Remote Sensing of Environment 170 (2015) 203–215

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Controlling for misclassified land use data: A post-classification latent multinomial logit approach

Raymundo Marcos Martinez *, Kenneth A. Baerenklau

Department of Environmental Sciences, University of California, Riverside, CA 92521, USA

ARTICLE INFO

Article history:
Received 29 January 2015
Received in revised form 1 August 2015
Accepted 25 September 2015
Available online 2 October 2015

Keywords:
Land use
Misclassification
Latent multinomial logit model
Expectation maximization algorithm
Transition rules
Landsat
Agroforests
Monte Carlo simulation

ABSTRACT

Terrain and landscape complexities can limit the accurate discrimination of land use categories with similar spectral signatures, as well as the accurate detection of land use change in temporal analyses of landscape dynamics. Studies based on misclassified land use data can generate biased parameter estimates and standard errors, inaccurate predictions, and incorrect policy recommendations. To address these challenges and improve the accuracy of land use analyses, we implement a post-classification strategy to detect misclassified land use observations using a latent multinomial logit model. This strategy is tested using both Monte Carlo simulations and a time series dataset based on supervised classification of remotely sensed data corresponding to land use decisions observed in a Mexican coffee growing region during the period 1984–2006. The results indicate that the strategy is useful for identifying land use observations with a high probability of being wrongly classified, even between categories with low discriminative spectral signatures. Reclassification of the land use data, based on the model results, increases the magnitudes of the marginal effects of the analyzed land use drivers in the theoretically expected directions, and in some cases improves the statistical significance of the parameter estimates.

1. Introduction

Classification errors are an intrinsic component of spatially explicit land use models that impact the accuracy of parameter estimates, predictions, and derived policy recommendations. Inaccuracies in land use and land cover (LULC) classification have several sources. In some cases, the resolution or quality of the remotely sensed data complicates the classification process. For instance, when the image has a high percentage of cloud cover or when the pixel size is very large that only coarse land use classification can be implemented. In other cases, terrain or landscape complexities complicate the discrimination of classes with similar plant functional types or low discriminative spectral signatures. For example, it is difficult to identify shrub lands from herbaceous crops in sparsely vegetated areas, or different types of forests (Fritz & See, 2008; Gao & Jia, 2013; Steele, Chris Winne, & Redmond, 1998). Furthermore, LULC classification errors can propagate in temporal analysis thereby reducing the precision of land use change detection procedures, particularly when more than two periods are considered (Yuan, Sawaya, Loeffelholz, & Bauer, 2005).

Transition probability matrices, expert rules, and change detection algorithms have been implemented to improve the accuracy of multi-period LULC classifications. For example Lehmann, Wallace, Caccetta, Furby, and Zdunic (2013) use transition probabilities and expert rules to improve the mapping of forested and non-forested areas in Australia for the period 1989–2006 using Landsat Thematic Mapper imagery. Kleynhans et al. (2010) tested a change detection procedure based on an extended Kalman filter to detect new rural settlements in South African savannas, grasslands and shrub lands using time series Moderate Resolution Imaging Spectroradiometer (MODIS) data. Fraser, Olthof, and Pouliot (2009) implemented change detection procedures, including expert rules constraining land cover transitions, to improve vegetation analysis in Canada's national park system for the period 1985 to 2005.

As an alternative approach to improving the accuracy of LULC datasets, in this paper we implement a post-classification strategy that simultaneously detects misclassified land use observations and incorporates corrections into a latent multinomial logit (LMNL) land use model. Because accurate classification of anthropogenic land uses is key for understanding landscape dynamics, we focus our analysis on land use classifications. Nevertheless, the method is also applicable to land cover classification.

A time series land use dataset based on supervised classification of remotely sensed data, controlled with transition rules to remove inter-temporal inconsistencies, is used to test the LMNL procedure. The dataset is derived from a Mexican coffee growing region in which the vegetation density of forested areas, agroforestry parcels, and abandoned lands produces similar spectral values that are difficult to discriminate even with state-of-the-art object-oriented classifiers. The results from the empirical application indicate that the LMNL model

[*] Corresponding author.
E-mail address: rmarc004@ucr.edu (R. Marcos Martinez).

http://dx.doi.org/10.1016/j.rse.2015.09.025
0034-4257/© 2015 Elsevier Inc. All rights reserved.

can be used to detect misclassified observations and to replace subjectively determined transition rules. This is useful for improving the accuracy of land use datasets and the robustness of related analyses. In our empirical study, the reconfiguration of the original dataset also is used to quantify the impact of such inaccuracies on the estimated marginal effects of land use drivers. Additional validation of the LMNL algorithm through Monte Carlo simulations indicates that the approach is highly accurate for detecting misclassified land use data.

2. Literature review

Since the seminal work of Dempster, Laird, and Rubin (1977), the expectation maximization (EM) algorithm has been used to generate parameter estimates in probabilistic models with incomplete or misclassified data. This is typically done by associating an incomplete data problem with a complete-data problem for which maximum likelihood estimation is tractable (McLachlan & Krishnan, 1997). An iterative process between the expectation step (E-step) and the maximization step (M-step) is the basis of the EM algorithm. The E-step computes the expectation of the missing/misclassified data conditional on the given set of incomplete information and initial values of the parameters to be estimated. The M-step uses those conditional expectations in the place of the missing/misclassified information to “complete” the dataset and estimate the parameters that maximize the likelihood function for the “complete-data” problem. The parameter estimates produced in the M-step are used as updated initial values of the coefficients in the E-step and the process is repeated until the likelihood converges to a local maximum (McLachlan & Krishnan, 1997; Zhai, 2007).

In the context of land use and land cover mapping the EM algorithm has been used to refine unsupervised classification methods (Chardin & Perez, 1999; Yang, Peng, Xia, & Zhang, 2013); to estimate the pixel values of portions of remotely sensed imagery that are missing due to the presence of clouds during the time of data collection (Melgani, 2006); and to improve the classification accuracy of pixels that include mixed information corresponding to more than one land use category (Susaki, Shibasaki, Susaki, & Shibasaki, 2000). To our knowledge the EM algorithm has not been used to analyze the impact of misclassified data on agent based land use analyses, a task that can be accomplished using a latent multinomial logit model.

The LMNL model uses a nesting structure to represent the *N* discrete choices in a dataset with *N* branches. The structure is nested because each branch contains a sub-structure with one stem representing accurately classified observations, and up to *N*-1 stems containing misclassified observations that should be classified into the other *N*-1 branches. For instance, consider a land use dataset classified into Cereals, Grasslands, and Forests with potential misclassifications between the first two categories. In the LMNL context, such a dataset can be represented by three branches (Cereals, Grasslands, and Forests), with each branch containing one stem that accounts for observations that are correctly classified; and with the Cereals and Grassland branches containing an additional stem that controls for misclassified [unknown] observations (Fig. 1).

Caudill (2006), describes the methodology that can be used to produce parameter estimates with a dataset containing misclassified dependent variables, as is the case studied here. The procedure is based on a transformation of the standard multinomial logit likelihood function into a missing data formulation to which the EM algorithm can be applied. The methodology has been used to identify misleading response rates in a survey used to collect information on cheating behavior (Caudill & Mixon, 2005); to estimate the proportion of fraudulent claims for car damage that are erroneously classified as honest by an insurance company (Caudill, Ayuso, & Guillen, 2005); and to estimate the impact of misclassified observations on an analysis of hidden unemployment in six European economies (Caudill, 2006). More recently, the study by Caudill, Groothuis, and Whitehead (2011) uses an unconstrained version of the LMNL model to analyze hypothetical bias (the situation in which stated willingness to pay is higher than the actual willingness to pay) in a contingent valuation problem. The LMNL methodology offers a straightforward procedure to handle misclassified land use information as described in the following section.

3. Empirical application

Spatially explicit models of land use decisions in rural areas typically focus on how the driving forces of deforestation reconfigure pristine landscapes and affect the provision of environmental services (Andersen, 1996; Chomitz & Gray, 1996; Geist & Lambin, 2002; Puri, 2006). Nevertheless, the growing recognition that agroforestry production systems can provide forest-like services as well as biodiversity corridors between patches of forested or protected areas has highlighted the need for understanding land use decisions in agroforests (Ávalos-Sartorio & Blackman, 2010; Bhagwat, Willis, Birks, & Whittaker, 2008; Dinata Putra, Verbist, & Budidarsono, 2005; Huang et al., 2002; Kursten, 2000; Schroth, 2004; Shanker & Solanki, 2000; Swallow, Boffa, & Scherr, 2006). Worldwide, shade-grown coffee plantations are one of the most important agroforestry production systems not only for their ability to provide livelihood opportunities to many farmers (Albers, Avalos-Sartorio, Batz, & Blackman, 2006; Aoki & Suvedi, 2012; Blackman, Ávalos-Sartorio, & Chow, 2012; Jordan-Garcia, Collazo, Borkhataria, & Groom, 2012; Oxfam, 2002), but also for their ecological services (Escamilla Prado, 2007; Messer, Kotchen, & Moore, 2000). In Mexico, small-scale farmers across the

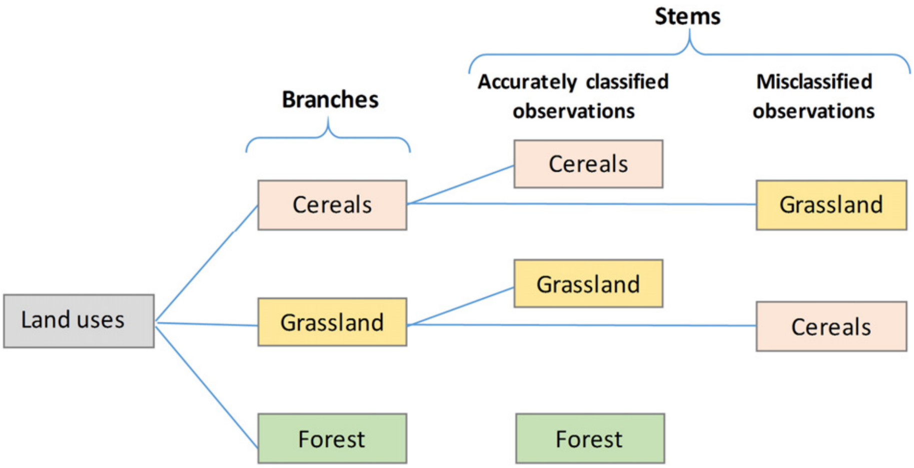

Fig. 1. Example of a latent multinomial logit nesting structure to control for misclassified observations in two out of three land use categories.

country depend upon shade grown crops, with coffee being the leader both in terms of cultivated land area and value of production. Escamilla Prado (2007) reports that around 3 million people in Mexico depend on coffee-related activities and that approximately 90% of the coffee-cultivated area lays under diversified shade. Unfortunately, the steady decline in the international coffee price during the 1990's and first years of the 2000's forced coffee farmers to find alternative sources of income. Some farmers opted for coffee certification schemes to obtain a price premium for implementing environmentally friendly production techniques, while others decided to clear their coffee plantations to transition to a different land use. In other cases, farmers abandoned their plantations to look for employment opportunities in other economic sectors and/or geographical locations (Blackman, Albers, Avalos-Sartorio, & Murphy, 2008; Lewis & Runsten, 2008; Nava-Tablada & Martínez-Camarillo, 2012).

In this paper, we utilize land use data from the low altitude zone of the municipality of Atzalan, Veracruz, Mexico (Fig. 2). Landscape metric and econometric analyses implemented by Ellis, Marcos-Martínez, and Chávez (2010) and Baerenklau, Ellis, and Marcos-Martínez (2012) indicate that this region registered a significant loss of tree canopy during the 1990s, mainly in coffee growing areas in response to the decline in the profitability of coffee-based agroforests during that decade. The study area consists of around 25,500 ha distributed across an altitudinal gradient that extends from 85 to 726 m above sea level. The landscape in that region has gradually reconfigured from secondary forest and coffee parcels to grasslands, citrus groves and banana plantations. Information collected in 2006 by the Mexican government (SAGARPA, 2006) indicates that, at the municipality level, citrus production was the main agricultural activity accounting for 68% of the agricultural GDP. Banana plantations contributed 12% of the production value; corn generated 9% and coffee production—after representing the main income source in the region during previous decades—only contributed 5% in that year. In aggregate around 89% of the agricultural GDP in the municipality is generated by agricultural systems that do not require tree canopy, which impacts the provision of environmental services.

3.1. Land use data

Land use information was obtained for the study region by classifying one Landsat Multispectral Scanner (MSS) image collected in 1973, six Landsat Thematic Mapper (TM) images for the years 1984, 1989, 1993, 1996, 2000, 2003, and one image collected by the *Satellite Pour l'Observation de la Terre* (SPOT-5) High Resolution Geometric sensor in 2006. All images were orthorectified and underwent radiometric calibration. Maximum likelihood supervised classification was applied using training samples to generate spectral signatures for each land use class. Training samples for the 2003 and 2006 images were produced using reference data. Mean values of the spectral signatures for the 2003 training samples estimated with the older images were computed, and their values compared with those obtained from the 2003 Landsat TM image. Training samples with similar signatures and located in visually similar and unchanged areas, relative to the 2003 image, were selected to classify the remaining Landsat MSS and TM imagery.

This process allowed the classification of the satellite imagery into three general land use categories: agroforestry (AG) which is composed of shade grown coffee plantations and secondary forest; perennial crops (PC), composed of citrus and banana plantations; and grasslands and cornfields (GC). The main criteria to construct the aggregated land use categories are that their components share similar biomass density, profitability and conversion costs. To assess the accuracy of the 2003 and 2006 classifications we used reference data from 165 and 168 locations, respectively. The 2006 classification presents an overall accuracy of 72% (Kappa-Cohen statistic of 0.58), while the 2003 classification has an overall accuracy of 68% (Kappa statistic of 0.52). Those accuracy levels are comparable to other studies implemented in regions with similar land uses (Cayuela, Benayas, & Echeverría, 2006; Ellis et al., 2010; Muñoz-Villers & López-Blanco, 2008).

The small number of citrus and banana plantations present in the study region in 1973, and the quality and resolution of the spectral information contained in the Landsat MSS image collected in that year, limited our ability to generate an adequate set of training polygons for the PC category in that period. This in fact prevented the identification of PC land in that image restricting our analysis to the period 1984–2006. Nevertheless, we used the AG and GC classification from the 1973 image to identify pixels (of 30 × 30 meters) that maintained the same land use during the period 1973–1984. Those pixels correspond to AG and GC land uses with an age of at least 11 years at the beginning of 1984. We used this approach to filter out new plantations that were potentially “locked” in a particular land use until recovering establishment costs, and to focus our analysis on land that could transition to a different use at the beginning of the study period without restrictions. Around 79% of the study area satisfied this criterion.

Land use change in most cases is a costly action since it requires the removal of the current land use, an up-front investment to establish a new crop, and the financial resources to implement maintenance activities during the growing period of the newly planted crops. Under some circumstances agents would prefer to abandon their lands during some periods and pursue employment in other sectors of the economy instead of changing their land use. To control for this type of decision, we constructed an additional category composed of abandoned lands (AB). This land use type was assigned to some pixels using a transition rule after analyzing the sequence of land use decisions produced with the remotely sensed data and maximum likelihood supervised classification. We considered that a land use transition that lasts at most six years (roughly two observation intervals) from GC or PC to AG and

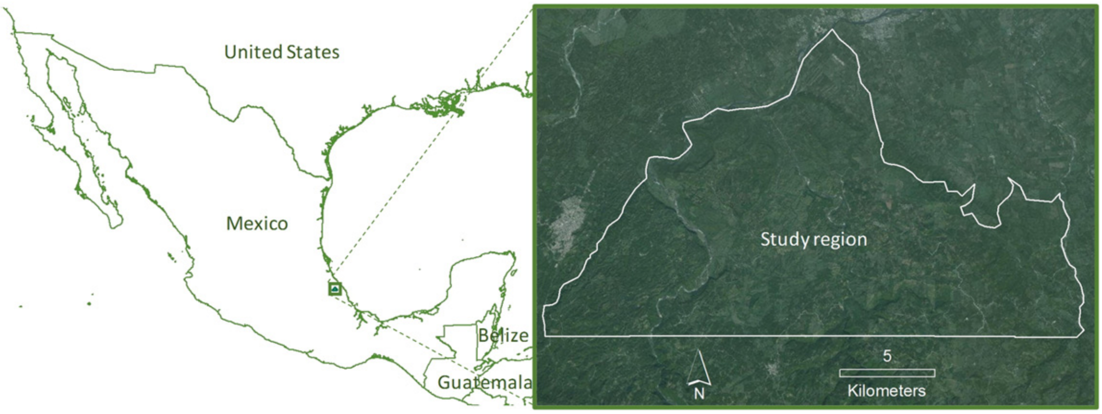

Fig. 2. Location of the study area (low altitude coffee growing region in Atzalan, Veracruz, Mexico).

then back to the previously observed land use indicates that that parcel was in fact abandoned during the period detected as AG. An example helps to clarify the procedure. Consider that the land use in parcel *s* is identified as GC during 1996, AG during 2000, and again GC in 2003. In general, this land use sequence is not logical either by economic or biological reasoning. In cases like this we consider that parcel *s* was in fact abandoned during 2000 and that the classifier algorithm categorized the land use as AG after detecting an increase in biomass that was likely generated because the landowner forwent maintenance activities in that parcel. Note that the transition GC–AG–GC is possible if AG is composed of only secondary forest. Nevertheless, secondary forest have been significantly reduced in the study region and the remaining portions are located in areas of difficult access with high slope that are not commonly used for agricultural purposes. Because temporary land use transitions between PC and AG represent less than 0.15% of the land use changes detected in the dataset; and given that land abandonment of those type of plantations is not common in the study region due to its significant impact on yield productivity, we focus our analysis on identifying misclassified observations in the AG, GC, and AB categories, as in the example.

To control for spatial autocorrelation, we generated a sample of spatially independent observations using a systematic random sampling procedure (see Dunn & Harrison, 1993 for a description of the method). Under such an approach each sampling point corresponds to a parcel with a land use value determined by the majority of the k-nearest neighboring cells. This is common in the discrete choice land use literature to approximate parcel-level land use data when parcel boundaries are not available (see for instance Blackman, Ávalos-Sartorio, & Chow, 2012; Chomitz & Gray, 1996; De Pinto & Nelson, 2008; Schmitt-Harsh, 2013). Here we set the neighborhood size k equal to 25 given that most of the small-scale farmers in this region own 1–2 ha parcels, and that the pixel size has a 30 m. resolution. This mechanism produced 210 sampling locations distributed across the study area. Fig. 3 shows the trends across the four land use categories in the sample data during the period 1984–2006, which is consistent with the trends observed in the complete dataset. The figure shows the decline in land allocated to AG, the increased proportion of PC, a slight decrease in the GC category and a more or less stable percentage of the land in AB status observed during the study period. The AG and PC proportions appear to follow complementary paths, i.e., at the time that one increases the other seems to decrease in a similar proportion. The same situation can be observed in trends corresponding to the GC and AB proportions. However, the data indicate that transitions occurred across all the land use categories and not exclusively within the classes with visually complementary paths.

3.2. Model description

There are undeniable complications in the transition rules that we use to construct the AB category. On the one hand, the procedure cannot be used to detect AG parcels that are in fact abandoned plots during any period. This is potentially a relevant issue, since Albers, Avalos-Sartorio, Batz, and Blackman (2006) report that at least 75% of farmers in a coffee growing region in Oaxaca, Mexico forwent maintenance activities during the coffee crisis period (1990–2004). On the other hand, the transition observed in some parcels between GC and AB may be part of a rotational production system used to recover soil productivity (Adiku, Kumaga, Tonyigah, & Jones, 2009; Kolawole, Salako, Idinoba, Kang, & Tian, 2005; Tian, Salako, Kolawole, & Kang, 1999). This means that it is possible that some of the parcels classified as AB are in fact GC fallowed as part of a rotational scheme and that the land use of those parcels has not actually changed. Alternatively, it is also possible that grasslands or cornfields with a relative increase in biomass are in fact parcels that have not received maintenance activities during the period in which the remotely sensed data was collected. Unfortunately, these types of misclassification problems cannot be addressed using algorithms based on spectral information or transition rules. Additionally, we cannot detect GC parcels that are AB in 1984. Nevertheless, we can use the LMNL model to estimate the probability that an AG parcel is actually abandoned as well as the probability that a parcel classified as AB is in fact a rotational GC plot.

The approach used to detect misclassified land use decisions is framed in the context of a discrete choice random utility model (see Ben-Akiva & Lerman, 1985; Train, 2009, for an in-depth review of the methodology and assumptions). These models posit that variations in socioeconomic, cultural and ecological factors influence land use changes through their impacts on the expected payoffs that landowners use to determine land use decisions (Chomitz & Gray, 1996; De Pinto & Nelson, 2008; Ellis et al., 2010; Lubowski, Plantinga, & Stavins, 2008). Let X_i represent a matrix of observable variables that determine the expected net revenue for each land use in the choice set J = {AG, PC, GC, AB} for agent i with $i = 1, \ldots, n$; β_j represent a vector of coefficients for the explanatory variables that affect the payoff of land use j; and α_j represent the constant term for alternative j; under the assumption that the unobservable components that determine land use j payoffs are independent extreme value type I (Gumbel) distributed variates, the probability of agent i selecting land use j, can be computed as

$$\Pr_{ij}(d_{ij} = 1 | \mathbf{X}_i, \boldsymbol{\beta}_j, \alpha_j) = \frac{e^{\alpha_j + \boldsymbol{\beta}'_j \mathbf{X}_i}}{\sum_{k \in J} e^{\alpha_k + \boldsymbol{\beta}'_k \mathbf{X}_i}} \forall j, k \in J$$

where $d_{ij} = 1$ if land use j is selected by agent i, and $d_{ij} = 0$ otherwise.

Defining $\tau \equiv \alpha_j \cup \beta_j \ \forall \ j \in \mathrm{J}$, the log-likelihood function under the assumption that all land use decisions N are accurately classified can be represented as:

$$\mathrm{LogL}(\tau) = \sum_{i=1}^{N} \sum_{j \in J} d_{ij} \ln \Pr_{ij} \quad \forall j \in \mathrm{J} = \{\mathrm{AG}, \mathrm{PC}, \mathrm{GC}, \mathrm{AB}\}.$$

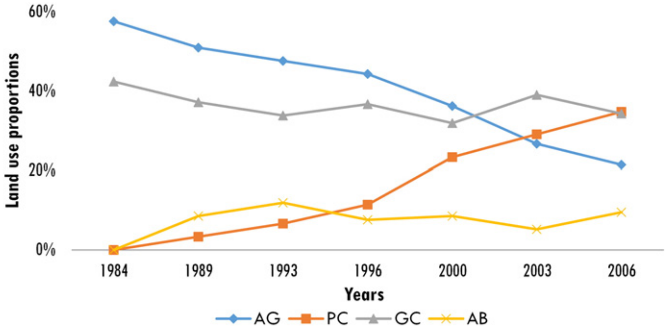

Fig. 3. Land use proportions in the sample data (1984–2006).

Considering that the set of parcels classified as AG may include a subset of misclassified AB parcels, and that this subset can have observations that should be in the GC category, following Caudill (2006) we can represent the log likelihood function using missing information indicators to represent the misclassification probabilities. Let $\mathrm{d}^*{}_{\mathrm{i,AG,AB}}$ indicate the probability that a land use observation in the AG category (branch) is actually a misclassified AB observation (stem), and $\mathrm{d}^*{}_{\mathrm{i,AG,AG}}$ represent the probability that it is accurately classified, thus satisfying the constraint $\mathrm{d}^*{}_{\mathrm{i,AG,AG}} + \mathrm{d}^*{}_{\mathrm{i,AG,AB}} = 1$; and similarly for $\mathrm{d}^*{}_{\mathrm{i,AB,AB}}$ and $\mathrm{d}^*{}_{\mathrm{i,AB,GC}}$. We can represent the log likelihood function as,

$$\mathrm{LogL}(\tau) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left(\begin{array}{l} \mathrm{d}^*{}_{\mathrm{i,AG,AG}} \mathrm{lnPr}_{\mathrm{i,AG,AG}} + \mathrm{d}^*{}_{\mathrm{i,AG,AB}} \mathrm{lnPr}_{\mathrm{i,AG,AB}} \\ + \mathrm{d}_{\mathrm{i,PC}} \mathrm{lnPr}_{\mathrm{i,PC}} \\ + \mathrm{d}_{\mathrm{i,GC}} \mathrm{lnPr}_{\mathrm{i,GC}} \\ + \mathrm{d}^*{}_{\mathrm{i,AB,AB}} \mathrm{lnPr}_{\mathrm{i,AB,AB}} + \mathrm{d}^*{}_{\mathrm{i,AB,GC}} \mathrm{lnPr}_{\mathrm{i,AB,GC}} \end{array} \right).$$

Because the probabilities of correct and incorrect classifications, $\mathrm{d}^*{}_{\mathrm{i,j,k}}$, are unknown we cannot identify the parameter estimates that maximize the log-likelihood function following the standard procedure. Nevertheless, we can replace the unknown probabilities by their conditional expectations (Caudill, 2006):

$$\mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AG,AG}} \middle| \mathrm{d}^*{}_{\mathrm{i,AG}}\right) = \frac{\exp\left(\alpha_{\mathrm{AG,AG}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AG}} \mathbf{X}_{\mathrm{i}}\right)}{\exp\left(\alpha_{\mathrm{AG,AG}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AG}} \mathbf{X}_{\mathrm{i}}\right) + \exp\left(\alpha_{\mathrm{AG,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AB}} \mathbf{X}_{\mathrm{i}}\right)}$$
$$\mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AG,AB}} \middle| \mathrm{d}^*{}_{\mathrm{i,AG}}\right) = \frac{\exp\left(\alpha_{\mathrm{AG,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AB}} \mathbf{X}_{\mathrm{i}}\right)}{\exp\left(\alpha_{\mathrm{AG,AG}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AG}} \mathbf{X}_{\mathrm{i}}\right) + \exp\left(\alpha_{\mathrm{AG,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AB}} \mathbf{X}_{\mathrm{i}}\right)}$$
$$\mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AB,AB}} \middle| \mathrm{d}^*{}_{\mathrm{i,AB}}\right) = \frac{\exp\left(\alpha_{\mathrm{AB,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AB,AB}} \mathbf{X}_{\mathrm{i}}\right)}{\exp\left(\alpha_{\mathrm{AB,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AB,AB}} \mathbf{X}_{\mathrm{i}}\right) + \exp\left(\alpha_{\mathrm{AB,GC}} + \boldsymbol{\beta}'{}_{\mathrm{AB,GC}} \mathbf{X}_{\mathrm{i}}\right)}$$
$$\mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AB,GC}} \middle| \mathrm{d}^*{}_{\mathrm{i,AB}}\right) = \frac{\exp\left(\alpha_{\mathrm{AB,GC}} + \boldsymbol{\beta}'{}_{\mathrm{AB,GC}} \mathbf{X}_{\mathrm{i}}\right)}{\exp\left(\alpha_{\mathrm{AB,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AB,AB}} \mathbf{X}_{\mathrm{i}}\right) + \exp\left(\alpha_{\mathrm{AB,GC}} + \boldsymbol{\beta}'{}_{\mathrm{AB,GC}} \mathbf{X}_{\mathrm{i}}\right)}$$

where $\mathrm{E}(\mathrm{d}^*{}_{\mathrm{i,AG,AG}}|\mathrm{d}^*{}_{\mathrm{i,AG}})$indicates the probability that parcel *i* classified as AG is actually an AG parcel, and $\mathrm{E}(\mathrm{d}^*{}_{\mathrm{i,AG,AB}}|\mathrm{d}^*{}_{\mathrm{i,AG}})$ represents the probability that a parcel classified as AG is in fact AB (the remaining conditional expectations have similar interpretations).

Defining

$$\begin{array}{l} \varphi \equiv \exp\left[\alpha_{\mathrm{AG,AG}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AG}} \mathbf{X}_i\right] + \exp\left[\alpha_{\mathrm{AG,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AB}} \mathbf{X}_i\right] + \\ \exp\left[\alpha_{\mathrm{PC}} + \boldsymbol{\beta}'{}_{\mathrm{PC}} \mathbf{X}_i\right] + \exp\left[\alpha_{\mathrm{GC}} + \boldsymbol{\beta}'{}_{\mathrm{GC}} \mathbf{X}_i\right] + \\ \exp\left[\alpha_{\mathrm{AB,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AB,AB}} \mathbf{X}_i\right] + \exp\left[\alpha_{\mathrm{AB,GC}} + \boldsymbol{\beta}'{}_{\mathrm{AB,GC}} \mathbf{X}_i\right] \end{array}$$

the probabilities that each observation is a member of each of the (now six) land use categories can be computed as

$$\mathrm{Pr}_{i,\mathrm{AG},j}\left(d_{i,\mathrm{AG},j} = 1 \middle| \mathbf{X}_i, \boldsymbol{\beta}_{AG,j}\right) = \frac{e^{\alpha_{\mathrm{AG},j} + \boldsymbol{\beta}'{}_{\mathrm{AG},j} \mathbf{X}_i}}{\varphi} \text{ for } j = \mathrm{AG, AB}$$
$$\mathrm{Pr}_{i,\mathrm{AB},k}\left(d_{i,\mathrm{AB},k} = 1 \middle| \mathbf{X}_i, \boldsymbol{\beta}_{\mathrm{AB},k}\right) = \frac{e^{\alpha_{\mathrm{AB},k} + \boldsymbol{\beta}'{}_{\mathrm{AB},k} \mathbf{X}_i}}{\varphi} \text{ for } k = \mathrm{AB, GC}$$
$$\mathrm{Pr}_{i,l}\left(d_{i,l} = 1 \middle| \mathbf{X}_i, \boldsymbol{\beta}_l\right) = \frac{e^{\alpha_l + \boldsymbol{\beta}'{}_l \mathbf{X}_i}}{\varphi} \text{ for } l = \mathrm{PC, GC}.$$

Under these modeling assumptions the log likelihood can be re-stated as,

$$\mathrm{LogL}\tau = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left(\begin{array}{l} \mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AG,AG}} \middle| \mathrm{d}^*{}_{\mathrm{i,AG}}\right) \ln \dfrac{\exp\left(\alpha_{\mathrm{AG,AG}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AG}} \mathbf{X}_{\mathrm{i}}\right)}{\varphi} + \\ \mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AG,AB}} \middle| \mathrm{d}^*{}_{\mathrm{i,AG}}\right) \ln \dfrac{\exp\left(\alpha_{\mathrm{AG,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AG,AB}} \mathbf{X}_{\mathrm{i}}\right)}{\varphi} + \\ \mathrm{d}_{\mathrm{i,PC}} \ln \dfrac{\exp\left(\alpha_{\mathrm{PC}} + \boldsymbol{\beta}'{}_{\mathrm{PC}} \mathbf{X}_{\mathrm{i}}\right)}{\varphi} + \\ \mathrm{d}_{\mathrm{i,GC}} \ln \dfrac{\exp\left(\alpha_{\mathrm{GC}} + \boldsymbol{\beta}'{}_{\mathrm{GC}} \mathbf{X}_{\mathrm{i}}\right)}{\varphi} + \\ \mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AB,AB}} \middle| \mathrm{d}^*{}_{\mathrm{i,AB}}\right) \ln \dfrac{\exp\left(\alpha_{\mathrm{AB,AB}} + \boldsymbol{\beta}'{}_{\mathrm{AB,AB}} \mathbf{X}_{\mathrm{i}}\right)}{\varphi} + \\ \mathrm{E}\left(\mathrm{d}^*{}_{\mathrm{i,AB,GC}} \middle| \mathrm{d}^*{}_{\mathrm{i,AB}}\right) \ln \dfrac{\exp\left(\alpha_{\mathrm{AB,GC}} + \boldsymbol{\beta}'{}_{\mathrm{AB,AB}} \mathbf{X}_{\mathrm{i}}\right)}{\varphi} \end{array} \right).$$

To avoid identification problems the parameters associated with the stems of the branches that contain misclassified information must be equivalent to the parameter estimates of the branches in which the parcels are accurately classified. Therefore we set $\boldsymbol{\beta}_{\mathrm{AG,AG}} = \boldsymbol{\beta}_{\mathrm{AG}}$; $\boldsymbol{\beta}_{\mathrm{AG,AB}} = \boldsymbol{\beta}_{\mathrm{AB}}$; $\boldsymbol{\beta}_{\mathrm{AB,AB}} = \boldsymbol{\beta}_{\mathrm{AB}}$; and $\boldsymbol{\beta}_{\mathrm{AB,GC}} = \boldsymbol{\beta}_{\mathrm{GC}}$. Additionally, Caudill (2006) highlights the relevance of the intercepts in the model since as $\alpha_{AG,AB} \rightarrow -\infty$the probability of identifying abandoned parcels that are misclassified as agroforestry goes to zero. Similar reasoning applies when $\alpha_{AB,GC} \rightarrow -\infty$. To test that the LMNL model can be used to detect misclassified observations, we estimate profile likelihood confidence intervals for those intercepts to check that they are statistically different from -∞. This also constitutes statistical evidence that the related branch has misclassified parcels. To compute profile likelihood confidence intervals for the intercepts $\alpha_{AG,AB}$ and $\alpha_{AB,GC}$ we use a grid search procedure described by Stryhn and Christensen (2003). The lower and upper bounds of a profile likelihood confidence interval for a parameter $\alpha_{j,k}$ satisfy the equation $LogL(\tau^*) - \frac{1}{2}\chi_1^2(0.95) \leq LogL(\tau_0)$, where τ^* is the maximum likelihood estimate of τ, $\chi_1^2(0.95)$indicates the 95% quantile of a chi-squared distribution with one degree of freedom, and τ_0 is a vector that contains the MLE of τ obtained after setting the parameter of interest to a fixed value x (i.e., $\alpha_{j,k} = x$), and treating the remaining parameters in the model as nuisance parameters.

The procedure to determine the probabilities of misclassified data and to compute the parameter estimates that maximize the likelihood function follows these steps:

1. Control for local maxima
   Set a global solver or grid search algorithm to define vectors of initial values for the alternative specific parameters that will be estimated. This step is necessary because this LMNL modeling approach is similar to a finite mixture model (Caudill et al., 2011), and thus during the computation of the parameter estimates we need to control for multiple local maxima of the likelihood function.
2. Expectation step
   Use the observed data $\mathbf{X}_{\mathrm{i}}$ and one of the vectors estimated in step 1 as initial values of the parameter estimates $\tau^{(0)}$ to compute the conditional expectations of the misclassified and accurately classified land use proportions, $\mathrm{d}_{\mathrm{i,jk}}{}^*$.
3. Maximization step
   Estimate the vector of parameters that maximize the likelihood function, τ^*, and the corresponding value of the likelihood function at that point $\mathrm{LogL}(\tau^*)$.
4. Iterate between the expectation and maximization steps using τ^* to update the conditional expectation of $\mathrm{d}_{\mathrm{i,jk}}{}^*$ and utilizing those values

to re-compute τ^* until the log-likelihood function convergences to a maximum value within a certain tolerance level.

5. Return to step 1 and repeat the process for a different vector of initial values $\tau^{(0)}$ until exhausting the set of defined vectors in step 1.
6. Identify the τ^* that produces the global maximum from the set of evaluated starting values.

3.3. Land use drivers

3.3.1. Revenue

Baerenklau et al. (2012) observe that a significant proportion of the agents in the study region replaced their coffee farms for citrus or banana plantations in response to low coffee prices. Given this evidence of price responsiveness, we use time series data on average market prices per ton of coffee, lemon, orange, tangerine, mandarin, grapefruit, banana, livestock, and corn received by farmers at the state level (SAGARPA, 2012) to construct land use-specific price indices. We also use historical productivity data (SAGARPA, 2012) and information from agronomists working in the study region to estimate the average productivity per hectare of shade grown coffee, banana, citrus, pasture, and corn. Price and productivity data is then used together to generate weighted revenue indexes for the land use categories considered in this study. Given that there is not commercial use of forested lands in the study region, and that the main component of the agroforestry production system is coffee, for the AG category we use the yearly average rural price per-ton received by coffee growers multiplied by the average productivity per hectare in coffee plantations to estimate an annual revenue index for this category. On the other hand, since the PC category is comprised of different citrus varieties, as well as banana plantations, we followed a two-step procedure to construct a price index for this category. In the first step prices of citrus varieties harvested in the study region were used to construct a weighted average price per-ton, with weights set according to the area harvested for each citrus type. Similar to the procedure followed to generate the revenue index for the AG category, we multiplied the citrus price index by the average productivity per hectare observed in the study area for this type of plantation to obtain an estimate of the average revenue per hectare. In the second step, a similar weighting process was implemented to merge this revenue index for citrus with time series data on yearly average revenue per hectare for banana plantations.

A different procedure was used to construct the price index corresponding to the GC category. Agricultural activities in the study area are undertaken with labor- and land-intensive production technologies that have not been significantly modified in decades. This is particularly true for cornfields and grasslands in which it is fair to assume that on average farmers get the same amount of grain and weight gain of livestock per hectare independently of the age of the land use. Therefore we use the average productivity of corn plantations (SAGARPA, 2012) and the average livestock weight gain per hectare observed in unfertilized grasslands in the state of Veracruz, Mexico (Tergas & Sanchez, 1979) to construct a per hectare weighted yearly revenue index for the GC category. Furthermore, considering that in the study area one person can complete all the required maintenance activities for a 2-hectare parcel without needing to hire additional labor, we homogenize the revenue indexes across all land use categories by assuming that each parcel in the sample data measures 2 ha.

Given the low educational level of farmers in the study region, few off-parcel employment options are available. Besides working land owned by other people, the most common alternative is to look for employment opportunities in Mexico City or as an illegal worker in the United States (Nava-Tablada & Martínez-Camarillo, 2012). Since the AB category does not involve crop production, to account for the monetary reward received by a farmer who decides to abandon his land we use the yearly minimum wage for construction workers.

3.3.2. Transportation costs

There are three main regional market centers in the proximity of the study area at which farmers can sell their products. Those three markets have similar prices for the produce generated from the land use categories under analysis. To compute the distance from each parcel to the nearest market we followed a three-stage process. First, the Euclidean distance from each sample parcel to the nearest road was computed using vector data (INEGI, 1999). Second, by using the network analysis ArcGIS extension and vector data of the road network in the area, we computed the most efficient route (in terms of distance) from each parcel's nearest road to each market center. Finally, the distances to each market were compared and the shortest was selected. This variable is assumed to be constant since the road network was not significantly changed during the period of analysis, despite improvements to the conditions of some of the main roads (e.g., changing from dirt roads to paved roads) that potentially reduced driving time but not driving distance to each market.

3.3.3. Socioeconomic land use drivers

Starting in 1995, every five years the Mexican Government computes a poverty index that uses data on education accessibility, housing conditions and monetary income at the community level. This index in general ranges from −2.37 to 4.49, with lower values corresponding to a better welfare status (CONAPO, 2006). A review of the statistics generated by CONAPO (1998, 2006, 2011) indicates that the poverty level in the 104 communities located either within the study area or up to 500 m outside its boundary, has not fluctuated significantly during the period 1995–2010. Considering the apparent static behavior of such variables, and given that data is unavailable for all the observation years, we used the 2005 version of the index to generate an interpolated surface using the Inverse Distance Weighting (IDW) method. This approach captures the effect of spatial differences in poverty on land use decisions. Statistics from CONAPO (1998, 2006, 2011) also are used to generate a population index because human settlements tend to generate more pressure on their surrounding environment and at the same time provide more labor to harvest the land. This index also is treated as static for each location (again using 2005 data) because the data indicate that the number of inhabitants in most of the communities has not significantly changed during the study window. Since population pressures diminish as the distance to the settlement increases we again use IDW interpolation to estimate values at the sample parcels.

3.3.4. Topographic land use drivers

To account for the effects of topographic variables in the land use decision process we use vector data of elevation level curves obtained from INEGI (1998) to construct a digital elevation model that was used to generate slope and elevation information. Finally, soil texture information from SEMARNAP (1998) was used as a proxy of soil quality. Table 1 presents a summary of the mean, minimum and maximum values of the land use drivers considered in the analysis.

4. Results and discussion

The model was implemented within the Matlab environment setting the coefficients of the PC category equal to zero for identification purposes. Table 2 shows the parameter estimates ordered by branches and stems as well as the sum of the probabilities in each stem that indicates the estimated number of observations accurately and inaccurately classified within each branch. For the AG and AB branches, the first (second) stem shows the number of observations and parameters estimates for the accurately classified (misclassified) observations. Recall that the coefficient estimates for the AB and GC stems are invariant to classification errors, as displayed in the table.

Table 1
Summary statistics for the parcel specific variables.

Variable	Description	Mean	Min	Max
AG Revenue	Mexican pesos (base 2000)	13,392	4,999	22,521
PC Revenue		26,376	12,825	57,645
GC Revenue		12,580	8,506	19,337
AB Revenue		16,020	10,730	33,552
Elevation	Meters above sea level	354	85	726
Slope	Degrees	10.49	0	60.09
Poverty	Index that uses education accessibility, housing conditions and monetary income data to measure the degree of poverty with lower values corresponding to a better welfare status	0.316	−0.798	2.109
Population	Index to measure labor availability	263	30	793
Soil texture	Soil texture of parcel (1 = fine, 2 = medium, 3 = coarse)	1.34	1.00	3.00
Distance to road	Euclidean distance from each parcel to the nearest road (m)	389	0	1,779
Distance to nearest market	Distance from each parcel to nearest market (km)	14.36	2.93	35.52

Overall the results indicate that an estimated 11% of the observations contained in the sample are misclassified. A total of 52 observations that are categorized as AG in the sample are more likely AB parcels. Those observations represent 8.7% of the parcels originally classified as AG during the study period. Similarly, the results indicate that the procedure used to construct the AB category is suspect because all the observations in the AB branch - AB stem are considered misclassified by the LMNL procedure. In other words, the analysis provides evidence that parcels that appear to be AB are actually part of a GC rotational production system, or are parcels that continue under cultivation but that did not receive maintenance activities during the time of the remotely sensed data collection.

To test whether the classification errors are statistically significant we compute profile likelihood confidence intervals for the intercepts $\alpha_{AG,AB}$ and $\alpha_{AB,GC}$ using the aforementioned Stryhn and Christensen (2003) grid search procedure. That procedure identifies the values of α for which the inequality $LogL(\tau^*) - \frac{1}{2}\chi_1^2(0.95) \leq LogL(\tau_0)$ holds. The profile likelihood confidence interval for $\alpha_{AG,AB}$ is [−17.1, 15.2] and for $\alpha_{AB,GC}$ is [−1.29, 0.77]. Clearly these intervals are bounded away from −∞, which provides evidence that the number of misclassified observations is statistically greater than zero. Fig. 4 shows the profile likelihood confidence intervals for both parameters of interest.

A depiction of the differences between the land use proportions in the sample data and the percentages estimated with the LMNL model is presented in Fig. 5. The results indicate that the AG category is overrepresented in the sample throughout the study period due to the presence of misclassified observations. On the other hand, the GC category is underrepresented in the sample since it should contain all the observations categorized as AB.

For the same reason, the AB category appears to be overrepresented throughout the period of analysis. A potential explanation for this finding is that small-landowners that rely primarily on household labor are less likely to abandon their plantations (Albers et al., 2006) specially if the current land use provides means to satisfy household subsistence constraints. To analyze the impacts of misclassified observations on the magnitudes and directions of the parameter estimates we use the original sample dataset and the reconstructed (corrected) sample based on the LMNL analysis to estimate a standard multinomial logit model of land use decisions. Table 3 shows the estimated coefficients, significance levels and standard errors. Overall the significance levels and values of the AG and GC parameter estimates are similar in the analysis of the two sample datasets. The values of the coefficients associated with the AB category appear to be significantly different in magnitude and in some cases the signs change using the LMNL-corrected sample. Given the significant reconfiguration of the AB category the difference in the corresponding parameter estimates is expected. Furthermore, McFadden's pseudo r-squared increases from 0.16 to 0.29, which is a significant improvement (McFadden, 1978).

To understand how changes in the independent variables affect land use proportions, we compute the change in the probability of observing

Table 2
Latent multinomial logit model parameter estimates.

Branch	Agroforestry		Abandoned		Grass and corn
Stem	Agroforestry	Abandoned	Abandoned	Grass and corn	Grass and corn
Land use observations	547	52	0	108	536
Revenue	0.1184 ***	0.1549 ***		0.1002 ***	
	(9.01)	(5.05)		(5.22)	
Slope	0.3549 ***	9.2943 ***		−0.2077	
	(4.14)	(5.30)		(−2.14)	
Distance to market	0.3760 **	66.6943 ***		0.9913 ***	
	(1.93)	(5.51)		(5.74)	
Distance to nearest road	1.3163 ***	46.7071 ***		1.2370 ***	
	(4.07)	(5.49)		(3.94)	
Poverty	0.0835	−156.9003		−0.1447	
	(0.44)	(−4.89)		(−0.70)	
Soil texture	0.0875	−180.2229		−0.5086	
	(0.48)	(−6.12)		(−3.18)	
Elevation	5.1065 ***	−221.7704		−2.4227	
	(7.12)	(−4.87)		(−3.50)	
Population	−0.2368	25.3460		−0.4954	
	(−3.34)	(4.78)		(−6.84)	
Constant	−3.3241	−9.0654		0.9822 **	
	(−5.97)	(−0.11)		(1.84)	

Notes: the parameter estimates are shown in bold numbers; the t-ratios are shown in parentheses. Significance codes: '***' significant at the 1% level; '**' significant at the 5% level; '*' significant at the 10%. For model identification the coefficients of stems with potential misclassified observations are equal to the coefficients of the branch-stem in which those observations should be classified.

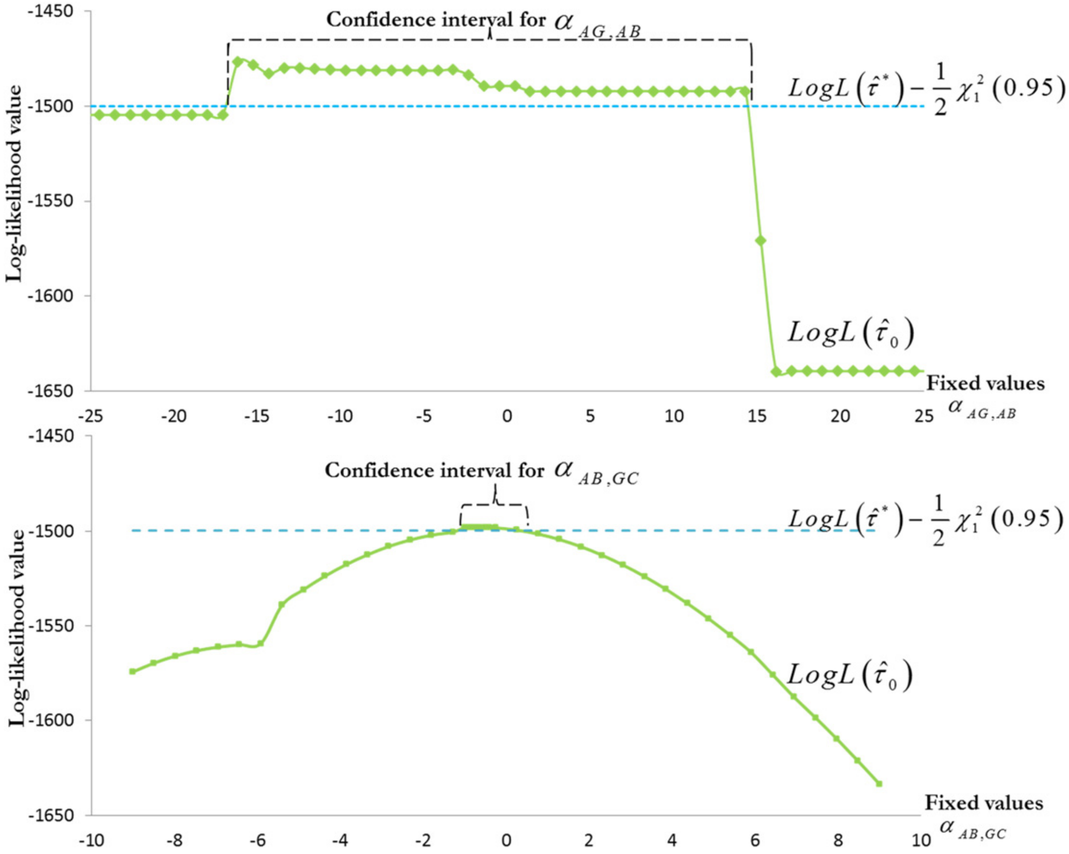

Fig. 4. Profile likelihood confidence intervals for $\alpha_{AG,AB}$ and $\alpha_{AB,GC}$.

land use j at each parcel i resulting from a marginal change in the observed magnitude of each of the independent k variables. The individual calculations are averaged across parcels and land uses and the results are shown in Table 4. In general, most of the marginal effects estimated with the two datasets have the expected directions. According to the analysis there is statistical evidence to argue that parcels with higher degrees of slope will be more likely to be used for agroforestry production, and areas with low slope are preferred for cornfields or grasslands.

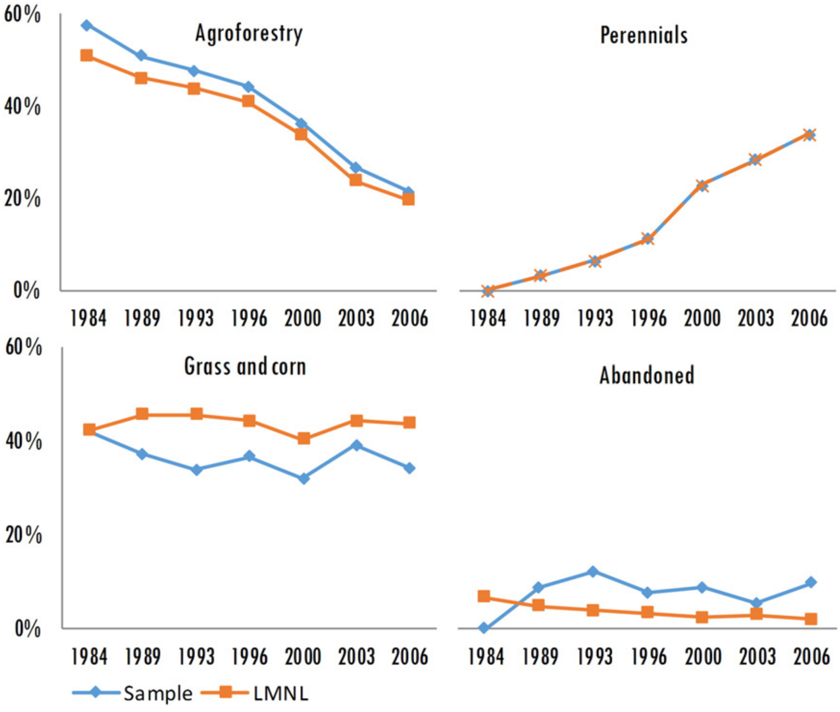

Fig. 5. Land use proportions in the sample data and estimated proportions using the LMNL model.

Table 3
Multinomial logit parameter estimates using the original sample data and the reconstructed sample data generated with the LMNL model.

		Original sample		Reconstructed sample		Difference
		Estimate (A)	Std. error	Estimate (B)	Std. error	Estimate (A–B)
Slope	AG	0.3789	0.0841 ***	0.3505	0.0848 ***	0.0284
	GC	−0.2557	0.0946 ***	−0.1956	0.0937 **	−0.0601
	AB	−0.1833	0.1389	3.1100	0.6253 ***	−3.2934
Distance to market	AG	0.5782	0.1685 ***	0.3214	0.1784 *	0.2567
	GC	0.8627	0.1713 ***	1.0540	0.1805 ***	−0.1913
	AB	0.9732	0.2086 ***	21.1698	4.2179 ***	−20.1966
Distance to road	AG	1.4016	0.3185 ***	1.3174	0.3267 ***	0.0841
	GC	1.2489	0.3119 ***	1.2305	0.3118 ***	0.0184
	AB	1.0312	0.3852 ***	16.1232	3.1221 ***	−15.0920
Poverty index	AG	−0.0154	0.2328	0.1164	0.2320	−0.1318
	GC	−0.2802	0.2344	−0.1733	0.2267	−0.1069
	AB	0.7179	0.3157 **	−45.4773	9.7809 ***	46.1952
Soil texture	AG	0.0896	0.1642	0.0728	0.1656	0.0168
	GC	−0.5437	0.1655 ***	−0.5168	0.1610 ***	−0.0268
	AB	−0.3428	0.2533	−62.3399	2366.38	61.9971
Elevation	AG	4.6620	0.7143 ***	5.2670	0.7357 ***	−0.6051
	GC	−1.9084	0.7206 ***	−2.6215	0.7295 ***	0.7130
	AB	−2.7255	1.1215 **	−64.3404	14.3039 ***	61.6149
Population	AG	−0.2124	0.0664 **	−0.2149	0.0674 ***	0.0025
	GC	−0.4901	0.0739 ***	−0.5067	0.0734 ***	0.0165
	AB	−0.5253	0.1291 ***	6.7800	1.5575 ***	−7.3054
Revenue	AG	0.1090	0.0127 ***	0.1196	0.0132 ***	−0.0105
	GC	0.0852	0.0188 ***	0.1013	0.0192 ***	−0.0161
	AB	−0.0123	0.0114	0.1846	0.0384 ***	−0.1969
Constant	AG	−3.3421	0.5539 ***	−3.3805	0.5630 ***	0.0384
	GC	1.2527	0.5424 **	1.1672	0.5365 **	0.0855
	AB	0.5276	0.8060	2.1882	2366.36	−1.6605
Log-likelihood:		−1492.4		−1187.4		
McFadden R^2:		0.16417		0.2939		

Notes: the coefficients of the Perennial Crops category were normalized to zero for model identification. Significance codes: '***' significant at the 1% level; '**' significant at the 5% level; '*' Significant at the 10%.

Table 4
Average marginal effects.

			Original sample (A)		Reconstructed sample (B)		Difference (A–B)
		Expected sign	Estimate	Standard error	Estimate	Standard error	
Slope	AG	+	0.103	0.032	0.077	0.056	0.026
	GC	−	−0.085	0.028	−0.089	0.067	0.003
	AB	±	−0.011	0.009	0.020	0.099	−0.031
	PC	−	−0.007	0.021	−0.009	0.021	0.002
Distance to nearest market	AG	−	−0.007	0.037	−0.116	0.275	0.109
	GC	+	0.070	0.033	0.071	0.442	0.000
	AB	+	0.021	0.015	0.132	0.647	−0.111
	PC	−	−0.084	0.051	−0.087	0.067	0.003
Distance to nearest road	AG	−	0.103	0.063	0.049	0.216	0.054
	GC	+	0.052	0.055	0.004	0.325	0.048
	AB	+	−0.005	0.012	0.097	0.473	−0.102
	PC	−	−0.150	0.089	−0.150	0.093	0.000
Poverty index	AG	−	0.013	0.014	0.144	0.601	−0.131
	GC	+	−0.076	0.038	0.130	0.946	−0.207
	AB	±	0.056	0.044	−0.294	1.440	0.350
	PC	−	0.008	0.008	0.020	0.098	−0.012
Soil texture	AG	+	0.084	0.032	0.215	0.821	−0.131
	GC	−	−0.104	0.031	0.139	1.296	−0.242
	AB	±	−0.006	0.009	−0.402	1.967	0.396
	PC	+	0.025	0.025	0.048	0.135	−0.023
Elevation	AG	+	1.140	0.354	1.378	0.936	−0.238
	GC	±	−0.793	0.303	−0.846	1.490	0.053
	AB	±	−0.208	0.137	−0.419	2.048	0.211
	PC	−	−0.140	0.239	−0.113	0.325	−0.027
Population	AG	+	0.025	0.021	0.007	0.097	0.019
	GC	−	−0.054	0.020	−0.092	0.145	0.039
	AB	−	−0.013	0.009	0.046	0.227	−0.059
	PC	+	0.041	0.026	0.039	0.033	0.002
Revenue	AG	+	0.011	0.005	0.009	0.006	0.003
	GC	+	0.005	0.006	0.003	0.006	0.001
	AB	+	−0.006	0.005	0.001	0.003	−0.006
	PC	+	−0.010	0.006	−0.013	0.008	0.002

Expected sign codes: '+' indicates that a positive marginal effect is expected, '–' indicates that a negative marginal effect is expected, '±' indicates that the marginal effects can go in either direction.

Table 5
First order autoregressive parameters used to estimate revenue paths.

Parameter	AG	PC	GC
Unconditional mean	12,058	4,796	21,480
Autocorrelation coefficient	0.7183	0.8695	0.9602
Standard deviation of the error	4439	1302	2146

The average marginal effects of the distance from a parcel to the nearest markets are statistically significant and have the expected signs. The probability of observing cash crops (AG or PC) decreases as the distance to a market increases.

On the other hand, the likelihood of an agent selecting the GC or AB category increases as the distance to the nearest market increases, which is consistent with the intuition that if a parcel is located far away from a market, transportation costs may reduce the profitability of some of the land uses thus limiting the choice set to subsistence crops (such as corn), or to land uses that require a large contiguous area (such a cattle ranching activities), or to land abandonment. A similar explanation applies to the average marginal effects of the variable measuring the distance from a parcel to the nearest road. Notably, these marginal effects have the expected directions only for GC and PC in the original sample, but for GC, AB, and PC in the reconstructed dataset.

The results corresponding to the poverty index are statistically significant only for the AB category. We would expect that richer areas have higher probability of selecting cash crops although this is not reflected in the results from the LMNL dataset. None of the parameter estimates for soil texture are statistically significant, which may reflect the difficulty in determining expected signs for all but the GC category (which should correlate with finer soils). All the parameter estimates for the elevation variable are statistically significant and the directions of the marginal effects of the AG and PC categories are consistent with the agroecological requirements of the crops in those land use classes. Because corn and grass can be produced in parcels located at different elevation gradients, the direction of the marginal effects could go in either direction depending on the location of the parcels in the dataset. The results for the original and reconstructed sample data indicate an inverse relationship between elevation and the probability of observing GC and AB. The parameter estimates corresponding to the population variable are statistically significant and the marginal effects have the expected signs indicating that higher population density may increase the probability of observing labor intensive land uses and vice versa.

Perhaps the most empirically relevant results are related to the statistical significance of the estimated coefficients of the revenue variables and the signs of the corresponding marginal effects across the two samples. The parameter estimates computed with the original dataset, that contains misclassified observations, are statistically significant at the 1% level for the AG and GC category and the marginal effects have theoretically consistent signs. However, the marginal effects of changes in revenue on the probability of an agent selecting the AB or PC categories indicate a counterintuitive direction. Those inconsistencies appear to be partially corrected in the reconstructed dataset using the results of the LMNL model. Specifically, the sign of the revenue-related marginal effect for the AB category has the expected sign although the multinomial logit model still cannot produce theoretically consistent parameter estimates for the PC category. A possible though speculative explanation is that this could be related to the associated price index which includes a variety of different tree crops in the calculation.

5. Model validation

The preceding are promising results but our empirical dataset does not allow us to validate the reconstructed sample due to lack of appropriate reference data. Therefore, to more rigorously test the performance of the LMNL model, we construct a simulated dataset. We simulate parcel specific characteristics and revenue data associated with four land use categories, and assume that unobservable land use drivers are independent and identically distributed extreme value type I variables. For consistency, our simulation uses the same land categories described in our empirical analysis, and the explanatory variables listed in Table 1. We use mean and standard deviation values from those variables to simulate location-specific characteristics defining a set of 500 artificial parcels. We simulate elevation, slope, population pressure, poverty, distance to the nearest road, and distance to the nearest market using pseudo-random draws from normal distributions fitted to our empirical dataset. To simulate soil texture values we use a discrete pseudo-random number generator constrained to the interval 1–3. To simulate annual revenue data for each of the four land use categories we estimate first-order autoregressive processes using our time series of revenue indices (results are shown in Table 5). For each land use category, the corresponding autoregressive equation is used to generate 100 revenue paths, each composed of 20 periods.

Table 6
Parameter estimates used to simulate land use decisions.

	Land use	Data Generating Parameters		Land use	Data Generating Parameters
Revenue	AGF	5.2563	Distance to the nearest market	AGF	−1.2901
	GC	2.0250		GC	0.6225
	AB	1.3141		AB	1.0927
Elevation	AGF	4.1728	Poverty	AGF	−4.2513
	GC	0.5218		GC	3.5240
	AB	−0.8536		AB	−5.4651
Slope	AGF	0.8163	Population	AGF	−0.0250
	GC	−0.6246		GC	−0.0540
	AB	0.6486		AB	0.0380
Soil texture	AGF	0.4176	Intercept	AGF	−11.7526
	GC	−1.9131		GC	−8.6200
	AB	−6.1708		AB	−20.296
Distance to the nearest road	AGF	−1.6161			
	GC	5.0377			
	AB	0.6053			

For each "parcel" $i = 1, 2, \ldots, 500$ and for each revenue path $r = 1, 2, \ldots, 100$, land use is estimated in each period $t = 1, 2, \ldots, 20$ using a standard multinomial logit model with randomly generated parameter values (shown in Table 6) that produce theoretically consistent marginal effects and land use proportions that mimic our empirical data (on average 34% of the simulated parcels were classified as agroforest, 28% as tree crops, 22% as grass and corn, and 16% as abandoned lands). This produces 1 million simulated land use decisions. To simulate misclassified land use observations, we next create three new datasets by randomly reclassifying 25%, 60%, and 95% of the "true" abandoned lands as agroforests. The LMNL is then applied to each dataset to test its ability to identify the misclassified observations and reconstruct the original dataset. The LMNL model estimation required around 34 h on a six-core 3.74 GHz Intel machine with 16 GB RAM, to complete the analysis at each misclassification level.

A useful baseline for contrasting the performance of the LMNL model can be established by converting the category-specific misclassification levels into global misclassification levels. The 25% misclassification of abandoned lands represents a global error of 3.94%. This error rate increases to 9.39% when 60% of those observations are misclassified, and reaches 14.87% at the 95% misclassification level. On average across all the simulations, the LMNL algorithm reduced these global errors to 1.26%, 1.34% and 1.41% respectively. The overall accuracy, and the user's and producer's accuracy for the abandoned lands category, during each of the one-hundred revenue path simulations are shown in Fig. 6. At the 25% misclassification level, the overall accuracy values vary within the interval 0.973–0.996 with a mean value of 0.987. The user's accuracy range from 0.920 to 0.987 with a mean of 0.963. Producer's accuracy values are observed in the interval 0.807–0.980 with an average of 0.923. The figure shows similar results for the 60% and 95% misclassification levels.

To further assess the performance of the LMNL model, the confusion matrices across all iterations were aggregated (Table 7). With that information we estimate Cohen's kappa values using only observations in the AG and AB categories (Table 8). We exclude observations classified as TC and GC, since we assume that those categories are correctly classified. Inclusion of those observations would further increase the reported

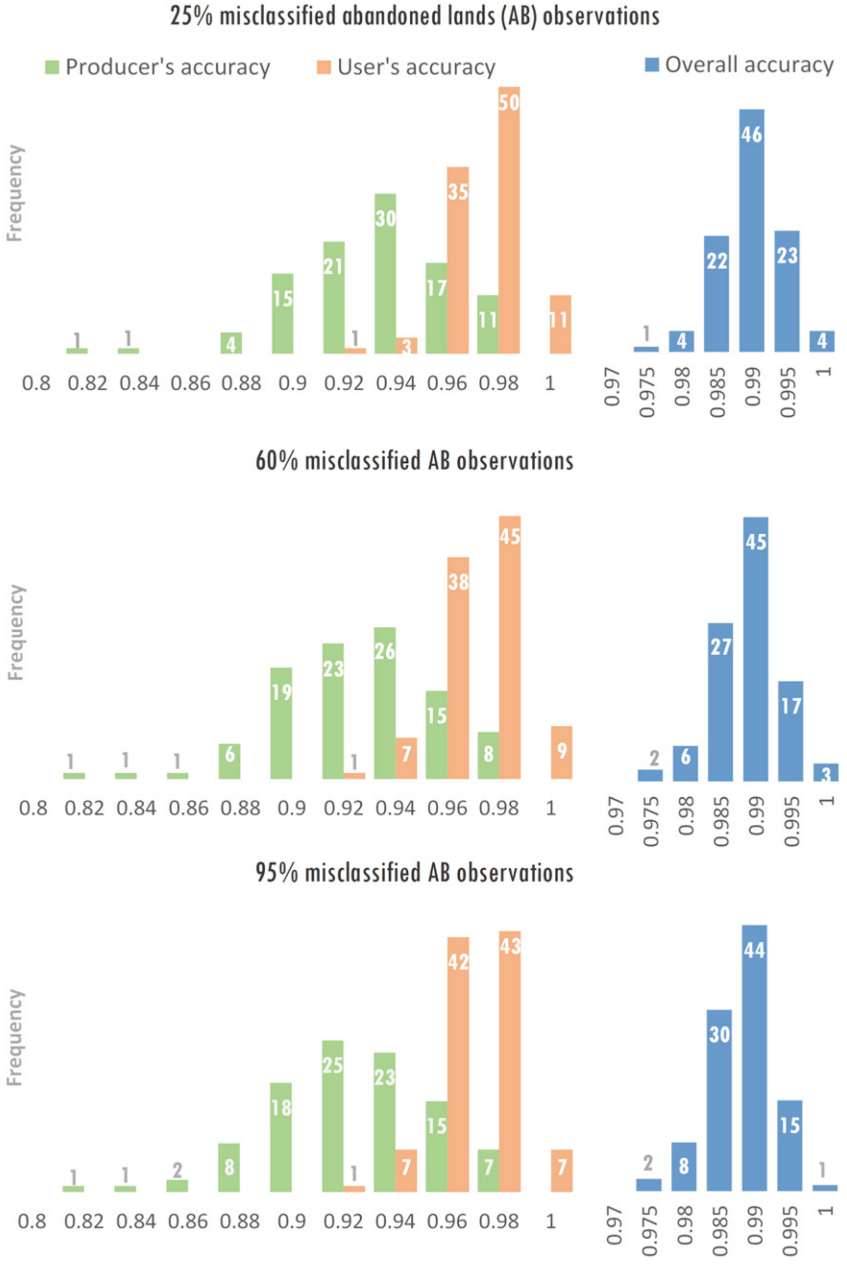

Fig. 6. Aggregated user's accuracy, producer's accuracy and overall accuracy for abandoned lands.

accuracy values. Similar to the results in Fig. 6, the Kappa statistic is almost the same across all t misclassification levels.

These results overwhelmingly validate the ability of the LMNL model to identify randomly misclassified parcels. Our modeling assumptions are standard for discrete choice land use models (Chomitz & Gray, 1996; De Pinto & Nelson, 2008; Ellis et al., 2010; Lubowski et al., 2008), and the results are essentially independent of the error rate in the misclassified category.

6. Conclusions

Given the limited availability of historical high resolution remotely sensed data, land use change analyses are often restricted to the study of transitions between a reduced set of choices. In some cases coarse datasets are enough to accomplish relevant research objectives, for instance in the study of deforestation processes. Nevertheless, in most of the spatially explicit land use analyses coarse land use classifications are implemented as a mechanism to reduce classification errors. Unfortunately, even in land use datasets composed of a reduced number of categories, misclassifications are still a potential modeling problem. Consider for example an analysis that uses only two categories, forested and agricultural lands, to study deforestation drivers in a particular region. In this case it is possible that some of the observations classified as forested areas are in fact fallowed parcels devoted to agricultural production, or even grasslands that have not received weed control activities during the time of data collection of the remotely sensed data. Unfortunately, those types of classification errors are difficult to reduce using only pixel-based algorithms, particularly if the available land use information is part of a time series dataset with many years of separation between the observed periods.

To reduce classification errors, this article implements a post-classification procedure to identify misclassified land use observations that cannot be detected using pixel-based classification algorithms. The Latent Multinomial Logit methodology has been implemented in several contexts to detect misclassified categorical data (Caudill, 2006; Caudill & Mixon, 2005; Caudill et al., 2005, 2011) but to our knowledge it has not been applied in the land use change literature. The analysis implemented here is based on land use information generated with remotely sensed data collected during seven points in time throughout the period 1984–2006, with a maximum separation of five years between observations. The data correspond to land use transitions observed in a Mexican coffee growing region in which relatively high rates of tree canopy removal were observed as a result of the clearing of shade-grown coffee plantations. We analyze land use dynamics between agroforestry parcels, perennial crops, grass and corn, and abandoned land. The category corresponding to abandoned lands was constructed analyzing the sequence of land use decisions observed in each parcel and assigning a parcel to the abandoned land category when the land use oscillated between grass and corn or perennial crops, and agroforests within a period of at most six years.

The implementation of the LMNL model provides statistical evidence to argue that the procedure used to construct the abandoned land category, while reasonable and objectively defensible, fails to recognize that temporary increases in biomass that appear to indicate a change in the corresponding land use classification to agroforestry may instead be the result of a production system that requires land fallowing as a mechanism to recover soil productivity; or simply an indication that the parcel has not been maintained during the time in which the remotely sensed data in that region were collected. The results also indicate that the LMNL procedure can be used to identify parcels within the agroforestry category that have a high likelihood of being abandoned without making any assumptions about the land use sequence followed by each landowner. With regard to the impact on the values and magnitudes of the parameter estimates and marginal effects, we can observe that in general the reclassification of the parcels based on the LMNL model increases the magnitudes of the marginal effects in the theoretically expected direction. Particularly, the marginal effect of changes in revenue associated with the abandoned land category becomes statistically significant with the theoretically expected sign.

Finally, the performance of the algorithm is assessed using artificially misclassified datasets generated through Monte Carlo simulations. The LMNL model is able to reconstruct the "true" dataset almost entirely, regardless of the error level in the misclassified category. Overall these results strongly suggest that the LMNL approach is a highly effective and beneficial method for controlling for misclassified land use data.

Table 7
Aggregated confusion matrices at different classification error levels.

% of misclassified AB observations	Land uses	AG	TC	GC	AB	User's accuracy
25%	AG	331,874	0	0	12,591	0.9634
	TC	0	279,116	0	0	1
	GC	0	0	219,864	0	1
	AB	5	0	0	156,550	0.9999
	Prod.'s Acc.	0.9999	1	1	0.9256	
60%	AG	331,106	0	0	13,359	0.9612
	TC	0	279,116	0	0	1
	GC	0	0	219,864	0	1
	AB	13	0	0	156,542	0.9999
	Prod.'s Acc.	0.9999	1	1	0.9214	
95%	AG	330,434	0	0	14,031	0.9593
	TC	0	279,116	0	0	1
	GC	0	0	219,864	0	1
	AB	21	0	0	156,534	0.9999
	Prod.'s Acc.	0.9999	1	1	0.9177	

References

Adiku, S.G.K., Kumaga, F.K., Tonyigah, A., & Jones, J.W. (2009). Effects of crop rotation and fallow residue management on maize growth, yield and soil carbon in a savannah-forest transition zone of Ghana. *The Journal of Agricultural Science*http://dx.doi.org/10.1017/S002185960900851X.

Albers, H.J., Avalos-Sartorio, B., Batz, M.B., & Blackman, A. (2006). Maintenance costs, price uncertainty, and abandonment in shade-grown coffee production: Coastal Oaxaca, Mexico. *environmental and resource economists 3rd world congress* (pp. 39) (http://www.webmeets.com/ERE/WC3/Prog/).

Andersen, L.E. (1996). The causes of deforestation in the Brazilian Amazon. In R. Dickenson (Ed.), *The Journal of Environment Development*, *5 (3)*. (pp. 309–328). World Bank Publications. http://dx.doi.org/10.1177/107049659600500304.

Aoki, K., & Suvedi, M. (2012). Coffee as a livelihood support for small farmers: A case study of Hamsapur Village in Nepal. *Journal of International Agricultural and Extension Education*, *19*, 16–29 (http://www.scopus.com/inward/record.url?eid=2-s2.0-84870033895&partnerID=40&md5=549570985480850ce41bbb87a8521b3c).

Ávalos-Sartorio, B., & Blackman, A. (2010). Agroforestry price supports as a conservation tool: Mexican shade coffee. *Agroforestry Systems*, *78*(2), 169–183. http://dx.doi.org/10.1007/s10457-009-9248-4.

Baerenklau, K., Ellis, E.A., & Marcos-Martínez, R. (2012). Economics of land use dynamics in two Mexican coffee agroforests: Implications for the environment and inequality. *Investigacion Economica*, *LXXI*(279), 93–124.

Ben-Akiva, M., & Lerman, S. (1985). *Discrete choice analysis: Theory and application to travel demand.* Cambridge, MA: MIT Press (http://books.google.com/books?hl=en&lr=&id=oLC6ZYPs9UoC&oi=fnd&pg=PR11&dq=ben-akiva+lerman&ots=nKgve-fkDd&sig=GTGBUFbpxmcL2SkRksZuwE9YZs8).

Bhagwat, S.A., Willis, K.J., Birks, H.J.B., & Whittaker, R.J. (2008). Agroforestry: a refuge for tropical biodiversity? *Trends in Ecology & Evolution (Personal Edition)*, *23*(5), 261–267.

Blackman, A., Albers, H.J., Avalos-Sartorio, B., & Murphy, L.C. (2008). Land cover in a managed forest ecosystem: Mexican shade coffee. *American Journal of Agricultural Economics*, *90*(1), 216–231. http://dx.doi.org/10.1111/j.1467-8276.2007.01060.x.

Blackman, A., Ávalos-Sartorio, B., & Chow, J. (2012). Land cover change in agroforestry: Shade coffee in El Salvador. *Land Economics*, *88*(1), 75–101 (http://www.

Table 8
Accuracy indicators at different classification error levels.

% misclassified AB observations	Observed accuracy	Expected accuracy	Kappa
25%	0.9749	0.5609	0.9427
60%	0.9733	0.5660	0.9393
95%	0.9720	0.5598	0.9363

Note: Since the TC and GC observations were not misclassified, accuracy indicators exclude those land uses.

scopus.com/inward/record.url?eid=2-s2.0-84855582595&partnerID=40&md5=5878b1594f6c4ddd816c0292b6344ec9).

Caudill, S.B. (2006). A logit model with missing information illustrated by testing for hidden unemployment in transition economies. *Oxford Bulletin of Economics and Statistics*, *68*(5), 665–677.

Caudill, S.B., & Mixon, F.G., Jr. (2005). Analysing misleading discrete responses: a logit model based on misclassified data. *Oxford Bulletin of Economics and Statistics*, *67*(1), 105–113.

Caudill, S.B., Ayuso, M., & Guillen, M. (2005). Fraud detection using a multinomial logit model with missing information. *The Journal of Risk and Insurance*, *72*, 539–550. http://dx.doi.org/10.1111/j.1539-6975.2005.00137.x.

Caudill, S.B., Groothuis, P.A., & Whitehead, J.C. (2011). The development and estimation of a latent choice multinomial logit model with application to contingent valuation. *American Journal of Agricultural Economics*, *93*(4), 983–992. http://dx.doi.org/10.1093/ajae/aar030.

Cayuela, L., Benayas, J.M.R., & Echeverría, C. (2006). Clearance and fragmentation of tropical montane forests in the highlands of Chiapas, Mexico (1975–2000). *Forest Ecology and Management*, *226*(1–3), 208–218. http://dx.doi.org/10.1016/j.foreco.2006.01.047.

Chardin, A., & Perez, P. (1999). Unsupervised image classification with a hierarchical EM algorithm. *Proceedings of the Seventh IEEE International Conference on computer vision*. Published by the IEEE Computer Societyhttp://dx.doi.org/10.1109/ICCV.1999.790353.

Chomitz, K.M., & Gray, D.A. (1996). Roads, land use, and deforestation: A spatial model applied to Belize. *The World Bank Economic Review*, *10*, 487–512. http://dx.doi.org/10.1093/wber/10.3.487.

CONAPO (1998). *Índices de Marginación, 1995* (1st ed.). Mexico City: Consejo Nacional de Población.

CONAPO (2006). *Índices de Marginación 2005* (1st ed.). Mexico City: Consejo Nacional de Población (http://www.conapo.gob.mx/es/CONAPO/Indices_de_marginacion_2005_).

CONAPO (2011). *Indice de Marginacion Por Entidad Federativa Y Municipio, 2010.* Mexico City: Consejo Nacional de Población.

De Pinto, A., & Nelson, G.C. (2008). Land use change with spatially explicit data: A dynamic approach. *Environmental and Resource Economics*, *43*(2), 209–229. http://dx.doi.org/10.1007/s10640-008–9232-x (Springer).

Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B Methodological, Series B*, *39*(1), 1–38. http://dx.doi.org/10.2307/2984875 (JSTOR).

Dinata Putra, A.E., Verbist, B., & Budidarsono, S. (2005). Factors driving land use change: effects on watershed functions in a coffee agroforestry system in Lampung, Sumatra. *Agricultural Systems*http://dx.doi.org/10.1016/j.agsy.2005.06.010.

Dunn, R., & Harrison, A.R. (1993). Two-dimensional systematic sampling of land use. *Journal of the Royal Statistical Society: Series C: Applied Statistics*, *42*(4), 585–601 (Blackwell Publishing for the Royal Statistical Society, http://www.jstor.org/stable/2986177).

Ellis, E.A., Baerenklau, K.A., Marcos-Martínez, R., & Chávez, E. (2010). Land use/land cover change dynamics and drivers in a low-grade marginal coffee growing region of Veracruz, Mexico. *Agroforestry Systems*, *80*(1), 61–84. http://dx.doi.org/10.1007/s10457-010–9339–2 (Springer Netherlands).

Escamilla Prado, E. (2007). *Influencia de Los Factores Ambientales, Genéticos, Agronómicos Y Sociales En La Calidad Del Café Orgánico En México.* INIFAP (http://www.biblio.colpos.mx:8080/jspui/handle/10521/1625).

Fraser, R.H., Olthof, I., & Pouliot, D. (2009). Monitoring land cover change and ecological integrity in Canada's National Parks. *Remote Sensing of Environment*, *113*(7), 1397–1409. http://dx.doi.org/10.1016/j.rse.2008.06.019 (Elsevier B.V.).

Fritz, S., & See, L. (2008). Identifying and quantifying uncertainty and spatial disagreement in the comparison of global land cover for different applications. *Global Change Biology*, *14*(5), 1057–1075.

Gao, H., & Jia, G. (2013). Assessing disagreement and tolerance of misclassification of satellite-derived land cover products used in WRF model applications. *Advances in Atmospheric Sciences*, *30*, 125–141.

Geist, H.J., & Lambin, E.F. (2002). Proximate causes and underlying driving forces of tropical deforestation. *BioScience*, *52*(2), 143–150. http://dx.doi.org/10.1641/0006–3568(2002)052[0143:PCAUDF]2.0.CO;2 (American Institute of Biological Sciences).

Huang, W., Luukkanen, O., Johanson, S., Kaarakka, V., Räisänen, S., & Vihemäki, H. (2002). Agroforestry for biodiversity conservation of nature reserves: functional group identification and analysis. *Agroforestry Systems*, *55*(1), 65–72. http://dx.doi.org/10.1023/A:1020284225155.

INEGI (1998). *Curvas de Nivel Para La República Mexicana.* Mexico City: Instituto Nacional de Estadística, Geografía e Informática.

INEGI (1999). *Cartas Vectoriales 1:20 000.* (Mexico City).

Jordan-Garcia, A., Collazo, J.A., Borkhataria, R., & Groom, M.J. (2012). Shade-grown coffee in Puerto Rico: opportunities to preserve biodiversity while reinvigorating a struggling agricultural commodity. *Agriculture, Ecosystems & Environment*http://dx.doi.org/10.1016/j.agee.2010.12.023.

Kleynhans, W., Olivier, K.J., Wessels, B.P., van den Bergh Salmon, F., Steenkamp, K., Olivier, J.C., ... Steenkamp, K. (2010O). Detecting land cover change using an extended Kalman filter on MODIS NDVI time series data. *IEEE Geoscience and Remote Sensing Letters*, *8*(3), 507–511 (papers3://publication/uuid/FA8F0FF5-127F-4A72-A95E-90AC1E768184, IEEE).

Kolawole, G.O., Salako, F.K., Idinoba, P., Kang, B.T., & Tian, G. (2005). Long-term effects of fallow systems and lengths on crop production and soil fertility maintenance in west Africa. *Nutrient Cycling in Agroecosystems*http://dx.doi.org/10.1007/s10705-004-1927-y.

Kursten, E. (2000). Fuelwood production in agroforestry systems for sustainable land use and CO2-mitigation. *Ecological Engineering*, *16*, S69–S72.

Lehmann, E.A., Wallace, J.F., Caccetta, P.A., Furby, S.L., & Zdunic, K. (2013). Forest cover trends from time series landsat data for the Australian continent. *International Journal of Applied Earth Observation and Geoinformation*, *21*(1), 453–462.

Lewis, J., & Runsten, D. (2008). Is fair trade-organic coffee sustainable in the face of migration? Evidence from a oaxacan community. *Globalizations*http://dx.doi.org/10.1080/14747730802057738.

Lubowski, R.N., Plantinga, A.J., & Stavins, R.N. (2008). What drives land-use change in the United States? A national analysis of landowner decisions. *Land Economics*, *84*(4), 529–550. http://dx.doi.org/10.3368/le.84.4.529 (University of Wisconsin Press).

McFadden, D. (1978). Quantitative methods for analyzing travel behaviour of individuals: Some recent developments. In D. Hensher, & P. Stopher (Eds.), *Behavioural travel modelling* (pp. 279–318). Croom Helm.

McLachlan, G.J., & Krishnan, T. (1997). The EM algorithm and extensions. In John Wiley Sons (Ed.), *Wiley Series in Probability and Statistics*, *vol. 274*, Wileyhttp://dx.doi.org/10.1002/9780470191613 (New York).

Melgani, F. (2006). Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, *44*http://dx.doi.org/10.1109/TGRS.2005.861929.

Messer, K.D., Kotchen, M.J., & Moore, M.R. (2000). Can shade-grown coffee help conserve tropical biodiversity? A market perspective. *Endangered Species Update*, *17*(6), 125–131.

Muñoz-Villers, L.E., & López-Blanco, J. (2008). Land use/cover changes using landsat TM/ETM images in a tropical and biodiverse mountainous area of Central-Eastern Mexico. *International Journal of Remote Sensing*http://dx.doi.org/10.1080/01431160701280967.

Nava-Tablada, M.E., & Martínez-Camarillo, E. (2012). International migration and change in land use in Bella Esperanza, Veracruz. *Tropical and Subtropical Agroecosystems*, *15*, S21–S29 (http://www.scopus.com/inward/record.url?eid=2-s2.0-84871062096&partnerID=40&md5=333ee4258fda4dd29ecbf054ac519b07).

Oxfam, M. (2002). *Poverty in your coffee cup.* Boston: Oxfam America (http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Poverty+in+your+coffee+cup#4).

Puri, J. (2006). Factors affecting agricultural expansion in forest reserves of Thailand: The role of population and roads. *UMD theses and dissertations agricultural & resource economics theses and dissertations* (http://drum.lib.umd.edu/bitstream/1903/3481/1/umi-umd-3308.pdf).

SAGARPA (2006). Servicio de Información Y Estadística Agroalimentaria Y Pesquera. Mexico City http://www.siap.gob.mx/index.php?option=com_content&view=article&id=44&Itemid=378

SAGARPA (2012). *Sistema de Información Agropecuaria de Consulta in Servicio de Información Y Estadística Agroalimentaria Y Pesquera.* Mexico City: Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (http://www.siap.gob.mx/index.php?option=com_content&view=article&id=181&Itemid=426).

Schmitt-Harsh, M. (2013). Landscape change in Guatemala: driving forces of forest and coffee agroforest expansion and contraction from 1990 to 2010. *Applied Geography*, *40*, 40–50. http://dx.doi.org/10.1016/j.apgeog.2013.01.007 (June, Elsevier Ltd).

Schroth, G. (2004). *Agroforestry and biodiversity conservation in tropical landscapes.* Island Presshttp://dx.doi.org/10.1007/s10457-006-9011-z.

SEMARNAP (1998). *Mapa de Suelos Dominantes de La República Mexicana.* Secretaría del Medio Ambiente, Recursos Naturales y Pesca.

Shanker, C., & Solanki, K.R. (2000). Agroforestry: An ecofriendly land-use system for insect management. *Outlook on Agriculture*http://dx.doi.org/10.5367/000000000101293095.

Steele, B.M., Chris Winne, J., & Redmond, R.L. (1998). Estimation and mapping of misclassification probabilities for thematic land cover maps. *Remote Sensing of Environment*, *66*(2), 192–202 (Elsevier Science Inc.).

Stryhn, H., & Christensen, J. (2003). Confidence intervals by the profile likelihood method, with applications in veterinary epidemiology. *Proceedings of the 10th International Symposium on veterinary epidemiology and economics, 0:208. Viña del Mar, Chile*http://people.upei.ca/hstryhn/stryhn208.pdf

Susaki, J., Shibasaki, R., Susaki, R., & Shibasaki, J. (2000). Maximum likelihood method modified in estimating a prior probability in improving misclassification errors. *International Archives of Photogrammetry and Remote Sensing*, *XXXIII*(B7) (Amsterdam).

Swallow, B., Boffa, J. -m., & Scherr, S.J. (2006). The potential for agroforestry to contribute to the conservation and enhancement of landscape biodiversity. In D. Garrity, A. Okono, M. Grayson, & S. Parrott (Eds.), *World agroforestry into the future* (pp. 95–101). World Agroforesty Centre.

Tergas, L.E., & Sanchez, P.A. (1979). *Produccion de Pastos En Suelos Acidos de Los Tropicos. Serie 03SG.*

Tian, G., Salako, F.K., Kolawole, G.O., & Kang, B.T. (1999). An improved cover crop-fallow system for sustainable management of low activity clay soils of the tropics. *Soil Science*http://dx.doi.org/10.1097/00010694-199909000-00007.

Train, K. (2009). Discrete choice methods with simulation. In Cambridge University Press (Ed.), (2nd. ed.). *Discrete choice methods with simulation*, *vol. 47*, . New York: Cambridge University Presshttp://dx.doi.org/10.1016/S0898-1221(04)90100-9 (New York).

Yang, H.L., Peng, J.H., Xia, B.R., & Zhang, D.X. (2013). An improved EM algorithm for remote sensing classification. *Chinese Science Bulletin*, *58*(9), 1060–1071. http://dx.doi.org/10.1007/s11434-012-5485-4.

Yuan, F., Sawaya, K.E., Loeffelholz, B.C., & Bauer, M.E. (2005). Land cover classification and change analysis of the twin cities (Minnesota) metropolitan area by multitemporal landsat remote sensing. *Remote Sensing of Environment*, *98*(2–3), 317–328.

Zhai, D. (2007). A note on the expectation–maximization (EM) Algorithm. http://times.cs.uiuc.edu/course/410s13/em-note.pdf