en el Oriente Ecuatoriano: ¿hacia nuevas formas de gobernanza energética? América Latina Hoy 67:119 . doi: 10.14201/alh201467119137
Kitchin R (2013) Big data and human geography: Opportunities, challenges and risks. Dialogues in Human Geography 3:262–267 . doi: 10.1177/2043820613513388
Lambin EF, Turner BL, Geist HJ, Agbola SB, Angelsen A, Bruce JW, Coomes OT, Dirzo R, Fischer G, Folke C, George PS, Homewood K, Imbernon J, Leemans R, Li X, Moran EF, Mortimore M, Ramakrishnan PS, Richards JF, Skånes H, Steffen W, Stone GD, Svedin U, Veldkamp TA, Vogel C, Xu J (2001) The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change 11:261–269 . doi: 10.1016/S0959-3780(01)00007-3
Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28:823–870 . doi: 10.1080/01431160600746456
Maestripieri N, Paegelow M (2013) Validation spatiale de deux modèles de simulation : l'exemple des plantations industrielles au Chili. Cybergeo. doi: 10.4000/cybergeo.26042
Mahmood R, Pielke RA, Hubbard KG, Niyogi D, Dirmeyer PA, McAlpine C, Carleton AM, Hale R, Gameda S, Beltrán-Przekurat A, Baker B, McNider R, Legates DR, Shepherd M, Du J, Blanken PD, Frauenfeld OW, Nair US, Fall S (2014) Land cover changes and their biogeophysical effects on climate: land cover changes and their biogeophysical effects on climate. International Journal of Climatology 34:929–953 . doi: 10.1002/joc.3736
Martin D, Bracken I (1993) The integration of socioeconomic and physical resource data for applied land management information systems. Applied Geography 13:45–53 . doi: 10.1016/0143-6228(93)90079-G
Mas J (2004) Modelling deforestation using GIS and artificial neural networks. Environmental Modelling & Software 19:461–471 . doi: 10.1016/S1364-8152(03)00161-0
Mas J-F (1999) Monitoring land-cover changes: A comparison of change detection techniques. International Journal of Remote Sensing 20:139–152 . doi: 10.1080/014311699213659
Mas J-F, Kolb M, Paegelow M, Camacho Olmedo MT, Houet T (2014) Inductive pattern-based land use/cover change models: A comparison of four software packages. Environmental Modelling & Software 51:94–111 . doi: 10.1016/j.envsoft.2013.09.010
Mas JF, Paegelow M, Camacho Olmedo MT (2018) LUCC Modeling Approaches to Calibration. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 11–25
Mas J-F, Soares Filho B, Pontius R, Farfán Gutiérrez M, Rodrigues H (2013) A Suite of Tools for ROC Analysis of Spatial Models. ISPRS International Journal of Geo-Information 2:869–887 . doi: 10.3390/ijgi2030869
Matthews RB, Gilbert NG, Roach A, Polhill JG, Gotts NM (2007) Agent-based land-use models: a review of applications. Landscape Ecology 22:1447–1459 .
Mena CF, Bilsborrow RE, McClain ME (2006) Socioeconomic Drivers of Deforestation in the Northern Ecuadorian Amazon. Environmental Management 37:802–815 . doi: 10.1007/s00267-003-0230-z

Morin L (2015) Diagnostic agraire d'un front pionnier en Amazonie équatorienne, Paroisse de Dayuma, province d'Orellana, Equateur. MSc thesis, SupAgro Montpellier IRC
National Research Council (1998) People and Pixels: Linking Remote Sensing and Social Science. National Academies Press, Washington, D.C.
Oliver TH, Morecroft MD (2014) Interactions between climate change and land use change on biodiversity: attribution problems, risks, and opportunities: Interactions between climate change and land use change. Wiley Interdisciplinary Reviews: Climate Change 5:317–335 . doi: 10.1002/wcc.271
Orme CDL, Davies RG, Burgess M, Eigenbrod F, Pickup N, Olson VA, Webster AJ, Ding T-S, Rasmussen PC, Ridgely RS, Stattersfield AJ, Bennett PM, Blackburn TM, Gaston KJ, Owens IPF (2005) Global hotspots of species richness are not congruent with endemism or threat. Nature 436:1016–1019 . doi: 10.1038/nature03850
Paegelow M (2018) Impact and Integration of Multiple Training Dates for Markov Based Land Change Modeling. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 121–138
Paegelow M, Camacho Olmedo MT (2010) Modelos de simulación espacio-temporal y Teledetección: el método de la segmentación para la cartografía cronológica de usos del suelo. Serie Geográfica – Universidad de Alcalá pp 19–34
Paegelow M, Camacho Olmedo MT, Mas J-F, Houet T (2014) Benchmarking of LUCC modelling tools by various validation techniques and error analysis. Cybergeo. doi: 10.4000/cybergeo.26610
Parker DC, Manson SM, Janssen MA, Hoffmann MJ, Deadman P (2003) Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. Annals of the Association of American Geographers 93:314–337 . doi: 10.1111/1467-8306.9302004
Perz SG, Walker RT (2002) Household Life Cycles and Secondary Forest Cover Among Small Farm Colonists in the Amazon. World Development 30:1009–1027
Pontius GR (2000) Quantification Error Versus Location Error in Comparison of Categorical Maps. Photogrammetric Engineering and Remote Sensing 66:1011–1016
Pontius RG, Huffaker D, Denman K (2004) Useful techniques of validation for spatially explicit land-change models. Ecological Modelling 179:445–461 doi: 10.1016/j.ecolmodel.2004.05.010
Pontius RG, Millones M (2011) Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing 32:4407–4429 . doi: 10.1080/01431161.2011.552923
Pontius RG, Schneider LC (2001) Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems & Environment 85:239–248 . doi: 10.1016/S0167-8809(01)00187-6
Preston SH (1996) The effect of population growth on environmental quality. Population Research and Policy Review 15:95–108 . doi: 10.1007/BF00126129
Rodrigues H, Soares-Filho B (2018) A Short Presentation of Dinamica EGO. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) Geomatic Approaches for Modeling Land Change Scenarios. Springer International Publishing, Cham, pp 493–498
Soares-Filho BS, 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 . doi: 10.1016/S0304-3800(02)00059-5
Tso B, Mather P (2009) Classification Methods for Remotely Sensed Data, Second Edition. CRC Press
Veldkamp A, Lambin E. (2001) Predicting land-use change. Agriculture, Ecosystems & Environment 85:1–6 . doi: 10.1016/S0167-8809(01)00199-2
Wasserstrom R, Southgate D (2013) Deforestation, Agrarian Reform and Oil Development in Ecuador, 1964-1994. Natural Resources 04:31–44 . doi: 10.4236/nr.2013.41004